Humu-2014-0549

Research Article

De novo heterozygous mutations in *SMC3* cause a range of Cornelia de Lange Syndrome-overlapping phenotypes

María Concepción Gil-Rodríguez[1*], Matthew A. Deardorff[2,3*#], Morad Ansari[4*], Christopher A. Tan[5], Ilaria Parenti[6,7], Carolina Baquero-Montoya[1,8], Lilian B. Ousager[9], Beatriz Puisac[1], María Hernández-Marcos[1], María Esperanza Teresa-Rodrigo[1], Iñigo Marcos-Alcalde[10], Jan-Jaap Wesselink[11], Silvia Lusa-Bernal[11], Emilia K. Bijlsma[12], Diana Braunholz[6], Inés Bueno-Martinez[1,13], Dinah Clark[2], Nicola S. Cooper[14], Cynthia J. Curry[15], Richard Fisher[16], Alan Fryer[17], Jaya Ganesh[2,3], Cristina Gervasini[7], Gabriele Gillessen-Kaesbach[18], Yiran Guo[19], Hakon Hakonarson[3,19], Robert J. Hopkin[20], Maninder Kaur[2], Brendan J. Keating[3,19], María Kibaek[21], Esther Kinning[22], Tjitske Kleefstra[23], Antonie D. Kline[24], Ekaterina Kuchinskaya[25], Lidia Larizza[7], Yun R. Li[19,26], Xuanzhu Liu[27], Milena Mariani[28], Jonathan D. Picker[29], Ángeles Pié[1], Jelena Pozojevic[6], Ethel Queralt[30], Julie Richer[31], Elizabeth Roeder[32], Anubha Sinha[14], Richard H. Scott[33,34], Joyce So[35,36,37], Katherine A. Wusik[20], Louise Wilson[33], Jianguo Zhang[27], Paulino Gómez-Puertas[10], César H. Casale[38], Lena Ström[39], Angelo Selicorni[28], Feliciano J. Ramos[1,13], Laird G. Jackson[40], Ian D. Krantz[2,3], Soma Das[5], Raoul C.M. Hennekam[41], Frank J. Kaiser[6], David R. FitzPatrick[4#], Juan Pié[1#].

*These authors contributed equally to this work.

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/humu.22761.

Addresses:

1 Unit of Clinical Genetics and Functional Genomics, Departments of Pharmacology-Physiology and Pediatrics, Medical School, University of Zaragoza, CIBERER-GCV and ISS-Aragon, Zaragoza E-50009, Spain.

2 Divisions of Genetics and Metabolism, Children's Hospital of Philadelphia, Philadelphia, PA, 19104, USA.

3 Department of Pediatrics, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA, 19104, USA.

4 MRC Human Genetics Unit, MRC Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh EH4 2XU, UK.

5 Department of Human Genetics, University of Chicago, Chicago, IL 60637, USA.

6 Sektion für Funktionelle Genetik am Institut für Humangenetik, Universität zu Lübeck, Lübeck D-23538, Germany.

7 Medical Genetics; Department of Health Sciences; Università degli Studi di Milano; Milan, Italy

8 Department of Pediatrics, Hospital Pablo Tobón Uribe, Medellín, 05001000, Colombia.

9 Departments of Clinical Genetics, Odense University Hospital, Odense DK-5000, Denmark.

10 Molecular Modelling Group, Centro de Biología Molecular "Severo Ochoa" (CSIC-UAM), Cantoblanco E-28049, Madrid, Spain.

11 Biomol-Informatics SL. Campus UAM, Madrid E-28049, Spain.

12 Department of Clinical Genetics, Leiden University Medical Centre, Leiden 2300 RC, The Netherlands.

13 Unidad de Genética Clínica, Servicio de Pediatría, Hospital Clínico Universitario "Lozano Blesa", CIBERER-GCV and ISS-Aragón, Zaragoza E-50009, Spain.

14 Clinical Genetics Unit, Birmingham Women's Hospital, Birmingham B15 2TG, UK.

15 Genetic Medicine Central California, University of California San Francisco, Fresno, CA 93701, USA.

16 Northern Genetics Service, Teesside Genetics Unit, The James Cook University Hospital, Middlesbrough TS4 3BW, UK.

17 Department of Clinical Genetics, Liverpool Women's Hospital and Alder Hey Children's Hospital, Liverpool L12 2AP, UK.

18 Institut für Humangenetik, Universität zu Lübeck, Lübeck D-23538, Germany.

19 Center for Applied Genomics, Children's Hospital of Philadelphia, Philadelphia, PA, 19104, USA.

20 Division of Human Genetics, Cincinnati Children's Hospital Medical Center, Cincinnati OH 45229, USA.

21 Department of Pediatrics, HC Andersen Children´s Hospital, Odense DK-5000, Denmark.

22 West of Scotland Genetics Service, Southern General Hospital, Glasgow G51 4TF, UK.

23 Department of Human Genetics, Radboud University Nijmegen Medical Centre, Nijmegen 6500 HB, The Netherlands.

24 The Harvey Institute for Human Genetics, Greater Baltimore Medical Center, Baltimore MD 21204, USA.

25 Department of Clinical Genetics, Linköping University Hospital, Linköping 581 85, Sweden

26 Medical Scientist Training Program, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA, 19104, USA.

27 BGI-Shenzhen, Shenzhen 518083, China.

28 Ambulatorio Genetica Clinica Pediatrica, Clinica Pediatrica Università Milano Bicocca, Fondazione MBBM AOS Gerardo, Monza 20900, Italy.

29 Departments of Genetics and Child Psychiatry, Boston Children's Hospital, Boston MA 02115, USA.

30 Cell Cycle Group, Cancer Epigenetics and Biology Program (PEBC), Institut d'Investigacions Biomèdica de Bellvitge (IDIBELL), L'Hospitalet de Llobregat, Barcelona E-08908, Spain.

31 Department of Medical Genetics, Children's Hospital of Eastern Ontario (CHEO) and University of Ottawa, Ottawa, ON, K1H 8L1, Canada.

32 Division of Genetics, University of Texas San Antonio, San Antonio, Texas 78229, USA.

33 Clinical Genetics Department, Great Ormond Street Hospital, London WC1N 3JH, UK.

34 Clinical and Molecular Genetics Unit, UCL Institute of Child Health, London WC1N 1EH, UK.

35 Department of Neuroscience Research, CAMH, Toronto M5T 1R8, Canada.

36 The Fred A. Litwin and Family Centre in Genetic Medicine, University Health Network and Mount Sinai Hospital, Toronto, ON M5T 3L9,Canada.

