Hnrnpul1 loss of function affects skeletal and limb development

Danielle L Blackwell[1,2], Sherri D Fraser[1,2], Oana Caluseriu[4], Claudia Vivori[5,6], Paul MK Gordon[2], Amanda V Tyndall[2,3], Ryan E Lamont[2,3], Jillian S Parboosingh[2,3], A Micheil Innes[2,3], François P Bernier[2,3], Sarah J Childs[1,2]

1 Department of Biochemistry and Molecular Biology

2 Alberta Children's Hospital Research Institute

3 Department of Medical Genetics, University of Calgary, 3330 Hospital Drive NW, Calgary, AB, Canada.

4 Department of Medical Genetics, University of Alberta, MSB 8-42B, Edmonton, AB, Canada

5 Centre for Genomic Regulation (CRG), The Barcelona Institute of Science and Technology, Dr. Aiguader 88, Barcelona 08003, Spain

6 Universitat Pompeu Fabra (UPF), Barcelona, Spain

***Corresponding author:**

schilds@ucalgary.ca,

403-220-8277

Running title (32 characters): Hnrnpul1 developmental anomalies

Keywords (6): Hnrnpul1, alternative splicing, zebrafish, fin, tendon, scoliosis

Summary statement: 15 - 30 words: A zebrafish model of loss of Hnrnpul1 shows alternative splicing defects and results in limb growth, craniofacial tendon, and skeletal anomalies.

Abstract: (180 words)

Mutations in RNA binding proteins can lead to pleiotropic phenotypes including craniofacial, skeletal, limb and neurological symptoms. Heterogeneous Nuclear Ribonucleoproteins (hnRNPs) are involved in nucleic acid binding, transcription and splicing through direct binding to DNA and RNA, or through interaction with other proteins in the spliceosome. Here, we show a developmental role for *hnrnpul1* in zebrafish fin and craniofacial development, and in adult onset scoliosis. Furthermore, we demonstrate a role of *hnrnpul1* in alternative splicing regulation. In two siblings with congenital limb malformations, whole exome sequencing detected a frameshift variant in *HNRNPUL1*; the developmental role of this gene in humans has not been explored. Our data suggest an important developmental role of hnRNPUL1 in both zebrafish and humans. Although there are differences in phenotypes between species, our data suggests potential conservation of ancient regulatory circuits involving hnRNPUL1 in these phylogenetically distant species.

Introduction

Mutations in proteins involved in alternative splicing (AS) lead to spliceosomopathies in humans. Despite being expressed ubiquitously, mutations in core and alternative splicing factors can result in tissue-restricted, cell-type specific phenotypes including craniofacial, limb, skeletal and neurological syndromes (Lehalle et al., 2015). Tissue-specificity can occur because of the sensitivity of individual tissues during embryonic development to AS, such as sensitivity of the neural crest in the case of craniofacial anomalies (Lehalle et al., 2015). AS is the process of producing multiple different mRNA transcripts and protein isoforms through the differential selection of splice sites within a single pre-mRNA molecule. AS of pre-mRNA is carried out by the spliceosome, which is a complex of small nuclear RNAs and proteins. AS events mainly include exon skipping, intron retention and alternative 5' or 3' splice sites. For example, mutations in *TXNL4A* (Wieczorek et al., 2014), *EIF4A3* (Favaro et al., 2014), *EFTUD2* (Lines et al., 2012), *SF3B4* (Bernier et al., 2012) and *SNRPB* (Lynch et al., 2014) cause human spliceosomopathies (Lehalle et al., 2015).

Members of the heterogeneous nuclear ribonucleoprotein (hnRNP) family are involved in nucleic acid binding, splicing and transcription. They are present in the spliceosome and contribute directly and indirectly to the processing of pre-mRNA into mature mRNA, with nearly all hnRNP proteins having RNA-binding motifs (Dreyfuss et al., 1993; Geuens et al., 2016). Pathogenic variants associated with human disease have been discovered in hnRNP family members *HNRNPK* (Au et al., 2015) and *HNRNPU* (Poot, 2019). In particular, the hnRNPU family often act as repressors of mRNA splicing (Matlin et al., 2005). hnRNPU proteins are also involved in DNA repair (Hegde et al., 2012; Polo et al., 2012) and U2 snRNP maturation (Xiao et al., 2012). hnRNPUL1 (also known as E1B-AP5), is also a transcriptional repressor (Kzhyshkowska et al., 2003). As little is known about hnRNPUL1 function *in vivo*, we studied its developmental role in the zebrafish model. Zebrafish is an ideal model for this study due to conservation of developmental processes and genetic networks with human, coupled with rapid development.

Zebrafish craniofacial tendons share patterning and developmental pathways with mammalian tendons. Bone, cartilage and tendons of the face are derived from the neural crest. Developing tendons require expression of *scleraxis (scxa)* for specification, followed by expression of differentiation markers *tenomodulin (tnmd)* and *collagen (col1a)* (Chen and Galloway, 2014). Initial tendon specification is independent of interactions with cartilage or muscle. However, fish lacking the myogenic regulators *myod1* and *myf5* fail to maintain *scxa* expression. Mutants that lack cartilage (*sox9a* and *sox9b)* also show abnormal *scxa* expression. Thus signals produced by the developing muscle and cartilage are required for correct patterning and maintenance of tendons (Chen and

Galloway, 2014). Zebrafish mutations leading to a disruption of craniofacial tendon differentiation have been linked to jaw closure defects (Mcgurk et al., 2017).

Paired fins of teleosts (including zebrafish) are ancestral to limbs in tetrapods (including humans) as they are derived from locomotive organs in common ancestral vertebrates (Shubin et al., 1997). Specification of limb and fin bud tissue is marked by expression of *tbx5* as early as 14- 16 hours post fertilisation (hpf) in the zebrafish (Gibert et al., 2006). Following specification, at approximately 23-26 hpf condensation of mesenchymal cells forms the fin bud. Similar to mammals, the fish fin develops an apical ectodermal ridge (AER) as a signalling center driving mesenchymal cell proliferation and growth. In fish, the AER transforms into the apical fold at approximately 36 hpf. The limb/fin bud actively grows as mesenchymal cells organise and begin to differentiate into muscle masses around 46-48 hpf (Grandel and Schulte-Merker, 1998).

In this study, we model loss of function of zebrafish *hnrnpul1* to understand its developmental role. We identify a multi-tissue phenotype involving limb, craniofacial and skeletal abnormalities including idiopathic scoliosis. Through RNA sequencing and alternative splicing analysis we show that alterative splicing events are disrupted in zebrafish *hnrnpul1* mutants. We also identify a homozygous frameshift variant in Heterogeneous Nuclear Ribonucleoprotein U Like 1 (*HNRNPUL1*) gene in two siblings with craniofacial and limb anomalies underscoring the importance to better understand the role of this gene across vertebrate development.

Results

Zebrafish *hnrnpul1* and *hnrnpul1l* loss of function mutations are homozygous viable

hnRNPUL1 has been studied in cell lines, however there are no data on its role at the organismal and developmental level. Therefore, due to our interest in RNA binding proteins in human craniofacial anomalies, we wanted to understand whether mutations in *hnrnpul1* in an animal model would result in developmental anomalies. We use the zebrafish model because of ease of genetic analysis. Due to the genome duplication in the teleost lineage, there are two closely related *HNRNPUL1* orthologues in zebrafish, *hnrnpul1* and *hnrnpul1l* and therefore we used a double knockout strategy. Human HNRNPUL1 protein sequence shows 65% similarity with zebrafish Hnrnpul1 (chromosome 18) and 67% with Hnrnpul1l (chromosome 5) respectively, with the DNA binding (SAP) and protein-protein interaction (SPRY) domains showing higher conservation (76% and 77% similarity for Hnrnpul1 and 83% and 78% similarity for Hnrnpul1l in the SAP and SPRY domains; Fig. 1A). Two siblings with craniofacial and limb malformations (see results below) have a homozygous frameshift variant (NM_007040.5:c.1673dup, p.(Glu560Argfs*17)); thus, we targeted CRISPR guides to make loss of

function alleles in zebrafish with mutations near the human mutation site. To ensure loss of function mutation, a homology-directed repair 'stop cassette' with stop codons in three reading frames was included in the CRISPR-Cas9 injections. *hnrnpul1l*^{Ca52} and *hnrnpul1*^{Ca53} and *hnrnpul1*^{Ca54} alleles were isolated with mutations at homologous locations to the human mutation in the DNA (Fig. S1) and protein (Fig. 1B). *hnrnpul1l*^{Ca52} has a 106-nucleotide insertion, resulting in a frame shift mutation and premature stop codon (Fig. S1B). This is predicted to produce a nonsense protein mutation resulting in truncation after 3 amino acids (Fig. 1B). *hnrnpul1*^{Ca53} has a 35-nucleotide insertion and *hnrnpul1*^{Ca54} has a 63-nucleotide insertion (Fig. S1C, D). Both mutations result in a frame shift and premature stop codon resulting in truncation after 6 amino acids for Ca53 and 11 amino acids for Ca54 (Fig. 1B). All three alleles are predicted to result in truncated protein. No phenotypic differences in the Ca53 and Ca54 *hnrnpul1* alleles were noted.

We next determined the expression pattern of both *hnrnpul1* and *hnrnpul1l* in embryos at 24 hpf. We find that both transcripts are expressed ubiquitously (Fig. 1C,D), in line with previous reports of ubiquitous expression of *hnrnpul1* from 1 cell to Pec fin stage (60 hpf; Thisse and Thisse, 2004). No previous expression analysis is available for *hnrnpul1l*. *hnrnpul1* and *hnrnpul1l* mutations result in nonsense-mediated decay of the transcript as shown by whole mount *in situ* hybridisation (WISH) against *hnrnpul1* and *hnrnpul1l* in wild type, heterozygous and homozygous mutants (Fig. 1C,D), suggesting both are null alleles. Analysis of the gross morphology of *hnrnpul1* and *hnrnpul1l* double homozygous mutants (hereafter *hnrnpul1/1l* mutants) showed low frequency developmental abnormalities including edema and embryo curvature (Fig. S2), however viable and fertile adults were obtained for all allelic combinations including *hnrnpul1/1l* mutants.

***hnrnpul1* mutation results in defects in fin growth but not fin specification**

To understand whether *hnrnpul1* is involved in fin development, a common system affected by disrupted AS, we first determined if there is normal specification of fins. We tested expression of the limb specification markers *gli3, hand2* and *tbx5* using WISH in *hnrnpul1/1l* mutants and found this is unchanged at 24-48 hpf (Fig. 2A-C'). We next analysed embryonic fin size by staining for *col1a1a* which localises to the apical fold of the developing fin bud at 48 hpf (Fig. 2D). *hnrnpul1/1l* mutants have significantly smaller fin bud area (1.0 ± 0.4 mm^2) compared to wild types (1.3 ± 0.4 mm^2, p=<0.0001; Fig. 2E). To ensure that decreased fin size was not due to defective overall growth we quantified eye size, but find no significant difference between *hnrnpul1/1l* mutants and wild types, suggesting that the growth defect, at this stage, is specific to the fin (Fig. 2F).

We next tested whether fin growth is deficient at larval stages. We used Alcian blue staining at 16 dpf. As larval fish differ in growth rates, we measured the fin length as a proportion of standard length (tip to tail) for each animal. We find that *hnrnpul1/1l* mutants have significantly shorter fins compared to wild type with 12.8 ± 0.7% of body length in wild type animals and 12.1 ± 0.7% of body length in *hnrnpul1/1l* mutant animals (p=0.0008; Fig. 2G,H). Taken together, our data suggest that *hnrnpul1* and *hnrnpul1l* play a role in regulating fin growth in embryos and larvae, but not in initial fin bud cell specification.

