DIAL

Digital access to libraries

"Loss-of-function HDAC8 mutations cause a phenotypic spectrum of Cornelia de Lange syndrome-like features, ocular hypertelorism, large fontanelle and X-linked inheritance"

Kaiser, Frank J. ; Ansari, Morad ; Braunholz, Diana ; Gil-Rodríguez, María Concepción ; Decroos, Christophe ; Wilde, Jonathan J. ; Fincher, Christopher T. ; Kaur, Maninder ; Bando, Masashige ; Amor, David J. ; Atwal, P.S. ; Bahlo, Melanie ; Bowman, Christine M. ; Bradley, Jacquelyn J. ; Brunner, Han G. ; Clark, Dinah ; Campo, Miguel Del ; Di Donato, Nataliya ; Diakumis, Peter ; Dubbs, Holly ; Dyment, David A. ; Eckhold, Juliane ; Ernst, Sarah ; Ferreira, Jose C. ; Francey, Lauren J. ; Gehlken, Ulrike ; Guillén-Navarro, Encarna ; Gyftodimou, Yolanda ; Hall, Bryan D. ; Hennekam, Raoul ; Hudgins, Louanne ; Hullings, Melanie ; Hunter, Jennifer M. ; Yntema, Helger ; Innes, A. Micheil ; Kline, Antonie D. ; Krumina, Zita ; Lee, Hane ; Leppig, Kathleen ; Lynch, Sally Ann ; Mallozzi, Mark B. ; Mannini, Linda ; Mckee, Shane ; Mehta, Sarju G. ; Micule, Ieva ; Consortium, Care Rare Canada ; Mohammed, Shehla ; Moran, Ellen ; Mortier, Geert R. ; Moser, Joe-Ann S. ; Noon, Sarah E. ; Nozaki, Naohito ; Nunes, Luis ; Pappas, John G. ; Penney, Lynette S. ; Pérez-Aytés, Antonio ; Petersen, Michael B. ; Puisac, Beatriz ; Revencu, Nicole ; Roeder, Elizabeth ; Saitta, Sulagna ; Scheuerle, Angela E. ; Schindeler, Karen L. ; Siu, Victoria M. ; Stark, Zornitza ; Strom, Samuel P. ; Thiese, Heidi ; Vater, Inga ; Willems, P. ; Williamson, Kathleen ; Wilson, Louise C. ; Hakonarson, Hakon ; Quintero-Rivera, Fabiola ; Wierzba, Jolanta ; Musio, Antonio ; Gillessen-Kaesbach, Gabriele ; Ramos, Feliciano J. ; Jackson, Laird G. ; Shirahige, Katsuhiko ; Pié, Juan ; Christianson, David W. ; Krantz, Ian D. ; Fitzpatrick, David R. ; Deardorff, Matthew A.

Abstract

Cornelia de Lange syndrome (CdLS) is a multisystem genetic disorder with distinct facies, growth failure, intellectual disability, distal limb anomalies, gastrointestinal and neurological disease. Mutations in NIPBL, encoding a cohesin regulatory protein, account for >80% of cases with typical facies. Mutations in the core cohesin complex proteins, encoded by the SMC1A, SMC3 and RAD21 genes, together account for ˜5% of subjects, often with atypical CdLS features. Recently, we identified mutations in the X-linked gene HDAC8 as the cause

Référence bibliographique

Loss-of-function *HDAC8* mutations cause a phenotypic spectrum of Cornelia de Lange syndrome-like features, ocular hypertelorism, large fontanelle and X-linked inheritance

Frank J. Kaiser[1], **Morad Ansari**[2], **Diana Braunholz**[1], **María Concepción Gil-Rodríguez**[1,3,4,5], **Christophe Decroos**[6], **Jonathan J. Wilde**[7], **Christopher T. Fincher**[7], **Maninder Kaur**[7], **Masashige Bando**[9], **David J. Amor**[10,11], **Paldeep S. Atwal**[14], **Melanie Bahlo**[15,12], **Christine M. Bowman**[6], **Jacquelyn J. Bradley**[7], **Han G. Brunner**[16], **Dinah Clark**[7], **Miguel Del Campo**[17,18,19], **Nataliya Di Donato**[20], **Peter Diakumis**[15,13], **Holly Dubbs**[7], **David A. Dyment**[21], **Juliane Eckhold**[1], **Sarah Ernst**[7], **Jose C. Ferreira**[22], **Lauren J. Francey**[7], **Ulrike Gehlken**[1], **Encarna Guillén-Navarro**[19,23,24], **Yolanda Gyftodimou**[25], **Bryan D. Hall**[26], **Raoul Hennekam**[27], **Louanne Hudgins**[14], **Melanie Hullings**[7], **Jennifer M. Hunter**[2], **Helger Yntema**[16], **A. Micheil Innes**[28], **Antonie D. Kline**[29], **Zita Krumina**[30], **Hane Lee**[31], **Kathleen Leppig**[32], **Sally Ann Lynch**[33], **Mark B. Mallozzi**[7], **Linda Mannini**[34], **Shane Mckee**[35], **Sarju G. Mehta**[36], **Ieva Micule**[30], **Care4Rare Canada Consortium**[37], **Shehla Mohammed**[38], **Ellen Moran**[39], **Geert R. Mortier**[40], **Joe-Ann S. Moser**[6], **Sarah E. Noon**[7], **Naohito Nozaki**[41], **Luis Nunes**[42], **John G. Pappas**[43], **Lynette S. Penney**[44], **Antonio Pérez-Aytés**[45], **Michael B. Petersen**[46], **Beatriz Puisac**[3,4], **Nicole Revencu**[47], **Elizabeth Roeder**[48], **Sulagna Saitta**[7,49,50,51], **Angela E. Scheuerle**[52], **Karen L. Schindeler**[53], **Victoria M. Siu**[54], **Zornitza Stark**[10], **Samuel P. Strom**[31], **Heidi Thiese**[32], **Inga Vater**[55], **Patrick Willems**[56], **Kathleen Williamson**[2], **Louise C. Wilson**[57], **University of Washington Center for Mendelian Genomics**[†], **Hakon Hakonarson**[7,8,58], **Fabiola Quintero-Rivera**[31], **Jolanta Wierzba**[59,60,61,62,63], **Antonio Musio**[34], **Gabriele Gillessen-Kaesbach**[1], **Feliciano J. Ramos**[3,4,5], **Laird G. Jackson**[64], **Katsuhiko Shirahige**[9,65], **Juan Pié**[3,4,5], **David W. Christianson**[6], **Ian D. Krantz**[7,58], **David R. Fitzpatrick**[2] and **Matthew A. Deardorff**[6,58,*]

[1]Sektion für Funktionelle Genetik am Institut für Humangenetik, Universität zu Lübeck, Lübeck 23538, Germany, [2]MRC Human Genetics Unit, IGMM, University of Edinburgh, Edinburgh EH4 2XU, UK, [3]Unit of Clinical Genetics and Functional Genomics, Department of Pharmacology-Physiology, [4]Department of Pediatrics, Medical School, University of Zaragoza, Zaragoza E-50009, Spain, [5]Genetics Clinic, Service of Pediatrics, University Clinic Hospital 'Lozano Blesa', Zaragoza E-50009, Spain, [6]Department of Chemistry, University of Pennsylvania, Philadelphia, PA 19104, USA, [7]Division of Genetics, [8]Center for Applied Genomics, Children's Hospital of Philadelphia, Philadelphia, PA 19104, USA, [9]Laboratory of Genome Structure and Function, Research Center for Epigenetic Disease, Institute of Molecular and Cellular Biosciences, The University of Tokyo, Tokyo 113-0032, Japan, [10]Victorian Clinical Genetics Services, Murdoch Childrens Research Institute, Parkville, VIC 3052, Australia, [11]Department of Paediatrics, [12]Department of Mathematics and Statistics and [13]Department of Medical Biology, University of Melbourne, Melbourne, VIC 3010, Australia, [14]Division

