

https://dspace.org/

Genetic factors influencing bone health in the black South African population

May, Andrew

2012
http://hdl.handle.net10539/13780

Downloaded from a DSpace repository. https://dspace.org/

Genetic factors influencing bone health in the black South African population

Andrew May

Supervised by Dr. Zané Lombard

Co-supervised by Professor Michèle Ramsay





A dissertation submitted to the Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, in fulfilment of the requirements for the degree of Master of Science (Medicine) in Human Genetics

Declaration

I, Andrew May, declare that this dissertation is my own, unaided work. It is being submitted for the degree of Master of Science (Medicine) in Human Genetics to the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination at any other University.

_____

Andrew May

_____

Date

Abstract

Maintaining suitable bone health is emerging as a serious point of concern worldwide, as the prevalence of skeletal disorders threatens to reach unmanageable proportions. Despite unfavourable environmental factors, black South Africans demonstrate elevated bone mass, especially at the femoral neck, when compared to whites. Genetic factors are thought to mediate this effect, which may have clinical or therapeutic value. Using a candidate gene approach, this study investigated associations of six candidate genes (*ESR1*, *TNFRSF11A*, *TNFRSF11B*, *TNFSF11*, *SOST* and *SPP1*) with bone mineral content amongst pre-pubertal black South African children that formed part of the longitudinal Birth to Twenty cohort. The GoldenGate genotyping assay with VeraCode microbeads was used to genotype 151 black children at 366 polymorphic loci, including 112 previously associated and 254 tagging SNPs. A linear regression approach was implemented to highlight significant associations whilst adjusting for height, weight, sex and bone area. Twenty seven markers (8 previously associated and 19 tag SNPs; $P < 0.05$) were found to link to either femoral neck (18) or lumbar spine (9) BMC. These signals derived from three genes, namely *ESR1* (17), *TNFRSF11B* (9) and *SPP1* (1). One marker (rs2485209) maintained its association with the femoral neck after correction for multiple testing ($P = 0.038$). These results fully support the existence of a strong genetic effect acting at the femoral neck in African ancestry individuals. Tagging SNP signals suggest the presence of a number of population specific variants that require further investigation. Combined, these markers may help to account for increased bone mass amongst black South Africans, when adjusted for covariates.

Acknowledgements

I'd like to extend my sincere gratitude to the following people:

- Firstly, to my supervisor Zané, for her constant support and guidance, given with unfaltering positivity. I am extremely grateful for her hard work and dedication which helped to keep me on track despite various lab work difficulties.
- To Michèle, who always made time to give freely of her unparalleled wealth of experience in the genetics field. Along with Zané, she helped to provide me with opportunities and experiences that few are lucky enough to enjoy.
- To Professors Shane Norris and John Pettifor, for lending their expertise and assisting with funding.
- To Punita, who was always prepared to help, and with much patience and guidance, taught me the techniques I needed to get by in the lab.
- To Jackie, Jaysen, Kenny, Michelle, Liesl, Shatadi, Thandi, and Sanam who all provided so much support, and helped to keep the mood light, when the going got tough.
- To Tyler, a friend I often don't feel worthy of having, who supports me in so many countless ways. His extensive computer knowledge has helped to elevate me from basic user to skilful adept, which has improved my life greatly.
- To the Carnegie Corporation Transformation Program for funding part of this study

The financial assistance of the National Research Foundation (NRF) towards this research is hereby acknowledged. Opinions expressed and conclusions arrived at, are those of the author and are not necessarily to be attributed to the NRF.

Table of Contents

Declaration i
Abstract ii
Acknowledgements iii
Table of Contents iv
List of Figures vii
List of Tables viii
Abbreviations ix
1. Introduction 1
1.1. Bones 2
1.1.1. Bone structure 2
1.1.1.1. Cortical bone 2
1.1.1.2. Trabecular bone 3
1.1.1.3. Osteoclasts 3
1.1.1.4. Osteoblasts 4
1.1.1.5. Osteocytes 4
1.1.2. Bone development and regulation 5
1.1.2.1. Initial bone development 5
1.1.2.2. Childhood 6
1.1.2.3. Puberty and early adulthood 6
1.1.2.4. Late adulthood 7
1.1.3. Mechanism 8
1.2. Bone health 11
1.2.1. Definition and measurement 12
1.2.2. Ethnicity and bone health 13
1.2.3. Bone disease 15
1.2.3.1. Osteoporosis 16
1.2.3.2. Prevalence of OP 17
1.2.3.3. Burden of disease 19
1.2.3.4. Prevention and treatment 20
1.3. Genetics of bone 22
1.3.1. Genetics of complex traits 22
1.3.2. Genetic studies on bone health 24

1.3.2.1. BMD heritability 24
1.3.2.2. Association studies 25
1.3.2.3. Candidate genes 27
1.3.2.4. Pleiotropic, gender and ethnic–specific effects 31
1.3.2.5. Further required bone health research 33
1.4. Bone health amongst South African populations 34
1.4.1. Birth to Twenty 35
1.4.2. Genetic research on black South Africans 37
1.5. Study aim and objectives 39
2. Materials and Methods 41
2.1. Subjects 41
2.2. Anthropometry 42
2.3. Measurement of bone 43
2.4. Sampling of blood and DNA extraction 43
2.5. Candidate gene selection 44
2.6. Tag SNP selection 44
2.7. Genotyping 45
2.8. Quality control 46
2.9. Data analysis 46
3. Results 49
3.1. Descriptive statistics 49
3.2. Genotyping 50
3.3. Association results 52
3.4. *ESR1* and *TNFRSF11B* 56
4. Discussion 63
4.1. Association results 64
4.2. Significant SNP markers 68
4.3. *ESR1* and *TNFRSF11B* 69
4.4. African genetics 72
4.5. Bone health genetics 75
4.6. Study limitations 76
4.7. Future studies 77
5. Conclusion 79
6. References 80

Appendix A – Ethical clearance 88
Appendix B – Equipment and software details 89
Appendix C – DNA concentrations 90
Appendix D – Full SNP panel 100
Appendix E – Full association results 110
Appendix F – Box-and-whisker plots 120

List of Figures

Figure 1.1: Bone structure. 5
Figure 1.2: Changes in bone mass with age. 8
Figure 1.3: Stages of a bone remodelling cycle. 10
Figure 1.4: Molecular regulation of osteoblast and osteoclast function. 11
Figure 1.5: BMD differences across ethnicity. 15
Figure 1.6: Comparison between healthy and osteoporotic bone. 16
Figure 1.7: Odds of osteoporosis for different ethnicities. 18
Figure 1.8: The rising prevalence of osteoporosis in America. 19
Figure 1.9: Differences in osteoporosis onset with varying levels of BMD. 21
Figure 1.10: Structure of the *ESR1* gene and ERα protein 30
Figure 1.11: Structure of the *TNFRSF11B* gene and OPG protein 31
Figure 2.1. Flow diagram indicating the selection of study participants, candidate genes and SNP markers. 42
Figure 2.2: Calling genotypes using BeadStudio 46
Figure 3.1: Comparison of bone measurements between genders, within an ethnic group. 51
Figure 3.2: Box-and-whisker plot of BMC distribution across the three genotype classes of rs2485209 52
Figure 3.3: Allele frequencies across several populations for rs248520 and rs6910500 54
Figure 3.4: LocusZoom plots for femoral neck association results within *ESR1*. 59
Figure 3.5: LocusZoom plots for lumbar spine association results within *ESR1*. 60
Figure 3.6: LocusZoom plots for femoral neck association results within *TNFRSF11B* 61
Figure 3.7: LocusZoom plots for lumbar spine association results within *TNFRSF11B* 62
Figure 4.1: The 17 known transcripts of *ESR1* 70

List of Tables

Table 1.1: Confirmed candidate genes influencing osteoporosis susceptibility..... 27

Table 2.1: Reliability of DXA technician 43

Table 3.1: Descriptive statistics for successfully genotyped participants, with inter-gender comparisons 50

Table 3.2: List of SNP markers generating significant association signals 53

Table 3.3: Allele frequencies for associated SNPs across several population groups, with statistical comparisons to YRI and CEU populations 55

Table 3.4: Comparison between present study results and those from the GEFOS meta-analysis 56

Table 3.5: Functions of significantly associated SNPs 58

Abbreviations

°C	degrees Celsius
µl	microliter
aBMD	areal bone mineral density
ADT	assay design tool
AF1	activation function 1
AF2	activation function 2
AIDS	acquired immune deficiency syndrome
ASO	allele specific oligonucleotide
BMC	bone mineral content
BMD	bone mineral density
BMP	bone morphogenetic protein
BRU	bone remodelling unit
Bt20	Birth to Twenty
CEU	Utah residents with Northern and Western European ancestry
CGAS	candidate gene association study
CHB	Han Chinese from Beijing, China
Chr	chromosome
cm^2	square centimetre
df	degrees of freedom
DNA	deoxyribonucleic acid
DXA	dual X-ray absorptiometry
EDTA	ethylenediaminetetraacetic acid
ERα	oestrogen receptor α
ESR1	oestrogen receptor 1 gene
ESR2	oestrogen receptor 2 gene
FN	femoral neck
g	gram
GEFOS	genetic factors for osteoporosis
GENOMOS	genetic markers for osteoporosis
GWAS	genome-wide association study
GH	growth hormone
HIV	human immunodeficiency virus
HWE	Hardy-Weinberg equilibrium

IGF1	insulin-like growth factor 1
JPT	Japanese from Tokyo, Japan
kg	kilogram
LD	linkage disequilibrium
LS	lumbar spine
LSO	locus specific oligonucleotide
LWK	Luhya from Webuye, Kenya
M-CSF	macrophage colony stimulating factor
MA-C	meta-analysis of candidate gene association studies
MA-G	meta-analysis of genome-wide association studies
MKK	Maasai from Kinyawa, Kenya
mm	millimetre
ml	millilitre
mRNA	messenger ribonucleic acid
MTC	multiple testing correction
n	sample size
ng	nanogram
NHLS	National Health Laboratory Service
NLS	nuclear localisation signal
OBL	osteoblasts
OCL	osteoclasts
OH	hydroxyl group
OMIM	online Mendelian inheritance in man
OP	osteoporosis
OPG	osteoprotegerin
OSC	osteocytes
P	previously associated SNP
P	probability value
PCR	polymerase chain reaction
PBM	peak bone mass
PF	proximal femur
PTH	parathyroid stimulating hormone
QCT	quantitative computed tomography
QTL	quantitative trait loci

RANK	receptor activator of nuclear factor kappa β
RANKL	receptor activator of nuclear factor kappa β ligand
SAS	statistical analysis software
SD	standard deviation
SE	standard error
SNP	single nucleotide polymorphism
SOST	sclerostin gene
SPP1	secreted phosphoprotein 1 gene
T	tag SNP
Taq	*Thermus aquaticus*
TE	Tris/EDTA
TNFRSF11A	tumour necrosis factor receptor superfamily 11A gene
TNFRSF11B	tumour necrosis factor receptor superfamily 11B gene
TNFSF11	tumour necrosis factor superfamily 11 gene
UK	United Kingdom
USA	United States of America
WHO	World Health Organization
vBMD	volumetric bone mineral density
YRI	Yoruba from Ibadan, Nigeria

1. Introduction

Maintaining suitable bone health is rapidly emerging as a serious point of concern worldwide, as the prevalence of skeletal disorders threatens to reach unmanageable proportions (WHO Scientific Group, 2003). Accumulated knowledge on bone development indicates that genetic factors exert a far greater influence than environmental ones, thus current strategies are aimed at elucidating genes and DNA variants that impact on bone regulation in order to find appropriate drug targets. Recent technological breakthroughs are greatly accelerating the pace at which this is achieved, but research focus is still largely aimed at Europeans, limiting our ability to account for persistent interethnic differences in bone physiology (Duncan and Brown, 2010). Individuals of African ancestry, for example, demonstrate superior bone health, even against environmental odds, suggesting the presence of a set of genetic variants that strongly promote bone maintenance (Section 1.4.2). Identifying these variants is likely to lead to advances in our understanding of bone formation and development, along with a wealth of subsequent opportunities for leveraging this information in drug design.

Similarly understudied are the genetic effectors mediating bone development during childhood, a period considered crucial for the determination of bone health during adult life (Duren et al., 2007). Genetic regulation of bone is suspected to differ substantially across different phases of life and thus knowledge of the processes and pathways that are most important during the early years may grant insight into preventative measures that will delay or inhibit the onset of skeletal problems (Duren et al., 2007). To better appreciate the value of these neglected areas of research, it is necessary to have a solid understanding of bone structure and physiology, as well as how these are affected by disease.

1.1. Bones

Bones are highly dynamic and metabolically active organs that together form the skeletons of humans and other vertebrates. Comprised of a specialised form of rigid connective tissue, bones primarily act to support and shape the body (Karasik, 2008; Moore, 1992). Their durable nature serves to protect the more delicate internal organs, whilst also providing a solid, yet flexible support for muscles. Aside from these more mechanical functions, bones undertake a diverse assortment of additional responsibilities including haematopoiesis, mineral and fatty acid storage, and alleviation of blood acidosis amongst others (Karasik, 2008). Furthermore, bones need to continuously adapt to changing conditions, and thus model (change the total amount of bone) and remodel (change bone size and shape without changing the total amount of bone) their shape with a precision that is unmatched by other regenerative organs such as the liver and skin (Bartl and Frisch, 2004).

1.1.1. Bone structure

Individual bones come in a wide variety of specific shapes and sizes but are linked together by a common group of structural features. Firstly, all bones have a similar architecture that is primarily comprised of two distinct types of bone, namely cortical and trabecular (**Figure 1.1**).

1.1.1.1. Cortical bone

The outer shell of all bones is formed by a crystalline layer of cortical bone that provides the necessary durability for bones to support and protect the body (de Villiers, 2009). Cortical bone is densely packed and highly calcified, having both a low surface to volume ratio and low metabolic rate. Bones forming the appendicular skeleton (limb bones) are typified by a large component of cortical bone (Compston, 2001).

1.1.1.2. Trabecular bone

Internal to the outer cortical layer lies a central mass of spongy (cancellous or trabecular) bone consisting of interconnecting plates and rods. This configuration creates a porous structure that has a high surface area, and allows bone to be light-weight without significantly penalising its strength (de Villiers, 2009). Bones of the axial skeleton (cranium, vertebral column, thorax and pelvis) are rich in trabecular bone (Bartl and Frisch, 2004).

Secondly, the molecular sub-structure of all bone (osseous) tissue is largely a composite of organic and inorganic extracellular substances. These include collagen (~40%), hydroxyapatite (~45%) and water (~15%) (Bauer and Link, 2009). Collagen provides flexibility, whilst crystals of combined calcium and phosphorous minerals (hydroxyapatite) enhance rigidity. Varying levels of these constituents can thus alter the strength and durability of bones (Pietschmann et al., 2009).

A final common feature between bones is their cellular composition. All bone surfaces, including the trabecular and inner (endosteum) and outer (periosteum) cortical surfaces are covered by osteoblast and osteoclast cells (de Villiers, 2009; Guggenbuhl, 2009; Lerner, 2006), although these are often in precursory forms. Embedded within the cortical bone matrix are osteocyte cells.

1.1.1.3. Osteoclasts

Osteoclast (OCL) cells are giant, multinucleated cells, derived from monocytes (Buckwalter et al., 1995; Compston, 2001), which can resorb aged or damaged bone. The portion of their cell membrane that faces the bone surface is ruffled, consisting of numerous folds. In between these folds, a mix of enzymes is secreted that degrade the bone matrix, creating a resorption cavity. Parts of the matrix that

aren't completely dissolved are instead phagocytosed and metabolised within the cytoplasm of the osteoclast (Bartl and Frisch, 2004).

1.1.1.4. Osteoblasts

Osteoblast (OBL) cells are responsible for the formation of new bone necessary for either growth or the reparation of resorption cavities. This is achieved through the secretion of osteoid, which is an appropriate mix of collagen type I and various proteins such as bone morphogenetic protein (BMP), osteocalcin and osteonectin (Bartl and Frisch, 2004). Pre-osteoblast cells are derived from mesenchymal stem cells (Buckwalter et al., 1995; Compston, 2001) which line bone surfaces in anticipation of the necessary maturation signals. Approximately one in ten osteoblasts becomes trapped within the bone matrix, and these cells subsequently develop into osteocytes (Bartl and Frisch, 2004).

1.1.1.5. Osteocytes

Osteocytes (OSC) are the most numerous of the bone cells and they occupy spaces known as lacunae within the bone matrix. Long, thin cellular processes stretch along canaliculi and thus join neighbouring osteoclasts as well as precursor osteoblasts lining the bone surface. These processes facilitate a complex network of cellular communication, suggesting that osteocytes act as mechanosensory cells that relay information about the status of the surrounding bone (Bartl and Frisch, 2004).

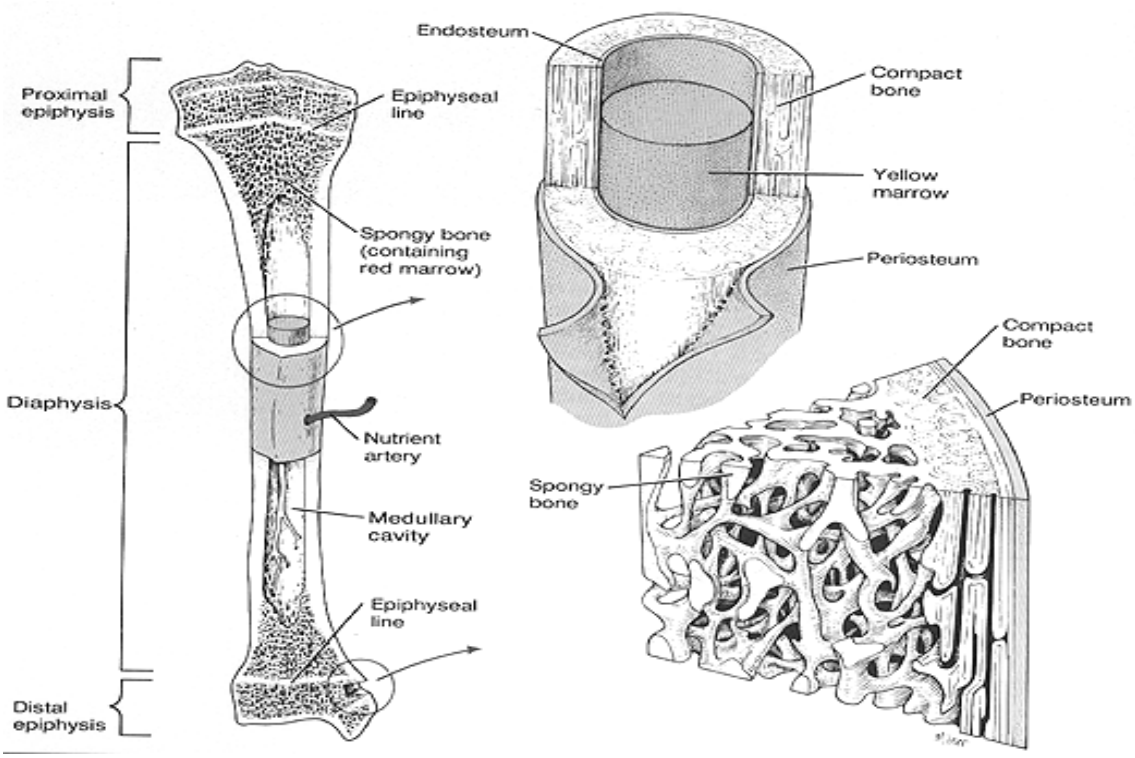


Figure 1.1: Bone structure. Regardless of shape, bones share a common structure that consists of an outer casing of compact, cortical bone which surrounds the inner trabecular (spongy) bone. The outer bone layer is known as the periosteum whilst the inner is referred to as the endosteum (Spence, 1990).

1.1.2. Bone development and regulation

As a result of their numerous functions, bones are constantly changing to meet the demands of the body. Besides the modelling during early life and remodelling during puberty, minor alterations in bone structure occur daily, as reparations are conducted and mineral homeostasis is maintained (de Villiers, 2009; Kim et al., 2007). The three cell types detailed above (**Sections 1.1.1.3 – 1.1.1.5**) work in concert to manage and respond to these changes, which are described in a chronological order below (**Figure 1.2**).

1.1.2.1. Initial bone development

Skeletal development, which begins *in utero*, is initialised by condensations of mesenchymal cells that secrete a matrix of components characteristic of

cartilaginous tissue, in a process termed chondrification. The matrix forms in the centre of these cellular aggregations, slowly causing individual cells to separate as it spreads to outline the required, genetically determined, bone shape (Bourne, 1972). The precursor cartilage model is then steadily replaced by bone, due to the actions of osteoblast cells which derive from both mesodermal and neural crest progenitor cells (Gafni and Baron, 2007). Complete replacement of cartilage is only achieved at skeletal maturity during early adulthood (Bartl and Frisch, 2004).

1.1.2.2. Childhood

Bone growth and development are strongly promoted throughout childhood, but a subtle difference between sexes begins to emerge (Martin and Seeman, 2008). A sustained level of growth hormone (GH) and insulin-like growth factor 1 (IGF1) maintains osteogenic stimulation throughout this period (Guggenbuhl, 2009), with the focus being on bone lengthening. Bone lengths are comparable between young boys and girls, but male bones are noticeably thicker, due to differences in sex hormone exposure *in utero*. Differences in length only become apparent nearer to the pubertal stage due to the different timing in puberty onset between boys and girls. Boys enter puberty two years later, thus enjoying an extended period of bone lengthening, especially of the appendicular skeleton (Guggenbuhl, 2009).

1.1.2.3. Puberty and early adulthood

During puberty, bone remodelling reaches its highest level and gender-based differences are accentuated. Both genders experience a characteristic growth spurt, brought about by the production of sex hormones and consequent increased levels of GH and IGF1 (Guggenbuhl, 2009; Hasselstrom et al., 2006), but the remodelling process is markedly different between male and female adolescents. In males, bone width and cortical thickness are emphasised in conjunction with

growth, whereas in females, no such thickening or widening of bone occurs. Conversely, some bones become narrower instead, leaving females with smaller, thinner bones (Guggenbuhl, 2009). Bone growth slows and ultimately ceases at around 20 years of age in females, but males can continue to make small gains in bone mass until 30, further magnifying gender differences (Rizzoli et al., 2010). Regardless, the final bone mass achieved upon skeletal maturity in either sex is referred to as the peak bone mass (PBM) (Bauer and Link, 2009; Rizzoli et al., 2010), and this state is thought to be maintained throughout early adulthood (Compston, 2009), although growing evidence suggests otherwise (Guggenbuhl, 2009).

1.1.2.4. Late adulthood

Most likely starting soon after the third decade of life (Nanes and Kallen, 2009), approximately 1% of bone material is lost with each passing year. The loss appears to be genetically prompted, and at a cellular level, is the consequence of increased osteoclastic activity coupled with incomplete refilling of resorption cavities (de Villiers, 2009). In men, the loss mostly occurs at a gradual, steady pace, but in women, bone loss accelerates rapidly in the 5-10 years following the menopause before steadying to again match the pace in men (Compston, 2009; Guggenbuhl, 2009; Nanes and Kallen, 2009; Pietschmann et al., 2009). Post-menopausal losses, which total almost half of all cancellous and a third of all cortical bone mass (Downey and Siegel, 2006), are the result of the substantive decline in levels of oestrogen; an important hormone in bone formation pathways (**Section 1.1.3**) (Compston, 2009; Nanes and Kallen, 2009). Indeed, bone loss in elderly men is also attributed to declining levels of oestrogen, further highlighting the importance of the hormone in bone health (Guggenbuhl, 2009).

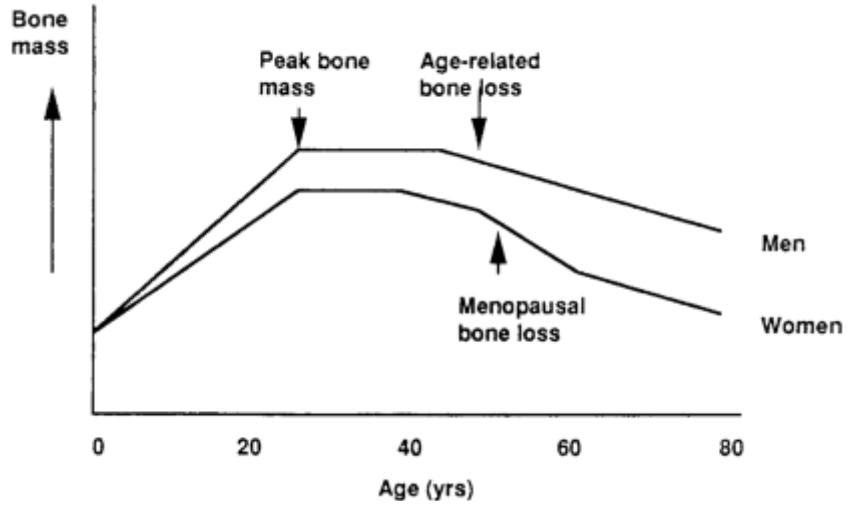


Figure 1.2: Changes in bone mass with age. Bone mass is rapidly gained for the first two decades of life, before reaching a peak at approximately 20 years of age. This peak bone mass is thought to be maintained for several years, although some evidence suggests otherwise. Regardless, bone mass starts to decrease with age during late adulthood, with a significant reduction during the period of the menopause. Males have noticeably increased bone mass in comparison to women, which persists throughout life (Compston, 2001).

Throughout the entirety of life, remodelling is conducted to facilitate the regular maintenance of bones. Daily activities are often damaging enough to cause microfractures that undermine the structural integrity of bone (Gallagher and Sai, 2010). Furthermore, minerals safely stored in the bone matrix are regularly needed to rectify homeostatic imbalances, especially with respect to calcium (de Villiers, 2009; Rizzoli et al., 2010). In both instances osteoclasts are enlisted to dissolve the necessary bone. Soon thereafter, provided that the cellular machinery is at normal functioning capacity, the resulting cavity is precisely filled by osteoid, thus maintaining the proper bone structure (Pederson et al., 2008).

1.1.3. Mechanism

The frequent modifications made to bone are dependent on tightly regulated cellular and molecular mechanisms (Kim et al., 2007; Pederson et al., 2008). Osteoblast and osteoclast cells do not operate independently, but rather form multicellular bone remodelling units (BRUs) in which they counter-regulate each

other (Bartl and Frisch, 2004; Gallagher and Sai, 2010; Livshits, 2006). When and where necessary, BRUs are signalled to begin a remodelling cycle (**Figure 1.3**), which lasts for between four to six months and is regulated primarily by the RANK/RANKL/OPG pathway (**Figure 1.4**) (Pietschmann et al., 2009).

The remodelling cycle can be separated into six stages, namely:

1. Quiescence
2. Resorption of bone (2-4 weeks)
3. Reversal of the resorption signalling pathway
4. Early bone formation ┐
5. Late bone formation ┘ (2-4 months)
6. Quiescence

Prior to active signalling, quiescent osteoblast progenitor cells line the bone surface (stage 1). In order to activate the maturation process of osteoclasts so that bone resorption can begin, two cytokine signals are required: macrophage colony stimulating factor (M-CSF) and receptor activator of nuclear factor kappa β ligand (RANKL) (Livshits, 2006). M-CSF helps with cell proliferation and survival, but it is RANKL that crucially initiates the maturation process. RANKL is secreted by both T-cells and osteoblasts, and it binds receptor activator of nuclear factor kappa β (RANK), expressed on the surface of osteoclast cells. Mature osteoclasts then begin bone resorption (stage 2) until signalling for osteoclast activity is reversed through the secretion of osteoprotegerin (OPG) by osteoblasts and local mesenchymal cells (stage 3). OPG functions as a decoy receptor, binding, and thus intercepting RANKL before it can bind RANK (Bartl and Frisch, 2004). Osteoblast precursors are then signalled to mature through a complex mix of hormones, growth factors and cytokines (Boyce and Xing, 2007; de Villiers, 2009). Of particular importance in the signalling process are oestrogen, vitamin D and the Wnt/β-

catenin pathway (Compston, 2009; Gafni and Baron, 2007; Livshits, 2006). Mature osteoblasts then synthesise sufficient osteoid to fill the resorption cavity (stage 4), into which calcium and phosphorous minerals are inserted (stage 5). The multicellular unit then returns to the quiescent phase until further signalling is received (stage 6). In contrast to the remodelling process, bone modelling during childhood is achieved by uncoupling osteoblasts and osteoclasts, such that resorption and deposition can occur independently to achieve the required changes to bone shape and size (Gafni and Baron, 2007; Martin and Seeman, 2008).

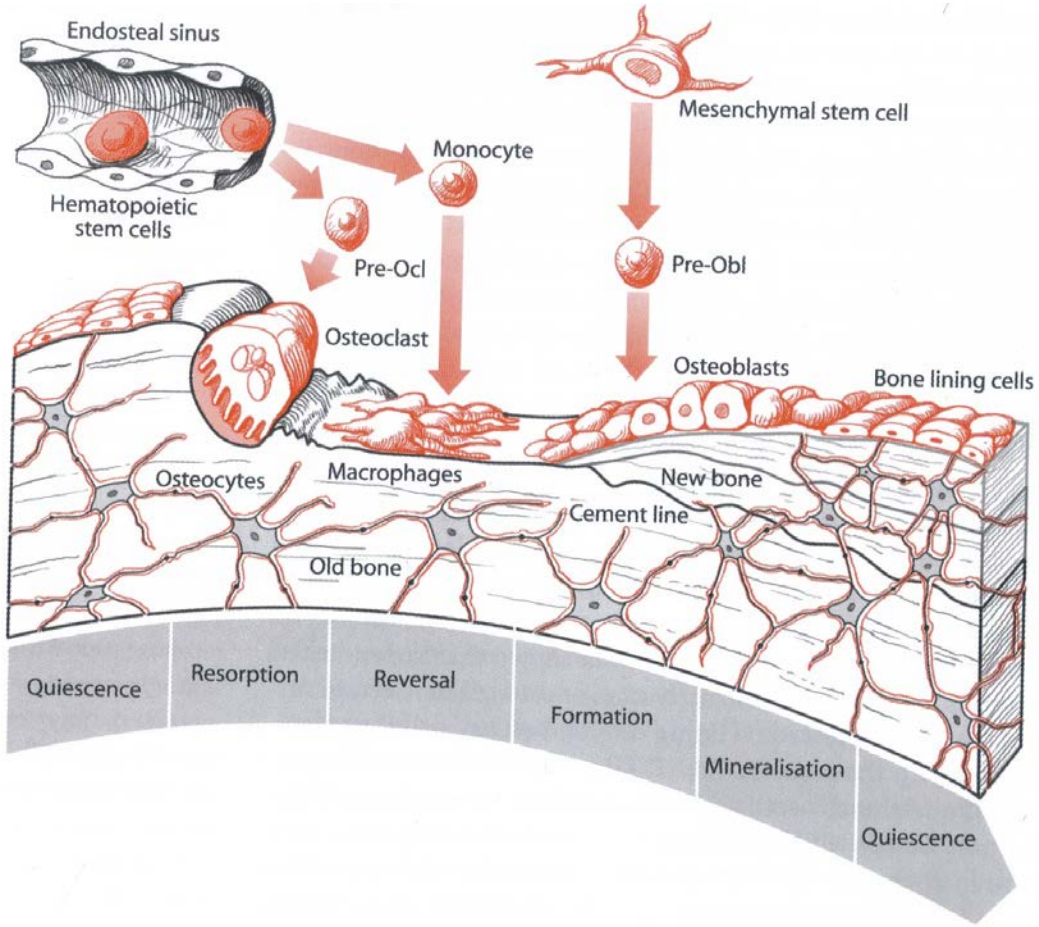


Figure 1.3: Stages of a bone remodelling cycle. A standard bone remodelling cycle consists of six stages. During quiescence, precursor osteoblasts line the bone surface and are known as bone lining cells. When necessary, pre-osteoclast cells attach to the surface and are triggered to mature in order to enable bone resorption. After sufficient resorption, signalling is reversed to prohibit further bone loss and instead promote osteoblast-mediated bone formation. Finally, newly deposited bone is mineralised over several months and a state of quiescence is once again maintained (Bartl and Frisch, 2004).

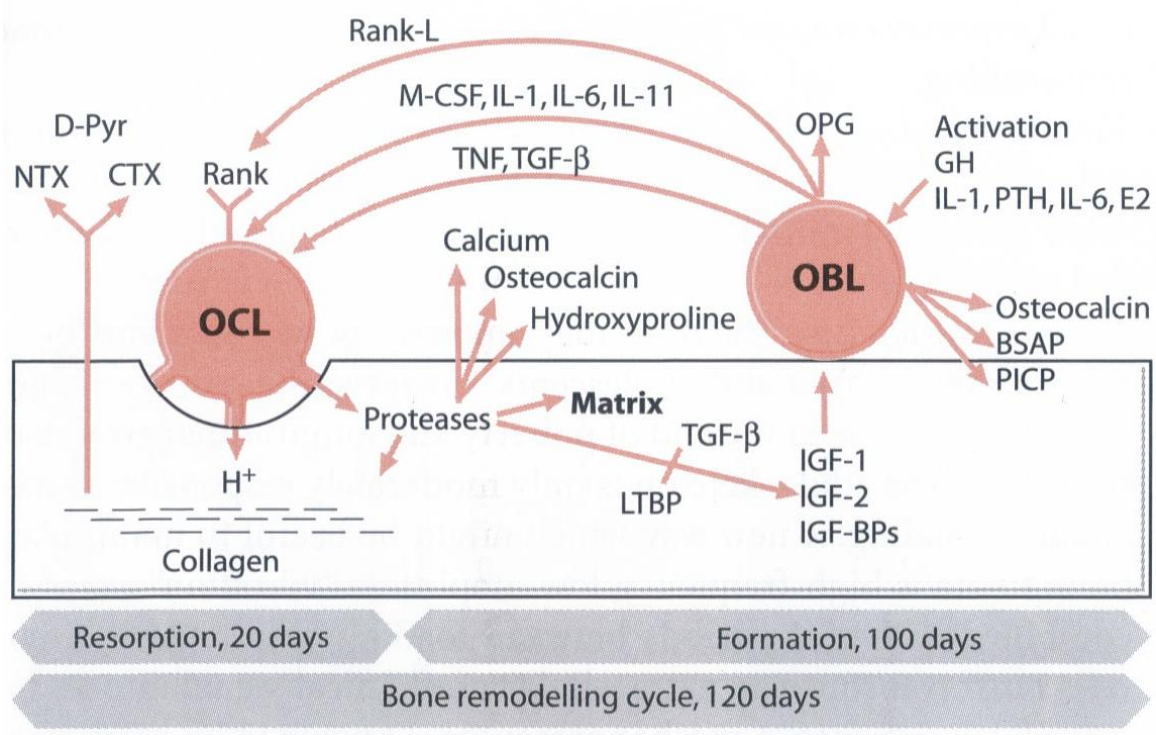


Figure 1.4: Molecular regulation of osteoblast and osteoclast function. OBL and OCL cells counter-regulate each other through a complex mixture of molecular signals in order to balance bone loss and formation. Most important are the RANKL molecule, its cognate receptor RANK and the decoy OPG. RANKL is secreted by OBL cells to encourage OCL maturation and consequent bone resorption. To reverse this signal, OBL cells secrete OPG instead which intercepts and binds RANKL before it can bind RANK. Bone formation then occurs to refill resorption cavities (Bartl and Frisch, 2004).

1.2. Bone health

Bone functions are central to the overall growth and well-being of humans. Unsurprisingly, bones can thus be implicated in, or affected by, almost every known medical condition. Diseases ranging from Crohn's disease to sarcoidosis and thalassemia negatively impinge on bone health whilst bones themselves are prone to a vast list of disorders including osteoarthritis, osteochondroma and osteomalacia (Bartl and Frisch, 2004). Ramifications of compromised bone health can be far-reaching, as is evident by the overall toll exacted by skeletal disorders, which has attained heights that have drawn the attention of the World Health Organisation (WHO). Indeed, the 2000-2010 period was declared as the "Bone and Joint Decade", in an effort to bolster research in this field (Bartl and Frisch, 2004;

Woolf, 2000). Maintaining and monitoring bone health is a thus a critically important part of the health care industry.

1.2.1. Definition and measurement

Bone health is a function of two main components, namely bone quality and bone mineral density (BMD) (Bauer and Link, 2009; Rizzoli et al., 2010). Bone quality is a broad-based term referring to a list of finer bone subtleties such as thickness of cortical bone, architecture of trabecular bone, rate of bone turnover, accumulation of microfractures and bone geometry, amongst others. BMD, on the other hand, is more simply defined as the amount of bone per unit area or volume (Kanis et al., 2008). The mineral density of any bone accounts for 50-80% of its strength, and is thus a more significant indicator of bone health than bone quality is. A particularly attractive feature of BMD is that it can be measured accurately and, unlike other facets of bone health, non-invasively (Bauer and Link, 2009; Gallagher and Sai, 2010). Consequently, BMD measurements have become an integral part of bone health monitoring in a clinical setting (Kanis et al., 2008).

The widely preferred technique for measuring BMD is dual X-ray absorptiometry (DXA), due to its ease of use (Rizzoli et al., 2010). The radiation generated by a pair of differentially intense X-ray beams targeted at a specific skeletal site is accurately interpreted as a measurement of areal bone mineral density (aBMD), based on the calculation of bone mineral content over area, in g/cm^2 (Rizzoli et al., 2010). Since BMD is not constant throughout the skeleton, several measurements are made at different skeletal sites to obtain a more comprehensive outlook. The DXA technique has been a cornerstone since 1988, and thus healthy ranges of BMD are well known (Bartl and Frisch, 2004). Average values per age as well as average values at

the point of skeletal maturity provide a comparative reference that can be used to gauge the bone health of individual patients (Kanis et al., 2008).

