



OPEN ACCESS

SHORT REPORT

Combined exome and whole-genome sequencing identifies mutations in *ARMC4* as a cause of primary ciliary dyskinesia with defects in the outer dynein arm

Alexandros Onoufriadis,¹ Amelia Shoemark,² Mustafa M Munye,³ Chela T James,⁴ Miriam Schmidts,¹ Mitali Patel,¹ Elisabeth M Rosser,⁵ Chiara Bacchelli,⁴ Philip L Beales,¹ Peter J Scambler,¹ Stephen L Hart,³ Jeannette E Danke-Roelse,⁶ John J Sloper,⁷ Sarah Hull,⁷ Claire Hogg,² Richard D Emes,^{8,9} Gerard Pals,¹⁰ Anthony T Moore,⁷ Eddie M K Chung,¹¹ UK10K,¹² Hannah M Mitchison,¹

► Additional material is published online only. To view please visit the journal online (<http://dx.doi.org/10.1136/jmedgenet-2013-101938>).

For numbered affiliations see end of article.

Correspondence to

Dr Hannah M Mitchison, Molecular Medicine Unit, University College London (UCL) Institute of Child Health, London WC1N 1EH, UK; h.mitchison@ucl.ac.uk

Received 19 July 2013

Revised 18 September 2013

Accepted 19 September 2013

Published Online First

7 November 2013

ABSTRACT

Background Primary ciliary dyskinesia (PCD) is a rare, genetically heterogeneous ciliopathy disorder affecting cilia and sperm motility. A range of ultrastructural defects of the axoneme underlie the disease, which is characterised by chronic respiratory symptoms and obstructive lung disease, infertility and body axis laterality defects. We applied a next-generation sequencing approach to identify the gene responsible for this phenotype in two consanguineous families.

Methods and results Data from whole-exome sequencing in a consanguineous Turkish family, and whole-genome sequencing in the obligate carrier parents of a consanguineous Pakistani family was combined to identify homozygous loss-of-function mutations in *ARMC4*, segregating in all five affected individuals from both families. Both families carried nonsense mutations within the highly conserved armadillo repeat region of *ARMC4*: c.2675C>A; pSer892* and c.1972G>T; p.Glu658*. A deficiency of *ARMC4* protein was seen in patient's respiratory cilia accompanied by loss of the distal outer dynein arm motors responsible for generating ciliary beating, giving rise to cilia immotility. *ARMC4* gene expression is upregulated during ciliogenesis, and we found a predicted interaction with the outer dynein arm protein DNAI2, mutations in which also cause PCD.

Conclusions We report the first use of whole-genome sequencing to identify gene mutations causing PCD. Loss-of-function mutations in *ARMC4* cause PCD with situs inversus and cilia immotility, associated with a loss of the distal outer (but not inner) dynein arms. This addition of *ARMC4* to the list of genes associated with ciliary outer dynein arm defects expands our understanding of the complexities of PCD genetics.

Primary ciliary dyskinesia (PCD; MIM244400) is a heterogeneous genetic disorder arising from ultrastructural defects that cause abnormal function of motile cilia and sperm flagella.¹ The disease has an autosomal recessive mode of inheritance and affects one in every 15 000–30 000 births. Motile cilia are hair-like organelles found on the epithelial surface of the respiratory airway tract, the brain

ependyma and fallopian tubes. Their axonemal structure consists of nine peripheral outer doublet microtubules surrounding a central microtubular pair (9+2 arrangement), a highly similar structure to that found in sperm tail flagella. During embryogenesis, motile monocilia at the node lack the central pair apparatus (9+0 arrangement). Microtubule-associated protein complexes are attached along the length of the axoneme at regular intervals, which regulate axonemal stability and ciliary motility. Of these, the inner and outer dynein arms (IDA and ODA) are responsible for beat generation together with radial spoke and nexin-dynein regulatory complexes.

Abnormal motility of respiratory cilia leads to congestion of the body's mucociliary clearance mechanism, causing a number of symptoms in PCD patients which include neonatal respiratory distress, chronic respiratory infections, sinusitis, otitis media and destructive lung disease (bronchiectasis).² Other features include subfertility in both sexes and left-right organ laterality abnormalities, predominantly situs inversus, and occasional hydrocephalus. In a proportion of patients, severe heterotaxic isomerisms and cardiac malformations can occur.³ So far, genetic studies have identified mutations in over 20 genes leading to various ultrastructural defects, including *RPGR* which causes a syndromic form of PCD.⁴ Large-scale transmission electron microscopy (TEM) studies in patients suggest that 65% of PCD arises from defects involving the outer dynein arms.^{5,6} Of the genes which are known to cause these defects (reduction or loss of the ODAs) when mutated, *DNAH5*, *DNAH11*, *DNAI1*, *DNAI2*, *DNAL1* and *NME8* (previously *TXNDC3*) encode subunits of the axonemal outer dynein arm components, and *CCDC114* encodes an outer dynein arm-docking complex component. Mutations causing PCD with ODA defects were also recently described by Hjeij *et al* in *ARMC4*, encoding a protein involved in assembling outer dynein arms into cilia which is likely involved in their targeting and/or anchoring onto microtubules.⁷ Mutations involving deficiency of both the outer and inner dynein arms have also



Open Access
Scan to access more
free content

To cite: Onoufriadis A, Shoemark A, Munye MM, *et al*. *J Med Genet* 2014;**51**:61–67.

