RESEARCH

A Nonsense Mutation in the Cathepsin K Gene Observed in a Family with Pycnodysostosis

Maureen R. Johnson,[1] Mihael H. Polymeropoulos,[2] Hans L. Vos,[3] Rosa Isela Ortiz de Luna,[4] and Clair A. Francomano[1,5]

[1]Medical Genetics Branch and [2]Laboratory of Genetic Disease Research, National Center for Human Genome Research, National Institutes of Health, Bethesda, Maryland 20892-1852; [3]Netherlands Cancer Institute, Amsterdam, The Netherlands; [4]Hospital Infantil de Mexico "Federico Gomez," Mexico City, D.F. Mexico

Pycnodysostosis (MIM 265800) is a rare, autosomal recessive skeletal dysplasia characterized by short stature, wide cranial sutures, and increased bone density and fragility. Linkage analysis localized the disease gene to human chromosome 1q21, and subsequently the genetic interval was narrowed to between markers DIS2612 and DIS2345. Expressed sequence tagged markers corresponding to cathepsin K, a cysteine protease highly expressed in osteoclasts and thought to be important in bone resorption, were mapped previously in the candidate region. We have identified a cytosine to thymidine transition at nucleotide 862 (GenBank accession no. S79895) of the cathepsin K coding sequence in the DNA of an affected individual from a large, consanguinous Mexican family. This mutation results in an arginine to STOP alteration at amino acid 241, predicting premature termination of cathepsin K mRNA translation. All affected individuals in this family were homozygous for the mutation, suggesting that this alteration may lead to pycnodysostosis. Recognition of the role of cathepsin K in the etiology of pycnodysostosis should provide insights into the pathogenesis and treatment of other disorders of bone remodeling, including osteoporosis.

Pycnodysostosis (MIM 265800) is a rare, autosomal recessive skeletal dysplasia characterized by short stature, brachycephaly, wide cranial sutures, osteosclerosis, and bone fragility. The phenotype is inherited with complete penetrance and many individuals affected with pycnodysostosis are the products of consanguinous matings. In 1965 it was suggested that the French Impressionist artist Henri de Toulouse-Lautrec was affected with the disorder (Maroteaux and Lamy 1965), leading to a lively debate questioning this diagnosis (Frey 1995a,b; Maroteaux 1995). The gene encoding the phenotype was mapped to human chromosome 1q21 by genetic linkage analysis (Gleb et al. 1995a; Polymeropoulos et al. 1995), and subsequent studies narrowed the localization to a 1-cM region (Gleb et al. 1995b; this report). Homology searches of expressed sequence tags (ESTs) mapped within the pycnodysostosis interval to DNA sequences of characterized genes identified two markers corresponding to the 5′ and the 3′ end of the cathepsin K gene. This cysteine protease is highly expressed in osteoclasts and cleaves collagen and osteonectin, implicating it as an important protease involved in bone resorption (Bromme and Okamoto 1995; Inaoka et al 1995; Li et al. 1995; Shi et al. 1995; Bossard et al. 1996; Drake et al. 1996). Thus, cathepsin K was an excellent candidate gene for pycnodysostosis because of its genomic localization and biological properties. We now report additional genetic mapping of the pycnodysosotosis locus and identify a mutation in cathepsin K, likely causing the disease phenotype.

[5]Corresponding author.
E-MAIL clairf@nchgr.nih.gov; FAX 301-496-7157.

RESULTS

We have used genetic linkage and haplotype analysis to narrow the region harboring the pycnodysostosis gene. A series of 12 genetic markers spanning the candidate interval from marker D1S514 to D1S305 on human chromosome 1q21 was analyzed in 24 members of a consanguinous Mexican pedigree, including 10 affected indi-

6:1051–1055 ©1996 by Cold Spring Harbor Laboratory Press ISSN 1054-9803/96 $5.00

viduals (Fig. 1). Recombinant events within the family define the interval to between markers D1S2612 and D1S2715. All affected individuals were homozygous for markers D1S498, D1S2347, D1S2345, D1S2343, FLG, LOR, and D1S1664. Data published previously (Gleb et al. 1995b) placed the boundaries of the disease locus between markers D1S2344 and D1S2345/D1S2343, moving the telomeric boundary of the candidate interval to markers D1S2345/D1S2343. The highest LOD score in our study was observed for marker D1S498 with $Z = 9.98$ at $\theta = 0.00$ (Table 1).

