Vera Kalscheuer ORCID iD: 0000-0001-6898-3259

Servi JC Stevens ORCID iD: 0000-0001-8769-3150

Raoul Hennekam ORCID iD: 0000-0002-6745-1522

Deleterious *de novo* variants of X-linked *ZC4H2* in females cause a variable phenotype with neurogenic arthrogryposis multiplex congenita

Suzanna G.M. Frints[1,2], Friederike Hennig[3], Roberto Colombo[4,5], Sebastien Jacquemont[6§], Paulien Terhal[7], Holly H. Zimmerman[8*], David Hunt[9], Bryce A. Mendelsohn[10], Ulrike Kordaß[11], Richard Webster[12], Margje Sinnema[1,2], Omar Abdul-Rahman[13], Vanessa Suckow[3], Deciphering Developmental Disorders (DDD) Study, Alberto Fernández-Jaén[14], Kees van Roozendaal[1], Servi J.C. Stevens[1,2], Merryn V.E. Macville[1,2], Salwan Al-Nasiry[15], Koen van Gassen[7], Norbert Utzig[16], Suzanne M. Koudijs[17], Lesley McGregor[18], Saskia M. Maas[19], Diana Baralle[9,20], Abhijit Dixit[21], Peter Wieacker[22], Marcus Lee[23], Arthur S. Lee[24,25], Elizabeth C. Engle[24,25,26,27], Gunnar Houge[28], Gyri A. Gradek[29], Andrew G.L. Douglas[9,30], Cheryl Longman[31], Shelagh Joss[31], Danita Velasco[32], Raoul C. Hennekam[33], Hiromi Hirata[34], Vera M. Kalscheuer[3]

[1]Department of Clinical Genetics, Maastricht University Medical Center[+], azM, Maastricht, The Netherlands

[2]Department of Genetics and Cell Biology, Faculty of Health Medicine Life Sciences, Maastricht University Medical Center[+], Maastricht University, Maastricht, The Netherlands

[3]Research Group Development and Disease, Max Planck Institute for Molecular Genetics, Berlin, Germany

[4]Catholic University of the Sacred Heart, Rome, Italy

[5]Center for the Study of Rare Inherited Diseases (CeSMER), Niguarda Ca' Granda Metropolitan Hospital, Milan, Italy

[6]Service de Génétique Médicale, CHUV, Lausanne, Switzerland

[7]Laboratories, Pharmacy and Biomedical Genetics Division, University Medical Centre Utrecht, Utrecht, The Netherlands

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/humu.23841.

[8]Department of Pediatrics, Division of Medical Genetics, University of Mississippi Medical Center, Jackson, Mississippi, USA

[9]Wessex Clinical Genetics Service, University Hospital Southampton NHS Foundation Trust, Southampton, UK

[10]Division of Genetics, Department of Pediatrics, University of California, San Francisco, USA

[11]MVZ für Humangenetik und Molekularpathologie GmbH, Zweigstelle Greifswald, Germany

[12]The Department of Neurology and Neurosurgery, The Children's Hospital at Westmead, Westmead NSW 2145, Australia

[13]Munroe-Meyer Institute for Genetics & Rehabilitation, University of Nebraska Medical Center, Omaha, NE 68198-5440, USA

[14]Hospital Universitario Quirón de Madrid, Madrid, Spain

[15]Department of Obstetrics and Gynecology, Prenatal Diagnostics & Therapy, Maastricht University Medical Center[+], azM, Maastricht, The Netherlands

[16]Universitätsmedizin Greifswald, Klinik für Kinder- und Jugendmedizin, Greifswald, Germany

[17]Department of Neurology, Maastricht University Medical Center[+], azM, Maastricht, The Netherlands

[18]SA Clinical Genetics Service, Women's and Children's Hospital, North Adelaide 5006, South Australia

[19]Department of Clinical Genetics, Amsterdam UMC location AMC, Amsterdam, The Netherlands

[20]Faculty of Medicine, University of Southampton, Southampton, United Kingdom

[21]Nottingham University Hospitals NHS Trust, City Hospital Campus, Nottingham, United Kingdom

[22]Institute of Human Genetics, Westfälische Wilhelms Universität Münster, Münster, Germany

[23]Department of Pediatrics, Division of Pediatric Neurology, University of Mississippi Medical Center, Jackson, Mississippi, USA

[24]Department of Neurology, Boston Children's Hospital and Harvard Medical School, Boston, MA, USA

[25]Broad Institute of MIT and Harvard, Cambridge, MA, USA

[26]Department of Ophthalmology, Boston Children's Hospital and Harvard Medical School, Boston, MA, USA

[27]Howard Hughes Medical Institute, Chevy Chase, MD, USA

[28]Center for Medical Genetics and Molecular Medicine, Haukeland University Hospital, Bergen, Norway

[29]Department of Medical Genetics, Haukeland University Hospital, Helse Bergen, Bergen, Norway

[30]Human Development and Health, Faculty of Medicine, University of Southampton, UK

[31]West of Scotland Regional Genetic Centre, Queen Elizabeth University Hospital, Glasgow, Scotland, UK

[32]Department of Pediatrics, Munroe-Meyer Institute for Genetics & Rehabilitation, University of Nebraska Medical Center, Omaha, NE 68198-5440, USA

[33]Department of Pediatrics, Amsterdam UMC location AMC, Amsterdam, The Netherlands

[34]Department of Chemistry and Biological Science, College of Science and Engineering, Aoyama Gakuin University, Sagamihara, Japan

[§]Current address: CHU Sainte-Justine, Montreal, Canada

*Current address: College of Allied Health Professions, University of Nebraska Medical Center, Omaha, NE, USA

Corresponding authors:

Dr. Suzanna G.M. Frints, MD, PhD, Clinical Geneticist

Maastricht University Medical Center+, azM

Department of Clinical Genetics

Postbox 5800

6202 AZ, Maastricht

The Netherlands

Phone: +31-43-3876543

Fax: +31-43-3875800

Email: s.frints@mumc.nl

Dr. Vera M. Kalscheuer, PhD

Max Planck Institute for Molecular Genetics

Ihnestrasse 63-73

D-14195 Berlin

Germany

Phone: +49-30-8413 1293

Email: kalscheu@molgen.mpg.de

Abstract

Pathogenic variants in the X-linked gene *ZC4H2*, which encodes a zinc-finger protein, cause an infrequently described syndromic form of arthrogryposis multiplex congenita (AMC) with central and peripheral nervous system involvement. We present genetic and detailed phenotypic information on 23 newly identified families and simplex cases that include 19 affected females from 18 families and 14 affected males from 9 families. Of note, the 15 females with deleterious *de novo ZC4H2* variants presented with phenotypes ranging from mild to severe, and their clinical features overlapped with those seen in affected males. By contrast, of the 9 carrier females with inherited *ZC4H2* missense variants that were deleterious in affected male relatives, 4 were symptomatic. We also compared clinical phenotypes with previously published cases of both sexes and provide an overview on 48 males and 57 females from 42 families. The spectrum of *ZC4H2* defects comprises novel and recurrent mostly inherited missense variants in affected males, and *de novo* splicing, frameshift, nonsense and partial *ZC4H2* deletions in affected females. Pathogenicity of two newly identified missense variants was further supported by studies in zebrafish. We propose *ZC4H2* as a good candidate for early genetic testing of males and females with a clinical suspicion of fetal hypo-/akinesia and/or (neurogenic) AMC.

Key words: ZC4H2-Associated Rare Disorders (ZARD), Xq11.2 microdeletion, *ZC4H2*, club foot/-feet, complicated spastic paraplegia/ spasticity, fetal hypo-/akinesia

1 Introduction

Arthrogryposis Multiplex Congenita (AMC) has a prevalence of one case per 3,000–5,000 newborns. It is defined as multiple joint contractures that involve at least two different body areas prior to birth. AMC is a descriptive term and present in over 400 specific conditions (Hall, 1984, 2014). One known causal factor is decreased fetal movement in utero (fetal hypo-/akinesia) (Hall, 2009). AMC can result from a single gene defect or a chromosomal abnormality and can be part of various syndromes. In these cases, autosomal dominant, autosomal recessive and X-linked inheritance is possible (Hall, 2014). There are currently over 800 genes connected with either arthrogryposis or other types of early contractures with over 150 forms due to deleterious defects of X-linked genes (Hunter et al., 2015).

One form of X-linked arthrogryposis was first described in 1985 in six men from three generations of one family and referred to as Wieacker-Wolff syndrome (WRWF; MIM# 314580) (Wieacker, Wolff, Wienker, & Sauer, 1985). All six affected had congenital contractures of the feet, slowly progressive predominantly distal muscle atrophy, visual dyspraxia, facial weakness and intellectual disability (ID). In 2013, this family was reported to carry a pathogenic missense variant of the X-linked gene *ZC4H2* (MIM* 300897) along with three additional unrelated families who carried *ZC4H2* missense variants in males and females, an unrelated male with a *de novo* chromosomal inversion that truncated *ZC4H2*, and two unrelated females harboring heterozygous *de novo ZC4H2* deletions (Hirata et al., 2013). In zebrafish, *zc4h2* knock-down caused abnormal swimming and

impaired alpha-motor neuron development, which could not be rescued by mutant proteins containing the pathogenic substitutions (Hirata et al., 2013).

