Inactivation of *AMMECR1* is associated with growth, bone and heart alterations

Running Title: *AMMECR1* loss-of-function

Mariana Moysés-Oliveira[1,2], Giuliana Giannuzzi[2], Richard J. Fish[3], Jill A. Rosenfeld[4], Florence Petit[5], Maria de Fatima Soares[6], Leslie Domenici Kulikowski[7], Adriana Di Battista[1], Malú Zamariolli[1], Fan Xia[4], Thomas Liehr[8], Nadezda Kosyakova[8], Gianna Carvalheira[1], Michael Parker[9], Eleanor G. Seaby[10], Sarah Ennis[10], Rodney D. Gilbert[11], R. Tanner Hagelstrom[12], Maria L. Cremona[12], Wenhui L. Li[12], Alka Malhotra[12], Anjana Chandrasekhar[12], Denise L. Perry[12], Ryan J. Taft[12], Julie McCarrier[13], Donald G. Basel[13], Joris Andrieux[14], Taiza Stumpp[15], Fernanda Antunes[16], Gustavo José Pereira[16], Marguerite Neerman-Arbez[3], Vera Ayres Meloni[1], Margaret Drummond-Borg[17], Maria Isabel Melaragno[1], Alexandre Reymond[2]

[1]Genetics Division, Department of Morphology and Genetics, Universidade Federal de São Paulo, 04023-900 São Paulo, Brazil

[2]Center for Integrative Genomics, University of Lausanne, 1015 Lausanne, Switzerland

[3]Department of Genetic Medicine and Development, 1211 University of Geneva Medical School, Geneva, Switzerland

[4]Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX 77030, USA

[5]Clinique de Génétique, CHU Lille - Hôpital Jeanne de Flandre, 59000 Lille, France

[6]Psychobiology Department, Universidade Federal de São Paulo, 04023-062 São Paulo, Brazil

[7]Department of Pathology, Laboratório de Citogenômica, LIM 03, Hospital das Clínicas, Faculdade de Medicina, Universidade de São Paulo, 05403-000 São Paulo, Brazil

[8]Universitätsklinikum Jena, Institut für Humangenetik, D-07740 Jena, Germany

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.23373.

[9]Sheffield Clinical Genetics Service, Sheffield Children's Hospital, Western Bank, S10 2TH Sheffield, United Kingdom

[10]Genomic Informatics Group, University Hospital Southampton, SO16 6YD Southampton, United Kingdom

[11]Southampton Children's Hospital, University Hospital Southampton, SO16 6HU Southampton, United Kingdom

[12]Illumina Clinical Services Laboratory, 5200 Illumina Way, San Diego CA 92122, USA

[13]Department of Pediatrics, Section of Genetics, Medical College of Wisconsin, Milwaukee, USA

[14]Institut de Génétique Médicale, CHU Lille - Hôpital Jeanne de Flandre, 59000 Lille, France

[15]Developmental Biology Division, Department of Morphology and Genetics, Universidade Federal de São Paulo, 04023-900 São Paulo, Brazil

[16]Department of Pharmacology, Universidade Federal de São Paulo, 04023-900, São Paulo, Brazil

[17]Cook Children's Genetic Clinic, Fort Worth, TX 76102, USA

Correspondence to

alexandre.reymond@unil.ch

melaragno.maria@unifesp.br

Summary

We report five individuals with loss-of-function of the X-linked *AMMECR1*: a girl with a balanced X-autosome translocation and inactivation of the normal X-chromosome; two boys with maternally-inherited and *de novo* nonsense variants; and two half-brothers with maternally inherited microdeletion variants. They present with short stature, cardiac and skeletal abnormalities and hearing loss. Variants of unknown significance in *AMMECR1* in four male patients from two families with

partially overlapping phenotypes were previously reported. *AMMECR1* is co-expressed with genes implicated in cell cycle regulation, five of which were previously associated with growth and bone alterations. Our knockdown of the zebrafish orthologous gene resulted in phenotypes reminiscent of patients' features. The increased transcript and encoded protein levels of *AMMECR1L*, an *AMMECR1* paralog, in the t(X;9) patient's cells indicate a possible partial compensatory mechanism. AMMECR1 and AMMECR1L proteins dimerize and localize to the nucleus as suggested by their nucleic acid-binding RAGNYA folds. Our results suggest that *AMMECR1* is potentially involved in cell cycle control and linked to a new syndrome with growth, bone, heart and kidney alterations with or without elliptocytosis.

Key words: Growth delay, bone dysplasia, heart alteration, *AMMECR1*, X-linked disease.

INTRODUCTION

AMMECR1 maps within the AMME (Alport syndrome, mental retardation, midface hypoplasia, and elliptocytosis) complex interval. AMME (MIM# 300194) is an Xq22.3 contiguous gene deletion syndrome encompassing about 20 genes (**Supp. Figure S1**) characterized by nephropathy (Alport syndrome, MIM# 301050), intellectual disability, midface hypoplasia and elliptocytosis. Some affected individuals also present with hearing loss, ocular abnormalities, limb malformations, cardiac arrhythmia and heart defects (Jonsson et al., 1998). The *COL4A5* gene has been implicated in Alport syndrome, hearing loss and eye abnormalities (Vaser et al., 2016; Vitelli et al., 1999), whereas variants in *ACSL4* were associated with X-linked intellectual disability 63 (MRX63; MIM# 300387) (Gazou et al., 2013; Meloni et al., 2002). The other clinical features of AMME, such as midface hypoplasia, elliptocytosis, limb malformations, cardiac arrhythmia and heart defects, could be associated with the remaining neighboring genes, including *AMMECR1* (**Supp. Figure S1, Table 1; Supp. Table S3**).

A family with two affected male individuals, the patient and one of his maternal uncles, who had partially overlapping phenotypes, i.e. midface hypoplasia, proportionate short stature, hearing loss and elliptocytosis with or without anemia, was recently described (Basel-Vanagaite et al., 2017). The causative nature of the identified stop gain variant in *AMMECR1* that induces skipping of 111 bp-long

exon 2 is not clear. The investigated patient's X-chromosome carries 12 more rare exonic variants in particular a frameshift and a stop gain variant in *MAGEB3* (pLI = 0.66, 0 loss of function (lof) variants identified compared to 3.3 expected according to ExAC) and the lof intolerant *SLC16A2* (pLI = 0.93, 0 identified/8.4 expected) genes, respectively. Two half-brothers presenting elliptocytosis, midface hypoplasia, early speech and language delay carrying a p.(Gly177Asp) missense variant in *AMMECR1* that causes mislocalization of the encoded protein were also reported (Andreoletti et al., 2016). These four cases were regrouped by OMIM under the identifier MFHEIN (midface hypoplasia, hearing impairment, elliptocytosis, and nephrocalcinosis, MIM# 300990).

Here we describe four new males and a female with variants that are predicted to lead to lof of the X-linked *AMMECR1* gene and present with short stature, cardiac and skeletal abnormalities and hearing loss. The female carries a balanced X-autosome translocation disrupting *AMMECR1* and silencing of the normal X, whereas the males are hemizygous for a nonsense variant or a deletion encompassing *AMMECR1*. Our results support the involvement of *AMMECR1* in a new syndrome and expand its associated phenotypes that group growth abnormalities with alterations of bones and heart.

MATERIALS AND METHODS

Enrollment

All samples used in this study were collected with written appropriate informed consent and approval of the local ethics committee.

Classical and Molecular Cytogenetic Evaluation

550 resolution G-banding was performed on lymphocyte cultures using standard methods. Genomic array was performed using the Affymetrix Genome-Wide Human SNP-Array 6.0 or Human Genome CGH (comparative genome hybridization) array 4x44k (Agilent Technologies). For patient 1, breakpoint localization was performed by FISH (fluorescence *in situ* hybridization) and array painting as described (Liehr et al., 2002; Moyses-Oliveira et al., 2015) and validated by FISH using BAC (bacterial artificial chromosome) probe CTD-3066N24 (Thermo Fisher Scientific) and whole-genome sequencing. X-chromosome inactivation was assessed as described (Sisdelli et al., 2016).

