

Molecular genetics and cell biology of ciliopathies

Katarzyna Szymanska

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

The University of Leeds
School of Medicine and Health

February, 2015

The candidate confirms that the work submitted is her own, except where work which has formed part of jointly authored publications has been included. The contribution of the candidate and the other authors to this work has been explicitly indicated below. The candidate confirms that appropriate credit has been given within the thesis where reference has been made to the work of others.

III

The work in Chapter 1 of the thesis has been published as follows: The transition zone: an essential functional compartment of cilia. Szymanska K, Johnson CA. *Cilia*. 2012 Jul 2;1(1):10. I was responsible for preparation of the manuscript. The contribution of the other authors was editing of the manuscript.

The work in Chapter 3.1 of the thesis has been published as follows: Founder mutations and genotype-phenotype correlations in Meckel-Gruber syndrome and associated ciliopathies. Szymanska K, Berry I, Logan CV, Cousins SR, Lindsay H, Jafri H, Raashid Y, Malik-Sharif S, Castle B, Ahmed M, Bennett C, Carlton R, Johnson CA. *Cilia*. 2012 Oct 1;1(1):18. I was responsible for performing experiments and preparation of the manuscript. The contribution of the other authors was to verify some of the mutations and do in silico analysis, as well as editing of the manuscript.

The work in Chapter 3.2 of the thesis has appeared in publication as follows: *TTC21B* contributes both causal and modifying alleles across the ciliopathy spectrum. Davis EE, Zhang Q, Liu Q, Diplas BH, Davey LM, Hartley J, Stoetzel C, Szymanska K, Ramaswami G, Logan CV, Muzny DM, Young AC, Wheeler DA, Cruz P, Morgan M, Lewis LR, Cherukuri P, Maskeri B, Hansen NF, Mullikin JC, Blakesley RW, Bouffard GG; NISC Comparative Sequencing Program, Gyapay G, Rieger S, Tönshoff B, Kern I, Soliman NA, Neuhaus TJ, Swoboda KJ, Kayserili H, Gallagher TE, Lewis RA, Bergmann C, Otto EA, Saunier S, Scambler PJ, Beales PL, Gleeson JG, Maher ER, Attié-Bitach T, Dollfus H, Johnson CA, Green ED, Gibbs RA, Hildebrandt F, Pierce EA, Katsanis N. *Nat Genet*. 2011 Mar;43(3):189-96. I was responsible for verification of mutations in *TTC21B* in MKS/JSRD Leeds patients cohort. The contribution of the other authors was to identify mutations in *TTC21B*, verify them in their patients cohorts, perform functional studies, prepare manuscript.

The work in Chapter 3.2 of the thesis has appeared in publication as follows: *TCTN3* mutations cause Mohr-Majewski syndrome. Thomas S, Legendre M, Saunier S, Bessières B, Alby C, Bonnière M, Toutain A, Loeuillet L, Szymanska K, Jossic F, Gaillard D, Yacoubi MT, Mougou-Zerelli S, David A, Barthez MA, Ville Y, Bole-Feysot C, Nitschke P, Lyonnet S, Munnich A, Johnson CA, Encha-Razavi F, Cormier-Daire V, Thauvin-Robinet C, Vekemans M, Attié-Bitach T. *Am J Hum Genet*. 2012 Aug 10;91(2):372-8. I was responsible for screening for mutations in *TCTN3* in MKS/JSRD Leeds patients cohort. The contribution of the other authors was to identify mutations in *TCTN3*, verify them in their patients cohorts, perform functional studies, prepare manuscript.

The work in Chapter 3.2 of the thesis has appeared in publication as follows: Mutations in *CSPP1*, encoding a core centrosomal protein, cause a range of ciliopathy phenotypes in humans. Shaheen R, Shamseldin HE, Loucks CM, Seidahmed MZ, Ansari S, Ibrahim Khalil M, Al-Yacoub N, Davis EE, Mola NA, Szymanska K, Herridge W, Chudley AE, Chodirker BN, Schwartzentruber J, Majewski J, Katsanis N, Poizat C, Johnson CA, Parboosingh J, Boycott KM, Innes AM, Alkuraya FS. *Am J Hum Genet.* 2014 Jan 2;94(1):73-9. I was responsible for screening for mutations in *CSPP1* in MKS/JSRD Leeds patients cohort. The contribution of the other authors was to identify mutations in *CSPP1*, verify them in their patients cohorts, perform functional studies, and prepare manuscript.

The work in Chapter 3.2 of the thesis has appeared in publication as follows: *TMEM237* is mutated in individuals with a Joubert syndrome related disorder and expands the role of the TMEM family at the ciliary transition zone. Huang L, Szymanska K, Jensen VL, Janecke AR, Innes AM, Davis EE, Frosk P, Li C, Willer JR, Chodirker BN, Greenberg CR, McLeod DR, Bernier FP, Chudley AE, Müller T, Shboul M, Logan CV, Loucks CM, Beaulieu CL, Bowie RV, Bell SM, Adkins J, Zuniga FI, Ross KD, Wang J, Ban MR, Becker C, Nürnberg P, Douglas S, Craft CM, Akimenko MA, Hegele RA, Ober C, Utermann G, Bolz HJ, Bulman DE, Katsanis N, Blacque OE, Doherty D, Parboosingh JS, Leroux MR, Johnson CA, Boycott KM. *Am J Hum Genet.* 2011 Dec 9;89(6):713-30. I was responsible for screening for mutations in *TMEM237* in MKS/JSRD Leeds patients cohort. I have done a series of functional experiments in *in vitro* systems and participated in manuscript preparation. The contribution of the other authors was to identify mutations in *TMEM237*, verify them in their patient cohorts, perform functional studies in zebrafish and *C.elegans* and prepare the manuscript.

The work in Chapter 3.3 of the thesis has appeared in publication as follows: Human Homolog of Drosophila Ariadne (HHARI) is a marker of cellular proliferation associated with nuclear bodies. Elmehdawi F, Wheway G, Szymanska K, Adams M, High AS, Johnson CA, Robinson PA. *Exp Cell Res.* 2013 Feb 1;319(3):161-72. I was responsible for performing an siRNA screen for ubiquitin proteasome system components. The contribution of the other authors was to do functional studies, characterise the HHARI protein and prepare manuscript.

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.

© 2015 The University of Leeds and Katarzyna Szymanska

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

Acknowledgements

For many hours of mentoring, discussing my experiments and future plans, as well for support, I would like to thank my supervisors – Prof. Colin Johnson, Dr Sandra Bell, Dr Carmel Toomes and the late Dr Philip Robinson, whose wise advice led me to interesting findings and taught me to ask questions.

I would like to thank Team Meckel members, as well as my friends from the Department of Ophthalmology and Neuroscience - Dr James Poulter, Dr Clare Logan, Dr Gabrielle Wheway, Dr Matthew Adams, Dr David Parry, Dr Verity Hartill, Dr Zakia Abdelhamed, Subashini Natarajan, Dr Urszula Burska, Ruth Binns and others for constant help, constructive scientific discussions, troubleshooting, support and many hours of laughter.

I would like to thank my husband for his support, patience and help with IT issues. I would like to thank very much to my family for standing by me and supporting my decisions. I also thank my friends Justyna, Beata, Ania and others for being there for me.

My contribution to this work included sequencing of known and candidate MKS/JSRD genes and optimisation and performance of the whole genome siRNA screen with Dr Gabrielle Wheway and Subaashini Natarajan. Set up and analysis of the whole exome sequencing data were done with the kind help of Dr Clare Logan and Dr David Parry. I would like to thank everybody else who contributed to the work presented in this thesis.

Finally I would like to thank the many MKS/JSRD families for participating in the study, collaborators as well as the ‘European Community Framework 7 project “Syscilia” for funding my research.

This thesis is dedicated to my late grandfather Maciej Swiatkowski, as this was as much his dream as mine.

Abstract

Defects in cilia structure and/or function are now known to be the cause of an important group of Mendelian developmental conditions called ciliopathies. Meckel-Gruber syndrome (MKS) and Joubert syndrome-related disorders (JBTS) are the focus of this work. The research comprised genetic screening of an established MKS/JSRD patient cohort for mutations in seven known genes, and different approaches to identify new causes for these disorders. The latter included whole exome sequencing (WES) of mutation-negative patients, and a high-throughput whole genome siRNA-based reverse genetics screen to identify novel ciliopathy genes and genes implicated in the process of ciliogenesis.

Mutation screening in the University of Leeds MKS/JSRD patient cohort showed that about 50% patients (n=29/65) were mutation-negative for known genes and confirmed mutations in *TMEM67* as a major cause of MKS/JSRD. WES gave a conclusive molecular diagnosis for n=4/7 families. WES allowed the identification of mutations in *TMEM237* as a new cause of JSRD. In vitro assays showed that the *TMEM237* protein is required for correct cilia formation and function. Loss of the protein in patient fibroblasts and after transcript knockdown caused defects in ciliogenesis and the Wnt signaling pathway.

The whole genome reverse genetics screen identified new functional modules that were not previously linked to cilia (components of the spliceosome and proteasome) or had a poorly characterized ciliary function (several neuroactive GPCRs). Cross-comparison of screen hits with available WES data allowed the prioritisation and confirmation of mutations in *PIBF1* and *C21orf2* as new causes of JBTS and the skeletal ciliopathy Jeune syndrome, respectively.

In summary, the multiple approaches presented in this work have allowed further insights into the structure and function of the primary cilia, as well as the disease mechanisms of human ciliopathies.

Contents

Acknowledgements	VI
Abstract	VII
Contents	VIII
Figures	XIII
Tables	XVI
Abbreviations	XVIII
1 Introduction.....	1
1.1 Cilia	1
1.1.1 Cilia types	1
1.1.1.1 Primary cilia	3
1.1.1.1.1 Basal body/centriole	3
1.1.1.1.2 The “ciliary gate”	3
1.1.1.1.2.1 Transition fibres (TF)	4
1.1.1.1.2.2 Transition zone (TZ).....	5
1.1.1.1.3 The axoneme	5
1.1.1.1.3.1 Intraflagellar transport (IFT).....	5
1.1.2 Transport to and from the cilium	6
1.1.2.1 Cilia in signalling	9
1.1.2.1.1 Sonic Hedgehog pathway (SHH)	9
1.1.2.1.2 Wnt	10
1.1.2.1.2.1 Canonical Wnt signalling	10
1.1.2.1.2.2 Non-canonical Wnt signalling.....	11
1.1.2.1.3 Other signalling pathways linked to the cilium.....	11
1.2 Ciliopathies	12
1.2.1 Motile cilia	13
1.2.2 Primary cilia	13
1.2.2.1 Meckel - Gruber syndrome (MKS) (MIM#249000)	14
1.2.2.1.1 Progress in gene identification	16
1.2.2.1.2 Molecular characterisation of MKS proteins	17
1.2.2.2 Joubert syndrome (JBTS) (MIM#614464)	17
1.2.2.2.1 Progress in gene identification	19
1.2.2.2.2 Molecular mechanisms of JBTS and MKS proteins	20
1.2.2.3 Future perspectives for research into MKS and JBTS	23
1.2.2.4 Skeletal ciliopathies	24
1.2.2.4.1 Short-rib thoracic dysplasia (SRTD) (MIM#208500).....	24

1.2.2.4.2	Cranioectodermal dysplasia (Sensenbrenner syndrome) (CED)	
(MIM#218330)		24
1.2.2.5	Nephronophthisis (NPHP) (MIM#256100)	25
1.2.2.6	Other ciliopathies	26
1.2.2.7	Inheritance pattern in ciliopathies	26
1.2.2.7.1	Variable expressivity in ciliopathies	27
1.3	Aims of the research	28
2	Materials and methods	30
2.1	Materials	30
2.1.1	General reagents	30
2.1.2	Reaction specific reagents	30
2.1.3	Buffers	34
2.2	Methods	35
2.2.1	Patient DNA	35
2.2.2	Patient fibroblasts	35
2.2.3	DNA extraction	36
2.2.3.1	Saliva	36
2.2.3.2	Cells	36
2.2.4	RNA extraction	36
2.2.5	Polymerase Chain Reaction (PCR)	36
2.2.6	Genotyping	39
2.2.6.1	Sample genotyping using SNPchip	40
2.2.7	Purification with ExoSAP	40
2.2.8	Sequencing	40
2.2.8.1	Sanger sequencing	40
2.2.8.2	WES	41
2.2.9	Quantitative real-time PCR (qRT-PCR)	41
2.2.10	Cells	42
2.2.11	Transfection	42
2.2.11.1	Over-expression	42
2.2.11.2	siRNA	42
2.2.12	IF and confocal microscopy	43
2.2.13	Whole cell extract preparation and WB	43
2.2.14	TopFlash assay	44
2.2.15	Immunoprecipitation (IP)	45
2.2.16	RhoA GST pull-down	45
2.2.17	Flow cytometry	45
2.2.18	The whole genome siRNA screen	46
2.2.18.1	Cell culture	46

2.2.18.2	siRNA	46
2.2.18.3	Transfection.....	47
2.2.18.4	Antibodies and staining reagents	48
2.2.18.5	High-throughput liquid handling	48
2.2.18.6	Fixation	48
2.2.18.7	IF staining	48
2.2.18.8	High-throughput imaging	49
2.2.18.9	Image analysis	49
3	Results	50
3.1	Mutation screening, founder mutations and genotype-phenotype correlations in MKS and associated ciliopathies	52
3.1.1	Research rationale	52
3.1.2	Microsatellite screening of MKS/JSRD loci.....	53
3.1.2.1	Genotyping – results.....	58
3.1.3	Sequencing of MKS/JSRD cohort.....	58
3.1.3.1	Sequencing results.....	59
3.1.3.2	Genotype – phenotype correlations.....	62
3.1.3.3	Conclusions about sequencing analysis.....	62
3.1.4	Characterisation of an allelic series in TMEM67	63
3.1.4.1	Reclassification of TMEM67 variants of unknown significance.....	73
3.1.4.1.1	Clinic – diagnostic laboratory – research laboratory workflow	73
3.1.4.1.1.1	Patient 387 phenotype.....	73
3.1.4.1.1.2	gDNA sequencing	73
3.1.4.1.1.3	In silico analysis.....	74
3.1.4.1.1.4	RNA/cDNA analyses	75
3.1.4.1.2	Patient 178.....	77
3.1.5	Discussion.....	77
3.2	Autozygosity mapping and candidate gene screening for new MKS/JSRD genes 82	82
3.2.1	Identification of a putative new MKS locus on chromosome 12q24.11-24.13 82	
3.2.2	Other ciliopathy candidate genes	84
3.2.2.1	CEP164.....	84
3.2.2.2	TTC21B.....	85
3.2.2.3	IFT genes.....	85
3.2.2.4	TMEM107	86
3.2.2.5	Other ciliopathy candidate genes.....	88
3.2.3	WES as a novel approach to identify causative mutations for MKS/JSRD....	89
3.2.3.1	Illumina GAllx platform.....	89

3.2.3.1.1	MKS family 157	89
3.2.3.2	Illumina MiSeq platform	91
3.2.3.2.1	MKS family 36A	91
3.2.3.2.2	MKS family 66F1	92
3.2.3.3	Illumina HiSeq2500 platform	93
3.2.3.3.1	MKS family 17	93
3.2.3.3.2	MKS family 325	94
3.2.3.3.3	MKS family 144	95
3.2.3.3.4	MKS family 351	96
3.2.3.4	CSPP1	97
3.2.4	Identification of mutations in TMEM237 causing Joubert syndrome related disorder	98
3.2.4.1	Genotyping of MKS/JSRD patient DNA	98
3.2.4.2	Sequencing	99
3.2.4.3	Patient phenotype	100
3.2.4.4	Subcellular localisation of TMEM237	100
3.2.4.5	TMEM237 function in ciliogenesis	101
3.2.4.6	TMEM237 in Wnt signalling	105
3.2.4.7	RPGRIP1L knockdown interferes with TMEM237 localisation	109
3.2.4.8	Loss of TMEM237 disrupts G1/S transition	110
3.2.5	Discussion	111
3.3	High-throughput genome-wide siRNA visual screen for effects on ciliogenesis	118
3.3.1	Screen set up	118
3.3.1.1	Cell line	118
3.3.1.2	siRNA library	118
3.3.1.3	Screen set up and analysis	119
3.3.1.3.1	Assay set up	119
3.3.1.3.2	Assay metrics	120
3.3.1.3.3	Assay analysis	121
3.3.1.3.4	Robust z score	122
3.3.1.3.5	Assay controls	122
3.3.2	Screen results	125
3.3.2.1	Quality control metrics and overall performance of the whole genome screen	126
3.3.2.2	Quality metrics	128
3.3.2.2.1	Off-target effects	128
3.3.2.2.2	MicroRNA – like effects	128
3.3.2.2.3	Partial on-target effects	128

3.3.2.2.4	Array CGH	128
3.3.2.2.5	RNA sequencing of IMCD3 transcripts	129
3.3.2.2.6	Sequence-independent off-target effects	130
3.3.2.2.7	Measurements of assay robustness	131
3.3.2.3	Positive hits.....	132
3.3.2.3.1	Primary screen	132
3.3.2.3.1.1	Data filtering – first pass	132
3.3.2.3.1.2	Data filtering: second pass	138
3.3.2.3.2	Secondary screen	144
3.3.2.3.3	Validation of secondary screen hits.....	149
3.3.2.3.3.1	Spliceosome components	150
3.3.2.3.3.2	GPCRs	154
3.3.2.3.4	Utility of screen results	155
3.3.3	Discussion.....	156
4	Discussion	160
4.1	Mutation screening, founder mutations and genotype – phenotype correlations in MKS and associated ciliopathies	160
4.2	Autozygosity mapping and candidate gene screening	163
4.3	Whole genome siRNA screen	167
4.4	Future plans	169
	References	174
	Appendix	193
1.	Ethical approval and amendments:.....	193
2.	Primer sequences of the screened genes:	204
3.	Homozygous regions in patient 227.	208
4.	Homozygous regions in patient 230.	209
5.	Homozygous regions in patient 261.	210
6.	Cilia recognition protocol:.....	211
7.	siRNA screen sequence-specific off-target effect gene list:.....	213
8.	siRNA screen partially targeted gene transcripts:	220
9.	siRNA screen – array CGH results.	230

Figures

Figure 1-1. Cross-section of different types of cilia.	3
Figure 1-2. Schematic representation of the ciliary structure with focus on the TZ.	4
Figure 1-3. Schematic representation of possible mechanisms of ciliary trafficking.	7
Figure 1-4. Schematic representation of allelism in ciliopathies.	14
Figure 1-5. Clinical features of MKS.	15
Figure 1-6. Clinical features of JBTS.	18
Figure 1-7. Schematic illustration of MKS/JBTS protein localisation in cilia.	21
Figure 1-8. Mendelian inheritance of a recessive mutation localised on an autosomal chromosome in a consanguineous family.	27
Figure 3-1. ClustalX analysis of the TMEM67 protein.	66
Figure 3-2. Summary of all reported <i>TMEM67</i> mutations, the clinical phenotype and the position in relation to domains in the TMEM67 protein.	67
Figure 3-3. Electropherograms presenting sequencing results of genomic DNA of MKS affected patient 387.	74
Figure 3-4. Electropherograms presenting sequencing results of cDNA of MKS affected patient 387.	76
Figure 3-5. RT-PCR primers for amplification of <i>TMEM67</i> cDNA.	77
Figure 3-6. Pie charts summarising mutation analysis in MKS and MKS-like patients in the study cohort.	79
Figure 3-7. Microsatellite genetic markers and mapping of a putative new MKS locus on chromosome 12.	83
Figure 3-8. IMCD3 cells overexpressing TMEM107.	87
Figure 3-9. Analysis of IMCD3 cells after <i>Tmem107</i> knock down.	88
Figure 3-10. MKS family 157 pedigree.	90
Figure 3-11. MKS family 36A pedigree.	91
Figure 3-12. MKS family 66F1 pedigree.	92
Figure 3-13. MKS family 17 pedigree.	93
Figure 3-14. MKS family 325 pedigree.	94
Figure 3-15. MKS family 144 pedigree.	95
Figure 3-16. MKS family 351 pedigree.	96
Figure 3-17. Chromosome 2 genotyping results from 18 MKS/JSRD DNA samples.	99
Figure 3-18. <i>TMEM237</i> genomic organization with domain structure of the TMEM237 transmembrane protein.	100
Figure 3-19. Subcellular localisation of TMEM237 in IMCD3s.	101

Figure 3-20. qRT-PCR quantification of TMEM237 levels between control (HDF) and JSRD2 RNA levels.....	102
Figure 3-21. Disruption of ciliogenesis in JSRD2 patient fibroblasts with the homozygous TMEM237 p.R18* nonsense mutation.	103
Figure 3-22. Tmem237 siRNA knockdowns in IMCD3 cells.	104
Figure 3-23. Measuring the effects of siRNA knockdown of TMEM237 on cilia length.	105
Figure 3-24. Deregulation of Wnt signalling following mutation or loss of TMEM237.	106
Figure 3-25. Pericentriolar and basolateral subcellular localisation of RhoA and increase of actin stress fibres in <i>TMEM237</i> -mutated patient fibroblast.	107
Figure 3-26. TOP Flash assays of canonical Wnt signalling.....	108
Figure 3-27. Ciliary TMEM proteins are anchored at the TZ by MKS5/RPGRIP1L.	110
Figure 3-28. FACS results showing that JSRD2 fibroblasts have a prolonged G1 phase and shortened S phase compared to control cells.....	111
Figure 3-29. Sample images of IMCD3 cells.	120
Figure 3-30. Schematic representation of focal planes imaged in IMCD3 and an example image of visualised cilia.	120
Figure 3-31. Example of image analysis in Columbus software.	121
Figure 3-32. Heatmaps representing results obtained from Columbus software.	121
Figure 3-33. Diagram representing knockdown efficiencies in a panel of positive controls.	123
Figure 3-34. Columbus images for siRNA screen controls.	124
Figure 3-35. Bar graphs showing efficiency of knockdowns for screen controls.	124
Figure 3-36. Histograms representing frequency of %CSC in run 1 and run 2 for the whole genome screen.....	126
Figure 3-37 Histograms representing frequency of WCN in run 1 and run 2.....	126
Figure 3-38. Scatter plot representing robust z scores for %CSC from a ten plate batch single run.	127
Figure 3-39. Correlation plot between two runs representing robust z scores for %CSC.	127
Figure 3-40. Array CGH analysis of mIMCD3 cell line.	129
Figure 3-41. SSMD values for all batches in the whole genome siRNA screen.	132
Figure 3-42. DAVID analysis using Swiss-Prot and Protein Information Resource keywords on the 2,174 primary screen hits.	134

Figure 3-43. DAVID analysis using Gene Ontology term search on the 2,174 primary screen hits.....	135
Figure 3-44. DAVID analysis using KEGG pathway term search on the 2,174 primary screen hits.....	135
Figure 3-45. Cartoon illustrating the enrichment of hits from the primary screen in the ribosome.....	136
Figure 3-46. Cartoon illustrating the enrichment of hits from the primary screen in the spliceosome.....	137
Figure 3-47. Cartoon illustrating the enrichment of hits from the primary screen in the proteasome.....	138
Figure 3-48. DAVID analysis using Swiss-Prot and Protein Information Resource keywords on the 154 filtered primary screen hits.....	143
Figure 3-49. DAVID analysis using GO terms for cellular components on the 154 filtered primary screen hits.....	143
Figure 3-50. DAVID analysis using KEGG pathway term search on the 154 filtered primary screen hits.....	143
Figure 3-51. Cartoon illustrating enrichment of the 154 filtered primary screen hits for neuroactive ligand-receptor interactions.....	144
Figure 3-52. Filtering strategy in the siRNA screen. Initially 18,690 Entrez RefSeq mouse genes were screened.....	145
Figure 3-53. Example of hits confirmed in the secondary screen.....	146
Figure 3-54. DAVID analysis of enriched KEGG pathways in the 68 hits from the secondary screen.....	146
Figure 3-55. Graph representation of qPCR results for selected secondary screen hits.....	150
Figure 3-56. Knockdown of Prpf6, Prpf8 and Prpf31.....	152
Figure 3-57. WB showing loss of protein after siRNA knockdown.....	153
Figure 3-58. Localisation of Prpf proteins to the base of the primary cilium.....	153
Figure 3-59. Localisation of Prpf proteins at the base of the connecting cilium in mouse retina.....	154
Figure 3-60. Localisation of GPCR to the base of cilium in differentiated SHSY5Y cells.....	154
Figure 3-61. GPCR protein localisation in mouse retina sections.....	155
Figure 3-62. Co-IP experiment between co-overexpressed TMEM237-FLAG and PIBF1-GFP in HEK293 cells.....	156

Tables

Table 1-1. List of genes identified with pathogenic mutations causing the MKS phenotype.....	16
Table 1-2. List of genes identified with pathogenic mutations causing the JBTS phenotype.....	19
Table 1-3. List of genes identified with pathogenic mutations causing nephronophthisis.....	25
Table 2-1. Cell lines.....	32
Table 2-2. List of primary antibodies.....	33
Table 2-3. List of MKS genes primers.....	38
Table 2-4. List of microsatellite primers flanking known <i>MKS</i> loci.....	39
Table 2-5. Sequences of siRNA duplexes used to target positive and negative control genes for ciliogenesis.....	47
Table 3-1. Microsatellite genetic markers for the indicated MKS and JSRD genes.....	53
Table 3-2. Table of genotyping results for MKS and JSRD genes for affected individuals and patients - <i>MKS1</i> , <i>TMEM216/TMEM138</i> , <i>TMEM67</i> , <i>RPGRIP1L</i> , <i>CEP290</i> , <i>CC2D2A</i> , <i>NPHP3</i> and <i>TMEM237</i>	57
Table 3-3. Summary of the genotyping results in Leeds MKS/JSRD patient cohort.....	58
Table 3-4. Clinical data and sequencing results of consanguineous and non-consanguineous patients with MKS and MKS-like phenotypes.....	61
Table 3-5. Table summarising all reported changes in <i>TMEM67</i> (NM_153704.5) to date.....	72
Table 3-6. Summary of <i>in silico</i> analyses on MKS patient 387 DNA.....	75
Table 3-7. Sequencing results for <i>TTC21B</i>	85
Table 3-8. List of changes identified in IFT genes in Leeds MKS/JSRD cohort.....	86
Table 3-9. List of variants detected in MKS family 157 sample after WES and variant filtering.....	90
Table 3-10. Summary of homozygous changes identified in WES of 36A sample.....	92
Table 3-11. WES results for 66F1 family.....	92
Table 3-12. Sequencing results of <i>TXNDC15</i> in MKS/JSRD panel of patients.....	93
Table 3-13. WES results for sample 17.....	93
Table 3-14. WES results for family 325. All homozygous changes predicted to be pathogenic are listed.....	95
Table 3-15. WES results for family 144.....	96
Table 3-16. WES results for family 351.....	97

Table 3-17. <i>CSPP1</i> changes identified in CiliaProteome v3 targeted capture sequencing.....	98
Table 3-18. Summary of expression levels of RISC (RNA Induced Silencing Complex) complex components.....	131
Table 3-19. List of genes that passed the second level of filtering.....	143
Table 3-20. List of genes confirmed in the secondary screen.....	149

Abbreviations

%CSC	% of cells with single cilium
AON	Antisense oligonucleotide
ATP	Adenosine triphosphate
BLAST	Basic local alignment search tool
bp	base pair
cDNA	Complementary DNA
cM	CentiMorgan
CNS	Central nervous system
CRD	Cysteine rich domain
dbSNP	Database SNP
del	Deletion
dH ₂ O	De-ionised water
DMSO	Dimethyl sulphoxide
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleotide triphosphate
EDTA	Ethylenediaminetetraacetic acid
ExoI	Exonuclease I
SAP	shrimp alkaline phosphatase
fs	frameshift
Gb	Gigabase
GPCR	G-coupled protein receptors
IBD	Identical by descent
IFT	intraflagellar transport
ins	Insertion
JBTS	Joubert syndrome
JSRD	Joubert syndrome related disorders

XIX

kb	Kilobase
LB	Luria-Bertani
LCA	Leber congenital amaurosis
M	Molar
Mb	Megabase
MKS	Meckel-Gruber syndrome
ml	Millilitre
mM	Millimolar
mRNA	Messenger RNA
ng	Nanogram
NGS	Next generation sequencing
NMD	nonsense-mediated decay
MIM	Mendelian inheritance in man
PAGE	Polyacrylamide gel electrophoresis
PCR	Polymerase chain reaction
pmol	Picomoles
μ M	Micromolar
TEM	transmission electron microscope
TF	transition fibres
TMEM	transmembrane
TZ	transition zone
WCE	whole cell extract
WCN	whole cell number
WES	whole exome sequencing
WGS	whole genome sequencing

1 Introduction

1.1 Cilia

Anthony van Leeuwenhoek was the first to describe a “little leg”-like structure on protozoa in 1675 ¹. The name “cilium” (meaning eyelash in Latin) was probably given in 1786 by Otto Muller ². The cilium is the oldest known cellular organelle and for decades was thought to be vestigial with no particular function in the cell. With advances in microscopy techniques, cilia structure was studied with no further insight into its function. However, the last two decades of research have shown that, not only can cilia be found on almost all cells in the human body, but that they are sophisticated, complex structures that are essential for developmental and homeostatic signalling pathways.

Cilia act as an “antenna” receiving mechanical and chemical signals from the extracellular environment and passing it to the cell body, performing various cellular functions including mechano-, chemo- and photosensation ³. Cilia have been shown to be crucial for normal organ function, and a range of developmental disorders collectively known as “ciliopathies” (section **1.2**) are caused by the defects in its structure and/or function. This highlights the importance of the cilium in normal embryonic development ⁴.

The cilium functions as a separate organelle but unlike mitochondria or endoplasmic reticulum (ER), it does not have a complete plasma membrane boundary with the cytoplasm. This determines a characteristic feature of the cilium, namely its compartmentalisation. Cilia also contain a basal body, derived from the mother and daughter centriole, as well as other structures such as transition fibres (TF) and the transition zone (TZ) that form the ciliary gate, and the ciliary axoneme. The ciliary gate appears to act as a boundary between cilium and the cytoplasm ⁵.

1.1.1 Cilia types

Cilia are microtubule-based, rod-like organelles that occur on the apical surface of most mammalian cells in G₀/G₁ of the cell cycle ⁶, with the exception of bone marrow-derived cells and liver hepatocytes ⁷. There are three main types of cilia that are distinguished on the basis of their microtubule structure observed on cross-section and the number carried by a cell.

Motile cilia contain the canonical “9+2” microtubule pattern (**Figure 1-1a**) with nine outer doublets (containing A and B tubules) and a central pair of microtubules.

Outer doublets contain dynein arms and are connected by nexin links. Radial spokes protrude from these doublets towards the central sheath and microtubule pair⁸. Inner and outer dynein arms are responsible for energy production⁹ that allows these cilia to perform “whip”-like movements that facilitate movement of mucus in the airways, eggs in the oviduct or cerebrospinal fluid in the brain ventricles^{10,11}. There are also specialised motile cilia crucial for the sense of hearing, localised in the Organ of Corti in the inner ear called kinocilium^{12,13}. The kinocilium seems to establish polarity in the inner and outer hair cells of the Organ of Corti, this then allows the formation of the hair cell bundles on the abneural side of the cell. These are actin-based and are the structures that detect sound waves and are called stereocilia.

Nodal cilia have a “9+0” microtubule pattern with outer microtubules, dynein arms and nexin connections present (**Figure 1-1b**). They are located at the embryonic node and they have the ability to mediate a leftward flow in a “whirlpool”-like manner. The leftward flow is thought to transport vesicles containing morphogens that are crucial for symmetry-breaking, and is an essential step during the correct development of the mammalian body plan⁷.

The third type of cilium is an immotile (“9+0”) primary cilium (**Figure 1-1c**). It contains only nine outer microtubule doublets and its movement depends on fluid flow. It is now known to participate in diverse roles in cell signalling, chemo-, mechano- and photosensation⁸ and these are discussed in more detail below (section 1.1.1.1). A specialised type of primary cilium can be found in the photoreceptors, where the TZ is hugely extended and is often known as the “connecting cilium”. In addition, the ciliary tip in photoreceptors is elongated and expanded, and is filled with stacked membrane discs containing cone opsin or rhodopsin¹⁴

The other specialised cilium with a hybrid character is the olfactory cilium. These cilia occur in multiple numbers on olfactory receptor neurons and are essential for the sense of smell. They have “9+2” microtubule pattern but no dynein arms that would facilitate their movement¹⁵.

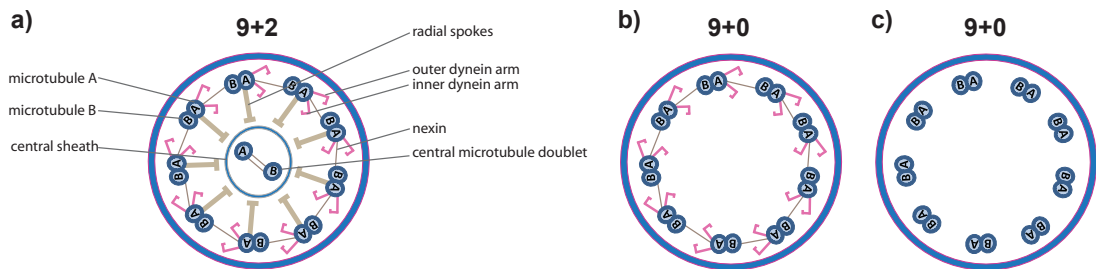


Figure 1-1. Cross-section of different types of cilia. a) motile cilia with “9+2” microtubule pattern, containing nine doublets of microtubules, dynein arms, nexin links, radial spokes, sheath and central microtubule pair, b) “9+0” nodal cilia, containing nine doublets of microtubules, dynein arms and nexin links, c) “9+0” primary cilia with nine doublets of microtubules.

1.1.1.1 Primary cilia

1.1.1.1.1 Basal body/centriole

In quiescent cells, a mother and daughter centrioles migrates to the apical surface of the cell and matures into the basal body (**Figure 1-2**, orange colour). The mother centriole consists of a barrel-shaped structure with nine triplets of microtubules, each triplet consisting of A-, B- and C-tubules. The basal body is built of mother and daughter centrioles, assembled at 90° in relation to each other. After migration to the apical cell surface, the basal body docks to the plasma membrane via fibrous distal and sub-distal appendages¹⁶ and the mother centriole subsequently acts as a matrix for microtubule nucleation during cilium formation.

1.1.1.1.2 The “ciliary gate”

The region just above the basal body is called the “ciliary gate” and consists of two structurally distinct sub-regions that include the TF and the TZ^{17,18} (**Figure 1-2**). The ciliary gate has been shown to form at the very early stages of ciliogenesis preceding ciliary axoneme extension and intraflagellar transport (IFT; section 1.1.1.1.3.1). The TF begin with the termination of the C-tubules of the basal body. The boundary between the axoneme and the TZ is known as the basal plate and, in motile cilia, is thought to take part in the nucleation of the central microtubules¹⁹. The basic components of the TZ appear to be conserved, but the structure varies between species and cell type²⁰, with the example of the connecting cilium in mammalian photoreceptor cells as the most elongated ciliary TZ²¹.

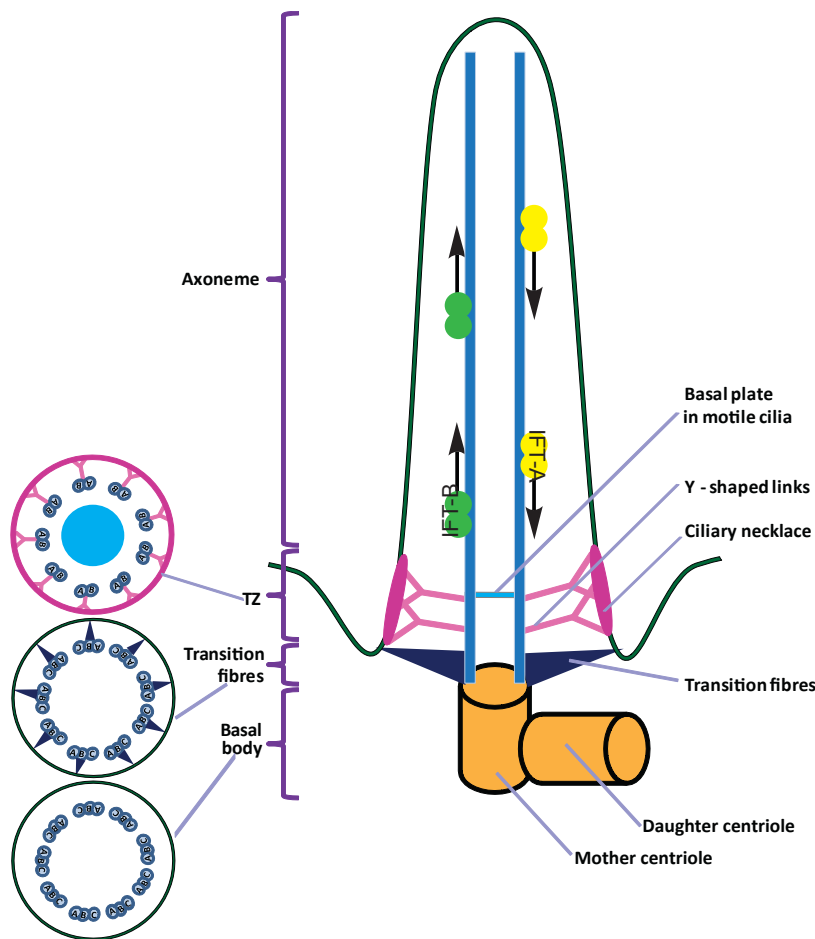


Figure 1-2. Schematic representation of the ciliary structure with focus on the TZ. The main sub-structures of cilia include the ciliary axoneme, TZ, TF and the basal body. The TZ compartment contains the ciliary necklace and Y-shaped links; the basal plate is only observed in motile cilia. The left side of the cartoon shows cross-sections of cilia at the basal body, TF and TZ level. Green circles represent anterograde transport by the intraflagellar transport IFT-B complex and the yellow ones retrograde transport by IFT-A. Image adapted from Szymanska and Johnson, 2012.

1.1.1.1.2.1 Transition fibres (TF)

The composition of TF is still largely unknown but they emerge from the B-tubules of the basal body triplet microtubules just before the end of the C-tubule and form a “pinwheel-like” structure on transmission electron microscope (TEM) cross-sections (**Figure 1-2**). TF, also known as distal appendages, are observed on the mature mother centriole and play a role in anchoring to the plasma membrane through CEP164²² and ODF2 (outer dense fibre 2)/cenexin²³. Some of the known components of the TF include the ciliopathy proteins OFD1²⁴ and CEP123²⁵. In addition, IFT52 was observed in *Chlamydomonas* on TF²⁶, suggesting that they have a role in docking the IFT and motor proteins required for ciliogenesis. It is also suggested that TF take part in creating a pore complex, similar to nuclear pores, that are required for transporting proteins in and out of cilia¹⁷.

1.1.1.1.2.2 Transition zone (TZ)

The TZ localises distally to the TF and contains characteristic structures of so-called “Y-shaped” linkers and the ciliary necklace (**Figure 1-2**). Y-shaped linkers are structures connecting the outer doublets of microtubules to the plasma membrane and the ciliary necklace²⁷. Their sophisticated structure is suggested to be built from known ciliary proteins. These include NPHP1, NPHP4 with a function of microtubule binding, CEP290, NPHP8 – acting as a connecting structures, B9 and C2 domain containing proteins like B9D1, B9D2, MKS1, CC2D2A, RPGRIP1L with a role of lipid binding at the plasma membrane, TMEM67, TMEM138, TMEM216, TCTN2, TCTN3 – components of ciliary necklace and TCTN1 an extracellular ciliary necklace protein²⁸. The ciliary necklace is a specialised structure that contains several parallel strands of intramembrane cell-specific particles. The identity of these strands is unknown, but they encircle the ciliary membrane spacing from the plasma/ciliary membrane boundary to the basal plate¹⁹. Y-shaped linkers and the ciliary necklace are especially visible in the elongated TZ structure of connecting cilia in photoreceptors²⁹. Similar to the TF, the TZ has been proposed to regulate ciliary protein composition in *Chlamydomonas*, *C. elegans* and mammalian cells^{5,30,31} by regulating intracellular trafficking to/from the cilium. However, the molecular details of protein sorting at the TZ remain to be discovered.

1.1.1.1.3 The axoneme

Matured centrioles dock at the cell apical surface via the distal appendages/TF³². The TZ is then established with motor and transport proteins recruited from the cytoplasm to nucleate the ciliary axoneme, containing the cilia-specific pattern of doublets of A- and B-tubules (**Figure 1-1**). The axoneme is enclosed by the ciliary membrane that is a continuation of the cell plasma membrane (**Figure 1-2**)¹⁸. Structural and functional components of the ciliary axoneme have to pass the ciliary gate and are transported along the axoneme by IFT.

1.1.1.1.3.1 Intraflagellar transport (IFT)

During ciliogenesis, IFT extends the centriolar doublets that contain A- and B-tubules to form the ciliary axoneme³³. Once the axoneme is assembled, kinesin motors (specifically kinesin-2) mediate anterograde transport along the microtubules towards the tip of the cilium by carrying IFT-complex B proteins and other cargo proteins (including LCA5, IFT20, IFT80, IFT88) (**Figure 1-2**). In turn, cytoplasmic dyneins (dynein-2) mediate retrograde movement of the IFT-complex A (including IFT122, IFT140, TTC21B/IFT139) towards the proximal regions of the

cilium. IFT mediates both the assembly and resorption of the cilium, and the processing of key intermediates of signalling cascades³⁴.

1.1.2 Transport to and from the cilium

There is no simple model for defining the mechanisms of protein targeting to cilia. Although the identity of ciliary cargo proteins is becoming clearer, the mechanisms of their intracellular transport are still largely unknown.

Mechanisms of transmembrane protein transport to cilia are of particular interest, since these proteins will then reside in the ciliary plasma membrane and act as receptors to receive extracellular signals. These proteins were shown to be packaged into vesicles at the trans Golgi network (TGN) and transported to cilia by the recognition of ciliary targeting sequences (CTS)³⁵⁻³⁷. There are many different CTS identified, suggesting either a lack of consensus sequence across ciliary proteins or a lack of CTS indicating that the ones identified so far are an incidental finding. Identified sequences include one in the N-terminus of polycystin-2 (PKD2) comprising RVxP (where x is any amino acid)³⁷, the C-terminal VxPx sequence of rhodopsin³⁶ or the Ax(S/A)xQ motif present in the third intracellular loop of G-protein-coupled receptors (GPCRs) such as SSTR3, HTR6 or MCHR1³⁸. These sequences are thought to be recognised by small GTPases including ARL, ARF and RAB which in turn recruit proteins that regulate membrane interactions³⁹.

Proteins are thought to be transported in vesicles from the Golgi apparatus to a specific docking site at the periciliary base. The exocyst complex is then thought to tether the vesicles, presumably directing the fusion of vesicles with the ciliary membrane mediated by soluble *N*-ethylmaleimide sensitive factor receptors (SNAREs) and the Rab family of small GTPases⁴⁰. SNAREs are present at the surface of the exocyst complex to allow proteins to pass the ciliary gate and to cross to the ciliary membrane⁴⁰, implying an active transport process⁴¹. This model of trafficking would therefore be analogous to that proposed for nuclear pores where importins and exportins are utilized. IFT proteins may mediate this transport machinery since they are enriched at the level of TF²⁶. Once trafficked through the ciliary gate, proteins may be then incorporated as cargo proteins in the IFT complexes and transported along the ciliary axoneme.

Other proposed mechanisms of ciliary protein transport include an interaction with IFT20 (**Figure 1-3a**). IFT20 is the only IFT protein observed outside of the cilium. It interacts with transmembrane proteins in the TGN facilitating their packaging into vesicles and chaperoning their transport to the ciliary base⁴². This

interaction has so far only been shown for polycystin-2, but may be a common mechanism for other ciliary membrane proteins.

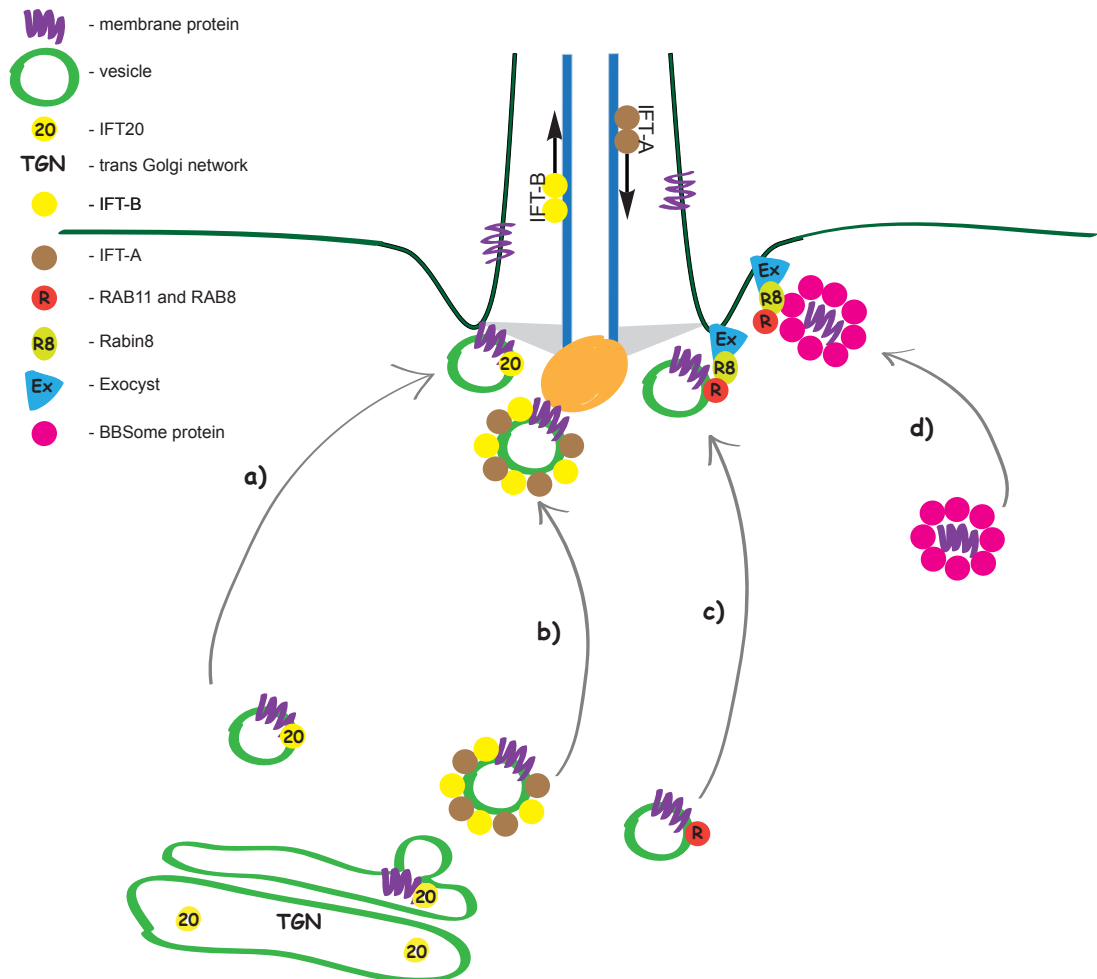


Figure 1-3. Schematic representation of possible mechanisms of ciliary trafficking. The cartoon illustrates four proposed mechanisms of ciliary protein trafficking. a) a vesicle is targeted to the ciliary base by IFT20; b) IFT proteins coat the vesicle and target it to the ciliary base; c) vesicle trafficking is facilitated by RAB11 and RAB8. At the periciliary membrane the vesicle binds the exocyst complex via Rabin8; d) a ciliary targeting sequence is recognised by the BBSome, a complex of proteins that interacts with the exocyst via Rabin8 at the periciliary membrane.

Another model suggests that IFT proteins create clusters in the TGN that coat vesicles and ensure targeted trafficking to the cilium (**Figure 1-3b**)⁴². These clusters would later become IFT complexes still attached to cargo proteins.

Compatible with these proposed mechanisms, or indeed as an independent process, it is thought that vesicles containing ciliary proteins bud from the TGN and are directed to the ciliary base where they tether to the ciliary membrane. This process is facilitated by coordinated interactions between Rab GTPases and the exocyst complex containing seven components: SEC3, SEC5, SEC6, SEC8, SEC15, EXO70 and EXO84 (**Figure 1-3c**). RAB11 and RAB8 mediate the transport of the vesicle to the base of the cilium, where Rabin8 (the guanine nucleotide exchange factor [GEF] RAB8 activator) binds to the exocyst complex protein

SEC15⁴³. The exocyst subsequently facilitates the tethering of the vesicle to the periciliary membrane^{44,45}. An additional downstream binding partner of RAB8 is CEP290, which has been shown to localise to the pericentrosomal compartment, basal body and the TZ^{46,47}. CEP290 plays an important role at the ciliary gate, where together with other proteins containing C2/B9-related domains is predicted to bind phospholipids⁴⁸ and to be involved in membrane/vesicle trafficking and fusion⁴⁶. In particular, CC2D2A, a C2 domain-containing protein, is now also suggested to mediate vesicle trafficking and fusion since it localizes to the photoreceptor connecting cilium/TZ, and appears to facilitate protein transport through Rab8-dependent processes⁴⁹.

Another possible mechanism of protein targeting to cilia is through the BBSome complex (**Figure 1-3d**). It consists of eight proteins with a coat-like structure and is proposed to mediate trafficking of transmembrane proteins to the ciliary membrane⁵⁰. The BBSome directly recognises cilia targeting sequences and is the major effector of Arl6/BBS3 (an Arf-like GTPase). Rabin8, interacting with the BBSome, regulates RAB8 in the same manner as the exocyst complex⁵⁰. The BBSome is not directly required for cilia formation in most tissues⁵¹ but its failure to deliver important receptors and transmembrane proteins to the cilium is thought to result in cell signalling failure and organ-specific pathological abnormalities. It is unclear how the actions of the BBSome and exocyst complex are coordinated. It was suggested that the BBSome may work as part of a separate mechanism. In this scenario BBSome transports a subset of ciliary proteins to the base of the cilium, and upon the entry to the cilium it acts as an IFT adaptor⁵².

Following tethering to the ciliary membrane, fusion of the vesicle with the periciliary membrane is facilitated by interaction between SNARE proteins on the vesicle (v-SNARE) and target SNARE (t-SNARE) regions of the membrane⁵³. Incorporated transmembrane proteins subsequently pass the diffusion barrier of the ciliary gate and fulfil their function in the ciliary membrane. The exact process of passing the ciliary gate is not well understood. Vesicles have never been observed in the cilium itself, suggesting that membrane proteins are being deposited before reaching the ciliary gate. Lateral diffusion processes have been observed as a method of translocation of activated Smoothed (the Sonic hedgehog signalling receptor) from the apical cell membrane via ciliary gate to the ciliary membrane⁵⁴. In photoreceptor cells, vesicles containing rhodopsin were shown to pass to the ciliary membrane by the periciliary ridge complex, a structure consisting of nine symmetrically arranged ridges and grooves that surrounds the connecting cilium and extends outwards along the plasma membrane⁴¹. Conversely, transmembrane proteins that have essential functions in ciliary signalling, are prevented from

premature exit from cilia by a septin diffusion barrier localised at the base of the primary cilium³⁰.

During the process of cilia disassembly or after fulfilling their function in cilia, proteins must be removed from cilia. The BBSome or the proteosomal degradation pathway can mediate this process. The BBSome functions as an adaptor for retrograde IFT at the base of the cilia, where it promotes the export of proteins out of the cilia into the cytoplasm where they can be targeted for degradation⁵². Ciliary proteins, destined to be removed from cilia, are transported to a proteosomal protein enriched area around the basal body. These proteins are then tagged with polyubiquitin chains and degraded by the proteasome⁵⁵.

1.1.2.1 Cilia in signalling

Primary cilia are sensory organelles that receive and transduce signals from the extracellular environment. Multiple receptors and ion channels are situated in the ciliary membrane and detect chemical messengers or mechanical stimuli to initiate a signalling cascade of signalling pathways such as Hedgehog, Wnt, mTOR, Notch, Hippo, JNK, FGF, PDGF and TGF β ³⁸.

1.1.2.1.1 Sonic Hedgehog pathway (SHH)

Probably the best-described pathway connected to the primary cilia is the Sonic Hedgehog (Shh) pathway. The Hedgehogs are a family of secreted proteins that participate in fate determination of the adjacent cells⁵⁶. Hedgehog (Hh) was identified in *D. melanogaster* genetic screens, and flies null for *hh* showed disrupted body hair pattern⁵⁷. Detailed analysis in mammalian cells showed three homologues of Hedgehog – Sonic (Shh), Indian (Ihh) and Desert (Dhh), all taking part in the organisation of body development^{56,58-60}. Shh is responsible for correct patterning of the limb bud, and it is also secreted from the notochord to the floor plate of the neural tube where it diffuses in a gradient to determine neuronal subtype identity. Ihh is partly redundant to Shh, and mediates cartilage and bone development. Dhh is essential for peripheral nerve sheath formation and the development of germ cells in testis.

The transduction of a signal from the extracellular environment to its downstream effectors in vertebrates is strictly regulated by the primary cilium. In normal conditions, in the presence of the cilium, the Shh morphogen binds to Patched (Ptch) receptors localised on the apical plasma membrane of cells and in the cilium⁶¹. The Ptch receptor represses the Smoothened (Smo) receptor, a G-protein-coupled seven transmembrane domain receptor, inhibiting its translocation

to the cilium. Upon binding Shh, Smo is released, allowing it to translocate to the cilium where it participates in the activation of the transcription mediator Gli⁶². Active Gli translocates to the nucleus where it acts as a transcription activator for the downstream effectors of Shh signalling. In the event of ciliary absence, Smo is unable to be translocated and the active form of Gli is not produced, therefore transcription of the downstream effectors is not activated.

Disruptions in Shh signalling causes defects in the neural tube development, defects in the mid-line often presented as holoprosencephaly, and pre-axial polydactyly⁶². Shh signalling is also crucial for stem cell homeostasis and its overexpression was noted in many cancer types including medulloblastoma and basal cell carcinoma⁶³. Somatic and germline mutations in Shh components cause cancer, for example Gorlin syndrome (an inherited form)⁶⁴ and sporadic forms of ependymoma⁶⁵.

1.1.2.1.2 Wnt

Wnt signalling was first investigated in *D. melanogaster*: flies that were null for *wnt* had under-developed wings, which led to their name - 'wingless' mutants. Multiple *wnt* genes were identified, along with their various functions and degree of importance, with some redundant to the other. It also became apparent that there are two distinguishable pathways in mammals driven by Wnt proteins: the canonical and non-canonical pathways. Wnt signalling was shown to be essential in embryonic development and adult tissue maintenance. It controls cell proliferation, cell fate and cell death⁶⁶.

1.1.2.1.2.1 Canonical Wnt signalling

In the absence of the Wnt ligand, the so-called "destruction complex" that includes axin, Adenomatous Polyposis Coli (APC), casein kinase 1 α (CK1 α) and GSK3 β is formed. Cytoplasmic β -catenin is constitutively recruited to the destruction complex, phosphorylated by CK1 α and GSK3 β , and this then drives its ubiquitination and subsequent proteolysis by the proteasome.

In the presence of the canonical Wnt ligand, Wnt binds to the Frizzled (Fz) seven transmembrane receptor and LRP5/6 co-receptors which drive recruitment of Dishevelled (Dvl) to the plasma membrane. Dvl binds Fz and this promotes phosphorylation of LRP by GSK3 β and CK1 γ , which in turn leads to axin binding and disassembling of the destruction complex. Non-phosphorylated β -catenin is then retained, enabling it to translocate to the nucleus and bind TCF/LEF proteins. These then act as transcription activators at the promoter sites of Wnt downstream effectors, such as cyclin D1, activating their transcription.

1.1.2.1.2.2 *Non-canonical Wnt signalling*

The non-canonical Wnt ligand binds the Frizzled receptor and Knypek co-receptor that triggers recruitment of Dvl to the plasma membrane. Dvl binds to the Fz receptor via the DEP domain, causing recruitment of the scaffolding protein Daam1 and activation of small GTPases such as RhoA. Alternatively, Dvl can also stimulate calcium flux and activation of calmodulin-dependent protein kinase II (CamKII) and calcium-sensitive kinases such as protein kinase C (PKC).

The non-canonical Wnt signalling is crucial in correct convergent-extension movement during gastrulation, regulation of dorsal axis formation, and the determination of ventral cell fate and heart formation ⁶⁶.

1.1.2.1.3 **Other signalling pathways linked to the cilium**

The role of cilia in mediating Wnt and Shh signalling has been the most extensively researched, but other signalling pathways are now emerging that have a requirement for cilia. These include the mTOR, Notch, Hippo, JNK, FGF, PDGF and TGF β pathways.

mTOR (mammalian target of rapamycin) pathway regulates cell size, proliferation, autophagy, protein translation, apoptosis and the actin cytoskeleton ⁶⁷. The presence of the primary cilium was shown to be necessary for mTOR signal transduction ^{68,69}. Treatment with rapamycin, an inhibitor of mTOR, caused reduction in number and size of kidney cysts, in animal models of ciliopathies. This suggests that defects in the mTOR pathway are an underlying cause of this common ciliopathy phenotype ⁷⁰.

Another explanation for the cystic kidney phenotype was researched by Du et al. in *Tmem67* mutated organisms. TMEM67 protein was shown to be required for cilia formation ⁷¹. Du et al. showed that cells overexpressing TMEM67 and *Tmem67* null mouse tissues present an activation of extracellular signal-regulated kinase (ERK) and c-jun N-terminal kinase (JNK) in Du et al. ⁷². They suggested that the cyst phenotype might be caused through ERK- and JNK-dependent signalling pathways.

The cilium also has essential functions during early embryogenesis and in the establishment of normal fetal growth. As discussed above (section 1.1.1), mammalian body symmetry is determined by the correct distribution of morphogens at the embryonic node by nodal cilia. Notch signalling components are localised to this cilium ⁷³ including the main effector, Nodal, a crucial left-side determinant ⁷⁴. Defects in Notch signalling were shown to cause laterality defects including congenital heart disease such as heterotaxia ⁷⁵. Limb development is also determined at this embryonic stage and depends on correct Shh and FGF signalling

^{76,77}. The latter is also required for the correct development of craniofacial structures ⁷⁸.

The Hippo pathway has been shown to be important in cell proliferation and organ size determination ⁷⁹. Recently, ciliopathy proteins such as NPHP3, NPHP4 and NPHP9 (also known as NEK8), were implicated in this pathway ⁸⁰⁻⁸². NPHP4 was identified as a negative regulator of the Hippo pathway, whilst NPHP3 and NPHP9 not only directly interacted with TAZ (transcriptional coactivator with PDZ-binding domain), a Hippo pathway effector, but also activated the whole pathway. Cell proliferation controlled by NPHP3, NPHP4 and NPHP9 was therefore suggested to be regulated by the Hippo pathway.

The presence of cilia is now thought to be required for TGF β signalling. For the TGF β signal to be transduced, clathrin-dependent endocytosis (CDE) of the receptor is thought to be necessary. CDE was proposed to be regulated by localisation to the ciliary pocket, a plasma membrane invagination at the base of the cilium ⁸³. TGF β receptors were shown to be present at the ciliary tip and the downstream effectors, such as SMAD2/3 and ERK1/2, were localized to the ciliary base ⁸³. The latter effector is shared with another pathway mediated by cilia: the Platelet-Derived Growth Factor (PDGF) signalling pathway. Platelet-Derived Growth Factor Receptor alpha (PDGFR α) is up-regulated and localised to the primary cilium, where the signal cascade is activated, initiating the AKT and ERK1/2 signalling cascades ⁸⁴, required for cell homeostasis.

1.2 Ciliopathies

Defects in the structure and/or function of cilia cause a suite of congenital developmental disorders collectively called the “ciliopathies”. They often present with polycystic kidney disease and other multi-organ phenotypes including central nervous system (CNS), eye and skeleton ⁸⁵. Many conditions are now classified as ciliopathies, such as Meckel-Gruber (MKS) or Joubert (JBTS) syndromes (sections **1.2.2.1** and **1.2.2.2**). Diseases that have some characteristic features of the classical ciliopathies, but unknown molecular causes of the phenotype or known links to cilia, are classified as “ciliopathy-like”. Ciliopathies are rare conditions but collectively, including polycystic kidney disease, they are quite common with a prevalence of 1 in 2000 births ²¹, highlighting the significance of cilia in human health and disease.

Many proteins that are mutated in ciliopathies are localized to the TZ^{86,87}. In particular, a protein complex at the TZ, known as the “MKS-JBTS module”, contains many of the proteins mutated in MKS and JBTS^{47,88}.

1.2.1 Motile cilia

Primary Ciliary Dyskinesia – PCD (MIM#244400) is an autosomal recessive heterogenous condition affecting motile cilia in respiratory tract, middle ear and oviducts. Its main phenotypic consequence is disrupted mucus movement and, in consequence, susceptibility to chronic recurrent respiratory infections. In combination with *situs inversus*, sinusitis and bronchiectasis, it is known as Kartagener syndrome^{89,90}. The phenotype is caused by defects in the main structural components of the ciliary axoneme including outer dynein arms (*DNAI1*, *DNAH5*, *DNAI2*, *TXNDC3*, *DNAL1*, *CCDC114*, *ARMC4*, *CCDC151*, *CCDC103*), inner and outer dynein arms (*KTU*, *LRRC50*, *DNAAF3*, *DYX1C1*, *HEATR2*, *LRRC6*, *ZMYND10*, *SPAG1*, *C21orf59*), microtubule disorganisation with inner dynein arm (*CCDC39* and *CCDC40*), nexin-dynein complex (*CCDC164* and *CCDC65*), microtubule disorganisation with central pair and radial spokes (*RSPH4A*, *RSPH9*, *RSPH1*), central microtubule pair (*HYDIN*) and no phenotype observed (*DNAH11*)⁹¹.

1.2.2 Primary cilia

In other ciliopathies, the structural as well as functional components of the primary cilia are affected. There is enormous heterogeneity in the “primary cilium ciliopathies” with an ever-increasing number of mutated genes being reported. Many phenotypes of previously unknown aetiology are now being classified as “ciliopathies”, following a plethora of research into protein function, cilia structure and the role of cilia in signalling pathways. In particular, the medical importance of the TZ is now clear. This sub-ciliary compartment acts as the ciliary gate (section 1.1.1.1.2.2), regulating the entry and exit of components to the cilium, and is a hub of proteins mutated in ciliopathies^{5,47}.

Genetic heterogeneity and pleiotropy causing both phenotypic overlap but also variable expressivity between various ciliopathies exists, yet the mechanistic basis of phenotypic variation has proven difficult to explain (**Figure 1-4**). Mutations in the same gene, for example *CEP290*, were shown to cause a relatively mild phenotype as Leber congenital amaurosis (LCA), through to nephronophthisis (NPHP) and severe, lethal MKS, with no clear correlation between the phenotype and genotype. Uncovering the genetic basis of these conditions, as well as

understanding the molecular function of the encoded protein, would benefit patients and their families, allowing better prognosis, diagnosis and treatment of the patients. The major interest in this thesis is in the genetic basis and mechanisms of disease pathogenesis of MKS and JBTS.

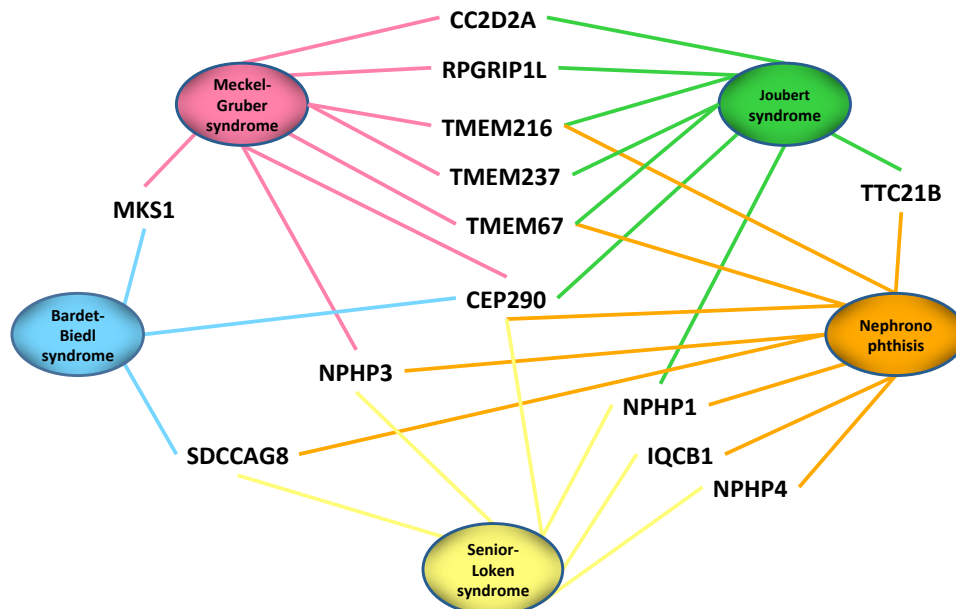


Figure 1-4. Schematic representation of allelism in ciliopathies. Mutations in the same gene can cause multiple phenotypes, represented by the coloured lines. For example mutations in *CEP290* can cause MKS, JBTS, Bardet-Biedl (BBS), Senior-Loken syndromes (SLS) or NPHP. Image adapted from Szymanska and Johnson, 2012.

1.2.2.1 Meckel - Gruber syndrome (MKS) (MIM#249000)

MKS is an autosomal recessive lethal condition characterised by central nervous system (CNS) defects including occipital meningoencephalocele (**Figure 1-5a, b and e**). Other features required for minimal diagnosis include bilateral polycystic kidneys (**Figure 1-5c and f**) and ductal plate malformation of the liver, leading to hepatic fibrosis and cysts (**Figure 1-5d**)⁹². Other frequently observed features may include post-axial polydactyly (**Figure 1-5a-b**), Dandy-Walker malformation (or other posterior fossa defects), bowing of long bones, cleft lip and/or palate, low set ears, microphthalmia, dextrocardia, *situs inversus* and iris coloboma⁹³⁻⁹⁵.

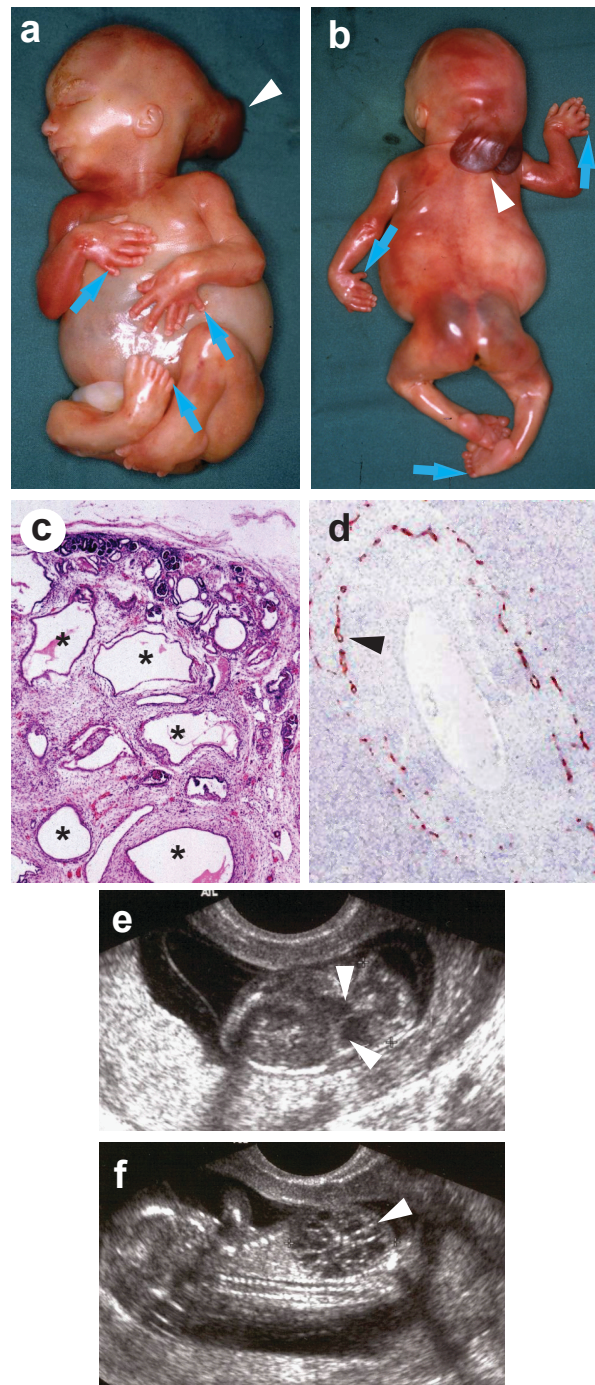


Figure 1-5. Clinical features of MKS. a) and b) typical external features for a fetus with MKS at 16 weeks of gestation showing massive flank masses due to bilateral renal cystic dysplasia, encephalocele (arrowheads) and post-axial polydactyly of all limbs (blue arrows); c) H&E staining of MKS fetal kidney at gestation age 18+/40 showing cystic dysplasia, comprising large, fluid-filled cysts (asterix), small cysts and cystic swelling of the proximal tubules and glomeruli, with the absence of normal renal parenchyma; d) IHC staining of MKS fetal liver at gestation age 18+/40 for cytokeratin-19, showing the retention of embryonic bile duct structures (the ductal plate malformation, arrowhead) without the formation of patent bile ducts; e) and f) ultrasound findings at 14+/40 weeks of gestation for MKS showing e) encephalocele (arrowheads) and f) large cystic kidneys (arrowhead). Images are used by kind permission of Dr Riitta Salonen (Rinnekoti Foundation, Helsinki, Finland) from the Robert J. Gorlin Slide Collection.

Johann F. Meckel first reported this condition in 1822 and subsequently C.B. Gruber in 1934⁹³. It was not until decades later that the condition was confirmed to be inherited in an autosomal recessive manner⁹⁶, and even later to identify the first causative mutations in genes^{97,98}. There are twelve genes identified to cause this syndrome to date (**Table 1-1**)⁹⁷⁻¹⁰⁸, but at the beginning of the research project described in this thesis only nine were reported. The extreme genetic heterogeneity of this syndrome has made the delineation of genotype-phenotype correlations difficult. Mutations in *TMEM67* were associated with liver phenotypes¹⁰⁹⁻¹¹¹, whilst polydactyly and occipital encephalocele were more often observed in patients with *MKS1* mutations¹¹². The incidence of MKS varies between 1:13,250 and 1:140,000 live births and depends on the ethnic background of the population, with Finnish and Gujarati-Indian populations having the highest incidence rate⁹⁵.

LOCUS	GENE	ENTREZ GENE ID	ALIASES	LOCATION	FOUNDER MUTATION	REFERENCE
MKS1	<i>MKS1</i>	54903	BBS13, MES, MKS, POC12	chr17:58205436-58219605	Finnish c.1408-35_1408-7del29	Kyttala et al. 2006
MKS2	<i>TMEM216</i>	51259	HSPC244	chr11:61392360-61398863	Ashkenazi p.R73L	Valente et al. 2010
MKS3	<i>TMEM67</i>	91147	JBTS6, MECKELIN, MKS3, NPHP11, TNEM67	chr8:93754844-93819234	Pakistani c.1575+1G>A	Smith et al. 2006
MKS4	<i>CEP290</i>	80184	3H11Ag, BBS14, CT87, JBTS5, LCA10, MKS4, NPHP6, POC3, SLSN6, rd16	chr12:88049013-88142216		Balaa et al. 2007
MKS5	<i>RPGRIP1L</i>	23322	CORS3, FTM, JBTS7, MKS5, NPHP8	chr16:53599906-53703859	Mixed European p.T625P	Delous et al. 2007
MKS6	<i>CC2D2A</i>	612013	JBTS9, MKS6	chr4:15469865-15601971	Finnish c.1762C>T	Tallila et al. 2008
MKS7	<i>NPHP3</i>	27031	MKS7, NPH3, RHPD, RHPD1, SLSN3	chr3:132680609-132722459		Bergmann et al. 2008
MKS8	<i>TCTN2</i>	79867	C12orf38, MKS8, TECT2	chr12:123671113-123708403		Shaheen et al. 2011
MKS9	<i>B9D1</i>	27077	B9, EPPB9, MKS9, MKSR1	chr17:19337250-19378182		Hopp et al. 2011
MKS10	<i>B9D2</i>	80776	ICIS-1, MKS10, MKSR2	chr19:41354417-41364540		Dowdle et al. 2011
MKS11	<i>TMEM231</i>	79583	UNQ870/PRO1886, ALYE870, JBTS20, MKS11, PRO1886	chr16:75538117-75556286		Shaheen et al. 2013
MKS12	<i>CSPP1</i>	79848	CSPP, JBTS21	chr8:67055392-67196614		Shaheen et al. 2014

Table 1-1. List of genes identified with pathogenic mutations causing the MKS phenotype. Twelve loci with identified genes are reported. The columns give the NCBI Entrez gene ID, other names, physical location (hg19), any reported founder mutations and key reference.

1.2.2.1.1 Progress in gene identification

The first MKS locus, *MKS1*, was mapped to chromosome 17q21-q24¹¹³ in endogamous Finnish families using a combination of homozygosity mapping and haplotype analysis. Homozygosity mapping also identified loci (*MKS2* and *MKS3*) on chromosomes 11q13¹¹⁴ and 8q24¹¹⁵ respectively, in consanguineous families from the Middle East and the Indian sub-continent. In 2006, Kyttälä et al. identified mutations in the *MKS1* gene as a cause of MKS in the Finnish population⁹⁸, whilst Smith et al. identified mutations in *TMEM67*⁹⁷ in the *MKS3* locus that encodes the *TMEM67/meckelin* orphan receptor (**Table 1-1**). Subsequently, homozygosity mapping identified point mutations in *CEP290*¹⁰⁰ and *RPGRIP1L*¹⁰¹ as causes of MKS, as well as microdeletions in *CEP290*¹¹⁶. The identification of mutations in *CC2D2A* for Finnish MKS families (excluded for mutations in *MKS1*), provided an

unusual example of a second major cause for a Finnish heritage disease ¹⁰². In contrast, microheterogeneity at the *MKS2* locus for two adjacent genes that both encode tetraspanin-like transmembrane proteins (*TMEM138* and *TMEM216*), prevented the identification of mutations in *TMEM216* as a cause of MKS until 2010 ⁹⁹. Interestingly, although mutations in *TMEM216* are allelic for both JBTS and MKS, mutations in *TMEM138* have not been described as a cause of MKS. Conversely, *MKS1* truncating mutations have only been described as a cause of MKS, whereas missense or hypomorphic mutations are a cause of JBTS ¹¹⁷.

Subsequent gene discovery studies for MKS have used whole exome sequencing (WES) to prioritize functional candidate genes, often based on sequence homology to known MKS or JBTS genes. For example, *TCTN1*, a known JBTS gene, was used to prioritize screening and then to identify a pathogenic private mutation in *TCTN2*, a paralogue and member of the Tectonic family of genes ¹⁰⁴. Furthermore, the MKS1 protein contains a B9 domain of unknown function that is also present in only two other proteins in the human genome (*B9D1* and *B9D2*). Mutations in *B9D1* were identified as a cause of MKS ¹⁰⁵, followed by the description of a family with a pathogenic private mutation in *B9D2* ¹⁰⁶. The most recently identified MKS genes include *TMEM231* ¹⁰⁷ and *CSPP1* ¹⁰⁸.

1.2.2.1.2 Molecular characterisation of MKS proteins

The protein products of the MKS genes have all been shown to be structural and functional components of the primary cilium. Most of these proteins localise to the TZ. Elaborate networks of protein-protein interactions have been described for these proteins, showing direct interactions with each other as well as with other ciliopathy proteins. This may suggest a hierarchy in TZ proteins, that could provide further insights into the mechanisms of variable ciliopathy expressivity. This is discussed in further detail in section **1.2.2.2.2**.

1.2.2.2 Joubert syndrome (JBTS) (MIM#614464)

JBTS is an autosomal recessive condition with a characteristic neurodevelopmental phenotype of the “molar tooth sign” (MTS) observed on axial magnetic resonance imaging (MRI) scans. This phenotype consists of cerebellar vermis hypoplasia or aplasia, abnormality of the fourth ventricle, elongated superior cerebellar peduncles and a deep interpeduncular fossa (**Figure 1-6a** and **b**). Together with developmental delay, it is necessary for JBTS diagnosis. Additional

features are reported in the affected patients and those include irregular breathing pattern or apnea, polydactyly, oculomotor apraxia and ataxia.

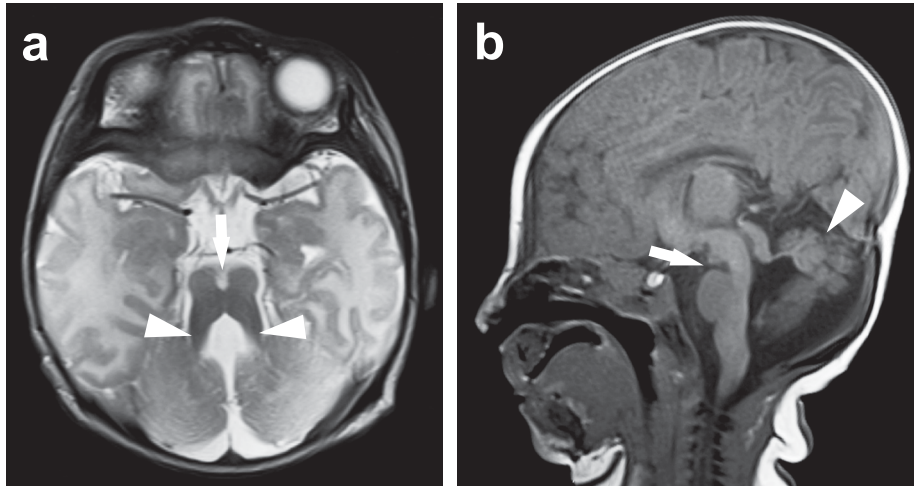


Figure 1-6. Clinical features of JBTS. a) axial T2-weighted (MRI) scan in a 5-year-old girl showing typical brain anomalies of JBTS, including a deepened interpeduncular fossa (arrow) and elongated superior cerebellar peduncles (arrowheads) that comprise the pathognomonic “molar tooth sign”, as well as cerebellar vermis hypoplasia; b) a sagittal T1-weighted MRI scan showing the deepened interpeduncular fossa (arrow) and cerebellar vermis hypoplasia (arrowhead). The cerebellar hemispheres and brainstem are also hypoplastic. Images are used by kind permission of Dr Daniel Doherty (Seattle Children's Hospital, University of Washington, USA).

Marie Joubert first described this condition in 1969 ¹¹⁸. It is a rare condition with an estimated prevalence of 1:100,000 ¹¹⁹. The pronounced phenotypic variability observed in this syndrome has caused it to be classified into classic JBTS and “Joubert syndrome-related disorders” (JSRD). The latter phenotype, in addition to the obligatory clinical features, can also include renal cysts, hepatic fibrosis, polydactyly, coloboma, retinal dystrophy and short oral frenula ¹²⁰. Another subclass of JBTS is COACH syndrome, characterised by the combination of Coloboma, Oligophrenia (mental retardation), Ataxia, Cerebellar vermis hypoplasia and Hepatic fibrosis ¹²¹. Additional ocular and renal phenotypes are also described as CORS (cerebellar-ocular-renal sndrome) ¹²², but the preferred nomenclature is now JSRD. Just as for MKS, JBTS also has extreme genetic heterogeneity with 22 genes identified to date (**Table 1-2**). All proteins encoded by these genes localise to the primary cilium, where they play crucial structural and functional roles.

LOCUS	GENE	ENTREZ GENE ID	ALIASES	LOCATION	FOUNDER MUTATION	REFERENCE
JBTS1	<i>INPP5E</i>	56623	CORS1, CPD4, JBTS1, MORMS, PPI5PIV	chr9:136428615-136439853		Bielas et al. 2009
JBTS2	<i>TMEM216</i>	51259	HSPC244	chr11:61392360-61398863	Ashkenazi p.R73L	Valente et al. 2010
JBTS3	<i>AHI1</i>	54806	RP1-32B1.2, AHI-1, JBTS3, ORF1, dJ71N10.1	chr6:135283532-135497765		Ferland et al. 2004
JBTS4	<i>NPHP1</i>	4867	JBTS4, NPH1, SLSN1	chr2:110123336-110205062		Parisi et al. 2004
JBTS5	<i>CEP290</i>	80184	3H11Ag, BBS14, CT87, JBTS5, LCA10, MKS4, NPHP6, POC3, SLSN6, rd16	chr12:88049013-88142216		Sayer et al. 2006
JBTS6	<i>TMEM67</i>	91147	JBTS6, MECKELIN, MKS3, NPHP11, TNEM67	chr8:93754844-93819234		Balaa et al. 2007
JBTS7	<i>RPGRIP1L</i>	23322	CORS3, FTM, JBTS7, MKS5, NPHP8	chr16:53599906-53703859		Delous et al. 2007
JBTS8	<i>ARL13B</i>	200894	ARL2L1, JBTS8	chr3:93980139..94055678		Cantagrel et al. 2008
JBTS9	<i>CC2D2A</i>	612013	JBTS9, MKS6	chr4:15469865-15601971		Noor et al. 2008
JBTS10	<i>OFD1</i>	8481	71-7A, CXorf5, JBTS10, RP23, SGBS2	chrX:13715430-13769361		Coene et al. 2009
JBTS11	<i>TTC21B</i>	79809	Nbla10696, ATD4, IFT139, JBTS11, NPHP12, SRTD4, THM1	chr2:165873362-165953838		Davis et al. 2011
JBTS12	<i>KIF7</i>	374654	UNQ340/PRO539, ACLS, HLS2, JBTS12, UNQ340	chr15:89617944-89655494		Dafinger et al. 2011
JBTS13	<i>TCTN1</i>	79600	UNQ9369/PRO34160, JBTS13, TECT1	chr12:110614027-110649130		Garcia-Gonzalo et al. 2011
JBTS14	<i>TMEM237</i>	65062	ALS2CR4, JBTS14	chr2:201620184-201643529		Huang et al. 2011
JBTS15	<i>CEP41</i>	95681	JBTS15, TSGA14	chr7:130393771-130441210		Lee et al. 2012
JBTS16	<i>TMEM138</i>	51524	HSPC196	chr11:61362001-61369505		Lee et al. 2012
JBTS17	<i>C5orf42</i>	65250	JBTS17	chr5:37067870-37249428	French Canadian p.Arg1336Trp, p.Ala1564Thr, c.7400+1G>A	Srouf et al. 2012
JBTS18	<i>TCTN3</i>	26123	RP11-7D5.3, C10orf61, JBTS18, OFD4, TECT3	chr10:95663396-95694143		Thomas et al. 2012
JBTS19	<i>ZNF423</i>	23090	Ebfaz, JBTS19, NPHP14, OAZ, Roaz, ZFP423, Zfp104	chr16:49490604-49857919		Chaki et al. 2012
JBTS20	<i>TMEM231</i>	79583	UNQ870/PRO1886, ALYE870, JBTS20, MKS11, PRO1886	chr16:75538117-75556286		Srouf et al. 2012
JBTS21	<i>CSPP1</i>	79848	CSPP, JBTS21	chr8:67055392-67196614		Tuz et al. 2014
JBTS22	<i>PDE6D</i>	5147	JBTS22, PDED	chr2:231732437-231781264		Thomas et al. 2014

Table 1-2. List of genes identified with pathogenic mutations causing the JBTS phenotype. Twenty two loci with identified genes are reported. The columns give the NCBI Entrez gene ID, other names, physical location (hg19), any reported founder mutations and key reference.

1.2.2.2.1 Progress in gene identification

JBTS was proposed to be an autosomal recessive disorder in 1977⁹⁶. The first JBTS locus was reported by Natacci et al.¹²³, who identified a patient affected with JBTS and a deletion at chromosome 17p11.2. Subsequently, loci were mapped to chromosome 9q¹²⁴, 6q23¹²⁵ and 3q24¹²⁶ using the specialised linkage analysis technique of homozygosity mapping (section 1.2.2.7) in extended consanguineous families. The locus for a variant form of JBTS phenotype that included retinal dysplasia and cystic kidneys (known as "JBTS type B", or "cerebello-oculo-renal syndrome"; CORS2) was initially mapped to 11p12-q13.3¹²⁷. The breakthroughs in gene identification came in 2004, when Parisi et al. identified a deletion involving *NPHP1* in a patient affected with JBTS¹²⁸ as well as nephronophthisis (a hereditary kidney disease now also described as a ciliopathy) (Table 1-2). In the same year, Ferland et al.¹²⁹, investigated JBTS patients with the classical phenotype and linkage to 6q23.2-q23.3, and identified pathogenic mutations in the *AHI1* gene encoding the protein called joubertin.

Subsequent genetic studies have provided many seminal insights into the JBTS phenotype and ciliopathies. In 2006, Valente et al.¹³⁰ and Sayer et al.¹³¹ identified mutations in *CEP290* as a cause of JBTS and substantiated the importance of primary cilia dysfunction. Allelism between JBTS and MKS was

demonstrated in 2007 with the identification of mutations in *TMEM67* as a cause of JBTS¹³², as well as the identification of mutations in *RPGRIP1L* to be a cause of both JBTS and MKS^{101,133}. Mutations in *ARL13B*, *CC2D2A* and *INPP5E* were shown to cause classical JBTS¹³⁴⁻¹³⁶. *ARL13B* is a small Arf-family GTPase and *INPP5E* encodes an inositol 1,4,5-trisphosphate (InsP3) 5-phosphatase. InsP3 5-phosphatases hydrolyze InsP3, which acts as a secondary messenger to mobilize calcium from intracellular stores. These findings implicate the dysregulation of embryonic signalling pathways as a cause of the ciliopathy phenotype (section **1.2.2.2.2**). Further allelism between an unusual X-linked form of JBTS¹³⁷ and oro-facial-digital (OFD) syndrome was demonstrated by identification of mutations in *OFD1*¹³⁸. Affected males presented with a typical JBTS phenotype, but with additional features, including coloboma, and without the typical MTS.

In recent years, advances in genetic technology (principally, WES) have enabled a renaissance in gene discovery, often without initial linkage analysis, that has now enabled the study of smaller, non-consanguineous families. A key biological insight from these studies has been the importance of the Tectonic (TCTN) family of transmembrane proteins and small tetraspanin-like transmembrane proteins (TMEMs) in the pathogenesis of the JBTS phenotype. Initial WES studies identified the JBTS genes *TMEM216*^{99,139}, *KIF7*¹⁴⁰, *TCTN1*^{47,141}, *CEP41*¹⁴² and *TMEM138*¹⁴³. More recent studies have used WES to identify mutations in *TMEM237*¹⁴⁴ and *C5ORF42* as a common cause in French-Canadian JBTS patients¹⁴⁵. *C5ORF42* was also reported to be mutated in a cohort of Saudi-Arabian JBTS patients¹⁴¹. The most recently identified JBTS genes include *EXOC8*¹⁴⁶, *TCTN3*¹⁴⁷, *TMEM231*¹⁴⁸, *PDE6D*¹⁴⁹, *CSPP1*^{108,150,151} and *TCTN2*¹⁵².

1.2.2.2.2 Molecular mechanisms of JBTS and MKS proteins

Following the initial gene discovery studies, contemporary research has focused on the delineation of possible cellular functions for JBTS and MKS ciliary proteins. Several recent studies have used biochemical assays and proteomic studies to delineate networks of protein-protein interactions and, in some instances, infer possible functions from those of other, better characterised members of a complex. Elegant transgenesis and localisation studies in animal models such as zebrafish (*Danio rerio*) and the nematode (*Caenorhabditis elegans*) have inferred genetic interactions between JBTS and MKS genes, although these do not always support the existence of biochemical interactions. However, in most of these studies the target ciliary protein is over-expressed either with a convenient epitope tag for biochemical purification, or a fluorescent protein reporter in genetic interaction experiments, and is not at physiological levels of expression. It is debateable if over-expressed proteins correctly model the localization and

interactions of the *in vivo* cognate protein, and these studies should be interpreted with some caution.

Despite these limitations, several studies have identified the protein complex known as the “MKS-JBTS module” at the ciliary TZ that contains many JBTS and MKS proteins ^{28,88}. These include all of the known small tetraspanin-like TMEMs mutated in ciliopathies (TMEM216, TMEM138, TMEM237 and TMEM231), as well as others (TMEM17 and TMEM107) that have yet to be implicated in a human ciliopathy (**Figure 1-7**). The Tectonic transmembrane proteins (TCTN1, TCTN2 and TCTN3) and the orphan receptor TMEM67/meckelin also localize to the TZ. All of these transmembrane proteins are thought to be inserted into the local plasma membrane at the TZ. This, in turn, may mediate interactions or regulate the function of membrane-targeting proteins with C2 domains such as RPGRIP1L and CC2D2A ^{28,88}. For example, a proteomics study of TCTN1 showed that this protein forms a complex at the TZ with MKS1, TMEM216, TMEM67, CEP290, B9D1, TCTN2 and CC2D2A ⁴⁷. Confirmation of the direct interaction between CC2D2A and CEP290 has come from yeast two-hybrid studies and GST pull-downs, with genetic interactions between these genes also demonstrated in zebrafish models ¹³⁵.

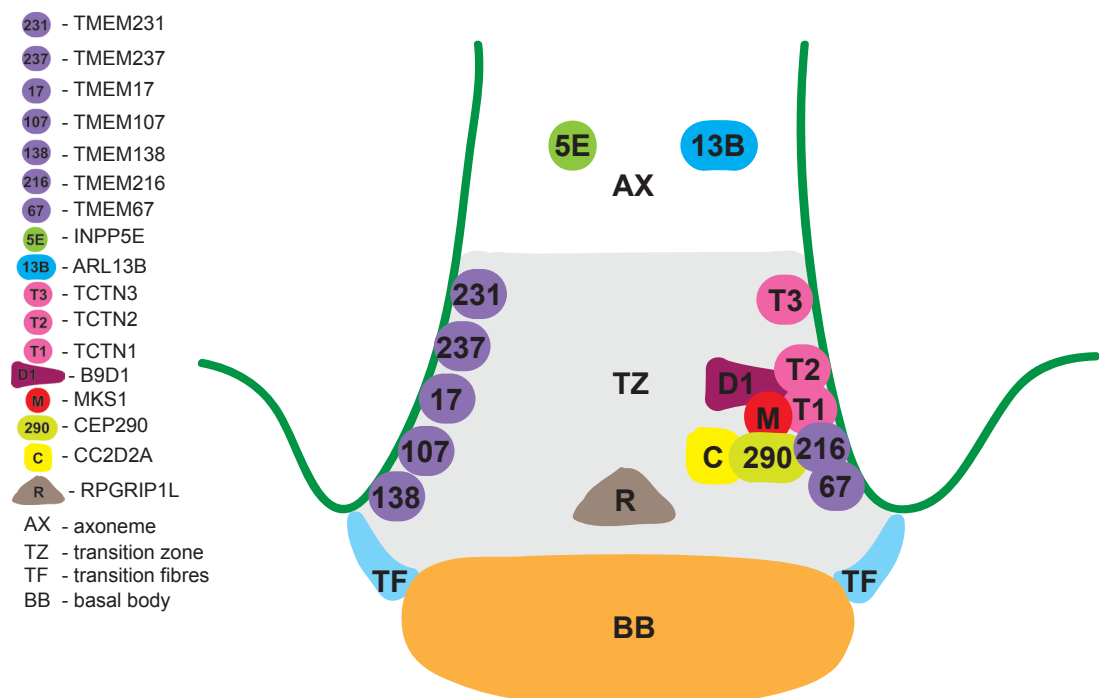


Figure 1-7. Schematic illustration of MKS/JBTS protein localisation in cilia. Transmembrane proteins including TMEM17, 67, 107, 138, 216, 231 and 237 that are localised to the TZ membrane as well as Tectonic proteins (TCTN1, 2 and 3). The cartoon focuses on the biochemical interactions of TCTN1 described in Garcia-Gonzalo et al. 2011.

The role of CEP290 and the B9 proteins at the TZ remains unclear, but they presumably act as linkers or mediators between the TMEMs and either the vesicular cargo that is targeted to the TZ during ciliogenesis, or the subsequent transport of cargo proteins within the cilium by IFT. The former model is supported by analogy from studies on AHI1, a JBTS ciliary protein that contains WD40 and Src Homology 3 (SH3) protein interaction domains. AHI1/joubertin interacts with RAB8A, a small GTP/GDP-binding protein that mediates the vesicular transport of proteins from the endoplasmic reticulum to the Golgi and the plasma membrane ¹⁵³. Loss of AHI1/joubertin causes defects in ciliogenesis and also defects in vesicle transport ¹⁵³.

The elaborate machinery of the TZ appears to be in place to maintain the integrity of the ciliary gate, and the disruption of this function may contribute to the ciliopathy phenotype by altering the composition of the ciliary membrane or axoneme ⁵. Several recent studies have confirmed that disruption of the TZ results in the incorrect movement of proteins in and out of the cilium, with particular emphasis on the transport of enzymes and transmembrane proteins that mediate intracellular signalling. For example, TMEM231 and TMEM17 localize to the TZ, with the localization of TMEM231 dependent on other TZ proteins (CC2D2A and B9D1) ¹⁵⁴. In turn, this regulated the transport of G-protein-coupled signalling receptors (GPCRs; specifically, somatostatin and serotonin receptors SSTR3 and HTR6) respectively into the ciliary membrane ¹⁵⁴. Loss of either TMEM231 or B9D2 in mice caused defects in ciliogenesis and Shh signalling ¹⁵⁴. Sang et al. showed that loss of TCTN1, TCTN2, TMEM67 or CC2D2A caused tissue-specific defects in ciliogenesis and ciliary membrane composition ¹⁵². In a separate study, tissues that lacked TCTN1 had normal ciliogenesis, but nevertheless had defects in the ciliary localization of adenylyl cyclase 3 (ACIII, an enzyme that catalyzes the formation of the secondary messenger cAMP), the transmembrane signalling proteins Smoothed and PKD2/polycystin-2 and ARL13B (a small Arf-family GTPase) ⁴⁷. Interestingly, ARL13B regulates the migration of interneurons in the developing brain, and this may provide a partial explanation for the neurological defects observed in JBTS patients ¹⁵⁵.

In contrast to other JBTS and MKS proteins that predominantly localise to the ciliary TZ, both ARL13B and INPP5E localise to the ciliary axoneme. The ciliary localisation of ARL13B depends on TZ function, as described above, but the axonemal localisation of INPP5E depends on the functions of both ARL13B and PDE6D (a phosphodiesterase that appears to act as a chaperone for prenylated ciliary and retinal proteins) ^{47,149,156}. An attractive hypothesis is that INPP5E dysfunction (due to either mutation or mislocalisation) causes alterations in ciliary

signalling through changes in the levels of the secondary messenger InsP3. It is probable that this is one of the fundamental pathogenic mechanisms in both JBTS and MKS, but it has yet to be formally tested. The JBTS phenotype has also been associated with alterations in the tubulin post-translational modifications of the ciliary axoneme, and this would presumably affect both the stability of the cilium and the trafficking of ciliary proteins. Specifically, the JBTS protein CEP41, encoding a centrosomal and microtubule-binding protein, regulates the entry of TLL6 (a tubulin polyglutamylase enzyme) to the ciliary axoneme¹⁴². This thus implicates tubulin post-translational modification and therefore the composition of the axoneme in the ciliopathy phenotype^{142,157}.

Finally, the loss or mislocalisation of many JBTS and MKS proteins cause defects in Shh signalling (section 1.1.2.1.1), as demonstrated in a number of mouse ciliopathy models^{140,154,158-160}. One explanation is that TZ disruption prevents the correct trafficking of KIF7, a ciliary-associated kinesin motor protein that regulates Shh signalling through altering the relative levels of the activator and repressor isoforms of GLI transcription factors^{140,161}. GLI proteins are the downstream effectors of the Shh signalling pathway, and KIF7 appears to act as both a negative regulator by preventing the inappropriate activation of GLI2 in the absence of ligand, and as a positive regulator by preventing the processing of GLI3 into its repressor form. Animal models with loss of TZ-associated JBTS and MKS proteins also have dysregulation of the “canonical” and “non-canonical” branches of the Wnt signalling pathway^{99,158,159,162,163}, but the mechanistic detail of how the ciliary TZ regulates this pathway is less clear than for the Shh pathway. One possibility, at least for the canonical β -catenin-mediated branch, is that AHI/jouberin directly interacts and sequesters the downstream effector β -catenin at the cilium^{164,165}. In turn, this would limit the nuclear entry of β -catenin and its availability to act as a transcription factor for Wnt-responsive genes.

1.2.2.3 Future perspectives for research into MKS and JBTS

To date, mutations in the known JBTS and MKS genes appear to account for no more than 60% of cases. The spectrum of causative genes for these conditions is therefore incomplete, but the remaining genes are likely to be uncommon or even harbour mutations that are private to a single family. However, with the widespread availability and affordability of WES and targeted clonal sequencing techniques such as MIPS (molecular inversion probe sequencing)¹⁶⁶, many researchers have recently reported rare mutations in genes in patients that have been previously identified as mutation negative. In addition, whole-genome sequencing (WGS) at low coverage depth now allows rapid copy number analysis¹⁶⁷. These studies are

likely to uncover copy number variations and intronic mutations, as well as changes in the promoter sequence or in *cis* regulatory elements as potential pathogenic causes. This will improve JBTS and MKS patient diagnosis, and, with the emerging genotype-phenotype correlations for JBTS variants such as COACH, some improvements in prognostic testing. However, as described above, the efforts to describe most of these correlations in anything but broad terms are confounded by both the allelism and unusual phenotypic variability for these conditions. Furthermore, the range of phenotypes associated with the ciliopathies continues to be broadened¹⁶⁸⁻¹⁷³. As many patients described in this work had either limited phenotype description or the phenotype was not clear JBTS, and for consistency purposes, these patients will be referred as JSRD.

1.2.2.4 Skeletal ciliopathies

1.2.2.4.1 Short-rib thoracic dysplasia (SRTD) (MIM#208500)

SRTD is a lethal, heterogeneous, autosomal recessive group of skeletal ciliopathies with presence/absence of polydactyly¹⁷⁴. It is characterised by short ribs, shortened tubular bones and constricted thoracic cage. For some types, the diagnosis is often based on the presence of 'trident' appearance of the acetabular roof in the hip bone, a feature that regresses with age¹⁷⁵. Additional features may include cleft lip/palate, retinal degeneration¹⁷⁶, hepatic fibrosis, bile duct proliferation^{177,178}, cysts in the kidneys, liver and pancreas¹⁷⁹ and hydrocephalus¹⁸⁰. Many of these features are shared with other ciliopathies and differential diagnoses for these conditions is frequent. SRTD encompasses many syndromes, including Jeune asphyxiating thoracic dystrophy (ATD), Ellis-van Creveld syndrome (EVC) (MIM#225500), short rib-polydactyly syndrome (SRPS), and Mainzer-Saldino syndrome (MZSDS). Genes mutated in this disease are reported to be involved in ciliary intraflagellar transport, such as *IFT80*¹⁸¹, *TTC21B*¹⁸², *IFT144*¹⁸³, *WDR35*¹⁸⁴, *IFT140*¹⁸⁵, *IFT172*¹⁸⁶; in ciliary transport motors, specifically *DYNC2H1*¹⁸⁷; and in ciliogenesis, including *WDR60*¹⁸⁸, *WDR34*¹⁸⁹ and *NEK1*¹⁹⁰.