been identified in genes that encode a group of cytoplasmic proteins (*DNAAF1/LRRC50*, *DNAAF2/KTU*, *DNAAF3*, *CCDC103*, *HEATR2*, *LRRC6* and *DYX1C1*), which are likely to play a role in the preassembly of the dynein arm components and/or in their axonemal transport.^{4–8} Additionally, mutations have been reported in genes that encode protein subunits of the radial spoke heads (*RSPH4A*, *RSPH9*), proteins linked to the nexin-dynein regulatory complexes (*CCDC39*, *CCDC40*, *CCDC164*) and the central pair apparatus (*HYDIN*).⁴

In order to determine the genetic basis of disease in PCD families, we have employed a next-generation sequencing approach. All patient samples in this study were obtained with informed consent according to the protocols approved by the ethical committees of the Institute of Child Health/Great Ormond Street Hospital (#08/H0713/82) and those of collaborating institutions. First, whole-exome sequencing was performed in one affected individual (II:1) of a consanguineous Turkish family PCD-221 at the Wellcome Trust Sanger Institute (Cambridge, UK) as part of the UK10K project.⁹ PCD-221 II:1 is one of two affected siblings who present with a classic clinical course, both having laterality defects, respiratory symptoms and recurrent chest infections since a young age, compromised lung function, chronic ear, nose and throat (ENT) symptoms including rhinitis and otitis media. PCD-221 II:1 suffers from frequent pneumonias and bronchiectasis, and II:2 has had surgery for hydronephrosis. Additionally, both siblings have intellectual and developmental delay, and an unusual ocular phenotype of bilateral ptosis, variable divergent strabismus and upgaze paresis with poor horizontal saccades, suggestive of a brain stem or cranial nerve disinnervation syndrome. Exome sequencing and variant calling was performed as previously described, using approximately 3 µg of genomic DNA and the Agilent Technologies Human All Exon 50 Mb kit.^{4–9} Over 3 Gb of sequence was generated, such that >68% of the target exome was present at greater than 20-fold coverage (see online supplementary table S1). Analysis of the exome variant profile was performed using the EVAR software tool V0.2.2 β. Copy number variations (CNV) were analysed from the exome data using ExomeDepth.¹⁰

To prioritise candidate genes, we based our analysis on the rare-recessive disease model, with the knowledge that PCD is largely caused by mutations affecting the protein-coding region of genes. Since the PCD-221 II:1 individual is the offspring of a consanguineous marriage, we focused on homozygous variants predicted to cause non-synonymous or splice-site substitutions or indels. We also filtered to prioritise only those that were either novel or present in the 1000 Genomes Project exome database with a frequency <0.01. We next used our in-house internal allele count data, removing variants detected more than 10 times across a database of 500 exomes available from the UK10K_RARE cohort (<http://www.uk10k.org/studies/rarediseases.html>), because PCD-causing mutations would not be predicted as likely to appear in multiple well-phenotyped non-PCD patients. This filtering strategy revealed eight homozygous variants of interest that met these criteria, which are listed in online supplementary table S2. We proceeded to search for the presence of these genes and their species-conserved homologues in the Cilia Proteome database,¹¹ and this identified three genes *EEF1D*, *MYO1D*, *ARMC4* (see online supplementary table S2). Of these, *EEF1D* encoding a translation elongation factor was excluded on gene function grounds¹² and also since further analysis showed that the missense change identified was only in a highly truncated transcript of unknown functional significance (ENST00000532400). The *MYO1D* variant was also

excluded based on putative gene function despite the suggested link between *MYO1D* and left-right asymmetry determination since it is a widely expressed cytoskeleton-associated unconventional myosin.^{13–14} Furthermore, the identified variant was a missense mutation scored as 'benign' (score 0.292) in Polyphen-2 and 'tolerated' in Sorting Intolerant From Tolerant software (SIFT) for its effect on protein function (c.2585A>T; p.His862Leu). This left a single homozygous protein-truncating nonsense variant (c.2675C>A; p.Ser892*) in *ARMC4*. The expected damaging effect of this variant, which was the only predicted null-effect allele in the final filtered set, provides strong support for its likely pathogenic role. Segregation analysis in all available members of the family including the affected sibling confirmed correct recessive inheritance of this variant (figure 1A and see online supplementary figure S1).

In parallel, whole-genome sequencing (WGS) was performed as part of the UK10K project in the two unaffected, obligate carrier parents of a consanguineous Pakistani PCD family PCD-141 (individuals I:1 and I:2 figure 1A), since material from their affected offspring was not sufficient for exome sequencing. In this family, there are three affected siblings, all of whom display situs inversus and classic disease symptoms similar to those described for PCD221, including repeated infections of the chest, chronic nasal discharges and bronchiectasis. Additionally, all three siblings have had surgery to remove nasal polyps. For WGS, approximately 3 µg of genomic DNA was sheared to 100–1000 bp (Covaris) and the sheared DNA subjected to Illumina Paired-end DNA library preparation. Following size selection (300–500 bp insert size), DNA libraries were sequenced as 100 bp paired-end reads on the HiSeq platform (Illumina). For each subject, more than 91% of genomic bases were represented by at least 24 reads (see online supplementary table S3). We filtered per chromosome, to identify protein-altering heterozygous variants shared by both parents with a MAF<0.01 in the 1000 Genomes Project exome database, using the same criteria as for the exome sequencing. 522 heterozygous variants meeting the filtering criteria were shared between the two parents, out of more than nearly three million heterozygous variants per parental sample. Of these, just 16 shared variants were found in genes represented in the Cilia Proteome database, and only one, *ARMC4*, had any functional annotation suggestive of a role in cilia motility. Thus, this strategy as detailed in online supplementary table S4 revealed a second *ARMC4* protein-truncating nonsense variant (c.1972G>T; p.Glu658*). Notably, of the 16 variants that were shared and homozygous, this was the only stop-gained effect allele, providing further support for a disease-causing effect. Segregation analysis by Sanger sequencing in all available family members confirmed recessive inheritance with consistent genotypes (figure 1A and see online supplementary figure S1). The lack of an ocular phenotype in anyone from family PCD-141 suggests that this finding in family PCD-221 is unconnected to *ARMC4* mutations, and further analysis of the PCD-221 exome data is ongoing to investigate potential other loci.