Exome and whole genome sequencing and analysis

Patient 2 DNA was whole exome sequenced using the VCRome capture reagent as described (Yang et al., 2014). Variants with MAF<0.05% in controls cohorts (dbSNP, 1000 genome project, NHLBI GO Exome, ExAC (Exome Aggregation Consortium)) and predicted to be deleterious by at least one of the predictor SIFT (Vaser et al., 2016), PolyPhen-2 (Adzhubei et al., 2013) and/or UMD-Predictor (Salgado et al., 2016) were prioritized. Segregation of variants was confirmed by Sanger sequencing. Patient 3 and his parents DNA samples were prepared using the Illumina TruSeq PCR-free library prep and sequenced on a HiSeqX sequencer with paired end 150 bp reads. Reads were aligned to human reference genome assembly GRCh37 using ISACC aligner and variants were filtered using inheritance patterns, allele frequency, and molecular consequences. Variants were manually curated, and classified according to the American College of Medical Genetics and Genomics (ACMG) guidelines (Richards et al., 2015). The identified pathological variants were submitted to ClinVar.

Cell Lines, RNA and Quantitative RT-PCR

We established lymphoblastoid cell lines (LCLs) from the Brazilian patient (patient 1), her mother and ten unrelated Brazilian female controls, seven of which are carriers of balanced X-autosome translocations with different breakpoints (karyotypes described in **Supp. Table S1**). RNA was extracted using Trizol (Invitrogen) or PAXgene Blood RNA MDx Kit (Qiagen-Sciences). Gene expression analysis was performed by real-time quantitative PCR (RT-qPCR) using TaqMan assays (Life Technologies) and SYBR green primer pairs (**Supp. Table S2)** assayed in Applied Biosystems 7900HT Sequence Detection System (Life Technologies).

Western Blot

Fractionated protein extraction was performed as described (Settembre et al., 2011). 30 µg of protein were resolved on 10% SDS–PAGE and wet-transferred to PVDF membrane (GE Healthcare Life Sciences). Membranes were blocked with PBS 5% milk, 0.1% Tween 20 for 30 minutes at room temperature and then probed overnight at 4°C with anti-AMMECR1/AMMECR1L (Abcam 155780,

1:300), anti-AMMECR1 (Santa Cruz sc-131384, 1:300) or anti-GFP (Santa Cruz sc-8334, 1:500) antibodies. After three washes, membranes were incubated with horseradish peroxidase-conjugated secondary antibodies (Santa Cruz Biotechnology). Lamin A/C (Santa Cruz sc-6215) and tubulin (Sigma Aldrich T5168) were used as nuclear and cytoplasmic markers, respectively. Immunoreactive bands were visualized with Western highligting plus-ECL (Perkin Elmer).

Morpholinos and Zebrafish Embryo Manipulations

Zebrafish animal experimentation was approved by the Canton of Geneva Animal Experimentation Veterinary authority. Wild-type TU zebrafish were maintained in standard conditions (27^{o}C, 500 mS, pH 7.5). Embryos obtained by natural intercrossing were staged according to morphology (Kimmel et al., 1995).

To model a lof of *AMMECR1*, we designed three splice-blocking MO targeting the zebrafish *ammecr1* pre-mRNA: (1) MO-E1I1-5'-CATAGCCATCTTGCTACTACTTACT-3', (2) MO-I2E3-5'-GCACTGCCAAGGGAACATGAAGACA-3' and (3) MO-E3I3-5'-TGGAAATAGCTTCTTACCTCCCAGT-3' (Gene Tools) that target the exon 1/intron 1, intron 2/exon3 and exon 3/intron 3 boundaries, respectively. The standard Ctrl-MO (5'-CCTCTTACCTCAGTTACAATTTATA-3') (Gene Tools), with no targets in the zebrafish genome was used for mock injections. MOs were dissolved in nuclease-free water and their concentrations determined with a NanoDrop. They were injected at 1–2 cell stages (1–2 nL). For rescue experiments, the human wild-type *AMMECR1* mRNA was *in vitro* transcribed from a pCS2+*AMMECR1* plasmid using the SP6 mMessage mMachine System (Ambion) and purified by lithium chloride extraction. 200 pg of the human wild-type mRNAs were coinjected with the MOs.

Non-injected, Standard Control MO and ammecr1 MO-injected embryos were collected at 3 dpf and total RNA was isolated using standard Trizol protocol (Invitrogen). 850 ng of total RNA from each sample was used to synthesize cDNA with the Superscript III kit with Oligo d(T)$_{12\text{-}19}$ primers (Invitrogen). Dilutions of cDNA were used in qPCR reactions with SYBR Green PCR Master Mix (Life Technologies). All samples of cDNA were assayed in technical triplicate with standard PCR

conditions of the Applied Biosystems 7900HT Sequence Detection System (Life Technologies). ΔCt values were calculated relative to *ef1-alpha*. Relative quantification (RQ) of transcript levels were quantified using the $2^{-\Delta\Delta Ct}$ method.

Functional Network Analysis

We utilized GIANT (Genomescale Integrated Analysis of gene Networks in Tissues) to assess tissue-specific functional interactions (Greene et al., 2015). Enrichment of biological processes and cellular components was determined using Enrichr (Chen et al., 2013). The same analysis was repeated for *GAPDH* and *RGAG1* genes.

Expression of GFP-AMMECR1 in HeLa Cells

pcDNA™6.2/EmGFP Gateway destination vector was used to generate GFP-*AMMECR1* NM_015365/ENST00000262844 isoform construct with mammalian expression according to the instructions of the manufacturer (Invitrogen). The construct was introduced into the human HeLa cell line by transient transfection using FuGENE HD Transfection Reagent (Promega). 48 hours after transfection, cells were fixed with 4% paraformaldehyde, mounted using Vectashield with DAPI and observed using a Microscope Zeiss Imager A1 microscope.

Gene Expression in Early Mouse Embryos

Balb/c mice experimentation was approved by the local committee for animal care. Immediately after the postnatal females euthanasia, 9.5 GD embryos were collected and fine dissected using a needle under a stereoscope. Material from 35 embryos were pooled in five samples per tissue (prosencephalon, heart and somite). RNA was isolated using TRIzol reagent (Thermo Fisher Scientific). *Ammecr1* and *Ammecr1l* gene expression analysis was performed by RT-qPCR with TaqMan assays (Life Technologies) (**Supp. Table S2**) assayed in ViiA 7 Real-Time PCR System (Thermo Fisher Scientific).

RESULTS

Clinical descriptions

The phenotypes of the five affected individuals identified in this report are compared to previously published MFHEIN cases and AMME deletion syndrome patients in **Table 1** and **Supp. Table S3**. Patient 1 is the first and only child of non-consanguineous and healthy Brazilian parents. During the second trimester of the pregnancy, her mother presented partial placenta previa. She was born at 39 weeks with a birth weight of 2.705 kg (3^{rd}-10^{th} centile), length of 44 cm ($<3^{rd}$ centile) and head circumference of 33 cm (25^{th} centile). The patient spent 29 days in neonatal intensive care because of anemia, jaundice and sepsis. She presented with normal developmental milestones. At 2 years and 8 months of age, serum dosage of IGF-1 was 37.8 ng/ml (reference 51-303 ng/ml), TSH was 0.47 µUI/ml (reference 0.35-5.5 µUI/ml) and T4L was 1.34 ng/dl (reference value 0.8-1.4 ng/dl). At the age of six, the clinical evaluation revealed weight of 12.2 kg ($<3^{rd}$), short stature (97 cm, $<1^{st}$), head circumference of 44 cm ($<3^{rd}$), triangular and flat face, midface hypoplasia, prominent forehead, esotropia, cupid's bow, low-set ears, anteverted nares, short neck, bilateral clinodactyly of the fifth finger, bilateral cubitus valgus and pes planus and cardiac arrhythmia (**Figure 1; Supp. Figure S2; Table 1**; **Supp. Table S3**). At the same age, whole-skeleton X-ray revealed wormian bones in the skull, bone age delay, thoracolumbar scoliosis, diffuse bone demineralization with peripheral osteopenia showing looser zones in the first metatarsals and metacarpals bilaterally and also in iliac bones. The distal phalanges were cone shaped with irregularities, and bones of hands, feet and long bones presented hyperdense epiphyses and metaphyses (**Supp. Figure S2; Table 1**; **Supp. Table S3**). At seven years of age, a brain MRI showed no abnormalities, and neuropsychological evaluation using the WISC IV revealed normal intelligence (IQ 91). At the same age, GH replacement therapy was initiated, and the patient responded to the treatment. The echocardiogram showed an ostium primum atrial septal defect with thin aneurysmal septal tissue. Different medical examinations at ages 3 and 9 revealed tachycardia at rest. The immunological evaluation at eight years of age showed normal immunoglobulin levels with positive response to pneumococcal vaccine and normal lymphocyte numbers (CD3, CD4 and CD19). The central memory and peripheral CD8+T cells

were reduced (1% and 10%, respectively). Facial features coarsened with age. At ten years of age bilateral neurosensorial hearing loss for high frequencies was detected, and CT scan showed normal temporal bones.

