BMC Biology

Regenerating zebrafish scales express a subset of evolutionary conserved genes involved in human skeletal disease

Dylan J. M. Bergen[1,2*] , Qiao Tong[1], Ankit Shukla[3], Elis Newham[1], Jan Zethof[4], Mischa Lundberg[5,6], Rebecca Ryan[1], Scott E. Youlten[7,8], Monika Frysz[2,9], Peter I. Croucher[7,8,10], Gert Flik[4], Rebecca J. Richardson[1], John P. Kemp[3,5,9†], Chrissy L. Hammond[1†] and Juriaan R. Metz[4*†]

Abstract

Background: Scales are mineralised exoskeletal structures that are part of the dermal skeleton. Scales have been mostly lost during evolution of terrestrial vertebrates whilst bony fish have retained a mineralised dermal skeleton in the form of fin rays and scales. Each scale is a mineralised collagen plate that is decorated with both matrix-building and resorbing cells. When removed, an ontogenetic scale is quickly replaced following differentiation of the scale pocket-lining cells that regenerate a scale. Processes promoting *de novo* matrix formation and mineralisation initiated during scale regeneration are poorly understood. Therefore, we performed transcriptomic analysis to determine gene networks and their pathways involved in dermal scale regeneration.

Results: We defined the transcriptomic profiles of ontogenetic and regenerating scales of zebrafish and identified 604 differentially expressed genes (DEGs). These were enriched for extracellular matrix, ossification, and cell adhesion pathways, but not in enamel or dentin formation processes indicating that scales are reminiscent to bone. Hypergeometric tests involving monogenetic skeletal disorders showed that DEGs were strongly enriched for human orthologues that are mutated in low bone mass and abnormal bone mineralisation diseases ($P < 2 \times 10^{-3}$). The DEGs were also enriched for human orthologues associated with polygenetic skeletal traits, including height ($P < 6 \times 10^{-4}$), and estimated bone mineral density (eBMD, $P < 2 \times 10^{-5}$). Zebrafish mutants of two human orthologues that were robustly associated with height (*COL11A2*, $P = 6 \times 10^{-24}$) or eBMD (*SPP1*, $P = 6 \times 10^{-20}$) showed both exo- and endo-skeletal abnormalities as predicted by our genetic association analyses; *col11a2*$^{Y228X/Y228X}$ mutants showed exoskeletal and endoskeletal features consistent with abnormal growth, whereas *spp1*$^{P160X/P160X}$ mutants predominantly showed mineralisation defects.

* Correspondence: dylan.bergen@bristol.ac.uk; juriaan.metz@ru.nl
†John P. Kemp, Chrissy L. Hammond and Juriaan R. Metz have contributed equally to this work and share last authorship.
[1]School of Physiology, Pharmacology, and Neuroscience, Faculty of Life Sciences, University of Bristol, Bristol, UK
[4]Department of Animal Ecology and Physiology, Radboud Institute for Biological and Environmental Sciences, Faculty of Science, Radboud University, Nijmegen, The Netherlands
Full list of author information is available at the end of the article

Conclusion: We show that scales have a strong osteogenic expression profile comparable to other elements of the dermal skeleton, enriched in genes that favour collagen matrix growth. Despite the many differences between scale and endoskeletal developmental processes, we also show that zebrafish scales express an evolutionarily conserved sub-population of genes that are relevant to human skeletal disease.

Keywords: Bone, Collagen, Osteoblast, Musculoskeletal disease, Zebrafish, Scales, Transcriptomics, Extracellular matrix, Ossification

Background

The vertebrate skeleton is composed of non-mineralised cartilage and mineralised bone, enamel, and dentine structures. Skeletal structures have a collagen-rich matrix that provides shape, organ protection, endocrine, and locomotive functions to the organism. Bone is formed by evolutionarily conserved processes via either endochondral or intramembranous (dermal) ossification. Endochondral bone formation occurs through progressive remodelling of a cartilage template into bone; intramembranous bone formation occurs by mesenchymal condensation of osteoblasts. Bones are dynamic tissues and the formation, maintenance and remodelling of the calcified extracellular matrix (ECM) are controlled via a tight coupling between the functional activity of the bone-building osteoblasts and bone-degrading osteoclasts [1]. In most vertebrates, osteoblasts can become entrapped within the bone matrix and differentiate into osteocytes forming a lacunocanalicular network [2]. A fundamental difference between tetrapods and most teleost fish is that bone is devoid of osteocytes in advanced (neo-)teleosts, including the medaka [3]. In zebrafish, most bony elements contain osteocytes while some do not. As osteocytes act as mechanosensors, recent studies showed that both teleost species dynamically express osteocyte marker sclerostin (*sost*) at sites of loading [4]. Zebrafish show a bone remodelling response depending on the degree of exercise-induced loading of bones [5, 6]. Osteoblasts therefore have potential to fulfil a mechano-sensing role to initiate functional coupling to osteoclasts in the bone remodelling cycle and that this can occur without or with reduced osteocyte function.

While skeletal regenerative capacity in mammals is limited to remodelling and injury healing, teleost exoskeletal structures composed of fin rays and elasmoid scales can undergo epimorphic regeneration and subsequent rapid regrowth of a mineralised tissue [7]. This process can therefore offer a potential route to identify new pathways and targets involving calcified matrix formation. Zebrafish are increasingly used as model systems for musculoskeletal (MSK) research, due to their genetic tractability, availability of transgenic reporter lines, and the dynamic imaging opportunities they offer; and because, while there are some differences, the molecular processes underpinning bone formation are comparable with those of terrestrial vertebrates [8]. In recent years, zebrafish fin regeneration has been extensively studied

and has revealed many of the mechanisms that underpin blastema formation, and skeletal dedifferentiation and re-differentiation during tissue regrowth [9].

The elasmoid scales of zebrafish are exoskeletal ectodermal appendages, part of the dermal skeleton, which function as a protective armour and a calcium reservoir [10, 11]. They have two layers: a calcified layer on the exterior and a type I collagen fibrillary plate to the inner side. Whilst lacking osteocytes, scales are decorated with matrix-building cells lining the internal (hyposquamal) and calcified external layer. The nomenclature involving matrix-building cells is diverse and they have interchangeably been called ameloblasts, elasmoblasts, scleroblasts, and osteoblasts [10, 12–15]. Developmental, molecular and genetic studies have helped establish that elasmoid scales, like other elements of the dermal skeleton, likely derive from ancestral odontodes [16, 17]. In fact, in the vertebrate common ancestor, the earliest skeletal tissues to biomineralise were likely to be part of protruding exoskeletal structures from the epidermis that provided protection to the organism. Exoskeletal elements have been strongly reduced or completely lost in terrestrial animals during evolution but have been retained in bony fish [18]. In contrast to lower jaw endochondral bones which are derived from the ectodermal neural crest, dermal scales are formed by mesoderm-derived mesenchymal precursor condensations and not (ectodermal) trunk neural crest cells during zebrafish development [17]. The subsequent epidermal placode (or papilla) formation is under spatiotemporal control of hedgehog, Wnt/β-catenin, Fgf, and Eda signalling pathways [19]. In common with intramembranous flat bones, the scales are then formed and mineralised directly by *de novo* differentiated *sp7*+ cells [20, 21]. On an ontogenetic scale, *sp7*+ cells are located at the hyposquamal side of the plate.

Natural shedding or mechanical removal of a scale initiates a wound healing phase followed by an epimorphic regeneration process and has recently gained attention in studies aiming at a deeper understanding of mechanisms of calcified tissue growth, regeneration, and repair. The regeneration phase leads to the formation of a small mineralised scale as early as two days post-harvest (dph) [15, 22, 23]. Expression of early osteoblast marker *sp7* is detected and increases during growth of the scale, that is

followed by increased expression of mineralisation marker *entpd5a* but also Wnt inhibitor and osteocyte marker *sost* [24]. Regenerating scales have a high density of ECM-secreting cells at the posterior edge of the plate forming the leading growth plane whose dynamics can be tracked *in toto* [12]. These cells form a monolayer and provide conditions that favour precipitation of apatites, such as hydroxyapatite, into a type I collagen-rich matrix. The hyper-mineralised layer on the exterior side can be resorbed by tartrate-resistant acid phosphatase (TRAP)-positive osteoclasts [25, 26]. Similar to endoskeletal bones in fish and terrestrial animals, scale matrix turnover responds to exposure of prednisolone, alendronate, chronic hyperglycaemia, and fatty acids (e.g. OMEGA-6) highlighting evolutionarily conserved mechanisms of calcified matrix metabolism [24, 27–29]. Moreover, the collagen matrix of scales is organised in a woven plywood manner resembling similarities to lamellar bone [30–32] and show a response to mechanical injury (performed *in vivo*) with recruitment of *trapc*-expressing cells to the site of damage [33].

All those features warrant a deeper analysis of gene expression and the pathways involved in dermal scale regeneration. In this study, we describe the transcriptome of ontogenetic and regenerating scales and show that scales have a strong osteogenic expression profile. We also show that differentially expressed genes (DEGs) are enriched for genes associated with different modes of bone formation and human MSK disease.

Results

The transcriptome of ontogenetic and regenerating scales contains bone growth and matrix genes

We aimed to define the transcriptome during peak growth of the regenerating scale. Regenerating scales showed a temporal increase of matrix growth and calcification markers (*sp7*, *entpd5a*, and *col1a1a*) and we observed that this peaked around 9 dph compared to original scales formed during development (ontogenetic) (Fig. 1A). At this time point, the number of *sp7* positive cells were visually increased (Fig. 1B) combined with an increased alkaline phosphatase (ALP) activity (Fig. 1C) and calcein green labelling (Fig. 1D), similar to previous reports [12, 24]. To identify the genes involved in calcified matrix growth during the regeneration process, we performed RNA-sequencing (RNA-seq) on ontogenetic and regenerating scales 9 days into their regeneration.

A total of 13,170 protein coding genes were consistently expressed (i.e. 51.4% of total protein coding genome) in both ontogenetic and regenerating scales (Additional file 1). Principal component analysis (PCA) of the two groups of all genes expressed showed that ontogenetic and regenerating scale groups cluster together (Additional file 2: Fig. S1A) and that there was a high level of correlation

(Pearson's correlation of 97.1%, Additional file 2: Fig. S1B), confirming similar expression patterns within the PCA groups. Setting the arbitrary threshold at 1.25 $\log_2$ fold change and a false discovery rate (FDR) of < 0.05 showed that 604 protein coding genes had substantial differences in gene expression (Additional file 3). We observed a skew towards upregulation of genes ($n = 514$ compared to $n = 90$ downregulated) (Fig. 2); consistent with regenerating tissues such as the caudal fin [34, 35]. The two top DEGs that were highly downregulated were also teleost specific, and a homology search revealed that both genes encoded different proteins [i.e. ENSDARG00000068621 (*si:ch211-181d7.3*) and ENSDARG00000088274 (*si:ch211-181d7.1*)] (Additional file 2: Fig. S2). Both genes have NOD-like receptor (NLR) domains that possess NACHT (NTPase) and leucine-rich repeat (LRR) protein domains, important for innate immune system antigen interactions [36]. To validate the RNA-seq dataset, quantitative real-time PCR (qRT-PCR) was performed on mineralised tissue growth and degradation markers and showed concordant data between RNA-seq and qRT-PCR assays of all selected amplicons (Fig. 3A).

Among the top DEGs, we noticed a high number of collagen and cell adhesion-related genes, but interestingly also *secretory calcium-binding phosphoprotein* (*scpp*) 5 and *scpp7* genes (Table 1). These two *scpp* genes are teleost specific that belong to the *SCCP* family of genes. In vertebrates, this so-called *SCPP* repertoire of genes tends to group together in the genome as two main clusters. During evolution, these two clusters developed high variability between species as a result of a high rate of gene gains and losses but are derived from two *SCPP* precursors (*SPARC* and *SPARCL1*) that originated together with the advent of biomineralisation in a vertebrate common ancestor [37]. In humans, the two clusters on chromosome 4 harbour genes that are predominantly expressed in either (hypermineralised) enamel matrix (*AMBN*, *ENAM*, *AMTN* and *ODAM*) or in bone and dentin matrix (*secreted phosphoprotein 1* (*SPP1*, commonly known as Osteopontin (*opn*)), *DSPP*, *DMP1*, *IBSP* and *MEPE*). In zebrafish, both *scpp* clusters have orthologues that show a form of genomic clustering. The zebrafish equivalent of 'enamel' cluster (chromosome 1) lacks an *AMTN* orthologue but contains *scpp5*, *scpp7* and *scpp9* genes in this region, whilst the orthologous 'bone/dentin' cluster (chromosome 10) only harbours an *SPP1* orthologue and a teleost annotated *scpp8* gene. Both *sparc* and *sparcl1* ancestral *scpp* repertoire genes are also present in the zebrafish genome. As elasmoidin matrix is composed of a lamellar matrix resembling similarity to enamel, we wondered whether the 'enamel' or 'bone/dentin' *scpp* repertoire gene clusters were differentially expressed in regenerating scales. A look up of the gene expression of all

Bergen *et al. BMC Biology* (2022) 20:21

Page 4 of 25
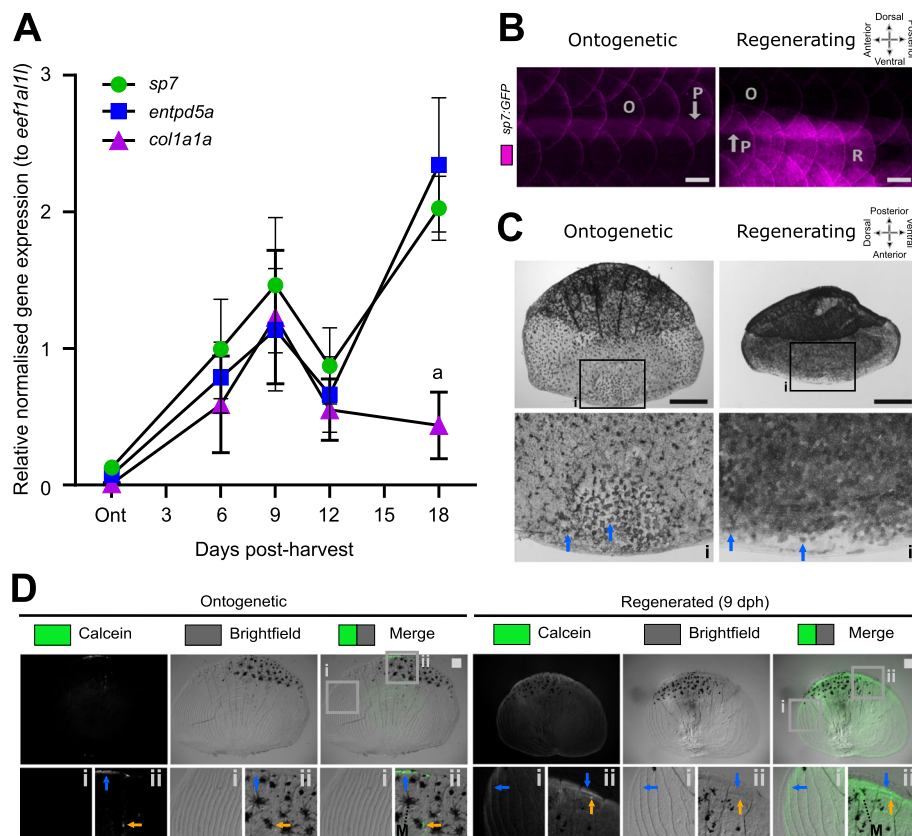
Fig. 1 Regenerating scales express metabolic active bone growth markers at 9 days post-harvest. **A** Quantitative real-time PCR of osteoblast (*sp7*), mineralisation (*entpd5a*), and matrix growth (*col1a1a*) amplicon expression (*n* = 4 fish). Two-way ANOVA analysis showed that there was no interaction factor variance (*p* = 0.133) but that time (*f*(4) = 9.47, *p* < 0.001) and amplicon (*f*(2) = 4.62, *p* < 0.05) had independent statistically significant effects. Tukey's multiple comparison determined that *col1a1a* was significantly lower expressed at 18 days post-harvest (dph) compared to other amplicons (a). **B** *In toto* fluorescent stereomicroscope images of flanks from *sp7:GFP* transgenic fish showing ontogenetic (O) and 9 dph regenerated scales (R) in an anterior (left) to posterior (right) direction. Note the auto-fluorescent pigment stripe (P) and the compass orientation is used to depict *in toto* scales in this paper with the distal (posterior) edge pointing caudally. **C** *In situ* alkaline phosphatase staining of a pre-collection and 9 dph scale with insets (i) showing anterior part of the scale. Blue arrows indicate ALP+ cells which appear to be larger and more extended in ontogenetic than regenerating scales. Compass orientation is used to depict *in situ* scales in this paper as the anterior region was closest to the scale dermal socket (pre-harvest). **D** Representative stereomicroscope images of live calcein green staining labelling newly deposited calcium phosphate in ontogenetic and regenerating scales (*n* = 4 fish each condition). Blue arrow indicates elevated levels of calcein green labelling compared to surrounding signal whereas orange arrows indicate small puncta of enhanced signal. Insets show lateral circuli (i) and posterior epidermal (containing pigmenting melanocytes (M)) region (ii). Scale bar: 100 μm

zebrafish orthologues including zebrafish specific *scpp* genes showed that the 'enamel' cluster, *ambn*, *enam* and *odam* were not or very lowly expressed whilst teleost specific *scpp5* and *scpp7* were differentially upregulated (Table 2). The 'bone/dentin' cluster *spp1* gene was substantially upregulated as was ancestral *SCPP sparc* (osteonectin (*osn*)) that is also associated with bone development (Table 2). We next assessed the temporal expression of DEG *SCPP* genes (*spp1*, *scpp5*, *scpp7* and *sparc*) during scale regeneration. This showed that these four genes peaked their expression at 9 dph (Fig. 3B), and this pattern was similar to *col1a1a* and not *sp7* or *entpd5a* (Fig. 1A). These observations indicate that the elasmoid scale growth relies on 'bone/dentin' type of *scpp* repertoire gene expression rather than 'enamel'

cluster genes. Scale growth therefore resembles an osteoanabolic process.

