(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
12 July 2018 (12.07.2018)

WIPO | PCT

(10) International Publication Number
WO 2018/128587 A1

(51) International Patent Classification:
C07K 14/47 (2006.01) *C12N 15/00* (2006.01)
A61K 38/17 (2006.01) *A61P 21/00* (2006.01)

(21) International Application Number:
PCT/SG2018/050008

(22) International Filing Date:
08 January 2018 (08.01.2018)

(25) Filing Language: English

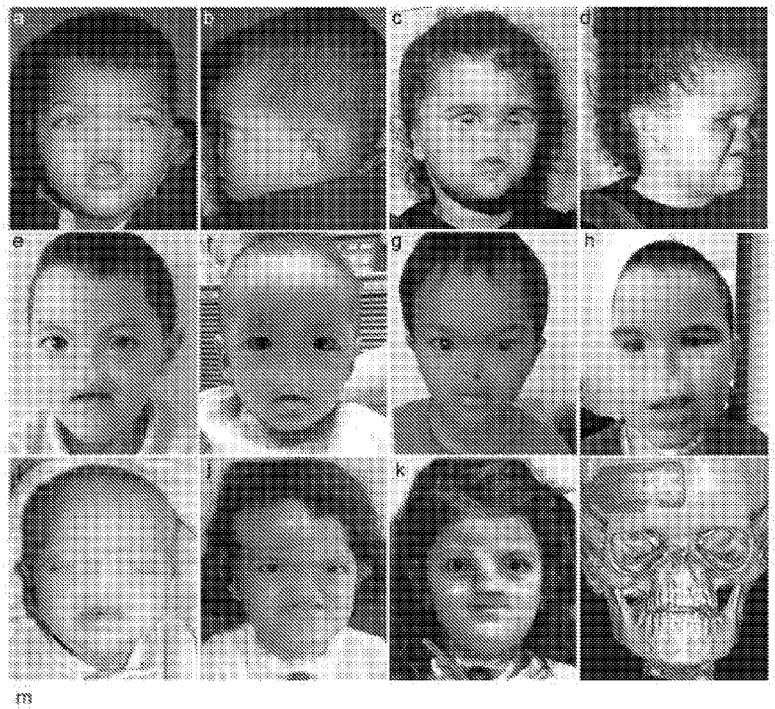
(26) Publication Language: English

(30) Priority Data:
10201700129W 06 January 2017 (06.01.2017) SG

(71) Applicants: **AGENCY FOR SCIENCE, TECHNOLOGY AND RESEARCH** [SG/SG]; 1 Fusionopolis Way, #20-10 Connexis North Tower, Singapore 138632 (SG). **THE WALTER AND ELIZA HALL INSTITUTE OF MEDICAL RESEARCH** [—/AU]; 1G Royal Parade, Parkville, Victoria 3052 (AU). **INSERM (INSITUT NATIONAL DE LA SANTÉ ET DE LA RECHERCHE MÉDICALE)** [—/FR]; 101, rue de Tolbiac, 75013 Paris (FR). **UNIVERSITÄTSMEDIZIN GÖTTINGEN** [—/DE]; Georg-August-Universität, Robert-Koch-Straße 40, 37075 Göttingen (DE).

(72) Inventors: **REVERSADE, Bruno**; c/o Institute of Medical Biology, 8A Biomedical Grove, #06-06 Immunos, Singapore 138648 (SG). **AMIEL, Jeanne**; c/o INSERM (Insitut National De La Santé Et De La Recherche Médicale), 101, rue de Tolbiac, 75013 Paris (FR). **BLEWITT, Marnie**; c/o The Walter and Eliza Hall Institute of Medical Research, 1G Royal Parade, Parkville, Victoria 3052 (AU). **MURPHY, James**; c/o The Walter and Eliza Hall Institute of Medical Research, 1G Royal Parade, Parkville, Victoria 3052 (AU). **WOLLNIK, Bernd**; c/o Universitätsmedizin Göttingen Georg-August-Universität, Robert-Koch-Straße 40, 37075 Göttingen (DE). **JAVED, Asif**; c/o Genome Institute of Singapore, 60 Biopolis Street, Genome, #02-01, Singapore 138672 (SG).

(54) Title: MUTANT SMCHD1 FOR THERAPY



(57) Abstract: Disclosed is an isolated variant of SMCHD1 peptide, or fragment thereof, comprising one or more mutation(s) that increases SMCHD1 activity in a cell compared to a cell with a wild- etype SMCHD1 peptide. Also disclosed are isolated nucleic acid encoding the variant SMCHD1 peptide as described herein, a vector comprising the nucleic acid as described herein, a pharmaceutical composition comprising the peptide, nucleic acid, or vector as described herein, and a host cell comprising the vector as described herein. Also disclosed are a method of treating facioscapulohumeral muscular dystrophy (FSHD) in a subject in need thereof and a method of screening an agent capable of increasing SMCHD1 activity.

WO 2018/128587 A1

[Continued on next page]

(74) **Agent: SPRUSON & FERGUSON (ASIA) PTE LTD**; P.O. Box 1531, Robinson Road Post Office, Singapore 903031 (SG).

(81) **Designated States** *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) **Designated States** *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:

— *of inventorship (Rule 4.17(iv))*

Published:

— *with international search report (Art. 21(3))*
— *with sequence listing part of description (Rule 5.2(a))*
— *in black and white; the international application as filed contained color or greyscale and is available for download from PATENTSCOPE*

MUTANT SMCHD1 FOR THERAPY

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] This application claims the benefit of priority of Singapore patent application No. 10201700129W, filed 6 January 2017, the contents of it being hereby incorporated by reference in its entirety for all purposes.

FIELD OF THE INVENTION

[0002] The present invention generally relates to a peptide or fragment thereof. In particular, the present invention relates to a variant peptide for use in therapy.

BACKGROUND OF THE INVENTION

[0003] Facioscapulohumeral muscular dystrophy (FSHD) is a disorder characterized by progressive weakness and wasting of muscles of face, neck, and shoulder girdle. FSHD has variable age onset of between 10 to 30 years of age and the severity of the condition varies from milder cases where dystrophy is not observed until later in life to rare severe cases where dystrophy is apparent in infancy and early childhood. In some subjects, muscle weakness can spread to abdominal muscles and sometimes hip muscles.

[0004] FSHD type 1 is an autosomal dominant inherited dystrophy and, in most cases, an affected subject inherits the altered chromosome from one affected parent. However, there were subject who have no history of the disorder in the family.

[0005] FSHD type 2 is a digenic inherited dystrophy where the subject must inherit a mutation in the SMCHD1 gene and a copy of the “permissive” chromosome 4.

[0006] Currently no cure is available and symptoms are managed by addressing each manifestation individually and with the hope that management strategy will delay the progression of the disorder. Thus, there is a clear need to provide for therapy for treating subjects having FSHD.

[0007] In view of the above, the present disclosure aims to provide a therapy for treating subjects having FSHD.

SUMMARY OF THE INVENTION

[0008] In one aspect, there is provided an isolated variant of SMCHD1 (structural maintenance of chromosomes flexible hinge domain containing 1) peptide or fragment

thereof, comprising one or more mutation(s) that increases SMCHD1 activity in a cell compared to a cell with a wild-type SMCHD1 peptide.

[0009] In one embodiment, the mutation is comprised in one or more regions selected from the group consisting of: the N-terminal motif I, which is the highly conserved region of GHKL-ATPases and participates in the coordination of the Mg2+-ATP complex during ATP hydrolysis, and the C-terminal to ATPase domain.

[0010] In one embodiment, the mutation is comprised in region between amino acid corresponding to residue 111 to amino acid corresponding to residue 702 of SMCHD1.

[0011] In one embodiment, the mutation is a substitution mutation.

[0012] In one embodiment, the mutation is at one or more position(s) selected from the group consisting of: amino acid residue corresponding to residue 134 of human SMCHD1, amino acid residue corresponding to residue 135 of human SMCHD1, amino acid residue corresponding to residue 136 of human SMCHD1, amino acid residue corresponding to residue 342 of human SMCHD1, amino acid residue corresponding to residue 348 of human SMCHD1, amino acid residue corresponding to residue 420 of human SMCHD1, amino acid residue corresponding to residue 518 of human SMCHD1, and amino acid residue corresponding to residue 552 of human SMCHD1 (wherein each residue position is referenced to SMCHD1 human protein UniProtKB identifier: A6NHR9).

[0013] In one embodiment, the mutation is one or more selected from the group consisting of: amino acid residue corresponding to residue 134 of human SMCHD1 has been substituted with an amino acid with a polar uncharged side chain, amino acid residue corresponding to residue 135 of human SMCHD1 has been substituted with an amino acid with polar uncharged side chain and/or an amino acid with sulfur side chain, amino acid residue corresponding to residue 136 of human SMCHD1 has been substituted with an amino acid with an aliphatic side chain, amino acid residue corresponding to residue 342 of human SMCHD1 has been substituted with an amino acid with polar uncharged side chain, amino acid residue corresponding to residue 348 of human SMCHD1 has been substituted with an amino acid with positively charged side chain, amino acid residue corresponding to residue 420 of human SMCHD1 has been substituted with an amino acid with hydrophobic side chain, amino acid residue corresponding to residue 518 of human SMCHD1 has been substituted with an amino acid with negatively charged side chain, and amino acid residue corresponding to residue 552 of human SMCHD1 has been substituted with an amino acid

with a negative charged side chain (wherein each residue position is referenced to SMCHD1 human protein UniProtKB identifier: A6NHR9).

[0014] In one embodiment, wherein the mutation is one or more selected from the group consisting of: amino acid residue corresponding to residue 134 of human SMCHD1 has been substituted with a serine residue, amino acid residue corresponding to residue 135 of human SMCHD1 has been substituted with a cysteine and/or a asparagine acid residue, amino acid residue corresponding to residue 136 of human SMCHD1 has been substituted with a glycine residue, amino acid residue corresponding to residue 342 of human SMCHD1 has been substituted with a serine residue, amino acid residue corresponding to residue 348 of human SMCHD1 has been substituted with an arginine residue, amino acid residue corresponding to residue 420 of human SMCHD1 has been substituted with a valine residue, amino acid residue corresponding to residue 518 of human SMCHD1 has been substituted with a glutamic acid residue, and amino acid residue corresponding to residue 552 of human SMCHD1 has been substituted with a glutamine residue (wherein each residue position is referenced to SMCHD1 human protein UniProtKB identifier: A6NHR9).

[0015] In yet another aspect, the present invention provides an isolated nucleic acid or fragment thereof encoding the isolated variant SMCHD1 peptide as described herein, wherein the nucleic acid comprises one or more mutation(s) that increases SMCHD1 protein activity in a cell compared to a cell with a wild-type SMCHD1 peptide.

[0016] In one embodiment, the SMCHD1 mutation is a missense mutation.

[0017] In one embodiment, the SMCHD1 mutation is at one or more position(s) selected from the group consisting of: c.407, c.403, c.404, c.1043, c.1259, c.1655, c.1552, c.1025, and c.400 (nucleic acid residue referenced to NCBI Reference Sequence: NM_015295.2).

[0018] In one embodiment, the SMCHD1 mutation is one or more selected from the group consisting of: c.407A>G, c.403A>T, c.404G>A, c.1043A>G, c.1259A>T, c.1655G>A, c.1552A>G, c.1025G>C, and c.400G>T (nucleic acid residue referenced to referenced to NCBI Reference Sequence: NM_015295.2).

[0019] In another aspect, the present invention provides a vector comprising the nucleic acid as described herein.

[0020] In another aspect, the present invention provides an isolated variant of SMCHD1 as described herein, or the isolated nucleic acid as described herein, or the vector as described herein for use in therapy.

[0021] In yet another aspect, the present invention provides a pharmaceutical composition comprising the isolated variant SMCHD1 peptide or fragment thereof as described herein, and/or an isolated nucleic acid or fragment thereof as described herein, or a vector as described herein, and a pharmaceutically acceptable excipient thereof.

[0022] In yet another aspect, the present invention provides a gene-therapy composition comprising an isolated nucleic acid as described herein or a vector as described herein.

[0023] In yet another aspect, the present invention provides a host cell comprising the vector as described herein.

[0024] In yet another aspect, the present invention provides a method of treating facioscapulohumeral muscular dystrophy (FSHD) in a subject in need thereof, comprising administering a therapeutically effective amount of an agent capable of increasing the expression level or activity of SMCHD1 peptide and/or nucleic acid to the subject in need thereof.

[0025] In one embodiment, the agent is one or more selected from the group consisting of an isolated variant SMCHD1 peptide or fragment thereof as described herein, a transgene encoding the isolated variant SMCHD1 peptide as described herein, an isolated nucleic acid or fragment thereof as described herein, a vector as described herein, a pharmaceutical composition as described herein, and a gene-therapy composition as described herein.

[0026] In one embodiment, the FSHD is a type 1 and/or type 2 FSHD.

[0027] In yet another aspect, the present invention provides a method of screening an agent (or a nucleotide mutation) capable of increasing SMCHD1 activity, comprising: a. introducing the agent (or nucleotide mutation) to an assay for determining SMCHD1 activity, b. determining the SMCHD1 activity the agent (or nucleotide mutation) elicits, and c. comparing the SMCHD1 activity the agent (or nucleotide mutation) elicits with the SMCHD1 activity elicited by one or more selected from the group consisting of an isolated variant SMCHD1 peptide or fragment thereof as described herein, and/or a transgene encoding the isolated variant SMCHD1 peptide as described herein, and/or an isolated nucleic acid or fragment thereof as described herein, and/or a vector as described herein, wherein when the SMCHD1 activity of the agent (or nucleotide mutation) is the same or more than the SMCHD1 activity of the one or more selected from the group consisting of the isolated variant SMCHD1 peptide or fragment thereof as described herein, and/or the transgene encoding the isolated variant SMCHD1 peptide as described herein, and/or the isolated nucleic acid or fragment thereof as described herein, and/or the vector as described herein,

the agent (or nucleotide mutation) is considered a suitable agent (or nucleotide mutation) for increasing SMCHD1 activity.

[0028] In another aspect, the present invention provides a use of one or more selected from the group consisting of an isolated variant SMCHD1 peptide as described herein, a transgene encoding the isolated variant SMCHD1 peptide as described herein, an isolated nucleic acid as described herein, a vector as described herein, a pharmaceutical composition as described herein, and a gene-therapy composition as described herein in the manufacture of a medicament for treating facioscapulohumeral muscular dystrophy (FSHD) in a subject in need thereof.

[0029] In yet another aspect, the present invention provides a kit for screening an agent capable of increasing SMCHD1 activity, comprising a reference control comprising one or more selected from the group consisting of an isolated variant SMCHD1 peptide as described herein, transgene encoding the isolated variant SMCHD1 peptide as described herein, an isolated nucleic acid as described herein, and a vector as described herein; and a reagent for determining SMCHD1 activity.

BRIEF DESCRIPTION OF THE DRAWINGS

[0030] The invention will be better understood with reference to the detailed description when considered in conjunction with the non-limiting examples and the accompanying drawings, in which:

[0031] **Figure 1** shows a series of photograph of patients having BAMS. (a,b) Patient 1. (c,d) Patient 12. (e) Patient 3. (f) Patient 9. (g) Patient 10. (h) Patient 6. (i-l) Patient 11, with forehead implant in preparation for rhinoplasty (rectangular box in j), 6 months post-operation (k) and computed tomography scan of the skull pre-operation (l). (m) Position of BAMS missense mutations (i.e. mutations listed below the bar) and heterozygous loss-of-function mutations from ExAC (i.e. mutations with asterisks (*)- indicated on above the bar) in SMCHD1. Figure 1 *SMCHD1* is mutated in Bosma arhinia microphthalmia syndrome (BAMS) and isolated arhinia. Figure 1 shows the physical manifestation of BAMS.

[0032] **Figure 2** shows graphs of the results of ATPase assays using the purified recombinant N-terminal region harboring BAMS or FSHD2 mutations. (a-e) ATPase assays performed using recombinant protein encompassing amino acids 111-702 of Smchd1. (a) wildtype, (b) p.Ala134Ser, (c) p.Ser135Cys, (d) p.Glu136Gly, (e) p.Tyr353Cys. The amount of ADP produced at each protein concentration (0.1, 0.2, 0.4 and 0.6 μM) and ATP

concentration (1, 2.5, 5 and 10 μM) was measured as described in the Online Methods. Data are displayed as mean ± s.d. of technical triplicates. Each plot is representative of at least two independent experiments using different batches of protein preparation. (f) Relative ATPase activities of the mutant proteins compared to wildtype protein. The amount of ADP produced by the mutant proteins was normalised to that of wildtype protein at each protein and substrate concentration as in (a-e). The normalised values are plotted as mean + s.d. of biological replicates (n=44 for p.Ala134Ser, n=24 for p.Ser135Cys, n=32 for p.Glu136Gly, p.Asp420Val, p.Tyr353Cys and p.Thr527Met). In addition to analyzing normalised fold changes, for each mutant the mean of the triplicates at each protein/ATP concentration was compared to that of wildtype using the Wilcoxon matched-pairs signed rank test; apart from p.Asp420Val with a p-value of 0.1776 (non-significant), all the other mutants had a p-value <0.0001 (significant). Thus, the biochemical assays in Figure 2 indicate that BAMS-associated mutations in SMCHD1 show increased ATPase activity.

[0033] **Figure 3** shows the results of *in vivo* functional assays in *Xenopus* embryos indicate that BAMS mutations behave as gain-of-function alleles. (a) qPCR shows that smchd1 is zygotically transcribed. Expression levels are shown relative to 18S rRNA. (b) In late tailbud stages, *smchd1* expression is restricted to the head region and the neural tube. (c) To target the head structures, the dorsal-animal blastomeres of the 8-cell stage *Xenopus* embryo were injected with synthesized mRNAs (120 pg for all panels except k). These cells are fated to give rise to head structures as revealed by Dextran lineage tracing. (d-f') Representative stage 45 tadpoles injected with SMCHD1A134S display craniofacial anomalies and smaller eyes compared to control and SMCHD1WT injected tadpoles. (g) Western blot of stage 12 embryonic extracts from control and injected embryos shows exogenous human SMCHD1 expression. (h) The eye diameter is significantly reduced in embryos overexpressing BAMS mutants (blue) relative to SMCHD1WT overexpressing siblings (black), or embryos overexpressing an FSHD2 mutant (open circles). n = at least 15 embryos for each condition. (i, j) *In situ* hybridization for rx2a, six1 and twist1, demarcating the eyes, placodes and neural crest respectively in embryos injected with SMCHD1WT (i) or SMCHD1A134S mRNA (j). Pictures are representative 453 of n= 9/10, 7/10, 10/10 embryos for each probe. Dotted lines outline the nasal placodes in middle panels and the eye in the right panels. Numbers label the streams of migrating cranial neural crest. (k) Measurements of eye diameter of *Xenopus* embryos injected with 0.5 ng or 1 ng wildtype or BAMS mutant SMCHD1 mRNA show that SMCHD1 overexpression causes dose-dependent craniofacial

anomalies. mRNAs injected for this panel did not contain a poly A signal and were polyadenylated in vitro, hence requiring higher RNA concentration to produce a phenotype (in other panels in this figure, and in Figure 13, the mRNAs contain a poly A signal allowing polyadenylation in vivo). Biological variation between clutches of frog tadpoles is seen in the data presented in panels h and k. n = 20 embryos for each condition. n.s. not significant, ***p< 0.001, ** p<0.01. Thus, Figure 3 are *in vivo* results that partially recapitulate the microphthalmia and facial hypoplasia seen in severe BAMS pateints, further support the notion that, in contrast to FSHD2 alleles, BAMS-associated missense mutations exhibit gain-of-function or neomorphic activity. In order words, Figure 3 provides support that the identification of gain-of-function mutations in SMCHD1 can be used in gene therapy approach for treatment of FSHD.

[0034] **Figure 4** shows photographic images of computed tomography and magnetic resonance imaging in BAMS. (a-c) Controls and (d-f) patient 1 at four years. Patient 1 displays maxillary hypoplasia and absent nasal bones (d and e). Olfactory bulbs and sulci (labelled with red and white arrows, respectively, on the left side in the control in c) are absent in patient 1 (f). Skeletal imaging of patients 14 (g,h) and 11 (i) indicates similar midface hyploplasia.

[0035] **Figure 5** shows BAMS pedigrees and Sanger sequencing chromatograms of SMCHD1 mutations. Individuals submitted for exome sequencing are indicated by a red asterisk. A question mark indicates no phenotypic information was available. Note Sanger sequencing was unavailable for individual 13.

[0036] **Figure 6** shows multiple sequence alignment of vertebrate SMCHD1orthologues and yeast Hsp90. Residues mutated in BAMS are indicated by pink arrows. Residues mutated in FSHD are indicated by purple arrows. Hs, Homo sapiens; Mm, Mus musculus; Bt, Bos taurus; Gg, Gallus gallus; Md, Monodelphis domestica; Cm, Chelonia mydas; Xt, Xenopus tropicalis; Dr, Danio rerio; Sc, Saccharomyces cerevisiae. FSHD mutation reference: LOVD SMCHD1 variant database, http://databases.lovd.nl/shared/variants/SMCHD1/unique.

[0037] **Figure 7** shows photographic images of the results of X-gal staining of mouse embryos expressing lacz from the Smchd1 locus. E, embryonic day. gt/+, embryos heterozygous for the Smchd1gt allele expressing lacz. +/+, wildtype embyros. hf, head folds. npl, nasal placode. ov, optic vesicle. npi, nasal pit. ne, nasal epithelium. f-i, coronal sections. r and s, transverse sections. An asterisk in panel p indicates deep nasal staining.

[0038] **Figure 8** shows bar graphs of results of sodium bisulfite sequencing in BAMS patients. a, individuals 1-6. b, individuals 8-11 and 14. The position of the three different regions analyzed within D4Z4 is indicated above the corresponding column (left, DR1; middle, 5'; right, Mid). For each sample, at least 10 cloned DNA molecules were analyzed by Sanger sequencing. Each histogram column corresponds to a single CpG. Black corresponds to the global percentage of methylated CpGs and white to the global percentage of unmethylated CpGs. The percentage of methylated CpGs among the total CpGs in each individual analyzed are given in Table 2.

[0039] **Figure 9** shows block graphs of comparison of D4Z4 methylation in BAMS or FSHD2 patients, BAMS patient relatives and controls. Distribution of methylation for the three different regions within the D4Z4 sequence (DR1, 5' and Mid) in control individuals, patients with FSHD2 carrying a SMCHD1 mutation and BAMS patients and their relatives. Means ± SEM are shown. A Kruskal-Wallis multiple comparisons test was performed, followed by a Dunn's test and Bonferroni correction, with $\alpha = 0.05$. ***, $p<0.0001$; **, $p<0.001$; *, $p<0.05$. Dot points indicate outliers. Crosses indicate medians. The level of methylation is statistically significantly different between controls and FSHD2 patients for the DR1 (** ; $p<0.001$) and the 5' (*** ; $p< 0.0001$) regions. The level of methylation is significantly different between controls and BAMS patients for the 5' region (*, $p<0.05$) and between BAMS patients and their relatives for the DR1 (*, $p<0.05$) and 5' (**, $p<0.001$) regions.

[0040] **Figure 10** shows Smchd1 structure modelling, based on the structure of Hsp90. Residues mutated in BAMS are indicated with an underline (i.e. W342, H346, R552, D420, K518, A134, E136, and S135). Residues mutated in FSHD are indicated in boxes (i.e. Y283, H263, L194, Y353, G425, G478, R479, T537, G137, P690, V615, and C492).

[0041] **Figure 11** shows fibroblasts derived from BAMS patients show no defects in NHEJ or in H2AX activation. In (a), a microhomology mediated end joining (MMEJ) assay was performed on wildtype (WT), XRCC4-deficient and case 1 and 2 fibroblasts. Whereas XRCC4-deficient fibroblasts show multiple smaller DNA bands after BstXI digestion indicating defects in NHEJ-mediated DNA repair and leading to preferential use of MMEJ-mediated DNA double strand repair, BAMS patient fibroblasts show no defects in NHEJ-mediated DNA repair pathways compared to wildtype. In (b), Western blot analysis of UV- and etoposide-induced phosphorylation of H2AX at Ser139 (γH2AX). Wildtype fibroblasts (WT) and fibroblasts derived from cases 1 and 2 were treated with UV-C (UV) or etoposide

(Eto) or left untreated as a control (-). Cells were lysed and subjected to Western blot analysis with an antibody against γH2AX. Equal protein loading was confirmed by reprobing of the membrane with an antibody against β-Actin. Wildtype and BAMS patient fibroblasts did not show significant differences in H2AX activation.

[0042] **Figure 12** shows ATPase assays performed using recombinant wildtype or mutant Smchd1 protein in the presence of radicicol. Data are displayed as mean ± s.d. of technical triplicates. The data are representative of at least two independent experiments using different batches of protein preparation. Figure 12 indicates that BAMS-associated mutations elevate the catalytic activity of SMCHD1.

[0043] **Figure 13** shows SMCHD1 overexpression in Xenopus causes dose dependent craniofacial anomalies. (a,b) Measurements of eye diameter of Xenopus embryos injected with 240 pg (a) or 500 pg (b) SMCHD1 mRNA. Y353C is an FSHD2 mutation. n = at least 20 embryos for each condition. (c-f) Representative Xenopus embryos injected with 500 pg of WT or FSHD2 mutant SMCHD1 or 120 pg of BAMS mutant mRNA show varying degrees of craniofacial abnormalities as compared to uninjected control tadpoles at 4 days post fertilization. n.s. not significant.

[0044] **Figure 14** shows SDS/PAGE gels showing the Purity of proteins used for ATPase assays. Purified recombinant wild type or mutant proteins were resolved by 4-20% (w/v) Tris-Glycine reducing SDS/PAGE and were stained with SimplyBlue SafeStain. Protein quantities loaded: left gel, 1.4 µg; middle gel, 1.05 µg; right gel, 0.7 µg. Molecular weight (MW) markers are as indicated on the left-hand side.