of Medical Genetics, Department of Pediatrics, Stanford University, Stanford, CA 94305-5208, USA, [15]Bioinformatics Division, Walter and Eliza Hall Institute, Parkville, VIC 3052, Australia, [16]Department of Human Genetics 855, Radboud University Medical Center, Nijmegen 6525GA, The Netherlands, [17]Department of Genetics, Hospital Vall d́Hebron, Barcelona E-08015, Spain, [18]Department of Life Science, Universitat Pompeu Fabra, Barcelona 08002, Spain, [19]CIBERER, Instituto de Salud Carlos III (ISCIII), Madrid 28029, Spain, [20]Institut für Klinische Genetik, Technische Universität Dresden, Dresden 01307, Germany, [21]Department of Genetics, Children's Hospital of Eastern Ontario, Ottawa, ON, Canada K1H 8L1, [22]Department of Obstetrics and Gynecology, Faculty of Medicine, Medical University of Warsaw, Warsaw 02-091, Poland, [23]Unidad de Genética Médica y Dismorfología, Servicio de Pediatría, Hospital Clínico Universitario Virgen de la Arrixaca, El Palmar, Murcia 30120, Spain, [24]Cátedra de Genética Médica. UCAM-Universidad Católica San Antonio de Murcia, Spain, [25]Department of Genetics, Institute of Child Health, Athens GR-11527, Greece, [26]Department of Pediatrics, University of Kentucky, Lexington, KY 40536, USA, [27]Department of Pediatrics, Academic Medical Center, Amsterdam, AZ 1105, The Netherlands, [28]Department of Medical Genetics, Alberta Children's Hospital, Calgary, AB, Canada T3B 6A8, [29]The Harvey Institute for Human Genetics, Greater Baltimore Medical Center, Baltimore, MD 21204, USA, [30]Medical Genetics, Children's University Hospital, Riga LV-1004, Latvia, [31]Department of Pathology and Laboratory Medicine, David Geffen School of Medicine at UCLA and the UCLA Clinical Genomics Center, University of California Los Angeles, Los Angeles, CA 90024, USA, [32]Genetic Services, Group Health Cooperative, Seattle, WA 98112, USA, [33]National Centre for Medical Genetics, Our Lady's Children's Hospital, Dublin, Ireland, [34]Istituto di Ricerca Genetica e Biomedica, C.N.R, Pisa 56124, Italy, [35]Department of Genetic Medicine, Belfast City Hospital, Belfast, Ireland, [36]Department of Clinical Genetics, Addenbrookes Hospital, Cambridge, Cambridge CB2 0QQ, UK, [37]Care4Rare Canada Consortium, Department of Genetics, 401 Smyth Road, Ottawa, ON, Canada K1H 8L1, [38]Clinical Genetics Service, Guy's Hospital, London SE1 9RT, UK, [39]NYU Hospital for Joint Diseases, New York, NY 10003, USA, [40]Department of Medical Genetics, Antwerp University Hospital and University of Antwerp, B-2650 Antwerp, Edegem, Belgium, [41]MAB Institute, Inc, Sapporo 001-0021, Japan, [42]Medical Genetics Department, Dona Estefânia Hospital, CHLC and Medical Sciences Faculty of Lisbon, UNL, Lisboa 1100, Portugal, [43]Department of Pediatrics, NYU School of Medicine, New York, NY 10016, USA, [44]Department of Pediatrics, Dalhousie University, Halifax, NS, Canada B3H 4R2, [45]Unidad de Dismorfología y Genética Reproductiva, Grupo de Investigacion en Perinatologia, Instituto de Investigacion Sanitaria Hospital Universitario LA FE, Valencia 46026, Spain, [46]Department of Clinical Genetics, Aalborg University Hospital, Aalborg 9000, Denmark, [47]Center for Human Genetics, Cliniques Universitaires St Luc, Université Catholique de Louvain, Brussels, Belgium, [48]Section of Genetics, Department of Pediatrics, Baylor College of Medicine, Children's Hospital of San Antonio, San Antonio, TX 78207, USA, [49]Medical Genetics Institute, Division of Genetics, Department of Pediatrics, [50]Department of Pathology and [51]Department of Laboratory Medicine, Cedars-Sinai Medical Center, Los Angeles, CA 90048, USA, [52]Tesserae Genetics, Dallas, TX 75230, USA, [53]Maritime Medical Genetics Service, IWK Health Centre, Halifax, NS, Canada B3K 6R8, [54]Division of Medical Genetics, Department of Pediatrics, Western University, London, ON, Canada N6A 3K7, [55]Institute of Human Genetics, University Hospital Schleswig-Holstein Campus Kiel/Christian-Albrechts University, Kiel, Germany, [56]Gendia, Antwerp 2020, Belgium, [57]Department of Clinical Genetics, Great Ormond Street Hospital, London WC1N 3JH, UK, [58]Department of Pediatrics, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA 19104, USA, [59]Department of Pediatrics, [60]Department of Hematology, [61]Department of Oncology, [62]Department of Endocrinology and [63]Department of General Nursery, Medical University of Gdansk, Gdansk, Poland, [64]Department of Obstetrics and Gynecology, Drexel University College of Medicine, Philadelphia, PA 19102, USA and [65]CREST, JST, K's Gobancho, 7, Gobancho, Chiyoda-ku, Tokyo 102-0076, Japan

---

Cornelia de Lange syndrome (CdLS) is a multisystem genetic disorder with distinct facies, growth failure, intellectual disability, distal limb anomalies, gastrointestinal and neurological disease. Mutations in *NIPBL*, encoding a cohesin regulatory protein, account for >80% of cases with typical facies. Mutations in the core cohesin complex proteins, encoded by the *SMC1A*, *SMC3* and *RAD21* genes, together account for ∼5% of subjects, often with atypical CdLS features. Recently, we identified mutations in the X-linked gene *HDAC8* as the cause of a small number of CdLS cases. Here, we report a cohort of 38 individuals with an emerging spectrum of features caused by *HDAC8* mutations. For several individuals, the diagnosis of CdLS was not considered prior to genomic testing. Most mutations identified are missense and *de novo*. Many cases are heterozygous females,

each with marked skewing of X-inactivation in peripheral blood DNA. We also identified eight hemizygous males who are more severely affected. The craniofacial appearance caused by *HDAC8* mutations overlaps that of typical CdLS but often displays delayed anterior fontanelle closure, ocular hypertelorism, hooding of the eyelids, a broader nose and dental anomalies, which may be useful discriminating features. *HDAC8* encodes the lysine deacetylase for the cohesin subunit SMC3 and analysis of the functional consequences of the missense mutations indicates that all cause a loss of enzymatic function. These data demonstrate that loss-of-function mutations in *HDAC8* cause a range of overlapping human developmental phenotypes, including a phenotypically distinct subgroup of CdLS.

INTRODUCTION

Cornelia de Lange syndrome and cohesin

Cornelia de Lange syndrome (CdLS; or Brachmann-de Lange syndrome, OMIM 122740, 300590, 610759 and 614701) is a genetically heterogeneous disorder with congenital malformations that include characteristic facial features, growth retardation, intellectual disability and limb anomalies. Typical CdLS is caused by *de novo*, heterozygous mutations in *NIPBL* (1–7), and recent data have demonstrated that as many as 20% of individuals may have mosaic mutations that are not readily detectable in blood (8). *NIPBL* encodes the human ortholog of the Drosophila Nipped-B (9) and yeast Scc2 proteins, key components of sister chromatid cohesion (10). NIPBL is required for the chromosomal loading of the cohesin complex (11,12), which comprises four highly conserved subunits: SMC1, SMC3, RAD21 (hSCC1) and STAG (SCC3/SA) (13). Approximately 5% of patients with CdLS, usually those with milder congenital anomalies or atypical facial features, have mutations in the human *SMC1A* gene (14,15). One patient has been reported with a mutation in the *SMC3* gene (16) and several have been noted with mutations in *RAD21* (17). However, 10–20% of patients with typical CdLS have no identified etiology. In addition, a high number of individuals with atypical CdLS or an overlapping phenotype have no identified cause.

Acetylation of SMC3 and the role of HDAC8

The acetylation of SMC3 at key lysine residues is required to establish the cohesiveness of chromatin-loaded cohesin during S-phase (18–20). In addition, SMC3 is rapidly deacetylated during mitosis (19–21) and was noted in yeast to be regulated by the class I deacetylase Hos1 (22–24). We recently identified HDAC8 as the vertebrate SMC3 deacetylase and demonstrated its role in recycling cohesin from one cell cycle to the next (25). Furthermore, we identified *HDAC8* mutations in six individuals with typical CdLS or an overlapping phenotype.

To investigate and fully characterize the range of clinical features in patients with *HDAC8* mutations, we screened an internationally assembled cohort of 586 patients with typical CdLS and overlapping clinical presentations who had no known molecular etiology. Here we report on a range of mutations, including chromosomal aberrations, nonsense and missense mutations that disrupt *HDAC8* and clarify the clinical spectrum of features in individuals with these mutations. This study demonstrates the range of phenotypes caused by *HDAC8* mutations in an emerging era of genomic/exomic sequencing, since seven subjects herein were identified by genomic studies (array and exome sequencing) independent of a CdLS diagnosis. This work serves to emphasize the range of features seen in individuals with *HDAC8* mutations and underscores their distinctness from typical CdLS. Furthermore, it portends the likelihood that additional subjects will be identified by NGS approaches.

RESULTS

We recently identified HDAC8 as the vertebrate deacetylase responsible for deacetylating cohesin to facilitate its recycling (25) and identified mutations in a small number of subjects from a cohort of children with CdLS. Also recently, a family with an X-linked pattern of obesity and mental retardation in males was noted to have an intronic *HDAC8* mutation that segregates with the phenotype (26). To further clarify the clinical consequences of *HDAC8* mutations in human developmental disorders, we assembled a large international cohort of individuals with CdLS and overlapping phenotypes and screened for mutations.