Indeed, amongst the DEGs we observed osteoblast factors such as *sp7*, *entpd5a/b*, *postnb*, *yap1*, *smpd3*, *phospho1*, *plod1a*, *plod2* and *plod3*; genes related to ECM growth, collagens (predominantly fibrillar collagens) and collagen remodelling including *bmp1a*, *mmp9*, *spp1*, *sparc*, *col1a1a/b*, *col1a2*, *col5a1*, *col5a2a/b*, *col10a1a/b*, *bgna/b*, *col11a1a/b* and *col11a2* [38]. We also observed elevated expression of genes normally associated with endochondral ossification: *col11a1a/b* and *col11a2*, *ihha*, *ptch2*, *dlx5a*, *ostn* (osteocrin) and *mef2cb*, suggesting that scale DEGs have a relationship with both modes of ossification [39]. Osteoclast and monocyte-related factors (e.g. *tnfrsf11a* (*RANK*), chloride channel *clc7n7*,

Fig. 2 (See legend on next page.)

Fig. 2 Volcano plot of RNA-seq data comparing expression between ontogenetic and 9 days post-harvest regenerating scales. Relative fold change ($\log_2$) and false discovery rate ($-\log_{10}$ converted) of coding RNA sequences that were expressed in both ontogenetic and regenerating scales. Green ($\geq \pm 1.25$ fold change, ≥ 1.3 FDR), magenta ($\geq \pm 1.25$ fold change, < 1.3 FDR), blue ($< \pm 1.25$ fold change, ≥ 1.3 FDR), and black (not passing any threshold) coloured dots mark the different criteria

cathepsin K (*ctsk*) and *acp5a/b* (*trapc*)) but also markers associated with osteocyte differentiation; *sost*, *wnt3a*, *lrp4*, *lrp5* and *dkk1a* were expressed in both ontogenetic and regenerating scales, but differential expression was not observed (Additional file 1). These results indicate osteoblast genes are elevated to a greater degree than those associated with osteoclasts concordant with an osteoanabolic process.

The collagen processing pathway is highly enriched and upregulated during scale regeneration

To identify biological pathways overrepresented among DEGs, we performed gene ontology (GO) enrichment analysis. We performed PANTHER GO analysis and in accordance with a regenerating growing tissue there were many GO terms that included 'morphogenesis', 'development', 'regeneration' and 'collagen' in their descriptions (Fig. 4A and Additional file 4). There were also bone-specific GO terms such as 'regulation of bone biomineralization' (GO:0110149; 10.9-fold, FDR=0.018), 'ossification' (GO:0001503; 6.0-fold, FDR=0.0019), and 'skeletal system development' (GO: 0001501; 3.6-fold, FDR=5.3×10^{-6}). Additional bone matrix associated processes such as 'calcium-ion binding', 'glycosaminoglycan binding' and 'matrix metalloprotease (MMP) activity' were identified (Additional file 2: Fig. S3 and S4). A separate GOrilla GO analysis broadly confirmed these

findings and visualised the GO terms by hierarchical clustering (Fig. 4B; Additional file 2: Fig. S3 and S4; Additional file 5).

We also observed cell adhesion-related terms and PANTHER 'Pathways' showed that the 'integrin signalling pathway' (P00034; 3.5-fold, FDR=0.016) was the only enriched signalling pathway, which is known to bind collagen during cell adhesion-related processes [40] (Additional file 5). To further identify potential clusters of genes, PANTHER 'Reactome pathways' protein-protein interactions (PPI) analysis was used. This showed consistent enrichment of players involved in collagen matrix biosynthesis (e.g. 'O-linked glycosylation' and 'ECM proteoglycans') but also ECM interactors such as neural cell adhesion molecule (NCAM) and insulin-like growth factor (IGF) (Fig. 4C and Additional file 6). STRING Network analysis showed that the DEGs PPI network has 6.6-fold higher number of nodes (connections) compared to the whole reactome ($P < 1.0 \times 10^{-16}$). There were ten clusters with distinct functions that were significantly enriched that were highly associated with collagen-rich ECM and cell adhesion, but also with IGF and hedgehog signalling (Fig. 5). Note that STRING added 10 'high-scoring' interactors to the network that were not DEGs (Additional file 2: Fig. S5A). Additional four unlinked clusters were related to cytoskeletal structure or cell motility (Fig. 5 and Additional file 2: Table
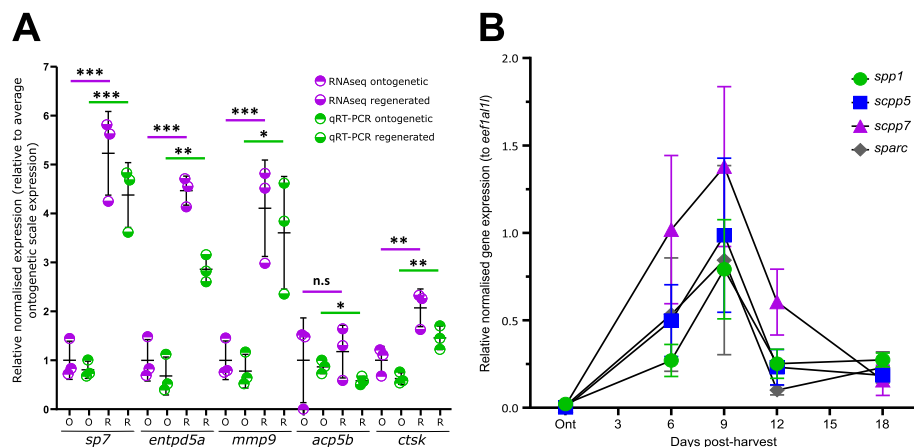
Fig. 3 Validation of RNA-sequencing dataset and SCPP repertoire gene expression during scale regeneration. **A** Comparison of relative expression levels of qRT-PCR (unpaired *t* test) and transcriptomic analysis (false discovery rate) of selected amplicons. Total RNA used for RNA-seq experiment was used. **B** Quantitative real-time PCR (qRT-PCR) of secretory calcium-binding protein (SCPP) family genes that were differentially expressed in the RNA-seq dataset (*n* = 4 fish). Two-way ANOVA analysis showed that there was no interaction factor variance (*p* = 0.882) but that time (*f*(4) = 11.05, *p*=0.001) was the only independent statistically significant effect on amplicon expression

Table 1 Top-40 smallest differentially regulated transcripts. Thirty-eight out of 40 of the top-40 differentially expressed genes are up (green) regulated, whilst only two were down (magenta) regulated. Log$_2$ fold change in regenerated versus ontogenetic scales with false discovery rate (FDR) are shown. Top-40 was selected based on the lowest FDR numerical

Ensembl ID	Gene symbol	Log$_2$ Fold Change	FDR	Gene description
ENSDARG00000068621	*si:ch211-181d7.3*	-6.767	2.4×10^{-98}	NOD-like receptor (NLR) with NACHT and LRR domain
ENSDARG00000074132	*scpp7*	3.756	2.8×10^{-70}	secretory calcium-binding phosphoprotein 7
ENSDARG00000093659	*entpd5b*	3.560	2.4×10^{-48}	ectonucleoside triphosphate diphosphohydrolase 5b
ENSDARG00000025089	*loxl4*	3.167	1.4×10^{-47}	lysyl oxidase-like 4
ENSDARG00000054753	*col10a1a*	3.610	3.1×10^{-43}	collagen, type X, alpha 1a
ENSDARG00000052633	*si:ch211-106n13.3*	4.676	7.8×10^{-39}	si:ch211-106n13.3
ENSDARG00000039999	*slc13a5b*	4.351	1.5×10^{-37}	solute carrier family 13 member 5b
ENSDARG00000012422	*col11a2*	3.484	6.8×10^{-37}	collagen, type XI, alpha 2
ENSDARG00000069216	*c6ast1*	4.275	1.0×10^{-33}	six-cysteine containing astacin protease 1
ENSDARG00000098226	*smpd3*	3.226	1.2×10^{-32}	sphingomyelin phosphodiesterase 3
ENSDARG00000004627	*panx3*	2.897	1.3×10^{-27}	pannexin 3
ENSDARG00000104267	*postnb*	2.775	1.2×10^{-26}	periostin, osteoblast specific factor b
ENSDARG00000024847	*col5a2b*	2.783	2.8×10^{-26}	collagen, type V, alpha 2b
ENSDARG00000009014	*col11a1b*	3.298	2.8×10^{-26}	collagen, type XI, alpha 1b
ENSDARG00000020007	*col1a2*	2.951	6.0×10^{-26}	collagen, type I, alpha 2
ENSDARG00000105153	*ifitm5*	3.133	1.3×10^{-25}	interferon induced transmembrane protein 5
ENSDARG00000042186	*paplnb*	3.405	7.4×10^{-25}	papilin b, proteoglycan-like sulfated glycoprotein
ENSDARG00000095259	*tmem119a*	2.622	6.1×10^{-24}	transmembrane protein 119a
ENSDARG00000088274	*si:ch211-181d7.1*	-4.630	6.1×10^{-24}	si:ch211-181d7.1
ENSDARG00000059746	*plod1a*	2.209	7.5×10^{-24}	procollagen-lysine, 2-oxoglutarate 5-dioxygenase 1a
ENSDARG00000115341	*col11a1a*	3.421	9.5×10^{-24}	collagen, type XI, alpha 1a
ENSDARG00000035809	*col1a1b*	2.926	1.3×10^{-23}	collagen, type I, alpha 1b
ENSDARG00000091667	*si:dkey-234i14.3*	3.501	4.0×10^{-23}	si:dkey-234i14.3
ENSDARG00000079572	*plcd3b*	3.515	1.5×10^{-22}	phospholipase C, delta 3b
ENSDARG00000058270	*si:ch211-114l13.7*	2.465	1.8×10^{-22}	si:ch211-114l13.7
ENSDARG00000055175	*c1qtnf6a*	3.482	2.6×10^{-22}	C1q and TNF related 6a
ENSDARG00000037145	*slc8a4b*	3.076	4.4×10^{-21}	solute carrier family 8 member 4b
ENSDARG00000101535	*col10a1b*	3.032	6.0×10^{-21}	collagen, type X, alpha 1b
ENSDARG00000078866	*sdk1a*	1.876	8.4×10^{-21}	sidekick cell adhesion molecule 1a
ENSDARG00000041703	*rrbp1b*	2.696	1.3×10^{-20}	ribosome binding protein 1b
ENSDARG00000040985	*itgbl1*	3.531	9.4×10^{-20}	integrin, beta-like 1
ENSDARG00000019949	*serpinh1b*	3.135	1.2×10^{-19}	serpin peptidase inhibitor, clade H (heat shock protein 47), member 1b
ENSDARG00000026348	*csad*	2.915	3.7×10^{-19}	cysteine sulfinic acid decarboxylase
ENSDARG00000012593	*col5a1*	2.447	5.7×10^{-19}	procollagen, type V, alpha 1
ENSDARG00000022437	*cd81b*	2.749	5.9×10^{-19}	CD81 molecule b
ENSDARG00000021265	*mybpc2b*	6.520	1.0×10^{-18}	myosin binding protein C, fast type b
ENSDARG00000009418	*mef2cb*	2.062	1.2×10^{-18}	myocyte enhancer factor 2cb
ENSDARG00000017884	*bgna*	2.896	3.3×10^{-18}	biglycan a
ENSDARG00000062363	*phex*	2.034	5.5×10^{-18}	phosphate regulating endopeptidase homolog, X-linked
ENSDARG00000079486	*fam20a*	2.263	1.1×10^{-17}	FAM20A Golgi associated secretory pathway pseudokinase

Table 2 Differential expression of scpp gene clusters in scales corresponding with human orthologue SCPP clusters associated with site-specific expression

Evolutionary origin	Zebrafish gene symbol	Chr #	Transcript start site	Log₂ Fold Change	False Discovery Rate	Mean NRC	Human gene symbol	Chr #	Transcript start site	Alias
Bone/dentin	*spp1*	10	15,120,711	2.29	1.6×10^{-7}	2459.1	*SPP1*	4	87,975,650	osteopontin
Bone/dentin	*N/A*						*DSPP*	4	87,608,529	
Bone/dentin	*N/A*						*DMP1*	4	87,650,280	
Bone/dentin	*N/A*						*IBSP*	4	87,799,554	
Bone/dentin	*N/A*						*MEPE*	4	87,821,411	
Bone/dentin?	*scpp8*	10	15,106,808	0.38	8.5×10^{-1}	79.3	N/A			
Enamel	*ambn*	1	43,933,271	NE	-	0	*AMBN*	4	70,592,256	scpp6
Enamel	*enam*	1	43,905,263	NE	-	0	*ENAM*	4	70,628,744	enamelin
Enamel	*odam*	1	43,958,198	0.21	9.4×10^{-1}	0.7	*ODAM*	4	70,196,496	scpp2
Enamel	*N/A*						*AMTN*	4	70,518,569	amelotin
Enamel?	*scpp5*	1	43,910,360	3.24	2.4×10^{-17}	203.9	*N/A*			
Enamel?	*scpp7*	1	43,917,966	3.76	2.8×10^{-70}	32496.3	*N/A*			
Enamel?	*scpp9*	1	43,969,797	NE	-	0	*N/A*			
Ancestral	*sparcl1*	1	37,404,990	1.16	4.6×10^{-1}	8.9	*SPARCL1*	4	87,473,335	
Ancestral & Skeletal	*sparc*	14	25,938,895	2.90	4.2×10^{-17}	80211.9	*SPARC*	5	151,661	osteonectin
Other	*scpp1*	1	37,380,662	1.29	3.6×10^{-3}	469.0	*N/A*			

SCPP cluster genes associated with enamel formation are barely expressed in the RNA-seq dataset. Abbreviations: *N/A* not annotated, *NE* not expressed, *NRC* normalised read count; ?: suggestive expression based on genomic clustering in zebrafish. Table separates human SCPP 'bone/dentin' genomic cluster, SCPP 'enamel' genomic cluster, and human 'ancestral' SCPP genes. Green numbering represents passing arbitrary Log₂ fold change thresholds. Mean NRC is mean read count taken from the three ontogenetic and regenerating samples. AMELX and AMELY have been omitted from analysis as zebrafish do not have sex chromosomes