Another study of mutations in *ARMC4* causing PCD recently reported that the protein is involved in outer dynein arm assembly into the cilia and probably has a role in their correct axonemal docking and targeting.⁷ Furthermore *ARMC4*, encoding the 1044 amino acid Armadillo repeat-containing protein 4, has previously been implicated in ciliogenesis.¹⁵ We also used the UMCG Groningen Gene Network tool which analyses data from 80 000 Gene Expression Omnibus microarrays to predict gene function in Gene Ontology Consortium terms, and found the top-scoring predictions for *ARMC4* were highly significant

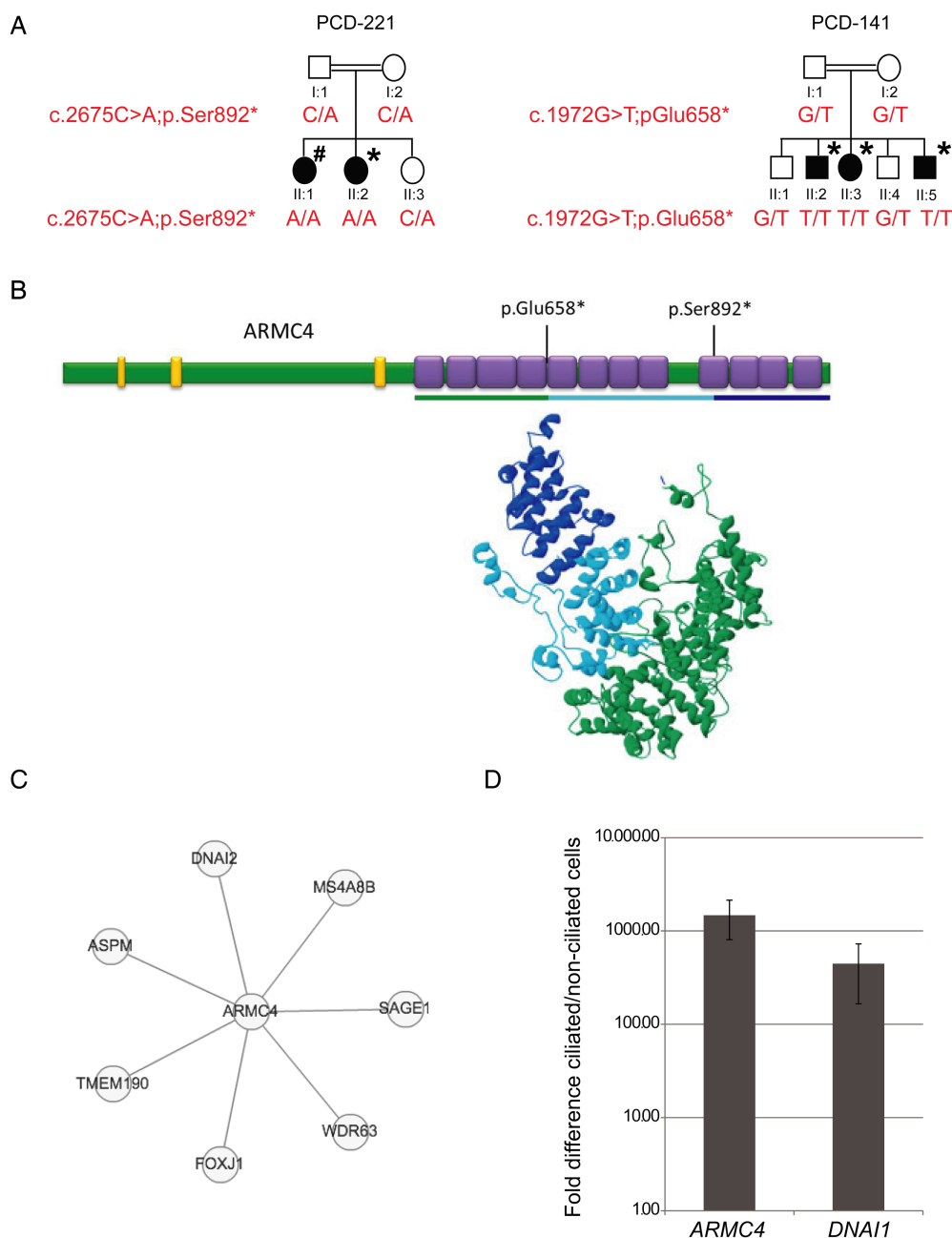


Figure 1 *ARMC4* mutations, protein analysis and upregulation during ciliogenesis. (A) Pedigree structure and segregation analysis of the two primary ciliary dyskinesia (PCD) families in whom *ARMC4* mutations were identified. Asterisk indicates situs inversus, hash indicates dextrocardia. (B) Model of *ARMC4* protein with ARM repeats in purple (residues 482–523, 524–564, 565–620, 621–661, 662–702, 703–744, 745–785, 786–826, 868–910, 911–951, 952–992, 993–1033) and low complexity regions in yellow (residues 73–84, 174–189, 424–438). The predicted ARM repeat superhelix is shown below the linear representation, with dark blue indicating the region lost by the Ser892* mutation, and dark plus light blue the region lost by the Glu658* mutation. (C) STRING predicted protein interactions for *ARMC4*. (D) *ARMC4* expression is induced upon ciliogenesis. A marked increase in *ARMC4* mRNA levels and the control PCD-associated gene *DNAI1* was observed in ciliated normal human bronchial epithelial cells compared to non-ciliated cells. Normal human bronchial epithelial cells (NHBE; Lonza) were cultured submerged on collagen-coated plates with Bronchial Epithelial Growth Media (Lonza), seeded onto transwells and cultured at an air–liquid interface for 28 days essentially as previously described,²⁶ then analysed by qPCR performed with TaqMan primer probes and PCR products detected with an Applied Biosystems (ABI) Prism 7000 Sequence Detection System (Applied Biosystems). The fold difference in gene expression between the two was assessed normalised to *GAPDH* as an endogenous control, using the delta-delta Ct method. The TaqMan probes were *DNAI1*, Hs00201755_m1; *ARMC4*, Hs00216318_m1; *GAPDH*, Hs02758991_g1. Data are plotted as the mean, and error bars represent the SEM of triplicate repeats experiments.

and all involved in cilia functions: ciliary or flagella motility ($p=6.10 \times 10^{-19}$), microtubule-based movement ($p=1.16 \times 10^{-9}$) and cilium assembly ($p=1.16 \times 10^{-9}$). Interestingly, *ARMC4* was previously proposed to be an axonemal protein equivalent to radial spoke protein 8 (RSP8) of *Chlamydomonas*, however, this was acknowledged to be a low-scoring homology.¹⁶ There

appears to be no misfunction of the radial spokes in *ARMC4*-deficient cilia, and this seems to be an erroneous homology, due to the multiple armadillo repeats present in *ARMC4* and RSP8 proteins.¹⁷

We proceeded to use protein modelling to further investigate *ARMC4* function. We could not detect a clear homologue in

the PCD model species, *Chlamydomonas*, but BLAST and SMART¹⁸ protein domain homology searches showed that there are 12 predicted ARM repeats at the C-terminus, in addition to three low amino acid complexity regions of unknown significance in the N-terminus (figure 1B). This contrasts with the study of Hjeij *et al*⁷ which predicts 10 ARM motifs and one HEAT repeat in ARMC4, as is annotated in the Swissprot database (<http://www.ncbi.nlm.nih.gov/protein/74744660>). Both models are based on predictions rather than experimental evidence. The N-terminus did not contain sufficient similarity to any known protein domains to allow modelling, however, a structural model could be generated for the C-terminus using I-TASSER 3D-structural model prediction. This predicts that the tandem ARM-repeat domains of ARMC4 fold together as a series of tandem helices forming a superhelix, that creates a surface or groove for protein interaction similar to that of the β -catenin ARM repeat structure (figure 1B).^{19–20} The two nonsense mutations identified in the PCD families are most likely to be non-functional through nonsense-mediated decay (NMD); as shown in figure 1B, they would both remove multiple ARM domains and that by inference would likely influence the protein binding capabilities of ARMC4 if a truncated form was present.

