The role of genetic variation in osteoporosis

Seyed Alireza Dastgheib

Thesis submitted in accordance with the requirements of The University of Sheffield for the degree of Doctor of Philosophy.

Department of Human Metabolism
Faculty of Medicine, Dentistry & Health

May 2015

Acknowledgements

Firstly, I would like to extend my deepest gratitude to my supervisors Dr. Dawn Teare and Dr. Alison Gartland for giving me an opportunity to do this project and for their excellent teaching and guidance throughout. I regard them as outstandingly good scientific supervisors. I am grateful to Dr. Mohammad Tabei as my remote supervisor for providing facilities for sample gathering and osteoclast culture in Iran. I would like to acknowledge Dr Santibanez-Koref for adapting his specific software to analyse data of AEI study. I also want to thank my colleagues in the Gartland bone group – Karan, Robin, Iain, Ankita, and Ning for helping me through the lab work session and making this a very enjoyable and memorable experience. Fondly remember my family members, my parents for their continuous belief and support and my wife Maryam for giving me constant support and encouragement.

Summary

There is accumulating evidence that genetic variations play an important role in osteoporosis by reducing bone mass and increasing the risk of fragility fractures. Many different genetic variants with mild effects contribute to the regulation of bone mineral density and fracture risk. Investigating the effect of genetic variants in osteoporosis may elucidate new mechanisms of bone homeostasis and possibly create new effective pharmaceutical targets for osteoporosis in the future.

In the present study, data from genome-wide association studies, candidate gene association studies, and meta-analysis studies have been used to select stronger candidate genetic variants for further study. Iranian population samples were utilized in this project to investigate the effect of these selected genetic variants. Functional studies examined the association with genetic variants and the function of osteoclast cells and in the allelic expression imbalance (AEI) in target cells.

The analysis of AEI was used to investigate the effect of cis-acting factors on the transcribed SNPs in the target cells via evaluation of cDNA from tissues expressing those transcribed SNPs in the heterozygous individual. The preliminary analysis of the Iranian population data showed features similar to other populations. There were also statistically significant associations between BMD and risk SNPs: VDR-rs731236, P2RX7-rs3751143, RANK-rs884205, RANKL-rs9594738, RANKL-rs9594759, LRP5-rs2306862, OPG-rs2073618, OPG-rs11995824, OPG-rs4355801, and ESR1-rs2077647. There were statistically significant associations between osteoclast number and risk allele of RANKL rs9594759 /rs1021188 SNPs. These associations are in the direction that would be expected from risk alleles of BMD. AIE analysis suggested evidence of cis acting effects in both osteoclast culture and blood samples for two risk SNPs, ESR1-rs2504063 and VDR-rs731236 which were each associated with BMD in the Iranian population. This study and other same studies helps to improve our knowledge about the role of genetic variation in osteoporosis with the goal of providing a genetic profile can be used for genetic screening tests for the identification of individual at risk of osteoporosis. A suitable genetic profiling test helps to find useful treatment. Novel pathways that contribute to control of BMD could be identified, improving our knowledge about the effect of genetic variation with the possibility of producing new therapeutic agents.

Table of contents

List of figures

List of Table

List of Abbreviations

α -MEM	α –Modification of minimum essential medium
APC	Adenomatosis polyposis coli
ASE	Allelic specific expression
AEI	Allelic Expression Imbalance
ATP	Adenosine triphosphate
BMD	Bone mineral density
BMI	Body mass index
BMU	Basic molecular unit
CATK	Cathepsin K
cDNA	Complementary DNA
CGAS	Candidate gene association study
COL1A1	Collagen, type I, alpha 1
COL1A2	Collagen, type I, alpha 2
CRTAP	Cartilage associated protein
CT	Cycle threshold
CYP17	Cytochrome P450c17alpha gene
DALYs	Disability-adjusted life years
DEXA	Dual-energy x-ray absorptiometry
DKK-1	Dickkopf-related protein 1
DSH	Dishevelled
EDTA	Ethylenediaminetetraacetic acid
ESR1	Estrogen receptor 1
FCS	Fetal calf serum
FRET	Fluorescent Resonance Energy Transfer
Fz	Frizzled

GWAS	Genome-wide association study
HBM	High bone mass
IL1	Interlukin-1
IL6	Interlukin-6
IMOS	Iranian Multi-centre Osteoporosis Study
KASPar	KBioscience Competitive Allele-Specific Polymerase chain reaction
LCLs	Lymphoblastoid cell lines
LD	Linkage disequilibrium
LEPRE1	Leucine proline-enriched proteoglycan 1
LPA	Lysophosphatidic acid
LS-BMD	Lumbar spine BMD
LRP5	Lipoprotein receptor-related protein 5
MA-C	Meta-analysis of candidate gene association study
MAF	Minor allele frequency
MA-G	Meta-analysis of genome-wide association study
M-CFS	Macrophage colony stimulating factor
Min	Minute
μ	Population mean
NHANES	Third national health and nutrition examination survey
NHGRI	National human genome research institute
OC	Osteoclast cell
Oligo (dt)	A short sequence of deoxy-thymine nucleotides
OPG	Osteoprotegerin
OPPS	Osteoporosis-pseudoglioma syndrome
OSTM1	Osteopetrosis-associated transmembrane protein 1
P2RX7	Purinergic receptor P2X, ligand-gated ion channel, 7
PBMCs	Peripheral blood mononuclear cells

PBS	Phosphate-buffered saline
PCR	Polymerase Chain Reaction
PIC	Polymorphism information content
PLD	Phospholipase D
PPi	Inorganic pyrophosphate
PTH	Parathyroid hormone
qRT-PCR	Quantitative reverse transcription polymerase chain reaction
RANK	Receptor activator of nuclear factor κB
RANKL	Receptor activator of nuclear factor κB ligand
RFLP	Restriction fragment length polymorphism
rRNA	Ribosomal RNA
RT	Reverse transcription
Runx2	Runt-related transcription factor 2
SD	Standard deviation
SDS	Sodium dodecyl sulfate
SEM	Standard error of the mean
SFRP	Secreted frizzled-related protein
SNP	Single nucleotide polymorphism
TCF/LEF	T-cell factor/lymphoid enhancing factor
TCIRG1	T-cell immune regulator 1
TNF	Tumour necrosis factor
Total hip BMD	Total hip BMD
TRAP	Tatrate resistant acid phosphatase
tRNA	Transfer RNA
TROCH	Trochanter
UTR	Un-translated region
VDR	Vitamin D receptor

Chapter 1 – Introduction

1.1 Bone

1.1.1 Bone structure and function

Bone is a dynamic and living tissue with the structural and materialistic properties to enable limb mobility and protection of internal organs. Bone functions as a reservoir for minerals, and modulates calcium and phosphorus homeostasis. As an endocrine organ, the cells of the bone respond to hormonal signals and secrete factors that regulate various cellular functions including pancreatic beta cell insulin production and kidney phosphorus excretion (Guntur and Rosen 2012). Bone also has an important role in blood production via the bone marrow (Zappitelli and Aubin 2014).

Bone is composed of cortical and trabecular bone, cartilage, hematopoietic tissue and connective tissue. The cortical bone is the hard compact outer layer of bones and accounts for approximately 80% of the total mass of bone in an adult skeleton. Trabecular bone accounts for the remaining 20% of bone mass and is a spongy or porous form of tissue filled with hematopoietic marrow, fatty marrow and blood vessels. The bones in vertebrates are either flat bones (hip, scapula and skull) or long bones (tibia, humerus and femur). Flat bones are made of trabecular tissue surrounded by layers of cortical bone. Long bones consist of the shaft or diaphysis, metaphysis and epiphysis regions. The diaphysis is composed mostly of cortical bone, while the metaphysis and epiphysis contain trabecular bone with a protective shell of cortical bone. The denser cortical bone is important for loading, whereas the trabecular bone has a larger surface area per unit volume for facilitating metabolic activity (Favus 2008).

Bone is mainly composed of the bone matrix, which consists of organic and inorganic matter. The organic matter is 90-95% type I collagen, providing tensile ability and a non-collagenous portion consisting of serum proteins. The inorganic component of the bone matrix is made up of calcium and phosphate in the form of hydroxyapatite, which fills the “holes” between the collagen fibrils providing strength under compression (Favus 2008). The bone marrow consists of a hematopoietic component known as the parenchyma and a vascular component known as the stroma. Within the parenchyma, hematopoietic stem cells and hematopoietic progenitor cells exist close to the endosteum of the bone and around blood vessels (Zhao, Xu et al. 2012). Bone tissue consists of osteoclasts, osteoblasts, osteocytes and bone lining cells.

These cells have fundamental roles in new bone tissue development, bone maintenance and mineral regulation (Morrison and Scadden 2014) .

1.1.2 Cells of the bone

Bone cells are responsible for the development of new bones, the maintenance of bones, and the regulation of minerals in the body. There are four main categories of bone cells which include: osteoclast, osteoblast, osteocyte, and lining cells (Zappitelli and Aubin 2014).

1.1.2.1 Osteoclasts

Osteoclasts are multinucleated cells that are responsible for the removal or resorption of bone. They are found in resorption pits also known as Howship's lacunae, within the bone surface. Osteoclasts are formed as a result of the fusion and subsequent differentiation of monocyte cells. The cytoplasm of osteoclast cells is homogeneous, "foamy" appearance, due to the large number of vesicles and vacuoles, which contain acid phosphatase. Osteoclasts are characterised by staining for high expression of tatrate resistant acid phosphatase (TRAP) and cathepsin K (Delaisse 2014). Monocytes differentiate to osteoclast and migrate to the resorption site, where the initial event in bone resorption occurs. The attachment between osteoclast and bone surface forms a special site, known as the "sealing" zone. In this area plasma membrane of osteoclast attaches firmly to the bone matrix and an isolated extracellular microenvironment is generated between the cell and bone matrix (Teitelbaum 2000). At the bone resorption site, osteoclast cells form specialised ruffled border membranes using a high number of intracellular acidic vesicles in the plasma membrane of the osteoclasts. The extensively folded borders of the ruffled membrane are morphologically characteristic of actively resorbing osteoclasts and facilitate bone removal by increasing the cell surface for secretion. Vesicles release a large quantity of acid, mostly crystalline hydroxyapatite, which functions to dissolve the bone mineral (Vaananen and Laitala-Leinonen 2008). Degradation of organic part and then removal of degradation products occur from the resorption site. At the end osteoclast returns to the non-resorbing stage or apoptosis occurs: these are the final events in the bone resorption by osteoclasts (Vaananen, Zhao et al. 2000).

The two important factors which drive osteoclastogenesis are macrophage colony stimulating factor (M-CSF) and receptor activator of nuclear factor κB ligand (RANKL), a member of tumour necrosis factor (TNF) family (Boyle, Simonet et al. 2003). In 1997, Simonet et al. showed that over-expression or administration of a protein, osteoprotegerin (OPG) inhibits the osteoclast formation. Today it is known that OPG competes with RANKL for Receptor Activator of Nuclear Factor κB (RANK) (Vaananen, Zhao et al. 2000).

One way to study the function of genes in osteoporosis is to focus on osteoclast activity. Osteoclasts can be generated *in vitro* from the blood of an individual identified as having specific alleles in the gene of interest. Osteoclast function can be measured by culturing peripheral blood mononuclear cells (PBMCs) on the dentine wafers (ivory wafers), and then by investigation of resorption lacunas, which are produced by differentiated osteoclasts (Chan, Gartland et al. 2007). The number of osteoclasts that are generated *in vitro* and the rates of resorption and apoptosis can be determined as osteoclast activity (Gartland, Buckley et al. 2003; Chan, Gartland et al. 2007).

1.1.2.2 Osteoblasts

Osteoblasts are mononuclear cells responsible for bone formation. Osteoblasts differentiate from mesenchymal precursor cells, which have the same origin as chondrocytes, tendon cells, myoblasts and adipocytes. Osteoblast cells have a large nucleus, extensive endoplasmic reticulum and enlarged Golgi. These cells have the ability to secrete type I collagen and specific proteins such as albumin, which form the organic matter of the bone matrix and regulate the bone matrix mineralisation (Harada and Rodan 2003). During the embryonic stage osteoblast differentiation occurs in two pathways: endochondral ossification and intramembranous ossification. During early embryogenesis, the collar bone surrounds the cartilage template. After invasion of vascular tissue, chondrocytes are replaced by osteoblasts and this process is called endochondral ossification. Endochondral ossification regulates the formation of all the skeletal bones with the exception of the skull, the mandibles and the clavicles. On the other hand, mesenchymal progenitor cells differentiate directly into osteoblasts, a process known as intramembranous ossification (Ducy, Schinke et al. 2000).

1.1.2.3 Osteocytes

Osteocytes are mononuclear cells comprising a thin ring of cytoplasm and numerous cellular extensions. They support bone structure and metabolic functions. Osteocytes are the terminal differentiation stage of osteoblasts. When osteoblasts are trapped in the matrix they are differentiated to osteocytes. Osteocyte cells are localized in lacunae and their cellular extensions occupy tiny canals called canaliculi. Osteocytes in the mineralized matrix have direct communication with each other and bone surface lining cells. They can live for long periods but in ageing bone the intra-cellular connections between the osteocytes are disrupted, leading to apoptosis (Favus 2008).

1.1.2.4 Bone lining cells

Bone lining cells are a subpopulation of the osteoblast family and are characterised by their long, flattened and spindly appearance. They are associated with the bone surface at sites where a thin non-mineralized collagen layer is existed. Although they do not produce an osteoid layer like an osteoblast cell does, bone lining cells still belong to the same lineage as osteoblasts as they are associated with the surface of bone, express alkaline phosphatase and are responsive to PTH. Bone lining cells regulate bone remodelling by preventing interaction of osteoclast precursors with the bone surface. Signals that initiate osteoclast formation may also stimulate the bone lining cells to prepare for bone resorption mediated through collagenase activity, which breaks down un-mineralised bone, revealing the mineralised matrix underneath. The bone lining cells are then able to migrate to cover the re-modelling area, mainly at sites next to osteoclasts, creating a microenvironment (in phagocytosis of collagen at the bone surface) for the link required during bone remodeling. It has been suggested that the bone lining cells are responsible for the cell to cell interactions between RANKL and RANK receptor on osteoclast precursors (Everts, Delaissé et al. 2002).

1.1.3 Bone remodelling

Bone remodelling is an important and lifelong metabolic process where bone is renewed continuously in response to a variety of stimuli. The most important of these stimuli are functional demands, which depend on density, strength, and shape of the bone in response to load alterations. The important functions of bone remodelling are maintaining anatomical and structural integrity of the skeleton and regulating serum calcium by supplying calcium from the skeleton to the serum. Bone remodelling is a complex process involving interactions between different cell phenotypes and is regulated by a variety of biochemical factors (Delaisse 2014).

Bone remodelling can be described as a 4-stage cyclic process involving osteoclasts, osteoblasts, osteocytes and bone lining cells: resting, resorption, reversal and formation (**Figure 1.1**). In the resorption phase, differentiated osteoclasts attach to the bone and digest it by acidification and proteolytic mechanisms. This phase is followed by the reversal stage. During reversal, osteoblasts are recruited to the site of resorption to start bone formation. Following this stage osteoblasts spread over the resorption areas and bone formation begins with the secretion of osteoid, which comprises matrix of collagen and other proteins. Finally, bone-lining cells cover the surface of the bone and osteocytes are entombed in the matrix (Delaisse 2014)

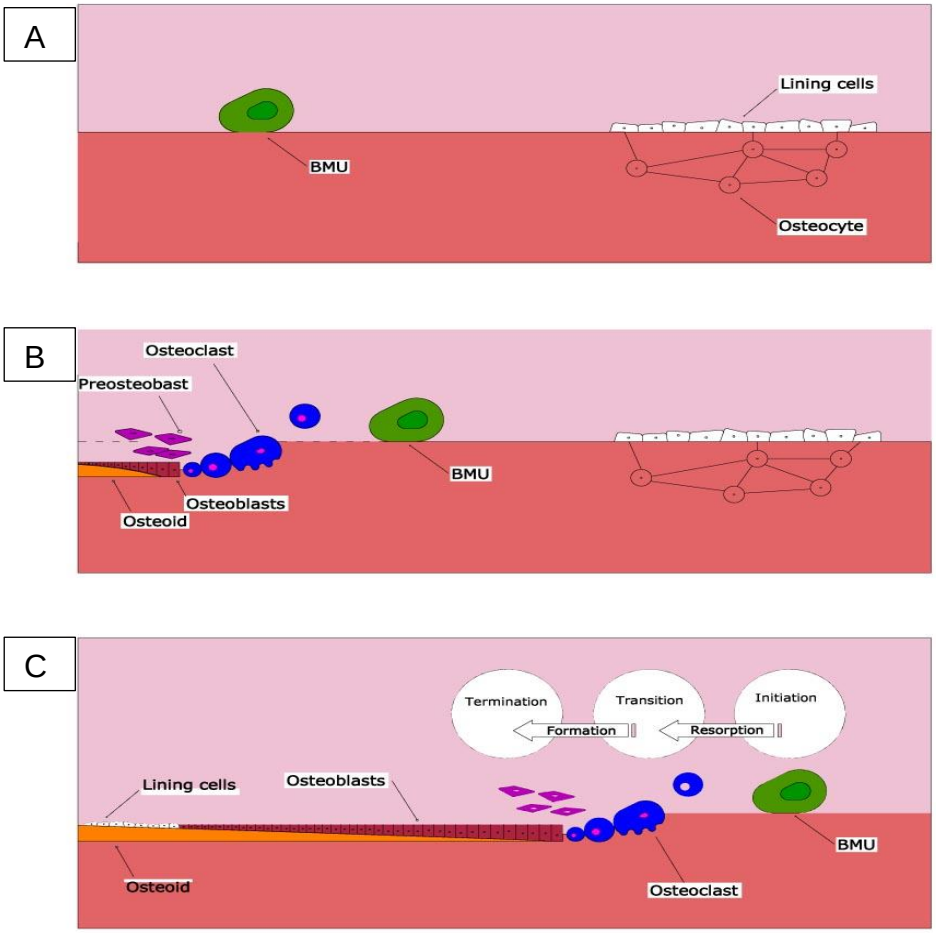


Figure 1-1: bone remodeling

A*: The basic molecular unit (BMU) moves along the trabecular surface.* ***B****: Complex BMU-lining cells attach themselves to exposed collagen and attract pre-osteoclast cells. Multinucleated cells differentiate themselves from pre-osteoclast cells and resorb a cavity.* ***C****: Pre-osteoblasts are changed to osteoblasts. Osteoblasts produce osteoid. Previous osteoid starts to mineralize (horizontal lines). Osteoblasts begin to flatten. Osteoblasts change to lining cells covering the surface of the bone; bone remodelling is now complete, but BMU is still advancing (to the right).*

Bone remodelling is tightly regulated by a number of factors including local cytokines such as RANKL, interlukin-1 (IL-1), interlukin-6 (IL-6), and systemic hormones such as calcitonin, parathyroid hormone (PTH), and 1,25–dihydroxy vitamin D3 (Vitamin D). RANKL is expressed on the cell surface of osteoblasts and attaches to the RANK. RANK is responsible for the differentiation and activation of osteoclasts from osteoclast precursors. As the decoy receptor to RANK, OPG produced by osteoblast and stromal cells, inhibits interaction of RANKL. Increased levels of RANK have an anabolic effect on bone (**Figure 1.2**) (Sambrook and Cooper 2006). Twenty five percent of trabecular bone and only three percent of cortical bone are replaced by this mechanism within the skeleton each year. Trabecular bones have more surface contact with bone marrow in comparison with cortical bones and this difference shows the role of local factors (Manolagas 1995)

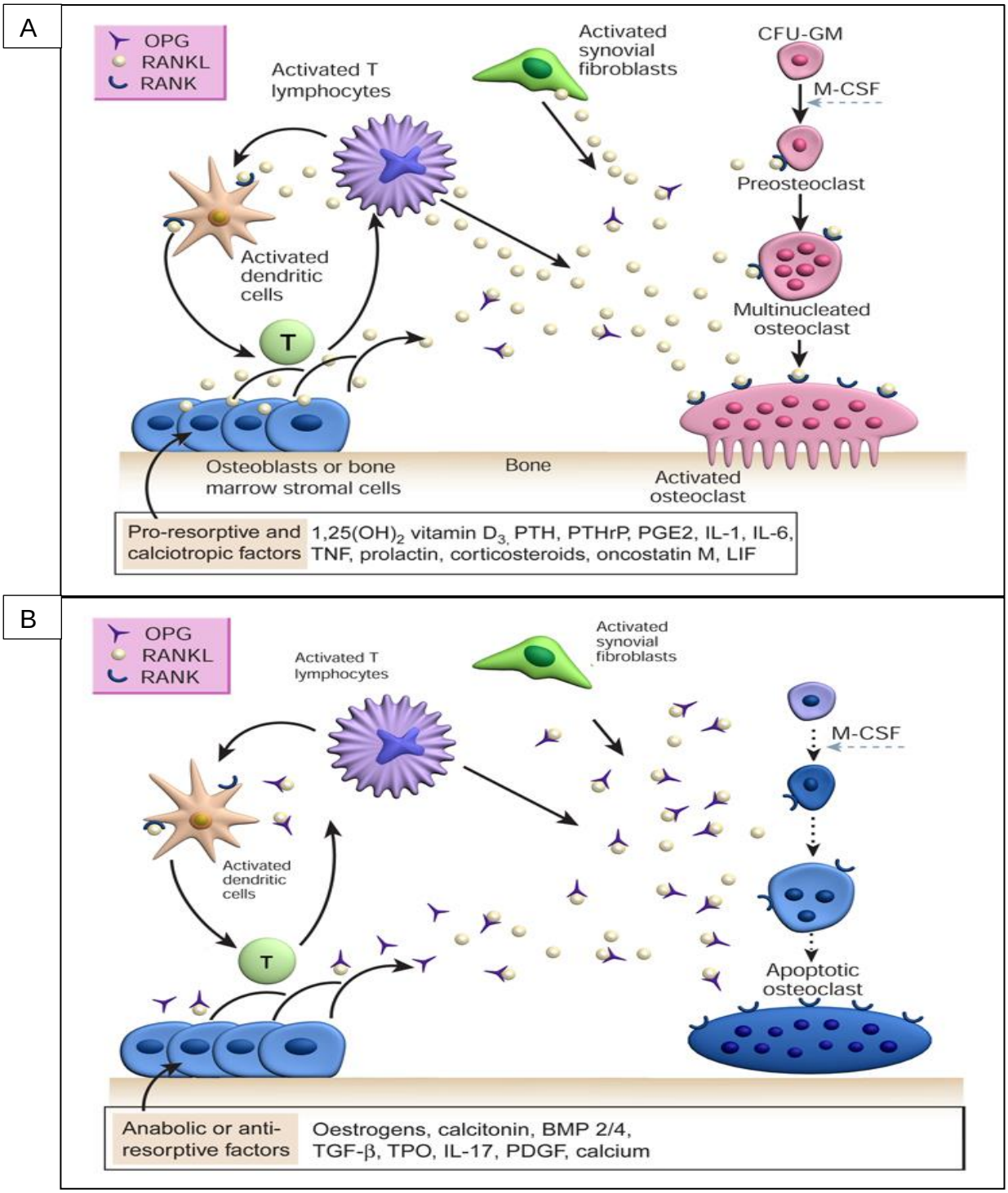


Figure 1-2: Hormonal control of bone resorption

***A**: Pro-resorptive and calcitropic factors. Expression of RANKL is induced in osteoblasts, activated T cells, synovial fibroblasts, and bone marrow stromal cells. It binds to membrane-bound receptor RANK to induce osteoclast differentiation, activation, and survival.* ***B**: Anabolic and anti-osteoclastic factors. OPG has anabolic effects. OPG binds to and neutralizes RANKL then induces block in osteoclastogenesis and decreases survival of pre-existing osteoclasts.*

1.1.3.1 Bone remodelling in disease

As the process of bone remodelling maintains skeletal integrity throughout life, imbalances in the underlying mechanisms can result in bone disorders. Around the fourth or fifth decade of life the processes of resorption and formation are not matched, and resorption activity is increased compared to formation in both women and men (Fauci 2008). There is increasing evidence to suggest that the cells involved in the remodelling of the skeleton under physiological conditions are the same cells involved in pathologic conditions. Abnormalities in the bone remodelling process can result in various skeletal disorders reflecting changes in the organic matrix, mineral phase and cellular mechanisms of remodelling. These skeletal disorders are either hereditary, acquired or both. The most common bone metabolism disorders affecting the skeleton are osteoporosis, osteomalacia, hyperparathyroidism and chronic renal failure and Paget's disease (Feng and McDonald 2011).

1.2 Osteoporosis bone disease

Osteoporosis is a skeletal disorder characterised by a reduction in the strength of bone with a consequent increase in bone fragility and susceptibility to fracture. Low bone mass is associated with micro-architectural deterioration of bone tissue (Fauci 2008). The World Health Organization (WHO) operationally defined osteoporosis based on measurement of bone mineral density (BMD) which is associated with bone fracture. This definition of osteoporosis is a bone mineral density (BMD) that is more than 2.5 SD below the young adult mean for the population (Kanis 1994). However, this definition only determines a gradual decline in mineralization and does not give a particular aspect of decline in micro-architecture (Cole, Dennison et al. 2008).

The loss of strength and bone mass in osteoporotic bone disease can be as a result of a multitude of factors including; the inability to reach optimal peak bone mass in early adulthood, increased resorption of bone having reached peak mass, or dysfunction of bone formation during bone remodelling. Osteoporosis is classified as primary or secondary, with the primary osteoporosis being the most common metabolic disorder of the skeleton. Primary osteoporosis is further categorised as type I and type II based on the etiology differences. Type I osteoporosis is also known as postmenopausal osteoporosis, and is seen commonly in postmenopausal women. Type I osteoporosis is thought to be caused by estrogen hormone deficiency associated with menopause. The type II osteoporosis also known as age-related osteoporosis is linked to ageing in males and female individuals. The less frequent secondary osteoporosis refers to disorders of the bone secondary to other medical conditions, alterations in the level and type of physical activity, or caused by therapeutic interventions (Sattui and Saag 2014).

It is expected that by 2050, the world population will increase, as a result it is also thought that the population of elderly people will also grow (Kassim Javaid, Chana et al. 2013). In the UK, by the year 2020, the over 85 years old age group is estimated to rise to 2.1 million, whereas this population was 1.2 million in 2002 (Clark 2002). This increase in the ageing population in the UK is predicted to cause double the number of osteoporotic fractures over the next 50 years (Kanis and Johnell 2005). The incidence of osteoporosis varies between male and female individuals. In males, the prevalence of osteoporosis based on total hip

BMD, ranged from 1% (in the UK) to 4% (in Japan). The prevalence, when based on spine BMD ranged from 3% (in Canada) to 8% (in France, Italy, Spain and Germany). In the female population, the prevalence of osteoporosis ranged from 9% (in the UK) to 15% (in France and Germany) calculated from total hip BMD data. Prevalence calculated from spine BMD data ranged from 16% (USA) to 38% (Japan)(Wade, Strader et al. 2014) .

The most prominent outcome of osteoporosis is the associated increase in hip fractures. In fact, it is estimated that by the year 2050, hip fracture incidences will increase from the 1.66 million in 1990 to 6.26 million (Dhanwal, Dennison et al. 2011). In the UK, one in two women and one in five men that are aged over 50 years will have suffered from an osteoporotic fracture in their remaining lifetime. In the United States approximately 10 million Americans have osteoporosis and an estimated 1.5 million fragility fractures occur each year (Kassim Javaid, Chana et al. 2013). In one study which was performed in the city of Shiraz in Iran, the incidence rate on hip fracture was higher for both genders compared with other Asian countries and this was similar to the rates reported in some European or American areas (Soveid, Serati et al. 2005).

There are three important adverse effects after osteoporotic fracture: mortality, disability, and cost that are correlated with type of fracture. Hip fracture has the most adverse outcome: the risk of death is 10-20% more than normal population in the first year after fracture but decreases over time (Sambrook and Cooper 2006). Hip fractures are the most important complication associated with osteoporosis and are linked to morbidity and sometimes mortality. Patients with hip fracture are prone to pressure sores and bronchopneumonia after fracture. Another fracture that occurs in osteoporotic patients is vertebral fracture. Vertebral deformities have been seen three times as frequently as hip fracture in studies using radiographic screening of populations. Kyphosis can be induced by multiple fractures, and in the acute stages of fracture severe back pain is the major symptom in the patients (Cole, Dennison et al. 2008). Adults with any type of fracture are at substantially greater risk of other types of fracture, particularly in the first year. The risk of further vertebral fracture increases ten-fold after a fracture in vertebrae. Furthermore, after a distal forearm fracture risk of hip fracture is increased 1.4 times in women and 2.7 times in men (Sambrook and Cooper 2006).

In the United States over 2 million osteoporotic fractures occurred in 2005 and hip fractures alone were responsible for 72% of the associated 20 billion dollar cost. In the United States,

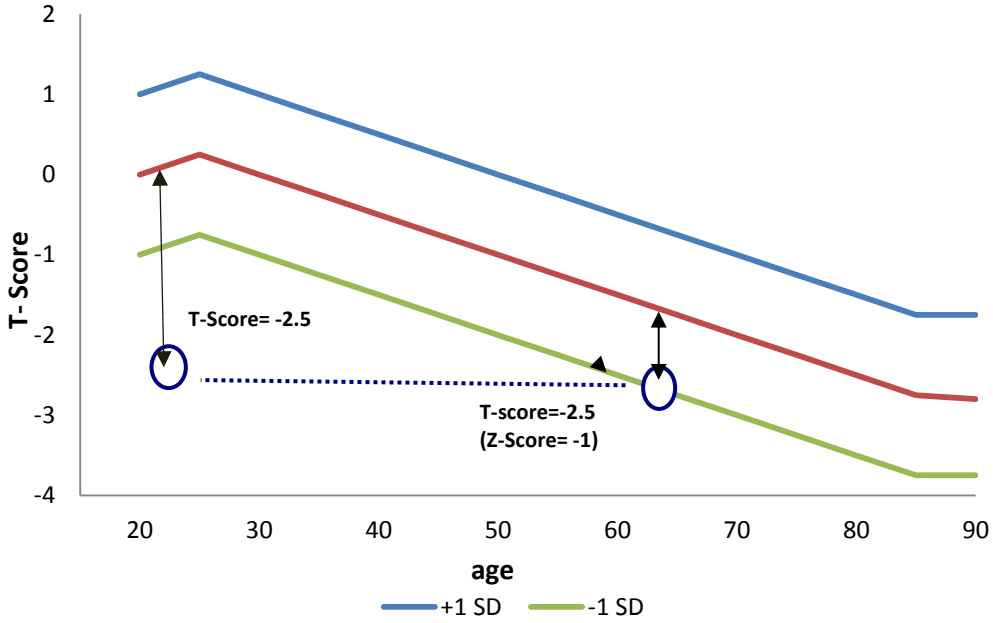
annual fractures and costs will rise by approximately 50% by the year 2025 in comparison with 2005. European countries are also estimated to spend an estimated 30 billion euros for all osteoporotic fractures every year (Sambrook and Cooper 2006). In comparison to other chronic non-communicable diseases, osteoporotic fractures induce more disability-adjusted life years (DALYs) in Europe (Cole, Dennison et al. 2008).

1.2.1 Clinical diagnosis

Bone mineral density is a quantity of bone mass per volume or area (amount of matter per cubic centimetre of bone) (Kanis, McCloskey et al. 2008). There are several techniques for measuring BMD but the gold standard method for diagnosis of osteoporosis is dual-energy x-ray absorptiometry (DXA) (Koh 2004). In this technique, two x-ray beams with different energy levels are used to determine the mineral content of bone tissue. Alternate techniques including quantitative ultrasound and computed tomography for assessing the condition of the skeleton are not as validated as DXA. Results produced by the DXA are compared to normal values by T and Z scores. The T-score is defined as the number of standard deviations above or below the mean BMD of a healthy 20 years old adult with the same patient's race and sex. To avoid an overestimation of bone mineral deficits, the results are also compared to the mean of a healthy individual with race, sex and age matched. The Z-score is defined as the number of standard deviations above or below the mean BMD of a healthy individual with the same patient's race, sex and age. For instance, a 65-year-old woman with a T-score of -2.5 (2.5 SD below mean for a young control group) has a Z- score of -1(1 SD below mean for age) **(Figure 1.3-A)** (Fauci 2008). DXA has the ability to measure bone mineral density at different sites, but usually is used at a specific site that has more validation for predicting risk of fracture. For example DXA of hip has the most validation for determining risk of fracture at the hip or even at other sites (Koh 2004). The reference measurement site and range which is recommended for a young population group is the femoral neck measurements in Caucasian women aged 20–29 years from NHANES III reference database (Third National Health and Nutrition Examination Survey).

A systematic review and meta-analysis of observational population-based studies using “absorptiometry” suggest that fracture risks increase approximately two-fold for every

standard deviation decrease in BMD (Kanis, Burlet et al. 2008). The predictive value of BMD for hip fracture is at least as good as that of blood pressure for stroke (Kanis 2002). The data from observational studies suggest that the absolute risk of fracture for any given BMD and age is similar in both genders. Thus, the same diagnostic cut-off values for BMD are now applied to in women and men (Kanis, Bianchi et al. 2011). There are four categories which describe diagnostic criteria for osteoporosis; normal, osteopenia, osteoporosis and severe osteoporosis (**figure 1.3-B**) (Kanis, McCloskey et al. 2008).



Diagnostic criteria for osteoporosis	
Normal Bone density is within 1 SD of the young adult reference mean (T-score=> -1 SD)	**Osteopenia** Bone density is within 1 to 2.5 SD below the adult reference mean (-2.5<T-score< -1)
Osteoporosis Bone density is 2.5 SD or more below the adult reference mean (T-score=<-2.5)	**Severe (established) osteoporosis** Bone density is more than 2.5 SD below the adult reference mean and there have been one or more osteoporotic fractures.

Figure 1-3: Criteria for osteoporosis

***A**: Relationship between Z-scores and T-scores in a 65-year-old woman. A value of -1 SD of Z-score in 65 years old is equal to -2.5 SD of T-score in 20 years old and less than this is defined as osteoporosis. **B**: Diagnostic criteria for osteoporosis. (Graph adapted from Fausi et al, Harrison's Principal of Internal Medicine, 17th edition).*

1.2.2 Molecular mechanisms involved in osteoporosis

The oestrogen-endocrine pathway is associated with the onset of osteoporosis in both women and men. Oestrogen is an essential component required for epiphyseal closure in puberty in women and men, and regulates bone turnover in both genders (Raisz 2005). Oestrogen is associated with increased osteoclast bone resorption activity and increased osteoclast numbers (**Figure 1.4**) (Manolagas 2000). Oestrogen deficiency is specifically a key player in the loss of bone in women, particularly those aged 70 years and older. The role of oestrogen in the regulation of bone metabolism is highlighted by findings that showed treatment rapidly reduced bone breakdown in older women (Raisz 2005). Oestrogen deficiency results in the release of cytokines by circulating monocytes, namely cytokines interlukin 1 (IL-1) and TNFα. These cytokines target mesenchymal cells, promoting the release of the lymphokine IL-7. IL-7 supresses both osteoblasts and T cells function via stimulation of pro-osteoclastogenic cytokines macrophage colony stimulating factor (M-CSF) and activated nuclear factor κ B ligand (RANKL) production. Interlukin 7 (IL-7) also acts to decrease osteoprotegerin (OPG) expression, the decoy receptor for RANKL. As a consequence of the latter changes, net osteoclastic activity is increased. Monocyte-derived IL-1 and tumour necrosis factor α (TNFα) synergize with RANKL to increase the loss of bone (Ross 2003).

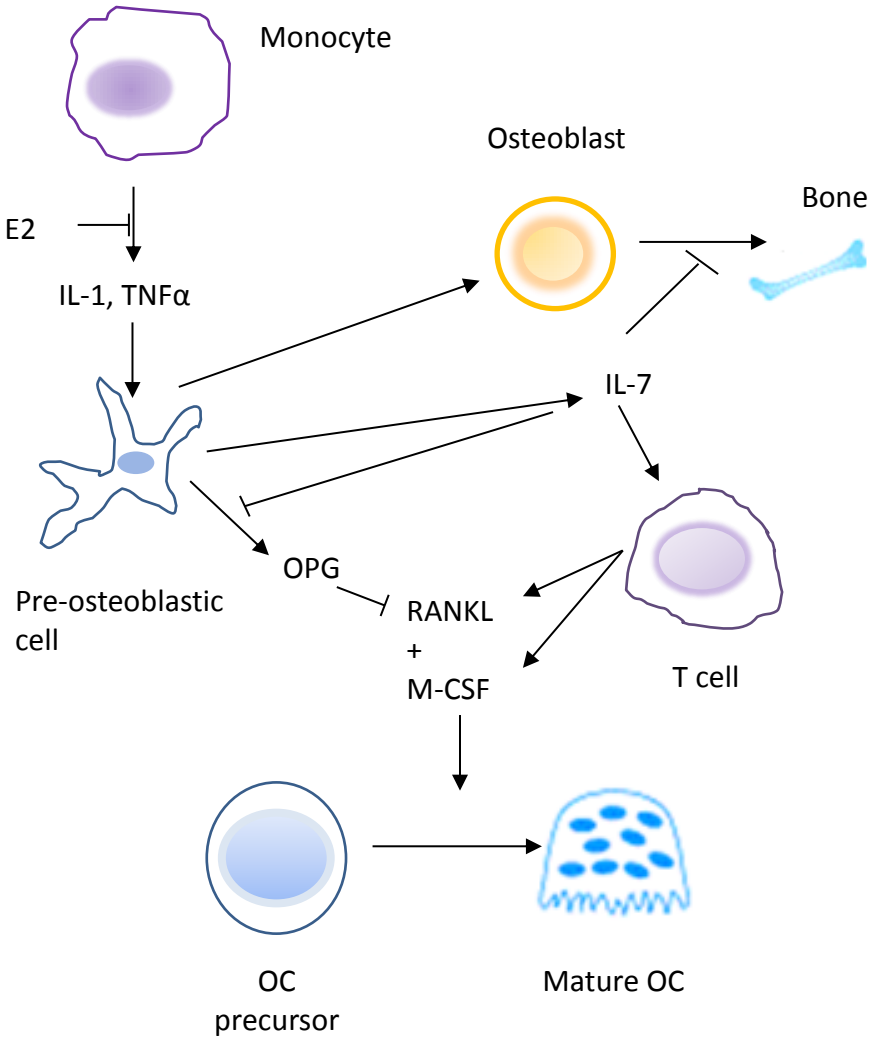


Figure 1-4: Estrogen-mediated bone loss in vivo

Oestrogen deficiency results in the release of IL-1 and TNFα by circulating monocytes. These cytokines stimulate mesenchymal cells to release IL-7, which induce bone loss by two pathways, suppression of bone formation and activation of bone resorption.

Understanding of the RANKL/RANK/OPG pathway has contributed to determining the mechanisms underlying the regulation of bone formation and resorption. RANKL and OPG are members of the TNF family and TNF receptor (TNFr) superfamilies, respectively, and binding to RANK regulates osteoclast formation, activation and survival in both normal bone modelling and remodelling processes. The RANK/RANKL/OPG axis is now also thought to be pivotal in several other pathologic conditions characterised by increased bone turnover (Ross 2003).

The canonical Wnt signalling pathway is also involved in the regulation of osteoblast differentiation and osteoclast suppression (**Figure 1.5**) (Canalis 2013). The Wnt proteins are a family of secreted glycoproteins that are involved in a multitude of cellular events including cell growth, differentiation, function, and apoptosis. Wnt signalling in an osteoblast is activated by binding of the Wnt proteins to their co-receptors; low-density lipoprotein receptor–related protein 5 (LRP5) and frizzled (Fz). Upon activation of Wnt signalling, a 'destructive' protein complex comprising: Adenomatosis polyposis coli (APC), Casein kinase 1α (CK1α), and Axin molecules become blocked downstream of the Wnt-receptor complex. The inhibition of the 'destructive protein complex' is regulated by the binding and subsequent phosphorylation of the dishevelled (dsh) protein to the Wnt-receptor complex. An important mediator of canonical Wnt signalling, B-catenin, is stabilised and accumulates in the cytoplasm of the cell, where it is phosphorylated and then translocates into the nucleus. Nuclear B-catenin binds to transcription factors belong to the T-cell factor/lymphoid enhancing factor (TCF/LEF) to induce the transcription of osteoblast-specific genes including Runt-related transcription factor 2 (Runx2), collagen, osteocalcin and osteonectin. In the absence of a Wnt protein, where Wnt signalling is not activated, dsh binds to the Axin component of the 'destructive protein complex' and this complex is then free to bind to B-catenin. B-catenin is degraded in the proteasome, resulting in low nuclear B-catenin levels and subsequently gene transcription is supressed. Antagonists of the Wnt pathway, such as Dickkopf-related protein, Secreted frizzled-related protein (SFRP) and sclerostin block Wnt signalling by interacting with Wnt proteins or transmembrane LRP/Frz receptors (Canalis 2013).

Our understanding of the molecular links between Wnt signalling and bone development have improved greatly since the first findings linking LRP5 mutations to alterations in human bone mass. It is now clear that Wnt/β-catenin signalling is responsible for increasing bone mass via numerous mechanisms involving osteoblastogenesis and suppression of osteoblast

and osteocyte apoptosis. Wnt/β-catenin is also known to increase the ratio of OPG to RANKL, repressing osteoclastogenesis (Krishnan, Bryant et al. 2006).

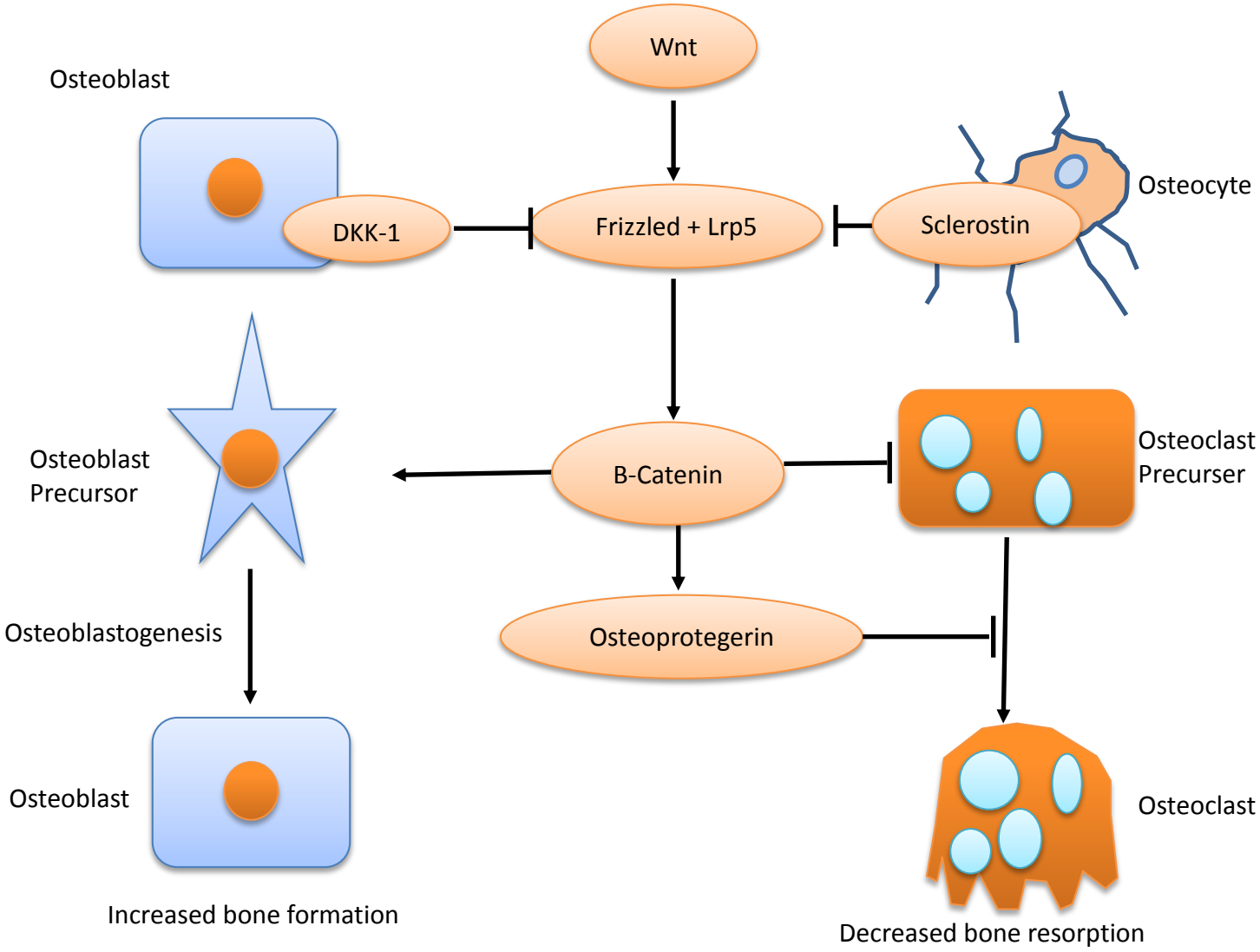


Figure 1-5: Canonical Wnt signaling and bone remodeling

Wnt/b-catenin induces osteoblastogenesis and increases bone formation. Canonical Wnt signaling suppresses osteoclastogenesis is suppressed by Wnt-induced OPG. Wnt signalling also suppresses osteoclast precursor's function/resorption by an OPG-independent mechanism. This co-regulatory effect of Wnt on osteoblast and osteoclast cells causes an overall increase in bone mass. Wnt antagonists sclerostin and Dkk-1 bind to Wnt co-receptor, thereby preventing Wnt-receptor interaction and signaling.

1.3 Effect of genetic variation in osteoporosis

Osteoporosis is one of the most widespread and costly diseases worldwide and the incidence of this disease is increasing due to the gradual ageing of the global population. Genetic factors are important in the pathogenesis of osteoporosis, with links to a reduction in BMD and an increased susceptibility to osteoporotic fractures (Cheung, Xiao et al. 2010). Family and twin studies indicate that more than 50% of the variance in peak bone mass is defined by genetic factors (Krall and Dawson-Hughes 1993; Gueguen, Jouanny et al. 1995). Although the role of genetic factors in the pathogenesis of bone loss remains unclear, heritability studies have determined that genetics does have a strong associations on a variety of osteoporotic fracture risk elements, which include quantitative ultrasound properties of bone, BMI and bone turnover markers (Kaprio, Rimpela et al. 1995; Arden, Baker et al. 1996; Hunter, De Lange et al. 2001). There is approximately a fourfold greater risk of having low whole body BMD in a son from a father with low BMD compared to a father with normal BMD and a five times higher risk in a daughter from a mother with low BMD compared to a mother with normal BMD (Zmuda, Sheu et al. 2006). Genetic effects on BMD are mediated by several genetic factors with mild influences (Uitterlinden, Ralston et al. 2006). More studies are needed for determining the relation between genetic factors and bone loss (Uitterlinden, Ralston et al. 2006). This information prompts us to identify the genetic factors that are associated with osteoporosis for therapeutic intervention.

It should be noted that sometimes patients have severe osteoporosis and fragility fractures or unusually high bone mass (HBM) that can be a characteristic of rare diseases which are genetic in nature and inherited in a classical monogenic mendelian manner such as: osteogenesis imperfecta, osteopetrosis, aromatase deficiency, oestrogen receptor deficiency, sclerosteosis, and high bone mass syndrome (**Table 1.1**) (Ralston 2010). In these diseases the effect of genes is severe. Despite the fact that these diseases are rare studies show that polymorphic variations in genes, which are mutated in this rare bone disease also exert influence on BMD in the general population (Uitterlinden, Ralston et al. 2006).

Disease	Phenotype	Genes
Osteogenesis Imperfect	Low BMD, fractures	COL1A1, COL1A2, CRTAP, LEPRE1
Osteopetrosis	High bone mass Fractures, bone marrow failure, blindness, osteoarthritis	TCIRG1, CATK, OSTM1, RANKL, RANK
Aromatase deficiency	Osteoporosis	CYP17
Oestrogen receptor deficiency	Osteoporosis, tall stature	ESR1
Sclerosteosis	High bone mass, bone overgrowth, nerve compression syndromes	SOST
High bone mass syndrome	High bone mass, torus palatines	LRP5

Table 1-1: Genes for monogenic bone diseases associated with abnormal bone mass

1.3.1 Osteoporosis Genome-wide association and candidate gene studies

Genetic association studies investigate correlation between genetic variation and disease risk to identify genome regions or candidate genes that contribute to a specific disease. A higher frequency of a SNP allele or genotype in a series of individuals affected with a disease is evidence that the tested variant increases the risk of a specific disease. Due to recent technical advances SNPs are the most commonly used in association studies. Association studies are used for identifying genes linked to the susceptibility to complex disorders. The term 'complex' is applied to traits and diseases which both multi genetic variants and environmental factors contribute to the risk of susceptibility. Role of different genetic variants has been studied in various illnesses like diabetes, heart disease, autoimmune diseases and psychiatric traits with each variant having only a subtle effect (Lewis and Knight 2012).

There are two different types of association study commonly used for the investigation of genetic variants in osteoporosis: genome-wide association study and candidate gene association study (Zmuda, Sheu et al. 2005; Richards, Zheng et al. 2012). In many different diseases associations with polymorphisms in candidate genes have been confirmed (Lohmueller et al. 2003). Many novel associations are being identified by GWAS in genes that had not been a priori strong candidate genes for the disease (Wellcome Trust Case Control Consortium 2007). The development of the GWAS over the last few years, has introduced a powerful tool for the investigation of the genetic architecture of human disease (Bush and Moore 2012). GWAS approach the analysis of complex traits by genotyping a very large number of markers across the genomic DNA of many people and relate these variants with observable traits such as: BMI, or blood pressure, or common diseases (Pearson 2008). . GWAS are particularly useful in complex disorders such as osteoporosis, asthma, cancer, diabetes, heart disease and mental illnesses because most genetic variants associated with these diseases in humans are believed to have mild effect on them (Stranger, Stahl et al. 2011).

The completion of the Human Genome Project in 2003 and the International Hap Map Project in 2005 provide facilities for studying the genetic contributions to common diseases. Human genome sequences and human genetic variation are available via computerized

databases. Whole-genome samples for genetic variations can be analysed quickly and accurately by current high-throughput sequencing technologies.

In the candidate gene approach genetic variants are tested based on a prior hypothesis regarding the role of that gene or gene product in the pathophysiology of the disease, but in the GWAS the whole genome can be investigated without a prior hypothesis of which genes or genetic variants have effect on the disease (Wilkening, Chen et al. 2009). Rather the hypothesis is that we believe there are risk genes to be found. The GWAS approach can be problematic because during this method very many statistics tests are performed that can be an unprecedented potential for false-positive results, so it has become common practice to follow GWAS positive signals with replication studies (Bush and Moore 2012).

The first influential GWAS was published in 2006 (Dewan, Liu et al. 2006); it determined the association between a gene and age-related macular degeneration (ARMD), the leading cause of acquired blindness in adults. Following this study, there has been a considerable increase in genomic discoveries involving complex diseases. Osteoporosis is a complex disorder influenced by multiple genes mostly with mild effect.

Candidate gene association studies aim to establish or characterise association between the polymorphic variants occurring within a specific gene or locus and a quantitative trait or disease of interest (Teare 2011). Candidate gene association studies have been used many times in the field of osteoporosis. Several polymorphisms associated with BMD, bone loss or osteoporotic fractures have been identified by performance of many candidate gene association studies but few findings have been replicated in independent studies (Ralston 2007). Meta-analysis is a statistical tool for combining results across genetic studies and is fast becoming the gold standard technique for resolving discrepancies in genetic association studies. The difficulties in producing robust, replicable results in genetic association studies are due to the small genetic effects and highlight the need for studies with many thousands of subjects (Munafo and Flint 2004). In one meta-analysis study by Richards et al, 150 human candidate genes were analysed for their associations with BMD or fractures. Results showed only 9 of 150 candidate genes were linked with the regulation of BMD (Richards, Kavvoura et al. 2009). As a result, candidate genes for this study were identified through a systemic review (explained in the method section 2.2.1) which was subsequently reported by candidate gene association study (CGAS), meta-analysis of candidate gene association study (MA-C), genome-wide association study (GWAS) and meta-analysis of genome-wide association

study (MA-G). Two extra candidate genes: VDR and P2RX7 were added to the list of selected candidate genes as explained in the method section 2.2.1.

1.3.2 Selected candidate genes from systemic review in addition two extra candidate genes

1.3.2.1 RANKL

As mentioned previously, RANKL had an essential regulatory role in osteoclastogenesis. Interaction between RANKL and its receptor RANK on osteoclast precursors, induces osteoclast differentiation and promotes bone resorption (Lau and Guo 2011). RANKL was initially analysed as a candidate gene associated with osteoporosis in 2006 (Li, Hou et al. 2010). Hsu et al assessed RANKL association with OPG and RANK in 1,120 Chinese with extremely low hip BMD or high hip BMD. This study found that the rs9594782 T>C polymorphism located in the 5-UTR of the gene, was associated with BMD in men. This association was not reflected in the female subjects. Hsu et al reported that men with the TC/CC genotypes were twice as likely to have extremely low hip BMD and also had lower whole body BMD (Hsu, Niu et al. 2006). The Deng group evaluated the association of RANKL with 19 candidate genes in a total of 1,873 individuals from 405 Caucasian nuclear families. Their data correlated with the Hsu et al, in that it supported the evidence that RANKL was strongly associated with hip BMD (Xiong, Shen et al. 2006).

According to the GWASs, RANKL SNPs rs9594759, rs9594738 and rs1021188 were selected (Paternoster, Lorentzon et al. 2013; Welter, MacArthur et al. 2014). The SNP rs9594759 located 113 kb upstream of RANKL, is known to be highly associated with spine BMD (Styrkarsdottir, Halldorsson et al. 2008). This association was replicated across the meta-analysis of five GWASs (Rivadeneira, Kavvoura et al. 2008). SNP rs9594738 was significantly associated with BMD in two GWASs (Styrkarsdottir, Halldorsson et al. 2008; Duncan, Danoy et al. 2011).

1.3.2.2 RANK

As the receptor for RANKL, RANK is an integral part of the RANK/RANKL/OPG axis and is involved in osteoclast differentiation and activation. The first RANK/BMD genetic

association studies by Choi et al reported no association between RANK SNP rs1805034 and BMD (Choi, Shin et al. 2006). However, in a Chinese study using a cohort of 1120 subjects, there was evidence of an association between SNP rs1805034 and hip BMD in men only (Hsu, Niu et al. 2006). The GWASs reported two RANK SNPs Rs3018362 and Rs884205 associated with BMD (Welter, MacArthur et al. 2014). In their initial GWAS, Styrkarsdottir and colleagues found that SNP rs3018362 located on 27 kb downstream of the *TNFRSF11A* gene, was consistently associated with hip BMD. However, this association did not reach genome-wide significance (Styrkarsdottir, Halldorsson et al. 2008). In their subsequent GWAS this association became significant genome-wide as a result of increased replication samples ($P = 5.4 \times 10^{8}$) and this study was the first to demonstrate that link between *RANK* and osteoporosis (Styrkarsdottir, Halldorsson et al. 2009). Two meta-analytic studies of GWAS data supported previous findings indicating RANK as an osteoporosis gene (Richards, Kavvoura et al. 2009; Rivadeneira, Styrkarsdottir et al. 2009). Furthermore, the results from two Genome wide meta-analysis studies supported evidence of the link between rs884205 and BMD (Rivadeneira, Styrkarsdottir et al. 2009; Estrada, Styrkarsdottir et al. 2012).

1.3.2.3 Osteoprotegerin

OPG as a member of the TNF receptor superfamily, is the soluble decoy receptor to RANKL which inhibits interaction between RANKL and RANK (Krakauer 2008). OPG has been considered an important gene candidate in osteoporosis due to findings that report polymorphisms in OPG are thought to genetically influence BMD (Pocock, Eisman et al. 1987; Hofbauer and Schoppet 2002; Langdahl, Carstens et al. 2002; Yamada, Ando et al. 2003; Arko, Prezelj et al. 2005; Vidal, Formosa et al. 2011). Lee et al collated the data from 8 studies and using an extensive literature search, evaluated the association between OPG polymorphisms and BMD (Garcia-Unzueta et al. 2008; Kim et al. 2007; Langdahl et al. 2002; MoVett et al. 2008; Ueland et al. 2007; Yamada et al. 2003; Zhao et al. 2005). The three OPG polymorphisms rs2073618, rs2073617, and rs3102735 were shortlisted as potential links to BMD. From these, SNP rs2073618 was shown to be associated with lumbar BMD in European and Asian populations. SNP rs2073618 was shown to be associated with femoral neck and total hip BMD only in the European population only (Lee, Woo et al. 2010). A

meta-analysis performed by Luo et al on the associations between OPG rs3102735 and rs2073618 polymorphisms and osteoporosis suggested that the G allele of the OPG rs3102735 could increase the risk of developing osteoporosis, whereas the C allele of rs2073618 was protective against osteoporosis (Luo, Hu et al. 2014).

GWASs reported two SNPs, rs64469804 and rs6993813 associated with both spine and hip bone BMD and they are in LD with the rs2073618 and rs3102735 polymorphisms in the OPG gene (Styrkarsdottir, Halldorsson et al. 2008). Another SNP, rs4355801, located in the 3`UTR of the OPG gene was found to be significantly associated with lumbar spine BMD ($P = 7.6 \times 10^{-10}$) and femoral neck BMD ($P = 3.3 \times 10^{-8}$) (Richards, Rivadeneira et al. 2008). Meta-analysis of five GWASs confirmed these findings (Rivadeneira, Styrkarsdottir et al. 2009). The significant association between OPG SNP rs2062377 and BMD was reported by three GWASs (Rivadeneira, Styrkarsdottir et al. 2009; Estrada, Styrkarsdottir et al. 2012; Paternoster, Lorentzon et al. 2013). In a study by Rivadeneira et al, findings suggested SNP rs11995824 association with total hip BMD (Rivadeneira, Styrkarsdottir et al. 2009). In addition, Hsu, Zillikens et al showed there are significant associations between rs2062375 and total hip BMD (Hsu, Zillikens et al. 2010).

1.3.2.4 ESR1

Estrogen plays an important role in the regulation of human bone mass and bone geometry (Gennari, Merlotti et al. 2005). Therefore, estrogen receptors are considered important therapeutic targets in osteoporosis bone disease (Ohta 2014). The impact of the gene encoding the estrogen receptor (ESR1) on bone growth was first demonstrated in a male subject who was homozygous for a loss-of-function mutation in ESR1. Sequencing of exon 2 showed that a cytosine-to-thymine transition at codon 157 of both alleles, resulting in a premature stop codon, showed a lack of ossification of the growth plate and low bone mass (Smith, Boyd et al. 1994).

The genetic link between BMD/fracture and ERS1 polymorphisms has been reported in numerous genetic studies. The ESR1 polymorphic sites reported are namely the XbaI (dbSNP: rs9340799), PvuII (dbSNP: rs2234693), and microsatellite thymine–adenine dinucleotide (TA) repeat polymorphisms. However, there are inconsistencies between the results reported in these studies (Ioannidis, Stavrou et al. 2002; Ioannidis, Ralston et al. 2004; Wang, Tang et al. 2007). The GWAS by Styrkarsdottir et al detected an association of the

ESR1 SNP rs1999805 with spine BMD ($P = 2\times10^{-8}$) and an association between SNPs rs1038304 and rs4870044 with both spine BMD (rs1038304, $P = 4\times10^{-11}$) (rs4870044, $P = 2\times10^{-11}$) and hip BMD (rs1038304, $P = 5\times10^{-9}$) (rs4870044, $P = 2\times10^{-7}$) in a European population (Styrkarsdottir, Halldorsson et al. 2008). In another study by this group an association of the SNP rs6929137 and spine BMD was detected ($P = 2\times10^{-10}$) (Styrkarsdottir, Halldorsson et al. 2009). The meta-analysis of five GWASs confirmed association between rs2941740 with Hip BMD ($P = 2\times10^{-10}$) and SNP rs2504063 with spine BMD ($P = 6\times10^{-11}$) (Rivadeneira, Styrkarsdottir et al. 2009).

1.3.2.5 LRP5

The LRP5 and LRP6 receptor proteins in the osteoblast cells are an important component of the Wnt signalling pathway involved in the regulation of bone mass and metabolism (Sassi, Sahli et al. 2014). LRP5 has a central role in the normal morphology, development and formation of bone (Cheung, Xiao et al. 2010). LRP5 protein transduces canonical signals downstream to stimulate the stem cell renewal, osteoblastogenesis and increases the ratio of OPG to RANKL (Urano and Inoue 2014). The key role of the *LRP5* gene in bone mass regulation was initially identified from the study of rare human monogenic skeletal diseases. Inactivating mutations in the human *LRP5* gene result in autosomal recessive osteoporosis-pseudoglioma syndrome characterised by a reduction in bone mass (Gong, Slee et al. 2001). Activating mutations in the *LRP5* gene cause autosomal-dominant bone disease characterised by an abnormal increase in bone mass (Van Wesenbeeck, Cleiren et al. 2003).

The most commonly studied non-synonymous LRP SNPs rs4988321 and rs3736228 were genotyped by GENOMOS investigators in 37,534 individuals across Europe and North America (van Meurs, Trikalinos et al. 2008). Results of these studies showed that both SNPs rs4988321 and rs3736228 were associated with reduced lumbar spine BMD, reduced femoral neck BMD and increased risk to vertebral fracture (van Meurs, Trikalinos et al. 2008). A Bayesian meta-analysis of ten studies constituting 16,705 subjects showed a modest effect of the SNPs rs4988321 on BMD (Tran, Nguyen et al. 2008). Although the molecular mechanism remains unclear, the Ala1330Val (rs4988321) mutants had significantly lower level of Wnt-signalling capacity when compared with the wild type molecule in two independent TCF-LEF reporter assays (Kiel et al. 2007; Urano et al. 2009).

The link between LRP5 SNP rs3736228 and BMD was highlighted by the GWAS (Richards, Rivadeneira et al. 2008) and the Genome-wide meta-analysis study (Estrada, Styrkarsdottir et al. 2012). GWASs also demonstrated an association between rs599083 and spine BMD ($P=5\times10^{-8}$) (Rivadeneira, Styrkarsdottir et al. 2009). A recent Genome-wide meta-analysis study by Zhang L et al showed a genetic link between LRP5 SNP rs525592 and spine BMD ($P = 3\times10^{-11}$) (Zhang, Choi et al. 2014).

1.3.2.6 Vitamin D receptor

The first candidate gene-linked to BMD was the vitamin D receptor (VDR) (Morrison, Qi et al. 1994). VDR is expressed in various tissues including the intestine, thyroid and kidney and has a pivotal role in calcium homeostasis. It is clear that the vitamin D endocrine system has an important regulatory role in the metabolism of bone. The specific interaction between VDR and 1, 25-dihydroxy vitamin D3 (1,25(OH)2D3) influences skeletal development, skeletal architecture, hormonal secretion and immune function (Singh, Singh et al. 2013). VDR represses the expression of the proximal activator of 1,25(OH)2D3, hydroxylase, and also stimulates the expression of the 1,25(OH)2D3 inactivating enzyme CYP24 (Christakos, Hewison et al. 2013).