Patient 2 was born in America. He presented with motor delay, speech delay, hypotonia, short stature and failure to thrive, dysmorphia, congenital heart disease with atrial septal defect, ureterocele and hypercalcemia, radioulnar synostosis, small penis, and failed hearing screen for high pitch (**Figure 1; Table 1; Supp. Table S3**). Patient 3 is of mixed African-American, Cuban, Polish and German ancestry. He presented with hypotonia, short stature and failure to thrive, dysmorphia included crowded midface, pectus excavatum, flared rib cage, chordee and hypospadia, and mild sensorineural hearing loss (**Figure 1; Table 1; Supp. Table S3**). The two remaining affected individuals are siblings of British origin. The older brother (patient 4) presented with mild intellectual disability, short stature, pes planus, whereas the younger brother (patient 5, DECIPHER 259774) exhibited neurodevelopmental and speech delay, otitis media with effusion and elliptocytosis (**Figure 1; Table 1; Supp. Table S3**).

Molecular findings

Chromosome and SNP-array analysis of patient 1 revealed a *de novo* balanced translocation involving chromosomes X and 9 (46,X,t(X;9)(q23;q11.2); **Supp. Figure S3a**). Array-CGH of the microdissected and amplified DNA from the derivative chromosomes mapped the Xq23 breakpoint to a 62 kb region (chrX:110,259,238-110,321,368; hg38; **Supp. Figure S3b**). The Xq23 and the 9q12 breakpoints were validated by FISH (**Supp. Figure S3d**). We then sequenced the patient's whole genome to finely map the breakpoints. Eleven chimeric inserts with paired-end reads mapping to different chromosomes and two soft clipped sequences, i.e. reads with portions of the read not part of the alignment (**Supp. Table S4**), and subsequent validation by Sanger sequencing, allowed mapping the Xq23 breakpoint between *AMMECR1* and *RGAG1* at coordinates chrX:110,318,754 (hg38) and chrX:110,318,762 (hg38) on the der(X) and der(9) chromosome, respectively, with a loss of seven nucleotides (**Figure 2a; Supp. Figure S4**). The 9q12 breakpoint affected a region of repeated nature that could not be unequivocally mapped; however the corresponding soft-clipped reads are consistent

with chromosome 9 nucleotide sequences. HUMARA and 5-ethynyl-2′-deoxyuridine (EdU) incorporation showed an extremely skewed inactivation (100:0) with the normal X-chromosome inactivated in all metaphases analyzed (**Supp. Figure S5**). As the *AMMECR1* and *RGAG1* genes do not generally escape X-chromosome inactivation (Carrel and Willard, 2005; Cotton et al., 2015), only the translocated allele would be transcriptionally active. We assessed expression of *AMMECR1* and *RGAG1* by RT-qPCR in LCLs from the patient, her mother and Brazilian female controls. Whereas *AMMECR1* was expressed in LCLs from the mother and ten controls (CT mean 30.1; CT SD 0.8, $\Delta CT_{AMMECR1\text{-}GAPDH}$ mean 13.3; $\Delta CT_{AMMECR1\text{-}GAPDH}$ SD 0.9), we failed to detect *AMMECR1* gene expression in the patient's cells (**Figure 2b**). On the contrary, we found no differences in expression levels of *RGAG1* in the assessed LCLs (**Figure 2b**). The absence of expression of *AMMECR1* in the patient and the presence of AMMECR1 expression in controls (CT mean 31.4; CT SD 0.4, $\Delta CT_{AMMECR1\text{-}GAPDH}$ mean 5.4; $\Delta CT_{AMMECR1\text{-}GAPDH}$ SD 0.3) were confirmed in a second RNA source (peripheral blood; **Figure 2b**). Such lack of expression could be linked to a breakpoint-induced separation of the promoter/enhancers and transcribed region of *AMMECR1* coupled to a skewed X-inactivation. In support of this hypothesis the Xq23 breakpoint maps 515 bp 5' to the *AMMECR1* transcriptional start site of two of the three coding transcripts (NM_015365/ENST00000262844 and NM_001025580/ENST00000372059) in a region enriched according to ENCODE (Consortium, 2012) in H3K4me3 and H3K27ac histone marks that are often found at promoters and regulatory regions. It disrupts the third *AMMECR1* coding transcript (NM_001171689/ENST00000372057) within its intron 2 (**Figure 2a**). Consistent with a possible causative role of *AMMECR1* in the patient's syndrome, the four individuals listed in ExAC (0.3.1) (Lek et al., 2016) with lof variants in that gene are heterozygous females and *AMMECR1* is «extremely intolerant» to lof with a pLI=0.91. These results prompted us to search for further individuals with lof variants in *AMMECR1*. We identified two males (patients 2 and 3) carrying a hemizygous *AMMECR1* nonsense variant, i.e. the maternally-inherited c.502C>T p.(Arg168Ter) (chrX:110,264,571 (hg38)) and the *de novo* c.429T>A p.(Tyr143Ter) (chrX:110,317,691 (hg38)), respectively. Additionally, we describe a three generation pedigree with two half-brothers, their mother, their maternal aunt and their maternal grandmother

carrying a 414 kb deletion from chrX:110,250,890 to chrX:110,665,082 (hg38) encompassing the *AMMECR1* and *RGAG1* genes (**Figure 3**). Patient 5 also carries a maternally inherited 1.25 Mb 8p22 duplication with unknown clinical significance encompassing *MSR1*, *SGCZ* and *TUSC3*. The overlap of clinical features of the female with the X-autosome balanced translocation, the nonsense variant hemizygous males, and the half-brothers with hemizygous deletion, all causing *AMMECR1* lof, is concordant with the involvement of *AMMECR1* in growth failure, heart malformations and skeletal dysplasia (**Table 1**, **Supp. Table S3**).

AMMECR1 interaction network

The AMMECR1 protein is conserved in vertebrates, whereas proteins containing the C-terminal part of human AMMECR1, named AMMECR1 domain, are present in eukaryotes, bacteria and archea (**Supp. Figure S6a**). The human AMMECR1 encodes a 33kDa protein of unknown function that harbors two RAGNYA folds known to interact with nucleic acid (Balaji and Aravind, 2007). The phylogeny of proteins containing RAGNYA traces this domain back to the last universal common ancestor of all extant life forms (Balaji and Aravind, 2007). An *AMMECR1* paralog, *AMMECR1L*, appeared in tetrapods. Like its homolog, the 68% similar to *AMMECR1* (**Supp. Figure S6b**) autosomal *AMMECR1L* is «extremely intolerant» to lof variation according to ExAC with a pLI=1.0. *AMMECR1L* is overexpressed in LCLs from patient 1 (**Figure 2b,** $p<0.05$). Of note, her mother presented a similarly elevated expression level when compared to female Brazilian controls (**Figure 2b**). This also impacted the level of the encoded protein as western blotting showed that AMMECR1L was higher in cells of patient 1 than in cells from a control individual (**Figure 2c, Supp. Figure S7**). Our results suggest that AMMECR1L could potentially partially compensate for loss of AMMECR1. Alternatively, the elevated *AMMECR1L* expression level of patient 1 could be caused by a coincidental expression quantitative trait loci (eQTL) passed down from her mother.

In line with nucleic acid binding, transfection assay in HeLa and HEK293 cells showed a nuclear sublocalization of GFP-tagged AMMECR1 (**Supp. Figure S8**). The western blotting results were less straightforward as they showed that the tagged AMMECR1 was present in both the cytoplasmic and

nuclear fractions (**Supp. Figure S7 and S9**). These results possibly reflect exogenous higher expression levels as the endogenous AMMECR1L paralog was exclusively identified in the nuclear extract of LCL cells of both patient 1 and a control individual (**Supp. Figure S7**). As the major antibody-reactive band ran slightly above 60kDa and as AMMECR1L is predicted to be 31kDa in size, we hypothesized that AMMECR1 proteins are creating dimers insensitive to SDS-PAGE (**Supp. Figure S7**). Consistent with our hypothesis a ≈30kDa reactive band could be identified upon longer exposure. This was confirmed when assessing transfected GFP-AMMECR1: whereas in most experiments the major reactive band ran at approximately 60kDa, in some experiments the reactive band corresponded to twice the singleton mass of this tagged proteins, i.e. 120kDa (**Supp. Figure S7**). Our results suggest that a fraction of AMMECR1 and AMMECR1L proteins are present in the nuclear fraction of the cells as dimers.