In 2015, May and colleagues identified *ZC4H2* variants in four additional families, including a large family with syndromic X-linked ID (XLID) (MRXS4), previously described as having Miles-Carpenter syndrome (MCS) (Miles & Carpenter, 1991) and characterized by XLID, exotropia, distal muscle wasting, microcephaly, congenital contractures and low digital arches (May et al., 2015). Most recently, three simplex females with heterozygous *de novo* deleterious *ZC4H2* were reported, including an early truncating nonsense variant and two microdeletions (Godfrey, Dowlatshahi, Martin, & Rothkopf, 2018; Okubo et al., 2018; Zanzottera et al., 2017). As a result of these findings, these various allelic syndromes are now referred to as ZC4H2-Associated Rare Disorders (ZARD).

We here report 23 additional ZARD families and simplex cases with inherited or *de novo* pathogenic *ZC4H2* variants plus three cases with publicly available information in DECIPHER (https://decipher.sanger.ac.uk/). We compared the genetic and clinical results with previously published families, thereby extending the molecular and clinical spectrum of ZARD throughout life, and discuss the broad clinical spectrum and its clinical variability in both males and females. Furthermore, we report a late adult-onset mild form of slowly progressive spastic paraplegia in *ZC4H2* carrier females of one previously published family (Hennekam, Barth, Van Lookeren Campagne, De Visser, & Dingemans, 1991; Hirata et al., 2013). We emphasize that the ZARD phenotype in females with a pathogenic *de novo ZC4H2* variant can be highly variable.

2 MATERIALS AND METHODS

2.1 Editorial Policies and Ethical Considerations

The study was carried out in accordance with the Declaration of Helsinki and the protocol approved by the local ethical committees for clinical genetic investigations. Written informed consent was obtained for molecular genetic analysis, publication of clinical, radiological data and photographs from all participants or their legal guardians.

2.2 Genetic studies

DNAs were extracted from peripheral blood, skin fibroblasts, buccal cells or umbilical cord using standard procedures.

Families 1, 14, 16 and 18 (DECIPHER Patient IDs: 263304, 263305, 260529, 276496 and 296515) were recruited to the Deciphering Developmental Disorders (DDD) study. DNA samples were analyzed by the Wellcome Sanger Institute using array-CGH and whole exome sequencing (WES) (Wright et al., 2015). For families 2, 8, 13, 15, 19 and 24 trio WES was performed as described in more detail in Supp. Materials and Methods, and for families 3 and 5 essentially as reported previously (Neveling et al., 2013). For families 4 and 6 all *ZC4H2* coding exons were amplified by PCR and Sanger sequenced with gene-specific primer pairs previously published (Hirata et al., 2013). Also for family 6, all *TYR* and *OCA2* coding exons were amplified by PCR and Sanger sequenced. For family 7, see (Hennekam et al., 1991; Hirata et al., 2013). For family 9 WES was performed for the index proband (Figure 1, V:1), his unaffected parents (III:1 and III:2), and two of his affected male relatives (III:5 and IV:2). For family 17 WES was performed for the index proband (Figure 1, II:1) in a research context. For families 3 (umbilical DNA of II:3), 10,

11, 12, 20 and 23 array-CGH was carried out, and for family 12 the breakpoints of the chromosomal deletion were fine-mapped by serial PCR and bidirectional Sanger sequencing (for further details see Supp. Information and Figure S1). For family 21, trio whole genome sequencing (WGS) was performed on peripheral blood DNA from the proband and both parents through the 100,000 Genomes Project (The 100,000 Genomes Project Protocol v3, Genomics England. doi:10.6084/m9.figshare.4530893.v3. 2017). For family 22, trio WGS was performed on saliva DNA at Baylor Genome Center as part of the NIH Gabriella Miller Kids First Research Program. For this family, the GeneMatcher and matchbox nodes of the Matchmaker exchange database were used to obtain collaborations using the search term ZC4H2 (Arachchi et al., 2018; Philippakis et al., 2015; Sobreira et al., 2017). For family 24 trio WES was performed at Fulgent Genetics, US.

The *ZC4H2* variants identified by WES and WGS were subsequently confirmed in the patients using Sanger sequencing. Segregation analysis of the variants was performed using standard Sanger sequencing with gene-specific primers and for family 9 using restriction digestion of the specific PCR product (see Supp. Information for more details).

The novel *ZC4H2* variants identified through this study have been submitted to the LOVD database (https://databases.lovd.nl/shared/genes/ZC4H2).

2.3 X-inactivation studies

For X-inactivation studies, DNA extracted from blood lymphocytes or skin fibroblasts was analyzed for the methylation sensitive site of *FMR1* exon 1 (Carrel & Willard, 1996), the CAG-repeat of the Androgen Receptor (*AR*) gene (Allen, Zoghbi, Moseley, Rosenblatt, & Belmont, 1992), or of *ZMYM3*.

2.4 In silico analysis

The pathogenicity of missense variants was assessed using *in silico* tools including Combined Annotation Dependent Depletion (CADD, http://cadd.gs.washington.edu/) score (Kircher et al., 2014) and Provean (http://provean.jcvi.org/index.php) (Choi & Chan, 2015).

2.5. Extraction of variants and clinical phenotypes from the literature and the DECIPHER database

We established a list of all published *ZC4H2* variants reported in the literature as of January 2019 and three cases from DECIPHER, and included available clinical information in our analysis.

2.6 Knock-down and rescue experiments in zebrafish

Zebrafish were bred and maintained according to approved guidelines prescribed by the Committee on Use and Care of Animals at Aoyama Gakuin University (Japan). Zebrafish *zc4h2* (GenBank NM_199642) cloned into pCR4-TOPO vector (Invitrogen) was used for cRNA synthesis as previously described (Hirata et al., 2013). Antisense morpholino oligonucleotides (MO) designed against the exon 2- intron 2 splice donor site (MO2) of zebrafish *zc4h2* were used for knocking down *zc4h2*. Zebrafish embryos were injected with 5 ng of MOs at 1-2 cell stages and studied as published before (Hirata et al., 2013). For rescue experiments, the missense variants were introduced into the wild-type mouse *Zc4h2* (RefSeq NM_001003916.2) pCS2+ expression construct by site-directed mutagenesis (primer sequences are provided in Supp. Table S1). Capped RNA was synthesized using mMESSAGE mMACHINE SP6 kit (Life Technologies) according to the manufac-

turer's protocol. Capped RNA (100 pg) was co-injected with MO2 (4 ng) into zebrafish embryos at 1-2 cell stages. At 2 days post-fertilization, normal zebrafish embryos swim away rapidly (>2 cm/s) upon tactile stimulation. The number of embryos that exhibited slow swimming (<2 cm/s) following touch was counted.

2.7. *ZC4H2* reference sequence

All *ZC4H2* variants reported are based on GRCh37/hg19 and transcript isoform 1 (RefSeq NM_018684.3) with nucleotide (cDNA) numbering using +1 as the A of the ATG translation initiation codon in the reference sequence, with the initiation codon as codon 1.

3 RESULTS

3.1 Clinical features in the new patient cohort and in previously published families

We report on 23 novel families and simplex cases with likely deleterious *ZC4H2* variants and on one previously published family, family 7 of this study (Hennekam, et al., 1991; Hirata, et al., 2013). The pedigrees of the novel families are shown in Figure 1 (Fam1-6, Fam8-24) and of family 7 in Supp. Figure S2. Detailed clinical descriptions are presented in Supp. Results. In the clinical evaluation we also included publicly accessible information of three females reported to DECIPHER and all families and simplex cases published to date (Godfrey et al., 2018; Hirata et al., 2013; Kondo et al., 2018; May et al., 2015; Okubo et al., 2018; Zanzottera et al., 2017). A compilation of the main (≥30%) clinical features of affected males and females from these 42 families is given in Table 1 and of the additional less common clinical features (<30%) in Supp. Table S2. Percentages are related to the total number of probands who were clinically evaluated for a given

feature (positive informative). In total, informative clinical data was collected for 87 out of 105 (83%) probands, including 44 out of 48 (92%) males and 43 out of 57 (75%) females using the Human Phenotype Ontology (https://hpo.jax.org/app/) scoring list. Their ages ranged from 23 weeks of gestation to >70 years at the time of clinical genetic investigation. Families were recruited worldwide and were of Caucasian and Asian ethnicity.

Prenatal presentation of five fetuses with ZARD (two males and three females), included club foot/feet, rocker bottom feet, fetal hypo/akinesia (mostly at the end of the second or third trimester of pregnancy, >18-26 weeks of gestation), contractures, AMC, nuchal and/or frontal head edema and (nearly) normal growth parameters (See Supp. Videos 1-3 and Supp. Figure S3). One affected male fetus presented with hypogenitalism (cryptorchidism and micropenis).

At birth and neonatal age, the clinical genetic diagnosis in males varied from descriptive clinical features e.g. AMC with or without additional clinical features including: psychomotor developmental delay, short stature, brain atrophy, microcephaly, tetraplegia, congenital general hypotonia, facial weakness – palsy, muscle weakness, complex spasticity, (flexion) contractures of large and/or small joints, congenital hip dislocation – subluxation, rocker bottom feet, club feet, camptodactyly, ptosis, to (un)recognizable syndromes such as cerebral palsy, MCS, WRWF with or without cleft palate, feeding difficulties, laryngomalacia, hyperinsulinemic hypoglycemia, mixed obstructive and central sleep apnea.

The clinical genetic features in 33 newly identified and published carrier females (from childhood till late adulthood) with a maternally inherited *ZC4H2* variant varied from normal phenotype (n=13) to mild (mainly hand-finger) flexion contractures, borderline to

mild ID (n=20) with facial dysmorphism such as ptosis, strabismus and long (flat) philtrum. Also, distal muscle weakness, urine (stress) incontinence, anosmia, and walking difficulties which were slowly progressive in a few (n=3) females in late adulthood were reported.