1.2.2. Ethnicity and bone health

A striking observation emerges when assessing bone health across different population groups, in that it varies significantly along ethnic lines (**Figure 1.5**). The interethnic comparison most extensively studied is that between African Americans and US whites, which has repeatedly shown greater aBMD and trabecular thickness in the former (Han et al., 1996; Nelson et al., 2008). On average, an aBMD 3-11% higher in African Americans has been reported (Hill et al., 2008), although some studies have demonstrated even greater density differences when compared to whites. For example, when measuring true volumetric density (vBMD), using quantitative computed tomography (QCT), the bones of African American women were found to be 40% denser (Kleerekoper et al., 1994). Conversely, research into Asian populations has revealed some of the lowest BMD values. In a longitudinal study of 423 Asian, Hispanic, white and black males and females, Asian participants ranked last for both femoral neck and whole body BMD (Bachrach et al., 1999). Similarly, Tobias and colleagues (1994) noted a lower femoral neck and lumbar spine BMD when comparing Asian women to both Caucasians and Afro-Caribbeans. Admixed populations, such as Hispanics, tend to possess BMD values that are not significantly different from whites, with the exception of African Americans (Bachrach et al., 1999; Nelson et al., 2008; Wang et al., 1997).

The interpretation of interethnic comparisons has, however, been strongly cautioned since covariates such as height and weight should be accurately corrected for, although this is infrequently done. Failing to account for these covariates often results in either over- or underestimations of BMD which may then

lead to erroneous conclusions (Vidulich et al., 2011). A study on white and Chinese women found that bone mass was greater amongst the Chinese woman when height and weight were factored into the comparison, thus defying the popular belief that Asians possess the lowest BMD values (Randolph et al., 2003). Indeed, most studies tend to report little to no difference between BMD measurements for whites, Hispanic and Asian adults after controlling for covariates (Nelson et al., 2008). Significant differences are still seen between African Americans and whites, but these are of a lesser degree following body size adjustments (Nelson et al., 2008). Once again, a level of caution must still be exercised even when considering adjusted BMD values, as consensus on the best method for correcting for covariates has not yet been reached (Vidulich et al., 2011). Consequently, some studies still report findings that adhere to the traditional idea that African Americans possess the greatest BMD followed by whites then Asians, as was seen in a cohort study on 197 848 postmenopausal women whose density measurements were adjusted for weight and other covariates (Barrett-Connor et al., 2005). These discrepancies in interethnic comparisons are most likely attributed to differences in study design, but also reflect the inherent difficulty of trying to sub-divide humans into discrete sub-populations based on socially defined concepts of race or ethnicity, which are not necessarily supported by genetic theory (Nelson et al., 2008). Ultimately, although it is difficult to properly quantify, a difference in bone health between currently defined ethnicities does exist and this is more clearly seen when considering the prevalence of bone disease, specifically osteoporosis, as well.

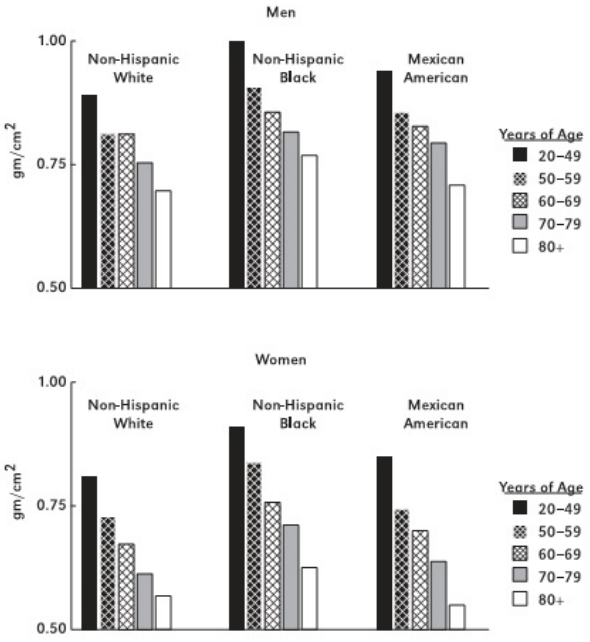


Figure 1.5: BMD differences across ethnicity. African ancestry individuals (labelled as Non-Hispanic Black) are repeatedly observed to possess the highest bone measurements compared to other ethnicities, regardless of sex or age (Looker et al., 1998).

1.2.3. Bone disease

Bone diseases are numerous and incorporate a wide variety of problems from, amongst others, bone cysts, hypomineralization, ectopic growth and even vanishing bones. Underscoring a notable majority of these diseases, as well as any other condition that indirectly hampers bone tissue, is the concern of fracture. Defined as an acute discontinuity in bone, fractures can pose serious threats to patient health and well-being (Bartl and Frisch, 2004). The basis for increased fracture risk lies in the loss of bone mineral mass that facilitates the development of structural weaknesses. In addition to occurring during numerous diseases, ailing bone structure is an inevitable part of late adulthood due to a steady decline of bone mass (**Section 1.1.2.4**). Beyond a certain threshold, low bone mass carries a risk of fracture at even mild levels of trauma; a state which is classified as osteoporosis (Xiong et al., 2009). Since the risk for osteoporosis lurks whenever

bone health is jeopardized, regardless of the exact cause, it is the only bone disease considered further.

1.2.3.1. Osteoporosis

Osteoporosis (OP; OMIM 166710) is defined as a systemic skeletal disease characterised by low bone mass, microarchitectural deterioration and reduced bone strength, all of which enhance the probability of low-trauma fractures (**Figure 1.6**)(Kanis et al., 2008; Nanes and Kallen, 2009). An estimated 200 million individuals are currently suffering from OP, making it the most common bone metabolic disorder (Xiong et al., 2009). OP is regarded as a "silent" disease, because until it manifests in a clinically significant fracture, it fails to present with any symptoms (de Villiers, 2009; Nanes and Kallen, 2009). Indeed, osteoporotic bones often appear entirely normal, lacking any distinguishable features with respect to the bone exterior. Internally, however, differences becomes clearly visible in the form of reduced cortical thickness, imbalanced cortical and trabecular bone proportions and a marked deterioration of the trabecular framework (Bartl and Frisch, 2004).

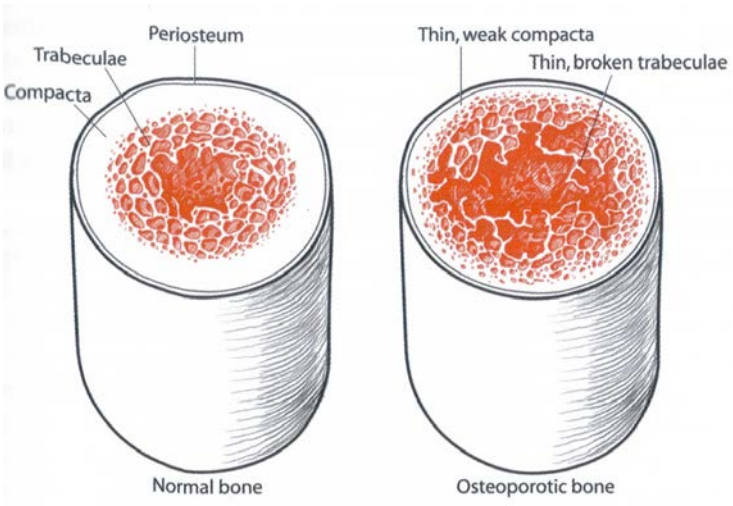


Figure 1.6: Comparison between healthy and osteoporotic bone. Normal, healthy bone consists of a thick cortical layer, with a strong trabecular framework. In those that suffer from osteoporosis, bone cortical layers are thin and trabecular frameworks are weak and degraded (Bartl and Frisch, 2004).

The loss of bone structural integrity has minimal effect on quality of life, and thus the true burden of osteoporosis is made manifest through fractures. OP sufferers are hallmarked by fragility, or low trauma fractures, such as falling from standing height, with each doubling the chance for an additional fracture, thus creating a continually escalating level of risk (de Villiers, 2009; Nanes and Kallen, 2009). Trabecular-rich bones are those most likely to break since they possess the largest surface area and are thus more prone to the effects of imbalanced bone turnover. Consequently, the three most common osteoporotic fractures are vertebral, hip and Colles' (distal radius) fractures (Li et al., 2010).

While fracture is the true phenotype of osteoporosis, BMD has been adopted as the surrogate phenotype, given both the ease of obtaining BMD measurements and the close relationship between fracture risk and bone density (Duncan and Brown, 2008; Kanis et al., 2008). Calculations have shown that every standard deviation below mean values for BMD increases the risk of any fracture nearly two-fold (Kanis et al., 2008), which is appropriate given the correlation of BMD with other bone parameters that determine fracture risk such as structure and strength (Pietschmann et al., 2009). Current diagnostic guidelines for OP are thus based on the BMD metric, which is also used for clinical and research based investigations of the disorder (Kanis et al., 2008; Richards et al., 2009).

1.2.3.2. Prevalence of OP

Due to the accumulating loss of bone per year after 30 years of age (**Section 1.1.2.4**), the prevalence of OP increases with age, thus representing a significant threat to the elderly population. Whilst OP affects both men and women of all backgrounds, prevalence levels between gender and ethnicity are noticeably different. OP is observed far more commonly in women than in men. Indeed, 1 in 2

women, but only 1 in 4 men over the age of 50 will experience an osteoporotic fracture (Gallagher and Sai, 2010). This is thought to be ascribable to events during female skeletal formation and growth (**Section 1.1.2**), as well as the dramatic effect of the menopause (Guggenbuhl, 2009). Such is the number of at-risk individuals that more women will experience an osteoporotic fracture than will a heart-attack, stroke and all female cancers combined (Bartl and Frisch, 2004).

When comparing OP prevalence across ethnic boundaries, Asians and Caucasians are observed to be far more prone to the disorder, whereas individuals of African ancestry are less commonly affected (**Figure 1.7**) (Miller et al., 2005; Paradowska and Lacki, 2007). Interestingly, OP and fracture prevalence rates inversely mirror average BMD measurements amongst ethnic groups (Nelson et al., 2008). Where average BMD measurements are high, such as in African individuals, OP prevalence is low, and vice-versa, further highlighting both the strong relationship between BMD and osteoporosis as well as the inherent interethnic differences in bone health previously outlined (**Section 1.2.2**).

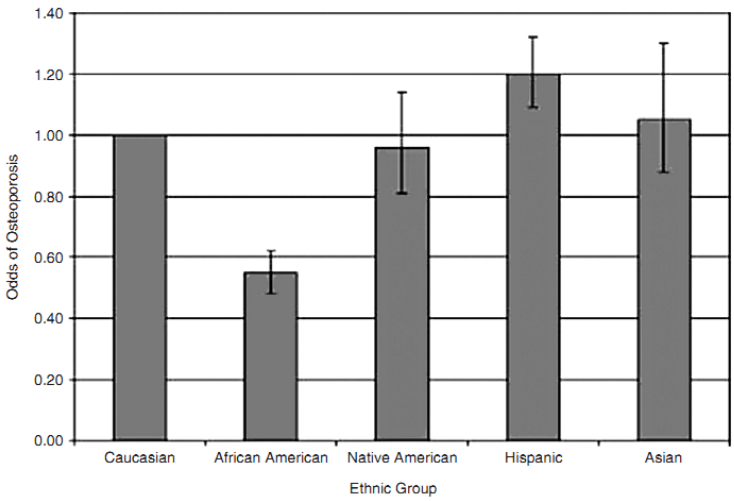


Figure 1.7: Odds of osteoporosis for different ethnicities. Based on official diagnostic criteria, African ancestry individuals have a very low risk of osteoporosis, which inversely correlates to their high average BMD values (Nelson et al., 2008)

Regardless of a differing prevalence, failing the development of an appropriate intervention strategy, it is estimated that OP will increase by approximately three-fold before 2035, mainly due to demographic changes in the number of elderly (**Figure 1.8**) (Gallagher and Sai, 2010).

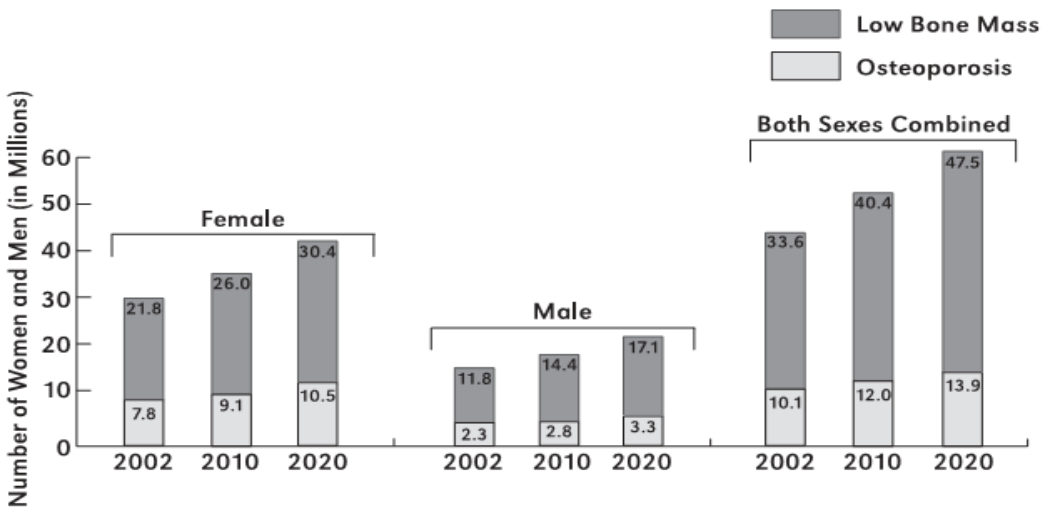


Figure 1.8: The rising prevalence of osteoporosis in America. As life expectancy increases in both developed and developing countries, the burden of osteoporosis is also projected to increase substantially, representing a serious threat (National Osteoporosis Foundation, 2002).

1.2.3.3. Burden of disease

Having such a high prevalence, OP places a massive burden on the health care industry. Clinically significant osteoporotic fractures, especially those of the hip, are a substantive source of morbidity and mortality (Compston, 2009; Karasik et al., 2010; Tanriover et al., 2010), even if surgically repaired. Approximately 36% of male and 21% of female hip fracture patients die within the first year following the fracture, and excess mortality persists for an additional four years thereafter (Guggenbuhl, 2009). Those that survive an osteoporotic fracture must often settle for a reduced quality of life due to poor mobility and kyphosis of the spine, *inter alia*. Furthermore, these physical complications frequently manifest into psychological and emotional problems including depression and lowered self-

esteem (Holroyd et al., 2008). The financial implications, as estimated in the USA in 2005, suggested that the totality of osteoporotic fracture cost exceeded the combined costs of breast, cervical and ovarian cancer (Nanes and Kallen, 2009). Worldwide, a conservative figure of US $131.5 billion was reportedly spent treating osteoporotic fractures in 1997 alone (Holroyd et al., 2008).

1.2.3.4. Prevention and treatment

Although several treatment options are available to ameliorate osteoporosis, including both anti-resorptive and anabolic agents, prevention is thought to remain a more efficacious approach (Fewtrell et al., 2009; Kalkwarf et al., 2007; Rizzoli et al., 2010). Following the attainment of peak bone mass, few measures can be taken to maintain bone health in the face of gradual losses. Accruing an appreciable PBM prior to skeletal maturity is thus widely believed to be the most effective positive intervention in reducing OP risk (Bianchi, 2007; Paradowska and Lacki, 2007; Rizzoli et al., 2010). Hitherto, the genetically programmed progression of bone formation and loss has been described, but environmental factors have been well documented to significantly modify this progression (Kalkwarf et al., 2007). According to Wolff's Law, bone will automatically adapt to varying skeletal loads, thus frequent physical activity and weight-bearing exercise can stimulate osteoblasts to further strengthen existing osseous tissue at affected bone sites, whereas an excessive lack of physical pressure can cause atrophy (Downey and Siegel, 2006; Looker et al., 2009; Wolff, 1986). This phenomenon is a common concern for astronauts as well as patients hospitalised for prolonged durations (Looker et al., 2009). Other well-known environmental factors include the amounts of dietary calcium and active metabolites of vitamin D. Calcium and vitamin D crucially assist in the proper formation of bone, thus plentiful amounts help to maintain bone structure (Bauer and Link, 2009). Conversely, some environmental

factors can enhance the rate of bone loss. Strong risk factors to this effect include smoking and alcohol consumption (de Villiers, 2009). Appropriate management of environmental factors, therefore, may allow for the attainment of PBM levels that guard against the development of osteoporosis, with every 10% increase above the young adult average reportedly delaying the onset of the disease by 13 years (**Figure 1.9**)(Rizzoli et al., 2010).

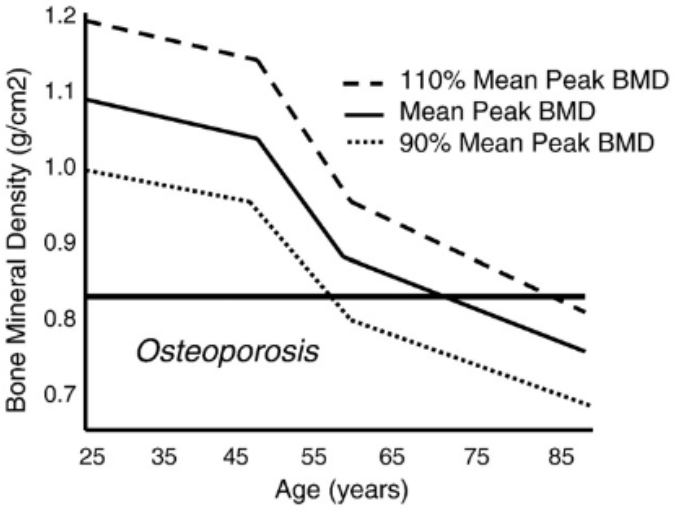


Figure 1.9: Differences in osteoporosis onset with varying levels of BMD. An individual that attains a PBM similar to the population mean is expected to experience disease onset near 70 years of age. Differences as small as 10% in mean PBM can alter this onset by over 10 years (Rizzoli et al., 2010).

Although environmental factors appear to powerfully modify bone health and subsequent risk for osteoporosis, growing evidence suggests that they fail to outweigh genetic influences. Some studies have demonstrated that positive interventions taken to raise BMD during childhood are not indefinitely reflected on bone health, and thus may in fact only be transient (Gafni and Baron, 2007). In tandem with the ability to accurately predict adult BMD in one year old infants (Guggenbuhl, 2009), this has raised the idea of a possible genetic “set-point” for bone health. Such a notion runs directly against the widely accepted importance of preventative measures outlined above, although further investigation into this

matter is required (Rizzoli et al., 2010). Regardless, it is becoming increasingly clear that preventative measures alone are unlikely to be sufficient in combating the rising prevalence of osteoporosis and therefore genetic knowledge must be leveraged through drug design if better progress is to be made (de Villiers, 2009). Bone phenotypes such as quality and BMD, however, are complex traits that arise from the interplay of numerous genetic factors, making it difficult to fully comprehend the molecular regulation of bone health.

1.3. Genetics of bone

1.3.1. Genetics of complex traits

The term “complex trait” is used to describe phenotypes with a polygenic mode of inheritance, combined with a susceptibility to environmental influences (Carlborg and Haley, 2004; Lango Allen et al., 2010). In addition to bone quality and BMD, complex traits include height, body mass index and skin colour, amongst others, and are frequently calculated to have high heritability (>0.5), which is a measure of the proportion of the phenotype that is determined by genetic factors (Peacock et al., 2002). To elucidate these underlying factors, statistically significant associations between genetic polymorphisms and phenotypic traits are sought in either genome-wide or candidate gene association studies (GWAS or CGAS), each of which offers separate limitations and benefits (Frazer et al., 2009; Lunetta, 2008).

A common starting point when assessing complex traits is to adopt a genome-wide approach, in which a dense set of genetic markers [predominantly single nucleotide polymorphisms (SNPs)] is used to interrogate the entire genome in order to unearth regions that may contribute to the phenotype of interest (Karasik et al., 2010). To enhance both the coverage and affordability of the approach, whilst limiting the

statistical computation, the occurrence of linkage disequilibrium (LD) is leveraged, which is the non-random association of alleles that thus occur at frequencies greater than expected (Hirschhorn and Daly, 2005; Teo et al., 2010). Within regions of LD, often known as LD bins, a single marker can serve as a proxy for all other markers within that region, affording the requirement of only one genotype to assess the entire region (Hirschhorn and Daly, 2005). The hypothesis-free nature of this approach is attractive, but genomic regions that are found to be associated are often large and contain numerous genes and their corresponding regulatory regions (Duncan and Brown, 2010; Lunetta, 2008), each of which must be investigated further if the susceptibility variants are to be found. Further complication arises from the fact that the human genome lacks full annotation (Church et al., 2011), so a thorough analysis of a genomic region may still fail to produce significant findings if it contains currently unknown genes or regulatory elements. In addition, results from GWA studies conducted on one population group are often not transferrable to other population groups due to differing allele and genotype frequencies (Frazer et al., 2009; Rosenberg et al., 2010; Teo et al., 2010).

Candidate gene association studies are more focused, but require a strong hypothesis to drive the investigation. Genetic markers within one or more genes are assessed for their correlation with phenotype, provided that these genes are known to play a prominent role in the establishment of that phenotype. An *a priori* hypothesis limits the application of this study design to the realm of our current knowledge on genes and gene function, but since each candidate gene is investigated with a higher density of SNPs across the region, this narrower focus is more likely to highlight a small sub-region of the gene, in which a causal variant is more easily discerned (Lunetta, 2008; Nguyen and Eisman, 2006).

Both GWA and candidate gene studies have generated important findings that have enhanced our understanding of complex traits, but the overall success of correlating genotype with phenotype has fallen well below expectation. Commonly, several appropriate loci have been linked to complex phenotypes at a statistically significant level, but there functional relevance is seldom known and together, these loci tend to make minimal contributions to the heritability of the trait (McCarthy et al., 2008). For example, human height (heritability ~0.8) has been investigated in extensive detail, with a total of 180 loci that have been implicated in determination of the trait. Combined, these loci account for just 10% of the overall heritability, leaving a substantial portion for which no currently known explanation exists (Lango Allen et al., 2010). This phenomenon is also observed when trying to elucidate the genetic mechanisms of common diseases, and has been dubbed as the "missing heritability" problem. Hypotheses as to what may account for the remaining heritability have been proposed, such as epigenetic factors, rare variants of major effect, gene-gene and gene-environment interactions but , to date, little evidence exists to support these claims (Frazer et al., 2009; Manolio et al., 2009; van Ijzendoorn et al., 2011).

1.3.2. Genetic studies on bone health

1.3.2.1. BMD heritability

Assessing the genetic determinants of bone health has largely been conducted by attempting to associate BMD with genotype, often in the context of osteoporosis. Other phenotypic measures such as bone quality or fracture have been considered as well, but these suffer limitations that make their use problematic (Duncan and Brown, 2008). Since BMD varies throughout the skeleton (**Section 1.2.1**) correlations to genotype need to be tested for in a site specific manner, often

involving sites that are prone to fracture such as the lumbar spine (LS) and the femoral neck (FN). Both twin and family studies have been performed in order to calculate the heritability of BMD, which was found to be high, attaining ranges of 60-90% in twins (80-90% when considering monozygotic twins only) and 45-70% in families (Li et al., 2010; Tranah et al., 2008). More specifically, femoral neck BMD along with the axial skeleton was estimated to be between 70-85%, whilst the appendicular skeleton was determined to be 50-60% heritable. Heritability has been detected both at a young age, when bones are still developing, as well as late into adulthood after the gradual decrease in bone mass has started (Peacock et al., 2002; Styrkarsdottir et al., 2003). Studies in other species have shown similarly high heritability for BMD, including in non-human primates (Duren et al., 2007).

Given their close relationship, it follows that risk for osteoporosis is as similarly heritable as BMD, when comparing osteoporotic parents and their offspring. Daughters of osteoporotic mothers have been found to have lower LS and FN BMD than their age-appropriate value, and a similar scenario was discovered in sons of fathers with idiopathic osteoporosis. Indeed, any family history of osteoporosis appears to confer increased risk for the disorder (Ferrari, 2008; Li et al., 2010; Paradowska and Lacki, 2007; Ralston and de Crombrugghe, 2006).

1.3.2.2. Association studies

With the burden of osteoporosis and other skeletal disorders being brought to the fore by the WHO (**Section 1.4**), along with technological advances in high-throughput genetic research, the past fifteen years has seen a wealth of BMD and osteoporosis association studies (Li et al., 2010; Nguyen and Eisman, 2006), but these have still been marred by the problems typical of complex trait genetics (**Section 1.5.1**). Linkage analyses and GWA studies have identified loci relevant to

BMD on every chromosome (Duncan and Brown, 2010; Ferrari, 2008). Many of these loci, however, have proved contentious and inconclusive, likely due to small sample sizes or poor study design (Ng et al., 2006; Tranah et al., 2008). In response to this problem, two large consortiums were established, namely GENOMOS (Genetic Markers for Osteoporosis) and GEFOS (Genetic Factors for Osteoporosis), which have sought to conduct more powerful investigations with larger sample sizes, or through meta-analytical approaches (Duncan and Brown, 2010; Rivadeneira et al., 2009; Uitterlinden et al., 2006). A recent GEFOS meta-analysis attempted to strengthen results pertaining to 150 osteoporosis candidate genes, only nine of which were discovered to reach statistical significance. These nine genes either housed, or were in close proximity to, SNP markers that were strongly linked ($P < 2.39 \times 10^{-6}$) to either LS or FN BMD, and in some cases, both (Richards et al., 2009). A further meta-analysis based on five GWA studies assisted in validating seven previously known BMD loci, whilst also identifying 13 new loci not originally linked to the trait (Rivadeneira et al., 2009). In the wake of these and numerous other association studies, Li and colleagues (2010) comprehensively reviewed the field, listing both repeatedly confirmed candidate genes as well as newly discovered genes that showed substantial promise. At the time of their research, a total of 15 genes were considered to be strong candidates, having been repeatedly associated to BMD, often in separate studies of different methodology (**Table 1.1**).

Table 1.1: Confirmed candidate genes influencing osteoporosis susceptibility

Gene	Chromosome	Pathway/function	Study confirmation[1]
VDR	12	Vitamin D endocrine pathway	CGAS; MA-C
ESR1	6	Oestrogen endocrine pathway	CGAS; GWAS; MA-C; MA-G
ESR2	14	Oestrogen endocrine pathway	CGAS
LRP5	11	Wnt/β-catenin pathway	CGAS; GWAS; MA-C; MA-G
LRP4	11	Wnt/β-catenin pathway	GWAS; MA-C; MA-G
SOST	17	Wnt/β-catenin pathway	CGAS; GWAS
GRP177	1	Wnt/β-catenin pathway	MA-G
TNFRSF11B	8	RANK/RANKL/OPG pathway	CGAS; GWAS; MA-C; MA-G
TNFRSF11A	18	RANK/RANKL/OPG pathway	CGAS; GWAS; MA-G
TNFSF11	13	RANK/RANKL/OPG pathway	CGAS; GWAS; MA-G
COL1A1	17	Structural protein	CGAS; MA-C
SPP1	4	Bone remodelling	CGAS; MA-C
ITGA1	5	Structural protein	CGAS; MA-C
SP7	12	Transcription factor	GWAS; MA-G
SOX6	11	Transcription factor	GWAS; MA-G

[1]CGAS – candidate gene association study; GWAS – genome-wide association study; MA-G – meta-analysis of GWAS; MA-C – meta-analysis of CGAS. Adapted from Li et al., 2010.

1.3.2.3. Candidate genes

In addition to achieving statistical significance, the 15 candidate genes most highly associated with BMD also display functional relevance. The structural protein collagen 1α1 is an important constituent of the bone matrix, coded for by *COL1A1*. Importantly, *COL1A1* mutations are well linked to 90% of osteogenesis imperfecta cases (Huang and Kung, 2006). *SPP1* encodes secreted phophoprotein 1 which helps with bone remodelling (Li et al., 2010) whilst *ITGA1* codes for a protein that assists in the proper formation and assembly of collagen fibres (Lee et al., 2007). Several components of the Wnt/β-catenin pathway have also been implicated including *LRP4* and *LRP5*, which both encode receptors that are part of the frizzled family of G protein-coupled receptors. Similarly, a negative regulator of the pathway, sclerostin is encoded by *SOST*, and a participating G-receptor protein by *GRP177*. Candidate genes *TNFRSF11A*, *TNFRSF11B* and *TNFSF11* code for

RANK, OPG and RANKL respectively, the three molecules that play the most central role in regulating bone health (**Section 1.1.3**). Mouse models have been used to develop scenarios in which OPG (*TNFRSF11B*) is either over expressed, resulting in an osteopetrotic phenotype (high bone mass), and under expressed, causing severe osteoporosis (Lee et al., 2010). Additionally, studies on monogenic bone disorders such as high bone mass syndrome, Paget's disease, sclerosteosis and osteoporosis-pseudoglioma have further implicated genes in the Wnt and RANK/RANKL/OPG signalling pathways (Huang and Kung, 2006; Rivadeneira et al., 2009). Unsurprisingly, given the dramatic effects of oestrogen on bone health (**Section 1.1.2-1.1.3**), the oestrogen receptor α gene (*ESR1*) has been validated using four different study designs (**Table 1.1**)(Li et al., 2010). Furthermore, in 1994, a 28 year old man with low BMD was found to be homozygous for an inactivating mutation in *ESR1*, and a similar scenario was observed in mice lacking a functional copy of the gene (Korach, 1994; Smith et al., 1994). Despite both a statistical and functional importance, however, the combined effects of the polymorphisms in these genes account for less than 5% of the heritability of BMD, with the remaining percentage contributing to the missing heritability problem (**Section 1.3.1**) (Huang and Kung, 2006; Rivadeneira et al., 2009).

Two candidate genes, and their respective proteins, warrant closer inspection due to the strength of their association with bone health, namely *ESR1* and *TNFRSF11B*. *ESR1*, located on chromosome 6, is 492.93kb in length, consisting of eight coding exons and producing 17 known transcripts (Ensembl release 65). Transcript diversity arises from the numerous different transcription start sites at the 5' end, and gene structure is further complicated by the presence of a multitude of non-coding exons (Weickert et al., 2008). The wild-type transcript is translated into oestrogen receptor α (ERα), with several other transcripts encoding truncated

isoforms of this protein (**Figure 1.10**). Determining the exact function of all known transcripts has been difficult, however, as they are often simultaneously expressed with the wild-type version (Deroo and Korach, 2006). Regardless, ERα is known to occur in the cell nucleus where it awaits binding by oestrogen or other selective oestrogen receptor modulators. Binding induces certain conformational changes that promote dimerization of the receptors that then attach to specific DNA sequences known as oestrogen response elements. Co-activators or co-repressors are then enlisted to either up- or down-regulate gene expression. Before 1987, oestrogen was not thought to influence bone cells, but it is now well known that OBL, OCL and OSC cells all express ERα, which is predominantly responsible for mediating oestrogen's effect on bone, especially cortical bone. Notably, trabecular bone relies more on oestrogen receptor β, which is encoded by *ESR2 (*Weitzmann and Pacifici, 2006).

TNFRSF11B, located on chromosome 8, covers 48.64kb and consists of five exons with three known transcripts (Ensembl release 65). The wild-type transcript encodes the cytokine OPG, which is a soluble, dimeric protein, part of the tumour necrosis receptor superfamily, and contains three major domains (**Figure 1.11**)(Jurado et al., 2010; Seremak-Mrozikiewicz et al., 2011; Theoleyre et al., 2004). At the N-terminal is a cysteine-rich domain that facilitates dimerization and is essential in inhibiting osteoclastogenesis. The heparin-binding domain interacts with various proteoglycans, whilst the death domain homologous regions are of currently unknown significance. The two remaining transcripts are splice variants that incorporate either all or a portion of intron 2. Along with RANK and RANKL, OPG is pivotal in bone health maintenance, but this molecular triad also has relevance to immune and vascular biology (Theoleyre et al., 2004).

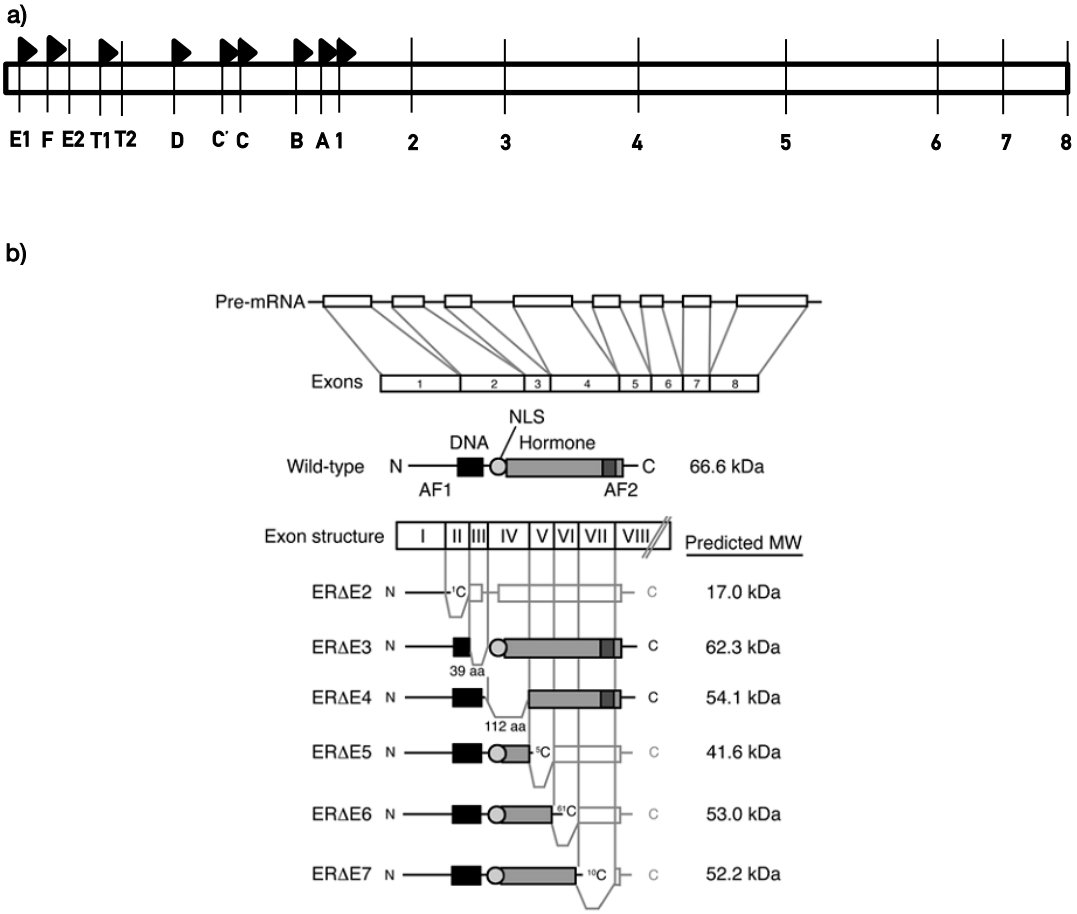


Figure 1.10: Structure of the *ESR1* gene and ERα protein. a) *ESR1* consists of eight coding exons (labelled 1-8; exons are displayed by vertical bars), but the 150kb region upstream of the first coding exon (not drawn to scale) contains numerous non-coding exons E1-A, many of which have their own unique promoter (black triangles). Through splice variation, some of these non-coding exons are incorporated into ERα and its isoforms. b) The wild-type transcript from *ESR1* consists of all eight coding exons and translates into a receptor protein consisting of a DNA-binding domain (DNA), a hormone-binding domain (Hormone), a nuclear localisation signal (NLS) and two activation functions (AF1 and AF2). Splice variants (E2-E7) exclude particular exons resulting in truncated proteins (Deroo and Korach, 2006; Stevens and Meech, 2006; Weickert et al., 2008).

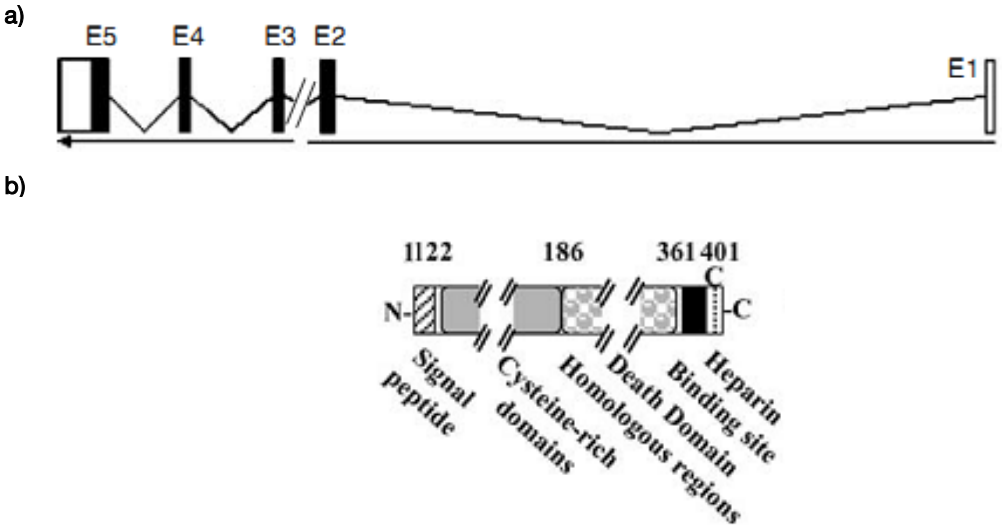


Figure 1.11: Structure of the *TNFRSF11B* gene and OPG protein. a) *TNFRSF11B* is coded on the minus strand and thus gene structure is represented from right to left. The gene has only five exons and three splice variants. b) The OPG protein is a soluble, decoy receptor for RANKL, and it consists of three main domains. The cysteine-rich domain facilitates dimerization and is also the domain that interacts with RANKL. The heparin-binding domain is important for proteoglycan interactions, whilst the functions of the death domain homologous regions are not well understood (Jurado et al., 2010; Theoleyre et al., 2004).

1.3.2.4. Pleiotropic, gender and ethnic–specific effects

Whilst currently identified candidate genes provide numerous avenues of exploration for potential drug targets, a couple of important caveats need to be borne in mind. As mentioned previously, most associations are detected through site-specific analyses that suggest different sets of genes regulate bone mass at separate skeletal sites, but there is also evidence of pleiotropic effects, implying a common set of a genes that determine bone mass across the skeleton. How these gene sets differ may be an important consideration (Garnero et al., 1996; Nguyen et al., 1998).