***hnrnpul1* mutation results in defects in craniofacial tendon development**

One of the most obvious phenotypes visible in *hnrnpul1/1l* mutant larvae is a gaping jaw at 8 dpf as compared wild types (Fig. 3A). We tested whether the gaping jaw is due to mispatterning of pharyngeal skeletal elements. Alcian blue staining showed cartilage is correctly patterned, but the lower jaw is ventrally displaced, leading to a significantly increased incidence of open jaw (wild type = 10%, $hnrnpul1^{+/+}; hnrnpul1l^{-/-}$ = 29%, $hnrnpul1^{+/-}; hnrnpul1l^{-/-}$ = 38%, $hnrnpul1^{-/-}; hnrnpul1l^{-/-}$ =39%; Fig. 3B-D).

As bone and cartilage development appeared normal, we analysed the development of tendons in the embryonic craniofacial region. Expression of *scleraxis* (*scxa*), a tendon specific marker, was used to visualise craniofacial tendons at 72 hpf. We find that *hnrnpul1/1l* mutants have a significantly shorter Sternohyoideus tendon (wild type = 0.66 ± 0.1 µm, *hnrnpul1/1l* = 0.47 ± 0.1 µm, p=<0.0001; Fig. 3E, F). The distance between the most anterior points of the Sternohyoideus tendons is also significantly narrower in *hnrnpul1/1l* mutants compared to wild types (wild type = 0.35 ± 0.09 µm, *hnrnpul1/1l* = 0.30 ± 0.05 µm, p=0.007; Fig. 3G). However, we find no difference in other craniofacial tendons, the Adductor Mandibulae or Palatoquadrate tendons (Fig. S3). Craniofacial tendons develop from neural crest cell (NCC) populations therefore we analysed the expression of the NCC markers *foxd3* and *sox10.* Expression of these specification markers at 12 hpf was normal in *hnrnpul1/1l* mutants (Fig. S5). Thus, the specification of craniofacial bone and cartilage appears normal, but shortening of the Sternohyoideus tendon may contribute to the gaping jaw phenotype by not allowing the mandible to close properly.

***hnrnpul1* mutation results in increased incidence of scoliosis in adult zebrafish**

We noted that young adult *hnrnpul1/1l* mutants show scoliosis at a high frequency. To determine whether this is congenital or idiopathic scoliosis we tested when scoliosis is first visible. 16 dpf larvae

were stained with Alcian blue to visualise the maturing spinal column. Although *hnrnpul1/1l* mutant larvae are significantly smaller (Fig. S4) the development of the spinal column and vertebral structure is normal and comparable to wild type fish at this stage (Fig. 4A). We next compared the incidence of scoliosis in mutants and wild types at 16 weeks, when they are sexually differentiated young adults. We carefully controlled housing density, a factor that influences growth rates. Fish were housed at a density of 10 per 3L tank from 4.5 weeks until 16 weeks of age, at which point they were sacrificed and processed for alizarin red staining to visualise bone. The relative severity of spinal curvature were scored as none, mild, moderate, or severe scoliosis (Fig. 4B). The incidence of scoliosis is significantly higher in *hnrnpul1/1l* mutants (76%) compared to wild types (28%, p=<0.0001; Fig. 4C). There is no difference in the incidence of mild or moderate scoliosis. However, the most severe phenotype only occurs in *hnrnpul1/1l* mutants.

Overall, our data suggest that mutation of *hnrnpul1* and *hnrnpul1l* contributes to idiopathic scoliosis that develops in the larval period, and is visible in the young adult.

***hnrnpul1/1l* mutation leads to disruptions in transcription and alternative splicing**

The hnRNP family regulates mRNA splicing, but knowledge of the specificity and targets of *hnrnpul1* is limited. Thus, we performed paired-end RNA sequencing (RNAseq) to identify differentially spliced events between wild type and *hnrnpul1/1l* mutant embryos at 3 dpf. This developmental stage was chosen as a stage when fins and tendons are actively differentiating. We used VAST-TOOLS to identify splice junctions and characterise splicing events (Irimia et al., 2014). We observed 76 alternative splicing (AS) events: 25 skipped exons, 33 retained introns, 7 alternative 3' splice site and 11 alternative 5' splice site changes (Fig. 5C,E. Table S1). The most differentially expressed exon in our AS analysis is exon 13 of *hnrnpul1l*, as expected from its excision via our CRISPR-Cas9 mutagenesis (Fig 5A,B).

One of the goals of this analysis is to identify potential transcripts that might explain the phenotypes of *hnrnpul1/1l* mutants. Of interest, *basigin (bsg)* shows a 33% reduction in exon 2 usage in *hnrnpul1/1l* mutants compared to wild type. *F-box and leucine rich repeat protein 3 (fbxl3b)* shows a 31% increase in exon 4 usage in *hnrnpul1/1l* mutants compared to wild type, while *telomeric repeat binding factor a (terfa)* shows a 17% increase in exon 12 usage (Fig. 5D). *Heat shock protein a 9 (hspa9)* shows a 37% increase in retention of intron 3-4, while *chromodomain helicase DNA binding protein 4a (chd4a)* showes a 14% increase in retention of intron 30–31 in *hnrnpul1/1l* mutants compared to wild type. *Pseudouridylate synthase 7 (pus7)* shows a 24% reduction in the retention of intron 13 -14 in *hnrnpul1/1l* mutants compared to wild type (Fig. 5F).

Differential gene expression analysis using RPKM from VAST-TOOLS output, identified 311 genes downregulated by at least -1 log_2 fold change and 116 genes upregulated by at least 1 log_2 fold change (Fig. 5G. Table S2). Ingenuity pathway analysis (QIAGEN, Redwood City, CA) identified the photoreceptor pathway as the top differentially expressed pathway, including the down regulation of *pde6c, prph2a* and *arl3l2.* We have not explored vision defects in our animals, as the patients we identified showed no eye deficiencies. We note that glypican 6a (*gpc6a*) expression is lower in *hnrnpul1/1l* mutants. *Gpc6a* is linked to omodysplasia in humans which includes shortening of extremities. Also of note is the lower expression of *growth differentiation factor 3* (*gdf3*) in *hnrnpul1/1l* mutants, an important growth factor in vertebral and skeletal development.

***HNRNPUL1* as a candidate gene for human developmental anomalies**

Whole exome sequencing was conducted on two similarly affected sisters with craniofacial and limb abnormalities. These sisters were born to consanguineous first cousin parents of Arab descent. Given the similarity between the two sisters and that there was no family history of craniofacial or limb abnormalities, this suggested autosomal recessive or biallelic inheritance patterns. Our analysis identified the variant c.1673dup, p.(Glu560Argfs*17) in the gene *HNRNPUL1.* This gene has not been previously associated with a known genetic condition. We did not detect any additional patients with variants in this gene using web-based tools such as Matchmaker exchange (Buske et al., 2015) or GeneMatcher, (https://www.genematcher.org/); (Sobreira et al., 2015); networks that facilitates the matching of phenotypes and genomics.

The probands both presented at birth with multiple skeletal malformations following uncomplicated pregnancies. The older sister was born to a 27-year-old G4P4 woman. Limb malformations were noted on a 27 week ultrasound. The child was born at 38 weeks gestation by spontaneous vaginal delivery with a birth weight of 3.8 kg (75-90th percentile). Outside of limb malformations and dysmorphic features, the child's neonatal examination and clinical course was unremarkable. She was born with bilateral short humeri, right absent ulna with two fixed in extension digits of the right hand; five rays present but with missing carpals, abnormal nails and dorsal creases on the left hand (Fig. 6A, B). Her lower limbs had mid shaft femoral pseudoarthroses, fused tibia to the femoral condyles, absent fibulas and abnormal toes (Fig. 6C, D). Her karyotype showed mosaic Turner syndrome, which is thought to be an incidental finding. Her course has been mostly uncomplicated except for orthopedic issues and intelligence is in the normal range. She is minimally dysmorphic with hypertelorism, upslanting palpebral fissures, prominent eyes and eyelashes and a high palate.

The younger sister was born at term to a 28-year-old G5P5 woman. Her birth weight was 4.03 kg (75-90th percentile) and a head circumference of 39.0 cm (+2.5 SD). She was stable as a neonate but briefly admitted to NICU for assessment of her malformations. She had bilateral fibular agenesis, short and bowed femurs and four metatarsals and tarsals bilaterally (Fig. 6F, G, and H). Upper limb development was normal (Fig. 6I) with the exception of the hands showing despite the presence of five rays, stiff index finger dorsal creases and abnormal nails progressively more severe from ray 5 to 1. Her echocardiogram and abdominal ultrasounds were normal. She was felt to be minimally dysmorphic with a prominent forehead, relative macrocephaly, bitemporal narrowing, hypertelorism with prominent eyes and heavy eyelashes. Her palate was high but there was no clefting. Her clinical course has been mostly unremarkable outside of orthopedic complications and her intelligence is felt to be within the normal range. In addition to their skeletal malformations, both had bicornuate uterus, hirsutism, the first sister showed bilateral hydronephrosis, and the second sister showed Dandy Walker malformation. Clinically the girls share some features with Al-Awadi/Raas-Rothschild/Furhmann syndrome (OMIM 276820), which arises due to homozygous variants in *WNT7A* however *WNT7A* sequencing was negative.

Whole exome sequence analysis was undertaken as part of the Finding of Rare Disease Genes (FORGE) Canada consortium in both affected siblings. Variants were filtered based on rarity in the FORGE internal variant database and ExAC, along with predicted effects on protein function (Beaulieu et al., 2014). Given the family structure, a recessive mode of inheritance was favoured, although heterozygous variants present in both siblings were assessed since germline mosaicism could not be excluded. Rare homozygous variants were identified in both affected siblings in the *PODXL, CFTR, HNRNPUL1, and PRKD2* genes. The rare homozygous variant in the Podocalyxin-like gene (Chr7(GRCh37):g.131195974C>T, NM_001018111.2(PODXL):c.319G>A, p.(Val107Met)) was not considered to be a likely causative candidate due to the fact that a biallelic loss of function variant in this gene was reported in an autosomal recessive juvenile Parkinson family (Sudhaman et al., 2016) and homozygous knockout in mouse results in perinatal lethality due to severe defects in kidney development and omphalocele. Limb development is normal in these embryos (Doyonnas et al., 2001). Likewise, the rare homozygous variant in protein kinase D2 gene (Chr19(GRCh37):g.47204104G>A, NM_016457.4(PRKD2):c.1073C>T, p.(Ala358Val)) was not considered to be candidate as *in silico* analysis (SIFT, PolyPhen2, alignGVGD, MutationTaster) did not predict any damaging effect on the protein and either heterozygous or homozygous loss of PRKD2 in Rhesus monkey and mouse, respectively, leads to hyperinsulinemia and insulin resistance without any reported congenital limb anomalies (Xiao et al., 2018).

The third homozygous variant was in the cystic fibrosis transmembrane conductance regulator gene (Chr7(GRCh37):g.117232214A>T, NM_000492.3(CFTR):c.1993A>T, p.(Thr665Ser)). This variant has been reported as disease-associated in the literature where is has been found in the heterozygous state in two clinically affected individuals of Tunisian and Egyptian descent (Messaond et al., 1996; Naguib et al., 2007). *In vitro* functional studies suggest potentially decreased chloride currents (Vankeerberghen et al., 1998) and a splice enhancer effects leading to partial exclusion of coding sequence (Aznarez et al., 2003). This finding was unexpected as neither affected individual presented with clinical features of cystic fibrosis, although this variant may be associated with a mild presentation. Given the paucity of clinical information on the impact of this variant in the literature, it is impossible to determine if homozygosity for this variant has any clinical impact. Nevertheless given the well studied nature of pathogenic variation in *CFTR* in humans, this gene is not a plausible candidate for the striking developmental anomalies in these siblings.