In marked contrast, the 25 females with a *de novo* pathogenic variant (from neonatal age till adulthood) of *ZC4H2* including Xq11.2 microdeletions showed high variation in clinical presentation and ranged from mildly to severely affected (Table 1, Supp. Table S2).

Overall, common clinical features in child- and adulthood in both males and females (30% or more positive informative in probands) included postnatal growth retardation, generalized hypotonia, motor delay, inability to walk, spasticity, hyperreflexia, urinary incontinence, dysarthria-deficit in expressive language, poor or absent speech, ID, drooling, dysphagia, chewing difficulties including oral motor dysfunction, feeding difficulties, facial weakness - palsy, high forehead, high anterior hairline, ocular motor apraxia, strabismus, anteverted nares, microretrognathia, AMC, limited shoulder movement, elbow, wrist contractures, metacarpophalangeal joint contractures, camptodactyly, radial deviation of any finger, knee flexion contractures, equinovarus deformity - club feet, Achilles tendon contracture, distal limb muscle atrophy or weakness and micropenis or cryptorchidism in males (Table 1, Supp. Table S2, Supp. Figures S4, S5).

Less commonly reported clinical features, which were present in <30% of the probands included short stature, microcephaly (more prominent in males (16/26) than in females (4/30)), round face, low-set (posteriorly rotated) ears, upslanting palpebral fissures, almond-shaped eyes, ptosis, deeply set eyes, pupillary dysfunction, nystagmus, short nose, short philtrum, broad alveolar ridges, high-arched palate, and (submucous) cleft palate

(Supp. Figures S3, S4). Furthermore, U-shaped upper lip vermillion/ carpshaped mouth and downturned corners of the mouth were reported. A short neck (with limited rotation) was more frequently reported for females with a *de novo* variant (12/14) than for affected males with an inherited or *de novo* variant (12/37) (Table 1, Supp. Table S2, Supp. Figure S5). Of note, neonatal respiratory distress, recurrent aspiration pneumonia, and (obstructive sleep) apnea can be a particular concern with apnea more frequently reported in males (11/16) than in affected females with a *de novo* variant (3/11). Additionally reported clinical features included narrow chest, narrow shoulders-thorax (more frequently reported in females with a *de novo* variant (11/15) than in affected males (18/36)) (Supp. Figure S5), umbilical hernia (Supp. Figure S3A, 6.III:1), cervical and/or thoracic kyphosis, scoliosis, narrow pelvis, congenital hip contracture, hip dislocation/subluxation, short limbs, proximally placed thumb, ulnar deviation of any finger, overlapping – proximally placed toe(s), edema of the dorsum of hands and feet (Supp. Figure S5), generalized hypotonia, absent speech, epileptic seizures, abnormal cortical gyration, delayed CNS myelination, global brain atrophy and ventriculomegaly (Supp. Figure S6). So far, impaired smell was only reported in carrier females in late adulthood. Upslanting palpebral fissures (Supp. Figure S4), arrhythmia, sick sinus syndrome, hypoglycemia and aggressive behavior were so far only reported in males (Table 1 and Supp. Table S2).

Besides the ZARD typical features such as AMC, hypo-/akinesia, club foot/feet and facial dysmorphy, associated sporadic clinical features which can be a clue to clinical diagnosis in probands with ZARD include pupillary dysfunction, optic disk cupping, short neck with limited rotation with narrow chest – thorax – shoulders with limited shoulder movements, malposition of the stomach, diaphragmatic eventration, focal autonomic seizures,

electrical status epilepticus during slow sleep (ESES) seizure pattern on electroencephalogram (EEG), pancreatic hypoplasia (and postprandial hypoglycemia), and congenital achalasia (Supp. Table S2).

Central and peripheral neurological findings in affected individuals included: ID, poor or absent speech, dysarthria, deficit in expressive language, dysphagia, drooling, and chewing difficulties including oral motor dysfunction, motor delay, inability to walk, spasticity, hyperreflexia, areflexia and dystonia. Various types of seizures, generalized hypotonia, impaired smell (3/13, adult carrier females) and urinary (stress) incontinence (16/28) were also mentioned.

MRI brain and spine images showed variable and global brain atrophy (11/31) (examples are given in Supp. Figure S6), delayed CNS myelination (9/30), abnormality of periventricular white matter (4/27), corpus callosum abnormality (7/26), abnormal cortical gyration (5/24), ventriculomegaly (8/20), polymicrogyria, anterior to posterior gradient (1/8 males), tethered cord (3/19) and hydromyelia (2/10 males). Abnormal peripheral nerve conduction was present in one affected (1/7 males). Behavioral phenotypes included emotional lability (5/14) and aggressive behavior (3/5 males) (Supp. Table S2).

Cardiovascular associated clinical features, e.g. arrhythmia (5/13), bradycardia (5/8), (congenital) sick sinus syndrome (5/13) and right ventricular hypertrophy (4/12) were so far reported in affected males only (Supp. Table S2).

3.2 ZC4H2 genetic characteristics in the new cohort and global mutational spectrum

Most of the novel families were investigated by WES, WGS and/or array-CGH. For two affected index males (family 4, II:2 and family 6, III:1) it was assumed that the phenotype

could be caused by a pathogenic *ZC4H2* variant and therefore all coding exons were screened by Sanger sequencing.

A total of eight males from six families (families 1, 3, 4, 5, 9, and 19) inherited the variant from a healthy carrier mother, while three males from three families (families 5, 6, and 9) inherited the variant from a mildly affected mother (Figure 1). For one affected male (family 9, II:5) the phenotype of his mother is unknown. In the moderately affected male from family 18 (Figure 1, II:1 and Supp. Figures S4, S5) and the severely affected male from family 24 (Figure 1, II:1) the missense variant occurred *de novo*.

In 15 females the variant occurred *de novo* (Figure 1), including individual II:1 from family 6 who presented with contractures of her metacarpophalangeal finger joints and is otherwise healthy. By contrast, the female fetus of family 3 (Figure 1, II:3) inherited the pathogenic variant from her healthy mother who is a mosaic and carries the variant in about 10% of her cells (Supp. Figure S7) and the two mildly affected females from family 9 (Figure 1, II:3 and II:6) carry a maternally inherited missense variant.

A total of 23 variants were identified (Figure 2a, b). Thirteen of the 23 were *de novo* loss-of-function variants (nonsense, splice-site, frameshift, CNVs) present in affected females and ten were missense variants identified in affected males which, except for the two *de novo* variants in affected males (families18 and 24), were inherited from their mildly affected or asymptomatic carrier mothers. Twenty of the 23 variants were novel including single nucleotide variants (SNVs) and five *de novo* microdeletions in females. The microdeletions removed the first exon of *ZC4H2* but did not affect any adjacent genes. Two missense variants were defined as recurrent. Families 4 and 24 of this study carry a previously reported p.(Arg198Gln) change in an unrelated family (Hirata et al., 2013). Family

18 of this study has a *de novo* p.(Arg213Trp) variant previously reported in three unrelated families (Hirata et al., 2013; May et al., 2015). Also, in two families in this study (families 5 and 6) the same amino acid (Ala200) was altered. Finally, the proline residue 201 mutated to serine in a published family (Hirata et al., 2013) was mutated to histidine in family 19 of this study.

None of the variants identified in this study were reported in publicly available population databases, including the 1000 Genomes project database and Genome Aggregation Database (gnomAD, http://gnomad.broadinstitute.org/). Also, there were no deletions reported in the database of genomic variants (DGV, http://dgv.tcag.ca/dgv/app/home?ref=GRCh37/hg19) overlapping *ZC4H2* exons, demonstrating that the variants identified in the patients are likely not common polymorphisms and supporting our suggestion that they are deleterious to protein function. All missense variants altered highly conserved amino acids (fully conserved from human to zebrafish; Figure 2c and Supp. Figure S8). Also, in silico predictions using CADD revealed high scores (>22) and using Provean all variants were predicted as deleterious, except for the p.(Lys217Arg) change, which perfectly co-segregated with the phenotype in a large family (Fam9). At the gene level, *ZC4H2* is highly constrained with a pLI score of 0.91 and zero known LOF mutations in gnomAD (Lek et al., 2016).

Combined with our results of this study, there are currently 31 unique *ZC4H2* variants known in affected males and females, including one *de novo* X-inversion interrupting *ZC4H2* in a male plus ten *de novo* Xq11.2 microdeletions in females (Figure 2a). An overview of all variants is provided in Supp. Table S3. Thus far, SNVs leading to mis-

sense changes cluster in the last exon of *ZC4H2*, which encodes the zinc-finger domain and the most C-terminal part of the protein.

The X-inactivation pattern in blood or skin fibroblasts in affected females varied from random to skewed and was not useful to predict the phenotype (Figure 1 and Supp. Results), e.g. one severely affected female fetus showed an Xi ratio of 61:39 in skin fibroblasts (Figure 1, Fam3, II:3). Other mild to severely affected females showed Xi ratios varying between 100:0 and 80:20 (Figure 1) in blood lymphocytes. Previous asymptomatic carrier females reported in the literature showed also a ratio of >95:5 (e.g. Hirata et al., 2013).

3.23 Zebrafish Studies

We investigated the potential effects of the p.(Ala200Val) and p.(His70Gln) missense variants which we identified early on in zebrafish morphants similar to our previous studies (Hirata et al., 2013). Following knock-down of *zc4h2*, zebrafish morphants showed impaired swimming capability at 2 days post-fertilization (27/34, 79% of morphants) due to compromised swimming contraction. This defect could be rescued with wild-type Zc4h2 (7/29, 24% of morphants), but not with constructs carrying the p.(Ala200Val) (27/35, 77% of morphants) or p.(His70Gln) (29/41, 71% of morphants) variants and thus the results strongly supported pathogenicity of the altered amino acids.

