QATAR UNIVERSITY

COLLEGE OF HEALTH SCIENCES

GENOME-WIDE ASSOCIATION STUDY (GWAS) TO UNCOVER GENETIC RISK FACTORS ASSOCIATED WITH LOW BONE MINERAL DENSITY AND OSTEOPOROSIS IN QATAR POPULATION

BY

NADIN NAGY YOUNES

A Thesis Submitted to

the Faculty of the College of Health Sciences

in Partial Fulfillment of the Requirements for the Degree of

Masters of Science in Biomedical Sciences

June 2019

COMMITTEE PAGE

The members of the Committee approve the Thesis of
Nadin defended on 17/04/2019.

Marawan Abou Madi
Thesis/Dissertation Supervisor

Nasser Rizk
Committee Member

Nahla Afifi
Committee Member

Mohammed Harris
Committee Member

Approved:

Asma Al-Thani, Dean, College of Health Science

ABSTRACT

YOUNES, NADIN, NAGY NY., Masters of Science : June : [2019:],
Biomedical Sciences

Title: Genome Wide Association Study (GWAS) to Uncover Genomic Risk Factors Associated with Low Bone Mineral Density and Osteoporosis in The Qatari population

Supervisor of Thesis: Marawan Abdelhamid Abou Madi.

Osteoporosis is an increasingly prevalent, global health burden characterized by low bone mineral density (BMD) and increased fracture risk. Despite the serious consequences of osteoporosis and the significant impact it can have on human health, the majority of affected individuals are unaware of the disease because of its asymptomatic 'silent' nature. Understanding the genetic basis of the Osteoporosis is crucial to fully elucidate the etiology of the disease. Towards this goal, genome-wide association studies (GWAS) have identified a number of promising genetic variants that are associated with osteoporosis and low BMD. Here, we undertook a genome-wide association study (GWAS) in 3000 healthy Qatari individuals from Qatar Biobank to identify risk genetic variants associated with low BMD in the Qatari population. 19 significant single-nucleotide polymorphisms (SNPs) have been identified to be associated with BMD ($P<5\times10^{-8}$). Of these, 6 SNPs were replicated and directionally consistent in UK Biobank, in which 2 of these SNPs were identified and known to be involved in the Wnt signaling pathways which is important in bone formation. The other 13 SNPs weren't associated to any diseases and thus were regarded as novel. 8 of these variants were intronic variants harbored in 8 gene loci; *MALAT1, MRPL39, FASLG, SAG, FAM189A2, RP11-15A1.7, LSAMP*, and *BMPR1B* and 5 were intergenic

variants. The finding of our study, which to our knowledge is the first GWAS of any form of bone disease in the Qatari population, provide new insights into the genetic architecture of BMD. Further studies are needed to identify the causal variants and their functional effects to unveil unknown players contributing to BMD variation and fracture susceptibility.

DEDICATION

To my beloved Mother, Father and Sister

ACKNOWLEDGMENTS

First of all, I am very thankful to Allah for his blessing to me with this great opportunity, who gave me the wellbeing, health, and the strength to finish this project for my master's thesis. Second, I would like to express my sincere appreciation to my supervisor Dr. Marawan Abu Madi and Dr. Nahla Afifi for their endless support and guidance throughout my Masters' project. Working with him helped me to improve my personal and professional skills, and most importantly to finish this work. I am very thankful for his guidance, unlimited patience and continues motivation. I appreciate the provision of my committee members; Dr. Nasser Rizk for their continuous support, guidance mentoring, encouragement and their insightful comments throughout my project. Special thanks to the Qatar Biobank for giving me this large number of participant's sample to assure the best quality of the project and for easing all the process for me to be able to continue and finish my project on time.

Last but not least, my sincere appreciation goes to my family and friends for their encouragement, prayers, cheerfulness. I wouldn't be able to complete this journey without my mother's prayers and the limitless spiritual support they all surrounded me with.

This project has been supported by Qatar University students grants. Ethical approval has been obtained from Qatar Biobank.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

BMD: Bone Mineral Density

BMU: Bone Multicellular Units

DPA: Dual energy photon absorptiometry

DXA: Dual-energy x-ray absorptiometry

GEFOS: Genetic Factors for Osteoporosis Consortium

GWAS: Genome Wide Association Study

HWE: Hardy-Weinberg Equilibrium

IBD: Identical by Descent

LD: Linkage Disequilibrium

MAF: Minor allele frequency

MRI: Magnetic resonance imaging

NCBI: National Center for Biotechnology Information

SNP: Single Nucleotide Polymorphism

Chapter 1: Introduction

Osteoporosis is a common systemic skeletal disease characterized by reduction of bone mass and density, resulting in increased susceptibility to bone fractures (Munch & Shapiro, 2006). Osteoporosis (meaning "porous bones") is the most common bone disease in humans, representing a serious public health problem. It has been estimated that, worldwide, more than 200 million people are suffering from osteoporosis (Sözen, Özışık, & Başaran, 2017).

Osteoporosis is described as a clinically silent disease because of its asymptomatic nature that makes it difficult to diagnose in its first stages, it is often recognized only when the first fracture occurs (Sharma, Tandon, Mahajan, Kour, & Kumar, 2006). Osteoporosis is responsible for ~9 million fractures each year worldwide (Kanis et al., 2012). Bone fractures occur mostly in the hip, wrist or spine, impairing the quality of life and conferring substantial risk for morbidity and mortality. Although affecting both genders, osteoporosis is more prevalent in postmenopausal women, this is in part, due to declining estrogen levels which have been shown to accelerate bone loss. However, the significant prevalence of the disease in old men and its absence in some postmenopausal women indicate that other factors are also involved in the development of osteoporosis. Factors such as tobacco and alcohol consumption, physical activity, and body weight have all been previously reported to influence bone mass (Tian et al., 2017). The high prevalence of osteoporosis poses a serious economic burden on patients, families, and nation. Early identification and management of people with low BMD is crucial to reduce the incidence of osteoporotic fractures. Bone mineral density (BMD) is widely recognized as one of the most important predictors of osteoporotic fractures and is the primary

clinical measurement used to diagnose osteoporosis. Dual-energy x-ray absorptiometry (DXA) is the gold standard for assessing BMD worldwide. Emerging evidence suggests a link between genetic variation and impairment of BMD, with heritability estimates of 0.6 to 0.8 (Peacock, Econs, Turner, & Foroud, 2002). The heritability is usually measured with twin studies, which can be determined by concordance rate (Sahu & Prasuna, 2016). Innovative advances in high-throughput genotyping and the genome-wide database of human genetic variation produced by the HapMap project have made the genome-wide association (GWA) studies technically feasible, unraveling complex associations between common genetic variants and a growing range of diseases and traits. A number of genome-wide association studies and their meta-analyses have been conducted to explore the relationship between genetic variation and bone mineral density (Karol Estrada et al., 2012; Kemp et al., 2017; Stuart K. Kim, 2018; Koller et al., 2010) identifying dozens of genomic loci. However, the cumulative effects of these identified loci account for only 5.8% of total BMD variation (K. Estrada et al., 2012) implying that many of the genes that influence BMD remain to be unveiled. Thus, Identification of genes regulating BMD, particularly at the most common skeletal fracture sites is critical to provide insights into the genetic architecture of osteoporosis and fracture risk.

