Supplemental Data

Bi-allelic *CSF1R* Mutations Cause Skeletal Dysplasia

of Dysosteosclerosis-Pyle Disease Spectrum and

Degenerative Encephalopathy with Brain Malformation

Long Guo, Débora Romeo Bertola, Asako Takanohashi, Asuka Saito, Yuko Segawa, Takanori Yokota, Satoru Ishibashi, Yoichiro Nishida, Guilherme Lopes Yamamoto, José Francisco da Silva Franco, Rachel Sayuri Honjo, Chong Ae Kim, Camila Manso Musso, Margaret Timmons, Amy Pizzino, Ryan J. Taft, Bryan Lajoie, Melanie A. Knight, Kenneth H. Fischbeck, Andrew B. Singleton, Carlos R. Ferreira, Zheng Wang, Li Yan, James Y. Garbern, Pelin O. Simsek-Kiper, Hirofumi Ohashi, Pamela G. Robey, Alan Boyde, Naomichi Matsumoto, Noriko Miyake, Jürgen Spranger, Raphael Schiffmann, Adeline Vanderver, Gen Nishimura, Maria Rita dos Santos Passos-Bueno, Cas Simons, Kinya Ishikawa, and Shiro Ikegawa

Supplemental case reports

Individual A-III-1.

The affected 5-year-old boy (**Figure 1A**, A-III-1) was the only child of healthy and non-consanguineous Brazilian parents. Fetal ultrasound (US) showed dilated ventricles and hypoplasia of the cerebellum suggestive of a Dandy-Walker malformation, which was confirmed by a fetal magnetic resonance image (MRI) study. The boy was born at 37 5/7 weeks by Cesarean section, with a birth weight of 4,270 g (p.>90 centile), length of 50 cm (p.75-90th centile), occipital-frontal circumference (OFC) of 40 cm (p.>98th centile) and Apgar scores of 8/9 in the 1st and 5th minutes. Brain imaging studies showed findings compatible with a Dandy-Walker malformation, scattered periventricular calcifications, corpus callosum agenesis and abnormal signal in the periventricular white matter (**Figure 2A-C**). Ventriculo-peritoneal shunt was placed at 3 months of age. He was hypotonic and developed focal seizures. He sat without support at 13 months of age and had his first steps at 20 months of age, but since age 2.5 years he regressed motorically. At the last visit at the age of 4.5 years, he was severely dysarthric with partial head control and unable to sit without support. A skeletal survey in the first months of life showed bone abnormalities. He had poor vision with nystagmus secondary to optic nerve atrophy. G-banded karyotype and chromosomal microarray were normal.

Physical examination at age 4 years and 4 months disclosed weight of 14.5 kg (Z-score: -1.31), length of 94.5 cm (Z-score -2.81), OFC of 51 cm (50th centile); facial dysmorphisms (long eyelashes, bilateral epicanthus, bulbous nose, sagging cheeks, tented upper lip, dysplastic ears, abnormal enamel in some of his teeth); narrow, bell-shaped thorax with pectus carinatum; joint restrictions at the elbows and ankle, and dorsal kyphosis. Serum calcium, phosphorus and PTH levels were normal. ALP was 72 U/L (normal values for ages 1-9: 142-335 U/L). A skeletal survey showed diffuse osteosclerosis of the cranio-facial bones, most prominent in the skull base (**Figure 2D**). The vertebral bodies of the lumbar spine and the neural arches of the thoracic spine were mildly flat with sclerotic margins. (**Figure 2E**). The ilia, pubes and ischia showed sclerosis (**Figure 2G**). Both the long and short tubular bones showed diaphyseal sclerosis and metaphyseal radiolucency with metaphyseal under-modeling, which were most prominent in the femora (Erlenmeyer flask deformity) (**Figure 2F, H**). The skeletal changes were compatible with that of DOS.

The parents and grandparents denied bone fractures and had normal height. The father had mild cortical hyperostosis in the long tubular bones (**Figure S1A-C, E-G**). Cranial computed tomography (CT) scan of the parents revealed no brain anomaly (**Figure S1D, H**). The 76-year-old paternal grandfather showed short-term memory loss from the age of 70. His

cranial CT scan showed one parietal parenchymal calcification (**Figure S1I**). Complete blood

count (CBC) and serum markers for bone metabolism showed normal results (**Table 1**).

