nature medicine

LETTERS

https://doi.org/10.1038/s41591-019-0353-2

Gain-of-function mutation of microRNA-140 in human skeletal dysplasia

Giedre Grigelioniene[1,2,3,15], Hiroshi I. Suzuki[4,15], Fulya Taylan[2], Fatemeh Mirzamohammadi[1], Zvi U. Borochowitz[5], Ugur M. Ayturk[6], Shay Tzur[7,8], Eva Horemuzova[9,10], Anna Lindstrand[2,3], Mary Ann Weis[11], Gintautas Grigelionis[2], Anna Hammarsjö[2,3], Elin Marsk[12], Ann Nordgren[2,3], Magnus Nordenskjöld[2,3], David R. Eyre[11], Matthew L. Warman[6], Gen Nishimura[13], Phillip A. Sharp[4,14] and Tatsuya Kobayashi[1*]

MicroRNAs (miRNAs) are post-transcriptional regulators of gene expression. Heterozygous loss-of-function point mutations of miRNA genes are associated with several human congenital disorders[1–5], but neomorphic (gain-of-new-function) mutations in miRNAs due to nucleotide substitutions have not been reported. Here we describe a neomorphic seed region mutation in the chondrocyte-specific, super-enhancer-associated *MIR140* gene encoding microRNA-140 (miR-140) in a novel autosomal dominant human skeletal dysplasia. Mice with the corresponding single nucleotide substitution show skeletal abnormalities similar to those of the patients but distinct from those of miR-140-null mice[6]. This mutant miRNA gene yields abundant mutant miR-140-5p expression without miRNA-processing defects. In chondrocytes, the mutation causes widespread derepression of wild-type miR-140-5p targets and repression of mutant miR-140-5p targets, indicating that the mutation produces both loss-of-function and gain-of-function effects. Furthermore, the mutant miR-140-5p seed competes with the conserved RNA-binding protein Ybx1 for overlapping binding sites. This finding may explain the potent target repression and robust in vivo effect by this mutant miRNA even in the absence of evolutionary selection of miRNA–target RNA interactions, which contributes to the strong regulatory effects of conserved miRNAs[7,8]. Our study presents the first case of a pathogenic gain-of-function miRNA mutation and provides molecular insight into neomorphic actions of emerging and/or mutant miRNAs.

Genetic skeletal disorders comprise a large group of approximately 500 clinically distinct and genetically heterogeneous conditions. Mutations in approximately 400 different genes are currently known to cause congenital skeletal diseases[9], but underlying molecular mechanisms are not elucidated for many of them. Herein, we describe a novel skeletal dysplasia caused by a single nucleotide substitution in the *MIR140* gene and the underlying mechanism.

A novel skeletal dysplasia was identified in two unrelated families in a project for molecular diagnosis of ultrarare congenital skeletal disorders (Fig. 1a,b). In family 1, individual II-2, patient 1, P1, is affected, as is one of her sons (III-2, P2). In family 2, II-1, P3 is the only affected individual. Clinical features of the skeletal dysplasia included disproportionate short stature with short limbs, small hands and feet, and midface hypoplasia with small nose. The radiological hallmarks were mild spondylar dysplasia, delayed epiphyseal ossification of the hip and knee, and severe brachydactyly with cone-shaped phalangeal epiphyses (Fig. 1c–p and auxology data in Supplementary Table 1). The spondylar and epiphyseal abnormalities evolved into premature spondylosis and degenerative joint disease in adulthood, respectively. P1 and P2 also suffered from frequent respiratory infections with prolonged cough and inspiratory stridor. P1 had a narrow larynx likely due to floppy cartilage (Extended Data Fig. 1a–c). All three affected individuals had normal intelligence, dentition, hearing, visual acuity, and basic blood tests.

Whole-genome sequencing (WGS) identified the same heterozygous nucleotide substitution (chr16:g.69967007A>G (hg19), *MIR140*:NR_029681.1:n.24A>G) in P1 and P2 (Fig. 1a,q, Extended Data Fig. 1d, and Methods). This nucleotide substitution occurred de novo in P1 and P3, and co-segregated with the skeletal dysplasia phenotype in family 1 (Fig. 1a and Extended Data Fig. 1d). In P3, since the phenotype was recognized as this novel clinical diagnostic entity, the pathogenic variant was first detected by Sanger sequencing and then confirmed retrospectively in the exome sequencing and WGS data (Extended Data Fig. 1d). This substitution is located at the first nucleotide of the seed sequence of the highly conserved miRNA miR-140-5p, encoded by *MIR140* (Fig. 1q,r and Extended Data Fig. 1e), and is not reported in gnomAD, a large-scale human genome sequence database[10]. *MIR140* is associated with a chondrocyte-specific super-enhancer both in humans and mice (Fig. 1s and Extended Data Fig. 2), consistent with this miRNA's high expression in chondrocytes. Super-enhancers commonly drive high-level

[1]Endocrine Unit, Massachusetts General Hospital and Harvard Medical School, Boston, MA, USA. [2]Department of Molecular Medicine and Surgery and Center for Molecular Medicine, Karolinska Institutet, Stockholm, Sweden. [3]Department of Clinical Genetics, Karolinska University Hospital, Stockholm, Sweden. [4]David H. Koch Institute for Integrative Cancer Research, Massachusetts Institute of Technology, Cambridge, MA, USA. [5]Rappaport Faculty of Medicine, Technion—Israeli Institute of Technology, Medical Genetics Clinics, Assuta Medical Center, Haifa, Israel. [6]Orthopaedic Research Labs, Department of Orthopaedic Surgery, Boston Children's Hospital, Boston, MA, USA. [7]Laboratory of Molecular Medicine, Rambam Health Care Campus, Haifa, Israel. [8]Genomic Research Department, Emedgene Technologies, Tel Aviv, Israel. [9]Department for Women's and Children's Health, Karolinska Institutet, Stockholm, Sweden. [10]Pediatric Endocrinology Unit, Karolinska University Hospital, Stockholm, Sweden. [11]Department of Orthopedics and Sports Medicine, University of Washington, Seattle, WA, USA. [12]Department of Otorhinolaryngology, Karolinska University Hospital, Stockholm, Sweden. [13]Center for Intractable Diseases, Saitama Medical University Hospital, Saitama, Japan. [14]Department of Biology, Massachusetts Institute of Technology, Cambridge, MA, USA. [15]These authors contributed equally: Giedre Grigelioniene, Hiroshi I. Suzuki. *e-mail: tkobayashi1@mgh.harvard.edu

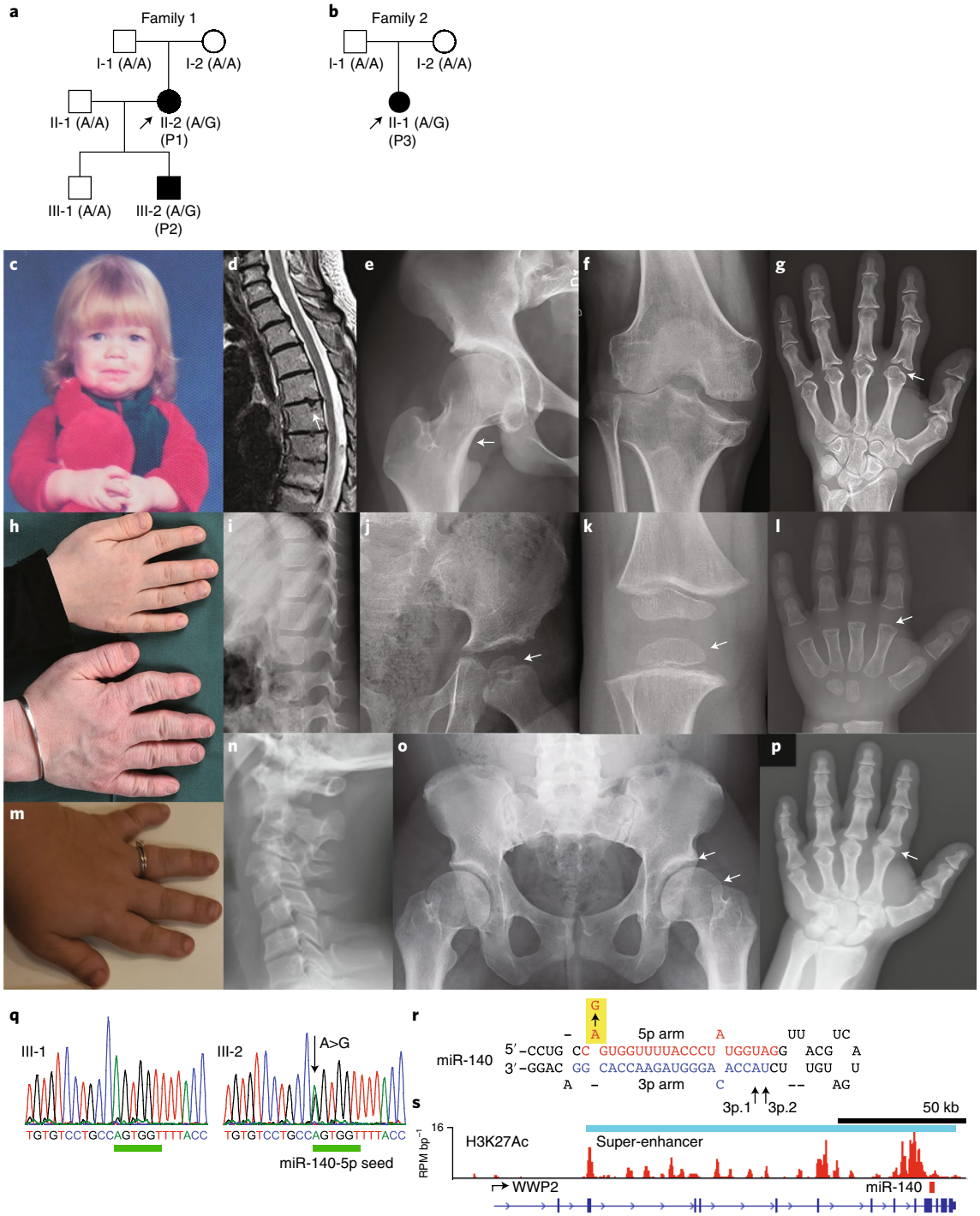


Fig. 1 | Novel skeletal dysplasia associated with miR-140-5p mutation. a,**b**, Pedigrees of family 1 (**a**) and family 2 (**b**). Black filled symbols indicate affected individuals; black arrows indicate probands. II-2 and III-2 in family 1 are P1 and P2, respectively. II-1 in family 2 denotes P3. The genotype of *MIR140* at the position of the mutation (A/A or A/G) is indicated for each individual in the study. **c**, Patient 1 at 2 years of age; note midface hypoplasia and short nose. The patient has kindly provided her childhood photo for publication, and the authors have obtained signed proof of informed consent. **d**–**g**, Radiograms of P1 at 40 years of age. Premature degeneration of intervertebral disks in the cervical spine (**d**); hypoplasia of the lower ilium, and short, constricted femoral neck (white arrow) (**e**); premature degenerative changes of the knee joint (**f**); brachydactyly with coning of the proximal phalangeal end, particularly of the second digit (white arrow) (**g**). **h**, Brachydactyly and broad thumbs in P1 and her son, P2. **i**–**l**, Radiograms of P2 at 4 years of age. Mild irregularity of vertebral endplates at the thoracolumbar junction (**i**); hypoplasia of the lower ilium, and delayed ossification of the proximal femoral epiphysis (white arrow) (**j**); small epiphyses of the distal femur and proximal tibia (white arrow) (**k**); brachydactyly with delayed epiphyseal maturation (white arrow) (**l**). **m**–**p**, P3 at 30 years of age: note brachydactyly and deformities of interphalangeal joints (**m**); premature disk degeneration in the upper cervical spine (**n**); hypoplasia of the lower ilium, and short femoral necks (white arrows) (**o**); brachydactyly with mild coning of the proximal phalangeal end of the index finger (white arrow) (**p**). **q**, Schematic representation of the mutation in miR-140-5p (nucleotides in the seed sequence are marked with green). The mutated nucleotide is indicated by a black arrow. **r**, Sequence of miR-140 precursor. Heterogeneity in Dicer processing yields two species of miR-140-3p (3p.1 and 3p.2). **s**, Histone 3 lysine 27 acetylation (H3K27Ac) ChIP–seq showing super-enhancers associated with *MIR140* in human chondrocytes. RPM, reads per million.

expression of tissue-specific genes, including miRNA genes, with evolutionarily conserved cell-type-specific function[11–13].

miR-140-5p and miR-140-3p (collectively miR-140) are abundantly and relatively specifically expressed in chondrocytes[14,15], and loss of miR-140, encoded by *Mir140*, in mice leads to short stature and craniofacial abnormalities[6,16]. This miR-140-null phenotype included mildly advanced terminal differentiation of growth plate chondrocytes of developing bones[6,17]. To investigate the possible disease-causing role of the mutant miR-140, we introduced this single-nucleotide substitution into the mouse genome using the CRISPR–Cas9 technology[18]. We then compared the skeletal phenotype of heterozygous (miR-140$^{G/+}$) and homozygous (miR-140$^{G/G}$) knock-in mice with those of wild-type (WT, miR-140$^{+/+}$) and the previously reported knockout (miR-140$^{-/-}$) mice (Supplementary Table 2)[6].

Although miR-140$^{G/G}$ and miR-140$^{-/-}$ mice were smaller and had shorter noses than WT littermates (Fig. 2a,b and Extended Data Fig. 3a), only miR-140$^{G/G}$ and miR-140$^{G/+}$ mice showed delayed secondary ossification of tubular and carpal bones (Fig. 2c–f), decreased expression of *Col10a1*, a maturation marker for hypertrophic chondrocytes (Extended Data Fig. 3b), and delayed cartilage maturation of the larynx, trachea, and anterior ribs (Extended Data Fig. 3c,d). Micro-computed tomography analysis showed severely decreased epiphyseal mineralization in miR-140$^{G/G}$ mice (Fig. 2g). The proliferation rate of tibial growth plate chondrocytes was not significantly different among miR-140$^{G/G}$, miR-140$^{G/+}$, and WT mice (Extended Data Fig. 3e). We did not find an increased number of apoptotic cells in the growth plate of miR-140$^{G/G}$ mice (Extended Data Fig. 3f). miR-140$^{G/G}$ and miR-140$^{G/+}$ mice showed mildly flat vertebral bodies (Extended Data Fig. 3g). The basal skull growth plate was wider owing to the increased length of all layers of chondrocytes in miR-140$^{G/G}$ and miR-140$^{G/+}$ mice, whereas miR-140$^{-/-}$ mice showed no significant changes (Fig. 2h). The proximal tibial growth plate also showed an expansion of the resting zone in miR-140$^{G/G}$ and miR-140$^{G/+}$ mice (Extended Data Fig. 3h). These bone abnormalities of miR-140$^{G/G}$ and miR-140$^{G/+}$ mice are consistent at all analyzed ages from postnatal day 7 (P7) to P56 when compared with WT mice (Supplementary Table 2) and are highly consistent with the skeletal dysplasia features of the patients, presenting with delayed secondary ossification, mild platyspondyly, small epiphyses, and scaphocephaly. Taken together, these differences in phenotype between miR-140$^{G/G}$ and miR-140$^{-/-}$ mice and delayed epiphyseal maturation in humans indicate that this single-base substitution causes a neomorphic, and not just loss-of-function, effect. Furthermore, the delayed skeletal maturation observed in humans and mice with the miR-140 substitution suggests a common pathogenic mechanism.

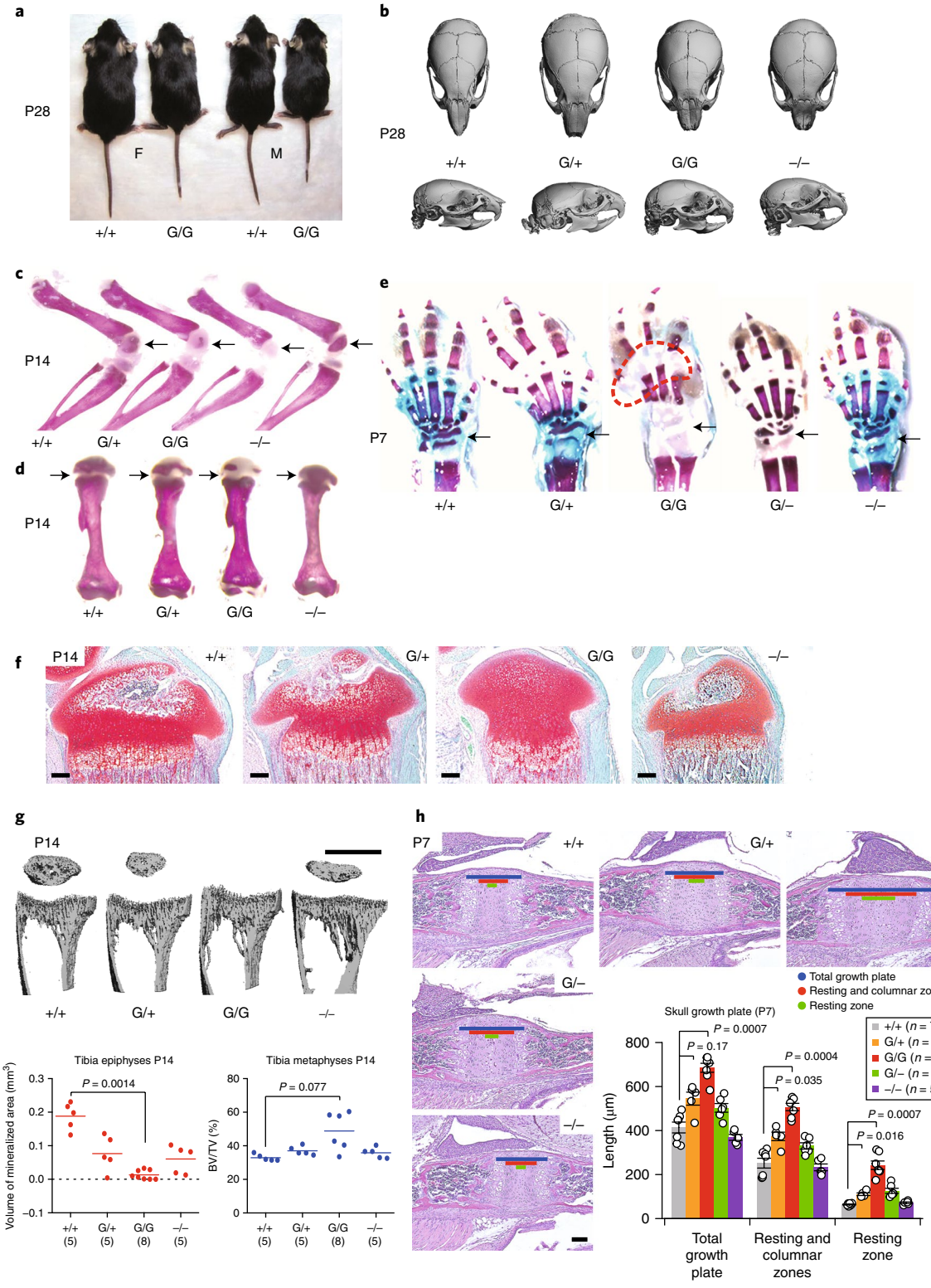
Small RNA-sequencing (RNA-seq) analysis in mouse chondrocytes revealed that the miR-140 A>G mutation yields slightly greater amounts of the mutant miR-140-5p (miR-140-5p-G) than WT miR-140-5p, while decreasing two miR-140-3p species (3p.1 and 3p.2) (Fig. 3a, Extended Data Fig. 4a–c, and Supplementary Table 3). Minor changes in Dicer processing sites and miRNA strand choice were also observed (Extended Data Fig. 4d,e); this finding is consistent with the critical role of the 5′-end structure of pre-miRNAs and miRNA duplexes in Dicer processing and in asymmetric miRNA strand selection[19,20]. These data indicate that the A>G mutation did not compromise miRNA processing and yielded abundant expression of miR-140-5p-G from primary miR-140 transcripts of which transcription is strongly activated by a super-enhancer in chondrocytes.