Aims:

The aim of our study is to investigate the genetic architecture of bone mineral density variations and to evaluate the role of genetic factors in the pathogenesis of osteoporosis, thereby discovering new genetic loci and the biological pathways, which may help identify drug targets for the prevention and treatment of fragility fracturs.

Specific Goals:

1) Investigate the association of BMD values with common genetic variations.

2) Explore the gender effect on BMD and the association of Vitamin D levels with BMD.

Chapter 2: Literature Review

2.1 Bone Physiology

This section discusses the gross anatomy and histology of bone tissue, as well as describes the crucial process of structural remodeling of the bone.

2.1.1 Bone Architecture

Human bones are complex structure that are critical for providing mobility, and protection for the whole body, and more importantly, providing a reservoir for storing essential minerals. Bone strength is composed of two main components: a) the bone mineral density (BMD) and b) the bone quality, which are the bone architecture, bone turnover, and the mineralization status (Lorincz, Manske, & Zernicke, 2009). The BMD is the amount of bone mineral in the bone tissue, it is the most commonly used expression that measures the overall bone strength in an individual. It is estimated to account for approximately 70% of bone strength (Sözen et al., 2017). Thus, BMD is essential for the diagnosis of osteoporosis and to provide overall information on the bone fracture risk.

Healthy human bones are composed of two different types of structural tissue: Cortical and Cancellous Bone. Cortical (compact) bone is the dense hard outer shell of the bones that forms the protective layer around the inner trabecular bone (Iolascon et al., 2013). It is critical for providing strength to all the long bones of the body, as well as providing sites for attachment of tendons and ligaments. The cortical bones makeup approximately 80% of the skeletal mass (Clarke, 2008) and are vital for supporting the body structure and weight as it is highly resistant to bending and torsion. The remaining 20 % of the skeletal mass are Cancellous bone (Clarke, 2008),

also referred to as spongy or trabecular bone. it is found mainly at the ends of long bones, as well as in the pelvis, ribs, skull, and the vertebral column. It consists of fine sponge-like lattice and has a much higher turnover rate than cortical bone, thus, it has a major role in metabolism. In addition, the cancellous bones are very porous and contain the red bone marrow thus, it is weaker and more prone to fracture compared to cortical bone. Moreover, it has a honeycombed or spongy appearance, in which the bone matrix is organized into a three-dimensional latticework, called trabeculae. The trabeculae contain three types of bone cells: a) the osteoblasts, b) the osteocytes and c) the osteoclasts (Florencio-Silva, Sasso, Sasso-Cerri, Simões, & Cerri, 2015). The osteoblasts are cells derived from mesenchymal stem cells, they produce bone extracellular matrix and are responsible for bone mineralization. They do this by depositing a protein mixture called "osteoid", which contains collagen, as well as depositing minerals including calcium into the osteoid to make new bones. Osteocytes are mature osteoblasts that have been embedded in the bone matrix. They are the most abundant cells in bone, comprising 90-95% of all bone cells (Schaffler & Kennedy, 2012). They act as sensory cells that are involved in signaling processes inside the bone. They form an extensive network through their projections that connect them to other bone cells, endothelial cells, and hematopoietic cells (Schaffler & Kennedy, 2012). Osteoclasts are large multinucleated cells responsible for bone resorption, they break down and destroy old or damaged bone tissue, initiating bone repair and replacement by osteoblasts. This destruction-rebuilding cycle is constantly ongoing to ensure the upkeep of the bone's structural integrity.

2.1.2 Differences between trabecular and cortical bone

The cortical bones are stiffer and more resistant to higher stresses compared to trabecular bone, however, they are more brittle (Carter Dr. & Hayes, 1977; Keaveny & Hayes, 1993). In vitro, the trabecular bone can withstand the strain up to 30% whilst, cortical bone fails to withstand strain of only 2% (Osterhoff et al., 2016). In addition, the biomechanical structure of cortical bone is uniform while trabecular bone shows a wide variation in strength and stiffness. To a large extent, this variability depends on the apparent density of the trabecular bones. Due to its heterogeneity, the trabecular bone modulus can vary 100-fold from one location to another within the same metaphysis (Keaveny & Hayes, 1993). Moreover, the stiffness and strength of cortical and trabecular bone depend on the loading direction, indicating its anisotropic microstructure (Galante J., Rostoker W., & Ray, 1970). Overall, the bones can resist higher compression loads than tension loads and to higher tension loads than shear loads(Carter Dr Fau - Hayes & Hayes, 1977).

The location of the bone in the human body and the forces acting on it determine its unique microstructure and composition compared to other bones in other locations. For instance, the vertebral column must resist the high and repeating axial compression loads but, it experiences a much less shear or tension loads (Mizrahi, Keaveny, Edwards, & Hayes, 1976). On the other hand, the femoral neck or the proximal humerus are mainly subjected to high shear and bending forces compared to the vertebral column, which creates a combination of compression, tension, and shear by which both show a distinct need for a cortical structure.

In the body, bones experience different loads from different directions and in different intensity and frequency over time. Bone has two main mechanical responses

to the changing loading patterns: by altering the structural density and by increasing the degree of structural orientation along the acting force vectors, i.e. anisotropy (Keaveny & Hayes, 1993; Nordin, 2012). These adaptive responses of the bones were made possible by the process of continuous remodeling (Seeman & Delmas, 2006), which will be discussed later.

2.1.3 Bone Matrix

Bone matrix mainly composed of Type 1 Collage fibers (2 α1 chain and 1 α 2 chain), which represents 90-95% of the organic composition of the whole bone tissue and non-collagenous in-organic constituents (hydroxyapatite and other salts of calcium and phosphate). The collagen fibers are essential for the bone's tensile strength, and the non-collagenous proteins are essential for the bone's compressive strength. These effects are synergistic.

2.1.4 Classification of Bones

Bones can be classified according to the arrangement of collagen into two categories; 'lamellar' bone and 'woven' bone (fibrous bone). Lamellar bone is the main type of bone in a mature skeleton. It is composed of an organized collagen arrangement that exhibits a lamellar pattern with circular layers of collagen alternating longitudinal ones. It is stress oriented, mechanically strong, and exhibits low flexibility. In contrast, woven bone is an immature or pathologic bone composed of loosely and randomly arranged collagen, it is non-stress oriented, mechanically weak, and highly flexible. Woven bone is formed when osteoblasts produce osteoid rapidly during the production of fetal bones. In adulthood, the woven bones are produced after

a bone fracture or in Paget's disease. Later these bones are replaced with more resilient lamellar bone.