37 Department of Laboratory Medicine and Pathobiology, University of Toronto, Toronto M5S 1A8, Canada.

38 Department of Molecular Biology, Science School, National University of Rio Cuarto, Córdoba 5800, Argentina.

39 Department of Cell and Molecular Biology, Karolinska Institutet, Stockholm 171 77, Sweden.

40 Department of Obstetrics and Gynecology, Drexel University College of Medicine, Philadelphia PA 19103, USA.

41 Department of Pediatrics, Academic Medical Center, University of Amsterdam, Amsterdam, 1105AZ, The Netherlands.

#Corresponding authors:

Juan Pié, MD PhD, Unit of Clinical Genetics and Functional Genomics, Department of Pharmacology and Physiology, University of Zaragoza, Medical School, c/ Domingo Miral s/n, Zaragoza E-50009, Spain. Tel.: +34 976761677; Fax: +34 976761700; Email: juanpie@unizar.es

David FitzPatrick, MD PhD, MRC Human Genetics Unit, MRC IGMM, University of Edinburgh, Edinburgh EH4 2XU, UK. Tel: 0131 467 8423; Email: davidfitzpatrick@nhs.net

Matthew A. Deardorff, MD PhD, ARC 1002B, 3615 Civic Center Boulevard, Philadelphia, PA 19104, USA. Tel: +1 2155903856; Fax: +1 2674268635; Email: deardorff@email.chop.edu

GRANT SPONSORS:

This work was supported by: The Spanish Ministry of Health - *Fondo de Investigación Sanitaria* (FIS) [Ref.# PI12/01318]; the *Diputación General de Aragón* [Grupo Consolidado B20], European Social Fund ("Construyendo Europa desde Aragón"); Spanish Ministerio de Economía y Competitividad [Ref.# IPT2011-0964-900000 and # SAF2011-13156-E, to P.G-P]; University of Zaragoza [Ref.# PIF-UZ_2009-BIO-02 to M.E.T-R]; the Fundació La Marató de TV3 [Ref.# 1013EXPFMTV3 to E.Q]. MCG-R, BP, MH-M, MET-R, IB-M, FJP and JP are members of "Grupo Clínico Vinculado al CIBER-ER" and ISS-Aragon at the University of Zaragoza Medical School and Hospital Clínico Universitario "Lozano Blesa". FJK is supported by intramural funding from the University of Lübeck [Schwerpunktprogramm, Medizinische Genetik: Von seltenen Varianten zur Krankheitsentstehung] and the German Federal Ministry of Education and Research (B.M.B.F.) under the frame of E-Rare-2 [TARGET-CdLS]. DRF and MA are supported by a program grant from the Medical Research Council (UK) to the MRC Human Genetics Unit. National Institutes of Health Grants [NICHD K08HD055488 to, M.A.D.], [NICHD P01 HD052860 to I.D.K.]; research grants from the USA CdLS Foundation; the Doris Duke Charitable Foundation Grant #:2012059 to M.A.D. Work at Biomol-Informatics is partially supported by Fundación Severo Ochoa and the European Social Fund.

Abstract

Cornelia de Lange syndrome (CdLS) is characterized by facial dysmorphism, growth failure, intellectual disability, limb malformations and multiple organ involvement. Mutations in five genes, encoding subunits of the cohesin complex (*SMC1A*, *SMC3*, *RAD21*) and its regulators (*NIPBL*, *HDAC8*), account for at least 70% of patients with CdLS or CdLS-like phenotypes. To date, only the clinical features from a single CdLS patient with *SMC3* mutation has been published. Here, we report the efforts of an international research and clinical collaboration to provide clinical comparison of sixteen patients with CdLS-like features caused by mutations in *SMC3*. Modelling of the mutation effects on protein structure suggests a dominant-negative effect on the multimeric cohesin complex. When compared to typical CdLS, many *SMC3*-associated phenotypes are also characterized by postnatal microcephaly but with a less distinctive craniofacial appearance, a milder prenatal growth retardation that worsens in childhood, few congenital heart defects and an absence of limb deficiencies. While most mutations are unique, two unrelated affected individuals shared the same mutation but presented with different phenotypes. This work confirms that *de novo SMC3* mutations account for ~1-2% of CdLS-like phenotypes.

Key Words: Cornelia de Lange Syndrome, CdLS, *SMC3*, cohesin complex, CdLS-overlapping phenotypes, CdLS-like

INTRODUCTION

Cornelia de Lange syndrome (CdLS; MIM#s 122470, 300590, 610759, 614701, 300882) is a multisystem developmental diagnosis characterized by distinctive facial dysmorphism, pre- and postnatal growth failure, intellectual disability, limb malformations, hypertrichosis and variable involvement of other organ systems (Kline, et al., 2007). The prevalence is estimated to be up to 1 in 15,000 births (Kline, et al., 2007). Almost all cases are sporadic with *de novo* heterozygous loss-of-function (LOF) mutations in *NIPBL* (MIM# 608667) being the most common genetic finding in typical CdLS (Gillis, et al., 2004; Krantz, et al., 2004; Pie, et al., 2010; Selicorni, et al., 2007; Tonkin, et al., 2004; Wierzba, et al., 2012). A proportion of the "*NIPBL*-negative" cases with typical CdLS have recently been shown to have mosaic *NIPBL* mutations, often undetected in the blood by Sanger based screening (Ansari, et al., 2014; Baquero-Montoya, et al., 2014; Braunholz, et al., 2014; Huisman, et al., 2013). Mutations in four other genes have been reported to account for a smaller proportion of mostly atypical cases; *SMC1A* (MIM# 300040) on chromosome Xp11 (~4-6%), *SMC3* (MIM# 606062) on chromosome 10q25 (<1%), *RAD21* (MIM# 606462) on chromosome 8q24 (<1%), and *HDAC8* (MIM# 300269) on chromosome Xq13 (4%) (Deardorff, et al., 2012a; Deardorff, et al., 2007; Deardorff, et al., 2012b; Kaiser, et al., 2014; Minor, et al., 2014; Musio, et al., 2006).

These five genes encode regulatory or structural components of the evolutionary conserved cohesin complex, which has been implicated in a wide range of functions including sister chromatid cohesion, DNA repair mechanisms, gene regulation and maintenance of genome stability (Revenkova, et al., 2009). Cohesin is a multimeric complex consisting of an SMC1A - SMC3 heterodimer and the two non-SMC subunits, RAD21 and a STAG protein. Each SMC protein folds upon itself so that the N- and C-termini meet to form a globular ATP-binding "head" domain separated from a globular "hinge" domain by

antiparallel coiled-coil segments. SMC3 and SMC1A interact via their respective hinge regions to form a bracelet-shaped heterodimer (Figure 1A). The two ATPase head domains further interact with the N- and C-termini of RAD21, creating a ring structure that is proposed to encircle sister chromatids (Nasmyth and Haering, 2009). NIPBL has been shown to facilitate loading of cohesin onto chromatin, and HDAC8 is involved in recycling of cohesin after its removal from chromatin (Deardorff, et al., 2012a).

To date, only the clinical features of the unique mildly affected CdLS male with *SMC3* mutation has been published (c.1464_1466del, p.(Glu488del)) (Deardorff, et al., 2007; Pie, et al., 2010). Subsequently, a missense *SMC3* mutation has been reported without clinical correlation in one patient within a large cohort of individuals with autism spectrum disorder (c.2413C>T; p.(Arg805Cys)) (Sanders, et al., 2012) and five additional mutations in a cohort of typical and atypical CdLS patients (Ansari, et al., 2014) with the detailed clinical descriptions of these cases documented for the first time in this manuscript.

Here, we report the clinical features of 16 unrelated *SMC3* individuals identified via a large international collaboration and assess the degree of overlap with typical CdLS associated with this gene. Of these, ten are unreported patients with novel or reported mutations in the *SMC3* gene and six individuals have only had molecular information previously published. Furthermore, we mapped all mutations to the known structure of the SMC complex to predict molecular/functional consequences. Our results clearly indicate that *SMC3* mutations result in a CdLS-like phenotype and account for a higher percentage of CdLS and CdLS-like cases than previously appreciated.

MATERIALS AND METHODS

Patient Recruitment

We screened for mutations in *SMC3* an internationally assembled cohort of 674 patients with typical CdLS and overlapping clinical presentations who had no known molecular etiology. All patients were enrolled in this study under institutionally-approved protocols of informed consent at the Odense University Hospital, University Hospital "Lozano Blesa" of Zaragoza, The Children's Hospital of Philadelphia, the UK (Scotland A) MREC Committee, the MET Committee at the Academic Medical Centre of the University of Amsterdam and University of Lübeck. Most individuals in this study were diagnosed by clinical geneticists due to clinical features consistent or overlapping with a CdLS phenotype.