Similarly, much of the genetic research to date has been performed using samples of Caucasian females, one of the population groups currently most susceptible to osteoporotic fracture and thus necessitating the highest level of priority (Guggenbuhl, 2009). Results have shown, however, that both sex and ethnic

specific effects play a role in the establishment of BMD (Huang and Kung, 2006; Karasik et al., 2010). A mouse model study demonstrated that only three of five quantitative trait loci (QTL) were common between both genders. Similarly, sex-specific QTLs have been reported in humans, including 10q21 which links to FN BMD in males under 50, whilst female specific loci include 18p11 (LS BMD in females over 50) and 20q13 (females 50 years or younger). Furthermore, heritability differences were seen between pairs of opposite sex and same sex individuals, suggesting a possible involvement for imprinting and other epigenetic phenomena (Karasik, 2008). These would do well to explain the BMD differences that are apparent between males and females starting from puberty (**Section 1.1.2**).

Where results from European studies have been tested amongst populations of varying ethnicity, a number of important differences have also been detected. Separate ethnicities have characteristic genetic backgrounds that differ in terms of allelic and genotypic frequencies. This was highlighted by the fact that the high risk allele, denoted “s”, at the Sp1 binding site of *COL1A1*, which was strongly associated with fracture, occurs at a frequency of 15 to 32% in Caucasians, but was not detected at all in an East Asian population (Lambrinoudaki and Kung, 2001; Lei et al., 2006). A similar situation was observed for the *BsmI* site in the Vitamin D receptor gene, which provides a strong linkage signal in Caucasians and Brazilian women, but has no heterozygosity in Oriental populations (Kung et al., 1998). Among African American women, the genotype distribution for a *BsmI* polymorphism was determined to be significantly different from white women, with African Americans more frequently possessing the high bone mass genotype (bb), but never the low bone mass alternative (BB) (Nelson et al., 2000). Along with many other examples, these results strongly incriminate genetics as the cause for bone

health differences between ethnicities, although possible local, population specific effects may also be important (Uitterlinden et al., 2006).

1.3.2.5. Further required bone health research

Despite substantial progress over the last few years, a wealth of research remains to be conducted on bone health, as well as its relation to osteoporosis. The pace at which new loci are linked to BMD is encouraging, but few studies have attempted characterise the functional impacts of associated variants (Li et al., 2010). It is also important to bear in mind that BMD still remains the only phenotype studied to an appreciable degree, and it may be worthwhile to focus more attention on the genetics of bone quality and fracture, despite the inherent difficulty. Although intimately tied to BMD, studies have suggested that bone quality and risk of fracture may be governed by different sets of genes that would not otherwise be highlighted in association studies aimed purely at BMD (Huang and Kung, 2006; Paradowska and Lacki, 2007; Peacock et al., 2002). Similarly, the genetic factors influencing loss of bone mass have not been suitably investigated, despite this being the primary cause for osteoporosis (Peacock et al., 2002).

Research focus also needs to be shifted away from Caucasian females. While they may represent one of the populations most severely affected by osteoporotic fracture, men and women of all ethnicities are experiencing a rise in prevalence of the disorder (Duncan and Brown, 2008). Moreover, identifying the genetic effectors that underlie consistent physiological differences between ethnicities may help to further elucidate biological pathways that regulate bone health (Duncan and Brown, 2010; McCarthy et al., 2008). The role of genetics in bone health during childhood also warrants further attention, since bone accrual at this stage is assumed to determine osteoporosis risk in later life. Heritability differences between pairs of

younger and older twins have suggested that childhood bone mass may be under greater genetic influence, an idea further supported by mouse model work (Duren et al., 2007).

1.4. Bone health amongst South African populations

Preceding sections have highlighted that much of bone health research has incorporated African Americans as representatives for African populations, despite the well understood inaccuracy of this approach. A significant body of literature has clearly demonstrated levels of Caucasian and other genetic admixture amongst current African American populations following years of integration into local US communities (Gong et al., 2006a; Parra et al., 1998). Since admixture appears to dilute effects that may appear stronger in more separately defined ethnicities (Hill et al., 2008), it has long been suggested that bone research should be conducted on individuals still resident on the African continent, where blacks free of this admixture may demonstrate even stronger bone phenotypes. Investigation into this claim has been performed in South Africa, with many significant contributions to the bone health field (Hill et al., 2008; Micklesfield et al., 2009; Vidulich et al., 2011). A study on the structure of the iliac crest revealed that black South Africans had both lower cortical porosity and greater cortical thickness. Furthermore, mineral apposition rate was found to correlate with osteoid thickness, supporting the argument that bone formation amongst blacks is more aggressive (Schnitzler and Mesquita, 2006). In line with research on African-Americans, Walker and colleagues found that the South African Bantu population had higher vertebral BMD than their white counterparts, whilst an investigation on female nurses suggested that proximal femur (PF) BMD was higher amongst black participants, however, no difference was observed at the lumbar spine (Daniels et al., 1995). Correspondingly, multiple studies have also demonstrated the lower prevalence of fracture amongst

African individuals (Schnitzler et al., 2009; Thandrayen et al., 2009). Indeed, a 10-fold lower prevalence of hip fractures is observed between black and white South Africans.

1.4.1. Birth to Twenty

More recently, extensive research into bone health in black Africans has been conducted as part of the Birth to Twenty study (Bt20). This longitudinal study consists of a cohort of individuals that were recruited following their birth within the six week period spanning April 23rd to June 8th 1990. Recruitment occurred at major public sector hospitals within the Soweto-Johannesburg metropolitan area, which are, by and large, frequented by black citizens. The initial motivation behind the cohort was to provide an unsurpassed 10 year analysis of the growth and development of the cohort members. This would include annual monitoring of both physical and mental progress, whilst documenting the environmental context in which the progress was made. At the end of the 10 year period, it was decided that the study should continue for a further 10 years, hence becoming the Birth to Twenty study. To enhance the strength of the cohort, the overrepresentation of black individuals amongst the 3273 original members was mitigated through the incorporation 120 white children who had also been born within the same timeframe. After 16 years, the study was still strongly supported by over 70% of members whom had continued to submit themselves to regular evaluation (Richter et al., 2007; Thandrayen et al., 2009).

A sub-cohort of Bt20 members was formed for the purposes of bone health analysis, designed to meet the growing demand for both African and childhood relevant data. In addition to the standard assessment, these individuals were subjected to annual DXA measurements. BMD measurements on children,

however, are a contentious issue since correcting for covariates such as height, weight, body and bone size is more complicated at a young age (Bianchi, 2007). Similarly, due to the lack of research on childhood bone health (**Section 1.3.2.5**), little standardised reference data for interpreting BMD is available, thus the amount of bone is frequently underestimated in children (Bachrach, 2007; Bianchi, 2007). It has thus been simpler to use the crude measure of bone mineral content (BMC) when analysing pre- and peri-pubertal children, although this is also subject to proper correction of covariates. Birth to Twenty studies have, therefore, adopted this approach, adjusting for height and weight in initial studies when cohort members were 10 years of age. Results were less than expected, however, with only a 6% greater density at the femoral neck for black children, but no difference at the lumbar spine; a finding similar to that amongst the adult female nurses mentioned previously. Since the vast array of correction methods for both BMC and BMD measurements are known to generate different sets of results that then undermine the value of any single corrective approach, the comparison between black and white children was recently revisited using multiple covariate corrections. Results showed that adjusting BMC for bone area, raised to a pre-calculated, population specific power coefficient, greatly magnified differences at the femoral neck, from 7 to 69% in girls and from 8 to 64% in boys. Significant lumbar spine differences were also seen for the first time (4 to 85% in girls; 3 to 34% in boys) (Vidulich et al., 2011). These results more appropriately accounted for observed fracture rate discrepancies, with white children experiencing ±53 fractures per 1000 children/year, but only ±18 amongst black children (Micklesfield et al., 2009).

Placing the differences between black and white bone health in further contrast have been environmental factors. The socioeconomic status of black individuals is generally poor and as a result, malnourishment is a common problem. Specifically,

the traditional diet of black South Africans lacks calcium, and daily intakes are well below the recommended daily allowance of 800mg (MacKeown et al., 2003). The same is true for vitamin D, which is further complicated by the limited ability of ultraviolet rays to activate vitamin D pathways in the body, due to the dark pigmentation of black skin (Fausto-Sterling, 2008; Thomas, 2007). A lack of vitamin D has notable knock-on effects, including an elevated level of parathyroid stimulating hormone (PTH) which encourages bone loss (Rizzoli et al., 2001). Furthermore, vitamin D deficiency results in muscle weakness, increasing the risk of falls (Pietschmann et al., 2009). In the aforementioned study on nurses in South Africa, black nurses were observed to have lower serum total calcium and 25-(OH) vitamin D (Daniels et al., 1995). Similarly, African Americans tend to be vitamin D deficient as well (Fausto-Sterling, 2008; Thomas, 2007). Combined, these two detrimental factors should reflect poorly upon bone health in black individuals, but no such consequence is seen. While it may be suggested that the rural lifestyle of most black South Africans entails more physical labour, and is thus the reason for their stronger bones, a Birth to Twenty investigation revealed that black children are more inactive than their white counterparts (McVeigh et al., 2004). Since lifetime bone health is largely established at this age, physical activity differences fail to account for differences in bone mass.

1.4.2. Genetic research on black South Africans

The apparent inability of adverse environmental factors to deter healthy bone development in black individuals, combined with recent findings that more appropriately highlight bone physiological differences between black and white populations, provides even greater incentive to investigate highly probable underlying genetic mediators. This has not yet been attempted amongst Birth to Twenty participants, or indeed amongst any non-admixed African population,

despite the fact that successfully identifying pertinent genetic factors is more likely when using an African sample (Gong et al., 2006a). Africa is considered the birthplace of anatomically modern humans based on the unparalleled levels of human genetic variation seen amongst indigenous Africans (Campbell and Tishkoff, 2010; Henn et al., 2011; Ramsay et al., 2011). This is best observed when considering regions of linkage disequilibrium, which are far smaller than in other population groups, reflecting the increased time period in which Africans have endured sexual recombination (Teo et al., 2010). Since the basis of cost-effective and informative association studies lies in selecting tagging markers in each of these LD regions that act as a proxy for other genetic variants (**Section 1.3.1**), causal variants can be traced to much smaller regions in Africans. This allows for fine mapping of association signals, greatly enhancing the likelihood of identifying true susceptibility loci (Duncan and Brown, 2010).

Several additional reasons also motivate for further research of bone health in black Africans. Most significantly, concerning links between HIV and bone health have recently emerged. HIV affected patients have been found comorbid with numerous bone diseases including osteonecrosis, osteopenia and osteoporosis (Mondy et al., 2003). Furthermore, the treatment of HIV has been demonstrated to be detrimental to bone health. Specifically, a history of protease inhibitor use has been linked to decreasing BMD at both the hip and the spine (Mondy et al., 2003). Since Africa suffers with the highest burden of HIV/AIDS and African men and African women represent the two most susceptible population groups (Shisana et al., 2009), low bone health and subsequent osteoporosis are poised to become significant problems. Moreover, with a poor socioeconomic background, many Africans are ill-equipped to deal with the consequences that accompany osteoporotic fracture (Duncan and Brown, 2008). Accordingly, studies in America have shown that

African American women have a higher mortality rate than that of Caucasian women following hip fracture (Miller et al., 2005; Mudano et al., 2003). Moreover, treatment options that are designed on the back of purely Caucasian data are likely to be ineffective for pharmacogenetic reasons (Lei et al., 2006; Massart, 2005; Nguyen and Eisman, 2006). Medical literature has repeatedly identified cases where African patients respond differently to medication (Alving et al., 1956; Oliveira et al., 2009; Suarez-Kurtz, 2005), because their high genetic diversity increases the likelihood of unforeseen drug reactions, highlighting the importance of understanding local genetic structure.

1.5. Study aim and objectives

The purpose of this study was to initiate a body of genetic knowledge relevant to bone health in black Africans. Since no prior knowledge exists on this matter, the main aim was to investigate the applicability of European results to a non-admixed African sample. Based on the results of the GEFOS consortium meta-analysis, and their repeated validation in other study designs (**Section 1.3.2.3**), six of the 15 known BMD candidate genes were selected for analysis, namely *ESR1*, *TNFRSF11A*, *TNFRSF11B*, *TNFSF11*, *SOST* and *SPP1*. These are amongst the most well validated candidate genes influencing BMD, which were thus prioritised as kit size limitations prohibited the assessment of additional genes. The research objectives were thus to:

- Obtain DNA samples and bone phenotypic data from the bone health sub-cohort of the Birth to Twenty study
- Select SNPs with the strongest association signals within the six chosen genes, according to European-based studies

- Select tag SNP markers that would enable the interrogation of the candidate genes for association signals that are possibly unique to Africans
- Genotype the selected panel of SNP markers within study participants
- Determine statistically if an association existed between SNP markers and bone mineral content, accounting for height, weight, bone area and sex as covariates.

It was hypothesised that not all European results would be fully replicated amongst an African cohort, and that several new association signals would be detected instead, reflecting the different genomic architecture between these two populations.

2. Materials and Methods

2.1. Subjects

Study participants initially included 359 black and 68 white individuals that derive from the larger bone health sub-cohort (n = 682) of the Birth to Twenty study (note that for the purposes of this study, “black” is used to define individuals of African ancestry without known admixture, whilst “white” denotes individuals of European descent). Cohort members had been enrolled based on the exclusion criteria of any pre-existing conditions that would impact on bone mineral content, including rickets, tuberculosis, cancer, a history of endocrine disease, diseases of the thyroid gland, gastrointestinal disease associated with malabsorption, and regular medication use, specifically corticosteroids, anti-epileptic drugs, calcium and/or vitamin D supplementation. Since participants were recruited from the Soweto-Johannesburg metropolitan area, black participants were all southern Bantu-speakers, belonging to the “S” group according to Guthrie classification (Guthrie, 1948). Unpublished data (Lombard et al.) in which 18 ancestry informative markers were measured amongst 1000 black Bt20 members confirmed that no high-level population sub-structure was evident, and thus no additional stratification of the study group was implemented. Although growth and development had been tracked since birth, this study used phenotypic data relevant to when cohort members were of mean age 10.6 years in order to avoid complications posed by puberty, such as rapid growth and bone remodelling that hinders direct comparison between individuals (Duren et al., 2007). By this criterion, however, 42 samples had to be excluded due to missing bone phenotypic data at this age (**Figure 2.1**). The power of the study, based on the sample size, was calculated using Quanto (Gauderman and Morrison, 2006).

Informed consent from a legal guardian was obtained prior to enrolment into the study. Ethical approval for this study was granted by the University of the Witwatersrand Human Research Ethics Committee (Medical); certificate number M10457 (**Appendix A**).

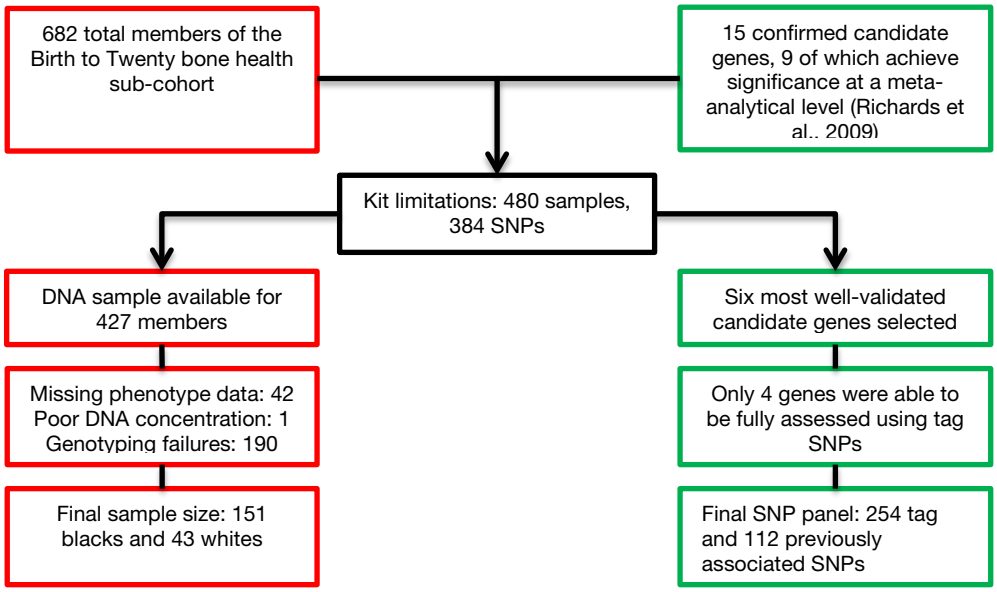


Figure 2.1. Flow diagram indicating the selection of study participants, candidate genes and SNP markers.

2.2. Anthropometry

Measurements of height, to the last completed 1mm, were previously taken using a wall-mounted stadiometer (Holtain, Crosswell, UK), whilst weight, to the nearest completed 0.1kg, was recorded using a digital electronic instrument (Dismed, Halfway House, South Africa) based on standardized protocols (Lohman et al., 1991). Both instruments were regularly calibrated, and subjects wore minimal clothing when being weighed.

2.3. Measurement of bone

Measurements of bone area and mineral content at the hip and L1 to L4 vertebrae were previously performed using the Hologic QDR-4500A DXA machine (Bedford, MA). Using a spinal phantom, the coefficient of variation was calculated as 0.44% for BMC (g); 0.36% for bone area (cm^2) and 0.37% for BMD (g/cm^2). A single technician conducted all readings to avoid an additional source of variation (de Villiers, 2009). Reliability of the technician is reported in **Table 2.1**. Participants were wearing only light clothing at the time of measurement, and all recordings were done within a timeframe that avoided possible seasonal variances in bone mass (Hasselstrom et al., 2006).

Table 2.1: Reliability of DXA technician

Variable (n=20)	Mean ± SD	Standard error of measurement
Spine: BMC (g)	60.9 ± 11.1	0.55
Spine: BA (cm^2)	59.4 ± 7.6	0.36
Spine: BMD (g/cm^2)	1.1 ± 0.1	0.003
Hip: BMC (g)	28.2 ± 4.25	0.54
Hip: BA (cm^2)	31.9 ± 3.3	0.56
Hip: BMD (g/cm^2)	0.9 ± 0.1	0.004

2.4. Sampling of blood and DNA extraction

A 5ml venous blood sample was drawn from each patient by a qualified phlebotomist. DNA was previously extracted from whole blood using the salting out procedure (Miller et al., 1988), and stored in Tris/EDTA (TE) buffer at 4°C, in the National Health Laboratory Service (NHLS) biobank. Sample concentrations were measured using the Tecan Infinite® 200 PRO NanoQuant (All equipment details are listed in **Appendix B**). For each sample, a volume of 2µl was loaded and concentration was determined using spectrophotometry based on the Beer-Lambert law: $A = \varepsilon c \ell$, where "A" is the absorbance at a given wavelength of light,

"ε" is the molar extinction coefficient, "c" is the molar concentration and "ℓ" is the path length (Bouguer, 1729). Dilutions, normalized to 50ng/µl, were prepared from stock solutions by means of a robotic pipetting system (Tecan Freedom EVO®) which automatically calculated required volumes of solvent (TE buffer). One sample was excluded at this point due to an insufficient concentration of DNA. **Appendix C** lists all study participants, and their corresponding stock DNA concentrations.

2.5. Candidate gene selection

Candidate genes were selected on the basis of the recent meta-analysis conducted by the GEFOS consortium (Richards et al., 2009). Of the nine genes strongly linked to either LS or FN BMD, six genes were selected for validation amongst black South Africans, namely *ESR1*, *TNFRSF11A*, *TNFRSF11B*, *TNFSF11*, *SOST* and *SPP1*. Kit size limitations and time restraints prevented the inclusion of additional candidate genes. A total of 189 SNPs, spanning across all six genes, were found to generate strong association signals amongst Europeans, and these were all considered for inclusion within this study ($P < 2.39 \times 10^{-6}$).

2.6. Tag SNP selection

In addition to assessing the validity of previously associated SNPs identified in populations of European ancestry, a panel of tag SNPs was selected for the purposes of identifying association signals that may be unique to this population group, but only within genes *ESR1*, *TNFRSF11A*, *TNFRSF11B*, and *TNFSF11* due to kit size limitations. Tagging SNPs are maximally informative markers that act as proxies for additional, nearby SNPs that share the same LD bin, thereby providing greater coverage of candidate genes and eliminating information redundancy (Hirschhorn and Daly, 2005). Tagger (de Bakker et al., 2005) was used to identify tag SNPs through a multi-marker tagging approach with a minor allele frequency of

1% and an r^2 cutoff of 0.8, based on Yoruba genetic information from The International Hapmap Project (Release 27, Phase II+III, February 2009, on NCBI B36 assembly, dbSNP b126).

All SNPs, both previously associated and tag, were assessed by the Illumina Assay Design Tool (ADT) for compatibility with the GoldenGate genotyping assay. Low scoring SNPs (<0.6 on a scale of 0-1) were excluded, and tag SNP selection was repeated until all chosen SNPs were amenable to the selected assay. The final SNP panel consisted of 112 previously associated and 254 tag SNPs, which provided a coverage range of 90-96% in genes *ESR1*, *TNFRSF11A*, *TNFRSF11B*, and *TNFSF11*. All SNPs are listed in **Appendix D**. Information on these SNPs, including reference allele designation and allele frequency amongst different populations was sourced using HapMart, an extension of BioMart (Smedley et al., 2009). Allele frequencies were rounded-off to two decimal places.

2.7. Genotyping

Genotyping was performed using the GoldenGate genotyping assay with VeraCode microbeads on the Illumina BeadXpress™ platform as described elsewhere (Lin et al., 2009). Briefly, sample DNA is biotinylated and attached to streptavidin-coated paramagnetic beads. SNPs are then interrogated by two allele specific oligonucleotides (ASOs) which are dye-labelled different colours. The correctly annealed ASO is extended and ligated to a downstream locus specific oligonucleotide (LSO). Universal primer sequences flanking assay oligonucleotides afford the use of a single polymerase chain reaction (PCR) to amplify all of the SNP-capturing templates. Fluorescent strands in the resultant PCR products are purified out and attached to specific VeraCode beads by virtue of unique address sequences embedded into each LSO. Finally, beads are scanned by the

BeadXpress™ reader and fluorescent ratios are converted into genotype data. Details of the *Taq* polymerase selected for this assay are shown in **Appendix B**. Assay kits, for the 384-plex level, were supplied by Illumina, and pre- and post PCR steps were conducted in separate areas. Sample plates were scanned within 24 hours of completing preparation in order to maximize fluorescent intensity.

2.8. Quality control

Raw data were examined using the genotyping module of Beadstudio (Framework version 3.1.3.0; module version 3.2.32). Illumina designed built-in assay controls were used as a measure of result quality and samples failing more than two such controls (out of five) were excluded. In terms of SNP performance, only SNPs with a Hardy-Weinberg equilibrium (HWE) *P*-value >0.05 were analysed further. The genotype calls for the remaining SNPs and samples were individually assessed and ambiguous calls that could not be suitably clarified were removed (**Figure 2.2**).

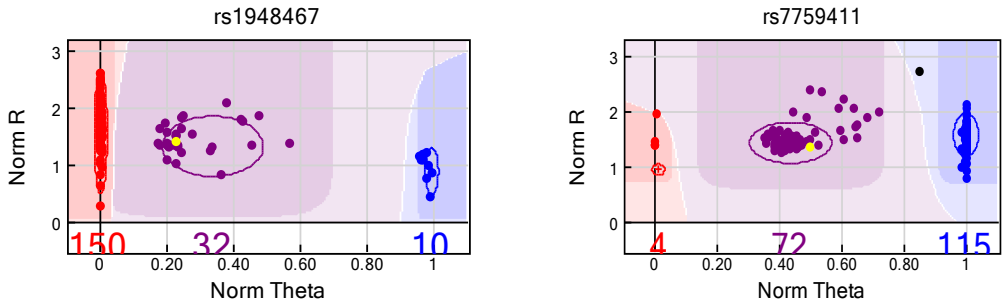


Figure 2.2: Calling genotypes using BeadStudio. Based on fluorescent intensity, one of three genotypes is assigned to individuals (red/blue – homozygous AA/BB; purple – heterozygous AB). Intensity readings that are ambiguous are not assigned a genotype (rs7759411 – black dot) and corresponding individuals are not included for analysis for that specific SNP.

2.9. Data analysis

SAS Enterprise Guide version 4.2 was used to assess the descriptive statistics of the cohort. Normality of the distribution for measured parameters was tested using a Chi-square test for goodness-of-fit. Where a normal distribution was applicable, a

two sample t-test was conducted in order to detect differences between participants, sub-divided by gender and ethnicity. In all other instances, a Wilcoxon-Mann-Whitney test was used to account for skew distributions. Where the equality of variance assumption was not met for t-tests, a Satterthwaite approach was adopted instead. Enterprise Guide was also used to generate all box-and-whisker plots.

gPLINK software (version 2.050)(Purcell et al., 2007) was used to perform a linear regression analysis in order to test for associations between BMC and genotype, using an additive genetic model. This model assumes each additional copy of the variant allele causes either an increase or decrease in BMC. Although offering less flexibility than a general genetic model, which carries no underlying assumptions but requires two degrees of freedom (*df*), an additive model requires only 1 *df*, and is thus more powerful. Sex, height, weight and bone area were entered as covariates, and a label-swapping permutation procedure (n = 1000) was used to correct for multiple testing. Permutation testing is less conservative than the widely implemented Bonferroni correction, and was thus a more suitable approach given the small sample size. A significance level threshold of 5% ($\alpha = 0.05$) was adopted for all tests, and *P*-values were rounded-off to three decimal places.

To further visualise the results for *ESR1* and *TNFRSF11B*, LocusZoom (version 1.1) was used (Pruim et al., 2010) which is available online at http://csg.sph.umich.edu/locuszoom/. This program is designed to generate graphs of association study results by plotting the $-\log_{10}$ (*P*-value) for a studied marker on the y-axis, and its genomic position on the x-axis. Linkage to nearby markers is also shown. These plots are based on data from the The International HapMap Project (Phase II data), the 1000 Genomes Project (August 2009 release),

and current builds of both dbSNP (135) and the UCSC genome browser (GRCh37/hg19), where applicable.

All software is detailed further in **Appendix B**, whilst web resources are listed in the references section (**Chapter 6**).

3. Results

Detailed below are the results for this study including the descriptive statistics for the successfully genotyped individuals as well as the significant results of the association tests. Refer to **Appendices E** and **F** for additional results.

3.1. Descriptive statistics

Table 3.1 shows the descriptive statistics for the study participants that were successfully genotyped (n = 194). These were not significantly different from the values reported elsewhere (Vidulich et al., 2011) for the bone health sub-cohort of the Bt20 study, except for LS BMC and bone area which had proportionally higher means across all four groups. Age and bone age were both comparable between ethnicities, sub-divided by gender, but black girls and boys were significantly shorter than their white counterparts. Black boys tended to be significantly lighter than whites (P=0.054), a trend which is in accordance with significant differences in other studies (Vidulich et al., 2011). Analysis of uncorrected bone measurements revealed significantly lower LS BMC in black boys, but no such difference when comparing the girls, or when comparing FN BMC in either gender. The bone area differed between ethnicity for the femoral neck and for the lumbar spine, but only amongst boys for the latter. These interethnic comparisons all yielded conclusions that were in line with those previously reported (Richter et al., 2007; Vidulich et al., 2011). When considering differences between gender, within an ethnic group, black boys were seen to have higher bone mineral content at the femoral neck, but lower BMC at the spine than black girls. No such difference was detected for whites (**Figure 3.1**).

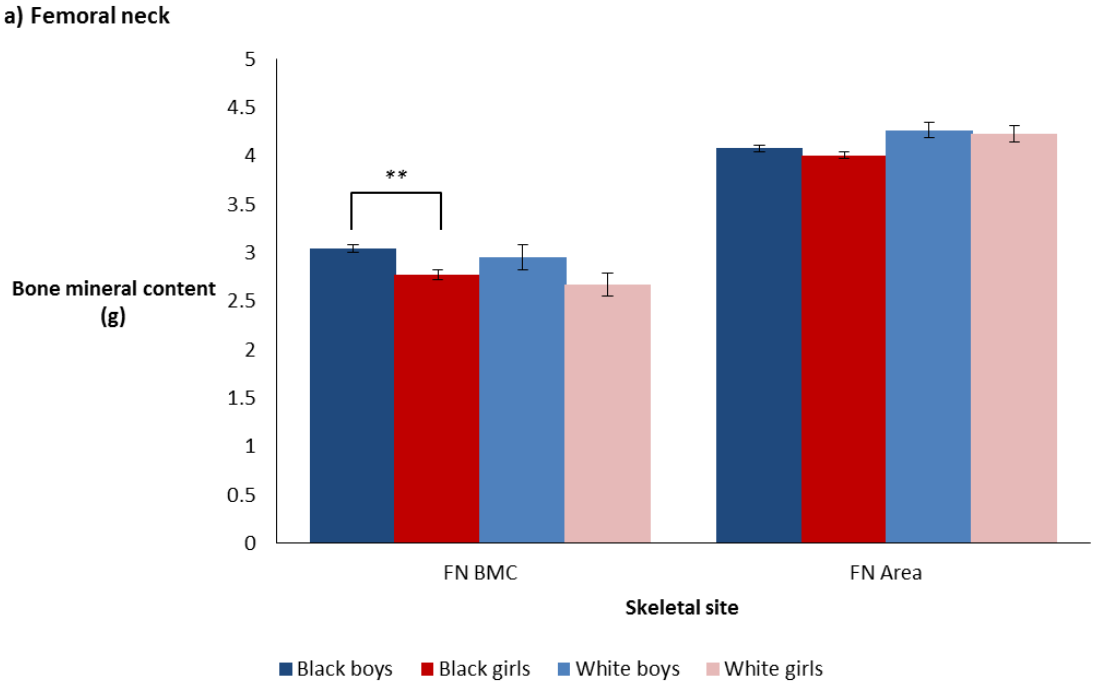
Table 3.1: Descriptive statistics for successfully genotyped participants, with inter-gender comparisons[◊]

	Black boys (n = 80)	White boys (n = 20)	*P*-value	Black girls (n = 71)	White girls (n = 23)	*P*-value
Age (years ± SE)	10.53 ± 0.03	10.59 ± 0.07	0.532	10.55 ± 0.04	10.55 ± 0.08	0.725
Bone age (years ± SE)	9.92 ± 0.09	9.99 ± 0.25	0.516	9.94 ± 0.15	9.99 ± 0.24	0.704
Height (mm ± SE)	1369.96 ± 7.02	1437.35 ± 21.06	0.006	1388.10 ± 7.10	1437.57 ± 17.75	0.015
Weight (kg ± SE)	31.74 ± 0.66	34.88 ± 1.62	0.054	34.69 ± 0.07	37.68 ± 1.91	0.199
FN BMC (g)	3.05 ± 0.04	2.96 ± 0.13	0.543[†]	2.77 ± 0.05	2.67 ±0.12	0.420
FN bone area (cm^2)	4.08 ± 0.03	4.27 ± 0.08	0.024	4.01 ± 0.03	4.23 ± 0.08	0.004[†]
LS BMC (g)	23.46 ± 0.42	26.11 ± 1.05	0.012	25.72 ± 0.67	25.81 ± 1.10	0.742
LS bone area (cm^2)	42.64 ± 0.46	46.27 ± 1.29	0.014[†]	42.95 ± 0.50	44.30 ± 0.83	0.156

† - The equality of variance assumption employed by the two sample t-test was not applicable, thus a Satterthwaite approach was adopted instead. [◊]Bone area and bone mineral content are not corrected for size.

3.2. Genotyping

Following quality checks for sample performance during the genotyping assay, 190 samples were excluded. Machine faults contributed to the majority of sample failures, and were thus not the result of methodology or handling errors. Based on the remaining sample size, the study power was calculated as 5.5%, using Quanto. A single SNP (rs9479134) failed to reach quality thresholds, likely due to incompatibility with the assay design, and was removed from further analysis. Twenty additional SNPs were also excluded as their genotype frequencies were not consistent with Hardy-Weinberg equilibrium (P <0.05). Lastly, 16 SNPs had over 10% missingness (the proportion of individuals unsuccessfully genotyped for that SNP divided by total study participants) and were also excluded. The final SNP panel for analysis consisted of 329 SNPs, out of the 366 initially chosen.



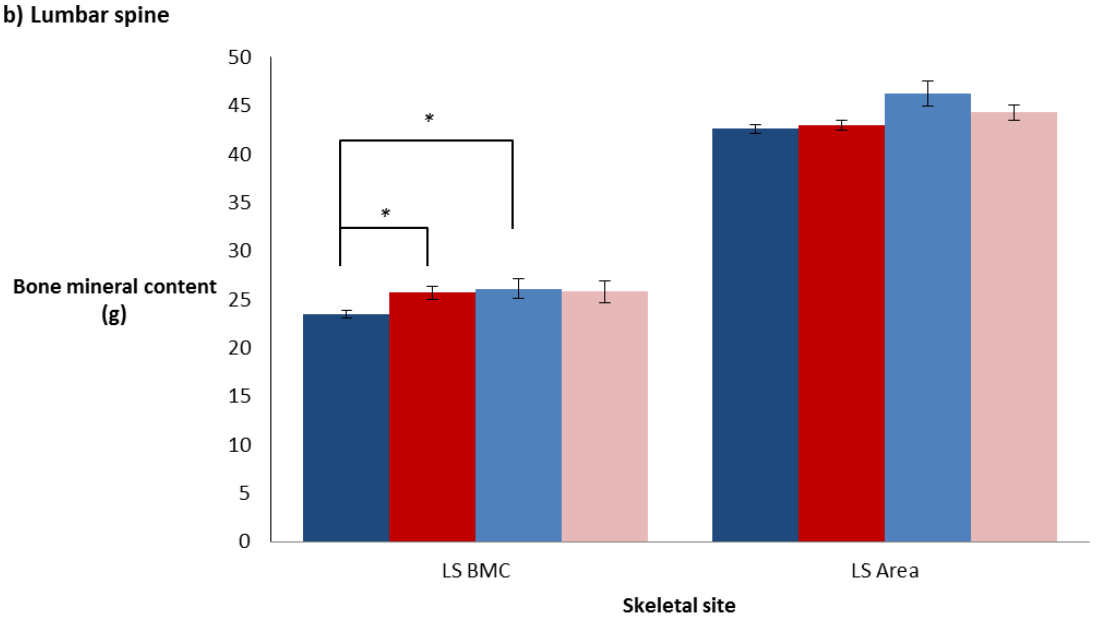


Figure 3.1: Comparison of bone measurements between genders, within an ethnic group. a) At the femoral neck, black boys were observed to have significantly higher BMC than girls (**P<0.01), b) Conversely, at the lumbar spine, BMC was significantly lower amongst black boys when compared to both black girls and white boys (*P<0.05). No significant differences in bone area measurements were seen between genders, within an ethnic group.

3.3. Association results

The number of successfully typed white participants (n = 43) was deemed too low for further analysis and thus association tests were only conducted in the 151 remaining black participants. **Appendix E** lists *P*-values for association with both LS and FN BMC for all typed SNPs. Significant associations are listed in **Table 3.2** below. All reported beta values indicate the effect on BMC [measured in grams (g)] with each additional minor allele, as determined by sample allele frequencies. Thus, a positive beta value indicates that the minor allele is associated with increased BMC, whereas a negative value indicates association with decreased BMC.

A total of 27 SNPs (8 previously associated; 19 tag) achieved pointwise estimates for association with BMC that were significant at the 5% level. Nine SNPs were linked to LS BMC and 18 to FN BMC, but no marker achieved significance for both skeletal sites. The majority of the associated SNPs were located in either *ESR1* (18) or *TNFRSF11B* (9), with one signal deriving from a SNP in *SPP1*. When adjusting *P*-values for the 1000 permutations conducted, across 151 samples, only one SNP marker (rs2485209) maintained significance at the 5% level (**Figure 3.2**).

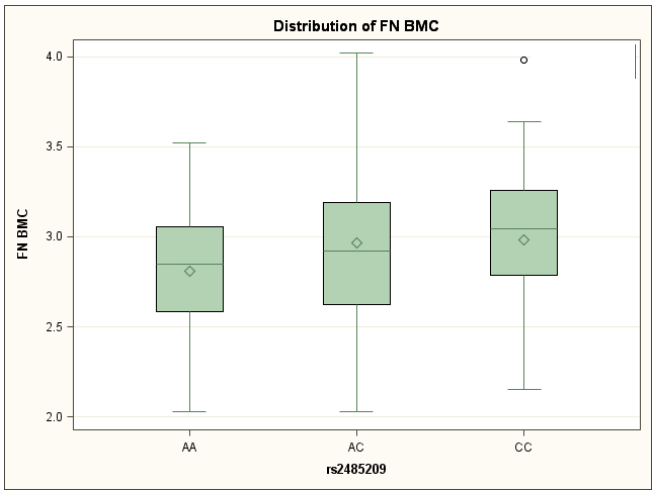


Figure 3.2: Box-and-whisker plot of BMC distribution across the three genotype classes of rs2485209. When raw, unadjusted FN BMC (measured in grams) values for black individuals were plotted, an increasing trend with each C allele at polymorphic site rs2485209 was observed. This was corroborated by the strong association signal detected following linear regression analysis with adjusted BMC values (P=0.002). Diamond – mean; Horizontal line – median.