Individual B-II-2.

The 37-year-old Japanese female was one of the two daughters of healthy and non-

consanguineous parents (**Figure 1C**). Her birth and development were uneventful. At age 5, she

fell down and suffered coccyx fracture. A skeletal survey revealed bone abnormalities and she

was diagnosed with osteopetrosis. She had no clear motor deficits. After she graduated from

college, she worked as an office clerk for seven years until her marriage at age 33. When she

was 28, her visual acuity gradually deteriorated in the left eye. She was found to have increased

intra-cranial pressure due to type 1 Chiari malformation and received foramen magnum

decompression. At age 31, she started to drag her right foot when walking. At age 33, she started

to experience difficulties in vocalization, which made her speak only a few words. Her walking

ability gradually worsened until at age 35 when she could only walk for short distances with

support. At age 36 she visited our clinic with a chief complaint of walking difficulties. Her

height was 152.7 cm (-1.0 SD) and body weight was 42.8 kg (-1.3 SD). All extremities were

short. Neurological examinations disclosed impaired cognitive function: Mini Mental State

Examination score was 21 points, which had been 26 points one year before, suggesting

progressive deterioration. She could only walk one or two steps holding onto others. She was severely hypophonic and had generalized spastic rigidity. Tendon reflexes were diffusely hyperactive, and pathological reflexes were elicited on the right. She exhibited inappropriate laughter.

Cranial CT scan showed dilated lateral ventricle and periventricular calcifications, particularly remarkable in the anterior horn and along the genu of corpus callosum (**Figure 2I**). FLAIR-weighted horizontal MRI disclosed multiple periventricular hyperintense areas and malformation of frontal lobe (**Figure 2J**). T1-weighted sagittal MRI showed hypogenesis of cerebral cortex and corpus callosum with enlarged ventricle (**Figure 2K**). MRI of the cervical cord showed bilateral T2-hyperintensity in the lateral funiculus. Bone survey disclosed diffuse osteosclerosis in craniofacial bones, particularly of the skull base (**Figure 2L**). The neural arches of the thoracolumbar spine were sclerotic. The vertebral bodies were concaved at their posterior thirds with endplate sclerosis (**Figure 2M**). The long and short tubular bones showed metaphyseal under-modeling with metaphyseal radiolucency and diaphyseal sclerosis (**Figure 2N, P**). The iliac bodies and proximal femoral diaphyses were sclerotic (**Figure 2O**). Laboratory examination including CBC and serum markers for bone metabolism were normal (**Table 1**). The skeletal and neurologic changes were similar to those of individual A-III-1 and the mice with bi-allelic *Csf1r* mutation, but osteosclerosis and brain malformation were milder.

The proband's father, aged 70, did not show any abnormalities in neurological examination, including brain MRI and bone survey. The proband's mother, aged 69, showed scoliosis, mild short stature and a facial appearance similar to the proband; however, brain MRI and bone survey did not show any obvious abnormalities. The older sister, aged 38, has no bone abnormality. She was diagnosed with multiple sclerosis, but no obvious lesions suggesting HDLS were seen.

Individual C-III-4.

The 14-year-old male had hydrocephalus diagnosed by fetal US. He was from an extended Chaldean family of double first cousins **(Figure 1E)**. He was a full-term baby with a birth weight of 3,300 g (+0.8 SD). He was hypotonic and developed focal seizures. He had developmental delay and intellectual disability (IQ 50) with superimposed progressive deterioration in cognitive function and vocalization. He showed rigido-spasticity with hyperactive tendon reflexes and pathological reflexes. His height was 148 cm (-2.4 SD) and body weight was 50 kg (-1.1 SD).

Brain horizontal CT disclosed dilated ventricles and symmetric patterns of clustered multiple punctuate calcifications in the cerebral hemispheric white matter, particularly in the gray-white junction without post-contrast enhancement **(Figure 2Q)**. Brain horizontal MRI

showed patchy and regional T2 and FLAIR signal white matter hyperintensities particularly in

a periventricular pattern, but also involving some subcortical white matter **(Figure 2R)**. The

lateral ventricles and third ventricle were prominent suggesting brain tissue loss **(Figure 2S)**.

