American Journal of Medical Genetics 110:144–152 (2002)

A 52-kb Deletion in the *SOST-MEOX1* Intergenic Region on 17q12-q21 Is Associated With van Buchem Disease in the Dutch Population

Karen Staehling-Hampton,[1] Sean Proll,[1] Bryan W. Paeper,[1] Lei Zhao,[1] Patrick Charmley,[1] Analisa Brown,[1] Jessica C. Gardner,[2] David Galas,[1] Randall C. Schatzman,[1] Peter Beighton,[2] Socrates Papapoulos,[3] Herman Hamersma,[4] and Mary E. Brunkow[1]*

[1]*Celltech R&D Inc., Bothell, Washington*
[2]*Department of Human Genetics, University of Cape Town Medical School, Observatory, South Africa*
[3]*Department of Endocrinology and Metabolic Diseases, Leiden University Medical Center, The Netherlands*
[4]*Otolaryngologist, Florida Park, Roodepoort, South Africa*

Van Buchem disease is an autosomal recessive sclerosing bone dysplasia characterized by skeletal hyperostosis, overgrowth of the mandible, and a liability to entrapment of the seventh and eighth cranial nerves. The genetic determinant maps to chromosome 17q12-q21. We refined the critical interval to the < 1-Mb region between *D17S2250* and *D17S2253* in 15 affected individuals, all of whom shared a common disease haplotype. Furthermore, we report here the identification of a 52-kb deletion located within the interval and encompassing *D17S1789* that is 100% concordant with the disorder. Although the deletion itself does not appear to disrupt the coding region of any known or novel gene(s), the closest flanking genes are *MEOX1* on the proximal side, and *SOST* on the distal side of the deletion. *MEOX1* is known to be important for the development of the axial skeleton, whereas the *SOST* gene is the determinant of sclerosteosis, a disorder that shares many features with van Buchem disease, thus raising the possibility that van Buchem disease results from dysregulation of the expression of one or both of these genes. © 2002 Wiley-Liss, Inc.

KEY WORDS: endosteal hyperostoses; sclerosteosis; chromosomal deletion; van Buchem disease; SOST/AF326739; MEOX1/U10492

INTRODUCTION

Van Buchem disease, or hyperostosis corticalis generalisata (MIM 239100), is an autosomal recessive sclerosing bone dysplasia, first described in 1955 [Van Buchem et al., 1955]. The prevalence of the disorder is extremely low; fewer than 30 individuals have been diagnosed, most of whom come from the Netherlands [reviewed in Van Buchem et al., 1976; Fryns and Vandenberghe, 1988; Van Hul et al., 1998]. An affected family in Britain was documented [Dixon et al., 1982], but additional reports emanating from France [Lapresle et al., 1976], England [Owen, 1976], and Canada [Worth and Wollin, 1966] probably represent instances of autosomal dominant endosteal hyperostosis, whereas a case from the United States [Eastman and Bixler, 1977] was seen in 1979 by H.H., who suggested a diagnosis of Camurati-Engelmann syndrome because of the severe involvement of the long bones.

Sclerosteosis (MIM 269500) is a closely related disorder found predominantly in the Afrikaner population of South Africa, where 66 affected persons have been documented to date [Beighton et al., 1976; Beighton and Hamersma, 1979; Beighton, 1988]. Isolated individuals or families with sclerosteosis have been reported from the United States [Higinbotham and Alexander, 1941; Stein et al., 1983], Germany [Pietruschka, 1958], Japan [Sugiura and Yasuhara, 1975], Brazil [Freire de Paes Alves et al., 1982], Spain [Bueno et al., 1994], and Senegal [Tacconi et al., 1998].

The two disorders share clinical and radiographic features, notably massive hyperostosis of the calvarium and mandible, mild sclerosis of the spine, and increased

Karen Staehling-Hampton, Sean Proll, and Bryan W. Paeper contributed equally to this work.

David Galas's present address is Keck Graduate Institute of Applied Life Sciences, Claremont, California.

*Correspondence to: Dr. Mary E. Brunkow, Celltech R&D Inc., 1631 220th Street SE, Bothell, WA 98021.
E-mail: mary.brunkow@sea.celltechgroup.com

Received 12 October 2001; Accepted 6 February 2002
DOI 10.1002/ajmg.10401

© 2002 Wiley-Liss, Inc.

radiographic density and cortical widening of the long bones of the limbs. Bone overgrowth often results in facial distortion together with deafness and facial palsy due to cranial nerve entrapment (Fig. 1). The manifestations of sclerosteosis are generally more severe than those of van Buchem disease; in particular, sclerosteosis is often associated with large stature and syndactyly of the second and third fingers. In addition, elevation of intracranial pressure can lead to sudden death by impaction of the brain in the foramen magnum. This seldom occurs in van Buchem disease.

Within the relatively homogeneous populations affected by van Buchem disease in the Netherlands and by sclerosteosis in South Africa, the recessive mutations resulting in these disorders are presumed to have arisen as founder mutations. Furthermore, because these two populations share ancestral links, it was proposed several years ago that van Buchem disease and sclerosteosis might be directly related at the molecular level and that observed differences in phenotypic severity between the disorders could be the result of epistatic influences [Beighton et al., 1984]. Subsequently, mapping studies localized the two disorders to an overlapping region on chromosome 17q12-21 [Van Hul et al., 1998; Balemans et al., 1999; Brunkow et al., 2001], and thereafter the sclerosteosis gene, termed *SOST*, was identified [Brunkow et al., 2001; Balemans et al., 2001]. The *SOST* gene mutation found in the Afrikaner population was not observed in the Dutch patients with van Buchem disease, thus ruling out the possibility that the two disorders are identical. The findings did not, however, resolve the question of whether the two disorders arise from different mutations in the same gene or in two different, closely linked genes.

By combining physical mapping with genetic analysis of affected individuals of Dutch descent, we have localized the van Buchem disease critical interval to the <1-Mb region between markers *D17S2250* and *D17S2253*, a region that also includes the *SOST* gene. Furthermore, we have now identified an approximately 52-kb chromosomal deletion that appears to be specific to the disorder. The deletion forms the basis of a simple polymerase chain reaction (PCR)-based diagnostic test that may now be used to identify carriers of the affected chromosome, thus facilitating genetic management of van Buchem disease. A potential pathogenetic mechanism whereby the deletion results in dysregulation of a nearby gene is discussed.