4 DISCUSSION

This study reports on the genetic and clinical findings of 23 novel families and simplex cases with deleterious inherited or *de novo ZC4H2* variants identified in males and females, including *de novo* partial *ZC4H2* microdeletions present in affected females only,

and a review of the ZARD literature. Moreover, we have investigated pathogenicity of two single amino acid substitutions in zebrafish as described before (Hirata et al., 2013).

Inherited and *de novo* variants are present in all coding exons and there is no hotspot in the gene. On the protein level, most of the missense changes lie within in the C-terminal part of ZC4H2 including the zinc-finger domain. All mutated amino acids of ZC4H2 are highly conserved and predicted to be functionally relevant. The two missense variants tested in zebrafish resulted in impaired swimming of the mutants, supporting functional importance of the mutated amino acids. Two of the *de novo* pathogenic missense variants identified in females (families 8 and 15, p.(Cys206Phe) and p.(Cys203Ser) respectively) altered one of the four cysteine residues of the Cys4His2 type zinc finger. The function of this zinc finger is currently unknown.

Most *ZC4H2* pathogenic variants identified in affected males are missense changes and, except for two *de novo* missense variants identified in affected males (families 18 and 24), were inherited from mostly asymptomatic mothers (X-recessive). This is well in line with previous findings. To the best of our knowledge, with only two exceptions there are no pathogenic variants reported in affected males predicted to lead to a truncated or completely absent ZC4H2 protein. The two exceptions are the *de novo* X-chromosome inversion which disrupted *ZC4H2* at one of the breakpoints in a severely affected boy (Hirata et al., 2013) and the affected boy from family 3 who we assume, based on clinical presentation, carried the maternally inherited p.(Glu92Vfs*7) variant but did not undergo genetic testing. He was born after Caesarean section because of fetal distress at 40 weeks pregnancy with multiple congenital malformations and died 10 hours after birth because of respiratory distress. He had short limbs, AMC and a cleft palate. Affected males can have

syndromic variable XLID (mild-moderate ID) with (mild) hypotonia, spasticity (as seen in SPG16 (Steinmuller et al., 1997)), mild facial dysmorphism e.g. high forehead with high frontal hairline, ocular dyspraxia, relative short stature and (relative) microcephaly. The differential diagnosis is still challenging and earlier reported XLID families could benefit from WES including *ZC4H2* analysis (Supp. Table S4).

In contrast to the mostly inherited missense variants in affected males, the spectrum of *de novo* pathogenic *ZC4H2* variants identified in affected females includes missense, early stop, frameshift and splicing variants, as well as Xq11.2 microdeletions which removed *ZC4H2* completely or exon 1 of all transcript isoforms. Thus, from the current data it seems that pathogenic variants predicted to lead to a complete loss of ZC4H2 protein function are very rare in males, while they usually occur *de novo* in females. Of note, none of the pathogenic variants identified in males were found *de novo* in female simplex cases. The *de novo* variants in the affected females are predicted to be loss-of function alleles, suggesting *ZC4H2* insufficiency as the most likely pathological mechanism leading to an X-linked dominant phenotype.

Our detailed clinical characterization (Table 1 and Supp. Table S2) confirms and extends clinical findings reported for the few published families and simplex cases caused by deleterious *ZC4H2* SNVs in males and *de novo ZC4H2* microdeletions in females. Furthermore, we report the first mosaic frameshift variant of *ZC4H2* in a family with AMC and fetal hypo-/akinesia. Combined, the results indicate that, in males, inherited forms of ZARD are associated with complete penetrance but variable expressivity, ranging from nonspecific XLID to XL-AMC with variable mostly moderate to severe forms of ID, fetal hypo-/akinesia, postnatal growth retardation, spasticity, tetraplegia, (relative) microceph-

aly and hypogenitalism. By contrast, in females, the inherited forms of ZARD are associated with incomplete penetrance and variable expressivity, ranging from unaffected carriers to XL-AMC with variable cognition, postnatal growth retardation, poor or absent speech, spasticity, inability to walk, distal muscle wasting/atrophy, short neck with limited rotation, narrow chest and limited shoulder movements. Moreover, clinical follow-up of family 7 for more than 25 years (Hennekam et al., 1991) indicated very mild progressive muscle weakness in two former asymptomatic carrier females and marked progression including loss of the ability to walk and distal muscle wasting in one carrier female (Supp. Figure S2, III:5, III:7 and III:14). She also had episodic periods of pain of unknown origin. All three females had inability to smell and urine stress incontinence. As differential diagnoses a late-onset mild progressive syndromic form of spastic paraplegia (SPG), multiple sclerosis (MS), amyotrophic lateral sclerosis (ALS), or Charcot-Marie-Tooth (CMT) were clinically (re)considered. While we cannot exclude co-morbidity with other X-linked forms of SPG (e.g. *PLP1*, *SLC16A2*, *TAF1*), ALS (*SMAX1*) and CMT (*GJB1*) which could not be tested genetically, the females are all carriers of a *ZC4H2* pathogenic variant causing ZARD in their male offspring. This makes it more likely that their late-onset phenotype belongs to ZARD's clinical variability. In contrast to this late-onset neurodegenerative adult phenotype reported so far in one family, there is no early-onset progressive neurodegenerative phenotype known in affected children.

Compared to males, females with a pathogenic *de novo* variant of *ZC4H2* presented with a more variable phenotype. Cognition ranged from normal (families 3, 6, and 23) to mild learning difficulties (families 8, 21, and 22), as reported for affected who carry missense, frameshift or splicing variants and a partial gene deletion, respectively, to moderate or

severe ID (families 10-17). Furthermore, clinical characteristics of the *de novo* Xq11.2 microdeletions present in females showed that these deletions are associated with a recognizable phenotype and that the overall clinical outcome varies between mildly and severely affected and thus cannot be predicted. Speech delay and poor speech is a remarkable clinical feature in the affected, so speech therapy should be started as early as possible to optimize communication leading to improved quality of life. Also, a large Xq11.2 deletion which, in addition to *ZC4H2*, removed *ARHGEF9, WTX* and *MTMR8* was reported in a female with severe ID and normal X-inactivation (Holman et al., 2013). The major clinical phenotype was caused by the *WTX* deletion leading to osteopathia striata with cranial sclerosis and it was suggested that deletion of *ZC4H2* contributed to her severe ID. Given the results obtained through this study it is possible that in addition to ID other clinical features reported for this female, such as clenched hands with immobile joints, early onset complex partial seizures (controlled on levetiracetam), dysphagia and lower extremity spasticity could be caused by ZARD.

Two of the pathogenic *ZC4H2* missense variants identified in this study are recurrent. Firstly, the maternally inherited amino acid change p.(Arg198Gln) identified in the affected males of one family (family 4) and *de novo* in the affected male of an unrelated family (family 24) has previously been reported in a large German family (Hirata et al., 2013) and there are phenotypic overlaps with affected males being severely affected. Secondly, the *de novo* p.(Arg213Trp) substitution identified in a moderately affected male (family 18) has previously been reported in three large families (Hirata et al., 2013; May et al., 2015). The clinical phenotype is variable within and between families, but all affected males (Family K8615, May, et al., 2015, damilies 4 and 5, Hirata, et al., 2013)

presented with moderate to severe ID with absent or poor speech. Mild facial dysmorphism was evident in one family only (family 5, Hirata, et al., 2013). Other common shared clinical features included motor delay, inability to walk, hyperreflexia, spasticity and seizures. Skeletal findings, e.g. AMC were reported in a few affected males (2/2 males of family 4 and 1/2 males of family 5, Hirata, et al., 2013). All female carriers, except one, of the three families had borderline or mild ID (family K8615, May, et al., 2015, families 4 and 5 Hirata, et al., 2013).

Overall, despite the growing number of families and simplex cases with deleterious *ZC4H2* variants, there is currently no evidence for a clear genotype-phenotype correlation. Also, as discussed above, clinical presentations of affected individuals who carry the same pathogenic variant leading to ZARD can vary within families and between families. Nevertheless, there are some gender specific clinical features, e.g. heart rhythm disturbances and hypogenitalism in males.

The striking phenotypic differences between females with inherited X-linked recessive *ZC4H2* missense variants versus *de novo* X-linked dominant deleterious *ZC4H2* nonsense, frameshifting or Xq11.2 microdeletions is similar to that seen for a few other X-linked genes, such as *ARX* (Bienvenu et al., 2002; Mattiske et al., 2017; Stromme et al., 2002), *HDAC8* (Kaiser et al., 2014), *PHF6* (Lower et al., 2002; Zweier et al., 2013), *IQSEC. 2* (Ewans et al., 2017; O'Rawe et al., 2015; Shoubridge et al., 2010; Zerem et al., 2016) and *CLCN4* (Hu et al., 2016; Palmer et al., 2018). Also gender-specific pathogenicity differences of inherited variants in males and *de novo* variants in females have been reported for other XLID genes, e.g. *DDX3X* (Dikow et al., 2017; Snijders Blok et al.,

2015), *KIAA2022* (Lorenzo et al., 2018; Van Maldergem et al., 2013), *MTM1* (Schara, Kress, Tucke, & Mortier, 2003) and *CASK* (Moog et al., 2011).