A skeletal survey showed milder bone sclerosis than those in Patient 1 and 2 (**Figure 2T-X**).

The cranial base showed mild sclerosis. Vertebral bodies were slightly flat at their posterior

thirds. There was mild marginal sclerosis in the vertebral and iliac bones. The tibiae were

mildly bowed. The long and short tubular bones showed metaphyseal under-modeling and

radiolucency and their diaphyseal cortices were only mildly thick. The bone changes

resembled those of Pyle disease. Iliac crest biopsies showed formation of calcified cartilage

islands in areas of woven bone (**Figure S4**). CBC and serum markers for bone metabolism

were within normal limits (**Table 1**).

Individual C-III-6.

The 23-year-old female was a cousin of individual C-III-4 **(Figure 1E)**. Her birth and

development were uneventful. She had a mild limitation in intellectual functioning (IQ 73) and

presented with a progressive deterioration in cognitive function. She had a progressive

disability in vocalization. Rigido-spasticity and hyperactive tendon reflexes were evident, and

pathological reflexes were elicited. Her height was 153.2 cm (-1.4 SD) and body weight was

45.9 kg (-0.8 SD).

Neuroimaging disclosed periventricular calcifications, dilated ventricle, and signal

hyperintensities in the white matter (**Figure S2A, B**). Skeletal imaging showed sclerotic skull,

mild platyspondyly and failure of metaphyseal modeling of tubular bones (**Figure S2C-F**),

similar to those of individual C-III-4. CBC and serum markers for bone metabolism were also

within normal limits (**Table 1**).

The Chaldean family had total five affected individuals (4 females and one male)

(**Figure 1E**). The course of the disease and clinical characteristics were similar in the five

patients with variable clinical severity. The phenotypes represented a rapidly progressive

encephalopathy with intellectual decline, pyramidal, extrapyramidal and cerebellar features,

ataxia, dysarthria, seizures and psychiatric symptoms, culminating in spastic quadriplegia,

mutism and persistent vegetative state. The onset of symptomatic neurological and skeletal

defects varied from the first decade of life in three patients and second decade in two patients.

Of the five, three attended special education, and two were delayed in walking and language

milestones. The premorbid cognitive and intellectual abilities of two affected females were

remarkable in that one completed college and one completed her first year before the

manifestations of ataxia and dysarthria. Two suffered from congenital hydrocephalus that

required ventricular peritoneal shunting. None had facial dysmorphism, oligodontia, scoliosis,

or fracture. The phenotypes were similar to those in Family B and C, but osteoscleorosis was