1.2.2.4.2 Cranioectodermal dysplasia (Sensenbrenner syndrome) (CED) (MIM#218330)

Sensenbrenner syndrome is an autosomal recessive, heterogenous condition with a high phenotypic overlap with SRTD. This condition is characterised by sagittal craniosynostosis, short stature, dolichocephaly, frontal bossing, widely spaced teeth, narrow thorax, nephronophthisis, joint laxity, brachydactyly and sparse hair¹⁹¹⁻¹⁹⁴. Four genes have been implicated in this syndrome and all of

them are members of the intraflagellar transport machinery: *WDR35*¹⁸⁴, *IFT122*¹⁹⁴, *IFT144*¹⁸³ and *IFT43*¹⁹⁵.

1.2.2.5 Nephronophthisis (NPHP) (MIM#256100)

Nephronophthisis is an autosomal recessive, heterogenous condition characterised by cystic kidneys leading to renal failure in childhood or adolescence. It is the most common genetic cause of renal failure in children. Additional features include *situs inversus*, liver fibrosis or cardiac malformations. It phenotypically overlaps with other disorders and can be split on the basis of additional clinical features into syndromes with, for example, retinitis pigmentosa (SLS) and cerebellar vermis hypoplasia (JBTS). Conversely, nephronophthisis can be an occasional phenotypic feature of MKS.

This condition also displays extreme genetic heterogeneity with 18 genes identified to date (**Table 1-3**) and allelism with other ciliopathies (**Figure 1-4**). The encoded proteins localise to cilia, many to the TZ^{5,144,152}. Sang et al. showed NPHP2, NPHP3 and NPHP9 to form a complex and localise to a discrete compartment located distal to the TZ, which they called the inversin compartment¹⁵².

LOCUS	GENE	ENTREZ GENE ID	ALIASES	LOCATION	FOUNDER MUTATION	REFERENCE
NPHP1	<i>NPHP1</i>	4867	FLJ97602, JBTS4, NPH1, SLSN1	chr2:110237177-110319883	<i>NPHP1</i> deletion	Konrad et al. 1996
NPHP2	<i>INVS</i>	27130	RP11-208F1.1, INV, KIAA0573, MGC133080, MGC133081, NPH2, NPHP2	chr9:101901332-102103247		Otto et al. 2003
NPHP3	<i>NPHP3</i>	27031	DKFZp667K242, DKFZp781K1312, FLJ30691, FLJ36696, KIAA2000, NPH3	chr3:133882144-133923966		Olbrich et al. 2003
NPHP4	<i>NPHP4</i>	261734	RP11-33M12.1, KIAA0673, SLSN4	chr1:5845457-5975118		Otto et al. 2002
NPHP5	<i>IQCB1</i>	9657	NPHP5, PIQ, SLSN5	chr3:122971531-123036616		Otto et al. 2005
NPHP6	<i>CEP290</i>	80184	3H11Ag, BBS14, FLJ13615, FLJ21979, JBTS5, JBTS6, KIAA0373, LCA10, MKS4, NPHP6, SLSN6, rd16	chr12:86966921-87060124		Sayer et al. 2006
NPHP7	<i>GLIS2</i>	84662	NKL; NPHP7; FLJ38247; GLIS2	chr16:42382225		Attanasio et al. 2007
NPHP8	<i>RPGRIPL</i>	23322	CORS3, DKFZp686C0668, JBTS7, KIAA1005, MKS5, NPHP8	chr16:52191319-52295272	p.T615P	Wolf et al. 2007
NPHP9	<i>NEK8</i>	284086	JCK; NPHP9; NEK12A; MGC138445; NEK8	chr17:27055832		Otto et al. 2008
NPHP10	<i>SDCCAG8</i>	10806	BBS16; CCCAP; SLSN7; NPHP10; hCCCAP; HSPC085; NY-CO-8	chr1:243419307-243663393		Otto et al. 2010
NPHP11	<i>TMEM67</i>	91147	JBTS6, MECKELIN, MGC26979, MKS3, TNEM67	chr8:94836321-94898323		Otto et al. 2009
NPHP12	<i>TTC21B</i>	79809	ATD4; THM1; JBTS11; NPHP12; FLJ11457; Nbla10696	chr2:166730453-166810348		Davis et al. 2011
NPHP13	<i>WDR19</i>	57728	ATD5; CED4; DYF-2; ORF26; Oseg6; PWDMP; IFT144; NPHP13; FLJ23127; KIAA1638	chr4:39184024-39287430		Bredrup et al. 2011
NPHP14	<i>ZNF423</i>	23090	OAZ; Roaz; Ebfaz; JBTS19; NPHP14; ZFP423; Zfp104	chr16:49490604-49857919		Chaki et al. 2012
NPHP15	<i>CEP164</i>	22897	NPHP15	chr11:117321778-117413266		Chaki et al. 2012
NPHP16	<i>ANKS6</i>	203286	PKDR1; SAMD6; NPHP16; ANKRD14	chr9:98732009-98796512		Hoff et al. 2013
NPHP17	<i>IFT172</i>	26160	SLB; wim; osm-1; NPHP17; SRTD10	chr2:27444373-27489784		Halbritter et al. 2013
NPHP18	<i>CEP83</i>	51134	CCDC41; NPHP18; NY-REN-58	chr12:94277758-94459988		Failler et al. 2014

Table 1-3. List of genes identified with pathogenic mutations causing nephronophthisis. Eighteen genes are identified to the date with *NPHP1* deletion being the main cause of this condition. Listed are loci, gene names, other names, chromosomal localisation (hg19), founder mutations and references.

About 80% of NPHP patients carry an *NPHP1* deletion¹⁹⁶, although this mutation is also found in JBTS and SLSN¹²⁸.

1.2.2.6 Other ciliopathies

There are many syndromic conditions that are now classified as ciliopathies. These include Bardet-Biedl, Alstrom or oro-facio-digital syndromes. There are also conditions affecting only one organ system but with features compatible with a ciliopathy such as nephronophthisis, polycystic kidney disease (PKD) or Leber congenital amaurosis. Proteins encoded by genes mutated in these conditions have also been shown to be involved in cilia structure and/or function. There are also many syndromes and phenotypes that have not yet been proved to be caused by defects in cilia through formal gene identification, although some evidence points towards ciliary dysfunction as an underlying cause. An example is Marden-Walker syndrome (MIM#248700)⁴ which shows multiple overlapping features of ciliopathies including microcephaly, cleft palate, low set ears, dextrocardia, hypoplastic lungs, cystic kidneys, Dandy-Walker malformation or hypotonia.

1.2.2.7 Inheritance pattern in ciliopathies

Ciliopathies are Mendelian conditions, which means that a particular genotype at a single locus is both necessary and sufficient for the phenotype to be expressed. That means that a given genetic state of one gene is enough to give a certain phenotype.

Most ciliopathies are inherited as autosomal recessive traits (**Figure 1-8**). This means that both copies of a gene, located on one of the autosomes, need to be mutated for the ciliopathy phenotype to be expressed. Heterozygous mutation carriers do not express the phenotype. However, the offspring of two heterozygous carriers have a 25% chance of inheriting both mutant alleles and being affected. Frequently, autosomal recessive traits are seen in the children of consanguineous families or in isolated populations due to a founder mutation. In these situations, a method called autozygosity mapping can be used to locate the mutated gene by identifying the regions of homozygosity in the affected family members that surround the causative mutation¹⁹⁷.

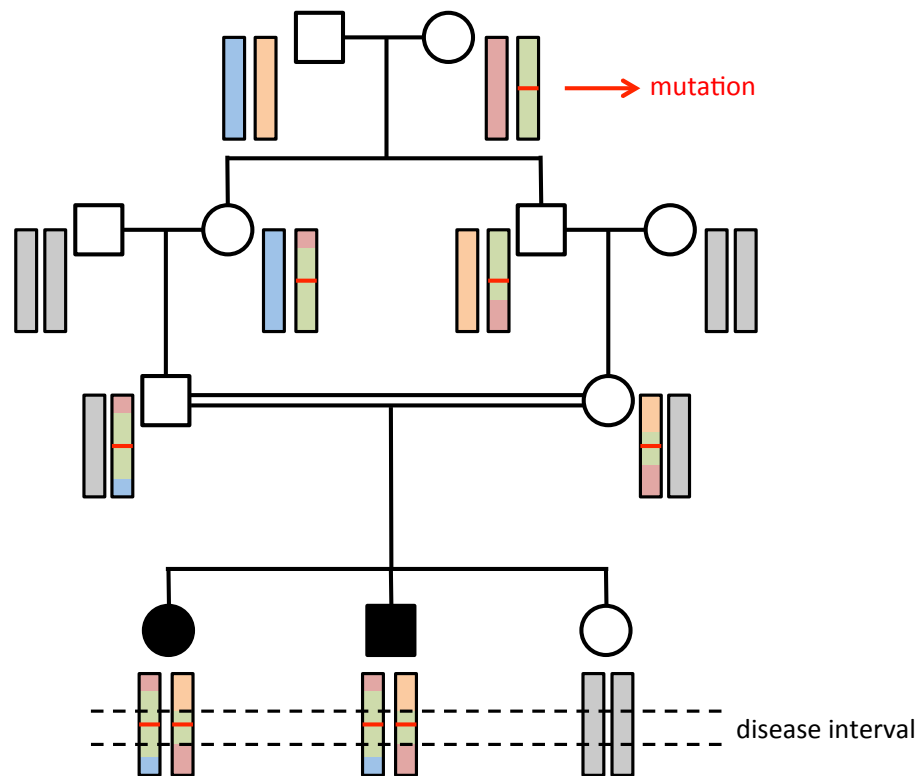


Figure 1-8. Mendelian inheritance of a recessive mutation localised on an autosomal chromosome in a consanguineous family. Mutation that occurred in the ancestor is being passed to offspring in heterozygous state. Heterozygous carriers' offspring has 25% chances to be healthy, 50% chances to be a carrier of the mutation and 25% chances to inherit mutation in homozygous state determining its expression. Recombination happening between generations decrease disease interval. Image adapted from autozygosity.org.

The exception to autosomal inheritance in ciliopathies is *OFD1*, which is an X-linked condition¹³⁸. Mutations in *OFD1* gene cause Oral-facial-digital syndrome 1 (MIM#311200), Joubert syndrome 10 (MIM#300804), Retinitis pigmentosa 23 (MIM#300424) and Simpson-Golabi-Behmel syndrome type 2 (MIM#300209). The phenotype, in families carrying mutations in this gene, is only expressed in male subjects.

1.2.2.7.1 Variable expressivity in ciliopathies

Mendelian inheritance of ciliopathies assumes that mutations in one gene are responsible for the phenotype. However, the phenotypic analysis of ciliopathy patients can demonstrate intrafamilial variability in the expressivity of the ciliopathy phenotype. In this situation, siblings carrying the same mutation in the same gene have differences in expressivity or even different clinical features within a broader spectrum of ciliopathy phenotypes. A leading hypothesis to explain this variation has been the existence of modifier alleles, which is now supported by several studies. It was also reported that some ciliopathies may be inherited in a non-Mendelian manner, where three alleles are required for the phenotype to be expressed¹⁹⁸⁻²⁰⁰. In particular, studies in BBS families showed that two alleles in

BBS2 were not sufficient for the phenotype manifestation and, only if inherited with third allele present in *BBS6*, manifested the disease phenotype. Also *BBS4* was shown to participate in phenotype inheritance with *BBS1* and *BBS2*. The triallelic inheritance remains controversial and needs further investigation^{201,202}.

Missense coding variants in the *RPGRIP1L* (JBTS7, MKS5, NPHP8) and *AHI1* (JBTS1) genes were associated with the expressivity of retinal degeneration in the ciliopathy phenotype. Thr229 in the common variant p.A229T in *RPGRIP1L* is associated with retinal degeneration in a range of ciliopathies including JBTS and MKS²⁰³, and Trp830 in the variant p.R830W in *AHI1* is associated with a more than sevenfold increase in relative risk of retinal degeneration within a cohort of individuals with the kidney ciliopathy nephronophthisis²⁰⁴. In addition T231S, A327S, R867H, c.2322+3A>G, K157E, P209L, Y255C, T1103R, I1208S, C552* changes in *TTC21B* (ATD4, JBTS11, NPHP12) were shown to modify the phenotype in a range of ciliopathies (MKS, BBS, NPHP) interacting in *trans* with mutations in other disease-causing genes¹⁸². Davies *et al.* showed that saturated resequencing combined with systematic functional studies may help in understanding the complexity of phenotypic variability.

The molecular mechanisms that cause such phenotypic variability remain largely unknown, but are likely to include the effect of modifier alleles in other ciliary-related genes and stochastic effects in signalling pathways during embryonic development. Exome sequencing, where all coding variants are collected, could allow further in-depth analysis of genotype-phenotype correlation and variable expressivity.

1.3 Aims of the research

The overall aim of this project is to identify and characterise new genes involved in ciliogenesis and human ciliopathies.

The first part of the project focused on the analysis of DNA samples from a cohort of MKS and JSRD patients established by Prof. Colin Johnson (University of Leeds) through national and international collaborations. At the start of this project, some families from this cohort had already been analysed and causative mutations had been identified in about 50% of subjects. However, the remaining cohort did not have a defined genetic cause for their disease. The aim of this study was to use DNA from mutation negative patients to identify new ciliopathy genes using new sequencing technologies and to characterise function of the proteins they encode. In addition, new ciliopathy genes identified during the course of this study by other

research teams would be screened in this cohort to aid in characterising the mutation spectrum and to investigate genotype-phenotype correlations.

The second part of this project focused on identifying proteins involved in ciliogenesis. Previous studies^{205,206} have shown that RNAi knockdown of the known MKS proteins, TMEM67 and MKS1, caused loss of primary cilia in cell lines. Based on these observations, it was hypothesized that by knocking-down every gene in the mouse genome by RNAi, it may be possible to identify novel candidates for human ciliopathies by screening for loss of primary cilia. This strategy is called “reverse genetics” since disruption of a known genetic locus would lead to an observable phenotype rather than the traditional “forward genetics” approach of using phenotypes to identify genes²⁰⁷. This strategy would identify functional candidate genes that could then be screened in patient DNA samples or be used to filter/prioritise WES and WGS data.

In summary, the specific aims of the project are therefore:

- To identify mutations in the known MKS/JSRD genes
- To identify new genes mutated in MKS/JSRD using multiple approaches
- To investigate the function of the proteins encoded by these new MKS/JSRD genes
- To perform a whole genome RNAi knockdown screen to identify novel ciliopathy genes and genes that mediate ciliogenesis

2 Materials and methods

For the purpose of this thesis “room temperature (RT)” was established as 22-24°C. Followed abbreviations were used: hours (h), minutes (min), seconds (sec), acceleration (g), revolutions per minute (rpm), 1000 (k).

2.1 Materials

2.1.1 General reagents

Water	Millipore
Ethanol	Sigma-Aldrich
Methanol	Sigma-Aldrich

2.1.2 Reaction specific reagents

- PCR
 - Thermoprime plus DNA polymerase ABgene
 - Hot-Shot mastermix Clent
 - Phusion high-fidelity polymerase NEB
 - Primers [25nmol] Sigma-Aldrich
 - dNTP [100mM] Thermo Scientific
 - TRIS base Sigma-Aldrich
 - Amonium sulphate Sigma-Aldrich
 - Magnesium chloride Sigma-Aldrich
 - Tween®20 Sigma-Aldrich
- Gel electrophoresis
 - Agarose Fisher Scientific
 - Ethidium bromide [10mg/ml] Sigma-Aldrich
 - EDTA Sigma-Aldrich
 - Trizma®base Sigma-Aldrich
 - Boric acid Sigma-Aldrich
 - Glycerol Sigma-Aldrich
 - Orange G Sigma-Aldrich
 - Xylene cyanol Sigma-Aldrich
 - Easy ladder Bioline
 - 1kb ladder Promega
- Genotyping

- GeneScan™ 500 ROX™ dye Size Standard Life Technologies
- Hi-Di Formamide Life Technologies
- 5'FAM labelled primers Life Technologies
- Sequencing
 - Exonuclease I (ExoI) Fermentas
 - Shrimp Alkaline Phosphatase (SAP) Fermentas
 - BigDye Terminator Cycle Sequencing Kit v3.0 Life Technologies
- WES
 - SureSelectXT All Exon V4 Agilent Technologies
- Cloning
 - Luria Broth (LB) Sigma-Aldrich
 - Agar Sigma-Aldrich
 - Ampicilin Sigma-Aldrich
 - “α-Select gold efficiency” competent *E. coli* cells Bioline
 - QIAquick mini/maxi prep kits QIAGEN
- Tissue culture
 - Dulbecco's Modified Essential Medium (DMEM) with L-glutamine Sigma-Aldrich
 - DMEM/F12 Life Technologies
 - Fetal bovine serum Life Technologies
 - Trypsine Life Technologies
 - PBS Sigma-Aldrich
 - Lipofectamine2000 Life Technologies
 - LipofectamineRNAiMax Life Technologies
 - OptiMEM Life Technologies
 - G418, Geneticin® [50mg/ml] Life Technologies
 - siRNA duplexes [5nmol] Ambion/Dharmacon
 - Thymidine [2mM] Sigma-Aldrich
 - Propidium iodide (PI) [1mg/ml] Sigma-Aldrich
 - RNase A [10mg/ml] Sigma-Aldrich
 - cell lines:

Cell line	Origin	Source	Reference
IMCD3	mouse inner medullary collecting duct	ATCC	Rauchman et al, 1993
HEK293	human embryonic kidney	ATCC	Graham et al, 1977
HDF neo	hTERT immortalised neonatal human dermal fibroblasts from foreskin of healthy control	Genlantis	
JSRD2	hTERT immortalised human dermal fibroblasts from a patient with TMEM237 mutation p.R18*	Children's Hospital of Eastern Ontario Research Institute	Huang et al, 2011

Table 2-1. Cell lines. Listed are cell lines used in the studies described in this thesis.

- Immunofluorescence (IF)
 - Paraformaldehyde Sigma-Aldrich
 - Non-fat skimmed dried milk Marvel
 - Triton X-100 Sigma-Aldrich
 - DAPI [5mg/ml] Sigma-Aldrich
 - AlexaFluor® (488nm, 568nm) goat anti mouse and goat anti rabbit secondary antibodies [2mg/ml] Life Technologies
 - TOTO3 [1mM] Life Technologies
 - Mowiol Sigma-Aldrich
 - Primary antibodies:

ANTIGEN	RAISED IN	SPECIES REACTIVITY	FIXATION-PFA	FIXATION-MtOH	IF DILUTION (1/X)	WB DILUTION (1/X)	PRODUCER
5HT1B	Rb	Hu, Ms	+		100	500	Novus Biologicals
ABC (active β -catenin)	Ms	Hu, Ms		+	100		Millipore
acetylated α -tubulin	Ms	Ms	+	+	2000	N/A	Sigma-Aldrich
ARL13B	Rb	Hu, Ms	+	+	1000		Proteintech
C21orf2	Rb	Hu, Ms		+	100	500	GeneTex
CRFR2	Rb	Hu, Ms	+	+	100	200	Novus Biologicals
cyclin D1	Ms	Hu, Ms		+	100		Santa Cruz Biotechnology
Dvl1	Ms	Hu, Ms	-	+	100	500	Santa Cruz Biotechnology
FLAG	Ms			+	1000	2000	Sigma-Aldrich
GPR173	Rb	Hu, Ms	+	+	100	200	Novus Biologicals
GPR20	Rb	Ms		+	100	500	Thermo Scientific
GT335	Ms	Ms		+	2000		Enzo LifeScience
HA	Ms	Ms	+		500		Cell Signalling Technology
KI67	Ms	Hu		+	200		
Living colours A.v. peptide (GFP)	Rb				200		Clontech
Living colours GFP monoclonal	Ms					1000	Clontech
MAS1	Rb	Hu, Ms	+	+	100	1000	Novus Biologicals
OPRL1	Rb	Hu, Ms		+	50	500	Abcam
P2RY14	Rb	Hu, Ms	+		100	200	Novus Biologicals
Phalloidin-488			+		200		Life Technologies
Phospho-myosin-2b	Rb	Hu, Ms		+	50	N/A	Cell Signalling Technology
Phospho- β -catenin	Rb	Hu, Ms		+	200	500	Cell Signalling Technology
PIBF1	Rb	Hu		+	100	500	Novus Biologicals
PLK4	Rb	Hu, Ms		+	100	500	Cell Signalling Technology
PRPF31	Gt	Hu		+	100	500	Abnova
PRPF6	Rb	Hu, Ms		+	100	500	Santa Cruz Biotechnology
PRPF8	Ms	Hu, Ms		+	100	200	Santa Cruz Biotechnology
PRPF8	Rb	Hu, Ms		+	100	200	Santa Cruz Biotechnology
Rapsyn	Ms	Ms		+	100	500	Thermo Scientific
RhoA	Ms	Ms		+	100	500	Cytoskeleton
TMEM237	Rb	Ms		+	100	500	kindly provided by Dr.C.Craft
TMEM67	Rb	Ms	+		25	500	GeneScript
Xpress	Ms					1000	Life Technologies
β -actin	Ms	Hu, Ms			N/A	10000	Abcam
β -catenin	Rb	Hu, Ms		+	100	1000	Cell Signalling Technology
γ -tubulin	Rb	Hu, Ms	-	+	1000	N/A	Abcam
γ -tubulin	Ms	Hu, Ms	-	+	200	N/A	Sigma-Aldrich
γ -tubulin C3		Ms	-/+	+	200		Sigma-Aldrich

Table 2-2. List of primary antibodies. Table contains used antibodies indicating optimized fixation method, species they were raised in, species reactivity, dilution for IF and WB and producer. Ms-mouse, Rb-rabbit, Hu-human.

- Western blotting (WB)
 - Protease/Phosphatase inhibitors Thermo Scientific
 - PMSF Sigma-Aldrich
 - NP40 (IGEPAL) Sigma-Aldrich
 - RC DC™ Protein Assay Bio-Rad
 - NuPage 4-12% MES SDS gels Life Technologies
 - NuPage MES running buffer Life Technologies

- NuPage transfer buffer Life Technologies
- Invitrolon™ PVDF filter paper sandwich Life Technologies
- SeeBlue2 prestained standard Life Technologies
- SuperSignal West Femto Substrate Thermo Scientific
- Sodium hydroxide Sigma-Aldrich
- Goat anti-rabbit HRP Dako
- Goat anti-mouse HRP Dako
- Immunoprecipitation
 - Protein A agarose Roche

2.1.3 Buffers

- Phosphate buffer saline (PBS) (1x)
- Tris – EDTA (TE)
 - 10mM Tris-HCl pH7.5
 - 1mM EDTA
- Tris/Borate/EDTA (TBE) (10x) pH8.0
 - 890mM Tris
 - 890mM Boric acid
 - 20mM EDTA
- PCR buffer (10x)
 - 750mM Tris-HCl pH8.8
 - 200mM (NH₄)₂SO₄
 - 0.1% Tween 20
 - 15mM MgCl₂
- Gel loading buffer (2x)
 - 50% Glycerol
 - 10% 10x TBE
 - 0.1% Orange G
 - 0.1% Xylene cyanol
- NP40 cell lysis buffer
 - 1% NP40
 - 50mM Tris–HCl pH8.0
 - 150mM NaCl
 - 1x Protease/Phosphatase Inhibitors
- IP incubation buffer (transmembrane proteins)
 - 10% glycerol

- 10% ethanol
- 20mM Tris – HCl pH8.0
- 25mM NaCl
- 2mM EDTA
- 0.5mM PMSF
- protease/phosphatase inhibitor (1x)
- IP wash buffer
 - 150mM NaCl
 - 50mM Tris pH8.0
 - 0.5mM EDTA
 - 0.1% NP40
- SDS loading buffer (2x)
 - 4% SDS
 - 20% glycerol
 - 20mM β – mercaptoethanol
 - 100mM Tris-HCl pH6.8
 - 0.004% bromophenol blue

2.2 Methods

2.2.1 Patient DNA

Families across the UK with offspring diagnosed with Meckel-Gruber or Meckel-Gruber-like syndrome were recruited to the study. Blood samples with informed consent were collected under ethical approval from South Yorkshire Local Research Ethics Committee (study title “Molecular genetic investigations of autosomal recessive conditions”, REC reference 11/H1310/1, **Appendix 1**). DNA was extracted using a standard phenol/chloroform protocol by the Yorkshire Regional Genetics Service (<http://www.leedsth.nhs.uk/sites/genetics/>). Other DNA samples were obtained from collaborators abroad.

2.2.2 Patient fibroblasts

4-6mm skin punch biopsies were obtained from affected individuals under local anaesthesia with 1% lidocaine (“EMLA” cream). Fibroblasts were dispersed and grown in DMEM tissue culture medium (Life Technologies) containing 10% foetal bovine serum and 0.4% penicillin/streptomycin. Once confluent cultures were established, primary cells were grown and cryopreserved by standard procedures.

Fibroblasts were immortalised using retrovirus-mediated transduction to express human telomerase reverse transcriptase (hTERT)²⁰⁸. Fibroblasts were immortalised by Dr Julie Burns, CR-UK Cancer Studies, University of Leeds.

2.2.3 DNA extraction

2.2.3.1 Saliva

Saliva samples were collected and extracted using Oragene® DNA sample collection kit (DNA Genotek) following the manufacturer's instructions. The DNA pellet was resuspended in 50µl 1x TE for long term storage. DNA concentration was determined using a NanoDrop 1000 spectrophotometer (Thermo Scientific) according to the manufacturer's protocol.

2.2.3.2 Cells

DNA was extracted from patients' fibroblasts following standard phenol/chlorophorm extraction protocol.

2.2.4 RNA extraction

For RNA extraction, cells were scraped off culture vessels and spun down for 5min at 200x g at RT. The cell pellet was then resuspended in 1x PBS and spun down again. The final cell pellet was resuspended in Trizol® reagent (Life Technologies) following the manufacturer's instructions.

Tissue samples were snap-frozen in liquid nitrogen, cut into fine pieces and ground into a fine powder using a chilled pestle and mortar. The powdered tissue sample was resuspended in Trizol® reagent (Life Technologies) following the manufacturer's instructions. All equipment and benches were cleaned with RNaseZap (Ambion) to remove RNase traces. All solutions were prepared with RNase-free water.

2.2.5 Polymerase Chain Reaction (PCR)

PCR primers were designed using Primer 3.0 online software (<http://bioinfo.ut.ee/primer3-0.4.0/>) covering all coding exons and flanking intronic regions (**Table 2-3** lists primers for MKS genes, and primers for all other investigated genes are listed in the **Appendix 2**).

PCR reactions were performed in a final volume of 10 μ l. Each PCR reaction contained 20ng genomic DNA, 2 μ M of 5' and 3' primers (**Table 2-3** and **Appendix 2**), 0.1 unit of Taq DNA polymerase (ABgene), 0.25 μ M of dATP, dGTP, dCTP and dTTP each (Thermo Scientific), in PCR buffer (75mM Tris-HCl buffer pH8.8, 20mM (NH₄)₂SO₄, 0.01% [w/v] Tween 20 and 1.5mM MgCl₂). Reactions were placed in a thermocycler (Bio-Rad) with the following reaction conditions: 96°C for 3min, followed by 30 to 45 cycles of 96°C for 30sec, 50-65°C for 30sec (**Table 2-3**), 72°C for 1min and then final extension step of 72°C for 5min. Amplified DNA products were subsequently analysed by agarose gel electrophoresis. Samples were mixed in a 1:1 [v/v] ratio with 2x loading buffer and run on a 1.2% [w/v] agarose gel containing 0.5 μ g/ml ethidium bromide with a DNA size standard (Easy Ladder I, Bionline). PCR products were visualised on a GelDoc station (Bio-Rad).

2.2.6 Genotyping

DNA was genotyped using primers with the 6-FAM fluorescent reporter dye attached by the hydroxy group at the 5' end of the forward primer. A standard PCR reaction was set up, followed by gel electrophoresis and the sample diluted (usually 1:25) before genotyping. 1µl of diluted PCR product was mixed with 8.5µl HiDi formamide and 0.5µl fluorescently-labelled size standard (ROX™ 500, Life Technologies) and run on an ABI3130xl genetic analyser using POP7 polymer and the default run module FragmentAnalysis36_POP7_1.

DNA from individuals was genotyped for microsatellite markers with heterozygosity values of >0.70 that flanked seven of the nine known MKS genes (at the time of the study: *MKS1*, *TMEM216*, *TMEM67*, *RPGRIP1L*, *CC2D2A*, *CEP290* and *TMEM237*) at <1cM genetic distance, or otherwise as close as possible (**Table 2-4**).

GENE	MARKER	FORWARD PRIMER (WITH 'FAM' DYE)	REVERSE PRIMER	ANNEALING TEMPERATURE
<i>MKS1</i>	D17S1606	TGGTATTCAATCCTGGAGC	TGATGAGTCTTCATAGCCCC	58.9
	D17S1290	GCCAACAGAGCAAGACTGTC	GGAAACAGTTAAATGGCCAA	58.9
<i>TMEM216/TMEM138</i>	D11S4191	GCAAGATGGCCAATTAGAAG	TTTTGGTTGGAATGTAGTTGTTTAT	58.9
	D11S4076	CATGAATGCTCTTGTC	AACCCCTGGAAAATAGACT	58.9
<i>TMEM67</i>	D8S1988	CCTTTGGACTCAGACCAGAA	TAGTCAGAGTCTCAGAGAAACA	58.9
	D8S1699	CAACCTGACCCTGCCA	CATGATGTTCTAAGCATATCTGC	58.9
<i>CEP290</i>	D12S1719	TCCTCCAGTTTCAGTAATGTTT	GGTGGTTGATGCCTGTAA	58.9
	D12S1710	AGGTTTCTGGGTTCCGATA	CCATAATCCGTAGGAGCAA	58.9
<i>RPGRIP1L</i>	D16S3034	TAATCTAGTTAAAGATGCAACTGCC	GCTCAGAAGTTTTGATGCC	58.9
	D16S771	GTCCAAAACACCACCTCTA	AAGTAGATCAGTCATCTTGCTGC	58.9
<i>CC2D2A</i>	D4S1511	AGCCTCTGTAATCTTGTTG	TCCATTACTCAGGGCTCTC	58.9
	D4S2960	AAGGCTTTATCATTAAAGAATCCTA	TGAGGGTATAGTTACCATCTTTT	58.9
<i>NPHP3</i>	D3S1596	ATCAATGCCCTGCTCATTAC	CCTGCATCATGTGCTCTC	58.9
	D3S1290	TTGCAGTAATGACCATAGGG	AACACTTAGGGTAATGGGGC	58.9
<i>TMEM237</i>	D2S2309	TGTCAGGCACTTCGCTA	TGCTTCTATTGTACCCAAA	58.9
	D2S1384	AATAGAGGGCCCTTGCTTAA	TTGGGATAAAAGGTATTTTGC	58.9

Table 2-4. List of microsatellite primers flanking known MKS loci. Listed are known MKS genes with flanking markers, their sequences and optimised annealing temperatures.

In samples from singleton or multiplex affected individuals of consanguineous origin, two homozygous markers indicated putative linkage to a locus under investigation, prioritizing the gene for subsequent sequencing. In non-consanguineous multiplex families, two or more affected individuals sharing haplotypes for flanking markers indicated putative linkage to a locus. Data were analysed using GeneMapper® (Life Technologies) and haplotypes were drawn in Excel.

2.2.6.1 Sample genotyping using SNPchip

1µg of DNA from consanguineous affected individuals were sent to AROS (Applied Biotechnology) for genotyping with the Affymetrix Genome-Wide Human SNP Array 6.0. Firstly, data were analysed for CNVs using Affymetrix Genotyping Console 3.0 software and subsequently genotypes were analysed using SnpViewer (Sourceforge, <http://sourceforge.net/projects/snpviewer/>). Autozygous regions were identified and further analysed using Excel.

2.2.7 Purification with ExoSAP

In order to be sequenced, PCR-amplified DNA products were treated with ExoSAP, a mixture of two enzymes: 20U Exonuclease I and 1U shrimp alkaline phosphatase (SAP; Fermentas). To each 2.5µl of PCR reaction, 1µl of the ExoSAP enzyme mix was added. Samples were subsequently incubated at 37°C for 15min followed by denaturation at 85°C for 15min.

2.2.8 Sequencing

2.2.8.1 Sanger sequencing

Enzymatically purified DNA samples were sequenced using the BigDye Terminator Kit v.3.1 (Life Technologies) following the manufacturer's instructions. 1µl DNA was added to 0.5µl BigDye, 2µl 5x ABI Sequencing Buffer, 5.5µl dH₂O and 1µl of the appropriate sequencing primer (**Table 2-1**) at a final concentration of 0.2µM. Reactions were incubated for 28 cycles at 96°C for 15sec, 50°C for 15sec, and 60°C for 4min.

Sequenced samples were subsequently precipitated using 5µl 0.125M EDTA pH8.0 and 60µl 100% [v/v] ethanol. The mixture was centrifuged at 1400x g for 30min at 4°C. Samples were then inverted and tapped onto paper towel to remove excess supernatant. Samples in 96-well plates were spun inverted on a paper towel for 15sec at 100 x g to remove the remaining liquid and 60µl of freshly prepared 70% [v/v] ethanol was added to each well. Samples were spun at 720x g for 15min at 4°C. Afterwards, plates were inverted on a paper towel, and then again centrifuged. Samples were air dried at RT for about 10min. Precipitated DNA was then resuspended in Hi-Di Formamide (Life Technologies), and sequencing products were separated by electrophoresis through polymer POP-7™ (Life Technologies), according to the manufacturer's instructions. A total of 188 coding

exons were amplified by standard PCR protocols with one additional *CEP290* intronic amplicon covering the common LCA change ²⁰⁹. The electropherogram trace was analysed using Sequencing Analysis, SeqScape (Life Technologies) or 4Peaks (Mek&Tosj.com).

2.2.8.2 WES

Initial WES sample preparation was done using SureSelectXT Human All Exon V4 kit (Agilent Technologies) by Dr Clare Logan following manufacturer's instructions. Samples were sequenced on an Illumina GAIIX (sample 157), Illumina MiSeq (samples 36A and 66F1) and Illumina HiSeq2500 instrument (samples 17, 144, 325 and 352). Fastq sequence files were aligned to GRCh37 (Genome Reference Consortium human reference assembly built 37) using the Novoalign aligner (<http://novocraft.com/>). SAM files were converted to BAM files using the Picard tool (<http://broadinstitute.github.io/picard/>), and using the same tool PCR duplicates were marked and removed. The Genome Analysis Toolkit (GATK; <http://www.broadinstitute.org/gatk/>) suite of tools were subsequently used to realign insertions/deletion and bases and identify variants. All variants were filtered out with poor quality scores (Phred-scaled $p\text{-value} > 60$, $\text{RMSMappingQuality} < 40$), and if they were present in dbSNP129 (<http://www.ncbi.nlm.nih.gov/projects/SNP/>) or higher, the 1000 Genome Project (<http://www.1000genomes.org/>) ^{210,211} and in the EVS server (<http://evs.gs.washington.edu/EVS/>) with $\text{MAF} \geq 0.01$. Remaining variants were further filtered using vcfhacks suite of tools (<http://sourceforge.net/projects/vcfhacks/>) to identify those that were biallelic since all patients sequenced were consanguineous, although biallelic heterozygous variants were also checked in case of compound heterozygosity. The pathogenic potential of putative missense mutations was assessed by analysis with PolyPhen2 (<http://genetics.bwh.harvard.edu/pph2/>) ²¹², SIFT (<http://sift.jcvi.org/>) ²¹³ or by manual comparison of CLUSTALX ²¹⁴ alignments of protein homologues to determine the phylogenetic conservation of mutated amino acid residues. The absence of novel mutant alleles was confirmed in a panel of 96 DNA samples from ethnically-matched normal control individuals.

2.2.9 Quantitative real-time PCR (qRT-PCR)

Primers were designed for amplification of human *TMEM237* and a reference gene - *36B4* (member of 40S ribosomal subunit) cDNA using Primer Express 3.0

software. Primers sequences were as follows: for *TMEM237* forward 5'GCAATGAGCCATCAACTAAAGAACT and reverse 5'GAGGAAGTCTCCAATTC AAGAGGTA; for *36B4* forward 5'AGATGCTGCAGATCCGCAT and reverse 5' ATATGAGGAGCAGTTTCTCCAG. RNA extracted from fibroblasts (section 2.2.4) was converted to cDNA using SuperScript III Reverse Transcriptase (Life Technologies) according to manufacturer's instructions. qRT-PCR reactions were set up using SYBR® GreenER™ qPCR SuperMix Universal Kit (Life Technologies) as follows: for total volume of 25µl, 12.5µl of 2x PCR mix was added to 2µl cDNA, 0.25µl 20µM primers, 0.05µl ROX dye and 10.2µl of DEPC H₂O. Samples were amplified on an ABI 7500 Real Time PCR System (Life Technologies). Initially samples were incubated at 50°C for 2min, followed by denaturation at 95°C for 10min and 40 cycles including denaturation at 95°C for 15sec and annealing/extension 60°C for 1min. Data were analysed with 7500 System software (Life Technologies).

2.2.10 Cells

Mouse inner medullary collecting duct (IMCD3) cells were grown in Dulbecco's minimum essential medium (DMEM)/Ham's F12 supplemented with 10% foetal bovine serum at 37°C/5% CO₂. Immortalised fibroblasts were grown in the same medium supplemented with 0.2mg/ml geneticin.

2.2.11 Transfection

2.2.11.1 Over-expression

For transfection with plasmids in 6-well plate, cells at 80% confluency were transfected using Lipofectamine2000 (Life Technologies). 6µl of Lipofectamine2000 was mixed with 250µl OptiMEM and after 5min incubation 1µg of plasmid DNA was added. Growing media was replaced with OptiMEM and transfection complexes were added and incubated for 4h. Then complexes were taken off and replaced with fresh growing media. Cells were assayed 48-72h later.

2.2.11.2 siRNA

For RNAi knock-down experiments, siRNA duplexes were designed against mouse *Tmem237* sequence using Ambion custom Silencer Select siRNA service (catalogue number: 4390771) and *Rpgrip1l* using the Dharmacon ON-TARGET

plus SMART pool (catalogue number: L-055668-01-0005). Antisense sequences were as follows: *Tmem237*: duplex 1 5'-GGAUCUUAGUGAAGAGUUATT, and duplex 2 5'-GAACGAAAACGGCAUUGAUTT; *Rpgrip1l* pool: duplex 1 5'-GGAUCAAGCUAUUCGACUU, duplex 2 5'-CAGCACAGAUUACGAAACA, duplex 3 5'-GAAUACUGGUUCCGAUUAA, duplex 4 5'-CAAUAAAGAUUCUAGACCGA. The medium or low GC non-targeting scrambled siRNA duplexes (Dharmacon) were used as negative controls. 100pmol of each siRNA was transfected into IMCD3 cells at 60-80% confluency in a 6-well plate. 6µl Lipofectamine RNAiMax (Life Technologies) was mixed with 125µl OptiMEM, 100pmol of siRNA was mixed with 125µl OptiMEM and incubated for 5min. Both solutions were then combined together and incubated for another 20min. Cells had their growing media replaced with OptiMEM and knockdown complexes were added. Cells were assayed 72h later.

2.2.12 IF and confocal microscopy

Cells were seeded at 2×10^4 cells/well on sterilised glass coverslips in 6-well plates and fixed with ice-cold methanol (5min at -20°C) or 2% paraformaldehyde (20min at RT, then permeabilised with 0.01% Triton X-100 for 5min at RT). Cells were blocked with 1% [w/v] Marvel dried milk for 5min at RT and then incubated for an 1h with diluted primary antibody in the blocking solution (**Table 2-2**) in a humidity chamber. Cells were then washed three times with 1x PBS and incubated for 1h with diluted secondary antibodies in the humidity chamber. Cells were then washed five times with 1x PBS and once with distilled water and were mounted to slides with Mowiol. Confocal images were obtained using a Nikon Eclipse TE2000-E system, controlled and processed by EZ-C1 3.50 (Nikon) software. Images were assembled using Adobe Illustrator CS2.

2.2.13 Whole cell extract preparation and SDS-PAGE western blotting

Whole cell extracts (WCE) were prepared from confluent cells in 90mm tissue culture dishes, or scaled as appropriate. All procedures were conducted at 4°C to prevent protein degradation by proteases. Cells were washed with cold 1x PBS once and lysed with 100µl ice-cold NP40 lysis buffer for 5min at 4°C . Cells were scraped off the surfaces using chilled plastic cell scrapers, placed in cold Eppendorf tubes. Cells were then spun down at $12k \times g$ for 5min and the supernatant was

transferred to new tubes. Protein concentration was measured using RC DC™ Protein Assay kit (Bio-Rad) following the manufacturer's instructions. Absorbance was measured on a spectrophotometer at 750nm. Proteins were mixed with 2x SDS loading buffer, boiled and run on 4-12% SDS-PAGE (sodium dodecyl sulfate polyacrylamide gel electrophoresis) acrylamide gradient gels (Life Technologies) alongside protein marker (SeeBlue2, Life Technologies) in MES running buffer (Life Technologies) for 1.5h at 120V. Proteins were subsequently immunoblotted onto polyvinylidene difluoride (PVDF) membrane (Life Technologies) in transfer buffer supplemented with 10% methanol for 1h at 30V. Membranes were blocked in 5% [w/v] Marvel dried milk in 1x PBS with Tween 20 (PBST) for 1h and incubated with primary antibodies (**Table 2-2**) in Falcon tubes on a roller for 1h. Membrane was then washed four times with 1x PBST and incubated for 1h with the appropriate HRP-conjugated secondary antibody (Dako) at a final dilution of x10000. The membrane was then washed four times with 1x PBST. For detection by the enhanced chemiluminescence, the Femto West immunoblot detection system (Thermo Scientific) was used. Visualised bands were processed in ImageLab software (Bio-Rad).

2.2.14 TopFlash assay

For luciferase assays of canonical Wnt activity, fibroblasts were grown in 48-well plates and co-transfected with 0.25µg Topflash firefly luciferase construct (or Fopflash, as a negative control); 0.25µg of expression constructs (pCMV-HA-TMEM67, pCMV-GFP-TMEM216, pEGFPN1-TMEM237, or empty pCMV-HA/pEGFPN1- vector; and 0.05µg of pRL-TK (Promega; Renilla luciferase construct used as an internal control reporter). Cells were treated with Wnt3A or Wnt5A conditioned media to stimulate or inhibit the canonical Wnt pathway. Wnt3A- or Wnt5A-conditioned media were obtained from stably-transfected L cells with Wnt3A or Wnt5A expression vectors. Control media was obtained from untransfected L cells. Activities from Firefly and Renilla luciferases were assayed with the Dual-Luciferase Reporter Assay system (Promega) on a Mithras LB940 (Berthold Technologies) luminometer. Minimal responses were noted with co-expression of the FopFlash negative control reporter construct. Raw readings were normalized with Renilla luciferase values.

2.2.15 Immunoprecipitation (IP)

WCE were prepared as in **2.2.12** with additional DNA shearing by several passes through a 19G needle. 1mg of protein was transferred to a fresh pre-chilled Eppendorf tube and the volume was adjusted to 1ml with IP incubation buffer. Protein A agarose beads were washed with 1x PBS and centrifuged at 1000x g for 1min at 4°C. The wash and spin were repeated for one more time and then restored to a 50% [v/v] suspension with 1x PBS. The cell lysate was pre-cleared with 70µl 50% [v/v] bead suspension by incubation on a rotator for 30min at 4°C, then spun down for 1min 1000x g at 4°C. The cell lysate was then transferred to a fresh tube. 1 to 8µg of the appropriate antibody was added to 150 to 500µg of soluble protein in the cell lysate and incubated overnight on a rotator at 4°C. 80µl of protein A or G suspension was added and incubated on a rotator at 4°C for 3h. Beads were then spun down at 1000x g for 1min at 4°C and the supernatant was discarded. Beads were washed three times with 500µl ice-cold IP wash buffer and spun down at 1000x g for 1min at 4°C in between the washes. After washing, 20µl 2% SDS was added to the bead pellet, mixed by inverting to elute the proteins and left to incubate for 30min at RT with gentle mixing every few minutes. The final spin was done at 14k x g for 2min at RT. The supernatant was transferred to a new tube and an equal volume of 2x SDS loading buffer was added. Samples were further analysed as described in **2.2.13**.

2.2.16 RhoA GST pull-down

The activated GTP-bound isoform of RhoA was specifically assayed in pull-down assays using a GST fusion protein of the Rho effector rhotekin (Cytoskeleton), using conditions recommended by the manufacturer. WCEs were processed as rapidly as possible at 4°C, and snap-frozen in liquid nitrogen. Total RhoA in WCEs and pull-down protein was detected on immunoblots using a proprietary anti-RhoA monoclonal antibody (Cytoskeleton). Immunoblot analysis for total RhoA and β -actin were used as loading controls.

2.2.17 Flow cytometry

Four T25 flasks of fibroblasts (two control and two patient fibroblasts) were seeded at the same cell density. At about 30% density 2mM thymidine was added to synchronise cell cycle and incubated for 18h. After reaching 50% confluency, cells from two flasks (one control and one patient fibroblasts) were harvested by

trypsinization. and the cells sedimented by centrifugation at 200x g for 5min. Media was then taken off and the pellet was resuspended in 1x PBS. After re-centrifugation at 200x g for 5min the PBS wash was discarded leaving 100µl to cover the cell pellet. 500µl of 70% [v/v] ethanol was then added dropwise to fix the cells and mixed. The same volume of ethanol was added twice more in the same manner. Samples were stored at -20°C. The remaining two flasks of cells were incubated for another 48h and then cells were collected as described above. Before samples were run into the flow cytometer, they were washed twice with 1x PBS, and then incubated for at least 30min in propidium iodide (PI) solution (20µg/ml; Sigma-Aldrich) containing RNAase A (0.1µg/ml; Sigma-Aldrich) to remove the RNA. Cells were then analysed on a LSRII FACS machine (Beckman Coulter) using ModFit software (Verity Software). Data from 5000 events were collected and analysed to identify cells in G1-S-G2/M phases of the cell cycle.

2.2.18 The whole genome siRNA screen

Cell culture and transfections were done by BSTG (BioScreening Technology Group) members (Mrs Julie Higgins and Dr Matthew Adams). IF staining, image acquisition and analysis were done by Dr Gabrielle Wheway and the author.

2.2.18.1 Cell culture

Mouse inner medullary collecting duct (mIMCD3) cells from American Type Culture Collection (ATCC) were maintained in DMEM/F12 medium supplemented with 10% foetal bovine serum (FBS), under standard conditions (37°C, 5% CO₂). Cells were passaged at a split ratio of 1:10 twice a week. Cells were obtained from ATCC at passage 13 and were used for screening purposes between passage 17 and 25.

2.2.18.2 siRNA

A Dharmacon mouse siGENOME siRNA library (Dharmacon) was used, plated in a standard 96-plate well format. Each gene was targeted by a pool of four duplexes per well, across 36 sub-library plates, 89 drug target sub-library plates, and 114 plates targeting the remaining genes in the mouse genome (the genome sub-library). The entire library targeted a total of 18960 Entrez RefSeq genes across 239 plates.

Library plates contained 80 targets per plate across columns 2-11. Lyophilised siRNA was re-constituted with 1x RNA buffer (Thermo Scientific) to a concentration

of 2 μ M using a Fluid-X XRD-384 dispenser, followed by 90min of agitation on a rotary shaker. Reconstituted siRNA plates were stored at -80°C.

Thawed siRNAs were directly aliquoted into 96-well assay plates (View Plates, PerkinElmer) using an Agilent Bravo liquid handling platform. Eight different control siRNAs (siGENOME, Dharmacon) at 2 μ M were added to plates using a Star-pet-E 8-channel electronic pipette (in duplicate to columns 1 and 12). An siRNA pool targeting mouse si*Plk1* was used as a positive transfection control. Mouse si*Mks1*, si*lft88* and si*Rpgrip1l* were used as positive controls for effects on ciliogenesis. A number of negative controls were used: a non-targeting siRNA against human si*MLNR* which has no target in the mouse genome, two duplicate Dharmacon scrambled non-targeting control siRNAs and a transfection reagent-only control. Sequences of control siRNA duplexes are given in **Table 2-5**.

Target gene	Gene ID	Gene accession	siRNA sequence
<i>Plk1</i>	18817	NM_011121	CCAACCAAAGUGGAAUAUGA
			GCAAUUACAUGAGUGAGCA
			GCAAGAUUCGUGCCUAAGUC
			UCACUCUCCUCAACUAUUU
<i>Mks1</i>	380718	NM_001039684	CGGCGAAUCUUCACUACA
			AGUUUGAAGUCGACCUGUA
			CAAUGUACAUCAUGGCGGA
			UGGCUGAGCGGAUGGCGAA
<i>Rpgrip1l</i>	244585	NM_173431	GGAUCAAGCUAUUCGACUU
			CAGCACAGAUUACGAAACA
			GAAUACUGGUUCCGAUUAA
			CAUAAAAGAUCUAGACCGA
<i>lft88</i>	21821	NM_009376	CGGAGAAUGUUGAAUGUUU
			GCUUGGAGCUUAUUACAUU
			CGUCAGCUCUCACUAAUAA
			GUAGCUAGCUGCUUUAGAAA
<i>MLNR</i>	2862	NM_001507	GAAGAUUCGCGGAUGAUGU
			CAUCGUCGCUCUGCAACUU
			GCGCAUCUAUCAACCCAUAU
			GCGCUAACGUGAAGACGAU

Table 2-5. Sequences of siRNA duplexes used to target positive and negative control genes for ciliogenesis.

2.2.18.3 Transfection

Reverse transfections were set up in batches of 20 plates. 2.5 μ l 2 μ M siRNA were added per well to optical-bottomed 96-well View Plates (PerkinElmer), followed by 0.2 μ l Lipofectamine RNAiMAX transfection reagent (Life Technologies) suspended in 17.5 μ l OptiMEM serum free media (Life Technologies) using a Fluid X XRD-384 dispenser at high speed (300rpm). Plates were gently mixed on a rotary

shaker at RT for a minimum of 20min. 80µl mIMCD3 cells at a density of 10^5 cells per ml in OptiMEM, in a uniform suspension maintained by constant magnetic stirring, were added per well to plates using an XRD-384 FluidX dispenser at high speed.

To minimize edge effects, cells and transfection complexes were left to rest in a laminar flow hood for 1h at RT before returning to a 37°C 5% CO₂ incubator without further changes of media in wells ²¹⁵. Cells were assayed after 72h.

2.2.18.4 Antibodies and staining reagents

Cilia were detected with a mouse monoclonal antibody against acetylated α -tubulin (Sigma-Aldrich) and an Alexa Fluor 488-conjugated secondary antibody (Life Technologies). DAPI (Sigma-Aldrich) was used with the far-red RNA stain TOTO3 (Life Technologies) to determine cell, cytoplasm and nuclear boundaries.

2.2.18.5 High-throughput liquid handling

Plates were processed for immunofluorescent staining in batches of 20. 72h after transfection media was removed from plates by inverting the plate and blotting on clean paper towels to remove excess liquid. Cells were washed three times with 100µl 1x PBS, using two XRD-384 FluidX dispensers working in parallel on slow speed (100rpm), dispensing to the left side of the well to minimise disturbance of cells. After fixation, all cell washes were performed at medium speed (150rpm). After incubation with primary antibody, cells were washed once using 100µl PBS on a XRD-384 FluidX dispenser and three times on a Biotek ELx405 Select CW plate-washer with a Biostack automated plate stacker that dispensed against the left side of the well. After incubation with secondary antibody and stains, cells were washed once using 100µl PBS on a XRD-384 FluidX dispenser and six times on a Biotek ELx405 Select CW plate-washer with a Biostack automated plate stacker. Plates were stored with 100µl PBS per well at 4°C for up to 1 week before imaging.

2.2.18.6 Fixation

50µl ice cold methanol was dispensed to the left side of each well using FluidX XRD-384 dispenser on slow speed (100rpm). Plates were then placed in a -20°C freezer for exactly 5min. After fixation, each plate was removed from the freezer, inverted and blotted to remove methanol and washed with 50µl PBS.

2.2.18.7 IF staining

Plates were blocked with 100µl 1% Marvel dried milk/PBS [w/v] previously clarified of particulates by centrifugation at 3000x g for 5min. All antibodies and

stains were diluted in blocking solution and clarified by centrifugation at 12k x g for 5min. at RT. Primary antibody (anti-acetylated α -tubulin) was diluted 1000x and cells were incubated with 50 μ l antibody solution per well for 1h at RT. Alexa 488 secondary antibody (2000x dilution), and DAPI and TOTO3 (both 5000x dilutions) were incubated for 1h at RT in the dark.

2.2.18.8 High-throughput imaging

Plates were processed using the PerkinElmer Operetta high-content wide-field fluorescence imaging system, coupled to Harmony software. Plates were automatically loaded onto the Operetta using a PerkinElmer Plate Handler II robotic arm, operated through Plateworks software. Plates were bar-coded with specific plate information, and barcodes were read by a SICK barcode reader before loading onto the Operetta. Wells were imaged using a 20x objective lens, detecting three colours in three separate focal planes to ensure the best resolution was obtained for each colour. The bottom of each well was detected automatically by the Operetta infra-red focusing laser, and focal planes of detection for each colour were calculated relative to this value. DAPI emission was detected for 60ms 6 μ m below the calculated bottom of the well, TOTO3 emission was detected for 500ms 1 μ m above the DAPI plane, and Alexa Fluor 488 emission was detected for 700ms 7 μ m above DAPI. Six fields of view (each 510x675 μ m) were imaged per well (with an identical pattern of fields in every well away from the dispense area), with an approximate total of 4000 cells detected and analysed per well.

2.2.18.9 Image analysis

Modified PerkinElmer image analysis algorithms were used throughout (**Appendix 6**). Nuclei were detected as blue (DAPI) fluorescent regions $>30\mu\text{m}^2$, with contrast >0.10 . Cytoplasm was detected as far-red (TOTO3) fluorescent regions around nuclei. Border objects were excluded so that only whole cells were analysed. Cilia on whole cells were detected using a modified 'find spots' algorithm, identifying green (Alexa Fluor 488) fluorescent spots with radius <3.8 pixels, contrast >0.1 , spot-to-region intensity >1.3 with a distance from all neighbouring spots >5.6 pixels. Any cells with more than one cilium in the cell area were excluded from analysis. Key output parameters were number of whole cells, and percentage of whole cells with a single cilium. Additional output parameters included number of whole cells with two or more cilia and mean cilium intensity, which could be used as an indicator of cilium length.

3 Results

The work that comprises the results chapter was divided into three subchapters. The first subchapter includes efforts to identify linkage of samples to known MKS/JBTS loci using microsatellite markers and subsequent mutation screening using Sanger sequencing (**Figure 3-1**, green background). No linkage to these loci would prioritise the sample for the screening of new functional candidate genes, SNIpchip analysis and/or whole exome sequencing. However if the sample showed linkage to the particular locus it was Sanger sequenced for the mutation identification. Lack of variant identification would subsequently prioritise the sample for the same analyses described for no linkage samples.

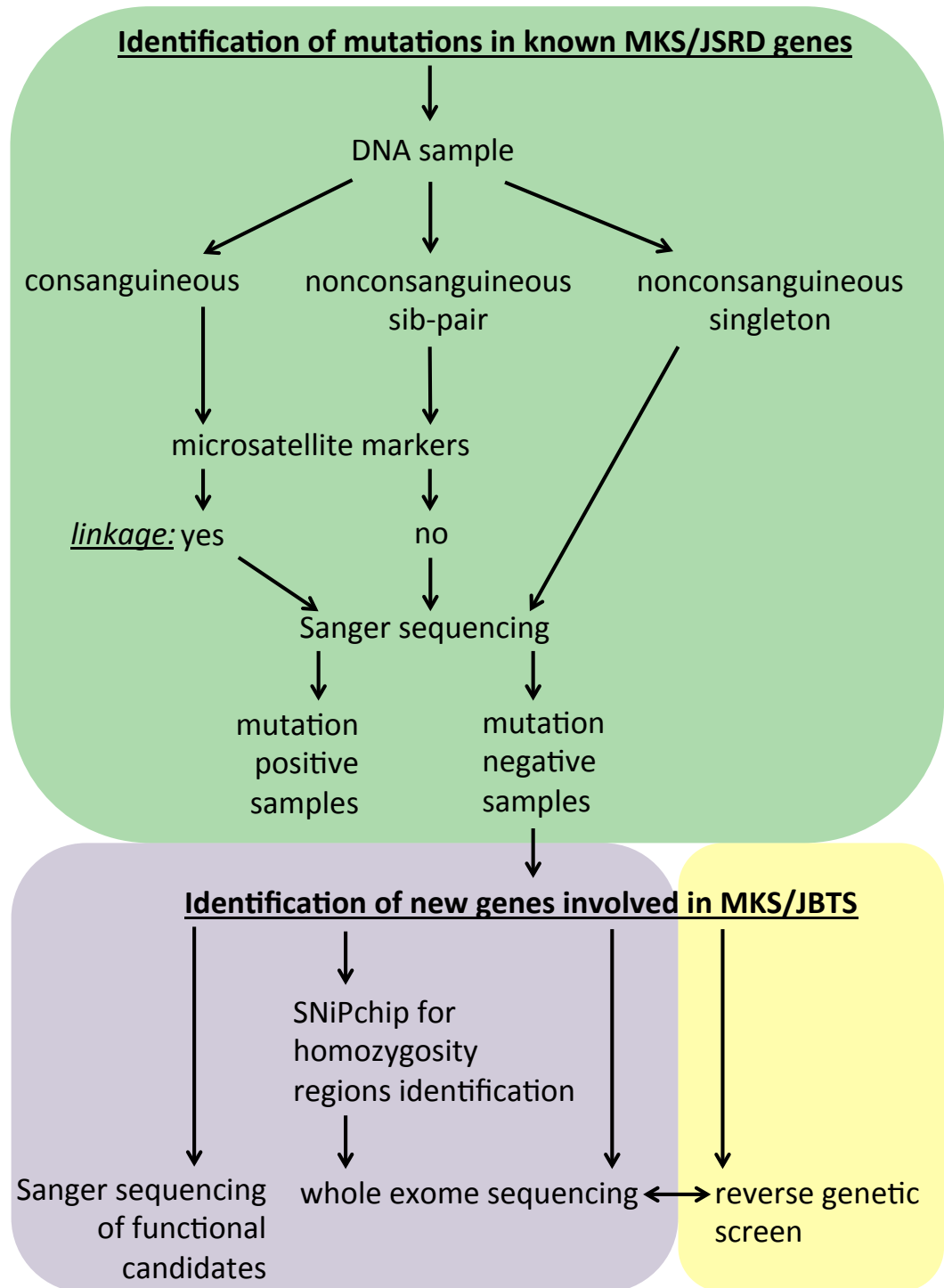


Figure 3-1. Flow diagram presenting strategy of mutation identification if known and novel MKS/JSRD genes. Green background represent strategy for subchapter 3-1, there samples were analysed for linkage to the locus using microsatellite markers and were Sanger sequenced for mutation identification. Purple background represents strategy for subchapter 3-2 where new functional candidate genes were Sanger sequenced, samples that showed no linkage to know loci or were mutation negative were prioritised for SNIpchip analysis and/or WES. Yellow background represents strategy described in chapter 3-3, a whole genome siRNA screen for defects in ciliogenesis. Hits from the screen were cross-compared to the WES results to prioritise variants.

The following two subchapters (3-2 and 3-3) describe efforts to identify new genes as a cause of the MKS/JSRD phenotype. In subchapter 3-2 samples that were

shown to have no linkage to known MKS/JSRD loci and/or were mutation negative in the screened genes were prioritised for Sanger sequencing of new functional candidates. Some of the samples were SNIpChipped for identification of homozygous regions and/or whole exome sequenced (**Figure 3-1**, purple background). Subchapter 3-2 focuses on a reverse genetics siRNA screen for identification of genes required for ciliogenesis/cilia maintenance. Highest confidence hits were used for filtering of variants identified in the WES (**Figure 3-1**, yellow background).

3.1 Mutation screening, founder mutations and genotype-phenotype correlations in MKS and associated ciliopathies

MKS (**1.2.2.1**) and JSRD (**1.2.2.2**) are heterogeneous conditions (**Table 1-1, 1-2**), which means that mutations in different genes can cause the same phenotype. They are also allelic to each other, meaning that mutations in the same gene can cause both phenotypes. To date, mutations in 12 genes have been identified to cause MKS (**Table 1-1**), whereas mutations in 22 genes are a cause of JSRD (**Table 1-2**). Although these genes have been screened in various cohorts by several different research groups, causative mutation identification was possible in only about 50% of investigated families. Therefore, there is still scientific and clinical benefit in identifying new loci linked to MKS and JSRD, as it is likely lots more genes will be identified as many of the new genes account for only one family

104,105,216

3.1.1 Research rationale

The genetic heterogeneity and phenotypic variability in MKS and JSRD have hindered the development of an evidence-based strategy for genetic diagnosis of families. To facilitate this process, the unequivocal identification of pathogenic variants, genotype-phenotype correlations and founder mutations in specific ethnic groups have important clinical utility.

The aim of this chapter was to present a developed strategy to identify causative mutation in MKS/JSRD families, founder mutations and genotype-phenotype correlations if present. The research was initially focused on MKS genes until, during the progress of this research, allelism for MKS and JSRD was demonstrated. Families with MKS and JSRD were ascertained and DNA was collected to form our screening patient cohort. Genotyping and sequencing were

started in the University of Birmingham, where the group was established, and carried on in the University of Leeds.

3.1.2 Microsatellite screening of MKS/JSRD loci

To identify causative mutations in 92 MKS families, genotyping analyses were first performed using microsatellite genetic markers. Microsatellites are short tandem repeats of 2 – 6 base pairs located in the non-coding part of the genome^{217,218}. Genetic markers were chosen based on their close proximity to the flanking ends of the gene of interest, minimizing the possibility of an intervening cross-over event (<http://genome.ucsc.edu/cgi-bin/hgGateway>). Markers were also chosen for their high heterozygosity values so that the identification of a homozygous variant would be of greater significance (<http://www.cephb.fr/en/cephdb/browser.php>). Primer sequences were obtained from <http://www.ncbi.nlm.nih.gov/unists/>. DNA from affected individuals, from all available families, were genotyped for markers flanking eight loci of interest (**Table 3-1**). All collected samples were genotyped, as in most cases clinical notes were either incomplete or absent. Also distant consanguinity could not be excluded in the reported non-consanguineous cases.

GENE	MARKER	FORWARD PRIMER (WITH 'FAM' DYE)	REVERSE PRIMER
<i>MKS1</i>	D17S1606	TGGTATTCAATCCTGGAGC	TGATGAGTCTTCATAGCCCC
	D17S1290	GCCAACAGAGCAAGACTGTC	GGAAACAGTTAAATGGCCAA
<i>TMEM216/TMEM138</i>	D11S4191	GCAAGATGGCCAATTAGAAG	TTTTGGTTGGAATGTAGTTGTTTAT
	D11S4076	CATGAATGCTCTGTCCC	AACCCCTGGAAAATAGACT
<i>TMEM67</i>	D8S1988	CCTTTGGACTCAGACCAGAA	TAGTCAGAGTCCTCAGAGAAACA
	D8S1699	CAACCTGACCCTGCCA	CATGATGTTCTAAGCATATCTGC
<i>CEP290</i>	D12S1719	TCCTCCAGTTTCAGTAATGTTT	GGTGGTTGATGCCTGTAA
	D12S1710	AGGTTTCTGGGTTCCCTGATA	CCATAATCCGTAGGAGCAA
<i>RPGRIP1L</i>	D16S3034	TAATCTAGTTAAAGATGCAACTGCC	GCTCAGAAGTTTTGATGCC
	D16S771	GTCCAAAACACCACCTCTA	AAGTAGATCAGTCATCTTGCTGC
<i>CC2D2A</i>	D4S1511	AGCCTCTGTAATCTTGTGTG	TCCATTACTCAGGGCTCTC
	D4S2960	AAGGCTTATCATTAAAGAATCCTA	TGAGGGTATAGTTACCATCTTTT
<i>NPHP3</i>	D3S1596	ATCAATGCCCTGCTCATTAC	CCTGCATCATGTGCTCTC
	D3S1290	TTGCAGTAATGACCATAGGG	AACAATTAGGGTAATGGGGC
<i>TMEM237</i>	D2S2309	TGTCAGGCACTTCGCTA	TGCTTCTTATTGTACCCAAA
	D2S1384	AATAGAGGGCCCTTGCTTAA	TTTGGATAAAAGGTATTTTGC

Table 3-1. Microsatellite genetic markers for the indicated MKS and JSRD genes. Note that *TMEM216* and *TMEM138* share a locus and therefore the same markers were used for genotyping.

Genes with two homozygous flanking markers in a consanguineous patient were subsequently Sanger sequenced as they were compatible with linkage to the locus. In non-consanguineous siblings shared haplotypes indicated potential linkage and were also sequenced. The results for analysed genetic markers are presented in **Table 3-2**.

Physical location (Mbp)	52.95	53.68	59.76	61.12	93.55	96.07	85.68	88.2	51.69	53.06	14.72	16.06	133.55	134.47	202.33	204.93
	MKS1 CHR.17	MKS1 CHR.17	TMEM216/1 38 CHR.11	TMEM216/1 38 CHR.11	TMEM67 CHR.8	TMEM67 CHR.8	CEP290 CHR.12	CEP290 CHR.12	RPGRIP1L CHR.16	RPGRIP1L CHR.16	CC2D2A CHR.4	CC2D2A CHR.4	NPHP3 CHR.3	NPHP3 CHR.3	TMEM237 CHR.2	TMEM237 CHR.2
ID	D17S1606	D17S1290	D11S4191	D11S4076	D8S1988	D8S1699	D12S1719	D12S1710	D16S3034	D16S771	D4S1511	D4S1567D4 S2960	D3S1596	D3S1290	D2S2309	D2S1384
3			x x	x x			x x	x x	x x	nsd nsd	150	x x				
10			105 124	157			231	nsd nsd	272	253	150 183	233 240	nsd nsd	220		
13			105 115	155 157			235	264 278	275	257	179	233	nsd nsd	204 220		
16			107	140 161			223 233	275 278	272 277	249 253	179 183	240 245	85 114	210	195	147
17			107	148 164			223 233	264 278	272 277	249 261	170 183	240 245	85	210 214	195	142 147
21			x x	x x			223	267 276	x x	x x	x x	x x				
25			106 107	163 167			223 235	267 278	276	267	179 183	233 242	85 116	214 218		
29A 33A (TMEM67)																
36 36A			105 128	155 157			223 225	274 278	275	253	179	233 238	85 105	214 220		
			106 107	166 167			223 235	273 275	272 275	253	160 183	233 243	85	220 222		
39 (CEP290)			106	142			223	278	272	266	x x	239 241	84	207 213		
42 43 (TMEM138)			109	157			223 236	267	273 275	253 257	150 179	240 245			195 205	138 147
70 (TMEM67)			109	157			236	267 273	273 275	253 257	150	240 245			192	143
78 388			122 128	153 157			232	275 277	272 275	253 261	150 175	233 242	104 116	209 219		
66F1 66F2	130 160 165 167	192 204 192 213	113	167			223 233	275 278	272	257 261	160 183	240 242	85 113	214 220	192	162
			107 113	157			233	275	272 277	253 264	150 183	240 242	113 116	212 220	195	142 155
67FB																
51 (TMTM67)			105 130	151 157			223 225	x x	273 277	253?	179	240				
73 (TMEM67)																
76 (TMEM67)																
P95 (TMEM67)																
102 103 244 270 (MKS1)	130	191			x x	x x	223 231		273	256 260					189 195	143 152
106 (MKS1)																
111	130 166	196 204	106	148 164	287 291	210 221	223 236	267 278	276	263 266	161	240	84 116	211		
112	152 164	196 200	106	148 154	287 291	210 221	223 235	267 278	271 275	253 261	151	240	84 116	211		
115	130 158	188 192	105 113	149 159			223 238	268 273	272	257 261	179	233 239	84 116	219		
121	x x	x x	x x	x x			x x	x x	x x	x x	150 179	x x				
128 (CC2D2A)	152 162	164 200	105 122	148 157	200 201	217	223 235	268 273	272	257	150	241				
134	154 160	184 204	105 100	156 160			233	275	272	263	150	233 241	85 114	214 224		
135	x x	x x	105 115	155	x x	206 210	223 233	x x	272	257	150	243	85	216 220		
144 145 (BBS12)			107 124	166 155			233 235	267 276	272 275	249 261	150 179	240	85	212 214		
			107 124	155			229 235	269 280	272	257 261	150 179	241 249	85	214 222		
150	152 173	168 204	107 120	148 155			233	273 275	273 275	253	150	238 245				
151	x x	188 200	107 109	153 157			223 236	277	272 275	261	179	233				
152	154 162	188	105 120	155 157			233	275	273 275	249 253	150 179	241				
153	nsd nsd	192 204	124 126	156 157	207 201	210	233	263 272	273 275	263 264	151 184	240 242	85 105	210 214		
154 (CEP290)	100 100	192 204	124 126	156 167	287 291	202 214	233	263 272	273 275	263 264	151 184	238 240	85	210		
157	130 158	188 196	105 107	153 155			223	273 278	277	257	150	233 240	85	214 220	180 192	138 147

Physical location (Mbp)	52.95	53.68	59.76	61.12	93.55	96.07	85.68	88.2	51.69	53.06	14.72	16.06	133.55	134.47	202.33	204.93
	MKS1 CHR.17	MKS1 CHR.17	TMEM216/1 38 CHR.11	TMEM216/1 38 CHR.11	TMEM67 CHR.8	TMEM67 CHR.8	CEP290 CHR.12	CEP290 CHR.12	RPGRIP1L CHR.16	RPGRIP1L CHR.16	CC2D2A CHR.4	CC2D2A CHR.4	NPHP3 CHR.3	NPHP3 CHR.3	TMEM237 CHR.2	TMEM237 CHR.2
ID	D17S1606	D17S1290	D11S4191	D11S4076	D8S1988	D8S1699	D12S1719	D12S1710	D16S3034	D16S771	D4S1511	D4S1567D4 S2960	D3S1596	D3S1290	D2S2309	D2S1384
157A					287 291	216		272 274			150 180	233 240	85	214 220		
158 (CC2D2A)	158 162	196	124 126	140	291 295	218 223	223	267 275	272	245 265	150	245	85 116	220 222	192	147 152
162 163 (TMEM216)	nsd nsd	188 196	104 120	148 156	287 291	217 210	233		272 275	245 253	150 183	233 243			192 195	147 152
166 (RPGRIP1L)	154 165	196 200	107 120	148 155	288 291	212 214	223 235	267 276	275	245 253	150	240				
167	130 152	172 188	107 120	148 155	288	210 216	233 237	x x	275	253 261	150 179	238 240				
168	163	168 200	105 132	155 157			233 235	274 276	272 275	257 261	150 179	240 246				
169	161 165	188 192	105	148 155			235 237	269 274	273 275	249 253	150	233 239				
170 (TMEM67)																
175			105 115	140 155			225 234	264 273	272 275	249 257	150 183	241 243	85	214 224	192 195	147 152
176 177 (TMEM67)																
178	158	192 196	107 109	157 159			233	273 276	273 275	257	150	237 241	114 116	nsd nsd		
179	154 173	192 200	122 126	148 157			223 233	267 273	273	253 257	150 183	241	x x	210 214		
180 (CC2D2A)	163	164	113 117	148 157			233	273	272	253	150	245	85 116	210 220		
183	165 169	168 192	105 130	157 159	291 295	216 220	231 237	273 275	272	245 257	150 183	238 242	85 114	210 220		
184	130 160	196 200	122	155	288 291	221 225	225 233	264 275	272	261	150 183	233 240	112 114	214 224		
185			105 109	157 161			233 239	267 274	272 275	253 261	150	234 240	85 116	212 214		
186 (TMEM67)																
202	130	161 196 204			292	204 216	223 237		272 277	260	151 179	96 108				
205 (TMEM67)	149	191 208			288	216	232		272 277	253						
206	153	189 192			292	216 210	223 233		272 275	253 257	175 179	98 100	112 114	210 214		
207 (RPGRIP1L)	155 159	184 196			284 292	216 218	223 237		275	257						
210 239 (CEP290)	130 155	192 200			288	214 216	223 238	271 278	276	253	151	96 100	85	210 224		
211	161 163	188			284 288	216 219	223 238	x x	272 275	253	179 183	98 100	85 103	nsd nsd		
212	153 161	188 192			288 292	x x	223 233		272 275	249 253			85 113	214 220		
213	130 152	167 183			288 296	x x	223 233		275	249 253			85 113	nsd nsd		
243 252 (NPHP3)	130 152	167 183			286	213 222	223 233		274 276	253 256	151 179	98 100	105 114	214 220		
217	130 152	167 183			286 294	211 217	233		272 276	253 256	151	98	105 114	214 220		
218	157 159	200 204			292	210 216	223 231		271 273	249 253	161 183	100 108	85	210 224		
219					288 292	214 217	223 233		275	261 264	151 179	96 98	105 116	210 220	192 195	152
221	160 162	192 196			288 292	214 216	233		271 275	x x	151 183	96 106	85 116	210 220		
222	130 165	191 199			282 286	216	223 233		271	260 267	nsd nsd	245	85	nsd nsd		
227	161	200			284 288	210 212	223 233		271 273	253	151 179	98 100	85 105	210		
230	152 156	171 195			286 290	210 216	223 233		271 273	260 263	151	96	85 116	214		
231	165 170	187 195			286 290	212 224	223 233		271	252 260	150	240 245	85	208	195	147 164
256	159 161	196 200			286	208 214	223 233		271 274	248 252	151 183	98	85	212 214	195	143 152

Physical location (Mbp)	52.95	53.68	59.76	61.12	93.55	96.07	85.68	88.2	51.69	53.06	14.72	16.06	133.55	134.47	202.33	204.93
	MKS1 CHR.17	MKS1 CHR.17	TMEM216/1 38 CHR.11	TMEM216/1 38 CHR.11	TMEM67 CHR.8	TMEM67 CHR.8	CEP290 CHR.12	CEP290 CHR.12	RPGRIP1L CHR.16	RPGRIP1L CHR.16	CC2D2A CHR.4	CC2D2A CHR.4	NPHP3 CHR.3	NPHP3 CHR.3	TMEM237 CHR.2	TMEM237 CHR.2
ID	D17S1606	D17S1290	D11S4191	D11S4076	D8S1988	D8S1699	D12S1719	D12S1710	D16S3034	D16S771	D4S1511	D4S1567D4 S2960	D3S1596	D3S1290	D2S2309	D2S1384
257	159 161	196 200			280 288	212 214	235	235	271 275	253 257	151	98 100	116	nsd nsd		
258	159 161	192 196			280	212 214	223 235	nsd nsd	275 277	257						
242	166 161	170 187			286 290	210 213	232	271 274	276	266 260			85	218 224		
247	120 120	190 190			288	196 208	223 231		271 273	264	170 98	98	85 113	214 224		
250	120 168	183 190			284 292	216 222	223 233		271 273	266	161 98	98	85 105	214		
251	129 138	187 199			292	214 222	223 233		271 273	252 256	151 96 98	98	85	210 226		
255	161 163	187			x x	210 222	223 236		273 275	260 263	151 175 96	96	105 116	210 214		
261	156 160	195 199			286	216 220	233 273	271 275	252 259	252 259	151 160 243	243	85 116	216 220	196	147
264 (MKS1)	168 170	187			286	211 218	223 233		271 273	252						
269	164 166	167 199			x x	x x	221		274 276	255 259	180 240	240	85 114	214 226		
276	152 167 ?	168 188			287 291	214 216	223 233	273 280	272 274	253 261	151 98	98	116 118	nsd nsd	191 195 138 142	
277	154 161 ?	181 188			287 294	214 218	223 237	264 273	272 274	253 257	151 98	98	nsd nsd	218 222	195 143 156	
281																
284	130	188 192			287	206 216	233	271 273	276	253 257	151 96 96	96	105 113	220 224		
287	152 159 ?	188			287 290	214 222	223 237	271 275	272 276	261	151 96	96	85 113	220 222		
290	160	168			287	214 218	223 237	273	276	257 261	161 96 98	98	85 116	210 214		
292 (CEP290)	156 165	188 192	107	155	283 287	206 218	223 275	272 276	253	253	150 150 232 239	239				
295	133 165	172 188 ?	122	Nsd	287 294	208 210	231	264 273	273 275	253 257	151 98	98	85 116	210 224		
296 (TMEM67)	153 159	204	105 111 142 156		290	218	223 235	278	275	257 261	151 179 108	108	84	212 222		
302 (TMEM67)	161 165	200	167 120 154 156		291 293	204 212	233	278 280	275 279	245 249	150 175 238 240	240	84 104	214 220		
308	x x	x x	x x	x x	x x	x x	x x	x x	x x	x x	x x	x x	x x			
311																
319 (TMEM67)	152 164	168 172	105 124	148 158	291	208	225 233	X x	273 275	253 257	150	239 241	210 214	98 105		
324	158 164	200 204	104 114	157	288 291	208 210	233 235	267 273	272	253 257	151 179 238	238	112 120	216 222		
325	130 164	176 204	121	155	291	208 216	223 233	267 273	272 274	253	168 179 x x	x	84 114	214 222		
326	157 163	192 204	104 121	155 157	291	208 214	223 235	267	272	245 253	168 179 x x	x	114 116	222		
330	162 164	196 204	109 113	154 156	291 293	204 212	233	278 280	275 279	245 249	150 233 240	240	84	214 222		
333 (CEP290)	152 162	192 200	104 148		288 291 244		233 274		271 273	263 257	137 177 x x	x	104 116	212 224		
336 (RPGRIP1L)	130	196	105 111	156	291 223		233 237	276 278	272	253	150 183	240	108	214		
347 (TMEM67)	157 159	187 191	104	152 156	287	208	232 234	266 272	271 275	256	150 183	232 241	84 116	219 221		
351	152	187 191	104 124	154 164	280 287	220 222	222 234 266 272	273 275	256 260	256 260	183	237	84 116	209		
352 (ALG9)	152 158	187	104 124	152 156	287 291	212 218	222 234 266 272	271 273	253 260	260	179 183	237	84 114	209 215		
353	nsd nsd	x x	104	142	nsd nsd	210 230	235	270	nsd nsd	x x	nsd nsd	236	85	221		
360	127 153	191 199	118 120	154	283 287	214 220	232 240	274	271	252 260	nsd nsd	238 242	nsd nsd	219 221		
361	nsd nsd	191 195	104 126	154	287 291	210 220	nsd nsd	268 272	nsd nsd	248 268	150 182	238	84 104	213 219		
366	159 169	nsd nsd	x x	x x	287 218		232 236	263 281	272 274	252	150 174	235 242	nsd nsd	221 223		
368	153 165	x x	nsd nsd	148 158	287	216	x x	258 263	nsd nsd	252 260	nsd nsd	x x	84 116	nsd nsd		
377	129 161	nsd nsd	112 118	nsd nsd	287 291	216 220	232	266-278	nsd nsd	248 264	150 238	238	116	213		

Physical location (Mbp)	52.95	53.68	59.76	61.12	93.55	96.07	85.68	88.2	51.69	53.06	14.72	16.06	133.55	134.47	202.33	204.93
	MKS1 CHR.17	MKS1 CHR.17	TMEM216/1 38 CHR.11	TMEM216/1 38 CHR.11	TMEM67 CHR.8	TMEM67 CHR.8	CEP290 CHR.12	CEP290 CHR.12	RPGRIP1L CHR.16	RPGRIP1L CHR.16	CC2D2A CHR.4	CC2D2A CHR.4	NPHP3 CHR.3	NPHP3 CHR.3	TMEM237 CHR.2	TMEM237 CHR.2
ID	D17S1606	D17S1290	D11S4191	D11S4076	D8S1988	D8S1699	D12S1719	D12S1710	D16S3034	D16S771	D4S1511	D4S1567D4 S2960	D3S1596	D3S1290	D2S2309	D2S1384
378	129 157	x x	108 120	148 154	291 294	216 212	230 232	268 274	272 274	x x	150 178	238	84	213 219		

Table 3-2. Table of genotyping results for MKS and JSRD genes for affected individuals and patients - MKS1, TMEM216/TMEM138, TMEM67, RPGRIP1L, CEP290, CC2D2A, NPHP3 and TMEM237. In the ID column patients are allocated a number to anonymize identity and are highlighted in red if a causative mutation has been identified (presented in brackets). Affected siblings are grouped in the same row. Consanguineous patients are indicated by underscoring ID number. Each column give genotyping data for a single microsatellite marker with the genetic distance, gene under investigation, chromosome and marker name listed in rows. In each column, pink highlighting indicates a locus with causative mutation(s) identified, yellow indicates potential linkage to the locus, vertical lines indicate linkage excluded previously by other researchers, and horizontal lines indicate excluded linkage based on Sanger sequencing data; Abbreviations: 'x' : genetic marker size could not be interpreted or there was no signal for that sample; 'nsd' no sizing data-(due to poor quality or no ladder was added, therefore the size of the PCR product could not be determined). Blank boxes indicate that genetic markers were not analysed for that sample.

3.1.2.1 Genotyping – results

	<i>MKS1</i>	<i>TMEM216/TMEM138</i>	<i>TMEM67</i>	<i>CEP290</i>	<i>RPGRIP1L</i>	<i>CC2D2A</i>	<i>NPHP3</i>	<i>TMEM237</i>
Linkage excluded by genotyping	69	19	46	71	70	48	66	7
Linkage excluded by sequencing	23	54	35	5	8	18	2	63
Linked to the locus and mutation confirmed	6	4	19	7	3	3	2	1
Linked to the locus but mutation excluded	5	11	5	12	15	21	3	2
Linked to the locus but not sequenced yet	0	1	1	5	1	3	6	0
Not analysed or didn't work	19	33	16	22	25	29	43	49

Table 3-3. Summary of the genotyping results in Leeds MKS/JSRD patient cohort. The table summarizes the numbers of families with/without linkage to particular loci.

Genotyping allowed exclusion of 396 cases from sequencing (**Table 3-3**), bearing in mind that about a fifth of markers were not genotyped because the PCR reaction failed or the data was not interpretable. *TMEM216/TMEM138* and *TMEM237* were reported as a novel genes involved in MKS/JSRD by Prof. Colin Johnson's group (Dr Clare Logan and Katarzyna Szymanska, respectively) and they were first Sanger sequenced in the whole MKS/JSRD cohort. Sequencing the *NPHP3* gene was not feasible, and all reported mutations were kindly provided by collaborator Prof. Friedhelm Hildebrandt (University of Michigan). Mutations in this gene have been reported to be a cause of so-called "Meckel-like syndrome", characterised by polycystic kidneys and liver fibrosis ¹⁰³.

3.1.3 Sequencing of MKS/JSRD cohort

Samples, where genotyping analysis showed compatible linkage to MKS/JSRD loci, were taken forward for direct sequencing. This analysis focused on the seven most frequently mutated genes in MKS/JSRD ¹¹¹: *MKS1*, *TMEM216*, *TMEM67*, *CEP290*, *RPGRIP1L*, *CC2D2A* and *TMEM237*. *TMEM138*, the JSRD16 gene was also sequenced. Three remaining genes (*TCTN2*, *B9D1* and *B9D2*) have only reported private mutations in single families ^{104,105,216} and were therefore not further investigated.

All known coding exons and additionally known intronic founder mutation in *CEP290* were Sanger sequenced for all the genes (**Table 2-3** and **Appendix 2**). Once a putative mutation was identified, the variant was first checked for segregation in the remaining family members if available. If the change was not previously reported in the literature it was checked for presence in the dbSNP, EVS and 1000Genome datasets, and sequenced in a cohort of the best available ethnically-matched controls (n=96). Pathogenicity of novel missense mutations was determined using online tools: PolyPhen2²¹², MutationTaster²¹⁹ and SIFT²¹³.

3.1.3.1 Sequencing results

To define the allelic series of pathogenic mutations for eight known MKS/JSRD genes, a cohort of 87 separate individuals affected with MKS/JSRD were screened. This group included all patients compatible with linkage to a particular locus (**Table 3-2**) and all non-consanguineous singletons. A total of 49 consanguineous and 18 non-consanguineous families were sequenced. Biallelic mutations were identified in 25 consanguineous and 13 non-consanguineous families (**Table 3-4**), including a total of 18 previously unreported mutations¹¹¹.

SAMPLE		MUTATION			PHENOTYPE						
ID	ETHNICITY	GENE	ALLELE 1	ALLELE 2	OE	PK	PD	DPM	CLP	DWM	OTHER
CONSANGUINEOUS											
102+ 103+ 244+ 270	Pakistani	<i>MKS1</i>	c.1448_1451dupCAGG	c.1448_1451dupCAGG	+	+	+		+		Short neck, low set ears, bilateral talipes, syndactyly, micropenis, situs inversus, congenital heart defect inc. dextrocardia, short femurs and short spindle-shaped tibiae, deformed tongue
264	Jordanian	<i>MKS1</i>	c.1408-35_1408-6del30 ^H	c.1408-35_1408-6del30 ^H							diagnosed with MKS
42+ 43	Pakistani	<i>TMEM138</i>	c. A287G p.H96R	c. A287G p.H96R	+	+		+			
29A+ 33A	Pakistani/ Mirpuri	<i>TMEM67</i>	c.1575+1G>A	c.1575+1G>A	+	+	+	+			
70	Pakistani/ Mirpuri	<i>TMEM67</i>	c.1575+1G>A	c.1575+1G>A		+				+	
76	Pakistani/ Mirpuri	<i>TMEM67</i>	c.1575+1G>A	c.1575+1G>A	+	+					
77117	Pakistani	<i>TMEM67</i>	c.1575+1G>A	c.1575+1G>A							diagnosed with MKS
51	Pakistani/ Mirpuri	<i>TMEM67</i>	c.870-2A>G	c.870-2A>G	+						
73	Pakistani/ Mirpuri	<i>TMEM67</i>	c.870-2A>G	c.870-2A>G	+	+		+			

SAMPLE		MUTATION			PHENOTYPE						
ID	ETHNICITY	GENE	ALLELE 1	ALLELE 2	OE	PK	PD	DPM	CLP	DWM	OTHER
319	British	<i>TMEM67</i>	c.1321C>T p.R441C	c.1321C>T p.R441C		+		+			some dilation of pancreatic ducts, hydrocephalus, posterior fossa cyst
347	Pakistani	<i>TMEM67</i>	c.1321C>T p.R441C	c.1321C>T p.R441C							diagnosed with MKS
67FB	Pakistani	<i>TMEM67</i>	c.647delA, p.E216fs*221	c.647delA, p.E216fs*221	+	+		+	+		
P95	Pakistani	<i>TMEM67</i>	c.1127A>C p.Q376P	c.1127A>C p.Q376P	+	+		+			
125	Omani	<i>TMEM67</i>	c.383_384delAC p.H128fs*140	c.383_384delAC p.H128fs*140	+	+	+			+	
170	Turkish	<i>TMEM67</i>	c.1674+1G>A ^N	c.1674+1G>A ^N							diagnosed with MKS
205	Chinese	<i>TMEM67</i>	c.1645C>T p.R549C ^N	c.1645C>T p.R549C ^N		+		+			hypoplastic cerebellum, small fourth ventricle with large cisterna magna, small defect in superior aspect of occipital bone
C28	Pakistani	<i>TMEM67</i>	c.274G>A p.G92R ^N	c.274G>A p.G92R ^N							MTS, coloboma, mental retardation
39	Pakistani/ Mirpuri	<i>CEP290</i>	c.1429C>T p.R477* ^N	c.1429C>T p.R477* ^N		+		+			
292	Pakistani	<i>CEP290</i>	c.954delT p.S318fs16* ^N	c.954delT p.S318fs16* ^N	+						
333	Pakistani	<i>CEP290</i>	c.5744insT p.G1915Ffs*1 ^N	c.5744insT p.G1915Ffs*1 ^N	+						
207	Pakistani	<i>RPGRIP1L</i>	c.1945C>T p.R649* ^N	c.1945C>T p.R649* ^N	+	+	+				small cerebellum
336	Pakistani	<i>RPGRIP1L</i>	c.1945C>T p.R649* ^N	c.1945C>T p.R649* ^N							diagnosed with MKS
158	Pakistani	<i>CC2D2A</i>	c.3540delA p.R1180Sfs*6 ^N	c.3540delA p.R1180Sfs*6 ^N	+	+	+	+	+		low set ears, pulmonary hypoplasia, intestinal malrotation, markedly enlarged pancreas- irregular ducts on histology, brain shows dilated fourth ventricle with small cerebellum, poorly developed pyramidal tracts and some possible dysplasia in the basal ganglia
180	Pakistani	<i>CC2D2A</i>	c.3540delA p.R1180Sfs*6 ^N	c.3540delA p.R1180Sfs*6 ^N	+	+	+			+	
261	Jordanian	<i>TMEM237</i>	c.1066_1067dupC p.Q356Pfs*23	c.1066_1067dupC p.Q356Pfs*23							maningomyelocele, developmental delay, cortical visual impairment
178	Pakistani/ Mirpuri	<i>TMEM67</i>	c.1645C>T p.R549C ^N	not detected							diagnosed with MKS
16+17	Pakistani	<i>CC2D2A</i>	c.685_687delGAA p.E229del	not detected		+		+			
66F1+ 66F2	Pakistani	<i>CC2D2A</i>	c.685_687delGAA p.E229del	not detected	+	+	+	+			Absent uterus, micrognathia, bilateral talipes, low set ears, wide spread eyes
NON-CONSANGUINEOUS											

SAMPLE		MUTATION			PHENOTYPE						
ID	ETHNICITY	GENE	ALLELE 1	ALLELE 2	OE	PK	PD	DPM	CLP	DWM	OTHER
106	British	<i>MKS1</i>	c.1408-35_1408-7del29	c.1408-35_1408-7del29	+	+	+	+			
77172	Finnish	<i>MKS1</i>	c.1408-35_1408-7del29	c.811delC p.H271fs*29 ^N							diagnosed with MKS
74699	British	<i>MKS1</i>	c.1408-35_1408-7del29	c.1408-35_1408-7del29							diagnosed with MKS
162+ 163	British	<i>TMEM216</i>	c.253C>T p.R85* [†]	c.253C>T p.R85* [†]	+	+	+	+	+		facial dysmorphism, postural deformities of limbs, small perimembranous ventricular septal defect, intestinal malrotation
176+ 177	British	<i>TMEM67</i>	c.1426C>T p.P476S ^{††}	c.2440-3C>A	+	+	+	+			flexion deformity of elbows and wrists, low set ears
186	British	<i>TMEM67</i>	c.755T>C p.M252T	c.653G>T p.G218V	+	+					
302	British	<i>TMEM67</i>	c.755T>C p.M252T	c.651+5G>A p.V217Vfs ^N		+		+		+	agenesis of corpus callosum
83527	Norwegian-Indian	<i>TMEM67</i>	c.755T>C p.M252T	c.2882C>A p.S961Y ^N	+	+		+			
74406a+b		<i>TMEM67</i>	c.1351C>T p.R451*	c.2018T>A p.V673A		+		+			mental retardation, retinal coloboma
210+ 239	Dutch	<i>CEP290</i>	c.679_680delGA p.E227Sfs*2	c.1984C>T p.Q662*		+		+		+	abnormal cerebellum, wide nasal bridge, extended abdomen, situs inversus thoracal and abdominal, intestinal rotation, small bladder, uterus duplex
153+ 154	French	<i>CEP290</i>	c.2251C>T p.R751*	c.4864insTdelCG p.R1622Ffs*9 ^N		+					
166	British	<i>RPGRIP1L</i>	c.1829A>C p.H610P	c.721_724delAATG p.N241fs*25	+	+	+			+	
128	British	<i>CC2D2A</i>	c.3544T>C p.W1182R	c.3774_3774insT p.E1259fs*1							diagnosed with MKS
36+ 36A	Pakistani/Gujarati	<i>RPGRIP1L</i>	c.466C>T p.R156C* ^N	not detected	+	+					
111+ 112	Portuguese	<i>CEP290</i>	c.1451delA p.K484fs*8	not detected	+	+	+				
202	British	<i>CC2D2A</i>	c.685_687delGAA p.E229del ^{**}	not detected	+	+					craniofacial abnormalities related to oligohydramnios, bone-cartilage junctions showed disarray

Table 3-4. Clinical data and sequencing results of consanguineous and non-consanguineous patients with MKS and MKS-like phenotypes. Abbreviations: OE: occipital encephalocele, PK: polycystic kidneys, PD: polydactyly, DPM: ductal plate malformation, CLP: cleft lip/palate, DWM: Dandy-Walker malformation; + indicates the presence of a clinical feature. * *in cis* with c.3790G>A het p.D1264N, ** *in cis* with c.3893T>A p.V1298D; [†] p.R85* allele was present in 2/10266 European/African/American controls in European Variant Server (EVS) database, ^{††} p.P476S allele present in 6/7012 European/American controls (EVS), ^{†††} p.R208* allele present in 8/7012 European/American controls (EVS). Remaining changes are excluded in about 10000 European/African/American controls (EVS). ^N indicates novel variants.

3.1.3.2 Genotype – phenotype correlations

There are previous reports of genotype-phenotype correlations in MKS^{110,220,221}. Some of these correlations were confirmed with the available clinical data for this cohort of MKS patients. Occipital encephalocele and polycystic kidneys were almost obligatory features for all patients. Individuals with *TMEM67* mutations frequently had a diagnosis of ductal plate malformation in the liver (n=10/19), but polydactyly was infrequent (n=3/19) compared to *RPGRIP1L* and *CC2D2A* mutated individuals (n=4/6; $p<0.001$, chi-squared test). The Dandy-Walker malformation (or a posterior fossa defect) was occasionally observed in patients with *TMEM67* mutations (n=3/19). Retinal colobomata were only observed for *TMEM67*-mutated individuals (n=2/19). Furthermore, *situs* or gut malrotation defects were never caused by *TMEM67* mutations (n=0/19), in contrast to the occasional manifestation of these clinical features with *MKS1*, *TMEM216*, *CEP290* or *RPGRIP1L* mutations (n=4/17; $p<0.05$, chi-squared test).

3.1.3.3 Conclusions about sequencing analysis

Mutation analysis in MKS/JSRD cohorts suggested that some common mutations have arisen from probable founder effects. These observations will allow initial prioritization of gene and exon screening in affected patients. Patients diagnosed with MKS, particularly those with the additional features of ductal plate malformation and/or retinal coloboma, should be tested for *TMEM67* mutations. Mutations in this gene are the most common cause of these clinical features and, in any case, MKS mutations are most frequent in this gene. In families of Pakistani origin, the *TMEM67* splice-site mutations c.1546+1G>A and c.870-2A>G should be prioritized. In addition, screening for missense mutations between amino acid residues 250 to 570 would detect a third (n=10/29) of all of the *TMEM67* mutations in this cohort and would involve sequencing 10 out of 28 exons only. It is likely that missense or nonsense mutations of conserved arginine residues in this region (for example, R441C, R451* and R549C), may indicate this protein region to be a mutational hotspot. In Pakistani families, the probable founder mutations *RPGRIP1L* c.1945C>T p.R649* and *CC2D2A* c.3540delA p.R1180Sfs*6 should also be prioritised. For families of northern European (including British) origin, testing the *TMEM67* missense mutation p.M252T may be useful, but the most common mutation observed was the *MKS1* “Finn major” mutation²²² (identified initially in the Finnish population). The results demonstrate the broad phenotypic variability in MKS and the lack of clear genotype-phenotype correlations to guide

diagnostic choices. Furthermore, some MKS mutations, such as the *TMEM67* p.R440Q missense mutation, are allelic for JSRD and other ciliopathies. For one patient, the phenotypic and genetic overlap between MKS and JSRD enabled the correct allocation of the causative mutation and the diagnosis. Patient 319 was originally reported as a MKS case and compatible with linkage to the *TMEM67* locus. Screening of the gene identified a homozygous mutation in *TMEM67*, and additional detailed follow-up of clinical features for this patient allowed re-diagnosis for JSRD.

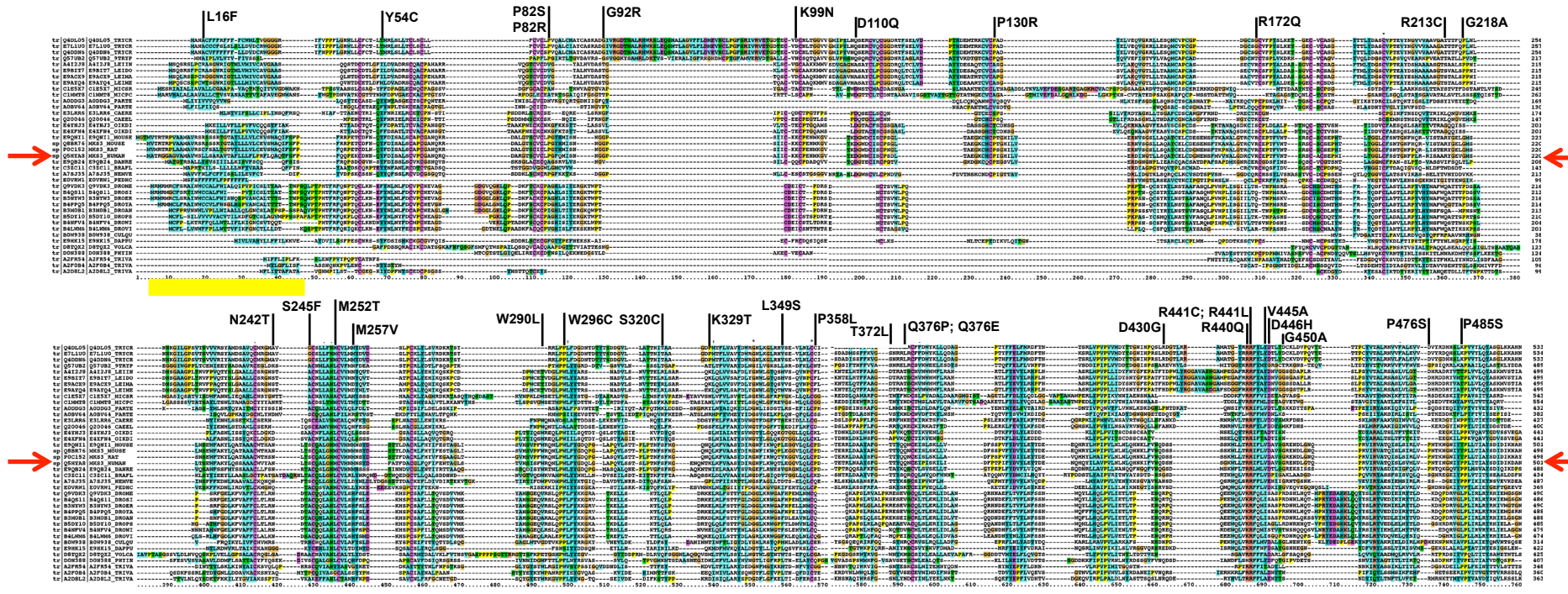
3.1.4 Characterisation of an allelic series in *TMEM67*

The most commonly mutated gene in MKS is *TMEM67*. Mutations in this gene cause not only MKS but are also reported in patients with JSRD, NPHP, SLS and may contribute to the BBS phenotype. A summary of the variants in this genes, that are potentially pathogenic is therefore a valuable clinical and diagnostic resource.

The *TMEM67* protein (also known as meckelin) contains a signal peptide at the *N*-terminus, followed by a cysteine rich domain, a predicted beta-sheet structure, seven predicted transmembrane helices and a coiled-coil domain at the *C*-terminus (**Figure 3-3**)¹⁶³. The protein is highly conserved across species and its orthologue is found in *D. melanogaster* (**Figure 3-2**). To assess its conservation and identify functional domains ClustalX analyses were performed. Multiple orthologous sequences of *TMEM67* were aligned, and functional domains were identified as presented in **Figure 3-2** by regions of amino acids sharing similar physicochemical properties. Missense mutations have also been overlaid on the ClustalX alignment in **Figure 3-2**, which allows an assessment of the degree of conservation for the mutated amino acids.