2.1.4.1 Osteoblast- Bone Formation

Osteoblasts are the cells responsible for bone formation. They synthesize and secrete the organic and inorganic constituents of bone matrix. Osteoblasts function in groups along the bone surface, lining on the layer of bone matrix that they are secreting. Osteoblasts originate from mesenchymal stem cells (MSC), which have the capacity to differentiate into osteoblasts, chondrocytes, muscle, fat, ligament and tendon cells (Bianco, Riminucci M., Gronthos, & Robey, 2001). In recent years, much progress has been made in understanding the factors that regulate the gene expression underlying the induction, proliferation, differentiation, and maturation of osteoblast (Jensen, Gopalakrishnan, & Westendorf, 2010). Recent studies have shown that deletion mutations in runt-related transcription factor 2 (Runx2) or osterix genes lead to abnormal bone development (Ducy et al., 1997). Toward the end of the matrix production, 15% of mature osteoblasts get embedded in the new bone matrix, differentiating into osteocytes, while the remaining osteoblasts stay on the bone surface and become flat lining cells.

The process of bone formation is called osteogenesis or ossification, it occurs in three subsequent phases: production and maturation of osteoid matrix, followed by mineralization of the matrix. In adults, these phases occur at the same rate, maintaining a balance between matrix production and mineralization. During osteogenesis, clusters of osteoblasts on the bone surface deposit collagen and other molecules to form an organic soft matrix referred to as "osteoid". This is followed by

another phase where there is an increase in the mineralization rate to equal that of collagen synthesis. In this phase, osteoblasts secrete alkaline phosphatase to create sites for the deposition of calcium and phosphate, allowing crystallization. Finally, the rate of collagen synthesis decreases, and mineralization continues until the osteoid becomes fully mineralized.

Osteoblasts produce a wide range of growth factors under a variety of stimuli including the insulin-like growth factors (IGF) (Canalis, J, Gabbitas, Rydziel, & Varghese, 1993), platelet-derived growth factor (PDGF) (Shikada et al., 2005) basic fibroblast growth factor (bFGF) (Globus, Plouet J., & Gospodarowicz, 1989), transforming growth factor-beta (TGF) (Canalis, J, & Varghese, 1993) and the bone morphogenetic proteins (BMP) (Chen, Zhao M., & Mundy, 2004). Osteoblast activity is known to be regulated in an autocrine and paracrine manner by a range of growth factors, whose receptors are found on osteoblasts for classical hormones such as parathyroid hormone, parathyroid hormone-related protein, thyroid hormone (Rizzoli, Poser, & Bürgi, 1986), growth hormone (Barnard, Ng Kw., Martin, & Waters, 1991), insulin (Levy, Murray, Manolagas, & Olefsky, 1986), progesterone (Wei, Leach Mw., Miner, & Demers, 1993), and prolactin (Clement-Lacroix et al., 1999) are located in osteoblasts as well. Osteoblastic nuclear steroid hormone receptors are members of the nuclear receptor subfamily 3 (NR3), which include receptors for estrogens (Eriksen et al., 1988), androgens (Colvard et al., 1989), vitamin D3, (Darwish & DeLuca, 1996) and retinoids (Kindmark, H., Johansson, Ljunghall, & Melhus, 1993).

2.1.4.2 Osteocytes

Osteocytes are the osteoblasts that have been embedded in the osteoid of the

bones. The metabolic activity of the osteoblast usually decreases once it is fully encased in bone matrix. However, osteoblasts are still capable to produce matrix proteins. The osteocytes are rich in microfilaments, which are organized during the matrix formation and before the calcification process. In addition, osteocytes form a network of thin canaliculi permeating the entire bone matrix. Interestingly, the osteocytes' morphology and functional activity vary according to the cells age. The young osteocytes have almost all the structural characteristics as the old osteocyte but with decreased cell volume and lower capacity of protein synthesis. The older osteocyte (located deep within the calcified bone) is characterized by accumulation of glycogen in the cytoplasm. During the process of osteoclastic bone resorption, these old osteocytes are eventually phagocytosed and digested (Elmardi, Katchburian Mv., & Katchburian, 1990). The exact function of these complex osteocytes network is still vague. However, it is likely that these osteocytes respond to bone tissue strain and stimulate bone remodeling activity by recruiting the osteoclasts to sites where remodeling of the bone is required (Lanyon, 1993).

2.1.4.3 Osteoclast–Bone Resorption

The osteoclast is giant multinucleated cells that reach up to 100 μm in diameter. They are derived from the hematopoietic cells of the mononuclear lineage (macrophage lineage). Mature monocytes and macrophages have the ability to differentiate into osteoclasts if the suitable microenvironment is available and prepared by bone marrow-derived stromal cells (Udagawa et al., 1990). Osteoclasts are the mediators of the continuous destruction of the bone tissue in response to different stimuli such as structural stress or the body requirement of calcium. They occupy small depressions

known as Howship lacunae on the surface of the bone. Osteoclasts are rich with Golgi apparatus, mitochondria, and transport vesicles which are loaded with lysosomal enzymes. These organelles are present in deep foldings of the osteoclast's plasma membrane in the area that is close to the bone matrix which is known by ruffled border and the surrounding zone of attachment which is known as the sealing zone. Lysosomal enzymes including tartrate-resistant acid phosphatase and cathepsin K are synthesized by the osteoclast. Then secreted via the ruffled border into the bone-resorbing compartment (Stenbeck, 2002). These lysosomal enzymes, mainly the acid phosphatase have the ability to dissolve the organic collagen and the inorganic minerals (calcium and phosphorus) of the bone. First of all, the mineralized bones are broken into fragments. Then, the osteoclast engulfs the fragments and digests them within cytoplasmic vacuoles. The liberated Calcium and phosphorus from the breakdown of the mineralized bone are released into the bloodstream. On the other hand, the unmineralized bones (osteoid) are protected from the osteoclastic resorption. This process of adhesion of the osteoclast with the bone matrix surface involves the binding of integrins expressed in osteoclasts to specific amino acid sequences on the surface of bone matrix (Davies et al., 1989). After osteoclast adhesion is complete, the binding of avb3 integrin activates the cytoskeletal reorganization inside the osteoclast (Reinholt, Hultenby K Fau - Oldberg, Oldberg A Fau - Heinegard, & Heinegard, 1990). This attachment usually occurs via structures known as podosomes, which are adhesive structures present at the ventral surface monocytic myeloid lineage cells. Through podosomes continual assembly and disassembly, they allow the movement of the osteoclasts across the bone surface during which bone resorption proceeds. Integrin

signaling and subsequent podosome formation are dependent on a number of adhesion kinases including the proto-oncogene src (Destaing et al., 2008).

Osteoclasts resorb the bone by acidification of the bone matrix. The first process during bone matrix resorption is the mobilization of the hydroxyapatite crystals, encapsulated within the sealing zone, by digestion of their link to the collagen fibers. The remaining collagen fibers are digested by proteases (cathepsins or collagenases) and the residues from this digestion process will either be internalized or transported across the cell to be released at the basolateral domain. The function of osteoclast is regulated by cytokines that act locally and by the systemic hormones. Osteoclastic receptors for calcitonin (Warshawsky H., Goltzman, Rouleau, & Bergeron, 1980), androgens (Mizuno et al., 1994),thyroid hormone (Abu, Bord S., Horner, Chatterjee, & Compston, 1997), insulin (Thomas et al., 1998), PTH (Teti, Rizzoli R Fau - Zambonin Zallone, & Zambonin Zallone, 1991), IGF-1 (Hou, Sato T., Hofstetter, & Foged, 1997), interleukin (IL)-1 (Xu et al., 1996), CSF-1, (Hofstetter et al., 1992) and PDGF (Z. Zhang, Chen J Fau - Jin, & Jin, 1998) have been demonstrated.

