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

Moysés-Oliveira, Mariana; Giannuzzi, Giuliana; Fish, Richard; Rosenfeld, Jill A; Petit, Florence; Soares, Maria de Fatima; Kulikowski, Leslie Domenici; Di-Battista, Adriana; Zamariolli, Malú; Xia, Fan; Liehr, Thomas; Kosyakova, Nadezda; Carvalheira, Gianna; Parker, Michael [**and 21 more**]

How to cite

MOYSÉS-OLIVEIRA, Mariana et al. Inactivation of AMMECR1 is associated with growth, bone, and heart alterations. In: Human Mutation, 2018, vol. 39, n° 2, p. 281–291. doi: 10.1002/humu.23373

This publication URL: https://archive-ouverte.unige.ch/unige:103594

Publication DOI: 10.1002/humu.23373

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

Mariana Moysés-Oliveira[1,2], Giuliana Giannuzzi[2], Richard J. Fish[3], Jill A. Rosenfeld[4], Florence Petit[5], Vera Ayres Meloni[1], 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], Joris Andrieux[9], Taiza Stumpp[10], Fernanda Antunes[11], Gustavo José Pereira[11], Marguerite Neerman-Arbez[3], Margaret Drummond-Borg[12], 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

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

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

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

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

Correspondence to

alexandre.reymond@unil.ch

Running Title: *AMMECR1* loss-of-function

Summary

We report a nine-year-old girl with a balanced X-autosome translocation 46,X,t(X;9)(q23;q12)dn, preferential inactivation of the normal X-chromosome, disproportionate short stature, atrial septal defect, scoliosis, bone dysplasia, high-frequency hearing loss, coarse face and normal cognition. The X-chromosomal breakpoint maps between *AMMECR1* and *RGAG1*. Whereas expression of the latter was unmodified, RT-qPCR showed absence of *AMMECR1* expression in cells of the proband. Concordantly, we identified a male hemizygote for a nonsense variant in *AMMECR1* presenting with short stature, failure to thrive, congenital heart disease with prominent aorta and atrial septal defect, radioulnar synostosis and high-frequency hearing loss. Four male individuals from two families with variants of unknown significance in *AMMECR1*, a missense and in-frame deletion, and partially overlapping phenotypes plus elliptocytosis and renal alterations were previously identified. *AMMECR1* is co-expressed with genes implicated in cell cycle and translation regulation, five of which were previously associated with growth and bone alteration syndromes. Our knockdown of the zebrafish orthologous gene resulted in animals with phenotypes such as shorter tails, thinner bodies, kinked tail-ends, poorly defined somites, pericardial edema, tachycardia and hydrocephaly reminiscent of probands' features.

The increased transcript and encoded protein levels of *AMMECR1L*, an *AMMECR1* paralog, in the cells of the t(X;9) proband indicate a partial compensatory mechanism. The AMMECR1/AMMECR1L proteins dimerize and localize to the nucleus as suggested by their nucleic acid-binding RAGNYA domains.

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 and elliptocytosis.

INTRODUCTION

Balanced chromosomal rearrangements have a prevalence of approximately one in 500 individuals in the general population (1). Whereas these structural abnormalities usually have no phenotypic consequences (2, 3), around 6% of *de novo* cases of apparently balanced reciprocal translocations are associated with clinical abnormalities (4) through either disruption of gene(s) mapping to the breakpoint(s) (5), separation of regulatory elements from their gene core (6), chromosome repositioning (7) and/or modification of the expression levels of gene(s) mapping to the rearranged chromosomes (8). Thus, high-resolution breakpoint definition in balanced chromosomal rearrangements associated with abnormal phenotype can be used as a tool to identify disease loci (9, 10).

The most frequent features of females with balanced reciprocal translocations involving an X-chromosome are premature ovarian failure, intellectual disability, congenital abnormalities and X-linked disorders (11-13). In the majority of cases, females with balanced X-autosome translocation present a skewed X-chromosome inactivation pattern to maintain gene dosage (14). In these instances, the normal X-chromosome is usually inactivated, whereas the derivative X-chromosome allele is active (15). Thus, expression impairment of an X-chromosome gene through translocation frequently results in absence of functional copies in females (16-18).

Here we describe a male and a female with variants that inactivate the X-linked *AMMECR1* gene and present with short stature, bone dysplasia and atrial septal defect. The female carries a balanced X-autosome translocation that disrupts the *AMMECR1* gene and silencing of the normal X, whereas the male is hemizygous for a nonsense variant in *AMMECR1*. Our results support the involvement of that gene in a new syndrome that groups growth abnormalities with alterations of bones and heart.

RESULTS

Proband 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 term (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 proband spent 29 days in neonatal intensive care unit 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 value 51-303 ng/ml), TSH was 0.47 µUI/ml (reference value 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 (<3rd centile), disproportionate short stature (97 cm, <1st centile), head circumference of 44 cm (<3rd centile), 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; Supplementary Figure S1; Supplementary Table S1**). 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 (**Figure 1; Supplementary Figure S1; Supplementary Table S1**). At seven years of age, a brain MRI showed no abnormalities, and neuropsychological evaluation using the WISC IV revealed normal intelligence levels (IQ 91). At the same age, GH replacement therapy was initiated, and the proband responded to the treatment. The echocardiogram showed an ostium primum atrial septal defect with thin aneurysmal septal tissue. Different medical examinations at the age of 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, CD8 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.

Chromosome analysis of the proband revealed a translocation involving chromosomes X and 9 (46,X,t(X;9)(q23;q11.2); **Figure 2a**) and a normal karyotype in her parents, demonstrating that the translocation was *de novo*. A SNP-array analysis showed neither pathogenic deletions or duplications, nor loss of genetic material in the rearrangement region suggesting a balanced translocation. 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; **Figure 2b**). This breakpoint localization

was validated by FISH, which showed splitting of the CTD-3066N24 BAC-probe signal in both derivative chromosomes (**Figure 2c**). The chromosome 9 breakpoint affects a heterochromatic region that cannot be evaluated by array-CGH. The autosomal breakpoint was further validated by FISH with chromosome 9 centromeric and pericentromeric probes that allowed confirmation of its localization to 9q12 (**Figure 2d**). We then sequenced the proband'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, 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 3a; Supplementary Figure S2**). The 9q12 breakpoint affected a region of repeated nature and could not be unequivocally mapped; however the corresponding soft-clipped reads are consistent with chromosome 9 nucleotide sequences on both sides of the breakpoint.

The HUMARA assay showed an extremely skewed inactivation (100:0) and 5-ethynyl-2′-deoxyuridine (EdU) incorporation indicated that the normal X-chromosome was inactivated in all metaphases analyzed (**Supplementary Figure S3**). As the *AMMECR1* and *RGAG1* genes do not generally escape X-chromosome inactivation independently of the tissue or cell type (19, 20), only the translocated allele would be transcriptionally active. We assessed the expression of *AMMECR1* and *RGAG1* by RT-qPCR in LCLs from the proband, 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 proband's cells (**Figure 3b**). On the contrary, we found no differences in expression levels of *RGAG1* in the assessed LCLs (**Figure 3b**). The absence of expression of *AMMECR1* in the proband 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 3b**). 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 (canonical transcript NM_015365/ENST00000262844 and NM_001025580/ENST00000372059) in a region enriched according to ENCODE (21) 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 3a**). Consistent with a possible causative role of *AMMECR1* in the proband's syndrome, the four individuals listed in ExAC (Exome Aggregation Consortium, Cambridge, MA; Version 0.3.1) with lof (loss of function) variants in that gene are heterozygous females (p.(Gln74ProfsTer125); prevalence 4/12,908). *AMMECR1* is «extremely intolerant» to lof variation with a pLI=0.91.

These results prompted us to search for further individuals with lof variants in *AMMECR1*. Among individuals undergoing clinical exome sequencing, we identified one male individual carrying a hemizygous *AMMECR1* nonsense variant (c.502C>T; p.(R168X)) at chX:110,264,571 (hg38), inherited from his mother. Proband 2 presents with motor delay, speech delay, hypotonia, short stature and failure to thrive, dysmorphia, congenital heart disease with prominent aorta and atrial septal defect, ureterocele and hypercalcemia, radioulnar synostosis, small penis, and failed hearing screen for high pitch. The overlap of clinical features of both the female with the X-autosome balanced translocation and the nonsense variant hemizygous male, both causing *AMMECR1* lof, is concordant with the involvement of *AMMECR1* in growth failure, bone and heart malformations (**Table 1**).

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 (**Supplementary Figure S4a**). The human AMMECR1 encodes a 33 kDa protein of unknown function that harbors two RAGNYA domains known to interact with nucleic acid (22). The phylogeny of proteins containing RAGNYA traces this domain back to the last universal common ancestor of all extant life forms (22). An *AMMECR1* paralog, *AMMECR1L*, appeared in tetrapods. Like its homolog, the autosomal *AMMECR1L* is «extremely intolerant» to lof variation according to ExAC with a pLI=1.0 (0 variants observed out of 15.8 expected). It encodes a protein 68% similar to *AMMECR1* (**Supplementary Figure S4b**). Measuring *AMMECR1L* expression level by RT-qPCR revealed that it was overexpressed in LCLs from proband 1 (**Figure 3b,** $p<0.05$). Of note, her mother presented a similarly elevated expression level when compared to female Brazilian controls (**Figure**

3b). This impacted also protein levels as western blotting with an antibody that recognizes both AMMECR1 and AMMECR1L paralogs showed that the level of AMMECR1L alone was higher in cells of proband 1 than AMMECR1 and AMMECR1L together in cells from a control individual (**Figure 3c, Supplementary Figure S5**). Our results suggest that AMMECR1L could potentially partially compensate for loss of AMMECR1. In line with nucleic acid binding, transfection assay in HeLa and HEK293 cells showed a nuclear sublocalization of GFP-tagged AMMECR1 (**Supplementary Figure S6**). Western blotting of cytoplasmic and nuclear extracts of these transfected cells and of LCLs from proband 1 and a control individual confirmed the nuclear localization of the AMMECR1 and AMMECR1L proteins (**Supplementary Figure S5**). As the major anti-AMMECR1/AMMECR1L antibody-reactive bands in GFP-AMMECR1 transfected ran at approximately 120kD (**Supplementary Figure S5**), we hypothesized that AMMECR1 proteins are creating dimers insensitive to SDS-PAGE. This was confirmed when assessing the endogenous AMMECR1/AMMECR1L proteins as the reactive band ran at a molecular weight corresponding to the double of the singleton mass of these proteins, i.e. slightly above 60kD (**Supplementary Figure S5**). Consistent with our hypothesis a ≈30kD reactive band could be identified upon longer exposure. Our results suggest that almost all AMMECR1/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) (23) to identify the AMMECR1 functional network. 77 genes are interacting with AMMECR1 directly or with a maximum of one intermediate node (**Supplementary Figure S7; Supplementary Table S3**). 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; **Supplementary Table S4-S5**). This analysis revealed many other GO Biological Process terms related to cell cycle control, such as DNA conformation change (GO:0071103), chromosome organization (GO:0051276), chromosome segregation (GO:0007059), DNA replication (GO:0006260), DNA packaging (GO:0006323) and regulation of transcription involved in G1/S transition of mitotic cell cycle (GO:0000083) (**Supplementary Table S4-S5**). 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 (**Supplementary Table S6-S8**). Random network models generated with the same settings do not present enrichment of cell cycle related biological processes and cellular compartments, confirming the specificity of the association between AMMECR1 and these GO terms (**Supplementary Table S9-S14**). 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 (OMIM#268300), Meier-Gorlin syndrome 5 (#613805), Neu-Laxova syndrome 2 (#616038), Weaver syndrome (#277590) and Smith-McCort dysplasia 2 (#615222), respectively. These results further support a causative role of lof of *AMMECR1* in growth and bone phenotypes.