S1). Allowing less stringent interaction evidence scores (≥ 0.4) showed that *sp7* and *spp1* factors have connections with both the collagen and bone factor clusters, the latter showing connections to adjacent osteogenic clusters that for example contain *phospo1* and *ostn*, but also to *sparc, scpp5, scpp7* and *col10a1a* (Additional file 2: Fig. S5B). Together, the GO and PPI analyses indicated that DEGs were enriched for biological pathways involved in bone formation and collagen-matrix synthesis, consistent with increased osteoblast activity in a re-generating tissue.
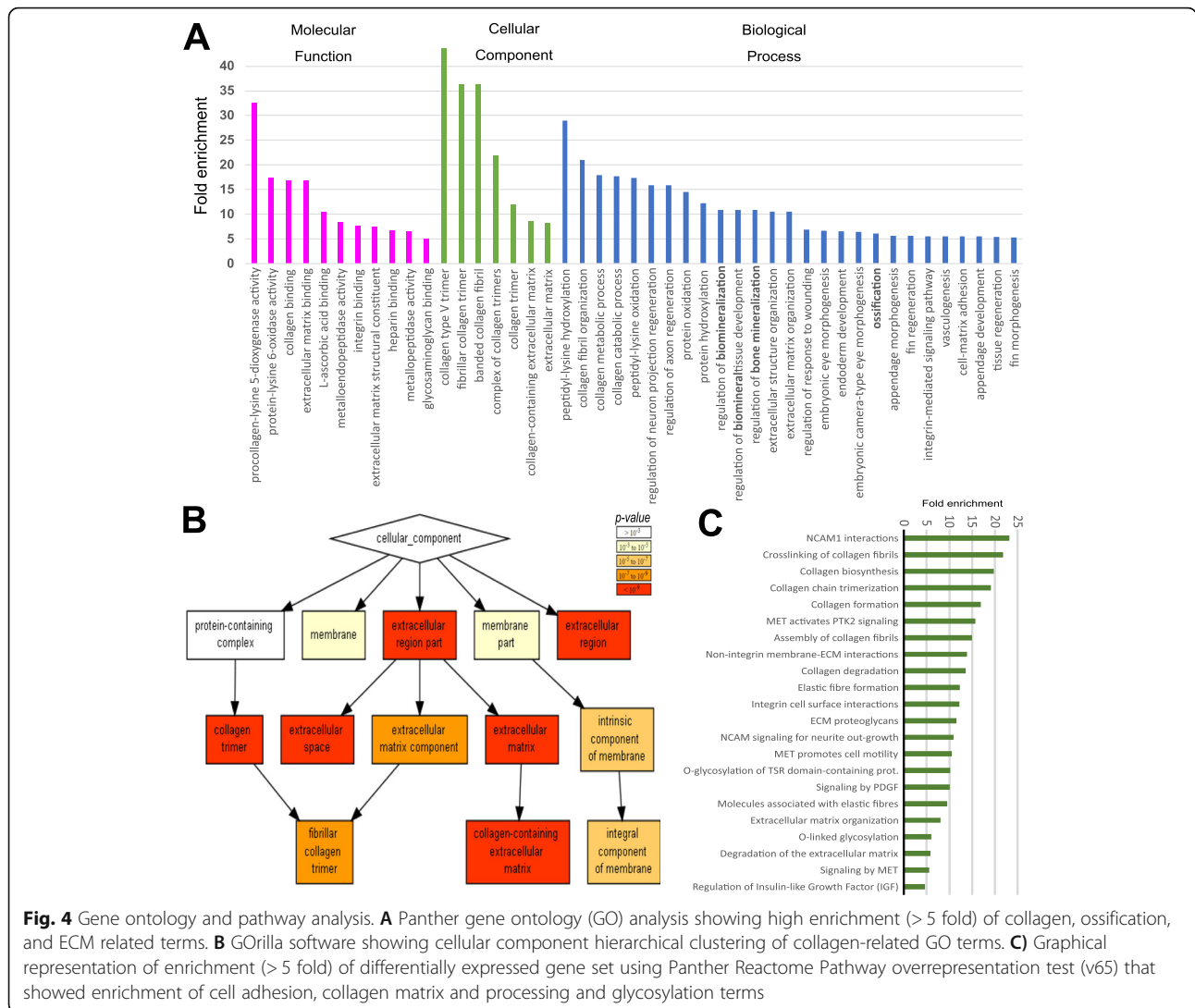
We next verified RNA expression of *col1a2,* osteoblast deposited collagen *col10a1a*, the collagen nucleation factor proteoglycan *bgna*, and hedgehog signalling ligand *ihha* that represent some of these clusters identified above. The expression of the gap junction gene *cx43* was also assessed as it regulates fin regeneration growth [41], and within our PPI network, it was connected to the IGF cluster known to modulate bone growth. These amplicons were assessed in regenerating scales from an independent experiment, and all showed similar trends between RNA-seq and qRT-PCR, validating these findings (Fig. 6A). Note, a subset of genes showed similar qRT-PCR profiles between the RNA-seq total RNA and the independent harvest total RNA samples (Additional file 2: Fig. S6). Ontogenetic scales showed weak *col1a1a* expression, predominantly located in the epidermis adjacent to *sp7* positive cells at the posterior edge of the scale (Fig. 6B). In regenerating scales, *col1a1a* promoter

activity was elevated along with expanded *sp7* expression covering most of the scale plate. We observed elevated activity at newly forming circuli associated with thickening of the calcified ECM (Figs. 2B and 6B). We performed the same procedure in *col10a1a:Citrine* and *col2a1a:mCherry* double transgenic zebrafish where the *col2a1a* reporter functions as a negative control as it was not a DEG and is associated with cartilage formation. This showed that ontogenetic reporter expression of both transgenes was absent. At 9 dph, only *col10a1a* reporter expression was observed, specifically at the posterior edge of the scale and interestingly also at the newly developing circuli (thicker ECM) located at the anterior region of the scale (Fig. 6C). These findings confirm that the RNA-seq dataset of regenerating scales contains networks with osteo-active factors that actively assemble (nucleate) a calcified collagen-rich matrix.

Differentially expressed genes are enriched for human orthologues that cause monogenetic skeletal disorders and that associate with polygenetic skeletal traits

We hypothesised that DEGs were likely to be enriched for human orthologues involved in skeletal disorders involving abnormal bone formation. Hypergeometric tests involving the ISDS Nosology and Classification of Skeletal Disorders database [42] revealed that human orthologues of DEGs were 2.8 times more likely to cause any monogenic skeletal dysplasia in humans than expected by chance ($P = 8 \times 10^{-11}$) (Fig. 7A). Orthologues of 47

Bergen *et al. BMC Biology*   (2022) 20:21

Page 9 of 25

Fig. 4 Gene ontology and pathway analysis. **A** Panther gene ontology (GO) analysis showing high enrichment (> 5 fold) of collagen, ossification, and ECM related terms. **B** GOrilla software showing cellular component hierarchical clustering of collagen-related GO terms. **C)** Graphical representation of enrichment (> 5 fold) of differentially expressed gene set using Panther Reactome Pathway overrepresentation test (v65) that showed enrichment of cell adhesion, collagen matrix and processing and glycosylation terms
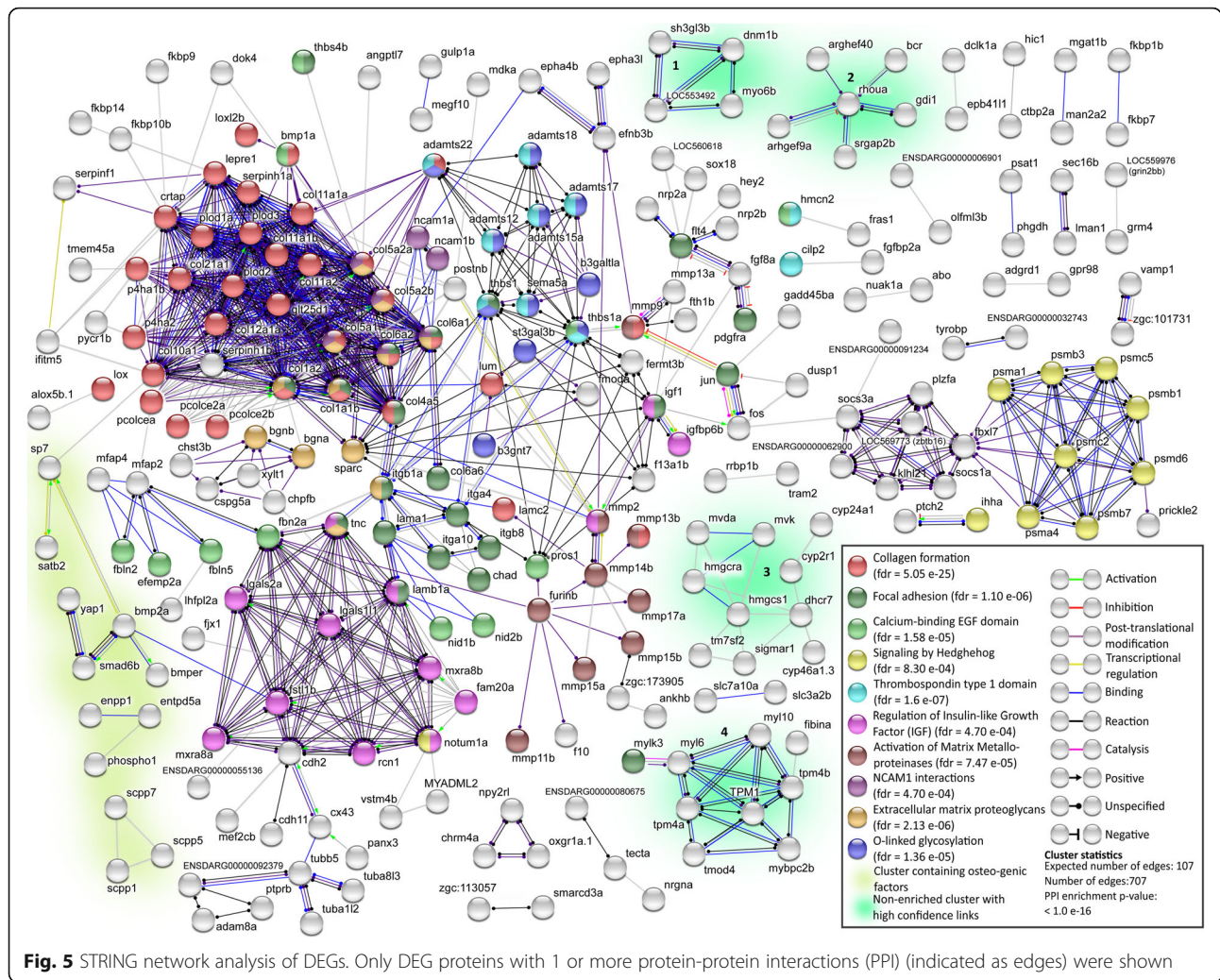
DEGs resulted in one or more primary bone dysplasias in humans when mutated (Additional file 7). Subgroup analysis revealed that DEGs were most strongly enriched for human genes involved in disorders characterised by low bone mass (i.e. Osteogenesis Imperfecta, $P=8.9 \times 10^{-10}$), and abnormal bone mineralisation ($P=1.6 \times 10^{-3}$). Compelling enrichment for 'collagen type 11' and 'metaphyseal dysplasia' groups was also detected; however, these observations did not meet our conservative *Bonferroni* corrected significance threshold ($P=2 \times 10^{-3}$) (Fig. 7B and Additional file 7).

MAGMA competitive gene set analysis (GSA) was used to investigate the relationship between genetic variation surrounding human orthologues of 482 mappable DEGs and estimated heel bone mineral density (eBMD) measured in 448,010 white European adults (Fig. 7A). Human orthologues of DEGs were more strongly associated with eBMD than all other human protein coding genes in the genome, suggesting that DEGs were enriched for eBMD associated genes ($P=1.6 \times 10^{-5}$, Table 3, Additional file 8). In a sensitivity analysis, we adjusted for potential confounders, including the set of orthologues that could not be mapped between fish and humans, and for the set of orthologues that was not expressed in ontogenetic and/or regenerating scales. Adjustment reduced the magnitude of enrichment by ~ 18%; however, the set of DEGs remained enriched for BMD associated orthologues ($P=4.4 \times 10^{-4}$). Post hoc permutation analysis suggested that enrichment was attributable to most, but not all DEGs. To further investigate this, DEGs were stratified according to whether they were up- or downregulated, and each set was re-analysed. Enrichment for eBMD associated orthologues was stronger for the 423 upregulated DEGs ($P=1.4 \times 10^{-6}$), whereas the 59 downregulated DEGs did not appear to be enriched (P=0.75) (Additional file 2: Fig.

Fig. 5 STRING network analysis of DEGs. Only DEG proteins with 1 or more protein-protein interactions (PPI) (indicated as edges) were shown

S7A). We repeated the analysis using height and OA and showed that DEGs were enriched for height associated genes ($P=6\times 10^{-4}$, Fig. 7A, Table 3 and Additional file 2: Fig. S7B). However, post hoc analysis revealed that enrichment was attributable to a small subset of DEGs, rather than all DEGs. Stratified analysis suggested that upregulated DEGs were more strongly enriched for height associated orthologues, whereas no enrichment was observed for downregulated DEGs (Additional file 2: Fig. S7B and Table 3). We observed some evidence of enrichment for OA associated genes ($P=0.02$); however, this observation did not meet our conservative *Bonferroni* corrected significance threshold.

Scale regeneration defects of *spp1* and *col11a2* mutants are consistent with results from human genetic association studies
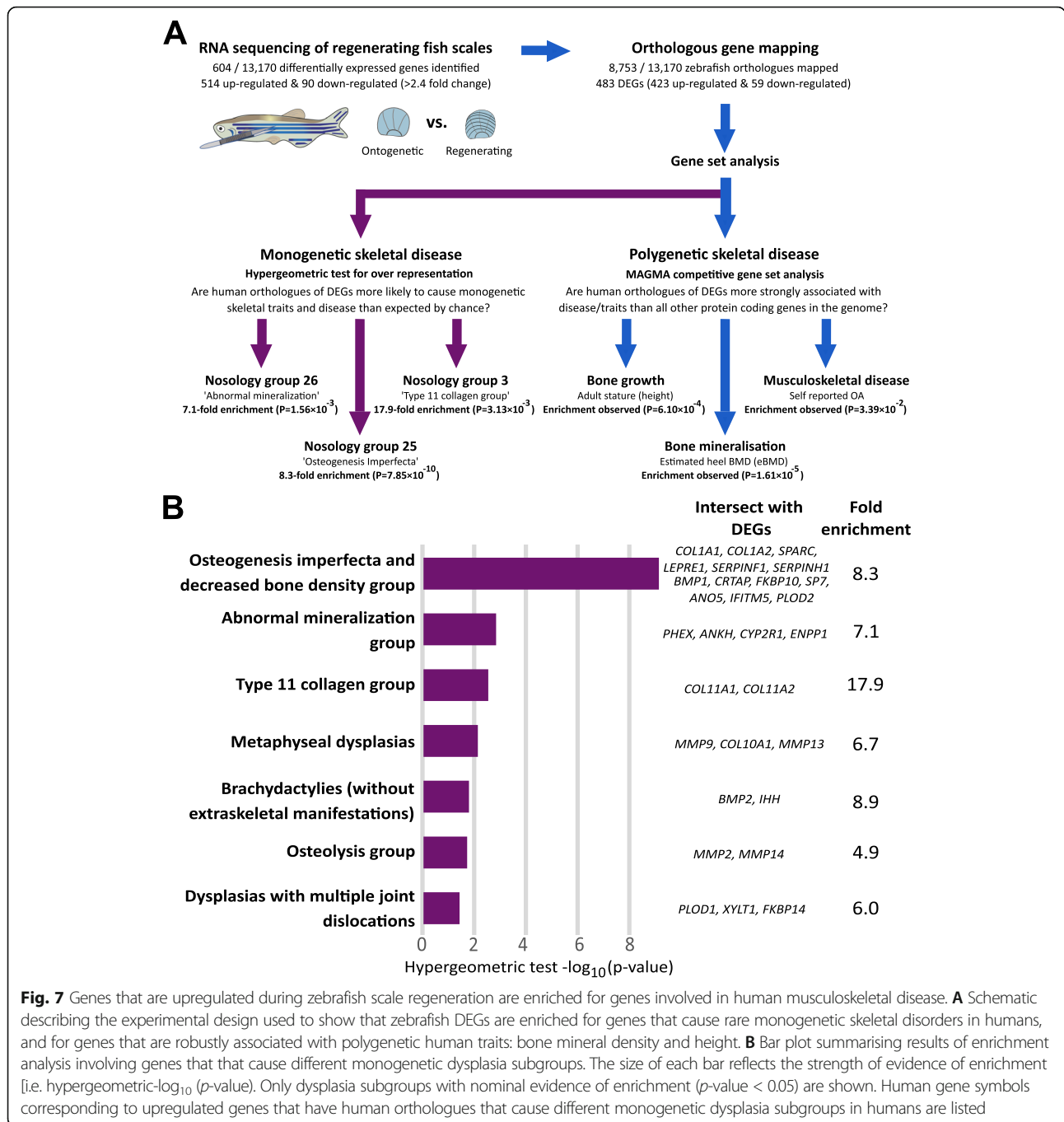
Our analysis showed that DEGs were enriched for human orthologues that cause monogenetic diseases and are associated with polygenetic traits involving bone growth and mineralisation. We therefore hypothesised that DEGs were also likely to be involved in bone formation during scale regeneration, and endoskeletal development in zebrafish. To investigate this, we selected *spp1* and *col11a2* that were robustly associated in our human genetic studies and validated their predicted role in the zebrafish. We hypothesised that *spp1* was most likely involved in biomineralisation, as *SPP1* was more strongly associated with eBMD ($P=6\times 10^{-20}$), as compared with height ($P=5\times 10^{-3}$) or OA ($P=0.15$). In contrast, we hypothesised that *col11a2* was involved in bone growth, size and shape, as *COL11A2* was robustly associated with height ($P=6\times 10^{-24}$) and OA ($P=1.9\times 10^{-6}$), but not with eBMD (i.e. $P=0.02$). In fact, the human orthologue of *col11a2* also cause type III Stickler syndrome when mutated, and this disorder is characterised by early-onset OA [43].

