

Loss-of-function mutations in the cathepsin C gene result in periodontal disease and palmoplantar keratosis

Carmel Toomes^{1,2*}, Jacqueline James^{3*}, A. Joseph Wood^{1*}, Chu Lee Wu^{1*}, Derek McCormick⁴, Nicholas Lench⁵, Chelsea Hewitt¹, Leanne Moynihan⁵, Emma Roberts⁵, C. Geoffrey Woods⁶, Alexander Markham⁵, Melanie Wong⁷, Richard Widmer⁸, Khaled Abdul Ghaffar⁹, Michael Pemberton³, Ibtesam Ramzy Hussein¹⁰, Samia A. Temtamy¹⁰, Robin Davies¹¹, Andrew P. Read¹, Philip Sloan³, Michael J. Dixon^{2,3} & Nalin S. Thakker^{1,3}

*These authors contributed equally to this work.

Papillon-Lefèvre syndrome, or keratosis palmoplantaris with periodontopathia (PLS, MIM 245000), is an autosomal recessive disorder that is mainly ascertained by dentists because of the severe periodontitis that afflicts patients^{1,2}. Both the deciduous and permanent dentitions are affected, resulting in premature tooth loss. Palmoplantar keratosis, varying from mild psoriasis-form scaly skin to overt hyperkeratosis, typically develops within the first three years of life. Keratosis also affects other sites such as elbows and knees. Most PLS patients display both periodontitis and hyperkeratosis. Some patients have only palmoplantar keratosis or periodontitis, and in rare individuals the periodontitis is mild and of late onset^{3–6}. The PLS locus has been mapped to chromosome 11q14–q21 (refs 7–9). Using homozygosity mapping in eight small consanguineous families, we have narrowed the candidate region to a 1.2-cM interval between *D11S4082* and *D11S931*. The gene (*CTSC*) encoding the lysosomal protease cathepsin C (or dipeptidyl aminopeptidase I) lies within this interval. We defined the genomic structure of *CTSC* and found mutations in all eight families. In two of these families we used a functional assay to demonstrate an almost total loss of cathepsin C activity in PLS patients and reduced activity in obligate carriers.

Previous linkage analysis in PLS families reported a maximum two-point lod score of 8.24 ($\theta=0$) for the microsatellite marker *D11S1367*, whereas multipoint analysis generated a lod score of 10.45, placing the PLS gene within a 4–5-cM genetic interval flanked by *D11S4197* and *D11S931* (ref. 9). To refine the location of the PLS gene, we performed homozygosity mapping and linkage analysis in eight families with affected children showing features typical of PLS (Fig. 1): three Egyptian and three Indian/Pakistani with known parental consanguinity, and two Lebanese where consanguinity was not reported. We genotyped the families with six microsatellite markers (*D11S1365*, *D11S1354*, *D11S4082*, *D11S1332*, *D11S1311*, *D11S919*) spanning approximately 11.3 cM on chromosome 11q14–q21. We demonstrated linkage with no evidence of genetic heterogeneity. A maximum two-point lod score of 9.14 ($\theta=0.05$) was observed with the marker *D11S1332*. Multipoint analysis gave a maximum lod score of 15.1 within the interval bounded by *D11S1365* and *D11S1332*. In all families, including those non-consanguineous, we saw a region of homozygosity at 11q14–q21. Families 2 and 8 define a

minimal region of homozygosity between markers *D11S4082* and *D11S1332* (~3.4 cM on the Marshfield map) common to all affected individuals (Fig. 2). As all genetic linkage studies performed so far indicate that PLS is genetically homogeneous^{7–9}, we combined our data with those of Hart *et al.*⁹, which further narrowed the minimal critical region to a 1.2-cM interval bounded proximally by *D11S4082* and distally by *D11S931* (Fig. 2).