Lastly, the rare homozygous variant in heterogeneous nuclear riboprotein U-like 1 (Chr19(GRCh37):g.41807595dup, NM_007040.5(HNRNPUL1):c.1673dup, p.(Glu560Argfs*17)) is predicted to result in a frameshift and premature stop approximately two-thirds of the way through the protein and may target the transcript for nonsense mediated decay. *HNRNPUL1* is depleted for loss of function variants in the general population. In the gnomAD database, *HNRNPUL1* has a probability of loss of function intolerance (pLI) of 1, suggesting that loss of function mutations in this gene may be poorly tolerated in humans (Karczewski et al., 2019) and no individuals with homozygous loss of function variants are reported. Segregation of this variant by Sanger sequence analysis of both affected individuals, both parents, and four unaffected siblings provided additional support to this variant where both parents were heterozygous for the c.1673dup variant while none of the four unaffected siblings were homozygous for this variant.

Discussion

Here we show that the phenotypes from loss of *hnrnpul1/1l* are pleiotropic in zebrafish, and affect multiple systems from craniofacial tendons to fin growth and skeletal morphology. We show that *hnrnpul1/1l* controls splicing and expression of mRNAs *in vivo*. We demonstrate that *hnrnpul1* and *hnrnpul1l* are ubiquitously expressed in embryonic zebrafish, but produce phenotypes in a tissue-specific manner. Tissue specific phenotypes likely occur through varying composition of components of the spliceosome within different tissues, and through varying target transcript expression in tissues (reviewed in (Baralle and Giudice, 2017)). Improvements in sequencing technology have helped identify many disease-causing AS mutations, consequently our understanding of how global genome mutations in splicing regulatory networks impact tissue-specific development and disease has also

expanded (Cieply and Carstens, 2015; Suzuki et al., 2019). We anticipate that each developmental phenotype we observe could be due to changes in multiple tissue-specific *hnrnpul1*, making mechanistic analysis complex, but affirming the multi-system role of AS genes.

Involvement of Hnrnpul1 in fin growth

The roles of genes that initiate limb specification, such as retinoic acid, *tbx5*, and *fgf10* are conserved across vertebrates (Mercader, 2007). However, in zebrafish *hnrnupl1/1l* mutants there is no change in expression or splicing of limb specification genes. Similarly, the siblings with the candidate variant have shortened limbs, rather than an absence of limbs. Our data suggest that the specification program is intact in zebrafish *hnrnpul1/1l* mutants and human patients, however the growth program may be impaired.

Growth of the limb occurs via proliferation of limb mesenchyme, driven by Fgf signals from the Apical Ectodermal Ridge (AER) in fish and tetrapods. Fgf10 induces Fgf8 via Wnt3 in chick, mouse and zebrafish (Mercader, 2007; Yano and Tamura, 2013). Although both fish and other vertebrates rely on Fgf signalling, the expression pattern of specific Fgfs is slightly different (reviewed in (Yano and Tamura, 2013)). We used RNA sequencing to identify transcripts and splice variants targeted by *hnrnpul1/1l* that might affect limb growth, and identified some interesting candidates. We identified differential alternative splicing between *hnrnpul1/1l* mutant and wild type fish in *basigin* (*bsg*; also known as CD147) exon 2. Bsg exon 2 encodes a 351 bp/ 117 amino acid immunoglobulin domain, one of 3 Ig domains present in this transmembrane glycoprotein. Exon 2 is also alternatively spliced in human Bsg (Karczewski et al., 2019). The effect of loss of exon 2 and this Ig domain has not been studied in any species. However Bsg/CD147 is stabilised by CBRN, a chaperone that can be bound by thalidomide, a known teratogen that reduces limb size (Eichner et al., 2016). Knockdown of *bsg/cd147* in zebrafish reduces pectoral fin size and phenocopies the teratogenic effects of thalidomide, suggesting that reduction in bsg expression impairs fin growth. Thalidomide is known to have anti-proliferative effects, while *bsg/CD147* promotes proliferation and invasiveness of cancer cells *in vitro* (Yang et al., 2017). It will require additional studies to determine whether AS of *bsg* exon2 modulates its function in proliferation and growth of the fin/limb.

Hnrnpul1 roles in craniofacial tendon development

hnrnpul1/1l mutants have an unusual craniofacial phenotype of an open jaw. Despite apparent displacement of skeletal elements, we observed no defects in skeletal development. We thus

examined the development of tendons. Muscular and tendon development is coupled, as muscle attachment is required for the normal maturation of craniofacial tendons (Chen and Galloway, 2014) and mechanical load is needed for their differentiation (Brunt et al., 2017). Zebrafish embryos that are anesthetised to prevent movement, show reduced jaw muscle activity and reduced tendon cell numbers in the jaw via reduced Wnt16 activity (Brunt et al., 2017). We observe a shorter Sternohyoideus tendon, a tendon that connects the Sternohyoideus muscle to the hyohyal junction at the second pharyngeal arch midline (Mcgurk et al., 2017). This region is referred to as the mandibulohyoid junction and is important for jaw opening. We find that tendons not associated with this junction (Adductor Mandibulae or Palatoquadrate tendons) show no difference in size, consistent with a specific defect in jaw tendon development at the midline as opposed to tendon development in general. A shorter Sternohyoideus tendon would hold the mandible open. Whether the Sternohyoideus is shorter due to decreased tenocyte specification and/or proliferation, or due to decreased activity and mechanical loading in *hnrnpul1/1l* mutants is unknown.

Development of Idiopathic scoliosis in *hnrnpul1* mutants

The human patients are now entering adolescence and have not developed scoliosis, but we will continue to monitor them given that we noted an unexpected scoliosis phenotype in early adult *hnrnpul1*/*1l* zebrafish mutants. Scoliosis is defined as a three-dimensional rotation of the spine to an angle greater than 10°. Congenital scoliosis (CS) is present at birth arising due to a developmental abnormality, while idiopathic scoliosis (IS) develops during childhood or adolescence with no known cause (Goldstein and Waugh, 1973). The etiology of IS remains unknown but it does not appear to be due to vertebral abnormalities (Wajchenberg et al., 2016). Neurological, muscular, growth and even hormonal abnormalities may be associated with IS (reviewed in (Latalski et al., 2017)), however, no conclusive cause has been established. The fact that concordance is much higher in monozygotic twins than dizygotic twins does suggest a genetic link (Kesling and Reinker, 1997). Human IS has been difficult to study in mammalian models because common models such as mouse are quadrupeds and show a difference in spine structure and gravitational load (Ouellet and Odent, 2013). The zebrafish has recently emerged as an excellent model for IS, due to a similar cranial to caudal spinal load and the ease of genetic manipulation, which makes them susceptible to late onset spinal curvatures (Gorman and Breden, 2009). Therefore, we searched for a link between Hnrnpul1 proteins and known genes leading to scoliosis in zebrafish. Zebrafish mutants in *ptk7*, a regulator of Wnt signalling required in ciliated lineages for cilia motility, show late onset scoliosis. Defective cerebrospinal fluid (CSF) flow as a result of *ptk7* mutation leads to scoliosis, potentially by inducing neuroinflammation (Grimes et al., 2016; Hayes et al., 2014; Van Gennip et al., 2018). CSF circulation

is also important for circulation of neurotensin neuropeptides (Zhang et al., 2018) and mechanosensation by Pkd2l1 (Sternberg et al., 2018) and mutants in these genes develop IS. Thus, disruption in CSF circulation leads to multiple downstream consequences that can result in the failure of spine straightening. While we did not find changes in expression or splicing of *ptk7*, *urps* or *pkd2l1* using whole embryo RNA sequencing, it is possible that signal from tissue specific differences may have masked by using bulk sequencing of whole embryo tissue. However, we do find decreased expression of *gdf3* in *hnrnpul1/1l* mutants. Mutations in human *GDF3* cause skeletal abnormalities including vertebral fusion and in some patients, mild scoliosis (Ye et al., 2009). Additionally, as idiopathic scoliosis is late-developing, the RNAseq we undertook at 3 dpf may have been too early to detect changes relevant to motile cilia and CSF circulation. We did however find a decrease in expression of *arl3l2*. ARL3 proteins are present in the cilia and mutations lead to Joubert syndrome in humans (Alkanderi et al., 2018; Powell et al., 2019). Whether *arl3l2* plays a role in cilia activity in the zebrafish is currently unknown but may provide a mechanistic link between *hnrnpul1/1l* and scoliosis.

Alternative splicing changes after loss of Hnrnpul1/1l

Hnrnpul1 has been implicated, but never demonstrated, to be a protein involved in alternative splicing. We provide the first evidence for its role in AS in zebrafish by demonstrating alternative exon usage, intron retention, and alternative 3' and 5' splice sites in *hnrnpul1/1l* mutants.

Interestingly, we show that *hnrnpul1/1l* mutants are smaller at larval stages. Some of the genes we found with alternative splicing changes in zebrafish have been associated with short stature in humans. For instance human *PUS7* (de Brouwer et al., 2018)*, CHD4A* (Weiss et al., 2016), and *FBXL3* are associated with short stature in humans, and are orthologs of *pus7, chd4a* and *fbxl3a* identified in our study (Ansar et al., 2019). The role of *hnrnpul1* in AS of *pus7* and *fbxl3a* should be investigated as possible causes of defective growth.

A large scale zebrafish mutagenesis screen identified *telemetric repeat factor a* (*terfa*) mutants as having a protruding jaw phenotype (Golling et al., 2002). Interestingly, *terfa* has altered AS in *hnrnpul1/1l* mutants. While the function of *terfa* in normal jaw development is not understood, it is possible that disrupted splicing of this gene could be contributing to the phenotype in *hnrnpul1/1l* mutants.

Identification of altered AS of *chd4a* in *hnrnpul1/1l* mutants is particularly interesting. CHD4, a chromodomain helicase DNA Binding protein also participates in DNA repair (Pan et al., 2012). Loss of CHD4 impairs recruitment of BRCA1 to sites of DNA damage. This is intriguing because a second known role for human HNRNPUL1 is directly in DNA repair. HNRNPUL1 is recruited by the MRN

complex to sites of DNA damage to promote DNA resection (Polo et al., 2012), a role separate from its AS and RNA activity that we have examined here. Control of DNA repair directly at the double strand break site as well as through AS of essential DNA repair regulators such as CHD4 highlights a potential dual role for HNRNPUL1 in DNA repair. Whether DNA repair defects contribute to any of the phenotypes in embryonic or adult zebrafish *hnrnpul1/1l* mutants could be investigated going forward.

Relationship of the zebrafish *hnrnpul1/1l* mutants to human variants in *HNRNPUL1*

We generated a zebrafish model for loss of *hnrnpul1*/*1l* genes prompted by the identification of two human siblings with a homozygous *HNRNPUL1* frameshift variant and skeletal and limb anomalies. As no animal loss of function model had been previously generated, our *hnrnpul1/1l* zebrafish mutants were used to test whether *hnrnpul1* mutants had disruptions in alternative splicing and conservation of phenotypic traits in limb and craniofacial development compared with the human patients with a variant in the orthologous gene. Our data suggests that *hnrnpul1* is a ubiquitously expressed gene that shows a remarkable conservation in function in zebrafish and humans. Despite tissue-specific differences, this suggests there are deep similarities in the regulatory circuits in which hnRNPUL1 is involved. Additional mechanistic experiments are needed to fully understand whether the patients' *HNRNPUL1* variant is the cause of their disease, as our analysis is correlative. The patients have variable loss and shortening of bones with some limbs unaffected and others severely affected. Overall, both parents show shortening of elements of the limb zeugopod (humerus and fibula) and variable agenesis of elements in the stylopod (ulna and tibia). The fish does not have correlates of these bones, however mesenchymal outgrowth that forms the fins and limbs (including bones) occurs via a homologous process.