We first assessed the expression of the two DEGs during scale regeneration. An independent experiment validated that both *spp1* and *col11a2* were upregulated during scale regeneration (Additional file 2: Fig. S8A). We had access
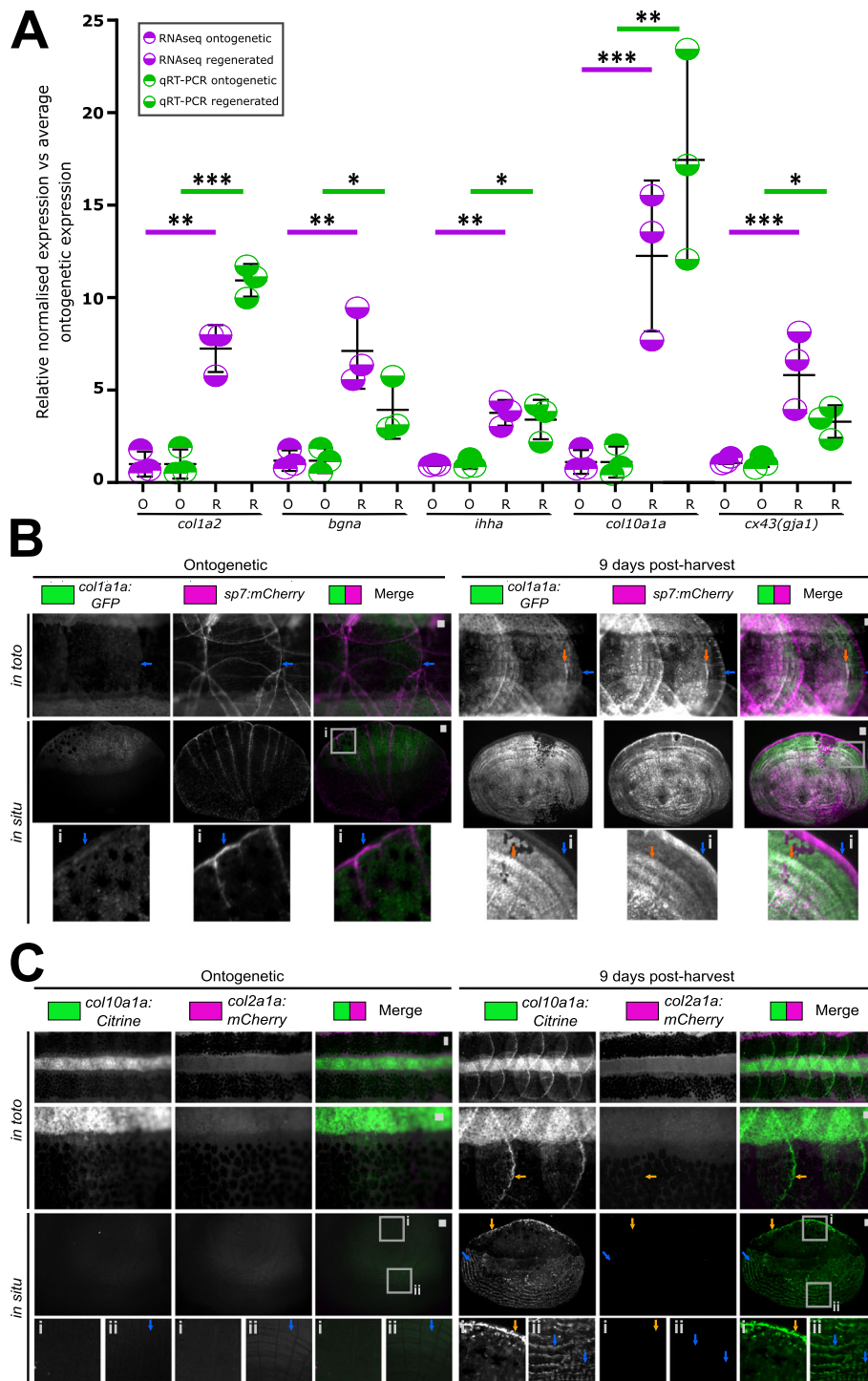
Fig. 6 Expression validation of gene ontology findings. **A** Relative expression levels of qRT-PCR (unpaired *t* test) and transcriptomic analysis of amplicons (false discovery rate) found in the collagen, proteoglycan and hedgehog signalling networks. **B** Images of *in toto* and harvested scales (*in situ*) of *col1a1a:GFP* and *sp7:mCherry-NTR* double transgenic fish (*n* = 4 each condition). Blue arrow indicates the posterior (distal) fringe with high expression of mCherry. Orange arrow points at co-expression of GFP and mCherry. **C** *In toto* and *in situ* stereomicroscope images of *col10a1a:Citrine* and *col2a1a:mCherry* double transgenic ontogenetic and regenerating scales (*n* = 4 fish each condition). Orange arrows indicate Citrine signal at the posterior distal edge (inset i) while blue arrow points at a newly forming lateral circulus (inset ii) of the scale. Scale bar: 100 μm

Fig. 7 Genes that are upregulated during zebrafish scale regeneration are enriched for genes involved in human musculoskeletal disease. **A** Schematic describing the experimental design used to show that zebrafish DEGs are enriched for genes that cause rare monogenetic skeletal disorders in humans, and for genes that are robustly associated with polygenetic human traits: bone mineral density and height. **B** Bar plot summarising results of enrichment analysis involving genes that that cause different monogenetic dysplasia subgroups. The size of each bar reflects the strength of evidence of enrichment [i.e. hypergeometric-$\log_{10}$ (*p*-value). Only dysplasia subgroups with nominal evidence of enrichment (*p*-value < 0.05) are shown. Human gene symbols corresponding to upregulated genes that have human orthologues that cause different monogenetic dysplasia subgroups in humans are listed

to a transgenic reporter of *spp1* (*spp1:mCherry*) and this showed that *spp1* was exclusively localised at the distal rim of the ontogenetic scale, similar to *sp7:GFP*-positive osteoblast localisation marking the leading edge (Fig. 8A). These *sp7*-positive (sub)marginal cells at the rim are classed as more mesenchymal osteoblasts involved in *de novo* bone formation in ontogenetic scales [44]. In regenerating scales, *sp7* expression was seen more broadly but elevated at the posterior edge; followed proximally by an increased *spp1* signal with reduced *sp7* expression in the

same area (Fig. 8A). The extent of *spp1* expression was increased compared to ontogenetic scales, implying an enhanced response to bone growth and mineralisation in the regenerating scale (Fig. 8B, C).

We stained regenerating scales of *col11a2*[Y228X] and *spp1*[P160X] mutant fish with calcein green. *In toto* images showed that *spp1* mutant regenerating scales bind less calcein compared to wildtype and *col11a2* mutant fish indicative of reduced mineralised matrix formation (Fig. 9A). Quantification of the calcein intensity revealed that both

Table 3 Results of MAGMA competitive gene set analysis involving human polygenetic traits and disease

Model	Trait/Disease	#Genes	BETA	BETA STD	SE	P-value
ALL DEGs						
Base	eBMD	482	0.33	0.05	0.079	1.61×10^{-5}
	Height	482	0.31	0.05	0.096	6.09×10^{-4}
	Osteoarthritis	465	0.09	0.01	0.047	3.39×10^{-2}
Sensitivity	eBMD	482	0.27	0.04	0.080	4.43×10^{-4}
	Height	482	0.21	0.03	0.097	1.70×10^{-2}
	Osteoarthritis	465	0.06	0.01	0.047	1.01×10^{-1}
Upregulated DEGs						
Base	eBMD	423	0.40	0.06	0.085	1.38×10^{-6}
	Height	423	0.36	0.05	0.103	2.06×10^{-4}
	Osteoarthritis	408	0.09	0.01	0.050	4.40×10^{-2}
Sensitivity	eBMD	423	0.33	0.05	0.085	4.52×10^{-5}
	Height	423	0.26	0.04	0.103	6.41×10^{-3}
	Osteoarthritis	408	0.06	0.01	0.051	1.18×10^{-1}
Downregulated DEGs						
Base	eBMD	59	-0.15	-0.01	0.223	7.48×10^{-1}
	Height	59	-0.05	0.00	0.268	5.74×10^{-1}
	Osteoarthritis	57	0.08	0.00	0.128	2.58×10^{-1}
Sensitivity	eBMD	59	-0.22	-0.01	0.223	8.36×10^{-1}
	Height	59	-0.16	-0.01	0.268	7.19×10^{-1}
	Osteoarthritis	57	0.06	0.00	0.128	3.25×10^{-1}

Model: Statistical models used, with (sensitivity) or without (base) adjustment for the set of genes that could be mapped between zebrafish and humans. Trait: the phenotype or disease trait investigated. #GENES: number of human orthologues of zebrafish DEGs in the gene set. BETA: regression coefficient. BETA_STD: the semi-standardised regression coefficient, corresponding to the predicted change in Z value given a change of one standard deviation in the predictor gene set/ gene covariate. SE standard error of BETA. P-value: Strength of evidence to reject the null hypothesis that on average, human orthologs of zebrafish DEGs genes exhibit stronger associations with the corresponding human polygenetic trait / disease than all other protein coding genes in the genome. Green numbers highlight gene sets that have nominal evidence of enrichment (i.e. $p < 0.05$). DEG differentially expressed genes. Full details can be found in the 'read me' tab of Additional file 8
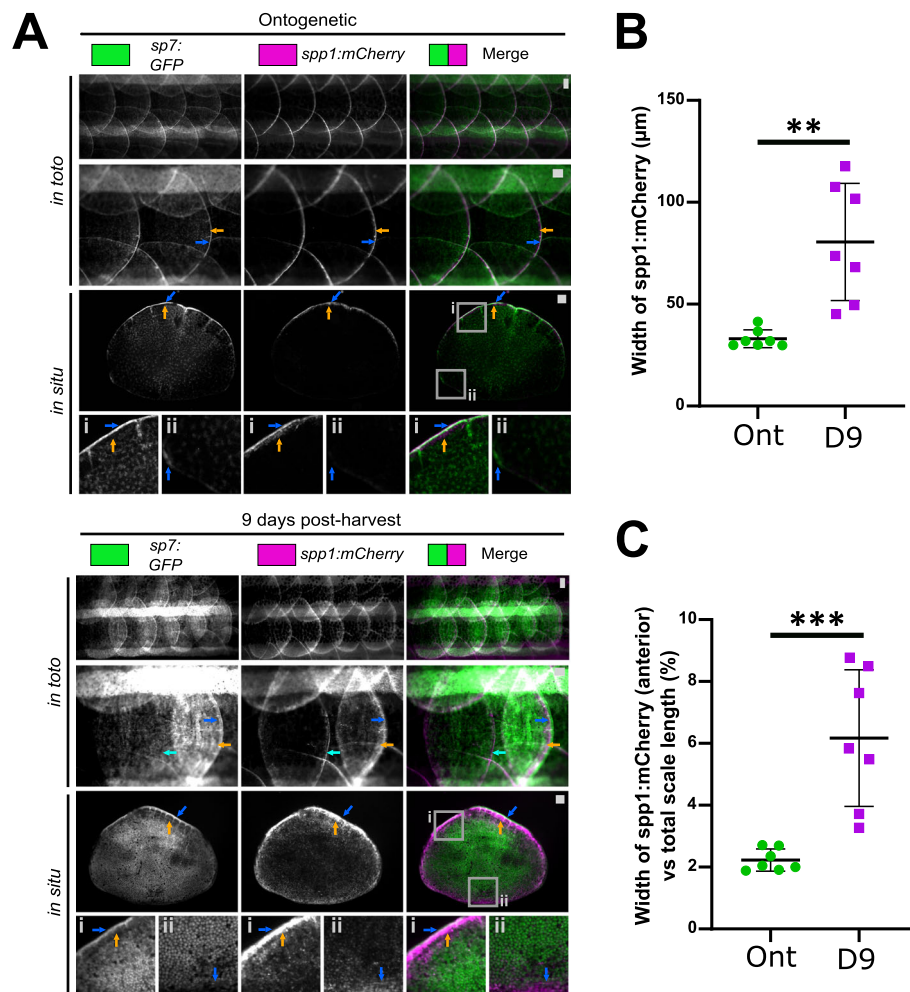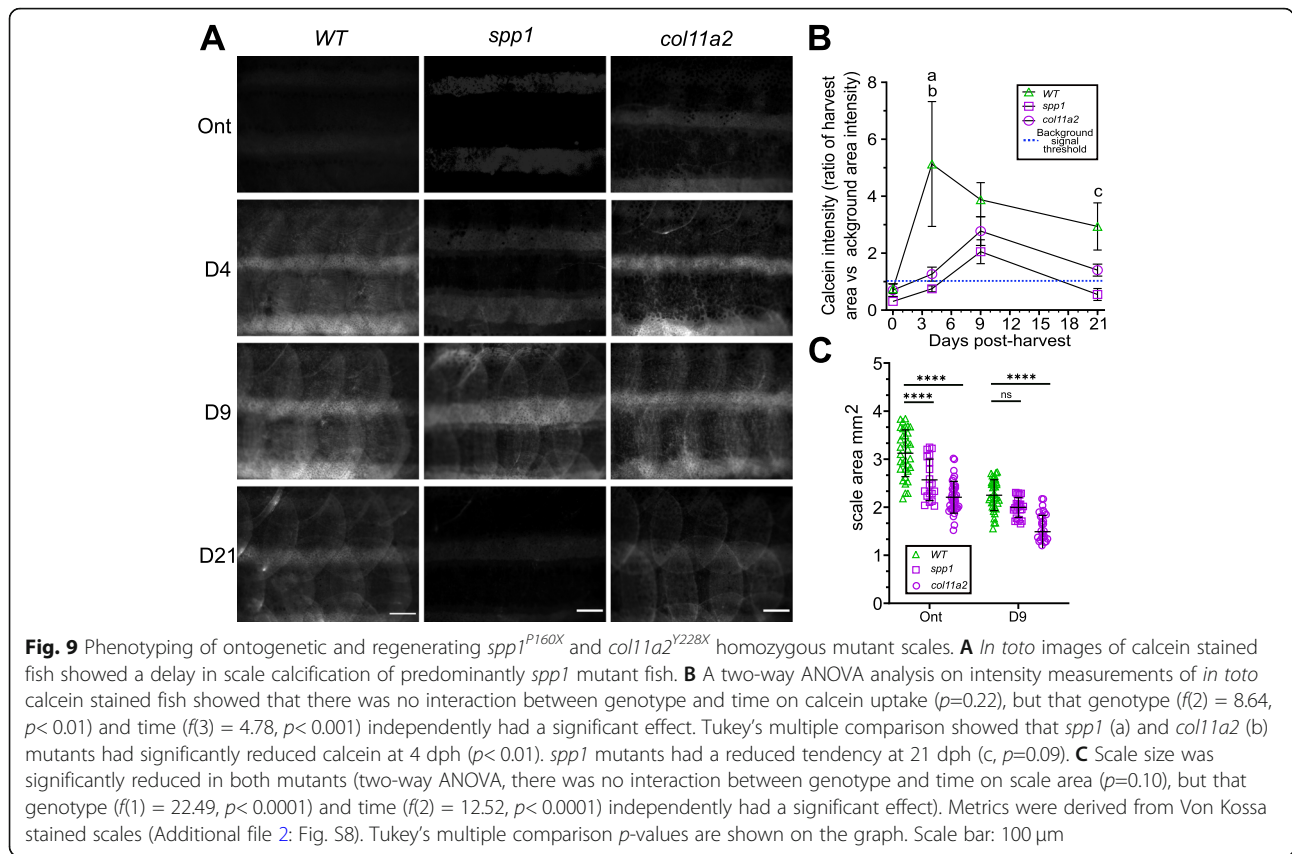
Fig. 8 Validation of genes associated with polygenic traits in humans. **A** Images of scales from *in toto* and harvested scales of *sp7:GFP* and *spp1:mCherry* double transgenic fish (ontogenetic: *n* = 5, regenerating: *n* = 4 fish). For overview and inset i, blue arrows indicate high GFP signal and orange the mCherry signal at the posterior edge (inset i) while light blue arrow indicates an ontogenetic scale next to a regenerating scale. Inset ii shows anterior region of the scale with opposite GFP and mCherry signals (blue arrow) in ontogenetic and regenerating conditions. **B** Quantification of width of the mCherry signal at the posterior edge from harvested scales (between blue and orange arrows in panel **C**). **C** Width of mCherry signal normalised by scale length (anterior to posterior). Unpaired two-tailed student t-test p-values are shown on the graph. Scale bar: 100 μm

spp1 and *col11a2* mutants showed a delay and a lower rate of calcein uptake (Fig. 9B). Von Kossa staining on regenerating scales confirmed that *spp1* mutants have mineralisation defects in both ontogenetic and regenerating scales (Additional file 2: Fig. S8B). The *col11a2* mutant scales exhibited uneven mineralisation in the (hypermineralised) epidermal area by displaying areas of hypomineralisation whilst the rim showed increased mineralisation (Additional file 2: Fig. S8B). Both *coll11a2* and *spp1* mutant ontogenetic scales were smaller than wildtype (Fig. 9C). However, while *col11a2* mutants showed a reduced body length and *spp1* mutants were normal length; therefore, when corrected for body size, *spp1* mutants had relatively smaller scales (Additional file 2: Fig. S8C and D). During regeneration, *col11a2* showed growth defects whereas *spp1* mutant scales showed a similar rate of growth at 9

dph (Fig. 9C). These data show that both mutants show exoskeletal phenotypes that are largely consistent with their GSA predictions.