A search for expressed sequences in the candidate region identified two partially sequenced cDNAs (GenBank accession nos. T67463 and G07268) that in homology sequence searches showed 100% identity with regions of the cathepsin K cDNA (GenBank accession nos. S79895, U13665, X82153, and U20280), an osteoclast-specific cysteine protease. In particular, marker WI12245 (GenBank accession no. T67463) has been radiation hybrid mapped to the pycnodysostosis interval (Hudson et al. 1995; Dib et al. 1996), and corresponds to cathepsin K coding sequence. Genomic DNA sequence analysis of the cathepsin K gene derived from an affected individual identified a cytosine (C) to thymidine (T) transition at nucleotide 862 (GenBank accession no. S79895), changing an in frame codon from CGA to TGA (Fig. 2). The mutation substitutes a STOP codon for an arginine at amino acid 241, predicting the premature termination of cathepsin K mRNA translation. The C to T transition occurs at a CG dinucleotide, a known mutational "hot spot" (Cooper and Krawczak 1990). Premature termination of translation at codon 241 would result in the deletion of the carboxy-terminal 89 amino acids, or 42% of the mature, catalytically active form of the enzyme.

The change from a C to a T at nucleotide 862 creates a *Dde*I restriction site, facilitating mutational analysis at the cathepsin K locus. *Dde*I digestion of a 166-bp amplified fragment including