milder.
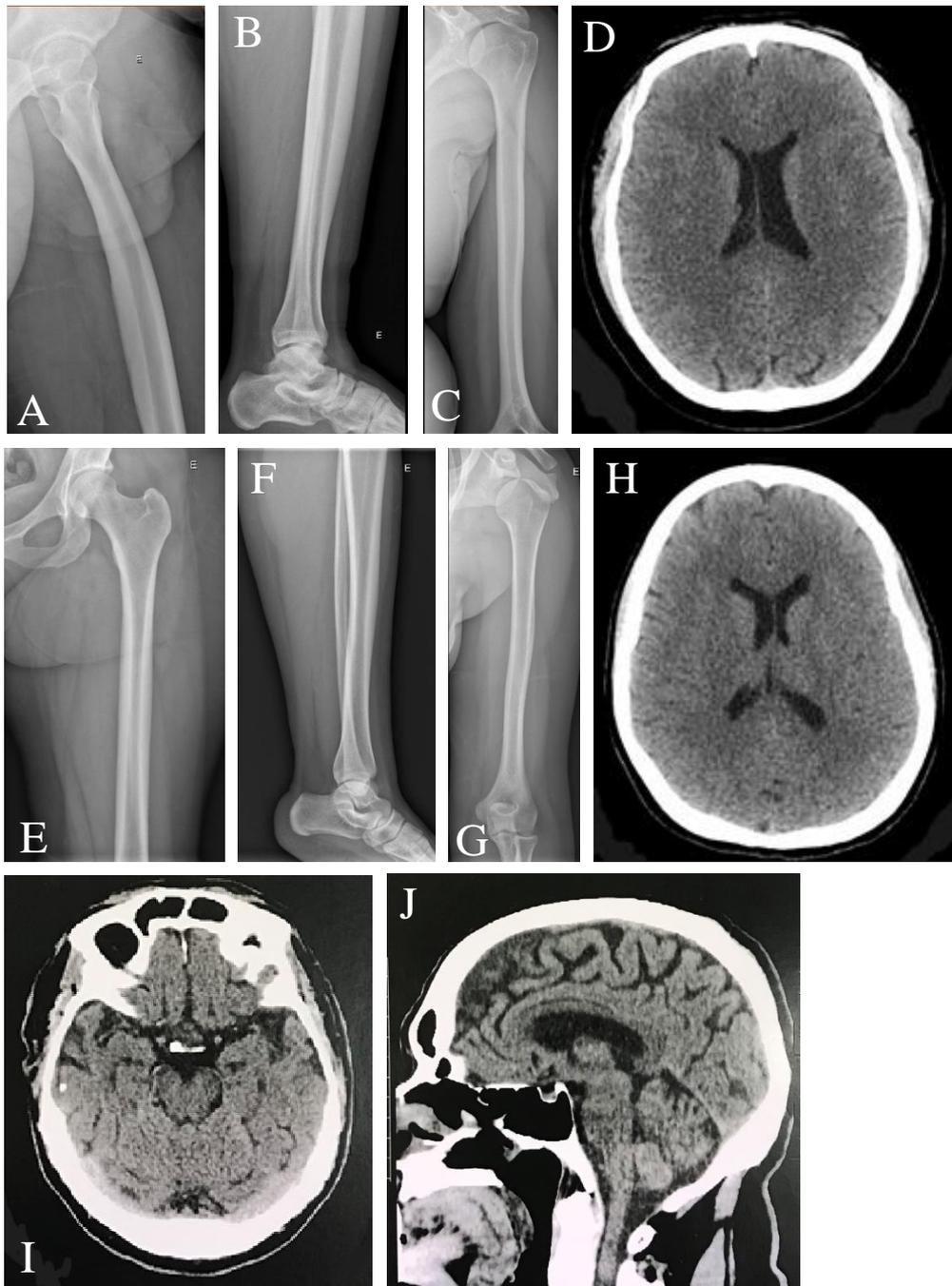
Figure S1. Brain and skeletal findings of the mutation carriers in Family A.
(A-D) The patient's father (A-II-1). Mild cortical hyperostosis in the long tubular bones of the lower extremities. No metaphyseal changes. No brain anomaly in cranial computed tomography (CT). (E-H) The patient's mother (A-II-2). Mild cortical thickening in femur. No brain anomaly in cranial CT. (I-J) The patient's paternal grandfather (A-I-1). One parietal parenchymal calcification in his cranial CT.

Figure S2. Brain and skeletal findings of Patient 4 (C-III-6) at age 23 years.
(A) Brain CT showing multiple calcification and enlargement of both lateral ventricles. (B) Brain MRI. Enlargement of both lateral ventricles and patchy hyperintensity signal in white matter. (C)-(F) Radiographs showing mild osteosclerosis. (C) Vertebral bodies are mildly concaved at their posterior thirds. Endplates show mild sclerosis. (D) Metaphyseal under-modeling in the femur. The metaphysis is radiolucent, while diaphysis is relatively sclerotic. Thin cortex in the expanded meta-diaphyseal segment. (E) Pelvic bones and proximal femora show mild sclerosis. (F) The left hand. Under-modeling, sclerosis of mid-diaphysis and sub-metaphyseal radiolucency similar to the long tubular bone.

Reference sequence You analyze your own sequence.   ☐ WT exon 14

```
  1 gctcagtgac aggctcatcc ctcctctctc tctctctgcc atctgtgtgt ctgcattttt ccttctcctt cttttggctt ctggtcactc cgggtcttgg
101 ggtatgccct gctttctccc ctgggtctct gcatttggtc cccatgtatc tgtgtggtgc tctctgtcct gccctctccc tgtctttggg actgtggttc
201 ttcctcccag CCACGGCCCA TGCTGATGAG AAGGAGGCCC TCATGTCCGA GCTGAAGATC ATGAGCCACC TGGGCCAGCA CGAGAACATC GTCAACCTTC
301 TGGGAGCCTG TACCCATGGA Ggtaagggcc ttggggttcc tggggccaag gtcttggggc ctctgg
```