Interestingly, the three strongest phenotypes in the zebrafish mutants (fin defects, craniofacial tendon development and scoliosis) all have links to Wnt signalling. The patients appear to have a very similar phenotype to Fuhrmann or Al-Awadi/Raas-Rothschild/Schinzel phocomelia syndromes, both caused by loss of WNT7A (Woods et al., 2006). Furthermore, pathogenic variants in WNT5A in humans lead to autosomal dominant Robinow syndrome, and mutations in its receptor ROR2 lead to recessive Robinow syndrome in humans (Person et al., 2010; Van Bokhoven et al., 2000). Robinow syndrome is characterised by craniofacial defects, short stature and vertebral segmentation defects. Although we have not detected members of the canonical Wnt pathway undergoing AS in *hnrnpul1* mutants, the phenotypes that we observe are very consistent with disruption of Wnt signalling. It is possible that Wnt is regulated secondarily through other *hnrnpul1* targets that are alternatively spliced, potentially including cilia genes such as *arl3l2*. Thus, even though the link is between *hnrnpul1* and regulation of

441 developmental processes also regulated by Wnt signaling is still unknown, the phenotypic similarities
442 and affected developmental structures suggest similar underlying mechanisms should be explored.

443

444

445

Materials and Methods

Animal and patient data

All Zebrafish (*Danio rerio*) strains were maintained and raised under established protocols (Westerfield, 2000) and all animal procedures were approved by the University of Calgary Animal Care Committee (protocol AC17-0189). Zebrafish embryos were collected and incubated at 28.5°C in E3 medium (5 mM NaCl, 170 µM KCl, 330 µM $CaCl_2$, 300 µM $MgSO_4$, 223µM Methylene blue) and staged in hours post fertilisation (hpf) or days post fertilisation (dpf). When required endogenous pigmentation was inhibited from 24 hpf by addition of 0.003% 1-phenyl-2-thiourea (PTU, Sigma Aldrich) in E3 medium.

This study was part of the Finding of Rare Disease Genes in Canada (FORGE Canada) consortium and approved by the University of Calgary Conjoint Health Research Ethics Board (REB# 23927).

DNA was extracted from all family members from whole blood using Puregene chemistry (Qiagen). Exome capture was undertaken in both affected individuals using the SureSelect 50 Mb All Exon Kit v3 (Agilent) followed by sequencing with a HiSeq2000 (Illumina). Variant calling and annotation were as described in Lynch et al. (2014). Confirmation and segregation of the *HNRNPUL1* c.1406dup variant was performed by PCR amplification (HotStar Taq Plus, Qiagen, Toronto, ON) from genomic DNA and Sanger sequencing with the ABI BigDye Terminator Cycle Sequencing Kit v1.1 (Life Technologies, Burlington, ON) on a 3130xl genetic analyzer (Life Technologies). Sequence subtraction and analysis was performed using Mutation Surveyor software (SoftGenetics, State College, PA).

Generation of *hnrnpul1* and *hnrnpul1l* mutant Zebrafish

CRISPR mutations were created in both the *hnrnpul1* and *hnrnpul1l* genes by injection of guide RNAs at the single cell stage in conjunction with Cas9 mRNA and a homology-directed repair STOP cassette sequence oligonucleotide (Table S3). Guide RNAs were designed to target a location close to the human mutation. Founders were identified by genomic PCR analysis using primers in Table S3. DNA from F1 heterozygotes was cloned into pCR™Blunt II-TOPO® vector (Thermofisher) and sequenced. Mutants are genotyped by PCR using primers described in Table S3, detailed protocol in Supplementary methods.

Whole mount *in situ* hybridisation (WISH)

Embryos were fixed in 4% paraformaldehyde in PBS with 0.1% Tween-20 (PFA) at 4°C overnight and stored in 100% methanol at -20°C until required. All whole-mount *in situ* hybridisation was carried out

478 according to standard protocols (Lauter et al., 2011). Antisense probes for *hnrnpul1*, *hnrnpul1l*, *scxa*,
479 *hand2, tbx5, foxd3* and *sox10* were produced by in vitro transcription using T7 polymerase (Roche), in
480 the presence of digoxigenin-11-UTP (Sigma), from PCR fragments amplified from embryonic zebrafish
481 cDNA (Table S3). Antisense probes for *gli3* and *col1a1a* were a gift from Peng Huang and produced
482 from plasmid clones.

483

484 **Alcian blue and Alizarin red staining**

485 Alcian blue staining of 16 dpf zebrafish was carried out as previously described (Walker and Kimmel,
486 2007). In brief fish were fixed in 4% PFA overnight at 4°C and stained in 0.04% Alcian blue in 100 mM
487 Tris-HCl/ 10 mM $MgCl_2$. Following staining, fish were washed in decreasing concentration of
488 Ethanol/100 mM Tris-HCl to remove excess stain. Fish were bleached in 3% H_20_2/ 0.5% KOH until
489 pigment was lost. Fish were then washed in increasing concentrations of glycerol in 1% KOH until
490 100% glycerol. Fish were imaged in 100% glycerol.

491 Alizarin red staining of adult zebrafish was carried out as previously described (Connolly and Yelick,
492 2010). In brief, adult zebrafish were eviscerated and fixed for 48hrs in 4% PFA at 4°C. Zebrafish were
493 bleached in 30% H_20_2/ 1% KOH for 2 hours followed by 2 hours in 15% H_20_2/ 0.5% KOH. Zebrafish
494 were cleared in 1% trypsin/2% borax solution overnight and stained in 1mg/ml alizarin red in 1% KOH
495 over night. Following staining fish were washed in increasing concentrations of glycerol in 1% KOH
496 until 100% glycerol. Fish were imaged in 100% glycerol.

497

498 **RNA sequencing**

499 Zebrafish were genotyped from excised tail tissue, while matching head tissue from the same embryo
500 was snap frozen at -80°C for RNA extraction after genotyping. Tails were exposed to 25mM NaOH at
501 55°C for 30 minutes then neutralised with 40mM Tris HCl pH5 to extract genomic DNA (Meeker et al.,
502 2007) followed by PCR genotyping. 8 embryos of each genotype were pooled per replicate and total
503 RNA was purified using RNeasy Plus mini kit (Qiagen). Three replicates each of wild type sibling and
504 *hnrnpul1/1l* mutants were sequenced using paired end reads on Illumina NextSeq500 to a read depth
505 of ~100M reads. RNA libraries were prepared using NEBNext Ultra II Directional RNA Library Prep kit
506 (New England Biolabs). Alternative splicing analysis of RNA sequencing data was completed using
507 the Vertebrate Alternative Splicing and Transcript Tools (VAST-TOOLS) v2.2.2, using genome release
508 danRer10 (Irimia et al., 2014). For all events, a minimum read coverage of 10 actual reads per sample
509 was required, as described (Irimia et al., 2014). PSI values for single replicates were quantified for all

types of alternative events, including single and complex exon skipping events (S, C1, C2, C3, ANN), microexons (MIC), alternative 5'ss and 3'ss (Alt5, Alt3) and retained introns (IR-S, IR-C). A minimum ΔPSI of 10% was required to define differentially spliced events upon each knockdown, as well as a minimum range of 5% between the PSI values of the two samples.

Differential gene expression analysis was performed using RPKM output from VAST-TOOL analysis. For each gene p-values were determined by Student's T-test of RPKM values from 3 biological replicates. Log_2 FC was calculated using the mean RPKM for each genotype. A log_2 FC of ≤-1 or ≥1 and a p-value of ≤0.05 were required to define a gene as differentially expressed.

Imaging and analysis

All images were taken on a Zeiss Stemi SV 11 microscope with a Zeiss Axiocam HRc camera. Area and length measurements were completed in ImageJ using the line and measure tools or Zen Blue (Zeiss) using the line tool.

Statistical analysis

All experiments were performed in at least 3 independent biological replicates. All quantitative data are presented as mean ± standard deviation. Statistical analysis was performed using PRISM Graph Pad Software. ns, P>0.05, *P≤0.05, **P≤0.01, ***P≤0.001, ****P≤0.0001.

Acknowledgements:

We would like to thank the patients and their family for participation in this research. We thank the Alberta Children's Hospital Research Institute for funding for DB and SF through the MORPH project. Operating funding was received from the Canadian Institute of Health Research Institute of Genetics Rare Disease Models and Mechanisms (180309-001-001) to SJC.

Patient sequencing was performed under the Care4Rare Canada Consortium funded by Genome Canada, the Canadian Institutes of Health Research, the Ontario Genomics Institute, Ontario Research Fund, Genome Alberta, Genome BC, Genome Quebec, and Children's Hospital of Eastern Ontario Foundation.

We would also like to thank the MORPH committee for guidance, and Juan Valcárcel Juárez and Wendy Dean for helpful comments on the manuscript.

Competing interests:

No competing interests declared.

Data availability:

All RNASeq reads have been deposited with the National Centre for Biotechnology Information in the Gene Expression Omnibus database and are available under the accession GSE144754.

References

Alkanderi, S., Molinari, E., Shaheen, R., Elmaghloob, Y., Stephen, L. A., Sammut, V., Ramsbottom, S. A., Srivastava, S., Cairns, G., Edwards, N., et al. (2018). ARL3 Mutations Cause Joubert Syndrome by Disrupting Ciliary Protein Composition. *Am. J. Hum. Genet.* **103**, 612–620.

Ansar, M., Paracha, S. A., Serretti, A., Sarwar, M. T., Khan, J., Ranza, E., Falconnet, E., Iwaszkiewicz, J., Shah, S. F., Qaisar, A. A., et al. (2019). Biallelic variants in FBXL3 cause intellectual disability, delayed motor development and short stature. *Hum. Mol. Genet.* **28**, 972–979.

Au, P. Y. B., You, J., Caluseriu, O., Schwartzentruber, J., Majewski, J., Bernier, F. P., Ferguson, M., Valle, D., Parboosingh, J. S., Sobreira, N., et al. (2015). GeneMatcher Aids in the Identification of a New Malformation Syndrome with Intellectual Disability, Unique Facial Dysmorphisms, and Skeletal and Connective Tissue Abnormalities Caused by De Novo Variants in HNRNPK. *Hum. Mutat.* **36**, 1009–1014.

Aznarez, I., Chan, E. M., Zielenski, J., Blencowe, B. J. and Tsui, L. C. (2003). Characterization of disease-associated mutations affecting an exonic splicing enhancer and two cryptic splice sites in exon 13 of the cystic fibrosis transmembrane conductance regulator gene. *Hum. Mol. Genet.* **12**, 2031–2040.

Baralle, F. E. and Giudice, J. (2017). Alternative splicing as a regulator of development and tissue identity. *Nat. Rev. Mol. Cell Biol.* **18**, 437–451.

Beaulieu, C. L., Majewski, J., Schwartzentruber, J., Samuels, M. E., Fernandez, B. A., Bernier, F. P., Brudno, M., Knoppers, B., Marcadier, J., Dyment, D., et al. (2014). FORGE Canada consortium: Outcomes of a 2-year national rare-disease gene-discovery project. *Am. J. Hum. Genet.* **94**, 809–817.

Bernier, F. P., Caluseriu, O., Ng, S., Schwartzentruber, J., Buckingham, K. J., Innes, A. M., Jabs, E. W., Innis, J. W., Schuette, J. L., Gorski, J. L., et al. (2012). Haploinsufficiency of SF3B4, a component of the pre-mRNA spliceosomal complex, causes nager syndrome. *Am. J. Hum. Genet.* **90**, 925–933.