MATERIALS AND METHODS

Patient Samples

We ascertained 15 affected individuals for this study. All but two have been documented previously [individuals 4–8 by Van Buchem, 1971; 12–14 by Van der Wouden, 1971; 1, 2, and 9–11 by Van Hul et al., 1998]. The identification numbers used here are as in the reports from Van Buchem and Van Hul; individuals 12–14 correspond to cases 2, 4, and 5, respectively, of Van der Wouden. Individuals 1–14 have all been examined by H.H., most of them on multiple occasions. During the most recent visit (1999), individual 3, who is a sibling of individual 2, was examined by H.H. and was determined to be affected, on the basis of the characteristic craniofacial involvement. Individual 15, not

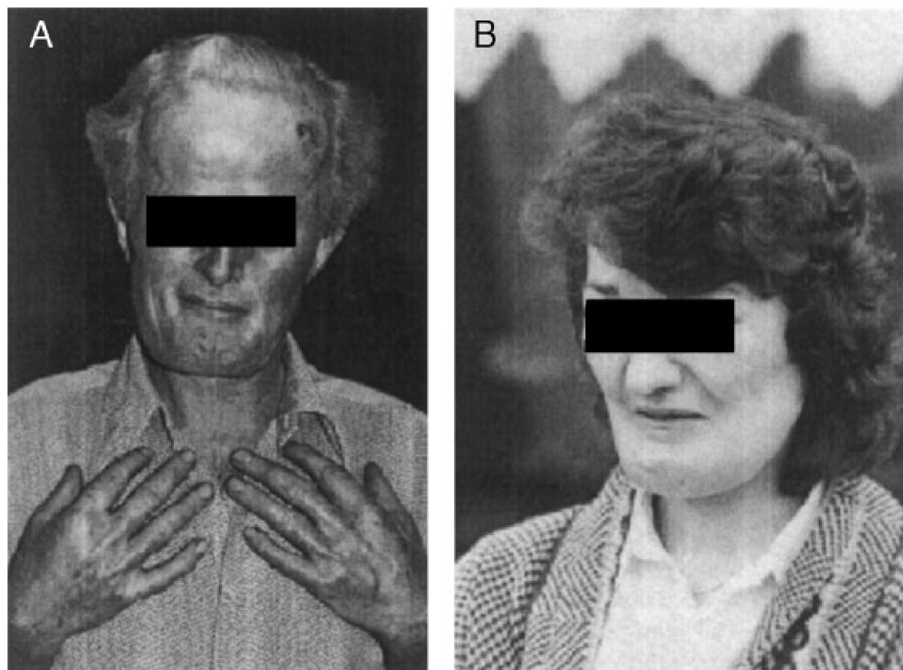

Fig. 1. Clinical manifestations of van Buchem disease. Craniofacial involvement in the disorder ranges from mild to moderate **(A)** to severe **(B)**. In contrast to sclerosteosis, syndactyly has never been observed in patients with van Buchem disease. Photographs appeared originally in Beighton et al., 1984, reprinted with permission of the publisher. ©1984 Munksgaard International Publishers, Ltd., Copenhagen, Denmark.

known to be related to any of the aforementioned individuals, was diagnosed by S.P. The patient, a Dutch boy born in 1984 from a consanguineous marriage, presented at the age of 3 years with facial palsy and progressive deafness requiring a hearing aid at the age of 10 years. He was tall for his age (above the 90th centile), but both parents were also tall; he had a large skull and no syndactyly. Radiographs of the skull showed a homogeneous thickening of the calvarium and of the base of the skull consistent with the diagnosis of van Buchem disease. Radiographs of the spine and the pelvis revealed diffuse sclerosis but no other abnormal findings. The radiographic appearance was confirmed by measurements of spine bone mineral density (BMD), which from the age of 10 to 17 years ranged between 4 and 6 SDs above the mean BMD of boys of his age (Z-score). Spine BMD of the parents was also on the high side; the mother had a Z-score of +0.98 and the father of +0.85. On obtaining informed consent, venous blood samples were drawn from all individuals (including both parents of individual 15), and genomic DNA was extracted by conventional methods.

Microsatellite and *SOST* SNP Genotyping

Patient DNA samples were genotyped using PCR amplification of polymorphic microsatellite markers. The microsatellite markers selected for this study (Table I) are described in public databases or were developed in-house using the SPUTNIK algorithm on sequence obtained from a BAC contig across the van Buchem disease/sclerosteosis region [Brunkow et al., 2001]. PCR products were labeled with infrared IRDyes (LI-COR, Lincoln, NE) using an M13 tailing approach [Oetting et al., 1995] and were resolved on a Li-Cor 4000 DNA sequencer. Allele calls were made using the SAGA genotyping analysis software (University of Washington) or by a trained eye. The C116T polymorphism in the *SOST* gene [described in Brunkow et al., 2001] was typed by direct amplification and sequencing from genomic DNA using primers CCGGGGCTGAGGGAAACAT and TCCGCCCGGTTCATGGTCTTGTTG (333 bp amplicon) and the following cycling conditions: 94°C for 3 min, followed by 40 cycles of 94°C for 30 sec, 55°C for 30 sec, 72°C for 30 sec, and a final extension at 72°C for 5 min.

Amplicon Walking

The following primers were used for PCR amplification of the *SOST* and *MEOX1* exons: *SOST*: exon1 GGTCACCTGGGAGTGCC and AAGTTTCTAAAACCTCCCCA, exon2 GCTTGGCAAAAGCAGGGCTG and TCTCACCTCTGCCCATTCAAA; *MEOX1*: exon1 TGCCAATGAGACAGAGAAGAA and AGAGAGGGTGAGTAACTTCC, exon 2 AATAAAAGAAAGTTTGGGGT and AGGTGGAGGTTACAGTAAGA, exon 3 GCAGAGTGCTTTTAGAACAT and AAGCAGTATCTCTGAAGCTG. In addition, the following primer pairs, derived from the *SOST-MEOX1* intergenic region; [Brunkow et al., 2001; and AF397423] were used to determine the extent of the van Buchem disease deletion: VB1 AGCGCATACGTGAGGACGAT and AGGTCAGAATGGTCGCCAGTT; VB2 ATAGGTGCCTACAAATAGCG and CCCACATTACAGGTATCTCTG; VB3 GATAAATGGATATGGCAAAG and GGTTT-

TABLE I. Marker Genotypes of Affected Individuals for Selected Markers at the van Buchem Disease Locus at 17q12-q21*

