

RESEARCH

Open Access



# Zebrafish as a model to investigate a biallelic gain-of-function variant in *MSGN1*, associated with a novel skeletal dysplasia syndrome

Asuman Koparir<sup>1†</sup>, Caroline Lekszas<sup>1†</sup>, Kemal Keseroglu<sup>2</sup>, Thalia Rose<sup>1</sup>, Lena Rapp<sup>1</sup>, Aboufazel Rad<sup>3</sup>, Reza Maroofian<sup>4</sup>, Nakul Narendran<sup>5</sup>, Atefeh Hasanzadeh<sup>6</sup>, Ehsan Ghayoor Karimiani<sup>7</sup>, Felix Boschann<sup>8,9</sup>, Uwe Kornak<sup>8,10</sup>, Eva Klopocki<sup>1</sup>, Ertuğrul M. Özbudak<sup>2,11</sup>, Barbara Vona<sup>10,12</sup>, Thomas Haaf<sup>1</sup> and Daniel Liedtke<sup>1\*</sup>

## Abstract

**Background/Objectives** Rare genetic disorders causing specific congenital developmental abnormalities often manifest in single families. Investigation of disease-causing molecular features are most times lacking, although these investigations may open novel therapeutic options for patients. In this study, we aimed to identify the genetic cause in an Iranian patient with severe skeletal dysplasia and to model its molecular function in zebrafish embryos.

**Results** The proband displays short stature and multiple skeletal abnormalities, including mesomelic dysplasia of the arms with complete humero-radio-ulna synostosis, arched clavicles, pelvic dysplasia, short and thin fibulae, proportionally short vertebrae, hyperlordosis and mild kyphosis. Exome sequencing of the patient revealed a novel homozygous c.374G > T, p.(Arg125Leu) missense variant in *MSGN1* (NM\_001105569). *MSGN1*, a basic-Helix-Loop-Helix transcription factor, plays a crucial role in formation of presomitic mesoderm progenitor cells/mesodermal stem cells during early developmental processes in vertebrates. Initial in vitro experiments show protein stability and correct intracellular localization of the novel variant in the nucleus and imply retained transcription factor function. To test the pathogenicity of the detected variant, we overexpressed wild-type and mutant *msgn1* mRNA in zebrafish embryos and analyzed *tbxta* (*T/brachyury/ntl*). Overexpression of wild-type or mutant *msgn1* mRNA significantly reduces *tbxta* expression in the tailbud compared to control embryos. Mutant *msgn1* mRNA injected embryos depict a more severe effect, implying a gain-of-function mechanism. In vivo analysis on embryonic development was performed by clonal *msgn1* overexpression in zebrafish embryos further demonstrated altered cell compartments in the presomitic mesoderm, notochord and pectoral fin buds. Detection of ectopic *tbx6* and *bmp2* expression in these embryos hint to affected downstream signals due to *Msgn1* gain-of-function.

**Conclusion** In contrast to loss-of-function effects described in animal knockdown models, gain-of-function of *MSGN1* explains the only mildly affected axial skeleton of the proband and rather normal vertebrae. In this context we observed notochord bending and potentially disruption of pectoral fin buds/upper extremity after overexpression of *msgn1* in zebrafish embryos. The latter might result from *Msgn1* function on mesenchymal stem cells

<sup>†</sup>Asuman Koparir and Caroline Lekszas contributed equally.

\*Correspondence:

Daniel Liedtke

Daniel.liedtke@uni-wuerzburg.de

Full list of author information is available at the end of the article



or on chondrogenesis in these regions. In addition, we detected ectopic *tbx6* and *bmp2a* expression after gain of *Msgn1* function in zebrafish, which are interconnected to short stature, congenital scoliosis, limb shortening and prominent skeletal malformations in patients. Our findings highlight a rare, so far undescribed skeletal dysplasia syndrome associated with a gain-of-function mutation in *MSGN1* and hint to its molecular downstream effectors.

## Introduction

Skeletal dysplasias are a highly genetically heterogeneous group of skeletal and cartilaginous tissue disorders. The current classification comprises 461 distinct disorders and 437 underlying genes [25, 41]. Because of genetic and phenotypic heterogeneity of skeletal dysplasias, it is often difficult to get an unambiguous clinical diagnosis without molecular analyses. Next generation sequencing (NGS) technologies allow elucidation of the genetic basis of skeletal dysplasias, enabling precise diagnostics and adapted treatments. This method can identify new genetic variants in small patient cohorts or even a single affected individual. However, many of these variants are of uncertain significance/unknown function.

Mesogenin1 (*MSGN1*, OMIM: \*612209), a basic-Helix–Loop–Helix transcription factor, is expressed in the presomitic mesoderm (PSM) only at early stages of vertebrate embryogenesis and plays a crucial role in formation of mesodermal progenitor cells during somitogenesis [47]. *Msgn1* null mouse embryos have been shown to be defective in somitogenesis and exhibit somite segmentation defects with absence of all structures arising from the paraxial mesoderm (e.g. musculo-skeletal trunk, trunk skeleton, parts of the skin and trunk vasculature). In addition, these mouse embryos displayed prominent tail segmentation defects. Reduced expression of components of the Delta-Notch signaling pathway and the segmentation clock oscillator was observed within the PSM [47]. In addition to the loss-of-function phenotype, ectopic expression of *Msgn1* in mice suppresses notochord cell differentiation and promotes a PSM stem-cell fate [2]. The notochord is a primary structural element, a source of patterning signals, and is thereby essential for locomotion by enabling early cartilage and vertebrae development during vertebrate embryogenesis [6, 37]. In accordance with mouse experiments, zebrafish embryos injected with *msgn1* mRNA exhibited severe truncation of tail structure, e.g. loss of tailbud and notochord cells [46]. This altered development is induced by disturbed expression of essential PSM regulators like *tbxta* (alias: *TBXT* (human), *T/brachyury/bra* (mouse), *no tail/ntla/zft/ta* (zebrafish)), which were significantly reduced in the posterior tailbud and in the axial region after gain of *Msgn1* function [46]. The correct function of the PSM during early development is the prerequisite to subsequent processes like somite formation, notochord

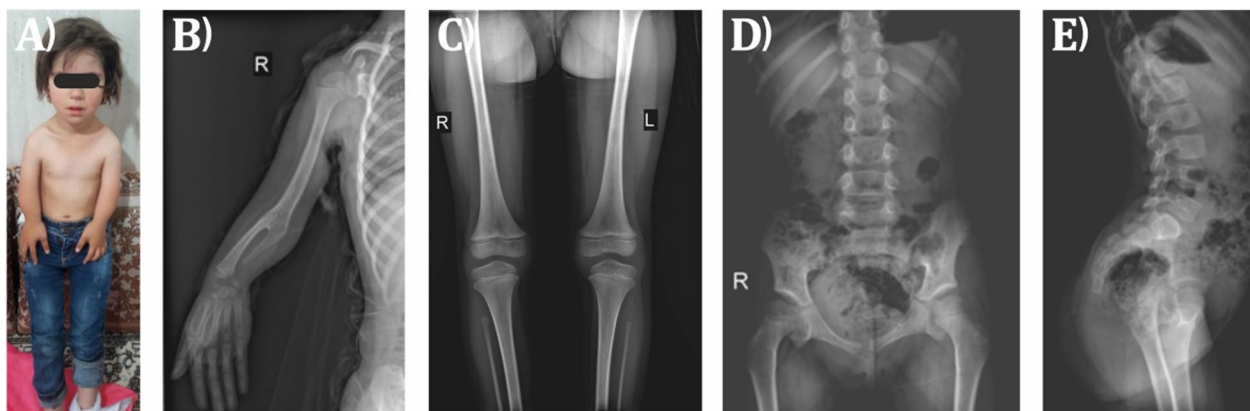
development and the segmentation clock itself. Its function involves a highly sophisticated network of different molecular factors in addition to *MSGN1*. These comprise WNT, FGF, retinoic acid, Delta-Notch signaling pathways and T-box transcription factors [23, 35] and are evolutionarily conserved among vertebrate species. *MSGN1* functions as a transcription factor, which is regulated by these factors, but moreover orchestrates correct transcription of several essential effectors within this network [2].

Here, we have investigated the molecular pathogenesis of a novel form of skeletal dysplasia in the offspring of consanguineous Iranian parents and identified a novel homozygous c.374G>T, p.(Arg125Leu) variant in *MSGN1* (NM\_001105569) in the affected child. To test the pathogenicity of this variant, we overexpressed either wild-type or mutant *msgn1* in zebrafish via mRNA injection and analyzed *tbxta* expression in the tailbud. Consequences of enhanced *Msgn1* activity on early development were further analyzed in vivo in zebrafish embryos. Mosaic transient-transgenic overexpression of *msgn1* showed spatial and temporal restricted consequences on tail development by influencing Delta-Notch active cells in the posterior progenitor zone and by inducing ectopic gene expression.

