Abstract
Osteoporosis is a complex metabolic disease with a strong genetic component which results in a decrease in bone mineral density and deterioration in the microarchitecture of bone, leading to an increased risk of fractures. During the last decade, a number of genetic studies of the Maltese population were performed to determine a potential genetic component which increases the individual’s susceptibility to osteoporosis. Both family and population case-control approaches were used. Linkage analysis using large affected families has identified a region on chromosome 11p12, following which functional studies have shown that variations within two genes found in that region namely, TRAF6 and CD44, affect gene expression, and, more specifically, pre-mRNA splicing in the case of CD44. This is the first report of strong suggestive linkage of 11p12 to osteoporosis. Using the candidate gene approach, selected candidate genes, which had been studied in other populations and been found to increase susceptibility to osteoporosis to varying degrees in some of them, included those coding for the vitamin D and oestrogen receptors (VDR & ER1), osteoprotegerin (TNFRSF11B), collagen type 1 (COL1A1), methylenetetrahydrofolate reductase (MTHFR) and lipoprotein receptor related protein (LRP)-5. In the Maltese population, the most significant result showed that a polymorphism in the promoter region of TNFRSF11B, which codes for osteoprotegerin, had a significant effect on bone mineral density, thus increasing the susceptibility to osteoporosis. These studies have shown that variations within three different genes, namely TRAF6, CD44 and TNFRSF11B, all involved in osteoclast biology, might increase the risk of osteoporosis and could therefore serve as targets for novel therapeutic agents.

Introduction
The identification of genes that are responsible for common human diseases such as diabetes, osteoporosis, heart disease, epilepsy, obesity and coeliac disease, among others, poses the greatest challenge to geneticists. The aetiology of these diseases is known to be complex and includes both genetic and environmental factors. Each of these factors alone has a modest effect, but then cumulative effects, especially when a certain threshold is exceeded, result in disease. Susceptibility to disease is increased by the interactions of variants within the same and/or different genes which are modified by various environmental factors, including diet and lifestyle habits. Other confounding factors, such as heterogeneity, phenocopies, genetic imprinting and penetrance, influence the effect of the genotype on the phenotype and further complicate the identification of susceptibility genes. These factors make it more difficult to reach conclusions and to replicate results using current approaches.

Three common approaches used for the identification of susceptibility genes are the case-control (studying candidate genes), the genomic linkage (also known as gene-mapping or family studies) and the genome-wide association study (GWAS) approach. For a case-control study, candidate genes are most often selected by knowledge of physiology or by prior evidence from association studies performed in other populations, thus assuming a priori that these could increase the risk of disease. Single nucleotide polymorphisms (SNPs), found within the selected genes, are studied in a random sample of affected individuals while a group of sex- and age-matched unaffected individuals is used as controls. Significant association is found if allele frequencies differ significantly between affected individuals while a group of sex- and age-matched unaffected individuals is used as controls. Significant association is found if allele frequencies differ significantly between affected individuals and normal controls, where an over-representation in the former group indicates a risk allele for the disease.

The second approach that has been used in the study of the Maltese population is the linkage approach or a family study. In genome-wide linkage analysis, no knowledge of candidate genes is required and the objective is to identify a novel gene responsible for the disease (and/or a disease-causing genetic variation) when located at a chromosomal region which is inherited with the phenotype. A number of polymorphic markers spaced across the genome are analysed within one or a group of
Genetic association studies

One of the most extensively studied genes is that coding for the vitamin D receptor (VDR), which is a nuclear receptor responsible for the initiation of transcription of target genes upon stimulation by the active form of vitamin D. The most common single nucleotide polymorphisms (SNPs) studied for an association with low bone mineral density include a start codon polymorphism found in exon 2 (Fok I), three SNPs identified by endonucleases BsmI, Apal and TaqI at the 3’ end of the gene and a functional binding site for the intestinal-specific transcriptional factor Cdx-2 in the 1a promoter region.11-14

Conflicting results were obtained from studies performed in different populations, where this lack of concordance can be attributed to a lack of sample power, population stratification, heterogeneity between different populations, effects of other confounding factors and technical problems, such as lack of harmonisation of methodologies used in different studies.

Calcium intake was observed to have an effect on the influence of various polymorphisms on BMD, including that of the B allele for the BsmI polymorphism of the VDR gene, which was correlated with a low BMD only in the presence of low calcium intake in the Dutch population.15 The same observation was made for the G allele (with low BMD) of the Cdx-2 polymorphism in the promoter region. No statistically significant association was observed between mean BMD and any of the studied SNPs within the VDR gene in the Maltese population16,17, but trends were similar to those reported in other studies.18 A significant odds ratio of 2.7 (95% CI: 1.7 - 4.5; p<0.001) was observed for the G allele (BsmI polymorphism), indicating that this allele could increase the risk of low BMD in the Maltese population.