Table 3.2: List of SNP markers generating significant association signals

SNP ID	Gene	T/P	Chr	A1 (Minor)	A2 (Major)	MAF (sample)	*P*-value (FN BMC)	*P*-Value (LS BMC)	Beta (Effect size[1])
rs11155823	*ESR1*	T	6	T	G	0.16	**0.032**	0.469	-0.100
rs2485209	*ESR1*	P	6	C	A	0.44	**0.002** †	0.717	0.112
rs2504069	*ESR1*	P	6	G	A	0.23	0.710	**0.028**	0.864
rs2982551	*ESR1*	P	6	C	A	0.42	0.398	**0.037**	-0.677
rs3020377	*ESR1*	T	6	A	G	0.15	**0.041**	0.401	0.085
rs3020396	*ESR1*	T	6	A	G	0.27	**0.013**	0.790	0.085
rs4458702	*ESR1*	T	6	A	G	0.10	**0.037**	0.258	-0.129
rs532010	*ESR1*	T	6	C	T	0.46	**0.033**	0.105	-0.065
rs6910500	*ESR1*	T	6	A	G	0.07	**0.005**	0.939	0.174
rs7744006	*ESR1*	T	6	A	G	0.13	0.917	**0.042**	-0.978
rs9322332	*ESR1*	T	6	C	A	0.45	**0.024**	0.337	-0.073
rs9340785	*ESR1*	T	6	C	G	0.05	**0.032**	0.962	-0.161
rs9340820	*ESR1*	T	6	G	T	0.16	**0.019**	0.955	0.095
rs9340939	*ESR1*	T	6	A	C	0.13	0.700	**0.043**	-0.981
rs9371234	*ESR1*	T	6	A	C	0.17	**0.030**	0.451	-0.087
rs9383598	*ESR1*	T	6	T	C	0.14	**0.046**	0.791	-0.102
rs9397453	*ESR1*	T	6	A	G	0.22	**0.034**	0.406	-0.080
rs2869722	*SPP1*	P	4	G	A	0.15	**0.030**	0.446	0.101
rs1032129	*TNFRSF11B*	T	8	C	A	0.33	0.466	**0.050**	0.712
rs11573938	*TNFRSF11B*	T	8	A	G	0.02	**0.022**	0.886	0.317
rs1804854	*TNFRSF11B*	T	8	C	T	0.33	0.373	**0.036**	-0.761
rs4520195	*TNFRSF11B*	T	8	G	A	0.27	0.200	**0.014**	-0.977
rs7464496	*TNFRSF11B*	P	8	T	C	0.06	**0.046**	0.946	0.141
rs882763	*TNFRSF11B*	T	8	T	C	0.33	0.356	**0.030**	-0.790
rs2035977	*TNFRSF11B*	P	8	G	C	0.42	**0.015**	**0.100**	0.088
rs2101752	*TNFRSF11B*	P	8	G	T	0.23	**0.056**	**0.027**	0.925
rs7839059	*TNFRSF11B*	P	8	C	A	0.33	**0.036**	0.208	0.074

[1] Effect of each added minor allele on BMC (g) at skeletal sites with significant association; T/P – tag/previously associated; Chr – chromosome; MAF – minor allele frequency; † A significant *P*-value of 0.038 was attained after correction

Allele frequencies of significantly associated SNPs are contrasted across several populations to examine whether the frequency trend correlates with observed interethnic differences in bone mass (**Table 3.3**). In other words, is the frequency of alleles associated with increased BMC higher in African populations? Highlighted rows indicate positive effect alleles that occur at a higher frequency amongst the

present sample, or negative effect alleles that occur at a lower frequency when compared to individuals of European ancestry. Only 4 of 27 markers (14.8%) followed this pattern at a significant level. Frequencies for the two SNPs attaining the lowest *P*-values for association are separately depicted in pie charts in **Figure 3.3**. The minor allele for rs2485209 is relatively common (>40%) in all populations with the exception of the Yoruba population ($P<0.001$), showing a highly significant difference between the two African populations. In contrast, the minor allele for rs6910500 appears to occur in a higher proportion of African ancestry individuals (this study and the Yoruba), and is non-polymorphic amongst tested Asian populations. No data were available for the Luhya or Maasai from Kenya.

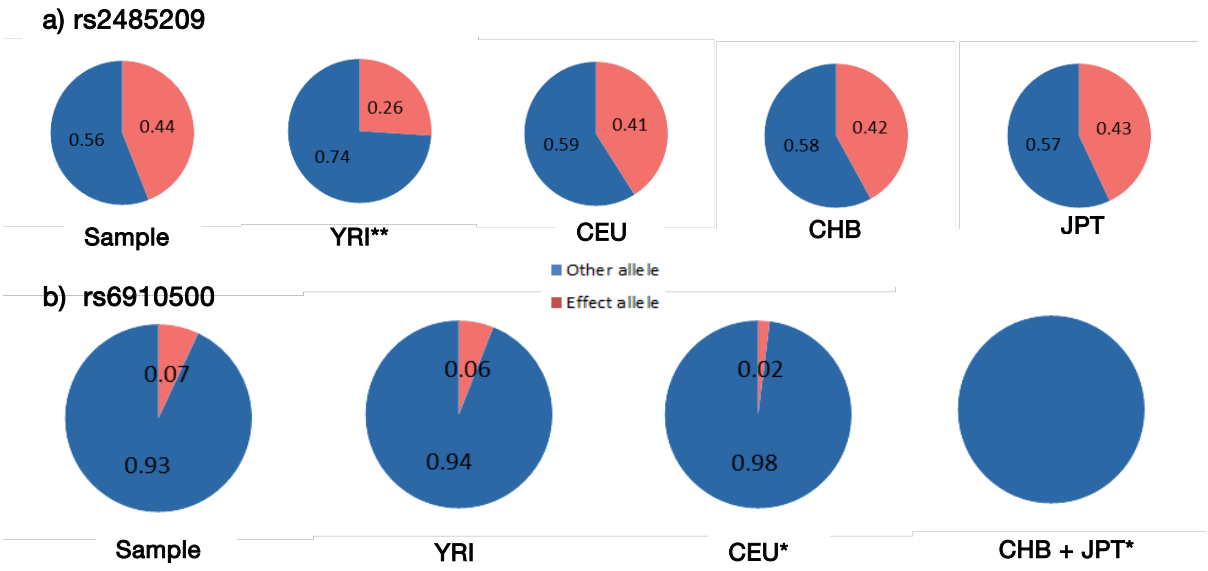


Figure 3.3: Allele frequencies across several populations for rs248520 and rs6910500. a) Alleles of rs248509 appear equally distributed across most tested populations, except for the Yoruba, who possess the effect allele (C) at a significantly smaller proportion. b) The effect allele (A) for rs6910500 is significantly more common amongst African ancestry individuals when compared to Europeans. The locus is non-polymorphic amongst tested Asian populations. **Significantly different from sample allele frequencies below the 1% level; *below the 5% level

Table 3.3: Allele frequencies for associated SNPs across several population groups, with statistical comparisons to YRI and CEU populations

SNP ID	Effect allele	Effect size	Sample	YRI	*P*-value†	CEU	*P*-value†	CHB	JPT	LWK	MKK
rs11155823	T	-0.100	0.16	0.06	0.008	0	<0.001	0.10	0.09		
rs2485209	C	0.112	0.44	0.26	<0.001	0.41	0.465	0.42	0.43	0.36	0.25
rs2504069	G	0.864	0.23	0.20	0.483	0.29	0.218	0.07	0.06		
rs2982551	C	-0.677	0.42	0.31	0.009	0.40	0.606	0.80	0.82	0.43	0.37
rs3020377	A	0.085	0.15	0.23	0.052	0.70	<0.001	0.19	0.21		
rs3020396	A	0.085	0.27	0.20	0.059	0.74	<0.001	0.38	0.36	0.29	0.41
rs4458702	A	-0.129	0.10	0.05	0.070	0.20	0.001	0.45	0.47	0.07	0.05
rs532010	C	-0.065	0.46	0.44	0.557	0.36	0.016	0.39	0.36	0.33	0.45
rs6910500	A	0.174	0.07	0.06	0.665	0.02	0.030	0	0		
rs7744006	A	-0.978	0.13	0.22	0.025	0	<0.001	0	0		
rs9322332	C	-0.073	0.45	0.64	<0.001	0.60	0.008	0.61	0.58		
rs9340785	C	-0.161	0.05	0.12	0.047	0	0.010	0	0		
rs9340820	G	0.095	0.16	0.17	0.883	0.02	<0.001	0.29	0.32	0.28	0.19
rs9340939	A	-0.981	0.13	0.22	0.011	0.01	<0.001	0.03	0	0.16	0.03
rs9371234	A	-0.087	0.17	0.10	0.082	0	<0.001	0	0.06		
rs9383598	T	-0.102	0.14	0.06	0.003	0	<0.001	0.08	0.08	0.04	0.06
rs9397453	A	-0.080	0.22	0.08	<0.001	0	<0.001	0.08	0.08	0.10	0.12
rs2869722	G	0.101	0.15	0.12	0.362	0.33	<0.001	0.26	0.23		
rs1032129	C	0.711	0.33	0.42	0.103	0.32	0.759	0.57	0.65		
rs11573938	A	0.317	0.02	0.06	0.010	0.14	<0.001	0.06	0	0.04	0.07
rs1804854	C	-0.761	0.33	0.28	0.250	0	<0.001	0	0	0.33	0.21
rs4520195	G	-0.977	0.27	0.23	0.318	0	<0.001	0	0	0.28	0.11
rs7464496	T	0.141	0.06	0.06	0.879	0.55	<0.001	0.36	0.22		
rs882763	T	-0.790	0.33	0.29	0.416	0	<0.001	0	0	0.33	0.21
rs2035977	G	0.088	0.42	0.44	0.671	0.57	0.006	0.44	0.38		
rs2101752	G	0.925	0.23	0.21	0.707	0.56	<0.001	0.36	0.25		
rs7839059	C	0.074	0.33	0.35	0.818	0.67	<0.001	0.80	0.76		

†*P*-value for difference from the present study sample allele frequency; YRI – Yoruba from Ibadan, Nigeria; CEU – Utah residents with Northern and Western European ancestry; CHB – Han Chinese from Beijing, China; JPT – Japanese from Tokyo, Japan; LWK – Luhya from Webuye, Kenya; MKK – Maasai from Kinyawa, Kenya

The results from this study were compared with those obtained from the GEFOS meta-analysis (**Table 3.4**) (Richards et al., 2009). Interestingly, only four of the eight SNPs achieving significant *P*-values in both studies were associated with the same skeletal site in both studies. Five markers were observed to have significantly different allele frequencies, including rs2869722, rs7464496, rs2035877, rs2101752

and rs7939059. Effect sizes are not listed, as the unit of measurement differed between studies (BMC versus BMD) as well as the average age of participants (10.5 years old versus 49+ years old).

Table 3.4: Comparison between present study results and those from the GEFOS meta-analysis

		South Africa			GEFOS meta-analysis (Richards et al., 2009)		
SNP ID	Gene	Allele	*P*-value	Skeletal site	Allele	*P*-value*	Skeletal site
rs2485209	*ESR1*	C	0.0020	FN	C	3.20E-07	LS
rs2504069	*ESR1*	G	0.0280	LS	G	1.10E-08	LS
rs2982551	*ESR1*	C	0.0370	LS	C	2.10E-09	LS
rs2869722	*SPP1*	G	0.0300	FN	G	9.60E-07	LS
rs7464496	*TNFRSF11B*	T	0.0460	FN	C	1.70E-09	LS
rs2035977	*TNFRSF11B*	G	0.0150	FN	C	6.40E-08	LS
rs7839059	*TNFRSF11B*	C	0.0360	FN	A	4.50E-08	FN
rs2101752	*TNFRSF11B*	G	0.0270	LS	T	5.00E-08	LS

**P*-value determined under a fixed effects model

3.4. *ESR1* and *TNFRSF11B*

Due to the number of positive associations, a closer investigation of both *ESR1* and *TNFRSF11B* was conducted. LocusZoom plots are displayed in **Figures 3.4-3.7**, with r^2 measures of linkage disequilibrium calculated based on either Yoruba or European HapMap population data. **Table 3.5** lists the reported functions of all associated SNPs according to data obtained from the UCSC human genome browser (Feb. 2009, GRCh37/hg19 assembly).

Leading association signals for LS and FN BMC within *ESR1* (rs2504069 and rs2485209 respectively) were both located towards the 5' end of the gene, within the 150kb region upstream of the first coding exon that contains numerous transcription start sites. Interestingly, these SNPs shared little linkage disequilibrium with other typed SNPs in close proximity, according to Yoruba

population data. In contrast, numerous SNPs share a moderate level of LD with these two markers amongst Europeans.

With respect to *TNFRSF11B*, the leading signal for FN BMC (rs2035977) was located well downstream of the gene, and displayed limited LD with other typed SNPs based on Yoruba data. Amongst Europeans, however, the marker was in LD with other SNPs that extended across the entire plotted region. Notably, rs4520195, which linked most significantly to LS BMC, was in tight LD (r^2>0.8) with two other associated SNPs, namely rs1804854 and rs882763, towards the 3' end of the gene also within an intron. As rs4520195 is non-polymorphic amongst tested Europeans, no linkage information could be generated for this SNP.

Table 3.5: Functions of significantly associated SNPs

rs number	Gene	Function[1]
rs11155823	*ESR1*	Intron
rs2485209	*ESR1*	Intron
rs2504069	*ESR1*	Intron
rs2982551	*ESR1*	Intron
rs3020377	*ESR1*	Intron
rs3020396	*ESR1*	Intron
rs4458702	*ESR1*	Intron
rs532010	*ESR1*	Intron
rs6910500	*ESR1*	Intron
rs7744006	*ESR1*	Intron
rs9322332	*ESR1*	Intron
rs9340785	*ESR1*	Intron
rs9340820	*ESR1*	Intron
rs9340939	*ESR1*	Intron
rs9371234	*ESR1*	Intron
rs9383598	*ESR1*	Intron
rs9397453	*ESR1*	Intron
rs2869722	*SPP1*	Unknown
rs1032129	*TNFRSF11B*	Intron
rs11573938	*TNFRSF11B*	Intron
rs1804854	*TNFRSF11B*	Coding - Synonymous
rs2035977	*TNFRSF11B*	Unknown
rs2101752	*TNFRSF11B*	Unknown
rs4520195	*TNFRSF11B*	Intron
rs7464496	*TNFRSF11B*	Intron
rs7839059	*TNFRSF11B*	Unknown
rs882763	*TNFRSF11B*	Intron

[1] "Intron" implies the variant is not functional, that is, it does not affect known splice consensus sites

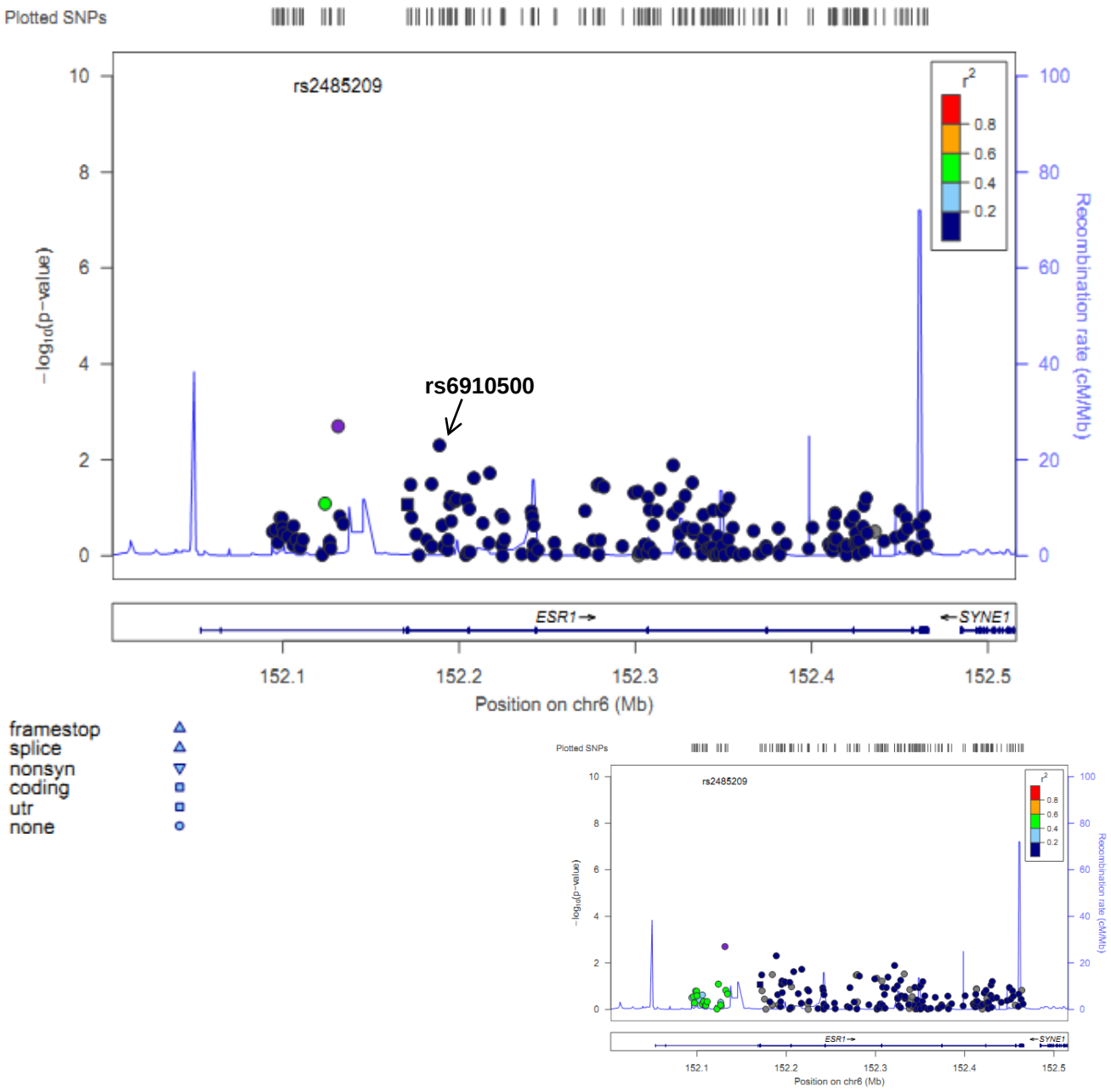


Figure 3.4: LocusZoom plots for femoral neck association results within *ESR1*.The SNP marker with the leading association signal (rs2485209) is coloured purple, with other SNP markers coloured according to the strength of r^2 linkage disequilibrium shared with the leading marker, based on Yoruba population data. rs2485209 maps to the 150kb region upstream of the first coding exon (denoted by a thinner section of blue line representing gene structure), which is replete with non-coding exons that have unique, individual promoters (see **Figure 1.10**). Only one other marker is reported to share moderate (0.4-0.6) linkage with the leading SNP. The smaller, inset diagram shows the same results, but LD is calculated using data from European ancestry individuals. Here, rs2485209 is in moderate LD with numerous other markers that fall between successive recombination peaks (blue line graph). Worthy of mention is rs6910500, which was the only other marker to achieve a *P*-value for association below the 1% level and maps to an intron between coding exons. Marker shapes reflect current knowledge on variant functions.

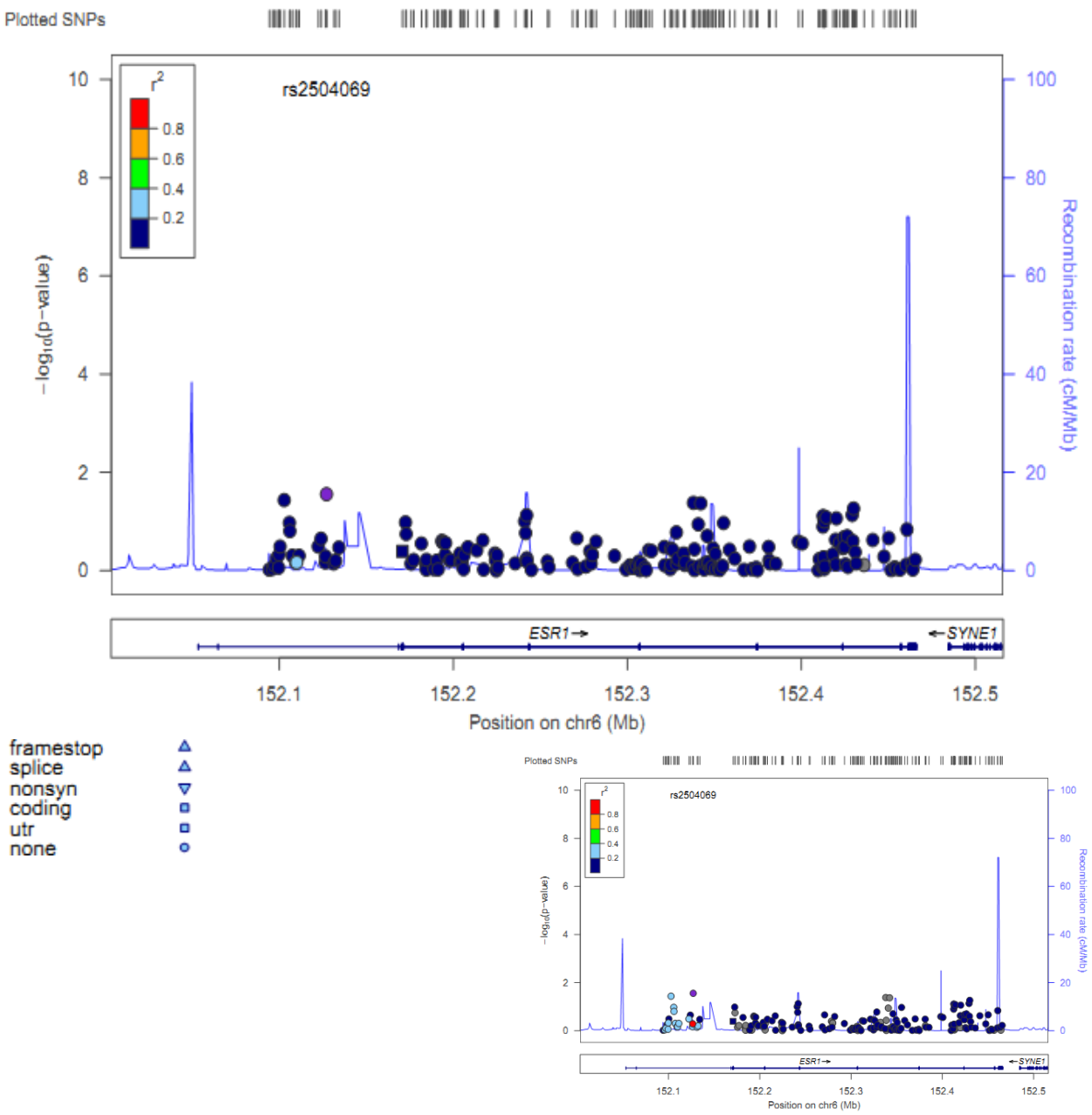


Figure 3.5: LocusZoom plots for lumbar spine association results within *ESR1*. Similar to **Figure 3.4**, the leading marker for association with lumbar spine BMC maps to the 150kb region upstream of the first coding exon. Limited LD was observed on a Yoruba population background, but appeared more extensive amongst European individuals (inset).

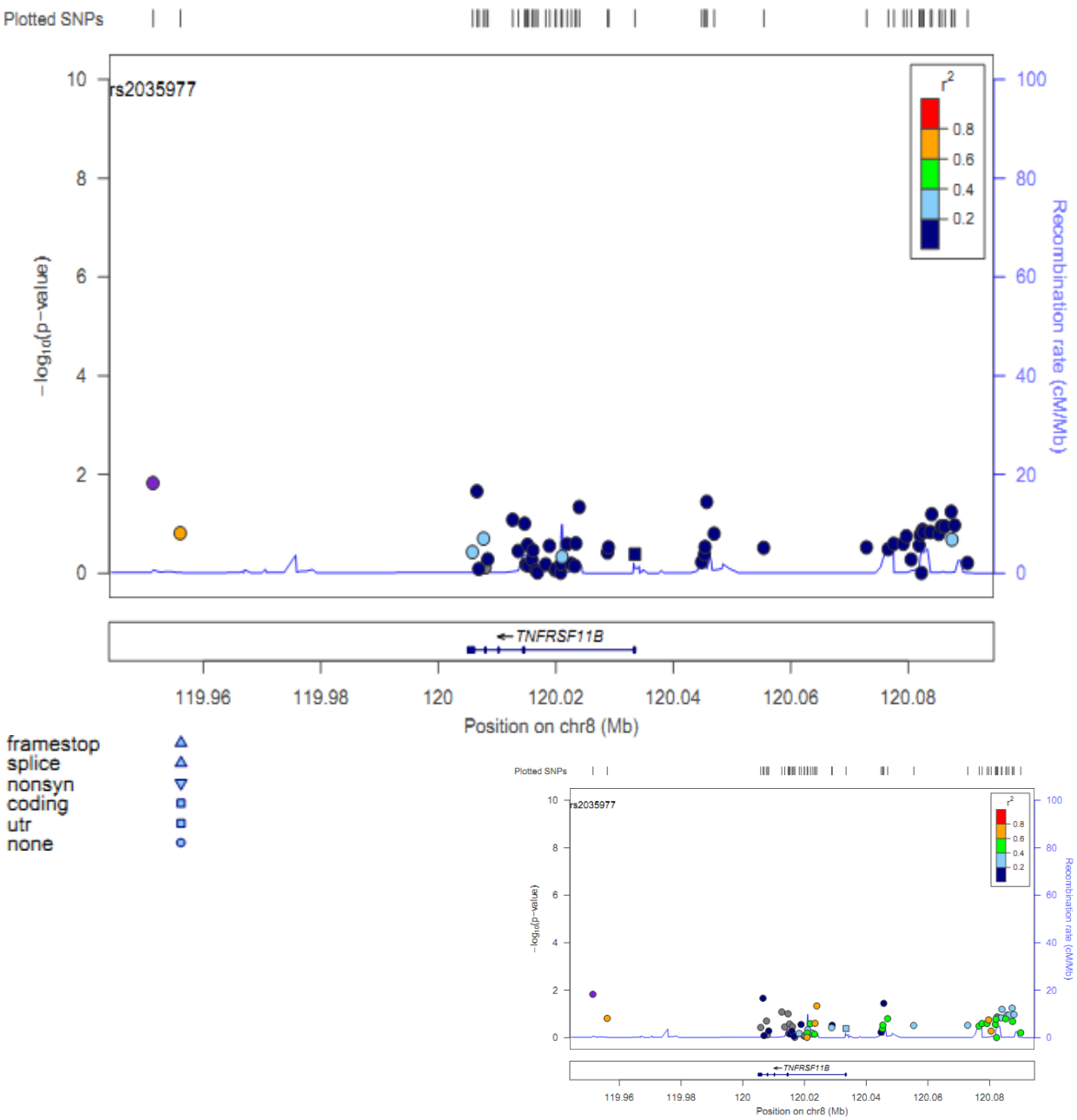


Figure 3.6: LocusZoom plots for femoral neck association results within *TNFRSF11B*. Leading association signal rs2035977 is located well downstream of *TNFRSF11B* which is coded on the minus strand. Moderate to strong (0.6-0.8) LD is shared with a nearby SNP, but only limited LD is detected with three SNPs occurring within the gene interior. In contrast, rs2035977 is in linkage with SNPs across the entire plotted region amongst individuals of European descent (inset). Notably, recombination rate appears minimal for this region.

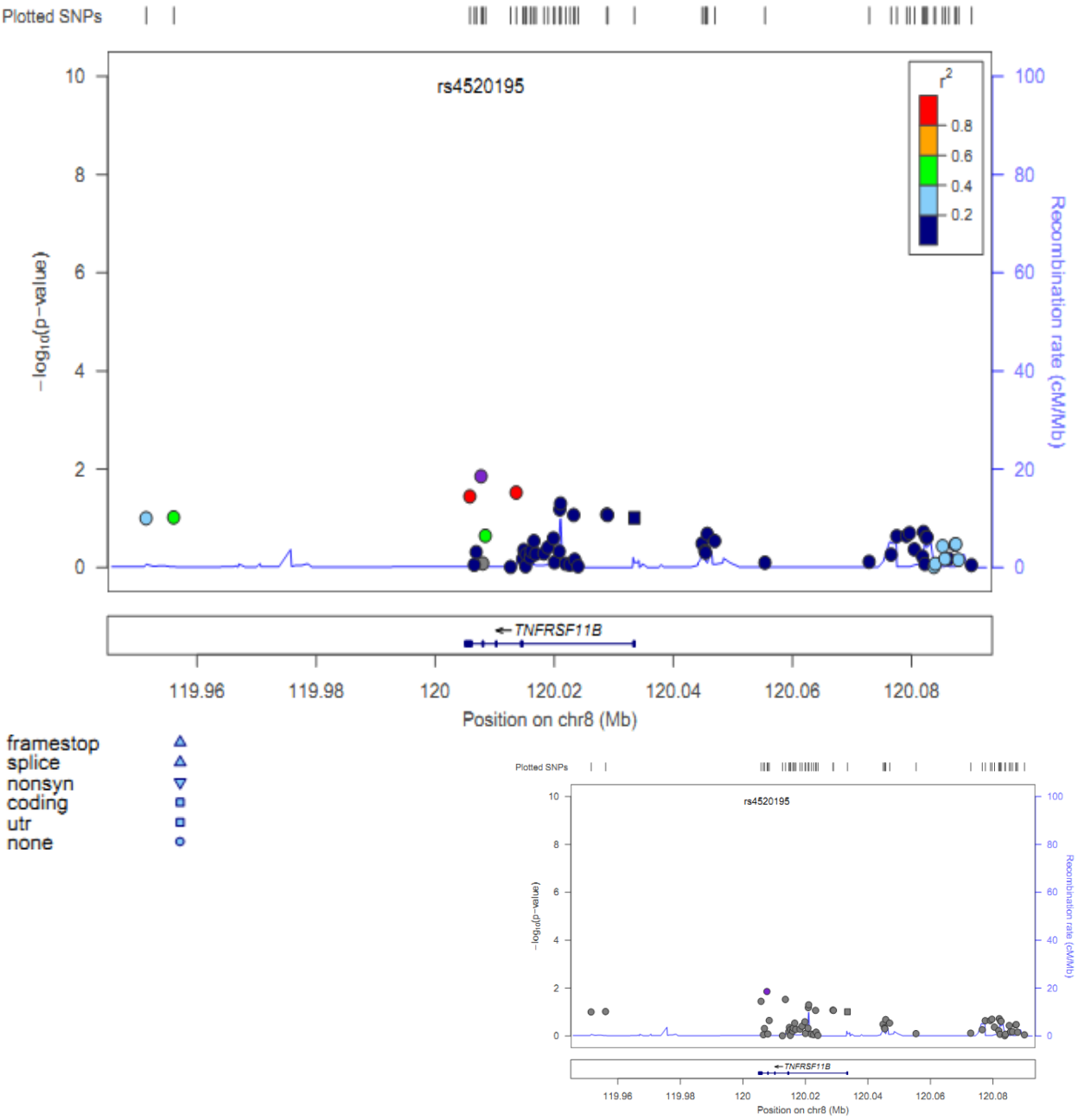


Figure 3.7: LocusZoom plots for lumbar spine association results within *TNFRSF11B*. Leading marker rs4520195 occurs within an intron near the 3' end of the gene. Interestingly, the marker is in strong (0.8-1) LD with two nearby SNPs; both of which are also associated with LS BMC. As rs4520195 is non-polymorphic in tested European populations, no LD data could be generated (inset).

4. Discussion

Genetic factors have been increasingly suspected to provide a powerful protective effect on bone health in black South Africans. Numerous research outputs from the Birth to Twenty bone health sub-cohort have demonstrated higher corrected BMC measurements in specific skeletal regions in black children, when compared to their white counterparts, especially when properly adjusted for covariates (Vidulich et al., 2011). These results have appeared counterintuitive in light of several unfavourable factors affecting the health of black cohort members. Vidulich and colleagues (2011) noted a poor socioeconomic status (median = 7 on a scale of 0 – 13), which was significantly lower than for white children (median = 12; $P<0.05$). Whether or not socioeconomic status has a direct influence on BMC within this cohort has not been properly investigated, but it is assumed that the poor socioeconomic status of blacks would provide a barrier to healthy bone mass acquisition. Similarly, McVeigh and colleagues (2004) observed significant differences in measures of physical activity ($P<0.001$), indicating black children were insufficiently privileged to enjoy a high level of activity that would promote bone mass, compared to their white counterparts in South Africa. MacKeown and colleagues (2003) documented calcium intake levels that were, for 70-90% of black participants, well below the recommended daily allowance across four measurement periods, including at ten years of age. In addition, descriptive statistics for the entire sub-cohort of the Bt20 study have revealed less than favourable anthropometry measurements, with black children being shorter and lighter; a build commonly associated with lower bone mass (Micklesfield et al., 2011; Richter et al., 2007; Vidulich et al., 2011).

This study represented the first investigation into possible genetic factors that would account for BMC measurements for black individuals in this sub-cohort. Without any pre-existing genetic information in non-admixed black populations, a recent meta-analysis by the GEFOS consortium served as a point of reference for selecting candidate genes and pertinent association signals, even though they had been observed through work on Caucasian adults (Richards et al., 2009). Despite numerous sample failures for reasons beyond control, the black individuals that were successfully typed presented with descriptive statistics (**Table 3.1**) that were representative of the entire bone health sub-cohort of the Bt20 study as a whole, and thus remained a suitable sample in which to correlate genotype with phenotype.

4.1. Association results

Using a linear regression technique, associations between genotypes at 329 polymorphic sites and bone mineral content at both the femoral neck and lumbar spine were sought, whilst adjusting for sex, height, weight and bone area, which are common covariates corrected for in similar studies in the literature (Richards et al., 2009; Vidulich et al., 2011). A total of 27 significant signals ($P<0.05$) were detected prior to multiple testing corrections (MTC). One signal, rs2485209, remained significant at the 5% level after MTC, suggesting a strong and truly associated marker. However, all significant associations prior to MTC are considered further in the discussion as the stringency of correction often leads to the loss of potentially true positive associations, which may be important given the small sample size of this study. Detailed discussion is largely reserved for only the most highly significant SNPs.

Observed association signals aligned well with current knowledge on bone health for black members of the Bt20 bone health sub-cohort. The majority of the associated markers (18 of 27) showed association with femoral neck BMC; the skeletal site at which interethnic bone differences have been most clearly seen within this cohort, as well as in other studies on black South Africans (Chantler et al., 2011; Daniels et al., 1995). Suitable adjustments for several different combinations of covariates such as age, body size and bone volume have all served to highlight significant differences at the femoral neck (McVeigh et al., 2004; Micklesfield et al., 2011; Micklesfield et al., 2009; Vidulich et al., 2011), and thus it appears fitting that numerous association signals should derive from this site. Nine markers were also notably linked to LS BMC. Although no significant interethnic differences in bone mass at the lumbar spine have been seen in the bone health sub-cohort of the Bt20 and other studies (Daniels et al., 1995; Vidulich et al., 2006), a more recent adjustment for bone mass by bone area raised to a population specific power coefficient has uncovered increases as large as 85% in girls and 34% in boys, in favour of black children (Vidulich et al., 2011). With no SNP marker associated to both skeletal sites, it is possible that those markers linked to FN BMC are specifically associated with cortical bone formation, whilst those linked to LS BMC are associated to trabecular bone.

Similarly appropriate was the recurrence of both *ESR1* and *TNFRSF11B* within the association results. As mentioned previously, *ESR1* has been linked to bone mineral density in four different study designs, namely candidate gene, genome-wide association, and meta-analytical studies for both approaches (Li et al., 2010). Oestrogen receptor α has the crucial responsibility of mediating the many beneficial effects oestrogen has on bone health in men and women (Compston, 2009), as has been noted in both human and mouse research (**Section 1.5.2.3**). Indeed,

oestrogen replacement therapy aids in preventing bone loss in postmenopausal women (Felson et al., 1993). Some studies have suggested that this gene may have particular importance for bone accretion during younger years (Wang et al., 2007). Of unarguably equal importance is osteoprotegerin (OPG), encoded by *TNFRSF11B*, which acts as a decoy receptor that successfully facilitates the shift from osteoclast resorption to new bone deposition by osteoblasts (**Section 1.1.3**). Mouse model work has clearly demonstrated that OPG overexpression leads to osteopetrosis (Simonet et al., 1997), whilst deficiency encourages early onset osteoporosis (Bucay et al., 1998; Mizuno et al., 1998). Present results have thus served to confirm the major roles these two genes play in the establishment of bone health regardless of ethnicity or age. Both genes are discussed in further detail in the next section.

The lack of prominent association with the four other selected candidate genes (although *SPP1* did generate a single association signal) can be explained by several possible reasons. These genes had been linked to bone phenotypes through research on mostly adult Caucasians. Thus a portion of the resulting association signals may have been unique to either ethnicity or to the age range of the study participants, in line with knowledge that association results are often not transferrable to other ethnicities (**Section 1.3.1**) and that gene sets regulating bone health have been suspected to differ according to age (Pietschmann et al., 2009). Since the present study sample consisted of black children, results may be indicative of a less important role for genes *TNFRSF11A*, *TNFSF11*, *SPP1*, and *SOST* within individuals of African ancestry, or during childhood, or for both reasons. Sample size limitations may have also meant that the study was underpowered to detect modest signals for small effect sizes within these genes.

One objective of this study was to attempt to replicate some of the most notable association signals observed by Richards and colleagues (2009) amongst their meta-analytical sample of 19 195 Caucasian male and female adults. Of 112 such signals, only eight SNPs were found to be significantly associated within this sample. Once again, a lack of replication may be explained by the same reasons outlined above. In addition, meta-analytical approaches are subject to their own set of pitfalls. It is difficult for such studies to fully account for certain biases, such as publication, information and selection biases (Kavvoura and Ioannidis, 2008), thus non-replication may simply be reflective of these problems. Of particular interest, however, was that four of the eight SNPs did not link to the same skeletal site between studies (**Table 3.4**). This may serve to further enforce the idea that gene effects at specific skeletal sites may be substantively different per age and ethnicity. It was also interesting to note that Richards et al. detected a majority of association signals at the lumbar spine (230 SNPs), but fewer at the femoral neck (100 SNPs), whereas the inverse was true for this study, possibly reflecting the importance of the femoral neck during childhood growth, in addition to the strong genetic influence that appears to act on this site amongst individuals of African ancestry. Any further in-depth comparisons between the present study and the GEFOS meta-analysis were avoided due to the sample composition and size differences.

The other main objective of the current study was to search for signals that were possibly unique to this population group by using tag SNPs that had not been associated previously and were based on LD patterns in an African population. Interestingly, twenty signals were generated by these tagging SNP markers. These findings validate the overall reasoning behind the study design, in that they suggest there may indeed be several specific DNA variants that, at least, partially account

for the unmatched bone health not only within this cohort, but potentially in other individuals of African ancestry. This idea is more fully explored below.

