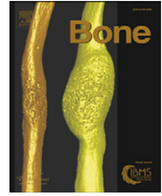
Bone 45 (2009) 736–742

Contents lists available at ScienceDirect

Bone

journal homepage: www.elsevier.com/locate/bone



Effects of a synonymous variant in exon 9 of the *CD44* gene on pre-mRNA splicing in a family with osteoporosis

Christopher Vidal [a,b], Adela Cachia [a], Angela Xuereb-Anastasi [a,b,*]

[a] *Department of Pathology, University of Malta, Medical School, G'Mangia, Malta*
[b] *Institute of Health Care, University of Malta, Mater Dei Hospital, Msida, Malta*

ARTICLE INFO

Article history:
Received 1 January 2009
Revised 23 June 2009
Accepted 25 June 2009
Available online 4 July 2009

Edited by: M. Noda

Keywords:
CD44
Osteoporosis
Linkage
Alternative splicing
RNA

ABSTRACT

In a previous linkage study, suggestive linkage to osteoporosis was observed in marker D11S1392 on chromosome 11p12. The *CD44* gene, found at this locus, was sequenced in one of the families studied. Sequencing all coding regions and promoter in affected and non-affected family members revealed a number of sequence variants, one of which was found to be linked and inherited identical by descent together with the linked STR allele. This G to A variant, which does not cause an amino acid change, was found in exon 9 of the *CD44* gene, 32 base pairs upstream from the exon–intron junction. Preliminary analysis using a bioinformatics tool suggested that the presence of the A allele abolished an exon splicing enhancer (ESE) site, thus possibly affecting RNA splicing. It was observed using an exon-trapping vector, that in the presence of the A allele, only one transcript was observed in RAW264.7 cells, as opposed to two transcripts transcribed in the presence of the G allele. These observations suggest that the linked synonymous variant found in exon 9 of the *CD44* gene might be increasing susceptibility to osteoporosis in this family by affecting the splicing mechanism.

© 2009 Elsevier Inc. All rights reserved.

Introduction

For the past 15 years, numerous studies have been performed in different populations using both linkage and association approaches, to try to identify genes that might increase the individual's susceptibility to osteoporosis [1,2]. Most of these studies did not give any conclusive results, and lacked replication due to reasons such as genetic heterogeneity between different populations and low sample power. Still such studies were useful in understanding better the complex pathophysiology of osteoporosis by identifying genes playing a role in various metabolic pathways and other mechanisms that might lead to disease. Linkage studies of monogenic bone diseases revealed several chromosomal regions putting into light various genes that were never thought to be involved in bone physiology such as the low density lipoprotein receptor-related protein (LRP)-5 and sclerostin (SOST) genes [3,4].

In a previous linkage study, performed on two Maltese families with a high incidence of osteoporosis, suggestive linkage to chromosome 11p12 was observed [5]. Fine-mapping reduced the linkage interval to a region between markers D11S1392 (50.64 cM) and D11S935 (52.94 cM), with highest total heterogeneity LOD and NPL (3.07 and 7.0, respectively) to marker D11S935. Although both families shared the same linkage interval, highest LOD scores were obtained to two different markers with a spacing of approximately 4 cM between them, showing also different inherited alleles, thus suggesting that different genes might be responsible for the disease in different families [5]. Highest LOD and NPL scores (1.77 and 5.9, respectively) were obtained for marker D11S1392 (50.64 cM) in Pedigree 2, while for Pedigree 1 highest scores were obtained to marker D11S4102 (54 cM). A total of 24 genes and four hypothetical proteins are known at the linkage interval from 49 to 55 cM on the deCode genetic map, of which tumour necrosis factor receptor-associated factor (TRAF)-6 and *CD44* genes were the most likely candidates based upon prior knowledge of physiology. Sequencing the *TRAF-6* gene, found between and approximately 1 cM away from both D11S935 and D11S4102, did not reveal any linked variations with disease. The other candidate gene at this locus is the *CD44* gene which is found 1 cM away from D11S1392 (~51 cM).

CD44 is a transmembrane glycoprotein existing in various isoforms, all consisting of a common hyaluronan-binding homology region, trans-membrane and cytoplasmic domains with most variant forms (CD44v) containing combinations of exons 6 to 15 (v1–v10) [6]. Osteoclast formation was inhibited by CD44 antibody suggesting its important role in bone physiology and as a potential therapeutic target for metabolic bone disease [7]. Expression of CD44 was observed to be induced at the onset of macrophage fusion where the intracellular domain of this molecule was found to play a major role in this process. Translocation of the CD44 intracellular domain to the nucleus also resulted in activation of NF-κβ which is essential for osteoclast activation [8]. Low molecular weight, but not high molecular weight,

* Corresponding author. Institute of Healthcare, University of Malta, Block A, Mater Dei Hospital, Msida, Malta.
E-mail address: angela.a.xuereb@um.edu.mt (A. Xuereb-Anastasi).

8756-3282/$ – see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.bone.2009.06.027

hyaluronic acid was observed to increase osteoclastogenesis induced by RANKL due to interactions with CD44 in bone marrow stromal cells which are involved in joint tissue destruction [9]. CD44 also plays an important role in cell migration, including that of osteoclasts, due to its interactions with osteopontin and other membrane proteins such as ezrin, radixin and moesin [10]. CD44 deficiency increased the severity of TNF-mediated arthritis due to deregulation of osteoclastogenesis [11]. Mice deficient in CD44 were more responsive to tumour necrosis factor (TNF) resulting in generalised osteopenia, while not affecting bone formation, suggesting the important role of CD44 in inflammatory bone loss. This increased responsiveness to TNF was associated with an increased activity of p38 mitogen-activated protein kinase which is negatively controlled by CD44 [12].