Identification of HDAC8 mutations in a large cohort of patients

We assessed 586 individuals with CdLS and overlapping phenotypes for whom screening for mutations in *NIPBL*, *SMC1A* and, in most cases, *SMC3* (see Materials and Methods) was negative. Among this cohort, we identified 25 probands, three affected family members and three unaffected mothers with mutations in *HDAC8*, revealing a cause in ∼4% of probands. In addition, two children were ascertained by discovery of *de novo HDAC8* microdeletions and five as a result of identifying a *de novo HDAC8* mutation using clinical exome sequencing, all these latter cases prior to considering the diagnosis of CdLS. In total, including the six previously reported individuals (25), this includes six different chromosomal microdeletions or microduplications, three nonsense mutations, one splice site and 16 different missense mutations (Fig. 1). Of the mutations, 23 occurred *de novo* in the affected child, six were maternally inherited (in four families) and in seven the inheritance could not be fully confirmed.

HDAC8 mutations in CdLS cause loss of enzyme activity

To obtain independent data regarding the effect and severity of these mutations on protein function, we mapped mutations onto the previously reported crystal structure of HDAC8 (27,30) (Fig. 1B) and expressed the wild type and each of the
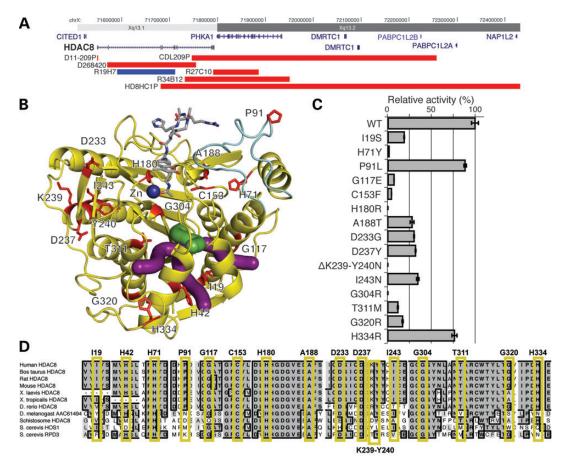
Figure 1. Mutations in *HDAC8*. (**A**) Copy number abnormalities that disrupt the *HDAC8* locus. Localization of deletions (red) and duplications (blue) on Xq13 are indicated by chromosome band and position (hg19). Gene locations are indicated in blue with gene names. *HDAC8* is indicated in bold. (**B**) Localization of HDAC8 missense mutations on the crystal structure (PDB accession code 2V5W (27)). Mutated residues are in red and labeled. The acetylated substrate is in gray and the active site zinc ion (Zn) is in dark blue. The light blue portion indicates the L2 substrate-binding loop. Green and purple tubes indicate the catalytically important main and branch exit channels, respectively, for the acetate hydrolysis product (28,29). (**C**) HDAC8 mutations disrupt deacetylase activity. Bar graphs demonstrate the effect of HDAC8 mutations on deacetylase relative activity relative to the wild-type enzyme activity (specific activities are noted in Supplementary Material, Table S1). All assays were performed in triplicate. Error bars indicate standard deviation. Enzymatic activity for all mutations is significantly less than wild type ($P < 0.05$) using a two-tailed unpaired *t*-test. Single letter amino acids are used for mutation notation to correlate with residues in (B and C). (**D**) Conservation of mutations in HDAC8. Alignment of HDAC8 and related deacetylases from 10 species, as indicated on the left. Residues identical to the human residue in a majority of species at each position are in bold, in dark gray and a black border. Those that are similar are shaded in lighter gray, without a border. Each mutated human residue is indicated at the top, with a yellow box surrounding the aligned residues. Note the conservation through yeast for all residues in which a missense mutation causes a complete loss of function.

mutated forms of HDAC8 in *Escherichia coli*. We purified each and assayed deacetylase activity using a commercially available HDAC8 assay (31) (Fig. 1C; Supplementary Material, Table S1).

Wild-type HDAC8 requires a single zinc ion for the hydrolysis of N6-acetyllysine side chains in acetylated protein substrates such as SMC3 (blue sphere in Fig. 1B). The side chains of D178, H180 and D267 coordinate to the zinc ion and play a critical role in stabilizing the active site structure and modulating the chemistry of the zinc ion for catalysis (32). X-ray crystal structures of inactive HDAC8 mutants complexed with peptide substrates reveal extended regions of the protein surface capable of interacting with protein substrates, including a portion of the L2 loop (light blue in Fig. 1B) that exhibits moderate flexibility in the binding of substrates and inhibitors (27,30,33,34). A catalytically important internal channel leading to possible exit channels for the acetate hydrolysis product has been identified in modeling and mutagenesis studies (green and purple tubes

in Fig. 1B) (28,29). This channel may facilitate catalysis when HDAC8 is bound to large protein substrates by providing a "back door" exit for the small product without first requiring dissociation of the HDAC8-substrate complex. Inspection of mutation sites mapped onto the HDCA8 structure reveals that most appear capable of compromising molecular features required for optimal catalytic function.

Based on structural location and enzymatic activity, mutations are classified into four groups. The first group of mutations (H180R, G304R, H71Y and ΔK239-Y240N) completely inactivate the enzyme. The H180R mutation destroys the binding site of the catalytically obligatory zinc ion. The G304R mutation introduces a large positively charged residue near the active site that might destabilize the protein structure and/or block substrate binding. The H71Y substitution disrupts hydrogen bonds with the backbone NH group of N156 and the backbone carbonyl groups of K145 and K146, which may destabilize the L2 loop and adjacent residues that accommodate substrate binding. Finally,

the positively charged amino group of K239 makes an electrostatic interaction with D355 and donates a hydrogen bond to the backbone carbonyl group of N357. While the side chain of adjacent residue Y240 does not make any hydrogen bond interactions with protein residues, it is close to the active site. The deletion of K239 and the simultaneous Y240N mutation suggests they perturb the protein scaffold in the vicinity of the active site to abolish catalytic activity.

The second mutation group, comprised of G117E and C153F, is characterized by residual enzyme activity less than 10% of that observed for the wild-type enzyme. The larger, negatively charged glutamate side chain of G117E would be oriented toward solvent. However, it could sterically perturb the conformation of the adjacent K60-A62 polypeptide segment, to destabilize the nearby L2 loop. C153 is located only ∼10 Å away from the catalytically obligatory zinc ion, and it is also close to H142 and H143, residues that are known to be involved in catalysis (30,32). The substitution by a larger hydrophobic phenylalanine is likely to dramatically remodel the active site and thus impair catalysis.

A third group of mutations (T311M, I19S, G320R, A188T, D233G, D237Y and I243N) exhibit ∼10–50% residual activity compared with the wild-type enzyme. The T311M mutation may perturb the putative internal channel for product acetate dissociation, and the I19S may perturb one of the possible acetate exit channels that branch away from the internal channel. Additionally, the T311M and G320R mutations may compromise the active site structure, since the helix containing these residues (helix H2) buttresses the β-sheet comprising one wall of the active site. The T311 hydroxyl group also donates a hydrogen bond to the backbone carbonyl group of N307, and disruption of this hydrogen bond by the T311M mutation would destabilize the protein. The guanidinium side chain in the G320R mutation would occupy a cleft on the surface of the protein, but it is difficult to explain how this particular substitution might influence the active site structure. A188 is in a hydrophobic environment, surrounded by F70, V159, L163, V185 and F189; mutation to a larger, partially hydrophilic residue might subtly affect packing interactions. D233 is located on the protein surface, but its carboxylate group accepts a salt-linked hydrogen bond from the side chain of K202. The deletion of the carboxylate group in the D233G mutant disrupts this interaction; being moderately close to the active site, this mutation might cause some subtle structural changes that compromise catalysis. The carboxylate side chain of D237 accepts hydrogen bonds from the side chain hydroxyl group and the backbone NH group of T280, and D237 is moderately close to the active site. Mutation of D237 to a less polar residue in the D237Y mutant probably causes a structural rearrangement that propagates through to the active site. Finally, while I243 is a partially conserved residue among the HDAC isozymes, analysis of the HDAC8 structure does not suggest any particular reason for the loss of activity resulting from this mutation.

A fourth group of mutations (P91L and H334R) exhibit relatively minor but significant losses of catalytic activity (residual activities >50% compared with the wild-type enzyme). H334 is located on the protein surface far from the active site, and P91L is in the portion of the L2 loop that is most distant from the active site.

The varying loss of activity of these mutations also correlates with the amino acid conservation at each site (Fig. 1D). Each of the substitution mutations that result in complete ablation of activity (H71Y, H180R and G304R) alters a residue that is identical from humans through yeast. Mutations with minor or intermediate losses of activity largely reside in residues that are less conserved in non-vertebrates, although nearly all mutations occur in residues that are nearly identical in vertebrates, consistent with the overall high conservation of sequence in vertebrate HDAC8 homologs.

HDAC8 mutations are associated with severely skewed X-inactivation

To assess the *HDAC8* mutations effect on RNA, we evaluated expression in cell lines available for several of the probands. We noted that for a male subject (case ID CDL426P), with a hemizygous missense mutation, the mutant transcript was detectable, as expected for this X-linked gene. However, for available lymphoblastoid cell lines from females (CDL016P, CDL209P, CDL248P, CDL317P and CDL603P) that harbored heterozygous missense mutations (Table 1), we were only able to detect the wild-type allele, even in the presence of the nonsense-mediated decay inhibitor cycloheximide. Since blood cell lines freely mix and would be expected to demonstrate biallelic expression, this inability to detect the mutant allele suggested selection for expression of the normal allele in these cell lines or in peripheral blood prior to their establishment. We then analyzed the allelic expression in a fibroblast cell line from the female probands CDL016P and CDL248P, the only two subjects for whom both lymphoblastoid and fibroblast cell lines were available. We reasoned that the line derived from a small skin biopsy had a better possibility of being clonally derived and would be less subject to selection. Consistent with this possibility, these cells expressed only the mutant allele, also supporting the likelihood of selection within the peripheral blood (25) (Table 1).