ZC4H2 is known as a gene which is subject to X-inactivation. Current results suggest that, in heterozygous carrier females, X-inactivation status in blood and skin fibroblasts does not predict the clinical outcome. This is very much in line with results obtained for other X-linked disease genes and also with recent findings suggesting that skewed X-inactivation is common in the general female population (Shvetsova et al., 2018). Still, the variable clinical manifestations of heterozygous carrier females with a *de novo* variant ranging from very mildly to severely affected may be partially explained by the X-inactivation status within specific affected cells and tissues. However, other genetic, environmental or stochastic factors may also be involved and could have an impact on phenotypic variability and severity of ZARD.

Similar to patients with other types of AMC who develop joint contractures during pregnancy, abnormal fetal movement due to ZARD can be identified using real time ultrasound prenatally, as we report here for affected fetuses from several families (for family 3 see Supp. Videos 1-3). An early diagnosis of AMC prenatally would allow parents and clinicians early decision making, e.g. the possibility of in utero therapy (increasing movement in utero), or early delivery at a time when lungs are mature but the contractures may not yet be so severe (Hall, Agranovich, Ponten, & van Bosse, 2015). However, AMC and fetal hypo-/akinesia can occur late in pregnancy, after routine prenatal ultrasound sonography has taken place between 18-22 weeks of gestation, thus the diagnosis of AMC and/or fetal hypo-/akinesia can be easily missed. Moreover, reduced fetal hand, finger and feet movements are not examined routinely by ultrasound and occasionally

club foot/feet is/are the only clinical feature which fetuses with a pathogenic *ZC4H2* variant presented prenatally or neonatally.

In conclusion, *ZC4H2* is one of the more commonly mutated XL-AMC and XLID genes. Up to now *de novo* pathogenic variants of *ZC4H2* have been more frequently seen in affected females than in males with ZARD and displayed a broad clinical spectrum ranging from mildly to severely affected patients with neurogenic AMC with or without CNS and PNS involvement (see Supp. Information). Our findings suggest including *ZC4H2* analysis retrospectively in AMC male and female cohorts and prospectively in prenatal and/or neonatal genetic diagnostic tests in male and female fetuses presenting with fetal hypo-/akinesia and/or AMC e.g. (only) club foot/feet with or without hypogenitalism in males. Long-term prospective clinical studies assessing the development of the ZARD phenotypes and investigations determining the molecular and cellular mechanisms underlying the disorders are required.

Acknowledgments

The authors would like to thank the individuals and their families who participated in this study. We also thank Andrew Green, Christina Fagerberg and Melissa Lees for giving permission to include publicly available DECIPHER entries on their patients. The DDD study presents independent research commissioned by the Health Innovation Challenge Fund [grant number HICF-1009-003], a parallel funding partnership between Wellcome and the Department of Health, and the Wellcome Sanger Institute [grant number WT098051]. The views expressed in this publication are those of the author(s) and not necessarily those of Wellcome or the Department of Health. The study has UK Research Ethics Committee approval (10/H0305/83, granted by the Cambridge South REC, and

GEN/284/12 granted by the Republic of Ireland REC). This study makes use of data generated by the DECIPHER community. A full list of centres who contributed to the generation of the data is available from http://decipher.sanger.ac.uk and via email from decipher@sanger.ac.uk. Funding for the project was provided by the Wellcome Trust. For one of the families this research was made possible through access to the data and findings generated by the 100,000 Genomes Project. The 100,000 Genomes Project is managed by Genomics England Limited (a wholly owned company of the Department of Health). The 100,000 Genomes Project is funded by the National Institute for Health Research and NHS England. The Wellcome Trust, Cancer Research UK and the Medical Research Council have also funded research infrastructure. The 100,000 Genomes Project uses data provided by patients and collected by the National Health Service as part of their care and support. Part of this study was supported by a Dutch NWO VENI grant (OND1312421 to S.F.), NEI grant R01EY027421 and NHLBI grant X01HL132377 (to ECE). The Broad Center for Mendelian Genomics (UM1 HG008900) is funded by the National Human Genome Research Institute with supplemental funding provided by the National Heart, Lung, and Blood Institute under the Trans-Omics for Precision Medicine (TOPMed) program and the National Eye Institute.

References

Allen, R. C., Zoghbi, H. Y., Moseley, A. B., Rosenblatt, H. M., & Belmont, J. W. (1992). Methylation of HpaII and HhaI sites near the polymorphic CAG repeat in the human androgen-receptor gene correlates with X chromosome inactivation. *Am J Hum Genet, 51*(6), 1229-1239.

Arachchi, H., Wojcik, M. H., Weisburd, B., Jacobsen, J. O. B., Valkanas, E., Baxter, S., . . . Rehm, H. L. (2018). matchbox: An open-source tool for patient matching via the Matchmaker Exchange. *Hum Mutat, 39*(12), 1827-1834. doi:10.1002/humu.23655

Bienvenu, T., Poirier, K., Friocourt, G., Bahi, N., Beaumont, D., Fauchereau, F., . . . Chelly, J. (2002). ARX, a novel Prd-class-homeobox gene highly expressed in the telencephalon, is mutated in X-linked mental retardation. *Hum Mol Genet, 11*(8), 981-991.

Carrel, L., & Willard, H. F. (1996). An assay for X inactivation based on differential methylation at the fragile X locus, FMR1. *Am J Med Genet, 64*(1), 27-30. doi:10.1002/(SICI)1096-8628(19960712)64:1<27::AID-AJMG3>3.0.CO;2-O

Choi, Y., & Chan, A. P. (2015). PROVEAN web server: a tool to predict the functional effect of amino acid substitutions and indels. *Bioinformatics, 31*(16), 2745-2747. doi:10.1093/bioinformatics/btv195

Dikow, N., Granzow, M., Graul-Neumann, L. M., Karch, S., Hinderhofer, K., Paramasivam, N., . . . Moog, U. (2017). DDX3X mutations in two girls with a phenotype overlapping Toriello-Carey syndrome. *Am J Med Genet A, 173*(5), 1369-1373. doi:10.1002/ajmg.a.38164

Ewans, L. J., Field, M., Zhu, Y., Turner, G., Leffler, M., Dinger, M. E., . . . Shoubridge, C. (2017). Gonadal mosaicism of a novel IQSEC. 2 variant

causing female limited intellectual disability and epilepsy. *Eur J Hum Genet, 25*(6), 763-767. doi:10.1038/ejhg.2017.29

Godfrey, N. D., Dowlatshahi, S., Martin, M. M., & Rothkopf, D. M. (2018). Wieacker-Wolff syndrome with associated cleft palate in a female case. *Am J Med Genet A, 176*(1), 167-170. doi:10.1002/ajmg.a.38527

Hall, J. G. (1984). An approach to research on congenital contractures. *Birth Defects Orig Artic Ser, 20*(6), 8-30.

Hall, J. G. (2009). Pena-Shokeir phenotype (fetal akinesia deformation sequence) revisited. *Birth Defects Res A Clin Mol Teratol, 85*(8), 677-694. doi:10.1002/bdra.20611

Hall, J. G. (2014). Arthrogryposis (multiple congenital contractures): diagnostic approach to etiology, classification, genetics, and general principles. *Eur J Med Genet, 57*(8), 464-472. doi:10.1016/j.ejmg.2014.03.008

Hall, J. G., Agranovich, O., Ponten, E., & van Bosse, H. J. (2015). Summary of the 2nd International Symposium on Arthrogryposis, St. Petersburg, Russia, September 17-19, 2014. *Am J Med Genet A, 167*(6), 1193-1195. doi:10.1002/ajmg.a.36938

Hennekam, R. C., Barth, P. G., Van Lookeren Campagne, W., De Visser, M., & Dingemans, K. P. (1991). A family with severe X-linked arthrogryposis. *Eur J Pediatr, 150*(9), 656-660.

Hirata, H., Nanda, I., van Riesen, A., McMichael, G., Hu, H., Hambrock, M., . . . Kalscheuer, V. M. (2013). ZC4H2 mutations are associated with arthrogryposis multiplex congenita and intellectual disability through impairment of central and peripheral synaptic plasticity. *Am J Hum Genet, 92*(5), 681-695. doi:10.1016/j.ajhg.2013.03.021

Holman, S. K., Morgan, T., Baujat, G., Cormier-Daire, V., Cho, T. J., Lees, M., . . . Robertson, S. P. (2013). Osteopathia striata congenita with cranial sclerosis and intellectual disability due to contiguous gene deletions involving the WTX locus. *Clin Genet, 83*(3), 251-256. doi:10.1111/j.1399-0004.2012.01905.x

Hu, H., Haas, S. A., Chelly, J., Van Esch, H., Raynaud, M., de Brouwer, A. P., . . . Kalscheuer, V. M. (2016). X-exome sequencing of 405 unresolved families identifies seven novel intellectual disability genes. *Mol Psychiatry, 21*(1), 133-148. doi:10.1038/mp.2014.193

Hunter, J. M., Kiefer, J., Balak, C. D., Jooma, S., Ahearn, M. E., Hall, J. G., & Baumbach-Reardon, L. (2015). Review of X-linked syndromes with arthrogryposis or early contractures-aid to diagnosis and pathway identification. *Am J Med Genet A, 167A*(5), 931-973. doi:10.1002/ajmg.a.36934

Kaiser, F. J., Ansari, M., Braunholz, D., Concepcion Gil-Rodriguez, M., Decroos, C., Wilde, J. J., . . . Deardorff, M. A. (2014). Loss-of-function HDAC8 mutations cause a phenotypic spectrum of Cornelia de Lange syndrome-

like features, ocular hypertelorism, large fontanelle and X-linked inheritance. *Hum Mol Genet, 23*(11), 2888-2900. doi:10.1093/hmg/ddu002

Kircher, M., Witten, D. M., Jain, P., O'Roak, B. J., Cooper, G. M., & Shendure, J. (2014). A general framework for estimating the relative pathogenicity of human genetic variants. *Nat Genet, 46*(3), 310-315. doi:10.1038/ng.2892