In this study, sequencing of the *CD44* gene found on chromosome 11p12 was performed in Pedigree 2 since this gene is found closer to D11S1392 which shows the highest LOD scores within this family. DNA sequencing of the *CD44* gene revealed a rare sequence variant that was linked with the inherited haplotype in the majority of affected members.

Materials and methods

Subjects

Sequencing of the *CD44* gene was performed in an extended family (Pedigree 2) participating in a previous linkage study [5] (Fig. 1 and Table 1). Affected status was defined according to WHO criteria (t-score < -2.5) for generation II, while z-scores were used for younger individuals in generation III. As shown in Fig. 1, a linked haplotype for markers D11S1392 and D11S935 (alleles 3 and 1, respectively) inherited identical by descent was observed in the majority of affected individuals.

A random sample of anonymous cord blood DNA ($n = 126$) obtained from newborns was used to estimate the frequency of the linked CD44 exon 9 variant in the population at birth. Cord blood was collected from neonates and taken to the Laboratory of Molecular Genetics, University of Malta, Malta for routine testing. Genomic DNA was extracted and purified using a salting out technique [13].

All affected members (9 individuals) from Pedigree 1 were also tested for the CD44 exon 9 variant.

Sequencing of CD44 gene

Oligonucleotide primers were designed to amplify all coding regions including intron–exon boundaries and promoter region of the *CD44* gene (MIM 107269) using transcript ENST00000278385) from the ENSEMBL database (http://www.ensembl.org) (Table 2). Initial sequencing was performed using BigDye® Terminator v3.1 cycle sequencing kit (Applied Biosystems, Foster City, CA, USA) in three affected (II:6, III:1, III:9) and one unaffected (III:6) individuals (with and without haplotype) and any sequence variants identified were further analysed in all family members by sequencing or by using restriction fragment length polymorphism (RFLP). The frequency of the identified linked variant in the *CD44* gene was determined in the general Maltese population using allele specific amplification.

Allele specific amplification

The linked variant identified in exon 9 of the *CD44* gene was analysed in the general population by polymerase chain reaction using confronting two-pair primers (PCR-CTPP) [14]. Two sets of primers were designed to amplify the specific alleles separately at optimum annealing temperature. A fragment of 163 bp was amplified using a forward primer with its last base at the 3′ end specific for the G allele while the A allele was detected by the amplification of 545 bp using a specific reverse primer. This reaction was carried out in a single tube where the forward primer used to amplify the A allele, together with the reverse primer of the G allele fragment, yielded another largest fragment of 666 bp that was used as an internal PCR control. The optimum PCR annealing temperature was that of 63 °C. The best results were obtained when using 15 pmol of each primer for allele G (163 bp), together with 25 pmol of primers used for the A allele. PCR thermal profile used was as follows: hot-start for 15 min at 95 °C followed by 30 cycles of 95 °C for 30 s, 63 °C for 15 s, 72 °C for 90 s. A final extension for 5 min at 72 °C was also performed. This method was validated against results obtained by direct DNA sequencing, while 20% of samples were repeated using separate PCR reactions for each allele.

RNA splicing analysis

An online application was used to identify any possible exon splicing enhancer (ESE) at this region as described by Cartegni et al. [15] (http://rulai.cshl.edu/tools/ESE2/index.html). The G/A variant was found to abolish an ESE motif (TGAGGA) for the SR protein (SRp55) with a score of 2.817 (threshold 2.676), in the presence of the A allele. The online application RESCUE-ESE (http://genes.mit.edu/fas-ess/) did not predict any possible ESEs at this locus [16]. A mini-gene system was used to test whether this linked variant found in exon 9 might affect splicing of pre-mRNA. Using this system, a hybrid construct

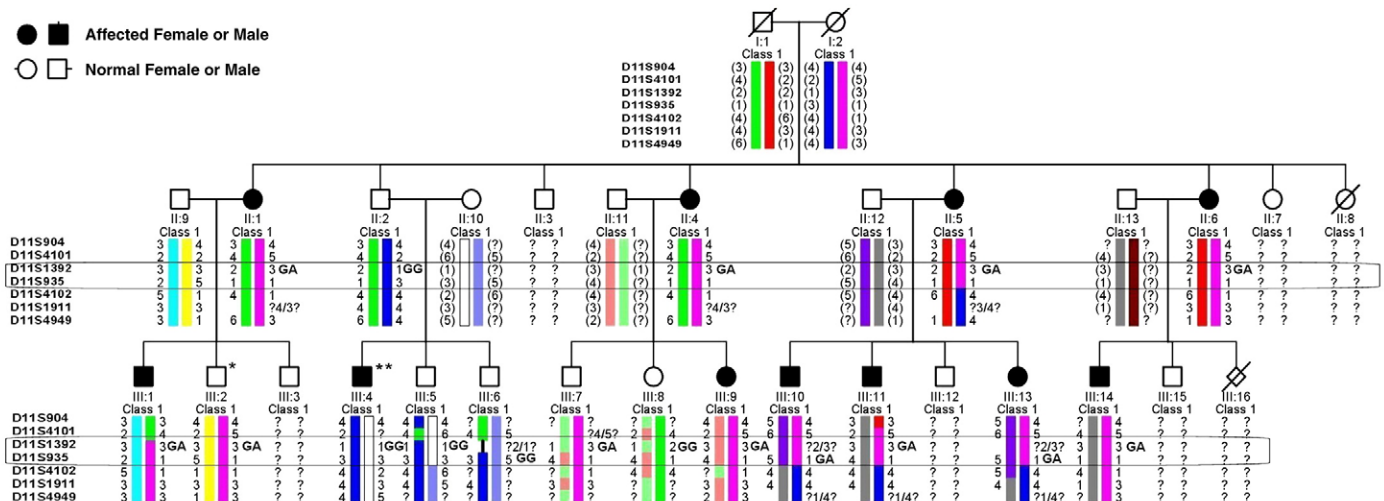


Fig. 1. Structure of Pedigree 2 showing linked haplotype and CD44 exon 9 variant.

made up of two vector β-globin exons flanking CD44 exon 9 and adjacent introns was created. Following transfection into mammalian cells, the construct was transcribed under the control of a SV40 promoter and spliced. The mRNA derived from this construct was extracted and reverse transcribed followed by specific amplification using cDNA as template and specific primers to β-globin exons (SD6 and SA2). The spliced transcripts were analysed by agarose gel.

