Mutations in *EXTL3* cause neuro-immuno-skeletal dysplasia syndrome

Running title: *EXTL3* is mutated in neuro-immuno-skeletal dysplasia

Machteld M. Oud,[1,26]* Paul Tuijnenburg,[2,3,26] Maja Hempel,[4,26] Naomi van Vlies,[5] Zemin Ren,[6] Sacha Ferdinandusse,[5] Machiel H. Jansen,[2,3] René Santer,[7] Jessika Johannsen,[7] Chiara Bacchelli,[8] Marielle Alders,[9] Rui Li,[10,11] Rosalind Davies,[8] Lucie Dupuis,[12] Catherine M. Cale,[13] Ronald J.A. Wanders,[5] Steven T. Pals,[6] Louise Ocaka,[8] Chela James,[8] Ingo Müller,[14] Kai Lehmberg,[14] Tim Strom,[15,16] Hartmut Engels,[17] Hywel J. Williams,[8] Phil Beales,[8] Ronald Roepman,[1] Patricia Dias,[18] Han G. Brunner,[1] Jan-Maarten Cobben,[19,20] Christine Hall,[21] Taila Hartley,[22] Polona Le Quesne Stabej,[8] Roberto Mendoza-Londono,[12] E. Graham Davies,[8,13] Sérgio B. de Sousa,[8,23,24] Davor Lessel,[4,27] Heleen H. Arts,[1,25,27] Taco W. Kuijpers,[2,27]**

[1]Department of Human Genetics, Radboud Institute for Molecular Life Sciences, Radboud University Medical Centre, PO Box 9101, 6500 HB Nijmegen, The Netherlands; [2]Department of Experimental Immunology, Academic Medical Centre, PO Box 22660, 1100 DD Amsterdam, The Netherlands; [3]Department of Pediatric Hematology, Immunology, and Infectious disease, Academic Medical Centre, PO Box 22660, 1100 DD Amsterdam, The Netherlands; [4]Institute of Human Genetics, University Medical Center Hamburg-Eppendorf, 20246 Hamburg, Germany; [5]Laboratory Genetic Metabolic Diseases, Academic Medical Centre, PO Box 22660, 1100 DD Amsterdam, The Netherlands; [6]Department of Pathology, Academic Medical Center, University of Amsterdam, PO Box 22660, 1100 DD Amsterdam, The Netherlands; [7]Department of Pediatrics, University Medical Center Hamburg-Eppendorf, 20246 Hamburg, Germany; [8]COSGene, UCL Great Ormond Street Institute of Child Health, WC1N 1EH London, UK; [9]Department of Clinical Genetics, Academic Medical Center, University of Amsterdam, PO Box 22660, 1100 DD Amsterdam, The Netherlands; [10]Department of Human Genetics, McGill University, H3A 1B1 Montreal, Québec, Canada; [11]McGill University and Génome Québec Innovation Centre, H3A 0G1 Montreal, Québec, Canada; [12]Division of Clinical and Metabolic

Genetics, Department of Peadiatrics, The Hospital for Sick Children and University of Toronto, M5G 1X8 Toronto, ON, Canada; [13]Department of Immunology, Great Ormond Street Hospital, WC1N 3JH London, UK; [14]Division of Pediatric Stem Cell Transplantation and Immunology, University Medical Center Hamburg-Eppendorf, 20246 Hamburg, Germany; [15]Institute of Human Genetics, Helmholtz Zentrum München, 85764 Neuherberg, Germany; [16]Institute of Human Genetics, Technische Universität München, 81675 München, Germany; [17]Institute of Human Genetics, University of Bonn, 53127 Bonn, Germany; [18]Serviço de Genética, Departamento de Pediatria, Hospital de Santa Maria, Centro Hospitalar Lisboa Norte, Centro Académico de Medicina de Lisboa, 1640-035 Lisboa, Portugal; [19]Department of Pediatrics, Academic Medical Center University Hospital, PO Box 22660, 1100 DD Amsterdam, The Netherlands; [20]Department of Clinical Genetics, St George's University Hospital, PO Box SW19 0ER London, UK; [21]Emerita, Department of Radiology, Great Ormond Street Hospital, WC1N 3JH London, UK; [22]Children's Hospital of Eastern Ontario Research Institute, K1H 8 L1 Ottawa, ON, Canada; [23]Medical Genetics Unit, Hospital Pediátrico, Centro Hospitalar e Universitário de Coimbra, 3000-602 Coimbra, Portugal; [24]Faculty of Health Sciences, University of Beira Interior, 6200-506 Covilhã, Portugal; [25]Department of Pathology and Molecular Medicine, McMaster University Medical Centre, L8S 4J9, Hamilton, ON, Canada; [26]First authors contributed equally to this work; [27]Senior authors contributed equally to this work; *Correspondence: machteld.oud@radboudumc.nl; **Correspondence: t.w.kuijpers@amc.uva.nl

Abstract

EXTL3 regulates the biosynthesis of heparan sulfate (HS), important for both skeletal development and hematopoiesis by the formation of HS proteoglycans (HSPGs). By whole-exome sequencing, we identified here homozygous missense mutations in *EXTL3*, c.1382C>T, c.1537C>T, c.1970A>G, and c.2008T>G, in nine affected individuals from five unrelated families. Notably, we found the identical homozygous missense mutation c.1382C>T (p.Pro461Leu) in four affected individuals from two unrelated families. Affected individuals presented with variable skeletal abnormalities and neurodevelopmental defects. Severe Combined Immunodeficiency (SCID) with a complete absence of T cells was observed in three families. EXTL3 shows highest protein levels in hematopoietic stem cells and early progenitor T cells, which is in line with a SCID phenotype at the level of early T cell development in the thymus. To provide further support for the hypothesis that mutations in *EXTL3* cause a neuro-immuno-skeletal dysplasia syndrome, and to gain insight into the pathogenesis of the disorder, we analyzed the localization of EXTL3 in fibroblasts derived from affected individuals and determined glycosaminoglycan concentrations in these cells as well as in urine and blood from the affected individuals. We observed abnormal glycosaminoglycan concentrations, an increased concentration of the non-sulfated chondroitin disaccharide D0a0 and of the disaccharide D0a4 in serum and urine of all analyzed affected individuals. In summary, we show that biallelic mutations in *EXTL3* disturb glycosaminoglycan synthesis, leading to a recognizable syndrome characterized by a variable expression of skeletal, neurological and immunological abnormalities.

Introduction

A number of complex immunodeficiency syndromes additionally include skeletal dysplasia as part of their variable phenotype. For example, Schimke immuno-osseous dysplasia [MIM 242900] (SIOD) and cartilage-hair hypoplasia [MIM 250250] (CHH) are (spondylo-epiphyseal-) metaphyseal dysplasias that can also result in disproportionate short stature.[1,2,5] Besides skeletal dysplasia, CHH is further characterized by T cell immunodeficiency, brittle hair, ligamentous laxity, hypoplastic anemia, and sometimes neuronal dysplasia of the intestine leading to the Hirschsprung's disease.[6] SIOD is commonly associated with skeletal dysplasia, nephropathy and T cell immunodeficiency.[7] Notably, both CHH and SIOD are recessive disorders caused by mutations in *RMRP* [MIM 157660] and *SMARCAL1* [MIM 606622], respectively[6,8] and affected individuals display both broad interfamilial and intrafamilial variability of the immunological abnormalities. [6,7]

Here we report on a neuro-immuno-skeletal disorder caused by pathogenic mutations in *EXTL3* [MIM 605744; exostosin-like glycosyltransferase 3], a gene not previously associated with a human disease. EXTL3 is a member of the exostosin (EXT) family of glycosyltransferases comprising EXT1, EXT2, EXTL1, and EXTL2. These enzymes regulate glycosylation, a process by which glycans are attached to both proteins and lipids in the endoplasmic reticulum (ER) or Golgi complex. EXTL3 and its family members are known to be involved in the biosynthesis of the glycosaminoglycan (GAG) heparan sulfate (HS), in a variety of species.[11-16] EXT family members exert an effect on many physiological activities by the covalent binding of HS chains to proteoglycans forming heparan sulfate proteoglycans (HSPGs). HSPGs are a major component of the extracellular matrix (ECM) in all organs in the human body and are involved in numerous physiological processes.[17] Notably, there are three subfamilies of HSPGs: membrane-spanning proteoglycans, glycophosphatidylinositol (GPI)-anchored proteoglycans, and secreted ECM proteoglycans,[17] all of which have been implicated in skeletogenesis and hematopoiesis.[18,19]

EXTL3 is a *N*-acetylglucosaminyltransferase (GlcNac transferase) that exhibits GlcNac-TI activity by catalyzing the GlcNac binding to a GAG protein linkage region, thereby initiating the biosynthesis of HS chains. EXTL2 shares the GlcNac-TI activity with EXTL3 but initiates HS chains on different core proteins than EXTL3, by discriminating the amino acid sequences.[14] EXTL1 is involved in the elongation of HS chains by exhibiting GlcNAcT-II activity,[14] whereas the exostosin family members EXT1 and EXT2 form a heterooligomeric complex with GlcNAc and GlcA transferase activity essential for the polymerization of HS. The crucial role of EXTL3 in general organogenesis is exemplified by embryonic lethality of mice homozygous for a null mutation.[16] Regarding human disorder, *EXT1* and *EXT2* are associated with autosomal dominant hereditary multiple exostoses (HME) [MIM 133700 and 133701].[20] In addition, autosomal recessive mutations in *EXT2* lead to seizures-scoliosis-macrocephaly syndrome [MIM 616682].[21] *EXTL3* mutations have not yet been connected to any disease. In this report, we describe nine individuals from five unrelated families with an autosomal-recessive neuro-immuno-skeletal dysplasia syndrome caused by biallelic missense mutations in *EXTL3*.