Brunt, L. H., Begg, K., Kague, E., Cross, S. and Hammond, C. L. (2017). Wnt signalling controls the response to mechanical loading during zebrafish joint development. *Development* **144**, 2798–2809.

Buske, O. J., Schiettecatte, F., Hutton, B., Dumitriu, S., Huang, L., Hartley, T., Girdea, M., Sobreira, N., Mungall, C. and Brudno, M. (2015). The Matchmaker Exchange API: automating patient matching throught the exchange of structured phenotypic and genotypic profiles. *Hum. Mutat.* **36**, 922–927.

Chen, J. W. and Galloway, J. L. (2014). The development of zebrafish tendon and ligament progenitors. *Development* **141**, 2035–2045.

Cieply, B. and Carstens, R. P. (2015). Functional roles of alternative splicing factors in human disease. *Wiley Interdiscip. Rev. RNA* **6**, 311–326.

Connolly, M. H. and Yelick, P. C. (2010). High-throughput methods for visualizing the teleost skeleton: Capturing autofluorescence of alizarin red. *J. Appl. Ichthyol.* **26**, 274–277.

de Brouwer, A. P. M., Abou Jamra, R., Körtel, N., Soyris, C., Polla, D. L., Safra, M., Zisso, A., Powell, C. A., Rebelo-Guiomar, P., Dinges, N., et al. (2018). Variants in PUS7 Cause

591 Intellectual Disability with Speech Delay, Microcephaly, Short Stature, and Aggressive Behavior.
592 *Am. J. Hum. Genet.* **103**, 1045–1052.

593 **Doyonnas, R., Kershaw, D. B., Duhme, C., Merkens, H., Chelliah, S., Graf, T. and McNagny, K.**
594 **M.** (2001). Anuria, omphalocele, and perinatal lethality in mice lacking the CD34-related protein
595 podocalyxin. *J. Exp. Med.* **194**, 13–27.

596 **Dreyfuss, G., Matunis, M., Piñol-Roma, S. and Burd, C.** (1993). hnRNP Proteins and the
597 Biogenesis of mRNA. *Annu. Rev. Biochem.* **62**, 289–321.

598 **Eichner, R., Heider, M., Fernández-Sáiz, V., Van Bebber, F., Garz, A. K., Lemeer, S., Rudelius,**
599 **M., Targosz, B. S., Jacobs, L., Knorn, A. M., et al.** (2016). Immunomodulatory drugs disrupt the
600 cereblon-CD147-MCT1 axis to exert antitumor activity and teratogenicity. *Nat. Med.* **22**, 735–743.

601 **Favaro, F. P., Alvizi, L., Zechi-Ceide, R. M., Bertola, D., Felix, T. M., De Souza, J., Raskin, S.,**
602 **Twigg, S. R. F., Weiner, A. M. J., Armas, P., et al.** (2014). A Noncoding expansion in EIF4A3
603 causes richieri-costa-pereira syndrome, a craniofacial disorder associated with limb defects. *Am.*
604 *J. Hum. Genet.* **94**, 120–128.

605 **Geuens, T., Bouhy, D. and Timmerman, V.** (2016). The hnRNP family: insights into their role in
606 health and disease. *Hum. Genet.* **135**, 851–867.

607 **Gibert, Y., Gajewski, A., Meyer, A. and Begemann, G.** (2006). Induction and prepatterning of the
608 zebrafish pectoral fin bud requires axial retinoic acid signaling. *Development* **133**, 2649–2659.

609 **Goldstein, L. A. and Waugh, T. R.** (1973). Classification and Terminology of Scoliosis. *Clin. Orthop.*
610 *Relat. Res.* **9**, 10–22.

611 **Golling, G., Amsterdam, A., Sun, Z., Antonelli, M., Maldonado, E., Chen, W., Burgess, S., Haldi,**
612 **M., Artzt, K., Farrington, S., et al.** (2002). Insertional mutagenesis in zebrafish rapidly identifies
613 genes essential for early vertebrate development. *Nat. Genet.* **31**, 135–140.

614 **Gorman, K. F. and Breden, F.** (2009). Idiopathic-type scoliosis is not exclusive to bipedalism. *Med.*
615 *Hypotheses* **72**, 348–352.

616 **Grandel, H. and Schulte-Merker, S.** (1998). The development of the paired fins in the Zebrafish
617 (Danio rerio). *Mech. Dev.* **79**, 99–120.

618 **Grimes, D. T., Boswell, C. W., Morante, N. F. C., Henkelman, R. M., Burdine, R. D. and Ciruna, B.**
619 (2016). Zebrafish models of idiopathic scoliosis link cerebrospinal fluid flow defects to spine
620 curvature. *Science* **352**, 1341–1344.

621 **Hayes, M., Gao, X., Yu, L. X., Paria, N., Henkelman, R. M., Wise, C. A. and Ciruna, B.** (2014). Ptk7
622 mutant zebrafish models of congenital and idiopathic scoliosis implicate dysregulated Wnt
623 signalling in disease. *Nat. Commun.* **5**, 1–11.

624 **Hegde, M. L., Banerjee, S., Hegde, P. M., Bellot, L. J., Hazra, T. K., Boldogh, I. and Mitra, S.**
625 (2012). Enhancement of NEIL1 protein-initiated oxidized DNA base excision repair by
626 heterogeneous nuclear ribonucleoprotein U (hnRNP-U) via direct interaction. *J. Biol. Chem.* **287**,
627 34202–34211.

628 **Irimia, M., Weatheritt, R. J., Ellis, J. D., Parikshak, N. N., Gonatopoulos-Pournatzis, T., Babor,**
629 **M., Quesnel-Vallières, M., Tapial, J., Raj, B., O'Hanlon, D., et al.** (2014). A highly conserved
630 program of neuronal microexons is misregulated in autistic brains. *Cell* **159**, 1511–1523.

631 **Karczewski, K. J., Francioli, L. C., Tiao, G., Cummings, B. B., Alföldi, J., Wang, Q., Collins, R. L.,**
632 **Laricchia, K. M., Ganna, A., Birnbaum, D. P., et al.** (2019). Variation across 141,456 human
633 exomes and genomes reveals the spectrum of loss-of-function intolerance across human protein-

coding genes. *bioRxiv* doi:10.1101/531210.

Kesling, K. L. and Reinker, K. A. (1997). Scoliosis in twins: A meta-analysis of the literature and report of six cases. *Spine* **22**, 2009–2015.

Kzhyshkowska, J., Rusch, A., Wolf, H. and Dobner, T. (2003). Regulation of transcription by the heterogeneous nuclear ribonucleoprotein E1B-AP5 is mediated by complex formation with the novel bromodomain-containing protein BRD7. *Biochem. J.* **371**, 385–393.

Latalski, M., Danielewicz-Bromberek, · A, Fatyga, · M, Latalska, · M, Kröber, · M and Zwolak, · P (2017). Current insights into the aetiology of adolescent idiopathic scoliosis. *Arch. Orthop. Trauma Surg.* **137**, 1327–1333.

Lauter, G., Söll, I. and Hauptmann, G. (2011). Two-color fluorescent in situ hybridization in the embryonic zebrafish brain using differential detection systems. *BMC Dev. Biol.* **11**,.

Lehalle, D., Wieczorek, D., Zechi-Ceide, R. M., Passos-Bueno, M. R., Lyonnet, S., Amiel, J. and Gordon, C. T. (2015). A review of craniofacial disorders caused by spliceosomal defects. *Clin. Genet.* **88**, 405–415.

Lines, M. A., Huang, L., Schwartzentruber, J., Douglas, S. L., Lynch, D. C., Beaulieu, C., Guion-Almeida, M. L., Zechi-Ceide, R. M., Gener, B., Gillessen-Kaesbach, G., et al. (2012). Haploinsufficiency of a spliceosomal GTPase encoded by EFTUD2 causes mandibulofacial dysostosis with microcephaly. *Am. J. Hum. Genet.* **90**, 369–377.

Lynch, D. C., Revil, T., Schwartzentruber, J., Bhoj, E. J., Innes, A. M., Lamont, R. E., Lemire, E. G., Chodirker, B. N., Taylor, J. P., Zackai, E. H., et al. (2014). Disrupted auto-regulation of the spliceosomal gene SNRPB causes cerebro-costo-mandibular syndrome. *Nat. Commun.* **5**, 1–6.

Matlin, A. J., Clark, F. and Smith, C. W. J. (2005). Understanding alternative splicing: Towards a cellular code. *Nat. Rev. Mol. Cell Biol.* **6**, 386–398.

Mcgurk, P. D., Swartz, M. E., Chen, J. W., Galloway, J. L. and Eberhart, J. K. (2017). In vivo zebrafish morphogenesis shows Cyp26b1 promotes tendon condensation and musculoskeletal patterning in the embryonic jaw. *Plos Genet.* **13**,.

Meeker, N. D., Hutchinson, S. A., Ho, L. and Trede, N. S. (2007). Method for isolation of PCR-ready genomic DNA from zebrafish tissues. *Biotechniques* **43**, 610–614.

Mercader, N. (2007). Early steps of paired fin development in zebrafish compared with tetrapod limb development. *Dev. Growth Differ.* **49**, 421–437.

Messaond, T., Verlingue, C., Denamur, E., Pascaud, O., Quere, I., Fattoum, S., Elion, J. and Ferec, C. (1996). Distribution of CFTR mutations in cystic fibrosis patients of Tunisian origin: Identification of two novel mutations. *Eur. J. Hum. Genet.* **4**, 20–24.

Naguib, M. L., Schrijver, I., Gardner, P., Pique, L. M., Doss, S. S., Abu Zekry, M. A., Aziz, M. and Nasr, S. Z. (2007). Cystic fibrosis detection in high-risk Egyptian children and CFTR mutation analysis. *J. Cyst. Fibros.* **6**, 111–116.

Ouellet, J. and Odent, T. (2013). Animal models for scoliosis research: State of the art, current concepts and future perspective applications. *Eur. Spine J.* **22**, S81–S95.

Pan, M. R., Hsieh, H. J., Dai, H., Hung, W. C., Li, K., Peng, G. and Lin, S. Y. (2012). Chromodomain helicase DNA-binding protein 4 (CHD4) regulates homologous recombination DNA repair, and its deficiency sensitizes cells to poly(ADP-ribose) polymerase (PARP) inhibitor treatment. *J. Biol. Chem.* **287**, 6764–6772.

Person, A. D., Beiraghi, S., Sieben, C. M., Hermanson, S., Neumann, A. N., Robu, M. E.,

Schleiffarth, J. R., Billington, C. J., Van Bokhoven, H., Hoogeboom, J. M., et al. (2010). WNT5A mutations in patients with autosomal dominant Robinow syndrome. *Dev. Dyn.* **239**, 327–337.

Polo, S. E., Blackford, A. N., Chapman, J. R., Baskcomb, L., Gravel, S., Rusch, A., Thomas, A., Blundred, R., Smith, P., Kzhyshkowska, J., et al. (2012). Regulation of DNA-End Resection by hnRNPU-like Proteins Promotes DNA Double-Strand Break Signaling and Repair. *Mol. Cell* **45**, 505–516.

Poot, M. (2019). HNRNPU: Key to Neurodevelopmental Disorders such as Intellectual Delay, Epilepsy, and Autism. *Mol. Syndromol.* **9**, 275–278.

Powell, L., Samarakoon, Y. H., Ismail, S. and Sayer, J. A. (2019). ARL3, a small GTPase with a functionally conserved role in primary cilia and immune synapses. *Small GTPases* doi:10.1080/21541248.2019.1703466.

Shubin, N., Tabin, C. and Carroll, S. (1997). Fossils, genes and the evolution of animal limbs. *Nature* **388**, 639–648.