DNA cloning

A 523 bp fragment containing the whole exon 9 and adjacent introns, was amplified by PCR using template DNA from a heterozygous family member for this variant, to isolate both mutated and normal alleles. The forward and reverse oligonucleotide primers were designed harbouring consensus sequences for endonucleases Xho I and Pst I, respectively at the 5′ end together with a non-complementary tail (Fwd: 5′-gTTCTCAATgCggAACTC-gAggTTTgTggTTCTgCCTTTgAC-3′ and Rev: 5′-gTTCTTgAACTTgCACTg-CAgAATTTgTgTgCgggTTTAgC-3′). The amplified fragment together with the exon-trapping vector pSPL3 (generously donated by Prof. Junko Oshima, Department of Pathology, University of Washington, Washington, USA) were digested and ligated overnight at 4 °C, using T4 ligase (Promega Corporation, Madison, USA) (Fig. 2). Transformation in DH5α was carried out using a heat shock protocol, first by incubating 25 µl cells with 2.5 µl ligation reaction for 20 min on ice. Following incubation on ice, the cells were subjected to heat shock treatment for 45–50 s at 42 °C in a heating block and placed again on ice for another 2 min. Ampicillin free Luria Bertani (LB) broth (475 µl) was added into each tube and the cell suspension was incubated at 37 °C for 1 h in a shaking incubator at 150 rpm. Following incubation, 150 µl from each suspension were plated onto $LBAmp^{100}$ selective plates and incubated at 37 °C overnight. Selected clones were screened by digestion and DNA sequencing to check the integrity of the sequence and to select the two alleles. Purified plasmid DNA for transfections was extracted from a 100 ml culture grown for 16 h at 37 °C using the PureYield™ Plasmid Midiprep System (Promega Corporation, Madison, USA). DNA was adjusted to a concentration of 200 ng/µl.

Cell culture

Three different cell lines were used for transfections, including COS-7, HeLa (DSMZ, Braunschweig, Germany) and RAW264.7 (InterLab Cell Line Collection, Institute of Cancer Research, Genova, Italy). All cells were cultured in a high glucose Dulbecco Modified Eagle's Medium (DMEM) supplemented with 10% heat-deactivated foetal calf serum (FCS) (Sigma, Saint Louise, Missouri, USA). Incubation was carried out at 37 °C in 5% CO_2, splitting cells at 80% confluency. The day before transfection, 5×10^4 cells were added into each well of a 24-well plate. Each sample was assayed in triplicate, including positive (native pSPL3) and negative controls (cells without the addition of any DNA). Transient transfection was performed using the cationic lipid transfection reagent Tfx™-20 (Promega Corporation, Madison, USA) using 1 µg of each DNA construct and a charge ratio of DNA to transfection reagent of 1:2 for all cell lines. The mixture consisting of DNA constructs, transfection reagent and serum free DMEM was prepared in 1.5 ml microcentrifuge tubes and incubated for 10–15 min at room temperature. Culture medium was aspirated from cells and 200 µl of transfection mixture was added into each well. Incubation was carried out at 37 °C for 1 h in 5% CO_2. After incubation, 500 µl of complete medium was added into each well and plates were incubated for 24 h at 37 °C in 5% CO_2. RAW264.7 cells were also incubated in the presence of 10 ng/ml M-CSF, with or without 50 ng/ml RANKL (R&D Systems Ltd, Abington, UK). Stimulation with these cytokines was initiated 48 h before transfection.

Reverse transcription and PCR

After incubation, the cells were lysed and RNA extracted using RNeasy® mini kit (Qiagen Gmbh, Germany) as instructed by the supplier. First strand cDNA was synthesised by reverse transcription using oligo-$(dT)_{16}$, followed by PCR using specific primers for the splice donor (SD6) and acceptor (SA2) of the pSPL3 vector. The products were analysed using a 2% agarose gel electrophoresis and visualised under UV. All transcripts were excised from the gel and sequenced.

Results

Sequencing results

Sequencing CD44 revealed a number of intronic sequence variants, including two A/G changes (rs4756196 and rs3736812) and an A/C transversion in intron 16, none of which was observed to be inherited with the linked marker. A number of other variants were found in coding regions, including an A/G (rs9666607) and C/T (rs11607491) change in exon 10, both resulting in an amino acid change, which was not linked with the inherited haplotype. Another C/T synonymous variant (rs35356320) was detected in three affected and one unaffec-

Table 1
BMD values of family.