Findings from recent studies have identified numerous polymorphisms in the VDR gene. Until now, research has primarily focused on the polymorphisms located on the limited area in the 3′ end of the VDR gene, namely; BsmI, ApaI, and TaqI (Marini and Brandi 2010). The Cdx-2 polymorphism was identified as another VDR polymorphism in the promoter region of the VDR gene, and has since been associated with BMD in Japanese women (Arai, Miyamoto et al. 2001). The cdx-2 polymorphism correlates with fracture in other populations but not specifically with BMD (Fang, van Meurs et al. 2003). Alleles that were identified in the promoter region and the 3' un-translated region were associated with an increased risk of fracture. It was revealed that the promoter haplotypes that increase fracture risk correlated with reduced VDR expression, whereas the 3' UTR risk haplotype was associated with increased degradation of VDR mRNA (Fang, van Meurs et al. 2005). Grunberg et al reconstructed three major haplotypes in the VDR 3'- un-translated region (UTR) baT, BAt, and bAT haplotypes that they named Hap1, Hap2, and Hap3, respectively. Their study aimed at determining any associations of Hap1, Hap2 and Hap3 with BMD and risk of fracture.

Their data showed that the Hap1 haplotype was associated with increased risk of vertebral fractures (OR, 1.655; 95% CI, 1.146–2.391; $p < 0.01$) (Grundberg, Lau et al. 2007).

In meta-analysis studies in which statistical procedures combine the results of several independent studies with a set of related research hypotheses, BsmI, ApaI, TaqI polymorphism were not found to be associated with BMD or fracture (Fang, Rivadeneira et al. 2006; Uitterlinden, Ralston et al. 2006; Qin, Dong et al. 2013; Ma, Zhou et al. 2014). Although some show an association between FOK1 and BMD (Wang, Liu et al. 2013) and BSMI and BMD (Jia, Sun et al. 2013) (Thakkinstian, D'Este et al. 2004).

1.3.2.7 P2RX7

Purinergic receptors are plasma membrane proteins involved in several cellular actions that have adenosine triphosphate (ATP) or other nucleotides as their natural ligands. To date three classes of purinergic receptors have been identified including P1 receptors, P2Y receptors, and P2X receptors. The P1 and P2Y molecules are G-protein-coupled receptors and P2X is a ligand-gated ion channel (Gartland 2012). Various P2Y and P2X receptor subtypes have been identified on bone cells. Eight recognised subtypes of P2Y and seven recognized subtypes of P2X are shown to be functionally expressed by bone cells (Wesselius, Bours et al. 2011). Activation of the P2-receptor subtypes by extracellular nucleotides can regulate several cellular actions such as bone formation, mineral deposition, cytokine release, bone resorption, and apoptosis. Thus bone formation and resorption in the bone microenvironment are modulated by these receptors(Rumney, Wang et al. 2012).

The P2RX7 gene is located in close proximity to the tip of the long arm of chromosome 12 (12q24), spans over 53kb and consists of 13 exons. P2RX7 seems to exert opposite effects on osteoblast and osteoclast function depending on extracellular ATP concentrations. Low extracellular concentrations of ATP are defined 0.2-2 μM and high extracellular concentrations of ATP are 200-2000 μM (Wesselius, Bours et al. 2011).Human osteoclasts, both *in vitro* and *in vivo* express P2X7 protein. Prolonged or repeated activation of the P2RX7 induces osteoclast death (Agrawal, Buckley et al. 2010). A study by Korcok et al showed that P2RX7 activation in osteoclasts that had low extracellular concentrations of ATP, enhanced multinucleated and mature osteoclast formation by increasing nuclear localisation of the transcription factor NF-κB, independently of RANKL. NF-κB is an important regulator of osteoclast formation and activity (Korcok, Raimundo et al. 2004).

Gartland et al showed for the first time that activation of P2RX7 in a sub-population of human osteoblasts that express functional P2RX7, leads to apoptosis in these cells (Gartland, Hipskind et al. 2001). In contrast P2RX7 activation stimulates osteogenesis via activation of alkaline phosphatase, which promotes mineralization by hydrolysing inorganic pyrophosphate (Zappitelli and Aubin) (Panupinthu, Rogers et al. 2008)

In other studies, it was shown that a 10-year fracture incidence in a cohort of approximately 1800 Danish postmenopausal women was significantly associated with both the rs3751143 and the rs1653624 polymorphism of P2RX7 (Ohlendorff, Tofteng et al. 2007). This study also showed that the rs3751143 polymorphism was closely related to ATP-induced osteoclast apoptosis *in vitro* as osteoclasts from individuals homozygous for risk allele were significantly less susceptible to apoptosis than osteoclasts from individuals carrying the wild type allele. Thus increased fracture incidence could be due to an increased number of osteoclasts. Gartland et al reported that post-menopausal women with the Gln^{307} mutation of rs28360457 polymorphism had a significantly lower lumbar spine BMD than women without this mutation. Also they showed that a combined group of heterozygous subjects who had any loss of function of 6 SNPs (rs35933842, rs208294, rs28360457, rs2230911, rs3751143, and rs1653624) had nearly nine-fold greater annualized percent change in LS-BMD than subjects who were wild-type at the SNP positions (Gartland, Skarratt et al. 2012). In addition, Wesselius et al further reported findings that suggested associations between BMD and four non-synonymous polymorphism (rs3751143, rs28360447, rs1262536, and rs2230912) of P2RX7 (Wesselius, Bours et al. 2013).

1.4 Allelic expression imbalance

Genome-wide association studies using SNP maps have revealed dozens of disease-associated loci and have revolutionised the mapping of common genetic loci determining susceptibility to a wide range of multifactorial disorders (Wellcome Trust Case Control 2007). Current evidence shows that only a small fraction of the causal loci consists of variants (non-synonymous SNPs) directly affect the protein amino-acid sequence. Therefore a large fraction of the loci is expected to have a regulatory role on gene expression via effects on transcription, splicing, and message stability and sometimes induces Allelic-specific gene expression (ASGE) (Heap, Yang et al. 2010).

ASGE or Allelic expression imbalance (AEI) is an important factor for phenotypic variability in humans and consequently for the development of common diseases (Schadt, Monks et al. 2003). The association between AEI and X-chromosome inactivation phenomena and genomic imprinting has been extensively studied (Constancia, Pickard et al. 1998) But new studies have demonstrated the variability of gene expression that exists within and between populations (Enard, Khaitovich et al. 2002; Cheung, Conlin et al. 2003) and now we understand that AEI is commonly present among non-imprinted genes (Cowles, Hirschhorn et al. 2002; Yan, Yuan et al. 2002). Allelic variation in gene expression and its relation with some human illnesses has been shown in association with certain genes (Yan, Dobbie et al. 2002; Yan, Yuan et al. 2002). Variation from gene expression may be caused by changes in the sequence of regulatory elements such as SNPs. Recent studies have demonstrated how this phenomenon is common through genome and tissues (Dixon, Liang et al. 2007; Stranger, Nica et al. 2007). 25-35% of differences seen between individuals in terms of allelic gene expression may be explained by these changes.

Therefore, in order to understand the extent of regulatory variation that is functionally important, we need AEI identification and characterization. Additionally, this will draw our attention to allelic differences in candidate genes that may guide us to relation between genetic variation and complex traits and common diseases (Palacios, Gazave et al. 2009).

On a basic level, gene expression is the process in which functional gene products are synthesised by using the information of genes. These products are proteins or non-protein coding genes, such as transfer RNA (tRNA) genes or ribosomal RNA (rRNA) genes. Any step of gene expression has the potential to be modulated from the DNA-RNA transcription step to post-translational modification of a protein. Recently several studies have shown that genetic variation at the sequence level can regulate the gene expression in a cell and this regulation is very important (Gaur, Li et al. 2013).

Most variation contributing to disease susceptibility, especially in complex disease, will alter expression rather than function of proteins (Williams, Chan et al. 2007). Such changes in gene expression may occur in two ways: by acting on genes at a distance through an expressed transcription factor (or non-coding RNA) on both the maternal and paternal chromosome (trans-acting factor), or by acting to regulate a gene on the same allele which resides on the same chromosome (cis-acting factor) (Monks, Leonardson et al. 2004; Cheung, Spielman et al. 2005). An inherited cis-acting variant, or less often, silencing of allele

expression from one parent (gene imprinting) induces allelic imbalance expression. Trans-acting factors and environmental backgrounds influence expression of both alleles: thus they are not responsible for allelic imbalance expression (Tung, Fedrigo et al. 2009).Cis-acting mutations may alter regulation for just one allele through a change to promoter, intron regions or enhancer elements that affect transcription factor-binding sites. The allele-specific expression is investigated by evaluation of cDNA from tissues expressing the genes of interest in the heterozygous individual because it is not possible to measure the relative allele expression in the homozygous person. This is instead of using traditional methods to identify SNPs *in vitro* which are time-consuming and challenging (Marsh 2007). A major advantage of this approach is that the two SNP alleles are measured in the context of the same trans-acting and environmental factors with likely cis-acting regulatory variants or parent of origin (imprinting) effects (Farh, Grimson et al. 2005).

Vidal et al demonstrated that at least 25 % of the human genes display Allelic Specific Expression (ASE) (Vidal, de Souza et al. 2011). The result of Heap et al showed that 4.6 % of heterozygous single-nucleotide polymorphism (SNP)–sample pairs have evidence of ASE ($P<0.001$). Using the Illumina Allele-Specific Expression platform, Serre et al investigate reported 2,968 SNPs located in 1,380 genes of more than 80 individuals and confirmed that differential allelic expression is a widespread phenomenon affecting the expression of 20 % of human genes. Although the complex regulatory mechanism remains unknown, it has been demonstrated that cis-regulation of some genes is shared across cell types (Emilsson et al. 2008; Myers et al. 2007; Schadt et al. 2008; Dimas et al. 2009). Consequently genetics have a significant role in variation on gene expression between individuals. GWAS show allelic imbalance expression in about 10–20% of expressed genes (if imbalance that sporadically occurs in one or only a few individuals is not considered) (Milani, Gupta et al. 2007; Serre, Gurd et al. 2008). These differences can correlate with phenotypic differences or even with risk of diseases (Teare, Heighway et al. 2006). Hence allelic imbalance expression analysis could be an interesting strategy for functional analysis of genetic factors. Grundberg et al. 2007 investigated the function of the VDR 3` haplotypes by studying allelic imbalance analysis of VDR transcripts in the normal chromosomal context of 70 unrelated human trabecular bone samples. This was performed by the quantitative genotyping of coding polymorphisms in RNA samples and in corresponding DNA samples. Both were sequenced in parallel from the trabecular bone samples. Finally they showed that over-expression of

VDR baT haplotype could be considered as a risk allele for osteoporosis (Grundberg, Lau et al. 2007).

1.5 Aims

The aims of the present study are as follows:

Aim 1: Investigate the genetic loci associated with BMD variation and osteoporosis in the Iranian population by bone mineral density, blood sampling, and genotyping on the recruited people. The extent of association between BMD and these candidates will be evaluated in the Iranian population sample.

Aim 2: Determine the effect of genetic variants in the function of osteoclast cells

Aim 3: Determine whether cis-acting factors of the genetic variations identified in aim 1 lead to allelic imbalance in different cellular populations, namely whole blood and differentiated osteoclasts.

Chapter 2 – Materials and Methods

2.1 Study participants and phenotypes

2.1.1 Iranian Multi-centre Osteoporosis Study cohort study

The Iranian Multi-centre Osteoporosis Study (IMOS) is a continuing study to determine reference values for bone densitometry in the female and male Iranian population from 2005. Five centres in southern, northern and central Iran (Tehran, Shiraz, Mashad, Booshehr, and Tabriz) take part in this project. Randomized clustered sampling was used to select healthy women and men aged 20 to 75 years of all areas of the corresponding cities. Data of the registered birth deliveries from all public and private hospitals of each city in a three consecutive day were gathered to define clusters. The distribution of the first deliveries has been shown to be highly comparable to the total population distributions concerning several factors such as, health status and socioeconomic (Larijani, Moayyeri et al. 2006). Fifty addresses of the first child deliveries were chosen as a basis for the selection of fifty clusters in each centre. First, the home address of the selected mother was found, and then twenty four neighbour houses were searched with a defined arrangement by proceeding round in a clockwise direction. Only one person from a randomly defined age range was selected for the study from each house. Hence in all of the five cities the number of invited subjects was 6000 persons, of whom 5339 persons accepted invitation from IMOS (response rate 88%) (Larijani, Moayyeri et al. 2006).

Subjects with the following conditions or diseases were excepted to confirm all participant are healthy subjects for this study: known history or evidence of hepatic or renal failure, rheumatoid arthritis, parathyroid or adrenal disease, thyroid disease, metabolic bone disease, sterility, diabetes mellitus, oligo-menorrhea, alcoholism, malignancy, mal-absorption, lactation, pregnancy, immobility for more than one week, medications influencing bone metabolism, and smoking more than 10 cigarettes per day. In sum 437 persons were omitted from the study. The study protocol was permitted by the research ethics committee of the Iranian Ministry of Health and Medical Education. Written informed consent got from all contributors.

Subjects for this PhD study were selected from the Shiraz centre. In the Shiraz centre 540 subjects were selected randomly from 1080 subjects who were registered in the IMOS cohort.

Invitations were sent to them to attend the Nemazi hospital for this study and study participant visited the hospital between May 2009 and February 2010. All participants provided written informed consent after being fully informed of the nature of the study.
Height was read from a free-standing height with the nearest 0.5cm. Subjects were weighed in light street clothes without shoes using Soehnle® digital scale with the nearest 100 gr. Body mass index (BMI) was calculated as weight (kg) divided by the square of height (m). Bone mineral density and blood samples were taken. Definitive data of bone mineral density measurements of the total proximal femur and lumbar vertebrae (L2–L4) in the antero-posterior position are available. Blood sampling was taken from 512 persons for DNA extraction. Bone mineral densitometry was done for 501 persons. Finally there were 501 persons who blood samples and BMD were available, as a major part of the work of this project. Each participant spent around two hours in Nemazi hospital for all the above work up.

2.1.2 Bone mineral density (BMD) measurement

The most precise method available for bone densitometry in clinical practice is dual-energy X-ray absorptiometry (DEXA). It is used to measure appendicular and axial sites and can be used to measure total body bone mass. Minimal radiation exposure was performed as produced in the method by Black et al (Blake and Fogelman 1997). In DEXA two discrete low energies are generated by an X-ray tube. Low energy beams are attenuated more in bone than in soft tissues. For calculating the attenuation caused by the bone and to neutralize the effect of soft tissue, two simultaneous absorption curves are generated, after measuring the absorption of each of the two X-ray beams. In this technology an area of bone density (gm/cm^2) is calculated from linear measurement of bone mineral content (gm or gm/cm) by dividing the bone mineral content by the area that was scanned (Blake and Fogelman 1997). In this study the recruited subjects were referred for BMD measurement using a GE Lunar DPX-IQ Bone densitometer (GE Healthcare, Madison, WI USA) at the endocrine research centre of Nemazi hospital. Definitive data of bone mineral density measurements of the total proximal femur and lumbar vertebrae (L2–L4) in the anterior-posterior position are available. The QC of BMD measurement is essential to be sure about the consistency of BMD measurement. For this purpose, Error ranges from measuring device, radio-technologist's experience, and patient's condition should be considered and minimized (Kim and Yang

2014). Daily QC is necessary for checking device condition. The tendency of BMD change is assessed by evaluating results of repeated scans of same subject which the average value, standard deviation and coefficient of variation can be found for each set of repeated measurements (Nagy and Clair 2000). For instance, if the accuracy of a Dual X-ray absorptiometry of total hip bone density measurement have a %CV of 2-5%, this indicated that the measurement of BMD tend to vary by 2-5% of the real BMD. In fact the %CV is describing the variability of the measurement about the true value. Therefore, the smaller the %CV means the more accurate device.

2.2 Genotyping

2.2.1 Systemic review of the literature for selection of candidate genes and SNPs of osteoporosis

Systematic reviews typically involve a detailed and comprehensive search strategy with the objective of reducing bias by identifying and extracting specific data sets from relevant studies (Uman 2011). This study required the systematic review of publications that were relevant to bone mineral density, which involved genetic association in human populations. As was discussed in the introduction section bone mineral density is considered an important risk factor to predict the likelihood of fragility fracture and is widely used in clinical practice to identify those at increased risk for fracture (Johnell, Kanis et al. 2005). Other risk factors that should be considered in assessment of fracture risk are such as: age, sex, race, country of residence, history of fragility fracture, corticosteroid use, and current smoking (Tremollieres, Pouilles et al. 2010). According to diagnostic criteria for osteoporosis was discussed in the introduction section, it is sensible to study association with BMD in osteoporosis study.

The method of systematic review was performed as previously outlined (Xu, Dong et al. 2010). The systemic review of the literature involved developing search strategies and locating the studies, selecting relevant studies and extracting data.

1. In the search strategy stage for the selection of genes and their SNPs associated with osteoporosis, the literature from 2000 to December 2010 was scanned through the

Web of Knowledge and NIH database for genes using the searching key words "BMD" "osteoporosis" in combination with "association" "polymorphisms" and "gene".

2. This search retrieved article studies reported by: candidate gene association study (CGAS), meta-analysis of candidate gene association study (MA-C), genome-wide association study (GWAS) and meta-analysis of genome-wide association study (MA-G).

In this review, due to space limitations, only the most representative results that have had an immediate influence on understanding and research of genetic mechanisms underlying osteoporosis are discussed. Based on this search *LRP5, RANK, RANKL, OPG and ESR1* were selected as genes that were associated with BMD in the literature reported by all CGAS, MA-C, GWAS and MA-G studies. These genes are also involved in three biological pathways, the Wnt/-catenin signalling pathway, the estrogen endocrine pathway, and the RANKL/RANK/OPG pathway.

GWAS review was used for the strategic selection of single nucleotide polymorphisms (SNPs) that were the most important in BMD. The catalogue of published GWAS in national human genome research institute (NHGRI) was used for searching risk SNPs of selected candidate genes (Table 2-1) (Hindorff, Sethupathy et al. 2009). There are some specific criteria for the selection of association studies in the national human genome research institute. Studies that searched at least 100,000 SNPs in the initial stages were chosen and the SNP-trait associations listed were limited to those with p-values of $< 1.0 \times 10^{-5}$. By the end of 2010, the database summarising the information comprising fourteen GWAS related to osteoporosis and BMD were gathered in the national human genome research institute of which one is was meta-analysis study of five GWAS.

Reported gene	Trait	Strongest SNP risk allele	Articles	P-value	Beta-coefficient and [95% CI]
OPG (The RANKL/RANK /OPG pathway) Established gene (CGAS+ MA-C +GWAS +MA-G)	BMD (hip)	rs11995824-G	(Rivadeneira, Styrkársdottir et al. 2009)	4×10^{-9}	.07 [0.05-0.09]
	BMD (spine)	rs2062377-T		4×10^{-16}	.09 [0.07-0.11]
	BMD	rs4355801-A	(Richards, Rivadeneira et al. 2008)	8×10^{-10}	.09 [NR]
	BMD (spine)	rs2062375-G	(Hsu, Demissie-Banjaw et al. 2008)	3×10^{-11}	.14 [NR]
	BMD (spine)	rs6469804-A	Styrkarsdottir, Halldorsson et al. 2008)	7×10^{-15}	.12 [0.09-0.15]
ESR1 (The estrogen endocrine pathway) Established gene (CGAS+ MA-C +GWAS +MA-G)	BMD (hip)	rs2941740-G	(Rivadeneira, Styrkársdottir et al. 2009)	2×10^{-10}	.07 [0.05-0.09]
		rs1038304-G	(Styrkarsdottir, Halldorsson et al. 2009)	4×10^{-10}	.08 [0.06-0.11]

Reported gene	Trait	Strongest SNP risk allele	Articles	P-value	Beta-coefficient and [95% CI]
		rs1038304-G	(Styrkarsdottir, Halldorsson et al. 2008)	5 x 10-9	.08 [0.06-0.11]
	BMD (spine)	rs4870044-T		2 x 10^{-11}	.11 [0.08-0.14]
	BMD (spine)	rs6929137-A	(Styrkarsdottir, Halldorsson et al. 2009)	2 x 10^{-10}	.1 [0.07-0.13]
	BMD (spine)	rs2504063-A	(Rivadeneira, Styrkársdottir et al. 2009)	6 x 10^{-11}	.08 [0.06-0.10]
	BMD (spine)	rs1999805-C	(Styrkarsdottir, Halldorsson et al. 2008)	2 x 10^{-8}	.09 [0.06-0.12]
LRP5 (The Wnt/β-catenin signalling pathway)	BMD (spine)	rs599083-G	(Rivadeneira, Styrkársdottir et al. 2009)	5 x 10^{-8}	.07 [0.05-0.09]
Established gene (CGAS+ MA-C +GWAS +MA-G)	BMD	rs3736228-T	(Richards, Rivadeneira et al. 2008)	6 x 10^{-12}	.13 [NR]
RANKL	BMD (hip)	rs9594738-T	(Styrkarsdottir,	2 x 10^{-8}	.1 [0.06-0.13]

Reported gene	Trait	Strongest SNP risk allele	Articles	P-value	Beta-coefficient and [95% CI]
(TNFRSF11) (The RANKL/RANK /OPG pathway) Established gene (CGAS+GWAS+MA-G)	BMD (spine)	rs9594759-T	Halldorsson et al. 2008)	2 x 10^{-21}	.17 [0.14-0.21]
	BMD	rs1021188-C	(Paternoster, Lorentzon et al. 2010)	2 x 10^{-14}	3.89 [1.91-5.87]
RANK (TNFRSF11) (The RANKL/RANK /OPG pathway) Established gene (CGAS+GWAS+MA-G)	BMD	rs884205-A	(Rivadeneira, Styrkársdottir et al. 2009)	6 x 10^{-9}	.09 [0.06-0.12]
	BMD (hip)	Rs3018362-A	(Styrkarsdottir, Halldorsson et al. 2008)	1 x 10^{-6}	.07[0.04-0.10]

Table 2-1: List of candidates arising from CGAS, MA-C, GWAS, and MA-G studies with their risk SNPs (Hindorff LA ; Li, Hou et al. 2010)

The VDR gene was the first candidate gene studies with respect to osteoporosis identifies in several CGAS and MA-C studies within the literature (Li, Hou et al. 2010). In Particular the VDR genes, as the most studied candidate gene related to osteoporosis in Iranian population (Hossein-nezhad, Varzaneh et al. 2011; Pouresmaeili, Jamshidi et al. 2013) were also added to the panel of BMD associated genes for genotyping. Additionally, Gartland *et al.* 2012 has data to show that P2RX7 is a novel candidate gene associated with BMD and also Ohlendorff *et al.* 2007 showed association between P2RX7 gene and osteoporotic fracture risk (Ohlendorff, Tofteng et al. 2007; Gartland, Skarratt et al. 2012). As a result this gene was also added to the panel of BMD associated genes for genotyping (Table 2.2)

Gene	Trait	Risk	Articles	P-value	beta-coefficient and (95% CI)
VDR (CGAS+ MA-C)	BMD (Spine)	Rs2228570-T	(Wang, Liu et al. 2013)	0.001	-0.87 (-1.38,0.35)
	BMD (Femoral neck)			0.001	-0.43 (-0.93, 0.06)
P2RX7 (CGAS)	BMD (Spine)	rs28360457-A	(Gartland, Skarratt et al. 2012)	0.004	-0.12 (NR)
	BMD	rs3751143	(Wesselius, Bours et al. 2013)	0.018	2.47(1.15, 5.32)

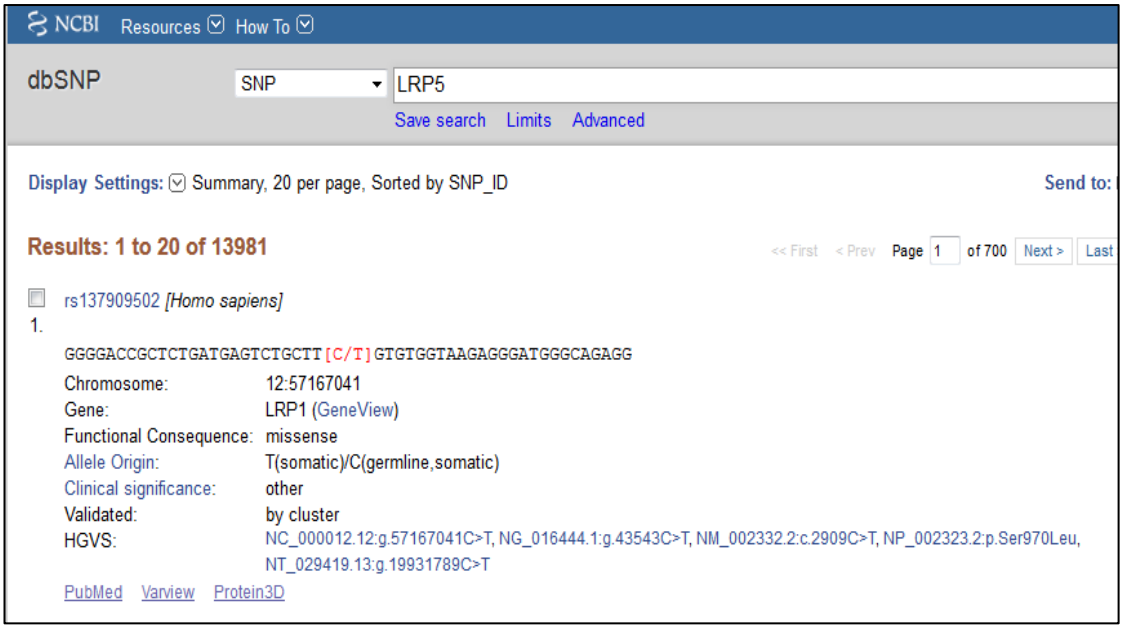
Table 2-2: List of candidates arising from CGAS and MA-C or just CGAS studies with their risk SNPs

2.2.2 Selection of transcribed SNPs using NCBI reference assembly database

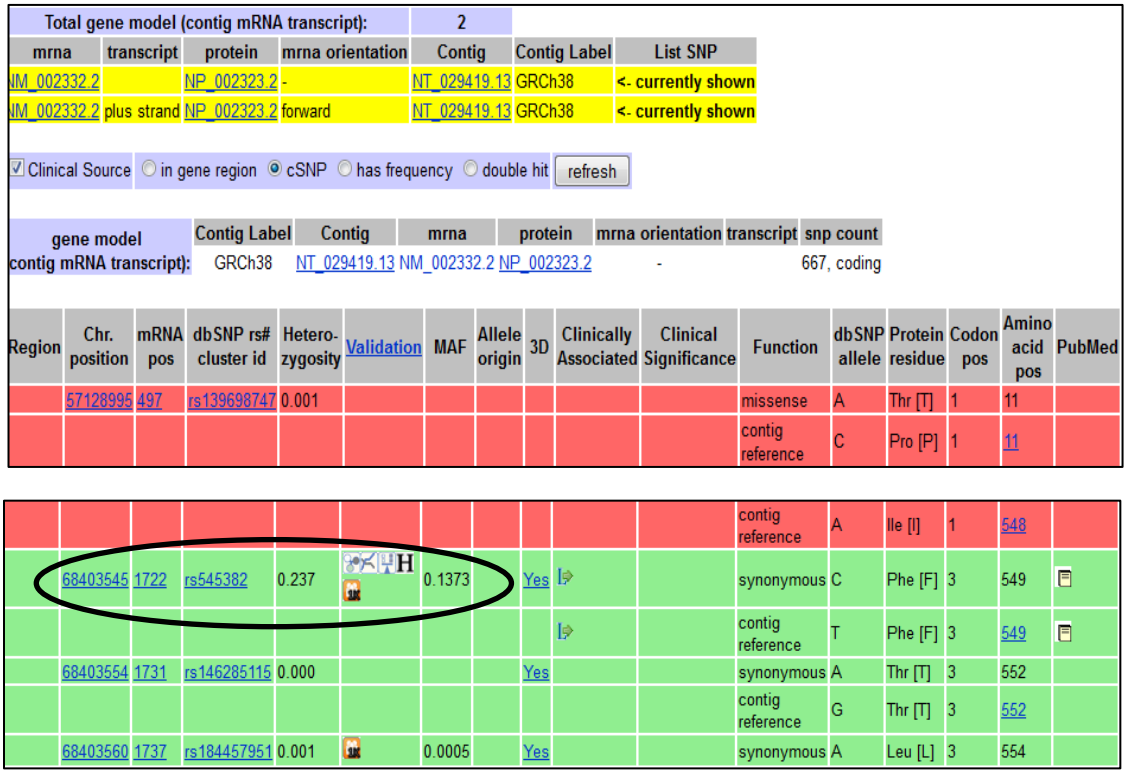
To perform Allelic Expression Imbalance (AEI) analysis, individuals heterozygous for transcribed SNP were chosen where the contribution of each of the two alleles assessed by quantifying the relative amount of transcripts from each. Unequal expression was then defined as AEI. Therefore, the transcribed SNPs with the highest heterozygosity for the candidate osteoporosis genes identified in **section 2.2.1** (LRP5, RANK, RANKL, P2RX7, ESR1, OPG and VDR) were determined using the NCBI reference assembly (http://www.ncbi.nlm.nih.gov/projects/SNP). The osteoporosis-gene candidate transcribed SNPs were selected using the NCBI database software as follows:

1. The search to identify a candidate transcribed SNP was initiated at the NCBI homepage by entering the gene name in 'SNP' search bar.
2. The SNP was selected from the retrieved results i.e reported and annotated SNP for the query gene was retrieved (**Figure 2.1 A**).
3. Under the 'Gene View' section cSNP was selected from the gene variation options and the resulting transcribed SNPs were displayed in a table. A range of information associated with the transcribed SNPs for the candidate gene was available including chromosome position, clinical association/significance and importantly, heterozygosity (**Figure 2.1 B**).
4. Following the transcribed SNP selection process, the SNP with the highest heterozygosity was selected from the complete transcribed SNP list retrieved by NCBI for future AEI study analysis. The criteria for deeming a transcribed SNP as 'highly heterozygous' was directly related to the associated heterozygosity value. For example a SNP that had a heterozygosity value of 0.237 was more heterozygous compared to a SNP with a heterozygosity value of 0.001 (**Figure 2.1 B**). The minor allele frequency (MAF) refers to the frequency at which the least common allele occurs in a given population. Where heterozygosity could not be defined for transcribed SNPs, the MAF could be a more useful clue to identifying transcribed SNPs with more chance to be heterozygous in the population table (**table 2.3**).

A



B



Total gene model (contig mRNA transcript):				2		
mrna	transcript	protein	mrna orientation	Contig	Contig Label	List SNP
NM_002332.2		NP_002323.2	-	NT_029419.13	GRCh38	<- currently shown
NM_002332.2	plus strand	NP_002323.2	forward	NT_029419.13	GRCh38	<- currently shown

Clinical Source | in gene region | cSNP | has frequency | double hit | refresh

gene model contig mRNA transcript):	Contig Label	Contig	mrna	protein	mrna orientation	transcript	snp count
	GRCh38	NT_029419.13	NM_002332.2	NP_002323.2	-		667, coding

Region	Chr. position	mRNA pos	dbSNP rs# cluster id	Hetero-zygosity	Validation	MAF	Allele origin	3D	Clinically Associated	Clinical Significance	Function	dbSNP allele	Protein residue	Codon pos	Amino acid pos	PubMed
	57128995	497	rs139698747	0.001							missense	A	Thr [T]	1	11	
											contig reference	C	Pro [P]	1	11	
											contig reference	A	Ile [I]	1	548	
	68403545	1722	rs545382	0.237		0.1373		Yes			synonymous	C	Phe [F]	3	549	
											contig reference	T	Phe [F]	3	549	
	68403554	1731	rs146285115	0.000				Yes			synonymous	A	Thr [T]	3	552	
											contig reference	G	Thr [T]	3	552	
	68403560	1737	rs184457951	0.001		0.0005		Yes			synonymous	A	Leu [L]	3	554	

Figure 2-1: Transcribed SNPs with the highest heterozygosity were selected from the retrieved results for the query gene

The SNPs associated with LRP5 are used as an example.

Transcribed SNP	MAF
ESR1	
2077647	0.4665
OPG	
2073618	0.3333
1804854	0.0881
LRP5	
556442	0.4207
2306862	0.1102
545382	0.1585
RANKL	
9562415	0.0415
RANK	
1805034	0.4058
35211496	0.0573
P2RX7	
2230911	0.1649
2230912	0.0693
208294	0.47
7958311	0.2558
1718119	0.3225
1621388	0.3203
VDR	
731236	0.2766

Table 2-3 : Transcribed SNPs of selected candidate genes with the highest MAF in the NCBI

2.3 DNA Extraction

Material	
ITEM	**SUPPLIER**
QIAamp 96 Plates 4 12 **S-Blocks** **Collection Microtubes (racked)** **Elution Microtubes CL (racked)** **Caps for Collection Microtubes** **Caps for Elution Microtubes** **AirPore Tape** **QIAGEN® Protease** **Protease Solvent** **Buffer AL** **Buffer AW1 (concentrate** **Buffer AW2 (concentrate)** **Buffer AE** **96-Well Plate Register** **Selection Guide**	Qiagen
Molecular-grade Ethanol	Sigma-Aldrich
Heratherm Incubator	Thermo Scientific

The QIAamp 96 DNA Blood Kit was used for the isolation of DNA from peripheral blood and using a fast 96-well-plate procedure. In principle, DNA binds specifically to the QIAamp silica-gel membrane while the contaminants pass through and are removed. Inhibitors of polymerase chain reaction including divalent cations and proteins are completely removed in wash steps. As a result pure nucleic acid is eluted in water or buffer ready to use in gene expression procedures. The purified DNA is up to 50 Kb in size and DNA of this length denatures completely. Typically a yield of 6 µg per 200 µl can be obtained from healthy whole blood.

Briefly, the protease stock solution at a volume of 20µl was pipetted into the bottom of the collection micro-tubes. 200µl of human peripheral blood was collected by brachial vein puncture method (Galena 1992; Okeson and Wulbrecht 1998). Blood was collected into micro-tubes by touching the insides of the tubes without wetting the rims. Buffer AL in the amount of 200 µL was added to each blood sample, without wetting the rims of the collection

micro-tubes for cell lysis. Collection micro-tubes were sealed by using the caps. The rack of micro-tubes was covered with the plastic cover provided and mixed thoroughly by shaking vigorously for 15 seconds. The rack was centrifuged briefly at 6000 x g to collect any solution from the caps and then incubated at 70°C for at least 10 min. The rack was centrifuged briefly at 6000 x g rpm to collect any lysate from the caps. Caps were removed and 200 µL of 100% ethanol added to each tube. The tubes were sealed by using new caps for collection micro-tubes and shaken vigorously for 15 seconds. The rack was centrifuged briefly at 6000 x g to collect any solution from the caps.

The QIAamp 96 plate was placed on top of an S-Block. The plate was marked for later identification. The mixture of 620 µL per collection micro-tube was applied carefully to the QIAamp 96 plate. The QIAamp 96 plate was sealed with an AirPore Tape sheet. The S-Block (96-well blocks with 2.2 ml wells), and QIAamp 96 plate were loaded onto the carrier, placed in the rotor bucket, and centrifuged at 6000 x g for 4 minutes. The tape was removed and 500 µl Buffer AW1 (wash buffer for removing all residual contaminants without affecting DNA binding) was added to each well. The QIAamp 96 plate was sealed with a new AirPore Tape sheet and centrifuged at 6000 x g for 2 minutes. The tape was removed and 500 µL Buffer AW2 was added carefully to each well for removing all residual contaminants without affecting DNA binding. Each well was centrifuged at full speed 20000 x g for 3 minutes.

The QIAamp 96 plate was placed on top of a rack of elution micro-tubes. For elution of DNA, 200 µL Buffer AE, equilibrated to room temperature, was added to each well using a multichannel pipette. The QIAamp 96 plate was sealed with a new AirPore tape sheet, incubated for 1 minute at room temperature and centrifuged at 6000 x g for 4 minutes. The wells of the micro-tubes were sealed with the caps and stored at -20°C.

2.4 Genotyping using Competitive Allele-Specific Polymerase chain reaction (KASPar) assay

Single nucleotide polymorphisms (SNPs) genotyping was performed by Competitive Allele-Specific Polymerase chain reaction (KASPar) assay by KBioscience services. SNP genotyping is the measurement of genetic variations of SNPs between members of a species using a SNP-based genotyping platform. KASPar assay SNP genotyping system from KBioscience is based on Fluorescent Resonance Energy Transfer (FRET) technology. In this SNP-genotyping platform, FRET is coupled with the power of competitive allele specific PCR, allowing for determination of SNP or insertion/deletion genotypes. Following completion of the reaction, the data reported were evaluated using the KlusterCaller 1.1 software (KBioscience).

Distribution of genotype frequencies was tested for consistency to Hardy–Weinberg equilibrium. Three SNPs showed a deviation from HWE (at the p=0.05 level) which were excluded. For quality control missing data for each SNP and in each sample were calculated. If missing data for each SNP were more than 25% of data that SNP was excluded which missing data for all SNPs were less than 25% of data in this genotyping. If missing data in each sample were more than 5% of data, that sample was excluded which ten samples were excluded.

2.5 Sample selection procedure for Allelic Imbalance Expression and osteoclast function study

In the first part of this study, 540 individuals were selected randomly and both blood sampling and BMD were done for 501 individuals as described in the **section 2.1.1.** Genotyping of the peripheral blood samples donated from these 501 volunteers were performed as described in the **section 2.4.** Due to the time and cost limitations of our core genomic facility in Shiraz, it was not possible to include all 501 samples within the AEI study. Thus, the minimum numbers of volunteer peripheral blood and osteoclast culture samples that were necessary to produce statistically viable results within the AEI study were calculated by "**sample size tables for clinical studies software".**

It was feasible and affordable to use 16 osteoclast cultures and perform AEI analysis on 52 blood samples. By using the above software, assuming a two-sided 0.05 significance level, 16 samples would provide 80% power to detect a true difference in the mean ratios of 0.8 of

standard deviation. Assuming 0.05 significant level, 52 samples would provide 80% power to detect a true difference in the mean ratios of 0.4 of a standard deviation

The above calculations showed that with these relatively modest sample sizes it was possible to detect medium effect sizes within the AEI study. There was a minimum requirement of 52 peripheral blood samples and 16 osteoclast culture samples. The total numbers of 52 subjects were selected to maximise the number of heterozygosity at each of the transcribed SNPs. This would ensure that the sample of 52 (16) would be maximally informative for the AEI for all selected genes. This was performed by sorting the samples according to the number of heterozygous SNPs of the six transcribed SNPs: rs556442, rs2073618, rs208294, rs2077647, rs1805034, rs731236. Each SNP was coded as heterozygous or homozygous according to genotyping result. First sixteen samples were selected for osteoclast culture (Table 2.4).

Sample ID	All	rs556442	rs2073618	rs208294	rs2077647	rs1805034	rs731236
A498AL	6	√	√	√	√	√	√
R255NA	6	√	√	√	√	√	√
K90HA	6	√	√	√	√	√	√
S539SI	6	√	√	√	√	√	√
T513MO	5	√	√	√	√	√	-
A266FA	5	√	√	√	√	√	-
R438AG	5	√	√	√	√	√	-
E335RA	5	-	√	√	√	√	√
S189ZO	5	√	√	√	√	√	-
S116KH	5	-	√	√	√	√	√
S200ZE	5	√	√	-	√	√	√
A2AT	5	√	-	√	√	√	√
G179AL	5	-	√	√	√	√	√
K455HO	5	√	-	√	√	√	√
M272MA	5	√	-	√	√	√	√
G472GO	5	√	√	-	√	√	√
G204OM	5	-	√	√	√	√	√
S137AF	5	√	-	√	√	√	√
E388FA	5	-	√	√	√	√	√
M345SH	5	-	√	√	√	√	√
S203ME	5	√	-	√	√	√	√
S475YA	5	√	-	√	√	√	√
S343ZE	5	-	√	√	√	√	-
M288PE	5	-	√	√	√	√	√
S206KH	5	√	-	√	√	√	√
E228NA	5	√	√	√	-	√	√
N479RE	5	√	√	√	-	√	√
M320AB	5	√	√	√	-	√	√
A555SH	5	√	√	√	-	√	√
A329AK	5	√	√	√	√	-	√
G383KE	5	√	√	√	√	-	√
S446HA	5	√	√	√	√	-	√
P571HA	5	√	√	√	√	-	√
R177KO	5	√	√	√	√	-	√
M450SA	5	√	√	√	√	-	√
G298LE	5	√	√	√	√	-	√
K339MO	5	√	√	√	√	-	√
G461BE	5	√	√	-	√	-	-
R119ES	5	√	√	√	√	-	√
G257SH	5	√	√	√	√	-	√
K551EB	5	√	√	√	-	√	√
P414GO	4	√	√	-	-	√	√
A360ZE	4	-	√	√	-	√	√
M426NO	4	√	-	√	-	√	√
A469AM	4	√	-	√	-	√	√
S181AL	4	√	√	-	-	√	√
S460MO	4	√	√	-	-	√	√
S378SO	4	-	√	√	-	√	√
G299HA	4	-	√	√	-	√	√
K282MO	4	√	-	√	-	√	√
N408MO	4	-	√	√	-	√	√
N575BE	4	√	√	√	-	√	-
Total	-	**40**	**42**	**46**	**36**	**41**	**45**

Table 2-4: Sample sorting according to the number of heterozygous SNPs of the six transcribed SNPs: rs556442, rs2073618, rs208294, rs2077647, rs1805034, rs731236.

The 52 subjects were given identification codes outlined in the first column. The ticks represent a heterozygous SNP for a particular subject. The total number of subjects out of the possible 52 subjects, heterozygous for each of the 6 SNPs could be determined.

2.6 Assessing genotype group counts between transcribed and risk SNPs using cross tabulate function

AEI is achieved in heterozygous transcribed samples and therefore should be defined taking into consideration the potential link between the heterozygous risk SNPs and heterozygous transcribed SNPs. A pivot table was used to cross tabulate the genotyping data of transcribed and risk SNPs within the 52 selected subject samples. For each of the six candidate genes, all genotyping data of transcribed and risk SNPs were transferred into excel worksheets to form the pivot table. Data collected from the transcribed SNPs were inserted into excel rows and the risk SNPs were inputted into excel columns of the pivot tables as described in the example illustrated in **Table 2.5**. Subsequently, the relationship between each of the transcribed SNPs and risk SNPs were assessed to determine a genetic association.

		Risk SNP rs3018362			
		A:A	G:A	G:G	Grand Total
Transcribed SNP rs1805034	C:C	0	3	3	6
	T:C	3	18	23	44
	T:T	0	0	2	2
	Grand Total	3	21	28	52

Table 2-5: An example of a pivot table to assess genotype group links of a transcribed SNP and risk SNP

The transcribed rs1805034 SNP with its three genotypes (C: C, C: T, and T: T) were inserted in the pivot table columns and the risk rs3018362 SNP inserted in the row columns. This pivot table analysis showed that in the 52 samples there were 44 samples which were heterozygous for rs1805034 with the T: C genotype and 18 out of these 44 samples were also heterozygous for rs3018362 with G: A genotype.

2.7 Cell culture

Material	
IITEM	**SUPPLIER**
Haemocytometer - Depth 0.1mm, 1/400mm^2	Hawksley
Cell culture flat bottom 96-well plates	Corning® Costar® (New York, USA)
BD Falcon™ Centrifuge tube – 15ml, 50ml	Fisher Scientific
Eppendorph Tubes (0.05 abd 1.5ml)	Starstedt
GPR Centrifuge	Beckman Coulter
Microcentrifuge	IEC Micromase
Water-cooled Labcut 1010 low-speed diamond saw	Agar Scientific Ltd (Stansted, UK)
Water Bath	Fisher Scientific
Vortex Mixer	VLP® Scientifica
Z2 Counter	Bechman Coulter Inc.
MS columns for magnetic selection of CD14+ monocyte/macrophage population	Miltenyi Biotec Ltd. (Bisley, UK)
Glass coverslips of 6mm diameter	Richardson's of Leicester (Leicester, UK)
Minimum Essential Media (MEM) -Alpha+ GlutaMAX	Gilbco™, Invitrogen (Paisley, UK)
Human recombinant RANKL	Insight Biotechnology Ltd. (Wembley, UK)
Macrophage Colony Stimulating Factor (M-CFS)	R&D (Minneapolis, USA)
Fetal calf serum (FCS)	Gilbco™, Invitrogen (Paisley, UK)
Penicillin: Streptomycin Solution (PenStrep)	Gilbco™, Invitrogen (Paisley, UK)
CD14 Microbeads	Miltenyi Biotec Ltd. (Bisley, UK)
Phosphate Buffered Saline (PBS) (pH-7.4 Ca2+ and Mg2+ free)	Sigma-Aldrich (Poole, UK)
UltraPure™ 0.5M, Ethylenediaminetetraacetic acid (EDTA)	Gilbco™, Invitrogen (Paisley, UK)

(pH 8.0)	
Trypan blue	Sigma-Aldrich (Poole, UK)
Lithium-Heparin and BD SST™ vacutainers	BD Biosciences (Oxford, UK)
Histopaque®-1077	Sigma-Aldrich (Poole, UK)

2.7.1 Isolation of Peripheral Blood Mononuclear Cells (PBMCs)

To study the effect of genetic variation on osteoclast and function *in vitro*, peripheral blood (up to 100mL) was collected in lithium-heparin vacutainers from healthy volunteers. Human peripheral blood was collected by brachial vein puncture method (Galena 1992, Okeson and Wulbrecht 1998). Subjects were recruited from Shiraz Medical University approved by the University of Shiraz Research Ethics Committee. Informed Ethical consent was obtained from the volunteers prior to sample collection. The collected blood was diluted 1:1 ice-cold PBS supplemented with 2mM EDTA (Buffer 1) and 8 mL of this mixture was layered over 5 mL of Histopaque®-1077 in 15 mL Falcon tubes. Blood was centrifuged at 400g for 30 minutes at room temperature without break. Following centrifugation, the upper layer of the blood mixture was removed and the opaque interface containing the monocytes and lymphocytes fraction of blood was collected by using fine pipettes (**Figure 2.2**). Subsequently, the PBMCs fraction was washed with Buffer 1 at 300g and washed again at 200g for 10 minutes.

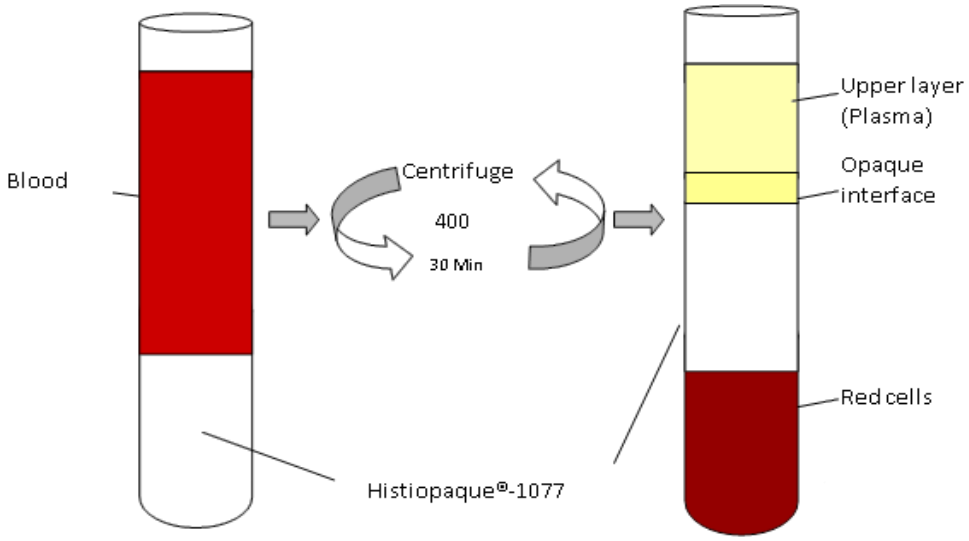
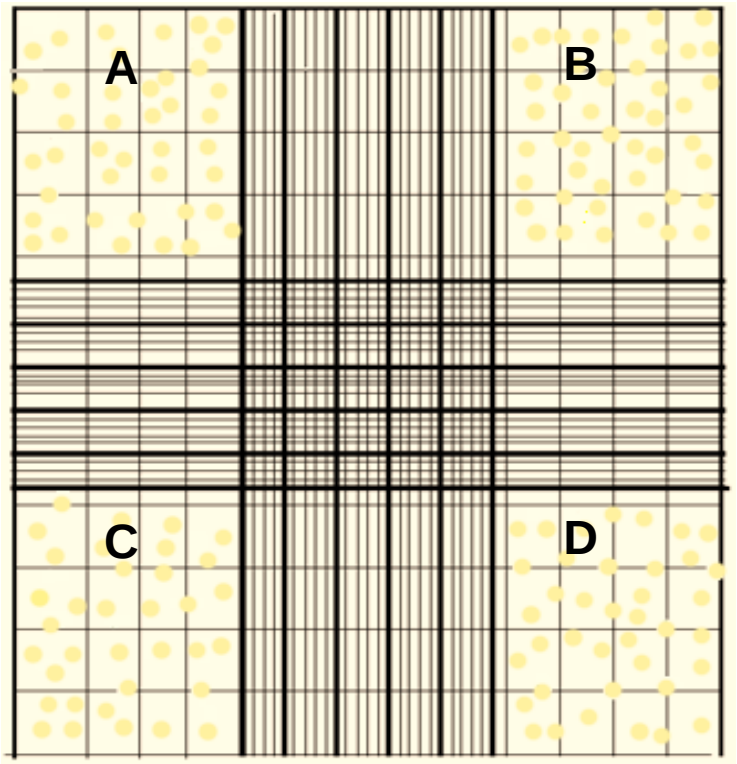


Figure 2-2: Isolation of Peripheral Blood Mononuclear Cells

2.7.2 Cell count using Haemocytometer

For cell counting, an aliquot of cell suspension was mixed 1:1 with 1% acetic acid in order to lyse any contaminating red cells. This mixture was applied carefully into a haemocytometer ensuring each chamber was properly filled. The number of cells was counted in 2×4 fields of counting surfaces. An average number of these fields were multiplied by 2 as dilution factor with acetic acid and by 10^4 because the volume under one square is equivalent to 10^{-4} mL (**Figure 2.3)**.



The average cell number from A, B, C and D was multiplied by a dilution factor of 2 and by 10,000 (10^4) to determine the number of cells per ml

Figure 2-3: Cell count by haemocytometer

Cells were counted from each corner square of the haemocytometer (A, B, C and D) and the average cell number calculated (cells/ml).

2.7.3 Osteoclast cultures

2.7.3.1 Isolation of $CD14^+$ monocytes from PBMCs

CD14-positive ($CD14^+$) monocytes were isolated from PBMCs using anti-CD14-coated beads (Miltenyi Biotec Ltd). The CD14 antigen belongs to the LPS receptor complex and is strongly expressed on most monocytes and macrophages. Osteoclast precursors derive from $CD14^+$ monocytes (Massey and Flanagan 1999; Nicholson, Malakellis et al. 2000). Following purification and count of PBMC as outlined in **sections 2.7.1** and **2.7.2** respectively**,** the PBMCs were washed with ice-cold PBS supplemented with 2mM EDTA (Buffer 1). Cells were centrifuged at 300g for 10min and the cell pellet resuspended in ice-cold PBS supplemented with 2mM EDTA + 0.5% BSA (Buffer 2) containing the magnetic microbeads conjugated with anti-human monoclonal CD14 antibody. PBMCs were resuspended at a cell density of 10^7 cells in 80µL of Buffer 2, supplemented with microbeads at ratio of 4:1. To allow antibody binding, the cell suspension was then incubated for 15 minutes at 4°C. The $CD14^+$ cell population was isolated from the unlabelled $CD14^-$ cells using a MS-column mounted on a magnetic separator.

2.7.3.2 Differentiation of CD14+ monocytes into mature osteoclasts

Following isolation, the $CD14^+$ monocytes were differentiated into mature osteoclasts by stimulation with macrophage colony-stimulating factor (M-CSF) and receptor activator of nuclear factor κB ligand (RANKL). Monocytes were cultured on sterile dentine disks or glass cover slips. Prior to the osteoclast cultures, the dentine disks were punched using a water-cooled Labcut 1010 low-speed diamond saw (Agar Scientific Ltd., Stansted, UK). The cut disks were 5mm in diameter and up 60µm in thickness. Dentine disks were sterilised by sonication for 10 minutes in distilled water repeated twice, Disks were then stored in 70% ethanol for a minimum of 24 hours until required for cell culture. Prior to cell culture procedure, the dentine disks were washed once in PBS and pre-wetted in α-MEM for one hour at 37°C. Glass coverslips with a diameter of 6mm (Richardson's of Leicester, Leicester, UK) were sterilized by dry heating at 175°C for 2 hours. The separated $CD14^+$ cells were counted according to **section 2.7.2** and seeded at a density of 4×10^4 per well on sterile

dentine disks or glass coverslips in the of a 96-well plate. CD14+ monocytes cells were cultured at 37˚C and 7% CO_2 for approximately 2-3 weeks, in α-MEM supplemented with 100 IU/mL penicillin, 100μg/mL streptomycin, 10% FCS, 30ng/mL RANKL and 25ng/mL M-CSF ; referred to as osteoclastogenic media unless otherwise specified. The osteoclastogenic media was replaced every 2–3 days.

2.7.4 Osteoclast Tartrate resistant acid phosphatase staining

Material	
IITEM	**SUPPLIER**
sodium nitrite	Sigma (Poole, UK)
Pararosaniline	Sigma (Poole, UK)
sodium acetate trihydrate	Sigma (Poole, UK)
Gill's haematoxylin	Sigma (Poole, UK)
sodium tartrate dihydrate	Sigma (Poole, UK)
Naphthol AS-BI phosphate	Sigma (Poole, UK)
Hydrochloric acid (10125 4H)	VWR (Lutterworth, UK)
Dimethylformamide	Fisher, (Loughborough, UK)
DPX mounting reagent	BDH Laboratory (Poole, UK)
Acetic acid (10001 CU)	VWR (Lutterworth, UK)

Osteoclast cells which were cultured on disks or coverslips were fixed in 10% formalin and then stained for tartrate resistant acid phosphatase (TRAP). The isoenzyme TRAP stimulates the hydrolysis of phosphoproteins and nucleotides (Hayman, et al 2000) and is characterised as a cytochemical and functional marker of osteoclasts (Baron, et al 1986, Cole and Walters 1987) (Minkin 1982). The TRAP staining assay involves hydrolysation of naphthol AS-BI phosphate by TRAP, which couples with the hexazotized dye, pararosaniline, to produce a coloured complex that precipitates at the site of osteoclast activity.

Fixed osteoclast cultures were treated with 0.1M Sodium tartrate in 0.2M sodium acetate buffer (pH 5.2) for 5 minutes at 37°C. The acetate buffer was removed and the osteoclast cultures incubated in 20mg/mL Naphthol AS-BI phosphate/dimethylformamide acetate-tartrate buffer for 30 minutes at 37°C. Cultures were further treated with acetate-tartrate buffer containing hexazotized pararosaniline solution for 15 minutes at 37°C. Osteoclast cultures were counterstained with Gill's haematoxylin for 30-40 seconds. Excess stain was removed by washing several times with tap water. Coverslips were transferred from the wells in the culture plates and air-dried. Finally, cultures were mounted in DPX reagent and visualised under a light microscope. The dentine disks were also air-dried and stored at room temperature in preparation for quantification using a light microscope.

2.7.5 Osteoclasts quantification

Osteoclast cultures were assessed for total osteoclast number, number of resorbing osteoclasts and percentage resorption using point counting. When osteoclasts are placed on dentine slices they form resorption pits. The presence of functional osteoclast cells can be determined by measuring the amount of resorption on the dentine surface. Osteoclasts were characterised by TRAP positive stained cells with three or more nuclei. Resorbing osteoclasts were characterised as cells on or in close proximity to resorption pits. The dentine disks are TRAP stained and counterstained by Gill's haemotoxylin. A 10x10 point grid was fit into an eye-piece of the microscope. Using 10X magnification, 9 fields could be obtained for each dentine disk. In each field the points of intersection that fall over a resorption pit were counted (2 points for the first pits and 6 points for the second pit, **Figure 2.4**). All the points for all 9 fields were summed to give the measure for resorption for a disk. Other parameters that were counted included the total osteoclast number and the number of resorbing osteoclasts for the entire dentine disk.

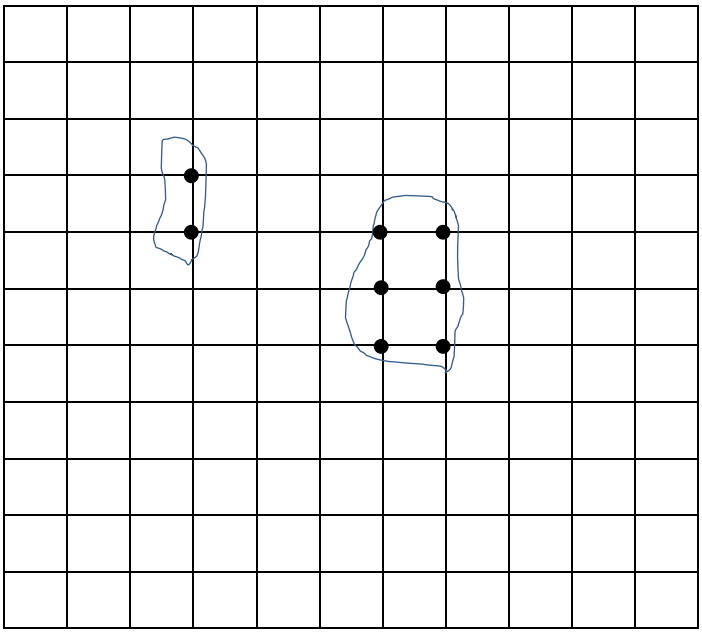


Figure 2-4: Counting the points for quantification resorption area

The points of intersection that fall over a resorption pit were counted

2.8 Osteoclast and Blood RNA Extraction

Material	
ITEM	**SUPPLIER**
ReliaPrep™ RNA Cell Miniprep **ReliaPrep™ Minicolumns (50/pack)** **Column Wash Solution (CWE)** **Elution tubes (50/pack)** **Collection tubes (50/pack)** **BL buffer** **1-Thioglycerol (TG)** **Nuclease-Free Water** **DNase I (lyophilized)** **Yellow Core Buffer** **RNA Wash Solution (RWA)** **MnCl2, 0.09M**	Progema
Microseal® 'B' Adhesive Seals	Bio Rad
Individual PCR Tubes, 0.5 ml and 1.5ml	Thermo-Scientific
384-Well standard PCR Plates	Thermo-Scientific
95% Molecular-grade Ethanol	Sigma-Aldrich
Molecular-grade Isopropanol	Sigma-Aldrich

RNA was extracted from cells cultured on dentine disks or coverslips using the column-based ReliaPrep RNA cell Miniprep kit (Promega). Dentine disks were placed into eppendorf tube containing 100µL of BL-TG buffer. To prepare the BL-TG buffer, 325µL of 1-Thioglycerol (TG) was added to 32.5ml of BL Buffer. RNA was sheared by repeat-pipetting of BL-TG buffer over the surface of the dentine disk. Lysates were pipetted into PCR-grade microcentrifuge eppendorfs and isopropanol was added in 1:3 ratios i.e. 35µL added to the 100µL of BL-TG buffer, followed by thorough mixing by vortex to assist RNA shearing.

Each cell lysate was transferred to a ReliaPrep™ Minicolumn inserted into a fresh collection tube. Lysates were centrifuged at 12,000 × g for 30 seconds at room temperature and the supernatant discarded. Next, 500µL of RNA wash solution containing 95% ethanol was added to the Minicolumn and again centrifuged at 12,000 × g for 30 seconds. To remove any genomic DNA contamination, RNA samples were treated with DNase enzyme. For each sample, 30µL of DNase I solution was prepared by mixing 24µL of Yellow Core buffer, 3µL

0.09M MnCl2 and 3µL of DNase I enzyme. Samples on the Minicolumn were treated with 30ul of the DNase I solution for 15 minutes at room temperature.

Minicolumns were washed with 200µL of column wash solution by centrifugation at 12,000 × g for 15 seconds. Minicolumns were then washed with 500µL of RNA wash solution was by centrifugation at 12,000 RPM for 30 seconds. The Minicolumn membrane was washed again with 300µL of RNA wash solution by centrifugation at 13,000 × g for 2 minutes. Minicolumns were transferred to fresh Elution tubes and 30 µL of nuclease-free water added to each membrane and centrifuged at 13,000 g for 1 minute.

The purity and nucleic acid concentration of RNA was determined using the NanoDrop-1000 spectrophotometer (Thermo scientific) at an absorbance of 260nM (A_{260}) and 280nM (A_{280}). The concentration of 1µL of nucleic acid was determined using the **Beer-Lambert law** $\boldsymbol{A = \varepsilon lc}$ which calculates the difference between light absorbance (***A***), sample concentration (***c***) specific extinction coefficient ($\boldsymbol{\varepsilon}$) and length of the light pathway which travels through the absorber (***l***). An A_{260}/A_{280} ratio of 1.8-2.2 was considered an indicator of uncontaminated RNA.

2.9 Reverse Transcription-Polymerase Chain Reaction (RT-PCR)

Material	
REAGENTS	SUPPLIER
RNAase-free eppendorfs	Alpha Laboratories
Full range100 BP DNA Ladder	Norgen
100mM DNTP mix	Thermo-scientific
RevertAid H Minus First Strand cDNA Synthesis Kit	Thermo-scientific
DEPC-treated Water	Ambion® Life Technologies
Quantity One Software	BioRad
Gel Doc XR+ System	BioRad
Taqman Gene expression Assays	Applied Biosystem
2X Taqman Universal Master Mix	Applied Biosystem
ABI 7900HT Platform	Applied Biosystem
Ethidium Bromide	BioRad
Agarose	Sigma-Aldrich
Tris/Borate/EDTA (TBE) buffer	Applichem
X5 DNA Loading buffer	Kappa Biosystems

2.9.1 Complementary DNA (cDNA) synthesis

The reverse transcription (RT) procedure is based the catalyzing activity of reverse transcriptase enzyme which converts RNA template to complementary DNA (cDNA). The cDNA is required for the polymerase chain reaction (PCR**)** process to evaluate gene expression**.** cDNA synthesis is performed using an oligo(dT), a short piece of single-stranded DNA, which anneals to the 3' end poly-A tail of mRNA molecules. In all of the study experiments, RNA samples were converted to cDNA using the RevertAid H Minus First Strand cDNA Synthesis Kit (Thermo Scientific) following manufacturer's instruction. The kit also includes RevertAid H Minus Reverse Transcriptase that has the ability to eliminate RNase H activity.

Briefly, first strand cDNA synthesis was performed by initially mixing and centrifuging all components of the kit and placing on ice. Prior RT reaction, RNA samples were placed on ice. To reduce risk of RNA contamination, all working environment was sterilised using RNase*Zap*® RNase decontamination solution and tubes and pipettes irradiated in a UV hood for 20 minutes prior to use. The following reagents were added into a sterile PCR grade eppendorf and maintained on ice: Template RNA poly(A) mRNA 1 µg 0.5-1 µg, primer oligo (dT) primer 1 µL, nuclease free water to make a total of 12 µL. RNA samples were then returned to be chilled on ice. Next, 4 µL of 5X reaction buffer, 1 µL of RiboLock RNase Inhibitor (20 u/µL), 2 µL of 10 mM dNTP Mix and 1 µL of RevertAid H Minus M-MuLV Reverse Transcriptase (200 u/µL) were added to the RNA mixture to give a total volume of 20 µL. The mixture was incubated for 60 min at 42°C. The reaction was terminated by heating at 70°C for 5 minutes. To detect any DNA contamination, for each RNA sample a corresponding no-RT sample was also prepared which did not contain RT.