Since miRNAs have global impacts on target messenger RNA (mRNA) expression, we next investigated transcriptome effects of the miR-140 mutation. Principal component analysis and hierarchical clustering analysis of the RNA-seq data clearly distinguished miR-140$^{G/G}$ and miR-140$^{G/+}$ mice from miR-140$^{-/-}$ and WT mice (Fig. 3b,c). Expression levels of genes within the neighborhood of the miR-140 gene locus were not largely affected (Extended Data Fig. 5a). We predicted potential de novo targets of miR140-5p-G using TargetScan (v7.0/1)[21] (Extended Data Fig. 5b and Supplementary Tables 4 and 5). Target sites of miR-140-5p-G seldom overlapped with those of WT miR-140 species (WT 5p, 3p.1, and 3p.2) and were less conserved than those of WT miR-140 species (Extended Data Fig. 5c–e). Both the miR-140 A>G mutation and miR-140 deletion caused derepression of conserved 8mer target genes for WT miR-140-5p and, to a lesser extent, for miR-140-3p species (Fig. 3d and Extended Data Fig. 6a), suggesting that miR-140-5p exerts greater target repression than miR-140-3p even though the latter is more abundant. Importantly, the widespread suppression of predicted miR-140-5p-G targets is only observed in miR-140$^{G/G}$ and miR-140$^{G/+}$ chondrocytes (Fig. 3d and Extended Data Fig. 6a). This suppressive effect of miR-140-5p-G is particularly strong for 8mer target sites (Extended Data Fig. 6b) and correlates with the predicted miRNA–target efficiency (Extended Data Fig. 6c), suggesting that the miR-140-5p-G–target interaction follows the molecular rules of conventional miRNA targeting. Accordingly, WT and the mutant miR-140-5p showed similar target-site efficiencies by TargetScan prediction (Extended Data Fig. 6d). A 3′ untranslated region (UTR) reporter assay, using minimal target sequences and endogenous 3′ UTR sequences, confirmed that miR-140-5p-G and WT miR-140-5p target their respective predicted 8mer target sites with similar strengths, and validated the efficiency of target sites in several genes, respectively (Fig. 3e,f and Supplementary Table 6).

The miR-140-5p-G targets include many genes important for skeletal development and homeostasis, including *Loxl3*, *Btg1*, and *Trps1*, and genes associated with various metabolic pathways

Fig. 2 | Skeletal phenotype of miR-140 mutant mice. a, Both female (F) and male (M) homozygous miR-140$^{G/G}$ (G/G) mice are smaller in size than WT (+/+) littermates at P28 (+/+, $n=6$; G/+, $n=10$; G/G, $n=5$). **b**, Micro-computed tomography scans of the skull at P28, note the shortening of the nasal bone in homozygous (G/G) and miR-140 knockout (−/−) mice. **c**–**e**, Skeletal preparation. Delayed secondary ossification (black arrows) in epiphyses of femura (**c**) and humeri (**d**) of homozygous (G/G) and heterozygous (G/+) mice (+/+, $n=4$; G/+, $n=7$; G/G, $n=9$; −/−, $n=2$). Note that epiphyseal ossification is normal in (−/−) mice. Delayed ossification is similarly present in carpal bones (black arrows) and epiphyses of metacarpals (area marked with a red dashed line) of (G/G), (G/+), and (G/−) mice (**e**) (+/+, $n=7$; G/+, $n=6$; G/G, $n=8$; −/−, $n=5$). **f**, Representative images from three independent experiments of safranin O-stained sections of the proximal tibia. Each experiment included duplicate sections of tibias from at least two animals of each genotype. Note the absence and delay in vascular invasion in (G/G) and (G/+) epiphyses, respectively. **g**, Micro-computed tomography analysis showing decreased epiphyseal mineralization in (G/G) mice (+/+, $n=5$; G/+, $n=5$; G/G, $n=8$; −/−, $n=5$). All animals used for micro-computed tomography analysis were males. Note that bone volume and total volume ratios (BV/TV) in the metaphyses do not differ among the different genotypes. Horizontal bars indicate mean values. **h**, H&E-stained sagittal sections of the cranial base growth plate between the basisphenoid and basioccipital bones. Note significant expansion of all zones (blue bars, total growth plate; red bars, the proliferating zone containing the resting and columnar zones; green bars the resting zone) in (G/G) compared with (+/+) mice. (G/+) and (G/−) mice also show expansion of the growth plate, although to a smaller extent than (G/G) mice. The bar graph shows mean ± s.e.m. (+/+, $n=7$; G/+, $n=5$; G/G, $n=7$; G/−, $n=6$; −/−, $n=5$). Postnatal ages (P) are indicated in days. Scale bars, 200 μm in microscope images and 1 mm in micro-computed tomography images. In **g** and **h**, statistical significance was assessed using Kruskal–Wallis one-way ANOVA and post hoc Steel–Dwass test.

(Fig. 3g and Supplementary Table 7)[22–24]. Furthermore, gene set enrichment analysis (GSEA) revealed that the hypoxia pathway is impaired in miR-140$^{G/G}$, but not in miR-140$^{-/-}$ chondrocytes (Fig. 3h,i)[25,26]. The coding region of *Hif1a* that encodes a master transcription factor controlling hypoxic response[27] contains a conserved sequence complementary to the miR-140-5p-G seed sequence. Hif1a protein expression was reduced in miR-140$^{G/G}$ chondrocytes (Extended Data Fig. 7a). Along with the mutation experiment of the target site (Extended Data Fig. 7b), this suggests that miR-140-G directly suppresses *Hif1a*. Consistent with these findings, mRNA expression of Hif1a and Hif1a-regulated molecules including lysyl oxidases, essential for collagen cross-linking, significantly decreased in miR-140$^{G/G}$ chondrocytes, with a concomitant decrease in collagen cross-linking in the rib cartilage (Fig. 3j and Extended Data Fig. 7c–e). These alterations in genes important for cartilage and chondrocyte functions may contribute to the skeletal abnormalities of miR-140$^{G/G}$ and miR-140$^{G/+}$ mice. In addition, expression of several extracellular matrix genes including *Col10a1* decreased in



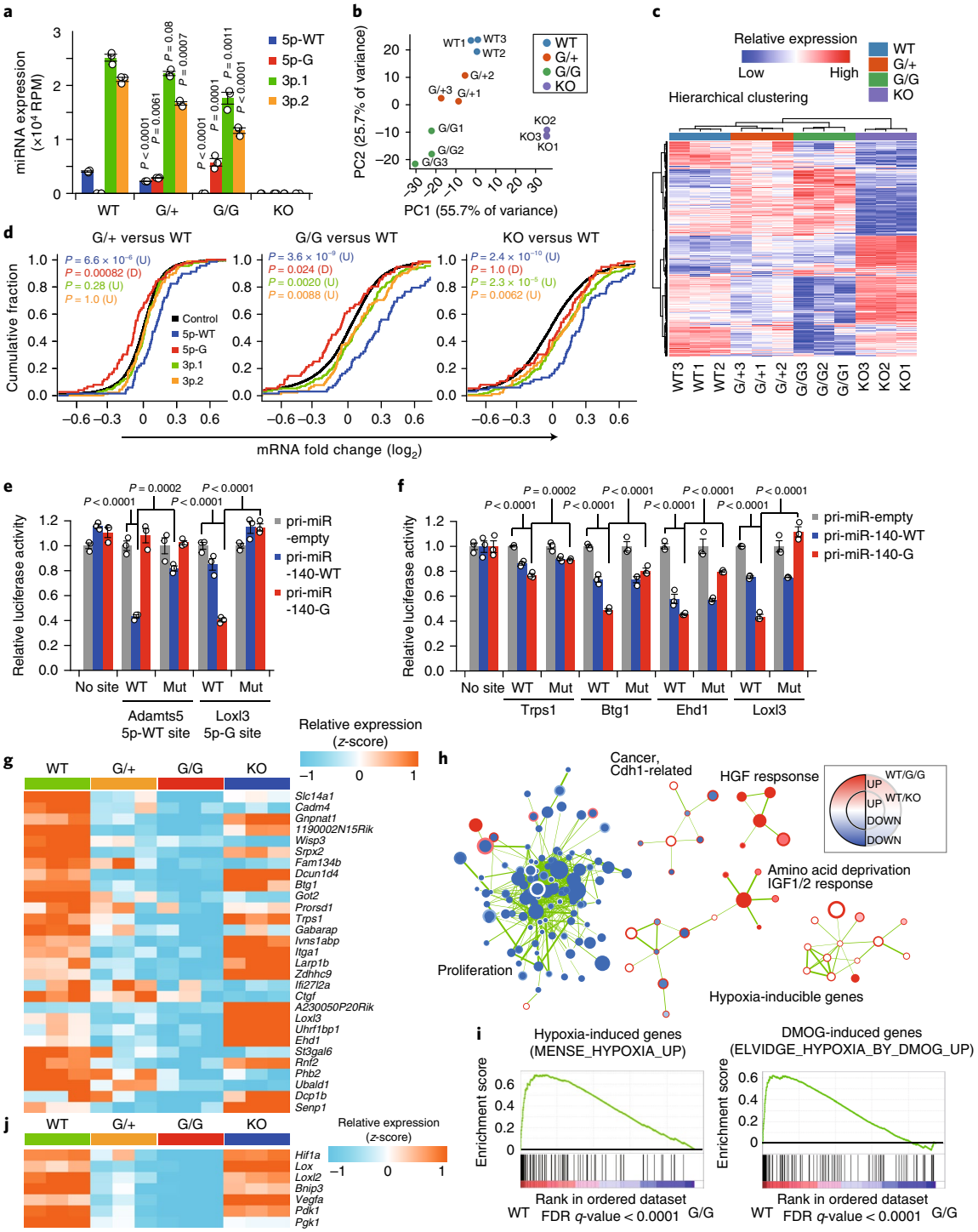


Fig. 3 | Loss-of-function and gain-of-function effects of miR-140-G mutation in chondrocyte transcriptome. a, Expression levels of miR-140-5p (WT and G mutant) and two species of miR-140-3p (3p.1 and 3p.2) in primary rib chondrocytes. Results represent mean ± s.e.m. ($n = 3$ animals). Statistical significance was assessed using one-way ANOVA and post hoc Tukey–Kramer test. *P* values indicate comparison versus WT. KO, knockout. **b**,**c**, Principal component (PC) analysis (**b**) and hierarchical clustering (**c**) based on 1,000 most variable genes in RNA-seq datasets ($n = 3$ animals for each group). **d**, Cumulative distributions of fold changes of mRNAs with conserved 8mer binding sites for miR-140-5p-WT, miR-140-5p-G, miR-140-3p.1, and miR-140-3p.2 in miR-140 mutant and null chondrocytes. *P* values (versus a control gene set) were calculated by one-sided Kolmogorov–Smirnov test for either direction: upregulation (U) or downregulation (D) with Bonferroni correction (control genes, $n = 11{,}472$; miR-140-5p-WT, $n = 68$; miR-140-5p-G, $n = 76$; miR-140-3p.1, $n = 173$; miR-140-3p.2, $n = 91$). **e**,**f**, Target verification of miR-140-5p-G target sites by 3′ UTR reporter assay in the context of minimal target sites (**e**) and the endogenous 3′ UTR context (**f**). Results represent mean ± s.e.m. ($n = 3$ biologically independent samples). Statistical significance was assessed using two-way ANOVA and post hoc Tukey–Kramer test. **g**, Heat map showing expression changes of representative genes with conserved miR-140-5p-G target sites and downregulation in miR-140$^{G/G}$ chondrocytes. Red denotes important regulators of skeletal development. **h**–**j**, Enrichment map analysis on GSEA results (false discovery rate *q* value < 0.1) showing affected cellular programs in miR-140$^{G/G}$ or miR-140-null chondrocytes (**h**), representative GSEA results for hypoxia-induced gene sets (**i**, DMOG: hypoxia mimic), and expression heat map for major Hif1 target genes (**j**). In **h**, enrichments in WT versus miR-140$^{G/G}$ and WT versus miR-140-null cells are mapped to the node borders and inner node area, respectively. Red and blue represent upregulation and downregulation in WT cells relative to miR-140$^{G/G}$ or miR-140-null cells, respectively. Pathway enrichment analysis and statistical analysis were performed using GSEA.

miR-140$^{G/G}$ and miR-140$^{G/+}$ chondrocytes (Extended Data Fig. 7f), as confirmed by in situ hybridization analysis for *Col10a1* (Extended Data Fig. 3b).

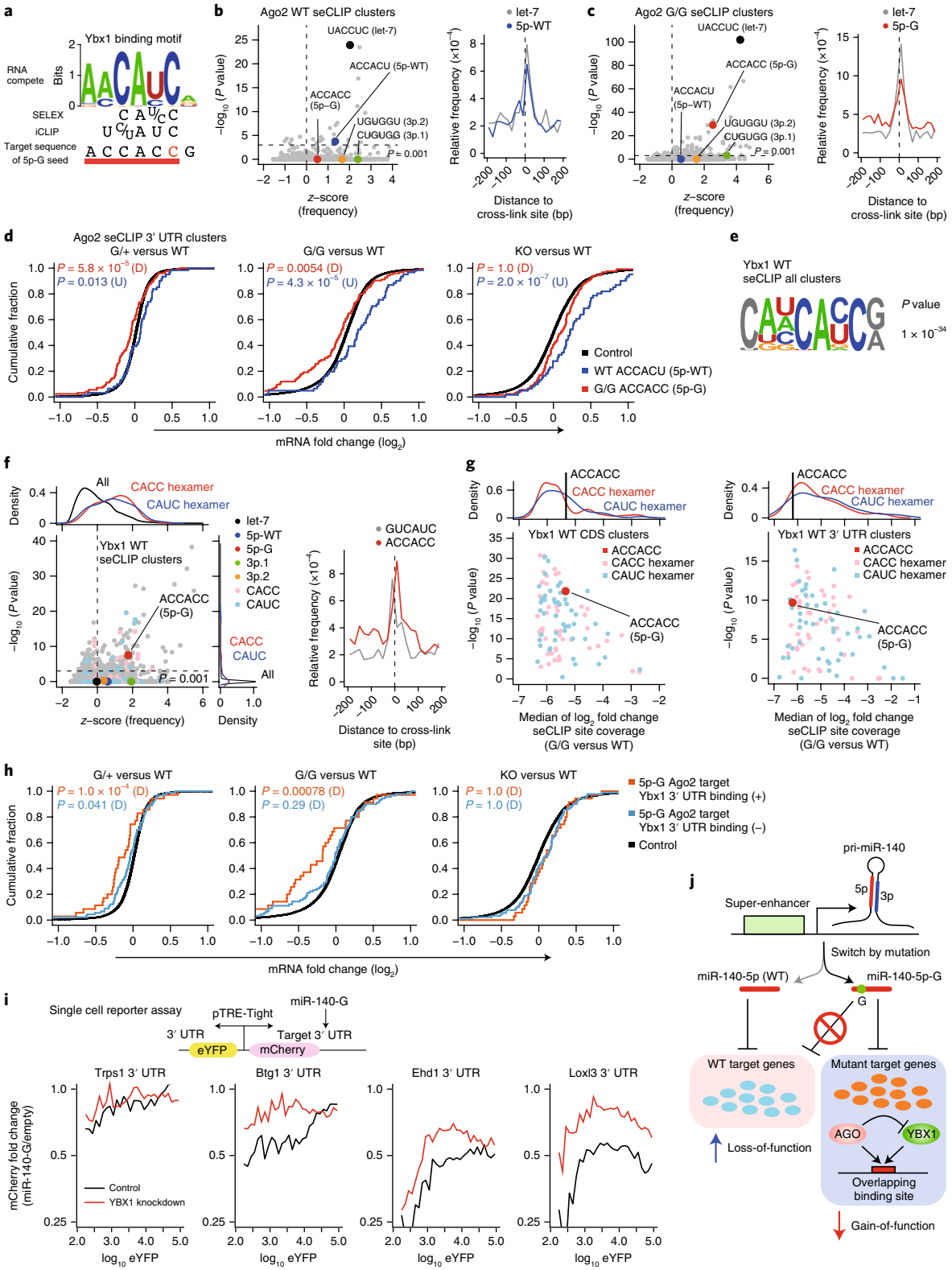
Gene regulation by conserved miRNAs has a significant impact on 3′ UTR evolution, and evolutionarily maintained target sites mediate actions of miRNAs more effectively than non-conserved target sites[28]. In light of this miRNA–target coevolution concept[7,8], an intriguing question is how 'newly emerged' miR-140-5p-G exerts such potent target repression and robust in vivo effects. Since a mutation in the miRNA seed region creates a new repertoire of target genes and the magnitude of such target suppression is typically small, a neomorphic mutant miRNA would need to target a biologically important regulatory network to produce a disease phenotype. Because the seed sequence of the mutant miR-140-5p-G is not shared with other known miRNAs and multiple RNA-binding proteins (RBPs) are reported to modulate miRNA function[29], we hypothesized that miR-140-5p-G conversely interferes with conserved RBP pathways. Computational analysis suggested that the miR-140-5p-G seed in the Ago2 complex and the RBP Ybx1 (also known as YB-1) potentially compete for the same RNA-binding sites (Fig. 4a and Extended Data Fig. 8a)[30–32]. This hypothesis is supported by the binding of YBX1 to several miR-140-5p-G target genes in a previously reported YBX1 individual nucleotide resolution cross-linking and immunoprecipitation (iCLIP) dataset (Extended Data Fig. 8b)[32]. YBX1 is a multifunctional RBP that stabilizes mRNAs and regulates RNA splicing[33,34]. We performed single-end enhanced CLIP (seCLIP) analysis of Ybx1 and Ago2, a central component of the miRNA pathway, in WT and miR-140$^{G/G}$ chondrocytes to test this hypothesis (Extended Data Fig. 8c–e)[35,36]. Ago2 seCLIP analysis revealed cross-link site-centric enrichment of hexamer sequences complementary to the seed sequences of WT miR-140-5p and miR-140-5p-G in Ago2 seCLIP clusters in WT and miR-140$^{G/G}$ chondrocytes, respectively (Fig. 4b,c and Extended Data Fig. 9a,b). Accordingly, genes with the identified 3′ UTR target sites of miR-140-5p-G were downregulated only in miR-140$^{G/G}$ and miR-140$^{G/+}$ chondrocytes, while genes with WT miR-140-5p target sites were upregulated in both miR-140-null and miR-140 mutant chondrocytes (Fig. 4d). This effect was not apparent for the target sites in coding sequences (CDS) (Extended Data Fig. 9c). Next, Ybx1 seCLIP analysis identified a binding motif similar to the Ybx1 binding motif reported previously[30–32] (Fig. 4e). It also demonstrated cross-link site-centric enrichment of a hexamer sequence ('ACCACC') complementary to the seed of miR-140-5p-G (Fig. 4f and Extended Data Fig. 10a,b). Analysis of Ybx1 seCLIP signals at the seCLIP clusters identified in WT chondrocytes showed that the miR-140 mutation preferentially attenuated seCLIP signals at the 3′ UTR sites with the ACCACC hexamer among the representative Ybx1 target sequences, CAUC/CACC hexamers (Fig. 4g). This trend was not observed for the CDS seCLIP clusters (Fig. 4g), suggesting a predominant competition between miR-140-5p-G and Ybx1 in 3′ UTRs. Notably, miR-140-5p-G Ago2 seCLIP target genes with Ybx1 binding in the 3′ UTR showed stronger repression in mutant chondrocytes than those without Ybx1 binding (Fig. 4h). Furthermore, a single-cell 3′ UTR reporter assay of several miR-140-5p-G target genes demonstrated that suppressive effects by miR-140-5p-G were dampened by Ybx1 inhibition (Fig. 4i and Extended Data Fig. 10c,d). In addition, previously reported YBX1-stabilized transcripts were preferentially repressed in miR-140$^{G/G}$ and miR-140$^{G/+}$ chondrocytes compared with miR-140$^{-/-}$ chondrocytes (Extended Data Fig. 10e)[34]. Furthermore, systematic RBP motif analysis suggested that Ybx1 activity was selectively suppressed in miR-140$^{G/G}$ but not in miR-140$^{-/-}$ chondrocytes (Extended Data Fig. 10f,g). These observations collectively suggest that the potent suppressive effect of miR-140-5p-G is partly attributable to the competition against Ybx1, especially in 3′ UTRs, and subsequent suppression of Ybx1 activity.