Participant	Age	BMI	Lumber BMD (g/cm^2)	LS *t*-score (*z*-score)	Femoral BMD (g/cm^2)	FN *t*-score (*z*-score)	Rs11033026
II: 1[a]	55	26.5	0.703	−3.47 (−1.68)	0.731	−1.57 (−0.56)	GA
II: 2	66	27.0	0.955	−1.19 (0.57)	0.811	−0.19 (0.62)	GG
II: 4	61	28.8	0.792	−2.66 (−0.78)	0.716	−1.70 (−0.37)	GA
II: 5	70	29.1	0.768	−2.88 (−0.78)	0.652	−2.24 (−0.65)	GA
II: 6	52	21.2	0.672	−3.75 (−2.18)	0.594	−2.72 (−1.86)	GA
III: 1	29	25.3	0.890	−2.00 (−2.02)	0.862	−1.22 (−1.09)	GA
III: 2	23	19.7	0.941	−1.58 (−0.87)	0.877	−1.10 (−0.36)	GA
III: 4	41	30.1	0.930	−1.67 (−1.42)	0.909	−0.86 (−0.45)	GG
III: 5	38	23.5	1.071	−0.13 (−0.03)	0.852	−0.57 (−0.37)	GG
III: 6	33	24.1	1.017	−0.62 (−0.60)	0.953	−0.27 (0.25)	GG
III: 7	33	31.8	1.079	−0.42 (−0.36)	0.989	−0.24 (0.04)	GA
III: 8	38	30.9	1.375	2.64 (2.42)	1.089	1.41 (1.56)	GG
III: 9	42	25.4	0.854	−2.10 (−1.62)	0.707	−1.77 (−1.40)	GA
III: 10	37	28.1	0.851	−2.32 (−2.19)	0.877	−1.10 (−0.76)	GA
III: 11	42	25.3	0.963	−1.39 (−1.13)	0.750	−2.08 (−1.66)	GA
III: 13	39	24.5	0.768	−1.46 (−1.19)	0.652	−0.56 (−0.33)	GA
III: 14	25	21.1	0.976	−1.28 (−1.36)	0.832	−1.45 (−1.43)	GA

Participant numbers refer to those in Fig. 1.
[a] Proband.

Table 2
Primers used for *CD44* gene sequencing.

	Primer sequences 5′-3′	Fragment size/bp	Annealing temp./°C
Promoter 1	AgA TCT gCT ggg TAg gAA AgA CCA TCC ACT AAC CAC TTA ggT C	376	64
Promoter 2	ATg gTg gAT ggT TgT ggT TT ggA gAg CTC ATT CTT TTT CCC	449	64
Promoter 3	gAA TgA gCT CTC CCT CTT TCC AC CTg CTg Agg CTg TAA ATA ATC gg	398	64
Promoter 4—Exon 1	gCA gCC CCg ATT ATT TAC AgC gTg CCA CCA AAA CTT gTC CA	359	64
Exon 1	AAA AAC TgC AgC CAA CTT CC TgC TTC CAC AgA CAC ATT CTC	417	64
Exon 2	ggA gTC TgT CCT AAA CTg AAC ACA ggT TgC AgA TTg ACT Tg	402	64
Exon 3	TCT CTT TCC CTC TCT ATg CTg T ggC TAC ATg AAT CCC CAC TTA	401	60
Exon 4	ATT AgA gAA ggA gCC gTg TCT A CAA AgT gCT ggg ATT ACA ggA	532	64
Exon 5	CTC CCA CCA CTg gAT AgA TAg gT CTT CCA gCT gTC TAg AgA CgA gAA	477	64
Exon 6	TAA gAT CCT TCC CTC TCT gA ggC TgT CAA gAA AAC ATA Cg	383	60
Exon 7	TTT ggg gTA TgC AAg TCA ACT C CCT TgC ATA AAT ggC TgT ACT g	305	64
Exon 8	CCA TgC AgC CAT CTA TAC AAC CT TgT CAA Agg CAg AAC CAC AAA C	499	64
Exon 9	gAg Tgg ATg CTC AgA ggT AAC A ACC gAg Tgg AAA TAg AAg ATg C	353	64
Exon 10	CCT Tgg TgC CTT TCT TTT CTA CC ggC TTg TAT CCA TTC CTg TTC Ag	454	64
Exon 11	CTg CgT TTA TgC AAC TTC CTT g CCC TCC TgC TAA ATT CCT TAC T	497	64
Exon 12	CTC AgT TCg ggA gAT ATA gTC Ag gAg ACC TCC ACA AgA AAT ACA gAC	412	64
Exon 13	CTC TgT gTT CTC Tgg TTT TCT ggA AgT gTA ACg gCA gCT AAT ggA Tg	533	64
Exon 14	gTg TTC Tgg AgA CAA gCA CAT AgT CCA ATC CCA gAC TAC ACA g	380	64
Exon 15	CTg Tgg TCA TCC TTC TTA gCC TTT ggA gAC ACT gAC TCA TAC TTC gTT g	407	64
Exon 16	CTT TAT gCA gCT CCA CAA ggA AC CTA gTT TgC AgA ACC CAg gAA gA	365	64
Exon 17	gCT CTT CCA gAA TTA CCT gCC TAT ggT gTC TCC TgT CTC TAA AAA CCg	500	64
Exon 18	Tgg TgC TCA gTA Agg ATg ATg ATC gTT AAg TgT CCC AgC TCC CTg TAA	355	64

ted individual. A non-synonymous variant found in exon 12 (rs1467558) was only found in two affected members of this family.

An interesting variant was detected in exon 9, a synonymous G/A transition (rs11033026), found 32 nucleotides upstream from the exon/intron junction. Sequencing the gene in all members of this family, revealed that all individuals carrying the linked STR allele 3 (Fig. 1) for marker D11S1392 were also heterozygous for this variant, suggesting that the two were linked. All affected members, with the exception of one (III:4), were heterozygous for both the STR allele and the A allele. This variant was not found in any of the non-affected family members, with the exception of two who also carried the linked STR allele (III:2 and III:7). Mean lumbar and femoral BMD (and *z*-scores) were compared between family members heterozygous for the A allele ($n=12$) and homozygous GG ($n=5$) (Table 3). The independent sample *t*-test was used to test for significance. As shown in Table 3, both lumbar and femoral BMD (including *z*-scores) were statistically significant when comparing GG with GA ($p<0.05$).