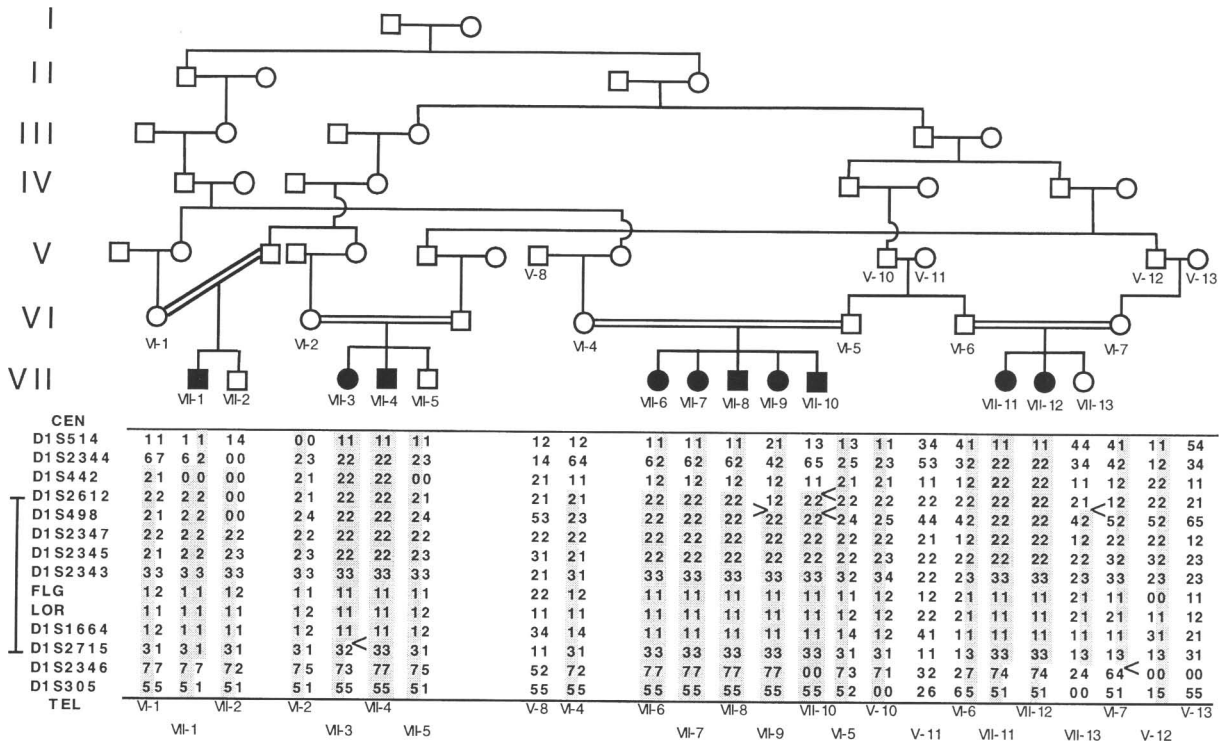

Figure 1 Haplotype analysis of the members of the pycnodysostosis family. Filled symbols represent affected individuals; unfilled symbols represent unaffected individuals. Double lines represent consanguinous matings. Numbers used to identify family members are those individuals whose DNA were analyzed. The shaded area of the haplotypes represents the ancestoral allele segregating with the disease. Arrowheads denote location of recombination events. The black bar on the *left* indicates the localization of the pycnodysostosis interval. Markers D1S2345 and D1S2343 were described previously (Gleb et al. 1995b) as AFMc002wf5 and AFMa046wd9.

Table 1. Two Point LOD Scores Between Chromosome 1 Markers and the Pycnodysostosis Locus

	Recombination fraction (θ)								
Locus	0.00	0.01	0.05	0.10	0.20	0.30	0.40	Z_{max}	θ_{max}
D1S514	-Infin	4.00	4.90	4.80	3.90	2.70	1.30	4.91	0.065
D1S2344	-Infin	3.61	4.96	5.00	4.12	2.82	1.39	5.06	0.075
D1S442	-Infin	2.84	3.70	3.68	2.99	2.01	0.98	3.74	0.070
D1S2612	-Infin	4.67	4.84	4.47	3.37	2.14	0.96	4.89	0.033
D1S498	9.98	9.79	9.04	8.07	6.06	3.98	1.93	9.98	0.001
D1S2347	4.50	4.41	4.03	3.55	2.57	1.59	0.71	4.50	0.001
D1S2345	7.34	7.19	6.58	5.80	4.19	2.58	1.13	7.34	0.001
D1S2343	8.23	8.08	7.48	6.69	5.05	3.34	1.63	8.23	0.001
FLG	5.79	5.67	5.15	4.50	3.20	1.96	0.87	5.79	0.001
LOR	6.80	6.66	6.10	5.38	3.94	2.52	1.19	6.80	0.001
D1S1664	6.77	6.66	6.17	5.52	4.11	2.66	1.26	6.77	0.001
D1S2715	-Infin	5.59	6.25	5.94	4.66	3.09	1.48	6.25	0.048
D1S2346	-Infin	4.99	5.62	5.36	4.21	2.79	1.34	5.63	0.050

the mutation results in two restriction fragments of 102 and 64 bp in all affected individuals in the family (Fig. 3). Obligate heterozygotes have an additional fragment of 166 bp, corresponding to the normal, undigested allele. *Dde*I restriction analysis of PCR products derived from 105 unaffected, unrelated individuals (210 chromosomes) from the U.S. Caucasian population revealed only the 166-bp fragment, indicating that this alteration is unlikely to be a neutral polymorphism (data not shown).

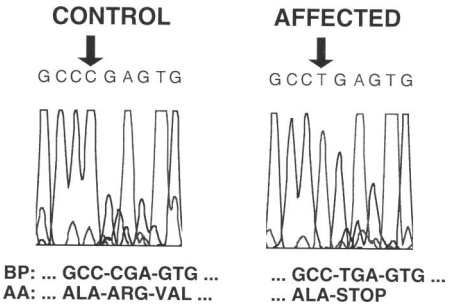

Figure 2 Cathepsin K DNA sequence analysis. The 166-nucleotide PCR product derived from affected individual VII-3 and a control sample were sequenced. The pycnodysostosis individual (affected) shows a T at the first nucleotide of codon 241 (arrow), whereas a C is seen in the control sample corresponding to the wild-type sequence. This mutation changes an arginine codon to a termination codon, as depicted in the *lower* part of the figure.