2.1.5 Bone Remodeling

Bone remodeling is a complex ongoing process by which old bones are continuously replaced by new tissue to repair microdamage and maintain bone strength. In other words, bone is being turned over, allowing the maintenance of the shape, quality, and size of the skeleton. The process of bone remodeling requires the interaction between different cell phenotypes and is regulated by a variety of biochemical and mechanical factors. It relies on the correct balance between bone resorption and bone formation. In the homeostatic equilibrium condition, bone

resorption and formation are balanced so that the old bones are continuously replaced by new tissue, adapting to mechanical load and strain. In 1990 Frost defined this phenomenon as bone remodeling (Frost, 1990).

The remodeling cycle is composed of five sequential phases: (1) activation, (2) resorption, (3) reversal and (4) formation and finally (5) quiescence. The first phase (activation) starts when pre-osteoclasts are attracted to the remodeling sites and fuse to form the multinucleated osteoclasts (Matsubara et al., 2012). Then, the second phase (resorption) begins when osteoclasts dig out a cavity, called a resorption pit, in spongy bone or burrow a tunnel in compact bone and calcium is released into the bloodstream to be used by different body organs. After the completion of osteoclastic resorption, there is a reversal phase when the mesenchymal stem cells, pre-cursors to osteoblasts, appear along the bone surface where they proliferate and differentiate into pre-osteoblast and prepare the bone surface for the new osteoblasts to begin the fourth phase, which is bone formation and provide signals for pre-osteoblast to mature release osteoid at the site, forming a new soft nonmineralized matrix (Delaisse, 2014). The new matrix is then mineralized with calcium and phosphorous. When the final phase is completed, the bone surface will be covered with a flattened cell lining and a prolonged resting period will begin until the next remodeling cycle is initiated (Quiescence phase). The phases of the remodeling cycle occur occurring over the course of 120–200 days in cortical and trabecular bone, respectively (Agerbaek, Eriksen Ef., Kragstrup, Mosekilde, & Melsen, 1991).

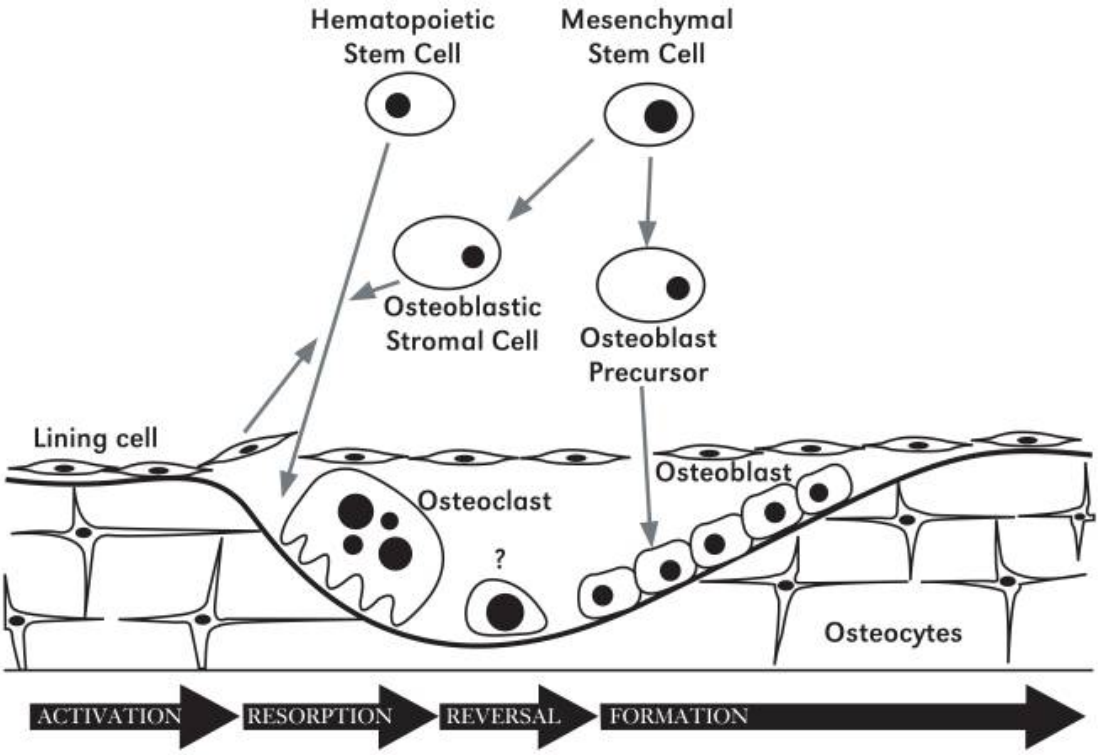

Figure 1 Representative figure of bone remodeling
Adapted from "Bone Health and Osteoporosis: A Report of the Surgeon General," by the U.S. Department of Health and Human Services, 2004.

2.2 Osteoporosis

2.2.1 Definition of Osteoporosis

Osteoporosis is one of the most common diseases that affect human bones. It is mainly characterized by decreased bone mineral density (BMD) and significantly increased predisposition to bone fracture. In other words, osteoporosis is the state of increased risk of bone fracture after minimal trauma due to decreased bone mineral density (Bonnick, 2010). The symptoms and pain associated with osteoporosis appear only when a fracture has occurred. If the treatment courses are not initiated, osteoporosis develops until there is bone breakage basically at the hip, spine, or wrist. Hip fractures interfere severely with a person's mobility and independence, while

vertebral fractures lead to height loss, stooped posture, and chronic pain (National Institutes of Health, 2017; Sözen et al., 2017).

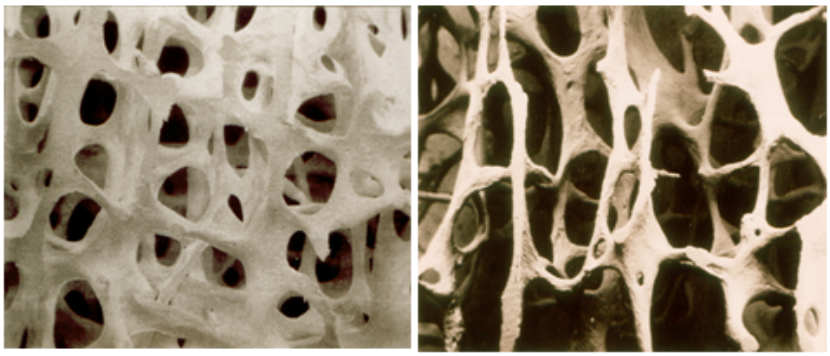

Figure 2 Normal and osteoporotic bones.