Additional cases of mutations in *SMC3* were referred from clinical colleagues who identified mutations by the use of different molecular analyses such as gene panel or exome-sequencing approaches. Most probands ascertained as CdLS were prescreened and found to be negative for mutations in *NIPBL* and *SMC1A*.

Mutation screening by Sanger sequencing

Genomic DNA was isolated from peripheral blood leukocytes using standard protocols. PCR primers flanking the entire coding region (exons 1-29) and flanking intron sequences of *SMC3* gene were used as previously described (Deardorff, et al., 2007; Pie, et al., 2010). The resulting PCR products were sequenced using the BigDye Terminator 3.1 reagents on an ABI 3730 analyzer. The *SMC3* reference sequence used was NM_005445.3, in which the A of the ATG translation initiation codon was nucleotide 1. Parental genotypes were screened to assess whether the variant was *de novo* or inherited when parental DNA was available.

Ion Torrent Semiconductor gene panel sequencing

Mutation analyses by Ion AmpliSeq-Ion PGM were performed as described previously (Ansari, et al., 2014; Baquero-Montoya, et al., 2014; Braunholz, et al., 2014). Briefly, 10-20 ng of genomic DNA was amplified using custom-designed gene panels (Ion AmpliSeq™, Life Technologies) to cover the coding exons of the known CdLS genes, including approximately 90% of the coding sequence of *SMC3* (NC_000010) and its splice junctions in particular. The DNA library was sequenced on an Ion PGM™ instrument (Life Technologies). Sequence alignment and variant calling were performed as described previously (Ansari, et al., 2014; Baquero-Montoya, et al., 2014; Braunholz, et al., 2014). Possible pathological variants found were assessed by Sanger sequencing.

Exome sequencing

For P7, exomes were captured with the Agilent SureSelect Human All Exon V4+UTR kit (Agilent Technologies, Santa Clara, CA, USA) and sequencing was performed on Illumina HiSeq 2000 machines using standard pair-end read sequencing protocol (Illumina, San Diego, CA, USA). Analysis was as per (Falk, et al., 2014; Li, et al., 2014). Possible pathological variants found were confirmed by Sanger sequencing.

Exome sequencing for P13 was performed clinically at the Baylor Whole Genome Lab. Briefly, exomes were captured using VCRome 2.1 in-solution capture, and sequenced on Illumina HiSeq using 100bp paired end reads. Data analysis and interpretation was as per (Yang, et al., 2013). Possible pathological variants found were confirmed by Sanger sequencing.

Exome sequencing was performed in the affected individual P14 as well as in the non-affected parents. Exomes were enriched in solution with SureSelect^{XT} Target Enrichment System (Agilent Technologies) or SeqCap EZ VCRome 2.0 (Roche NimbleGen) and sequenced as 100 bp paired-end runs on a HISeq2000 or HISeq 2500 system (Illumina).

Mutation modelling

Three dimensional models of the HEAD and HINGE domains of the human SMC1A/SMC3 dimer, for wild-type (wt) and mutant proteins, were generated using homology modelling procedures and the coordinates of the mouse HINGE domain ((Kurze, et al., 2011); PDB code: 2WD5) and yeast HEAD domain -SMC1 homodimer-((Haering, et al., 2004); PDB code: 1W1W) as templates. Model coordinates were built using the SWISS-MODEL server (Guex, et al., 1999; Peitsch, 1996; Schwede, et al., 2003) available at http://swissmodel.expasy.org/, and their structural quality was checked using the analysis programs provided by the same server (Anolea/Gromos/QMEAN4; (Benkert, et al., 2011)) being within the range of those accepted for homology-based structure models. To optimize geometries, models were energy minimized using the GROMOS 43B1 force field implemented in DeepView (http://spdbv.vital-it.ch/), using 500 steps of steepest descent minimization followed by 500 steps of conjugate-gradient minimization. Coiled-coil predictions were calculated using COILS server with a window of 28 residues (http://www.ch.embnet.org; (Lupas, et al., 1991)). Multiple sequence alignment of proteins from the SMC3 family was generated using TCOFFEE (http://www.tcoffee.org/) (Notredame, et al., 2000). Functional prediction for nonsynonymous or *indel* variants were obtained using PolyPhen-2 (http://genetics.bwh.harvard.edu/pph/) (Adzhubei, et al., 2010), SIFT (http://sift.jcvi.org/) (Ng and Henikoff, 2001), PROVEAN (http://provean.jcvi.org/index.php) (Choi, et al., 2012), Mutation Taster (http://www.mutationtaster.org/) (Schwarz, et al., 2010) and the Biomol-Informatics exome analysis system (http://results.genoma4u.com/).

Reference sequences

SMC3 accession numbers used include NM_005445.3 (mRNA) and NP_005436.1 (RefSeq protein). SMC3 protein sequences (UniProt) for human (Q9UQE7), *Pongo abelii* (Q5R4K5), *Rattus norvegicus* (P97690), *Mus musculus* (Q9CW03), *Bos taurus* (O97594), *Xenopus laevis* (O93309), *Saccharomyces cerevisiae* (P47037) and *Plasmodium falciparum* (Q8I1U7).

RESULTS

Intragenic mutations in *SMC3* in a large cohort of patients

Sequence analysis of patients with CdLS and CdLS-like phenotypes for mutations in *SMC3* identified 15 different intragenic mutations in 16 unrelated individuals. Six of 15 mutations have been previously described (Ansari, et al., 2014; Deardorff, et al., 2007), therefore here we report ten individuals with nine new mutations (Table 1). Seven of the ten individuals had both parents available for testing and in each case these mutations occurred *de novo*. One in-frame *de novo* deletion of three nucleotides (c.1464_1466del; p.(Glu488del)) was also identified in the first reported individual (Deardorff, et al., 2007). Three of these are caused by in-frame mutations that retain the open reading frame (one duplication and two deletions of one or two residues) and seven mutations were missense (Table 1, Figure 1, Supp. Figure S1). All variants have been added to a publicly accessible LOVD database (http://www.LOVD.nl/SMC3). None of these mutations were seen in 100 control alleles or publicly available repositories of sequence variation.

In silico Analyses of Missense and In-frame Mutations

The predicted functional effect of each mutation is summarized in Table 1 and the cross-species alignment showing the degree of evolutionary conservation of the residues involved in the missense and in-frame variants is shown in Figure 1B. Figure 1A indicates the location of each variant with regard to the known functional domains of SMC3.

Gly655 localizes to the SMC3 hinge domain and the substitution with aspartic acid is predicted to structurally destabilize the domain core. Thr235, Arg236, Arg839 and His917 localize to the N- and the C-terminal coiled-coil structures respectively and their deletion or substitution is predicted to displace the two antiparallel helices (Supp. Figure S2).

In the globular ATP-binding head domain Phe47 is located in the alpha helices. Gln1147 is within the functional motif D-loop, close to both the gamma-phosphate of ATP and the interface between the head domains of SMC3 and SMC1A. Substitution of this polar residue Gln1147 by a negatively charged glutamate residue could alter the ATPase activity of the active site of the heterodimer as well as alter the essential interaction between SMC1A and SMC3 at the head interface (Figure 1C). Thr1215 is located in an apparently non-structured region close to the C-terminus and the effect of the isoleucine substitution at this residue is not clear, although it cannot be excluded a putative role in the SMC3-RAD21 interaction.