We used the GIANT database (Genome-scale Integrated Analysis of gene Networks in Tissues) (Greene et al., 2015) to predict the AMMECR1 functional network. 77 genes are interacting with AMMECR1 directly or with a maximum of one intermediate node (**Supp. Figure S10; Supp. Table S5**). This list of genes was enriched for Gene Ontology (GO) Biological Process term "mitotic cell cycle" (GO:0000278; Fisher exact test $p=2.1x10^{-21}$; z-score=-2.3; combined score=93.78) and GO Cellular Component term "nucleoplasm" (GO:0005654; $p=6.2x10^{-14}$; z-score=-2.26; combined score=57.71; **Supp. Table S6-S7**). The list of genes that interact with *AMMECR1L*, the *AMMECR1* paralog in humans, presents "regulation of transcription involved in G1/S transition of mitotic cell cycle" (GO:0000083) as the most enriched GO term (**Supp. Table S8-S10**). Other network models generated as controls with the same settings do not present enrichment of cell cycle related biological processes and cellular compartments, suggesting that the association between AMMECR1 and these GO terms is specific (**Supp. Table S11-S16**). Five of the 77 members of the *AMMECR1* functional network, *ZWINT, CDC6, PSAT1*, *EZH2* and *SMC2*, encode genes associated with syndromes characterized with growth and bone alterations, namely Roberts syndrome (MIM# 268300), Meier-Gorlin syndrome 5 (MIM# 613805), Neu-Laxova syndrome 2 (MIM# 616038), Weaver syndrome

(MIM# 277590) and Smith-McCort dysplasia 2 (MIM# 615222), respectively. These results further support a causative role of lof of *AMMECR1* in growth and bone phenotypes.

AMMECR1 knockdown in zebrafish

The zebrafish embryo is an emerging *in vivo* model for translational medicine (Chetaille et al., 2014; Golzio et al., 2012; Payne et al., 2014; Stewart et al., 2014). To further assess the consequences of absence of *AMMECR1* expression, we used a morpholino (MO)-induced knockdown of its zebrafish ortholog *ammecr1*. *ammecr1* is expressed in the whole zebrafish embryo according to ZFIN. Injection of early zebrafish embryos with 2.1, 4.2 and 8.4 ng of MO-I2E3 resulted in a 50, 72 and 84% reduction, respectively, of *ammecr1* transcripts through exon 3 skipping and frameshift, seen at 3 dpf (days post-fertilization) when compared with non-injected fish (**Supp. Figure S11**). Morphants had shorter tails, thinner bodies, kinked tail-ends, poorly defined somites, pericardial edema, tachycardia and hydrocephaly (**Figure 4**), phenotypes reminiscent of the patients' features, such as short stature, bone dysplasia, atrial septal defect and tachycardia. The severity of the phenotype increased according to the degree of *ammecr1* knock-down (**Supp. Figure S11-S12**). These features were not triggered by mechanical stress since mock injections with the same doses of standard Ctrl-MO, with no targets in the zebrafish transcriptome, induced no phenotypes. Aware that nonspecific effects have been reported when using MOs (Kok et al., 2015), we designed two other MOs that target different splice junctions, i.e. exon 3/intron 3 (MO-E3I3) and exon 1/intron 1 (MO-E1I1). Injections of 8.4 ng of MO-E3I3 induced the modification of 90% of the *ammecr1* transcripts. 65% of the mRNAs suffered exon 3 skipping and frameshift, whereas the remaining 25% included a 24 bp deletion within exon 3, probably resulting from a new splicing site (**Supp. Figure S13**). The outcome was heterogeneous with two classes of severity. Both classes presented phenotypic effects comparable to those observed in the MO-I2E3 experiments with small body size, kinked tail-ends, poorly defined somites, pericardial edema and tachycardia (**Figure 4**). Injections of MO-E1I1 produced what appears to be a hypomorphic transcript lacking the last 6 codons encoded by exon 1 (**Supp. Figure S14**). Embryos injected with 8.4 ng of this third MO presented tachycardia (**Figure 4b**) but no morphological alteration. To further substantiate the specificity of the observed features we tested if the knockdown

phenotype could be rescued by co-injection of mRNA from the human ortholog, *AMMECR1*. Co-injection of 1.26 ng of MO-I2E3 and 200 pg of *AMMECR1* mRNA restored tail length ($p<0.05$) and resulted in less impaired animals when compared with MO-I2E3 injection (**Figure 4b-d**). These results indicate that the function of *AMMECR1* was conserved from teleosts to primates.

The correspondence of the phenotypic outcomes using three different MOs targeting *ammecr1* and the phenotypic rescue generated by co-injection of the human *AMMECR1* mRNA suggest an association between the observed features and a decrease in *ammecr1* expression rather than "off target" interference. The zebrafish knockdowns present more severe phenotypes than the patients, i.e. greater morphological alterations and larger range of affected tissues. This could be potentially linked to the absence of an *ammecr1* paralog in non-tetrapod vertebrates. Consistent with this hypothesis, *AMMECR1L* is expressed at higher levels than *AMMECR1* despite largely overlapping expression pattern according to GTEx. We similarly found that *Ammecr1l* was more expressed than *Ammecr1* in mice embryonic tissues affected in the patients (**Supp. Figure S15**).

DISCUSSION

We found through breakpoint mapping of an X-autosome balanced translocation, exome and genome sequencing and array-CGH that lof of *AMMECR1* is associated with a novel syndrome that groups heart, bone and growth alterations and in one family elliptocytosis. Whereas failure to thrive and skeletal alterations are common to our patients and the four previously reported MHFEIN individuals (Basel-Vanagaite et al., 2017) (Andreoletti et al., 2016), psychomotor delay, elliptocytosis, cardiac and renal alterations for example are present only in a subset of affected individuals (**Table 1, Supp. Table S3**). The description of the p.(Gly177Asp) half-brothers (Andreoletti et al., 2016) encouraged us to search for individuals with overlapping phenotypes and predicted-to-be-deleterious missense variants in *AMMECR1*. Whereas we identified three such male individuals, as none of their AMMECR1 variants (p.Pro72Leu, p.Pro73Arg and p.Arg313Cys) induced mislocalization of the encoded protein in cellular assays (not shown) and as some of their features are dissimilar to those of

the patients described here, future work is warranted to confirm a possible association between these genetic changes and their diverse phenotypes.

Although little is known about the function of AMMECR1 and AMMECR1L, available information suggests roles in control of cell division. *AMMECR1* expression is upregulated in response to estrogen and activin A, a member of the TGFβ family, in breast cancer cells and embryonic stem cells, respectively (Tan et al., 2014; Tsai et al., 2011). It is targeted by ELAV1, a protein that regulates the stability and translation of a variety of mRNAs encoding stress-response and proliferative proteins (Abdelmohsen et al., 2009). The silencing of the *AMMECR1-like* paralog in HeLa cells induces a G1 arrest phenotype (Kittler et al., 2007), whereas expression of both the *Saccharomyces cerevisiae* (YOR289W) and *Schizosaccharomyces pombe* (SPAC688.03c.1) orthologs peaks in G1 (Marguerat et al., 2012; Santos et al., 2015). YOR289W is sensitive to drugs that block translation, is interacting genetically with translation regulators and is synthetic lethal with *GCL7*, a spindle checkpoint regulator gene (Cherry et al., 2012). Correspondingly, the *AMMECR1* functional network established using GIANT is enriched for genes associated with the mitotic cell cycle, the G1/S transition in particular, and DNA replication and packaging and contains five genes (*ZWINT, CDC6, PSAT1*, *EZH2* and *SMC2)* associated with perturbations of intrauterine growth and bone maturation. Our findings suggest that *AMMECR1* is potentially involved in cell cycle control and linked to a new syndrome with growth, bone and heart alterations.

CONFLICT OF INTEREST

The authors declare that they have no competing interests.