2.9.2 End-point Polymerase Chain Reaction

Polymerase chain reaction (PCR) is the enzymatic amplification of specific DNA sequences. PCR involves multiple cycles of template denaturation, primer annealing, and primer elongation. During denaturation the two cDNA strands are cleaved, requiring a temperature of 94ºC to denature the DNA. In the annealing stage, the primers anneal with their specific target sequences on each of the single stranded DNA templates. Annealing temperatures dependent on the length and guanine/cytosine (G/C) ratio of the primer and can range between 53ºC and 65ºC.

End-point PCR was performed using Promega GoTaq® DNA polymerase. For each 50 µL reaction, the reaction mixture was prepared as below (**Table 2.3**). Product with no cDNA templates were also used for detection contamination in samples. The baseline PCR amplification reactions were pre-incubated at 94°C for 2 minutes, followed by 34 cycles of denaturation at 94°C for 60 seconds, anneal at 55°C for 30 seconds and extension at 72°C for 30 seconds. End-point PCR products were visualized on an ethidium bromide-stained agarose gel. 2g of agarose (Sigma Aldrich) was dissolved by heating in 100ml of 1X

Tris/Borate/EDTA (TBE) buffer to produce a 2% gel. Ethidium bromide (10µL) was mixed into agarose solution and the agarose allowed setting for at least 1 hour at room temperature. The agarose gel was transferred to an electrophoresis tank containing TBE buffer and loaded with 10µl of DNA ladder (full range100 BP Norgen) and 25µL of PCR product containing X5 DNA loading buffer. Electrophoresis was performed for 30 minutes at 100V. DNA fragments were detected by UV light using the Gel Doc XR+ System and the Quantity one software.

Component	Amount per reaction
5X Green GoTaq® Reaction Buffer (1.5mM MgCl2)	10µL
Forward primer (0.2µM)	1µL
Reverse primer (0.2µM)	1µL
dNTP mix (0.2mM)	1µL
GoTaq® DNA Polymerase (5u/µl)	0.25µL
Nuclease free water.	35.75µL
template DNA	1µL

Table 2-6: The PCR settings for each 25µL reaction

2.9.3 Real Time quantitative Reverse Transcription-PCR (qRT-PCR)

Quantitative reverse tanscription PCR (qRT-PCR) exponentially measures products produced during PCR cycles using TaqMan oligonucleotide probes. In principle, the sequence-specific TaqMan probe has a fluorescent probe also known as a Reporter bound to the 5' end, and a 'quencher' molecule attached to the 3' end of the sequence. The probes hybridize to the target sequences in the amplicon during PCR amplification. The close proximity of the reporter and quencher molecules prevents detection of fluorescence by Fluorescence resonance energy transfer (FRET). Uncoupling of the reporter and quencher as a result of polymerase activity results in increases in fluorescence intensity that correlates to the number of the probe cleavage cycles.

Human TaqMan® assays were used to determine *GAPDH* expression. A 7µl reaction mix was prepared per sample by mixing 5µl of TaqMan® gene expression mastermix, 0.5µl TaqMan® GAPDH assay and 1.5µl nuclease-free water. To this reaction, 3µl of cDNA template was added and loaded into a 384 well PCR plate. As a negative control for PCR reaction mix contamination, reaction mix was loaded with 3µl of nuclease-free water instead of cDNA. Optical adhesive covers plates were used to seal PCR plates and PCR performed using the Applied Biosystems 7900HT Real-Time PCR system. Data were analysed using the SDS 2.2.1 software and the measured cycle threshold (CT) values were representative of the number of cycles at which each specific reaction crossed a selected threshold. The baseline of a PCR reaction was set to remove any background fluorescence. High CT values correlated with less cDNA template and an increased number of amplification cycles to reach the fluorescence intensity threshold.

2.9.4 Primer design

As each candidate gene was known to have a different transcript variant, the Ensemble genome browser software (http://www.ensembl.org) was used to identify all of the existing variants. The Cluster omega software (http://www.ebi.ac.uk/Tools/msa/clustalo/) was further used to identify specific sequential regions of an exon shared by all of the transcript variants. The sequences which spanned the exons were preferential for primer design and the Primer 3 software (http://primer3.ut.ee/) was used to design primers. In-Silico PCR (http://genome.ucsc.edu/cgi-bin/hgPcr?command=start) was used to determine whether the designed primers were able to polymerase the specific regions of the cDNA. Primers were designed for determining the expression of candidate genes: *ESR1, TNFRSR11B, LRP5, VDR* and *P2RX7,* with the common part of the transcribed variant of each candidate genes. The lyophilised primers were dissolved in DEPC-treated water when was arrived to the lab to reach a concentration of 10µM suitable for working (Table 2.7).

Candidate Gene/SNP (Rs)	Sequence (5’ -> 3’)	Melting Temperature °C
RANK		
F1	ATCCCCTGGGTGGTTTTATC	59.88
R1	CCAGTGATCCACAGTCATGC	60.12
P2RX7		
F1	GCAGCTGCAGTGATGTTTTC	58.31
R1	CTGCTATCCCCTTCACCTTG	57.66
LRP5		
F1	CCTGGATTTGAACTCGGACT	59.14
R1	TTACAGGGGCACAGAGAAGC	60.40
VDR		
F1	GCCCACCATAAGACCTACGA	59.96
R1	AGATTGGAGAAGCTGGACGA	59.95
ESR1		
F1	ATGATCAACTGGGCGAAGAG	60.22
R1	GATCTCCACCATGCCCTCTA	60.03
OPG		
F1	GGGGACCACAATGAACAACT	59.68
R1	ACACGGTCTTCCACTTTGCT	59.77

Table 2-7: The sequences and the melting temperature for the forward and reverse primers specific to ESR1, TNFRSR11B, LRP5, VDR and P2RX7

2.10 Sequenom

Allele specific expression analysis was performed by MassARRAY ™ platform by Wellcome Trust Centre for Human Genetics service of the Oxford University. The procedure for investigating allele-specific expression was similar to that used when performing gene screens/disease-association SNP, with pooled DNA populations (Herbon, Werner et al. 2003; Tang, Oeth et al. 2004), with the exception that cDNA was used as a template for PCR as opposed to genomic DNA. In fact the ratio of alleles from one group is compared to another to define if any statistically significant difference in allele frequency exists between them. The cDNA of interest was amplified by PCR. MassARRAY Typer version 3.1 software was used for the MassARRAYsystem to achieve quantitative processing for Allele-Specific Expression Analysis.

2.11 Statistical analysis

Graph Pad Prism 6 was used to analyse data. For sample sizes below 50 data was inspected for being consistent with normality and non-parametric statistical hypothesis tests were used if this was not a fair assumption. For comparing the means of two groups, a student's t-test was used. For small samples (below 50) where data were suspected of being non-normal and the student's t-test resulted in a p-value below 0.05 the Mann Whitney test was then used to check that a significant difference remained. If the result then become non-significant this suggested that the conclusions are sensitive to the normality assumption and the evidence is weakened but here, in subsequent chapters when the student's t-test resulted in a p-value below 0.05 in non-parametric data, the Mann Whitney test was also in a p-value below 0.05 . For comparing more than two groups, One- way ANOVA was used. For analysing correlation between two continues variables where at least one was not consistent with the normal distribution, Pearson correlation coefficient was computed. For checking association between SNPs and BMD linear regression analysis by using IBM® SPSS® statistics was performed. Genotype specific effects on BMD was analysed by univariate analysis of variance. Dr Santibanez-Koref kindly adapted his specific software to analyse data of AEI study which is named as MSK-test in this thesis. This software enables to do the analysis (compare the mean allelic expression ratio in the two phases) when haplotypes are unknown. Outliers in the data were removed using the Grubb's outlier test (http://graphpad.com/quickcalcs/Grubbs1.cfm).

Chapter 3 – Genetic SNP association with Bone Mineral Density in an Iranian population

3.1 Hypothesis and Objectives

3.1.1 Hypothesis

There are associations between specific transcribed and risk SNPs and bone mineral density (BMD) within the Iranian population.

To test this hypothesis the following objectives were determined:

3.1.2 Specific objectives

- Determine the BMD of a random selection of subjects from an Iranian population.
- Evaluate the genetic associations between the selected SNPs and BMD.
- Determine whether the genetic associations observed in this Iranian population study are consistent with results from other populations.

3.2 Introduction

Having recently identified 62 genome-wide-significant loci, the genome-wide association studies have now provided clearer insights into the allelic architecture osteoporosis (Richards, Zheng et al. 2012). However, genetic factors linked to BMD are also dependent on sex and age (Karasik and Ferrari 2008). There are limited studies investigating the effect of genetic factors on bone mineral density (BMD) in the Iranian population. The articles which have studied the association between genetic factor and BMD in the Iranian population were searched using PubMed software using the systematic article review, as performed earlier in the thesis as outlined in the **method section 2.2.1**. The search terms were as follows:

1) Genetic association study
2) Iranian
3) bone mineral density

Results of the systemic article review process on NOV 2012 showed in first step five article. Of these, after deleting unrelated articles there were only three that studied the genetic associations of osteoporosis and BMD in the Iranian population (Golbahar, Hamidi et al. 2004; Hossein-nezhad, Varzaneh et al. 2011; Pouresmaeili, Jamshidi et al. 2013). In order to determine whether the genetic factors involved in the pathogenesis of osteoporosis and alterations in BMD within the Iranian population are comparable to that of western populations, genetic studies of the Iranian population must first be performed.

The approaches to identify and characterise genes involved in the pathogenesis of osteoporosis usually involve establishing links between phenotypic characteristics and polymorphic genetic markers in the same way as candidate gene association studies (Stewart and Ralston 2000). The phenotypic characteristic could be a continuous variable such as BMD and the genetic marker could be such as single nucleotide polymorphisms (SNPs).

This chapter consider, the polymorphic variants occurring within seven candidate genes; ESR, LRP5, RANK, RANKL, OPG, P2RX2 and VDR known to be associated with BMD, which is measured using dual energy X-ray absorptiometry as outlined in the **method section 2.1.2**. The selection of these genes and SNPs selection was performed using the systematic review processes outlined in **section 2.2.1** and transcribed SNPs selection outlined in the **method section 2.2.2.** Phenotypic and genetic analysis of the Iranian population sample was

performed using bone densitometry, blood sampling, and genotyping. Per allele and per genotype specific analysis have been used to evaluate association between the candidate SNPs (risk and transcribed) and the BMD of our population.

The IMOS-AD study population consisted of a total of 501 healthy men and women subjects aged between 20 and 75 years of age recruited from the south of Iran. This sample size allowed the statistical evaluation of any genetic association between BMD and these candidates in the Iranian sample. Assuming 0.5 significant level, 500 samples would provide %80 power to detect theory difference in the mean ratio of 0.25 of standard deviation. The main aim of this chapter was to describe the features of the Iranian population and determine whether the genetic associations observed with BMD within this Iranian population study were comparable with other populations (figure 3.1). BMD was selected in preference to fractures in this study because there were not a clear past history in sample population to distinguish osteoporotic fracture from other types of bone fractures.

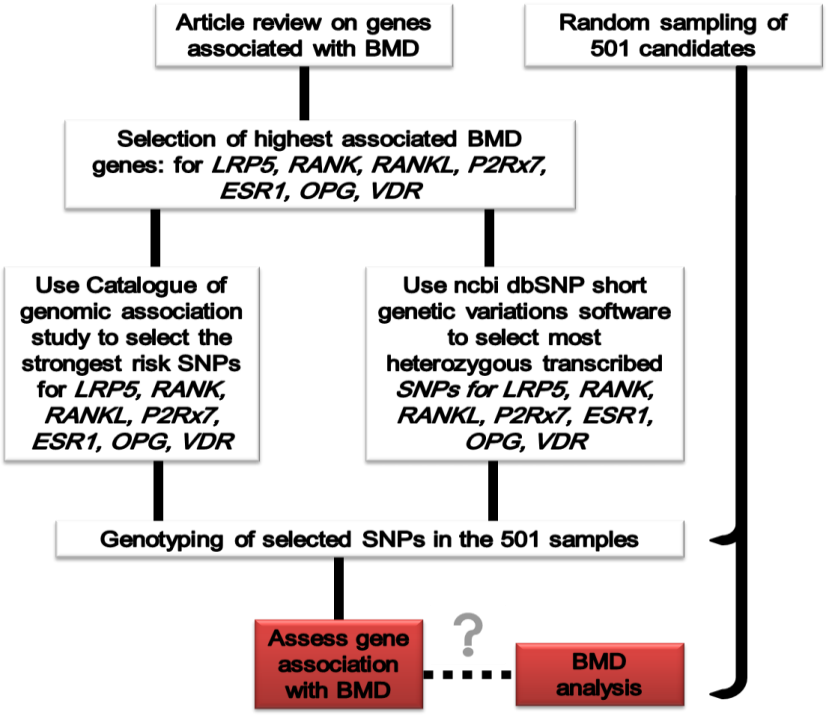


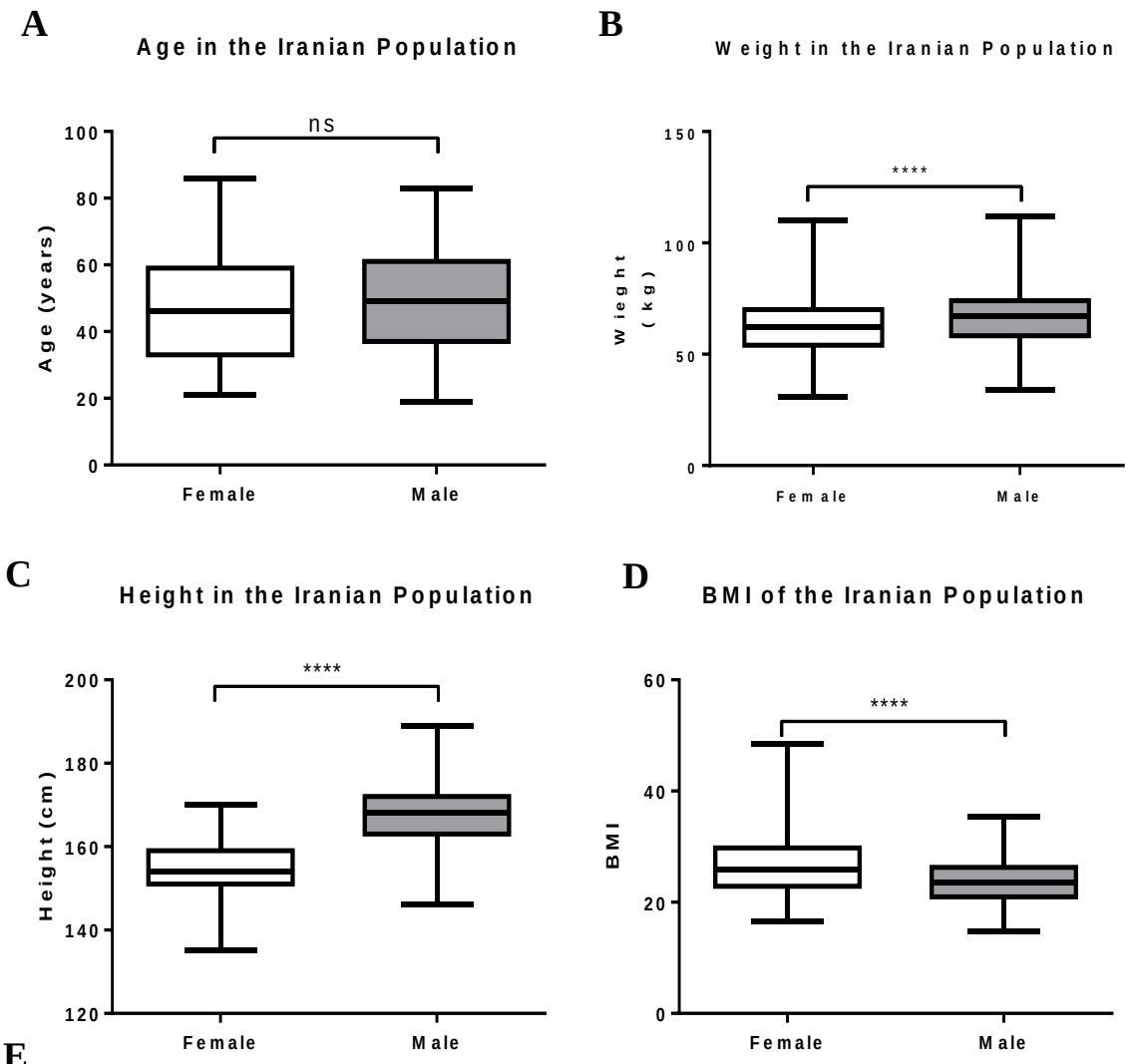
Figure 3-1: Flow chart to illustrate the processes involved in assessing gene associations with BMD

3.3 Results

3.3.1 The BMI of the males and females within a random selection of subjects from an Iranian population

Descriptive statistics and sample quality control of 501 randomly selected Iranians were determined in the first part of this study. Each subject was supplied with a questionnaire from which descriptive analysis factors for this study including age, height, weight and BMI were determined. Data were collected from each subject as specified in the **method section 2.1** and split into male and female. The data were analysed separately by gender as BMI is known to be compatibly higher in females than in males in the Iranian population (Bahrami, Sadatsafavi et al. 2006). Statistical differences between genders were formally compared using a student *t*-test.

Statistical analysis showed that there were no significant differences in age between the male and female subjects of our sample group (**Figure 3.2 A**). Data showed that the male subjects within our sample population had significantly higher weight and height metrics compared to the female subjects (*****P<0.001, Figure 3.2 B, C**). The calculated BMI was higher for females than for the male subjects within the selected Iranian population sample (*****P<0.001, Figure 3.2 D**).



E

Statistical evaluation of the demographic characteristics of female and male subjects				
Variable	**Total**	**Female**	**Male**	
	Mean±SD	**Mean±SD**	**Mean±SD**	**p-value**
AGE	47.48±16	46.7±15.7	48.3±16.3	0.19
HEIGHT	161.2±9.6	154.5±5.8	168.4±7.4	<.001
WEIGHT	64.5±11.8	62.2±11.8	66.9±11.4	<.001
BMI	24.8±4.4	26.05±4.6	23.6±3.67	<.001

Figure 3-2: Physical characteristics of both male and female study subjects

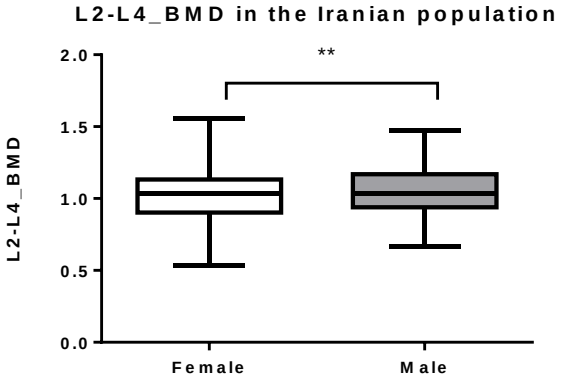
*The age, height, weight and BMI of female and males were determined. (A) There were no gender differences in age. (B, C,) The weight and height of the males within the subject population were higher than that of the female subjects. (D) The male subjects had lower BMI values compared to female subjects. N=501, Student's unpaired t-test. ***P<0.001.*
(box-and-whisker plot: the marked line shows the median, and the bottom and top of the box show the lower and upper quartiles.)

3.3.2 The BMD of IMOS-AD subjects

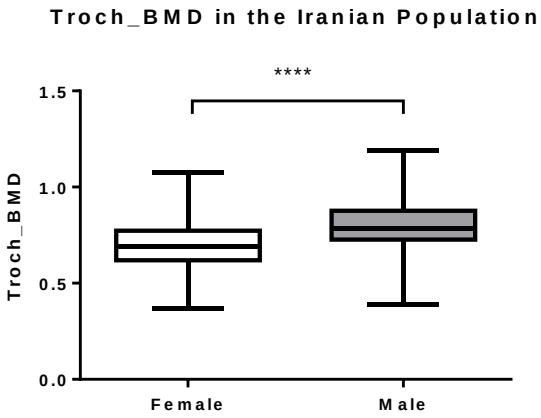
The L2-L4 (L2-L4_BMD), neck of femur (Neck_BMD), femur trochanter (Troch_BMD) and total hip BMD (Total_BMD) of IMOS-AD sample consisting of 501 individuals was assessed using the bone Scanner according to the **method section 2.1.2.** Data from the whole sample population were separated into male and female groups to determine any gender associations with BMD and any statistical relationships determined using a student's *t*-test.

Analysis of the Iranian sample population showed that the male subjects had significantly higher BMD means associated with L2-L4_BMD, neck_BMD, troch_BMD and total_BMD compared to that of the female population (*****P<0.001**, **Figure 3.3, A-D**). The data consisting of the BMD values of all participants were also summarised in **Figure 3.3 E**.

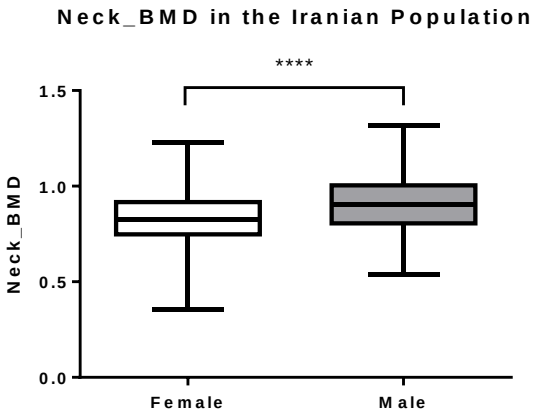
A



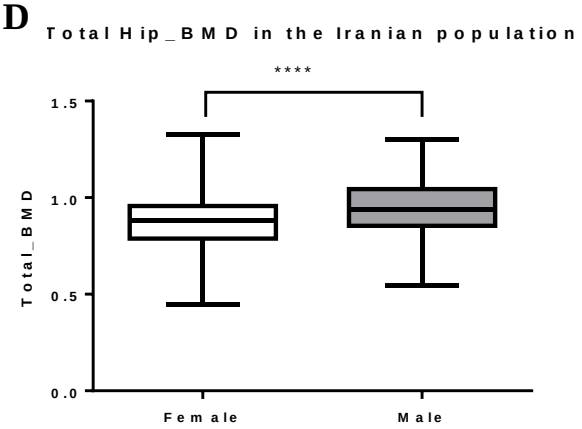
B



C



D



E

BMD of participants				
Variable	**Total**	**Female**	**Male**	
	Mean±SD	**Mean±SD**	**Mean±SD**	**p-value**
L2_L4_BMD	1.031±0.17	1±0.17	1.06±0.15	<.001
NECK_BMD	0.870±0.15	0.827±0.14	0.918±0.14	<.001
TROCH_BMD	0.741±0.14	0.689±0.13	0.798±0.13	<.001
TOTAL HIP_BMD	0.906±0.15	0.864±0.15	0.952±0.13	<.001

Figure 3-3: The total hip BMD was higher in male subjects

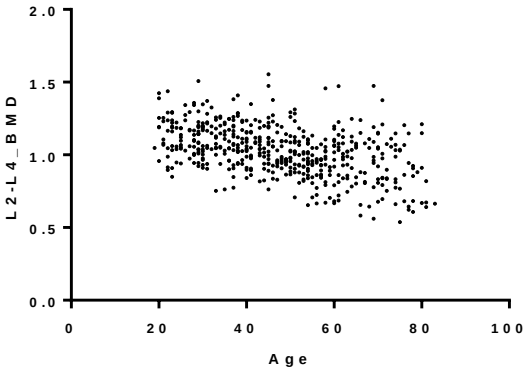
*The L2-L4, Neck, Troch and total hip BMD of female and males were determined. (**A, B, C, D**) There were gender differences in L2-L4, troch, neck, and total hip BMD (**E**) the female subjects had lower BMD values compared to male subjects. N=501, Student's unpaired t-test. **P<0.01,****P<0.001. (box-and-whisker plot: the marked line shows the median, and the bottom and top of the box show the lower and upper quartiles.)*

3.3.3 Correlation between BMD and BMI of a random selection of subjects from an Iranian population

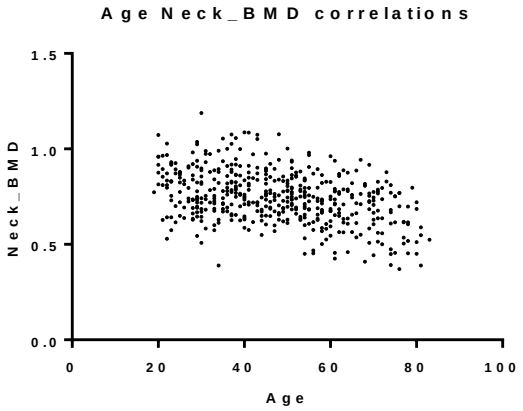
To determine any correlations between BMD and BMI factors which included age, height and weight, statistical analysis was performed using a Pearson's correlation coefficient test. BMD and BMI data collected from **sections 3.3.1** and **3.3.2** were exported into the Graphpad software and analysed for statistical correlations.

There were statistically significantly negative correlations between age with L2-L4_BMD, age with neck BMD, and age with total hip BMD. In addition analysis showed that there were positive correlations between both height and weight with L2-L4 BMD, neck BMD, and total hip BMD (**Figure 3.4**).

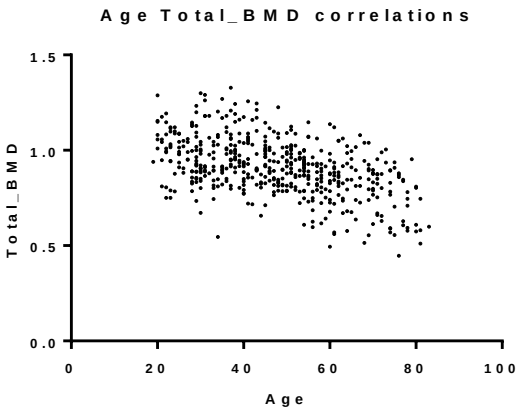
A


Age L2-L4_BMD correlations
L2-L4_BMD
2.0
1.5
1.0
0.5
0.0
0
20
40
60
80
100
Age

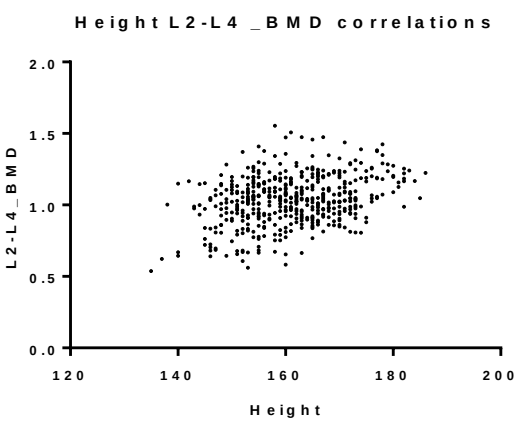
B


Age Neck_BMD correlations
Neck_BMD
1.5
1.0
0.5
0.0
0
20
40
60
80
100
Age

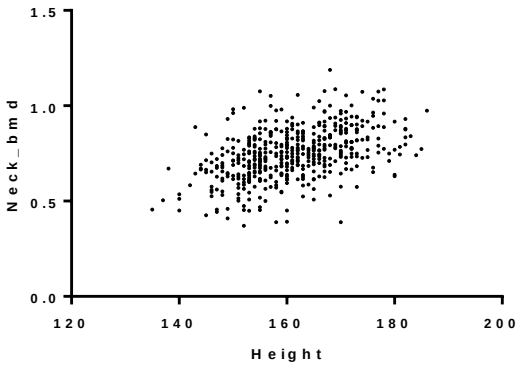
C


Age Total_BMD correlations
Total_BMD
1.5
1.0
0.5
0.0
0
20
40
60
80
100
Age

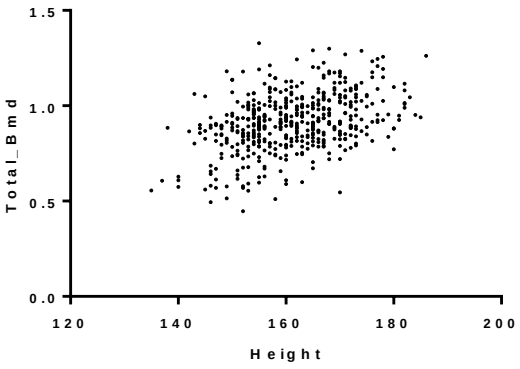
D


Height L2-L4 _BMD correlations
L2-L4_BMD
2.0
1.5
1.0
0.5
0.0
120
140
160
180
200
Height

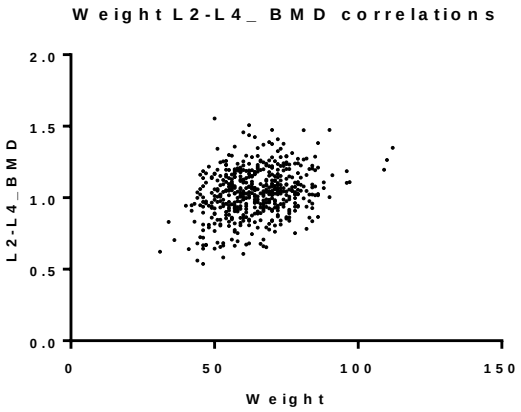
E


Height neck-BMD correlations
Neck_bmd
1.5
1.0
0.5
0.0
120
140
160
180
200
Height

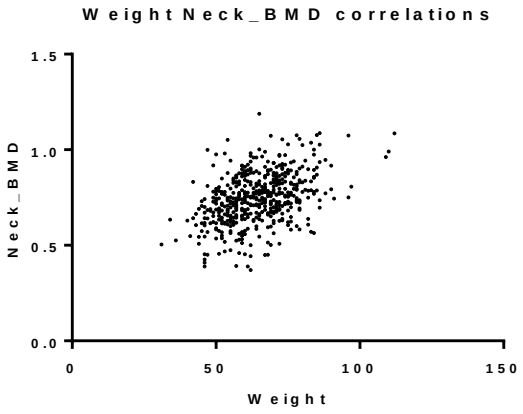
F


Height Total_ BMD correlations
Total_Bmd
1.5
1.0
0.5
0.0
120
140
160
180
200
Height

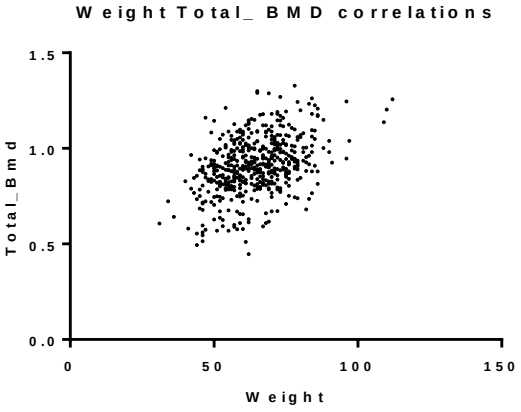
G



H



I



J

Statistical analysis of the correlations between BMD and age, height and weight within the Iranian population study subjects				
Variable		**L2_L4 BMD**	**NECK BMD**	**TOTAL HIP BMD**
AGE	***Correlation Coefficient** **P-value**	**-0.455** **< 0.0001**	**-0.388** **< 0.0001**	**-0.492** **< 0.0001**
HEIGHT	***Correlation Coefficient** **P-value**	**0.332** **< 0.0001**	**0.464** **< 0.0001**	**0.161** **< 0.0001**
WEIGHT	***Correlation Coefficient** **P-value**	**0.085** **< 0.0001**	**0.187** **< 0.0001**	**0.185** **< 0.0001**

Figure 3-4: Correlations were determined between the BMD and age, height and weight of the study subjects

*(**A, B, C**) There were negative correlations between age and L2-L4 BMD, neck BMD, and total hip BMD. (**D, E, F**) There were positive correlations between height and L2-L4 BMD, neck BMD, and total hip BMD. (**G, H, I**) There were positive correlations between weight and L2-L4 BMD, neck BMD, and total hip BMD. (**J**) There were statistically significantly negative correlations between age and L2-L4 BMD, neck BMD, and total hip BMD. The correlations of both height and weight with L2-L4 BMD, neck BMD, and total hip BMD were significantly positive as would be expected. *Pearson's r*

3.3.4 Analysis of genotyping data of seven selected candidate genes and BMD in the Iranian subjects

LRP5, RANK, RANKL, ESR1, VDR, P2XR7 and OPG were selected as genes according to the method outlined in the **method section 2.2.1**. In this section the risk SNPs associated with BMD were determined using the catalogue of GWAS, as outlined in the **method section 2.2.1.** The transcribed SNPs with highest heterozygosity for the *LRP5, RANK, RANKL, P2RX7, ESR1, OPG* and *VDR* genes were determined using NCBI reference assembly as outlined in the **method section 2.2.2.** Data of genotyping, were performed for each SNP as outlined in **section 2.4**., were first qualified as outlined in **section 2.4** and then except data of ten samples and two SNPs which were excluded, other data were used for genotyping specific analysis and per allele specific analysis.

3.3.4.1 Analysis of genotyping data of ESR candidate gene and BMD in the IMOS-AD study

Within the ESR gene, a total of 7 SNPs was identified for this study: Rs2077647, Rs2504063, Rs1999805, Rs4870044, Rs1038304, Rs2941740 and Rs6929137. Genotyping specific analysis determined an existing association between Rs2077647 and total hip BMD (**P-value<0.000, Table 3.1**). Analysis of per allele specific analysis data also showed a significant association between Rs2077647 and L2-L4 BMD (**P-value<0.036, Table 3.1**). No other associations were statistically significant between any of the BMD factors and Rs2504063, Rs1999805, Rs4870044, Rs1038304, Rs2941740 and Rs6929137 SNPs.

ESR1			
SNPs/genotyping	L2-L4 BMD Coefficient (CI)	NECK BMD Coefficient (CI)	Total Hip BMD Coefficient (CI)
Rs 2504063 Baseline genotyping: GG Risk Allele: A Risk allele frequency: 45.62%			
AG/GG	.014(-.019_.046)	0.013(-0.013-0.04)	.014(-.011_.040)
AA/GG	.004(-.035_.042)	-0.04(-0.03-0.03)	-.014(-.044_.017)
P-value	.67	0.403	.128
Per risk allele effect	.003(-.017_.022)	-0.001(-0.02-0.014)	-.006(-.021_.010)
P-value	**.795**	**0.86**	**.465**
Rs 1999805 Baseline genotyping: GG Risk Allele: A Risk allele frequency:46.61%			
AG/GG	-.001(-.033_.031)	.004(-.021_.029)	.010(-.015_.035)
AA/GG	-.010(-.047_.028)	-.004(-.033_.025)	.008(-.022_.037)
P-value	.860	.851	.722
Per risk allele effect	-.005(-.023_.014)	-.002(-.016_.013)	.004(-.010_.019)
P-vaue	**.628**	**.819**	**.567**
Rs 4870044 Baseline genotyping:C:C Risk Allele:T Risk allele frequency:32.26%			
T:C/C:C	.002(-.028_.031)	.003(-.020_.026)	.000(-.023_.023)
T:T/C:C	-.004(-.049_.041)	.000(-.035_.036)	-.001(-.036_.035)
P-value	.970	.967	.999
Per risk allele effect	-.001(-.020_.021)	.001(-.015_.017)	-.000(-.016_.016)
P-vaue	**.935**	**.897**	**.967**
Rs 1038304 Baseline genotyping:A:A Risk Allele:G Risk allele frequency: 57.95%			
G:A/A:A	.023(-.016_.063)	.016(-.016_.047)	.028(-.004_.060)
G:G/A:A	.033(-.009_.075)	.025(-.008_.059)	.035(.001_.069)
P-value	.309	.327	.118
Per risk allele effect	.015(-.005_.036)	.012(-.004_.028)	.016(-.001_.032)
P-vaue	**.144**	**.141**	**.062**
Rs 2941740 Baseline genotyping:A:A			

Risk Allele:G Risk allele frequency: 26.31%			
G:A/A:A	.014(-.016_.044)	.006(-.018_.030)	.015(-.009_.038)
G:G/A:A	.027(-.021_.075)	.002(-.036_.039)	.003(-.035_.041)
P-value	.445	.881	.476
Per risk allele effect	.013(-.007_.034)	.003(-.014_.019)	.006(-.010_.023)
P-vaue	**.203**	**.747**	**.448**
Rs 6929137 Baseline genotyping:G:G Risk Allele:A Risk allele frequency: 38.4%			
A:G/G:G	.028(-.002_.058)	.019(-.005_.043)	.027(.003_.051)
A:A/G:G	.035(-.008_.079)	.008(-.026_.042)	.022(-.013_.056)
P-value	.122	.282	.079
Per risk allele effect	.020(-.000_.041)	.007(-.009_.024)	.015(-.002_.031)
P-vaue	**.054**	**.363**	**.077**
Rs 2077647 Baseline genotyping:T:T Risk Allele:C Risk allele frequency:48.67%			
C:T/T:T	.018(-.015_.051)	.013(-.012_.039)	.027(.001_.053)
C:C/T:T	.041(.003_.079)	.017(-.013_.047)	.024(-.006_.054)
P-value	.109	.478	.000
Per risk allele effect	.020 (.001_.039)	.009 (-.006_.023)	.012 (-.003_.027)
P-value	**.036**	**.258**	**.108**

Table 3-1: SNP genotyping for the ESR1 gene

Genotype specific analysis and per allele specific analysis were performed for each SNP, adjusting for age and sex. Genotype specific analysis showed that Rs2077647 is associated with total hip BMD, P-value=0.000. Per allele specific analysis shows an association between risk allele of Rs2077647 and L2-L4 BMD. N=501, linear regression.

3.3.4.2 Analysis of genotyping data of OPG candidate gene and BMD in the IMOS-AD study

For the OPG gene, a total of 7 SNPs were identified for this study; Rs2073618, Rs1804854, Rs6469804, Rs2062375, Rs4355801, Rs2062377, and Rs11995824. Genotyping specific analysis showed a significant association between Rs4355801 and L2-L4 BMD (**P-value<0.010, Table 3.2**). Per allele specific analysis also showed an association between Rs4355801 and L2-L4 BMD (**P-value<0.029, Table 3.2**).

Both genotype specific analysis and per allele specific analysis also showed an association between Rs11995824 and L2-L4 BMD (**P-value<0.032 and P-value<0.044, Table 3.2).** Additionally **data** produced from per allele specific analysis showed an association between Rs2073618 and L2-L4 BMD (**P-value<0.022, Table 3.2**). Using genotype specific analysis and per allele specific analysis, no other associations were determined between any of the BMD factors and Rs1804854, Rs6469804, Rs2062375, Rs4355801 and Rs2062377 SNPs.

OPG			
SNPs/genotyping	L2-L4 BMD Coefficient (CI)	NECK BMD Coefficient (CI)	Total hip BMD Coefficient (CI)
Rs: 6469804 Baseline genotyping:G:G Risk Allele:A Risk allele frequency:69.87%			
G:A/G:G	-.064(-.110_-.018)	-.040(-.076_-.003)	-.038(-.075_-.001)
A:A/G:G	-.066(-.112_-.021)	-.036(-.072_.000)	-.038(-.074_-.002)
P-value	.013	.099	.099
Per risk allele effect	-.023(.043_-.003)	-.011(-.027_.005)	-.012(.029_.004)
P-vaue	**.026**	**.195**	**.129**
Rs: 4355801 Baseline genotyping:G:G Risk Allele:A Risk allele frequency:70.99%			
G:A/G:G	-.071(-.120_-.022)	-.018(-.057_.021)	-.026(-.065_.013)
A:A/G:G	-.073(-.121_-.025)	-.024(-.063_.014)	-.033(-.072_.005)
P-value	**.010**	.450	.233
Per risk allele effect	-.023 (-.044_-.002)	-.010 (-.026_.007)	-.013 (-.029_.003)
P-vaue	**.029**	**.240**	**.119**
Rs: 2062375 Baseline genotyping:C:C Risk Allele:G Risk allele frequency:33.62%			
G:C/C:C	.009(-.021_.039)	.001(-.023_.025)	.000(-.024_.023)
G:G/C:C	.027(-.017_.070)	.009(-.025_.044)	.007(-.027_.041)
P-value	.473	.857	.908
Per risk allele effect	.012(.008_.032)	.004(-.012-.020)	.002(-.013_.018)
P-vaue	**.235**	**.638**	**.761**
Rs: 2062377 Baseline genotyping:A:A Risk Allele:T Risk allele frequency:28.68%			
A:T/A:A	.006(-.023_.035)	-.004(-.027_.018)	-.001(-.024_.022)
T:T/A:A	.046(-.002_.094)	.026(-.012_.064)	.030(-.009_.068)
P-value	.172	.308	.274
Per risk allele effect	.017(-.004_.037)	.006(-.010_.023)	.009(-.008_.025)
P-vaue	**.118**	**.452**	**.301**
Rs :11995824 Baseline genotyping:C:C Risk Allele:G Risk allele frequency:46.62%			

G:C/C:C	-.058(-.105_-.010)	-.042(-.079_-.004)	-.042(-.080_-.005)
G:G/C:C	-.061(-.107_-.014)	-.037(-.074_.000)	-.041(-.078_-.004)
P-value	**.032**	.086	.071
Per risk allele effect	-.021 (-.042_-.001)	-.011 (-.027_.006)	-.013 (-.029_.005)
P-vaue	**.044**	**.199**	**.125**
Rs: 2073618 Baseline genotyping:C:C Risk Allele:G Risk allele frequency:35.83%			
C:G/C:C	.003(-.027_.032)	.004(-.019_.028)	.000(-.023_.023)
G:G/C:C	.057(.016_.098)	.031(-.002_.063)	.032(-.001_.064)
P-value	.016	.165	.122
Per risk allele effect	.023 (.003_.042)	.013 (-.002_.028)	.012 (-.003_.027)
P-vaue	**.022**	**.099**	**.120**
Rs: 1804854 Baseline genotyping:A:A Risk Allele:G Risk allele frequency:16.23%			
G:A/A:A	.043(-.070_.157)	.009(-.081_.099)	.013(-.078_.104)
P-value	.452	.844	.776
Per risk allele effect	.043(-.070_.157)	.009(-.081_.099)	.013(-.078_.104)
P-vaue	**.452**	**.844**	**.776**

Table 3-2: SNP genotyping for the OPG gene

Genotype specific analysis and per allele specific analysis were performed for each SNP, adjusting for age and sex. Both genotype specific analysis and per allele specific analysis show Rs4355801is associated with L2-L4 BMD P-value=0.010, P-value=.029. There is also an association between Rs11995824 and L2-L4 BMD, P-value=.032 and P-value= .044. There is evidence of per allele specific analysis to show an association between Rs2073618 and L2-L4 BMD, P-value=.022. N=501, linear regression.

3.3.4.3 Analysis of genotyping data of LRP5 candidate gene and BMD in the Iranian population

For the LRP5 gene, a total of 5 SNPs was identified for this study: Rs545382, Rs2306862, Rs556442, Rs599083, and Rs3736228. Per allele specific analysis just showed an association between Rs2306862 and L2-L4 BMD (**P-value<0.042, Table 3.3**). No other associations were determined between any of the BMD factors and Rs545382, Rs556442, Rs599083, and Rs3736228 SNPs.

LRP5			
SNPs/genotyping	L2-L4 BMD Coefficient (CI)	NECK BMD Coefficient (CI)	Total Hip BMD Coefficient (CI)
Rs: 599083 Baseline genotyping:T:T Risk Allele:G Risk allele frequency:32%			
G:T/T:T	-.012(-.042_.018)	.003(-.020_.027)	.003(-.021_.026)
G:G/T:T	-.027(-.066_.013)	.001(-.030_.033)	.005(-.026_.037)
P-value	.400	.961	.943
Per risk allele effect	-.013(-.032_.006)	.001(-.014_.016)	.003(-.012_.018)
P-vaue	**.177**	**.868**	**.732**
Rs: 3736228 Baseline genotyping:C:C Risk Allele:T Risk allele frequency:11%			
T:C/C:C	-.024(-.058_.010)	.016(-.011_.043)	-.046(-.153_.061)
T:T/C:C	-.066(-.199_.068)	-.049(-.155_.057)	.023(-.004_.050)
P-value	.258	.311	.167
Per risk allele effect	-.026(-.057_.005)	.008(-.016_.033)	.014(-.011_.039)
P-vaue	**.103**	**.503**	**.269**
Rs: 545382 Baseline genotyping:T:T Risk Allele:C Risk allele frequency:59.92%			
T:C/C:C	-.030(-.137_.078)	-.077(-.162_.007)	-.074(-.159_.010)
T:T/C:C	-.027(-.132_.078)	-.074(-.156_.008)	-.066(-.149_.016)
P-value	.864	.193	.222
Per risk allele effect	-.002(-.030_.027)	-.007(-.030_.015)	-.003(-.025_.020)
P-vaue	**.914**	**.523**	**.810**
Rs:2306862 Baseline genotyping:C:C Risk Allele:T Risk allele frequency:54.92%			
T:C/C:C	-.025(-.060_.010)	.004(-.024_.032)	.009(-.019_.037)
C:C/C:C	-.086(-.186_.014)	-.021(-.100_.059)	-.031(-.111_.049)
P-value	.107	.830	.587
Per risk allele effect	-.031(-.060_-.001)	-.001(-.024_.023)	.001(-.022_.025)
P-vaue	**.042**	**.966**	**.902**
Rs: 556442 Baseline genotyping:G:G			

Risk Allele:A Risk allele frequency:11.93%			
A:G/G:G	.007(-.032_.047)	-.002(-.033_.029)	-.009(-.040_.022)
A:A/G:G	.023(-.017_.064)	-.007(-.038_.024)	-.013(-.044_.019)
P-value	.428	.873	.727
Per risk allele effect	.013(-.007_.032)	-.004(-.019_.011)	-.006(-.021_.009)
P-value	**.204**	**.618**	**.446**

Table 3-3: SNP genotyping for the LRP5 gene

Genotype specific analysis and per allele specific analysis were performed for each SNP, adjusting for age and sex. Genotyping analysis showed that for the LRP5gene the risk SNP2306862 (highlighted in red) was associated with L2-L4 BMD. N=501, linear regression.

3.3.4.4 Analysis of genotyping data of RANKL candidate gene and BMD in the IMOS-AD study

For the RANKL gene, a total of 4 SNPs was identified for this study: Rs9562415, Rs9594738, Rs9594759, and Rs1021188. Per allele specific analysis as an allele specific analysis showed a significant association between Rs9594738 and L2-L4 BMD (**P-value<0.019, Table 3.4**). Per allele specific analysis also showed an association between Rs9594759 and L2-L4 BMD (**P-value<0.032, Table 3.4**), no other associations were determined between any of the BMD factors and Rs9562415 and Rs1021188 SNPs.

RANKL			
SNPs/genotyping	L2-L4 BMD Coefficient (CI)	NECK BMD Coefficient (CI)	Total Hip BMD Coefficient (CI)
Rs : 9594738 Baseline genotyping: C:C Risk Allele: T Risk allele frequency:25.37%			
T:C/C:C	.029(.000_.058)	-.003(-.026_.020)	-.006(-.029_.017)
T:T/C:C	.046(-.006_.097)	.039(-.001_.079)	.044(.004_.085)
P-value	.061	.133	.063
Per risk allele effect	.025 (.004_.047)	.010 (-.007_.027)	.010 (-.006_.027)
P-vaue	**.019**	.236	.225
Rs : 9594759 Baseline genotyping: C:C Risk Allele: T Risk allele frequency:39.71			
T:C/C:C	.017(-.014_.048)	.006(-.018_.031)	-.001(-.025_.024)
T:T/C:C	.042(.004_.079)	.012(-.017_.042)	.008(-.022_.038)
P-value	.097	.704	.187
Per risk allele effect	.020(.002_.039)	.006(-.008_.021)	.003(-.011_.018)
P-vaue	**.032**	.402	.645
Rs : 1021188 Baseline genotyping: T:T Risk Allele: C Risk allele frequency:14.05%			
T:C/T:T	-.020(-.052_.013)	-.001(-.027_.025)	-.010(-.036_.016)
C:C/T:T	-.038(-.125_.049)	-.068(-.137_.001)	-.070(-.140_-.001)
P-value	.982	.154	.120
Per risk allele effect	-.019(-.046_.008)	-.012(-.034_.009)	-.018(-.040_.003)
P-vaue	.161	.261	.095
Rs :9562415 Baseline genotyping: T:T Risk Allele:C Risk allele frequency:2%			
T:C/T:T	-.009(-.122_.104)	.000(-.088_.089)	-.010(-.100_.080)
C:C/T:T	-.217(-.515_.081)	-.129(-.363_.106)	-.189(-.426_.048)
P-value	.357	.559	.286
Per risk allele effect	-.045(-.135_.045)	-.023(-.094_.048)	-.041(-.113_.031)
P-vaue	.325	.521	.264

Table 3-4: SNP genotyping for the RANKL gene

Genotype specific analysis and per allele specific analysis were performed for each SNP, adjusting for age and sex. Genotyping analysis showed that for the RANKL gene the risk SNP 9594736 and Rs9594759 (highlighted in red) are associated with L2-L4 BMD. N=501, linear regression.

3.3.4.5 Analysis of genotyping data of RANK candidate gene and BMD in the IMOS-AD study

For the RANK gene, a total of 4 SNPs was identified for this study: Rs1805034, Rs35211496, Rs884205, and Rs3018362. Both genotype specific analysis and per allele specific analysis showed an association between Rs884205 and total hip BMD (**P-value<0.042, Table 3.5).** No other associations were determined between any of the BMD factors and Rs1805034, Rs35211496, and Rs3018362 SNPs.

RANK			
SNPs/genotyping	L2-L4 BMD Coefficient (CI)	NECK BMD Coefficient (CI)	Total Hip BMD Coefficient (CI)
Rs : 884205 Baseline genotyping: C:C Risk Allele: A Risk allele frequency:42.62%			
A:C/C:C	-.020(-.049_.010)	-.009(-.032_.015)	-.015(-.038_.009)
A:A/C:C	.019(-.057_.096)	-.060(-.121_.000)	-.061(-.122_.000)
P-value	.342	.129	**.088**
Per risk allele effect	-.009(-.034_.015)	-.016(-.036_.003)	-.020(-.040_-.001)
P-vaue	.464	.101	**.042**
Rs : 3018362 Baseline genotyping: G:G Risk Allele: A Risk allele frequency:31.36%			
A:G/G:G	-.013(-.042_.016)	-.006(-.029_.016)	-.003(-.026_.020)
A:A/G:G	.025(-.025_.076)	.022(-.018_.061)	.022(-.018_.062)
P-value	1.216	.381	.476
Per risk allele effect	-.002(-.024_.019)	-.004(-.021_.013)	-.005(-.022_.012)
P-vaue	.846	.655	.555
Rs:1805034 Baseline genotyping:C:C Risk Allele: T Risk allele frequency:33.11%			
T:C/C:C	.010(-.019_.040)	-.016(-.039_.007)	-.007(-.030_.016)
T:T/C:C	.026(-.021_.072)	.002(-.034_.039)	.003(-.034_.040)
P-value	.523	.319	.786
Per risk allele effect	.012(-.009_.033)	-.005(-.021_.012)	-.001(-.018_.015)
P-vaue	.259	.582	.888
Rs : 35211496 Baseline genotyping: C:C Risk Allele: T Risk allele frequency:9.27%			
T:C/C:C	-.014(-.050_.021)	-.031(-.059_-.004)	-.031(-.059_-.002)
T:T/C:C	-.006(-.178_.167)	.020(-.114_.153)	.032(-.103_.168)
P-value	.731	.079	.089
Per risk allele effect	-.013(.046_-.021)	-.026(-.052_-.000)	-.024(-.050_.002)
P-vaue	.452	.052	.073

Table 3-5: SNP genotyping for the RANK gene

Genotype specific analysis and per allele specific analysis were performed for each SNP, adjusting for age and sex. Genotyping analysis showed that for the RANK gene the risk SNP 884205 (highlighted in red) is associated with total hip BMD. N=501, linear regression.

3.3.4.6 Analysis of genotyping data of P2RX7 candidate gene and BMD in the Iranian population

For the P2RX7 gene a total of 8 SNPs was identified for this study: Rs28360457, Rs2230911, Rs2230912, Rs208294, Rs7958311, Rs1718119, Rs1621388, and Rs3751143. Per allele specific analysis as an allele specific analysis showed a significant association between Rs3751143 and NECK BMD (**P-value<0.024, Table 3.6**). Both genotype specific analysis and per allele specific analysis showed an association between Rs3751143 and total hip BMD (**P-value<0.020 and P-value<0.045, Table 3.6),** no other associations were determined between any of the BMD factors and Rs28360457, Rs2230911, Rs2230912, Rs208294, Rs7958311, Rs1718119, and Rs1621388 SNPs.

P2RX7			
SNPs/genotyping	L2-L4 BMD Coefficient (CI)	NECK BMD Coefficient (CI)	Total Hip BMD Coefficient (CI)
Rs: 3751143 Baseline genotyping: T:T Risk Allele: G Risk allele frequency:23.74%			
T:G/G:G	-.023(-.053_.006)	-.026(-.049_-.003)	-.023(-.046_.000)
T:T/G:G	-.037(-.093_.020)	-.029(-.073_.014)	-.027(-.071_.018)
P-value	.186	.060	**.020**
Per risk allele effect	-.021 (-.043_.002)	-.020 (-.037_-.003)	-.018 (-.035_.000)
P-vaue	.069	**.024**	**.045**
Rs: 28360457 Baseline genotyping: G:G Risk Allele: A Risk allele frequency:5%			
A:G/G:G	.088(-.209_.386)	.078(-.156_.312)	.073(-.164_.310)
P-value	.559	.511	.547
Per risk allele effect	.088(-.209_.386)	.078(-.156_.312)	.073(-.164_.310)
P-vaue	.559	.511	.547
Rs: 2230911 Baseline genotyping: G:G Risk Allele: C Risk allele frequency:5.23%			
C:G/G:G	.041(-.004_.087)	.020(-.015_.056)	.024(-.012_.060)
C:C/G:G	-.211(-.508_.085)	.024(-.209_.258)	.047(-.188_.282)
P-value	.074	.527	.400
Per risk allele effect	.029(-.015_.073)	.020(-.015_.054)	.024(-.011_.059)
P-vaue	.191	.261	.176
Rs: 2230912 Baseline genotyping: A:A Risk Allele: G Risk allele frequency:14.88%			
A:G/A:A	-.002(-.033_.030)	.007(-.018_.032)	.004(-.021_.029)
G:G/A:A	-.049(-.136_.037)	-.012(-.081_.056)	-.018(-.087_.052)
P-value	.534	.783	.830
Per risk allele effect	-.009(-.036_.017)	.003(-.018_.024)	-.000(-.022_.021)
P-vaue	.487	.797	.977
Rs: 208294 Baseline genotyping: G:G Risk Allele: A Risk allele frequency:53.07%			
G:A/G:G	.007(-.028_.041)	.002(-.025_.028)	.002(-.025_.029)
A:A/G:G	-.023(-.060_.015)	-.017(-.046_.012)	-.020(-.049_.010)
P-value	.200	.306	.209
Per risk allele	-.012(-.031_.007)	-.009(-.024_.005)	-.010(-.025_.004)

effect			
P-value	.203	.219	.163
Rs: 7958311 Baseline genotyping: G:G Risk Allele: A Risk allele frequency:21.83%			
G:A/G:G	-.004(-.033_.024)	.012(-.010_.034)	.017(-.006_.039)
A:A/G:G	.037(-.032_.106)	.035(-.019_.088)	.028(-.027_.082)
P-value	.511	.308	.260
Per risk allele effect	.005(-.019_.028)	.014(-.004_.033)	.016(-.003_.034)
P-value	.695	.133	.102
Rs: 1718119 Baseline genotyping: C:C Risk Allele: T Risk allele frequency:36.78%			
C:T/C:C	-.001(-.030_.029)	-.007(-.031_.017)	-.012(-.036_.012)
T:T/C:C	.013(-.029_.055)	-.006(-.040_.027)	-.014(-.048_.020)
P-value	.804	.835	.540
Per risk allele effect	.005(-.015_.024)	-.004(-.020_.012)	-.008(-.024_.008)
P-value	.643	.618	.305
Rs: 1621388 Baseline genotyping: G:G Risk Allele: A Risk allele frequency:37.86%			
G:A/G:G	.004(-.027_.034)	-.004(-.027_.020)	-.009(-.033_.014)
A:A/G:G	.021(-.021_.063)	.007(-.025_.040)	-.001(-.034_.032)
P-value	.620	.789	.725
Per risk allele effect	.009(-.011_.029)	.002(-.014_.017)	-.003(-.018_.013)
P-value	.387	.803	.747

Table 3-6: SNP genotyping for the P2RX7 gene

Genotype specific analysis and per allele specific analysis were performed for each SNP, adjusting for age and sex. Genotyping analysis showed that for the P2RX7 gene the risk SNP Rs3751143 is associated with both neck and total hip BMD. N=501, linear regression.

3.3.4.7 Analysis of genotyping data of VDR candidate gene and BMD in the IMOS-AD study

For the VDR gene two SNPs were identified for this study: Rs731236 (Taq1) and Rs2228570 (Fok1). Per allele specific analysis as an allele specific analysis showed a significant association between Rs731236 and L2-L4 BMD (**P-value<0.038, Table 3.7**). Both genotype specific analysis and per allele specific analysis also showed an association between Rs731236 and NECK BMD (**P-value<0.002 and P-value<0.001, Table 3.7).** Both genotype specific analysis and per allele specific analysis also showed an association between Rs731236 and total hip BMD (**P-value<0.000 and P-value<0.002, Table 3.7)**. No other associations were determined between any of the BMD factors and Rs2228570 SNP.

VDR			
SNPs/genotyping	L2-L4 BMD Coefficient (CI)	NECK BMD Coefficient (CI)	Total Hip BMD Coefficient (CI)
Rs :731236 Baseline genotyping: T:T Risk Allele: C Risk allele frequency:32.67%			
T:C/T:T	.022(-.007_.050)	.037(.015_.059)	.037(.014_.060)
C:C/T:T	.045(-.003_.092)	.044(.007_.081)	.050(.012_.087)
P-value	.116	**.002**	**.001**
Per risk allele effect	.022 (.001_.043)	.027(.011_.044)	.029(.013_.046)
P-vaue	**.038**	**.001**	**.000**
Rs 2228570 Baseline genotyping: T:T Risk Allele: C Risk allele frequency:73.44%			
T:C/T:T	.017(-.038_.072)	.013(-.030_.056)	-.028(-.071_.015)
C:C/T:T	.031(-.023_.084)	.022(-.020_.064)	-.016(-.058_.027)
P-value	.430	.507	.362
Per risk allele effect	.014(-.007_.036)	.010(-.007_.027)	.001(-.016_.018)
P-vaue	**.195**	**.248**	**.918**

Table 3-7: SNP genotyping for the VDR gene

Genotype specific analysis and per allele specific analysis were performed for each SNP, adjusting for age and sex. Genotyping analysis showed that for the VDR gene, the risk SNP 731236 (highlighted in red) is associated with L2-L4, neck, and total hip BMD. N=501, linear regression.

3.3.4.8 A summary of significant associations determined between BMD and selected SNPs in this study

List of SNPs with significant P-value in genotype specific analysis and per allele analysis						
	Genotype specific analysis P-value			Per Allele analysis P-value		
SNPs	**L2-L4**	**NECK**	**Total**	**L2-L4**	**NECK**	**Total**
VDR						
Rs:731236	.116	**.002**	**.001**	**.038**	**.001**	**.000**
T:C/T:T	.022(-.007_.050)	.037(.015_.059)	.037(.014_.060)	.022 (.001_.043)	.027(.011_.044)	.029(.013_.046)
C:C/T:T	.045(-.003_.092)	.044(.007_.081)	.050(.012_.087)			
P2RX7						
Rs:3751143	.186	.060	**.020**	.069	**.024**	**.045**
T:G/G:G	-.023(-.053_.006)	-.026(-.049_-.003)	-.023(-.046_.000)	-.021 (-.043_.002)	-.020 (-.037_-.003)	-.018 (-.035_.000)
T:T/G:G	-.037(-.093_.020)	-.029(-.073_.014)	-.027(-.071_.018)			
RANK						
Rs:884205	.342	.129	.088	.464	.101	**.042**
A:C/C:C	-.020(-.049_.010)	-.009(-.032_.015)	-.015(-.038_.009)	-.009(-.034_.015)	-.016(-.036_.003)	-.020(-.040_-.001)
A:A/C:C	.019(-.057_.096)	-.060(-.121_.000)	-.061(-.122_.000)			
RANKL						
Rs:9594738	.061	.133	.063	**.019**	.236	.225
T:C/C:C	.029(.000_.058)	-.003(-.026_.020)	-.006(-.029_.017)	.025 (.004_.047)	.010 (-.007_.027)	.010 (-.006_.027)
T:T/C:C	.046(-.006_.097)	.039(-.001_.079)	.044(.004_.085)			
Rs:9594759	.097	.704	.187	**.032**	.402	.645
T:C/C:C	.017(-.014_.048)	.006(-.018_.031)	-.001(-.025_.024)			
T:T/C:C	.042(.004_.079)	.012(-.017_.042)	.008(-.022_.038)			
OPG						
Rs:2073618	**.016**	.165	.122	**.022**	.099	.120
C:G/C:C	.003(-.027_.032)	.004(-.019_.028)	.000(-.023_.023)	.023 (.003_.042)	.013 (-.002_.028)	.012 (-.003_.027)
G:G/C:C	.057(.016_.098)	.031(-.002_.063)	.032(-.001_.064)			
Rs:11995824	**.032**	.086	.071	**.044**	.199	.125
G:C/C:C	-.058(-.105_-.010)	-.042(-.079_-.004)	-.042(-.080_-.005)	-.021 (-.042_-.001)	-.011 (-.027_.006)	-.013 (-.029_)

G:G/C:C	-.061(-.107_-.014)	-.037(-.074_.000)	-.041(-.078_-.004)			
Rs:4355801	**.010**	.450	.233	**.029**	.240	.119
G:A/G:G	-.071(-.120_-.022)	-.018(-.057_.021)	-.026(-.065_.013)	-.023 (-.044_-.002)	-.010 (-.026_.007)	-.013 (-.029_.003)
A:A/G:G	-.073(-.121_-.025)	-.024(-.063_.014)	-.033(-.072_.005)			
ESR1						
Rs:2077647	.109	.478	**.000**	**.036**	.258	.108
C:T/T:T	.018(-.015_.051)	.013(-.012_.039)	.027(.001_.053)	.020 (.001_.039)	.009 (-.006_.023)	.012 (-.003_.027)
C:C/T:T	.041(.003_.079)	.017(-.013_.047)	.024(-.006_.054)			
LRP5						
RS:2306862	.107	.830	.587	**.042**	.966	.902
T:C/C:C	-.025(-.060_.010)	.004(-.024_.032)	.009(-.019_.037)	-.031(-.060_-.001)	-.001(-.024_.023)	.001(-.022_.025)
T:T/C:C	-.086(-.186_.014)	-.021(-.100_.059)	-.031(-.111_.049)			

Table 3-8: Summary of associations between BMD and SNPs

3.3.4.9 A summary of genetic BMD associations study in the Iranian population and current study

Hossein-Nezhad et al 2009 showed there is a significant association between different genotypes of rs2228570 and spine BMD in premenopausal women CC (1.20± 0.17), CT (1.14± 0.05), and TT (1.03± 0.13) with P-value=0.001and also postmenopausal women CC (1.20± 0.17), CT (1.14± 0.05), and TT (1.03± 0.13) with P-value=0.001. Hossein-Nezhad et al 2009 also showed there is a significant association between different genotypes of rs2228570 and total hip BMD in premenopausal women CC (1± 0.1), CT (0.96± 0.05), and TT (0.90±0.12) with P-value=0.001 but there is no association in postmenopausal women. The present study showed that there is no association between rs2228570 and BMD and revealed there is an association between rs731236 of VDR and spine 0.022 (95%CI 0.001, 0.43; p = 0.038), neck BMD 0.027(95%CI 0.011, 0.44; p = 0.001), and total hip BMD 0.029(95%CI 0.013, 0.46; p = 0.000). Golbahar et.al 2004 showed there is no association between rs1801133 and both spine and femoral neck BMD in post-menopausal women (Table 3.9).