Besides gene dosage abnormalities[37,38], single nucleotide substitutions in two miRNA genes have been identified in human diseases: *MIR96* mutations in autosomal dominant deafness 50 (OMIM #613704) and *MIR184* mutations in endothelial dystrophy, iris hypoplasia, congenital cataract, and stromal thinning (EDICT) syndrome (OMIM #614303) and familial and sporadic cases of keratoconus[1–5,39]. However, these mutations, occurring at multiple different nucleotides of single miRNAs, are thought to primarily cause loss of function (Supplementary Table 8). In contrast, the miR-140 mutation found in this study affected the same nucleotide in both families and resulted in gain of a new function as well as loss of the original function, as shown by the transcriptome analysis revealing repression of new target genes and derepression of original targets. Although it is unusual that a 'newly emerged' miRNA exerts such a potent biological effect, in this instance, the disease phenotype may be at least partially explained by potential competition with the RBP, Ybx1, and by the super-enhancer-driven high-level miRNA expression (Fig. 4j). This finding suggests that sequence-dependent cross-talks with RBPs may confer extended biological properties to miRNAs and other types of regulatory RNAs. Such sequence-dependent competition between a miRNA seed and an RBP would add a new mechanism of transcript regulation. Our subsequent systematic analysis of siRNA off-target data and genome-scale RNAi-based loss-of-function screens for essential genes in cancer cells has recently confirmed and reinforced this concept (cross-talk with endogenous RBPs; ceRBP) in general[40]. We observed that

Fig. 4 | Competition between miR-140-5p-G and Ybx1. a, Overlap between previously reported Ybx1 binding motifs and the complementary sequence of miR-140-5p-G seed region. **b**,**c**, Hexamer enrichment analysis in non-intronic Ago2-bound seCLIP clusters identified in WT (**b**) and mutant (**c**) chondrocytes (WT, $n=4,469$ clusters; G/G, $n=6,663$ clusters). The *x* axis shows *z*-scores of hexamer frequency. The *y* axis shows negative $\log_{10}$ *P* values of two-sided Fisher's exact test with Bonferroni correction of the hexamer enrichment above background. Right panels show positional enrichment of indicated motifs around the cross-link site. **d**, Cumulative distributions of fold changes of mRNAs with Ago2-bound seCLIP 3′ UTR clusters with ACCACC motif in mutant chondrocytes (red) or Ago2-bound seCLIP 3′ UTR clusters with ACCACU motif in WT chondrocytes (blue). *P* values (versus a control gene set) were calculated by one-sided Kolmogorov–Smirnov test for either direction: upregulation (U) or downregulation (D) with Bonferroni correction (control genes, $n=9,428$; miR-140-5p-WT, $n=61$; miR-140-5p-G, $n=125$). **e**, A representative motif enriched in all Ybx1-bound seCLIP clusters in WT chondrocytes ($n=8,354$). Statistical analysis was performed using HOMER. **f**, Hexamer enrichment analysis in non-intronic Ybx1-bound seCLIP clusters identified in WT chondrocytes ($n=5,202$). The graphs were plotted as in **b**. **g**, Comparison of seCLIP signals at Ybx1-bound CDS seCLIP clusters (left, $n=2,537$) and Ybx1-bound 3′ UTR seCLIP clusters (right, $n=1,649$) containing individual CAUC or CACC motif, identified in WT chondrocytes, between the miR-140$^{G/G}$ and WT chondrocytes. The *x* axis shows the median of $\log_2$ fold changes of seCLIP signals between mutant and WT samples. The *y* axis shows negative $\log_{10}$ *P* values of two-sided Wilcoxon rank sum test with Bonferroni correction. **h**, Cumulative distributions of fold changes of expression levels of miR-140-5p-G target genes, identified by the Ago2 seCLIP, with or without Ybx1 binding in 3′ UTRs (control genes, $n=9,428$; miR-140-5p-G seCLIP target genes with Ybx1 binding, $n=35$; miR-140-5p-G seCLIP target genes without Ybx1 binding, $n=90$). Graphs are drawn as in **d**. **i**, Single-cell 3′ UTR reporter assay in HEK293T cells transfected with Ybx siRNA, eYFP and mCherry dual reporter, and pri-miRNA expression plasmids. Fold changes of mean mCherry signal by miR-140-G overexpression across eYFP signal bins are shown. Suppressive effects by miR-140-G were dampened by Ybx1 knockdown. **j**, Summary of the present study.

ceRBP effects are widespread in RNAi and contributes to off-target effects and phenotype modulation[40]. Therefore, avoiding RBP binding motif overlap with seed sequences may further minimize RNAi off-target effects.

Overall, the present study expands our knowledge on how mutations in miRNAs contribute to human disease development and thus suggests the possibility that some of the currently unresolved congenital disorders may be caused by neomorphic miRNA mutations. In addition, given the expansion of miRNA genes in vertebrates[7,8], the molecular mechanism revealed in this present study also has broader implications in understanding how function of new miRNA genes might evolve in humans.



Online content

Any methods, additional references, Nature Research reporting summaries, source data, statements of data availability and associated accession codes are available at https://doi.org/10.1038/s41591-019-0353-2.

Received: 4 December 2017; Accepted: 11 January 2019;
Published online: 25 February 2019

References

1. Mencia, A. et al. Mutations in the seed region of human miR-96 are responsible for nonsyndromic progressive hearing loss. *Nat. Genet.* **41**, 609–613 (2009).
2. Solda, G. et al. A novel mutation within the *MIR96* gene causes non-syndromic inherited hearing loss in an Italian family by altering pre-miRNA processing. *Hum. Mol. Genet.* **21**, 577–585 (2012).
3. Hughes, A. E. et al. Mutation altering the miR-184 seed region causes familial keratoconus with cataract. *Am. J. Hum. Genet.* **89**, 628–633 (2011).
4. Iliff, B. W., Riazuddin, S. A. & Gottsch, J. D. A single-base substitution in the seed region of miR-184 causes EDICT syndrome. *Invest. Ophthalmol. Vis. Sci.* **53**, 348–353 (2012).
5. Lechner, J. et al. Mutational analysis of *MIR184* in sporadic keratoconus and myopia. *Invest. Ophthalmol. Vis. Sci.* **54**, 5266–5272 (2013).
6. Nakamura, Y., Inloes, J. B., Katagiri, T. & Kobayashi, T. Chondrocyte-specific microRNA-140 regulates endochondral bone development and targets *Dnpep* to modulate bone morphogenetic protein signaling. *Mol. Cell. Biol.* **31**, 3019–3028 (2011).
7. Chen, K. & Rajewsky, N. The evolution of gene regulation by transcription factors and microRNAs. *Nat. Rev. Genet.* **8**, 93–103 (2007).
8. Berezikov, E. Evolution of microRNA diversity and regulation in animals. *Nat. Rev. Genet.* **12**, 846–860 (2011).
9. Bonafe, L. et al. Nosology and classification of genetic skeletal disorders: 2015 revision. *Am. J. Med. Genet. A* **167A**, 2869–2892 (2015).
10. Lek, M. et al. Analysis of protein-coding genetic variation in 60,706 humans. *Nature* **536**, 285–291 (2016).
11. Suzuki, H. I., Young, R. A. & Sharp, P. A. Super-enhancer-mediated RNA processing revealed by integrative microrna network analysis. *Cell* **168**, 1000–1014 e1015 (2017).
12. Whyte, W. A. et al. Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell* **153**, 307–319 (2013).
13. Hnisz, D. et al. Super-enhancers in the control of cell identity and disease. *Cell* **155**, 934–947 (2013).
14. Wienholds, E. et al. MicroRNA expression in zebrafish embryonic development. *Science* **309**, 310–311 (2005).
15. Tuddenham, L. et al. The cartilage specific microRNA-140 targets histone deacetylase 4 in mouse cells. *FEBS Lett.* **580**, 4214–4217 (2006).
16. Miyaki, S. et al. MicroRNA-140 plays dual roles in both cartilage development and homeostasis. *Genes Dev.* **24**, 1173–1185 (2010).
17. Papaioannou, G. et al. MicroRNA-140 provides robustness to the regulation of hypertrophic chondrocyte differentiation by the PTHrP-HDAC4 pathway. *J. Bone Miner. Res.* **30**, 1044–1052 (2015).
18. Ran, F. A. et al. Genome engineering using the CRISPR–Cas9 system. *Nat. Protoc.* **8**, 2281–2308 (2013).
19. Park, J. E. et al. Dicer recognizes the 5′ end of RNA for efficient and accurate processing. *Nature* **475**, 201–205 (2011).
20. Suzuki, H. I. et al. Small-RNA asymmetry is directly driven by mammalian Argonautes. *Nat. Struct. Mol. Biol.* **22**, 512–521 (2015).
21. Agarwal, V., Bell, G. W., Nam, J. W. & Bartel, D. P. Predicting effective microRNA target sites in mammalian mRNAs. *eLife* **4**, e05005 (2015).
22. Zhang, J. et al. Loss of lysyl oxidase-like 3 causes cleft palate and spinal deformity in mice. *Hum. Mol. Genet.* **24**, 6174–6185 (2015).
23. Tijchon, E. et al. Targeted deletion of btg1 and btg2 results in homeotic transformation of the axial skeleton. *PLoS ONE* **10**, e0131481 (2015).
24. Napierala, D. et al. Uncoupling of chondrocyte differentiation and perichondrial mineralization underlies the skeletal dysplasia in tricho-rhino-phalangeal syndrome. *Hum. Mol. Genet.* **17**, 2244–2254 (2008).
25. Subramanian, A. et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl Acad. Sci. USA* **102**, 15545–15550 (2005).
26. Merico, D., Isserlin, R., Stueker, O., Emili, A. & Bader, G. D. Enrichment map: a network-based method for gene-set enrichment visualization and interpretation. *PLoS ONE* **5**, e13984 (2010).
27. Wan, C. et al. Role of HIF-1α in skeletal development. *Ann. N. Y. Acad. Sci.* **1192**, 322–326 (2010).
28. Grimson, A. et al. MicroRNA targeting specificity in mammals: determinants beyond seed pairing. *Mol. Cell* **27**, 91–105 (2007).
29. van Kouwenhove, M., Kedde, M. & Agami, R. MicroRNA regulation by RNA-binding proteins and its implications for cancer. *Nat. Rev. Cancer* **11**, 644–656 (2011).
30. Ray, D. et al. A compendium of RNA-binding motifs for decoding gene regulation. *Nature* **499**, 172–177 (2013).
31. Wei, W. J. et al. YB-1 binds to CAUC motifs and stimulates exon inclusion by enhancing the recruitment of U2AF to weak polypyrimidine tracts. *Nucleic Acids Res.* **40**, 8622–8636 (2012).
32. Wu, S. L. et al. Genome-wide analysis of YB-1-RNA interactions reveals a novel role of YB-1 in miRNA processing in glioblastoma multiforme. *Nucleic Acids Res.* **43**, 8516–8528 (2015).
33. Lyabin, D. N., Eliseeva, I. A. & Ovchinnikov, L. P. YB-1 protein: functions and regulation. *Wiley Interdiscip. Rev. RNA* **5**, 95–110 (2014).
34. Goodarzi, H. et al. Endogenous tRNA-derived fragments suppress breast cancer progression via ybx1 displacement. *Cell* **161**, 790–802 (2015).
35. Van Nostrand, E. L. et al. Robust transcriptome-wide discovery of RNA-binding protein binding sites with enhanced CLIP (eCLIP). *Nat. Methods* **13**, 508–514 (2016).
36. Van Nostrand, E. L. et al. Erratum to: robust, cost-effective profiling of RNA binding protein targets with single-end enhanced crosslinking and immunoprecipitation (seCLIP). *Methods Mol. Biol.* **1648**, E1 (2017).
37. de Pontual, L. et al. Germline deletion of the miR-17 approximately 92 cluster causes skeletal and growth defects in humans. *Nat. Genet.* **43**, 1026–1030 (2011).
38. Henrion-Caude, A., Girard, M. & Amiel, J. MicroRNAs in genetic disease: rethinking the dosage. *Curr. Gene. Ther.* **12**, 292–300 (2012).
39. Lewis, M. A. et al. An ENU-induced mutation of miR-96 associated with progressive hearing loss in mice. *Nat. Genet.* **41**, 614–618 (2009).
40. Suzuki, H. I., Spengler, R. M., Grigelioniene, G., Kobayashi, T. & Sharp, P. A. Deconvolution of seed and RNA-binding protein crosstalk in RNAi-based functional genomics. *Nat. Genet.* **50**, 657–661 (2018).

Acknowledgements

We thank patients and their family members for their participation in the study, M. Mannstadt, K. Leuter, M. Wein, H. Kronenberg, H. Jueppner, G. Björk, and A. Merker for advice and chemicals, J. Lundin for assistance with the human comparative genomic hybridization array databases, and Science for Life Laboratory, the National Genomics Infrastructure (NGI), Sweden, SNIC through UPPMAX under project b2014231 for providing assistance in massive parallel DNA sequencing and computational infrastructure. We also thank Center for Skeletal Research Core (NIH P30 AR066261) for access to histological analysis equipment; Harvard GMF for generation of knock-in mice; and DFCI and MGH Sequencing Cores for assistance with RNA and Sanger sequencing; the Robert A. Swanson (1969) Biotechnology Center at the Koch Institute for Integrative Cancer Research at Massachusetts Institute of Technology for technical support, specifically S. Levine and the staff of the BioMicro Center/KI Genomic Core Facility and G. Paradis, M. Jennings, and M. Saturno-Condon of the Flow Cytometry Core Facility. This project was supported by grants provided by the Stockholm County Council (ALF projects 20150143 and 20130315 to G. Grigelioniene. and A.N.) and by the National Institutes of Health (National Institute of Arthritis and Musculoskeletal and Skin Diseases grant R01-AR056645 to T.K., National Institute of General Medical Sciences grant R01-GM034277 and National Cancer Institute grant R01-CA133404 to P.A.S., and National Cancer Institute grant P30-CA14051 to the Koch Institute Core Facility). G.Grigelioniene was supported by a grant of the Sabbatical Leave Programme of the European Society for Paediatric Endocrinology through an educational grant from Eli Lilly International Corporation, travel grants from Fernström Foundation, Karolinska Institutet and Swedish Society of Medicine, scholarship from Stiftelsen Samariten, Stockholm, Sweden, Research Funds from Promobilia and Frimurare Barnhuset Stockholm, and project grant from Swedish Research Council 2018-03046. H.I.S. is supported by the Uehara Memorial Foundation Research Fellowship and the Osamu Hayaishi Memorial Scholarship for Study Abroad.

Author contributions

G.Grigelioniene, H.I.S., and T.K. designed the study and wrote the manuscript. G.Grigelioniene, H.I.S., T.K., F.T., F.M., U.M.A., S.T., A.L., M.A.W., G.Grigelionis, A.H., D.R.E., M.L.W., and P.A.S. conducted experiments and data collection. E.M., A.N., M.N., G.N., Z.U.B., and E.H. performed clinical characterization. All authors contributed to data interpretation and revised the manuscript.

Competing interests

S.T. works at Emedgene Technologies, and P.A.S. is a board member of Syros Pharmaceuticals.

Additional information

Extended data is available for this paper at https://doi.org/10.1038/s41591-019-0353-2.

Supplementary information is available for this paper at https://doi.org/10.1038/s41591-019-0353-2.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to T.K.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Methods

Ethical approvals. The study was approved by the regional ethical review board in Stockholm (2014/983-31/1, 2012/2106-31/4) and Israel Research Board no. 0072-13-BNZ. Animal studies were approved by the Institutional Animal Care and Use Committee (IACUC) of the Massachusetts General Hospital (2009N000217). Written informed consent for research participants was obtained from the patients and from the parents of the involved minors. P1 (currently 45 years old) has kindly provided her childhood photo for publication and the authors have obtained signed proof of informed consent. The patients were examined at the Karolinska University Hospital, Stockholm, Sweden (family 1, P1 and P2) and at the Assuta Medical Center, Haifa, Israel (family 2, P3) due to unknown ultrarare skeletal dysplasia.

Clinical diagnosis of spondyloepiphyseal dysplasia *MIR140* type Nishimura. One of the authors (G.N.) recognized radiographic similarities between the unrelated individuals, diagnosed the disorder in the first family (P1 and P2) as a novel skeletal dysplasia, and then diagnosed P3 as having the same disorder. Because the phenotype has not been reported previously and these patients were found to have the same variant in the *MIR140* gene, we propose the eponym of spondyloepiphyseal dysplasia (SED) *MIR140* type Nishimura. The present disorder should be differentiated from acrodysostosis. Both disorders share several features, including midface hypoplasia and brachydactyly with cone-shaped epiphyses. However, epiphyseal maturation is delayed in SED *MIR140* type Nishimura, whereas epiphyseal ossification, particularly carpal ossification, is advanced in acrodysostosis. In addition, epiphyseal dysplasia, a hallmark in SED *MIR140* type Nishimura, is not a feature in acrodysostosis. Furthermore, the presented patients did not have mutations in the *PDE4D* and *PRKAR1A* genes known to be associated with acrodysostosis. They did not show any of the endocrine abnormalities seen in acrodysostosis, such as hypocalcemia, hyperphosphatemia, elevated levels of parathyroid hormone and/or thyroid stimulating hormone, and decreased insulin-like growth factor 1. P1 had measurement of bone density by dual-energy X-ray absorptiometry at the age of 43. She had normal bone density for the age, which suggests that the heterozygous *MIR140* mutation in humans does not lead to bone mass abnormalities.