ACKNOWLEDGMENTS

We thank the patients and families for their contribution to this study, Beatriz Costa-Carvalho, Magnus Dias, Ravi Savarirajan, Viesturs Simanis and Taru Tukiainen for discussions and the members of the Lausanne Genomic Technologies Facility for technical help. We are grateful to Jennifer E. Posey, Ender Karaca, Zeynep Akdemir and James R. Lupski who shared unpublished data

with us. MMO is recipient of a FAPESP fellowship (Fundação de Amparo à Pesquisa do Estado de São Paulo). GG is a grantee of the Pro-Women program of the Faculty of Biology and Medicine, University of Lausanne. This work was supported by grants from the Swiss National Science Foundation [31003A_160203 to A.R.]; the Lithuanian-Swiss cooperation program to A.R.; and the FAPESP [2014/11572-8] to MIM. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

REFERENCES

Abdelmohsen, K, Srikantan, S, Yang, X, Lal, A, Kim, HH, Kuwano, Y, Galban, S, Becker, KG, Kamara, D, de Cabo, R, Gorospe, M. 2009. Ubiquitin-mediated proteolysis of HuR by heat shock. *Embo J, 28*(9), 1271-1282.

Adzhubei, I, Jordan, DM, Sunyaev, SR. 2013. Predicting functional effect of human missense mutations using PolyPhen-2. *Curr Protoc Hum Genet, Chapter 7*, Unit7 20.

Andreoletti, G, Seaby, EG, Dewing, JM, O'Kelly, I, Lachlan, K, Gilbert, RD, Ennis, S. 2016. AMMECR1: a single point mutation causes developmental delay, midface hypoplasia and elliptocytosis. *J Med Genet*.

Balaji, S, Aravind, L. 2007. The RAGNYA fold: a novel fold with multiple topological variants found in functionally diverse nucleic acid, nucleotide and peptide-binding proteins. *Nucleic Acids Res, 35*(17), 5658-5671.

Basel-Vanagaite, L, Pillar, N, Isakov, O, Smirin-Yosef, P, Lagovsky, I, Orenstein, N, Salmon-Divon, M, Tamary, H, Zaft, T, Bazak, L, Meyerovitch, J, Pelli, T, Botchan, S, Farberov, L, Weissglas-Volkov, D, Shomron, N. 2017. X-linked elliptocytosis with impaired growth is related to mutated AMMECR1. *Gene, 606*, 47-52.

Carrel, L, Willard, HF. 2005. X-inactivation profile reveals extensive variability in X-linked gene expression in females. *Nature, 434*(7031), 400-404.

Chen, EY, Tan, CM, Kou, Y, Duan, Q, Wang, Z, Meirelles, GV, Clark, NR, Ma'ayan, A. 2013. Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinformatics, 14*, 128.

Cherry, JM, Hong, EL, Amundsen, C, Balakrishnan, R, Binkley, G, Chan, ET, Christie, KR, Costanzo, MC, Dwight, SS, Engel, SR, Fisk, DG, Hirschman, JE, Hitz, BC, Karra, K, Krieger, CJ, Miyasato, SR, Nash, RS, Park, J, Skrzypek, MS, Simison, M, Weng, S, Wong, ED. 2012. Saccharomyces Genome Database: the genomics resource of budding yeast. *Nucleic Acids Res, 40*(Database issue), D700-705.

Chetaille, P, Preuss, C, Burkhard, S, Cote, JM, Houde, C, Castilloux, J, Piche, J, Gosset, N, Leclerc, S, Wunnemann, F, Thibeault, M, Gagnon, C, Galli, A, Tuck, E, Hickson, GR, El Amine, N, Boufaied, I, Lemyre, E, de Santa Barbara, P, Faure, S, Jonzon, A, Cameron, M, Dietz, HC, Gallo-McFarlane, E, Benson, DW, Moreau, C, Labuda, D, Consortium, FC, Zhan, SH, Shen, Y, Jomphe, M, Jones, SJ, Bakkers, J, Andelfinger, G. 2014. Mutations in SGOL1 cause a novel cohesinopathy affecting heart and gut rhythm. *Nat Genet, 46*(11), 1245-1249.

Consortium, EP. 2012. An integrated encyclopedia of DNA elements in the human genome. *Nature, 489*(7414), 57-74.

Cotton, AM, Price, EM, Jones, MJ, Balaton, BP, Kobor, MS, Brown, CJ. 2015. Landscape of DNA methylation on the X chromosome reflects CpG density, functional chromatin state and X-chromosome inactivation. *Hum Mol Genet, 24*(6), 1528-1539.

Gazou, A, Riess, A, Grasshoff, U, Schaferhoff, K, Bonin, M, Jauch, A, Riess, O, Tzschach, A. 2013. Xq22.3-q23 deletion including ACSL4 in a patient with intellectual disability. *Am J Med Genet A, 161A*(4), 860-864.
Golzio, C, Willer, J, Talkowski, ME, Oh, EC, Taniguchi, Y, Jacquemont, S, Reymond, A, Sun, M, Sawa, A, Gusella, JF, Kamiya, A, Beckmann, JS, Katsanis, N. 2012. KCTD13 is a major driver of mirrored neuroanatomical phenotypes of the 16p11.2 copy number variant. *Nature, 485*(7398), 363-367.
Greene, CS, Krishnan, A, Wong, AK, Ricciotti, E, Zelaya, RA, Himmelstein, DS, Zhang, R, Hartmann, BM, Zaslavsky, E, Sealfon, SC, Chasman, DI, FitzGerald, GA, Dolinski, K, Grosser, T, Troyanskaya, OG. 2015. Understanding multicellular function and disease with human tissue-specific networks. *Nat Genet, 47*(6), 569-576.
Jonsson, JJ, Renieri, A, Gallagher, PG, Kashtan, CE, Cherniske, EM, Bruttini, M, Piccini, M, Vitelli, F, Ballabio, A, Pober, BR. 1998. Alport syndrome, mental retardation, midface hypoplasia, and elliptocytosis: a new X linked contiguous gene deletion syndrome? *J Med Genet, 35*(4), 273-278.
Kimmel, CB, Ballard, WW, Kimmel, SR, Ullmann, B, Schilling, TF. 1995. Stages of embryonic development of the zebrafish. *Dev Dyn, 203*(3), 253-310.
Kittler, R, Pelletier, L, Heninger, AK, Slabicki, M, Theis, M, Miroslaw, L, Poser, I, Lawo, S, Grabner, H, Kozak, K, Wagner, J, Surendranath, V, Richter, C, Bowen, W, Jackson, AL, Habermann, B, Hyman, AA, Buchholz, F. 2007. Genome-scale RNAi profiling of cell division in human tissue culture cells. *Nat Cell Biol, 9*(12), 1401-1412.
Kok, FO, Shin, M, Ni, CW, Gupta, A, Grosse, AS, van Impel, A, Kirchmaier, BC, Peterson-Maduro, J, Kourkoulis, G, Male, I, DeSantis, DF, Sheppard-Tindell, S, Ebarasi, L, Betsholtz, C, Schulte-Merker, S, Wolfe, SA, Lawson, ND. 2015. Reverse genetic screening reveals poor correlation between morpholino-induced and mutant phenotypes in zebrafish. *Dev Cell, 32*(1), 97-108.
Lek, M, Karczewski, KJ, Minikel, EV, Samocha, KE, Banks, E, Fennell, T, O'Donnell-Luria, AH, Ware, JS, Hill, AJ, Cummings, BB, Tukiainen, T, Birnbaum, DP, Kosmicki, JA, Duncan, LE, Estrada, K, Zhao, F, Zou, J, Pierce-Hoffman, E, Berghout, J, Cooper, DN, Deflaux, N, DePristo, M, Do, R, Flannick, J, Fromer, M, Gauthier, L, Goldstein, J, Gupta, N, Howrigan, D, Kiezun, A, Kurki, MI, Moonshine, AL, Natarajan, P, Orozco, L, Peloso, GM, Poplin, R, Rivas, MA, Ruano-Rubio, V, Rose, SA, Ruderfer, DM, Shakir, K, Stenson, PD, Stevens, C, Thomas, BP, Tiao, G, Tusie-Luna, MT, Weisburd, B, Won, HH, Yu, D, Altshuler, DM, Ardissino, D, Boehnke, M, Danesh, J, Donnelly, S, Elosua, R, Florez, JC, Gabriel, SB, Getz, G, Glatt, SJ, Hultman, CM, Kathiresan, S, Laakso, M, McCarroll, S, McCarthy, MI, McGovern, D, McPherson, R, Neale, BM, Palotie, A, Purcell, SM, Saleheen, D, Scharf, JM, Sklar, P, Sullivan, PF, Tuomilehto, J, Tsuang, MT, Watkins, HC, Wilson, JG, Daly, MJ, MacArthur, DG. 2016. Analysis of protein-coding genetic variation in 60,706 humans. *Nature, 536*(7616), 285-291.
Liehr, T, Heller, A, Starke, H, Claussen, U. 2002. FISH banding methods: applications in research and diagnostics. *Expert Rev Mol Diagn, 2*(3), 217-225.
Marguerat, S, Schmidt, A, Codlin, S, Chen, W, Aebersold, R, Bahler, J. 2012. Quantitative analysis of fission yeast transcriptomes and proteomes in proliferating and quiescent cells. *Cell, 151*(3), 671-683.
Meloni, I, Vitelli, F, Pucci, L, Lowry, RB, Tonlorenzi, R, Rossi, E, Ventura, M, Rizzoni, G, Kashtan, CE, Pober, B, Renieri, A. 2002. Alport syndrome and mental retardation: clinical and genetic dissection of the contiguous gene deletion syndrome in Xq22.3 (ATS-MR). *J Med Genet, 39*(5), 359-365.
Moyses-Oliveira, M, Guilherme, RD, Dantas, AG, Ueta, R, Perez, AB, Haidar, M, Canonaco, R, Meloni, VA, Kosyakova, N, Liehr, T, Carvalheira, GM, Melaragno, MI. 2015. Genetic mechanisms leading to primary amenorrhea in balanced X-autosome translocations. *Fertil Steril*.
Payne, F, Colnaghi, R, Rocha, N, Seth, A, Harris, J, Carpenter, G, Bottomley, WE, Wheeler, E, Wong, S, Saudek, V, Savage, D, O'Rahilly, S, Carel, JC, Barroso, I, O'Driscoll, M, Semple, R. 2014. Hypomorphism in human NSMCE2 linked to primordial dwarfism and insulin resistance. *J Clin Invest, 124*(9), 4028-4038.
Richards, S, Aziz, N, Bale, S, Bick, D, Das, S, Gastier-Foster, J, Grody, WW, Hegde, M, Lyon, E, Spector, E, Voelkerding, K, Rehm, HL. 2015. Standards and guidelines for the interpretation of sequence

variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med, 17*(5), 405-424.
Salgado, D, Desvignes, JP, Rai, G, Blanchard, A, Miltgen, M, Pinard, A, Levy, N, Collod-Beroud, G, Beroud, C. 2016. UMD-Predictor: A High-Throughput Sequencing Compliant System for Pathogenicity Prediction of any Human cDNA Substitution. *Hum Mutat, 37*(5), 439-446.
Santos, A, Wernersson, R, Jensen, LJ. 2015. Cyclebase 3.0: a multi-organism database on cell-cycle regulation and phenotypes. *Nucleic Acids Res, 43*(Database issue), D1140-1144.
Settembre, C, Di Malta, C, Polito, VA, Garcia Arencibia, M, Vetrini, F, Erdin, S, Erdin, SU, Huynh, T, Medina, D, Colella, P, Sardiello, M, Rubinsztein, DC, Ballabio, A. 2011. TFEB links autophagy to lysosomal biogenesis. *Science, 332*(6036), 1429-1433.
Sisdelli, L, Vidi, AC, Moyses-Oliveira, M, Di Battista, A, Bortolai, A, Moretti-Ferreira, D, Dias da Silva, MR, Melaragno, MI, Carvalheira, G. 2016. Incorporation of 5-ethynyl-2'-deoxyuridine (EdU) as a novel strategy for identification of the skewed X inactivation pattern in balanced and unbalanced X-rearrangements. *Hum Genet, 135*(2), 185-192.
Stewart, AM, Braubach, O, Spitsbergen, J, Gerlai, R, Kalueff, AV. 2014. Zebrafish models for translational neuroscience research: from tank to bedside. *Trends Neurosci, 37*(5), 264-278.
Tan, S, Ding, K, Li, R, Zhang, W, Li, G, Kong, X, Qian, P, Lobie, PE, Zhu, T. 2014. Identification of miR-26 as a key mediator of estrogen stimulated cell proliferation by targeting CHD1, GREB1 and KPNA2. *Breast Cancer Res, 16*(2), R40.
Tsai, ZY, Chou, CH, Lu, CY, Singh, S, Yu, SL, Li, SS. 2011. Proteomic comparison of human embryonic stem cells with their differentiated fibroblasts: Identification of 206 genes targeted by hES cell-specific microRNAs. *Kaohsiung J Med Sci, 27*(8), 299-306.
Vaser, R, Adusumalli, S, Leng, SN, Sikic, M, Ng, PC. 2016. SIFT missense predictions for genomes. *Nat Protoc, 11*(1), 1-9.
Vitelli, F, Piccini, M, Caroli, F, Franco, B, Malandrini, A, Pober, B, Jonsson, J, Sorrentino, V, Renieri, A. 1999. Identification and characterization of a highly conserved protein absent in the Alport syndrome (A), mental retardation (M), midface hypoplasia (M), and elliptocytosis (E) contiguous gene deletion syndrome (AMME). *Genomics, 55*(3), 335-340.
Yang, Y, Muzny, DM, Xia, F, Niu, Z, Person, R, Ding, Y, Ward, P, Braxton, A, Wang, M, Buhay, C, Veeraraghavan, N, Hawes, A, Chiang, T, Leduc, M, Beuten, J, Zhang, J, He, W, Scull, J, Willis, A, Landsverk, M, Craigen, WJ, Bekheirnia, MR, Stray-Pedersen, A, Liu, P, Wen, S, Alcaraz, W, Cui, H, Walkiewicz, M, Reid, J, Bainbridge, M, Patel, A, Boerwinkle, E, Beaudet, AL, Lupski, JR, Plon, SE, Gibbs, RA, Eng, CM. 2014. Molecular findings among patients referred for clinical whole-exome sequencing. *JAMA, 312*(18), 1870-1879.

Table 1. General clinical description of the patients with variants involving AMMECR1

	Present work					**Andreoletti et al, 2016**		**Basel-Vanagaite et al, 2017**		**Jonsson et al, 1998**		**Gazou et al, 2013**
	P 1	**P 2**	**P 3**	**P 4**	**P 5**	**P 1**	**P 2**	**P**	**P's uncle**	**P 850**	**P 850's brother**	**P**
Variant		**LoF**				**Missense and in-frame deletion**				**AMME deletion syndrome**		
Gender	female	male	male	male	male	male	male	male	male	male	male	male
Failure to thrive	+	+	+	+	-	+	+	+	+	+	+	-
Short stature	+	+	+	+	-	-	+	+	+	-	-	-
Psychomotor delay	-	+	-	-	+	+	-	+	-	+	+	+
Intellectual disability	-	+	-	+	+	+	-	-	-	+	+	-
Prominent forehead	+	+	+	+	+	+	-	-	-	-	-	-
Flat face	+	+	+	+	+	+	+	-	-	-	-	-
Malar flattening	+	+	+	+	+	+	+	-	+	-	-	-
Midface hypoplasia	+	+	+	-	+	+	-	+	-	+	+	-
Skeletal alterations	+	+	+	+	-	+	+	+	+	+	-	-
Cardiac alterations	+	+	+	NI	NI	+	-	+	-	+	+	-
Renal alterations	-	+	-	-	-	+	+	+	-	+	-	-
Hearing loss	+	+	+	NI	NI	+	+	+	+	+	+	+
Elliptocytosis	-	-	-	-	+	+	-	+	+	+	+	-

+: present; -: absent; **NI**: not investigated; **P:** patient; **LoF**: Loss of function.