4.2. Significant SNP markers

Having discussed association results in general, attention is now turned towards the more specific details of significant SNP markers. The strongest observed signal was generated by rs2485209, which had been previously reported by Richards et al. (2009). However, the associated skeletal site differed (**Table 3.4**), as well as the direction of effect (not shown), possibly implying that this marker is tagging different causal variants between African and European descent populations. Given that this signal maintained significance after MTC even within this small sample size was promising, suggesting a prominent genetic effect. However, no obvious functional roles for rs2485209 or any of the other 26 associated SNP markers were predicted (**Table 3.5**). Whilst a more detailed laboratory and *in silico* experimentation and analysis may be required to adequately confirm a lack of function for these variants, this result largely implies that these markers are in linkage disequilibrium with true causal variants.

Allele frequency differences across several populations were contrasted in **Table 3.3**. Surprisingly, the studied population group was not more likely to possess positive effect alleles at a higher frequency, and negative effect alleles at a lower frequency, than in European ancestry individuals, with only 4 of 27 markers following this pattern with a significant difference in allele frequencies. Although counterintuitive, little meaning is derived from this observation considering there may still be a wealth of other association signals in different candidate genes that need to be factored into the equation. Additionally, as is common with all studies on complex traits (**Section 1.3.1**), these associations account for a tiny proportion

of the heritability of BMC, and thus the unexpected frequency clines fail to generate any immediate concern.

Given their significant *P*-values, allele frequencies for rs2485209 and rs6910500 were separately depicted in **Figure 3.3**. Both alleles of rs2485209 are common (frequency of >0.40) in the study sample, Europeans and Asians, but the allele associated with higher BMC is significantly lower in the Yoruba population ($P < 0.001$). The significance of this observation is likely dependent on the causal variation that is being tagged by this marker, which may be different in separate populations, and thus remains to be seen. Although of small overall effect and low frequency, the higher occurrence of the minor allele for rs6910500 (**Figure 3.3**) amongst the current sample and the Yoruba may represent one minor factor accounting for increased BMC at the femoral neck amongst individuals of African ancestry.

4.3. *ESR1* and *TNFRSF11B*

LocusZoom plots for *ESR1* and *TNFRSF11B* highlighted a number of important observations (**Figures 3.4-3.7**). Both leading association signals for FN and LS BMC within *ESR1* mapped to an intronic region near the 5' end of the gene. It is most significant that two leading association signals should derive from this section of the gene as it is known to be the location of a multitude of transcription start sites (**Figure 1.10**). Variation in this region may well be influencing the binding of transcription factors and thus the whole transcriptional machinery for this gene. In turn, this is likely to have consequences for the expression levels of one or several mRNA transcripts, which would ultimately modify the influence of oestrogen on bone, hence offering a feasible explanation for interethnic bone health differences. Above, it was mentioned that rs2485209 appeared to be tagging different causal

variants between African and European descent populations, based on skeletal site and direction of effect differences. If true, this would further highlight the degree of variation that may be present in this region of the gene and the important ways in which these DNA differences may alter the overall effect on bone health. Furthermore, according to Ensembl release 65 (December 2011), this region itself is also subject to transcription, but as it only consists of non-coding exons, no protein product is translated (**Figure 4.1**). What purpose this transcript has, if any, remains to be clarified. Regardless, the 150kb area upstream of the first coding exon for *ESR1* should be prioritized for future research. Also worth mentioning is rs6910500, which was the only other marker to achieve a significant association below the 1% level. Similarly, this marker maps to an intronic region, but further downstream between coding exons and thus may be tagging a causal variant that affects splicing behaviour.

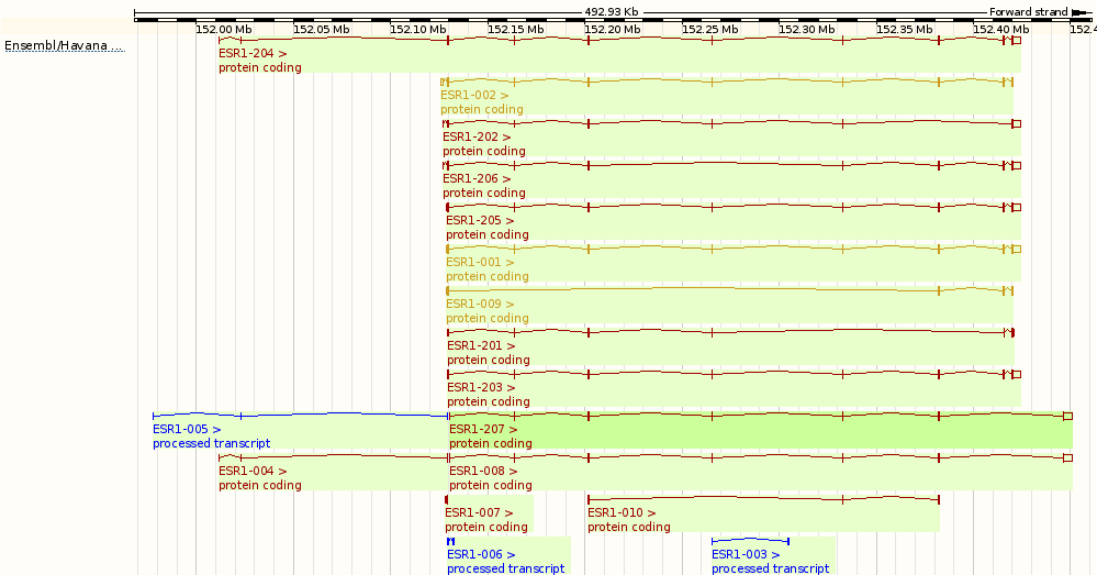


Figure 4.1: The 17 known transcripts of *ESR1*. *ESR1* gene structure begins with several non-coding exons, each with a unique promoter. Transcripts starting from these various promoters usually splice to coding exons that are located further downstream, to create protein-coding mRNA (red and yellow transcripts). Notably, ESR1-005 is a transcript that covers only the non-coding exons and hence lacks an open reading frame. (Ensembl release 65)

Of the markers interrogating *TNFRSF11B*, the strongest association with FN BMC was generated by a SNP downstream of the stop codon. This marker, rs2035977, is currently of unknown function (**Table 3.5**), but may be implicating a nearby regulatory region that moderates *TNFRSF11B* expression. This would further highlight the importance of expression level differences in creating phenotypic diversity, extending to interethnic bone mass differences. Encouragingly, the marker with the greatest affinity to LS BMC (rs4520195) was in strong LD with two other associated SNPs in an intronic region, possibly reflecting their shared inheritance with a single causal variant on the same haplotype background.

LD comparison of Yoruba and European data for highly significant SNPs served to confirm the differences in genomic architecture between these two populations. Amongst Europeans, associated markers were in LD with numerous other surrounding SNPs, reflecting the large LD blocks that are characteristic of these individuals. In contrast, very limited LD was observed when r^2 measurements were based on the Yoruba peoples, in line with the knowledge that African populations have endured more extensive sexual recombination over a longer time period which has eroded LD structure (Teo et al., 2010). This observation has important implications, in that it confirms that identifying causal variation relevant to bone health would be far easier within a non-admixed African sample, as a much smaller area surrounding the associated tagging marker needs to be surveyed.

As previously outlined, the detection of numerous association signals within *ESR1* and *TNFRSF11B* was unsurprising given their functional relevance to bone health. Indeed, these represent two of the most thoroughly investigated and well-linked candidate genes (Li et al., 2010). Whilst present results also highlight the importance of these two genes, associated SNP markers differ to other studies by

some extent. Beyond the eight SNPs replicated from the GEFOS meta-analysis, we failed to find noteworthy overlap with other reported results. For example, rs9340799, rs1999805, rs2073618 and rs6993813 demonstrated no association with BMC in the studied cohort, but have been linked to bone phenotypes elsewhere (Ioannidis et al., 2002; Lee et al., 2010; Li et al., 2010; Liu et al., 2010; Styrkarsdottir et al., 2008). Although this was by no means an extensive investigation, given the volume of available literature, it further enforces the idea that association signals are likely differing in accordance with ethnicity and/or age.

4.4. African genetics

Considering present study results in the wider context of African genetics merits further discussion. As previously highlighted in chapter one, the Bt20 bone health sub-cohort has not been the only study to demonstrate a greater corrected bone mass amongst African ancestry individuals. This has been repeatedly observed in studies incorporating African Americans (Micklesfield et al., 2011), but also less admixed populations including Tobagonian men (Hill et al., 2008) and Somali women (Melton III et al., 2002). Although mostly limited to the femoral neck, a common underlying genetic mechanism is implied through this observation. Most strikingly within the Bt20 study was that this elevated bone mass was maintained in spite of so many unfavourable factors (detailed above) many of which may be shared by other African populations, such as an insufficient level of vitamin D which has also been observed amongst African Americans (Fausto-Sterling, 2008; Thomas, 2007). These conflicting observations potentially speak to the strength of the genetic mechanism underpinning elevated bone mass at the femoral neck in black Africans, but also to the notion of a genetic "set-point" for bone health. It has been previously suggested that interventions taken to improve bone health during childhood are only transient (**Section 1.2.3.4**), with bone mass ultimately returning

to an undefined homeostatic level. Thus, the potential ability of bone-related genetic mechanisms to override environmental influences would justify bone measurements in black individuals, even amongst those that are socioeconomically disadvantaged. Despite offering an alluringly simple explanation, this theory has been largely ignored. Gafni and Baron (2007) cite multiple instances where childhood bone mass accrual was unequivocally stated to influence bone health in later life, without any mention of evidence to the contrary. Results from the present study, however, fall effortlessly in line with this "set-point" theory, given both the extent of detrimental environmental influences as well as the detection of a reasonably strong association signal amongst a small sample size. Further signals from 20 tagging SNPs, although not as strong, suggests the presence of additional DNA variants that may participate in this powerful genetic maintenance of bone health in black individuals.

Notably, not all study results support the idea of a genetic set-point for bone mass, suggesting that the importance of environmental influences cannot yet be side-lined. Certain studies on black individuals in South Africa, the Gambia and one study on recent Sudanese immigrants to the USA have shown that bone mass at many skeletal sites is not significantly greater than population matched individuals of European descent. Indeed, in some cases, bone measurements were found to be lower in Africans. This may be indicative of population-specific environmental factors that do manage to exert a substantial effect on bone mass, either alone or on a certain genetic background. For example, total body and hip BMD for Sudanese immigrants was found to correlate with milk intake, as well as their length of stay in the United States (Gong et al., 2006b). Additionally, calcium absorption and retention have been shown to be improved over populations of European ancestry in studies on both African American adults and black female

children (Abrams et al., 1995; Nelson et al., 2008). Hence, calcium intakes may be better utilised amongst blacks, potentially undermining conclusions about bone health that are made based on interethnic comparisons of only dietary amounts of calcium. It has been previously noted that factors such as fat distribution, vitamin D intake, hormones, adipokines and insulin levels may also exert there effect on BMC in relation to the underlying genetic architecture (Casazza et al., 2010; Nguyen and Eisman, 2006). Hence, whilst genetic influences are currently suspected to account for the majority of interethnic site-specific bone variance, it appears important to remain mindful of other possible factors.

On balance, it is with some caution that these results should be extrapolated to all individuals of African ancestry. Assuming genetic similarity based on a broad categorisation of ethnicity is a flawed approach, particularly amongst Africans who display such rich genetic diversity (Campbell and Tishkoff, 2010; Ramsay et al., 2011). Whilst some of these findings may be replicated amongst independent samples of black Africans, it is more salient to consider their implication within the context of this Bt20 cohort alone. Documented environmental variables appear to be largely ineffectual at influencing bone health, thus we must currently conclude that a very strong genetic effect maintains bone health within these individuals. Supportive of this idea is the fact that association signals were detected, despite a restrictive sample size, although this may also be reflective of the strength of a cohort study design. Whether this effect continues throughout life or is only present during childhood remains to be investigated. Of similar interest, given the transition to a more urban sedentary lifestyle and diet, would be research into the selective pressures that are maintaining this effect, if any. However, there may still be other, currently unknown, environmental factors or population-specific effects that need to be accounted for.

4.5. Bone health genetics

Beyond African genetics, it is prudent to consider the implications of the present study on bone health genetics in general. As outlined previously, suspicions over the possibility of differential genetic effects per skeletal site at different ages and within varied ethnicities are fully supported here. Similarly, it has been postulated that genetic effects on bone accrual may be easier to unmask at a young age, when these effects are strongest and when environmental influences are less complex (Duren et al., 2007). The ability to demonstrate a reasonably strong genetic signal within a small sample of black children may supply at least some evidence for this claim. As alluded to previously, this may also be demonstrative of the strength of cohort study designs, as opposed to those that involve random sampling from the population (Duncan and Brown, 2008).

Taken in summation, these findings should strongly motivate for bone health research focus to be shifted away from adult Caucasian females. Greater advances in our knowledge stand to be gained from the incorporation of varied ethnicities and population isolates, whose diversity of genetics will yield important insights into overall bone maintenance pathways. A greater range of ages should also be considered in further studies. This wider perspective need also be adopted for pharmacogenetic reasons. Five of the eight markers replicated in this study showed significant allele frequency differences between European ancestry and African ancestry individuals. We also noted significant allele frequency differences between the present study sample and the Yoruba population (for example: rs2485209 and rs9322332), despite shared ethnicity. Failure to account for these genetic differences between population groups may limit the applicability of future medications, which would hamper overall progress towards the improved ability to manage bone health and prevent disease.

One final caveat must be borne in mind when considering all studies on bone health genetics. Currently, there is a lack of consensus on many issues, including, but not limited to: most suitable phenotype to use for genetic association studies assessing bone health; best correction method for bone measurements; pertinent environmental factors and ethnic classification. This has resulted in substantial confusion as is easily evident in the literature. For example, there have been conflicting reports about bone differences between males and females prior to puberty (Guggenbuhl, 2009), as well as between separate ethnicities (Nelson et al., 2008). Lack of replication of associated genetic signals, most likely due to study design differences, is a frequent occurrence (Tranah et al., 2008). Overall, this vastly limits the usefulness of inter-study comparisons and facilitates the arrival at erroneous conclusions. Above discussions should be considered carefully in light of this.

4.6. Study limitations

Several important limitations need to be borne in mind when considering the above findings. Most importantly, the final sample size was lower than anticipated due to machine errors, and this impacted substantially on the overall power to detect association signals. Of the 28 SNP markers that failed quality control, 20 were tag SNPs, causing a loss of coverage in genes *ESR1*, *TNFRSF11B*, *TNFRSF11A* and *TNFSF11*. The remaining eight markers were previously reported associated SNPs and would thus need to be retested in order to validate or refute their applicability to this population group. As with the majority of association studies, only SNP variants were tested, disregarding structural variants such as copy number variants, inversions and deletions. Similarly, only six candidate genes were investigated, leaving nine other confirmed candidate genes and many other potential candidates unexplored. All associations were tested under an additive

genetic model, thus missing out on possible dominant or recessive effects. Although previous unpublished work on 18 ancestry informative markers failed to demonstrate significant population sub-structure within this sample, a lack of understanding of more subtle population stratification in this sample may still have influenced the results. In general, candidate gene association study designs are criticised for non-replication and a lack of thorough gene examination, which are two issues this study is also susceptible to (Tabor et al., 2002). Lastly, exclusions were made where gross deviations in HWE were observed, as this may be suggestive of genotyping error, although some have argued that such markers may in fact play a pertinent role in associations (Balding, 2006; Lunetta, 2008).

4.7. Future studies

There are a number of suggestions for future research that are brought to light through the present study. Foremostly, these results should be replicated within an independent study sample. Thereafter, a larger and better powered investigation that includes additional candidate genes should be the main priority for this population group. The Bt20 study has the additional advantage of having recorded a wealth of lifestyle and other physiological variables which may potentially be incorporated to create a sophisticated study design. *ESR1* and *TNFRSF11B* should be subjected to the greatest level of scrutiny, considering their strong showing in this study. Functional analysis, laboratory experimentation and *in silico* analysis on rs2485209, rs6910500 and other highly associated SNPs is fully warranted, and surrounding regions should also be investigated more thoroughly for the presence of causal variants. Additionally, correction for other combinations of covariates may yield more insightful results, such as bone age, fat and lean mass and especially bone area raised to a population specific power coefficient, given the clear interethnic differences highlighted in previous work when correcting for

this factor. As more information on African genetic variation is generated through efforts such as the 1000 Genomes Project and the International Hapmap Project, it may be worthwhile to consider the potential of rarer variants which may have greater effect sizes (Uitterlinden et al., 2006). The study of other bone phenotypes such as fracture and bone strength may also warrant consideration. Finally, beyond simple correlation between genotype and phenotype, it is highly recommended that other genetic mechanisms are investigated such as epistasis, epigenetics, microRNAs and gene-gene and gene-environment interactions.

5. Conclusion

Although based on a small, exploratory study, these results encouragingly suggest the presence of a number of genetic factors that may aid in explaining observed interethnic differences in bone health. Eight previously reported associations amongst populations of European origin were replicated, pointing to the possible existence of a number shared genetic influences on bone across ethnic boundaries. Twenty tagging SNP markers also generated signals that potentially link to causal variants uniquely accounting for elevated bone measurements within the Bt20 bone health sub-cohort, with possible extrapolation to other individuals of African ancestry.

We have successfully supported the idea that genetic effects may differ per skeletal site according to age and/or ethnicity. Similarly, this study has promoted the use of children in research on bone, in line with previous recommendations, as genetic influences appear stronger and thus easier to detect during this phase of life. Lastly, further evidence for the importance of *ESR1* and *TNFRSF11B* in bone health has been generated. Specifically, the 5' region of *ESR1* in which several transcription start sites are located appears to be of particular relevance to understanding interethnic differences in bone regulation.

Ultimately, these results represent only the first steps in African bone health genetics. It is strongly recommended that research continue on samples of non-admixed indigenous Africans, not only due to the potential of expediting the discovery of causal variants but also to ensure we are suitably equipped to efficiently deal with bone disease within this population.

6. References

Abrams, S. A., O'Brien K, O., Liang, L. K. and Stuff, J. E. (1995). Differences in calcium absorption and kinetics between black and white girls aged 5-16 years. *J Bone Miner Res*, **10** (5), 829-833.

Alving, A. S., Carson, P. E., Flanagan, C. L. and Ickes, C. E. (1956). Enzymatic deficiency in primaquine-sensitive erythrocytes. *Science*, **124** (3220), 484-485.

Bachrach, L. K. (2007). Osteoporosis in children: still a diagnostic challenge. *J Clin Endocrinol Metab*, **92** (6), 2030-2032.

Bachrach, L. K., Hastie, T., Wang, M. C., Narasimhan, B. and Marcus, R. (1999). Bone mineral acquisition in healthy Asian, Hispanic, black, and Caucasian youth: a longitudinal study. *J Clin Endocrinol Metab*, **84** (12), 4702-4712.

Balding, D. J. (2006). A tutorial on statistical methods for population association studies. *Nat Rev Genet*, **7** (10), 781-791.

Barrett-Connor, E., Siris, E. S., Wehren, L. E., Miller, P. D., Abbott, T. A., Berger, M. L., Santora, A. C., et al. (2005). Osteoporosis and fracture risk in women of different ethnic groups. *J Bone Miner Res*, **20** (2), 185-194.

Bartl, R. and Frisch, B. (2004). Osteoporosis: Diagnosis, Prevention, Therapy: A Practical Guide for all Physicians - from Pediatrics to Geriatrics. Springer-Verlag, Berlin, Germany.

Bauer, J. S. and Link, T. M. (2009). Advances in osteoporosis imaging. *Eur J Radiol*, **71** (3), 440-449.

Bianchi, M. L. (2007). Osteoporosis in children and adolescents. *Bone*, **41** (4), 486-495.

Bouguer, P. (1729). Essai d'Optique, sur la gradation de la lumiere. Claude Jombert, Paris, France.

Bourne, G. H. (1972). The Biochemistry and Physiology of Bone. 2nd edition. Volume I. Academic Press, London, England.

Boyce, B. F. and Xing, L. (2007). Biology of RANK, RANKL, and osteoprotegerin. *Arthritis Res Ther*, **9 Suppl 1**, S1.

Bucay, N., Sarosi, I., Dunstan, C. R., Morony, S., Tarpley, J., Capparelli, C., Scully, S., et al. (1998). Osteoprotegerin-deficient mice develop early onset osteoporosis and arterial calcification. *Genes Dev*, **12** (9), 1260-1268.

Buckwalter, J. A., Glimcher, M. J., Cooper, R. R. and Recker, R. (1995). Bone Biology Part I: Structure, Blood Supply, Cells, Matrix, and Mineralization. *J Bone Joint Surg Am*, **77**, 1256-1275.

Campbell, M. C. and Tishkoff, S. A. (2010). The evolution of human genetic and phenotypic variation in Africa. *Curr Biol*, **20** (4), R166-173.

Carlborg, O. and Haley, C. S. (2004). Epistasis: too often neglected in complex trait studies? *Nat Rev Genet*, **5** (8), 618-625.

Casazza, K., Thomas, O., Dulin-Keita, A. and Fernandez, J. R. (2010). Adiposity and genetic admixture, but not race/ethnicity, influence bone mineral content in peripubertal children. *J Bone Miner Metab*, **28** (4), 424-432.

Chantler, S., Dickie, K., Goedecke, J. H., Levitt, N. S., Lambert, E. V., Evans, J., Joffe, Y., et al. (2011). Site-specific differences in bone mineral density in black and white premenopausal South African women. *Osteoporos Int*, **23** (2), 533-542.

Church, D. M., Schneider, V. A., Graves, T., Auger, K., Cunningham, F., Bouk, N., Chen, H. C., et al. (2011). Modernizing reference genome assemblies. *PLoS Biol*, **9** (7), e1001091.

Compston, J. (2009). Clinical and therapeutic aspects of osteoporosis. *Eur J Radiol*, **71** (3), 388-391.

Compston, J. E. (2001). Sex steroids and bone. *Physiol Rev*, **81** (1), 419-447.

Daniels, E. D., Pettifor, J. M., Schnitzler, C. M., Russell, S. W. and Patel, D. N. (1995). Ethnic differences in bone density in female South African nurses. *J Bone Miner Res*, **10** (3), 359-367.

de Bakker, P. I., Yelensky, R., Pe'er, I., Gabriel, S. B., Daly, M. J. and Altshuler, D. (2005). Efficiency and power in genetic association studies. *Nat Genet*, **37** (11), 1217-1223.

de Villiers, T. J. (2009). Bone health and osteoporosis in postmenopausal women. *Best Pract Res Clin Obstet Gynaecol*, **23** (1), 73-85.
Deroo, B. J. and Korach, K. S. (2006). Estrogen receptors and human disease. *J Clin Invest*, **116** (3), 561-570.
Downey, P. A. and Siegel, M. I. (2006). Bone biology and the clinical implications for osteoporosis. *Phys Ther*, **86** (1), 77-91.
Duncan, E. L. and Brown, M. A. (2008). Genetic studies in osteoporosis--the end of the beginning. *Arthritis Res Ther*, **10** (5), 214.
Duncan, E. L. and Brown, M. A. (2010). Mapping genes for osteoporosis--old dogs and new tricks. *Bone*, **46** (5), 1219-1225.
Duren, D. L., Sherwood, R. J., Choh, A. C., Czerwinski, S. A., Chumlea, W. C., Lee, M., Sun, S. S., et al. (2007). Quantitative genetics of cortical bone mass in healthy 10-year-old children from the Fels Longitudinal Study. *Bone*, **40** (2), 464-470.
Fausto-Sterling, A. (2008). The Bare Bones of Race. *Social Studies of Science*, **38** (5), 657-694.
Felson, D. T., Zhang, Y., Hannan, M. T., Kiel, D. P., Wilson, P. W. and Anderson, J. J. (1993). The effect of postmenopausal estrogen therapy on bone density in elderly women. *N Engl J Med*, **329** (16), 1141-1146.
Ferrari, S. (2008). Human genetics of osteoporosis. *Best Pract Res Clin Endocrinol Metab*, **22** (5), 723-735.
Fewtrell, M. S., Williams, J. E., Singhal, A., Murgatroyd, P. R., Fuller, N. and Lucas, A. (2009). Early diet and peak bone mass: 20 year follow-up of a randomized trial of early diet in infants born preterm. *Bone*, **45** (1), 142-149.
Frazer, K. A., Murray, S. S., Schork, N. J. and Topol, E. J. (2009). Human genetic variation and its contribution to complex traits. *Nat Rev Genet*, **10** (4), 241-251.
Gafni, R. I. and Baron, J. (2007). Childhood bone mass acquisition and peak bone mass may not be important determinants of bone mass in late adulthood. *Pediatrics*, **119 Suppl 2**, S131-136.
Gallagher, J. C. and Sai, A. J. (2010). Molecular biology of bone remodeling: implications for new therapeutic targets for osteoporosis. *Maturitas*, **65** (4), 301-307.
Garnero, P., Arden, N. K., Griffiths, G., Delmas, P. D. and Spector, T. D. (1996). Genetic influence on bone turnover in postmenopausal twins. *J Clin Endocrinol Metab*, **81** (1), 140-146.
Gauderman, W. J. and Morrison, J. M. (2006). QUANTO 1.1: A computer program for power and sample calculations for genetic-epidemiology studies. http://hydra.usc.edu/gxe.
Gong, G., Haynatzki, G., Haynatzka, V., Howell, R., Kosoko-Lasaki, S., Fu, Y. X., Yu, F., et al. (2006a). Bone mineral density-affecting genes in Africans. *J Natl Med Assoc*, **98** (7), 1102-1108.
Gong, G., Haynatzki, G., Haynatzka, V., Kosoko-Lasaki, S., Howell, R., Fu, Y. X., Gallagher, J. C., et al. (2006b). Bone mineral density of recent African immigrants in the United States. *J Natl Med Assoc*, **98** (5), 746-752.
Guggenbuhl, P. (2009). Osteoporosis in males and females: Is there really a difference? *Joint Bone Spine*, **76** (6), 595-601.
Guthrie, M. (1948). The classification of the Bantu languages. Oxford University Prss for the International African Institute, London, England.
Han, Z. H., Palnitkar, S., Rao, D. S., Nelson, D. and Parfitt, A. M. (1996). Effect of ethnicity and age or menopause on the structure and geometry of iliac bone. *J Bone Miner Res*, **11** (12), 1967-1975.
Hasselstrom, H., Karlsson, K. M., Hansen, S. E., Gronfeldt, V., Froberg, K. and Andersen, L. B. (2006). Sex differences in bone size and bone mineral density exist before puberty. The Copenhagen School Child Intervention Study (CoSCIS). *Calcif Tissue Int*, **79** (1), 7-14.
Henn, B. M., Gignoux, C. R., Jobin, M., Granka, J. M., Macpherson, J. M., Kidd, J. M., Rodriguez-Botigue, L., et al. (2011). Hunter-gatherer genomic diversity suggests a southern African origin for modern humans. *Proc Natl Acad Sci U S A*, **108** (13), 5154-5162.

Hill, D. D., Cauley, J. A., Sheu, Y., Bunker, C. H., Patrick, A. L., Baker, C. E., Beckles, G. L., et al. (2008). Correlates of bone mineral density in men of African ancestry: the Tobago bone health study. *Osteoporos Int*, **19** (2), 227-234.
Hirschhorn, J. N. and Daly, M. J. (2005). Genome-wide association studies for common diseases and complex traits. *Nat Rev Genet*, **6** (2), 95-108.
Holroyd, C., Cooper, C. and Dennison, E. (2008). Epidemiology of osteoporosis. *Best Pract Res Clin Endocrinol Metab*, **22** (5), 671-685.
Huang, Q. Y. and Kung, A. W. (2006). Genetics of osteoporosis. *Mol Genet Metab*, **88** (4), 295-306.
Ioannidis, J. P., Stavrou, I., Trikalinos, T. A., Zois, C., Brandi, M. L., Gennari, L., Albagha, O., 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.
Jurado, S., Nogues, X., Agueda, L., Garcia-Giralt, N., Urreizti, R., Yoskovitz, G., Perez-Edo, L., et al. (2010). Polymorphisms and haplotypes across the osteoprotegerin gene associated with bone mineral density and osteoporotic fractures. *Osteoporos Int*, **21** (2), 287-296.
Kalkwarf, H. J., Zemel, B. S., Gilsanz, V., Lappe, J. M., Horlick, M., Oberfield, S., Mahboubi, S., et al. (2007). The bone mineral density in childhood study: bone mineral content and density according to age, sex, and race. *J Clin Endocrinol Metab*, **92** (6), 2087-2099.
Kanis, J. A., McCloskey, E. V., Johansson, H., Oden, A., Melton, L. J., 3rd and Khaltaev, N. (2008). A reference standard for the description of osteoporosis. *Bone*, **42** (3), 467-475.
Karasik, D. (2008). Osteoporosis: an evolutionary perspective. *Hum Genet*, **124** (4), 349-356.
Karasik, D., Dupuis, J., Cho, K., Cupples, L. A., Zhou, Y., Kiel, D. P. and Demissie, S. (2010). Refined QTLs of osteoporosis-related traits by linkage analysis with genome-wide SNPs: Framingham SHARe. *Bone*, **46** (4), 1114-1121.
Kavvoura, F. K. and Ioannidis, J. P. (2008). Methods for meta-analysis in genetic association studies: a review of their potential and pitfalls. *Hum Genet*, **123** (1), 1-14.
Kim, K. S., Kim, G. S., Hwang, J. Y., Lee, H. J., Park, M. H., Kim, K. J., Jung, J., et al. (2007). Single nucleotide polymorphisms in bone turnover-related genes in Koreans: ethnic differences in linkage disequilibrium and haplotype. *BMC Med Genet*, **8**, 70.
Kleerekoper, M., Nelson, D. A., Peterson, E. L., Flynn, M. J., Pawluszka, A. S., Jacobsen, G. and Wilson, P. (1994). Reference data for bone mass, calciotropic hormones, and biochemical markers of bone remodeling in older (55-75) postmenopausal white and black women. *J Bone Miner Res*, **9** (8), 1267-1276.
Korach, K. S. (1994). Insights from the study of animals lacking functional estrogen receptor. *Science*, **266** (5190), 1524-1527.
Kung, A. W., Yeung, S. S. and Lau, K. S. (1998). Vitamin D receptor gene polymorphisms and peak bone mass in southern Chinese women. *Bone*, **22** (4), 389-393.
Lambrinoudaki, I. and Kung, A. W. (2001). Absence of high-risk "s" allele associated with osteoporosis at the intronic SP1 binding-site of collagen Ialpha1 gene in Southern Chinese. *J Endocrinol Invest*, **24** (7), 499-502.
Lango Allen, H., Estrada, K., Lettre, G., Berndt, S. I., Weedon, M. N., Rivadeneira, F., Willer, C. J., et al. (2010). Hundreds of variants clustered in genomic loci and biological pathways affect human height. *Nature*, **467** (7317), 832-838.
Lee, H. J., Kim, S. Y., Koh, J. M., Bok, J., Kim, K. J., Kim, K. S., Park, M. H., et al. (2007). Polymorphisms and haplotypes of integrinalpha1 (ITGA1) are associated with bone mineral density and fracture risk in postmenopausal Koreans. *Bone*, **41** (6), 979-986.
Lee, Y. H., Woo, J. H., Choi, S. J., Ji, J. D. and Song, G. G. (2010). Associations between osteoprotegerin polymorphisms and bone mineral density: a meta-analysis. *Mol Biol Rep*, **37** (1), 227-234.

Lei, S. F., Chen, Y., Xiong, D. H., Li, L. M. and Deng, H. W. (2006). Ethnic difference in osteoporosis-related phenotypes and its potential underlying genetic determination. *J Musculoskelet Neuronal Interact*, **6** (1), 36-46.
Lerner, U. H. (2006). Bone Remodeling in Post-menopausal Osteoporosis. *Journal of Dental Research*, **85** (7), 584-595.
Li, W. F., Hou, S. X., Yu, B., Li, M. M., Ferec, C. and Chen, J. M. (2010). Genetics of osteoporosis: accelerating pace in gene identification and validation. *Hum Genet*, **127** (3), 249-285.
Lin, C. H., Yeakley, J. M., McDaniel, T. K. and Shen, R. (2009). Medium- to High-Throughput SNP Genotyping Using VeraCode Microbeads. In P. Bugert (Eds.), *DNA and RNA Profiling in Human Blood: Methods and Protocols*. Volume 496. Humana Press, New York, USA,
Liu, J. M., Zhang, M. J., Zhao, L., Cui, B., Li, Z. B., Zhao, H. Y., Sun, L. H., et al. (2010). Analysis of recently identified osteoporosis susceptibility genes in Han Chinese women. *J Clin Endocrinol Metab*, **95** (9), E112-120.
Livshits, G. (2006). Quantitative genetics of circulating molecules associated with bone metabolism: a review. *J Musculoskelet Neuronal Interact*, **6** (1), 47-61.
Lohman, T., Roche, A. and Martorell, R. (1991). Anthropometric standardization reference manual. Human Kinetics, Champaign, IL.
Looker, A. C., Melton, L. J., 3rd, Harris, T., Borrud, L., Shepherd, J. and McGowan, J. (2009). Age, gender, and race/ethnic differences in total body and subregional bone density. *Osteoporos Int*, **20** (7), 1141-1149.
Looker, A. C., Wahner, H. W., Dunn, W. L., Calvo, M. S., Harris, T. B., Heyse, S. P., Johnston, C. C., Jr., et al. (1998). Updated data on proximal femur bone mineral levels of US adults. *Osteoporos Int*, **8** (5), 468-489.
Lunetta, K. L. (2008). Genetic association studies. *Circulation*, **118** (1), 96-101.
MacKeown, J. M., Cleaton-Jones, P. E. and Norris, S. A. (2003). Nutrient intake among a longitudinal group of urban black South African children at four interceptions between 1995 and 2000. *Nutrition Research*, **23**, 185-197.
Manolio, T. A., Collins, F. S., Cox, N. J., Goldstein, D. B., Hindorff, L. A., Hunter, D. J., McCarthy, M. I., et al. (2009). Finding the missing heritability of complex diseases. *Nature*, **461** (7265), 747-753.
Martin, T. J. and Seeman, E. (2008). Bone remodelling: its local regulation and the emergence of bone fragility. *Best Pract Res Clin Endocrinol Metab*, **22** (5), 701-722.
Massart, F. (2005). Human races and pharmacogenomics of effective bone treatments. *Gynecological Endocrinology*, **20** (1), 36-44.
McCarthy, M. I., Abecasis, G. R., Cardon, L. R., Goldstein, D. B., Little, J., Ioannidis, J. P. and Hirschhorn, J. N. (2008). Genome-wide association studies for complex traits: consensus, uncertainty and challenges. *Nat Rev Genet*, **9** (5), 356-369.
McVeigh, J. A., Norris, S. A., Cameron, N. and Pettifor, J. M. (2004). Associations between physical activity and bone mass in black and white South African children at age 9 yr. *J Appl Physiol*, **97** (3), 1006-1012.
Melton III, L. J., Marquez, M. A., Achenbach, S. J., Tefferi, A., O'Connor, M. K., O'Fallon, W. M. and Riggs, B. L. (2002). Variations in bone density among persons of African heritage. *Osteoporos Int*, **13** (7), 551-559.
Micklesfield, L. K., Norris, S. A. and Pettifor, J. M. (2011). Ethnicity and bone: a South African perspective. *J Bone Miner Metab*, **29** (3), 257-267.
Micklesfield, L. K., Norris, S. A., van der Merwe, L., Lambert, E. V., Beck, T. and Pettifor, J. M. (2009). Comparison of site-specific bone mass indices in South African children of different ethnic groups. *Calcif Tissue Int*, **85** (4), 317-325.
Miller, R. G., Ashar, B. H., Cohen, J., Camp, M., Coombs, C., Johnson, E. and Schneyer, C. R. (2005). Disparities in osteoporosis screening between at-risk African-American and white women. *J Gen Intern Med*, **20** (9), 847-851.
Miller, S. A., Dykes, D. D. and Polesky, H. F. (1988). A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res*, **16** (3), 1215.