The positions of all reported mutations in *TMEM67* are presented in **Figure 3-3**. These mutations appear to be fairly evenly spread across the coding regions of the gene. However, missense mutations in the cysteine-rich domain are only observed in JSRD patients (**Figure 3-3**, mutations coloured green with the exception of S245F and M252T that are also reported in MKS). MKS mutations are grouped mainly in the cysteine-rich domain and the *C*-terminus of the protein (amino acids 128-252 and 786-979, respectively). Mutations reported in the JSRD group are localized to the amino acids 82-213 and 637-964, whilst NPHP mutations (coloured blue) are spread infrequently but evenly across the protein. BBS changes (grey) are only reported in the heterozygous state and are suggested to contribute to the BBS phenotype as potential modifier alleles rather than cause of the phenotype. Only one change (D430G) was reported to cause SLS (black). Mutations in the regions amino acids 82-110 and 670-728 appear to be specific to JSRD patients.

This summary of all known and reported mutations in *TMEM67* were collected and submitted as an update to the LOVD (Leiden Open Variation Database), an open source DNA variation database system (<http://www.lovd.nl/3.0/home>). This database is free access and uses a standard nomenclature for the unambiguous reporting of variants and mutations that is approved by the Human Genome Variation Society. The summary of the mutations, and the primary publications that reported them, are presented in **Table 3-5**.



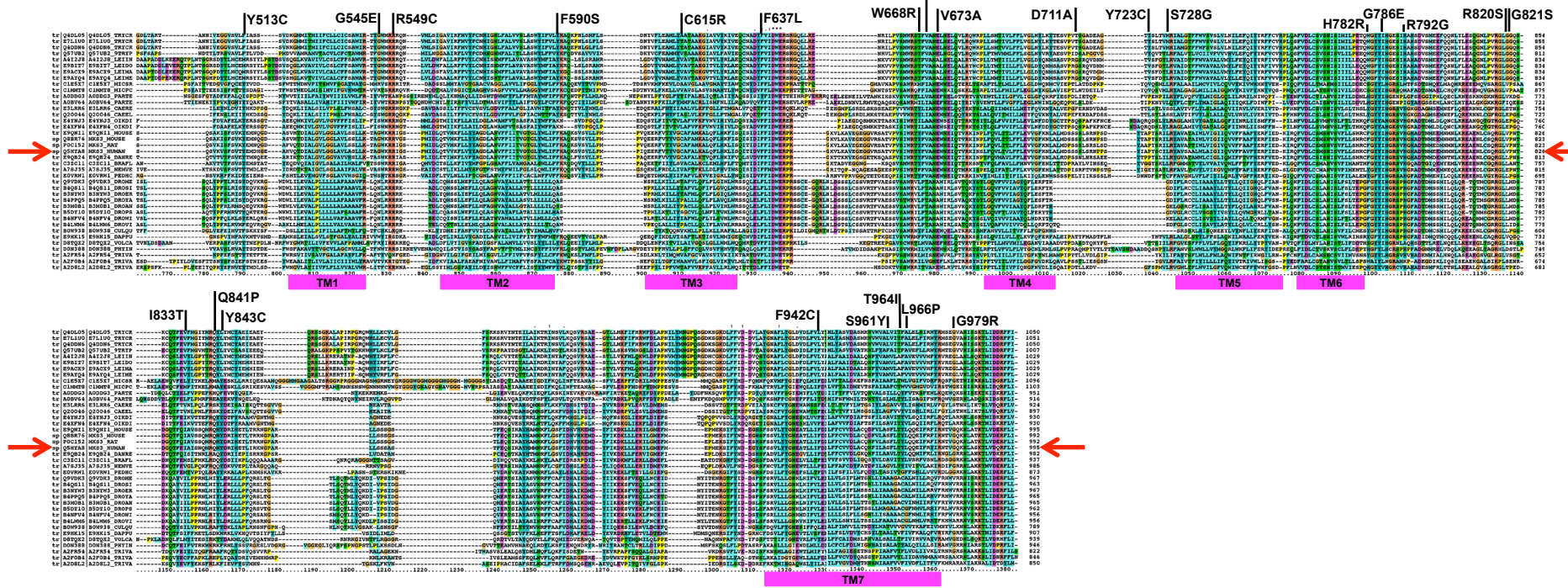


Figure 3-2. ClustalX analysis of the TMEM67 protein. Pink bars indicate transmembrane domains, yellow bar – signal peptide, red arrows indicate human TMEM67 sequence. Missense mutations are placed above the sequences. In the alignment cyan colour represents conserved non-polar residues; green = polar; red = basic; purple = acidic; orange = Glycine a non chiral amino acid; yellow = Proline amino acid. Abbreviations for other animal species with TMEM67 sequences aligned here: TRTCR- *Trypanosoma cruzi*, 9TRYP- *Trypanosoma brucei brucei*, LEIIN – *Leishmania infantum*, LEIDO – *Leishmania donovani*, LEIMA – *Leishmania major*, LEIME – *Leishmania mexicana*, MICSR – *Micromonas sp.*, MICPC – *Micromonas pusilla*, PARTE – *Paramecium tetraurelia*, CAERE – *Caenorhabditis remanei*, CAEEL – *Caenorhabditis elegans*, OIKDI - *Oikopleura dioica*, DANRE – *Danio rerio*, BRAFL – *Brachiosotma floridae*, NEMVE – *Nematostella vectensis*, PEDHC – *Pediculus humanus subsp. corporis*, DROME – *Drosophila melanogaster*, DROSI – *Drosophila simulans*, DROER – *Drosophila erecta*, DROYA – *Drosophila yakuba*, DROAN – *Drosophila ananassae*, DROPS – *Drosophila pseudoobscura pseudoobscura*, DROWI – *Drosophila willistoni*, DROVI – *Drosophila virilis*, CULQU – *Qulex quinquefasciatus*, DAPPU – *Daphnia pulex*, VOLCA – *Volvox carter*, PHYIN - *Phytophthora infestans*, TRIVA – *Trichomonas vaginalis*.

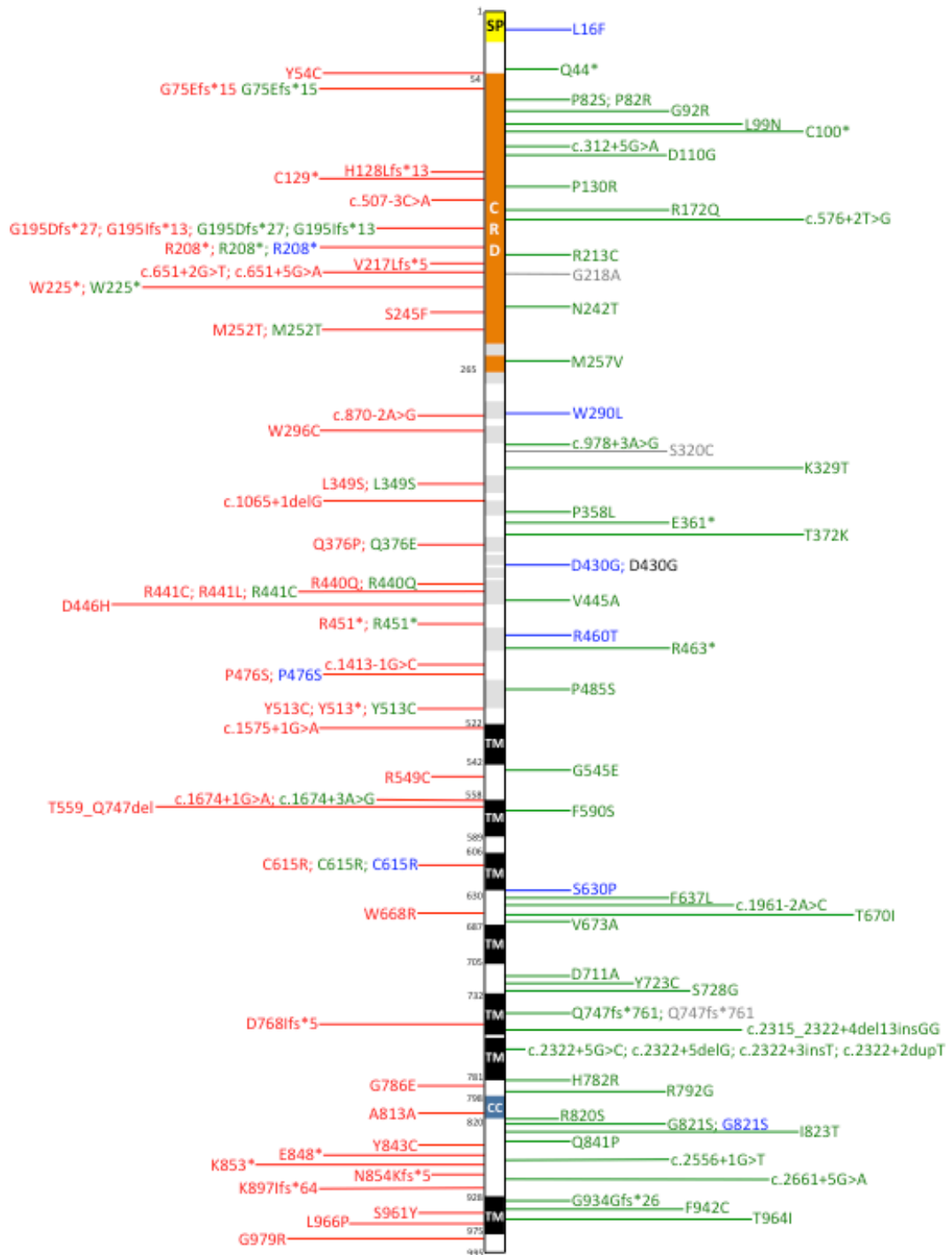


Figure 3-3. Summary of all reported *TMEM67* mutations, the clinical phenotype and the position in relation to domains in the *TMEM67* protein. The mutations are represented as predicted functional changes at the protein level. Red, mutations causing MKS; green, JSRD and COACH; blue, NPHP; grey, reported in BBS; black, SLS. Yellow block (SP), signal peptide; Orange (CRD), cysteine rich domain; black (TM), transmembrane domain. Grey blocks indicate a predicted beta-sheet region.

	DNA change (cDNA)	Published as	Protein	Variant remarks	Reference	dbSNP ID
1	c.46C>T	c.46C>T p.L16F	p.(Leu16Phe)	1 European family with NPHP (het)	Helbritter et al.2012	
1	c.130C>T	c.130C>T p.Q44*	p.(Gln44*)	1 family with JSRD (com-het with c.2461G>C)	Chaki et al.2011	
1	c.161A>G	c.161A>G p.Tyr54Cys	p.(Tyr54Cys)	1 European MKS family (com-het)	Tallila et al. 2009	-
1i	c.224-2delA	IVS1-2delA; 224del89; G75fs*89	p.(Gly75Glufs*15)	1 German-Polish MKS family (com-het), 1 family with JS (com-het with c.1843T>C)	Consugar et al. 2007, Gunay-Aygun et al. 2009	-
2	c.244C>T	c.244C>T p.P82S	p.(Pro82Ser)	1 American family with JSRD (com-het with c.579_580delAG)	Doherty et al.2010	
2	c.245C>G	c.245C>G p.P82R	p.(Pro82Arg)	1 American family with COACH (com-het with c.755T>C)	Doherty et al.2010	
2	c.274G>A	c.274G>A p.G92R	p.(Gly92Arg)	1 Pakistani family with JSRD (hom)	Szymanska et al.2012	
2	c.297G>T	c.297G>T p.K99N	p.(Lys99Asn)	1 American family with COACH (com-het with c.2322+3insT)	Doherty et al.2010	
2	c.300C>A	c.300C>A p.C100*	p.(Cys100*)	1 American family with COACH (com-het with c.2498T>C)	Doherty et al.2010	
2i	c.312+5G>A	c.G312+5G>A	p.?	1 Croatian family with COACH (com-het with c.2498T>C)	Brancati et al.2009	
3	c.329A>G	c.329A>G p.D110G	p.(Asp110Gly)	1 Japanese family with JSRD (com-het with c.2322+5delG)	Tsurusaki et al.2012	
3	c.383_384delAC	383-384delAC H128fs*140	p.(His128Leufs*13)	1 Omani MKS family (hom), 1 Omani MKS family (hom)	Smith et al. 2006, Szymanska et al.2012	-
3	c.387T>A	c.387T>A p.C129*	p.(Cys129*)	1 Italian MKS family (com-het)	Iannicelli et al. 2010	-
3	c.389C>G	c.389C>G p.P130R	p.(Pro130Arg)	1 American family with COACH (com-het with c.675G>A)	Brancati et al.2009, Doherty et al.2010	
4i	c.507-3C>A	c.507-3C>A Splicing	p.?	1 British family with MKS (com-het with c.2935G>A)	Szymanska et al.2013	
5	c.515G>A	c.515G>A p.R172Q	p.(Arg172Gln)	1 American family with COACH (com-het with c.769A>G)	Doherty et al.2010	
5i	c.576+2T>G	IVS6+2T>G	p.?	1 French family with JS (com-het with p.G545E and p.Q747fs*761)	Baala et al.2007	
6	c.579delA	579delA T193fs*221	p.(Gly195Aspfs*27)	2 European MKS families (1 hom and 1 com-het)	Consugar et al. 2007, Iannicelli et al. 2010	-
6	c.579_580delAG	c.579_580delAG p.G195Ifs*13	p.(Gly195Ilefs*13)	1 Moroccan MKS family (hom), 1 Italian family with COACH (com-het with c.1769T>C), 1 American family with COACH (com-het with c.244C>T)	Brancati et al.2009, Iannicelli et al. 2010, Doherty et al.2010	-
6	c.622A>T	622A>T R208*	p.(Arg208*)	4 European MKS families (all com-het), 1 American family with COACH (com-het with c.2522A>C), 2 German families with JSRD (com-het with c.2168A>G, com-het with c.1843T>C), 1 Italian family with SLS (com-het with c.1289A>G), 1 German family with NPHP (com-het with c.2498T>C), 1 family with MKS (com-het with c.2168A>G), 1 family with JSRD (com-het with c.2498T>C), 1 family with JSRD (com-het with c.622A>T), 1 family with JSRD (com-het with c.1843T>C)	Consugar et al. 2007, Kahaddour et al. 2007, Doherty et al.2010, Otto et al.2011, Chaki et al.2011, Halbritter et al.2012	rs137853108
6	c.637C>T	c.637C>T	p.Arg213Cys	1 French family with JS (com-het with c.2131A>C)	Baala et al.2007	
6	c.648delA	c.647delA E216fs*221	p.(Val217Leufs*5)	1 Pakistani MKS family (hom), 1 Pakistani MKS family (hom)	Smith et al. 2006, Szymanska et al.2012	-
6		p.G218A	p.(Gly218Ala)	1 family with BBS (het in cis with p.(S320C), CEP290 p.E1903* (hom))	Leitch et al. 2008	
6i	c.651+2T>G	c.651+2G>T Splicing	p.?	1 French MKS family (com-het)	Kahaddour et al. 2007	-
6i	c.651+5G>A	c.651+5G>A p.V217Vfs	p.?	1 British family with MKS (com-het with c.755T>C)	Szymanska et al.2012	

	DNA change (cDNA)	Published as	Protein	Variant remarks	Reference	dbSNP ID
7	c.675G>A	c.675G>A p.W225*	p.(Trp225*)	1 Italian MKS family (com-het), 1 Italian family with COACH (com-het with c.389C>G), 1 American family with COACH (com-het with c.389C>G)	Brancati et al.2009, Iannicelli et al. 2010, Doherty et al.2010	-
7	c.725A>G	c.725A>G p.N242T	p.(Asn242Thr)	1 American family with COACH (hom)	Doherty et al.2010	
7	c.734C>T	c.734C>T p.Ser245Phe	p.(Ser245Phe)	1 European MKS family (com-het)	Tallila et al. 2009	-
7	c.755T>C	755T>C M252T	p.(Met252Thr)	5 European MKS families (all com-het), 2 American families with COACH (het, com-het with c.245C>G), 3 German families with JSRD (com-het with c.1843T>C, com-het with c.2498T>C), 2 British families with MKS (com-het with c.653G>T, com-het with c.651+5G>A), 1 Norwegian-Indian family with MKS (com-het with c.2882C>A), 3 families with JSRD (com-het with c.1843T>C), 1 family with JSRD (com-het with c.2498T>C)	Consugar et al. 2007, Kahaddour et al. 2007, Tallila et al. 2009, Iannicelli et al. 2010, Doherty et al.2010, Otto et al.2011, Chaki et al.2011, Szymanska et al.2012	-
7	c.769A>G	c.769A>G p.M257V	p.(Met257Val)	1 American family with COACH (com-het with c.515G>A)	Doherty et al.2010	
8	c.869G>T	c.869G>T p.W290L	p.(Trp290Leu)	1 German family with NPHP (com-het with c.1843T>C), 1 family with NPHP/JSRD (com-het with c.1843T>C)	Otto et al. 2009, Chaki et al.2011	
8i	c.870-2A>G	INV8-2A>G; c.870-2A>G Splice change	p.?	1 Pakistani MKS family (hom), 1 Pakistani MKS family (hom), 2 Pakistani families with MKS (hom)	Smith et al. 2006, Kahaddour et al. 2007, Szymanska et al.2012	-
9	c.888G>T	c.888G>T p.Trp296Cys	p.(Trp296Cys)	1 European MKS family (com-het)	Tallila et al. 2009	-
9	c.958A>T	c.958A>T p.S320C	p.(Ser320Cys)	1 family with BBS (het in cis with p.(G218A), CEP290 p.E1903* (hom))	Leitch et al. 2008	rs111619594
9i	c.978+3 A>G	c.978+3 A>G	p.?	1 American family with COACH (com-het with c.2825T>G)	Doherty et al.2010	
10	c.986A>C	c.986A>C p.K329T	p.(Lys329Thr)	1 American family with JSRD (com-het with c.2556+1G>A), 1 family with JSRD (com-het with c.2556+1G>A)	Otto et al.2011, Chaki et al.2011	
10	c.1046T>C	c.1046T>C p.L349S	p.(Leu349Ser)	2 French (com-het) and 1 Algerian MKS family (hom), 2 American families with COACH (com-het with c.2498T>C, com-het with 1843T>C), 1 German family with JSRD (com-het with c.1843T>C), 1 family with JSRD (com-het with c.1843T>C)	Kahaddour et al. 2007, Iannicelli et al. 2010, Doherty et al.2010, Otto et al.2011, Chaki et al.2011	-
10i	c.1065+1delG	c.1065+1delG splice	p.?	1 Palestinian MKS family (hom)	Kahaddour et al. 2007	-
11	c.1073T>C	c.1073T>C p.P358L	p.(Pro358Leu)	1 American family with COACH (com-het with c.2661+5G>A)	Doherty et al.2010	
11	c.1083G>T	c.108G>T p.E361*	p.(Glu361*)	2 American family with COACH (com-het with c.1911A>C)	Doherty et al.2010	
11	c.1115C>A	c.1115C>A p.T372K	p.Thr372Lys	1 Italian family with COACH (com-het with c.2345A>G), 1 American family with COACH (hom)	Brancati et al.2009, Doherty et al.2010	
11	c.1126C>G	c.1126C>G p.Q376E	p.(Gln376Glu)	1 American family with COACH (com-het with c.1674+3A>G)	Doherty et al.2010	
11	c.1127A>C	c.1127A>C Q376P	p.(Gln376Pro)	1 Pakistani MKS family (hom), 1 Pakistani MKS family (hom)	Smith et al. 2006, Szymanska et al.2012	rs137853106
13	c.1289A>G	c.1289A>G p.D430G, c.1259A>G p.D420G	p.(Asp430Gly)	1 Italian family with SLS (com-het with c.622A>T), 1 European family with NPHP (het)	Halbritter et al.2012	
13	c.1319G>A	1319G>A R440Q	p.(Arg440Gln)	4 European MKS families (1 hom and 3 com-het), 1 Italian family with COACH (com-het with c.2182A>G)	Consugar et al. 2007, Kahaddour et al. 2007, Brancati et al.2009, Tallila et al. 2009, Iannicelli et al. 2010	-
13	c.1321C>T	c.1321C>T p.R441C	p.(Arg.441Cys)	1 American family with COACH (com-het with c.1453C>T), 1	Doherty et al.2010,	

	DNA change (cDNA)	Published as	Protein	Variant remarks	Reference	dbSNP ID
				British family with MKS (hom), 1 Pakistani family with MKS (hom)	Szymanska et al.2012	
13	c.1322G>T	c.1322G>T p.R441L	p.(Arg441Leu)	1 French MKS family (com-het)	Iannicelli et al. 2010	-
13	c.1334T>C	c.1334T>C p.V445A	p.(Val445Ala)	1 German family with JSRD (het)	Otto et al.2011	
13	c.1336G>C	c.1336G>C p.D446H	p.(Asp446His)	1 Moroccan MKS family (com-het)	Kahaddour et al. 2007	-
13	c.1351C>T	c.1351C>T R451*	p.(Arg451*)	1 Swedish-German MKS family (com-het), 1 American family with COACH (com-het with c.2498T>C), 1 British family with JSRD (com-het with c.2018T>C), 1 family with JSRD (com-het with c.2018T>C)	Consugar et al. 2007, Doherty et al.2010, Otto et al.2011, Chaki et al.2011	rs116647652
13	c.1379G>C	c.1349G>C p.R450T	p.(Arg460Thr)	1 European family with NPHP (het)	Halbritter et al.2012	
13	c.1387C>T	c.1387C>T p.R463*	p.(Arg463*)	1 German family with JSRD (com-het with c.2891C>T), 1 family with JSRD (com-het with c.2891C>T)	Otto et al.2011, Chaki et al.2011	
13i	c.1413-1G>C	c. 1413-1G>C Splice	p.?	1 French MKS family (com-het)	Iannicelli et al. 2010	-
14	c.1426C>T	c.1426C>T p.P476S; c.1396C>T p.P466S	p.(Pro476Ser)	1 British family with MKS (com-het with c.2440-3C>A), 1 European family with NPHP (het)	Szymanska et al.2012, Halbritter et al.2012	rs145236803
14	c.1453C>T	c.1453C>T p.P485S	p.(Pro485Ser)	1 American family with COACH (com-het with c.1321C>T)	Doherty et al.2010	
15	c.1538A>G	c.1538A>G p.Tyr513Cys, c.1438A>G p.Y513C	p.(Tyr513Cys)	1 European MKS family (com-het), 1 French family with JS (two affected sibs, com-het with c.2315_2323+4del13insGG), 1 American family with COACH (com-het with c.1843T>C), 1 American family with COACH (com-het with c.2497T>C)	Tallila et al. 2009, Baala et al. 2007, Doherty et al.2010	rs137853107
15	c.1538_1539delAT	c.1538_1539delAT p.Y513*	p.(Tyr513*)	1 Senegal MKS family (hom)	Iannicelli et al. 2010	-
15i	c.1575+1G>A	INV15+1G>A Splice change	p.?	1 Pakistani MKS family (hom), 2 Pakistani MKS families (hom), 4 Pakistani families with MKS (hom)	Smith et al. 2006, Kahaddour et al. 2007, Szymanska et al.2012	-
16	c.1634G>A	c.1634G>A	p.(Gly545Glu)	1 French family with JS (com-het with c.576+2T>G and in cis with p.Q747fs*761)	Baala et al.2007	
16	c.1645C>T	c.1615C>T p.R549C	p.(Arg549Cys)	1 Chinese family with MKS (hom), 1 Pakistani family with MKS (het)	Szymanska et al.2012	
16i	c.1674+3A>G	c.1674+3A>G splice	p.?	1 American family with COACH (com-het with c.1126C>G)	Doherty et al.2010	
16i	c.1674+1G>A	c.1674+1G>A splice	p.?	1 Turkish family with MKS (hom)	Szymanska et al.2012	
16i-21i	c.1675-?_2241+?del	c.1675-?_2241+?del p.T559_Q747del	p.Thr559_Gln747del	1 Ivory Coast MKS family (hom)	Kahaddour et al. 2007	-
17	c.1769T>C	c.1769T>C p.F590S	p.(Phe590Ser)	2 Italian families with COACH (com-het with c.G1961-2A>C) (com-het with c.579_580delAG), 1 family with COACH (com-het with c.1961-2A>C)	Brancati et al.2009, Chaki et al.2011	-
18	c.1843T>C	c.1843T>C p.Cys615Arg	p.(Cys615Arg)	1 European MKS family (com-het), 1 Turkish family with NPHP (hom), 1 German family with NPHP (hom), 1 family with JS (com-het with c.224-2delA), 2 American families with COACH (com-het with c.1046T>C, com-het with c.1538A>G), 5 German families with JSRD (com-het with c.755T>C, co-het with c.1911C>A, com-het with c.1045T>C, com-het with c.622A>T), 1 family with JSRD (com-het with c.622A>T), 3 families with JSRD (com-het with c.755T>C), 1 family with NPHP/JSRD (com-het	Tallila et al. 2009, Otto et al. 2009, Gunay-Aygun et al. 2009, Doherty et al.2010, Otto et al.2011, Chaki et al.2011	-

	DNA change (cDNA)	Published as	Protein	Variant remarks	Reference	dbSNP ID
				with c.868G>T), 1 family with JSRD (com-het with c.1045T>C), 2 families with JSRD (hom), 1 family with JSRD (com-het with c.1911C>A)		
19	c.1888T>C	c.1888T>C p.S630P	p.(Ser630Pro)	1 family with NPHP (hom)	Chaki et al.2011	
19	c.1911C>A	c.1911C>A p.F637L	p.(Phe637Leu)	1 American family with COACH (com-het with c.1073T>C), 2 German families with JSRD (het, com-het with c.1843T>C), 1 family with JSRD (com-het with c.1843T>C)	Doherty et al.2010, Otto et al.2011, Chaki et al.2011	
19i	c.1961-2A>C	c.1961-2A>C	p.?	1 Italian family with COACH (com-het with c.1769T>C), 1 family with COACH (com-het with c.1769T>C)	Brancati et al.2009, Chaki et al.2011	
20	c.2002T>C	c.2002T>C p.W668R	p.(Trp668Arg)	1 French MKS family (com-het)	Iannicelli et al. 2010	-
20	c.2009C>T	c.2009C>T p.T670I	p.(Thr670Ile)	1 Italian family with JSRD (het)	Otto et al.2011	
20	c.2018T>C	c.2018T>C p.V673A	p.(Val673Ala)	1 British family with JSRD (com-het with c.1351C>T), 1 family with JSRD (com-het with c.1351C>T)	Otto et al.2011, Chaki et al. 2011	
21	c.2132A>C	c.2132A>C	p.(Asp711Ala)	1 French family with JS (com-het with c.637C>T)	Baala et al.2007	
21	c.2168A>G	c.2168A>G p.Y723C	p.(Tyr723Cys)	1 German family with JSRD (com-het with c.622A>T), 1 family with MKS (com-het with c.622A>T)	Otto et al.2011, Chaki et al.2011	
21	c.2182A>G	c.2182A>G p.S728G	p.(Ser728Gly)	1 Italian family with COACH (com-het with c.1319G>A)	Brancati et al.2009	
21	c.2241G>A	c.2241G>A	p.Gln747fs*761	1 French family with JS (com-het with c.576+2T>G), 1 family with BBS (single het with BBS9: c2804G>A hom)	Baala et al. 2007, Leitch et al. 2008	
22	c.2301delT	c.2301delT p.D768Iifs*5	p.(Asp768Ilefs*5)	1 French MKS family (com-het)	Iannicelli et al. 2010	-
22-22i	c.2315_2322+4del13insGG	c.2315_2323+4del13insGG	p.?	1 French family with JS (two affected sibs, com-het with p.(Tyr513Cys))	Baala et al.2007	
22i	c.2322+5G>C	IVS23+5G>C	p.(Ile775_Ala813delLeu)	1 Algerian family with JS (hom)	Baala et al. 2007	
22i	c.2322+5delG	c.2322+5delG	p.?	1 Japanese family with JSRD (com-het with c.329A>G)	Tsurusaki et al.2012	
22i	c.2322+3insT	c.2322+3insT spl	p.?	2 American families with COACH (com-het with c.115C>A, com-het with c.297G>T)	Doherty et al.2010	
22i	c.2322+2dupT	c.2322+2dupT Splice	p.?	1 American MKS family (com-het)	Iannicelli et al. 2010	-
23	c.2345A>G	c.2345A>G p.H782R	p.(His782Arg)	1 Italian family with COACH (com-het with c.1115C>A)	Brancati et al.2009	
23	c.2357G>A	c.2357G>A p.G786E	p.(Gly786Glu)	1 French MKS family (com-het)	Iannicelli et al. 2010	-
23	c.2374A>G	c.2374A>G p.R792G	p.(Arg792Gly)	1 Swiss family with JSRD (het)	Otto et al.2011	
23	c.2439G>A	c.2439G>A p.A813A splice site	p.[=, ?]	1 Moroccan MKS family (com-het). Last nucleotide of exon 24 before 5' splice donor site.	Kahaddour et al. 2007	-
24	c.2460A>C	c.2460A>C p.R820S	p.(Arg820Ser)	1 Croatian family with COACH (het)	Brancati et al.2009	
24	c.2461G>A	c.2461G>A p.G821S	p.(Gly821Ser)	1 Turkish family with NPHP (hom), 1 German family with NPHP (hom), 1 Turkish family with JSRD (het), 1 Egyptian family with JSRD (het), 2 families with NPHP (hom)	Otto et al. 2009, Otto et al.2011, Chaki et al.2011	
24	c.2461G>C	c.2461G>C p.G821R	p.(Gly821Arg)	1 family with JSRD (com-het with c.130C>T)	Chaki et al.2011	
24	c.2498T>C	c.2498T>C p.I833T, c.2468T>C p.I823T	p.(Ile833Thr)	1 Croatian family with COACH (com-het with c.312+5G>A), 1 Belgian family with COACH (com-het with c.2556+1G>T), 3 American families with COACH (com-het), 1 American family with COACH (com-het with c.1538A>G), 1 European family with NPHP (het)c.2498T>C, com-het with c.1351C>T, com-het with c.300C>A), 1 German family with JSRD (com-het with	Brancati et al.2009, Doherty et al.2009, Doherty et al.2010, Otto et al.2011, Chaki et al.2011, Halbritter et al.2012	

	DNA change (cDNA)	Published as	Protein	Variant remarks	Reference	dbSNP ID
				c.755T>C), 1 German family with NPHP (com-het with c.622A>T), 1 family with JSRD (com-het with c.755T>C)		
24	c.2522A>C	c.2522A>C p.Q841P	p.(Gln841Pro)	1 American family with COACH (com-het with c.622A>T)	Doherty et al.2010	
24	c.2528A>G	c.2528A>G p.Y843C	p.(Tyr843Cys)	1 Italian MKS family (com-het)	Iannicelli et al. 2010	-
24	c.2542G>T	c.2542G>T p.E848*	p.(Glu848*)	1 French MKS family (com-het)	Iannicelli et al. 2010	-
24	c.2556+1G>T	c.2556+1G>T	p.?	1 Belgian family with COACH (com-het with c.2498T>C), 1 American family with JSRD (com-het with c.986A>C), 1 family with JSRD (com-het with c.986A>C)	Brancati et al.2009, Otto et al.2011, Chaki et al.2011	
25	c.2557A>T	c.2557A>T p.K853*	p.(Lys853*)	1 French MKS family (com-het).	Kahaddour et al. 2007	-
25	c.2561dupA	c.2561dupA p.N854Kfs*5	p.(Asn854Lysfs*5)	1 American MKS family (com-het)	Iannicelli et al. 2010	-
25i	c.2661+5G>A	c.2661+5G>A splice	p.?	1 American family with COACH (com-het with c.1083G>T)	Doherty et al.2010	
26	c.2689_2690insTA	c.2689_2690insTA p.K897lfs*64	p.(Lys897lfs*64)	1 French MKS family (com-het)	Iannicelli et al. 2010	-
27	c.2802delA	c.2802delA p.G934Gfs*26	p.(Gly934Gfs*26)	1 American family with COACH (het)	Doherty et al.2010	
27	c.2825T>G	c.2825T>G p.F942C	p.(Phe942Cys)	1 American family with COACH (com-het with c.978+3A>G)	Doherty et al.2010	
27	c.2882C>A	c.2882C>A p.S961Y	p.(Ser961Tyr)	1 Norwegian-Indian family with MKS (com-het with c.755T>C)	Szymanska et al.2012	
27	c.2891C>T	c.2891C>T p.T964I	p.(Thr964Ile)	1 German family with JSRD (com-het with c.1387C>T), 1 family with JSRD (com-het with c.1387C>T)	Otto et al.2011, Chaki et al, 2011	
27	c.2897T>C	c.2897T>C p.L966P	p.(Leu966Pro)	2 European MKS families (both com-het)	Consugar et al. 2007, Tallila et al. 2009	-
28	c.2935G>A	c.2935G>A p.G979R	p.(Gly979Arg)	1 British family with MKS (com-het with c.507-3C>A)	Szymanska et al.2013	

Table 3-5. Table summarising all reported changes in TMEM67 (NM_153704.5) to date. Changes are presented in the form they were reported in the original publication and in the corrected nomenclature, if necessary, to be compatible with LOVD following guidelines from HGVS. Mutation status and the reference publication are indicated.

3.1.4.1 Reclassification of TMEM67 variants of unknown significance

Further research studies were performed to determine the pathogenic potential of variants of unknown significance (VOUS) in *TMEM67*. These were originally identified during routine service testing of *TMEM67* by Dr Ian Berry, in the Diagnostic DNA lab, Yorkshire Regional Genetics Service, Leeds Teaching Hospitals NHS Trust. Investigations into two unrelated MKS patients are presented as an example for our routine procedures.

3.1.4.1.1 Clinic – diagnostic laboratory – research laboratory workflow

3.1.4.1.1.1 Patient 387 phenotype

A couple was referred to the genetics clinic following termination of a pregnancy for suspected diagnosis of MKS. The couple was non-consanguineous and there was no other phenotype of relevance reported in the extended family. This was the first pregnancy of the couple. During the pregnancy, the antenatal scan at 18 weeks gestation showed occipital encephalocele and bilateral multicystic, enlarged kidneys. A clinical diagnosis of MKS was suspected and a termination of pregnancy was performed. Post-mortem examination showed bilateral syndromic cystic renal dysplasia, hepatic fibrosis and an occipital encephalocele. There was no evidence of polydactyly.

3.1.4.1.1.2 gDNA sequencing

Initial direct Sanger sequencing excluded mutations in the *MKS1* gene. Analysis of changes in *TMEM67* revealed two heterozygous changes: c.507-3C>A in intron 4 and c.2935G>A p.G979R in exon 28 (**Figure 3-4**). Both changes were not previously reported in the scientific literature. Both changes were excluded in online available databases for common variants (dbSNP, the 1000 Genomes Project and EVS). Parental DNA sequencing revealed the changes were inherited *in trans*: c.507-3C>A was inherited from the father and c.2935G>A from the mother. The c.507-3C>A change suggested disruption in splicing of an intron/exon boundary and c.2935G>A was localised in the last exon of *TMEM67*. *In silico* analyses were performed to predict the effects of these variants on the splicing of the *TMEM67* transcript.

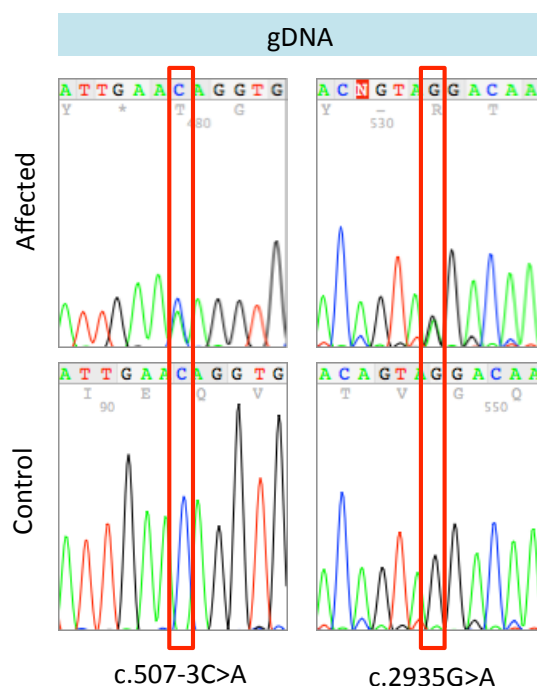


Figure 3-4. Electropherograms presenting sequencing results of genomic DNA of MKS affected patient 387. In red frames (highlighted) are changes identified in the affected patient compared to a normal control individual at the genomic (gDNA) level.

3.1.4.1.1.3 *In silico* analysis

The c.507-3C>A variant is located in intron 4 just before the beginning of the fifth exon. Although it is not localised within a consensus acceptor site, its close localisation to an intron/exon boundary suggests that it has an effect on splicing. To analyse the predicted effect on splicing the following algorithms were used: SpliceSiteFinder-like²²³, MaxEntScan²²⁴, NNSplice (also known as the BDGP fruitfly algorithm)²²⁵, GeneSplicer²²⁶ and HSF (Human Splicing Finder)²²⁷ (**Table 3-6**). These were all interrogated through the Alamut v2.2 software package. Predictions from these algorithms suggested that c.507-3C>A may have an effect on the 3' (splice acceptor/end of intron) consensus site (**Table 3.1-9**). NetGene2 (<http://www.cbs.dtu.dk/services/NetGene2/>) was also used, and predicted a weak acceptor splice site (confidence = 0.00) in the wild-type, with abolition of the acceptor site in the mutant sequence (analysis done by Dr Ian Berry).

Splicing algorithm:	c.507-3C>A		c.2935G>A		
	Wild-type confidence score 3':	Mutant confidence score 3':	Wild-type confidence score 5':	Mutant confidence score 5':	Mutant confidence score 3':
SpliceSiteFinder-like (0-100)	80.4	70.4	63.2	73.3	96.1
MaxEntScan (0-16)	8	4.8	1.7	6	10.9
NNSplice (fruitfly) (0-1)	0.7	0	0	0.8	1
Genesplicer	2.3	1.1	0	0	4.2

(0-15)					
Human Splicing Finder (0-100)	84.2	74.9	73	81.4	91.8

Table 3-6. Summary of *in silico* analyses on MKS patient 387 DNA. Multiple online tools were used, integrated into Alamut v2.2 software. Both changes reported in case 387 were analysed, since the missense change was located in the last exon of *TMEM67*. For both changes, a possible splice site effect was noted.

All the algorithms predicted that the junction between c.507-1 and c.507 could function as both a 5' and a 3' splice junction site. They also all predicted a reduction in confidence of the 3' acceptor splice site. This reduction varied from being fairly minimal (SpliceSiteFinder-like & HSF, both of which lack specificity for predicting putative splice junctions) to being a complete abolition (or effective abolition) of the 3' junction (NNsplice and NetGene).

Glycine at position 979 is conserved in vertebrate orthologs **Figure 3-2**. *D. melanogaster* and *C. elegans* orthologues also have non-polar, hydrophobic residues at this position (alanine and valine). ClustalX alignment (**Figure 3-2**) also shows leucine (non-polar, hydrophobic) in some species, as well as a couple of polar uncharged amino acids (glutamine and serine). There are no charged amino acids at this position for any species in this alignment, but the missense mutation is glycine to arginine (charged, basic). The Grantham distance (physicochemical difference) between glycine and arginine is large (-125). BLOSUM scores (45:-2, 62: -2, 80: -8) suggest a below average probability that this change has occurred by chance. Position 979 is just beyond the seventh transmembrane domain in the protein. Since this residue is localised in the last exon and close to the C-terminus, the same splice-site predictions as for the c.507-3 position were performed. The algorithm predicted the possibility of the change introducing a cryptic donor (5') splice junction between c.2931 and c.2932. For this splice site to cut out the end of exon 28, there would have to be a putative acceptor splice-site within the remaining portion of transcribed material. According to the NM_153704.5 transcript, *TMEM67* has quite a large 3' UTR and the closest "strong" option is after c.*128.

The c.2935G>A p.G979R change also has a possible pathogenic effect on the structure and/or function of the protein caused by amino acid missense change. *In silico* protein prediction tools leaned towards a potential pathogenic interpretation: MutationTaster ²¹⁹ gave a score of 0.99, which is predicted to be damaging based on the amino acid change and possible splicing change. Polyphen2 ²¹² (PSIC score 1.00, Sensitivity 0.00, Specificity 1.00), predicted this change was probably damaging. SIFT ²¹³ suggested the change might be tolerated (with a score of 0.11, sequence conservation of 3.34).

3.1.4.1.1.4 RNA/cDNA analyses

To establish the pathogenic potential of these possible splice-site changes, total RNA was extracted from foetal tissue. To investigate the c.507-3C>A change in intron 4, five primers sets were designed to span between exon 2 and 8 of *TMEM67*. Forward primers were localised in exon 2, 3, 4 and reverse primers in exon 7 and 8. Following RT-PCR of patient cDNA, all primer sets showed the wild-type sequence with no effect on splicing (**Figure 3-5**).

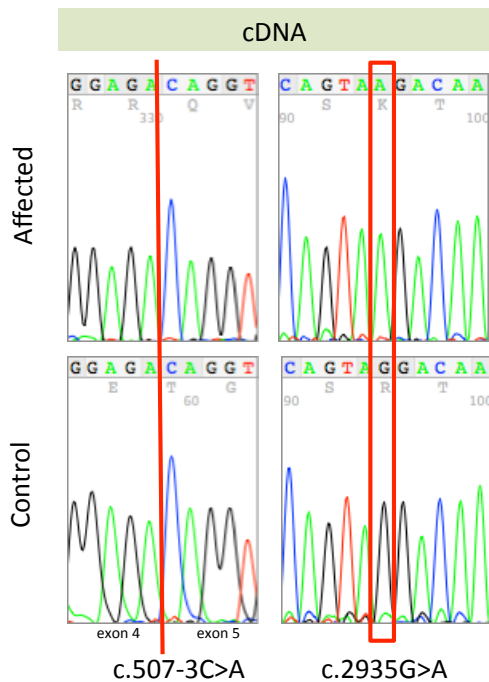


Figure 3-5. Electropherograms presenting sequencing results of cDNA of MKS affected patient 387. In red frame (highlighted) is change identified in the affected patient compared to a normal control individual at coding DNA (cDNA) level. The red line for c.507-3C>A indicates an exon/exon boundary with no changes observed in the affected patient sample. The cDNA for change c.2935G>A shows homozygous change in affected patient sample.

To investigate the c.2935G>A change, primers were designed to span between exon 27 and a region in the 3'UTR after c.*128. These analyses of patient cDNA showed that this change does not affect the splice-site but revealed the variant at position c.2935 to be in the homozygous (AA) state (**Figure 3-5**).

Homozygosity of c.2935G>A change may suggest that c.507-3C>A is essential for correct splicing, and that the RNA product transcribed from the paternal chromosome was lost through a nonsense-mediated decay (NMD) and was therefore not detected by cDNA sequencing. This mutation could either directly abolish the splice site, or that there would be a competition between the putative 5' and 3' sequences at this locus for the splicing machinery²²⁸. The c.2935G>A p.G979R change at the transmembrane boundary may affect the membrane topology or affect protein-protein interaction with cytoplasmic proteins.

3.1.4.1.2 Patient 178

Ultrasound scan of the pregnancy showed enlarged polycystic kidneys, occipital encephalocele and polydactyly, suggestive of MKS as a diagnosis. The pregnancy was terminated and a cord blood sample was collected for gDNA extraction. An additional skin biopsy was also collected. As the patient was of Pakistani ethnic origin, *TMEM67* gene sequencing was prioritised. Initial analysis revealed a single heterozygous change c.1615C>T p.R549C. As the patient originated from a consanguineous family, it was expected that the causative mutation would be homozygous. The c.1615C>T p.R549C mutation in the homozygous state was previously reported in a consanguineous Chinese family¹¹¹, is present at a very low allele frequency in the available variant databases, hence is very likely to be causal and pathogenic. Since the gDNA under test was from a cord blood sample, it was possible that maternal contamination could explain the heterozygous state of c.1615C>T p.R549C in patient 178. Fibroblasts originating from patient 178 were therefore grown, RNA was extracted and cDNA was synthesised. Primers spanning across the whole coding sequence (**Figure 3-6**) of *TMEM67* for patient 178 and control sample were used for amplification of *TMEM67* from cDNA. No differences in product sizes were observed between patient 178 and control samples. cDNA was subsequently sequenced, which confirmed the heterozygous c.1615C>T change, however no other mutations in the amplified PCR products were identified. As this sample did not show linkage to any other known MKS loci (**Table 3-2**), it should be prioritised for WES for gene discovery of new MKS genes.

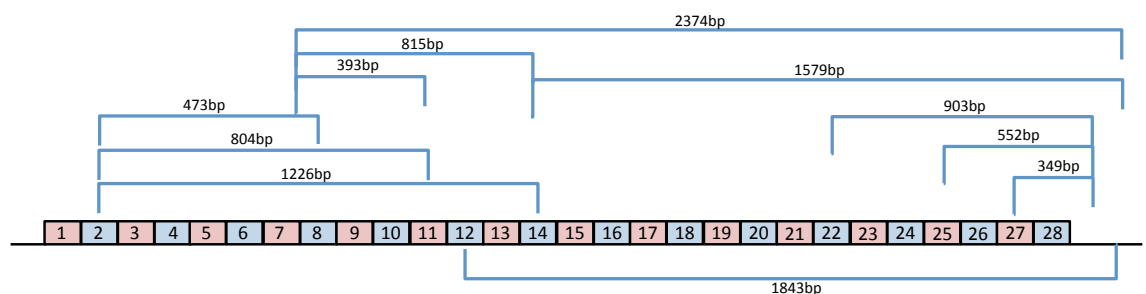


Figure 3-6. RT-PCR primers for amplification of *TMEM67* cDNA. Coloured blocks represent coding exons of *TMEM67*.

3.1.5 Discussion

As the first step in efforts to identify causative mutations in known MKS/JSRD genes, genotyping analyses were completed in cohort of MKS and JSRD patients. This method allows fast identification of linked loci by the analysis of genetic markers flanking genes of interest. There are four steps involved in the procedure:

PCR, gel electrophoresis, sample analysis on an ABI 3730xl genetic analyser and data analysis on appropriate software (GeneMapper). This strategy reduces the costs of sample analysis since it does not require a sequencing step, which is the main outlay in expense and time. Genotyping allows fast analysis of the patient DNA and prioritisation of the samples with linkage to the certain loci to be taken forward for sequencing analysis. Linkage was excluded in just over half of our cohort and those samples were not sequenced for the known MKS genes. This also allowed a reduction in the amount of patient DNA to be used, as only 10-20ng of DNA was required for the analysis of one microsatellite marker.

Linkage analysis was mainly performed for consanguineous patients, where homozygous markers flanking the gene would suggest linkage to the locus. Non-consanguineous patients in multiplex families with two or more affected individuals were also analysed, as haplotypes shared between these affected individuals could also suggest linkage. This method was robust and allowed prioritisation of many samples for direct Sanger sequencing and linkage exclusion in the others. However the latter could lead to false negative results as observed in the case of sample 264. This consanguineous patient had the *MKS1* marker D17S1606 in the heterozygous state, whilst D17S1290 was homozygous. This sample was not sequenced for changes in *MKS1* as the markers suggested no linkage. Conversely, this patient showed linkage to the *RPGRIP1L* locus but Sanger sequencing revealed no mutations. In the light of no causative mutations identified and putative linkage to other loci, this patient was prioritised for SNPchip genotyping analysis in efforts to identify a new MKS locus by homozygosity mapping. The whole genome genotyping data showed a large segment of homozygosity across the *MKS1* locus and the *MKS1* gene was screened revealing a pathogenic mutation. This highlighted the importance of careful data analysis and suggested that choosing a marker as close as possible to the gene, or even within the gene, is very important to prevent false negative findings.

Mutations were identified in 57% (38/67) of families that were recruited to the study (**Table 3-4**). Out of all families with identified mutations, 19/38 (50%) had changes in *TMEM67* (**Figure 3-7**), which highlights the prevalence of *TMEM67* mutations as a major cause of MKS. The second most commonly mutated genes were *CEP290* and *MKS1* (each 5/38; 13.1%), then *CC2D2A* and *RPGRIP1L* (each 3/38; 7.9%). *TMEM216*, *TMEM138* and *TMEM237* each had mutations in only one family (2.7% each), confirming that these were uncommon causes of the MKS phenotype.

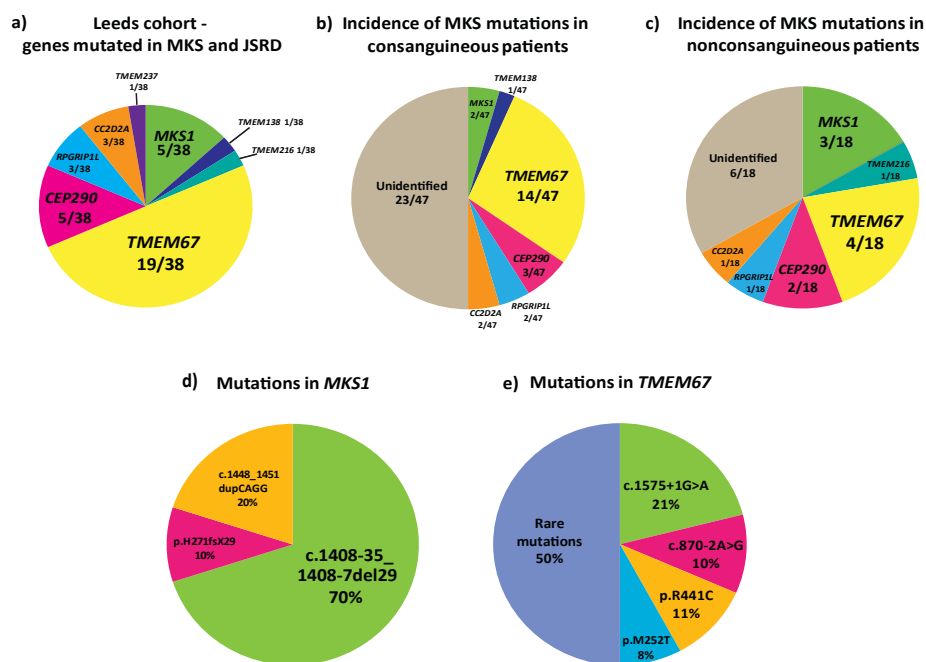


Figure 3-7. Pie charts summarising mutation analysis in MKS and MKS-like patients in the study cohort. a) frequency of genes mutated in MKS and MKS-like phenotype; b) frequency of MKS genes mutations in consanguineous patients; c) frequency of MKS genes mutations in non-consanguineous patients; d) common mutations in *MKS1* in Leeds cohort study; e) common mutations in *TMEM67* in Leeds cohort study. Image adapted from Szymanska et al. 2012.

Homozygous mutations predicted to be pathogenic were identified in 50% of consanguineous families (**Figure 3-7b**). Two families had mutations in *MKS1*. Patient 264 had a homozygous *MKS1* mutation c.1408-35_1408-6del30 (**Figure 3-7d**), which overlaps with the “Finn major” Finnish founder mutation (c.1408-35_1408-7del29)²²². Since family 264 is of Jordanian origin, and therefore has a different genetic background to northern European patients with the “Finn major” mutation, this finding suggests a mutation hot spot in this intronic region of the *MKS1* gene. Three different homozygous mutations in *CEP290* were identified in Pakistani families. Two of these were frameshift mutations and one was a nonsense mutation and all are predicted to cause NMD.

The majority of identified mutations were found in *TMEM67* (**Figure 3-7b**), comprising n=14/47 (30%) families. Two splice-site mutations and one missense mutation were identified multiple times and may therefore be founder mutations (**Figure 3-7e**). In patients of Pakistani origin two *TMEM67* splice-site mutations were identified, c.1546+1G>A and c.870-2A>G, which have previously been reported as common mutations⁹⁷. The homozygous missense mutation p.R441C was identified in two families (319 and 347), a mutation reported previously in the heterozygous state for patients with COACH syndrome²²⁹. A missense mutation affecting the same amino acid residue, p.R441L, has also been reported previously

in an MKS patient¹¹⁰. As families 319 and 347 have different ethnic origins (British and Pakistani, respectively), this emphasizes the mutability of arginine residues and their importance to the function of the protein since the neighbouring residue p.R440 is also mutated in MKS and MKS-associated ciliopathies^{112,230}.

Probable founder mutations in both *RPGRIP1L* and *CC2D2A* for families of Pakistani ethnic origin were identified. The *RPGRIP1L* nonsense mutation c.1945C>T p.R649* was observed in families 207 and 336 which are reported to be unrelated. The frame-shift mutation in *CC2D2A* c.3540delA p.R1180Sfs*6 occurred in the unrelated families 158 and 180, with polydactyly noted as an obligatory feature in all affected individuals.

Applied methodology was useful to prioritise the known (at the time of this study) MKS/JSRD genes for sequencing, but establishing common haplotypes was difficult as samples were run on different days, with different batches of polymer and sometimes even with different sequencing arrays causing slight shifts in the peaks. The exception to this was observed for *CC2D2A* for which patients 158 and 180 shared the haplotype 150/150-245/245 (for markers D4S1511 and D4S2960, respectively) and carry the same homozygous mutation c.3540delA (**Table 3-4**). Therefore consanguineous patients originating from Pakistan with this particular set of microsatellite markers could be prioritised for sequencing for this mutation, with a view to reducing the cost and time of the analyses.

Two-thirds (n=13/19) of non-consanguineous families had their causative mutations identified (**Figure 3-7c**), with the majority of mutations (n=10/13) in the compound heterozygous state. The majority of identified mutations were found in the *TMEM67* gene, with mutations in *MKS1* and *CEP290* the next frequent. In our cohort, the “Finn major” mutation was found in all *MKS1*-mutated patients, either in the homozygous state (families 106 and 74699 of British origin) or as a compound heterozygous mutation *in trans* with the frameshift mutation p.H271fs*29 (family 77172 of Finnish origin). Overall, the *MKS1* “Finn major” mutation was the most frequent (**Figure 3-7d**). The heterozygous missense mutation p.M252T accounted for 30% of identified alleles in *TMEM67* in non-consanguineous patients. The previously reported common Finnish *CC2D2A* mutation was absent in screened cohort¹⁰², even though the *MKS1* “Finn major” Finnish mutation seemed frequent. This suggests that the latter is more widespread throughout European populations, whereas the *CC2D2A* mutation is less common outside Scandinavia. In addition, six families were identified with a single heterozygous mutation in an MKS gene with a second pathogenic variant to be identified. *CC2D2A* p.E229del was probably a common variant, as it was detected as a single heterozygous variant in two

families of Pakistani origin. In sibpair 36-36A and patient 202, two changes in the same gene were detected but segregation analysis in parental samples showed that these were inherited *in cis* from the paternal line, therefore the pathogenic potential of these variants was unclear. Patient 178 has the single heterozygous missense mutation p.R549C in *TMEM67*. Any other potential pathogenic changes in any of the seven MKS genes that were screened for these patients were not detected.

The molecular basis of the phenotypic variability in MKS/JSRD may arise from oligogenic inheritance²³¹, where a third allele modifies the phenotypic effect of two recessive alleles. It is interesting to note that many ciliopathy and ciliary-related proteins interact and are reported to create functional modules that are localized to discrete structural regions of the cilium such as the TZ^{47,108,152}. The effect of modifier alleles may be to abrogate interactions between components of a functional module, which may disrupt protein complexes or signalling pathways giving rise to the ciliopathy phenotype. Four different heterozygous changes in six patients were identified in MKS/JSRD genes, in the absence of a second detectable pathogenic mutation in the same gene or any other mutations in remaining MKS genes. These heterozygous alleles could be potential modifier alleles, but the possibility of microdeletion or a second pathogenic mutation occurring deep within introns or regulatory elements of the same MKS/JSRD gene were not exhaustively excluded. These patients were also not excluded for the occurrence of biallelic mutations in all known MKS/JSRD genes, and the heterozygous variants could be incidental findings. Furthermore, the incomplete nature of some of the clinical details raises the possibility of differential diagnoses.

With the ever-increasing power and affordability of genetic sequencing technologies, there is now the clear opportunity for the further rapid and robust identification of mutations in patients referred for a defined condition. As a prerequisite, there remains a pressing clinical need for the dissemination of mutations identified on a research basis, and the establishment of databases that provide detailed clinical phenotypes and allelic series for specific genes.

3.2 Autozygosity mapping and candidate gene screening for new MKS/JSRD genes

The extensive genetic heterogeneity identified in the severe ciliopathies (**Table 1-1** and **1-2**) still does not explain the aetiology for as many as half of families with these phenotypes. This chapter describes the work to identify mutations in new genes causing MKS/JSRD.

In the first instance, DNA samples from multiplex (multiple affected offspring) consanguineous families with MKS/JSRD were analysed for shared autozygous haplotypes (section **1.2.2.7**) using microarray-based SNP genotyping. Autozygous regions were identified using established software programs AutoSNPa²³², IBDfinder (<http://autozygosity.org/>)²³³ and SnpViewer (<http://sourceforge.net/>).

Microsatellite genetic markers were also often used to exclude linkage to homozygous regions spanning across known MKS and/or JSRD loci, in order to prioritize known genes for conventional Sanger sequencing. For potentially novel loci, the best candidate genes were initially prioritised for Sanger sequencing based on the putative function related to the phenotype, similar expression patterns or homology to a known gene.

3.2.1 Identification of a putative new MKS locus on chromosome 12q24.11-24.13

All mapping in this chapter is based on GRCh37.p13.

Autozygosity mapping of six samples – 158, 227, 230, 261, 330 - identified a potential homozygous region of interest on chromosome 12. SNP chip genotyping (interval rs4475967, 108472383bp to rs7954961, 120286192bp) and genotyping of microsatellite genetic markers determined secure interval boundaries from D12S353 to D12S395 (**Figure 3-8**). There was no shared haplotype observed between the samples, additionally these samples contained additional homozygous regions at different loci (**Appendix 3, 4, 5**).

Mbp	marker	affected 158	affected 227	affected 230	affected 261	affected 330
85.68	D12S1719	224 224	222 232	224 234	232 232	234 234
88.2	D12S1710	268 274	274 278	278 278	274 274	274 274
95.07	D12S348	294 294	290 294	296 296	290 290	294 294
95.56	D12S1300	119 131	115 119	127 127	115 119	127 127
97.39	D12S1706	120 124	120 132	118 118	120 120	130 130
100.29	PAH	239 239	239 239	239 239	239 239	239 239
102.78	D12S78	174 190	182 182	188 188	194 186	184 184
106.55	D12S353	88 90	88 88	90 90	98 88	94 98
107.54	D12S84	214 214	198 198	214 214	216 216	214 214
108.3	D12S1583	224 224	224 224	234 234	244 244	238 238
108.52	D12S1645	216 216	216 216	238 238	212 212	216 216
108.55	D12S1328	268 268	268 268	268 268	268 268	270 270
110.18	D12S1332	204 204	210 210	206 206	200 200	200 200
110.8	D12S1333	226 226	228 228	240 240	236 236	240 240
111.45	D12S821	354 354	338 338	354 354	358 358	358 358
112.38	D12S811	316 316	332 332	344 344	324 324	354 358
114.54	D12S79	178 178	NSD NSD	162 162	162 162	170 120
116.63	D12S1720	96 96	114 114	98 98	96 96	104 104
117.19	D12S395	242 242	226 226	226 226	226 226	244 248
119.17	D12S1721	288 286	278 278	286 286	284 276	274 286
124.4	D12S342	234 234	222 234	228 228	224 234	224 224
124.94	D12S2078	282 270	270 274	260 260	264 270	274 274

Figure 3-8. Microsatellite genetic markers and mapping of a putative new MKS locus on chromosome 12. All genotyped samples are singleton consanguineous affected individuals with MKS. In pink the minimal homozygous region is highlighted, defined by markers genotyped in sample 330. Homozygous markers are indicated by bold frames. All samples, with exception of individual 261 who is of Jordanian origin, are from Pakistan.

Within the minimal interval 160 genes were present and two genes (*TMEM116* and *TMEM119*) were prioritised for screening based on topological homology to a known MKS/JSRD gene, *TMEM216*. These genes were subsequently sequenced in samples 158, 227, 230, 261 and 330 but no putative pathogenic mutations were identified. Based on the predicted protein function *ERP29*, *CCDC63*, *DYNLL1*, *TMEM233*, *RAB35*, *CORO1C* and *VPS29* were also sequenced in the same sample set. *ERP29* functions in the endoplasmic reticulum and is thought to process secretory proteins. Choosing this gene as a functional candidate was based on a prediction that the function of *TMEM67* is dependent on correct processing in the ER²³⁴. *CCDC63* contains coiled-coil domain commonly observed in multiple ciliary proteins. *DYNLL1* is a dynein light chain protein and dyneins are known as motile cilia components as well as for their function in microtubule association. *TMEM233* was chosen based on its similarity to the other ciliary transmembrane proteins. While other genes were chosen based on their involvement in endocytosis (*RAB35*, *VPS29*) and actin cytoskeleton remodelling (*RAB35*, *CORO1C*). There were no variants identified that could be pathogenic mutations in these candidate genes, with the exception of *CCDC63* in which a homozygous missense variant, p.K33E, was identified in sample 158. This change was predicted to be possibly damaging with a PolyPhen2 score of 0.925. However, in sample 158 a homozygous causative pathogenic change in *CC2D2A*, p.R1180Sfs*6 was identified in a parallel study, which was sufficient to explain the MKS phenotype in the patient. The *CCDC63* missense variant is therefore unlikely

to be the primary causative locus in this patient, and may instead be a benign polymorphism.

Other genes selected for screening based on their protein function were *TCTN1* and *TCTN2*. There was a strong evidence of *TCTN1* function in Shh signalling²³⁵ and its importance in embryonic development. *TCTN2* belongs to the same group of homologous tectonic proteins and, although its function was not known at the time of this study, it was also prioritised. The following samples were screened for mutations in *TCTN1*: 111, 112, 128, 168, 217, 218, 222, 242, 153, 157, 158, 227, 230 and 261 (sample 330 arrived after *TCTN1* screening); and *TCTN2* (this gene was localised just outside the original chromosome 12 locus, ~124Mbp, and only three samples showed linkage): 227, 230 and 330. Screened samples were used to define the chromosome 12 locus along with non-consanguineous samples with no causative mutation identified. Two heterozygous variants were identified in *TCTN1*: p.G466C in sample 128 and a frameshift change p.G567Afs*107 in sample 168. No changes were identified in *TCTN2*.

3.2.2 Other ciliopathy candidate genes

A group of collaborators with a research focus on ciliopathies was established by Prof. Colin Johnson. To support each other in identification of mutations in new candidate genes and unravelling the complex genetics of ciliopathies, genes proposed by collaborators were screened in MKS and JSRD cohorts on an *ad hoc* and collaborative basis.

3.2.2.1 CEP164

One of the proposed genes to screen was *CEP164* requested by Prof. Friedhelm Hildebrand (Harvard Medical School). Centrosomal protein 164 kDa was a good functional candidate since it is a component of the centrosome which is a matrix for ciliary axoneme nucleation. Samples from consanguineous patients with homozygous regions over the *CEP164* locus were sequenced (158, 175, 179, 269 and 352; homozygosity was established based on SNPchip 10k analysis for samples 158, 175, 179 and 269 [homozygous regions between rs1074480 and rs557940 were observed] and microsatellites analysis for sample 352 [D11S1986 and D11S1998]). One heterozygous change was identified in sample 175: c.3268T>C p.Y1090H, though this change was predicted to be benign by PolyPhen2. Results from this study were published in the following article Chaki et al.²³⁶.

3.2.2.2 TTC21B

Mutations in *TTC21B*, also known as *IFT139*, are a cause of sporadic nephronophthisis and syndromic Jeune asphyxiating thoracic dystrophy. The cohort of 18 MKS patients was screened for this gene in Prof F. Hildebrand's laboratory (University of Michigan), but no biallelic changes were identified. Multiple heterozygous changes were identified and confirmed by the author in stock DNA, but these were either *de novo*, had no second change identified, or were inherited in *cis* with another variant as presented in **Table 3-7**.

SAMPLE	STATUS	CHANGE
34	father	TTC21B c.1040A>G p.Y347C het, c.3121G>A p.D1041N het
35	mother	
36	affected	TTC21B c.1040A>G p.Y347C het, c.3121G>A p.D1041N het
36A	affected	TTC21B c.1040A>G p.Y347C het, c.3121G>A p.D1041N het
37	mother	TTC21Bc.2530A>G p.M844V het
38	father	
39*	affected	TTC21B c.2530A>G p.M844V het
63	mother	TTC21B c.1846C>T p.R616C het
64	father	
65	sister	TTC21B c.1846C>T p.R616C het
66F1	affected	TTC21B c.1846C>T p.R616C het
66F2	affected	TTC21B c.1846C>T p.R616C het
240	father	
241	mother	
242	affected	c.2588G>A p.R863Q het

Table 3-7. Sequencing results for TTC21B. Separate MKS families are separated with thick lines. Empty spaces indicate that there was no change identified in the sample. * - in this family had *CC2D2A* causative mutation identified in a parallel study. Abbreviation: ex - exon; het - heterozygous.

Screening of this gene was also carried out in collaboration with Prof. Nicolas Katsanis (Duke University Medical Centre, USA), and is described in the next subchapter.

3.2.2.3 IFT genes

The IFT machinery is well-described to function within ciliary axoneme, and some proteins that are components of the IFT complexes are already reported to be involved in ciliopathy phenotypes^{183,237}. Consequently, the remaining components of both the IFT-A and IFT-B complexes were sequenced by collaborator Prof. Nicolas Katsanis (Duke University Medical Centre, USA) (screened samples: 1, 2, 17, 23, 24, 33A, 36, 39, 42, 43, 51, 66F1, 66F2, 67, 70, 74, 75, 78, 102, 104, 105, 111, 112, 113, 114, 124, 128, 144, 145, 156F, 156M, 158, 162, 163, 166, 168, 169, 170, 173, 174, 175, 176, 178, 179, 180, 181, 182, 184, 185, 200, 201, 205, 206,

207, 211, 217, 221, 222, 227, 230, 231, 232, 233, 234, 235, 239, 242, 244, 247, 250, 251, 252, 255, 261, 262, 263, 265, 266, 269, 270, 276, 282, 283, 292, 296, 319, 324, 325, 326, 333, 356, 357, C14). Any detected variants were verified by the author in the stock DNA. No biallelic changes inherited in *trans* were observed, but as with the example of *TTC21B* (section 3.2.2.2), multiple heterozygous changes were detected (Table 3-8). The high frequency of those heterozygous changes may suggest that they act as modifier alleles on primary causal mutations of the ciliopathy phenotype, as suggested by a number of previous publications^{182,203}, but this hypothesis needs further investigation. Results from this investigation will be analysed in collaboration with Dr Erica Davies and Prof. Nicolas Katsanis (Duke University).

SampleName	Zygoty	Sanger	IFT	Gene	DNA var.	ref aa	aa	alt aa
17	het	variant inherited form mother	A	IFT144_WDR19	c.3565+4A>C		Intron	
36	het	variants in <i>cis</i> , inherited from father	A	IFT139_TTC21B		D	1041	N
36	het	variants in <i>cis</i> , inherited from father	A	IFT139_TTC21B		Y	347	C
36	het	variant inherited form father	A	IFT139_TTC21B		D	1041	N
36	het	variant inherited form father	A	IFT139_TTC21B		Y	347	C
42	hom	WT in the other affected, parents hets	B	IFT54_TRAF3IP1		E	132	D
51	het	variant inherited form father	B	IFT172		N	478	I
67	het	variant inherited form mother	A	IFT121_WDR35		Q	527	R
75	het	mother is het, father is WT	B	IFT88		D	544	V
102	het	variant inherited form mother	B	IFT88		M	383	K
102	het	variant inherited form mother	B	IFT88	c.C2265T		Exon	
115	het	variant inherited form father	A	IFT144_WDR19		D	523	N
115	het	variant inherited form father	A	IFT144_WDR19		D	523	N
176	het	no parental samples	B	IFT80		D	368	H
244	het	variant inherited form mother	B	IFT88		M	383	K
244	het	variant inherited form mother	B	IFT88	c.C2265T		Exon	
256	het	het in father, WT in mother	B	IFT80		E	500	G
257	het	variant inherited form father	B	IFT80		R	719	H
264	het	variant inherited form mother	B	IFT172		R	1134	L
264	het	variant inherited form mother	A	IFT139_TTC21B		D	242	N
264	het	variant inherited form mother	A	IFT144_WDR19		E	1003	G
264	het	variant inherited form mother	B	IFT46_C11ORF60		P	152	A
324	het	variant inherited form father	B	IFT88		M	383	K
324	het	variant inherited form father	B	IFT88	c.C2265T		Exon	
325	het	variant inherited form mother	B	IFT74	c.G727C	E	243	Q
66F1	het	variant inherited form mother	A	IFT139_TTC21B		R	616	C
66F2	het	variant inherited form mother	A	IFT139_TTC21B		R	616	C

Table 3-8. List of changes identified in IFT genes in Leeds MKS/JSRD cohort.

3.2.2.4 TMEM107

Phylogenetic analyses by Prof. Martijn Huynen (Radboud University Nijmegen Medical Centre) suggested that the hypothetical transmembrane protein TMEM107 was conserved in ciliates and had a probable functional role in ciliogenesis. This protein occurs in ciliated species but is not universal within it, suggesting an accessory role. It contains an X-box motif characteristic for ciliary genes. The protein is predicted to contain four transmembrane domains. The *schlei* mouse model has a missense mutation in *Tmem107*, and was shown to have defects in ciliogenesis and Shh signaling²³⁸. Thirty-five MKS/JSRD samples, compatible with linkage to *TMEM107* loci (markers: D17S1353 and D17S786), were sequenced but no pathogenic mutations were identified. Several other collaborators screened more

than 300 patients with diverse ciliopathy phenotypes but no pathogenic variants were identified. The EVS database shows that changes in *TMEM107* are very rare, and analysis of the UK10k Consortium database revealed a single heterozygous variant of unknown significance c.74T>A, p.L25* in ciliopathy patient UK10K_CIL5236555. No validated antibodies recognizing *TMEM107* protein were commercially available at the time of the study, therefore a construct containing the *TMEM107* ORF fused to a C-terminal GFP sequence was used to analyse its cellular localization. **Figure 3-9** shows IMCD3 cells over-expressing *TMEM107*-GFP (green) co-localising with the ciliary axoneme marker acetylated α -tubulin (arrowheads). Only cells over-expressing a moderate amount of *TMEM107*-GFP showed co-localisation to the cilium (**Figure 3-9**, white arrows), whilst those cells with high levels of overexpressed protein did not have such a localisation (**Figure 3-9**, yellow arrow).

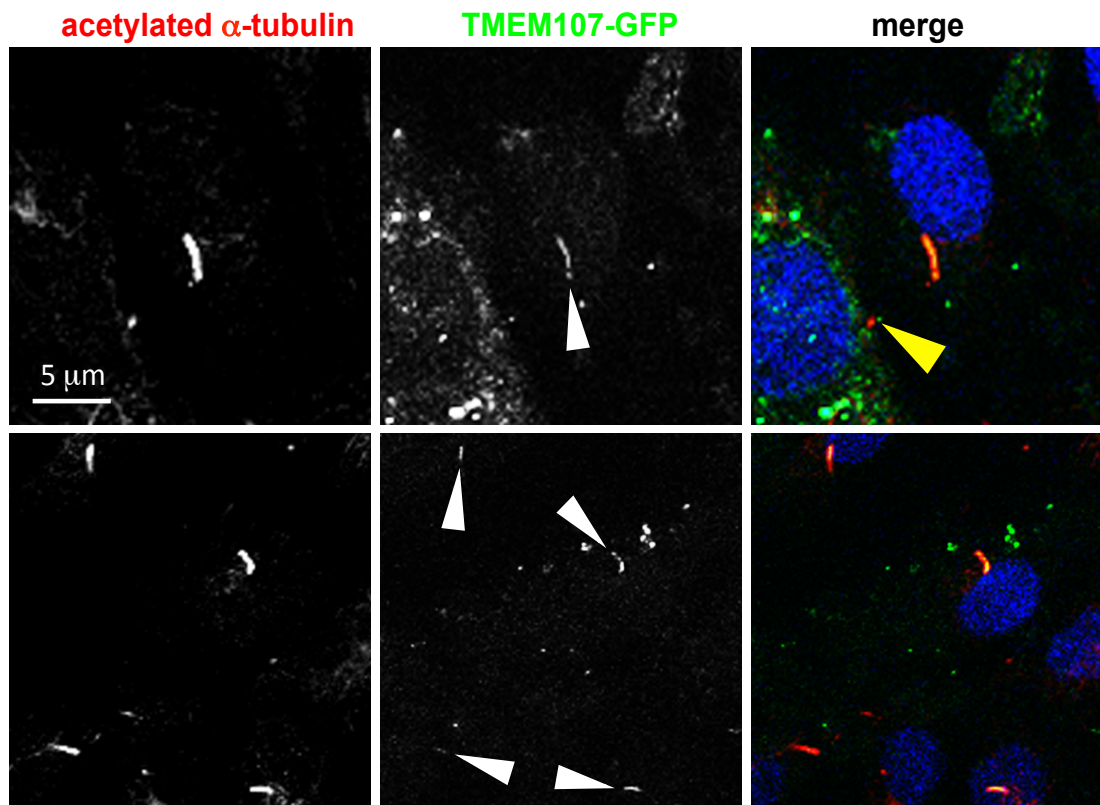


Figure 3-9. IMCD3 cells overexpressing *TMEM107*. White arrowheads point at ciliary localisation of *TMEM107*-GFP, while yellow one points at cilia without *TMEM107*-GFP co-localisation in cell with high level of exogenous protein level expression. Ciliary marker – acetylated α -tubulin was stained in red, nucleus marked by DAPI staining in blue. Scale bare = 5 μ m.

To further elucidate the function of *TMEM107* in cilia structure, an siRNA experiment was designed to transiently remove *Tmem107* transcript in IMCD3 cells. A pool of 4 si*Tmem107* duplexes (Dharmacon) versus scrambled siRNA as a

negative control were transfected into sub-confluent IMCD3 cells and incubated for 72 hours in serum-free media to induce ciliogenesis. Analysis of cilia number, length and cell number were done at the final concentration of siRNA of 50nM (**Figure 3-10**). A significant difference was observed in cilia length (t-test 0.02) and cilia number (t-test 0.0075) between control and *Tmem107* knockdown. After normalisation for cell number the difference in cilia number was not significant and may be explained by decreased cell number in the *Tmem107* knockdown. Multiple approaches to identify this protein's function were taken, and five separate groups have collaborated on this project with a publication in preparation.

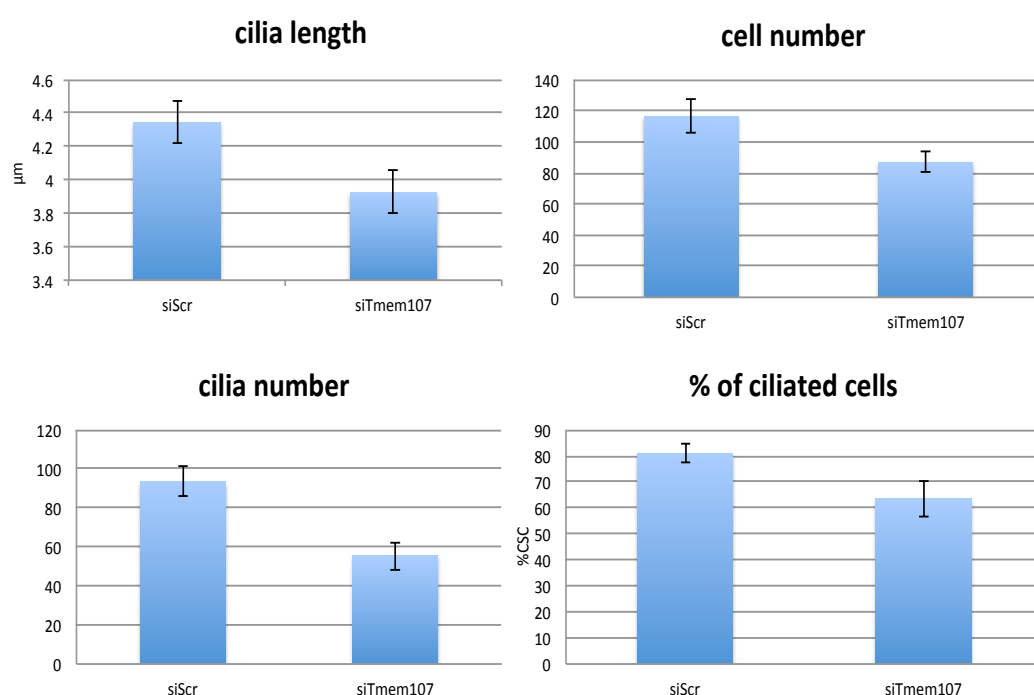


Figure 3-10. Analysis of IMCD3 cells after *Tmem107* knock down. Unpaired t-test with Welch's correction was calculated - cilia length:0.02; cell number:0.0604; cilia number:0.0075; % of ciliated cells:0.0708. Number of cells analysed: siScrambled=583, siTmem107=435. Number of replicates=5.

3.2.2.5 Other ciliopathy candidate genes

Other genes screened for possible mutations include: *PROM1* (one heterozygous change found in affected patient 106, variant c.55T>A p.S19A), and *CENPF* (samples 115 and 178 were screened, but no mutations were identified). The full cohort of 80 MKS and 16 JSRD patients was screened for variants in: *TCTN3*²³⁹ and *PDE6D*²⁴⁰ but all were mutation negative.

3.2.3 WES as a novel approach to identify causative mutations for MKS/JSRD

With the rapid progress of developing genetic technologies, positional cloning was succeeded by 'next – generation sequencing' which includes WES, WGS and transcriptome sequencing (RNAseq). In the case of consanguineous samples with SNP genotyping data and identified homozygous regions and prioritised candidate genes screened and no causative mutations identified, the next step would have been to screen of all genes in the homozygous regions. This would have involved hundreds of time-consuming PCR optimisations, costly Sanger sequencing and significant quantities of precious DNA. Alternatively, WES allows rapid cost-effective screening of candidate genes in autozygous regions. Homozygous regions shared only by affected individuals were checked for changes that were subsequently filtered based on zygosity, minor allele frequency (MAF), and presence in available on-line variant databases such as dbSNP, 1000 Genomes and EVS.

All in-house sample preparation and sequencing was done by Dr Clare Logan, data were subsequently analysed by Dr David Parry.

3.2.3.1 Illumina GAllx platform

3.2.3.1.1 MKS family 157

This family is of Pakistani origin with known consanguinity (**Figure 3-11**). Two affected siblings, presented with typical MKS phenotype, were negative for mutations in known MKS genes. DNA from one of the affected cases was analysed by microarray SNP genotyping. This revealed four regions of homozygosity: chromosome 4: 106657900 – 115232500 bp; chromosome 6: 162174100 – 170760000 bp; chromosome 11: 121476000 – 130075400 bp; chromosome 16: 10639860 – 20822750 bp (hg18). From these regions, the following functional candidates were prioritised for Sanger sequencing: *HYLS1* and *TMEM45B*. *HYLS1* encodes a gene mutated in hydrolethalus syndrome 1 MIM#236680²⁴¹, a condition with some features that are compatible with a ciliopathy, while *TMEM45B* showed similarity to other ciliary transmembrane proteins. Both genes were shown to have no pathogenic changes.

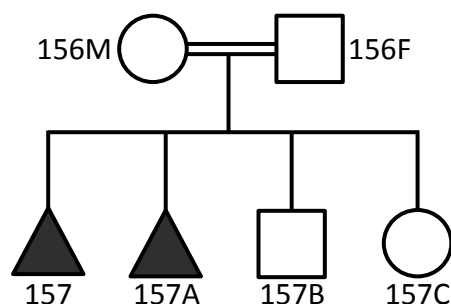


Figure 3-11. MKS family 157 pedigree. Consanguineous parents 156M (mother) and 156F (father) had two affected 157 and 157A and two unaffected 157B and 157C offsprings.

Whole exome capture and sequencing were done on DNA from one affected individual, 157 (section 2.2.8.2).

Gene	Class	RefSeqID -> Mutation	PolyPhen2	Chrom	Pos	SNP_ID
DLL1	missense	NM_005618 -> c.950G>A,p.R317Q	0.85	chr6	170594424	.
ROBO4	missense	NM_019055 -> c.1817G>A,p.R606H	0.004	chr11	124761326	.
ROBO4	missense	NM_019055 -> c.869C>T,p.P290L	0.968	chr11	124765520	.
DCPS	non-frameshift_insertion	NM_014026 -> c.215insTGGGGA,p.G77_E78insDG		chr11	126176478	.
ABCC6	missense	NM_001171 -> c.1171A>G,p.R391G	1	chr16	16295863	rs72653762

Table 3-9. List of variants detected in MKS family 157 sample after WES and variant filtering.

Table 3-9 shows a list of variants that were filtered based on their location within a homozygous region, their homozygous state and absence from the available on-line databases for common variants (dSNP, 1000Genomes, EVS). The variant detected in *ABCC6* was not taken into consideration as it was reported as a cause of an existing autosomal recessive condition, *pseudoxanthoma elasticum* (MIM#177850). This condition is characterised by *peau d'orange* changes in the retina, chroidal neovascularization, congestive heart failure and gastrointestinal hemorrhages, which do not resemble a ciliopathy phenotype. The variant in *DCPS* was an in-frame insertion, and this change was not predicted to have a significant pathogenic effect. One of the changes in *ROBO4* was predicted to be benign by PolyPhen2 and was not further investigated. *ROBO4* p.P290L on the other hand was predicted to be probably damaging and did segregate with the disease within the family. However, the function of the ROBO4 protein did not seem to fit with those of previously reported ciliary proteins as it is involved in angiogenesis and vascular patterning and acts as a receptor for SLIT proteins. The focus therefore shifted toward the change in *DLL1*. The p.R317Q missense variant did segregate with the disease within the family and was excluded in 96 Pakistani controls. The same change was observed in the heterozygous state in one of the UK10k Consortium ciliopathy patients (UK10K_CIL5002423). Another heterozygous change, c.1802_1804del a non-frameshift deletion, was identified in ciliopathy patient UK10K_CIL5236547, as well as in patient 287 from the Leeds MKS/JSRD

patient cohort. A further heterozygous change was detected in sample C18 (c.1622G>A p.G541E). However, the lack of other biallelic changes, also not reported in collaborator's cohorts, made the possibility of pathogenic mutation in *DLL1* as a cause of MKS to be unconvincing. Further functional work should help to support or reject hypothesis of mutation in this gene being a true cause of MKS phenotype.

3.2.3.2 Illumina MiSeq platform

3.2.3.2.1 MKS family 36A

This family (**Figure 3-12**) originates from the Gujarat state in India and although parents were reported to be unrelated, distant consanguinity was not excluded.

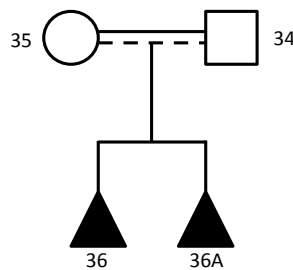


Figure 3-12. MKS family 36A pedigree. Parents 34 and 35 had two affected offspring, 36 and 36A. Dashed line between parents indicates possible consanguinity.

Patient DNA had whole exome library prepared and was sequenced on an Illumina MiSeq sequencer. Homozygous changes were identified as shown in **Table 3-10**. Out of these six, only *MACF1* was localised within a region of homozygosity. This variant segregated with the disease within the family and was not present in databases of common variants. Although *MACF1* was a good functional candidate with a putative function in cytoskeleton organisation, the large size of the gene, lack of microtubule-associated functions in known ciliary genes and the low read coverage across the exome precluded extensive further study of this gene.

Gene	Class	RefSeqID -> Mutation	Chrom	Pos	SNP_ID	Variant Percent Confidence	MKS36A Allele Depth
<i>MACF1</i>	missense	NM_012090 -> c.6797G>A,p.R2266Q	1	39835746	.	100	G=0/A=18
<i>LEFTY1</i>	nonsense	NM_020997 -> c.76C>T,p.Q26*	1	226076691	.	77.03851	G=0/A=1
<i>RASGRP3</i>	missense	NM_170672 -> c.1058A>T,p.N353I, NM_015376 -> c.1058A>T,p.N353I, NM_001139488 -> c.1058A>T,p.N353I	2	33752454	.	77.03851	A=0/T=1
<i>TGFBR2</i>	missense	NM_001024847 -> c.1616G>A,p.C539Y, NM_003242 -> c.1541G>A,p.C514Y	3	30732928	.	68.8111	G=0/A=1
<i>PRB3</i>	intronic/splice_consensus	NM_006249 -> c.725-3del	12	11420333	rs140904217	100	AG=0/A=10
<i>ALOX5AP</i>	splicing	NM_001204406 -> c.116+1insTA	13	31287979	.	100	G=0/GTA=11

Table 3-10. Summary of homozygous changes identified in WES of 36A sample.

3.2.3.2.2 MKS family 66F1

This consanguineous family originates from Pakistan (**Figure 3-13**).

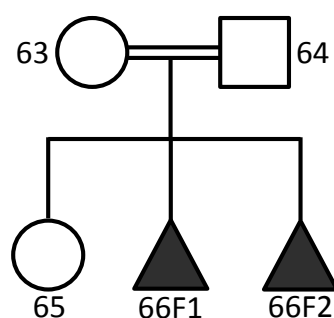


Figure 3-13. MKS family 66F1 pedigree. Consanguineous parents 63 and 64 had two affected offsprings 66F1 and 66F2 and one unaffected child 65.

Parents of the affected cases are first cousins and only homozygous changes in homozygous regions were taken into consideration – **Table 3-11**. All of these changes were confirmed by Sanger sequencing and segregation was checked within the family. Only the change in *TXNDC15* segregated with the phenotype.

Gene	Class	RefSeqID -> Mutation	Chrom	Pos	SNP_ID	PolyPhen2 score
<i>TXNDC15</i>	frameshift_insertion	NM_024715 -> c.955insT,p.S321Kfs*15	5	134235247	.	
<i>EYS</i>	intronic/splice_consensus	NM_001142800 -> c.6078-4_-3del	6	65016977	rs66531247	
<i>CTTNBP2</i>	missense	NM_033427 -> c.2750C>G,p.A917G	7	117417593	.	1
<i>PHGR1</i>	missense	NM_001145643 -> c.44G>A,p.G15D	15	40648299	.	0.998

Table 3-11. WES results for 66F1 family.

This change was excluded from databases of common variants and 96 Pakistani controls. Subsequently, a panel of ninety MKS/JSRD patients was screened for additional mutations in this gene (**Table 3-12**).

sample	exon	cDNA	protein	status	confirmed	segregate	controls	EVS (~13k)	PolyPhen2
227	1	15C>A	A5A	het	not done	not done	not done	not present	
377	2	C>T	S110S	homo	not done	not done	not done	not present	
206	3	703C>T	R235W	het	yes	no family members	not done	not present	probably damaging 1.0
378	5	1043G>A	R348Q	het	yes	no family members	not done	A=1/G=13005 rs149940297	probably damaging 1.0
66F1	5	955insT	S321Kfs*15	homo	yes	yes	yes (excluded)	yes (excluded)	
11	5	955insT	S321Kfs*15	het	yes	not passed to affected 13	yes (excluded)	yes (excluded)	

Table 3-12. Sequencing results of *TXNDC15* in MKS/JSRD panel of patients.

No other biallelic changes were identified, although in Shaheen et al, 2013 the authors report a non-frameshift variant in *TXNDC15* for an Arabic MKS patient. This change was not chosen by the authors as the causative mutation, since a missense variant in *EXOC4* was judged to be a better candidate for the cause of the MKS phenotype. No further functional work was performed for this gene.

3.2.3.3 Illumina HiSeq2500 platform

3.2.3.3.1 MKS family 17

This family originated from Pakistan and the parents are first cousins (**Figure 3-14**).

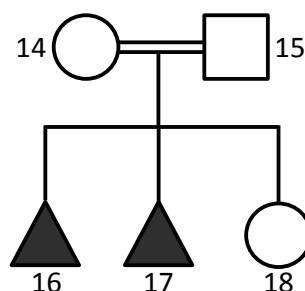


Figure 3-14. MKS family 17 pedigree. Consanguineous parents 14 and 15 had two affected offsprings 16 and 17 and one unaffected child 18.

DNA for the affected individual was sequenced on an Illumina HiSeq2500 sequencer. Only homozygous changes were prioritised for further analysis (**Table 3-13**).

Gene	Class	RefSeqID -> Mutation	Chrom	Pos	SNP_ID	MKS17 Allele Depth	PolyPhen2
<i>EXOC3L4</i>	nonsense	NM_001077594 -> c.232C>T,p.R78*	14	103566788	rs189674968	C=0/T=111	-
<i>AHNAK2</i>	missense	NM_138420 -> c.6122C>T,p.P2041L	14	105415666	rs150446570	G=0/A=2	NM_138420,P>L,probably damaging(1.0)
<i>AHNAK2</i>	missense	NM_138420 -> c.5526C>G,p.I1842M	14	105416262	.	G=0/C=6	NM_138420,I>M,benign(0.268)
<i>NUDT14</i>	missense	NM_177533 -> c.10A>G,p.I4V	14	105647537	.	T=0/C=16	NM_177533,I>V,benign(0.0)

Table 3-13. WES results for sample 17.

Out of these changes *EXOC3L4* was the best functional candidate gene, as the encoded protein is involved in exocytosis, process previously associated with ciliary function²⁴². The variant segregated with the disease within the family and

was excluded in databases of common variants and 96 Pakistani controls. Only two additional heterozygous changes were identified in the MKS/JSRD cohort: c.893A>T, p.Q298L in family 157 and c.1937G>A, p.R646Q in sample 168. Functional analysis of the EXOC3L4 protein will be performed by Dr David Parry in a zebrafish model.

3.2.3.3.2 MKS family 325

This family is of Pakistani origin with known consanguinity (**Figure 3-15**).

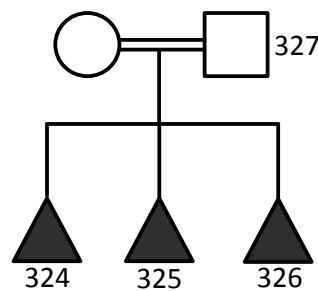


Figure 3-15. MKS family 325 pedigree. Consanguineous parent 327 (mother's sample was unavailable) had three affected offsprings 324, 325 and 326.

DNA samples for the family included three affected siblings and paternal DNA. SNP chip analysis of pooled and unpooled samples revealed no shared regions of homozygosity between the affected cases. Patient 325 was sequenced on an Illumina HiSeq2500 platform, but filtering for homozygous variants revealed a long list of possible variants – **Table 3-14**. Out of these, only *SPTBN4* was located in a homozygous region, but was excluded as a causal mutation following segregation analysis since it was not observed in other family members.

Gene	Class	RefSeqID -> Mutation	Chrom	Pos	SNP_ID	MKS325 Allele Depth	PolyPhen2
<i>ST3GAL5</i>	splicing	NM_003896 -> c.82+1G>T	2	86115946	.	C=0/A=2	-
<i>MFSD9</i>	missense	NM_032718 -> c.368A>G,p.Y123C	2	103343363	.	T=1/C=4	NM_032718,Y>C,probably damaging(1.0)
<i>RYK</i>	frameshift_insertion	NM_001005861 -> c.10insG,p.R4Afs*18, NM_002958 -> c.10insG,p.R4Afs*18	3	133969487	.	G=0/GC=5	
<i>FAM170A</i>	missense	NM_182761 -> c.755G>T,p.R252I, NM_001163991 -> c.614G>T,p.R205I	5	118970198	.	G=0/T=84	BC128243,R>I,probably damaging(1.0)/NM_001163991,R>I,probably damaging(0.999)
<i>HLA-B</i>	missense	NM_005514 -> c.356T>G,p.L119R	6	31324207	rs146092816	A=0/C=11	NM_005514,L>R,probably damaging(0.996)/AK310586,L>R,p probably damaging(0.998)
<i>MUC12</i>	missense	NM_001164462 -> c.5359A>C,p.S1787R	7	100639203	.	A=0/C=9	-
<i>LOC642236</i>	intronic/splice_consequence	NR_033907 -> 3'UTR+3A>G	9	68433462	rs111392331	T=0/C=2	-
<i>P2RY11,PPAN-P2RY11</i>	missense,non-coding,missense	NM_002566 -> c.814G>A,p.V272M, NM_001040664 -> c.2074G>A,p.V692M, NM_001198690 -> 3'UTR	19	10225103	rs147142449	G=1/A=161	NM_002566,V>M,probably damaging(0.999)
<i>KRTDAP</i>	missense	NM_207392 -> c.185A>T,p.E62V, NM_001244847 -> c.143A>T,p.E48V	19	35979371	.	T=0/A=26	NM_207392,E>V,possibly damaging(0.665)
<i>SPTBN4</i>	missense	NM_020971 -> c.2425C>A,p.R809S	19	41019121	.	C=0/A=80	AF324064,R>S,possibly damaging(0.924)
<i>CEACAM21</i>	missense	NM_033543 -> c.70C>A,p.L24I, NM_001098506 -> c.70C>A,p.L24I	19	42083557	.	C=0/A=21	NM_001098506,L>I,probably damaging(0.998)/NM_033543,L>I, probably damaging(0.997)
<i>ASMTL</i>	missense	NM_004192 -> c.1372G>A,p.V458M, NM_001173473 -> c.1198G>A,p.V400M, NM_001173474 -> c.1324G>A,p.V442M	X	1537881	rs151271783	C=0/T=2	NM_004192,V>M,probably damaging(0.973)/NM_001173474,V>M,probably damaging(0.985)/AK297805,V>M, probably damaging(0.973)
<i>ABCD1</i>	missense	NM_000033 -> c.1748T>A,p.V583E	X	153006141	rs79383557	T=0/A=2	NM_000033,V>E,probably damaging(1.0)

Table 3-14. WES results for family 325. All homozygous changes predicted to be pathogenic are listed.

The sequencing results in this family were inconclusive and no good functional candidates were identified. Ambiguity in sample labelling at the time of ascertainment also precluded further analysis of this family.

3.2.3.3.3 MKS family 144

This consanguineous family is of Moroccan origin, and was diagnosed with MKS (**Figure 3-16**).

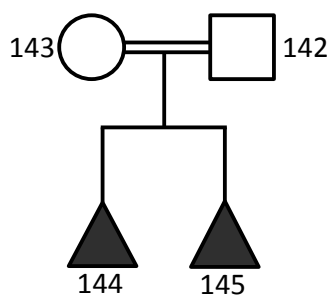


Figure 3-16. MKS family 144 pedigree. Consanguineous parents 142 and 143 had two affected offsprings 144 and 145.

WES was done on an Illumina HiSeq2500 sequencer, and homozygous changes in homozygous regions were identified (**Table 3-15**).

Gene	Class	RefSeqID -> Mutation	Chrom	Pos	SNP_ID	MKS144 Allele Depth	PolyPhen2
<i>GBP4</i>	missense	NM_052941 -> c.1090C>G,p.L364V	1	89655828	.	G=0/C=65	NM_052941,L>V,possibly damaging(0.529)
<i>ITIH3</i>	missense	NM_002217 -> c.1828G>A,p.E610K	3	52837989	rs190544531	G=0/A=85	NM_002217,E>K,possibly damaging(0.9)
<i>BBS12</i>	nonsense	NM_001178007 -> c.2023C>T,p.R675*, NM_152618 -> c.2023C>T,p.R675*	4	123665070	.	C=0/T=67	-
<i>BTBD11</i>	intronic,missense	NM_001017523 -> c.61G>A,p.G21R, NM_001018072 -> c.1491-185G>A	12	107974744	rs147063979	G=0/A=84	NM_001017523,G>R,probably damaging(1.0)

Table 3-15. WES results for family 144.

Amongst the homozygous variants, a nonsense variant was identified in the *BBS12* gene. It remains uncertain if MKS and BBS are allelic, and this result could have been a useful extension of the mutation spectrum in this gene, as well as definitive proof of allelism between these different ciliopathies. However, further clinical information on the patient phenotype confirmed a diagnosis of BBS with antenatal presentation rather than MKS.

This example shows how important a detailed clinical phenotype is for genetic research studies, and the possibility of differential diagnosis between ciliopathies.

3.2.3.3.4 MKS family 351

This consanguineous family from Southern India was referred for genetic analysis with a diagnosis of MKS (**Figure 3-17**).

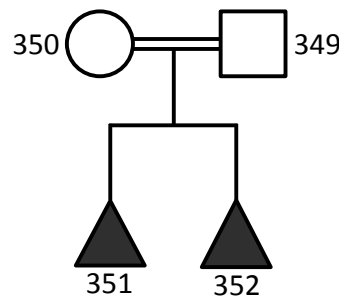


Figure 3-17. MKS family 351 pedigree. Consanguineous parents 349 and 350 had two affected offsprings 351 and 352.

WES on an Illumina HiSeq2500 sequencer was performed and only homozygous changes in regions of homozygosity shared by affected siblings were taken forward for further analysis (**Table 3-16**).

Gene	Class	RefSeqID -> Mutation	Chrom	Pos	SNP_ID	MKS351 Allele Depth	PolyPhen2
<i>FAM3D</i>	missense	NM_138805 -> c.47T>C,p.I16T	3	58639475	.	A=0/G=27	uc003dkq.2, >T, benign(0.001)
<i>LRCH4</i>	missense	NM_002319 -> c.853G>C,p.A285P	7	100175877	rs149693391	C=1/G=112	uc003uvj.2, A>P, benign(0.049)
<i>ZAN</i>	missense	NM_173059 -> c.6053C>T,p.P2018L, NM_003386 -> c.6053C>T,p.P2018L	7	100373317	.	C=1/T=121	-
<i>EMID2</i>	missense	NM_133457 -> c.257C>T,p.P86L	7	101063356	rs79106047	C=0/T=53	uc003uyo.1, P>L, probably damaging(1.0)/uc010lhy.1, P>L, probably damaging(1.0)
<i>ALG9</i>	nonsense	NM_001077690 -> c.764C>A,p.S255*, NM_001077691 -> c.251C>A,p.S84*, NM_001077692 -> c.251C>A,p.S84*, NM_024740 -> c.764C>A,p.S255*	11	111724397	rs17113312	G=0/T=42	-
<i>DIXDC1</i>	missense	NM_001037954 -> c.14T>G,p.L5R	11	111808237	.	T=0/G=126	uc001pmk.2, L>R, possibly damaging(0.571)/uc001pml.2, L>R, probably damaging(0.998)

Table 3-16. WES results for family 351.

Initially, as the best functional candidate, *DIXDC1* was prioritised for sequencing in MKS/JSRD cohort as the encoded protein was shown to localise at the centrosome²⁴³ and be involved in Wnt signalling²⁴⁴. Sequencing identified no other variants. Further investigation of the patient phenotype showed that affected individuals did not present with a 'typical' MKS phenotype. The patient phenotype included cystic dysplastic kidneys, skeletal anomalies, pre-axial limb malformation, cardiac defects, ductal plate malformation, pancreatic ductal ectasia, cerebral ventricular dilatation, pulmonary isomerism, hypertelorism, oligohydramnios with limb contractures (Potter sequence), limb shortening, hypoplastic vertebrae, pelvis, gracile ribs and acral anomalies. This phenotype, although unusually severe, is consistent with a congenital disorder of glycosylation type 1L, caused by mutations in the *ALG9* gene (MIM#608776). Segregation of the change with the phenotype was confirmed in the family and the change was excluded from databases of common variants. The MKS/JSRD cohort was screened for mutations in this gene but no other changes were identified, enabling a differential diagnosis of CDG type 1L to be excluded in the remaining MKS/JSRD patients.