Database analyses indicated the presence of several genes and ESTs within the *D11S4082*–*D11S931* interval. One of these genes, *CTSC*, is near *D11S931* and encodes an oligomeric lysosomal protease capable of removing dipeptides from the amino terminus of protein substrates^{10,11}. It also appears to have an endopeptidase activity¹². *CTSC* is expressed at high levels in tissues including various epithelia and cells involved in the immune response, such as polymorphonuclear leukocytes, alveolar macrophages and their precursors¹⁰. The main functions of cathepsin C are thought to be protein degradation and proenzyme activation¹⁰. *CTSC* was selected as a candidate for further analysis because other conditions with lysosomal dys-

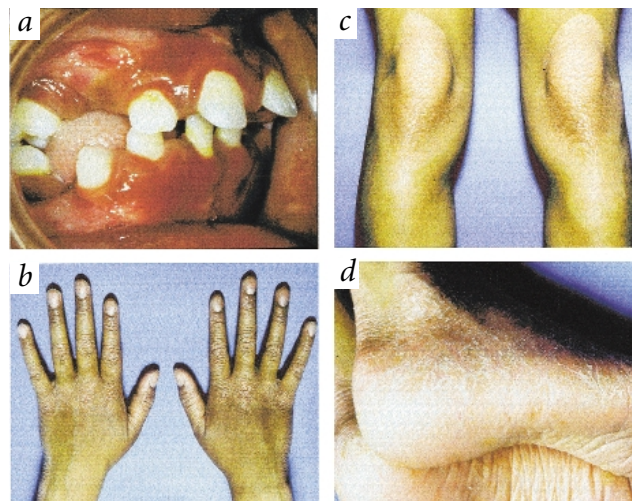
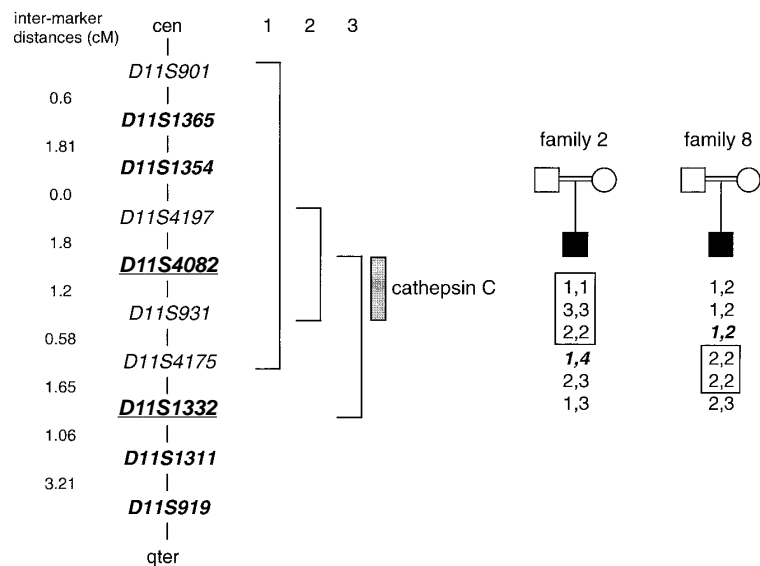


Fig. 1 Clinical features of PLS. **a**, Erythematous swollen gingivae indicative of severe periodontitis. **b,c,d**, Diffuse hyperkeratosis affecting the hands, knees and feet, respectively.

¹Department of Medical Genetics, St. Mary's Hospital; ²School of Biological Sciences, University of Manchester, Stopford Building; and ³Department of Dental Medicine and Surgery, Turner Dental School, University of Manchester, Manchester, UK. ⁴Department of Oncology, Queens University of Belfast, Belfast City Hospital Tower, Belfast, UK. ⁵Molecular Medicine Unit and ⁶Department of Clinical Genetics, St. James's University Hospital, Leeds, UK. ⁷Department of Immunology and Infectious Diseases, New Children's Hospital, and ⁸Department of Paediatric Dentistry, Westmead Hospital Dental Clinical School, Westmead, New South Wales, Australia. ⁹Department of Oral Diagnosis and Periodontology, Eims-Shams University, Cairo, Egypt. ¹⁰Department of Human Genetics, National Research Centre, Cairo, Egypt. ¹¹Dental Health Unit, Skelton House, Manchester, UK. Correspondence should be addressed to N.S.T. (e-mail: nthakker@man.ac.uk).