Endoskeletal phenotypes of *spp1* and *col11a2* mutants are largely consistent with results from human genetic association studies

As the exoskeletal collagen template growth process in the form of scales was one of the first biomineralisation processes in evolution to occur, we hypothesised to translate the contrasting patterns of association and scale regenerating defects to endoskeletal sites. Previous adult murine mutants of *Spp1* and *Col11a2* have shown complex (and sometimes conflicting) endoskeletal mineralisation defects and cartilage OA phenotypes respectively [45–48]. However, the adult zebrafish *spp1* and

Fig. 9 Phenotyping of ontogenetic and regenerating *spp1*[P160X] and *col11a2*[Y228X] homozygous mutant scales. **A** *In toto* images of calcein stained fish showed a delay in scale calcification of predominantly *spp1* mutant fish. **B** A two-way ANOVA analysis on intensity measurements of *in toto* calcein stained fish showed that there was no interaction between genotype and time on calcein uptake (*p*=0.22), but that genotype ($f(2)$ = 8.64, *p*< 0.01) and time ($f(3)$ = 4.78, *p*< 0.001) independently had a significant effect. Tukey's multiple comparison showed that *spp1* (a) and *col11a2* (b) mutants had significantly reduced calcein at 4 dph (*p*< 0.01). *spp1* mutants had a reduced tendency at 21 dph (c, *p*=0.09). **C** Scale size was significantly reduced in both mutants (two-way ANOVA, there was no interaction between genotype and time on scale area (*p*=0.10), but that genotype ($f(1)$ = 22.49, *p*< 0.0001) and time ($f(2)$ = 12.52, *p*< 0.0001) independently had a significant effect). Metrics were derived from Von Kossa stained scales (Additional file 2: Fig. S8). Tukey's multiple comparison *p*-values are shown on the graph. Scale bar: 100 μm

col11a2 mutant endoskeletal phenotypes have not been described. To consolidate evolutionary conservation of both between aquatic and terrestrial species and between exoskeletal and endoskeletal growth processes we chose to analyse 3D micro-CT renders of both zebrafish mutants. As a proxy for skeletal growth, we evaluated the length of the fish from these micro-CT renders which showed that *col11a2* but not *spp1* mutants were shorter than wildtype fish (Additional file 2: Fig. S9A). Next, we segmented the lower jaw (mandibular arch) and caudal vertebrae of these 3D micro-CT renders and measured several histomorphological parameters using element landmarks as set out in Additional file 2: Fig. S9B-C. These measurements of the lower jaw demonstrated that *col11a2* fish have reduced width and length in the lower jaw whereas *spp1* fish did not show significant changes (Fig. 10A). When we calculated the ratio between length and width, both mutant lines showed a mild alteration in lower jaw element proportion (Fig. 10B). Interestingly, in contrast to wildtype and *spp1* mutants, the *col11a2* mutants exhibited altered joint shape consistent with the predicted OA phenotype in adulthood and as seen in larval jaws previously (Fig. 10C) [49].

We observed a physical thickening of the vertebral arches and anterior shaft of the vertebral centrum in *spp1* mutants (Fig. 10D). However, histomorphological measurements of the centrum, neural arch, and haemal arch of the vertebra did not show significant changes in either mutant line (Additional file 2: Fig. S9D-K). To test whether *spp1* mutants have a mineralisation phenotype, we quantified mean volumetric estimated tissue mineral density TMD (a proxy for BMD) from sites formed via different modes of ossification: the parietal plate of the cranium and the first caudal vertebra as these are formed by intramembranous ossification and by ossification directed by the notochord sheath, respectively. We observed an increase in TMD for *spp1* mutants in the skull, and an increased tendency in vertebrae TMD was seen for both *col11a2* and *spp1* mutants (Fig. 10E). Taken together, our data suggest that genes that are differentially expressed during scale regeneration play a role in wider regulation of skeletal homeostasis at other skeletal sites in the zebrafish, and that mutants in these genes have phenotypes that show consistency with the results from human genetic association studies.

Discussion

In this study we define, for the first time, the transcriptome of ontogenetic and regenerating zebrafish scales and show that DEGs are enriched for biological pathways involved in osteoblast-mediated bone formation, many of which are conserved in humans. Although fish
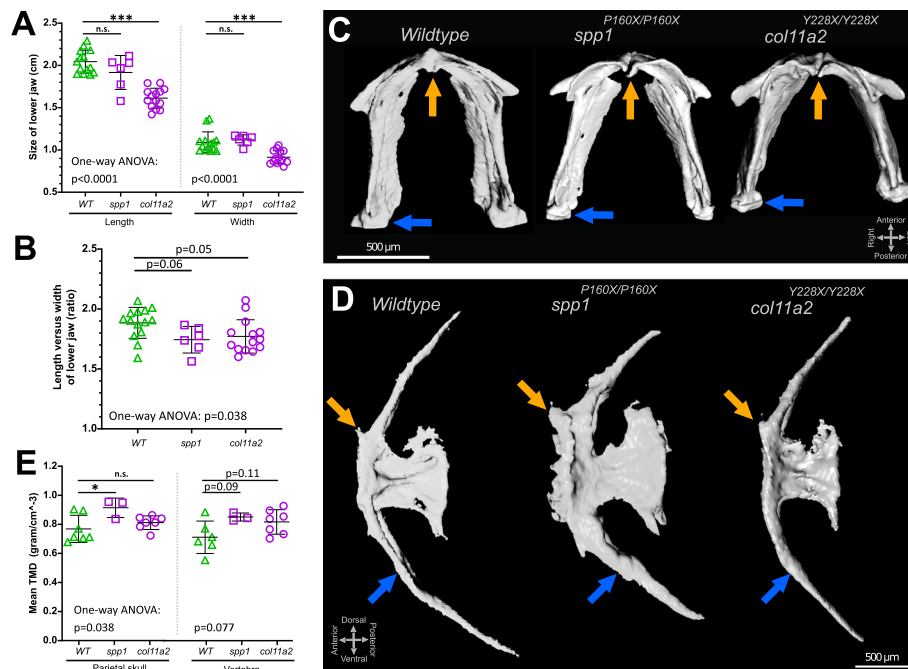
Fig. 10 Histomorphology and tissue mineral density measurements on 3D micro-CT images of *spp1*^{P160X} and *col11a2*^{Y228X} homozygous mutants revealed altered bone structures. **A** Lower jaw size was reduced in *col11a2* mutants (one-way ANOVA: length: ($f(2) = 33.38$, $p < 0.0001$), width: ($f(2) = 15.06$, $p < 0.0001$)). **B** A one-way ANOVA ($f(2) = 3.636$, $p < 0.05$) analysis indicated that lower jaw element proportions showed an altered tendency in both mutants. **C** Ventral view of the segmented lower jaw images. Orange arrow indicates anterior mandibular arch joint and blue arrow shows mandibular arch—palatoquadrate (not visible) joint. **D** Lateral view of segmented images of the first caudal vertebra showing the anterior facet (orange arrow) and haemal arch (blue arrow). **E** Mean tissue mineral density (TMD) calculations and one-way ANOVA analysis of dermal parietal skull ($f(2) = 4.144$, $p < 0.05$) and notochord sheath derived vertebral ($p = 0.077$) bone. Tukey's multiple comparison *p*-values are shown on the graphs

scales and endoskeletal bone are evolutionary distantly related, based on our findings we conclude that their tissue morphogenesis shares a similar expression profile harbouring osteoanabolic pathways. These matrix-building and calcifying cells (in the literature also referred to as scleroblasts, elasmoblasts, or ameloblasts) that decorate the scale therefore resemble types of osteoblasts [15]. Indeed, recent publications have acknowledged these cells as osteoblast or osteoblast-like cells [13, 19, 44]. By integrating the transcriptomic data with large-scale human genetic association studies, we found that DEGs were enriched for human orthologues that cause monogenetic skeletal diseases, and that associate with polygenetic skeletal disease traits. Our findings suggest that zebrafish scale regeneration has the potential to help us better understand biological function and pathways relevant to bone growth that is often dysregulated in human skeletal diseases.

Regenerating scales have an expression profile enriched in genes encoding matrix growth and mineralisation

Transcriptomic profiling revealed that regenerating scales upregulate gene expression of many osteoblast genes. The high expression of genes involved in ECM deposition, principally genes regulating collagen synthesis, processing, and deposition fits with the regeneration of the collagen-rich matrix of the scale. We demonstrated that while *col1a1a* is expressed throughout the regenerating scale it is largely excluded from the leading edge labelled by *sp7*. By contrast, *col10a1a* is strongly expressed in scale regions associated with thickening of the calcified matrix in accordance with the role of type X collagen as an early marker of ossifying tissues in the zebrafish [50–52]. Amongst DEG collagens, *col11a2* was one of the highest DEGs in this profile. Type XI collagen is more frequently associated with cartilage matrix formation; interacting with type II collagen to regulate spacing and nucleation of fibrils in the ECM [53, 54]. As well as the collagens themselves, matrix processing genes coding for MMPs were also upregulated. MMPs breakdown the dense type I collagen matrix, which is crucial for tissue homeostasis, release of signalling molecules, and cell migration [55]. We have previously shown that *mmp9* is upregulated during scale regeneration and *mmp9* expression was observed adjacent to newly deposited matrix and TRAP-positive cells [26].

Moreover, we observed transcriptional upregulation of ECM proteoglycans that bind collagens, such as biglycan *(bgna/b)*, which are important in regulating collagen nucleation and remodelling [56]. Biglycan has been shown to modulate ECM accessibility, and in regulating bone strength in mice [57, 58]. In addition to collagen related

factors, cell adhesion genes, in particular integrin signalling, were enriched and connected to the collagen and MMP networks. Cell adhesion and the ECM are interlinked and both must be tightly regulated during rapid growth [40]. Cell adhesions are heavily modulated to allow cell rearrangements and tissue mechanics in regenerative tissue expansion contexts. For example, in a mammalian epidermal wound healing response, it is crucial for epithelial regeneration [59].

Pathway analysis revealed signalling pathways related to ECM formation and expansion, such as integrin, HH and IGF. The absence of enrichment (note these were also not under-represented) of Wnt, Bmp, and Fgf pathway genes is likely to be associated with the timing of the transcriptomic profile (9 dph). For example, it has been shown that Wnt and Fgf signalling are required for the initiation of scale regeneration, with inhibiting Fgf or canonical Wnt signalling leading to a failure to form scales [19]. HH signalling controls morphogenesis of the scale over a prolonged time period, where the HH ligand secreted from the epidermis regulates the rate of ECM deposition by scale osteoblasts [19, 44].

Our study shows that scales expressed *scpp* repertoire genes that are involved in dentin and endoskeletal bone formation (*spp1* and *sparc*), while *scpp* repertoire and tooth regeneration genes that are associated with enamel matrix hypermineralisation and tooth formation were barely or not differentially expressed [37]. We identified three upregulated teleost-specific *scpp* repertoire genes, *scpp1*, *scpp5* and *scpp7*. Previous reports indicated that *scpp1/scpp5* expression is not restricted to exoskeletal elements; expression was also seen the in jaw and dental tissues, with *scpp1* also expressed in osteoblasts and osteocytes [60]. Whilst the *spp1* genomic region lacks terrestrial SCPP 'bone/dentin' repertoire *MEPE*, *DSPP*, *SMP1*, and *IBSP* orthologues, we showed that *spp1*, *sparc* and teleost specific *scpp1*, *scpp5* and *scpp7* peaked their temporal expression around 9 dph, coinciding with *col1a1a* transcriptional peaking. This indicates that there is a conserved regulation of expression of SCPP repertoire genes during bone growth and that scales lowly express enamel-related genes.

Osteoblastic evolutionary conserved gene sets involved in human endoskeletal bone formation are expressed during scale regeneration

The pathways identified are largely conserved in humans, prompting us to investigate whether DEGs were enriched for human orthologues that cause human monogenic skeletal disorders, and that associated with related polygenetic skeletal disease endophenotypes. Hypergeometric tests involving human monogenetic skeletal disorders complemented findings from our gene ontology analysis, suggesting that DEGs were most

strongly enriched for human orthologues that skeletal disorders that were broadly characterised by abnormal or decreased bone mineralisation. Suggestive evidence of enrichment for orthologues involved in abnormal cartilage formation and bone growth was also observed. Further analysis involving corresponding human polygenetic skeletal disease traits provided additional evidence that upregulated (but not downregulated) DEGs, were highly enriched for human orthologues associated with eBMD as captured by heel bone ultrasound. Enrichment for height associated genes was also observed for upregulated DEGS; however, post hoc permutation analysis revealed that enrichment was not attributable to all upregulated DEGs, but rather a nested subset of genes that may correspond to one or more biological pathway(s) involved in human bone growth. One possible explanation for the heterogeneity is that DEGs were identified using bulk RNA-sequencing, a method that captures a heterogenous mixture of transcriptome signature profiles that define different cell types, states, and their biological pathways, some of which may or may not be involved in bone growth during scale regeneration. Based on the collective analyses described above, we show that the zebrafish scale harbours *bona fide* osteoblast cell populations. Future studies may therefore benefit by focusing on transcriptomic profiling of separate cell populations from scales at different stages of regeneration to dissect the different pathways responsible for osteo-anabolic growth.

Detailed skeletal phenotyping on zebrafish mutants revealed that *spp1* fish had increased BMD, whereas *col11a2* mutants and developed a lower jaw phenotype that was consistent with OA. These observations were largely consistent with our genetic association analysis that showed that *SPP1* was robustly associated with eBMD, but not height or OA, and *COL11A2* was robustly associated with height and OA, but not with eBMD. We do however note that *col11a2* mutants had elevated vertebral TMD which was not consistent with results from out genetic association studies. Notably, in humans, mutations in *COL11A2* lead to Stickler syndrome, a condition associated with craniofacial dysplasias, and joint abnormalities that lead to premature OA [43] whilst in mutant mice the cartilage degradation phenotype appears to be milder [45]. Note that COL11A2 heterotrimer partner *COL11A1* is associated with both eBMD and OA traits [61, 62] and causes Stickler syndrome with fragility fractures [63]. The potentially type XI chain α1 and α2 heterotrimer complex disruption could provide an explanation for the uneven mineralisation of the scales and elevated tendency of endoskeletal BMD phenotypes in the *col11a2* mutants. A recent study identified the *col11a2* promotor to show enhanced activity in the regenerating caudal fin bony

segments, implying a more global role in (dermal) bone growth besides its known function in cartilage [64].