The zebrafish embryo is an emerging *in vivo* model for translational medicine (24-27). 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 (**Supplementary Figure S8**). Morphants had shorter tails, thinner bodies, kinked tail-ends, poorly defined somites, pericardial edema, tachycardia and hydrocephaly (**Figure 4**), phenotypes reminiscent of the probands' 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 (**Supplementary Figure S8-S9**). 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 (28), 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 (**Supplementary Figure S10**). 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 (**Supplementary Figure S11**). 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 probands, 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 probands (**Supplementary Figure S12**).

DISCUSSION

We found through breakpoint mapping of an X-autosome balanced translocation and exome sequencing that lof of *AMMECR1* is associated with a novel syndrome that groups heart, bone and growth alterations. The *AMMECR1* gene maps within the AMME (Alport syndrome, mental retardation, midface hypoplasia, and elliptocytosis) complex interval. AMME (OMIM#300194) is an Xq22.3 contiguous gene deletion syndrome encompassing about 20 genes (**Supplementary Figure S13, Supplementary Table S15**) characterized by nephropathy (Alport syndrome, OMIM#301050), intellectual disability, midface hypoplasia and elliptocytosis. Some affected individuals also present with hearing loss, ocular abnormalities, limb malformations, cardiac arrhythmia and heart

defects (29). The *COL4A5* gene was shown to be the causative gene for Alport syndrome, hearing loss and eye abnormalities (30, 31), whereas variants in *ACSL4* were associated with X-linked intellectual disability 63 (MRX63; OMIM#300387) (32-35). 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 (**Supplementary Figure S13, Supplementary Table S15**).

A family with two affected male individuals, the proband and one of his maternal uncles, and partially overlapping phenotypes, i.e. midface hypoplasia, proportionate short stature, hearing loss and elliptocytosis with or without anemia, was recently described (36) (**Table 1**). 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 proband chromosome X carries 12 more rare exonic variants in particular a frameshift and a stop gain variant in the lof intolerant *MAGEB3* (pLI = 0.66, 0 lof variants identified compared to 3.3 expected) and *SLC16A2* genes, respectively (pLI = 0.93, 0 identified/8.4 expected). Two half-brothers presenting elliptocytosis, midface hypoplasia, early speech and language delay carrying a p.G177D missense variant in *AMMECR1* that causes mislocalization of the encoded protein were also described (37). These four cases were regrouped by OMIM under the identifier MFHEIN (midface hypoplasia, hearing impairment, elliptocytosis, and nephrocalcinosis (MIM #300990). Their phenotypes and that of AMME deletion syndrome patients are compared to those of our probands in **Table 1**. Whereas failure to thrive, skeletal alterations and hearing impairment are common to both our two probands and the four MHFEIN individuals, psychomotor delay, cardiac and renal alterations for example are present only in a subset of affected individuals. The description of the half-brothers 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.P72L, p.P73R and p.R313C) induced mislocalization of the encoded protein in cellular assays (not shown) and as some of their features are dissimilar to those of probands 1 and 2, 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 (38, 39). It is targeted by ELAV1, a protein that regulates the stability and translation of a variety of mRNAs encoding stress-response and proliferative proteins (40). The silencing of the *AMMECR1-like* paralog in HeLa cells induces a G1 arrest phenotype (41), whereas expression of both the *Saccharomyces cerevisiae* (YOR289W) and *Schizosaccharomyces pombe* (SPAC688.03c.1) orthologs peaks in G1 (42, 43). 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 (44). 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. Although further studies are warranted to establish definitive causality, our findings suggest that *AMMECR1* is potentially involved in cell cycle control and linked to a new syndrome with growth, bone and heart alterations.

MATERIALS AND METHODS

Enrollment

All samples used in this study were collected with written appropriate informed consent and approval of the local ethics committee (CONEP 36019314.9.0000.5505 and Baylor College of Medicine IRB). The clinical evaluation included medical history interviews, a physical examination and review of medical records.

Classical and Molecular Cytogenetic Evaluation

550 resolution G-banding was performed on lymphocyte cultures using standard methods in a total of 40 analyzed metaphases per person. DNA for genomic array was isolated from peripheral blood using Gentra Puregene Kit (Quiagen-Sciencesm, Germantown, USA). Genomic array was performed using the Affymetrix Genome-Wide Human SNP-Array 6.0 (Affymetrix Inc., Santa Clara, CA, USA) or Human Genome CGH array 4x44k (Agilent Technologies, Palo Alto, CA, USA) and was analyzed, respectively, by the Genotyping Console 3.0.2 and

Chromosome Analysis Suite (ChAS) software (Affymetrix), and the CytoGenomics software (Agilent), both according to GRCh37/hg19.

Breakpoint mapping

FISH (fluorescence *in situ* hybridization) was performed in lymphocyte metaphase cells using chromosome 9 centromeric and pericentromeric probes to determine the autosome breakpoint as described (45). The sex chromosome breakpoint was ascertained by microdissection and amplification of the derivative chromosomes followed by array-CGH (comparative genome hybridization) as described (46). Breakpoint localization was validated by FISH using BAC (bacterial artificial chromosome) probe CTD-3066N24 (Thermo Fisher Scientific) and whole-genome sequencing. 2 µg of genomic DNA were sheared using Covaris with a target fragment size of 550 bp. The sequencing library was prepared using Tru-Seq DNA PCR-free Sample Prep Kit (Illumina) and 100 bp paired-end reads sequenced on a HiSeq 2500 platform (Illumina). The PCR-free kit was used to prepare the library in order to avoid PCR duplicates. Sequence-control, software real-time analysis and CASAVA software v1.8.2 (Illumina) were used for image analysis and base calling. Burrows-Wheeler Aligner (BWA-MEM) v7.10 with default parameters was used to map the data to hg38. We reached an 4.5x coverage and a sequence yield of 14.4 Gb. The mean insert size was 606 bp, alignment rate 97.8%, proper pair coverage 94.5%, duplicate rate 4.82% and chimera rate 0.34%. Aligned reads adjacent to the X-chromosome breakpoint were visualized and evaluated using Integrative Genomics Viewer (IGV) in search of chimeric inserts.

X-inactivation pattern evaluation

The ratio of the X-chromosome inactivation pattern of the Brazilian family was assessed by HUMARA (Human Androgen Receptor Assay) (47). To visualize if the normal or the derivative X-chromosome was inactivated we used EdU incorporation assay as described (48) and 93 metaphases analyzed were informative.

Cell lines and RNA isolation

We established lymphoblastoid cell lines (LCLs) from the Brazilian proband, her mother and ten unrelated Brazilian female controls, seven of which are carriers of balanced X-autosome translocations with different breakpoints (karyotypes described in **Supplementary Table S16**) by transforming peripheral blood mononuclear cells with Epstein-Barr virus (EBV). In brief, the mononuclear cell fractions were isolated from heparinized whole-blood samples by Ficoll-Paque Premium (GE Healthcare Bio-Sciences AB) centrifugation and incubated with EBV. Upon transformation, the cell lines were grown in RPMI 1640 medium (GIBCO) with addition of 10% fetal calf serum (PAA, GE Healthcare) and 1% penicillin-streptomycin (GIBCO), at 37°C and 5% CO_2, until stable and exponentially growing cell lines were established.

LCLs and whole blood RNA were extracted from cell pellets using standard Trizol protocol (Invitrogen) and the PAXgene Blood RNA MDx Kit (Qiagen-Sciences), respectively.

Quantitative RT-PCR

Gene expression analysis was performed by real-time quantitative PCR (RT-qPCR) for *AMMECR1, AMMECR1-like* and *RGAG1* genes using TaqMan assays (Life Technologies) and SYBR green primer pairs. *GAPDH* gene expression was used as an internal control. Primers and assays are detailed in **Supplementary Table S17**. The cDNA was synthetized using High Capacity cDNA Reverse Transcription (Life Technologies) and SuperScript III RT (Thermo Fisher Scientific) kits. All cDNA samples were assayed in technical triplicate with standard PCR conditions of the 96-Well Fast Thermal Cycling (Life Technologies) and Applied Biosystems 7900HT Sequence Detection System (Life Technologies). A qualitative analysis was performed to assess presence or absence of expression in proband and controls. The presence of expression in controls was considered only when the TaqMan assay showed amplification in all control individuals with a maximum cycle threshold (CT) value of 38 and maximum CT standard deviation of 1.0.

Exome sequencing and analysis

DNA from proband 2 was subjected to whole exome sequencing using the VCRome capture reagent in Human Genome Sequencing at Baylor College of Medicine according to the previously described protocol (49). Variants with

MAF<0.05% in controls cohorts (dbSNP, the 1000 genome project, NHLBI GO Exome Sequencing Project, the Exome Aggregation Consortium database and our in-house databases) and predicted to be deleterious by SIFT (30), PolyPhen-2 (34) and/or UMD-Predictor (49) were prioritized. Segregation of the *AMMECR1* variant in the family was confirmed by Sanger sequencing.

Western blot

Overexpressing GFP-AMMECR1 HEK-293 cells and LCLs from proband 1 and a control individual were lysed in 10mM HEPES pH7.9, 10 mM KCl, 0.1mM EDTA, 0.4% Nonidet P40 with protein inhibitors cocktail, kept on ice for 30 minutes vortexing every 10 minutes, centrifuged at 4°C for 1 minute at top speed. This first supernatant was used the cytoplasmic fraction. The remaining pellet was washed twice with the lysis buffer, resuspended in 20 mM HEPES pH 7.9, 0.4 M NaCl, 1.0 mM EDTA with protein inhibitors cocktail, shaken vigorously for 15 minutes and centrifuged at 4°C for 10 minutes at top speed. This second supernatant corresponded to the nuclear extract. After protein quantification by bicinchoninic acid assay (Bio-Rad), equal amounts of protein (30 µg) were resolved on 10% SDS–PAGE and wet-transferred to PVDV 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 (Abcam 155780) or anti-GFP (Santa Cruz sc-8334) primary antibodies diluted 1:300 and 1:500, respectively. After three washes, membranes were incubated with horseradish peroxidase-conjugated secondary antibodies diluted 1:10,000 (Santa Cruz Biotechnology) for 1 hour at room temperature. Lamin A/C (Santa Cruz sc-6215) and tubulin (Sigma Aldrich T5168) were used as nuclear and cytoplasmic markers, respectively. Immunoreactive bands were visualized and analyzed using the enhanced chemiluminescence detection reagents (Western highligting plus-ECL, Perkin Elmer), according to the manufacturer's instructions.

Morpholinos and Zebrafish Embryo Manipulations

Zebrafish animal experimentation was approved by the Ethical Committee for Animal Experimentation of the Geneva University Medical School and 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 (50).

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) using phenol red in 1– tracer in Danieau buffer. 8.2 ng of MO-E1I1, 1.26, 2.1, 4.2 and 8.4 ng of MO-I2E3, 8.4 ng of MO-E3I3 and 2.1 ng, 4.2 ng and 8.4 ng of Ctrl-MO were injected into wild-type zebrafish embryos, respectively. 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)$_{12-19}$ 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 (51).

Functional network analysis

We utilized GIANT (Genomescale Integrated Analysis of gene Networks in Tissues) to assess tissue-specific functional interactions (23). Interactions were evaluated in affected tissues/cells (heart, bone, osteoblast, cartilage,

chondrocyte), tissues/cells with high expression of *AMMECR1* according to GTEx (52) and Genevestigator (53) (skin, ovary, bone marrow, bronchial epithelial cells, hematopoietic stem cells, endothelium and placenta) and non-specific options, such as fetus, embryo and “all tissues”. *AMMECR1L* interactions were evaluated in heart, bone, osteoblast, cartilage, chondrocyte, fetus, embryo and “all tissues”. Enrichment of biological processes and cellular components was determined using Enrichr tool (54). In order to ensure the specify of the biological processes and cellular components enriched in AMMECR1 related genes, molecular interaction networks in the same tissues/cells were constructed for GAPDH and RGAG1 genes, and the resultant gene sets were submitted to Enrichr.