## Results

### Clinical description

The patient was the first child born to a gravida 2, para 1 mother. The healthy parents were second-degree cousins. On physical examination, the female proband had mild facial dysmorphism including hypertelorism, epicanthus, midface hypoplasia, mild retrognathia, and short stature with noticeable mesomelic shortening of the arms. She also presented dysplastic neck-thorax region with a prominent trapezoid muscle and arched clavicles (Fig. 1A). Radiographic images revealed humero-radial and radio-ulnar synostosis. The radii were severely shortened and hypoplastic, the ulnae were rudimentary with only their proximal part present. The scaphoid bone was only present as a tiny nucleus on the left side, while it was ossified on the right side (Fig. 1B). The diaphysis of the femora was very thin. The fibulae were clearly shortened and too thin (Fig. 1C). The iliac bones were somewhat narrow, the ramus inferior of the pubic bone was missing,



**Fig. 1** **A** Short stature and mesomelic shortening of arms in the patient. **B–E** Skeletal X-ray surveys of the patient's right upper extremity (**B**), lower extremities (**C**), anterior view of the lower spine and pelvis (**D**), and lateral view of the lower spine and pelvis (**E**)

and the ischium was only rudimentary (Fig. 1D). In addition, there was a hyperlordosis and mild scoliosis of the spine (Fig. 1E).

#### Whole exome sequencing revealed a likely disease-causing variant in *MSGN1*

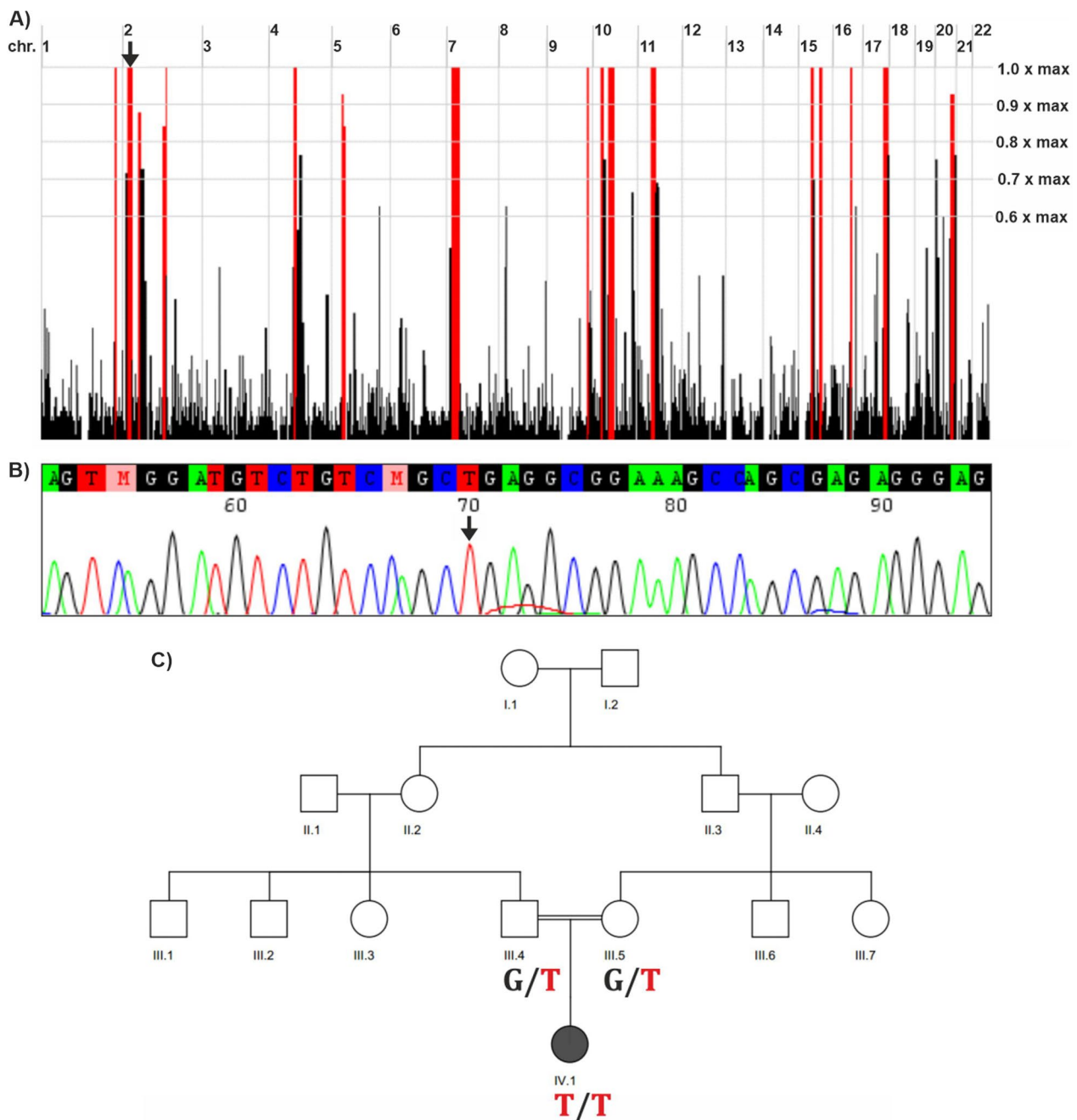
Whole exome sequencing (WES) analysis detected a rare homozygous missense variant in *MSGN1* (NM\_001105569.3: c.374G > T, p.(Arg125Leu); NC\_000002.11:g.17998159G > T). This variant is listed 11 times in heterozygous but not in homozygous state in the population database gnomAD v.4.0.0. The variant affects a conserved amino acid residue within a highly conserved bHLH domain in *MSGN1* (Additional file 2: Fig. S1 and Additional file 3: Fig. S2). Different *in-silico* prediction tools such as CADD (Combined Annotation Dependent Depletion; [15, 30]), REVEL [11] and AlphaMissense [5] classified this variant as pathogenic (Additional file 8: Table S1). Moreover, the variant was predicted by SpliceAI to have no splicing effect [12]. Furthermore, we visualized the theoretical three-dimensional (3D) representation of both wild-type and mutant forms of the *MSGN1* protein, to see any alteration in the structure and conformation of the protein (Additional file 2: Fig. S1A, B). The affected Arginine 125 within the *MSGN1* protein is located at the second position of the evolutionary highly conserved basic helix-loop-helix (bHLH) protein domain at the beginning of the larger, first alpha-helix. The exchange from Arginine (negative hydrophathy index, amphipathic and polar, positive charged side chain) to Leucine (positive hydrophathy index, aliphatic and nonpolar, no charge) is biochemically profound. However, the exchange is predicted to not abrogate the helix structure but to

cause misfolding of the encoded protein at the transition between the alpha-helix and the N-terminus (Additional file 2: Fig. S1A, B). According to HomozygosityMapper [34], the detected *MSGN1* variant lies within an approximately 13.2 Mb homozygous interval on chromosome 2 (Fig. 2A). Sanger sequencing analysis confirmed WES findings and demonstrated that both parents are heterozygous carriers for the variant (Fig. 2B, C). An additional WGS analysis did not detect other candidate variants with higher pathogenic potential in known skeletal dysplasia genes and surrounding non-coding sequences.

#### In vitro analysis of intracellular *MSGN1* localization

At first, we tested whether the newly identified *MSGN1* variant and its predicted misfolding results in retention of *MSGN1* protein in the cellular cytoplasm and nuclear exclusion, consistent with a loss-of-function of the transcription factor. HEK 293T cells were transfected with CMV-promoter driven expression plasmids at different concentrations incorporating either wild-type or c.374G > T, p.(Arg125Leu) *MSGN1*-FLAG-tag variants. Subsequent protein visualization was performed after 48 h via immunofluorescence by utilizing the protein tagged (Myc-DDK/FLAG). Transfected HEK 293T cells showed expression of both *MSGN1* variants at different plasmid concentrations and did not imply loss of *MSGN1* protein or restrained protein localization (Fig. 3A; controls Additional file 4: Fig. S3A).

To further quantify intracellular localization of the *MSGN1* proteins in the cytoplasm or in the nucleus of transfected cells we subsequently quantified