No clear human phenotype for *SPP1* has been defined yet, but as a *SCPP* repertoire gene, it plays a complex intertwined role in regulating post-inflammatory tissue repair and regeneration of multiple organ systems in vertebrates. We show that *spp1:mCherry* was up-regulated in the regenerating scale adjacent to pre-osteoblastic sp7+ cells which is consistent with previous reports that expression of *spp1* is up-regulated during bone remodelling in both zebrafish fins and mammals [65–67]. Moreover, *spp1* mutants have elevated TMD in the cranial and axial endoskeleton, in line with murine studies where loss-of *Spp1* leads to enhanced mineral content in some parts of trabecular bone [47] and increased cortical thickness [68]. Mouse studies have shown that *Spp1* determines mineralisation rate in the skeleton through regulating *Phospho1* expression [46, 69]. For example, Osteopontin expression is detected during cranial suture closure in both mice and zebrafish, suggesting a role in appropriate timing of pre-osteoblast differentiation as it is often expressed in similar regions positive for *sp7*, *col1a1a1* and *col10a1a* [66, 70, 71]. Outside of its function in skeletal tissue, mammalian *Spp1* is functionally diverse, as it promotes angiogenesis and its expression is also triggered during cutaneous wound healing to control the rate of repair [72, 73]. As Osteopontin is a component of the ECM and possesses integrin binding domains, it can bind various integrins to modulate cell adhesion to the collagen matrix important during i.e. tissue growth [74]. As we see high *spp1:mCherry* expression adjacent to *sp7+* pre-osteoblasts in the regenerating scale, it implies that *spp1* could play a role in differentiation of these pre-osteoblasts and therefore timing of scale mineralisation and the association with human eBMD is suggestive of a conserved function in control of BMD across species. Indeed, spp1 mutants have mineralisation defects in both ontogenetic and regenerating scales. However, endoskeletal TMD appears to be increased, implying that *spp1* plays an additional role in promoting osteoclastic bone resorption in line with its role in regulating the inflammation response in for example wound healing.

Scales offer an alternative model to discover novel evolutionarily conserved osteoanabolic genes

The discovery of the WNT pathway inhibitor Sclerostin (SOST) as a high bone mass gene demonstrated the translation health sciences value of understanding the genetic control of bone growth and osteoblast differentiation processes [75]. The combination of human genetic linkage studies subsequent successful modelling in mice and cells showing a transient increase of *Sost* expression in osteoblasts differentiating into osteocytes [75, 76]. This prompted the development of a humanised monoclonal antibody (Romozumab) against SOST which is now an osteo-anabolic osteoporosis therapeutic that reduces fracture risk [77]. Interestingly, previous reports showed that zebrafish scales express osteocyte marker *sost* [24], and we indeed observed expression of *sost*, *wnt3a*, *lrp4*, *lrp5* and *dkk1a* 'osteocyte markers' in both ontogenetic and regenerating scale transcriptomes albeit not classed as differentially expressed. Future experiments *in situ* should confirm which cells express these osteocyte markers. While zebrafish scales do not have cells that display cytomorphological characteristics resembling osteocytes, the teleost evolutionary relative bigeye tuna (*Thunnus obesus*) does have thicker trabeculated osteocytic scales that display a lacunocanalicular network [78]. This indicates that it is likely that a subpopulation of the zebrafish scale osteoblasts already express 'osteocyte markers' as other previous studies also have shown that there is transcriptional heterogeneity among scale osteoblasts [13, 44].

The common vertebrate ancestors date back more than 400 million years ago allowing morphological and functional divergence between exoskeletal scales and terrestrial endoskeletal bones. In fact, despite this considerable evolutionary distance between teleosts and terrestrial vertebrates (tetrapods), our data not only indicate that the distinct gene expression profiles of scale cells are indeed osteoblast-like profiles, but also demonstrate a striking conservation of osteogenic pathways from teleosts to mammals. When these morphological and functional differences between elasmoid scales and mammalian endoskeletal bones are considered, scales offer an alternative avenue to study a subset of genes involved in bone growth. Since the prevalence of diseases with pronounced bone fragility phenotypes is increasing due to an ageing population there is a need to discover and rapidly test new bone growth candidate genes that could act as drug targets. These multi-factorial diseases have complex genetic and physiological underpinnings that are not well understood. The regenerating zebrafish scale could therefore function as an additional model to study and discover osteo-anabolic factors relevant to human skeletal diseases. The relative abundance of teleost scales (hundreds per fish) and their amenability for imaging make them an ideal model for regenerative studies, and unlike caudal fins, scales can be treated as independent bone units, each harbouring their native complex tissue environment retaining vital intercellular interactions (between osteoblasts and osteoclasts). Scales can be cultured *ex vivo* in a semi-high throughput multi-well format, offering a path to test compounds that could complement established tissue culture and *in vivo* pharmacology studies [32, 79].

Conclusions

Zebrafish regenerating scales express osteoanabolic gene sets that are evolutionarily conserved and despite the evolutionary distance between exoskeletal and endoskeletal elements, scales show transcriptional similarities with mammalian endoskeletal bone growth. They express human orthologues involved in skeletal disease making the scales an attractive model to identify novel players and their pathways in matrix formation and calcification. Moreover, our analyses provide strong evidence that zebrafish scales are decorated with *bona fide* osteoblast cell populations. We show that integrative analysis involving zebrafish transcriptomics and human genetic association studies is feasible, and that future studies involving zebrafish scales have potential to better our understanding of bone formation in general that could aid studies involving skeletal disease pathogenesis.

Methods

Zebrafish husbandry, mutant and transgenic lines

Zebrafish were maintained under normal husbandry conditions [80]. Wildtype AB/TL (University of Bristol, UK) and AB (Radboud University, Nijmegen, NL) strains were used. Mutant lines (AB/TL background, UoB) have been previously described; *col11a2*^sa18324^ carries a nonsense mutation causing a premature stop codon at tyrosine position 228 (ENSDART00000151138.3), henceforth called *col11a2*^Y228X^ [49] and *spp1*^CGAT327-330del^ carrying a deletion leading to a frameshift resulting in a premature stop codon nonsense mutation at proline position 160 (ENSDART00000101261.6), henceforth called *spp1*^P160X^ [81]. Transgenic lines were kept in an AB/TL background and full nomenclature as follows: *spp1:mCherry [TgBAC(spp1: mCherry)*^bsl374^ [81]], *col2a1a:mCherry [Tg(Col2a1aBAC: mCherry)* [50]], *sp7:GFP [Tg(Ola.sp7:NLS-GFP)* [82]], *sp7: mCherry [Tg(osterix:mCherry-NTRo)*^pd46^ [83]], *col1a1a: GFP [Tg(col1a1a:EGFP)*^zf195tg^ [84]], and *col10a1a:Citrine [TgBAC(col10a1a:Citrine)* [85]].

Elasmoid scale harvesting and imaging

Anaesthetised fish (0.05% (v/v) tricaine methanesulfonate (MS-222) (UoB), 0.1% (v/v) 2-phenoxyethanol (RU)) were put on a wet tissue containing system water and anaesthetic, and scales plucked under a microscope with a watchmaker's tweezers from the midline of the lateral flanks near the dorsal fin. Fluorescent microscope images of flanks (*in toto*) and single harvested scales (*in situ*) were acquired on a fluorescent stereomicroscope (Leica Microsystems, Germany), using ×2, ×4 and ×8 magnification.

In vivo calcein staining

Adult fish were immersed in 40 µM Calcein (Sigma-Aldrich, cat# 154071-48-4) Danieau's buffer solution (pH 7.4) for two hours and washed in system water for at least 15 minutes prior to imaging. Calcein intensity was measured from *in toto* images of ontogenetic scales and regenerating scales (4, 9. 21 dph). For each image, pixel intensity was measured in ImageJ at locations where two adjacent scales overlap with each other for both non-harvest area (background) and harvest area (4 regions per image). Calcein intensity was calculated using the ratio of harvest area versus background area intensity.

Alkaline phosphatase staining

Scales were collected in ALP buffer (100 mM Tris-HCl (pH 9.5), 100 mM NaCl, and 50 mM MgCl$_2$) and stained in ALP buffer containing 2% (v/v) NBT/BCIP (Sigma, cat# 11681451001). After a brief wash in deionised water, scales were mounted on a microscope slide containing Mowiol® 4-88 (10% w/v, Sigma-Aldrich cat# 9002-89-5) in glycerol (25% v/v) solution. Images were taken on an upright microscope (Leica).

Von Kossa staining and scale morphometric analysis

Plucked scales were washed in deionised water and then incubated in silver nitrate (5% w/v) for 1 h in strong light. After deionised water washes the scales were fixed in sodium thiosulphate (5% w/v) and mounted on a microscope slide containing Mowiol® 4-88 solution.

RNA isolation, RNA sequencing, transcriptomic mapping and analysis

Approximately 40 scales were collected from a standardised area on the left flank of 1-year-old male zebrafish; the area that extends from just behind the operculum to the implant of the dorsal fin. The area included multiple rows of scales of similar size and shape. Total RNA was isolated from ontogenetic and regenerating scales (*n*=3 fish per group, RU) by using Trizol (Invitrogen) for RNA-sequencing (RNA-seq) and downstream qRT-PCR testing.

RNA-seq was performed by ZF-GENOMICS (Leiden, NL) and involved quality control of total RNA libraries on 2100 expert bioanalyzer (Agilent) that resulted in RIN scores of > 9.2. Illumina RNAseq library preparation involved standard 6 nucleotide adaptor ligation. Paired single read 1 × 50 nucleotide runs (10 million reads; 0.5 Gb per sample) were performed on an Illumina Hiseq2500 system.

Transcriptome mapping and differential expression analysis

Raw reads were mapped to the GRCz11 primary assembly (Ensembl version 99) [86] using STAR (version STAR_2.5.4b) software pipeline [87]. The read count table for all genes mapped was obtained from the mapping step and filtered to leave out lowly expressed genes by only keeping genes that had at least 5 mapped reads

over all samples. The differential gene expression analysis was performed using the R-package DESeq2 (version 1.28.1) [88] including the medians of ratio normalisation step to account for the bias in sequencing depth/coverage and RNA composition of samples. For determining DEGs, we used a threshold of 1.25 $\log_2$ fold (2.4-fold) change and a false discovery rate (FDR) of < 0.05. Further details are in Additional file 9.

For downstream analyses, genes were classed as 'expressed' in scales when all three ontogenetic and regenerating scale samples produced a > 5 normalised read count (background gene list). Arbitrary threshold for differential expression was set at ±1.25 $\log_2$ fold and an adjusted p value (padj) of ≤0.05 (DEG list).

Gene ontology enrichment and STRING network analysis
DEG and background expression gene lists' gene symbols were uploaded to GOrilla [89] using *Danio rerio* and 'two unranked lists of genes' as settings. The hierarchical images and Microsoft Excel files were used for figure making. Ensembl IDs of DEGs were analysed with an 'Overrepresentation Test' (Released 20200407) using Fisher's exact test and False Discovery Rate correction in PANTHER Gene Ontology (release 2020-06-01), and PANTHER 'Pathways' and 'Reactome pathways' (PANTHER version 15.0) software [90].

For STRING (v11) gene network analysis [91], the DEG set (zebrafish gene symbols) was uploaded and interaction score (high, ≥0.7 or medium ≥0.4), number of interactions (one shell with max. 10 interactions), and active interaction sources (all were on) were set. More details and other downstream procedures [92–96] are in Additional file 9.

Gene set enrichment analysis involving monogenic and polygenic skeletal traits and disease
We used hypergeometric tests in conjunction with ISDS Nosology and Classification of Skeletal Disorders database [42] to identify whether DEGs were enriched for human orthologues of skeletal disease-causing genes. We repeated the analysis focussing on 42 broad skeletal disease classes that had been characterised by the nosology based on having similar clinical, radiographic, and molecular phenotypes. Only 24 of the 42 disorder groups were considered as no DEGs were present in the remaining disorder groups. A Bonferroni corrected significance threshold of $p < 2 \times 10^{-3}$, was used to declare statistical significance (total of 25 tests, assuming a type-1 error rate of $\alpha = 0.05$). MAGMA competitive gene set analysis (GSA) was used to investigate whether DEGs were enriched for orthologues of human genes that were associated with eBMD, height and any form of self-reported or hospital defined OA. GSA encompassed three stages. (Stage 1) Orthologous gene mapping:

Zebrafish – human orthologues of all protein coding genes were mapped using dbOrtho (BioDBnet) ortholog converter (https://biodbnet-abcc.ncifcrf.gov/db/dbOrtho.php) using Danio rerio (GRCZ11) as input ('Gene Symbol' or 'Ensembl ID') and Homo sapiens as output ('Ensembl ID') [94]. Only Ensembl IDs generating 1:1 results were kept, and duplicates of zebrafish paralogs (2:1 conversion) were removed leading to a list of single human Ensembl IDs (see Additional file 9). (Stage 2) Gene-based association testing: The association between human protein coding genes and eBMD, height and OA was evaluated separately using the weighted average of two different gene-based tests implemented in MAGMA v1.08 [97]. A Bonferroni corrected gene-wide significance threshold of $p < 2.5 \times 10^{-6}$ was used (19,856 tests, α = 0.05). MAGMA gene-based tests were conducted on summary results statistics from a GWAS performed inhouse using BOLT-LMM v2.3.4 [98], correcting each trait for age, sex, genotyping array and ancestry (Additional file 2: Fig. S10 and S11) informative principal components 1 - 20 as previously described (61). GWAS involved high quality genome-wide imputed v3 genetic data (~ 12 million SNPs, INFO > 0.9, MAF > 0.05%) measured in 448,010 white Europeans from the UK-Biobank Study that had both eBMD and height measured [99]. Gene-based tests were conducted on OA using GWAS summary results statistics from the arcO-GEN Consortium [62]. (Stage 3) MAGMA competitive GSA: The strength of association of human orthologues of DEGs with each trait was compared to the strength of association of all other protein coding genes in the genome. For each trait, GSA was conducted on all DEGs, upregulated DEGs, and downregulated DEGs. Evidence of enrichment was quantified using a one-sided test of statistical significance, together with a *Bonferroni* corrected $p < 5 \times 10^{-3}$ (i.e. 9 tests, α=0.05). Sensitivity analysis was performed by further adjusting for the set of human genes that could not be mapped between human and zebrafish and for the set of human orthologues that was not expressed in our zebrafish experiments. Post hoc permutation analysis was performed for analyses that were suggestive of enrichment. Refer to Additional file 9 for an in-depth description of all methods and datasets used [98, 100–104].

Quantitative real-time PCR
Five hundred nanograms of total RNA was treated with DNase (1 unit) and reverse-transcribed (random hexamer primers) with SuperScript II (Invitrogen 100 units). iQ SYBR Green Supermix (Biorad) containing 350 nM primer and cDNA was used for amplification (primers and PCR conditions in Additional file 9). Relative expression was calculated based on a normalisation index of two reference genes: *eef1a1l1* and *rpl13*. For

comparison of qRT-PCR and RNA-seq expression of amplicons, the average of ontogenetic normalised expression (qRT-PCR) or read count (RNA-seq) were taken and every individual value was compared to the average ontogenetic expression (e.g. read count regenerating scales individual 2/average read count ontogenetic). The p values presented in the figures were derived from a two-tailed t test (qRT-PCR) and p-adjusted (padj) from the DESEQ2 analysis (RNA-seq).

Micro-computed tomography and tissue mineral density calculations

MicroCT was performed as previously described [105], with estimated tissue mineral density (TMD) calculated as previously described [106]. Briefly, $col11a2^{Y228X/Y228X}$ ($n=7$), $spp1^{P160X/P160X}$ ($n=3$) mutant and age-matched WT ($n=7$) zebrafish (1-year-old, all AB/TL strain) were fixed, dehydrated to 70% EtOH and scanned at 21 μm voxel size (scan settings 130 kV, 150 μA, 0.5-s exposure, 3141 projections). Images were reconstructed using NRecon software (version 1.7.1.0), with dynamic ranges calibrated against a scan of hydroxyapatite phantoms (0.25 g cm^{-3} and 0.75 g cm^{-3}) scanned with identical parameters. For TMD calculations Aviso software (Avison2020.2; Thermo Fisher Scientific) was used to isolate pixel greyscale values for the skull (parietal) and vertebrae (vertebrae 11–13) and calibrated against the greyscale values of known hydroxyapatite density phantoms (0.25 g cm^{-3} and 0.75 g cm^{-3}) to estimate mean TMD values and their standard deviations. Histomorphological assessment of segmented jaw and vertebrae (11–13) elements were determined by measuring the width, length, and depth of the elements (N.B. left and right side of jaw elements were considered as independent measurements) in AVIZO between element landmarks as shown in Additional file 2: Fig. S8A and 8B.

Graphs and statistical testing

All graphs and data of RNA-seq and downstream functional data were analysed in GraphPad Prism (v 8.4.3). One-way ANOVA (wildtype against mutant conditions), two-way ANOVA (when multiple genotypes or amplicons and time points are involved) or unpaired t test (comparing ontogenetic against regenerating scale qRT-PCR expression data) statistical testing was performed. For ANOVA, results of the interaction factor (e.g. time × genotype) and/or independent factors were reported in the figure legends are as follows: 'f(degrees of freedom) = F ratio (DFd), p -value' when 'statistically significant (set at $p \leq 0.050$)' or just p value alone if not. A post hoc Tukey's (α = 0.05) multiple comparison test determined statistical difference between specific samples. All graphs show mean with standard deviation. Type

of tests performed is presented in the figure legends. P values in figures as follows: n.s. > 0.05, * ≤0.05, ** ≤0.01, *** ≤0.001. RNA-seq, GO, STRING pathway, and GSEA statistical analyses are described in their relevant sections.