Fig. 2 Fine-mapping of the PLS locus. The genetic map shows relative order of 10 microsatellite DNA markers at 11q14–q21 and inter-marker distances (http://www.marshmed.org/genetics/). Bars represent limits of linkage intervals for the PLS gene as reported by Fischer *et al.*⁷ (1), Hart *et al.*⁹ (2) and our study (3). Markers used here are shown in bold. Annotated pedigrees for families 2 and 8 with genotypes are shown with homozygous regions boxed and flanking markers underlined (*D11S4082–D11S1332*). Combining data from all studies indicates the minimal critical region containing the PLS gene (shaded box).



function such as Chediak-Higashi syndrome (MIM 214500) also feature severe early onset periodontitis^{13,14}.

CTSC was reported to consist of two exons¹⁰. Amplification of exon 2 for mutation analysis using the reported sequence was successful. Our attempts to amplify exon 1 of *CTSC* using a variety of exonic and flanking intronic primers repeatedly failed. Thus, we re-characterized the genomic organization of *CTSC*. Sequence analysis of a *CTSC*-containing BAC clone, 133M2, revealed that the cDNA sequence previously referred to as 'exon 1' is actually divided into six exons. *CTSC* is therefore encoded by seven exons that are separated by six introns, all of which fall in positions identical to those described for the mouse gene¹⁵. All of the splice donor and acceptor sites conformed to the published consensus sequences¹⁶ (Table 1).

Mutation analysis of the eight families used for homozygosity mapping revealed homozygous sequence changes in all affected individuals (Table 2). In each case the parents were heterozygous; the other non-affected members were either heterozygous or homozygous for the wild-type allele (Fig. 3). The mutations include a nonsense mutation in family 7, mutation of an AG acceptor splice site in family 3 and six missense mutations. The missense mutations were not present in 100 Egyptian, 50 Pakistani/Indian and 50 European controls. All missense changes are in the region encoding the mature protein and each replaces an amino acid that is highly conserved across species. The Q252L and G301S mutations in families 1 and 6, respectively, replace amino acids that are conserved in all cysteine proteinases, and the V249F mutation in family 4 replaces an amino acid that is conserved in all but one cysteine proteinase¹¹. These three amino acids are probably part of substrate-binding sites in *CTSC* (ref. 17).

To confirm that the mutations were pathological, we undertook functional studies. We were able to obtain fresh blood from

affected individuals in families 4 and 5. Using a fluorometric assay and synthetic *CTSC*-specific fluorogenic substrates, we measured *CTSC* activity in peripheral blood leukocytes in at least one affected and one unaffected member of each family together with ethnically matched controls (Fig. 4). The mean activity of *CTSC* in seven healthy control individuals was 733.98 $\mu\text{mol}/\text{mg}/\text{min}$ (1s.d. \pm 48.37). *CTSC* activity was reduced in the two unrelated affected individuals, with a mean of 12.89 $\mu\text{mol}/\text{mg}/\text{min}$ (1s.d. \pm 7.39). This residual activity is probably not due to *CTSC*, as it was unaffected by the *CTSC* inhibitors guanidinium chloride or E-64 (ref. 18, and data not shown). In contrast, activity in control and heterozygote samples was reduced by approximately 30% with E-64 and 50% with guanidinium chloride (data not shown). Most likely, the affected individuals tested here have total loss of cathepsin C activity. As expected, three obligate or known (by mutation analysis) heterozygotes also had reduced activity. The mean activity of *CTSC* in this group was 268.33 $\mu\text{mol}/\text{mg}/\text{min}$ (1s.d. \pm 69.38).