Material and methods

Collection of samples and informed consent

For all five families, DNA samples were extracted from blood according to standard laboratory practice and used for whole-exome sequencing (WES). Fibroblast cell lines, obtained from skin biopsy material of the affected individuals A:II-1, B:II-1, B:II-2, C:II-1, D:III-1, D:IV-1, and E:II-1 and three in-house control fibroblast cell lines were used for GAG analysis and immunocytochemistry. Heparinized peripheral blood samples and urine were collected from healthy age-matched control individuals as well as from affected individuals during routine diagnostic procedures. Urine collected from affected individuals A:II-1, B:II-1, and B:II-2, as well as, serum collected from the blood of these individuals and parents B:I-1 and B:I-2 was used for GAG analysis. After identification of the *EXTL3* mutations in family A, other genetic

laboratories were contacted using GeneMatcher and Matchmaker Exchange that linked our EXTL3 submission to PhenomeCentral resulting in the ascertainment of families B-E.[22-24] All participants in this study gave written informed consent and all human material was collected after approval of the local ethic committees; Family A NL40332.078.12, family B PV3802, family C 1000029424, family D 09CM32 and family E 06/Q0508/16.

Whole-exome sequencing

Genomic DNA from the unaffected parents and affected individuals II-1 from family A, II-1 and II-2 from family B, II-1 from family C and her unaffected parents, III-1, III-2, and IV-1 from family D, and II-1 and II-2 and unaffected parents from family E were used for WES. WES experiments were performed in different centers with slightly different procedures that have essentially been described before (family A,[25-28] family B,[29] family C,[30,31] family D and E [32]). Briefly, exome enrichment was performed using a SureSelect Human All Exon kit v4 (family A, D, and E) or v5 (family B and C) 50Mb kit (Agilent, Santa Clara, USA), followed by sequencing on a SOLiD 5500xl System (family A) (Life Technologies/Thermo Fisher Scientific, Waltham, MA, USA) and a HiSeq2500 (family B) or HiSeq2000 (families C, D, and E) system (Illumina, San Diego, CA, USA). Read mapping and single base-pair variant and indel calling were performed using the LifeScope Software v2.1 (Life Technologies) for family A. For family B, C, D and E, reads were aligned to the human genome assembly HG19 (UCSC Genome Browser) with Burrows-Wheeler Aligner (BWA, v 0.5.87.5 for family B and 0.7.7 for family C), and detection of genetic variation was performed using SAMtools (v 0.1.18), PINDEL (v 0.2.4t), and ExomeDepth (v 1.0.0). For family C, indel realignment and base recalibration was performed using GATK 3.1.1. Duplicated reads were removed by Picard 1.108 and genetic variation was detected using SAMtools (v 1.2) and BCFtools (v 1.2).[33] Variant calling for families D and E was performed using GATK base quality score recalibration, indel realignment, duplicate removal, and performed SNP and indel discovery and genotyping using

standard hard filtering parameters or variant quality score. The variant annotation and interpretation analyses were generated through the use of Ingenuity Variant Analysis software (v 3.1.20140902) from Ingenuity Systems/Qiagen.

Copy number variation (CNV) detection was performed on WES data of each affected individual. For family A the CoNIFER computational pipeline[34] was used as described previously.[35] For family C, FishingCNV software[36] was used to call CNVs and compare the read depth of these CNVs against the distribution of reads in more than 50 control samples without related genetic disorders (all samples were sequenced and aligned using the same protocol). In addition, the eXome-Hidden Markov Model (XHMM) algorithm was applied to detect rare CNVs.[37-39] In addition, a SNP-array was performed in affected individuals from family B as previously described.[40]

Bioinformatic filtering

Variant annotation was performed using four independent in-house annotation pipelines. To prioritize variants in the sequencing data, variants that were non-genic, intronic (except for canonical splice sites) or synonymous were excluded. Further, we excluded variants that were present in dbSNPv135 at a frequency >1% or >1% in the in-house variants database (for family A the in-house database contains data from 2096 in-house analyzed exomes, for family C the in-house database contains ~3000 exomes). For family C, variants with a minor allele frequency (MAF) >5% in either 1000 Genomes database[41] (October 2014), NHLBI GO Exome Sequencing Project (ESP) exomes (v 0.0.30, November 3, 2014), or Exome Aggregation Consortium (ExAC) data 0.2 (October 29, 2014) were excluded. Variants excluded in family D and E had a MAF >1% in NHLBI ESP and ExAC. The NHLBI GO ESP database contains data from more than 200,000 individuals. The ExAC database includes sequences of 60,706 unrelated individuals from different ethnic backgrounds at the time of this study. A quality filter

was used to exclude variants with <5 reads or <20% variant reads. Candidate variants were validated by Sanger sequencing and segregation analysis was conducted in all five families.

Immunophenotyping

Peripheral blood mononuclear cells (PBMCs) were isolated using standard density gradient centrifugation techniques using Lymphoprep (Nycomed, Oslo, Norway). For immunophenotyping PBMCs were resuspended in phosphate buffered saline (PBS), containing 0.5% (w/v) bovine serum albumin (BSA) and 0.01% sodium azide and incubated with saturating concentrations of fluorescently labeled conjugated monoclonal antibodies. Analysis of cells was performed using a FACSCanto-II flow cytometer and FlowJo software (Ashland, OR, USA). Samples from the affected individuals were analyzed simultaneously with PBMCs from healthy controls. The following conjugated monoclonal antibodies were used: CD4 PE-Cy7, CD8 PerCP-Cy5.5, CD20 PerCP-Cy5.5, CD27 APC, CD38 PE-Cy7, CD45 APC-H7 and IgD PE from BD Biosciences (San Jose, CA, USA), CD3 Alexa 700 and CD19 Alexa 700 from eBioscience (San Diego, California, CA, USA), CD27 FITC from Sanquin (Amsterdam, The Netherlands), CD45RA (2H4-RD1) PE from Beckman Coulter (Brea, California, CA, USA), IgM FITC and IgG FITC from Dako (Glostrup, Denmark), IgA FITC from Miltenyi Biotec (Bergisch Gladbach, Germany).

In order to analyze the *in vitro* activation of T and B cells, PBMCs were re-suspended in PBS at a concentration of $5-10\times10^6$ cells/ml and labeled with 0.5μM carboxyfluorescein succinimidyl ester (CFSE) (Molecular Probes/Thermo Fisher Scientific) in PBS for 10 minutes at 37°C under constant agitation. Cells were washed and subsequently resuspended in Iscove's Modified Dulbecco's medium (IMDM) supplemented with 10% fetal calf serum (FCS) (BioWhittaker), antibiotics, and 3.57×10^{-4}% (v/v) β-mercaptoethanol (Merck). Labeled PBMCs containing a fixed number of B cells (2×10^4 per well) were cultured in 96-well flat-bottomed plates for 6 days at 37°C and stimulated with saturating amounts of anti-IgM mAb (clone MH15;

Sanquin), anti-CD40 mAb (clone 14G7; Sanquin) and 20ng/ml IL-21 (Invitrogen), or 1μg/ml CpG oligodeoxynucleotide 2006 (Invivogen) and 100U/ml IL-2 (R&D Systems), or anti-CD3 (clone 1xE; Sanquin) and anti-CD28 (clone 15E8; Sanquin). Proliferation of B and T cells was assessed by measuring CFSE dilution by flow cytometry.

The secretion of IgG and IgM by mature B cells was assessed by analyzing supernatants for secreted IgM and IgG with an in-house ELISA assay using polyclonal rabbit anti-human IgM, and IgG reagents and a serum protein calibrator. All reagents were from Dako (Glostrup, Denmark), as described previously[42,43].