Sobreira, N., Schiettecatte, F., Boehm, C., Valle, D. and Hamosh, A. (2015). New tools for mendelian disease gene identification: PhenoDB variant analysis module; and genematcher, a web-based tool for linking investigators with an interest in the same gene. *Hum. Mutat.* **36**, 425–431.

Sternberg, J. R., Prendergast, A. E., Brosse, L., Cantaut-Belarif, Y., Thouvenin, O., Orts-Del'Immagine, A., Castillo, L., Djenoune, L., Kurisu, S., McDearmid, J. R., et al. (2018). Pkd2l1 is required for mechanoception in cerebrospinal fluid-contacting neurons and maintenance of spine curvature. *Nat. Commun.* **9**, 1–10.

Sudhaman, S., Prasad, K., Behari, M., Muthane, U. B., Juyal, R. C. and Thelma, B. K. (2016). Discovery of a frameshift mutation in podocalyxinlike (PODXL) gene, coding for a neural adhesion molecule, as causal for autosomal-recessive juvenile Parkinsonism. *J. Med. Genet.* **53**, 450–456.

Suzuki, H., Kumar, S. A., Shuai, S., Diaz-Navarro, A., Gutierrez-Fernandez, A., De Antonellis, P., Cavalli, F. M. G., Juraschka, K., Farooq, H., Shibahara, I., et al. (2019). Recurrent noncoding U1 snRNA mutations drive cryptic splicing in SHH medulloblastoma. *Nature* **574**, 707–711.

Thisse, B. and Thisse, C. (2004). hnrnpul1. *Fast Release Clones A High Throughput Expr. Anal. ZFIN Direct Data Submiss.*

Van Bokhoven, H., Celli, J., Kayserili, H., Van Beusekom, E., Balci, S., Brussel, W., Skovby, F., Kerr, B., Percin, E. F., Akarsu, N., et al. (2000). Mutation of the gene encoding the ROR2 tyrosine kinase causes autosomal recessive Robinow syndrome. *Nat. Genet.* **25**, 423–426.

Van Gennip, J. L. M., Boswell, C. W. and Ciruna, B. (2018). Neuroinflammatory signals drive spinal curve formation in zebrafish models of idiopathic scoliosis. *Sci. Adv.* **4**, 1–11.

Vankeerberghen, A., Wei, L., Jaspers, M., Cassiman, J. J., Nilius, B. and Cuppens, H. (1998). Characterization of 19 disease-associated missense mutations in the regulatory domain of the cystic fibrosis transmembrane conductance regulator. *Hum. Mol. Genet.* **7**, 1761–1769.

Wajchenberg, M., Astur, N., Kanas, M. and Martins, D. E. (2016). Adolescent idiopathic scoliosis: Current concepts on neurological and muscular etiologies. *Scoliosis Spinal Disord.* **11**, doi: 10.1186/s13013-016-0066-y.

Walker, M. B. and Kimmel, C. B. (2007). A two-color acid-free cartilage and bone stain for zebrafish larvae. *Biotech. Histochem.* **82**, 23–28.

Weiss, K., Terhal, P. A., Cohen, L., Bruccoleri, M., Irving, M., Martinez, A. F., Rosenfeld, J. A., Machol, K., Yang, Y., Liu, P., et al. (2016). De Novo Mutations in CHD4 , an ATP-Dependent Chromatin Remodeler Gene , Cause an Intellectual Disability Syndrome with Distinctive Dysmorphisms. *Am. J. Hum. Genet.* **99**, 934–941.

Westerfield, M. (2000). *The zebrafish book. A guide for the laboratory use of zebrafish (Danio rerio).* 4th ed. Univ. of Oregon Press, Eugene.

Wieczorek, D., Newman, W. G., Wieland, T., Berulava, T., Kaffe, M., Falkenstein, D., Beetz, C., Graf, E., Schwarzmayr, T., Douzgou, S., et al. (2014). Compound Heterozygosity of Low-Frequency Promoter Deletions and Rare Loss-of-Function Mutations in TXNL4A Causes Burn-McKeown Syndrome. *Am. J. Hum. Genet.* **95**, 698–707.

Woods, C. G., Stricker, S., Seemann, P., Stern, R., Cox, J., Sherridan, E., Roberts, E., Springell, K., Scott, S., Karbani, G., et al. (2006). Mutations in WNT7A cause a range of limb malformations, including fuhrmann syndrome and Al-Awadi/Raas-Rothschild/Schinzel phocomelia syndrome. *Am. J. Hum. Genet.* **79**, 402–408.

Xiao, R., Tang, P., Yang, B., Huang, J., Zhou, Y., Shao, C., Li, H., Sun, H., Zhang, Y. and Fu, X. D. (2012). Nuclear Matrix Factor hnRNP U/SAF-A Exerts a Global Control of Alternative Splicing by Regulating U2 snRNP Maturation. *Mol. Cell* **45**, 656–668.

Xiao, Y., Wang, C., Chen, J. Y., Lu, F., Wang, J., Hou, N., Hu, X., Zeng, F., Ma, D., Sun, X., et al. (2018). Deficiency of PRKD2 triggers hyperinsulinemia and metabolic disorders. *Nat. Commun.* **9**, 1–11.

Yang, S., Qi, F., Tang, C., Wang, H., Qin, H., Li, X., Li, J., Wang, W., Zhao, C. and Gao, H. (2017). CD147 promotes the proliferation, invasiveness, migration and angiogenesis of human lung carcinoma cells. *Oncol. Lett.* **13**, 898–904.

Yano, T. and Tamura, K. (2013). The making of differences between fins and limbs. *J. Anat.* **222**, 100–113.

Ye, M., Berry-Wynne, K. M., Asai-Coakwell, M., Sundaresan, P., Footz, T., French, C. R., Abitbol, M., Fleisch, V. C., Corbett, N., Allison, W. T., et al. (2009). Mutation of the bone morphogenetic protein GDF3 causes ocular and skeletal anomalies. *Hum. Mol. Genet.* **19**, 287–298.

Zhang, X., Jia, S., Chen, Z., Chong, Y. L., Xie, H., Feng, D., Wu, X., Song, D. Z., Roy, S. and Zhao, C. (2018). Cilia-driven cerebrospinal fluid flow directs expression of urotensin neuropeptides to straighten the vertebrate body axis. *Nat. Genet.* **50**, 1666–1673.

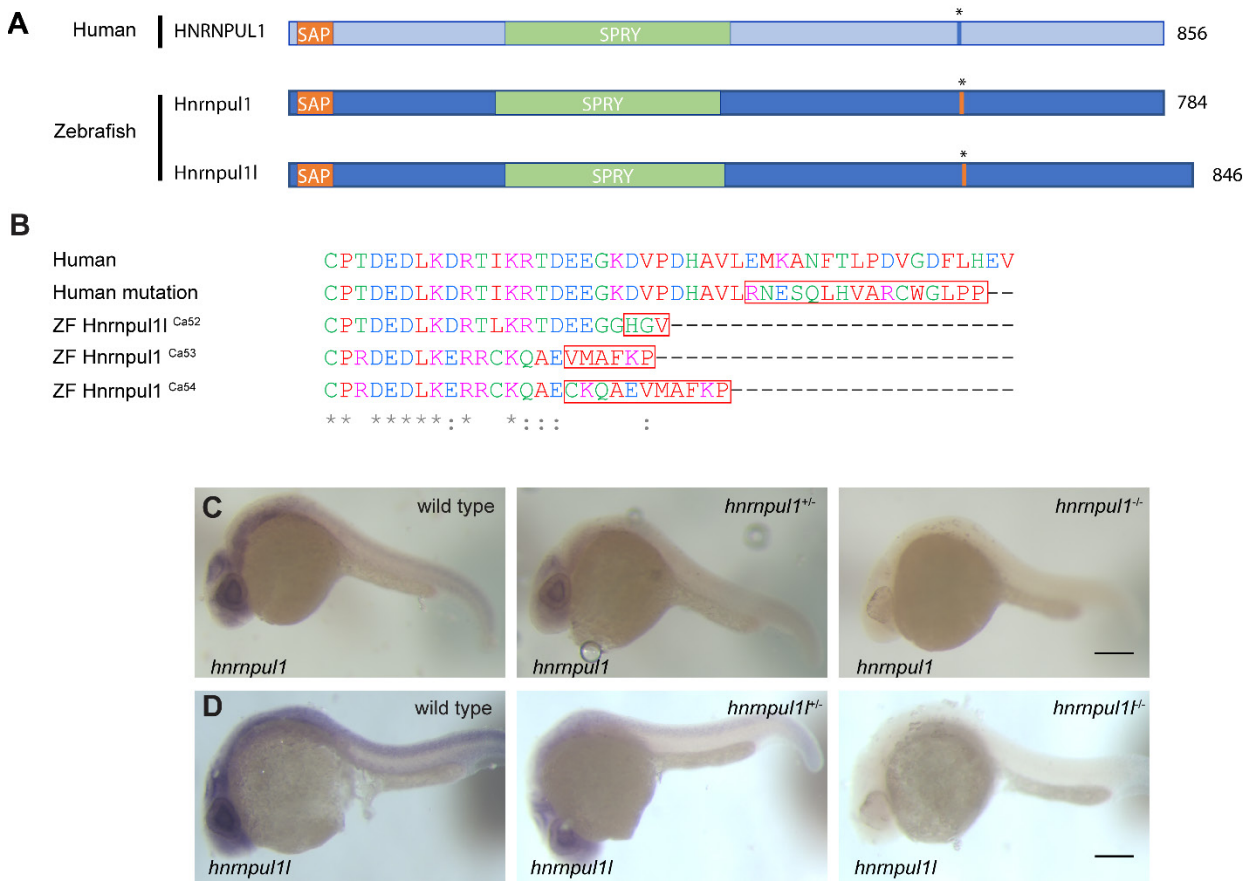

Figure 1 – Mutation of human *HNRNPUL1* and zebrafish *hnrnpul1* and *hnrnpul1l*

A) Schematic showing domains of human HNRNPUL1 and zebrafish Hnrnpul1 and Hnrnpul1l proteins. The mutation location in human HNRNPUL1 and equivalent sequence in zebrafish is marked by *. B) Amino acid sequence of mutations, red boxes indicate nonsense sequence. C,D) whole mount *in situ* hybridisation (WISH) staining for *hnrnpul1*^{Ca52} (C) and *hnrnpul1l*^{Ca53} (D), at 24 hpf reveals ubiquitous expression of both genes in wild types and nonsense mediated mRNA decay of mutant transcripts. Scale bars = 200 µm.

