Topological data analysis reveals genotype–phenotype relationships in primary ciliary dyskinesia

SHAREABLE PDF

Amelia Shoemark[1,2,30], Bruna Rubbo[3,4,30], Marie Legendre[5,6], Mahmoud R. Fassad[7,8], Eric G. Haarman[9], Sunayna Best[7,10], Irma C.M. Bon[9], Joost Brandsma[4], Pierre-Regis Burgel[11,12], Gunnar Carlsson[13], Siobhan B. Carr[1], Mary Carroll[3,4], Matt Edwards[14], Estelle Escudier[5,6], Isabelle Honoré[11], David Hunt[15], Gregory Jouvion[5,6], Michel R. Loebinger[16,17], Bernard Maitre[18,19], Deborah Morris-Rosendahl[14], Jean-Francois Papon[20,21,22,23], Camille M. Parsons[24], Mitali P. Patel[7], N. Simon Thomas[25,26], Guillaume Thouvenin[4,27,28], Woolf T. Walker[3,4], Robert Wilson[16], Claire Hogg[1], Hannah M. Mitchison[7,29,31] and Jane S. Lucas[3,4,31]

[1]PCD Diagnostic Centre and Dept of Paediatric Respiratory Medicine, Royal Brompton and Harefield NHS Trust, London, UK. [2]Division of Molecular and Clinical Medicine, University of Dundee, Ninewells Hospital and Medical School, Dundee, UK. [3]Primary Ciliary Dyskinesia Centre, University Hospital Southampton NHS Foundation Trust, Southampton, UK. [4]School of Clinical and Experimental Sciences, University of Southampton Faculty of Medicine, Southampton, UK. [5]Département de Génétique Médicale, Hôpital Trousseau, Assistance Publique–Hôpitaux de Paris (AP–HP), Paris, France. [6]Sorbonne Université, Institut National de la Santé et de la Recherche Médicale (INSERM) U933, Hôpital Trousseau, Paris, France. [7]Genetics and Genomic Medicine Dept, University College London, UCL Great Ormond Street Institute of Child Health, London, UK. [8]Dept of Human Genetics, Medical Research Institute, Alexandria University, Alexandria, Egypt. [9]Dept of Pediatric Pulmonology, Emma Children's Hospital, Amsterdam UMC, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands. [10]Leeds Institute of Medical Research, Faculty of Medicine and Health, University of Leeds, Leeds, UK. [11]Service de Pneumologie, Hôpital Cochin, Assistance Publique–Hôpitaux de Paris (AP–HP), Paris, France. [12]Université de Paris, Institut National de la Santé et de la Recherche Médicale (INSERM) U1016, Institut Cochin, Paris, France. [13]Dept of Mathematics, Stanford University, Stanford, CA, USA. [14]Clinical Genetics and Genomics, Royal Brompton and Harefield NHS Foundation Trust, London, UK. [15]Wessex Clinical Genetics Service, University Hospitals Southampton, Princess Anne Hospital, Southampton, UK. [16]Host Defence Unit, Dept of Respiratory Medicine, Royal Brompton and Harefield NHS Foundation Trust, London, UK. [17]National Heart and Lung Institute (NHLI), Imperial College, London, UK. [18]Service de Pneumologie, DHU A-TVB, Centre Hospitalier Intercommunal de Créteil, Université Paris Est, Créteil, France. [19]Université Paris Est, Institut National de la Santé et de la Recherche Médicale (INSERM) U955, Institut Mondor de Recherche Biomédicale (IMRB), Créteil, France. [20]Service d'ORL et Chirurgie Cervico-Faciale, Hôpital Kremlin-Bicêtre, Assistance Publique–Hôpitaux de Paris (AP–HP), Le Kremlin-Bicêtre, France. [21]Faculté de Médecine, Université Paris-Saclay, Le Kremlin-Bicêtre, France. [22]Centre national de la recherche scientifique (CNRS) ERL 7240, Créteil, France. [23]Institut National de la Santé et de la Recherche Médicale (INSERM) U955, Créteil, France. [24]Medical Research Council (MRC) Lifecourse Epidemiology Unit, University of Southampton, Southampton, UK. [25]Wessex Regional Genetics Laboratory, Salisbury NHS Foundation Trust, Salisbury, UK. [26]Human Genetics and Genomic Medicine, University of Southampton Faculty of Medicine, Southampton, UK. [27]Service de Pneumologie Pédiatrique, Hôpital Trousseau, Assistance Publique–Hôpitaux de Paris (AP–HP), Paris, France. [28]Sorbonne Université, Institut National de la Santé et de la Recherche Médicale (INSERM) U938, Centre de Recherche Saint-Antoine, Paris, France. [29]National Institute for Health Research (NIHR) Great Ormond Street Hospital Biomedical Research Centre, London, UK. [30]Equal first author contribution. [31]H.M. Mitchison and J.S. Lucas contributed equally to this article as lead authors and supervised the work.

Corresponding author: Jane Lucas (jlucas1@soton.ac.uk)

Check for updates

Shareable abstract (@ERSpublications)
Topological data analysis of 396 primary ciliary dyskinesia patients shows genetic mutations of worse (*CCDC39*), variable (*DNAH5*) and milder (*DNAH11*) effects on lung function, offering the potential for more accurately targeted disease management https://bit.ly/3oL5r64

Cite this article as: Shoemark A, Rubbo B, Legendre M, *et al*. Topological data analysis reveals genotype–phenotype relationships in primary ciliary dyskinesia. *Eur Respir J* 2021; 58: 2002359 [DOI: 10.1183/13993003.02359-2020].

This single-page version can be shared freely online.

Abstract

Background Primary ciliary dyskinesia (PCD) is a heterogeneous inherited disorder caused by mutations in approximately 50 cilia-related genes. PCD genotype–phenotype relationships have mostly arisen from

This article has supplementary material available from erj.ersjournals.com

This article has an editorial commentary: https://doi.org/10.1183/13993003.00392-2021

Received: 16 June 2020
Accepted: 24 Dec 2020

small case series because existing statistical approaches to investigating relationships have been unsuitable for rare diseases.

Methods We applied a topological data analysis (TDA) approach to investigate genotype–phenotype relationships in PCD. Data from separate training and validation cohorts included 396 genetically defined individuals carrying pathogenic variants in PCD genes. To develop the TDA models, 12 clinical and diagnostic variables were included. TDA-driven hypotheses were subsequently tested using traditional statistics.

Results Disease severity at diagnosis, measured by forced expiratory volume in 1 s (FEV_1) z-score, was significantly worse in individuals with *CCDC39* mutations (compared to other gene mutations) and better in those with *DNAH11* mutations; the latter also reported less neonatal respiratory distress. Patients without neonatal respiratory distress had better preserved FEV_1 at diagnosis. Individuals with *DNAH5* mutations were phenotypically diverse. Cilia ultrastructure and beat pattern defects correlated closely to specific causative gene groups, confirming these tests can be used to support a genetic diagnosis.

Conclusions This large scale, multi-national study presents PCD as a syndrome with overlapping symptoms and variations in phenotype according to genotype. TDA modelling confirmed genotype–phenotype relationships reported by smaller studies (*e.g.* FEV_1 worse with *CCDC39* mutation) and identified new relationships, including FEV_1 preservation with *DNAH11* mutations and diversity of severity with *DNAH5* mutations.