Marker	Dutch van Buchem disease genotypes Affected individual												Consensus van Buchem disease haplotype (Dutch)	Consensus sclerosteosis haplotype (Afrikaner)
	1		4		5		9		10		11			
D17S1787	4	4	4	4	4	4	4	8	4	8	4	8	4	1
D17S2231	4	4	4	4	4	4	4	11	4	11	4	11	4	6
D17S1793	4	4	4	4	4	4	4	7	4	7	4	7	4	4
D17S2247	11	11	11	11	11	11	11	12	11	12	11	12	11	11
D17S2248	2	2	2	2	2	2	2	3	2	3	2	3	2	2
D17S855	6	6	6	6	6	6	6	8	6	8	6	8	6	7
D17S2249	1	1	1	1	1	1	1	3	1	3	1	3	1	4
D17S2250	3	3	3	3	3	3	3	1	3	1	3	1	3	3
D17S2251	2	2	2	2	2	2	2	2	2	2	2	2	2	3
D17S2252	3	3	3	3	3	3	3	3	3	3	3	3	3	3
D17S1789	Δ	Δ	Δ	Δ	Δ	Δ	Δ	Δ	Δ	Δ	Δ	Δ	Δ	8
D17S951	1	1	1	1	1	1	1	1	1	1	1	1	1	3
SOST	C	C	C	C	C	C	C	C	C	C	C	C	C	T
D17S2253	1	7	1	1	1	1	1	1	1	1	1	1	1	4
D17S2234	2	1	2	2	2	2	2	2	2	2	2	2	2	2
D17S1860	7	11	7	7	7	7	7	7	7	7	7	7	7	10
D17S2254	2	6	2	2	2	2	2	2	2	2	2	2	2	4
D17S2255	10	1	10	1	10	1	10	10	10	10	10	10	10	8
D17S2235	3	1	3	1	3	1	3	3	3	3	3	3	3	3
D17S1861	8	13	8	2	8	2	8	8	8	8	8	8	8	8

*Alleles are number coded, and shown in order from centromere (*D17S1787*) to telomere (*D17S1861*). Additional markers in the region were included in the analysis; however, only those whose order was definitively established are shown here. The Δ indicates lack of amplification of the corresponding marker. The marker "SOST" is defined by a single nucleotide polymorphism in the sclerosteosis gene (C116T in GenBank acc. AF326739) that distinguishes affected Afrikaners from unaffected individuals. Underlined alleles indicate those that are divergent from the shared haplotype. Shading corresponds to the region of common disease haplotype, defined by the proximal recombination sites in individuals 9, 10, and 11, and by the distal recombination site in individual 1. In the last two columns are consensus haplotypes for van Buchem disease and sclerosteosis in the Dutch and Afrikaner populations, respectively.

CAACTAGTTCTGGTG; VB4 GGTTTATAATTTGCA-ACCAG and CTAGGGCTTAGAAGTTTCCT; VB5 TCCTGTGATCGCATTGAGAC and CCCTGCCATTC-TGGATAGTTT; VB6 TTTAGACCTATCACTCCCAA and ATTCCCTAAGAGATTTGTCC; VB7 CAGTGGCT-TTATTTTCCTAA and GAAGCTTCTCCATGTTCTTA; VB8 CAACTCAATCTTTTGGTGTT and CAAAGTGG-CTCTGATTATTT; VB9 GAGACCTCTCCTCTTTGAAT and GTATCACCAGTGAAGTTGGT. The following primer pairs were used to "walk" to the deletion breakpoints: VB5a AGGGTTCACACCATATCAGAA and CACAGCTGGAGACATGTTACA; VB5b CAACACGAC-ATGAATGGACT and GAGCTGAGATCGCACCACTT; VB5c CCAGTGGAAGAGACAGGTGA and ACCGTG-CAGAGGTAGATGGTA; VB5d CGGATATTTGTCTGT-GATACG and ACGGTGTACACATTTGGTTAG; VB9a ACGCTGCTGTTAAGGTCCA and TGCCAATTAGCC-ACACTCTTC; VB9b GTGCAAAGTGCCTTACACAG and GAGGTTAGACGGGTCTGAGTT; VB9c TGGCA-GGCAGTAGTAACTCTG and CTGGGATTACAGGTG-TCTGG; VB9d TGAGCTGTTCCCACACCACAT and TCAGGACGTTGCACTTTGACA; VB9e GAATGCTGG-ATGTGGATTGAG and GAGCAGAAGGCCTTGACTGA.

Diagnostic PCR Assay

PCR amplification with the following primers distinguish a normal, non-van Buchem chromosome 17 from one bearing the van Buchem disease deletion: VBspan1 (GAATTACTGGCTGAGGCAACC) and VBspan2 (TAC-TACTGGGCCTGGGATGTA) are located outside of the 52-kb region deleted from the affected chromosome, and amplify a 642-bp fragment from affected individuals; VBspan2 and VBint1 (TAGAGAAAGACCTCGT-TATTGG) flank the proximal deletion endpoint and amplify a 720-bp fragment from unaffected individuals. In a multiplexed PCR reaction, both the affected and unaffected chromosome may be detected. Twenty-nanogram genomic DNA was used in a 50 μl reaction containing 0.4 μM VBspan1 and VBint1, 0.8 μM VBspan2, 0.2 mM each dNTP, 60 mM Tris-HCl pH 8.5, 2.5 mM $MgCl_2$, 15 mM ammonium sulfate and 2.5 U Taq polymerase (Roche Molecular Biochemicals, Indianapolis, IN). Cycling conditions were 94°C for 3 min, followed by 40 cycles of 94°C for 30 sec, 63°C for 30 sec, 72°C for 2 min, and a final extension of 72°C for 5 min. Amplicons were resolved on a 2.5% agarose gel stained with ethidium bromide. This assay was applied to normal control DNAs derived from 50 unrelated Dutch individuals; 42 unrelated South African individuals; and 270 samples from the NIGMS Human Variation Collection, panels HD01-HD09, and the NIGMS DNA Polymorphism Discovery Resource, panels 1 and 2. NIGMS DNAs represent 10 different ethnicities (Coriell Cell Repositories, Camden, NJ).