Using allele specific amplification, 252 chromosomes were analysed in a random sample from the general Maltese population. Only 3 A alleles were identified in 3 heterozygous individuals, thus giving a population frequency of 2.38% and a minor allele frequency of 0.012 (1.19%). None of the nine affected members from the other pedigree (Pedigree 1), used in the previous linkage study [5], was found to carry the A allele.

RNA splicing analysis

Fig. 3 shows the results obtained from RT-PCR from cell lysates following transfections with the exon-trapping vector pSPL3. As shown in all three cell lines, two transcripts were amplified in the presence of the G allele (378 bp and 261 bp), while for the A allele the smaller transcript (261 bp) was weakly amplified in both COS-7 and HeLa cells but it was completely absent in RAW264.7 macrophages, even when stimulated by m-CSF, with or without RANKL. The sequencing of transcripts showed that the bigger fragment of 378 bp consisted of two vector exons, also harbouring exon 9 of CD44. The smaller fragment did not contain any portion of CD44 exon 9, and was entirely made up of vector exon sequences, suggesting skipping of exon 9. The presence of the G and A allele was confirmed when sequencing the bigger fragments.

Discussion

A large number of genes known to be involved in bone physiology have been studied to determine whether any variations within them might increase the individual's susceptibility to osteoporosis [1,17]. In this study, the effects on pre-mRNA splicing of a known synonymous variant, found in exon 9 of the *CD44* gene (rs11033026), are being reported. The A allele is in linkage with a haplotype inherited identical by descent in all affected family members, previously studied using qualitative linkage analysis [5]. Two clinically normal individuals (III:2 and III:7) also carried this allele in the heterozygous state, in linkage with the STR marker, suggesting incomplete penetrance possibly partly due to their relatively young age (23- and 33-year-old males). The BMD of individual III:7 might also have been positively influenced by his high body mass index (BMI) of 31.8 kg/cm^2, although one cannot exclude the possibility of other gene modifiers being involved. Conversely, individual III:4 was affected, but did not carry the CD44 A allele or the linked STR allele, suggesting that low BMD in this individual might be due to environmental influences (phenocopy).

This variant was not found in nine affected individuals from Pedigree 1 used in the previous linkage study [5], suggesting that it is increasing susceptibility to the disease only in Pedigree 2. Different mutations within the same or different gene/s present within the same linkage interval might be increasing the susceptibility to disease in these two families (showing allelic heterogeneity at the same locus).

The G/A variant (rs11033026) was only found in 2.38% of the general local population. When searching the NCBI SNP database (http://www.ncbi.nlm.nih.gov/SNP/snp_ref.cgi?rs=11033026), the HapMap data, revealed that this variant was not reported in the European Caucasian population, but was found in Sub-Saharan Africans, African-Americans and Han Chinese from Beijing (minor allele frequencies 0.336, 0.115, and 0.012, respectively). This suggests a founder effect in the Maltese population, complementing other previously reported studies on other human diseases [18,19].

Since this variant does not result in an amino acid change, the hypothesis that it could affect pre-mRNA splicing was tested using an exon-trapping vector. In all three cell lines used, two transcripts were amplified in the presence of the wild-type G allele, one transcript harbouring the complete CD44 exon 9 and a smaller one consisting entirely of the two vector exons. In the presence of the A allele, the quantity of the shorter transcript relative to the longer one was less in both epithelial cells (HeLa and COS-7) but was completely absent in murine macrophages, suggesting that this variant could affect different tissues in a specific manner.

The presence of more than one transcript was expected since this gene is known to be transcribed and translated into a number of isoforms, due to alternative splicing of a large number of its exons [20].