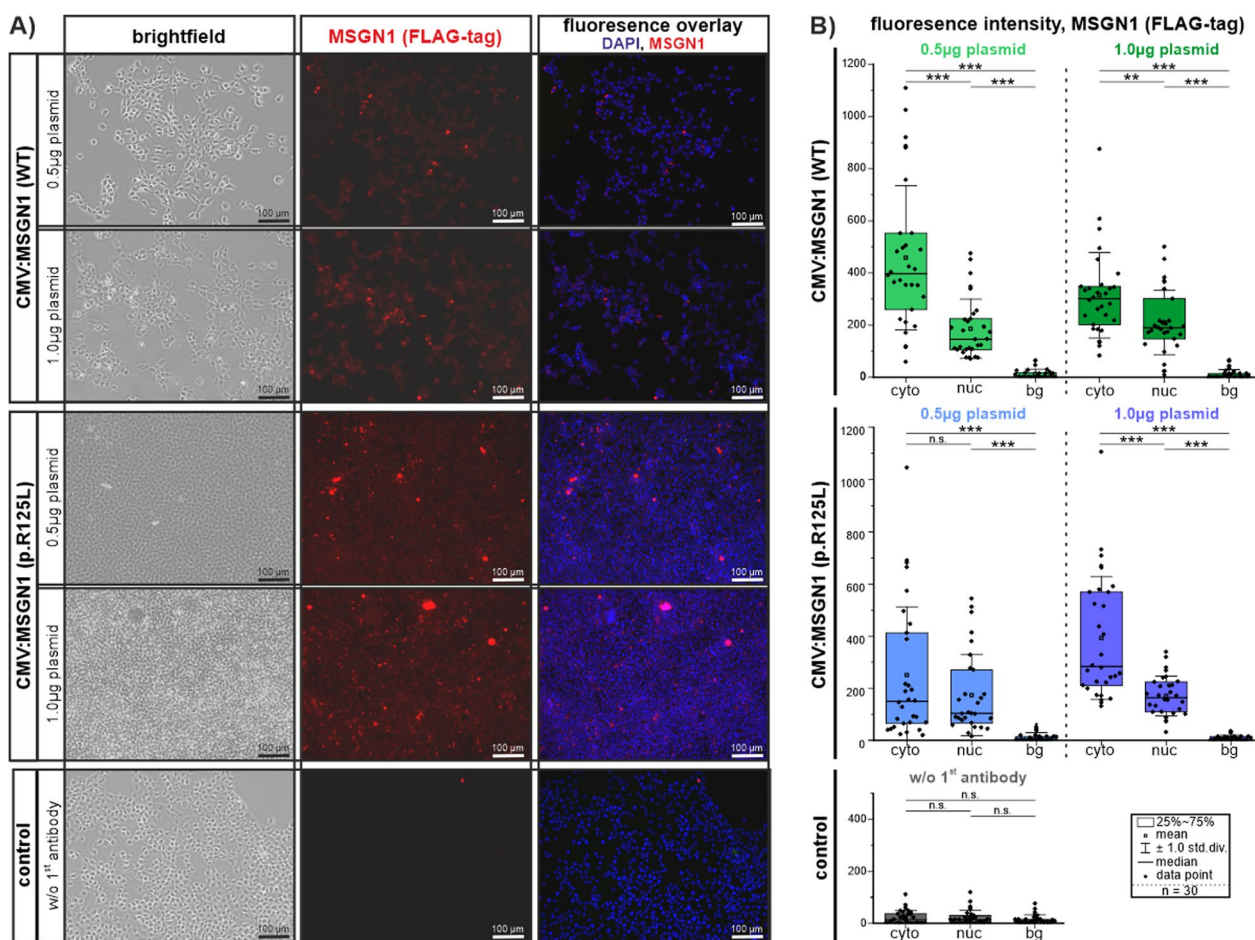


**Fig. 2** Genetic results of the family. **A** Homozygous intervals in the patient's genome, as determined by HomozygosityMapper. *MSGN1* resides in a ~13.2 Mb homozygous interval on chromosome 2 (black arrow). **B** Sanger sequencing electropherogram of the patient's variant in *MSGN1*. The mutated base is indicated by a black arrow. **C** Segregation of the *MSGN1* variant within the family

fluorescence signal intensities of *MSGN1*-FLAG-tag and DAPI by confocal laser-scanning microscopy (LSM; Fig. 3B and Additional file 4: Fig. S3B). Measurement of *MSGN1* fluorescence signal intensity did not reveal any difference in localization between *MSGN1* (WT)-FLAG-tag or *MSGN1* (Arg127Leu)-FLAG-tag variants.

#### Gain of *msgn1* activity results in developmental defects during early stages of zebrafish development

Next we investigated the possible pathogenic effects of the variant on Mesogenin1 activity in zebrafish a *in vivo* model. To this end, we overexpressed either wild-type or the variant version (p.(Arg71Leu) in zebrafish corresponding to p.(Arg125Leu) in humans) of *Msgn1*

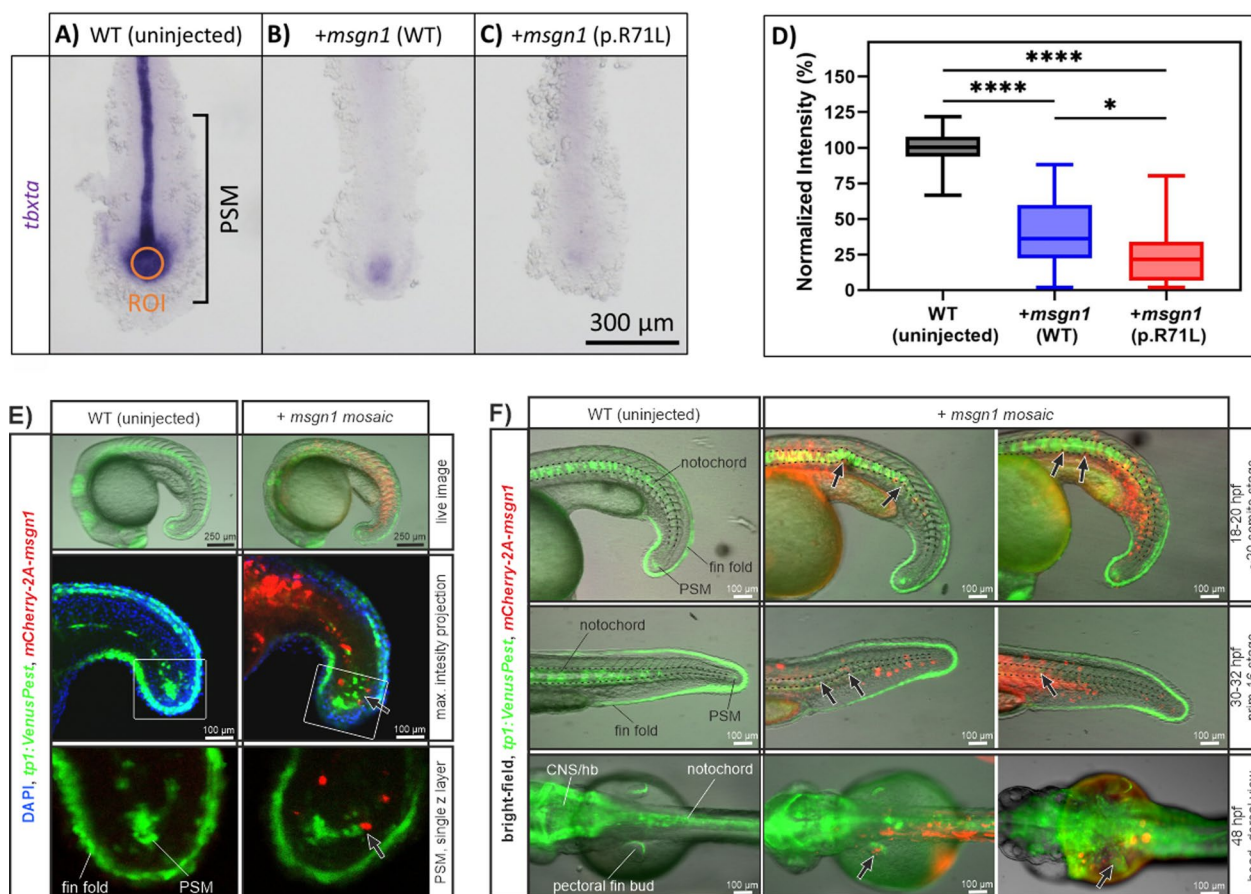


**Fig. 3** In vitro transfection of HEK 293T cells with CMV:MSGN1 (WT)-FLAG-tag and CMV:MSGN1 p.(Arg125Leu)-FLAG-tag plasmids at different concentrations. **A** Immunofluorescence showing expression of tagged MSGN1 proteins (WT) or p.(Arg125Leu) after 48h of transfection. Corresponding experimental controls are shown in Additional file 4: Fig. S3A. **B** Quantification of MSGN1 localization in transfected cells. Fluorescence signals in a by single z-plane were visualized by confocal laser-scanning microscopy and subsequently signal intensity was measured in single cells (cyto: cytoplasm; nuc: nucleus) and outside of cells (bg: background). Graphs show signal intensity measurements of the RFP channel (MSGN1 Flag-Tag) of 30 cells per experimental group and 30 background positions. Signal intensity measurements of the corresponding DAPI channel for nucleus identification are given in Additional file 4: Fig S3B. Values are given in Additional file 1: Excel file S1

via mRNA injection at one-cell stage. We examined *tbxta* expression, which is regulated by *Msgn1* [4, 46], in the tailbud at 8-somite stage (Fig. 4A–C). We found that overexpression of both wild-type and *msgn1* p.(Arg71Leu) mRNA significantly reduced *tbxta* expression in the tailbud compared to uninjected wild-type embryos ( $P < 0.0001$ ). In addition, we also observed a more severe effect by *msgn1* p.(Arg71Leu), compared to *msgn1* wild-type mRNA ( $P = 0.0186$ ), suggesting that the variant leads to gain-of-function in Mesogenin1 activity (Fig. 4D).