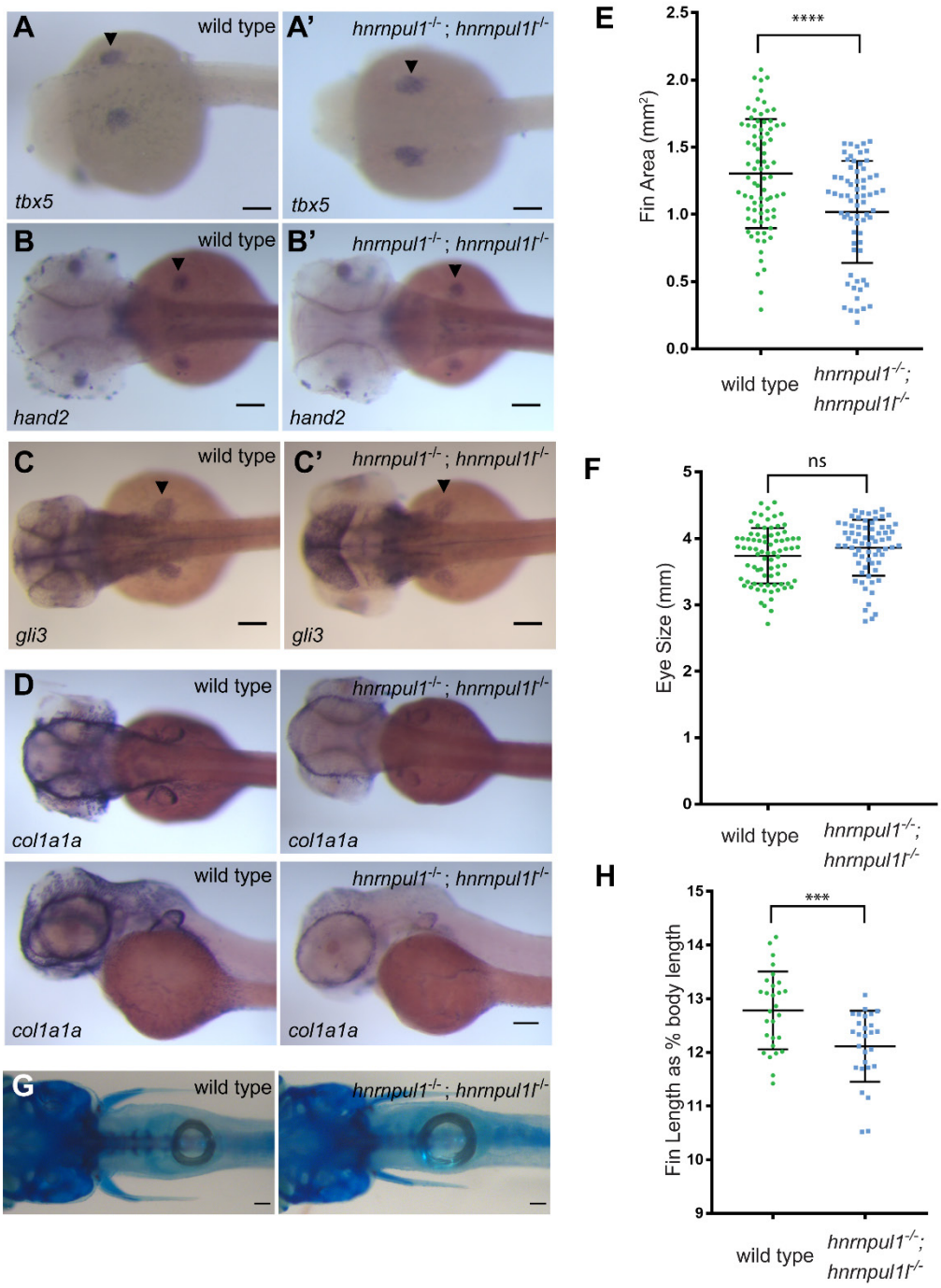

Figure 2 – Loss of *hnrnpul1* and *hnrnpul1l* does not affect fin specification, but leads to decreased fin growth in embryos and larvae

A-C') WISH staining for fin specification markers *tbx5* (A, A') at 24 hpf, *hand2* (B, B') and *gli3* (C, C') at 48 hpf in wild type (A-C) and *hnrnpul1*$^{-/-}$;*hnrnpul1l*$^{-/-}$ double mutant (A'-C') embryos. D) WISH staining for *col1a1a* in wild type and *hnrnpul1*$^{-/-}$;*hnrnpul1l*$^{-/-}$ double mutant embryos at 49

hpf. E and F) Quantification of fin area or eye size in wild type and *hnrnpul1*$^{-/-}$;*hnrnpul1l*$^{-/-}$ double mutant *col1a1a* stained embryos at 49 hpf. Wild type n= 79, *hnrnpul1*$^{-/-}$;*hnrnpul1l*$^{-/-}$ n= 68, from 2 trials. G) Alcian blue cartilage staining of wild type and *hnrnpul1*$^{-/-}$;*hnrnpul1l*$^{-/-}$ double mutant fish at 16 dpf. H) Quantification of fin length at 16 dpf as a percentage of body length. Wild type n= 28, *hnrnpul1*$^{-/-}$;*hnrnpul1l*$^{-/-}$ n= 27, from 2 trials. Scale bars = 100 µm. ***=P≤0.001, **** = P≤0.0001, ns = P>0.05, determined using Student's T-test.

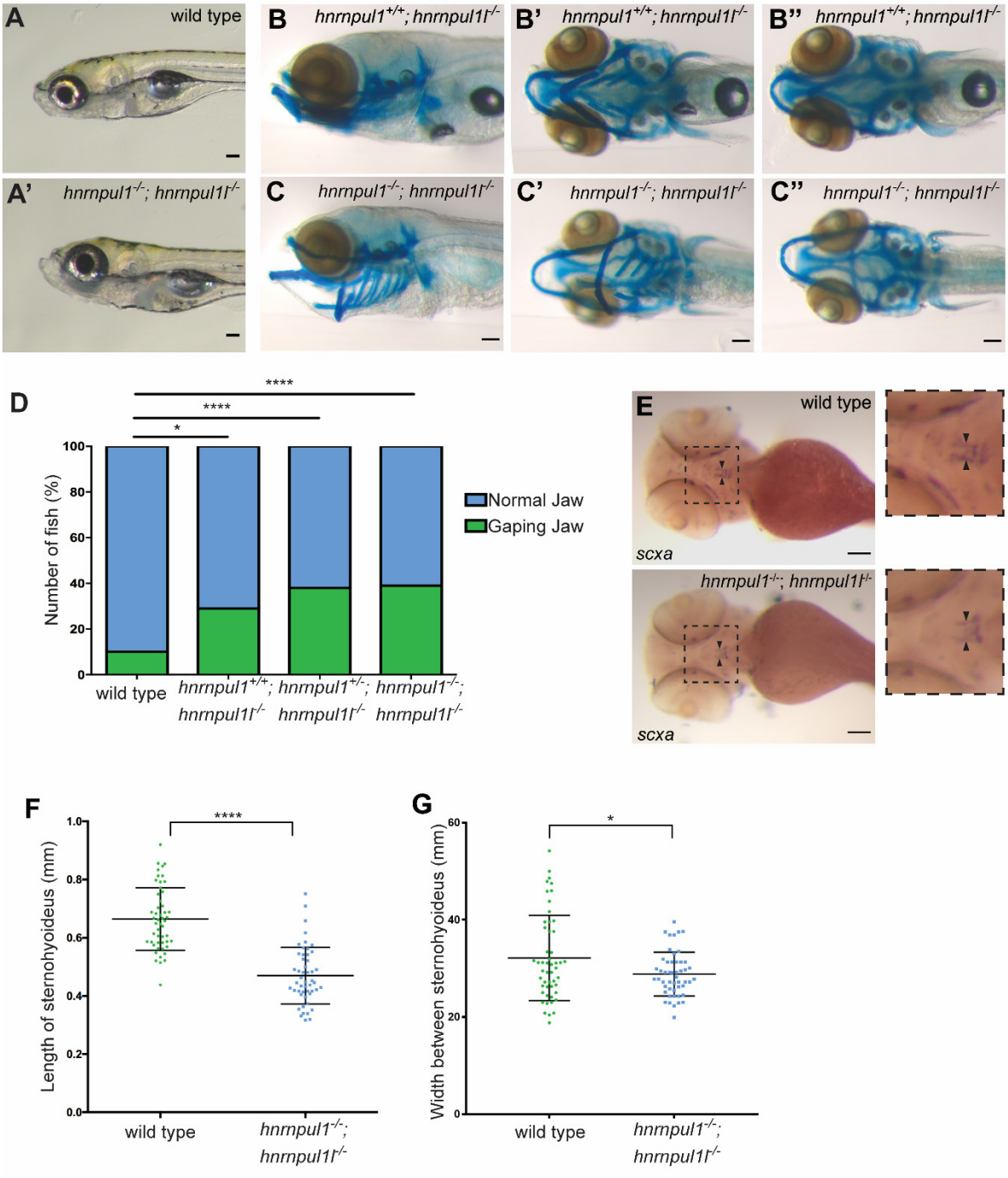

Figure 3 – *hnrnpu1/1l* double mutants show a craniofacial phenotype due to shortened Sternohyoideus tendon

A) Images of live 8 dpf wild type and *hnrnpul1*$^{-/-}$;*hnrnpul1l*$^{-/-}$ double mutant larvae. B-C") Alcian blue staining at 8 dpf. Lateral (B, C), ventral (B', C') and dorsal (B", C") views shown. Example of

a normal jaw phenotype in *hnrnpul1l*$^{-/-}$ single mutant (B) and a gaping jaw phenotype in *hnrnpul1*$^{-/-}$;*hnrnpul1l*$^{-/-}$ double mutant (C). D) Quantification of the proportion of fish showing a gaping jaw phenotype. Wild type n=283, *hnrnpul1l*$^{-/-}$; *hnRNPUL1*$^{+/+}$ n=24, *hnrnpul1l*$^{-/-}$; *hnrnpul1*$^{+/-}$ n=64, *hnrnpul1*$^{-/-}$;*hnrnpul1l*$^{-/-}$ n=84 from 5 trials. * = P≤0.05, **** = P≤0.0001, determined by Fisher's test. E) WISH staining for *scleraxis* (*scxa*) in the Sternohyoideus tendon (arrow heads) in wild type and *hnrnpul1*$^{-/-}$;*hnrnpul1l*$^{-/-}$ double mutant embryos at 72 hpf. F, G) Quantification of the length of, and width between the Sternohyoideus tendons in wild type (n= 51) and *hnrnpul1*$^{-/-}$; *hnrnpul1l*$^{-/-}$ double mutant (n= 50) embryos, from 3 trials. Scale bars = 100 µm. * = P≤0.05, **** = P≤0.0001 determined by Student's T-test.