Lack of functional *CTSC* in PLS may be associated with a reduced host response against bacteria in dental plaque and possibly at other sites¹⁹. *CTSC* has an essential role in the activation of granule serine proteases expressed in bone marrow-derived effector cells of both myeloid and lymphoid series²⁰. These proteases are implicated in a wide variety of immune and inflammatory processes, including phagocytic destruction of bacteria and local activation or deactivation of cytokines and other inflammatory mediators. *CTSC* is also required for processing and activation of the T-lymphocyte granzymes A and B, the key agents of T-cell-mediated cell killing²¹. The lack of a generalized T-cell immunodeficiency in PLS suggests that other pathways can compensate for loss of cathepsin C in most tissues.

The PLS phenotype also indicates a role for cathepsin C in epithelial differentiation or desquamation. This identifies a new group of candidates for other mendelian conditions involving hyperkeratosis. Aberrant epithelial differentiation may also affect the junctional epithelium that binds the gingiva to the tooth surface, possibly weakening the mechanical barrier to periodontal pathogens.

Table 1 • Intron-exon boundary sequences of *CTSC**

Exon	cDNA position	Splice acceptor	Splice donor
1	5' UTR–172	5' UTR-ATGGGTGCT	CGGTTATGGgtaagccgc
2	173–318	tttcttagGACCAACAAG	TTTTTTAAGgtagtttt
3	319–485	tgtttcagTATAAAGAA	TCAGGAAAAGtgagttgc
4	486–641	ttggggcagGTATTCTAA	AATCCCAAAGgtaatacaag
5	642–757	cctttctagGCCCCAACC	GAAACCAAGgtaaaaaaa
6	758–889	catgcagCATCCTGTG	ATGCTCAAAGgtaagtgtt
7	890–3' UTR	aatcttcagGCTGTGAAG	AAATTGTAG–3' UTR

*Intronic sequence is indicated in lower-case letters, exonic sequence in upper-case.

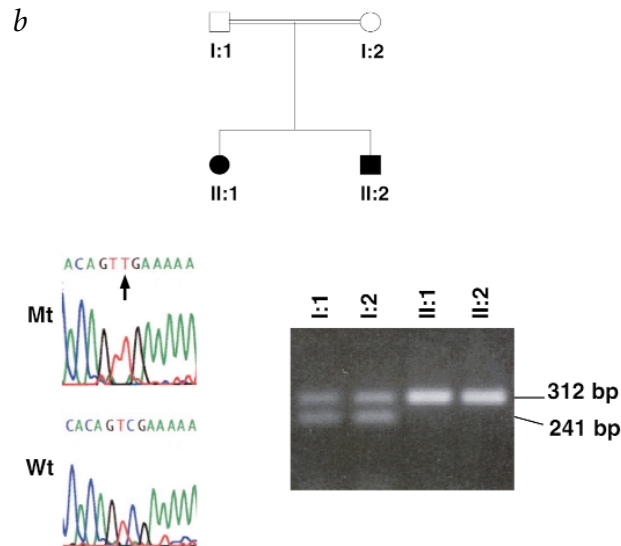
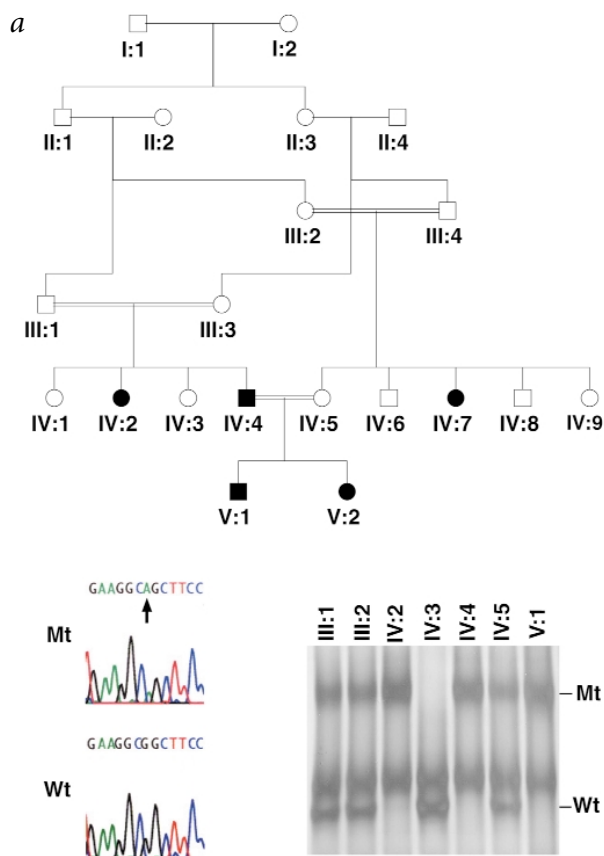