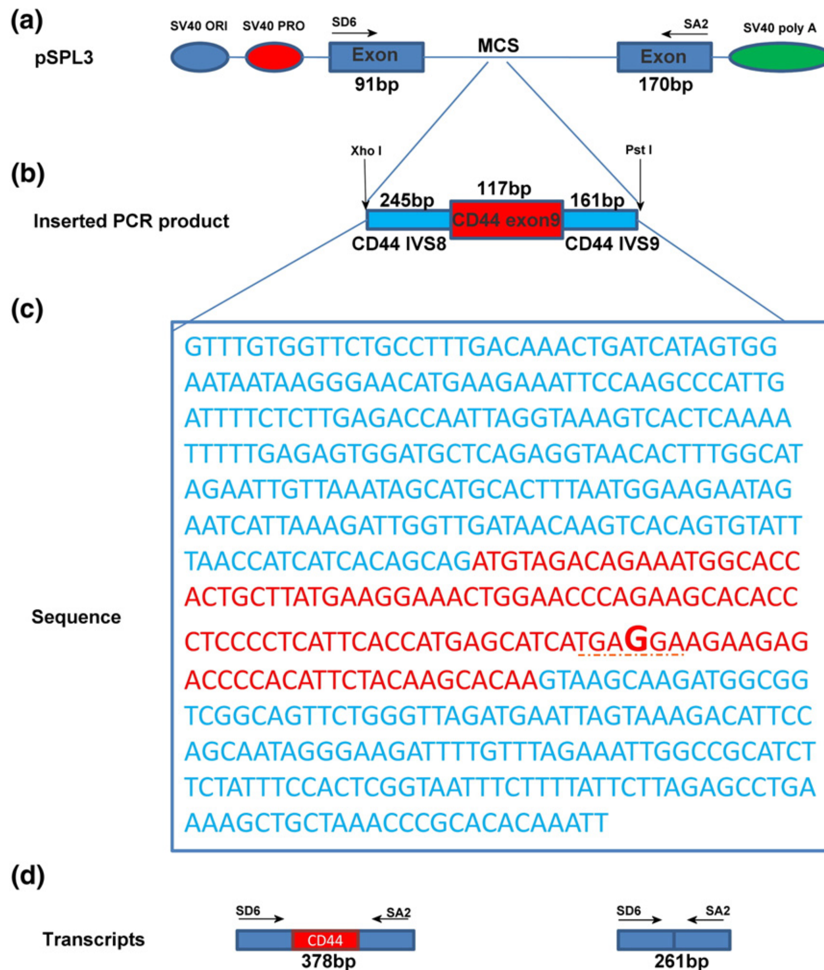


Fig. 2. Schematic diagram of the pSPL3 construct. (a) Basic structure of the exon-trapping vector (pSPL3) showing SV40 origin of replication (ORI) and promoter (PRO) together with two exons, one on each side of the multiple cloning site (MCS). Primer binding sites used for specific amplification following reverse transcription are marked as SD6 and SA2. An SV40 poly A tail is also shown. This diagram does not show the ampicillin resistance gene used for selection. (b) Schematic diagram of the amplified PCR product inserted into the MCS using restriction enzymes Xho I and Pst I. The diagram shows exon 9 and parts of adjacent introns or intervening sequences (IVS 8 and IVS 9). The sequence of the PCR product is shown in (c). The exon is coloured red with the identified variant marked with a bigger font. The identified ESE is underlined. (d) shows a schematic diagram of the expected transcripts following amplification using SD6 and SA2. The longer transcript harbouring exon 9 and two plasmid exons while the shorter one consists only of plasmid exons.

Different CD44 isoforms were shown to have various functional roles in epidermal Langerhans and dendritic cells, such as in cell migration and antigen presentation. A pattern in expression was also observed during migration of these cells from the epidermis to the lymph nodes, where up to 100% of Langerhans cells in the dermis expressed v5, v6 and v9 while those in the lymph nodes expressed v4, v5 and v6 [21]. Shifts in expression between different CD44 isoforms were also reported to be useful as prognostic markers for various carcinomas, including osteosarcoma [22].

Table 3
Intra-familial analysis for rs11033026.

	Rs11033026	n	Mean	p-value (t-test)
BMI (kg/cm^2)	GG	5	27.1	0.453
	GA	12	25.6	
Lumbar BMD (g/cm^2)	GG	5	1.069	0.015*
	GA	12	0.868	
Lumbar z-score	GG	5	0.188	0.010*
	GA	12	−1.278	
Femoral BMD (g/cm^2)	GG	5	0.923	0.037*
	GA	12	0.787	
Femoral z-score	GG	5	0.322	0.004*
	GA	12	−0.869	

* Statistically significant at level of significance $p<0.05$.

The important role of the *CD44* gene in the pathophysiology of osteoporosis is being documented for the first time. The role of CD44 in macrophage fusion is important since this process occurs during the differentiation of osteoclasts. Since macrophages are the precursors of osteoclasts, it could be hypothesised that this variant shifts the balance between different CD44 isoforms to that harbouring exon 9 (CD44v4) in osteoclasts, thus affecting their function and mechanisms, such as cell migration and fusion. Such a change in function could increase the risk of osteoporosis. The CD44v4 isoform (harbouring exon 9) was shown to be a major E-selectin ligand in breast cancer cells facilitating cell migration across the endothelium and hence enhancing metastasis [23]. Our study is of particular importance since a synonymous variant, located within the coding region of a gene, was found to increase the risk of osteoporosis in one family by affecting RNA splicing mechanism. This indicates that one cannot underestimate DNA variations found in non-coding regions and synonymous exonic variants. To our knowledge, to date, no variations that result in alternative splicing have been associated with osteoporosis, except for one variant found within intron 4 of the calcitonin-related peptide in an isolated case of an osteoporotic young male with very low calcitonin levels [24]. A number of variants were found to affect splicing in several genes that cause disease, such as the cystic fibrosis transmembrane conductance regulator (*CFTR*) gene [25], neurofibromatosis type 1 [26], type 5 collagen and Ehler–Danlos syndrome [27], superoxide dismutase

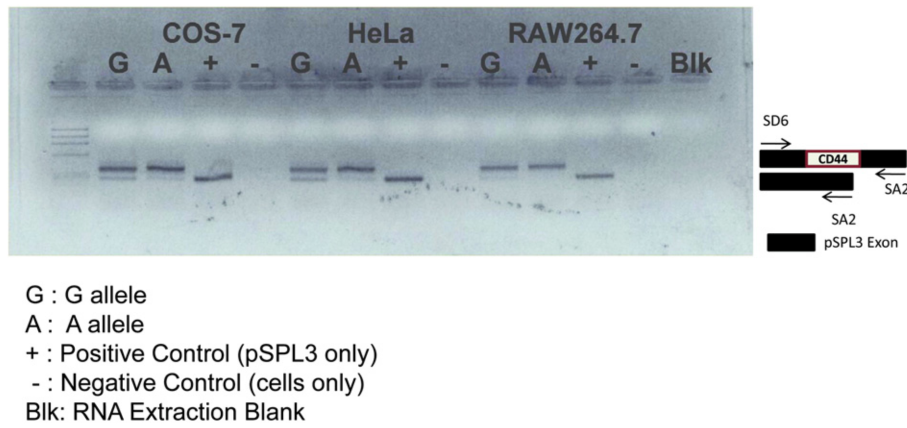


Fig. 3. Agarose gel electrophoresis of CD44 pSPL3 transcripts amplified by reverse transcriptase PCR.