Total sequence length: 366 nucleotides

Mutant sequence Substitution at position 92 (g>A)   ☐ Extended exon 14

c.1859-119G>A
↓

```
  1 gctcagtgac aggctcatcc ctcctctctc tctctctgcc atctgtgtgt ctgcattttt ccttctcctt cttttggctt ctggtcactc cAggtcttgg
101 ggtatgccct gctttctccc ctgggtctct gcatttggtc cccatgtatc tgtgtggtgc tctctgtcct gccctctccc tgtctttggg actgtggttc
201 ttcctcccag CCACGGCCCA TGCTGATGAG AAGGAGGCCC TCATGTCCGA GCTGAAGATC ATGAGCCACC TGGGCCAGCA CGAGAACATC GTCAACCTTC
301 TGGGAGCCTG TACCCATGGA Ggtaagggcc ttggggttcc tggggccaag gtcttggggc ctctgg
```

Total sequence length: 366 nucleotides

The sequences analyzed in HSF are underlined.

Sequence Position	cDNA Position	Splice site type	Motif	New splice site	Wild Type	Mutant	If cryptic site use, exon length variation	Variation (%)
82	-129	Acceptor	tggtcactccgggt	tggtcactccagGT	61.74	90.69	+117	New site +46.89

Figure S3. *In silico* **evaluation of c.1859-119G>A.**
The variant is predicted to generate a cryptic splice acceptor site by Human Splicing Finder 3.0,
which causes splicing aberrance to result in extension of exon 14.
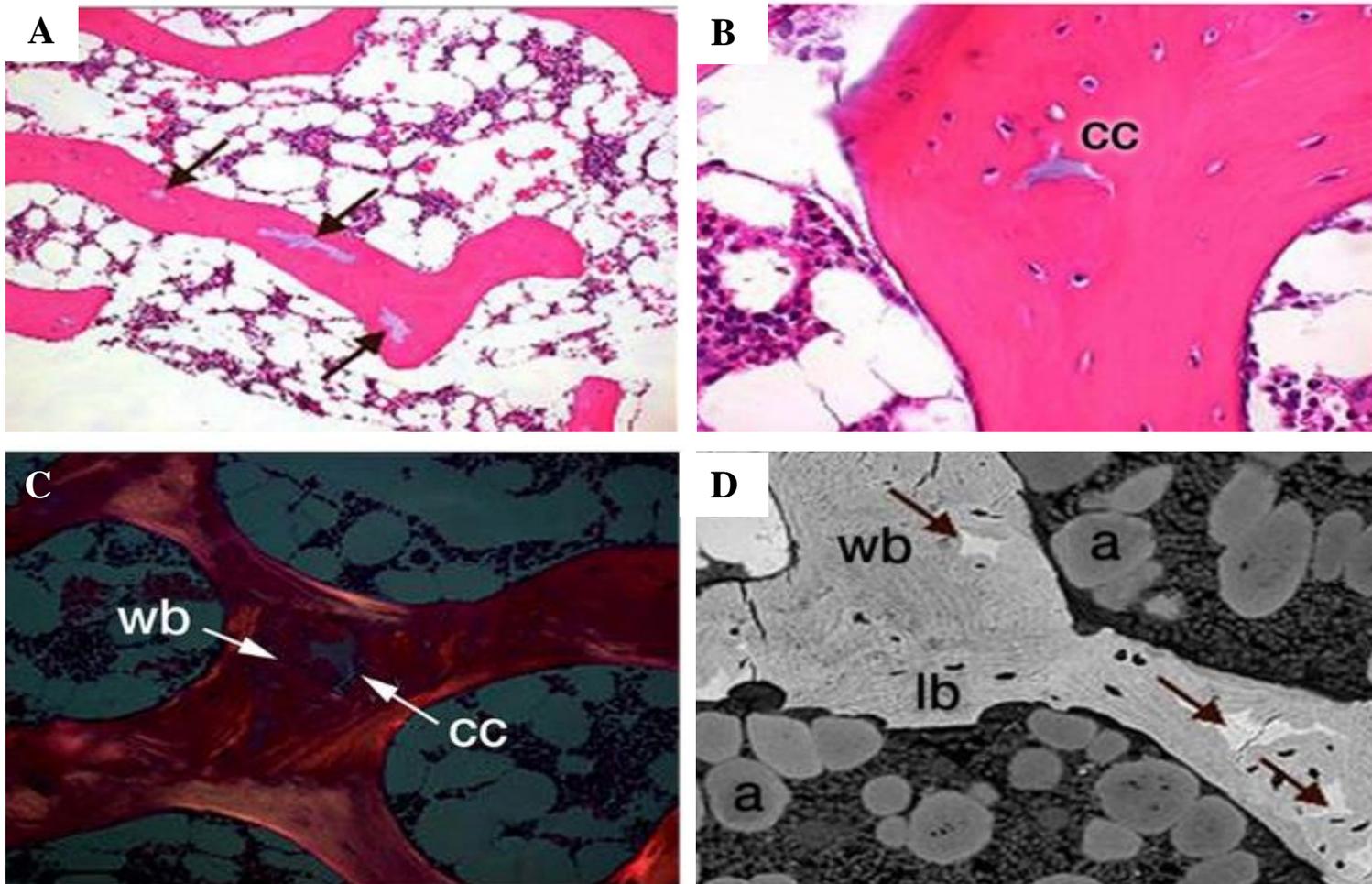
Figure S4. Trans-iliac Crest Biopsies Showing islands of calcified cartilage
(A-B) Sample sections from a patient in Family C showed sporadic islands of calcified cartilage (arrow in A, cc in B) throughout the entire length of the biopsy which was 13 mm away from the end of the growth plate. (C) The sample section under polarized light microscopy showed woven bone (wb) surrounding areas of calcified cartilage (cc). (D) By back scattered electron microscopy, islands of calcified cartilage (arrows), which were more highly mineralized than bone, were found in areas of woven bone (wb), away from areas of more lamellar bone (lb). The processing of this non-decalcified specimen also allowed for the visualization of marrow adipocytes (a).