EXTL3 antibody validation

A commercial antibody targeting EXTL3 (goat; AF2635; R&D Systems, Minneapolis, MN, USA) was used to detect the EXTL3 level in different cell lines and human tissue. In order to ensure the specificity of the antibody we knocked out *EXTL3* in Human Embryonic Kidney 293T cells (HEK293T) using the CRISPR/Cas9 technique. Three different single guide RNA's (sgRNA); 1. CTTCTCTATAACGTCAGTAC, 2. GGGACTGGCTTTCGGCCTAT, and 3. AGAAGCTCGGGACCGCATCG, were each inserted into a pLentiCRISPR v2 vector (52961; Addgene, Middlesex, UK) as described by Sanjana *et al.*[44] The pLentiCRISP v2 construct containing EXTL3 single guide RNA or the empty vector was transfected into HEK293T cells using Genius DNA transfection reagent (Westburg, Leusden, The Netherlands) according to the manufacturer's instructions. After transfection, the cells were cultured in standard cell culture medium containing 2.5 μg/ml puromycin for three days to select only the transfected cells. Subsequently, cell lysates were prepared and analyzed by western blot to test the specificity of the EXTL3 antibody.

Analysis of EXTL3 in lymphocytes and thymus tissue

Western blotting was used to assess the protein levels of EXTL3 in different lymphocyte cell lines. Protein lysates were made from diffuse large B cell lymphoma cell lines OCI-Ly1, OCI-Ly7, and U2932; myeloma cell lines ANBL6 and RPMI-8226; leukemia derived T cell lines Jurkat, CCRF-CEM, and Molt4; T cell lymphoma cell line SupT1; Burkitt lymphoma cell line Namalwa; lymphoblast cell line Ramos; mobilized peripheral blood (MPB) CD34$^+$ progenitor cells; CD4$^-$/CD8$^-$ cells, CD4$^+$/CD8$^+$ cells, CD4$^-$/CD8$^+$ cells, and CD4$^+$/CD8$^-$ cells sorted from thymocytes. All cell lines were cultured in IMDM containing 10% Fetal Bovine Serum (FBS), 100 units/ml penicillin. Protein lysates were separated by 10% sodium dodecyl sulfate (SDS) polyacrylamide gel electrophoresis and subsequently immunoblotted using a standard procedure. The immunoblots were incubated in PBS containing 5% milk and 0.2% tween with antibodies targeting human EXTL3 (goat; AF2635; R&D Systems, Minneapolis, MN, USA) and β-actin (clone AC-15, mouse, A5441, Sigma Aldrich, St. Louis, MO, USA). Primary antibodies were detected by the following horseradish peroxidase (HRP) –conjugated secondary antibodies: donkey-anti-goat (605-703-125; Rockland Immunochemicals Inc) and rabbit-anti-mouse (P0161; DAKO, Glostrup, Denmark), followed by detection using Amersham ECL Western Blotting Detection Reagent (GE Healthcare, Little Chalfont, UK).

Paraffin-embedded thymus tissue for immunohistochemical investigations was obtained from a neonate deceased due to asphyxia at birth, who had a post-mortem examination unrelated to this study. Thymus tissue sections of 4 μm thickness were treated with Tris ethylenediaminetetraacetic acid (EDTA) pH 9 for 20 minutes at 121ºC for antigen retrieval to ensure optimal staining of EXTL3. Sections were incubated overnight at 4ºC with an antibody targeting human EXTL3 (goat; AF2635; R&D Systems, Minneapolis, MN, USA). Subsequently, the tissues were washed with PBS and incubated with rabbit-anti-goat antibody (6160-01; Southern Biotech) for 30 minutes at room temperature. EXTL3 was detected by poly-HRP-anti-rabbit IgG (DPVR110HRP; Immunologic, Duiven, The Netherlands) and followed by detection using Ultra DAB (Immunologic, Duiven, The Netherlands).

Immunocytochemistry in human fibroblasts

The localization of EXTL3 was assessed in primary fibroblasts derived from A:II-1, B:II-1 and B:II-2 and compared to fibroblasts from two healthy controls. The fibroblasts were cultured in Dulbecco's Modified Eagle's medium (DMEM) with 20% FCS on sterile coverslips for 72 hours, followed by immunostaining. For this purpose, cells were washed in PBS and fixed in 2% paraformaldehyde (PFA) in PBS for 20 minutes. In order to enable antibody binding cells were incubated in 10mM sodium citrate buffer for 30 minutes at 60ºC. Cells were then permeabilized in 1% triton-X-100 in PBS for 5 minutes and blocked for 30 minutes in 2% BSA in PBS. The following primary antibodies were used targeting EXTL3 at 1:100 dilution (goat; AF2635; R&D Systems, Minneapolis, MN, USA) and the Golgi marker targeting GM130 at 1:250 dilution (mouse; 610822; BD Biosciences, San Jose, CA, USA). Primary antibodies were diluted in 2% BSA in PBS and incubated for 1 hour at room temperature. After washing with PBS, cells were incubated with fluor-labeled secondary antibodies for 30 minutes. These include donkey-anti-goat IgG Alexa Fluor 488 (1:400) and donkey-anti-mouse IgG Alexa Fluor 568 (1:400), obtained from Life Technologies (Bleiswijk, The Netherlands). After washing cover glasses in PBS, a drop of DAPI containing Fluoromount-G (Southern Biotech, Birmingham, AL, USA) was added to the cells. Cover glasses were placed on microscopic glass slides. Microscopic analysis was conducted with a Zeiss Axio Imager Z2 fluorescence microscope (Zeiss, Sliedrecht, The Netherlands) equipped with an ApoTome slider. Images were obtained with ZEN 2012 software (Zeiss) and processed with Photoshop CS4 (Adobe Systems, San Jose, CA, USA) and freely available FIJI software.

Heparan sulfate, dermatan sulfate, and chondroitin sulfate analysis

GAG concentrations were determined in serum, urine, and fibroblasts as described previously *(Langereis et al. 2015, Kingma et al. 2014)*.[45,46] In brief, homogenates were made from fibroblasts by disrupting the cells in PBS using a Vibra Cell sonicator (Sonics & Materials Inc.,

Newtown, CT, USA). Protein concentrations were measured in whole cell lysates as described by Lowry *et al.* To obtain GAG disaccharides, 50 µL serum, 50 µL urine or 25 µg of fibroblast protein were enzymatically digested in a mixture containing 100 mM NH_4Ac (pH 7.0), 10 mM $Ca(Ac)_2$, 2 mM DTT, 5 mIU each of heparinase I, II, III and 50 mIU chondroitinase B, or 50 mIU chondroitinase ABC in a final volume of 150 µL. After 2 hours of incubation at 30°C, 15 µL of 150 mM EDTA (pH 7.0) was added along with 125 ng of the internal standard, 4UA-2S-GlcNCOEt-6S (HD009, Iduron), and the reaction was stopped and boiled for 5 minutes to denature the proteins. The reaction mixture was centrifuged at 20,000× *g* for 5 minutes at room temperature. Subsequently, the supernatant was applied to an Amicon Ultra 30K centrifugal filter (Millipore) and centrifuged at 14,000 × *g* for 15 minutes at 25°C. The filtrate was stored at -20°C until analysis. HS, dermatan sulfate (DS), and chondroitin sulfate (CS) levels were

quantified on a Waters Quattro Premier XE (tandem) mass spectrometer (Waters Corporation, Milford, MA, USA) coupled to an Acquity UPLC system (UPLC-MS/MS), using a Thermo Hypercarb HPLC column (100 × 2.1 mm, 5 µm) with a mobile phase consisting of 10 mM NH_4HCO_3 (pH 10). Disaccharides were eluted with an acetonitrile gradient of 0% to 20% for 2.5 minutes, held at 20% for the next 2.5 minutes, with 2 minutes of equilibration at 0% before the next injection; the flow rate was 0.2 mL/min, and the total run time was 7.1 minutes. All disaccharides were detected and quantified in MRM acquisition mode, using the transitions m/z 378.1>175.1 for D0A0, 458.0>299.9 for D0a4, 378.0>175.0 for D0a0 and 458.0>97.0 for D0a6 (following the nomenclature of Lawrence et al.) and 472.0>97.0 for the 4UA-2S-GlcNCOEt-6S internal standard. The concentration of each disaccharide was calculated using a calibration curve for each of the disaccharides.

All samples were digested and analyzed in duplicates. For each subject, all serum and urine GAG levels were compared to age-matched controls (n=4). The normal range for each GAG was defined as mean ± 2x standard deviation of the control samples.

Results

Clinical report

We report on nine affected individuals from five unrelated families with neuro-immuno-skeletal dysplasia syndrome. Families A and B were consanguineous, suggesting autosomal recessive inheritance of the disorder (Figure 2A, Table 1, and Table S1-S2). Family C was from a geographically isolated district in Colombia (South America) with a possibility of distant degree of consanguinity. Family D was the offspring of families who lived for several generations in two small neighboring Portuguese villages. Family E was of Indian descent with no history of consanguinity.