Kondo, D., Noguchi, A., Takahashi, I., Kubota, H., Yano, T., Sato, Y., . . . Takahashi, T. (2018). A novel ZC4H2 gene mutation, K209N, in Japanese siblings with arthrogryposis multiplex congenita and intellectual disability: characterization of the K209N mutation and clinical findings. *Brain Dev, 40*(9), 760-767. doi:10.1016/j.braindev.2018.05.003

Lek, M., Karczewski, K. J., Minikel, E. V., Samocha, K. E., Banks, E., Fennell, T., . . . Exome Aggregation, C. (2016). Analysis of protein-coding genetic variation in 60,706 humans. *Nature, 536*(7616), 285-291. doi:10.1038/nature19057

Lorenzo, M., Stolte-Dijkstra, I., van Rheenen, P., Smith, R. G., Scheers, T., & Walia, J. S. (2018). Clinical spectrum of KIAA2022 pathogenic variants in males: Case report of two boys with KIAA2022 pathogenic variants and review of the literature. *Am J Med Genet A, 176*(6), 1455-1462. doi:10.1002/ajmg.a.38667

Lower, K. M., Turner, G., Kerr, B. A., Mathews, K. D., Shaw, M. A., Gedeon, A. K., . . . Gecz, J. (2002). Mutations in PHF6 are associated with Borjeson-

Forssman-Lehmann syndrome. *Nat Genet, 32*(4), 661-665. doi:10.1038/ng1040

Mattiske, T., Moey, C., Vissers, L. E., Thorne, N., Georgeson, P., Bakshi, M., & Shoubridge, C. (2017). An Emerging Female Phenotype with Loss-of-Function Mutations in the Aristaless-Related Homeodomain Transcription Factor ARX. *Hum Mutat, 38*(5), 548-555. doi:10.1002/humu.23190

May, M., Hwang, K. S., Miles, J., Williams, C., Niranjan, T., Kahler, S. G., . . . Kim, C. H. (2015). ZC4H2, an XLID gene, is required for the generation of a specific subset of CNS interneurons. *Hum Mol Genet, 24*(17), 4848-4861. doi:10.1093/hmg/ddv208

Miles, J. H., & Carpenter, N. J. (1991). Unique X-linked mental retardation syndrome with fingertip arches and contractures linked to Xq21.31. *Am J Med Genet, 38*(2-3), 215-223.

Moog, U., Kutsche, K., Kortum, F., Chilian, B., Bierhals, T., Apeshiotis, N., . . . Uyanik, G. (2011). Phenotypic spectrum associated with CASK loss-of-function mutations. *J Med Genet, 48*(11), 741-751. doi:10.1136/jmedgenet-2011-100218

Neveling, K., Feenstra, I., Gilissen, C., Hoefsloot, L. H., Kamsteeg, E. J., Mensenkamp, A. R., . . . Nelen, M. R. (2013). A post-hoc comparison of the utility of sanger sequencing and exome sequencing for the diagnosis of heterogeneous diseases. *Hum Mutat, 34*(12), 1721-1726. doi:10.1002/humu.22450

O'Rawe, J. A., Wu, Y., Dorfel, M. J., Rope, A. F., Au, P. Y., Parboosingh, J. S., . . . Lyon, G. J. (2015). TAF1 Variants Are Associated with Dysmorphic Features, Intellectual Disability, and Neurological Manifestations. *Am J Hum Genet, 97*(6), 922-932. doi:10.1016/j.ajhg.2015.11.005

Okubo, Y., Endo, W., Inui, T., Suzuki-Muromoto, S., Miyabayashi, T., Togashi, N., . . . Haginoya, K. (2018). A severe female case of arthrogryposis multiplex congenita with brain atrophy, spastic quadriplegia and intellectual disability caused by ZC4H2 mutation. *Brain Dev, 40*(4), 334-338. doi:10.1016/j.braindev.2017.11.011

Palmer, E. E., Stuhlmann, T., Weinert, S., Haan, E., Van Esch, H., Holvoet, M., . . . Kalscheuer, V. M. (2018). De novo and inherited mutations in the X-linked gene CLCN4 are associated with syndromic intellectual disability and behavior and seizure disorders in males and females. *Mol Psychiatry, 23*(2), 222-230. doi:10.1038/mp.2016.135

Philippakis, A. A., Azzariti, D. R., Beltran, S., Brookes, A. J., Brownstein, C. A., Brudno, M., . . . Rehm, H. L. (2015). The Matchmaker Exchange: a platform for rare disease gene discovery. *Hum Mutat, 36*(10), 915-921. doi:10.1002/humu.22858

Schara, U., Kress, W., Tucke, J., & Mortier, W. (2003). X-linked myotubular myopathy in a female infant caused by a new MTM1 gene mutation. *Neurology, 60*(8), 1363-1365.

Shoubridge, C., Tarpey, P. S., Abidi, F., Ramsden, S. L., Rujirabanjerd, S., Murphy, J. A., . . . Gecz, J. (2010). Mutations in the guanine nucleotide exchange factor gene IQSEC. 2 cause nonsyndromic intellectual disability. *Nat Genet, 42*(6), 486-488. doi:10.1038/ng.588

Shvetsova, E., Sofronova, A., Monajemi, R., Gagalova, K., Draisma, H. H. M., White, S. J., . . . Go, N. L. c. (2018). Skewed X-inactivation is common in the general female population. *Eur J Hum Genet*. doi:10.1038/s41431-018-0291-3

Snijders Blok, L., Madsen, E., Juusola, J., Gilissen, C., Baralle, D., Reijnders, M. R., . . . Kleefstra, T. (2015). Mutations in DDX3X Are a Common Cause of Unexplained Intellectual Disability with Gender-Specific Effects on Wnt Signaling. *Am J Hum Genet, 97*(2), 343-352. doi:10.1016/j.ajhg.2015.07.004

Sobreira, N. L. M., Arachchi, H., Buske, O. J., Chong, J. X., Hutton, B., Foreman, J., . . . Matchmaker Exchange, C. (2017). Matchmaker Exchange. *Curr Protoc Hum Genet, 95*, 9 31 31-39 31 15. doi:10.1002/cphg.50

Steinmuller, R., Lantigua-Cruz, A., Garcia-Garcia, R., Kostrzewa, M., Steinberger, D., & Muller, U. (1997). Evidence of a third locus in X-linked recessive spastic paraplegia. *Hum Genet, 100*(2), 287-289.

Stromme, P., Mangelsdorf, M. E., Shaw, M. A., Lower, K. M., Lewis, S. M., Bruyere, H., . . . Gecz, J. (2002). Mutations in the human ortholog of

Aristaless cause X-linked mental retardation and epilepsy. *Nat Genet, 30*(4), 441-445. doi:10.1038/ng862

Van Maldergem, L., Hou, Q., Kalscheuer, V. M., Rio, M., Doco-Fenzy, M., Medeira, A., . . . Man, H. Y. (2013). Loss of function of KIAA2022 causes mild to severe intellectual disability with an autism spectrum disorder and impairs neurite outgrowth. *Hum Mol Genet, 22*(16), 3306-3314. doi:10.1093/hmg/ddt187

Wieacker, P., Wolff, G., Wienker, T. F., & Sauer, M. (1985). A new X-linked syndrome with muscle atrophy, congenital contractures, and oculomotor apraxia. *Am J Med Genet, 20*(4), 597-606. doi:10.1002/ajmg.1320200405

Wright, C. F., Fitzgerald, T. W., Jones, W. D., Clayton, S., McRae, J. F., van Kogelenberg, M., . . . study, D. D. D. (2015). Genetic diagnosis of developmental disorders in the DDD study: a scalable analysis of genome-wide research data. *Lancet, 385*(9975), 1305-1314. doi:10.1016/S0140-6736(14)61705-0

Zanzottera, C., Milani, D., Alfei, E., Rizzo, A., D'Arrigo, S., Esposito, S., & Pantaleoni, C. (2017). ZC4H2 deletions can cause severe phenotype in female carriers. *Am J Med Genet A, 173*(5), 1358-1363. doi:10.1002/ajmg.a.38155

Zerem, A., Haginoya, K., Lev, D., Blumkin, L., Kivity, S., Linder, I., . . . Lerman-Sagie, T. (2016). The molecular and phenotypic spectrum of IQSEC. 2-related epilepsy. *Epilepsia, 57*(11), 1858-1869. doi:10.1111/epi.13560

Zweier, C., Kraus, C., Brueton, L., Cole, T., Degenhardt, F., Engels, H., . . . Wieczorek, D. (2013). A new face of Borjeson-Forssman-Lehmann syndrome? De novo mutations in PHF6 in seven females with a distinct phenotype. *J Med Genet, 50*(12), 838-847. doi:10.1136/jmedgenet-2013-101918

Figures

Figure 1. Pedigrees of the families with pathogenic *ZC4H2* variants discovered by candidate gene approach, whole exome/genome sequencing and array-CGH. Families 1, 4-6, 8, 9, 15, 18, 19 and 24 with likely pathogenic *ZC4H2* missense variants, famillies 2, 3, 13, 14, 16, 17, 21, and 22 with splicing, frameshift and stop codon variants and families 10-12, 20, and 23 with a microdeletion removing the 5' part of *ZC4H2*, Fam, family; *, variant present; wt, wild-type. For family7 see Supp. Figure S2.