DISCUSSION

We have identified a mutation in the cathepsin K

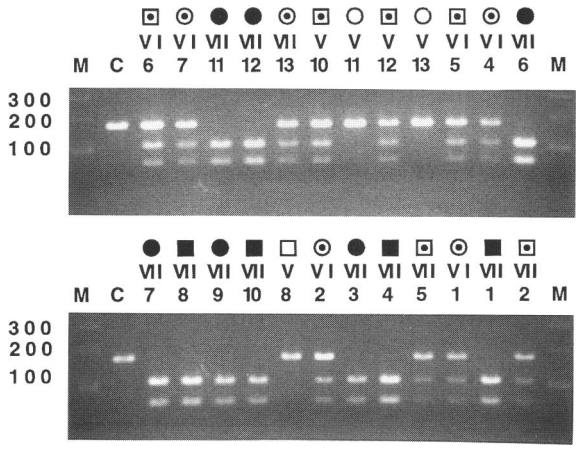

Figure 3 *Dde*I restriction endonuclease analysis. M denotes 100-bp marker (Life Technologies, Inc.), and C is a control DNA sample. Numbering of family members corresponds to Fig. 1. Undigested 166-nucleotide fragment is derived from noncarrier individual DNA (open symbols). DNA samples that showed both the undigested 166-nucleotide fragment and the digested products of 102 and 64 nucleotides is derived from individuals who are carriers of the mutant allele (dotted symbols). The presence of only the digested products of 102 and 64 nucleotides in affected individuals demonstrates autosomal recessive transmission in this consanguinous family (filled symbols).

gene resulting in an arginine to STOP alteration at amino acid 241, predicting premature termination of cathepsin K mRNA translation. Cathepsin K is a member of the papain protease family known to have a conserved triad of amino acids involved in proteolytic catalysis, at amino acids CYS_{139}, HIS_{276}, and ASN_{296} (Bossard et al. 1996). Therefore, deletion of 89 carboxy-terminal amino acids will disrupt the catalytic site, because of the lack of two of the three amino acids necessary for catalytic activity, as well as lack of two of the six cysteines thought to be important in the formation of intramolecular disulfide bonds (Bossard et al. 1996). In the highly related cysteine protease cathepsin L, a single amino acid substitution in the most carboxy-terminal, conserved cysteine residue, as well as the deletion of the carboxy-terminal 22 amino acids, resulted in almost complete loss of activity (Smith and Gottesman 1989). This provides further evidence that the deletion of the carboxy-terminal 89 amino acids in cathepsin K in this family with pycnodysosotosis would produce an inactive protease. While this manuscript was under review, the R241X mutation as well as two additional cathepsin K mutations in pycnodysostosis families were reported (Gleb et al. 1996). These mutations confirm the etiological role of cathepsin K alterations in this disease.

Skeletal tissue normally undergoes cycles of remodeling resulting from osteoblast deposition of new bone matrix and osteoclast resorption of tissue. Many studies have speculated that cysteine proteases may be involved in the degradation of the protein matrix because these enzymes digest collagen (Maciewicz and Etherington 1988), which composes up to 90% of the skeletal matrix. The critical protease necessary for proper degradation of the matrix during bone resorption may be the cysteine protease cathepsin K, which is highly expressed in osteoclasts and cleaves collagen and osteonectin (Bromme and Okamoto 1995; Inaoka et al 1995; Li et al. 1995; Shi et al. 1995; Bossard et al. 1996; Drake et al. 1996). Our finding of a mutation in the cathepsin K gene in a family with pycnodysostosis further supports the role of cathepsin K in bone remodeling. Investigation into abnormal cathepsin K activity is warranted in the analysis of other bone disorders where anomalies of bone resorption are considered. These diseases include osteoporosis, which could result from excess cathepsin K activity or increased susceptibility of the collagen matrix to enzymatic degradation. Cysteine protease inhibitors have been shown to prevent osteoclast-specific bone resorption (Delaisse et al. 1980, 1984; Lerner and Grubb 1992; Hill et al. 1994), suggesting the use of cathepsin K inhibitors as a therapeutic strategy for the treatment of osteoporosis, a significant health risk to postmenopausal women as well as elderly men.

METHODS

Clinical Samples

As reported previously (Polymeropoulos et al. 1995), a large, consanguinous family with pycnodysostosis was ascertained after the proband of the family was evaluated for short stature at the Hospital Infantil de Mexico. The clinical features leading to the diagnosis of pycnodysostosis included frequent fractures, short stature, frontal prominence, persistence of open fontanelles, facial hypoplasia, small mandible, dental malocclusion, and increased bone density. Informed consent was obtained from all individuals.