All nine affected individuals presented with various skeletal dysplasias. In four families (families, A, C, D, and E) X-rays revealed severe platyspondyly leading to disproportionate stature (Figure 1). Lumbar gibbus and kyphoscoliosis were observed in families C and D, respectively. Three cases (from families D and E) had cervical malformation with hypoplastic odontoid peg/cervical instability. Short stature was also present in families A and B but appears due to shortening of the limbs. Metaphyseal abnormalities of the long bones were found in families A and E, while families B, C, and D showed epiphyseal abnormalities (Figure 1). Brachydactyly was observed in affected individuals of families A, B, and D. In addition to the skeletal abnormalities all affected individuals in families B and C showed intellectual disability (ID). For individuals A:II-1, E:II-1, and E:II-2 this could not be assessed due to their early death at age 7 weeks, 6 months, and 10 months, respectively. Both siblings of family B are severely disabled, unable to speak and are currently wheelchair bound. Dysmorphic facial features were non-discriminatingly mild, but include coarse facial features, prominent nose, and full cheeks in almost all affected individuals (Figure S1). A single palmar crease was also observed in six out of nine affected individuals. A striking phenotype present in three of the five families is a T cell-negative SCID. Affected individuals from families A and B showed a lack of T cells while B cell numbers were normal (Figure 4). In family E there was an idiopathic $CD4^+$ lymphopenia with an absolute lack of naïve T cells (both $CD4^+CD45RA^+$ and $CD8^+ CD45RA^+$ T cells were missing). In families B and E, a typical skin rash prompted further evaluation of the blood cells.

Maternal T cells were excluded, and the oligoclonal T cell reaction found in the presence of an eosinophilia was compatible with an Omenn-like presentation of SCID.[1,2] Worth mentioning are also the liver cysts seen in affected individuals from families A and E. Although the phenotypes are widely variable and partly contrasting (severe ID versus no ID; severe platyspondyly versus mild; epiphyseal changes versus metaphyseal), all major malformations and features occur in three organ systems only: the immune system, the skeleton, and the intellectual disability. Taken together we classify these five families as phenotypic variability of an EXTL3 related neuro-immuno-skeletal dysplasia phenotype. The affected individuals of families A and E died in early infancy, whereas the others are currently between 2 and 38 years of age. Individual III-1 from family D died at 30 years of age due to a unknown cause. Table 1 gives an overview of the most affected organ systems in all affected individuals. Detailed case report and clinical data for each affected individual can be found in the Supplemental Note: Case Reports.

Whole-exome sequencing

Massive parallel sequencing of genomic DNA of nine affected individuals from five unrelated families was performed in four different centers (Figure 2A). WES data filtering was applied using a recessive disease model, since families A and B show consanguineous bonds and families B, D, and E have more than one affected individual in the family. The severity of the phenotype, especially in families A and E, suggests that the disease is caused by a rare genetic mutation. We therefore prioritized homozygous and compound heterozygous variants, nonsense variants, and variants with a phyloP score >2.7 as the majority of Human Gene Mutation Database (HGMD) annotated variants have such scores.[47] Applying these analysis rules in family A only two out of 37 candidate variants segregated with the disease in the family (Table 2, Table S3). These two mutations were homozygous missense mutations in *EXTL3* (c.1537C>T, p.Arg513Cys) [HG19 RefSeq NM_001440.3] and *CHST9* [MIM 610191, HG19

RefSeq NM_031422] (c.1133A>G, p.Asn378Ser) (Figure 2B). Being left with two candidate genes, we made use of the GeneMatcher tool, a web-based tool for researchers and clinicians working on identical genes.[22] A match was found in four independent families from different centers. WES data from these families revealed likely pathogenic homozygous missense mutations in *EXTL3*. Data from these families also revealed one or more genes carrying biallelic variants that segregated with the disease, however, in each family the variant in *EXTL3* was the most likely candidate based on the position and predicted impact of the variant together with the segregation analysis (Table 2, Table S3-7). In total we found four homozygous missense mutations; c.1382C>T (p.Pro461Leu), c.1537C>T (p.Arg513Cys), c.1970A>G (p.Asn657Ser), and c.2008T>G (p.Tyr670Asp). Variant analysis in the ExAC database showed that homozygous non-synonymous alterations in coding regions of *EXTL3* are not tolerated. Only one such alteration was present (c.1324G>C, p.Val442Leu), which is predicted to be tolerated by SIFT prediction software (analysis performed in October 2016). Furthermore, variants p.Arg513Cys and p.Tyr670Asp were both absent from the ExAC database. The other two variants p.Pro461Leu and p.Asn657Ser are only present in a heterozygous state, with detected MAF of 1.6×10^{-5} and 5.8×10^{-5}, respectively. Copy number variation (CNV) analysis, including the SNP-array in family B did not reveal any likely pathogenic alterations (data not shown).

Taken together, we found homozygosity for four missense mutations in *EXTL3* in five unrelated families as a cause of a recessive form of neuro-immuno-skeletal dysplasia.

EXTL3 antibody validation

In order to validate the specificity of the antibody targeting EXTL3 we made use of CRISPR/Cas9 techniques to knock out EXTL3 in HEK293T cells and study the protein lysate on western blot. A band of approximately 130 kDa was clearly visible in lanes 1, 2, and 4, corresponding to cell lysates from HEK293T cells transfected with sgRNA 1, sgRNA 2, and

the empty vector pLentiCRISP v2, respectively. The absence of a band in lane 3 corresponding to cell lysate from HEK293T cells transfected with sgRNA 3 validates the specificity of the antibody targeting EXTL3 (goat; AF2635; R&D Systems, Minneapolis, MN, USA) (Figure 3D).

EXTL3 is present during early T cell development

Investigation of immune cells detected complete or nearly complete absence of T cells but normal counts of B and NK cells in affected individuals from three families (Families A, B and E) (Table 3). There are only a limited number of gene defects known that result in T-minus SCID.[1,2] The disease is most often caused by the lack of signaling of the T cell receptor to propagate the development of thymocytes toward the single CD4+ and CD8+ T cell stage.[48] There is no published information about the actual level of EXTL3 in T cells. To elucidate the disease mechanism underlying the T cell deficiency seen in three of these five families we studied the levels of EXTL3 in both T and B lymphocytes of normal controls. EXTL3 was not observed by western blotting in circulating T or B lymphocyte cell types (Figure 3A). After strong activation by a combination of CD3 and CD28 during culture for 6 days into T cell blasts (CD4+CD27−; CD8+CD27−), low levels of EXTL3 were detected (Figure 3A). This was far less if any at all in case of activated B cells differentiating for 6 days into plasmablasts (CD38+/CD138+) (Figure 3A).

These findings suggest that the effect of the selective T cell defect that we observed in EXTL3 deficiency may depend on the early protein levels during hemato- and/or lymphopoietic development prior to the circulating stage in the peripheral blood. Indeed, EXTL3 was demonstrated on protein level in purified CD34+ hematopoietic stem cells (HSCs) (Figure 3B), as well as in the immature (leukemic) T cell lines (Figure 3A). As was expected from the analysis of T and B cell subsets from blood or upon *in-vitro* activation, in none of the more mature T or B cell lines, or any additional hematopoietic cell lines tested (including Jurkat,

Ramos, EBV B cell blasts, KG1a, HL-60, U937) demonstrated any EXTL3 on protein level (Figure 3A).

For assessment of thymic tissue from neonates for the level of EXTL3, we fractionated the thymocytes into early double-negative, double positive and late single-positive thymocytes. EXTL3 was observed in all fractions, including the single-positive CD4$^+$ or CD8$^+$ thymocytes (Figure 3B).

Upon immunohistochemistry staining of thymus tissue from a normal neonate that was obtained at post-mortem examination (cause of death asphyxia) EXTL3 showed moderate localization, throughout the thymus, in both the cortical and extramedullary niches (Figure 3C). These findings suggest that the level of EXTL3 is downregulated when T cells enter the circulation.

By determining the immunophenotype of circulating B and T cells in families A and B we confirmed the T cell selective effect in the *EXTL3*-mutated individuals with a so-called T negative SCID phenotype. Skewing of T cell memory subset distribution towards a highly differentiated phenotype was observed in individual B:II-1, which was not observed in the transplanted sibling B:II-2 (Figure 4A). Upon stimulation with anti-CD3/anti-CD28 there was an impaired proliferative response of both CD4$^+$ and CD8$^+$ T cells in individual B:II-1 compared to the transplanted sibling B:II-2 or healthy controls (Figure 4C). Furthermore, there was an inability of the T cells in individual B:II-1 to induce B cell proliferation (Figure 4C).

To exclude a functional B cell defect due to the early lack of EXTL3 at the HSC and common lymphoid progenitors (CLP) stage within the bone marrow, we generated plasmablasts and tested antibody release into the supernatant upon stimulation with CpG/IL-2 in families A and B. The memory B cell subsets distribution were normal for their age (Figure 4A). Functionally, these data showed a normal B cell activation and plasmablast formation with (low-normal for age) IgM and IgG release in individual A:II-1 (Figure 4B, D-F). The low *in-vitro* IgG production

in individual B:II-2 can be ascribed to the fact that she had been transplanted only one year earlier with few circulating donor memory B cells at this stage, whereas the older sibling B:II-1 shows 5% CD27+ memory B cells, plasmablast formation and production of IgG and IgM (Figure 4B, D-F).