3.2.3.4 CSPP1

A homozygous mutation in *CSPP1* (NM_024790.6:c.2244_2247delAAGA, p.E750Lfs*7) was identified in consanguineous Canadian Hutterite family with JSRD by a collaborator, Prof. Micheil Innes (Medical Genetics University of Calgary, Canada), using an autozygosity mapping and WES strategy. Two patients from Leeds MKS/JSRD cohort that were compatible with linkage to this locus (microsatellites tested: D8S1785 and D8S1840) were sequenced but this analysis revealed no variants that could be interpreted as pathogenic mutations. *CSPP1* encodes centrosome and spindle pole associated protein 1 and was therefore an excellent candidate to cause a ciliopathy phenotype. However, the absence of other families with mutations in *CSPP1* prevented further progress on this project until a second collaborator, Prof. Fowzan Alkuraya (Alfaisal University, Saudi Arabia), reported a second family with a homozygous *CSPP1* mutation. Genotyping analysis of the *CSPP1* locus were done for the remaining MKS/JSRD cohort and patients

compatible with linkage were Sanger sequenced. No variants were identified that could be interpreted as pathogenic mutations. Furthermore, 96 samples with MKS/JSRD were shared with Prof. Nicolas Katsanis for targeted resequencing of the ciliome. Targeted capture was performed for over 700 genes, from the CiliaProteome v3²⁴⁵ or known to cause a ciliopathy or be involved in ciliary function. During the analysis, three variants of unknown significance in MKS/JSRD cohort were detected (**Table 3-17**), but these are unlikely to be disease-causing. All those changes were Sanger sequenced by the author in the stock DNA and all were detected to be in heterozygous. Sample 166 already had a causative mutation detected in *RPGRIP1L*, 175 inherited the change from mother and there is no DNA from an affected child for parent 356. Sequencing of the whole *CSPP1* gene in the other parental sample to 356 did not reveal any changes. These results were incorporated into a subsequent publication¹⁰⁸.

Sample		Other Ciliopathy Mutations	Gene	annotation	exon	rsnumber	ESP6500	MAF	PolyPhen
166	Affected	RPGRIP1L [c.1829A>C p.H610P][c.721_724delAATG p.N241fsX25]	CSPP1	c.1376 C>G p.Ser459Cys hom	11	rs146431326	G=54/C=12018	0.004473	Probably damaging with a score of 0.989 (sensitivity: 0.72 ; specificity: 0.97)
175	Affected		CSPP1	c.2219 G>G/A p.Arg740His het	17		-	0	Probably damaging with a score of 1.000 (sensitivity: 0.00 ; specificity: 1.00)
356	Parent		CSPP1	c.2966 G>G/A p.Arg989Gln het	24		A=1/G=11827	0.000085	Probably damaging with a score of 0.999 (sensitivity: 0.14 ; specificity: 0.99)

Table 3-17. CSPP1 changes identified in CiliaProteome v3 targeted capture sequencing.

3.2.4 Identification of mutations in TMEM237 causing Joubert syndrome related disorder

3.2.4.1 Genotyping of MKS/JSRD patient DNA

Based on a personal communication (Prof. Kym Boycott, University of Ottawa) a novel locus for JSRD was identified on chromosome 2q33.1. Informative genetic markers (D2S2309 and D2S1384) were chosen that spanned approximately 2Mb across this locus. DNA samples from 18 MKS/JSRD patients, already determined to be mutation negative for all other known MKS genes, were chosen to be genotyped (**Figure 3-18**).

	OFD1	16	17	66F1	66F2	102	157	158	163
D2S2309	190 196	196 196	196 196	192 192	196 196	190 196	190 192	192 192	192 196
D2S1384	134 148	148 148	142 148	152 152	142 156	144 152	138 148	148 152	148 152
	175	218	230	231	276	277	42	43	261
D2S2309	192 196	192 196	196 196	196 196	192 196	196 196	196 206	192 192	196 196
D2S1384	148 152	152 152	148 164	144 152	138 142	143 156	138 148	144 144	147 148

Figure 3-18. Chromosome 2 genotyping results from 18 MKS/JSRD DNA samples. Non-consanguineous samples are coloured yellow while consanguineous samples are represented in red. Genotypes homozygous for both markers are highlighted in green.

DNA samples 16/17, 66F1/66F2 and 42/43 are from pairs of affected siblings, but none of them shared haplotypes for this region. The only sample that was compatible with linkage to chromosome 2q33.1 was from the singleton patient 261, of a consanguineous Jordanian origin.

3.2.4.2 Sequencing

Next generation sequencing was performed by collaborators in Prof. Kym Boycott's group in Canada. As a result, mutations were identified in the previously uncharacterised gene *TMEM237* (also known as *ALS2CR4*). Ninety MKS/JSRD patients from the University of Leeds patients cohort were then subsequently Sanger sequenced for mutations in *TMEM237* by the author. One additional family with a mutation in *TMEM237* was identified. In the consanguineous Jordanian patient 261 homozygous across the *TMEM237* locus, a homozygous *TMEM237* frame-shift mutation in exon 12 (c.1066dupC) (**Figure 3-19**) was identified. The frame-shift mutation is predicted to result in a truncated protein (p.Q356Pfs*23) and was absent in one hundred and five normal Jordanian control individuals that were Sanger sequenced for the variant.

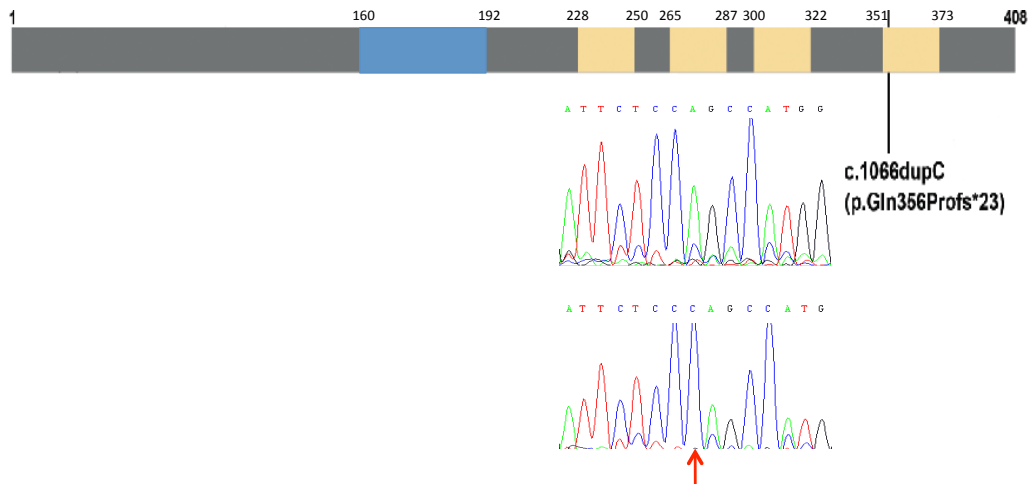


Figure 3-19. *TMEM237* genomic organization with domain structure of the *TMEM237* transmembrane protein. Yellow blocks indicate four predicted transmembrane domains in the *TMEM237* protein, blue tetratricopeptide domain (predicted using TMHMM online tool: <http://www.cbs.dtu.dk/services/TMHMM-2.0/>). Both *N*- and *C*-terminal domains are predicted to be intracellular. The black line indicates exon 12 with the homozygous c.1066dupC frameshift mutation in patient 261. The red arrow shows the identified mutation on an electrophoretogram (lower panel), compared to a normal control (upper panel). Image adapted from Huang et al. 2011.

TMEM237 spans 23 kb on chromosome 2q33 and contains 14 exons. Two alternative *TMEM237* transcripts (1 [NM_001044385.1] and 2 [NM_152388.2]), translating into two protein isoforms, a [NP_001037850.1] and b [NP_689601.2], have been proposed in humans. Each transcript utilizes one of the two alternative exons, exon 1 or exon 2. These are spliced in a mutually exclusive manner. All positional information in this thesis refers to transcript 1, and isoform a.

3.2.4.3 Patient phenotype

The mutation in *TMEM237* was identified in patient 261 who was initially reported to have MKS. Further detailed information obtained from their clinician though suggested JSRD with described occipital meningocele, cortical visual impairment, mild dilatation of pelvic ureter, sloping forehead, low set ears, developmental delay with no sitting and no speech.

3.2.4.4 Subcellular localisation of *TMEM237*

To elucidate the potential roles for *TMEM237*, the subcellular localization of the encoded protein was first determined. Two affinity-purified polyclonal antibodies raised against mouse *Tmem237* amino acids 76-88 (“FLJ-FM”) and 390-403 (“FLJ-LG”) (kindly provided by Dr Cheryl Craft [University of California Santa Cruz]) in the *N*- and *C*-terminal domains, respectively, were used (**Figure 3-20**)²⁴⁶. These

antibodies were used for immunocytochemical staining of a polarized monolayer of ciliated mouse inner medullary collecting duct (IMCD3) cells (**Figure 3-20**). Discrete signals at the proximal region of primary cilia were observed, consistent with localization to the TZ.

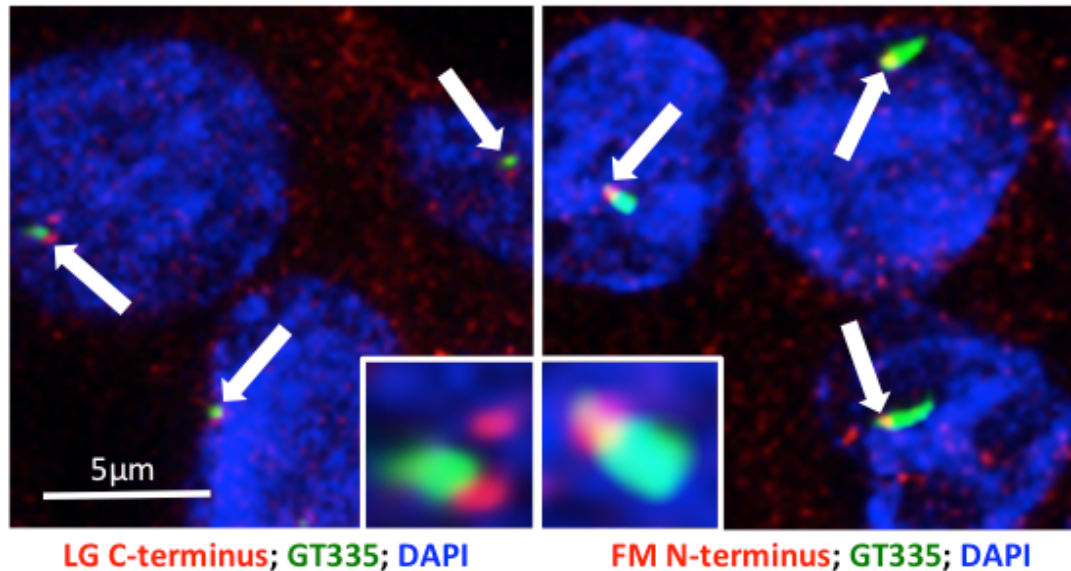


Figure 3-20. Subcellular localisation of TMEM237 in IMCD3s. On the left hand side cells were stained with anti-C-terminus, on the right with anti-N-terminus TMEM237 antibodies (red). White arrows indicate TZ localisation of TMEM237. For both assays polyglutamylated tubulin (GT335 in green) was used as a cilia marker. DAPI (blue) was used as a nucleus marker. Image adapted from Huang et al. 2011.

3.2.4.5 TMEM237 function in ciliogenesis

Fibroblasts derived from a patient with a homozygous nonsense mutation (p.R18*, line JSRD2) in *TMEM237* were obtained from collaborator Prof. Kym Boycott. As a control neonatal human dermal fibroblasts (HDF) obtained from Dr Jacqueline Bond (University of Leeds) were used. Quantitative real-time (qRT) PCR was used to measure *TMEM237* transcript expression since the existing mouse antibodies were not specific to human TMEM237. *TMEM237* transcript levels in the JSRD2 line were reduced by about 97% (**Figure 3-21**).

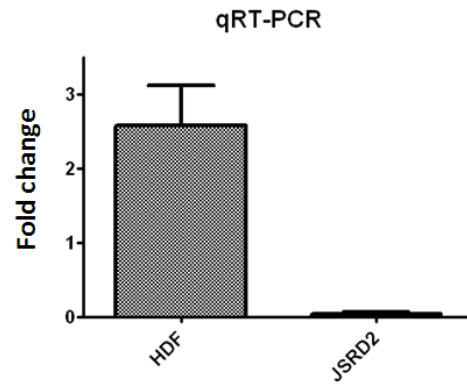


Figure 3-21. qRT-PCR quantification of TMEM237 levels between control (HDF) and JSRD2 RNA levels. Error bars represent standard error of the mean (n=3).

A failure in ciliogenesis was observed in patient fibroblasts, following 48h serum starvation (**Figure 3-22**) compared to control fibroblasts as seen previously in cells deficient for TMEM216 and TMEM67^{99,205}.

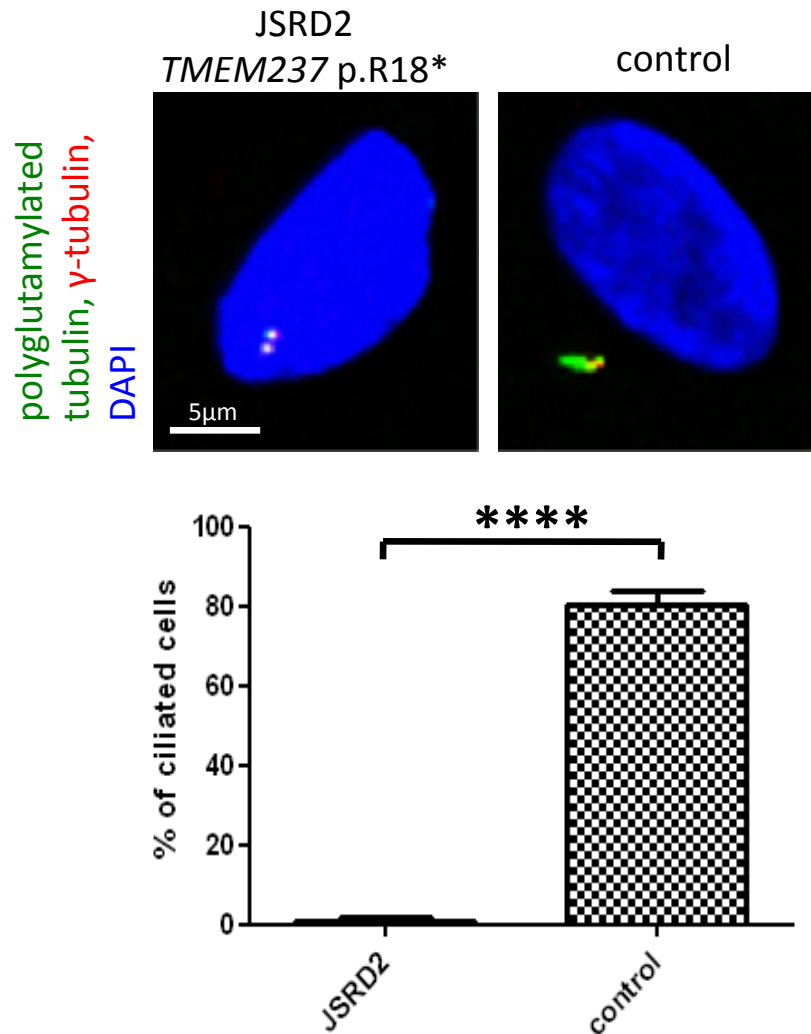


Figure 3-22. Disruption of ciliogenesis in JSRD2 patient fibroblasts with the homozygous *TMEM237* p.R18* nonsense mutation. Images show fibroblasts stained with polyglutamylated tubulin (green), γ -tubulin (red) and nuclear marker DAPI (blue); scale bar=5 μ m. In JSRD2, a lack of cilia was observed, compared to control fibroblasts with cilia (defined as >1 μ m in length). The bar graph shows the comparison of % of cells with cilia (n=300), demonstrating a statistically significant loss of cilia in JSRD2 (**** represents $p < 0.0001$, Student's t-test, paired two-tailed). Error bars represents standard error of the mean. Figure adapted from Huang et al. 2011.

These data were quantified by analysing the percentage of cells with evident cilia (defined as > 1 μ m length) versus those without cilia (defined as < 1 μ m length), with 300 cells analysed for each condition. For independent confirmation of these results, ciliogenesis defects were tested in IMCD3 cells with transient si*Tmem237* knockdown. The expected 45kDa protein was recognized by immunoblotting with the "FLJ-LG" antibody²⁴⁶ in cells transfected with siScr (**Figure 3-23a**, first lane). The staining was lost following si*Tmem237* knockdown (**Figure 3-23a**). Consistent with the results obtained from patient fibroblasts, transfection with two separate

si*Tmem237* duplexes (#1 and #2) impaired ciliogenesis in polarized cells (**Figure 3-23b**).

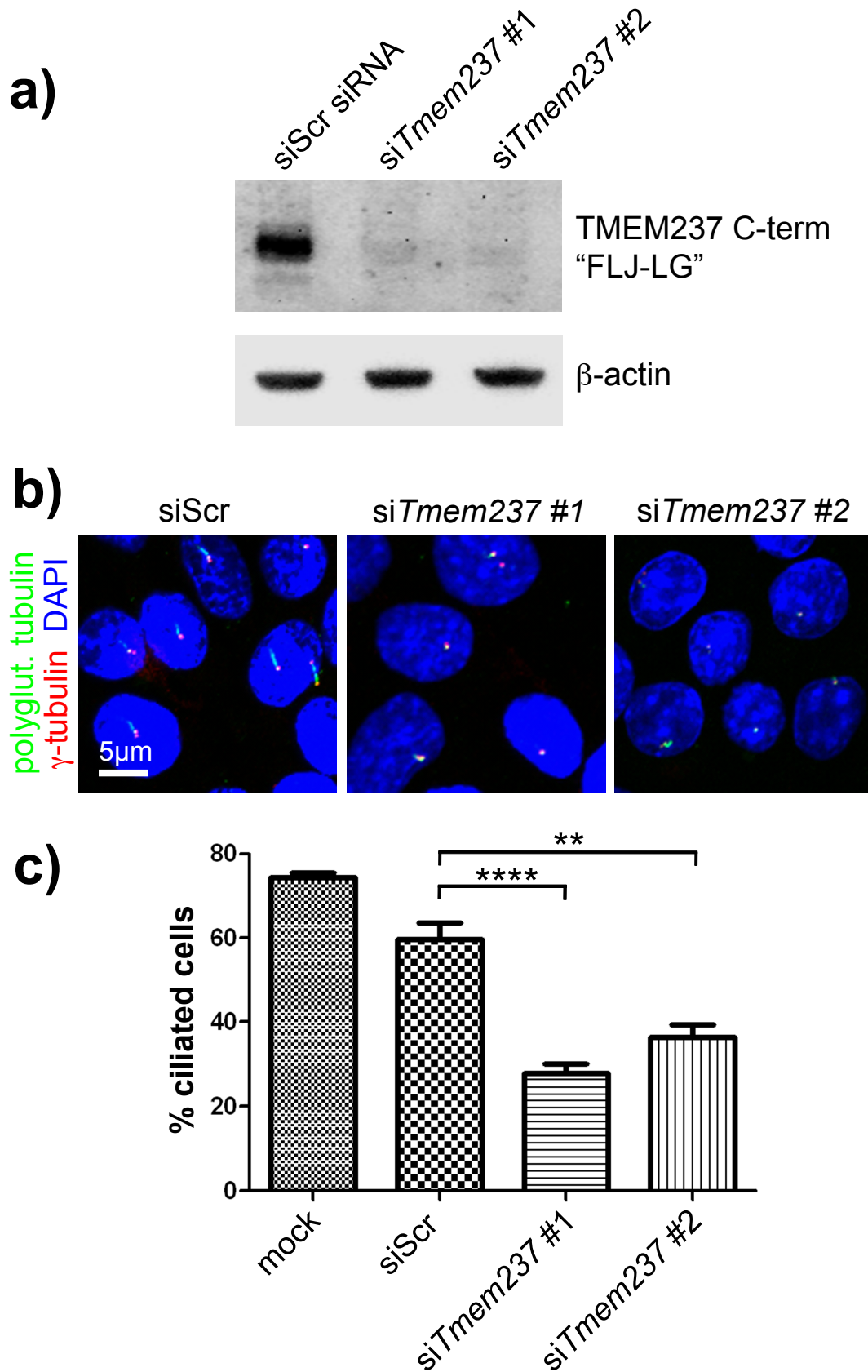


Figure 3-23. *Tmem237* siRNA knockdowns in IMCD3 cells. a) Western blot analysis demonstrating loss of protein in lysates from IMCD3 cells transfected with

si*Tmem237* compared to siScr; b) immunofluorescence staining of IMCD3 cells transfected with si*Tmem237* demonstrating loss and shortening of cilia compared to siScr. Green - cilia marker polyglutamylated tubulin, red - basal body marker γ -tubulin, blue - nucleus marker DAPI; c) bar graph showing loss of cilia in cells transfected with si*Tmem237* compared to siScr, for mock reaction transfection reagent only was used (** represents $p < 0.01$ and **** represents $p < 0.0001$, Student's t-test, for pairwise comparisons, two-tailed; error bars represents standard error of the mean). Figure adapted from Huang et al. 2011.

Knockdown of *Tmem237* caused not only cilia loss but also shortening of cilia (Figure 3-24), a phenotype often observed following loss or mutation of proteins involved in ciliogenesis²².

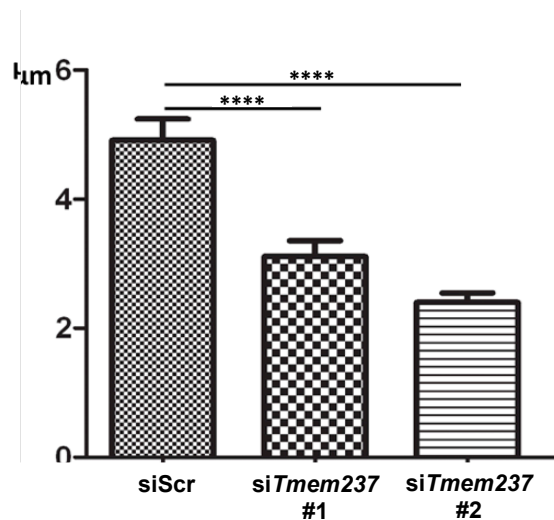


Figure 3-24. Measuring the effects of siRNA knockdown of *Tmem237* on cilia length. Diagram showing shortening of cilia length after treatment with si*Tmem237* compared to siScr treatment. Length (μm) of cilia for >300 cells were measured, and pairwise comparisons (Student t-test, **** represents $p < 0.0001$) are indicated. Error bars represents standard error of the mean.

3.2.4.6 TMEM237 in Wnt signalling

Many aspects of actin-dependent polarized cell behaviour, including morphogenetic cell movements²⁴⁷ and ciliogenesis²⁴⁸, are mediated by the planar cell polarity (PCP) pathway of non-canonical Wnt signalling (section 1.1.2.1.2). Perturbation of non-canonical Wnt signalling is implicated in the pathogenesis of MKS and deregulation of the canonical β -catenin pathway is implicated in the ciliopathy disease state^{99,159,162}. To assess if these pathways were perturbed following *TMEM237* mutation or loss, levels of key mediators in patient and control fibroblasts were determined. Immunoblot analysis of patient fibroblasts protein extract demonstrated constitutive phosphorylation and hyperactivation of Dishevelled-1 (Dvl1), a core Wnt signalling protein, and an increase in both soluble and total levels of β -catenin (Figure 3-25a). Levels of downstream effectors for both

canonical and non-canonical Wnt signalling, cyclin-D1 and phosphorylated myosinIIB respectively, were also assessed using immunoblot analysis. A slight increase in phosphorylated myosinIIB levels and a striking decrease in the amount of cyclin-D1 in “JSRD2” patient cells was observed (**Figure 3-25a**). RhoA activation was then assayed as the Rho family of small GTPases are key PCP mediators. Consistent with previous results following *TMEM67* or *TMEM216* mutation or knockdown, RhoA signalling was found to be activated in *TMEM237*-mutated fibroblasts (**Figure 3-25a**) or following *Tmem237* knockdown (**Figure 3-25b**) despite normal total amounts of RhoA in these cells.

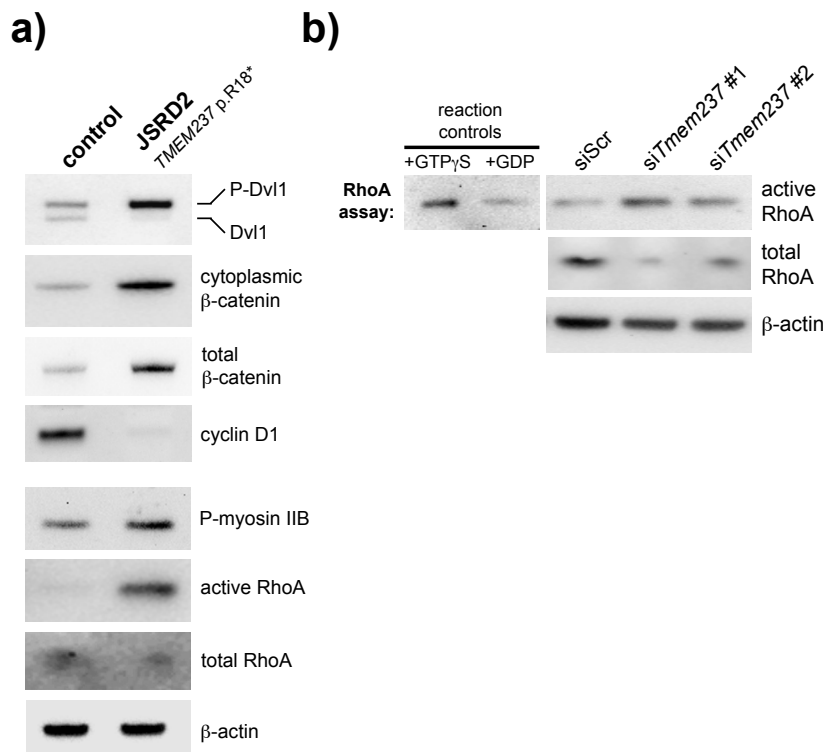


Figure 3-25. Deregulation of Wnt signalling following mutation or loss of *TMEM237*. a) Western blots showing the relative levels of the indicated endogenous mediators of both canonical and non-canonical/PCP Wnt signalling in normal control fibroblasts compared to the *TMEM237*-mutated patient fibroblast line (JSRD2 p.R18* line). Levels of β -actin and total RhoA indicate relative protein loading; b) Increase in levels of activated RhoA-GTP following si*Tmem237* knockdown with duplexes #1 or #2, compared to negative control (siScr). Levels of β -actin are shown as loading controls. Positive control for the assay (loading with non-hydrolyzable GTP γ S) and a negative control (loading with GDP) are shown on the left.

RhoA localized to the basal body in confluent IMCD3 cells (**Figure 3-26**), supporting a role in mediating centrosome docking at the apical cell surface prior to ciliogenesis. However, following *Tmem237* knockdown, RhoA was mislocalized to peripheral regions of the basal body and to basolateral cell-cell contacts (**Figure 3-26**), consistent with translocation of ectopically-activated RhoA to the cytosol. Since RhoA modulates the actin cytoskeleton in the PCP pathway, the cytoskeletal phenotype of *TMEM237* patient fibroblast lines was evaluated. Strikingly, prominent

actin stress fibres were seen in the patient fibroblasts but not in normal control fibroblasts (**Figure 3-26**).

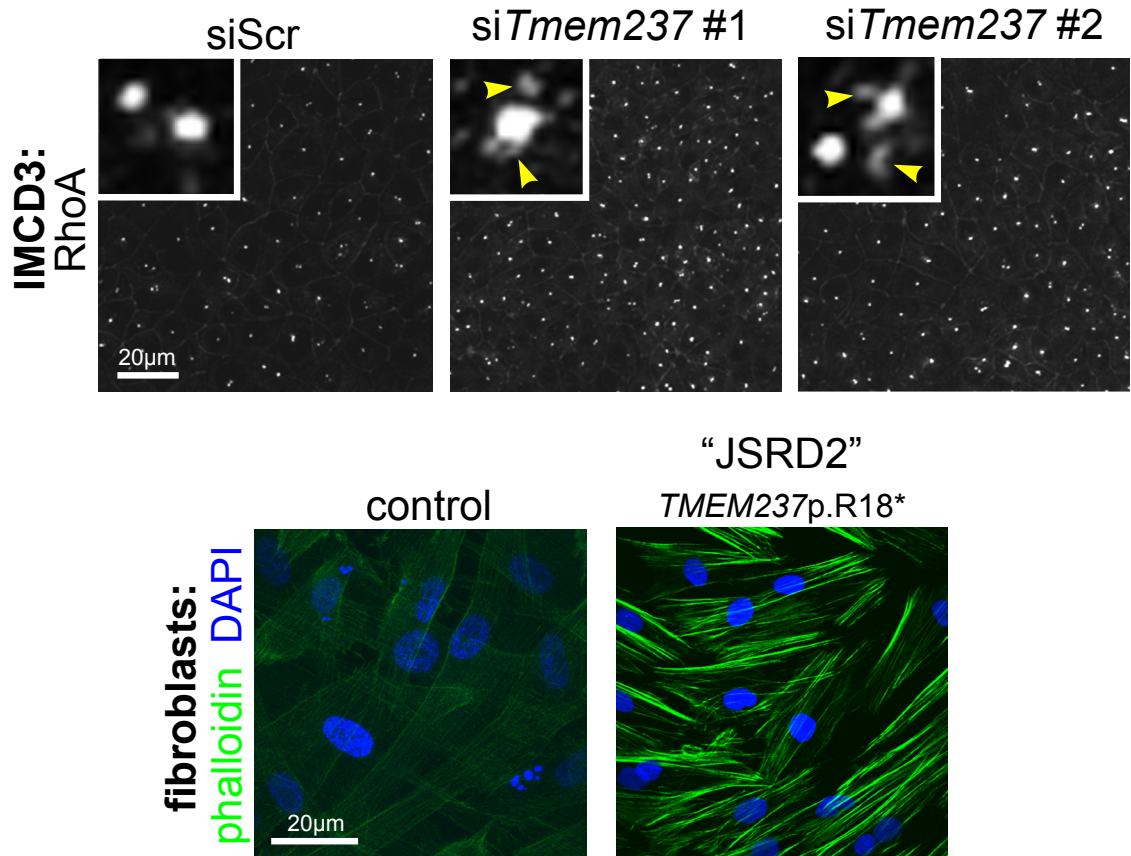


Figure 3-26. Disruption of the noncanonical Wnt signalling effectors subcellular localisation. Top panel – RhoA localisation to the basal body after treatment with siScr. This localisation was disrupted after treating cells with si*Tmem237* (#1 and #2), where RhoA was mislocalized to the peripheral regions of the basal body (yellow arrows). Bottom panel – actin marker staining (phalloidin) showed an increase in actin stress fibres in JSRD2 patient fibroblast compared to the control cells. Scale bar=20 μ m. Image adapted from Huang et al. 2011.

To assay levels of Wnt signalling in those cells TOPflash assay was done. In this assay luciferase levels are regulated by nuclear β -catenin binding to TCF/LEF regulatory elements. Therefore the more nuclear β -catenin the more luciferase expressed. JSRD2 patient fibroblasts demonstrated dysregulated canonical Wnt signalling (over 5-fold basal levels) compared to control fibroblasts upon stimulation with Wnt3A-conditioned media (a canonical Wnt signalling ligand) (**Figure 3-27a**). Treatment with Wnt5A (a noncanonical Wnt signalling ligand) had no effect on activation; however, Wnt5A suppressed the activation by Wnt3A (**Figure 3-27a**).

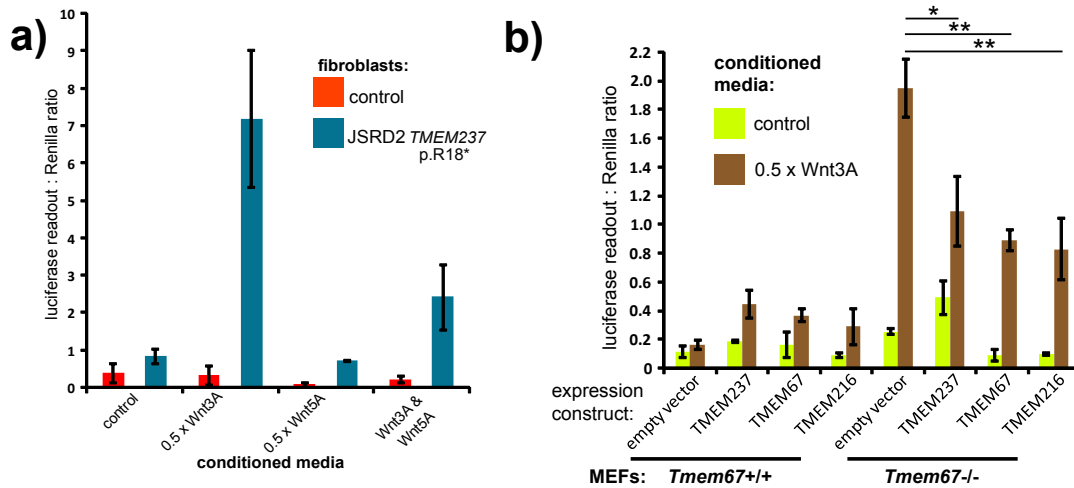


Figure 3-27. TOP Flash assays of canonical Wnt signalling. a) Assay showed deregulation in patient fibroblasts (JSRD2 p.R18*, red) compared to normal control cells (blue) following treatment with 0.5x L cell control conditioned media (control), and conditioned media containing expressed Wnt3A and/or Wnt5A, as indicated. Activity is expressed as ratios of luciferase reporter construct expression, normalized for loading by measurement of a Renilla construct expression. Values shown are means of three independent replicates, with error bars indicating standard error of the mean (s.e.m.); b) TOP Flash assays of canonical Wnt signalling, following cotransfection of *Tmem67*^{+/+} and *Tmem67*^{-/-} MEFs with reporter constructs and empty vector control, wild-type GFP-*TMEM237*, HA-MKS3 or GFP-*TMEM216* as indicated. The empty vector results combine the data from co-transfections with pCMV-HA and pGFPN1. Responses are shown to L cell control conditioned media (control, green bars, values are means of three independent replicates) and conditioned media containing expressed Wnt3A (0.5x Wnt3A, brown bars, values are means of four independent replicates). Error bars indicate s.e.m. Statistical significance of pairwise comparisons are shown (* p < 0.05, ** p < 0.01, Student t-test). Figure adapted from Huang et al. 2011.

Previously transmembrane proteins involved in MKS/JSRD have been shown to interact⁹⁹. N-terminal part of *TMEM67* contains a leucine-rich domain, which is a characteristic feature of the Frizzled receptors, which are the receptors for the Wnt ligands. It was therefore suggested that *TMEM67* may be a Frizzled-like receptor and its interaction with other MKS/JSRD transmembrane proteins, acting as a co-receptors, facilitates the correct Wnt signalling. Consequently any interaction between *TMEM237* and other transmembrane proteins, like *TMEM67*, have been investigated using mouse embryonic fibroblasts (MEFs, obtained from *Tmem67*^{-/-} and *Tmem67*^{+/+} mice). Deregulated Wnt signalling activation by Wnt3A was also apparent in MEFs derived from wild-type *Tmem67*^{+/+} and knock-out mutant *Tmem67*^{-/-} E18 embryos (**Figure 3-27b**). Interestingly, responses were attenuated by expression of *TMEM237*, *TMEM67* and *TMEM216*, indicating either partial complementation of *TMEM67* loss by other TZ-localized members of the *TMEM* family or Wnt signal compensation from another cellular site (or process).

3.2.4.7 RPGRIP1L knockdown interferes with TMEM237 localisation

RPGRIP1L was shown to be crucial TZ component, having an important function in correct localisation of other TZ components ²⁴⁹. It shares subcellular localisation with Tmem237 and further interaction between those proteins was investigated. Knockdowns of *Mks5/Rpgrip1l* in IMCD3 cells were performed and cells were stained for Tmem237 and the basal body markers γ -tubulin (**Figure 3-28a**) or Dvl1 (**Figure 3-28b**), and RhoA (**Figure 3-28c**). Following scrambled siRNA transfection, basal body architecture was observed to be intact (**Figure 3-28b**), and Tmem237 was just distal to the basal body (**Figure 3-28a** and **b**; marked by either γ -tubulin or Dvl-1, respectively). This localization to the TZ was disrupted following knockdown of *Mks5/Rpgrip1l*. RhoA was also mislocalized to peripheral regions of the basal body and to basolateral cell-cell contacts (**Figure 3-28c**), identical to the cellular phenotype following *Tmem237* knockdown (**Figure 3-28c**) as well as *Tmem216* knockdown ⁹⁹, indicating Rpgrip1l plays an important role in Tmem237 localisation ²⁴⁹.

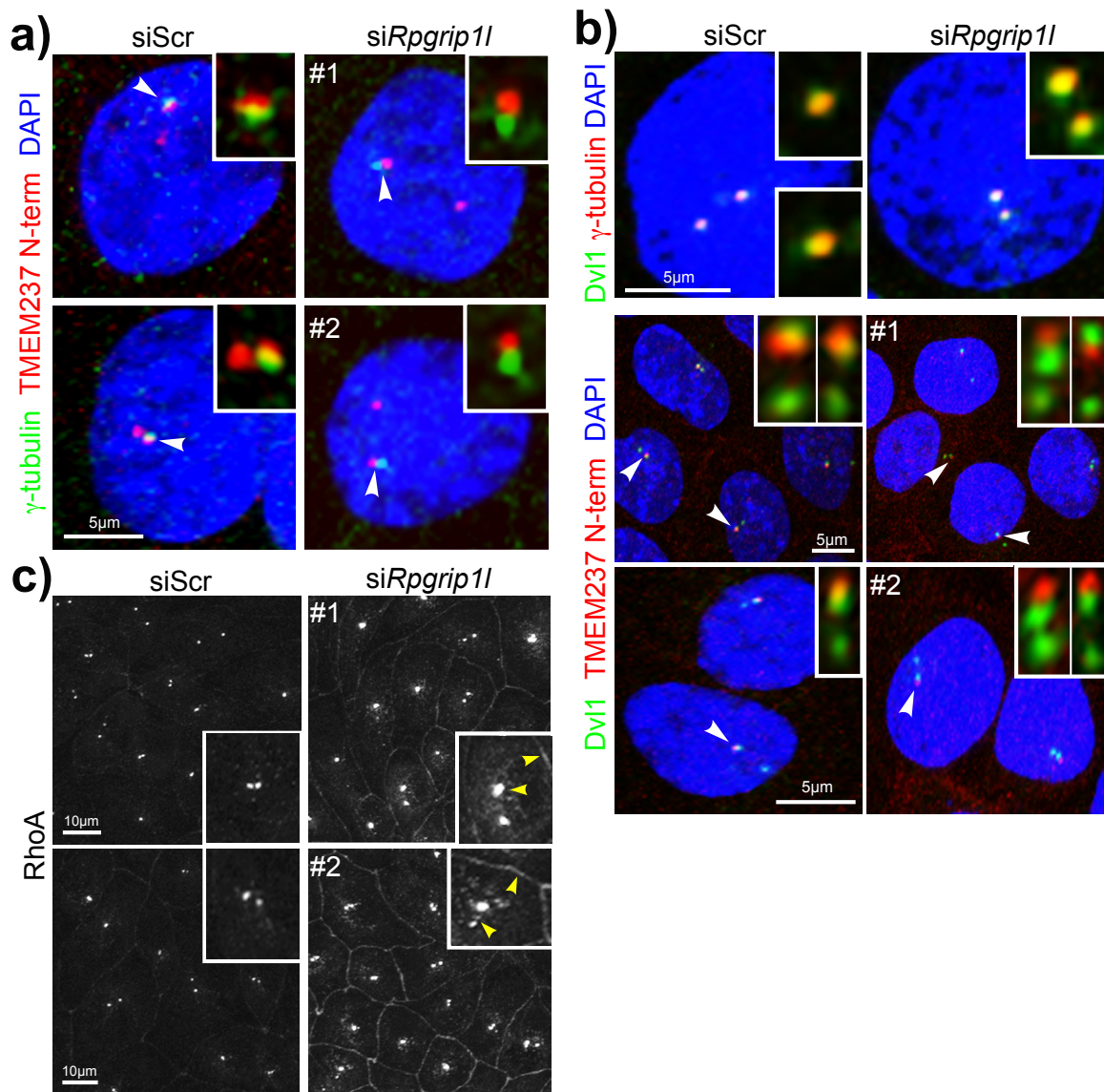


Figure 3-28. Ciliary TMEM proteins are anchored at the TZ by MKS5/RPGRIP1L.
 a) In siScr-treated IMCD3 cells, Tmem237 (red) is at proximal regions of the cilium and partially colocalizes with γ -tubulin (green) at the basal body, consistent with Tmem237 localization to the TZ. Colocalization with γ -tubulin is disrupted following siRpgr11 knockdown. Insets show magnified regions (indicated by arrowheads on the main image) for the red and green channels only; scale bars 5 μ m; b) Top panels: Dvl1 (green) is a basal body that colocalizes with γ -tubulin (red) in IMCD3 cells, as described previously. siRpgr11 does not disrupt overall basal body architecture. Bottom panels: in siScr-treated IMCD3 cells Tmem237 partially colocalizes with Dvl1, but this is disrupted following siRpgr11 treatment; c) As above, but with RhoA staining; magnified insets indicated by white frames.

3.2.4.8 Loss of TMEM237 disrupts G1/S transition

During the analysis of cyclin D1 levels in JSRD2 (*TMEM237*-mutated) cells, loss of expression of cyclin D1 protein was observed (**Figure 3-25a**). This was unexpected as cyclin D1 acts downstream from β -catenin, and in normal cells, its levels are observed to be elevated during up-regulation of canonical Wnt signalling.

Propidium iodide staining and FACS analysis was performed to understand the effect of the absence of cyclin D1 expression on the cell cycle after loss of TMEM237. Patient and control fibroblasts were cultured in normal growth media and had cell cycle synchronised by treatment with thymidine for 18h hours. They confirmed that JSRD2 cells had a prolonged G1 phase and very short S phase, although G2 remained similar to the one observed in the control fibroblasts (**Figure 3-29**). Proliferation of both control and patient fibroblasts remained at a similar level, based on cell number count, and there is an evidence in the literature that in the absence of cyclin D1, cyclin D2 levels are increased to compensate and maintain proliferation²⁵⁰.

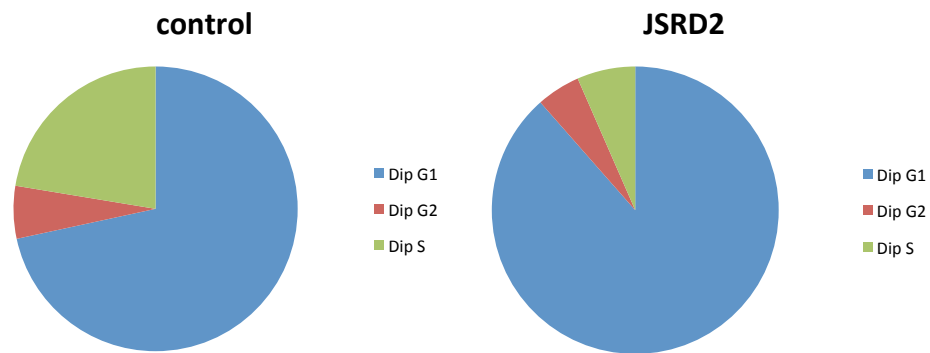


Figure 3-29. FACS results showing that JSRD2 fibroblasts have a prolonged G1 phase and shortened S phase compared to control cells. Blue, the proportion of cells in G1 phase; green, S phase cells; and red, cells in G2 phase. Abbreviation: Dip - diploid cells.

3.2.5 Discussion

With a vibrant development of new technologies allowing fast and cost-effective causative mutation discovery, it is expected to identify all genes implicated in a certain syndrome (<http://www.omim.org/statistics/update>). In this chapter the author focused on this task.

Using SNP chip analysis and genotyping, a putative locus on chromosome 12 was identified. No causative mutations were identified in the best functional candidates within this locus, *TCTN1* and *TCTN2*. These genes were later proved to be good functional candidates, as mutations were identified in *TCTN2* causing MKS¹⁰⁴ and in *TCTN1* causing JSRD⁴⁷. None of the samples had shared haplotypes that were compatible with linkage to this locus, although some of the individuals shared the same ethnic background. It is also known from a parallel study that the cause of the phenotype in patient 158 is a mutation in *CC2D2A* (c.3540delA p.R1180Sfs*6), and for patient 261 a mutation in *TMEM237* (c.1066_1067dupC

p.Q356Pfs*23). This could suggest that this chromosomal region is highly homozygous in consanguineous populations (Dr Ian Carr, University of Leeds, personal communication). Consanguineous first cousins, in theory, share 6% (1/16) of their genome (though in practise those numbers are shown to be higher ²⁵¹), therefore large homozygous chromosomal segments may contain genes with disease causing mutations. Those regions are called “identical by descent” (IBD) or “autozygous” and strategy of disease gene identification based on this assumption is called “autozygosity mapping”. This strategy was used in this study yet findings suggest that the locus on chromosome 12 is likely to be a false positive finding. The other samples homozygous at this locus still do not have any causative mutations identified despite additional screening of new MKS genes. It is therefore possible that another gene in this locus may have a causative mutation, or these patients actually harbour mutations in the non-coding regions of *TCTN1* or *TCTN2*. The latter possibility could be investigated by either whole genome sequencing (WGS) to look at deep intronic mutations, changes in the regulatory elements or copy number variants (CNV), or in combination with RNA sequencing to look at transcript levels ²⁵², or alternatively by long-range PCR followed by clonal sequencing ²⁵³.

Collaborative work between groups interested in ciliopathies have proved to be fruitful and very efficient, especially in the event of monoallelic mutation identification. The best examples are the *CEP164* and *TTC21B* projects. Although there were no biallelic mutations identified in our cohort in these genes, it allowed the community to better understand the frequency of mutations in those genes and phenotypes they are and are not involved in ^{182,236}.

The collaboration with the group at Duke University highlighted the importance of data quality and filtering methods in variant calling, as the false positive rate in those results reached 45%. This project was an opportunity to have all of the CiliaProteome v3 genes sequenced in these samples, and the maximum available amount of DNA was sent (often 5µg). However, even these amounts of DNA were insufficient for the targeted capture protocol and WGA was often performed. This primer extension reaction allows the amplification of short fragments of genomic DNA using random primers. It uses a polymerase that lacks proof-reading activity and is prone to introducing single base-pair changes. The results obtained from sample processed with WGA should be therefore interpreted with some caution. Obtained results indicate that mutations in IFT genes are not the cause of MKS/JSRD. Heterozygous changes were identified and they may contribute to the severity of the phenotype ¹⁸². In two cases (76 and 256) homozygous changes were observed in patient stock DNA but segregation showed that either only one of the parents carried the mutated allele (76) or affected siblings were only heterozygous

for the mutation (256). These findings could be caused by poor quality DNA or allelic drop-out. However, in case 76 a causative mutation was identified in *TMEM67* (c.1575+1G>A) excluding variant in *IFT88* as a possible cause of the phenotype. Further analysis of this study is required and extensive genetic and proteomic data comparisons are being undertaken to understand the genetic and phenotypic variability in ciliopathy patients and the involvement of changes identified in IFT genes.

One further collaboration concerned the understanding of the function of the *TMEM107* protein and verifying whether it is involved in a human ciliopathy phenotype. Sequencing of patients with linkage compatible to the *TMEM107* locus did not reveal any variants that could be interpreted as pathogenic and extensive screening of many ciliopathy patient cohorts did not identify mutations. This suggested that mutations in this gene are either not a cause of human ciliopathies, cause another disorder that was not represented by the screened cohort or they affect early embryonic development stages and are not compatible with life. *TMEM107*-GFP was shown to localise to the base of the cilium and ciliary axoneme (**Figure 3-9**) and loss of *Tmem107* caused shortening of the cilium and affected cilia and cell number (**Figure 3-10**). To further elucidate the function of this protein, work in *C. elegans* was undertaken by a collaborator as well as the identification of novel protein-protein interactions using the TAP assay. The predicted tetraspanin-like structure of *TMEM107* resembles that of other *TMEMs* implicated in ciliopathies. It has been postulated that these proteins interact together to function at the base of the ciliary membrane (at the so-called “ciliary necklace” of the transition zone) and regulate the trafficking of other proteins to and from the cilium (Dr Oliver Blacque, University of Dublin, personal communication).

WES was used to identify mutations in consanguineous families with an MKS/JSRD phenotype. Three different sequencing platforms were used with varying output data quality. DNA from family 157 was sequenced on the Illumina GAllx platform and biallelic homozygous changes in homozygous regions shared between two affected individuals were prioritised. There were no obvious functional candidates, but the p.R317Q variant in *DLL1* segregated with the phenotype in the family. Screening of this gene in an MKS/JSRD cohort revealed only heterozygous changes and an additional heterozygous change was identified in a UK10k sample. This change was identical to the one found in family 157 and it may suggest that this change is a rare variant for the Pakistani population. It is possible that the true mutation was not detected during this analysis or it was present in the intronic sequences, as seen in *OFD1*²⁵⁴ and *CEP290*²⁰⁹ or indeed the cause of the phenotype was a large insertion/deletion. The possibility that the change in *DLL1* is

the true cause of the phenotype was not excluded especially because ENU induced mice with a homozygous missense mutation in *Dll1* (p.E26G) showed a phenotype resembling ciliopathies including *situs inversus*, ectopic neural tube and short tail²⁵⁵. Further analysis should be taken to dissect the cause of the phenotype in this family, which could include WES of the second affected and unaffected siblings, as well as WGS to look at possible CNV and deep intronic/regulatory elements changes.

Samples sequenced on the Illumina MiSeq platform showed very low read coverage per base and it made interpretation of the results very difficult. In family 36A, five homozygous variants were identified that passed analysis filters but only one of them was localised in a homozygous region and showed quite good read depth (RD=18) (average RD for this sample was 7). The variant in *MACF1* segregated with the phenotype in the family but no further analysis was done as this gene is large (93 exons) and the possibility of a true variant being missed during data analysis and filtering was quite high. With no previous reports of microtubulins proteins being directly involved in ciliopathies it was decided that these samples will have to be re-sequenced to get much higher base coverage.

The second family (66F1) sequenced on the MiSeq had a mutation identified in *TXNDC15*. The change segregated with the disease in the family and sequencing of this gene in an MKS/JSRD cohort revealed additional changes but none were homozygous (**Table 3-12**). A variant in this gene was reported in another MKS patient²⁵⁶, but the authors did not suggest that it was causative, as it was an in-frame deletion and better functional candidate had mutations. Another thioredoxin gene (*TXNDC3*) had a single mutation identified as a cause of primary ciliary dyskinesia²⁵⁷. Altogether this may suggest that the change in *TXNDC15* is the true cause of the MKS in family 66F1, but further functional studies are essential to dissect out the function of *TXNDC15* in cilia.

Results from the Illumina HiSeq2500 sequencing platform provided much better data quality than those from the MiSeq. A mutation in *EXOC3L4* was identified in family 17 and the change segregated with the disease in the family. Sequencing of this gene in the MKS/JSRD cohort revealed only two additional heterozygous changes. The function of *EXOC3L4* in ciliogenesis is unknown but the excocyst complex has been strongly linked to trafficking to the primary cilium²⁵⁸. The identified variant in *EXOC3L4* has been assigned an rs number indicating that it may be a benign variant. However, the MAF for this change is 0.0005, therefore this variant is very rare and could be pathogenic. This scenario was observed in the study that identified mutations in *SLC38A8* as a cause of Foveal Hypoplasia and Optic-Nerve-Decussation Defects. Poulter et al. found a variant

with rs number (rs149592537) p.Q200*²⁵⁹, which was observed at a very low frequency in the databases (1/12,999 in the EVS and 1/4,545 in dbSNP), suggesting that it could be pathogenic and therefore causative for the phenotype.

Family 325 has three affected offspring but SNP chip analysis on both pooled and unpooled samples showed no shared homozygous regions. There was confusion in sample information as sample 325 was reported as a male, but SNP chip data clearly showed it was a female. This could have also suggested that naming on the samples was mixed up, as for this family no maternal DNA was provided. Further information obtained from the clinician confirmed that initial sample calling was correct with the exception of the sample gender. There was no clear functional candidate amongst those with homozygous changes identified in WES (**Table 3-14**). The possibility was not excluded that the cause of the phenotype in this patient could be a deep intronic mutation or big deletion/insertion.

In family 144 a change was identified in *BBS12* p.R675* and it was thought that this result will extend the allelism seen between the ciliopathies. However, further phenotypic investigation of the patient showed that they presented a typical BBS phenotype and had been mis-diagnosed initially. This example highlighted the possibility of a differential diagnosis in the Leeds cohort and it should be considered during further studies.

This point was particularly valid during the investigation of sample 351 where initially a change in *ALG9* (c.764C>A, p.S255*) was filtered out as it was at the same position as a SNP (p.S225L). *DIXDC1* was screened in the whole MKS/JSRD cohort with no mutations identified. Only when the analysis was repeated it was noticed that a novel nonsense mutation in *ALG9* was present which segregated with the disease in the family. Mutations in *ALG9* are associated with Congenital disorder of glycosylation (CDG) type 1I²⁶⁰ and further investigation of the patient's phenotype revealed that it was not typical MKS but it actually resembled CDG phenotype. *ALG9* was screened in the whole MKS/JSRD cohort to exclude from further studies other samples misdiagnosed for CDG.

Identification of private mutations makes it difficult to establish if the change is the true cause of the phenotype. On average 66 homozygous rare pathogenic coding variants are being identified in first cousin consanguineous individuals (figure based on in-house WES data) that need to be prioritized by literature mining for the known protein function and interactions, structure analysis and animal models. The evidence suggesting that a particular variant is the cause of the phenotype has to be compelling before embarking on additional functional studies. With the rapid progress of gene discovery, the identification of the causes of Mendelian conditions has expanded vastly, leaving rare, often private variants to be

identified. Many samples analysed during this study proved this, and only sequencing of large cohorts and identification of a second case with mutations enables researchers to gain confidence in the true pathogenic nature of the variant. This was shown in the case of *CSPP1* when a single homozygous private mutation was identified by a Canadian group who were unable to find additional patients with mutations. The project was not taken forward and only after the identification of another variant by collaborators in Saudi Arabia, the evidence sufficient to be reasonably certain that mutations in this gene were a cause of MKS¹⁰⁸. Surprisingly, other groups identified additional mutations in their patient cohorts resulting in three back-to-back publications reporting mutations in *CSPP1*^{108,150,151}. It remains possible that the variants found in patients 157, 17, 36A, 66F1 may in the future be confirmed as pathogenic by independent replication in other cohorts.

TMEM237 mutations were initially identified by a group in Canada in two Hutterite families. Sequencing of the Leeds cohort of MKS/JSRD patients revealed an additional mutation in sample 261 (p.Gln356Profs*23). Patient 261 was initially diagnosed with MKS but a detailed phenotype provided by the clinician indicated JSRD. Functional analysis of *TMEM237* showed that the protein localises to the base of the cilium, the TZ. In mouse proteomic studies, *Tmem237* was found in the photoreceptor connecting cilium complex and outer segments²⁴⁶. *TMEM237* is a predicted tetraspan transmembrane protein with both amino and carboxyl termini directed to the cytoplasmic side of cells. Analysis of patient cell lines showed that the *TMEM237* transcript was at very low level compared to the control (**Figure 3-21**) and cilia were lost in those cells. Knockdown of *TMEM237* in mIMCD3 cells also caused loss of cilia and cilia shortening. *TMEM237* is hypothesised to be a co-receptor for *TMEM67* and *TMEM216*, and consistent with this hypothesis patient cell-lines for *TMEM67* and *TMEM216* showed deregulated canonical and non-canonical Wnt signalling^{99,206}. Downstream effectors for these pathways were also analysed in *TMEM237* patient cell line and showed prominent up-regulation suggesting that *TMEM237* is a negative regulator of Wnt signalling. Unexpectedly, cyclin D1 levels were shown to be lower than in the control sample. The increased levels of β -catenin, a transcriptional regulator of cyclin D1, would suggest that high level of cyclin D1 should had been observed. This may indicate that *TMEM237* plays a role in the correct localisation of β -catenin to the nucleus. The blocking of the *TMEM237* patient cell line in the G1 phase of the cell cycle confirms loss of cyclin D1 as it is a key regulator of G1/S transition²⁶¹. As *TMEM237* showed localisation to the TZ, it was investigated if its localisation would be affected by loss of the TZ component *RPGRIP1L*. *TMEM237* showed mislocalisation after loss of *RPGRIP1L* suggesting that this protein is crucial for the correct localisation of

TMEM237 to the TZ. Further studies to elucidate the consequence of TMEM237 loss were not conducted.

Other members of the transmembrane protein (TMEM) family, TMEM216/MKS2 and TMEM67/meckelin, have both been implicated in JSRDs and MKS and both localize to the ciliary basal body or TZ region^{99,152,262}. Since other tetraspan transmembrane proteins function through the formation of complexes with each other, and other membrane proteins such as Frizzled receptors²⁶³, it suggests that TMEM237, TMEM216 and TMEM67 may cooperate in maintaining normal ciliary function.

The addition of TMEM237 to the suite of TMEMs mutated in ciliopathy patients, highlights the importance of this group of proteins in the function and structure of the primary cilia. Mutations in *TMEM237* were recently reported to cause MKS²⁵⁶ and other transmembrane protein: for example, TMEM231 was implicated in the MKS phenotype¹⁰⁷. Further investigation into the function of these particular transmembrane proteins in the cilia structure and function could help in prioritisation of further candidate genes responsible for MKS/JSRD phenotypes.

Small nuclear families, poor quality DNA and incomplete phenotypic descriptions are all issues that were faced in this study. Analysis of large patient cohorts shows that the most common causes of MKS/JSRD have already been identified¹¹¹ and recent reports of new genes involve one family with one mutation (so-called “private” mutations)^{104,105,216}. Therefore, the possibility that the remaining causative mutations could be private was taken into consideration in this study. Big insertion/deletions and deep intronic changes, known to cause a disease phenotype^{209,254} were also considered. The technology moves now to WGS and this method has become more validated, robust and cost effective. It was shown not only to allow identification of CNVs but also structural DNA variations like inverted duplications or SNV²⁶⁴.

New technologies may bring advances in techniques and a reduction in the cost of the analysis, but they do not always provide an ‘easy’ answer. The high variant/noise ratio is still of a great consideration during data interpretation and careful analysis have to be applied. The above data shows that it is only one of many obstacles in the way to get a definitive molecular diagnosis. However, a careful and detailed clinical summary of the patient phenotype is of a great value to the researcher. For example, differential diagnoses were made for families 144 and 351 this approach showed the effectiveness of WES analysis and enabled the reclassification of the patient phenotypes.

3.3 High-throughput genome-wide siRNA visual screen for effects on ciliogenesis

The identification of functional candidate genes that cause a certain phenotype is not always possible or accurate. In the case of ciliopathies, candidate genes are prioritised based on known or predicted protein function and possible links to cilia or signalling pathways linked to cilia. This can be a useful strategy to filter variants obtained from a WES or WGS experiment in order to identify causative mutations. However, the function of many proteins and their subcellular localisation is often unknown or has not been previously investigated in the context of ciliary biology. To identify proteins that are implicated in cilia structure or mechanisms important in ciliary assembly/disassembly on a global level, an unbiased whole mouse genome siRNA screen for ciliogenesis was performed.

3.3.1 Screen set up

3.3.1.1 Cell line

To obtain meaningful results from visual screening strategies, cilia needed to be reliably and accurately resolved in sufficient numbers to generate a statistically powerful assay. *In vivo*, the primary cilium is a structure formed on epithelial cells during G0 and G1. However, *in vitro*, most ciliated cell lines only form a cilium during G0 phase, which must be induced by growing cells to confluence followed by serum starvation. Mouse inner medullary collecting duct (mIMCD3) cells are unusual in their ability to readily polarize and form single, long (up to 10µm) cilia in culture in both G0 and G1, particularly after serum starvation. For this reason, mIMCD3 cells were chosen for this assay over alternative ciliated cell lines such as human immortalized retinal pigment epithelial (hTERT-RPE1) or canine Madin-Darby Canine Kidney epithelial (MDCK) cells.

3.3.1.2 siRNA library

The siGENOME mouse siRNA library from Dharmacon/ThermoFisher Scientific was used, and 19,097 genes were screened in duplicate with the methodologies described in section 2.2.18. The library was divided into 10 sub-libraries: G-protein-coupled receptors (GPCRs), protein kinases, ion channels, proteases, phosphatases, three sub-libraries targeting components of ubiquitin

proteasome system, and two large sub-libraries for “druggable” targets and the remaining genes in the mouse genome.

3.3.1.3 Screen set up and analysis

3.3.1.3.1 Assay set up

Seeding 8,000 cells per well in 96-well optical bottom plates resulted in a uniform monolayer of 95-100% confluency after 72 hours of incubation and reverse transfection of siRNAs. This confluency was optimal for the detection of cilia at the same single focal plane in every well, since over-confluent cells grew in multiple layers. The proportion of ciliated cells was consistently above 60%, which is comparable with the levels reported previously^{99,205}.

Acetylated α -tubulin was used as a ciliary axoneme marker, due to its specificity for the entire length of the primary ciliary axoneme and low background staining. This antibody only recognises specific modified stable microtubules in the ciliary axoneme and minimally binds to cell body microtubules. Optimization experiments in pilot screens showed that separate primary and secondary antibody binding steps were essential for high-quality cilia staining, and that ice-cold methanol fixation gave much better resolution of cilia staining and lower background compared to other fixation methods.

To maintain simplicity of the assay each cell was imaged for three markers. The green channel imaged acetylated α -tubulin (a ciliary axoneme marker), the blue channel was used to visual nuclei with DAPI, and the far red channel used TOTO3 to segregate cytoplasmic and nuclear staining (**Figure 3-30**).

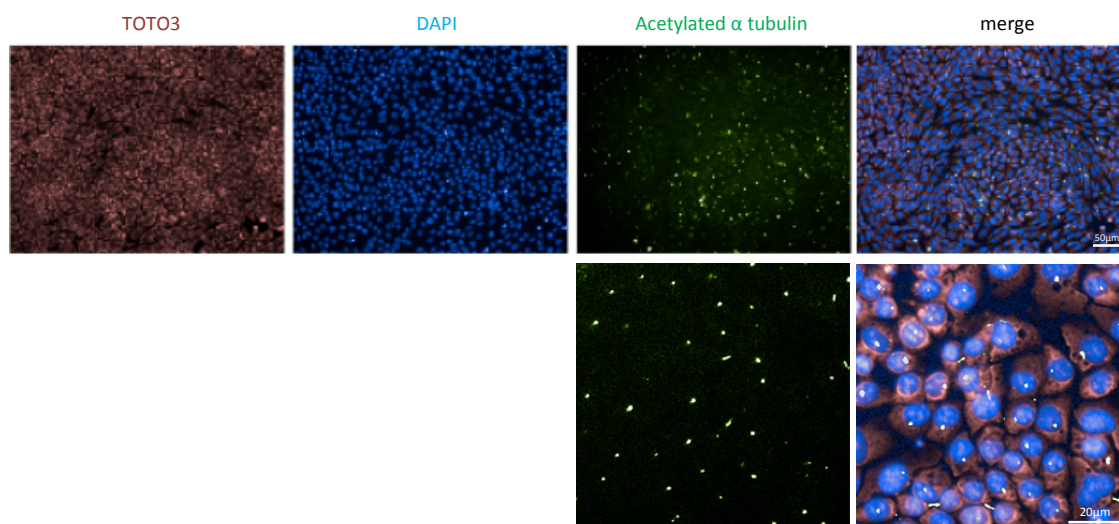


Figure 3-30. Sample images of IMCD3 cells. Cells were stained in far red for cytoplasmic marker (TOTO3), nucleus – blue (DAPI) and cilia marker – green (acetylated α -tubulin). Top panel scale bar = 50 μ m. Bottom panel figures show magnification of cilia staining and merged channels. Bottom panel scale bar=20 μ m.

Cilia in this assay pointed straight up (**Figure 3-31** right panel), therefore after imaging at the appropriate focal plane allowed the cilia to be visualized as spots of a characteristic size and intensity (**Figure 3-31** left panel).

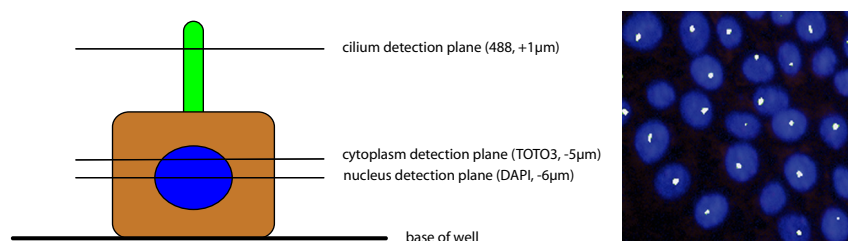


Figure 3-31. Schematic representation of focal planes imaged in IMCD3 (left panel) and an example image of visualised cilia (right panel). Three focal planes were imaged: one for the nucleus (DAPI, in blue), a second 1 μ m above it for visualization of the cytoplasm using TOTO3, and at 7 μ m above the nucleus a third focal plane to image cilia using acetylated α -tubulin (AaT, green) (left panel). Cilia were visualised as spots in siScrambled (negative control)-treated cells using Columbus visualization software (right panel).

3.3.1.3.2 Assay metrics

Two measurements were analysed: ‘whole cell number’ (WCN) and ‘% of cells with a single cilium’ (%CSC) (**Figure 3-32**). WCN was obtained in the far-red channel using TOTO3 staining and %CSC from the green channel using an Alexa 488 secondary antibody bound to acetylated α -tubulin. The same six fields of view per well were imaged across the whole screen. These fields of view were localised close to the centre of each well since a ‘doughnut effect’ of cell loss in the middle of the well was often observed.

Cells on the border of each field of view were excluded from further analysis, since cilia for these particular cells might have been missed causing underrepresentation in results and increasing false positive rates (**Figure 3-32**).

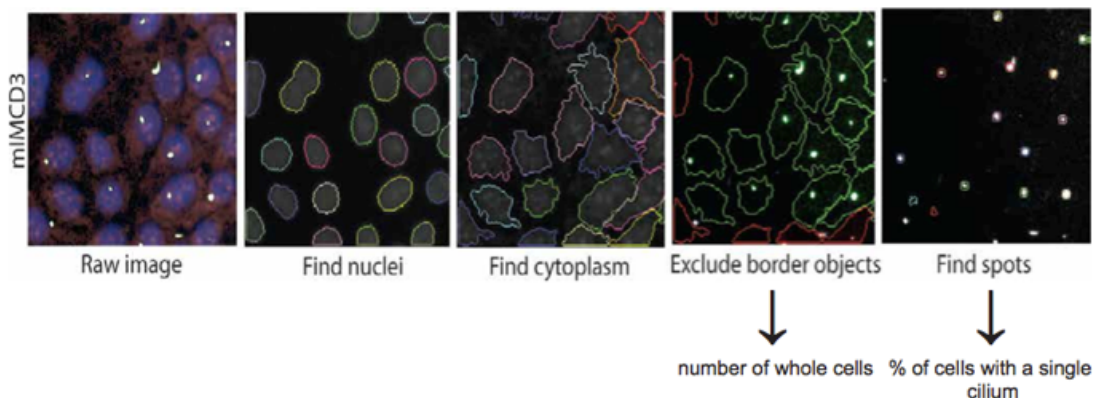


Figure 3-32. Example of image analysis in Columbus software. The input image was analysed using an algorithm for object recognition. The software recognised nuclei, cytoplasm, cilia (with a separate spot-recognition protocol), and excluded border objects from further analysis.

3.3.1.3.3 Assay analysis

Images in TIFF format were exported from an Operetta automated high-content imager and uploaded to Columbus on-line image analysis software. A script written for the purpose of this screen was applied (Perl script written by Dr David Parry) and the analysed data were exported in an Excel format and/or represented in the form of heatmaps (**Figure 3-33**).



Figure 3-33. Heatmaps representing results obtained from Columbus software. Top heatmap shows results for WCN and the bottom one for %CSC. Colour scale indicates that the darker the colour the less WCN/%CSC was observed in the well.

Measurements for WCN and %CSC were transferred from the Columbus output Excel files into template Excel files for the analysis of ten plates in duplicate. The screen was performed in batches of 10 plates (or less) as this was the maximum number which could be processed accurately by the available equipment. Duplicate plates were biological replicates and were set up in an identical manner for cells with a different passage number. Data from the biological replicates were inputted into a template file that was subsequently analysed using a Perl script (written by Dr David Parry). Robust z scores were automatically calculated and appropriate cut-offs were applied.

3.3.1.3.4 Robust z score

The use of median and median absolute deviation rather than mean and standard deviation reduces the effect of outliers ²⁶⁵. Therefore the robust z score was chosen as a statistical measure of cellular phenotypes compared to a series of positive and negative controls.

$$\text{robust z score} = (x - \mu) / \sigma$$

x value per well

μ median of values of all negative controls

σ MAD (Median Absolute Deviation)

where MAD is calculated:

$$MAD = \text{median}_i(|y_i - \mu|)$$

y_i negative control value per well

μ median of values of all negative controls

To ensure normalisation of the data and to exclude batch-specific effects, data were analysed within batches rather than across batches. Measurements of cilia number were normalised against whole cell number to remove systematic errors, plate-specific effects and allow comparison of results between different plates and batches.

3.3.1.3.5 Assay controls

A series of positive controls were optimised and evaluated to identify the most appropriate ones for this screen. These controls included siRNAs targeting known ciliopathy or ciliary-related genes including *Lca5*, *Ift80*, *Mks1*, *Tmem67*, *Ift88*, *Ttc21b*, *Pla2g3*, *Rpgrip11* and *Kif3a* (**Figure 3-34**).

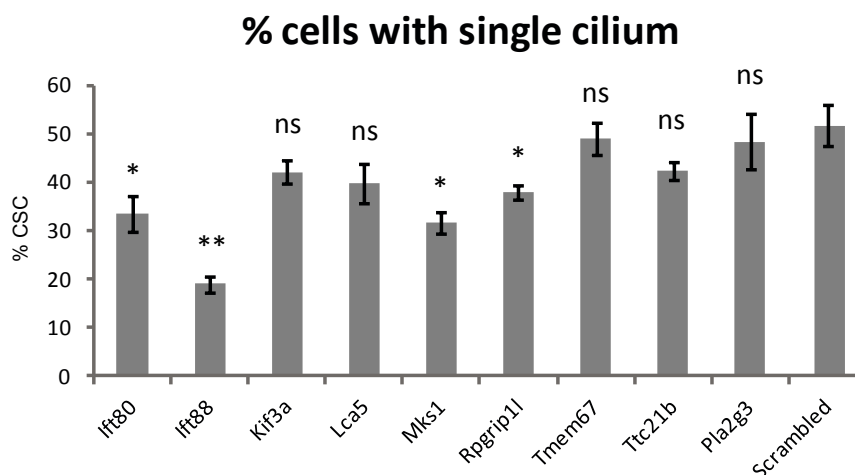


Figure 3-34. Diagram representing knockdown efficiencies in a panel of positive controls. %CSC was calculated in a panel of positive controls compared to scrambled negative control siRNA. The greatest effect on %CSC was observed after knockdown of *lft80*, *lft88*, *Mks1* and *Rpgr11*. Statistical significance of comparisons between individual knockdown experiments and negative control (Scrambled) is as follows: *lft80* – 0.031 (*), *lft88* – 0.002 (**), *Kif3a* – 0.121 (ns), *Lca5* – 0.108 (ns), *Mks1* – 0.013 (*), *Rpgr11* – 0.036 (*), *Tmem67* – 0.654 (ns), *Ttc21b* – 0.113 (ns), *Pla2g3* – 0.669 (ns); ns, not significant; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; **** $p < 0.0001$, for Student's t-test (unpaired, two-tailed). Error bars indicate s.d. for n=3 biological replicates.

Knockdown of these genes showed a variable degree of cilia loss (**Figure 3-34**). Although most of these genes had shown reductions in cilia number after knockdown in other studies, incubation times and reagent origin have differed to variable extents. Three positive controls were chosen, with pooled siRNA duplexes targeting three ciliary genes: *Mks1*, *Rpgr11* and *lft88* (example of *silft88* shown in **Figure 3-35**). Mutations in all of these genes are known to cause a ciliopathy phenotype in human patients or mouse models, and knockdowns of their transcripts have shown variable but significant cilia losses that accord well with the expected cellular phenotype.

The negative control siRNAs included a non-targeting or 'scrambled' siRNA (**Figure 3-35**), an siRNA targeted against human motilin receptor (*MLNR*) that has no sequence homology in the mouse genome, and a mock transfection control (transfection reagent only). These controls had no significant effect on the percentage of cells with a single cilium, as previously reported^{99,205} (**Figure 3-36b and d**).

In addition to assay controls, a transfection control was required for visual monitoring of transfection efficiency during high-throughput screening. Since the image analysis protocol automatically evaluated the number of nuclei within an image, this was used as a surrogate value for cell number. Following knockdown of *Plk1*, a gene essential for cell growth and proliferation²⁶⁶, >95% of cells were lost

when compared to non-targeting siRNA indicating a very high transfection efficiency (Figure 3-36a and c).

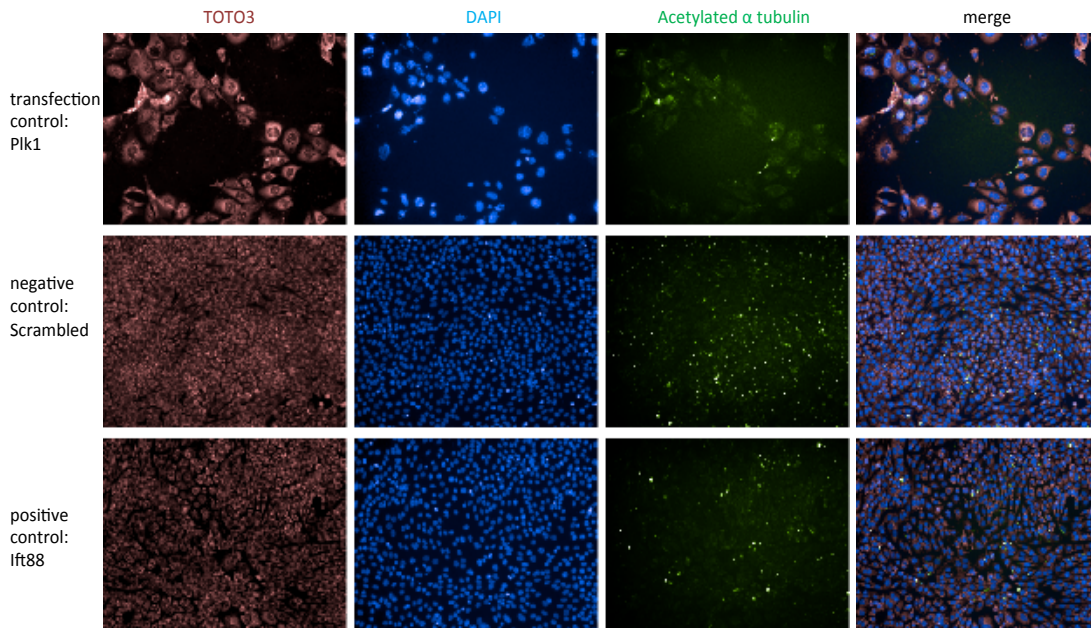


Figure 3-35. Columbus images for siRNA screen controls. Top panel shows transfection efficiency control *siPlk1* with striking cell loss (DAPI, blue and TOTO3, red staining) compared to negative control (*siScrambled*, middle panel). The positive control for cilia loss (*siIft88*, bottom panel) shows fewer cilia (acetylated α -tubulin, green) than the negative control.

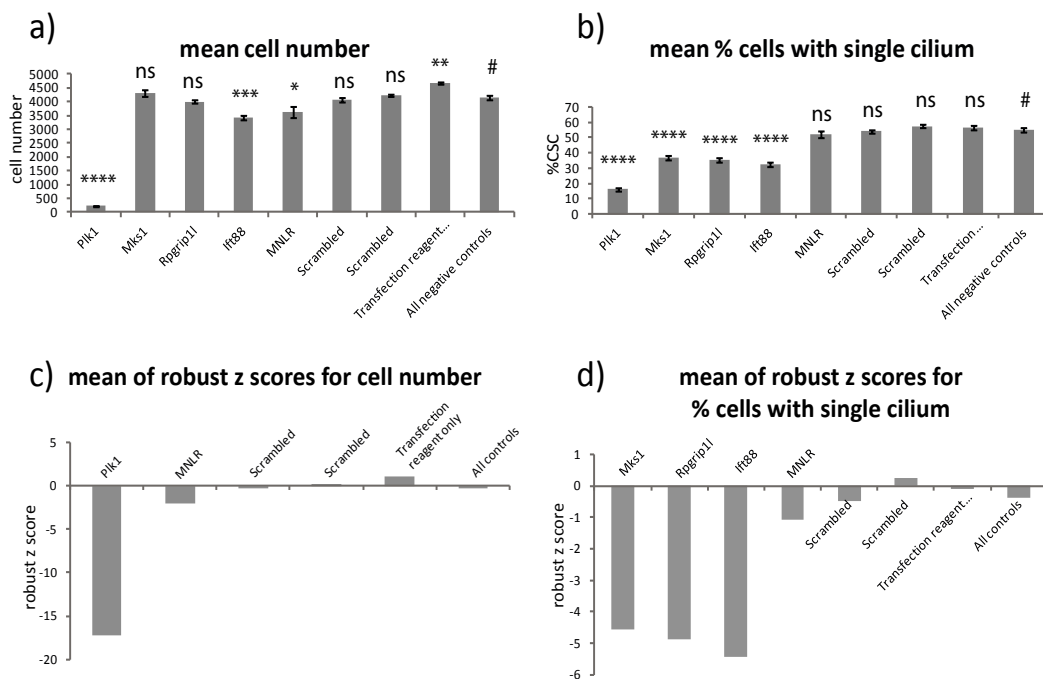


Figure 3-36. Bar graphs showing efficiency of knockdowns for screen controls. a) mean cell number in a series of knockdowns showing highly significant cell loss after *Plk1* knockdown, used in this study as a transfection efficiency indicator. There was moderate

cell loss observed after *Ift88* and *MLNR* knockdowns, and increased cell number after treatment with transfection reagent compared to all negative controls. b) mean % of cells with single cilium showing efficient cilia loss after knockdown of *Plk1*, *Mks1*, *Rpgrip1l* and *Ift88*. %CSC was not affected in cells treated with si*MLNR*, siScrambled and transfection reagent only. c) mean of robust z scores for cell number showing a low value for si*Plk1*. d) mean of robust z scores for %CSC showing high scores for si*Mks1*, si*Rpgrip1l* and si*Ift88*. Statistical significance of pair-wise comparisons between individual knockdown experiments and all negative controls (#) is indicated by: ns, not significant; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; **** $p < 0.0001$, for Student's t-test (paired, two-tailed). Error bars indicate s.d. for n=32 biological replicates.

A test assay of 128 positive control and 128 negative control knockdowns across 16 plates in two separate batches (each control repeated 32 times) showed a highly reproducible, statistically significant effect of *Plk1* knockdown on cell number (mean cell number 226.4) compared to all negative controls (mean cell number 4094.7), $p < 0.0001$ (paired two-tailed Student's t-test) (**Figure 3-36a**). *Plk1*, *Mks1*, *Rpgrip1l* and *Ift88* knockdowns were found to have highly reproducible, statistically significant effects on the percentage of cells with a single cilium (16.2, 36.9, 35.2 and 32.4% respectively) compared to all negative controls (55.0%), $p < 0.0001$, two-tailed Student's t-test in all cases (**Figure 3-36b**). However, since *Plk1* siRNA had a significant effect on cell number (**Figure 3-36a** and **c**), the cilia loss observed (**Figure 3-36b** and **d**) is a secondary effect due to loss of cells. This result indicates a potential source of false positive hits in the discovery phase of the whole genome screen that needs to be filtered from subsequent analyses and hit validation.

The mean of robust z scores for *Plk1* knockdown, quantitated for cell number, was < -17 (**Figure 3-36c**). All negative control siRNAs for cell number effects had robust z scores between -2 and 1 (**Figure 3-36c**). The mean of all robust z scores for positive controls, quantitated for the percentage cells with a single cilium, gave a value of < -2 (**Figure 3-36d**), indicating the suitability of these siRNAs as positive controls for effects on ciliogenesis. The individual mean robust z scores, for the %CSC, was $z = -4.48$ after *Mks1* knockdown, $z = -4.79$ for *Rpgrip1l* and $z = -5.39$ for *Ift88*. Calculation of robust z scores allowed to set robust and meaningful statistical cut-offs for identification of significant 'hits' affecting the %CSC. These cut-offs were based on the median and median absolute deviation of positive controls per ten plate batch ²⁶⁷.

3.3.2 Screen results

Robust z scores for WCN and %CSC were calculated for all results, using the median and median absolute deviation for all biological replicates of negative

controls (siScrambled n=40, siMLNR n=20, mock transfection n=20 in a ten plate batch).

3.3.2.1 Quality control metrics and overall performance of the whole genome screen

The frequency of the z score distribution for the whole genome screen was analysed and is shown in **Figure 3-37** for %CSC and **Figure 3-38** for WCN for the two runs.

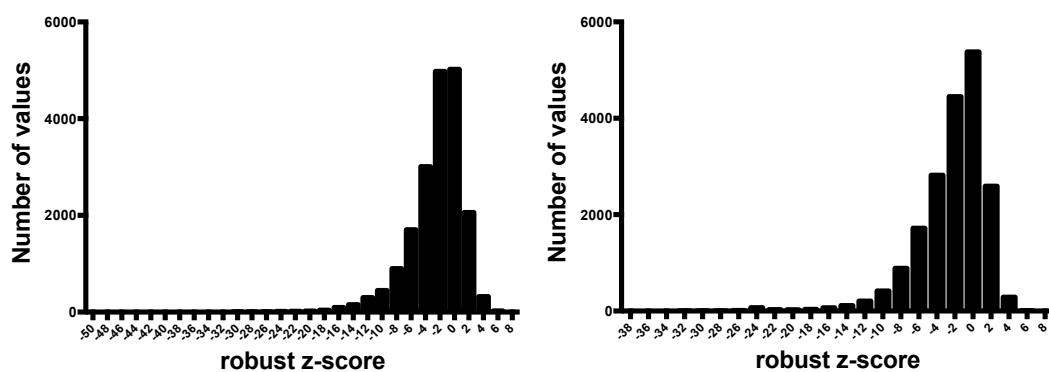


Figure 3-37. Histograms representing frequency of %CSC in run 1 (left histogram) and run 2 (right histogram) for the whole genome screen. The x axis represents bins for robust z scores of %CSC.

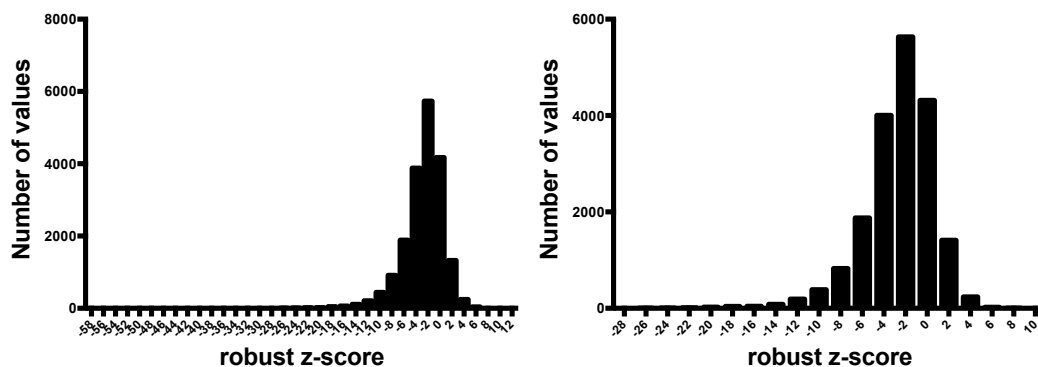


Figure 3-38 Histograms representing frequency of WCN in run 1 (left histogram) and run 2 (right histogram). The x axis represents bins for robust z score for WCN.

Robust z scores from both runs of the screen for WCN and %CSC showed skewed (non-symmetric) distribution of negative values (**Figure 3-37** and **3-38**). This may be a result of start-up effects due to initial failures caused by difficulties in scaling up the set up protocol or the design of the screen when loss of cilia was of main interest, therefore any gain in cilia/cell number would be less sensitive.

To visualise overall screen performance, robust z scores for %CSC from one run of a ten plate batch were plotted (**Figure 3-39**). Controls from each plate were coloured and clear separation of positive (green) and negative (purple) controls was visible (**Figure 3-39**). Data from the transfection control (si*Plk1*) is represented in red.

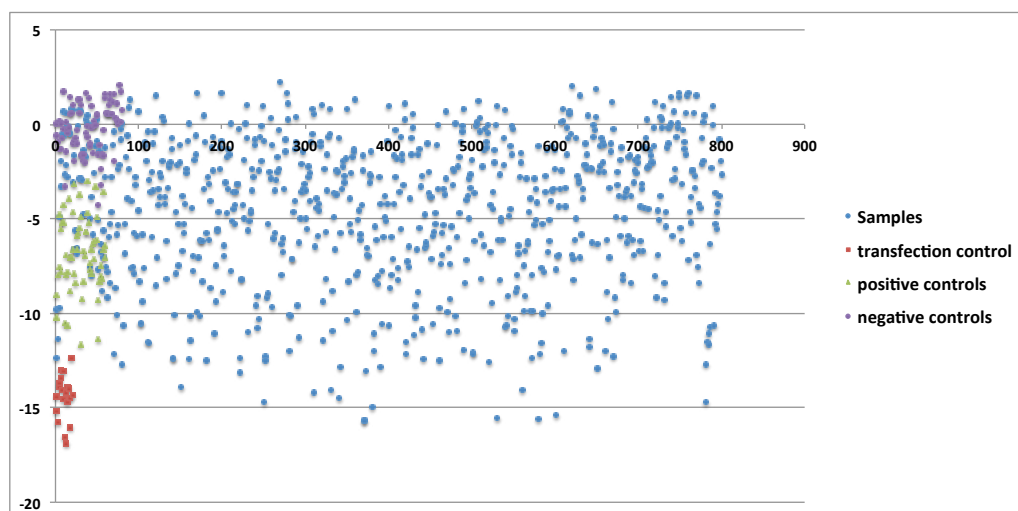


Figure 3-39. Scatter plot representing robust z scores for %CSC from a ten plate batch single run. Negative control values are shown in purple, positive controls in green and si*Plk1* transfection control in red. Clear separation between positive and negative controls is evident. Experimental samples are shown in blue. The x axis represents robust z score, while y axis represents sample number (the 10 plate batch investigated 80 target siRNAs).

Robust z scores for %CSC of ten plate batch for two biological replicates were plotted against each other and showed good correlation (Pearson coefficient of correlation, $R^2=0.60266$) (**Figure 3-40**).

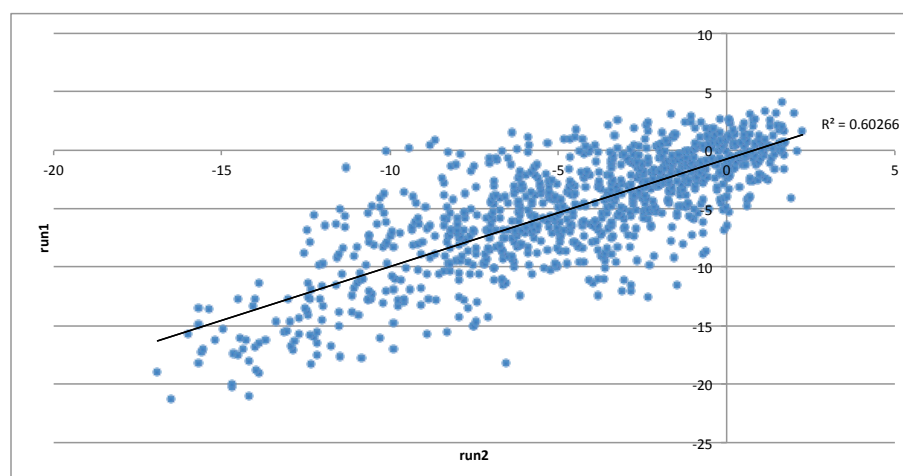


Figure 3-40. Correlation plot between two runs representing robust z scores for %CSC. x and y axes show values of the robust z scores. The black line indicates the trendline between the two runs with a good correlation score (Pearson coefficient of correlation, $R^2=0.60266$).

3.3.2.2 Quality metrics

The data illustrate a unique, efficient and reliable cilia detection algorithm, and highly reproducible positive controls that enable statistical verification of their efficiency and suitability in this screen. However, false positives and false negatives results are common in siRNA screens and need to be closely controlled.

3.3.2.2.1 Off-target effects

Potential sources of false positive hits are sequence-specific off-target effects (OTEs) due to the targeting of more than one transcript in the genome by siRNAs. To account for this, BLASTN alignments were used to identify all individual siRNAs duplexes with potential OTEs. BLASTN analysis of every siRNA in the Dharmacon siGENOME SMARTpool mouse siRNA library identified 832 siRNAs which had potential sequence-specific off-target effects (**Appendix 7**), and these were subsequently excluded from further analysis. These alignments also confirmed that the remaining siRNAs used in the whole genome screen had specific homology to the target transcripts.

3.3.2.2.2 MicroRNA – like effects

To test the possibility that siRNAs in the Dharmacon siGENOME library may have microRNA-like effects, leading to further off-target effects, genome-wide enrichment of seed sequence matches (GESS) analysis was carried out ²⁶⁸. This tests for 7-mer seed sequence complementarity to 3'UTR sequences in a mouse genome database. GESS analysis of siRNA sequences showed that none of the seed regions for the siRNAs duplexes were likely to cause off-target effects.

3.3.2.2.3 Partial on-target effects

Potential false negative results may arise from incomplete targeting of all alternative transcripts of particular mRNA by the four siRNAs in each Dharmacon siGENOME siRNA pool. BLASTN alignments of siRNAs sequences identified 2030 siRNAs that only partially targeted their gene transcripts (**Appendix 8**), with at least one transcript not being targeted. These siRNAs were not excluded from further screen analysis, but this data provides the basis for potential weighting of results in further analyses. BLASTN analysis confirmed that the *Plk1*, *Mks1*, *Rpgrip1l* and *Iff88* siRNA pools targeted every Ensembl annotated and protein coding alternative transcript of the target genes, with the exception of one incomplete *Rpgrip1l* transcript (ENSMUST00000132757) encoding a predicted 78 amino acid isoform of the protein.

3.3.2.2.4 Array CGH

False negatives may arise due to imbalances in genetic copy number in the cell-line used. To control for this, array CGH was carried out in IMCD3 cells. Cells were treated to best represent their state in the screen set up. Samples were prepared by Dr Gabrielle Wheway and run and analysed by Dr Grischa Toedt in EMBL, Heidelberg, Germany. 5274 genes were identified with abnormal copy number (**Figure 3-41**), including 3 genes with total biallelic loss in IMCD3 cells (*Olf1r543*, *Igf1r*, *Cd244*) (**Appendix 9**). The three genes with biallelic deletion were excluded from further studies. Abnormal copy number data provides the basis for potential weighting of results in further analyses.

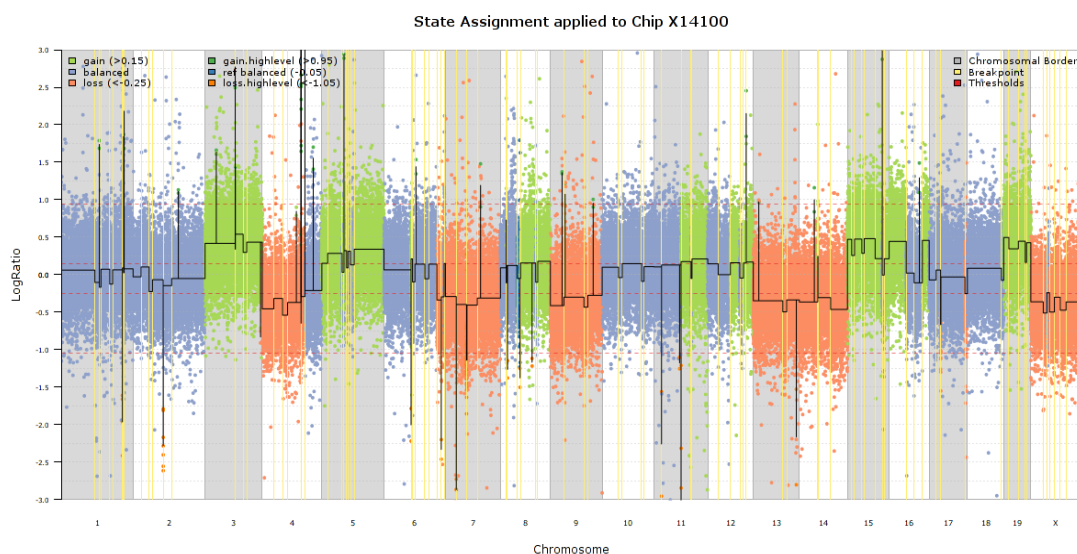


Figure 3-41. Array CGH analysis of mIMCD3 cell line. Copy number analysis shows extensive genetic imbalance in the IMCD3 cell-line. Blue indicates genetically balanced regions, green indicates allelic gain and orange indicates allelic loss, black line shows trendline, yellow – breakpoints, red hatched lines indicate thresholds, grey – chromosomal borders.

3.3.2.2.5 RNA sequencing of IMCD3 transcripts

Variations in expression levels of different transcripts could be a potential source of false positive or false negative hits. Expression levels of transcripts in IMCD3 cells were determined by RNA sequencing. Dr Gabrielle Wheway prepared sample and analysed the data. Samples were sequenced in the EMBL Genomics Facility in Heidelberg by collaborators Dr Grischa Toedt and Dr Toby Gibson.

For this analysis, the cells were plated and treated in conditions identical to those for negative control knockdowns. RNA was extracted using standard protocols (section 2.2.4). Total RNA was purified and paired-end library was prepared following manufacturer’s protocols (Zymo and Agilent Technologies). Prepared libraries were run on an Illumina HiSeq 2500 clonal sequencer. Sequences were aligned using the “splice aware” TopHat2 algorithm^{269,270} with the gene reference .gff files downloaded for the Ensembl mouse genome mm19 build.

Transcript assembly and differential transcript expression was analysed using Cufflinks or Cuffdiff algorithms ²⁷¹. FPKM (fragments per kilobase of exon per million fragments mapped) were calculated for all alternative transcripts listed in the Cuffdiff database, representing levels of expressed transcript (example in **Table 3-18**).

In summary, 96,253 transcripts were aligned and 72,898 had specific UniGene IDs. This covered a total of 21,488 mouse genes, out of which 3,893 had at least one transcript not expressed and 7,150 were not expressed in IMCD3s and were excluded from further studies.

3.3.2.2.6 Sequence-independent off-target effects

Other possible sources of false positive and negative hits were sequence-independent off-target effects due to siRNA competition with, or saturation of, endogenous RNAi machinery components, especially if the cell-line used has under-expressed rate-limiting components of RNAi such as exportin-5, TRBP, Dicer and Ago1-4 ²⁷²⁻²⁷⁵. To control for these effects, the normal copy number of rate-limiting components of RNAi in IMCD3 cells was confirmed by arrayCGH. *Xpo5* had normal copy number, *Tarbp2* +1 gain and *Dicer1* +1 gain (**Table 3-18**). There were no appropriate SNP markers for *Eif2c1* so there is no copy number information for this gene, but both *Eif2c3* and *Eif2c4* are adjacent on mouse chromosome 4 and had a normal copy number therefore *Eif2c1* is likely to have a normal copy number. *Eif2c2* had a +1 gain and both *Eif2c3* and *Eif2c4* were normal. Importantly, there were no genomic losses in the genes encoding these crucial parts of RNAi machinery. RNAseq of the IMCD3 cell line confirmed these results (**Table 3-18**). With the exception of transcript NM_001253795 for the *Tarbp2* gene that was not detected, all remaining transcripts of crucial RNAi machinery components were expressed at detectable levels. Robust negative controls including 'scrambled' controls (eight negative controls per plate) were also used to control for any sequence-independent off-target effects.

REF. GENE ID	TRANSCRIPT ID	CUFFDIFF ID	FPKM	ARRAY CGH
Xpo5	NM_028198	NM_028198	15.453292	0
Tarbp2	NM_001253795	NM_001253795	0	1
Tarbp2	NM_009319	NM_009319	35.585083	1
Dicer1	NM_148948	CUFF.20614.1	2.421229	1
Dicer1	NM_148948	NM_148948	7.212395	1
Eif2c1	NM_153403	CUFF.61275.1	0.20074	-
Eif2c1	NM_153403	CUFF.61277.1	0.963916	-
Eif2c1	NM_153403	CUFF.61276.1	0.974174	-
Eif2c1	NM_153403	CUFF.61278.1	2.229429	-
Eif2c1	NM_153403	NM_153403	5.181405	-
Eif2c2	NM_153178	CUFF.29857.1	8.447386	1
Eif2c3	NM_153402	CUFF.61252.1	0.154763	0
Eif2c3	NM_153402	CUFF.61252.2	0.344294	0
Eif2c3	NM_153402	CUFF.61269.1	0.382929	0
Eif2c3	NM_153402	CUFF.61268.1	0.487889	0
Eif2c3	NM_153402	CUFF.61259.1	0.64435	0
Eif2c3	NM_153402	CUFF.61258.1	0.714686	0
Eif2c3	NM_153402	CUFF.61266.1	0.958926	0
Eif2c3	NM_153402	CUFF.61263.1	1.06833	0
Eif2c3	NM_153402	CUFF.61256.1	1.111508	0
Eif2c3	NM_153402	CUFF.61273.1	1.180859	0
Eif2c3	NM_153402	CUFF.61251.1	1.904965	0
Eif2c3	NM_153402	CUFF.61267.1	2.462731	0
Eif2c3	NM_153402	CUFF.61253.1	3.059563	0
Eif2c3	NM_153402	CUFF.61262.1	3.882692	0
Eif2c3	NM_153402	CUFF.61255.1	5.060067	0
Eif2c3	NM_153402	CUFF.61260.1	23.023834	0
Eif2c3	NM_153402	CUFF.61257.1	31.806826	0
Eif2c3	NM_153402	CUFF.61261.1	33.267184	0
Eif2c3	NM_153402	CUFF.61264.1	39.290785	0
Eif2c3	NM_153402	CUFF.61270.1	41.785199	0
Eif2c3	NM_153402	CUFF.61265.1	351889.8994	0
Eif2c3	NM_153402	CUFF.61272.1	351889.8994	0
Eif2c3	NM_153402	CUFF.61254.1	422267.8793	0
Eif2c3	NM_153402	CUFF.61271.1	774157.7786	0
Eif2c4	NM_153177	NM_153177	0.014123	0
Eif2c4	NM_153177	CUFF.61279.3	0.023801	0
Eif2c4	NM_153177	CUFF.61279.2	0.077269	0

Table 3-18. Summary of expression levels of RISC (RNA Induced Silencing Complex) complex components. All canonical and non-canonical transcripts had RNA levels verified by RNAseq and are represented by scores in the FPKM column. The higher the number the higher the expression level of the transcript. Results from the array CGH analysis are also shown: 0 indicates genomic stability (2 alleles per gene), 1 indicates allelic gain, “-“ no data available.

