Identification of a 52 kb deletion downstream of the SOST gene in patients with van Buchem disease


Van Buchem disease (MIM 239100) is a disorder of the skeleton inherited as an autosomal recessive trait which, according to Beighton,1 belongs to the group of craniofibular hyperostoses. The condition was first described by van Buchem et al2 in 1955 as “hyperostosis corticalis generalisata familiaris”. But Fosmoe et al3 subsequently introduced the eponym “van Buchem disease”. Radiological examination of patients diagnosed with van Buchem disease shows a generalised hyperostosis mainly involving the skull, mandible, clavicles, ribs, and the diaphysis of the long bones (fig 1A). Clinical complications such as facial nerve palsy, optic atrophy, and impaired hearing occur in most patients. These features are very similar to those of sclerosteosis and the two conditions are only differentiated by the hand malformations and the tall stature appearing in sclerosteosis. Using an extended Dutch inbred van Buchem family and two inbred sclerosteosis families, we mapped both disease genes to the same region on chromosome 17q12-q21, supporting the hypothesis that van Buchem disease and sclerosteosis are caused by mutations in the same gene. In a previous study, we positionally cloned a novel gene, called SOST, from the linkage interval and identified three different, homozygous mutations in the SOST gene in sclerosteosis patients leading to loss of function of the underlying protein. The present study focuses on the identification of a 52 kb deletion in all patients from the van Buchem family. The deletion, which results from a homologous recombination between Alu sequences, starts approximately 35 kb downstream of the SOST gene. Since no evidence was found for the presence of a gene within the deleted region, we hypothesise that the presence of the deletion leads to a down regulation of the transcription of the SOST gene by a cis regulatory action or a position effect.

Van Buchem disease is an autosomal recessive skeletal dysplasia characterised by generalised bone overgrowth, predominantly in the skull and mandible. Clinical complications including facial nerve palsy, optic atrophy, and impaired hearing occur in most patients. These features are very similar to those of sclerosteosis and the two conditions are only differentiated by the hand malformations and the tall stature appearing in sclerosteosis. Using an extended Dutch inbred van Buchem family and two inbred sclerosteosis families, we mapped both disease genes to the same region on chromosome 17q12-q21, supporting the hypothesis that van Buchem disease and sclerosteosis are caused by mutations in the same gene. In a previous study, we positionally cloned a novel gene, called SOST, from the linkage interval and identified three different, homozygous mutations in the SOST gene in sclerosteosis patients leading to loss of function of the underlying protein. The present study focuses on the identification of a 52 kb deletion in all patients from the van Buchem family. The deletion, which results from a homologous recombination between Alu sequences, starts approximately 35 kb downstream of the SOST gene. Since no evidence was found for the presence of a gene within the deleted region, we hypothesise that the presence of the deletion leads to a down regulation of the transcription of the SOST gene by a cis regulatory action or a position effect.
PATIENTS AND METHODS

Patients

Patients from the extended Dutch van Buchem family have been described elsewhere. Three other patients are also of Dutch origin. Patient 1 is a 68 year old male. The father of his maternal grandmother and the father of his paternal grandmother were brothers. On examination, macrocephaly and an enlarged mandible (fig 1B), narrow external ear canals, and bilateral facial paresis (right side grade II HB, left side grade V HB) were evident. Audiometry showed a 40-60 dB sensorineural hearing loss in the right ear and 40-100 dB hearing loss in the left ear. Radiological examination of the skull showed a markedly dense skull and mandible (fig 1A). The patient has no history of fractures. Patients 2 and 3 are a sister and brother aged 58 and 65, respectively, with consanguineous parents, second cousins once removed. Around puberty, enlargement of the skull and mandible and thickening of the clavicles and phalanges started to appear in both patients. At the age of 16, the sister showed facial paresis, is now deaf and almost completely blind with exophthalmos and nystagmus, and suffers from disturbances of balance. The brother showed no abnormalities. Radiological examination showed a generalised endosteal hyperostosis from the skull to the lower extremities in both patients.