BRIEF DESCRIPTION OF THE TABLES

[0045] The invention will be better understood with reference to the detailed description when considered in conjunction with the non-limiting examples and the accompanying Tables, in which:

[0046] **Table 1** SMCHD1 mutations identified in patients 1-14.

[0047] **Table 2** DNA methylation analysis in BAMS probands and family members. Three different regions within the D4Z4 macrosatellite repeat were analyzed: DR1 (as described in Hartweck et al., Neurology, 2013); 5' and Mid (as described in Gaillard et al., Neurology, 2014). The Mid region corresponds to the DUX4 promoter. % M+ indicates the percentage of methylated CpGs among the total CpGs for a given region. X indicates samples

that were not analyzed. All samples were obtained from peripheral blood leukocytes except for individual 4's brother and sister and individual 14, which were from saliva.

[0048] **Table 3** Primers used for sodium bisulfite PCR.

[0049] **Table 4** Primers used for cloning and mutagenesis of recombinant murine *Smchd1*.

DETAILED DESCRIPTION OF THE PRESENT INVENTION

[0050] Bosma arhinia microphthalmia syndrome (BAMS) is an extremely rare and striking condition characterized by complete absence of the nose with or without ocular defects. The present disclosure shows that missense mutations in the extended ATPase domain of the epigenetic regulator *SMCHD1* causes BAMS in all 14 cases studied. All mutations were *de* novo where parental DNA was available. Biochemical tests and *in vivo* assays in *Xenopus* embryos suggest that these mutations behave as gain-of-function alleles. This is in contrast to loss-of-function mutations in SMCHD1 that have been associated with facioscapulohumeral muscular dystrophy (FSHD).

[0051] Given that loss-of-function mutations in SMCHD1 are associated with FSHD2, BAMS and FSHD2 represent a rare example of gain- versus loss-of-function mutations in the same gene leading to vastly different human disorders, in terms of the affected tissue and age of onset. Thus, the present disclosure shows that since FSHD is caused in part by a loss of SMCHD1, augmentation of the expression or activity of SMCHD1 in affected muscles can surprisingly be developed as a form of treatment. The identification of gain-of-function mutations in SMCHD1 in the present disclosure can inform gene therapy approaches, or in combination with future structural studies on the effect of these mutations on the ATPase domain, aid the design of drugs that reproduce the gain-of-function effect, for treatment of FSHD. Therefore, the present disclosure establishes SMCHD1 in its enzymatic function that can be exploited for the development of therapeutics for FSHD.

[0052] In the present disclosure, there is provided a method of treatment for facioscapulohumeral muscular dystrophy comprising administering a polypeptide or nucleotide sequence encoding SMCHD1 or a functional fragment thereof into the subject, or administering to the subject a compound capable of modifying the expression levels or activity of SMCHD1. In particular, the inventors of the present invention have identified mutations that can increase SMCHD1 activity. Therefore mimicking these mutations, or targeting these amino acids directly can be used as a way of increasing SMCHD1 activity.

[0053] Thus, in one aspect, there is provided an isolated variant of SMCHD1 (structural maintenance of chromosomes flexible hinge domain containing 1) peptide or fragment thereof, comprising one or more mutation(s) that increases SMCHD1 activity in a cell compared to a cell with a wild-type SMCHD1 peptide.

[0054] The term "isolated" is used herein to refer to polynucleotides, polypeptides and proteins that are extracted or obtained from other cellular proteins and is meant to encompass both purified and recombinant polypeptides. In some examples, the term "isolated" and “recombinant” are interchangeable and means separated from constituents, cellular and otherwise, in which the cell, tissue, polynucleotide, peptide, polypeptide, protein, antibody or fragment(s) thereof, which are normally associated in nature. For example, an isolated cell is a cell that is separated from tissue or cells of dissimilar phenotype or genotype. An isolated polynucleotide is separated from the 3' and 5' contiguous nucleotides with which it is normally associated in its native or natural environment, e.g., on the chromosome.

[0055] The term “mutation” as used herein is meant to include all kinds of nuclear and/or mitochondrial gene mutations, including any alteration in the base sequence of a DNA strand compared to the wild-type reference strand. In the present disclosure, the mutation is comprised in one or more regions selected from the group consisting of: the N-terminal motif I, which is the highly conserved region of GHKL-ATPases and participates in the coordination of the Mg^{2+}-ATP complex during ATP hydrolysis, and the C-terminal to ATPase domain.

[0056] In some examples, the mutation is comprised on both of the regions of the N-terminal motif I, which is the highly conserved region of GHKL-ATPases and participates in the coordination of the Mg^{2+}-ATP complex during ATP hydrolysis, and the C-terminal to ATPase domain.

[0057] In some examples, the mutation is comprised in region between amino acid corresponding to residue 111 to amino acid corresponding to residue 702 of SMCHD1 (wherein each residue position is referenced to SMCHD1 human protein UniProtKB identifier: A6NHR9, or SEQ ID NO: 1).

[0058] In some examples, the wildtype SMCHD1 has the sequence SEQ ID NO: 1, as follows:

MAAADGGGPGGASVGTEEDGGGVGHRTVYLFDRREKESELGDRPLQVGERSDYAGFRACVCQ
TLGISPEEKFVITTTSRKEITCDNFDETVKDGVTLYLLQSVNQLLLTATKERIDFLPHYDTL
VKSGMYEYYASEGQNPLPFALAELIDNSLSATSRNIGVRRIQIKLLFDETQGKPAVAVIDNG
RGMTSKQLNNWAVYRLSKFTRQGDFESDHSGYVRPVPVPRSLNSDISYFGVGGKQAVFFVGQ

SARMISKPADSQDVHELVLSKEDFEKKEKNKEAIYSGYIRNRKPSDSVHITNDDERFLHHLI
IEEKEKDSFTAVVITGVQPEHIQYLKNYFHLWTRQLAHIYHYYIHGPKGNEIRTSKEVEPFN
NIDIEISMFEKGKVPKIVNLREIQDDMQTLYVNTAADSFEFKAHVEGDGVVEGIIRYHPFLY
DRETYPDDPCFPSKLKDEDDEDDCFILEKAARGKRPIFECFWNGRLIPYTSVEDFDWCTPPK
KRGLAPIECYNRISGALFTNDKFQVSTNKLTFMDLELKLKDKNTLFTRILNGQEQRMKIDRE
FALWLKDCHEKYDKQIKFTLFKGVITRPDLPSKKQGPWATYAAIEWDGKIYKAGQLVKTIKT
LPLFYGSIVRFFLYGDHDGEVYATGGEVQIAMEPQALYDEVRTVPIAKLDRTVAEKAVKKYV
EDEMARLPDRLSVTWPEGDELLPNEVRPAGTPIGALRIEILNKKGEAMQKLPGTSHGGSKKL
LVELKVILHSSSGNKEIISHISQHGGKWPYWFKKMENIQKLGNYTLKLQVVLNESNADTYAG
RPLPSKAIKFSVKEGKPEKFSFGLLDLPFRVGVPFNIPLEFQDEFGHTSQLVTDIQPVLEAS
GLSLHYEEITKGPNCVIRGVTAKGPVNSCQGKNYNLKVTLPGLKEDSQILKIRLLPGHPRRL
KVKPDSEILVIENGTAFPFQVEVLDESDNITAQPKLIVHCKFSGAPNLPVYVVDCSSSGTSI
LTGSAIQVQNIKKDQTLKARIEIPSCKDVAPVEKTIKLLPSSHVARLQIFSVEGQKAIQIKH
QDEVNWIAGDIMHNLIFQMYDEGEREINITSALAEKIKVNWTPEINKEHLLQGLLPDVQVPT
SVKDMRYCQVSFQDDHVSLESAFTVRPLPDEPKHLKCEMKGGKTVQMGQELQGEVVIIITDQ
YGNQIQAFSPSSLSSLSIAGVGLDSSNLKTTFQENTQSISVRGIKFIPGPPGNKDLCFTWRE
FSDFIRVQLISGPPAKLLLIDWPELKESIPVINGRDLQNPIIVQLCDQWDNPAPVQHVKISL
TKASNLKLMPSNQQHKTDEKGRANLGVFSVFAPRGEHTLQVKAIYNKSIIEGPIIKLMILPD
PEKPVRLNVKYDKDASFLAGGLFTDFMISVISEDDSIIKNINPARISMKMWKLSTSGNRPPA
NAETFSCNKIKDNDKEDGCFYFRDKVIPNKVGTYCIQFGFMMDKTNILNSEQVIVEVLPNQP
VKLVPKIKPPTPAVSNVRSVASRTLVRDLHLSITDDYDNHTGIDLVGTIIATIKGSNEEDTD
TPLFIGKVRTLEFPFVNGSAEIMSLVLAESSPGRDSTEYFIVFEPRLPLLSRTLEPYILPFM
FYNDVKKQQQMAALTKEKDQLSQSIVMYKSLFEASQQLLNEMKCQVEEARLKEAQLRNELKI
HNIDIPTTQQVPHIEALLKRKLSEQEELKKKPRRSCTLPNYTKGSGDVLGKIAHLAQIEDDR
AAMVISWHLASDMDCVVTLTTDAARRIYDETQGRQQVLPLDSIYKKTLPDWKRSLPHFRNGK
LYFKPIGDPVFARDLLTFPDNVEHCETVFGMLLGDTIILDNLDAANHYRKEVVKITHCPTLL
TRDGDRIRSNGKFGGLQNKAPPMDKLRGMVFGAPVPKQCLILGEQIDLLQQYRSAVCKLDSV
NKDLNSQLEYLRTPDMRKKKQELDEHEKNLKLIEEKLGMTPIRKCNDSLRHSPKVETTDCPV
PPKRMRREATRQNRIITKTDV* (SEQ ID NO: 1)

[0059] As used herein, the term “mutation” is intended to include not only naturally occurring mutations, but also artificially introduced mutations. For example, the mutation can include random or site-specific mutations, including, but is not limited to, missense mutation, nonsense mutation, insertion, deletion, duplication, frameshift mutation, repeat expansion, and the like. In some examples, the mutation is missense mutation. In some example, the mutation leads to a substitution mutation. In some example, the mutation is a substitution mutation.

[0060] In some examples, the mutation as described herein is at one or more position(s) selected from the group consisting of: amino acid residue corresponding to residue 134 of human SMCHD1, amino acid residue corresponding to residue 135 of human SMCHD1, amino acid residue corresponding to residue 136 of human SMCHD1, amino acid residue corresponding to residue 342 of human SMCHD1, amino acid residue corresponding to residue 348 of human SMCHD1, amino acid residue corresponding to residue 420 of human SMCHD1, amino acid residue corresponding to residue 518 of human SMCHD1, and amino

acid residue corresponding to residue 552 of human SMCHD1 (wherein each residue position is referenced to SMCHD1 human protein UniProtKB identifier: A6NHR9 or SEQ ID NO: 1).

[0061] In some examples, the variant SMCHD1 as described herein may comprise at least one, or at least two, or at least three, or at least four, or at least five, or at least six, or at least seven, or at least eight mutations as described herein.

[0062] In some example, the mutation is one or more selected from the group consisting of: amino acid residue corresponding to residue 134 of human SMCHD1 has been substituted with an amino acid with a polar uncharged side chain, amino acid residue corresponding to residue 342 of human SMCHD1 has been substituted with an amino acid with polar uncharged side chain, amino acid residue corresponding to residue 348 of human SMCHD1 has been substituted with an amino acid with positively charged side chain, amino acid residue corresponding to residue 518 of human SMCHD1 has been substituted with an amino acid with negatively charged side chain, and amino acid residue corresponding to residue 552 of human SMCHD1 has been substituted with an amino acid with a negative charged side chain (wherein each residue position is referenced to SMCHD1 human protein UniProtKB identifier: A6NHR9; SEQ ID NO: 1).

[0063] In some examples, the mutation is one or more selected from the group consisting of: amino acid residue corresponding to residue 134 of human SMCHD1 has been substituted with a serine residue, amino acid residue corresponding to residue 342 of human SMCHD1 has been substituted with a serine residue, amino acid residue corresponding to residue 348 of human SMCHD1 has been substituted with an arginine residue, amino acid residue corresponding to residue 518 of human SMCHD1 has been substituted with a glutamic acid residue, and amino acid residue corresponding to residue 552 of human SMCHD1 has been substituted with a glutamine residue (wherein each residue position is referenced to SMCHD1 human protein UniProtKB identifier: A6NHR9; SEQ ID NO: 1).

[0064] In some example, the mutation is one or more selected from the group consisting of: amino acid residue corresponding to residue 134 of human SMCHD1 has been substituted with an amino acid with a polar uncharged side chain, amino acid residue corresponding to residue 135 of human SMCHD1 has been substituted with an amino acid with polar uncharged side chain and/or an amino acid with sulfur side chain, amino acid residue corresponding to residue 136 of human SMCHD1 has been substituted with an amino acid with an aliphatic side chain, amino acid residue corresponding to residue 342 of human SMCHD1 has been substituted with an amino acid with polar uncharged side chain, amino

acid residue corresponding to residue 348 of human SMCHD1 has been substituted with an amino acid with positively charged side chain, amino acid residue corresponding to residue 420 of human SMCHD1 has been substituted with an amino acid with hydrophobic side chain, amino acid residue corresponding to residue 518 of human SMCHD1 has been substituted with an amino acid with negatively charged side chain, and amino acid residue corresponding to residue 552 of human SMCHD1 has been substituted with an amino acid with a negative charged side chain (wherein each residue position is referenced to SMCHD1 human protein UniProtKB identifier: A6NHR9; SEQ ID NO: 1).

[0065] In some examples, the mutation is one or more selected from the group consisting of: amino acid residue corresponding to residue 134 of human SMCHD1 has been substituted with a serine residue, amino acid residue corresponding to residue 135 of human SMCHD1 has been substituted with a cysteine and/or a asparagine acid residue, amino acid residue corresponding to residue 136 of human SMCHD1 has been substituted with a glycine residue, amino acid residue corresponding to residue 342 of human SMCHD1 has been substituted with a serine residue, amino acid residue corresponding to residue 348 of human SMCHD1 has been substituted with an arginine residue, amino acid residue corresponding to residue 420 of human SMCHD1 has been substituted with a valine residue, amino acid residue corresponding to residue 518 of human SMCHD1 has been substituted with a glutamic acid residue, and amino acid residue corresponding to residue 552 of human SMCHD1 has been substituted with a glutamine residue (wherein each residue position is referenced to SMCHD1 human protein UniProtKB identifier: A6NHR9; SEQ ID NO: 1).

[0066] In some examples, the amino acid residue corresponding to residue 134 of human SMCHD1 has been substituted with a serine residue is SEQ ID NO: 2 (substituted residue underlined and in italics), having the following sequence:

MAAADGGGPGGASVGTEEDGGGVGHRTVYLFDRREKESELGDRPLQVGERSDYAGFRACVCQ
TLGISPEEKFVITTTSRKEITCDNFDETVKDGVTLYLLQSVNQLLLTATKERIDFLPHYDTL
VKSGMYEYY***S***SEGQNPLPFALAELIDNSLSATSRNIGVRRIQIKLLFDETQGKPAVAVIDNG
RGMTSKQLNNWAVYRLSKFTRQGDFESDHSGYVRPVPVPRSLNSDISYFGVGGKQAVFFVGQ
SARMISKPADSQDVHELVLSKEDFEKKEKNKEAIYSGYIRNRKPSDSVHITNDDERFLHHLI
IEEKEKDSFTAVVITGVQPEHIQYLKNYFHLWTRQLAHIYHYYIHGPKGNEIRTSKEVEPFN
NIDIEISMFEKGKVPKIVNLREIQDDMQTLYVNTAADSFEFKAHVEGDGVVEGIIRYHPFLY
DRETYPDDPCFPSKLKDEDDEDDCFILEKAARGKRPIFECFWNGRLIPYTSVEDFDWCTPPK
KRGLAPIECYNRISGALFTNDKFQVSTNKLTFMDLELKLKDKNTLFTRILNGQEQRMKIDRE
FALWLKDCHEKYDKQIKFTLFKGVITRPDLPSKKQGPWATYAAIEWDGKIYKAGQLVKTIKT
LPLFYGSIVRFFLYGDHDGEVYATGGEVQIAMEPQALYDEVRTVPIAKLDRTVAEKAVKKYV
EDEMARLPDRLSVTWPEGDELLPNEVRPAGTPIGALRIEILNKKGEAMQKLPGTSHGGSKKL
LVELKVILHSSSGNKEIISHISQHGGKWPYWFKKMENIQKLGNYTLKLQVVLNESNADTYAG

RPLPSKAIKFSVKEGKPEKFSFGLLDLPFRVGVPFNIPLEFQDEFGHTSQLVTDIQPVLEAS
GLSLHYEEITKGPNCVIRGVTAKGPVNSCQGKNYNLKVTLPGLKEDSQILKIRLLPGHPRRL
KVKPDSEILVIENGTAFPFQVEVLDESDNITAQPKLIVHCKFSGAPNLPVYVVDCSSSGTSI
LTGSAIQVQNIKKDQTLKARIEIPSCKDVAPVEKTIKLLPSSHVARLQIFSVEGQKAIQIKH
QDEVNWIAGDIMHNLIFQMYDEGEREINITSALAEKIKVNWTPEINKEHLLQGLLPDVQVPT
SVKDMRYCQVSFQDDHVSLESAFTVRPLPDEPKHLKCEMKGGKTVQMGQELQGEVVIIITDQ
YGNQIQAFSPSSLSSLSIAGVGLDSSNLKTTFQENTQSISVRGIKFIPGPPGNKDLCFTWRE
FSDFIRVQLISGPPAKLLLIDWPELKESIPVINGRDLQNPIIVQLCDQWDNPAPVQHVKISL
TKASNLKLMPSNQQHKTDEKGRANLGVFSVFAPRGEHTLQVKAIYNKSIIEGPIIKLMILPD
PEKPVRLNVKYDKDASFLAGGLFTDFMISVISEDDSIIKNINPARISMKMWKLSTSGNRPPA
NAETFSCNKIKDNDKEDGCFYFRDKVIPNKVGTYCIQFGFMMDKTNILNSEQVIVEVLPNQP
VKLVPKIKPPTPAVSNVRSVASRTLVRDLHLSITDDYDNHTGIDLVGTIIATIKGSNEEDTD
TPLFIGKVRTLEFPFVNGSAEIMSLVLAESSPGRDSTEYFIVFEPRLPLLSRTLEPYILPFM
FYNDVKKQQQMAALTKEKDQLSQSIVMYKSLFEASQQLLNEMKCQVEEARLKEAQLRNELKI
HNIDIPTTQQVPHIEALLKRKLSEQEELKKKPRRSCTLPNYTKGSGDVLGKIAHLAQIEDDR
AAMVISWHLASDMDCVVTLTTDAARRIYDETQGRQQVLPLDSIYKKTLPDWKRSLPHFRNGK
LYFKPIGDPVFARDLLTFPDNVEHCETVFGMLLGDTIILDNLDAANHYRKEVVKITHCPTLL
TRDGDRIRSNGKFGGLQNKAPPMDKLRGMVFGAPVPKQCLILGEQIDLLQQYRSAVCKLDSV
NKDLNSQLEYLRTPDMRKKKQELDEHEKNLKLIEEKLGMTPIRKCNDSLRHSPKVETTDCPV
PPKRMRREATRQNRIITKTDV*

[0067] In some example, the isolated variant of SMCHD1 comprises an amino acid having at least 75%, or at least 80%, or at least 85%, or at least 90%, or at least 91%, or at least 92%, or at least 93%, or at least 94%, or at least 95%, or at least 96%, or at least 97%, or at least 98%, or at least 99%, or identical to SEQ ID NO: 2.

[0068] In some examples, the amino acid residue corresponding to residue 135 of human SMCHD1 has been substituted with a cysteine residue is SEQ ID NO: 3 (substituted residue underlined and in italics), having the following sequence:

MAAADGGGPGGASVGTEEDGGGVGHRTVYLFDRREKESELGDRPLQVGERSDYAGFRACVCQ
TLGISPEEKFVITTTSRKEITCDNFDETVKDGVTLYLLQSVNQLLLTATKERIDFLPHYDTL
VKSGMYEYYA*C*EGQNPLPFALAELIDNSLSATSRNIGVRRIQIKLLFDETQGKPAVAVIDNG
RGMTSKQLNNWAVYRLSKFTRQGDFESDHSGYVRPVPVPRSLNSDISYFGVGGKQAVFFVGQ
SARMISKPADSQDVHELVLSKEDFEKKEKNKEAIYSGYIRNRKPSDSVHITNDDERFLHHLI
IEEKEKDSFTAVVITGVQPEHIQYLKNYFHLWTRQLAHIYHYYIHGPKGNEIRTSKEVEPFN
NIDIEISMFEKGKVPKIVNLREIQDDMQTLYVNTAADSFEFKAHVEGDGVVEGIIRYHPFLY
DRETYPDDPCFPSKLKDEDDEDDCFILEKAARGKRPIFECFWNGRLIPYTSVEDFDWCTPPK
KRGLAPIECYNRISGALFTNDKFQVSTNKLTFMDLELKLKDKNTLFTRILNGQEQRMKIDRE
FALWLKDCHEKYDKQIKFTLFKGVITRPDLPSKKQGPWATYAAIEWDGKIYKAGQLVKTIKT
LPLFYGSIVRFFLYGDHDGEVYATGGEVQIAMEPQALYDEVRTVPIAKLDRTVAEKAVKKYV
EDEMARLPDRLSVTWPEGDELLPNEVRPAGTPIGALRIEILNKKGEAMQKLPGTSHGGSKKL
LVELKVILHSSSGNKEIISHISQHGGKWPYWFKKMENIQKLGNYTLKLQVVLNESNADTYAG
RPLPSKAIKFSVKEGKPEKFSFGLLDLPFRVGVPFNIPLEFQDEFGHTSQLVTDIQPVLEAS
GLSLHYEEITKGPNCVIRGVTAKGPVNSCQGKNYNLKVTLPGLKEDSQILKIRLLPGHPRRL
KVKPDSEILVIENGTAFPFQVEVLDESDNITAQPKLIVHCKFSGAPNLPVYVVDCSSSGTSI
LTGSAIQVQNIKKDQTLKARIEIPSCKDVAPVEKTIKLLPSSHVARLQIFSVEGQKAIQIKH
QDEVNWIAGDIMHNLIFQMYDEGEREINITSALAEKIKVNWTPEINKEHLLQGLLPDVQVPT
SVKDMRYCQVSFQDDHVSLESAFTVRPLPDEPKHLKCEMKGGKTVQMGQELQGEVVIIITDQ

YGNQIQAFSPSSLSSLSIAGVGLDSSNLKTTFQENTQSISVRGIKFIPGPPGNKDLCFTWRE
FSDFIRVQLISGPPAKLLLIDWPELKESIPVINGRDLQNPIIVQLCDQWDNPAPVQHVKISL
TKASNLKLMPSNQQHKTDEKGRANLGVFSVFAPRGEHTLQVKAIYNKSIIEGPIIKLMILPD
PEKPVRLNVKYDKDASFLAGGLFTDFMISVISEDDSIIKNINPARISMKMWKLSTSGNRPPA
NAETFSCNKIKDNDKEDGCFYFRDKVIPNKVGTYCIQFGFMMDKTNILNSEQVIVEVLPNQP
VKLVPKIKPPTPAVSNVRSVASRTLVRDLHLSITDDYDNHTGIDLVGTIIATIKGSNEEDTD
TPLFIGKVRTLEFPFVNGSAEIMSLVLAESSPGRDSTEYFIVFEPRLPLLSRTLEPYILPFM
FYNDVKKQQQMAALTKEKDQLSQSIVMYKSLFEASQQLLNEMKCQVEEARLKEAQLRNELKI
HNIDIPTTQQVPHIEALLKRKLSEQEELKKKPRRSCTLPNYTKGSGDVLGKIAHLAQIEDDR
AAMVISWHLASDMDCVVTLTTDAARRIYDETQGRQQVLPLDSIYKKTLPDWKRSLPHFRNGK
LYFKPIGDPVFARDLLTFPDNVEHCETVFGMLLGDTIILDNLDAANHYRKEVVKITHCPTLL
TRDGDRIRSNGKFGGLQNKAPPMDKLRGMVFGAPVPKQCLILGEQIDLLQQYRSAVCKLDSV
NKDLNSQLEYLRTPDMRKKKQELDEHEKNLKLIEEKLGMTPIRKCNDSLRHSPKVETTDCPV
PPKRMRREATRQNRIITKTDV*

[0069] In some example, the isolated variant of SMCHD1 comprises an amino acid having at least 75%, or at least 80%, or at least 85%, or at least 90%, or at least 91%, or at least 92%, or at least 93%, or at least 94%, or at least 95%, or at least 96%, or at least 97%, or at least 98%, or at least 99%, or identical to SEQ ID NO: 3.