On the other hand, in one of the two families with the p.Pro461Leu variant (family C), that were not prone to infections, normal lymphocyte counts in their initial white blood cell differentials was found. We confirmed that the lymphocyte numbers, maturation, subset distribution (naïve, central-memory and effector-memory T cells; naïve, non-switched and switched memory B cells; and NK cells) and proliferative capacity were completely normal (Figure S2).

Taken together, these immunological data indicate that a severe T cell defect is most apparent although variable, whereas the other hematopoietic cell lineages seem to develop normally.

EXTL3 mislocalizes in fibroblasts derived from affected individuals

In order to determine the effects of the herein identified missense variants, we compared EXTL3 localization in fibroblasts derived from A:II-1, B:II-1, and B:II-2 to cells from two unaffected individuals, used as controls, by immunofluorescence microscopy. Overlap of the EXTL3 signal with that of GM130, a Golgi marker, indicates that EXTL3 is localized in the Golgi complex in control fibroblasts. In contrast to the control cell lines, no EXTL3 was detected in the Golgi complex in fibroblasts derived from individual A:II-1 (Figure 5), further supporting the causality of EXTL3 mutations in this individual. Noteworthy however, we could detect EXTL3 in the Golgi complex in fibroblasts from family B affected individuals. We suspect that the p.Arg513Cys variant, which is located near the exostosin domain (Figure 2B), causes degradation of EXTL3 leading to its absence from the Golgi complex in cells from individual A:II-1. On the other hand, we speculate that the p.Tyr670Asp variant detected in family B does not affect EXTL3 localization, but rather disrupts its GlcNAc transferase activity as this variant lies in the glycosyl transferase family 64 region of the protein (Figure 2B).

Heparan sulfate levels are altered in material derived from affected individuals

We then determined whether the identified mutations in affected individuals A:II-1, B:II-2, C:II-1, D:IV-1, and E:II-1 affect GAG synthesis as EXTL3 is known to be essential for HS biosynthesis.[14] GAG concentrations were measured in serum and urine of A:II-1, B:II-1, and B:II-2, and compared to the concentration in aged-matched controls (Figure 6B-C). In addition, the GAG concentrations were measured in fibroblasts from all families. As the levels of most disaccharides were below the limit of quantification in several samples, we only present results for the most abundant disaccharides for each GAG; D0A0 for HS, D0a4 for DS, and D0a6 and D0a0 for CS. The concentrations of these disaccharides strongly correlate with total GAG concentrations.[45]

HS concentrations in fibroblasts of all the affected individuals except affected individual C:II-1 were markedly reduced compared to the levels in control fibroblasts. For the affected individuals in family B, the HS concentrations were also reduced in serum and urine, whereas both parents in family B had serum HS concentrations within normal range. Surprisingly, for individual A:II-1 HS concentrations were normal in serum and urine.

DS concentrations, as measured by the disaccharide D0a4, showed more variable results. DS concentrations were within normal range in fibroblasts from family B and D, reduced in families A and E, and increased in family C. In serum and urine, DS concentrations were elevated in all three affected individuals studied (A:II-1, B:II-1, B:II-2). The DS concentration was also elevated in the serum of the father in family B (B:I-1) and was below the limit of detection in serum of the mother (B:I-2).

The CS concentrations in urine and serum showed an increase of the non-sulfated CS disaccharide D0a0. The other CS disaccharide, D0a6, was only elevated in urine and serum of individual B:II-2, while this disaccharide was in normal range in the samples from both A:II-1 and B:II-1.

In summary, although the findings differ among affected individuals and among the biological materials investigated, the results show abnormal glycosylation patterns in the affected individuals with generally lower HS and elevated DS and CS concentrations.

Discussion

In this study we identified homozygous missense mutations in *EXTL3* in nine neuro-immuno-skeletal dysplasia individuals with a variable phenotype from five unrelated families. EXTL3 as well as the other four exostosin family members are involved in the biosynthesis of HS.[14,49] Specifically, EXTL3 regulates the initiation of HS chain formation by enzymatically catalyzing the binding of GlcNac to the GAG linkage region of a specific protein.[14] This process is crucial in early development as complete *Extl3* knockdown in mice is lethal around 9 days postcoitum.[16] Zebrafish extl3 box with point mutation tm70g have been reported to have penetrant pectoral fin and jaw phenotypes and lack swim bladders that can be completely rescued with wildtype extl3.[50] This may turn out to be a comparable phenotype to the human. Because the affected individuals reported in the current study survived past the neonatal period, we expect that the homozygosity for the variants; p.Pro461Leu, p.Arg513Cys, p.Asn657Ser, and p.Tyr670Asp leads to diminished EXTL3 activity. The mutations probably have different effects on EXTL3 function. This is supported by the immunofluorescence data (Figure 5) showing that EXTL3 is only absent from the Golgi complex in fibroblasts derived from individual A:II-1, while it was still detectable in fibroblasts derived from siblings B:II-1 and B:II-2. The localization defect in A:II-1 may be explained by degradation of EXTL3 due to the p.Arg513Cys variant in this family, whereas the variant in family B presumably affects the enzymatic function of EXTL3 without disrupting its localization to the Golgi complex. The latter speculation is based on the presence of the p.Tyr670Asp variant in the glycosyl transferase family 64 domain of EXTL3 (Figure 2B). The pathophysiological effect of the variants in EXTL3 identified in families C and D are clearly different from A, B, and E, further functional reconstitution assays should formally prove the effects on functional activity. The p.Pro461Leu

variant detected in families C and D, is located within the exostosin domain. This domain is shared between the EXT family members, and appears to be associated with skeletal phenotypes,[51] whereas the severe immunological phenotype seems to be associated with improper GAG modifications that rely on intact function of EXTL3. Which proteins, either cellular or extracellular,[52] will be most affected by the perturbed GAG production in individuals with EXTL3 variants remains unclear.

All four described EXTL3 variants show an effect on GAG concentrations, where in general all HS concentrations are lower and both DS and CS concentrations are elevated. This data supports the hypothesis that improper function of EXTL3 leads to an unbalanced GAG synthesis resulting in less HS and more DS and CS. The only unexpected findings were normal HS concentrations in urine and serum of individual A:II-1, who did have markedly reduced HS concentrations in fibroblasts, and normal HS concentrations in urine and serum in fibroblasts of affected individual C:II-1. At the moment we cannot explain these results and they require further investigation.

Despite the common neuro-immuno-skeletal dysplasia phenotype of the affected individuals, there clearly is phenotypic variability between the reported families with an immense impact on survival because of the presence or absence of the concomitant SCID phenotype. The skeletal phenotype is mainly spondyloepy(meta)physeal dysplasia, overlapping with the growing clinical spectrum of other linkeropathies (e.g. associated malformations, dislocations, etc). Linkeropathies are associated with short-trunk/short stature resulting from spinal malformations and platyspondyly together with joint laxity, dislocations and osteoporosis. Furthermore, extra-skeletal features such as intellectual disability, facial dysmorphisms and congenital anomalies are common in linkeropathies and are also present in the herein described affected individuals. Motor development appears to be seriously affected by the EXTL3-related neuro-skeletal features in most of the reported affected individuals (Table 1 and Table S1). Intellectual disability appears not to be a constant feature. Whether these

phenotypic differences are due to the different positions of the mutations in the gene, or caused by other as of yet unknown modifiers, remains unclear. That said, it is worth noting that a variable penetrance of a T cell immunodeficiency, as observed in the five presented families, has previously been reported in Schimke immuno-osseous dysplasia (SIOD) and cartilage-hair hypoplasia (CHH).[6,7]

With respect to the SCID phenotype in three families, we have demonstrated that the defect of EXTL3 is at the level of early T cell development in the thymus. According to current models of hematopoiesis, lymphoid-primed multi-potent progenitors (LMPPs) (Lin$^-$Sca-1$^+$c-Kit$^+$CD34$^+$Flt3^{hi}) and common myeloid progenitors (CMPs) (Lin$^-$Sca-1$^+$c-Kit$^+$CD34$^+$CD41^{hi}) establish an early branch point for separate lineage-commitment from hematopoietic stem cells. While the common lymphoid progenitor compartment seemed to be involved our further exploration of this cellular complexity in immunological development provided support for EXTL3 playing an important role at the thymic stage of T cell development. It has been described that HSPGs influence the hematopoietic progenitor cell development[18,19] in keeping with our finding that the EXTL3 level is highest in the hematopoietic stem cells and early progenitor cells.

Hematopoietic stem cell transplantation (HSCT) in one of the EXTL3 deficient SCID cases, resulted in a complete recovery of normal T cell development without affecting either the non-hematopoietic defects nor the abnormal GAG findings in blood and urine samples.