Identification of the *MIR140* mutation. Exome sequencing analysis was initially performed for both families in quartet for family 1 (the proband, P1, her unaffected parents, and her affected son, P2) and in trio for family 2 (the proband, P3, and her unaffected parents). Since the disease was inherited in an autosomal dominant way in family 1, we searched for de novo variants in P1. For family 2, exome sequencing was initially performed independently from family 1, searching for de novo or autosomal recessively inherited mutations. No potential disease-causing variants were found in the protein-coding genes.

Next, WGS was performed to search for variants in both protein-coding and protein-non-coding genes for family 1 in quartet. A single nucleotide substitution in chr16: 69967007A>G in the *MIR140* gene (+24A>G) in P1 (II-2) and her affected son P2 (III-2) was detected and confirmed by Sanger sequencing (Fig. 1q and Extended Data Fig. 1d). This single nucleotide substitution occurred de novo in P1 and was transmitted to her son. This variant segregated with the skeletal dysplasia phenotype in family 1. The same de novo single nucleotide substitution was detected using Sanger sequencing in P3, and the mutation was retrospectively confirmed in exome sequencing data of family 2 (Extended Data Fig. 1d).

Screening P1, P2, and P3 for possible disease-causing variants in 378 known skeletal dysplasia genes using WGS data, and screening P2 for gene dose abnormalities of the above-mentioned skeletal dysplasia genes using custom-designed comparative genome hybridization analysis was all negative (data not shown), excluding involvement of these genes in the skeletal dysplasia phenotype of our patients.

An extended search in copy number variation (CNV) database regarding human phenotypes associated with CNVs involving the *MIR140* gene revealed a single individual with heterozygous 1.1 Mb deletion with isolated Arnold–Chiari malformation[41]. No individuals with short stature and heterozygous deletions of greater than 2 Mb size including the *MIR140* gene were reported in the CNV databases (https://decipher.sanger.ac.uk/). This result suggests that haploinsufficiency of *MIR140* does not lead to skeletal abnormalities in humans, consistent with the normal phenotype previously reported in heterozygous miR-140 knockout mice[6], supporting the hypothesis that the heterozygous nucleotide substitution (chr16:g.69967007A>G (hg19), *MIR140*:NR_029681.1:n.24A>G) is a neomorphic mutation.

Clinical data. The information regarding the symptoms and features was collected from the patient records. The skeletal dysplasia was identified as a unique entity by an expert pediatric radiologist (G.N.) in the first family and was recognized by the same radiologist in the second patient. In auxology analysis, the *z*-scores at birth and at follow up were calculated using the Swedish growth reference data[42] and WHO (World Health Organization) auxology calculator (http://www.childgrowthcalculator.com/).

DNA extraction and analyses. DNA was extracted from peripheral venous blood samples (2–10 ml) from the patients and their family members. For DNA extraction from the samples of family 1, we used the Qiagen Puregen Blood Core kit C (Qiagen). A phenol-chloroform DNA extraction protocol was applied for DNA extraction from members of family 2.

The following methods were applied for DNA examination of family 1: whole exome sequencing and WGS for I-1, I-2, II-2, and III-2, comparative genomic hybridization-array for III-2, and Sanger sequencing of *MIR140* for all family members; and for family 2: whole exome sequencing for I-1, I-2, and II-1, WGS for II-1, and Sanger sequencing for *MIR140* for all family members.

Exome sequencing. Libraries for sequencing on Illumina HiSeq2500 (Illumina) were prepared from genomic DNA. Exome sequences were enriched with the Agilent SureSelect XT All Exon V5 target enrichment kit (Agilent Technologies), according to the manufacturer's instruction. On average, this resulted in >64 million mapped unique sequences with a mean coverage of 135, that is, 10× coverage for 97.9% of target sequences. Post-capture libraries were sequenced as 2×101 base pair (bp) paired-end reads, which were base-called using CASAVA (bcl2Fastq v1.8.3, Illumina). For bioinformatic analysis, the following steps were performed as described previously[43]. Reads were mapped to the human reference genome (hg19) using MOSAIK (2.2.3). Duplicates were marked with Picard 1.92. Variants were called using the Genome Analysis Tool (GATK v.3.2.0), annotated using both ANNOVAR (version 2014 July 14) and Variant Effect Predictor (VEP), and loaded into the GEMINI (v0.16.0) database. Variants that were rare, with minor allele frequency less than 1% in the 1000 Genomes Project (2014 Oct version) all populations, 6500 NHLBI-GO Exome Sequencing Project (EVS), and Exome Aggregation Consortium (ExAC v0.2)[10], were considered for further analysis. The Single Nucleotide Polymorphism Database (dbSNP) build 138 (dbSNP138) nonflagged was used to annotate known single-nucleotide polymorphisms. Combined annotation dependent depletion (CADD) was used to score the deleteriousness of single nucleotide variants.

Whole-genome sequencing. Libraries for sequencing on Illumina HiSeqX Ten (Illumina) were prepared from the genomic DNA using the Illumina TruSeq PCR-free kit with a mean insert size of >350 bp. On average, this resulted in >620 million mapped unique sequences with a mean coverage of >29, that is, 19× coverage for 80% of reference sequences. An in-house pipeline developed by Science For Life Laboratory, Stockholm, Sweden was used to map reads to the human reference genome (hg19). Data were aligned to the reference genome using bwa (v0.7.12)[44]. The raw alignments were then deduplicated, recalibrated, and cleaned using GATK (v3.3-0-geee94ec)[45]. The quality control information was gathered using Qualimap (v2.0)[46]. Single nucleotide variants and indels were called using the HaplotypeCaller in GATK. The variants were further analyzed using the following steps as described previously[43]. The variants were processed with GenotypeGVCFs, VariantRecalibrator, ApplyRecalibration, and SelectVariants tools in GATK, then functionally annotated using VEP, and loaded into a database using GEMINI (v0.16.0). The variants were explored in the database using built-in tools in GEMINI.

Comparative genomic hybridization array. CNV screening using a custom-designed exon targeted array. Comparative genomic hybridization was performed using the DNA sample of P2 targeting 872 genes involved in skeletal dysplasias, ciliary disorders, and malformation syndromes (Agilent Technologies). The custom array design was created using e-Array (https://earray.chem.agilent.com/earray/), Agilent's online resource for custom array design. Briefly, we used a 2×4,000k array format to create a design with an average coverage of one probe per 100 bp in coding sequences and one probe per 500 bp in non-coding sequences in the target genes and at the same time to retain a genome-wide background resolution of 50 kb across the genome. Experiments were performed according to the manufacturer's recommendation with minor modifications as described previously[47]. The copy number data analysis was performed using the Agilent Genomic Workbench software (Agilent Technologies).

PCR, Sanger sequencing, and genotyping. To confirm the WGS and exome sequencing findings, we used PCR primers: MIR140-F 5′-CCGTGGATGGATGTTCCTTTTT-3′ and MIR140-R 5′-GGCTTGAGCTAAACCAGCAAG-3′ to amplify and sequence the *MIR140* gene. The primers were used to amplify 281 bp fragments of the *MIR140* locus of the genomic DNA from the patients and their family members. The obtained PCR products were purified with ExoSAP-IT (Thermo Fisher Scientific) and sequenced at the MGH DNA core facility.

Mice. The animal study was approved by the Institutional Animal Care and Use Committee of the Massachusetts General Hospital and performed in accordance with the regulations and guidelines. miR-140-null mice were previously described[6]. For generation of knock-in mice with the A>G substitution in the *Mir140* gene, we used in vivo CRISPR–Cas9 editing in mouse zygotes according to an established protocol[48]. Wild-type Cas9 mRNA, a guide RNA with the Cas9 handle, and a single-stranded oligomer DNA repair template were injected into one-cell stage C57/B6 mouse embryos. The sequence of the guide RNA is 5′-CATAGGGTAAAACCACTGGC-3′. For the repair template, a 140-base-long

single-stranded oligomer DNA with the A>G single nucleotide substitution (corresponding to the mutation) at the 70th nucleotide from the 5′-end was synthesized. This substitution creates a novel restriction enzyme site (*Msp* I; CCGG). Founder mice were screened by PCR followed by *Msp* I digestion. For genotyping, a 216 nt genomic sequence containing the mutation site was amplified using the PCR primers, Mir140-F 5′-TCTGTGTTCATCCCATCCTG-3′ and Mir140-R 5′-ATGGAGTCCTTCGATGCAGA-3′ and digested with *Msp* I. The PCR amplicon containing the A>G substitution is digested into 98-bp- and 119-bp-long fragments. Out of 61 mice embryos injected with the guide RNA, repair template and Cas9, 11 mice were positive for the knock-in mutation. Among them, 5 mice were selected for further characterization of the genomic modification. All of these mice were compound heterozygous with one allele mutated as desired and the other with various size insertion–deletion mutations. After confirming the absence of additional mutations in the 2 kb region surrounding the *Mir140* gene using Sanger sequencing, we selected two founder lines, crossed them to WT C57/B6 mice to breed out the opposite allele to generate heterozygous F1 lines (miR-140$^{G/+}$). F1 lines were further intercrossed to generate homozygous mutants (miR-140$^{G/G}$). The consistency of the phenotype was confirmed in two lines. The sample size was estimated based on our previous experience of performing similar sets of experiments. The number of examined mice is shown in Supplementary Table 2.

Skeletal preparation and histology. For skeletal examination, one side of the extremities was taken for whole-mount staining with Alizarin red with or without Alcian blue staining, and the contralateral extremities were fixed in formalin, decalcified, embedded in paraffin, sectioned, and stained with hematoxilin and eosin or safranin O for microscopic examination for one-, two-, three-, four-, and eight-week-old mice. The number of examined mice is summarized in Supplementary Table 2. Tibias of two-week-old mice and skulls of four-week-old animals were examined by micro-computed tomography analysis to determine the formation of the secondary ossification and details in skull anatomy. For two-week-old animals, bones were also stained with Alizarin red. The growth plate in the basal skull of seven-day-old mice was sectioned, stained with hematoxylin and eosin according to the standard protocol, and histologically examined. As we found no phenotypic differences between males and females, males and females were not discriminated for the analysis. Lengths of the growth plate and its zones were measured using the image processor, Fiji. Lengths of total and different zones of tibial growth plates at P56 were also measured. Tibia growth plates are irregular, thus the measurements were performed by selecting a rectangle area in the middle flattest part of each growth plate section, averaging the height of each zone in the middle and at the edges. The trachea and larynx of 56-day-old animals were examined by whole-mount staining followed by histological analysis.

Primary rib chondrocyte culture. Primary rib chondrocytes were isolated from the frontal part of the rib cage of seven-day-old mice. Muscle, soft tissues, and mineralized rib and sternal bones were removed using fine tweezers and surgical scissors under a dissection microscope. Dissected costal cartilage rods were then incubated in a digestion medium containing DMEM medium, 10% FCS, and 0.2% (approximately 600 U ml^{-1}) collagenase type II (Worthington) at 37 °C for 6 h. Cells were released from the cartilage matrix by gentle pipetting, passed through nylon mesh strainers (BD Falcon), spun to remove collagenase, counted, plated at the concentration of 1×10^5 cells ml^{-1} in the DMEM medium containing 10% FCS and cultured overnight.

Total RNA, including miRNAs, was isolated using the Direct-zol RNA MiniPrep kit (Zymo Research). For western blot analysis, cells were lysed in 2× Laemmli sample buffer. Proteins were resolved in 4–20% gradient acrylamide/MOPS gels, transferred on nitrocellulose membrane and incubated with indicated antibodies (anti-Hif-1a antibody (NB100-449) from Novus Biologicals, anti-beta actin antibody (no. 4970S) and anti-Ybx1 antibody (no. 9744S) from Cell Signaling Technology). Uncropped blots are shown in Source Data.

BrdU labeling and detection. For BrdU labeling, 50 µg of BrdU per gram of body weight was given to mice intraperitoneally 2 h before sacrifice. Tissues were fixed in 10% formalin solution, processed, and sectioned using standard procedures. BrdU was detected using the BrdU-staining kit (Life Sciences). The BrdU-labeling index was calculated as the ratio of BrdU-positive nuclei over total nuclei in a standardized region of the growth plate.

TUNEL staining. Apoptosis was evaluated in proximal tibial growth plates of seven-day-old miR-140$^{G/G}$ and miR-140$^{+/+}$ mice using the In Situ Cell Detection Kit (Sigma Aldrich) according to the manufacturer's protocol.

Micro-computed tomography analysis. Assessment of the bone morphology and microarchitecture was performed by high-resolution micro-computed tomography (mCT40; Scanco Medical) as previously described[49]. In brief, the proximal tibial epiphysis and metaphysis were scanned using the condition of 70 kVp peak X-ray tube potential, 113 mAs X-ray tube current, 200 ms integration time, and 10 mm isotropic voxel size. For the epiphysis, the total mineralized tissue volume was evaluated. Cancellous bone was assessed in the proximal tibial metaphysis. The epiphysis and primary spongiosa were identified by semi-manual contouring of the regions of interest. Micro-computed tomography analysis was performed in a blinded manner with all mice assigned to coded sample numbers.

Collagen cross-link analysis. Costal rib cartilage was isolated from four-week-old mice. Cartilage was scraped clean, defatted with chloroform/methanol (3:1 v/v), hydrolyzed in 6 N HCl, dried, dissolved in 1% (v/v) *n*-heptafluorobutyric acid, and analyzed by C18 reverse-phase HPLC as described[50]. Lysyl pyridinoline and hydroxylysyl pyridinoline were quantified. Statistical analysis was performed using one-way analysis of variance (ANOVA) and post hoc Tukey–Kramer test.

Super-enhancer analysis. Four H3K27Ac chromatin immunoprecipitation followed by sequencing (ChIP–seq) datasets in human chondrocytes (donor id 57: GSM916058, GSM1112777; donor id 58: GSM1112794, GSM772791; donor id 59: GSM1112787, GSM670044; donor id 60: GSM1112786, GSM772781) were derived from the ENCODE project. Human datasets were aligned to the version hg19 of the human genome. Mouse H3K27Ac and Sox9 ChIP–seq profiles were previously described[51]. Mouse datasets were aligned to the version mm9 of the mouse genome using bowtie 1.0.1[52] with the following parameters: -n 2, -e 70, -m 1, -k 1. Identification of ChIP–seq peaks were performed using MACS 1.4.2.[53] with parameters -p 1e-9, –keep-dup = auto, -w –S –space = 50, and -g hs (or mm). We used ROSE (https://bitbucket.org/young_computation/rose) for identification of super-enhancers according to previous reports[12,13]. The stitching distance and transcription start site (TSS) exclusion zone size were set to 12,500 bp and 2,500 bp, respectively.

RNA sequencing and analysis. Libraries were prepared using the Illumina TruSeq Stranded mRNA sample preparation kit from 500 ng of purified total RNA according to the manufacturer's protocol, and sequenced on Illumina NextSeq500. Single-end 75-bp sequences were read. Sequenced reads were aligned to the mm9 reference genome assembly and gene counts were quantified using STAR (v2.5.1b)[54] and Cufflinks (2.2.0)[55]. In Extended Data Fig. 7f, the reads were mapped to the mouse genome with RNA-seq Unified Mapper (RUM), and gene expression changes was calculated in R by the counting and normalization of uniquely mapped reads and by using edgeR[56,57]. For differential expression (Extended Data Fig. 7d,f), statistical analysis was performed using one-way ANOVA and post hoc Tukey–Kramer test.

Small RNA sequencing and analysis. Total RNA was extracted from chondrocytes using the Direct-zol RNA MiniPrep kit (Zymo Research). Small RNA libraries were constructed with the NEBNext Small RNA Library Prep Set for Illumina (New England Biolabs) and analyzed by Illumina HiSeq 2000. Obtained sequences were processed for adapter removal and size exclusion of sequences <15 nt with Cutadapt. Filtered reads were mapped to the mm9 genome assembly with bowtie 1.0.1, allowing two mismatches, and further quantitated using miRBase v21 as previously described[52,58]. Drosha and Dicer cleavage sites were inferred from the 5′-ends of 5p and 3p miRNAs, respectively. Selection of a single mature miRNA strand from miRNA duplex is determined by combination of 5′-end nucleotide identity and thermodynamic asymmetry of miRNA duplex termini[20]. Ago2 prefers strands with 5′-uridine or 5′-adenosine and thermodynamically unstable 5′-ends. This mechanistically explains abundant expression of miR-140-3p compared with miR-140-5p in chondrocytes. Prediction of the 5p/3p ratio of WT and mutant miRNA duplexes was performed using 5′-trinucleotide sequences according to our previous report[20]. The A>G substitution is predicted to destabilize the 5′-end of the 5p arm and thus to increase the 5p/3p ratio for both 5p/3p.1 and 5p/3p.2 duplexes, consistent with the increase in miR-140-5p and decrease in miR-140-3p in chondrocytes from miR-140$^{G/+}$ and miR-140$^{G/G}$ mice. Sequence analysis of miR-140 is summarized in Supplementary Table 3.

Target identification of miR-140-5p-G and bioinformatics. Identification of miR-140-5p-G was performed using stand-alone computation pipelines of the latest version of TargetScan (v7.0)[21] released in September 2016. We performed target predictions for both mouse and human 3′ UTR annotations and analyzed target site conservation and cumulative weighted context++ score (CWCS) using TargetScanMouse and TargetScanHuman, respectively. We used TargetScanMouse predictions for subsequent analyses. These predictions are included in Supplementary Tables 4 and 5. Datasets of target genes of all miRNAs were downloaded from the TargetScan database (v7.1; targetscan.org) in September 2016. After filtering out genes with a maximum count across all samples less than 1, normalized fragments per kilobase of transcript per million data were used for subsequent analyses. For target gene expression changes, *P* values were calculated by one-sided Kolmogorov–Smirnov test for either direction: upregulation or downregulation. GO analysis was carried out using Database for Annotation, Visualization, and Integrated Discovery (DAVID; https://david.abcc.ncifcrf.gov).

GSEA and RBP motif analysis. GSEA was performed with the GSEA software (http://software.broadinstitute.org/gsea/index.jsp) using C2CGP gene set collections[25]. Enrichment map analysis was further applied for GSEA results[26]. In Fig. 3h, enrichments in WT versus miR-140$^{G/G}$ and WT versus miR-140-null cells

are mapped to the node borders and inner node area, respectively. Red and blue represent upregulation and downregulation in WT cells relative to miR-140$^{G/G}$ or miR-140-null cells, respectively. Node size represents the number of genes overlapped. For motif analysis in Fig. 4a and Extended Data Fig. 8a, we performed RBP motif analysis using the CISBP-RNA database (http://cisbp-rna.ccbr.utoronto.ca/index.php)[30]. Ybx1 binding motifs were previously investigated by RNAcompete, SELEX, and iCLIP in human glioblastoma cells[30–32]. In Extended Data Fig. 10e, YBX1 target gene analysis was performed using GSE63562 datasets[34]. We selected genes downregulated by YBX1 knockdown ($P < 0.05$, fold change > 2) in MDA-MB-231 cells as YBX1-regulated genes and used them for subsequent analysis. To perform systematic RBP motif analysis, we collected mouse representative 3′ UTR sequences of all genes from TargetScanMouse (v7.1) and predicted all possible bindings sites of 94 human/mouse RBPs using RBPmap (http://rbpmap.technion.ac.il/) and a high stringency option[59]. For each gene, the RBP target score (RTS) for each RBP was calculated as follows: RTS = Σ (z-score of each binding site). Then, for each RBP, genes with top 200 RTS were considered as strong possible RBP targets and used as the gene set in GSEA in Extended Data Fig. 10f,g.

qRT–PCR analysis. qRT–PCR was performed on complementary DNA synthesized by the Verso cDNA Synthesis Kit (Thermo Fisher Scientific) using EvaGreen qPCR Supermix (Solis Biodyne). Primer sequences are described in Supplementary Table 9. Statistical analysis was performed using one-way ANOVA and post hoc Tukey–Kramer test.