We then performed a STRING search to look for predicted protein–protein interactions for ARMC4. This interactome analysis generated a small network (figure 1C) including a predicted direct interaction between ARMC4 and the dynein intermediate chain protein DNAI2, a known component of the outer dynein arm located at the ODA base, mutations in which are also responsible for causing PCD with outer dynein arm defects and cilia immotility.²¹ Notably, a direct interaction with FOXJ1 the master regulator of motile ciliogenesis was also predicted in this analysis. FOXJ1 is essential for assembly of motile cilia in vertebrates, through the regulation of genes specific to motile cilia.²² Last, we investigated ARMC4 expression during ciliogenesis of normal human ciliated bronchial epithelial (NHBE) cells by TaqMan qPCR. ARMC4 transcript levels were undetectable in non-ciliated basal NHBE cells, whereas after ciliogenesis, a significant upregulation (over 1000-fold) of ARMC4 expression was observed. Similar results were obtained for DNAI1, a known ciliary gene (figure 1D).

To understand the impact of mutations in ARMC4, we examined transmission electron microscopy (TEM) of respiratory cilia cross-sections from nasal samples collected from the two families carrying ARMC4 mutations. TEM was performed as previously described.⁶ All affected individuals displayed a loss of the outer dynein arms, as demonstrated in figure 2A for the affected individuals PCD-221 II:1 and PCD-141 II:5. In order to further investigate ARMC4 function, we analysed its subcellular localisation by high-resolution immunofluorescence microscopy in the patient's ciliated epithelial cells as previously described.²³ In control cells, ARMC4 protein was observed to localise along the whole length of the cilia axoneme (figure 2B). However, in the PCD patient, PCD-221 II:1, severely reduced levels of ARMC4 along the cilia axoneme were observed in contrast to the control individual, in comparison to the axonemal marker acetylated α -tubulin which was unaffected (figure 2B). There was a notable accumulation of ARMC4 staining in the cell body in the patient's cells, apparently clustered directly underneath the ciliary base. This needs further investigation but could conceivably represent an accumulation of truncated protein not subject to NMD, and it is also seen in the previous study of Hjeij *et al*⁷ in patients carrying ARMC4 termination mutations. We next used two well-

established diagnostic markers of axoneme integrity, DNAH5 and DNALI1, to examine cilia structure in ARMC4 patients. DNAH5 detects both the different types of outer dynein arms that have previously been described in respiratory cilia: one class at the distal half of the axoneme (DNAH5-positive, DNAH9-positive) and one class at the proximal end (DNAH5-positive, DNAH9-negative).²⁴ DNALI1 stains the inner dynein arms along the entire axoneme length.²¹ This analysis confirmed the presence of the IDAs in cilia of PCD-221 II:1 (see online supplementary figure S2). However, in contrast there was a complete loss of DNAH5 staining at the distal ends of the cilia, but a retention of a reduced level of DNAH5 staining in the proximal ends of cilia (figure 2B). This indicates the loss of distal DNAH5-positive ODAs but retention of proximal DNAH5-positive ODAs which is in agreement with previous observations by Hjeij *et al*⁷ in ARMC4-deficient cilia.