In summary, we present nine affected individuals from five independent families that share a neuro-skeletal dysplasia phenotype. The immunological features are variable, three out of five families presented with a severe T cell immunodeficiency caused by a lack of CD4$^+$ and CD8$^+$ T cells. WES analysis revealed all affected individuals harbored mutations in *EXTL3.* A remedy for this disorder is not available, however, the T cell-negative (S)CID phenotype seen in five of the nine affected individuals presented in this report gives an indication to take preventive

measures against infections and to maybe consider a HSCT procedure to cure the severe T cell defect in these affected individuals.

Supplemental Data

The supplemental data includes: nine case reports, two figures, and seven tables.

Acknowledgements

Web resources

1000 Genomes project, http://www.internationalgenome.org/

Exome Aggregation Consortium (ExAC), http://exac.broadinstitute.org

GeneMatcher, http://genematcher.org

Matchmaker Exchange, http://www.matchmakerexchange.org

NHLBI GO Exome Sequencing Project (ESP), https://esp.gs.washington.edu/drupal/

Online Mendelian Inheritance in Man (OMIM), http://www.omim.org/

PhenomeCentral, http://www.phenomecentral.org

RefSeq, http://www.ncbi.nlm.nih.gov/RefSeq

Uniprot, http://www.uniprot.org/

References

1. Bousfiha, A., Jeddane, L., Al-Herz, W., Ailal, F., Casanova, J.L., Chatila, T., Conley, M.E., Cunningham-Rundles, C., Etzioni, A., Franco, J.L., et al. (2015). The 2015 IUIS Phenotypic Classification for Primary Immunodeficiencies. J Clin Immunol *35*, 727-738.
2. Picard, C., Al-Herz, W., Bousfiha, A., Casanova, J.L., Chatila, T., Conley, M.E., Cunningham-Rundles, C., Etzioni, A., Holland, S.M., Klein, C., et al. (2015). Primary Immunodeficiency Diseases: an Update on the Classification from the International Union of Immunological Societies Expert Committee for Primary Immunodeficiency 2015. J Clin Immunol *35*, 696-726.
3. Kuijpers, T.W., Ijspeert, H., van Leeuwen, E.M., Jansen, M.H., Hazenberg, M.D., Weijer, K.C., van Lier, R.A., and van der Burg, M. (2011). Idiopathic CD4+ T lymphopenia without autoimmunity or granulomatous disease in the slipstream of RAG mutations. Blood *117*, 5892-5896.
4. Walter, J.E., Rosen, L.B., Csomos, K., Rosenberg, J.M., Mathew, D., Keszei, M., Ujhazi, B., Chen, K., Lee, Y.N., Tirosh, I., et al. (2015). Broad-spectrum antibodies against self-antigens and cytokines in RAG deficiency. J Clin Invest *125*, 4135-4148.
5. Warman, M.L., Cormier-Daire, V., Hall, C., Krakow, D., Lachman, R., LeMerrer, M., Mortier, G., Mundlos, S., Nishimura, G., Rimoin, D.L., et al. (2011). Nosology and classification of genetic skeletal disorders: 2010 revision. Am J Med Genet A *155A*, 943-968.
6. Ridanpaa, M., van Eenennaam, H., Pelin, K., Chadwick, R., Johnson, C., Yuan, B., vanVenrooij, W., Pruijn, G., Salmela, R., Rockas, S., et al. (2001). Mutations in the RNA component of RNase MRP cause a pleiotropic human disease, cartilage-hair hypoplasia. Cell *104*, 195-203.

7. Morimoto, M., Lewis, D.B., Lucke, T., and Boerkoel, C.F. (1993). Schimke Immunoosseous Dysplasia. In GeneReviews(R), R.A. Pagon, M.P. Adam, H.H. Ardinger, S.E. Wallace, A. Amemiya, L.J.H. Bean, T.D. Bird, C.T. Fong, H.C. Mefford, R.J.H. Smith, et al., eds. (Seattle (WA).

8. Boerkoel, C.F., Takashima, H., John, J., Yan, J., Stankiewicz, P., Rosenbarker, L., Andre, J.L., Bogdanovic, R., Burguet, A., Cockfield, S., et al. (2002). Mutant chromatin remodeling protein SMARCAL1 causes Schimke immuno-osseous dysplasia. Nat Genet *30*, 215-220.

9. Hennet, T. (2012). Diseases of glycosylation beyond classical congenital disorders of glycosylation. Biochim Biophys Acta *1820*, 1306-1317.

10. Freeze, H.H. (2013). Understanding human glycosylation disorders: biochemistry leads the charge. J Biol Chem *288*, 6936-6945.

11. Holmborn, K., Habicher, J., Kasza, Z., Eriksson, A.S., Filipek-Gorniok, B., Gopal, S., Couchman, J.R., Ahlberg, P.E., Wiweger, M., Spillmann, D., et al. (2012). On the roles and regulation of chondroitin sulfate and heparan sulfate in zebrafish pharyngeal cartilage morphogenesis. J Biol Chem *287*, 33905-33916.

12. Izumikawa, T., Egusa, N., Taniguchi, F., Sugahara, K., and Kitagawa, H. (2006). Heparan sulfate polymerization in Drosophila. J Biol Chem *281*, 1929-1934.

13. Kitagawa, H., Izumikawa, T., Mizuguchi, S., Dejima, K., Nomura, K.H., Egusa, N., Taniguchi, F., Tamura, J., Gengyo-Ando, K., Mitani, S., et al. (2007). Expression of rib-1, a Caenorhabditis elegans homolog of the human tumor suppressor EXT genes, is indispensable for heparan sulfate synthesis and embryonic morphogenesis. J Biol Chem *282*, 8533-8544.

14. Kim, B.T., Kitagawa, H., Tamura, J., Saito, T., Kusche-Gullberg, M., Lindahl, U., and Sugahara, K. (2001). Human tumor suppressor EXT gene family members EXTL1 and EXTL3 encode alpha 1,4- N-acetylglucosaminyltransferases that likely are involved in heparan sulfate/ heparin biosynthesis. Proc Natl Acad Sci U S A *98*, 7176-7181.

15. Kaidonis, X., Liaw, W.C., Roberts, A.D., Ly, M., Anson, D., and Byers, S. (2010). Gene silencing of EXTL2 and EXTL3 as a substrate deprivation therapy for heparan sulphate storing mucopolysaccharidoses. Eur J Hum Genet *18*, 194-199.

16. Takahashi, I., Noguchi, N., Nata, K., Yamada, S., Kaneiwa, T., Mizumoto, S., Ikeda, T., Sugihara, K., Asano, M., Yoshikawa, T., et al. (2009). Important role of heparan sulfate in postnatal islet growth and insulin secretion. Biochem Biophys Res Commun *383*, 113-118.

17. Bishop, J.R., Schuksz, M., and Esko, J.D. (2007). Heparan sulphate proteoglycans fine-tune mammalian physiology. Nature *446*, 1030-1037.

18. Rodgers, K.D., San Antonio, J.D., and Jacenko, O. (2008). Heparan sulfate proteoglycans: a GAGgle of skeletal-hematopoietic regulators. Dev Dyn *237*, 2622-2642.

19. Bruno, E., Luikart, S.D., Long, M.W., and Hoffman, R. (1995). Marrow-derived heparan sulfate proteoglycan mediates the adhesion of hematopoietic progenitor cells to cytokines. Exp Hematol *23*, 1212-1217.

20. McCormick, C., Duncan, G., Goutsos, K.T., and Tufaro, F. (2000). The putative tumor suppressors EXT1 and EXT2 form a stable complex that accumulates in the Golgi apparatus and catalyzes the synthesis of heparan sulfate. Proc Natl Acad Sci U S A *97*, 668-673.

21. Farhan, S.M., Wang, J., Robinson, J.F., Prasad, A.N., Rupar, C.A., Siu, V.M., Consortium, F.C., and Hegele, R.A. (2015). Old gene, new phenotype: mutations in heparan sulfate synthesis enzyme, EXT2 leads to seizure and developmental disorder, no exostoses. J Med Genet *52*, 666-675.

22. Sobreira, N., Schiettecatte, F., Valle, D., and Hamosh, A. (2015). GeneMatcher: a matching tool for connecting investigators with an interest in the same gene. Hum Mutat *36*, 928-930.

23. Philippakis, A.A., Azzariti, D.R., Beltran, S., Brookes, A.J., Brownstein, C.A., Brudno, M., Brunner, H.G., Buske, O.J., Carey, K., Doll, C., et al. (2015). The Matchmaker Exchange: a platform for rare disease gene discovery. Hum Mutat *36*, 915-921.

24. Buske, O.J., Girdea, M., Dumitriu, S., Gallinger, B., Hartley, T., Trang, H., Misyura, A., Friedman, T., Beaulieu, C., Bone, W.P., et al. (2015). PhenomeCentral: a portal for phenotypic and genotypic matchmaking of patients with rare genetic diseases. Hum Mutat *36*, 931-940.

25. Becker, J., Semler, O., Gilissen, C., Li, Y., Bolz, H.J., Giunta, C., Bergmann, C., Rohrbach, M., Koerber, F., Zimmermann, K., et al. (2011). Exome sequencing identifies truncating mutations in human SERPINF1 in autosomal-recessive osteogenesis imperfecta. Am J Hum Genet *88*, 362-371.