Other candidate genes studied in the Maltese population were those encoding for the oestrogen receptor (ER)-1, collagen type 1A1, methylenetetrahydrofolate reductase (MTHFR) and lipoprotein receptor-related protein (LRP)-5. None of the polymorphisms studied in these genes were significantly associated with an increased risk of osteoporosis.

Significant differences were observed in the Maltese population in the distribution of genotype frequencies between women with low BMD and normal controls for the T950C polymorphism found in the promoter region of the TNFRSF11B gene, which codes for the decoy receptor osteoprotegerin (OPG). OPG is a secreted glycoprotein of the tumour necrosis factor (TNF) superfamily that controls osteoclastogenesis by binding to RANKL thus preventing its interaction with RANK on osteoclast progenitors and promoting apoptosis (Figure 1).19

Genotype frequencies for the T950C polymorphism were similar to those observed in the Danish, Slovenian and Swedish populations.20-22 In the Maltese population, the frequency of the TT genotype was significantly over-represented in individuals having a low BMD (83%) while the CC genotype was found in 59% of the normal controls.23 Also for another polymorphism in the first exon of the gene (G1181C), the GG genotype was found more frequently in the group of osteopenic/osteoporotic women, although this was not statistically significant. Similar observations were reported in the Danish population where the G allele was found to be significantly more common in osteoporotic patients.20 Significant odds ratios were also observed for the T950C and G1181C polymorphisms, when comparing genotypes in affected and normal postmenopausal women. The A-T-G haplotype was also significantly associated with an increased risk of a low BMD while the A-C-C haplotype was shown to have a protective role. It is not known yet whether this polymorphism in the promoter region of the TNFRSF11B gene is functional or whether it is in linkage disequilibrium with another as yet unknown functional variant found elsewhere within this gene.

Linkage studies

The other approach used in our genetic studies, which yielded valuable and new data, was the linkage approach. A family based linkage analysis was performed using 27 members from two Maltese families (15 females and 12 males) with
with a SNP within the LRP-4 gene, only showed weak evidence where a non-significant association to fractures was observed and a recent genome-wide analysis using over 300,000 SNPs, which included a meta-analysis of nine genome-wide scans. Subsequently other studies, 11p12 (OMIM: 611739; BMD8; http://www.ncbi.nlm.nih.gov/ the first time that strong suggestive linkage was reported to Densitometry.

The initial genome-wide scan was performed using a set of 400 polymorphic markers spread across the 22 autosomes and X-chromosome. The osteoporosis phenotype was defined according to WHO criteria using a t-score of < -2.5 as threshold for post-menopausal women and older men (≥50 years) while z-scores were used for the younger generations as suggested by the International Society of Clinical Densitometry. Following the initial analysis, fine-mapping was performed at the indicated chromosomal regions only, by increasing the number of tested markers to be able to identify any suspected genes.

Strong suggestive linkage was observed to chromosome 11p12 where haplotypes, inherited in an autosomal dominant fashion with a number of recombination events occurring close to this region were observed in each family. This was the first time that strong suggestive linkage was reported to 11p12 (OMIM: 611739; BMD8; http://www.ncbi.nlm.nih.gov/entrez/dispmim.cgi?id=611739). Subsequently other studies, which included a meta-analysis of nine genome-wide scans and a recent genome-wide analysis using over 300,000 SNPs, where a non-significant association to fractures was observed with a SNP within the LRP-4 gene, only showed weak evidence of linkage to this region. Strong suggestive linkage was also observed to chromosome 11p12 where haplotypes, inherited in an autosomal dominant fashion including rheumatoid arthritis, coeliac and inflammatory bowel disease. The same locus was also linked with coeliac disease in a Maltese family where linked variants were identified by sequencing genes at this locus. The above mentioned diseases were excluded from the two osteoporotic Maltese families used in this study. Such observations could support the hypothesis that a common inflammatory process controlled by genes in this region might be responsible for the onset of these diseases.

Twenty-two genes were found at this locus, with the best candidate being those coding for the TNF receptor associated factor (TRAF)-6 (MIM 602355) and CD44 (MIM 107269), with TRAF6 being closer to the indicated marker. TRAF6 plays a very important role in the differentiation and activation of osteoclasts by interacting with the cytoplasmic domain of RANK and activating transcription mediated through at least three signal transduction cascades (Figure 1). TRAF6 might also be a key molecule involved in the activation of osteoclasts following stimulation by a number of inflammatory cytokines, thus further supporting the hypothesis that osteoporosis results from a mild inflammatory process. RANK/RANKL and TRAF6 were reported to play a very important role in what is known as osteoimmunology, where both bone physiology and immunology interact and share various molecules.