To investigate additional consequences of enhanced *Msgn1* activity on embryonic development we utilized a mosaic transient-transgenic overexpression method. In short, microinjection of a DNA plasmid containing

a *mCherry-2A-msgn1* coding sequence (driven by the zebrafish *msgn1* promoter) into single cells of early blastula stage zebrafish embryos drives clonal *msgn1* expression during development. Ectopic *msgn1* expressing cells are marked by mCherry fluorescence. By application of this method in the *tp1:VenusPEST* transgenic zebrafish background, we were able to investigate *msgn1* overexpressing cells and Notch signaling during development simultaneously by in vivo fluorescence visualization. We observed that mCherry positive/*msgn1* expressing cells disturb Notch positive cells within the PSM (18–20 hpf; Fig. 4E). The cell population resembles *tbxta* expressing posterior PSM cells at the maturation zone, giving rise to notochord and somite tissues [8, 17]. Other Notch positive tissues, e.g. dorsal fin fold cells, were not influenced



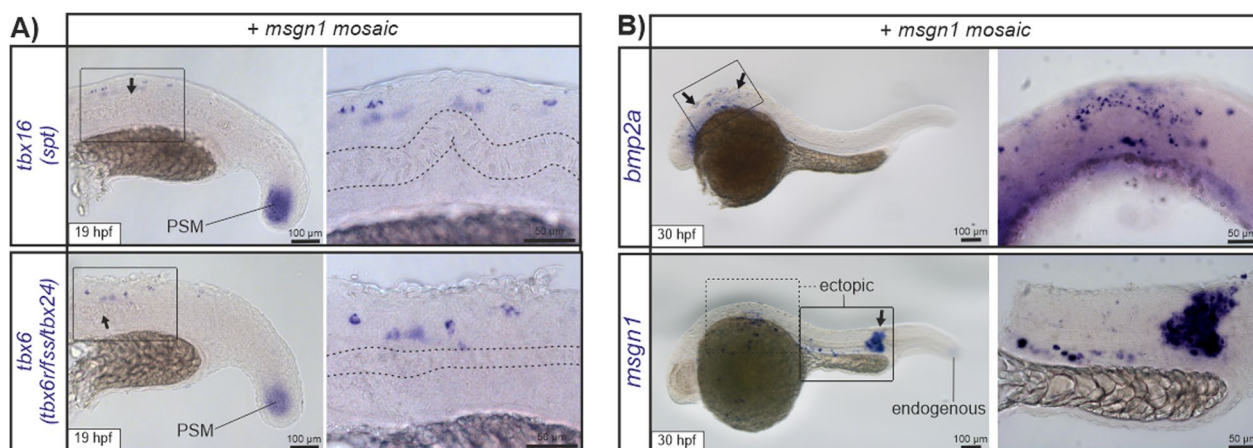
**Fig. 4** In situ hybridization images of *tbxta* (*no tail*, *Brachyury*) expression in uninjected wild-type controls (**A**), wild-type *msgn1* RNA (**B**) and mutant *msgn1* RNA p.(Arg71Leu) in zebrafish ( $\approx$  p.(Arg125Leu) in humans) injected embryos (**C**). **D** Quantification of *tbxta* ISH signal in posterior PSM cells (orange ROI mark) is shown by normalized intensity comparison between the three groups of three independent experiments (WT (uninjected): 49 embryos; +*msgn1* (WT) RNA: 58 embryos; +*msgn1* p.(Arg71Leu) RNA: 54 embryos). Mosaic transient-transgenic overexpression of wild-type *msgn1* results in partial disruption of Notch signals in PSM cells (**E**) and influences normal notochord (marked with dashed lines) and pectoral fin bud development (marked with arrow). **F** In vivo images show representative embryos of three independent injections experiments (overall: 37 imaged embryos; WT: 5 embryos; +*msgn1* mosaic: 32 embryos). ISH signal quantification values and injection statistics are given in Additional file 1: Excel file S1

by *msgn1* implying PSM specific effects. PSM disruption is associated with developmental defects whose severity is linked to the amount of *msgn1* overexpressing cells in the PSM region (Additional file 5: Fig S4). Investigation of these embryos at later developmental stages further showed that mosaic embryos displayed altered notochord and pectoral fin structures (e.g. wavy notochord outgrowth, disruption of Notch signals in notochord cells and in pectoral fin buds, marked by arrows in Fig. 4F).

To further decipher potential downstream effects of ectopic *Msgn1* activity, we investigated expression of previously described *MSGN1* downstream target genes in mosaic zebrafish embryos [2]. Visualization of *tbx6* and *tbx16* via in situ hybridization showed that both factors are endogenously expressed within the PSM and

are additionally expressed within the trunk at regions depicting notochord alterations (19 hpf; Fig. 5A). Moreover, strong ectopic expression of *bmp2a* and *msgn1* was detected within the head and trunk regions of 30 hpf embryos (short staining time with low levels of endogenous ISH transcript signal visible; Fig. 5B). Detection of *msgn1* expression via in situ hybridization was in accordance with *mCherry-2a-msgn1* fluorescence expression patterns (see Figs. 3F and 5E) and probably detects primarily transgene induced *msgn1* expression in these mosaic embryos.

General overexpression either of human *MSGN1* (WT) or *MSGN1* p.(Arg125Leu) variants in zebrafish embryos resulted in similar developmental consequences during early developmental stages as gain of zebrafish *msgn1*



**Fig. 5** **A** *In situ* hybridization images of *tbx6* (*tbx6/rfss/fused somites/tbx24*) and *tbx16* (*spt/spatetail*) expression in injected embryos 19 hpf showing mosaic transient-transgenic overexpression of wild-type *msgn1*. Expression of *tbx6* and *tbx16* was detected in PSM cells (endogenous expression) and within the trunk, colocalizing with notochord alterations (shown in higher magnification). **B** In 30 hpf old *msgn1* mosaic embryos, ectopic expression of *bmp2a* was detected in the head region, consistent with strong ectopic *msgn1* expression in trunk and head regions of injected embryos at this stage

in mosaic embryos, e.g. notochord bending and disruption of PSM cells (used plasmids: CMV:MSGN1-Flag-tag (WT) and CMV:MSGN1-Flag-tag p.(Arg125Leu), Additional file 6: Fig. S5 and Additional file 1: Excel file S1). This observation implies retained function of human MSGN1 and of the patient derived MSGN1 variant also in zebrafish. Further tissue-restricted mosaic expression of zebrafish *msgn1* p.(Arg71Leu) and human *MSGN1* p.(Arg125Leu) variants in PSM and trunk cells under *msgn1* promoter control in *tp1:VenusPEST* embryos resulted in similar developmental consequences like *msgn1* wild-type injected embryos, but display more severe manifestations and higher rates of severe phenotypes in injected embryos (*msgn1*:mCherry-2A-*msgn1* vector: 15% (n=197) and 20% (n=92); *msgn1*:mCherry-2A-*msgn1* p.(Arg71Leu) vector: 32% (n=50); *msgn1*:mCherry-2A-MSGN1 p.(Arg125Leu) vector: 27% (n=104); Additional file 7: Fig. S6 and Additional file 1: Excel file S1).

## Discussion

In our study, we identified a homozygous missense variant c.374G>T, p.(Arg125Leu) in *MSGN1* via exome sequencing to be potentially associated with prominent skeletal abnormalities in a single patient. *MSGN1* missense variants have not yet been associated with a human Mendelian disease. Previously, eleven *MSGN1* missense variants have been reported in ClinVar, but all lack detailed zygosity and phenotype information or segregation analyses and are all classified as variants of uncertain significance (Tab. S3). Several structural variants (gains

and losses) including the *MSGN1* locus are additionally described in ClinVar, but all these alterations affect larger genomic regions incorporating several coding genes and therefore do not allow the conclusion that *MSGN1* is disease causing (Tab. S4). The aim of this study was to provide molecular insights into our patient's phenotype and to link the observed malformations to known *MSGN1* functions during development using zebrafish.

In contrast to loss of function phenotypes in *MSGN1* animal models (short tail, prenatal lethal, absent notochord), we did not observe prominent vertebral segmentation or axial skeleton defects in our patient. We therefore propose the detected *MSGN1* variant has a gain of function effect, which could explain only mildly affected vertebrae of the proband. Initial *in vitro* investigations clarified, that the corresponding protein of the novel *MSGN1* p.(Arg125Leu) variant is still able to localize into the nucleus and therefore may act as a transcription factor. Previously published results [46] in zebrafish showed that elevated *msgn1* activity reduces *tbxta* expression in the PSM. Our own experimental data are in accordance with this observation and further demonstrate that the corresponding *msgn1* p.(Arg71Leu) variant results in a stronger gain of function effect hinting to raised Mesogenin1 activity (Fig. 4D and Additional file 7: Fig. S6). TBXT, the prototypical T-box transcription factor, is essential for the formation and differentiation of posterior mesoderm and for axial development in all vertebrates. Homozygous variants in *TBXT* orthologues cause embryonic lethality, while heterozygous variants are associated with short tail mice [16, 33] and zebrafish (*no tail* mutant; [27, 38, 42]). In humans, a homozygous

*TBXT* variant (c.512A>G, p.(His171Arg), NM\_003181), has been associated with sacral agenesis, abnormal ossification of vertebral bodies and persistent notochord canal [29]. In chicken *T* is expressed in the lateral plate mesoderm at the onset of limb bud formation [22]. *T* controls *TBX6* expression, which is known to regulate activation of *MSGN1* [45]. *TBX6* variants cause spondylocostal dysostosis in human [36]. However, ectopic *Tbx6* gain-of-function within the segmented paraxial mesoderm and its derivatives in mice are associated with appendicular skeletal defects including hypoplastic scapula, small limbs, shortened and malformed humerus, radius, ulna, femur, tibia and fibula, resembling *Tbx15* null mouse embryos and Cousin syndrome in humans (OMIM #260660) [18, 43]. Since gain of *Tbx6* function in mice results in short stature and severe limb defects [45], we hypothesize that *MSGN1* gain of function is responsible for short stature and limb abnormalities in our patient. The expression of T-box transcription factors like *tbxt*, *tbx6*, *tbx16*, and *tbx16l* in zebrafish is regulated by *Msgn1* [4, 13, 24]. Our transient mosaic experiments in zebrafish and previously published zebrafish experiments further show ectopic *tbx6* expression via gain of *Msgn1* function (Fig. 4 and [46]). To our knowledge, no skeletal dysplasia has been associated with biallelic pathogenic gain of function variants. However, there are some examples in the literature for biallelic pathogenic gain of function variants, which affect other physiological systems e.g. the immune system (*NLRP1*, *STAT2*) [7, 9].