Computational Analysis of *SOST-MEOX1* Intergenic Region

BAC clone 209m4 (human CITB library; Research Genetics) was sequenced to completion to obtain the genomic sequence of the *SOST-MEOX1* intergenic region [Brunkow et al., 2001; and AF397423]. BLAST analysis of this region was performed to identify known genes and/or expressed sequence tags. GENSCAN [Burge and Karlin, 1997] analysis was performed to identify potential coding sequences. We screened the CITB mouse BAC library (Research Genetics, Huntsville, AL) for clones containing both the *Sost* and *Meox1* genes using the following primers: *Sost*, GCTGGCA-TACTGTGTAAGTCC and CTCCACAGCTCACCA-AAGAT (amplicon located ~3 kb 5′ of *Sost* coding sequence); *Meox1*, TGTGAAGTTGCCCAGTATGTG and GGAACACGCAGGATAGGTC (amplifies from *Meox1* 3′ UTR). Clone 266f4 was found to contain both the *Sost* and *Meox1* genes and so was sequenced to completion. BLAST and GENSCAN analyses were performed on the *Sost-Meox1* intergenic region as already described. BLAST was also used to identify conserved sequences in the human and mouse intergenic regions, and putative transcription factor binding sites of the TRANSFAC database [Wingender et al., 2000] were predicted within these conserved elements of 100–400 bp using the Match program (BIOBASE Biological Databases GmbH, Braunschweig, Germany).

RESULTS

Van Buchem Disease Locus Refined to > 1 Mb Region

We obtained DNA samples from 15 individuals affected with van Buchem disease. Individuals 1–11 belong to a 10-generation pedigree described previously [Van Buchem et al., 1976; Van Hul et al., 1998]. This large extended family comes from a small, isolated region in the Netherlands, and the autosomal recessive disease determinant is presumed to result from a founder mutation. Individuals 12–15 in this study are also of Dutch origin but have no known direct relation to the large pedigree. We used a number of microsatellite markers from chromosome 17q12-21 to analyze the DNA samples. More specifically, we focused on the region spanning *D17S1787* to *D17S1861*, because this was reported to contain the disease locus [Van Hul et al., 1998]. Using a large set of novel polymorphic markers derived from a BAC contig across this region [Brunkow et al., 2001] we observed, as expected, a large region of homozygosity across the disease locus. Furthermore, all affected individuals (including 12–15) shared a common disease haplotype (Table I). Individuals 2, 3, 6–8, and 12–15 were nonrecombinant across the entire *D17S1787* to *D17S1861* interval, whereas individuals 1, 4, 5, and 9, 10, 11 carried recombinant chromosomes that allowed us to refine the disease locus to the <1 Mb region between *D17S2250* and *D17S2253* (Table I).

Sclerosteosis in Afrikaners and van Buchem Disease Haplotypes Are Distinct

We next asked whether the Dutch van Buchem disease haplotype was related directly to the disease haplotype found in Afrikaner individuals affected

with sclerosteosis. To address this question, we typed sclerosteosis DNA samples with the same set of markers already described here and compared these directly to van Buchem disease DNA samples. The analysis showed marked differences between the two disease haplotypes (Table I). We then typed all of the van Buchem disease samples for a single nucleotide polymorphism specific for the Afrikaner *SOST* gene mutation [Brunkow et al., 2001], and found that they all carried the normal "C" allele, thus ruling out the possibility that the two disorders arose on a common genetic background.

Van Buchem Disease Critical Region Contains Chromosomal Deletion

During the course of the genotype analysis already described, we observed a lack of amplification of *D17S1789* from the van Buchem disease chromosome, suggesting a possible chromosomal aberration affecting this locus. We localized this marker to a BAC clone (209m4) also carrying *D17S951*, which amplified normally from patient DNA. Examination of the complete sequence of BAC 209m4 indicated that the closest known genes to *D17S1789* were *MEOX1* and *SOST*, located 14 kb proximally and 79 kb distally, respectively. We determined that the coding sequences of these two genes were intact in affected individuals by amplifying each exon directly from genomic DNA. To characterize further the chromosomal region surrounding *D17S1789*, we designed nine PCR primer pairs across the 93 kb *SOST-MEOX1* intergenic region (AF397423) and attempted to amplify short (200–400 bp) sequences from affected and unrelated unaffected DNA samples. We found that four of these amplicons (VB6-VB9) were absent from the patient DNA, suggesting a deletion in the region. By "amplicon walking" we were able to localize roughly the deletion endpoints, and found that primers VB5-forward and VB9c-reverse, located 54 kb apart in normal genomic sequence (Fig. 2A), amplified a 2.3-kb junction fragment from patient DNA. Comparison of the sequence of this junction fragment to the normal genomic sequence obtained from BAC 209m4 allowed the exact localization of the deletion endpoints, normally located 51.7 kb apart (Fig. 2A). At each endpoint was an identical 16-bp sequence (Fig. 2B), included within complete Alu repeat elements, suggesting a possible mechanism for generating the deletion via homologous recombination.

To determine whether the approximately 52-kb deletion was associated only with van Buchem disease or represented a rare polymorphism in the normal population, we designed an assay with which to distinguish deletion carriers from noncarriers (Fig. 3A). A multiplex PCR reaction, including two primers flanking the deletion and a third primer located within the deleted sequence, amplified a deletion-specific and/or wild-type-specific fragment from genomic DNA of affected (deletion-specific fragment only), unaffected heterozygotes (both fragments), and normal (wild-type-specific fragment only) individuals. We found that all 15 patients with van Buchem disease were homozygous for the deletion (three representative samples in Fig. 3B). In contrast, an analysis of 362 normal control individuals, including 50 from the Netherlands, 42 from South Africa, and 270 representing 10 different ethnicities, failed to detect a deletion carrier, thus demonstrating 100% concordance of the deletion with van Buchem disease.