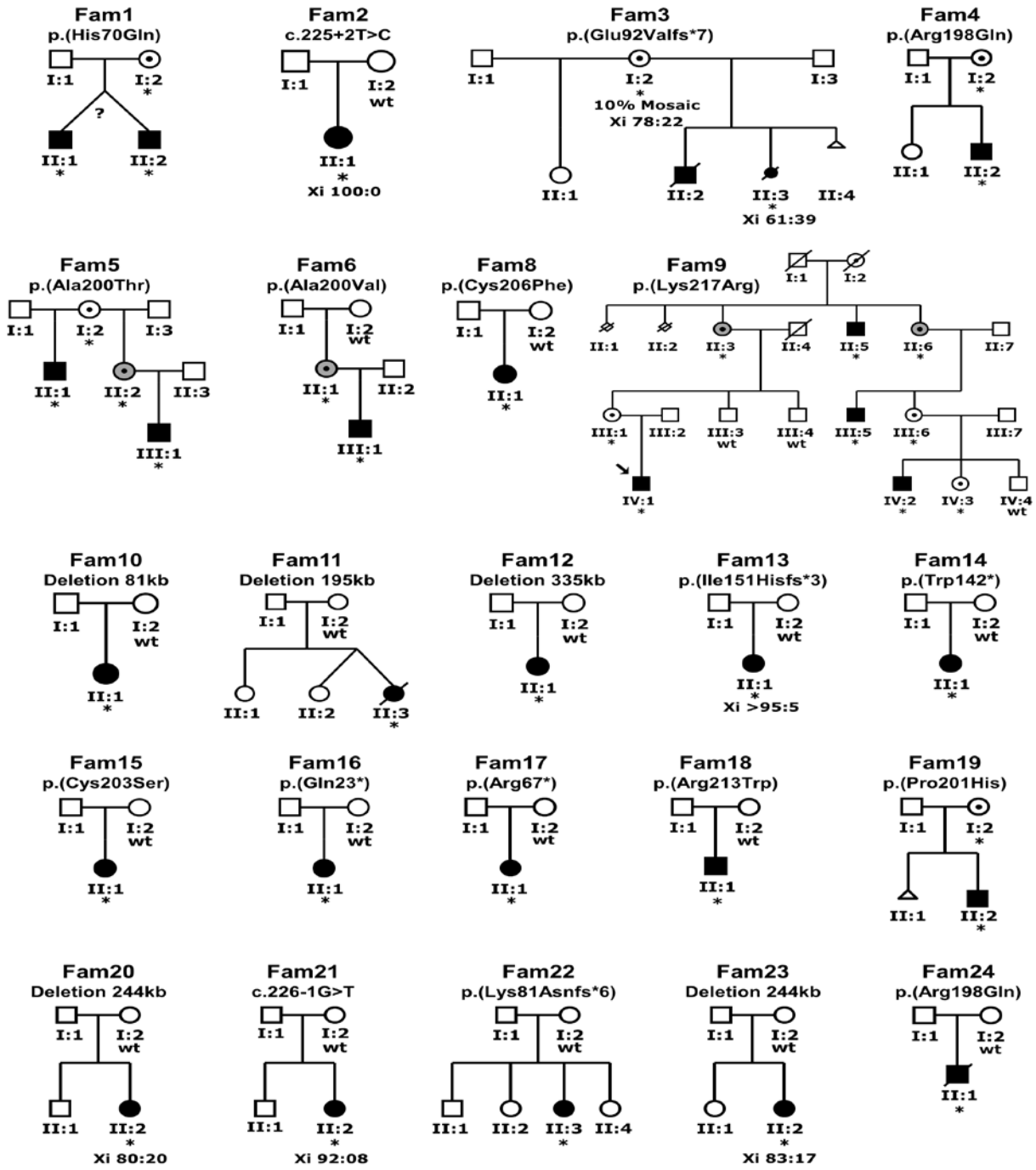
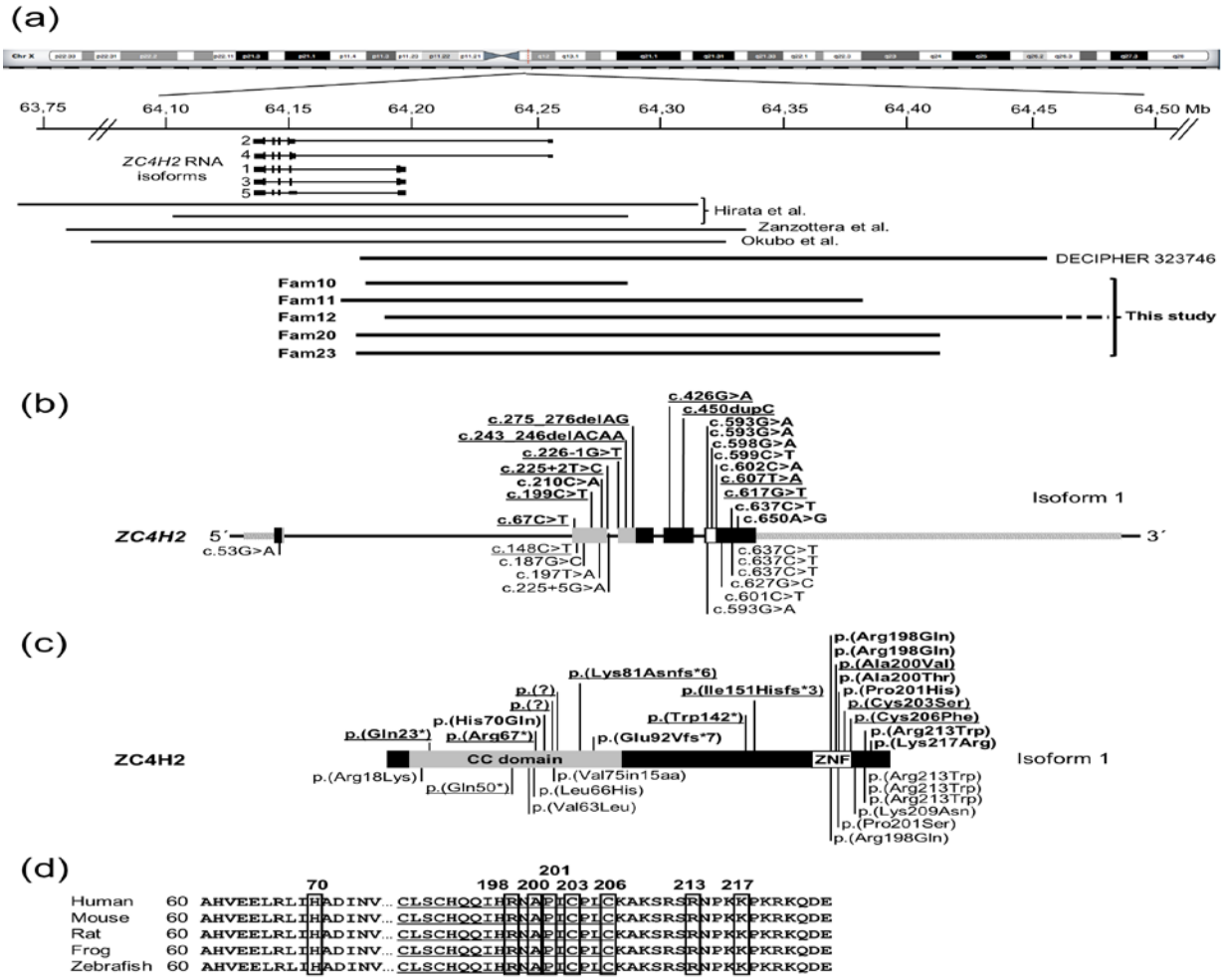


Figure 2. Overview of genetic results in ZARD affected males and females. An overview of newly identified likely pathogenic variants of *ZC4H2* described here and previously reported in the literature. (a) Schematic view of the Xchromosome with the location of *ZC4H2* and the five known RNA isoforms. Horizontal bars indicate the *de novo* complete *ZC4H2* deletions in affected females published previously (Hirata et al., 2013; Okubo et al., 2018; Zanzottera et al., 2017), DECIPHER case 323746 and the newly identified *de novo ZC4H2* microdeletions removing part of *ZC4H2* identified in this study (families 10-12, 20, and 23). All deletions removed *ZC4H2* exon 1 but no other known gene. (b) The structure of human *ZC4H2* (Isoform 1; NM_018684.3) with non-translated sequences grey-striped, exons encoding the coiled-coil domain in grey and exons encoding the zinc finger domain in white. Newly identified pathogenic *ZC4H2* variants identified in this study are depicted above the gene with *de novo* variants in affected females underlined. Likely pathogenic variants from the literature are depicted below the gene and protein. (c) Schematic representation of ZC4H2 protein isoform 1 (NP_061154.1) with its functional domains (CC domain: Coiled coil domain; ZNF: zinc finger domain). ZC4H2 variants newly identified in this study are depicted above the protein with *de novo* vari-

ants in females underlined and variants from the literature are depicted below the protein.

(d) Multiple sequence alignment of ZC4H2 protein sequences in five species showing 100% conservation of the newly identified mutated amino acids (boxed). Amino acids of the ZC4H2 zinc-finger domain in the C-terminal part of the protein are underlined.



Table

Table 1. Summary of the variable ZARD clinical phenotype present in 30% or more of affected males and females from 42 families.