Expression of GFP-AMMECR1 in HeLa cells

pcDNA™6.2/EmGFP Gateway destination vector was used to generate GFP-*AMMECR1* construct with mammalian expression according to the instructions of the manufacturer (Invitrogen). The canonical *AMMECR1* transcript (ENST00000262844) was used to generate the vector. The construct was introduced into the human HeLa cell line by transient transfection using FuGENE HD Transfection Reagent following the manufacturer’s instructions (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

Animal housing, maintenance, and experimentation were approved by the local committee for animal care (Comissão de Ética no Uso de Animais da Universidade Federal de São Paulo/Hospital São Paulo N 9672261113). Balb/c mice were acquired from Center for Development of Experimental Models for Biology and Medicine (CEDEME), Universidade Federal de São Paulo. Animals were kept in plastic cages under a 12/12-hour light/dark cycle at 23 to 25 °C. Food and water were allowed ad libitum. The animals were mated overnight and pregnancy was detected at 0.5 gestation day (GD) by the observation of the mating plug. The postnatal females were euthanized by CO_2 inhalation at 9.5 GD and the embryos were collected. To avoid bias related to different developmental stages, only embryos with 16-18 somites were selected for the study.

Immediately after the euthanasia, embryos were fine dissected using a needle under a stereoscope and the prosencephalon, somites and heart were collected separately. Prior to RNA isolation, tissue from five to eight embryos were pooled in one tube to ensure a sufficient final RNA per sample. A total of 35 embryos were dissected and distributed in five pooled samples per tissue. RNA was isolated using TRIzol reagent (Thermo Fisher Scientific) according to the manufacturer's guidelines. RNA concentration and purity were assessed using NanoDrop ND-1000 Spectrophotometer (Thermo Fisher Scientific) and integrity was assessed by agarose gel electrophoresis. The cDNA was synthetized using High Capacity cDNA Reverse Transcription (Thermo Fisher Scientific) using 580 ng of total RNA of each sample. The cDNA was diluted 100-fold in RNAse free water for subsequent qRT-PCR analysis. *Ammecr1* and *Ammecr1l* gene expression analysis was performed by RT-qPCR with TaqMan assays (Life Technologies) and using *Actb*, *Gapdh* and *Tbp* as internal controls. Assays are detailed in Supplementary Table S17. RT-qPCR was performed using ViiA 7 Real-Time PCR System (Thermo Fisher Scientific). cDNA samples were assayed in technical triplicate in 384-Well PCR plates (Thermo Fisher Scientific) with standard conditions. *Ammecr1* and *Ammecr1l* relative expressions were calculated by the Relative Standard Curve Method, in which the Ct values were converted to cDNA quantities using a standard curve as a normalizer.

Funding

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 Fundação de Amparo à Pesquisa do Estado de São Paulo [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.

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.

Author contributions

MMO mapped the breakpoints and performed the molecular analyses with the help of GG, ADB, TL and NK. MMO, RJF and MNA modeled the disease in zebrafish. FP, VAM, MFS, LDK, FA, GP, TS, GC, JA and JAR phenotyped and genotyped individuals. MMO and AR wrote the manuscript. MIM and AR designed the study and obtained the necessary funding. All authors commented on and approved the manuscript.

Competing interest

The authors declare that they have no competing interests.

Abbreviations

Loss of function (lof)

Gene Ontology (GO)

Morpholino (MO)

Chromosome Analysis Suite (ChAS)

Fluorescence *in situ* hybridization (FISH)

Bacterial artificial chromosome (BAC)

Burrows-Wheeler Aligner (BWA-MEM)

Integrative Genomics Viewer (IGV)

Human Androgen Receptor Assay (HUMARA)

5-ethynyl-2′-deoxyuridine (EdU)

Lymphoblastoid cell line (LCL)

Epstein-Barr virus (EBV)

Real-time quantitative PCR (RT-qPCR)

Cycle threshold (CT)

Relative quantification (RQ)

The Web Resources section

Clustal Omega: http://www.ebi.ac.uk/Tools/msa/clustalo/

Cyclebase: http://www.cyclebase.org/

Enrichr: http://amp.pharm.mssm.edu/Enrichr/

ExAC: http://exac.broadinstitute.org/

GENEVESTIGATOR: https://genevestigator.com

gFeatBrowser: http://www.gfeatbrowser.com

GIANT: http://giant.princeton.edu/

GTEx: http://www.gtexportal.org/

Needle: http://www.ebi.ac.uk/Tools/psa/emboss_needle/

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

PomBase: http://www.pombase.org/

SGD Project: http://www.yeastgenome.org

SWISS-MODEL: http://swissmodel.expasy.org/interactive

ZFIN (ammecr1): https://zfin.org/ZDB-FIG-060710-830

References

1 Ogilvie, C.M., Braude, P. and Scriven, P.N. (2001) Successful pregnancy outcomes after preimplantation genetic diagnosis (PGD) for carriers of chromosome translocations. *Human fertility*, **4**, 168-171.
2 Clifford, K., Rai, R., Watson, H. and Regan, L. (1994) An informative protocol for the investigation of recurrent miscarriage: preliminary experience of 500 consecutive cases. *Human reproduction (Oxford, England)*, **9**, 1328-1332.
3 Patsalis, P.C., Evangelidou, P., Charalambous, S. and Sismani, C. (2004) Fluorescence in situ hybridization characterization of apparently balanced translocation reveals cryptic complex chromosomal rearrangements with unexpected level of complexity. *Eur J Hum Genet*, **12**, 647-653.
4 Warburton, D. (1984) Outcome of cases of de novo structural rearrangements diagnosed at amniocentesis. *Prenatal diagnosis*, **4 Spec No**, 69-80.

5 Curran, M.E., Atkinson, D.L., Ewart, A.K., Morris, C.A., Leppert, M.F. and Keating, M.T. (1993) The elastin gene is disrupted by a translocation associated with supravalvular aortic stenosis. *Cell*, **73**, 159-168.
6 Lupianez, D.G., Kraft, K., Heinrich, V., Krawitz, P., Brancati, F., Klopocki, E., Horn, D., Kayserili, H., Opitz, J.M., Laxova, R. *et al.* (2015) Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions. *Cell*, **161**, 1012-1025.
7 Harewood, L., Schutz, F., Boyle, S., Perry, P., Delorenzi, M., Bickmore, W.A. and Reymond, A. (2010) The effect of translocation-induced nuclear reorganization on gene expression. *Genome research*, **20**, 554-564.
8 Ricard, G., Molina, J., Chrast, J., Gu, W., Gheldof, N., Pradervand, S., Schutz, F., Young, J.I., Lupski, J.R., Reymond, A. *et al.* (2010) Phenotypic consequences of copy number variation: insights from Smith-Magenis and Potocki-Lupski syndrome mouse models. *PLoS biology*, **8**, e1000543.
9 Talkowski, M.E., Ernst, C., Heilbut, A., Chiang, C., Hanscom, C., Lindgren, A., Kirby, A., Liu, S., Muddukrishna, B., Ohsumi, T.K. *et al.* (2011) Next-generation sequencing strategies enable routine detection of balanced chromosome rearrangements for clinical diagnostics and genetic research. *American journal of human genetics*, **88**, 469-481.
10 Schluth-Bolard, C., Labalme, A., Cordier, M.P., Till, M., Nadeau, G., Tevissen, H., Lesca, G., Boutry-Kryza, N., Rossignol, S., Rocas, D. *et al.* (2013) Breakpoint mapping by next generation sequencing reveals causative gene disruption in patients carrying apparently balanced chromosome rearrangements with intellectual deficiency and/or congenital malformations. *Journal of medical genetics*, **50**, 144-150.
11 Kalz-Fuller, B., Sleegers, E., Schwanitz, G. and Schubert, R. (1999) Characterisation, phenotypic manifestations and X-inactivation pattern in 14 patients with X-autosome translocations. *Clinical genetics*, **55**, 362-366.
12 Leppig, K.A. and Disteche, C.M. (2001) Ring X and other structural X chromosome abnormalities: X inactivation and phenotype. *Seminars in reproductive medicine*, **19**, 147-157.
13 Waters, J.J., Campbell, P.L., Crocker, A.J. and Campbell, C.M. (2001) Phenotypic effects of balanced X-autosome translocations in females: a retrospective survey of 104 cases reported from UK laboratories. *Human genetics*, **108**, 318-327.
14 Gartler, S.M. and Sparkes, R.S. (1963) The Lyon-Beutler Hypothesis and Isochromosome X Patients with Turner Syndrome. *Lancet*, **2**, 411.
15 Mattei, M.G., Mattei, J.F., Ayme, S. and Giraud, F. (1982) X-autosome translocations: cytogenetic characteristics and their consequences. *Human genetics*, **61**, 295-309.
16 Cantagrel, V., Lossi, A.M., Boulanger, S., Depetris, D., Mattei, M.G., Gecz, J., Schwartz, C.E., Van Maldergem, L. and Villard, L. (2004) Disruption of a new X linked gene highly expressed in brain in a family with two mentally retarded males. *Journal of medical genetics*, **41**, 736-742.
17 Kalscheuer, V.M., Tao, J., Donnelly, A., Hollway, G., Schwinger, E., Kubart, S., Menzel, C., Hoeltzenbein, M., Tommerup, N., Eyre, H. *et al.* (2003) Disruption of the serine/threonine kinase 9 gene causes severe X-linked infantile spasms and mental retardation. *American journal of human genetics*, **72**, 1401-1411.
18 Najm, J., Horn, D., Wimplinger, I., Golden, J.A., Chizhikov, V.V., Sudi, J., Christian, S.L., Ullmann, R., Kuechler, A., Haas, C.A. *et al.* (2008) Mutations of CASK cause an X-linked brain malformation phenotype with microcephaly and hypoplasia of the brainstem and cerebellum. *Nature genetics*, **40**, 1065-1067.
19 Carrel, L. and Willard, H.F. (2005) X-inactivation profile reveals extensive variability in X-linked gene expression in females. *Nature*, **434**, 400-404.