[0070] In some examples, the amino acid residue corresponding to residue 135 of human SMCHD1 has been substituted with a asparagine residue is SEQ ID NO: 4 (substituted residue underlined and in italics), having the following sequence:

MAAADGGGPGGASVGTEEDGGGVGHRTVYLFDRREKESELGDRPLQVGERSDYAGFRACVCQ
LGISPEEKFVITTTSRKEITCDNFDETVKDGVTLYLLQSVNQLLLTATKERIDFLPHYDTLV
KSGMYEYYA***N***EGQNPLPFALAELIDNSLSATSRNIGVRRIQIKLLFDETQGKPAVAVIDNGR
GMTSKQLNNWAVYRLSKFTRQGDFESDHSGYVRPVPVPRSLNSDISYFGVGGKQAVFFVGQS
ARMISKPADSQDVHELVLSKEDFEKKEKNKEAIYSGYIRNRKPSDSVHITNDDERFLHHLII
EEKEKDSFTAVVITGVQPEHIQYLKNYFHLWTRQLAHIYHYYIHGPKGNEIRTSKEVEPFNN
IDIEISMFEKGKVPKIVNLREIQDDMQTLYVNTAADSFEFKAHVEGDGVVEGIIRYHPFLYD
RETYPDDPCFPSKLKDEDDEDDCFILEKAARGKRPIFECFWNGRLIPYTSVEDFDWCTPPKK
RGLAPIECYNRISGALFTNDKFQVSTNKLTFMDLELKLKDKNTLFTRILNGQEQRMKIDREF
ALWLKDCHEKYDKQIKFTLFKGVITRPDLPSKKQGPWATYAAIEWDGKIYKAGQLVKTIKTL
PLFYGSIVRFFLYGDHDGEVYATGGEVQIAMEPQALYDEVRTVPIAKLDRTVAEKAVKKYVE
DEMARLPDRLSVTWPEGDELLPNEVRPAGTPIGALRIEILNKKGEAMQKLPGTSHGGSKKLL
VELKVILHSSSGNKEIISHISQHGGKWPYWFKKMENIQKLGNYTLKLQVVLNESNADTYAGR
PLPSKAIKFSVKEGKPEKFSFGLLDLPFRVGVPFNIPLEFQDEFGHTSQLVTDIQPVLEASG
LSLHYEEITKGPNCVIRGVTAKGPVNSCQGKNYNLKVTLPGLKEDSQILKIRLLPGHPRRLK
VKPDSEILVIENGTAFPFQVEVLDESDNITAQPKLIVHCKFSGAPNLPVYVVDCSSSGTSIL
TGSAIQVQNIKKDQTLKARIEIPSCKDVAPVEKTIKLLPSSHVARLQIFSVEGQKAIQIKHQ
DEVNWIAGDIMHNLIFQMYDEGEREINITSALAEKIKVNWTPEINKEHLLQGLLPDVQVPTS
VKDMRYCQVSFQDDHVSLESAFTVRPLPDEPKHLKCEMKGGKTVQMGQELQGEVVIIITDQY
GNQIQAFSPSSLSSLSIAGVGLDSSNLKTTFQENTQSISVRGIKFIPGPPGNKDLCFTWREF
SDFIRVQLISGPPAKLLLIDWPELKESIPVINGRDLQNPIIVQLCDQWDNPAPVQHVKISLT
KASNLKLMPSNQQHKTDEKGRANLGVFSVFAPRGEHTLQVKAIYNKSIIEGPIIKLMILPDP
EKPVRLNVKYDKDASFLAGGLFTDFMISVISEDDSIIKNINPARISMKMWKLSTSGNRPPAN
AETFSCNKIKDNDKEDGCFYFRDKVIPNKVGTYCIQFGFMMDKTNILNSEQVIVEVLPNQPV
KLVPKIKPPTPAVSNVRSVASRTLVRDLHLSITDDYDNHTGIDLVGTIIATIKGSNEEDTDT

PLFIGKVRTLEFPFVNGSAEIMSLVLAESSPGRDSTEYFIVFEPRLPLLSRTLEPYILPFMF
YNDVKKQQQMAALTKEKDQLSQSIVMYKSLFEASQQLLNEMKCQVEEARLKEAQLRNELKIH
NIDIPTTQQVPHIEALLKRKLSEQEELKKKPRRSCTLPNYTKGSGDVLGKIAHLAQIEDDRA
AMVISWHLASDMDCVVTLTTDAARRIYDETQGRQQVLPLDSIYKKTLPDWKRSLPHFRNGKL
YFKPIGDPVFARDLLTFPDNVEHCETVFGMLLGDTIILDNLDAANHYRKEVVKITHCPTLLT
RDGDRIRSNGKFGGLQNKAPPMDKLRGMVFGAPVPKQCLILGEQIDLLQQYRSAVCKLDSVN
KDLNSQLEYLRTPDMRKKKQELDEHEKNLKLIEEKLGMTPIRKCNDSLRHSPKVETTDCPVP
PKRMRREATRQNRIITKTDV*

[0071] In some example, the isolated variant of SMCHD1 comprises an amino acid having at least 75%, or at least 80%, or at least 85%, or at least 90%, or at least 91%, or at least 92%, or at least 93%, or at least 94%, or at least 95%, or at least 96%, or at least 97%, or at least 98%, or at least 99%, or identical to SEQ ID NO: 4.

[0072] In some examples, the amino acid residue corresponding to residue 136 of human SMCHD1 has been substituted with a glycine residue is SEQ ID NO: 5 (substituted residue underlined and in italics), having the following sequence:

MAAADGGGPGGASVGTEEDGGGVGHRTVYLFDRREKESELGDRPLQVGERSDYAGFRACVCQ
TLGISPEEKFVITTTSRKEITCDNFDETVKDGVTLYLLQSVNQLLLTATKERIDFLPHYDTL
VKSGMYEYYAS*G*GQNPLPFALAELIDNSLSATSRNIGVRRIQIKLLFDETQGKPAVAVIDNG
RGMTSKQLNNWAVYRLSKFTRQGDFESDHSGYVRPVPVPRSLNSDISYFGVGGKQAVFFVGQ
SARMISKPADSQDVHELVLSKEDFEKKEKNKEAIYSGYIRNRKPSDSVHITNDDERFLHHLI
IEEKEKDSFTAVVITGVQPEHIQYLKNYFHLWTRQLAHIYHYYIHGPKGNEIRTSKEVEPFN
NIDIEISMFEKGKVPKIVNLREIQDDMQTLYVNTAADSFEFKAHVEGDGVVEGIIRYHPFLY
DRETYPDDPCFPSKLKDEDDEDDCFILEKAARGKRPIFECFWNGRLIPYTSVEDFDWCTPPK
KRGLAPIECYNRISGALFTNDKFQVSTNKLTFMDLELKLKDKNTLFTRILNGQEQRMKIDRE
FALWLKDCHEKYDKQIKFTLFKGVITRPDLPSKKQGPWATYAAIEWDGKIYKAGQLVKTIKT
LPLFYGSIVRFFLYGDHDGEVYATGGEVQIAMEPQALYDEVRTVPIAKLDRTVAEKAVKKYV
EDEMARLPDRLSVTWPEGDELLPNEVRPAGTPIGALRIEILNKKGEAMQKLPGTSHGGSKKL
LVELKVILHSSSGNKEIISHISQHGGKWPYWFKKMENIQKLGNYTLKLQVVLNESNADTYAG
RPLPSKAIKFSVKEGKPEKFSFGLLDLPFRVGVPFNIPLEFQDEFGHTSQLVTDIQPVLEAS
GLSLHYEEITKGPNCVIRGVTAKGPVNSCQGKNYNLKVTLPGLKEDSQILKIRLLPGHPRRL
KVKPDSEILVIENGTAFPFQVEVLDESDNITAQPKLIVHCKFSGAPNLPVYVVDCSSSGTSI
LTGSAIQVQNIKKDQTLKARIEIPSCKDVAPVEKTIKLLPSSHVARLQIFSVEGQKAIQIKH
QDEVNWIAGDIMHNLIFQMYDEGEREINITSALAEKIKVNWTPEINKEHLLQGLLPDVQVPT
SVKDMRYCQVSFQDDHVSLESAFTVRPLPDEPKHLKCEMKGGKTVQMGQELQGEVVIIITDQ
YGNQIQAFSPSSLSSLSIAGVGLDSSNLKTTFQENTQSISVRGIKFIPGPPGNKDLCFTWRE
FSDFIRVQLISGPPAKLLLIDWPELKESIPVINGRDLQNPIIVQLCDQWDNPAPVQHVKISL
TKASNLKLMPSNQQHKTDEKGRANLGVFSVFAPRGEHTLQVKAIYNKSIIEGPIIKLMILPD
PEKPVRLNVKYDKDASFLAGGLFTDFMISVISEDDSIIKNINPARISMKMWKLSTSGNRPPA
NAETFSCNKIKDNDKEDGCFYFRDKVIPNKVGTYCIQFGFMMDKTNILNSEQVIVEVLPNQP
VKLVPKIKPPTPAVSNVRSVASRTLVRDLHLSITDDYDNHTGIDLVGTIIATIKGSNEEDTD
TPLFIGKVRTLEFPFVNGSAEIMSLVLAESSPGRDSTEYFIVFEPRLPLLSRTLEPYILPFM
FYNDVKKQQQMAALTKEKDQLSQSIVMYKSLFEASQQLLNEMKCQVEEARLKEAQLRNELKI
HNIDIPTTQQVPHIEALLKRKLSEQEELKKKPRRSCTLPNYTKGSGDVLGKIAHLAQIEDDR
AAMVISWHLASDMDCVVTLTTDAARRIYDETQGRQQVLPLDSIYKKTLPDWKRSLPHFRNGK
LYFKPIGDPVFARDLLTFPDNVEHCETVFGMLLGDTIILDNLDAANHYRKEVVKITHCPTLL
TRDGDRIRSNGKFGGLQNKAPPMDKLRGMVFGAPVPKQCLILGEQIDLLQQYRSAVCKLDSV

NKDLNSQLEYLRTPDMRKKKQELDEHEKNLKLIEEKLGMTPIRKCNDSLRHSPKVETTDCPV
PPKRMRREATRQNRIITKTDV*

[0073] In some example, the isolated variant of SMCHD1 comprises an amino acid having at least 75%, or at least 80%, or at least 85%, or at least 90%, or at least 91%, or at least 92%, or at least 93%, or at least 94%, or at least 95%, or at least 96%, or at least 97%, or at least 98%, or at least 99%, or identical to SEQ ID NO: 5.

[0074] In some examples, the amino acid residue corresponding to residue 342 of human SMCHD1 has been substituted with a serine residue is SEQ ID NO: 6 (substituted residue underlined and in italics), having the following sequence:

AAADGGGPGGASVGTEEDGGGVGHRTVYLFDRREKESELGDRPLQVGERSDYAGFRACVCQT
LGISPEEKFVITTTSRKEITCDNFDETVKDGVTLYLLQSVNQLLLTATKERIDFLPHYDTLV
KSGMYEYYASEGQNPLPFALAELIDNSLSATSRNIGVRRIQIKLLFDETQGKPAVAVIDNGR
GMTSKQLNNWAVYRLSKFTRQGDFESDHSGYVRPVPVPRSLNSDISYFGVGGKQAVFFVGQS
ARMISKPADSQDVHELVLSKEDFEKKEKNKEAIYSGYIRNRKPSDSVHITNDDERFLHHLII
EEKEKDSFTAVVITGVQPEHIQYLKNYFHL***S***TRQLAHIYHYYIHGPKGNEIRTSKEVEPFNN
IDIEISMFEKGKVPKIVNLREIQDDMQTLYVNTAADSFEFKAHVEGDGVVEGIIRYHPFLYD
RETYPDDPCFPSKLKDEDDEDDCFILEKAARGKRPIFECFWNGRLIPYTSVEDFDWCTPPKK
RGLAPIECYNRISGALFTNDKFQVSTNKLTFMDLELKLKDKNTLFTRILNGQEQRMKIDREF
ALWLKDCHEKYDKQIKFTLFKGVITRPDLPSKKQGPWATYAAIEWDGKIYKAGQLVKTIKTL
PLFYGSIVRFFLYGDHDGEVYATGGEVQIAMEPQALYDEVRTVPIAKLDRTVAEKAVKKYVE
DEMARLPDRLSVTWPEGDELLPNEVRPAGTPIGALRIEILNKKGEAMQKLPGTSHGGSKKLL
VELKVILHSSSGNKEIISHISQHGGKWPYWFKKMENIQKLGNYTLKLQVVLNESNADTYAGR
PLPSKAIKFSVKEGKPEKFSFGLLDLPFRVGVPFNIPLEFQDEFGHTSQLVTDIQPVLEASG
LSLHYEEITKGPNCVIRGVTAKGPVNSCQGKNYNLKVTLPGLKEDSQILKIRLLPGHPRRLK
VKPDSEILVIENGTAFPFQVEVLDESDNITAQPKLIVHCKFSGAPNLPVYVVDCSSSGTSIL
TGSAIQVQNIKKDQTLKARIEIPSCKDVAPVEKTIKLLPSSHVARLQIFSVEGQKAIQIKHQ
DEVNWIAGDIMHNLIFQMYDEGEREINITSALAEKIKVNWTPEINKEHLLQGLLPDVQVPTS
VKDMRYCQVSFQDDHVSLESAFTVRPLPDEPKHLKCEMKGGKTVQMGQELQGEVVIIITDQY
GNQIQAFSPSSLSSLSIAGVGLDSSNLKTTFQENTQSISVRGIKFIPGPPGNKDLCFTWREF
SDFIRVQLISGPPAKLLLIDWPELKESIPVINGRDLQNPIIVQLCDQWDNPAPVQHVKISLT
KASNLKLMPSNQQHKTDEKGRANLGVFSVFAPRGEHTLQVKAIYNKSIIEGPIIKLMILPDP
EKPVRLNVKYDKDASFLAGGLFTDFMISVISEDDSIIKNINPARISMKMWKLSTSGNRPPAN
AETFSCNKIKDNDKEDGCFYFRDKVIPNKVGTYCIQFGFMMDKTNILNSEQVIVEVLPNQPV
KLVPKIKPPTPAVSNVRSVASRTLVRDLHLSITDDYDNHTGIDLVGTIIATIKGSNEEDTDT
PLFIGKVRTLEFPFVNGSAEIMSLVLAESSPGRDSTEYFIVFEPRLPLLSRTLEPYILPFMF
YNDVKKQQQMAALTKEKDQLSQSIVMYKSLFEASQQLLNEMKCQVEEARLKEAQLRNELKIH
NIDIPTTQQVPHIEALLKRKLSEQEELKKKPRRSCTLPNYTKGSGDVLGKIAHLAQIEDDRA
AMVISWHLASDMDCVVTLTTDAARRIYDETQGRQQVLPLDSIYKKTLPDWKRSLPHFRNGKL
YFKPIGDPVFARDLLTFPDNVEHCETVFGMLLGDTIILDNLDAANHYRKEVVKITHCPTLLT
RDGDRIRSNGKFGGLQNKAPPMDKLRGMVFGAPVPKQCLILGEQIDLLQQYRSAVCKLDSVN
KDLNSQLEYLRTPDMRKKKQELDEHEKNLKLIEEKLGMTPIRKCNDSLRHSPKVETTDCPVP
PKRMRREATRQNRIITKTDV*

[0075] In some example, the isolated variant of SMCHD1 comprises an amino acid having at least 75%, or at least 80%, or at least 85%, or at least 90%, or at least 91%, or at

least 92%, or at least 93%, or at least 94%, or at least 95%, or at least 96%, or at least 97%, or at least 98%, or at least 99%, or identical to SEQ ID NO: 6.

[0076] In some examples, the amino acid residue corresponding to residue 348 of human SMCHD1 has been substituted with a arginine residue is SEQ ID NO: 7 (substituted residue underlined and in italics), having the following sequence:

AAADGGGPGGASVGTEEDGGGVGHRTVYLFDRREKESELGDRPLQVGERSDYAGFRACVCQT
LGISPEEKFVITTTSRKEITCDNFDETVKDGVTLYLLQSVNQLLLTATKERIDFLPHYDTLV
KSGMYEYYASEGQNPLPFALAELIDNSLSATSRNIGVRRIQIKLLFDETQGKPAVAVIDNGR
GMTSKQLNNWAVYRLSKFTRQGDFESDHSGYVRPVPVPRSLNSDISYFGVGGKQAVFFVGQS
ARMISKPADSQDVHELVLSKEDFEKKEKNKEAIYSGYIRNRKPSDSVHITNDDERFLHHLII
EEKEKDSFTAVVITGVQPEHIQYLKNYFHLWTRQLA*R*IYHYYIHGPKGNEIRTSKEVEPFNN
IDIEISMFEKGKVPKIVNLREIQDDMQTLYVNTAADSFEFKAHVEGDGVVEGIIRYHPFLYD
RETYPDDPCFPSKLKDEDDEDDCFILEKAARGKRPIFECFWNGRLIPYTSVEDFDWCTPPKK
RGLAPIECYNRISGALFTNDKFQVSTNKLTFMDLELKLKDKNTLFTRILNGQEQRMKIDREF
ALWLKDCHEKYDKQIKFTLFKGVITRPDLPSKKQGPWATYAAIEWDGKIYKAGQLVKTIKTL
PLFYGSIVRFFLYGDHDGEVYATGGEVQIAMEPQALYDEVRTVPIAKLDRTVAEKAVKKYVE
DEMARLPDRLSVTWPEGDELLPNEVRPAGTPIGALRIEILNKKGEAMQKLPGTSHGGSKKLL
VELKVILHSSSGNKEIISHISQHGGKWPYWFKKMENIQKLGNYTLKLQVVLNESNADTYAGR
PLPSKAIKFSVKEGKPEKFSFGLLDLPFRVGVPFNIPLEFQDEFGHTSQLVTDIQPVLEASG
LSLHYEEITKGPNCVIRGVTAKGPVNSCQGKNYNLKVTLPGLKEDSQILKIRLLPGHPRRLK
VKPDSEILVIENGTAFPFQVEVLDESDNITAQPKLIVHCKFSGAPNLPVYVVDCSSSGTSIL
TGSAIQVQNIKKDQTLKARIEIPSCKDVAPVEKTIKLLPSSHVARLQIFSVEGQKAIQIKHQ
DEVNWIAGDIMHNLIFQMYDEGEREINITSALAEKIKVNWTPEINKEHLLQGLLPDVQVPTS
VKDMRYCQVSFQDDHVSLESAFTVRPLPDEPKHLKCEMKGGKTVQMGQELQGEVVIIITDQY
GNQIQAFSPSSLSSLSIAGVGLDSSNLKTTFQENTQSISVRGIKFIPGPPGNKDLCFTWREF
SDFIRVQLISGPPAKLLLIDWPELKESIPVINGRDLQNPIIVQLCDQWDNPAPVQHVKISLT
KASNLKLMPSNQQHKTDEKGRANLGVFSVFAPRGEHTLQVKAIYNKSIIEGPIIKLMILPDP
EKPVRLNVKYDKDASFLAGGLFTDFMISVISEDDSIIKNINPARISMKMWKLSTSGNRPPAN
AETFSCNKIKDNDKEDGCFYFRDKVIPNKVGTYCIQFGFMMDKTNILNSEQVIVEVLPNQPV
KLVPKIKPPTPAVSNVRSVASRTLVRDLHLSITDDYDNHTGIDLVGTIIATIKGSNEEDTDT
PLFIGKVRTLEFPFVNGSAEIMSLVLAESSPGRDSTEYFIVFEPRLPLLSRTLEPYILPFMF
YNDVKKQQQMAALTKEKDQLSQSIVMYKSLFEASQQLLNEMKCQVEEARLKEAQLRNELKIH
NIDIPTTQQVPHIEALLKRKLSEQEELKKKPRRSCTLPNYTKGSGDVLGKIAHLAQIEDDRA
AMVISWHLASDMDCVVTLTTDAARRIYDETQGRQQVLPLDSIYKKTLPDWKRSLPHFRNGKL
YFKPIGDPVFARDLLTFPDNVEHCETVFGMLLGDTIILDNLDAANHYRKEVVKITHCPTLLT
RDGDRIRSNGKFGGLQNKAPPMDKLRGMVFGAPVPKQCLILGEQIDLLQQYRSAVCKLDSVN
KDLNSQLEYLRTPDMRKKKQELDEHEKNLKLIEEKLGMTPIRKCNDSLRHSPKVETTDCPVP
PKRMRREATRQNRIITKTDV*

[0077] In some example, the isolated variant of SMCHD1 comprises an amino acid having at least 75%, or at least 80%, or at least 85%, or at least 90%, or at least 91%, or at least 92%, or at least 93%, or at least 94%, or at least 95%, or at least 96%, or at least 97%, or at least 98%, or at least 99%, or identical to SEQ ID NO: 7.

[0078] In some examples, the amino acid residue corresponding to residue 420 of human SMCHD1 has been substituted with a valine residue is SEQ ID NO: 8 (substituted residue underlined and in italics), having the following sequence:

MAAADGGGPGGASVGTEEDGGGVGHRTVYLFDRREKESELGDRPLQVGERSDYAGFRACVCQ
TLGISPEEKFVITTTSRKEITCDNFDETVKDGVTLYLLQSVNQLLLTATKERIDFLPHYDTL
VKSGMYEYYASEGQNPLPFALAELIDNSLSATSRNIGVRRIQIKLLFDETQGKPAVAVIDNG
RGMTSKQLNNWAVYRLSKFTRQGDFESDHSGYVRPVPVPRSLNSDISYFGVGGKQAVFFVGQ
SARMISKPADSQDVHELVLSKEDFEKKEKNKEAIYSGYIRNRKPSDSVHITNDDERFLHHLI
IEEKEKDSFTAVVITGVQPEHIQYLKNYFHLWTRQLAHIYHYYIHGPKGNEIRTSKEVEPFN
NIDIEISMFEKGKVPKIVNLREIQDDMQTLYVNTAADSFEFKAHVEG*V*GVVEGIIRYHPFLY
DRETYPDDPCFPSKLKDEDDEDDCFILEKAARGKRPIFECFWNGRLIPYTSVEDFDWCTPPK
KRGLAPIECYNRISGALFTNDKFQVSTNKLTFMDLELKLKDKNTLFTRILNGQEQRMKIDRE
FALWLKDCHEKYDKQIKFTLFKGVITRPDLPSKKQGPWATYAAIEWDGKIYKAGQLVKTIKT
LPLFYGSIVRFFLYGDHDGEVYATGGEVQIAMEPQALYDEVRTVPIAKLDRTVAEKAVKKYV
EDEMARLPDRLSVTWPEGDELLPNEVRPAGTPIGALRIEILNKKGEAMQKLPGTSHGGSKKL
LVELKVILHSSSGNKEIISHISQHGGKWPYWFKKMENIQKLGNYTLKLQVVLNESNADTYAG
RPLPSKAIKFSVKEGKPEKFSFGLLDLPFRVGVPFNIPLEFQDEFGHTSQLVTDIQPVLEAS
GLSLHYEEITKGPNCVIRGVTAKGPVNSCQGKNYNLKVTLPGLKEDSQILKIRLLPGHPRRL
KVKPDSEILVIENGTAFPFQVEVLDESDNITAQPKLIVHCKFSGAPNLPVYVVDCSSSGTSI
LTGSAIQVQNIKKDQTLKARIEIPSCKDVAPVEKTIKLLPSSHVARLQIFSVEGQKAIQIKH
QDEVNWIAGDIMHNLIFQMYDEGEREINITSALAEKIKVNWTPEINKEHLLQGLLPDVQVPT
SVKDMRYCQVSFQDDHVSLESAFTVRPLPDEPKHLKCEMKGGKTVQMGQELQGEVVIIITDQ
YGNQIQAFSPSSLSSLSIAGVGLDSSNLKTTFQENTQSISVRGIKFIPGPPGNKDLCFTWRE
FSDFIRVQLISGPPAKLLLIDWPELKESIPVINGRDLQNPIIVQLCDQWDNPAPVQHVKISL
TKASNLKLMPSNQQHKTDEKGRANLGVFSVFAPRGEHTLQVKAIYNKSIIEGPIIKLMILPD
PEKPVRLNVKYDKDASFLAGGLFTDFMISVISEDDSIIKNINPARISMKMWKLSTSGNRPPA
NAETFSCNKIKDNDKEDGCFYFRDKVIPNKVGTYCIQFGFMMDKTNILNSEQVIVEVLPNQP
VKLVPKIKPPTPAVSNVRSVASRTLVRDLHLSITDDYDNHTGIDLVGTIIATIKGSNEEDTD
TPLFIGKVRTLEFPFVNGSAEIMSLVLAESSPGRDSTEYFIVFEPRLPLLSRTLEPYILPFM
FYNDVKKQQQMAALTKEKDQLSQSIVMYKSLFEASQQLLNEMKCQVEEARLKEAQLRNELKI
HNIDIPTTQQVPHIEALLKRKLSEQEELKKKPRRSCTLPNYTKGSGDVLGKIAHLAQIEDDR
AAMVISWHLASDMDCVVTLTTDAARRIYDETQGRQQVLPLDSIYKKTLPDWKRSLPHFRNGK
LYFKPIGDPVFARDLLTFPDNVEHCETVFGMLLGDTIILDNLDAANHYRKEVVKITHCPTLL
TRDGDRIRSNGKFGGLQNKAPPMDKLRGMVFGAPVPKQCLILGEQIDLLQQYRSAVCKLDSV
NKDLNSQLEYLRTPDMRKKKQELDEHEKNLKLIEEKLGMTPIRKCNDSLRHSPKVETTDCPV
PPKRMRREATRQNRIITKTDV*

[0079] In some example, the isolated variant of SMCHD1 comprises an amino acid having at least 75%, or at least 80%, or at least 85%, or at least 90%, or at least 91%, or at least 92%, or at least 93%, or at least 94%, or at least 95%, or at least 96%, or at least 97%, or at least 98%, or at least 99%, or identical to SEQ ID NO: 8.