*MSGN1* plays an important role in notochord maintenance and development, that expresses genes such as *SOX9* which are characteristic of chondrogenesis [37]. Prominent functions of Delta-Notch signals during early posterior tail development have been intensively investigated and are necessary to synchronize PSM cells within the somitogenesis clock of vertebrates [19–21]. We visualized Delta-Notch signalling for in vivo observation of developmental defects in the posterior PSM maturation zone and in the notochord. Our zebrafish results imply that gain of *msgn1* function in single cells of the posterior progenitor zone (maturation zone in the posterior PSM) disturbs the structure of Delta-Notch positive cells which give rise to somite, paraxial mesoderm and notochord tissues during later developmental stage [8, 31]. We subsequently observed notochord bending and potential disruption of pectoral fin buds/upper extremity. The later might result from *msgn1* ectopic function on mesenchymal stem cells or on chondrogenesis in this region [3, 28, 40]. Interestingly within this context, we found that transcriptional *Msgn1* activity results in ectopic *bmp2a* expression (besides action on *tbx6* and *tbxta*). *Bmp2* is described as a direct *MSGN1* target [2] and is known

to be associated with prominent skeletal anomalies in patients, like brachydactyly type A2 (OMIM #112600), as well as short stature, facial dysmorphism, and skeletal anomalies with or without cardiac anomalies (OMIM #112261). Further studies will have to clarify, if the long-term ectopic expression of downstream effectors of *MSGN1* is also the driver of long-term developmental alterations in the current patient, since spatio-temporal *MSGN1* expression is usually tightly regulated during development.

Collectively, the studied patient presents a novel *MSGN1*-associated skeletal syndrome in humans, highlighting the importance of the temporally and spatially appropriate *MSGN1* activity in various developmental pathways. However, additional patients/families are needed to unequivocally link *MSGN1* gain-of-function to the described syndrome.

## Materials and methods

### Whole exome sequencing and database research

Exome capture was performed according to the Illumina Nextera Rapid Capture Enrichment library preparation protocol (individuals IV.1, III.4 and III.5) using 50 ng of genomic DNA. Paired-end sequencing of the libraries was performed with a NextSeq500 sequencer and the v2 reagent kit (Illumina, San Diego, California, USA). Sequences were mapped to the human genome reference (NCBI build 37/hg19 version) using the Burrows-Wheeler Aligner. Aligned reads ranged between 85,736,827 and 99,719,268. The mean coverage was  $\geq 50$  with 91.2%. 99% of the exome were covered at least 10x. A total of 678,344–961,783 variants per sample were called and analyzed using GensearchNGS software (PhenoSystems SA, Braine le Chateau, Belgium). Variants with a coverage of  $\leq 10$ , a Phred-scaled quality of  $\leq 15$ , a frequency of  $\leq 15$ , and a MAF of  $\geq 2\%$  were neglected. Six control samples from healthy individuals were used for filtering out platform artefacts. Alamut Visual (Interactive Biosoftware, Rouen, France) software including prediction tools like SIFT, MutationTaster, PolyPhen-2, CADD-, and REVEL-Score was used for variant prioritization. Potential effects of a variant on pre-mRNA splicing were evaluated by SpliceAI, SpliceSiteFinder-like, MaxEntScan, NNSPLICE, GeneSplicer, Human Splicing Finder, ESEfinder, RESCUE-ESE, and EX-SKIP. Population databases like gnomAD v4.0.0, and GME revealed whether a variant has been previously found. Protein expression, structure, and functional aspects were investigated with UniProt and The Human Protein Atlas. Information on mouse and zebrafish models was retrieved from the MGI and ZFIN database, respectively.



### Sanger sequencing

*MSGN1* exon one was amplified by a touchdown PCR program using primers in the flanking introns (forward: 5'-GGTGGACTACAATATGTTAGCTTTCC-3' and reverse: 5'-TAGACAGGTGGCAGGTAATTCC-3'). A clean-up step with ExoSAP-IT (Applied Biosystems, Foster City, USA) was followed by the sequencing reaction using the BigDye Terminator Cycle Sequencing Kit v1.1 (Applied Biosystems, Waltham, USA). Sequencing was conducted on a 3130XL capillary sequencer (Applied Biosystems, Waltham, USA) and data analysis was performed with Gensearch (PhenoSystems SA, Braine le Chateau, Belgium).

### Whole genome sequencing

Short read genome sequencing was performed to exclude structural variants and non-coding variants in known skeletal dysplasia genes. In brief, genomic DNA was isolated from peripheral blood and sequenced at 30× coverage using the Illumina TruSeq PCR-free protocol at the West German Genome Center (WGCC). Reads were aligned to the human genome build GRCh37/hg19 using BWA-MEM 0.7.17. The VarFish software was used for filtering and interpretation of variants including SVs (caller: Delly v.0.8.1) according to an in-house SOP [10].

### In vitro experiments and intracellular *MSGN1* localization measurements

Transfections were performed with  $2.5 \times 10^5$  HEK 293T cells per well in 12-well plates. HEK 293T cell were initially grown on glass cover slides and subsequently transfected with two different amounts of DNA plasmids containing either the coding sequence of wild-type *MSGN1* (WT) or of the *MSGN1* p.(Arg125Leu) variant coupled to a FLAG-tag under control of the *CMV* promoter. 0.5 µg and 1.0 µg plasmid DNA per 100 µl total transfection volume was used with 2 to 3 µl FuGene HD transfection agent per well (Promega, Madison, USA, product-nr. E2311). 48 h after transfection, cells were paraformaldehyde fixed and immunofluorescence staining of Flag-tag was performed by standard protocols for cell HEK 293T cell cultures (primary antibody: monoclonal DYKDDDDK Tag Recombinant, Thermo Fisher Scientific, Waltham, USA, product-nr. 701629, RRID: AB\_2532497, 1:250 dilution (2.5 µg/ml); secondary antibody: goat anti-rabbit-Alexa 594, Invitrogen/Thermo Fisher Scientific, product-nr. A-11012, RRID: AB\_2534079, 1:1000 dilution (1 µg/ml)) and DAPI (Sigma-Aldrich, St. Louis, USA; product-nr. D9542, 1:5000 dilution (4 µg/ml)). Corresponding controls were performed by transfection of an *CMV:GFP-Tag* plasmid (positive control), by including non-transfected control cells, by omitting primary or secondary antibody incubations (negative

controls). Images were taken with a Keyence BZ-X800 fluorescence microscope (Keyence, Osaka, Japan).

Intracellular localization of *MSGN1* proteins was quantified by analysis of single z-level immunofluorescence images acquired by laser scanning confocal microscopy (Nikon A1+, Nikon NIS-Elements software, Nikon Instruments, Tokyo, Japan) via ImageJ/Fiji software (<https://fiji.sc/>) [32]. In contrast to analysis by classical fluorescence microscope techniques, usage of LSM enables high resolution optical sectioning of fluorescence signals in a narrow single z-level at height of the nucleus and removes out-of-focus fluorescence signals outside of the focal plane in conjunction with strict pinhole settings. Images were taken at 600-fold magnification, in two channels (*MSGN1*-tag: 561 nm laser excitation, 610–620 nm emission; DAPI in nucleus: 405 nm laser excitation, 470–480 nm emission), at a resolution of 1024×1024 pixel corresponding to a 205×205 µm area. To circumvent light scattering, smallest pinhole settings per laser channel were chosen. For each transfection group, three non-overlapping images were acquired, subsequent 10 cells per image were randomly selected and fluorescence intensity values were measured in ImageJ/Fiji software by straight line histograms (line length: 10 µm (50 pixel), line thickness: 2 µm (10 pixels), mean value per pixel position in length were measured at 50 positions per line and channel). Line positions were adjusted to start in the cytoplasm (low DAPI intensity signal) and end in the nucleus (high DAPI intensity signal), with line midpoint corresponding to nucleus border (position #25). Cytoplasmic values were defined as position #15, 2.5 µm away from nucleolus border. Nuclear values were defined as position #35, 2.5 µm away from nucleolus border. Background values were defined as position #15 and #35 values measured outside of cell bodies at five positions per image.