Genetic BMD associations study in the Iranian population					
Author	**Sample size**	**Sex**	**Candidate gene/ polymorphism**	**Phenotypes**	**P-value**
(Hossein-nezhad, Varzaneh et al. 2011)	312	Female Pre-menopausal	VDR gene Rs2228570 CT/CC vs. TT	Spine BMD	0.001
				Total hip BMD	0.001
		Female Post-menopausal		Spine BMD	0.001
				Total hip BMD	0.50
(Golbahar, Hamidi et al. 2004)	271	Female Post-menopausal	MTHFR* gene Rs1801133 TC/TT vs. CC	Spine BMD	0.806
				Femoral neck BMD	0.802
Current study	501	Male &Female	VDR gene Rs2228570 CT/CC vs T:T	Spine BMD	0.195
				Neck BMD	0.248
				Total hip BMD	0.918
			VDR gene Rs731236 CT/CC vs T:T	Spine BMD	0.116
				Neck BMD	0.002
				Total hip BMD	0.001

Table 3-9: Summary of associations between BMD and SNPs in the Iranian population

Data of the all genetic association studies between BMD and SNPs in the Iranian population (according to the systematic review explained in the introduction part) were summarized and compared with data of available same candidate gene in the current study. It is showed that the result of association study between rs2228570 and BMD in one of those papers was not the same as in the current study. *methylene tetra hydro folate reductase*

3.4 Discussion

There is a multitude of factors that can influence the risk of developing osteoporosis, including age, gender, physical activity levels, diet and medication. However, it has been established that one of the key clinical risk factors of osteoporosis is a positive family history, highlighting the significance of genetics in the aetiology of osteoporosis (Li, Hou et al. 2010). Currently, BMD is the most commonly used predictor of osteoporotic fractures, and as a result is often used as a marker phenotype for osteoporosis (Li, Hou et al. 2010). This chapter looked at whether the BMD of a random selection of subjects from an Iranian population was genetically associated with a selection of candidate genes known to be associated with BMD. Next, the most associated SNPs for the *LRP5, RANK, RANKL, P2RX7, ESR1, OPG* and *VDR* genes were selected for genotyping, and any genetic associations between the specific SNPs and BMD were evaluated. The final objective of this part of the study was to determine whether the genetic associations observed in our Iranian population study were consistent with other demographic populations. Therefore we hypothesise that there are genomic associations between specific SNPs and BMD within the Iranian population.

Differences in sex, age and BMI are all possible causes of slight differences in findings between studies (Johnson et al. 2009; Shen et al. 2005), which also give rise to variations in genetic association studies. Producing a descriptive analysis of a sample group prior to genetic studies is therefore important for determining genetic associations. In this study the descriptive analysis of our random selection of subjects from the Iranian population showed that there were no significant differences between the mean ages of the males and females within the group. Additionally the weight and height of the males were higher than that of the females, as observed in similar subject groups of other similar Iranian studies (Bahrami, Sadatsafavi et al. 2006). These data are an indication that as the descriptive characteristics of the randomly selected sample group was comparable to samples within other studies; further genetic studies would also be comparable to the literature. Interestingly data showed that our male subjects had a lower BMI range compared to the female subjects and these data were consistent with findings from Bahrami *et al* (Bahrami, Sadatsafavi et al. 2006).

Data collected from the BMD analysis of our Iranian population samples showed that the total hip BMD of the male subjects was higher than the female subjects. The average BMD of our male sample group was 0.95 mg/cm^2 and the females had an average BMD value of 0.86

mg/cm2. These data were consistent with BMD data obtained from other Iranian groups (Larijani, Moayyeri et al. 2006),(Omrani 2006) European groups (Styrkarsdottir, Halldorsson et al. 2010) and American groups (Mazess and Barden 1999).These data suggest that the BMD pattern observed within our sample group was comparable to that of others used in similar genetic association studies (Kiel, Ferrari et al. 2007) (Shang, Lin et al. 2013). As a result this increases the viability of data obtained from future genetic association studies.

As mentioned previously, a low BMD is associated with the risk of developing osteoporosis (Cummings, Nevitt et al. 1995). The relationship between BMD and BMI, including its weight and height constituents is widely reported (Felson, Zhang et al. 1993; Ravn, Cizza et al. 1999; Nguyen, Center et al. 2000). In the literature BMD is reported to be lower in lean postmenopausal women in the majority of studies. However, some studies have also reported increased BMD in lean postmenopausal women (Salamat, Salamat et al. 2013). Analysis to determine the existence of any correlation between BMD and BMI constituents in our random selection of Iranian subjects showed that correlations of height and weight with L2-L4 BMD, neck BMD, and total hip BMD were significantly positive. This was not unexpected, and it was shown by other researchers (Munasinghe, Botea et al. 2002; Mascarenhas, Negreiro et al. 2003; Lei, Deng et al. 2004; Yeh, Chen et al. 2004). There were also significant negative correlations between age and BMD, which were consistent with other reports (Nguyen, Center et al. 2000).

It is expected that many different genetic variants contribute to the regulation of BMD. Most of the variants are associated with small effect size, but there is evidence that rare variants of large effect size also contribute in some individuals (Ralston 2010). In the past 10 years there have been advances in the identification and validation of osteoporosis susceptibility loci via large-scale association studies, meta-analyses and GWAS of SNPs. A thorough review of latter developments has revealed that more than 15 genes can be assigned as confirmed BMD genes linked to osteoporosis (Li, Hou et al. 2010). The *LRP5, RANK, RANKL, P2RX7, ESR1, OPG* and *VDR* genes selected for this study are known to have important roles in four well defined biological pathways; the oestrogen endocrine pathway, the Wnt/β-catenin signalling pathway, the RANKL/RANK/OPG pathway and the vitamin D synthesis pathway (Li, Hou et al. 2010). There are limited data related to effective genetic variants in osteoporosis, particularly in respect to the mentioned genes, in the Iranian population. Genotyping specific analysis and allele specific analysis were performed to determine

significant associations between BMD and our 37 selected SNPs (**section 3.3.4**). Associations were calculated between each of the 37 SNPs and L2-L4 BMD, NECK_BMD and Total_ BMD. For this association analysis, a total of 37 x 6 (222) hypothesis tests were performed. Therefore, due to repeated testing we would expect to find 11 of the tests to be significant under the null by chance (P<0.05). Genotype specific analysis and per allele specific analysis of the BMDs of our Iranian population sample and selected SNPs showed that 19 associations were P<0.05 (**Table 3.10**). Hence we do see some evidence of associations with this set of genetic loci in the Iranian sample. However, the SNPs selected for analysis in this cohort are already established as risk SNPs so in this evaluation we are mainly interested to know if the overall pattern of association is similar to other populations or may be different. Such differences may be due to different population, alleles, or the impact of a distinct environment; hence it is more the estimate and 95%CI that is of interest.

Number of significant P-value for association between SNPs and BMD in both genotype and per allele specific analysis models				
Analysis	**L2-L4 BMD**	**Neck BMD**	**Total Hip BMD**	**Total number of associations**
Univariate Analysis	2/37	1/37	4/37	
Per allele specific analysis	7/37	2/37	3/37	
Total	9/74	3/74	7/74	**19/229**

Table 3-10: BMD and SNP association analysis

Genotype specific analysis and per allele specific analysis were performed to determine significant genetic associations between each of the 37 selected SNPs and BMD parameters. Genotyping data showed a total of 19 significant associations between selected SNPs and the BMD of our Iranian population sample. N=501, linear regression, P<0.05.

Out of the 37 selected SNPs our genotyping data identified associations between BMD and 8 of these SNPs: Rs731236, Rs3751143, Rs884205, Rs9594738, Rs2073618, Rs4355801, and Rs2077647 (**Table 3.11**). Genotyping analysis of SNPs belonging to the ESR gene showed a

significant association between BMD (L2-L4 and Total) and RS 2077647 in our Iranian population sample. To date associations between SNP RS 2077647 and BMD have not been demonstrated in any Iranian population prior to this study. Interestingly findings from Sonoda *et al* were consistent with our data, in that they showed a significant increase in BMD with minor alleles of rs2077647 located in the first exon of ESR1 (Sonoda, Takada et al. 2012). On the other hand in 2009 Rivadeneira *et al* group performed meta-analysis study of five genome-wide association studies of lumbar spine and femoral neck BMD. This group identified 20 BMD loci that reached genome-wide significance level (GWS; $P<10^{8}$), where ESR1 was one of the known BMD loci. In comparison to the genotyping data obtained from our select Iranian sample, Rivadeneira *et al* showed rs2504063 (P-value=$6.1E^{-11}$) and rs2941740 (P-value=$2.0E^{-09}$) were associated with lumbar and femoral BMD (Rivadeneira, Styrkarsdottir et al. 2009). In contrast, Holliday *et al* findings showed associations of rs1999805 not only with femoral BMD and spinal BMD but also with calcaneus BMD 0.10 (95%CI 0.05, 0.16; p = 0.0001) and trabecular BMD at the distal radius 0.18 (95%CI 0.06, 0.29; p = 0.003) (Holliday, Pye et al. 2011).

In another large GWAS analysis between SNPs most implicated in the discovery set and BMD of the lumbar spine and hip was performed in Icelandic (4169), Danish (2269), and Australian (1491) subjects as replication sets (Styrkarsdottir, Halldorsson et al. 2008). In this study sequence variants in five genomic regions were significantly associated with BMD in the discovery set and in the replication sets. The *ESR1* gene was one of the five genomic regions which: rs1999805: −0.11 (95%CI −0.15, −0.06), rs6929137: −0.10 (95%CI −0.15, −0.05), rs1038304: −0.10 (95%CI −0.14, −0.06), and rs4870044: −0.11 (95%CI −0.16, −0.07) of this locus were associated with BMD (Styrkarsdottir, Halldorsson et al. 2008).

For the RANK gene, the present study showed that SNP rs884205 was associated with total hip BMD, and the same association was observed in studies by Styrkarsdottir et al within a European and East Asian population (Estrada, Styrkarsdottir et al 2012; Rivadenier, Styrkarsdottir et al 2009). The present study further showed that RANKL SNPs rs9594738 and rs9594759 were also linked closely to spine BMD, similar to studies by Duncan et al and Styrkarsdottir et al in the European population (Duncan, Danoy et al 2001; Styrkarsdottir, Halldorsson et al 2008). The OPG SNP rs2073618, rs1995824 and rs4355801 associations with spine BMD determined in the present were also observed in other studies specific to spine and hip BMD in post/pre-menopausal women and the European population respectively

(Mencej-Bedrač, Preželj et al. 2011; (Shang, Lin et al. 2013). The ESR1 SNP 2077647 associated with spine BMD in the present study was also associated with post-menopausal osteoporosis in Japanese women in a recent study (Sonoda, Takada et al 2012). LRP5 SNP rs2306862 was associated with BMD in men and women in the present study as well as in a larger study carried out by Kiel et al. Interestingly, where we observed an association with spine BMD, the Kiel study shown an association of rs2306862 with femoral neck BMD in men only. Although there are few data on P2RX7, one study did correlate with ours in that SNP rs3751143 was associated with BMD in a similar number of subjects (Wesselius, Bours et al. 2013). From our data we infer that BMD is genetically dependent and these findings correlated with the literature (Arden, Baker et al. 1996).

Genetic BMD associations worldwide compared to the present Iranian study					
Author	**Sample size**	**sex**	**Gene/ Polymorphism**	**Phenotypes**	**P-value Coefficient (CI)**
RANK					
Dastgheib et al	501	Male & Female	Rs884205 A-allele	Total hip BMD	0.042 -.020(-.040_-.001)
(Estrada, Styrkarsdottir et al. 2012)	Initial: up to 32,961 European and East Asian ancestry individuals Replication: Up to 50,933 European and East Asian ancestry individuals	Male & Female	Rs884205 A-allele	Total hip BMD	2×10^{-17}* .05 (NA)
(Rivadeneira, Styrkarsdottir et al. 2009)	19,195 European ancestry individuals	Male & Female	Rs884205 A-allele	Spine BMD	9×10^{-9}* .08 (0.05-0.11)
RANKL					
Dastgheib et al	501	Male & Female	Rs:9594738 T-Allele	Spine BMD	0.019 .025 (.004_.047)
(Duncan, Danoy et al. 2011)	Initial: European ancestry 900 cases with low hip BMD, 1,055 cases with high hip BMD Replication: 20,898 European ancestry	Female	Rs:9594738 T-Allele	Spine BMD	4×10^{-6}* 0.15 (NA)
(Styrkarsdottir, Halldorsson et al. 2008)	Initial:5,861 European ancestry individuals Replication: 7,925 European ancestry individuals	Male & Female	Rs:9594738 T-Allele	Spine BMD	2×10^{-8}* 0.1 (0.06-0.13)
Dastgheib et al	501	Male & Female	Rs:9594759 T-Allele	Spine BMD	.032 .020(0.002_.039)

(Styrkarsdottir, Halldorsson et al. 2009)	Initial:6,865 European ancestry individuals Replication: 8,510 European ancestry individuals	Male & Female	Rs:9594759 T-Allele	Spine BMD	2 x10^{-17*} .12 (.09-.15)
(Styrkarsdottir, Halldorsson et al. 2008)	Initial: 5,861 European ancestry individuals Replication: 7,925 European ancestry individuals	Male & Female	Rs:9594759 T-Allele	Spine BMD	2 x10^{-21*} .17 (0.14-0.21)
OPG					
Dastgheib et al	501	Male & Female	Rs:2073618 G Allele	Spine BMD	.022 .023 (.003_.042)
(Mencej-Bedrač, Preželj et al. 2011)	478 post-menopausal women of Slovenia	Female	Rs:2073618 C:G vs C:C	Spine BMD	.006 NA
(Shang, Lin et al. 2013)	108 pre-menopausal and 127 post-menopausal women	Female	Rs:2073618 C:G/G:G vs C:C	Total hip	<0.05 NA
Dastgheib et al	501	Male & Female	Rs:11995824 G Allele	Spine BMD	0.044 -.021 (-.042_-.001)
(Rivadeneira, Styrkarsdottir et al. 2009)	19,195 European ancestry individuals	Male & Female	Rs:11995824 G Allele	Hip BMD	7 x10^{-9*} .07 (0.05-0.09)
Dastgheib et al	501	Male & Female	Rs:4355801 A Allele	Spine BMD	0.029 -.023 (-.044_-.002)
(Richards, Rivadeneira et al. 2008)	Initial:2,094 European ancestry women Replication: 6,463 European ancestry individuals	Female	Rs:4355801 A Allele	Femoral neck and spine BMD	8x10^{-10*} .09 (NA)
ESR1					
Dastgheib et al	501	Male & Female	Rs:2077647 C Allele	Spine BMD	0.036 .020 (.001_.039)

(Sonoda, Takada et al. 2012)	Cases were 114 Japanese women with a confirmed diagnosis of postmenopausal osteoporosis and controls were 171 general Japanese women	Female	Rs:2077647 C:T vs T:T	Osteoporosis	<0.0001 3.27 (1.83–5.85)
			Rs:2077647 C:C vs T:T		0.015 2.48 (1.20–5.16)
LRP5					
Dastgheib et al	501	Male & Female	RS:2306862 T Allele	Spine BMD	0.042 -.031(-.060_-.001)
(Kiel, Ferrari et al. 2007)	subset of 1797 unrelated individuals	Men	RS:2306862 C:T/T:T vs C:C	Neck BMD	0.01 NA
P2RX7					
Dastgheib et al	501	Male & Female	Rs:3751143 T Allele	Neck BMD	0.024 -.020 (-.037_-.003)
				Total hip BMD	0.045 -.018 (-.035_.000)
(Wesselius, Bours et al. 2013)	Six hundred ninety women and 231 men aged ≥50 years	Male & Female	Rs:3751143 T:T vs G:G	Hip BMD	0.018

Table 3-11 : Summary of genetic BMD associations worldwide compared to the present Iranian study

*Significant associations determined between BMD and selected SNPs in this study were compared with other populations. *P-value for genome wide association studies*

In addition to age and BMI, the sample size and ethnicity issue are possible causes of inconsistent findings between one study and another (Johnson et al. 2009; Shen et al. 2005). Although many association studies, including the present study, have used sample sizes ranging between 500-600 samples (Wesselius, Bours et al. 2013), there are obvious benefits to using a larger sample size group. For example, GWAS use large sample size groups in their studies, with replications in a range of ethnic groups (Estrada, Styrkarsdottir et al. 2012). Some of the genetic associations that were observed between BMD and specific SNPs in our

Iranian sample differed from the associations observed in other populations. This could be due to a number of underlying reasons, including variations in population parameters such as age, weight and height, and lack of power.

Taking into account the extensive set of possible genetic associations investigated worldwide, the most important findings represent an extreme sample (Hong and Park 2012). Usually in any disease genetic associations are modest and single studies do not have the power to detect these associations effectively (Ioannidis, Ntzani et al. 2001).

Finally the main findings from this chapter are as follows:

Analysis of the Iranian sample population showed that the male subjects had significantly higher BMD means associated with L2-L4_BMD, neck_BMD, troch_BMD and total_BMD compared to that of the female population.

There were statistically significantly negative correlations between age with L2-L4_BMD, age with neck BMD, and age with total hip BMD. In addition analysis to determine the existence of any correlation between BMD and BMI constituents in our random selection of Iranian subjects showed that correlations of height and weight with L2-L4 BMD, neck BMD, and total hip BMD were significantly positive.

Genotyping data analysis identified associations between BMD and risk SNPs: VDR-rs731236, P2RX7-rs3751143, RANK-rs884205, RANKL-rs9594738, RANKL-rs9594759, LRP5-rs2306862, OPG-rs2073618, OPG-rs11995824, OPG-rs4355801, and ESR1-rs2077647. From our data we infer that BMD is genetically dependent. Interestingly, these finding are similar to results from genome wide association studies in other population.

Chapter 4 – Role of genetic variants in osteoclast function

4.1 Hypothesis and Objectives

4.1.1 Hypothesis

There are genomic associations between BMD and osteoclast function and between specific SNPs and osteoclast function within the Iranian population.

To test this hypothesis the following objectives were determined:

4.1.2 Specific objectives

- Determine the number of osteoclasts, the number of osteoclasts in an active resorption state, and the resorption area on the dentine disc.
- Evaluate any genetic associations between BMD and osteoclast function
- Evaluate any genetic associations between the specific SNPs and osteoclast function

4.2 Introduction

In the previous chapter the data consistent with literature indicated that variations in BMD are genetically dependent (Arden, Baker et al. 1996). Polymorphisms which affect osteoclast function may also contribute to the heritability of bone mass and diseases such as osteoporosis (Boyle, Simonet et al. 2003). Meta-analysis studies have suggested that BMD is also linked to genetic markers within the RANK/RANKL/OPG system (Richards, Rivadeneira et al. 2008; Styrkarsdottir, Halldorsson et al. 2008; Styrkarsdottir, Halldorsson et al. 2009). The RANK, RANKL and OPG molecules are key regulators of osteoclastogenesis (Boyle, Simonet et al. 2003). In addition, a further GWA meta-analysis study has since identified several important BMD-related pathways, including the RANK/RANKL/OPG signalling pathway (Rivadeneira, Styrkarsdottir et al. 2009). The expression of RANK and RANKL was shown in osteoclast cultures (CD14+ monocyte cells) (www.biogps.org). OPG, which encodes the OPG protein, is mainly expressed by osteoblast cells rather than osteoclast cells (www.biogps.org), so association with osteoclast function in osteoclast culture is not expected.

There are limited data on the association between osteoclast activity and gene polymorphisms. The majority of studies examining genetic influences on bone have mainly focused on BMD due to availability of DXA scanners. However, BMD is effected by skeletal parameters including periosteal expansion, cortical density/thickness, and trabecular number/thickness, which may be biologically and genetically influenced (Bachrach 2001). Therefore assessing genetic associations in osteoclast function in correlation with BMD values could provide new insight into the genetics of osteoporosis. In the present study it was determined whether osteoclast number and resorption function were associated with variations in BMD and whether there were any genetic associations with osteoclast activity in a selection of an Iranian population sample.

4.3 Results

4.3.1 Determining osteoclast number in cultures obtained from a selected Iranian population

Osteoclast cultures were isolated from peripheral blood samples obtained from the selected Iranian population samples needed for allelic imbalance experiments later in the thesis as described in the **method section 2.7**. Osteoclast number was determined per culture using TRAP staining and osteoclast point counting, as outlined in the **method section 2.7**. The osteoclast differentiation process was monitored for approximately three weeks. In the first day of culture cells were small, spherical, and mononuclear (**Figure 4.1 A**). Within seven days of culture cells grew with spread cytoplasm but were still mononuclear. Around the second week cells were larger and began to display an osteoclast-like appearance, some possibly being multinucleated. They became typically huge and round with extended cytoplasm. It seems from these experiments that multinucleated cells first appear in the second week, but larger multinucleated osteoclasts are seen from the fourteenth day onward (**Figure 4.1 B**). Counts showed that there was variation in the number of osteoclasts between individual biological samples with an average osteoclast number of 218 in the sample population **(Figure 4.2).**

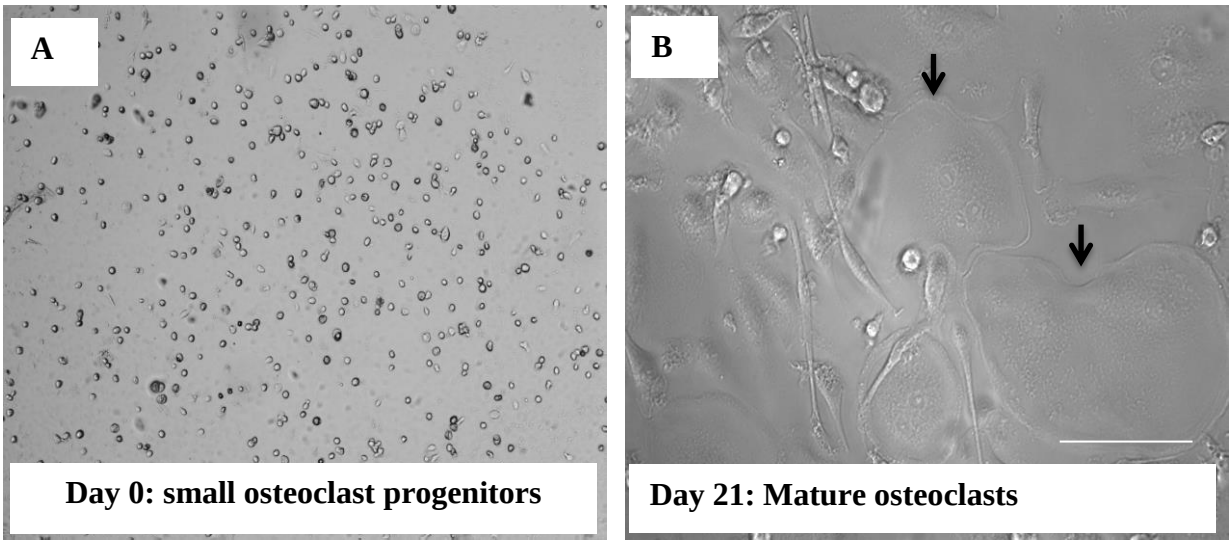


Figure 4-1: Osteoclast formation in culture medium supplemented with recombinant RANKL and MCSF

*Images are representative of cell appearance using phase-contrast microscopy. (**A**) D 0: cells are small, spherical and mononuclear, distributed in the culture medium. (**B**) D 21: cells are larger, multinucleated, with osteoclast shape (black arrows). Scale bar = 100 µm.*

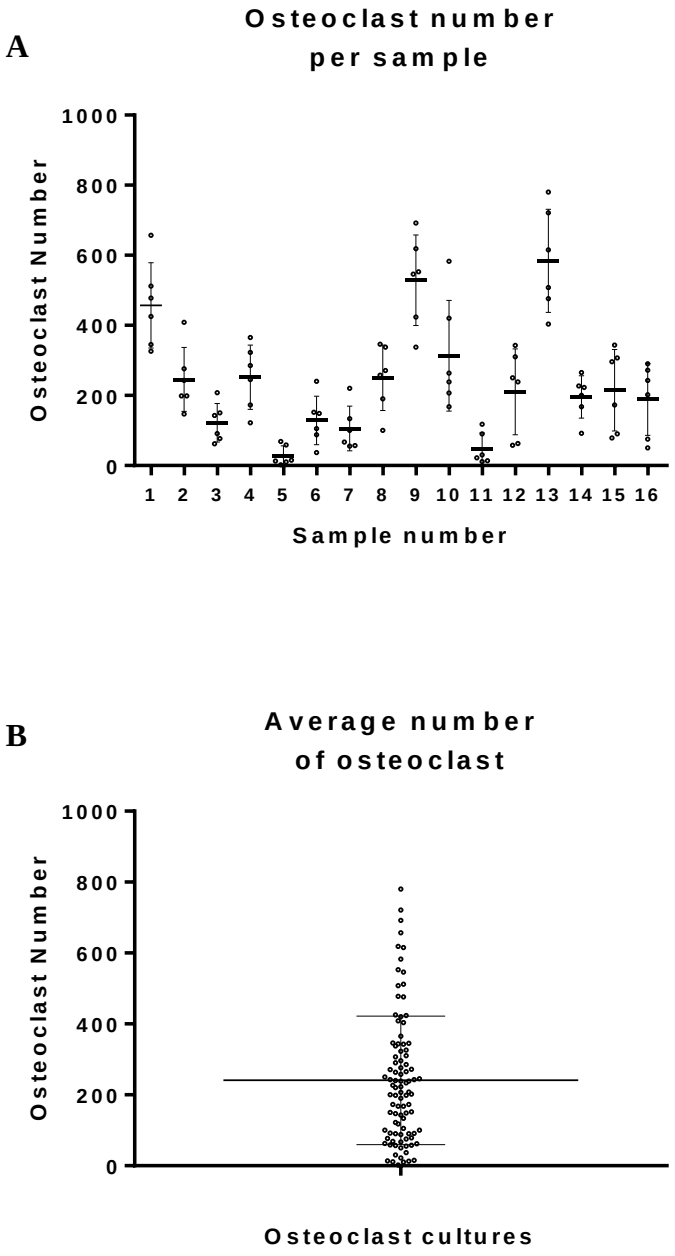


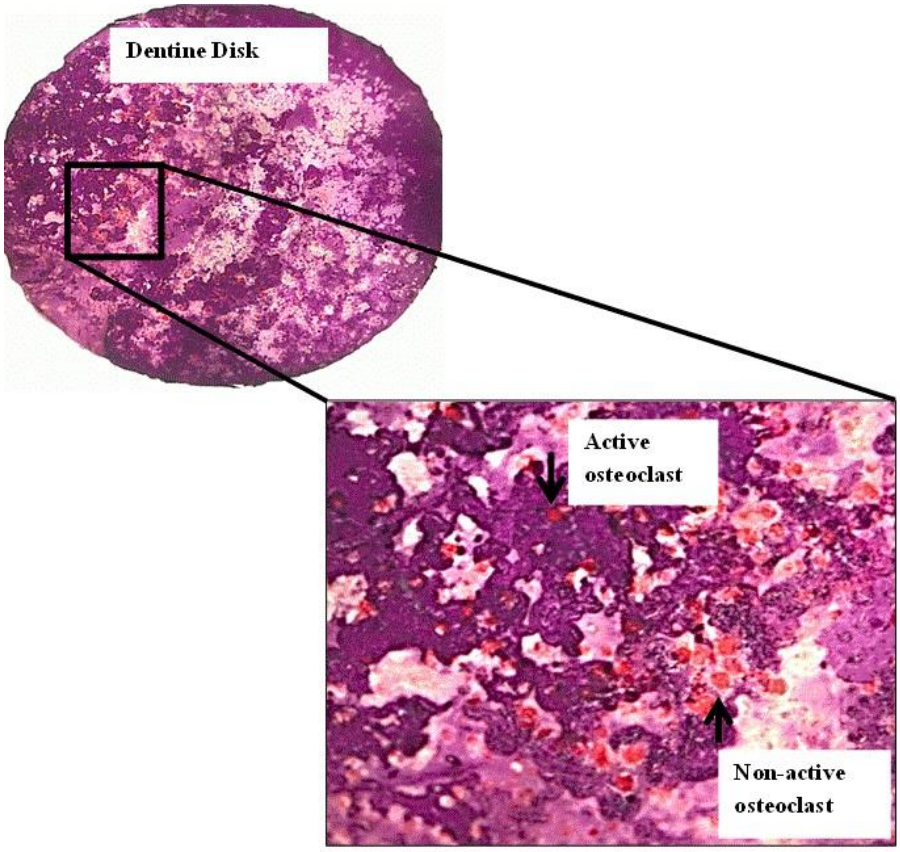
Figure 4-2: The average number of osteoclasts obtained from primary human samples was determined using Trap staining

Osteoclast cultures were isolated from human peripheral blood samples. Osteoclast numbers were determined per culture using Trap staining and osteoclast point counting. ***(A)*** *Data showed variations in the number of osteoclasts across individual biological samples.* ***(B)*** *The average osteoclast number was 218 per culture. N=16 individual biological cases with 6 replicates in each. The marked line shows the mean, and the bottom and top line show SD.*

4.3.2 Determining the number of active resorption osteoclast cells

Osteoclasts that demonstrated resorption activity were used as a marker of osteoclast function. Osteoclasts were isolated from the Iranian population samples as outlined in the **section 4.3.1**. Trap staining of osteoclasts on dentine disks was performed, and non-active and active osteoclasts (**Figure 4.3 A**) were counted as outlined in the **method sections 2.7.4 and 2.7.5**. Each osteoclast culture was maintained on 6 dentine disks, to give 6 replicates of the same biological sample. The percentage of resorbing osteoclasts was calculated based on the total number of osteoclasts. Osteoclast data from all donor samples within the selected Iranian population demonstrated that 19 – 83% of osteoclast cells were actively resorbing (**Figure 4.3 B).**

A



B

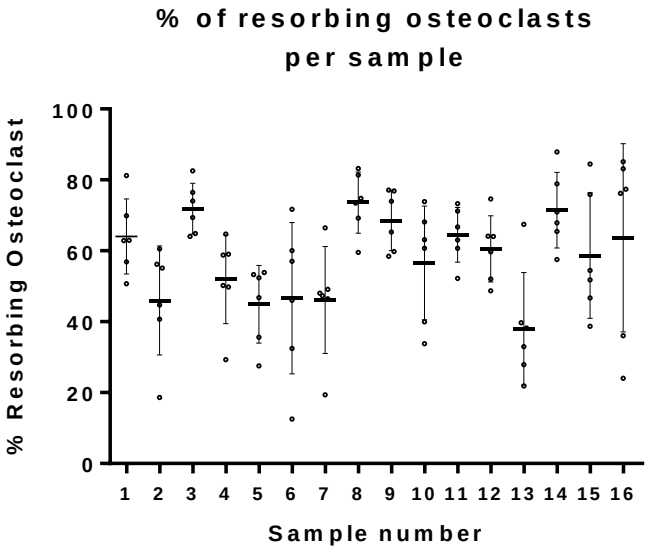


Figure 4-4: The % of actively resorbing osteoclasts from the total osteoclast population was determined using Trap staining

(A) *Osteoclasts cultured on dentine disks were either in an activated or non-activated resorption state.* ***(B)*** *The proportion (%) of actively resorbing osteoclasts in cultures. N=16 individual biological cases with 6 replicates in each. The marked line shows the mean, and the bottom and top line show SD.*

4.3.3 Determining the osteoclast resorption area

Isolated monocytes seeded on dentine discs were cultured in the presence of M-CSF and RANKL for 21 days. After two weeks resorption pits appeared on the dentine discs. After 21 days dentine discs were fixed and stained with toluidine blue staining to validate the Trap staining showing resorption. Resorption pits were visualised under the light microscope in fewer than 21 days. More defined pits were visible following three weeks in culture. The patterns of resorption pits produced by human osteoclast cells were often observed to be "snake like" (**Figure 4.4 A**). Osteoclasts formed authentic resorption lacunae on the dentine discs, determined by adjusting the focus of the light microscope (**Figure 4.4 B and C**). Commonly osteoclast cells could be seen located at the end of the resorption pit. The percentage of resorption area on dentine discs with osteoclast cells was determined in 16 Iranian population samples. Data showed that percentage (%) resorption area of the osteoclast cultures varied across biological samples, ranging between approximately 1-33 % resorption per disc (**Figure 4.4 D**).

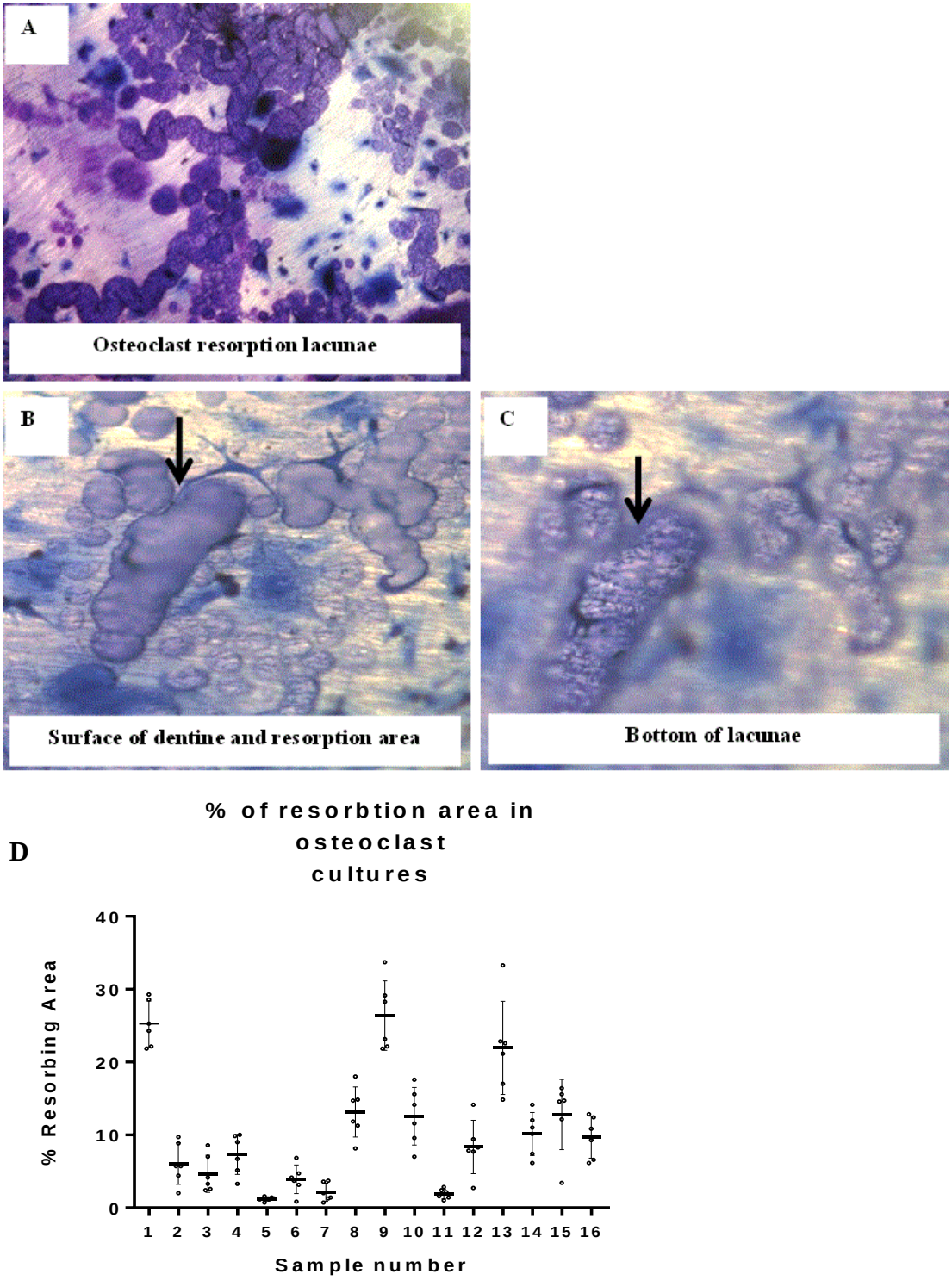


Figure 4-6: Determining osteoclast resorption activity using toluidine blue staining

(A) *snake-like resorption lacunae on dentine discs produced by osteoclast cells resorption lacunae on dentine disc after 3 weeks* ***(B)*** *Surface of dentine and resorption area.* ***(C)*** *Bottom of pit showing the depth of lacunae.* ***(D)*** *The proportion of the disc with resorption varied across biological samples. N=16 individual biological cases with 6 replicates in each. The marked line shows the mean, and the bottom and top line show SD.*

4.3.4 Osteoclast number and function were increased in osteopenia and osteoporotic patients

The objective of this part of the study was to assess any links between osteoclast number/activity and BMD in the Iranian population sample. First the BMD value of each of the donor samples was determined. BMD values were determined according to the protocol outlined in the **method section 2.1.2**. Subjects were placed into normal, osteopenia and osteoporosis subtypes, based on the L2-L4 and total hip BMD (total hip BMD) values. Subjects that had a BMD within 1 SD of the adult reference mean (T-score=> -1 SD) were classified as 'normal' in relation to bone disease. Any subjects with a BMD within 1 to 2.5 SD below the adult reference mean (-2.5<T-score< -1) were placed into the 'osteopenia' category and all subjects with BMD of <-2.5 SD below the adult reference mean (T-score=<-2.5) were classed as diagnosed with osteoporosis (Kanis 1994). The osteoclast resorption area was determined for each of the normal, osteopenic and osteoporotic samples.

Data showed that there were no differences in the osteoclast number within osteoporotic samples with low L2-L4 BMD compared to normal samples (**Figure 4.5 A**) and also was not observed in relation to total hip BMD-associated bone disease (**Figure 4.5 B).** The osteoclast numbers increased in osteopenic or osteoporotic donor samples with a L2-L4 BMD of >1 SD below the adult reference mean (**Figure 4.5 C,** *P<0.05).

Next, the proportions of resorbing osteoclasts calculated as a percentage of the total osteoclast population were compared between normal, osteopenic and osteoporotic sample groups. Data showed that there was no difference in the % of resorbing osteoclasts within osteoporotic samples with low L2-L4 BMD compared to that of normal samples (**Figure 4.6 A**). Data analysis also determined no difference in the % of resorbing osteoclasts within osteopenic patient samples with low total hip BMD compared to that of normal samples (**Figure 4.6 B**).

There were no difference in the % resorption area in osteoclast culture isolated from osteopenic and osteoporotic samples compared to normal samples. (**Figure 4.7 A, B**). Data showed that the % of osteoclast resorption area increased in sample donors with a L2-L4 BMD of >1 SD below the adult reference mean, i.e. those classified as osteopenic or osteoporotic (**Figure 4.7 C,** *P<0.05).

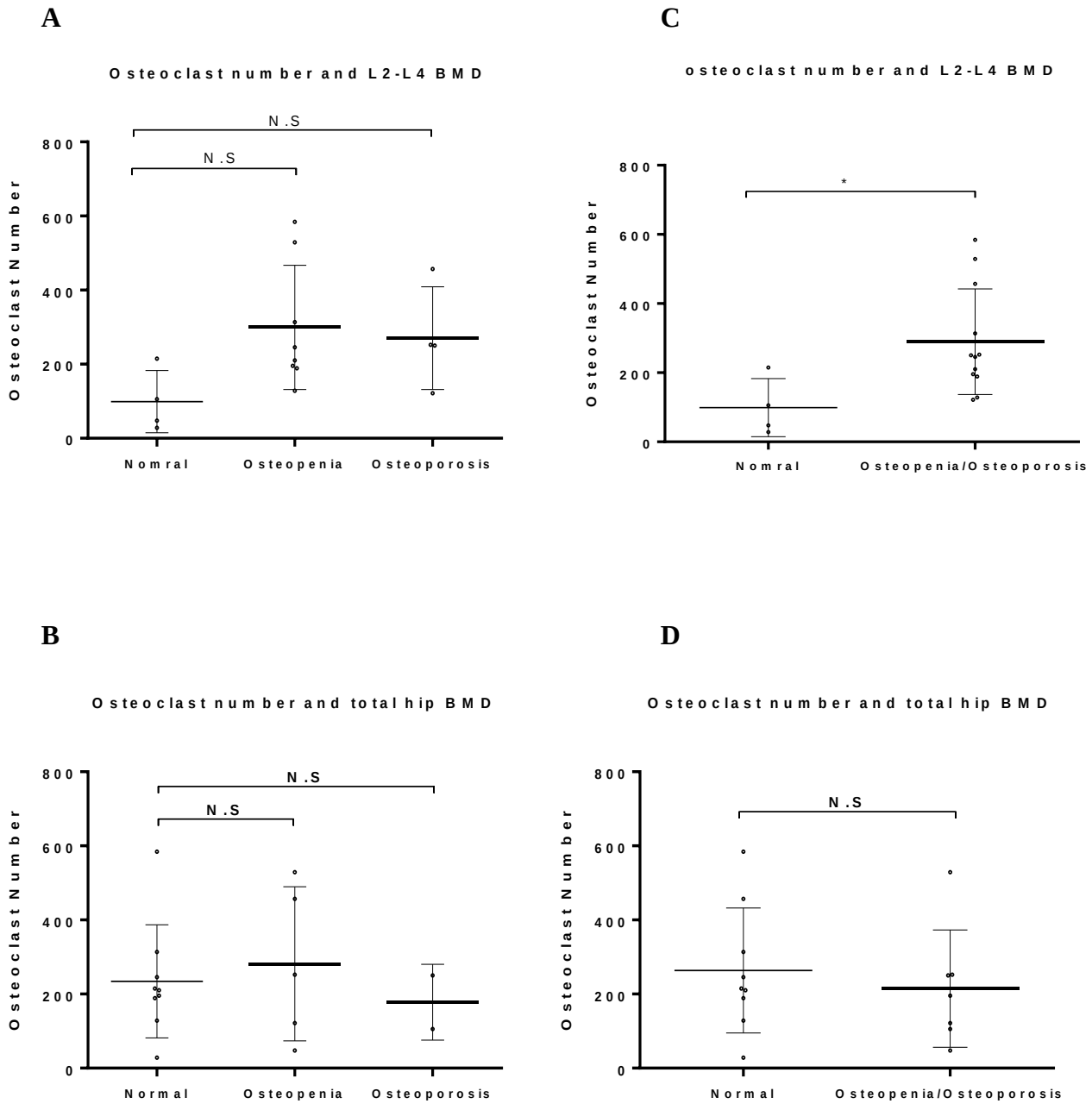


Figure 4-7: Osteoclast number was increased in patients with low L2-L4 BMD corresponding to osteopenia and osteoporosis

The osteoclast number determined from osteopenic and osteoporotic samples was compared to normal samples. ***(A, B)*** *There were no statistically significant differences in the osteoclast number between normal and osteopenic or osteoporotic samples.* ***(C, D)*** *The osteoclast number increased in osteopenic or osteoporotic sample donors with a L2-L4 BMD of >1 SD below the adult reference mean. N=16* biological cases*. The marked line shows the mean, and the bottom and top line show SD. Students t-test and Man-Whitney test, *P<0.05.*

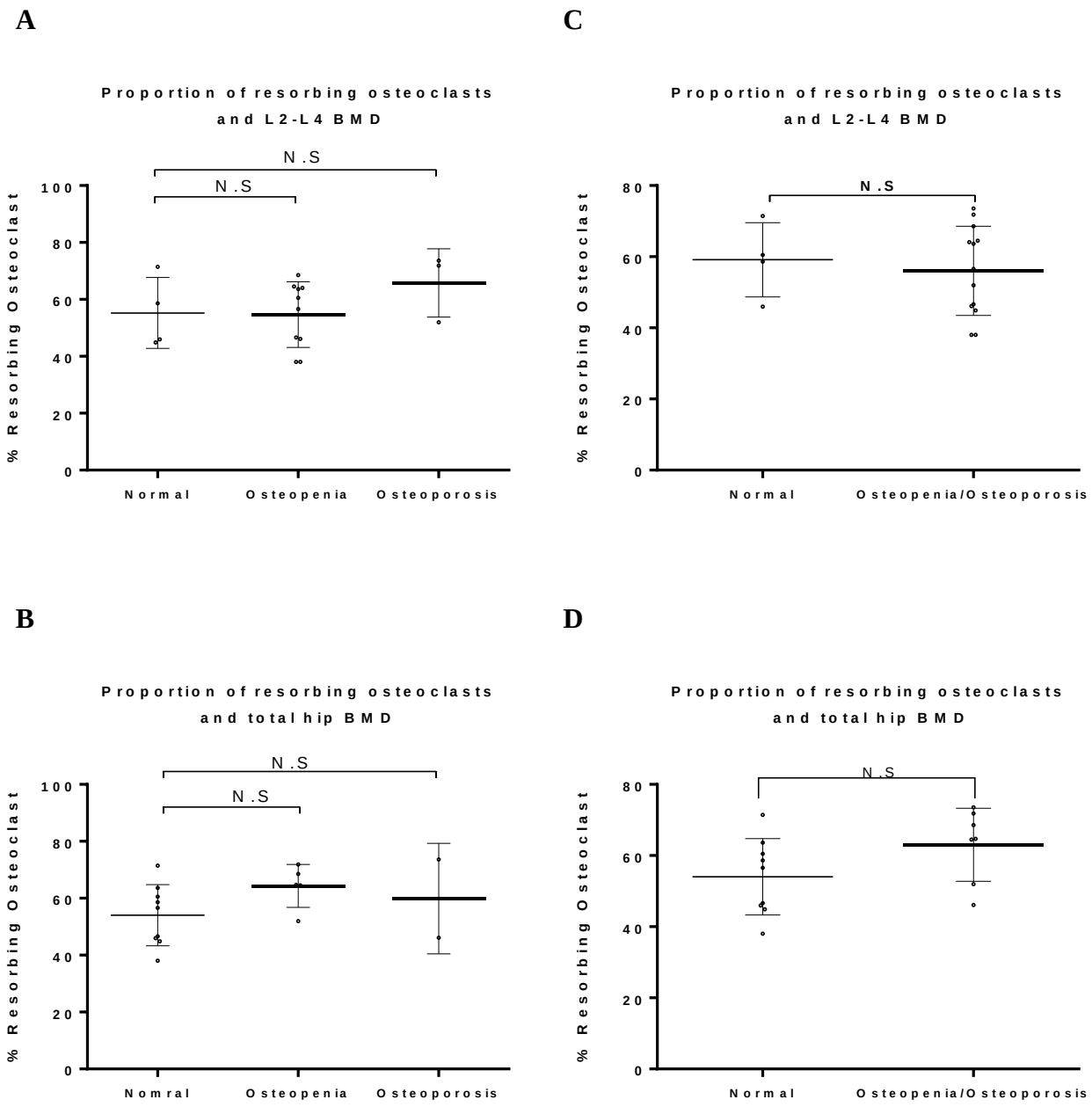


Figure 4-8: The proportion of resorbing osteoclasts increased in patients with low BMD corresponding to osteopenia and osteoporosis

The percentage of resorbing osteoclasts taken from whole osteoclast cultures was compared between normal, osteopenic and osteoporotic sample groups. ***(A, B,C,D)*** *There were no statistically significant differences in the % of resorbing osteoclast numbers between normal and osteopenic or osteoporotic samples. N=16 biological cases. The marked line shows the mean, and the bottom and top line show SD.*

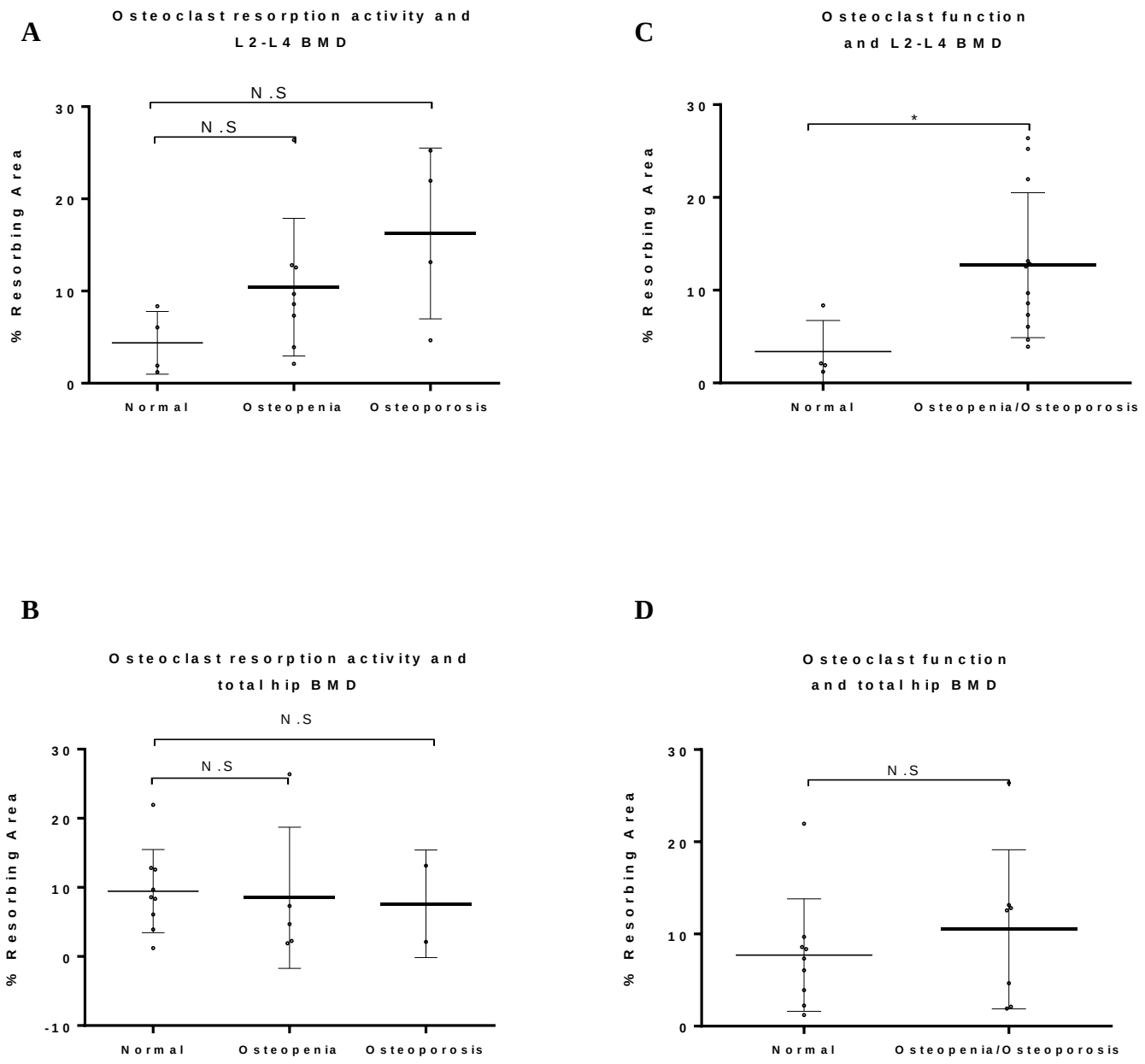


Figure 4-9: Osteoclast function was increased in patients with low L2-L4 BMD corresponding to osteopenia and osteoporosis

*The % osteoclast resorbing area in osteoclast culture isolated from osteopenic and osteoporotic samples was compared to normal samples. (A, B) There were no statistically significant differences in the % of osteoclast resorption between normal and osteopenic or osteoporotic samples. (C, D) The % of osteoclast resorption area increased in osteopenic or osteoporotic sample donors with a L2-L4 BMD of >1 SD below the adult reference mean. N=16 biological cases. The marked line shows the mean, and the bottom and top line show SD. Students t-test and Man-Whitney test, *P<0.0.*

4.3.5 Osteoclast number were genetically associated by SNPs in a selected Iranian sample

Specific genes including *RANK*, *RANKL* and *OPG* have been reported in the literature as genes which regulate osteoclastogenesis (Kikuchi, Matsuguchi et al. 2001). In the **method section 2.2.1,** the systemic review was identified risk SNPs associated with RANK, RANKL and OPG genes. In this study, we use the term 'risk' refer to SNPs which were associated with BMD. The following risk SNPs were identified for each of the genes:

- RANK - rs884205 and rs30183625
- RANKL - rs9594738, rs9594759 and rs1021188
- OPG - rs6469804, rs2062375, rs11995824, 4355801, and rs2062377

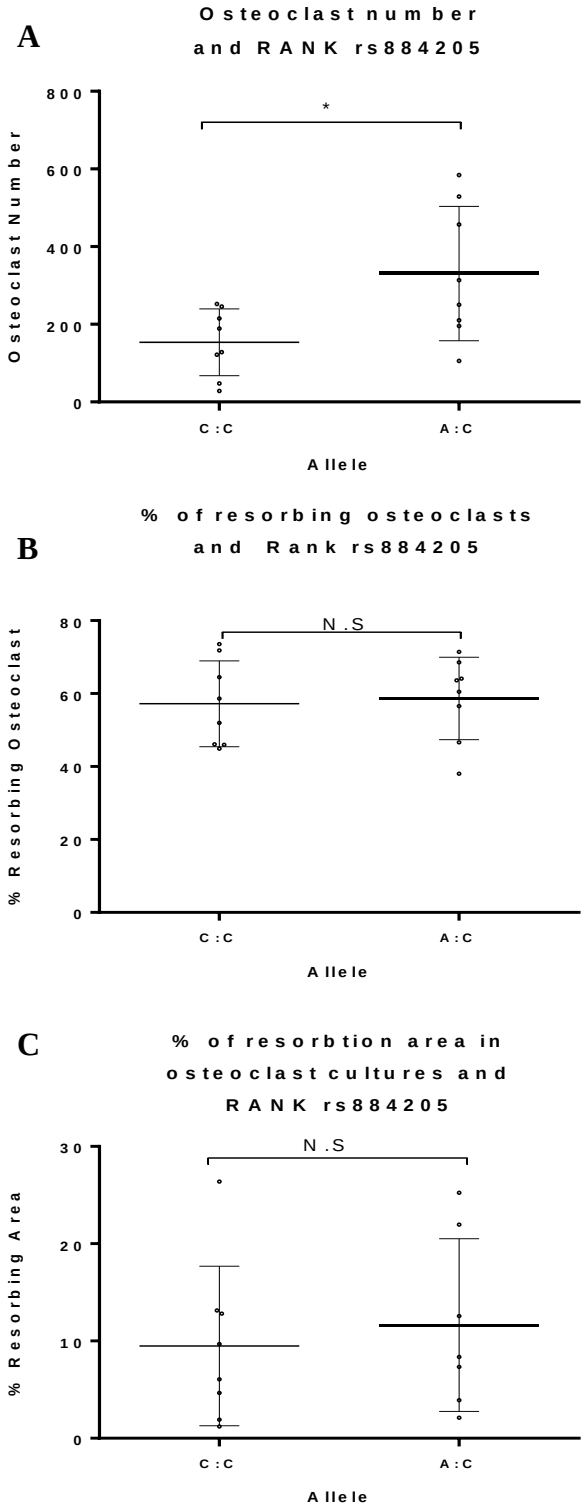
In this part of the study any association between the risk SNPs associated with *RANK*, *RANKL* and *OPG* genes and osteoclast number/function in the 16 samples (selected in the section 2.5) were determined. To perform DNA genotyping, DNA was isolated from the peripheral blood of the corresponding osteoclast study volunteers and extraction was performed as described in **section 2.3.** DNA samples were sent to K-Bioscience for genotyping as outlined in **section 2.4**. The osteoclast number, the % of resorbing osteoclast numbers and % resorbing area were determined as outlined in **section 2.7**.5. Any statistical genetic associations between the *RANK*, *RANKL* and *OPG* genes, risk SNPs, and osteoclast number/resorption were assessed. Genetic association between the RANK, RANKL and OPG genotypes and osteoclast number/resorption were analysed, specifically looking at whether increasing copies of the risk allele (either homozygous or heterozygous, depending on the SNP) is associated with decreasing BMD. According to the genotyping that was performed, Due to the small number of osteoclast samples, subjects were classified determined into have either two genotypes (homozygous wild type or risk and heterozygous risk) or three genotypes (homozygous wild type, heterozygous risk and homozygous risk).

Analysis of the genetic associations between the RANK gene SNPs rs884205 and rs30183625 and osteoclast number/ resorption activity showed that there was an increase in the osteoclast number (**Figure 4.8 A**, *P<0.05) in those with risk heterozygous genotype (A: C) compared to wild homozygous genotypes (C: C) of the rs884205 SNP. There were no differences observed in the % resorption area and % of resorbing osteoclasts between the

heterozygous and homozygous genotypes (**Figure 4.8 B, C**). Data also showed an increase in the osteoclast in those with risk heterozygous genotype (A: C) compared to wild homozygous genotypes (C: C) of the rs30183625 SNP (**Figure 4.9A** *P<0.05). No differences were determined in the % resorption area and % of resorbing osteoclasts between heterozygous and homozygous genotypes of the rs30183625 SNP (**Figure 4.9 B, C**).

The analysis for the genetic associations between the RANKL-associated risk SNPs rs9594759, rs9594738, and rs1021188 and osteoclast number/ resorption activity showed that there were no differences in the osteoclast number or % of resorbing osteoclasts between heterozygous and homozygous genotypes of rs9594738 (**Figure 4.10**). Statistical analysis showed that there were no differences in the osteoclast number or % of resorbing osteoclasts between heterozygous and homozygous genotypes of rs9594759 (**Figure 4.11).** Statistical analysis also showed that there were no differences in the osteoclast number or % of resorbing osteoclasts between heterozygous and homozygous genotypes of the rs1021188 SNP (**Figure 4.12**).

Further analysis for the genetic associations between the OPG-associated risk SNPs rs6469804, rs2062375, rs11995824 and rs2062377 and osteoclast number/ resorption activity showed that there were no differences in the osteoclast number, % of resorbing osteoclasts, and % resorption area between wild type and risk alleles in any of the tested rs6469804, rs2062375, rs11995824, or rs2062377 SNPs (**Figures 4.13 – 4.17).**



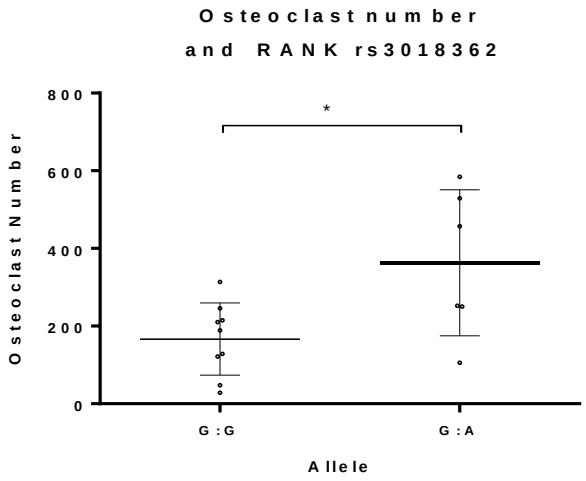
C: C Homozygous for the wild allele – C allele is not associated with decreasing BMD

A: C Heterozygous for the risk allele – A allele is associated with decreasing BMD

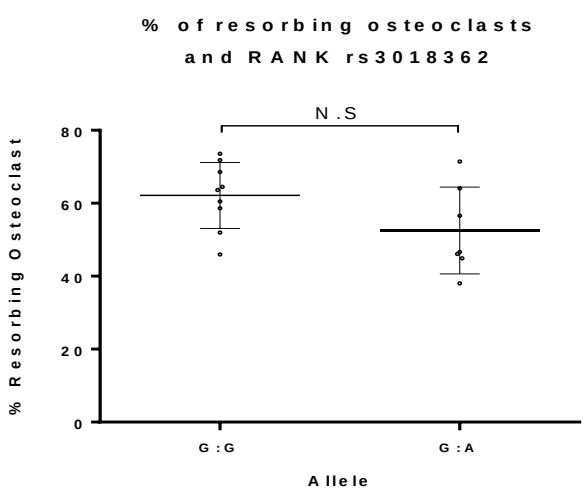
Figure 4-10: Osteoclast number and function were genetically associated by RANK SNP rs884205 genotype

*Osteoclast number, proportion of resorbing osteoclasts and the % resorption area were assessed in association with genotyping of RANK SNPs. (A) There was statistically significant differences in the osteoclast number in those homozygous for the C (wild) allele (C: C) compared to heterozygous for the A (risk) allele (A: C) in the rs884205 SNP. (B, C) There were no statistical significant difference in % resorption area and % of resorbing osteoclasts between heterozygous and homozygous genotypes allele of rs884205 SNP. The marked line shows the mean, and the bottom and top line show SD. Student's t-Test and Man-Whitney test, *P<0.05.*

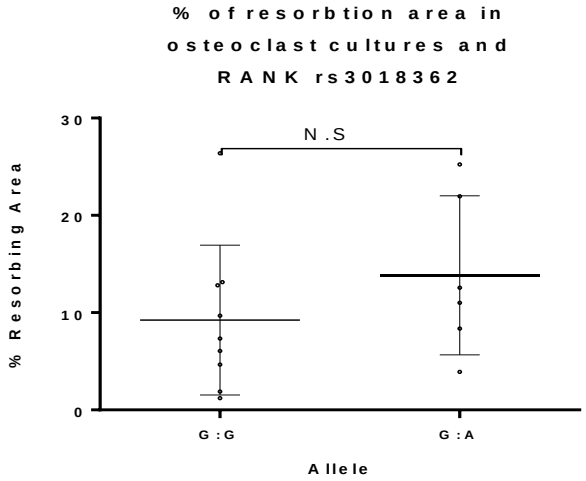
A



B



C



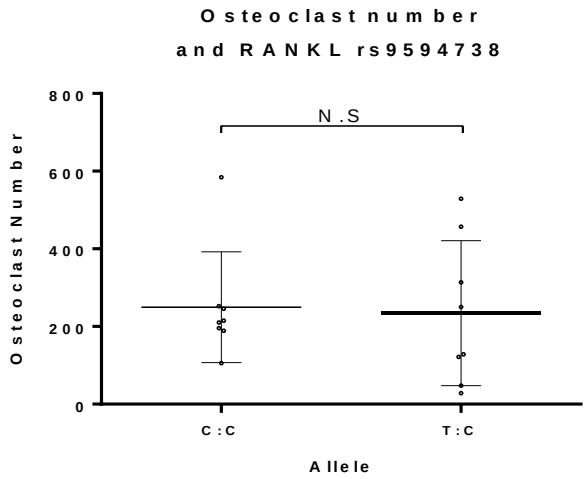
G: G Homozygous for the wild type allele – G allele is not associated with decreasing BMD

G: A Heterozygous for the risk allele – A allele is associated with decreasing BMD

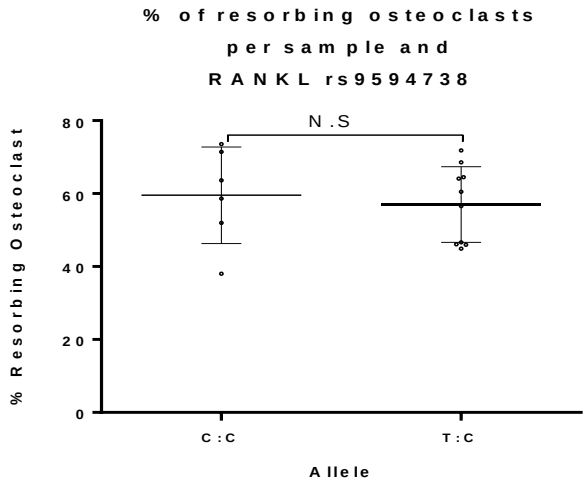
Figure 4-11: Osteoclast number and function were genetically associated by RANK SNP rs3018362 genotype

*Osteoclast number, proportion of resorbing osteoclasts and the % resorption area were assessed in association with genotyping of RANK SNPs. (A) There was statistically significant differences in the osteoclast number in those homozygous for the G (wild) allele (G: G) compared to heterozygous for A (risk) allele (G: A) in the rs30183625 SNP. (B, C) There were no statistically significant difference in % resorption area and % of resorbing osteoclasts between heterozygous and homozygous genotypes of rs30183625 SNP. The marked line shows the mean, and the bottom and top line show SD. Student's t-Test and Man-Whitney test, *P<0.05.*

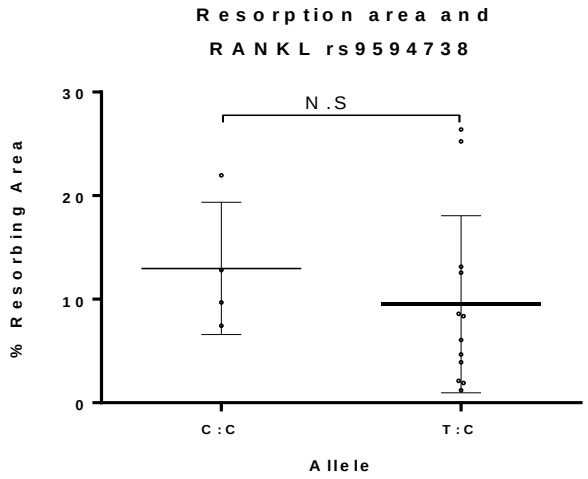
A



B



C



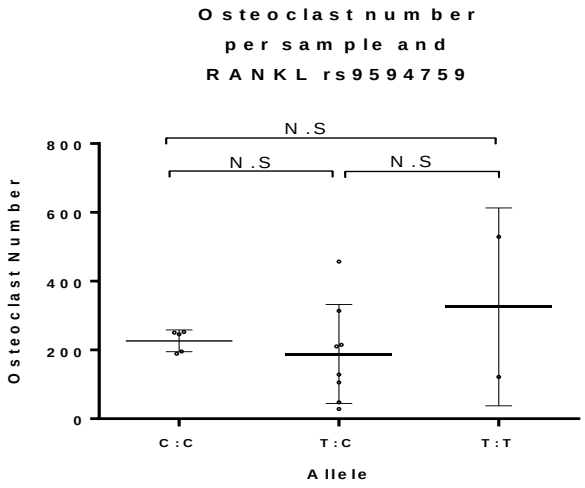
C: C Homozygous for the wild type allele – C allele is not associated with decreasing BMD

T: C Heterozygous for the risk allele – T allele is associated with decreasing BMD

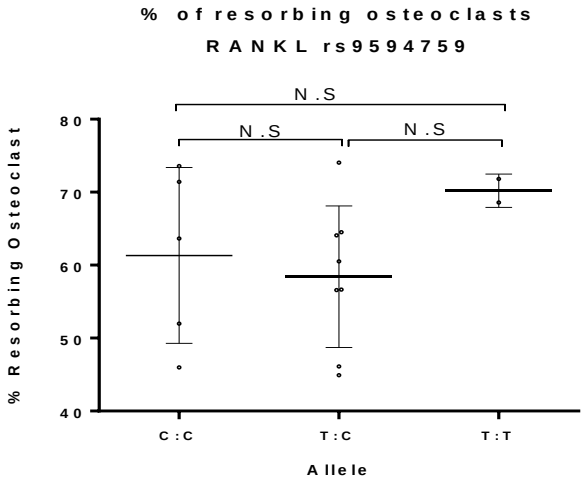
Figure 4-12: Osteoclast number and function were genetically associated by RANKL SNP rs9594738 genotype

Osteoclast number, proportion of resorbing osteoclasts and the % resorption area were assessed in association with genotyping of RANKL SNPs. (A, B, C) There were no statistically significant differences in the osteoclast number, % of resorbing osteoclasts, or % resorption area between heterozygous and homozygous genotypes of rs9594738 SNP. The marked line shows the mean, and the bottom and top line show SD. Student's t-Test.