Osteoporosis affects the mechanical structural properties of bone, resulting in cortical thinning and trabecular bone loss. The structure of bone tissue is well-documented in the research literature; however, the bone remodeling process and its molecular mechanism continues to evolve.

2.2.2 Classification of Osteoporosis

Osteoporosis has been classified into two categories; primary and secondary osteoporosis. Primary osteoporosis is the most common form of diseases. It can further classify into two types; postmenopausal osteoporosis, which is type 1 and senile osteoporosis, which is type 2. On the other hand, secondary osteoporosis is characterized mainly by the clear definable etiology. Primary type 1 osteoporosis,

post-menopausal osteoporosis, develop in women when the amount of estrogen and androgen in the body decreases, which leads to an increase in the bone turnover, with resorption of bone exceeding bone formation and most importantly the loss of trabecular bone predominant compared with cortical bone. On the other hand, primary type 2 osteoporosis represents the gradual age-related bone loss found in both sexes caused by systemic senescence. It is induced by the loss of stem-cell precursors, with a predominant loss of cortical bone (Riggs Bl Fau. & Melton, 1983).

Human bone mass peak at the age of 30 for both sexes. Men and women lose bone at a rate of approximately 0.3% to 0.5% per year, respectively (Dobbs, Buckwalter, & Saltzman, 1999). After menopause, bone loss in female accelerates further at a rate of 2% to 3% per year due to estrogen deficiency up to 6 years after menopause (Dobbs et al., 1999). Due to the fact that age-related bone loss is a natural phenomenon in humans, any limitations in the individual's ability to maximize peak adult bone mass leads to an increased risk of developing osteoporosis later in life. In addition, since there are no safe and effective methods to rebuild and repair the osteoporotic skeleton, developing prevention strategies is crucial (Cosman et al., 2014; Riggs BL. & Melton LJ., 1992). Therefore, a knowledge of preventive approaches is essential, including the efficacy and safety of estrogen and progestin therapy, intake of calcium and vitamin D, exercise, bisphosphonates (Cosman et al., 2014).

2.2.3 Prevalence of osteoporosis

Osteoporosis is a chronic disease that has escalated to what is considered a major public health concern in developed countries. Approximately 200 million people suffer from osteoporosis and cause approximately 8.9 million fractures

annually resulting in an osteoporotic fracture every 3 seconds (Hernlund et al., 2013). These fractures occur mainly at the hip, vertebrae, and distal forearm (Minisola, Cipriani, Occhiuto, & Pepe, 2017) and are associated with significant morbidity, mortality, and reduced quality of life, attributed not only to the fracture itself but also to the high prevalence of comorbidities in this population of patients (Hernlund et al., 2013; Minisola et al., 2017). Between 1990 and 2000, osteoporosis caused a 25% worldwide increment in hip fractures. By 2050, the worldwide incidence of hip fracture in men is projected to increase by 310% and 240% in women, compared to rates in 1990 (Gullberg, Johnell O., & Kanis, 1997). The peak for hip or other fracture types occurs for both women and men aged 75-79 years and 50-59 years, respectively (Johnell & Kanis, 2006). Vertebral fracture due to osteoporosis is very common as well with one occurring every 22 seconds worldwide in men and women over age 50 (Johnell & Kanis, 2006).

Even though osteoporosis is a devastating disease, it remains a neglected health priority in the Arab world (Sweileh, Al-Jabi, Zyoud, Sawalha, & Ghanim, 2014). A study published on the Saudi population showed that the prevalence of osteoporosis was 34% among 5160 healthy women aged 50–79. In addition, they showed that the prevalence of osteopenia and osteoporosis among men 46.3% and 30.7%, respectively (Sadat-Ali, Al-Habdan, Al-Turki, & Azam, 2012). Another study was conducted in Jordon reported that the prevalence of osteoporosis among post-menopausal women was 12.3% (Shilbayeh, 2003). Only one study was conducted in Qatar on the prevalence of osteoporosis and they reported that the prevalence of osteoporosis among post-menopausal women is 12.3% (Bener, Hammoudeh, & Zirie, 2007). According to the

Qatar Biobank report, osteopenia represented 66% of the newly diagnosed diseases in their participants and osteoporosis represented 5% (Biobank, 2017).

The prevalence of osteoporosis, as determined by bone mineral density measurements by Dual-energy X-ray Absorptiometry (DXA), has increased massively with age. It is thought to be an unavoidable consequence of aging in women, however, extensive research in the past 2 decades reported that it is a disease that affects both women and men. A study conducted in Qatar showed that BMD values differ widely among the Gulf States, with Qatari women having low BMD values compared to Kuwaitis, but higher BMD values compared to Lebanese and similar BMD values compared to Saudis. However, femur BMD values were higher in Qatari women in the age group 40-59 years compared to Kuwaitis, Saudis, and Lebanese, but lower in the age group 60-69 years (Hammoudeh, Al-Khayarin, Zirie, & Bener, 2005).

Osteoporosis represents a major concern of the health care systems because of its growing economic burden (Kuo & Chen, 2017). In the United States, costs related to osteoporosis fractures were estimated at $13.8 billion (O'Neill & Roy, 2005).

2.2.4 Clinical Consequences and Economic Burden

The fractures that occur as complication of the sever weakening of the bones is the clinical significance of osteoporosis. These fractures may occur not only from a fall from standing height but also from simple movements such as normal lifting and bending. It is estimated that at least 1 in 3 females and 1 in 5 males will experience an osteoporotic fracture during their remaining lifetime (Kanis et al., 2000; Melton, Atkinson Ej Fau ., O'Connor, O'Fallon, & Riggs, 1998; Melton, Chrischilles Ea Fau.,

Cooper, Lane, & Riggs, 1992). About 80% of the age-related fractures in elder people are due to osteoporosis, yet, less than 30% get diagnosed and treated (Papaioannou et al., 2008). Osteoporosis takes a huge personal and economic toll. In Europe, the disability due to osteoporosis is greater than that caused by cancers (with the exception of lung cancer) and is comparable or greater than that lost to a variety of chronic noncommunicable diseases, such as rheumatoid arthritis, asthma and high blood pressure related heart diseases (Johnell & Kanis, 2006). Hip fractures are invariably associated with chronic pain, reduced mobility, disability, and an increasing degree of dependence (Keene, Parker, & Pryor, 1993). After sustaining a hip fracture 10-20% of formerly community-dwelling patients require long term nursing care (Autier et al., 2000; Cree et al., 2000; Kiebzak, Perser, Ambrose Cg., & Heggeness, 2002) with the rate of nursing home admission rising with age (Cree et al., 2000; Kiebzak et al., 2002). Similar to hip fractures, vertebral fractures can lead to back pain, loss of height, deformity, immobility, increased number of bed days, and even reduced pulmonary function (Nevitt et al., 1998). Their impact on quality of life can be profound as a result of the loss of self-esteem, distorted body image and depression (Gold, 2001; Robbins, Hirsch C., Whitmer R., Cauley J., & Harris, 2001; Tosteson et al., 2001). Vertebral fractures also significantly impact activities of daily living (Adachi et al., 2002; Hall, Criddle Ra., Comito Tl., & Prince, 1999)