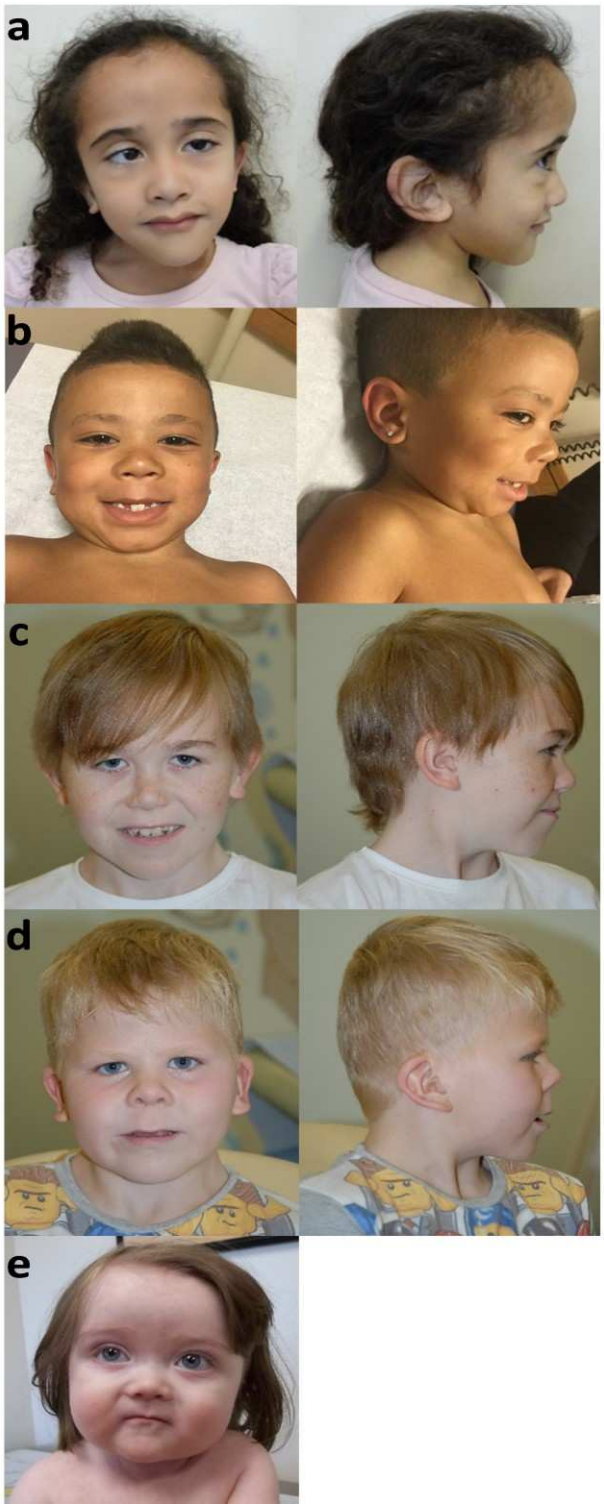

Figure 1: Facial features of patients with *AMMECR1* loss-of-function. (a) Patient 1, carrier of a balanced X-autosome translocation, at 6 years of age. **(b)** Patient 3, carrier of a maternally inherited nonsense variant, at 17 months of age. **(c-d)** Patients 4 and 5, carriers of a maternally inherited microdeletion, at 13 and 9 years of age, respectively. **(e)** Patient 2, carrier of a *de novo* nonsense variant, at 8 years of age.

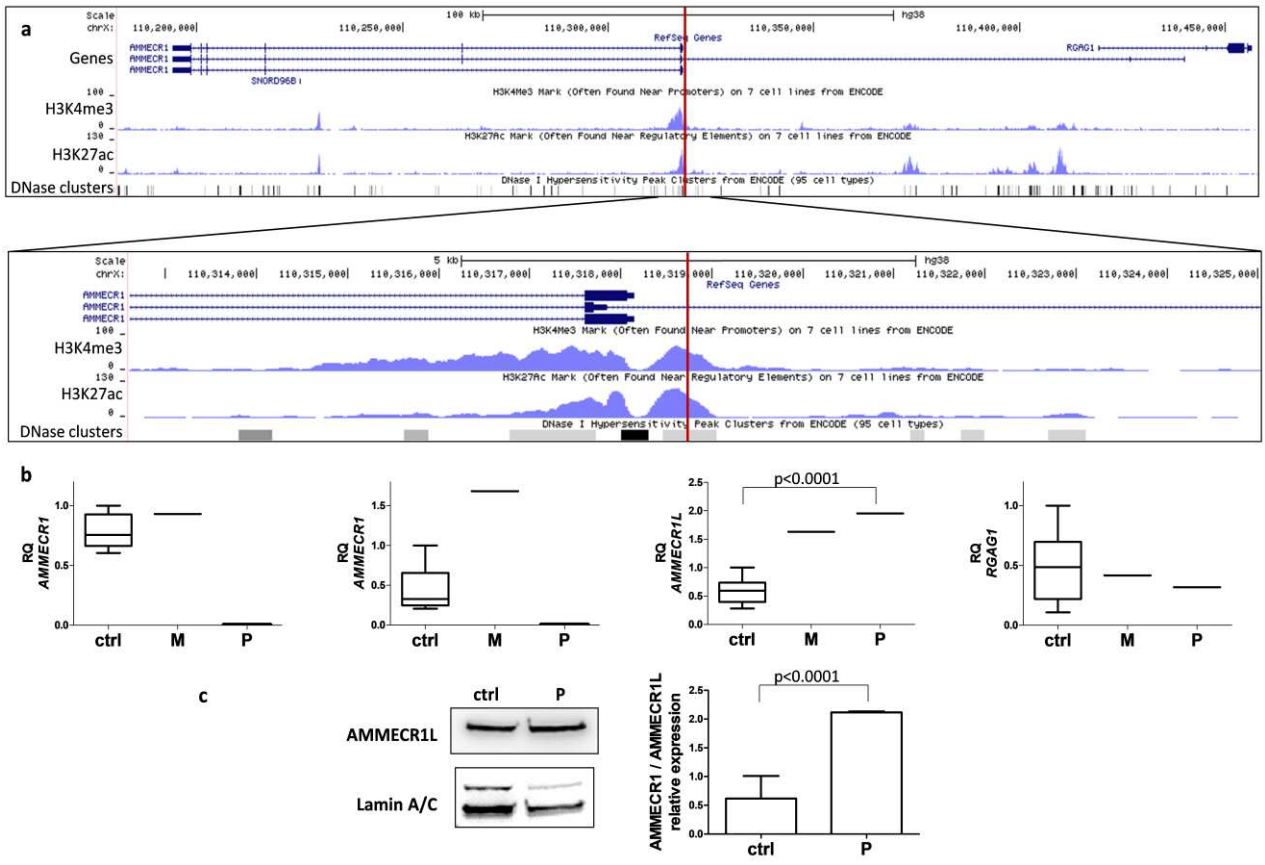

Figure 2: *AMMECR1* expression. (a) UCSC genome browser view of chromosome X region 110.18-110.46Mb (top) and 110.31-110.32Mb (bottom) showing the localization of the breakpoint between *AMMECR1* and *RGAG1* genes (vertical red line). The gene transcripts (blue), the H3K4me3 and H3K27ac histone marks (both blue) according to ENCODE (Consortium, 2012) are shown top to bottom. The Xq23 breakpoint maps 515 bp 5' to the *AMMECR1* transcriptional start site of transcripts NM_015365/ENST00000262844 and NM_001025580/ENST00000372059. It disrupts transcript NM_001171689/ENST00000372057 within its intron 2. **(b)** From left to right, RT-qPCR expression levels of *AMMECR1* in whole blood and of *AMMECR1*, *AMMECR1L* and *RGAG1* in the lymphoblastoid cells of n=10 Brazilian female controls (ctrl), the mother (M) and patient 1 (P). Note the absence of *AMMECR1* expression in the patient 1. Student's t-test was performed to compare gene expression in patient vs. controls and the corresponding p value is denoted on the bar graph. **(c)** Representative image of western blot in LCL from a control individual (ctrl) and patient (P) incubated with anti-AMMECR1/AMMECR1L (top left panel) and anti-lamin A/C (bottom left panel) as normalizer. The level of AMMECR1L was higher in cells of patient 1 than in cells from a control individual (right panel). More details are presented in **Supp. Figure S7**.