[0080] In some examples, the amino acid residue corresponding to residue 518 of human SMCHD1 has been substituted with a glutamic acid residue is SEQ ID NO: 9 (substituted residue underlined and in italics), having the following sequence:

AAADGGGPGGASVGTEEDGGGVGHRTVYLFDRREKESELGDRPLQVGERSDYAGFRACVCQT
LGISPEEKFVITTTSRKEITCDNFDETVKDGVTLYLLQSVNQLLLTATKERIDFLPHYDTLV

KSGMYEYYASEGQNPLPFALAELIDNSLSATSRNIGVRRIQIKLLFDETQGKPAVAVIDNGR
GMTSKQLNNWAVYRLSKFTRQGDFESDHSGYVRPVPVPRSLNSDISYFGVGGKQAVFFVGQS
ARMISKPADSQDVHELVLSKEDFEKKEKNKEAIYSGYIRNRKPSDSVHITNDDERFLHHLII
EEKEKDSFTAVVITGVQPEHIQYLKNYFHLWTRQLAHIYHYYIHGPKGNEIRTSKEVEPFNN
IDIEISMFEKGKVPKIVNLREIQDDMQTLYVNTAADSFEFKAHVEGDGVVEGIIRYHPFLYD
RETYPDDPCFPSKLKDEDDEDDCFILEKAARGKRPIFECFWNGRLIPYTSVEDFDWCTPPKK
RGLAPIECYNRISGALFTND***E***FQVSTNKLTFMDLELKLKDKNTLFTRILNGQEQRMKIDREF
ALWLKDCHEKYDKQIKFTLFKGVITRPDLPSKKQGPWATYAAIEWDGKIYKAGQLVKTIKTL
PLFYGSIVRFFLYGDHDGEVYATGGEVQIAMEPQALYDEVRTVPIAKLDRTVAEKAVKKYVE
DEMARLPDRLSVTWPEGDELLPNEVRPAGTPIGALRIEILNKKGEAMQKLPGTSHGGSKKLL
VELKVILHSSSGNKEIISHISQHGGKWPYWFKKMENIQKLGNYTLKLQVVLNESNADTYAGR
PLPSKAIKFSVKEGKPEKFSFGLLDLPFRVGVPFNIPLEFQDEFGHTSQLVTDIQPVLEASG
LSLHYEEITKGPNCVIRGVTAKGPVNSCQGKNYNLKVTLPGLKEDSQILKIRLLPGHPRRLK
VKPDSEILVIENGTAFPFQVEVLDESDNITAQPKLIVHCKFSGAPNLPVYVVDCSSSGTSIL
TGSAIQVQNIKKDQTLKARIEIPSCKDVAPVEKTIKLLPSSHVARLQIFSVEGQKAIQIKHQ
DEVNWIAGDIMHNLIFQMYDEGEREINITSALAEKIKVNWTPEINKEHLLQGLLPDVQVPTS
VKDMRYCQVSFQDDHVSLESAFTVRPLPDEPKHLKCEMKGGKTVQMGQELQGEVVIIITDQY
GNQIQAFSPSSLSSLSIAGVGLDSSNLKTTFQENTQSISVRGIKFIPGPPGNKDLCFTWREF
SDFIRVQLISGPPAKLLLIDWPELKESIPVINGRDLQNPIIVQLCDQWDNPAPVQHVKISLT
KASNLKLMPSNQQHKTDEKGRANLGVFSVFAPRGEHTLQVKAIYNKSIIEGPIIKLMILPDP
EKPVRLNVKYDKDASFLAGGLFTDFMISVISEDDSIIKNINPARISMKMWKLSTSGNRPPAN
AETFSCNKIKDNDKEDGCFYFRDKVIPNKVGTYCIQFGFMMDKTNILNSEQVIVEVLPNQPV
KLVPKIKPPTPAVSNVRSVASRTLVRDLHLSITDDYDNHTGIDLVGTIIATIKGSNEEDTDT
PLFIGKVRTLEFPFVNGSAEIMSLVLAESSPGRDSTEYFIVFEPRLPLLSRTLEPYILPFMF
YNDVKKQQQMAALTKEKDQLSQSIVMYKSLFEASQQLLNEMKCQVEEARLKEAQLRNELKIH
NIDIPTTQQVPHIEALLKRKLSEQEELKKKPRRSCTLPNYTKGSGDVLGKIAHLAQIEDDRA
AMVISWHLASDMDCVVTLTTDAARRIYDETQGRQQVLPLDSIYKKTLPDWKRSLPHFRNGKL
YFKPIGDPVFARDLLTFPDNVEHCETVFGMLLGDTIILDNLDAANHYRKEVVKITHCPTLLT
RDGDRIRSNGKFGGLQNKAPPMDKLRGMVFGAPVPKQCLILGEQIDLLQQYRSAVCKLDSVN
KDLNSQLEYLRTPDMRKKKQELDEHEKNLKLIEEKLGMTPIRKCNDSLRHSPKVETTDCPVP
PKRMRREATRQNRIITKTDV*

[0081] In some example, the isolated variant of SMCHD1 comprises an amino acid having at least 75%, or at least 80%, or at least 85%, or at least 90%, or at least 91%, or at least 92%, or at least 93%, or at least 94%, or at least 95%, or at least 96%, or at least 97%, or at least 98%, or at least 99%, or identical to SEQ ID NO: 9.

[0082] In some examples, the amino acid residue corresponding to residue 552 of human SMCHD1 has been substituted with a glutamine residue is SEQ ID NO: 10 (substituted residue underlined and in italics), having the following sequence:

AAADGGGPGGASVGTEEDGGGVGHRTVYLFDRREKESELGDRPLQVGERSDYAGFRACVCQT
LGISPEEKFVITTTSRKEITCDNFDETVKDGVTLYLLQSVNQLLLTATKERIDFLPHYDTLV
KSGMYEYYASEGQNPLPFALAELIDNSLSATSRNIGVRRIQIKLLFDETQGKPAVAVIDNGR
GMTSKQLNNWAVYRLSKFTRQGDFESDHSGYVRPVPVPRSLNSDISYFGVGGKQAVFFVGQS
ARMISKPADSQDVHELVLSKEDFEKKEKNKEAIYSGYIRNRKPSDSVHITNDDERFLHHLII
EEKEKDSFTAVVITGVQPEHIQYLKNYFHLWTRQLAHIYHYYIHGPKGNEIRTSKEVEPFNN
IDIEISMFEKGKVPKIVNLREIQDDMQTLYVNTAADSFEFKAHVEGDGVVEGIIRYHPFLYD
RETYPDDPCFPSKLKDEDDEDDCFILEKAARGKRPIFECFWNGRLIPYTSVEDFDWCTPPKK

RGLAPIECYNRISGALFTNDKFQVSTNKLTFMDLELKLKDKNTLFTRILNGQEQQMKIDREF
ALWLKDCHEKYDKQIKFTLFKGVITRPDLPSKKQGPWATYAAIEWDGKIYKAGQLVKTIKTL
PLFYGSIVRFFLYGDHDGEVYATGGEVQIAMEPQALYDEVRTVPIAKLDRTVAEKAVKKYVE
DEMARLPDRLSVTWPEGDELLPNEVRPAGTPIGALRIEILNKKGEAMQKLPGTSHGGSKKLL
VELKVILHSSSGNKEIISHISQHGGKWPYWFKKMENIQKLGNYTLKLQVVLNESNADTYAGR
PLPSKAIKFSVKEGKPEKFSFGLLDLPFRVGVPFNIPLEFQDEFGHTSQLVTDIQPVLEASG
LSLHYEEITKGPNCVIRGVTAKGPVNSCQGKNYNLKVTLPGLKEDSQILKIRLLPGHPRRLK
VKPDSEILVIENGTAFPFQVEVLDESDNITAQPKLIVHCKFSGAPNLPVYVVDCSSSGTSIL
TGSAIQVQNIKKDQTLKARIEIPSCKDVAPVEKTIKLLPSSHVARLQIFSVEGQKAIQIKHQ
DEVNWIAGDIMHNLIFQMYDEGEREINITSALAEKIKVNWTPEINKEHLLQGLLPDVQVPTS
VKDMRYCQVSFQDDHVSLESAFTVRPLPDEPKHLKCEMKGGKTVQMGQELQGEVVIIITDQY
GNQIQAFSPSSLSSLSIAGVGLDSSNLKTTFQENTQSISVRGIKFIPGPPGNKDLCFTWREF
SDFIRVQLISGPPAKLLLIDWPELKESIPVINGRDLQNPIIVQLCDQWDNPAPVQHVKISLT
KASNLKLMPSNQQHKTDEKGRANLGVFSVFAPRGEHTLQVKAIYNKSIIEGPIIKLMILPDP
EKPVRLNVKYDKDASFLAGGLFTDFMISVISEDDSIIKNINPARISMKMWKLSTSGNRPPAN
AETFSCNKIKDNDKEDGCFYFRDKVIPNKVGTYCIQFGFMMDKTNILNSEQVIVEVLPNQPV
KLVPKIKPPTPAVSNVRSVASRTLVRDLHLSITDDYDNHTGIDLVGTIIATIKGSNEEDTDT
PLFIGKVRTLEFPFVNGSAEIMSLVLAESSPGRDSTEYFIVFEPRLPLLSRTLEPYILPFMF
YNDVKKQQQMAALTKEKDQLSQSIVMYKSLFEASQQLLNEMKCQVEEARLKEAQLRNELKIH
NIDIPTTQQVPHIEALLKRKLSEQEELKKKPRRSCTLPNYTKGSGDVLGKIAHLAQIEDDRA
AMVISWHLASDMDCVVTLTTDAARRIYDETQGRQQVLPLDSIYKKTLPDWKRSLPHFRNGKL
YFKPIGDPVFARDLLTFPDNVEHCETVFGMLLGDTIILDNLDAANHYRKEVVKITHCPTLLT
RDGDRIRSNGKFGGLQNKAPPMDKLRGMVFGAPVPKQCLILGEQIDLLQQYRSAVCKLDSVN
KDLNSQLEYLRTPDMRKKKQELDEHEKNLKLIEEKLGMTPIRKCNDSLRHSPKVETTDCPVP
PKRMRREATRQNRIITKTDV*

[0083] In some example, the isolated variant of SMCHD1 comprises an amino acid having at least 75%, or at least 80%, or at least 85%, or at least 90%, or at least 91%, or at least 92%, or at least 93%, or at least 94%, or at least 95%, or at least 96%, or at least 97%, or at least 98%, or at least 99%, or identical to SEQ ID NO: 10.

[0084] As discussed in detail below, the inventors of the present disclosure found gain-of-function mutation in SMCHD1 as disclosed herein can be useful in treating subjects having FSHD. Accordingly, in another aspect, the present invention provides an isolated variant of SMCHD1 as described herein for use in therapy or medicine.

[0085] It is also envisaged that the present invention also provides an isolated nucleic acid or fragment thereof encoding the isolated variant SMCHD1 peptide as described herein, wherein the nucleic acid comprises one or more mutation(s) that increases SMCHD1 protein activity in a cell compared to a cell with a wild-type SMCHD1 peptide.

[0086] As discussed above, the SMCHD1 mutation that caused the substitution mutation is a missense mutation. Thus, in some examples, the SMCHD1 mutation is a missense mutation. In some examples, the SMCHD1 mutation is at one or more position(s), including, but is not limited to c.407, c.403, c.404, c.1043, c.1259, c.1655, c.1552, c.1025, and c.400 (nucleic acid residue referenced to NCBI Reference Sequence: NM_015295.2). IN some

example, the SMCHD1 mutation may be at least two, at least three, at least four, at least five, at least six, at least seven, at least eight, or all of the mutations such as c.407, c.403, c.404, c.1043, c.1259, c.1655, c.1552, c.1025, and c.400 (nucleic acid residue referenced to NCBI Reference Sequence: NM_015295.2).

[0087] In some examples, SMCHD1 mutation may be one or more including, but not limited to, c.407A>G, c.403A>T, c.404G>A, c.1043A>G, c.1259A>T, c.1655G>A, c.1552A>G, c.1025G>C, c.400G>T, and the like (nucleic acid residue referenced to referenced to NCBI Reference Sequence: NM_015295.2). As would be understood by the person skilled in the art, the mutation is denoted by the alphabet recitation. That is, c.406A>G refers to a missense mutation at position 406 of the nucleic acid residue referenced at NCBI reference sequence NM_015295.2 and the nucleic acid adenine (A) has been replaced by guanine (G). In some examples, as used herein, the term "nucleic acid" is used in accordance to the common use in the art and is referred to include the various nucleic acids known in the art. For example, the nucleic acid includes, but is not limited to, a DNA, an mRNA, and an RNA.

[0088] In some examples, the present disclosure also provides for an isolated nucleic acid as described herein for use in therapy (such as gene-therapy).

[0089] In some examples, the present disclosure provides for one or more vector(s) comprising the nucleic acid as described herein. The vectors as used herein refer to any vectors that would be recognized by the person skilled in the art to be capable of being used in gene therapy, nucleic acid, or peptide expression. Accordingly, in some examples, the vector may include, but is not limited to, an adeno-associated virus based vector, a retro-associated virus based vector, and the like.

[0090] Since the peptides, nucleic acids, or vectors as described herein is envisaged to be used in therapy or medicine, the present disclosure also provides a pharmaceutical composition. Thus, in another aspect, there is provided a pharmaceutical composition comprising the isolated variant SMCHD1 peptide or fragment thereof as described herein, and/or an isolated nucleic acid or fragment thereof as described herein, or a vector as described herein, and a pharmaceutically acceptable excipient thereof.

[0091] In another aspect, the present invention provides a gene-therapy composition comprising an isolated nucleic acid as described herein or a vector as described herein.

[0092] In another aspect, the present invention also provides one or more host cell(s) comprising the vector as described herein.

[0093] In some examples, there is provided an SMCHD1 mutein comprising an amino acid change of at least one of the following: 134Ser, 135Asn, 135Cys, 136Gly, 342Ser, 348Arg, 420Val, 518Glu, and 552Gln.

[0094] As illustrated in Figure 3, the variant SMCHD1 peptide or fragment as described herein has been illustrated to be useful increasing the expression of SMCHD1 *in vivo*. Thus, the experimental data provides proof that the variant SMCHD1 peptide could be used to treat FSHD.

[0095] Thus, in one aspect, there is provided a method of treating facioscapulohumeral muscular dystrophy (FSHD) in a subject in need thereof, comprising administering a therapeutically effective amount of an isolated variant SMCHD1 peptide or fragment thereof as described herein, and/or a transgene encoding the isolated variant SMCHD1 peptide as described herein, and/or an isolated nucleic acid or fragment thereof as described herein, and/or a vector as described herein, and/or a pharmaceutical composition as described herein, or a gene-therapy composition as described herein, to the subject in need thereof.

[0096] In another aspect, there is provided use of an isolated variant SMCHD1 peptide as described herein, and/or a transgene encoding the isolated variant SMCHD1 peptide as described herein, and/or an isolated nucleic acid as described herein, and/or a vector as described herein, and/or a pharmaceutical composition as described herein, or a gene-therapy composition as described herein in the manufacture of a medicament for treating facioscapulohumeral muscular dystrophy (FSHD) in a subject in need thereof

[0097] In some examples, the isolated peptide, and/or transgene, and/or nucleic acid, and/or vector, and/or pharmaceutical composition may be delivered via a route known in the art. For example, the route of delivery or administration includes, but is not limited to, intravascularly, intramuscularly, intradermally, intraperitoneally, intraarterially, and the like.

[0098] In another aspect, the present invention also provides a method of treating facioscapulohumeral muscular dystrophy (FSHD) in a subject in need thereof, comprising administering a therapeutically effective amount of an agent capable of increasing the expression level or activity of SMCHD1 peptide and/or nucleic acid to the subject in need thereof.

[0099] In some examples, the agent capable of increasing the expression level or activity of SMCHD1 peptide and/or nucleic acid to the subject may be the variant SMCHD1 peptide and/or nucleic acid as described herein.

[00100] In some examples, the isolated peptide, and/or transgene, and/or nucleic acid, and/or vector, and/or pharmaceutical composition, and/or agent capable of increasing the expression level or activity of SMCHD1 peptide and/or nucleic acid to the subject may be delivered via a route known in the art. For example, the route of delivery or administration includes, but is not limited to, intravascularly, intramuscularly, intradermally, intraperitoneally, intraarterially, and the like.

[00101] In some examples, the FSHD as discussed in the present disclosure may be a type 1 and/or type 2 FSHD. In some examples, the FSHD as discussed in the present disclosure may be a type 2 FSHD.

[00102] In some examples, the subject is an animal such as, but is not limited to, a human and a non-human mammal. In some examples, the subject is a human.

[00103] The present inventors also surprisingly found that the nucleotide mutations or muteins as disclosed herein can be used to screen compounds that can modulate SMCHD1. Thus, in some examples, there is provided a method of screening using the nucleotide mutations or muteins as described herein to identify compounds that can modulate SMCHD1 activity. For example, the gain of function (GOF) mutations or muteins as described herein may be used as a) positive controls in drug screens for increasing SMCHD1 activity; b) possible therapeutic in the form of introducing these gain of function (GOF) SMCHD1 protein or gene in FSHD patients; c) tools in developing molecules that mimic the gain of function (GOF) activity of these mutations or increase the activity of endogenous SMCHD1.

[00104] It would also be appreciated by the person skilled in the art that tools for developing gain of function activity may be developed by understanding the structural changes that occur with the mutations, which would then lead to the design molecules that mimic these change vis a vis chromatin architecture and downstream targets, either by directly acting on SMCHD1 or as a functional mimic of SMCHD1.

[00105] Thus, in one aspect, there is provided a method of screening an agent (or a nucleotide mutation) capable of increasing SMCHD1 activity, comprising: a. introducing the agent (or nucleotide mutation) to an assay for determining SMCHD1 activity, b. determining the SMCHD1 activity the agent (or nucleotide mutation) elicits, and c. comparing the SMCHD1 activity the agent (or nucleotide mutation) elicits with the SMCHD1 activity elicited by an isolated variant SMCHD1 peptide or fragment thereof as described herein, and/or a transgene encoding the isolated variant SMCHD1 peptide as described herein, and/or an isolated nucleic acid or fragment thereof as described herein, and/or a vector as described

herein, wherein when the SMCHD1 activity of the agent (or nucleotide mutation) is the same or more than the SMCHD1 activity of the isolated variant SMCHD1 peptide or fragment thereof as described herein, and/or the transgene encoding the isolated variant SMCHD1 peptide as described herein, and/or the isolated nucleic acid or fragment thereof as described herein, and/or the vector as described herein, the agent (or nucleotide mutation) is considered a suitable agent (or nucleotide mutation) for increasing SMCHD1 activity.

[00106] In some examples, the assay for determining SMCHD1 activity may include, but is not limited to, an ATPase activity assay, assays capable of determining and/or measuring the expression and activity of a downstream target of SMCHD1 (such as DUX4; and such assays may be a cell-based assay and/or a reporter-based assay), assays capable of determining the in vivo activity of SMCHD1 (for example microinjection of the agent into Xenopus embryos, wherein mutant agent that causes increase in SMCHD1 activity may cause craniofacial abnormalities in Xenopus embryos during development), and the like.

[00107] In some examples, the present disclosure also provides a kit for treating facioscapulohumeral muscular dystrophy (FSHD) in a subject in need thereof, comprising a therapeutically effective amount of an isolated variant SMCHD1 peptide as described herein, and/or a transgene encoding the isolated variant SMCHD1 peptide as described herein, and/or an isolated nucleic acid as described herein, and/or a vector as described herein, and/or a pharmaceutical composition as described herein, or a gene-therapy composition as described herein.

[00108] In some examples, the present disclosure also provides a kit for screening an agent capable of increasing SMCHD1 activity, comprising a reference control comprising an isolated variant SMCHD1 peptide as described herein, and/or a transgene encoding the isolated variant SMCHD1 peptide as described herein, and/or an isolated nucleic acid as described herein, and/or a vector as described herein, and a reagent for determining SMCHD1 activity.

[00109] The reagent for determining SMCHD1 activity as used herein are reagents that are commonly used in the art of determining the activity of a gene in a human or animal body. For example, the reagents may include primers as used in the present disclosure, for example primers described in the Experimental section. For example, the reagents for determining SMCHD1 activity may include at least one or more of primers selected from the group consisting of SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO:

20, SEQ ID NO: 21, SEQ ID NO: 22, SEQ ID NO: 23, SEQ ID NO: 24, SEQ ID NO: 25, SEQ ID NO: 26, SEQ ID NO: 27, SEQ ID NO: 28, SEQ ID NO: 29, SEQ ID NO: 30, SEQ ID NO: 31, SEQ ID NO: 32, SEQ ID NO: 33, SEQ ID NO: 34, SEQ ID NO: 35, and SEQ ID NO: 36. The methods of using the primer/primer pairs as disclosed herein are commonly known in the art and one exemplary method is discussed in the Experimental section of the present disclosure.

[00110] As used in this application, the singular form “a,” “an,” and “the” include plural references unless the context clearly dictates otherwise. For example, the term “a peptide” includes a plurality of peptides (or polypeptides), including mixtures and combinations thereof.

[00111] As used herein, the terms “increase” and “decrease” refer to the relative alteration of a chosen trait or characteristic in a subset of a population in comparison to the same trait or characteristic as present in the whole population. An increase thus indicates a change on a positive scale, whereas a decrease indicates a change on a negative scale. The term “change”, as used herein, also refers to the difference between a chosen trait or characteristic of an isolated population subset in comparison to the same trait or characteristic in the population as a whole. However, this term is without valuation of the difference seen.

[00112] As used herein, the term “about” in the context of concentration of a substance, size of a substance, length of time, or other stated values means +/- 5% of the stated value, or +/- 4% of the stated value, or +/- 3% of the stated value, or +/- 2% of the stated value, or +/- 1% of the stated value, or +/- 0.5% of the stated value.

[00113] Throughout this disclosure, certain embodiments may be disclosed in a range format. It should be understood that the description in range format is merely for convenience and brevity and should not be construed as an inflexible limitation on the scope of the disclosed ranges. Accordingly, the description of a range should be considered to have specifically disclosed all the possible sub-ranges as well as individual numerical values within that range. For example, description of a range such as from 1 to 6 should be considered to have specifically disclosed sub-ranges such as from 1 to 3, from 1 to 4, from 1 to 5, from 2 to 4, from 2 to 6, from 3 to 6 etc., as well as individual numbers within that range, for example, 1, 2, 3, 4, 5, and 6. This applies regardless of the breadth of the range.

[00114] The invention illustratively described herein may suitably be practiced in the absence of any element or elements, limitation or limitations, not specifically disclosed herein. Thus, for example, the terms "comprising", "including", "containing", etc. shall be

read expansively and without limitation. Additionally, the terms and expressions employed herein have been used as terms of description and not of limitation, and there is no intention in the use of such terms and expressions of excluding any equivalents of the features shown and described or portions thereof, but it is recognized that various modifications are possible within the scope of the invention claimed. Thus, it should be understood that although the present invention has been specifically disclosed by preferred embodiments and optional features, modification and variation of the inventions embodied therein herein disclosed may be resorted to by those skilled in the art, and that such modifications and variations are considered to be within the scope of this invention.

[00115] The invention has been described broadly and generically herein. Each of the narrower species and subgeneric groupings falling within the generic disclosure also form part of the invention. This includes the generic description of the invention with a proviso or negative limitation removing any subject matter from the genus, regardless of whether or not the excised material is specifically recited herein.

[00116] Other embodiments are within the following claims and non- limiting examples. In addition, where features or aspects of the invention are described in terms of Markush groups, those skilled in the art will recognize that the invention is also thereby described in terms of any individual member or subgroup of members of the Markush group.

EXPERIMENTAL SECTION

[00117] **Material and Methods**

[00118] **Subjects**

[00119] In all cases informed consent was obtained from the families for genetic testing. For patients in Figure 1, consent to publish photos was obtained.

[00120] **Whole-exome sequencing (WES)**

[00121] WES was conducted in accordance with approved institutional ethical guidelines (Comité de Protection des Personnes Ile-de-France II; Ethics Committee of the University Hospital Cologne, Germany; National University of Singapore Institutional Review Board).

[00122] For trio whole-exome sequencing (WES) of case 1, Agilent SureSelect libraries were prepared from 3 µg of genomic DNA from each individual and sheared with a Covaris S2 Ultrasonicator. Exome capture was performed with the 51 Mb SureSelect Human All Exon kit V5 (Agilent technologies). Sequencing was carried out on a pool of barcoded exome libraries using a HiSeq 2500 instrument (Illumina), generating 100 + 100 bp paired-end

reads. After demultiplexing, paired-end sequences were mapped to the reference human genome (GRCh37/hg19 assembly, NCBI) using Burrows- Wheeler Aligner (BWA). The mean depth of coverage obtained for the three samples from case 1 was 123-, 149- and 150-fold, and 98% of the exome was covered by at least 15-fold. Downstream processing was performed using the Genome Analysis Toolkit (GATK), SAMtools and Picard. Variant calls were made with the GATK Unified Genotyper. All calls with read coverage ≤2-fold or a Phred-scaled SNP quality score of ≤20-fold were removed from consideration. Variant annotation was based on Ensembl release 71. Variants were filtered against publicly available SNPs plus variant data from more than 7,000 in-house exomes (Institut Imagine).

[00123] For trio WES of case 2, exonic and adjacent intronic sequences were enriched from genomic DNA using the NimbleGen SeqCap EZ Human Exome Library v2.0 enrichment kit and probes were run on an Illumina HiSeq2000 sequencer at the Cologne Center for Genomics (CCG). Data analysis and filtering of mapped target sequences was performed with the "Varbank" exome and genome analysis pipeline v.2.1 (CCG) and data were filtered for high-quality (coverage of more than 6 reads, a minimum quality score of 10), rare (MAF < 0.5%) autosomal recessive and de novo variants.

[00124] For cases 9 and 11 trios and 10 and 12 quartets, WES was performed at the Genome Institute of Singapore. Barcoded libraries were prepared for each individual by shearing 1ug of genomic DNA, followed by end-repair, A-tailing, adaptor ligation and PCR enrichment, then pooled and hybridized with NimbleGen SeqCap EZ Human Exome Library v3.0 probes. Captured DNA targets were purified and PCR amplified, then sequenced on Illumina HiSeq 2500 (cases 9 and 11) or HiSeq 4000 (cases 10 and 12) sequencers. Variant calling was performed following GATK (v3.4.46) recommended best practices. Reads were mapped to GRCh37/hg19 using BWA and the aligned files pre-processed by Picard and GATK. All samples were sequenced at mean coverage of 75X or higher. The variants were called using GATK HaplotypeCaller along with in-house exomes sequenced with the same chemistry. The variants were recalibrated, annotated and filtered against in-house data plus common publicly available databases. Each family was independently analyzed using Phen-Gen for *de novo* and recessive disease inheritance patterns. Variants with alternate allele frequency ≤ 10 or coverage ≤ 20 were not considered.