To demonstrate the effect of ARMC4 deficiency on ciliary beat frequency, we also performed high-speed videomicroscopic analysis of patient's nasal cilia as previously described.²³ In both the affected siblings PCD-221 II:1 and II:2, the cilia were immotile compared with controls, which is a consistent pattern as seen in many other patients with outer dynein arm defects due to mutations in various different genes. The occasional twitch was seen in some cilia but in the majority they were completely static (see online supplementary videos 1–2). Taken together, these findings suggest that genetic defects in ARMC4 result in loss of selected ODAs.

In summary, here we show that loss-of-function mutations in ARMC4 cause PCD associated with left-right axis defects and a loss of the cilia's distal outer dynein arms. Both the nonsense mutations identified result in a deficiency of the ARMC4 protein along the length of the ciliary axoneme with an accompanying loss of the distal but not the proximal outer dynein arms, as defined by DNAH5 immunostaining, and this is associated with cilia immotility. Both mutations affect the highly conserved ARM repeat superhelix at the protein's C-terminus, likely disrupting its interactions with other protein partners. ARMC4 is, therefore, the eighth gene to be associated with a deficiency of outer dynein arms, and retention of the inner dynein arms. A previous report⁷ has shown that ARMC4 is not likely to be an integral component of the outer dynein arm, but is more probably involved in ODA targeting, docking or attachment. It is known that armadillo repeat-containing proteins can have more than one function in cells, potentially interacting with different protein partners.¹⁹ The putative interaction we detected between ARMC4 and DNAI2 requires further experimental proof, however if true, it would potentially localise ARMC4 close to DNAI2 within the axonemal outer dynein arm structures, towards the base of the outer dynein arm. Notably, Hjeij have shown in ARMC4-deficient cilia that DNAI2 similarly to DNAH5 is present proximally but absent distally from the ciliary axonemes.

Our findings have clinical application since we have demonstrated for the first time that mutations in genes causing PCD can be identified by combining whole-exome and whole-genome sequencing data. Furthermore, we conclude that the downstream analysis of WGS data can enable the identification of mutations in genes causing disease in cases where genetic material of patients is unavailable or hard to obtain. Our analysis of parental sequencing data of PCD patients shows this can be used to assist the genetic diagnosis of the disease. ARMC4 joins an important group of highly conserved ARM repeat-containing proteins associated with the ciliary axoneme that play roles in