(SOD2) [28] and the ATP-binding cassette transporter A1 (ABCA1) [29]. RNA splicing can be affected by variations that alter the two dimensional structure of pre-mRNA molecule affecting its availability to interact with splicing factors or by increasing/decreasing recognition sites in both introns or exons for splicing enhancers (ESE) or silencers (ESS) [30]. A possible ESE motif that was predicted using a bioinformatics tool in the presence of the G allele was absent in the A allele (TGAGGA >TGAAGA), possibly affecting a splice site enhancer binding site that resulted in mRNA only harbouring exon 9 of this gene. The deleterious effects of missense mutations and the mechanisms by which these could result in aberrant splicing were described by other investigators. It was suggested that there might be a higher risk that functional ESEs cause aberrant splicing when these are found closer to the 3′ end of exons [31].

A common difficulty encountered when studying putative splicing mutations is to obtain RNA directly from the appropriate tissues to exclude cell specific effects, in this case, bone tissue. An alternative option is to use an exon-trapping system, as was used in this study which although reliable for initial evaluation of the splicing mechanism, does not always reflect the *in vivo* splicing mechanism [32]. A disadvantage of this system is that it is a hybrid, where in the ideal situation all exons and introns should come from the same gene. A major advantage, however, is that very short DNA fragments can be inserted avoiding the problem of handling thousands of nucleotides usually making up most mammalian introns [33].

The importance of synonymous DNA variations together with those found in non-coding regions of genes, and the mechanisms by which such variants could be affecting susceptibility to disease, is slowly emerging. More evidence is showing that these variants can still affect the phenotype not only by altering RNA splicing but also by affecting protein folding, without a change in the amino acid sequence. Co-translational protein folding might be affected by timing as a result of a synonymous variant changing a commonly used codon into a rarer one, also coding for the same amino acid [34]. The time of translation is dependent upon the abundance of tRNA present in cells which correlates with codon usage in a tissue specific manner [35]. The rarer the codon, the less abundant is the corresponding tRNA, and thus the longer the time taken for translation to progress. Recent studies have shown that the enzyme specificity of the *MDR1* gene was changed by the presence of a synonymous variant (C3435T), indicating that both function and 3-D structure of a protein could be altered without a corresponding change in the amino acid sequence [34,36]. Such studies addressing the effects of these variants on protein function are important not only in the understanding of the pathophysiology of disease but also in the pharmacokinetics of different proteins. This could contribute towards the development of better targeted and individualised treatment regimens. These findings show that it is not enough to consider the DNA sequence of a gene. Studies have to investigate the functional role of variants at the RNA and protein levels, and also the epigenetic control of gene expression and post-translational modifications of proteins.

Further studies are needed to evaluate the role of the *CD44* gene in bone physiology, osteoporosis, inflammatory bone loss and other bone diseases, including its possible role in metastasis of bone tumours.

Acknowledgments

We would like to thank Dr. Anthony Fenech Ph.D., B.Pharm. (Hons.) for the use of facilities at the Department of Pharmacology and Clinical Therapeutics, University of Malta, and Prof. Junko Oshima, Department of Pathology, University of Washington, Washington, USA, who generously donated the pSPL3 plasmid. We thank MLS BioDNA Ltd. for the use of facilities for DNA sequencing. This project was approved by the Research Ethics Committee, and supported by the Research Fund Committee of the University of Malta.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.bone.2009.06.027.

References

[1] Styrkarsdottir U, Halldorsson BV, Gretarsdottir S, Gudbjartsson DF, Walters GB, Ingvarsson T, et al. Genetic variation in candidate osteoporosis genes, bone mineral density and fracture risk: the study of osteoporotic fractures. Calcif Tissue Int 2008;83:155–66.
[2] Kaufman JM, Ostertag A, Saint-Pierre A, Cohen-Solal M, Boland A, Van Pottelbergh I, et al. Genome-wide linkage screen of bone mineral density (BMD) in European pedigrees ascertained through a male relative with low BMD values: evidence for quantitative trait loci on 17q21–23, 11q12–13, 13q12–14, and 22q11. J Clin Endocrinol Metab 2008;93:3755–62.
[3] Gong Y, Slee RB, Fukai N, Rawadi G, Roman-Roman S, Reginato AM, et al. Osteoporosis-Pseudoglioma Syndrome Collaborative Group. LDL receptor-related protein 5 (LRP5) affects bone accrual and eye development. Cell 2001;107:513–23.
[4] Brunkow ME, Gardner JC, Van Ness J, Paeper BW, Kovacevich BR, Proll S, et al. Bone dysplasia sclerosteosis results from loss of the *SOST* gene product, a novel cystine knot-containing protein. Am J Hum Genet 2001;68:577–89.
[5] Vidal C, Galea R, Brincat M, Xuereb-Anastasi A. Linkage to chromosome 11p12 in two Maltese families with a highly penetrant form of osteoporosis. Eur J Hum Genet 2007;15:800–9.
[6] Thorne RF, Legg JW, Isacke CM. The role of CD44 transmembrane and cytoplasmic domains in co-ordinating adhesive and signalling events. J Cell Sci 2004;117:373–80.
[7] Kania JR, Kehat-Stadler T, Kupfer SR. CD44 antibodies inhibit osteoclast formation. J Bone Miner Res 1997;12:1155–64.
[8] Cui W, Zhang Ke J, Zhang Q, Ke HZ, Chalouni C, Vignery A. The intracellular domain of CD44 promotes the fusion of macrophages. Blood 2006;107:796–805.
[9] Ariyoshi W, Takahashi T, Kanno T, Ichimiya H, Takano H, Koseki T, et al. Mechanisms involved in enhancement of osteoclast formation and function by low molecular weight hyaluronic acid. J Biol Chem 2005;280:18967–72.