[00125] For case 13 WES, a library was prepared using the SureSelect XT Human All Exon V5 kit (Agilent Technologies) according to the manufacturer's instructions, followed by sequencing on a HiSeq2500 (Illumina). Raw data files were converted to FASTQ files with

the bcl2fastq software package version 1.8.4 (Illumina). FASTQ files were mapped by Novoalign version 3 (Novocraft) to the hg19 human reference genome sequence. In this step, SNV information in dbSNP build 138 was used for base quality score recalibration. Marking of PCR duplicates and position-wise sorting was performed with Novosort version 3 (Novocraft). Calling of SNVs and small indels was performed using GATK version 3.4-46. A GATK workflow was used in which local realignment and variant calling were performed by IndelRealigner and HaplotypeCaller, respectively. Low quality SNV and small indel calls were removed using the following criteria: QD<2.0, MQ<40.0, FS>60.0, MQRankSum<-12.5, ReadPosRankSum<-8.0 for SNVs; QD<2.0, ReadPosRankSum<-20.0, FS>200.0 for small indels. SNVs and small indels were annotated with the ANNOVAR software package using the following datasets and programs: gene information from GENCODE (version 19); allele frequencies of the 1000 Genome Project (version August, 2015), ExAC (version 0.3; see URLs), EVS (release ESP6500SI-V2; see URLs) and an in-house database; and predictions of protein damage by PolyPhen-2 and SIFT via dbNSFP (version 3.0).

[00126] **DNA methylation analysis**

[00127] **Sodium bisulfite sequencing.**

[00128] DNA methylation was analyzed at single base resolution after sodium bisulfite modification, PCR amplification, cloning and Sanger sequencing. Briefly, 2 μg of genomic DNA was denatured for 30 minutes at 37°C in NaOH 0.4N and incubated overnight in a solution of sodium bisulfite 3M pH 5 and 10 mM hydroquinone using a previously described protocol. Converted DNA was then purified using the Wizard DNA CleanUp kit (Promega) following the manufacturer's recommendations and precipitated by ethanol precipitation for 5 hours at -20°C. After centrifugation, DNA pellets were resuspended in 20 μL of water and stored at -20°C until use. Converted DNA was then amplified using primer sets (Table 3) designed with the MethPrimer software avoiding the presence of CpGs in the primer sequence in order to amplify methylated and unmethylated DNA with the same efficiency. Amplification was carried out using High Fidelity Taq polymerase (Roche) according to the manufacturer's instructions. After initial denaturation at 94°C for two minutes, amplification was done at 94°C for 20 seconds, 54°C for 30 seconds and 72°C for one minute for 10 cycles, then at 94°C for 20 seconds, 54°C for 30 seconds, followed by an extension step of 4 minutes and 30 seconds for the first cycle and an increment of 30 seconds at each subsequent cycle for 25 cycles. At the end of the program, a final extension step at 72°C for 7 minutes was performed. PCR products were then purified using the Wizard SV gel and PCR Purification

system (Promega), resuspended in 50 µl of water and cloned using the pGEM®-T Easy Vector cloning kit (Promega). Colonies were grown overnight at 37°C with ampicillin selection and randomly selected colonies were PCR amplified directly using T7 or SP6 primers. For each sample and region, at least ten randomly cloned PCR products were sequenced according to Sanger's method by Eurofins MWG Operon (Ebersberg, Germany) with either SP6 or T7 primers. Sequences were analyzed using the BiQ Analyser software and the average methylation score was calculated as the number of methylated CpGs for the total number of CpGs in the reference sequence.

[00129] **Statistics and subjects.** The average methylation level of each group of samples (FSHD2 patients carrying a SMCHD1 mutation, control individuals and BAMS patients and their relatives) was compared using the Kruskal-Wallis non-parametric multiple comparisons test followed by a Dunn's comparison and Bonferroni correction, with $\alpha = 0.05$. Control individuals (n=8) were healthy donors that have been previously reported. The FSHD2 patients carrying a SMCHD1 mutation have been previously reported and comprise n=8 for the DR1 region and n=15 for each of the 5' and Mid regions, while for the DR1 region 21 additional patients for whom sodium bisulfite sequencing data exists in the Leiden Open variation Database (LOVD) SMCHD1 variant database were included.

[00130] **Smchd1-Hsp90 structure modeling and multiple sequence alignment**

[00131] A homology model of the N-terminal region of Smchd1 was generated using the online server Phyre2 (Protein Homology/Analogy Recognition Engine 2). Protein sequence of 111-702 aa of mouse Smchd1 was submitted as the input sequence and intensive modelling mode was selected. The second highest scoring model with the most sequence alignment coverage based on the crystal structure of yeast Hsp90 (PDB: 2CG9) was elected for further evaluation. The model was visualized in PyMOL. The multiple sequence alignment was generated using CLUSTAL W (via the PBIL server) and ESPript 3.0.

[00132] **ATPase assays**

[00133] Cloning, expression and purification of recombinant mouse Smchd1 protein was performed as previously described, and the primers used for cloning and mutagenesis are provided in Table 4. The purity of the protein preparations was judged by migration of samples on 4-20% (w/v) Tris-Glycine reducing SDS/PAGE gels followed by staining with SimplyBlue SafeStain (Thermo Fisher Scientific, USA) (Figure 14). The ATPase assay was performed with the Transcreener ADP2 fluorescence polarization assay kit (BellBrook Labs) as previously described. Briefly, 10 µl reactions in triplicate were set up in 384-well (low

volume, black) plates, containing 7 µl reaction buffer (50 mM HEPES pH 7.5, 4 mM $MgCl_2$ and 2 mM EGTA), 1 µl of recombinant Smchd1 111-702 aa protein at concentrations ranging from 0.1-0.6 µM or buffer control, 1 µl of radicicol or solvent control and 1 µl of 10 µM ATP substrate or nuclease-free water control. Hsp90 inhibitor radicicol (Sigma-Aldrich) was dissolved in 70% ethanol and further diluted to a final concentration ranging from 0.1 nM to 10 µM. A 12-point 10 µM ADP/ATP standard curve was set up in parallel. Reactions were incubated at room temperature for 1 hour in the dark before addition of 10 µl of detection mix (1X Stop & Detection Buffer B, 23.6 µg/ml ADP2 antibody) for a further hour of incubation. Fluorescence polarization readings were performed with an Envision plate reader (PerkinElmer Life Sciences) following the manufacturer's instructions. The amount of ADP present in each reaction was estimated by using the standard curve following the manufacturer's instructions.

[00134] Table 3. Primers used for sodium bisulfite PCR

Name	**Size (bp)**	**# of CpG**	**Forward**	**Reverse**
DR1	255	31	GAAGGTAGGGAGGAAAAG (SEQ ID NO: 11)	ACTCAACCTAAAAATATACAATCT (SEQ ID NO: 12)
5'	275	21	AAATATGTAGGGAAGGGTGTAAGTT (SEQ ID NO: 13)	GGAGAGAGGGTTTGGTATATTTAAG (SEQ ID NO: 14)
Mid	354	31	ATTCATGAAGGGGTGGAGCCT (SEQ ID NO: 15)	CAGAGAACGGCTGGCCCAGGCCAT (SEQ ID NO: 16)

Table 4. Primers used for cloning and mutagenesis of recombinant murine Smchd1

Primers for generation of expression constructs for recombinant murine Smchd1 protein		SEQ ID NO:
5' BamHI 111 aa	CGC GGATCC acg aaa gaa aga att gac ttt cta cct c	17
3' EcoRI 702 aa	CGGAATTCA ttc atc tcc ttc agg cca agt tac aga c	18
Primers for oligonucleotide-directed mutagenesis		
p. Ala134Ser Forward	atg tat gag tat tat Tcg agt gaa gga cag aat	19
p. Ala134Ser Reverse	att ctg tcc ttc act cgA ata ata ctc ata cat	20
p. Ser135Cys Forward	g tat gag tat tat gcg Tgt gaa gga cag aat cct	21
p. Ser135Cys Reverse	agg att ctg tcc ttc acA cgc ata ata ctc ata c	22
p. Glu136Gly Forward	gag tat tat gcg agt gGa gga cag aat cct ttg	23
p. Glu136Gly Reverse	caa agg att ctg tcc tCc act cgc ata ata ctc	24

p. Asp420Val Forward	cac gtt gaa gga gTc ggt gta gtg gaa g	25
p. Asp420Val Reverse	c ttc cac tac acc gAc tcc ttc aac gtg	26
p. Tyr353Cys Forward	cat att tat cat tac tGt att cat gga cca aaa g	27
p. Tyr353Cys Reverse	c ttt tgg tcc atg aat aCa gta atg ata aat atg	28
p. Thr527Met Forward	c agc aca aat aaa ctg aTG ttt atg gat ctt gag ctg	29
p. Thr527Met Reverse	cag ctc aag atc cat aaa CAt cag ttt att tgt gct g	30

[00135] **Mouse embryo dissection and X-gal staining**

[00136] Mice were housed and mouse work approved under the Walter and Eliza Hall Institute of Medical Research Animal Ethics Committee approval (AEC 2014.026). Embryos were produced from C57BL/6 Smchd1$^{gt/+}$ congenic strain sires mated with C57BL/6 dams, with embryo ages ranging from embryonic day 8.5 to embryonic day 12.55. All embryos analyzed were female. No randomization or blinding was used during the experimental procedure. Embryos were briefly fixed in 2 % paraformaldehyde/0.2 % glutaraldehyde and stained in 1 mg/ml X-gal for several hours. Cryosections were cut at 12 µm.

[00137] **Xenopus embryological assays**

[00138] Xenopus *laevis* were used according to guidelines approved by the Singapore National Advisory Committee on Laboratory Animal Research. Protocols for fertilization, injections and whole mount in situ hybridization are according to methods known in the art. Human SMCHD1 (Origene) was cloned into expression vector pCS2+, linearized with NotI and transcribed with mMESSAGE mMachine SP6 transcription kit (Thermo Fisher). Transcribed mRNA was column purified and its concentration measured using a Nanodrop. The mRNA contains a poly A signal that allows for polyadenylation *in vivo*. To specifically target the cells destined to contribute to anterior head tissue, the two dorsal-animal blastomeres were injected at the 8-cell stage with the synthesized mRNA. Embryos were allowed to develop at room temperature until stage 45-46 (4 days post fertilization) and fixed. Eye diameter was measured using a Leica stereomicroscope with a DFC 7000T digital camera. No statistical method was used to predetermine sample size. No randomization or blinding was used. Embryos that died before gastrulation were excluded. Statistics were calculated using 1 way ANOVA, followed by Dunn's multiple comparison test. A p value of less than 0.05 was considered significant. Results are given as means ± s.d. Injections were performed on multiple clutches to reduce clutch-specific bias. Embryonic extracts were prepared by lysing Stage 12 embryos in CelLytic Express (Sigma) on ice, followed by

centrifugation to remove yolk proteins. Extracts were analyzed by Western blot with anti-SMCHD1 (Atlas HPA039441) and anti-β-Actin antibodies (clone AC-74, Sigma). cDNA was made from RNA extracted from *Xenopus laevis* embryos of various stages using iScript reverse transcriptase (Bio-Rad). qPCR was performed using the following primers, xsmchd1_qPCR_F 5'- CAGTGGGTGTCATGGATGCT (SEQ ID NO: 31), xsmchd1_qPCR_R 5'- TCCATGGCTAGACCACTTGC (SEQ ID NO: 32), XL_18S_F 5'-GCAATTATTTCCCATGAACGA (SEQ ID NO: 33), XL_18S_R 5'-ATCAACGCGAGCTTATGACC (SEQ ID NO: 34). In situ hybridization probe for *smchd1* was amplified from stage 20 cDNA using primers 5'-CGAATGCAAAGTCCTTGGGC (SEQ ID NO: 35) and 5'-GCATCCATGACACCCACTGA (SEQ ID NO: 36), cloned into pGEM-T, linearized and transcribed using DIG-labelling mix (Roche) according to manufacturer's guidelines.

[00139] **DNA damage response assays**

[00140] **Cell lines and cell cultures.** XRCC4-deficient cells and primary fibroblast cell lines established from cases 1 and 2 were cultured in Dulbecco's modified Eagle medium (DMEM, Gibco) supplemented with 10% fetal calf serum (FCS, Gibco), and antibiotics. Testing for mycoplasma contamination was negative. For H2AX activation, cells were either irradiated with 100 J/m^2 UV-C or treated with 50 μM etoposide (Sigma-Aldrich, USA) for 1 hour. Drugs were then washed out, fresh media was added, and cells were incubated for 6 hrs and then subjected to Western blot analysis.

[00141] Protein isolation and analysis. Cells were solubilized by using ice-cold RIPA buffer (10 mM Tris, pH: 8.0; 150 mM NaCl; 1 mM EDTA; 10 mM NaF; 1 mM Na_3VO_4; 10 μM Na_2MoO_4; 1% NP-40; 0.25% SDS; protease inhibitors P 2714 [Sigma-Aldrich, USA]). The total protein concentration of extracts was determined using the BCA Protein Assay Kit (Thermo Fisher Scientific, USA). 10 μg of total cell lysates were separated by 4-12 % SDS-PAGE (Invitrogen, Germany) and blotted onto nitrocellulose membranes (GE Healthcare, Germany). Protein detection was performed using antibodies specific for phosphorylation of H2AX at Ser139 (γH2AX) (clone 20E3, Cell Signaling Technology, USA). Anti-β-Actin antibodies were purchased from Sigma-Aldrich (clone AC-74). Secondary antibodies conjugated to peroxidase (Santa Cruz Biotechnology Inc., USA) were used and blots were developed using an enhanced chemiluminescence system, ECL Plus (GE Healthcare), followed by detection on autoradiographic films.

[00142] **Microhomology-mediated End-Joining (MMEJ) assay.** The MMEJ assay using linearized pDVG94 plasmid was performed as previously described. In brief, cells were transfected with 2 μg EcoRV/AfeI (Thermo Fisher Scientific, Germany; New England Biolabs, Germany)-linearized pDVG94 and extrachromosomal DNA was isolated 48 h after transfection. PCR analysis was performed, PCR products were digested using BstXI, separated by gel electrophoresis and visualized by ethidium bromide staining.

[00143] **Results**

[00144] Congenital absence of the nose (arhinia) is a rare and striking condition with less than 50 patients reported to date. Arhinia is variably associated with absent paranasal sinuses, hypertelorism, microphthalmia, colobomas, nasolacrimal duct abnormalities, mid-face hypoplasia, high-arched palate, absent olfactory bulbs and defects of the reproductive axis in males. In its most severe presentation, consisting of nasal, ocular and reproductive defects, it is referred to as Bosma arhinia microphthalmia syndrome (BAMS) (OMIM 603457). Arhinia is presumed to result from a specific defect of the nasal placodes or surrounding neural crest-derived tissues during embryonic development, but a genetic cause has not been established.

[00145] Fourteen unrelated individuals with isolated arhinia or a syndromic presentation compatible with BAMS were investigated (Figure 1a-l, Figure 4). Trio or quartet whole-exome sequencing (WES) for cases 1, 2 and 9-12 led to the identification of de novo heterozygous missense mutations in the Structural Maintenance of Chromosomes Flexible Hinge Domain Containing 1 (SMCHD1) gene in all six cases (Figure 1m, Table 1 and Table 2), which were confirmed by Sanger sequencing (Figure 5). Singleton WES for case 13 also identified a SMCHD1 mutation. Sanger sequencing of SMCHD1 was then performed to the remaining seven BAMS patients. Heterozygous missense mutations were identified in all. In total, 10 out of 14 variants were de novo, suggesting germline mutations in parental gametes, while in four cases parental DNA was not available (Figure 1 and Table 1). None of the identified mutations have been reported in the ExAC, EVS or dbSNP144 databases (accessed via the UCSC browser, November 2016), all mutations affect highly conserved residues (Figure 6) and all are predicted damaging by PolyPhen-2 (Table 1). Remarkably, all 14 mutations are located in exons 3, 8-10, 12 and 13 of SMCHD1 (48 exons total); these exons code for the ATPase domain of SMCHD1 and an associated region immediately C-terminal (see further below). Notably, six of the 14 patients had mutations affecting three adjacent amino acids: Ala134, Ser135 and Glu136, while p.His348Arg and p.Asp420Val were identified in three and two independent patients respectively, suggesting possible hotspots

(Figure 1m). Mutations in SMCHD1 in arhinia patients have also been identified in an independent study that includes six of the cases analyzed here (cases 2, 4, 5, 6, 7 and 13; Shaw et al, accompanying manuscript).

[00146] Table 1 SMCHD1 mutations identified in patients 1 to 14

Geographic origin	**Nucleotide change (1)**	**Amino acid change**	**Predicted functional effect (Polyphen-2 score) (2)**	**Mutation origin**
1 Morocco	c.407A>G	p.Glu136Gly	0.999	de novo
2 Germany	c.403A>T	p.Ser135Cys	1.000	de novo
3 North Africa	c.404G>A	p.Ser135Asn	0.997	de novo
4 Ireland	c.403A>T	p.Ser135Cys	1.000	de novo
5 (Chinese, by name)	c.1043A>G	p.His348Arg	0.998	unknown (parental DNA unavailable)
6 Scotland	c.1259A>T	p.Asp420Val	0.877	de novo
7 Japan	c.1655G>A	p.Arg552Gln	1.000	de novo
8 Wales	c.1552A>G	p.Lys518Glu	0.976	unknown (parental DNA unavailable)
9 Thailand	c.1259A>T	p.Asp420Val	0.877	de novo
10 Thailand	c.1025G>C	p.Trp342Ser	0.999	de novo
11 Turkey	c.400G>T	p.Ala134Ser	0.999	de novo
12 Turkey	c.400G>T	p.Ala134Ser	0.999	de novo
13 Norway	c.1043A>G	p.His348Arg	0.998	unknown (parental DNA unavailable)
14 Ukraine	c.1043A>G	p.His348Arg	0.998	unknown (parental DNA unavailable)

(1) NCBI Reference Sequence: NM_015295.2

(2) UniProtKB identifier: A6NHR9

[00147] During craniofacial development, the olfactory placode ectoderm thickens and invaginates to form the olfactory epithelium within the nasal cavity, a process that depends on cross-talk between the placodal epithelium and the underlying cranial neural crest-derived mesenchyme. For example, ablation of the nasal placode epithelium in chick embryos disrupts development of adjacent nasal skeletal elements. Strong X-gal staining was observed in the developing face of mouse embryos expressing lacz from the Smchd1 locus, including in the nasal placodes and optic vesicles at E9.5 and nasal epithelium at E12.5 (Figure 7).

Eurexpress *in situ* hybridization data indicates regional expression of Smchd1 in the nasal cavity at E14.5, while transcriptional profiling of post-natal olfactory epithelium demonstrated that Smchd1 is specifically expressed in immature olfactory sensory neurons. These data are consistent with roles for SMCHD1 during early nasal development. Gonadotropin-releasing hormone (GnRH) neurons migrate from the olfactory placode along olfactory axon tracts to the hypothalamus, where they regulate reproductive hormone release from the pituitary gland. Defects of the reproductive axis have occasionally been reported in males with arhinia; this finding confirmed and also report pubertal delay or anomalies of menarche in all three post-pubertal age females in the series disclosed herein. The reproductive axis defects associated with arhinia are likely secondary to a defect in GnRH neuron production in, or migration from, the olfactory placode.

[00148] *Smchd1* was identified as a modifier of transgene silencing in mice and was subsequently shown to be involved in X chromosome inactivation, being required for CpG island (CGI) methylation on the inactive X (Xi), CGI-independent silencing of some X chromosome genes, and Xi compaction. In addition, *Smchd1* functions as an epigenetic repressor at various autosomal loci, with dysregulation of imprinted and monoallelically-expressed gene clusters observed in mutant mice. A requirement for SMCHD1 in repair of DNA double-strand breaks has also been demonstrated. Whereas female mice null for Smchd1 display midgestation lethality due to derepression of inactive X chromosome genes, male mutant mice display perinatal lethality of undescribed causes in certain strains or viability on the FVB/n background. Strikingly, craniofacial abnormalities have not been documented in Smchd1 loss-of-function mice regardless of their sex.

[00149] Recently, haploinsufficiency of SMCHD1 was reported as a cause of facioscapulohumeral muscular dystrophy (FSHD) type 2 (FSHD2) (OMIM 158901). FSHD has a prevalence of 1/20,000, with FSHD type 1 (FSHD1) and FSHD2 accounting for ~95% and ~5% of cases, respectively. FSHD results from pathogenic misexpression of the transcription factor DUX4 (encoded by an array of D4Z4 repeats on chromosome 4q) in skeletal muscle. In FSHD1 (OMIM 158900), D4Z4 repeat contraction leads to hypomethylation of the locus and derepression of DUX4 expression on a permissive haplotype (4qA) that harbors a stabilizing polyadenylation signal for DUX4 mRNA. FSHD2 occurs in individuals harboring loss-of-function SMCHD1 mutations and the permissive 4qA allele, without the requirement for D4Z4 repeat contraction, although SMCHD1 mutations can also modify the severity of FSHD1. SMCHD1 is thought to function as a silencer at the

4q locus via binding to the D4Z4 repeats. Over 80 unique, putatively pathogenic SMCHD1 variants have been reported in FSHD2 patients (LOVD SMCHD1 variant database). These mutations, which include clear loss-of-function alleles, occur throughout the protein, and are not clustered in specific domains. Several loss- of-function mutations have also been reported in ExAC (Figure 1m), and over 60 deletions affecting SMCHD1 have been reported in the DECIPHER database (available phenotypic information does not indicate arhinia). The methylation status of D4Z4 repeats was analyzed in peripheral blood leukocytes in BAMS patients by sodium bisulphite sequencing (Table 2 and Figures 8 and 9). Although a trend for hypomethylation was noted for BAMS patients relative to controls or unaffected family members, depending on the site tested within D4Z4, some BAMS patients were normally methylated. A large variability in D4Z4 methylation has also been observed in controls and FSHD patients, and is not an absolute indicator of FSHD. Moreover, an important argument against BAMS and FSHD2 mutations acting in the same direction is the absence of BAMS and FSHD co-occurring in the same patient in the literature. None of the BAMS patients reported here have signs of muscular dystrophy, including both the individuals (2 and 12) older than the average age of FSHD2 onset of 26 years, and none of the BAMS missense mutations identified in the present disclosure have been associated with FSHD2.

[00150] **Table 2** DNA methylation analysis in BAMS probands and family members

Sample ID	**Gender**	**Clinical Status**	**Mutation**	**DR1 %M+**	**5' %M+**	**Mid %M+**
Indiv-1	M	Proband	E136G	5	54	73
Indiv-1 father	M	CTRL	-	69	72	64
Indiv-1 mother	F	CTRL	-	56	73	73
Indiv-2	F	Proband	S135C	19	48	60
Indiv-2 mother	F	CTRL	-	74	86	71
Indiv-2 father	M	CTRL	-	45	77	X
Indiv-3	M	Proband	S135N	22	44	62
Indiv-3 mother	F	CTRL	-	58	82	83
Indiv-3 father	M	CTRL	-	62	75	75
Indiv-4	F	Proband	S135C	21	44	70
Indiv-4 father	M	CTRL	-	44	69	79
Indiv-4 mother	F	CTRL	-	31	50	X

Indiv-4 sister	F	CTRL	-	18	62	X
Indiv-4 brother	M	CTRL	-	26	70	X
Indiv-5	M	Proband	H348R	66	56	58
Indiv-6	M	Proband	D420V	19	41	68
Indiv-6 father	M	CTRL	-	50	67	50
Indiv-6 mother	F	CTRL	-	X	87	69
Indiv-8	F	Proband	K518E	33	74	71
Indiv-14	F	Proband	H348R	35	49	63
Indiv-9	F	Proband	D420V	25	50	66
Indiv-9 father	M	CTRL	-	29	74	75
Indiv-9 mother	F	CTRL	-	76	81	69
Indiv-11 father	M	CTRL	-	50	72	71
Indiv-11 mother	F	CTRL	-	80	81	69
Indiv-11	F	Proband	A134S	17	63	68
Indiv-10	F	Proband	W342S	54	69	X
Indiv-10 mother	F	CTRL	-	27	50	73
Indiv-10 sister	F	CTRL	-	86	75	68

[00151] Proteins of the SMC family are involved in chromatid cohesion, condensation of chromosomes and DNA repair. SMCHD1 is considered a non-canonical member of the family, with a C-terminal chromatin-binding hinge domain and an N-terminal GHKL (gyrase, Hsp90, histidine kinase, and MutL) ATPase domain21 (Figure 1m). Potentially, SMCHD1 uses energy obtained from ATP hydrolysis to manipulate chromatin ultrastructure and interactions. Using small angle X-ray scattering, the purified Smchd1 ATPase domain and an adjacent C-terminal region (amino acids 111-702 for the two regions combined; denoted "N-terminal region" in Figure 1m) have been shown to adopt a structural conformation similar to Hsp9021. Consistent with this, the Hsp90 inhibitor radicicol decreased the ATPase activity of Smchd1. Mapping of the SMCHD1 amino acids mutated in BAMS and FSHD2 to the homology model of Smchd1 based on the Hsp90 crystal structure indicates that the major cluster of BAMS mutations (amino acids 134-136) is situated immediately N-terminal to Motif I, which is highly conserved among the GHKL-ATPases and participates in coordination of the Mg^{2+}-ATP complex during ATP hydrolysis23 (Figure 6 and 10). The finding of other BAMS mutations in the region immediately C terminal to the ATPase

domain supports the idea that this extended region has a function intimately associated with that of the ATPase domain. Without wishing to be bound by theory, given that (i) loss-of-function of SMCHD1 causes FSHD2, (ii) FSHD is not known to co-occur with arhinia, (iii) there are no visible craniofacial anomalies in Smchd1 null mice, (iv) the mutations in BAMS patients are clustered in the extended ATPase domain and (v) in contrast to SMCHD1 depletion, BAMS mutations do not cause DNA damage response alterations or impaired non-homologous end joining (Figure 11), the inventors believed that the BAMS mutations may result in a gain- rather than a loss-of-function of the SMCHD1 protein. To test this belief, ATPase assays were conducted using the purified recombinant N-terminal region harboring BAMS or FSHD2 mutations. Compared to wildtype, hydrolysis of ATP was increased for the N-terminal region containing the mutations p.Ala134Ser, p.Ser135Cys or p.Glu136Gly, strongly or slightly decreased for the FSHD2 mutations p.Tyr353Cys15 or p.Thr527Met18, respectively, and unchanged for the BAMS mutation p.Asp420Val (Figure 2a-f). The half-maximal inhibitory concentration (IC50) of radicicol was similar for BAMS mutant and wildtype recombinant protein ATPase activities (Figure 12), indicating that the mutants retain an intact ATP-binding site. These results indicates that BAMS-associated mutations elevate the catalytic activity of SMCHD1.