Table S2. List of the primers used in this study

Name	Sequence (5'-3')*
Sanger sequencing	
Sanger seq confirm f1	GCTCAGTGACAGGCTCATCC
Sanger seq confirm r1	ATGGGTACAGGCTCCCAGAAG
Sanger seq confirm f2	TGTGACCTCAGATGAATGATCAGCC
Sanger seq confirm r2	TCAGGGCACTGCATTGATAGTCC
Sanger seq confirm f3	GATAGGTAAGTGGGCTCCACTCAC
Sanger seq confirm r3	GTGGACCTTGGTACTGCTAGGATC
RT-PCR	
c.1859-119G>A RT-PCR F1	ACAAGTATAAGCAGAAGCCCAAGTACC
c.1859-119G>A RT-PCR R1	CCCGGTGGATGCAATTCTTGG
c.1969+115_1969+116delAG RT-PCR F1	GTCCTGAAGGTGGCTGTGAA
c.1969+115_1969+116delAG RT-PCR R1	AGAAGTGGAGACAGGCCTCA
Plasmid construction	
CSF1R WT cds EcoR1 f1	CGCGAATTCAATGGGCCCAGGAGTTCTGCTG
CSF1R WT cds Xba1 r1	TCCTCTAGATCAGCAGAACTGATAGTTGTTGGGCTG
Inverse PCR c.395C>t f1	GAAGCAGGCGTCTCGCTGGT
Inverse PCR c.395C>t r1	CAGCACCaGGTCTGTGAGCA
Inverse PCR c.1441C>t f1	TCTAGAGGATCCCGGGTGGC
Inverse PCR c.1441C>t r1	GGTTTaGTTGTGCTCTAAGGTCTCAAC
Inverse PCR c.1879_1881del f1	GAGGCCCTCATGTCCGAGCT
Inverse PCR c.1879_1881del r1	CTCATCAGCATGGGCCGTGG
Infusion insert inverse PCR f1	TGAAGATGCTGAAGTGTCTTGGGGTATGCCCTGCT
Infusion insert inverse PCR r1	CAGCATGGGCCGTGGCTGGGAGGAAGAACCACAGTCCC
Infusion vector inverse PCR f1	CCACGGCCCATGCTGATGA
Infusion vector inverse PCR r1	ACTTCAGCATCTTCACAGCCACC
Nonsense-mediated mRNA decay	
NMD PCR f1	CCGCCTGAAGCCCTCTGAG
NMD PCR r1	TGCAGAGATGGGTATGAAGGCC
NMD Sanger seq f1	CTGCAGTGCAGTGGCCAC

*Bases for the mutations are in lower cases.