Mizuno, A., Amizuka, N., Irie, K., Murakami, A., Fujise, N., Kanno, T., Sato, Y., et al. (1998). Severe osteoporosis in mice lacking osteoclastogenesis inhibitory factor/osteoprotegerin. *Biochem Biophys Res Commun*, **247** (3), 610-615.
Mondy, K., Yarasheski, K., Powderly, W. G., Whyte, M., Claxton, S., DeMarco, D., Hoffmann, M., et al. (2003). Longitudinal evolution of bone mineral density and bone markers in human immunodeficiency virus-infected individuals. *Clin Infect Dis*, **36** (4), 482-490.
Moore, K. L. (1992). Clinically oriented anatomy. 3rd edition. Williams and Wilkins, Baltimore, USA.
Mudano, A. S., Casebeer, L., Patino, F., Allison, J. J., Weissman, N. W., Kiefe, C. I., Person, S., et al. (2003). Racial Disparities in Osteoporosis Prevention in a Managed Care Population. *Southern Medical Journal*, **96** (5), 455-451.
Nanes, M. S. and Kallen, C. B. (2009). Clinical assessment of fracture risk and novel therapeutic strategies to combat osteoporosis. *Fertil Steril*, **92** (2), 403-412.
National Osteoporosis Foundation (2002). America's bone health: The state of osteoporosis and low bone mass in our nation. National Osteoporosis Foundation, Washington, DC.
Nelson, D. A., Pettifor, J. M. and Norris, S. (2008). Chapter 26: Race, Ethnicity and Osteoporosis. In R. Marcus, D. Feldman, D. A. Nelson and C. J. Rosen (Eds.), *Osteoporosis*. 3rd edition. Volume I. Elsevier Academic Press, London, UK,
Nelson, D. A., Vande Vord, P. J. and Wooley, P. H. (2000). Polymorphism in the vitamin D receptor gene and bone mass in African-American and white mothers and children: a preliminary report. *Ann Rheum Dis*, **59** (8), 626-630.
Ng, M. Y., Sham, P. C., Paterson, A. D., Chan, V. and Kung, A. W. (2006). Effect of environmental factors and gender on the heritability of bone mineral density and bone size. *Ann Hum Genet*, **70** (Pt 4), 428-438.
Nguyen, T. V. and Eisman, J. A. (2006). Pharmacogenomics of osteoporosis: opportunities and challenges. *J Musculoskelet Neuronal Interact*, **6** (1), 62-72.
Nguyen, T. V., Howard, G. M., Kelly, P. J. and Eisman, J. A. (1998). Bone mass, lean mass, and fat mass: same genes or same environments? *Am J Epidemiol*, **147** (1), 3-16.
Oliveira, E., Pereira, R., Amorim, A., McLeod, H. and Prata, M. J. (2009). Patterns of pharmacogenetic diversity in African populations: role of ancient and recent history. *Pharmacogenomics*, **10** (9), 1413-1422.
Paradowska, A. and Lacki, J. K. (2007). Genetic aspects of osteoporosis. *Centr Eur J Immunol*, **32** (3), 172-180.
Parra, E. J., Marcini, A., Akey, J., Martinson, J., Batzer, M. A., Cooper, R., Forrester, T., et al. (1998). Estimating African American admixture proportions by use of population-specific alleles. *Am J Hum Genet*, **63** (6), 1839-1851.
Peacock, M., Turner, C. H., Econs, M. J. and Foroud, T. (2002). Genetics of osteoporosis. *Endocr Rev*, **23** (3), 303-326.
Pederson, L., Ruan, M., Westendorf, J. J., Khosla, S. and Oursler, M. J. (2008). Regulation of bone formation by osteoclasts involves Wnt/BMP signaling and the chemokine sphingosine-1-phosphate. *Proc Natl Acad Sci U S A*, **105** (52), 20764-20769.
Pietschmann, P., Rauner, M., Sipos, W. and Kerschan-Schindl, K. (2009). Osteoporosis: an age-related and gender-specific disease--a mini-review. *Gerontology*, **55** (1), 3-12.
Pruim, R. J., Welch, R. P., Sanna, S., Teslovich, T. M., Chines, P. S., Gliedt, T. P., Boehnke, M., et al. (2010). LocusZoom: regional visualization of genome-wide association scan results. *Bioinformatics*, **26** (18), 2336-2337.
Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D., Maller, J., et al. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*, **81** (3), 559-575.
Ralston, S. H. and de Crombrugghe, B. (2006). Genetic regulation of bone mass and susceptibility to osteoporosis. *Genes Dev*, **20** (18), 2492-2506.
Ramsay, M., Tiemessen, C. T., Choudhury, A. and Soodyall, H. (2011). Africa: the next frontier for human disease gene discovery? *Hum Mol Genet*, **20** (R2), R214-220.
Randolph, J. F., Jr., Sowers, M., Gold, E. B., Mohr, B. A., Luborsky, J., Santoro, N., McConnell, D. S., et al. (2003). Reproductive hormones in the early menopausal

transition: relationship to ethnicity, body size, and menopausal status. *J Clin Endocrinol Metab*, **88** (4), 1516-1522.
Richards, J. B., Kavvoura, F. K., Rivadeneira, F., Styrkarsdottir, U., Estrada, K., Halldorsson, B. V., Hsu, Y. H., et al. (2009). Collaborative meta-analysis: associations of 150 candidate genes with osteoporosis and osteoporotic fracture. *Ann Intern Med*, **151** (8), 528-537.
Richter, L., Norris, S., Pettifor, J., Yach, D. and Cameron, N. (2007). Cohort Profile: Mandela's children: the 1990 Birth to Twenty study in South Africa. *Int J Epidemiol*, **36** (3), 504-511.
Rivadeneira, F., Styrkarsdottir, U., Estrada, K., Halldorsson, B. V., Hsu, Y. H., Richards, J. B., Zillikens, M. C., 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.
Rizzoli, R., Bianchi, M. L., Garabedian, M., McKay, H. A. and Moreno, L. A. (2010). Maximizing bone mineral mass gain during growth for the prevention of fractures in the adolescents and the elderly. *Bone*, **46** (2), 294-305.
Rizzoli, R., Bonjour, J. P. and Ferrari, S. L. (2001). Osteoporosis, genetics and hormones. *J Mol Endocrinol*, **26** (2), 79-94.
Rosenberg, N. A., Huang, L., Jewett, E. M., Szpiech, Z. A., Jankovic, I. and Boehnke, M. (2010). Genome-wide association studies in diverse populations. *Nat Rev Genet*, **11** (5), 356-366.
Schnitzler, C. M. and Mesquita, J. M. (2006). Cortical bone histomorphometry of the iliac crest in normal black and white South African adults. *Calcif Tissue Int*, **79** (6), 373-382.
Schnitzler, C. M., Mesquita, J. M. and Pettifor, J. M. (2009). Cortical bone development in black and white South African children: iliac crest histomorphometry. *Bone*, **44** (4), 603-611.
Seremak-Mrozikiewicz, A., Barlik, M., Drews, K., Bogacz, A., Tatusko, J., Piotrowska, A., Spaczynski, M., et al. (2011). The genetic variants of the RANKL/RANK/OPG signla trial in postmenopausal women with osteopenia and osteoporosis. *Archives of the Perinatal Medicine*, **17** (2), 72-80.
Shisana, O., Rehle, T., Simbayi, L. C., Zuma, K., Jooste, S., Pillay-van-Wyk, V., Mbelle, N., et al. (2009). South African national HIV prevalence, incidence, behaviour and communication survey 2008: A turning tide among teenagers? HSRC Press, Cape Town.
Simonet, W. S., Lacey, D. L., Dunstan, C. R., Kelley, M., Chang, M. S., Luthy, R., Nguyen, H. Q., et al. (1997). Osteoprotegerin: a novel secreted protein involved in the regulation of bone density. *Cell*, **89** (2), 309-319.
Smedley, D., Haider, S., Ballester, B., Holland, R., London, D., Thorisson, G. and Kasprzyk, A. (2009). BioMart--biological queries made easy. *BMC Genomics*, **10**, 22.
Smith, E. P., Boyd, J., Frank, G. R., Takahashi, H., Cohen, R. M., Specker, B., Williams, T. C., 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.
Spence, A. P. (1990). Basic Human Anatomy. 3rd edition. The Benjamin-Cummings Publishing Company, California, USA.
Stevens, T. A. and Meech, R. (2006). BARX2 and estrogen receptor-alpha (ESR1) coordinately regulate the production of alternatively spliced ESR1 isoforms and control breast cancer cell growth and invasion. *Oncogene*, **25** (39), 5426-5435.
Styrkarsdottir, U., Cazier, J. B., Kong, A., Rolfsson, O., Larsen, H., Bjarnadottir, E., Johannsdottir, V. D., et al. (2003). Linkage of osteoporosis to chromosome 20p12 and association to BMP2. *PLoS Biol*, **1** (3), E69.
Styrkarsdottir, U., Halldorsson, B. V., Gretarsdottir, S., Gudbjartsson, D. F., Walters, G. B., Ingvarsson, T., Jonsdottir, T., et al. (2008). Multiple genetic loci for bone mineral density and fractures. *N Engl J Med*, **358** (22), 2355-2365.
Suarez-Kurtz, G. (2005). Pharmacogenomics in admixed populations. *Trends Pharmacol Sci*, **26** (4), 196-201.

Tabor, H. K., Risch, N. J. and Myers, R. M. (2002). Candidate-gene approaches for studying complex genetic traits: practical considerations. *Nat Rev Genet*, **3** (5), 391-397.
Tanriover, M. D., Oz, S. G., Tanriover, A., Kilicarslan, A., Turkmen, E., Guven, G. S., Saracbasi, O., et al. (2010). Hip fractures in a developing country: osteoporosis frequency, predisposing factors and treatment costs. *Arch Gerontol Geriatr*, **50** (3), e13-18.
Teo, Y. Y., Small, K. S. and Kwiatkowski, D. P. (2010). Methodological challenges of genome-wide association analysis in Africa. *Nat Rev Genet*, **11** (2), 149-160.
Thandrayen, K., Norris, S. A. and Pettifor, J. M. (2009). Fracture rates in urban South African children of different ethnic origins: the Birth to Twenty cohort. *Osteoporos Int*, **20** (1), 47-52.
Theoleyre, S., Wittrant, Y., Tat, S. K., Fortun, Y., Redini, F. and Heymann, D. (2004). The molecular triad OPG/RANK/RANKL: involvement in the orchestration of pathophysiological bone remodeling. *Cytokine Growth Factor Rev*, **15** (6), 457-475.
Thomas, P. A. (2007). Racial and ethnic differences in osteoporosis. *J Am Acad Orthop Surg*, **15 Suppl 1**, S26-30.
Tobias, J. H., Cook, D. G., Chambers, T. J. and Dalzell, N. (1994). A comparison of bone mineral density between Caucasian, Asian and Afro-Caribbean women. *Clin Sci (Lond)*, **87** (5), 587-591.
Tranah, G. J., Taylor, B. C., Lui, L. Y., Zmuda, J. M., Cauley, J. A., Ensrud, K. E., Hillier, T. A., et al. (2008). Genetic variation in candidate osteoporosis genes, bone mineral density, and fracture risk: the study of osteoporotic fractures. *Calcif Tissue Int*, **83** (3), 155-166.
Uitterlinden, A. G., van Meurs, J. B., Rivadeneira, F. and Pols, H. A. (2006). Identifying genetic risk factors for osteoporosis. *J Musculoskelet Neuronal Interact*, **6** (1), 16-26.
van Ijzendoorn, M. H., Bakermans-Kranenburg, M. J., Belsky, J., Beach, S., Brody, G., Dodge, K. A., Greenberg, M., et al. (2011). Gene-by-environment experiments: a new approach to finding the missing heritability. *Nat Rev Genet*, **12** (12), 881; author reply 881.
Vidulich, L., Norris, S. A., Cameron, N. and Pettifor, J. M. (2006). Differences in bone size and bone mass between black and white 10-year-old South African children. *Osteoporos Int*, **17** (3), 433-440.
Vidulich, L., Norris, S. A., Cameron, N. and Pettifor, J. M. (2011). Bone mass and bone size in pre- or early pubertal 10-year-old black and white South African children and their parents. *Calcif Tissue Int*, **88** (4), 281-293.
Wang, C. L., Tang, X. Y., Chen, W. Q., Su, Y. X., Zhang, C. X. and Chen, Y. M. (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, M. C., Aguirre, M., Bhudhikanok, G. S., Kendall, C. G., Kirsch, S., Marcus, R. and Bachrach, L. K. (1997). Bone mass and hip axis length in healthy Asian, black, Hispanic, and white American youths. *J Bone Miner Res*, **12** (11), 1922-1935.
Weickert, C. S., Miranda-Angulo, A. L., Wong, J., Perlman, W. R., Ward, S. E., Radhakrishna, V., Straub, R. E., et al. (2008). Variants in the estrogen receptor alpha gene and its mRNA contribute to risk for schizophrenia. *Hum Mol Genet*, **17** (15), 2293-2309.
Weitzmann, M. N. and Pacifici, R. (2006). Estrogen deficiency and bone loss: an inflammatory tale. *J Clin Invest*, **116** (5), 1186-1194.
WHO Scientific Group (2003). The burden of musculoskeletal conditions at the start of the new millennium. *World Health Organ Tech Rep Ser*, **919**, 1-218.
Wolff, J. (1986). The Law of Bone Remodeling. Springer, Berlin Heidelberg New York.
Woolf, A. D. (2000). The bone and joint decade 2000-2010. *Ann Rheum Dis*, **59** (2), 81-82.
Xiong, D. H., Liu, X. G., Guo, Y. F., Tan, L. J., Wang, L., Sha, B. Y., Tang, Z. H., et al. (2009). Genome-wide association and follow-up replication studies identified ADAMTS18 and TGFBR3 as bone mass candidate genes in different ethnic groups. *Am J Hum Genet*, **84** (3), 388-398.

Web resources

Ensembl (release 65, December 2011) - http://www.ensembl.org/index.html

UCSC human genome browser (Feb. 2009, GRCh37/hg19 assembly) - http://genome.ucsc.edu/cgi-bin/hgGateway

The International HapMap Project (Release 27, Phase II+III, February 2009, on NCBI B36 assembly, dbSNP b126) - http://hapmap.ncbi.nlm.nih.gov/

dbSNP (b126) - http://www.ncbi.nlm.nih.gov/projects/SNP/

LocusZoom (version 1.1) - http://csg.sph.umich.edu/locuszoom/

Online Mendelian Inheritance in Man® - http://www.ncbi.nlm.nih.gov/omim

Appendix A – Ethical clearance

UNIVERSITY OF THE WITWATERSRAND, JOHANNESBURG
Division of the Deputy Registrar (Research)

HUMAN RESEARCH ETHICS COMMITTEE (MEDICAL)
R14/49 Mr Andrew May

CLEARANCE CERTIFICATE	**M10457**
PROJECT	Investigating Genetic Factors Influencing Bone Health in Black South Africans
INVESTIGATORS	Mr Andrew May.
DEPARTMENT	School of Pathology
DATE CONSIDERED	30/04/2010
DECISION OF THE COMMITTEE*	Approved unconditionally

Unless otherwise specified this ethical clearance is valid for 5 years and may be renewed upon application.

DATE 03/05/2010 **CHAIRPERSON** .....
(Professor PE Cleaton-Jones)

*Guidelines for written 'informed consent' attached where applicable
cc: Supervisor : Dr Z Lombard

DECLARATION OF INVESTIGATOR(S)

To be completed in duplicate and **ONE COPY** returned to the Secretary at Room 10004, 10th Floor, Senate House, University.
I/We fully understand the conditions under which I am/we are authorized to carry out the abovementioned research and I/we guarantee to ensure compliance with these conditions. Should any departure to be contemplated from the research procedure as approved I/we undertake to resubmit the protocol to the Committee. **I agree to a completion of a yearly progress report.**

PLEASE QUOTE THE PROTOCOL NUMBER IN ALL ENQUIRIES...

Appendix B – Equipment and software details

List of equipment used

Equipment/Product name	Manufacturer/Supplier	Country
384-plex BeadXpress kit	Illumina	USA
Allegra® X-15R Centrifuge	Beckman Coulter	USA
C1000 Thermal Cycler	Bio-Rad	USA
Centrifuge 5810	Eppendorf	Germany
Freedom EVO®	Tecan	Switzerland
Heat Sealer	Eppendorf	Germany
Hybex Microsample Incubator	SciGene	USA
Millivac™ Maxi SD1P014M04	Millipore	USA
Infinite® 200 PRO NanoQuant	Tecan	Switzerland
TITANIUM™ Taq DNA polymerase	Clontech	USA
Vortemp™ 56 shaking incubator	Labnet	USA

Software

Developer	Product	Version	Country
Gauderman and Morrison, 2006	QUANTO	1.2	-
Illumina	Beadstudio	3.1.3.0	USA
Illumina	Beadstudio genotyping module	3.2.32	USA
SAS Institute Inc.	SAS Enterprise Guide	4.22.0.9238	USA
Shaun Purcell	gPLINK	2.050	-
RJ Pruim et al., 2010	LocusZoom	1.1	-

DNA concentrations for all study participants

UID	BTT ID	BH ID	Concentration (ng/ul)
DNA002496	3054225	104	551.3
DNA002497	3056273	105	259.2
DNA002498	3445333	109	1008.9
DNA002499	3494429	110	618
DNA002505	5976570	123	158.8
DNA002510	6008210	129	221.8
DNA002511	6017210	131	365.5
DNA002512	6026210	133	267.6
DNA002515	6027170	134	171.2
DNA002517	6079402	140	232.7
DNA002522	6094330	144	793.7
DNA002528	6159381	155	239.1
DNA002532	6185357	161	274.8
DNA002533	6224328	164	422.6
DNA002534	6234832	165	308.8
DNA002535	6262248	166	164.1
DNA002537	6267368	167	363.5
DNA002538	6567201	174	288.6
DNA002539	6593593	175	172
DNA002541	7460169	180	387.2
DNA002544	7464529	181	159.5
DNA002546	6108237	145	382.1
DNA002549	7742383	186	228.1
DNA002551		188	625.7
DNA002553	8070656	189	199
DNA002555	8206715	194	989.1
DNA002557	9810267	199	308.8
DNA002559	9907438	201	611.5
DNA002561		203	206.4
DNA002563		207	1146
DNA002565	7480153	209	329.4
DNA002566		349	101.6
DNA002570	3702667	355	945.7
DNA002572	1803288	356	229.5
DNA002573	6021202	358	334.9
DNA002576	5820279	362	317
DNA002577	1059107	363	368.3
DNA002580	1082747	365	1180.1
DNA002583	5258909	370	812
DNA002587	4268386	373	842.8

UID	BTT ID	BH ID	Concentration (ng/ul)
DNA002589	3584104	374	234.1
DNA002592	7585140	375	664.5
DNA002594	5067575	377	293.6
DNA002597	5717273	380	236.3
DNA002599	6366971	383	380.7
DNA002602	2629432	385	247.5
DNA002604	5681665	388	197.8
DNA002606	2111137	389	66
DNA002611	8360569	392	392.5
DNA002613	1032123	395	556.5
DNA002614	1366956	396	705.3
DNA002616	3658483	397	1538.7
DNA002617	9099459	398	84.3
DNA002620	1133166	399	225
DNA002623	1121638	402	343.2
DNA002626	1172246	404	133.4
DNA002628	1288361	407	273.2
DNA002632	1449647	409	426.3
DNA002636	1516007	411	53.7
DNA002638	1611250	413	312.7
DNA002640	1635594	415	329.3
DNA002644	1771776	419	513.5
DNA002646	1802264	420	354.5
DNA002649	1930043	425	253.5
DNA002652	2027598	426	832.6
DNA002653	2039710	427	339.1
DNA002654	2105153	429	312.7
DNA002657	2155201	431	174.4
DNA002660	2216228	434	912.1
DNA002662	2225340	435	285.4
DNA002666		439	985.9
DNA002667	2375375	441	2376
DNA002669	2403855	442	355
DNA002672	2606312	445	528.9
DNA002674	2621757	446	215.9
DNA002675	2644360	448	186.4
DNA002679	2758256	450	152
DNA002681	2845307	452	442.7
DNA002683	2937478	453	669.2
DNA002686	3077345	458	285.1
DNA002691	3476045	461	238.9
DNA002693	3485525	462	246.8
DNA002695	3588200	464	247.5
DNA002698	3693187	467	942.9
DNA002699	3893417	468	285.7
DNA002701	8759549	551	63.7

UID	BTT ID	BH ID	Concentration (ng/ul)
DNA002708	4100692	474	223.6
DNA002710	4194723	475	101.2
DNA002713	4299898	479	272.8
DNA002714	4299899	480	335.9
DNA002718	3914425	647	232.9
DNA002719	4614595	484	434.6
DNA002722	4810465	485	133.1
DNA002724	4929028	488	891.4
DNA002725	5012567	489	153.8
DNA002726	5059703	490	478.9
DNA002728	5082759	491	1.7
DNA002730	5092679	492	256.2
DNA002731	5128818	493	111
DNA002733	5162642	494	168.1
DNA002735	5192866	495	1553
DNA002737	5203370	496	546
DNA002739	5339552	499	150.9
DNA002740	5349032	500	395.7
DNA002742	5367992	501	112.8
DNA002744	5369016	502	576.1
DNA002747	5375424	503	67
DNA002749	5377472	504	307.6
DNA002752	5440443	506	448.4
DNA002756	5455643	508	160.9
DNA002760	5472979	509	204.1
DNA002761	5483483	510	322.7
DNA002762	5489627	511	674.9
DNA002765	5496035	512	314.1
DNA002767	5502179	513	429
DNA002771	5516019	517	174.5
DNA002774	5539075	519	147.7
DNA002776	5575654	521	474
DNA002777	5590254	522	1127.3
DNA002779	5746737	524	521.9
DNA002781	5750833	525	255.6
DNA002782	5885582	526	177.2
DNA002785	5885991	527	227
DNA002789	6498086	530	1041.1
DNA002792	6586161	532	212.9
DNA002794	6635609	533	140.2
DNA002798	6660660	535	392.6
DNA002802	6684740	537	307.5
DNA002803	6760316	538	445.3
DNA002807	6832503	539	60.3
DNA002810	7566180	540	1311
DNA002813		542	252.9

UID	BTT ID	BH ID	Concentration (ng/ul)
DNA002814	7688839	543	204.8
DNA002815	7730855	544	270.2
DNA002818	7956453	546	262.9
DNA002819	8320918	547	244.1
DNA002821	8352377	548	81.2
DNA002824	8640735	550	876.2
DNA002828	8797925	552	245.4
DNA002834	9067739	558	189.7
DNA002835	9097411	559	127.8
DNA002838	9102102	560	308.4
DNA002840	9174318	561	967.3
DNA002845	9256345	565	476.1
DNA002849	9272297	567	345.6
DNA002852	9315256	571	576.5
DNA002854	1211105	573	114.6
DNA002855	6689860	574	164
DNA002856	1008635	575	165.5
DNA002858	1101707	576	533.8
DNA002863	1177366	579	773.2
DNA002865	1191270	580	990.2
DNA002867	1315777	583	285.2
DNA002868	1337492	584	371.4
DNA002870	1361836	585	199.5
DNA002877	1494999	590	342.2
DNA002879	1517031	592	488.9
DNA002880	1753893	594	86.1
DNA002881	1799952	596	511.6
DNA002883	1865288	598	322.8
DNA002886		600	77.5
DNA002887	1956435	601	272.8
DNA002890	2107201	604	268.8
DNA002893	2126761	606	913.9
DNA002897	2312444	609	527.9
DNA002900	2460154	614	241.4
DNA002901	2508578	615	234.6
DNA002903		616	180.1
DNA002904	2583077	617	279.1
DNA002906	2620733	618	1007
DNA002908	2663384	620	311.3
DNA002909	2834803	621	198.7
DNA002910	2872459	622	410.2
DNA002912	2889054	623	347.9
DNA002913	2893414	624	475.7
DNA002915	2918518	625	251.6
DNA002917	3074273	628	113.8
DNA002922	3110600	630	365.1

UID	BTT ID	BH ID	Concentration (ng/ul)
DNA002924	3129300	631	1113
DNA002925	3338354	632	71.1
DNA002929	3399498	633	730.8
DNA002930	3558736	635	354.7
DNA002932	3586152	636	456.5
DNA002933	3601776	638	439.7
DNA002936	3618688	640	103.8
DNA002937	3654296	642	108.3
DNA002940		920	503.7
DNA002941	3821990	644	703.6
DNA002948	3974113	649	457
DNA002950	4016076	650	193.1
DNA002954		653	389.3
DNA002956	4195175	654	592.2
DNA002958	4334429	655	459.3
DNA002961	4578725	658	192.7
DNA002964	4790481	661	696.1
DNA002965	4895468	664	270.1
DNA002967	5058095	665	250.2
DNA002968	5061167	666	102.2
DNA002972	5313477	670	222.7
DNA002975	5376448	671	693.3
DNA002979	5432040	673	121.7
DNA002983		676	246.2
DNA002985	5564126	678	2345.3
DNA002986	5587182	679	124
DNA002988	5832500	681	170.7
DNA002989	6000194	682	72.9
DNA002992	6008650	684	423.8
DNA002994	6022490	685	369.5
DNA002996	6066554	687	310.6
DNA003000	6489630	691	419.8
DNA003001	6554649	692	498.9
DNA003002	6844295	693	274.5
DNA003005	6874519	694	244.7
DNA003006	7793826	696	337.3
DNA003009	7909053	697	2377.1
DNA003011	8287878	698	390.4
DNA003016	8319894	700	547.7
DNA003017	8485540	701	178.2
DNA003018	9081243	703	194.1
DNA003020		999	1266.2
DNA003023	9117246	704	250.4
DNA003025	9441255	705	138.7
DNA003026	9492399	706	1510.6
DNA003028	1596650	707	1049.9

UID	BTT ID	BH ID	Concentration (ng/ul)
DNA003031	1774848	708	347.9
DNA003035	1857856	711	723.6
DNA003036	2050214	713	184.1
DNA003040	2242252	717	588.3
DNA003043	2273764	719	82.6
DNA003047	2799955	722	112.6
DNA003048	2839923	724	414.1
DNA003050	2975398	726	216.6
DNA003052	3493981	727	446.8
DNA003054	3556688	729	431.3
DNA003057	3902897	731	737.2
DNA003058	4021196	732	282.6
DNA003061	4145727	736	207.6
DNA003063	4146761	737	610.5
DNA003067	4365941	740	200.2
DNA003072	5087119	745	448.6
DNA003075	5335456	748	208
DNA003077	5675257	750	1623.5
DNA003079	5869396	751	860.4
DNA003080		752	42.2
DNA003081	5955423	753	440.5
DNA003083	6461454	755	319.1
DNA003085	6518070	756	499
DNA003087	7552604	757	330.2
DNA003089	7982581	759	148.7
DNA003090		760	689.8
DNA003092		761	9.5
DNA003095		765	2021.6
DNA003097		766	295.3
DNA003099		769	149.6
DNA003102		774	353.6
DNA003106		779	686.7
DNA003108		780	374.3
DNA003113		781	259.6
DNA003114		783	258.4
DNA003117		786	188.4
DNA003121		789	876.8
DNA003123		793	419.6
DNA003125		794	921.3
DNA003128		795	1015
DNA003129		796	144.4
DNA003131		800	214.4
DNA003133		802	258.3
DNA003135		803	291.9
DNA003137		807	189.1
DNA003138		808	1003.5

UID	BTT ID	BH ID	Concentration (ng/ul)
DNA003140		814	472.7
DNA003142	8201331	815	268.2
DNA003144		819	183.8
DNA003145		820	394
DNA003147		825	172.3
DNA003148		826	1087.3
DNA003150		827	227.3
DNA003153		829	324.3
DNA003155		832	218.6
DNA003158		833	619.2
DNA003161		834	414.2
DNA003162		835	1085.6
DNA003165		837	759
DNA003167		839	291.6
DNA003170	6250456	840	414.2
DNA003172		842	196.3
DNA003175		843	556.3
DNA003176		846	543.7
DNA003179		850	815.6
DNA003181		853	414.9
DNA003184		855	242.6
DNA003187		859	167.2
DNA003189		860	221.9
DNA003191		862	79.8
DNA003193		865	536.8
DNA003196		869	1143
DNA003200		874	240
DNA003201		880	106.7
DNA003203		882	1053
DNA003205		886	59.3
DNA003208		888	1571
DNA003210		892	467.4
DNA003212		899	513.2
DNA003214		900	543.7
DNA003215		903	116.4
DNA003217		909	445.9
DNA003219		910	2340.1
DNA003225	6022050	921	948.6
DNA003227		927	971.7
DNA003514	7457097	179	138.9
DNA003515		928	241.7
DNA003516		801	170.3
DNA003517	4056044	734	60.6
DNA003519	9325276	568	139.5
DNA003521		916	217.3
DNA003523	6302976	754	356.85

UID	BTT ID	BH ID	Concentration (ng/ul)
DNA003524	7448641	176	396.1
DNA003526	8304270	699	325.9
DNA003529		768	259.1
DNA003530	6075306	139	247.5
DNA003534	9089435	364	182.4
DNA003535	6646113	534	187.5
DNA003537	3764438	379	204
DNA003539	2343255	387	237.4
DNA003541	2308348	720	130.6
DNA003542	2455034	612	140.1
DNA003543	3096369	459	130.2
DNA003544	4249426	738	165.3
DNA003546	6085330	142	333.6
DNA003548	8739501	352	349.5
DNA003549	6308360	369	227.7
DNA003551	2393351	611	300.6
DNA003554	6163261	158	113.2
DNA003556	1615610	414	135.7
DNA003557	5456067	674	212.5
DNA003560	4469640	361	154
DNA003561	1716237	593	365.2
DNA003562		824	434.8
DNA003565	4044076	472	52.6
DNA003566		767	369.9
DNA003567		762	233.2
DNA003572	4905972	742	222.4
DNA003577	3652248	641	294.9
DNA003578	1616634	384	705.4
DNA003580	6156309	154	171.4
DNA003583	3092637	107	308.3
DNA003584	2246348	718	165.3
DNA003585		868	272.4
DNA003588	1244297	405	284.3
DNA003590	5506539	515	177.9
DNA003591	3473997	460	266.9
DNA003596	6398483	529	128.8
DNA003597	5034015	360	294.2
DNA003598	8826432	553	159.8
DNA003599	5452995	507	208.1
DNA003600		918	256.4
DNA003606	4815849	486	112.6
DNA003609	8397469	549	187.3
DNA003610	2495002	570	399.5
DNA003613	4092236	390	175.7
DNA003618	6038114	376	312.6
DNA003620	6092242	143	268.4

UID	BTT ID	BH ID	Concentration (ng/ul)
DNA003621	5013591	743	277.5
DNA003622	1254801	581	151.2
DNA003626	8890456	554	401.2
DNA003628	1064491	400	555.3
DNA003629	6671164	536	347.4
DNA003630	2140337	430	409.3
DNA003633	6027610	528	323.3
DNA003634	1836435	424	703.8
DNA003635	4509608	657	446.2
DNA003636	5407960	505	318.5
DNA003637	3888033	646	602.2
DNA003638	3467589	634	350.7
DNA003639	5512683	516	534
DNA003640		487	157.4
DNA003641	3822321	368	342.2
DNA003645	3650200	466	256.2
DNA003647	2213865	433	312
DNA003648		784	138.6
DNA003649	4238922	476	379
DNA003651	3612280	465	11.9
DNA003652	5501155	675	159
DNA003653	2012998	602	341.8
DNA003654	9024177	557	235.5
DNA003656	3001473	455	435.7
DNA003659	9385420	569	393.5
DNA003662	2286316	438	335.9
DNA003663	2341855	721	289
DNA003664	3034009	457	326.2
DNA003665	2277860	608	138.2
DNA003666	2556685	444	375.9
DNA003667	4297850	478	238.4
DNA003671	4376445	481	109.9
DNA003672	9233225	563	369.1
DNA003675	4087556	652	41.5
DNA003676	6007626	683	227.2
DNA003679	1124174	577	782.4
DNA003681	2234820	437	58
DNA003682	4721366	660	400.2
DNA003683	1093251	401	472.2
DNA003684		797	373
DNA003686	3029649	456	711.9
DNA003687	1283241	582	140.5
DNA003689	2661336	449	218.1
DNA003690	4015052	118	237.1
DNA003691	1377460	408	84.9
DNA003692	2627141	447	492.3

UID	BTT ID	BH ID	Concentration (ng/ul)
DNA003693	6076034	688	315.6
DNA003694	1826411	422	287.2
DNA003695	3039297	102	443.3
DNA003697	4975404	372	612.1
DNA003698	2332207	610	214.4
DNA003699	3438381	354	102.6
DNA003700	9259417	566	237.3
DNA003701	4020172	212	237.6
DNA003702	5093263	746	506.9
DNA003704	5989666	127	533.8
DNA003709	3579528	112	235
DNA003710		863	284.1
DNA003711	4033572	471	264.4
DNA003712	1164678	578	362.4
DNA002812	7587188	541M	139.9
DNA002971	5087559	668M	352.3
DNA002921	3085321	629M	469.1
DNA003034	1798928	709M	75.1
DNA002873	1443503	586M	108.3

Details of all SNPs selected for this study

SNP rs ID	Gene	T/P	Chr	Position	Strand	Alleles	Reference allele
rs10002782	*SPP1*	P	4	89067649	+	A/G	A
rs10086835	*TNFRSF11B*	P	8	120109362	+	C/T	C
rs10098408	*TNFRSF11B*	P	8	120108902	+	C/T	T
rs1022926	*TNFSF11*	T	13	42056734	+	C/G	C
rs1032128	*TNFRSF11B*	T	8	120020954	+	A/G	G
rs1032129	*TNFRSF11B*	T	8	120021081	+	A/C	A
rs10484862	*ESR1*	T	6	152338654	+	A/G	A
rs10505344	*TNFRSF11B*	T	8	120014727	+	A/G	G
rs10505345	*TNFRSF11B*	T	8	120014911	+	A/G	G
rs10505351	*TNFRSF11B*	P	8	120103179	+	C/T	C
rs1054016	*TNFSF11*	T	13	42080002	+	G/T	G
rs1062577	*ESR1*	T	6	152465598	+	A/T	T
rs10955911	*TNFRSF11B*	T	8	120006922	+	C/T	C
rs10955919	*TNFRSF11B*	P	8	120090099	+	C/G	C
rs10955924	*TNFRSF11B*	P	8	120122524	+	A/C	C
rs11155814	*ESR1*	T	6	152192877	+	A/G	A
rs11155819	*ESR1*	T	6	152241052	+	C/T	T
rs11155823	*ESR1*	T	6	152279828	+	G/T	G
rs1124674	*ESR1*	P	6	152122428	+	G/T	T
rs11573828	*TNFRSF11B*	T	8	120028994	+	C/T	C
rs11573880	*TNFRSF11B*	T	8	120020054	+	G/T	T
rs11573894	*TNFRSF11B*	T	8	120016919	+	A/G	A
rs11573896	*TNFRSF11B*	T	8	120016611	+	A/T	T
rs11573899	*TNFRSF11B*	T	8	120016139	+	A/C	C
rs11573903	*TNFRSF11B*	T	8	120015216	+	A/G	G
rs11573911	*TNFRSF11B*	T	8	120012684	+	A/G	G
rs11573930	*TNFRSF11B*	T	8	120008017	+	C/T	C
rs11573938	*TNFRSF11B*	T	8	120006595	+	C/T	C
rs11664594	*TNFRSF11A*	T	18	58169186	+	A/T	A
rs11872145	*TNFRSF11A*	T	18	58159914	+	A/T	T
rs11873955	*TNFRSF11A*	T	18	58192230	+	A/G	G
rs11876459	*TNFRSF11A*	T	18	58198987	+	A/T	A
rs11877530	*TNFRSF11A*	T	18	58166915	+	A/G	G
rs11964865	*ESR1*	T	6	152352739	+	A/T	A
rs11968373	*ESR1*	T	6	152436095	+	G/T	T
rs11992136	*TNFRSF11B*	P	8	120072908	+	C/G	C

SNP rs ID	Gene	T/P	Chr	Position	Strand	Alleles	Reference allele
rs11995824	*TNFRSF11B*	P	8	120081881	+	C/G	C
rs12154178	*ESR1*	T	6	152292773	+	A/C	C
rs12664544	*ESR1*	T	6	152350666	+	C/G	G
rs12956925	*TNFRSF11A*	T	18	58164620	+	A/G	G
rs12959396	*TNFRSF11A*	T	18	58190289	+	G/T	T
rs12970081	*TNFRSF11A*	T	18	58179362	+	A/G	G
rs13203975	*ESR1*	T	6	152374797	+	A/G	G
rs13219476	*ESR1*	T	6	152367197	+	G/T	T
rs13262276	*TNFRSF11B*	P	8	120103099	+	A/T	T
rs13264172	*TNFRSF11B*	P	8	120082042	+	A/T	T
rs13264791	*TNFRSF11B*	T	8	120082271	+	C/T	T
rs1336981	*ESR1*	P	6	152124062	+	C/T	C
rs13381269	*TNFRSF11A*	T	18	58162786	+	C/T	C
rs1381948	*SPP1*	P	4	89054974	+	C/T	T
rs1415193	*ESR1*	P	6	152134331	+	A/T	T
rs1463090	*SPP1*	P	4	89055940	+	G/T	G
rs1471403	*SPP1*	P	4	88994267	+	C/T	T
rs1485287	*TNFRSF11B*	T	8	120019878	+	C/T	C
rs1485289	*TNFRSF11B*	P	8	120023460	+	A/G	G
rs1485295	*TNFRSF11B*	P	8	120108643	+	A/G	A
rs1485302	*TNFRSF11B*	P	8	120045325	+	A/G	G
rs1485303	*TNFRSF11B*	P	8	120045437	+	A/G	G
rs1485305	*TNFRSF11B*	P	8	120079701	+	A/T	A
rs1485310	*TNFRSF11B*	P	8	120120837	+	A/T	A
rs1485312	*TNFRSF11B*	P	8	120113116	+	C/T	T
rs1493934	*TNFRSF11B*	P	8	119915847	+	C/T	T
rs1514347	*ESR1*	T	6	152271138	+	C/T	T
rs1514348	*ESR1*	T	6	152224008	+	G/T	T
rs1564860	*TNFRSF11B*	P	8	120044851	+	C/T	C
rs1569788	*ESR1*	T	6	152370309	+	C/T	T
rs1643821	*ESR1*	T	6	152225244	+	A/G	G
rs17069898	*TNFRSF11A*	T	18	58180261	+	A/G	A
rs17069901	*TNFRSF11A*	T	18	58182164	+	A/G	A
rs17069902	*TNFRSF11A*	T	18	58182579	+	C/T	C
rs17069904	*TNFRSF11A*	T	18	58183929	+	A/G	G
rs17081740	*ESR1*	T	6	152191360	+	C/T	T
rs17081821	*ESR1*	T	6	152255045	+	G/T	T
rs17082028	*ESR1*	T	6	152381551	+	C/T	T
rs1709182	*ESR1*	T	6	152217050	+	C/T	T
rs1709183	*ESR1*	T	6	152235689	+	C/T	C
rs17720953	*TNFRSF11A*	T	18	58165098	+	A/G	G