The economic consequences of the morbidity and mortality rates due to osteoporotic fractures are shocking. It is estimated that the annual cost for osteoporosis treatment including; care costs, prescription of drugs, outpatient care, and indirect costs, to the Canadian health care system is about 2.3 billion dollars per year. Indeed, these costs rise when taking into consideration patients living in long

term care, as it rises to around 3.9 billion dollars per year (Tarride et al., 2012). A longitudinal study done in Canada evaluated the individual one year societal cost for a patients over 50 years suffering from hip fracture who were admitted to acute care facility to be about $21,285 after hospitalization, and $44,156 if the patient was institutionalized (Wiktorowicz, Goeree R Fau - Papaioannou, Papaioannou A Fau - Adachi, Adachi Jd Fau - Papadimitropoulos, & Papadimitropoulos, 2001). The economic burden of osteoporosis worldwide is similar to that seen in Canada. In the US, osteoporosis- related fractures are responsible for an estimate of 19 billion dollars, with men accounting for 25% of the total burden (National Osteoporosis Foundation., 2012). Researchers predict that the health care costs of osteoporosis and its fractures will grow to more than 48% by 2025 (Burge et al., 2005).

2.2.5 Diagnosis of Osteoporosis

2.2.5.1 Dual Energy X-ray Absorptiometry

DXA is an X-ray imaging technique, that is used mainly to derive the mass of one material in the presence of another through knowledge of their unique X-ray attenuation at different energies. It is a mean of measuring BMD. It is a two X-ray beam, with different levels of energy, aimed at the patient bone. By subtracting the absorption of soft tissues out, the BMD can be measured from the absorption of each beam. The DXA systems were first introduced in the late 1980s (Kelly, Dm., & Neer, 1989).

DXA scan is an extension of dual energy photon absorptiometry (DPA), which is an earlier imaging technique. However, DXA differs from the DPA technique as DPA uses a monochromatic emission from a radioisotope whereas DXA uses

polychromatic X-ray beams with different energies. DXA's main application is to measure BMD, which is essential for assessing the fracture risk and most importantly in diagnosing osteoporosis. For osteoporosis diagnosis, the lumbar spine, proximal hip and, sometimes, the distal forearm are scanned. The whole body of patients can be scanned to measure the total bone mass (Laskey, 1996). DXA scan is regarded as the gold standard for quantifying BMD in vivo and for diagnosing osteoporosis (Stuart K. Kim, 2018).

2.2.5.2 Measurements from DXA scans

DXA scan measures the content of minerals in the bone (g) representing the total mass of an area in the bones scanned. The value for BMD (g/ cm2) is derived when the content of minerals in the bones (g) is divided by the area measured. In this context, DXA scan does not measure the size of the bone or assess it is quality (e.g. microarchitecture, bone turnover) that are also main elements of bone strength.

The World Health Organization (WHO) has defined the criteria for diagnosing osteoporosis and for assessing osteoporotic fracture risk using a DXA scan. A BMD value that is more than 2.5 SDs below the optimal mean for healthy young individuals of the same race and gender defines an individual as having osteoporosis (T-score ≤ -2.5). On the other hand, osteopenia was defined by WHO as a BMD t score that ranges between 1.0 and 2.5 SDs below the optimal mean (-2.5 < T-score < -1.0).

2.2.6 Osteoporosis Risk Factors

Osteoporotic risk factors are genetic, nutritional, hormonal, and lifestyle.

win and family studies have shown that genetic factors play an important role in

regulating bone mineral density and other determinants of osteoporotic fracture risk, such as skeletal geometry and bone turnover. Genetic risk factors are more significant for osteoporosis than all the others combined—nutritional, hormonal, lifestyle and environmental factors (Cohen & Roe, 2000). The clinical definition of osteoporosis takes account of BMD, a highly heritable trait. It has been estimated in cohort studies that bone density has an estimated 0.78 heritability at the lumbar spine, and 0.84 at the femoral neck (Arden, Baker J., Hogg, Baan, & Spector, 1996). These population-based studies have demonstrated that having a first degree relative with a hip fracture predicts future hip fractures (Cummings et al., 1995).

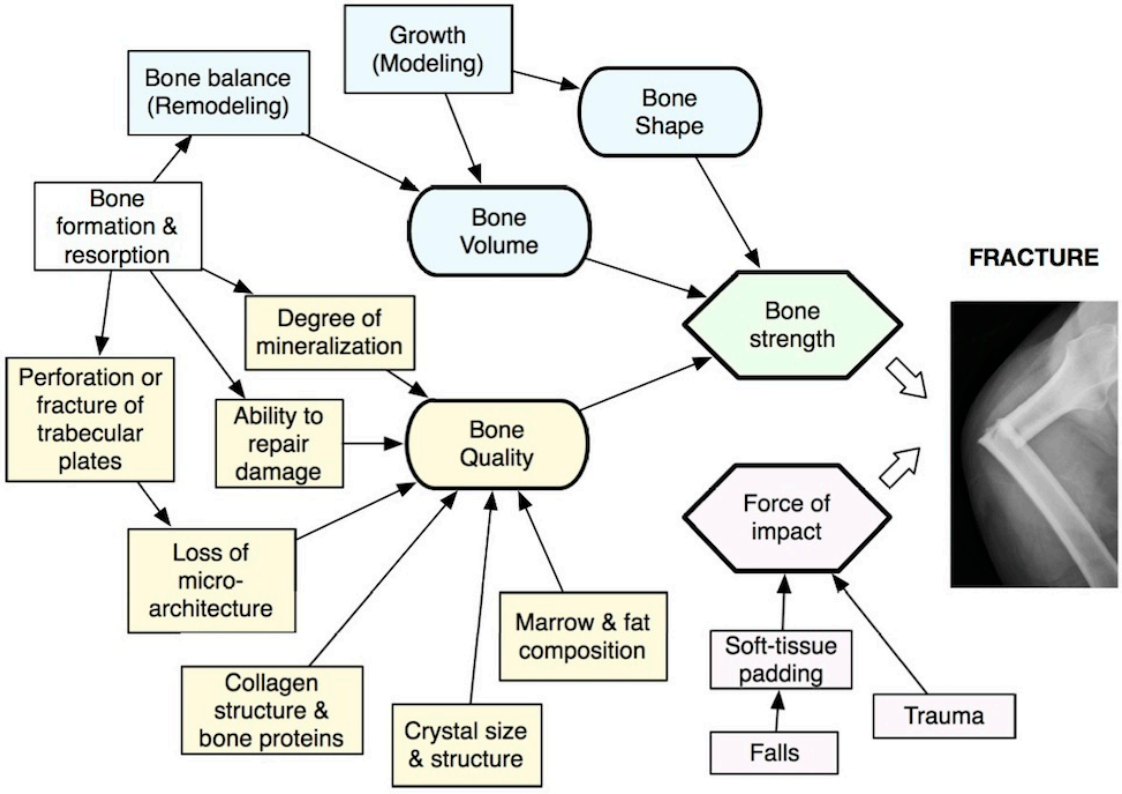

Figure 3 Risk Factors of osteoperotic fractures. (retrieved from https://courses.washington.edu/bonephys/oprisk.html).