Plasmids. A primary miRNA transcript (pri-miRNA) expression vector was prepared by inserting short fragments of pri-miR-140 into pcDNA6.2-GW/EmGFP-miR (Thermo Fisher Scientific) as previously described[20]. 3′ UTR reporter vectors were generated by inserting miRNA-target site sequences or 3′ UTRs of miR-140-5p-G target genes into the 3′ UTR of the luciferase gene in the psiCHECK-2 dual luciferase reporter vector (Promega). For the Hif1a CDS expression plasmid, a Flag-tagged mouse Hif1a cDNA was cloned into the pcDNA3 vector. The 3′ UTR sequences of mouse *Btg1*, *Ehd1*, *Trps1*, and *Loxl3* were inserted into the 3′ UTR of mCherry in the bidirectional pTRE-Tight-BI (Clontech) enhanced yellow fluorescent protein (eYFP) and mCherry reporter vector. Mutations were introduced by a PCR-based approach. Primer sequences are given in Supplementary Table 9.

Luciferase reporter assay. HEK293T cells were obtained from American Type Culture Collection and cultured in DMEM (Thermo Fisher Scientific) supplemented with 10% FBS, penicillin, and streptomycin. HEK293T cells were tested for mycoplasma contamination. Cells were transfected with luciferase reporter and pri-miRNA expression plasmids using TransIT-LT1 (Mirus). Cell extracts were collected 48 h after transfection. The ratio between firefly and *Renilla* luciferase was determined using Dual-Luciferase Reporter Assay System (Promega). We independently repeated the molecular biology experiment at least twice, and all attempts to reproduce the results were successful. The sample size was estimated based on our previous experience of similar sets of experiments.

Hif1a target site analysis. Forty-eight hours after transfection with Flag-Hif1a and pri-miRNA expression plasmids, HEK293T cells were lysed with a buffer containing 1% Nonidet P-40, 20 mM Tris-HCl, pH 7.4, 150 mM NaCl, 5 mM EDTA, and cOmplete, Mini, EDTA-free Protease Inhibitor Cocktail Tablet (Sigma) and subjected to western blot analysis using anti-Flag M2 (F1804, Sigma) and anti-α-tubulin DM-1A (T9026, Sigma) antibodies. Uncropped blots are shown in Source Data.

seCLIP analysis. seCLIP was performed in mouse WT and mutant primary chondrocytes using Ago2 and Ybx1 antibodies, as previously described[35,36]. Anti-Ago2 antibody (2D4), and anti-Ybx1 antibody (D2A11, no. 9744S) were from Wako and Cell Signaling Technology, respectively. seCLIP libraries were analyzed by Illumina HiSeq 2000. Identification of seCLIP clusters enriched against size-matched input (SMinput) was performed as previously described[35]. First, obtained sequences were processed for adaptor removal, and reads mapped to repetitive elements were removed. Remaining reads were aligned to the mm9 reference genome assembly using STAR (v2.5.1b), and seCLIP clusters for IP samples were identified using the CLIPper peak calling algorithm (https://github.com/yeolab/eclip) with the following parameters: –bonferrroni –superlocal –threshold binomial. Finally, peak normalization versus SMinput was performed, and enriched seCLIP clusters were extracted. Hexamer enrichment analysis was performed using the 40 nt sequences flanking the center of seCLIP clusters on both directions, all possible hexamers, and background sequence sets generated by HOMER (v4.10)[60]. In Fig. 4b,c,f, non-intronic seCLIP clusters were analyzed, as enrichment of miRNA motifs and Ybx1 motifs was not observed in intronic clusters (Extended Data Figs. 9b and 10b) and majority of intronic clusters was considered to be non-specific due to the low sensitivity of seCLIP for small starting materials like primary chondrocytes. In Fig. 4e, motif prediction was performed using all Ybx1 seCLIP clusters in WT chondrocytes, the 20 nt sequences flanking the center of seCLIP clusters on both directions, and HOMER. In Fig. 4g, the seCLIP site coverage was calculated using the window of seCLIP cluster regions and 20 nt extension on both directions.

Single-cell dual fluorescence reporter assay. HEK293T cells were first transfected with control or YBX1 siRNA (Dharmacon, siGENOME SMARTpool siRNA) using Lipofectamine RNAiMAX (Thermo Fisher Scientific). Twenty-four hours later, eYFP and mCherry reporter, rtTA expression vector, and pri-miRNA expression plasmids were subsequently cotransfected using TransIT-LT1 (Mirus). Four hours after the second transfection, HEK293T cells were treated with 1 μg ml^{-1} doxycycline (Sigma). Forty-eight hours after the second transfection, flow cytometry analysis was carried out with BD FACS Celesta (BD Biosciences). Data collection was performed using FACS Diva Version 8.0.1. FlowJo version 10.4.1. and R was used for data analysis. After selection of single-cell populations by sequential gating by SSC and FSC (Extended Data Fig. 10d), we analyzed about 20,000 eYFP-positive cells (P3 population in Extended Data Fig. 10d). Each experiment included non-transfection controls. The eYFP and mCherry signals of each cell were background-normalized by subtracting the mean value plus two standard deviation of signal in non-transfected samples, and binned by eYFP signal levels, as previously described[40].

Statistical analysis. Sample sizes were determined based on our previous experience of performing similar sets of experiments. The number of examined mice is shown in Supplementary Table 2. Statistical tests were performed using R (3.5.0). In Fig. 2g,h and Extended Data Fig. 3h, differences among multiple groups were analyzed using Kruskal–Wallis one-way ANOVA and post hoc Steel–Dwass test. In Fig. 3a,e,f and Extended Data Figs. 3a and 7c–f, statistical analysis was performed using one-way or two-way ANOVA and post hoc Tukey–Kramer test. In Figs. 3d, 4d, and 4h and Extended Data Figs. 6a,b, 9c, and 10e, P values were calculated by one-sided Kolmogorov–Smirnov test for either direction: upregulation (U) or downregulation (D). For Kolmogorov–Smirnov tests, Bonferroni correction was performed in Figs. 3d, 4d, and 4h and Extended Data Figs. 6a,b and 9c. In Fig. 4b,c,f and Extended Data Figs. 9a and 10a, statistical significance was assessed using two-sided Fisher's exact test with Bonferroni correction. In Extended Data Figs. 9b and 10b, statistical significance was assessed with one-sided Fisher's exact test with Bonferroni correction. In Fig. 4g, statistical significance was assessed with two-sided Wilcoxon rank sum test with Bonferroni correction. In molecular biology experiments, results are representative of more than two independent and reproducible experiments. In all bar graphs, data are expressed as mean ± s.e.m. In all box plots, center lines show medians; box limits indicate the 25th and 75th percentiles; whiskers extend to 1.5× the interquartile range.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

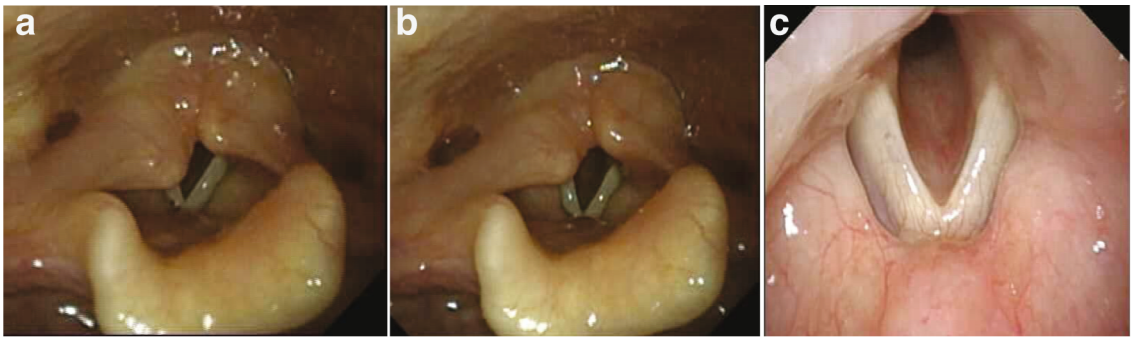
Data availability

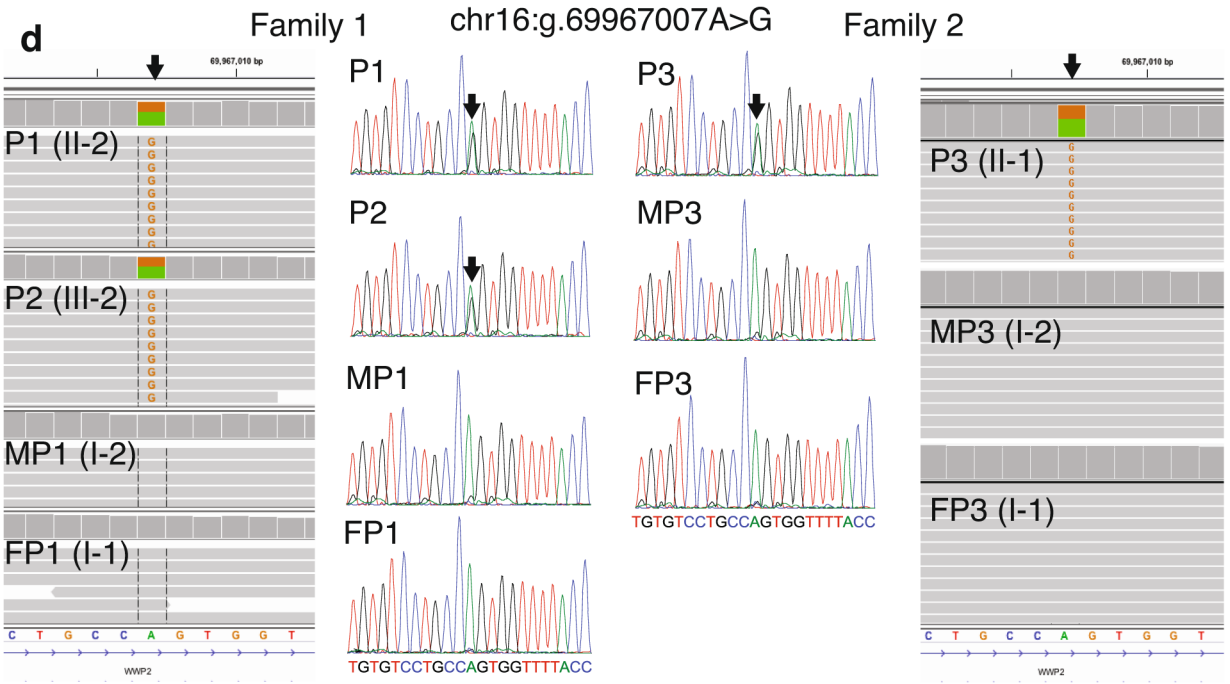
Data generated during this study are available in the Gene Expression Omnibus under accession number GSE98309. The human variant is deposited in the ClinVar database (SCV000586692.1). Human genome data from the individuals participating in the study is protected by Swedish law (2006:351), and raw Sanger sequencing data within the area of interest is available upon request. All other data will be made available upon request to the corresponding author.

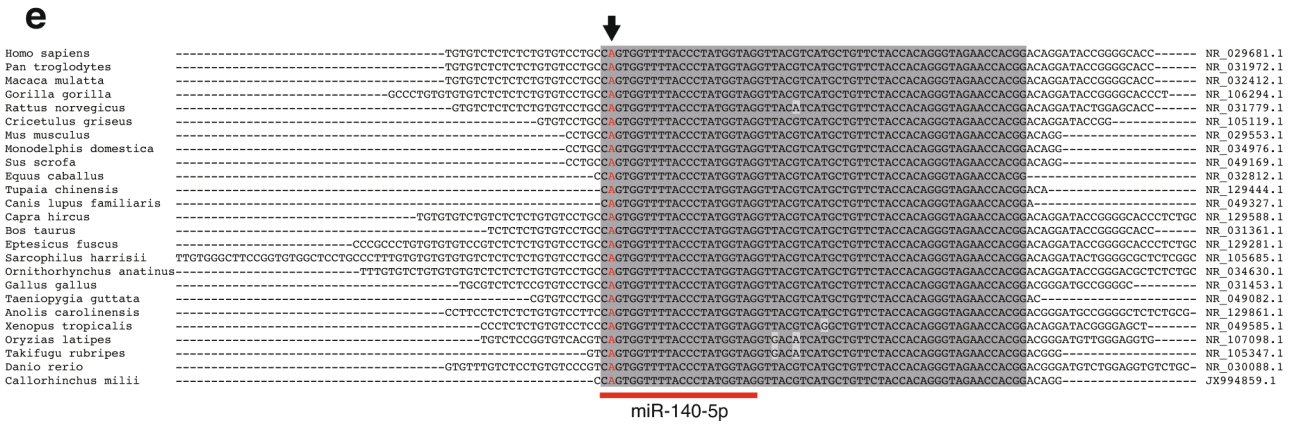
References

41. Kaminsky, E. B. et al. An evidence-based approach to establish the functional and clinical significance of copy number variants in intellectual and developmental disabilities. *Genet. Med.* **13**, 777–784 (2011).
42. Niklasson, A. & Albertsson-Wikland, K. Continuous growth reference from 24th week of gestation to 24 months by gender. *BMC Pediatr.* **8**, 8 (2008).
43. Kvarnung, M. et al. Genomic screening in rare disorders: new mutations and phenotypes, highlighting ALG14 as a novel cause of severe intellectual disability. *Clin. Genet.* **94**, 528–537 (2018).
44. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
45. McKenna, A. et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–1303 (2010).
46. Okonechnikov, K., Conesa, A. & Garcia-Alcalde, F. Qualimap 2: advanced multi-sample quality control for high-throughput sequencing data. *Bioinformatics* **32**, 292–294 (2016).
47. Lindstrand, A. et al. Improved structural characterization of chromosomal breakpoints using high resolution custom array-CGH. *Clin. Genet.* **77**, 552–562 (2010).
48. Yang, H., Wang, H. & Jaenisch, R. Generating genetically modified mice using CRISPR/Cas-mediated genome engineering. *Nat. Protoc.* **9**, 1956–1968 (2014).
49. Wein, M. N. et al. SIKs control osteocyte responses to parathyroid hormone. *Nat. Commun.* **7**, 13176 (2016).
50. Eyre, D. Collagen cross-linking amino acids. *Methods Enzymol.* **144**, 115–139 (1987).

51. Ohba, S., He, X., Hojo, H. & McMahon, A. P. Distinct transcriptional programs underlie Sox9 regulation of the mammalian chondrocyte. *Cell Rep.* **12**, 229–243 (2015).
52. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome. Biol.* **10**, R25 (2009).
53. Zhang, Y. et al. Model-based analysis of ChIP-Seq (MACS). *Genome. Biol.* **9**, R137 (2008).
54. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29, (15–21 (2013).
55. Trapnell, C. et al. Transcript assembly and quantification by RNA-seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* **28**, 511–515 (2010).
56. Grant, G. R. et al. Comparative analysis of RNA-Seq alignment algorithms and the RNA-Seq unified mapper (RUM). *Bioinformatics* **27**, 2518–2528 (2011).
57. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).
58. Zamudio, J. R., Kelly, T. J. & Sharp, P. A. Argonaute-bound small RNAs from promoter-proximal RNA polymerase II. *Cell* **156**, 920–934 (2014).
59. Paz, I. et al. RBPmap: a web server for mapping binding sites of RNA-binding proteins. *Nucleic Acids Res.* **42**, W361–W367 (2014).
60. Heinz, S. et al. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol. Cell* **38**, 576–589 (2010).

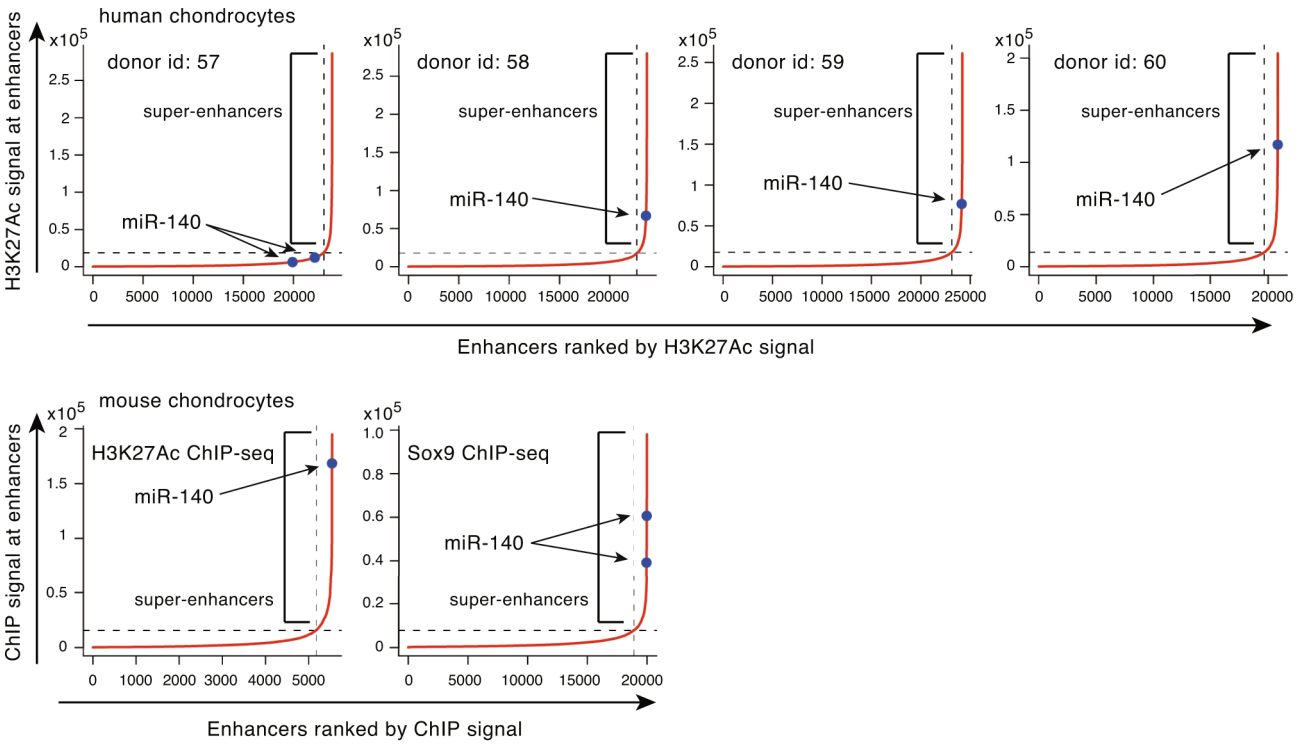






Extended Data Fig. 1 | Human phenotype and genotype. a–**c**, Fiberscopy findings in P1, 44 years of age, consistent with laryngomalacia. **a**, The right arytenoid cartilage is enlarged with redundant soft tissues and it prolapses antero-medially over the larynx during inspiration. **b**, Reduced laryngeal contraction is observed during expiration. **c**, A close view of the larynx and vocal folds shows dorsal narrowing of the subglottic region. **d**, The pathogenic variant in *MIR140* (chr16:g.69967007A>G) (indicated by black arrows) in WGS and exome sequencing data (left and right panels), and its Sanger sequencing confirmation (middle panel) in individuals P1, P2, and P3 as in Fig. 1. Position of the mutation is according to GRCh37 (hg19). MP1, mother of P1; FP1, father of P1; MP3, mother of P3; FP3, father of P3. **e**, Evolutionarily conservation of *MIR140* in different species; miRNA hairpin sequence is highlighted in gray. The second nucleotide in miR-140-5p is conserved from human to Australian ghostshark (indicated by a black arrow).