Clinical features of individuals with *SMC3* mutations

The clinical features in the 16 individuals with mutations involving *SMC3* are summarized in Table 2 and Supp. Table S1. Figure 2 shows facial and limb findings. Many patients have CdLS-like craniofacial features including brachycephaly (73%, (11/15)), low anterior hairline (50%, (7/14)), arched eyebrows (93%, (14/15)), synophrys (73%, (11/15)), long eyelashes (94%, (15/16)), ptosis (27%, (4/15)), depressed nasal bridge (47%, (7/15)), anteverted nostrils (57%, (8/14)), long philtrum (67%, (10/15)), thin upper lip vermilion (81%, (13/16)), downturned corners of the mouth (60%, (9/15)), high palate (45%, (5/11)), dental anomalies (38%, (5/13)) and micrognathia (40%, (6/15)) (Table 2). Although often long, the philtrum is typically not smooth in these individuals and only one patient had a cleft palate. Major limb malformations were not observed. Intellectual disability was a prominent

feature, although behavioral problems were not frequently reported and many were described as having friendly personalities.

DISCUSSION

To further characterize the nature of *SMC3* gene mutations and the range of resulting clinical features, we utilized an international cooperative research and clinical effort coupled with standard sequencing and Next Generation Sequencing (NGS) strategies. This enabled us to identify 16 probands with 15 different intragenic mutations in *SMC3*, including the previously reported individuals (Ansari, et al., 2014; Deardorff, et al., 2007). Based on these numbers, we could estimate that individuals with *SMC3* mutations comprise approximately ~1-2% of patients with features suggestive of CdLS or overlapping phenotypes.

Typically, SMC3 mutations identified in these CdLS-like patients are missense or in-frame insertions or deletions, similar to CdLS-causing mutations found in the SMC1A protein (Deardorff, et al., 2007; Gimigliano, et al., 2012; Liu, et al., 2009; Mannini, et al., 2010; Musio, et al., 2006; Revenkova, et al., 2009). Nine of fifteen *SMC3* mutations identified predict amino acid alterations in the coiled-coil domain (Figure 1A, Supp. Figure S2). In the *SMC1A*-associated CdLS-like disorder, 69% of the disease causing mutations (all missense/in-frame) are also identified in the cognate coiled-coil domain (Gervasini, et al., 2013). The similarity of structure and function of the two SMC proteins, as well as the mutation spectrum, suggests that SMC3 missense/in-frame mutations may act via a dominant negative effect as has been previously suggested for other mutations in the SMC1A protein (Deardorff, et al., 2007; Mannini, et al., 2013).

Several craniofacial features commonly seen in typical CdLS (>80% of the CdLS patients, reviewed in Kline et al., (2007)) are absent or infrequent in this *SMC3* cohort. For example, while the eyebrows may be highly arched and the eyelashes long, synophrys is often absent or subtle. The nasal bridge is less frequently depressed, and the nasal tip is often

broad or bulbous, unlike the small triangular shaped nose in typical CdLS. Furthermore, the nostrils are not typically anteverted in this cohort, as is seen in CdLS caused by mutations in *NIPBL* (Rohatgi, et al., 2010). The philtrum may be long but is often well formed in this cohort and infrequently flat, as in typical CdLS. Thin upper lips vermilion are observed but the downturned mouth often seen in typical CdLS is uncommon.

Congenital heart defects (CHD) are common in CdLS (13-70%) with isolated defects seen in 86% (PS, VSD and ASD) and multiple defects in 14% (Selicorni, et al., 2009). Consistent with this, *SMC3* probands appear to have cardiac malformations (56%). For example a number of individuals presented with some degree of pulmonic stenosis, one of the most frequent findings in CdLS (Chatfield, et al., 2012; Selicorni, et al., 2009). In addition, two individuals showed with aortic stenosis with bicuspid aortic valve and one with Tetralogy of Fallot. While this frequency and severity of cardiac anomalies can be seen in CdLS caused by mutations in *NIPBL*, they are infrequent in patients with *SMC1A* mutations (Chatfield, et al., 2012), suggesting that SMC3 is important for the normal development of the heart.

Clinical comparison between two individuals (P6 and P7) which carried the same deletion of three nucleotides, c.1464_1466del (Deardorff, et al., 2007), showed a similar craniofacial appearance during their newborn period, even though this evolved with time differently (Figure 2). In addition, these patients had markedly different cognitive and developmental impairment and musculoskeletal involvement, with one working as an adult and the other nonverbal and nonambulatory (Figure 2, Supp. Table S1). This emphasizes that, phenotypes associated with the identical mutations are likely variable, which indicates the influence of other factors in the manifestation of CdLS, as it has been reported for other CdLS genes (Gillis, et al., 2004; Pie, et al., 2010).

In general, *SMC3* probands present with a mild to severe phenotype that differs from typical CdLS that is frequently caused by *NIPBL* mutations. Clinical features of patients with *SMC3*-mutations are more similar to those of patients with mutations in *SMC1A* (Borck, et al., 2007; Deardorff, et al., 2007; Gervasini, et al., 2013; Liu, et al., 2009; Mannini, et al., 2010; Musio, et al., 2006). Thus the craniofacial phenotype of patients with mutations in *SMC1A* and *SMC3* genes do show overlapping features such as broader, fuller less arched eyebrows and a more prominent nasal bridge (Deardorff, et al., 2007; Rohatgi, et al., 2010). In addition, both groups of patients seem to have less growth restriction than typically seen in patients with mutations in *NIPBL*. However, this is fairly difficult to generalize, given the variability in the range of severity and the small number of patients with *SMC3* mutations.

Interestingly, several individuals from this cohort were ascertained independently of a diagnosis of CdLS (e.g. P7 and P13). Although they have some CdLS-overlapping features, they were felt to be divergent enough from CdLS to pursue exome-based testing rather than CdLS gene panel testing. In addition, an *SMC3* mutation has been reported in a patient with autism spectrum disorder, but to our knowledge has no obvious CdLS phenotype (Sanders, et al., 2012). These findings are consistent with an emerging range of clinical phenotypes caused by mutations in the cohesin complex, as is supported by the finding of an *HDAC8* mutation in a family with Wilson-Turner syndrome (intellectual disability, truncal obesity, hypogonadism and distinctive facial features) (Harakalova, et al., 2012) and an *SGOL1* mutation in 17 patients with CAID syndrome (Chronic Atrial and Intestinal Dysrhythmia) (Chetaille, et al., 2014). These findings indicate that the range of clinical phenotypes caused by alterations in cohesin may be significantly broader than previously appreciated.

In conclusion, we report a series of *SMC3* mutations that provide a significant advance in our understanding of the clinical and molecular basis of human disorders of cohesin. Although this cohort represents ~1-2% of individuals with CdLS-like phenotypes, they provide us novel insight into the understanding of cohesin in health and disease.

ACKNOWLEDGMENTS

We sincerely thank the patients' families for participating in this study. DRF and MA would like to thank the CdLS Foundation of UK and Ireland for their long-term help and support. MAD and IDK are indebted to the USA Cornelia de Lange Syndrome Foundation for their continued support. JP and FJR would like to thank the Spanish Cornelia de Lange Syndrome Foundation.

AUTHOR CONTRIBUTIONS

M.A.D., A.S., F.J.R., L.G.J., I.D.K., R.C.M.H., F.J.K., D.R.F and J.P. initiated the human studies.

M.A.D., M.A., C.A.T., C.B-M., L.B.O., E.K.B., I.B-M., D.C., N.S.C., C.J.C, R.F., A.F., J.G., C.G., G.G-K., R.J.H., M.K., E.K., T.K., A.D.K., E.K., L.L., M.M., J.D.P., A.P., J.R., E.R., A.S., R.H.S., J.S., K.A.W., L.W., F.J.R., L.G.J., I.D.K., S.D., R.C.M.H., D.R.F. and J.P. identified, characterized and provided patient data and samples.