20 Cotton, A.M., Price, E.M., Jones, M.J., Balaton, B.P., Kobor, M.S. and Brown, C.J. (2015) Landscape of DNA methylation on the X chromosome reflects CpG density, functional chromatin state and X-chromosome inactivation. *Human molecular genetics*, **24**, 1528-1539.
21 Consortium, E.P. (2012) An integrated encyclopedia of DNA elements in the human genome. *Nature*, **489**, 57-74.
22 Balaji, S. and 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 research*, **35**, 5658-5671.
23 Greene, C.S., Krishnan, A., Wong, A.K., Ricciotti, E., Zelaya, R.A., Himmelstein, D.S., Zhang, R., Hartmann, B.M., Zaslavsky, E., Sealfon, S.C. *et al.* (2015) Understanding multicellular function and disease with human tissue-specific networks. *Nature genetics*, **47**, 569-576.
24 Chetaille, P., Preuss, C., Burkhard, S., Cote, J.M., Houde, C., Castilloux, J., Piche, J., Gosset, N., Leclerc, S., Wunnemann, F. *et al.* (2014) Mutations in SGOL1 cause a novel cohesinopathy affecting heart and gut rhythm. *Nature genetics*, **46**, 1245-1249.
25 Golzio, C., Willer, J., Talkowski, M.E., Oh, E.C., Taniguchi, Y., Jacquemont, S., Reymond, A., Sun, M., Sawa, A., Gusella, J.F. *et al.* (2012) KCTD13 is a major driver of mirrored neuroanatomical phenotypes of the 16p11.2 copy number variant. *Nature*, **485**, 363-367.
26 Payne, F., Colnaghi, R., Rocha, N., Seth, A., Harris, J., Carpenter, G., Bottomley, W.E., Wheeler, E., Wong, S., Saudek, V. *et al.* (2014) Hypomorphism in human NSMCE2 linked to primordial dwarfism and insulin resistance. *The Journal of clinical investigation*, **124**, 4028-4038.
27 Stewart, A.M., Braubach, O., Spitsbergen, J., Gerlai, R. and Kalueff, A.V. (2014) Zebrafish models for translational neuroscience research: from tank to bedside. *Trends in neurosciences*, **37**, 264-278.
28 Kok, F.O., Shin, M., Ni, C.W., Gupta, A., Grosse, A.S., van Impel, A., Kirchmaier, B.C., Peterson-Maduro, J., Kourkoulis, G., Male, I. *et al.* (2015) Reverse genetic screening reveals poor correlation between morpholino-induced and mutant phenotypes in zebrafish. *Developmental cell*, **32**, 97-108.
29 Jonsson, J.J., Renieri, A., Gallagher, P.G., Kashtan, C.E., Cherniske, E.M., Bruttini, M., Piccini, M., Vitelli, F., Ballabio, A. and Pober, B.R. (1998) Alport syndrome, mental retardation, midface hypoplasia, and elliptocytosis: a new X linked contiguous gene deletion syndrome? *Journal of medical genetics*, **35**, 273-278.
30 Vaser, R., Adusumalli, S., Leng, S.N., Sikic, M. and Ng, P.C. (2016) SIFT missense predictions for genomes. *Nature protocols*, **11**, 1-9.
31 Vitelli, F., Piccini, M., Caroli, F., Franco, B., Malandrini, A., Pober, B., Jonsson, J., Sorrentino, V. and 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**, 335-340.
32 Gazou, A., Riess, A., Grasshoff, U., Schaferhoff, K., Bonin, M., Jauch, A., Riess, O. and Tzschach, A. (2013) Xq22.3-q23 deletion including ACSL4 in a patient with intellectual disability. *American journal of medical genetics. Part A*, **161A**, 860-864.
33 Yang, Y., Muzny, D.M., Reid, J.G., Bainbridge, M.N., Willis, A., Ward, P.A., Braxton, A., Beuten, J., Xia, F., Niu, Z. *et al.* (2013) Clinical whole-exome sequencing for the diagnosis of mendelian disorders. *The New England journal of medicine*, **369**, 1502-1511.
34 Adzhubei, I., Jordan, D.M. and Sunyaev, S.R. (2013) Predicting functional effect of human missense mutations using PolyPhen-2. *Current protocols in human genetics*, **Chapter 7**, Unit7 20.

35 Meloni, I., Vitelli, F., Pucci, L., Lowry, R.B., Tonlorenzi, R., Rossi, E., Ventura, M., Rizzoni, G., Kashtan, C.E., Pober, B. *et al.* (2002) Alport syndrome and mental retardation: clinical and genetic dissection of the contiguous gene deletion syndrome in Xq22.3 (ATS-MR). *Journal of medical genetics*, **39**, 359-365.
36 Basel-Vanagaite, L., Pillar, N., Isakov, O., Smirin-Yosef, P., Lagovsky, I., Orenstein, N., Salmon-Divon, M., Tamary, H., Zaft, T., Bazak, L. *et al.* (2017) X-linked elliptocytosis with impaired growth is related to mutated AMMECR1. *Gene*, **606**, 47-52.
37 Andreoletti, G., Seaby, E.G., Dewing, J.M., O'Kelly, I., Lachlan, K., Gilbert, R.D. and Ennis, S. (2016) AMMECR1: a single point mutation causes developmental delay, midface hypoplasia and elliptocytosis. *Journal of medical genetics*, in press.
38 Tan, S., Ding, K., Li, R., Zhang, W., Li, G., Kong, X., Qian, P., Lobie, P.E. and Zhu, T. (2014) Identification of miR-26 as a key mediator of estrogen stimulated cell proliferation by targeting CHD1, GREB1 and KPNA2. *Breast cancer research : BCR*, **16**, R40.
39 Tsai, Z.Y., Chou, C.H., Lu, C.Y., Singh, S., Yu, S.L. and Li, S.S. (2011) Proteomic comparison of human embryonic stem cells with their differentiated fibroblasts: Identification of 206 genes targeted by hES cell-specific microRNAs. *The Kaohsiung journal of medical sciences*, **27**, 299-306.
40 Abdelmohsen, K., Srikantan, S., Yang, X., Lal, A., Kim, H.H., Kuwano, Y., Galban, S., Becker, K.G., Kamara, D., de Cabo, R. *et al.* (2009) Ubiquitin-mediated proteolysis of HuR by heat shock. *The EMBO journal*, **28**, 1271-1282.
41 Kittler, R., Pelletier, L., Heninger, A.K., Slabicki, M., Theis, M., Miroslaw, L., Poser, I., Lawo, S., Grabner, H., Kozak, K. *et al.* (2007) Genome-scale RNAi profiling of cell division in human tissue culture cells. *Nature cell biology*, **9**, 1401-1412.
42 Santos, A., Wernersson, R. and Jensen, L.J. (2015) Cyclebase 3.0: a multi-organism database on cell-cycle regulation and phenotypes. *Nucleic acids research*, **43**, D1140-1144.
43 Marguerat, S., Schmidt, A., Codlin, S., Chen, W., Aebersold, R. and Bahler, J. (2012) Quantitative analysis of fission yeast transcriptomes and proteomes in proliferating and quiescent cells. *Cell*, **151**, 671-683.
44 Cherry, J.M., Hong, E.L., Amundsen, C., Balakrishnan, R., Binkley, G., Chan, E.T., Christie, K.R., Costanzo, M.C., Dwight, S.S., Engel, S.R. *et al.* (2012) Saccharomyces Genome Database: the genomics resource of budding yeast. *Nucleic acids research*, **40**, D700-705.
45 Liehr, T., Heller, A., Starke, H. and Claussen, U. (2002) FISH banding methods: applications in research and diagnostics. *Expert review of molecular diagnostics*, **2**, 217-225.
46 Moyses-Oliveira, M., Guilherme, R.S., Meloni, V.A., Di Battista, A., de Mello, C.B., Bragagnolo, S., Moretti-Ferreira, D., Kosyakova, N., Liehr, T., Carvalheira, G.M. *et al.* (2015) X-linked intellectual disability related genes disrupted by balanced X-autosome translocations. *American journal of medical genetics. Part B, Neuropsychiatric genetics : the official publication of the International Society of Psychiatric Genetics*, **168**, 669-677.
47 Hickey, T., Chandy, A. and Norman, R.J. (2002) The androgen receptor CAG repeat polymorphism and X-chromosome inactivation in Australian Caucasian women with infertility related to polycystic ovary syndrome. *The Journal of clinical endocrinology and metabolism*, **87**, 161-165.
48 Sisdelli, L., Vidi, A.C., Moyses-Oliveira, M., Di Battista, A., Bortolai, A., Moretti-Ferreira, D., Dias da Silva, M.R., Melaragno, M.I. and 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. *Human genetics*, **135**, 185-192.
49 Salgado, D., Desvignes, J.P., Rai, G., Blanchard, A., Miltgen, M., Pinard, A., Levy, N., Collod-Beroud, G. and Beroud, C. (2016) UMD-Predictor: A High-Throughput Sequencing Compliant System for Pathogenicity Prediction of any Human cDNA Substitution. *Human mutation*, **37**, 439-446.
50 Kimmel, C.B., Ballard, W.W., Kimmel, S.R., Ullmann, B. and Schilling, T.F. (1995) Stages of embryonic development of the zebrafish. *Developmental dynamics : an official publication of the American Association of Anatomists*, **203**, 253-310.
51 Livak, K.J. and Schmittgen, T.D. (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods*, **25**, 402-408.
52 Carithers, L.J. and Moore, H.M. (2015) The Genotype-Tissue Expression (GTEx) Project. *Biopreservation and biobanking*, **13**, 307-308.
53 Hruz, T., Laule, O., Szabo, G., Wessendorp, F., Bleuler, S., Oertle, L., Widmayer, P., Gruissem, W. and Zimmermann, P. (2008) Genevestigator v3: a reference expression database for the meta-analysis of transcriptomes. *Advances in bioinformatics*, **2008**, 420747.
54 Chen, E.Y., Tan, C.M., Kou, Y., Duan, Q., Wang, Z., Meirelles, G.V., Clark, N.R. and Ma'ayan, A. (2013) Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC bioinformatics*, **14**, 128.
55 Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Soding, J. *et al.* (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular systems biology*, **7**, 539.
56 McWilliam, H., Li, W., Uludag, M., Squizzato, S., Park, Y.M., Buso, N., Cowley, A.P. and Lopez, R. (2013) Analysis Tool Web Services from the EMBL-EBI. *Nucleic acids research*, **41**, W597-600.
57 Arnold, K., Bordoli, L., Kopp, J. and Schwede, T. (2006) The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics*, **22**, 195-201.

Figure Legend

Figure 1: Description of proband 1's phenotype. Proband's facial features in front **(a)** and lateral view **(b)** at 6 years of age. Proband's skeleton X-ray showing skull's wormian bones (white arrow) **(c)**, thoracolumbar scoliosis and enlarged heart **(d)**, hyperdense epiphyses and metaphyses in the hands (white arrows) **(e)** and feet (white arrows) **(f)**. Further pictures of proband's features are presented in **Supplementary Figure S1**.

Figure 2: Cytogenomic characterization of the rearrangement. Partial G-banded karyotype **(a)** and ideograms **(b)** of the chromosomes involved in the translocation. The localization of BAC CTD-3066N24 is indicated with red boxes. **(c)** FISH of metaphase cells using BAC CTD-3066N24 to validate the chromosome X breakpoint localization, showing three hybridization signals, including split signals in the derivative chromosomes. **(d)** FISH of metaphase cells with 9p12 (yellow), 9q13 (yellow), 9p11.2 (magenta), chromosome 9 centromere (cen 9: green), 9q12 (magenta and blue), and 9q21.1 (red) probes to determine the autosomal breakpoint at 9q12. The localization of these probes is indicated on the right along an ideogram of chromosome 9.

Figure 3: *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 light blue line). The gene transcripts (blue), the H3K4me3 and H3K27ac histone marks (both blue) according to ENCODE (21) 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 proband 1 (P). Note the absence of *AMMECR1* expression in the proband 1. Student's t-test was performed to compare gene expression in proband 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 proband (P) incubated with anti-AMMECR1/AMMECR1L (top left panel) and anti-lamin A/C (bottom left panel) as normalizer. The level of AMMECR1L alone was higher in cells of proband 1 than AMMECR1 and

AMMECR1L together in cells from a control individual (right panel). More details are presented in Supplementary Figure S6.

Figure 4: *ammecr1* knockdown in zebrafish results in phenotypic features reminiscent of proband 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.

Supplementary Figure S1: Description of proband 1's phenotype. Proband's facial features in front (left panels) and lateral view (right panles) at 3 **(a)**, 6 **(b)** and 8 **(c)** years of age. Proband's dysmorphisms of hands **(d)** and feet **(e)**. Proband's skeleton X-ray showing enlarged heart, toraco-lombar scoliosis and looser zone in the pelvis **(f)**, sacrum flattening **(g)**, left lower limb shortening and coxa valga **(h)**, growth recovery lines in tibia bilaterally and hyperdense metaphysis and epiphysis of the long bones **(i)**.

Supplementary Figure S2: Integrative Genomics Viewer (IGV) view of the chimeric inserts. (a) IGV screenshot of whole genome sequencing pair-end reads of proband 1. The X-chromosome breakpoint region shows 11 reads belonging to chimeric inserts (dark grey) and 2 split reads (red). The sequence of the 2 split reads (reads 1 and 2, red arrows) pinpoints the der(9) junction point sequence. The vertical dashed line indicates the chromosomes X breakpoint. The full nucleotide sequences of the reads numbered 1 to 4 are presented in **Supplementary Table S2**. **(b)** Sanger sequencing electropherogram of the der(X) junction point. The breakpoint is indicated by an arrowhead. **(c)** Alignment

of the reads 1 and 2 and the mate of read 4 from panel **a** to the chromosome X reference and Sanger sequencing of the der(X) junction point. Sequences from the der(9) and der(X) junction points aligned with the chromosome X reference sequence and Sanger sequencing of the der(X) junction point. Sequences from chromosome X and 9 are in red and black, respectively. The chrX 8bp deletion of proband 1 and the sequence from the mate of read 4 that aligns with the der(X) junction point are highlighted in bold.