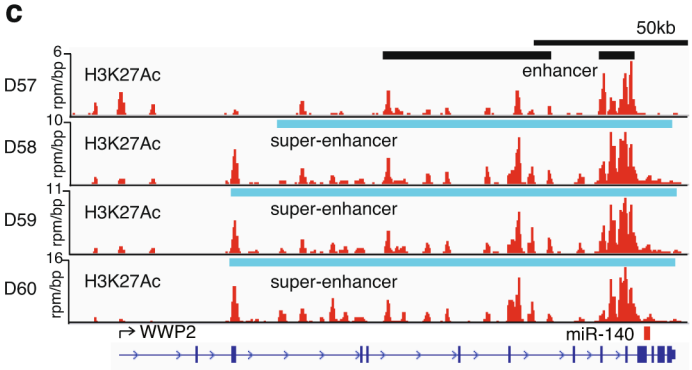
a



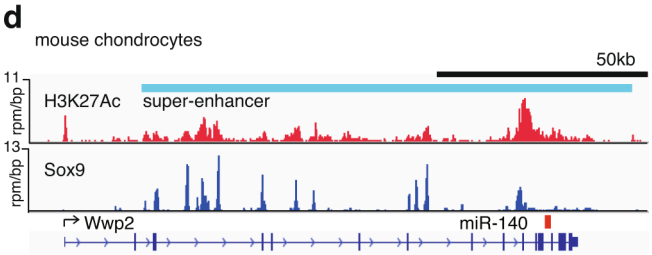
b

samples	SE number	miR-140 enhancer rank	classification
human chondrocyte, H3K27Ac, ENCODE			
donor id: 57 (D57)	825	1728, 3947	top TE
donor id: 58 (D58)	979	81	SE
donor id: 59 (D59)	1060	66	SE
donor id: 60 (D60)	1162	12	SE
mouse chondrocyte			
H3K27Ac	365	3	SE
Sox9	1112	10, 28	SE

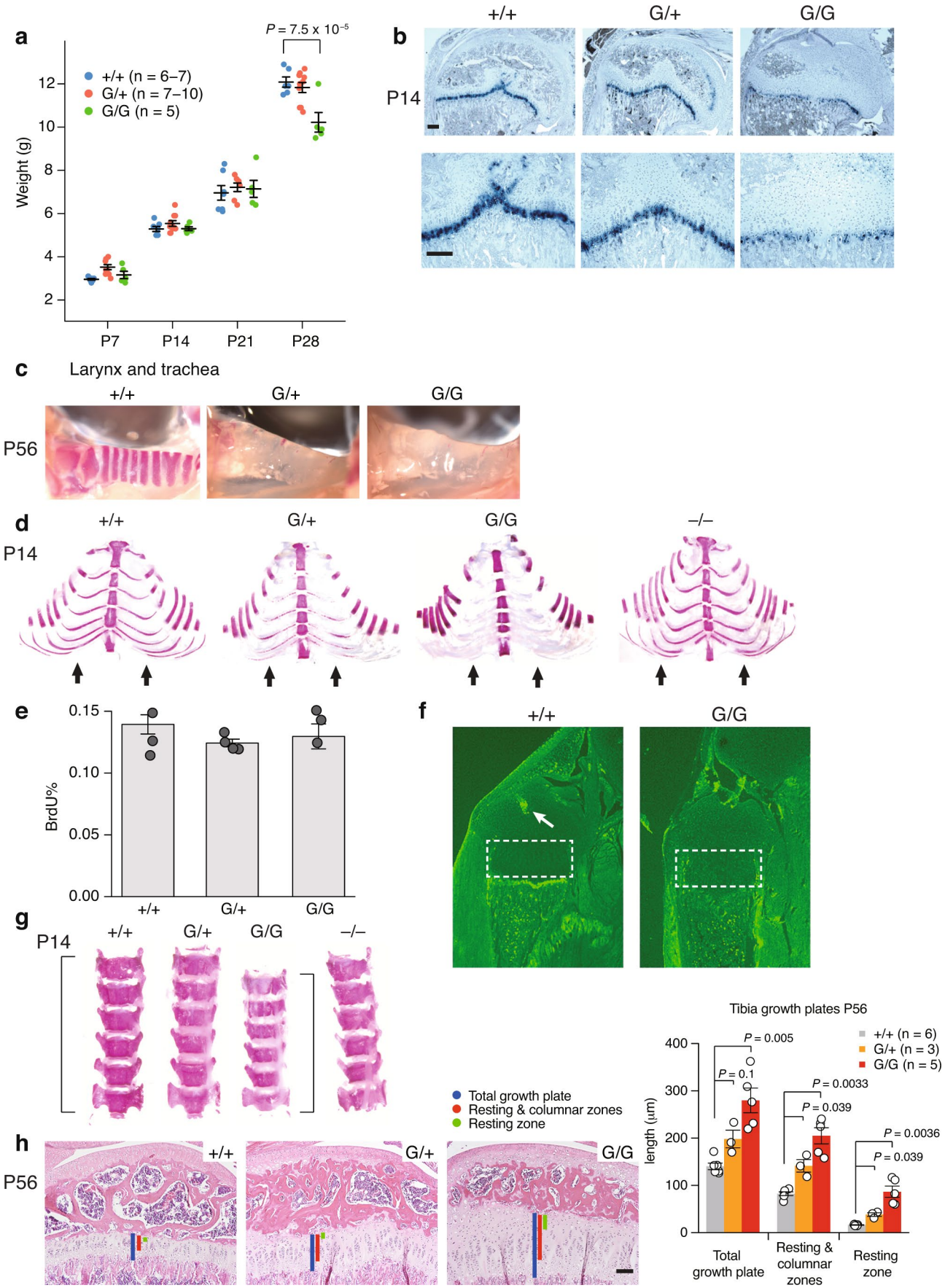
c



d

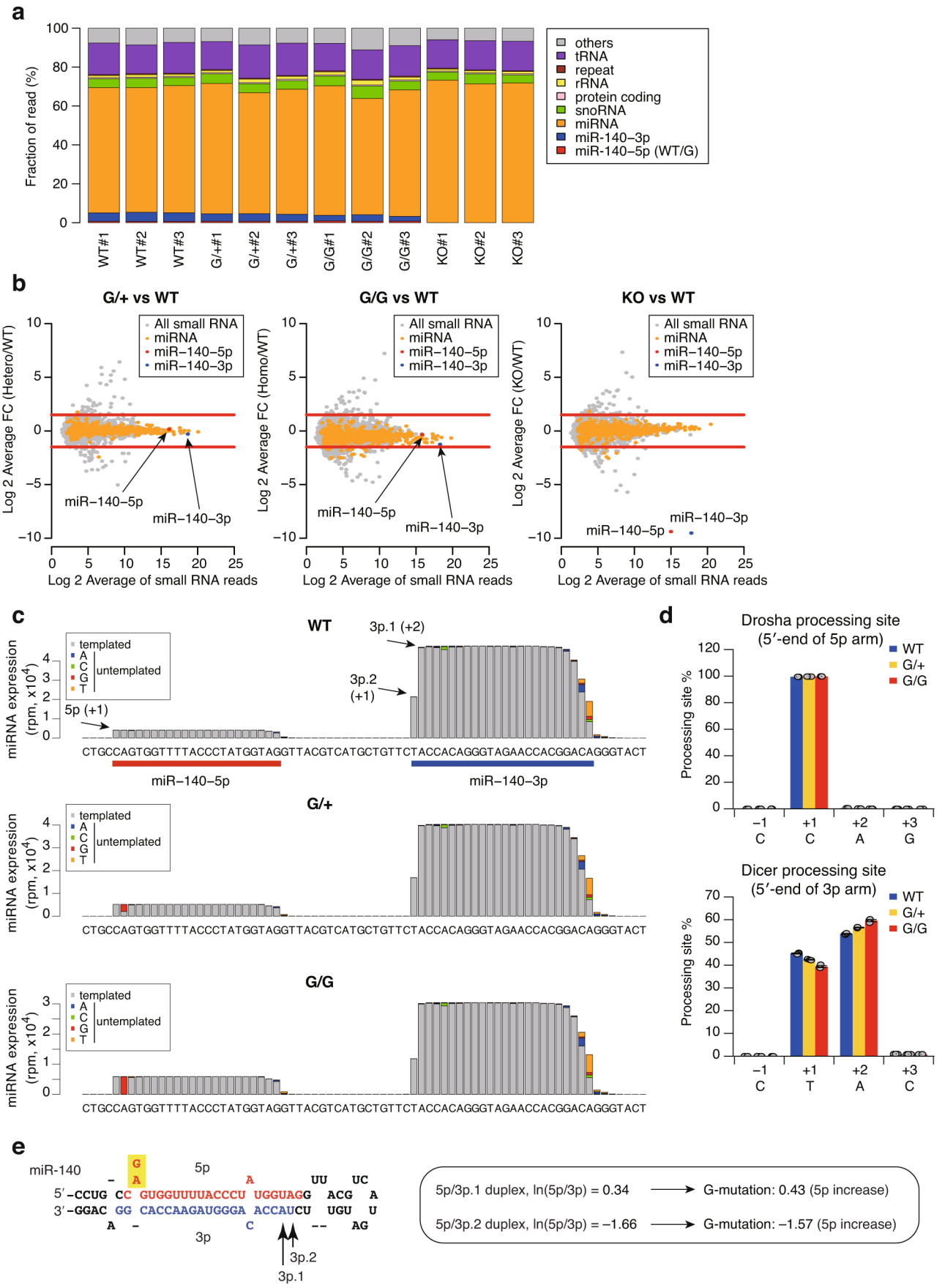


Extended Data Fig. 2 | The miR-140 gene-associated super-enhancers in human and mouse chondrocytes. a, Distribution of ChIP-seq signals at enhancers in human and mouse chondrocytes. The miR-140 gene-associated enhancers are highlighted. The miR-140 locus is associated with high ChIP-seq signals of H3K27Ac and Sox9 in mouse chondrocytes. **b**, Summary of association between miR-140 and super-enhancers (SEs). TE, typical enhancer. **c**, H3K27Ac ChIP-seq profiles at the miR-140 gene locus in human chondrocytes. The H3K27Ac profile of the donor id: 60 is shown in Fig. 1s. **d**, H3K27Ac and Sox9 ChIP-seq profiles of the miR-140 gene locus in mouse chondrocytes.



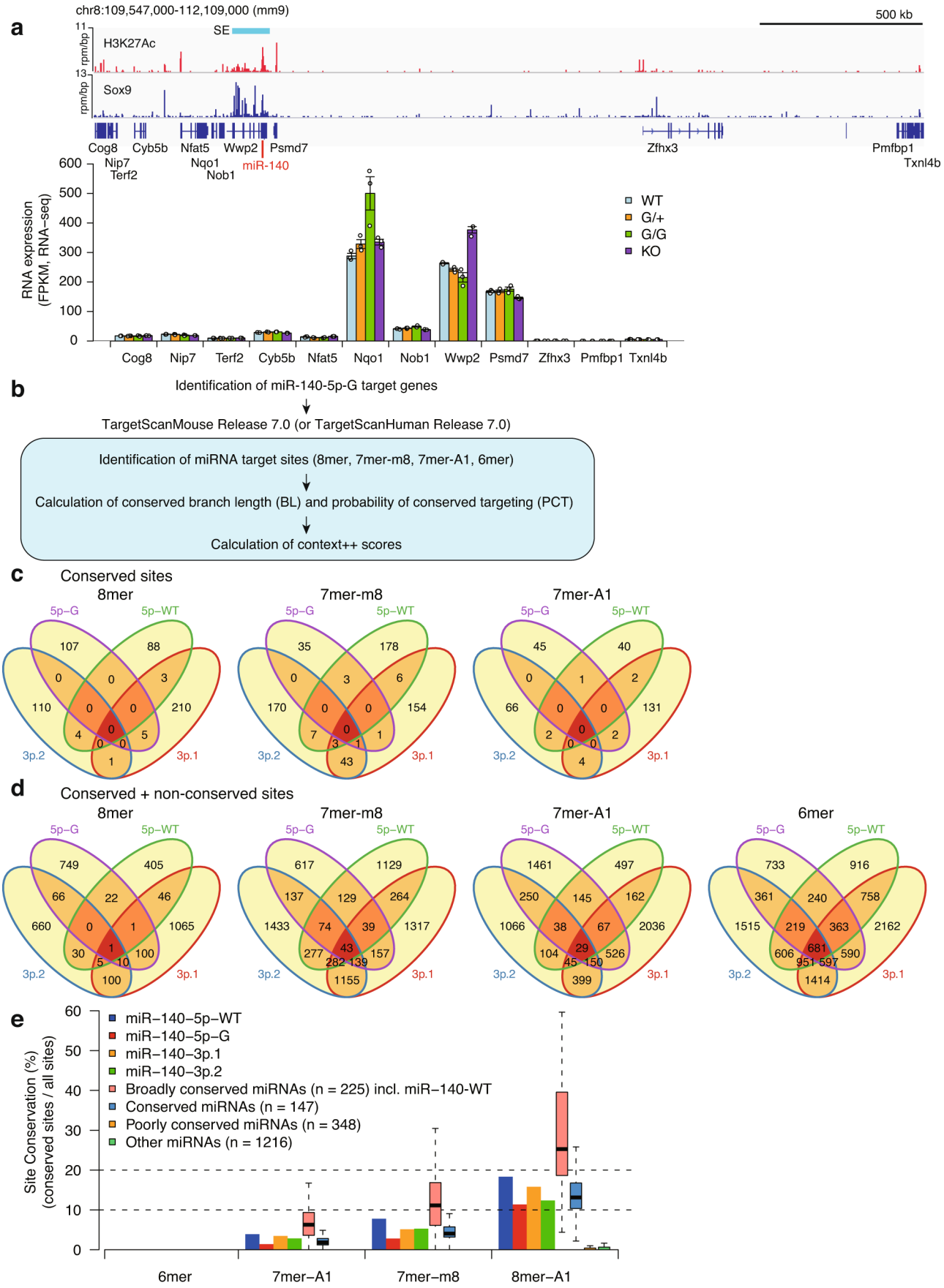
Extended Data Fig. 3 | See next page for caption.

Extended Data Fig. 3 | Skeletal phenotype of miR-140 A>G mutant mice. a, miR-140$^{G/G}$ mice are smaller. The *y* axis and *x* axis indicate weight in grams and days of postnatal age, respectively. Numbers of evaluated mice with indicated genotypes are shown in parentheses. Results represent mean ± s.e.m. Statistical significance was assessed using two-way ANOVA and post hoc Tukey–Kramer test. **b**, Representative images from two independent experiments (at least three sections of each genotype) of in situ hybridization in epiphyseal growth plates of proximal humeri. Decreased expression of *Col10a1* is observed in miR-140$^{G/G}$ and miR-140$^{G/+}$ mice. **c**, Representative images of Alizarin red staining of the trachea (+/+, $n=6$; G/+, $n=3$; G/G, $n=5$). Cartilage mineralization is absent or severely impaired in miR-140$^{G/+}$ and miR-140$^{G/G}$ mice. **d**, Representative images of Alizarin red staining of the thoracic cage (+/+, $n=4$; G/+, $n=7$; G/G, $n=9$; −/−, $n=2$). Delay in chondrocyte mineralization in the rib cartilage is observed in miR-140$^{G/G}$ and miR-140$^{G/+}$ mice. **e**,**f**, Proliferation rate assessed by BrdU labeling (**e**) and apoptosis assessed by TUNEL staining (**f**) in P7 tibial proximal growth plates. The fraction of BrdU-positive cells was calculated in the columnar proliferating chondrocytes. Results represent mean ± s.e.m. (+/+, $n=3$; G/+, $n=4$; G/G, $n=3$). Apoptotic cells were not observed in the growth plate of miR-140$^{G/G}$ or miR-140$^{+/+}$ mice (**f**). White arrow indicates apoptotic cells associated with blood vessels invaded into the epiphysis of WT mice. White dotted boxes indicate the growth plates (excluding the epiphyseal cartilage), and there were no TUNEL-positive cells in the growth plate cartilage in both miR-140$^{+/+}$ and miR-140$^{G/G}$ mice. **g**, Representative images of Alizarin red staining of vertebral bodies (+/+, $n=4$; G/+, $n=7$; G/G, $n=9$; −/−, $n=2$). Vertebral bodies of miR-140$^{G/G}$ mice are smaller compared with miR-140$^{G/+}$, miR-140$^{+/+}$, and miR-140$^{-/-}$ mice. **h**, Delayed maturation of the proximal tibial epiphysis in P56-old mutant mice. Scale bars, 200 μm. Measurements of tibial growth plates of P56 mice show significant expansions in miR-140$^{G/G}$ and miR-140$^{G/+}$ mice. The bar graphs show mean ± s.e.m. (+/+, $n=6$; G/+, $n=3$; G/G, $n=5$). Statistical significance was assessed using Kruskal–Wallis one-way ANOVA and post hoc Steel–Dwass test.