Southern blot analysis

Genomic DNA from the van Buchem patients, carriers, and controls was digested with EcoRI and HindIII. Fragments were separated on a 0.8% agarose gel in 0.5 × TBE and transferred to a Hybond-N+ nylon membrane (Amersham). Hybridisation was carried out overnight at 65°C with probes derived from PCR products of ESTs and STSs located within the region between D17S1326 and D17S1860. Probes were radioactively labelled using α<sup>32</sup>P-dCTP and α<sup>32</sup>P-dATP (ICN) and purified on a Sephadex<sup>®</sup> G50 fine column (Amersham).

Control samples

A set of 50 Dutch control and 50 random DNA samples was used to check whether the deletion found in the van Buchem patients exists as a polymorphism in the population. We amplified genomic DNA using three primers: primer 1 on the proximal side outside the deletion (del-F1: 5′-cAaAGAcAgCcAgcATTGg-3′), primer 2 on the proximal side within the deletion (del-R1: 5′-AgAgccATcTcAgcTTg-3′), and primer 3 on the distal side outside the deletion (del-R2: 5′-AggTgggAAccTATccgTgc-3′) using a standard multiplex PCR protocol. PCR products were run on a 1% agarose gel in 0.5 × TBE. This resulted in a fragment of 0.7 kb when the deletion was present and a fragment of 1.2 kb when the person did not carry the deletion. In heterozygotes for the deletion both the 0.7 kb and 1.2 kb fragment can be seen.

Identification of transcribed sequences

The complete sequence of the deletion was subjected to the BLAST algorithm<sup>29</sup> and to different exon prediction programs using the NIX program (http://www.hgmp.mrc.ac.uk/Bioinformatics/) to identify putative exons and transcribed sequences. Based on the alignment of the results from the different programs, we selected for further investigation those exons which were predicted by more than two programs.

Tissue expression and library screening

We designed primers located within the predicted exons, MTO sequences, BCC8, and B169 to perform RT-PCR analysis on RNA from 16 different tissues (Human Multiple Tissue cDNA panels I and II, Clontech) and bone tissue. The ICRF human fetal brain
Identification of conserved sequences

The human sequence of the deletion was compared to the homologous mouse sequence located within genomic clone RP23-346P7 (Genbank accession No AC012296) using the "dotter" program by E Sonnhammer. The program creates dot plots for orthologous chromosome regions. Visual inspection allows a straightforward identification of conserved regions.

Quantitative TaqMan analysis

The expression levels of VHR phosphatase in van Buchem patients versus controls were determined using a quantitative RT-PCR method (TaqMan). Fifty ng of RNA derived from lymphocytes was subjected to reverse transcription using MMLV reverse transcriptase according to the manufacturer's protocol (Advantage RT for PCR kit, Clontech). The RT-PCR products were used in a quantitative real time PCR system (ABI PRISM 7700 Sequence Detection System) with the following primer set: VHRtaq-F (5'-gccAgcccTgcAAcgA-3') and VHRtaq-R (5'-cgTTcAgcAcATgggTgATg-3'). PCR products were detected using a VHR specific TaqMan fluorogenic probe (5'-ccTgAcccAcAgATtgccAcgtAgATcc-3'). Experiments were done in duplicate and expression levels were normalised to GAPDH expression as an endogenous control.

RESULTS

Mutation analysis

We performed sequence analysis to check for sequence variations in SOST in patients from the extended Dutch van Buchem family reported previously45 and the three additional Dutch patients reported here. Sequence comparison between

Figure 2  Map showing the van Buchem linkage interval between D17S1326 and D17S1860, and a detailed map of the region between MOXI and VHR with the position of the deletion (white box). The axis on the right represents the distances in kb.
van Buchem patients and a control did not result in the identification of any disease causing mutation, although the entire 5 kb of the gene, including the two exons, the complete intron sequence, and the 5′ UTR, and approximately 1 kb upstream of the gene was sequenced.