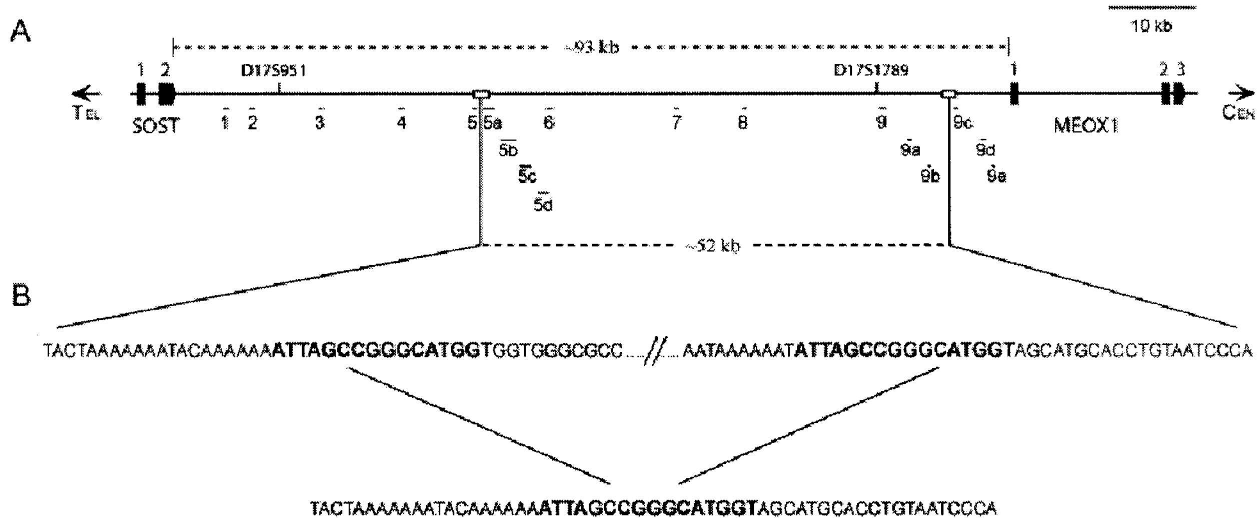

Fig. 2. The van Buchem disease chromosome carries a 52-kb deletion. **A:** The normal *SOST-MEOX1* locus spans the region indicated by light horizontal line. Solid boxes represent exons of the *SOST* (two exons) and *MEOX1* (three exons) genes, the transcriptional orientation of the genes is indicated by the shape of the terminal exons. Short horizontal bars below the line represent amplicons VB1–VB9e, used to determine the extent of the deletion. Deletion endpoints are located between VB5/VB5a and VB9b/VB9c; the Alu elements present at each endpoint are represented by open boxes. **B:** Identical 16-bp sequence (in bold type) is present at each endpoint, as shown on top. A 2.3-kb junction fragment was amplified from patient DNA, using primers VB5-forward and VB9c-reverse, and sequenced. The sequence at the breakpoint, including the single 16-bp element, is most likely the result of homologous recombination.

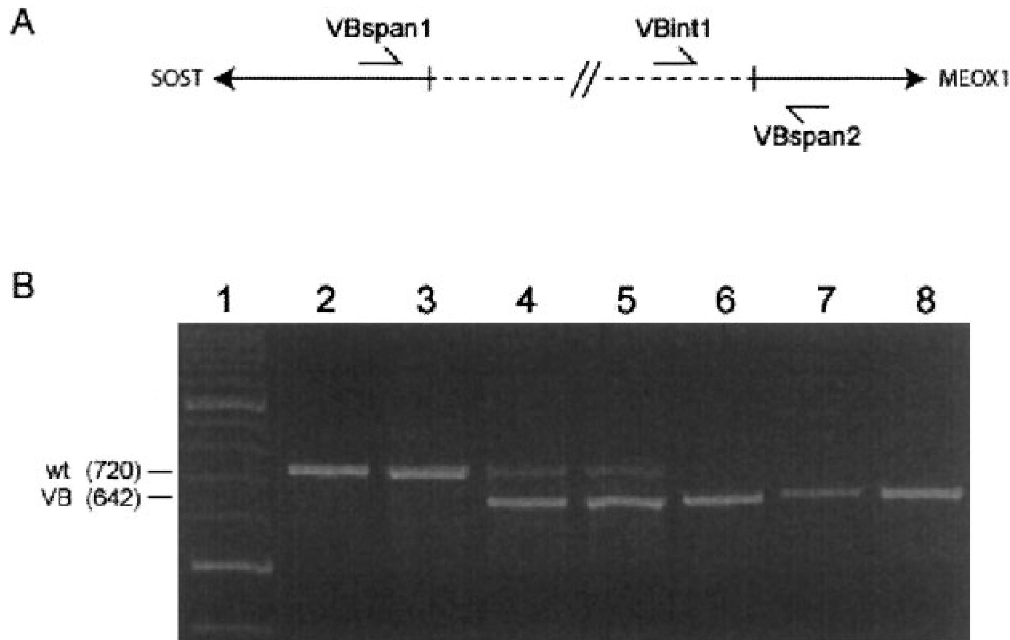

Fig. 3. Diagnostic polymerase chain reaction (PCR) assay detects the van Buchem disease deletion. **A:** Chromosomal region flanking the van Buchem disease deletion (dashed line) is shown, along with PCR primers used in multiplex reaction to detect both affected and unaffected chromosomes. **B:** Genomic DNA from affected and unaffected individuals was amplified and PCR products resolved on an agarose gel. The 642-bp fragment spanning the deletion (del) and the 720-bp fragment obtained from the normal chromosome (wt) are indicated. Lanes: (1) 100 bp ladder; (2) normal control; (3) normal control; (4) father of individual 15; (5) mother of individual 15; (6) VB patient 15; (7) VB patient 1; (8) VB patient 2.

Is the Deletion Disrupting a Novel "van Buchem Disease Gene"?

To determine whether the disease-specific deletion resulted in disruption of a functional gene, we applied extensive computational analyses to the complete sequence of the 93 kb *SOST-MEOX1* intergenic region. In addition, we attempted to obtain evidence for the presence of expressed sequences through exon trapping, Northern blot analysis, cDNA library screening, and reverse transcriptase (RT)-PCR approaches. BLAST [Altschul et al., 1990] analysis of public sequence databases identified a number of potential cDNAs of unknown function, all of which were originally identified through cDNA library screening, direct cDNA selection, or exon trapping procedures [Jones et al., 1994; Brody et al., 1995; Friedman et al., 1995; Osborne-Lawrence et al., 1995]. Our analyses indicated that most of these short "expressed" sequences, especially those identified through exon trapping, represent nonspecific, low-abundant repetitive elements, because they are similar to sequences found in other chromosomal regions. We could not confirm that they represent authentic transcribed genes through direct experimental approaches (see below). Two sequences previously reported during the course of the characterization of the *BRCA1* region, called BCC8 [U70074; Jones et al., 1994] and B169 [U17907; Friedman et al., 1995], did warrant further examination. The 1.1-kb sequence called BCC8 was originally identified on the basis of a short region at one end that is conserved in vertebrates; a probe overlapping this conserved sequence hybridizes strongly to a transcript in skeletal muscle, and less strongly to one in heart [Jones et al., 1994]. We analyzed the sequence of the BCC8 cDNA isolated from a retina cDNA library (U70074), and found that approximately 40% of the sequence is homologous to SINE and LINE elements. BLAST analysis of the unique regions in this clone did not identify any related sequences in public databases, including the EST database, which is well populated with cDNAs from skeletal muscle libraries. Additional observations that led us to surmise that BCC8 does not represent a bona fide transcription unit are the lack of any exon breaks within the 1.1-kb cDNA, and the fact that, aside from the very short (230 bp) conserved element at one end, the sequence is not conserved in mouse (see below). The sequence called B169 was originally identified by direct screening of cDNA libraries with YACs or cosmids from 17q21 [Friedman et al., 1995]. Our analysis of the sequence led us to believe that it is actually a pseudogene, a conclusion confirmed by a recent report on the identification of the *WHSC1L1* gene (corresponding to B169) located on 8p11.2 and the related, nonexpressed pseudogene *WHSC1L2P* located on 17q21 [Stec et al., 2001].