M.C.G-R., M.A.D., M.A., I.P., B.P., M.H-M., M.E.T-R, J-J.W., S.L-B., D.B., Y.G., H.H., M.K., B.J.K., Y.R.L., X.L., J.P., E.Q., J.Z., C.H.C., L.S., A.S., F.J.K., D.R.F., and J.P. performed array analysis, mutation screening and/or exome analysis.

M.C.G-R., M.A.D., I.M.A., P.G.-P., and J.P. performed structural analysis.

M.C.G-R., M.A.D., M.A., F.J.K., D.R.F. and J.P. drafted the manuscript.

All authors analyzed data, discussed the results and were provided opportunity to comment on the manuscript.

CONFLICT OF INTEREST STATEMENT

No authors have declared conflicts of interest regarding this work.

REFERENCES

Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, Kondrashov AS, Sunyaev SR. 2010. A method and server for predicting damaging missense mutations. Nat Methods 7:248-249.

Ansari M, Poke G, Ferry Q, Williamson K, Aldridge R, Meynert AM, Bengani H, Chan CY, Kayserili H, Avci S, Hennekam RC, Lampe AK et al. 2014. Genetic heterogeneity in Cornelia de Lange syndrome (CdLS) and CdLS-like phenotypes with observed and predicted levels of mosaicism. J Med Genet 51(10):659-668.

Baquero-Montoya C, Gil-Rodriguez MC, Braunholz D, Teresa-Rodrigo ME, Obieglo C, Gener B, Schwarzmayr T, Strom TM, Gomez-Puertas P, Puisac B, Gillessen-Kaesbach G, Musio A et al. 2014. Somatic mosaicism in a Cornelia de Lange syndrome patient with *NIPBL* mutation identified by different next generation sequencing approaches. Clin Genet 86:595-597.

Benkert P, Biasini M, Schwede T. 2011. Toward the estimation of the absolute quality of individual protein structure models. Bioinformatics 27:343-350.

Borck G, Zarhrate M, Bonnefont JP, Munnich A, Cormier-Daire V, Colleaux L. 2007. Incidence and clinical features of X-linked Cornelia de Lange syndrome due to *SMC1L1* mutations. Hum Mutat 28:205-206.

Braunholz D, Obieglo C, Parenti I, Pozojevic J, Eckhold J, Reiz B, Braenne I, Wendt KS, Watrin E, Vodopiutz J, Rieder H, Gillessen-Kaesbach G et al. 2014. Hidden

Mutations in CdLS - Limitations of Sanger Sequencing in Molecular Diagnostics. Hum Mutat 36(1):26-29.

Chatfield KC, Schrier SA, Li J, Clark D, Kaur M, Kline AD, Deardorff MA, Jackson LS, Goldmuntz E, Krantz ID. 2012. Congenital heart disease in Cornelia de Lange syndrome: phenotype and genotype analysis. Am J Med Genet A 158A:2499-2505.

Chetaille P, Preuss C, Burkhard S, Cote JM, Houde C, Castilloux J, Piche J, Gosset N, Leclerc S, Wunnemann F, Thibeault M, Gagnon C et al. 2014. Mutations in *SGOL1* cause a novel cohesinopathy affecting heart and gut rhythm. Nat Genet 46:1245-1249.

Choi Y, Sims GE, Murphy S, Miller JR, Chan AP. 2012. Predicting the functional effect of amino acid substitutions and indels. PloS one 7:e46688.

Deardorff MA, Bando M, Nakato R, Watrin E, Itoh T, Minamino M, Saitoh K, Komata M, Katou Y, Clark D, Cole KE, De Baere E et al. 2012a. HDAC8 mutations in Cornelia de Lange syndrome affect the cohesin acetylation cycle. Nature 489:313-317.

Deardorff MA, Kaur M, Yaeger D, Rampuria A, Korolev S, Pie J, Gil-Rodriguez C, Arnedo M, Loeys B, Kline AD, Wilson M, Lillquist K et al. 2007. Mutations in cohesin complex members SMC3 and SMC1A cause a mild variant of cornelia de Lange syndrome with predominant mental retardation. Am J Hum Genet 80:485-494.

Deardorff MA, Wilde JJ, Albrecht M, Dickinson E, Tennstedt S, Braunholz D, Monnich M, Yan Y, Xu W, Gil-Rodriguez MC, Clark D, Hakonarson H et al. 2012b. *RAD21* mutations cause a human cohesinopathy. Am J Hum Genet 90:1014-1027.

Falk MJ, Li D, Gai X, McCormick E, Place E, Lasorsa FM, Otieno FG, Hou C, Kim CE, Abdel-Magid N, Vazquez L, Mentch FD, Vazquez L, Mentch FD et al. 2014. AGC1 Deficiency Causes Infantile Epilepsy, Abnormal Myelination, and Reduced N-Acetylaspartate. JIMD Rep 14:119.

Gervasini C, Russo S, Cereda A, Parenti I, Masciadri M, Azzollini J, Melis D, Aravena T, Doray B, Ferrarini A, Garavelli L, Selicorni A et al. 2013. Cornelia de Lange individuals with new and recurrent *SMC1A* mutations enhance delineation of mutation repertoire and phenotypic spectrum. Am J Med Genet A 161A:2909-2919.

Gillis LA, McCallum J, Kaur M, DeScipio C, Yaeger D, Mariani A, Kline AD, Li HH, Devoto M, Jackson LG, Krantz ID. 2004. *NIPBL* mutational analysis in 120 individuals with Cornelia de Lange syndrome and evaluation of genotype-phenotype correlations. Am J Med Genet 75:610-623.

Gimigliano A, Mannini L, Bianchi L, Puglia M, Deardorff MA, Menga S, Krantz ID, Musio A, Bini L. 2012. Proteomic Profile Identifies Dysregulated Pathways in Cornelia de Lange Syndrome Cells with Distinct Mutations in SMC1A and SMC3 Genes. J Proteome Res 11:6111-6123.

Guex N, Diemand A, Peitsch MC. 1999. Protein modelling for all. Trends Biochem Sci 24:364-367.

Haering CH, Schoffnegger D, Nishino T, Helmhart W, Nasmyth K, Lowe J. 2004. Structure and stability of cohesin's Smc1-kleisin interaction. Mol Cell 15:951-964.

Harakalova M, van den Boogaard MJ, Sinke R, van Lieshout S, van Tuil MC, Duran K, Renkens I, Terhal PA, de Kovel C, Nijman IJ, van Haelst M, Knoers NV et al. 2012. X-exome sequencing identifies a *HDAC8* variant in a large pedigree with X-linked intellectual disability, truncal obesity, gynaecomastia, hypogonadism and unusual face. J Med Genet 49:539-543.

Huisman SA, Redeker EJ, Maas SM, Mannens MM, Hennekam RC. 2013. High rate of mosaicism in individuals with Cornelia de Lange syndrome. J Med Genet 50:339-344.

Jackson L, Kline AD, Barr MA, Koch S. 1993. de Lange syndrome: a clinical review of 310 individuals. Am J Med Genet 47:940-946.

Kaga K, Tamai F, Kitazumi E, Kodama K. 1995. Auditory brainstem responses in children with Cornelia de Lange syndrome. Int J Pediatr Otorhi 31:137-146.

Kaiser FJ, Ansari M, Braunholz D, Gil-Rodriguez MC, Decroos C, Wilde JJ, Fincher CT, Kaur M, Bando M, Amor DJ, Atwal PS, Bahlo M et al. 2014. Loss of Function *HDAC8* Mutations Cause a Phenotypic Spectrum of Cornelia de Lange Syndrome-like Features, Ocular Hypertelorism, Large Fontanelle and X-linked Inheritance. Hum Mol Genet 23, 2888-2900.