Supplementary Figure S3: Skewed X-chromosome inactivation. (a) HUMARA capillary electrophoresis showing different peaks pattern between undigested (top panel) and digested samples (bottom panel). Digested samples show only the major allele peak (231 bp), a typical extremely skewed inactivation pattern. **(b)** Proband 1's EdU metaphase showing the normal X-chromosome inactive (Xi) and the der(X) active [der(X)a]. X-chromosomes were marked with X-centromere probe (red signal).

Supplementary Figure S4: AMMECR1 protein conservation. (a) AMMECR1 protein conservation in vertebrates and AMMECR1 domain-containing protein conservation in eukaryotes, bacteria and archea. Multialignment of AMMECR1 from 17 vertebrates (*Homo sapiens* NP_056180.1, *Macaca mulatta* XP_014983496.1, *Mus musculus* NP_062369, *Bos Taurus* XP_015325851.1, *Canis lupus familiaris* XP_549183.4, *Myotis lucifugus* XP_006082298.1, *Dasypus novemcinctus* XP_004471886.2, *Monodelphis domestica* XP_001375926.2, *Gallus gallus* XP_015134088.1, *Chrysemys picta belli* XP_005291870.1, *Danio rerio* NP_956875.1, *Cynoglossus semilaevis* XP_008330387.1, *Astyanax mexicanus* XP_007247120.1, *Larimichthys crocea* KKF29168.1, *Oncorhynchus mykiss* CDQ63682.1, *Anolis carolinensis* XP_008104682.1 and *Xenopus tropicalis* NP_001116966.1) and AMMECR1 domain-containing proteins from *Drosophila melanogaster* Q9VCF0, *Caenorhabditis elegans* NP_496270.1, *Saccharomyces cerevisiae* YOR289W, *Trypanosoma grayi* XP_009307024.1, *Arabidopsis thaliana* NP_001031509.1, *Zea mays* NP_001143656.1, three bacteria (*Treponema caldarium* AEJ19136.1, *Rhodospirillum rubrum* ABC21252.1 and *Thermotoga* ACB09558.1) and three archea (*Methanobrevibacter smithii* EFC93356.1, *Thermofilum carboxyditrophus* AJB41163.1 and *Halopiger xanaduensis* AEH36405.1) using Clustal Omega (1.2.1) (55) with default parameters. **(b)** Comparison of the sequences of human AMMECR1 and AMMECR1L proteins. Protein sequences were aligned using Needle (56) with default parameters.

Supplementary Figure S5: AMMECR1 nuclear localization. (a) DAPI, **(b)** GFP and **(c)** merged signals of HeLa cells transfected with plasmids encoding the GFP-AMMECR1 fusion protein.

Supplementary Figure S6. AMMECR1 and AMMECR1L protein expression pattern in nuclear and cytoplamatic compartments. (a) Representative image of western blot in HEK-293 cells overexpressing GFP-AMMECR1 (27 + 33 kDa, GFP-A) incubated with anti-GFP antibody. Tubulin (55 kDa, T) and Lamin A/C (65/72 kDa, L) were used as loading controls of cytoplasmatic (C) and nuclear (N) extracts, respectively. **(b)** Nuclear/cytoplasmatic expression ratio is plotted for the monomer (60 kDa) and the dimer (120 kDa) of GFP-AMMECR1. **(c)** Dimer/Monomer relative expression ratio is plotted for nuclear (N) and cytoplasmatic (C) compartments. **(d)** Representative image of western blot in LCL from a control individual (Ctrl) and patient with the balanced translocation t(X;9) (P (t X;9)) incubated with anti-AMMECR1/AMMECR1L (33/31 kDa, A/AL) antibody. AMMECR1/AMMECR1L expression is detected only in the nuclear compartment and AMMECR1/AMMECR1L is more frequently observed as a dimer than as a monomer. **(e)** AMMECR1/AMMECR1L monomer was observed when the membrane had the strong signal from AMMECR1/AMMECR1L dimer removed (cut at 40 kDa) and was overexposed. **(f)** AMMECR1/AMMECR1L dimer relative expression in the nuclear extract in control's and patient's cells. The higher expression in the patient can be related to a compensatory mechanism with the overexpression of AMMECR1L to attenuate the consequences of AMMECR1 LoF.

Supplementary Figure S7: Co-expression-based molecular interactions. Tissue-specific functional interaction network built using GIANT (Genome-scale Integrated Analysis of gene Networks in Tissues, http://giant.princeton.edu/) in "all tissues" (top panel), bone (central panel) and cartilage (bottom panel) with minimum relationship confidence = 0.6. Genes highlighted in blue in the top panel appeared as *AMMECR1* interacting gene in more than one queried tissue.

Supplementary Figure S8: *ammecr1* MO-I2E3 knockdown strategy. (a) Schematic representation of the exon-intron structure of the zebrafish *ammecr1* gene. The site targeted by MO-I2E3 (red bar), the impact of MO on the transcript (i.e. exon 3 skipping and frameshift; red lines) and the position of primers pairs E1 F – E4 R (red arrows) and E2/E3 F and E3/E4 R (green arrows) are indicated. **(b)** From left to right, RT-PCR with *ammecr1* E1 F – E4 R primers in non-

injected, standard control MO and MO-I2E3 injected embryos. The 354 bp and the 239 bp bands correspond to the amplimers from the wild type transcript and the transcript lacking exon 3, respectively. **(c)** From left to right, RT-qPCR-determined expression levels of wild type *ammecr1* transcript with E2/E3 F – E3/E4 R primers in non-injected, standard control MO and 2.1 ng, 4.2 ng and 8.4 ng of MO-I2E3 injected embryos. The silencing of the *ammecr1* wild type transcript is correlated to the dose of *ammecr1* MO-I2E3.

Supplementary Figure S9: MO-I2E3 dose increase and phenotype. From top to bottom, lateral views of representative 3 dpf embryos injected with 8.4 ng of standard control MO and decreasing doses of ammecr1 MO-I2E3 (8.4ng, 4.2 ng, 2.1 ng and 1.26 ng, respectively). Note the correlation between the dose of MO injected and the severity of the phenotype.

Figure S10: *ammecr1* MO-E3I3 knockdown strategy. (a) Schematic representation of the exon-intron structure of the zebrafish *ammecr1* gene. The site targeted by MO-E3I3 (red bar), the impact of MO on the transcript (i.e. exon 3 skipping or new splice site in exon 3; red lines and the position of primers pairs E1/E2 F – E4 R (red arrows), E2/E3 F and E3/E4 R (green arrows) and E2/E3 F and E4 R (yellow arrows) are indicated. **(b)** From left to right, RT-PCR with *ammecr1* E1/E2 F – E4 R primers in non-injected, standard control MO and MO-E3I3 injected embryos. The 314 bp, 290 bp and 199 bp bands correspond to the amplimers from the wild type transcript, the transcript lacking the last 8 codons of exon 3 and the transcript lacking exon 3, respectively. **(c)** Left panel**,** RT-qPCR-determined expression levels of the wild type *ammecr1* transcript with E2/E3 F – E3/E4 R primers in non-injected, standard control MO and MO-E3I3-injected embryos. Right panel, RT-qPCR-determined expression levels of full length transcript and transcript lacking the last 8 codons of exon 3 using E2/E3 F – E4 R primers in non-injected, standard control MO and MO-E3I3-injected embryos. **(d)** Comparison of 3D models of full-length *ammecr1* zebrafish protein (left) and *ammecr1* protein lacking the last 8 amino-acids encoded by exon 3. Red arrows indicate the differences in protein structure. The shorter peptide lacks one of the RAGNYA domains and is thus predicted to functionally impaired. Models were built in SWISS-MODEL Workspace (57).

Figure S11: *ammecr1* MO-E1I1 knockdown strategy. (a) Schematic representation of the exon-intron structure of the zebrafish *ammecr1* gene. The site targeted by MO-E1I1 (red bar), the impact of MO on the transcript (i.e. new

splice site in exon 1; red lines), and primers pair E1 F – E2 R (red arrows) are indicated. **(b)** From left to right, RT-PCR with *ammecr1* E1 F – E2 R primers in non-injected, standard control MO and MO-E1I1 injected embryos. The 102 bp and 84 bp bands correspond to the amplimers from the wild type transcript and the transcript lacking the last 6 codons of exon 1, respectively. **(c)** Comparison of 3D models of full-length *ammecr1* zebrafish protein (left) and *ammecr1* protein lacking the last 6 amino-acids encoded by exon 1. Red arrows indicate the differences in protein structure. Note the modified external fold highlighted in grey. Models were built in SWISS-MODEL Workspace (57). **(d)** Lateral views of representative 3 dpf embryos injected with 8.4 ng of standard control MO and MO-E1I1. No morphological differences were observed.

Supplementary Figure S12. *Ammecr1* and *Ammecr1l* expression in early mouse embryos. (a) 9.5 GD mouse embryo with the dissected structures schematically marked: heart in pink, forebrain in gray and somites in blue. **(b)** Relative expression of *Ammecr1* and *Ammecr1l* in heart (red), prosencephalon (black) and somite (blue).

Figure S13: CNVs in Xq22.3q23 region. UCSC genome browser view of chromosome X region 108.0-110.6Mb showing the localization of the deletions reported in patients with AMME syndrome (horizontal red bars) and the transcripts of the genes mapping to this region (blue). Note that the patients described in Jonsson et al. (1998), and Gazou et al (2013) have CNVs that encompass *AMMECR1* gene.

List of Tables

Table 1. Clinical description of the patients with variants involving *AMMECR1*.

Supplementary Table S1: Clinical features of the proband carrying a translocation t(X;9).

Supplementary Table S2: Nucleotide sequence of the reads nubered according to Figure S2 and the reference sequence for chrX:110318645-110318862 (hg38) interval.

Supplementary Table S3: List of the 77 genes interacting directly or with a maximum of one node with *AMMECR1* according to GIANT.

Supplementary Table S4: GO Biological Process terms enriched in *AMMECR1* functional network according to Enrichr.

Supplementary Table S5: GO Cellular Component terms enriched in *AMMECR1* functional network according to Enrichr.

Supplementary Table S6: List of 88 genes interacting directly or with a maximum of one node with AMMECR1L according to GIANT.

Supplementary Table S7: GO Biological Process terms enriched in *AMMECR1L* functional network according to Enrichr.

Supplementary Table S8: GO Cellular Component terms enriched in *AMMECR1L* functional network according to Enrichr.

Supplementary Table S9: List of the 49 genes interacting directly or with a maximum of one node with *GAPDH* according to GIANT.

Supplementary Table S10: GO Biological Process terms enriched in *GAPDH* functional network according to Enrichr.

Supplementary Table S11: GO Cellular Component terms enriched in *GAPDH* functional network according to Enrichr.

Supplementary Table S12: List of the 101 genes interacting directly or with a maximum of one node with *RGAG1* according to GIANT.

Supplementary Table S13: GO Biological Process terms enriched in *RGAG1* functional network according to Enrichr.