### Zebrafish animal maintenance and lines

Laboratory zebrafish embryos (*Danio rerio*) of the *AB/TU* and *AB/AB* wild-type strain (ZDB-GENO-010924-10; ZDB-GENO-960809-7) and transgenic *Tg(EPV.Tp1-Mmu.Hbb:Venus-Mmu.Odc1)* (ZDB-TGCON-STRCT-120419-4, *tp1:VenusPEST*, [26]) were maintained as previously described under standard aquatic conditions at an average of 24 °C water temperature [1, 44]. Embryos were staged by morphological characteristics according to Kimmel et al. [14]. “hpf” and “dpf” indicate embryonic development in hours/days-post fertilization at 28.5 °C incubation temperature, respectively. All procedures involving experimental animals were performed in compliance with local animal welfare laws, guidelines, and policies. All presented experiments have been performed in zebrafish embryos and larvae younger than 5

dpf, before free-swimming and independent feeding, and thus are not regulated as animal experiments in Germany under current legislation.

#### Mosaic transient-transgenic *msgn1* overexpression and in vivo visualization

Mosaic transient-transgenic overexpression was performed by microinjection of a previously published plasmid containing a zebrafish *msgn1* promoter fragment driving 2A-coupled expression of *mCherry* fluorescence protein and wild-type zebrafish *msgn1* [46]. Injection of this plasmid results in prominent and severe tail malformations due to gain of *Msgn1* function in mesodermal progenitor cells. Injection of *msgn1* injection solution (DNA plasmid was diluted in water to a final concentration of 25 ng/ $\mu$ l and admixed with 0.05% Phenol-red (pH 7.0) for visualization of injection solution) into only one cell of early blastula zebrafish embryos (4-cell stage/1.0 hpf up to 16-cell stage/1.5 hpf) resulted in a fraction of *msgn1* overexpressing cells within developing embryos. In addition to zebrafish wild-type *msgn1* (*msgn1*:mCherry-2A-*msgn1*), plasmids containing either a zebrafish *msgn1* p.(Arg72Leu) variant (*msgn1*:mCherry-2A-*msgn1* (p.R71L)), or a human *MSGN1* p.(Arg125Leu) variant (*msgn1*:mCherry-2A-*MSGN1* (p.R125L)) were injected in a similar way into zebrafish embryos and investigated.

Injected embryos were in vivo analyzed under a Fluorescence Stereomicroscope (Leica S8 APO equipped with Leica GFP and DSR filter sets (filter nr. 10447408 and 10447412), Leica Microsystems, Wetzlar, Germany) at different time points during embryonic and early larval development (approximately between 16 and 48 hpf) for cells showing *msgn1* overexpression by mCherry fluorescence and for resulting developmental consequences. Injection was performed in embryos of a *tp1*:*VenusPEST* transgenic line [26], which enables in vivo visualization of Notch signaling due to expression of a short-half-life version of the fluorescence protein Venus under the *tp1* promoter element. The *tp1* element contains 12 EBV terminal protein 1 (TP1) gene promoter fragments for endogenous Notch (NICCD) and RBPJ/CB1/Su(H) co-factor binding. Detailed microscopic investigation was performed with Zeiss Imager A1 (in situ hybridizations, Carl Zeiss AG, Jena, Germany) or a Nikon A1+Laser scanning confocal microscope (in vivo, Nikon Corporation, Tokyo, Japan). For detailed microscopic investigations embryos were short time fixed in 4% paraformaldehyde/PBS for 30 min and were mounted in Mowiol. Nuclei were stained by DAPI incubation (1  $\mu$ g/mL in PBST; 15 min) before mounting. Images were analyzed with ImageJ/Fiji (<https://fiji.sc/>) and arranged with CorelDraw Graphics Suite (Alludo, Canada) software.

#### Plasmid vector cloning and mutagenesis

In vitro HEK 293T transfection experiments were performed with CMV promoter driven human tagged *MSGN1* (Origene, Rockville, USA, product-nr. RC225212, CDS: NM\_001105569). Patient *MSGN1* variant was introduced into the plasmid by site-directed mutagenesis (Q5 Site-Directed Mutagenesis Kit, New England Biolabs/NEB, Ipswich, USA, product-nr. E0554S) and was validated via Sanger sequencing.

Zebrafish injection experiments were performed with *msgn1* promoter driven zebrafish *msgn1* plasmid (sk-tol2-*msgn1*:mCherry-2A-*msgn1*) [46]. Human *MSGN1* c.374G>T, p.(Arg125Leu) and zebrafish *msgn1* c.211AG>CT p.(Arg71Leu) missense variants were introduced by site-directed mutagenesis in subcloned coding sequences without start codons (Q5 Site-Directed Mutagenesis and PCR cloning Kit, New England Biolabs/NEB, product-nr. E0554S and E1202S) and were validated via Sanger sequencing. Subsequently, the newly established variant coding sequences were used to replace zebrafish *msgn1* (WT) in sk-tol2-*msgn1*:mCherry-2A-*msgn1* by restriction site cloning. Primer sequences and used plasmids are listed in Additional file 8: Table S2.

#### Generation of mutant *msgn1*<sup>Arg71Leu</sup> mRNA and overexpression in zebrafish

A cDNA library was extracted from embryos and wild-type *msgn1* was amplified with PCR using primers listed in Additional file 8: Table S2. Mutant *msgn1* p.(Arg71Leu) was created using overlap extension primers. The wild-type *msgn1* and mutant *msgn1* p.(Arg71Leu) cDNA were cloned into separate pCS2+ vectors between *Bam*HI and *Eco*RI restriction sites. In vitro transcription was used to create and isolate mRNA (mMESSAGE mMACHINE™ SP6 Transcription Kit, Invitrogen/Thermo Fisher Scientific, Cat# AM1340). Wild-type embryos were either injected with 200 pg wild-type *msgn1* mRNA (58 embryos) or *msgn1* p.(Arg71Leu) mRNA (54 embryos) at one-cell stage and compared with untreated ones (49 embryos) in three independent experiments. The embryos were fixed at 8-somite stage in 4% paraformaldehyde before in situ hybridization.

#### Zebrafish embryo RNA in situ hybridization

RNA in situ hybridization was performed according to standard protocols [39]. RNA probes were synthesized from cloned partial mRNA sequences of target genes using the DIG or FLU RNA Labeling Kit (Roche, Basel, Switzerland, product-nr. 11685619910 and 11175025910). All detected expressions patterns with newly established RNA probes were in accordance with previously published and ZFIN database patterns. Used in situ probes were targeted against: *tbxta*

(ZDB-GENE-980526-437; synonyms: *T/ta/brachyury/no tail/ntl*), *msgn1* (ZDB-GENE-030722-1; synonyms: *mespo*), *bmp2a* (ZDB-GENE-980526-388), *tbx6* (ZDB-GENE-020416-5; synonyms: *tbx6r/fss/fused somites/tbx24*), *tbx16* (ZDB-GENE-990615-5; synonyms: *spt/spa-detail*). In-situ experiments were performed two times independently and included  $n \geq 10$  embryos per sample and condition. Detailed information and primers used for probe cloning are listed in Additional file 8: Table S2.

### Image and statistical analysis of RNA in situ hybridization embryos

The embryos were flat mounted via dissection of the yolk sac and imaged under a Nikon SMZ1500 stereomicroscope (HR Plan Apo 1X WD 54), Nikon DS-Ri1 digital camera with reflected light at 23 °C room temperature. FIJI software (ImageJ 1.54f) [32] was used to assess intensity of *tbxta* staining in the tailbud using a standardized circular region (120  $\mu\text{m}$  in diameter, ROI seen in Fig. 4A). Images were first inverted, then the mean intensity of anterior tissue background (17  $\mu\text{m}$  in diameter) was subtracted from the mean intensity of tailbud (ROI). Then each intensity was normalized to the average intensity of uninjected embryos in an entire experiment.

We used unpaired two-tailed Kruskal–Wallis nonparametric test without equal standard deviation assumption in Fig. 4D. The statistical tests and distribution calculations (median and quartiles, confidence intervals) were performed in GraphPad Prism 9.5.0 software.

### Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s40246-024-00593-w>.