Kline AD, Krantz ID, Sommer A, Kliewer M, Jackson LG, FitzPatrick DR, Levin AV, Selicorni A. 2007b. Cornelia de Lange syndrome: clinical review, diagnostic and scoring systems, and anticipatory guidance. Am J Med Genet A 143A:1287-1296.

Krantz ID, McCallum J, DeScipio C, Kaur M, Gillis LA, Yaeger D, Jukofsky L, Wasserman N, Bottani A, Morris CA, Nowaczyk MJ, Toriello H et al. 2004. Cornelia de Lange syndrome is caused by mutations in *NIPBL*, the human homolog of Drosophila melanogaster Nipped-B. Nat Genet 36:631-635.

Kurze A, Michie KA, Dixon SE, Mishra A, Itoh T, Khalid S, Strmecki L, Shirahige K, Haering CH, Lowe J, Nasmyth K. 2011. A positively charged channel within the Smc1/Smc3 hinge required for sister chromatid cohesion. Embo J 30:364-378.

Li Q, Brodsky JL, Conlin LK, Pawel B, Glatz AC, Gafni RI, Schurgers L, Uitto J, Hakonarson H, Deardorff MA, Levine MA. 2014. Mutations in the *ABCC6* gene as a cause of generalized arterial calcification of infancy: genotypic overlap with pseudoxanthoma elasticum. J Invest Dermatol 134:658-665.

Liu J, Feldman R, Zhang Z, Deardorff MA, Haverfield EV, Kaur M, Li JR, Clark D, Kline AD, Waggoner DJ, Das S, Jackson LG et al. 2009. *SMC1A* expression and mechanism of pathogenicity in probands with X-Linked Cornelia de Lange syndrome. Hum Mutat 30:1535-1542.

Lupas A, Van Dyke M, Stock J. 1991. Predicting coiled coils from protein sequences. Science 252:1162-1164.

Luzzani S, Macchini F, Valade A, Milani D, Selicorni A. 2003. Gastroesophageal reflux and Cornelia de Lange syndrome: typical and atypical symptoms. Am J Med Genet A 119A:283-287.

Mannini L, Cucco F, Quarantotti V, Krantz ID, Musio A. 2013. Mutation spectrum and genotype-phenotype correlation in Cornelia de Lange syndrome. Hum Mutat 34:1589-1596.

Mannini L, Liu J, Krantz ID, Musio A. 2010. Spectrum and consequences of *SMC1A* mutations: the unexpected involvement of a core component of cohesin in human disease. Hum Mutat 31:5-10.

Minor A, Shinawi M, Hogue JS, Vineyard M, Hamlin DR, Tan C, Donato K, Wysinger L, Botes S, Das S, Del Gaudio D. 2014. Two novel *RAD21* mutations in patients with mild Cornelia de Lange syndrome-like presentation and report of the first familial case. Gene 537:279-284.

Musio A, Selicorni A, Focarelli ML, Gervasini C, Milani D, Russo S, Vezzoni P, Larizza L. 2006. X-linked Cornelia de Lange syndrome owing to *SMC1L1* mutations. Nat Genet 38:528-530.

Nallasamy S, Kherani F, Yaeger D, McCallum J, Kaur M, Devoto M, Jackson LG, Krantz ID, Young TL. 2006. Ophthalmologic findings in Cornelia de Lange syndrome: a genotype-phenotype correlation study. Arch Ophthalmol-Chic 124:552-557.

Nasmyth K, Haering CH. 2009. Cohesin: its roles and mechanisms. Annu Rev Genet 43:525-558.

Ng PC, Henikoff S. 2001. Predicting deleterious amino acid substitutions. Genome Res 11:863-874.

Notredame C, Higgins DG, Heringa J. 2000. T-Coffee: A novel method for fast and accurate multiple sequence alignment. J Mol Biol 302:205-217.

Peitsch MC. 1996. ProMod and Swiss-Model: Internet-based tools for automated comparative protein modelling. Biochem Soc Trans 24:274-279.

Pie J, Gil-Rodriguez MC, Ciero M, Lopez-Vinas E, Ribate MP, Arnedo M, Deardorff MA, Puisac B, Legarreta J, de Karam JC, Rubio E, Bueno I et al. 2010. Mutations and variants in the cohesion factor genes *NIPBL*, *SMC1A*, and *SMC3* in a cohort of 30 unrelated patients with Cornelia de Lange syndrome. Am J Med Genet A 152A:924-929.

Revenkova E, Focarelli ML, Susani L, Paulis M, Bassi MT, Mannini L, Frattini A, Delia D, Krantz I, Vezzoni P, Jessberger R, Musio A. 2009. Cornelia de Lange syndrome mutations in SMC1A or SMC3 affect binding to DNA. Hum Mol Genet 18:418-427.

Rohatgi S, Clark D, Kline AD, Jackson LG, Pie J, Siu V, Ramos FJ, Krantz ID, Deardorff MA. 2010. Facial diagnosis of mild and variant CdLS: Insights from a dysmorphologist survey. Am J Med Genet A 152A:1641-1653.

Sanders SJ, Murtha MT, Gupta AR, Murdoch JD, Raubeson MJ, Willsey AJ, Ercan-Sencicek AG, DiLullo NM, Parikshak NN, Stein JL, Walker MF, Ober GT et al. 2012. *De novo*

mutations revealed by whole-exome sequencing are strongly associated with autism. Nature 485:237-241.

Schwarz JM, Rodelsperger C, Schuelke M, Seelow D. 2010. MutationTaster evaluates disease-causing potential of sequence alterations. Nat Methods 7:575-576.

Schwede T, Kopp J, Guex N, Peitsch MC. 2003. SWISS-MODEL: An automated protein homology-modeling server. Nucleic Acids Res 31:3381-3385.

Selicorni A, Colli AM, Passarini A, Milani D, Cereda A, Cerutti M, Maitz S, Alloni V, Salvini L, Galli MA, Ghiglia S, Salice P et al. 2009. Analysis of congenital heart defects in 87 consecutive patients with Brachmann-de Lange syndrome. Am J Med Genet A 149A:1268-1272.

Selicorni A, Russo S, Gervasini C, Castronovo P, Milani D, Cavalleri F, Bentivegna A, Masciadri M, Domi A, Divizia MT, Sforzini C, Tarantino E et al. 2007. Clinical score of 62 Italian patients with Cornelia de Lange syndrome and correlations with the presence and type of *NIPBL* mutation. Clin Genet 72:98-108.

Tonkin ET, Wang TJ, Lisgo S, Bamshad MJ, Strachan T. 2004. *NIPBL*, encoding a homolog of fungal Scc2-type sister chromatid cohesion proteins and fly Nipped-B, is mutated in Cornelia de Lange syndrome. Nat Genet 36:636-641.

Wierzba J, Gil-Rodriguez MC, Polucha A, Puisac B, Arnedo M, Teresa-Rodrigo ME, Winnicka D, Hegardt FG, Ramos FJ, Limon J, Pié J. 2012. Cornelia de Lange syndrome with *NIPBL* mutation and mosaic Turner syndrome in the same individual. BMC Med Genet13:43.

Wygnanski-Jaffe T, Shin J, Perruzza E, Abdolell M, Jackson LG, Levin AV. 2005. Ophthalmologic findings in the Cornelia de Lange Syndrome. J AAPOS 9:407-415.

Yang Y, Muzny DM, Reid JG, Bainbridge MN, Willis A, Ward PA, Braxton A, Beuten J, Xia F, Niu Z, Hardison M, Person R et al. 2013. Clinical whole-exome sequencing for the diagnosis of mendelian disorders. N Engl J Med 369:1502-1511.