Because lymphoblastoid and fibroblast cell lines are available for only a small subset of patients, we analyzed for skewing of X-inactivation as a marker of *HDAC8* allele selection in the peripheral blood. To accomplish this, we utilized a well-characterized assay that measures the methylation status of a microsatellite repeat in the 5′UTR of the androgen receptor (AR) of the inactivated X chromosome (35). For 23 heterozygous females for whom peripheral blood genomic DNA or cell lines were available, 20 were skewed >95:5, one was skewed between 95:5 and 70:30, one showed no skewing and two were uninformative due homozygous polymorphisms at the testing locus. Of note, the three samples that were skewed <95:5 were all from girls aged <4 years when the sample was obtained, which supports the hypothesis that selection of the normal allele occurs after birth.

Clinical features of individuals with *HDAC8* mutations

CdLS is noted for growth, cognitive delays, limb anomalies and distinctive facial features. Of the 35 affected individuals identified in this work with mutations in *HDAC8*, all have clinical features that overlap those seen in CdLS, although few subjects are fully consistent with that diagnosis. Figure 2 demonstrates facial and limb findings, Supplementary Material, Table S2 details the

Table 1. Skewed X-inactivation with HDAC8 mutations

Patient	DNA source	Gender	Allele 1:Allele 2	cDNA allele expressed
C0007	Blood	Female	97:3	N/A
C0084	Blood	Female	98:2	N/A
C0146	Blood (3 years 7 months)	Female	16:84	N/A
CDL016P	Blood	Female	100:0	100% wt
CDL016P	Fibroblast	Female	0:100	100% mut
CDL179P	Blood	Male	100	N/A
CDL209P	Blood	Female	100:0	100% wt
CDL231P	Blood	Female	0:100	N/A
CDL233P	Blood	Male	100	N/A
CDL248P	Blood	Female	UTD	100% wt
CDL248P	Fibroblast	Female	UTD	100% mut
CDL279P	Blood (infant)	Female	59:41	N/A
CDL317P	Blood	Female	100:0	100% wt
CDL326AS	Blood	Female	100:0	N/A
CDL326M	Blood	Female	UTD	N/A
CDL326P	Blood	Male	100	N/A
CDL347P	Blood	Male	100	N/A
CDL423M	Blood	Female	100:0	N/A
CDL423P	Blood	Male	100	N/A
CDL426P	Blood	Male	100	100% mut
CDL603P	Blood	Female	100:1	100% wt
CDLG13952	Blood (3 years)	Female	8:92	N/A
D11-209M	Blood	Female	2:98	N/A
DECIPHER 268420	Blood	Female	100:0	N/A
HD8HC01P	Blood (1 year)	Female	100:0	N/A
R19E9	Blood	Female	95:5	N/A
R19H7	Blood	Female	2:98	N/A
R23F4	Blood	Female	98:2	N/A
R24E11	Blood	Female	3:97	N/A
R34B12	Blood (15 months)	Female	85:15	N/A
R34B12	Blood (5 years)	Female	100:0	N/A
R27C10	Blood	Female	1:99	N/A

Patient ID, source of DNA, age (in parentheses where relevant) and sex are indicated. The ratio of the activity is noted for each allele. In the allele ratio column, 'UTD' indicates that skewing cannot be determined due to homozygosity of the AR microsatellite allele tested. In this column, samples are shaded to indicate whether they are male (blue), female skewed toward the mutated allele (red), female skewed toward the normal allele (green), female skewed toward unknown allele (gray) or unskewed (unshaded). Note that in some familial cases (CDL326, CDL423), the determination of the wild-type and mutant alleles is inferred from an affected male family member. nd, not determined; N/A, not available; wt, wild-type allele; mut, mutated allele.

clinical features for each individual and Table 2 demonstrates the summary of clinical findings in this cohort. The range of phenotypes was consistent with that seen for an X-linked disorder, with males being more affected. In multiplex families, we observed apparently unaffected carrier females. This sex-weighted effect is illustrated by the differences in average height, weight and head circumference between the two sexes in affected individuals (Table 3).

In this cohort of affected individuals, many features are quite similar to those observed in typical CdLS. These include non-specific findings such as postnatal growth retardation (28%), microcephaly (29%), hearing loss (59%), gastroesophageal reflux (67%), feeding difficulties (86%), cardiac defects (36%), genito-urinary anomalies (44%) and intellectual disabilities (100%).

Many also display craniofacial features suggestive of CdLS, including brachycephaly (70%), arched eyebrows (88%),

synophrys (90%), long eyelashes (45%), depressed nasal bridge (45%), anteversion of the nares (76%), long philtrum (57%), downturned corners of the mouth (57%), small widely spaced teeth (61%), micrognathia (59%) and cleft palate without cleft lip (18%). Other findings suggestive of CdLS include cutis marmorata (18%), hirsutism (65%), small hands (96%) and feet (89%), toe syndactyly (44%), verbal performance that is worse than motor performance and development of behavioral problems with age. Only one patient demonstrates a major limb anomaly.

However, several features are unique in these patients that are not seen in typical CdLS. These include delayed fontanelle closure (~50%), ocular hypertelorism (47%) and/or telecanthus (64%), hooding or redundant overfolded skin of the upper eyelids (46%), a broad or bulbous nasal tip (66%), a prominent gap between the central incisors (50%), dental anomalies (~50%, including oligodontia and fused incisors), nevus flammeus (58%), mosaic skin pigmentation (~25%), limb length discrepancies in the absence of major malformations ($n = 5$) and happy or friendly personalities as younger children (~50%).

Finally, several features are absent in this cohort that are commonly seen in typical CdLS. For example, the growth tends to be less disrupted with lower frequency of postnatal growth retardation (28%) and microcephaly (20%). In the face, the philtrum is often long but it is infrequently flat as in typical CdLS. In addition, the vermillion border of the upper lip tends to be thin less frequently. Only one subject has major cardiac defect, none have diaphragmatic defects and only a minority of individuals have cutis marmorata (17%), single palmar creases (20%) or restriction of elbow range of motion (4%).

DISCUSSION

Variable clinical severity caused by HDAC8 mutations

In an ongoing effort to understand the phenotypic consequences of mutations in *HDAC8*, we screened a broad cohort of subjects with CdLS or features that overlap the diagnosis and have sought to identify patients with *HDAC8* mutations ascertained independently of a diagnosis of CdLS. In total, we have identified 38 individuals with mutations in *HDAC8*. These range from genomic copy number abnormalities through single nucleotide substitutions that cause missense mutations. Based on these numbers, individuals with *HDAC8* mutations comprise ~4% of patients ascertained (25 probands of 586 screened). However, the identification of *HDAC8* mutations by genomic testing in seven individuals without a diagnosis of CdLS affirms that this phenotype differs from typical CdLS and may ultimately have wider ranging features than even the patients reported here. Furthermore, the severity of clinical phenotypes caused by mutations in *HDAC8*, especially in females, is heavily influenced by random X-inactivation. We would suspect that females with unfavorable Lyonization in the brain would display a more affected phenotype, but additional investigation would be needed to test this model.

HDAC8 mutations in CdLS cause reduced enzymatic activity

Although we have noted a range of clinical features, genetic and biochemical analysis of all *HDAC8* mutations to date causing CdLS or overlapping phenotypes predict or demonstrate loss
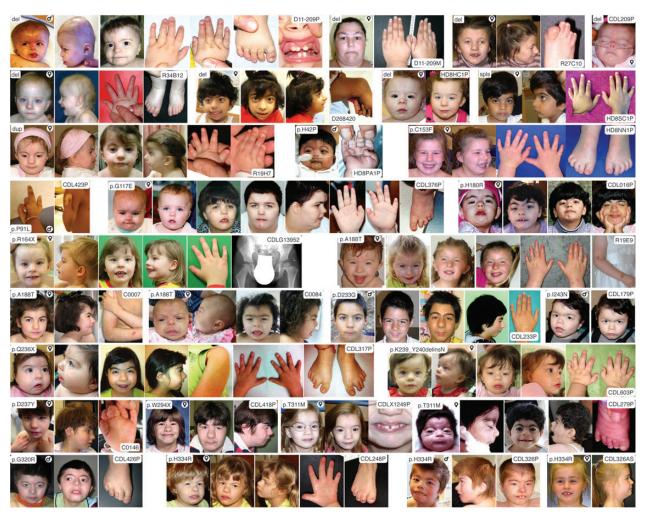
Figure 2. Individuals with *HDAC8* mutations. Facial features of individuals with *HDAC8* mutations labeled with corresponding mutation and sex; ♂, male; ♀, female. Photos for individual patients are grouped, with the first image for each containing the type of mutation and sex, and the last image containing the subject identifier.

of function. Several mutations we have identified result from microdeletions or microduplications that disrupt the gene and several result from nonsense substitutions. Furthermore, enzymatic assays of all missense mutations demonstrated reduced or absent activity. This correlates with *in silico* mapping of the missense mutations using the known crystal structure that shows these are often located in or around residues critical for HDAC8 function. Although the numbers are quite small, there is a weak trend in males toward a milder phenotype with higher residual activity of HDAC8. However, a significant caveat exists for this observation, since this assay was performed *in vitro* with standardized amounts of purified HDAC8, and does not represent *in vivo* levels of protein. Unfortunately, we have cell lines for very few of the affected boys to determine stability of each mutation *in vivo* and further studies would be needed to more clearly correlate genotype and clinical severity.