Table 1 Risk Factors of Osteoporosis

Age Related	1. Each decade beyond the fourth decade is 1.5-fold risk 2. Reduction in absorption of calcium 3. Rise in parathyroid hormone levels 4. Decline in calcium
Genetic	1. Women more than men 2. Familial prevalence 3. High concordance in monozygotic twins
Ethnicity	1. White, Asian, Latino and Black
Nutritional	1. Low calcium intake 2. Vitamin D deficiency 3. High alcohol 4. High caffeine 5. High sodium 6. High animal protein
Lifestyle	1. Cigarette use 2. Low physical activity
Endocrine	1. Menopausal age 2. Obesity 3. Exercise-induced amenorrhea
Iatrogenic factors	1. Glucocorticoids 2. Cyclosporine 3. Anticonvulsant 4. Thyroxin 5. Aluminum 6. Lithium

2.3 Genome-Wide Association Studies (GWAS)

2.3.1 GWAS Overview

Genome-wide association study (GWAS) is an observational study of a genome-wide set used to detect associations between genetic variants and phenotypes in a sample from populations. The main goal of these observational studies is to aid in understanding the etiology and biology of diseases, which will lead to a better understanding and eventually prevention or better treatment. The fundamental principle underlying Genome-wide Association Studies (GWAS) is that it is an agnostic scan of germline variants across DNA samples from a sufficiently large number of case and control subjects using genome-wide SNP microarrays. The primary aim is to discover genomic regions that harbor genetic variation that could influence susceptibility to common complex diseases such as Osteoporosis. The path between GWAS results and the biology of the disease is not a straightforward path, because the association between a single nucleotide polymorphism (SNP) at a genomic locus and a clinical phenotype is not directly informative with respect to the target gene or the mechanism whereby the variant is associated with phenotypic differences. However, recent discoveries of new data reports, new molecular technologies, and new statistical analysis methodologies have provided the opportunities to bridge our gap in knowledge from sequence to consequence. In addition, GWASs have also been very successful in helping scientists to implement better definition of the relative role of genes and the environment in risk of developing certain diseases, which is helpful in risk prediction, which was very helpful in the development and investigating natural selection and population phenotypical differences.

2.3.4 Single Nucleotide Polymorphism

Single nucleotides polymorphisms, frequently known as SNPs, re the most common genetic variant in among a population. In other words, SNPs are single base-pair changes that occur in a high frequency in the genome (Abecasis et al., 2010). They are typically used as genomic markers; however, most of SNPs do not have a significant impact on the human physiological systems. There are usually only two alleles at a SNP locus. The term SNPs, which are common genetic variants, lie in complete contrast to rare genetic polymorphisms associated with rare genetic disorders, such as cystic fibrosis (Jackson, Marks, May, & Wilson, 2018). Such genetic disorders are caused by rare genetic variants that usually result in a deleterious alteration tin the protein function, thus, causing a disease. Variants that occur at low frequency are known as mutation.

2.3.5 common Variation

2.3.5.1 The Human Haplotype Map Project

Up to date, researchers have identified more than a thousand genes that contribute to rare, heritable genetic disorders that follow 'Mendelian' heritability. However, challenges remain in studying common disorders as they are a result of combined DNA variants interacting with environmental factors. To investigate the hypothesis of common variant lead to common disease for a specific phenotype, a systematic approach is required to interrogate the huge amount of common variation in the human whole genome. First of all, the location and the density of a common SNPs is required to identifying the genomic regions and individual sites that must be

tested by the genetic studies. Secondly, the difference in the genetic variance between different populations must be cataloged to ease the study of specific phenotypes in different populations with the accurate design The International HapMap Project was implanted to help researchers in identify common genetic variation across the genome that might help in identifying genetic risk factors for common diseases (International HapMap, 2005).

Up to date, the International HapMap Project has successfully discovered and cataloged SNPs in the European populations, the Yoruba population (African origin), Han Chinese individuals from Beijing, and Japanese individuals from Tokyo (Olivier, 2003; Ritchie et al., 2010) The project has since been expanded to include 22 human populations. In addition, the HapMap genotype data helped in the examination of *linkage disequilibrium.*

2.3.5.2 linkage disequilibrium

Linkage disequilibrium (LD) is the nonrandom linkage of nearby variants such as alleles at different loci in a specific population. This non-random association in a population is a sensitive indicator of the population genetic forces that structure their genome (Slatkin, 2008). Linkage disequilibrium term was first introduced in 1960 (Lewontin & Kojima, 1960). In the 1980s, LD attracted wide attention and the importance of LD for assisting in gene mapping became widely evident. LD throughout our whole genome represent the population's different history, the pattern of geographic subdivision and most importantly, the breeding system in the population (Slatkin, 2008). In this context, it reflects the population's history of natural selection, mutation and other factors that impacted the evolution of gene-

frequency. In addition, linkage disequilibrium can be influenced by many contributing factors, including natural selection, genetic recombination rate, mutation rate, the system of mating in the population, genetic drift, the structure of the population, and finally, genetic linkage (Barton, 2010).

The recent systematic studies of the population's common genetic variation including GWAS are aided by the fact that individuals who carry a particular SNP at a particular locus often carry a specific SNP at another nearby locus, which can be predicted (Stadler et al., 2010). A particular combination of alleles at different loci that are inherited as a block along a chromosome is known as a haplotype (International HapMap, 2005).

This linkage of common variants along a chromosome that is inherited together came to existence due to the shared ancestry of the chromosomes. When a new variant arises through mutation; SNP, insertion or deletion, it initially occurs in a unique chromosome, which is marked by a distinct combination of genetic variants. Subsequently, the natural process of recombination and mutations during the division process act to erode this mutation, however, the process is very slow. This association between the mutations and the haplotypes have served as a tool for human genetic research by finding its association to a haplotype and then identifying the causal variant of the phenotype. The association studies were first conducted on the *HLA* region identify causal genes for certain Mendelian diseases such as cystic fibrosis (Kerem et al., 1989) and diastrophic dysplasia (Hastbacka et al., 1992)) and complex disorders such macular degeneration associated with age (Edwards et al., 2005; Haines et al., 2005; Klein et al., 2005).