Fig. 3 Segregation of *CTSC* mutations in PLS families 6 and 7. **a**, Segregation of 901G→A mutation in family 6. This mutation is detected as a SSCP variant. Part of exon 7 was amplified and samples size-fractionated on a non-denaturing gel. III:1, III:2 and IV:5 are heterozygous carriers with both the wild-type (Wt) and mutant (Mt) alleles; IV:2, IV:4 and V:1 are affected and homozygous for the mutant allele; IV:3 is unaffected and homozygous for the wild-type allele. **b**, Segregation of the 628C→T mutation in family 7. The mutation abolishes a *TaqI* site in exon 4; this exon was amplified from genomic DNA and products digested with *TaqI*. I:1 and I:2 are heterozygous carriers and have both the wild-type (241 and 71 bp) and the mutant alleles (312 bp); II:1 and II:2 are both affected and have only the mutant allele. The 71-bp band of the wild-type allele is not shown.

Methods

Patient selection. Families with PLS were ascertained on the basis of at least one affected member with typical manifestations of PLS. We collected venous blood samples with informed consent in EDTA and lithium heparin for DNA extraction and *CTSC* functional analysis, respectively.

Genotyping. We extracted DNA from peripheral blood by conventional automated procedures. We genotyped DNA samples from PLS families at six microsatellite loci (*D11S1365*, *D11S1354*, *D11S4082*, *D11S1332*, *D11S1311*, *D11S919*) mapping to chromosome 11q14–q21. Samples were PCR amplified using unlabelled primers and size-fractionated by denaturing polyacrylamide gel electrophoresis. Products were visualized by silver staining or PCR amplified with fluorescence dye-labelled primers and resolved on an ABI 377 automated DNA sequencer and analysed by Genescan and Genotyper software. We then constructed haplotypes manually to define the critical PLS region.

Linkage analysis. DNA marker information was obtained from CEPH-Généthon (<http://www.cephb.fr/cgi-bin/wdb/ceph/systeme/form>). We obtained inter-marker distances and map orders from the Marshfield web site (<http://www.marshmed.org/genetics/>) and the Genetics Location Database (<http://cedar.genetics.soton.ac.uk>). We generated two-point lod scores using the LINKAGE analysis programs²² assuming a fully penetrant

autosomal recessive mode of inheritance with a disease gene frequency of 0.001. Multipoint lod scores were generated using Genehunter Plus²³.

Analysis of the genomic organization of *CTSC*. We used primers (exon 7-1F, 5'-CGGCTTCCTGGTAATTCTTC-3'; exon 7-1R, 5'-GTAGTGGAGGAAGTCATCATATAC-3') in a PCR assay to isolate clone 133M2 from a BAC library (Research Genetics). This clone was digested with *AluI*, *BamHI/BglII*, *PstI*, *Sau3A1* or *SstI* and the resulting fragments cloned into pBluescript. We screened recombinant colonies with end-labelled oligonucleotides designed using the *CTSC* cDNA sequence. Genomic and cDNA sequences were compared and intron-exon boundaries identified by comparison with the published consensus sequence¹⁶. Alternatively, flanking intronic sequence was determined by direct sequencing of BAC 133M2.