3.3.2.2.7 Measurements of assay robustness

Analyses of signal-to-background ratio, coefficient of variation and Z'-factor have shown in previous studies that siRNA reverse genetics screens are less robust than small molecule screens²⁷⁶ using these standard measurements. To account for the limitations of controls with moderate effect, the strictly standardised mean of differences (SSMD) is an alternative metric often used to assess the performance of siRNA screens^{265,277}. SSMD is a ratio between the difference of means and the standard deviation of the difference between two populations (the negative and positive control groups):

$$SSMD = (\mu_P - \mu_N) / \sqrt{(\sigma_P^2 + \sigma_N^2)}$$

μ_P – mean of positive controls

μ_N – mean of negative controls

σ_P – standard deviation of positive controls

σ_N – standard deviation of negative controls

Within the first batch of controls, the SSMD of cell number in *Plk1* knockdown and all negative controls was 7.93, and in the second batch of controls the SSMD was 8.34 (the higher the number the better control), reflecting the suitability of *Plk1* as a positive control for effects on cell number. The SSMD of the percentage of cells with a single cilium in *Mks1*, *Ift88* and *Rpgrip1l* knockdowns, compared to all negative controls, was >1 for most duplicate batches of controls (**Figure 3-42**) with an average SSMD value for all batches of 1.717 (**Figure 3-42**), indicating an acceptable consistency and quality for the screen

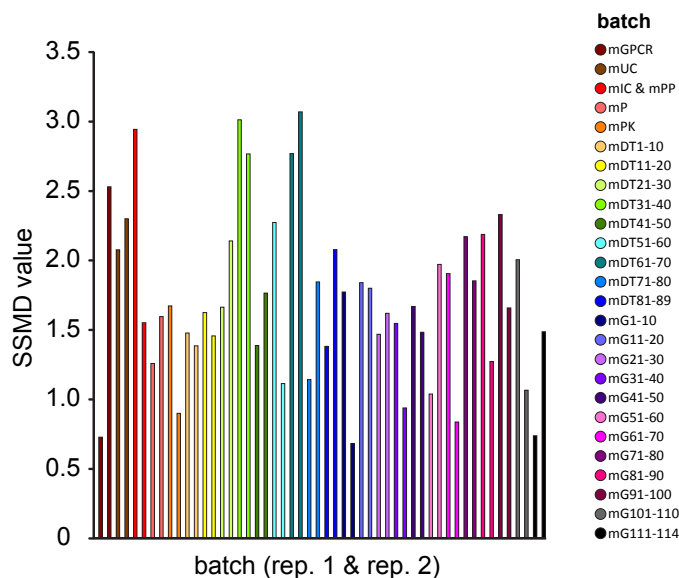


Figure 3-42. SSMD values for all batches in the whole genome siRNA screen. Each batch (ten or less plates) was assigned a colour which is used for all batch duplicates in the screen.

3.3.2.3 Positive hits

3.3.2.3.1 Primary screen

243 96-well plates containing 19,097 siRNA duplexes targeting 18,690 Entrez RefSeq mouse genes were screened in duplicate. Each plate contained two transfection efficiency controls (*siPlk1*), six positive controls for a ciliogenesis defect (*siMks1*, *siRpgrip1l* and *siIft88*) and eight negative controls (four siScrambled, two *siMLNR* and two mock transfections).

3.3.2.3.1.1 Data filtering – first pass

All results for WCN and %CSC were fed into a template spreadsheet that had a macro for robust z scores calculations in ten plate batches for both replicates. This file was subsequently uploaded into a Perl script called 'Dharmascript' (written by Dr David Parry) that calculated robust z scores, organised results based on the average highest effect on %CSC, and filtered for hits with a robust z score greater than the median robust z score of the positive controls in a ten plate batch for both runs of %CSC.

A list of 2,174 genes was obtained after this filtering method, which comprises a 11.38% hit rate. siRNA duplexes targeting 133 genes from this list had potential off-target effects and were excluded from further studies. 288 genes had only partial on-target effects. There were no hits with biallelic copy number loss for any gene, (according to the results from array CGH). 1,956 genes had human homologues and mutations in 360 of these are known to be implicated in or causative of a Mendelian condition (OMIM, July 2013). Interestingly, most of those conditions are not known ciliopathies. 369 of the hits were listed in the CiliaProteome v3 database. The list of "hit" genes was submitted to the DAVID program (an on-line functional annotation and enrichment tool). Based on the Swiss-Prot and Protein Information Resource keywords, the terms with the most statistically significant enrichment were 'ribonucleoprotein', 'ribosomal protein', 'acetylation' and 'mRNA splicing' (**Figure 3-43**). There was also significant enrichment for the Gene Ontology terms 'ribonucleoprotein complex', 'ribosome' and 'spliceosome' (**Figure 3-44**). A KEGG pathway (Kyoto Encyclopedia of Genes and Genomes) search showed enrichment for three functional modules: ribosome, spliceosome and proteasome (**Figure 3-45**). **Figures 3-46, -47 and -48** illustrate these modules and red stars highlight the position of hits from the screen. None of those terms have been associated previously with ciliary processes, which indicates either a high false positive rate or new functional modules that are important for ciliogenesis.

Sublist	Category	Term	RT	Genes	Count	%	P-Value	Benjamini
SP_PIR_KEYWORDS		ribonucleoprotein	RT		85	7.5	4.0E-40	1.6E-37
SP_PIR_KEYWORDS		ribosomal protein	RT		61	5.4	7.7E-31	1.5E-28
SP_PIR_KEYWORDS		acetylation	RT		255	22.6	1.3E-26	1.8E-24
SP_PIR_KEYWORDS		mRNA splicing	RT		55	4.9	1.4E-23	1.4E-21
SP_PIR_KEYWORDS		Spliceosome	RT		42	3.7	4.2E-23	3.3E-21
SP_PIR_KEYWORDS		mRNA processing	RT		55	4.9	1.3E-18	8.5E-17
SP_PIR_KEYWORDS		ribosome	RT		17	1.5	1.9E-15	1.1E-13
SP_PIR_KEYWORDS		protein biosynthesis	RT		31	2.7	2.1E-10	1.0E-8
SP_PIR_KEYWORDS		mRNA-binding	RT		65	5.8	3.0E-10	1.3E-8
SP_PIR_KEYWORDS		nucleus	RT		298	26.4	9.6E-10	3.8E-8
SP_PIR_KEYWORDS		phosphoprotein	RT		435	38.5	6.0E-7	2.1E-5
SP_PIR_KEYWORDS		mRNA transport	RT		16	1.4	9.6E-7	3.1E-5
SP_PIR_KEYWORDS		nuclear pore complex	RT		10	0.9	4.0E-5	1.2E-3
SP_PIR_KEYWORDS		viral nucleoprotein	RT		9	0.8	9.2E-5	2.6E-3
SP_PIR_KEYWORDS		mRNA-binding	RT		7	0.6	1.1E-4	2.9E-3
SP_PIR_KEYWORDS		translocation	RT		13	1.2	2.7E-4	6.7E-3
SP_PIR_KEYWORDS		protein transport	RT		45	4.0	5.9E-4	1.4E-2
SP_PIR_KEYWORDS		ribosome biogenesis	RT		10	0.9	8.6E-4	1.9E-2
SP_PIR_KEYWORDS		tpr repeat	RT		19	1.7	1.1E-3	2.3E-2
SP_PIR_KEYWORDS		lipid-binding	RT		13	1.2	5.5E-3	1.0E-1
SP_PIR_KEYWORDS		mRNA processing	RT		10	0.9	6.0E-3	1.1E-1
SP_PIR_KEYWORDS		isopeptide bond	RT		27	2.4	8.9E-3	1.5E-1
SP_PIR_KEYWORDS		DNA-binding	RT		101	8.9	1.2E-2	1.9E-1
SP_PIR_KEYWORDS		nonsense-mediated mRNA decay	RT		6	0.5	1.2E-2	1.8E-1
SP_PIR_KEYWORDS		Initiation factor	RT		9	0.8	1.2E-2	1.8E-1
SP_PIR_KEYWORDS		mitosis	RT		19	1.7	1.4E-2	1.9E-1
SP_PIR_KEYWORDS		transcription regulation	RT		109	9.7	1.6E-2	2.0E-1
SP_PIR_KEYWORDS		repressor	RT		34	3.0	1.6E-2	2.0E-1
SP_PIR_KEYWORDS		Transcription	RT		122	10.8	2.0E-2	2.4E-1
SP_PIR_KEYWORDS		cell division	RT		24	2.1	2.1E-2	2.4E-1
SP_PIR_KEYWORDS		cytoplasmic vesicle	RT		22	1.9	2.5E-2	2.8E-1

Figure 3-43. DAVID analysis using Swiss-Prot and Protein Information Resource keywords on the 2,174 primary screen hits. Listed are enriched terms with associated P values and Benjamini-Hochberg q values (to control for the false discovery rate due to multiple testing). Terms with $q < 0.05$ are listed. Image modified from <http://david.abcc.ncifcrf.gov/>.

Sublist	Category	Term	RT	Genes	Count	%	P-Value	Benjamini
GOTERM_CC_FAT		ribonucleoprotein complex	RT		126	11.2	2.3E-50	9.8E-48
GOTERM_CC_FAT		ribosome	RT		61	5.4	5.8E-28	1.2E-25
GOTERM_CC_FAT		spliceosome	RT		44	3.9	2.3E-22	3.3E-20
GOTERM_CC_FAT		intracellular non-membrane-bounded organelle	RT		195	17.3	3.7E-15	3.9E-13
GOTERM_CC_FAT		non-membrane-bounded organelle	RT		195	17.3	3.7E-15	3.9E-13
GOTERM_CC_FAT		cytosolic ribosome	RT		13	1.2	2.4E-11	2.0E-9
GOTERM_CC_FAT		nuclear lumen	RT		94	8.3	2.7E-8	1.9E-6
GOTERM_CC_FAT		ribosomal subunit	RT		19	1.7	3.2E-8	1.9E-6
GOTERM_CC_FAT		small ribosomal subunit	RT		12	1.1	3.9E-7	2.1E-5
GOTERM_CC_FAT		intracellular organelle lumen	RT		107	9.5	1.2E-6	5.5E-5
GOTERM_CC_FAT		organelle lumen	RT		107	9.5	1.3E-6	5.7E-5
GOTERM_CC_FAT		membrane-enclosed lumen	RT		108	9.6	3.6E-6	1.4E-4
GOTERM_CC_FAT		nuclear pore	RT		15	1.3	6.6E-6	2.3E-4
GOTERM_CC_FAT		nucleolus	RT		38	3.4	3.9E-5	1.3E-3
GOTERM_CC_FAT		pore complex	RT		15	1.3	8.4E-5	2.6E-3
GOTERM_CC_FAT		nucleoplasm	RT		59	5.2	1.5E-4	4.2E-3
GOTERM_CC_FAT		cytosolic small ribosomal subunit	RT		5	0.4	1.7E-4	4.5E-3
GOTERM_CC_FAT		nuclear speck	RT		16	1.4	2.5E-4	6.2E-3
GOTERM_CC_FAT		nucleoplasm part	RT		51	4.5	3.6E-4	8.4E-3
GOTERM_CC_FAT		cytosolic large ribosomal subunit	RT		5	0.4	3.7E-4	8.3E-3
GOTERM_CC_FAT		eukaryotic translation initiation factor 3 complex	RT		6	0.5	4.1E-4	8.6E-3
GOTERM_CC_FAT		nuclear body	RT		20	1.8	7.3E-4	1.5E-2
GOTERM_CC_FAT		cytosolic part	RT		12	1.1	1.1E-3	2.1E-2
GOTERM_CC_FAT		microtubule cytoskeleton	RT		44	3.9	1.4E-3	2.5E-2
GOTERM_CC_FAT		spindle microtubule	RT		6	0.5	2.5E-3	4.3E-2
GOTERM_CC_FAT		nuclear envelope	RT		19	1.7	4.2E-3	7.0E-2
GOTERM_CC_FAT		endomembrane system	RT		48	4.3	4.4E-3	7.0E-2
GOTERM_CC_FAT		microtubule	RT		24	2.1	1.6E-2	2.2E-1
GOTERM_CC_FAT		coated membrane	RT		9	0.8	2.1E-2	2.7E-1
GOTERM_CC_FAT		membrane coat	RT		9	0.8	2.1E-2	2.7E-1
GOTERM_CC_FAT		large ribosomal subunit	RT		7	0.6	2.6E-2	3.2E-1

Figure 3-44. DAVID analysis using Gene Ontology term search on the 2,174 primary screen hits. Listed are enriched terms with associated *P* values and Benjamini-Hochberg *q* values (to control for the false discovery rate due to multiple testing). Terms with *q* < 0.05 are listed. Image modified from <http://david.abcc.ncifcrf.gov/>.

Sublist	Category	Term	RT	Genes	Count	%	P-Value	Benjamini
KEGG_PATHWAY		Ribosome	RT		56	5.0	1.5E-44	2.5E-42
KEGG_PATHWAY		Spliceosome	RT		50	4.4	1.2E-27	9.4E-26
KEGG_PATHWAY		Proteasome	RT		7	0.6	7.5E-2	9.8E-1

Figure 3-45. DAVID analysis using KEGG pathway term search on the 2,174 primary screen hits. Listed are enriched terms with associated *P* values and Benjamini-Hochberg *q* values (to control for the false discovery rate due to multiple testing). Terms with *q* < 0.05 are listed. Image modified from <http://david.abcc.ncifcrf.gov/>.

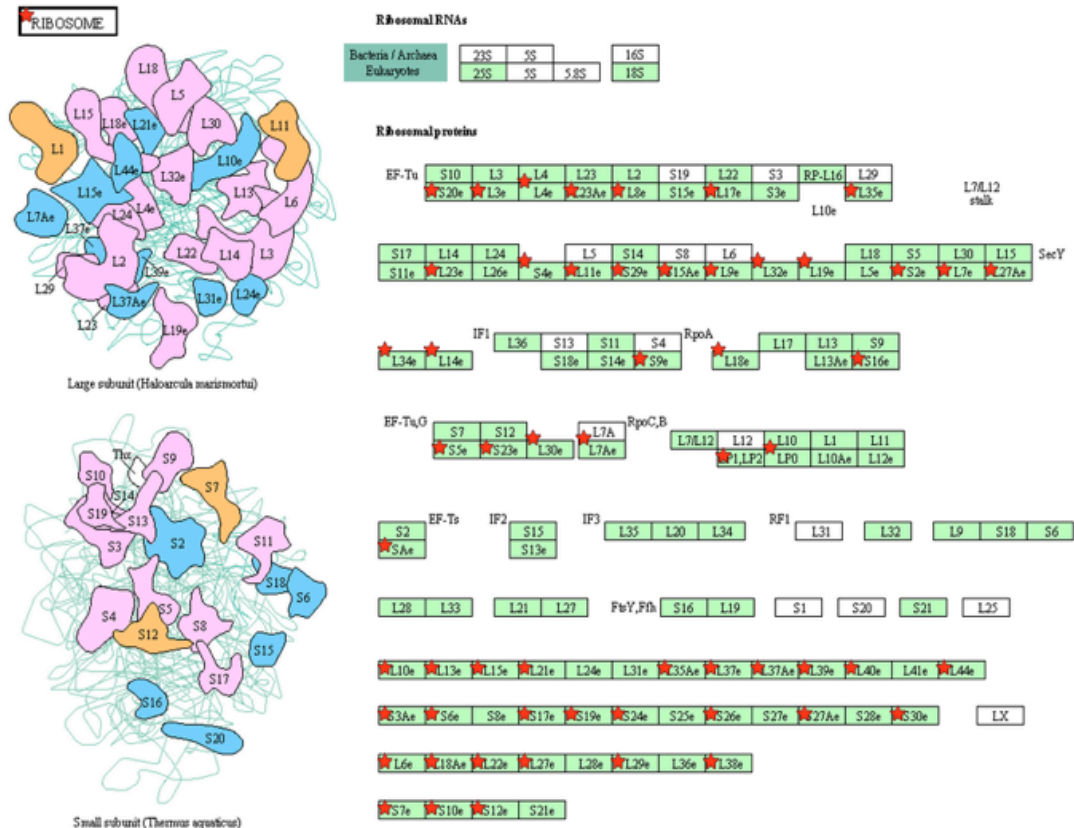


Figure 3-46. Cartoon illustrating the enrichment of hits from the primary screen in the ribosome. Red stars indicate proteins encoded by genes that were hits in the screen or interacting partners for those proteins. Image modified from <http://david.abcc.ncifcrf.gov/>.

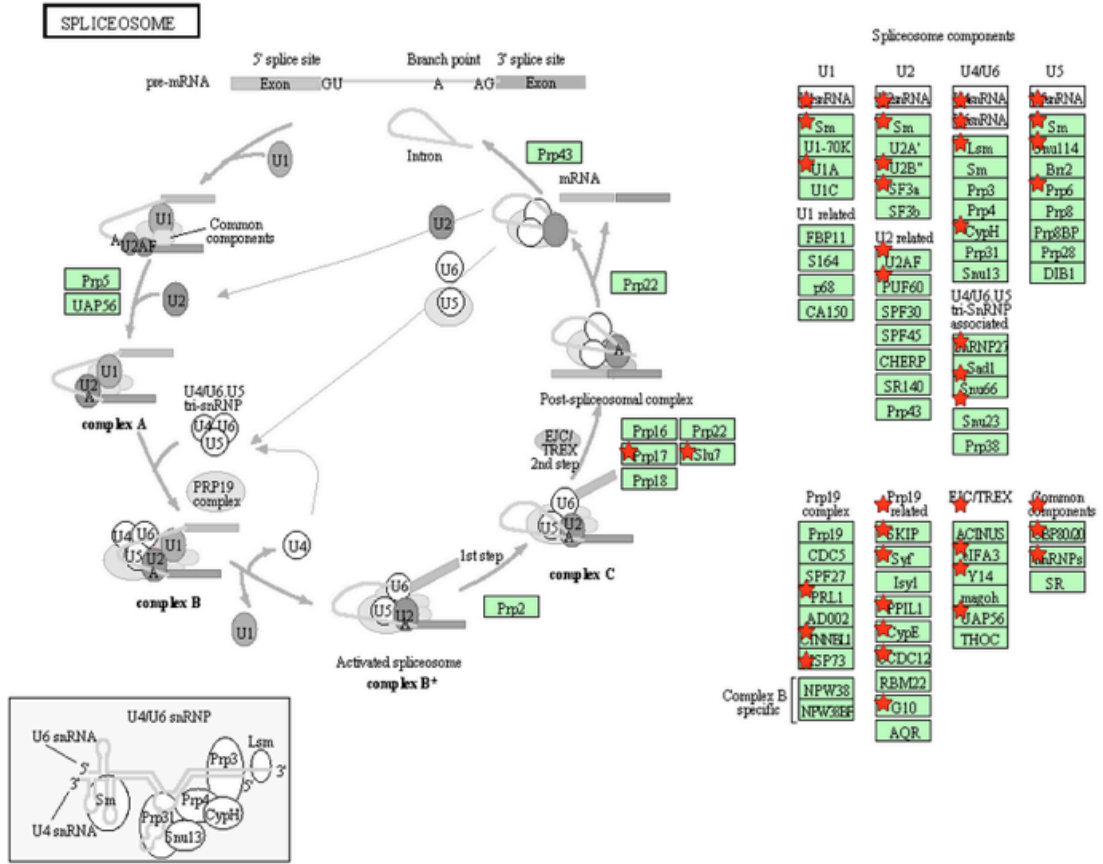


Figure 3-47. Cartoon illustrating the enrichment of hits from the primary screen in the spliceosome. Red stars indicate proteins encoded by genes that were hits in the screen or interacting partners for those proteins. Image modified from <http://david.abcc.ncifcrf.gov/>.

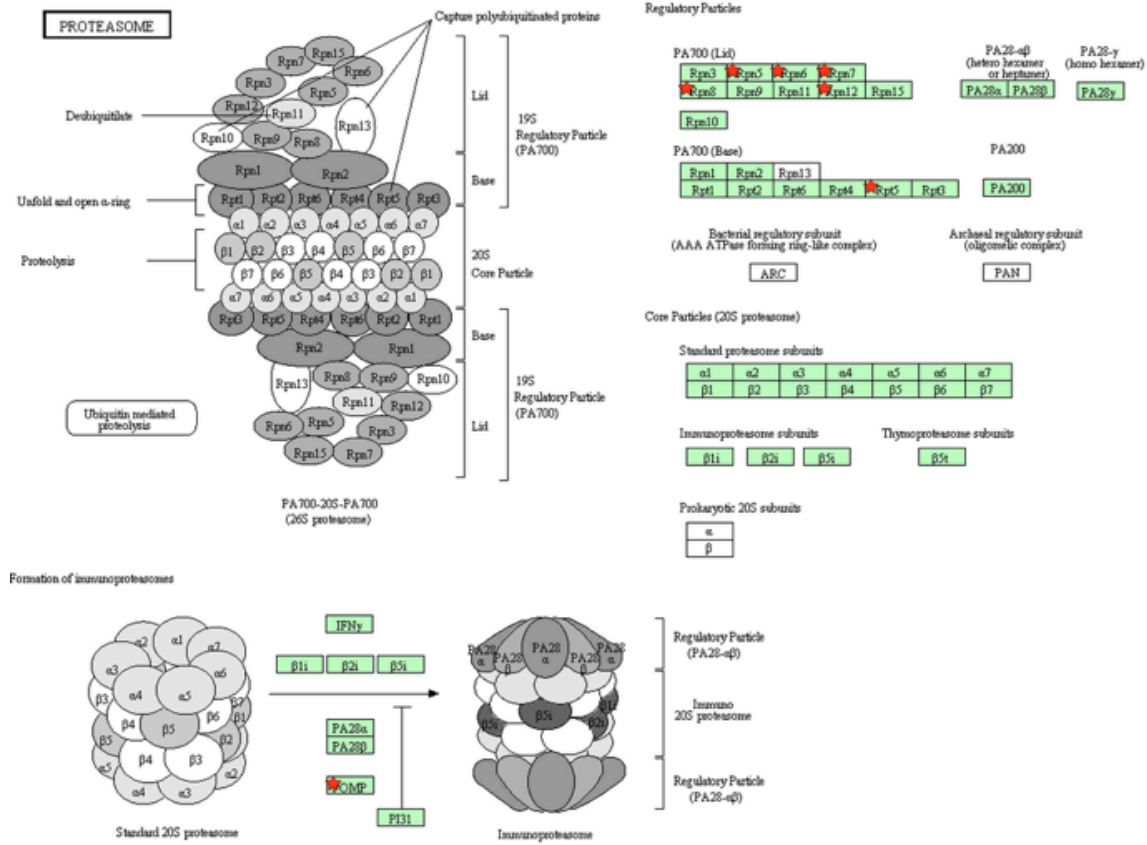


Figure 3-48. Cartoon illustrating the enrichment of hits from the primary screen in the proteasome. Red stars indicate proteins encoded by genes that were hits in the screen or interacting partners for those proteins. Image modified from <http://david.abcc.ncicrf.gov/>.

3.3.2.3.1.2 Data filtering: second pass

Although 11.38% was an acceptable hit rate for the whole genome screen, further filters needed to be applied to reduce the number of hits to those with the highest confidence. As the %CSC filter was shown to be efficient, effects on WCN were assessed as a potential additional filter, since many siRNAs in the primary screen had an effect on cell number. A decreasing cell number may be caused by either cell death or a delay in the length of the cell cycle. If the cell cycle is delayed, then cilia formation may be subsequently delayed and the number of cilia may be under-represented in knockdowns of those specific genes. This could be an important cause of potential false positive results due to non-specific effects that could cause defects in ciliogenesis due to secondary processes. For this reason, a stringent WCN filter was therefore applied. All hits were removed if their robust z score was lower than the median robust z score - 2MAD of the positive controls in each ten plate batch for both runs. MAD was applied to increase the stringency of the cell number cut-off based on positive controls. In particular, *si/ft88* affected cell

number compared to the negative controls (**Figure 3-35a**) and the -2MAD filter removed this effect.

This filtering strategy resulted in 174 hits. Out of these 154 had human orthologues, 8 had off-target effects, 17 were only partially targeted and 31 had an OMIM annotation. Known ciliopathy genes included *OFD1*, which is mutated as a cause of Oro-Facio-Digital type 1 and Joubert syndrome type 10, *USH1C*, which is mutated in Usher syndrome type 1c, *CEP120*, mutated in Jeune asphyxiating thoracic dystrophy²⁷⁸ and *PLK4*, mutations in which cause ciliopathy-like phenotype – microcephaly with growth failure and retinopathy²⁷⁹. All of the hits with a human orthologue were taken forward for the secondary screen (**Table 3-19**). This also included seven hits with predicted off-target effects to investigate if single duplexes of siRNAs produced with different chemistry and targeting different sequences of RNA would still have an off-target effect.

The list of genes was analysed for enrichment of functional annotations using DAVID. Swiss-Prot and Protein Information Resource databases keyword searches identified statistically significant enrichment of the terms ‘G-protein coupled receptors’ and ‘transducers’ (**Figure 3-49**). Gene Ontology searches of cellular components showed the most significantly enriched terms were ‘plasma membrane’, ‘nonmotile primary cilium’ and ‘photoreceptor outer segment components’ (**Figure 3-50**), all of which are either directly or indirectly linked to the primary cilium. Transmembrane proteins such as TMEM67, TMEM138, TMEM216 or TMEM237 are known to be mutated in ciliopathies and localise to the ciliary transition zone^{99,144,206,280}. Also modified TZ in photoreceptor cells (known as the connecting cilium), links the inner and outer segments of photoreceptor cells. KEGG pathway annotations showed enrichment for the terms ‘neuroactive ligand-receptor interaction’ (**Figure 3-51** and **3-52**). This finding was particularly interesting, as there are few reports about cilia involvement in neurotransmission, although primary cilia are well-known to be present on the surface of differentiated neuronal cells²⁸¹.

%CSC run1	%CSC run2	%CSC average	WCN average	Mouse GeneID	Mouse Gene Symbol	Human GeneID	Human Gene Symbol	OMIM
-7.35	-9.56	-8.46	-3.76	15551	<i>Htr1b</i>	3351	<i>HTR1B</i>	
-7.97	-8.78	-8.37	-4.18	80910	<i>Gpr84</i>	53831	<i>GPR84</i>	
-6.74	-7.33	-7.03	-4.63	14428	<i>Galr2</i>	8811	<i>GALR2</i>	
-6.47	-7.25	-6.86	-1.47	12057	<i>Opn1sw</i>	611	<i>OPN1SW</i>	Colorblindness, tritan, 19090
-6.12	-6.91	-6.52	-3.87	140795	<i>P2ry14</i>	9934	<i>P2RY14</i>	
-6.74	-6.08	-6.41	-3.08	215859	<i>Taar8a</i>	83551	<i>TAAR8</i>	
-6.66	-5.85	-6.26	-3.01	70771	<i>Gpr173</i>	54328	<i>GPR173</i>	
-4.76	-7.66	-6.21	-1.73	78560	<i>Gpr124</i>	25960	<i>GPR124</i>	
-4.76	-7.62	-6.19	-3.32	83603	<i>Elovl4</i>	6785	<i>ELOVL4</i>	Macular dystrophy, autosomal dominant, chromosome 6-linked, 600110; Ichthyosis, spastic quadriplegia, and mental

								retardation, 614457; Stargardt disease 3, 600110
-4.75	-7.60	-6.18	-2.03	76206	<i>Gpr165</i>	392486	<i>GPR165P</i>	
-6.48	-5.81	-6.14	-4.13	17171	<i>Mas1</i>	4142	<i>MAS1</i>	
-3.90	-8.13	-6.02	-1.98	12768	<i>Ccr1</i>	1230	<i>CCR1</i>	
-5.77	-6.09	-5.93	-2.29	236781	<i>Gpr119</i>	139760	<i>GPR119</i>	
-3.11	-8.57	-5.84	-3.18	14539	<i>Opn1mw</i>	2652	<i>OPN1MW</i>	Colorblindness, deutan, 303800; Blue cone monochromacy, 303700
-5.75	-5.86	-5.80	-0.23	13491	<i>Drd4</i>	1815	<i>DRD4</i>	Autonomic nervous system dysfunction
-3.87	-7.53	-5.70	-2.50	17202	<i>Mc4r</i>	4160	<i>MC4R</i>	Obesity, autosomal dominant, 601665
-4.51	-6.82	-5.67	-2.35	57260	<i>Ltb4r2</i>	56413	<i>LTB4R2</i>	
-3.51	-7.53	-5.52	-2.92	11555	<i>Adrb2</i>	154	<i>ADRB2</i>	Beta-2-adrenoreceptor agonist, reduced response to
-3.62	-6.84	-5.23	-4.09	239530	<i>Gpr20</i>	2843	<i>GPR20</i>	
-3.80	-6.59	-5.19	-3.57	14715	<i>Gnrhr</i>	2798	<i>GNRHR</i>	Fertile eunuch syndrome, 228300; Hypogonadotropic hypogonadism
-3.39	-6.96	-5.17	-2.27	81006	<i>Gpr63</i>	81491	<i>GPR63</i>	
-3.33	-6.57	-4.95	-2.18	94043	<i>Tm2d1</i>	83941	<i>TM2D1</i>	
-3.62	-6.16	-4.89	-1.60	18389	<i>Opr1</i>	4987	<i>OPRL1</i>	
-3.89	-5.67	-4.78	-3.63	107934	<i>Celsr3</i>	1951	<i>CELSR3</i>	
-3.11	-6.29	-4.70	-1.76	20607	<i>Sstr3</i>	6753	<i>SSTR3</i>	
-3.43	-5.78	-4.60	-2.29	12922	<i>Crhr2</i>	1395	<i>CRHR2</i>	
-3.24	-5.86	-4.55	-1.58	15552	<i>Htr1d</i>	3352	<i>HTR1D</i>	
-3.06	-5.77	-4.42	-1.71	11540	<i>Adora2a</i>	135	<i>ADORA2A</i>	
-14.96	-10.10	-12.53	-2.38	107932	<i>Chd4</i>	1108	<i>CHD4</i>	
-10.44	-7.71	-9.07	-2.29	66105	<i>Ube2d3</i>	7323	<i>UBE2D3</i>	
-8.86	-6.51	-7.69	-2.56	225055	<i>Fbxo11</i>	80204	<i>FBXO11</i>	
-8.86	-6.45	-7.65	-3.12	19400	<i>Rapsn</i>	5913	<i>RAPSN</i>	Myasthenic syndrome, congenital, associated with facial dysmorphism and acetylcholine receptor deficiency, 608931; Myasthenic syndrome, congenital, associated with acetylcholine receptor deficiency, 608931; Fetal akinesia deformation sequence, 208150
-7.95	-7.03	-7.49	-3.55	72973	<i>Fbxo47</i>	494188	<i>FBXO47</i>	
-8.23	-6.68	-7.45	-2.78	78938	<i>Fbxo34</i>	55030	<i>FBXO34</i>	
-7.52	-6.44	-6.98	-3.12	50759	<i>Fbxo16</i>	157574	<i>FBXO16</i>	
-6.60	-6.92	-6.76	-2.16	94094	<i>Trim34a</i>	53840	<i>TRIM34</i>	
-7.70	-5.38	-6.54	-1.31	72323	<i>Asb6</i>	140459	<i>ASB6</i>	
-11.65	-7.47	-9.56	-1.24	269180	<i>Inpp4a</i>	3631	<i>INPP4A</i>	
-8.87	-9.81	-9.34	-2.40	16535	<i>Kcnq1</i>	3784	<i>KCNQ1</i>	Short QT syndrome-2, 609621; Long QT syndrome-1, 192500; Jervell and Lange-Nielsen syndrome, 220400; Atrial fibrillation familial 3, 607554
-8.72	-8.90	-8.81	-3.13	68440	<i>Dusp23</i>	54935	<i>DUSP23</i>	
-9.22	-7.50	-8.36	-2.56	11928	<i>Atp1a1</i>	476	<i>ATP1A1</i>	
-7.91	-5.96	-6.94	-2.45	227292	<i>Ctdsp1</i>	58190	<i>CTDSP1</i>	
-6.16	-8.35	-7.25	-1.64	224697	<i>Adamts10</i>	81794	<i>ADAMTS10</i>	Weill-Marchesani syndrome 1, recessive, 277600
-4.99	-5.68	-5.34	-1.78	30806	<i>Adamts8</i>	11095	<i>ADAMTS8</i>	
-4.75	-5.88	-5.31	-0.62	11490	<i>Adam15</i>	8751	<i>ADAM15</i>	
-5.94	-6.19	-6.07	-1.96	268822	<i>Adck5</i>	203054	<i>ADCK5</i>	
-6.42	-5.39	-5.90	-1.48	14933	<i>Gyk</i>	2710	<i>GK</i>	Glycerol kinase deficiency, 307030
-4.97	-5.44	-5.21	-2.64	77976	<i>Nuak1</i>	9891	<i>NUAK1</i>	
-4.70	-5.40	-5.05	-1.57	192292	<i>Nrbp1</i>	29959	<i>NRBP1</i>	
-5.05	-4.86	-4.95	-1.53	20873	<i>Plk4</i>	10733	<i>PLK4</i>	
-4.22	-4.43	-4.33	-1.60	69635	<i>Dapk1</i>	1612	<i>DAPK1</i>	
-4.98	-3.03	-4.00	-0.92	110355	<i>Adrbk1</i>	156	<i>ADRBK1</i>	
-4.89	-2.73	-3.81	-2.35	58864	<i>Tssk3</i>	81629	<i>TSSK3</i>	
-2.29	-5.35	-3.82	-0.38	109594	<i>Lmo1</i>	4004	<i>LMO1</i>	Leukemia, T-cell acute lymphoblastic
-2.27	-5.09	-3.68	-1.71	59057	<i>Zfp191</i>	7572	<i>ZNF24</i>	
-2.61	-4.58	-3.60	-1.28	17349	<i>Mif1</i>	4291	<i>MLF1</i>	Leukemia, acute myeloid, 601626
-2.59	-4.31	-3.45	-1.82	19387	<i>Rangap1</i>	5905	<i>RANGAP1</i>	

-6.53	-6.45	-6.49	-0.55	17160	<i>Man2b2</i>	23324	<i>MAN2B2</i>	
-6.87	-5.87	-6.37	0.01	12479	<i>Cd1d1</i>	912	<i>CD1D</i>	
-7.65	-4.74	-6.19	0.18	17751	<i>Mt3</i>	4504	<i>MT3</i>	
-5.63	-6.00	-5.82	-1.29	103711	<i>Pnpo</i>	55163	<i>PNPO</i>	Pyridoxamine 5'-phosphate oxidase deficiency, 610090
-4.85	-6.18	-5.51	-1.63	11910	<i>Atf3</i>	467	<i>ATF3</i>	
-4.61	-4.94	-4.77	-1.76	226049	<i>Dmrt2</i>	10655	<i>DMRT2</i>	
-4.25	-4.97	-4.61	0.63	240690	<i>St18</i>	9705	<i>ST18</i>	
-4.95	-4.09	-4.52	4.49	245867	<i>Pcmdt2</i>	55251	<i>PCMTD2</i>	
-4.83	-4.10	-4.46	-1.85	227327	<i>B3gnt7</i>	93010	<i>B3GNT7</i>	
-3.77	-4.20	-3.98	-0.85	234889	<i>Gucy1a2</i>	2977	<i>GUCY1A2</i>	
-3.93	-4.01	-3.97	0.61	242341	<i>Atp6v0d2</i>	245972	<i>ATP6V0D2</i>	
-4.16	-3.69	-3.93	-0.11	17826	<i>Fam89b</i>	23625	<i>FAM89B</i>	
-4.18	-3.57	-3.87	-0.47	56085	<i>Ubqln1</i>	29979	<i>UBQLN1</i>	
-7.33	-9.85	-8.59	-0.28	192113	<i>Atp12a</i>	479	<i>ATP12A</i>	
-8.55	-6.46	-7.51	-0.71	17754	<i>Mtap1a</i>	4130	<i>MAP1A</i>	
-6.55	-8.28	-7.42	-2.08	56445	<i>Dnaja2</i>	10294	<i>DNAJA2</i>	
-6.06	-8.21	-7.14	-1.95	12419	<i>Cbx5</i>	23468	<i>CBX5</i>	
-6.34	-7.64	-6.99	-1.58	67738	<i>Ppid</i>	5481	<i>PPID</i>	
-5.27	-7.28	-6.28	-2.46	18584	<i>Pde8a</i>	5151	<i>PDE8A</i>	
-6.05	-6.30	-6.18	-2.42	54215	<i>Cd160</i>	11126	<i>CD160</i>	
-11.36	-9.66	-10.51	1.85	77766	<i>Elp4</i>	26610	<i>ELP4</i>	
-8.92	-8.54	-8.73	-1.80	232984	<i>B3gnt8</i>	374907	<i>B3GNT8</i>	
-9.98	-12.22	-11.10	-0.93	12351	<i>Car4</i>	762	<i>CA4</i>	Retinitis pigmentosa 17, 600852
-6.20	-12.01	-9.11	-1.39	20526	<i>Slc2a2</i>	6514	<i>SLC2A2</i>	Fanconi-Bickel syndrome, 227810
-6.34	-8.82	-7.58	-3.72	114886	<i>Cygb</i>	114757	<i>CYGB</i>	
-4.61	-9.68	-7.15	-4.18	14595	<i>B4galt1</i>	2683	<i>B4GALT1</i>	Congenital disorder of glycosylation, type IId, 607091
-4.51	-6.89	-5.70	-1.81	57138	<i>Slc12a5</i>	57468	<i>SLC12A5</i>	
-8.32	-24.29	-16.31	-1.56	237222	<i>Ofd1</i>	8481	<i>OFD1</i>	Oral-facial-digital syndrome 1, 311200; Joubert syndrome 10, 300804; Simpson-Golabi-Behmel syndrome, type 2, 300209
-7.82	-24.02	-15.92	-1.83	216227	<i>Slc17a8</i>	246213	<i>SLC17A8</i>	Deafness, autosomal dominant 25, 605583
-7.26	-24.49	-15.87	-1.84	14057	<i>Sfxn1</i>	94081	<i>SFXN1</i>	
-7.59	-22.05	-14.82	-2.41	108682	<i>Gpt2</i>	84706	<i>GPT2</i>	
-13.88	-11.96	-12.92	-3.12	70208	<i>Med23</i>	9439	<i>MED23</i>	Mental retardation, autosomal recessive 18, 614249
-13.95	-11.10	-12.52	-3.03	21922	<i>Clec3b</i>	7123	<i>CLEC3B</i>	
-10.39	-11.89	-11.14	-1.61	56847	<i>Aldh1a3</i>	220	<i>ALDH1A3</i>	
-13.39	-8.45	-10.92	-2.95	217166	<i>Nr1d1</i>	9572	<i>NR1D1</i>	
-7.78	-13.70	-10.74	-2.08	56324	<i>Stam2</i>	10254	<i>STAM2</i>	
-13.32	-7.69	-10.51	-0.14	17300	<i>Foxc1</i>	2296	<i>FOXC1</i>	Axenfeld-Rieger syndrome, type 3, 602482; Iridogoniodysgenesis, type 1, 601631; Iris hypoplasia and glaucoma, 601631; Rieger or Axenfeld anomalies, 602482
-11.27	-9.39	-10.33	-2.09	237940	<i>Aoc2</i>	314	<i>AOC2</i>	
-11.57	-8.90	-10.24	-1.39	52206	<i>Anapc4</i>	29945	<i>ANAPC4</i>	
-11.68	-8.57	-10.12	-2.14	16779	<i>Lamb2</i>	3913	<i>LAMB2</i>	Nephrotic syndrome, type 5, with or without ocular abnormalities, 614199; Pierson syndrome, 609049
-9.32	-10.21	-9.76	-1.66	14583	<i>Gfpt1</i>	2673	<i>GFPT1</i>	
-6.89	-9.73	-8.31	-2.34	224656	<i>Zfp523</i>	7629	<i>ZNF76</i>	
-8.14	-7.90	-8.02	-1.28	72088	<i>Ush1c</i>	10083	<i>USH1C</i>	Deafness, autosomal recessive 18, 602092; Usher syndrome, type 1C, 276904
-7.40	-7.42	-7.41	-2.80	78829	<i>Tsc22d4</i>	81628	<i>TSC22D4</i>	
-12.42	-12.54	-12.48	-1.45	16691	<i>Krt8</i>	3856	<i>KRT8</i>	Cirrhosis, cryptogenic
-7.95	-12.52	-10.24	-0.16	53791	<i>Tlr5</i>	7100	<i>TLR5</i>	
-8.70	-8.89	-8.79	-2.31	14563	<i>Gdf5</i>	8200	<i>GDF5</i>	Brachydactyly, type C, 113100; Brachydactyly, type A2, 112600; Acromesomelic dysplasia, Hunter-Thompson type, 201250; Chondrodysplasia, Grebe type, 200700; Fibular hypoplasia and complex brachydactyly, 228900; Multiple synostoses syndrome 2, 610017; Symphalangism, proximal, 185800

-7.67	-8.68	-8.18	-0.78	19414	<i>Rasa3</i>	22821	<i>RASA3</i>	
-4.23	-8.88	-6.56	-1.59	15107	<i>Hadh</i>	3033	<i>HADH</i>	3-hydroxyacyl-CoA dehydrogenase deficiency, 231530; Hyperinsulinemic hypoglycemia, familial, 4, 609975
-3.20	-8.97	-6.09	-1.06	276770	<i>Eif5a</i>	1984	<i>EIF5A</i>	
-3.89	-8.28	-6.09	-1.48	16663	<i>Krt13</i>	3860	<i>KRT13</i>	White sponge nevus, 193900
-4.14	-7.04	-5.59	-0.57	74760	<i>Rab3il1</i>	5866	<i>RAB3IL1</i>	
-4.64	-6.13	-5.38	-1.21	11768	<i>Ap1m2</i>	10053	<i>AP1M2</i>	
-3.69	-6.62	-5.15	-0.17	214498	<i>Cdc73</i>	79577	<i>CDC73</i>	Hyperparathyroidism-jaw tumor syndrome, 145001; Hyperparathyroidism, familial primary, 145000; Parathyroid adenoma with cystic changes, 145001; Parathyroid carcinoma, 608266
-3.07	-6.47	-4.77	-2.04	15959	<i>Ifit3</i>	3437	<i>IFIT3</i>	
-3.09	-6.35	-4.72	-0.95	21810	<i>Tgfb1</i>	7045	<i>TGFBI</i>	Corneal dystrophy, Thiel-Behnke type, 602082; Corneal dystrophy, Reis-Bucklers type, 608470; Corneal dystrophy, Groenouw type I, 121900; Corneal dystrophy, Avellino type, 607541; Corneal dystrophy, epithelial basement membrane, 121820; Corneal dystrophy, lattice type I, 122200; Corneal dystrophy, lattice type IIIA, 608471
-4.98	-4.38	-4.68	-1.10	13074	<i>Cyp17a1</i>	1586	<i>CYP17A1</i>	17,20-lyase deficiency, isolated, 202110; 17-alpha-hydroxylase/17,20-lyase deficiency, 202110
-3.41	-5.77	-4.59	-1.99	16647	<i>Kpna2</i>	3838	<i>KPNA2</i>	
-3.20	-5.27	-4.23	-0.49	207839	<i>Galnt6</i>	11226	<i>GALNT6</i>	
-3.13	-5.24	-4.18	-2.16	58250	<i>Chst11</i>	50515	<i>CHST11</i>	
-2.82	-4.42	-3.62	-0.65	230673	<i>Ipo13</i>	9670	<i>IPO13</i>	
-8.64	-2.52	-5.58	-2.02	18857	<i>Pmp2</i>	5375	<i>PMP2</i>	
-7.30	-3.10	-5.20	-1.89	384009	<i>Glpr2</i>	152007	<i>GLIPR2</i>	
-8.24	-10.47	-9.36	-0.35	69029	<i>1500032L24Rik</i>	91689	<i>C22orf32</i>	
-8.50	-6.81	-7.66	0.61	76719	<i>1700081L11Rik</i>	284058	<i>KANSL1</i>	
-5.08	-6.33	-5.71	-1.54	227736	<i>1700019L03Rik</i>	286207	<i>C9orf117</i>	
-5.10	-5.94	-5.52	-0.98	67902	<i>Sumf2</i>	25870	<i>SUMF2</i>	
-4.18	-5.00	-4.59	-2.27	208606	<i>Rsrc2</i>	65117	<i>RSRC2</i>	
-4.40	-3.75	-4.08	2.36	70381	<i>Tecpr1</i>	25851	<i>TECPR1</i>	
-5.67	-5.31	-5.49	-2.07	78887	<i>Sfi1</i>	9814	<i>SFI1</i>	
-4.94	-3.88	-4.41	-1.60	67809	<i>Fam82a2</i>	55177	<i>FAM82A2</i>	
-5.12	-3.44	-4.28	-0.67	72373	<i>Psca</i>	8000	<i>PSCA</i>	
-5.73	-2.65	-4.19	-0.49	74716	<i>Wbp2nl</i>	164684	<i>WBP2NL</i>	
-4.69	-3.63	-4.16	-1.14	116905	<i>Dph1</i>	1801	<i>DPH1</i>	
-4.92	-3.38	-4.15	0.70	67878	<i>Tmem33</i>	55161	<i>TMEM33</i>	
-12.08	-3.57	-7.83	-1.50	258961	<i>Olfr631</i>	390059	<i>OR51M1</i>	
-11.77	-9.49	-10.63	1.36	51812	<i>Mcrs1</i>	10445	<i>MCRS1</i>	
-12.98	-8.17	-10.58	-0.19	73242	<i>Atat1</i>	79969	<i>ATAT1</i>	
-6.67	-7.30	-6.98	-0.77	68118	<i>9430023L20Rik</i>	60673	<i>C12orf44</i>	
-6.99	-5.72	-6.36	0.15	19934	<i>Rpl22</i>	6146	<i>RPL22</i>	
-19.21	-4.16	-11.69	-1.14	69612	<i>2310037I24Rik</i>	54934	<i>KANSL2</i>	
-7.40	-2.26	-4.83	-0.07	108671	<i>Dnajc9</i>	23234	<i>DNAJC9</i>	
-5.78	-2.91	-4.35	-2.57	320840	<i>Negr1</i>	257194	<i>NEGR1</i>	
-13.57	-9.60	-11.59	-1.29	66592	<i>Stoml2</i>	30968	<i>STOML2</i>	
-10.32	-4.32	-7.32	-1.32	66373	<i>Lsm5</i>	23658	<i>LSM5</i>	
-13.50	-3.70	-8.60	-0.65	52023	<i>Pibf1</i>	10464	<i>PIBF1</i>	
-12.95	-3.92	-8.43	-1.89	83701	<i>Srrt</i>	51593	<i>SRRT</i>	
-8.67	-2.43	-5.55	1.58	71617	<i>9130011E15Rik</i>	79591	<i>C10orf76</i>	
-8.68	-17.61	-13.14	-0.42	77048	<i>Ccdc41</i>	51134	<i>CCDC41</i>	
-21.69	-4.64	-13.16	-1.61	27756	<i>Lsm2</i>	57819	<i>LSM2</i>	
-17.49	-6.77	-12.13	-3.61	404315	<i>Olfr372</i>	284383	<i>OR2Z1</i>	
-17.64	-4.43	-11.04	-0.45	67884	<i>1810043G02Rik</i>	755	<i>C21orf2</i>	
-16.73	-4.43	-10.58	-2.41	225523	<i>Cep120</i>	153241	<i>CEP120</i>	
-11.66	-9.48	-10.57	-1.16	319158	<i>Hist1h4i</i>	8294	<i>HIST1H4I</i>	
-16.05	-5.06	-10.55	-3.31	171506	<i>H1foo</i>	132243	<i>H1FOO</i>	
-10.95	-8.52	-9.74	2.12	67217	<i>2810055F11Rik</i>	112849	<i>C14orf149</i>	
-13.35	-5.97	-9.66	-2.77	93714	<i>Pcdhga6</i>	56109	<i>PCDHGA6</i>	

Table 3-19. List of genes that passed the second level of filtering. 154 genes had robust z score lesser than the median robust z score of positive controls for %CSC, and robust z score greater than the median robust z score reduced by two MADs of the positive controls for WCN in both runs of a ten plate batch and had homologues in human genome.

Sublist	Category	Term	RT	Genes	Count	%	P-Value	Benjamini
SP_PIR_KEYWORDS		g-protein coupled receptor	RT		33	19.0	1.5E-6	3.2E-4
SP_PIR_KEYWORDS		transducer	RT		33	19.0	3.4E-6	3.6E-4
SP_PIR_KEYWORDS		cell membrane	RT		36	20.7	4.8E-6	3.3E-4
SP_PIR_KEYWORDS		receptor	RT		44	25.3	1.8E-5	9.2E-4
SP_PIR_KEYWORDS		G protein-coupled receptor	RT		6	3.4	1.8E-4	7.3E-3
SP_PIR_KEYWORDS		glycoprotein	RT		53	30.5	3.1E-4	1.1E-2
SP_PIR_KEYWORDS		disulfide bond	RT		36	20.7	5.4E-3	1.5E-1
SP_PIR_KEYWORDS		membrane	RT		67	38.5	6.5E-3	1.6E-1
SP_PIR_KEYWORDS		lipoprotein	RT		12	6.9	2.0E-2	3.8E-1
SP_PIR_KEYWORDS		neurotransmitter receptor	RT		3	1.7	5.2E-2	6.7E-1
SP_PIR_KEYWORDS		retinal protein	RT		2	1.1	7.1E-2	7.6E-1
SP_PIR_KEYWORDS		potassium transport	RT		4	2.3	7.6E-2	7.5E-1
SP_PIR_KEYWORDS		photoreceptor protein	RT		2	1.1	8.8E-2	7.7E-1
SP_PIR_KEYWORDS		potassium	RT		4	2.3	9.5E-2	7.8E-1
SP_PIR_KEYWORDS		chromophore	RT		2	1.1	9.7E-2	7.6E-1

Figure 3-49. DAVID analysis using Swiss-Prot and Protein Information Resource keywords on the 154 filtered primary screen hits. Listed are enriched terms with associated *P* values and Benjamini-Hochberg *q* values (to control for the false discovery rate due to multiple testing). Terms with *q* < 0.05 are listed. Image modified from <http://david.abcc.ncifcrf.gov/>.

Sublist	Category	Term	RT	Genes	Count	%	P-Value	Benjamini
GOTERM_CC_FAT		plasma membrane	RT		52	29.9	1.5E-6	2.9E-4
GOTERM_CC_FAT		nonmotile primary cilium	RT		4	2.3	3.4E-3	2.8E-1
GOTERM_CC_FAT		intrinsic to membrane	RT		72	41.4	4.3E-3	2.4E-1
GOTERM_CC_FAT		photoreceptor outer segment	RT		3	1.7	1.8E-2	5.8E-1
GOTERM_CC_FAT		cell projection	RT		12	6.9	2.1E-2	5.5E-1
GOTERM_CC_FAT		integral to membrane	RT		66	37.9	3.0E-2	6.2E-1
GOTERM_CC_FAT		anchored to membrane	RT		6	3.4	4.8E-2	7.3E-1
GOTERM_CC_FAT		chromatin	RT		5	2.9	7.3E-2	8.3E-1
GOTERM_CC_FAT		sarcolemma	RT		3	1.7	7.4E-2	8.0E-1

Figure 3-50. DAVID analysis using GO terms for cellular components on the 154 filtered primary screen hits. Listed are enriched terms with associated *P* values and Benjamini-Hochberg *q* values (to control for the false discovery rate due to multiple testing). Terms with *q* < 0.05 are listed. Image modified from <http://david.abcc.ncifcrf.gov/>.

Sublist	Category	Term	RT	Genes	Count	%	P-Value	Benjamini
KEGG_PATHWAY		Neuroactive ligand-receptor interaction	RT		15	8.6	1.8E-7	1.3E-5

Figure 3-51. DAVID analysis using KEGG pathway term search on the 154 filtered primary screen hits. Listed are enriched terms with associated *P* values and Benjamini-Hochberg *q* values (to control for the false discovery rate due to multiple testing). Terms with *q* < 0.05 are listed. Image modified from <http://david.abcc.ncifcrf.gov/>.

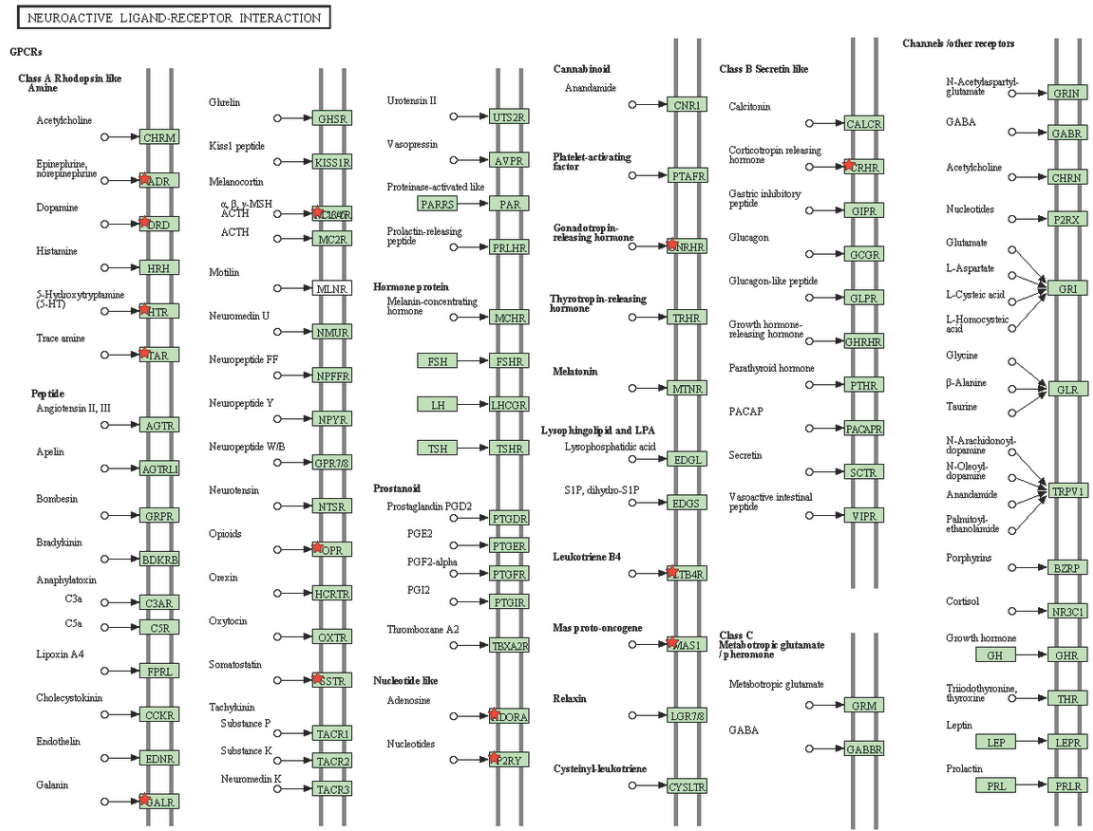


Figure 3-52. Cartoon illustrating enrichment of the 154 filtered primary screen hits for neuroactive ligand-receptor interactions. Red stars indicate proteins encoded by genes that were hits in the screen or interacting partners for those proteins. Image modified from <http://david.abcc.ncifcrf.gov/>.

3.3.2.3.2 Secondary screen

2,174 hits from the primary screen were filtered to obtain a list of genes with the most significant effect on ciliogenesis (%CSC) that could be taken forward for validation in a secondary screen. Two different filtering strategies were applied in parallel (**Figure 3-53**). In the first instance, out of 2,174 hits, those with a statistically significant effect on cell number were removed. Out of remaining 174 hits, any that did not have a human orthologue were removed. This comprised 154 target genes (**Table 3-19**) that were taken forward to the secondary screen. The other, less stringent filtering method, included all genes with an effect on cell number (2,000 hits out of 2,174; 1,802 with human homologue) but were components that were enriched in discrete functional modules (**Figure 3-45**). From this group, 38 hits were taken forward for the secondary screen.

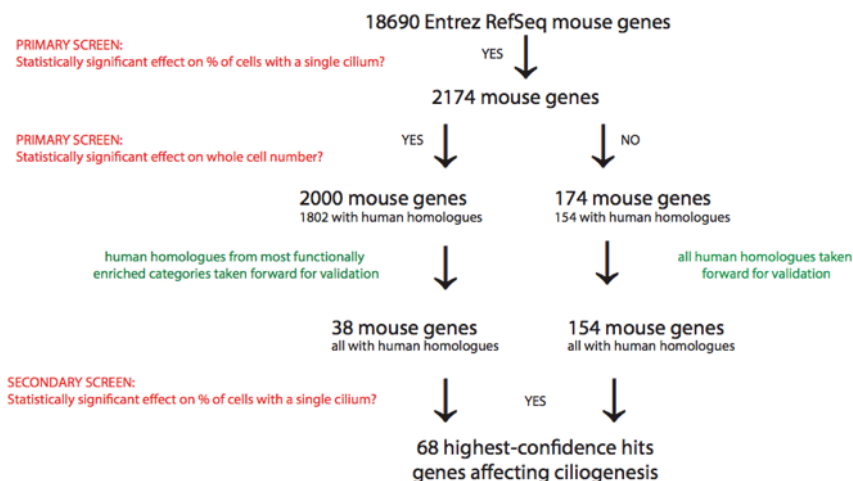


Figure 3-53. Filtering strategy in the siRNA screen. Initially 18,690 Entrez RefSeq mouse genes were screened. Out of those 2,174 had a significant effect of %CSC. Two parallel filters were further applied: genes within the most functionally enriched categories having effect on both WCN and %CSC (38 targets) and genes with effect on %CSC but no effect on WCN (154 targets). Only hits with human homologues were analysed in the secondary screen. Out of 192 targets, 68 were confirmed in the secondary screen, an overall hit rate of 35%.

192 hits, that passed filtering after primary screen, were targeted for validation in the secondary screen. IMCD3s were used as the preferred model system, as they proved to be reliable and robust in the set-up of the primary screen, but siRNA duplexes of a different chemistry were used to those in the primary screen. The primary screen used pooled Dharmacon “siGENOME” reagents, whereas the secondary screen used four individual Dharmacon “ON-TARGET PLUS” unpooled siRNA duplexes which have modifications that decrease off-target effects and have an improved design to target all known transcripts of the target gene. The same protocol for set up, staining, imaging and analysis was used as in the primary screen to maintain uniformity between the two screens.

To maintain consistency in the analysis, the 154 screened targets were filtered for an effect on ciliogenesis (%CSC) but without effect on cell number (WCN). 60 targets were confirmed in the secondary screen. No hits gave a phenotype for all four duplexes, but seven hits gave a phenotype with three duplexes, eighteen hits with two duplexes and thirty five hits with one duplex (**Table 3-20**, top 60 hits).

Out of the 38 genes shortlisted from the enriched functional modules, eight were confirmed in the secondary screen (**Table 3-20**, yellow shaded). Hits were filtered for an effect on cilia number but had no filter based on cell number. One of the hits had all four duplexes giving a ciliogenesis defect phenotype, one showed a phenotype with three targeting duplexes, three hits with two duplexes and four with one duplex.

The variability of %CSC phenotype between duplexes targeting one gene are shown in **Figure 3-54**. Genes in which the %CSC value was confirmed in the secondary screen, and therefore by duplexes using a different chemistry, were designated as positive hits. The functional annotation modules enriched in these 68 hits include centrosome/microtubule organisation, spliceosome/RNA processing, GPCR, UPS/proteasome components (**Figure 3-54** and **3-55**).

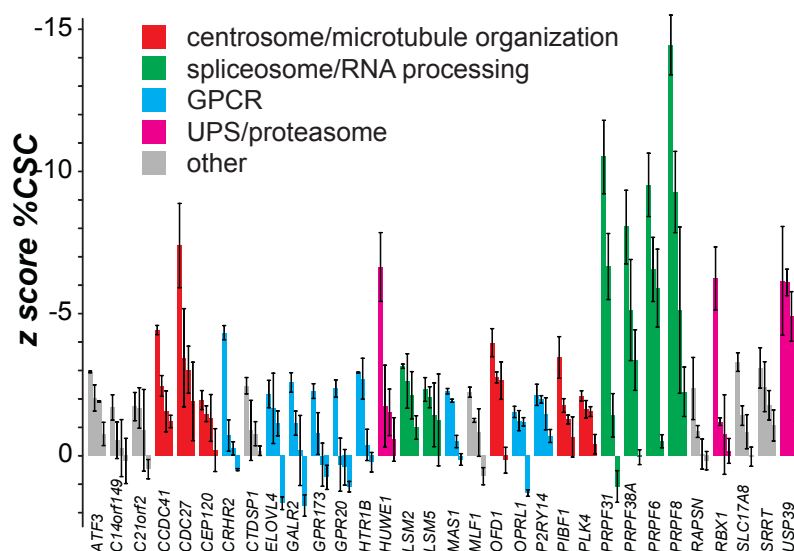


Figure 3-54. Example of hits confirmed in the secondary screen. Robust z scores for %CSC for selected validated genes from the secondary screen, each assayed in two replicates with four individual siRNAs in mouse IMCD3 cells. Error bars indicate the range in values for the replicates. Cut-off values are based on the median z of positive controls per batch (-1.44, -0.81, -2.89, -4.08) for each batch. Colours indicate selected functional annotations, including centrosome/microtubule organisation genes (red), spliceosome/RNA processing genes (green), GPCRs (blue), UPS/proteasome genes (pink) and others (grey).

Sublist	Category	Term	RT	Genes	Count	%	P-Value	Benjamini
	KEGG_PATHWAY	Spliceosome	RT		7	10.3	2.5E-5	7.9E-4
	KEGG_PATHWAY	Neuroactive ligand-receptor interaction	RT		6	8.8	7.2E-3	1.1E-1

Figure 3-55. DAVID analysis of enriched KEGG pathways in the 68 hits from the secondary screen. Statistically significant enrichment of spliceosome components and GPCRs were identified. Image modified from <http://david.abcc.ncifcrf.gov/>.

The 68 targets confirmed in the screen represent a 35% confirmatory hit rate for the secondary screen and an overall 0.4% hit rate for the whole genome screen (**Table 3-20**).

%CSC run1	%CSC run2	%CSC Average	WCN Average	Mouse GeneID	Mouse Gene Symbol	Human GeneID	Human Gene Symbol	OMIM
-1.85	-1.70	-1.78	-3.20	30806	<i>Adamts8</i>	11095	<i>ADAMTS8</i>	
-3.55	-1.02	-2.28	0.13	268822	<i>Adck5</i>	203054	<i>ADCK5</i>	
-1.95	-2.36	-2.15	-5.64	11768	<i>Ap1m2</i>	10053	<i>AP1M2</i>	

-1.54	-1.69	-1.62	-2.25	11768	<i>Ap1m2</i>	10053	<i>AP1M2</i>	
-2.98	-2.90	-2.94	-6.59	11910	<i>Atf3</i>	467	<i>ATF3</i>	
-2.68	-1.38	-2.03	-5.05	11910	<i>Atf3</i>	467	<i>ATF3</i>	
-1.87	-1.94	-1.90	-6.28	11910	<i>Atf3</i>	467	<i>ATF3</i>	
-1.83	-0.85	-1.34	-5.30	242341	<i>Atp6v0d2</i>	245972	<i>ATP6V0D2</i>	
-4.56	-1.32	-2.94	-1.18	14595	<i>B4galt1</i>	2683	<i>B4GALT1</i>	Congenital disorder of glycosylation type IId, 607091
-5.80	-2.60	-4.20	-9.59	71617	<i>9130011E15Rik</i>	79591	<i>C10orf76</i>	
-2.32	-1.06	-1.69	-1.76	67217	<i>2810055F11Rik</i>	112849	<i>C14orf149</i>	
-2.42	-1.02	-1.72	-8.70	67884	<i>1810043G02Rik</i>	755	<i>C21orf2</i>	
-2.14	-2.96	-2.55	-2.74	12419	<i>Cbx5</i>	23468	<i>CBX5</i>	
-1.99	-3.08	-2.53	-3.06	12419	<i>Cbx5</i>	23468	<i>CBX5</i>	
-4.64	-4.18	-4.41	-0.66	77048	<i>Ccdc41</i>	51134	<i>CCDC41</i>	
-2.97	-1.93	-2.45	-7.86	77048	<i>Ccdc41</i>	51134	<i>CCDC41</i>	
-1.51	-0.89	-1.20	-1.59	77048	<i>Ccdc41</i>	51134	<i>CCDC41</i>	
-2.31	-2.40	-2.35	-2.55	214498	<i>Cdc73</i>	79577	<i>CDC73</i>	Hyperparathyroidism-jaw tumor syndrome, 145001; Hyperparathyroidism familial primary, 145000; Parathyroid adenoma with cystic changes, 145001; Parathyroid carcinoma, 608266
-2.42	-1.48	-1.95	-3.23	225523	<i>Cep120</i>	153241	<i>CEP120</i>	
-1.86	-1.08	-1.47	-1.57	225523	<i>Cep120</i>	153241	<i>CEP120</i>	
-3.97	-4.67	-4.32	0.98	12922	<i>Crhr2</i>	1395	<i>CRHR2</i>	
-2.86	-2.05	-2.46	-9.67	227292	<i>Ctdsp1</i>	58190	<i>CTDSP1</i>	
-2.85	-1.48	-2.16	-1.88	83603	<i>Elovl4</i>	6785	<i>ELOVL4</i>	Macular dystrophy autosomal dominant chromosome 6-linked, 600110; Ichthyosis spastic quadriplegia and mental retardation, 614457; Stargardt disease 3, 600110
-3.05	-2.09	-2.57	0.41	14428	<i>Galr2</i>	8811	<i>GALR2</i>	
-4.49	-2.85	-3.67	-3.12	14563	<i>Gdf5</i>	8200	<i>GDF5</i>	Brachydactyly, type C, 113100; Brachydactyly, type A2, 112600; Acromesomelic dysplasia, Hunter-Thompson type, 201250; Chondrodysplasia, Grebe type, 200700; Fibular hypoplasia and complex brachydactyly, 228900; Multiple synostoses syndrome 2, 610017; Symphalangism proximal, 185800
-1.78	-0.96	-1.37	0.69	384009	<i>Glipr2</i>	152007	<i>GLIPR2</i>	
-1.90	-2.63	-2.27	-0.25	70771	<i>Gpr173</i>	54328	<i>GPR173</i>	
-1.93	-2.78	-2.36	-1.01	239530	<i>Gpr20</i>	2843	<i>GPR20</i>	
-4.46	-1.02	-2.74	-1.01	15107	<i>Hadh</i>	3033	<i>HADH</i>	3-hydroxyacyl-CoA dehydrogenase deficiency, 231530; Hyperinsulinemic hypoglycemia familial, 4, 609975
-2.83	-1.57	-2.20	-4.56	15107	<i>Hadh</i>	3033	<i>HADH</i>	3-hydroxyacyl-CoA dehydrogenase deficiency, 231530; Hyperinsulinemic hypoglycemia familial, 4, 609975
-2.94	-2.90	-2.92	-0.71	15551	<i>Htr1b</i>	3351	<i>HTR1B</i>	
-3.72	-1.69	-2.70	-4.54	15551	<i>Htr1b</i>	3351	<i>HTR1B</i>	
-2.73	-2.83	-2.78	-8.56	15959	<i>Ifit3</i>	3437	<i>IFIT3</i>	
-1.83	-1.88	-1.86	-2.70	15959	<i>Ifit3</i>	3437	<i>IFIT3</i>	
		-2.13	-3.30	76719	<i>1700081L11Rik</i>	284058	<i>KANSL1</i>	
-3.37	-2.56	-2.97	-1.45	69612	<i>2310037I24Rik</i>	54934	<i>KANSL2</i>	
-2.44	-2.59	-2.52	-4.07	69612	<i>2310037I24Rik</i>	54934	<i>KANSL2</i>	
-2.40	-2.16	-2.28	-1.16	69612	<i>2310037I24Rik</i>	54934	<i>KANSL2</i>	

-3.39	-1.39	-2.39	-8.97	16535	<i>Kcnq1</i>	3784	<i>KCNQ1</i>	Short QT syndrome-2, 609621; Long QT syndrome-1, 192500; Jervell and Lange-Nielsen syndrome, 220400; Atrial fibrillation familial, 3, 607554
-3.77	-2.67	-3.22	0.12	16663	<i>Krt13</i>	3860	<i>KRT13</i>	White sponge nevus, 193900
-1.53	-1.01	-1.27	-0.43	16663	<i>Krt13</i>	3860	<i>KRT13</i>	White sponge nevus, 193900
-2.23	-3.16	-2.70	-4.65	109594	<i>Lmo1</i>	4004	<i>LMO1</i>	
-1.48	-2.51	-1.99	-11.32	109594	<i>Lmo1</i>	4004	<i>LMO1</i>	
-3.25	-3.03	-3.14	-3.43	27756	<i>Lsm2</i>	57819	<i>LSM2</i>	
-3.97	-1.27	-2.62	-2.69	27756	<i>Lsm2</i>	57819	<i>LSM2</i>	
-3.29	-0.95	-2.12	-1.56	27756	<i>Lsm2</i>	57819	<i>LSM2</i>	
-2.91	-1.73	-2.32	-8.30	66373	<i>Lsm5</i>	23658	<i>LSM5</i>	
-2.57	-1.53	-2.05	-0.71	66373	<i>Lsm5</i>	23658	<i>LSM5</i>	
-1.82	-1.93	-1.88	-1.80	17754	<i>Mtap1a</i>	4130	<i>MAP1A</i>	
-2.40	-2.13	-2.26	1.18	17171	<i>Mas1</i>	4142	<i>MAS1</i>	
-1.85	-2.01	-1.93	0.34	17171	<i>Mas1</i>	4142	<i>MAS1</i>	
-4.35	-3.57	-3.96	-1.69	51812	<i>Mcrs1</i>	10445	<i>MCRS1</i>	
-1.90	-2.85	-2.38	-11.98	51812	<i>Mcrs1</i>	10445	<i>MCRS1</i>	
-1.98	-2.48	-2.23	-7.46	17349	<i>Mif1</i>	4291	<i>MLF1</i>	Leukemia acute myeloid, 601626
-3.37	-0.90	-2.13	-4.84	17751	<i>Mt3</i>	4504	<i>MT3</i>	
-1.71	-2.67	-2.19	-3.84	217166	<i>Nr1d1</i>	9572	<i>NR1D1</i>	
-2.67	-2.84	-2.76	0.38	192292	<i>Nrbp1</i>	29959	<i>NRBP1</i>	
-4.02	-2.89	-3.45	-6.61	77976	<i>Nuak1</i>	9891	<i>NUAK1</i>	
-4.66	-3.27	-3.97	0.36	237222	<i>Ofd1</i>	8481	<i>OFD1</i>	Oral-facial-digital syndrome 1, 311200; Joubert syndrome 10, 300804; Simpson-Golabi-Behmel syndrome, type 2, 300209
-2.48	-3.02	-2.75	-1.36	237222	<i>Ofd1</i>	8481	<i>OFD1</i>	Oral-facial-digital syndrome 1, 311200; Joubert syndrome 10, 300804; Simpson-Golabi-Behmel syndrome, type 2, 300209
-3.56	-1.71	-2.63	-2.21	237222	<i>Ofd1</i>	8481	<i>OFD1</i>	Oral-facial-digital syndrome 1, 311200; Joubert syndrome 10, 300804; Simpson-Golabi-Behmel syndrome, type 2, 300209
-1.44	-1.28	-1.36	-0.55	14539	<i>Opn1mw</i>	2652	<i>OPN1MW</i>	Colorblindness deutan, 303800; Blue cone monochromacy, 303700
-2.44	-2.27	-2.35	-4.65	12057	<i>Opn1sw</i>	611	<i>OPN1SW</i>	Colorblindness tritan, 190900
-1.81	-1.26	-1.54	3.01	18389	<i>Opr1</i>	4987	<i>OPRL1</i>	
-1.91	-1.23	-1.57	-0.69	404315	<i>Olfir372</i>	284383	<i>OR2Z1</i>	
-4.02	-2.64	-3.33	-8.95	258961	<i>Olfir631</i>	390059	<i>OR51M1</i>	
-1.45	-1.43	-1.44	-5.22	258961	<i>Olfir631</i>	390059	<i>OR51M1</i>	
-2.66	-1.61	-2.14	2.55	140795	<i>P2ry14</i>	9934	<i>P2RY14</i>	
-2.16	-1.79	-1.98	-0.27	140795	<i>P2ry14</i>	9934	<i>P2RY14</i>	
-2.90	-3.07	-2.98	-7.32	93714	<i>Pcdhga6</i>	56109	<i>PCDHGA6</i>	
-4.48	-2.42	-3.45	0.59	52023	<i>Pibf1</i>	10464	<i>PIBF1/CEP90</i>	
-2.10	-1.42	-1.76	-0.57	52023	<i>Pibf1</i>	10464	<i>PIBF1/CEP90</i>	
-1.49	-1.05	-1.27	0.46	52023	<i>Pibf1</i>	10464	<i>PIBF1/CEP90</i>	
-1.85	-2.35	-2.10	-1.88	20873	<i>Plk4</i>	10733	<i>PLK4</i>	
-2.06	-1.20	-1.63	-3.28	20873	<i>Plk4</i>	10733	<i>PLK4</i>	
-1.92	-2.07	-2.00	-4.10	74760	<i>Rab3il1</i>	5866	<i>RAB3IL1</i>	
-3.50	-1.92	-2.71	-11.58	19387	<i>Rangap1</i>	5905	<i>RANGAP1</i>	
-1.85	-1.21	-1.53	-4.93	19387	<i>Rangap1</i>	5905	<i>RANGAP1</i>	

-3.90	-0.81	-2.36	-6.62	19400	<i>Rapsn</i>	5913	<i>RAPSN</i>	Myasthenic syndrome congenital, associated with facial dysmorphism and acetylcholine receptor deficiency, 608931; Myasthenic syndrome congenital associated with acetylcholine receptor deficiency, 608931; Fetal akinesia deformation sequence, 208150
-3.74	-2.82	-3.28	-6.28	216227	<i>Slc17a8</i>	246213	<i>SLC17A8</i>	Deafness autosomal dominant 25, 605583
-1.89	-0.92	-1.41	-2.33	216227	<i>Slc17a8</i>	246213	<i>SLC17A8</i>	Deafness autosomal dominant 25, 605583
-1.87	-1.33	-1.60	-2.16	20526	<i>Slc2a2</i>	6514	<i>SLC2A2</i>	Fanconi-Bickel syndrome, 227810
-4.08	-2.08	-3.08	-1.07	83701	<i>Srrt</i>	51593	<i>SRRT</i>	
-3.66	-1.18	-2.42	0.42	83701	<i>Srrt</i>	51593	<i>SRRT</i>	
-2.50	-1.04	-1.77	-0.25	83701	<i>Srrt</i>	51593	<i>SRRT</i>	
-4.63	-2.60	-3.62	-10.04	66592	<i>Stoml2</i>	30968	<i>STOML2</i>	
-3.59	-1.07	-2.33	-6.17	67902	<i>Sumf2</i>	25870	<i>SUMF2</i>	
-2.16	-1.71	-1.93	-6.93	67902	<i>Sumf2</i>	25870	<i>SUMF2</i>	
-1.67	-2.44	-2.05	-1.42	78829	<i>Tsc22d4</i>	81628	<i>TSC22D4</i>	
-2.11	-0.94	-1.53	0.58	78829	<i>Tsc22d4</i>	81628	<i>TSC22D4</i>	
-1.59	-2.42	-2.01	-1.21	56085	<i>Ubqln1</i>	29979	<i>UBQLN1</i>	
-3.19	-1.06	-2.12	0.54	74716	<i>Wbp2nl</i>	164684	<i>WBP2NL</i>	
-3.12	-11.63	-7.38	-10.64	217232	<i>Cdc27</i>	996	<i>CDC27</i>	
-3.15	-10.10	-6.62	-5.02	59026	<i>Huwe1</i>	10075	<i>HUWE1</i>	Mental retardation X-linked syndromic Turner type, 300706
-6.72	-14.28	-10.50	-5.50	68988	<i>Prpf31</i>	26121	<i>PRPF31</i>	Retinitis pigmentosa 11, 600138
-5.00	-8.28	-6.64	-9.23	68988	<i>Prpf31</i>	26121	<i>PRPF31</i>	Retinitis pigmentosa 11, 600138
-4.48	-11.67	-8.07	-4.28	230596	<i>Prpf38a</i>	84950	<i>PRPF38A</i>	
-6.32	-12.70	-9.51	-5.28	68879	<i>Prpf6</i>	24148	<i>PRPF6</i>	Retinitis pigmentosa 60, 613983
-4.94	-8.14	-6.54	-3.90	68879	<i>Prpf6</i>	24148	<i>PRPF6</i>	Retinitis pigmentosa 60, 613983
-3.95	-7.82	-5.89	-2.99	68879	<i>Prpf6</i>	24148	<i>PRPF6</i>	Retinitis pigmentosa 60, 613983
-11.11	-17.74	-14.42	-15.39	192159	<i>Prpf8</i>	10594	<i>PRPF8</i>	Retinitis pigmentosa 13, 600059
-5.28	-13.24	-9.26	-7.82	192159	<i>Prpf8</i>	10594	<i>PRPF8</i>	Retinitis pigmentosa 13, 600059
-4.66	-7.79	-6.22	-12.88	56438	<i>Rbx1</i>	9978	<i>RBX1</i>	
-5.43	-6.74	-6.08	-10.09	28035	<i>Usp39</i>	10713	<i>USP39</i>	
-3.66	-6.12	-4.89	-6.59	28035	<i>Usp39</i>	10713	<i>USP39</i>	

Table 3-20. List of genes confirmed in the secondary screen. 68 genes were confirmed in the secondary screen with at least one siRNA duplex having an effect on %CSC. Targets from the enriched functional modules were not filtered for cell number (yellow shaded cells). 20 hits had OMIM annotations and 18 were present in the CiliaProteome v3.

3.3.2.3.3 Validation of secondary screen hits

Hits listed in **Table 3-20** included genes identified previously from functional genomics analyses of centriole biogenesis in human cells²⁸² including *PLK4* and *CEP120*, and other hits with known ciliary roles including *CCDC41*²⁸³ and *OFD1*²⁸⁴, suggested that the screen had high specificity for ciliary processes. Several PRPFs and ribonucleoproteins, including PRPF8, PRPF19 and LSM2, were also hits that

have been previously identified in a genome-wide screen of mediators of DNA-damage response²⁸⁵. Interestingly, our hits PRPF8 and PRPF38A have also been implicated in the process of centriolar under-duplication²⁸².

Based on significance, functional categorization (**Figure 3-55**), relevant literature information and the availability of validated antibodies, a total of 15 genes encoding GPCRs, PRPFs and predicted centrosomal proteins were chosen for further study (Prpf6, Prpf8, Prpf31, Prpf38a, Pibf1, C21orf2, Gpr20, Gpr173, Rapsn, Mas1, Oplr1, Crhr2, Plk4, Htr1b, P2ry14). For each of these genes a series of validation experiments was designed and conducted in collaboration with colleagues in the laboratory of Prof. Colin Johnson (University of Leeds). The reduction of mRNA transcript levels was assayed in IMCD3 cells by quantitative PCR if the negative control (siScr) cycle threshold (Ct) value was < 30 (performed by Subaa Natarajan). Transcripts quantity was compared between knockdown of specific gene and scrambled knockdown and unpaired two-tailed Student's t-test was calculated (**Figure 3-56**). A decrease in protein level was assayed by western immunodetection (with assistance of Warren Herridge) if suitable antibodies were available. The decrease in cilia number and a reduction in immunostaining of the cognate protein by IF microscopy was also confirmed where suitable antibodies were available (the author and Dr Gabrielle Wheway). Finally, expression of selected proteins (principally, GPCRs) in embryonic neocortex was assayed in mouse tissue sections (Dr Zakia Abdelhamed).

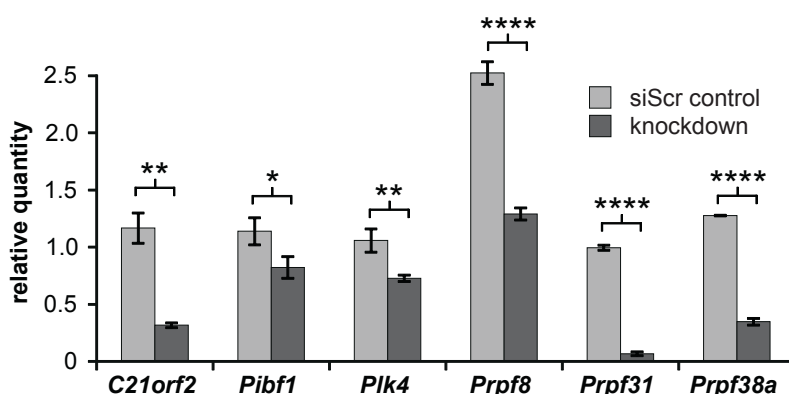


Figure 3-56. Graph representation of qPCR results for selected secondary screen hits. Relative quantities of the transcripts are shown of siScr versus knockdowns of selected genes. P values were as follows: *C21orf2* – 0.0034, *Pibf1* – 0.0225, *Plk4* – 0.0057, *Prpf8* - <0.0001, *Prpf31* - <0.0001, *Prpf38a* - <0.0001. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; **** $p < 0.0001$, for Student's t-test (unpaired, two-tailed). Error bars indicate s.d. for n=3 biological replicates.

3.3.2.3.3.1 Spliceosome components

Although Prpf6, Prpf8 and Prpf31 predominantly localised to nuclear speckles as expected (**Figure 3-57** and **3-59**), these proteins were also shown to co-localise to the base of the cilium in diverse human and mouse ciliated cell lines (**Figure 3-59**) and to proximal regions of the connecting cilium in the adult mouse retina (**Figure 3-60**; image obtained from Dr Zakia Abdelhamed). Reduction of the protein (**Figure 3-58** and **3-59**) causes cilia loss and, as expected, a decrease in the level of protein present at the nuclear speckles and base of the cilium (**Figure 3-57** and **3-59**).

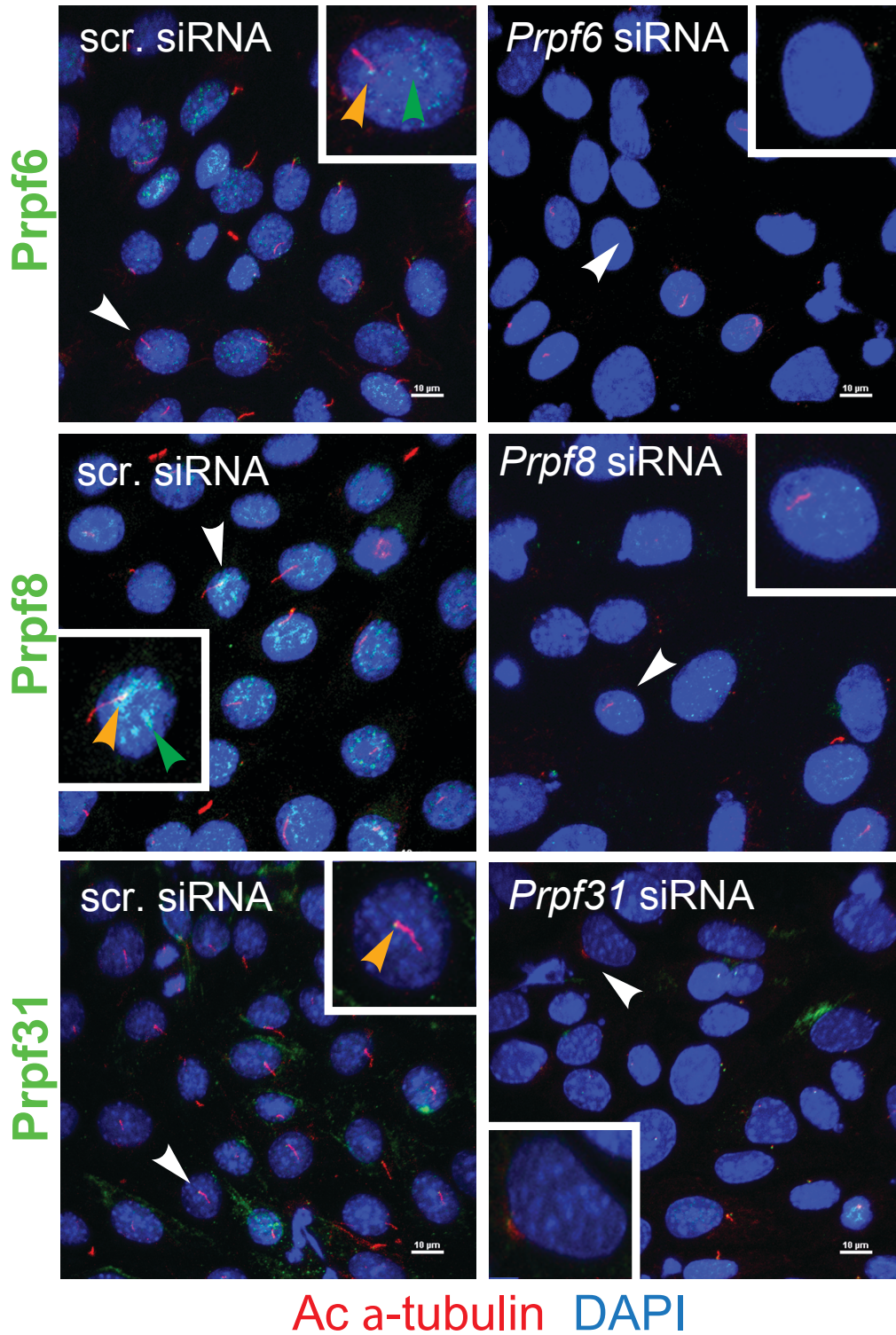


Figure 3-57. Knockdown of Prpf6, Prpf8 and Prpf31. Prpf proteins are stained in green and in siScrambled are shown to localise to nuclear speckles (green arrowheads) and at the base of cilium (orange arrowheads). Knockdown of these proteins shows a decrease in intensity at the nuclear speckles and at the base of the cilium. Cells pointed with white arrowheads were enlarged to show protein localisation to the base of the cilium (white squares).

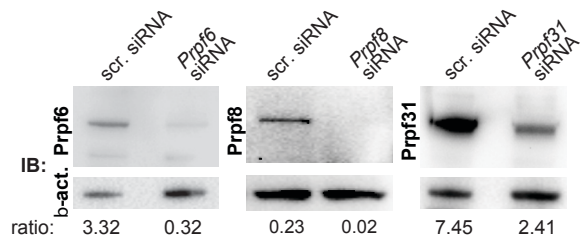


Figure 3-58. WB showing loss of protein after siRNA knockdown. Efficient knockdown of the indicated Prpf proteins was shown after siRNA knockdown in IMCD3 cells. The target to loading control (β -actin) ratio is shown below each panel, to express the efficiency of protein knockdown.

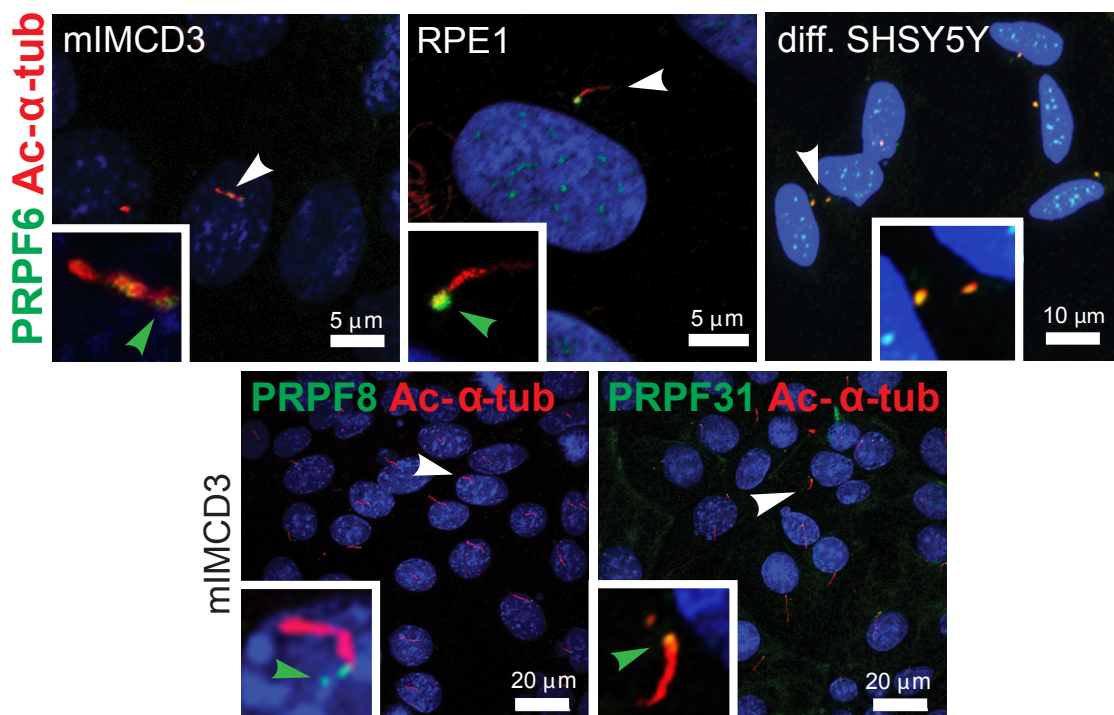


Figure 3-59. Localisation of Prpf proteins to the base of the primary cilium. Prpf proteins (green, indicated by green arrowheads in magnified insets) localised to the base of the cilium, visualized by staining for acetylated α -tubulin (red) in ciliated mouse IMCD3, human hTERT-RPE1 and differentiated human neuronal SHSY5Y cells. PRPF6 antibody was specific to mouse and human epitopes.

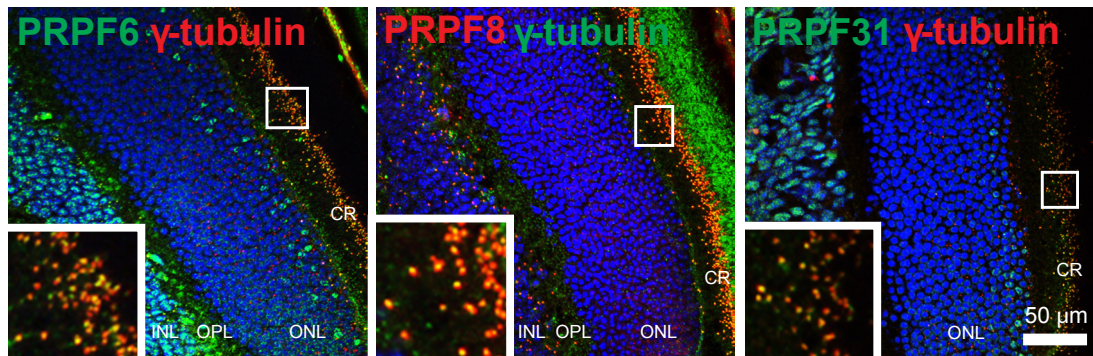


Figure 3-60. Localisation of Prpf proteins at the base of the connecting cilium in mouse retina. Prpf proteins (green or red as indicated) co-localised with the basal body marker γ -tubulin in mouse retinal sections. INL, inner nuclear layer; OPL, outer plexiform layer; ONL, outer nuclear layer; CR, connecting cilia. Scale bar = 50 μ m.

3.3.2.3.3.2 GPCRs

Many of the human GPCRs identified as hits in the secondary screen were neuroactive GPCRs, including serotonin receptor 1b (HTR1B), corticotropin releasing hormone receptor 2 (CRHR2) and the nociceptin receptor (opiate receptor-like 1; OPRL1). The localisation of these proteins was investigated in a human ciliated neuroblastoma cell-line (SHSY5Y) with neuronal-like characteristics that differentiate by extending neurites. The localisation of most of these GPCRs (HTR1B, P2RY14, MAS1, OPRL1, CRHR2) was shown to be at the base of the cilium in differentiated SHSY5Y cells (**Figure 3-61**), and to the connecting cilium in adult mouse retina (**Figure 3-62**; image obtained from Dr Zakia Abdelhamed). The results from the screen suggest that these GPCRs are required not only for ciliary signalling but also for cilia assembly.

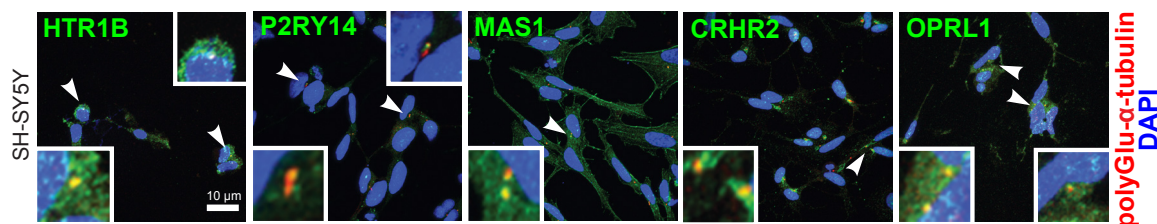


Figure 3-61. Localisation of GPCR to the base of cilium in differentiated SHSY5Y cells. Localization of the selected indicated GPCRs (green) to proximal or basal regions of primary cilia (polyglutamylated α -tubulin; red) in differentiated SHSY5Y neuronal cells. Magnified insets for selected cells (white arrowheads) are shown in white frames. Scale bar = 10 μ m.

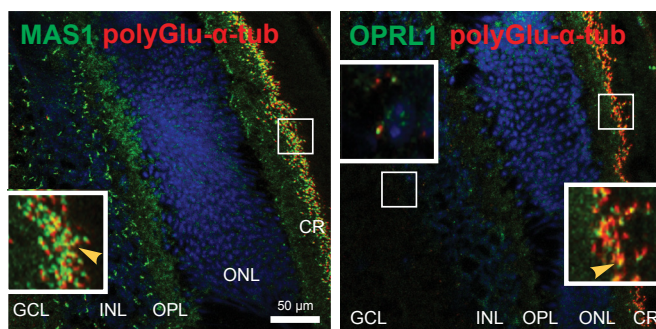


Figure 3-62. GPCR protein localisation in mouse retina sections. Localization of MAS1 and OPRL1 (green) to proximal regions of the connecting cilium (red; yellow arrowheads in magnified insets) in adult mouse retina. INL, inner nuclear layer; OPL, outer plexiform layer; ONL, outer nuclear layer; CR, connecting cilia. Scale bar = 50μm.

3.3.2.3.4 Utility of screen results

It was next investigated whether the list of validated ciliogenesis effectors (**Table 3-20**) could be used to prioritize predicted pathogenic variants identified from WES of ciliopathy patients. Mutations in two genes were identified by collaborators: mutations in *PIBF1* as a cause of JSRD (Dr Daniel Doherty, University of Washington) and mutations in *C21ORF2* as a cause of Jeune syndrome (Dr Hannah Mitchinson, University College London).

PIBF1 has independently been implicated in ciliogenesis²⁸⁶. Yeast two-hybrid (Y2H) analysis with the N-terminal part of TMEM237 as a bait identified *PIBF1* as a potential interactor (personal communication with Prof. Kym Boycott, University of Ottawa). To validate this interaction co-immunoprecipitation (coIP) assays of co-overexpressed epitope-tagged *PIBF1*-GFP and TMEM237C-terminal-deletion-FLAG were performed in HEK293 cells. There was a weak interaction observed between these proteins, when TMEM237 antibody pulled down *PIBF1*-GFP (**Figure 3-63** lane 6). *PIBF1*-GFP was also pulled down by GFP and *PIBF1* antibodies but not by the irrelevant antibody – MICU3 (**Figure 3-63** lanes 7, 8 and 9 respectively).

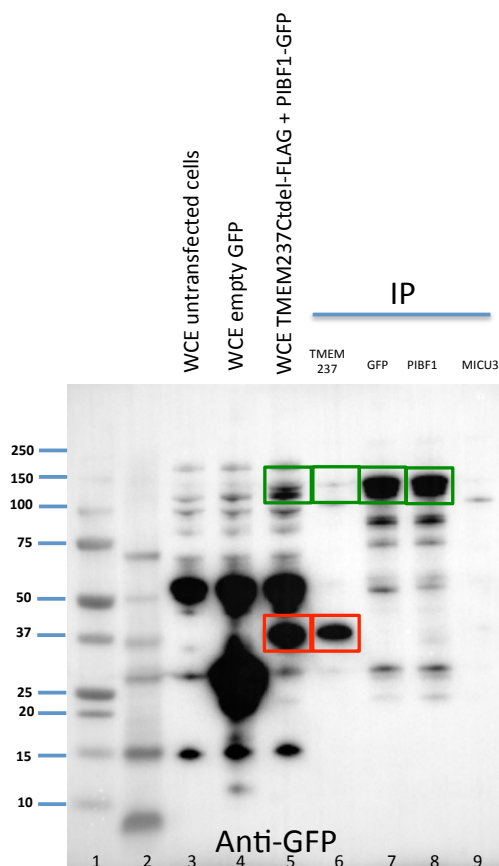


Figure 3-63. Co-IP experiment between co-overexpressed TMEM237-FLAG and PIBF1-GFP in HEK293 cells. TMEM237, PIBF1 and GFP antibodies pulled down PIBF1-GFP (green boxes). Bands in red boxes could indicate a shorter form of PIBF1-GFP, but lack of this band in lanes 7 and 8 disprove this. Lane 1 and 2 – protein markers, lane 3 – whole cell extract (WCE) of untransfected HEK293 cells, lane 4 – WCE of HEK293 cells transfected with empty GFP, lane 5 – WCE of HEK293 cells co-transfected with TMEM237Cterminal deletion-FLAG and PIBF1-GFP, lane 6 – pull down with TMEM237 antibody, lane 7 – pull down with GFP antibody, lane 8 – pull down with PIBF1 antibody, lane 9 – pull down with MICU3 antibody. Membrane was stained with mouse anti-GFP antibody.

3.3.3 Discussion

A whole mouse genome siRNA reverse genetics screen was conducted in ciliated IMCD3 cells. Cells were analysed for transcript expression (RNAseq) and copy number (arrayCGH). The siRNAs duplexes were analysed by BLAST in order to identify potential OTE and partial on-target effects, and were also analysed by GESS for predicted microRNA like effects. The controls used in the screen gave statistically significant results and were reliable and reproducible across the whole screen. Transcript loss of many genes was identified to have an effect on cilia as well as on cell number. The data was filtered in a number of ways to enable a meaningful and strict selection for hit validation in secondary screen studies. The different chemistry of siRNA duplexes allowed confirmation of 68 out of 192 hits in

the secondary screen. The selected hits taken forward to a tertiary screen were confirmed manually (confocal IF microscopy in 6 well plates instead of high content imaging in 96 well plates) and all of the selected proteins were shown to localise to the base of cilia. The utility of the screen data was confirmed by identification of *PIBF1* and *C21ORF2* to be mutated in known human ciliopathies.

Mouse inner medullary collecting duct cells were chosen as the model system for the screen, as these cells are widely used in the cilia biology research community and widely accepted as a model for ciliogenesis. Their advantages include good levels of ciliation, with one cilium per cell, with a greater cilia length (~8µm) compared to other model systems such as human hTERT-RPE1 cells (~4µm). The mIMCD3 cells are adherent and easy to culture, although cell number prior to siRNA transfection had to be strictly controlled. Too many seeded cells caused the cells to quickly become over-confluent. Since mIMCD3s cells have reduced contact inhibition and can grow in multiple layers and even begin to differentiate to form tubule-like structures. Over-confluency disturbed the cilia recognition high content imaging protocol that involved scanning the cell monolayer at specific, pre-defined focal planes. Although on coverslips these cells appear to have flat long cilia, in the 96-well format the cilia were perpendicular to their substrate, which allowed the easy recognition of cilia as a spot over a cell.

The methodology for IF staining in the screen was robust once initially optimised for high throughput cilia detection, but difficulties were encountered when a batch of acetylated α -tubulin mouse monoclonal antibody was exhausted and the new batch showed differences in staining. This required antibody re-optimisation and a delay in screening. A further delay was caused by the loss of transfection efficiency caused by a new batch of serum-free OptiMEM medium, which was observed by the loss of effect following *Pik1* knockdown (the transfection positive control). It was concluded that this batch of OptiMEM contained foetal bovine serum leading to a sudden drop in transfection efficiency. These obstacles led to the conclusion that all reagents for a screen should be ordered beforehand and originate from one manufacturing batch.