**Additional file 1. Excel file S1:** Raw Data Fig. 3, ISH signal quantification values and injection statistics

**Additional file 2. Figure S1:** MSGN1 protein structure computer predictions indicate a conformational change, which can be distinguished between the normal protein structure (A) and the p.(Arg125Leu) variant (B). Amino acid alignment of six vertebrate species indicates evolutionary conservation of Arg125 within the basic helix-loop-helix (bHLH) protein domain (C). Numbers indicate position within the human amino acid sequence. Uniport amino acid sequence IDs are given for different species

**Additional file 3. Figure S2:** Next generation sequencing reads of new MSGN1 missense variant of the affected patient

**Additional file 4. Figure S3:** Additional data to *in vitro* transfection of HEK 293T cells with CMV:MSGN1-Tag and CMV:MSGN1-Arg125Leu plasmids at different concentrations presented in Fig. 3. (A) Corresponding immunofluorescence controls after 48h of transfection are shown. (B) Additional data to quantification of MSGN1 localization in transfected cells. Fluorescence signals in a single z-plane were visualized by confocal laser-scanning microscopy and subsequently signal intensity was measured in single cells (cyto: cytoplasm; nuc: nucleus) and outside of cells (bg: background). Graphs show signal intensity measurements of the DAPI channel (nucleus) of 30 cells per experimental group and 30 background positions. Values are given in Excel file S1

**Additional file 5. Figure S4:** Examples of different phenotypes of *msgn1* mosaic zebrafish embryos. Injection with the *msgn1:mCherry-2A-msgn1* plasmid results in different amounts of mCherry positive/*msgn1* expressing cells within the trunk and tailbud region of *tp1:VenusPEST* transgenic embryos. Phenotype severity in the PSM (marked by arrows) 24 h after injection is correlated with *msgn1* overexpressing cell amount and position

**Additional file 6. Figure S5:** Different phenotypes of CMV:MSGN1 plasmid injected zebrafish embryos. Injection with the CMV:MSGN1-FLAG-tag (WT) or with the CMV:MSGN1-FLAG-tag p.(Arg125Leu) plasmids, which have been used for *in vitro* cell transfections, results in alteration of tail development also within the trunk and tailbud region of zebrafish embryos at 18 hpf. Injected embryos display axis bending or cell aggregations in the trunk and PSM regions (marked by black arrows). Number of injected embryos are given in Excel file S1.

**Additional file 7. Figure S6:** Examples of different phenotypes of *msgn1* p.(Arg71Leu) and *MSGN1* p.(Arg125Leu) mosaic *tp1:VenusPEST* zebrafish embryos *in vivo*. Injection with the *sk-tol2-msgn1:mCherry-2A-msgn1* p.(Arg71Leu) or with the *sk-tol2-msgn1:mCherry-2A-MSGN1* p.(Arg125Leu) plasmid results in mCherry positive/*msgn1* variant expressing cells within the trunk and tailbud region of *tp1:VenusPEST* transgenic embryos. Phenotype severity in the PSM (marked by arrows) in a time frame between 16 and 35 hpf after injection is shown. The observed phenotypes correlate with *msgn1* overexpressing cell amount, position and partly differ in severity between zebrafish and human CDS versions. Number of injected embryos are given in Excel file S1.

**Additional file 8. Table S1:** Pathogenicity prediction of *MSGN1* c.374G>T, p.(Arg125Leu) variant. **Table S2:** PCR Primers and DNA plasmids used in this study. **Table S3:** ClinVar missense *MSGN1* VUS variants. **Table S4:** ClinVar structural *MSGN1* variants.

### Acknowledgements

First of all, we thank the family for their participation and for sharing relevant information. The authors additionally thank PD Dr. Christina Lillesaar and cooperating zebrafish laboratories for sharing the *tp1:VenusPEST* transgenic line and hinting us to investigating Notch signaling, Frederik Tessmer for cloning of *tbx* in situ probes, Angela Borst and Alice Schaaf for help with cell culture experiments. The *sk-tol2-msgn1:mCherry-2A-msgn1* plasmid was a kind gift from the Shinji Takada lab (ExCELLS, National Institutes of Natural Sciences, Okazaki, Japan). U.K. and B.V. are members of the European Reference Network on Rare Congenital Malformations and Rare Intellectual Disability (ERN-ITHACA) [EU Framework Partnership Agreement ID: 3HP-HP-FPA ERN-01-2016/739516]. F.B. is a participant in the Clinician Scientist Program (CS4RARE) funded by the Alliance4Rare and associated with the BIH Charité Clinician Scientist Program.

### Author contributions

AK, CL, TH, and DL designed and organized the study. BV and EK did formal analysis and supervised the study. KK, TR, LR, NN, EMÖ, and DL designed, visualized, and performed functional studies. AR, RM, AH, and EGH contributed to the clinical data. FB and UK performed genome sequencing analysis. AK and DL wrote the manuscript, and all the co-authors reviewed the manuscript.

### Funding

Open Access funding enabled and organized by Projekt DEAL. B.V. is supported by the German Research Foundation (DFG) VO 2138/7-1 grant 469177153. D.L. is supported by the German Research Foundation (DFG) LI 2411/2-2 grant 397519724.

### Availability of data and materials

The dataset(s) supporting the conclusions of this article is(are) included within the article and its additional files.

### Declarations

#### Ethical approval and consent to participate

Patient consent to participate in this study was obtained prior to investigation by their parents and/or legal guardians. Ethics approval to this study was given

to Dr. Ehsan Ghayoor Karimiani by the regional Ethics Committee in the field of Research in Medical Sciences at Mashhad University of Medical Sciences (Reference number IR.MUMS.REC.1395.40). All procedures involving experimental animals were performed in compliance with local animal welfare laws (Tierschutzgesetz §11, Abs. 1, Nr. 1, husbandry permit number 568/300-1870/13), European Union animal welfare guidelines (EU directive 2010/63/EU), and best-practice scientific policies (ARRIVE guidelines). All presented experiments have been performed in zebrafish embryos younger than 2 dpf and have been terminated timely before free-swimming and independent feeding larval stages (5 dpf) and thereby do not fall into the German regulatory frameworks dealing with animal experimentation. Procedures involving transgenic animals were in addition performed according to local regulations on this topic (Gentechnikrecht/Gentechnik-Sicherheitsverordnung).

### Competing interests

The authors declare no competing interests.

### Author details

<sup>1</sup>Institute of Human Genetics, Julius-Maximilians-Universität Würzburg, Biozentrum, Am Hubland, 97074 Würzburg, Germany. <sup>2</sup>Division of Developmental Biology, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, USA. <sup>3</sup>Cellular and Molecular Research Centre, Sabzevar University of Medical Sciences, Sabzevar, Iran. <sup>4</sup>Department of Neuromuscular Disorders, UCL Queen Square Institute of Neurology, London WC1N 3BG, UK. <sup>5</sup>University of Cincinnati College of Medicine, Cincinnati, OH, USA. <sup>6</sup>Cellular and Molecular Research Centre, Sabzevar University of Medical Sciences, Sabzevar 009851, Iran. <sup>7</sup>Department of Medical Genetics, Next Generation Genetic Polyclinic, Mashhad, Iran. <sup>8</sup>Institute for Medical Genetics and Human Genetics, Charité - Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt-Universität Zu Berlin, Berlin, Germany. <sup>9</sup>Berlin Institute of Health at Charité - Universitätsmedizin Berlin, Berlin, Germany. <sup>10</sup>Institute of Human Genetics, University Medical Center Göttingen, Göttingen, Germany. <sup>11</sup>Department of Pediatrics, University of Cincinnati College of Medicine, Cincinnati, OH, USA. <sup>12</sup>Institute for Auditory Neuroscience and InnerEarLab, University Medical Center Göttingen, Göttingen, Germany.