26. Gilissen, C., Arts, H.H., Hoischen, A., Spruijt, L., Mans, D.A., Arts, P., van Lier, B., Steehouwer, M., van Reeuwijk, J., Kant, S.G., et al. (2010). Exome sequencing identifies WDR35 variants involved in Sensenbrenner syndrome. Am J Hum Genet *87*, 418-423.

27. Hoischen, A., van Bon, B.W., Gilissen, C., Arts, P., van Lier, B., Steehouwer, M., de Vries, P., de Reuver, R., Wieskamp, N., Mortier, G., et al. (2010). De novo mutations of SETBP1 cause Schinzel-Giedion syndrome. Nat Genet *42*, 483-485.

28. Hoischen, A., van Bon, B.W., Rodriguez-Santiago, B., Gilissen, C., Vissers, L.E., de Vries, P., Janssen, I., van Lier, B., Hastings, R., Smithson, S.F., et al. (2011). De novo nonsense mutations in ASXL1 cause Bohring-Opitz syndrome. Nat Genet *43*, 729-731.

29. Hempel, M., Cremer, K., Ockeloen, C.W., Lichtenbelt, K.D., Herkert, J.C., Denecke, J., Haack, T.B., Zink, A.M., Becker, J., Wohlleber, E., et al. (2015). De Novo Mutations in CHAMP1 Cause Intellectual Disability with Severe Speech Impairment. Am J Hum Genet *97*, 493-500.

30. Mendoza-Londono, R., Fahiminiya, S., Majewski, J., Care4Rare Canada, C., Tetreault, M., Nadaf, J., Kannu, P., Sochett, E., Howard, A., Stimec, J., et al. (2015). Recessive osteogenesis imperfecta caused by missense mutations in SPARC. Am J Hum Genet *96*, 979-985.

31. Srour, M., Hamdan, F.F., McKnight, D., Davis, E., Mandel, H., Schwartzentruber, J., Martin, B., Patry, L., Nassif, C., Dionne-Laporte, A., et al. (2015). Joubert Syndrome in French Canadians and Identification of Mutations in CEP104. Am J Hum Genet *97*, 744-753.

32. Le Quesne Stabej, P., Williams, H.J., James, C., Tekman, M., Stanescu, H.C., Kleta, R., Ocaka, L., Lescai, F., Storr, H.L., Bitner-Glindzicz, M., et al. (2016). STAG3 truncating variant as the cause of primary ovarian insufficiency. Eur J Hum Genet *24*, 135-138.

33. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and Genome Project Data Processing, S. (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics *25*, 2078-2079.

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

35. de Ligt, J., Boone, P.M., Pfundt, R., Vissers, L.E., Richmond, T., Geoghegan, J., O'Moore, K., de Leeuw, N., Shaw, C., Brunner, H.G., et al. (2013). Detection of clinically relevant copy number variants with whole-exome sequencing. Hum Mutat *34*, 1439-1448.

36. Shi, Y., and Majewski, J. (2013). FishingCNV: a graphical software package for detecting rare copy number variations in exome-sequencing data. Bioinformatics *29*, 1461-1462.

37. Fromer, M., Moran, J.L., Chambert, K., Banks, E., Bergen, S.E., Ruderfer, D.M., Handsaker, R.E., McCarroll, S.A., O'Donovan, M.C., Owen, M.J., et al. (2012). Discovery and statistical genotyping of copy-number variation from whole-exome sequencing depth. Am J Hum Genet *91*, 597-607.

38. Fromer, M., and Purcell, S.M. (2014). Using XHMM Software to Detect Copy Number Variation in Whole-Exome Sequencing Data. Curr Protoc Hum Genet *81*, 7 23 21-21.

39. Poultney, C.S., Goldberg, A.P., Drapeau, E., Kou, Y., Harony-Nicolas, H., Kajiwara, Y., De Rubeis, S., Durand, S., Stevens, C., Rehnstrom, K., et al. (2013). Identification of small exonic CNV from whole-exome sequence data and application to autism spectrum disorder. Am J Hum Genet *93*, 607-619.

40. Engels, H., Wohlleber, E., Zink, A., Hoyer, J., Ludwig, K.U., Brockschmidt, F.F., Wieczorek, D., Moog, U., Hellmann-Mersch, B., Weber, R.G., et al. (2009). A novel microdeletion syndrome involving 5q14.3-q15: clinical and molecular cytogenetic characterization of three patients. Eur J Hum Genet *17*, 1592-1599.

41. Genomes Project, C., Abecasis, G.R., Auton, A., Brooks, L.D., DePristo, M.A., Durbin, R.M., Handsaker, R.E., Kang, H.M., Marth, G.T., and McVean, G.A. (2012). An integrated map of genetic variation from 1,092 human genomes. Nature *491*, 56-65.

42. aan de Kerk, D.J., Jansen, M.H., ten Berge, I.J., van Leeuwen, E.M., and Kuijpers, T.W. (2013). Identification of B cell defects using age-defined reference ranges for in vivo and in vitro B cell differentiation. J Immunol *190*, 5012-5019.

43. Kuijpers, T.W., van Leeuwen, E.M., Barendregt, B.H., Klarenbeek, P., aan de Kerk, D.J., Baars, P.A., Jansen, M.H., de Vries, N., van Lier, R.A., and van der Burg, M. (2013). A reversion of an IL2RG mutation in combined immunodeficiency providing competitive advantage to the majority of CD8+ T cells. Haematologica *98*, 1030-1038.

44. Sanjana, N.E., Shalem, O., and Zhang, F. (2014). Improved vectors and genome-wide libraries for CRISPR screening. Nat Methods *11*, 783-784.

45. Langereis, E.J., Wagemans, T., Kulik, W., Lefeber, D.J., van Lenthe, H., Oussoren, E., van der Ploeg, A.T., Ruijter, G.J., Wevers, R.A., Wijburg, F.A., et al. (2015). A Multiplex Assay for the Diagnosis of Mucopolysaccharidoses and Mucolipidoses. PLoS One *10*, e0138622.

46. Kingma, S.D., Wagemans, T., L, I.J., Wijburg, F.A., and van Vlies, N. (2014). Genistein increases glycosaminoglycan levels in mucopolysaccharidosis type I cell models. J Inherit Metab Dis *37*, 813-821.

47. Vissers, L.E., de Ligt, J., Gilissen, C., Janssen, I., Steehouwer, M., de Vries, P., van Lier, B., Arts, P., Wieskamp, N., del Rosario, M., et al. (2010). A de novo paradigm for mental retardation. Nat Genet *42*, 1109-1112.

48. Rothenberg, E.V., Ungerback, J., and Champhekar, A. (2016). Forging T-Lymphocyte Identity: Intersecting Networks of Transcriptional Control. Adv Immunol *129*, 109-174.

49. Lind, T., Tufaro, F., McCormick, C., Lindahl, U., and Lidholt, K. (1998). The putative tumor suppressors EXT1 and EXT2 are glycosyltransferases required for the biosynthesis of heparan sulfate. J Biol Chem *273*, 26265-26268.

50. Lee, J.S., von der Hardt, S., Rusch, M.A., Stringer, S.E., Stickney, H.L., Talbot, W.S., Geisler, R., Nusslein-Volhard, C., Selleck, S.B., Chien, C.B., et al. (2004). Axon sorting in the optic tract requires HSPG synthesis by ext2 (dackel) and extl3 (boxer). Neuron *44*, 947-960.

51. Peterson, H.A. (1989). Multiple hereditary osteochondromata. Clin Orthop Relat Res, 222-230.

52. Forsberg, E., and Kjellen, L. (2001). Heparan sulfate: lessons from knockout mice. J Clin Invest *108*, 175-180.

53. Pedersen, L.C., Dong, J., Taniguchi, F., Kitagawa, H., Krahn, J.M., Pedersen, L.G., Sugahara, K., and Negishi, M. (2003). Crystal structure of an alpha 1,4-N-acetylhexosaminyltransferase (EXTL2), a member of the exostosin gene family involved in heparan sulfate biosynthesis. J Biol Chem *278*, 14420-14428.

Titles and legends of figures

Figure 1. Clinical features of the skeleton of all five families A-E.

(A) Lateral spine radiographs arranged chronologically. The affected individuals show severe platyspondyly (red boxes) with progressively severe kyphosis with increasing age (visualized for C:II-1, D:III-1, and D:IV-1), except for the affected siblings of family B who do not show severe platyspondyly (green boxes). Affected individual C:II-1 shows an additional abnormality of absent ossification of the anterior parts of the vertebral bodies in the upper lumbar region. Furthermore, wide intervertebral spaces are visible and the vertebral bodies show some posterior constriction giving a 'pear-shape'. (B) Hand radiographs arranged chronologically. Up to the age of about 6 years there is delay in carpal bone ossification. The 1st metacarpals (arrowhead) and the II middle phalanges (broad arrow) are short. The V middle phalanges are short or absent in A:II-1, E:II-1, and E:II-2 (newborns). Affected individual B:II-2 has soft tissue preaxial polydactyly. (C) The pelvic radiographs show a variable degree of sloping acetabula with shallow lateral notches, coxa valga, and small capital femoral epiphyses. Elbow radiographs showing dislocated radial heads with dysplastic radial heads.