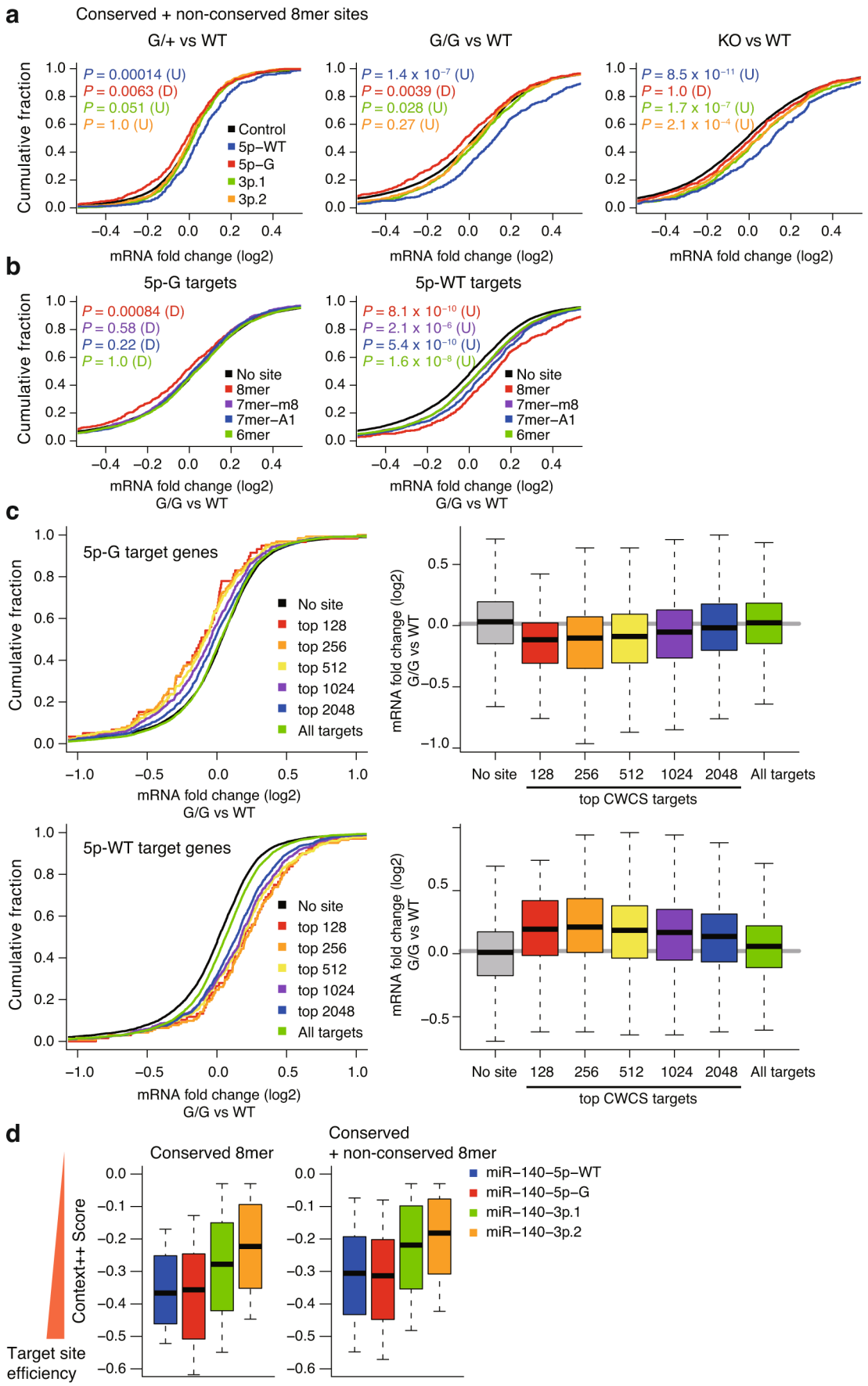
Extended Data Fig. 4 | See next page for caption.

Extended Data Fig. 4 | Small RNA profiles in miR-140-G mutant chondrocytes. a, Proportion of small RNA species of mouse primary chondrocytes with indicated genotypes. **b**, Relative fold changes and mean expression levels of small RNAs in WT chondrocytes versus mutant chondrocytes. Note that miR-140-5p and miR-140-3p are highly abundant miRNAs in chondrocytes, reflecting super-enhancer-mediated high transcriptional activity. Red lines denote twofold cutoffs. Significant changes of other miRNAs are not observed. **c**, Pileup plots of miR-140-5p and miR-140-3p in WT and mutant chondrocytes. **d**, Comparison of Drosha and Dicer processing sites in WT and mutant chondrocytes. Drosha and Dicer cleavage sites were inferred from the 5′-ends of miR-140-5p and miR-140-3p, respectively. Results represent mean ± s.e.m. ($n = 3$ animals). **e**, Predicted effects of the A>G substitution on miR-140 5p/3p ratio. Selection of a single mature miRNA strand from miRNA duplex is determined by combination of 5′-end nucleotide identity and thermodynamic asymmetry of miRNA duplex termini[20]. Ago2 prefers strands with 5′-uridine or 5′-adenosine and thermodynamically unstable 5′-ends. This mechanistically explains abundant expression of miR-140-3p compared with miR-140-5p in chondrocytes. Prediction of 5p/3p ratio was performed according to our previous report[20]. Note that the A>G substitution is predicted to destabilize the 5′-end of the 5p arm and thus to increase the 5p/3p ratio for both 5p/3p.1 and 5p/3p.2 duplexes, consistent with the increase in miR-140-5p and decrease in miR-140-3p in chondrocytes from miR-$140^{G/+}$ and miR-$140^{G/G}$ mice.



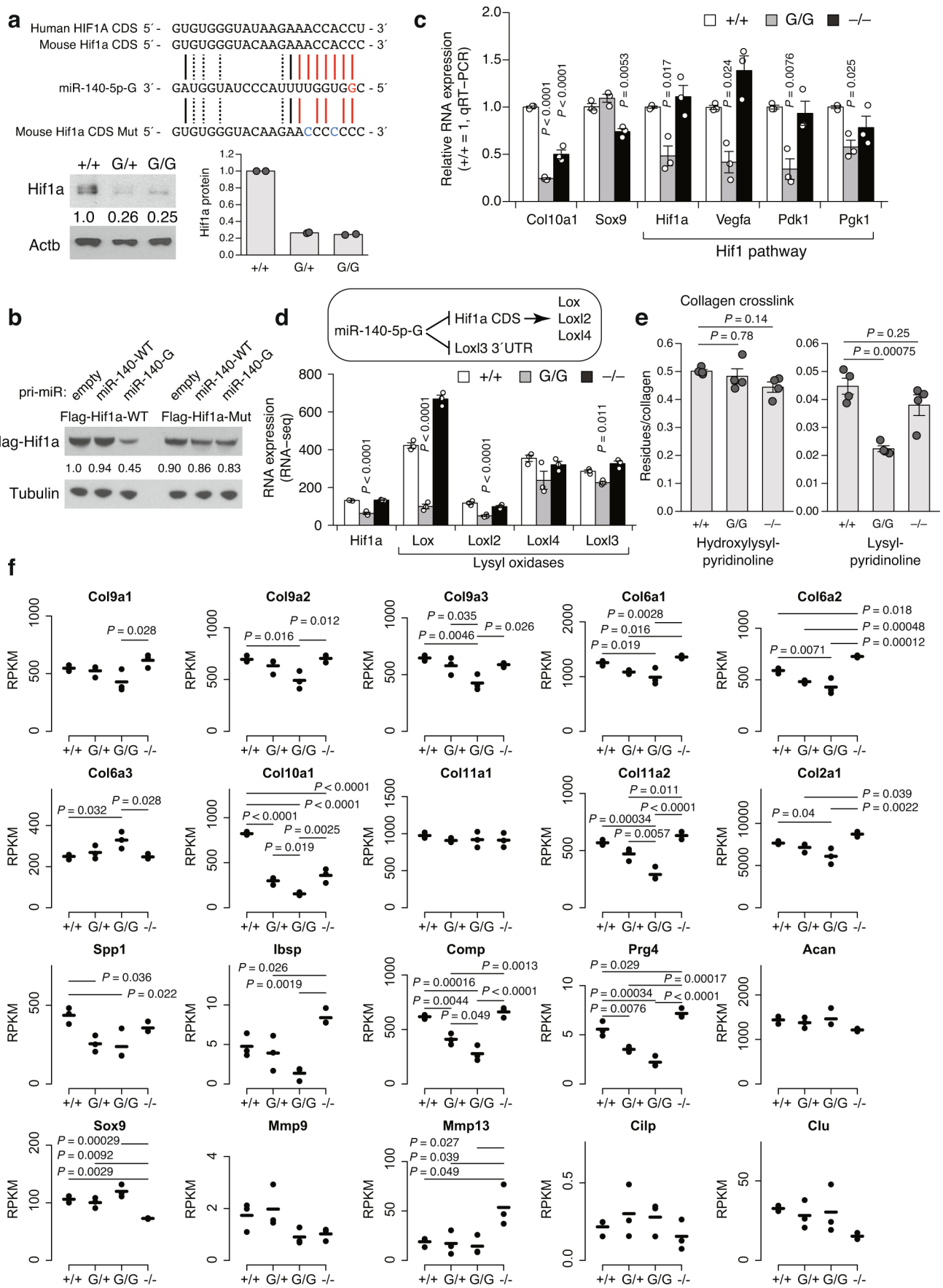
Extended Data Fig. 5 | See next page for caption.

Extended Data Fig. 5 | Transcriptome analysis by RNA-seq and identification of miR-140-5p-G targets by TargetScan. a, Expression levels of genes within the neighborhood of the miR-140 gene locus. Results represent mean ± s.e.m. ($n = 3$ animals). **b**, Workflow of target prediction by TargetScan v7.0. **c**,**d**, Venn diagrams showing overlaps of genes with conserved sites (**c**) and genes with conserved and non-conserved sites (**d**) for miR-140-5p-WT, -5p-G, -3p.1, and -3p.2. **e**, Conservation frequency of target sites for miR-140-5p-WT, -5p-G, -3p.1, and -3p.2. The target sites of miR-140-5p-G are less conserved than those of WT miR-140 species. Classification of miRNAs (broadly conserved, conserved, poorly conserved, and others) is based on the TargetScan information. Center lines show medians; box limits indicate the 25th and 75th percentiles; whiskers extend to 1.5× the interquartile range.



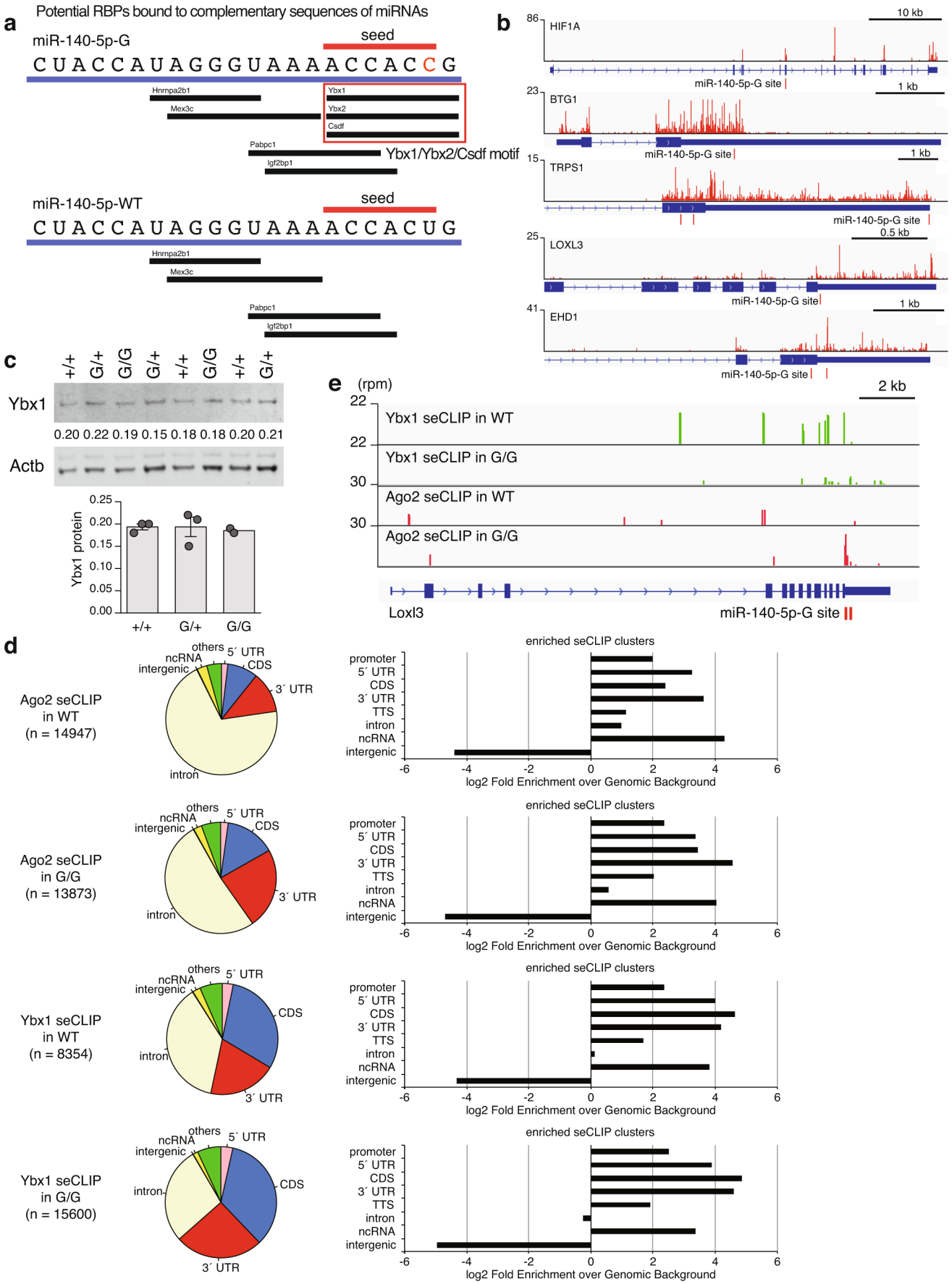
Extended Data Fig. 6 | See next page for caption.

Extended Data Fig. 6 | Deregulation of miR-140-5p-G targets in miR-140 mutant chondrocytes. a, Cumulative distributions of fold changes of mRNAs with conserved and non-conserved 8mer target sites for miR-140-5p-WT, -5p-G, -3p.1, and -3p.2 between miR-140 mutant or null chondrocytes versus WT chondrocytes. *P* values (versus a control gene set) were calculated by one-sided Kolmogorov–Smirnov test for either direction: upregulation (U) or downregulation (D) with Bonferroni correction (control genes, $n = 9{,}706$; miR-140-5p-WT, $n = 333$; miR-140-5p-G, $n = 608$; miR-140-3p.1, $n = 903$; miR-140-3p.2, $n = 603$). **b**, Cumulative distributions of fold changes of mRNAs with all 8mer, 7mer-m8, 7mer-A1, and 6mer target sites for miR-140-5p-G (left) and miR-140-5p-WT (right) when comparing miR-140-G homozygous chondrocytes with WT chondrocytes. *P* values (versus a control gene set) were calculated by one-sided Kolmogorov–Smirnov test for either direction: upregulation (U) or downregulation (D) with Bonferroni correction. Left (miR-140-5p-G): no site, $n = 7{,}684$; 8mer, $n = 608$; 7mer-m8, $n = 862$; 7mer-A1, $n = 1{,}741$; 6mer, $n = 2{,}497$. Right (miR-140-5p-WT): no site, $n = 7{,}453$; 8mer, $n = 333$; 7mer-m8, $n = 1{,}507$; 7mer-A1, $n = 710$; 6mer, $n = 3{,}093$. **c**, Performance of TargetScan v7.0 for prediction of miR-140-5p-G targeting. Cumulative distribution plots (left) and box plots (right) of fold-change expression of predicted target gene sets grouped by CWCS. Expression levels of predicted target genes of miR-140-5p-G (top; no site, $n = 7{,}684$; top 128, $n = 59$; top 256, $n = 129$; top 512, $n = 275$; top 1,024, $n = 559$; top 2,048, $n = 1178$; all targets, $n = 4{,}187$) and miR-140-5p-WT (bottom; no site, $n = 7{,}453$; top 128, $n = 68$; top 256, $n = 142$; top 512, $n = 311$; top 1,024, $n = 637$; top 2,048, $n = 1{,}235$; all targets, $n = 4{,}418$) were analyzed. Gene expression of miR-140-G homozygous chondrocytes and WT chondrocytes was compared. All predictions were considered. High CWCS was associated with stronger target gene regulation for both miR-140-5p-G and miR-140-5p-WT, suggesting that miR-140-5p-G follows mechanisms of conventional miRNA targeting as well as other endogenous miRNAs. In box plots, center lines show medians; box limits indicate the 25th and 75th percentiles; whiskers extend to 1.5× the interquartile range. **d**, Box plots showing context++ scores for conserved (left) and all (right) predicted target sites for miR-140-5p-WT, -5p-G, -3p.1, and -3p.2. Consistent with the effects on target gene expression, miR-140-5p-WT and miR-140-5p-G show similar and higher predicted target site efficiency than miR-140-3p.1 and miR-140-3p.2. Center lines show medians; box limits indicate the 25th and 75th percentiles; whiskers extend to 1.5× the interquartile range. Left (conserved 8mer): miR-140-5p-WT, $n = 96$; miR-140-5p-G, $n = 114$; miR-140-3p.1, $n = 222$; miR-140-3p.2, $n = 115$. Right (conserved and non-conserved 8mer): miR-140-5p-WT, $n = 524$; miR-140-5p-G, $n = 1{,}002$; miR-140-3p.1, $n = 1{,}405$; miR-140-3p.2, $n = 930$.



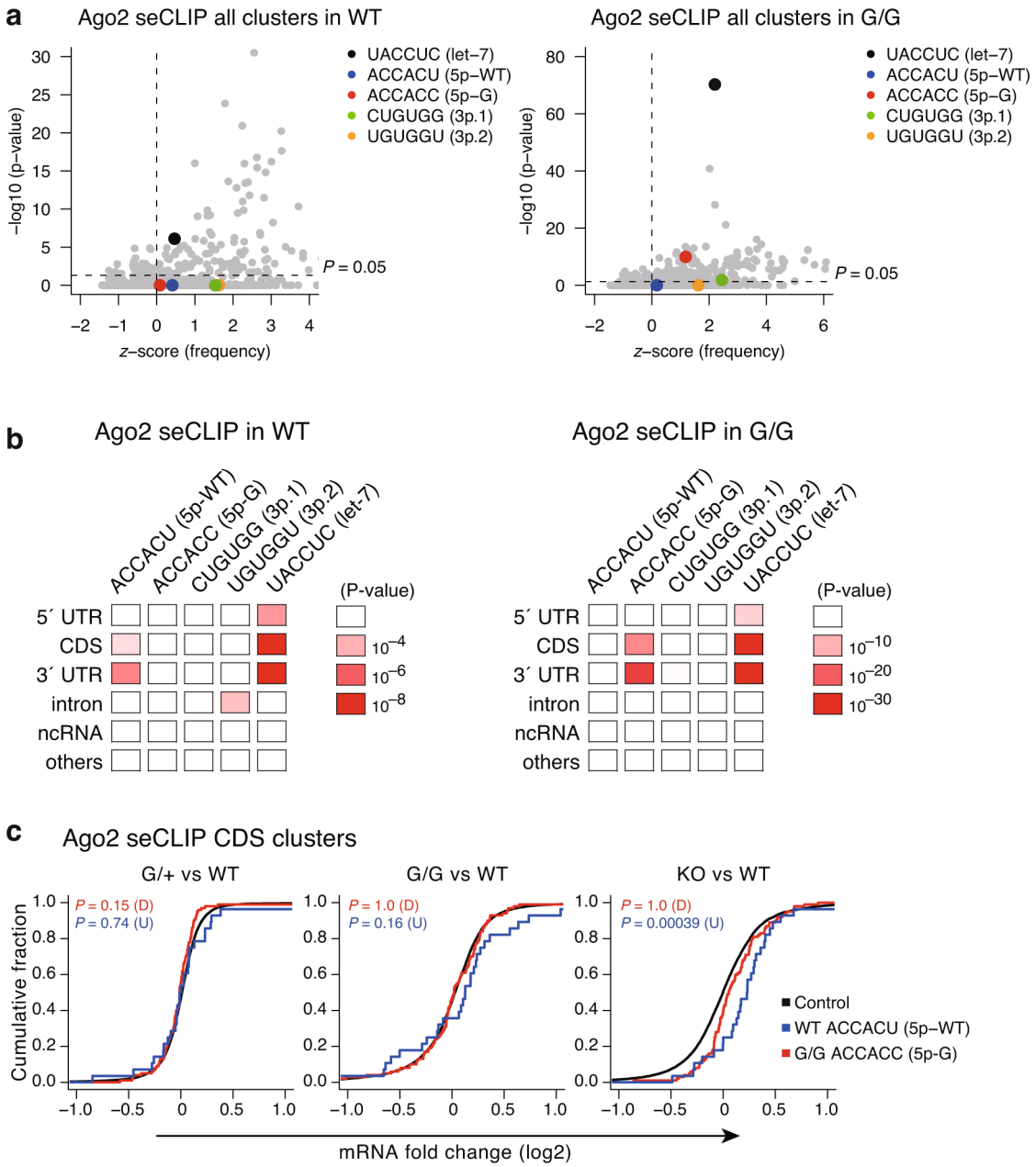
Extended Data Fig. 7 | See next page for caption.