Complete loss of HDAC8 function appears to be viable in humans, as evidenced by a male with a partial deletion of *HDAC8* (D11−209P) and another with undetectable levels of HDAC8 protein in cells (CDL426P). This is supported by the finding that mice with a targeted null mutation can be viable, albeit at lower frequencies than wild type (38). This differs from the survivability seen for loss of other cohesin genes (39−41) in which homozygous knockouts result in embryonic lethality. This finding suggests that the endogenous function of HDAC8 may be at least partially redundant, or that other proteins, such as related deacetylases, may be upregulated to compensate for HDAC8 loss of function.

Overlapping and unique features caused by *HDAC8* mutations

Although the facial features certainly overlap those of patients with typical CdLS caused by *NIPBL* mutations, there are notable features in individuals with *HDAC8* mutations. Several of the most distinct likely derive from an abnormality of skull formation. Quite notably, this effect is directly comparable with the *Hdac8* mouse knockout, which demonstrates extreme growth failure and markedly delayed closure of skull sutures (38). In a large percentage of individuals with *HDAC8* mutations, delayed, sometimes marked, closure of the anterior fontanelle has been observed. We hypothesize that this effect represents an embryonic anomaly that results in delayed development of midline structures of the face to result in hypertelorism,

Table 2. Clinical feature frequency in individuals with HDAC8 mutations

Category	Feature	Percent (no. observed/no. assessed)	Details
Head	Brachycephaly	70% (24/34)	
	Low anterior hairline	67% (23/34)	
	Skull		Prominent metopic suture (1), open/delayed fontanelles (13), asymmetry (2)
Eyes	Arched eyebrows	88% (30/34)	
	Synophrys	90% (30/33)	
	Long eyelashes	45% (14/31)	
	Hypertelorism	47% (16/34)	
	Telecanthus	64% (22/34)	
	Ptosis	18% (6/32)	
	Myopia	44% (11/25)	
	Lacrimal duct obstruction	24% (6/25)	
	Hooding of lids	46% (15/32)	
Nose	Depressed nasal bridge	45% (15/33)	
	Anteverted nostrils	76% (26/34)	
	Long philtrum	57% (19/33)	
	Featureless philtrum	30% (10/33)	
	Broad/bulbous nasal tip	66% (22/33)	
Mouth	Thin upper lip	29% (10/34)	
	Downturned corners of mouth	57% (19/33)	
	Palate—high arch	33% (7/21)	
	Palate—cleft	18% (5/27)	
	Widely spaced teeth	61% (16/26)	
	Gap between upper incisors	50% (13/26)	
	Other teeth	(19)	Congenital abnormalities (10), oligodontia (3), fused (4), delayed (6)
	Micrognathia/retrognathia	59% (19/32)	
	Short neck	48% (12/25)	
	Other facial		Facial asymmetry (2), grimacing smile (1), small low-set ears (5), macrostomia (1), premature aging (1)
Skin	Cutis marmorata	18% (5/27)	
	Hirsutism	65% (21/32)	
	Nevus flammeus	58% (17/29)	
	Other skin		Pigmentary mosaicism (6), slow-growing hair and nails (1)
Hands	Small hands	96% (28/29)	
	Proximally set thumbs	85% (23/27)	
	Clinodactyly 5th finger	50% (14/28)	
	Short 5th finger	84% (21/25)	
	Single Palmar crease	20% (5/24)	
Feet	Small feet	89% (26/29)	
	Syndactyly of toes	44% (12/27)	
Arms	Restriction of elbow movements	4% (1/23)	
	Other skeletal		Atypical/absent patella (2)
Cardiac	Cardiac defects		Leg asymmetry (5), Short 4/5 metacarpals, 5th fingers and distal (6), phalanges of thumbs
Genitourinary	Genitourinary defects	36% (11/30)	
Gastrointestinal	GER	44% (13/29)	VUR/hydronephrosis (6), renal cysts (2), dysplasia (2), chronic disease (2), cryptorchidism or hypoplastic genitalia (5/8 males) PCOS (2)
	Feeding problems in infancy	67% (19/28)	
	GI anomalies	86% (25/29)	
	Other gastrointestinal		
Otolaryngologic	Hearing loss	59% (16/27)	
CNS	CNS anomalies	22% (6/27)	
	Seizures	28% (8/28)	
	Other		Hypertonia (3), hypotonia (1), gait apraxia (1)
Cognitive	Intellectual disability	100% (30/30)	Carrier mothers not carefully evaluated
	Behavior, personality	(23)	Autistic (3), friendly (14), hyperactive (3), aggressive (3 older), self-injurious (3)

Clinical features are summarized by category. Percent and fractional data are indicated. Single numbers in parentheses indicate the number assessed or noted with the feature. M, male; F, female; PCOS, polycystic ovarian syndrome; GER, gastroesophageal reflux; CNS, central nervous system; VUR, vesicoureteral reflux.

Table 3. Comparison of quantitative features between males and females with HDAC8 mutations and classical CdLS

Parameter	HDAC8 males	HDAC8 females	Classical CdLS
Birth weight	-1.7 (-3 to $^-0.5$, $n = 6$)	-1.6 (-3.6 to 0, $n = 21$)	-1.9 (-3.0 to 0.0)
Birth length	-2.4 (-4 to -1, $n = 5$)	-1.9 (-4 to $+0.5$, $n = 17$)	-1.5 (-4.3 to $+0.7$)
Birth head circumference	-1.7 (-2.8 to -0.5, $n = 2$)	-1.8 (-3 to $+1.5$, $n = 12$)	-2.0 (-2.8 to -0.9)
Subsequent weights	-2.8 (-6.5 to -0.3, $n = 7$)	-1.4 (-6.5 to $+3.0$, $n = 22$)	-4.8 (-10.1 to -1.8)
Subsequent heights	-2.8 (-6.6 to $+0.4$, $n = 7$)	-2.0 (-4.5 to $+0.4$, $n = 22$)	-4.1 (-8.0 to -1.2)
Subsequent head circumference	-2.5 (-6.6 to -0.4, $n = 7$)	-1.9 (-5 to $+1$, $n = 21$)	-3.8 (-5.6 to -1.5)

Growth parameters are noted at the top. Z-scores were calculated using the LMS values from 2000 United States CDC Growth Percentile Data Files (36). Data within the *HDAC8* categories are presented as: (Z-score average, minimum-to-maximum range and the number or individuals with available data). Z-scores for classical CdLS data are derived from Kline *et al.* (37) and represents an average of both sexes at birth and 3 years for the mean and 5th–95th percentile range.

telecanthus and/or forehead nevus flammeus. Furthermore, several additional facial features, including hooding of eyelids and dental anomalies, are distinctive in this cohort. These findings, in combination with a more well-formed philtrum and upper lip, suggest that the face may be useful in defining a clinically recognizable subgroup of individuals. While all affected cases demonstrate some dysmorphic facial features, there is a significant diversity of facial appearance within the *HDAC8*-mutated subjects. This diversity may reflect level of mutation load in the embryonic face (in the context of unaffected mothers) or an evolving phenotype with age. However, it may also provide clues to domain-specific functions of the protein.

Of note, a family was recently reported by Harakalova *et al.* (26), in which affected males demonstrate intellectual disability, truncal obesity, hypogonadism and distinctive facial features caused by a splice site mutation in *HDAC8*. This family does not demonstrate facial features or a body build of typical CdLS, suggesting that the mechanism of this mutation may differ from that for cases of CdLS initially noted with *HDAC8* mutations (25). In support of this, our cohort had only three individuals for whom obesity was a significant feature. However, several of the boys reported by Harakalova *et al.* (26) demonstrate some weak similarity with hooding of the eyelids and arched eyebrows. In combination with the wide range of phenotypes seen in our cohort, it is possible that this family represents another instance of an emerging spectrum of phenotypes caused by *HDAC8* mutations. Further analysis of the mutation identified in this family and the role of HDAC8 in growth and metabolism will be useful to help further understand these observations.