SNP rs ID	Gene	T/P	Chr	Position	Strand	Alleles	Reference allele
rs1804854	*TNFRSF11B*	T	8	120005850	+	A/G	A
rs1828720	*SOST*	P	17	39153218	+	A/C	A
rs1856057	*ESR1*	P	6	152109562	+	A/G	A
rs1881107	*SOST*	P	17	39141602	+	A/G	A
rs1884049	*ESR1*	T	6	152329060	+	C/T	T
rs1884051	*ESR1*	T	6	152324972	+	A/G	G
rs1884053	*ESR1*	T	6	152333159	+	C/T	C
rs1884054	*ESR1*	T	6	152333259	+	A/C	C
rs1905779	*TNFRSF11B*	P	8	120110346	+	C/T	C
rs1905783	*TNFRSF11B*	P	8	120084017	+	C/T	C
rs1905784	*TNFRSF11B*	P	8	120083804	+	A/G	G
rs1905786	*TNFRSF11B*	P	8	120020873	+	C/T	T
rs1948467	*TNFRSF11B*	P	8	119905753	+	C/T	T
rs1983490	*SOST*	P	17	39138483	+	A/G	G
rs1999805	*ESR1*	P	6	152110057	+	A/G	G
rs1999807	*ESR1*	P	6	152105892	+	C/G	C
rs2035977	*TNFRSF11B*	P	8	119951468	+	C/G	C
rs2055101	*TNFRSF11B*	P	8	119956104	+	C/T	C
rs2062305	*TNFSF11*	P	13	41950880	+	A/G	G
rs2062375	*TNFRSF11B*	P	8	120046973	+	C/G	G
rs2062376	*TNFRSF11B*	P	8	120079190	+	C/G	C
rs2062377	*TNFRSF11B*	P	8	120076601	+	A/T	T
rs2073617	*TNFRSF11B*	P	8	120033464	+	A/G	G
rs2073618	*TNFRSF11B*	P	8	120033233	+	C/G	G
rs2077647	*ESR1*	T	6	152170770	+	C/T	T
rs2101752	*TNFRSF11B*	P	8	119928856	+	G/T	T
rs2144025	*ESR1*	T	6	152349399	+	C/T	T
rs2148072	*TNFSF11*	T	13	42060063	+	A/G	G
rs2148073	*TNFSF11*	T	13	42061799	+	C/G	G
rs2200287	*TNFSF11*	T	13	42066660	+	A/G	G
rs2207231	*ESR1*	T	6	152371577	+	A/G	A
rs2207232	*ESR1*	T	6	152381981	+	C/T	T
rs2207396	*ESR1*	T	6	152424075	+	A/G	G
rs2228480	*ESR1*	T	6	152461788	+	A/G	G
rs2277439	*TNFSF11*	T	13	42053443	+	A/G	G
rs2324851	*TNFSF11*	T	13	42058930	+	A/G	A
rs2326193	*TNFRSF11B*	P	8	120098890	+	A/C	A
rs2347867	*ESR1*	T	6	152271543	+	A/G	G
rs238223	*TNFSF11*	T	13	42077386	+	A/G	A
rs2474148	*ESR1*	T	6	152454597	+	G/T	G
rs2485209	*ESR1*	P	6	152131461	+	A/C	C

SNP rs ID	Gene	T/P	Chr	Position	Strand	Alleles	Reference allele
rs2504063	*ESR1*	P	6	152132400	+	A/G	A
rs2504069	*ESR1*	P	6	152127210	+	C/T	C
rs2504070	*ESR1*	P	6	152126666	+	A/G	A
rs2504071	*ESR1*	P	6	152126555	+	C/T	T
rs2869722	*SPP1*	P	4	89042645	+	C/T	C
rs2875845	*TNFRSF11B*	T	8	120018949	+	A/G	A
rs2957137	*TNFRSF11A*	P	18	58217868	+	A/T	A
rs2982551	*ESR1*	P	6	152102903	+	G/T	G
rs2982554	*ESR1*	P	6	152099703	+	A/C	A
rs2982556	*ESR1*	P	6	152098535	+	A/G	G
rs2982557	*ESR1*	P	6	152098062	+	A/G	G
rs2982560	*ESR1*	P	6	152097299	+	A/T	A
rs2982561	*ESR1*	P	6	152094345	+	C/T	C
rs2982575	*ESR1*	P	6	152111484	+	A/G	A
rs2982684	*ESR1*	T	6	152347897	+	A/C	A
rs2982708	*ESR1*	T	6	152397913	+	C/T	T
rs2982709	*ESR1*	T	6	152398437	+	C/G	G
rs2982734	*ESR1*	T	6	152409847	+	A/G	G
rs2982896	*ESR1*	T	6	152441186	+	C/T	C
rs3003917	*ESR1*	T	6	152308161	+	A/G	G
rs3003925	*ESR1*	T	6	152326151	+	A/G	G
rs3020300	*ESR1*	P	6	152100537	+	A/T	T
rs3020304	*ESR1*	P	6	152106157	+	C/T	C
rs3020306	*ESR1*	T	6	152107579	+	C/T	C
rs3020314	*ESR1*	T	6	152312365	+	C/T	C
rs3020343	*ESR1*	P	6	152096056	+	C/T	C
rs3020349	*ESR1*	P	6	152099961	+	A/C	C
rs3020366	*ESR1*	T	6	152410451	+	C/T	T
rs3020368	*ESR1*	T	6	152412883	+	C/T	C
rs3020371	*ESR1*	T	6	152425513	+	C/T	C
rs3020375	*ESR1*	T	6	152431661	+	A/C	A
rs3020377	*ESR1*	T	6	152314091	+	A/G	G
rs3020381	*ESR1*	T	6	152453771	+	A/T	T
rs3020396	*ESR1*	T	6	152321571	+	A/G	G
rs3020401	*ESR1*	T	6	152324737	+	A/G	G
rs3020407	*ESR1*	T	6	152348954	+	A/G	G
rs3020410	*ESR1*	T	6	152308070	+	A/C	A
rs3020411	*ESR1*	T	6	152385456	+	A/G	A
rs3020434	*ESR1*	T	6	152400633	+	C/T	C
rs3102724	*TNFRSF11B*	T	8	120015988	+	A/G	G
rs3103969	*TNFRSF11B*	P	8	119928984	+	A/G	A

SNP rs ID	Gene	T/P	Chr	Position	Strand	Alleles	Reference allele
rs3103971	*TNFRSF11B*	P	8	119927307	+	C/T	C
rs3133585	*TNFRSF11B*	P	8	119906332	+	A/G	A
rs3134053	*TNFRSF11B*	T	8	120015322	+	C/T	T
rs3134054	*TNFRSF11B*	P	8	120018292	+	A/G	A
rs3134057	*TNFRSF11B*	P	8	120022649	+	A/G	A
rs3134058	*TNFRSF11B*	P	8	120023289	+	A/G	G
rs3134063	*TNFRSF11B*	P	8	120028838	+	C/T	C
rs346588	*TNFSF11*	T	13	42063468	+	A/G	A
rs346589	*TNFSF11*	T	13	42064284	+	A/G	G
rs346591	*TNFSF11*	T	13	42065474	+	A/T	A
rs3742257	*TNFSF11*	T	13	42071198	+	C/T	T
rs3778080	*ESR1*	T	6	152426929	+	A/C	A
rs3778082	*ESR1*	T	6	152429357	+	A/G	G
rs3778084	*ESR1*	T	6	152429732	+	C/T	T
rs3778093	*ESR1*	T	6	152452872	+	A/G	G
rs3778099	*ESR1*	T	6	152460268	+	C/T	T
rs3798571	*ESR1*	T	6	152426493	+	C/T	T
rs3798573	*ESR1*	T	6	152431055	+	A/G	A
rs3798577	*ESR1*	T	6	152462823	+	C/T	T
rs3798758	*ESR1*	T	6	152463547	+	A/C	C
rs3822990	*ESR1*	T	6	152447658	+	C/T	C
rs3826618	*TNFRSF11A*	T	18	58166522	+	C/T	T
rs3826620	*TNFRSF11A*	T	18	58172484	+	G/T	T
rs3844508	*ESR1*	T	6	152181735	+	A/G	A
rs4263037	*TNFRSF11A*	T	18	58167213	+	A/G	G
rs4303637	*TNFRSF11A*	T	18	58182743	+	C/T	T
rs4307369	*TNFRSF11B*	P	8	120106134	+	A/G	A
rs4335155	*TNFRSF11B*	P	8	120082650	+	A/C	A
rs4354338	*TNFRSF11B*	P	8	120085241	+	A/C	A
rs4401893	*TNFRSF11B*	P	8	120113564	+	C/T	C
rs4424291	*TNFRSF11B*	P	8	120087916	+	A/G	A
rs4424296	*TNFRSF11B*	P	8	120082457	+	C/T	C
rs4426449	*TNFRSF11A*	T	18	58193797	+	C/T	C
rs443188	*TNFSF11*	T	13	42075113	+	A/G	G
rs4458702	*ESR1*	T	6	152282116	+	A/G	A
rs4485469	*TNFRSF11A*	T	18	58150553	+	A/G	G
rs4520195	*TNFRSF11B*	T	8	120007728	+	A/G	A
rs4524034	*TNFRSF11A*	T	18	58178988	+	A/G	A
rs4524035	*TNFRSF11A*	T	18	58179224	+	A/G	A
rs4583998	*ESR1*	T	6	152302361	+	A/G	A
rs4615609	*TNFRSF11B*	P	8	120104674	+	A/G	A

SNP rs ID	Gene	T/P	Chr	Position	Strand	Alleles	Reference allele
rs4792909	*SOST*	P	17	39154350	+	G/T	G
rs4793018	*SOST*	P	17	39136595	+	C/T	C
rs4793019	*SOST*	P	17	39136686	+	G/T	G
rs4870056	*ESR1*	T	6	152203920	+	A/G	G
rs4870057	*ESR1*	T	6	152213591	+	A/G	A
rs4870062	*ESR1*	T	6	152279311	+	G/T	G
rs4876876	*TNFRSF11B*	P	8	120055416	+	C/T	T
rs4941125	*TNFRSF11A*	T	18	58148664	+	A/G	A
rs4941433	*TNFSF11*	T	13	42057135	+	G/T	G
rs4942143	*TNFSF11*	T	13	42052806	+	A/G	G
rs532010	*ESR1*	T	6	152172611	+	A/G	A
rs6469783	*TNFRSF11B*	T	8	120008446	+	C/T	C
rs6469788	*TNFRSF11B*	P	8	120021931	+	A/C	A
rs6469792	*TNFRSF11B*	P	8	120077552	+	C/T	T
rs6469795	*TNFRSF11B*	P	8	120085650	+	A/G	G
rs6469797	*TNFRSF11B*	P	8	120087340	+	C/T	C
rs6469798	*TNFRSF11B*	P	8	120087472	+	C/T	C
rs6469805	*TNFRSF11B*	P	8	120117330	+	C/T	C
rs6557171	*ESR1*	T	6	152276286	+	C/T	T
rs6557193	*ESR1*	T	6	152414365	+	C/T	T
rs6567270	*TNFRSF11A*	T	18	58177985	+	A/T	T
rs6567271	*TNFRSF11A*	T	18	58178129	+	A/T	T
rs6567274	*TNFRSF11A*	T	18	58188780	+	G/T	G
rs6846509	*SPP1*	P	4	89059410	+	C/T	C
rs6901451	*ESR1*	T	6	152412629	+	A/G	G
rs6902725	*ESR1*	T	6	152412561	+	C/T	C
rs6903763	*ESR1*	T	6	152194849	+	A/G	G
rs6909023	*ESR1*	T	6	152195390	+	A/G	G
rs6910500	*ESR1*	T	6	152188978	+	A/G	G
rs6911230	*ESR1*	T	6	152305053	+	A/G	A
rs6912184	*ESR1*	T	6	152301899	+	A/G	G
rs6913408	*ESR1*	T	6	152419805	+	C/T	C
rs6914211	*ESR1*	T	6	152310233	+	A/T	T
rs6915819	*ESR1*	T	6	152305321	+	A/G	A
rs6920483	*ESR1*	T	6	152173014	+	A/G	G
rs6927072	*ESR1*	T	6	152299380	+	G/T	T
rs6930355	*ESR1*	T	6	152413084	+	C/T	T
rs6932902	*ESR1*	T	6	152418217	+	A/G	G
rs6937568	*ESR1*	T	6	152195657	+	C/T	T
rs6940499	*ESR1*	T	6	152419876	+	C/T	T
rs6984675	*TNFRSF11B*	P	8	120114408	+	G/T	T

SNP rs ID	Gene	T/P	Chr	Position	Strand	Alleles	Reference allele
rs6993813	*TNFRSF11B*	P	8	120121419	+	C/T	T
rs6996754	*TNFRSF11B*	P	8	120104051	+	C/T	T
rs7004052	*TNFRSF11B*	P	8	120100210	+	A/T	T
rs7010043	*TNFRSF11B*	P	8	120080512	+	G/T	G
rs7010267	*TNFRSF11B*	T	8	120015837	+	A/C	C
rs7013722	*TNFRSF11B*	P	8	120116802	+	A/T	A
rs7013731	*TNFRSF11B*	P	8	120107546	+	C/T	T
rs7018198	*TNFRSF11B*	P	8	120117214	+	A/G	A
rs7220711	*SOST*	P	17	39145491	+	A/G	A
rs722210	*ESR1*	T	6	152364756	+	G/T	G
rs7231887	*TNFRSF11A*	T	18	58182469	+	A/G	G
rs7235803	*TNFRSF11A*	T	18	58151359	+	A/G	A
rs7236060	*TNFRSF11A*	T	18	58174262	+	A/G	G
rs7237982	*TNFRSF11A*	T	18	58155233	+	A/G	A
rs7239261	*TNFRSF11A*	T	18	58156026	+	A/C	A
rs726281	*ESR1*	T	6	152344271	+	A/G	G
rs728523	*ESR1*	T	6	152344884	+	A/T	T
rs7322806	*TNFSF11*	P	13	41944397	+	A/G	G
rs7327904	*TNFSF11*	P	13	41952094	+	C/T	C
rs7334187	*TNFSF11*	T	13	42061641	+	G/T	T
rs7464496	*TNFRSF11B*	P	8	120024023	+	C/T	T
rs7683960	*SPP1*	P	4	89053478	+	A/G	A
rs768625	*TNFSF11*	P	13	41946491	+	C/T	C
rs7743290	*ESR1*	T	6	152350825	+	G/T	T
rs7744006	*ESR1*	T	6	152338111	+	A/G	G
rs7748205	*ESR1*	T	6	152321384	+	C/T	T
rs7754762	*ESR1*	T	6	152353230	+	A/T	T
rs7755185	*ESR1*	T	6	152381308	+	A/G	A
rs7757956	*ESR1*	T	6	152358833	+	A/T	T
rs7759411	*ESR1*	T	6	152190563	+	C/T	C
rs7759554	*ESR1*	T	6	152338727	+	C/T	C
rs7760755	*ESR1*	T	6	152341394	+	C/T	T
rs7761133	*ESR1*	T	6	152193556	+	C/T	T
rs7761846	*ESR1*	T	6	152254201	+	C/T	T
rs7766762	*ESR1*	T	6	152174302	+	A/G	G
rs7768330	*ESR1*	T	6	152343905	+	A/G	A
rs7772475	*ESR1*	T	6	152268615	+	C/T	C
rs7775047	*ESR1*	T	6	152193886	+	C/G	G
rs7813486	*TNFRSF11B*	P	8	120101706	+	C/T	T
rs7822098	*TNFRSF11B*	P	8	120089310	+	A/T	A
rs7839059	*TNFRSF11B*	P	8	120045723	+	A/C	C

SNP rs ID	Gene	T/P	Chr	Position	Strand	Alleles	Reference allele
rs7983721	*TNFSF11*	P	13	41963232	+	A/G	G
rs7990107	*TNFSF11*	T	13	42075252	+	C/T	C
rs7994018	*TNFSF11*	T	13	42079102	+	C/T	T
rs8002449	*TNFSF11*	T	13	42062541	+	A/G	G
rs8083511	*TNFRSF11A*	T	18	58179635	+	A/C	A
rs8086340	*TNFRSF11A*	T	18	58157958	+	C/G	C
rs8089829	*TNFRSF11A*	T	18	58182884	+	A/G	A
rs8092023	*TNFRSF11A*	T	18	58145764	+	A/T	T
rs8094884	*TNFRSF11A*	T	18	58179108	+	A/G	G
rs827421	*ESR1*	T	6	152198815	+	A/G	G
rs827423	*ESR1*	T	6	152197890	+	A/G	G
rs875625	*TNFSF11*	T	13	42072100	+	A/G	A
rs882763	*TNFRSF11B*	T	8	120013632	+	C/T	C
rs884205	*TNFRSF11A*	P	18	58205837	+	A/C	A
rs900797	*TNFSF11*	T	13	42072489	+	A/C	C
rs922996	*TNFSF11*	T	13	42076671	+	A/G	G
rs926777	*ESR1*	T	6	152346740	+	A/C	A
rs9322331	*ESR1*	T	6	152204010	+	C/T	C
rs9322332	*ESR1*	T	6	152208494	+	A/C	C
rs9322335	*ESR1*	T	6	152241822	+	C/T	T
rs9322337	*ESR1*	T	6	152242332	+	A/C	C
rs9322338	*ESR1*	T	6	152242442	+	A/G	G
rs9322352	*ESR1*	T	6	152413336	+	A/G	G
rs9322356	*ESR1*	T	6	152450352	+	A/G	G
rs9322359	*ESR1*	T	6	152451715	+	A/T	A
rs932477	*ESR1*	T	6	152346289	+	A/G	A
rs9340779	*ESR1*	T	6	152177302	+	C/T	C
rs9340784	*ESR1*	T	6	152184340	+	A/T	T
rs9340785	*ESR1*	T	6	152184552	+	C/G	G
rs9340787	*ESR1*	T	6	152184636	+	A/G	A
rs9340799	*ESR1*	T	6	152205074	+	A/G	A
rs9340807	*ESR1*	T	6	152205979	+	A/T	A
rs9340811	*ESR1*	T	6	152206340	+	A/C	C
rs9340820	*ESR1*	T	6	152217518	+	G/T	T
rs9340831	*ESR1*	T	6	152224306	+	A/G	G
rs9340833	*ESR1*	T	6	152224773	+	A/C	A
rs9340835	*ESR1*	T	6	152241624	+	A/G	G
rs9340838	*ESR1*	T	6	152242202	+	G/T	T
rs9340888	*ESR1*	T	6	152280024	+	A/C	A
rs9340901	*ESR1*	T	6	152306704	+	A/G	G
rs9340933	*ESR1*	T	6	152341662	+	C/G	C

SNP rs ID	Gene	T/P	Chr	Position	Strand	Alleles	Reference allele
rs9340939	*ESR1*	T	6	152342320	+	A/C	C
rs9340941	*ESR1*	T	6	152354839	+	C/T	C
rs9340944	*ESR1*	T	6	152355411	+	A/G	G
rs9340954	*ESR1*	T	6	152361865	+	G/T	T
rs9340971	*ESR1*	T	6	152374211	+	A/G	A
rs9341004	*ESR1*	T	6	152411941	+	C/G	C
rs9341008	*ESR1*	T	6	152413401	+	A/G	G
rs9341009	*ESR1*	T	6	152413550	+	A/C	C
rs9341021	*ESR1*	T	6	152425071	+	C/T	C
rs9341049	*ESR1*	T	6	152456886	+	A/G	A
rs9341062	*ESR1*	T	6	152460772	+	A/G	A
rs9341074	*ESR1*	T	6	152464148	+	A/T	T
rs9371234	*ESR1*	T	6	152332439	+	A/C	C
rs9371564	*ESR1*	T	6	152303806	+	A/G	A
rs9371575	*ESR1*	T	6	152420245	+	C/T	T
rs9383598	*ESR1*	T	6	152301627	+	C/T	C
rs9383599	*ESR1*	T	6	152310910	+	C/G	C
rs9383607	*ESR1*	T	6	152422011	+	A/T	T
rs9383951	*ESR1*	T	6	152337306	+	C/G	G
rs9397453	*ESR1*	T	6	152278572	+	A/G	G
rs9397459	*ESR1*	T	6	152307352	+	A/G	G
rs9397462	*ESR1*	T	6	152340913	+	A/T	T
rs9397463	*ESR1*	T	6	152346021	+	C/T	C
rs9478248	*ESR1*	T	6	152225823	+	A/G	G
rs9478251	*ESR1*	T	6	152244972	+	C/G	C
rs9479120	*ESR1*	T	6	152175760	+	A/G	A
rs9479134	*ESR1*	T	6	152230688	+	A/G	G
rs9479188	*ESR1*	T	6	152430137	+	A/G	G
rs9525643	*TNFSF11*	T	13	42057516	+	C/T	T
rs9525644	*TNFSF11*	T	13	42057804	+	A/G	G
rs9562415	*TNFSF11*	T	13	42079024	+	C/T	T
rs9594766	*TNFSF11*	P	13	41938043	+	A/G	G
rs9646629	*TNFRSF11A*	T	18	58202179	+	C/G	C
rs9650075	*TNFRSF11B*	P	8	120086260	+	A/G	A
rs974277	*ESR1*	T	6	152424514	+	C/T	C
rs985191	*ESR1*	T	6	152325151	+	A/C	A
rs985694	*ESR1*	T	6	152328318	+	C/T	T
rs985695	*ESR1*	T	6	152328398	+	C/T	C
rs9909172	*SOST*	P	17	39143047	+	C/T	T
rs9911277	*SOST*	P	17	39136010	+	A/G	G
rs9951012	*TNFRSF11A*	T	18	58174774	+	A/G	G

SNP rs ID	Gene	T/P	Chr	Position	Strand	Alleles	Reference allele
rs9960450	*TNFRSF11A*	T	18	58167855	+	C/T	T
rs9961156	*TNFRSF11A*	T	18	58150614	+	C/T	T

T/P – tag/previously associated; Chr - chromosome

Association results for all SNPs typed in this study

SNP ID	Gene	T/P	Ref.	A1 (minor)	A2 (major)	MAF (sample)	*P*-value (FN BMC)	*P*-Value (LS BMC)
rs10484862	*ESR1*	T	A	G	A	0.1325	0.1409	0.3706
rs1062577	*ESR1*	T	T	A	T	0.1589	0.5794	0.6044
rs11155814	*ESR1*	T	A	A	G	0.4259	N/A	N/A
rs11155819	*ESR1*	T	T	C	T	0.06333	0.1189	0.0999
rs11155823	*ESR1*	T	G	T	G	0.1556	0.03197	0.4685
rs1124674	*ESR1*	P	T	A	C	0.4139	0.958	0.3247
rs11964865	*ESR1*	T	A	T	A	0.1887	0.4486	0.8941
rs11968373	*ESR1*	T	T	G	T	0.009934	0.3137	0.7742
rs12154178	*ESR1*	T	C	A	C	0.4371	0.6324	0.5045
rs12664544	*ESR1*	T	G	C	G	0.2133	0.968	0.4605
rs13203975	*ESR1*	T	G	A	G	0.2417	0.7303	0.988
rs13219476	*ESR1*	T	T	G	T	0.1225	0.3027	0.969
rs1336981	*ESR1*	P	C	G	A	0.2063	0.08192	0.2258
rs1415193	*ESR1*	P	T	A	T	0.03311	0.2168	0.3417
rs1514347	*ESR1*	T	T	A	G	0.09272	0.8332	0.2208
rs1514348	*ESR1*	T	T	A	C	0.2483	0.1419	0.4515
rs1569788	*ESR1*	T	T	T	C	0.1987	0.9121	0.3247
rs1643821	*ESR1*	T	G	C	T	0.25	0.1618	0.4895
rs17081740	*ESR1*	T	T	C	T	0.0298	0.6474	0.9491
rs17081821	*ESR1*	T	T	G	T	0.2752	0.9331	0.8611
rs17082028	*ESR1*	T	T	C	T	0.149	0.8152	0.6214
rs1709182	*ESR1*	T	T	C	T	0.1633	0.5445	0.2428
rs1709183	*ESR1*	T	C	A	G	0.4272	0.9251	0.7163
rs1856057	*ESR1*	P	A	C	T	0.4305	0.6653	0.6883
rs1884049	*ESR1*	T	T	C	T	0.4933	0.2597	0.5784
rs1884051	*ESR1*	T	G	G	A	0.4834	0.3197	0.3586
rs1884053	*ESR1*	T	C	T	C	0.07285	0.3397	0.7692
rs1884054	*ESR1*	T	C	A	C	0.07333	0.2767	0.6823
rs1999805	*ESR1*	P	G	T	C	0.43	0.6693	0.6873
rs1999807	*ESR1*	P	C	G	C	0.3709	0.5924	0.1069
rs2077647	*ESR1*	T	T	A	G	0.4764	0.08492	0.4066
rs2144025	*ESR1*	T	T	T	C	0.4503	0.6214	0.3566
rs2207231	*ESR1*	T	A	G	A	0.1358	0.8761	0.9421
rs2207232	*ESR1*	T	T	C	T	0.1391	0.9461	0.7183
rs2207396	*ESR1*	T	G	A	G	0.3255	0.1518	0.3497
rs2228480	*ESR1*	T	G	A	G	0.1814	N/A	N/A

SNP ID	Gene	T/P	Ref.	A1 (minor)	A2 (major)	MAF (sample)	*P*-value (FN BMC)	*P*-Value (LS BMC)
rs2347867	*ESR1*	T	G	A	G	0.2351	0.1159	0.9281
rs2474148	*ESR1*	T	G	G	T	0.4558	0.2697	0.9211
rs2485209	*ESR1*	P	C	C	A	0.4404	0.03796[†]	0.7173
rs2504063	*ESR1*	P	A	G	A	0.2119	0.1508	0.6344
rs2504069	*ESR1*	P	C	G	A	0.2315	0.7103	0.02797
rs2504070	*ESR1*	P	A	T	C	0.3767	0.5045	0.5175
rs2504071	*ESR1*	P	T	G	A	0.2417	0.6464	0.6873
rs2982551	*ESR1*	P	G	C	A	0.4205	0.3976	0.03696
rs2982554	*ESR1*	P	A	G	T	0.3179	0.1648	0.8761
rs2982556	*ESR1*	P	G	T	C	0.2947	0.1628	0.5614
rs2982557	*ESR1*	P	G	T	C	0.08108	N/A	N/A
rs2982560	*ESR1*	P	A	A	T	0.3423	0.5295	0.7962
rs2982561	*ESR1*	P	C	A	G	0.2682	0.3137	0.959
rs2982575	*ESR1*	P	A	C	T	0.2033	0.4615	0.5015
rs2982684	*ESR1*	T	A	A	C	0.1678	0.8801	0.9101
rs2982708	*ESR1*	T	T	T	C	0.5	N/A	N/A
rs2982709	*ESR1*	T	G	C	G	0.09333	0.7083	0.2587
rs2982734	*ESR1*	T	G	G	A	0.1486	0.5724	0.986
rs2982896	*ESR1*	T	C	T	C	0.1467	0.4945	0.2398
rs3003917	*ESR1*	T	G	A	G	0.4503	0.1109	0.7333
rs3003925	*ESR1*	T	G	G	A	0.3046	0.6903	0.3786
rs3020300	*ESR1*	P	T	A	T	0.2285	0.3447	0.3257
rs3020304	*ESR1*	P	C	T	C	0.2639	0.2408	0.1578
rs3020306	*ESR1*	T	C	T	C	0.2033	0.4615	0.5015
rs3020314	*ESR1*	T	C	T	C	0.1722	0.1149	0.3876
rs3020343	*ESR1*	P	C	T	C	0.3311	0.2837	0.9221
rs3020349	*ESR1*	P	C	A	C	0.2483	0.2577	0.4745
rs3020366	*ESR1*	T	T	T	C	0.3212	0.6583	0.5874
rs3020368	*ESR1*	T	C	T	C	0.1656	0.4705	0.5385
rs3020371	*ESR1*	T	C	C	T	0.298	0.5095	0.7512
rs3020375	*ESR1*	T	A	A	C	0.1126	0.3377	0.7183
rs3020377	*ESR1*	T	G	A	G	0.1523	0.04096	0.4006
rs3020381	*ESR1*	T	T	T	A	0.2417	0.1598	0.8651
rs3020396	*ESR1*	T	G	A	G	0.2682	0.01299	0.7902
rs3020401	*ESR1*	T	G	A	G	0.3113	0.0969	0.9281
rs3020407	*ESR1*	T	G	A	G	0.06291	0.1159	0.8372
rs3020410	*ESR1*	T	A	A	C	0.2282	0.6494	0.8771
rs3020411	*ESR1*	T	A	A	G	0.1722	0.5714	0.7293
rs3020434	*ESR1*	T	C	T	C	0.1667	0.2597	0.2817
rs3778080	*ESR1*	T	A	C	A	0.1258	0.4835	0.8432
rs3778082	*ESR1*	T	G	A	G	0.4868	0.2478	0.07393

SNP ID	Gene	T/P	Ref.	A1 (minor)	A2 (major)	MAF (sample)	*P*-value (FN BMC)	*P*-Value (LS BMC)
rs3778084	*ESR1*	T	T	T	C	0.4735	0.08991	0.2517
rs3778093	*ESR1*	T	G	A	G	0.1422	N/A	N/A
rs3778099	*ESR1*	T	T	C	T	0.231	0.7453	0.7702
rs3798571	*ESR1*	T	T	T	C	0.3986	0.9321	0.1998
rs3798573	*ESR1*	T	A	G	A	0.4895	0.06294	0.4226
rs3798577	*ESR1*	T	T	C	T	0.3226	N/A	N/A
rs3798758	*ESR1*	T	C	T	G	0.2119	0.3706	0.7193
rs3822990	*ESR1*	T	C	T	C	0.2947	0.4216	0.5225
rs3844508	*ESR1*	T	A	T	C	0.4833	0.4715	0.2827
rs4458702	*ESR1*	T	A	A	G	0.09524	0.03696	0.2577
rs4583998	*ESR1*	T	A	G	A	0.4801	0.8312	0.8661
rs4870056	*ESR1*	T	G	G	A	0.3974	0.965	0.6673
rs4870057	*ESR1*	T	A	G	A	0.2567	0.2108	0.3896
rs4870062	*ESR1*	T	G	G	T	0.447	0.956	0.7223
rs532010	*ESR1*	T	A	C	T	0.4636	0.03297	0.1049
rs6557171	*ESR1*	T	T	T	C	0.404	0.4805	0.8292
rs6557193	*ESR1*	T	T	T	C	0.4933	0.4545	0.08292
rs6901451	*ESR1*	T	G	G	A	0.4967	0.5544	0.07792
rs6902725	*ESR1*	T	C	C	T	0.4934	0.2228	0.1279
rs6903763	*ESR1*	T	G	A	G	0.1921	0.08392	0.2747
rs6909023	*ESR1*	T	G	A	G	0.3212	0.05994	0.4835
rs6910500	*ESR1*	T	G	A	G	0.07285	0.004995	0.9391
rs6911230	*ESR1*	T	A	G	A	0.2789	N/A	N/A
rs6912184	*ESR1*	T	G	G	A	0.1755	0.982	0.8032
rs6913408	*ESR1*	T	C	T	C	0.2781	0.5944	0.2368
rs6914211	*ESR1*	T	T	A	T	0.1656	0.2288	0.8791
rs6915819	*ESR1*	T	A	G	A	0.2329	0.5874	0.7782
rs6920483	*ESR1*	T	G	A	G	0.2086	0.1588	0.1808
rs6927072	*ESR1*	T	T	G	T	0.21	0.04895	0.9451
rs6930355	*ESR1*	T	T	T	C	0.4901	0.6064	0.0959
rs6932902	*ESR1*	T	G	A	G	0.3874	0.7003	0.4725
rs6937568	*ESR1*	T	T	C	T	0.1979	0.1898	0.2757
rs6940499	*ESR1*	T	T	C	T	0.1026	0.968	0.7562
rs722210	*ESR1*	T	G	T	G	0.1125	N/A	N/A
rs726281	*ESR1*	T	G	A	G	0.08054	0.1129	0.95
rs728523	*ESR1*	T	T	A	T	0.3013	0.954	0.7053
rs7743290	*ESR1*	T	T	G	T	0.3907	0.0959	0.9361
rs7744006	*ESR1*	T	G	A	G	0.1291	0.9171	0.04196
rs7748205	*ESR1*	T	T	C	T	0.3	0.1339	0.3327
rs7754762	*ESR1*	T	T	A	T	0.3758	0.06394	0.968
rs7755185	*ESR1*	T	A	A	G	0.3179	0.2677	0.3327

SNP ID	Gene	T/P	Ref.	A1 (minor)	A2 (major)	MAF (sample)	*P*-value (FN BMC)	*P*-Value (LS BMC)
rs7757956	*ESR1*	T	T	A	T	0.4133	0.976	0.3776
rs7759411	*ESR1*	T	C	T	C	0.22	0.2328	0.6314
rs7759554	*ESR1*	T	C	T	C	0.0894	0.4456	0.979
rs7760755	*ESR1*	T	T	C	T	0.0777	0.2827	0.8242
rs7761133	*ESR1*	T	T	T	C	0.2318	0.7552	0.2488
rs7761846	*ESR1*	T	T	C	T	0.298	0.5385	0.6533
rs7766762	*ESR1*	T	G	G	A	0.232	N/A	N/A
rs7768330	*ESR1*	T	A	G	A	0.2567	0.6753	0.8551
rs7772475	*ESR1*	T	C	T	C	0.2219	0.7572	0.6963
rs7775047	*ESR1*	T	G	G	C	0.2367	0.4675	0.2667
rs827421	*ESR1*	T	G	T	C	0.2752	0.06593	0.6164
rs827423	*ESR1*	T	G	T	C	0.2781	0.07293	0.6234
rs926777	*ESR1*	T	A	C	A	0.3792	0.3926	0.6953
rs9322331	*ESR1*	T	C	T	C	0.08456	0.06893	0.4505
rs9322332	*ESR1*	T	C	C	A	0.4503	0.02398	0.3367
rs9322335	*ESR1*	T	T	T	C	0.3311	0.8681	0.6284
rs9322337	*ESR1*	T	C	A	C	0.2667	0.5874	0.5644
rs9322338	*ESR1*	T	G	G	A	0.3642	0.2338	0.6513
rs9322352	*ESR1*	T	G	G	A	0.5	0.4366	0.1009
rs9322356	*ESR1*	T	G	A	G	0.27	0.1149	0.2188
rs9322359	*ESR1*	T	A	T	A	0.1358	0.3686	0.981
rs932477	*ESR1*	T	A	A	G	0.1467	0.9421	0.7003
rs9340779	*ESR1*	T	C	T	C	0.04305	0.98	0.6124
rs9340784	*ESR1*	T	T	A	T	0.1887	0.6523	0.7612
rs9340785	*ESR1*	T	G	C	G	0.05333	0.03197	0.962
rs9340787	*ESR1*	T	A	G	A	0.02055	0.6354	0.6054
rs9340799	*ESR1*	T	A	G	A	0.4007	0.8422	0.6583
rs9340807	*ESR1*	T	A	T	A	0.04636	0.1059	0.9391
rs9340811	*ESR1*	T	C	A	C	0.01987	0.8382	0.4855
rs9340820	*ESR1*	T	T	G	T	0.1623	0.01898	0.955
rs9340831	*ESR1*	T	G	A	G	0.06	0.5924	0.6224
rs9340833	*ESR1*	T	A	0	A	0	1	1
rs9340835	*ESR1*	T	G	A	G	0.3444	0.1558	0.1728
rs9340838	*ESR1*	T	T	G	T	0.07947	0.977	0.07493
rs9340888	*ESR1*	T	A	C	A	0.1225	0.4755	0.4825
rs9340901	*ESR1*	T	G	A	G	0.03311	0.8531	0.6623
rs9340933	*ESR1*	T	C	G	C	0.08929	N/A	N/A
rs9340939	*ESR1*	T	C	A	C	0.1333	0.7003	0.04296
rs9340941	*ESR1*	T	C	T	C	0.1033	0.8192	0.8102
rs9340944	*ESR1*	T	G	A	G	0.1424	0.2617	0.1079
rs9340954	*ESR1*	T	T	T	G	0.2167	0.9031	0.5734

SNP ID	Gene	T/P	Ref.	A1 (minor)	A2 (major)	MAF (sample)	*P*-value (FN BMC)	*P*-Value (LS BMC)
rs9340971	*ESR1*	T	A	G	A	0.1067	0.6364	0.8781
rs9341004	*ESR1*	T	C	G	C	0.1126	0.7882	0.8372
rs9341008	*ESR1*	T	G	A	G	0.1225	0.1299	0.9271
rs9341009	*ESR1*	T	C	A	C	0.006623	0.1379	0.8322
rs9341021	*ESR1*	T	C	T	C	0.02649	0.3357	0.3287
rs9341049	*ESR1*	T	A	G	A	0.04967	0.6563	0.953
rs9341062	*ESR1*	T	A	G	A	0.03642	0.2208	0.1469
rs9341074	*ESR1*	T	T	A	T	0.0298	0.1508	0.957
rs9371234	*ESR1*	T	C	A	C	0.1689	0.02997	0.4505
rs9371564	*ESR1*	T	A	G	A	0.3467	0.8511	0.8871
rs9371575	*ESR1*	T	T	T	C	0.4603	0.6234	0.08691
rs9383598	*ESR1*	T	C	T	C	0.1391	0.04595	0.7912
rs9383599	*ESR1*	T	C	G	C	0.1391	0.8811	0.99
rs9383607	*ESR1*	T	T	T	A	0.4895	0.1958	0.2717
rs9383951	*ESR1*	T	G	C	G	0.1921	0.6164	0.8272
rs9397453	*ESR1*	T	G	A	G	0.2148	0.03397	0.4056
rs9397459	*ESR1*	T	G	A	G	0.2	0.06094	0.983
rs9397462	*ESR1*	T	T	A	T	0.1589	0.7083	0.1149
rs9397463	*ESR1*	T	C	T	C	0.1987	0.961	0.1988
rs9478248	*ESR1*	T	G	A	G	0.1846	0.4565	0.8691
rs9478251	*ESR1*	T	C	G	C	0.2417	0.7552	0.97
rs9479120	*ESR1*	T	A	G	A	0.05298	0.3596	0.7213
rs9479134	*ESR1*	T	G	N/A	N/A	N/A	N/A	N/A
rs9479188	*ESR1*	T	G	G	A	0.3867	0.8102	0.05495
rs974277	*ESR1*	T	C	C	T	0.4735	0.5485	0.2128
rs985191	*ESR1*	T	A	C	A	0.09603	0.3506	0.7692
rs985694	*ESR1*	T	T	T	C	0.2086	0.8422	0.1678
rs985695	*ESR1*	T	C	T	C	0.4128	0.05594	0.7243
rs1828720	*SOST*	P	A	T	G	0.1567	0.7073	0.5734
rs1881107	*SOST*	P	A	A	G	0.1954	0.3347	0.7782
rs1983490	*SOST*	P	G	C	T	0.03333	N/A	N/A
rs4792909	*SOST*	P	G	G	T	0.3245	0.7682	0.4196
rs4793018	*SOST*	P	C	C	T	0.3311	0.7792	0.6873
rs4793019	*SOST*	P	G	G	T	0.106	0.1548	0.8312
rs7220711	*SOST*	P	A	A	G	0.1954	0.5514	0.5055
rs9909172	*SOST*	P	T	T	C	0.1788	0.8242	0.8332
rs9911277	*SOST*	P	G	G	A	0.3311	0.7792	0.6873
rs10002782	*SPP1*	P	A	A	G	0.1722	0.3327	0.7962
rs1381948	*SPP1*	P	T	A	G	0.2152	0.4356	0.3666
rs1463090	*SPP1*	P	G	C	A	0.2152	0.3137	0.4995
rs1471403	*SPP1*	P	T	T	C	0.34	0.06893	0.4156