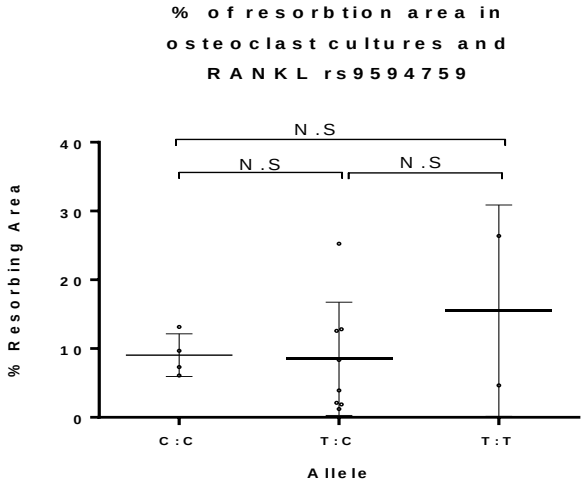
A



B



C



C: C Homozygous for the wild type allele – C allele is not associated with decreasing BMD

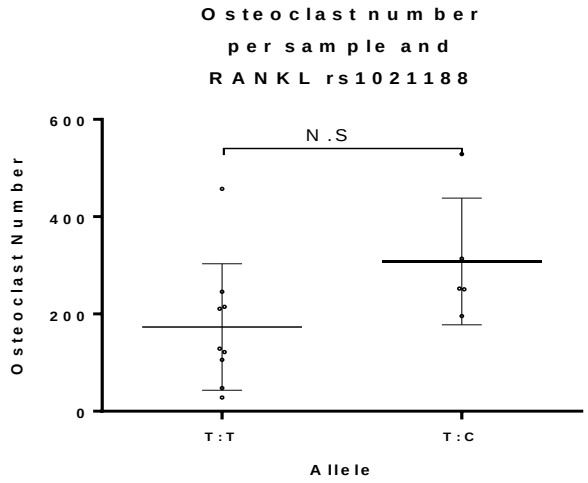
T: C Heterozygous for the risk allele – T allele is associated with decreasing BMD

T: T Homozygous for the risk allele

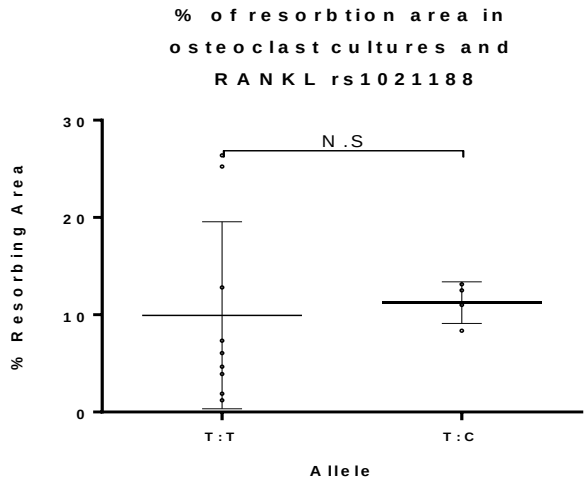
Figure 4-13: Osteoclast number and function were genetically associated by RANKL SNPs rs9594759

Osteoclast number, proportion of resorbing osteoclasts and the % resorption area were assessed in association with genotyping of RANKL SNPs. (A, B, C) There were no statistically significant increases in the osteoclast number, % of resorbing osteoclasts and % resorption area in those homozygous for the T (risk) allele (T: T) compared to those heterozygous for T allele (T: C)or those homozygous for the C allele in the rs9594759 SNP. The marked line shows the mean, and the bottom and top line show SD. One-way ANOVA.

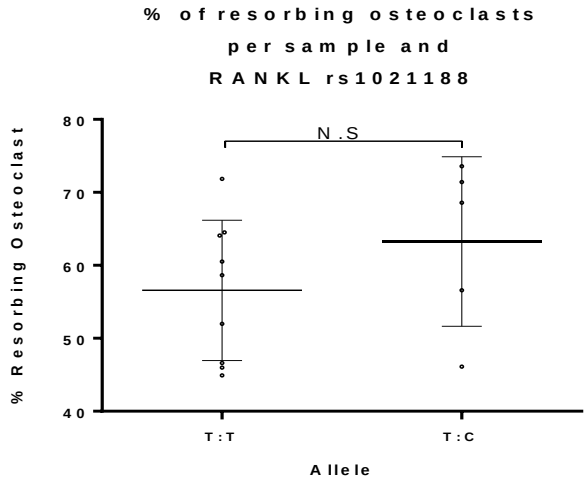
A



B



C



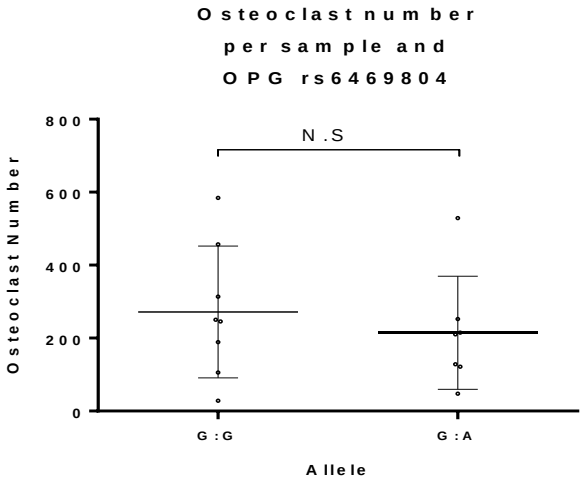
T: T Homozygous for the wild type allele – T allele is not associated with decreasing BMD

T: C Heterozygous for the risk allele – C allele is associated with decreasing BMD

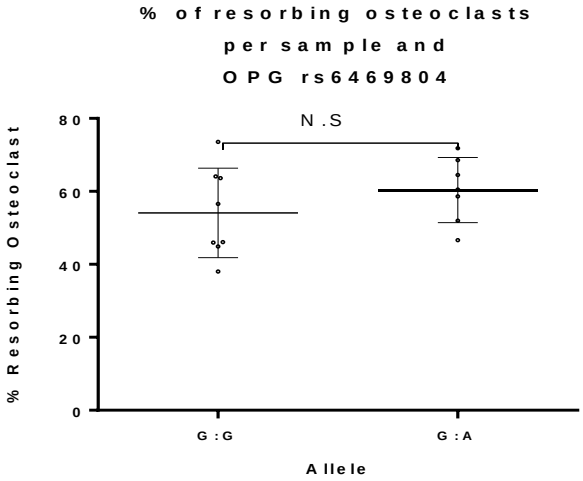
Figure 4-14: Osteoclast number and function were genetically associated by RANKL SNP rs1021188

Osteoclast number, proportion of resorbing osteoclasts and the % resorption area were assessed in association with genotyping of RANKL SNPs. (A, B, C) There were no statistically significant increases in the osteoclast number, % of resorbing osteoclasts and % resorption area in those heterozygous for C (risk) allele compared to those homozygous for T (wild) allele in the rs1021188 SNP. The marked line shows the mean, and the bottom and top line show SD. Student's t-Test.

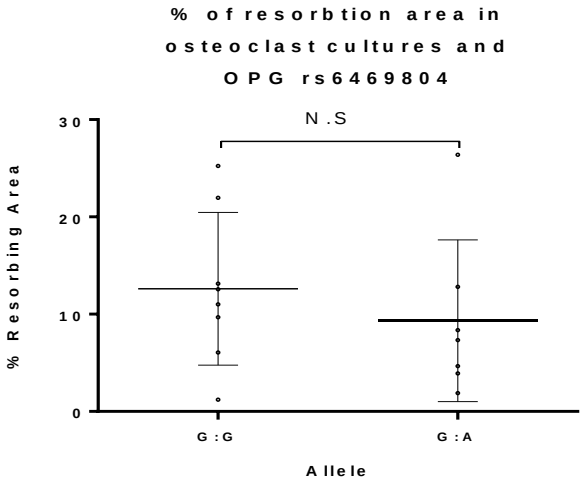
A



B



C



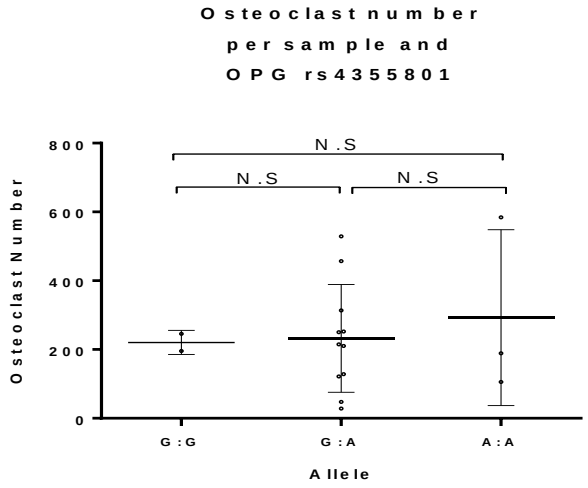
G: G Homozygous for the wild type allele – G allele is not associated with decreasing BMD

G: A Heterozygous for the risk allele – A allele is associated with decreasing BMD

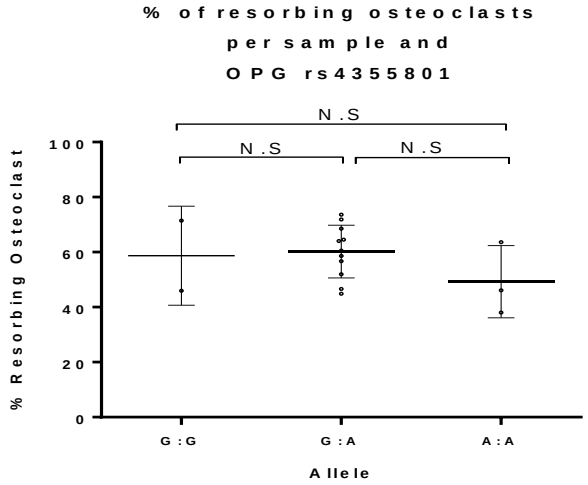
Figure 4-15: Osteoclast number and function were genetically associated by OPG SNP rs6469804

Osteoclast number, proportion of resorbing osteoclasts and the % resorption area were assessed in association with genotyping of OPG SNPs. (A, B, C) There were no statistically significant differences in osteoclast number, % of resorbing osteoclasts or % resorption area between heterozygous and homozygous genotypes of the rs6469804 SNP. The marked line shows the mean, and the bottom and top line show SD. Student's t-Test.

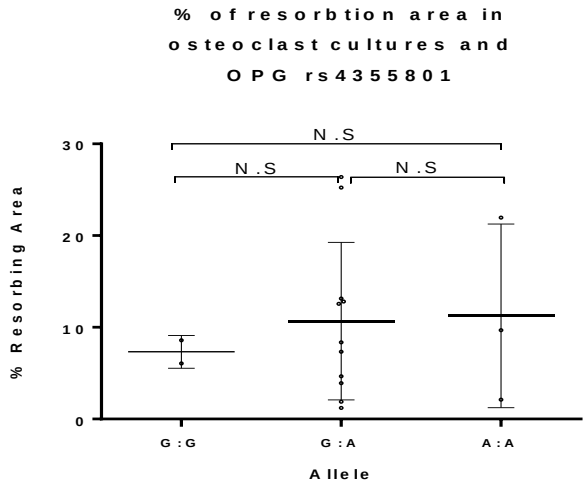
A



B



C



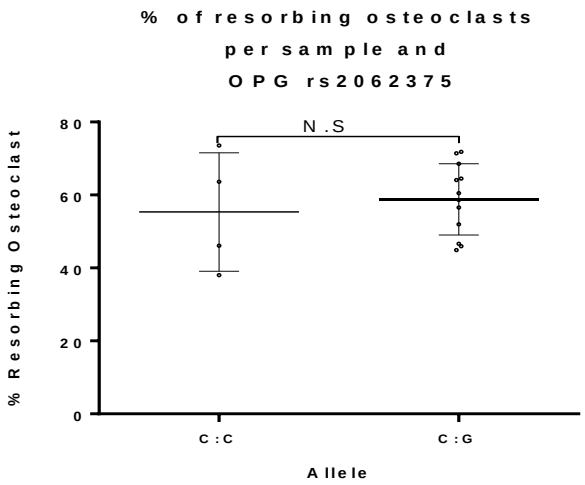
G: G Homozygous for the wild type allele – G allele is not associated with decreasing BMD

G: A Heterozygous for the risk allele – A allele is associated with decreasing BMD

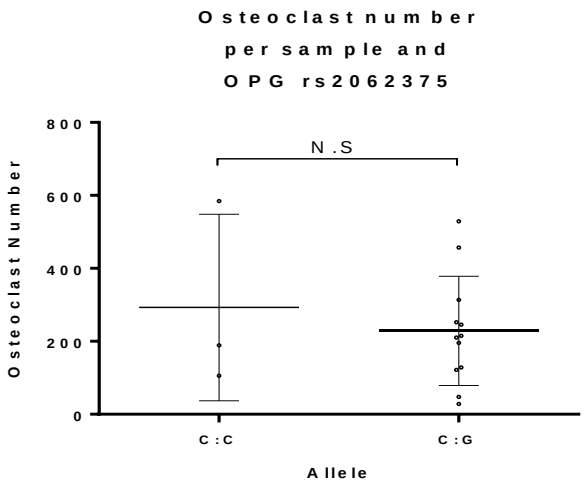
Figure 4-16: Osteoclast number and function were genetically associated by OPG SNP rs4355801

Osteoclast number, proportion of resorbing osteoclasts and the % resorption area were assessed in association with genotyping of OPG SNPs. (A, B, C) There were no statistically significant differences in osteoclast number, % of resorbing osteoclasts or % resorption area between heterozygous and homozygous genotypes of the rs6469804 SNP. The marked line shows the mean, and the bottom and top line show SD. One-way ANOVA.

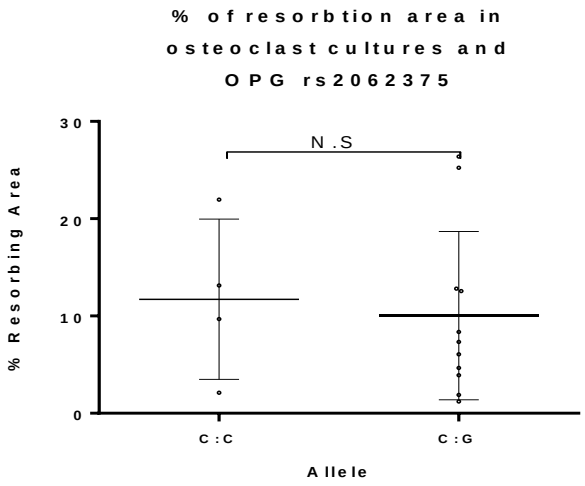
A



B



C



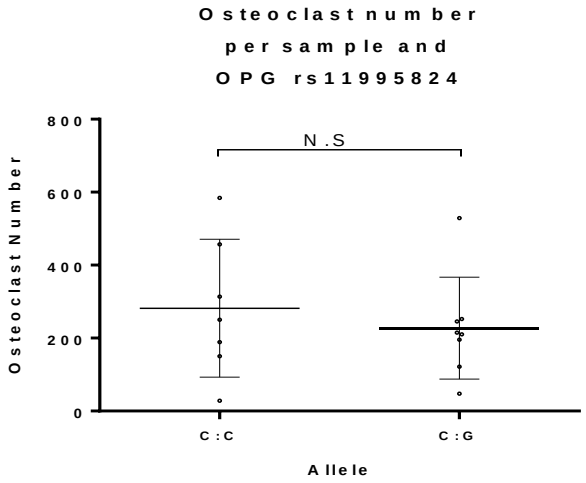
C: C Homozygous for the wild type allele – C allele is not associated with decreasing BMD

C: G Heterozygous for the risk allele – G allele is associated with decreasing BMD

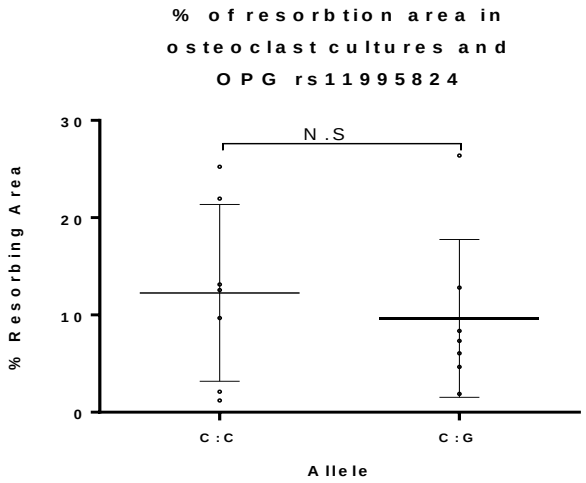
Figure 4-17: Osteoclast number and function were genetically associated by OPG SNP rs2062375

Osteoclast number, proportion of resorbing osteoclasts and the % resorption area were assessed in association with genotyping of OPG SNPs. (A, B, C) There were no statistically significant differences in osteoclast number, % of resorbing osteoclasts or % resorption area between heterozygous and homozygous genotypes of the rs2062375 SNP. The marked line shows the mean, and the bottom and top line show SD. Student's t-Test.

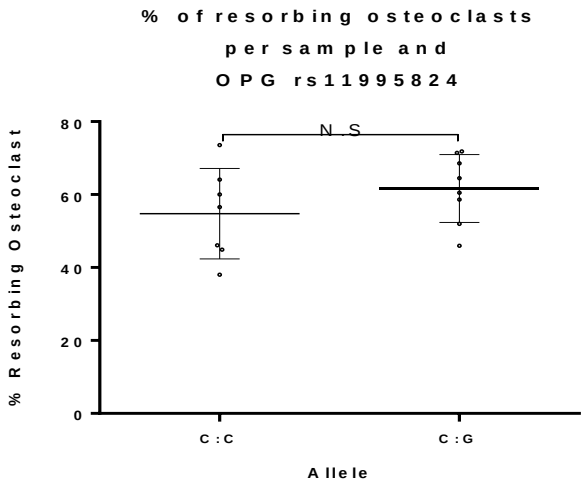
A



B



C



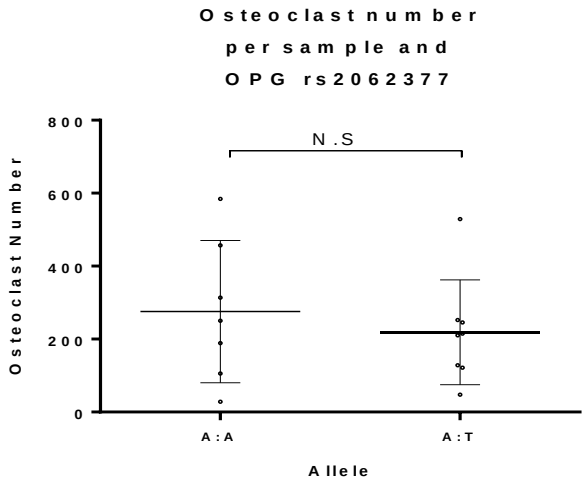
C: C Homozygous for the wild type allele – C allele is not associated with decreasing BMD

C: G Heterozygous for the risk allele – G allele is associated with decreasing BMD

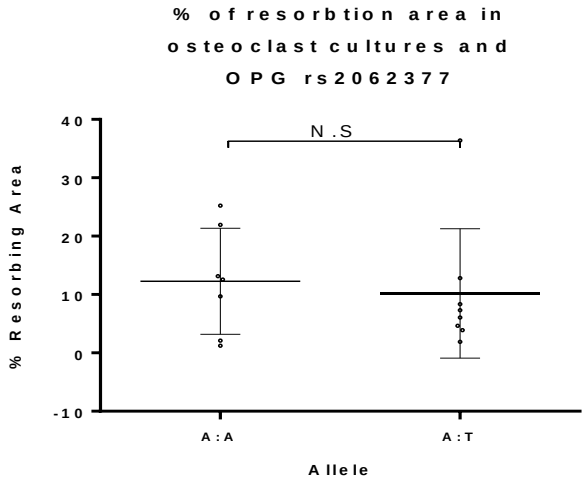
Figure 4-18: Osteoclast number and function were genetically associated by OPG SNP rs11995824 genotype

Osteoclast number, proportion of resorbing osteoclasts and the % resorption area were assessed in association with genotyping of OPG SNPs. (A, B, C) There were no statistically significant differences in osteoclast number, % of resorbing osteoclasts or % resorption area between heterozygous and homozygous genotypes of the rs11995824 SNP. The marked line shows the mean, and the bottom and top line show SD. Student's t-Test.

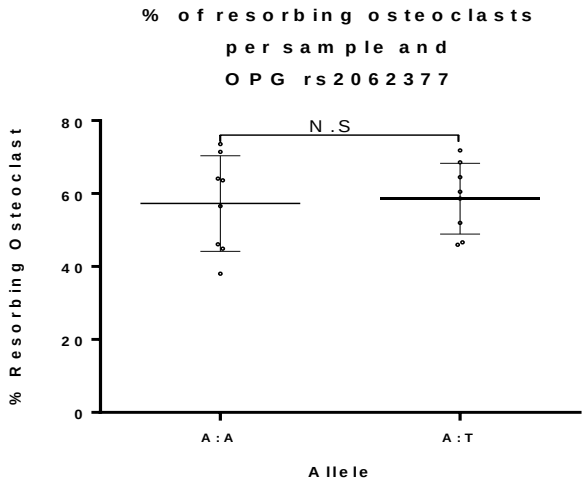
A



B



C



A: A Homozygous for the wild type allele – A allele is not associated with decreasing BMD

A: T Heterozygous for the risk allele – T allele is associated with decreasing BMD

Figure 4-19: Osteoclast number and function was genetically associated by OPG SNP rs2062377 genotype

Osteoclast number, proportion of resorbing osteoclasts and the % resorption area were assessed in association with genotyping of OPG SNPs. (A, B, C) There were no statistically significant differences in osteoclast number, % of resorbing osteoclasts or % resorption area between heterozygous and homozygous genotypes of the rs2062377 SNP. The marked line shows the mean, and the bottom and top line show SD. Student's t-Test.

4.4 Discussion

In the previous chapter data are consistent with the larger body of published evidence, suggested that variations in BMD are genetically dependent (Arden, Baker et al. 1996). Polymorphisms which effect osteoclast function may also contribute to the heritability variation of bone mass and diseases such as osteoporosis. The formation of osteoclasts is centred on the key osteoclastogenic cytokine RANKL. Although there are many inflammatory cytokines known to promote the formation and activity of osteoclasts, RANKL is considered to be the final downstream effector cytokine that drives osteoclastogenesis and regulates osteoclastic bone resorption (Weitzmann 2013). The RANKL/RANK/OPG axis is now recognised as the pivotal regulator of osteoclastogenesis and mutations in the *RANK*, *OPG*, and *RANKL* genes cause bone disease, including osteoporosis (Kobayashi, Udagawa et al. 2009). In this chapter, we wanted to test the hypothesis that there are genomic associations between specific RANK, RANKL and OPG risk polymorphisms and osteoclast function within the Iranian population. The aim of this study was to firstly determine the number of osteoclasts and their level of resorption activity within the sample group. Next, any associations between these risk SNPs and osteoclast number/activity was assessed.

A role for increased osteoclastic bone resorption observed in osteoporotic bone disease has been suggested by studies which have shown an increase in biochemical markers of bone resorption in patients with low skeletal BMD (Gough, Sambrook et al. 1998). However, there are limited data on the genetic associations of osteoclast number/activity and BMD or RANK, RANKL and OPG risk SNPs, particularly in the Iranian population. Therefore, to investigate osteoclast activity in relation to BMD and genetic associations, we first characterised our human osteoclast model *in-vitro*.

The ability to study the human osteoclasts formation and function of *in-vitro* has been limited. Our group has provided a reliable guide to isolating and culturing primary osteoclasts from human peripheral blood, with effective methodology that enables us to characterise osteoclasts and quantify resorption activity (Agrawal, Gallagher et al. 2012). However, as with all methods and protocols, there are some limitations. The initial attempt to culture osteoclasts from our human samples resulted in cells in the culture media that did not differentiate into mature osteoclast lineage. One possible reason for this problem may

have been the technical limitation of using the opaque interface containing the monocytes and lymphocytes fraction of blood for osteoclast isolation. An alternate method of osteoclast extraction derived from CD14+ monocytes isolated from peripheral blood (Sorensen, Henriksen et al. 2007), proved to be far more reliable in osteoclast cell differentiation within this study. Cultured mature osteoclast cell morphology and the presence of resorption pits were in agreement with findings from our group and others (Schilling, Linhart et al. 2004; Kreja, Liedert et al. 2007; Agrawal, Gallagher et al. 2012).

As it was indicated before a single osteoclast culture could be performed to fulfil a dual purpose of functionally characterising the osteoclast cells in vitro and determining allelic imbalance expression. The limitations for this part of the study were the high cost of material for culture in remote sites, and the need for the large blood volume sample required from each subject. Within our project budget, approximately 16 cultures were affordable in a twelve months period in Iran. These samples are selected according to high heterozygosity for transcribed SNPs and are not random sampling but because they are not selected according to risk SNPs so the results are affected less.

Having successfully cultured mature osteoclasts, TRAP staining of osteoclast cultures showed that there were variations in osteoclast numbers across donor samples in our selected Iranian population. Osteoclasts that demonstrated resorption activity were used as a marker of osteoclast function. These data showed that the proportion of resorbing osteoclasts taken from a total osteoclast population also varied greatly between donors. This data showed that although there was variability in the number of actively resorbing osteoclasts across donors, at least 1/3 of all the osteoclasts obtained from our human donors were indeed actively functioning. Subsequently data showed that the percentage area of resorption of the osteoclast cultures obtained from each donor varied, in the approximate range 1-33 % resorption area. To assess possible associations between osteoclast number/activity and BMD in our Iranian population sample, the BMD values of each of our 16 donor samples was determined and the donors placed into normal, osteopenia and osteoporosis subtypes. Our data suggested that individuals classified as having osteopenia or osteoporosis in L2-L4 also, were associated with increased osteoclast number and resorption activity compared to normal individuals. Interestingly, this effect was observed relative to L2-L4 BMD and not total hip BMD. One possible explanation for this may be that from the total number of patients from our 16 donors only 2 were osteoporotic compared with 5 that were osteopenic in total hip region. These ratios in fact mirror data from the entire population from other

larger Iranian studies (Omrani 2006). A focus of interest within this study is that most of the associations observed with osteoclastogenesis and BMD appear to be evident in osteopenic and osteoporotic patients with low L2-L4 BMD. This is consistent with other studies within the literature that also those genetic associations with BMD are dependent on the bone site (Tran, Nguyen et al. 2008). The majority of subjects used in GWA studies related to BMD consisted of postmenopausal women and elderly men (Paternoster, Ohlsson et al. 2010). Therefore, it is difficult to determine whether the genetic effects reported are actually related to the control of peak bone gain in early life or bone loss that is age-related.

A further analysis of gene polymorphisms associated with osteoclastogenesis regulators RANK, RANKL and OPG (Kikuchi, Matsuguchi et al. 2001) allowed us to identify the risk SNPs associated with these genes. Our objective was to then determine any genetic links between the risk SNPs associated with *RANK, RANKL* and *OPG* genes and osteoclast number or function. Our data showed that out of the three genes investigated, RANK polymorphisms had the most genetic association with osteoclast number.

Data show that there were statistically significant associations between increasing osteoclast number and risk allele of RANK rs884205/rs3018362 SNPs. These associations were in the direction that would be expected from risk alleles of BMD. Analysis of genetic associations between RANKL rs9594738/rs9594759/rs1021188 risk SNPs and osteoclast number and resorption activity showed that there were no statistically significant association between increasing in the osteoclast number and activity and risk alleles of rs9594759 /rs1021188/rs9594738 SNPs. No genetic associations were determined between osteoclast number/function and OPG SNPs rs6469804, rs2062375, 2062377, 4355801 and rs11995824.

We have shown that osteoclast number and activity appear to be reduced in individuals with low L2-L4 BMD. It was of particular interest to us that the 16 individuals selected on the basis of strongest genetic heterozygosity from our study cohort, >70% were either osteopenic or osteoporotic in L2-L4 BMD. Taken together, these findings suggest that there may be genetic associations between specific RANK/RANKL polymorphisms and osteoclast number and function.

Finally the main findings from this chapter are as follows:

The osteoclast numbers increased in osteopenic or osteoporotic donor samples with a L2-L4 BMD of >1 SD below the adult reference mean.

The % of osteoclast resorption area increased in sample donors with a L2-L4 BMD of >1 SD below the adult reference mean.

There were statistically significant associations between osteoclast number and risk allele of RANKL rs9594759 /rs1021188 SNPs according to genotyping and osteoclast function assay. These associations are in the direction that would be expected from risk alleles of BMD.

Chapter 5 – Allelic Expression Imbalance

5.1 Hypothesis and Objectives

5.1.1 Hypothesis

That association between candidate risk SNPs and variation in BMD may be mediated through allelic specific imbalance expression in different cell types.

To test this hypothesis and to determine whether cis-acting effects of the risk SNPs in the selected candidate genes lead to allelic imbalance in different cellular populations, namely whole blood and differentiated osteoclasts the following objective was determined:

5.1.2 Specific objective

Determine level of allelic expression in whole blood and differentiated osteoclasts by using transcribed SNPs with Sequenom technology

5.2 Introduction

Gene expression is the process in which functional gene products are synthesised by using the information of genes. These products are proteins or non-protein coding genes such as transfer RNA (tRNA) genes or ribosomal RNA (rRNA) genes. Gene expression modulates a whole host of functions ranging from DNA-RNA transcription to post-translational modification of proteins. Recently, several studies have shown that genetic variation at a sequence level can result in the modulation of gene expression in a cell (Tung, Fedrigo et al. 2009). The variations associated with disease susceptibility, particularly in complex diseases, frequently alter expression rather than function of proteins (Freedman, Monteiro et al. 2011). Such changes in gene expression may occur in two ways: by acting on genes at a distance through an expressed transcription factor (or non-coding RNA) on both the maternal and paternal chromosome (trans-acting factor), or by acting to regulate a gene on the same short range haplotype (allele) which resides on the same chromosome (cis-acting factor) (Monks, Leonardson et al. 2004; Cheung, Spielman et al. 2005). Cis-acting variants induce allelic expression imbalance (AEI). Trans-acting factors and environmental backgrounds influence expression of both alleles and are therefore not responsible for AEI (Tung, Fedrigo et al. 2009). Cis-acting mutations may alter the regulation of a single allele through a change to the promoter regions, intron regions or enhancer elements that affect transcription factor-binding sites.

The allele-specific expression is investigated by evaluation of cDNA from tissues expressing the genes of interest in the heterozygous (for the putative cis-acting variant) individual as it is not technically possible to measure the relative allele expression in a homozygous individual for the cis-acting variant. A major advantage of this approach is that the two SNP alleles are measured in the context of the same trans-acting and environmental factors with likely cis-acting regulatory variants or parent of origin (imprinting) effects (Farh, Grimson et al. 2005). GWAS show AEI in approximately 20% of expressed genes (Milani, Gupta et al. 2007; Serre, Gurd et al. 2008). Consequently genetic variation has a significant role in the variation on gene expression between individuals. These differences can correlate with phenotypic differences or even with risk of diseases (Teare, Heighway et al. 2006).

The most comprehensible way of observing allelic expression is by determining whether the imbalance observed from a set of transcribed SNPs across different heterozygous individuals

is consistent with a risk-acting effect. In the simplest scenario we consider that only one candidate polymorphism is tested and one transcribed polymorphism is used to measure technically the expression of the two alleles to find the ratio of one allele to the other. For an individual that is heterozygous at the transcribed and the risk-acting sites, the overexpressed transcribed allele is on the same chromosome (haplotype) as the risk-acting allele causing overexpression. A haplotype is a collection of DNA variation or polymorphisms that are likely to be inherited together. The phase between alleles at the two sites may vary between individuals. Determining the effect of the risk-acting polymorphism requires identifying the linkage disequilibrium (dissociation/association) between the transcribed and risk-acting sites. When there is complete linkage disequilibrium between both polymorphisms, only two distinct haplotype exists so those that are heterozygous should show systematic overexpression of the same transcribed allele.

Within the current study, I aimed to investigate evidence of cis acting effects which may account for some of the association between BMD variation and genetic variations. Allele specific expression can only be detected technically where the transcribed alleles can be distinguished in some way (such as heterozygous for the transcribed regions) **(Figure 5.1).** I am interested in studying allelic expression imbalance (AEI) in cDNA obtained from osteoclast cultures and peripheral blood tissue samples. These data support the ideology that there may be a potential genetic association between non-transcribed SNPs and allelic imbalances in relation to BMD. This potential concept is further described in **Figure 5.2**, in which, using the TNSFR11B gene encoding OPG as an example, we propose that allelic imbalances in risk SNPs on non-transcribed loci have an effect on the regulation of transcribed SNPs.

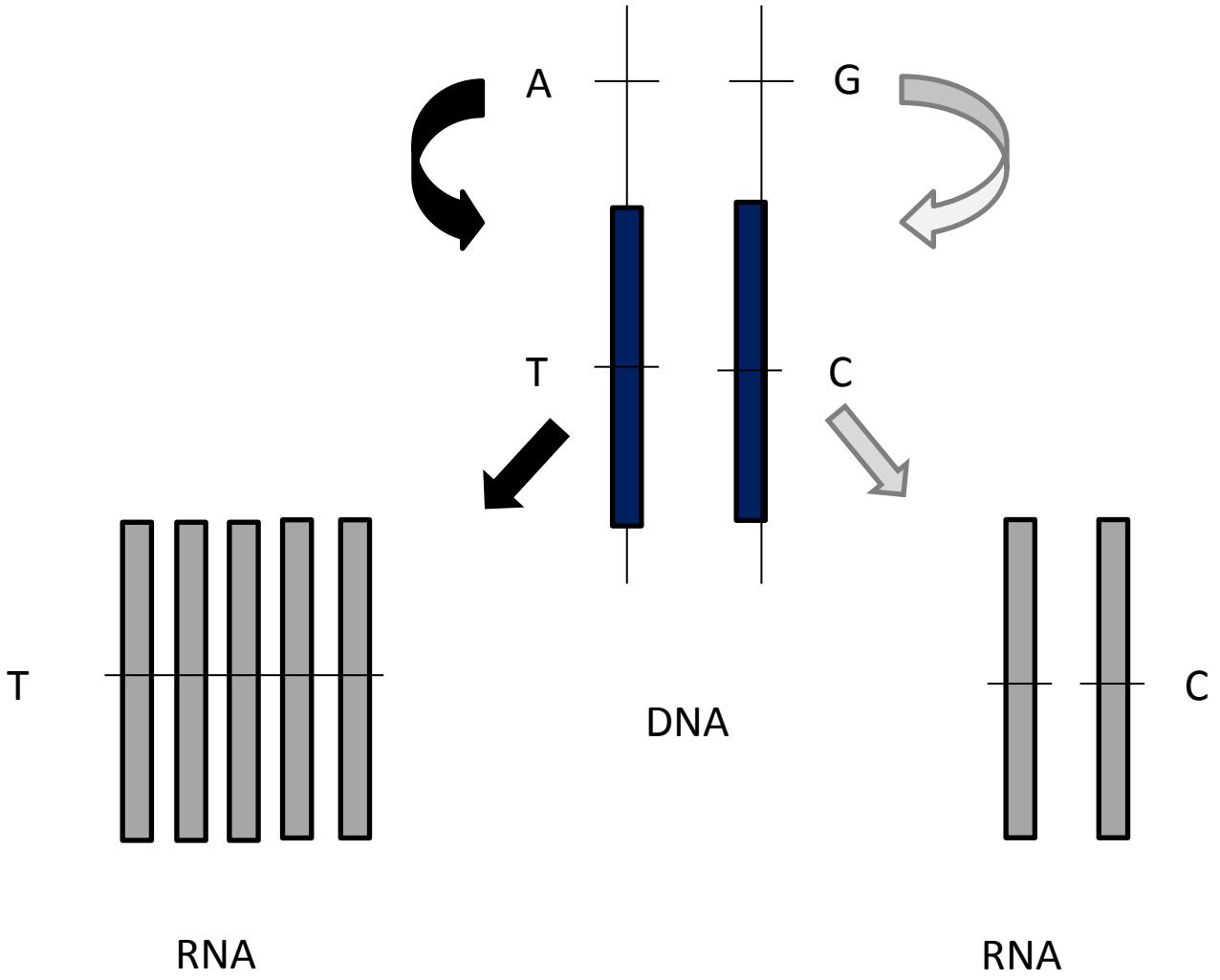


Figure 5-1: Effect of a cis acting polymorphism on allelic expression

The situation of a cis acting SNP with allele A and G, transcribed SNP with allele T and C, and affected transcript in an individual who is heterozygous for cis acting and transcribed SNPs. As it is shown A as a cis acting SNP have effected on one of the haplotype and induced more transcript which these affected transcribed are identified by allele T of transcribed SNP. Correlations between cis acting SNPs and AEI of transcribed SNPs in the samples which were heterozygous for transcribed SNPs were then determined. One of the most important goals of genetics is to identify (*cis*-) regulatory elements for gene expression and heritable variation of them, especially those that impact clinical phenotypes and common diseases (Teare, Pinyakorn et al. 2011).

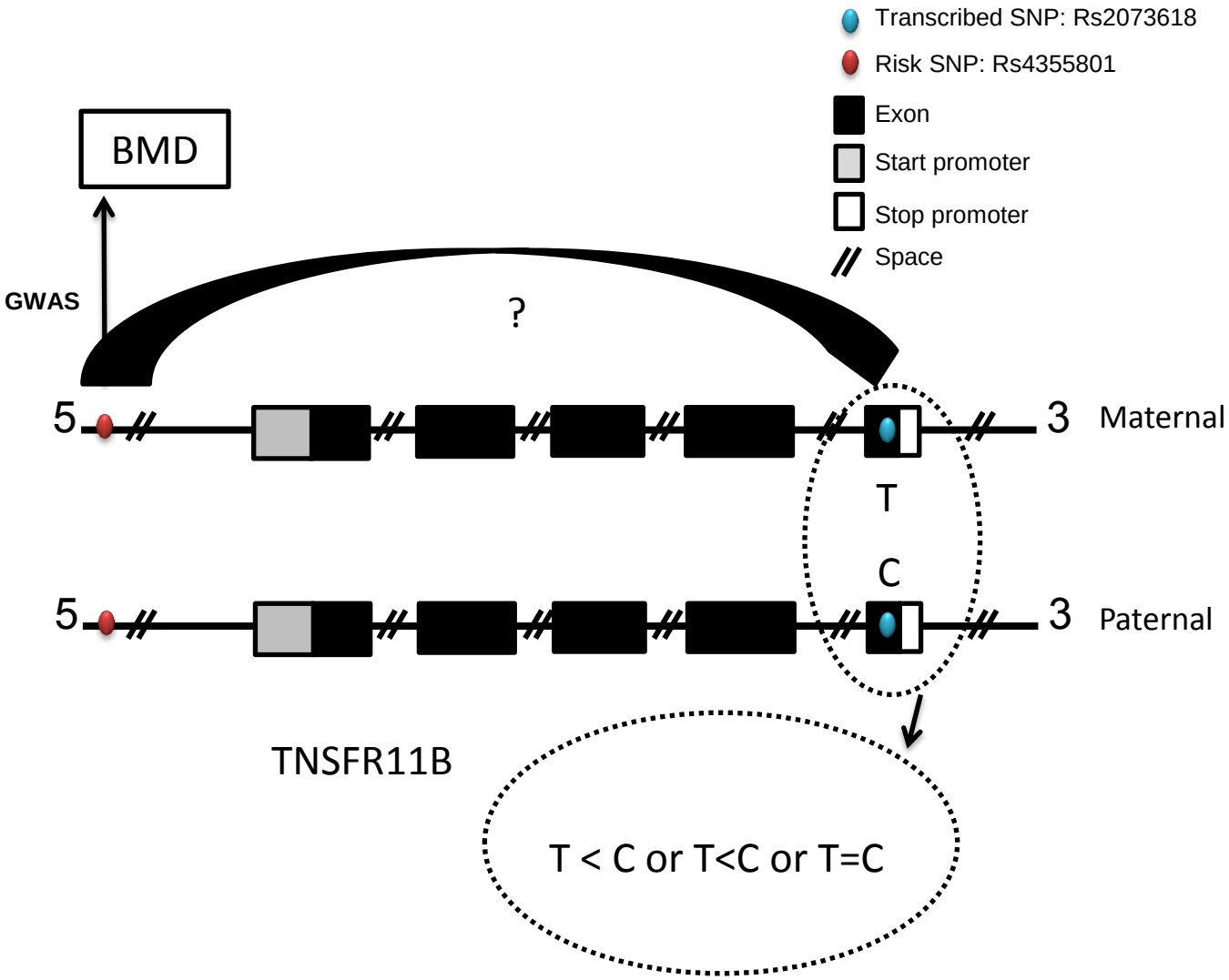


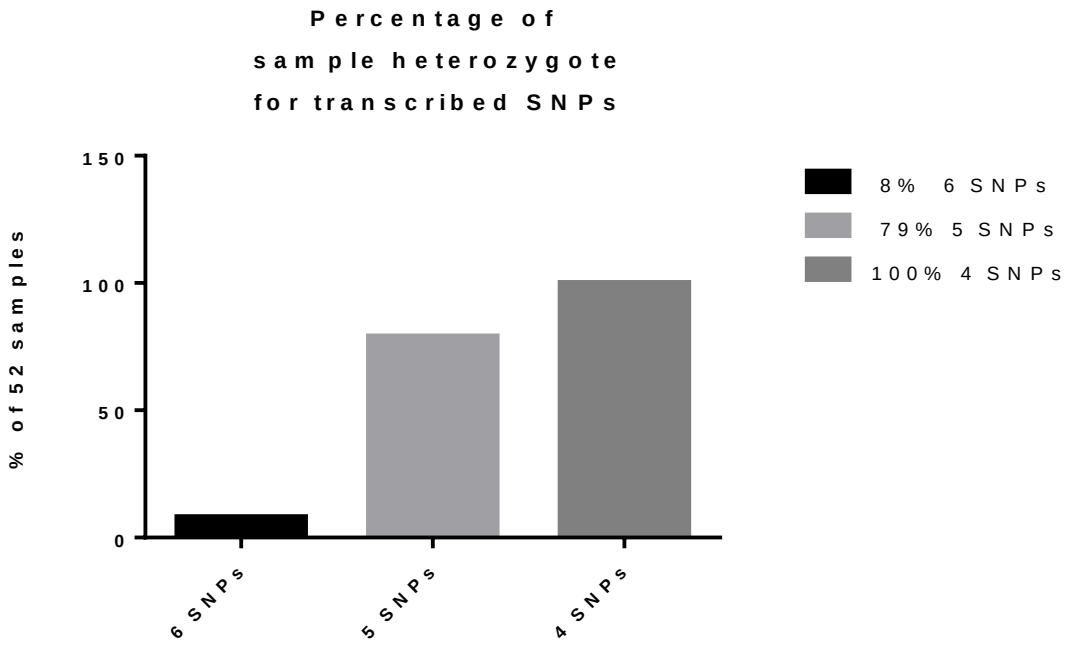
Figure 5-2: Risk SNPs on non-transcribed loci may have an effect on BMD by allelic imbalance expression

Rs4355801 as a risk SNP is associated with BMD according to GWAS. Rs2073618 as a transcribed SNP is as a label to investigate allelic imbalance expression

5.3 Results

5.3.1 Fifty-two candidate samples were selected to be used in the allelic imbalance expression study based on highest heterozygosity

In the first part of this study, blood sampling and BMD were done for 501 randomly selected individuals as described in the **method section 2.1.1.** Genotyping of the peripheral blood samples donated from these 501 volunteers was performed as described in the method **section 2.4.** The **method section 2.5** demonstrates what size of an effect would be detected (with 80% power) with sample sizes of 16 osteoclast cultures and 52 peripheral blood samples. Out of the 501 available volunteer samples, a total number of 52 subjects for blood sampling and 16 subjects for osteoclast culture were selected based on having the highest level of heterozygosity at each of the specified transcribed SNPs. This was performed by sorting the samples according to the number of heterozygous SNPs of the six transcribed SNPs in each sample. Six transcribed SNPs are: rs2077647, rs2073618, rs556442, rs1805034, rs208294 and rs731236, as described in the **method section 2.5** (Table 2.4). Analysis of the genotyping data showed that 8% of the total selected 52 subjects were heterozygous for all 6 transcribed SNPs. A larger proportion 79% of the total sample population was heterozygous for a minimum of 5 transcribed SNPs and all 52 subjects were heterozygous for a minimum of 4 transcribed SNPs (**Figure 5.3**). Data showed that out of our selected 52 subjects, 36 subjects were heterozygous for ESR1 rs2077647 SNP, 42 were heterozygous for TNFRSF11 rs2073618 SNP, 40 were heterozygous for LRP5 rs556442 SNP, 46 were heterozygous for P2RX7 rs208294 and 41 heterozygous for the RANK rs1805034 SNP (**Figure 5.3**). Transcribed SNP heterozygosity is important in AEI as it allows the expression of each allele to be quantified separately.



Transcribed SNP for candidate gene	Number of the 52 subjects who are heterozygous for each listed SNP
ESR1	
2077647	42
OPG	
2073618	40
LRP5	
556442	37
RANK	
1805034	44
P2RX7	
208294	38
VDR	
731236	38

Figure 5-3: All the sample candidates selected were heterozygous for at least 4 transcribed SNPs

(A) *Genotyping analysis showed that out of the 52 selected candidate samples, all were heterozygous for a minimum of 4 transcribed SNPs. A smaller population were heterozygous for all 6 SNPs.* ***(B)*** *The number of heterozygotes in the sample for selected transcribed SNPs of ESR1, OPG, LRP5, RANK, P2RX7 and VDR.*

The 52 samples were selected for efficiency. With more resources the total sample would have been extended to ensure there was sufficient power (80%) to detect cis-acting effects on AEI at all the candidate loci of interest. With these limited resources the samples were selected to maximise the number of loci that could be examined. Hence subjects who were heterozygous at more of the transcribed SNPs were more desirable. However the test for cis-acting effects relies on there being homozygous and heterozygous at the putative cis-acting loci, so before the allelic expression imbalance measurements were taken the breakdown of the genotype counts for each risk SNP in each candidate gene was examined to ensure there were sizeable numbers of homozygotes and heterozygotes for each risk SNP in those heterozygous for the transcribed SNPS. The next section shows this breakdown by candidate gene.

5.3.2 The joint distribution between the transcribed SNP and risk SNP heterozygosity

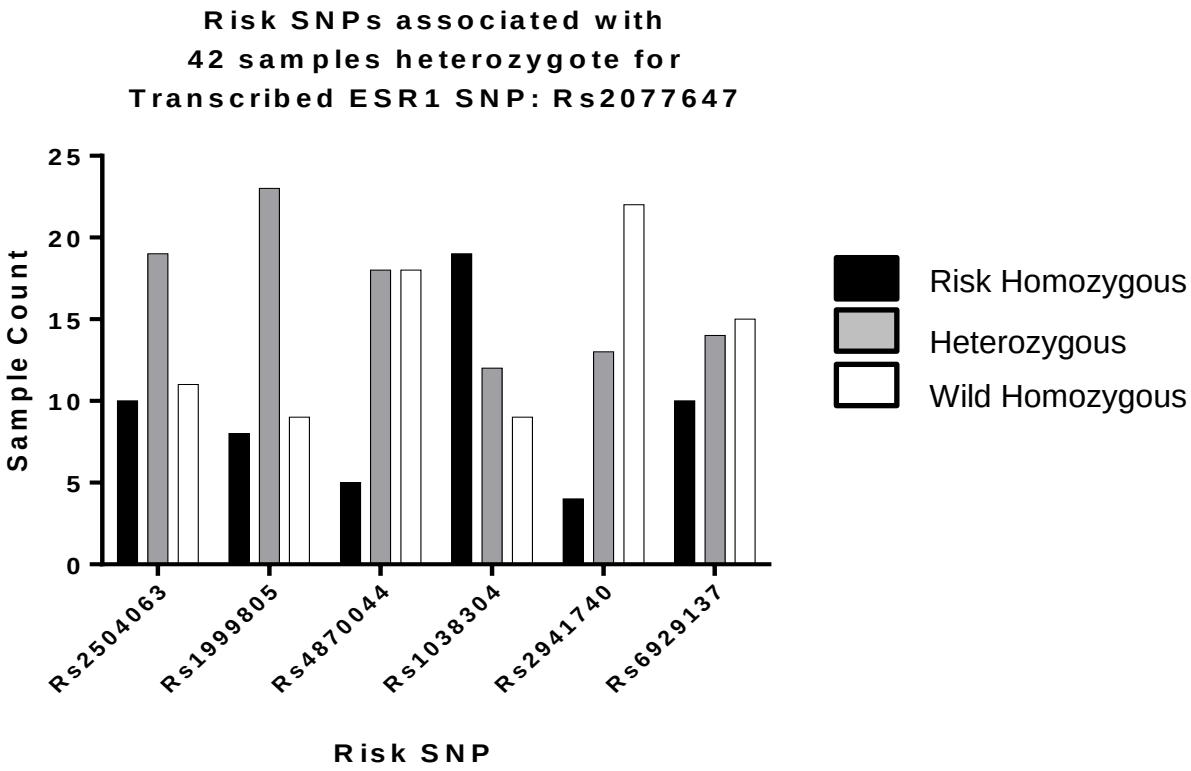
AEI is only measurable in subjects that are heterozygous for the transcribed SNP. The power of the test to look for evidence for the cis-acting effects comes from comparing the distributions of the AEI by risk SNP genotype. An excel pivot table was used to cross tabulate the genotyping data of selected transcribed and risk SNPs within the 52 selected subject samples, as outlined in the **method section 2.6**. This shows how many of those who are heterozygous for the transcribed SNP belong to each risk SNP genotype group. For each of the ESR1, OPG, LRP5, RANK, P2RX7 and VDR candidate genes, all genotyping data of transcribed and risk SNPs were transferred into excel worksheets to form the pivot table. The numbers of those with specific risk SNP genotypes between each of the transcribed SNPs and risk SNPs genotype were then assessed.

The genotype count breakdown of the 52 samples was used to ensure there were sufficient numbers of subjects who were heterozygous for both transcribed and risk SNPs. Here we firstly identified, out of our 52 candidate samples, how many were heterozygous for the transcribed SNP of our candidate gene. Next we constructed a contingency table to illustrate how many samples were risk allele homozygous, heterozygous or wild-type homozygous. Where genotyping was unsuccessful due to methodological issues, the risk SNP was

categorised as having an 'undetermined' genotype. The number (count) of samples that was either risk homozygous, heterozygous or wild homozygous within our sample population was summarised. A profile was then created for each of the candidate transcribed SNPs in relation to the genotype of the associated risk SNPs.

The genotype profile for the candidate samples heterozygous for transcribed ESR rs2077347 SNP and risk SNPs rs2504063, rs1999805, rs4870044, rs1038304, rs2941740 and rs6929137 is presented. Data showed that for the transcribed SNP rs2077647 the candidate sample population were heterozygous, risk homozygous and wild homozygous for all risk SNPs within this gene (Figure 5.4). The genotype profile for the candidate samples heterozygous for transcribed OPG rs2073618 SNP and risk SNPs rs6469804 rs4355801, rs2062375, rs2062377 and 11995824 is presented in (Figure 5.5). The majority of samples that were heterozygous for transcribed SNP rs2073618 were also heterozygous for rs6469804, rs4355801, rs2062375, rs2062377 and rs11995824. Figure 5.6 shows the candidate samples heterozygous for transcribed LRP5 rs556442 SNP and risk SNPs rs599083 and rs3736228; no samples were risk homozygous and the majority of samples were heterozygous.

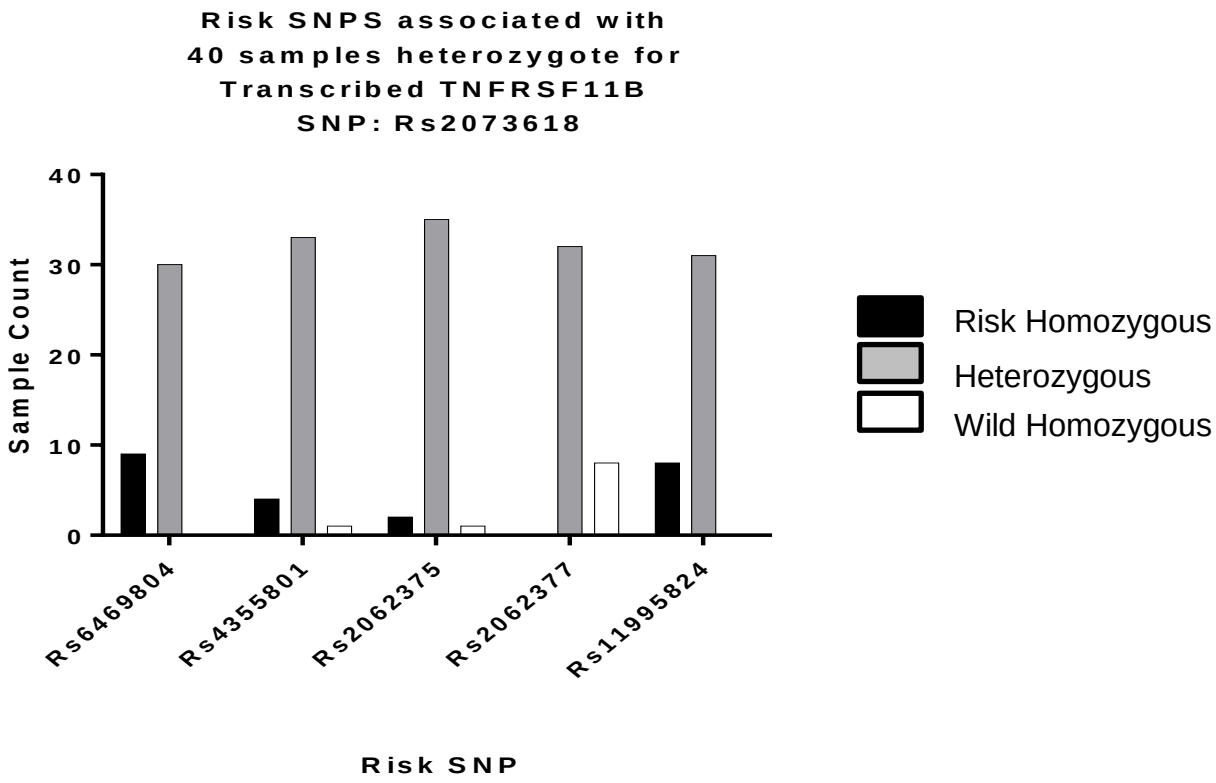
The genotype profile for the candidate samples heterozygous for transcribed RANK rs1805034 SNP and risk SNPs rs884205 and rs3018362 was presented (Figure 5.7). For the transcribed SNP rs1805034, the majority of samples were wild homozygous for risk SNPs rs884205 and rs3018362. For P2RX7, genotyping analysis showed that for the transcribed SNPs rs7958311 and rs1718119, the candidate sample population were either heterozygous or wild homozygous for the risk SNP rs3751143 (Figure 5.8). The genotype profile for the candidate samples heterozygous for transcribed VDR rs731236 SNP and risk SNP rs2228570 showed that for the transcribed SNP rs731236, the candidate sample population were either heterozygous or risk homozygous for risk SNP 2228570 (Figure 5.9).



ESR1					
Number of Heterozygous samples for transcribed Rs2077647 SNP	**Risk SNP**	**Risk homozygous**	**Heterozygous**	**Wild homozygous**	**Undetermined**
42	**2504063**	10	19	11	2
	1999805	8	23	9	2
	4870044	5	18	18	1
	1038304	19	12	9	2
	2941740	4	13	22	3
	6929137	10	14	15	3

Figure 5-4:The genotype profile for the candidate samples heterozygous for transcribed ESR rs2077347 SNP and risk SNPs rs2504063, rs1999805, rs4870044, rs1038304, rs2941740 and rs6929137

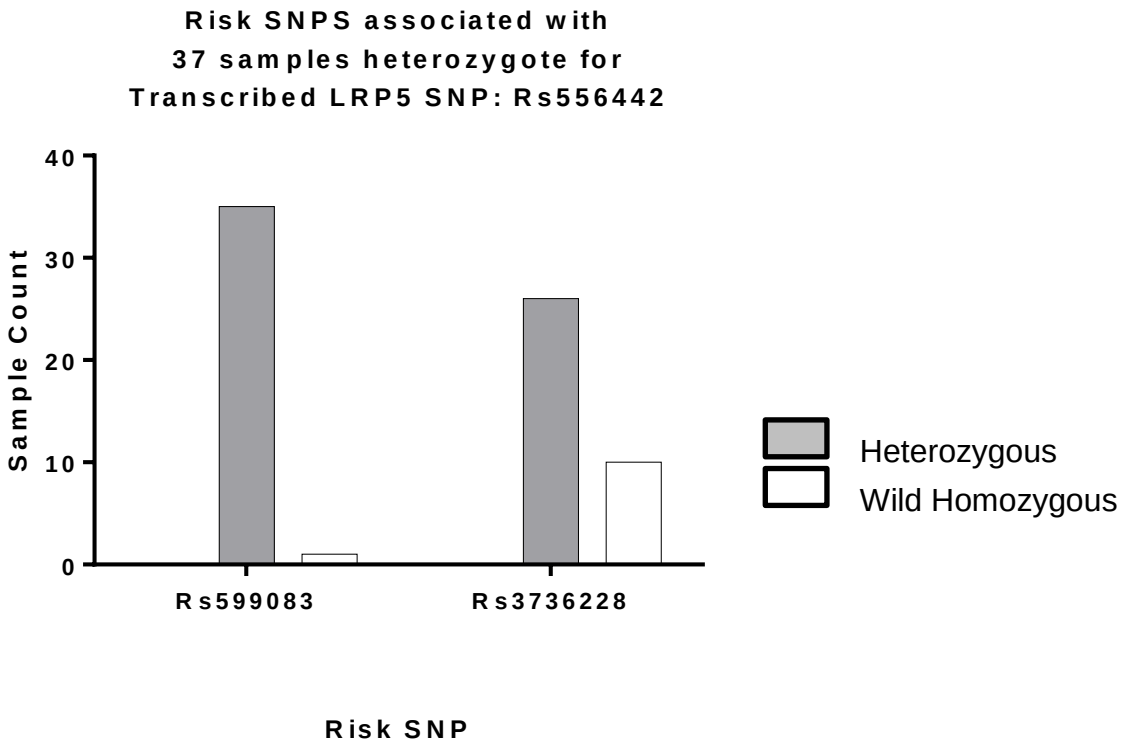
Genotyping analysis showed that for the transcribed SNP rs2077647, the candidate sample population were subdivided into heterozygous, risk homozygous and wild homozygous.