Mutation analysis. We screened all exons for sequence variations by direct cycle sequencing of both forward and reverse strands on an ABI 377 sequencer using the Applied Biosystems DyeDeoxy terminator kit. The PCR primers for each exon were as follows: exon 2F, 5'-GACTGTGCTCAA ACTGGGTAG-3'; exon 2R, 5'-CTACTAATCAGAAGAGGTTTCAG-3'; exon 3F, 5'-GGGGCACATTTACTGTGAATG-3'; exon 3R, 5'-CGTATGTCTCATTTGTAGCAAC-3'; exon 4F, 5'-GTACCACTTTCACCTTAGCA-3'; exon 4R, 5'-GGAGGATGGTATTCAGCATTC-3'; exon 5F, 5'-CCTAGCTAGTCTGGTAGCTG-3'; exon 5R, 5'-GTATCCCCGAAATCCATCACA-3'; exon 6F, 5'-CTCTGTGAGGCTTCAGATGTC-3'; exon 6R, 5'-CAACAGCCAGCTGCACACAG-3'; exon 7-1F, 5'-CGGCTTCCTGTGTAATTCTTC-3'; exon 7-1R, 5'-GTAGTGGAGGAAGTCATCATATAC-3'; exon 7-2F, 5'-CAATGAAGCCCTGATCAAGC-3'; exon 7-2R, 5'-CTTCTGAGATTGCTGCTGAAAAG-3'.

We tested for the segregation of the mutations within families and for the presence of the mutation in controls by PCR amplification of the exons, followed by either a mutation-specific restriction-fragment-polymorphism (RFLP) analysis or a combined single-strand-conformation heteroduplex (SSCP-HD) analysis. For mutation-specific RFLP analysis, PCR products were restriction digested and size fractionated by agarose gel electrophoresis. For SSCP-HD analysis, the PCR products were denatured,

Table 2 • Mutations in *CTSC*

Family	Site	Mutation	Amino acid	Predicted effect
1	exon 5	755A→T	Q252L	glutamine→leucine
2	exon 7	1015C→T	R339C	arginine→cysteine
3	intron 3	IVS3-1G→A	–	altered splicing
4	exon 5	745G→T	V249F	valine→phenylalanine
5	exon 7	1040A→G	Y347C	tyrosine→cysteine
6	exon 7	901G→A	G301S	glycine→serine
7	exon 4	628C→T	R210X	arginine→stop
8	exon 6	815G→C	R272P	arginine→proline