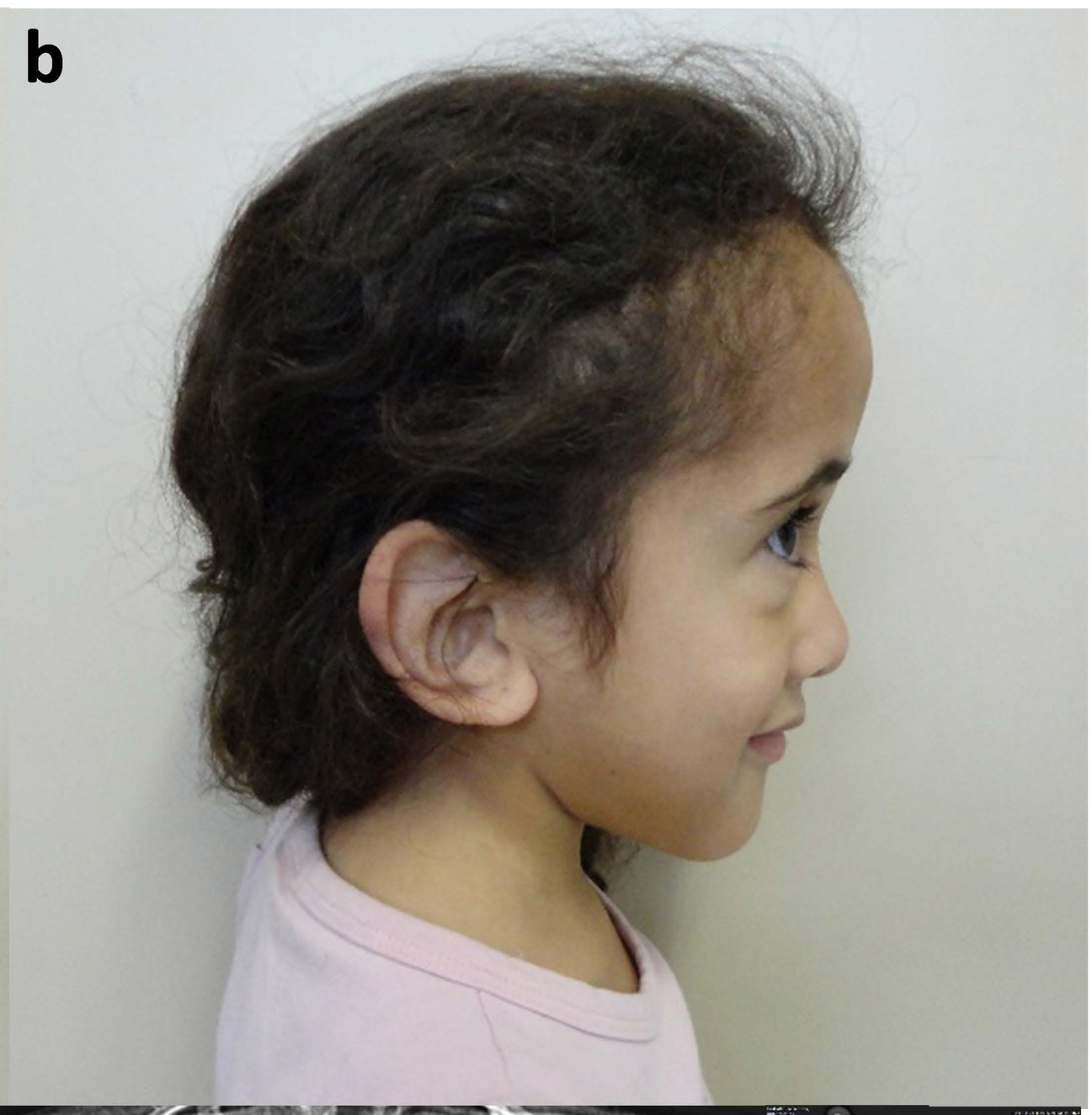
Supplementary Table S14: GO Cellular Component terms enriched in RGAG1 functional network according to Enrichr.

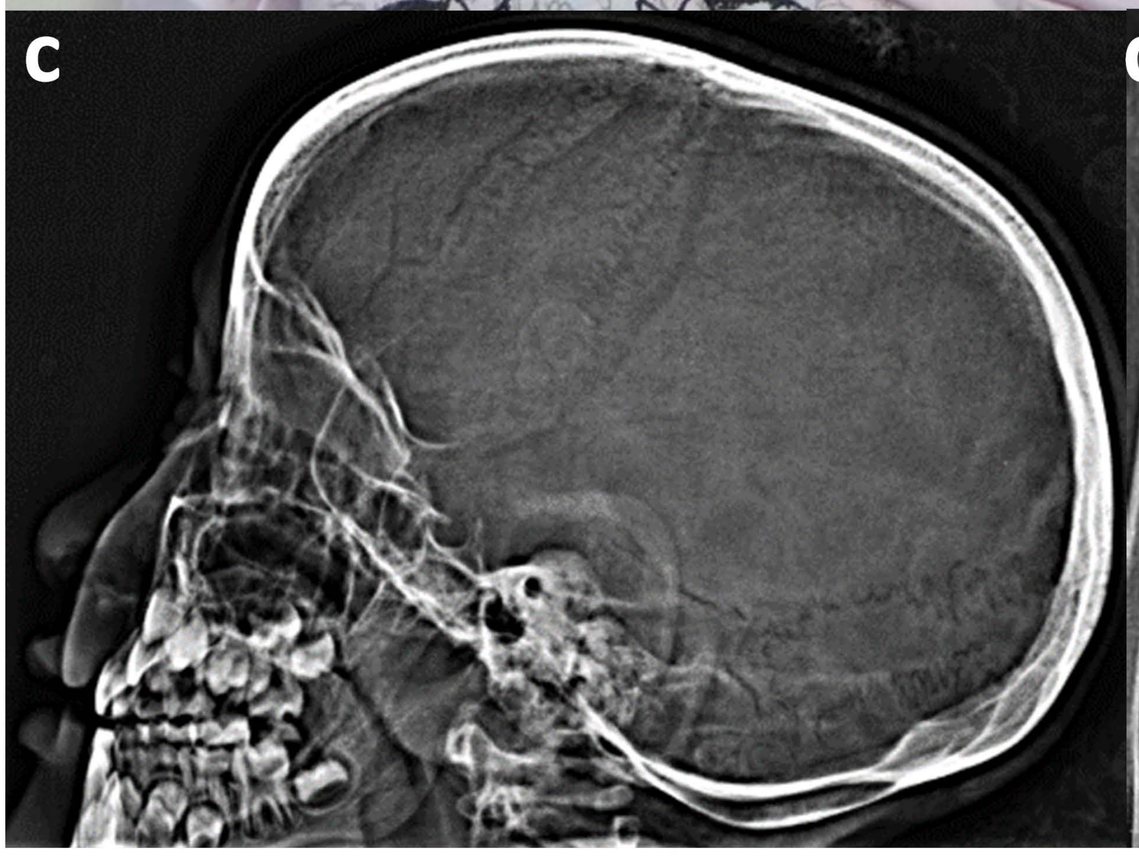
Supplementary Table S15. General clinical features observed in patients with AMMECR1 loss of function (LoF), AMME deletion syndrome, missense variants and in-frame transcript deletion in AMMECR1.

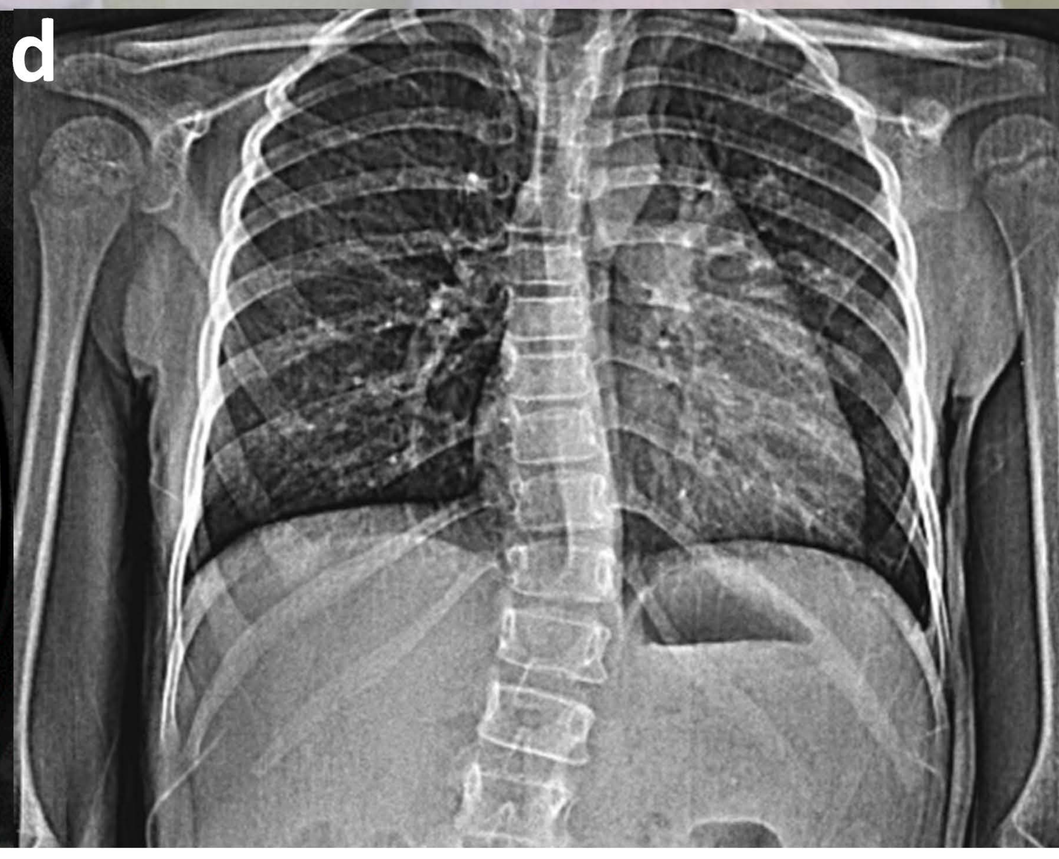
Supplementary Table S16: Karyotypes of lymphoblastoid cell lines used in this study

Supplementary Table S17: Sequence of the primers and PCR assays used in this study.