Figure 1:

(A) Schematic representation of the SMC1A-SMC3 heterodimer of the cohesin complex and the locations of SMC3 mutations in coiled-coil, hinge and head domains. Position of mutated residues in CdLS patients, described in the text, are indicated by red dots.

(B) Multiple sequence alignment of several proteins homologous to SMC3 in the areas surrounding mutated residues Phe47, Thr235, Arg236, Glu287, Lys400_Ser401, Glu488, Gly655, Gly666, Leu832_Asn833, Arg 839, His917, Gln1147 and Thr1215. Represented sequences are: *Homo sapiens* (SMC3_HUMAN), *Pongo abelii* (SMC3_PONGO), *Rattus norvegicus* (SMC3_RAT), *Mus musculus* (SMC3_MOUSE), *Bos taurus* (SMC3_BOVIN), *Xenopus laevis* (SMC3_XENOPUS), *Saccharomyces cerevisiae* (SMC3_YEAST) and *Plasmodium falciparum* (SMC3_PLASMOD). Residues are colored according to conservation.

(C) Left: Predicted structure of SMC3 head domain in the neighborhood of the ATPase active centre. Interaction surface of SMC3 to SMC1A has been colored according to electrostatic characteristics (red: negative; blue: positive; white: neutral). Positions of ATP, Mg^{++} atom and Residue Q1147 are indicated. Right: Predicted surface for Q1147E mutant. The negatively charged patch that appeared close to gamma phosphate of ATP and in the interaction surface to SMC1A is highlighted.

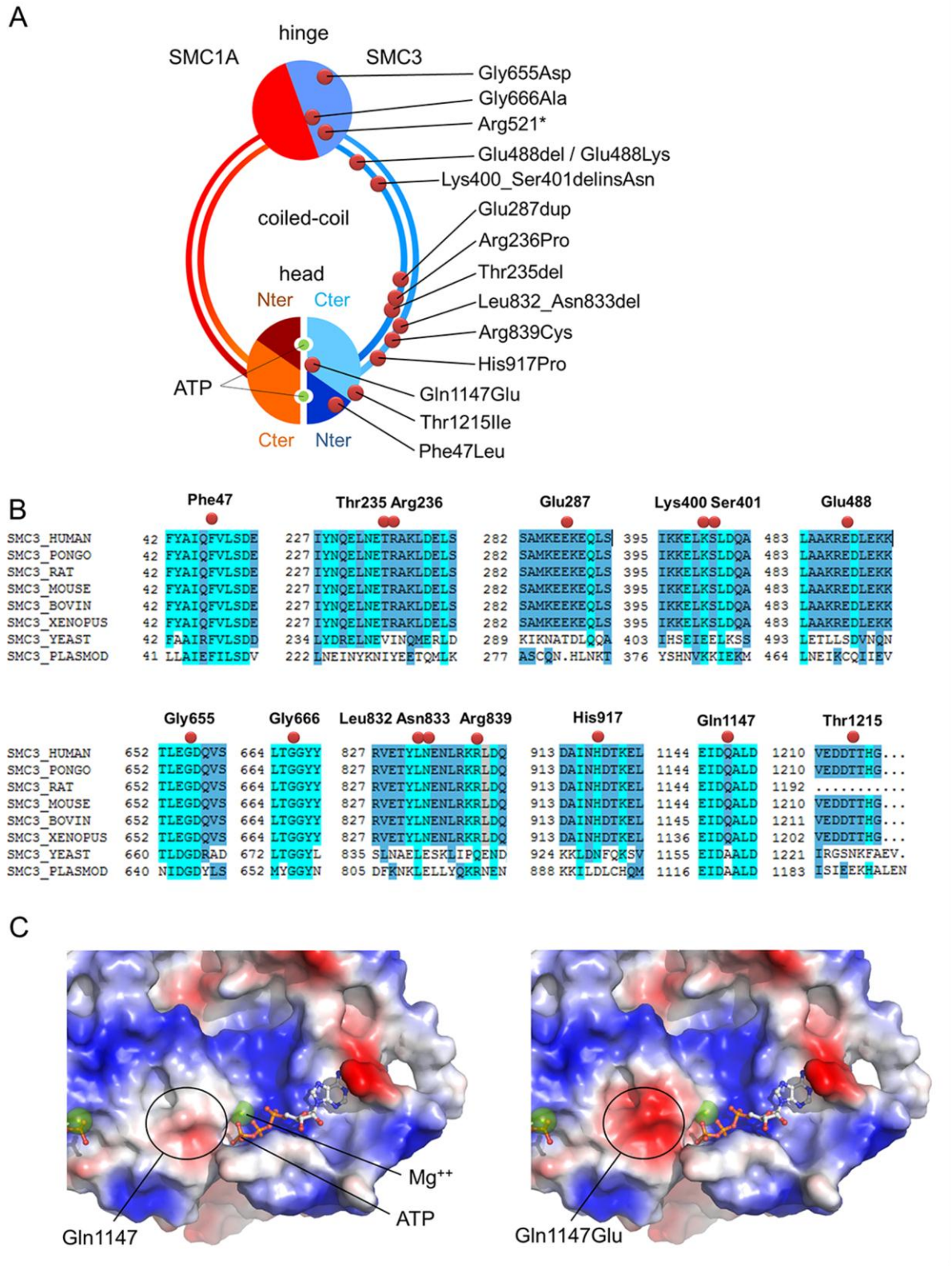

Figure 2: Clinical Photographs of individuals with *SMC3* mutations.

Photos for individual patients are grouped ((i-iv) frontal view at different ages, hands and feet, when they are available) and labelled with corresponding identifier, mutation and sex; ♂=male, ♀=female.

Table 1. *SMC3* mutations identified

ID	MUTATION	*De novo*	EXON	PREDICTED PROTEIN CHANGE	PROTEIN DOMAIN	*IN SILICO* FUNCTIONAL PREDICTION		Reference
						SIFT/Provean	PolyPhen-2	
1	c.139T>C	n/a	4	p.(Phe47Leu)	Head	Damaging: 0.01	Probably damaging: 1	(Ansari, et al., 2014)
2	c.[=/703_705del] mosaic	+	9	p.[=/Thr235del]	Coiled coil	Deleterious: -10.683	n/a	(Ansari, et al., 2014)
3	c.707G>C	+	9	p.(Arg236Pro)	Coiled coil	Damaging: 0.04	Probably damaging: 0.998	This study
4	c.859_861dup	n/a	11	p.(Glu287dup)	Coiled coil	Deleterious: -9.076	n/a	This study
5	c.1200_1202delGTC	n/a	13	p.(Lys400_Ser401delinsAsn)	Coiled coil	Deleterious: -13.196	n/a	(Ansari, et al., 2014)
6	c.1464_1466delAGA	+	15	p.(Glu488del)	Coiled coil	Deleterious: -8.108	n/a	(Deardorff, et al., 2007)
7	c.1464_1466delAGA	+	15	p.(Glu488del)	Coiled coil	Deleterious: -8.108	n/a	This study
8	c.1462G>A	+	15	p.(Glu488Lys)	Coiled coil	Tolerated: 0.2	Possibly damaging: 0.851	This study
9	c.1561C>T	n/a	16	p.(Arg521*)	Hinge	n/a	n/a	(Ansari, et al., 2014)
10	c.1964G>A	+	19	p.(Gly655Asp)	Hinge	Damaging: 0	Probably damaging: 1	This study
11	c.1997G>C	+	19	p.(Gly666Ala)	Hinge	Damaging: 0.01	Probably damaging: 1	This study
12	c.2494_2499del	+	22	p.(Leu832_Asn833del)	Coiled coil	Deleterious: -11.538	n/a	This study
13	c.2515C>T	n/a	22	p.(Arg839Cys)	Coiled coil	Damaging: 0.01	Probably damaging: 1	This study
14	c.2750A>C	+	24	p.(His917Pro)	Coiled coil	Tolerated: 0.08	Possibly damaging: 0.820	This study
15	c.3439C>G	+	27	p.(Gln1147Glu)	Head	Damaging: 0	Probably damaging: 0.998	(Ansari, et al., 2014)
16	c.3644C>T	n/a	29	p.(Thr1215Ile)	Head	Damaging: 0	Probably damaging: 1	This study

The on-line predicted functional effect of nonsynonymous or *indel* variants have been determined by SIFT or Provean programs respectively. The *SMC3* reference sequence used was NM_005445.3, in which the A of the ATG translation initiation codon was nucleotide 1.