OPG					
Number of Heterozygous samples for transcribed Rs2073618 SNP	**Risk SNP**	**Risk homozygous**	**Heterozygous**	**Wild homozygous**	**undetermined**
40	**6469804**	**9**	**30**	**0**	**1**
	4355801	**4**	**33**	**1**	**2**
	2062375	**2**	**35**	**1**	**2**
	2062377	**0**	**32**	**8**	**0**
	11995824	**8**	**31**	**0**	**1**

Figure 5-5: The genotype profile for the candidate samples heterozygous for transcribed OPG rs2073618 SNP and risk SNPs rs6469804 rs4355801, rs2062375, rs2062377 and 11995824

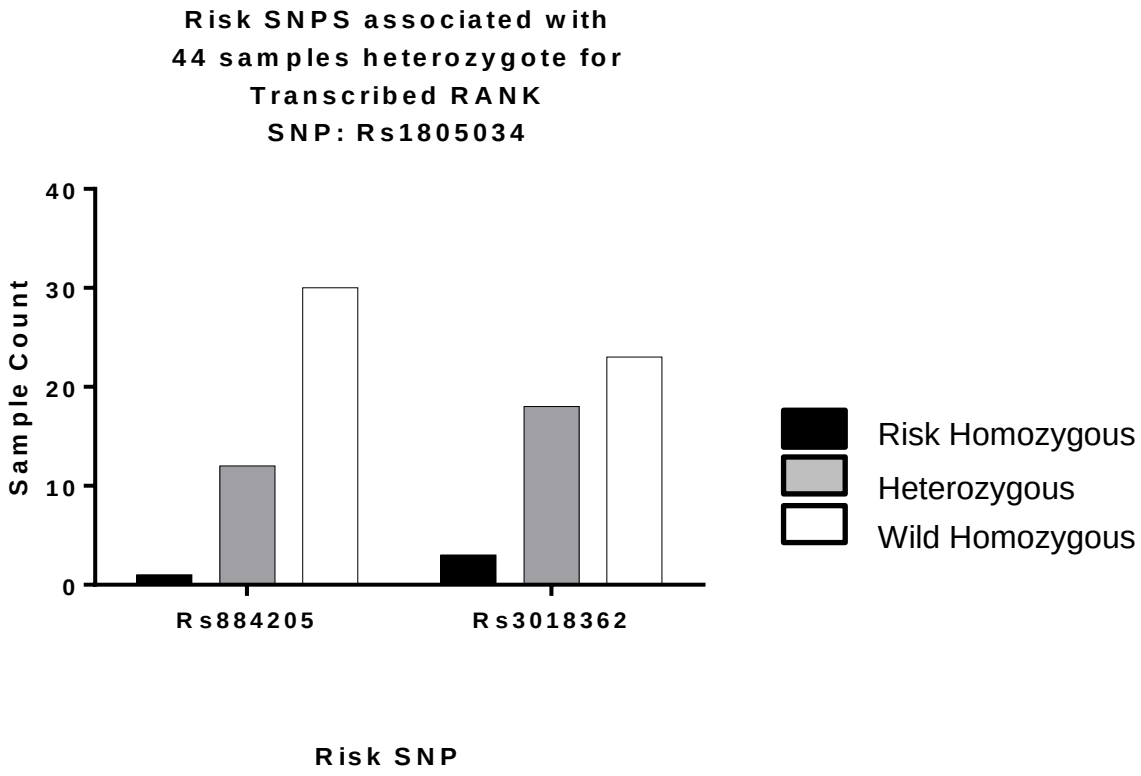
Genotyping analysis showed that for the transcribed SNP rs2073618, the candidate sample population were subdivided into heterozygous, risk homozygous and wild homozygous. The majority of samples were heterozygous for transcribed SNP rs2073618 were also heterozygous for rs6469804, rs4355801, rs2062375, rs2062377, and 11995824.



LRP5					
Number of Heterozygous samples for transcribed Rs556442 SNP	**Risk SNP**	**Risk homozygous**	**Heterozygous**	**Wild homozygous**	**Undetermined**
37	**599083**	**0**	**35**	**1**	**1**
	3736228	**0**	**26**	**10**	**1**

Figure 5-6: The genotype profile for the candidate samples heterozygous for transcribed LRP5 rs556442 SNP and risk SNPs sn599083 and 3736228

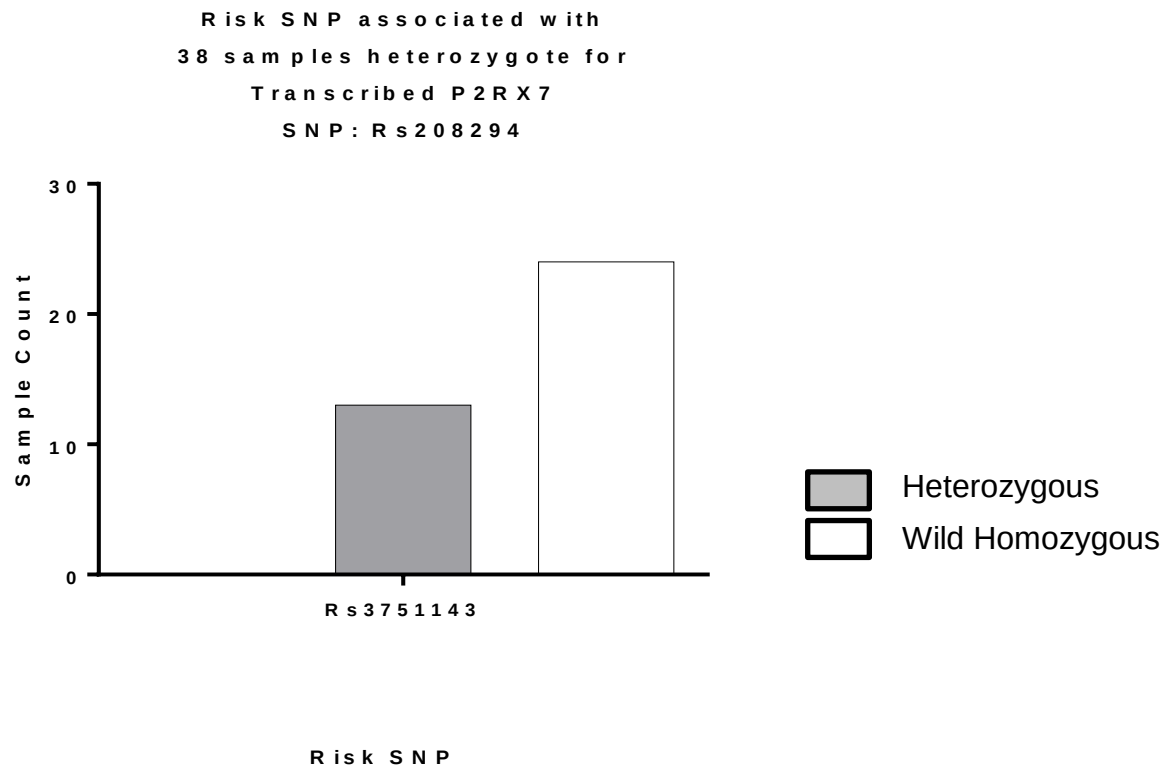
Genotyping analysis showed that for the transcribed SNP rs556442, no samples were risk homozygous and the majority of samples were heterozygous.



RANK					
Number of Heterozygous samples for transcribed Rs1805034 SNP	**Risk SNP**	**Risk homozygous**	**heterozygous**	**Wild homozygous**	**undetermined**
44	**884205**	**1**	**12**	**30**	**1**
	3018362	**3**	**18**	**23**	**0**

Figure 5-7: The genotype profile for the candidate samples heterozygous for transcribed RANK rs1805034 SNP and risk SNPs rs884205 and rs3018362

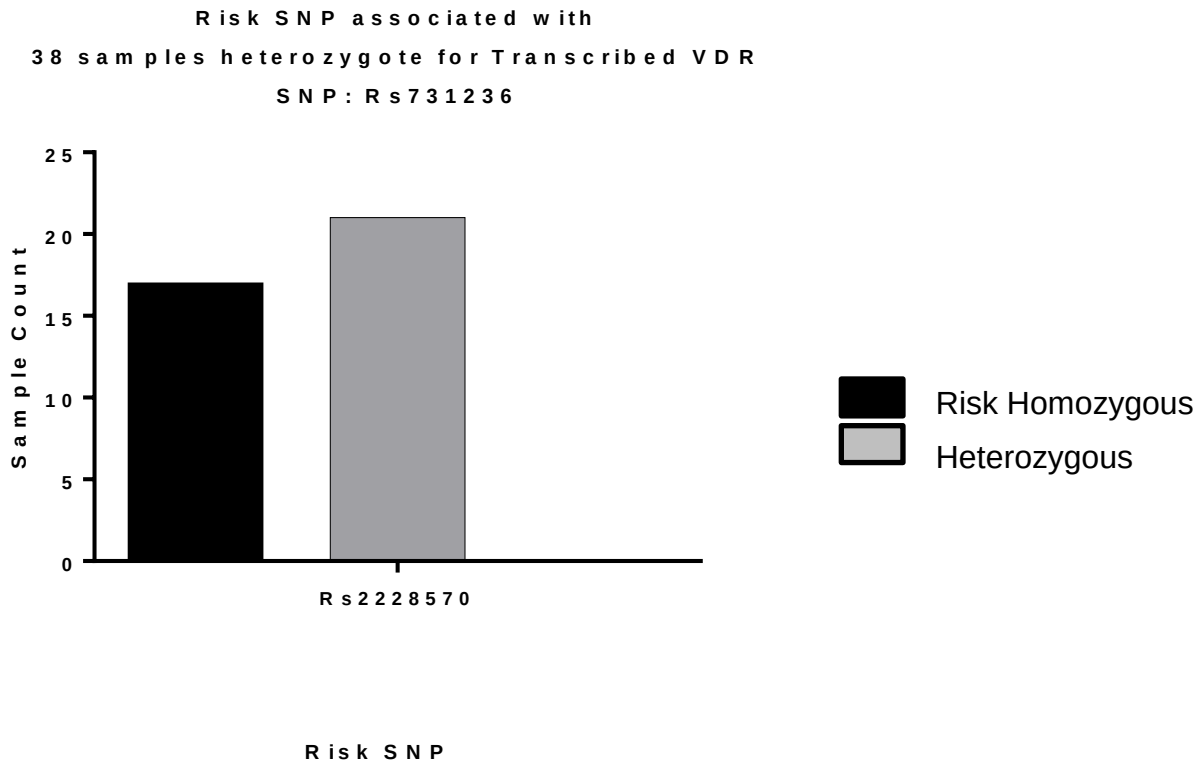
Genotyping analysis showed that for the transcribed SNP rs1805034, the candidate sample population could be subdivided into heterozygous, risk homozygous and wild homozygous. The majority of samples were wild homozygous for risk SNPs rs884205 and rs3018362.



P2RX7					
Number of Heterozygous samples for transcribed Rs208294 SNP	**Risk SNP**	**Risk homozygous**	**heterozygous**	**Wild homozygous**	**undetermined**
38	**3751143**	**0**	**13**	**24**	**1**

Figure 5-8 : The genotype profile for the candidate samples heterozygous for transcribed P2RX7 rs208294 and risk SNP rs3751143

Genotyping analysis showed that for the transcribed SNPs rs298294, the candidate sample population were either heterozygous or wild homozygous the risk SNP rs3751143.



VDR				
Number of Heterozygous samples for transcribed Rs731236 SNP	**Risk SNP**	**Risk homozygous**	**Heterozygous**	**Wild homozygous**
38	**2228570**	**17**	**21**	**0**

Figure 5-9: The genotype profile for the candidate samples heterozygous for transcribed VDR rs731236 SNP and risk SNP rs2228570

Genotyping analysis showed that for the transcribed SNP rs731236, the candidate sample population were either heterozygous or risk homozygous for risk SNP 2228570

The data presented in the above tables shows the number of subjects in each genotype for risk SNPs (homozygous, heterozygous) for individuals who are heterozygous for transcribed SNPs, which is essential to detect cis-acting effect. Knowledge of the presence of linkage disequilibrium between risk SNP and transcribed SNPs is also important and influences the method of analysis to detect cis-acting effects (Teare, Pinyakorn et al. 2011).

5.3.3 Linkage disequilibrium between selected ESR1, RANK, VDR, LRP5, OPG and P2RX7 SNPs

Linkage disequilibrium (LD) occurs when there is non-random association of alleles at two or more loci (Slatkin 2008). In populations where the occurrence of some combinations of alleles is more frequent than would be expected from a random formation of haplotypes from alleles based on their frequencies were said to be in linkage disequilibrium.

For example, let us assume there are two loci: locus 1 has alleles G and C, locus 2 has alleles A and T. It is considered that p1 and p2 are frequency of the alleles at locus 1 and q1 and q2 are frequency of alleles at locus 2.

SNP1		SNP2	
Allele	Frequency	Allele	Frequency
G	p1	A	q1
C	P2	T	q2

It is supposed that haplotype frequencies are as follows:

Haplotype	Frequency	Haplotype	Frequency
GA	p11	GT	p12
CA	p21	CT	q22

When haplotype frequencies are equal to their corresponding allele frequencies, it shows the loci are in the linkage equilibrium situation.

Haplotype frequency		Product of allele frequency
p11	=	p1q1
p12	=	p1q2
p21	=	p2q1
p22	=	p2q2

We now illustrate this phenomenon with a hypothetical example. It is assumed that samples of 100 subjects were taken and the following haplotype counts have been found:

GA: 25; GT: 50; CA: 60; CT: 65. Total=200

Haplotype frequency and allele frequency are as below:

Haplotype frequencies Allele frequencies

GA= 25/200= .125 p1= .375

GT=50/200= .25 p2= .625

CA=60/200= .3 q1= .325

CT=65/200= .325 q2= .575

After inputting values in the equation for D to calculate linkage disequilibrium we will have the result below:

D = (P11* P22) – (P12* P21)

D= (.25*.325) - (.125*.3) = .04375

Depending on allele frequency of two loci sometimes the value of D can be negative but in reality frequencies cannot be negative so standardization methods have been proposed to overcome this issue. A relative measure of disequilibrium (D) compared to its maximum is used as a standardization method. The first of these measures is D’ where D is divided by the maximum value it could take (Dmax).

To estimate Dmax allele frequencies and value for D were inputted in the following equation:

Dmax = min [(p1q2) or (p2q1)]

Dmax = (.375*.575) = .215

D'= D/Dmax

D' =.04375/.215= .2

Pearson coefficient of correction (r) is another commonly used measured to calculate LD between loci. However squared coefficient of correlation (R^2) is used to remove arbitrary sign.

$R= D/ (p1p2q1q2)^{1/2}$

$R= .043/ (.375*.625*.325*.575)^{½}$

R= .043/.2= .215

D' and R^2 values describe the extent of the correlation between a pair of loci, and range from 0 to 1.00, where for D', 0 means that there is complete linkage equilibrium and 1.00 means complete LD. The R^2 values are also ranged from 0 to 1.00, where 0 means that there is complete linkage equilibrium and 1.00 means perfect LD. In this study we were not able to count haplotypes exactly so we used the software Haploview software version 4.0 (Barrett, Fry et al. 2005) to calculate D' and R^2 for this study by using genotyping data. Haploview software reports D' and R^2 values as percentage. When D'= 1 is referred to as Complete LD, there are at most 3 of the 4 possible haplotypes present in the populations. It means in complete LD, two loci are not being separated by a recombination in population since at least one of the haplotypes does not occur in the population. When R^2= 1 is referred to as Perfect LD, there are exactly 2 of the 4 possible haplotypes present in the population, and as a result, the two loci also have the same allele frequencies. Loci which are in perfect LD are also in complete LD (VanLiere and Rosenberg 2008).

If two loci are independent (not co-inherent) then both the D´ and R^2 values will be 0.0 regardless of either allele frequency. If two polymorphisms are in total disequilibrium both with a 50% allele frequency then both the D´ and R^2 values would be 1.0 but when the allele frequencies are not the same the result is different. For example, if there are two polymorphisms in total disequilibrium, one with a 50% allele frequency and the other with a

1% allele frequency the D´ value would be 1.0 but the R^2 value would only be 0.01. So ranges of R^2 value can depend on the frequencies of alleles at the loci under consideration. The range of D' is frequency-independent (Hedrick 1987).

The data in the previous section was shown to ensure that there were sufficient heterozygotes in the sample which was analysed for allelic expression imbalance. This was a small subset of the full population studied and the extent of linkage disequilibrium between the pairs of SNP in each candidate region can be examined in the full set of 501 samples.

The form the statistical test takes depends upon the extent of LD between the transcribed and the risk SNP. Four scenarios are shown in Figure 5.5. As AEI cannot be measured in homozygous transcribed SNPs these figures include only individuals who are heterozygous. Figure 5.10.A shows what we might see if two loci were only in weak or no LD so that all four haplotypes exist and are directly measureable. In this scenario there is a cis-acting effect. When 'c' is in phase with 'm' the ratio tends to be below μ_0 and when 'c' is in phase with 'M' the ratio tends to be above μ_0. In fact we cannot distinguish between the two phases so when a cis-acting effect exists there tends to be a wider spread of values in the heterozygous group. The program implemented by MSK jointly estimates the haplotype phase and the parameter μ when phase is not known.

Fig 5.10.B depicts a scenario where there is perfect disequilibrium between both polymorphisms or the cis-acting and the transcribed polymorphisms are one and the same (such as rs731236 in this study) when there is perfect LD and a cis-acting effect then one specific allele is overexpressed consistently. Investigating the effect shows a systematic deviation in one direction from balanced expression. In figure 5.10 C a complete disequilibrium situation is shown. As explained above, in complete disequilibrium there are at most 3 of the 4 possible haplotypes present in the populations. In this figure a functional polymorphism is expected to make a systematic overexpression of one and the same transcribed allele among heterozygotes. This overexpression is not seen in homozygotes.

If phase is known the genotype groups can be rearranged to place the two types of heterozygotes in groups either side of the homozygotes as in Figure 5.10.D. When rearranged to this scheme the parameter μ_{cis} can be estimated by simple linear regression. So, μ_{cis} is the average increase in allelic expression when the cis-acting element is in phase with the transcribed allele. When there is only partial LD then it is not generally possible to classify the phase of double heterozygotes, so a simple group comparison or linear regression cannot

be done. Hence when the phase of compound heterozygotes cannot be determined a specialist program is required which can estimate the phase of the haplotypes and estimate the μ_{cis} at the same time. In this study we had measured genotypes only, so we conducted an analysis of linkage disequilibrium to see what methods would be necessary in the following the AEI measurements.

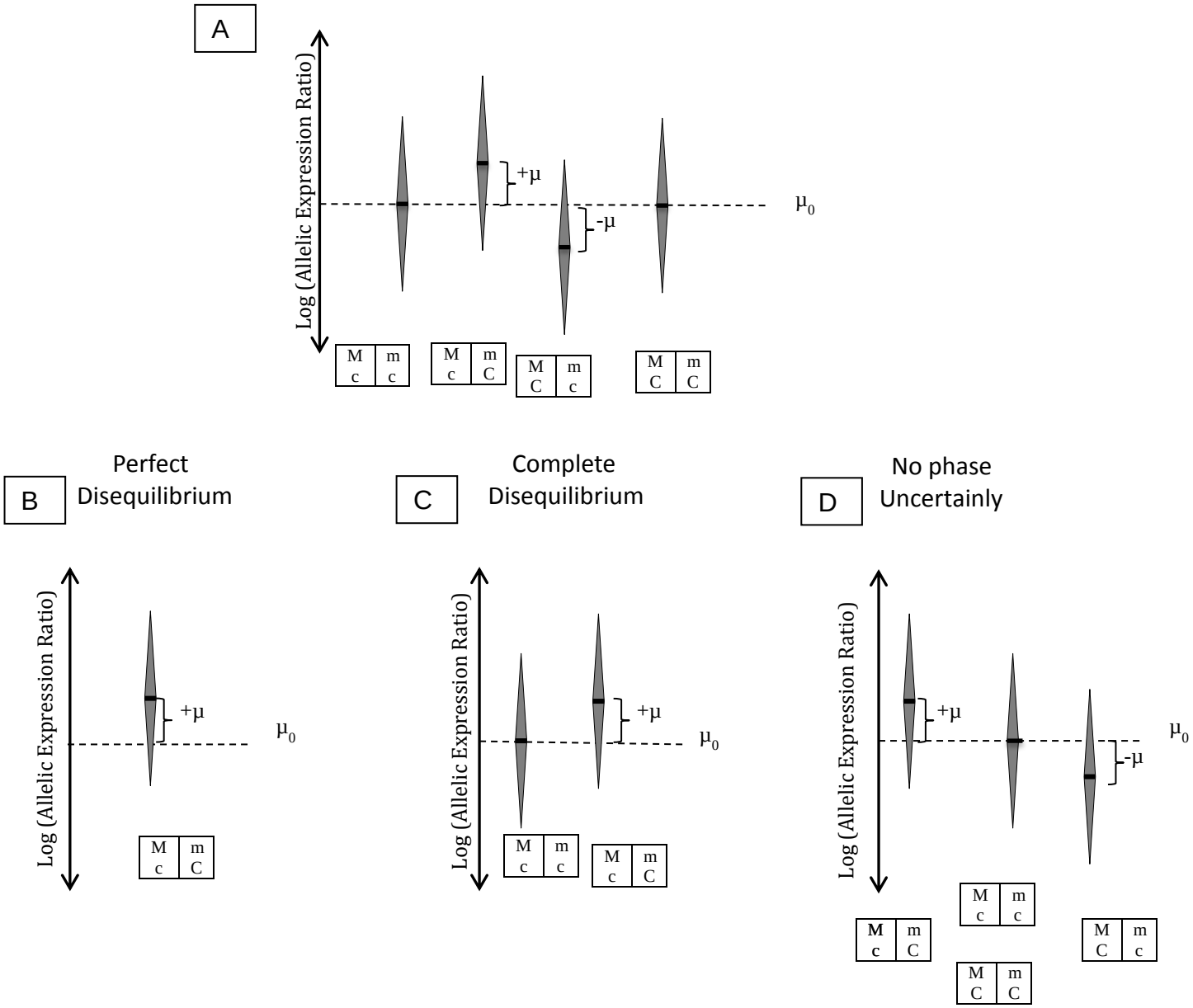


Figure 5-10: Different scenario for testing an association between two allelic polymorphisms and allelic expression

The distribution of allelic expression ratios across a population is shown. Two polymorphisms are considered here: 1) Cis-acting SNP with alleles: C and c. 2) Transcribed SNP with two alleles: M and m, used as a marker to determine allelic expression. Each diamond represents the spread and the mean of the AEI measurements by specific genotypes. A) All possible conditions. B) Perfect disequilibrium (R^2 = 1, D'=1) between the transcribed and the cis-acting polymorphism, in this situation only two distinct haplotypes happen. C) Complete disequilibrium (R^2< 1, D'=1), in this situation three distinct haplotypes happen. D) Situation when there is no phase between cis and transcribed alleles or it is unknown (Teare, Pinyakorn et al. 2011). This figure is explanatory without data.

The allelic expression imbalance was only assessed on a maximum of 51 samples. These were a selected sample group from the larger population so the linkage disequilibrium analysis was performed on the full set of 501 genotyped samples obtained from volunteers as described in **section 2.4**. In this part of the population genetics study, the extent of LD between selected SNPs within our candidate genes: ESR1, OPG, LRP5, RANK, P2RX7 and VDR was summarised. For each candidate gene region, pairwise LD was summarised with D' and R^2 statistic. Though D' and R^2 are constrained to take values between 0 and 1, the Haploview software reports them as a percentage.

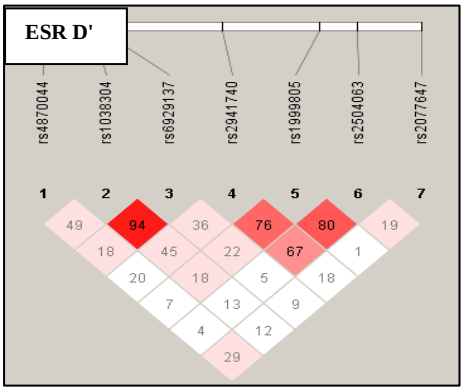
The pairwise LD was calculated for the ESR risk SNPs rs4870044, rs1038304, rs6929137, rs2941740, rs1999805, rs2504063 and transcribed SNP rs2077647 (**Figure 5.11 A**). The D' values are less than 29 between the risk SNPs and transcribed SNPs. The R^2 values are less than 4 between the risk SNPs and transcribed SNPs. So there is no complete or perfect LD between the transcribed SNP and risk SNPs. The risk SNPs rs1999805 and rs2504063 had the highest R^2 value of 50 and therefore were associated with an increased level of LD. The R^2 value calculated for the RANK between the transcribed SNP rs1805034 and the risk SNPs were less than 16 and the D' values between the transcribed SNP rs1805034 and the risk SNPs were less than 91 which show there is no complete or perfect LD between transcribed and risk SNPs (**Figure 5.11 B**). The D' and R^2 calculated for VDR SNPs rs731236 and rs228570 were 1 and 27 and these were associated with low LD between VDR SNPs (**Figure 5.11 C**).

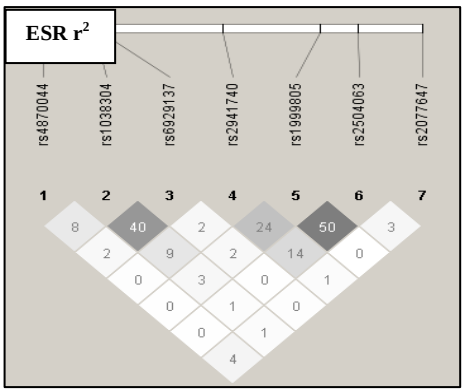
The LD was also calculated for the OPG risk SNPs: rs4355801, rs2062375, rs2062377, rs11995824, rs6469804 and transcribed SNP rs4355801 (**Figure 5.11 D**). Interestingly, the D' and R^2 values calculated showed high LD between the transcribed SNP rs4355801 and risk SNPs. The highest D' value was observed between transcribed SNP rs4355801 and risk SNPs: rs2062375 and rs2062375 with a calculated D' value of 95.The highest R^2 value was observed between transcribed SNP rs4355801 and risk SNP rs2062375 with a calculated r^2 value of 83. Because both D' and R^2 values are below 100 again there is no perfect or complete LD between transcribed and risks SNPs.

For the P2RX7 gene, the LD was determined for risk SNPs rs795811, rs1718119, rs2230911, rs22309912, rs3751143 and rs1621388 and transcribed SNP rs208294 (**Figure 5.11 E**). The D' and R^2 values calculated showed D' $\leq$ 76 and R^2 $\leq$13 between the transcribed SNP rs208294 and any of the risk SNPs. However, a perfect LD represented by a D' and R^2 value

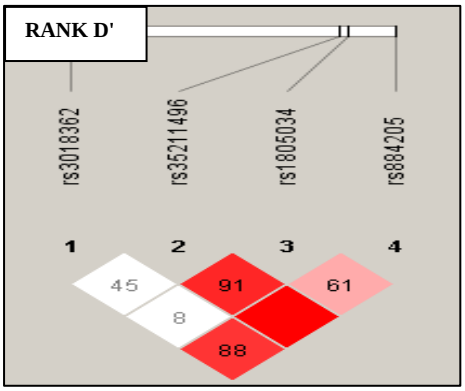
of =100, was observed between risk SNPs rs1718119 and rs1621388. Also risk SNPs: rs1718119, rs2230911, 2230912, 3751143, and 1621388 are in complete LD with D' =100. Finally the LD associations were determined for LRP5 risk SNPs rs599083, rs3736228, and transcribed SNP rs545382 (**Figure 5.11 F**). Here D' calculations determined a complete LD between transcribed SNPs rs545382 and risk SNPs599083 (D'=100). However, R^2 value between the transcribed SNP rs545382 and either risk SNP were low (R^2 of $\leq$24).

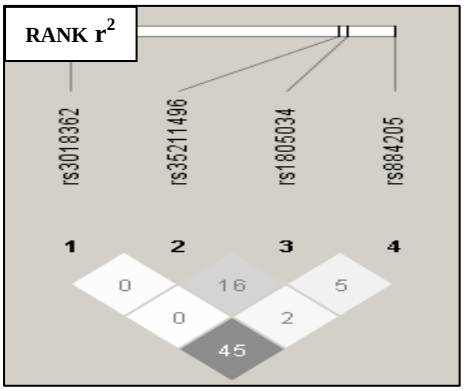
A


ESR D'
rs4870044
rs1038304
rs6929137
rs2941740
rs1999805
rs2504063
rs2077647
1 2 3 4 5 6 7
49 94 36 76 80 19
18 45 22 67 1
20 18 5 18
7 13 9
4 12
29

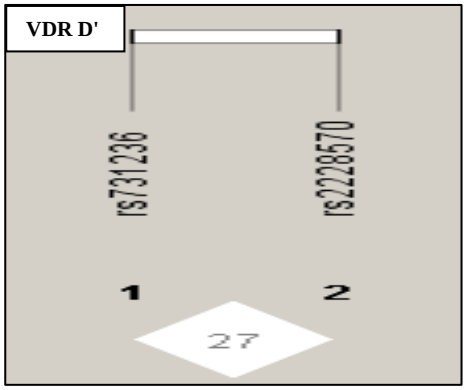

ESR r2
rs4870044
rs1038304
rs6929137
rs2941740
rs1999805
rs2504063
rs2077647
1 2 3 4 5 6 7
8 40 2 24 50 3
2 9 2 14 0
0 3 0 1
0 1 0
0 1
4

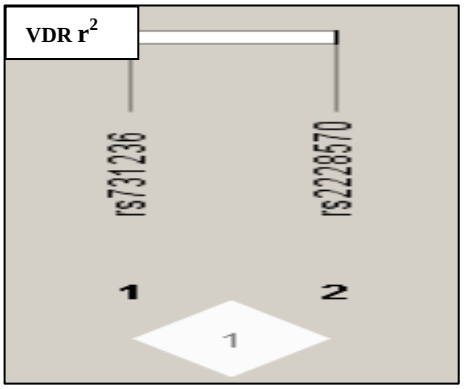
B


RANK D'
rs3018362
rs35211496
rs1805034
rs884205
1 2 3 4
45 91 61
8 88

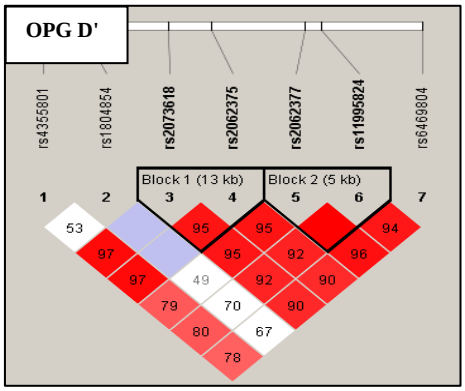

RANK r2
rs3018362
rs35211496
rs1805034
rs884205
1 2 3 4
0 16 5
0 2
45

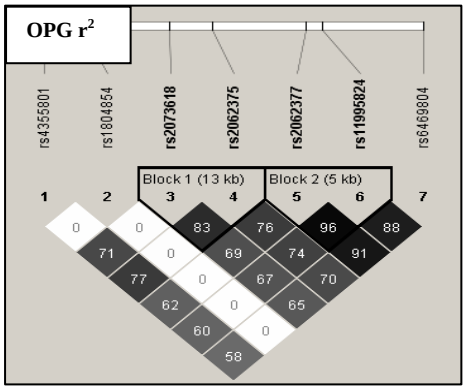
C


VDR D'
rs731236
rs2228570
1 2
27

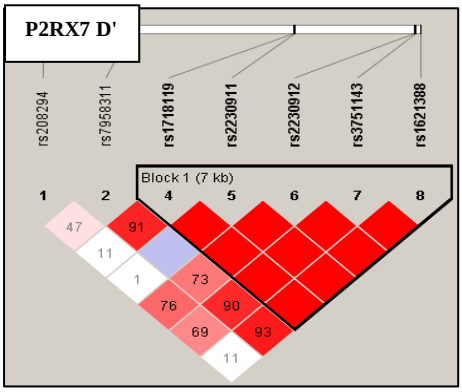

VDR r2
rs731236
rs2228570
1 2
1

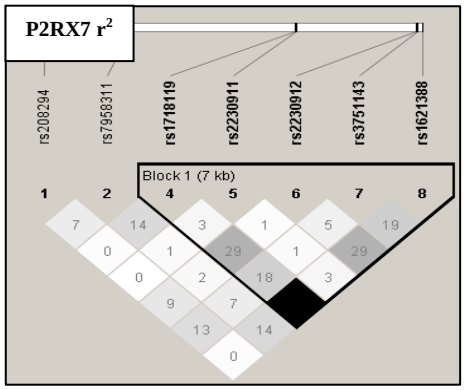
D



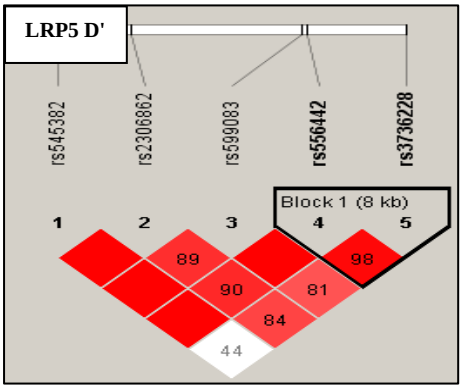


E





F



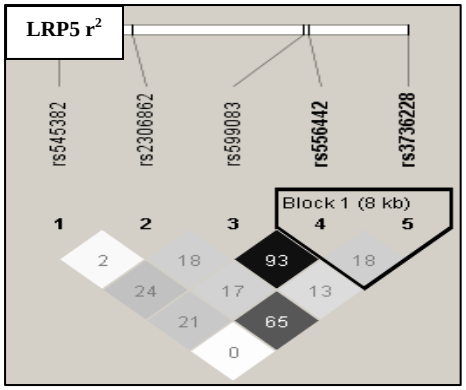


Figure 5-11: Linkage disequilibrium associations between transcribed and risk SNPs with D' and R^2 values.

5.3.4 White blood cells and osteoclasts express *ESR1, OPG, LRP5, RANK, P2RX7* and *VDR*

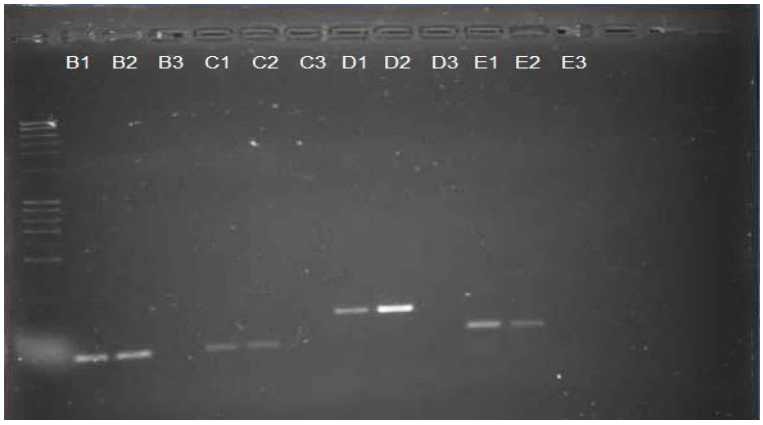
A number of gene markers including *ESR1, OPG, LRP5, RANK, P2RX7, and VDR* have been reported in the literature as genes which are expressed by white blood cells and *ESR1, LRP5, RANK, P2RX7, and VDR* by osteoclast cells. In this study I wanted to determine the expression of these genes in the white blood cell and osteoclast cultures collected from the volunteer samples by RT-PCR, prior to performing the AEI analysis. Positive expression of these genes in our candidate samples was an essential prerequisite for AEI.

RNA was extracted from the peripheral blood and osteoclast culture of a healthy human volunteer. RNA extraction and cDNA synthesis were performed as described in the **method sections 2.8 and 2.9.1**. End point PCR was performed to determine the expression of *ESR1, OPG, LRP5, RANK, P2RX7* and *VDR* in cDNA isolated from white blood cell and osteoclast cell cultures using Promega GoTaq® DNA polymerase as described in the **method section 2.9.2**. Amplified PCR products were detected on an agarose gel using electrophoresis as described in the method section. Detection of End-point PCR products showed the presence of *ESR1, OPG, LRP5, RANK, P2RX7* and *VDR (***Figure 5.12).**

A

RANK Expression

B



P2RX7, LRP5, VDR and ESR Expression

C

OPG Expression

Figure 5-12:ESR1, OPG, LRP5, RANK, P2RX7, and VDR were detected in osteoclast and white blood cell cultures by end-point PCR

PCR products were visualised alongside a 100bp ladder. Gel images are representative of three independent experimental repeats. (A1) Osteoclast cells were positive for Rank expression in comparison to the RT –ve control (A2). (A3) The cDNA isolated from white blood cell cultures was also positive for Rank expression compared to its corresponding RT-ve control (A4). (A5) No band was detected in no-template control. Data showed that osteoclast cells and white blood cell cDNA were positive P2RX2 **(***B1, B2), LRP5 (C1, C2), VDR (D1, D2) and ESR (E1, E2). No bands were detected for the corresponding negative cDNA controls for P2RX2 (B3), LRP5 (C3), VDR (D3) and ESR (E3). (F1) Osteoclast cells were negative for OPG expression and white blood cell cultures were positive for OPG expression compared to its corresponding negative cDNA control (F3).*

5.3.5 Validation of all blood and osteoclast cDNA samples using qRT-PCR and GAPDH

In this study we wanted to use Sequenom technology to determine AEI in the white blood cell and osteoclast culture samples obtained from our candidates selected from the genotyping study outlined in the **method section 2.5**. The RNA was extracted from the 52 individual white blood cell samples and 16 individual osteoclast culture samples, according to the method outlined in the main **method section 2.8.** Then cDNA was synthesised from 500ng of RNA, as described in the **method section 2.9.1** and stored at -20°C until required. The quality of the cDNA was checked by real time PCR and using GAPDH according to **method section 2.9.3**. The cycle threshold (Ct) values of 52 cDNA of blood samples and 16 cDNA from osteoclast culture were determined. Ct value less than 29 is indicative of abundant target nucleic acid in the sample (Mylvaganam, Zhang et al. 2010). The qRT-PCR analysis showed that that GAPDH expression levels within the blood and osteoclast cells were between 17-25 Ct (**Figures 5.13**), indicating that the level of cDNA was viable for further gene expression analysis.

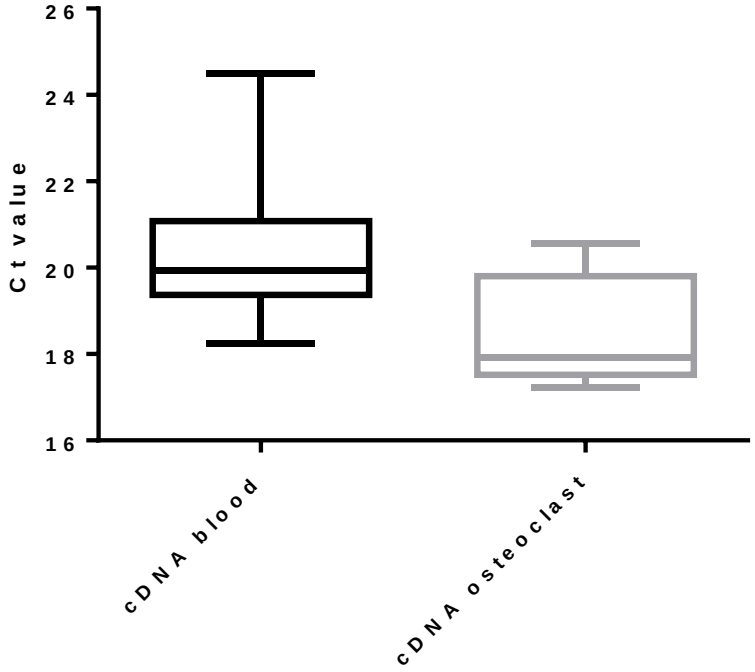


Figure 5-13: The Ct values of 52 cDNA of blood samples and 16 cDNA from osteoclast cultures. **(box-and-whisker plot: the marked line shows the median, and the bottom and top of the box show the lower and upper quartiles.)**

5.3.6 Allelic expression ratio within ESR1, RANK, VDR, LRP5 and OPG SNPs

The MassArray system from Sequenom (Sequenom, California, USA) was used for genotyping and determination of allelic expression by Wellcome Trust Centre for Human Genetics service of Oxford University as outlined in the **method section 2.10.** Originally, AEI analysis was to be performed at the University of Sheffield and primers were extensively designed and optimised (Appendix 1). However, it was concluded that the Sequenom technology due to its higher sensitivity and efficiency would be a better choice of method. Genotyping was performed on the 53 cDNA samples derived from blood and 16 cDNA samples obtained from osteoclast cultures as described in the **method section 2.8 and 2.9.1**. Allelic expression imbalance was assessed using Sequenom assays for seven transcript SNPs; rs2228570, Rs731236, Rs208294, Rs1805034, Rs2073618, Rs2077647, and Rs556442. All 53 cDNA samples were heterozygous for four or more of the seven SNPs. Allelic expression ratios were estimated as the ratios of the area under the G, A, C and T peaks. All analyses used the logarithm of the ratios. The error bars represent the standard error (of the mean of the log ratios) as estimated from up to four replicates.

Two systematic quality control criteria for the allelic expression measurement at the transcribed SNPs were used. First all four replicates for each sample were checked with Grubb's test for finding outliers (http://graphpad.com/quickcalcs/Grubbs1.cfm). Simply all four replicates were copied to this website: it was calculated that any replicated which is a significant outlier would be deleted. Once all single samples had been checked with Grubb's test and any outliers removed then the mean and standard error of the mean (SEM) was calculated for each sample. Any samples with an SEM more than 0.15 were deleted (http://hdl.handle.net/10443/1083)

Results from the amplification of genomic DNA (25 samples) were used as a reference to normalise the blood cDNA and osteoclast cDNA values. Allelic expression ratios of the blood and osteoclast samples were normalised by subtracting the mean ratio of the 25 DNA samples after the removal of outliers.

To illustrate the quality control and data processing steps we show the data for the ESR transcribed SNP rs2077647. The plot **(Figure 5.14)** shows the mean log ratio and SE for each sample. The data (x-axis) are sorted by size of the mean and this type of graph is referred to as a caterpillar plot. **Figure 5.14A** shows two very clear outliers in the raw data of blood

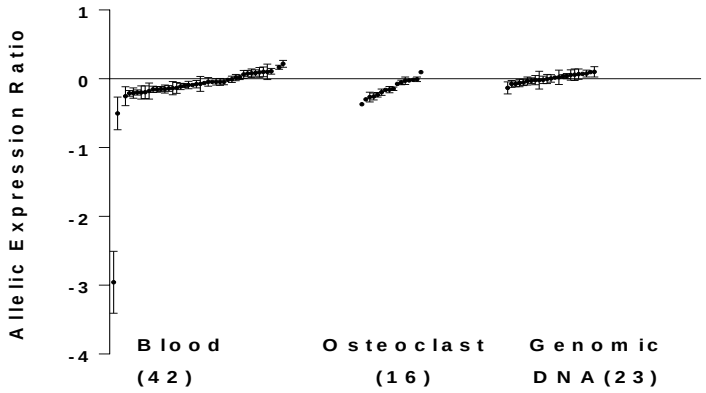
samples. Using Grubb's test and the SE check, two outlier samples were detected, and removed. Once the outliers were removed the range of ratios was smaller in genomic DNA than in the cDNAs (**Figure 5.14B**).

Figure 5.15A shows the raw data: clear outliers can be seen in the blood and osteoclast samples for the OPG transcribed SNP rs2073618. First Grubb's test was applied to each sample and seventeen outlier samples (twelve in genomic and five in blood samples) were detected within these results, which were removed and accounted for as a result of technical error. The range of individual allelic expression ratios was greater in the blood samples compared to osteoclast culture (**Figure 5.1B**).

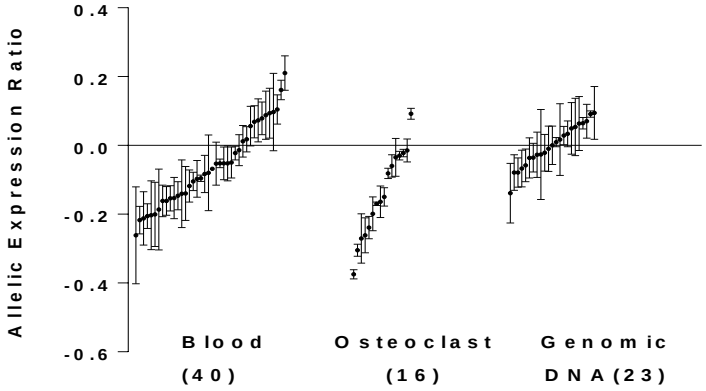
Figure 5.16A shows the raw data: clear outliers can be seen in the blood and genomic samples for the RANK transcribed SNP rs1805034. First Grubb's test was applied to each sample: five outlier samples (three in genomic and two in blood samples) were detected within these results, which were removed and accounted for as a result of technical error. The range of individual allelic expression ratios was greater in the blood samples compared to osteoclast cultures (**Figure 5.16B**).

Figure 5.17A shows the raw data: clear outliers can be seen in the blood and osteoclast samples for the VDR transcribed SNP rs731236. First Grubb's test was applied to each sample and twelve outlier samples (ten in blood and two in osteoclast samples) were detected within these results, which were removed and accounted for as a result of technical error. The range of individual allelic expression ratios was greater in the osteoclast samples compared to genomic DNA (**Figure 5.17B**).

A



B



C

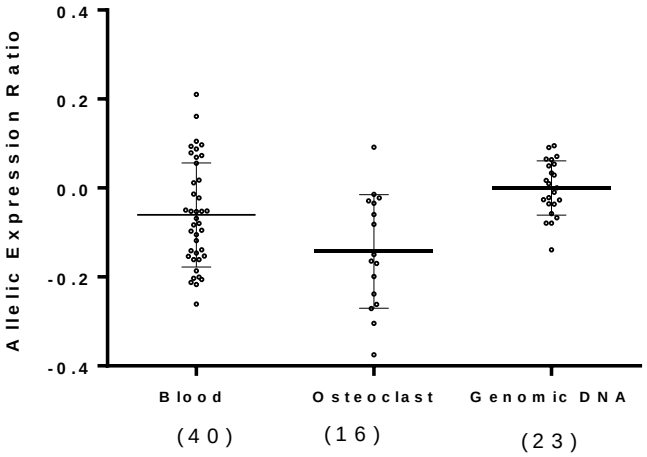
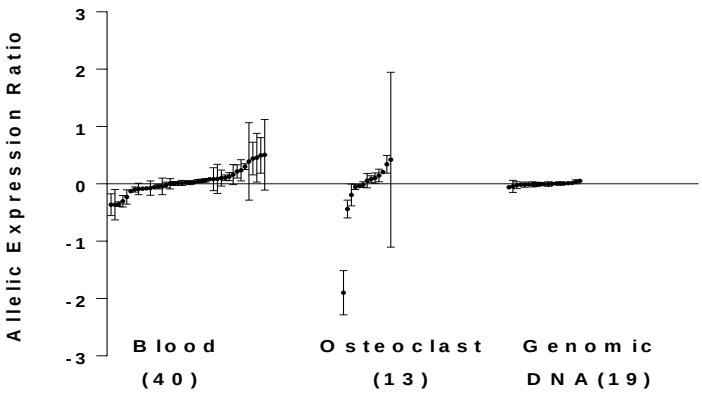


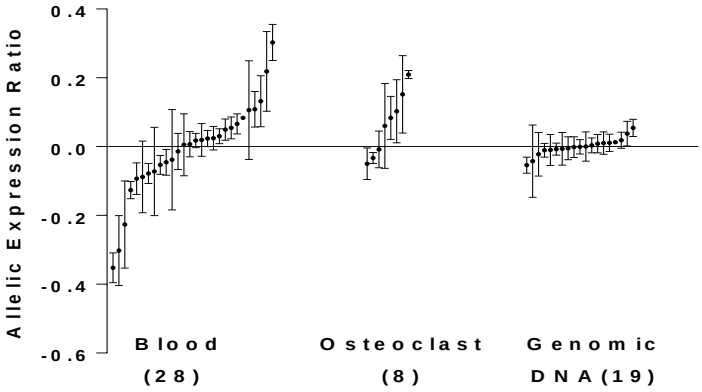
Figure 5-14: Allelic expression ratio for ESR transcribed SNP rs2077647

(A) *Raw allelic ratio data shown* ***(B)*** *Outliers were removed before Allelic expression ratios of the blood and osteoclast samples were normalized by subtracting the mean ratio of the DNA samples. This figure shows that only 1/16 of the osteoclast samples have a ratio above 0, and 11/41 of the bloods show a ratio above 0. So this is consistent with the transcribed SNP itself having a cis-acting effect or this transcribed SNP is in LD with a cis-acting variant.* ***(C)****Mean of allelic imbalance expression ratio from up to four replicates shown.*

A



B



C

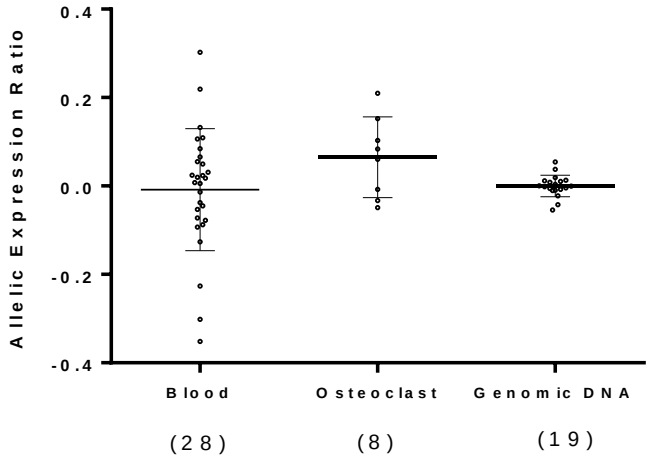
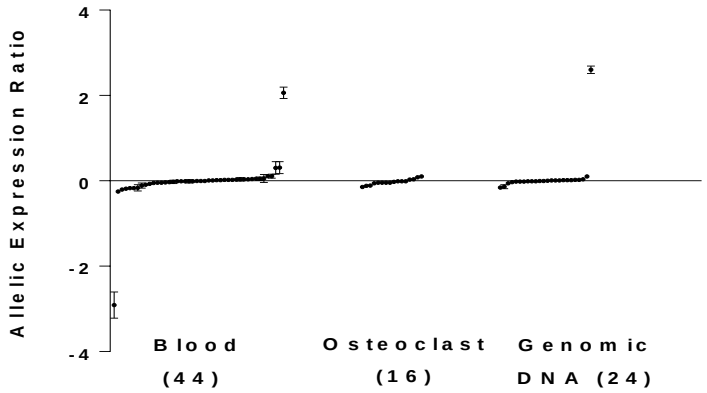


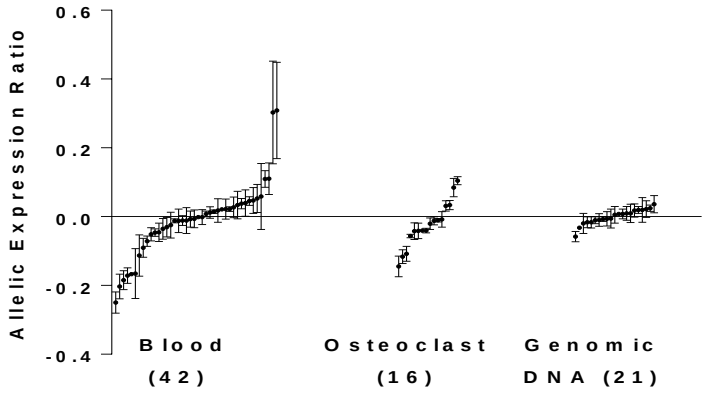
Figure 5-15: Allelic expression ratio for OPG transcribed SNP rs2073618

(A) *Raw allelic ratio data shown* ***(B)*** *Outliers were removed before Allelic expression ratios of the blood and osteoclast samples were normalized by subtracting the mean ratio of the DNA samples.* ***(C)****Mean and SD of allelic imbalance expression ratio from up to four replicates shown.*

A



B



C

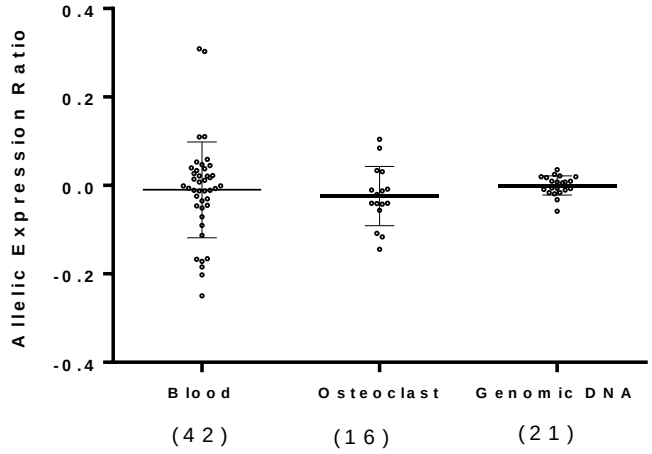
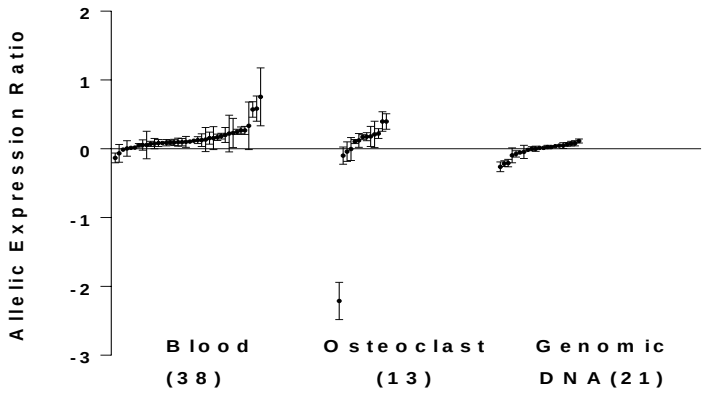


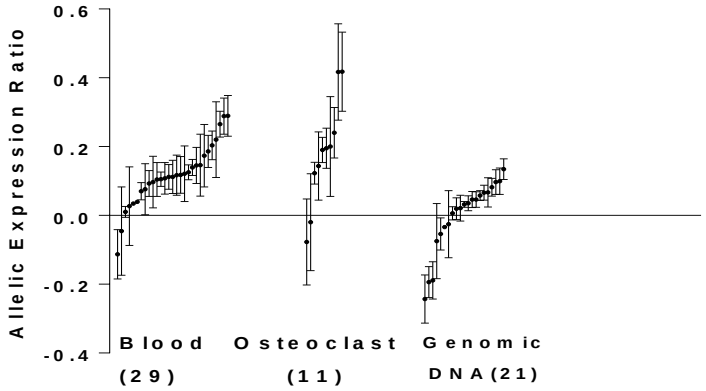
Figure 5-16: Allelic expression ratio for RANK transcribed SNP rs1805034

(A) *Raw allelic ratio data shown* ***(B)*** *Outliers were removed before Allelic expression ratios of the blood and osteoclast samples were normalized by subtracting the mean ratio of the DNA samples.* ***(C)*** *Mean and SD of allelic imbalance expression ratio from up to four replicates shown.*

A



B



C

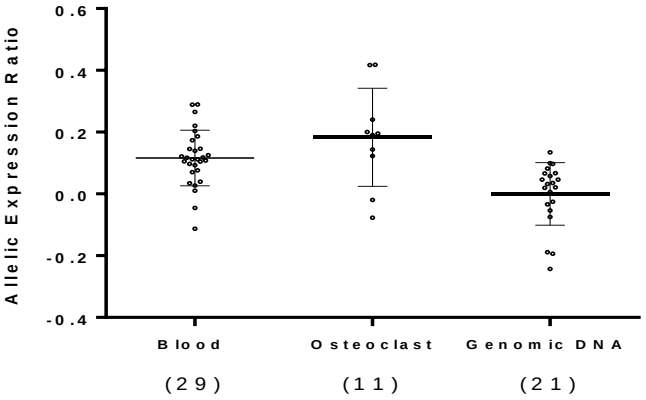


Figure 5-17: Allelic expression ratio for VDR SNP rs731236

(A) *Raw allelic ratio data shown* ***(B)*** *Outliers were removed before Allelic expression ratios of the blood and osteoclast samples were normalized by subtracting the mean ratio of the DNA samples. This figure shows that only 2/29 of the osteoclast samples have a ratio above 0, and 2/11 of the bloods show a ratio above 0. So this is consistent with the transcribed SNP itself having a cis-acting effect or this transcribed SNP is in LD with a cis-acting variant.* ***(C)*** *Mean and SD of allelic imbalance expression ratio from up to four replicates shown.*

5.3.7 Effect of risk SNP genotype on allelic expression ratio of transcribed SNPs

To assess the association between transcribed SNPs and allelic expression ratio, we used an extension of the method published previously by Teare et al., 2006. Allelic expression ratio across the blood and osteoclast was used to evaluate associations between selected risk and transcribed SNPs for osteoporosis ESR1, OPG, RANK and VDR candidate genes. In these experiments the expression ratios of the transcribed SNPs were then grouped by specific risk SNP genotype for each subject within our Iranian population. The allelic expression ratio of the transcribed SNPs was determined for genomic DNA samples (gDNA) and used as an assay control. All allelic expression data are presented as the log ratios and are a mean of four replicates. All tests for cis-acting effect on allelic imbalance expression were done using Dr Santibanez-Koref method (MSK test).

There is evidence of a cis-acting effect associated with the transcribed SNP rs2077647 of the ESR candidate gene, as after normalisation by the genomic sample mean both the blood and osteoclast cDNA ratios show a tendency for over expression of one of the alleles (Figure 5.14) (one-sample t-test, blood cDNA μ_{cis} =-0.07 p=0.0004; osteoclast μ_{cis} = -0.14 p=0.0003). The cis-acting effect could be due to the transcribed variant or one in LD with the transcribed SNP. A total of six risk SNPs were tested for evidence of cis-acting effects on the blood and osteoclast samples. A significant result was observed for the two SNPs rs2941740 and rs2504063. For the SNP rs294170 (Figure 5.22A) the distribution of the AE ratio looks to be centred lower for the AG group compared to the AA group, the cis-acting effect reducing the ratio by -0.12 (p=0.02). This effect cannot be seen on the osteoclast analysis (p=0.69), but this may be due to lack of power, only two heterozygotes were available in the osteoclast sample.

A more consistent result is observed with the SNP rs2504063. The anticipated pattern associated with a cis-acting effect is observed in the osteoclast with the heterozygous group showing a wider range of values than the two homozygotes. The anticipated pattern is not so obvious in the blood sample, with the GG genotype group having a wider range of values than the AG group. However, as the MSK test relies on a test of the means there is sufficient statistical evidence that the estimated known-phase haplotype groups do differ from the homozygous means (μ_{cis} =-0.1, p=0.007) (Figure 5.18). So overall, within ESR, evidences of

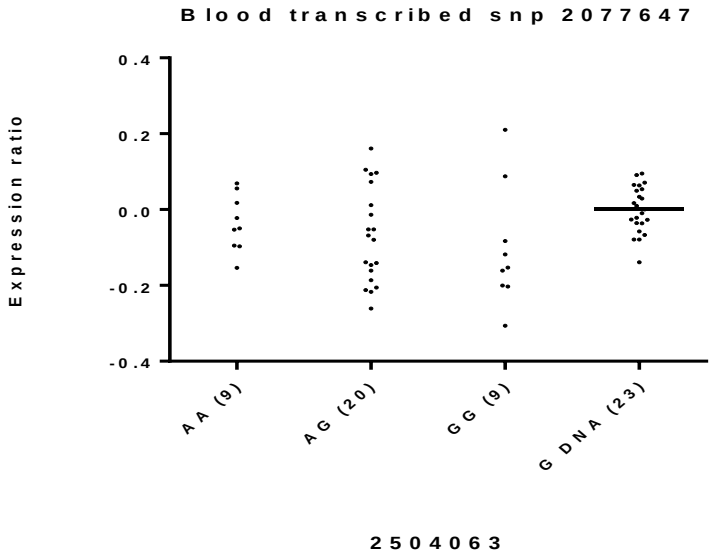
a cis-acting effect with the transcribed SNP 2077647 were seen. One statistically strong cis-acting effect was seen for rs2504063 in both blood and osteoclast cDNA. A weaker but still statistically significant association was seen for SNP rs2941740 in the blood sample only.

A total of five risk SNPs was tested for evidence of cis-acting effects on the blood and osteoclast samples for the rs2073618 transcribed SNP of the OPG candidate gene. A significant result was observed in osteoclast cells for two SNPs rs4355801 and rs2062375. For the SNP rs4355801 (Figure 5.25B) there is just one homozygous sample and others are heterozygous for rs4355801. The heterozygous group shows a wider value than the G DNA samples, the cis-acting effect increasing the ratio by 0.07 (p=0.02). For the SNP rs2062375 (Figure 5.26B) all osteoclast samples are heterozygous. The heterozygous group shows a wider value than the G DNA samples, the cis-acting effect increasing the ratio by 0.06 (p=0.05).

Two risk SNPs were tested for evidence of cis-acting effects on the blood and osteoclast samples for the rs1805034 transcribed SNP of the RANK candidate gene. A significant result was observed in osteoclast cells for the risk SNPs rs3018362, the cis-acting effect increasing the ratio by 0.06 (p=0.02) (Figure 5.30B).

There is evidence of a cis-acting effect associated with the transcribed SNP rs731236 of the VDR, as after normalisation by the genomic sample mean both the blood and osteoclast cDNA ratios show a tendency for over-expression of one of the alleles (Figure 5.17) (one-sample t-test, blood cDNA μ_{cis} = 0.116 p=< 0.0001; osteoclast μ_{cis} = 0.183 p=0.005). The cis-acting effect could be due to the transcribed variant or one in LD with the transcribed SNP. Two risk SNPs in total were tested for evidence of cis-acting effects on the blood and osteoclast samples. A significant result was observed for the SNPs rs731236 which is also transcribed SNPs so it is in perfect disequilibrium (Figure 5.10.B). In the blood samples (Figure 5.32A) the distribution of the AE ratio looks to be centred upper for the CT group compared to the GDNA group, the cis-acting effect increasing the ratio by 0.12 ($p=1.8e^{-08}$). In the osteoclast samples (Figure 5.32B) the distribution of the AE ratio looks to be centred upper for the CT group compared to the G DNA group, the cis-acting effect increasing the ratio by 0.18 (p=0.002). So overall within VDR evidences of a cis-acting effect with the transcribed SNP 731236 were seen.