a

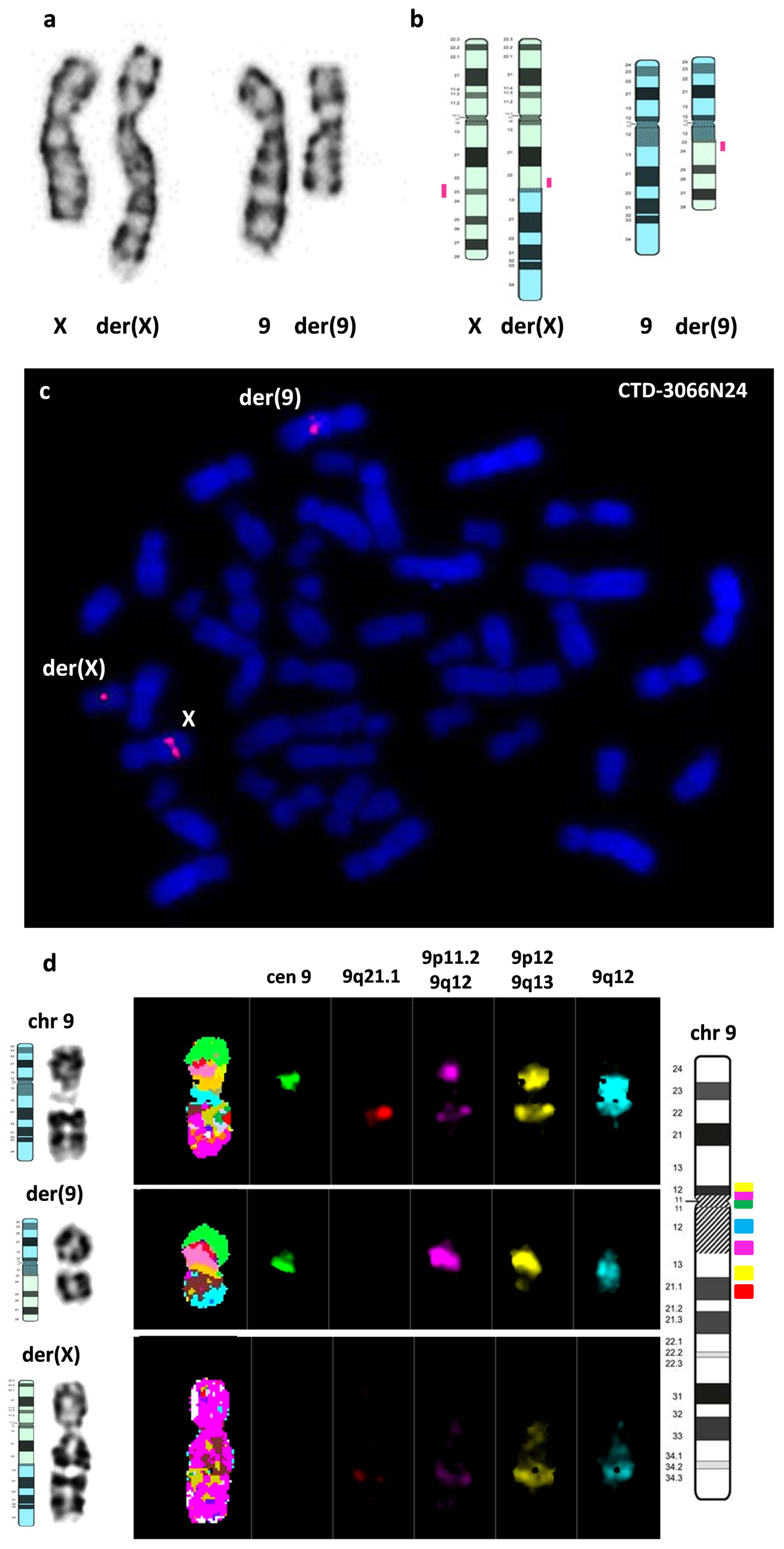

b

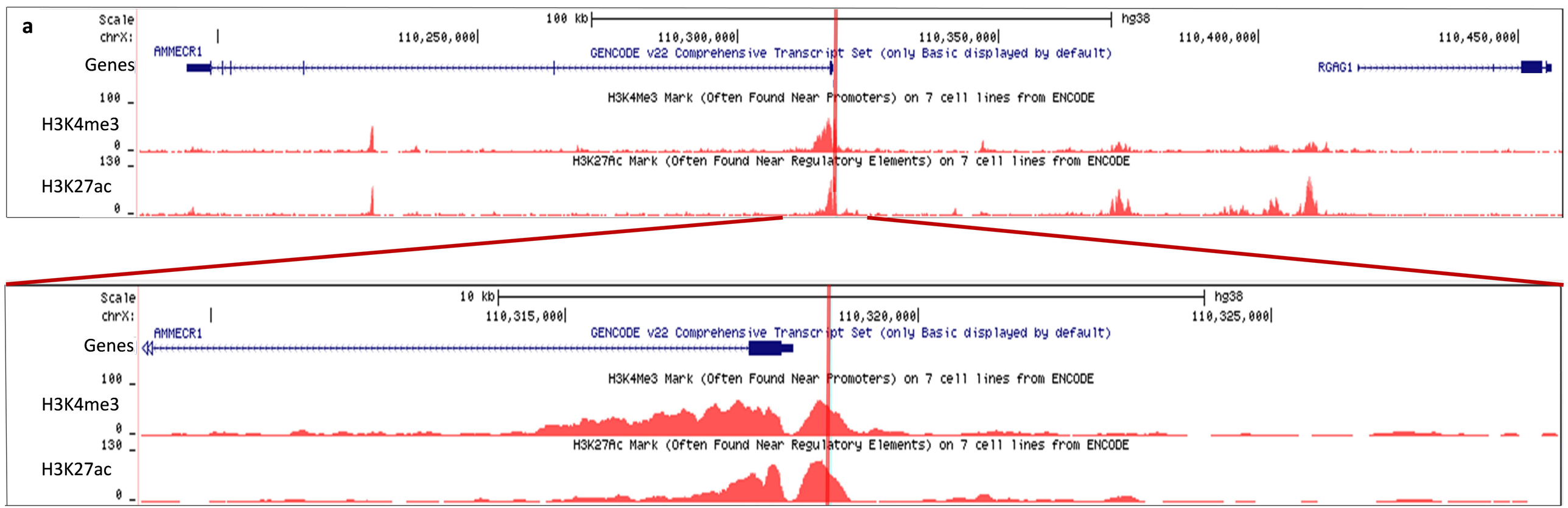

c

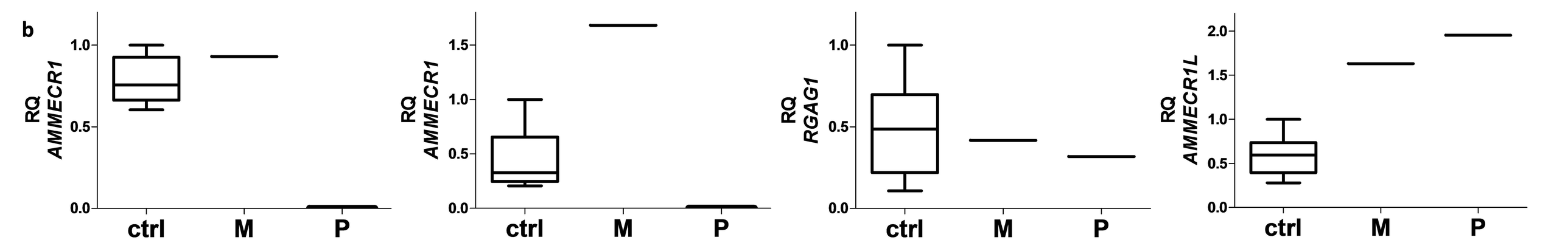

d


e

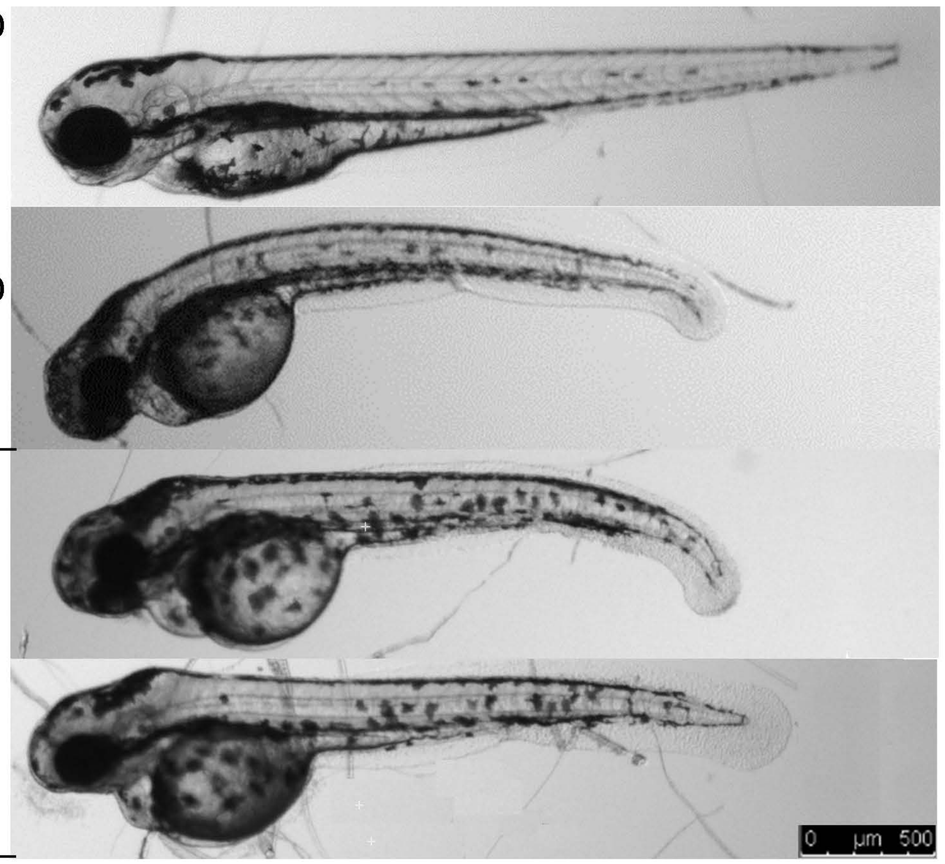

f


a
X
der(X)
9
der(9)
b
X
der(X)
9
der(9)
c
der(9)
CTD-3066N24
der(X)
X
d
cen 9
9q21.1
9p11.2
9q12
9p12
9q13
9q12
chr 9
der(9)
der(X)
chr 9

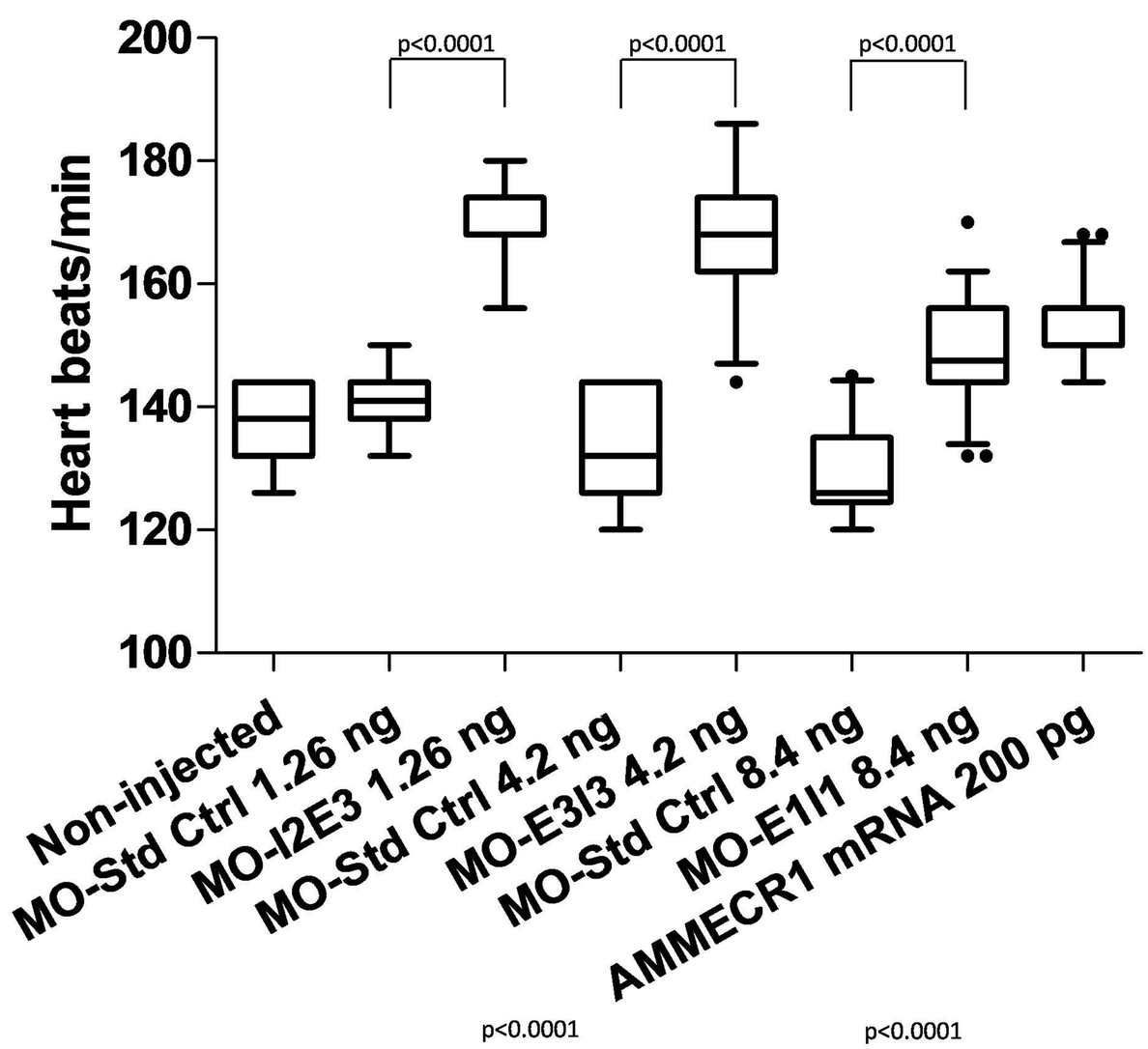

a
Scale
chrX:
100 kb
hg38
110,250,000
110,300,000
110,350,000
110,400,000
110,450,000
AMMECR1
GENCODE v22 Comprehensive Transcript Set (only Basic displayed by default)
Genes
RGAG1
100
H3K4Me3 Mark (Often Found Near Promoters) on 7 cell lines from ENCODE
H3K4me3
0
130
H3K27Ac Mark (Often Found Near Regulatory Elements) on 7 cell lines from ENCODE
H3K27ac
0
Scale
chrX:
10 kb
hg38
110,315,000
110,320,000
110,325,000
AMMECR1
GENCODE v22 Comprehensive Transcript Set (only Basic displayed by default)
Genes
100
H3K4Me3 Mark (Often Found Near Promoters) on 7 cell lines from ENCODE
H3K4me3
0
130
H3K27Ac Mark (Often Found Near Regulatory Elements) on 7 cell lines from ENCODE
H3K27ac
0


b
RQ AMMECR1
1.0
0.5
0.0
ctrl
M
P
RQ AMMECR1
1.5
1.0
0.5
0.0
ctrl
M
P
RQ RGAG1
1.0
0.5
0.0
ctrl
M
P
RQ AMMECR1L
2.0
1.5
1.0
0.5
0.0
ctrl
M
P