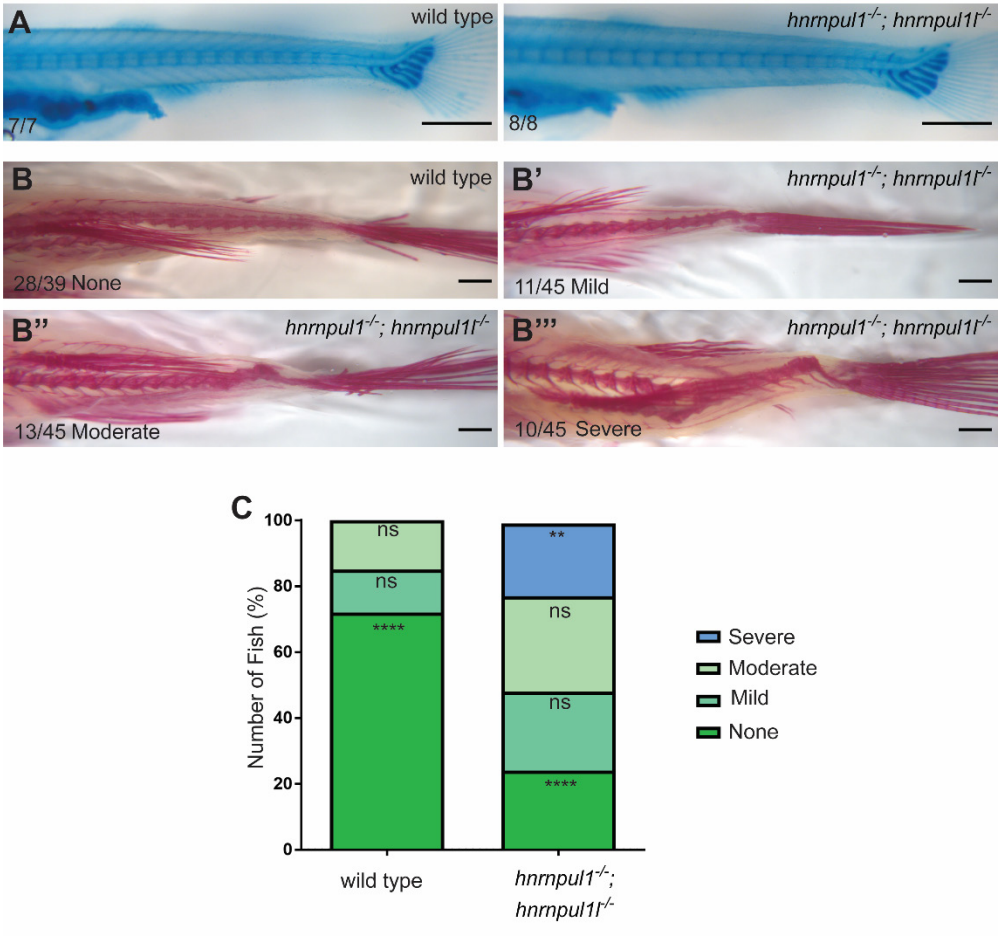

Figure 4 – Loss of *hnrnpu1l; hnrnpul1* leads to idiopathic scoliosis

A) Alcian blue cartilage staining of wild type and *hnrnpul1*$^{-/-}$; *hnrnpul1l*$^{-/-}$ double mutant fish at 16 dpf showing normal larval spinal development. B-B''') Alizarin red bone staining of wild type (B) and *hnrnpul1*$^{-/-}$; *hnrnpul1l*$^{-/-}$ double mutant (B'-B''') fish at 16 weeks of age. Example images of relative mild (B'), moderate (B'') and severe (B''') scoliosis. C) Quantification of the proportion of fish with none, mild, moderate or severe scoliosis. Wild type n=18, *hnrnpul1*$^{-/-}$; *hnrnpul1l*$^{-/-}$ double mutant n=26, from 3 trials. ** = P≤0.01, **** = P≤0.0001 determined by Fisher's test. Scale bar = 500 µm (A) 1000 µm (B-B''').

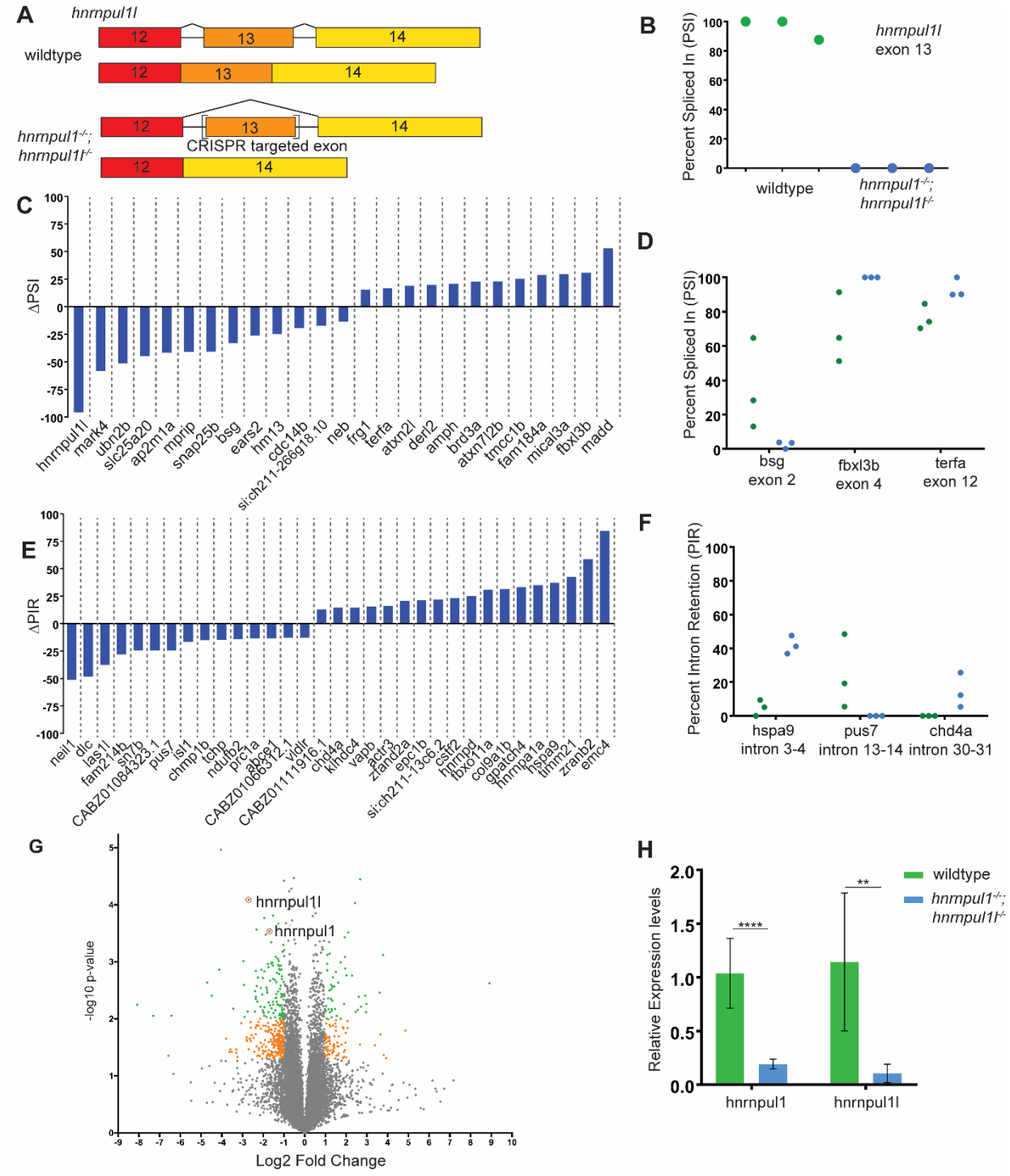

Figure 5 – Loss of *hnrnpu1l*; *hnrnpul1* leads to differential splicing

A) Schematic showing the processing of *hnrnpul1l* to form standard transcript and alternative splicing of exon 13 of the *hnrnpul1l* gene as a result of CRISPR-Cas9 targeted mutatgenesis. B) Percent Spilced In (PSI) for exon 13 of *hnrnpul1l* in wild type and *hnrnpul1*$^{-/-}$; *hnrnpul1l*$^{-/-}$ double

mutant embryos at 3 dpf. C) Change in PSI of all exon skipping events in *hnrnpul1*$^{-/-}$; *hnrnpul1l*$^{-/-}$ double mutant embryos compared to wild type. D) Detailed view of PSI of genes associated with phenotypes, points represent each biological replicate. E) Change in Percent Intron Retention (PIR) of all intron retention events in *hnrnpul1*$^{-/-}$; *hnrnpul1l*$^{-/- -}$ double mutant embryos compared to wild type. F) Detailed view of PIR of genes associate with phenotypes, points represent each biological replicate. Details of affected exon/intron in Table S1. G) Volcano plot showing all differentially expressed genes. Grey points = 1>Log2 FC>-1 or P>0.05, orange points = 1<Log2 FC<-1 and P≤0.05, green points = 1<Log2 FC<-1 and P≤0.01. H) qPCR validation of knock down of *hnrnpul1* and *hnrnpul1l* expression in *hnrnpul1*$^{-/-}$; *hnrnpul1l*$^{-/-}$ double mutant embryos at 3 dpf compared to wild type. ** = P≤0.01, **** = P≤0.0001.

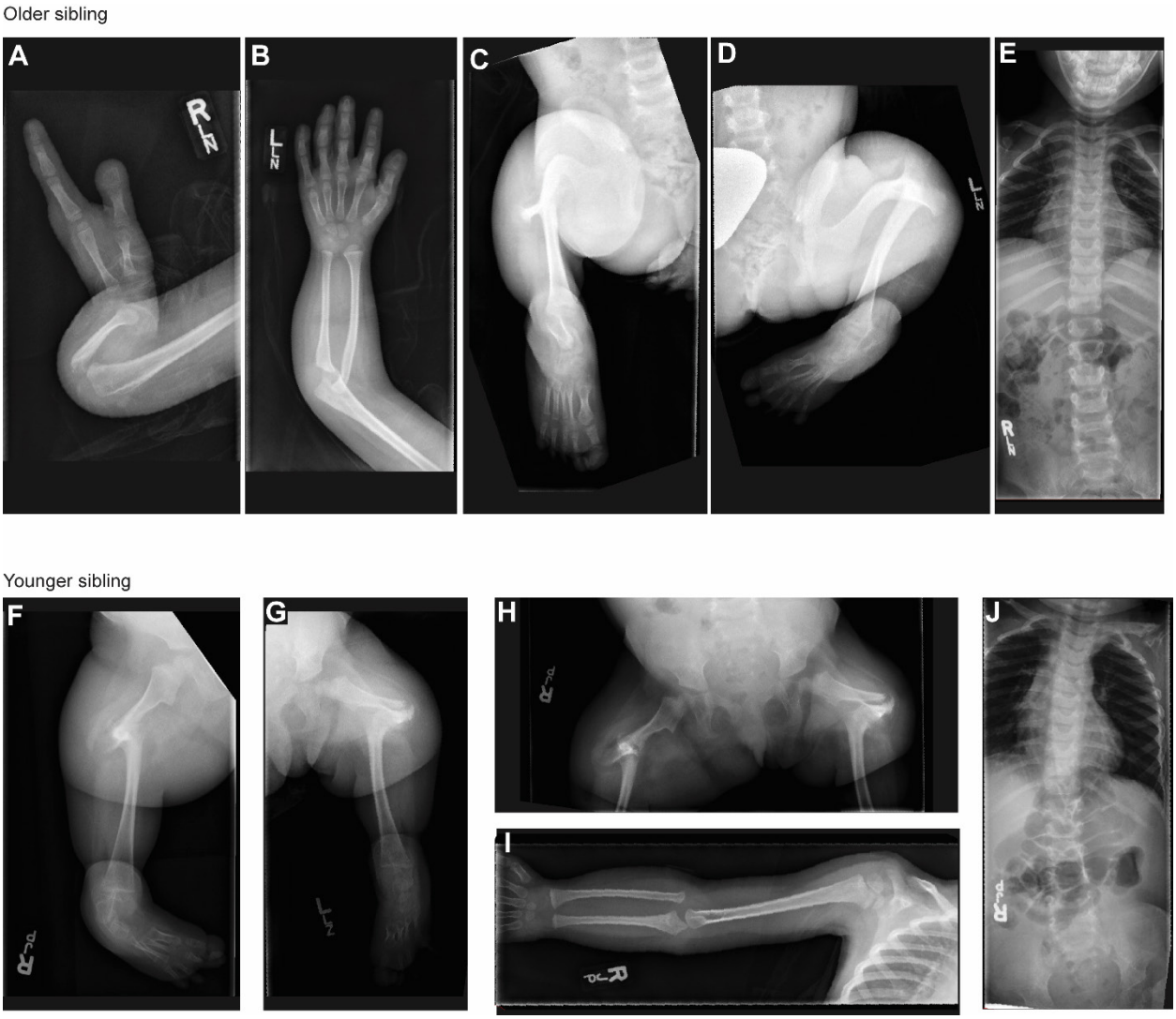

Figure 6 - Radiographic features of affected siblings.

A-E) Xray images of older sibling. Right arm showing short humerus and absent ulna with 2 fixed in extension digits of the right hand (A). Left arm showing short humerus and normal upper arm (B). Right (C) and Left (D) legs showing mid shaft femoral pseudoarthroses, fused tibia to the femoral condyles, absent fibulas and abnormal toes. F-J) X-ray images of the younger sibling Right (F) and Left (G) legs showing bilateral fibular agenesis, short and bowed femurs and four metarsals and tarsals (H). Right arm (I) showing normal upper limb development.