SNP ID	Gene	T/P	Ref.	A1 (minor)	A2 (major)	MAF (sample)	*P*-value (FN BMC)	*P*-Value (LS BMC)
rs2869722	*SPP1*	P	C	G	A	0.15	0.02997	0.4456
rs6846509	*SPP1*	P	C	C	T	0.22	0.9041	0.4426
rs7683960	*SPP1*	P	A	A	G	0.2152	0.4356	0.3666
rs11664594	*TNFRSF11A*	T	A	A	T	0.3146	0.973	0.2927
rs11872145	*TNFRSF11A*	T	T	A	T	0.07616	0.5644	0.6793
rs11876459	*TNFRSF11A*	T	A	T	A	0.04967	0.6194	0.8432
rs11877530	*TNFRSF11A*	T	G	A	G	0.394	0.3796	0.3686
rs12956925	*TNFRSF11A*	T	G	A	G	0.06623	0.7742	0.8621
rs12959396	*TNFRSF11A*	T	T	G	T	0.3113	0.7263	0.5155
rs12970081	*TNFRSF11A*	T	G	A	G	0.3533	0.991	0.6983
rs13381269	*TNFRSF11A*	T	C	T	C	0.106	0.8352	0.7932
rs17069898	*TNFRSF11A*	T	A	G	A	0.271	N/A	N/A
rs17069901	*TNFRSF11A*	T	A	G	A	0.1358	0.8881	0.7932
rs17069902	*TNFRSF11A*	T	C	T	C	0.1159	N/A	N/A
rs17069904	*TNFRSF11A*	T	G	A	G	0.1367	0.4925	0.9171
rs17720953	*TNFRSF11A*	T	G	A	G	0.1126	0.9271	0.07393
rs3826618	*TNFRSF11A*	T	T	C	T	0.2133	0.6014	0.5914
rs3826620	*TNFRSF11A*	T	T	G	T	0.1477	N/A	N/A
rs4263037	*TNFRSF11A*	T	G	G	A	0.3808	0.7762	0.3836
rs4303637	*TNFRSF11A*	T	T	C	T	0.04305	0.8611	0.968
rs4426449	*TNFRSF11A*	T	C	T	C	0.1812	0.7782	0.5984
rs4485469	*TNFRSF11A*	T	G	G	A	0.3533	0.3916	0.2607
rs4524034	*TNFRSF11A*	T	A	G	A	0.2667	N/A	N/A
rs4524035	*TNFRSF11A*	T	A	G	A	0.3245	0.9251	0.6384
rs4941125	*TNFRSF11A*	T	A	G	A	0.009934	0.2268	0.1818
rs6567270	*TNFRSF11A*	T	T	T	A	0.4367	0.3417	0.4076
rs6567271	*TNFRSF11A*	T	T	T	A	0.2715	0.2448	0.3886
rs6567274	*TNFRSF11A*	T	G	T	G	0.3079	0.5215	0.4386
rs7231887	*TNFRSF11A*	T	G	A	G	0.27	N/A	N/A
rs7235803	*TNFRSF11A*	T	A	A	G	0.3741	0.3367	0.1968
rs7236060	*TNFRSF11A*	T	G	G	A	0.2872	0.4985	0.6394
rs7237982	*TNFRSF11A*	T	A	G	A	0.3709	0.4016	0.7183
rs7239261	*TNFRSF11A*	T	A	A	C	0.2583	0.6244	0.7243
rs8083511	*TNFRSF11A*	T	A	C	A	0.4073	0.1928	0.5375
rs8086340	*TNFRSF11A*	T	C	C	G	0.3907	0.4136	0.5335
rs8089829	*TNFRSF11A*	T	A	A	G	0.3179	0.6324	0.5085
rs8092023	*TNFRSF11A*	T	T	A	T	0.3046	0.3746	0.9111
rs8094884	*TNFRSF11A*	T	G	A	G	0.3907	0.2218	0.5125
rs9646629	*TNFRSF11A*	T	C	C	G	0.2174	0.2058	0.1608
rs9951012	*TNFRSF11A*	T	G	A	G	0.1107	0.8142	0.5704
rs9960450	*TNFRSF11A*	T	T	C	T	0.09603	0.7273	0.6693

SNP ID	Gene	T/P	Ref.	A1 (minor)	A2 (major)	MAF (sample)	*P*-value (FN BMC)	*P*-Value (LS BMC)
rs9961156	*TNFRSF11A*	T	T	C	T	0.1722	0.8432	0.8501
rs11873955	*TNFRSF11A*	T	G	A	G	0.08278	0.3546	0.8352
rs2957137	*TNFRSF11A*	P	A	A	T	0.3267	0.5524	0.9291
rs884205	*TNFRSF11A*	P	A	T	G	0.03	0.6763	0.8242
rs1032128	*TNFRSF11B*	T	G	A	G	0.3212	0.6454	0.06593
rs1032129	*TNFRSF11B*	T	A	C	A	0.3322	0.4655	0.04995
rs10505344	*TNFRSF11B*	T	G	A	G	0.1887	0.0989	0.6733
rs10505345	*TNFRSF11B*	T	G	A	G	0.2252	0.6693	0.4436
rs10955911	*TNFRSF11B*	T	C	A	G	0.1225	0.8222	0.4875
rs11573828	*TNFRSF11B*	T	C	A	G	0.07616	0.2987	0.08591
rs11573880	*TNFRSF11B*	T	T	C	A	0.06291	0.8242	0.7902
rs11573894	*TNFRSF11B*	T	A	C	T	0.03311	0.965	0.5335
rs11573896	*TNFRSF11B*	T	T	T	A	0.13	0.8691	0.2917
rs11573899	*TNFRSF11B*	T	C	T	G	0.1126	0.3447	0.4685
rs11573903	*TNFRSF11B*	T	G	T	C	0.3209	0.2667	0.9461
rs11573911	*TNFRSF11B*	T	G	T	C	0.1821	0.08292	0.981
rs11573930	*TNFRSF11B*	T	C	A	G	0.06291	0.7582	0.8292
rs11573938	*TNFRSF11B*	T	C	A	G	0.01656	0.02198	0.8861
rs1485287	*TNFRSF11B*	T	C	T	C	0.3709	0.8581	0.2567
rs1485289	*TNFRSF11B*	P	G	G	A	0.03642	0.2488	0.6983
rs1804854	*TNFRSF11B*	T	A	C	T	0.33	0.3736	0.03596
rs1905786	*TNFRSF11B*	P	T	T	C	0.3146	0.973	0.4735
rs2073617	*TNFRSF11B*	P	G	C	T	0.277	0.4106	0.0979
rs2073618	*TNFRSF11B*	P	G	C	G	0.1075	N/A	N/A
rs2875845	*TNFRSF11B*	T	A	G	A	0.03974	0.2787	0.3936
rs3102724	*TNFRSF11B*	T	G	A	G	0.4669	0.5345	0.6254
rs3134053	*TNFRSF11B*	T	T	T	C	0.1788	0.6933	0.5405
rs3134054	*TNFRSF11B*	P	A	G	A	0.4964	0.6603	0.5185
rs3134057	*TNFRSF11B*	P	A	G	A	0.404	0.6464	0.8781
rs3134058	*TNFRSF11B*	P	G	G	A	0.3027	0.7143	0.08591
rs3134063	*TNFRSF11B*	P	C	C	T	0.2781	0.3806	0.08392
rs4520195	*TNFRSF11B*	T	A	G	A	0.2682	0.1998	0.01399
rs6469783	*TNFRSF11B*	T	C	C	T	0.3958	0.5225	0.2268
rs6469788	*TNFRSF11B*	P	A	A	C	0.01325	0.2577	0.8402
rs7010267	*TNFRSF11B*	T	C	A	C	0.3093	N/A	N/A
rs7464496	*TNFRSF11B*	P	T	T	C	0.06291	0.04595	0.9461
rs882763	*TNFRSF11B*	T	C	T	C	0.3333	0.3556	0.02997
rs10086835	*TNFRSF11B*	P	C	C	T	0.2086	0.1808	0.952
rs10098408	*TNFRSF11B*	P	T	T	C	0.3013	0.1129	0.6693
rs10505351	*TNFRSF11B*	P	C	C	T	0.3079	0.1109	0.6484
rs10955919	*TNFRSF11B*	P	C	C	G	0.2248	0.6214	0.8981

SNP ID	Gene	T/P	Ref.	A1 (minor)	A2 (major)	MAF (sample)	*P*-value (FN BMC)	*P*-Value (LS BMC)
rs10955924	*TNFRSF11B*	P	C	C	A	0.3079	0.06094	0.7762
rs11992136	*TNFRSF11B*	P	C	C	G	0.2119	0.2997	0.7672
rs11995824	*TNFRSF11B*	P	C	C	G	0.06623	0.2727	0.6034
rs13262276	*TNFRSF11B*	P	T	T	A	0.09934	0.3926	0.7163
rs13264172	*TNFRSF11B*	P	T	T	A	0.2119	0.1668	0.1898
rs13264791	*TNFRSF11B*	T	T	T	C	0.0596	0.988	0.8531
rs1485295	*TNFRSF11B*	P	A	A	G	0.1391	0.1608	0.3696
rs1485302	*TNFRSF11B*	P	G	G	A	0.3642	0.4156	0.4645
rs1485303	*TNFRSF11B*	P	G	G	A	0.3675	0.2947	0.4995
rs1485305	*TNFRSF11B*	P	A	T	A	0.2351	0.1788	0.2008
rs1485310	*TNFRSF11B*	P	A	A	T	0.4272	0.6673	0.4715
rs1485312	*TNFRSF11B*	P	T	A	G	0.3079	0.06094	0.7762
rs1493934	*TNFRSF11B*	P	T	C	T	0.09524	N/A	N/A
rs1564860	*TNFRSF11B*	P	C	T	C	0.4305	0.5944	0.3257
rs1905779	*TNFRSF11B*	P	C	C	T	0.3079	0.1189	0.7393
rs1905783	*TNFRSF11B*	P	C	G	A	0.3113	0.06394	0.8462
rs1905784	*TNFRSF11B*	P	G	C	T	0.3199	0.1469	0.974
rs1948467	*TNFRSF11B*	P	T	G	A	0.06954	0.5604	0.963
rs2035977	*TNFRSF11B*	P	C	G	C	0.4195	0.01499	0.0999
rs2055101	*TNFRSF11B*	P	C	G	A	0.4868	0.1548	0.0959
rs2062375	*TNFRSF11B*	P	G	G	C	0.07285	0.1588	0.2897
rs2062376	*TNFRSF11B*	P	C	G	C	0.2185	0.2537	0.2278
rs2062377	*TNFRSF11B*	P	T	A	T	0.1689	0.3257	0.5485
rs2101752	*TNFRSF11B*	P	T	G	T	0.2252	0.05594	0.02697
rs2326193	*TNFRSF11B*	P	A	A	C	0.09272	0.2997	0.5375
rs3103969	*TNFRSF11B*	P	A	T	C	0.4596	0.9401	0.1708
rs3103971	*TNFRSF11B*	P	C	A	G	0.1338	0.8581	0.1389
rs3133585	*TNFRSF11B*	P	A	G	A	0.2848	0.1568	0.05694
rs4307369	*TNFRSF11B*	P	A	A	G	0.09934	0.3876	0.5514
rs4335155	*TNFRSF11B*	P	A	A	C	0.2848	0.1479	0.2448
rs4354338	*TNFRSF11B*	P	A	A	C	0.1391	0.1608	0.3696
rs4401893	*TNFRSF11B*	P	C	C	T	0.31	0.06294	0.8242
rs4424291	*TNFRSF11B*	P	A	A	G	0.3033	0.1069	0.7033
rs4424296	*TNFRSF11B*	P	C	C	T	0.1987	0.1329	0.2418
rs4615609	*TNFRSF11B*	P	A	A	G	0.303	N/A	N/A
rs4876876	*TNFRSF11B*	P	T	T	C	0.2133	0.3057	0.8002
rs6469792	*TNFRSF11B*	P	T	T	C	0.22	0.2537	0.2318
rs6469795	*TNFRSF11B*	P	G	G	A	0.3046	0.1129	0.6693
rs6469797	*TNFRSF11B*	P	C	C	T	0.3533	0.05694	0.3457
rs6469798	*TNFRSF11B*	P	C	C	T	0.1429	0.2088	0.3357
rs6469805	*TNFRSF11B*	P	C	C	T	0.003472	0.993	0.3826

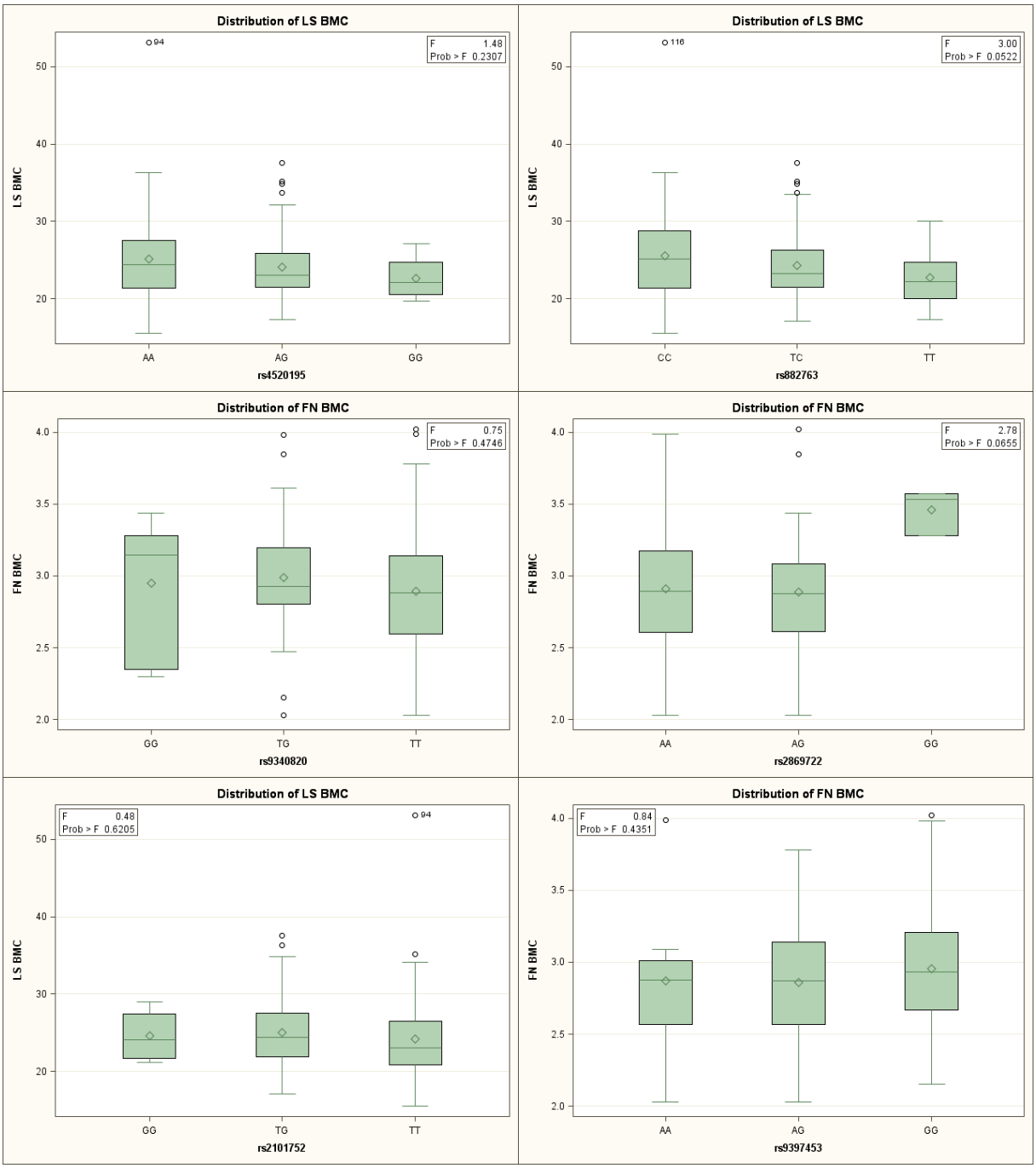
SNP ID	Gene	T/P	Ref.	A1 (minor)	A2 (major)	MAF (sample)	*P*-value (FN BMC)	*P*-Value (LS BMC)
rs6984675	*TNFRSF11B*	P	T	T	G	0.2552	N/A	N/A
rs6993813	*TNFRSF11B*	P	T	T	C	0.2351	0.05495	0.5914
rs6996754	*TNFRSF11B*	P	T	T	C	0.1391	0.1608	0.3696
rs7004052	*TNFRSF11B*	P	T	T	A	0.31	0.06294	0.8242
rs7010043	*TNFRSF11B*	P	G	G	T	0.1821	0.5275	0.4296
rs7013722	*TNFRSF11B*	P	A	A	T	0.3113	0.06394	0.8462
rs7013731	*TNFRSF11B*	P	T	T	C	0.2963	N/A	N/A
rs7018198	*TNFRSF11B*	P	A	A	G	0.3085	0.1878	0.9191
rs7813486	*TNFRSF11B*	P	T	T	C	0.3384	N/A	N/A
rs7822098	*TNFRSF11B*	P	A	A	T	0.1225	N/A	N/A
rs7839059	*TNFRSF11B*	P	C	C	A	0.3333	0.03596	0.2078
rs9650075	*TNFRSF11B*	P	A	A	G	0.3046	0.1129	0.6693
rs1022926	*TNFSF11*	T	C	C	G	0.1159	0.3227	0.7163
rs1054016	*TNFSF11*	T	G	T	G	0.1858	0.9001	0.5245
rs2148072	*TNFSF11*	T	G	A	G	0.2815	0.5095	0.2757
rs2148073	*TNFSF11*	T	G	G	C	0.1093	0.3047	0.6683
rs2200287	*TNFSF11*	T	G	T	C	0.2349	0.7862	0.1469
rs2277439	*TNFSF11*	T	G	G	A	0.1093	0.8062	0.9491
rs2324851	*TNFSF11*	T	A	T	C	0.1291	0.1029	0.6683
rs238223	*TNFSF11*	T	A	T	C	0.1333	N/A	N/A
rs346588	*TNFSF11*	T	A	C	T	0.05298	0.6064	0.8571
rs346589	*TNFSF11*	T	G	T	C	0.08278	0.5485	0.8621
rs346591	*TNFSF11*	T	A	A	T	0.07947	0.4236	0.7722
rs3742257	*TNFSF11*	T	T	C	T	0.4238	0.6314	0.1918
rs443188	*TNFSF11*	T	G	T	C	0.06623	0.05794	0.8482
rs4941433	*TNFSF11*	T	G	T	G	0.07947	0.969	0.7672
rs4942143	*TNFSF11*	T	G	A	G	0.06623	0.3856	0.4785
rs7334187	*TNFSF11*	T	T	G	T	0.1921	0.8252	0.3137
rs7990107	*TNFSF11*	T	C	T	C	0.1854	0.991	0.3297
rs7994018	*TNFSF11*	T	T	C	T	0.1887	0.967	0.4915
rs8002449	*TNFSF11*	T	G	G	A	0.1	0.1598	0.6873
rs875625	*TNFSF11*	T	A	G	A	0.3775	0.973	0.9161
rs900797	*TNFSF11*	T	C	A	C	0.06291	0.05594	0.7083
rs922996	*TNFSF11*	T	G	T	C	0.3867	0.6753	0.3916
rs9525643	*TNFSF11*	T	T	C	T	0.2483	0.4895	0.9341
rs9525644	*TNFSF11*	T	G	A	G	0.457	0.4605	0.4146
rs9562415	*TNFSF11*	T	T	C	T	0.09055	N/A	N/A
rs2062305	*TNFSF11*	P	G	A	G	0.2318	0.4715	0.6414
rs7322806	*TNFSF11*	P	G	A	G	0.3576	0.9301	0.7493
rs7327904	*TNFSF11*	P	C	T	C	0.4205	0.7013	0.8112
rs768625	*TNFSF11*	P	C	A	G	0.3477	0.9351	0.9111

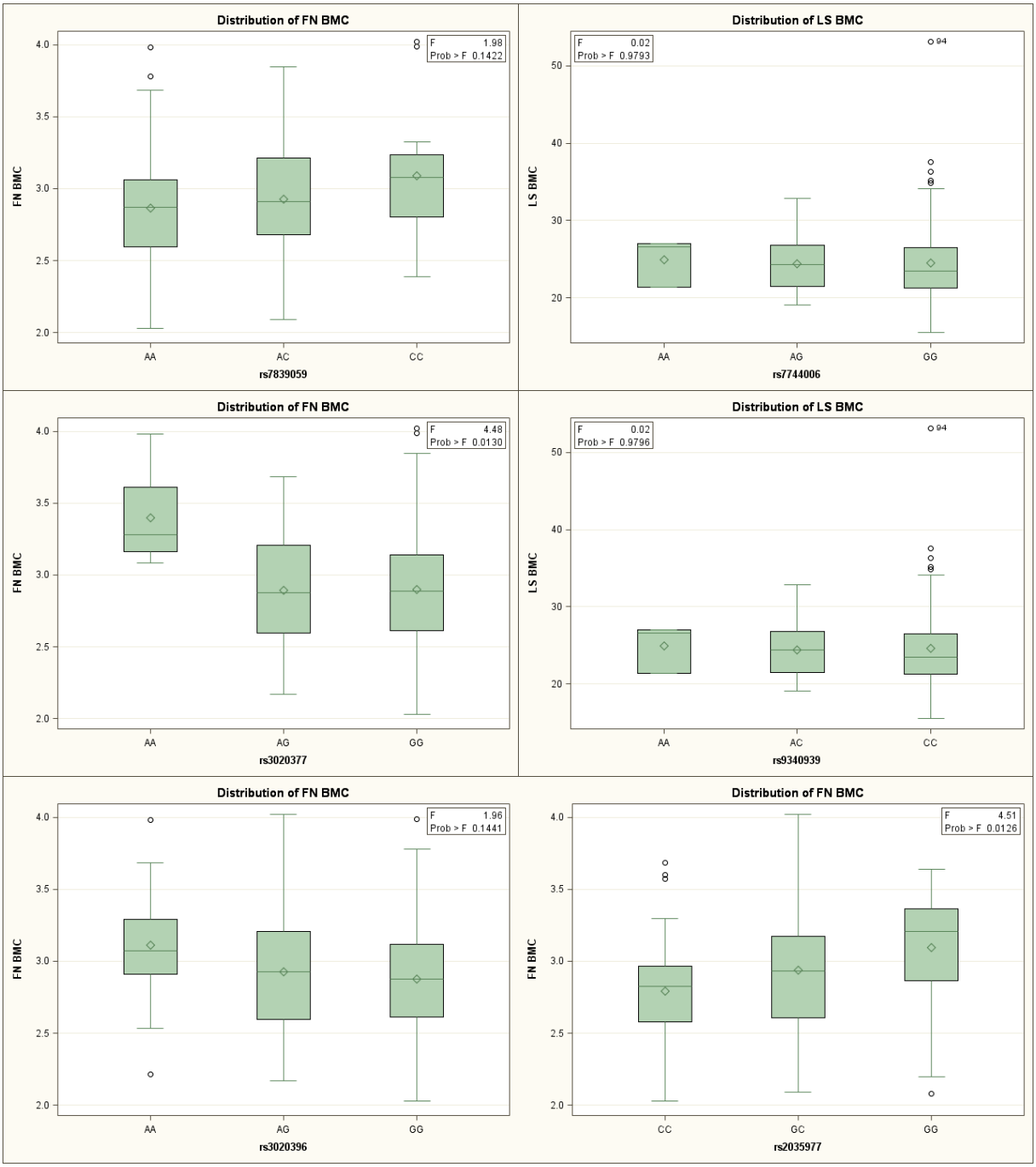
SNP ID	Gene	T/P	Ref.	A1 (minor)	A2 (major)	MAF (sample)	*P*-value (FN BMC)	*P*-Value (LS BMC)
rs7983721	*TNFSF11*	P	G	A	G	0.35	0.5045	0.8032
rs9594766	*TNFSF11*	P	G	A	G	0.1771	0.3477	0.2238

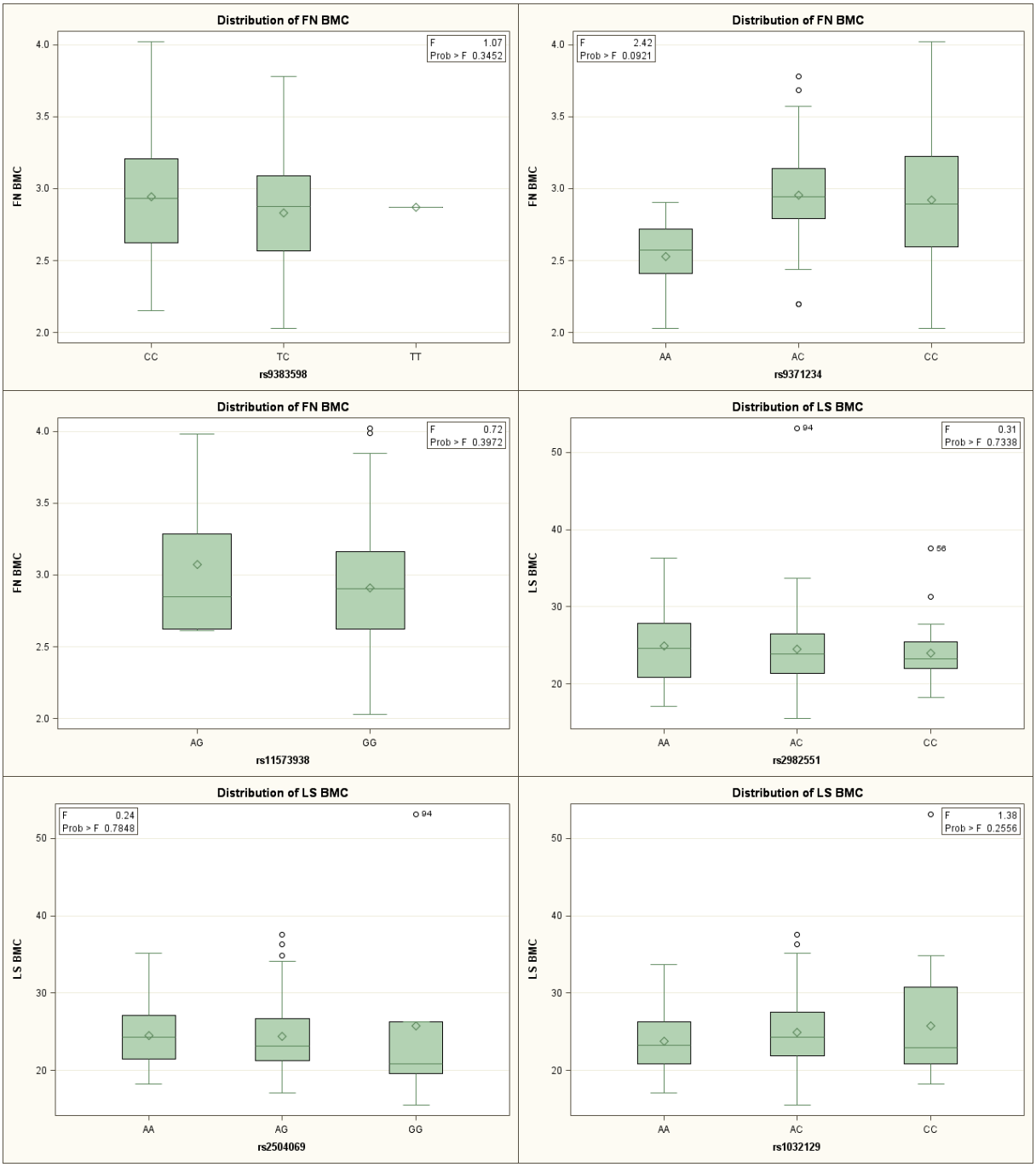
† *P* = 0.001998 prior to correction for multiple testing; T/P – tag/previously associated

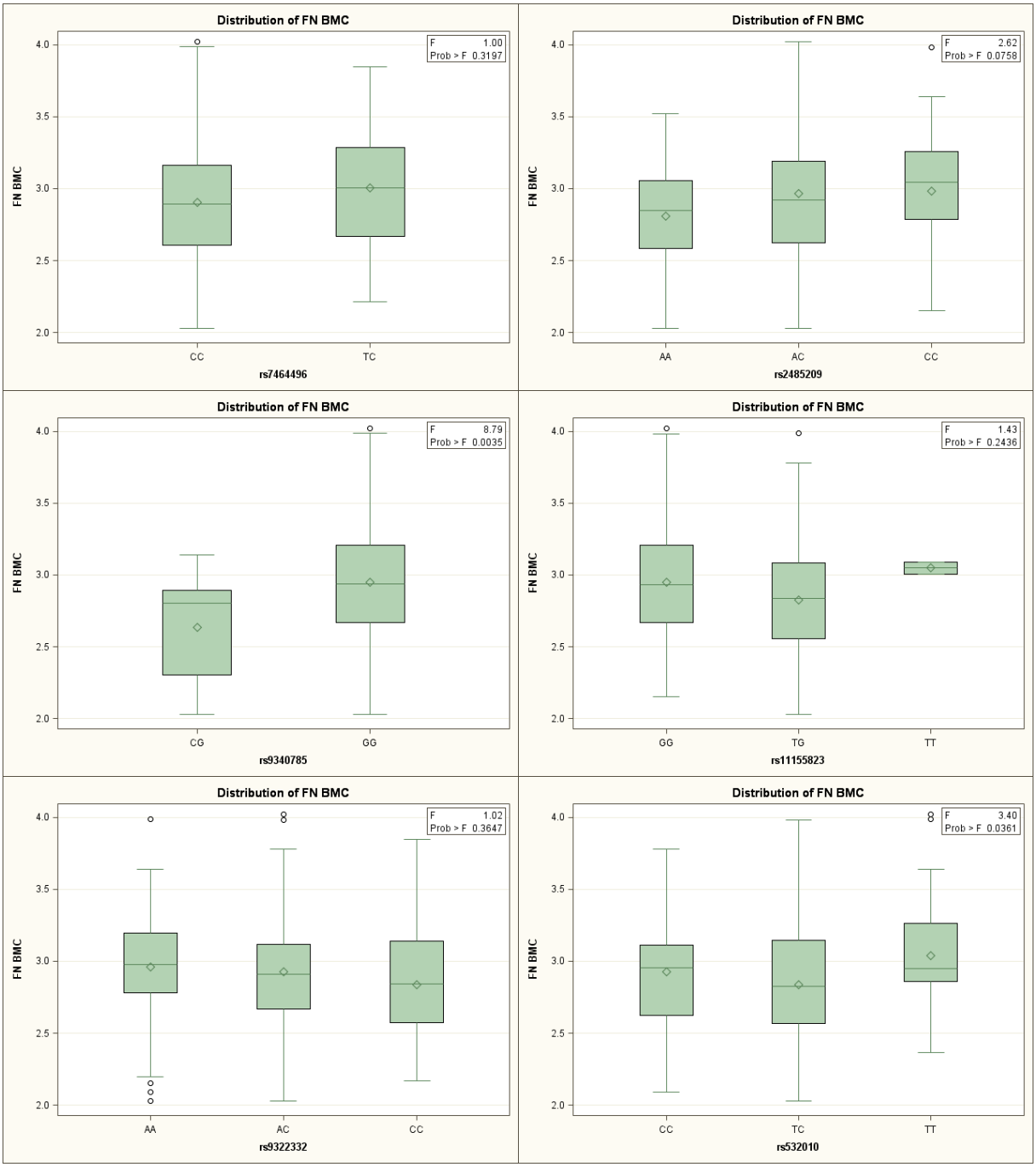
Appendix F – Box-and-whisker plots

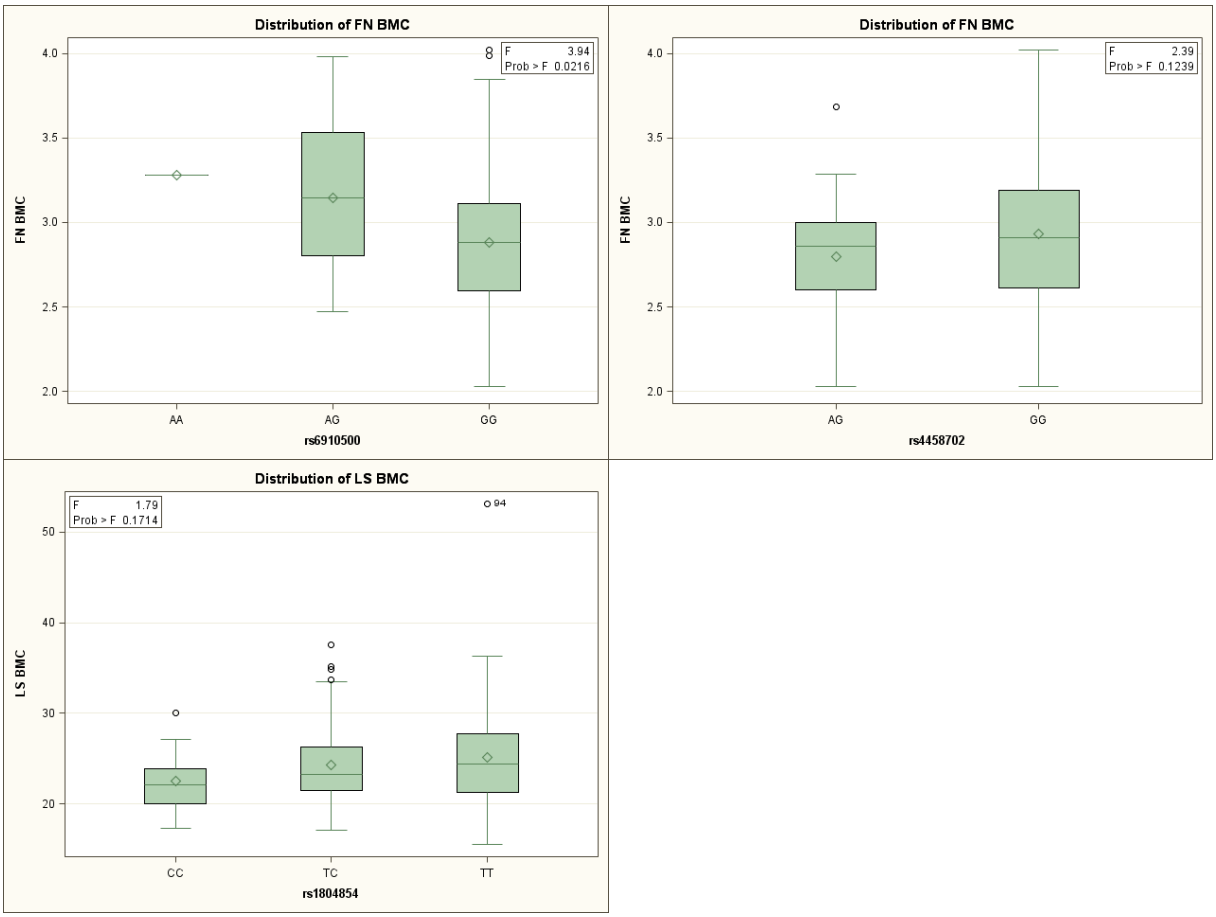
Box-and-whisker plots for all associated SNPs




Distribution of FN BMC
F 1.98
Prob > F 0.1422
FN BMC
AA
AC
CC
rs7839059
Distribution of LS BMC
F 0.02
Prob > F 0.9793
LS BMC
AA
AG
GG
rs7744006
Distribution of FN BMC
F 4.48
Prob > F 0.0130
FN BMC
AA
AG
GG
rs3020377
Distribution of LS BMC
F 0.02
Prob > F 0.9796
LS BMC
AA
AC
CC
rs9340939
Distribution of FN BMC
F 1.96
Prob > F 0.1441
FN BMC
AA
AG
GG
rs3020396
Distribution of FN BMC
F 4.51
Prob > F 0.0126
FN BMC
CC
GC
GG
rs2035977


Distribution of FN BMC
F 1.07
Prob > F 0.3452
FN BMC
4.0
3.5
3.0
2.5
2.0
CC
TC
TT
rs9383598
Distribution of FN BMC
F 2.42
Prob > F 0.0921
FN BMC
AA
AC
CC
rs9371234
Distribution of FN BMC
F 0.72
Prob > F 0.3972
FN BMC
AG
GG
rs11573938
Distribution of LS BMC
F 0.31
Prob > F 0.7338
LS BMC
50
40
30
20
AA
AC
CC
rs2982551
Distribution of LS BMC
F 0.24
Prob > F 0.7848
LS BMC
AA
AG
GG
rs2504069
Distribution of LS BMC
F 1.38
Prob > F 0.2556
LS BMC
AA
AC
CC
rs1032129


Distribution of FN BMC
F 1.00
Prob > F 0.3197
FN BMC
4.0
3.5
3.0
2.5
2.0
CC
TC
rs7464496
Distribution of FN BMC
F 2.62
Prob > F 0.0758
FN BMC
AA
AC
CC
rs2485209
Distribution of FN BMC
F 8.79
Prob > F 0.0035
FN BMC
CG
GG
rs9340785
Distribution of FN BMC
F 1.43
Prob > F 0.2436
FN BMC
GG
TG
TT
rs11155823
Distribution of FN BMC
F 1.02
Prob > F 0.3647
FN BMC
AA
AC
CC
rs9322332
Distribution of FN BMC
F 3.40
Prob > F 0.0361
FN BMC
CC
TC
TT
rs532010


Distribution of FN BMC
F 3.94
Prob > F 0.0216
FN BMC
4.0
3.5
3.0
2.5
2.0
AA
AG
GG
rs6910500
Distribution of FN BMC
F 2.39
Prob > F 0.1239
FN BMC
4.0
3.5
3.0
2.5
2.0
AG
GG
rs4458702
Distribution of LS BMC
F 1.79
Prob > F 0.1714
LS BMC
50
40
30
20
CC
TC
TT
rs1804854