Genotype Analysis

Microsatellite markers were obtained from the Genethon collection of markers (Dib et al. 1996). Polymorphisms within the profilaggrin gene, FLG (Gan et al. 1990), and the second glycine loop domain of the loricrin gene, LOR (Yoneda et al. 1992) were also analyzed in this family. One primer from each marker pair was radiolabeled with [γ-^{32}P]ATP using T4 polynucleotide kinase (Life Technologies). Fifty nanograms of genomic DNA was amplified in a total volume of 10 µl using 0.6 pmole of the radiolabeled primer, 4 pmole of the nonradiolabeled primer, 200 µM of each dNTP, and 0.5 units of *Taq* polymerase (Perkin Elmer) in a buffer containing 10 mM Tris (pH 8.3), 50 mM KCl, 1.5 mM $MgCl_2$, and 0.01% gelatin. PCR was performed in a Hybaid OmniGene thermocycler using the following parameters: one cycle denaturing at 94°C for 8 min; 30 cycles denaturing at 94°C for 30 sec, annealing at 55°C for 30 sec, and extending at 72°C for 30 sec, followed by a final extension at 72°C for 10 min. Ten microliters of a formamide-containing loading buffer was added, the samples were heated for 5 min at 95°C, then 6 µl of the sample was electrophoresed on a 6% denaturing polyacrylamide gel. The dried gel was exposed to X-ray film from 1 to 24 hr.

DNA Sequence Analysis

Primer pairs were designed from the cDNA sequence (Bromme and Okamoto 1995; Inaoka et al. 1995; Li et al. 1995; Shi et al. 1995; Drake et al. 1996) to amplify regions of the cathepsin K gene. The genomic structure of the highly related cysteine protease cathepsin L (Chauhan et al. 1993) was used to infer the genomic structure of cathepsin K, which is not published. It should be noted that no mutations occurring within the PCR primer sequence will be identified with these primer pairs, nor will any splice alterations be detected, as the primers were con-

structed from coding sequence. The following primers were synthesized (the number succeeding the name is the nucleotide number from GenBank accession no. S79895): pair A, CK1F (101), 5′-GCTGCCGAAACGAAGCCAGAC-3′, and CK17R (261), 5′-CTTGTTGTTATATTGCTTCCTGTGGG-3′; pair B, CK18F (262), 5′-GTGGATGAAATCTCTCGGCG-3′, and CK18R (384), 5′-CATGTCCCCCAGGTGGTTC-3′; pair C, CK4F (314), 5′-CCATCCATAACCTTGAGGCTTC-3′, and CK5R (482), 5′-CTACCTTCCCATTCTGGG-3′; pair D, CK5F (384), 5′-GACCAGTGAAGAGGTGGTTC-3′, and CK9R (754), 5′-CCACATATGGGTAGGCATC-3′; pair E, CK20F (541), 5′-GGTCAGTGTGGTTCCTGTTG-3′, and CK20R (759), 5′-CTGTCCCACATATGGGTAG-3′; pair F, which detected the mutation, CK21F (760), 5′-GAAGAGAGTTG-TATGTACAACCC-3′, and CK22R (925), 5′-CTTTGCTG-TAAAACTGGAAGG-3′; pair G, CK23F (926), 5′-GTG-TGTATTATGATGAAAGCTGC-3′, and CK23R (1031), 5′-CTGTTTTTAATTATCCAGTGC-3′; and pair H, CK15F (1029), 5′-CAGCTGGGGAGAAAACTGGGG-3′, and CK15R (1190), 5′-CGTTACACTGCACCATCGTGG-3′.

Standard conditions for the PCR consisted of amplifying 750 ng of genomic DNA in a total volume of 200 µl using 80 pmole of primers, 200 µM of each dNTP, and 10 units of *Taq* polymerase (Perkin Elmer) in a buffer containing 10 mM Tris (pH 8.3), 50 mM KCl, 1.5 mM $MgCl_2$, and 0.01% gelatin, except that the final concentration of magnesium chloride in primer pair C reaction was 2.0 mM. PCR was performed in a Hybaid OmniGene thermocycler using the cycle profile described above, with the following annealing temperatures: primer pairs D, E, and F, 52°C; primer pairs B, C, and G, 55°C; and primer pairs A and H, 58°C. The PCR products were gel purified (Promega Wizard PCR preps) then sequenced with the amplification primers using an Applied Biosystems 373A DNA sequencer. Of the entire cathepsin K coding sequence, 86% was analyzed from DNA from an affected individual, and no additional base pair changes were identified when compared with the control sample.