Out of the many positive controls for ciliogenesis defects, those targeting *Mks1*, *Rpgrip11* and *Ift88* were chosen (**Figure 3-34**). These genes are known to be involved in spectrum of ciliopathies including MKS, JSRD and Jeune syndrome and their protein function is well described. Their knockdown had graduated effects on %CSC with *siIft88* having the most significant effect on ciliogenesis.

Only one time point (72h siRNA incubation) was used to visualise cilia and this may not be reflected by a loss of protein levels since different proteins have different half-life. This could be a significant source of false negatives seen in this

screen, since many genes known to be important for ciliogenesis were not detected in the screen. For example many IFT genes were filtered out early in the analysis since they not only affect cilia number but also cell number. Although this was a consideration taken into account during the design of the screen, the optimal transfection time was chosen to ensure the optimal siRNA transfection efficiency (as measured by the effect of the positive control si*Plk1*). The massive format of the screen did not allow multiple incubation time points, but follow-up secondary screens with different siRNA incubation times are being considered.

Ciliary length is one of the metrics of cilia defects but was not assayed in this screen, as cilia were poking straight up and the length measurement was not possible. Imaging at another focal plane would have involved the storage and analysis of additional large data-sets and this was limited by the capacity of the existing screening facilities. Cilia intensity measured at the 'cilia' focal plane could be considered as a proxy indicator of ciliary length, but this metric needs further validation before it can be considered for future secondary screens.

In both primary and secondary screens, an enrichment for components of the spliceosome and GPCRs was observed. These novel findings implicate these functional modules in ciliogenesis and cilia maintenance. There are only few reports for the function of GPCRs in cilia^{281,287} and unveiling their function is of much interest. GPCRs have had ciliary localisation sequences identified, making them very interesting candidates for ciliary trafficking of membrane proteins³⁸. **Figures 3-61** and **3-62** also demonstrate localisation of selected GPCRs, that were hits in the siRNA screen, to the base of the primary cilium and to the connecting cilium in mouse retinal sections. Previous studies have hypothesised that GPCRs may play a sensory or signalling role in the ciliary vesicle or procilium^{288,289} prior to ciliogenesis, rather than during ciliogenesis itself. Their function is also known in Hedgehog signalling crucial for correct embryonic development^{287,290}. Transcript and protein levels for all GPCR genes in the studied model systems were both very low, precluding further validation by either Western blotting or qRT-PCR. One of the reasons for this could have been the model system that was used, namely kidney cells. The low transcript levels were confirmed in the RNAseq study and further methods to investigate this group of proteins are necessary. For further investigation of GPCRs more relevant model system should be used, such as differentiated neuronal-like cells from the SHSY5Y neuroblastoma cell-line. Further characterisation of the GPCR hits from this siRNA screen may contribute to understanding the complex ciliary functions in different organs.

The function of PRPFs is well-known and well-described in the spliceosome, but they were not previously reported to localise to or have any function in the

primary cilium. The data suggests that the PRPFs are required for ciliogenesis or cilia maintenance, but it remains unclear if there is any further function for those proteins at the base of the cilium. Unpublished protein-protein interaction data has shown that PRPF8 was pulled down in a series of TAP experiments by the IFT proteins IFT27 and IFT52 (n=3 biological replicates), and PRPF31 by the ciliary proteins IFT20, IFT27, IFT46, IFT52, TTC21B and CC2D2A (n=3; experiments performed by collaborators Dr Dorus Mans, Radboud University Nijmegen Medical Centre and Dr Karsten Boldt, University of Tübingen). This protein localisation data and the ciliogenesis defect caused by protein loss, suggests that PRPF6, PRPF8, PRPF31 and PRPF38A could be crucial structural components of the primary cilium. *PRPF6*, *PRPF8* and *PRPF31* are all mutated in autosomal dominant retinitis pigmentosa (RP types 60 MIM#613983, 13 MIM#600059 and 11 MIM#600138, respectively). The pathogenic mechanism for these RNA splicing-factor forms of RP remains poorly understood, and none have been characterised as retinal ciliopathies. Further functional work to investigate the function of PRPF proteins at the basal body is on-going and is being conducted by the author's colleague, Dr Gabrielle Wheway.

In summary, the siRNA screen represents a systematic and unbiased strategy to assess the contribution of every gene in the genome to the processes of ciliogenesis. The elaborate steps to fully characterise the cell-line and the siRNA library, as well as the careful optimisation of the transfection, staining and imaging assay protocols, were important to minimize the false positive rate for hits. In addition, the characterisation of the IMCD3 cell-line is a useful resource for further studies. The various protocols can be adapted to suit screens in other ciliated cell-lines, and have the potential for use in other high-throughput visual screens of, for example, small molecule effectors of cilium structure and length. Most importantly, this represents a novel and useful methodology which is a high-quality example of the application of the hypothesis-neutral strategy of reverse-genetic screening. The data highlights the opportunities of a synergistic analysis of large data-sets. The combination of genome-wide siRNA screening, WES and proteomics should lead to the identification of new disease-causing genes for diagnostic benefit, better refinement of patient phenotypes as an aid to prognosis and individualised clinical care, and to the discovery of new disease pathways that provide deeper insights into cilia biology.

4 Discussion

4.1 Mutation screening, founder mutations and genotype – phenotype correlations in MKS and associated ciliopathies

DNA samples were collected from ciliopathy patients and family members which led to a large patient cohort to be established. The initial investigation to identify causative mutations in the genes was undertaken at the University of Birmingham, where Prof. Colin Johnson's group was established. Part of the work presented in section 3.1 was a continuation of this work. The gene identification workflow included initial linkage analysis using microsatellite markers. These markers were chosen based on their proximal physical position to the locus under investigation with the markers chosen to be no further than 1Mbp from the coding sequence of the gene (**Table 3-2**). A high (>0.70) heterozygosity value was a second criterion determining the choice of the marker and was verified on the CEPH website. Fulfilling those two criteria was desirable but not always possible for each microsatellite. Crossovers between the marker and gene could also not be excluded, though highly unlikely.

Sanger sequencing of the known and candidate genes, at the time, revealed mutation in about 57% (n=38/67) of Leeds cohort families. This unselected MKS/JSRD cohort, including both consanguineous and non-consanguineous families from a range of ethnicities, suggests that other genes remain to be discovered. Sanger sequencing is a robust and reliable method of mutation identification but it is time consuming, costly and requires relatively large amounts of DNA. Prior genotyping allows linkage confirmation/exclusion that can decrease cost but may generate false negative/positive results. *TMEM67* was also initially sequenced in Birmingham/Leeds for only the first eighteen exons (out of total of twenty eight), as most of the reported mutations (at the time) were located in this part of the gene. Later this assumption was proved to be wrong as mutations in the C-terminal part of *TMEM67* were shown to be important in mediating protein-protein interactions²⁰⁶.

Genotype-phenotype correlations in the Leeds cohort confirmed previously reported observations^{110,221,291}. Patients with a liver phenotype should have *TMEM67* screening prioritised, while those with occipital encephalocele, bone dysplasia, laterality defects and polydactyly should be screened for *MKS1* first. Those with bone dysplasia, craniofacial abnormalities including micrognathia should

have *RPGRIP1L* screened first. More extensive correlations were often not possible since patient phenotypes were sometimes incompletely documented. The importance of detailed phenotype information should therefore be emphasized with referring clinicians. Extensive patient phenotyping would allow not only correlation with the genetic causes of the phenotype to be made, but would also decrease the time and cost of such analyses.

Mutation analysis in the Leeds cohort revealed common mutations in the known MKS genes. Pakistani patients with either linkage to the *MKS3* locus or with a liver phenotype should have mutational screening prioritised for two mutations: c.1575+1G>A and c.870-2A>G. For the same ethnic background, c.954delT in *CEP290*, c.1945C>T in *RPGRIP1L* or c.3540delA in *CC2D2A* should also be selected for screening. Patients of European origin should be initially screened for the “Finn major” *MKS1* common Finnish mutation c.1408-35_1408-7_del29. Also, the c.755T>C mutation in *TMEM67* appears to be common in combination with another missense change.

Yorkshire Regional Genetics Service (YRGS) prior to 2013 was offering NHS service testing of all exons of *MKS1*, *TMEM216* and *TMEM67*. Pre-screening was, and still is, the first pass of the screening process and common mutations were also prioritised. This included sequencing of the common Finnish mutation in *MKS1* and *CC2D2A* in Caucasians, and for Pakistani and other Asian patients common mutations in *TMEM67*, *RPGRIP1L*, *CC2D2A* and *NPHP3*. In practice, this strategy was not efficient and the focus of testing has moved to the use of WES of all samples (Dr Ian Berry, personal communication).

Particular interest was taken into *TMEM67*, as mutations in this gene are a major cause of MKS and MKS-like phenotype (**Figure 3-7a**). Reported mutations are quite evenly spread across the protein (**Figure 3-3**), but those specific to JSRD are located in amino acids 82-110 and 670-728. There are also changes that were reported in different ciliopathies, for example p.R208* was found in MKS, JSRD and NPHP patients. This truncating mutation occurs quite early in the protein sequence and would probably cause NMD and loss of the protein²⁹². The molecular outcome and patient phenotype would therefore be expected to be the same. However, differences in patient phenotype are observed that could suggest the modulating function of an additional modifier allele or alleles²³¹.

It is interesting to note that many ciliopathy and ciliary-related proteins interact and are reported to create functional modules that are localized to discrete structural regions of the cilium such as the TZ^{47,144,152,249}. The effect of modifier alleles may be to abrogate interactions between components of a functional

module, which may disrupt protein complexes or signalling pathways giving rise to the ciliopathy phenotype.

Four different heterozygous changes in six patients were identified, in the absence of a second detectable pathogenic mutation in the same gene or any other mutations in other MKS genes. These heterozygous alleles could be potential modifier alleles, but the possibility that a second pathogenic mutation occurred deep within introns or regulatory elements of the same MKS gene was not excluded. The molecular mechanism of additive ciliary protein mutations should therefore be explored to explain observed human phenotypes.

Many genetic changes are being identified while searching for true disease causing mutations and particular care should be taken to verify the pathogenicity of these changes. In section 3.1.4.1.1 it was described how mutations are being identified and verified in the YRGS while further confirmatory work was undertaken in the Leeds Institutes of Molecular Medicine. This is not a standard workflow for variant interpretation, although it shows how these two institutions can complement each other. Currently, the standard practice for YRGS is to conduct whole exome sequencing for each patient filtering for variants only in the known ciliopathy genes. If no mutation is identified, information sheets and consent forms for analysis on a research basis are sent to the referring clinician to be offered to the families. Appropriate DNA samples are then submitted to the Sir Jules Thorn autozygosity mapping project (<http://autozygosity.org/>) for WES, or the WES data files (fastq files) are supplied from the diagnostic lab for alignment and variant calling on a research basis. If a variant is predicted to be causative, and it is absent in databases of common variants, then segregation is checked in available family members. The variant is also excluded in a panel of ethnically-matched control DNA samples, and other affected patients presenting the same phenotype are screened for mutations in the gene. The sub-cellular localisation of the encoded protein can also be investigated and RNA levels can be checked (for predicted null alleles such as stop, frame-shift or splice mutations).

There is a need to establish protocols for robust, high-throughput mutation identification in the known MKS/JSRD genes. To enable that technologies like MIPS¹⁶⁶ or targeted capture of ciliopathy genes^{293,294} should be considered. Databases, like one available for *CEP290* (<http://www.retina-international.org/files/sci-news/cep290mu.htm>), should be established to allow easy access to combined genetic and phenotypic information to research and diagnostic laboratories, as well as to clinicians.

4.2 Autozygosity mapping and candidate gene screening

Ciliopathies are very genetically heterogeneous conditions, and only a small proportion of genes with causative mutations have been identified to date. However, in some populations (for example, in Bradford which has a sizable population of individuals of Pakistani ancestry), the incidence of the most severe ciliopathy, MKS, can be as high as 1 in 3000 births. The identification of new ciliopathy genes is therefore clinically essential to provide options for accurate, molecular diagnostic testing, including carrier testing, prenatal diagnosis, accurate gene counselling and even, in the future, pre-nuptial testing. These general aims take advantage of modern technologies that allow high-throughput sequencing of genes and therefore quicker mutation identification. The molecular basis of ciliopathies is still unclear, and the investigation of proteins involved in primary cilia function and maintenance would help to understand this complicated and crucial process during embryonic development.

Based on SNP whole-genome genotyping studies obtained from five consanguineous patients, a shared homozygosity region was identified on chromosome 12. This locus was also confirmed by microsatellite data (**Figure 3-8**). In both analyses no shared haplotypes were found despite the same ethnical background for some of the patients. Candidate genes within the minimum region were prioritised and sequenced for variant identification. No changes were identified, although genes chosen for sequencing were later shown to be a cause of MKS/JSRD (*TCTN1*⁴⁷, *TCTN2*¹⁰⁴ and *TCTN3*¹⁴⁷).

The most common molecular causes of MKS were already identified and MKS7-12 were found to be a rare cause of this phenotype (<1%). To improve mutation identification in new genes that cause the MKS phenotype, a network of collaborators was established. Private mutations in single cases for putative functional candidate genes are difficult to prove to be truly pathogenic without extensive functional work. Therefore, the sequence analyses collected by the University of Leeds and elsewhere were combined with those collected by collaborators to improve the chances of independent replication. No pathogenic biallelic variants were identified in *CEP164*²³⁶, *TTC21B*¹⁸², *TMEM107*, *PROM1*, *CENPF*²⁹⁵, *TCTN3* (although a mutation in sample 330 in this gene was later identified)¹⁴⁷, *PDE6D*¹⁴⁹ or *CSPP1*^{108,150,151}. Mutations in these genes were either already reported to cause a ciliopathy phenotype or have manuscripts in preparation, and these represent the true functional candidates, and therefore were screened in the University of Leeds cohort.

WES to identify new causes of MKS/JSRD did not reveal mutations in any strong functional candidate genes. Samples sequenced on an Illumina MiSeq showed too low read depth to make any conclusions, and these samples will have to be re-sequenced on an Illumina HiSeq2500. The apparent absence of putative pathogenic variants could be explained by errors made during data analysis, but this is unlikely as the same analysis protocol was applied to other samples and it allowed successful pathogenic variant identification. A second explanation could be that the phenotype was caused by a microdeletion or insertion that was not mapped during sequence alignment by the Noalign program. It is therefore advisable to re-analyze the original .bam file for CNVs in the known MKS/JSRD genes using, for example, the Integrative Genomics Viewer (IGV, Broad Institute). If this proved negative, then low read-depth WGS could be done for these samples in order to identify potential copy number variants (CNVs) or other genomic variants that include the non-coding regions of the known MKS/JSRD genes. CNVs at the resolution of arrayCGH (40kb) could also be analysed in order to identify potential larger deletions/insertions.

The list of MKS/JSRD candidate genes found in the University of Leeds patients was passed to collaborators to look into their WES data, but no additional potential pathogenic variants were identified. The remaining MKS/JSRD families with no causative mutations found to date should be exome sequenced to identify mutations in new candidate genes. Although this task has proved to be more difficult than expected, the molecular causes of the phenotype should be identified in these families. The new “SureSelectQXT” reagent (Agilent Technologies) allows the use of as little as 50ng of genomic DNA (2µl of 25ng/µl), enabling the investigation of samples that have had too little DNA to be included in previous studies. Also current kits for WGS, for CNV investigations, require only 200ng of DNA or less (TruSeq Nano DNA Sample Prep Kit, Illumina).

Many of the known ciliopathy proteins from vertebrates or other species such as *C. elegans* have now ascribed roles in cilium biogenesis and signalling at the ciliary TZ^{47,49,152,249}. To gain further insights into the molecular aetiology of ciliopathies, the identification of mutations in a novel and previously uncharacterized gene, *TMEM237*, was performed. Although mutations in *TMEM237* only account for <1% cases of our JSRD/MKS cohort, the high carrier frequency of the c.52C>T mutation in a Canadian Hutterite population (approximately 1:15) means that carrier testing, cascade testing in families, and prenatal diagnosis would be of high clinical utility in this population¹⁴⁴. Families with JSRD will also benefit from improved diagnosis and accurate genetic counselling.

The function of the TZ-localized TMEM237 protein was then characterised. Loss of human *TMEM237* results in a failure of ciliogenesis and de-regulation of both canonical and non-canonical/PCP Wnt signalling pathways. These findings are strikingly similar to previous studies of both TMEM216 and TMEM67^{99,205}, two additional transmembrane proteins associated with JSRD and MKS. These membrane proteins have been suggested to be non-canonical Wnt receptors that regulate the RhoA pathway and thus mediate the cytoskeleton rearrangements required for basal body docking at the apical region of the cells prior to ciliogenesis⁹⁹. Importantly, in a complementary zebrafish model system, disruption of TMEM237 was demonstrated to cause developmental (convergent extension) phenotypes comparable to those obtained upon abrogation of TZ-localized proteins, including TMEM216 and TMEM67¹⁴⁴. Hence, the similarities between TMEM237, TMEM216 and TMEM67 in cellular localization, protein structure, role in cilium formation/function, and clinical phenotypes indicate that TMEM237 likely functions in the same pathway to regulate ciliogenesis and signalling.

One of the central questions that remain about the molecular aetiology and underlying phenotypic variability of the JBTS-MKS-NPHP ciliopathy spectrum is whether the ciliary defects arise from disruption in one or more macromolecular assemblies, or modules, comparable to that of the 'BBSome'⁵⁰, an oligomeric protein complex containing proteins associated with BBS. Support for this model comes from the study of *C. elegans* which has suggested the existence of two genetically-defined modules: one, being an "MKS" module (consisting of MKS-1, MKSR-1/B9D1, MKSR-2/B9D2, MKS-3, and MKS-6), and the other, an "NPHP" module, consisting of NPHP-2 and NPHP-4²⁴⁹. Specifically, disruption of any combination of genes within either module abrogates a "ciliary gate" function but does not significantly impair the structure of most cilia. In contrast, disrupting any combination of two genes, one from each module, causes TZ structure anomalies which are concomitant with loss of basal body/TZ anchoring to the membrane and ciliary axoneme structure defects. Additionally, disruption of individual mammalian TZ proteins (such as MKS1 and TMEM67) in patients or mouse models is on the whole phenotypically less severe than that observed for disruption of IFT, which causes global defects in ciliogenesis. This suggests a functional redundancy in mammals and could be exposed by analysis of mouse double knockouts.

Previous studies^{47,152,249,296} support the concept of human ciliopathies being caused by sorting defects at the TZ, and the ciliary gate playing a crucial role in cilia assembly and selective regulation of cilia protein content. What remains unexplained is the function of ciliary modules in mediating cilia trafficking, how these could regulate the signalling cascades that are mediated by cilia, and the

connection with other complexes such as the inversin module and the BBSome. It seems likely that the elucidation of these mechanistic details will begin to explain the phenotypic variability and pleiotropy of human ciliopathies. These could arise from either the diverse requirements of the protein composition of the TZ in different tissues, or the influence of modifier alleles in interacting components of individual functional modules. Another study by Li et al.²⁹⁶ highlighted the complex, dynamic nature of the TZ and a possible role of this region of the cilium in G₁/S checkpoint control. Linking the cilium, cell cycle control and extracellular cues of signalling pathways will be a further field of intensive future work, and will no doubt bring further surprises in understanding of the complex ultrastructure of the primary cilium.

With the ever-increasing power and affordability of genetic sequencing technologies, there is now the clear opportunity for the further rapid and robust identification of mutations in patients referred for a defined condition. As a prerequisite, there remains a pressing clinical need for the dissemination of mutations identified on a research basis, and the establishment of databases that provide detailed clinical phenotypes and allelic series for specific genes, as described in section 3.1.4 and presented in **Table 3-5**. Those databases should use a standardised terminology, for both variant calling and phenotype description, as uniform databases can be easily searched (<http://dgv.tcag.ca/dgv/app/home>, <http://www.ncbi.nlm.nih.gov/dbvar/>, <http://www.lovd.nl/3.0/home>, <http://evs.gs.washington.edu/EVS/>, <https://decipher.sanger.ac.uk/>, <http://exac.broadinstitute.org/>).

Increased power of variant filtering was shown when family trios were whole exome sequenced²⁹⁷⁻³⁰⁰. Parental samples allow the exclusion of about 78% of biallelic variants and filtering based on variants observed in unaffected sibling will exclude 57% of changes. However parental and siblings variants together will eliminate about 83% of biallelic variants, highlighting the power of this strategy (numbers obtained from Dr David Parry, University of Leeds, personal communication).

Large number of the obtained possible causative variants emphasizes the difficulties in assessing the pathogenic potential of variants of unknown significance (VOUS) in both basic research and clinical diagnosis of not only the ciliopathies but, more generally, for autosomal recessive and more so for autosomal dominant conditions³⁰¹. This is a key challenge that needs to be addressed to prevent false positive results from hindering the translation of research findings into clinical diagnostic testing and to enable the further biological understanding of disease mechanisms. For JSRD and MKS, the major causative genes are now known and

there are good insights into the function for some of the encoded proteins. In interpreting VOUS in the ciliopathies, researchers and clinical scientists can now take full advantage of public data-sets of genomic variation, functional genomic data and model-organism phenotypes. However, it remains important that variants identified from the many targeted screening and WES experiments for JSRD and MKS are reported as public data-sets with a standard nomenclature that follows published guidelines. Leiden Open Variation Database (LOVD) v3.0 would be a good database to use since it provides both gene- and patient-specific data storage, including datasets from WES and WGS³⁰².

4.3 Whole genome siRNA screen

The other approach to identify new causes of the MKS/JSRD phenotype was to apply an unbiased 'reverse genetics' whole genome siRNA screen. Loss of primary cilia was a known cellular phenotype, observed consistently in ciliopathy patient cells, and was assessed in ciliated mIMCD3 cells. All annotated mouse genes were knocked down and cilia number per cell was assessed. This methodology involved the screening of over nineteen thousands mouse genes in duplicate and, by necessity, had to be limited to answer a simple biological question: are cilia present or not? Although mRNA turnover varies between different genes, the high-throughput methodology used a single siRNA concentration and incubation time. This could lead to false negative results, as more rapid mRNA transcription would preferentially deplete the siRNAs and return the cell to a normal transcript level. For these targets shorter incubation times and/or increase in siRNA concentration would minimize the potential for false negative results. This could be tested by qRT-PCRs, but the necessity to do this in high throughput for all targeted transcripts precludes this approach on grounds of time, cost and available resources. The screen methodology also only used three out of a possible four fluorescence channels on grounds of data storage space constrains. About five hundred 96-well plates were scanned, excluding those that failed and were used for optimisation purposes. For each well, six fields of view were imaged with three fluorescence channels. Each plate contained about 5Gb of image data, and the whole screen was therefore about 3TB of data storage. If storage constraints were not an issue, the fourth channel could have been used to assay an additional marker of ciliary dysfunction such as Smoothened or active β -catenin.

The screen led to the identification of unexpected connections between cilia biology and other cellular or molecular processes³⁰², demonstrating the power of an unbiased screen. GPCRs have not been ascribed a role in ciliogenesis by previous

screens and biochemical approaches, due to the technical constraints of studying transmembrane proteins. Neuroactive GPCRs, such as those identified in the screen, may play sensory or signalling roles in the ciliary vesicle or procilium at the earliest stages of ciliogenesis both in the developing brain and in other tissues ³⁰³. In support of this hypothesis, adenylyl cyclase III co-localises with GPCRs at proximal ciliary regions in the developing neocortex, and over-expression of specific ciliary GPCRs including the 5-HT₆ serotonin receptor in cortical neurons causes cilia elongation ³⁰⁴.

Although the role of the PRPFs in the spliceosome is well-known and well-studied, a group of seven splicing factors including PRPF6, PRPF8, PRPF31 and PRPF38a were identified to be required for ciliogenesis. PRPFs mutated in retinitis pigmentosa localise specifically to the base of primary cilia and the photoreceptor connecting cilium (**Figure 3-60**). Some of the splicing factors could therefore fulfil an additional ciliary function independent of their nuclear role in splicing. In support of this notion, two previous studies of the centrosomal proteome have suggested that splicing factors, including two of the splicing factors identified in our screen (PRPF6 and PRPF8), may be true centrosomal proteins ^{305,306}. However, it seems less speculative to hypothesize that these proteins may be required for the correct splicing of an unidentified subset of genes that are involved in cilia formation. In support of the possible involvement of pre-mRNA processing factors in ciliogenesis, previous screens have identified that several splicing factor hits are important in related cellular processes of microtubule formation and regulation, namely centriolar biogenesis (PRPF8, PRPF38A) ²⁸² and cell division (PRPF6, PRPF8) ³⁰⁷.

In relation to this, several hits in the screen have a known function in the ubiquitin proteasome system (UPS), reflecting the importance of specific proteostasis mechanisms in mediating or regulating ciliogenesis which could be mediated through, for example, ubiquitylation or SUMOylation of ciliary transcription factors or chaperoning of ciliary proteins to the base of the cilium. It is known that proteasome function is necessary for cilia maintenance. Appropriate protein levels are necessary to maintain cell homeostasis and uninterrupted ciliogenesis, therefore a discrete balance between ciliary protein ubiquitination ³⁰⁸⁻³¹⁰ and de-ubiquitination ³¹¹ has to be maintained. Remarkably, 5/7 of the splicing factor hits (LSM2, PRPF6, PRPF8, PRPF31 and USP39) are implicated in the ubiquitin-dependent regulation of the spliceosome ³¹². It is interesting to note that the interaction of PRPF8 with ubiquitinated PRPF3 is regulated by the deubiquitinating enzyme USP4, and that loss of USP4 prevents the correct splicing of mRNAs including those for α -tubulin ³¹². Thus, UPS and/or PRPF proteins could act as multifunctional “nexus molecules”

that are involved in multiple aspects of proteostasis of ciliary proteins, or, more specifically, ensure the correct splicing of transcripts encoding proteins important for ciliogenesis including structural components of the cilium such as α -tubulin^{312,313}.

The specificity and clinical utility of the screen suggests that it is a useful tool for disease-gene discovery. When combined with large variant datasets, for example WES of ciliopathy patients, the functional data from the siRNA screen allowed the filtering and prioritization of variants to identify pathogenic mutations. Although previous studies have identified both *PIBF1* and *C21orf2* as plausible functional candidate genes for ciliopathies^{286,314,315}, this screen demonstrated the utility and validity of a systems biology approach by the identification of mutations in these genes in JSRD and Jeune syndrome patients, respectively. Recently, another gene from the list of highest confidence hits (*CEP120*) was shown to be mutated as a cause of Jeune asphyxiating thoracic dystrophy²⁷⁸.

4.4 Future plans

In future investigations the other datasets obtained from the whole genome siRNA screen could be further explored. So far only hits with cilia loss and no effect on cell number were analysed (with exception to splicosome and UPS hits). This allowed a focus on those hits that were relevant to ciliogenesis. It has been reported that loss of cilia proteins can also affect cell number, as these two processes are tightly correlated^{316,317}. Hits with cilia z score < cell number z score by -2 were shown to be highly enriched in known ciliopathy and cilia-related genes including *CEP164*, *CEP290*, *GLIS2*, *MKKS*, *PKHD1*, *TTC21B*, *TULP3* and others. Therefore this subset of genes could be particularly enriched in good ciliopathy candidate genes and could be used as a filter in WES variant prioritisation. Furthermore, these hits would be excellent candidates for investigating ciliary function and dissection of ciliary biological process.

A further interesting subset included hits that increased cilia number. This included *ROCK2*, *POC5*, *GATA4*, *AKAP5* and *AKAP8*. DAVID analysis showed that this subset is enriched in microtubule and cytoskeleton components, suggesting their function in determining correct cilia number. The underlying mechanisms could involve correct microtubule organisation and dynamics as well as actin cytoskeleton remodelling for correct centriole division and apical cell surface transport. Actin stress fibres were observed in ciliopathy patients fibroblasts^{99,205} and may be

formed by ROCK2, a direct target of the small GTPase Rho, that in turn is a downstream target of non-canonical Wnt signalling. No multiciliated cells were observed in patients fibroblasts, although this phenotype has been observed in TMEM67-mutated patient kidney cells³¹⁸. POC5, on the other hand, was reported to have a crucial role in centriole maturation³¹⁹, but no links of this protein loss and multiciliation have been reported previously. The understanding of centriole biogenesis, actin cytoskeleton remodelling and microtubule dynamics would be therefore an interesting subject of the future study.

The whole genome siRNA screen did not assess the efficiency of knockdown since this was impossible to measure it for each gene. Protein levels were only assayed in the subsequent validation steps. This protein level represented a population of cells but not each cell individually, although the cilia presence/absence was investigated per each cell in a field of view on 96-well plate. The kinetics of the reagents was also not known per each target and the amount of siRNA duplexes and incubation times were generalised over the whole mouse genome. These problems suggest that future screens should use new genomic editing methodologies. The CRISPR-Cas9 (Clustered Regularly Interspaced Palindromic Repeats) system could introduce in-del mutations on a genomic level, allowing the loss of the mRNA due to NMD. The CRISPR-Cas9 system was originally identified in bacteria as a defence mechanism against foreign pathogens³²⁰. It is an RNA-guided nuclease system for the targeted introduction of double-stranded breaks in DNA^{321,322}. This system requires two components: a guide RNA that binds specifically to the target sequence and Cas9, an endonuclease that introduces double stranded breaks. Specificity of this method is determined by guide RNA, which consists of twenty nucleotides and binds to the DNA immediately upstream of a Protospacer Adjacent Motif (PAM). The PAM is a three-nucleotide sequence (NGG, when N is any base pair) specific to *Streptococcus pyogenes* Cas9 (the one used most commonly for this method) that is present in human genome at high frequency of, on average, once every twelve nucleotides³²².

CRISPR-Cas9 technology relies on two mechanisms for double stranded break repair: non-homologous end joining (NHEJ) and homology-directed repair (HDR). In the former, due to mistakes made by the polymerase, a small number of nucleotides may be inserted or deleted causing a frame-shift and a premature stop codon. It is therefore advised to design targets at the beginning of the coding sequence to maximize the possibility of NMD. HDR, on the other hand, requires a repair template that is co-transfected into cells. This template contains two arms of about fifty nucleotides with a high degree of homology directly upstream and downstream of the predicted double-strand break. The repair template also

contains a specific mutation that will be inserted into the genome. This method allows the introduction of a specific mutation into a cellular or animal model system, enabling the further investigation of the mutated protein in a biologically relevant system at physiological levels of expression. However, both methods of repair require careful optimization and selection of cells transfected with components of the CRISPR-Cas9 system. Current methods are facilitated by either selecting cells transfected with a guide RNA construct containing either an antibiotic resistance gene or GFP allowing cell sorting using FACS. Monoclonal cells containing the genomic change are then grown, and the function of the protein or phenotype of interest can be further investigated. This methodology would allow the better understanding of protein function using a reverse genetic modification system that is not dependent on the protein level, unlike siRNA knockdown which is transient and variable in efficiency. Furthermore, this system also enables the modelling of specific patient mutations that could elucidate the pathological mechanisms of a specific disease.

This method could be used in screens, for example for chemical components rescuing cilia loss to facilitate the need for future possible ciliopathy therapeutics. In order to achieve this the cell line would be treated with CRISPR-Cas9 targeting for example, *CEP290* transcript (*CEP290* is mutated in a large spectrum of ciliopathies). Cells would be selected for those with successful transfection and introduction of frameshift mutation. Few monoclonal cell colonies would be grown and genetic and phenotypic state would be characterised. Cells with the most robust phenotype would be chosen for the screen and treated with available chemical library screen for cilia rescue. Similarly, the same model system could be used for a screen of genetic interactants by transient knockdown of potential TZ components. It is known that CEP290 localises to the basal body/TZ region³²³ and the list of predicted TZ components was obtained from collaborator Maritjn Huijnen (Radboud University Nijmegen Medical Centre). Furthermore, screens could be designed for the effects on the downstream effectors of cilogenesis defects like nuclear localisation of β -catenin or ciliary localisation of Smothened.

Research into disease mechanisms should be conducted in the best available model system. Often the only accessible patient cells are dermal fibroblasts, which although a useful model system, are not derived from a tissue that is affected in most ciliopathy phenotypes. One solution to this potential problem would be to reprogram fibroblasts into inducible pluripotent stem cells (iPSCs), and then differentiate them into the desired cell types³²⁴⁻³²⁷. Although this method is time consuming, taking several months to obtain cultures or organoids of the required

cell type, it could be used in combination with genomic editing to create the most relevant model system for physiological processes investigation³²⁷.

The further search for human mutations using WES and WGS should also be a priority, since a description of the genetic causes of MKS/JSRD is still incomplete. The identification of new disease genes and pathways may enable the future rational in design of therapeutics to modify or treat cystic kidney disease, retinal degeneration³²⁸ or ciliopathy disease progression, or improve the long-term outlook of patients with these conditions. Since JSRD and MKS are predominantly autosomal recessive, they are caused by the absence of normal protein (rather than the presence of an abnormal protein) so they can, in principle, be corrected by gene-replacement. In the first demonstration of this strategy in a ciliopathy mouse model, McIntyre et al. used the well-established *Ift88*^{Tg737Rpw} mouse mutant with many typical phenotypic features including anosmia³²⁹. Remarkably, the adenoviral-mediated expression of IFT88 (a protein essential for IFT in cilia) in fully-differentiated olfactory sensory neurons of mutant mice was sufficient to restore both ciliary structures and rescue olfactory function³²⁹. Patients in whom mutations are found can therefore be given a clearer prognosis and prioritised for these new treatments, making inherited disease a top priority for further characterization.

The further convergence of genetic data is envisaged with other independent lines of evidence that assess the pathogenic potential of a variant. These will include comparative genomic approaches and bioinformatic datasets, although the experimental validation of the damaging impact of a candidate variant still provides the most definitive proof³³⁰⁻³³². Future studies should use assays of patient-derived cells or tissues, as well as well-established cell or animal models of gene function^{333,334}. Not only will these lead to improvements in the diagnosis and clinical management of ciliopathy patients, but they will also provide pre-clinical models to test future therapeutic interventions³³⁵.

In the future, targeted therapies such as antisense oligonucleotide (AON) therapy³³⁶ and stop-codon read-through therapy³³⁷ may be beneficial for patients with suitable splice-site or nonsense mutations. AON therapy was applied to correct a hypomorphic 126bp exon insertion in *CEP290* caused by the deep intronic mutation (c.2991+1655A>G) in LCA patients. This insertion caused the introduction of a nonsense codon and loss of transcript. AON targeting this insertion prevented its introduction into the transcript and enabled the production of an in-frame mRNA and protein³³⁸. This makes the JSRD and MKS group of ciliopathies a top priority for further genetic and functional characterization in order to prioritize patients for these potential treatments.

In conclusion, this thesis highlight the important clinical utility of combining different systems biology approaches, namely high content functional genomic screening and WES. The major findings of the chosen approach are the identification of new ciliary roles for well-studied proteins, identified new disease-causing genes for diagnostic benefit, allowed the refinement of patient phenotypes as an aid to prognosis and individualized clinical care, and has highlighted new potential disease pathways that should provide deeper insights into cilia biology.

References

1. Dobell, C. *Antony van Leeuwenhoek and his 'Little Animals'*, (Harcourt, Brace and Co, New York, 1932).
2. Muller, O.F. *Animalcula infusoria; fluvia tilia et marina, que detexit, systematice descripsit et ad vivum delineari curavit.*, (Typis N. Molleri, Havniae, 1786).
3. Singla, V. & Reiter, J.F. The primary cilium as the cell's antenna: signaling at a sensory organelle. *Science* **313**, 629-33 (2006).
4. Badano, J.L., Mitsuma, N., Beales, P.L. & Katsanis, N. The ciliopathies: An emerging class of human genetic disorders. *Annual Review of Genomics and Human Genetics* **7**, 125-148 (2006).
5. Williams, C. *et al.* MKS and NPHP modules cooperate to establish basal body/transition zone membrane associations and ciliary gate function during ciliogenesis. *J Cell Biol* **192**, 1023 - 1041 (2011).
6. Pan, J. & Snell, W. The Primary Cilium: Keeper of the Key to Cell Division. *Cell* **129**, 1255-1257 (2007).
7. Praetorius, H.A. & Spring, K.R. A physiological view of the primary cilium. *Annual Review of Physiology* **67**, 515-529 (2005).
8. Fliegauf, M., Benzing, T. & Omran, H. When cilia go bad: cilia defects and ciliopathies. *Nat Rev Mol Cell Biol* **8**, 880-893 (2007).
9. Badano, J.L., Teslovich, T.M. & Katsanis, N. The centrosome in human genetic disease. *Nat Rev Genet* **6**, 194-205 (2005).
10. Ishikawa, H. & Marshall, W.F. Ciliogenesis: building the cell's antenna. *Nat Rev Mol Cell Biol* **12**, 222-234 (2011).
11. Banizs, B. *et al.* Dysfunctional cilia lead to altered ependyma and choroid plexus function, and result in the formation of hydrocephalus. *Development* **132**, 5329-39 (2005).
12. Fisch, C. & Dupuis-Williams, P. Ultrastructure of cilia and flagella - back to the future! *Biol Cell* **103**, 249-70 (2011).
13. Spoon, C. & Grant, W. Biomechanics of hair cell kinocilia: experimental measurement of kinocilium shaft stiffness and base rotational stiffness with Euler-Bernoulli and Timoshenko beam analysis. *J Exp Biol* **214**, 862-70 (2011).
14. Wheway, G., Parry, D.A. & Johnson, C.A. The role of primary cilia in the development and disease of the retina. *Organogenesis* **10**, 69-85 (2014).
15. Louvi, A. & Grove, E.A. Cilia in the CNS: the quiet organelle claims center stage. *Neuron* **69**, 1046-60 (2011).
16. Paintrand, M., Moudjou, M., Delacroix, H. & Bornens, M. Centrosome organization and centriole architecture: Their sensitivity to divalent cations. *Journal of Structural Biology* **108**, 107-128 (1992).
17. Rosenbaum, J.L. & Witman, G.B. Intraflagellar transport. *Nat Rev Mol Cell Biol* **3**, 813-825 (2002).
18. Satir, P. & Christensen, S.T. Overview of Structure and Function of Mammalian Cilia. *Annual Review of Physiology* **69**, 377-400 (2007).
19. Gilula, N.B. & Satir, P. The ciliary necklace. *The Journal of Cell Biology* **53**, 494-509 (1972).

20. Fisch, C. & Dupuis-Williams, P. Ultrastructure of cilia and flagella – back to the future! *Biology of the Cell* **103**, 249-270 (2011).
21. Quinlan, R.J., Tobin, J.L. & Beales, P.L. Modeling ciliopathies: primary cilia in development and disease. *Mouse Models of Developmental Genetic Disease* **84**, 249-+ (2008).
22. Graser, S. *et al.* Cep164, a novel centriole appendage protein required for primary cilium formation. *The Journal of Cell Biology* **179**, 321-330 (2007).
23. Ishikawa, H., Kubo, A., Tsukita, S. & Tsukita, S. Odf2-deficient mother centrioles lack distal/subdistal appendages and the ability to generate primary cilia. *Nat Cell Biol* **7**, 517-524 (2005).
24. Singla, V., Romaguera-Ros, M., Garcia-Verdugo, J.M. & Reiter, J.F. Odf1, a human disease gene, regulates the length and distal structure of centrioles. *Dev Cell* **18**, 410-24 (2010).
25. Sillibourne, J.E. *et al.* Assessing the localization of centrosomal proteins by PALM/STORM nanoscopy. *Cytoskeleton (Hoboken)* **68**, 619-27 (2011).
26. Deane, J.A., Cole, D.G., Seeley, E.S., Diener, D.R. & Rosenbaum, J.L. Localization of intraflagellar transport protein IFT52 identifies basal body transitional fibers as the docking site for IFT particles. *Current Biology* **11**, 1586-1590 (2001).
27. O'Toole, E.T., Giddings Jr, T.H. & Dutcher, S.K. Understanding Microtubule Organizing Centers by Comparing Mutant and Wild-Type Structures with Electron Tomography. in *Methods in Cell Biology*, Vol. Volume 79 (ed. McIntosh, J.R.) 125-143 (Academic Press, 2007).
28. Garcia-Gonzalo, F.R. & Reiter, J.F. Scoring a backstage pass: mechanisms of ciliogenesis and ciliary access. *J Cell Biol* **197**, 697-709 (2012).
29. Horst, C.J., Forestner, D.M. & Besharse, J.C. Cytoskeletal-membrane interactions: a stable interaction between cell surface glycoconjugates and doublet microtubules of the photoreceptor connecting cilium. *The Journal of Cell Biology* **105**, 2973-2987 (1987).
30. Hu, Q. *et al.* A Septin Diffusion Barrier at the Base of the Primary Cilium Maintains Ciliary Membrane Protein Distribution. *Science* **329**, 436-439 (2010).
31. Craige, B. *et al.* CEP290 tethers flagellar transition zone microtubules to the membrane and regulates flagellar protein content. *The Journal of Cell Biology* **190**, 927-940 (2010).
32. Sillibourne, J.E. *et al.* Primary ciliogenesis requires the distal appendage component Cep123. *Biol Open* **2**, 535-45 (2013).
33. Kozminski, K.G., Johnson, K.A., Forscher, P. & Rosenbaum, J.L. A motility in the eukaryotic flagellum unrelated to flagellar beating. *Proc Natl Acad Sci USA* **90**, 5519-23 (1993).
34. Pedersen, L.B., Veland, I.R., Schroder, J.M. & Christensen, S.T. Assembly of primary cilia. *Developmental Dynamics* **237**, 1993-2006 (2008).
35. Deretic, D. *et al.* Rhodopsin C terminus, the site of mutations causing retinal disease, regulates trafficking by binding to ADP-ribosylation factor 4 (ARF4). *Proceedings of the National Academy of Sciences of the United States of America* **102**, 3301-3306 (2005).
36. Mazelova, J. *et al.* Ciliary targeting motif VxPx directs assembly of a trafficking module through Arf4. *EMBO J* **28**, 183-192 (2009).
37. Geng, L. *et al.* Polycystin-2 traffics to cilia independently of polycystin-1 by using an N-terminal RVxP motif. *Journal of Cell Science* **119**, 1383-1395 (2006).

38. Berbari, N.F., Johnson, A.D., Lewis, J.S., Askwith, C.C. & Mykytyn, K. Identification of Ciliary Localization Sequences within the Third Intracellular Loop of G Protein-coupled Receptors. *Molecular Biology of the Cell* **19**, 1540-1547 (2008).
39. Guo, W., Sacher, M., Barrowman, J., Ferro-Novick, S. & Novick, P. Protein complexes in transport vesicle targeting. *Trends Cell Biol* **10**, 251-5 (2000).
40. Jahn, R. & Scheller, R.H. SNAREs [mdash] engines for membrane fusion. *Nat Rev Mol Cell Biol* **7**, 631-643 (2006).
41. Nachury, M.V., Seeley, E.S. & Jin, H. Trafficking to the Ciliary Membrane: How to Get Across the Periciliary Diffusion Barrier? *Annual Review of Cell and Developmental Biology* **26**, 59-87 (2010).
42. Follit, J.A., Tuft, R.A., Fogarty, K.E. & Pazour, G.J. The Intraflagellar Transport Protein IFT20 Is Associated with the Golgi Complex and Is Required for Cilia Assembly. *Molecular Biology of the Cell* **17**, 3781-3792 (2006).
43. Chiba, S., Amagai, Y., Homma, Y., Fukuda, M. & Mizuno, K. NDR2-mediated Rabin8 phosphorylation is crucial for ciliogenesis by switching binding specificity from phosphatidylserine to Sec15. *EMBO J* **32**, 874-85 (2013).
44. Heider, M.R. & Munson, M. Exorcising the exocyst complex. *Traffic* **13**, 898-907 (2012).
45. Knödler, A. *et al.* Coordination of Rab8 and Rab11 in primary ciliogenesis. *Proceedings of the National Academy of Sciences* **107**, 6346-6351 (2010).
46. Kim, J., Krishnaswami, S.R. & Gleeson, J.G. CEP290 interacts with the centriolar satellite component PCM-1 and is required for Rab8 localization to the primary cilium. *Human Molecular Genetics* **17**, 3796-3805 (2008).
47. Garcia-Gonzalo, F. *et al.* A transition zone complex regulates mammalian ciliogenesis and ciliary membrane composition. *Nat Genet* **43**, 776 - 784 (2011).
48. Nalefski, E.A. & Falke, J.J. The C2 domain calcium-binding motif: Structural and functional diversity. *Protein Science* **5**, 2375-2390 (1996).
49. Bachmann-Gagescu, R. *et al.* The ciliopathy gene cc2d2a controls zebrafish photoreceptor outer segment development through a role in Rab8-dependent vesicle trafficking. *Human Molecular Genetics* **20**, 4041-4055 (2011).
50. Nachury, M.V. *et al.* A Core Complex of BBS Proteins Cooperates with the GTPase Rab8 to Promote Ciliary Membrane Biogenesis. *Cell* **129**, 1201-1213 (2007).
51. Jin, H. *et al.* The Conserved Bardet-Biedl Syndrome Proteins Assemble a Coat that Traffics Membrane Proteins to Cilia. *Cell* **141**, 1208-1219 (2010).
52. Lehtreck, K.-F. *et al.* The *Chlamydomonas reinhardtii* BBSome is an IFT cargo required for export of specific signaling proteins from flagella. *The Journal of Cell Biology* **187**, 1117-1132 (2009).
53. Cai, H., Reinisch, K. & Ferro-Novick, S. Coats, tethers, Rabs, and SNAREs work together to mediate the intracellular destination of a transport vesicle. *Dev Cell* **12**, 671-82 (2007).
54. Milenkovic, L., Scott, M.P. & Rohatgi, R. Lateral transport of Smoothed from the plasma membrane to the membrane of the cilium. *J Cell Biol* **187**, 365-74 (2009).
55. Huang, K., Diener, D.R. & Rosenbaum, J.L. The ubiquitin conjugation system is involved in the disassembly of cilia and flagella. *The Journal of Cell Biology* **186**, 601-613 (2009).

56. Ingham, P.W. & McMahon, A.P. Hedgehog signaling in animal development: paradigms and principles. *Genes Dev* **15**, 3059-87 (2001).
57. Nusslein-Volhard, C. & Wieschaus, E. Mutations affecting segment number and polarity in *Drosophila*. *Nature* **287**, 795-801 (1980).
58. Riddle, R.D., Johnson, R.L., Laufer, E. & Tabin, C. Sonic hedgehog mediates the polarizing activity of the ZPA. *Cell* **75**, 1401-16 (1993).
59. Echelard, Y. *et al.* Sonic hedgehog, a member of a family of putative signaling molecules, is implicated in the regulation of CNS polarity. *Cell* **75**, 1417-30 (1993).
60. Krauss, S., Concordet, J.P. & Ingham, P.W. A functionally conserved homolog of the *Drosophila* segment polarity gene *hh* is expressed in tissues with polarizing activity in zebrafish embryos. *Cell* **75**, 1431-44 (1993).
61. Nakano, Y. *et al.* A protein with several possible membrane-spanning domains encoded by the *Drosophila* segment polarity gene *patched*. *Nature* **341**, 508-13 (1989).
62. Hui, C.C. & Angers, S. Gli proteins in development and disease. *Annu Rev Cell Dev Biol* **27**, 513-37 (2011).
63. Beachy, P.A., Karhadkar, S.S. & Berman, D.M. Tissue repair and stem cell renewal in carcinogenesis. *Nature* **432**, 324-31 (2004).
64. Oro, A.E. *et al.* Basal cell carcinomas in mice overexpressing sonic hedgehog. *Science* **276**, 817-21 (1997).
65. Modena, P. *et al.* Identification of tumor-specific molecular signatures in intracranial ependymoma and association with clinical characteristics. *J Clin Oncol* **24**, 5223-33 (2006).
66. Liu, S., Dontu, G. & Wicha, M.S. Mammary stem cells, self-renewal pathways, and carcinogenesis. *Breast Cancer Res* **7**, 86-95 (2005).
67. Boletta, A. Emerging evidence of a link between the polycystins and the mTOR pathways. *Pathogenetics* **2**, 6 (2009).
68. Bell, P.D. *et al.* Loss of primary cilia upregulates renal hypertrophic signaling and promotes cystogenesis. *J Am Soc Nephrol* **22**, 839-48 (2011).
69. Boehlke, C. *et al.* Primary cilia regulate mTORC1 activity and cell size through *Lkb1*. *Nat Cell Biol* **12**, 1115-22 (2010).
70. Zullo, A. *et al.* Kidney-specific inactivation of *Odf1* leads to renal cystic disease associated with upregulation of the mTOR pathway. *Hum Mol Genet* **19**, 2792-803 (2010).
71. Dawe, H.R. *et al.* The Meckel-Gruber Syndrome proteins MKS1 and meckelin interact and are required for primary cilium formation. *Human Molecular Genetics* **16**, 173-186 (2007).
72. Du, E. *et al.* Evidence that TMEM67 causes polycystic kidney disease through activation of JNK/ERK-dependent pathways. *Cell Biol Int* **37**, 694-702 (2013).
73. Ezratty, E.J. *et al.* A role for the primary cilium in Notch signaling and epidermal differentiation during skin development. *Cell* **145**, 1129-41 (2011).
74. Nakaya, M.A. *et al.* Wnt3a links left-right determination with segmentation and anteroposterior axis elongation. *Development* **132**, 5425-36 (2005).
75. Boskovski, M.T. *et al.* The heterotaxy gene GALNT11 glycosylates Notch to orchestrate cilia type and laterality. *Nature* **504**, 456-9 (2013).
76. O'Rourke, M.P., Soo, K., Behringer, R.R., Hui, C.C. & Tam, P.P. Twist plays an essential role in FGF and SHH signal transduction during mouse limb development. *Dev Biol* **248**, 143-56 (2002).

77. Miller, K.A. *et al.* Cauli: a mouse strain with an Ift140 mutation that results in a skeletal ciliopathy modelling Jeune syndrome. *PLoS Genet* **9**, e1003746 (2013).
78. Tabler, J.M. *et al.* Fuz mutant mice reveal shared mechanisms between ciliopathies and FGF-related syndromes. *Dev Cell* **25**, 623-35 (2013).
79. Badouel, C., Garg, A. & McNeill, H. Herding Hippos: regulating growth in flies and man. *Curr Opin Cell Biol* **21**, 837-43 (2009).
80. Habbig, S. *et al.* NPHP4, a cilia-associated protein, negatively regulates the Hippo pathway. *J Cell Biol* **193**, 633-42 (2011).
81. Habbig, S. *et al.* The ciliopathy disease protein NPHP9 promotes nuclear delivery and activation of the oncogenic transcriptional regulator TAZ. *Hum Mol Genet* **21**, 5528-38 (2012).
82. Frank, V. *et al.* Mutations in NEK8 link multiple organ dysplasia with altered Hippo signalling and increased c-MYC expression. *Hum Mol Genet* **22**, 2177-85 (2013).
83. Clement, C.A. *et al.* TGF-beta signaling is associated with endocytosis at the pocket region of the primary cilium. *Cell Rep* **3**, 1806-14 (2013).
84. Clement, D.L. *et al.* PDGFRalpha signaling in the primary cilium regulates NHE1-dependent fibroblast migration via coordinated differential activity of MEK1/2-ERK1/2-p90RSK and AKT signaling pathways. *J Cell Sci* **126**, 953-65 (2013).
85. Adams, M., Smith, U.M., Logan, C.V. & Johnson, C.A. Recent advances in the molecular pathology, cell biology and genetics of ciliopathies. *Journal of Medical Genetics* **45**, 257-267 (2008).
86. Szymanska, K. & Johnson, C.A. The transition zone: an essential functional compartment of cilia. *Cilia* **1**, 10 (2012).
87. Reiter, J.F., Blacque, O.E. & Leroux, M.R. The base of the cilium: roles for transition fibres and the transition zone in ciliary formation, maintenance and compartmentalization. *EMBO Rep* **13**, 608-18 (2012).
88. Blacque, O.E. & Sanders, A.A. Compartments within a compartment: what *C. elegans* can tell us about ciliary subdomain composition, biogenesis, function, and disease. *Organogenesis* **10**, 126-37 (2014).
89. Leigh, M.W. *et al.* Clinical and genetic aspects of primary ciliary dyskinesia/Kartagener syndrome. *Genetics in Medicine* **11**, 473-487 10.1097/GIM.0b013e3181a53562 (2009).
90. Knowles, M.R., Daniels, L.A., Davis, S.D., Zariwala, M.A. & Leigh, M.W. Primary ciliary dyskinesia. Recent advances in diagnostics, genetics, and characterization of clinical disease. *Am J Respir Crit Care Med* **188**, 913-22 (2013).
91. Kurkowiak, M., Zietkiewicz, E. & Witt, M. Recent advances in primary ciliary dyskinesia genetics. *J Med Genet* (2014).
92. Salonen, R. The Meckel syndrome: clinicopathological findings in 67 patients. *Am J Med Genet* **18**, 671-89 (1984).
93. Salonen, R. & Paavola, P. Meckel syndrome. *J Med Genet* **35**, 497-501 (1998).
94. Alexiev, B.A., Lin, X., Sun, C.-C. & Brenner, D.S. Meckel-Gruber Syndrome: Pathologic Manifestations, Minimal Diagnostic Criteria, and Differential Diagnosis. *Archives of Pathology & Laboratory Medicine* **130**, 1236-1238 (2006).
95. Chen, C.P. Meckel syndrome: genetics, perinatal findings, and differential diagnosis. *Taiwan J Obstet Gynecol* **46**, 9-14 (2007).

96. Boltshauser, E. & Isler, W. Joubert syndrome: episodic hyperpnea, abnormal eye movements, retardation and ataxia, associated with dysplasia of the cerebellar vermis. *Neuropadiatrie* **8**, 57-66 (1977).
97. Smith, U.M. *et al.* The transmembrane protein meckelin (MKS3) is mutated in Meckel-Gruber syndrome and the wpk rat. *Nat Genet* **38**, 191-6 (2006).
98. Kyttila, M. *et al.* MKS1, encoding a component of the flagellar apparatus basal body proteome, is mutated in Meckel syndrome. *Nat Genet* **38**, 155-7 (2006).
99. Valente, E.M. *et al.* Mutations in TMEM216 perturb ciliogenesis and cause Joubert, Meckel and related syndromes. *Nature Genetics* **42**, 619-U100 (2010).
100. Baala, L. *et al.* Pleiotropic effects of CEP290 (NPHP6) mutations extend to Meckel syndrome. *American Journal of Human Genetics* **81**, 170-179 (2007).
101. Delous, M. *et al.* The ciliary gene RPGRIP1L is mutated in cerebello-oculorenal syndrome (Joubert syndrome type B) and Meckel syndrome. *Nat Genet* **39**, 875-81 (2007).
102. Tallila, J., Jakkula, E., Peltonen, L., Salonen, R. & Kestila, M. Identification of CC2D2A as a Meckel syndrome gene adds an important piece to the ciliopathy puzzle. *American Journal of Human Genetics* **82**, 1361-1367 (2008).
103. Bergmann, C. *et al.* Loss of nephrocystin-3 function can cause embryonic lethality, meckel-gruber-like syndrome, situs inversus, and renal-hepatic-pancreatic dysplasia. *American Journal of Human Genetics* **82**, 959-970 (2008).
104. Shaheen, R. *et al.* A TCTN2 mutation defines a novel Meckel Gruber syndrome locus. *Hum Mutat* **32**, 573-8 (2011).
105. Hopp, K. *et al.* B9D1 is revealed as a novel Meckel syndrome (MKS) gene by targeted exon-enriched next-generation sequencing and deletion analysis. *Human Molecular Genetics* **20**, 2524-2534 (2011).
106. Dowdle, W.E. *et al.* Disruption of a ciliary B9 protein complex causes Meckel syndrome. *Am J Hum Genet* **89**, 94-110 (2011).
107. Shaheen, R., Ansari, S., Mardawi, E.A., Alshammari, M.J. & Alkuraya, F.S. Mutations in TMEM231 cause Meckel-Gruber syndrome. *J Med Genet* **50**, 160-2 (2013).
108. Shaheen, R. *et al.* Mutations in CSPP1, encoding a core centrosomal protein, cause a range of ciliopathy phenotypes in humans. *Am J Hum Genet* **94**, 73-9 (2014).
109. Otto, E.A. *et al.* Hypomorphic mutations in meckelin (MKS3/TMEM67) cause nephronophthisis with liver fibrosis (NPHP11). *J Med Genet* **46**, 663-70 (2009).
110. Iannicelli, M. *et al.* Novel TMEM67 mutations and genotype-phenotype correlates in Meckelin-related ciliopathies. *Hum Mutat* **31**, E1319 - E1331 (2010).
111. Szymanska, K. *et al.* Founder mutations and genotype-phenotype correlations in Meckel-Gruber syndrome and associated ciliopathies. *Cilia* **1**, 18 (2012).
112. Consugar, M.B. *et al.* Molecular diagnostics of Meckel-Gruber syndrome highlights phenotypic differences between MKS1 and MKS3. *Human Genetics* **121**, 591-599 (2007).

113. Paavola, P., Salonen, R., Weissenbach, J. & Peltonen, L. The locus for Meckel syndrome with multiple congenital anomalies maps to chromosome 17q21-q24. *Nat Genet* **11**, 213-5 (1995).
114. Roume, J. *et al.* A gene for Meckel syndrome maps to chromosome 11q13. *Am J Hum Genet* **63**, 1095-101 (1998).
115. Morgan, N.V. *et al.* A novel locus for Meckel-Gruber syndrome, MKS3, maps to chromosome 8q24. *Hum Genet* **111**, 456-61 (2002).
116. Molin, A. *et al.* 12q21 Microdeletion in a fetus with Meckel syndrome involving CEP290/MKS4. *Eur J Med Genet* **56**, 580-3 (2013).
117. Romani, M. *et al.* Mutations in B9D1 and MKS1 cause mild Joubert syndrome: expanding the genetic overlap with the lethal ciliopathy Meckel syndrome. *Orphanet J Rare Dis* **9**, 72 (2014).
118. Joubert, M., Eisenring, J.J., Robb, J.P. & Andermann, F. Familial agenesis of the cerebellar vermis. A syndrome of episodic hyperpnea, abnormal eye movements, ataxia, and retardation. *Neurology* **19**, 813-25 (1969).
119. Parisi, M.A. Clinical and molecular features of Joubert syndrome and related disorders. *American Journal of Medical Genetics Part C: Seminars in Medical Genetics* **151C**, 326-340 (2009).
120. Parisi, M.A., Doherty, D., Chance, P.F. & Glass, I.A. Joubert syndrome (and related disorders) (OMIM 213300). *Eur J Hum Genet* **15**, 511-21 (2007).
121. Satran, D., Pierpont, M.E. & Dobyns, W.B. Cerebello-oculo-renal syndromes including Arima, Senior-Loken and COACH syndromes: more than just variants of Joubert syndrome. *Am J Med Genet* **86**, 459-69 (1999).
122. Valente, E.M. *et al.* Description, nomenclature, and mapping of a novel cerebello-renal syndrome with the molar tooth malformation. *Am J Hum Genet* **73**, 663-70 (2003).
123. Natacci, F. *et al.* Patient with large 17p11.2 deletion presenting with Smith-Magenis syndrome and Joubert syndrome phenotype. *Am J Med Genet* **95**, 467-72 (2000).
124. Saar, K. *et al.* Homozygosity mapping in families with Joubert syndrome identifies a locus on chromosome 9q34.3 and evidence for genetic heterogeneity. *Am J Hum Genet* **65**, 1666-71 (1999).
125. Lagier-Tourenne, C. *et al.* Homozygosity mapping of a third Joubert syndrome locus to 6q23. *J Med Genet* **41**, 273-7 (2004).
126. Bennett, C.L. *et al.* Joubert syndrome: a haplotype segregation strategy and exclusion of the zinc finger protein of cerebellum 1 (ZIC1) gene. *Am J Med Genet A* **125A**, 117-24; discussion 117 (2004).
127. Keeler, L.C. *et al.* Linkage analysis in families with Joubert syndrome plus oculo-renal involvement identifies the CORS2 locus on chromosome 11p12-q13.3. *Am J Hum Genet* **73**, 656-62 (2003).
128. Parisi, M.A. *et al.* The NPHP1 gene deletion associated with juvenile nephronophthisis is present in a subset of individuals with Joubert syndrome. *Am J Hum Genet* **75**, 82-91 (2004).
129. Ferland, R.J. *et al.* Abnormal cerebellar development and axonal decussation due to mutations in AH11 in Joubert syndrome. *Nat Genet* **36**, 1008-13 (2004).
130. Valente, E.M. *et al.* Mutations in CEP290, which encodes a centrosomal protein, cause pleiotropic forms of Joubert syndrome. *Nat Genet* **38**, 623-5 (2006).

131. Sayer, J.A. *et al.* The centrosomal protein nephrocystin-6 is mutated in Joubert syndrome and activates transcription factor ATF4. *Nat Genet* **38**, 674-81 (2006).
132. Baala, L. *et al.* The Meckel-Gruber syndrome gene, MKS3, is mutated in Joubert syndrome. *American Journal of Human Genetics* **80**, 186-194 (2007).
133. Arts, H.H. *et al.* Mutations in the gene encoding the basal body protein RPGRIP1L, a nephrocystin-4 interactor, cause Joubert syndrome. *Nat Genet* **39**, 882-8 (2007).
134. Cantagrel, V. *et al.* Mutations in the cilia gene ARL13B lead to the classical form of Joubert syndrome. *Am J Hum Genet* **83**, 170-9 (2008).
135. Gorden, N.T. *et al.* CC2D2A is mutated in Joubert syndrome and interacts with the ciliopathy-associated basal body protein CEP290. *Am J Hum Genet* **83**, 559-71 (2008).
136. Bielas, S.L. *et al.* Mutations in INPP5E, encoding inositol polyphosphate-5-phosphatase E, link phosphatidylinositol signaling to the ciliopathies. *Nat Genet* **41**, 1032-6 (2009).
137. Kroes, H.Y. *et al.* Cerebral, cerebellar, and colobomatous anomalies in three related males: Sex-linked inheritance in a newly recognized syndrome with features overlapping with Joubert syndrome. *Am J Med Genet A* **135**, 297-301 (2005).
138. Coene, K.L. *et al.* OFD1 is mutated in X-linked Joubert syndrome and interacts with LCA5-encoded lebercilin. *Am J Hum Genet* **85**, 465-81 (2009).
139. Edvardson, S. *et al.* Joubert syndrome 2 (JBTS2) in Ashkenazi Jews is associated with a TMEM216 mutation. *Am J Hum Genet* **86**, 93-7 (2010).
140. Dafinger, C. *et al.* Mutations in KIF7 link Joubert syndrome with Sonic Hedgehog signaling and microtubule dynamics. *J Clin Invest* **121**, 2662-7 (2011).
141. Alazami, A.M. *et al.* Molecular characterization of Joubert syndrome in Saudi Arabia. *Hum Mutat* **33**, 1423-8 (2012).
142. Lee, J.E. *et al.* CEP41 is mutated in Joubert syndrome and is required for tubulin glutamylation at the cilium. *Nature Genetics* **advance online publication**(2012).
143. Lee, J.H. *et al.* Evolutionarily assembled cis-regulatory module at a human ciliopathy locus. *Science* **335**, 966-9 (2012).
144. 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* **89**, 713 - 730 (2011).
145. Srour, M. *et al.* Mutations in C5ORF42 cause Joubert syndrome in the French Canadian population. *Am J Hum Genet* **90**, 693-700 (2012).
146. Dixon-Salazar, T.J. *et al.* Exome sequencing can improve diagnosis and alter patient management. *Sci Transl Med* **4**, 138ra78 (2012).
147. Thomas, S. *et al.* TCTN3 mutations cause Mohr-Majewski syndrome. *Am J Hum Genet* **91**, 372-8 (2012).
148. Srour, M. *et al.* Mutations in TMEM231 cause Joubert syndrome in French Canadians. *J Med Genet* **49**, 636-41 (2012).
149. Thomas, S. *et al.* A homozygous PDE6D mutation in Joubert syndrome impairs targeting of farnesylated INPP5E protein to the primary cilium. *Hum Mutat* **35**, 137-46 (2014).
150. Akizu, N. *et al.* Mutations in CSPP1 Lead to Classical Joubert Syndrome. *The American Journal of Human Genetics* **94**, 80-86 (2014).

151. Tuz, K. *et al.* Mutations in CSPP1 Cause Primary Cilia Abnormalities and Joubert Syndrome with or without Jeune Asphyxiating Thoracic Dystrophy. *The American Journal of Human Genetics* **94**, 62-72 (2014).
152. Sang, L. *et al.* Mapping the NPHP-JBTS-MKS protein network reveals ciliopathy disease genes and pathways. *Cell* **145**, 513 - 528 (2011).
153. Hsiao, Y.-C. *et al.* Ahi1, whose human ortholog is mutated in Joubert syndrome, is required for Rab8a localization, ciliogenesis and vesicle trafficking. *Human Molecular Genetics* **18**, 3926-3941 (2009).
154. Chih, B. *et al.* A ciliopathy complex at the transition zone protects the cilia as a privileged membrane domain. *Nat Cell Biol* **14**, 61-72 (2012).
155. Higginbotham, H. *et al.* Arl13b in primary cilia regulates the migration and placement of interneurons in the developing cerebral cortex. *Dev Cell* **23**, 925-38 (2012).
156. Humbert, M.C. *et al.* ARL13B, PDE6D, and CEP164 form a functional network for INPP5E ciliary targeting. *Proc Natl Acad Sci U S A* **109**, 19691-6 (2012).
157. Sergouniotis, P.I. *et al.* Biallelic variants in TLL5, encoding a tubulin glutamylase, cause retinal dystrophy. *Am J Hum Genet* **94**, 760-9 (2014).
158. Cui, C. *et al.* Disruption of Mks1 localization to the mother centriole causes cilia defects and developmental malformations in Meckel-Gruber syndrome. *Dis Model Mech* **4**, 43-56 (2011).
159. Abdelhamed, Z.A. *et al.* Variable expressivity of ciliopathy neurological phenotypes that encompass Meckel–Gruber syndrome and Joubert syndrome is caused by complex de-regulated ciliogenesis, Shh and Wnt signalling defects. *Human Molecular Genetics* **22**, 1358-1372 (2013).
160. Weatherbee, S.D., Niswander, L.A. & Anderson, K.V. A mouse model for Meckel syndrome reveals Mks1 is required for ciliogenesis and Hedgehog signaling. *Human Molecular Genetics* **18**, 4565-4575 (2009).
161. Putoux, A. *et al.* KIF7 mutations cause fetal hydrolethals and acrocallosal syndromes. *Nat Genet* **43**, 601-6 (2011).
162. Wheway, G. *et al.* Aberrant Wnt signalling and cellular over-proliferation in a novel mouse model of Meckel–Gruber syndrome. *Developmental Biology* **377**, 55-66 (2013).
163. Leightner, A.C. *et al.* The Meckel syndrome protein meckelin (TMEM67) is a key regulator of cilia function but is not required for tissue planar polarity. *Human Molecular Genetics* **22**, 2024-2040 (2013).
164. Lancaster, M.A. *et al.* Impaired Wnt-beta-catenin signaling disrupts adult renal homeostasis and leads to cystic kidney ciliopathy. *Nat Med* **15**, 1046-54 (2009).
165. Lancaster, M.A., Schroth, J. & Gleeson, J.G. Subcellular spatial regulation of canonical Wnt signalling at the primary cilium. *Nat Cell Biol* **13**, 700-7 (2011).
166. Hiatt, J.B., Pritchard, C.C., Salipante, S.J., O'Roak, B.J. & Shendure, J. Single molecule molecular inversion probes for targeted, high-accuracy detection of low-frequency variation. *Genome Res* **23**, 843-54 (2013).
167. Wang, H., Nettleton, D. & Ying, K. Copy number variation detection using next generation sequencing read counts. *BMC Bioinformatics* **15**, 109 (2014).
168. Ozyurek, H., Kayacik, O.E., Gungor, O. & Karagoz, F. Rare association of Hirschsprung's disease and Joubert syndrome. *Eur J Pediatr* **167**, 475-7 (2008).

169. Koutsouraki, E., Markou, E., Karlovasitou, A., Costa, V. & Baloyannis, S. Clinical case: vermis hypoplasia with features of Smith-Lemli-Opitz syndrome. *Int J Neurosci* **117**, 443-51 (2007).
170. Casamassima, A.C., Mamunes, P., Gladstone, I.M., Jr., Solomon, S. & Moncure, C. A new syndrome with features of the Smith-Lemli-Opitz and Meckel-Gruber syndromes in a sibship with cerebellar defects. *Am J Med Genet* **26**, 321-36 (1987).
171. Lindhout, D., Barth, P.G., Valk, J. & Boen-Tan, T.N. The Joubert syndrome associated with bilateral chorioretinal coloboma. *Eur J Pediatr* **134**, 173-6 (1980).
172. Lopez, E. *et al.* C5orf42 is the major gene responsible for OFD syndrome type VI. *Hum Genet* **133**, 367-77 (2014).
173. Moerman, P., Pauwels, P., Vandenberghe, K., Lauweryns, J.M. & Fryns, J.P. Goldston syndrome reconsidered. *Genet Couns* **4**, 97-102 (1993).
174. Shokeir, M.H., Houston, C.S. & Awen, C.F. Asphyxiating thoracic chondrodystrophy. Association with renal disease and evidence for possible heterozygous expression. *J Med Genet* **8**, 107-12 (1971).
175. Maroteaux, P. & Savart, P. [Asphyxiating Throacic Dystrophy. Radiological Study and Relation to the Ellis-Van Creveld Syndrome]. *Ann Radiol (Paris)* **7**, 332-8 (1964).
176. Allen, A.W., Jr., Moon, J.B., Hovland, K.R. & Minckler, D.S. Ocular findings in thoracic-pelvic-phalangeal dystrophy. *Arch Ophthalmol* **97**, 489-92 (1979).
177. Labrune, P. *et al.* Jeune syndrome and liver disease: report of three cases treated with ursodeoxycholic acid. *Am J Med Genet* **87**, 324-8 (1999).
178. Schmidts, M. *et al.* Exome sequencing identifies DYNC2H1 mutations as a common cause of asphyxiating thoracic dystrophy (Jeune syndrome) without major polydactyly, renal or retinal involvement. *J Med Genet* **50**, 309-23 (2013).
179. Hopper, M.S., Boulton, J.E. & Watson, A.R. Polyhydramnios associated with congenital pancreatic cysts and asphyxiating thoracic dysplasia. A case report. *S Afr Med J* **56**, 32-3 (1979).
180. Singh, M., Ray, D., Paul, V.K. & Kumar, A. Hydrocephalus in asphyxiating thoracic dystrophy. *Am J Med Genet* **29**, 391-5 (1988).
181. Beales, P.L. *et al.* IFT80, which encodes a conserved intraflagellar transport protein, is mutated in Jeune asphyxiating thoracic dystrophy. *Nat Genet* **39**, 727-9 (2007).
182. Davis, E.E. *et al.* TTC21B contributes both causal and modifying alleles across the ciliopathy spectrum. *Nature Genetics* **43**, 189-196 (2011).
183. Bredrup, C. *et al.* Ciliopathies with Skeletal Anomalies and Renal Insufficiency due to Mutations in the IFT-A Gene WDR19. *The American Journal of Human Genetics* **89**, 634-643 (2011).
184. Gilissen, C. *et al.* Exome sequencing identifies WDR35 variants involved in Sensenbrenner syndrome. *Am J Hum Genet* **87**, 418-23 (2010).
185. Perrault, I. *et al.* Mainzer-Saldino syndrome is a ciliopathy caused by IFT140 mutations. *Am J Hum Genet* **90**, 864-70 (2012).
186. Halbritter, J. *et al.* Defects in the IFT-B component IFT172 cause Jeune and Mainzer-Saldino syndromes in humans. *Am J Hum Genet* **93**, 915-25 (2013).
187. Dagoneau, N. *et al.* DYNC2H1 mutations cause asphyxiating thoracic dystrophy and short rib-polydactyly syndrome, type III. *Am J Hum Genet* **84**, 706-11 (2009).

188. McInerney-Leo, A.M. *et al.* Short-rib polydactyly and Jeune syndromes are caused by mutations in WDR60. *Am J Hum Genet* **93**, 515-23 (2013).
189. Schmidts, M. *et al.* Mutations in the gene encoding IFT dynein complex component WDR34 cause Jeune asphyxiating thoracic dystrophy. *Am J Hum Genet* **93**, 932-44 (2013).
190. Thiel, C. *et al.* NEK1 mutations cause short-rib polydactyly syndrome type majewski. *Am J Hum Genet* **88**, 106-14 (2011).
191. Levin, L.S. *et al.* A heritable syndrome of craniosynostosis, short thin hair, dental abnormalities, and short limbs: cranioectodermal dysplasia. *J Pediatr* **90**, 55-61 (1977).
192. Zaffanello, M. *et al.* Sensenbrenner syndrome: a new member of the hepatorenal fibrocystic family. *Am J Med Genet A* **140**, 2336-40 (2006).
193. Fry, A.E. *et al.* Connective tissue involvement in two patients with features of cranioectodermal dysplasia. *Am J Med Genet A* **149A**, 2212-5 (2009).
194. Walczak-Sztulpa, J. *et al.* Cranioectodermal Dysplasia, Sensenbrenner syndrome, is a ciliopathy caused by mutations in the IFT122 gene. *Am J Hum Genet* **86**, 949-56 (2010).
195. Arts, H.H. *et al.* C14ORF179 encoding IFT43 is mutated in Sensenbrenner syndrome. *J Med Genet* **48**, 390-5 (2011).
196. Saunier, S. *et al.* Characterization of the NPHP1 locus: mutational mechanism involved in deletions in familial juvenile nephronophthisis. *Am J Hum Genet* **66**, 778-89 (2000).
197. Lander, E.S. & Botstein, D. Homozygosity mapping: a way to map human recessive traits with the DNA of inbred children. *Science* **236**, 1567-70 (1987).
198. Katsanis, N. *et al.* Triallelic inheritance in Bardet-Biedl syndrome, a Mendelian recessive disorder. *Science* **293**, 2256-9 (2001).
199. Katsanis, N. *et al.* BBS4 is a minor contributor to Bardet-Biedl syndrome and may also participate in triallelic inheritance. *Am J Hum Genet* **71**, 22-9 (2002).
200. Beales, P.L. *et al.* Genetic interaction of BBS1 mutations with alleles at other BBS loci can result in non-Mendelian Bardet-Biedl syndrome. *Am J Hum Genet* **72**, 1187-99 (2003).
201. Abu-Safieh, L. *et al.* In search of triallelism in Bardet-Biedl syndrome. *Eur J Hum Genet* **20**, 420-7 (2012).
202. Mykytyn, K. *et al.* Identification of the gene (BBS1) most commonly involved in Bardet-Biedl syndrome, a complex human obesity syndrome. *Nat Genet* **31**, 435-8 (2002).
203. Khanna, H. *et al.* A common allele in RPGRIP1L is a modifier of retinal degeneration in ciliopathies. *Nature Genetics* **41**, 739-745 (2009).
204. Louie, C.M. *et al.* AHI1 is required for photoreceptor outer segment development and is a modifier for retinal degeneration in nephronophthisis. *Nat Genet* **42**, 175-80 (2010).
205. Dawe, H.R. *et al.* Nesprin-2 interacts with meckelin and mediates ciliogenesis via remodelling of the actin cytoskeleton. *Journal of Cell Science* **122**, 2716-2726 (2009).
206. Adams, M. *et al.* A meckelin–filamin A interaction mediates ciliogenesis. *Human Molecular Genetics* **21**, 1272-1286 (2012).
207. DiTommaso, T. *et al.* Identification of genes important for cutaneous function revealed by a large scale reverse genetic screen in the mouse. *PLoS Genet* **10**, e1004705 (2014).

208. Chapman, E.J. *et al.* Expression of hTERT immortalises normal human urothelial cells without inactivation of the p16/Rb pathway. *Oncogene* **25**, 5037-45 (2006).
209. den Hollander, A.I. *et al.* Mutations in the CEP290 (NPHP6) gene are a frequent cause of Leber congenital amaurosis. *Am J Hum Genet* **79**, 556-61 (2006).
210. Genomes Project, C. *et al.* A map of human genome variation from population-scale sequencing. *Nature* **467**, 1061-73 (2010).
211. Genomes Project, C. *et al.* An integrated map of genetic variation from 1,092 human genomes. *Nature* **491**, 56-65 (2012).
212. Adzhubei, I.A. *et al.* A method and server for predicting damaging missense mutations. *Nat Meth* **7**, 248-249 (2010).
213. Kumar, P., Henikoff, S. & Ng, P.C. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat. Protocols* **4**, 1073-1081 (2009).
214. Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F. & Higgins, D.G. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res* **25**, 4876-82 (1997).
215. Lundholt, B.K., Scudder, K.M. & Pagliaro, L. A simple technique for reducing edge effect in cell-based assays. *J Biomol Screen* **8**, 566-70 (2003).
216. Dowdle, William E. *et al.* Disruption of a Ciliary B9 Protein Complex Causes Meckel Syndrome. *The American Journal of Human Genetics* **89**, 94-110 (2011).
217. Litt, M. & Luty, J.A. A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene. *Am J Hum Genet* **44**, 397-401 (1989).
218. Taylor, G.R., Noble, J.S., Hall, J.L., Stewart, A.D. & Mueller, R.F. Hypervariable microsatellite for genetic diagnosis. *Lancet* **2**, 454 (1989).
219. Schwarz, J.M., Rodelsperger, C., Schuelke, M. & Seelow, D. MutationTaster evaluates disease-causing potential of sequence alterations. *Nat Meth* **7**, 575-576 (2010).
220. Mougou-Zerelli, S. *et al.* CC2D2A Mutations in Meckel and Joubert Syndromes Indicate a Genotype-Phenotype Correlation. *Human Mutation* **30**, 1574-1582 (2009).
221. Khaddour, R. *et al.* Spectrum of MKS1 and MKS3 mutations in Meckel syndrome: a genotype-phenotype correlation. *Hum Mutat* **28**, 523 - 524 (2007).
222. Auber, B. *et al.* A disease causing deletion of 29 base pairs in intron 15 in the MKS1 gene is highly associated with the campomelic variant of the Meckel-Gruber syndrome. *Clinical Genetics* **72**, 454-459 (2007).
223. Zhang, M.Q. Statistical Features of Human Exons and Their Flanking Regions. *Human Molecular Genetics* **7**, 919-932 (1998).
224. Yeo, G.B., C. B. Maximum Entropy Modeling of Short Sequence Motifs with Applications to RNA Splicing Signals. *Journal of Computational Biology* **11**, 377-394 (March 2004).
225. Reese, M.G.E., F.H.; Kulp, D.; Haussler D. Improved splice site detection in Genie. *J Comput Biol.* **4**, 311-23 (1997).
226. Pertea, M., Lin, X. & Salzberg, S.L. GeneSplicer: a new computational method for splice site prediction. *Nucleic Acids Research* **29**, 1185-1190 (2001).

227. Desmet, F.-O. *et al.* Human Splicing Finder: an online bioinformatics tool to predict splicing signals. *Nucleic Acids Research* **37**, e67 (2009).
228. Zhang, C., Hastings, M.L., Krainer, A.R. & Zhang, M.Q. Dual-specificity splice sites function alternatively as 5' and 3' splice sites. *Proceedings of the National Academy of Sciences* **104**, 15028-15033 (2007).
229. Doherty, D. *et al.* Mutations in 3 genes (MKS3, CC2D2A and RPGRIP1L) cause COACH syndrome (Joubert syndrome with congenital hepatic fibrosis). *Journal of Medical Genetics* **47**, 8-21 (2010).
230. Brancati, F. *et al.* RPGRIP1L mutations are mainly associated with the cerebello-renal phenotype of Joubert syndrome-related disorders. *Clinical Genetics* **74**, 164-170 (2008).
231. Chen, J. *et al.* Molecular Analysis of Bardet-Biedl Syndrome Families: Report of 21 Novel Mutations in 10 Genes. *Investigative Ophthalmology & Visual Science* **52**, 5317-5324 (2011).
232. Carr, I.M., Flintoff, K.J., Taylor, G.R., Markham, A.F. & Bonthron, D.T. Interactive visual analysis of SNP data for rapid autozygosity mapping in consanguineous families. *Human Mutation* **27**, 1041-1046 (2006).
233. Carr, I.M., Sheridan, E., Hayward, B.E., Markham, A.F. & Bonthron, D.T. IBDfinder and SNPsetter: Tools for pedigree-independent identification of autozygous regions in individuals with recessive inherited disease. *Human Mutation* **30**, 960-967 (2009).
234. Wang, M., Bridges, J.P., Na, C.L., Xu, Y. & Weaver, T.E. Meckel-Gruber Syndrome Protein MKS3 Is Required for Endoplasmic Reticulum-associated Degradation of Surfactant Protein C. *Journal of Biological Chemistry* **284**, 33377-33383 (2009).
235. Reiter, J.F. & Skarnes, W.C. Tectonic, a novel regulator of the Hedgehog pathway required for both activation and inhibition. *Genes & Development* **20**, 22-27 (2006).
236. Chaki, M. *et al.* Exome Capture Reveals ZNF423 and CEP164 Mutations, Linking Renal Ciliopathies to DNA Damage Response Signaling. *Cell* **150**, 533-548 (2012).
237. Cavalcanti, D.P. *et al.* Mutation in IFT80 in a fetus with the phenotype of Verma-Naumoff provides molecular evidence for Jeune-Verma-Naumoff dysplasia spectrum. *Journal of Medical Genetics* **48**, 88-92 (2011).
238. Christopher, K.J., Wang, B., Kong, Y. & Weatherbee, S.D. Forward genetics uncovers Transmembrane protein 107 as a novel factor required for ciliogenesis and Sonic hedgehog signaling. *Developmental Biology* **368**, 382-392 (2012).
239. Thomas, S. *et al.* TCTN3 Mutations Cause Mohr-Majewski Syndrome. *The American Journal of Human Genetics* **91**, 372-378 (2012).
240. Thomas, S. *et al.* A Homozygous PDE6D Mutation in Joubert Syndrome Impairs Targeting of Farnesylated INPP5E Protein to the Primary Cilium. *Human Mutation* **35**, 137-146 (2014).
241. Paetau, A. *et al.* Hydroletharus syndrome: neuropathology of 21 cases confirmed by HYL51 gene mutation analysis. *J Neuropathol Exp Neurol* **67**, 750-62 (2008).
242. Barral, D.C. *et al.* Arl13b regulates endocytic recycling traffic. *Proc Natl Acad Sci U S A* **109**, 21354-9 (2012).
243. Wu, Y. *et al.* DIXDC1 co-localizes and interacts with gamma-tubulin in HEK293 cells. *Cell Biol Int* **33**, 697-701 (2009).

244. Singh, K.K. *et al.* Dixdc1 is a critical regulator of DISC1 and embryonic cortical development. *Neuron* **67**, 33-48 (2010).
245. Gherman, A., Davis, E.E. & Katsanis, N. The ciliary proteome database: an integrated community resource for the genetic and functional dissection of cilia. *Nat Genet* **38**, 961-2 (2006).
246. Zuniga, F.I. & Craft, C.M. Deciphering the Structure and Function of Als2cr4 in the Mouse Retina. *Investigative Ophthalmology & Visual Science* **51**, 4407-4415 (2010).
247. Wallingford, J.B. *et al.* Dishevelled controls cell polarity during *Xenopus* gastrulation. *Nature* **405**, 81-85 (2000).
248. Park, T.J., Haigo, S.L. & Wallingford, J.B. Ciliogenesis defects in embryos lacking inturned or fuzzy function are associated with failure of planar cell polarity and Hedgehog signaling. *Nature Genetics* **38**, 303-311 (2006).
249. Williams, C.L. *et al.* MKS and NPHP modules cooperate to establish basal body/transition zone membrane associations and ciliary gate function during ciliogenesis. *The Journal of Cell Biology* **192**, 1023-1041 (2011).
250. Carthon, B.C. *et al.* Genetic Replacement of Cyclin D1 Function in Mouse Development by Cyclin D2. *Molecular and Cellular Biology* **25**, 1081-1088 (2005).
251. Woods, C.G. *et al.* Quantification of homozygosity in consanguineous individuals with autosomal recessive disease. *Am J Hum Genet* **78**, 889-96 (2006).
252. Wang, K. *et al.* Whole-genome DNA/RNA sequencing identifies truncating mutations in RBCK1 in a novel Mendelian disease with neuromuscular and cardiac involvement. *Genome Med* **5**, 67 (2013).
253. Flanagan, S.E. *et al.* Next-generation sequencing reveals deep intronic cryptic ABCC8 and HADH splicing founder mutations causing hyperinsulinism by pseudoexon activation. *Am J Hum Genet* **92**, 131-6 (2013).
254. Webb, T.R. *et al.* Deep intronic mutation in OFD1, identified by targeted genomic next-generation sequencing, causes a severe form of X-linked retinitis pigmentosa (RP23). *Hum Mol Genet* **21**, 3647-54 (2012).
255. Wansleben, C. *et al.* An ENU-mutagenesis screen in the mouse: identification of novel developmental gene functions. *PLoS One* **6**, e19357 (2011).
256. Shaheen, R. *et al.* Genomic analysis of Meckel-Gruber syndrome in Arabs reveals marked genetic heterogeneity and novel candidate genes. *Eur J Hum Genet* **21**, 762-8 (2013).
257. Duriez, B. *et al.* A common variant in combination with a nonsense mutation in a member of the thioredoxin family causes primary ciliary dyskinesia. *Proceedings of the National Academy of Sciences* **104**, 3336-3341 (2007).
258. Das, A. & Guo, W. Rabs and the exocyst in ciliogenesis, tubulogenesis and beyond. *Trends in Cell Biology* **21**, 383-386 (2011).
259. Poulter, J.A. *et al.* Recessive mutations in SLC38A8 cause foveal hypoplasia and optic nerve misrouting without albinism. *Am J Hum Genet* **93**, 1143-50 (2013).
260. Frank, C.G. *et al.* Identification and functional analysis of a defect in the human ALG9 gene: definition of congenital disorder of glycosylation type II. *Am J Hum Genet* **75**, 146-50 (2004).
261. Resnitzky, D., Gossen, M., Bujard, H. & Reed, S.I. Acceleration of the G1/S phase transition by expression of cyclins D1 and E with an inducible system. *Mol Cell Biol* **14**, 1669-79 (1994).

262. Brancati, F. *et al.* MKS3/TMEM67 Mutations Are a Major Cause of COACH Syndrome, a Joubert Syndrome Related Disorder with Liver Involvement. *Human Mutation* **30**, E432-E442 (2009).
263. Wu, J. & Mlodzik, M. The Frizzled Extracellular Domain Is a Ligand for Van Gogh/Stbm during Nonautonomous Planar Cell Polarity Signaling. *Developmental Cell* **15**, 462-469 (2008).
264. Nishiguchi, K.M. *et al.* Whole genome sequencing in patients with retinitis pigmentosa reveals pathogenic DNA structural changes and NEK2 as a new disease gene. *Proc Natl Acad Sci U S A* **110**, 16139-44 (2013).
265. Birmingham, A. *et al.* Statistical methods for analysis of high-throughput RNA interference screens. *Nat Meth* **6**, 569-575 (2009).
266. Reagan-Shaw, S. & Ahmad, N. Silencing of polo-like kinase (Plk) 1 via siRNA causes induction of apoptosis and impairment of mitosis machinery in human prostate cancer cells: implications for the treatment of prostate cancer. *Faseb Journal* **19**, 611-+ (2005).
267. Chung, N. *et al.* Median Absolute Deviation to Improve Hit Selection for Genome-Scale RNAi Screens. *Journal of Biomolecular Screening* **13**, 149-158 (2008).
268. Sigoillot, F.D. *et al.* A bioinformatics method identifies prominent off-targeted transcripts in RNAi screens. *Nature Methods* **9**, 363-U66 (2012).
269. Trapnell, C., Pachter, L. & Salzberg, S.L. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* **25**, 1105-11 (2009).
270. Kim, D. *et al.* TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol* **14**, R36 (2013).
271. Trapnell, C. *et al.* Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotechnol* **28**, 511-5 (2010).
272. Grimm, D. *et al.* Argonaute proteins are key determinants of RNAi efficacy, toxicity, and persistence in the adult mouse liver. *Journal of Clinical Investigation* **120**, 3106-3119 (2010).
273. Vickers, T.A., Lima, W.F., Nichols, J.G. & Crooke, S.T. Reduced levels of Ago2 expression result in increased siRNA competition in mammalian cells. *Nucleic Acids Research* **35**, 6598-6610 (2007).
274. Ohrt, T., Merkle, D., Birkenfeld, K., Echeverri, C.J. & Schwillle, P. In situ fluorescence analysis demonstrates active siRNA exclusion from the nucleus by Exportin 5. *Nucleic Acids Research* **34**, 1369-1380 (2006).
275. Koller, E. *et al.* Competition for RISC binding predicts in vitro potency of siRNA. *Nucleic Acids Research* **34**, 4467-4476 (2006).
276. Birmingham, A. *et al.* Statistical methods for analysis of high-throughput RNA interference screens. *Nat Methods* **6**, 569-75 (2009).
277. Zhang, X.D. A pair of new statistical parameters for quality control in RNA interference high-throughput screening assays. *Genomics* **89**, 552-561 (2007).
278. Shaheen, R. *et al.* A founder CEP120 mutation in Jeune asphyxiating thoracic dystrophy expands the role of centriolar proteins in skeletal ciliopathies. *Hum Mol Genet* (2014).
279. Martin, C.A. *et al.* Mutations in PLK4, encoding a master regulator of centriole biogenesis, cause microcephaly, growth failure and retinopathy. *Nat Genet* **46**, 1283-92 (2014).
280. Lee, J.H. *et al.* Evolutionarily Assembled cis-Regulatory Module at a Human Ciliopathy Locus. *Science* **335**, 966-969 (2012).

281. Loktev, Alexander V. & Jackson, Peter K. Neuropeptide Y Family Receptors Traffic via the Bardet-Biedl Syndrome Pathway to Signal in Neuronal Primary Cilia. *Cell Reports* **5**, 1316-1329 (2013).
282. Balestra, Fernando R., Strnad, P., Flückiger, I. & Gönczy, P. Discovering Regulators of Centriole Biogenesis through siRNA-Based Functional Genomics in Human Cells. *Developmental Cell* **25**, 555-571 (2013).
283. Joo, K. *et al.* CCDC41 is required for ciliary vesicle docking to the mother centriole. *Proceedings of the National Academy of Sciences* **110**, 5987-5992 (2013).
284. Ferrante, M.I. *et al.* Oral-facial-digital type I protein is required for primary cilia formation and left-right axis specification. *Nat Genet* **38**, 112-7 (2006).
285. Paulsen, R.D. *et al.* A Genome-wide siRNA Screen Reveals Diverse Cellular Processes and Pathways that Mediate Genome Stability. *Molecular Cell* **35**, 228-239 (2009).
286. Kim, K., Lee, K. & Rhee, K. CEP90 is required for the assembly and centrosomal accumulation of centriolar satellites, which is essential for primary cilia formation. *PLoS One* **7**, e48196 (2012).
287. Mukhopadhyay, S. *et al.* The Ciliary G-Protein-Coupled Receptor Gpr161 Negatively Regulates the Sonic Hedgehog Pathway via cAMP Signaling. *Cell* **152**, 210-223 (2013).
288. Arellano, J.I., Guadiana, S.M., Breunig, J.J., Rakic, P. & Sarkisian, M.R. Development and distribution of neuronal cilia in mouse neocortex. *The Journal of Comparative Neurology* **520**, 848-873 (2012).
289. Guadiana, S.M. *et al.* Arborization of Dendrites by Developing Neocortical Neurons Is Dependent on Primary Cilia and Type 3 Adenylyl Cyclase. *The Journal of Neuroscience* **33**, 2626-2638 (2013).
290. Bassilana, F. *et al.* Target identification for a Hedgehog pathway inhibitor reveals the receptor GPR39. *Nat Chem Biol* **10**, 343-9 (2014).
291. Mougou-Zerelli, S. *et al.* CC2D2A mutations in Meckel and Joubert syndromes indicate a genotype-phenotype correlation. *Hum Mutat* **30**, 1574 - 1582 (2009).
292. Cheng, J. & Maquat, L.E. Nonsense codons can reduce the abundance of nuclear mRNA without affecting the abundance of pre-mRNA or the half-life of cytoplasmic mRNA. *Mol Cell Biol* **13**, 1892-902 (1993).
293. Halbritter, J. *et al.* Identification of 99 novel mutations in a worldwide cohort of 1,056 patients with a nephronophthisis-related ciliopathy. *Hum Genet* **132**, 865-84 (2013).
294. Halbritter, J. *et al.* High-throughput mutation analysis in patients with a nephronophthisis-associated ciliopathy applying multiplexed barcoded array-based PCR amplification and next-generation sequencing. *J Med Genet* **49**, 756-67 (2012).
295. Waters, A.M. *et al.* The kinetochore protein, CENPF, is mutated in human ciliopathy and microcephaly phenotypes. *J Med Genet* (2015).
296. Li, A. *et al.* Ciliary transition zone activation of phosphorylated Tctex-1 controls ciliary resorption, S-phase entry and fate of neural progenitors. *Nat Cell Biol* **13**, 402-411 (2011).
297. Lee, H. *et al.* Clinical exome sequencing for genetic identification of rare Mendelian disorders. *JAMA* **312**, 1880-7 (2014).
298. Parry, D.A. *et al.* SAMS, a syndrome of short stature, auditory-canal atresia, mandibular hypoplasia, and skeletal abnormalities is a unique

- neurocristopathy caused by mutations in Goosecoid. *Am J Hum Genet* **93**, 1135-42 (2013).
299. O'Roak, B.J. *et al.* Exome sequencing in sporadic autism spectrum disorders identifies severe de novo mutations. *Nat Genet* **43**, 585-9 (2011).
 300. Renton, A.E. & Traynor, B.J. CRESTing the ALS mountain. *Nat Neurosci* **16**, 774-5 (2013).
 301. MacArthur, D.G. *et al.* Guidelines for investigating causality of sequence variants in human disease. *Nature* **508**, 469-76 (2014).
 302. Fokkema, I.F. *et al.* LOVD v.2.0: the next generation in gene variant databases. *Hum Mutat* **32**, 557-63 (2011).
 303. Arellano, J.I., Guadiana, S.M., Breunig, J.J., Rakic, P. & Sarkisian, M.R. Development and distribution of neuronal cilia in mouse neocortex. *J Comp Neurol* **520**, 848-73 (2012).
 304. Guadiana, S.M. *et al.* Arborization of dendrites by developing neocortical neurons is dependent on primary cilia and type 3 adenylyl cyclase. *J Neurosci* **33**, 2626-38 (2013).
 305. Andersen, J.S. *et al.* Proteomic characterization of the human centrosome by protein correlation profiling. *Nature* **426**, 570-4 (2003).
 306. Jakobsen, L. *et al.* Novel asymmetrically localizing components of human centrosomes identified by complementary proteomics methods. *EMBO J* **30**, 1520-35 (2011).
 307. Neumann, B. *et al.* Phenotypic profiling of the human genome by time-lapse microscopy reveals cell division genes. *Nature* **464**, 721-7 (2010).
 308. Yim, H., Sung, C.K., You, J., Tian, Y. & Benjamin, T. Nek1 and TAZ interact to maintain normal levels of polycystin 2. *J Am Soc Nephrol* **22**, 832-7 (2011).
 309. Borgal, L. *et al.* The ciliary protein nephrocystin-4 translocates the canonical Wnt regulator Jade-1 to the nucleus to negatively regulate beta-catenin signaling. *J Biol Chem* **287**, 25370-80 (2012).
 310. Kasahara, K. *et al.* Ubiquitin-proteasome system controls ciliogenesis at the initial step of axoneme extension. *Nat Commun* **5**, 5081 (2014).
 311. Li, J. *et al.* USP33 regulates centrosome biogenesis via deubiquitination of the centriolar protein CP110. *Nature* **495**, 255-9 (2013).
 312. Song, E.J. *et al.* The Prp19 complex and the Usp4Sart3 deubiquitinating enzyme control reversible ubiquitination at the spliceosome. *Genes Dev* **24**, 1434-47 (2010).
 313. Pelisch, F., Risso, G. & Srebrow, A. RNA metabolism and ubiquitin/ubiquitin-like modifications collide. *Brief Funct Genomics* **12**, 66-71 (2013).
 314. Lai, C.K. *et al.* Functional characterization of putative cilia genes by high-content analysis. *Mol Biol Cell* **22**, 1104-19 (2011).
 315. Abu-Safieh, L. *et al.* Autozygome-guided exome sequencing in retinal dystrophy patients reveals pathogenetic mutations and novel candidate disease genes. *Genome Res* **23**, 236-47 (2013).
 316. Slaats, G.G. *et al.* Nephronophthisis-Associated CEP164 Regulates Cell Cycle Progression, Apoptosis and Epithelial-to-Mesenchymal Transition. *PLoS Genet* **10**, e1004594 (2014).
 317. Jonassen, J.A., SanAgustin, J., Baker, S.P. & Pazour, G.J. Disruption of IFT complex A causes cystic kidneys without mitotic spindle misorientation. *J Am Soc Nephrol* **23**, 641-51 (2012).

318. Tammachote, R. *et al.* Ciliary and centrosomal defects associated with mutation and depletion of the Meckel syndrome genes MKS1 and MKS3. *Human Molecular Genetics* **18**, 3311-3323 (2009).
319. Azimzadeh, J. *et al.* hPOC5 is a centrin-binding protein required for assembly of full-length centrioles. *J Cell Biol* **185**, 101-14 (2009).
320. Barrangou, R. *et al.* CRISPR provides acquired resistance against viruses in prokaryotes. *Science* **315**, 1709-12 (2007).
321. Sternberg, S.H., Redding, S., Jinek, M., Greene, E.C. & Doudna, J.A. DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. *Nature* **507**, 62-7 (2014).
322. Ran, F.A. *et al.* Genome engineering using the CRISPR-Cas9 system. *Nat Protoc* **8**, 2281-308 (2013).
323. Zhang, Y. *et al.* BBS mutations modify phenotypic expression of CEP290-related ciliopathies. *Hum Mol Genet* **23**, 40-51 (2014).
324. Takahashi, K. & Yamanaka, S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126**, 663-76 (2006).
325. Polo, J.M. *et al.* A molecular roadmap of reprogramming somatic cells into iPS cells. *Cell* **151**, 1617-32 (2012).
326. Okita, K. *et al.* A more efficient method to generate integration-free human iPS cells. *Nat Methods* **8**, 409-12 (2011).
327. Miyaoka, Y. *et al.* Isolation of single-base genome-edited human iPS cells without antibiotic selection. *Nat Methods* **11**, 291-3 (2014).
328. Sarra, G.M. *et al.* Gene replacement therapy in the retinal degeneration slow (rds) mouse: the effect on retinal degeneration following partial transduction of the retina. *Hum Mol Genet* **10**, 2353-61 (2001).
329. McIntyre, J.C. *et al.* Gene therapy rescues cilia defects and restores olfactory function in a mammalian ciliopathy model. *Nat Med* **18**, 1423-8 (2012).
330. Nho, K. *et al.* Identification of functional variants from whole-exome sequencing, combined with neuroimaging genetics. *Mol Psychiatry* **18**, 739 (2013).
331. Zuchner, S. *et al.* Whole-exome sequencing links a variant in DHDDS to retinitis pigmentosa. *Am J Hum Genet* **88**, 201-6 (2011).
332. Mirzaa, G.M. *et al.* De novo CCND2 mutations leading to stabilization of cyclin D2 cause megalencephaly-polymicrogyria-polydactyly-hydrocephalus syndrome. *Nat Genet* **46**, 510-5 (2014).
333. Tse, H.F. *et al.* Patient-specific induced-pluripotent stem cells-derived cardiomyocytes recapitulate the pathogenic phenotypes of dilated cardiomyopathy due to a novel DES mutation identified by whole exome sequencing. *Hum Mol Genet* **22**, 1395-403 (2013).
334. Tucker, B.A. *et al.* Exome sequencing and analysis of induced pluripotent stem cells identify the cilia-related gene male germ cell-associated kinase (MAK) as a cause of retinitis pigmentosa. *Proc Natl Acad Sci U S A* **108**, E569-76 (2011).
335. Freedman, B.S. *et al.* Reduced ciliary polycystin-2 in induced pluripotent stem cells from polycystic kidney disease patients with PKD1 mutations. *J Am Soc Nephrol* **24**, 1571-86 (2013).
336. Aartsma-Rus, A. Antisense-mediated modulation of splicing: therapeutic implications for Duchenne muscular dystrophy. *RNA Biol* **7**, 453-61 (2010).
337. Nagel-Wolfrum, K., Moller, F., Penner, I. & Wolfrum, U. Translational read-through as an alternative approach for ocular gene therapy of retinal

- dystrophies caused by in-frame nonsense mutations. *Vis Neurosci* **31**, 309-16 (2014).
338. Collin, R.W. *et al.* Antisense Oligonucleotide (AON)-based Therapy for Leber Congenital Amaurosis Caused by a Frequent Mutation in CEP290. *Mol Ther Nucleic Acids* **1**, e14 (2012).