[00152] These biochemical results were then validated *in vivo* using full-length SMCHD1 protein. In *Xenopus laevis*, the expression of *smchd1* begins zygotically, and rises steadily after gastrulation (Figure 3a). Endogenous smchd1 is strongly enriched in the head region and the neural tube (Figure 3b). To faithfully recapitulate this expression pattern, the two dorsal-animal blastomeres of 8-cell stage Xenopus embryos were micro-injected with 120 pg of capped mRNA encoding either wildtype or mutant human SMCHD1 (Figure 3c). Each set of injected embryos produced comparable human SMCHD1 protein levels (Figure 3g). However, only tadpoles overexpressing SMCHD1 mRNA with BAMS mutations showed noticeable craniofacial anomalies (Figure 3d-f, Figure 13), including microphthalmia and in severe cases, anophthalmia (Figure 3f'). At 4 days post fertilization, quantification of the eye size showed a marked reduction in the eye diameter in tadpoles overexpressing BAMS mutants whereas tadpoles overexpressing wildtype SMCHD1 or p.Tyr353Cys, an FSHD2 mutation, were indistinguishable from control uninjected embryos (Figure 3h). One of the BAMS mutants with phenotypic effects in this assay, p.Asp420Val, showed no change in ATPase activity *in vitro* (Figure 2), showing higher sensitivity of the *in vivo* assay. Whole mount *in situ* hybridization showed a decrease in the size of the eye and nasal placodes,

marked by rx2a and six1 respectively, upon overexpression of a BAMS mutant (Figure 3i,j). In contrast, migration of cranial neural crest, marked by twist1, was largely unaffected. Craniofacial anomalies were dose-dependent for both wildtype and BAMS mutant SMCHD1 injections, while overexpression of the FSHD2 mutant p.Tyr353Cys was without effect regardless of dose (Figure 3k, Figure 13). The finding that wildtype SMCHD1, when overexpressed at a sufficiently high dose, acts in the same phenotypic direction as the BAMS mutants shows that these mutants may at least in part act by augmenting the normal activity of the protein. These *in vivo* results, which partially recapitulate the microphthalmia and facial hypoplasia seen in severe BAMS patients, further support the notion that, in contrast to FSHD2 alleles, BAMS associated missense mutations exhibit gain-of-function or neomorphic activity.

[00153] In conclusion, the present inventors have surprisingly identified *de novo* missense gain-of-function mutations restricted to the extended ATPase domain of SMCHD1 as the cause of isolated arhinia and BAMS. It will be of great interest to explore the epistatic relationships between SMCHD1 and known regulators of nasal development, such as PAX6 and FGF and BMP signaling, as well as to uncover other potential human-specific nasal regulators. Nose shape and size vary greatly between human populations and even more drastically among animal species, the elephant's trunk being an extreme example. As such, it will be interesting to determine the role of SMCHD1 in controlling nose size from an evolutionary perspective.

[00154] Given that loss-of-function mutations in SMCHD1 are associated with FSHD2, BAMS and FSHD2 represent a rare example of gain- versus loss-of-function mutations in the same gene leading to vastly different human disorders, in terms of the affected tissue and age of onset. Thus, the present disclosure shows that since FSHD is caused in part by a loss of SMCHD1, augmentation of the expression or activity of SMCHD1 in affected muscles can surprisingly be developed as a form of treatment. The identification of gain-of-function mutations in SMCHD1 in the present disclosure can inform gene therapy approaches, or in combination with future structural studies on the effect of these mutations on the ATPase domain, aid the design of drugs that reproduce the gain-of-function effect, for treatment of FSHD. Importantly for such an approach, the deleterious consequences of SMCHD1 gain-of-function appear restricted to a narrow window of human embryonic development.

[00155] **References**

1. Brasseur, B., Martin, C. M., Cayci, Z., Burmeister, L. & Schimmenti, L. A. Bosma arhinia microphthalmia syndrome: Clinical report and review of the literature. Am. J. Med. Genet. A. 170, 1302–1307 (2016).

2. Graham, J. M. & Lee, J. Bosma arhinia microphthalmia syndrome. Am. J. Med. Genet. A. 140, 189–193 (2006).

3. Forni, P. E. & Wray, S. GnRH, anosmia and hypogonadotropic hypogonadism- -where are we? Front. Neuroendocrinol. 36, 165–177 (2015).

4. Szabo-Rogers, H. L. et al. Novel skeletogenic patterning roles for the olfactory pit. Development 136, 219–29 (2009).

5. Blewitt, M. E. et al. SmcHD1, containing a structural-maintenance-of chromosomes hinge domain, has a critical role in X inactivation. Nat. Genet. 40, 663–669 (2008).

6. Nickell, M. D., Breheny, P., Stromberg, A. J. & McClintock, T. S. Genomics of mature and immature olfactory sensory neurons. J. Comp. Neurol. 520, 2608–2629 (2012).

7. Tryggestad, J. B., Li, S. & Chernausek, S. D. Hypogonadotropic hypogonadism presenting with arhinia: a case report. J. Med. Case Reports 7, 52 (2013).

8. Gendrel, A.-V. et al. Smchd1-dependent and -independent pathways determine developmental dynamics of CpG island methylation on the inactive X chromosome. Dev. Cell 23, 265–279 (2012).

9. Gendrel, A.-V. et al. Epigenetic functions of smchd1 repress gene clusters on the inactive X chromosome and on autosomes. Mol. Cell. Biol. 33, 3150–3165 (2013).

10. Nozawa, R.-S. et al. Human inactive X chromosome is compacted through a PRC2-independent SMCHD1-HBiX1 pathway. Nat. Struct. Mol. Biol. 20, 566–573 (2013).

11. Mould, A. W. et al. Smchd1 regulates a subset of autosomal genes subject to monoallelic expression in addition to being critical for X inactivation. Epigenetics Chromatin 6, 19 (2013).

12. Chen, K. et al. Genome-wide binding and mechanistic analyses of Smchd1- mediated epigenetic regulation. Proc. Natl. Acad. Sci. U. S. A. 112, E3535–3544 (2015).

13. Coker, H. & Brockdorff, N. SMCHD1 accumulates at DNA damage sites and facilitates the repair of DNA double-strand breaks. J. Cell Sci. 127, 1869–1874 (2014).

14. Tang, M. et al. Structural maintenance of chromosomes flexible hinge domain containing 1 (SMCHD1) promotes non-homologous end joining and inhibits homologous recombination repair upon DNA damage. J. Biol. Chem. 289, 34024–34032 (2014).

15. Lemmers, R. J. L. F. et al. Digenic inheritance of an SMCHD1 mutation and an FSHD-permissive D4Z4 allele causes facioscapulohumeral muscular dystrophy type 2. Nat. Genet. 44, 1370–1374 (2012).

16. Hewitt, J. E. Loss of epigenetic silencing of the DUX4 transcription factor gene in facioscapulohumeral muscular dystrophy. Hum. Mol. Genet. 24, R17–23 (2015).

17. Lemmers, R. J. L. F. et al. A unifying genetic model for facioscapulohumeral muscular dystrophy. Science 329, 1650–1653 (2010).

18. Sacconi, S. et al. The FSHD2 gene SMCHD1 is a modifier of disease severity in families affected by FSHD1. Am. J. Hum. Genet. 93, 744–751 (2013).

19. Lemmers, R. J. L. F. et al. Inter-individual differences in CpG methylation at D4Z4 correlate with clinical variability in FSHD1 and FSHD2. Hum. Mol. Genet. 24, 659–669 (2015).

20. de Greef, J. C. et al. Clinical features of facioscapulohumeral muscular dystrophy 2. Neurology 75, 1548–1554 (2010).

21. Chen, K. et al. The epigenetic regulator Smchd1 contains a functional GHKL type ATPase domain. Biochem. J. (2016). doi:10.1042/BCJ20160189

22. Brideau, N. J. et al. Independent Mechanisms Target SMCHD1 to Trimethylated Histone H3 Lysine 9-Modified Chromatin and the Inactive X Chromosome. Mol. Cell. Biol. 35, 4053–4068 (2015).

23. Dutta, R. & Inouye, M. GHKL, an emergent ATPase/kinase superfamily. Trends Biochem. Sci. 25, 24–28 (2000).

24. McKenna, A. et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res 20, 1297–303 (2010).

25. Li, H. et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25, 2078–9 (2009).

26. Flicek, P. et al. Ensembl 2013. Nucleic Acids Res 41, D48–55 (2013). 27. DePristo, M. A. et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat. Genet. 43, 491–498 (2011).

28. Van der Auwera, G. A. et al. From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. Curr. Protoc. Bioinforma. Ed. Board Andreas Baxevanis Al 43, 11.10.1–33 (2013).

29. Javed, A., Agrawal, S. & Ng, P. C. Phen-Gen: combining phenotype and genotype to analyze rare disorders. Nat. Methods 11, 935–937 (2014).

30. Sherry, S. T. et al. dbSNP: the NCBI database of genetic variation. Nucleic Acids Res. 29, 308–311 (2001).

31. Mishima, H., Sasaki, K., Tanaka, M., Tatebe, O. & Yoshiura, K.-I. Agile parallel bioinformatics workflow management using Pwrake. BMC Res. Notes 4, 331 (2011).

32. Wang, K., Li, M. & Hakonarson, H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res. 38, e164 (2010).

33. Harrow, J. et al. GENCODE: the reference human genome annotation for The ENCODE Project. Genome Res. 22, 1760–1774 (2012).

34. 1000 Genomes Project Consortium et al. A global reference for human genetic variation. Nature 526, 68–74 (2015).

35. Adzhubei, I. A. et al. A method and server for predicting damaging missense mutations. Nat. Methods 7, 248–249 (2010).

36. Kumar, P., Henikoff, S. & Ng, P. C. Predicting the effects of coding non synonymous variants on protein function using the SIFT algorithm. Nat. Protoc. 4, 1073–1081 (2009).

37. Liu, X., Jian, X. & Boerwinkle, E. dbNSFP: a lightweight database of human nonsynonymous SNPs and their functional predictions. Hum. Mutat. 32, 894–899 (2011).

38. Liu, X., Wu, C., Li, C. & Boerwinkle, E. dbNSFP v3.0: A One-Stop Database of Functional Predictions and Annotations for Human Nonsynonymous and Splice-Site SNVs. Hum. Mutat. 37, 235–241 (2016).

39. Magdinier, F. et al. Regional methylation of the 5' end CpG island of BRCA1 is associated with reduced gene expression in human somatic cells. FASEB J. Off. Publ. Fed. Am. Soc. Exp. Biol. 14, 1585–1594 (2000).

40. Li, L.-C. & Dahiya, R. MethPrimer: designing primers for methylation PCRs. Bioinforma. Oxf. Engl. 18, 1427–1431 (2002).

41. Bock, C. et al. BiQ Analyzer: visualization and quality control for DNAmethylation data from bisulfite sequencing. Bioinforma. Oxf. Engl. 21, 4067–4068 (2005).

42. Gaillard, M.-C. et al. Differential DNA methylation of the D4Z4 repeat in patients with FSHD and asymptomatic carriers. Neurology 83, 733–742 (2014).

43. Gaillard, M.-C. et al. Segregation between SMCHD1 mutation, D4Z4 hypomethylation and Facio-Scapulo-Humeral Dystrophy: a case report. BMC Med. Genet. 17, 66 (2016).

44. Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N. & Sternberg, M. J. E. The Phyre2 web portal for protein modeling, prediction and analysis. Nat. Protoc. 10, 845–858 (2015).

45. Thompson, J. D., Higgins, D. G. & Gibson, T. J. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22, 4673–4680 (1994).

46. Robert, X. & Gouet, P. Deciphering key features in protein structures with the new ENDscript server. Nucleic Acids Res. 42, W320–324 (2014).

47. Rosin, N. et al. Mutations in XRCC4 cause primary microcephaly, short stature and increased genomic instability. Hum. Mol. Genet. 24, 3708–3717 (2015).

48. Verkaik, N. S. et al. Different types of V(D)J recombination and end-joining defects in DNA double-strand break repair mutant mammalian cells. Eur. J. Immunol. 32, 701–709 (2002).

Claims

1. An isolated variant of SMCHD1 (structural maintenance of chromosomes flexible hinge domain containing 1) peptide or fragment thereof, comprising one or more mutation(s) that increases SMCHD1 activity in a cell compared to a cell with a wild-type SMCHD1 peptide.

2. The isolated variant of SMCHD1 peptide of claim 1 or fragment thereof, wherein the mutation is comprised in one or more regions selected from the group consisting of:

the N-terminal motif I, which is the highly conserved region of GHKL-ATPases and participates in the coordination of the Mg^{2+}-ATP complex during ATP hydrolysis, and

the C-terminal to ATPase domain.

3. The isolated variant of SMCHD1 peptide of claim 1 or 2, wherein the mutation is comprised in region between amino acid corresponding to residue 111 to amino acid corresponding to residue 702 of SMCHD1.

4. The isolated variant of SMCHD1 of any one of claims 1 to 3, wherein the mutation is a substitution mutation.

5. The isolated variant of SMCHD1 of any one of claims 1 to 4, wherein the mutation is at one or more position(s) selected from the group consisting of:

amino acid residue corresponding to residue 134 of human SMCHD1,

amino acid residue corresponding to residue 135 of human SMCHD1,

amino acid residue corresponding to residue 136 of human SMCHD1,

amino acid residue corresponding to residue 342 of human SMCHD1,

amino acid residue corresponding to residue 348 of human SMCHD1,

amino acid residue corresponding to residue 420 of human SMCHD1,

amino acid residue corresponding to residue 518 of human SMCHD1, and

amino acid residue corresponding to residue 552 of human SMCHD1.

6. The isolated variant of SMCHD1 of any one of the preceding claims, wherein the mutation is one or more selected from the group consisting of:

amino acid residue corresponding to residue 134 of human SMCHD1 has been substituted with an amino acid with a polar uncharged side chain,

amino acid residue corresponding to residue 135 of human SMCHD1 has been substituted with an amino acid with polar uncharged side chain and/or an amino acid with sulfur side chain,

amino acid residue corresponding to residue 136 of human SMCHD1 has been substituted with an amino acid with an aliphatic side chain,

amino acid residue corresponding to residue 342 of human SMCHD1 has been substituted with an amino acid with polar uncharged side chain,

amino acid residue corresponding to residue 348 of human SMCHD1 has been substituted with an amino acid with positively charged side chain,

amino acid residue corresponding to residue 420 of human SMCHD1 has been substituted with an amino acid with hydrophobic side chain,

amino acid residue corresponding to residue 518 of human SMCHD1 has been substituted with an amino acid with negatively charged side chain, and

amino acid residue corresponding to residue 552 of human SMCHD1 has been substituted with an amino acid with a negative charged side chain.

7. The isolated variant of SMCHD1 of any one of the preceding claims, wherein the mutation is one or more selected from the group consisting of:

amino acid residue corresponding to residue 134 of human SMCHD1 has been substituted with a serine residue,

amino acid residue corresponding to residue 135 of human SMCHD1 has been substituted with a cysteine and/or a asparagine acid residue,

amino acid residue corresponding to residue 136 of human SMCHD1 has been substituted with a glycine residue,

amino acid residue corresponding to residue 342 of human SMCHD1 has been substituted with a serine residue,

amino acid residue corresponding to residue 348 of human SMCHD1 has been substituted with an arginine residue,

amino acid residue corresponding to residue 420 of human SMCHD1 has been substituted with a valine residue,

amino acid residue corresponding to residue 518 of human SMCHD1 has been substituted with a glutamic acid residue, and

amino acid residue corresponding to residue 552 of human SMCHD1 has been substituted with a glutamine residue.

8. An isolated nucleic acid or fragment thereof encoding the isolated variant SMCHD1 peptide according to any one of claims 1 to 7, wherein the nucleic acid comprises one or more mutation(s) that increases SMCHD1 protein activity in a cell compared to a cell with a wild-type SMCHD1 peptide.

9. The isolated nucleic acid of claim 8, wherein the SMCHD1 mutation is a missense mutation.

10. The isolated nucleic acid of claim 8 or 9, wherein the SMCHD1 mutation is at one or more position(s) selected from the group consisting of: c.407, c.403, c.404, c.1043, c.1259, c.1655, c.1552, c.1025, and c.400.

11. The isolated nucleic acid of any one of claims 8 to 10, wherein the SMCHD1 mutation is one or more selected from the group consisting of: c.407A>G, c.403A>T, c.404G>A, c.1043A>G, c.1259A>T, c.1655G>A, c.1552A>G, c.1025G>C, and c.400G>T (nucleic acid residue referenced to referenced to NCBI Reference Sequence: NM_015295.2).

12. A vector comprising the nucleic acid of any one of claims 8 to 11.

13. The isolated variant of SMCHD1 of any one of claims 1 to 7, or the isolated nucleic acid of any one of claims 8 to 11, or the vector of claim 12 for use in therapy.

14. A pharmaceutical composition comprising the isolated variant SMCHD1 peptide or fragment thereof of any one of claims 1 to 7, and/or an isolated nucleic acid or fragment thereof of any one of claims 8 to 11, or a vector of claim 12, and a pharmaceutically acceptable excipient thereof.

15. A gene-therapy composition comprising an isolated nucleic acid of any one of claims 8 to 11 or a vector of claim 12.

16. A host cell comprising the vector of claim 12.

17. A method of treating facioscapulohumeral muscular dystrophy (FSHD) in a subject in need thereof, comprising administering a therapeutically effective amount of an agent capable of increasing the expression level or activity of SMCHD1 peptide and/or nucleic acid to the subject in need thereof.

18. The method of claim 17, wherein the agent is one or more selected from the group consisting of an isolated variant SMCHD1 peptide or fragment thereof of any one of claims 1 to 7, a transgene encoding the isolated variant SMCHD1 peptide of any one of claims 1 to 7, an isolated nucleic acid or fragment thereof of any one of claims 8 to 11, a vector of claim 12, a pharmaceutical composition of claim 14, and a gene-therapy composition of claim 15.

19. The method of any one of claim 17 or 18, wherein the FSHD is a type 1 and/or type 2 FSHD.

20. A method of screening an agent capable of increasing SMCHD1 activity, comprising:

a. introducing the agent to an assay for determining SMCHD1 activity,

b. determining the SMCHD1 activity the agent elicits, and

c. comparing the SMCHD1 activity the agent elicits with the SMCHD1 activity elicited by one or more selected from the group consisting of an isolated variant SMCHD1 peptide or fragment thereof of any one of claims 1 to 7, a transgene encoding the isolated variant SMCHD1 peptide of any one of claims 1 to 7, an isolated nucleic acid or fragment thereof of any one of claims 8 to 11, and a vector of claim 12,

wherein when the SMCHD1 activity of the agent (or nucleotide mutation) is the same or more than the SMCHD1 activity of the one or more selected from the group consisting of the isolated variant SMCHD1 peptide or fragment thereof of any one of claims 1 to 7, the transgene encoding the isolated variant SMCHD1 peptide of any one of claims 1 to 7, the isolated nucleic acid or fragment thereof of any one of claims 8 to 11, and the vector of claim 12, the agent is considered a suitable agent for increasing SMCHD1 activity.

21. Use of one or more selected from the group consisting of an isolated variant SMCHD1 peptide of any one of claims 1 to 7, a transgene encoding the isolated variant SMCHD1 peptide of any one of claims 1 to 7, an isolated nucleic acid of any one of claims 8 to 11, a vector of claim 12, a pharmaceutical composition of claim 14, and a gene-therapy composition of claim 15 in the manufacture of a medicament for treating facioscapulohumeral muscular dystrophy (FSHD) in a subject in need thereof.

22. A kit for screening an agent capable of increasing SMCHD1 activity, comprising

a reference control comprising one or more selected from the group consisting of an isolated variant SMCHD1 peptide of any one of claims 1 to 7, a transgene encoding the isolated variant SMCHD1 peptide of any one of claims 1 to 7, an isolated nucleic acid of any one of claims 8 to 11, and a vector of claim 12; and

a reagent for determining SMCHD1 activity.

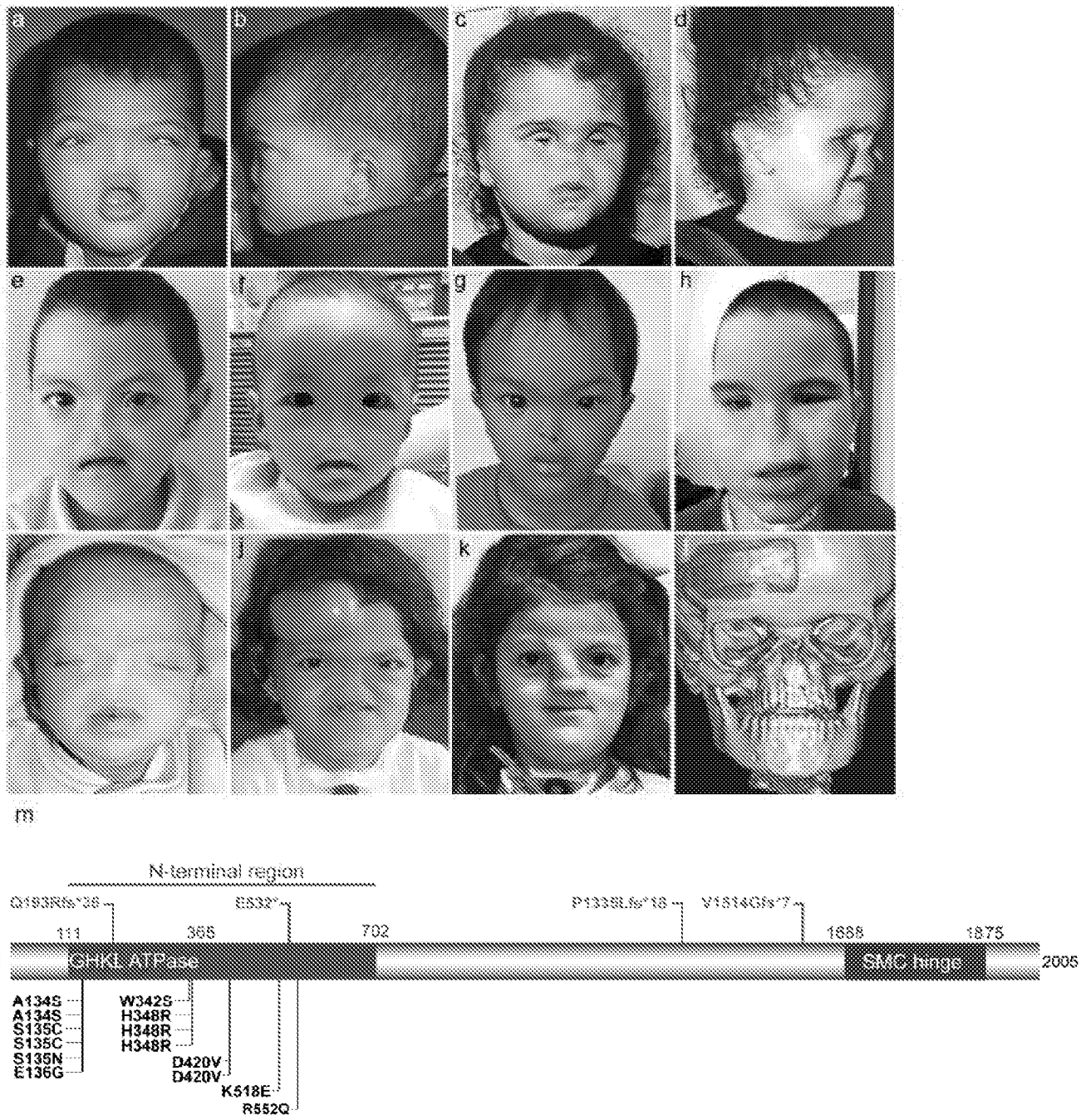

a
b
c
d
e
f
g
h
i
j
k
m
N-terminal region
Q193RfsX38
E532X
P1338LfsX18
V1514GfsX7
1
111
365
702
1688
1875
2005
GHKL ATPase
SMC hinge
A134S
A134S
S135C
S135C
S135N
E136G
W342S
H348R
H348R
H348R
D420V
D420V
K518E
R552Q