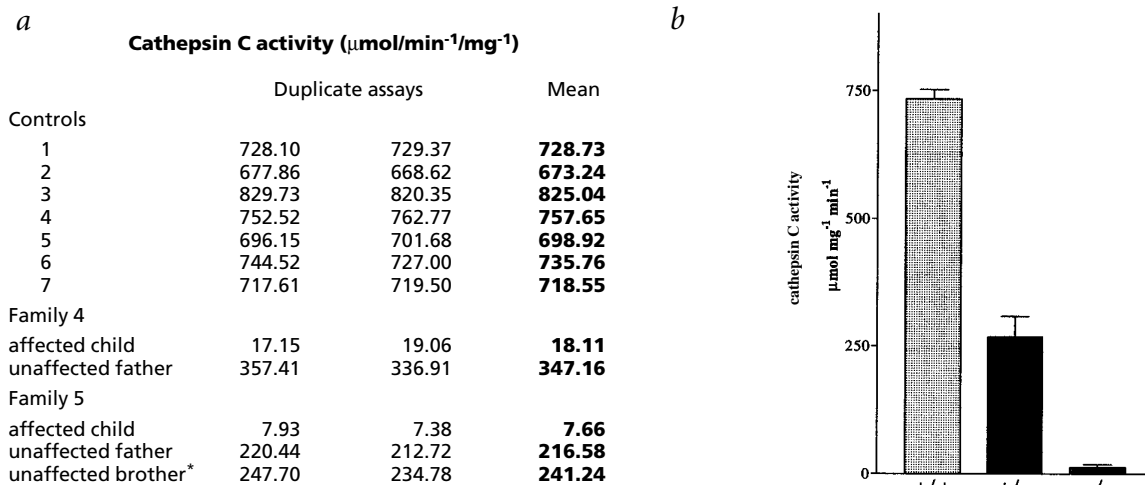


Fig. 4 Reduction of cathepsin C activity in PLS. Activity in peripheral blood leukocytes was measured by hydrolysis of fluorogenic substrate H-gly-arg-NHMec. Cathepsin C activity is expressed as μmol NHMec produced/min/mg protein ($\mu\text{mol min}^{-1} \text{mg}^{-1}$). **a**, Data from duplicate analysis of samples from seven healthy control individuals, three heterozygotes and two affected individuals. *Known heterozygote by mutation analysis. **b**, The columns show mean \pm s.e. of cathepsin C activity ($\mu\text{mol min}^{-1} \text{mg}^{-1}$) in control (+/+), heterozygote (+/-) and affected (-/-) groups. There is nearly a complete loss of activity in affected individuals and a reduction of activity in heterozygotes compared with controls.

electrophoresed through a 0.7 \times Hydrolink-MDE gel (AT Biochem) at 250 V for 13–16 h and visualized by silver staining.

Functional analysis of CTSC. We prepared viable leukocyte pellets for the functional assay as described²⁴ and stored them at -70°C . We resuspended leukocyte pellets in 500 μl sodium phosphate buffer (0.1 M, pH 6.5) containing 0.1% Triton X-100 (Sigma). The suspension was sonicated on ice for 5 s using an MSE Soniprep 150. We determined the protein content of the preparation by the bicinchoninic acid method (Pierce).

We determined the activity of leukocyte CTSC by measuring the amount of NHMec released by hydrolysis of the synthetic substrate glycyl-L-arginine-7-amido-4-methylcoumarin (H-gly-arg-NHMec, Bachem) using a modified protocol²⁵. The amount of NHMec produced from the substrate (5 mM) by leukocyte sonicate (20 μl) was measured in 200 μl sodium phosphate buffer (0.1 M, pH 6.5) containing NaCl (2 mM) and dithiothreitol (2 mM; Sigma). Substrate hydrolysis was monitored for 1 h at 25°C in the microtitre plate reader (Perkin Elmer LS50B luminescence

spectrometer) at 370-nm excitation and 460-nm emission. The fluorescence measurement was converted to mmol NHMec using a calibration line obtained from NHMec (Bachem) standards under identical conditions. Each assay included controls in which either substrate or cell sonicate was omitted from the reaction mixture. We calculated cathepsin C specific activity as μmol NHMec produced/min/mg protein.

Acknowledgements

We thank the patients and their families for help with this study, and A. Naylor, A. Wallace, A. Carradice, G. Karbani and A. Jackson for contribution to this work. This study was funded in part by the MRC, Wellcome Trust (grant 058423) and Northern and Yorkshire Regional Health Authority.

Received 31 August; accepted 19 October 1999.