Table S3. Throughput of the whole exome sequencing (WES) and whole genome sequencing (WGS)

Method	Sample	Read		≥ 10 reads coverage
		Total (Gbp)	Mean depth	(%)
WES	Patient A-III-1	4.9	×80.3	97.9
	Patient A-II-1	6.6	×108.8	98.6
	Patient A-II-2	4.8	×78.5	97.4
	Patient B-II-1	2.2	×66.4	97.2
WGS	Patient C-III-1	126	×29.9	97.2
	Patient C-III-3	128	×30.0	97.3
	Patient C-III-4	132	×29.5	97.4

Calculated based on the entire coding regions or the entire genomic sequence of human genome browser, hg19.

Supplemental methods

Whole exome sequencing and variant calling

The study protocol was approved by the ethical committee of RIKEN and participating institutions. Peripheral blood was obtained from the family members after the informed consent. Genomic DNA was extracted from the blood using QIAamp DNA Blood Midi Kit (Qiagen, Germany). DNA concentration was measured by using a Qubit V.2.0 Fluorometer (Life Technologies, USA).

Whole exome sequencing was performed as previously described.[1–3] Briefly, DNA (3 μg) was sheared with S2 Focused-ultrasonicator (Covaris, Wobum, USA) and processed by SureSelectXT Human All Exon V5 or V6 (Agilent Technologies, USA). Captured DNA was sequenced using HiSeq 2000 (Illumina, USA) with 101-bp paired-end reads or HiSeq 2500 with 150-bp paired-end reads. Image analysis and base calling were performed using HCS, RTA and CASAVA software (Illumina, USA). Reads were mapped to the reference human genome (hg19) by Novoalign-3.02.04 or BWA-mem.[4] Aligned reads were processed by Picard to remove PCR duplicates. Variants were called by GATK v2.7-4 following GATK Best Practice Workflow v3 (ref. 5) and annotated by ANNOVAR.[6]

PCR and Sanger sequencing

The identified *CSF1R* variants and their transmission in the pedigrees were confirmed by Sanger sequencing. The genomic segment harboring the variants were amplified by PCR and sequenced both strands. The PCR primer sets is shown in **Table S1**. A 3730 DNA analyzer (Life Technologies, USA) was used for the Sanger sequencing. Sequencher V.4.7 (Gene Codes, USA) and Genetyx (Genetyx, Japan) were used for aligning sequencing chromatographs to reference sequences.

Evaluation of the mutation identified in *CSF1R*

The variants were evaluated by using four databases, dbSNP, 1000 genomes, gnomAD, and HGMD. The potential pathogenicity in splicing was evaluated by several online prediction programs including HSF, NNSPLICE, SplicePort, ASSP, and MaxEntScan.

Reverse transcription polymerase chain reaction (RT-PCR)

Total RNA was extracted from the blood (Family A), an immortalized patient-derived lymphoblastoid cell line (Family B), and patient-derived fibroblasts using SV Total RNA Isolation System (Promega, USA). Total RNAs (500 ng) were used to synthesize cDNA with a PrimeScript RT reagent Kit (Takara Bio, Japan). The cDNA region of the heterozygotes for p.Gln481* in Family A including partial exon8, exon9, and partial exon10 was amplified by PCR. The cDNA region of Patient B-II-1 bracing partial exon 12, exon 11-17, and partial exon 18 was amplified by PCR. The cDNA region of Patient C-III-4 and C-III-6 bracing partial exon

13, exon 14, exon 15 and partial exon 16 was amplified by PCR. All the primers were shown in **Table S1**. The electrophoretic bands of the PCR products were cut from the gel, purified, and sequenced using the PCR primers.