*Dde*I Restriction Endonuclease Analysis

Genomic DNA (250 ng) was PCR amplified in a total volume of 50 µl as described above. Ten microliters of PCR product was digested with 10 units of *Dde*I (Life Technologies, Inc.) and electrophoresed on a 1% agarose, 2% NuSieve GTG agarose (FMC BioProducts) gel.

ACKNOWLEDGMENTS

We thank Jinny K. Szabo, Christiane M. Robbins, and Beth Helmbold for technical assistance, and Drs. Paul Bornstein, Iain McIntosh, Douglas Wilkin, and Fred Drake for their helpful discussions.

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC section 1734 solely to indicate this fact.

REFERENCES

Bossard, M.J., T.A. Tomaszek, S.K. Thompson, B.Y. Amegadzie, C.R. Hanning, C. Jones, J.T. Kurdyla, D.E. McNulty, F.H. Drake, M. Gowen, and M.A. Levy. 1996. Proteolytic activity of human osteoclast cathepsin K. *J. Biol. Chem.* **271:** 12517–12524.

Bromme, D. and K. Okamoto. 1995. Human cathepsin O2, a novel cysteine protease highly expressed in osteoclastomas and ovary molecular cloning, sequencing and tissue distribution. *Biol. Chem. Hoppe-Seyer* **376:** 379–384.

Chauhan, S.S., N.C. Popsecu, D. Ray, R. Fleischmann, M.M. Gottesman, and B.R. Troen. 1993. Cloning, genomic organization, and chromosomal localization of human cathepsin L. *J. Biol. Chem.* **268:** 1039–1045.

Cooper, D.N. and M. Krawczak. 1990. The mutational spectrum of single base pair substitutions causing human genetic disease: Patterns and predictions. *Hum. Genet.* **85:** 55–74.

Delaisse, J.M., Y. Eeckhout, and G. Vaes. 1980. Inhibition of bone resorption in culture by inhibitors of thiol proteinases. *Biochem. J.* **192:** 365–367.

———. 1984. In vivo and in vitro evidence for the involvement of cysteine proteinases in bone resorption. *Biochem. Biophys. Res. Commun.* **125:** 441–447.

Dib, C., S. Faure, C. Fizames, D. Samson, N. Drouot, A. Vignal, P. Millasseau, S. Marc, J. Hazan, E. Seboun, M. Lathrop, G. Gyapay, J. Morissette, and J. Weissenbach. 1996. A comprehensive genetic map of the human genome based on 5,264 microsatellites. *Nature* **380:** 152–154.

Drake, F.H., R.A. Dodds, I.E. James, J.R. Connor, C. Debouck, S. Richardson, E.L. Lee-Rykaczewski, L. Coleman, D. Reiman, R. Barthlow, G. Hastings, and M. Gowen. 1996. Cathepsin K, but not cathepsins B, L, or S, is abundantly expressed in human osteoclasts. *J. Biol. Chem.* **271:** 12511–12516.

Frey, J. 1995a. What dwarfed Toulouse-Lautrec? *Nature Genet.* **10:** 128–130.

Frey, J.B. 1995b. Toulouse-Lautrec's diagnosis. *Nature Genet.* **11:** 363.

Gan, S.-Q., O.W. McBride, W.W. Idler, N. Markowva, and P.M. Steinert. 1990. Organization, structure and polymorphisms of the human profilaggrin gene. *Biochem.* **29:** 9432–9440.

Gleb, B.D., J.G. Edelson, and R.J. Desnick. 1995a. Linkage of pycnodysostosis to chromosome 1q21 by homozygosity mapping. *Nature Genet.* **10:** 235–237.

Gleb, B.D., E. Spencer, S. Obad, G. Edelson, S. Faure, J. Weissenbach, and R.J. Desnick. 1995b. The pycnodysostosis locus resides in a 2 cM region at chromosome 1q21. *Am. J. Hum. Genet.* **57:** A259.

Gleb, B.D., G.-P. Shi, H.A. Chapman, and R.J. Desnick. 1996. Pycnodysostosis, a lysosomal disease caused by cathepsin K deficiency. *Science* **273:** 1236–1238.