**Analysis of the linkage interval**

As no mutation could be found in the SOST gene in genomic DNA of van Buchem patients, we searched for other interesting candidate genes in the linkage interval between D17S1326 (proximal) and D17S1860 (distal). About 18 known genes, including RNU2, DX8X, MOXI1, VHR phosphatase, dg3, PPY, PYY, and glucose-6-phosphatase, map within this region, which is estimated to be approximately 1 Mb. However, based on the function of the encoded proteins, none of these positional candidates could be considered as a strong functional candidate gene for van Buchem disease. We also looked for chromosomal rearrangements in the linkage interval between D17S1326 and D17S1860. Southern blot analysis was carried out on genomic DNA from patients and carriers of the Dutch van Buchem family using probes derived from ESTs and STSs located within this linkage interval. We identified a homozygous deletion in patients with probes derived from MTO-156 (Genbank accession No T27203) and MTO-206 (Genbank accession No T27197). Combined Southern blot and PCR analyses on genomic DNA of van Buchem patients carrying the deletion resulted in the positioning of the centromeric and telomeric breakpoints. The centromeric deletion breakpoint is located just telomeric of the MOXI1 gene (fig 2) within genomic sequence hCIT.501_O_10 (Genbank accession No AC004149). The telomeric deletion breakpoint is located within EST sequence B169 (Genbank accession Nos U17906-U17908). Southern blot analysis of the deletion was carried out on genomic DNA from patients and carriers of the Dutch van Buchem family using probes derived from ESTs and STSs located within this linkage interval. We identified a homozygous deletion in patients with probes derived from MTO-156 (Genbank accession No T27203) and MTO-206 (Genbank accession No T27197). Combined Southern blot and PCR analyses on genomic DNA of van Buchem patients carrying the deletion resulted in the positioning of the centromeric and telomeric breakpoints. The centromeric deletion breakpoint is located just telomeric of the MOXI1 gene (fig 2) within genomic sequence hCIT.501_O_10 (Genbank accession No AC004149). The telomeric deletion breakpoint is located within EST sequence B169 (Genbank accession Nos U17906-U17908). Sequence information of the deletion was available from genomic clone hCIT.501_O_10 (Genbank accession No AC004149), which includes the centromeric 28 kb of the deletion. For the remaining sequence, a cloning and sequencing effort was performed. The complete deletion spans 51.7 kb and starts approximately 35 kb downstream of the SOST gene (fig 2). PCR analysis with primers flanking the proximal and distal breakpoints showed that all patients from the extended Dutch van Buchem family and three additional Dutch van Buchem patients are homozygous for the deletion. Additionally, the parents of the van Buchem patients were shown to be heterozygous. Construction of haplotypes using genetic markers from the chromosome 17 linkage region showed a common haplotype in the three van Buchem patients and the patients from the Dutch van Buchem family, indicating the existence of a common ancestor for all patients (data not shown). We did not find this deletion in DNA samples from 50 random subjects and 50 Dutch controls.

**Mechanism causing the deletion**

Analysis of the sequences flanking the deletion breakpoints showed the presence of Alu repeats on each side. The Alu sequences on the centromeric and telomeric deletion breakpoint are both dimers and are oriented in the same direction. The deletion is most likely the result of an Alu mediated, unequal homologous recombination event between two 16 bp repeats within the Alu sequences which show 100% identity (fig 3).

**Characterisation of the deletion**

In an effort to isolate transcribed sequences from the deletion, we initially performed in silico analyses on the deletion sequence using the BLAST algorithm, the NIX package (http://www.hgmp.mrc.ac.uk/Bioinformatics), and the “dot-ter” program. This resulted in the identification of the cDNA sequences BCC8 and B169, seven MTO sequences already described by Brody et al., 10 predicted exons, and two highly conserved regions (fig 4).

The BCC8 cDNA sequence (Genbank accession No U70074) was initially indicated by zoo blot experiments, whereas northern blot analysis showed a transcript of 2.4 kb in heart and skeletal muscle. In addition, the isolation of a cDNA clone from a retina cDNA library has been reported. However, our NIX analysis did not predict this BCC8 sequence to be a putative exon and the dotter program showed no interspecies sequence homology. Moreover, there is no clear open reading frame in the BCC8 sequence and we were not able to amplify BCC8 by RT-PCR from a cdna panel derived from 16 different tissues. Our data suggest that it is very unlikely that BCC8 is derived from a gene.

B169 (Genbank accession Nos U17906-U17908) is a gene isolated by Friedman et al. 32 after screening fibroblast and ovarian cDNA libraries. In lymphoblasts, transcripts of 3 and 3.8 kb were detected. Homology searches using the BLAST program showed several EST sequences. We isolated one cDNA clone from the ICRF human fetal brain cDNA library (ICRFp5071228) using B169 sequence derived probes. However, comparison of the sequences from this cDNA clone and from the ESTs with the genomic sequence of the deletion showed several nucleotide differences. Since the sequence derived from the cDNA clone did not contain an open reading frame and was not found in any of the EST sequences, it can be concluded that the gene in the deletion is a pseudogene.