Table 2. Frequency of clinical features in individuals with *SMC3* mutations compared with classical CdLS

Category		Feature	Frequency in Classical CdLS *	*SMC3* Percent (# observed/ # assessed)	*SMC3* Details (number of patients with finding)
Craniofacial findings	Head	Brachycephaly		73% (11/15)	
		Low anterior hairline	92%	50% (7/14)	
		Skull			Congenital (5) and/or postnatal (12) microcephaly, plagiocephaly (1), flat facies (1), facial asymmetry (1), frontal bossing (1), posterior hair whorl on left side (1), sparse temporal hair (1), delayed closure of anterior fontanelle (1).
	Eyes	Arched eyebrows		93% (14/15)	
		Synophrys	99%	73% (11/15)	
		Thick eyebrows		69% (9/13)	
		Long eyelashes	99%	94% (15/16)	
		Hooding of lids		15% (2/13)	
	Nose	Depressed nasal bridge	83%	47% (7/15)	
		Anteverted nostrils	88%	57% (8/14)	
		Long and/or featureless philtrum	94%	67% (10/15)	
		Broad/bulbous nasal tip		86% (12/14)	
	Mouth	Thin upper lip vermilion	94%	81% (13/16)	
		Downturned corners of mouth	94%	60% (9/15)	
		Palate - high	86%	45% (5/11)	
		Palate - cleft	20%	7% (1/14)	
		Small/Widely spaced teeth	86%	22% (2/9)	
		Dental anomalies		38% (5/13)	Delayed with irregular eruption (1), not secondary (1), dysmorphic teeth (1), pegged incisors (1).
		Micrognathia/retrognathia	84%	40% (6/15)	
	Neck	Short neck		46% (6/13)	
	Other facial				Lateral extension eyebrows (1), almond shaped (1), deep-set eyes (1). Prominent supraorbital ridges (1). Low-set ears (6), posteriorly rotated ears (3), large ears (2). Small mouth (1), prognathism (2). Low posterior hairline (2), webbed neck (1).
Musculos-	Hands	Small hands	93%	79% (11/14)	
		Proximally set thumbs	72%	75% (12/16)	
		Short first metacarpal		79% (11/14)	
		Clinodactyly 5th finger	74%	64% (9/14)	
		Short 5th finger		69% (9/13)	

		Single palmar crease	51%	36% (5/14)	
	Feet	Small feet	93%	85% (11/13)	
		Syndactyly of toes	86%	29% (4/14)	
	Arms	Restriction of elbow movements	64%	45% (5/11)	
	Other skeletal				Tapered fingers (2), syndactyly 2^{nd}-3^{rd} (1) and 3^{rd}-4^{th} (1) fingers, hyoplastic distal phalanges (1). Joint laxity with flexible fingers (1). Madelung deformity (1). Tapered 1^{st} toes, short 4^{th} metatarsal (1), gap between 1^{st}-2^{nd} toes (1), pes cavus (2) and metatarsus adductus (1). Pectus excavatum (1), short sternum (1), scoliosis (1), cleft and butterfly vertebrae (1), Klippel-Feil (1). Dysplastic hip (1). Sacral dimple (1). Leg length discrepancy (1). Delayed skeletal maturity (1) and decreased muscle bulk (1). Extension defect of Achilles tendon (1). Bunions (1).
Cardiac system		Cardiac defects	13-70%	56% (9/16)	PDA+ASD (1), PS+VSD (1), ASD+ AS+BAV (1), ASD (PFO) (1), pulmonary artery dysplasia (1), PS+AS+BAV (1), PPS (1), ASD+VSD (1), TOF+PS+main pulmonary artery hypoplasia (1).
Gastrointestinal system		GERD	65%	67% (10/15)	
		Feeding problems in infancy		79% (11/14)	
		other gastrointestinal			Hiatal hernia (1), pyloric stenosis (1), malrotation (1).
Genitourinary system		Genitourinary defects	40% - 57%	40% (6/15)	Amenorrhea (1), cryptorchidism (2), hypoplastic genitalia (1), inguinal hernia (2). Bilateral megaureter (1), VUR (2), small kidneys (1).
ENT		Hearing loss	60%	54% (7/13)	
Ophthalmic system		Ptosis	44 - 46%	27% (4/15)	
		Myopia	57 - 58%	45% (5/11)	
		Lacrimal duct obstruction		33% (4/12)	
		Other			Upward deviation of gaze + amblyopia (1), astigmatism (1), exotropia (1), esotropia + cortical visual impairment + sensitivity to light (photophobic) (1), exotropia + astigmatism (1), microphthalmia, Peter's anomaly, congenital cataracts, and glaucoma (1).
Skin		Cutis marmorata	60%	31% (4/13)	
		Hirsutism	78%	93% (14/15)	
		Nevus flameus		8% (1/12)	

	Other skin			Hemangioma (1), abnormal dermatoglyphics (1).
Neurologic findings and Cognitive profile	CNS anomalies		36% (4/11)	Porencephalic cyst (1). Absence of the splenium of the corpus callosum, a large septum cavum pellucidum and cavum verge (1). Mildly coarse gyral pattern (1). Very small corpus callosum, cysts of right frontal region (1).
	Seizures	23%	25% (3/12)	
	Other			Hypertonia (1), hypotonia (3), autonomic dysfunction: apnea, bradycardia, temperature instability.
	Intellectual Disability		100% (13/13)	
	Behavior, personality			Friendly (6), Sociable (3), Extremely active (1), Affectionate (1), Fussy (1), Interactive (2), Decreased eye contact (1), Attention Deficit Disorder (1), Autistic-like features (1), Autism (1), Aggression (2) and Self-injurious behavior (2), Shy (1).

Clinical features are summarized by category. For the classical CdLS feature frequencies, percent frequencies are noted. For SMC3 features, percents are noted and in parentheses, fractional data. In the comments column, single numbers in parentheses indicate the number of subjects noted with the feature. Abbreviations: Ear-Nose-Throat (ENT); gastroesophageal reflux disease (GERD); central nervous system (CNS); patent ductus arteriosus (PDA), atrial septal defect (ASD), pulmonary stenosis (PS), ventricular septal defect (VSD), patent foramen ovale (PFO), aortic stenosis (AS), bicuspid aortic valve (BAV), peripheral pulmonic stenosis (PPS); tetralogy of Fallot (TOF), vesicoureteral reflux (VUR). *These frequencies in classical CdLS of these clinical features are compiled from different sources (Chatfield, et al., 2012; Jackson, et al., 1993; Kaga, et al., 1995; Kline, et al., 2007; Luzzani, et al., 2003; Nallasamy, et al., 2006; Selicorni, et al., 2009; Wygnanski-Jaffe, et al., 2005).