ZARD			MALES TOTAL			FEMALES						FEMALES TOTAL			MALES & FEMALES TOTAL		
ZC4H2 alterations			*De novo* and inherited			*De novo*			Inherited			*De novo* and inherited			*De novo* and inherited		
Category	Features	HPO	TOTAL positive informative	TOTAL informative	% *	TOTAL positive informative	TOTAL informative	% *	TOTAL positive informative	TOTAL informative	% *	TOTAL positive informative	TOTAL informative	% *	TOTAL positive informative	TOTAL informative	% *
	Maximum informative		44	48	92%	23	25	92%	20	32	63%	43	57	75%	87	105	83%
Growth	Short stature <3rd percentile	43222	19	37	51%	10	18	56%	2	16	13%	12	34	35%	31	71	44%
	Postnatal growth retar-	88897	7	12	58%	10	14	71%	0	1		10	15	67%	17	27	63%

	dation																
Head and neck	Microcephaly <3rd percentile	252	16	26	62%	3	15	20%	1	15	7%	4	30	13%	20	56	36%
	Facial weakness-palsy	10628	11	22	50%	9	14	64%	0	3		9	17	53%	20	39	51%
	High forehead	348	6	11	55%	7	12	58%	1	2	50%	8	14	57%	14	25	56%
	Low-set ears	369	12	19	63%	8	15	53%	1	9	11%	9	24	38%	21	43	49%
	Low-set, posteriorly rotated ears	368	8	17	47%	8	15	53%	0	2		8	17	47%	16	34	47%
	Upslanting palpebral fissures	582	10	21	48%	1	12	8%	0	1		1	13	8%	11	34	32%
	Ptosis	14888	26	44	59%	4	14	29%	9	22	41%	13	36	36%	39	80	49%

	Deeply set eyes	490	6	11	55%	5	13	38%	0	1		5	14	36%	11	25	44%
	Ocular motor apraxia	657	8	12	67%	4	10	40%	1	2	50%	5	12	42%	13	24	54%
	Strabismus	486	13	25	52%	12	16	75%	5	14	36%	17	30	57%	30	55	55%
	Antevertednares	463	13	22	59%	9	13	69%	3	9	33%	12	22	55%	25	44	57%
	Microretrognathia	308	11	20	55%	14	17	82%	1	9	11%	15	26	58%	26	46	57%
	Long (flat) philtrum	343	17	40	43%	6	14	43%	12	18	67%	18	32	56%	35	72	49%
	Broad alveolar ridges	187	12	21	57%	2	9	22%	2	9	22%	4	18	22%	16	39	41%
	High-arched palate	218	16	33	48%	3	10	30%	3	17	18%	6	27	22%	22	60	37%
	U-shaped	108	20	42	48%	5	13	38%	0	11		5	24	21%	25	66	38%

	upper lip ver-mil-ion - carps hape d mout h	06															
	Cleft pal-ate	175	1	8	13%	7	17	41%	0	2		7	19	37%	8	27	30%
	Dow nturn ed cor-ners of mout h	2714	4	11	36%	9	14	64%	0	2		9	16	56%	13	27	48%
	Short neck (with lim-ited rota-tion)	470	12	37	32%	12	14	86%	1	19	5%	13	33	39%	25	70	36%
Res pir ato ry	Neo-natal res-pira-tory dis-tress	2643	16	34	47%	5	16	31%	0	13		5	29	17%	21	63	33%
	Re-cur-rent aspi-ration pneu mo-nia	2100	3	9	33%	5	11	45%	0	3		5	14	36%	8	23	35%

	Apnea	2104	11	16	69%	3	11	27%	0	3		3	14	21%	14	30	47%
Chest	Limited shoulder movement	6467	7	16	44%	12	14	86%	1	2	50%	13	16	81%	20	32	63%
	Narrow chest - narrow shoulders-thorax	774	18	36	50%	11	15	73%	5	20	25%	16	35	46%	34	71	48%
Abdomen	Feeding difficulties - poor feeding in infancy	8872	19	30	63%	12	17	71%	2	13	15%	14	30	47%	33	60	55%
	Enco copresis - bowel incontinence	40183	6	7	86%	5	12	42%	2	9	22%	7	21	33%	13	28	46%
Genitouri	Micropenis	54	7	12	58%	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	7	12	58%

nary	Cryptorchidism	28	9	17	53%	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	9	17	53%
	Urinary incontinence	20	2	4	50%	9	12	75%	5	12	42%	14	24	58%	16	28	57%
Skeletal	Cervical kyphosis	2947	20	40	50%	7	18	39%	5	19	26%	12	37	32%	32	77	42%
	Thoracic kyphosis	2942	19	40	48%	7	18	39%	5	18	28%	12	36	33%	31	76	41%
	Scoliosis	2650	22	40	55%	7	19	37%	5	18	28%	12	37	32%	34	77	44%
	Hip contracture	3273	15	32	47%	15	19	79%	2	21	10%	17	40	43%	32	72	44%
	Congenital hip dislocations - subluxation	1374	14	32	44%	11	19	58%	2	21	10%	13	40	33%	27	72	38%
	Short limbs	98	7	17	41	3	14	21	6	10	60	9	24	38%	16	41	39

		26			%			%			%						%
	Arthrogrypo-sis multiplex con-genita	28004	31	45	69%	23	25	92%	8	18	44%	31	43	72%	62	88	70%
	Knee flex-ion con-tracture	63800	26	38	68%	13	19	68%	10	21	48%	23	40	58%	49	78	63%
	El-bow flex-ion con-tractures	29887	19	35	54%	10	17	59%	10	21	48%	20	38	53%	39	73	53%
	Wrist con-tractures	12399	4	6	67%	11	18	61%	0	3		11	21	52%	15	27	56%
	Prox-imally place-ment of thumb	96223	5	19	26%	7	17	41%	0	3		7	20	35%	12	39	31%
	Met-acar-poph alan-geal joint	60700	2	5	40%	10	16	63%	0	3		10	19	53%	12	24	50%

	con-tractures																
	Camp-to-dac-tyly	12385	20	33	61%	16	19	84%	12	24	50%	28	43	65%	48	76	63%
	Ulnar devi-ation of any fin-ger	94665	15	32	47%	11	18	61%	2	20	10%	13	38	34%	28	70	40%
	Radi-al devi-ation of any fin-ger	94666	6	9	67%	12	18	67%	0	3		12	21	57%	18	30	60%
	Over-lap-ping toe(s)	18445	4	10	40%	4	14	29%	0	1		4	15	27%	8	25	32%
	Prox-imal-ly place-d toes	17800	9	21	43%	5	14	36%	0	3		5	17	29%	14	38	37%
	Rock-er bot-tom feet	18388	13	18	72%	9	13	69%	0	12		9	25	36%	22	43	51%
	Equi-no-varus de-	81110	34	45	76%	14	20	70%	5	23	22%	19	43	44%	53	88	60%

	formity - club feet																
	Achilles tendon contracture	1771	21	28	75%	5	13	38%	0	3		5	16	31%	26	44	59%
Skin, nails, hair	High anterior hairline	9890	14	21	67%	8	14	57%	0	2		8	16	50%	22	37	59%
Muscle, soft tissue	Distal muscle weakness	2460	29	36	81%	18	20	90%	8	23	35%	26	43	60%	55	79	70%
	Distal limb muscle atrophy	3693	15	18	83%	13	17	76%	0	3		13	20	65%	28	38	74%
	Edema of the dorsum of hands and feet	7514	8	19	42%	8	15	53%	2	11	18%	10	26	38%	18	45	40%
Neurolo-	Motor delay	127	44	48	92%	19	21	90%	3	12	25%	22	33	67%	66	81	81%

gy		0															
	Inability to walk	2540	15	18	83%	17	18	94%	0	3		17	21	81%	32	39	82%
	Generalized hypotonia	1290	15	29	52%	7	14	50%	1	13	8%	8	27	30%	23	56	41%
	Intellectual disability	1249	44	48	92%	16	20	80%	20	32	63%	36	52	69%	80	100	80%
	Dysarthria, deficit in expressive language	1260	11	13	85%	7	10	70%	3	9	33%	10	19	53%	21	32	66%
	Poor speech	2465	24	28	86%	11	15	73%	2	5	40%	13	20	65%	37	48	77%
	Absent speech	1344	11	15	73%	3	16	19%	0	5		3	21	14%	14	36	39%
	Drooling	2307	26	33	79%	8	14	57%	1	19	5%	9	33	27%	35	66	53%
	Dysphagi	20	9	10	90	8	12	67	0	2		8	14	57%	17	24	71

	a	15			%			%								%
	Chewing difficulties including oral motor dysfunction	5216	12	16	75%	6	9	67%	2	11	18%	8	20	40%	20	36	56%
	Spasticity	1257	27	35	77%	12	17	71%	1	15	7%	13	32	41%	40	67	60%
	Seizures	1250	21	42	50%	4	14	29%	2	14	14%	6	28	21%	27	70	39%
	Delayed CNS myelination	2188	5	19	26%	4	11	36%	0	0		4	11	36%	9	30	30%
	Global brain atrophy	2283	8	18	44%	3	13	23%	0	0		3	13	23%	11	31	35%
	Ventriculomegaly	2119	3	7	43%	5	13	38%	0	0		5	13	38%	8	20	40%
	Hy-	1	18	26	6	6	9	6	0	11		6	20	30	24	46	5

	perre re-flexia	34 7			9 %			7 %						%			2 %
	Emo-tional labil-ity	71 2	4	6	67 %	1	7	14 %	0	1		1	8	13 %	5	14	36 %

ZARD-related clinical features are listed according to the nomenclature/systematics of the OMIM “Clinical synopsis” and were mapped to the Human Phenotype Ontology (HPO). Only positive informative clinical features were scored. Additional less frequent (<30%) ZARD-related clinical features noticed in this study and reported in the literature are listed in Supp. Table S2. Abbreviations: nr, number; y, years; w, weeks; HPO, Human Phenotype Ontology; %, percentage; *, positive informative / total informative; n/a, not applicable.