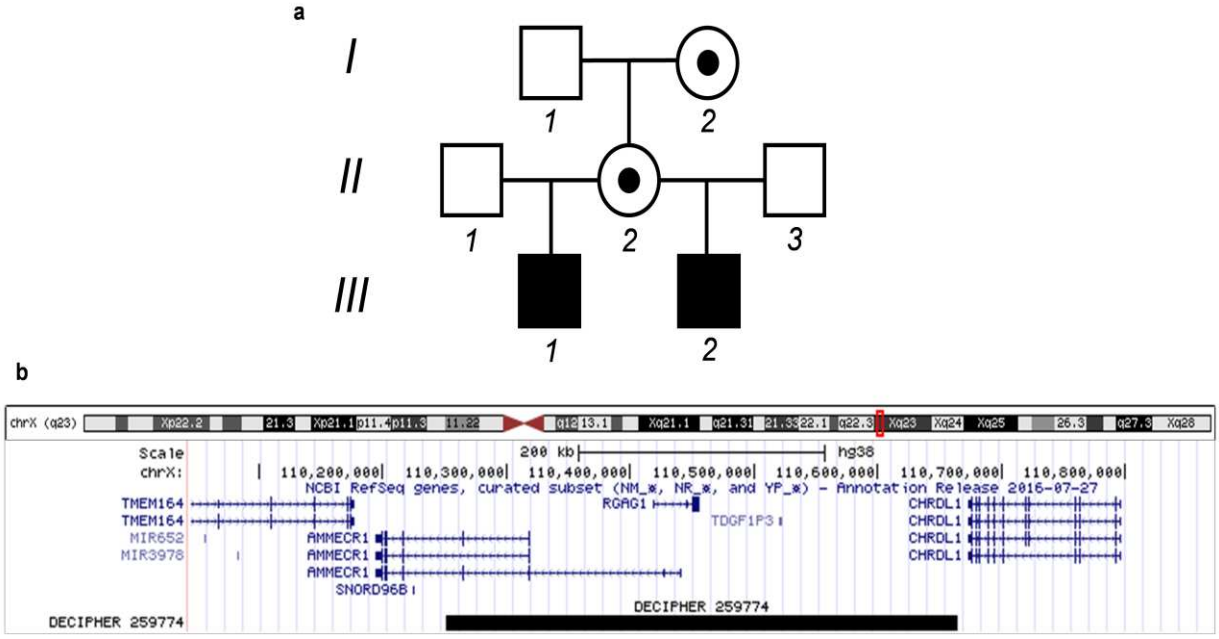

Figure 3: Pedigree of the family with Xq23 deletion encompassing *AMMECR1*. (a) Three generations pedigree with patients 4 and 5 (individuals III.1 and III.2; DECIPHER 259774). Both patients, their mother and maternal grandmother are carriers of the deletion. **(b)** UCSC genome browser view of chromosome X region 110.043-110.872Mb (hg38) showing the localization of the Xq23 414 kb deletion (horizontal black bar) encompassing the *AMMECR1* and *RGAG1* genes (blue).

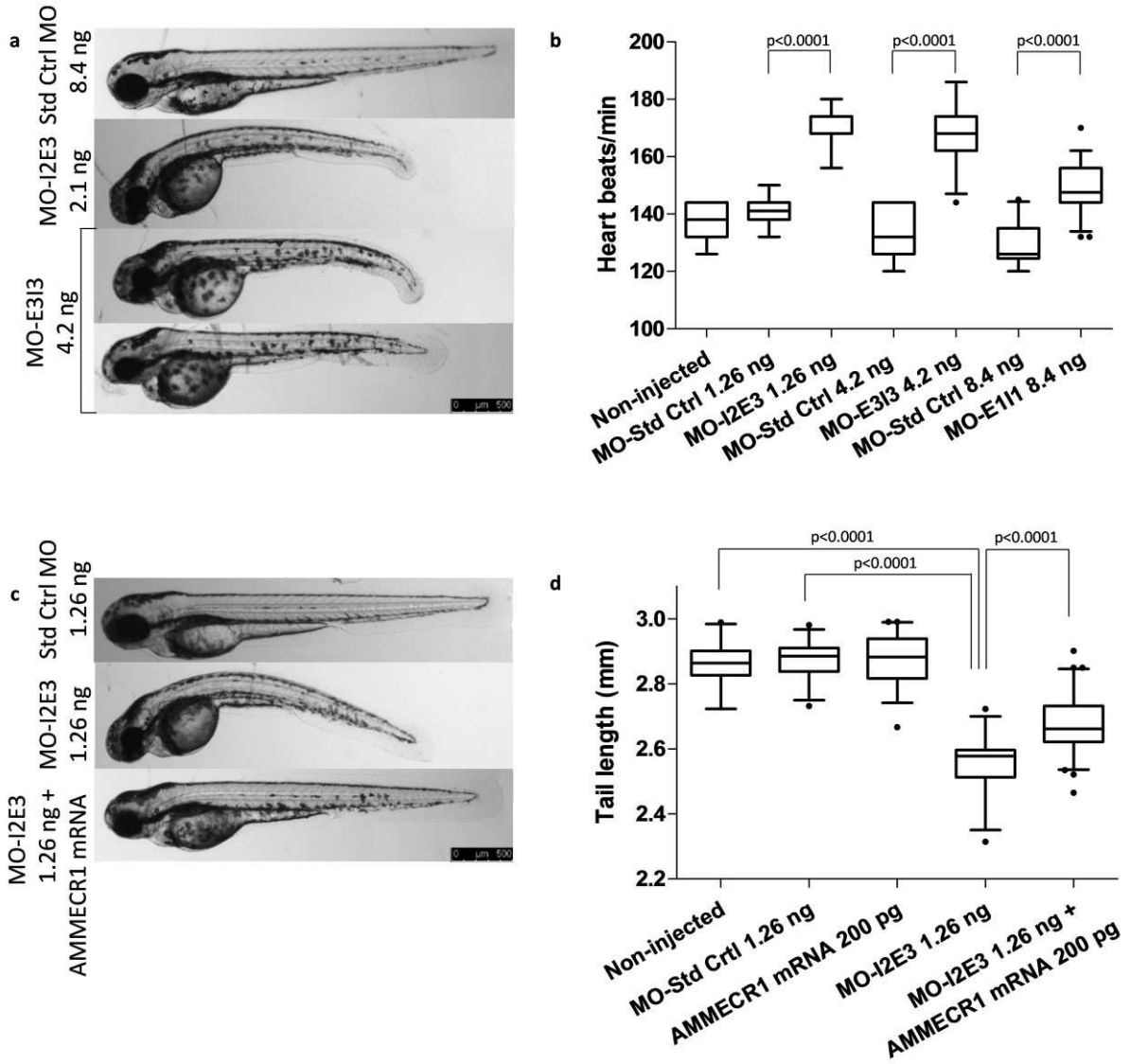

Figure 4: *ammecr1* knockdown in zebrafish results in phenotypic features reminiscent of patient 1`s clinical alterations. (a) From top to bottom, lateral views of representative 3 dpf zebrafish embryos injected with 8.4 ng of standard control MO (morpholino), 2.1 ng of MO-I2E3 and 4.2 ng of MO-E3I3 targeting *ammecr1*. The last injection resulted in two classes of phenotype (see text for details). **(b)** The heart rate of non-injected, standard control MO-injected and *ammecr1* MO-I2E3-, MO-E3I3-, and MO-E1I1-injected embryos are plotted. Student's t-test was performed and the corresponding p value is denoted on the bar graph. **(c)** From top to bottom, embryos injected with 1.26 ng of standard control MO, 1.26 ng of *ammecr1* MO-I2E3 and co-injected with 1.26 ng of MO-I2E3 and 200 pg of human *AMMECR1* mRNA. **(d)** The tail length (from otic vesicle to end of the tail following the body curvature) of non-injected, standard control MO-, human *AMMECR1* mRNA-, and

ammecr1 MO-I2E3-injected embryos are plotted. Co-injection of MO-I2E3 and *AMMECR1* mRNA partially rescued the short tail length observed in the morphants. The rescue experiment was performed twice and one representative example is shown.

LIST OF TABLES

Supp. Table S11: List of the 49 genes interacting directly or with a maximum of one node with *GAPDH* according to GIANT.

Supp. Table S12: GO Biological Process terms enriched in *GAPDH* functional network according to Enrichr.

Supp. Table S13: GO Cellular Component terms enriched in *GAPDH* functional network according to Enrichr.

Supp. Table S14: List of the 101 genes interacting directly or with a maximum of one node with *RGAG1* according to GIANT.

Supp. Table S15: GO Biological Process terms enriched in *RGAG1* functional network according to Enrichr.

Supp. Table S16: GO Cellular Component terms enriched in RGAG1 functional network according to Enrichr.