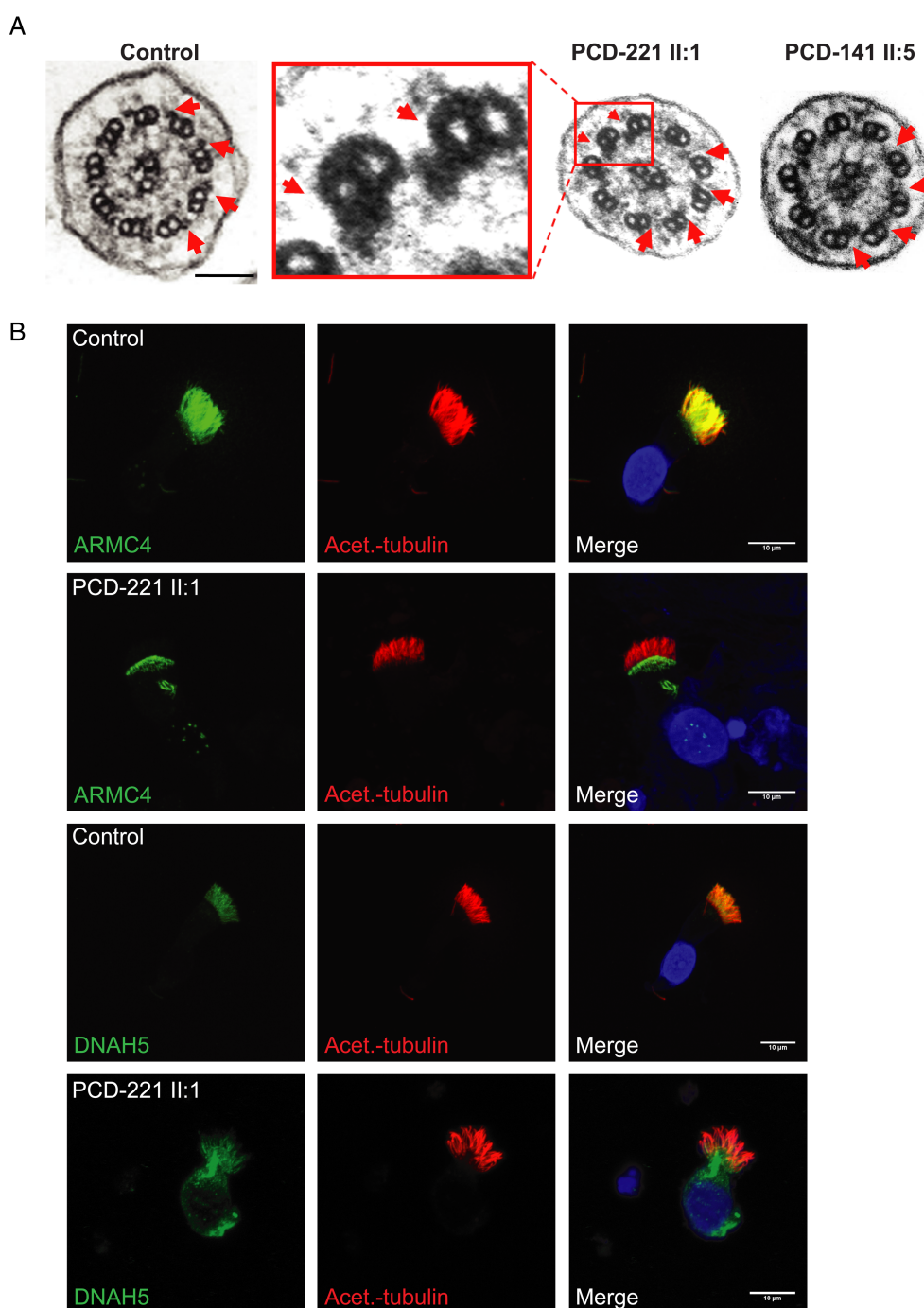


Figure 2 *ARM4* mutations result in outer dynein arms defects. (A) Transmission electron microscopy of nasal respiratory epithelial cell cilia demonstrates loss of the outer dynein arms (indicated by the arrows) in primary ciliary dyskinesia (PCD) patients carrying *ARM4* mutations compared to controls. Scale bar, 100 nm. A close-up of the peripheral outer doublet microtubules in the PCD-221 II:1 patient is indicated in a red box, to highlight the loss of the outer arms with retention of inner arms. (B) In healthy individuals, both *ARM4* and *DNAH5* are localised along the length of the axoneme of respiratory epithelial cells. In individual PCD-221 II:1, *ARM4* is markedly reduced along the full length of the axoneme and *DNAH5* is present only in the proximal portion of the cilia closest to the basal body, but significantly reduced at the distal end of the cilia. Cilia axonemal immunostaining was performed using antisera against *ARM4* (Sigma HPA037829) and the outer dynein arm protein *DNAH5* (Sigma HPA037470) (both green). Axoneme-specific anti-acetylated- α -tubulin antibody (Sigma) was used as a control to stain the entire axoneme (red), and (DAPI is 4',6-diamidino-2-phenylindole) (Invitrogen) was used to stain DNA (blue). Scale bars, 10 μ m.

motility which includes RSP8, RSP14 and PF16.^{17 19 25} The exact nature of the essential role of *ARM4* in targeting the outer dynein arms to cilia remains to be fully characterised, however, these results further expand our understanding of the molecular genetic basis of PCD, and facilitate the rapidly growing application of genetics in PCD diagnostics.

WEB RESOURCES

1000 Genomes Project, <http://www.1000genomes.org>
 OMIM, <http://www.omim.org>
 Polyphen-2, <http://genetics.bwh.harvard.edu/pph2/index.shtml>
 UMCG Groningen Gene Network, <http://genenetwork.nl:8080/GeneNetwork/>

BLAST, <http://blast.ncbi.nlm.nih.gov/Blast.cgi>

SMART, <http://smart.embl-heidelberg.de/>

I-TASSER, <http://zhanglab.ccmb.med.umich.edu/I-TASSER/>

STRING, <http://string-db.org/>

SWISSPROT, <http://www.uniprot.org/>

Author affiliations

¹Molecular Medicine Unit, Birth Defects Research Centre, Institute of Child Health, University College London, London, UK

²Department of Paediatric Respiratory Medicine, Royal Brompton and Harefield NHS Trust, London, UK

³Molecular Immunology Unit, Institute of Child Health, University College London, London, UK

⁴Centre for Translational Genomics-GOSgene, Institute of Child Health, University College London, London, UK

⁵Clinical Genetics Unit, Great Ormond Street Hospital, London, UK

⁶Department of Pediatrics, Atrium Medical Center, Heerlen, The Netherlands

⁷UCL Institute of Ophthalmology, Moorfields Eye Hospital, London, UK

⁸School of Veterinary Medicine and Science, University of Nottingham, Nottingham, Leicestershire, UK

⁹Advanced Data Analysis Centre, University of Nottingham, Sutton Bonington Campus, Nottingham, Leicestershire, UK

¹⁰Department of Clinical Genetics, VU University Medical Center, Amsterdam, The Netherlands

¹¹General and Adolescent Paediatric Unit, Institute of Child Health, University College London, London, UK

¹²<http://www.uk10k.org/>

Acknowledgements We would like to thank the PCD families for their participation in the study, and the PCD Family Support Group. We are also grateful to the physicians involved in analysis of the families especially Alison Male and Siobhan Carr. We thank Sarah Olsson and Andrew Rogers for light and electron microscopy. The Centre for Translational Genomics-GOSgene at the UCL Institute of Child Health is supported by the National Institute for Health Research Biomedical Research Centre at Great Ormond Street Hospital for Children NHS Foundation Trust and UCL Institute of Child Health. We are grateful to the UK10K consortium in particular the Rare Diseases Group for making this study possible; a full list of the UK10K investigators is available at http://www.uk10k.org/publications_and_posters.html.