We next used the exon prediction program GENSCAN [Burge and Karlin, 1997] to identify any potentially novel coding sequences in the 93-kb intact intergenic region, as well as the intergenic region with the 52-kb deletion. The results from this analysis suggested the presence of a single transcription unit comprising 12 exons. However, the probability scores for individual exons ranged from only 0.005 to 0.781

(8 of 12 exons had scores < 0.5), and the predicted open reading frame showed no homology to any other sequences in public databases. The approximately 52-kb deletion did not affect significantly the exon predictions in the flanking regions. We carried out a number of RT-PCR experiments in which primers specific for GENSCAN-predicted exons with probability scores greater than 0.5, as well as for the chromosome 17-derived elements matching the putative cDNAs mentioned herein, were used in numerous combinations against a panel of first-strand cDNAs from 36 different human tissues including whole long bone, cartilage, differentiated osteoblasts, as well as skeletal muscle and heart. In no instance were we able to obtain authentic cDNA products. Likewise, of 27 potentially novel exons that we obtained by exon trapping of BAC 209m4, none appeared to correspond to authentic expressed sequences as determined by BLAST, RT-PCR, and/or Northern blot analysis (not shown).

Finally, we determined the complete sequence of the analogous mouse *Sost-Meox1* intergenic region (AF405242) for further computational analyses. As with the human sequence, GENSCAN predicted primarily exons with low probability scores, and, strikingly, none of the predicted exons were orthologous to those predicted in the human intergenic region. A direct comparison of the complete human and mouse intergenic regions revealed several short conserved sequences (100–400 bp in length), all of which (with the exception of the 230-bp region at the end of BCC8) were distinct from the BLAST hits, or trapped or predicted exons described herein. Six of these conserved regions (defined by $> 75\%$ similarity over 100 bp) were located within the 52-kb interval deleted in van Buchem disease. Again, attempts to identify transcripts corresponding to these regions, using RT-PCR as well as cDNA library screening by hybridization, were unsuccessful. In addition, although these elements were strongly conserved at the nucleotide level, they did not share related open reading frames, because of numerous frameshifts. To address the possibility that the conserved elements are involved in transcriptional regulation, we used the TRANSFAC database [Wingender et al., 2000], and were able to identify potential transcription factor consensus binding sites within some of these elements. Whether these consensus sites are truly relevant to gene expression, however, will require functional analysis in cell culture and/or animal models. These studies are currently in progress, and the results will be presented elsewhere.

DISCUSSION

In this study, we showed that four unrelated Dutch families affected with van Buchem disease carry a common disease chromosome that is distinct from that found in Afrikaner families affected with sclerosteosis, a genetically linked and phenotypically similar disorder. This observation indicates that, despite the common ancestry of these two populations, the two very rare disorders arose independently. In addition, during the course of refining the critical van Buchem disease interval, we identified an approximately 52-kb chromosomal deletion that includes *D17S1789*. The presence of the same 16-bp sequence (subregion of complete Alu elements) at each endpoint suggests a mechanism for generation of the deletion through homologous recombination. This disease-specific deletion provides a straightforward means to identify heterozygous gene carriers in Dutch families related to the individuals described here. We have developed a PCR-based diagnostic for this purpose; although distinct causal mutations may exist in other unrelated affected populations, we believe that the assay described here would be useful for characterizing any family within which members have been diagnosed with van Buchem disease.

Although the evidence strongly suggests a causal link to the approximately 52-kb deletion, the exact mechanism for how the deletion results in the disorder remains unclear. Three possible models include one in which the "van Buchem disease gene" is located in the *SOST-MEOX1* intergenic region and is itself disrupted by the deletion; the second in which the deletion results in dysregulation of a nearby gene. A third possibility is that the deletion in affected individuals of Dutch descent is actually in linkage disequilibrium with the causal mutation located elsewhere in the 1-Mb disease interval. We tested the first model by a number of computational and molecular approaches, including BLAST analysis, exon prediction, human versus mouse sequence comparison, exon trapping, and direct RNA analyses. Although all of these methods identified potential coding sequences on their own, comparison of sequence elements identified by each method showed a very striking lack of correspondence. A particularly powerful indicator of coding sequence is cross-species conservation. With the complete sequence of the *SOST-MEOX1* intergenic region from both human and mouse, we were able to identify six conserved regions located within the sequence deleted in van Buchem disease. As with potential coding sequences identified through other computational or experimental approaches, however, we could not find corresponding transcripts through direct analysis of RNA from 36 human tissues. Although we cannot formally rule out the existence of a gene whose expression is exquisitely tissue specific and/or stage specific, our data indicate a lack of additional transcribed sequences in the 93-kb *SOST-MEOX1* intergenic region.