Figure 1

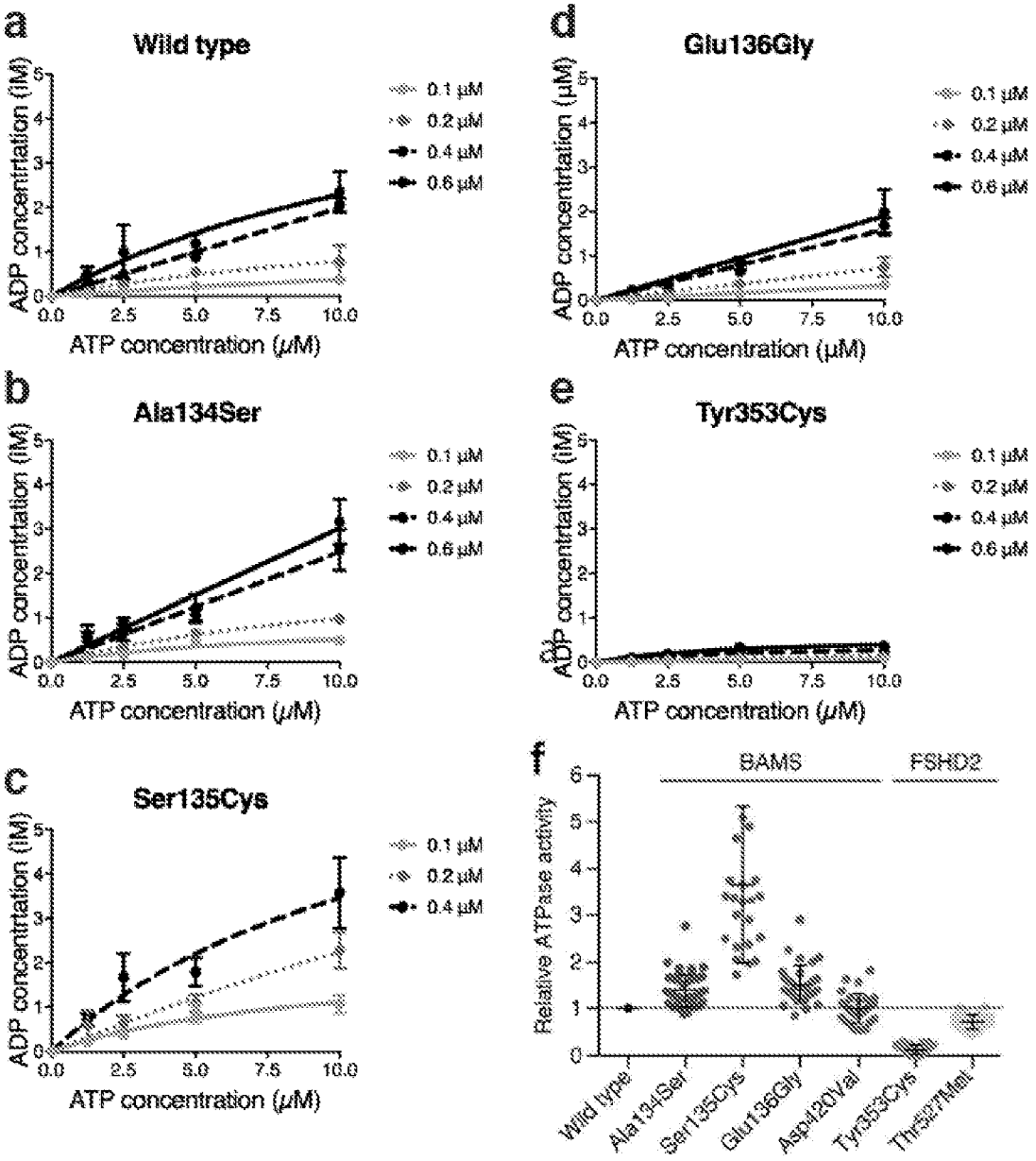

a
Wild type
ADP concentration (μM)
ATP concentration (μM)
0.1 μM
0.2 μM
0.4 μM
0.6 μM
d
Glu136Gly
b
Ala134Ser
e
Tyr353Cys
c
Ser135Cys
f
BAMS
FSHD2
Relative ATPase activity
Wild type
Ala134Ser
Ser135Cys
Glu136Gly
Asp420Val
Tyr353Cys
Thr527Met

Figure 2

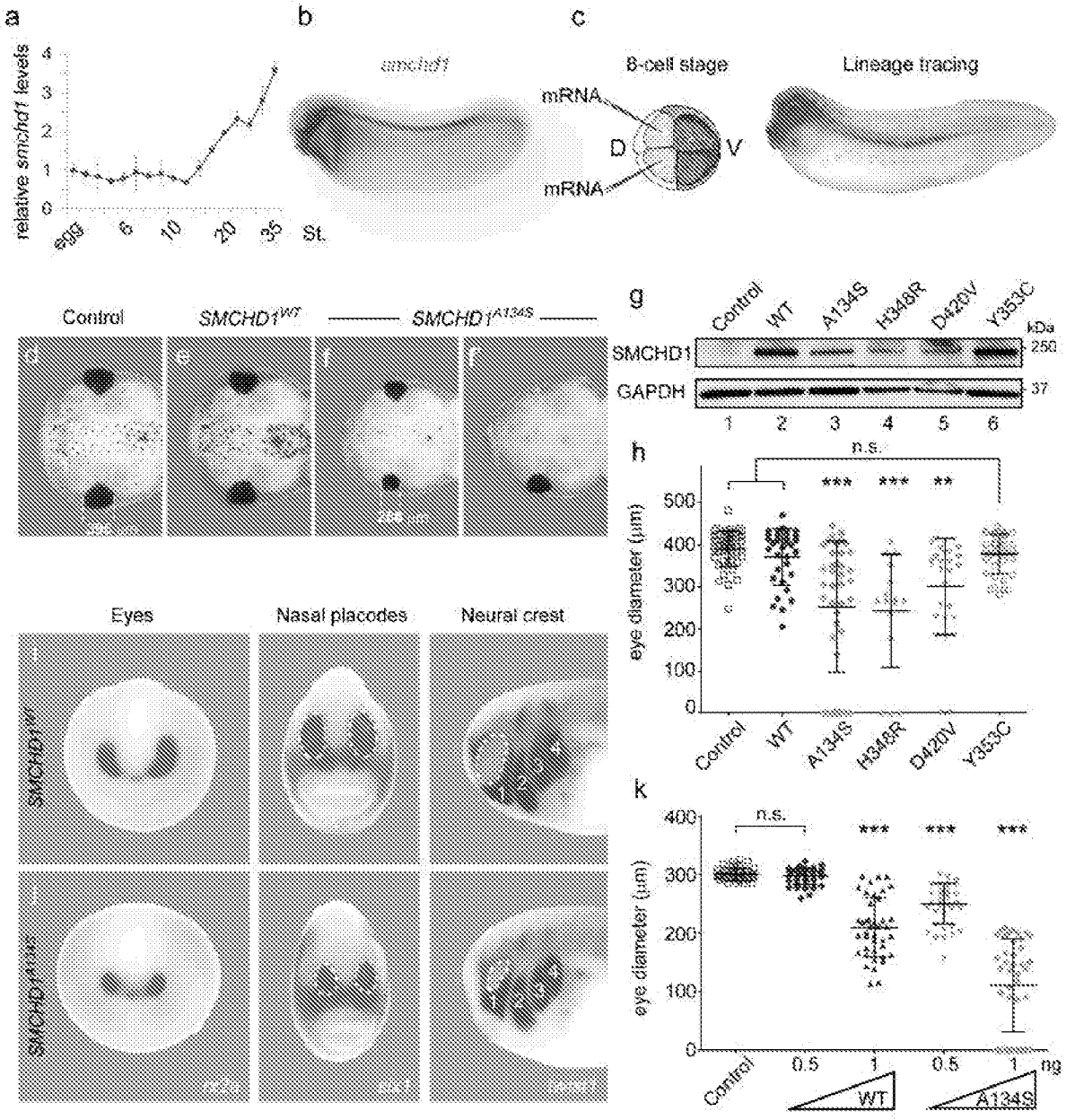

a
b
c
relative smchd1 levels
egg
6
10
20
35
St.
smchd1
8-cell stage
mRNA
D
V
mRNA
Lineage tracing
Control
SMCHD1WT
SMCHD1A134S
Eyes
Nasal placodes
Neural crest
SMCHD1WT
SMCHD1A134S
g
Control
WT
A134S
H348R
D420V
Y353C
kDa
SMCHD1
250
GAPDH
37
1 2 3 4 5 6
h
n.s.
***
***
**
eye diameter (μm)
500
400
300
200
100
0
Control
WT
A134S
H348R
D420V
Y353C
k
n.s.
***
***
***
eye diameter (μm)
400
300
200
100
0
Control
0.5
1
0.5
1
ng
WT
A134S

Figure 3

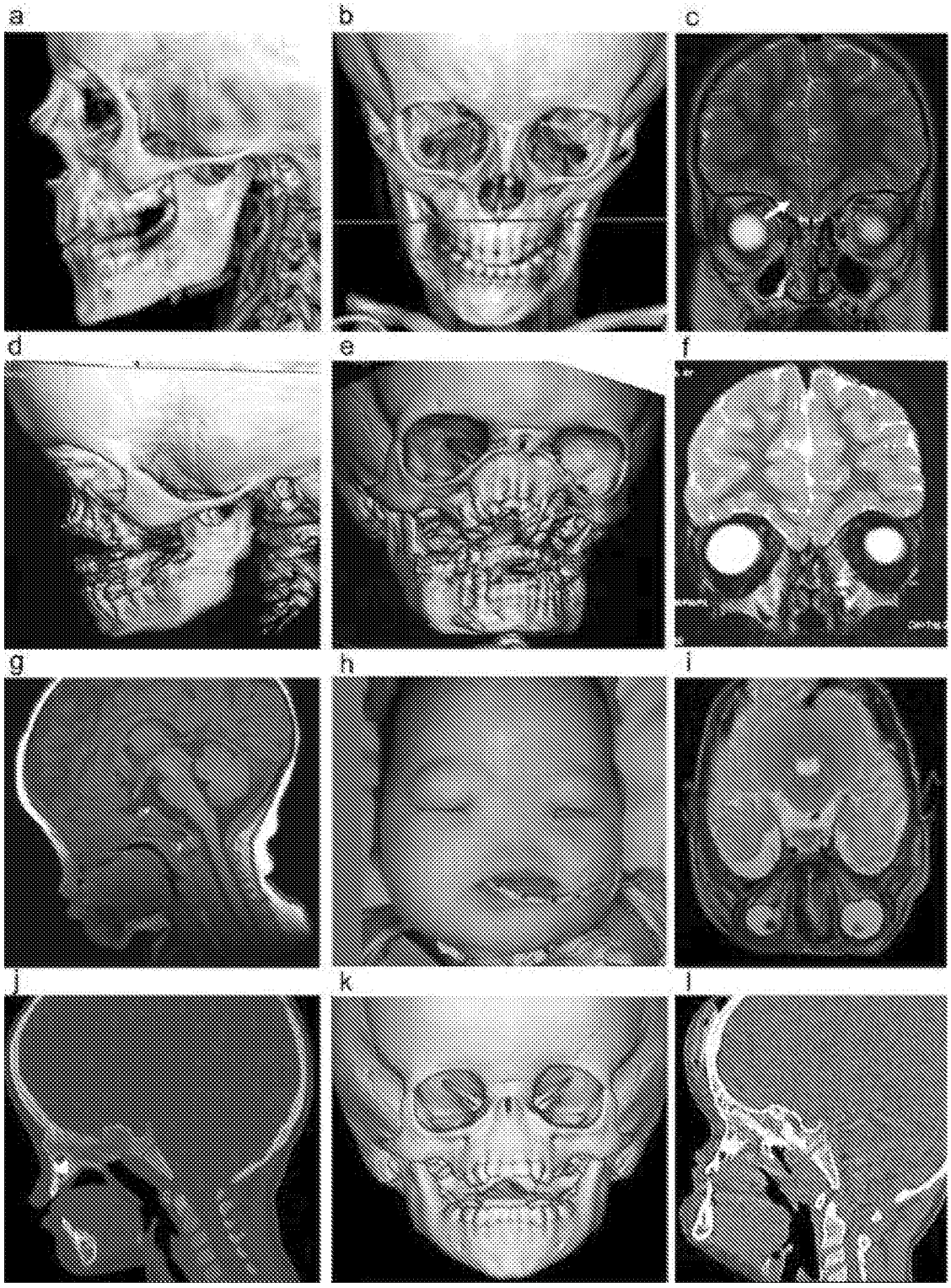


Figure 4

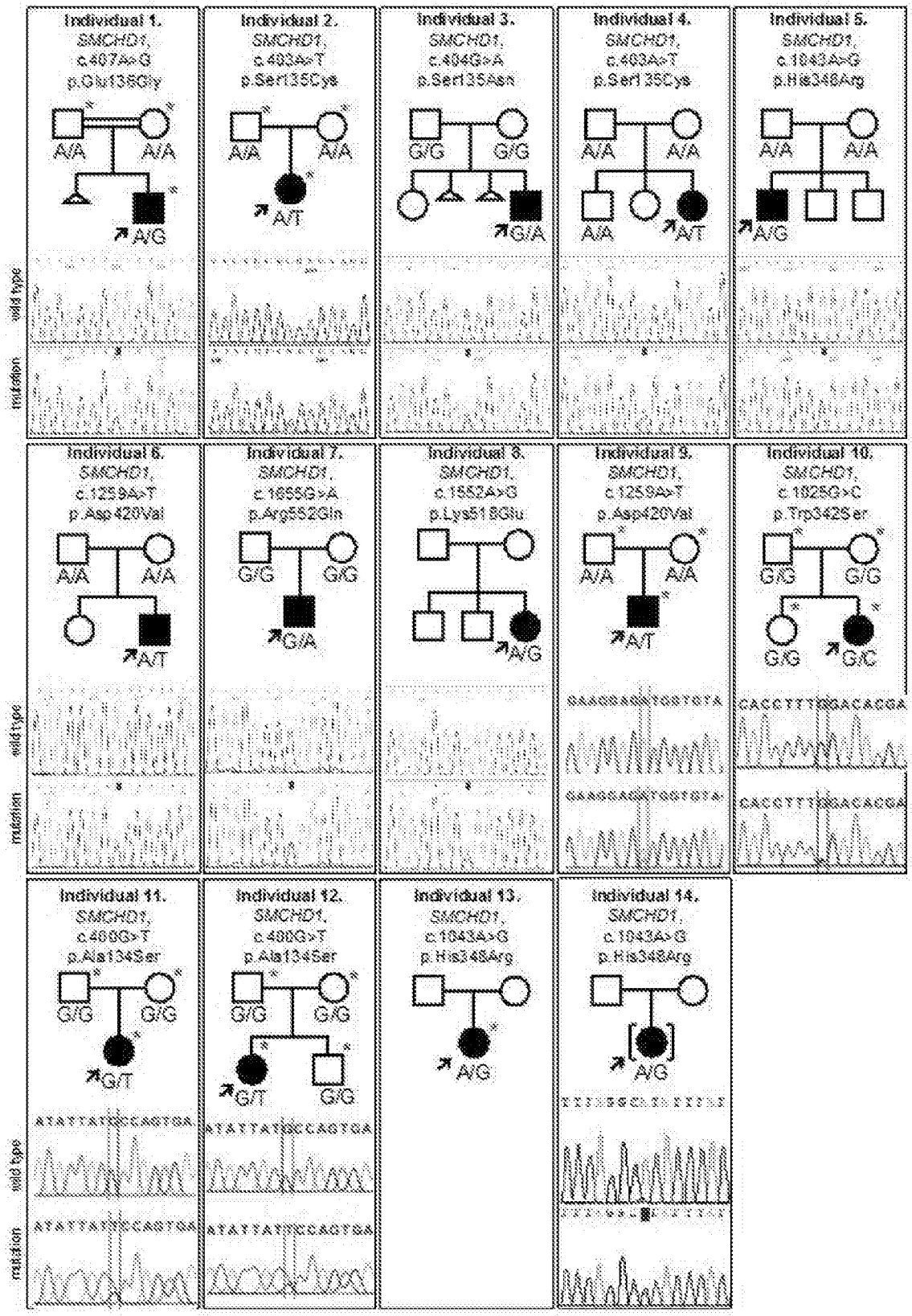

Individual 1.
SMCHD1,
c.407A>G
p.Glu136Gly
A/A
A/A
A/G
Individual 2.
SMCHD1,
c.403A>T
p.Ser135Cys
A/A
A/A
A/T
Individual 3.
SMCHD1,
c.404G>A
p.Ser135Asn
G/G
G/G
G/A
Individual 4.
SMCHD1,
c.403A>T
p.Ser135Cys
A/A
A/A
A/T
Individual 5.
SMCHD1,
c.1043A>G
p.His348Arg
A/A
A/A
A/G
wild type
mutation
Individual 6.
SMCHD1,
c.1259A>T
p.Asp420Val
A/A
A/A
A/T
Individual 7.
SMCHD1,
c.1655G>A
p.Arg552Gln
G/G
G/G
G/A
Individual 8.
SMCHD1,
c.1552A>G
p.Lys518Glu
A/G
Individual 9.
SMCHD1,
c.1259A>T
p.Asp420Val
A/A
A/A
A/T
Individual 10.
SMCHD1,
c.1025G>C
p.Trp342Ser
G/G
G/G
G/G
G/C
wild type
mutation
Individual 11.
SMCHD1,
c.400G>T
p.Ala134Ser
G/G
G/G
G/T
Individual 12.
SMCHD1,
c.400G>T
p.Ala134Ser
G/G
G/G
G/T
G/G
Individual 13.
SMCHD1,
c.1043A>G
p.His348Arg
A/G
Individual 14.
SMCHD1,
c.1043A>G
p.His348Arg
A/G
wild type
mutation

Figure 5

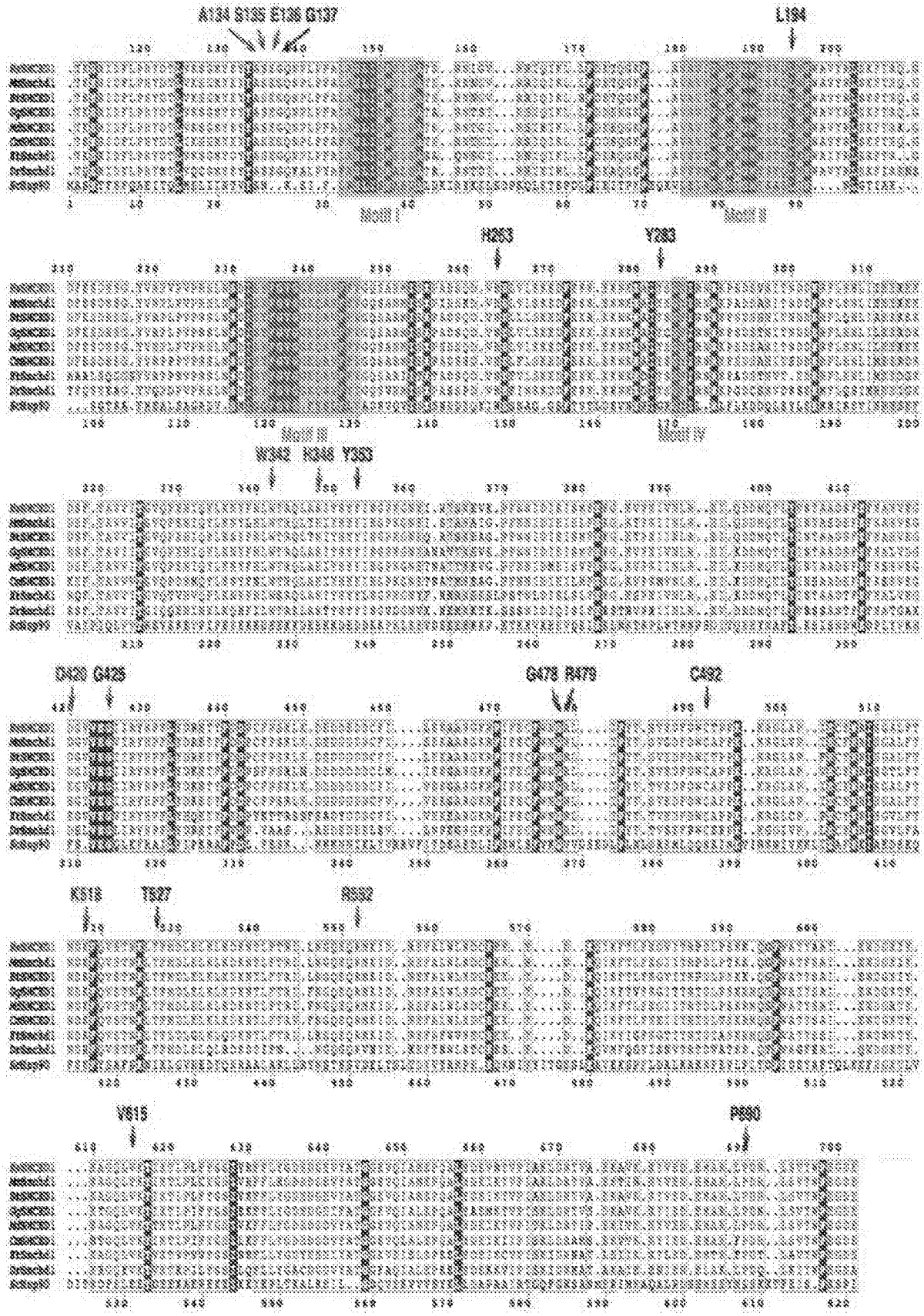

A134 S135 E136 G137
L194
Motif I
Motif II
H253
Y283
Motif III
Motif IV
W342 N346 Y353
D420 G425
G478 R479
C492
K518
T527
R532
V615
P680

Figure 6

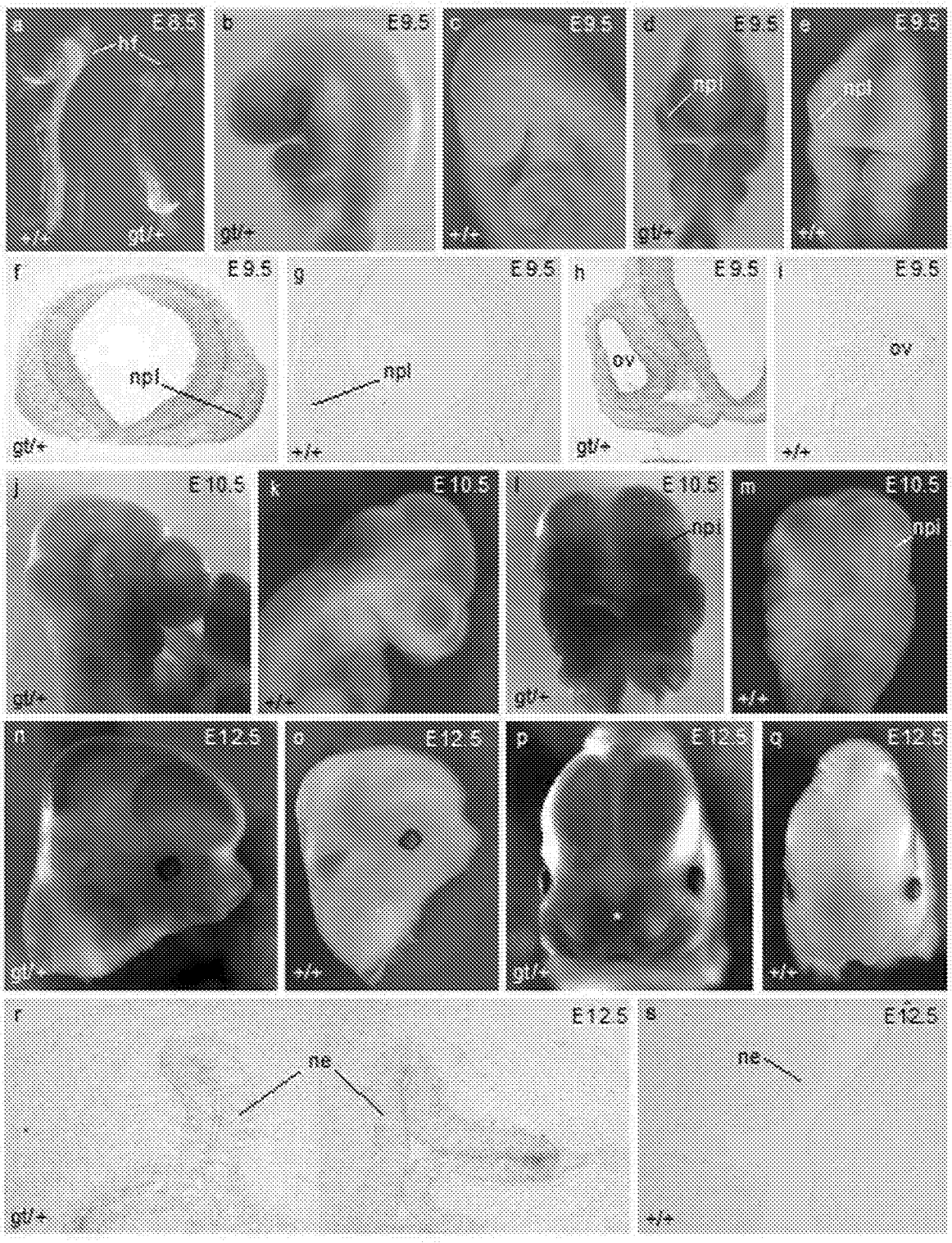

a E9.5
nf
gt/+
+/+
b E9.5
gt/+
c E9.5
+/+
d E9.5
npl
gt/+
e E9.5
npl
+/+
f E9.5
npl
gt/+
g E9.5
npl
+/+
h E9.5
ov
gt/+
i E9.5
ov
+/+
j E10.5
gt/+
k E10.5
+/+
l E10.5
npl
gt/+
m E10.5
npl
+/+
n E12.5
gt/+
o E12.5
+/+
p E12.5
gt/+
q E12.5
+/+
r E12.5
ne
gt/+
s E12.5
ne
+/+