- Gorlin, R.J., Sedano, H. & Anderson, V.E. The syndrome of palmar-plantar hyperkeratosis and premature periodontal destruction of teeth. *J. Pediatr.* **65**, 895–908 (1964).
- Hart, T.C. & Shapira, L. Papillon-Lefèvre syndrome. *Periodontol.* **2000** **6**, 88–100 (1994).
- Brown, R.S. *et al.* A possible late onset variation of Papillon-Lefèvre syndrome: report of 3 cases. *J. Periodontol.* **64**, 379–386 (1993).
- Bullon, P. *et al.* Late onset Papillon-Lefèvre syndrome? *J. Clin. Periodontol.* **20**, 662–667 (1993).
- Aubrey Soskolne, W., Stabholz, A., Van Dyke, T.E., Hart, T.C. & Meyle, J. Partial expression of Papillon-Lefèvre syndrome in 2 unrelated families. *J. Clin. Periodontol.* **23**, 764–769 (1996).
- Fardal, O., Drangsholt, E. & Olsen, I. Palmar plantar keratosis and unusual periodontal findings. *J. Clin. Periodontol.* **25**, 181–184 (1998).
- Fischer, J. *et al.* Mapping of Papillon-Lefèvre syndrome to the chromosome 11q14 region. *Eur. J. Hum. Genet.* **51**, 56–60 (1997).
- Laass, M.W. *et al.* Localisation of a gene for Papillon-Lefèvre syndrome to chromosome 11q14–q21 by homozygosity mapping. *Hum. Genet.* **101**, 376–382 (1997).
- Hart, T.C. *et al.* Sublocalization of the Papillon-Lefèvre syndrome locus on 11q14–q21. *Am. J. Med. Genet.* **79**, 134–139 (1998).
- Rao, N.V., Rao, G.V. & Hoidal, J.R. Human dipeptidyl-peptidase I. Gene characterization, localization, and expression. *J. Biol. Chem.* **272**, 10260–10265 (1997).
- Paris, A. *et al.* Molecular cloning and sequence analysis of human preprocathepsin C. *FEBS Lett.* **369**, 326–330 (1995).
- Kuribayash, M., Yamada, H., Ohmori, T., Yanai, M. & Imoto, T. Endopeptidase activity of cathepsin C, dipeptidyl aminopeptidase I, from bovine spleen. *J. Biochem.* **113**, 441–449 (1993).
- Tempel, T.R., Kimball, H.R., Kakenashi, S. & Amen, C.R. Host factors in periodontal disease: periodontal manifestations of Chediak-Higashi syndrome. *J. Periodontol.* **7** (suppl.), 26–27 (1972).
- Hamilton, R.E. Jr & Giansanti, J.S. The Chediak-Higashi syndrome. Report of a case and review of the literature. *Oral Surg. Oral Med. Oral Pathol. Oral Radiol. Endod.* **37**, 754–761 (1974).
- Pham, C.T.N. *et al.* Molecular cloning, chromosomal localization, and expression of murine dipeptidyl peptidase I. *J. Biol. Chem.* **272**, 10695–10703 (1997).
- Breathnach, R. & Chambon, P. Organization and expression of eukaryotic split genes coding for proteins. *Annu. Rev. Biochem.* **50**, 349–383 (1981).
- McGrath, M.E. The lysosomal cysteine proteases. *Annu. Rev. Biophys. Biomol. Struct.* **28**, 181–204 (1999).
- Gigic, B. & Pain, R.H. Competitive inhibition of cathepsin C by guanidinium ions and reexamination of substrate inhibition. *Biochem. Biophys. Res. Commun.* **258**, 6–10 (1999).
- Oguzkurt, P., Tanyel, F.C., Buyukpamukcu, N. & Hicsonmez, A.J. Increased risk of pyogenic liver abscess in children with Papillon-Lefèvre syndrome. *J. Pediatr. Surg.* **31**, 955–956 (1996).
- McGuire, M.J., Lipsky, P.E. & Thiele, D.L. Generation of active myeloid and lymphoid granule serine proteases requires processing by the granule thiol protease dipeptidyl peptidase I. *J. Biol. Chem.* **268**, 2458–2467 (1993).
- Pham, C.T. & Ley, T.J. Dipeptidyl peptidase I is required for the processing and activation of granzymes A and B in vivo. *Proc. Natl Acad. Sci. USA* **96**, 8627–8632 (1999).
- Lathrop, G.M., Lalouel, J.M., Julier, C. & Ott, J. Strategies for multilocus linkage analysis in humans. *Proc. Natl Acad. Sci. USA* **81**, 3443–3446 (1984).
- Kruglyak, L. & Lander, E.S. Faster multipoint linkage analysis using Fourier transforms. *J. Comput. Biol.* **5**, 1–7 (1998).
- Wallace, I.J.C., McCusker, C.A. & McCormick, D. The biochemical diagnosis of lysosomal storage diseases—a review of five years experience. *Ir. J. Med. Sci.* **159**, 203–209 (1990).
- Smyth, M. & O’Cuinn, G. Dipeptidyl aminopeptidase III of guinea-pig brain: specificity for short oligopeptide sequences. *J. Neurochem.* **63**, 1439–1445 (1994).