BLAST searches on the deletion sequence also showed the presence of seven MTO sequences previously isolated by exon amplification experiments by Brody et al. Four MTO sequences (MTO-154, MTO-144, MTO-153, MTO-143) were too small to use for further experiments. BLAST searches did not result in other ESTs derived from these sequences and the dotter program did not show any sequence conservation between man and mouse. RT-PCR analysis using RNA from multiple tissues on the other MTO sequences (MTO-156, MTO-206, MTO-167) did not provide evidence for the expression of any of these. Negative results were also obtained from BLAST searches and the dotter program. These results suggest that the MTO sequences within the deletion do not represent coding sequences.

Ten exons were predicted within the deletion by several exon prediction programs using the NIX package. These sequences were subjected to RT-PCR analysis and BLAST homology searches but negative results were obtained. Therefore, we concluded that the sequences do not represent novel genes. The only putative exon which was positive in RT-PCR
analysis and shows homology with ESTs is EXPR-7. However, further analysis of this sequence showed the presence of an Alu repeat sequence which accounts for the amplicon observed by RT-PCR and also the homology with ESTs. Only one predicted exon (EXPR-10) showed significant conservation between human and mouse, 51% identity over 57 amino acids. However, no homology with any EST was found and screening of the ICRF human fetal cDNA library and the Stratagene Human Universal cDNA library did not result in the isolation of a cDNA clone from EXPR-10, nor was it positive in our RT-PCR experiments.

Finally, we checked the complete sequence of the deletion for the presence of conserved sequences by performing a human-mouse cross comparison of the genomic sequences using the dotter program developed by E Sonnhammer. This resulted in the identification of two highly conserved regions (fig 4). The first region (Conserved Non-coding Sequence) CNS-1 is 157 bp and shows 84% identity with the mouse.
sequence. The second region (CNS-2) is 480 bp and has a degree of conservation between man and mouse of 91%. However, sequence analysis of these regions did not show an open reading frame, thereby making it unlikely that these regions are coding.

Expression levels of VHR phosphatase in van Buchem patients

To look for a possible influence of the deletion on the expression of neighbouring genes, quantitative TaqMan analysis was set up. Because of the lack of expression of the SOST gene in readily available tissues, this analysis was performed for the VHR phosphatase gene, located approximately 10 kb upstream of the SOST gene. In lymphoblastoid cell lines from van Buchem patients carrying the deletion and from control samples, the expression levels of the VHR phosphatase gene were determined. We were not able to detect an altered level of VHR expression in patients as compared to controls (data not shown) since statistical analysis using Student’s t test did not show a significant difference between the two groups (p=0.16).

DISCUSSION

We previously mapped van Buchem disease and sclerosteosis to the same chromosomal region in 17q12-q21,16,7 supporting the hypothesis of Beighton et al46 that both conditions could be caused by mutations in the same gene. Recently, we refined the linkage interval using additional microsatellite markers to a region of approximately 1 Mb between D17S1326 and D17S1860. Subsequently, we and others were able to positionally clone a new gene, SOST, from this region and we identified three different SOST loss of function mutations in sclerosteosis patients.58,27

Extensive sequence analyses in patients from the Dutch van Buchem family used to localise the van Buchem disease gene did not result in the identification of a disease causing mutation in the SOST gene. This suggested that van Buchem disease in this family might be caused by a regulatory mutation influencing the expression of SOST, or by a mutation in another (related) gene located within the linkage interval. Since no interesting functional candidate gene for van Buchem disease was identified in this region, we analysed the complete linkage region between D17S1326 and D17S1860 for chromosomal rearrangements. Southern blot experiments on genomic DNA from patients from the Dutch van Buchem family showed the presence of a 52 kb deletion starting approximately 35 kb downstream of the SOST gene. Three other Dutch patients with the typical clinical features of van Buchem disease are also homozygous for the deletion. The absence of this deletion in a set of Dutch and random control DNA samples provided evidence that the deletion was not a polymorphism, but a disease causing mutation.