The TRAF6 gene was sequenced and three polymorphisms were identified, two of which were found in introns, and an A to T variation found at position -721 upstream from the transcriptional start site.

This polymorphism was further tested for its possible functional effects on gene expression because it was found in 3 (33%) of affected family members from one family, but only in 1% of the general population. Reporter gene assays showed that gene expression was significantly increased in the presence of the T allele when compared to wild-type A allele, thus suggesting increased osteoclastogenesis and therefore increased bone resorption. Further experiments were performed to test whether other transcriptional factors that bind further upstream in the promoter region might affect expression in the absence or presence of this variation. GATA-1 was observed to be a potential enhancer of gene expression in wild-type alleles but not in the presence of the T allele.

Another candidate at 11p12 is the CD44 gene, which consists of a total of 18 exons and shows extensive heterogeneity when it is expressed due to alternative splicing. This molecule was found to have an important role in the fusion of macrophages to form multinucleated cells, such as osteoclasts, and was also found to activate signal transduction. The CD44 gene was sequenced and six known SNPs were found in both introns and exons, but none of these was linked to the inherited haplotype. The most significant finding was an already known G/A synonymous SNP (rs11033026) found 32 nucleotides upstream from the exon-intron junction in exon 9. This SNP was found linked to the STR allele in all affected family members (from one family), with the exception of one member. In the general Maltese population at birth, a frequency of 2.38% with a minor allele frequency of 0.012 was determined. This polymorphism was reportedly absent in European Caucasians but was found in Sub-Saharan Africans, African-Americans and Asians (minor allele frequencies 0.336, 0.115, 0.012, respectively) (http://www.ncbi.nlm.nih.gov/SNP/snp_ref.cgi?rs=11033026).

Since this polymorphism does not result in an amino acid change, it was hypothesised that it could affect pre-messenger RNA (pre-mRNA) splicing during transcription. Similar variations affecting splicing were reported to cause other diseases, such as cystic fibrosis and neurofibromatosis type 1. In vitro experiments performed to test this hypothesis showed that the variant affected RNA splicing in mouse macrophages, thus possibly increasing risk of osteoporosis in this family.
Macrophages are the pre-cursors of osteoclasts, and therefore a shift in the balance between different CD44 isoforms might affect their function and mechanisms, possibly cell fusion and migration. DNA variations found in non-coding regions of genes should not be underestimated. More supporting evidence is emerging to indicate that these variants can still affect the phenotype by affecting either RNA splicing or protein folding without changing the amino acid sequence.39

**General conclusions**

A number of polymorphisms were identified in three different genes, the TNFRSF11B, TRAF6 and CD44, all involved in osteoclast biology. Two of these polymorphisms found within the TRAF6 and CD44 genes were also tested functionally and were found to affect gene expression and pre-mRNA splicing, thus possibly increasing the risk of disease in affected individuals. In different ways, the protein products of these genes are all involved in osteoclast differentiation and activation. TRAF6 is a key player in the activation of various signal transduction cascades leading to increased osteoclast activation following stimulation through RANK. Increased expression of TRAF6 (or decreased degradation) results in increased osteoclastogenesis due to activation of different cascades. Conversely OPG (TNFRSF11B) is a negative regulator of osteoclastogenesis and thus decreased quantities or decreased binding to RANKL will lead to increased osteoclast activity. In a different way, CD44 isoforms can also affect osteoclast activity possibly by affecting fusion during their formation or even their migration. These findings show the importance of the RANK/RANKL/OPG system involved in osteoclast differentiation and activation. Both TRAF6 and OPG are targets for treatments aimed at controlling bone resorption, decreasing inflammatory bone loss and osteoporosis. Exogenous OPG has been used successfully in postmenopausal women as an antiresorptive agent showing its potential therapeutic use.39 TRAF6 inhibitors have already been suggested as having a potential therapeutic use to control inflammation not only in osteoporosis but also in other diseases such as periodontitis, osteolytic conditions, cystic fibrosis, viral infections and connective tissue destruction.40 Lately, biphenylcarboxylic acid was reported to inhibit TRAF recruitment to pro-inflammatory cytokine receptors and was found to prevent translocation of TRAF6 to the cell membrane.41 The role of CD44 in osteoporosis needs to be further studied. It can, however, be a potential therapeutic target with different isoforms being used for diagnostics purposes, possibly also with a prognostic value for bone tumours.42

Genetic studies together with functional studies of potential candidates will help to identify new metabolic pathways that might be involved in the pathogenesis of disease, thus assisting in the development of new and more effective treatments. Extending these genetic studies to patients, who have suffered fractures, irrespective of their BMD status, might help elucidate further the genetic basis of bone disease in the Maltese population.