Extended Data Fig. 7 | Deregulation of the Hif1a pathway and the cartilage markers in miR-140 mutant chondrocytes. a, Top: sequence alignments between miR-140-5p-G and its putative binding sites in the coding sequence of mouse and human *Hif1a*. Bottom: representative images from two independent experiments of western blot analysis of Hif1a protein expression in primary rib chondrocytes at the normoxic condition. Hif1a protein was detected by western blot analysis using anti-Hif1a antibody (Novus Biologicals NB100-449). Bar graphs represent mean ($n=2$ animals). **b**, Validation of the Hif1a target site. HEK293T cells were transfected with WT or mutant mouse Hif1a expression plasmids and pri-miRNA expression plasmids, and subjected to western blot analysis. Representative images from two independent experiments are shown. A mutation of the target site is shown in **a**. **c**, Relative expression of *Hif1a* and its target genes was determined by qRT–PCR in primary rib chondrocytes isolated from P7 mice. Results represent mean ± s.e.m. ($n=3$ animals). Statistical significance was assessed using one-way ANOVA and post hoc Tukey–Kramer test. *P* values indicate comparison versus WT. **d**, Expression of lysyl oxidases in P7 primary rib chondrocytes assessed by RNA-seq. Results represent mean ± s.e.m. ($n=3$ animals). Statistical significance was assessed using one-way ANOVA and post hoc Tukey–Kramer test. *P* values indicate comparison versus WT. **e**, Reduced collagen cross-linking in miR-140$^{G/G}$ rib cartilage cells. Rib cartilage of four-week-old mice was subjected to collagen cross-link analysis by mass spectrometry. Hydroxylysyl pyridinoline and lysyl pyridinoline were quantified. Results represent mean ± s.e.m. ($n=4$ animals). Statistical significance was assessed using one-way ANOVA and post hoc Tukey–Kramer test. *P* values indicate comparison versus WT. **f**, RNA expression of cartilage markers. The *x* axis and *y* axis indicate miR-140 genotypes and normalized expression (reads per kilobase per million mapped reads, RPKM), respectively. Each dot represents the expression level in an individual sample ($n=3$ animals), and the thick bars indicate the mean value. Statistical significance was assessed using one-way ANOVA and post hoc Tukey–Kramer test. Connector lines at the top of each panel indicate significant changes in expression ($P<0.05$).

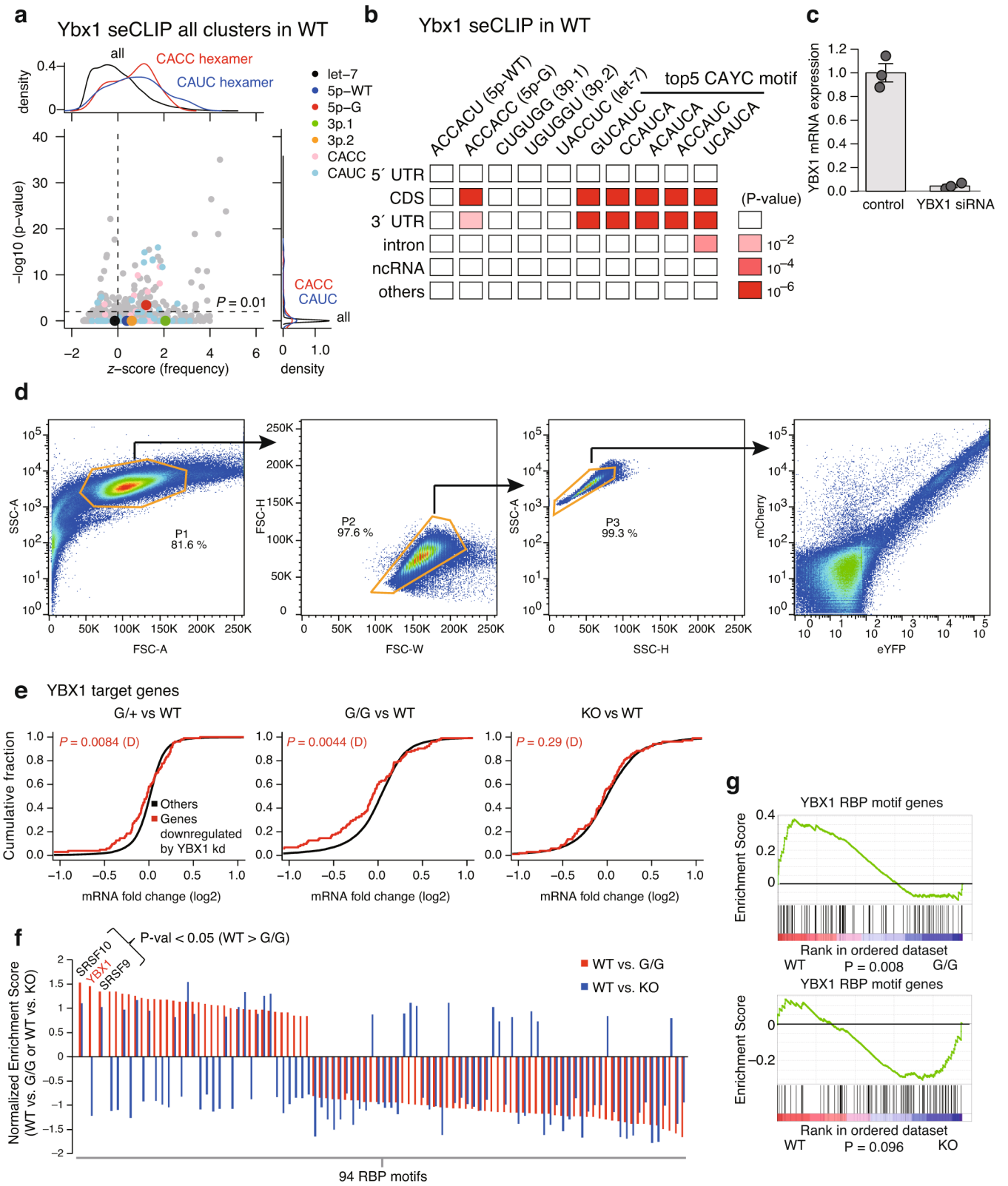


Extended Data Fig. 8 | See next page for caption.

Extended Data Fig. 8 | A relationship between miR-140-5p-G and Ybx1 and Ago2 and Ybx1 seCLIP analysis in mouse primary chondrocytes. a, Potential RBPs that are predicted to bind to complementary sequences of miR-140-5p-G (left) and miR-140-5p-WT (right). Motif analysis was performed using the CISBP-RNA database. **b**, Genome browser shot of YBX1 iCLIP for human *HIF1A*, *BTG1*, *TRPS1*, *LOXL3*, and *EHD1* using the previously reported YBX1 iCLIP data in human glioblastoma cells[32]. We observed that multiple conserved miR-140-5p-G target genes were bound by YBX1 in the YBX1 iCLIP dataset. **c**, Expression levels of Ybx1 protein in mouse primary chondrocytes. Representative images from three independent experiments are shown. Note that Ybx1 protein expression levels normalized to those of Actb did not differ between WT and mutant chondrocytes. Normalized values of densitometric analysis are indicated. Results represent mean ± s.e.m. (+/+, $n=3$; G/+, $n=3$; G/G, $n=2$). **d**, Genomic annotation of Ago2 and Ybx1 seCLIP clusters. Left pie charts shows genomic distribution of Ago2 and Ybx1 seCLIP clusters. Right panels show the fold enrichment of various genomic annotations over the background. **e**, Genome browser shot of the Loxl3 locus.



Extended Data Fig. 9 | Analysis of Ago2 seCLIP clusters. a, Hexamer enrichment analysis in all Ago2-bound seCLIP clusters including intronic clusters identified in WT (left) and mutant (right) chondrocytes (WT, $n=14,947$ clusters; G/G, $n=13,873$ clusters). The x axis shows z-scores of hexamer frequency. The y axis shows the negative $\log_{10}$ P value from two-sided Fisher's exact test with Bonferroni correction of the hexamer enrichment above background. **b**, Hexamer enrichment analysis for various genomic annotations. The negative $\log_{10}$ P values from one-sided Fisher's exact test with Bonferroni correction of the hexamer enrichment above background are shown. Left (Ago2 WT): 5′ UTR clusters, $n=279$; CDS clusters, $n=1,315$; 3′ UTR clusters, $n=1,782$; intronic clusters, $n=10,402$; non-coding RNA region clusters, $n=434$; others, $n=735$. Right (Ago2 G/G): 5′ UTR clusters, $n=303$; CDS clusters, $n=2,006$; 3′ UTR clusters, $n=3,229$; intronic clusters, $n=7,091$; non-coding RNA region clusters, $n=332$; others, $n=912$. **c**, Cumulative distributions of fold changes of mRNAs with Ago2-bound seCLIP CDS clusters with the ACCACC motif in mutant chondrocytes (red) or Ago2-bound seCLIP CDS clusters with the ACCACU motif in WT chondrocytes. P values (versus a control gene set) were calculated by one-sided Kolmogorov–Smirnov test for either direction: upregulation (U) or downregulation (D) with Bonferroni correction (control genes, $n=10,093$; miR-140-5p-WT, $n=28$; miR-140-5p-G, $n=100$).



Extended Data Fig. 10 | See next page for caption.

Extended Data Fig. 10 | Analysis of Ybx1 seCLIP clusters. a, Hexamer enrichment analysis in all Ybx1-bound seCLIP clusters including intronic clusters, identified in WT chondrocytes ($n = 8,354$). The *x* axis shows the *z*-scores of hexamer frequency. The *y* axis shows the negative $\log_{10}$ *P* value of two-sided Fisher's exact test with Bonferroni correction of the hexamer enrichment above background. **b**, Hexamer enrichment analysis for various genomic annotations (5′ UTR clusters, $n = 265$; CDS clusters, $n = 2,537$; 3′ UTR clusters, $n = 1,649$; intronic clusters, $n = 3,143$; non-coding RNA region clusters, $n = 191$; others, $n = 569$). The negative $\log_{10}$ *P* values from one-sided Fisher's exact test with Bonferroni correction of the hexamer enrichment above background are shown. **c**, siRNA-mediated knockdown of YBX1 in HEK293T cells for single-cell reporter analysis, confirmed by qRT–PCR analysis. Results represent mean ± s.e.m. ($n = 3$ biologically independent samples). **d**, Flow cytometry gating strategies for single-cell reporter analysis. Single-cell populations (P3) were selected by sequential gating by SSC and FSC and subjected to two-color analysis in Fig. 4i. An example of the flow cytometry gating strategy is shown. **e**, Cumulative distributions of fold changes of YBX1-stablized gene sets between miR-140 mutant or null chondrocytes versus WT chondrocytes. *P* values (versus a control gene set) were calculated by one-sided Kolmogorov–Smirnov test for downregulation (D) (other genes, $n = 11,768$; YBX1-regulated genes, $n = 103$). **f**,**g**, Systematic RBP target gene analysis using GSEA. For 94 RBPs, possible RBP target genes were predicted using RBPmap[59], and used for GSEA analysis to assess which RBP activity changes in miR-140-G homozygous and miR-140-null chondrocytes relative to WT chondrocytes ($n = 3$ animals for each group). In **f**, normalized enrichment scores for 94 RBPs are shown. Among three significantly suppressed motifs in miR-140-G homozygous cells (SRSF10, YBX1, and SRSF9), only the YBX1 RBP motif gene set was selectively suppressed in miR-140-G homozygous but not miR-140-null cells. In **g**, GSEA enrichment plots for YBX1 RBP motif genes are shown. Enrichment analysis and statistical analysis were performed using GSEA.

nature research

Corresponding author(s): Kobayashi Tatsuya

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	An indication of whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistics including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated
☐	☒	Clearly defined error bars *State explicitly what error bars represent (e.g. SD, SE, CI)*

Our web collection on statistics for biologists may be useful.

Software and code

Policy information about availability of computer code

Data collection	1000 Genomes Project (2014 Oct version), 6500 NHLBI-GO Exome Sequencing Project (EVS), Exome Aggregation Consortium (ExAC v0.2), dbSNP138, e-Array, ENCODE, TargetScan (v7.0)
Data analysis	CASAVA (bcl2Fastq v1.8.3), MOSAIK (2.2.3), Picard (1.92), Genome Analysis Tool (GATK v.3.2.0), ANNOVAR (version 2014 July 14), Variant Effect Predictor (VEP), Fiji (https://imagej.net/Fiji/Downloads), GEMINI (v0.16.0), Combined Annotation Dependent Depletion (CADD), ExomeDepth, bwa (v0.7.12), GATK (v3.3-0-geee94ec), Qualimap (v2.0), HaplotypeCaller (GATK), Agilent Genomic Workbench software (Agilent Technologies), bowtie (1.0.1), MACS 1.4.2., ROSE (https://bitbucket.org/young_computation/rose), STAR (v2.5.1b), Cufflinks (2.2.0), RNA-seq Unified Mapper, edgeR, Cutadapt, TargetScan (v7.0), DAVID (https://david.abcc.ncifcrf.gov), GSEA (http://software.broadinstitute.org/gsea/index.jsp), CISBP-RNA (http://cisbp-rna.ccbr.utoronto.ca/index.php), RBPmap (http://rbpmap.technion.ac.il/), CLIPper (https://github.com/yeolab/eclip), HOMER (v4.10), FACS Diva (8.0.1), FlowJo (10.4.1), GraphPad Prism (8), R (3.5.0).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Data generated during this study are available in the Gene Expression Omnibus under accession number GSE98309. The human variant is deposited in the ClinVar database (SCV000586692.1). Human genome data from the individuals participating in the study is protected by Swedish law (2006:351), and raw Sanger sequencing data within the area of interest is available upon request. All other data will be made available upon request to the corresponding author.

Field-specific reporting

Please select the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/authors/policies/ReportingSummary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Specific samples sizes are indicated in figure legends and in Supplemental Table 2. In animal study, the sample size was estimated based on our prior experience of performing similar sets of experiments. In molecular biology experiments, n=3 was chosen because it is sufficient to conduct statistical tests to generate p-values to determine if results are significant.
Data exclusions	No data were excluded.
Replication	Human DNA samples were sequenced using ES and WGS in one experiment, and the mutation was confirmed using Sanger sequencing in three independent experiments. Reproducibility of animal studies was confirmed by examination of multiple animals as described in Supplementary Table 2. In molecular biology experiments, we independently repeated the experiments at least twice, and all attempts to reproduce the results were successful.
Randomization	Randomizaton was not performed. Sample randomization is not applicable for cell culture experiments where batches of homogeneous cultures can be tested in parallel.
Blinding	Blinding was performed for micro-CT analysis.

Reporting for specific materials, systems and methods

Materials & experimental systems

n/a	Involved in the study
☐	☒ Unique biological materials
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology
☐	☒ Animals and other organisms
☐	☒ Human research participants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☐	☒ Flow cytometry
☒	☐ MRI-based neuroimaging

Unique biological materials

Policy information about availability of materials

Obtaining unique materials	All unique materials are available from the corresponding author upon reasonable request (except human DNA samples).

Antibodies

Antibodies used	anti-Hif-1a antibody, Novus Biologicals, NB100-449, dilution (WB) 1:1000 anti-beta actin antibody, Cell Signaling Technology, clone 13E5, Cat#4970S, dilution (WB) 1:1000 anti-Ybx1 antibody, Cell Signaling Technology, clone D2A11, Cat#9744S, dilution (WB) 1:1000 anti–Flag antibody, Sigma, clone M2, Cat#F1804, dilution (WB) 1:1000 anti–α-tubulin, Sigma, clone DM-1A, Cat# T9026, dilution (WB) 1:1000 anti-Ago2 antibody, Wako, clone 2D4, Cat#014-22023
Validation	All antibodies are well-validated by the manufacturer and are widely used in the scientific community. In addition, we observed expected bands and strong signals in mouse and human cells.

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	HEK293T cells were obtained from American Type Culture Collection.
Authentication	Cell lines were not authenticated.
Mycoplasma contamination	All cell lines were tested negative for mycoplasma contamination.
Commonly misidentified lines (See ICLAC register)	No commonly misidentified cell lines were used.

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals	Mir140-G knock-in mice were generated and kept in the C57/B6 background. Mir140 KO mice were kept in a C57/B6 dominant (>75%) mixed background due to the perinatal lethality in the pure C57/B6 background. Both males and females from P7 to P56 (P7, P14, P21, P25, P28, and P56) were used for the study. Number of animals used for skeletal phenotyping (P7, P14, P21, P25, P28, and P56) is summarized in Supplementary Table 2.
Wild animals	This study did not involve wild animals.
Field-collected samples	This study did not involve samples collected from the field.

Human research participants

Policy information about studies involving human research participants

Population characteristics	Three patients and 6 healthy family members. All were Causasians, family one is of Swedish origin, and family two is of Israeli origin. The genetic variants were checked using a local population database at Clinical Genomics, KI, and using Exome Aggregation Consortium (ExAC v0.2).
Recruitment	The patients were recruited to the study at the Karolinska University Hospital, Stockholm, Sweden and at the Assuta Medical Center, Haifa, Israel as a part of research studies in human skeletal dysplasias. Clinical phenotypes were evaluated by experienced human geneticists and clinicians (GG; AN; ZB; EM; MN; EH, MW) and experts in human skeletal dysplasias (GN; EH; GG; MW). The clinical phenotype and preliminary data was presented in 13th Meeting International Skeletal Dysplasia Society (ISDS), 2017, in Brudges, Belgium and evaluated by an independent scientific committee of skeletal dysplasia experts. We therefore consider that there is no clinical selection bias in our study.

Flow Cytometry

Plots

Confirm that:

☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☒ All plots are contour plots with outliers or pseudocolor plots.

☒ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	HEK293T cells were transfected with reporter plasmids as described in Method section.
Instrument	FACS Celesta.
Software	FACSDiva for collection and FlowJo (version 10) and R for analysis.
Cell population abundance	About 20,000 cells.
Gating strategy	Gating strategies are shown in Supplementary Information.

☒ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.