Received: 26 January 2024 Accepted: 29 February 2024

Published online: 06 March 2024

### References

- Alestrom P, D'Angelo L, Midtlyng PJ, Schorderet DF, Schulte-Merker S, Sohm F, Warner S. Zebrafish: housing and husbandry recommendations. *Lab Anim*. 2020;54:213–24.
- Chalamalasetty RB, Garriock RJ, Dunty WC Jr, Kennedy MW, Jailwala P, Si H, Yamaguchi TP. Mesogenin 1 is a master regulator of paraxial presomitic mesoderm differentiation. *Development*. 2014;141:4285–97.
- Chen S, Tao J, Bae Y, Jiang MM, Bertin T, Chen Y, Yang T, Lee B. Notch gain of function inhibits chondrocyte differentiation via Rbpj-dependent suppression of Sox9. *J Bone Miner Res*. 2013;28:649–59.
- Chen W, Liu J, Yuan D, Zuo Y, Liu Z, Liu S, Zhu Q, Qiu G, Huang S, Giampietro PF, et al. Progress and perspective of TBX6 gene in congenital vertebral malformations. *Oncotarget*. 2016;7:57430–41.
- Cheng J, Novati G, Pan J, Bycroft C, Zengulyte A, Applebaum T, Pritzel A, Wong LH, Zielinski M, Sargeant T, et al. Accurate proteome-wide missense variant effect prediction with AlphaMissense. *Science*. 2023;381:eadg7492.
- Corallo D, Trapani V, Bonaldo P. The notochord: structure and functions. *Cell Mol Life Sci*. 2015;72:2989–3008.
- Drutman SB, Haerynck F, Zhong FL, Hum D, Hernandez NJ, Belkaya S, Rapaport F, de Jong SJ, Creyten D, Tavernier SJ, et al. Homozygous NLRP1 gain-of-function mutation in siblings with a syndromic form of recurrent respiratory papillomatosis. *Proc Natl Acad Sci U S A*. 2019;116:19055–63.
- Fior R, Maxwell AA, Ma TP, Vezzano A, Moens CB, Amacher SL, Lewis J, Saude L. The differentiation and movement of presomitic mesoderm progenitor cells are controlled by Mesogenin 1. *Development*. 2012;139:4656–65.
- Gruber C, Martin-Fernandez M, Ailal F, Qiu X, Taft J, Altman J, Rosain J, Buta S, Bousfiha A, Casanova JL, et al. Homozygous STAT2 gain-of-function mutation by loss of USP18 activity in a patient with type I interferonopathy. *J Exp Med*. 2020. <https://doi.org/10.1084/jem.20192319>.
- Holtgrewe M, Stolpe O, Nieminen M, Mundlos S, Knaus A, Kornak U, Seelow D, Segebrecht L, Spielmann M, Fischer-Zirnsak B, et al. VarFish: comprehensive DNA variant analysis for diagnostics and research. *Nucleic Acids Res*. 2020;48:W162–9.
- Ioannidis NM, Rothstein JH, Pejaver V, Middha S, McDonnell SK, Baheti S, Musolf A, Li Q, Holzinger E, Karyadi D, et al. REVEL: an ensemble method for predicting the pathogenicity of rare missense variants. *Am J Hum Genet*. 2016;99:877–85.
- Jaganathan K, Kyriazopoulou Panagiotopoulou S, McRae JF, Darbandi SF, Knowles D, Li Yi, Kosmicki JA, Arbelaez J, Cui W, Schwartz GB, et al. Predicting splicing from primary sequence with deep learning. *Cell*. 2019;176(535–548):e524.
- Kimelman D. Tales of tails (and trunks): forming the posterior body in vertebrate embryos. *Curr Top Dev Biol*. 2016;116:517–36.
- Kimmel CB, Ballard WW, Kimmel SR, Ullmann B, Schilling TF. Stages of embryonic development of the zebrafish. *Dev Dyn*. 1995;203:253–310.
- Kircher M, Witten DM, Jain P, O'Roak BJ, Cooper GM, Shendure J. A general framework for estimating the relative pathogenicity of human genetic variants. *Nat Genet*. 2014;46:310–5.
- Kispert A, Herrmann BG. Immunohistochemical analysis of the Brachyury protein in wild-type and mutant mouse embryos. *Dev Biol*. 1994;161:179–93.
- Latimer AJ, Appel B. Notch signaling regulates midline cell specification and proliferation in zebrafish. *Dev Biol*. 2006;298:392–402.
- Lausch E, Hermanns P, Farin HF, Alanay Y, Unger S, Nikkel S, Steinwender C, Scherer G, Spranger J, Zabel B, et al. TBX15 mutations cause craniofacial dysmorphism, hypoplasia of scapula and pelvis, and short stature in Cousin syndrome. *Am J Hum Genet*. 2008;83:649–55.
- Lewis J, Hanisch A, Holder M. Notch signaling, the segmentation clock, and the patterning of vertebrate somites. *J Biol*. 2009;8:44.
- Liao BK, Jorg DJ, Oates AC. Faster embryonic segmentation through elevated Delta-Notch signalling. *Nat Commun*. 2016;7:11861.
- Liao BK, Oates AC. Delta-Notch signalling in segmentation. *Arthropod Struct Dev*. 2017;46:429–47.
- Liu C, Nakamura E, Knezevic V, Hunter S, Thompson K, Mackem S. A role for the mesenchymal T-box gene Brachyury in AER formation during limb development. *Development*. 2003;130:1327–37.
- McMillen P, Holley SA. The tissue mechanics of vertebrate body elongation and segmentation. *Curr Opin Genet Dev*. 2015;32:106–11.
- Morrow ZT, Maxwell AM, Hoshijima K, Talbot JC, Grunwald DJ, Amacher SL. tbx6l and tbx16 are redundantly required for posterior paraxial mesoderm formation during zebrafish embryogenesis. *Dev Dyn*. 2017;246:759–69.
- Mortier GR, Cohn DH, Cormier-Daire V, Hall C, Krakow D, Mundlos S, Nishimura G, Robertson S, Sangiorgi L, Savarirayan R, et al. Nosology and classification of genetic skeletal disorders: 2019 revision. *Am J Med Genet A*. 2019;179:2393–419.
- Ninov N, Borius M, Stainier DY. Different levels of Notch signaling regulate quiescence, renewal and differentiation in pancreatic endocrine progenitors. *Development*. 2012;139:1557–67.
- Odenthal J, Haffter P, Vogelsang E, Brand M, van Eeden FJ, Furutani-Seiki M, Granato M, Hammerschmidt M, Heisenberg CP, Jiang YJ, et al. Mutations affecting the formation of the notochord in the zebrafish, *Danio rerio*. *Development*. 1996;123:103–15.
- Pan Y, Liu Z, Shen J, Kopan R. Notch1 and 2 cooperate in limb ectoderm to receive an early Jagged2 signal regulating interdigital apoptosis. *Dev Biol*. 2005;286:472–82.
- Postma AV, Alders M, Sylva M, Bilardo CM, Pajkrt E, van Rijn RR, Schulte-Merker S, Bulk S, Stefanovic S, Ilgun A, et al. Mutations in the T (brachyury) gene cause a novel syndrome consisting of sacral agenesis, abnormal ossification of the vertebral bodies and a persistent notochordal canal. *J Med Genet*. 2014;51:90–7.
- Rentsch P, Witten D, Cooper GM, Shendure J, Kircher M. CADD: predicting the deleteriousness of variants throughout the human genome. *Nucleic Acids Res*. 2019;47:D886–94.
- Row RH, Tsotras SR, Goto H, Martin BL. The zebrafish tailbud contains two independent populations of midline progenitor cells that

- maintain long-term germ layer plasticity and differentiate in response to local signaling cues. *Development*. 2016;143:244–54.
32. Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, Preibisch S, Rueden C, Saalfeld S, Schmid B, et al. Fiji: an open-source platform for biological-image analysis. *Nat Methods*. 2012;9:676–82.
  33. Searle AG. Curtailed, a new dominant T-allele in the house mouse. *Genet Res*. 1966;7:86–95.
  34. Seelow D, Schuelke M, Hildebrandt F, Nurnberg P. HomozygosityMapper—an interactive approach to homozygosity mapping. *Nucleic Acids Res*. 2009;37:W593–599.
  35. Simsek MF, Chandel AS, Saparov D, Zinani OQH, Clason N, Ozbudak EM. Periodic inhibition of Erk activity drives sequential somite segmentation. *Nature*. 2023;613:153–9.
  36. Sparrow DB, McInerney-Leo A, Gucev ZS, Gardiner B, Marshall M, Leo PJ, Chapman DL, Tasic V, Shishko A, Brown MA, et al. Autosomal dominant spondylocostal dysostosis is caused by mutation in TBX6. *Hum Mol Genet*. 2013;22:1625–31.
  37. Stemple DL. Structure and function of the notochord: an essential organ for chordate development. *Development*. 2005;132:2503–12.
  38. Stemple DL, Solnica-Krezel L, Zwartkruis F, Neuhauss SC, Schier AF, Malicki J, Stainier DY, Abdelilah S, Rangini Z, Mountcastle-Shah E, et al. Mutations affecting development of the notochord in zebrafish. *Development*. 1996;123:117–28.
  39. Thisse C, Thisse B. High-resolution in situ hybridization to whole-mount zebrafish embryos. *Nat Protoc*. 2008;3:59–69.
  40. Tian Y, Xu Y, Fu Q, Chang M, Wang Y, Shang X, Wan C, Marymont JV, Dong Y. Notch inhibits chondrogenic differentiation of mesenchymal progenitor cells by targeting Twist1. *Mol Cell Endocrinol*. 2015;403:30–8.
  41. Unger S, Ferreira CR, Mortier GR, Ali H, Bertola DR, Calder A, Cohn DH, Cormier-Daire V, Girisha KM, Hall C, et al. Nosology of genetic skeletal disorders: 2023 revision. *Am J Med Genet A*. 2023;191:1164–209.
  42. van Eeden FJ, Granato M, Schach U, Brand M, Furutani-Seiki M, Haffter P, Hammerschmidt M, Heisenberg CP, Jiang YJ, Kane DA, et al. Mutations affecting somite formation and patterning in the zebrafish, *Danio rerio*. *Development*. 1996;123:153–64.
  43. Wehn AK, Chapman DL. Tbx18 and Tbx15 null-like phenotypes in mouse embryos expressing Tbx6 in somitic and lateral plate mesoderm. *Dev Biol*. 2010;347:404–13.
  44. Westerfield M. *The Zebrafish book: a guide for the laboratory use of Zebrafish (Danio Rerio)*. Eugene: University of Oregon Press; 2000.
  45. Wittler L, Shin EH, Grote P, Kispert A, Beckers A, Gossler A, Werber M, Herrmann BG. Expression of *Msgn1* in the presomitic mesoderm is controlled by synergism of WNT signalling and Tbx6. *EMBO Rep*. 2007;8:784–9.
  46. Yabe T, Takada S. Mesogenin causes embryonic mesoderm progenitors to differentiate during development of zebrafish tail somites. *Dev Biol*. 2012;370:213–22.
  47. Yoon JK, Wold B. The bHLH regulator pMesogenin1 is required for maturation and segmentation of paraxial mesoderm. *Genes Dev*. 2000;14:3204–14.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.