A



B

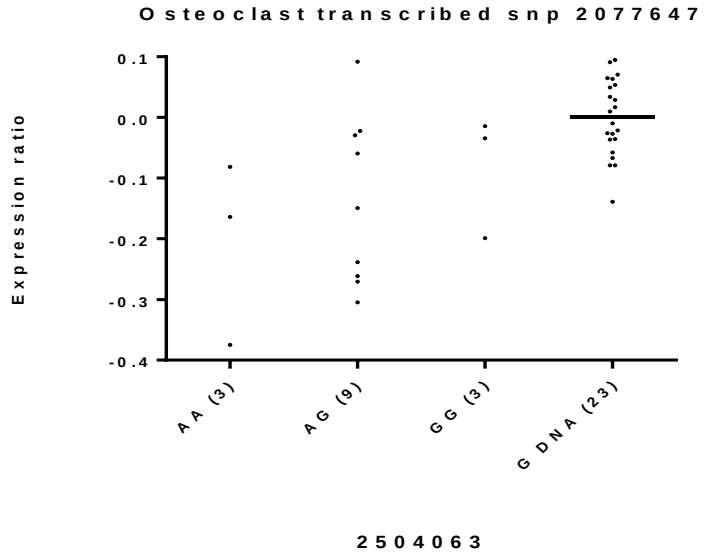
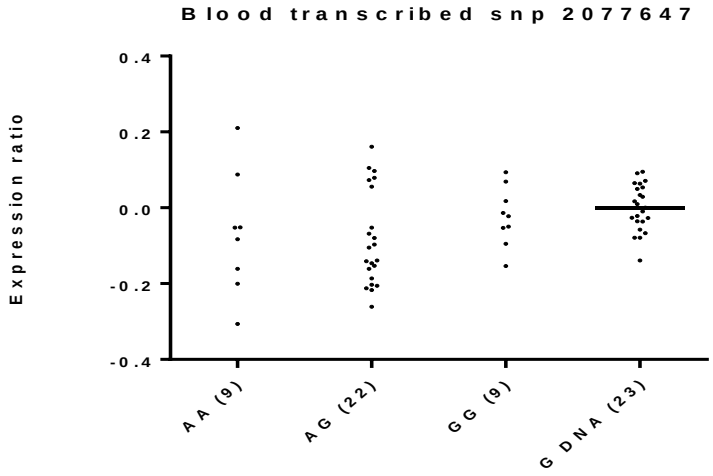


Figure 5-18: Effect of genotype at rs2504063 on allelic expression ratio of transcribed ESR SNP rs2077647

Each point of the graph is representative of the allelic expression ratio for the transcribed SNP rs2077647 for each individual: the data are grouped into homozygous risk (A:A), heterozygous (A:G) or homozygous wild type (G:G) of the risk SNP rs2504063. The genomic DNA sample mean was used as a normalization control. ***(A)*** *The blood sample genotype of SNP rs2504063 for the transcribed SNP rs2077647 (μ_{cis}=-0.10, P-value=0.007 for the comparison with genomic DNA between different genotypes of rs2504063).* ***(B)*** *The osteoclast sample genotype of SNP rs2504063 for the transcribed SNP rs2077647 (μ_{cis}=-0.15, P-value=0.02 MSK test for checking the evidence of a cis-acting effect with the transcribed SNP rs2077647).*

A



B

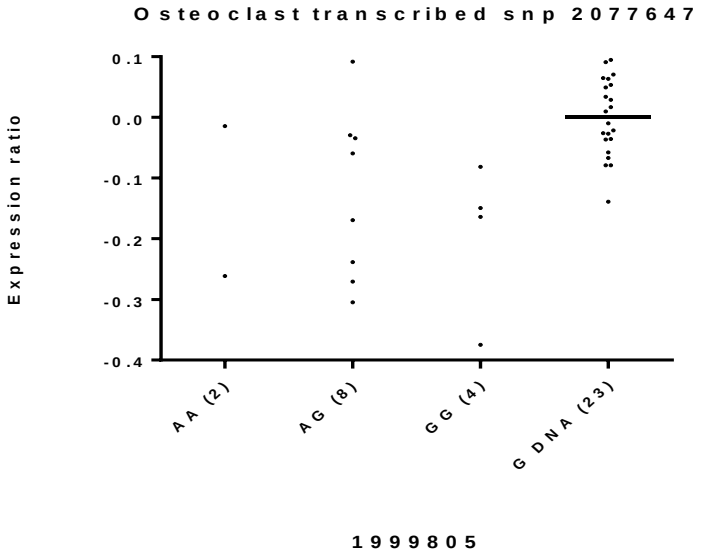
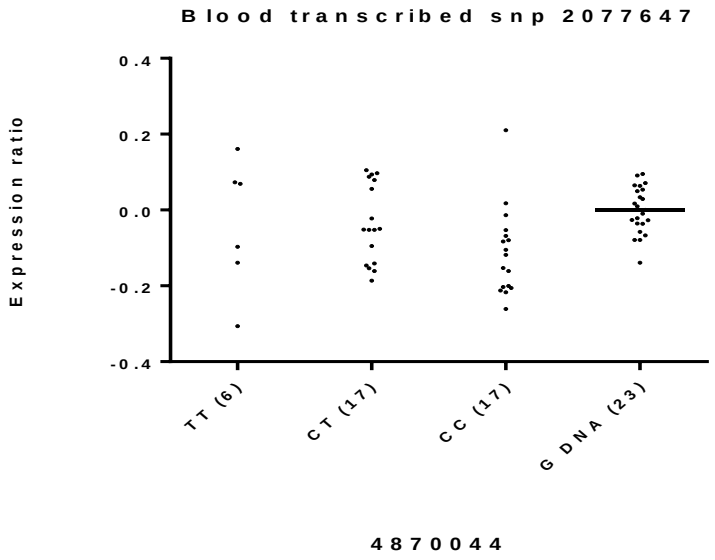


Figure 5-19: Effect of genotype at rs1999805 on allelic expression ratio of transcribed ESR SNP rs2077647

Each point of the graph is representative of the allelic expression ratio for the transcribed SNP rs2077647 for each individual: the data are grouped into homozygous risk (A:A), heterozygous (A:G) or homozygous wild type (G:G) of the risk SNP 1999805. The genomic DNA sample mean was used as a normalization control. ***(A)*** *The blood sample genotype of SNP rs1999805 for the transcribed SNP rs2077647 (μ_{cis}=0.13, P-value=0.13 for the comparison with genomic DNA between different genotypes of rs1999805).* ***(B)*** *The osteoclast sample genotype of SNP rs1999805 for the transcribed SNP rs2077647 (μ_{cis}=0.1, P-value=0.18 MSK test for checking the evidence of a cis-acting effect with the transcribed SNP rs2077647).*

A



B

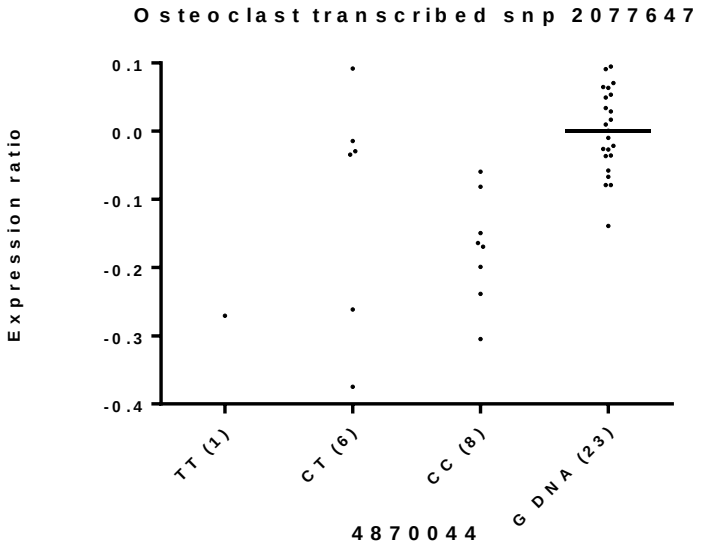
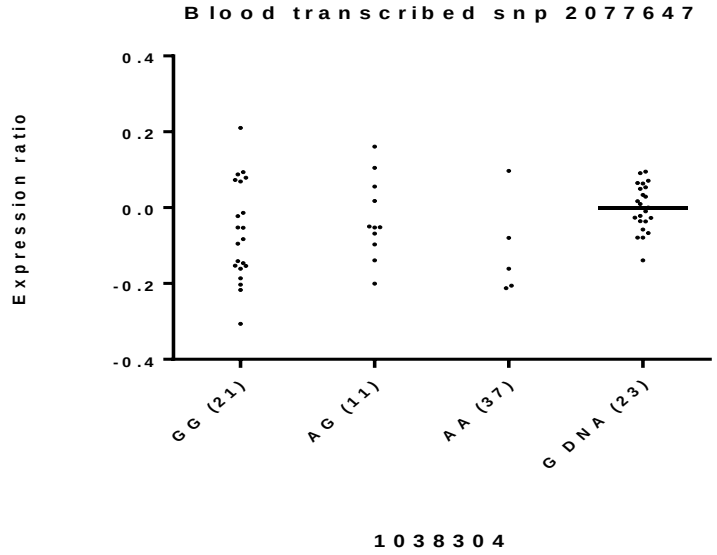


Figure 5-20: Effect of genotype at rs4870044 on allelic expression ratio of transcribed ESR SNP rs2077647

Each point of the graph is representative of the allelic expression ratio for the transcribed SNP rs2077647 for each individual: the data are grouped into homozygous risk (T: T), heterozygous (C:T) or homozygous wild type (C:C) of the risk SNP 4870044. The genomic DNA sample mean was used as a normalization control. ***(A)*** *The blood sample genotype of SNP rs4870044 for the transcribed SNP rs2077647 (μ_{cis}=0.053, P-value=0.15 for the comparison with genomic DNA between different genotypes of rs4870044).* ***(B)*** *The osteoclast sample genotype of SNP rs4870044 for the transcribed SNP rs2077647 (μ_{cis}=-0.01, P-value=0.90 MSK test for checking the evidence of a cis-acting effect with the transcribed SNP rs2077647).*

A



B

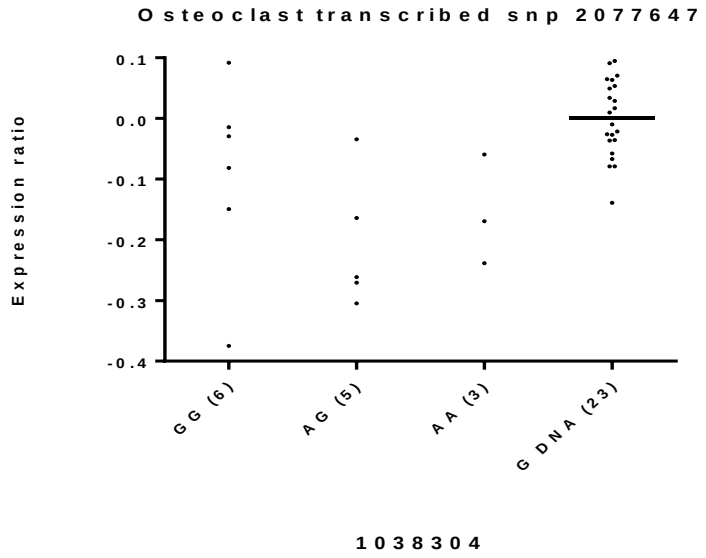
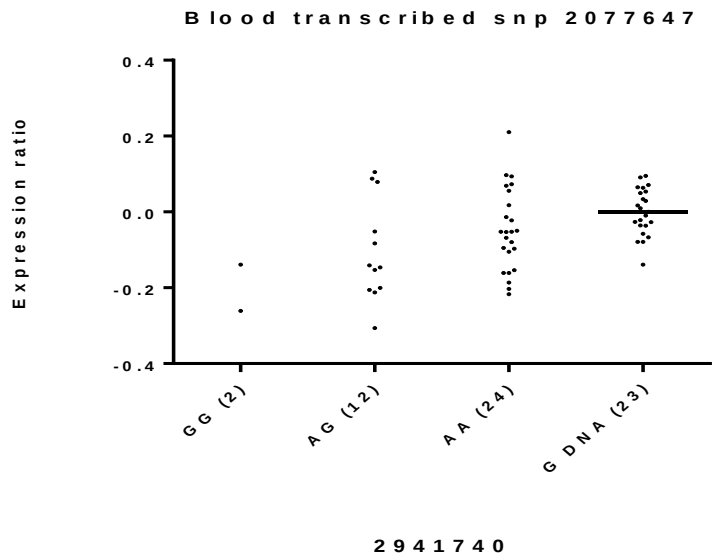


Figure 5-21: Effect of genotype at rs1038304 on allelic expression ratio of transcribed ESR SNP rs2077647

Each point of the graph is representative of the allelic expression ratio for the transcribed SNP rs2077647 for each individual: the data are grouped into homozygous risk (G:G), heterozygous (A:G) or homozygous wild type (A:A) of the risk SNP rs1038304. The genomic DNA sample mean was used as a normalization control. ***(A)*** *The blood sample genotype of SNP rs1038304 for the transcribed SNP rs2077647 (μ_{cis}=0.045, P-value=0.45 for the comparison with genomic DNA between different genotypes of rs1038304).* ***(B)*** *The osteoclast sample genotype of SNP rs1038304 for the transcribed SNP rs2077647 (μ_{cis}=0.15, P-value=0.53 MSK test for checking the evidence of a cis-acting effect with the transcribed SNP rs2077647).*

A



B

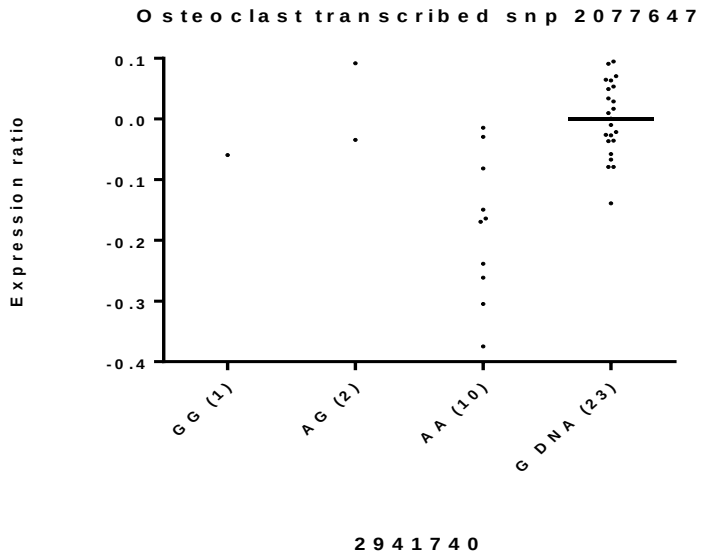
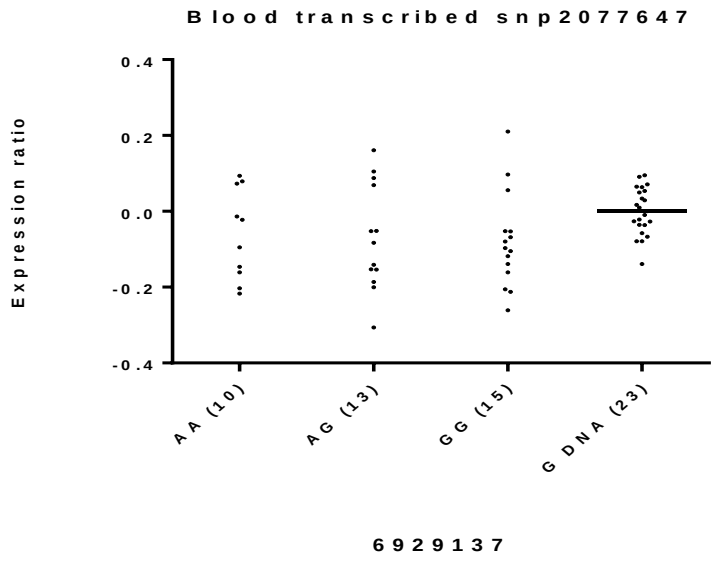


Figure 5-22: Effect of genotype at rs2941740 on allelic expression ratio of transcribed ESR SNP rs2077647

Each point of the graph is representative of the allelic expression ratio for the transcribed SNP rs2077647 for each individual: the data are grouped into homozygous risk (G:G), heterozygous (A:G) or homozygous wild type (A:A) of the risk SNP rs2941740. The genomic DNA sample mean was used as a normalization control. ***(A)*** *The blood sample genotype of SNP rs2941740 for the transcribed SNP rs2077647 (μ_{cis}=-0.12, P-value=0.02 for the comparison with genomic DNA between different genotypes of rs2941740).* ***(B)*** *The osteoclast sample genotype of SNP rs2941740 for the transcribed SNP rs2077647 (μ_{cis}=-0.05, P-value=0.69 MSK test for checking the evidence of a cis-acting effect with the transcribed SNP rs2077647).*

A



B

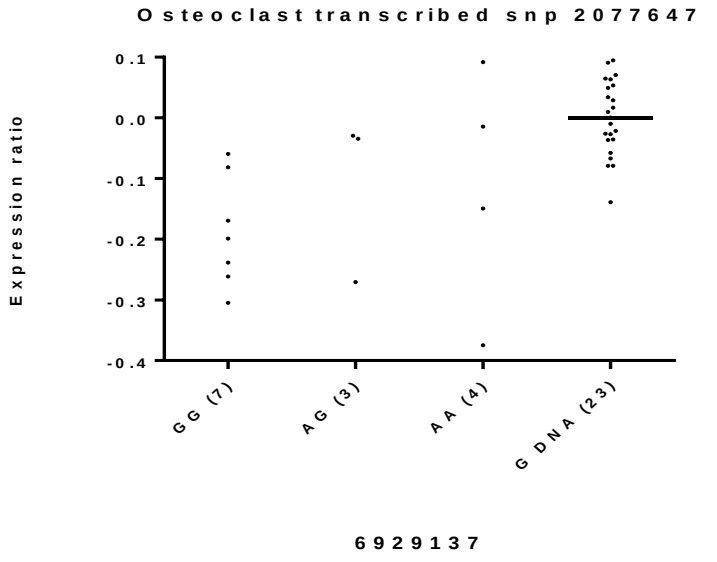
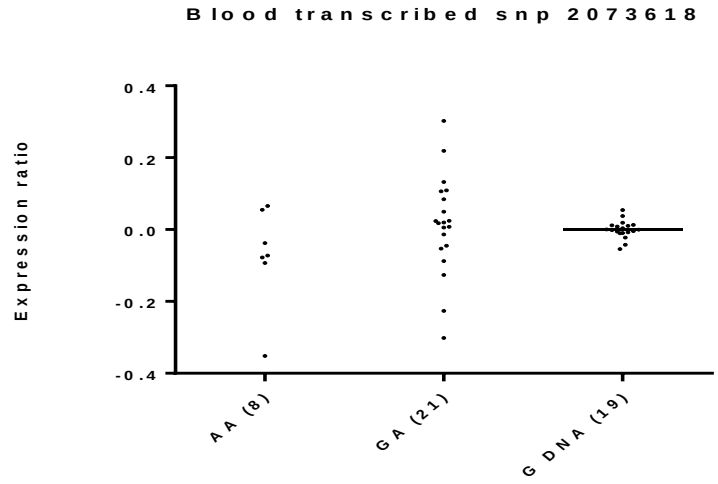


Figure 5-23: Effect of genotype at rs6929137 on allelic expression ratio of transcribed ESR SNP rs2077647

Each point of the graph is representative of the allelic expression ratio for the transcribed SNP rs2077647 for each individual: the data are grouped into homozygous risk (A:A), heterozygous (A:G) or homozygous wild type (G:G) of the risk SNP rs6929137. The genomic DNA sample mean was used as a normalization control. ***(A)*** *The blood sample genotype of SNP rs6929137 for the transcribed SNP rs2077647 (μ_{cis}=-0.07, P-value=0.18 for the comparison with genomic DNA between different genotypes of rs6929137).* ***(B)*** *The osteoclast sample genotype of SNP rs6929137 for the transcribed SNP rs2077647 (μ_{cis}=0.06, P-value=0.60 MSK test for checking the evidence of a cis-acting effect with the transcribed SNP rs2077647).*

A



B

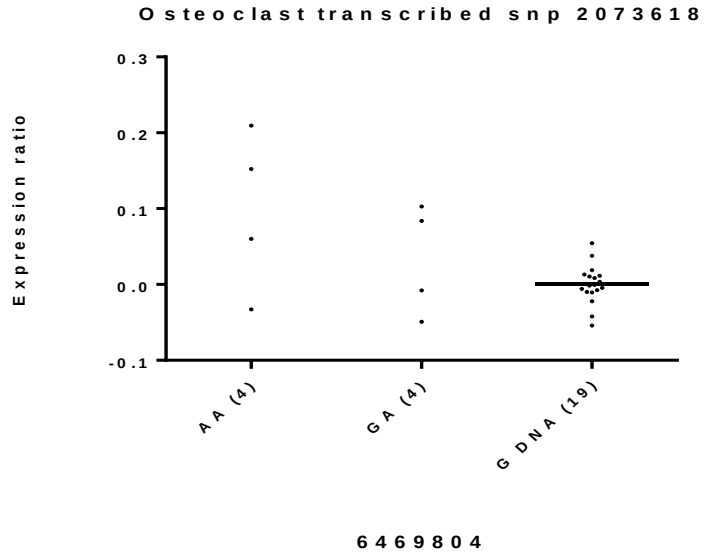
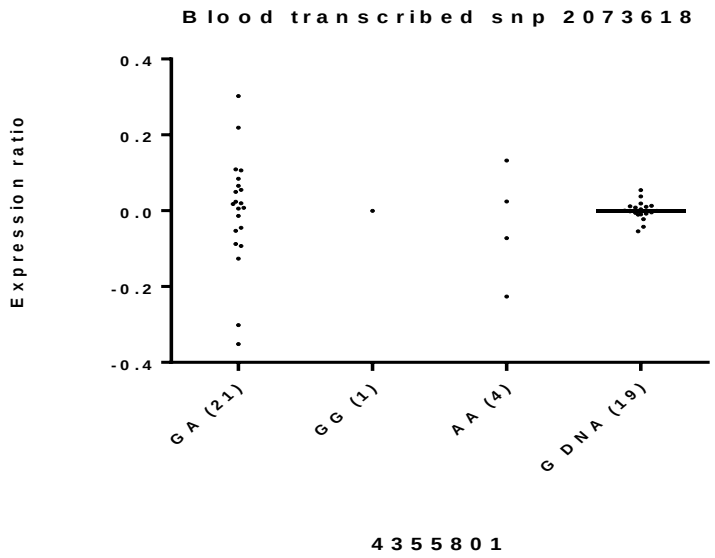


Figure 5-24: Effect of genotype at rs6469804 on allelic expression ratio of transcribed OPG SNP rs2073618

Each point of the graph is representative of the allelic expression ratio for the transcribed SNP rs2073618 for each individual: the data are grouped into homozygous risk (A:A) and heterozygous (G:A) of the risk SNP rs6469804. The genomic DNA sample mean was used as a normalization control. ***(A)*** *The blood sample genotype of SNP rs6469804 for the transcribed SNP rs2073618 (μ_{cis} =0.01, P-value=0.61 for the comparison with genomic DNA between different genotypes of rs6469804).* ***(B)*** *The osteoclast sample genotype of SNP rs6469804 for the transcribed SNP rs2073618 (μ_{cis} =0.03, P-value=0.54 MSK test for checking the evidence of a cis-acting effect with the transcribed SNP rs2073618).*

A



B

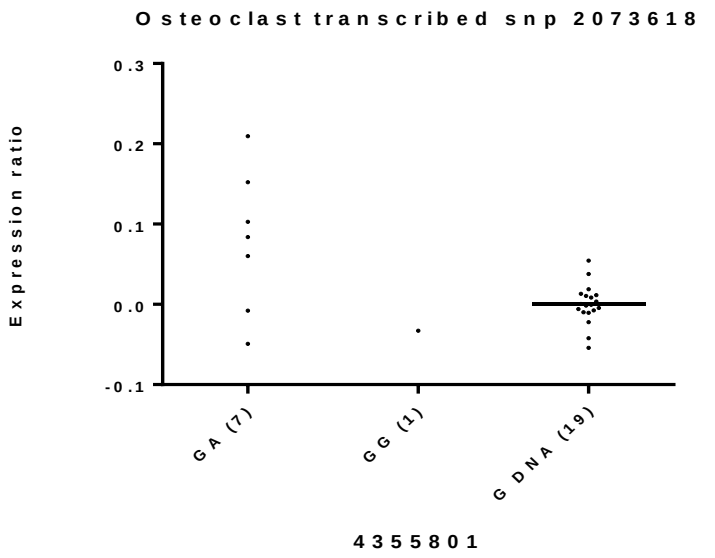
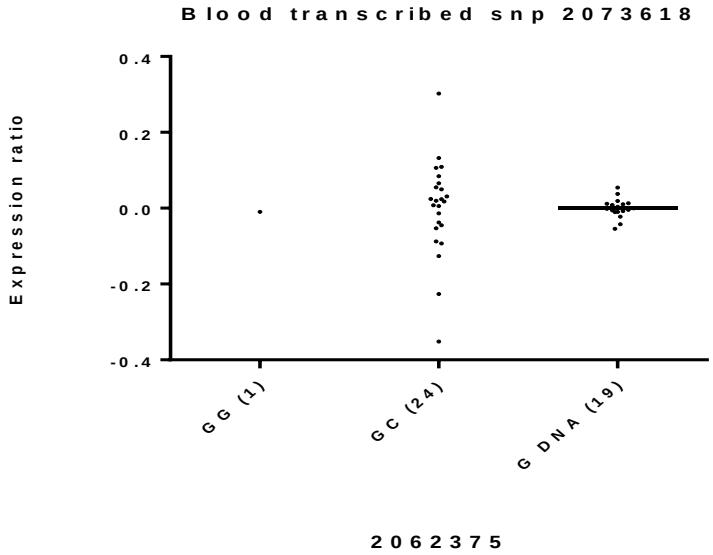


Figure 5-25: Effect of genotype at rs4355801 on allelic expression ratio of transcribed OPG SNP rs2073618

Each point of the graph is representative of the allelic expression ratio for the transcribed SNP rs2073618 for each individual: the data are grouped into homozygous risk (AA), heterozygous (G:A) and homozygous wild type (G:G) of the risk SNP rs4355801for blood samples and all except homozygous risk (AA) for osteoclast cells. The genomic DNA sample mean was used as a normalization control. ***(A)*** *The blood sample genotype of SNP rs4355801 for the transcribed SNP rs2073618 (μ_{cis} =0.00, P-value=0.98 for the comparison with genomic DNA between different genotypes of rs4355801).* ***(B)*** *The osteoclast sample genotype of SNP rs4355801 for the transcribed SNP rs2073618 (μ_{cis} =0.07, P-value=0.02 MSK test for checking the evidence of a cis-acting effect with the transcribed SNP rs2073618).*

A



B

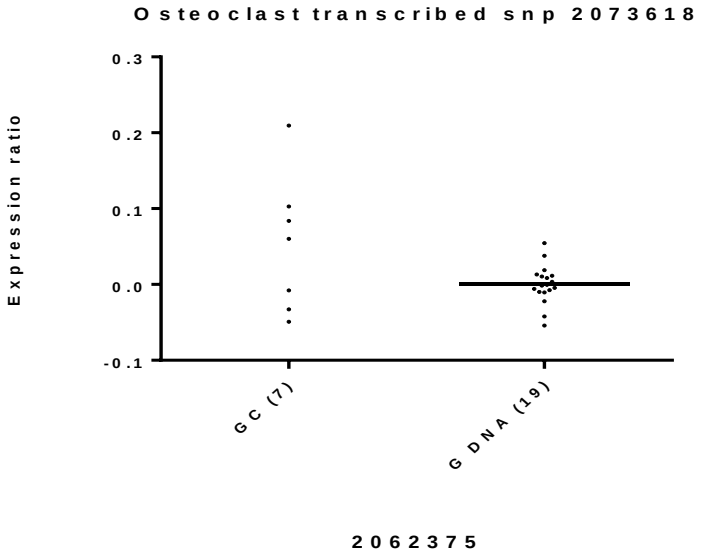
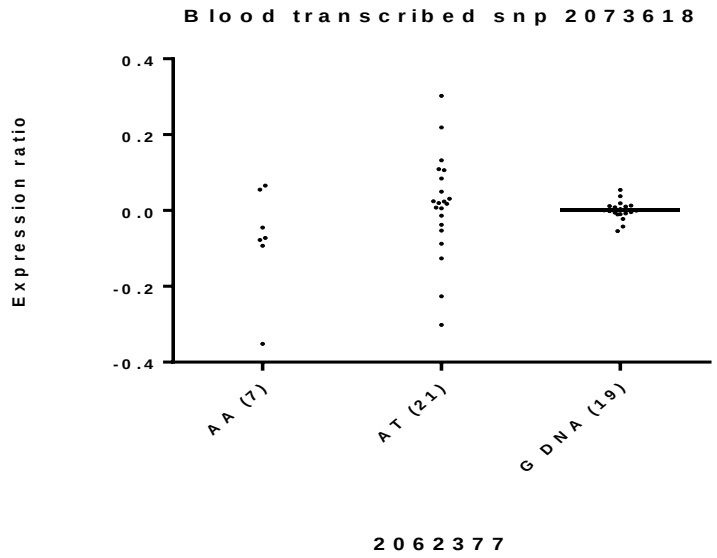


Figure 5-26: Effect of genotype at rs2062375 on allelic expression ratio of transcribed OPG SNP rs2073618

Each point of the graph is representative of the allelic expression ratio for the transcribed SNP rs2073618 for each individual: the data are grouped into homozygous risk (G:G) and heterozygous (G:C) of the risk SNP 2062375 for blood samples and just heterozygous (GC) for osteoclast cells. The genomic DNA sample mean was used as a normalization control. ***(A)*** *The blood sample genotype of SNP rs2062375 for the transcribed SNP rs2073618 (μ_{cis}=0.00, P-value=0.96 for the comparison with genomic DNA between different genotypes of rs2062375).* ***(B)*** *The osteoclast sample genotype of SNP rs2062375 for the transcribed SNP rs2073618 (μ_{cis}=0.06, P-value=0.05 MSK test for checking the evidence of a cis-acting effect with the transcribed SNP rs2073618).*

A



B

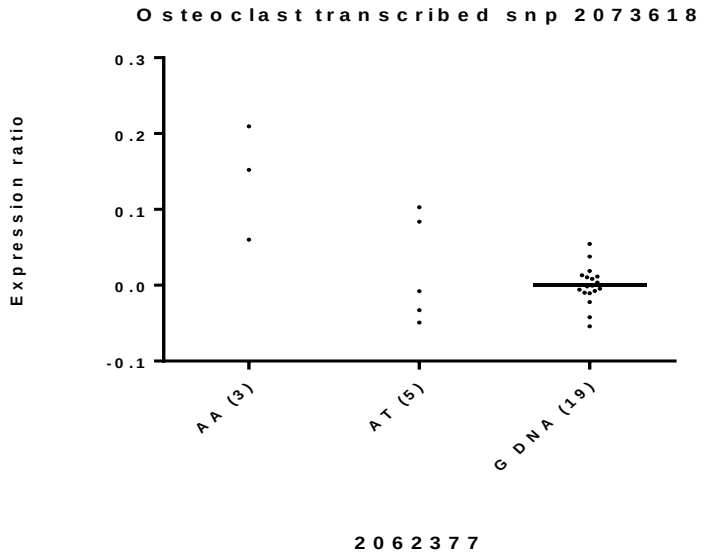


Figure 5-27:Effect of genotype at rs2062377 on allelic expression ratio of transcribed OPG SNP rs2073618

Each point of the graph is representative of the allelic expression ratio for the transcribed SNP rs2073618 for each individual: the data are grouped into heterozygous (A: T) and homozygous wild type (A:A) of the risk SNP rs2062377. The genomic DNA sample mean was used as a normalization control. ***(A)*** *The blood sample genotype of SNP rs2062377 for the transcribed SNP rs2073618 (μ_{cis}=0.01, P-value=0.60 for the comparison with genomic DNA between different genotypes of rs2062377).* ***(B)*** *The osteoclast sample genotype of SNP rs2062377 for the transcribed SNP rs2073618 (μ_{cis}=0.02, P-value=0.61 MSK test for checking the evidence of a cis-acting effect with the transcribed SNP rs2073618).*

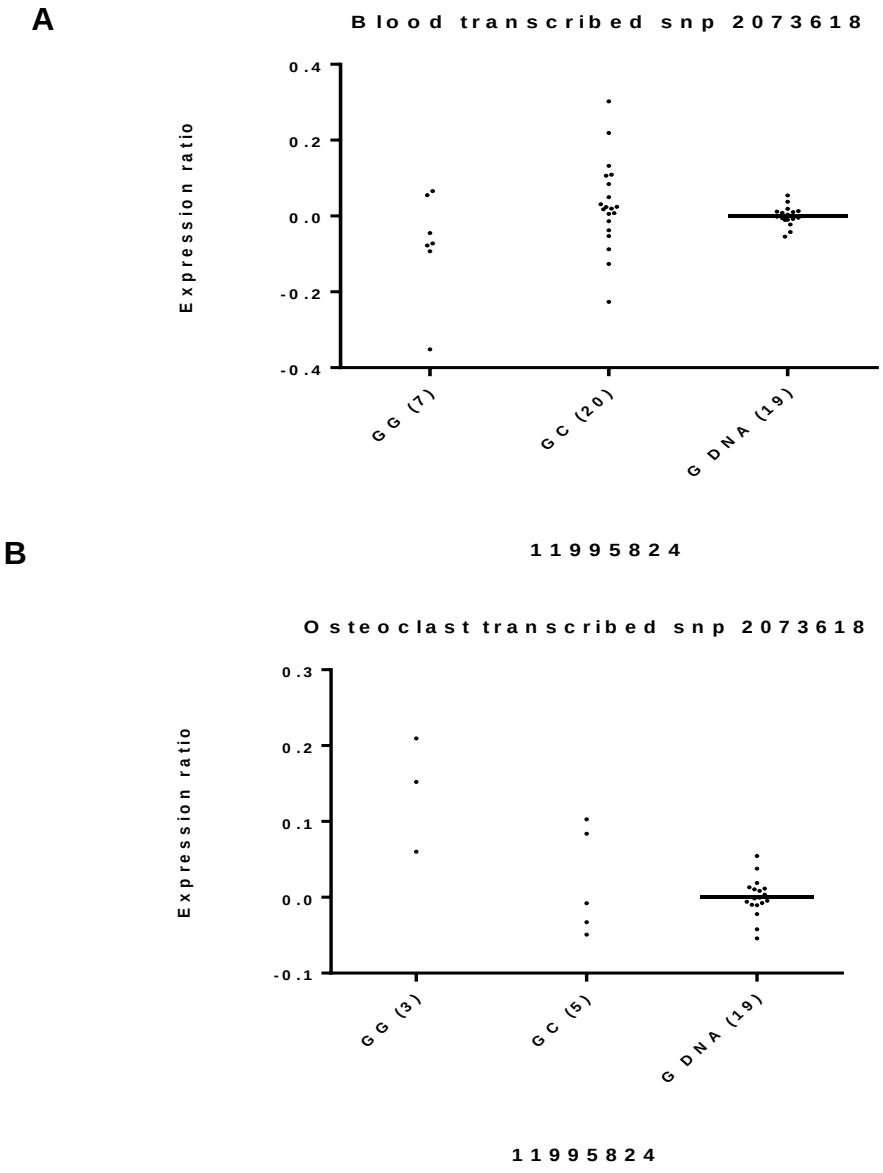
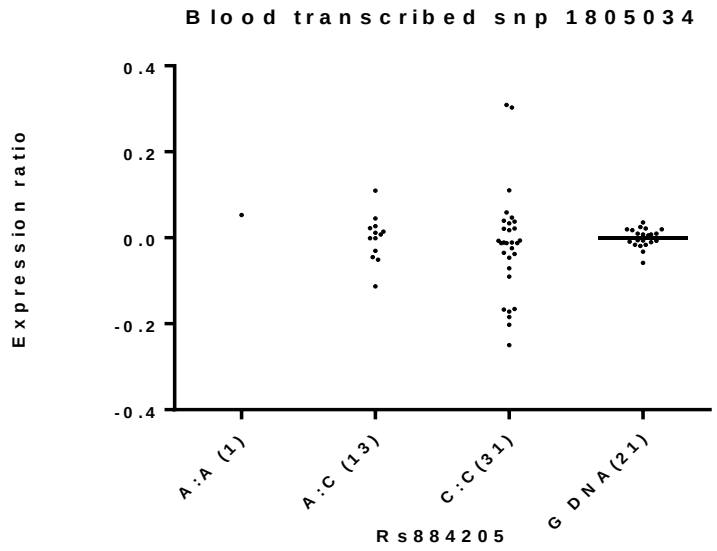


Figure 5-28:Effect of genotype at rs11995824 on allelic expression ratio of transcribed OPG SNP rs2073618

Each point of the graph is representative of the allelic expression ratio for the transcribed SNP rs2073618 for each individual: the data are grouped into homozygous risk (G: G) and heterozygous (G: C) of the risk SNP rs11995824. The genomic DNA sample mean was used as a normalization control. ***(A)*** *The blood sample genotype of SNP rs11995824 for the transcribed SNP rs2073618 (μ_{cis}=0.01, P-value=0.60 for the comparison with genomic DNA between different genotypes of rs11995824).* ***(B)*** *The osteoclast sample genotype of SNP rs11995824 for the transcribed SNP rs2073618 (μ_{cis}=0.02, P-value=0.61 MSK test for checking the evidence of a cis-acting effect with the transcribed SNP rs2073618).*

A



B

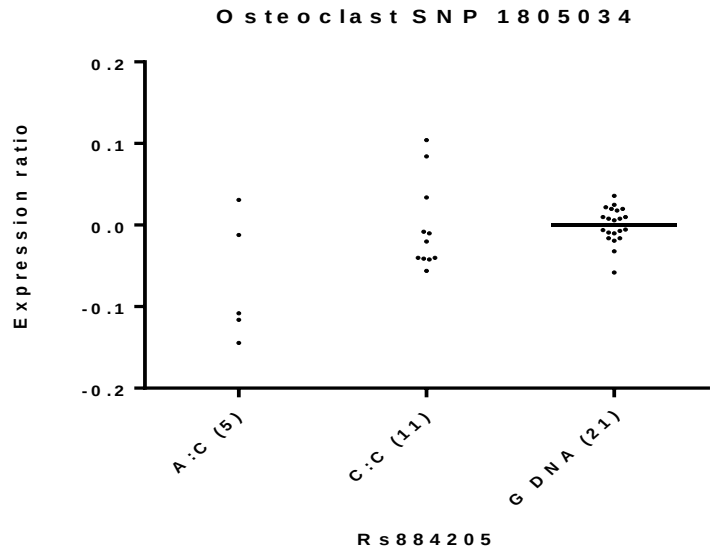
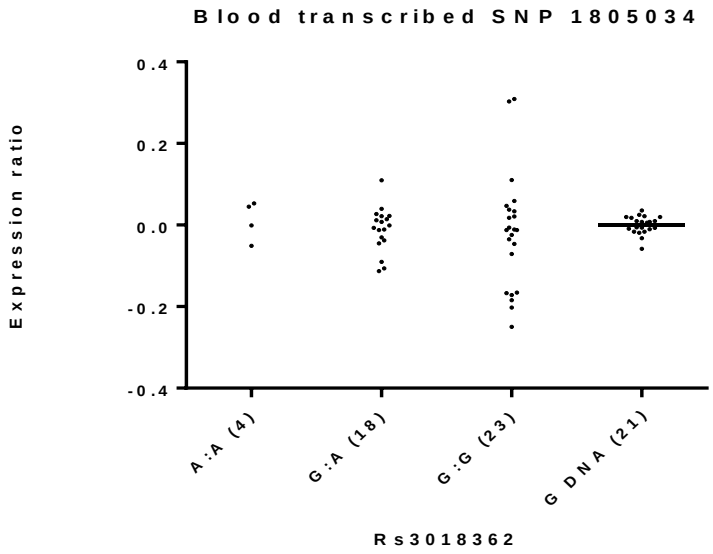


Figure 5-29: Effect of genotype at rs884205 on allelic expression ratio of transcribed RANK SNP rs1805034

Each point of the graph is representative of the allelic expression ratio for the transcribed SNP rs1805034 for each individual: the data are grouped into homozygous risk (A:A) and heterozygous (A:C) and homozygous wild type (C:C) of the risk SNP rs884205for blood samples and all except homozygous risk (A:A) for osteoclast cells. Genomic DNA was used as a normalization control. ***(A)*** *The blood sample genotype of SNP rs884205 for the transcribed SNP rs1805034 (μ_{cis} =0.00, P-value=0.95 for the comparison with genomic DNA between different genotypes of rs884205).* ***(B)*** *The osteoclast sample genotype of SNP rs884205 for the transcribed SNP rs1805034 (μ_{cis} =0.06, P-value=0.1495 MSK test for checking the evidence of a cis-acting effect with the transcribed SNP rs1805034).*

A



B

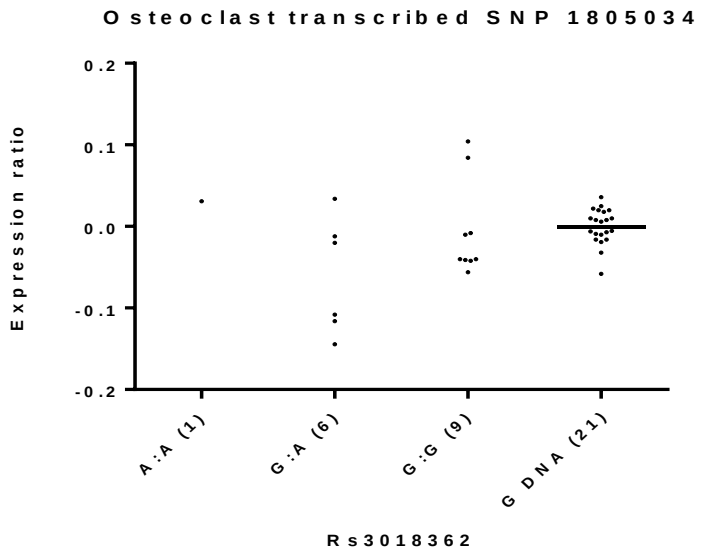
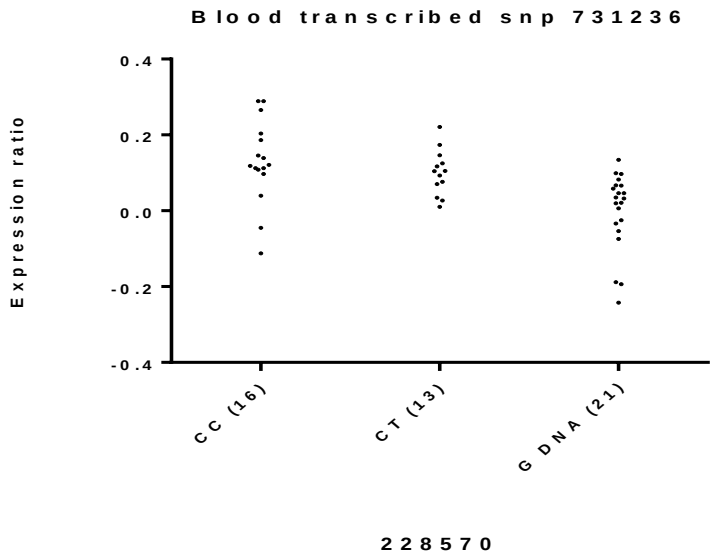


Figure 5-30: Effect of genotype at rs3018362 on allelic expression ratio of transcribed RANK SNP rs1805034

Each point of the graph is representative of the allelic expression ratio for the transcribed SNP rs1805034 for each individual: the data are grouped into homozygous risk (A:A), heterozygous (G:A) or homozygous wild type (G:G) of the risk SNP rs3018362. The genomic DNA sample mean was used as a normalization control. ***(A)*** *The blood sample genotype of SNP rs3018362 for the transcribed SNP rs1805034 (μ_{cis} ==0.001, P-value=0.94 for the comparison with genomic DNA between different genotypes of rs3018362).* ***(B)*** *The osteoclast sample genotype of SNP rs3018362 for the transcribed SNP rs1805034 (μ_{cis} =0.06, P-value=0.02 MSK test for checking the evidence of a cis-acting effect with the transcribed SNP rs1805034).*

A



B

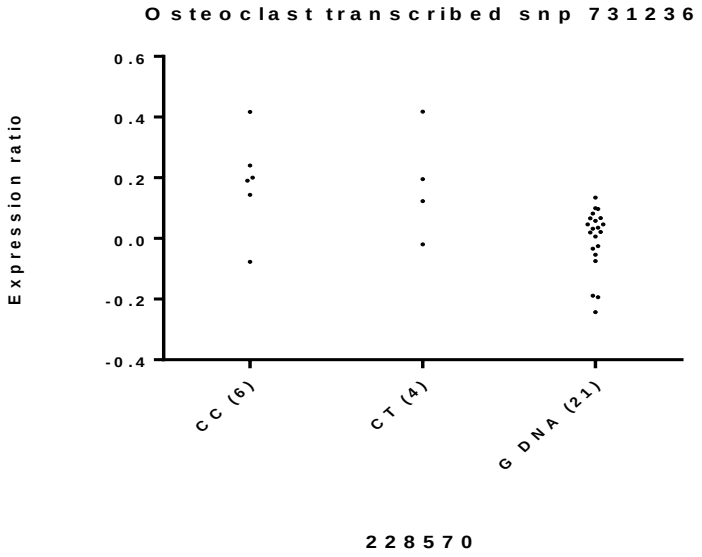
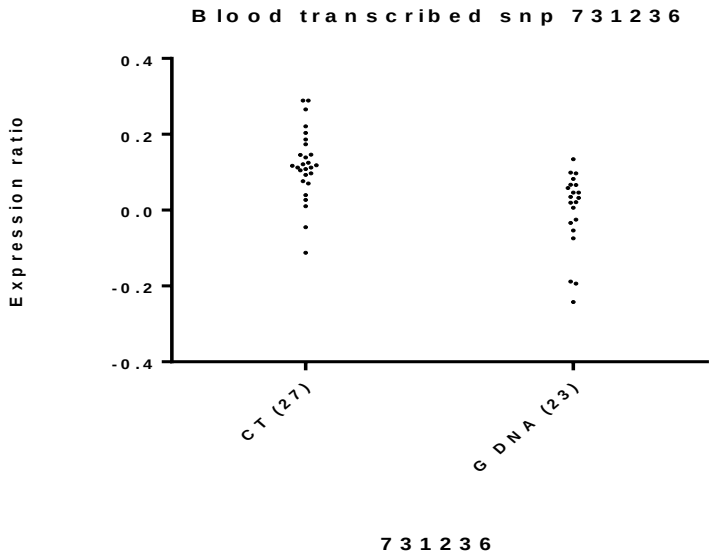


Figure 5-31: Effect of genotype at rs228570 on allelic expression ratio of transcribed VDR SNP rs731236

Each point of the graph is representative of the allelic expression ratio for the transcribed SNP rs731236 for each individual: the data are grouped into homozygous risk (C:C) and heterozygous (A:C) of the risk SNP rs228570. The genomic DNA sample mean was used as a normalization control. ***(A)*** *The blood sample genotype of SNP rs228570 for the transcribed SNP rs731236 (μ_{cis}=0.06, P-value=0.28 for the comparison with genomic DNA between different genotypes of rs228570).* ***(B)*** *The osteoclast sample genotype of SNP rs228570 for the transcribed SNP rs731236 (μ_{cis}=.01, P-value=0.37 MSK test for checking the evidence of a cis-acting effect with the transcribed SNP rs731236).*

A



B

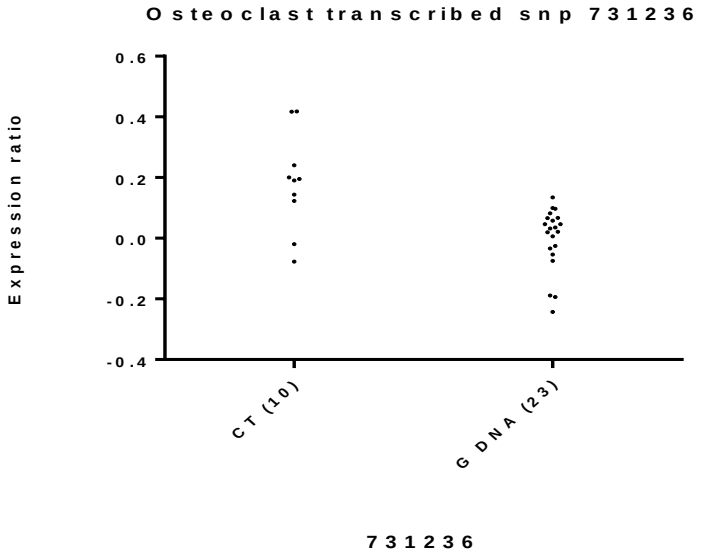


Figure 5-32: Effect of genotype at rs731236 on allelic expression ratio of transcribed VDR SNP rs731236

Each point of the graph is representative of the allelic expression ratio for the transcribed SNP rs731236 for each individual. Transcribed and risk SNPs are same and data of heterozygous (C: T) samples have been shown. The genomic DNA sample mean was used as a normalization control. ***(A)*** *the blood sample genotype of SNP rs731236 for the transcribed SNP rs731236 (μ_{cis}=0.12, P-value=$1.8e^{-08}$) for the comparison with genomic DNA between different genotypes of rs731236).* ***(B)*** *The osteoclast sample genotype of SNP rs731236 for the transcribed SNP rs731236 (μ_{cis}=0.18, P-value=0.002 MSK test for checking the evidence of a cis-acting effect with the transcribed SNP rs731236).*

5.4 Discussion

In the previous chapter our data suggested that polymorphisms which have an effect on the variation of bone mass and diseases such as osteoporosis may also contribute to osteoclast function. GWA studies now commonly use SNP maps and this method has revolutionised the mapping of common genetic loci. As a result of these mapping methods, the susceptibility to a wide range of common, multifactorial disorders can now be determined (Wellcome Trust Case Control 2007). Current evidence suggests that few of the causal loci within the genome consist of variants which directly affect the protein amino-acid sequence. Whereas coding variants are rare, regulatory SNPs appear to be common (Heap, Yang et al. 2010).

Variation in gene expression can result from alterations in the sequence of regulatory elements including SNPs, and recent reports indicate that this phenomenon is widespread throughout the genome and tissues (Cheung, Conlin et al. 2003) (Morley, Molony et al. 2004). These alterations could explain up to 25-35% of the inter-individual variations between allelic gene expression (Pastinen and Hudson 2004). Hence identification and characterisation of AEI can aid in understanding the level of regulatory variation that is functionally important. Identifying candidate SNPs with variations in allelic expression may provide a means to determine significant links between genetic variation and complex traits, including osteoporosis (Palacios, Gazave et al. 2009).

This chapter looked at initially selecting candidate samples to be used in the AEI study based on highest heterozygosity. Time and cost limitations are always a major limitation in independent genetic correlation studies outside of the GWAS and these problems were applicable within the present study. Careful selection of samples which had a high heterozygous genotype was essential for the present study, as expression imbalances of each allele could only be defined in individuals heterozygous for that allele. Correlations between risk SNPs and AEI of transcribed SNPs in the samples which were heterozygous for transcribed SNPs were then determined.

The transcribed SNPs which had the highest heterozygosity were initially identified from the literature from previously studied populations for selected candidate genes. The transcribed SNPs associated with ESR1, OPG, LRP5, RANK, RANKL, P2RX7 and VDR candidate genes were identified as having the highest heterozygosity according to NCBI search

(**Section 2.1**). The selected transcribed SNPs for each candidate gene were determined as: ESR1-rs2077647, OPG-rs2073618, LRP5-556442, RANK-rs1805034, RANKL-rs9562415, P2RX7-rs208294, and VDR-rs731236. Genotyping of the 501 samples within the present study showed that our population was not highly heterozygous for RANKL-rs9562415. Out of the 501 samples, only 1% was heterozygous for RANKL-rs9562415 and therefore this SNP was unsuitable for the AEI study and was omitted. Genotyping for ESR1-rs2077647, OPG-rs556442, RANK-rs2073618, P2RX7-rs208294, LRP5-rs1805034 and VDR-rs731236 showed that approximately 20% of our 501 sample population were heterozygous for at least four of these transcribed SNPs.

Out of the 501 available volunteer samples that we had access to, a total number of 52 subjects were selected, based on having the highest level of heterozygosity at each of the specified transcribed SNPs. We determined that the genotyping data obtained from 52 candidates would be preliminarily sufficient to produce statistically viable results within the AEI study. These statistics were based on computing the effective sample size and statistical power using sample size tables for clinical studies software.

We also showed the expression of ESR1, LRP5, OPG, RANK, P2RX7 and VDR in the candidate blood samples and this was in line with findings in the literature, which have shown white blood cells express the latter genes (www.biogps.org). The expression of ESR1, LRP5, RANK, P2RX7 and VDR genes was also determined in osteoclast cultures with the exception of OPG. This was an expected result as OPG, which encodes the OPG protein, is mainly expressed by osteoblast cells rather than osteoclast cells (www.biogps.org). The integrity of the cDNA samples determined using quantification of GAPDH expression showed that our cDNA isolated from candidate samples was viable to perform further gene expression analysis. These validator experiments of determining expression of the genes of interest and ensuring that the cDNA samples were viable in our candidate samples were essential for performing AEI analysis.

Allele specific expression is used to compare the expression of the two alleles in the same individual as an alternative method to the mapping of regulatory variants. For this approach the gene under study should be polymorphisms in the transcript, such as SNPs. The relative expression of the two alleles is quantified by using these polymorphisms in heterozygous individuals. The expression of the gene is under *cis*-regulation when the expression of the alleles is not equal (allelic imbalance). As both allele expression are measured in the context

of the same genetic (e.g. *trans*-acting effects) technical, and environmental factors relative abundance of allele specific transcript will reflect *cis*-acting effects only (Campino, Forton et al. 2008)

Following the selection of the 52 individuals with the heterozygosity, correlation analysis was performed between heterozygous transcribed SNPs and risk SNPs. Analysis for each heterozygous transcribed SNP determined the number of risk SNPs which were homozygous (risk or wild) and which were heterozygous. These results were particularly important as they could determine any associations between risk SNPs and transcribed SNPs following AEI analysis. Correlation between risk SNPs and transcribed SNPs in OPG, LRP5, and VDR showed the majority of samples heterozygous for selected risk SNPs were heterozygous for selected transcribed SNPs.

In order to study AEI across our sample population we performed a study in which genotyping was coupled with an analysis of allele-specific gene expression by screening four transcribed SNPs to identify differential allelic expression using Sequenom technology. The cDNA isolated from osteoclast and blood tissue was synthesised one base pair at a time and Sequenom was used to determine which specific base had been added at each individual step of the sequencing procedure. Therefore we hypothesise that the association between candidate risk SNPs on non-coding DNA sequences and BMD is mediated through specific allelic imbalance expression of transcribed SNPs.

As mentioned previously, sample size was the main limiting factor in this study. The sample size for detecting associations between disease and SNP markers is known to be highly affected by disease allele frequency, disease prevalence, effect size of the genetic variants (e.g., relative risk, odds ratio, etc.), and linkage disequilibrium (LD) (Pfeiffer and Gail 2003; Scherag, Muller et al. 2003). Previous studies have presented that a population-based design can be more powerful than a family-based study design in investigating genes influencing human complex traits (Risch and Teng 1998) as in our study, which was cohort study.

In the next part of the population genetics study, LD was used to determine specific associations between SNPs associated with ESR1, OPG, LRP5, RANK, P2RX7 and VDR candidate genes. Of particular interest to this study were the LD values determined between risk and transcribed SNPs. A complete LD between risk and transcribed SNPs is indicative that there is a known phase between alleles of risk and transcribed SNPs. Here the parameter μ_{cis} can be estimated by a simple group comparison (simple t-test). However, when the phase

cannot be determined a specialist program is required which can estimate the phase of the haplotypes and estimate the μ_{cis} at the same time, as was the case in the present study, which used the MSK test. According to the LD results, a complete LD between transcribed SNPs rs545382 and risk SNPs599083 (D'=100) was present. In rs731236 of VDR gene, the risk and transcribed SNP are the same, resulting in perfect disequilibrium. For other risk and transcribed SNPs there were no phase between cis and transcribed alleles or it was unknown and MSK test was used for AEI analysis.

Following on from the LD analysis, Sequenom technology was used to determine the allelic expression ratios for four transcript SNPs: rs2077647, rs2073618, rs1805034, and rs731236 in 53 cDNA samples derived from blood and 16 cDNA obtained from osteoclast cultures. The allelic expression ratios for G, A, C and T were performed in quadruples. Here it was important to determine the SEM within the quadruple analysis data points for each individual so that the data obtained could be deemed reliable. Due to the sensitive nature of the Sequenom technology, small technical errors could potentially result in greater inter-individual data variances. Sequenom analysis showed that for all OPG, ESR, RANK and VDR, the inter-individual variances represented by the SEM were within a statistically acceptable range of <0.15. In addition the outlier samples were removed from the set of data as these values often distort the distribution of allelic expression ratios.

It is important to acknowledge that the LRP5 were excluded from the Sequenom analysis due to technical issues and the results of P2RX7 were not useable. The cDNA was extracted from subjects in Iran and, following transportation to the UK, quantified. Based on the Sequenom requirements, the cDNA concentration was sufficient for a total of quadruple replicates performed in the assay. After careful consideration it was decided that OPG, ESR, RANK, P2RX**7** and VDR would be analysed for allelic imbalance expression and the LRP5 would be omitted from analysis. The reasoning for this omission was that the LRP5 required a new multiplexes assay. This was an added cost and it was decided to make the analyse in a separate study at a future date. The results from P2RX7 assay were not reliable, as data showed the primer was extending even when there was no DNA present (in the blanks and primer only samples). Therefore the results of P2RX7 were excluded from analysis: this gene was omitted from the current study, scheduled for re-analysed in a separate study at a future date with LRP5.

We would expect that allelic expression ratio between genomic DNA samples would be close to the log of 1 (expression ratio=0). Therefore any differences in expression ratio observed in blood or osteoclast samples compared to our genomic DNA control could potentially be an indication of allelic expression imbalances. To ensure that each expression ratio observed was a true value, expression ratios were matched and normalised to their corresponding genomic DNA controls. Our data showed that there were consistent inter-individual variations within the expression ratios of the four replications. There was also consistent variation between the allelic expression ratios for individuals within all three blood, osteoclast and genomic sample types. Interestingly, most of the expression ratios of ESR1-rs2077647 in osteoclast culture and blood samples were below zero. In the VDR-rs731236 the expression ratios were above zero. These data indicated that these two transcribed SNPs have a cis-acting effect.

To assess further the cis-acting effect with the transcribed SNPs for ESR1, OPG, RANK and VDR genes, the expression ratios of the transcribed SNPs were determined for the various allelic genotypes for a specific risk SNP for each individual. The various genotypes included homozygous risk, heterozygous and homozygous wild type. The study examined the association of risk variants of low BMD with allelic expression of transcribed SNP 2077647 of ESR gene. From the present study, there was evidence of a cis-acting effect with the transcribed SNP 2077647 of ESR. One statistically significant cis-acting effect was seen for risk SNP rs2504063 in both blood and osteoclast cDNA and statistically significant association was observed for risk SNP rs2941740 in the blood sample only. Cis-acting effects were also seen for both risk SNPs rs4355801 and rs2062375 in osteoclast cDNA with the transcribed SNP rs2073618 of OPG. There was evidence of a cis-acting effect for risk SNP rs3018362 in osteoclast cells with the transcribed SNP rs1805034 of RANK.

There was evidence of a cis-acting effect associated with the transcribed SNP rs731236 of the VDR. Interestingly this cis-acting effect was observed for the risk SNPs rs731236, which is also a transcribed SNP. From this data there is compelling evidence for a correlation between the BMD-associated SNP rs731236 and AEI of VDR candidate gene in both blood and osteoclast samples. Allelic imbalance expression of rs731236 was also shown by Grundberg et al in 2007. They investigated the function of the VDR 3` haplotypes by studying allelic imbalance analysis of VDR transcripts in the normal chromosomal context of 70 unrelated human trabecular bone samples. This was performed by the quantitative genotyping of coding polymorphisms in RNA samples and in corresponding DNA samples. Both were

sequenced in parallel from the trabecular bone samples. Grundberg et al also showed that over-expression of VDR baT haplotype could be considered as a risk allele for osteoporosis. One of the SNPs in the 3` end of the VDR gene, which Grundberg et al defined as baT haplotype, is rs731236 (Grundberg, Lau et al. 2007). This is consistent with the present study which showed cis acting effect with rs731236. There is no other paper which investigates cis acting effect of osteoporotic risk SNPs. One of the most important goals of genetics is to identify (*cis*-) regulatory elements for gene expression and heritable variation of them, especially those that impact clinical phenotypes and common diseases (Sun, Southard et al. 2010). In fact AEI, seems to be an important factor in the variation of human phenotypic and as a consequence, for the development of complex traits and common diseases (Palacios, Gazave et al. 2009).

This study examined the association of risk variants with expression of ESR1, RANK, OPG and VDR genes in both blood samples and osteoclast culture. Allelic imbalance expression in osteoclast cells as an important cell in osteoporosis was not previously investigated, and this study showed cis-acting effect in osteoclast cells for the first time. Such *in vitro* studies have limitations, especially in osteoclast cultures, first results depend on the methods and material that are used. Second expression is investigated outside of the normal body context, which may not show the same expression as in complex tissues *in vivo* (Pastinen and Hudson 2004; Cirulli and Goldstein 2007). Allelic expression analysis in blood samples allows i*n vivo* assessment of RNA transcripts in their regulatory context and native environment, which has higher sensitivity to detect cis-acting effects. Expression could be tested in other cells, which are potentially relevant to BMD regulation. However, many cis-acting effects on gene expression are predicted to be the same in various cell types (Mahr, Burmester et al. 2006). This study analysed the SNPs influencing expression in IMOS-D cohort of Iranian population for the first time. Although this cohort is of mixed ethnicity, the SNPs associated with expression may vary in other populations due to differing LD patterns and allele frequencies (Goddard, Hopkins et al. 2000). Further studies will be necessary to consider the association of other SNPs at present and other candidate genes associated with BMD in Iranian and other populations.

Chapter 6 – General discussion

Summary and Discussion

Genetic variation plays an important role in osteoporosis by reducing bone mass and increasing the risk of fragility fractures. Many different genetic variants contribute to the regulation of BMD and other determinants of fracture risk, including ultrasound properties of bone, skeletal geometry, and bone turnover. It is becoming increasingly obvious that most common traits are influenced by large numbers of distinct loci that individually have small effects. It is now known that the majority of complex trait alleles regulate gene expression by modulating transcription or affecting the stability of the transcript (Ralston 2010; Li, Hou et al. 2010).

Over the past ten years, there has been extensive progress in identifying links between genetics and osteoporosis on an international level, particularly following numerous GWAS. However, the Iranian population has been studied less than some other populations for effective genetic variants in osteoporosis. Due to population differences in allele frequency, the Iranian population may have different relative risk factors than other populations causing different results. There is little known in relation to the effect of genetic variants in the function of osteoclast cells and especially the effect of risk-acting factors on the transcribed SNPs in osteoclast cells. The present study investigated the effect of specific genetic variants in the BMD and functions of osteoclast cells of the Iranian population. This study went on to investigate the effect of risk-acting factors on the transcribed SNPs in osteoclast cultures and peripheral blood within our population.

To develop pharmaceutical agents that arrest the progression of osteoporosis, it is essential to determine the mechanism by which osteoclasts are formed and how they destroy bone. In the past years there has been great progress in understanding the molecular mechanisms involved in osteoclastogenesis. Interestingly this understanding has come mainly from results of genetic studies of abnormal bone phenotypes in animal models and humans. From these genetic observations there is an array of new osteoporosis therapies for both prevention and treatment of osteoporosis (Teitelbaum and Ross 2003). One way to study the function of genes in osteoporosis is to focus on the osteoclast cell itself with regard to its central role in the pathogenesis of osteoporosis (Teitelbaum and Ross 2003).

GWASs have successfully detected associations between hundreds of common SNPs and complex traits in humans (Altshuler, Daly et al. 2008; Stranger, Stahl et al. 2011). While this gene catalogue has uncovered important biological insights, the discovered associations are known to be dependent on variations in sex, age and BMI (Johnson et al. 2009; Shen et al. 2005). This corresponds to the differences seen between findings in genetic association studies. In the present study the descriptive analysis, which included age, sex, and BMI, of the subjects from our Iranian population cohort, was comparable to those observed in other similar studies using Iranian populations (Bahrami, Sadatsafavi et al. 2006). Similarly the BMD trends of our Iranian population samples showing a high male BMD compared to females was consistent with other BMD data obtained from other Iranian (Larijani, Moayyeri et al. 2006; Omrani et al 2006), European (Styrkarsdottir, Halldorsson et al. 2010) and American groups (Mazess and Barden 1999). Taken together, our randomly selected subjects showed similarities to other populations in their descriptive factors data, thus increasing the viability of future genetic association studies.