Abbreviations
ALP: Alkaline phosphatase; DEG: Differentially expressed gene; Dph: Days post-harvest; eBMD: Estimated bone mineral density; ECM: Extracellular matrix; FDR: False discovery rate; GO: Gene ontology; GSA: Gene set analysis; IGF: Insulin-like growth factor; LRR: Leucine-rich repeat; MMP: Matrix metalloprotease; MS-222: Tricaine methanesulfonate; MSK: Musculoskeletal; NCAM: Neural cell adhesion molecule; NLR: NOD-like receptor; OA: Osteoarthritis; PCA: Principal component analysis; PPI: Protein-protein interactions; qRT-PCR: Quantitative real-time polymerase chain reaction; RNA-seq: RNA-sequencing; SCPP: Secretory calcium-binding phosphoprotein; SOST: Sclerostin; SPP: Secreted phosphoprotein; SPP1: Osteopontin; TMD: Tissue mineral density; TRAP: Tartrate-resistant acid phosphatase

Supplementary Information
The online version contains supplementary material available at https://doi.org/10.1186/s12915-021-01209-8.

Additional file 1. List of protein coding genes consistently expressed in ontogenetic and regenerating scales. Excel file containing all protein coding genes and expression values that were captured during RNA-seq experiment.

Additional file 2: Additional Tables (S1) and Figures (S1-S11). PDF file containing the following: **Table S1.** DEGs identified belonging to isolated clusters (1-4) with no identified gene ontology; **Figure S1.** Ontogenetic and regenerating differential expression values are clustered together; **Figure S2.** Two top down regulated genes are on the same genomic contig but are not the same protein; **Figure S3.** Hierarchical clustering of 'Molecular Function' enriched gene ontology terms using GOrilla web interface; **Figure S4.** Hierarchical clustering of 'Biological Process' enriched gene ontology terms; **Figure S5.** STRING Network analysis showing high protein-protein interaction connectivity of DEGs; **Figure S6.** Quantitative Real-Time PCR analysis of RNA expression of bone markers; **Figure S7.** MAGMA competitive gene set analysis involving human polygenetic traits and disease; **Figure S8.** Von Kossa staining of spp1 and col11a2 mutant ontogenetic and regenerating scales; **Figure S9.** Histomorphological measurements of skeletal elements from wild-type; **Figure S10.** Scatterplots describing pairwise comparisons of ancestry informative UMAP components; **Figure S11.** Scatterplots describing pairwise comparisons of ancestry informative PCA components.

Additional file 3. List of differentially expressed genes. Excel file containing list of genes classed as differentially expressed.

Additional file 4. Gene Ontology data file of DEG using GOrilla. Excel file with several tabs containing outputs from GOrilla GO analyses.

Additional file 5. Data output files of PANTHER Overrepresentation Test of process / function / component GO accession numbers and PANTHER Pathway analysis of DEGs. Excel file with several tabs containing outputs of various PANTHER analyses.

Additional file 6. Data file containing the PANTHER reactome of DEGs. Excel file with output from PANTHER reactome with DEGs.

Additional file 7. Summary of results of gene-set enrichment analysis involving DEGs and human orthologues that cause different skeletal dysplasia subgroups in humans when mutated. Excel file containing a table with nosology enrichment analysis.

Additional file 8. MAGMA gene level analysis results quantifying the strength of association between human orthologues of DEG's and estimated bone mineral density, height and osteoarthritis.

Additional file 9. Supplement to the Methods. PDF file containing extended methods.

Availability of data and materials
The raw RNA-sequencing FASTQ data files supporting the conclusions of this article are available in the European Nucleotide Archive (ENA) repository under accession number PRJEB39971 [107]. The downstream analysed RNA-sequencing datasets, GSA datasets, and *in vivo* data supporting the conclusions of this article are included in this published article and its additional information files. Summary statistics used for GSA studies are available in the cited studies in the relevant Method sections. Raw data files can be found in the data.bris repository by searching with the DOI number of this article [108].

Declarations

Ethics approval and consent to participate
Human genetic and animal research that is presented in this work has been conducted with ethical approval following the ARRIVE guidelines. Experiments were locally ethically reviewed (by AWERB at University of Bristol (Bristol, UK) (UoB) and Radboud University (Nijmegen, NL) (RU) respectively) and performed under UK Home Office project licence 30/3408 (UoB) or RU-DEC2014-059 (RU). Research that has been conducted using the UK Biobank Resource has been approved under accession ID: 53641.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Author details
[1]School of Physiology, Pharmacology, and Neuroscience, Faculty of Life Sciences, University of Bristol, Bristol, UK. [2]Musculoskeletal Research Unit, Translational Health Sciences, Bristol Medical School, Faculty of Health Sciences, University of Bristol, Bristol, UK. [3]Institute for Molecular Bioscience, The University of Queensland, Brisbane, Queensland 4072, Australia. [4]Department of Animal Ecology and Physiology, Radboud Institute for Biological and Environmental Sciences, Faculty of Science, Radboud University, Nijmegen, The Netherlands. [5]The University of Queensland Diamantina Institute, The University of Queensland, QLD, Woolloongabba 4102, Australia. [6]Transformational Bioinformatics, Commonwealth Scientific and Industrial Research Organisation, Sydney, New South Wales, Australia. [7]Bone Biology, Garvan Institute of Medical Research, Sydney, New South Wales, Australia. [8]St Vincent's Clinical School, Faculty of Medicine, UNSW Australia, Sydney, New South Wales, Australia. [9]Medical Research Council Integrative Epidemiology Unit, Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK. [10]School of Biotechnology and Biomolecular Sciences, UNSW Australia, Sydney, New South Wales, Australia.

References
1. Kenkre JS, Bassett JHD. The bone remodelling cycle. Ann Clin Biochem. 2018;55(3):308–27. https://doi.org/10.1177/0004563218759371.
2. Tiede-Lewis LM, Dallas SL. Changes in the osteocyte lacunocanalicular network with aging. Bone. 2019;122:101–13. https://doi.org/10.1016/j.bone.2019.01.025.
3. Witten PE, Harris MP, Huysseune A, Winkler C. Chapter 13 - Small teleost fish provide new insights into human skeletal diseases. In: Detrich HW, Westerfield M, Zon LI, editors. Methods in Cell Biology, vol. 138: Academic Press; 2017. p. 321–46.
4. Ofer L, Dean MN, Zaslansky P, Kult S, Shwartz Y, Zaretsky J, et al. A novel nonosteocytic regulatory mechanism of bone modeling. PLoS Biol. 2019; 17(2):e3000140. https://doi.org/10.1371/journal.pbio.3000140.
5. Suniaga S, Rolvien T, vom Scheidt A, IAK F, Bale HA, Huysseune A, et al. Increased mechanical loading through controlled swimming exercise induces bone formation and mineralization in adult zebrafish. Sci Rep. 2018; 8(1):3646.
6. Khajuria DK, Karasik D. Novel model of restricted mobility induced osteopenia in zebrafish. J Fish Biol. 2021;98(4):1031–8. https://doi.org/10.1111/jfb.14369.
7. Zhao A, Qin H, Fu X. What Determines the Regenerative Capacity in Animals? BioScience. 2016;66(9):735–46. https://doi.org/10.1093/biosci/biw079.
8. Lleras-Forero L, Winkler C, Schulte-Merker S. Zebrafish and medaka as models for biomedical research of bone diseases. Dev Biol. 2020;457(2):191–205. https://doi.org/10.1016/j.ydbio.2019.07.009.
9. Sehring IM, Weidinger G. Recent advancements in understanding fin regeneration in zebrafish. WIREs Dev Biol. 2020;9(1):e367. https://doi.org/10.1002/wdev.367.
10. Sire JY, Donoghue PC, Vickaryous MK. Origin and evolution of the integumentary skeleton in non-tetrapod vertebrates. J Anat. 2009;214(4): 409–40. https://doi.org/10.1111/j.1469-7580.2009.01046.x.
11. Yasuo M. The source of calcium in regenerating scales of the goldfish, Carassius auratus. Comp Biochem Physiol Part A: Physiology. 1980;66(3):521–4. https://doi.org/10.1016/0300-9629(80)90202-9.
12. Cox BD, De Simone A, Tornini VA, Singh SP, Di Talia S, Poss KD. In Toto Imaging of Dynamic Osteoblast Behaviors in Regenerating Skeletal Bone. Curr Biol. 2018;28(24):3937–47.e4.
13. De Simone A, Evanitsky MN, Hayden L, Cox BD, Wang J, Tornini VA, et al. Control of osteoblast regeneration by a train of Erk activity waves. Nature. 2021;590(7844):129–33. https://doi.org/10.1038/s41586-020-03085-8.
14. Metz JR, de Vrieze E, Lock EJ, Schulten IE, Flik G. Elasmoid scales of fishes as model in biomedical bone research. J Appl Ichthyology. 2012;28(3):382–7. https://doi.org/10.1111/j.1439-0426.2012.01990.x.
15. Sire JY, Akimenko MA. Scale development in fish: a review, with description of sonic hedgehog (shh) expression in the zebrafish (Danio rerio). Int J Dev Biol. 2004;48(2-3):233–47. https://doi.org/10.1387/ijdb.15272389.
16. Dhouailly D, Godefroit P, Martin T, Nonchev S, Caraguel F, Oftedal O. Getting to the root of scales, feather and hair: As deep as odontodes? Exp Dermatol. 2019;28(4):503–8. https://doi.org/10.1111/exd.13391.
17. Shimada A, Kawanishi T, Kaneko T, Yoshihara H, Yano T, Inohaya K, et al. Trunk exoskeleton in teleosts is mesodermal in origin. Nat Commun. 2013; 4(1):1639. https://doi.org/10.1038/ncomms2643.
18. Donoghue PCJ, Sansom IJ. Origin and early evolution of vertebrate skeletonization. Microsc Res Tech. 2002;59(5):352–72. https://doi.org/10.1002/jemt.10217.
19. Aman AJ, Fulbright AN, Parichy DM. Wnt/β-catenin regulates an ancient signaling network during zebrafish scale development. eLife. 2018;7:e37001. https://doi.org/10.7554/eLife.37001.

20. Sire JY, Allizard F, Babiar O, Bourguignon J, Quilhac A. Scale development in zebrafish (Danio rerio). J Anatomy. 1997;190((Pt 4)(Pt 4)):545–61.

21. Pasqualetti S, Banfi G, Mariotti M. The zebrafish scale as model to study the bone mineralization process. J Mol Histology. 2012;43(5):589–95. https://doi.org/10.1007/s10735-012-9425-z.

22. Richardson R, Slanchev K, Kraus C, Knyphausen P, Eming S, Hammerschmidt M. Adult Zebrafish as a Model System for Cutaneous Wound-Healing Research. J Invest Dermatol. 2013;133(6):1655–65. https://doi.org/10.1038/jid.2013.16.

23. Bereiter-Hahn J, Zylberberg L. Regeneration of teleost fish scale. Comp Biochem Physiol Part A: Physiology. 1993;105(4):625–41. https://doi.org/10.1016/0300-9629(93)90262-3.

24. de Vrieze E, van Kessel MA, Peters HM, Spanings FA, Flik G, Metz JR. Prednisolone induces osteoporosis-like phenotype in regenerating zebrafish scales. Osteoporos Int. 2014;25(2):567–78. https://doi.org/10.1007/s00198-013-2441-3.

25. Guellec DL, Zylberberg L. Expression of Type I and Type V Collagen mRNAs in the Elasmoid Scales of a Teleost Fish as Revealed by In Situ Hybridization. Connect Tissue Res. 1998;39(4):257–67. https://doi.org/10.3109/03008209809021501.

26. de Vrieze E, Sharif F, Metz JR, Flik G, Richardson MK. Matrix metalloproteinases in osteoclasts of ontogenetic and regenerating zebrafish scales. Bone. 2011;48(4):704–12. https://doi.org/10.1016/j.bone.2010.12.017.

27. Carnovali M, Luzi L, Banfi G, Mariotti M. Chronic hyperglycemia affects bone metabolism in adult zebrafish scale model. Endocrine. 2016;54(3):808–17. https://doi.org/10.1007/s12020-016-1106-3.

28. Carnovali M, Ottria R, Pasqualetti S, Banfi G, Ciuffreda P, Mariotti M. Effects of bioactive fatty acid amide derivatives in zebrafish scale model of bone metabolism and disease. Pharmacol Res. 2016;104:1–8. https://doi.org/10.1016/j.phrs.2015.12.009.

29. Pasqualetti S, Congiu T, Banfi G, Mariotti M. Alendronate rescued osteoporotic phenotype in a model of glucocorticoid-induced osteoporosis in adult zebrafish scale. Int J Exp Pathol. 2015;96(1):11–20. https://doi.org/10.1111/iep.12106.

30. Giraud-Guille MM. Twisted plywood architecture of collagen fibrils in human compact bone osteons. Calcif Tissue Int. 1988;42(3):167–80. https://doi.org/10.1007/BF02556330.

31. Bigi A, Burghammer M, Falconi R, Koch MHJ, Panzavolta S, Riekel C. Twisted Plywood Pattern of Collagen Fibrils in Teleost Scales: An X-ray Diffraction Investigation. J Struct Biol. 2001;136(2):137–43. https://doi.org/10.1006/jsbi.2001.4426.

32. Bergen DJM, Kague E, Hammond CL. Zebrafish as an Emerging Model for Osteoporosis: A Primary Testing Platform for Screening New Osteo-Active Compounds. Front Endocrinol (Lausanne). 2019;10:6.

33. Kobayashi-Sun J, Yamamori S, Kondo M, Kuroda J, Ikegame M, Suzuki N, et al. Uptake of osteoblast-derived extracellular vesicles promotes the differentiation of osteoclasts in the zebrafish scale. Commun Biol. 2020;3(1):190. https://doi.org/10.1038/s42003-020-0925-1.

34. Padhi BK, Joly L, Tellis P, Smith A, Nanjappa P, Chevrette M, et al. Screen for genes differentially expressed during regeneration of the zebrafish caudal fin. Dev Dyn. 2004;231(3):527–41. https://doi.org/10.1002/dvdy.20153.

35. Schmidt JR, Geurtzen K, von Bergen M, Schubert K, Knopf F. Glucocorticoid Treatment Leads to Aberrant Ion and Macromolecular Transport in Regenerating Zebrafish Fins. Front Endocrinol. 2019;10:674.

36. Wu XM, Chen WQ, Hu YW, Cao L, Nie P, Chang MX. RIP2 Is a Critical Regulator for NLRs Signaling and MHC Antigen Presentation but Not for MAPK and PI3K/Akt Pathways. Front Immunol. 2018;9:726.

37. Kawasaki K, Buchanan AV, Weiss KM. Gene Duplication and the Evolution of Vertebrate Skeletal Mineralization. Cells Tissues Organs. 2007;186(1):7–24. https://doi.org/10.1159/000102678.

38. Ricard-Blum S. The collagen family. Cold Spring Harb Perspect Biol. 2011;3(1):a004978. https://doi.org/10.1101/cshperspect.a004978.

39. Long F, Ornitz DM. Development of the endochondral skeleton. Cold Spring Harb Perspect Biol. 2013;5(1):a008334. https://doi.org/10.1101/cshperspect.a008334.

40. Zeltz C, Gullberg D. The integrin–collagen connection – a glue for tissue repair? J Cell Sci. 2016;129(4):653–64. https://doi.org/10.1242/jcs.180992.

41. Bhattacharya S, Hyland C, Falk MM, Iovine MK. Connexin 43 gap junctional intercellular communication inhibits evx1 expression and joint formation in regenerating fins. Development. 2020;147(13):dev190512.

42. Mortier GR, Cohn DH, Cormier-Daire V, Hall C, Krakow D, Mundlos S, et al. Nosology and classification of genetic skeletal disorders: 2019 revision. Am J Med Genet Part A. 2019;179(12):2393–419. https://doi.org/10.1002/ajmg.a.61366.

43. Couchouron T, Masson C. Early-onset progressive osteoarthritis with hereditary progressive ophtalmopathy or Stickler syndrome. Joint Bone Spine. 2011;78(1):45–9. https://doi.org/10.1016/j.jbspin.2010.03.012.

44. Iwasaki M, Kuroda J, Kawakami K, Wada H. Epidermal regulation of bone morphogenesis through the development and regeneration of osteoblasts in the zebrafish scale. Dev Biol. 2018;437(2):105–19. https://doi.org/10.1016/j.ydbio.2018.03.005.