Figure 7

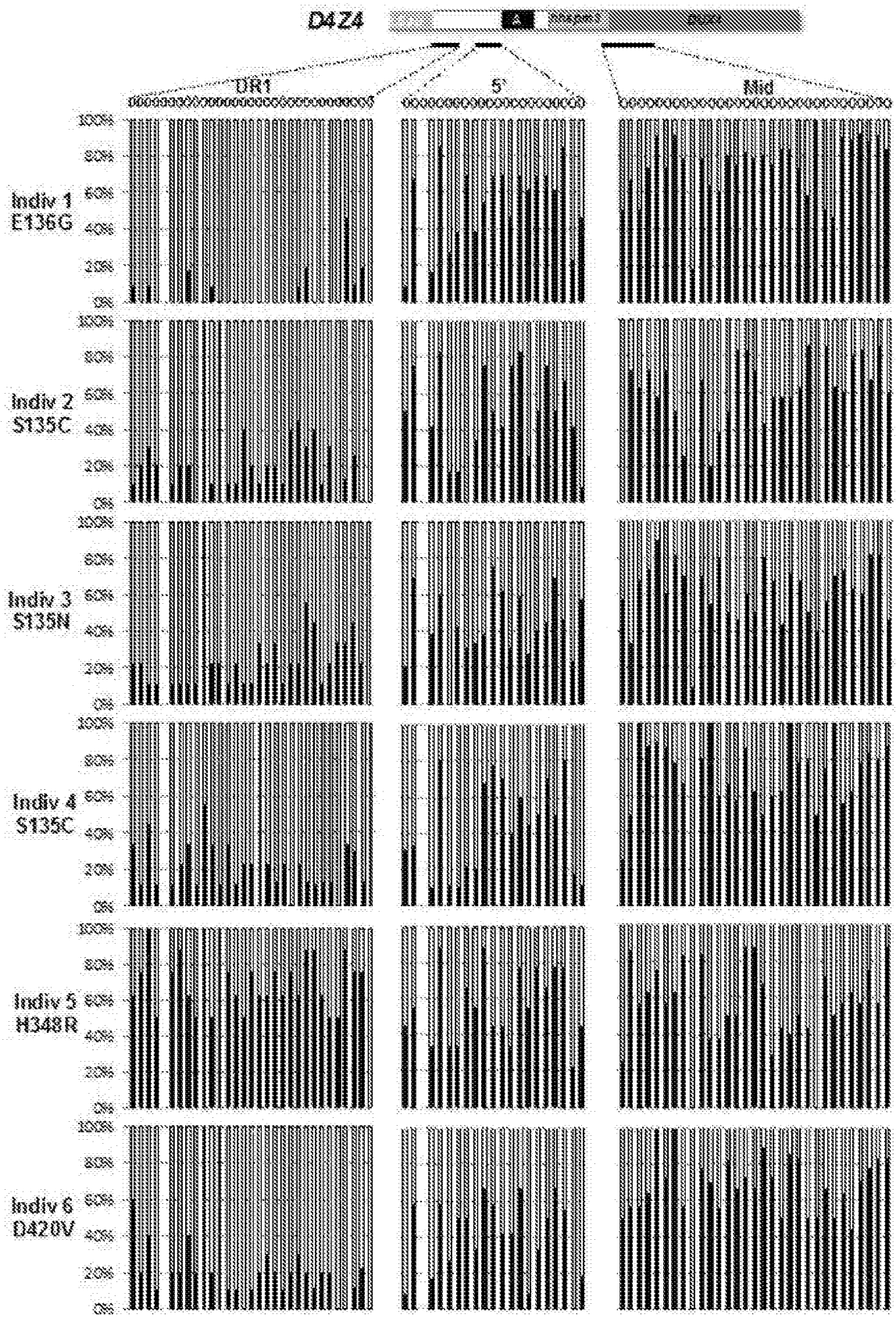


Figure 8 (continued)

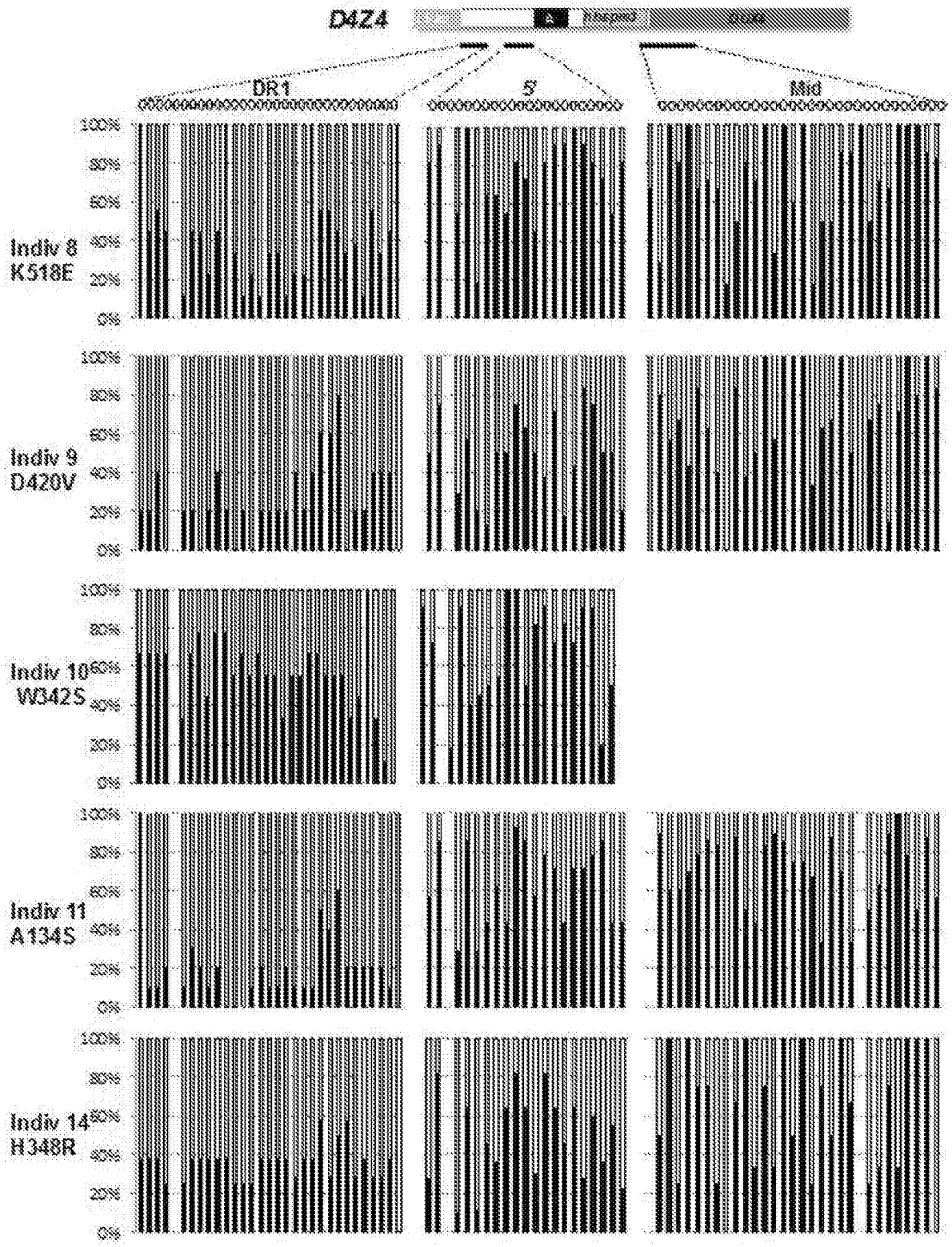

D4Z4
DR1
5'
Mid
100%
80%
60%
40%
20%
0%
Indiv 8
K518E
Indiv 9
D420V
Indiv 10
W342S
Indiv 11
A134S
Indiv 14
H348R

Figure 8

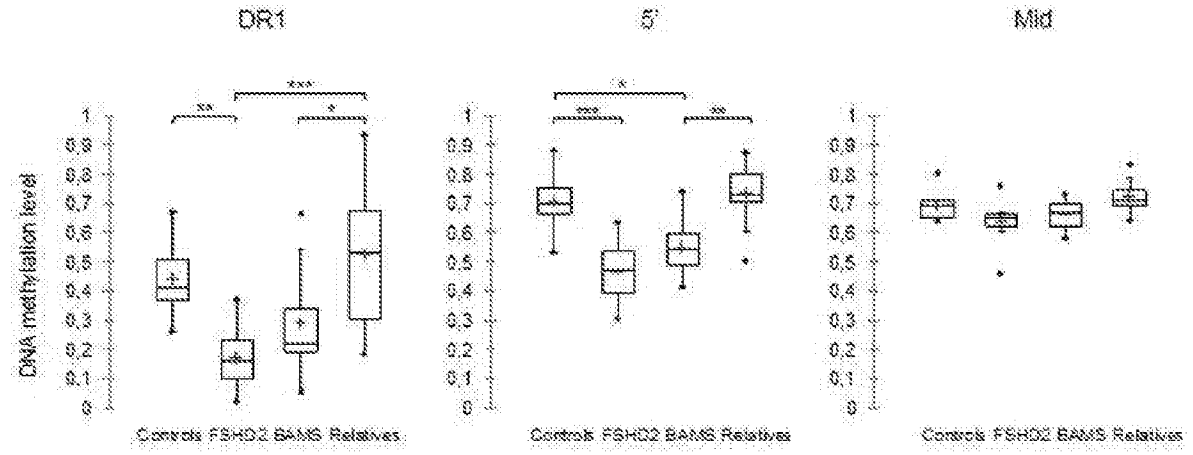


Figure 9

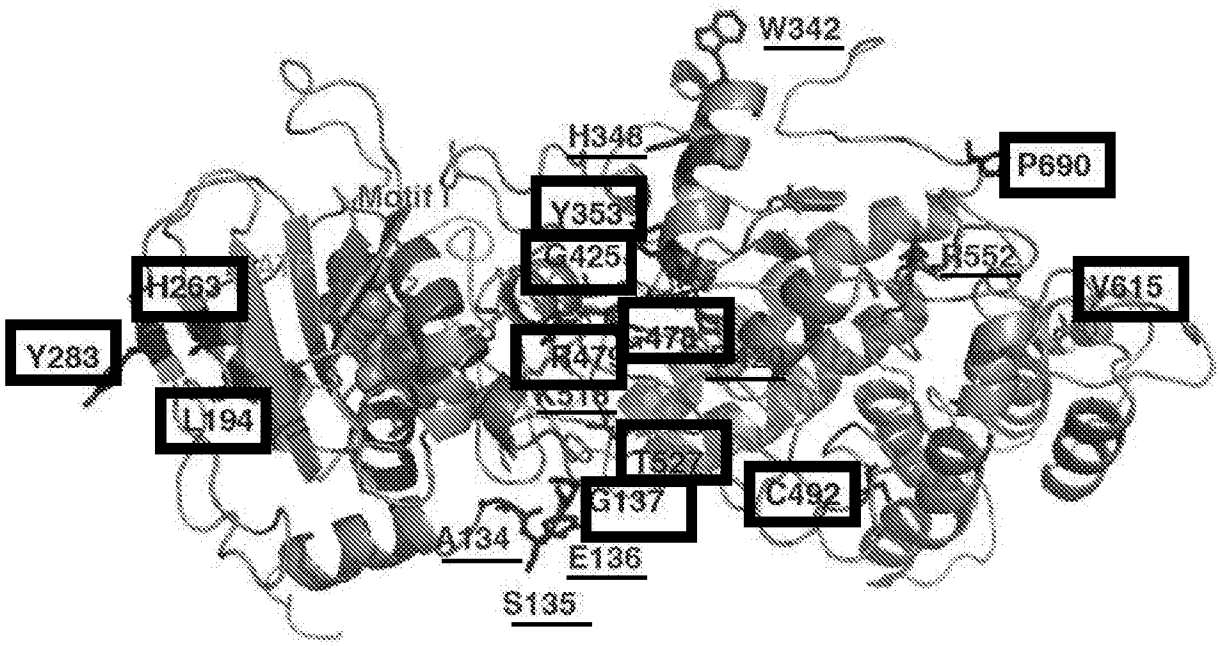


Figure 10

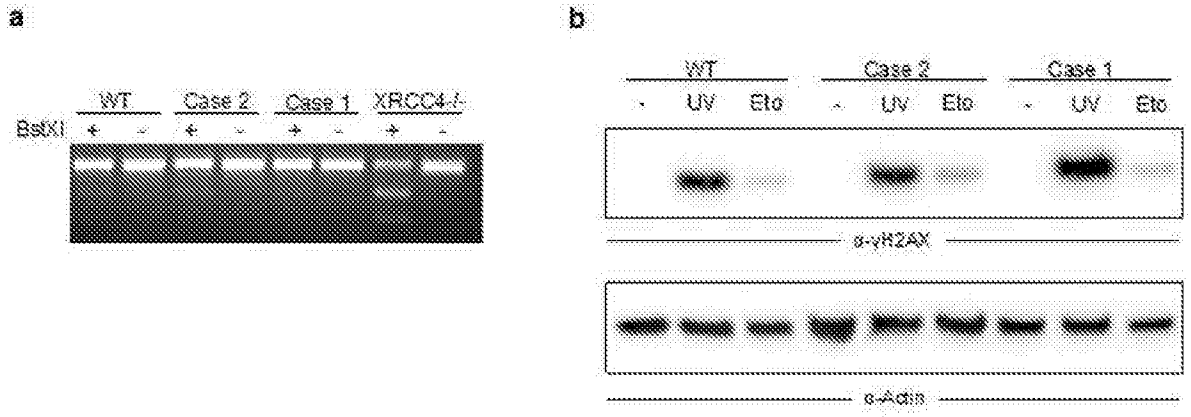


Figure 11

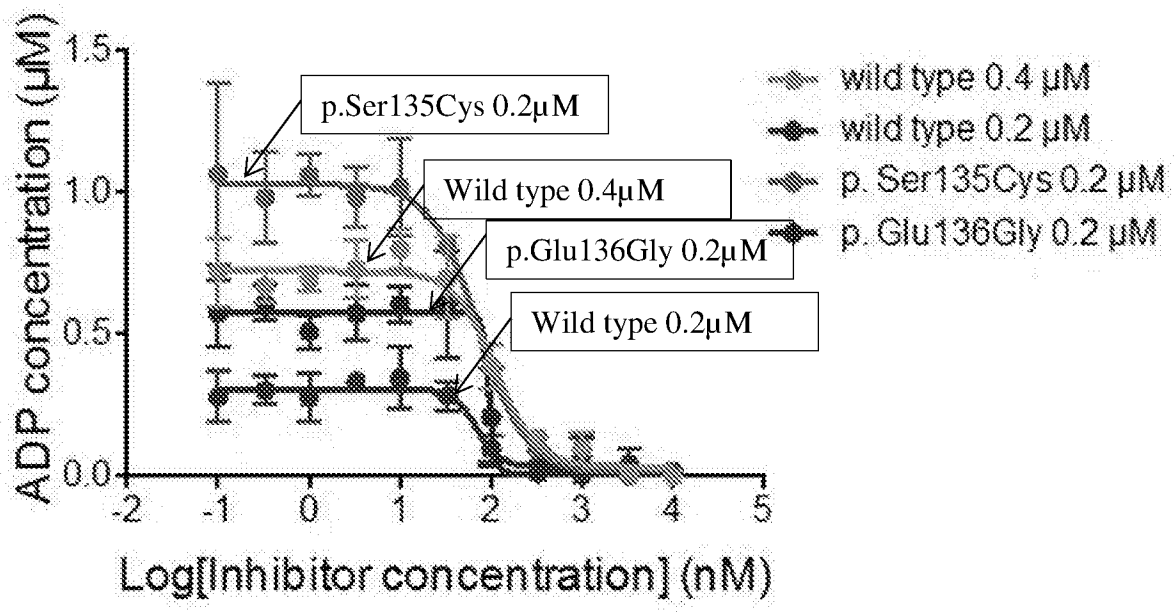


Figure 12

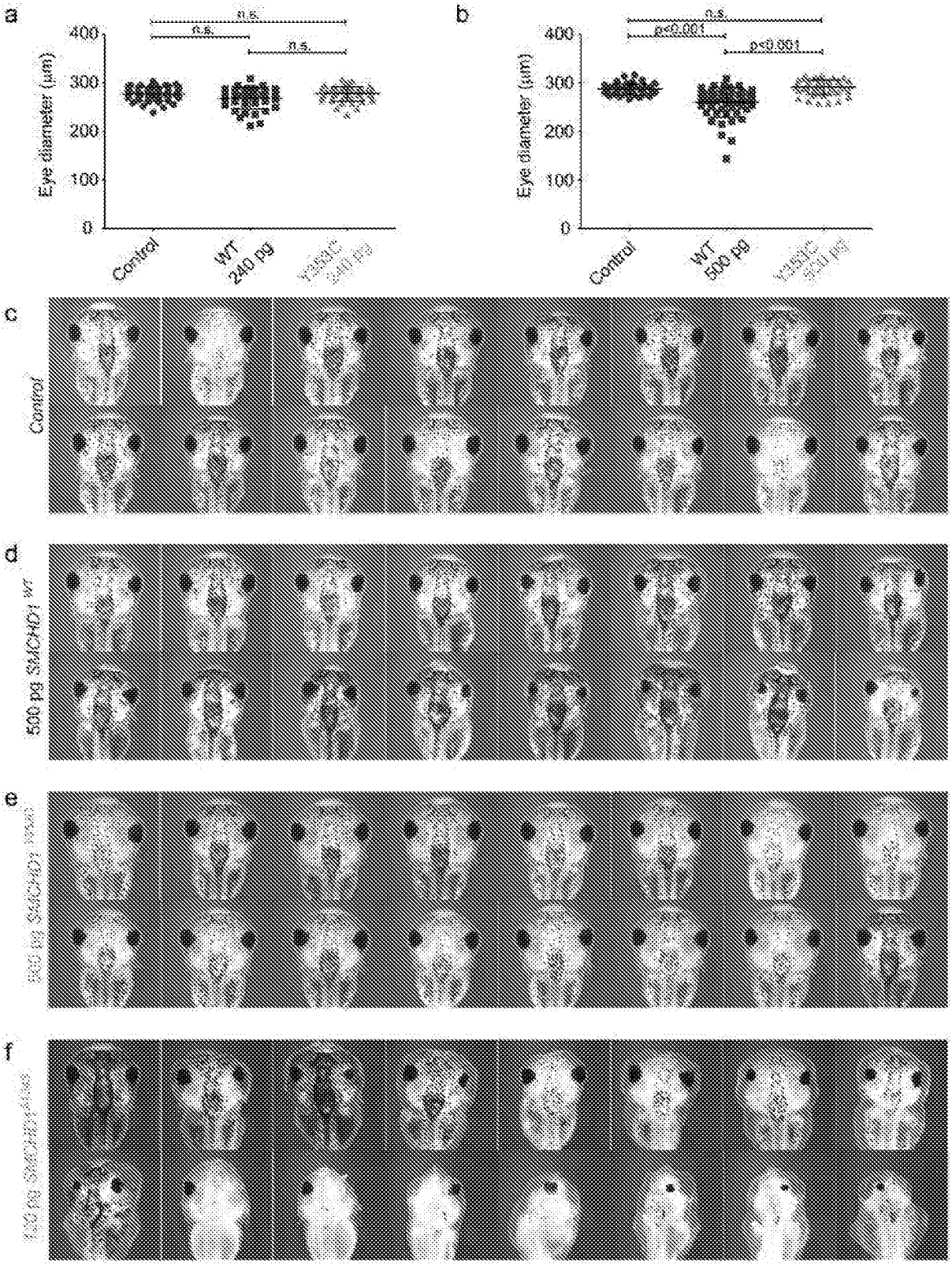

a
b
Eye diameter (µm)
400
300
200
100
0
n.s.
n.s.
n.s.
p<0.001
p<0.001
Control
WT
240 pg
WT
500 pg
c
Control
d
500 pg SMCHD1 WT
e
500 pg SMCHD1 mut
f
500 pg SMCHD1 del36

Figure 13

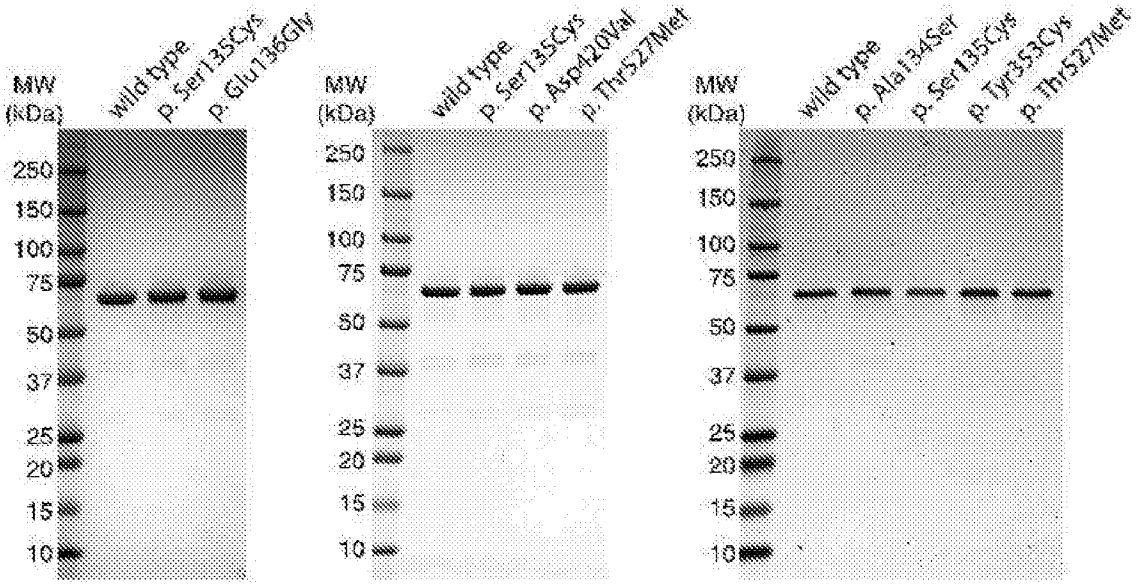

MW (kDa)
Wild type
P. Ser135Cys
P. Glu136Gly
250
150
100
75
50
37
25
20
15
10
MW (kDa)
Wild type
P. Ser135Cys
P. Asp420Val
P. Thr527Met
250
150
100
75
50
37
25
20
15
10
MW (kDa)
Wild type
P. Ala134Ser
P. Ser135Cys
P. Tyr353Cys
P. Thr527Met
250
150
100
75
50
37
25
20
15
10

Figure 14

INTERNATIONAL SEARCH REPORT

International application No.
PCT/SG2018/050008

A. CLASSIFICATION OF SUBJECT MATTER

C07K 14/47 (2006.01) ***A61K 38/17 (2006.01)*** ***C12N 15/00 (2006.01)*** ***A61P 21/00 (2006.01)***

According to International Patent Classification (IPC)

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

C07K A61K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EPODOC/WPI/CAPLUS/EMBASE/BIOSIS/MEDLINE: SMCHD1, structural maintenance of chromosomes flexible hinge domain-containing protein 1, SMC hinge domain-containing protein 1, KIAA0650 and similar terms.

REGISTRY: SEQ ID NO.: 1

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X A	WO 2016/195493 A1 (ACADEMISCH ZIEKENHUIS LEIDEN H.O.D.N. LUMC) 8 December 2016 Whole document, particularly page 5, line 5 to page 20.	17 & 19 1-16, 18 & 20-22
X A	WO 2013/192185 A1 (FRED HUTCHINSON CANCER RESEARCH CENTER ET AL.) 27 December 2013 Whole document, particularly, paragraphs [0028], [00115]-[00145].	17 & 19 1-16, 18 & 20-22
A	BLEWITT, M.E. ET AL., SmcHD1, containing a structural-maintenance-of-chromosomes hinge domain, has a critical role in X inactivation. *Nat Genet*, 20 April 2008, Vol. 40, No. 5, pages 663-669 [Retrieved on 2018-02-28] <DOI: 10.1038/NG.142> Abstract.	1-16, 18 & 20-22

☒ Further documents are listed in the continuation of Box C. ☒ See patent family annex.

*Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
28/02/2018 *(day/month/year)*	23/03/2018 *(day/month/year)*

Name and mailing address of the ISA/SG

IPOS INTELLECTUAL PROPERTY OFFICE OF SINGAPORE

Intellectual Property Office of Singapore
51 Bras Basah Road
#01-01 Manulife Centre
Singapore 189554

Email: pct@ipos.gov.sg

Authorized officer

Teclise Ng (Dr)

IPOS Customer Service Tel. No.: (+65) 6339 8616

INTERNATIONAL SEARCH REPORT

International application No.
PCT/SG2018/050008

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WINSTON, J. ET AL., Identification of two novel *SMCHD1* sequence variants in families with FSHD-like muscular dystrophy. *Eur J Hum Genet*, 23 April 2014, Vol. 23, No. 1, pages 67-71 [Retrieved on 2018-02-28] <DOI: 10.1038/EJHG.2014.58> Whole document.	1-16, 18 & 20-22
A	CHEN, K. ET AL., The epigenetic regulator Smchd1 contains a functional GHKL-type ATPase domain. *Biochem J*, 8 April 2016, Vol. 473, No. 12, pages 1733-1744 [Retrieved on 2018-02-28] <DOI: 10.1042/BCJ20160189> Abstract.	1-16, 18 & 20-22
A	SACCONI, S. ET AL., The FSHD2 gene *SMCHD1* is a modifier of disease severity in families affected by FSHD1. *Am J Hum Genet*, 26 September 2013, Vol. 93, No. 4, pages 744-751 [Retrieved on 2018-02-28] <DOI: 10.1016/J.AJHG.2013.08.004> Whole document.	1-16, 18 & 20-22
A	DAXINGER, L. ET AL., Genetic and epigenetic contributors to FSHD. *Curr Opin Genet Dev*, August 2015, Vol. 33, pages 56-61 [Retrieved on 2018-02-28] <DOI: 10.1016/J.GDE.2015.08.007> Whole document.	1-16, 18 & 20-22
P,X	GORDON, C.T. ET AL., *De novo* mutations in *SMCHD1* cause Bosma arhinia microphthalmia syndrome and abrogate nasal development. *Nat Genet*, 9 January 2017, Vol. 49, No. 2, pages 249-255 [Retrieved on 2018-02-28] <DOI: 10.1038/NG.3765> Whole document.	1-22

INTERNATIONAL SEARCH REPORT

International application No.
PCT/SG2018/050008

Box No. I Nucleotide and/or amino acid sequence(s) (Continuation of item 1.c of the first sheet)

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of a sequence listing:

   a. ☒ forming part of the international application as filed:

      ☒ in the form of an Annex C/ST.25 text file.

      ☐ on paper or in the form of an image file.

   b. ☐ furnished together with the international application under PCT Rule 13*ter*.1(a) for the purposes of international search only in the form of an Annex C/ST.25 text file.

   c. ☐ furnished subsequent to the international filing date for the purposes of international search only:

      ☐ in the form of an Annex C/ST.25 text file (Rule 13*ter*.1(a)).

      ☐ on paper or in the form of an image file (Rule 13*ter*.1(b) and Administrative Instructions, Section 713).

2. ☐ In addition, in the case that more than one version or copy of a sequence listing has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that in the application as filed or does not go beyond the application as filed, as appropriate, were furnished.

3. Additional comments:

   Since only one version or copy of a sequence listing has been filed or furnished, the statements under item 2 are not required.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.

PCT/SG2018/050008

Note: This Annex lists known patent family members relating to the patent documents cited in this International Search Report. This Authority is in no way liable for these particulars which are merely given for the purpose of information.

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2016/195493 A1	08/12/2016	NONE	
WO 2013/192185 A1	27/12/2013	US 2015/0159216 A1 EP 2861738 A1	11/06/2015 22/04/2015