Figure 2. Four homozygous missense mutations in *EXTL3* in five unrelated families

(A) Pedigrees of five affected families. All affected individuals carry homozygous missense mutations in *EXTL3* (NM_001440.3). In family A II-1 is homozygous for c.1537C>T (M1). In family B, II-1 and II-2 are homozygous for c.2008T>G (M2). Families C and D segregate c.1382C>T. Family E segregates c.1970A>G (M4). (B) Schematic overview of *EXTL3* and corresponding protein (NP_001431) showing the position of the missense mutations and changes at the protein level. All four mutations are in exon 3. EXTL3 consists of two annotated domains, a transmembrane domain (amino acids 29-51) and a coiled coil domain (amino acids 86-148). In addition, there are two predicted Pfam domains; a conserved Exostosin domain (amino acids 190-500), and glycosyl transferase family 64 domain (amino acids 663-904) that

is conserved between enzymes that catalyse the transfer reaction of *N*-acetylglucosamine and *N*-acetylgalactosamine.[53]

Figure 3. EXTL3 levels in lymphocytes

EXTL3 levels in T and B lymphocyte cell lines. (A) EXTL3 is absent in any B-cell lymphoma cells, myeloid cells, or circulating B and T lymphocyte cell lines. However, strong T cell stimulation seemed to induce low levels of EXTL3. (B) Purified CD34+ hematopoietic stem cells as well as thymocyte cells are positive for EXTL3. This is true for early double-negative, double positive and late single-positive thymocyte cells. (C) Immunohistochemistry staining of normal neonate tissue for localization of EXTL3 shows a moderate localization throughout the thymus. (D) EXTL3 antibody validation. The EXTL3 antibody shows a band for cells transfected with sgRNA 1 and 2, and the empty vector indicating the presence of EXTL3 in the cell lysate. Upon transfection with sgRNA 3 an EXTL3 knock out cell line is obtained and there is no band visible, indicating the specificity of the antibody.

Figure 4. Immunophenotype and *in vitro* proliferation, plasmablast formation and immunoglobulin production

Data are shown for individuals A:II-1, B:II-1, B:II-2 and a healthy adult control. (A) T cell subsets gated on CD3⁺CD4⁺ or CD3⁺CD8⁺ lymphocytes and B cell subsets gated on CD19⁺CD20⁺ lymphocytes. Numbers indicate percentages in corresponding gates. (B) CFSE-labeled PBMCs were cultured with either medium or CpG/IL-2 and gated on CD19⁺CD20⁺/⁻ lymphocytes. Gates were set for either CD27^bright (left) or CFSE⁺/⁻ and CD38⁺/⁻ (right) B cells to analyze cell division and plasmablast formation. (C) CFSE-labeled PBMCs were cultured with either medium or anti-CD3/anti-CD28 and were gated on CD3⁺CD4⁺ or CD3⁺CD8⁺ T lymphocytes and CD19⁺CD20⁺/⁻ B lymphocytes to analyze direct (T cell) or indirect (B cell)

proliferation upon TCR stimulation, *ND* not determined. (D, E) Levels of IgG and IgM were determined by ELISA in the supernatants after 6 days of culture with either medium, CpG/IL-2 or anti-IgM/anti-CD40/IL-21.

Figure 5. Localization of EXTL3 to the Golgi complex is lost in p.Arg513Cys mutated cells

The localization of EXTL3 in fibroblasts derived from affected individuals is compared to two healthy controls. The cells are stained for EXTL3 (green), Golgi marker GM130 (red), and DAPI for the cell nucleus (blue). The localization of EXTL3 to the Golgi complex seen in control fibroblasts is lost in cells derived from affected individual A:II-1, but not in both fibroblast cell lines from affected siblings B:II-1 and B:II-2.

Figure 6. Mutations in EXTL3 affect glycosaminoglycan concentrations

GAG concentrations in urine, serum, and fibroblasts from affected individuals and aged-matched control samples. D0A0 concentrations for heparan sulfate (HS), D0a4 for dermatan sulfate (DS), and D0a6 (sulfated) and D0a0 (non-sulfated) for chondroitin sulfate (CS). Error bars indicate standard deviation (SD). (A) GAG concentrations in fibroblasts. Reduced HS concentrations in fibroblasts from all affected individuals except for C:II-1. The DS concentrations were within normal range for the affected family members of family B and of D:IV-1, increased in C:II-1, and reduced in A:II-1 and E:II-1. (B) GAG concentration in serum from affected individuals and both parents of family B. HS concentrations were normal in affected individual A:II-1 and in both parents of family B (B:I-1 and B:I-2), but, decreased in the siblings of family B (B:II-1 and B:II-2). DS concentrations were elevated in all affected family members of A and B, as well as in parent B:I-1. Non-sulfated CS concentration was elevated in individuals B:II-1 and B:II-2 and the sulfated CS concentration was markedly increased in

B:II-2. (C) GAG concentrations in urine. HS concentrations for A:II-1 were normal, whilst decreased for both B:II-1 and B:II-2. DS was elevated in all three affected individuals. The non-sulfated CS concentration was elevated in all affected individuals, whereas, D0a6 was only elevated in B:II-2.

Table 1. Overview of the most affected organ systems in all affected individuals.

	A:II-1	B:II-1	B:II-2	C:II-1	D:III-1	D:III-2	D:IV-1	E:II-1	E:II-2
Neurological abnormalities	n.a.	+++	+++	+	+[a]	-	-	n.a.	n.a.
Immunological abnormalities	+++	+++	+++	-	-	-	-	++	++
Skeletal abnormalities	+++	+	+	+++	+++	+++	+++	+++	+++

[a]Normal/borderline cognition

Number of plus signs indicates the severity of the abnormality

n.a. means not assessed due to early death of the affected individuals (A:II-1 at 6 weeks, E:II-1 at 6 months, E:II-2 at 10 months of age)

Table 2. Causative *EXTL3* (NM_001440.3) mutations detected by WES in the five families.

Family	Chr	HG19 position	Variant	Amino acid change	ExAC frequency	PhyloP	SIFT score	PolyPhen-2 score
A	Chr8	g.28575113	c.1537C>T	p.Arg513Cys	n.p.	6.238	0	1
B	Chr8	g.28575584	c.2008T>G	p.Tyr670Asp	n.p.	7.984	0	1
C,D	Chr8	g.28574958	c.1382C>T	p.Pro461Leu	1.65E-05	7.727	0	1
E	Chr8	g.28575546	c.1970A>G	p.Asn657Ser	5.80E-05	9.206	0	1

n.p. means not present in database

Table 3. Lymphocyte counts in *EXTL3*-mutated individuals compared to controls (< 2 years).

Case	CD3[+] T cell (per µl) [a]	CD3[+]CD4[+] T cells (per µl)	CD3[+]CD8[+] T cells (per µl)	CD19[+] B cells (per µl)	CD56[+] NK cells (per µl)	IgG (g/L)	IgA (g/L)	IgM (g/L)
A:II-1	0	0	0	1436	279	1.0	<0.1	0.3
B:II-1 (7 yrs)	367 (25%)	157 (8%)	210 (11%)	656 (34%)	262 (13%)	5.65	0.39	0.24
(8 yrs)	488 (25%)	176 (9%)	293 (15%)	819 (42%)	585 (30%)	4.52	0.29	<0.2
B:II-2 [b] (3 mo)	2442 (51%)	2442 (51%)	0	1436 (30%)	575 (12%)	12.9	0.2	2.5
C:II-1 (12 yrs)	819 (44%)	348 (19%)	221 (12%)	557 (30%)	453 (25%)	n.a.	n.a.	n.a.
D:III-1	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
D:III-2	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
D:IV-1 (12 yrs)	600 (36%)	382 (23%)	206 (12%)	538 (32%)	513 (31%)	n.a.	n.a.	n.a.
E:II-1 [c] (2.5 mo)	1800 (40%)	500 (11%)	1310 (25%)	2030 (45%)	360 (8%)	<1.6	<0.4	0.5
E:II-2 (8 mo)	10 (0.2%)	0	0	1950 (69%)	760 (27%)	4.7	<0.06	0.11
Pediatric controls	700-1500	400-900	400-1100	300-1500	50-400	5-10	0.7-1.5	0.3-1.2

[a] All values were obtained pre-HSCT

[b] B:II-2: IgG before administration of IVIG and before HSCT. Ig subclasses (g/l): IgG1 11.9, IgG2 1.69, IgG3 0.57, IgG4 0.19).

[c] Already developed Omenn syndrome

In brackets the percentage of the lymphocyte subtype within the total lymphocyte count. n.a. means no data available