In the second model, the van Buchem disease–specific deletion results in transcriptional dysregulation of a nearby gene or genes, most likely through the removal of a critical regulatory sequence(s). Others have shown that cross-species sequence comparisons can lead to the identification of critical control elements in addition to coding regions [Loots et al., 2000; Dehal et al., 2001]. Indeed, the short conserved sequences we identified in this interval are good candidates for such transcriptional regulators. As a preliminary assessment of the probability that these 100–400-bp elements might have such a function, we used the TRANSFAC database [Wingender et al., 2000] to find potential transcription factor binding sites, and indeed found several different consensus sites in some of the

elements. Because most knowledge of the transcription factors found in this database comes from analyses of more proximal promoter regions, however, the relevance of this simple sequence analysis is unclear in the absence of a complete characterization at the functional level of potential *cis*-acting regulators of genes in the region. The two closest genes flanking the deletion are *MEOX1* and *SOST*, one endpoint being located 35 kb downstream of the *SOST* gene and the other only 6 kb from the 5′ end of *MEOX1*. The *MEOX1* gene encodes a homeodomain-containing protein; its mouse ortholog *Meox1* is expressed, along with the related family member *Meox2*, at high levels in mesodermally derived tissues of the developing embryo as well as several tissues in the adult [Candia et al., 1992; Candia and Wright, 1996]. In the mouse, loss of *Meox1* gene function results in defects in axial skeleton development during embryogenesis [Stamataki et al., 2001]. These vertebral defects are manifested as hemivertebrae, tail kinks, and craniovertebral fusions. Loss-of-function mutations in the *SOST* gene, however, result not in developmental defects, but in the very specific dysregulation of bone density seen in sclerosteosis [Brunkow et al., 2001; Balemans et al., 2001]. Because of the striking phenotypical similarities between van Buchem disease and sclerosteosis, it is tempting to speculate that dysregulation of the *SOST* gene is the underlying cause of the first disorder. Because the approximately 52-kb deletion does not directly affect *SOST* coding sequences, the less severe clinical manifestations of van Buchem disease could then be explained by the retention of some low-level expression of the normal protein product. However, we cannot rule out involvement of the *MEOX1* gene. We have been unable to examine effects on gene expression directly, because of the difficulty in obtaining patient material. Future studies designed to assess the regulatory activity of the deletion region, in cell culture and/or animal models, will further our understanding of the exact molecular mechanism of the disorder.

The molecular characterization of van Buchem disease represents a further step in the understanding of the sclerosing bone disorders. These advances now play a role in the genetic management of affected individuals. In addition, better understanding of the mechanism(s) leading to altered regulation of bone density in these disorders may have wider implications for the development of therapeutic agents for the treatment of osteoporosis.

ACKNOWLEDGMENTS

The authors sincerely thank the patients and their families for their continuing cooperation and interest throughout this project. We thank especially H. Jonkhoff and P. Devriese of Amsterdam for help in revisiting the families and obtaining samples; M. Karperian and C. Lowik of Leiden University and the Celltech Core Sequencing group (especially J. McCuaig and D. Hueglin) for technical assistance; and J. Mulligan, J. Van Ness, J. Latham, J. Howbert, and R. Cottingham for their support and helpful discussions.

REFERENCES

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. J Mol Biol 215:403–410.

Balemans W, Van Den Ende J, Freire Paes-Alves A, Dikkers FG, Willems PJ, Vanhoenacker F, de Almeida-Melo N, Alves CF, Stratakis CA, Hill SC, Van Hul W. 1999. Localization of the gene for sclerosteosis to the van Buchem disease-gene region on chromosome 17q12-q21. Am J Hum Genet 64:1661–1669.

Balemans W, Ebeling M, Patel N, Van Hul E, Olson P, Dioszegi M, Lacza C, Wuyts W, Van Den Ende J, Willems P, Paes-Alves AF, Hill S, Bueno M, Ramos FJ, Tacconi P, Dikkers FG, Stratakis C, Lindpaintner K, Vickery B, Foernzler D, Van Hul W. 2001. Increased bone density in sclerosteosis is due to the deficiency of a novel secreted protein (SOST). Hum Mol Genet 10:537–543.

Beighton P. 1988. Sclerosteosis. J Med Genet 25:200–203.

Beighton P, Hamersma H. 1979. Sclerosteosis in South Africa. S Afr Med J 55:783–788.

Beighton P, Durr L, Hamersma H. 1976. The clinical features of sclerosteosis. A review of the manifestations in twenty-five affected individuals. Ann Intern Med 84:393–397.

Beighton P, Barnard A, Hamersma H, Van der Wouden A. 1984. The syndromic status of sclerosteosis and van Buchem disease. Clin Genet 25: 175–181.

BLAST, http://www.ncbi.nlm.nih.gov/BLAST

Brody LC, Abel KJ, Castilla LH, Couch FJ, McKinley DR, Yin GY, Ho PP, Merajver S, Chandrasekharappa SC, Xu J, Cole JL, Struewing JP, Valdes JM, Collins FS, Weber BL. 1995. Construction of a transcription map surrounding the *BRCA1* locus of human chromosome 17. Genomics 25:238–247.

Brunkow ME, Gardner JC, Van Ness J, Paper BW, Kovacevich BR, Proll S, Skonier JE, Zhao L, Sabo PJ, Fu Y-H, Alisch RS, Gillett L, Colbert T, Tacconi P, Galas D, Hamersma H, Beighton P, Mulligan JT. 2001. The bone dysplasia sclerosteosis results from loss of the *SOST* gene product, a novel cystine knot-containing protein. Am J Hum Genet 68:577–589.

Bueno M, Olivan G, Jimenez A, Garragori JM, Sarria A, Bueno AL, Bueno MJ, Ramos FJ. 1994. Sclerosteosis in a Spanish male: first report in a person of Mediterranean origin. J Med Genet 31:976–977.

Burge C, Karlin S. 1997. Prediction of complete gene structures in human genomic DNA. J Mol Biol 268:78–94.

Candia AF, Hu J, Crosby J, Lalley PA, Noden D, Nadeau JH, Wright CVE. 1992. Mox-1 and Mox-2 define a novel homeobox gene subfamily and are differentially expressed during early mesodermal patterning in mouse embryos. Development 116:1123–1136.

Candia AF, Wright CVE. 1996. Differential localization of Mox-1 and Mox-2 proteins indicates distinct roles during development. Int J Dev Biol 40: 1179–1184.

Dehal P, Predki P, Olsen AS, Kobayashi A, Folta P, Lucas S, Land M, Terry A, Zhou CLE, Rash S, Zhang Q, Gordon L, Kim J, Elkin C, Pollard MJ, Richardson P, Rokhsar D, Uberbacher E, Hawkins T, Branscomb E, Stubbs L. 2001. Human chromosome 19 and related regions in mouse: conservative and lineage-specific evolution. Science 293:104–111.

Dixon JM, Cull RE, Gamble P. 1982. Two cases of van Buchem's disease. J Neurol Neurosurg Psychiatry 45:913–918.