45. Li S-W, Takanosu M, Arita M, Bao Y, Ren Z-X, Maier A, et al. Targeted disruption of Col11a2 produces a mild cartilage phenotype in transgenic mice: Comparison with the human disorder otospondylomegaepiphyseal dysplasia (OSMED). Dev Dyn. 2001;222(2):141–52. https://doi.org/10.1002/dvdy.1178.

46. Yadav MC, Huesa C, Narisawa S, Hoylaerts MF, Moreau A, Farquharson C, et al. Ablation of osteopontin improves the skeletal phenotype of phospho1(-/-) mice. J Bone Miner Res. 2014;29(11):2369–81. https://doi.org/10.1002/jbmr.2281.

47. Boskey AL, Spevak L, Paschalis E, Doty SB, McKee MD. Osteopontin Deficiency Increases Mineral Content and Mineral Crystallinity in Mouse Bone. Calcif Tissue Int. 2002;71(2):145–54. https://doi.org/10.1007/s00223-001-1121-z.

48. Steitz SA, Speer MY, McKee MD, Liaw L, Almeida M, Yang H, et al. Osteopontin Inhibits Mineral Deposition and Promotes Regression of Ectopic Calcification. Am J Pathol. 2002;161(6):2035–46. https://doi.org/10.1016/S0002-9440(10)64482-3.

49. Lawrence EA, Kague E, Aggleton JA, Harniman RL, Roddy KA, Hammond CL. The mechanical impact of col11a2 loss on joints; col11a2 mutant zebrafish show changes to joint development and function, which leads to early-onset osteoarthritis. Philos Trans R Soc B: Biol Sci. 2018;373(1759):20170335. https://doi.org/10.1098/rstb.2017.0335.

50. Hammond CL, Schulte-Merker S. Two populations of endochondral osteoblasts with differential sensitivity to Hedgehog signalling. Development. 2009;136(23):3991–4000. https://doi.org/10.1242/dev.042150.

51. Debiais-Thibaud M, Simion P, Ventéo S, Muñoz D, Marcellini S, Mazan S, et al. Skeletal Mineralization in Association with Type X Collagen Expression Is an Ancestral Feature for Jawed Vertebrates. Mol Biol Evol. 2019;36(10):2265–76. https://doi.org/10.1093/molbev/msz145.

52. Eames BF, Amores A, Yan Y-L, Postlethwait JH. Evolution of the osteoblast: skeletogenesis in gar and zebrafish. BMC Evol Biol. 2012;12(1):27. https://doi.org/10.1186/1471-2148-12-27.

53. Gregory KE, Oxford JT, Chen Y, Gambee JE, Gygi SP, Aebersold R, et al. Structural Organization of Distinct Domains within the Non-collagenous N-terminal Region of Collagen Type XI. J Biol Chem. 2000;275(15):11498–506. https://doi.org/10.1074/jbc.275.15.11498.

54. Li Y, Lacerda DA, Warman ML, Beier DR, Yoshioka H, Ninomiya Y, et al. A fibrillar collagen gene, Col11a1, is essential for skeletal morphogenesis. Cell. 1995;80(3):423–30. https://doi.org/10.1016/0092-8674(95)90492-1.

55. Page-McCaw A, Ewald AJ, Werb Z. Matrix metalloproteinases and the regulation of tissue remodelling. Nat Rev Mol Cell Biol. 2007;8(3):221–33. https://doi.org/10.1038/nrm2125.

56. Douglas T, Heinemann S, Bierbaum S, Scharnweber D, Worch H. Fibrillogenesis of collagen types I, II, and III with small leucine-rich proteoglycans decorin and biglycan. Biomacromolecules. 2006;7(8):2388–93. https://doi.org/10.1021/bm0603746.

57. Chen X-D, Fisher LW, Robey PG, Young MF. The small leucine-rich proteoglycan biglycan modulates BMP-4-induced osteoblast differentiation. FASEB J. 2004;18(9):948–58. https://doi.org/10.1096/fj.03-0899com.

58. Xu T, Bianco P, Fisher LW, Longenecker G, Smith E, Goldstein S, et al. Targeted disruption of the biglycan gene leads to an osteoporosis-like phenotype in mice. Nat Genet. 1998;20(1):78–82. https://doi.org/10.1038/1746.

59. Mosaffa P, Tetley RJ, Rodríguez-Ferran A, Mao Y, Muñoz JJ. Junctional and cytoplasmic contributions in wound healing. J R Soc Interface. 2020;17(169):20200264. https://doi.org/10.1098/rsif.2020.0264.

60. Kawasaki K. The SCPP gene repertoire in bony vertebrates and graded differences in mineralized tissues. Dev Genes Evol. 2009;219(3):147–57. https://doi.org/10.1007/s00427-009-0276-x.

61. Morris JA, Kemp JP, Youlten SE, Laurent L, Logan JG, Chai RC, et al. An atlas of genetic influences on osteoporosis in humans and mice. Nat Genet. 2019;51(2):258–66. https://doi.org/10.1038/s41588-018-0302-x.

62. Tachmazidou I, Hatzikotoulas K, Southam L, Esparza-Gordillo J, Haberland V, Zheng J, et al. Identification of new therapeutic targets for osteoarthritis through genome-wide analyses of UK Biobank data. Nat Genet. 2019;51(2): 230–6. https://doi.org/10.1038/s41588-018-0327-1.

63. Vogiatzi MG, Li D, Tian L, Garifallou JP, Kim CE, Hakonarson H, et al. A novel dominant COL11A1 mutation in a child with Stickler syndrome type II is associated with recurrent fractures. Osteoporosis Int. 2018;29(1):247–51. https://doi.org/10.1007/s00198-017-4229-3.

64. Lee HJ, Hou Y, Chen Y, Dailey ZZ, Riddihough A, Jang HS, et al. Regenerating zebrafish fin epigenome is characterized by stable lineage-specific DNA methylation and dynamic chromatin accessibility. Genome Biol. 2020;21(1):52. https://doi.org/10.1186/s13059-020-1948-0.

65. Sousa S, Valerio F, Jacinto A. A new zebrafish bone crush injury model. Biol Open. 2012;1(9):915–21. https://doi.org/10.1242/bio.2012877.

66. Morinobu M, Ishijima M, Rittling SR, Tsuji K, Yamamoto H, Nifuji A, et al. Osteopontin Expression in Osteoblasts and Osteocytes During Bone Formation Under Mechanical Stress in the Calvarial Suture In Vivo. J Bone Miner Res. 2003;18(9):1706–15. https://doi.org/10.1359/jbmr.2003.18.9.1706.

67. Terai K, Takano-Yamamoto T, Ohba Y, Hiura K, Sugimoto M, Sato M, et al. Role of Osteopontin in Bone Remodeling Caused by Mechanical Stress. J Bone Miner Res. 1999;14(6):839–49. https://doi.org/10.1359/jbmr.1999.14.6.839.

68. Bouleftour W, Juignet L, Verdière L, Machuca-Gayet I, Thomas M, Laroche N, et al. Deletion of OPN in BSP knockout mice does not correct bone hypomineralization but results in high bone turnover. Bone. 2019;120:411–22. https://doi.org/10.1016/j.bone.2018.12.001.

69. Holm E, Gleberzon Jared S, Liao Y, Sørensen Esben S, Beier F, Hunter Graeme K, et al. Osteopontin mediates mineralization and not osteogenic cell development in vitro. Biochem J. 2014;464(3):355–64. https://doi.org/10.1042/BJ20140702.

70. Kim H-J, Lee M-H, Park H-S, Park M-H, Lee S-W, Kim S-Y, et al. Erk pathway and activator protein 1 play crucial roles in FGF2-stimulated premature cranial suture closure. Dev Dyn. 2003;227(3):335–46. https://doi.org/10.1002/dvdy.10319.

71. Topczewska JM, Shoela RA, Tomaszewski JP, Mirmira RB, Gosain AK. The Morphogenesis of Cranial Sutures in Zebrafish. PLoS One. 2016;11(11): e0165775-e.

72. Dai J, Peng L, Fan K, Wang H, Wei R, Ji G, et al. Osteopontin induces angiogenesis through activation of PI3K/AKT and ERK1/2 in endothelial cells. Oncogene. 2009;28(38):3412–22. https://doi.org/10.1038/onc.2009.189.

73. Mori R, Shaw TJ, Martin P. Molecular mechanisms linking wound inflammation and fibrosis: knockdown of osteopontin leads to rapid repair and reduced scarring. J Exp Med. 2008;205(1):43–51. https://doi.org/10.1084/jem.20071412.

74. Lamort A-S, Giopanou I, Psallidas I, Stathopoulos GT. Osteopontin as a Link between Inflammation and Cancer: The Thorax in the Spotlight. Cells. 2019; 8(8):815. https://doi.org/10.3390/cells8080815.

75. Balemans W, Ebeling M, Patel N, Van Hul E, Olson P, Dioszegi M, et al. Increased bone density in sclerosteosis is due to the deficiency of a novel secreted protein (SOST). Hum Mol Genet. 2001;10(5):537–44. https://doi.org/10.1093/hmg/10.5.537.

76. Li X, Ominsky MS, Niu Q-T, Sun N, Daugherty B, D'Agostin D, et al. Targeted Deletion of the Sclerostin Gene in Mice Results in Increased Bone Formation and Bone Strength. J Bone Miner Res. 2008;23(6):860–9. https://doi.org/10.1359/jbmr.080216.

77. Cosman F, Crittenden DB, Adachi JD, Binkley N, Czerwinski E, Ferrari S, et al. Romosozumab Treatment in Postmenopausal Women with Osteoporosis. N Engl J Med. 2016;375(16):1532–43. https://doi.org/10.1056/NEJMoa1607948.

78. Wainwright DK, Ingersoll S, Lauder GV. Scale diversity in bigeye tuna (Thunnus obesus): Fat-filled trabecular scales made of cellular bone. J Morphol. 2018;279(6):828–40. https://doi.org/10.1002/jmor.20814.

79. de Vrieze E, Zethof J, Schulte-Merker S, Flik G, Metz JR. Identification of novel osteogenic compounds by an ex-vivo sp7:luciferase zebrafish scale assay. Bone. 2015;74(Supplement C):106–13.

80. Alestrom P, D'Angelo L, Midtlyng PJ, Schorderet DF, Schulte-Merker S, Sohm F, et al. Zebrafish: Housing and husbandry recommendations. Lab Anim. 2020;54(3):213–24. https://doi.org/10.1177/0023677219869037.

81. Bevan L, Lim ZW, Venkatesh B, Riley PR, Martin P, Richardson RJ. Specific macrophage populations promote both cardiac scar deposition and subsequent resolution in adult zebrafish. Cardiovasc Res. 2020;116(7):1357–71. https://doi.org/10.1093/cvr/cvz221.

82. DeLaurier A, Eames BF, Blanco-Sánchez B, Peng G, He X, Swartz ME, et al. Zebrafish sp7:EGFP: A transgenic for studying otic vesicle formation, skeletogenesis, and bone regeneration. Genesis. 2010;48(8):505–11.

83. Singh Sumeet P, Holdway Jennifer E, Poss KD. Regeneration of Amputated Zebrafish Fin Rays from De Novo Osteoblasts. Dev Cell. 2012;22(4):879–86. https://doi.org/10.1016/j.devcel.2012.03.006.

84. Kague E, Gallagher M, Burke S, Parsons M, Franz-Odendaal T, Fisher S. Skeletogenic fate of zebrafish cranial and trunk neural crest. PLoS One. 2012; 7(11):e47394-e.

85. Mitchell RE, Huitema LF, Skinner RE, Brunt LH, Severn C, Schulte-Merker S, et al. New tools for studying osteoarthritis genetics in zebrafish. Osteoarthr Cartil. 2013;21(2):269–78. https://doi.org/10.1016/j.joca.2012.11.004.

86. Cunningham F, Achuthan P, Akanni W, Allen J, Amode MR, Armean IM, et al. Ensembl 2019. Nucleic Acids Res. 2018;47(D1):D745–D51. https://doi.org/10.1093/nar/gky1113.

87. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 2013;29(1):15–21. https://doi.org/10.1093/bioinformatics/bts635.

88. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15(12):550. https://doi.org/10.1186/s13059-014-0550-8.

89. Eden E, Navon R, Steinfeld I, Lipson D, Yakhini Z. GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. BMC Bioinformatics. 2009;10:48.

90. Mi H, Muruganujan A, Ebert D, Huang X, Thomas PD. PANTHER version 14: more genomes, a new PANTHER GO-slim and improvements in enrichment analysis tools. Nucleic Acids Res. 2019;47(D1):D419–26. https://doi.org/10.1093/nar/gky1038.

91. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, et al. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Res. 2019;47(D1):D607–D13. https://doi.org/10.1093/nar/gky1131.

92. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol Syst Biol. 2011;7:539.

93. The UPC. UniProt: a worldwide hub of protein knowledge. Nucleic Acids Res. 2018;47(D1):D506–D15. https://doi.org/10.1093/nar/gky1049.

94. Mudunuri U, Che A, Yi M, Stephens RM. bioDBnet: the biological database network. Bioinformatics. 2009;25(4):555–6. https://doi.org/10.1093/bioinformatics/btn654.

95. Howe K, Clark MD, Torroja CF, Torrance J, Berthelot C, Muffato M, et al. The zebrafish reference genome sequence and its relationship to the human genome. Nature. 2013;496(7446):498–503. https://doi.org/10.1038/nature12111.

96. Meyer A, Schartl M. Gene and genome duplications in vertebrates: the one-to-four (-to-eight in fish) rule and the evolution of novel gene functions. Curr Opin Cell Biol. 1999;11(6):699–704. https://doi.org/10.1016/S0955-0674(99)00039-3.

97. de Leeuw CA, Mooij JM, Heskes T, Posthuma D. MAGMA: Generalized Gene-Set Analysis of GWAS Data. PLoS Comput Biol. 2015;11(4):e1004219. https://doi.org/10.1371/journal.pcbi.1004219.

98. Loh P-R, Tucker G, Bulik-Sullivan BK, Vilhjálmsson BJ, Finucane HK, Salem RM, et al. Efficient Bayesian mixed-model analysis increases association power in large cohorts. Nat Genet. 2015;47(3):284–90. https://doi.org/10.1038/ng.3190.

99. Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, et al. UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. PLoS Med. 2015;12(3):e1001779. https://doi.org/10.1371/journal.pmed.1001779.

100. Auton A, Abecasis GR, Altshuler DM, Durbin RM, Abecasis GR, Bentley DR, et al. A global reference for human genetic variation. Nature. 2015; 526(7571):68–74. https://doi.org/10.1038/nature15393.

101. Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: a tool for genome-wide complex trait analysis. Am J Hum Genet. 2011;88(1):76–82. https://doi.org/10.1016/j.ajhg.2010.11.011.

102. Altshuler DM, Gibbs RA, Peltonen L, Altshuler DM, Gibbs RA, Peltonen L, et al. Integrating common and rare genetic variation in diverse human populations. Nature. 2010;467(7311):52–8. https://doi.org/10.1038/nature09298.

103. Panoutsopoulou K, Southam L, Elliott KS, Wrayner N, Zhai G, Beazley C, et al. Insights into the genetic architecture of osteoarthritis from stage 1 of the

arcOGEN study. Ann Rheumatic Dis. 2011;70(5):864–7. https://doi.org/10.1136/ard.2010.141473.

104. Kemp JP, Morris JA, Medina-Gomez C, Forgetta V, Warrington NM, Youlten SE, et al. Identification of 153 new loci associated with heel bone mineral density and functional involvement of GPC6 in osteoporosis. Nat Genet. 2017;49(10):1468–75. https://doi.org/10.1038/ng.3949.

105. Kague E, Hughes SM, Lawrence EA, Cross S, Martin-Silverstone E, Hammond CL, et al. Scleraxis genes are required for normal musculoskeletal development and for rib growth and mineralization in zebrafish. FASEB J. 2019;33(8):9116–30. https://doi.org/10.1096/fj.201802654RR.

106. Stevenson NL, DJM B, REH S, Kague E, Martin-Silverstone E, Robson Brown KA, et al. Giantin knockout models reveal a feedback loop between Golgi function and glycosyltransferase expression. J Cell Sci. 2017;130(24):4132–43.

107. Bergen DJM, Metz JR. RNA-sequencing data of ontogenetic and 9-days post-harvest zebrafish scale regeneration. ENA https://identifiers.org/ena.embl:PRJEB39971. 2021.

108. Bergen DJM, Hammond CL, Metz JR. https://data.bris.ac.uk/data/. University of Bristol Research Data Repository. 2021.

Publisher's Note