Construction of expression plasmids

A full-length cDNA clone of human *CSF1R* (GenBank: NM005211) was PCR-amplified using KOD plus (Toyobo, Japan). The PCR amplicons were cloned into the *EcoR*I and *Xba*I sites of the pFLAG CMV-4 expression vector (Sigma–Aldrich, USA). The mutations were generated by inverse PCR-based site-directed mutagenesis kit (Toyobo, Japan) and In-Fusion HD Cloning Kit (Takara Bio, Japan). All the PCR primers are shown in **Table S1**.

Cell culture and transfection

The HEK293 cells were cultured in Dulbecco's Modified Eagle Medium (DMEM) supplemented with 10% fetal bovine serum (FBS). Cells were seeded in a six-well plate and transfected with FuGENE 6 Transfection Reagent (Promega, USA) according to the manufacturer's protocol. One day after transfection, cells were serum starved (0.2% FBS) for 24 hours, and treated with recombinant human macrophage colony-stimulating factor (M-CSF; R&D Systems, USA).

Western blot

The HEK293 cells were treated with M-CSF (20 ng/ml) for 24 min, and harvested with 200 μl of RIPA Lysis and Extraction Buffer (Thermo Fisher Scientific, USA). Western blots were performed by standard procedures using anti-CSFR1 antibody (HPA012323, Sigma-Aldrich, USA), anti-JNK antibody (#9252, Cell signaling Technology, USA), anti-phospho-JNK antibody (#9251, Cell signaling Technology, USA), anti-β-Actin-HRP antibody (PM053-7, MBL, Japan), and anti-rabbit IgG–peroxidase antibody (A0545, Sigma-Aldrich, USA). Specifically, cells were washed using ice-cold PBS after discarding the medium, and then scraped in ice-cold lysis buffer using cold plastic cell scraper. Cells were collected in microfuge tubes and agitated for 30 min at 4°C. After centrifuging the tubes at 16000G for 20 min at 4°C, collect the supernatant in fresh tube and place on ice. The total protein concentration was measured using BCA Protein Assay Kit (Thermo Fisher Scientific, USA). Cell lysate was mixed with loading buffer and boiled at 95 °C for 5 mins. Load equal amounts of protein into the wells of SDS-PAGE gel, along with molecular weight markers. Run the gel for 30 min at 200 V. Place the gel into the transfer sandwich and make sure no air bubbles are trapped in the sandwich. Place the cassette in the transfer tank and place an ice block in the tank. Transfer to PVDF memebrane for one hour at 350mA. Wash the membrane with TBST and block it in 3% skim milk for two hours. Incubate the membrane with the anti-phospho-JNK antibody (1/500) for three hours at room temperature. After washing with TBST,

incubate the membrane with anti-rabbit IgG–peroxidase antibody (1/20000) for one hour.

Apply the chemiluminescent substrate from ECL^{TM} Prime Western Blotting Detection

Reagents（GE Healthcare, UK）to the blot according to the manufacturer's recommendation.

Capture the chemiluminescent signals using films. Use image analysis software (ImageJ) to

read the band intensity. The phosphorylation state was quantified as the ratio between

phosphorylated and total levels of JNK. Results are expressed as arbitrary units where the

mean of the ratio in wildtype group is set at 1. Data were presented as mean ± S.E.M. of three

independent experiments. Statistical significance was tested with one-way ANOVA by SPSS

13.0 for Windows. *P*<0.05 was considered statistically significant.

Supplemental References

1.   Wang, Z. *et al.* Axial spondylometaphyseal dysplasia is also caused by NEK1

mutations. *J. Hum. Genet.* **62,** 503–506 (2017).

2.   Guo, L. *et al.* Dysosteosclerosis is also caused by TNFRSF11A mutation. *J. Hum.*

Genet. **63,** 769–774 (2018).

3.   Guo, L. *et al.* Identification of biallelic EXTL3 mutations in a novel type of spondylo-

epi-metaphyseal dysplasia. *J. Hum. Genet.* **62,** 797–801 (2017).

4.   Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler

transform. *Bioinformatics* **25,** 1754–1760 (2009).

5.   McKenna, A. *et al.* The Genome Analysis Toolkit: a MapReduce framework for

analyzing next-generation DNA sequencing data. *Genome Res.* **20,** 1297–303 (2010).

6.   Wang, K., Li, M. & Hakonarson, H. ANNOVAR: functional annotation of genetic

variants from high-throughput sequencing data. *Nucleic Acids Res.* **38,** e164 (2010).