Hill, P.A., D.J. Buttle, S.J. Jones, A. Boyde, M. Murata, J.J. Reynnolds, and M.C. Meikle. 1994. Inhibition of bone resorption by selective inactivators of cysteine proteinases. *J. Cell. Biochem.* **56:** 118–130.

Hudson, T., L. Stein, S. Gerety, J. Ma, A. Castle, J. Silva, D. Slonim, R. Baptista, L. Kruglyak, S. Xu, X. Hu, A. Colbert, C. Rosenberg, M.P. Reeve-Daly, S. Rozen, L. Hui, X. Wu, C. Vestergaard, K. Wilson, J. Bae, S. Maitra, S. Ganiatsas, C. Evans, M. DeAngelis, K. Ingalls, R. Nahf, L. Horton, M. Oskin, A. Collymore, W. Ye, V. Kouyoumjian, I. Zernsteva, J. Tarn, R. Devine, D. Courtney, M. Renaud, H. Nguyen, T. O'Connor, C. Fizames, S. Faure, G. Gyapay, C. Dib, J. Morissette, J. Orlin, B. Birren, N. Goodman, J. Weissenbach, T. Hawkins, S. Foote, D. Page, and E. Lander. 1995. An STS-based map of the human genome. *Science* **270:** 1945–1954; with supplementary data from the Whitehead Institute/MIT Center for Genome Research, Human Genetic Mapping Project, Data Release 10 (May 1996), http://www-genome.wi.mit.edu.

Inaoka, T., G. Bilbe, O. Ishibashi, K. Tezuka, M. Kumegawa, and T. Kokubo. 1995. Molecular cloning of human cDNA for cathepsin K: Novel cysteine proteinase predominantly expressed in bone. *Biochem. Biophys. Res. Commun.* **206:** 89–96.

Lerner, U.H. and A. Grubb. 1992. Human cystatin C, a cysteine proteinase inhibitor, inhibits bone resorption in vitro stimulated by parathyroid hormone and parathyroid hormone-related peptide of malignancy. *J. Bone Miner. Res.* **7:** 433–440.

Li, Y.-P., M. Alexander, A.L. Wucherpfennig, P. Yelick, W. Chen, and P. Stashenko. 1995. Cloning and complete coding sequence of a novel human cathepsin expressed in giant cells of osteoclastomas. *J. Bone Miner. Res.* **10:** 1197–1202.

Maciewicz, R.A. and D.J. Etherington. 1988. A comparison of four cathepsins (B, L, N and S) with collagenolytic activity from rabbit spleen. *Biochem. J.* **256:** 433–440.

Maroteaux, P. 1995. Toulouse-Lautrec's diagnosis. *Nature Genet.* **11:** 363–364.

Maroteaux, P. and M. Lamy. 1965. The malady of Toulouse-Lautrec. *J. Am. Med. Assoc.* **191:** 715–717.

Polymeropoulos, M.H., R.I. Ortiz De Luna, S.E. Ide, R. Torres, J. Rubenstein, and C.A. Francomano. 1995. The gene for pycnodysostosis maps to human chromosome 1cen-21. *Nature Genet.* **10:** 238–239.

Shi, G.-P., H.A. Chapman, S.M. Bhairi, C. DeLeeuw, V.Y. Reddy, and S.J. Weiss. 1995. Molecular cloning of human cathepsin O, a novel endoproteinase and homologue of rabbit OC2. *FEBS Lett.* **357:** 129–134.

Smith, S.M. and M.M. Gottesman. 1989. Activity and deletion analysis of recombinant human cathepsin L expressed in Escherichia coli. *J. Biol. Chem.* **264:** 20487–20495.

Yoneda, K., D. Hohl., O.W. McBride, M. Wang, K.U. Cehrs, W.W. Idler, and P.M. Steinert. 1992. The human loricrin gene. *J. Biol. Chem.* **267:** 18060–18066.

Received August 29, 1996; accepted in revised form October 3, 1996.

A nonsense mutation in the cathepsin K gene observed in a family with pycnodysostosis.

M R Johnson, M H Polymeropoulos, H L Vos, et al.

Genome Res. 1996 6: 1050-1055
Access the most recent version at doi:10.1101/gr.6.11.1050