Updated clinical diagnostic approach for CdLS

Based on genotype–phenotype correlations, we currently recommend the following diagnostic approach for patients with features suggestive of CdLS due to atypical growth, cognition, facies or limbs. For those with classical facial features, *NIPBL* should be assessed first, preferably in buccal or skin cells as suggested by Huisman *et al.* (8). If a multiple cohesin gene panel is not available and the facial features are atypical, X-inactivation studies may be useful. In the setting of atypical facies, a wide fontanelle, hooding of the eyelids and/or heavily skewed X-inactivation studies, *HDAC8* assessment would be increased in priority. For those with atypical features that include broader fuller less arched eyebrows or a broader nasal tip, *SMC1A*, and perhaps *SMC3*, assessment would be higher priority than *HDAC8* (14). If the above testing is negative, *RAD21* should be considered. As exome

sequencing coverage increases, exome-based exon-level copy number analysis improves, and the cost of exome sequencing falls, it will become reasonable to consider exome sequencing as first line testing for this heterogenous disorder.

Unique clinical recommendations for these patients

For patients in whom an *HDAC8* mutation is suspected or identified, we would include a workup similar to patients with a diagnosis of classical CdLS (42,43), which typically includes cardiac, renal and gastroenterologic studies, palate evaluation, hearing studies, speech, physical and occupational therapy assessments, molecular studies to confirm a diagnosis, and assessment of platelet levels and immune function. In addition, a careful maternal family history should be assessed and molecular testing of all mothers of children for whom *HDAC8* mutations have been found should be considered. If a high suspicion of an *HDAC8* mutation exists but none is identified, X-inactivation studies could help to increase or decrease suspicion of a mutation in the region.

In conclusion, mutations in *HDAC8* comprise ∼4% of mutations in individuals with features suggestive of CdLS and result in an atypical clinical phenotype that has elements of overlap with patients harboring *NIPBL* mutations. However, in general, there are several exceptions to the phenotypic similarity that include wide skull fontanelles, hooded eyelids, a slightly broader nasal base and a notably pleasant personality. This work emphasizes the role of acetylation of non-histone targets in human development and stresses the need for additional studies to understand the mechanism of cohesin regulation in the context of human developmental disorders.

MATERIALS AND METHODS

Subjects

All individuals in this study were diagnosed by clinical geneticists to have clinical features consistent with or overlapping a diagnosis of CdLS or were found to have *HDAC8* mutations identified by genome-wide analyses. All subjects and family members were enrolled in the study under an institutionally approved protocol of informed consent at The Children's Hospital of Philadelphia, the Institut für Humangenetik Lübeck or the UK (Scotland A) MREC Committee. All probands ascertained as CdLS were previously screened and found negative for mutations in *NIPBL* and *SMC1A*. Most were previously

screened for mutations in *SMC3*, *RAD21* and other cohesin genes. Subjects with *HDAC8* mutations found by genome-wide methods were enrolled following their identification.

Mutation screening

The coding exons and intron–exon boundaries of *HDAC8* were amplified by PCR of genomic DNA and Sanger sequenced. Primers were designed using ExonPrimer (http://ihg.gsf.de/ihg/ExonPrimer.html) (44). Primer sequences and PCR conditions (25) are available upon request. Sequencing was performed using BigDye Terminator v3.1 cycle sequencing and analyzed on an ABI 3730 (Applied Biosystems, Carlsbad, CA, USA).

Genome-wide copy number analysis

Whole-genome SNP genotyping was performed using Illumina (San Diego, CA, USA) Infinium HumanHap550 or Omni 1 M Beadchip or Affymetrix (Fremont, CA, USA) Genome-Wide Human SNP 6.0 arrays according to the manufacturer's protocols. Copy number calling was performed using published algorithms (45) and PennCNV (46). Inspection of copy number variants was performed for Xq13.1 by analyzing allele frequency and log R ratio values using Illumina BeadStudio (ver. 3.1.3) or Affymetrix Chromosome Analysis Suite (version 1.0) software as described (47). For case R19H7, genome-wide analysis of DNA copy number aberrations was carried out using the Roche Nimblegen 3×720K whole-genome array (median probe spacing of ~ 2 kb) according to the manufacturer's instructions, modified as follows: 500 ng of genomic DNA from patient and sex-matched control samples (pool of five samples) was labeled for 2 h followed by hybridization for 72 h. After washing, slides were scanned using the Roche MS200 scanner, and analyzed using the software NimbleScan (Roche Nimblegen). The CGH-segMNT module of NimbleScan was used for the analysis with a minimum segment length of five probes and an averaging window of 20 kb.

Exome sequencing

Exome capture and sequencing was performed in several clinically validated diagnostic labs. Subject HD8HC01P exome sequencing and copy number analysis were performed using the UW CMG CoNIFER pipeline (48).

Reference sequences and HDAC8 conservation analysis

HDAC8 accession numbers used to include NM_018486.2 (mRNA), NP_060956 (RefSeq protein) and Q9By41 (UniProt). HDAC8 protein sequences for human (AAF73428), *Bos taurus* (DAA12953), rat (AAI62023), mouse (CAM17598), *Danio rerio* (NP_998596), *Xenopus laevis* (NP_001085711), *Xenopus tropicalis* deduced from mRNA BC161282, *Drosophila melanogaster* (AAC61494) and for *Saccharomyces cerevisiae* HDAC8-like proteins, HOS1 (Q12214) and RPD3 (AAT92832) were aligned by the ClustalW method (49) using MacVector software (Accelrys Corp, San Diego, CA, USA).

Mapping mutations to HDAC8 structure

Identified HDAC8 mutations were mapped onto the crystal structure data of human HDAC8-substrate complex (PDB accession code 2V5W) (27) to visualize location with respect to the enzyme active site using Cn3D (50) and PyMol (51) software.

Recombinant HDAC8 purification from *E. coli*

Missense human *HDAC8* mutations were introduced into a previously described HDAC8–6His-pET20b construct (30) using QuikChange site-directed mutagenesis (Agilent Genomics, Santa Clara, CA, USA). Oligonucleotide sequences are available upon request. Recombinant wild-type HDAC8 and mutants I19S, H42P, H71Y, P91L, G117E, C153F, H180R, A188 T, D233G, D237Y, ΔK239-Y240N, I243N, G304R, T311M, G320R and H334R were expressed in BL21(DE3) *E. coli* cells and purified according to a previously published procedure (30), with minor modifications. Briefly, 50 ml cultures (LB media supplemented with 100 μg/l ampicillin) were grown overnight and used to inoculate 1 l flasks (minimal media supplemented with 100 μg/l ampicillin). Typically, 2–6 l were expressed for each mutant. Cells were grown at 37°C until OD$_{600}$ $\sim$0.5, at which point the temperature was lowered to 18°C. After 30 min, cells were induced by the addition of isopropyl β-D-thiogalactopyranoside (0.4 mM final concentration) and zinc chloride (100 μM final concentration), and grown overnight at 18°C. The cells were pelleted by centrifugation and kept at -80°C until purification. After thawing, cells were resuspended in $\sim$ 30 ml of lysis buffer (50 mM Tris (pH 8.0), 500 mM KCl, 5% glycerol, 3 mM β-mercaptoethanol (BME) and 115 μM phenylmethanesulfonyl fluoride), and lysed by sonication on ice. Cell debris was pelleted by centrifugation and the supernatant was purified by affinity chromatography (Talon resin, Clontech Labs) using an imidazole gradient. After loading the cell-free extract in buffer A (50 mM Tris (pH 8.0), 500 mM KCl, 5% glycerol, 3 mM BME), the column was washed with 4% buffer B (50 mM Tris (pH 8.0), 500 mM KCl, 250 mM imidazole, 5% glycerol, 3 mM BME) and the protein was eluted with a 4–80% linear gradient of buffer B. Fractions containing the protein were pooled and concentrated to $\sim$ 500 μl by ultrafiltration. The protein was then dialyzed at 4°C against buffer C (50 mM Tris (pH 8.0), 150 mM KCl, 5% glycerol, 1 mM dithiothreitol). Protein concentrations were determined from the absorbance using a calculated extinction coefficient: 50 240 M^{-1} cm^{-1} for the wild type and all the mutants except H71Y (51 520 M^{-1}cm^{-1}), C153F (50 180 M^{-1} cm^{-1}), D237Y (51 520 M^{-1} cm^{-1}) and ΔK239-Y240N (48 960 M^{-1} cm^{-1}) (52). The H42P mutant did not express a soluble protein under these conditions.

HDAC8 activity assays

Enzymatic activity of purified proteins was measured using the commercially available Fluor-de-Lys HDAC8 deacetylase substrate (BML-KI178-0005) and Developer II (BML-KI176-1250) from Enzo Life Sciences. In this assay, deacetylation of the tetrapeptide substrate (acetyl)-L-Arg-L-His-L-Lys(ε-acetyl)-L-Lys(ε-acetyl)-aminomethylcoumarin allows a protease developer to cleave the amide bond linking the C-terminal aminomethylcoumarin to the peptide backbone, which results in a fluorescence

shift. All assays were run at 25°C in assay buffer (25 mM Tris (pH = 8.2), 137 mM NaCl, 2.7 mM KCl, 1 mM MgCl$_2$) and contained 150 μM substrate with the following enzyme concentrations: 0.5 μM (wild type, P91L, H334R), 1.0 μM (A188 T), 1.5 μM (I19S, I243N, D233G, D237Y, G320R), 2.5 μM (H71Y, G117E, H180R, ΔK239-Y240N, G304R, T311M) or 3 μM (C153F) enzyme. After 30 min, reactions were quenched by the addition of the HDAC8 inhibitor M344 (Sigma-Aldrich, 100 μM) and Developer II. The fluorescence of both remaining substrate and formed product was measured using a Fluoroskan II plate reader (excitation = 355 nm, emission = 460 nm). Product concentration was calculated from raw data using standard curves for substrate and product (Aminomethylcoumarin, Enzo Life Sciences), with the total concentration of substrate and product kept at 150 μM (before dilution with the developer solution). All assays were performed in triplicate.