Contributors Exome and genome sequencing data is from the UK10K project. Analysis of the sequence data was performed by AO, MS, CTJ and CB. Gene expression analysis was by MMM and SLH. Immunofluorescence and other molecular analysis was performed by AO, MS and MP. Clinical studies are from EMR, JED-R, CH, JJS, SH, GP, ATM and EMKC. Electron microscopy and video analysis by AS and CH. AO coordinated and performed molecular studies, he and HMM designed the study and wrote the manuscript.

Funding Wellcome Trust, Milena Carvajal Pro-Kartagener Foundation, Action Medical Research, Newlife Foundation.

Competing interests Funding for UK10K was provided by the Wellcome Trust under award WT091310. JJS, SH and ATM are supported by the Moorfields Eye Hospital Biomedical Research Centre. P.J.S. is supported by the Wellcome Trust and the British Heart Foundation. P.L.B. is a Wellcome Trust Senior Fellow. MS is supported by an Action Medical Research UK Clinical Training Fellowship. MMM, PLB, PJS, SLH and HMM are supported by the Great Ormond Street Hospital Children's Charity and a Child Health Research Appeal Trust funded PhD studentship to MMM (SLH). GP is supported by the Dutch patient organisation PCD Belangengroep, funded by 'It Krysteam' (Friesland). EMKC and HMM are supported by grants from the Milena Carvajal Pro-Kartagener Foundation, Action Medical Research (GN1773, GN2101) and Newlife Foundation for Disabled Children UK (10-11/15).

Ethics approval Ethical Committee of the Institute of Child Health/Great Ormond Street Hospital (#08/H0713/82).

Provenance and peer review Not commissioned; externally peer reviewed.

Open Access This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 3.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/3.0/>

REFERENCES

1 Bush A, Chodhari R, Collins N, Copeland F, Hall P, Harcourt J, Hariri M, Hogg C, Lucas J, Mitchison HM, O'Callaghan C, Phillips G. Primary ciliary dyskinesia: current state of the art. *Arch Dis Child* 2007;92:1136–40.