The deletion is caused by an unequal homologous recombination event between Alu sequences, which are located at the deletion breakpoints. A 16 bp sequence at the junction of the deletion is 100% identical at the proximal and distal sides, further substantiating our hypothesis of Alu mediated recombination. Alu repetitive sequences are frequently involved in homologous and non-homologous recombinations leading to disease.9,10 In silico analysis of the 52 kb deleted sequence by BLAST searches and exon prediction programs showed 19 putative coding sequences. However, we were unable to confirm experimentally that any of these represent coding sequences, even after extensive analysis. We therefore concluded that no genes are located within the deleted sequence. The absence of genes within the deleted region was verified and confirmed by comparison of the human genomic sequence of the deletion with the homologous mouse sequence. The absence of a putative coding sequence within the deletion raises the possibility that the deletion has an effect on the transcription of the SOST gene in the van Buchem patients.

Two different mechanisms could account for altered transcription: an alteration in the chromatin structure because of the deletion25 or the presence of a cis regulatory element within the deleted region. Mutations in cis regulatory elements located more than 10 kb from the disease gene have previously been shown to be responsible for other monogenic conditions, such as aniridia (PAX6), campomelic dysplasia (SOX9), and X linked deafness (POU3F4).13-29 Interestingly, using man-mouse sequence comparison, we identified two highly conserved non-coding regions, CNS-1 and CNS-2. Despite the fact that these sequences did not show any similarity with known regulatory elements, it has already been shown that conserved non-coding sequences influence the expression of genes. Transgenic mice carrying a deletion of a CNS located between the interleukin-4 (IL-4) and interleukin-13 (IL-13) genes showed a twofold to threefold reduction of the expression of IL-4, IL-13, and IL-5.19,30 Since SOST is not expressed in lymphocytes or other readily available tissue, we were not able to test its transcription in the van Buchem patients. We therefore measured the expression level of VHR phosphatase, a gene located approximately 10 kb distally from SOST (fig 2). No significant difference was found between deletion patients and controls, indicating that if the deletion alters the SOST transcription, it does so without involving the VHR phosphatase gene. In light of the findings reported by Loots et al,29 that deleting a CNS did not alter the expression of RAD50, a gene located between IL-13 and IL-5, it would not be surprising that the CNS elements within our chromosome 17 deletion regulate only a subset of genes that are able to the region.

In summary, we identified a deletion downstream of the SOST gene in Dutch van Buchem patients caused by homologous recombination between Alu sequences. In this deletion, no coding sequences could be identified. We propose a regulatory effect of the deletion on the SOST gene as the mechanism underlying van Buchem disease in these patients.

ACKNOWLEDGEMENTS

We thank the patients for contributing to this project. This study was supported by a concerted action grant from the University of Antwerp to WW and a grant (G.0404.00) from the “Fonds voor Wetenschappelijk Onderzoek (FWO)” to WW. WB holds a predoctoral research position with the “Vlaams Instituut ter Bevordering van Wetenschappelijk-Technologisch Onderzoek in de Industrie (IWT)”. WW is a postdoctoral researcher for the FWO.

Authors’ affiliations

W Balemans, E Van Hul, W Wuys, W Van Hul, Department of Medical Genetics, University of Antwerp and University Hospital Antwerp, Universiteitsplein 1, 2610 Antwerp, Belgium
N Patel, C Lacza, M Dioszegi, B Vickery, Musculoskeletal Research, Inflammatory Disease Unit, Roche Bioscience, Palo Alto, CA, USA
M Ebeling, Pharmaceuticals Division, F Hoffmann-La Roche Ltd, Basel, Switzerland
F G Dikkers, Department of Otorhinolaryngology, University Hospital Groningen, Groningen, The Netherlands
P Hildering, General Practitioner, Uik, The Netherlands
P J Williems, Department of Clinical Genetics, Erasmus University, Rotterdam, The Netherlands
J B G M Verheij, Department of Clinical Genetics, University Hospital Groningen, Groningen, The Netherlands
K Lindpaintner, D Foernzler, Roche Genetics, F Hoffmann-La Roche Ltd, Basel, Switzerland

REFERENCES

Deletion in van Buchem patients


36 Hardison RC. Conserved noncoding sequences are reliable guides to regulatory elements. TIG 2000; 16: 369-72.