Tissue culture

Lymphoblastoid cell lines (LCLs) were cultured in RPMI 1640 supplemented to 20% FBS and 2% 20 mM L-glutamine, 100 units/ml penicillin and 100 μg/ml streptomycin at 37°C in 5% CO$_2$. HeLa cells were cultured at 37°C and 5% CO$_2$ in DMEM supplemented to 10% FBS. Nonsense-mediated decay was inhibited using 1 mg/ml cycloheximide (Sigma-Aldrich, St. Louis, MO, USA) in culture media for 6 h prior to harvesting RNA.

Assessment of X-inactivation

X chromosome inactivation (XCI) was determined by evaluating the methylation status of the CAG microsatellite locus at the 5′ end of the AR gene as previously described with modifications (35,53). Briefly, genomic DNA was isolated from peripheral blood or cell lines and for each sample, two reaction digests were performed. In one reaction, 1 μg of DNA was digested in 25 μl with the methylation-sensitive restriction enzyme *Hpa*II (New England Biolabs), to cut the active unmethylated allele. In the other reaction, DNA was incubated in enzyme digest buffer without enzyme. After a 16 h incubation at 37°C, digestion was terminated by incubation at 65°C for 20 min. From each reaction, 2 μl was then amplified by PCR with primers flanking the polymorphic androgen-receptor CAG repeat as described (35); forward 5′-CTGTGAAGGTTGCTGTTCCTCAT-3′; reverse 5′-FAM-TCCAGAATCTGTTCCAGAGCGTGC-3′). ABI3730 Genetic Analyzer and GeneMapper V4.0 software (Applied Biosystems) was used for genotyping analysis. Percent X-inactivation was calculated by dividing the ratio of the allele peak volumes in the *Hpa*II-treated by the ratio of the allele peak volumes in the untreated sample. We considered alleles separated by more than two trinucleotide repeats as informative (54) and used a cutoff of >80:20% for skewed XCI and >95:5% for extremely skewed XCI.

SUPPLEMENTARY MATERIAL

Supplementary Material is available at *HMG* online.

AUTHORS' CONTRIBUTIONS

F.J.K, G.G.-K., R.H., J.W., F.J.R, L.G.J., J.P., D.R.F, I.D.K. and M.A.D initiated the human studies. D.J.A., P.S.A., H.G.B., D.C., M.dC.-C., N.D., H.D., J.C.F, E.G.-N., Y.G., B.D.H., R.H., L.H., H.I., A.D.K., Z.K., K.L., S.A.L., G.R.M., S.G.M., I.M., S.M., E.M., G.R.M., L.N., A.P.-A., L.S.P., M.B.P., A.P., B.P., F.Q.-R., N.R., E.R., S.S., A.E.S., K.L.S., V.M.S., Z.S., H.T.,

REFERENCES

1. Bhuiyan, Z., Klein, M., Hammond, P., Mannens, M.M., Van Haeringen, A., Van Berckelaer-Onnes, I. and Hennekam, R.C. (2005) Genotype–phenotype correlations of 39 patients with Cornelia de Lange syndrome: the Dutch experience. *J. Med. Genet.*, **43**, 568–575.

2. Gillis, L.A., McCallum, J., Kaur, M., DeScipio, C., Yaeger, D., Mariani, A., Kline, A.D., Li, H.H., Devoto, M., Jackson, L.G. *et al.* (2004) NIPBL mutational analysis in 120 individuals with Cornelia de Lange syndrome and evaluation of genotype-phenotype correlations. *Am. J. Hum. Genet.*, **75**, 610–623.

3. Krantz, I.D., McCallum, J., DeScipio, C., Kaur, M., Gillis, L.A., Yaeger, D., Jukofsky, L., Wasserman, N., Bottani, A., Morris, C.A. *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.

4. Tonkin, E.T., Wang, T.J., Lisgo, S., Bamshad, M.J. and 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.

5. Yan, J., Saifi, G.M., Wierzba, T.H., Withers, M., Bien-Willner, G.A., Limon, J., Stankiewicz, P., Lupski, J.R. and Wierzba, J. (2006) Mutational and genotype-phenotype correlation analyses in 28 Polish patients with Cornelia de Lange syndrome. *Am. J. Med. Genet. A*, **140**, 1531–1541.

6. Selicorni, A., Russo, S., Gervasini, C., Castronovo, P., Milani, D., Cavalleri, F., Bentivegna, A., Masciadri, M., Domi, A., Divizia, M.T. *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.

7. Pie, J., Gil-Rodriguez, M.C., Ciero, M., Lopez-Vinas, E., Ribate, M.P., Arnedo, M., Deardorff, M.A., Puisac, B., Legarreta, J., de Karam, J.C. *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.

8. Huisman, S.A., Redeker, E.J., Maas, S.M., Mannens, M.M. and Hennekam, R.C. (2013) High rate of mosaicism in individuals with Cornelia de Lange syndrome. *J. Med. Genet.*, **50**, 339–344.

9. Rollins, R.A., Morcillo, P. and Dorsett, D. (1999) Nipped-B, a Drosophila homologue of chromosomal adherins, participates in activation by remote enhancers in the cut and Ultrabithorax genes. *Genetics*, **152**, 577–593.

10. Ciosk, R., Shirayama, M., Shevchenko, A., Tanaka, T., Toth, A. and Nasmyth, K. (2000) Cohesin's binding to chromosomes depends on a separate complex consisting of Scc2 and Scc4 proteins. *Mol. Cell*, **5**, 243–254.

11. Gillespie, P.J. and Hirano, T. (2004) Scc2 couples replication licensing to sister chromatid cohesion in Xenopus egg extracts. *Curr. Biol.*, **14**, 1598–1603.

12. Takahashi, T.S., Yiu, P., Chou, M.F., Gygi, S. and Walter, J.C. (2004) Recruitment of Xenopus Scc2 and cohesin to chromatin requires the pre-replication complex. *Nat. Cell Biol.*, **6**, 991–996.

13. Nasmyth, K. and Haering, C.H. (2009) Cohesin: its roles and mechanisms. *Annu. Rev. Genet.*, **43**, 525–558.

14. Rohatgi, S., Clark, D., Kline, A.D., Jackson, L.G., Pie, J., Siu, V., Ramos, F.J., Krantz, I.D. and Deardorff, M.A. (2010) Facial diagnosis of mild and variant CdLS: insights from a dysmorphologist survey. Am. J. Med. Genet. A, *152A*, 1641–1653.

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

16. Deardorff, M.A., Kaur, M., Yaeger, D., Rampuria, A., Korolev, S., Pie, J., Gil-Rodriguez, C., Arnedo, M., Loeys, B., Kline, A.D. *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.

17. Deardorff, M.A., Wilde, J.J., Albrecht, M., Dickinson, E., Tennstedt, S., Braunholz, D., Monnich, M., Yan, Y., Xu, W., Gil-Rodriguez, M.C. *et al.* (2012) RAD21 mutations cause a human cohesinopathy. *Am. J. Hum. Genet.*, **90**, 1014–1027.

18. Zhang, J., Shi, X., Li, Y., Kim, B.J., Jia, J., Huang, Z., Yang, T., Fu, X., Jung, S.Y., Wang, Y. *et al.* (2008) Acetylation of Smc3 by Eco1 is required for S phase sister chromatid cohesion in both human and yeast. *Mol. Cell*, **31**, 143–151.

19. Rolef Ben-Shahar, T., Heeger, S., Lehane, C., East, P., Flynn, H., Skehel, M. and Uhlmann, F. (2008) Eco1-dependent cohesin acetylation during establishment of sister chromatid cohesion. *Science*, **321**, 563–566.

20. Unal, E., Heidinger-Pauli, J.M., Kim, W., Guacci, V., Onn, I., Gygi, S.P. and Koshland, D.E. (2008) A molecular determinant for the establishment of sister chromatid cohesion. *Science*, **321**, 566–569.

21. Sutani, T., Kawaguchi, T., Kanno, R., Itoh, T. and Shirahige, K. (2009) Budding yeast Wpl1(Rad61)-Pds5 complex counteracts sister chromatid cohesion-establishing reaction. *Curr. Biol.*, **19**, 492–497.

22. Beckouet, F., Hu, B., Roig, M.B., Sutani, T., Komata, M., Uluocak, P., Katis, V.L., Shirahige, K. and Nasmyth, K. (2010) An Smc3 acetylation cycle is essential for establishment of sister chromatid cohesion. *Mol. Cell*, **39**, 689–699.