Appendix

1. Ethical approval and amendments:



National Research Ethics Service

South Yorkshire Research Ethics Committee

Millside
Mill Pond Lane
Meanwood
Leeds
LS6 4RA

Telephone: 0113 305 0128

18 February 2011

Prof Colin A Johnson
Professor of Medical & Molecular Genetics
Section of Ophthalmology and Neuroscience
Leeds Inst of Molecular Medicine
St James University Hospital
Leeds LS9 7TF

Dear Prof Johnson

Study Title: Molecular genetic investigations of autosomal recessive conditions
REC reference number: 11/H1310/1

Thank you for your letter of 8th February 2011, responding to the Committee's request for further information on the above research and submitting revised documentation.

The further information has been considered on behalf of the Committee by the Chair.

Confirmation of ethical opinion

On behalf of the Committee, I am pleased to confirm a favourable ethical opinion for the above research on the basis described in the application form, protocol and supporting documentation as revised, subject to the conditions specified below.

Ethical review of research sites

The favourable opinion applies to all NHS sites taking part in the study, subject to management permission being obtained from the NHS/HSC R&D office prior to the start of the study (see "Conditions of the favourable opinion" below).

Conditions of the favourable opinion

The favourable opinion is subject to the following conditions being met prior to the start of the study.

Management permission or approval must be obtained from each host organisation prior to the start of the study at the site concerned.

For NHS research sites only, management permission for research ("R&D approval") should be obtained from the relevant care organisation(s) in accordance with NHS research governance arrangements. Guidance on applying for NHS permission for research is available in the Integrated Research Application System or at <http://www.rdforum.nhs.uk>.

Where the only involvement of the NHS organisation is as a Participant Identification Centre (PIC), management permission for research is not required but the R&D office should be notified of the study and agree to the organisation's involvement. Guidance on procedures for PICs is available in IRAS. Further advice should be sought from the R&D office where necessary.

Sponsors are not required to notify the Committee of approvals from host organisations.

It is the responsibility of the sponsor to ensure that all the conditions are complied with before the start of the study or its initiation at a particular site (as applicable).

Approved documents

The final list of documents reviewed and approved by the Committee is as follows:

Document	Version	Date
Investigator CV	1	14 December 2010
Covering Letter		08 February 2011
Letter from Sponsor		
REC application		14 December 2010
Response to Request for Further Information		
Participant Information Sheet: for children, aged 8 years or younger (Mirpuri Urdu translation)	1	14 December 2010
Evidence of insurance or indemnity		
Participant Information Sheet: Children aged 12 - 15 years	2	08 February 2011
Letter of invitation to participant	1	14 December 2010
Participant Information Sheet: for parents (Mirpuri Urdu translation)	1	14 December 2010
Participant Information Sheet: Relatives that are adults or young persons	2	08 February 2011
Participant Information Sheet: for children aged 8 year or younger	1	14 December 2010
Participant Information Sheet: Parents	2	08 February 2011
Participant Information Sheet: for children aged 8-12 years	1	14 December 2010
Participant Consent Form: Assent form for older children	1	14 December 2010
Protocol	1	14 December 2010

Statement of compliance

The Committee is constituted in accordance with the Governance Arrangements for Research Ethics Committees (July 2001) and complies fully with the Standard Operating Procedures for Research Ethics Committees in the UK.

After ethical review

Now that you have completed the application process please visit the National Research Ethics Service website > After Review

You are invited to give your view of the service that you have received from the National Research Ethics Service and the application procedure. If you wish to make your views known please use the feedback form available on the website.

The attached document "After ethical review – guidance for researchers" gives detailed guidance on reporting requirements for studies with a favourable opinion, including:

- Notifying substantial amendments
- Adding new sites and investigators
- Progress and safety reports
- Notifying the end of the study

The NRES website also provides guidance on these topics, which is updated in the light of changes in reporting requirements or procedures.

We would also like to inform you that we consult regularly with stakeholders to improve our service. If you would like to join our Reference Group please email referencegroup@nres.npsa.nhs.uk.

11/H1310/1

Please quote this number on all correspondence

With the Committee's best wishes for the success of this project

Yours sincerely




Ms Jo Abbott
Chair

Email: Sinead.audsley@leedspft.nhs.uk

Enclosures: *After ethical review – guidance for researchers*

Copy to: *Mrs Rachel E de Souza, University of Leeds*
Mrs Anne Gowing, Leeds Teaching Hospitals NHS Trust

The Leeds Teaching Hospitals 
NHS Trust

Ref: Amanda Burd

29/04/2011

Dr Colin A. Johnson

Section of Ophthalmology and Neurosciences,
Wellcome Trust Brenner Bui
Leeds Institute of Molecular Medicine,
St. James's University Hospital
LS9 7TF

Research & Development

Leeds Teaching Hospitals NHS Trust

34 Hyde Terrace
Leeds
LS2 9LN

Tel: 0113 392 2878
Fax: 0113 392 6397

r&d@leedsth.nhs.uk
www.leedsth.nhs.uk

Dear Dr Johnson

**Re: NHS Permission at LTHT for: Molecular genetic investigations of
autosomal recessive conditions
LTHT R&D Number: CG11/9764 (53788/WY)
REC: 11/H1310/1**

I confirm that *NHS Permission for research* has been granted for this project at The Leeds Teaching Hospitals NHS Trust (LTHT). NHS Permission is granted based on the information provided in the documents listed below. All amendments (including changes to the research team) must be submitted in accordance with guidance in IRAS. Any change to the status of the project must be notified to the R&D Department.

Permission is granted on the understanding that the study is conducted in accordance with the *Research Governance Framework for Health and Social Care*, ICH GCP (if applicable) and NHS Trust policies and procedures available at http://www.leedsth.nhs.uk/sites/research_and_development/.

This permission is granted only on the understanding that you comply with the requirements of the *Framework* as listed in the attached sheet "Conditions of Approval".

If you have any queries about this approval please do not hesitate to contact the R&D Department on telephone 0113 392 2878.

Indemnity Arrangements

Chairman Mike Collier CBE Chief Executive Maggie Boyle

The Leeds Teaching Hospitals incorporating:

Chapel Allerton Hospital Leeds Dental Institute Seacroft Hospital
St James's University Hospital The General Infirmary at Leeds Wharfedale Hospital



HT/288

The Leeds Teaching Hospitals NHS Trust participates in the NHS risk pooling scheme administered by the NHS Litigation Authority 'Clinical Negligence Scheme for NHS Trusts' for: (i) medical professional and/or medical malpractice liability; and (ii) general liability. NHS Indemnity for negligent harm is extended to researchers with an employment contract (substantive or honorary) with the Trust. The Trust only accepts liability for research activity that has been managerially approved by the R&D Department.

The Trust therefore accepts liability for the above research project and extends indemnity for negligent harm to cover you as investigator and the researchers listed on the Site Specific Information form. Should there be any changes to the research team please ensure that you inform the R&D Department and that s/he obtains an appropriate contract, or letter of access, with the Trust if required.

Yours sincerely



Dr D R Norfolk
Associate Director of R&D

Approved documents

The documents reviewed and approved are listed as follows

<i>Document</i>	<i>Version</i>	<i>Date of document</i>
NHS R&D Form	3.1	02/03/2011
SSI Form	3.1	08/03/2011
Directorate Approval		08/03/2011
Radiology Approval		N/A
Pharmacy Approval		N/A
Protocol	1.0	14/12/2010
REC Letter confirming favourable opinion		18/02/2011
Evidence of Insurance		Not dated
Patient information sheet (REC Approved) Parent	2.0	08/02/2011
Patient information sheet (REC Approved) 8/younger URDU	1.0	14/12/2010
Patient information sheet (REC Approved) Parent URDU	1.0	14/12/2010
Patient information sheet (REC Approved) 12-15 yrs	2.0	08/02/2011
PIS (REC Approved) Adult relatives/young Person	2.0	08/02/2011
PIS (REC Approved) Children aged 8 or younger	1.0	14/12/2010
PIS (REC Approved) Children aged 8 – 12 years	1.0	14/12/2010
Informed Consent (REC Approved) Assent older children	1.0	14/12/2010
Letter of Invitation to participants (REC Approved)	1.0	14/12/2010

Conditions of NHS Permission for Research:

- Permission from your Directorate must be obtained before starting the study.
- Favourable Opinion of the appropriate Research Ethics Committee, where necessary, must be obtained before starting the study.
- Arrangements must be made to ensure that all members of the research team, where applicable, have employment contracts with the Trust (either full or honorary).
- Agreements must be in place with appropriate support departments regarding the services required to undertake the project and arrangements must be in place to recompense them for the costs of their services.
- Arrangements must be in place for the management of financial and other resources provided for the study, including intellectual property arising from the work.
- Priority should be given at all times to the dignity, rights, safety and well being of participants in the study
- Healthcare staff should be suitably informed about the research their patients are taking part in and information specifically relevant to their care arising from the study should be communicated promptly.
- Each member of the research team must be qualified by education, training and experience to discharge his/her role in the study. Students and new researchers must have adequate supervision, support and training.
- The research must follow the protocol approved by the relevant research ethics committee. Any proposed amendments to or deviations from the protocol must be submitted for review by the Research Ethics Committee, the research sponsor, regulatory authority and any other appropriate body. The R&D Department should be informed where the amendment has resource implications within the Directorate and the Directorate research lead/clinical director notified.
- Adverse Events in clinical trials of investigational medicinal products must be reported in accordance with the Medicines for Human Use (Clinical Trials) Regulations 2004.
- Complete and return Study Status Reports, when requested, to the R&D Department within 28 days of receipt as requested. (NB Failure to comply to such request with the requirement will lead to suspension of NHS Permission.)
- Procedures should be in place to ensure collection of high quality, accurate data and the integrity and confidentiality of data during processing and storage.

- Arrangements must be made for the appropriate archiving of data when the research has finished. Records must normally be kept for 15 years.
- All data and documentation associated with the study must be available for audit at the request of the appropriate auditing authority. Projects are randomly selected for audit by the R&D Department. You will be informed by letter if your study is selected.
- Findings from the study should be disseminated promptly and fed back as agreed to research participants.
- Findings from the study should be exposed to critical review through accepted scientific and professional channels.
- All members of the research team must ensure that the process of informed consent adheres to the standards GCP outlined in the UK Clinical Trials Regulations. Investigators are directed to the R&D website for further information and training availability.
- Where applicable, this NHS Permission includes aspects of the study previously covered by the NRES Site Specific Assessment (SSA) process.

Commercially Sponsored Trials

If the study is commercially sponsored, NHS Permission is given subject to provision of the following documents.

- Clinical Trials Agreement - agreed and signed off by the R&D Department (on behalf of the Leeds Teaching Hospitals NHS Trust) and the Sponsor. Investigators do not have the authority to sign contract on behalf of the Trust.
- Indemnity agreement, if not included in the Clinical Trials Agreement- (standard ABPI no fault arrangements apply) signed by the R&D Department and the Sponsor

It is essential that all the responsibilities set out in the Research Governance Framework, including those outlined above are fulfilled. The Trust reserves the right to withdraw NHS Permission where the above criteria are not being met. The Trust will not accept liability for any activity where NHS Permission has not been granted.

Carbon Copy: (PI) Dr Eamonn Sheridan



Health Research Authority

NRES Committee Yorkshire & The Humber - South Yorkshire

North East REC Centre
Unit 002, TEDCO Business Centre
Rolling Mill Road
Jarrow
Tyne and Wear
NE32 3DT

Tel: 0191 428 3387
Fax: 0191 428 3432

14 March 2013

Ms Miranda Squires
Senior Research Co-ordinator
Department of Clinical Genetics
Third Floor, Chapel Allerton Hospital
Chapelton Road
Leeds
LS7 4SA

Dear Ms Squires

Study title: Molecular genetic investigations of autosomal recessive conditions
REC reference: 11/H1310/1
Amendment number: Amendment 1
Amendment date: 10 December 2012
IRAS project ID: 53788

The above amendment was reviewed by the Sub-Committee in correspondence.

Ethical opinion

Members requested a copy of the consent forms used in study 10/H0106/8 to confirm that the participants had given consent for their samples to be used in future research.

You provided a copy of the consent forms.

The members of the Committee taking part in the review gave a favourable ethical opinion of the amendment on the basis described in the notice of amendment form and supporting documentation.

Approved documents

The documents reviewed and approved at the meeting were:

Document	Version	Date
SOP for Dealing with Requests for Samples Collected During "The National Study of MPGN"	Final	23 November 2012
Lab Sample Processing SOP	2.0	30 June 2012
Protocol	2	18 February 2013
Notice of Substantial Amendment (non-CTIMPs)	Amendment 1	10 December 2012
Covering Letter	Email from	20 February 2013

	Jackie Mitchell	
Participant Consent Form: Assent Form (REC Ref 10/H0106/8)	1	14 December 2010
Participant Consent Form: Consent Form (REC Ref 10/H0106/8)	1	14 December 2010

Membership of the Committee

The members of the Committee who took part in the review are listed on the attached sheet.

R&D approval

All investigators and research collaborators in the NHS should notify the R&D office for the relevant NHS care organisation of this amendment and check whether it affects R&D approval of the research.

Statement of compliance

The Committee is constituted in accordance with the Governance Arrangements for Research Ethics Committees and complies fully with the Standard Operating Procedures for Research Ethics Committees in the UK.

We are pleased to welcome researchers and R & D staff at our NRES committee members' training days – see details at <http://www.hra.nhs.uk/hra-training/>

11/H1310/1:	Please quote this number on all correspondence
--------------------	---

Yours sincerely



pp
Mr Neil Marsden
Chair


E-mail: nrescommittee.yorkandhumber-southyorks@nhs.net

Enclosures: List of names and professions of members who took part in the review

*Copy to: Prof Colin A Johnson, Leeds Teaching Hospitals NHS Trust
R&D Office, Leeds Teaching Hospitals NHS Trust
Clare Skinner, University of Leeds*

NRES Committee Yorkshire & The Humber - South Yorkshire**Attendance at Sub-Committee of the REC meeting**

<i>Name</i>	<i>Profession</i>	<i>Capacity</i>
Mr Neil Marsden	Police Staff	Lay Plus
Mrs Carole Taylor	Deputy Chief Pharmacist	Expert

The Leeds Teaching Hospitals 
NHS Trust

Ref: Amanda Burt

14/03/2013

Research & Development Directorate

Dr Eamonn Sheridan
Department of Clinical Genetics
Third Floor
Chapel Allerton Hospital
LS7 4SA

34 Hyde Terrace
Leeds
LS2 9LN

Tel: 0113 392 2878
Fax: 0113 392 6397

www.leedsth.nhs.uk/sites/research_and_development

Dear Dr Eamonn Sheridan

Re: **LTHT R&D Number: CG11/9764 (53788/WY) Molecular genetic investigations of autosomal recessive conditions**
REC: 11/H1310/1

Thank you for your E mail dated 14/03/2013 regarding amendment 1 to the above research study.

The amendment may be implemented with immediate effect in the Leeds Teaching Hospitals NHS Trust under the existing NHS Permission. Please note that you may only implement the changes described in the amendment notice or letter

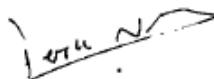
Continued NHS Permission for the project is subject to the following conditions:

- Research Ethics Committee approval/regulatory approval for the amendment, if required, has been obtained
- Any contractual arrangements relating to this change have been addressed
- The Research Lead/Clinical Director for the Directorate has approved any resource implications for the Directorate

If you have any queries about this acknowledgement please do not hesitate to contact the R&D Department on (0113) 392 2878.

With kind regards

Yours sincerely



Dr. D R Norfolk
Associate Director of R&D

The documents reviewed and approved are listed as follows

<i>Document</i>	<i>Version</i>	<i>Date of document</i>
SCP for dealing with requests for Samples Collected during "the study of MPGN"	Final	23/11/2012
MREC Approval Letter	AMD 1	14/03/2013
Lab Sample Processing SOP	2.0	30/06/2012
Protocol	2	18/02/2013
Notice of Substantial Amendment Form	AMD 1	10/12/2012
Cover Letter - E mail		20/02/2013
Participant Consent Form: Assent Form (10/H0106/8)	1	14/12/2010
Participant Consent Form: Consent Form: (10/H0106/8)	1	14/12/2010

Cc: Miranda Squires

2. Primer sequences of the screened genes:

Name	Sequence	Product size (bp)	Name	Sequence	Product size (bp)
CEP164_ex4F	ttttccctgtaagcttg	238	CSPP1_ex9R	caacccttattaaccatgtca	400
CEP164_ex4R	aaaaataaaccagacaatccctg	238	CSPP1_ex10F	aaggctgtaatgtcagtaaagg	278

				aa	
CEP164_ex5F	tcattgatggagagaaaagacaaa	299	CSPP1_ex10R	ctctatccccacctgccac	278
CEP164_ex5R	gaaaagctggccagttcaaaa	299	CSPP1_ex11F	tcttaagtgctgtggtgat	398
CEP164_ex6F	cctttccaccctactctct	400	CSPP1_ex11R	gaataacgctatccgcttca	398
CEP164_ex6R	gcagacacctagagttccc	400	CSPP1_ex12F	tttggtgaatgattggagca	375
CEP164_ex7F	aatttgagcgtgcaggattt	383	CSPP1_ex12R	agcaaagttgtgctttctca	375
CEP164_ex7R	cgaactctgaggactccacc	383	CSPP1_ex13F	atgcaccttttctggagga	244
CEP164_ex8F	ctgggagacattgacatagagagt	399	CSPP1_ex13R	ggatgggtgcttttaggtatg	244
CEP164_ex8R	gggtcacacattgtgaaggt	399	CSPP1_ex14F	gtcaaagcaaaggcaggaag	385
CEP164_ex9F	cactttccctctgcacacc	244	CSPP1_ex14R	ctctgagctctaagggacca	385
CEP164_ex9R	gcattacaacaaagtagcaggg	244	CSPP1_ex15F	gccacgacacctgacctaaa	464
CEP164_ex10_pa rt_1F	gtgggagtcagggaaatagca	495	CSPP1_ex15R	catgctgtctatagaggct	464
CEP164_ex10_pa rt_1R	gtgacttactggctggcgt	495	CSPP1_ex16F	gggcaccaacatatcattgaa	400
CEP164_ex10_pa rt_2F	aggcagaccctacaggcag	248	CSPP1_ex16R	gccaattctcaagatccca	400
CEP164_ex10_pa rt_2R	aagatgtgggatagacc	248	CSPP1_ex17F	ttagcaaaactgttacttttaactg	478
CEP164_ex11F	aattccctgggtaggacc	242	CSPP1_ex17R	cacctagcctgggtagaca	478
CEP164_ex11R	ctgctgatgctaggcaaat	242	CSPP1_ex18F	aggaaactgcaaaaagaggg	371
CEP164_ex12F	gtagtgaagagccctgggtg	236	CSPP1_ex18R	tgtactccaactgttctcctt	371
CEP164_ex12R	agggtcagttctgggtca	236	CSPP1_ex19F	aaggcagcacattttgaat	329
CEP164_ex13F	agcaggaggatggaggga	241	CSPP1_ex19R	caggaacaaaagacctctct	329
CEP164_ex13R	ggccaccaaagccaagaat	241	CSPP1_ex20F	aactataaggggtgtgtgtgt	259
CEP164_ex14F	tccacacactgtactccca	343	CSPP1_ex20R	gaaaattctctcagctccc	259
CEP164_ex14R	acccccatctctcagcac	343	CSPP1_ex21F	tgtaggtttggttctcctt	293
CEP164_ex15F	cccaaatggcgtacatctta	299	CSPP1_ex21R	acacagcacctaaccgctct	293
CEP164_ex15R	ccaccaagagtgaataggc	299	CSPP1_ex22F	gtcaggtgtgttttaagttagtc	289
CEP164_ex16F	ttccatctgtatgttccgt	373	CSPP1_ex22R	tgaattagctctgtcccaaac	289
CEP164_ex16R	agaggggtctccctggctc	373	CSPP1_ex23F	tacaagtgtagccacagca	385
CEP164_ex17F	attggaggtgagagccata	298	CSPP1_ex23R	ttccactttagaaaaaggtaca aa	385
CEP164_ex17R	cacagtggatggtaacggg	298	CSPP1_ex24F	ttttgtgctattttactgtcct	398
CEP164_ex18F	tgtgcttacctgtgagctgc	387	CSPP1_ex24R	atgccttgggaatgaatcaa	398
CEP164_ex18R	ttgcccttctctctgtg	387	CSPP1_ex25F	ttccagtagagagctagcaaaa aa	285
CEP164_ex19F	ggtcggtagtctctctggct	243	CSPP1_ex25R	tccactttactgataagacttagt tc	285
CEP164_ex19R	gagaggctcgctgcctagta	243	CSPP1_ex26F	aacagtggtgaaactgtgcc	232
CEP164_ex20F	cagaacacatccccacacag	278	CSPP1_ex26R	cagtactggttttaagtgtggag	232
CEP164_ex20R	cggctcagctgactgtaa	278	CSPP1_ex27F	ggctctggttagctctgg	250
CEP164_ex21F	accttcaggtgcagcagg	294	CSPP1_ex27R	agcacaggtcagaaccca	250
CEP164_ex21R	catgtgctgcatgtgtg	294	CSPP1_ex28F	ggccctgattcactgatttg	296
CEP164_ex22F	ggggagctgtgattttgtg	285	CSPP1_ex28R	cctcagacgtagtgaagcc	296
CEP164_ex22R	gtctgctccagcaccttc	285	CSPP1_ex29F	tggactacaagagacctgcg	378
CEP164_ex23- 24F	agcccagagtggaggtgt	463	CSPP1_ex29R	aggggcaaaaagtaggacat	378
CEP164_ex23- 24R	tgaacaaaatgtgggaaca	463	TCTN1_ex1_par t_1F	agcttcacaccgctcacta	494
CEP164_ex25F	ccttgcttccctaccctct	338	TCTN1_ex1_par t_1R	actgtgactccaggagctg	494
CEP164_ex25R	tgacacaagcagcagaggtc	338	TCTN1_ex1_par t_2F	caacgcgctgtccatgtc	342
CEP164_ex26F	tggtctgaccacttcacc	288	TCTN1_ex1_par t_2R	attattactgctcaccggg	342
CEP164_ex26R	tggaagtactcccctgatgc	288	TCTN1_ex2F	acctcggaaacttcccgt	232
CEP164_ex27F	gctttgggtcttgaacct	235	TCTN1_ex2R	tggacgtttatggaactcag	232
CEP164_ex27R	ccgtctcctcactcagctc	235	TCTN1_ex3F	tgtgctgtcctcacactttatt	250
CEP164_ex28F	tctggctgtgctcagggft	400	TCTN1_ex3R	cttgaggccaggagttca	250
CEP164_ex28R	ccaaaggactgtttgctga	400	TCTN1_ex4F	atggtacactgtggtggcag	293
CEP164_ex29F	aatggcagggttgggact	247	TCTN1_ex4R	gggttttaagcctgggaaa	293
CEP164_ex29R	actcaacccaatcccaagaa	247	TCTN1_ex5F	acagtacaagcatctgacagtttt	299
CEP164_ex30F	agtgtcccagccacttcta	360	TCTN1_ex5R	aatctgacacattttccataact	299
CEP164_ex30R	agcacctagtccagacaccg	360	TCTN1_ex6F	gcccagccaagacactatt	378
CEP164_ex31F	agctgtctggtgaatggtc	500	TCTN1_ex6R	aggaaccagtttctctggt	378
CEP164_ex31R	caggagaaagaggtagaagaa aa	500	TCTN1_ex7F	gtgggtgcccagtaagtgt	298
CEP164_ex32F	aggcaagaggtgctgggtg	234	TCTN1_ex7R	caaatgtaaagtattaagtgggt caa	298
CEP164_ex32R	gaggtctgctcccctcaga	234	TCTN1_ex8F	tggaggggatttactcatt	300
CEP164_ex33F	gatgtgatgacctgtgtc	299	TCTN1_ex8R	tggcaactaaaaagtttaacta	300

CEP164_ex33R	aaaggggtcaagggctgaaca	299	TCTN1_ex9F	gg tgccattggaataatgaaaa	294
CEP164_ex34F	cctttgacccttcatggc	250	TCTN1_ex9R	aaactcgagtgcaatcatgtact	294
CEP164_ex34R	agatggaagaaggcaggcag	250	TCTN1_ex10F	caactcgagatctgaaaaagca	213
CCDC63_ex2F	caggggtggcttacaagttc	220	TCTN1_ex10R	gcaacaagggtcccctctaa	213
CCDC63_ex2R	cgtgtaaagtgatggcaca	220	TCTN1_ex11F	ggaggttcttccagttggt	300
CCDC63_ex3F	cccactgcccctcactctta	331	TCTN1_ex11R	gagggcgtccagttacc	300
CCDC63_ex3R	agacaagcaacatgaagccc	331	TCTN1_ex12F	gatttgctatctatgggtgtttc	298
CCDC63_ex4F	gaaattctctctcccggtc	376	TCTN1_ex12R	gaaattcagttttatccactgaga a	298
CCDC63_ex4R	ggcctaaggcatcacgaata	376	TCTN1_ex13F	aagaatttttctcctgccatt	230
CCDC63_ex5F	tgctatctgagcacacttcagtt	291	TCTN1_ex13R	gataaagagcagtgagccc	230
CCDC63_ex5R	gctgggcagacatcatttt	291	TCTN1_ex14F	caagtgcactgtcttctactctg a	295
CCDC63_ex6F	gcaagacctgacagccttc	349	TCTN1_ex14R	ttgattaaatggggaaaagtga	295
CCDC63_ex6R	tgcccagagtaggaagcta	349	TCTN1_ex15F	gagctatgatgggtgccactg	389
CCDC63_ex7F	cattctctagcaggaaggc	375	TCTN1_ex15R	ggctgcctcatcacacc	389
CCDC63_ex7R	acaaggagacttggcacct	375	TCTN1_ex16F	tttggggcattgagttacttt	399
CCDC63_ex8F	tcacgtactgtctgctg	370	TCTN1_ex16R	acctgtcacctaccacacttg	399
CCDC63_ex8R	ctaggagactgcccggaaa	370	TCTN1_ex17F	tgaaaagtgtctcattgtgtca	298
CCDC63_ex9F	agggcagactatttttgg	236	TCTN1_ex17R	ttcccaggagactctagcc	298
CCDC63_ex9R	cggacagacagacagaagctc	236	TCTN1_ex18F	ggggaaactgagacgaagcta	273
CCDC63_ex10F	gtttgaaaaccgcagtttg	370	TCTN1_ex18R	ctcccacccacactgtatc	273
CCDC63_ex10R	gcaaggcctatgtctcagtc	370	TCTN2_ex1F	ctgctgctgtttctgtct	239
CCDC63_ex11F	atcccgcacctctgtactt	395	TCTN2_ex1R	agtccaagtctggccctttt	239
CCDC63_ex11R	ggacacggctccccaga	395	TCTN2_ex2F	agtcgacacgccaaagc	282
CCDC63_ex12F	aaaccagccttagcccca	295	TCTN2_ex2R	aaggatgccacctctccag	282
CCDC63_ex12R	atgtgtgtgacaggcagcag	295	TCTN2_ex3F	gcagtctgactcaatgcacc	218
CORO1C_ex1F	tcagatgactggagagcgg	399	TCTN2_ex3R	gagagcactgcagaaacct	218
CORO1C_ex1R	cacatgctcagaggacaggt	399	TCTN2_ex4F	ataaaatgaacggaggctga	383
CORO1C_ex2F	cagggaacgagtggtgg	476	TCTN2_ex4R	gctcaagcaatctcctacc	383
CORO1C_ex2R	agagctcgtttctcattc	476	TCTN2_ex5F	ctggccgataattcagcttt	269
CORO1C_ex3F	tcctttgtctttcacacc	383	TCTN2_ex5R	agctgttggtggaactgctg	269
CORO1C_ex3R	tggtctgttctaattcctt	383	TCTN2_ex6F	agtgtcctgctggccttaa	397
CORO1C_ex4F	gcagacaaaatgccatacca	391	TCTN2_ex6R	ggttgggaaaacgtgacc	397
CORO1C_ex4R	tgaactgtcaatctattaagga a	391	TCTN2_ex7F	agtgaacaaagatcacgcca	368
CORO1C_ex5F	aatgagaaatggcctgaga	277	TCTN2_ex7R	tgcaactgagcctaggaggt	368
CORO1C_ex5R	accacacacatgctttga	277	TCTN2_ex8F	ggctgcagaaagaccagttt	398
CORO1C_ex6F	gcaaatctttatagtgatgtca	297	TCTN2_ex8R	tgcaatgtgtcttacgga	398
CORO1C_ex6R	caaaacacattgtctccc	297	TCTN2_ex9F	cggtcaccataattgtcttg	295
CORO1C_ex7F	cagccgcatcttttctct	398	TCTN2_ex9R	gcagatcttttaactattgtgaac c	295
CORO1C_ex7R	tgcttaaaaagcagaagctc	398	TCTN2_ex10F	acgcaagcagagagaacctc	387
CORO1C_ex8F	gaatggcagttggtgaggtt	288	TCTN2_ex10R	ccttaaatgtcattttgtacac	387
CORO1C_ex8R	cactggaagacctgctcaga	288	TCTN2_ex11F	tgatgacaaaatgcaatttaag g	247
CORO1C_ex9F	aaaaatggctcgttaattgg	296	TCTN2_ex11R	catcctgatggctgggag	247
CORO1C_ex9R	cattgcccgtcctctaaatc	296	TCTN2_ex12F	ttgtattattgacagatgagggc	241
CORO1C_ex10F	tccaactactgctgggttt	358	TCTN2_ex12R	ttgattgcttgattagattcaaaa	241
CORO1C_ex10R	tctgtctcagaatggtcagat	358	TCTN2_ex13F	tgttttcacggaaaactgaa	376
CORO1C_ex11F	ggggatgggtaaagaagcac	248	TCTN2_ex13R	tttccctcatattgatgttgaa	376
CORO1C_ex11R	ggcctgggattttgaattt	248	TCTN2_ex14F	cggacattcactggaatga	250
CORO1C_ex12F	gactgttcccaggaaggtga	495	TCTN2_ex14R	atgctgcgactacagttg	250
CORO1C_ex12R	ttactgaaaaccctgagca	495	TCTN2_ex15F	tcattaatgtgatgctgccc	340
CORO1C_ex13F	gcttcccgttgcaaaata	300	TCTN2_ex15R	ttccctattgttggatgac	340
CORO1C_ex13R	cacaccaataaccagctccc	300	TCTN2_ex16-17F	gattgcatccgttacctgct	564
DYNLL1_ex3F	gggtggggcagttagtg	280	TCTN2_ex16-17R	ttcaagtggcaatattatccg	564
DYNLL1_ex3R	cttaggaaagcaggatcggg	280	TCTN2_ex18F	gtctcgaactggcctcaagt	388
DYNLL1_ex4F	acaggctctggttatcaa	290	TCTN2_ex18R	agtctccggaggctgag	388
DYNLL1_ex4R	tctggtatttggaaattagctg	290	TCTN3_ex1F	aaccctcttcgattggtt	499
ERP29_ex1F	tcactgaccgctgactc	294	TCTN3_ex1R	gctttgtggcagctcaact	499
ERP29_ex1R	aatccagcgtctcccac	294	TCTN3_ex2F	ggaagaaggagggtgagca	293
ERP29_ex2F	ccctgggaagggaatta	281	TCTN3_ex2R	ggtgggaggcacaagacta	293
ERP29_ex2R	ctgttccaagtgggaacga	281	TCTN3_ex3F	tcctcttggcatcatctgg	291
ERP29_ex3_part1F	ccctcagttcagctagttcc	467	TCTN3_ex3R	aatgtggccagggaaga	291
ERP29_ex3_part1R	tgtcatctctgatgctggga	467	TCTN3_ex4F	acacttaagattgcaccctgt	392
ERP29_ex3_part	atcagggcctctggtgtg	389	TCTN3_ex4R	aagggcatgattacgtggac	392

2F					
ERP29_ex3_part_2R	accactttcccactacccc	389	TCTN3_ex5-6F	gtccacgtaatcatgcccctt	496
RAB35_ex1F	tgtttgttcgggaagtggat	390	TCTN3_ex5-6R	ccacaaccacacacagtggag	496
RAB35_ex1R	gaacaggcgcatgactg	390	TCTN3_ex7F	cgccaaaccaagcctatt	211
RAB35_ex2F	tctcccgaagaagggct	248	TCTN3_ex7R	caacttagactaaaattgcctcaga	211
RAB35_ex2R	tggcaaccagaatgagacag	248	TCTN3_ex8F	ggtcaagggggaagggaaaaat	293
RAB35_ex3F	tgtgtgtgtgcagcgtg	300	TCTN3_ex8R	gcaaacaaaattcagttgggt	293
RAB35_ex3R	gaccaccagtgacatttcc	300	TCTN3_ex9F	cacccttgaagacagaaaatct	344
RAB35_ex4-5F	ctcacctttccgggtca	572	TCTN3_ex9R	gaggcctcaatcgacagact	344
RAB35_ex4-5R	aggcactcaataaatggcag	572	TCTN3_ex10F	gtacggtgaagccaagcagt	283
RAB35_ex6F	acaggtagaagagcctgggc	331	TCTN3_ex10R	aaaaattcctcactttggc	283
RAB35_ex6R	aacggcacgaaactgagact	331	TCTN3_ex11F	aacataaatttggcaatgctgc	279
TMEM116_ex1F	gctccatctgtctagggtg	457	TCTN3_ex11R	ttctgactagcattttccg	279
TMEM116_ex1R	agcctggccaaaaagctc	457	TCTN3_ex12F	gctaagagtttctggcaattgtt	399
TMEM116_ex2F	tgtttgagattaggcttgaaga	244	TCTN3_ex12R	atggagacaaggctggtttt	399
TMEM116_ex2R	aaggggaaatgtggaaac	244	TCTN3_ex13F	tggcttagtactgtgattga	388
TMEM116_ex3F	agcctggctgacatagcaa	393	TCTN3_ex13R	tcagtaatcaggcaggggtga	388
TMEM116_ex3R	aagggagccctactgacat	393	TCTN3_ex14F	cagggaaagtaataataacagcaa	410
TMEM116_ex4F	aggccccagaaaaatgatgt	342	TCTN3_ex14R	tcatgagcaggtagggttct	410
TMEM116_ex4R	gtctccactcccaaggtga	342	TMEM107_ex1F	cttgcggggagacttcag	250
TMEM116_ex5F	gcatgtgtcttagttcaggc	218	TMEM107_ex1R	agggtaagacactgggaggg	250
TMEM116_ex5R	gagggctgtctctgtctc	218	TMEM107_ex2-3F	ccctccagtgcttaccct	482
TMEM116_ex6F	gcctccaagttcaagtgag	374	TMEM107_ex2-3R	tttgaactggaaggattggc	482
TMEM116_ex6R	ggtttgggttttcaatgga	374	TMEM107_ex4F	caccactggcctttctgac	243
TMEM116_ex7F	atggttttgctggtatgga	375	TMEM107_ex4R	ctggtaggggaaaaacctt	243
TMEM116_ex7R	gaattcctgggcataagcaa	375	TMEM107_ex5F	agaggtgggtctctggttt	227
TMEM116_ex8F	aagcaactggcctttggaa	241	TMEM107_ex5R	cctatgcttctcttcca	227
TMEM116_ex8R	cctattgttcccactgaa	241	EXOC3L4_ex1_part_1F	acagggccacaacaggttt	462
TMEM116_ex9F	tttgaggagattggtggat	280	EXOC3L4_ex1_part_1R	tcgggttctagttcctcaga	462
TMEM116_ex9R	tggaaagtcttttccagctc	280	EXOC3L4_ex1_part_2F	gaccaggtctccaaggaa	297
TMEM116_ex10F	gtctgtgggggagagacaag	381	EXOC3L4_ex1_part_2R	gagctagggaggacagcctt	297
TMEM116_ex10R	cacaagaagtgctctctgc	381	EXOC3L4_ex2_part_1F	agacaatccagccccgat	567
TMEM116_ex11F	tgataggagatagacctggg	298	EXOC3L4_ex2_part_1R	aaccaagccaccagctc	567
TMEM116_ex11R	ccttcttagtgagaccattgc	298	EXOC3L4_ex2_part_2F	gacagcgaggtgtggac	486
TMEM116_ex12F	ctcagacgtgtgtggaggaa	233	EXOC3L4_ex2_part_2R	gcagctgtcctcgccact	486
TMEM116_ex12R	tcatccttaccgagtaaagtagt	233	EXOC3L4_ex2_part_3F	aggctatctgctgcttcc	217
TMEM116_ex13F	gggagtgagaattattccctgg	352	EXOC3L4_ex2_part_3R	agaagaaggtccggtgctc	217
TMEM116_ex13R	tccttcccaacatttcttcc	352	EXOC3L4_ex3F	gcgttagactttgagcctgg	392
TMEM119_ex1F	cccaccccaaacctctta	292	EXOC3L4_ex3R	gtggtgagaccctccaga	392
TMEM119_ex1R	caacagtcaccagtcctaa	292	EXOC3L4_ex4F	ctgtgaccggacagggc	291
TMEM119_ex2_part_1F	ggctggtgagcctgttct	397	EXOC3L4_ex4R	gaattcgttctggaaatcgc	291
TMEM119_ex2_part_1R	gagcccaccacagcaatc	397	EXOC3L4_ex5-6F	gaggggccaagactgaca	573
TMEM119_ex2_part_2F	acgttctggaggatgtgg	572	EXOC3L4_ex5-6R	cacacactcctcacctctgc	573
TMEM119_ex2_part_2R	ccatgtccctggacttct	572	EXOC3L4_ex7F	ctgtgtccccgctctat	279
TMEM119_ex2_part_3F	ccccagaagaagtacgtgg	558	EXOC3L4_ex7R	aatgaatgctgtgacctggg	279
TMEM119_ex2_part_3R	gagggagtgtcaggaagcag	558	EXOC3L4_ex8F	gagggaaaggggtcaggagt	287
TMEM233_ex1F	ctttcagagcctcgccac	375	EXOC3L4_ex8R	agggtcagggtcagggtc	287
TMEM233_ex1R	aggaagggagccgagag	375	EXOC3L4_ex9F	ctcatggacaaggtggtgac	376
TMEM233_ex2F	cctttcatcaggacagcct	294	EXOC3L4_ex9R	cgtaaggccctagtactg	376
TMEM233_ex2R	accctaacagattccaacgg	294	EXOC3L4_ex10F	ctgacctgcactgacctc	286
TMEM233_ex3F	ataaggaaggtcagctg	225	EXOC3L4_ex10R	ggggtgtcctaataaccag	286

			R		
TMEM233_ex3R	acctgctctggagagtct	225	EXOC3L4_ex11_F	catcctggaatgcagagtga	394
VPS29_ex1F	tacggcaattctgtctccc	218	EXOC3L4_ex11_R	cagagtgactccctgggct	394
VPS29_ex1R	atttccaattcctcgcc	218	TXNDC15_ex1F	ctgccagggttaagatggc	372
VPS29_ex2F	cgtaccggtttaactttgcc	250	TXNDC15_ex1R	ggagagaagtcggggagac	372
VPS29_ex2R	cactctcccagaaaaatgtgt	250	TXNDC15_ex2_part_1F	gggaactgtaattctttgggg	490
VPS29_ex3F	tgctcatttggatgggttt	248	TXNDC15_ex2_part_1R	tttggggatttcagactttca	490
VPS29_ex3R	tcggttactctcttcaacaga	248	TXNDC15_ex2_part_2F	gtcacctgtggtgctggag	398
VPS29_ex4F	atttctgggattcgggatct	397	TXNDC15_ex2_part_2R	tttctactgaagcagccatt	398
VPS29_ex4R	ccagttgagaaaccctggtc	397	TXNDC15_ex3F	ctaccaccctctctctc	397
VPS29_ex5F	gggcaacatagcaacactctg	491	TXNDC15_ex3R	cttccatcccaggacacagt	397
VPS29_ex5R	gcccagccacatttcttta	491	TXNDC15_ex4F	aatctccgtaggtcacacgc	318
VPS29_ex6F	cagatttctgaaatcacaatagtc	380	TXNDC15_ex4R	gaaccacagaaaggcaca	318
VPS29_ex6R	tttacaggaagcttgagca	380	TXNDC15_ex5F	aggactgcatgtcattttgc	396
CSPP1_ex1F	cacagcaggagaacgagttg	491	TXNDC15_ex5R	ttcaagtcaacacgtcactgg	396
CSPP1_ex1R	ccagagcctgtaactctggc	491	DLL1_ex1F	cttttctgccacgctcc	294
CSPP1_ex2F	tggtgactaagcctacatgttgat	283	DLL1_ex1R	ccccccgggattcatctc	294
CSPP1_ex2R	tctcacatagaagaggatgtttg	283	DLL1_ex2F	ctgagccctccaggctct	564
CSPP1_ex3F	tggctaattcctcaaaccttaca	320	DLL1_ex2R	ctgctgggctggagtcct	564
CSPP1_ex3R	tgacaagacaaaacctcttcag	320	DLL1_ex3F	gaatgactgctttttgacct	239
CSPP1_ex4F	gtgctttgttacaggaagtagaact	295	DLL1_ex3R	cacgtgcagaatgaaagctg	239
CSPP1_ex4R	tttgctagtaccactgaaaacct	295	DLL1_ex4F	cttccctgctgaatgtctc	466
CSPP1_ex5F	cccatgctttcagtgacat	399	DLL1_ex4R	ttcccaccagaaacatctct	466
CSPP1_ex5R	tttggttaattggctaattttg	399	DLL1_ex5F	gagttgtctctggcccc	243
CSPP1_ex6_part_1F	ttgcatggctgtatcagctt	673	DLL1_ex5R	ttcgaatgatcacctagggc	243
CSPP1_ex6_part_1R	tctcgaaagtctctatcaaaatca	673	DLL1_ex6-7F	tctagggtgagaatgtccactg	589
CSPP1_ex6_part_2F	gaagaagtgggcatttccaa	389	DLL1_ex6-7R	agcccacacactccattca	589
CSPP1_ex6_part_2R	cccctcagtgggacataaaa	389	DLL1_ex8F	ccctgggttgaatggagtg	381
CSPP1_ex7F	gccaacatcaaatgtaaaaaca	257	DLL1_ex8R	cgaggtcactcacaatgctt	381
CSPP1_ex7R	ttagccagtttaggcacattc	257	DLL1_ex9_part_1F	ggactcattcaggccacaga	592
CSPP1_ex8F	cattttactctgaaatttgctcc	441	DLL1_ex9_part_1R	ggacgcagaccaccacag	592
CSPP1_ex8R	cacctacggaaagcacatca	441	DLL1_ex9_part_2F	actgccagttctctctcc	598
CSPP1_ex9F	ggcatgggcatgattactg	400	DLL1_ex9_part_2R	tgcatgagaacatttgggaa	598
			DLL1_ex10-11F	gaaccactgctccgtttctc	394
			DLL1_ex10-11R	cctctcttcagcagcattc	394

3. Homozygous regions in patient 227.

Coordinates	rsIDs	Size (Mb)
chr1:51557175-52906454	RS11577254;RS269319	1.35
chr1:60297645-74832901	RS17561081;RS696694	14.54
chr1:75397053-82088328	RS1405308;RS709707	6.69
chr1:83474316-89665975	RS12403322;RS4495683	6.19
chr1:92744416-107291289	RS4970705;RS6699202	14.55
chr1:197584242-200239744	RS491369;RS10800806	2.66
chr2:43053652-51252624	RS7576359;RS17577972	8.20
chr2:52113108-58019972	RS4535080;RS2717031	5.91
chr2:60087448-70396890	RS12622148;RS1382457	10.31
chr2:71123543-74323642	RS10202354;RS828902	3.20
chr2:74850052-76229219	RS6740991;RS10204325	1.38
chr2:76821152-79860775	RS4853267;RS4852519	3.04
chr2:163111549-164190114	RS11884875;RS11902715	1.08
chr2:203291499-204401981	RS7571219;RS1968351	1.11
chr3:50346436-52114950	RS2236947;RS13071976	1.77

chr3:97790745-98834034	RS7631577;RS301932	1.04
chr3:139005299-141942918	RS2024369;RS9811699	2.94
chr3:142545854-145812915	RS6765238;RS965343	3.27
chr3:148824362-151887660	RS6796711;RS2116682	3.06
chr3:188201492-192234575	RS11711828;RS6782827	4.03
chr3:193156389-195951280	RS2133607;RS9881429	2.79
chr4:82179489-86572891	RS4389567;RS340203	4.39
chr4:98459674-107142009	RS11737720;RS17036483	8.68
chr5:11481523-13990476	RS31897;RS17278234	2.51
chr5:23764265-25447737	RS6875698;RS16894372	1.68
chr5:26547908-35493048	RS10069706;RS284728	8.95
chr5:36239831-41581643	RS10941278;RS6895481	5.34
chr5:42929682-52039426	RS33814;RS351932	9.11
chr5:53603843-65690100	RS10077344;RS17199162	12.09
chr5:137465939-138513764	RS449965;RS11242450	1.05
chr6:48627226-49789805	RS1361864;RS11759134	1.16
chr7:4350952-6594767	RS4723657;RS3088114	2.24
chr7:8965712-11479011	RS16874906;RS7811562	2.51
chr7:32207484-48162114	RS6462341;RS7458899	15.95
chr7:68370453-69685519	RS1718768;RS719114	1.32
chr7:110158999-111166818	RS12155444;RS2074115	1.01
chr7:118112850-119649537	RS13226251;RS2190183	1.54
chr7:139401288-142726613	RS7810536;RS10216140	3.33
chr7:144659900-146794202	RS6972400;RS11772135	2.13
chr7:147896977-148899655	RS7777242;RS4727092	1.00
chr7:149943582-151190466	RS1916028;RS12703164	1.25
chr8:47966439-49201285	RS7018026;RS7014689	1.23
chr8:99560087-101183757	RS7008395;RS2045677	1.62
chr9:4401273-9015146	RS10974520;RS324490	4.61
chr9:25554892-29266724	RS2498716;RS6476127	3.71
chr9:30052517-37638490	RS1857663;RS4490927	7.59
chr9:38396178-70357832	RS10973789;RS11795256	31.96
chr9:72663416-82235931	RS17555916;RS1833050	9.57
chr10:33965720-37724570	RS2076951;RS2505682	3.76
chr10:44732650-49492963	RS7923091;RS3844493	4.76
chr10:50164791-56074211	RS4838379;RS7905280	5.91
chr10:134403775-135534747	RS2148666;RS10745303	1.13
chr12:105809739-110614582	RS805505;RS659964	4.80
chr12:111833253-115441807	RS3741981;RS2657300	3.61
chr12:116297543-123303371	RS12305372;RS7305702	7.01
chr14:90861514-97475452	RS10147454;RS9635183	6.61
chr14:99096470-101876044	RS2400744;RS2896439	2.78
chr16:31800420-45213017	RS11645176;RS1865837	13.41
chr16:70010107-71459375	RS12444714;RS4788668	1.45
chr18:60402783-61778096	RS9959555;RS7234383	1.38
chr19:19274092-20312172	RS17751061;RS12609651	1.04
chr20:25198577-28084878	RS8184820;RS2379798	2.89

4. Homozygous regions in patient 230.

Coordinates	rsIDs	Size (Mb)
chr1:1-3904110	RS10458597;RS11583257	3.90
chr1:12487152-13668036	RS10864554;RS3013090	1.18
chr1:35151521-36340823	RS4652869;RS7553155	1.19
chr1:72481630-73793151	RS1026566;RS9661557	1.31
chr1:92261111-93322845	RS1927999;RS580828	1.06
chr1:102516897-104176993	RS10874571;RS11185202	1.66
chr1:153063896-154267970	RS16836414;RS7513082	1.20
chr1:212714251-213905713	RS17790562;RS11120594	1.19
chr2:26995019-28157217	RS9309559;RS2337699	1.16
chr2:187236880-188345166	RS3816386;RS1356873	1.11
chr2:193711027-194955833	RS12619354;RS1597691	1.24
chr3:48319771-49853117	RS4336143;RS2271961	1.53
chr3:130227425-152667552	RS395020;RS6790903	22.44
chr4:1-7599724	RS2859203;RS11723455	7.60
chr4:32986285-34044964	RS3098933;RS6846999	1.06
chr4:132715547-133754504	RS350993;RS12510292	1.04
chr4:151691113-152808889	RS12647325;RS6816002	1.12
chr5:35122726-38527160	RS43215;RS3756419	3.40

chr5:44740897-50098368	RS2218081;RS10512906	5.36
chr5:109566164-110570753	RS7380840;RS252858	1.00
chr5:129601565-131364004	RS971891;RS17132288	1.76
chr5:167697178-173144733	RS17731499;RS791362	5.45
chr5:174400961-180915260	RS2036563;RS1298854	6.51
chr6:58477173-64035369	RS2185929;RS6924050	5.56
chr6:83352795-84391944	RS6915395;RS217345	1.04
chr6:87612757-88633520	RS1321317;RS4311490	1.02
chr7:160848-15691629	RS11970804;RS10486772	15.53
chr7:78516290-92925480	RS4437570;RS6970701	14.41
chr7:98281694-99407330	RS7801382;RS4215	1.13
chr7:108526577-109681019	RS1513913;RS4141364	1.15
chr7:118305715-120502390	RS1554904;RS10953925	2.20
chr8:1-1801549	RS7462951;RS11136437	1.80
chr8:50140977-51534019	RS341815;RS1383819	1.39
chr9:43493496-70375735	RS11261805;RS9644996	26.88
chr9:71632360-83498768	RS10511972;RS2796441	11.87
chr9:90504265-101822934	RS1573235;RS3739795	11.32
chr9:103258996-116826994	RS2183745;RS7035322	13.57
chr9:136242947-141213431	RS7027150;RS4295734	4.97
chr10:149404-10442547	RS4880750;RS10905715	10.29
chr10:29439845-68245644	RS7901884;RS7085996	38.81
chr10:73640843-76886733	RS7073342;RS11001397	3.25
chr10:79941492-97970638	RS12218276;RS17111461	18.03
chr10:130223939-134766111	RS3858296;RS7918862	4.54
chr11:64593946-65837787	RS566908;RS10791867	1.24
chr11:72761528-73809544	RS7130689;RS3867273	1.05
chr11:84053389-85256952	RS1940065;RS11234454	1.20
chr12:33716344-37733974	RS7954221;RS1562727	4.02
chr12:59171931-60232119	RS17123871;RS7979611	1.06
chr12:88242227-89249513	RS11105260;RS825962	1.01
chr12:94908476-133851895	RS17676826;RS10747098	38.94
chr13:55705346-57113143	RS9569446;RS9527675	1.41
chr13:82970570-84730686	RS9575317;RS7329244	1.76
chr14:21266050-38127860	RS1018345;RS11624377	16.86
chr14:58873880-60922603	RS7143698;RS12100914	2.05
chr14:62352482-63724184	RS11848904;RS1255984	1.37
chr14:95013301-100542369	RS8020368;RS3809403	5.53
chr15:40113551-66639394	RS7163310;RS2120859	26.53
chr15:91953918-102531392	RS17658377;RS8029360	10.58
chr16:35063218-46743874	RS2200012;RS11861127	11.68
chr16:65787923-66941540	RS3730406;RS11075672	1.15
chr16:68287927-69300526	RS1437134;RS3813909	1.01
chr17:25002820-26626536	RS3098949;RS11870910	1.62
chr17:55393675-56715030	RS1024637;RS9915205	1.32
chr18:14983209-18031438	RS2872415;RS10454095	3.05
chr21:46389340-48129895	RS9978174;RS15047	1.74
chr22:39453741-41178298	RS6002083;RS5758727	1.72

5. Homozygous regions in patient 261.

Coordinates	rsIDs	Size (Mb)
chr1:27394091-28477260	RS7542139;RS7532379	1.08
chr1:35151521-36614436	RS4652869;RS3007220	1.46
chr1:49133207-50560129	RS12043418;RS7339939	1.43
chr1:148504248-149507166	RS2298161;RS11204791	1.00
chr2:189590310-205981949	RS10175784;RS849268	16.39
chr2:227240708-230006547	RS10184436;RS12694813	2.77
chr3:97568370-98847354	RS1877807;RS301947	1.28
chr3:103526734-104572931	RS7616427;RS7653853	1.05
chr4:8135460-13547648	RS6447849;RS493284	5.41
chr4:33257915-34401003	RS10025535;RS16990221	1.14
chr4:81368193-82558540	RS6824301;RS6823416	1.19
chr5:21942850-23191837	RS6452029;RS310912	1.25
chr5:44939787-49859810	RS9791164;RS7708902	4.92
chr6:110283483-111541603	RS12529570;RS9384787	1.26
chr6:145505951-146708451	RS617778;RS362836	1.20
chr7:12652719-20723067	RS6972755;RS6972085	8.07
chr7:68351859-69474084	RS7787868;RS12698901	1.12

chr7:98062972-99298043	RS817771;RS472660	1.24
chr7:118134126-120087274	RS1912492;RS17595350	1.95
chr7:132055624-138046665	RS17166719;RS12707395	5.99
chr7:151190466-153776526	RS12703164;RS2337377	2.59
chr8:89088414-90442285	RS1961470;RS6984875	1.35
chr10:73783339-76029081	RS9415068;RS7088329	2.25
chr11:46767492-56050851	RS11038993;RS658845	9.28
chr11:103340453-105029591	RS881364;RS10502059	1.69
chr12:78047312-79189125	RS12316236;RS2400697	1.14
chr12:85144563-86171734	RS12311272;RS7316333	1.03
chr12:106999780-118407800	RS10746115;RS518202	11.41
chr13:76438356-77560694	RS9530615;RS2775133	1.12
chr13:94926029-96096938	RS17268449;RS523268	1.17
chr14:1-20565973	RS12895974;RS768531	20.57
chr15:19131850-20575519	RS12594669;RS17137361	1.44
chr15:25753962-27216309	RS8035334;RS7167473	1.46
chr15:40884944-42262695	RS10467975;RS12908646	1.38
chr15:58994347-67170519	RS4378570;RS4777108	8.18
chr17:2266925-4688780	RS12452567;RS12450045	2.42
chr17:29006254-30020866	RS280046;RS2190980	1.01
chr18:35218493-36256659	RS16971892;RS9304221	1.04
chr19:1-2971691	RS8100066;RS7249809	2.97
chr19:20424240-21578647	RS10424235;RS2102920	1.15
chr20:5726406-9700408	RS6085278;RS6056855	3.97
chr20:42480762-51238361	RS11086925;RS757365	8.76
chr22:29591756-30994241	RS5749182;RS10439908	1.40
chr22:39380817-40540931	RS133055;RS139568	1.16

6. Cilia recognition protocol:

1. Input image
 - a. Nuclei – DAPI
 - b. Cilia – Alexa Fluor 488
 - c. Cell –DRAQ5 (actually TOTO3)
2. Find nuclei
 - a. Channel: DAPI
 - b. Method: C
 - i. Common threshold: 0.40
 - ii. Area: $>30\mu\text{m}^2$
 - iii. Split factor: 7.0
 - iv. Individual threshold: 0.40
 - v. Contrast: >0.10
 - c. Output population: All nuclei
3. Find cytoplasm
 - a. Channel: DRAQ5
 - b. Nuclei: all nuclei
 - c. Method: D
 - i. Individual threshold: 0.15
 - d. Output population: Cell, Cytoplasm, Membrane
4. Select population (1)
 - a. Population: all nuclei
 - b. Method: common filters
 - i. Remove border objects
 - c. Output population: Whole cells
5. Find spots
 - a. Channel: Alexa Fluor 488
 - b. Population: whole cells

- c. Region: cell
- d. Method: C
 - i. Radius: <3.8px
 - ii. Contrast >0.1
 - iii. Spot to region intensity: >1.3
 - iv. Distance: >5.6px
 - v. Spot peak radius: 0
- e. Output population: Cilia
- 6. Calculate intensity properties (2)
 - a. Channel: Alexa 488
 - b. Population: Whole cells
 - c. Region: Cilia
 - d. Method: Standard
 - i. Mean
 - e. Output population: Intensity of cilia
- 7. Select population (2)
 - a. Population: Whole cells
 - b. Method: Filter by property
 - i. Number of spots on cell = 1
 - c. Output population: One cilium on cell
- 8. Select population (3)
 - a. Population: Whole cells
 - b. Method: Filter by property
 - i. Number of spots on cell ≥ 2
 - c. Output population: Two or more cilia on cell

List of outputs:

Whole cells

-number of objects

-mean intensity of cilia

Cilia

-number of objects

Cells with single cilium

-number of objects

Cells with two or more cilia

-number of objects

Formula output:

$(a/b)*100$

a – one cilium on cell, number of objects (sum)

b – whole cells, number of objects (sum)

output: % cells with single cilium

$(a/b)*100$

a – two or more cilia on cell, number of objects (sum)

b – whole cells, number of objects (sum)

output: % cells with two or more cilia

7. siRNA screen sequence-specific off-target effect gene list:

GENE ID	Ensembl Gene Name	GENE ID	Ensembl Gene Name
11306	Abcb7,Relt,Abcb7	67983	Pdzd9,Pdzd9,Uqcr2
11519	Add2,Add2,Dlg5	68070	Pdzd2,Mfhas1,Pdzd2
11542	Adora3,Gm12824,Adora3	68198	Ndufb2,Ndufb2,Dennd5b
11593	Aga,Aga,Myo3a	68291	Mto1,Mto1,Gm17324
11639	Ak4,Ak4,Foxred1	68318	Aph1c,Aph1b
11656	Alas2,Alas2,Apex2	68401	G6pc3,G6pc3,Ttn
11758	Prdx6,Prdx6,Dsc3	68431	Fbxl15,Cuedc2,Fbxl15
11777	Ap3s1,Ap3s1-ps2,Ap3s1	68553	Col6a4,Col6a4,Dclre1c
11792	Apex1,Tmem55b	68563	Dpm3,Dpm3,Itsn2
11798	Xiap,Xiap,Bcl7a	68591	Mocos,Brox,Mocos
11801	Cd5l,Cd5l,Dach2	68842	Gm2792,Tulp4,Tulp4
11806	Apoa1,Apoa1,Mtmr1	68968	Cdan1,Stard9,Cdan1
11819	Nr2f2,Nr2f2,Papss2	69101	Ydjc,Ydjc,Ccdc116
11829	Aqp4,Aqp4,Olfir911-ps1	69482	Gm4353,Nup35
11840	Arf1,Arf1,Sic24a4	69536	Hemk1,Atp2b2,Hemk1
11854	Rhod,Rhod,Grif1	69663	Ddx51,Ddx51,Mapre3
11863	Arnt,Phrf1,Arnt	69802	Tom11l,Cox11,Cox11
11886	Asah1,Asah1,Pcdh19	69976	Galk2,Galk2,Insr
11951	Atp5g1,Atp5g1,Dst	70047	Trnt1,Trnt1,Zxdb
11975	Atp6v0a1,Prtg,Atp6v0a1	70101	Cyp4f16,Gm9705,Cyp4f16,Cyp4f16,Gm9705
11991	Hnrnpd,Hnrnpd,Cblb	70103	Znhit1,Znhit1,Plod3
12000	Avpr2,Arhgap4	70118	Srrd,Srrd,Tfip11,Tfip11,Srrd
12007	Azgp1,Azgp1,Mcoln3	70208	Dgki,Med23,Med23
12116	Bhmt,Bhmt-ps1	70225	Ppil3,Ppil3,Sgcg
12124	Bik,Bik,Qrs1	70568	Cpne3,Cpne3,Klri2,Klri1
12177	Bnip3l,Fbxl20,Bnip3l	70611	Fbxo33,Fbxo33,Gm16496
12192	Zfp36l1,Zfp36l1,Ptprs	70620	Gm10088,Ube2v2
12228	Btg3,Gm7334	70645	Nusap1,Oip5,Oip5
12263	C2,Gm20547,C2	70673	Prdm16,Prdm16,Mecom
12288	Cacna1c,Cacna1c,Gpr114	70997	Spef1,Spef1,Csnka2ip
12292	Cacna1s,Cacna1s,Megf9	71149	4933413G19Rik,Gnaq,4933413G19Rik
12385	Ctnna1,Ctnna1,Hook3	71336	Rbks,Rbks,Plekhg1
12390	Cav2,Fgfr4,Cav2	71351	5430402E10Rik,Gm14744,5430402E10Rik,Gm14744,Igsf10,Med12l
12399	Runx3,Rab27b,Runx3	71382	Pex1,Pex1,Grid1
12417	Gm6984,Gm10068,Cbx3,Cbx3,Gm6984	71562	Afmid,RP23-268N22.1,Afmid
12457	Ccrn4l,Ccrn4l,Tbl3	71679	Gm10250,Atp5h,Gm4953,Gm10250,Atp5h
12497	Entpd6,Fam5b,Entpd6	71702	Cdc5l,Gm9049,Cdc5l,Cdc5l,Gm9049,Lhx8
12505	Cd44,Cd44,Icam5	71710	Lrrcc1,Lrrcc1,Gnb4
12517	Cd72,Cd72,Mki67	71816	Rnf180,Rnf180,Ctnna1
12560	Cdh3,Cdh3,C030017K20Rik	71876	Mlf1ip,Ccdc111,Mlf1ip
12593	Cdyl,Cdyl,Vps8	71883	Coq2,Gm6728
12617	Cenpc1,Gtdc1,Cenpc1	71898	Apol9b,Apol9b,Apol9a
12627	Cfc1,Prss40,Cfc1	71946	Endod1,Endod1,Fech
12628	Cfh,Cfh,Cfhr2	72053	Tmub2,Atxn713
12684	Cideb,2610027L16Rik,Cideb	72085	Osgepl1,Osgepl1,Tmprss9
12715	A930016O22Rik,Ckm,Ckm	72129	Pex13,Zc3h15,Pex13

12722	Ciqa2,Ciqa1,Ciqa1
12737	Cldn1,Usp34,Cldn1
12824	Col2a1,Col2a1,Dis3
12914	Crebbp,Crebbp,Bptf
12995	Csnk2a1,Csnk2a1,Ii7
13011	Cst7,Cst7,Plcb4
13034	Ctse,Homer1,Ctse
13048	Cux2,Cux2,Slc35f4
13078	Cyp1b1,Col4a3,Cyp1b1
13089	Cyp2b13,Cyp2b13,Dct
13096	Cyp2c54,Cyp2c37,Cyp2c37,Cyp2c50,Cyp2c37
13098	Cyp2c39,Cyp2c39,Cyp2c38
13114	Cyp3a16,Prl8a2,Cyp3a16
13117	Cyp4a31,Cyp4a10,Cyp4a32,Cyp4a10,Cyp4a32,Cyp4a32,Cyp4a31,Cyp4a10
13396	Dlx6,Dlx6,Atg9a
13433	Dnmt1,Ildr2,Dnmt1
13445	Cdk2ap1,Gm12184,Cdk2ap1,Trim7,Gm12184
13510	Dsg1b,Dsg1a,Dsg1a,Dsg1a,Dsg1b,Dsg1b,Dsg1a,Dsg1c
13525	Adam26a,Adam26b
13531	Dub1,Dub1,Thrb
13537	Dusp2,Dusp1,Dusp2
13610	S1pr3,S1pr3,Hist1h1c
13709	Elf1,Elf1,Slc30a7
13713	Elk3,Cdk17,Elk3,Elk3,Tex2,Drp2
13796	Emx1,Emx1,Gm5878
13801	Enam,Enam,A930011G23Rik
13806	Eno1,Gm5506,Eno1
13821	Epb4.1i1,Epb4.1i1,Scml4
13999	Gm14288,Gm14440,Zfp931,Etohi1,Gm14326,Gm14327,Gm14410,Gm14419,Gm14295,Gm14305,Gm14306,Gm14308,Gm14435,Gm14434,Gm14436,Gm2026,Gm14288,Gm4724,Gm4723,Gm14440,Gm14399,Gm14420,Gm14431,Gm8898,Gm14288,Gm4245,Gm4723,Gm14440,Gm14435,Gm14288,Gm14440
14009	Etv1,Gm5454
14017	Evi2a,Evi2a,Rapgef6
14027	Evpl,Evpl,Zxdb
14057	Sfxn1,Setmar,Sfxn1
14168	Fgf13,Akap5,Fgf13
14175	Fgf4,Fgf4,Olf156
14283	Ccdc85b,Fosl1
14359	Dnajc19,Fxr1,Fxr1
14381	G6pdx,G6pdx,Ikbkg
14415	Gad1,Gad1-ps,Gad1
14462	Gata3,Gata3,Gata2
14464	Gata5,Gm14318,Gata5
14534	Kat2a,Kat2a,Myct1
14658	Glrb,Glrb,Pitpnc1
14664	Slc6a9,Ccdc24,Slc6a9
14674	Gna13,Gna13,Foxn2
14676	Gna15,Gna15,Gna11
14680	Mppe1,Gnal,Gnal
14683	AL593857.1,Gnas
14686	Ppat,Gnat2,Gnat2
14706	Gng4,Gng4,Bmx
14710	Gngt2,Gngt2,Fam38b
14719	Gm10874,Got2
14728	Lilrb4,Gp49a,Lilrb4,Gp49a,Cdc27

72199	Mms19,Ii1rap,Mms19
72303	Cyp2c65,Cyp2c66,Cyp2c65
72401	Slc43a1,Nrxn3,Slc43a1
72585	Lypd1,Lypd1,Gpr39
72749	Tonsl,Tonsl,Fnbp1
73122	Tgfbrap1,Tgfbrap1,Pik3c3
73250	Ceacam5,Gm5155,Ceacam5,Ceacam5,Gm5155
73347	1700042B14Rik,Gm16390,Gm4937,1700042B14Rik,Gm16390
73381	Cmtm1,Cmtm2a,Cmtm2a
73382	Prss52,Prss52,Chl1
73523	Pebp4,Pebp4,Cdkn2d,Gm4694
73542	Tssk5,Tssk5,Nup210
73626	Gm2663,1810009J06Rik
74019	Traf3ip1,Traf3ip1,Egfr
74154	Unkl,Gm5819,Unkl
74281	Spatc1,B4galnt3,Spatc1
74563	Cebpz,Rasgef1c,Rasgef1c
74585	Sppl3,Gm10401,Sppl3
74708	4930521A18Rik,E230019M04Rik,4930521A18Rik
74763	RP23-312H15.9,Nat15,Nat15
75051	4930578N16Rik,4930578N16Rik,Pla2g16
75199	Rhox2g,Rhox2e,Rhox2d,Gm20464,Rhox2c,Rhox2b,Rhox2a,Rhox2h,Rhox2g,Rhox2d,Rhox2a,Rhox2f,Rhox2e,Gm20464,Rhox2c,Rhox2b,Rhox2a,Rhox2e,Rhox2d,Rhox2a
75599	Pcdh1,Pcdh1,Ii1r1
75645	1700011F14Rik,1700011F14Rik,Cdan1
75677	Cldn22,Cldn24,Cox10,Cldn22,Cldn24,Cldn22
75692	Nr2c2ap,Rfxank,Nr2c2ap
75778	Them4,Zfp191,Them4
75788	Smurf1,Smurf1,Gatad2a
75847	Ispd,Ispd,D630036H23Rik
75860	4930588N13Rik,Akap6,4930588N13Rik
76089	Rapgef2,Rapgef6,Rapgef2
76367	Trp53rk,2810408M09Rik
76390	Zfp735,Zfp616,Zfp735
76571	Styx11,Styx11,Ptpn13
76688	Arfrp1,Nmnat2
76770	2010005H15Rik,Gm5483,2010005H15Rik
76826	Nubpl,Gm7073,Nubpl
76927	1700021C14Rik,Rhbg,1700021C14Rik
76933	Ifi2712a,Ifi2712b,Ifi2712a,Ifi2712a,Ifi2712b
76964	2610028H24Rik,Ybey,2610028H24Rik
77577	Spns3,Spns3,Zbtb8b
77864	Ypel2,Ypel2,Phactr4
78304	Lsmd1,Ttc7,Lsmd1
78376	Ng23,Ng23,D17H6S56E-3
78655	Eif3j,Gm9781
78889	Wsb1,Wsb1,Fgf7
78891	Scyl1,Scyl1,Accs
78921	9130019O22Rik,Zfp747,9130019O22Rik
79043	Eme2,Spsb3,Spsb3
79410	Klra7,Cngb3
80797	Ciqa2,Ciqa2,Ciqa1
83672	Sytl3,Gm2792,Gm4778,Sytl3,Gm2792,Sytl3
83767	Wasf1,Wasf1,Smarce1

14751	Gm1840,Gpi1,Gpi1
14858	Gsta2,Gm10639,Trpc3,Gsta2,Gsta2,Gsta2,Gm10639,Gm3776,Gsta1
14859	Synpr,Gsta3,Gsta3
14863	Gstm2,Gm6665,Gstm2
14869	Gstp2,Gstp1
14870	Gstp2,Gstp1,Gstp1
14963	H2-BI,H2-BI,1700013N18Rik
14990	H2-M2,Vwde,H2-M2
14999	H2-DMb1,H2-DMb1,H2-DMb2
15000	H2-DMb1,H2-DMb2,H2-DMb2
15130	Hbb-b1,Hbb-b2
15203	Heph,Heph,Gas211
15216	Hfe,Hfe,Tcl1
15228	Foxg1,Myo1b,Foxg1,3110039M20Rik,Foxg1
15233	Hgd,Hgd,Kcng1
15251	Hif1a,Tmed10,Hif1a
15273	Hivep2,Hivep2,Wapal
15398	Hoxa13,Hoxa13,Braf
15416	Hoxb8,Hoxb7,Hoxb8
15436	Hoxd4,Hoxd3
15478	Hs3st3a1,Hs3st3a1,Hs3st3b1
15486	Hsd17b2,Etl4,Hsd17b2
15493	Hsd3b2,Hsd3b3,Hsd3b2
15494	Hsd3b3,Hsd3b3,Hsd3b2,Hsd3b6,Hsd3b3,Hsd3b1
15507	Gm9817,Hspb1,Mon2,Hspb1,Gm9817,Hspb1
15525	Hspa4,Hspa4,Arid1b
15567	Slc6a4,Slc6a4,Gm10410,Gm10339,Gm10251,Gm11109,1700110I01Rik
15904	Id4,Olf727,Olf726,Olf725,Id4
15959	Ifit3,I830012O16Rik,I830012O16Rik,Ifit3
15964	Ifna11,Ifna11,A930038C07Rik
15976	Ifnar2,RP23-190G10.4,Ifnar2
16000	Igf1,Igf1,Rnf111
16155	Il10rb,RP23-190G10.4
16157	Il11ra2,Gm2002,Gm13305,Il11ra1,Il11ra1,Il11ra2,Gm2002,Gm13305
16158	Il11ra2,Gm2002,Gm13305,Il11ra1
16170	Il16,Stard5,Il16
16178	Il1r2,Il1r2,Dock3
16198	Il9,Ccnt1,Il9
16202	Taf10,Ilk
16403	Itga6,Itga6,Naa30
16451	Jak1,Jak1,Ranbp17
16453	Nr3c1,Jak3,Jak3
16551	Kif11,Kif11,Mrps9
16571	Kif4,Kif4,Pdcd11
16573	Kif5b,Kif5b,Zfyve26
16576	9330171B17Rik,Kif7,Kif7
16612	Klk1,Klk1,Klk1b3,Klk1b5,Klk1b9
16638	Klra7,Klra17,Klra7
16642	Klrc2,Klrc3,Klrc2
16679	Krt86,Krt83,Krt86,Dock9,Krt86,Krt81,Krt86
16687	Krt6a,Krt6a,Krt6b
16688	Krt6b,Krt6a,Krt6b
16785	Rpsa-ps10,Rpsa
16834	Cog1,D11Wsu99e,Cog1
16924	Lnx1,Lnx1,Fip111
16997	Ltbp2,D030025P21Rik,Ltbp2
17079	Cd180,Lrrc40,Cd180
17105	Lyz2,Lyz1,Lyz2
17110	Lyz1,Lyz2,Lyz1
17125	Smad1,Smad1,B3galnt2
17258	Mef2a,Mef2a,Wdr26
17259	Mef2b,2310045N01Rik
17354	Mllt10,Dnajc1,Mllt10
17364	Trpm1,AC139849.1
17389	Mmp16,Ccdc17,Mmp16

83984	Tssk6,Tssk6,Ndufa13
83996	Mmp1b,Mmp12,Mmp1b
84004	Mcam,Gm12886,Mcam
84092	Usp8,Usp8,Gpr155
84111	Gpr87,Gpr87,Wdr82
93694	Clec2d,Clec2g,Clec2d,Rad18,Clec2d
93765	Ube2n,Ube2n,Ube2n
93841	Uchl4,Uchl4,Uchl3
93968	Klra10,Klra8,Klra9,Klra9,Klra8,Klra8
94067	Peo1,Mrpl43,Mrpl43
98660	Atp1a2,Atp1a2,Igsf8
99439	Duox1,Lpp,Duox1
100559	Ugt2b38,Ugt2b38,Ugt2b5
102570	Slc22a13b,Slc22a13,Slc22a13
102680	Slc6a20a,Slc6a20a,Slc6a20b
103142	Rdh9,Rdh9,Rdh16,Rdh1
103324	Gm5506,Eno1
104111	Adcy3,Cep250,Adcy3
105663	Thtpa,Ap1g2,Thtpa
105689	Mycbp2,Mycbp2,Usp48
106795	Tcf19,Zhx3,Tcf19
106878	2010002N04Rik,2010002N04Rik,Bcl2l1
106957	Slc39a6,Slc39a6,Heatr6
107141	Cyp2c50,Cyp2c50,Cyp2c54
107272	Psat1,Psat1,Zfp658
107328	Fermt3,Trpt1,Trpt1
107885	Mthfs,Mthfs,Gm2382,Nova1,Mthfs,Gm2382
107986	Ddb2,Ddb2,Acp2
108079	Prkaa1,Prkaa2,Prkaa2
108148	Gm20388,Galnt2
108989	Tpr,Tpr,Cacna2d1
109075	Exosc4,Exosc4,Oplah
109113	Uhrf2,Uhrf2,Ktn1
109910	Zfp91,Bcl9,Zfp91
110312	Pmch,4930547N16Rik,Pmch
110460	Acat2,Acat3,Acat2
110842	EtfA,EtfA,Chpt1
112406	Egln2,Fnip2,Egln2
113868	Acaa1a,Acaa1b
114228	Prss1,Gm5771,Prss1,Gm5771,Try5,Prss1,Prss1,Prss2,Try10,Try5,Try4
114255	Dok4,Dok4,Polr2c
114671	4930444G20Rik,4930444G20Rik,Vegfa
114893	Tes3-ps,Dcun1d1
117158	Scgb3a2
117160	Ttyh2,Ttyh2,Dach1
140571	Plxn3,Zfp384,Plxn3
140780	Bmp2k,Paqr3,Bmp2k
140858	Wdr5,Wdr5,Lamb2
170789	Acot8,Acot8,Snx21
170930	Sumo2,Samd8,Sumo2
171180	Syt12,Syt12,Kcng2
171196	Vmn1r22,Vmn1r22,Galntl2
171200	Vmn1r19,Ankrd33b,Vmn1r19
171202	Vmn1r16,Vmn1r16,Vmn1r12,Vmn1r16,Tnfsf10
171210	Acot1,Acot2,Acot2
171259	Vmn1r193,Vmn1r193,Atp2b1
192159	Prpf8,Zkscan2,Prpf8
192176	Flna,Flna,Zfp184
195349	Btnl7,Btnl7,Btnl4
208117	Aph1b,Aph1c,Aph1b,Aph1b,Aph1c
208595	Gm9897,Mterf
208666	Diras1,Diras2,Diras1
209186	Acnat1,Acnat2,Acnat2
211151	Fntb,Churc1,Churc1
213211	Gm9008,Rnf26

17523	Mpo,Mpo,Rasgrf2
17755	Mtap1b,Mtap1b,Urb2
17874	Myd88,Acaa1a,Myd88
17876	Gm9833,Myef2,Myef2
17885	Myh8,Myh4,Myh8,BC125332,Myh8
17938	Naca,Ric3,Naca
17951	Naip5,Naip5,Naip6
17952	Naip6,Naip5,Naip6,Naip6,Naip5
17967	Ncam1,Gm11149,Ncam1
17977	Ncoa1,Ncoa1,Dak
17999	Nedd4,Nedd4,Bhlhe22
18032	Nfix,G430095P16Rik,Nfix
18038	Gm16181,Nfkbil1,Nfkbil1
18039	Nefl,Nefm,Nefl
18102	Nme1,Gm20390,Nme1
18115	Iqcf1,Nnt
18127	Nos3,Nos3,Aanat
18130	Ints6,Ints6,Ptgr2
18162	Npr3,Npr3,Mklin1
18241	Gpr143,Fam126b,Gpr143
18263	Odc1,Odc1,Lmln
18301	Fxyd5,Fxyd5,Fndc3b
18312	Olf15,Olf15,Sez6l2
18317	Olf2,Olf2,Zranb2
18328	Olf3,5730522E02Rik,Olf3
18359	Olf59,Olf59,Olfr406-ps,Olfr59
18386	Oprd1,Oprd1,Akap17b,AI314180
18477	Prdx1,Gm7204,Prdx1
18521	Pcbp2,Map3k12,Pcbp2
18567	Tbp,Pdcd2
18574	Pde1b,Dera,Pde1b
18575	Pde1c,Fndc3b,Pde1c
18582	Pde6d,Pde6d,Ccdc38
18606	Enpp2,Erbp4,Enpp2
18648	Gm8910,Pgam1,Pgam1,Gm8910
18701	Pigf,Prkaa2,Pigf
18707	Pik3cd,Pik3cd,2410004P03Rik
18711	Pikfyve,Pikfyve,Gm15140,Gm10057,4921511M17Rik,Gm15144
18724	Pira2,Gm15448,Gm14548,Pira2,Gm14548,Lilrb3,Lilra6,Gm15448,Gm14548
18781	Pla2g2c,Pla2g2c,Ubxn10
18782	Pla2g2d,Pla2g2d,Mepce
18857	Pmp2,Dcaf17,Pmp2
18950	Pnp2,Pnp,Pnp,Pnp2
18973	Pole,Pole,Bche
19046	Ppp1cb,Ppp1cb,Apbb2
19128	Pros1,Golim4,Pros1
19139	Prps1,Prps1,Prps113,Prps113,Prps1
19156	Psap,Psap,Slc1a6
19188	Psme2,Psme2b-ps
19230	Twf1,Twf1,Irak4
19250	Ptpn14,Ptpn14,Ankrd55
19265	Ptprcap,Ptprcap,Nlr1
19377	Rai1,Dnahc1,Rai1
19654	Rbm6,Rbm6,Mfsd2b
19664	Rbpsuh-rs3,Rbpj
19668	Rbpjl,Rbpjl,Trim41
19696	Rel,Rel,Gm16503
20022	Polr2j,Polr2j,Gpt2
20185	Ncor1,Ncor1,Smarca5
20190	Ryr1,Ryr1,Ryr3,Ryr2
20209	Saa2,Saa2,Saa1
20256	Clec11a,Clec11a,Plcx3
20274	Scn9a,Scn9a,Scn1a
20276	Scnn1a,Scnn1a,4930597O21Rik
20306	Ccl7,Ccl7,Irgb8
20341	Selenbp1,Selenbp2
20342	Selenbp1,Selenbp2

213311	Fbxl21,Fbxl21,Ift88
213409	Lemd1,Lemd1,Irga9
213980	Fbxw10,Trim16,Fbxw10
214254	Nudt15,Gm5519,Nudt15,Gm5519,Ets1
214766	Mmp21,Ar,Mmp21
215095	Astl,Astl,1700028116Rik
215384	Fcgbp,9530053A07Rik,Fcgbp
215418	Csrnp1,Ipo4,Csrnp1
215476	C330019G07Rik,RP23-118L1.1,C330019G07Rik
215929	AI317395,G630090E17Rik,AI317395
216150	Cdc34,Cdc34-ps,Cdc34
216233	Gm9847,Socs2
216635	Hbq1a,Hbq1b,Hbq1a,Pygo1,Hbq1a,Hbq1b
216848	Chd3,Gm17305,Chd3
216965	Taok1,Taok1,Abhd15
217166	Nr1d1,Nr1d1,Thra
217369	Uts2r,Uts2r,Catsperg1
217734	Pomt2,Pomt2,Scrn3
217866	Cdc42bpb,Cdc42bpb,Ociad2
218268	Eif4e1b,Eif4e1b,Gm5705
223254	Farp1,Farp1,Galnt9
223649	Nrbp2,Nrbp2,Psd3
223922	Atf7,Atf7,Atf2
223978	Cpped1,Cpped1,Dyrk1a
224530	Acat2,Acat3,Acat3
224624	Rab40c,Rab40c,Ehhadh
224762	Trim31,Trim31,I830134H01Rik
225152	Gjd4,Gjd4,Bai2
225256	Dsg1b,Dsg1a,Dsg1b
225579	Slc27a6,Stim2,Slc27a6
225872	Npas4,Npas4,Gm5936,Gm5640
226356	Gm101,Gm101,Rbms2
226418	Yod1,Pfkfb2,Yod1
226695	Ifi205,Ifi205,Mnda,Mndal
226791	Lyplal1,9630033F20Rik,Lyplal1
227099	Pms1,Pms1,Pkhd111
227648	Sec16a,Flrt1,Sec16a
228536	Bahd1,Bahd1,Lin7a
228576	Mall,Mall,4930415O20Rik
229574	Fig2,Gm11744,Fig2
229780	Ccdc76,Lrrc39,Ccdc76
229782	Slc35a3,Slc35a3,Hiat1
230396	Ifna13,Ifna13,Specc1
231103	Gckr,Gckr,Thoc2
231162	Cyt11,Cyt11,Sh2d4b
231510	Agpat9,Agpat9,Ank2
231580	Gak,Gak,Tax1bp1
231637	Ssh1,Ssh2,Ssh1
231889	Bud31,Ptcd1,Bud31
233001	Nlrp9a,Nlrp9a,Nlrp9b
233210	Ppargc1b,Prr12,Prr12
233870	Gm9755,Tufm
233879	Asphd1,Rbfox1,Asphd1
234366	Gatad2a,Gatad2a,Dnajb14
234733	Ddx19b,Ddx19a,Ddx19b
234779	Plcg2,Plcg2,C1qa
235344	Sik2,Sik2,Fry
235623	Scap,5830462I19Rik,Scap
235674	Acaa1b,Acaa1a,Acaa1b
236193	Zfp709,Gm17576,Zfp709
236266	Alms1,Alms1,Zmynd19
236539	Gm6756,Phgdh,Phgdh
237928	Phospho1,Abi3,Phospho1
238130	Dock4,Dock4,1810041L15Rik
238405	Adam6b,Adam6b,Adam6a
238406	Adam6a,Adam6b,Adam6a
238829	1700042B14Rik,Gm16390,Gm4937,Gm4937

20371	Foxp3,Foxp3,Dcakd	241159	Neu4,E330009J07Rik,Neu4
20383	Gm12355,Srsf3	241431	Xirp2,4930523C07Rik,Xirp2
20390	Sftpd,Sftpd,Ggn	242202	Pde5a,Pde5a,1600015110Rik
20431	Pmel,Pmel,Cep290	242700	Il28ra,Arhgap20
20437	Siah1b,Siah1a,Siah1a,Siah1b	242726	Padi6,Tanc1,Padi6,Tmtc4,Padi6
20438	Siah1b,Siah1a,Siah1b	243537	Uroc1,Uroc1,Sema5a
20505	Slc34a1,Slc34a1,Pfn3	243755	Slc13a4,Csk,Slc13a4
20527	Slc2a3,Slc2a3,Lingo2	243967	Ntn5,Ntn5,Sec1,Ntn5,Gm7092
20534	Slc4a1ap,Slc4a1ap,Ubr3	244202	Nlrp10,Nlrp10,Gm20516
20557	Slfn3,Slfn3,Slfn4	244667	Disc1,AC168060.2
20583	Snai2,Snai2,Wapal	245509	4932429P05Rik,4932429P05Rik,4930415L06Rik
20597	Smpd1,Ppic,Smpd1	245615	Kir3dl2,Kir3dl1,Kir3dl2
20620	Plk2,Plk2,Vmn1r33	245616	Kir3dl1,Kir3dl2,Kir3dl1
20667	Sox12,Sox12,Ankrd27	246257	Dph1,Ovca2,Ovca2
20684	Sp100,AC132444.3,AC168977.2,AC133103.7,Sp100,AC125149.8	246730	Oas1a,Oas1a,Oas1g
20700	Serpina1c,Serpina1a,Serpina1e,Serpina1c,Serpina1a,Serpina1d,Serpina1e,Serpina1c,Serpina1a,Serpina1d,Serpina1b	246779	Il27,Ttn,Il27
20701	Serpina1e,Serpina1a,Serpina1d,Serpina1b,Serpina1c,Serpina1b,Serpina1e,Serpina1c,Serpina1a,Serpina1d,Serpina1b	246791	Obox3,Amh,Obox3
20702	Serpina1e,Serpina1c,Serpina1a,Serpina1d,Serpina1b,Serpina1c,Serpina1a,Serpina1b,Serpina1c,Serpina1a,Serpina1d,Serpina1b,Serpina1e	252912	Vmn1r188,Vmn1r188,Gm
20703	Serpina1e,Serpina1d,Serpina1b,Serpina1c,Serpina1a,Serpina1e,Serpina1c,Serpina1a,Serpina1d,Serpina1d	257899	Olfr1000,Olfr1000,Olfr996
20704	Serpina1e,Serpina1d,Serpina1b,Serpina1c,Serpina1a,Serpina1e,Serpina1c,Serpina1a,Serpina1d,Serpina1b,Serpina1c,Serpina1e	257947	Olfr543,Olfr543,Synm
20768	Sephs2,Sephs2,Gm16485	257951	Olfr988,Olfr987
20775	Sqle,Sqle,Diap2	258067	Olfr1359,Olfr1360
20878	Aurka,Aurka,Rprd2	258101	Olfr1137,Olfr1137,Olfr1013
21367	Cntn2,Olfr394,Olfr381,Cntn2	258155	Olfr1425,Olfr1425,Olfr1424,Olfr1426,Olfr1425,Olfr1424
21390	Tbxa2r,Tbxa2r,Mxi1	258156	Olfr605,Olfr605,Sorbs1
21415	Gm15401,Tcf7l1,Tcf7l1	258160	Olfr685,Olfr685,Cars2,Sp8
21665	Tdg,Gm5806,Tdg	258166	Olfr988,Olfr987,Olfr988
21683	Tecta,Tecta,Zbtb42	258181	Olfr59,Olfr406-ps
21749	Terf1,Errf1,Terf1	258706	Olfr43,Olfr403
21754	Cd72,Tesk1,Tesk1	259097	Olfr558,Olfr558,Gm8251
21802	Tgfa,Tgfa,Epb4.1	259101	Olfr628,Olfr628,Olfr243
21823	Th,Lifr,Th	260298	Fev,Fev,Scn10a
21925	Tnnc2,Ube2c	266744	Lgsn,Lgsn,4931428L18Rik
21968	Gm5884,Tom1	268482	Krt12,Krt12,Krt42
22018	Tpo,Deptor,Tpo	269378	Gm4737,Ahcy
22038	Plscr1,Plscr1,Plscr4	269695	Rnf2,Fgf1,Rnf2
22139	Ttr,Ttr,Vmn2r102	270599	Gm648,Gm648,Gm10477
22142	Tuba1c,Tuba1a,Tuba1b,Slc7a11,Tuba1a,Tuba1c,Tuba1b,Tuba1a,Slc7a11,Tuba1c,Tuba1a,Tuba1b	271457	Rab5a,Galnt4,Rab5a
22143	Tuba1c,Tuba1a,Tuba1b,Tuba1b,Tuba1c,Tuba1a,Slc7a11,Tuba1c,Tuba1a,Tuba1b	276865	Olfr1371,Olfr1373,Fmn1,Olfr1371,Olfr1371,Olfr1373,Olfr1380,Olfr1381,Olfr1382,Olfr1371,Olfr1373,Olfr1380,Olfr1381
22144	Tuba3a,Tuba3b,Tuba3a	277973	Slc9a5,Prpf18,Slc9a5
22147	Tuba3b,Tuba3a,Tuba3b	278174	Ssxb3,Ssxb3,Ssxb9,Ssxb5,Ssxb3,Marcks
22151	Tubb2a,Tubb2b,Tubb2a	280645	B3gat2,B3gat2,Sertad2
22186	Gm11808,2810422J05Rik,Uba52,Gm5239,Gm11808,Gm7866,2810422J05Rik,Uba52,Gm5239,Gm11808,2810422J05Rik,Uba52	317677	Gm5077,Gm5077,C1s
22214	Ube2h,Gm2058	319195	Rpl17,Gm10268,Rpl17,Rpl17-ps3,Gm10362,Rpl17,Rpl17-ps3
22272	Uqcrcq,Epb4.1,Uqcrcq	319387	Lphn3,Synrg,Lphn3
22287	Scgb1a1,Oxtr,Scgb1a1	320484	Rasal3,A530088E08Rik,Rasal3
22290	Uty,Slc4a1ap,Uty	320615	Pgm3,Dopey1,Dopey1
22325	Vav2,Vav2,Mixip	320795	Pkn1,Pkn1,Ptger1
22340	Vegfb,Vegfb,Hk2	320864	Krt26,Krt26,Rnf213
22346	Vhl,Vhl,Ros1	320995	Rfx6,Rfx6,Rnf169
22350	Ezr,Ezr,Liph	327957	A430084P05Rik,A430084P05Rik,Pign
22362	Vpreb1,Vpreb2,Vpreb1	327978	Slfn5,Slfn5,Aak1

22363	Vpreb1,Vpreb2,Vpreb2,Vpreb1
22410	Wnt10b,Wnt10b,Atp10b
22590	Xpa,Xpa,Ttn
22598	Slc6a18,Ccdc36,Slc6a18
22599	Slc6a20b,Slc6a20b,Slc6a20a,Slc6a20a,Slc6a20b
22612	Yes1,Fyn,Yes1
22639	Zfa,Zfx,Zfa
22673	Zfp185,Zfp185,5430403G16Rik
22764	Zfa,Zfx,Zfx
23832	Xcr1,Dcdc5,Xcr1
23980	Pebp1,Pebp1,Rgl2
23983	Pcbp1,Pcbp1,Pcbp2
24046	Scn11a,Osbpl3,Scn11a
24053	Sgcg,Sacs,Sgcg
24075	Taf10,Gm4799
26386	Hsf4,Cep250,Hsf4
26425	Nubp1,Fam18a,Nubp1
26426	Nubp2,Nubp2,Clec2h
26442	Gm8394,Psma5,Psma5,Gm8394
26446	Gm4950,Psmb3,Psmb3,Gm4950,Gm5356,Psmb3
26888	Clec4a2,Clec4a2,Clec4a4
26909	Exo1,Exo1,Nsun2
26945	Tpsg1,Tpsg1,Elk4
27027	Tspan32,Tspan32,Ptger3
27223	Trp53bp1,Trappc8,Trp53bp1
27384	Akr1c13,Akr1c12,Akr1c13
27388	Ptdss2,Fstl4,Ptdss2
28080	Atp5o,Gm5436,Atp5o
28185	Tom70a,Cnrm4,Tom70a
28250	Slco1a4,Slco1a4,Gm5724
29849	Olf159,Olf159,Ung
30843	Fbxl12,Fbxl12,Akap6
50720	Sacs,Sacs,Vmn2r42
50772	Mapk6,Hadhb,Mapk6,Rnf169,Mapk6
50790	Acsl4,Acsl4,Dlc1
50797	Copb2,Copb2,Tinf2
50908	C1s,Gm5077,C1s
50911	Ccna2,Exosc9,Exosc9
50933	Uchl3,Uchl3,Uchl4,Uchl4,Uchl3
50997	Mpp2,Nebi,Mpp2
51791	Rgs14,Rgs14,Arl4a
51795	Srpx,Micalcl,Srpx
51885	Tubgcp4,Tubgcp4,Trp53bp1
52830	Pnrc2,Pnrc2,Mettl16
52838	Dnlz,Dnlz,Gpsm1
52855	Lair1,Lair1,Sh2d1b2
53314	Batf,Batf,Cadm2
53622	5430421N21Rik,Krt86,Krt81,Krt83,ltpr3
53625	Comm1d1,B3gnt2
53880	Naip5,Naip6
53893	Nudt5,Gm13199,Nudt5
54124	Cks1b,Cks1b,Gm6531

328424	Kcnrg,Psd3,Kcnrg
329384	Pthr1,Pthr1,Wdr4
329559	Zfp335,Zfp335,Pcif1
330390	Gm765,Gm765,Gabbr2
330450	Far2,Col8a1,Far2
330627	Trim66,Trim66,Dnm3
330814	Lphn1,Gm10644,Lphn1
330830	Ccdc135,Ccdc135,Zfp442
330962	Ostb,Gcc2,Ostb
331188	Zfp781,Gm3055
338354	Zfp780b,Zfp780b,1700049G17Rik
338521	Fa2h,Fa2h,Tln2
353187	Nr1d2,Nr1d2,Med13
380773	1810035L17Rik,1810035L17Rik,Vamp7,1810035L17Rik,2700073G19Rik
380855	Zfp429,Zfp456,Rsl1,Rsl1
380928	Lmo7,Lmo7,4933414I15Rik
381107	Tmem232,Tmem232,Lrrtm4
381236	Gm5097,Lipo1,Lipo4,Lipo2,Gm8975,Lipo2,Lipo1,Lipo4,Lipo1
381246	Xkr9,Xkr9,Apc
381287	AC133103.2,A530032D15Rik,AC125149.7,AC132444.4,AC125149.7,Arhgef18,AC132444.4,A530032D15Rik,AC125149.7,AC132444.4,AC125149.7
381308	Mnda,lfi205,Mnda,lfi205,Mnda,Mndal
381406	Trp53rk,2810408M09Rik
381417	Gm14085,Slc28a2,Gm14085
381531	Mup21,Mup3,Mup21
381823	Apold1,Apold1,C130039O16Rik
381852	Gm5155,Ceacam5,Gm5155
381936	Gm5890,Gm5891,Gm4545,Gm6882,Gm10668,Gm5156,Gm6902,Gm10662,Gm4567,Gm5890,Gm5891,Gm4545,Gm6882,Gm10668,Gm5156,Gm6902,Gm10662,Gm4567,Gm5878,Gm6176,Gm5890,Gm4545,Gm10668,Gm5156,Gm6902,Gm4567,Gm6176
382007	Adam26a,Adam26b,Adam26b
382045	Gpr114,Actn2,Gpr114
382156	Fbxw22,Fbxw16,Fbxw22
382209	Rhox3h,Rhox3f,Rhox3e,Rhox3c,Rhox3a,Rhox3h,Rhox3f,Rhox3e,Rhox3a
382864	Colq,Colq,Ryr3
384220	Vmn2r16,Mad11,Vmn2r117,Vmn2r16,Vmn2r14,Vmn2r16
384806	4930529F22Rik,4930529F22Rik,Adam25,4930529F22Rik,Adam25,Adam39
404195	Cyp2c54,Cyp2c50,Cyp2c54
406219	Krt83,Krt83,Krt86
407243	Tmem189,Gm20431,Hivep1,Tmem189,Gm20431,Tmem189
415115	Ctsa,Neur12,Neur12
432467	Hnrnp33,Hnrnp33,Rufy2
432479	4930404N11Rik,4930404N11Rik,Mfsd12
432736	Vmn1r209,Vmn1r206,Vmn1r209,Vmn1r209,C77370
432779	Lrrc14b,Ccdc127
432879	Gm5465,Kbtbd7,Gm5465
433016	2010005H15Rik,Gm5483,Taf6,Gm5483,2010005H15Rik,2010005H15Rik,Gm5483,Fbxl20,Gm5483
433182	Gm5506,Eno1
433215	BC048609,BC048609,1110014N23Rik
433416	Gm13547,Trub2,Gm13547
433638	I830077J02Rik
433759	Hdac1,Gm10093
433801	Gm13157,Gm13212,Gm13157,Gm13154,Gm13212,Gm13157,Gm13154,Gm13051,Gm13251,Gm13212,Gm13139
434171	Gm5591,Gm5591,Tmem164
434325	Tmem221,Tmem221,TEX14

54140	Avpr1a,Avpr1a,Ssh1
54196	Pabpn1,Gm20521
54393	Gabbr1,Gabbr1,Mtap1b
54709	Eif3i,Eif3i,Gm749
55927	Hes6,Hes6,Svil
55981	Pigb,Pigb,Ccpg1
55991	Panx1,Panx1,Zfp507
56173	Cldn14,Cldn14,Qars
56210	Rev1,Rev1,Eif5b
56275	Rbm14,Rbm14,Nlpr1c-ps
56289	Rassf1,Rassf1,Eif2c1
56291	Gm14698,Styx
56297	Arl6,Arl6,Rab22a
56309	Mycbp,Kif21a,Mycbp
56327	Arl2,Arl2,Esco1
56335	Mettl3,Tox4,Mettl3
56378	Arpc3,Arpc3,Gpr126
56388	Cyp3a25,Cyp3a59,Cyp3a25,Cyp3a25,Cyp3a59
56455	Dynll1,Gm6788,BC048507,Gm6788,Dynll1,BC048507,Dynll1
56456	Mrpl47,Actl6a,Actl6a
56489	Ikbke,0610009O20Rik,Ikbke
56495	Asna1,4930556J24Rik,Asna1
56506	Cib2,Cib2,Slc9a6
56532	Ptger2,Ripk3
56752	Aldh9a1,Myst4,Aldh9a1
57349	Pbbp,BC025920,Pbbp
57430	Sult3a1,Sult3a1,Limk1
57442	Kcne3,Kcne3,Pi4ka
57773	Wdr4,Wdr18,Wdr4
57914	Crlf2,Myo7b,Crlf2
57916	Tnfrsf13b,Tnfrsf13b,Gabpb2
58179	Klrc3,Klrc2,Klrc3
58521	Eid1,Shc4,Eid1
58810	Akr1a1,Ccpg1,Akr1a1
58988	Rps6kb2,lpo11,Rps6kb2
59001	Pole3,Pole3,Pik3c2a
59015	Nup160,Nup160,Phf16
59020	Pdzk1,Pdzk1,Gpr89
59021	Rab2a,Rab2a,Gm5113
59044	Rnf130,Rnf130,Sepsecs
59044	Rnf130,Rnf130,Sepsecs
59054	Mrps30,Mrps30,ltp1
59092	Pcbp4,Pcbp4,Fam159a,Pcbp4,Chdh