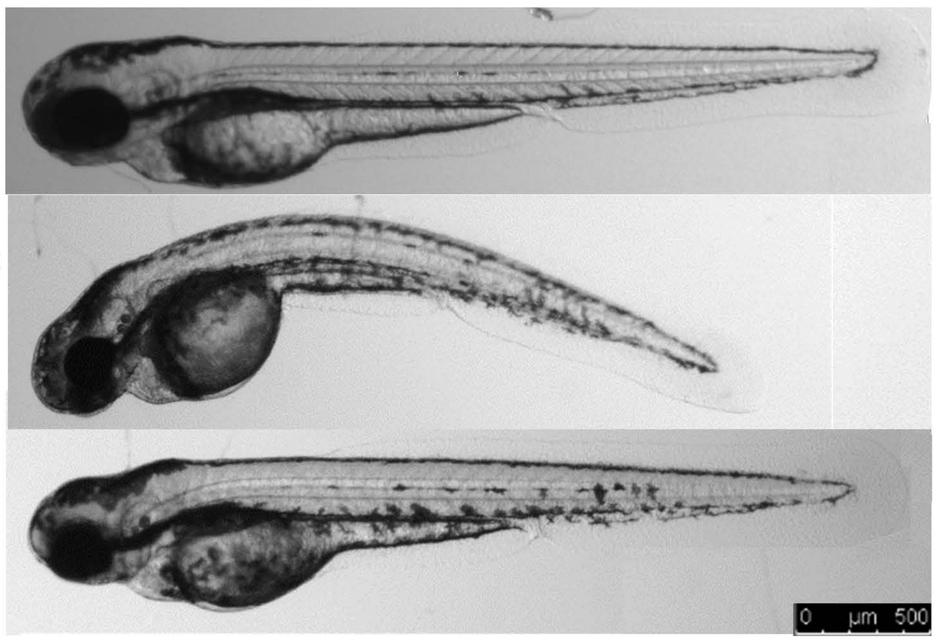
a


Std Ctrl MO
8.4 ng
MO-I2E3
2.1 ng
MO-E3I3
4.2 ng
0 µm 500

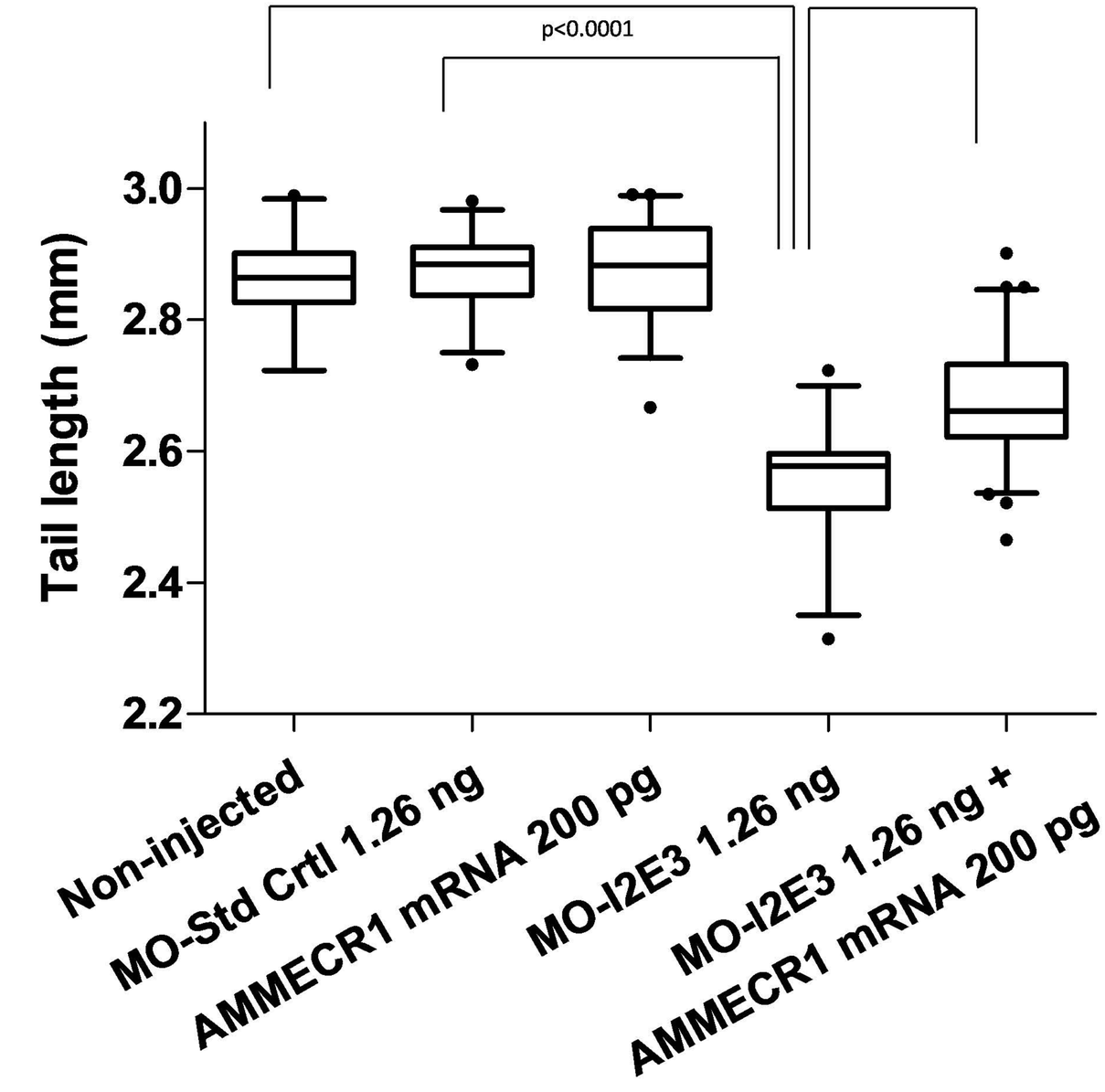
b


Heart beats/min
200
180
160
140
120
100
p<0.0001
p<0.0001
p<0.0001
p<0.0001
p<0.0001
Non-injected
MO-Std Ctrl 1.26 ng
MO-I2E3 1.26 ng
MO-Std Ctrl 4.2 ng
MO-E3I3 4.2 ng
MO-Std Ctrl 8.4 ng
MO-E1I1 8.4 ng
AMMECR1 mRNA 200 pg

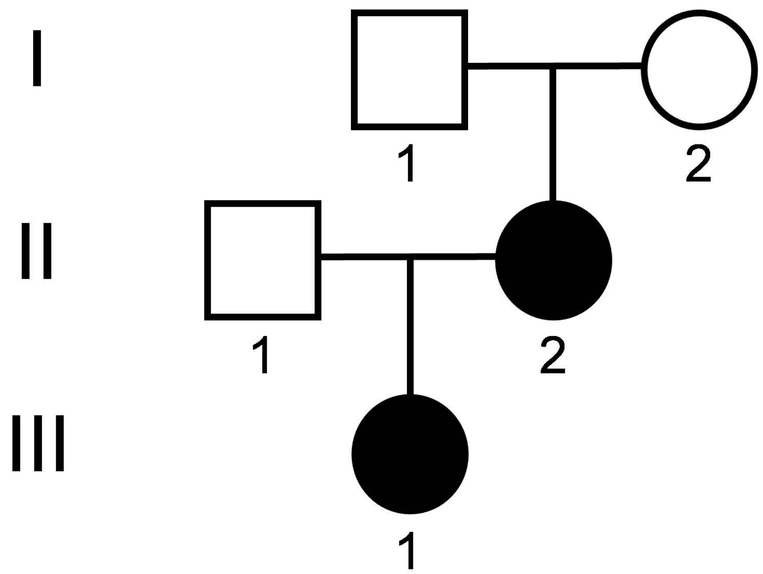
c


Std Ctrl MO
1.26 ng
MO-I2E3
1.26 ng
MO-I2E3
1.26 ng +
AMMECR1 mRNA
0 µm 500

d

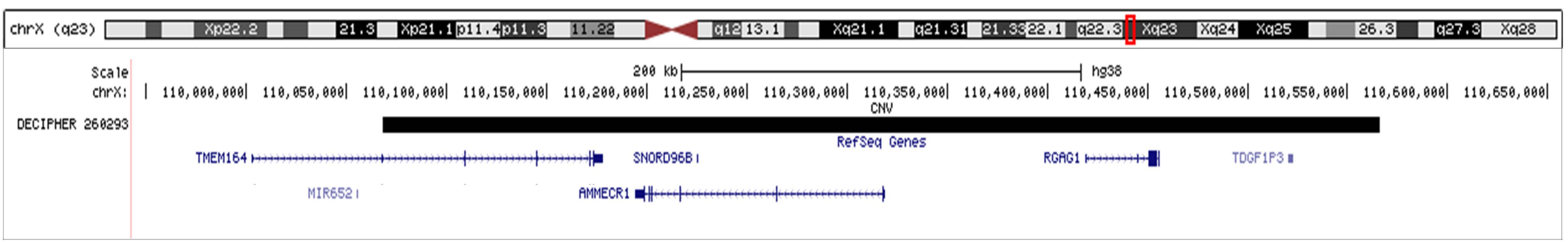

Tail length (mm)
3.0
2.8
2.6
2.4
2.2
p<0.0001
p<0.0001
Non-injected
MO-Std Crtl 1.26 ng
AMMECR1 mRNA 200 pg
MO-I2E3 1.26 ng
MO-I2E3 1.26 ng +
AMMECR1 mRNA 200 pg

a


I
1
2
II
1
2
III
1

b

c


chrX (q23)
Xp22.2
21.3
Xp21.1
p11.4
p11.3
11.22
q12
13.1
Xq21.1
q21.31
21.33
22.1
q22.3
Xq23
Xq24
Xq25
26.3
q27.3
Xq28
Scale
200 kb
hg38
chrX:
110,000,000
110,050,000
110,100,000
110,150,000
110,200,000
110,250,000
110,300,000
110,350,000
110,400,000
110,450,000
110,500,000
110,550,000
110,600,000
110,650,000
CNV
DECIPHER 260293
RefSeq Genes
TMEM164
SNORD96B
RGAG1
TDGF1P3
MIR652
AMMECR1