In the first part of the study, using a random selection of subjects from an Iranian population, we looked at whether the BMD used here as the most common predictor of osteoporosis bone disease was associated with a selection of candidate genes previously associated with BMD. A thorough review of osteoporosis susceptibility loci via large-scale association studies, meta-analyses and GWAS of SNPs revealed 15 BMD genes linked to osteoporosis (Li, Hou et al. 2010). In the present study, LRP5, RANK, RANKL, P2RX7, ESR1, OPG and VDR genes were selected due to their important roles in oestrogen endocrine signalling, Wnt/β-catenin signalling, RANKL/RANK/OPG pathway, and the vitamin D synthesis pathway (Li, Hou et al. 2010).

There is a lack of data demonstrating the effect of genetic variants in osteoporosis-associated LRP5, RANK, RANKL, P2RX7, ESR1, OPG and VDR genes in the Iranian population. We sought to identify the most highly associated SNPs for the osteoporosis-associated LRP5, RANK, RANKL, P2RX7, ESR1, OPG and VDR genes to perform genotyping and hypothesised that there are genomic associations between specific SNPs and BMD within the Iranian population. Using genotyping specific and allele specific analysis in the present study, genetic associations were determined between BMD and 37 selected candidate SNPs. From these associations 10 SNPs were significantly associated with BMD ($P<0.05$), thus suggesting there is evidence of links with this set of genetic loci in the Iranian sample.

For the RANK gene the present study showed that SNP rs884205 was associated with total hip BMD: the same association was observed in studies by Styrkarsdottir et al within a European and East Asian population (Estrada, Styrkarsdottir et al 2012; Rivadenier, Styrkarsdottir et al 2009). The present study further showed that RANK SNPs rs9594738 and rs9594759 were also linked closely to spine BMD, similar to studies by Duncan et al and Styrkarsdottir et al in the European population (Duncan, Danoy et al 2001; Styrkarsdottir, Halldorsson et al 2008). The OPG SNP rs2073618, rs1995824 and rs4355801 associations with spine BMD determined in the present study were also observed in other studies specific to spine and hip BMD in post/pre-menopausal women and the European population respectively (Mencej-Bedrač, Preželj et al. 2011; (Shang, Lin et al. 2013).

The ESR1 SNP 2077647 associated with spine BMD in the present study was also associated with post-menopausal osteoporosis in Japanese women in a recent study (Sonoda, Takada et al 2012). LRP5 SNP rs2306862 was associated with BMD in men and women in the present study as well as in a larger study carried out by Kiel et al. Interestingly, where we observed an association with spine BMD, and the Kiel study showed an association of rs2306862 with femoral neck BMD in men only. Although there are few data on P2RX7, one study did correlate with ours in that SNP rs3751143 was associated with BMD in a similar number of subjects (Wesselius, Bours et al. 2013). From our data we infer that BMD is genetically dependent, and these findings correlated with the literature (Arden, Baker et al. 1996).

In the second part of the study we predicted that there are genomic associations between specific SNPs and osteoclast function within our Iranian population sample. Polymorphisms affecting osteoclast activity could potentially be linked to osteoporosis. The RANKL/RANK/OPG axis is now recognised as the pivotal regulator of osteoclastogenesis, and mutations in the RANK, OPG, and RANKL genes cause osteoporosis (Kobayashi, Udagawa et al. 2009). Increased osteoclastic bone resorption and increased bone resorption markers have been linked to osteoporosis characterised by low skeletal BMD (Gough, Sambrook et al. 1998). However, there are limited data on the genetic associations of osteoclast function and BMD or RANK, RANKL and OPG associated SNPs, particularly in the Iranian population.

The *in vitro* assessment of osteoclast number, resorption activity and apoptosis rates allowed investigation of the relationship between genetic variants and osteoclast function (Gartland, Buckley et al. 2003; Chan, Gartland et al. 2007). The effective culturing of human osteoclasts

in vitro is technically demanding: our group has invested in optimising the method for primary osteoclast isolation and culture from human peripheral blood (Agrawal, Gallagher et al. 2012). In the present study we were able to characterise osteoclasts and quantify the resorption activity of osteoclasts in agreement with the literature (Schilling, Linhart et al. 2004, Kreja, Liedert et al. 2007).

Based on this optimised osteoclast model we were able to determine the osteoclast number and active resorption state of the cultures isolated from the sample group, and assess genetic associations between osteoclast function and the specific SNPs. Finally we sought to determine potential genetic links between the osteoclast genes, associated polymorphisms, and BMD. A single osteoclast culture could be performed to fulfil a dual purpose of functionally characterising the osteoclast cells *in vitro* and determining allelic imbalance expression. The limitations for this part of the study were the high cost of material for culture in remote sites, and the need for the large blood volume sample required from each subject. Within our project budget, approximately 16 cultures were affordable in a twelve months period in Iran.

Osteoclast functional studies showed that at least 30% of all the osteoclasts obtained from our human donors were actively resorbing. The individuals classified as having osteoporosis in L2-L4 also had increased osteoclast number and resorption activity compared to normal individuals. Of particular interest is that our findings suggest that the associations observed with osteoclastogenesis and BMD appear to be specific to low L2-L4 BMD. These findings are consistent with other studies that demonstrate that genetic associations with BMD are bone site dependent (Tran, Nguyen et al. 2008). One limitation of most BMD GWASs is that subjects are usually postmenopausal women and elderly men (Paternoster, Ohlsson et al. 2010), meaning that the genetic effects reported could actually be related to age-related bone loss or to the control of peak bone gain in early life. It could be reasoned that genetic variation in the regulatory RANK/RANKL/OPG mechanisms involved in osteoclastogenesis may influence the rate of bone loss in older subjects.

Identifying the risk SNPs associated with osteoclastogenesis regulators RANK, RANKL and OPG enabled the identification of related gene polymorphisms (Kikuchi, Matsuguchi et al. 2001). Genotyping and osteoclast function assays showed that RANK polymorphisms had a significant genetic association with osteoclast number. An increase in the number of osteoclast number was linked to heterozygous alleles of RANK rs884205 and rs30183625

SNPs compared to homozygous wild type allele. There were no genetic associations between osteoclast number/function and OPG or RANKL SNPs.

The osteoclast number and activity was lower in individuals with low L2-L4 BMD but not total hip BMD. Again, it is particularly interesting that out of the 16 individuals selected on the basis of strongest genetic heterozygosity from our study cohort, >70% were either osteopenic or osteoporotic. Taken together, these findings suggest that there may be genetic associations between specific RANK/RANKL polymorphisms and osteoclast number and function.

In the final part of the present study, transcribed SNPs and risk-acting SNPs associated with the LRP5, RANK, P2RX7, ESR1, OPG and VDR candidate genes were identified. Samples from each subject were genotyped according to the percentage of heterozygosity of the defined transcribed SNPs within our population. This study facilitated the investigation of the risk-variant in the transcribed SNP and was important as most of the variations contributing to disease susceptibility, especially in complex disease, may alter the expression rather than the function of proteins (Leonardson et al. 2004). Identifying candidate SNPs with variations in allelic expression may provide a means to determine an important link between individual genetic variation and complex traits including osteoporosis (Palacios, Gazave et al. 2009). Allele-specific expression was investigated by evaluation of cDNA from tissues expressing the genes of interest in the heterozygous individual (Marsh 2007), in both osteoclast culture and peripheral blood. The relative amount of transcript arising from each allele in heterozygous individuals was measured in this technique. Risk-acting effects on gene expression are shown by unequal amounts of transcript from each allele (Cunnington, Kay et al. 2009).

The data from chapter 3 showed that both ESR1-rs2077647 and VDR-rs731236 were associated with BMD. Interestingly, AEI analysis of these genes also indicated that most of the expression ratios of transcribed SNPs ESR1-rs2077647 and VDR-rs731236 in osteoclast culture and blood samples had a cis-acting effect. The expression ratios of the transcribed SNPs were determined for the various allelic genotypes for a specific risk SNP for each individual. Statistically cis-acting effect was evident for risk SNPs rs2504063 and rs2941740 on ESR1 and for risk SNP rs731236 on VDR. Cis-acting effects were also seen only in osteoclast cells for both risk SNPs rs4355801 and rs2062375 with OPG and for risk SNP rs3018362 with RANK.

Although several AEI studies have been conducted to date, according to our knowledge only one study has investigated AEI in osteoporosis. This AEI study has focused on one haplotype in the 3` end of the VDR gene (Grundberg, Lau et al. 2007). Grundberg *et al* showed over-expression of VDR baT haplotype compared with other haplotypes in 70 unrelated human trabecular bone samples, and also showed that over-expression of VDR baT haplotype could be considered a risk allele for osteoporosis. One of the SNPs in the 3` end of the VDR gene, which Grundberg *et al* defined as baT haplotype, is rs731236 (Grundberg, Lau et al. 2007). Interestingly these findings are consistent with the present study, in which we showed cis acting effect with rs731236.

The present study extends our understanding of the AEI in some important candidate genes in BMD. It was shown that there was a relationship between some risk variants and expression of ESR1, RANK, OPG and VDR genes. Identifying (cis-) regulatory elements for gene expression, especially those that impact clinical phenotypes and common diseases is one of the most important goals of genetics (Sun, Southard et al. 2010). In fact, one of the important factors in the variation of human phenotypes and, as a consequence, for the development of complex traits and common diseases is AEI (Palacios, Gazave et al. 2009).

In the present study osteoclast cells were used for AEI for the first time. The AEI in osteoclasts, cells known for their functional importance in osteoporosis, has not been previously investigated. For the first time the findings from the present data provided evidence to suggest a cis-acting effect in osteoclast cells. Although such *in vitro* studies have limitations, the results and their interpretation primarily depend on the methods and material that are used, and also have to take into consideration that the expression outside of the normal body context may not necessarily reflect the same expression as in complex tissues *in vivo* (Pastinen and Hudson 2004; Cirulli and Goldstein 2007).

An interesting approach employed in the current studies was the use of Iranian ancestry for association analyses. The Caucasian populations have primarily been studied for characterisation of genetic variation, such as SNPs and genomic sequence, and populations of European ancestry have been studied for most association studies performed to date (Manolio, Brooks et al. 2008). Investigation of diverse populations has the significant ability to develop the mapping of the genetic determinants of complex diseases for the whole human population. By investigating less studied populations it is possible to consider new findings that are perhaps of a higher prevalence in non-European populations, or establish the

generality of findings obtained initially in the European ancestry (Rosenberg, Huang et al. 2010).

Suggestions for Future Study:

The key issue for determining statistically significant outcomes from any meta-analysis or genetic-association study is heavy reliance on the total number of subjects participating. In the present study there were issues in recruiting and obtaining enough samples from the donors. Resolving restricted factors, including the recruitment of human donors, would allow further for the selection of more subjects for osteoclast cultures. As a result it would be possible to increase the power of statistical tests to detect a theoretical difference less than 0.8. This would mean that, based on more reliable results, a clinical outcome would be more likely. To build on the data for the present study, recruiting strategies would have to be improved, and the study extended nationally within Iran for both higher subject numbers and to obtain a better clinical reflection of the whole population.

Bone is a dynamic tissue with interactions between osteoblasts, osteoclasts and osteocytes vital for its maintenance, and for responses to strain and stress. This study explores the effects of genetic variation on osteoclast bone cells in isolation, ignoring the regulatory effects that bone cells have on each other. This could be overcome to a certain extent by evaluating the changes in gene expression of several regulatory molecules in these cells (according to genetic variations) in the context of bone tissue (Raggatt and Partridge 2010). As discussed previously, osteoblast cells have a critical role in bone maintenance and remodelling (Crockett, Rogers et al. 2011). Consequently dysfunction of osteoblast cell activity can result in development and progression of osteoporosis. The present study has focused on building upon the knowledge we have of the genetic association between osteoporosis and osteoclast function, in particular with regard to the Iranian population. However, due to the coupling relationship between osteoclast and osteoblast cells, it is of equal importance to study also BMI-osteoporosis genetic associations in osteoblast cells, with a particular interest in exploring the AEI of candidate genes in osteoblast cells.

As previously mentioned, the results from the P2RX7 assay were not reliable, as data showed the primers were extending even in the absence of DNA. As part of future study plans, these primers must be to re-designed and optimised prior to performing Seqeunom assay. Also LRP5, which was omitted from the Sequenom assay due to added cost, would be scheduled for re-analysis in a separate study at a future date alongside P2RX7.

This study analysed for the first time the SNPs influencing expression in an IMOS-D cohort of Iranian population. The SNPs associated with expression may vary in other populations due to differing allele frequencies and LD patterns (Goddard, Hopkins et al. 2000). Future studies will be necessary to investigate the SNPs influencing expression of selected candidate genes in other populations. This project provides an insight into the role of genetic variation in BMD and also analyses the SNPs influencing expression in the Iranian population. This study and other similar studies help to improve our knowledge about the role of genetic variation in osteoporosis, with the goal of providing a genetic profile that can be used for genetic screening tests for the identification of individuals at risk of osteoporosis (Wu, Guo et al. 2013). At the present time performance of genetic profiling tests is weak in complex diseases because associations between component SNPs and diseases are weak. Genetic profiling tests could be useful in screening. This would be after extending our knowledge in different populations to find either an extremely large panel of SNPs, each with a modest association with disease, or a few SNPs with very strong associations with disease (Simmonds 2011). A suitable genetic profiling test helps to find useful treatment. Novel pathways that contribute to control of BMD could be identified, improving our knowledge about the effect of genetic variation with the possibility of producing new therapeutic agents (Duncan and Brown 2008).

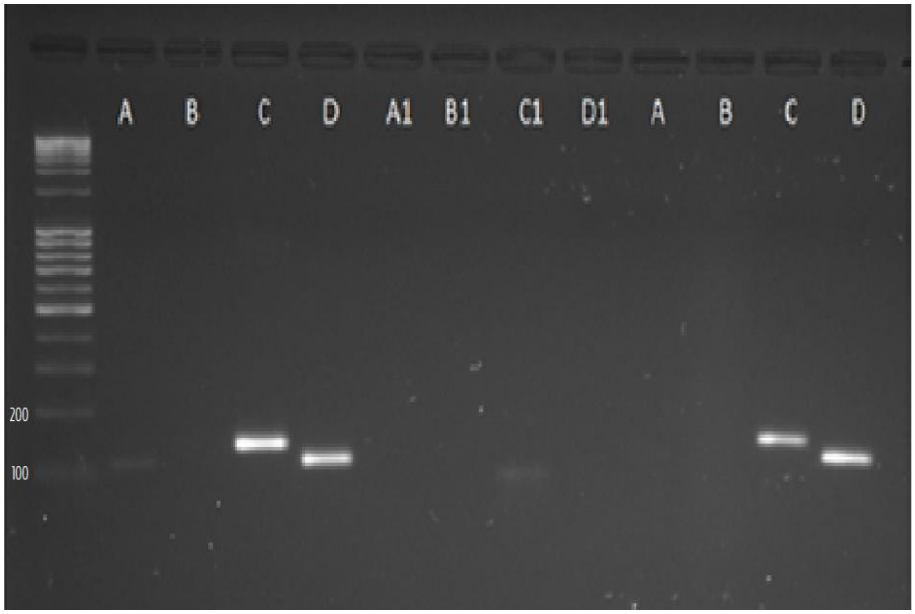
Appendix

Primers of the selected SNPs to be used in pyrosequencing were optimised in different concentration of MgCl2

To perform pyrosequencing, Dr Reynard kindly designed the primers specific to the candidate SNPs to be used subsequently in the AEI study (**Table A.1**). End-point PCR was performed to optimise these primers using Promega GoTaq® DNA polymerase, as outlined in the **method section 2.9.2.** Product with no cDNA templates were also used for detection of contamination in samples. The baseline PCR amplification reactions were pre-incubated at 94°C for 2 minutes, followed by 34 cycles of denaturation at 94°C for 60 seconds, anneal at 55°C for 30 seconds and extension at 72°C for 30 seconds. The PCR amplified products were visualised on an electrophoretic 2% agarose gel stained with ethidium bromide and product bands detected using the Gel Doc XR System and Quantity One Software as described in the **method section 2.9.2.**

The optimized conditions for each primer were determined by visualising the amplified product under varying annealing temperatures and MgCl2 concentrations. As it is shown in the table A.1, different melting temperature (TM) and MgCl2 concentrations were checked for primers A to E to find the optimised state for performing end-point PCR.

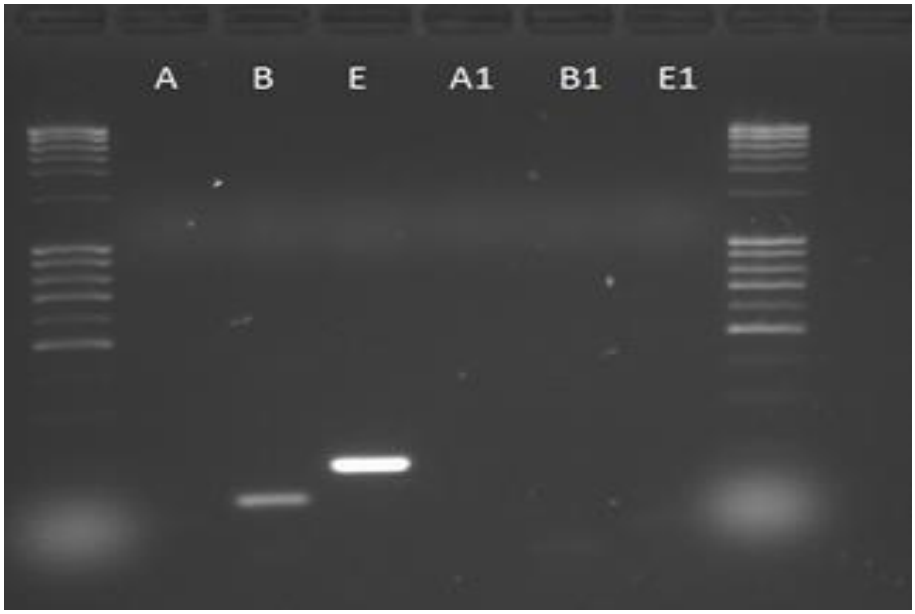
The optimised states for performing end-point PCR for primers of OPG and LRP5 were 1.5Mm MgCl2 and 55°C TM. The optimised states for performing end-point PCR for primers VDR and ESR-Rs2077647 (F4, R4) of ESR1 were 2Mm MgCl2 and 58°C TM. The PCR amplified product of the Rs2077647 (F1, R1) primer of ESR1 optimised by using 2Mm MgCl2 and 55°C TM. Finally the optimised states for performing end-point PCR for P2RX7/SNP primers were 2.0mM MgCl2 and 55°C TM (**Figure A.1-3** and **Table A.1**).



Candidate gene/SNP	ID
ESR1	
Rs2077647	A
Rs2077647	B
OPG	
Rs2073618	C
LRP5	
Rs556442	D

Figure A.1: **End-point PCR optimisation using 1.5Mm MgCl2 and 55°C TM**

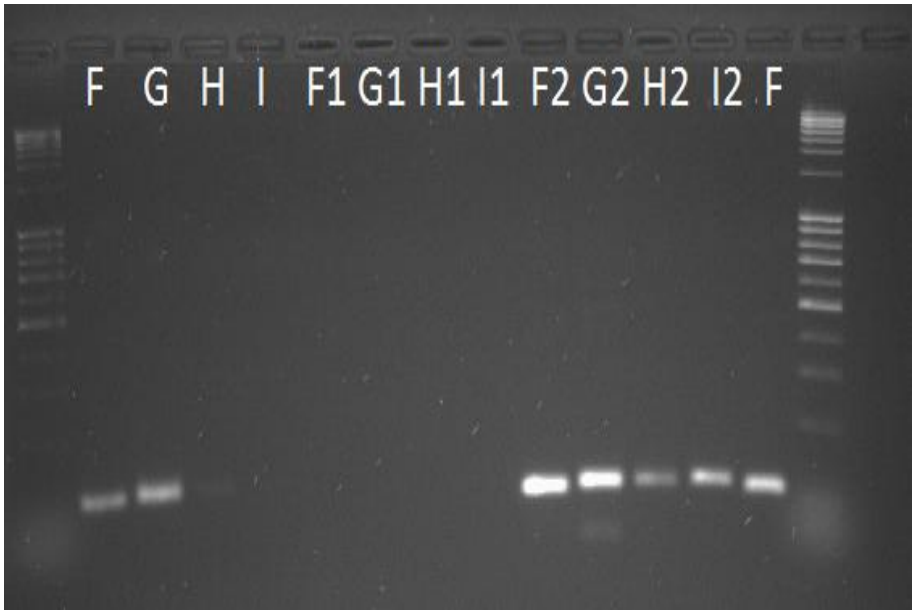
A1, B1, C1, and D1 were –ve control by water indicating no contamination in samples. Two bright bands were detected between 100 and 200 base pairs for primers C and D. The bands for primers A and B were not visible under the same PCR conditions. Both A and B are different primer for rs2077647.



Candidate gene/SNP	ID
ESR1	
Rs2077647	A
Rs2077647	B
VDR	
Rs731236	E

Figure A.2: **End point PCR optimisation using 2Mm MgCl2 and 58°C TM**

Two bright bands were visible between 100 and 200 base pairs for primers B and E. The band using primer was not visible. A1, B1, and E1 were –ve control by water for DNA contamination. Both A and B are different primer for rs2077647.



Candidate gene/SNP	ID
P2RX7	
Rs7958311	F
Rs7958311	G
Rs1718119	H
Rs1718119	I

Figure A.3: End point PCR optimisation using 1.5Mm or 2.0mM MgCl2 and 55°C TM for P2RX7/SNP primers

F, G, H, and I represent bands optimised using 1.5mM MgCl2 and F2, G2, H2, and I2 using 2mM MgCl2. Band representing PCR products appeared brighter using 2mM MgCl2. F1, G1, H1, and I1 were –ve control by water for DNA contamination. Both F and G are different primer for rs7958311.

SNP	ID	Primer	Tm (°C)	Mg^{2+} (mM)
Rs2077647 F1	A	5’-ACCATGACCCTCCACACCA-3’	55	2.0mM
Rs2077647 R1	A	5’-CTCCAG GGGGATCTTGAGCT-3’	55	2.0mM
Rs2077647 F4	B	5’-GGGGACACG GTCTGCACC-3’	58	2.0mM
Rs2077647 R4	B	5’-TTGAGCTGCGGACGGTTC-3’	58	2.0mM
Rs2073618 F1	C	5’-GAGCGCGCAGCACAG CAA-3’	55	1.5mM
Rs2073618 R1	C	5’-CTGTTG CCGGGACGCTAT-3’	55	1.5mM
Rs556442 F1	D	5’-CGAGCGCGAGGTCCTCTT-3’	55	1.5mM
Rs556442 R1	D	5’-CAATGCGCTTCAGGTCCG-3’	55	1.5mM
Rs731236 F1	E	5’-GGAAGGAGAGGCAGCGGT ACT-3’	58	2.0mM
Rs731236 r1	E	5’-TCCTGG GGTGCAGGACGC-3’	58	2.0mM
Rs7958311 F1	F	5’-ACTGGGACTGCAACCTAGACCG-3’	55	2.0mM
Rs7958311 R1	F	5’-TAGCCAGGGTACAAGGACACGT-3’	55	2.0mM
Rs7958311 F14	G	5’-ATCTACTGGGACTGCAACCTAGAC-3’	55	2.0mM
Rs1718119 F1	H	5’-TTTGGAGAAGGAAGTGACTAACG-3’	55	2.0mM
Rs1718119 R1	H	5’-ATGGGAGCGACAGCAGTTACT-3’	55	2.0mM
Rs1718119 F2	I	5’-GTTTGGAGAAGGAAGTGACTAACG-3’	55	2.0mM
Rs1718119 R2	I	5’-ATATGGGAGCGACAGCAGTTAC-3’	55	2.0mM

Table A.1: Optimization states for primers A-I

The optimized conditions for each primer including annealing temperature (TM) and MgCl2 concentration (mM) are summarized.

Bibliography

Agrawal, A., K. A. Buckley, et al. (2010). "The effects of P2X7 receptor antagonists on the formation and function of human osteoclasts in vitro." Purinergic Signal **6**(3): 307-315.

Agrawal, A., J. A. Gallagher, et al. (2012). "Human osteoclast culture and phenotypic characterization." Methods Mol Biol **806**: 357-375.

Altshuler, D., M. J. Daly, et al. (2008). "Genetic mapping in human disease." Science **322**(5903): 881-888.

Arai, H., K. I. Miyamoto, et al. (2001). "The polymorphism in the caudal-related homeodomain protein Cdx-2 binding element in the human vitamin D receptor gene." J Bone Miner Res **16**(7): 1256-1264.

Arden, N. K., J. Baker, et al. (1996). "The heritability of bone mineral density, ultrasound of the calcaneus and hip axis length: a study of postmenopausal twins." J Bone Miner Res **11**(4): 530-534.

Arko, B., J. Prezelj, et al. (2005). "Association of the osteoprotegerin gene polymorphisms with bone mineral density in postmenopausal women." Maturitas **51**(3): 270-279.

Bachrach, L. K. (2001). "Acquisition of optimal bone mass in childhood and adolescence." Trends Endocrinol Metab **12**(1): 22-28.

Bahrami, H., M. Sadatsafavi, et al. (2006). "Obesity and hypertension in an Iranian cohort study; Iranian women experience higher rates of obesity and hypertension than American women." BMC Public Health **6**: 158.

Barrett, J. C., B. Fry, et al. (2005). "Haploview: analysis and visualization of LD and haplotype maps." Bioinformatics **21**(2): 263-265.

Boyle, W. J., W. S. Simonet, et al. (2003). "Osteoclast differentiation and activation." Nature **423**(6937): 337-342.

Bush, W. S. and J. H. Moore (2012). "Chapter 11: Genome-wide association studies." PLoS Comput Biol **8**(12): e1002822.

Campino, S., J. Forton, et al. (2008). "Validating discovered Cis-acting regulatory genetic variants: application of an allele specific expression approach to HapMap populations." PLoS One **3**(12): e4105.

Canalis, E. (2013). "Wnt signalling in osteoporosis: mechanisms and novel therapeutic approaches." Nat Rev Endocrinol **9**(10): 575-583.

Chan, B. Y., A. Gartland, et al. (2007). "PPAR agonists modulate human osteoclast formation and activity in vitro." Bone **40**(1): 149-159.

Cheung, C. L., S. M. Xiao, et al. (2010). "Genetic epidemiology of age-related osteoporosis and its clinical applications." Nat Rev Rheumatol **6**(9): 507-517.

Cheung, V. G., L. K. Conlin, et al. (2003). "Natural variation in human gene expression assessed in lymphoblastoid cells." Nat Genet **33**(3): 422-425.

Cheung, V. G., R. S. Spielman, et al. (2005). "Mapping determinants of human gene expression by regional and genome-wide association." Nature **437**(7063): 1365-1369.

Choi, J. Y., C. S. Shin, et al. (2006). "Single-nucleotide polymorphisms and haplotypes of bone morphogenetic protein genes and peripheral bone mineral density in young Korean men and women." Calcif Tissue Int **78**(4): 203-211.

Christakos, S., M. Hewison, et al. (2013). "Vitamin D: beyond bone." Ann N Y Acad Sci **1287**: 45-58.
Cirulli, E. T. and D. B. Goldstein (2007). "In vitro assays fail to predict in vivo effects of regulatory polymorphisms." Hum Mol Genet **16**(16): 1931-1939.
Clark, S. (2002). "Osteoporosis--the disease of the 21st century?" Lancet **359**(9319): 1714.
Cole, Z. A., E. M. Dennison, et al. (2008). "Osteoporosis epidemiology update." Curr Rheumatol Rep **10**(2): 92-96.
Constancia, M., B. Pickard, et al. (1998). "Imprinting mechanisms." Genome Res **8**(9): 881-900.
Cowles, C. R., J. N. Hirschhorn, et al. (2002). "Detection of regulatory variation in mouse genes." Nat Genet **32**(3): 432-437.
Crockett, J. C., M. J. Rogers, et al. (2011). "Bone remodelling at a glance." J Cell Sci **124**(Pt 7): 991-998.
Cummings, S. R., M. C. Nevitt, et al. (1995). "Risk factors for hip fracture in white women. Study of Osteoporotic Fractures Research Group." N Engl J Med **332**(12): 767-773.
Delaisse, J. M. (2014). "The reversal phase of the bone-remodeling cycle: cellular prerequisites for coupling resorption and formation." Bonekey Rep **3**: 561.
Dewan, A., M. Liu, et al. (2006). "HTRA1 promoter polymorphism in wet age-related macular degeneration." Science **314**(5801): 989-992.
Dhanwal, D. K., E. M. Dennison, et al. (2011). "Epidemiology of hip fracture: Worldwide geographic variation." Indian J Orthop **45**(1): 15-22.
Dixon, A. L., L. Liang, et al. (2007). "A genome-wide association study of global gene expression." Nat Genet **39**(10): 1202-1207.
Ducy, P., T. Schinke, et al. (2000). "The osteoblast: a sophisticated fibroblast under central surveillance." Science **289**(5484): 1501-1504.
Duncan, E. L. and M. A. Brown (2008). "Genetic studies in osteoporosis--the end of the beginning." Arthritis Res Ther **10**(5): 214.
Duncan, E. L., P. Danoy, et al. (2011). "Genome-wide association study using extreme truncate selection identifies novel genes affecting bone mineral density and fracture risk." PLoS Genet **7**(4): e1001372.
Enard, W., P. Khaitovich, et al. (2002). "Intra- and interspecific variation in primate gene expression patterns." Science **296**(5566): 340-343.
Estrada, K., U. Styrkarsdottir, et al. (2012). "Genome-wide meta-analysis identifies 56 bone mineral density loci and reveals 14 loci associated with risk of fracture." Nat Genet **44**(5): 491-501.
Everts, V., J. M. Delaissé, et al. (2002). "The bone lining cell: its role in cleaning Howship's lacunae and initiating bone formation." J Bone Miner Res **17**(1): 77-90.
Fang, Y., F. Rivadeneira, et al. (2006). "Vitamin D receptor gene BsmI and TaqI polymorphisms and fracture risk: A meta-analysis." Bone **39**(4): 938-945.
Fang, Y., J. B. van Meurs, et al. (2003). "Cdx-2 polymorphism in the promoter region of the human vitamin D receptor gene determines susceptibility to fracture in the elderly." J Bone Miner Res **18**(9): 1632-1641.
Fang, Y., J. B. van Meurs, et al. (2005). "Promoter and 3'-untranslated-region haplotypes in the vitamin d receptor gene predispose to osteoporotic fracture: the rotterdam study." Am J Hum Genet **77**(5): 807-823.

Farh, K. K., A. Grimson, et al. (2005). "The widespread impact of mammalian MicroRNAs on mRNA repression and evolution." Science **310**(5755): 1817-1821.
Fauci, A. S., E. Braunwald, et al (2008). Harrison's principles of internal medicine, 17th edition.
Favus, M. J. (2008). Primer on the metabolic bone diseases and disorders of mineral metabolism. Chicago, American society for bone and mineral research.
Felson, D. T., Y. Zhang, et al. (1993). "Effects of weight and body mass index on bone mineral density in men and women: the Framingham study." J Bone Miner Res **8**(5): 567-573.
Feng, X. and J. M. McDonald (2011). "Disorders of bone remodeling." Annu Rev Pathol **6**: 121-145.
Freedman, M. L., A. N. Monteiro, et al. (2011). "Principles for the post-GWAS functional characterization of cancer risk loci." Nat Genet **43**(6): 513-518.
Galena, H. J. (1992). "Complications occurring from diagnostic venipuncture." J Fam Pract **34**(5): 582-584.
Gartland, A. (2012) "P2X receptors in bone

Wiley Interdisciplinary Reviews: Membrane Transport and Signaling Volume 1, Issue 2." Wiley Interdisciplinary Reviews: Membrane Transport and Signaling **1**, 221-227.
Gartland, A., K. A. Buckley, et al. (2003). "Blockade of the pore-forming P2X7 receptor inhibits formation of multinucleated human osteoclasts in vitro." Calcif Tissue Int **73**(4): 361-369.
Gartland, A., R. A. Hipskind, et al. (2001). "Expression of a P2X7 receptor by a subpopulation of human osteoblasts." J Bone Miner Res **16**(5): 846-856.
Gartland, A., K. K. Skarratt, et al. (2012). "Polymorphisms in the P2X7 receptor gene are associated with low lumbar spine bone mineral density and accelerated bone loss in post-menopausal women." Eur J Hum Genet **20**(5): 559-564.
Gaur, U., K. Li, et al. (2013). "Research progress in allele-specific expression and its regulatory mechanisms." J Appl Genet **54**(3): 271-283.
Gennari, L., D. Merlotti, et al. (2005). "Estrogen receptor gene polymorphisms and the genetics of osteoporosis: a HuGE review." Am J Epidemiol **161**(4): 307-320.
Goddard, K. A., P. J. Hopkins, et al. (2000). "Linkage disequilibrium and allele-frequency distributions for 114 single-nucleotide polymorphisms in five populations." Am J Hum Genet **66**(1): 216-234.
Golbahar, J., A. Hamidi, et al. (2004). "Association of plasma folate, plasma total homocysteine, but not methylenetetrahydrofolate reductase C667T polymorphism, with bone mineral density in postmenopausal Iranian women: a cross-sectional study." Bone **35**(3): 760-765.
Gong, Y., R. B. Slee, et al. (2001). "LDL receptor-related protein 5 (LRP5) affects bone accrual and eye development." Cell **107**(4): 513-523.
Gough, A., P. Sambrook, et al. (1998). "Osteoclastic activation is the principal mechanism leading to secondary osteoporosis in rheumatoid arthritis." J Rheumatol **25**(7): 1282-1289.
Grundberg, E., E. M. Lau, et al. (2007). "Vitamin D receptor 3' haplotypes are unequally expressed in primary human bone cells and associated with increased fracture risk: the MrOS Study in Sweden and Hong Kong." J Bone Miner Res **22**(6): 832-840.

Gueguen, R., P. Jouanny, et al. (1995). "Segregation analysis and variance components analysis of bone mineral density in healthy families." J Bone Miner Res **10**(12): 2017-2022.
Guntur, A. R. and C. J. Rosen (2012). "Bone as an endocrine organ." Endocr Pract **18**(5): 758-762.
Harada, S. and G. A. Rodan (2003). "Control of osteoblast function and regulation of bone mass." Nature **423**(6937): 349-355.
Heap, G. A., J. H. Yang, et al. (2010). "Genome-wide analysis of allelic expression imbalance in human primary cells by high-throughput transcriptome resequencing." Hum Mol Genet **19**(1): 122-134.
Hedrick, P. W. (1987). "Gametic disequilibrium measures: proceed with caution." Genetics **117**(2): 331-341.
Herbon, N., M. Werner, et al. (2003). "High-resolution SNP scan of chromosome 6p21 in pooled samples from patients with complex diseases." Genomics **81**(5): 510-518.
Hindorff LA, J. H., Hall PN, Mehta JP, and Manolio TA. A Catalog of Published Genome-Wide Association Studies. Available at: www.genome.gov/gwastudies.
Hindorff, L. A., P. Sethupathy, et al. (2009). "Potential etiologic and functional implications of genome-wide association loci for human diseases and traits." Proc Natl Acad Sci U S A **106**(23): 9362-9367.
Hofbauer, L. C. and M. Schoppet (2002). "Osteoprotegerin gene polymorphism and the risk of osteoporosis and vascular disease." J Clin Endocrinol Metab **87**(9): 4078-4079.
Holliday, K. L., S. R. Pye, et al. (2011). "The ESR1 (6q25) locus is associated with calcaneal ultrasound parameters and radial volumetric bone mineral density in European men." PLoS One **6**(7): e22037.
Hong, E. P. and J. W. Park (2012). "Sample size and statistical power calculation in genetic association studies." Genomics Inform **10**(2): 117-122.
Hossein-nezhad, A., F. N. Varzaneh, et al. (2011). "Association of the FokI polymorphism in the vitamin D receptor gene with vertebral fracture in Iranian postmenopausal women." Endocr Pract **17**(5): 826-827.
Hsu, Y. H., S. Demissie-Banjaw, et al. (2008). "Genome-wide association study of BMD and hip geometry indices. The Framingham Osteoporosis Study." Bone **43**: S34-S34.
Hsu, Y. H., T. Niu, et al. (2006). "Variation in genes involved in the RANKL/RANK/OPG bone remodeling pathway are associated with bone mineral density at different skeletal sites in men." Hum Genet **118**(5): 568-577.
Hsu, Y. H., M. C. Zillikens, et al. (2010). "An integration of genome-wide association study and gene expression profiling to prioritize the discovery of novel susceptibility Loci for osteoporosis-related traits." PLoS Genet **6**(6): e1000977.
Hunter, D., M. De Lange, et al. (2001). "Genetic contribution to bone metabolism, calcium excretion, and vitamin D and parathyroid hormone regulation." J Bone Miner Res **16**(2): 371-378.
Ioannidis, J. P., E. E. Ntzani, et al. (2001). "Replication validity of genetic association studies." Nat Genet **29**(3): 306-309.
Ioannidis, J. P., S. H. Ralston, et al. (2004). "Differential genetic effects of ESR1 gene polymorphisms on osteoporosis outcomes." JAMA **292**(17): 2105-2114.

Ioannidis, J. P., I. Stavrou, et al. (2002). "Association of polymorphisms of the estrogen receptor alpha gene with bone mineral density and fracture risk in women: a meta-analysis." J Bone Miner Res **17**(11): 2048-2060.
Jia, F., R. F. Sun, et al. (2013). "Vitamin D receptor BsmI polymorphism and osteoporosis risk: a meta-analysis from 26 studies." Genet Test Mol Biomarkers **17**(1): 30-34.
Johnell, O., J. A. Kanis, et al. (2005). "Predictive value of BMD for hip and other fractures." J Bone Miner Res **20**(7): 1185-1194.
Kanis, J. A. (1994). "Assessment of fracture risk and its application to screening for postmenopausal osteoporosis: synopsis of a WHO report. WHO Study Group." Osteoporos Int **4**(6): 368-381.
Kanis, J. A. (2002). "Diagnosis of osteoporosis and assessment of fracture risk." Lancet **359**(9321): 1929-1936.
Kanis, J. A., G. Bianchi, et al. (2011). "Towards a diagnostic and therapeutic consensus in male osteoporosis." Osteoporos Int **22**(11): 2789-2798.
Kanis, J. A., N. Burlet, et al. (2008). "European guidance for the diagnosis and management of osteoporosis in postmenopausal women." Osteoporos Int **19**(4): 399-428.
Kanis, J. A. and O. Johnell (2005). "Requirements for DXA for the management of osteoporosis in Europe." Osteoporos Int **16**(3): 229-238.
Kanis, J. A., E. V. McCloskey, et al. (2008). "A reference standard for the description of osteoporosis." Bone **42**(3): 467-475.
Kaprio, J., A. Rimpela, et al. (1995). "Common genetic influences on BMI and age at menarche." Hum Biol **67**(5): 739-753.
Karasik, D. and S. L. Ferrari (2008). "Contribution of gender-specific genetic factors to osteoporosis risk." Ann Hum Genet **72**(Pt 5): 696-714.
Kassim Javaid, M., J. Chana, et al. (2013). "Hip fracture as the tracer condition." Best Pract Res Clin Rheumatol **27**(6): 711-715.
Kiel, D. P., S. L. Ferrari, et al. (2007). "Genetic variation at the low-density lipoprotein receptor-related protein 5 (LRP5) locus modulates Wnt signaling and the relationship of physical activity with bone mineral density in men." Bone **40**(3): 587-596.
Kim, H. S. and S. O. Yang (2014). "Quality Control of DXA System and Precision Test of Radio-technologists." J Bone Metab **21**(1): 2-7.
Kobayashi, Y., N. Udagawa, et al. (2009). "Action of RANKL and OPG for osteoclastogenesis." Crit Rev Eukaryot Gene Expr **19**(1): 61-72.
Koh, L. K. H. (2004). "Osteoporosis: assessment for diagnosis, evaluation and treatment." jmhg, 1.
Korcok, J., L. N. Raimundo, et al. (2004). "Extracellular nucleotides act through P2X7 receptors to activate NF-kappa B in osteoclasts." Journal of Bone and Mineral Research **19**(4): 642-651.
Krakauer, T. (2008). "Nuclear factor-kappaB: fine-tuning a central integrator of diverse biologic stimuli." Int Rev Immunol **27**(5): 286-292.
Krall, E. A. and B. Dawson-Hughes (1993). "Heritable and life-style determinants of bone mineral density." J Bone Miner Res **8**(1): 1-9.
Kreja, L., A. Liedert, et al. (2007). "Influence of receptor activator of nuclear factor (NF)-kappaB ligand (RANKL), macrophage-colony stimulating factor (M-CSF) and fetal calf serum on human osteoclast formation and activity." J Mol Histol **38**(4): 341-345.

Krishnan, V., H. U. Bryant, et al. (2006). "Regulation of bone mass by Wnt signaling." J Clin Invest **116**(5): 1202-1209.
Langdahl, B. L., M. Carstens, et al. (2002). "Polymorphisms in the osteoprotegerin gene are associated with osteoporotic fractures." J Bone Miner Res **17**(7): 1245-1255.
Larijani, B., A. Moayyeri, et al. (2006). "Peak bone mass of Iranian population: The Iranian Multicenter Osteoporosis Study." Journal of Clinical Densitometry **9**(3): 367-374.
Lau, R. Y. and X. Guo (2011). "A review on current osteoporosis research: with special focus on disuse bone loss." J Osteoporos **2011**: 293808.
Lee, Y. H., J. H. Woo, et al. (2010). "Associations between osteoprotegerin polymorphisms and bone mineral density: a meta-analysis." Mol Biol Rep **37**(1): 227-234.
Lei, S. F., F. Y. Deng, et al. (2004). "Bone mineral density in elderly Chinese: effects of age, sex, weight, height, and body mass index." Journal of Bone and Mineral Metabolism **22**(1): 71-78.
Lewis, C. M. and J. Knight (2012). "Introduction to genetic association studies." Cold Spring Harb Protoc **2012**(3): 297-306.
Li, W. F., S. X. Hou, et al. (2010). "Genetics of osteoporosis: accelerating pace in gene identification and validation." Hum Genet **127**(3): 249-285.
Luo, Y., Z. Hu, et al. (2014). "Significant associations between the A163G and G1181C polymorphisms of the osteoprotegerin gene and risk of osteoporosis, especially in postmenopausal women: a meta-analysis." Genet Test Mol Biomarkers **18**(3): 211-219.
Ma, C., Q. L. Zhou, et al. (2014). "Association of vitamin D receptor BsmI gene polymorphism with risk of low bone mineral density in post-menopausal women: a meta-analysis." Genet Mol Res **13**(3): 7791-7799.
Mahr, S., G. R. Burmester, et al. (2006). "Cis- and trans-acting gene regulation is associated with osteoarthritis." Am J Hum Genet **78**(5): 793-803.
Manolagas, S. C. (2000). "Birth and death of bone cells: basic regulatory mechanisms and implications for the pathogenesis and treatment of osteoporosis." Endocr Rev **21**(2): 115-137.
Manolagas, S. C. J., Robert L. (1995). "Mechanisms of Diseases: Bone Marrow, Cytokines, and Bone Remodeling -- Emergency Insights into the Pathophysiology of Osteoporosis." The New England Journal Of Medicin.
Manolio, T. A., L. D. Brooks, et al. (2008). "A HapMap harvest of insights into the genetics of common disease." J Clin Invest **118**(5): 1590-1605.
Marini, F. and M. L. Brandi (2010). "Genetic determinants of osteoporosis: common bases to cardiovascular diseases?" Int J Hypertens **2010**.
Mascarenhas, M. R., F. Negreiro, et al. (2003). "Age, weight and height: How to predict the low bone mineral density risk in males?" Osteoporosis International **14**: P6.
Massey, H. M. and A. M. Flanagan (1999). "Human osteoclasts derive from CD14-positive monocytes." Br J Haematol **106**(1): 167-170.
Mazess, R. B. and H. Barden (1999). "Bone density of the spine and femur in adult white females." Calcif Tissue Int **65**(2): 91-99.
Mencej-Bedrač, S., J. Preželj, et al. (2011). "TNFRSF11B gene polymorphisms 1181G > C and 245T > G as well as haplotype CT influence bone mineral density in postmenopausal women." Maturitas **69**(3): 263-267.

Milani, L., M. Gupta, et al. (2007). "Allelic imbalance in gene expression as a guide to cis-acting regulatory single nucleotide polymorphisms in cancer cells." Nucleic Acids Research **35**(5): e34.
Monks, S. A., A. Leonardson, et al. (2004). "Genetic inheritance of gene expression in human cell lines." American Journal of Human Genetics **75**(6): 1094-1105.
Morley, M., C. M. Molony, et al. (2004). "Genetic analysis of genome-wide variation in human gene expression." Nature **430**(7001): 743-747.
Morrison, N. A., J. C. Qi, et al. (1994). "Prediction of bone density from vitamin D receptor alleles." Nature **367**(6460): 284-287.
Morrison, S. J. and D. T. Scadden (2014). "The bone marrow niche for haematopoietic stem cells." Nature **505**(7483): 327-334.
Munafo, M. R. and J. Flint (2004). "Meta-analysis of genetic association studies." Trends Genet **20**(9): 439-444.
Munasinghe, R., V. Botea, et al. (2002). "Association between age, height, weight and bone mass index with discordant changes in bone mineral density detected on sequential bone densitometry scans." Journal of Bone and Mineral Research **17**: WG8.
Mylvaganam, S., L. Zhang, et al. (2010). "Hippocampal seizures alter the expression of the pannexin and connexin transcriptome." J Neurochem **112**(1): 92-102.
Nagy, T. R. and A. L. Clair (2000). "Precision and accuracy of dual-energy X-ray absorptiometry for determining in vivo body composition of mice." Obes Res **8**(5): 392-398.
Nguyen, T. V., J. R. Center, et al. (2000). "Osteoporosis in elderly men and women: effects of dietary calcium, physical activity, and body mass index." J Bone Miner Res **15**(2): 322-331.
Nicholson, G. C., M. Malakellis, et al. (2000). "Induction of osteoclasts from CD14-positive human peripheral blood mononuclear cells by receptor activator of nuclear factor kappaB ligand (RANKL)." Clin Sci (Lond) **99**(2): 133-140.
Ohlendorff, S. D., C. L. Tofteng, et al. (2007). "Single nucleotide polymorphisms in the $P2X_7$ gene are associated to fracture risk and to effect estrogen treatment." Pharmacogenet Genomics **17**(7): 555-567.
Ohta, H. (2014). "[Updates on Lifestyle-Related Diseases and Bone Metabolism. Efficacy of selective estrogen receptor modulators (SERMs) in lifestyle-related osteoporosis]." Clin Calcium **24**(11): 1679-1687.
Okeson, G. C. and P. H. Wulbrecht (1998). "The safety of brachial artery puncture for arterial blood sampling." Chest **114**(3): 748-751.
Omrani, G. R. (2006). Bone mineral density in the normal Iranian population: a comparison with American reference data.
Palacios, R., E. Gazave, et al. (2009). "Allele-specific gene expression is widespread across the genome and biological processes." PLoS One **4**(1): e4150.
Panupinthu, N., J. T. Rogers, et al. (2008). "P2X7 receptors on osteoblasts couple to production of lysophosphatidic acid: a signaling axis promoting osteogenesis." J Cell Biol **181**(5): 859-871.
Pastinen, T. and T. J. Hudson (2004). "Cis-acting regulatory variation in the human genome." Science **306**(5696): 647-650.
Paternoster, L., M. Lorentzon, et al. (2013). "Genetic determinants of trabecular and cortical volumetric bone mineral densities and bone microstructure." PLoS Genet **9**(2): e1003247.
Paternoster, L., M. Lorentzon, et al. (2010). "Genome-wide association meta-analysis of cortical bone mineral density unravels allelic heterogeneity at the

RANKL locus and potential pleiotropic effects on bone." Plos Genetics **6**(11): e1001217.
Paternoster, L., C. Ohlsson, et al. (2010). "OPG and RANK polymorphisms are both associated with cortical bone mineral density: findings from a metaanalysis of the Avon longitudinal study of parents and children and gothenburg osteoporosis and obesity determinants cohorts." J Clin Endocrinol Metab **95**(8): 3940-3948.
Pearson, T. A. (2008). "How to interpret a genomewide association study (vol 299, pg 1335, 2008)." Jama-Journal of the American Medical Association **299**(18): 2150-2150.
Pfeiffer, R. M. and M. H. Gail (2003). "Sample size calculations for population- and family-based case-control association studies on marker genotypes." Genet Epidemiol **25**(2): 136-148.
Pocock, N. A., J. A. Eisman, et al. (1987). "Genetic determinants of bone mass in adults. A twin study." J Clin Invest **80**(3): 706-710.
Pouresmaeili, F., J. Jamshidi, et al. (2013). "Association between Vitamin D Receptor Gene BsmI Polymorphism and Bone Mineral Density in A Population of 146 Iranian Women." Cell J **15**(1): 75-82.
Qin, G., Z. Dong, et al. (2013). "Association of vitamin D receptor BsmI gene polymorphism with risk of osteoporosis: a meta-analysis of 41 studies." Mol Biol Rep **40**(1): 497-506.
Raggatt, L. J. and N. C. Partridge (2010). "Cellular and molecular mechanisms of bone remodeling." J Biol Chem **285**(33): 25103-25108.
Raisz, L. G. (2005). "Pathogenesis of osteoporosis: concepts, conflicts, and prospects." J Clin Invest **115**(12): 3318-3325.
Ralston, S. H. (2007). "Genetics of osteoporosis." Proceedings of the Nutrition Society **66**: 158-165.
Ralston, S. H. (2010). "Genetics of osteoporosis." Skeletal Biology and Medicine **1192**: 181-189.
Ralston, S. H. (2010). "Genetics of osteoporosis." Ann N Y Acad Sci **1192**: 181-189.
Ravn, P., G. Cizza, et al. (1999). "Low body mass index is an important risk factor for low bone mass and increased bone loss in early postmenopausal women. Early Postmenopausal Intervention Cohort (EPIC) study group." J Bone Miner Res **14**(9): 1622-1627.
Richards, J. B., F. K. Kavvoura, et al. (2009). "Collaborative Meta-analysis: Associations of 150 Candidate Genes With Osteoporosis and Osteoporotic Fracture." Annals of Internal Medicine **151**(8): 528-U532.
Richards, J. B., F. Rivadeneira, et al. (2008). "Bone mineral density, osteoporosis, and osteoporotic fractures: a genome-wide association study." Lancet **371**(9623): 1505-1512.
Richards, J. B., H. F. Zheng, et al. (2012). "Genetics of osteoporosis from genome-wide association studies: advances and challenges." Nat Rev Genet **13**(8): 576-588.
Risch, N. and J. Teng (1998). "The relative power of family-based and case-control designs for linkage disequilibrium studies of complex human diseases I. DNA pooling." Genome Res **8**(12): 1273-1288.
Rivadeneira, F., F. Kavvoura, et al. (2008). "Large-scale Meta-analysis of Genome-wide Association (GWA) Scans for Osteoporosis Traits: the GEFOS Consortium." Journal of Bone and Mineral Research **23**: S27-S27.

Rivadeneira, F., U. Styrkársdottir, et al. (2009). "Twenty bone-mineral-density loci identified by large-scale meta-analysis of genome-wide association studies." Nat Genet **41**(11): 1199-1206.
Rosenberg, N. A., L. Huang, et al. (2010). "Genome-wide association studies in diverse populations." Nat Rev Genet **11**(5): 356-366.
Ross, F. P. (2003). "Interleukin 7 and estrogen-induced bone loss." Trends Endocrinol Metab **14**(4): 147-149.
Rumney, R. M., N. Wang, et al. (2012). "Purinergic signalling in bone." Front Endocrinol (Lausanne) **3**: 116.
Salamat, M. R., A. H. Salamat, et al. (2013). "Relationship between Weight, Body Mass Index, and Bone Mineral Density in Men Referred for Dual-Energy X-Ray Absorptiometry Scan in Isfahan, Iran." J Osteoporos **2013**: 205963.
Sambrook, P. and C. Cooper (2006). "Osteoporosis." Lancet **367**(9527): 2010-2018.
Sassi, R., H. Sahli, et al. (2014). "Association of LRP5 genotypes with osteoporosis in Tunisian post-menopausal women." BMC Musculoskelet Disord **15**: 144.
Sattui, S. E. and K. G. Saag (2014). "Fracture mortality: associations with epidemiology and osteoporosis treatment." Nat Rev Endocrinol **10**(10): 592-602.
Schadt, E. E., S. A. Monks, et al. (2003). "Genetics of gene expression surveyed in maize, mouse and man." Nature **422**(6929): 297-302.
Scherag, A., H. H. Muller, et al. (2003). "Data adaptive interim modification of sample sizes for candidate-gene association studies." Hum Hered **56**(1-3): 56-62.
Schilling, A. F., W. Linhart, et al. (2004). "Resorbability of bone substitute biomaterials by human osteoclasts." Biomaterials **25**(18): 3963-3972.
Serre, D., S. Gurd, et al. (2008). "Differential allelic expression in the human genome: a robust approach to identify genetic and epigenetic cis-acting mechanisms regulating gene expression." Plos Genetics **4**(2): e1000006.
Shang, M., L. Lin, et al. (2013). "Association of genetic polymorphisms of RANK, RANKL and OPG with bone mineral density in Chinese peri- and postmenopausal women." Clin Biochem **46**(15): 1493-1501.
Simmonds, M. (2011). "Genetic profiling tests in screening for cardiovascular disease." J Med Screen **18**(2): 56-57.
Singh, M., P. Singh, et al. (2013). "Vitamin D receptor (VDR) gene polymorphism influences the risk of osteoporosis in postmenopausal women of Northwest India." Arch Osteoporos **8**(1-2): 147.
Slatkin, M. (2008). "Linkage disequilibrium--understanding the evolutionary past and mapping the medical future." Nat Rev Genet **9**(6): 477-485.
Smith, E. P., J. Boyd, et al. (1994). "Estrogen resistance caused by a mutation in the estrogen-receptor gene in a man." N Engl J Med **331**(16): 1056-1061.
Sonoda, T., J. Takada, et al. (2012). "Interaction between ESRα polymorphisms and environmental factors in osteoporosis." J Orthop Res **30**(10): 1529-1534.
Sorensen, M. G., K. Henriksen, et al. (2007). "Characterization of osteoclasts derived from CD14+ monocytes isolated from peripheral blood." J Bone Miner Metab **25**(1): 36-45.
Soveid, M., A. R. Serati, et al. (2005). "Incidence of hip fracture in Shiraz, Iran." Osteoporos Int **16**(11): 1412-1416.
Stewart, T. L. and S. H. Ralston (2000). "Role of genetic factors in the pathogenesis of osteoporosis." J Endocrinol **166**(2): 235-245.
Stranger, B. E., A. C. Nica, et al. (2007). "Population genomics of human gene expression." Nat Genet **39**(10): 1217-1224.

Stranger, B. E., E. A. Stahl, et al. (2011). "Progress and promise of genome-wide association studies for human complex trait genetics." Genetics **187**(2): 367-383.
Styrkarsdottir, U., B. V. Halldorsson, et al. (2008). "Multiple genetic loci for bone mineral density and fractures." N Engl J Med **358**(22): 2355-2365.
Styrkarsdottir, U., B. V. Halldorsson, et al. (2009). "New sequence variants associated with bone mineral density." Nature Genetics **41**(1): 15-17.
Styrkarsdottir, U., B. V. Halldorsson, et al. (2010). "European bone mineral density loci are also associated with BMD in East-Asian populations." PLoS One **5**(10): e13217.
Sun, C., C. Southard, et al. (2010). "Allelic imbalance (AI) identifies novel tissue-specific cis-regulatory variation for human UGT2B15." Hum Mutat **31**(1): 99-107.
Tang, K., P. Oeth, et al. (2004). "Mining disease susceptibility genes through SNP analyses and expression profiling using MALDI-TOF mass spectrometry." J Proteome Res **3**(2): 218-227.
Teare, M. D. (2011). "Candidate gene association studies." Methods Mol Biol **713**: 105-117.
Teare, M. D., J. Heighway, et al. (2006). "An expectation-maximization algorithm for the analysis of allelic expression imbalance." American Journal of Human Genetics **79**(3): 539-543.
Teare, M. D., S. Pinyakorn, et al. (2011). "Comparing methods for mapping cis acting polymorphisms using allelic expression ratios." PLoS One **6**(12): e28636.
Teitelbaum, S. L. (2000). "Bone resorption by osteoclasts." Science **289**(5484): 1504-1508.
Thakkinstian, A., C. D'Este, et al. (2004). "Meta-analysis of molecular association studies: Vitamin D receptor gene polymorphisms and BMD as a case study." Journal of Bone and Mineral Research **19**(3): 419-428.
Tran, B. N. H., D. Nguyen, et al. (2008). "Association between LRP5 polymorphism and bone mineral density: a Bayesian meta-analysis." Bmc Medical Genetics **9**.
Tremollieres, F. A., J. M. Pouilles, et al. (2010). "Fracture risk prediction using BMD and clinical risk factors in early postmenopausal women: sensitivity of the WHO FRAX tool." J Bone Miner Res **25**(5): 1002-1009.
Tung, J., O. Fedrigo, et al. (2009). "Genomic features that predict allelic imbalance in humans suggest patterns of constraint on gene expression variation." Molecular Biology and Evolution **26**(9): 2047-2059.
Uitterlinden, A. G., S. H. Ralston, et al. (2006). "The association between common vitamin D receptor gene variations and osteoporosis: a participant-level meta-analysis (vol 145, pg 255, 2006)." Annals of Internal Medicine **145**(12): 936-936.
Uman, L. S. (2011). "Systematic reviews and meta-analyses." J Can Acad Child Adolesc Psychiatry **20**(1): 57-59.
Urano, T. and S. Inoue (2014). "Genetics of osteoporosis." Biochem Biophys Res Commun **452**(2): 287-293.
Vaananen, H. K. and T. Laitala-Leinonen (2008). "Osteoclast lineage and function." Arch Biochem Biophys **473**(2): 132-138.
Vaananen, H. K., H. Zhao, et al. (2000). "The cell biology of osteoclast function." Journal of Cell Science **113**(3): 377-381.

van Meurs, J. B., T. A. Trikalinos, et al. (2008). "Large-scale analysis of association between LRP5 and LRP6 variants and osteoporosis." JAMA **299**(11): 1277-1290.
Van Wesenbeeck, L., E. Cleiren, et al. (2003). "Six novel missense mutations in the LDL receptor-related protein 5 (LRP5) gene in different conditions with an increased bone density." Am J Hum Genet **72**(3): 763-771.
VanLiere, J. M. and N. A. Rosenberg (2008). "Mathematical properties of the r2 measure of linkage disequilibrium." Theor Popul Biol **74**(1): 130-137.
Vidal, C., R. Formosa, et al. (2011). "Functional polymorphisms within the TNFRSF11B (osteoprotegerin) gene increase the risk for low bone mineral density." J Mol Endocrinol **47**(3): 327-333.
Vidal, D. O., J. E. de Souza, et al. (2011). "Analysis of allelic differential expression in the human genome using allele-specific serial analysis of gene expression tags." Genome **54**(2): 120-127.
Wade, S. W., C. Strader, et al. (2014). "Estimating prevalence of osteoporosis: examples from industrialized countries." Arch Osteoporos **9**: 182.
Wang, C. L., X. Y. Tang, et al. (2007). "Association of estrogen receptor alpha gene polymorphisms with bone mineral density in Chinese women: a meta-analysis." Osteoporos Int **18**(3): 295-305.
Wang, D., R. Liu, et al. (2013). "Vitamin D receptor Fok I polymorphism is associated with low bone mineral density in postmenopausal women: a meta-analysis focused on populations in Asian countries." Eur J Obstet Gynecol Reprod Biol **169**(2): 380-386.
Weitzmann, M. N. (2013). "The Role of Inflammatory Cytokines, the RANKL/OPG Axis, and the Immunoskeletal Interface in Physiological Bone Turnover and Osteoporosis." Scientifica (Cairo) **2013**: 125705.
Wellcome Trust Case Control, C. (2007). "Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls." Nature **447**(7145): 661-678.
Welter, D., J. MacArthur, et al. (2014). "The NHGRI GWAS Catalog, a curated resource of SNP-trait associations." Nucleic Acids Res **42**(Database issue): D1001-1006.
Wesselius, A., M. J. Bours, et al. (2011). "Role of purinergic receptor polymorphisms in human bone." Front Biosci (Landmark Ed) **16**: 2572-2585.
Wesselius, A., M. J. Bours, et al. (2013). "Association of P2X7 receptor polymorphisms with bone mineral density and osteoporosis risk in a cohort of Dutch fracture patients." Osteoporos Int **24**(4): 1235-1246.
Wilkening, S., B. W. Chen, et al. (2009). "Is there still a need for candidate gene approaches in the era of genome-wide association studies?" Genomics **93**(5): 415-419.
Williams, R. B., E. K. Chan, et al. (2007). "The influence of genetic variation on gene expression." Genome Res **17**(12): 1707-1716.
Wu, X., S. Guo, et al. (2013). "Screening of osteoprotegerin-related feature genes in osteoporosis and functional analysis with DNA microarray." Eur J Med Res **18**: 15.
Xiong, D. H., H. Shen, et al. (2006). "Robust and comprehensive analysis of 20 osteoporosis candidate genes by very high-density single-nucleotide polymorphism screen among 405 white nuclear families identified significant association and gene-gene interaction." J Bone Miner Res **21**(11): 1678-1695.

Xu, X. H., S. S. Dong, et al. (2010). "Molecular genetic studies of gene identification for osteoporosis: the 2009 update." Endocr Rev **31**(4): 447-505.
Yamada, Y., F. Ando, et al. (2003). "Association of polymorphisms of the osteoprotegerin gene with bone mineral density in Japanese women but not men." Mol Genet Metab **80**(3): 344-349.
Yan, H., Z. Dobbie, et al. (2002). "Small changes in expression affect predisposition to tumorigenesis." Nat Genet **30**(1): 25-26.
Yan, H., W. Yuan, et al. (2002). "Allelic variation in human gene expression." Science **297**(5584): 1143.
Yeh, L.-R., C. K. H. Chen, et al. (2004). "Normal bone mineral density in anteroposterior, lateral spine and hip of Chinese men in Taiwan: effect of age change, body weight and height." J Chin Med Assoc **67**(6): 287-295.
Zappitelli, T. and J. E. Aubin (2014). "The "connexin" between bone cells and skeletal functions." J Cell Biochem **115**(10): 1646-1658.
Zhang, L., H. J. Choi, et al. (2014). "Multistage genome-wide association meta-analyses identified two new loci for bone mineral density." Hum Mol Genet **23**(7): 1923-1933.
Zhao, E., H. Xu, et al. (2012). "Bone marrow and the control of immunity." Cell Mol Immunol **9**(1): 11-19.
Zmuda, J. M., Y. T. Sheu, et al. (2005). "Genetic epidemiology of osteoporosis: past, present, and future." Curr Osteoporos Rep **3**(3): 111-115.
Zmuda, J. M., Y. T. Sheu, et al. (2006). "The search for human osteoporosis genes." J Musculoskelet Neuronal Interact **6**(1): 3-15.