Eastman JR, Bixler D. 1977. Generalized cortical hyperostosis (Van Buchem disease): nosologic considerations. Radiology 125:297–304.

Expressed Sequence Tags database, http://www.ncbi.nlm.nih.gov/dbEST/index.html

Freire de Paes Alves A, Rubim JLC, Cardoso L, Rabelo MM. 1982. Sclerosteosis: a marker of Dutch ancestry? Rev Brasil Genet 5:825–834.

Friedman LS, Ostermeyer EA, Lynch ED, Welcsh P, Szabo CI, Meza JE, Anderson LA, Dowd P, Lee MK, Rowell SF, Ellison J, Boyd J, King M-C. 1995. 22 genes from chromosome 17q21: cloning, sequencing, and characterization of mutations in breast cancer families and tumors. Genomics 25:256–263.

Fryns JP, Vandenberghe H. 1988. Facial paralysis at the age of 2 months as a first clinical sign of van Buchem disease (endosteal hyperostosis). Eur J Pediatr 147:99–100.

GenBank, http://www.ncbi.nlm.nih.gov/Genbank/index.html (for *D17S2247* [accession number AF397414], for D17S2248 [accession number AF397415], for *D17S2249* [accession number AF397416], for *D17S2250* [accession number AF397417], for *D17S2251* [accession number AF397418], for *D17S2252* [accession number AF397419], for

D17S2253 [accession number AF397420], for *D17S2254* [accession number AF397421], for *D17S2255* [accession number AF397422], for human *SOST-MEOX1* intergenic region [accession number AF397423], for mouse *Sost-Meox1* intergenic region [accession number AF405242])

Higinbotham NL, Alexander SF. 1941. Osteopetrosis, four cases in one family. Am J Surg 53:444–454.

Jones KA, Black DM, Brown MA, Griffiths BL, Nicolai HM, Chambers JA, Bonjardim M, Xu C-F, Boyd M, McFarlane R, Korn B, Poustka A, North MA, Schalkwyk L, Lehrach H, Solomon E. 1994. The detailed characterisation of a 400 kb cosmid walk in the *BRCA1* region: identification and localisation of 10 genes including a dual-specificity phosphatase. Hum Mol Genet 3:1927–1934.

Lapresle J, Maroteaux P, Kuffer R, Said G, Meyer O. 1976. Hyperostose cortical generalisee dominante avec atteinte multiple des nerfs craniens. Nouv Presse Med 5:2703–2706.

Loots GG, Locksley RM, Blankespoor CM, Wang ZE, Miller W, Rubin EM, Frazer KA. 2000. Identification of a coordinate regulator of interleukins 4, 13, and 5 by cross-species sequence comparisons. Science 288:136–140.

Oetting WS, Lee HK, Flanders DJ, Wiesner GL, Sellers TA, King RA. 1995. Linkage analysis with multiplexed short tandem repeat polymorphisms using infrared fluorescence and M13 tailed primers. Genomics 30:450–458.

Online Mendelian Inheritance in Man (OMIM), http://www.ncbi.nlm.nih.gov/Omim (for Van Buchem disease/hyperostosis corticalis generalisata [MIM 239100] and sclerosteosis [MIM 269500])

Osborne-Lawrence S, Welcsh PL, Spillman M, Chandrasekharappa SC, Gallardo TD, Lovett M, Bowcock AM. 1995. Direct selection of expressed sequences within a 1-Mb region flanking *BRCA1* on human chromosome 17q21. Genomics 25:248–255.

Owen RH. 1976. Van Buchem's disease. Br J Radiol 49:126–132.

Pietruschka G. 1958. Weitere Mitteilungen uber die Marmorknochenkrankheit (Albers-Schönbergsche Krankheit) nebst Bemerkungen zur Differentiagnose. Klin Monatbl 132:509–525.

SPUTNIK, http://abajian.net/sputnik/index.html

Stamataki D, Kastrinaki M, Mankoo BS, Pachnis V, Karagogeos D. 2001. Homeodomain proteins Mox1 and Mox2 associate with Pax1 and Pax3 transcription factors. FEBS Lett 499:274–278.

Stec I, Van Ommen G-JB, den Dunnen JT. 2001. WHSC1L1, on human chromosome 8p11.2, closely resembles WHSC1 and maps to a duplicated region shared with 4p16.3. Genomics 76:5–8.

Stein SA, Witkop C, Hill S, Fallon MD, Viernstein L, Gucer G, McKeever P, Long D, Altman J, Miller NR, Teitelbaum SL, Schlesinger S. 1983. Sclerosteosis: neurogenetic and pathophysiologic analysis of an American kinship. Neurology 33:267–277.

Sugiura Y, Yasuhara T. 1975. Sclerosteosis: a case report. J Bone Surg [Am] 57:273–276.

Tacconi P, Ferrigno P, Cocco L, Cannas A, Tamburini G, Bergonzi P, Giagheddu M. 1998. Sclerosteosis: report of a case in a black African man. Clin Genet 53:497–501.

Van Buchem FSP. 1971. Hyperostosis corticalis generalisata. Acta Med Scand 189:257–267.

Van Buchem FSP, Hadders HN, Ubbens R. 1955. An uncommon familial systemic disease of the skeleton. Hyperostosis corticalis generalisata familiaris. Acta Radiol 44:109–119.

Van Buchem FSP, Prick JJG, Jaspar HHJ. 1976. Hyperostosis corticalis generalisata familiaris (van Buchem disease). Amsterdam: Excerpta Medica.

Van der Wouden A. 1971. Bone diseases of the temporal bone with hearing impairment. Ph.D. Thesis, University of Leiden, Leiden.

Van Hul W, Balemans W, Van Hul E, Dikkers FG, Obee H, Stokroos RJ, Hildering P, Vanhoenaker F, Van Camp G, Willems PJ. 1998. van Buchem disease (hyperostosis corticalis generalisata) maps to chromosome 17q12-q21. Am J Hum Genet 62:391–399.

Wingender E, Chen X, Hehl R, Karas H, Liebich I, Matys V, Meinhardt T, Pruss M, Reuter I, Schacherer F. 2000. TRANSFAC: an integrated system for gene expression regulation. Nucleic Acids Res 28:316–319.

Worth HM, Wollin DG. 1966. Hyperostosis corticalis generalisata congenita. J Can Assoc Radiol 17:67–74.