Supplementary Methods

Subjects and clinical evaluation. We obtained blood samples following informed consent from individuals and/or their parents in accordance with a protocol approved by the Institutional Review Board of the University of Pennsylvania School of Medicine. More detailed descriptions of the families are provided in **Supplementary Notes** online.

Linkage studies. We isolated genomic DNA directly from blood samples, buccal swabs, or from lymphoblast cell lines (LCLs) using QIAamp DNA Blood reagents (Qiagen) and standard protocols. EBV-transformed LCLs were established by standard protocols. Some samples were whole genome-amplified using Repli-G reagents (Molecular Staging Inc.) and standard protocols. We conducted genome wide linkage analysis through the University of Utah School of Medicine Genomics Core Facility using an ABI 3130xl Genetic Analyzer and scoring with ABI GeneMapper v.3. All families were genotyped using 400 microsatellite markers from the ABI Prism linkage mapping set v.2.5 (medium density 10 cM set). For fine mapping, we used markers selected from the Marshfield genetic map. Family A genotyping data from an earlier study[1] was integrated into this analysis. We used Genehunter v.2 to calculate multipoint parametric LOD scores. A model of an autosomal dominant trait with 100% penetrance of the FOP gene was assumed, with a population prevalence of FOP of 1 per million. SNP genotyping was conducted through the University of Pennsylvania School of Medicine Microarray Core Facility using the Affymetrix GeneChip Mapping 10K Array and Genespring GT software (Agilent Technologies).

Mutation analysis. We screened for mutations in *ACVR1* by PCR-amplification of genomic DNA from blood or LCLs corresponding to the 9 exons containing protein coding sequences (ACVR1 Transcript Report, Ensembl v35), using exon-flanking primers (see **Supplementary Table 1** online for primer sequences). We conducted DNA sequence analysis of genomic DNA on an ABI3730 sequencer through the University of Pennsylvania School of Medicine DNA Sequencing Facility. Sequence data were analyzed using 4Peaks software v.1.6 (http://www.mekentosj.com/4peaks/).

Differences in restriction endonuclease recognition sites were identified using MacVector v.7.2 software (ABI). We amplified 0.1 ug of genomic DNA using primers for protein coding exon 4 (**Supplementary Notes** and **Supplementary Table 1** online). Following agarose gel electrophoresis, we recovered the PCR products (350 bp) from agarose using QIAquick Gel Extraction reagents (Qiagen). Purified PCR product was digested with either *Hph*I (5 U/ul) or *Cac*8I (4 U/ul) (both from New England Biolabs) at 37°C for 2 hours. Fragments were resolved on 3% NuSieve 3:1 agarose (FMC BioProducts) gels with 100 bp ladder (New England Biolabs) as size markers.

Cell culture and RNA analysis. LCLs from 4 FOP patients and 4 controls were grown in RPMI 1640 media with 15% FBS. We extracted total RNA from 10^7 cells using RNeasy reagents (Qiagen) and performed reverse transcription using SuperScript III (Invitrogen). We PCR-amplified the region corresponding to protein coding exon 4 with specific primers (see

Supplementary Table 1 online for primer sequences) and Taq DNA polymerase (Invitrogen) then directly sequenced the amplified cDNA as described above.

Molecular modeling of protein structure.

Structural protein homology modeling was based on the PDB structure for type I TGFβ receptor kinase which is 66% identical to ACVR1 residues 178-498.[2,3] This region includes the serine/threonine protein kinase catalytic domain and the GS motif with arginine residues at ACVR1 positions 202 and 206. ACVR1 amino acid 178-498 sequence was submitted to PredictProtein (http://www.embl-heidelberg.de/predictprotein/submit_def.html), CPHmodels homology-modeling server (http://www.cbs.dtu.dk/services/CPHmodels/) and the SWISS-MODEL homology-modeling server (http://swissmodel.expasy.org/). Visualization used the DeepView Swiss PDB Viewer.

Gene and protein analysis. Genes in the linked region were identified through the National Center for Biotechnology Information Entrez Map Viewer and the UCSC Genome Browser. We obtained the intron-exon boundaries of the *ACVR1* gene through GenBank, Ensembl Human Genome Server, and the University of Santa Cruz. Genomic DNA positions of markers and the *ACVR1* gene are from the UCSC Genome Browser (May 2004, Build 35). Transcript and exon information is from Ensembl (Gene ID ENSG00000115170; transcript ID ENST00000263640) which reports 11 exons for ACVR1 (exons 1 and 2 contain only 5'UTR; the protein start site is in exon 3), consistent with GenBank BC033867, full length cDNA clone). All databases are consistent for the sequence information for the 9 exons containing protein-coding sequences, however, additional/alternate exons containing 5'UTRs are reported. ACVR1 protein ID is

Q04771 (Pfam, SWISSPROT). We used Clustal W (with the MacVector v 7.2 software program) for multiple protein alignment using sequences from GenBank.

GenBank Accession numbers.

ACVR1 cDNA, NM_001105; ACVR1 coding region, NT_005403. ACVR1 protein for Homo sapiens (human), NP_001096; Mus musculus (mouse), NP_031420; Rattus norvegicus (rat), NP_077812; Canis familiaris (dog), XP_856152; Bos taurus (cow), NP_788836; Gallus gallus (chick), NP_989891; Xenopus laevis (frog), AAH88947; Danio rerio (zebrafish), NP_571420. Fugu rubripes (pufferfish) sequence was from Ensembl prediction SINFRUG00000134562.

References

1. Feldman, G. et al. Fibrodysplasia ossificans progressiva, a heritable disorder of severe heterotopic ossification, maps to human chromosome 4q27-31. *Am. J. Hum. Genet.* **66**, 128-135 (2000).
2. Huse, M., Chen, Y.G., Massagué, J. & Kuriyan, J. Crystal structure of the cytoplasmic domain of the type I TGF beta receptor in complex with FKBP12. *Cell* **96**, 425-436 (1999).
3. Huse, M. et al. The TGF beta receptor activation process: an inhibitor-to substrate-binding switch. *Molecular Cell* **8**, 671-682 (2001).