[10] Denhardt DT, Noda M, O Regan AW, Pavlin D, Berman JS. Osteopontin as means to cope with environmental insults: regulation of inflammation, tissue remodelling and cell survival. J Clin Invest 2001;107:1055–61.

[11] Hayer S, Goertz B, Zwerina J, Redlich K, Hoffmann O, Amling M, et al. CD44 regulates bone erosion and osteoclastogenesis in arthritis. Arthritis Res Ther 2004;6:S17.

[12] Hayer S, Steiner G, Görtz B, Reiter E, Tohidast-Akrad M, Amling M, et al. CD44 is a determinant of inflammatory bone loss. J Exp Med 2005;201:903–14.

[13] Miller LJ, Dykes DD, Polesky HF. A simple salting out procedure for extracting DNA from human nucleated cells. Nucleic Acids Res 1988;16:1215.

[14] Hamajima N, Saito T, Matsuo K, Tajima K. Competitive amplification and unspecific amplification in polymerase chain reaction with confronting two-pair primers. J Mol Diagn 2002;4:103–8.

[15] Cartegni L, Wang J, Zhu Z, Zhang MQ, Krainer AR. ESE finder a web resource to identify exonic splicing enhancers. Nuc Acid Res 2003;31:3568–71.

[16] Wang Z, Rolish ME, Yeo G, Tung V, Mawson M, Burge CB. Systematic identification and analysis of exonic splicing silencers. Cell 2004;119:831–45.

[17] Huang QY, Kung AWC. Genetics of osteoporosis. Mol Genet Metab 2006;88: 295–306.

[18] Farrugia R, Scerri CA, Attard Montalto S, Parascandalo R, Neville BGR, Felice AE. Molecular genetics of the tetrahydrobiopterin (BH4) deficiency in the Maltese population. Mol Genet Metab 2007;90:277–83.

[19] Koziell A, Grech V, Hussain S, Lee G, Lenkkeri U, Tryggvason K, et al. Genotype/phenotype correlations of NPHS1 and NPHS2 mutations in nephrotic syndrome advocate a functional inter-relationship in glomerular filtration. Hum Mol Genet 2002;11:379–88.

[20] Screaton GR, Bell MV, Jackson DG, Cornelis FB, Gerth U, Bell JI. Genomic structure of DNA encoding the lymphocyte homing receptor CD44 reveals at least 12 alternatively spliced exons. Proc Natl Acad Sci 1992;89:12160–4.

[21] Weiss JM, Sleeman J, Renkl AC, Dittmar H, Termeer CC, Taxis S, et al. An essential role for CD44 variant isoforms in epidermal Langerhans cells and dendritic cell function. J Cell Biol 1997;137:1137–47.

[22] Kim HS, Park YB, Oh JH, Jeong J, Kim CJ, Lee SH. Expression of CD44 isoforms correlates with the metastatic potential of osteosarcoma. Clin Orthop Relat Res 2002;396:184–90.

[23] Zen K, Liu DQ, Guo YL, Wang C, Shan J, Fang M, et al. CD44v4 is a major E-selectin ligand that mediates breast cancer cell transendothelial migration. PLoS One 2008;3:e1826.

[24] Alevizaki M, Stevenson JC, Girgis SI, MacIntyre I, Legon S. Altered calcitonin gene in young patient with osteoporosis. British Med J 1989;298:1215–6.

[25] Hefferon TW, Groman GD, Yurk CE, Cutting GR. A variable dinucleotide repeat in the *CFTR* gene contributes to phenotype diversity by forming RNA secondary structures that alter splicing. Proc Natl Acad Sci 2004;101:3504–9.

[26] Bottillo I, De Luca A, Schirinzi A, Guida V, Torrente I, Calvieri S, et al. Functional analysis of splicing mutations in exon 7 of *NFI* gene. BMC Med Genet 2007;8:4.

[27] Burrows NP, Nicholls AC, Richards AJ, Luccarini C, Harrison JB, Yates JRW, et al. A point mutation in an intronic branch site results in aberrant splicing of COL5A1 and in Ehler–Danlos syndrome type II in two British families. Am J Hum Genet 1998;63:390–8.

[28] Shao J, Chen L, Marrs B, Lee L, Huang H, Manton KG, et al. SOD2 polymorphisms: unmasking the effect of polymorphism on splicing. BMC Med Genet 2007;8:7.

[29] Hong SH, Rhyne J, Miller M. A novel polypyrimidine variation (IVS46: del T−39...−46) in ABCA1 causes exon skipping and contributes to HDL cholesterol deficiency in a family with premature coronary disease. Circulation Res 2003;93:1006–12.

[30] Baralle D, Baralle M. Splicing in action: assessing disease causing sequence changes. J Med Genet 2005;42:737–48.

[31] Garlov IP, Garlova OY, Frazier ML, Amos CI. Missense mutations in the hMLH1 and hMLH2 are associated with exonic splicing enhancers. Am J Hum Genet 2003;73: 1157–61.

[32] Schneider B, Koppius A, Sedlmeier R. Use of an exon-trapping vector for the evaluation of splice site mutations. Mamm Genome 2007;18:670–6.

[33] Baralle D, Baralle M. Splicing in action: assessing disease causing sequence changes. J Med Genet 2005;42:737–48.

[34] Kimchi-Sarfaty C, Oh JM, Kim IW, Sauna ZE, Calcagno AM, Ambudkar SV, et al. A “silent” polymorphism in the *MDR1* gene changes substrate specificity. Science 2007;315:525–8.

[35] Dittmar KA, Goodenbour JM, Pan T. Tissue-specific differences in human transfer RNA expression. PLOS Genet 2006;2:2107–15.

[36] Komar AA. Silent SNPs: impact on gene function and phenotype. Pharmacogenomics 2007;8:1075–80.