23. Xiong, B., Lu, S. and Gerton, J.L. (2010) Hos1 is a lysine deacetylase for the Smc3 subunit of cohesin. *Curr. Biol.*, **20**, 1660–1665.

24. Borges, V., Lehane, C., Lopez-Serra, L., Flynn, H., Skehel, M., Rolef Ben-Shahar, T. and Uhlmann, F. (2010) Hos1 deacetylates Smc3 to close the cohesin acetylation cycle. *Mol. Cell*, **39**, 677–688.

25. Deardorff, M.A., Bando, M., Nakato, R., Watrin, E., Itoh, T., Minamino, M., Saitoh, K., Komata, M., Katou, Y., Clark, D. *et al.* (2012) HDAC8 mutations in Cornelia de Lange syndrome affect the cohesin acetylation cycle. *Nature*, **489**, 313–317.

26. Harakalova, M., van den Boogaard, M.J., Sinke, R., van Lieshout, S., van Tuil, M.C., Duran, K., Renkens, I., Terhal, P.A., de Kovel, C., Nijman, I.J. *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.

27. Vannini, A., Volpari, C., Gallinari, P., Jones, P., Mattu, M., Carfi, A., De Francesco, R., Steinkuhler, C. and Di Marco, S. (2007) Substrate binding to histone deacetylases as shown by the crystal structure of the HDAC8-substrate complex. *EMBO Rep.*, **8**, 879–884.

28. Wang, D.F., Helquist, P., Wiech, N.L. and Wiest, O. (2005) Toward selective histone deacetylase inhibitor design: homology modeling, docking studies, and molecular dynamics simulations of human class I histone deacetylases. *J. Med. Chem.*, **48**, 6936–6947.

29. Haider, S., Joseph, C.G., Neidle, S., Fierke, C.A. and Fuchter, M.J. (2011) On the function of the internal cavity of histone deacetylase protein 8: R37 is a crucial residue for catalysis. *Bioorg. Med. Chem. Lett.*, **21**, 2129–2132.

30. Dowling, D.P., Gantt, S.L., Gattis, S.G., Fierke, C.A. and Christianson, D.W. (2008) Structural studies of human histone deacetylase 8 and its site-specific variants complexed with substrate and inhibitors. *Biochemistry*, **47**, 13554–13563.

31. Wegener, D., Hildmann, C., Riester, D. and Schwienhorst, A. (2003) Improved fluorogenic histone deacetylase assay for high-throughput-screening applications. *Anal. Biochem.*, **321**, 202–208.

32. Lombardi, P.M., Cole, K.E., Dowling, D.P. and Christianson, D.W. (2011) Structure, mechanism, and inhibition of histone deacetylases and related metalloenzymes. *Curr. Opin. Struct. Biol.*, **21**, 735–743.

33. Somoza, J.R., Skene, R.J., Katz, B.A., Mol, C., Ho, J.D., Jennings, A.J., Luong, C., Arvai, A., Buggy, J.J., Chi, E. *et al.* (2004) Structural snapshots of human HDAC8 provide insights into the class I histone deacetylases. *Structure*, **12**, 1325–1334.

34. Cole, K.E., Dowling, D.P., Boone, M.A., Phillips, A.J. and Christianson, D.W. (2011) Structural basis of the antiproliferative activity of largazole, a depsipeptide inhibitor of the histone deacetylases. *J. Am. Chem. Soc.*, **133**, 12474–12477.

35. Allen, R.C., Zoghbi, H.Y., Moseley, A.B., Rosenblatt, H.M. and Belmont, J.W. (1992) Methylation of HpaII and HhaI sites near the polymorphic CAG repeat in the human androgen-receptor gene correlates with X chromosome inactivation. *Am. J. Hum. Genet.*, **51**, 1229–1239.

36. Kuczmarski, R.J., Ogden, C.L., Guo, S.S., Grummer-Strawn, L.M., Flegal, K.M., Mei, Z., Wei, R., Curtin, L.R., Roche, A.F. and Johnson, C.L. (2002) 2000 CDC Growth Charts for the United States: methods and development. *Vital Health Stat*. **11**, 1–190.

37. Kline, A.D., Barr, M. and Jackson, L.G. (1993) Growth manifestations in the Brachmann-de Lange syndrome. *Am. J. Med. Genet.*, **47**, 1042–1049.

38. Haberland, M., Mokalled, M.H., Montgomery, R.L. and Olson, E.N. (2009) Epigenetic control of skull morphogenesis by histone deacetylase 8. *Genes Dev.*, **23**, 1625–1630.

39. Kawauchi, S., Calof, A.L., Santos, R., Lopez-Burks, M.E., Young, C.M., Hoang, M.P., Chua, A., Lao, T., Lechner, M.S., Daniel, J.A. *et al.* (2009) Multiple organ system defects and transcriptional dysregulation in the Nipbl(+/−) mouse, a model of Cornelia de Lange syndrome. *PLoS Genet.*, **5**, e1000650.

40. Xu, H., Balakrishnan, K., Malaterre, J., Beasley, M., Yan, Y., Essers, J., Appeldoorn, E., Tomaszewski, J.M., Vazquez, M., Verschoor, S. *et al.* (2010) Rad21-cohesin haploinsufficiency impedes DNA repair and enhances gastrointestinal radiosensitivity in mice. *PLoS ONE*, **5**, e12112.

41. Remeseiro, S., Cuadrado, A., Carretero, M., Martinez, P., Drosopoulos, W.C., Canamero, M., Schildkraut, C.L., Blasco, M.A. and Losada, A. (2012) Cohesin-SA1 deficiency drives aneuploidy and tumourigenesis in mice due to impaired replication of telomeres. *EMBO J.*, **31**, 2076–2089.

42. Kline, A.D., Krantz, I.D., Sommer, A., Kliewer, M., Jackson, L.G., FitzPatrick, D.R., Levin, A.V. and Selicorni, A. (2007) Cornelia de Lange syndrome: clinical review, diagnostic and scoring systems, and anticipatory guidance. *Am. J. Med. Genet. A*, **143**, 1287–1296.

43. Deardorff, M.A., Clark, D.M. and Krantz, I.D. (2011), *In GeneReviews at GeneTests: Medical Genetics Information Resource (Database Online)*. University of Washington, Seattle, in press.

44. Strom, T. (2006), ExonPrimer (database online).

45. Shaikh, T.H., Gai, X., Perin, J.C., Glessner, J.T., Xie, H., Murphy, K., O'Hara, R., Casalunovo, T., Conlin, L.K., D'Arcy, M. *et al.* (2009) High-resolution mapping and analysis of copy number variations in the human genome: a data resource for clinical and research applications. *Genome Res.*, **19**, 1682–1690.

46. Wang, K., Li, M., Hadley, D., Liu, R., Glessner, J., Grant, S.F., Hakonarson, H. and Bucan, M. (2007) PennCNV: an integrated hidden Markov model designed for high-resolution copy number variation detection in whole-genome SNP genotyping data. *Genome Res.*, **17**, 1665–1674.

47. Peiffer, D.A., Le, J.M., Steemers, F.J., Chang, W., Jenniges, T., Garcia, F., Haden, K., Li, J., Shaw, C.A., Belmont, J. *et al.* (2006) High-resolution genomic profiling of chromosomal aberrations using Infinium whole-genome genotyping. *Genome Res.*, **16**, 1136–1148.

48. Krumm, N., Sudmant, P.H., Ko, A., O'Roak, B.J., Malig, M., Coe, B.P., Project, N.E.S., Quinlan, A.R., Nickerson, D.A. and Eichler, E.E. (2012) Copy number variation detection and genotyping from exome sequence data. *Genome Res.*, **22**, 1525–1532.

49. Eddy, S.R. (1995) Multiple alignment using hidden Markov models. *Proc. Int. Conf. Intell. Syst. Mol. Biol.*, **3**, 114–120.

50. Wang, Y., Geer, L.Y., Chappey, C., Kans, J.A. and Bryant, S.H. (2000) Cn3D: sequence and structure views for Entrez. *Trends Biochem. Sci.*, **25**, 300–302.

51. Jessberger, R. (2002) The many functions of SMC proteins in chromosome dynamics. *Nat. Rev. Mol. Cell. Biol.*, **3**, 767–778.

52. Gill, S.C. and von Hippel, P.H. (1989) Calculation of protein extinction coefficients from amino acid sequence data. *Anal. Biochem.*, **182**, 319–326.

53. Wang, X., Sutton, V.R., Omar Peraza-Llanes, J., Yu, Z., Rosetta, R., Kou, Y.C., Eble, T.N., Patel, A., Thaller, C., Fang, P. *et al.* (2007) Mutations in X-linked PORCN, a putative regulator of Wnt signaling, cause focal dermal hypoplasia. *Nat. Genet.*, **39**, 836–838.

54. Amos-Landgraf, J.M., Cottle, A., Plenge, R.M., Friez, M., Schwartz, C.E., Longshore, J. and Willard, H.F. (2006) X chromosome-inactivation patterns of 1,005 phenotypically unaffected females. *Am. J. Hum. Genet.*, **79**, 493–499.