- 2 Knowles MR, Daniels LA, Davis SD, Zariwala MA, Leigh MW. Primary Ciliary dyskinesia: recent advances in diagnostics, genetics, and characterization of clinical disease. *Am J Respir Crit Care Med* 2013. [Epub ahead of print].
- 3 Kennedy MP, Omran H, Leigh MW, Dell S, Morgan L, Molina PL, Robinson BV, Minnick SL, Olbrich H, Severin T, Ahrens P, Lange L, Morillas HN, Noone PG, Zariwala MA, Knowles MR. Congenital heart disease and other heterotaxic defects in a large cohort of patients with primary ciliary dyskinesia. *Circulation* 2007;115:2814–21.
- 4 Onoufriadis A, Paff T, Antony D, Shoemark A, Micha D, Kuyt B, Schmidts M, Petridi S, Dankert-Roelse JE, Haarman EG, Daniels JM, Emes RD, Wilson R, Hogg C, Scambler PJ, Chung EM, Pals G, Mitchison HM. Splice-site mutations in the axonemal outer dynein arm docking complex gene CCDC114 cause primary ciliary dyskinesia. *Am J Hum Genet* 2013;92:88–98.
- 5 Papon JF, Coste A, Roudot-Thoraval F, Boucherat M, Roger G, Tamalet A, Vojtek AM, Amselem S, Escudier E. A 20-year experience of electron microscopy in the diagnosis of primary ciliary dyskinesia. *Eur Respir J* 2010;35:1057–63.
- 6 Shoemark A, Dixon M, Corrin B, Dewar A. Twenty-year review of quantitative transmission electron microscopy for the diagnosis of primary ciliary dyskinesia. *J Clin Pathol* 2012;65:267–71.
- 7 Hjejri R, Lindstrand A, Francis R, Zariwala MA, Liu X, Li Y, Damerla R, Dougherty GW, Abouhamed M, Olbrich H, Loges NT, Pennekamp P, Davis EE, Carvalho CMB, Pehlivan D, Werner C, Raidt J, Köhler G, Häffner K, Reyes-Mugica M, Lupski JR, Leigh MW, Rosenfeld M, Morgan LC, Knowles MR, Lo CW, Katsanis N, Omran H. (2013) ARMC4 mutations cause primary ciliary dyskinesia with randomization of left/right body asymmetry. *Am J Hum Genet* 2013;93:357–67.
- 8 Tarkar A, Loges NT, Slagle CE, Francis R, Dougherty GW, Tamayo JV, Shook B, Cantino M, Schwartz D, Jahnke C, Olbrich H, Werner C, Raidt R, Pennekamp P, Abouhamed M, Hjejri R, Köhler G, Griese M, Li Y, Lemke K, Kléna N, Liu X, Gabriel G, Tobita K, Jaspers M, Morgan LC, Shapiro AJ, Letteboer SJF, Mans DA, Carson JL, Leigh MW, Wolf WE, Chen S, Lucas JS, Onoufriadis A, Plagnol V, Schmidts M, Boldt K, Roepman R, Zariwala M, Lo CW, Mitchison HM, Knowles MR, Burdine RD, LoTurco JJ, Omran H; UK10K. DYX1C1 is required for axonemal dynein assembly and ciliary motility. *Nat Genet* 2013;45:995–1003.
- 9 Olbrich H, Schmidts M, Werner C, Onoufriadis A, Loges NT, Raidt J, Banki NF, Shoemark A, Burgoyne T, Al TS, Hurler ME, Kohler G, Schroeder J, Nurnberg G, Nurnberg P, Chung EM, Reinhardt R, Marthin JK, Nielsen KG, Mitchison HM, Omran H. Recessive HYDIN mutations cause primary ciliary dyskinesia without randomization of left-right body asymmetry. *Am J Hum Genet* 2012;91:672–84.
- 10 Plagnol V, Curtis J, Epstein M, Mok KY, Stebbings E, Grigoriadou S, Wood NW, Hambleton S, Burns SO, Thrasher AJ, Kumararatne D, Doffinger R, Nejentsev S. A robust model for read count data in exome sequencing experiments and implications for copy number variant calling. *Bioinformatics* 2012;28:2747–54.
- 11 Gherman A, Davis EE, Katsanis N. The ciliary proteome database: an integrated community resource for the genetic and functional dissection of cilia. *Nat Genet* 2006;38:961–2.
- 12 Ong LL, Er CP, Ho A, Aung MT, Yu H. Kinectin anchors the translation elongation factor-1 delta to the endoplasmic reticulum. *J Biol Chem* 2003;278:32115–23.
- 13 Hozumi S, Maeda R, Taniguchi K, Kanai M, Shirakabe S, Sasamura T, Speder P, Noselli S, Aigaki T, Murakami R, Matsuno K. An unconventional myosin in *Drosophila* reverses the default handedness in visceral organs. *Nature* 2006;440:798–802.
- 14 Speder P, Adam G, Noselli S. Type ID unconventional myosin controls left-right asymmetry in *Drosophila*. *Nature* 2006;440:803–7.
- 15 Loneragan KM, Chari R, Deleeuw RJ, Shadoo A, Chi B, Tsao MS, Jones S, Marra M, Ling V, Ng R, Macaulay C, Lam S, Lam WL. Identification of novel lung genes in bronchial epithelium by serial analysis of gene expression. *Am J Respir Cell Mol Biol* 2006;35:651–61.
- 16 O'Toole ET, Giddings TH Jr, Porter ME, Ostrowski LE. Computer-assisted image analysis of human cilia and *Chlamydomonas* flagella reveals both similarities and differences in axoneme structure. *Cytoskeleton (Hoboken)* 2012;69:577–90.
- 17 Yang P, Diener DR, Yang C, Kohno T, Pazour GJ, Dienes JM, Agrin NS, King SM, Sale WS, Kamiya R, Rosenbaum JL, Witman GB. Radial spoke proteins of *Chlamydomonas* flagella. *J Cell Sci* 2006;119(Pt 6):1165–74.
- 18 Schultz J, Copley RR, Doerks T, Ponting CP, Bork P. SMART: a web-based tool for the study of genetically mobile domains. *Nucleic Acids Res* 2000;28:231–4.
- 19 Tewari R, Bailes E, Bunting KA, Coates JC. Armadillo-repeat protein functions: questions for little creatures. *Trends Cell Biol* 2010;20:470–81.
- 20 Xu W, Kimelman D. Mechanistic insights from structural studies of beta-catenin and its binding partners. *J Cell Sci* 2007;120(Pt 19):3337–44.
- 21 Loges NT, Olbrich H, Fenske L, Mussaffi H, Horvath J, Fliegau M, Kuhl H, Baktai G, Peterffy E, Chodhari R, Chung EM, Rutman A, O'Callaghan C, Blau H, Tiszlavicz L, Voelkel K, Witt M, Zietkiewicz E, Neesen J, Reinhardt R, Mitchison HM, Omran H. DNAI2 mutations cause primary ciliary dyskinesia with defects in the outer dynein arm. *Am J Hum Genet* 2008;83:547–58.
- 22 Roy S. The motile cilium in development and disease: emerging new insights. *Bioessays* 2009;31:694–9.

- 23 Antony D, Becker-Heck A, Zariwala MA, Schmidts M, Onoufriadis A, Forouhan M, Wilson R, Taylor-Cox T, Dewar A, Jackson C, Goggin P, Loges NT, Olbrich H, Jaspers M, Jorissen M, Leigh MW, Wolf WE, Daniels ML, Noone PG, Ferkol TW, Sagel SD, Rosenfeld M, Rutman A, Dixit A, O'Callaghan C, Lucas JS, Hogg C, Scambler PJ, Emes RD, Chung EM, Shoemark A, Knowles MR, Omran H, Mitchison HM; UK10K. Mutations in CCDC39 and CCDC40 are the major cause of primary ciliary dyskinesia with axonemal disorganization and absent inner dynein arms. *Hum Mutat* 2013;34:462–72.
- 24 Fliegauf M, Olbrich H, Horvath J, Wildhaber JH, Zariwala MA, Kennedy M, Knowles MR, Omran H. Mislocalization of DNAH5 and DNAH9 in respiratory cells from patients with primary ciliary dyskinesia. *Am J Respir Crit Care Med* 2005;171:1343–9.
- 25 Zhang Z, Jones BH, Tang W, Moss SB, Wei Z, Ho C, Pollack M, Horowitz E, Bennett J, Baker ME, Strauss JF. 3rd. Dissecting the axoneme interactome: the mammalian orthologue of Chlamydomonas PF6 interacts with sperm-associated antigen 6, the mammalian orthologue of Chlamydomonas PF16. *Mol Cell Proteomics* 2005;4:914–23.
- 26 Hirst RA, Rutman A, Williams G, O'Callaghan C. Ciliated air-liquid cultures as an aid to diagnostic testing of primary ciliary dyskinesia. *Chest* 2010;138:1441–7.