434758	Rhox3h,Rhox3e,Rhox3a,Rhox3h,Rhox3h,Rhox3f,Rhox3e,Rhox3c,Rhox3a
434764	Rhox2f,Rhox2e,Rhox2c,Rhox2b,Rhox2h,Rhox2f,Rhox2f,Rhox2e,Gm20464,Rhox2c,Rhox2b,Rhox2a,Rhox2f
434766	Rhox2g,Rhox2g,Jarid2,Rhox2h,Rhox2g,Rhox2d,Rhox2g,Rhox2d,Rhox2c
434768	Rhox8,Rhox8,Ctdp1
434903	Mageb4,Mageb10-ps,Mageb4
435529	Gpr111,Gpr111,Olfr988,Olfr987
435889	Try5,Try4,Try5
436002	Olfr628,Olfr243
436240	Foxr2,Foxr2,Slc9a2
436523	Gm5771,Prss1,Gm5771,Gm5771,Prss2,Prss1,Try10,Try5,Try4
474156	Ggnbp1,Zbtb9
544696	D630037F22Rik,D630037F22Rik,Hoxd11
544748	Olfr765,Olfr767,Olfr765,Olfr765,Camk2d
545288	Cyp2c67,Cyp2c69,Cyp2c40,Cyp2c68,Cyp2c68,Cyp2c67,Cyp2c67,Cyp2c69,Cyp2c67,Ankrd13d,Tll6,Cyp2c40,Cyp2c68
545366	Cfh,Cfhr2,Cfhr2
545471	Zfp345,Gm14124,Zfp345,Gm14139
545725	Gm9897,Mterf
546049	C330021F23Rik,C330021F23Rik,Cir1
546118	Ubtfl1,Ubtfl1,Naalad2
546546	Serpina3i,Dlgap1
546912	Vmn2r22,Vmn2r21,Zfp711,Vmn2r21,Vmn2r20,Vmn2r22,Vmn2r21,Vmn2r20
546913	Vmn2r22,Vmn2r21,Zfp711,Vmn2r22,Vmn2r22,Vmn2r21,Vmn2r20
546980	Vmn2r74,Vmn2r74,Vmn2r73,Vmn2r70,Vmn2r68-ps,Vmn2r67,Vmn2r66
547127	AC175035.2,Tmem181c-ps
547347	Gm7030,Gm19684,Gm11127,Gm6034,H2Q1,Gm7030,Gm19684,Gm11127,Gm6034
574418	Serinc4,2310003F16Rik,Serinc4,Sec11c,2310003F16Rik,Serinc4,Serf2,Serinc4
594844	Tceal3,Tceal6,Tceal3
619310	Eny2,Zfp872
619331	Zfp551,Zfp551,Syne1
621697	Gm10064,Rpl32,Rpl32-ps,Gm4987,Rpl32,Gm10064,Rpl32
621852	Rhox3h,Rhox3e,Rhox3a,Rhox3h,Rhox3h,Rhox3f,Rhox3e,Rhox3c,Rhox3a
622301	Rhox2h,Rhox2h,Rhox2g,Rhox2d,Rhox2h,Rhox2g,Rhox2d,Rhox2a
622402	Akr1c12,Akr1c12,Akr1c13
623131	Prr19,Rxra,Prr19
623661	Lipt1,Mitd1,Lipt1
623734	Vmn2r85,Vmn2r86,Vmn2r85,Vmn2r84,Vmn2r87,Vmn2r85,Vmn2r87,Vmn2r86,Vmn2r84,Vmn2r87,Vmn2r85,Vmn2r84,Vmn2r86
624681	Btnl6,Btnl6,Btnl7
624860	Gm12253,Gm12253,Flna
625068	Vmn2r86,Vmn2r84,Vmn2r85,Vmn2r86,Vmn2r85,Vmn2r84,Vmn2r87,Vmn2r84
626316	Gm13051,Gm13154,Gm13151,Gm13225,Gm13154,Gm13151,Gm13051,Zfp534,Gm13151,Gm13051,Gm13235,Gm13157,Gm13154,Gm13051,Gm13251,Gm13212,Gm13139
627111	Vmn2r92,Vmn2r91,Vmn2r107,Vmn2r104,Vmn2r101,Vmn2r100,Vmn2r99,Vmn2r98,Vmn2r97,Vmn2r96,Vmn2r95,Vmn2r93,Vmn2r94,Vmn2r-ps113,Vmn2r92,Vmn2r92,Vmn2r91,Rab40b
627576	Vmn2r101,Vmn2r101,Ube2e3
627805	Vmn2r108,Vmn2r108,Vmn2r109,Vmn2r10

11306	4
11350	2
11423	5
11426	4
11464	1
11472	1
11477	1
11482	3
11487	1
11491	1
11492	1
11496	11
11497	1
11517	1
11518	1
11519	5
11538	1
11564	1
11593	1
11596	1
11639	2
11647	2
11656	6
11670	1
11684	1
11699	1
11717	1
11733	3
11735	1
11758	1
11769	1
11775	1
11781	1
11784	2
11789	1
11790	1
11792	7
11793	1
11796	2
11798	3
11801	4
11804	1
11806	7
11810	1
11816	2
11818	2
11836	1
11854	1
11855	1
11863	6
11870	1
11877	1
11886	2
11890	1
11908	1
11909	4
11920	2
11937	1
11951	3
11958	1
11964	1
11975	1
11980	3
11987	1
11988	1
11991	3
11994	2
12000	8
12006	1
12007	1
12014	1

21823	2
21827	1
21843	1
21885	1
21886	2
21915	1
21925	1
21928	1
21929	2
21945	5
21955	1
21974	1
22017	1
22018	3
22021	1
22022	1
22031	1
22040	1
22041	1
22042	1
22061	1
22062	1
22064	4
22121	1
22129	5
22138	17
22138	17
22186	5
22193	1
22196	2
22230	2
22240	2
22249	2
22272	4
22276	3
22278	1
22284	2
22290	3
22295	1
22321	1
22337	2
22339	1
22340	1
22352	1
22359	2
22371	4
22410	1
22433	1
22590	17
22612	7
22629	1
22631	2
22634	3
22639	2
22668	1
22673	1
22688	2
22715	5
22719	2
22722	2
22724	4
22757	1
22770	2
22773	1
22778	2
23790	2
23792	1
23802	1
23805	3
23827	1
23832	2

101437	3
101471	1
101488	2
101497	1
101502	1
101568	3
101772	3
102098	1
102141	1
102414	1
102570	1
102626	3
102632	1
102857	1
103324	2
103406	1
103537	1
103573	1
103677	1
103737	2
103988	1
104027	1
104111	3
104248	3
104681	1
105148	1
105239	5
105377	2
105670	6
105689	10
106200	2
106344	1
106585	1
106633	1
106795	3
106878	7
106957	1
107260	1
107272	2
107448	1
107476	3
107522	2
107581	3
107589	1
107686	1
107817	1
107823	6
107831	1
107932	1
107971	1
107986	2
108000	1
108058	1
108073	2
108075	1
108078	1
108096	1
108099	1
108655	7
108857	3
108903	1
108960	1
108989	3
109032	1
109054	1
109075	3
109263	1
109331	1
109333	1
109658	1
109674	4

12029	1
12035	1
12036	1
12040	1
12045	6
12047	6
12054	1
12057	1
12121	1
12122	1
12153	1
12175	1
12177	4
12180	2
12192	1
12263	1
12266	2
12287	3
12288	2
12292	1
12293	1
12296	2
12305	1
12322	2
12335	1
12337	2
12355	2
12361	2
12385	1
12387	1
12390	1
12396	2
12399	2
12402	1
12411	1
12421	4
12457	2
12495	1
12544	1
12545	1
12554	1
12593	4
12617	6
12627	1
12628	2
12655	1
12704	3
12705	1
12721	2
12733	1
12737	1
12738	1
12747	1
12748	1
12753	1
12778	1
12788	1
12790	1
12804	3
12816	1
12818	1
12819	2
12822	1
12824	1
12827	1
12830	2
12833	1
12870	2
12896	2
12908	6
12912	3

23850	1
23871	1
23917	2
23950	1
23965	3
23980	1
23983	1
23984	1
23997	1
24012	2
24046	4
24053	7
24086	2
24100	1
24105	1
24136	6
26357	3
26373	1
26386	2
26404	3
26422	1
26423	1
26431	2
26462	1
26885	1
26889	1
26909	1
26912	1
26945	2
26965	2
27027	2
27053	3
27054	1
27056	1
27057	3
27103	1
27223	2
27389	1
27399	1
27411	1
28015	1
28240	1
28250	1
28254	1
29809	1
29856	2
29861	1
29871	1
30838	1
30841	3
30843	1
30941	1
30956	1
50496	1
50523	1
50721	1
50757	1
50762	1
50768	2
50769	2
50770	2
50780	5
50789	1
50790	3
50797	3
50873	1
50908	2
50915	2
50935	3
50995	1
50997	5

109689	2
109700	1
109900	8
109910	1
109929	1
109934	3
109979	2
110094	1
110157	1
110310	1
110312	3
110385	1
110542	1
110593	2
110651	2
110695	2
110784	2
110789	1
110821	3
110842	5
110854	2
110876	1
110911	1
110948	1
113858	3
113868	1
114142	5
114255	1
114600	1
114643	2
114671	3
114716	1
114873	2
114875	1
116837	1
116873	1
117147	1
117149	1
117160	1
117600	1
140571	2
140630	1
140780	2
140781	1
170638	1
170648	1
170707	7
170722	2
170738	2
170743	2
170744	1
170789	3
170791	2
170828	1
170930	1
171166	1
171180	3
171183	14
171196	1
171200	2
171202	1
171207	2
171259	1
171580	1
192119	1
192159	2
192166	1
192176	3
192185	3
192285	4
192292	3

12914	2	51789	1	192786	1
12916	2	51792	1	193116	1
12955	1	51885	3	194227	4
12967	1	52397	8	207165	1
12972	3	52563	1	207181	1
12977	2	52679	2	207596	1
12995	2	52685	1	207777	1
13003	2	52830	4	207806	1
13011	2	52838	2	208084	1
13017	12	52855	1	208092	1
13034	12	53325	1	208266	1
13047	1	53333	1	208449	1
13048	4	53357	1	208643	5
13075	1	53376	1	208650	1
13077	4	53413	1	208748	1
13078	1	53417	1	208869	3
13086	1	53418	1	208967	1
13089	1	53619	1	208982	1
13096	1	53625	4	209011	1
13117	1	53860	1	209018	1
13132	1	53870	1	209027	2
13170	2	53883	3	209186	1
13360	1	53885	1	209446	1
13385	3	53896	1	209462	1
13386	1	54139	1	209837	1
13390	2	54140	1	210162	2
13396	1	54153	1	210356	1
13405	3	54201	2	210710	1
13409	1	54338	1	211007	3
13419	1	54357	1	211151	1
13424	1	54369	1	211429	1
13426	1	54380	3	211535	4
13429	2	54383	2	211651	2
13433	5	54393	3	211660	8
13435	5	54403	1	211673	1
13445	4	54411	1	211712	1
13446	1	54427	3	212032	1
13487	1	54446	1	212285	2
13496	1	54563	1	212427	1
13506	1	54607	1	212712	2
13521	1	54611	1	212862	3
13527	2	54613	2	212933	1
13528	2	54614	1	213011	1
13548	1	54633	1	213053	1
13549	2	54652	2	213084	3
13557	1	54667	1	213402	1
13601	1	54722	1	213409	2
13605	4	55927	3	213819	1
13619	5	55981	3	213980	3
13631	1	55983	2	213990	1
13636	1	55991	1	214137	1
13654	1	55992	2	214254	3
13666	1	56046	1	214384	1
13667	1	56055	2	214444	2
13690	3	56075	1	214642	1
13713	6	56086	1	214855	1
13728	6	56173	3	214952	2
13733	3	56215	1	215113	1
13801	2	56218	1	215332	1
13806	3	56248	1	215476	1
13808	2	56275	2	216021	1
13821	4	56289	1	216152	1
13831	1	56299	1	216164	1
13848	2	56309	1	216233	2
13852	1	56314	1	216443	1
13858	1	56315	1	216616	1
13860	1	56324	1	216739	2
13864	2	56347	1	216742	1
13990	1	56361	2	216805	1
13999	39	56365	1	216810	1
14009	2	56375	1	216848	3

14011	1
14017	10
14030	1
14043	2
14050	1
14068	1
14069	2
14081	1
14088	2
14102	1
14107	1
14126	1
14134	1
14161	1
14163	1
14164	1
14168	4
14180	1
14182	2
14183	1
14186	1
14194	1
14219	1
14232	1
14256	2
14260	1
14261	1
14264	1
14275	6
14312	1
14359	3
14375	2
14381	14
14387	3
14391	3
14394	2
14407	1
14412	1
14423	1
14429	1
14463	3
14567	1
14585	4
14598	3
14634	1
14645	1
14658	1
14660	2
14664	6
14674	1
14680	1
14681	2
14683	5
14686	1
14710	2
14728	3
14747	1
14748	1
14755	3
14783	3
14794	1
14814	1
14815	4
14823	1
14824	1
14852	1
14853	1
14854	1
14859	1
14885	3
14886	2

56380	2
56389	1
56403	2
56405	1
56409	1
56422	1
56454	3
56460	1
56463	1
56484	2
56489	1
56506	1
56542	2
56717	1
56727	1
56734	2
56738	3
56749	1
56752	2
56839	1
56873	1
56874	3
57138	2
57278	1
57349	2
57357	1
57430	1
57442	1
57738	2
57742	1
57751	1
57810	1
57916	4
58186	1
58194	1
58226	3
58230	1
58231	1
58234	1
58521	1
58800	1
58800	1
58802	3
58810	4
58988	1
58992	1
58998	4
59004	1
59010	2
59015	1
59020	3
59024	1
59026	2
59030	1
59047	1
59047	1
59049	1
59054	1
59056	1
59057	1
59092	3
59126	4
63959	3
64009	2
64113	1
64291	4
64337	3
64652	6
64685	2
64705	1
64918	1

216856	1
216860	1
216867	1
216964	1
216965	1
217166	4
217198	1
217258	2
217316	2
217353	3
217356	1
217369	2
217734	1
217826	1
217893	1
218214	1
218268	2
218454	1
218885	1
223254	1
223272	1
223435	1
223649	2
223664	2
223669	1
223770	2
223774	1
223922	8
223978	3
224170	1
224619	1
224727	3
224742	1
224836	1
225010	1
225115	1
225579	2
225600	2
225876	1
225888	3
226169	4
226243	1
226251	2
226356	1
226418	5
226519	1
226747	2
226861	2
226970	2
227099	4
227377	1
227399	1
227634	3
227648	1
227659	1
227696	1
227720	1
227738	2
228005	1
228061	1
228366	1
228608	1
228684	1
228829	1
228839	1
228850	1
228880	1
228960	1
229003	2
229285	2
229574	1

14923	1
14924	1
14950	1
14962	2
15114	1
15170	1
15184	3
15201	1
15203	3
15216	3
15228	1
15233	1
15273	4
15275	14
15277	1
15357	3
15387	2
15398	3
15410	1
15427	1
15434	1
15436	5
15445	1
15451	1
15465	1
15476	1
15481	2
15483	2
15486	11
15493	3
15516	1
15525	2
15547	1
15569	2
15586	1
15900	1
15958	1
15975	1
16004	2
16150	1
16151	1
16169	2
16170	3
16178	3
16179	4
16190	1
16198	1
16328	3
16329	2
16330	2
16331	2
16336	4
16372	1
16392	1
16403	1
16409	2
16410	3
16412	1
16414	2
16418	1
16430	1
16439	1
16443	5
16451	3
16453	6
16468	3
16516	2
16517	2
16531	1
16532	1
16549	1

64945	1
65111	2
65254	1
66065	1
66078	1
66177	11
66179	4
66204	1
66209	1
66213	2
66234	2
66251	1
66354	3
66369	1
66399	1
66416	1
66425	4
66435	1
66495	11
66505	2
66510	2
66531	4
66552	1
66576	1
66589	2
66596	5
66656	3
66659	1
66680	1
66813	1
66848	1
66854	1
66881	1
66904	1
66914	2
66930	1
66942	1
66958	1
66985	1
66993	3
67016	6
67063	2
67071	1
67141	4
67155	1
67260	2
67269	7
67283	3
67292	1
67300	1
67323	4
67331	1
67338	1
67420	1
67426	6
67452	1
67469	1
67492	1
67533	1
67534	1
67569	3
67573	1
67575	1
67579	5
67665	1
67772	1
67873	6
67911	4
67941	1
67949	1
68021	1

229603	2
229603	2
229615	1
229715	1
229725	1
229782	1
230098	1
230099	1
230101	1
230396	1
230484	1
230582	2
230587	1
230594	2
230700	3
230737	1
230753	6
230779	1
230815	1
230857	1
230861	4
231103	5
231252	2
231510	7
231549	2
231580	3
231637	3
231798	2
231889	2
231915	1
232078	1
232232	1
232341	8
232791	1
232798	4
232807	1
232811	1
232933	1
232944	1
233001	3
233040	1
233046	1
233060	1
233328	2
233406	4
233575	4
233805	2
233833	1
233865	3
233876	1
233878	1
233904	3
234135	2
234214	8
234258	1
234311	2
234515	1
234582	1
234734	1
234788	1
235180	1
235312	3
235344	2
235380	1
235431	2
235574	7
235582	1
235587	2
235610	1
235611	2
235623	2

16551	2
16559	1
16561	2
16563	2
16568	1
16571	1
16574	1
16580	3
16594	1
16634	1
16644	1
16649	1
16653	1
16679	1
16706	2
16728	2
16777	1
16796	1
16801	1
16818	1
16818	1
16834	2
16847	1
16871	1
16875	1
16881	4
16882	2
16885	1
16890	1
16905	1
16912	1
16924	6
16969	2
16971	3
16973	1
16976	1
16997	1
17079	2
17125	1
17127	1
17160	1
17165	1
17171	1
17179	1
17220	1
17222	1
17250	1
17251	9
17258	1
17259	2
17279	2
17283	1
17295	1
17299	1
17346	1
17354	3
17364	4
17385	1
17389	1
17523	4
17524	2
17535	2
17536	3
17540	2
17685	1
17686	1
17756	6
17762	3
17764	1
17765	1
17769	2

68070	3
68135	1
68198	2
68279	1
68291	1
68294	1
68379	3
68401	19
68420	1
68490	1
68505	1
68553	3
68556	2
68563	1
68591	1
68655	2
68703	1
68708	1
68729	1
68732	3
68750	4
68775	2
68778	1
68795	2
68842	2
68846	1
68879	1
68922	1
68926	3
68938	4
68968	4
69024	1
69101	2
69192	1
69207	1
69282	2
69398	1
69536	2
69574	1
69634	1
69662	2
69663	2
69684	5
69719	1
69724	2
69743	1
69802	4
69807	1
69837	1
69847	1
69863	3
69922	1
69940	1
69976	1
70103	1
70118	1
70120	1
70208	3
70225	2
70297	2
70300	1
70351	4
70354	1
70355	1
70358	1
70450	2
70530	1
70561	1
70611	1
70645	1
70661	1

235626	1
235674	1
236798	1
237175	1
237178	1
237500	2
237615	1
237806	2
237831	2
237860	2
237898	1
237928	1
237940	1
238023	1
238055	1
238247	2
239099	1
239122	1
239134	1
239337	1
239420	1
239556	3
239719	2
240034	1
240880	2
241066	1
241112	1
241159	1
241201	1
241274	3
241431	1
241489	1
241494	1
241589	2
241694	1
242202	1
242362	1
242409	1
242443	1
242502	1
242669	1
242726	5
242939	1
242960	1
243277	1
243407	1
243537	5
244059	3
244329	2
245000	1
245109	1
245269	2
245670	1
245827	1
245860	1
245880	1
246179	1
246190	2
246257	1
246313	1
246727	1
246728	1
246779	16
252912	1
252973	1
257979	2
258156	1
258925	1
260296	1
260315	1
263803	1

17773	1
17776	2
17827	1
17857	1
17873	4
17874	5
17879	1
17883	1
17884	1
17885	3
17886	2
17896	1
17906	1
17909	6
17919	1
17921	2
17933	1
17937	2
17938	1
17951	2
17952	2
17966	2
17967	1
17968	1
17996	2
18000	2
18004	1
18022	3
18023	2
18029	1
18036	1
18039	1
18127	1
18128	1
18130	3
18143	1
18174	2
18181	2
18189	1
18190	2
18191	1
18193	1
18195	1
18201	1
18220	1
18222	1
18226	1
18241	3
18301	1
18312	1
18386	6
18412	1
18424	3
18452	1
18479	1
18481	1
18484	4
18504	1
18508	2
18510	1
18514	1
18519	2
18521	3
18550	2
18567	9
18573	1
18575	1
18578	2
18583	1
18585	1
18606	2

70673	6
70676	1
70696	1
70827	1
70846	2
70894	1
71007	3
71059	1
71147	1
71148	2
71302	1
71330	3
71336	1
71351	2
71354	2
71355	1
71365	1
71382	1
71389	2
71435	2
71458	2
71544	1
71567	1
71602	1
71609	1
71623	1
71675	2
71702	4
71710	2
71713	1
71715	1
71722	4
71729	1
71752	2
71760	1
71770	2
71790	1
71862	1
71868	1
71876	1
71902	1
71946	1
71991	1
71999	2
72061	2
72145	2
72149	1
72151	1
72194	1
72199	3
72401	4
72475	1
72479	1
72560	3
72605	3
72667	1
72668	2
72720	1
72749	14
72754	1
72823	2
72828	2
72831	4
72843	1
72999	5
73078	1
73106	1
73122	2
73166	2
73254	1
73261	1

264064	2
266692	1
266744	1
266781	2
268379	1
268482	1
268782	2
268857	1
268878	1
268903	1
268977	1
269113	2
269224	1
269254	2
269275	1
269400	4
269437	2
269582	1
269587	4
269615	3
269682	1
269695	4
269823	1
270076	1
270118	1
270120	1
270627	1
270757	1
271209	1
271564	1
271639	1
271786	1
276865	4
277468	2
277939	1
278174	1
280645	1
317677	2
319188	1
319217	1
319504	2
319586	1
319636	1
319748	1
320091	1
320100	1
320119	2
320226	3
320244	4
320267	2
320495	1
320595	3
320615	1
320634	1
320661	1
320718	1
320790	1
320799	1
320806	1
320864	4
320951	1
321006	1
327655	2
327957	4
327978	1
328424	3
328572	1
328699	1
328795	1
329251	2
329384	1

18607	1
18611	1
18647	2
18670	1
18675	1
18703	1
18705	1
18707	3
18709	1
18710	1
18724	1
18725	1
18739	1
18741	2
18759	1
18762	3
18770	1
18781	2
18786	1
18795	1
18803	2
18806	1
18807	2
18810	2
18822	2
18828	1
18854	2
18857	5
18861	2
18971	1
18973	2
18984	1
18986	4
18987	1
19025	1
19039	2
19046	13
19073	2
19085	4
19087	1
19116	7
19125	1
19128	2
19167	1
19201	1
19205	1
19206	1
19230	3
19250	4
19256	1
19266	2
19267	1
19270	1
19272	1
19273	1
19277	2
19280	1
19283	1
19300	1
19360	1
19377	1
19395	2
19668	3
19679	1
19682	3
19691	3
19699	1
19699	1
19700	1
19701	1
19702	3

73332	5
73340	4
73381	5
73382	1
73430	1
73542	2
73647	1
73674	1
73690	1
73750	3
73804	1
73945	1
73998	3
74007	1
74019	3
74044	1
74103	3
74104	1
74112	1
74121	1
74145	2
74178	1
74180	1
74198	1
74203	2
74204	2
74205	1
74270	1
74309	2
74340	3
74347	1
74351	1
74365	1
74370	1
74374	1
74442	1
74525	2
74552	1
74616	1
74646	4
74760	2
74762	2
74769	1
74776	1
74782	2
74840	1
75051	3
75079	1
75212	1
75273	1
75288	1
75317	1
75388	1
75475	2
75497	1
75570	1
75578	2
75580	3
75599	4
75671	1
75770	3
75778	3
75860	1
76014	1
76089	3
76113	1
76233	2
76251	1
76265	1
76295	1
76306	1

329421	1
329828	1
329910	2
330064	1
330177	1
330189	1
330260	1
330409	2
330627	4
330830	1
330962	2
331535	1
332221	1
333564	3
333654	3
333883	2
338354	1
338365	2
338521	4
353187	1
353234	1
353236	1
360213	1
380711	1
380753	5
380855	1
380969	1
381107	5
381199	2
381246	2
381280	1
381287	3
381405	1
381406	2
381569	1
381591	2
381626	2
381628	1
381680	1
381686	1
381801	2
381833	1
381924	2
382045	2
382066	1
382113	1
382209	1
382236	1
382252	2
382253	1
382864	3
383619	1
384220	1
394430	5
394433	5
394434	5
404710	1
406218	1
408065	1
414089	2
414872	1
415115	2
432442	3
432467	4
432508	1
432555	2
432736	1
432779	2
432825	3
433016	5
433182	2

19714	1
19726	1
19730	2
19731	1
19876	2
20112	1
20181	1
20185	7
20190	2
20216	2
20226	1
20230	2
20239	3
20256	2
20265	1
20269	1
20271	2
20274	1
20300	1
20301	4
20317	1
20346	1
20353	2
20356	1
20371	1
20382	1
20390	2
20391	1
20393	1
20397	2
20429	1
20438	1
20466	2
20480	2
20498	1
20504	1
20505	1
20509	1
20512	1
20527	4
20534	3
20535	1
20536	1
20541	1
20557	1
20562	1
20563	1
20568	1
20583	4
20587	1
20591	1
20595	2
20602	2
20603	1
20620	1
20623	3
20624	1
20658	1
20681	1
20683	3
20684	1
20700	1
20701	1
20703	1
20704	1
20715	2
20719	6
20740	1
20775	5
20778	1
20815	1

76367	2
76376	1
76390	1
76484	2
76551	2
76571	1
76574	1
76582	1
76612	1
76686	2
76719	2
76781	1
76826	1
76843	1
76884	1
76886	2
76927	1
77011	1
77044	1
77097	3
77531	2
77559	1
77569	2
77593	1
77626	2
77632	1
77669	1
77683	2
77864	4
77945	1
77980	1
78038	1
78283	1
78304	2
78376	1
78459	1
78473	3
78771	2
78787	1
78816	1
78829	2
78891	3
78911	1
78926	1
78933	2
79043	1
79202	2
79221	1
79565	1
80290	1
80879	1
80890	4
80904	1
80906	1
80911	1
80915	1
81840	2
83560	3
83672	1
83679	3
83703	5
83704	1
83766	3
83925	1
83996	3
84092	3
84585	1
93691	1
93694	2
93697	4
93736	2

433215	3
433247	1
433416	2
433466	2
433700	1
433719	8
433801	5
434171	5
434179	1
434325	3
434446	2
434459	1
434674	2
434689	1
434766	5
435285	1
435529	2
474156	2
504193	4
544678	1
544696	1
544748	3
544817	3
544988	43
545007	43
545288	3
545370	1
545388	3
545389	1
545428	1
545471	1
545554	1
545637	1
545693	1
545762	3
545812	2
545913	1
546118	4
546144	2
546546	5
546912	1
546913	1
550619	1
574418	4
574438	2
594844	1
619331	4
620078	5
622976	1
623131	4
623272	1
623281	1
623474	1
623661	1
624860	4
624866	2
625321	2
626316	9
627081	2
627111	2
627576	2
627914	4
630146	4
630579	1
632687	1
637027	1
637515	5
639774	3
639781	1
641376	2
654362	1

20852	2	93742	3	664723	1
20869	1	93760	1	664725	1
20878	1	93761	2	664987	10
20912	1	93841	1	665210	2
20916	1	93871	2	665229	4
20918	3	93968	1	665902	1
20972	1	94067	1	665943	3
20981	2	94089	3	666060	1
21335	2	94176	1	666168	2
21343	2	94178	1	666173	1
21355	1	94180	1	667214	4
21356	1	94187	1	668030	5
21374	1	94190	1	668039	4
21391	1	94191	1	672911	1
21411	1	94217	1	751865	1
21415	2	94282	2	791260	8
21417	4	97130	2	100037282	1
21423	1	97287	2	100038804	1
21425	7	98660	3	100038948	1
21453	1	98999	1	100038993	4
21454	1	99151	1	100039008	1
21462	2	99152	1	100039089	9
21665	1	99371	1	100039672	2
21676	1	99377	2	100039948	1
21677	1	99633	2	100040276	1
21679	2	99683	1	100040591	10
21683	2	100072	4	100040861	1
21685	1	100090	1	100040972	1
21744	1	100213	7	100041077	1
21752	1	100561	2	100041379	1
21754	4	100683	1	100041574	2
21770	1	100705	1	100041586	2
21787	1	100710	1	100041621	2
21802	4	100737	1	100041639	2
21807	1	100978	1	100041658	1
21814	1	100986	2	100043100	5
21815	2	101100	1	100048885	1

9. siRNA screen – array CGH results.

-2 indicates loss of two alleles, -1 loss of one allele, 0 allelic balance, 1 gain of one allele, 2 gain of two alleles:

GENE ID	GENOMIC COPY NUMBER	GENE ID	GENOMIC COPY NUMBER	GENE ID	GENOMIC COPY NUMBER
11298	1	22334	-1	80910	1
11302	1	22337	1	80911	1
11303	-1	22339	-1	80979	1
11304	1	22344	1	80981	1
11306	-1	22350	1,0	80987	-1
11364	1	22354	-1	81000	-1
11416	1	22355	1	81004	1
11418	1	22359	1	81535	1
11419	1	22362	1	81600	1
11428	-1	22363	1	81703	1
11429	1	22365	1	81840	1
11430	1	22367	1	81897	-1
11434	1	22375	1	81904	-1
11441	-1	22376	-1	81905	-1,0
11471	-1	22378	1	81906	-1
11472	-1	22384	1	83379	1
11474	1	22385	1	83383	1
11475	1	22388	1	83396	1
11479	1	22393	1	83398	1
11481	-1	22401	1	83454	-1
11482	1	22402	1	83456	1

11487	-1
11488	1
11489	-1
11490	1
11495	-1
11496	1
11500	-1
11501	-1
11504	1
11512	1
11513	1
11514	1
11515	1
11518	1
11519	-1,0
11522	1
11529	1
11532	1
11534	-1
11542	-1,0
11544	1
11549	-1
11551	1
11565	1
11569	-1
11595	-1
11600	1
11605	-1
11607	-1
11608	1
11624	-1
11632	1
11639	-1
11651	1
11652	-1
11655	-1
11656	-1
11668	1
11694	1
11699	-1
11717	-1
11720	-1
11722	1
11732	1
11736	1
11744	-1
11745	1
11747	1
11750	-1
11752	-1
11754	1
11765	1
11766	-1
11767	1
11768	-1
11769	1
11771	-1
11772	-1
11773	1
11774	-1
11775	-1
11778	-1
11781	1
11784	-1
11785	-1
11792	-1,0
11796	-1
11797	-1
11798	-1,1
11799	1
11801	-1,1

22410	1
22411	-1
22412	1
22414	1
22415	1
22416	1
22418	-1
22422	1
22423	1
22439	-1
22589	-1
22590	-1,0
22594	-1
22598	-1
22599	-1
22601	-1
22608	-1,0
22619	-1
22627	1
22628	1
22629	1
22632	1
22639	-1
22644	1
22658	1
22661	1
22668	1
22680	1
22688	-1
22691	-1
22697	1
22719	-1
22724	1
22756	-1
22757	1
22758	-1
22762	1
22764	-1
22770	1
22773	-1
22780	1
22787	-1
22788	1
23790	1
23794	1
23802	1
23806	-1
23806	-1
23807	-1
23821	-1
23831	1
23834	1
23849	-1
23857	1
23859	-1
23871	-1
23873	-1
23881	1
23888	-1
23908	1
23912	1
23920	1
23921	1
23923	1
23938	-1
23947	-1
23948	1
23950	1
23955	-1
23959	-1
23960	1

83486	-1
83490	1
83493	-1
83560	1
83563	-1
83603	-1
83669	-1
83674	1
83679	1
83702	-1
83704	1,0
83762	1
83766	1
83767	1,0
83797	1
83925	1
83946	-1
83984	1
83996	-1
84004	-1
84095	1
84111	-1
84112	1
84505	1
84585	-1
85031	1
93679	1
93681	-1
93683	-1
93684	1
93685	1
93686	1
93697	-1
93721	1
93732	-1
93742	1
93746	-1
93747	-1
93761	-1
93762	1
93835	1
93836	-1
93837	-1
93841	-1
93871	1
93961	1
94043	-1
94045	1,0
94061	1
94067	1
94088	-1
94093	1
94093	1
94112	1
94180	-1
94181	-1
94187	1
94190	-1
94191	-1
94216	-1
94219	1
94245	-1
94246	-1
94253	-1
94279	1
94280	1
94281	1
94315	1
94353	-1
97064	1
97212	1

11803	-1
11804	-1
11806	-1,0
11812	-1
11815	1
11816	-1
11818	1
11819	1
11820	1
11830	1
11831	1
11832	-1
11835	-1
11836	-1
11840	1
11841	1
11842	1
11843	-1
11847	1
11848	-1
11854	-1,1
11856	-1
11857	-1
11859	-1
11863	-1,1
11864	-1
11865	-1
11877	1
11878	-1
11881	-1
11883	1
11886	-1,0
11891	-1
11906	1
11908	1
11920	-1
11928	1
11933	-1
11936	-1
11937	-1
11938	1
11944	-1
11950	1,0
11951	1,0
11957	1
11958	1
11966	1
11972	1
11975	-1,1
11977	-1
11980	1
11982	-1
11987	1
11991	1,0
11992	-1
12006	1
12007	1
12012	-1
12013	1
12014	-1
12015	1
12017	-1
12020	1
12023	-1
12028	-1
12032	1
12035	-1
12036	-1
12038	1
12039	-1
12040	-1

23961	1
23962	1
23963	-1
23970	1
23971	1
23972	1
23983	1
23986	-1
23988	-1
23989	1
23993	-1
23999	-1
24000	1
24004	-1
24017	1
24018	-1
24044	-1
24045	1
24046	-1,0
24050	1
24051	1
24053	-1
24055	1
24056	-1
24058	-1
24059	-1
24060	-1
24061	-1
24064	-1
24069	1
24075	-1
24086	1
24087	1
24109	1
24115	1
24116	1
24127	-1
24131	-1
26358	1
26359	1
26362	-1
26363	-1
26364	1
26379	1
26380	1
26385	-1
26395	-1
26395	-1
26399	1
26401	-1
26401	-1
26403	1
26404	1
26406	1
26409	-1
26411	-1
26413	1
26414	1
26417	-1
26419	-1
26422	1
26425	1
26430	-1
26431	1
26432	-1
26433	1
26440	-1
26442	1
26446	1
26448	1
26457	1

97243	1
97484	1
97487	1
97541	-1
97863	-1
97884	-1
99439	1,0
99470	1
99526	1
99571	1
99586	1
99633	1
99683	1
99696	1
99709	1
99712	1
99738	1
99889	1
99929	1
100019	-1
100102	-1
100121	-1
100182	-1
100213	-1
100434	-1
100561	1
100609	1
100637	1
100678	1
100683	1
100705	1
100710	1
100715	-1
100732	1
100737	1
100756	1
100763	1
100910	1
100972	1
100978	1
100986	1
101437	-1
101471	-1
101488	-1
101490	-1
101497	-1
101540	-1
101568	-1
101592	-1
101613	-1
101631	-1
101700	-1
101744	-1
101772	-1
101809	-1
102093	1
102182	1
102339	1
102414	-1
102423	-1
102436	-1
102448	-1
102502	-1
102545	-1
102570	-1
102607	-1
102626	-1
102680	-1
102774	-1
102791	-1
103537	1

12041	-1
12053	1
12054	1
12062	1
12070	-1
12091	-1
12116	-1
12123	1
12124	1,0
12143	-1
12144	-1
12145	-1
12153	-1
12155	-1
12159	-1
12160	-1
12161	-1
12166	-1
12167	1
12169	-1
12173	-1
12175	-1
12176	-1
12177	-1,1
12180	1
12182	1
12189	1
12190	1
12192	1,0
12209	-1
12212	-1
12224	-1
12229	-1
12234	1
12257	1
12265	1
12274	1
12279	1
12282	-1
12286	1
12288	1,0
12289	-1
12291	1
12292	-1,0
12293	1
12294	-1
12295	1
12297	1
12300	1
12306	-1
12307	-1
12317	1
12319	-1
12325	-1
12332	1
12333	1
12337	-1
12338	-1
12339	-1
12340	1
12348	-1
12349	1
12350	1
12351	1
12354	1
12361	-1
12363	-1
12364	-1
12368	1
12369	1
12373	1

26459	-1
26462	1
26465	-1
26557	-1
26559	1
26563	-1
26564	-1
26565	1
26570	1
26757	-1
26875	1
26876	1
26879	1
26885	-1
26887	1
26896	-1
26905	-1
26909	-1,0
26912	1
26914	-1
26918	-1
26926	-1
26931	1
26932	1
26939	-1
26942	1
26944	-1
26946	-1
26951	-1
26992	1
27008	1
27015	-1
27027	-1,1
27029	1
27050	-1
27052	-1
27057	-1
27060	1
27206	-1
27225	1
27260	1
27267	-1
27277	1
27281	1
27356	1
27357	1
27359	-1
27360	1
27364	1
27371	1
27373	1
27376	1,0
27384	-1
27388	-1,0
27399	-1
27401	1
27404	1
27405	1
27406	1
27414	-1
27416	1
27419	1
27421	-1
27426	1
27681	1
27801	1
27965	-1
27979	1
28000	1
28018	-1
28030	1

103677	1
103743	1
104099	-1
104110	-1
104112	1
104156	1
104158	1
104174	1
104175	-1
104184	1
104245	-1
104263	1
104318	1
104382	1
104383	1
104384	-1
104394	1
104416	-1
104443	1
104445	1
104759	1
104776	1
104816	1
104831	-1
104859	1
104884	1
104910	1
105148	-1
105245	-1
105355	-1
105377	-1
105428	-1
105445	-1
105446	-1
105501	-1
105504	-1
105522	-1
105590	-1
105663	-1
105670	-1
105675	-1
105689	-1,0
105727	1
105785	1
105787	1
105837	1
105853	1
105988	1
106039	1
106200	1
106298	1
106344	1
106347	1
106369	1
106389	1
106407	1
106957	1,0
107141	1
107173	1
107182	1
107221	1
107260	1
107272	1
107321	1
107328	1
107338	1
107351	1
107358	1
107368	1
107371	1
107375	1

12374	1
12380	-1
12387	-1
12390	-1
12394	1
12395	-1
12400	1
12402	-1
12406	-1
12408	1
12412	1
12416	1
12418	1
12419	1
12424	-1
12425	1
12426	-1
12442	-1
12447	-1
12453	1
12454	1
12455	1
12457	1
12460	1
12462	1
12465	1
12467	1
12476	-1
12478	-1
12481	1
12483	-1
12491	1
12492	1
12494	1
12495	1
12499	1
12505	-1,0
12507	1
12508	1
12511	1
12517	-1
12518	-1
12519	1
12520	-1
12522	-1
12524	1
12526	1
12530	-1
12539	-1
12545	1
12550	1
12554	1
12556	1
12557	-1
12560	1,0
12562	1
12563	1
12565	1
12568	1
12571	1
12578	-1
12585	-1
12591	1
12592	-1
12593	-1,1
12614	1
12630	1
12647	-1
12649	-1
12652	1
12655	1

28035	1
28042	1
28080	1
28114	-1
28135	-1
28200	-1
28248	-1
28250	-1
28253	-1
28254	-1
29806	-1
29810	-1
29811	-1
29815	1
29817	1
29820	-1,0
29849	1
29857	1
29858	1
29859	1
29861	-1
29864	-1
29867	1
29871	-1
29875	-1
30046	-1
30049	1
30054	-1
30806	-1
30838	1
30841	1
30926	-1
30942	1
30945	1
30946	-1
30947	1
30949	-1
30952	-1
30959	-1
30962	-1
50490	-1
50505	1
50523	-1
50524	-1
50527	-1
50528	1
50706	1
50720	-1
50759	-1
50760	-1
50765	1
50769	-1
50772	-1,1
50776	1
50779	1
50780	-1
50786	-1
50787	-1,0
50788	1,0
50789	-1
50790	-1,0
50791	1
50794	-1
50796	1
50797	-1
50798	-1
50868	-1
50877	-1
50883	1
50887	-1
50905	-1

107448	-1
107476	1
107503	-1
107522	1
107568	-1
107589	1
107650	1
107656	1
107684	-1
107702	-1
107751	-1
107753	1
107767	-1
107815	-1
107817	1
107823	1
107831	1
107868	1
107869	1
107889	-1
107934	-1
107975	1
107976	1
107993	-1
108012	-1
108015	-1
108058	1
108067	-1
108069	1
108071	-1
108075	-1
108079	-1,1
108083	1
108086	1
108096	-1
108097	1
108098	-1
108099	1
108100	1
108101	1
108105	1
108116	-1
108124	-1
108138	-1
108148	1
108153	-1
108154	-1
108155	-1
108156	1
108672	-1
108682	1
108723	1
108737	-1
108760	1
108837	-1
108902	1
108903	1
108927	1
108989	1,0
109052	1
109075	1
109093	-1,0
109113	-1,1
109135	-1
109136	1
109151	1
109263	-1
109264	-1
109270	1
109305	1
109331	-1

12660	1
12662	-1
12669	1
12671	-1
12677	1
12684	-1
12704	1
12709	1
12715	-1
12722	1
12724	1
12725	1
12727	-1
12728	-1
12729	-1
12737	1,0
12738	-1
12745	1
12748	1,0
12752	-1
12753	1
12757	-1
12759	-1,0
12763	-1
12764	-1
12769	-1
12771	-1
12773	-1
12775	1
12780	1
12788	1
12801	-1
12804	-1
12807	1
12814	1
12816	-1
12818	1
12819	-1
12821	1,0
12824	1
12830	-1
12836	-1
12840	-1
12842	1
12845	1
12846	1
12856	1
12858	-1
12864	1
12869	-1
12873	1
12874	1
12876	1
12889	1
12891	-1
12896	-1
12904	1
12909	1
12914	1
12919	-1
12921	1
12927	1
12931	1
12933	1
12934	-1
12961	1
12962	1
12970	1
12972	1
12985	1
12988	-1

50907	1
50911	1
50916	-1
50931	1
50933	-1
50934	-1
50995	-1
50997	1,0
51789	1
51795	-1
51797	-1
51798	-1
51799	1
51813	-1
51886	1
52024	1
52028	1
52033	-1
52064	1
52118	-1
52150	-1
52187	-1
52206	1
52335	1
52389	-1
52397	1
52398	1
52432	-1
52585	-1
52615	1
52686	1
52793	1
52830	1
52897	1
53310	-1
53313	1
53314	1,0
53319	1
53320	-1
53321	1
53322	-1
53323	1
53333	-1
53334	1
53357	1
53378	-1
53380	-1
53381	-1
53412	1
53413	1
53417	-1
53419	1
53420	-1
53611	1
53612	1
53614	-1
53618	1
53622	1,0
53623	-1
53627	-1
53859	1
53860	1
53861	1
53867	-1
53868	1
53880	-1
53883	1
53892	1
53970	1
54004	-1
54006	-1

109333	1,0
109552	1
109594	-1
109620	-1
109637	-1
109652	-1
109674	1
109689	-1
109700	-1
109711	1
109731	-1
109754	1
109801	-1
109857	1
109900	1,0
109904	-1
109905	1
109910	1
109929	1
109934	1
109979	1
110006	1
110033	-1
110052	-1
110075	1
110082	1
110094	-1
110095	1
110115	1
110119	-1
110135	1
110173	1
110197	1
110213	1
110265	-1
110308	1
110310	1
110323	-1
110350	-1
110355	1
110380	-1
110382	-1
110385	1
110391	1
110446	-1
110521	-1
110524	1
110596	-1
110606	1
110616	1
110637	-1
110639	-1
110651	-1
110749	1
110784	1
110789	-1
110821	-1
110826	-1
110834	-1
110835	-1
110842	-1,0
110855	1
110862	1
110877	1
110880	1
110886	-1
110891	-1
110902	-1
110948	1
112405	1
112406	-1,1

12995	1,0
13000	1
13003	-1
13009	-1
13016	1
13018	1
13026	1
13030	-1
13033	-1
13034	-1,0
13036	-1
13040	1
13043	-1
13047	1
13048	-1,1
13051	-1
13056	1
13058	-1
13072	1
13075	-1
13076	-1
13088	-1
13089	-1
13090	-1
13095	1
13096	1
13097	1
13098	1
13101	1
13107	-1
13108	-1
13110	-1
13113	1
13114	1
13116	1
13117	-1
13120	-1
13123	1
13131	-1
13132	1
13134	-1
13135	-1
13138	-1
13142	1
13143	-1
13162	-1
13171	1
13172	-1
13178	1
13190	-1
13194	1
13205	-1
13206	-1
13209	-1
13350	1
13353	1
13356	1
13358	1
13360	-1
13361	-1
13363	1
13370	-1
13383	1
13383	1
13384	1
13386	1
13389	-1
13393	1
13394	1
13400	-1
13401	-1

54120	1
54122	-1
54124	1
54128	1
54130	1
54135	-1
54140	1
54152	1
54153	1
54156	-1
54169	-1
54195	1
54196	-1
54200	-1
54201	1
54214	-1
54215	1
54326	-1
54343	-1
54352	1
54357	-1
54366	-1
54367	1
54369	-1
54375	1
54376	-1
54377	1
54378	-1
54381	1
54391	1
54393	-1,0
54403	1
54446	1
54447	1
54473	-1
54524	1
54525	1
54526	1
54608	-1
54611	-1
54614	1
54616	-1
54631	-1
54632	-1
54633	-1
54645	-1
54652	-1
54672	1
54673	1
54683	1
54698	-1
54720	1
55925	-1
55936	-1
55944	1
55946	-1
55947	1
55960	1
55980	1
55981	-1
55982	1
55984	1
55988	-1
55991	-1
55992	-1
55993	1
55994	1
56032	-1
56041	1
56043	-1
56044	-1

112419	1
113868	-1
114230	1
114255	1
114304	-1
114479	1
114642	1
114643	1
114664	1
114671	-1
114674	1
114713	-1
114714	1
114741	-1
114873	-1
114874	-1
114875	-1
114886	1
114893	1
116701	1
116810	1
116852	-1
116870	1
116873	1
116904	-1
116939	1
116940	-1
116972	1
117109	1
117146	1
117147	-1
117149	-1
117160	-1,1
117589	-1
117590	1
117591	1
118446	1
118453	1
140475	-1
140477	-1
140481	-1
140493	1,0
140498	1
140557	1
140577	-1
140580	-1
140630	-1
140723	1
140780	1
140781	-1
140806	-1
140858	-1,0
140859	1
140887	1
140919	-1
142688	-1
170439	1
170472	1
170574	1
170625	-1
170644	1
170722	-1
170733	-1
170737	1
170743	-1
170744	-1
170745	-1
170749	1
170750	1
170752	-1
170757	1

13404	1
13405	-1
13409	1
13411	1
13417	-1
13421	-1
13424	1
13430	-1
13433	-1,0
13446	-1
13447	1
13449	-1
13478	-1
13483	1
13486	1
13489	-1
13497	-1
13522	-1
13531	-1
13544	1
13548	1
13549	1
13557	-1
13595	-1
13601	1
13602	1
13605	1
13607	-1
13610	-1
13612	-1
13614	-1
13631	-1
13636	1,0
13639	1
13641	-1
13644	-1
13645	1
13663	-1
13665	1
13666	1
13669	1
13682	1
13684	1
13709	-1,1
13712	-1
13713	0,1,-1
13716	1
13717	1
13722	1
13728	1
13797	1
13799	1
13803	-1
13809	1
13813	-1
13841	-1
13845	1
13846	1
13850	-1
13852	1
13856	1
13858	-1
13859	1
13860	-1
13866	1
13870	-1
13871	-1
13874	1
13875	-1
13876	1
13877	1

56048	-1
56068	-1
56077	1
56078	-1
56085	-1
56092	-1
56094	-1
56149	1
56173	-1,1
56175	1
56183	1
56189	-1
56196	-1
56209	-1
56212	-1
56215	-1
56216	-1
56217	1
56228	-1
56233	1
56274	1
56275	1
56278	-1
56280	-1
56289	-1
56291	-1
56294	-1
56295	-1
56305	1
56309	1
56312	-1
56314	1
56315	-1
56316	1
56317	1
56318	-1
56321	1
56322	1
56325	1
56330	-1
56335	-1
56347	-1
56349	-1
56350	1
56356	1
56357	2
56358	1
56373	-1
56375	1
56378	1,0
56380	-1
56382	-1
56388	1
56389	1,0
56401	-1
56403	-1
56404	-1
56405	1
56417	1
56419	-1
56420	-1
56421	-1
56426	1
56427	1
56430	1
56433	1
56438	1
56440	-1
56445	1
56447	1
56448	1

170762	1
170767	1
170770	-1
170790	1
170822	1
170828	-1
170930	-1
171166	1
171180	1
171196	-1
171200	1
171202	1
171209	1
171210	1
171211	-1
171238	-1
171261	-1
171286	1
171429	-1,0
171504	-1
191578	1
192119	1
192157	1
192159	-1,1
192167	1
192173	1
192176	-1,0
192195	1
192232	1
192236	1
192287	-1
192654	1
192656	-1
192657	-1
192663	-1
192775	1
192897	1
192970	1
192986	1
193034	1
193385	-1
194590	-1
194744	-1
195018	1
195646	-1
195727	-1
207151	1
207165	1
207181	-1
207212	-1
207214	1
207215	1
207352	-1
207425	-1
207521	1
207565	1
207596	-1
207728	-1
207742	1
207777	1
207785	1
207839	1
207965	1
208080	-1
208084	-1
208092	1
208104	1
208117	-1
208144	1
208146	1
208188	1

13884	1,0
13885	-1
13897	1
13909	1
13983	1
13984	-1
14000	1
14011	-1
14026	1
14027	1
14042	1
14055	1
14057	-1
14062	-1
14066	1
14071	-1
14073	-1
14083	1
14084	-1
14085	-1
14086	1
14088	-1
14102	1,0
14104	1
14114	1
14129	1
14132	-1
14137	-1
14148	-1
14149	1
14155	-1
14159	-1
14161	1
14163	-1
14165	-1
14167	1
14168	-1
14169	-1
14173	1
14174	-1
14175	-1
14176	1
14180	-1
14183	-1
14186	-1
14198	-1
14204	-1
14205	-1
14208	1
14211	-1
14218	1
14230	1
14232	1
14238	-1
14247	-1
14254	1
14255	1
14263	1
14265	-1
14266	-1
14275	-1
14276	-1
14282	-1
14283	1
14297	1
14313	-1
14314	1
14339	1
14344	-1
14348	-1
14357	1

56449	-1
56451	1
56452	1
56453	1
56454	1,0
56456	1
56458	1
56459	-1
56469	-1
56473	1
56494	1
56495	1,0
56501	-1,0
56503	-1
56506	-1
56527	1
56529	-1
56532	-1
56542	-1
56543	1
56546	-1
56551	1
56612	1
56613	1
56615	-1
56626	1
56637	1
56642	1
56643	-1
56696	1
56703	-1
56705	-1
56706	1
56708	1
56710	-1
56711	-1,0
56715	1
56722	1
56727	1
56734	-1
56739	-1
56747	1
56749	1
56752	-1,0
56773	1
56774	-1
56788	-1
56790	1
56794	-1
56805	-1
56808	-1
56839	1
56843	-1
56847	-1
56857	-1
56873	1
56874	1
57014	-1
57080	1
57230	1
57246	-1
57257	1
57258	-1
57259	1
57265	1
57266	-1
57278	-1
57279	-1
57321	1
57342	-1
57344	1

208194	-1
208211	1
208440	-1
208449	1
208501	1
208583	-1
208595	1
208606	1
208628	1
208643	1
208666	-1,0
208691	1
208846	1
208869	-1
208890	-1
208898	-1
208936	1
208943	-1
208982	-1
208994	-1
209005	-1
209018	1
209039	1
209091	-1
209186	-1
209195	1
209224	-1
209239	1
209318	1
209354	1
209357	1
209446	-1
209588	1
209707	1
209737	-1
209776	-1
210004	1
210009	-1
210044	-1
210106	-1
210108	-1
210126	1
210135	-1
210162	-1
210274	-1
210297	-1
210376	-1
210529	1
210530	1
210544	1
210992	-1
211064	1
211134	1
211147	1
211151	1
211208	-1
211253	-1
211286	-1
211480	-1
211535	-1
211548	-1
211578	-1
211586	-1
211612	-1
211666	1
211712	-1,0
211770	1
211949	-1
211978	1
212032	-1
212111	-1

14359	1
14362	1
14365	-1
14368	1
14377	1
14387	1
14388	1
14389	-1
14402	-1
14407	-1
14408	-1
14409	-1
14422	1
14429	1
14450	1
14453	-1
14460	-1
14463	-1
14468	1
14469	1
14473	1
14527	1
14528	-1
14534	1
14537	1
14538	-1
14544	1
14560	-1
14569	-1
14573	1
14580	1
14581	1
14585	1
14586	-1
14595	-1
14600	1
14605	-1
14611	-1
14613	1
14615	1,0
14618	-1
14623	-1
14629	-1
14630	1
14634	-1
14635	1
14652	-1
14658	1
14661	-1
14664	-1
14673	1
14674	1,0
14675	1
14677	1
14678	-1
14679	1
14681	1
14682	1
14686	1
14696	1
14697	-1
14702	-1
14706	-1
14709	-1
14710	1,0
14712	1
14718	1
14719	1
14728	1,0
14731	1
14732	1

57357	1
57376	1
57377	2
57385	-1
57423	1
57430	1,0
57434	1
57441	-1
57442	1
57444	-1
57740	-1
57742	1,0
57746	-1
57748	-1
57749	1
57765	1
57775	-1
57776	-1
57810	-1
57811	-1
57813	1
57837	1
57912	1
57914	1,0
57916	1,0
58170	1
58176	1
58178	1
58179	-1
58193	1
58194	-1
58200	1
58206	-1
58210	1
58222	1
58233	-1
58234	1
58235	-1
58242	-1
58245	-1
58249	1
58522	1
58805	1
58807	-1
58810	-1
58861	-1
58865	-1
58988	-1,1
58992	-1
58994	1
59001	-1
59007	-1
59008	1
59009	1
59015	-1,0
59020	1,0
59021	-1
59024	-1
59026	-1
59031	1
59033	1
59035	-1
59036	1
59040	1
59042	1
59044	1,0
59044	1,0
59045	1
59046	-1
59047	-1
59047	-1

212167	1
212285	1
212390	-1
212391	1
212503	-1
212528	1
212880	-1
212919	1
212974	-1
212986	1
212989	1
212996	1
212999	1
213011	-1
213053	-1
213054	1
213119	1
213121	1
213208	-1
213211	-1
213311	-1
213409	-1,0
213435	1
213438	-1
213439	-1
213498	1
213582	1
213948	1
213990	1
214084	1
214111	-1
214137	1
214162	-1
214254	-1
214292	1
214425	-1
214444	-1
214523	-1
214531	-1
214572	1
214579	-1
214639	-1
214663	1
214669	1
214766	-1
214897	-1
215061	1
215113	1
215114	1
215201	-1
215351	-1
215384	-1
215418	-1
215446	-1
215476	1
215512	1
215654	1
215707	1
216635	-1
216964	1
216965	1
216974	1
217026	1
217030	1
217031	1
217057	1
217069	1
217071	1
217109	1
217124	1
217127	1

14734	-1
14735	-1
14739	-1
14745	-1
14747	1
14751	-1
14756	-1
14760	-1
14764	1
14766	1
14772	1
14773	1
14784	1
14786	1
14800	1
14802	-1
14803	-1
14805	1
14809	-1
14811	1
14812	-1
14813	1
14814	-1
14824	1
14829	-1
14836	1
14843	1
14852	1
14858	-1,1
14859	-1,0
14860	-1
14863	1
14869	1
14870	1
14874	1
14884	-1
14886	1
14904	1
14917	-1
14918	-1
14933	-1
14936	-1
14938	-1
14945	-1
15107	1
15114	1
15117	1
15163	1
15165	-1
15168	1
15184	1
15185	-1
15191	1
15194	1
15201	1
15203	-1
15204	-1
15209	-1
15211	-1
15212	-1
15216	1
15218	1
15220	-1
15221	-1
15233	1,0
15234	1
15247	1
15251	1
15257	1
15270	-1
15273	-1,0

59048	-1
59049	-1
59056	1
59079	-1
59092	-1
59095	-1
59126	-1,0
59289	-1
60322	-1
60367	-1
60504	-1
60505	1
60510	-1
60527	1
60534	-1
60595	-1
60596	1
60599	-1
60613	-1
63828	1
63857	1
63859	-1
63873	1
64008	-1
64011	-1
64051	1
64113	1
64176	-1
64209	1
64293	1
64296	1
64297	-1
64340	1
64381	1
64383	-1
64384	-1
64424	-1
64580	1
64602	-1
64652	-1
64661	-1
64705	1
64706	1
64833	-1
64918	-1,0
64929	-1
64945	1
65079	1
65086	1
65098	-1
65107	-1
65114	1
65221	1
65246	-1
65254	1
65972	1
65973	-1
66046	1
66053	1
66065	-1
66069	-1
66073	-1
66082	-1
66101	-1
66105	1,0
66120	1
66138	1
66143	-1
66152	1
66156	1
66165	-1

217143	1
217166	1
217198	1
217207	1
217217	1
217232	1
217258	1
217262	1
217265	1
217302	1
217304	1
217316	1
217331	1
217333	1
217342	1
217344	1
217351	1
217353	1
217356	1
217364	1
217369	-1
217666	1
217716	1
217718	1
217721	1
217734	1,0
217826	1
217837	1
217843	1
217864	1
217866	1
217882	1
217944	1
217980	-1
218030	-1
218035	-1
218103	-1
218138	-1
218194	-1
218203	-1
218210	-1,0
218214	-1
218215	-1
218232	-1
218236	-1
218268	-1
218271	-1
218294	-1
218343	-1
218397	-1
218440	-1
218441	-1
218442	-1
218454	-1
218461	-1
218490	-1
218624	-1
218629	-1
218630	-1
218699	-1
218756	-1
218772	-1
218793	-1
218811	-1
218832	-1
218850	-1
218865	-1
218877	-1
218885	-1
218914	-1
218977	-1

15285	1
15288	-1
15357	-1
15360	1
15368	1
15369	1
15371	1
15377	-1
15408	1
15410	1
15416	1
15417	1
15422	1
15424	1
15425	1
15427	1
15439	1
15442	1
15450	-1
15451	-1
15452	-1
15458	-1
15460	-1
15461	-1
15464	-1
15466	-1
15467	1
15469	-1
15476	1
15486	1,0
15487	-1
15493	1
15494	1
15496	1
15499	1
15505	1
15519	1
15558	-1
15560	-1
15561	-1
15563	1
15566	1
15567	1
15569	-1
15571	-1
15572	-1
15894	-1
15898	-1
15904	-1
15925	1
15931	-1
15932	1
15957	1
15958	1
15959	1
15972	-1
15976	1
15980	1
15985	1
16000	-1,0
16001	-2
16002	-1
16007	1
16012	1
16154	-1
16155	1
16156	-1
16157	-1
16158	-1
16161	1
16164	-1

66171	1
66177	-1,0
66179	1
66185	-1
66190	-1
66194	1
66200	-1
66204	1
66209	-1
66212	-1
66214	-1
66222	-1
66233	-1
66234	1
66235	-1
66245	-1
66246	-1
66251	1
66313	1
66326	1,0
66333	-1
66335	1
66350	1
66354	1
66355	-1
66357	1
66362	-1
66368	1
66369	1
66401	-1
66408	-1
66413	-1
66422	-1
66427	1
66435	-1
66446	-1
66475	1,0
66482	-1
66498	1
66500	1
66505	-1
66513	1
66514	1
66522	1
66525	-1
66531	1
66549	-1
66556	1
66569	1
66583	1
66587	1
66588	-1
66590	1
66596	1
66622	1
66656	1
66661	1
66671	-1
66673	1
66681	1
66694	-1
66711	1
66772	1,0
66812	-1
66813	-1
66824	-1
66830	1
66834	-1
66841	1
66849	1
66853	-1

219024	-1
219103	-1
219114	-1
219135	-1
219150	-1
219151	-1
219228	-1
223254	-1,1
223255	-1
223272	-1
223435	1
223453	1
223455	1
223604	1
223649	1
223664	1
223669	1
223690	1
223693	1
223701	1
223722	1
223739	1
223753	1
223770	1
223774	1
223776	1
223843	1
223870	1
223881	1
223915	1
223922	1,0
223978	1
223989	1
224020	1
224079	1
224088	1
224105	1
224129	1
224132	1
224432	1
224624	1,0
225579	1,0
225849	1
225861	1
225865	1
225872	1
225876	1
225888	1
225922	1
225994	1
225997	1
225997	1
225998	1
226016	1
226025	1
226041	1
226075	1
226090	1
226098	1
226101	1
226105	1
226143	1
226151	1
226153	1
226169	1
226243	1
226251	1
226255	1
226265	1
227099	1,0
228576	1,0

16165	-1
16168	1
16170	-1
16173	-1
16178	-1,0
16180	1
16183	1
16188	-1
16194	1
16195	-1
16196	1
16197	1
16198	1
16201	-1
16202	-1
16211	1
16323	-1
16332	-1
16336	1
16341	1
16348	-1
16364	-1
16371	-1
16373	1
16391	-1
16392	-1
16398	-1
16399	1
16400	1
16402	1
16403	-1,0
16407	1
16408	-1
16409	-1
16411	-1
16412	1
16416	1
16419	1
16421	1
16430	-1
16432	-1
16439	-1
16451	-1,0
16452	1
16453	1,0
16468	-1
16480	1
16490	1
16497	1
16502	-1
16504	-1
16509	1
16511	1
16512	1
16516	1
16517	1
16520	1
16521	-1
16522	1
16523	-1
16525	1
16527	1
16529	-1
16531	-1
16535	-1
16542	1
16551	1,0
16552	-1
16553	-1
16563	-1
16571	-1

66854	-1
66863	1
66866	1
66867	-1
66884	1
66887	1
66889	-1
66890	-1
66894	1
66897	-1
66902	-1
66904	-1
66913	1
66914	1,0
66916	1
66922	-1
66925	-1
66930	-1
66939	-1
66945	-1
66948	-1
66949	1
66970	-1
66988	1
66993	1
67011	-1
67013	-1
67016	-1
67041	1
67042	1
67053	-1
67062	-1
67071	-1
67073	1
67075	-1
67085	-1
67089	-1
67103	-1
67109	-1
67111	1
67118	1
67123	-1
67128	1
67144	1
67150	-1
67151	1
67153	-1
67155	1
67161	1
67164	-1
67184	1
67201	1,0
67203	1
67213	-1
67235	1
67263	-1,0
67269	-1
67283	1
67285	-1
67287	-1
67295	-1
67296	-1
67298	-1
67300	1
67305	-1
67317	-1
67323	1
67338	1
67375	-1
67378	1
67381	-1

229211	1
229214	1
229285	1
229302	1
229320	1
229357	1
229363	1
229445	1
229474	1
229499	1
229504	1
229512	1
229521	1
229534	1
229542	1
229574	1
229584	1
229588	1
229589	1
229603	1
229603	1
229644	1
229663	1
229665	1
229681	1
229697	1
229699	1
229706	1
229709	1
229715	1
229722	1
229725	1
229776	1
229780	1
229782	1
229791	1
229841	1
229877	1
229898	1
229900	1
229905	1
229906	1
229937	1
229949	1
230025	-1
230027	-1
230073	-1
230099	-1
230101	-1
230103	-1
230125	-1
230145	-1
230162	-1
230233	-1
230393	-1
230459	-1
230558	-1
230582	-1
230587	-1
230594	-1
230597	-1
230598	-1
230648	-1
230661	-1
230673	-1
230674	-1
230700	-1
230709	-1
231042	1
231044	1
231050	1

16573	1,0
16576	-1
16578	-1
16581	1
16582	1
16590	1
16591	1
16593	1
16594	1
16596	1,0
16597	-1
16598	1
16599	1
16600	-1
16601	1
16612	-1
16636	-1
16637	-1
16638	-1
16642	-1
16644	1
16646	1
16648	-1
16649	1
16651	-1
16653	-1,0
16656	-1
16667	1
16669	1
16678	1
16679	-1,1
16681	1
16682	1
16687	1
16688	1
16691	1
16706	1
16709	-1
16779	-1
16784	-1
16785	-1
16790	-1
16795	1
16796	1
16800	1
16801	-1
16825	1
16826	1
16828	-1
16832	-1
16833	-1
16834	1
16835	-1
16842	1
16847	-1
16848	1
16854	-1
16855	-1
16859	1
16869	1
16873	1
16875	1
16881	-1
16882	1
16885	1
16889	1
16890	-1
16905	1
16924	1
16939	1
16949	-1

67397	1
67398	-1
67402	-1
67414	1
67417	-1
67420	-1
67437	1
67441	-1
67442	1
67451	1
67460	-1
67464	-1
67469	-1
67474	1
67486	-1
67487	1
67488	1
67498	1
67511	-1
67528	1
67533	-1
67542	1
67547	1
67549	1
67553	1
67573	1
67579	1,0
67608	1
67610	1
67615	-1
67618	-1
67622	1
67623	-1
67657	1
67667	-1
67684	1
67689	1
67711	-1
67712	-1
67731	1
67733	-1
67738	1
67758	1
67768	1
67772	-1
67800	-1
67801	1
67834	-1
67845	1
67848	1
67866	1
67870	1,0
67873	1,0
67902	1
67903	1
67905	-1
67909	1
67916	-1
67943	-1,0
67952	1
67956	1
67963	1
67967	-1
67983	-1
67994	-1
68015	1
68018	-1,0
68021	-1
68038	-1
68039	-1
68043	-1

231051	1
231086	1
231103	-1,1
231128	1
231162	-1
231225	1
231252	1
231326	1
231329	1
231430	1
231464	1
231510	1
231549	1
231580	1,0
231600	1
231605	1
231630	1
231637	1
231655	1
231659	1
231670	1
231672	1
231769	1
231834	1
231842	1
231861	1
231863	1
231871	1
231876	1
231889	1
231903	1,0
231912	1
232078	1
232089	1
232415	-1
232430	-1
232431	-1
232441	-1
232449	-1
232493	-1
232533	-1
232791	-1
232798	-1
232807	-1
232811	-1
232827	-1
232878	-1
232889	-1
232906	-1
232910	-1
232933	-1
232943	-1
232944	-1
232970	-1
232975	-1
232989	-1
233001	-1
233011	-1
233016	-1
233020	-1
233038	-1
233040	-1
233046	-1
233060	-1
233071	-1
233080	-1
233107	-1
233208	-1
233210	-1,0
233221	-1
233230	-1

16956	1
16970	-1
16973	1
16974	-1
16975	-1
16976	1
16985	-1
16997	1
16998	1
17002	-1
17025	-1
17079	-1,1
17083	-1
17096	-1
17101	-1
17120	1
17121	-1
17122	1
17125	-1,1
17127	-1
17129	-1
17130	-1
17132	1
17133	1
17134	1
17135	1
17152	-1
17156	1
17159	1
17160	1
17161	-1
17165	1
17169	1,0
17174	1
17178	-1
17187	1,0
17217	1
17218	1
17220	1
17237	1
17250	1
17251	1
17252	1
17258	-1,0
17259	1
17260	-1
17261	1
17274	1
17279	-1
17285	1
17293	-1
17305	1
17309	1
17314	-1
17318	-1
17330	1
17345	-1
17346	-1
17349	1
17350	-1
17355	1
17364	-1
17380	1
17381	-1
17388	1
17389	-1
17390	1
17394	-1
17420	1
17425	1
17428	1

68048	-1,0
68055	1
68058	1
68059	-1
68070	1,0
68077	-1
68079	-1
68097	1
68119	1
68133	1
68134	-1
68135	1
68137	-1
68147	1
68149	1
68178	-1
68183	1
68194	1
68196	1
68203	-1
68229	1
68255	-1
68267	-1
68273	-1
68279	1
68291	-1
68292	-1
68312	1
68318	-1
68328	1
68332	-1
68346	-1
68401	1,0
68420	1
68431	1
68449	-1
68458	-1
68472	-1
68490	-1
68497	1
68505	1
68519	1
68525	1
68537	1
68553	-1,0
68558	-1
68563	1,0
68564	1
68566	-1
68572	1
68581	1
68603	1
68606	1
68607	1
68614	1
68616	-1
68644	-1
68646	1
68666	1
68667	-1
68682	-1
68705	-1
68708	1
68729	1
68732	-1
68742	-1
68743	-1
68750	-1
68753	1
68770	1
68801	-1

233274	-1
233328	-1
233332	-1
233405	-1
233406	-1
233410	-1
233437	-1
233489	-1
233529	-1
233532	-1
233537	-1
233571	-1
233575	-1
233649	-1
233726	-1
233733	-1
233781	-1
233789	-1
233801	-1
233805	-1
233824	-1
233826	-1
233833	-1
233863	-1
233865	-1
233870	-1
233878	-1
233879	1
233900	-1
233902	-1
233908	-1
233977	-1
233979	-1
234309	1
234311	1
234329	1
234356	1
234366	1
234374	1
234388	1
234404	1
234515	1
234582	1
234593	1
234594	1
234595	1
234663	1
234664	1
234673	1
234684	1
234695	1
234724	1
234730	1
234733	1
234734	1
234736	1
234740	1
234779	1,0
234788	1
234875	1
234878	1
234889	-1
234915	-1
234959	-1
234967	-1
235033	-1
235040	-1
235041	-1
235072	-1
235086	-1
235106	-1

17436	-1
17444	1
17448	1
17454	1
17463	1
17470	1
17476	1
17480	-1
17523	-1,0
17532	-1
17534	1
17535	-1
17537	-1
17540	-1
17686	-1
17692	-1
17698	-1
17751	1
17755	-1
17762	1
17765	1
17776	-1
17777	1
17827	1
17855	1
17857	1,0
17858	1,0
17859	1
17869	1
17874	-1
17880	1
17886	1
17888	-1
17906	1
17909	1
17918	-1
17920	-1
17921	-1
17925	1
17938	-1,0
17939	1
17940	-1
17951	-1
17952	-1
17961	1
17962	1
17966	1
17967	-1
17969	1
17973	-1
17977	1,0
17984	-1
17986	-1
17993	-1
17999	-1
18002	-1
18004	1
18007	-1
18011	1
18013	1
18015	1
18021	1
18022	1
18023	1
18027	-1
18028	-1
18030	-1
18032	1
18033	1
18036	-1
18039	-1

68813	-1
68815	-1
68837	1
68839	1
68854	-1
68861	-1
68889	-1
68922	-1
68926	-1
68927	-1
68938	1
68939	1
68947	-1
68961	-1
68966	-1
68995	-1
68999	1
69019	-1
69024	1
69029	1
69032	-1
69035	-1
69047	1
69048	-1
69060	1
69089	-1
69091	-1
69094	-1
69101	1
69104	1
69123	-1
69131	1
69137	-1
69147	-1
69150	1
69159	1
69168	1
69207	1
69219	1
69227	1
69237	-1
69257	1
69276	1
69299	-1
69305	-1
69315	-1
69354	1
69398	-1
69462	1
69480	1
69520	1
69528	1
69536	-1,0
69544	1
69562	-1
69574	1
69587	1
69602	1
69606	-1
69608	1
69617	-1
69627	1
69632	-1
69634	-1
69635	-1
69639	1
69656	-1
69663	1
69666	-1
69672	-1
69684	1

235130	-1
235134	-1
235180	-1
235281	-1
235320	-1
235323	-1
235339	-1
235344	-1,1
235380	-1
235431	-1
235435	-1
235441	-1
235442	-1
235461	-1
235497	-1
235504	-1
235527	-1
235534	-1
235574	-1
235584	-1
235604	-1
235606	-1
235611	-1,0
235623	-1
235626	-1
235631	-1
235674	-1
235682	-1
235973	-1
236193	1
236539	1
236690	-1
236727	-1
236733	-1
236749	-1
236781	-1
236794	-1
236798	-1
236848	-1
236899	-1
236900	-1
236915	-1
237038	-1
237175	-1
237178	-1
237213	-1
237222	-1
237831	1
237847	1
237859	1
237860	1
237868	1
237877	1
237886	1
237891	1
237898	1
237911	1
237926	1
237928	1
237930	1
237958	1
237988	1
238021	1
238023	1
238024	1
238130	1,0
238247	1
238266	1
238271	1
238323	1
238331	1

18046	-1
18049	1
18053	1
18073	-1
18074	-1
18081	-1
18087	-1
18099	1
18102	1
18106	-2
18107	1
18113	-1
18115	-1,0
18121	1
18124	-1
18125	1
18127	1
18129	1
18130	-1,1
18133	1
18142	-1
18160	1
18162	1,0
18163	1
18164	1
18167	1
18168	1
18171	1
18174	1
18185	-1
18186	1
18189	-1
18190	1
18193	-1
18195	1
18198	-1
18201	-1
18203	1
18211	1
18212	-1
18213	-1
18218	-1
18222	1
18230	1
18241	-1,0
18256	1
18260	-1
18263	1
18300	-1
18301	-1,1
18312	-1
18315	-1
18317	1
18359	1
18383	1
18386	-1,0
18399	1
18400	-1
18412	1
18414	1
18415	1
18416	-1
18417	1
18419	-1
18420	-1
18424	-1
18426	1
18431	-1
18436	1
18438	1
18439	1

69706	1
69714	-1
69716	-1
69719	1
69721	-1
69724	1
69727	1
69740	1
69745	1
69748	-1
69757	-1
69774	1
69787	1,0
69790	1
69792	1
69806	1
69833	1
69863	-1
69865	1
69888	1
69890	-1,0
69906	1
69908	-1
69930	-1
69940	1
69955	-1
69983	1
70008	-1
70021	-1
70028	1
70031	-1
70052	-1
70059	1
70080	-1
70086	-1
70093	1
70099	1
70103	1
70110	1
70118	1
70120	1
70127	1
70144	1
70178	-1
70202	-1
70225	-1
70235	-1
70238	1
70300	-1
70315	-1
70316	-1
70348	-1
70350	1
70355	1
70356	1
70358	1
70359	1
70363	1
70381	1
70382	1
70394	-1
70415	-1
70426	1
70430	-1
70439	1
70450	1
70461	-1
70472	1
70478	-1
70484	-1
70495	-1,0

238377	1
238384	1
238405	1
238406	1
238463	-1
238505	-1
238673	-1
238680	-1
238683	-1
238690	-1
238693	-1
238722	-1
238799	-1
238803	-1
238829	-1
238831	-1
238871	-1
238944	-1
238988	-1
239083	-1
239122	-1
239134	-1
239273	-1
239283	-1
239337	1
239393	1
239405	1
239420	1
239436	1
239510	1
239528	1,0
239530	1
239556	1
239559	1
239570	1
239591	1
239659	1
239691	1
239706	1
239719	1
239731	1
239759	1
239789	1
239827	1
239833	1
239845	1
240549	1
240590	1
240595	1
240641	1
240665	1
240667	1
240672	1
242037	1
242083	1
242202	1
242259	1
242274	1
242286	-1
242291	-1
242341	-1
242362	-1
242409	-1
242425	-1
242443	-1
242505	-1
242546	-1
242557	-1
242585	-1
242607	-1
242700	-1,0

18440	1
18442	-1
18453	1
18458	1
18472	1
18477	-1
18479	-1
18481	-1
18491	-1
18504	1
18507	-1
18518	-1
18521	1
18548	-1
18550	-1
18554	-1
18555	-1
18563	1
18569	1
18570	-1
18571	-1
18574	-1,1
18575	1,0
18576	-1
18577	-1
18578	-1
18583	1
18584	-1
18587	1
18588	1
18590	1
18591	1
18595	1
18597	-1
18604	1
18606	1,0
18612	1
18613	1
18616	-1
18619	-1
18624	-1
18631	-1
18636	-1
18639	-1
18639	-1
18642	1
18645	1
18647	1
18648	1
18654	1
18655	-1
18667	-1
18669	1
18670	1
18671	1
18673	1
18675	-1
18676	-1
18679	-1
18682	1
18685	1
18693	1
18700	-1
18701	-1,0
18704	-1
18705	-1
18706	1
18708	-1
18709	1
18710	-1
18715	-1

70551	-1
70560	1
70561	-1
70567	1
70568	-1
70573	1
70575	1
70584	-1
70591	1
70598	-1
70601	-1
70620	1
70625	1
70646	-1
70661	-1
70673	1,0
70686	-1
70696	-1
70701	1
70737	1
70757	1
70762	1
70771	-1
70784	-1
70790	1
70797	1
70823	1
70834	1
70893	-1
70894	1
70930	-1
70967	1
70974	-1
70993	1
70998	-1
71007	-1
71037	-1
71041	1
71046	1
71059	1
71078	1
71091	1
71093	1
71101	1
71116	1
71146	1
71148	-1
71149	1,0
71164	-1
71200	-1
71213	-1
71240	1
71241	-1
71268	-1
71275	1
71276	1
71310	1
71330	-1
71340	-1
71351	1
71354	-1
71355	1
71375	1
71382	-1,1
71393	-1
71412	-1
71458	-1
71472	-1
71492	1
71519	1
71521	1

242726	-1,0
242851	1
242860	1
242864	1
242894	1
242939	1
242960	1
243168	1
243272	1
243277	1
243312	1
243537	1,0
243659	-1
243755	-1,0
243842	-1
243853	-1
243864	-1
243867	-1
243874	-1
243881	-1
243912	-1
243931	-1
243967	-1
243979	-1,0
243983	-1
244049	-1
244059	-1
244141	-1
244152	-1
244202	-1
244238	-1
244550	1
244556	1
244562	1
244595	1
244631	1
244646	1
244650	1
244666	1
244667	1
244668	1
244701	-1
244879	-1
244882	-1
244895	-1
244962	-1
245000	-1
245007	-1
245038	-1
245049	-1
245509	-1
245526	-1
245527	-1
245537	-1
245555	-1
245572	-1
245578	-1
245596	-1
245610	-1
245615	-1
245616	-1
245638	-1
245650	-1
245666	-1
245671	-1
245688	-1
245841	1
245857	1
245880	1
245902	-1
246103	-1

18719	1
18720	1
18724	-1,0
18738	1
18739	1
18741	1
18746	-1
18747	1
18749	1
18750	1
18751	-1
18752	-1
18753	-1
18755	1
18759	1
18764	1
18766	1
18770	1
18787	1
18789	1
18793	-1
18797	1
18799	-1
18805	1
18807	-1
18810	1
18817	-1
18823	-1
18828	-1
18854	-1
18935	1
18938	1
18946	1
18947	1
18950	-1
18952	1
18968	-1
18969	1
18971	-1
18973	1
18975	-1
18985	-1
18987	-1
18988	-1
19011	1
19012	-1
19013	1
19014	1
19017	1
19023	1
19024	-1
19042	1
19046	1
19047	1
19055	1
19056	-1
19057	-1
19060	-1
19062	1
19065	-1
19079	1
19084	1
19085	1
19087	-1
19089	-1
19090	1
19091	1
19092	1
19094	1
19108	-1
19109	-1

71522	1
71538	-1
71544	-1
71562	1
71591	1
71599	-1
71601	-1
71602	-1
71609	1
71667	1
71679	1
71704	-1
71710	1
71711	1
71722	-1
71723	-1
71726	1
71729	1
71732	-1
71743	1
71752	1
71755	-1
71756	1
71760	1
71764	-1
71770	1
71780	1
71782	1
71790	1
71793	1
71795	1
71807	1
71816	-1,0
71827	1
71833	1
71838	-1
71841	-1
71862	1
71868	1
71883	1
71898	1
71903	1
71911	1
71914	1
71918	-1
71919	1
71920	1
71929	-1
71934	1
71946	-1,0
71949	1
71956	1
71966	1
71967	-1
71974	-1
71978	-1
71985	1
71985	1
71990	1
71991	-1
71999	-1
72000	-1
72003	-1
72014	1
72026	1
72027	1
72033	1
72039	1
72041	1
72047	1
72053	1

246104	1
246133	1
246179	-1
246190	-1
246198	1
246257	1
246694	-1
246710	-1
246727	1
246728	1
246730	1
246779	-1,0
246787	-1
246788	1
252870	1
252912	1
252967	1
252972	1
252973	1
257632	1
257635	1
257943	1
257947	-2
258155	1
258156	1
258181	1
258196	-1
259097	-1
259101	-1
259277	-1
260298	-1
260302	1
260408	-1
264064	1
264895	1
266690	-1
266815	-1
268448	1
268470	1
268480	1
268482	1
268491	1
268498	1
268512	1
268566	1
268656	-1
268663	-1
268697	-1
268749	-1
268756	-1
268759	-1
268782	1
268783	1
268816	1
268822	1
268857	1
268860	1
268878	1
269053	1
269113	1
269424	1
269437	1
269473	1
269514	-1
269523	-1
269536	-1
269629	1,0
269642	1
269643	1
269682	1
269695	1

19116	1
19125	1
19126	1
19128	1,0
19134	-1
19139	-1
19141	1
19141	1
19142	1
19143	-1
19144	-1
19153	-1
19155	1
19157	1
19158	-1
19159	1
19164	1
19166	-1
19173	-1
19181	1
19188	-1
19193	1
19200	-1
19202	-1
19206	-1
19207	-1
19210	-1
19211	1
19217	-1
19218	1
19218	1
19219	1
19220	1
19228	-1
19229	-1,0
19245	1
19247	1
19248	1
19249	1
19250	-1,0
19256	-1
19259	-1
19260	1
19265	-1
19266	-1
19267	-1
19268	-1
19270	-1
19276	1
19277	-1
19289	-1
19294	-1
19299	1
19300	1
19302	1
19309	1
19340	-1
19345	1
19346	-1
19354	1
19355	1
19356	-1
19357	1
19358	1
19359	-1
19363	1
19364	1
19366	-1
19367	1
19377	-1,0
19378	-1

72054	1
72065	-1
72074	-1
72082	1
72088	-1
72133	1
72135	-1
72141	-1
72145	1
72149	1
72151	1
72155	1
72157	-1
72162	1
72168	1
72169	-1
72179	-1
72190	1
72193	1
72194	1
72195	1
72199	1
72258	1
72281	1
72297	1
72303	1
72310	-1
72318	1
72333	1
72341	-1
72349	1
72354	-1
72361	1
72388	1
72399	1
72400	-1
72404	-1
72413	1
72429	-1
72431	-1
72433	-1
72459	-1
72461	-1
72469	1
72475	-1
72479	-1
72486	-1
72502	1
72508	1
72535	-1
72552	1
72560	-1
72562	-1
72584	-1
72587	1
72590	-1
72599	1
72607	1
72615	1
72634	1
72640	1
72667	-1
72739	-1
72749	1,0
72776	1
72828	-1
72831	-1
72844	1
72852	-1
72873	1
72898	1

269702	1
269713	1
269870	-1
269881	-1
269941	-1
269951	-1
269954	-1
269955	-1
269966	-1
269994	-1
270004	-1
270058	1
270084	1
270097	1
270118	-1
270120	-1
270150	-1
270152	-1
270156	-1
270160	-1
270166	-1
270190	-1
270198	-1
270198	-1
270599	-1
270627	-1
270893	1
271127	-1
271144	-1
271209	-1
271305	1
271564	1
271970	1
271981	1
272158	1
272322	-1
272359	-1
272396	-1
272411	-1
272428	-1
276905	1
276952	1
277939	-1
277973	1,0
278097	-1
278240	-1
279653	-1
280408	1
280662	1
286940	-1
286942	1
317717	1
319146	-1
319197	-1
319207	-1
319239	-1
319293	-1
319387	1,0
319446	1
319478	1
319480	-1
319481	1
319518	1
319594	1
319636	-1
319670	1
319748	-1
319758	-1
319765	1
319804	1
319924	1

19387	1
19395	1
19401	1
19411	1
19415	1
19417	-1
19645	-1
19647	-1
19654	-1,0
19659	-1
19660	-1
19662	1
19664	1
19679	1
19684	-1
19687	1
19691	-1
19697	1
19698	-1
19699	1
19699	1
19704	1
19708	1
19712	1
19719	1
19720	-1
19724	1
19726	1
19739	1
19744	1
19766	-1
19771	-1
19773	1
19777	-1
19820	-1
19879	1
19882	-1
19883	-1
19885	1
19889	-1
19892	1
19894	1
19895	1
19983	1
20019	1
20021	1
20022	1
20024	1
20132	1
20133	-1
20147	-1
20168	1
20170	1
20174	-1
20184	-1
20185	1,0
20187	-1
20190	-1,0
20191	-1
20194	1
20200	1
20202	1
20211	-1
20226	1
20227	1
20238	-1
20239	1
20249	1
20250	1
20255	-1
20256	1

72925	1
72935	-1
72938	-1
72960	1
72961	-1
72962	1
72973	1
72981	-1
73032	-1
73062	1
73067	1
73078	1
73086	1
73095	1
73102	-1
73121	1
73166	1
73173	1
73181	-1
73246	1
73250	-1
73251	1
73254	1
73274	-1
73296	-1
73332	-1
73333	1
73341	-1
73347	-1
73381	1
73382	-1,0
73412	-1
73447	-1
73456	-1
73469	-1
73523	-1
73547	-1
73647	1
73658	-1
73670	-1
73699	-1
73707	1
73744	-1
73748	-1
73750	-1
73804	-1
73808	-1
73830	-1
73845	-1
73945	1
73991	1
74006	1
74008	1
74032	1
74039	1
74041	-1
74044	1
74051	1
74052	-1
74055	1
74080	-1
74087	-1
74090	-1
74096	1
74105	-1
74108	1
74125	-1
74127	1
74129	-1
74133	1
74134	-1

319944	1
319945	1
319953	1
319955	-1
320022	1
320024	1
320100	-1
320118	1
320129	1
320213	1
320214	1
320226	-1
320244	1
320277	1
320376	-1
320452	-1
320472	1
320534	1
320571	-1
320595	-1
320615	-1
320634	-1
320655	1
320661	1
320717	1
320790	-1
320795	1
320806	-1
320864	1
320873	1
320878	-1
320910	1
320940	-1
320995	-1,0
321003	1
321006	-1
326618	1
326623	-1
327958	1
327987	1
327992	1
328035	1
328162	1
328232	-1
328234	-1
328250	-1
328424	1
328572	1
328657	1
328660	1
329015	1
329047	1
329064	1,0
329065	1
329693	1
329777	1
329828	-1
329877	-1
329910	-1
330064	1
330119	1
330171	1
330177	1
330189	1
330390	-1
330470	-1
330474	-1
330490	-1
330503	-1
330627	-1,0
330662	-1

20257	1
20264	-1
20266	-1
20271	-1
20273	1
20277	-1
20280	-1
20286	1
20290	1
20292	1
20295	1
20299	1
20306	1,0
20308	1
20309	1
20311	1
20317	1
20319	1
20333	1
20340	1
20341	1
20342	1
20345	1
20346	1
20348	1
20350	-1
20354	-1
20356	1
20364	-1
20371	1
20379	-1
20388	1
20390	-1
20391	1
20400	-1
20404	-1
20408	-1
20409	1
20410	-1
20418	-1
20429	1
20437	-1,1
20438	-1,1
20439	1
20440	1
20441	-1,0
20442	1
20443	-1
20444	1
20446	1
20447	1
20449	-1
20450	-1
20454	1
20460	-1
20465	1,0
20467	1
20469	1
20474	1
20480	-1
20482	1
20492	1
20493	1
20497	1
20498	1
20499	-1
20500	1
20501	1
20502	-1
20504	-1
20505	-1

74136	1
74140	-1
74143	1
74145	-1
74147	1
74148	1
74152	-1
74153	-1
74156	-1
74164	-1
74167	1
74174	1
74180	-1
74195	-1
74197	1
74198	1
74204	-1
74211	-1
74221	1
74222	1
74237	-1
74249	-1
74251	1
74252	1
74256	1
74257	-1
74277	1
74281	1,0
74287	1
74306	-1
74318	1
74325	-1
74337	1
74338	-1
74343	1
74347	1
74351	1
74356	1
74365	-1
74370	1
74374	1
74388	-1
74405	-1
74407	1
74413	1
74414	1
74427	-1
74438	-1
74442	1
74443	-1
74451	1
74464	-1
74469	-1
74480	-1
74488	1
74513	1
74519	-1
74521	1
74525	-1
74570	1
74585	1
74596	1
74610	1
74617	1
74637	1
74666	1
74685	1
74686	1
74691	1
74708	-1
74718	1

330671	-1
330812	1
330812	1
330814	1
330830	1
330836	1
330914	-1
330921	-1
330953	-1
330962	-1,0
331046	-1
331374	-1
331416	-1
331493	-1
332110	1
332131	1
332396	1
332934	-1
333048	1
333050	1
333307	1
333424	-1
333433	-1
333564	-1
333605	-1
333654	-1
333789	1
338337	-1
338351	-1
338354	-1
338370	-1
338372	1
338467	1
338521	-1,1
353187	-1,1
353346	-1
360213	1
380614	1
380711	1
380712	1
380713	1
380718	1
380728	1
380730	1
380732	1
380773	1
380787	1
380836	-1
380840	-1
380842	-1
380863	-1
380912	-1
380916	-1
380921	-1
380928	-1
380969	1
381199	1
381204	1
381217	1
381236	1
381484	1
381485	1
381489	1
381510	-1
381511	-1
381534	-1
381538	-1
381549	-1
381591	-1
381626	1
381668	1

20510	1
20512	1
20514	-1
20523	-1
20524	1
20525	-1
20526	1
20527	-1
20529	-1
20530	-1
20531	1
20533	1
20534	1,0
20535	1
20537	1
20538	1
20540	-1
20556	1
20557	1
20562	1
20563	1
20583	-1,1
20585	1
20586	-1
20588	-1
20589	1
20591	-1
20592	1
20595	-1
20597	-1,0
20602	1
20603	-1
20607	1
20620	-1
20623	-1
20646	-1
20657	1
20658	1
20660	-1
20661	1
20663	1
20667	-1
20679	-1
20683	1
20688	1
20700	2
20701	2
20702	2
20703	2
20704	2
20713	1
20716	1
20719	-1
20723	-1
20733	-1
20737	-1
20741	1
20743	1
20768	-1
20773	1
20775	-1,1
20778	1
20788	1
20817	1
20840	-1
20848	1
20850	1
20851	1
20867	1
20873	1
20874	1

74734	1
74735	-1
74737	-1
74753	-1
74754	-1
74760	1
74763	1
74769	-1
74775	1
74776	1
74838	1
74996	-1
74998	1
75002	1
75051	1,0
75079	1
75106	-1,0
75137	1
75199	-1
75221	1
75288	-1
75339	-1
75404	-1
75410	-1
75420	-1
75465	1
75497	1
75502	1
75504	1
75512	-1
75524	-1
75540	1
75547	-1
75553	1
75578	-1
75596	-1
75597	-1
75607	-1
75622	1
75624	1
75627	1
75645	1
75646	1
75669	-1
75678	-1
75687	1
75691	-1
75692	1
75698	-1
75705	1
75706	1
75717	-1
75717	-1
75723	-1
75735	1
75751	-1
75769	1
75770	-1
75778	1,0
75784	1
75788	1
75805	-1
75811	-1
75826	1
75860	1,0
75871	1
75901	-1
75974	-1
75985	-1
76007	-1
76055	1

381677	1
381678	1
381680	1
381686	1
381693	1
381694	1
381738	1
381823	1
381852	-1
381853	-1
381903	-1
381924	-1
381925	-1
381974	-1
381979	-1
381983	-1
381983	-1
382018	1
382030	1
382044	1
382045	-1,1
382051	1
382056	1
382066	-1
382089	-1
382090	-1
382097	-1
382106	-1
382111	-1
382113	-1
382156	-1
382207	-1
382209	-1
382236	-1
382239	-1
382252	-1
382253	-1
382282	-1
382543	1
382620	1
382864	-1,0
382867	-1
382913	-1
382985	1
384009	-1
384198	1
384214	1
384220	1
384569	-1
384763	-1
384997	1
385643	1
385674	1
387285	-1
387565	1
387609	1
399510	1
399548	-1
403187	-1
403395	1
404195	1
404242	-1
404710	1
406219	1
406220	1
406222	1
406223	1
407243	-1,0
407786	-1
408067	-1
414872	-1

20878	1,0
20907	1
20908	1
20909	-1
20916	-1
20924	-1
20926	1
20927	-1
20928	-1
20963	-1
20964	-1
20972	1
20977	-1
20981	-1
21335	1
21338	1
21341	1
21343	1
21345	-1
21349	-1
21350	-1
21351	-1
21357	1
21371	-1
21384	1
21385	1
21387	1
21388	1
21390	1,0
21405	1
21406	-1
21410	1
21411	1
21415	1
21416	1
21418	-1
21422	1
21429	1
21432	1
21667	-1
21672	1
21673	1
21676	-1
21677	-1
21681	1
21682	1
21683	-1
21685	1
21687	-1
21744	1
21745	-1
21750	1
21752	-1
21754	-1
21762	1
21771	1
21803	-1
21804	-1
21807	-1
21809	1
21810	-1
21812	-1
21813	-1
21814	1
21816	-1
21819	1
21823	-1
21827	1,0
21828	-1
21832	1
21833	1

76072	-1
76073	1
76074	1
76089	1,0
76108	-1
76131	1
76137	-1
76184	1
76206	-1
76238	-1
76251	-1
76257	-1
76265	1
76267	1
76279	1
76295	1
76299	-1
76303	1
76308	1
76332	1
76338	-1
76355	-1,0
76366	1
76376	-1
76390	1
76408	1
76459	-1
76484	-1
76485	-1
76491	-1
76499	-1
76500	-1
76560	-1
76561	1
76571	1
76582	-1
76608	-1
76612	-1
76630	1
76645	1
76652	1
76686	-1
76703	1
76713	-1
76719	1
76722	-1
76742	1
76763	-1
76770	1
76773	1
76775	1
76781	-1
76787	-1
76792	1
76800	1
76856	-1
76858	-1
76863	-1
76872	1
76886	-1
76895	-1
76898	-1
76900	1
76927	1
76933	1
76947	-1
76952	1
76954	-1
76971	-1
77011	-1
77036	1

431706	-1
432582	1
432611	1
432677	1
432720	-1
432731	-1
432736	-1
432763	-1
432769	-1
432779	-1
432838	-1
432867	-1
432879	-1
432940	1
433016	1
433215	1
433256	1
433586	1
433638	1
433667	1
433693	-1
433700	-1
433899	1
433904	1
433926	1
433931	1
433938	1
433956	1
434171	-1
434203	-1
434232	-1
434234	-1
434246	-1
434325	1
434436	-1
434437	-1
434438	-1
434439	-1,0
434440	-1
434446	-1,0
434756	-1
434758	-1
434764	-1
434766	-1
434769	-1
434778	-1
434784	-1,0
434903	-1
435286	1
435350	-1
435391	-1
435766	1
435791	-1
435965	-1
436002	-1
436022	-1
436240	-1,0
446099	-1
448987	1
450219	1
541610	-1
544748	1
544817	1
544922	-1
544963	-1
544971	-1
544988	-1
545085	1
545123	1
545136	1
545140	1

21834	-1
21843	-1
21844	1
21846	-1
21847	1
21849	-1
21849	-1
21871	1
21872	-1
21873	1
21881	-1
21885	-1
21887	-1
21888	1
21892	1
21897	1
21898	-1
21899	1
21912	-1
21916	-1
21922	-1
21923	-1
21924	-1
21927	1
21928	1
21933	-1
21935	1
21943	-1
21946	-1
21947	-1
21949	-1
21954	-1
21955	-1
21968	1
21973	1
21974	-1
21976	1
21981	1
21985	1
21990	-1
22003	-1
22004	-1
22017	-1,0
22018	1,0
22021	1
22022	1
22031	1
22035	1
22036	-1
22038	-1
22042	1
22045	1
22061	1
22063	-1
22065	1
22066	1
22067	-1
22068	-1
22088	-1
22116	-1
22122	1
22129	1
22141	-1
22142	1
22143	1
22151	-1
22165	1
22166	-1
22171	1
22173	-1
22177	-1

77041	-1
77044	1
77045	1
77097	1
77113	1
77116	-1
77125	1
77219	1
77254	-1
77286	-1
77305	-1
77318	-1
77407	1
77519	-1
77559	1
77569	1
77573	1
77577	1,0
77578	1
77582	-1
77591	-1
77593	-1
77613	-1
77622	-1
77626	1
77634	-1
77669	1
77697	1
77706	1
77739	-1
77772	1
77782	1
77803	-1
77864	1
77891	-1
77914	1
77945	-1
77974	1
77980	1
78038	-1
78070	-1
78244	1
78267	-1
78283	-1
78317	1
78317	1
78339	1
78394	1
78408	1
78455	1
78459	-1
78473	1
78523	1
78541	1
78558	1
78610	-1
78618	1
78753	1
78754	-1
78767	1
78771	-1
78784	1
78787	-1
78796	1
78816	1
78826	-1
78829	1
78832	1
78889	1
78891	1
78894	1

545276	1
545279	1
545288	1
545291	1
545527	1
545622	-1
545667	-1
545725	1
545798	1
545902	-1
546118	-1
546134	-1
546144	-1
546161	-1
546519	1
546726	1
546837	-1
546912	-1
546913	-1
546980	-1
547176	-1
547253	1
554292	-1,0
594844	-1
606496	-1
619287	-1
619548	-1
620592	-1
620779	-1
621852	-1
622127	1
622139	1
622301	-1
622402	-1
622434	1
622665	-1
622675	1
623474	-1
624855	1
624866	1
625321	-1,0
625716	-1
626359	-1
627191	1
627280	-1
627585	1
629016	-1
629378	-1
630146	1,0
632687	1
632708	-1
634104	1
636104	-1
636931	-1
637027	-1,0
639774	-1
639781	-1,0
654362	1,0
654458	-1
654465	-1
654470	1
654801	-1
664723	-1,0
664725	-1
664994	-1
665095	-1
665203	-1
665229	-1
665563	1
666048	-1
666060	-1

22186	1
22194	-1
22195	1
22201	-1
22215	-1
22217	1
22223	1
22227	1
22228	-1
22229	-1
22232	-1
22234	-1
22235	1
22239	1
22241	1
22247	1
22249	-1
22253	1
22255	1
22256	1
22258	-1
22260	-1
22262	1
22273	-1
22276	-1
22284	-1
22287	1,0
22290	1
22293	1
22320	1
22325	1,0
22330	-1

78912	1
78913	1
78914	-1
78919	1
78920	1
78921	-1
78933	-1
78935	-1
78938	-1
78943	1
79264	1
79362	-1
79410	-1,0
79456	1
79464	1
79565	1
80291	1
80334	1
80707	1
80720	1
80751	1
80794	-1
80797	1
80837	1
80877	1
80879	1
80888	1
80890	1
80891	1
80893	-1
80898	-1
80902	-1

666168	-1
666173	1
666244	-1
666279	1
666348	1
666468	-1
666528	-1
667034	-1
667370	1
667736	-1
668110	1
668178	1
668257	-1
668303	1
670533	1
791260	-1
100009600	-1
100038804	1,0
100038993	-1
100039436	-1,0
100039672	-1
100040276	1
100040843	-1
100040937	-1
100040972	-1
100041449	1
100041581	-1
100041835	1
100042265	-1,0
100042314	-1
100042555	-1
100043803	1