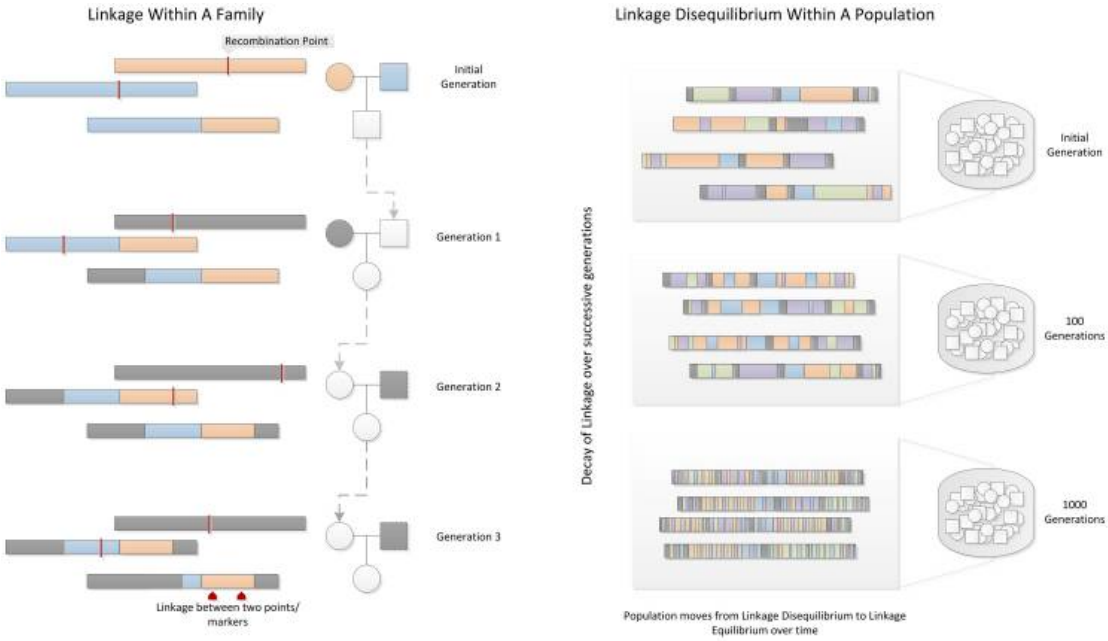

Figure 4 Representative figure demonstrating linkage disequilibrium.

Linkage occur when the two markers remain linked together after meiosis instead of being separated by recombination (red line). Retrieved from (Bush & Moore, 2012).

2.3.5.3 Indirect Association

Two possible outcomes will be present from the existence of Linkage Disequilibrium in a genetic association study. The first outcome is that the SNP that influences the biological mechanism and ultimately leads to the phenotype under investigation is genotyped in the GWAS directly (*direct association)*. In this case, the directly genotyped SNP is known as a *functional SNP*. The second possible outcome is that the functional SNP was not directly typed, however, a tag SNP which is in a high LD haplotype with the functional SNP is typed (*indirect association)*. A tag SNP is a representative SNP in a region of the genome with high linkage disequilibrium that represents a group of SNPs.

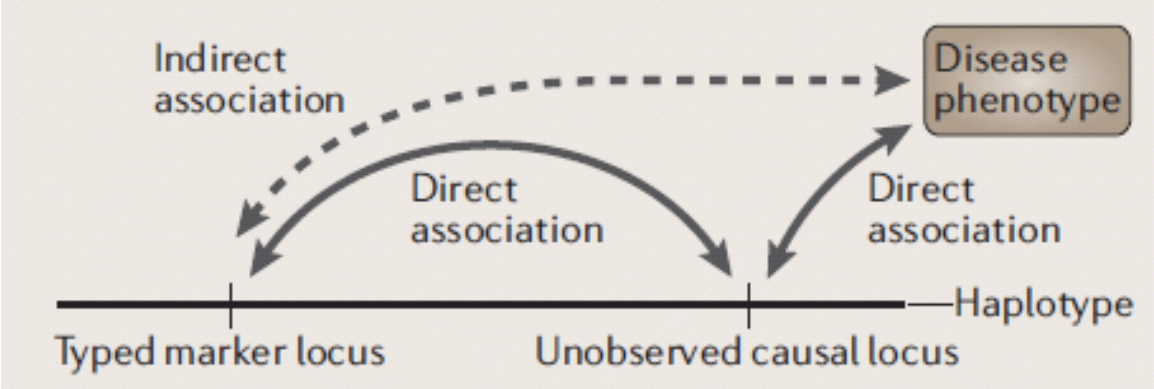

Figure 5 Indirect association testing in GWAS.

GWAS investigates the "indirect" association between the typed marker locus on the array and the disease phenotype. The unobserved causal locus, which is directly associated disease phenotype, is directly associated with the typed marker locus.

2.3.7 Genotyping Technologies

GWA studies are now feasible by the availability of microarrays for assaying millions of SNPs at a reasonable cost. Two primary microarrays platforms have been used by most of the GWA studies, which are Illumina (San Diego, CA) and Affymetrix (Santa Clara, CA). Both microarrays offer different approaches for measuring SNP variations. To elaborate, the Affymetrix platform microarray prints short DNA sequences as a spot on the chip that recognizes a specific SNP allele in which the alleles are detected by differential hybridization of the sample DNA. On the other hand, the Illumina microarrays platform uses a bead-based technology with slightly longer DNA sequences to detect alleles. The Illumina chips offer better specificity; however, it is more expensive to make.

Another important consideration aside from which platform should be used for the GWAS, is the SNPs that each platform has selected for assay. This is a very important

aspect that should be kept in mind depending on which specific human population will be studied.

2.3.5 *Study Designs*

2.3.5.1 Case-Control design

The case-control studies are the classical epidemiological designs used. Basically, it is composed of subjects who have the disease and investigate if there are any specific characteristics of these patients that differ from those who do not have the disease. In GWAS of case-control design, scientists compare the frequency of alleles or genotypes between the cases and controls. A difference in the frequency of an allele or genotype of the SNP under investigation between the cases and control groups indicates that the genetic marker may increase the risk of the disease or likelihood of the trait. Haplotypes can also show an association with a specific trait. One of the earliest successes in the genetic field was identifying a single base nucleotide mutation in the non-coding region of the APOC3 gene (apolipoprotein C3 gene) which is found to be associated with higher risks of hypertriglyceridemia and atherosclerosis (Rees A., Shoulders, Stocks, Galton, & Baralle, 1983) using a case-control design.

The major drawback for case-control study design is that genotype and haplotype frequencies vary between ethnic or geographic populations. If the recruited case and negative control populations were not well matched and counted for different ethnicities or geographic origin, then the false positive association can occur because of the confounding effects of population stratification. Another major limitation of this design is the lack of well-defined case and control groups especially in complex diseases.

2.3.5.2 Family-based design

The family-based association study designs main aim was to avoid any potential confounding effects that could result from population stratification. This study design recruits the parents or an unaffected sibling as controls for the case. Two tests are most commonly used by this study design, which are the transmission disequilibrium test (TDT) and the haploid-relative-risk (HRR). Both of these tests measure the association of a specific genetic marker in the affected families by transmission from parent to offspring. If an allele increases the risk of having a disease, then that allele is expected to be transmitted from parent to offspring more often in populations with the disease.

2.3.5.3 Quantitative trait association

From statistician prospective, GWA studies with quantitative phenotype are preferred because they can easily improve the power of genetic effect detection, and often have a more interpretable findings in contrast to case-control studies. For some quantitative traits, genomic risk SNPs have already been identified. As an example, the high-density and the low-density lipoproteins levels are very strong markers for cardiac disease. Thus, genetic studies of cardiac disease can be conducted by investigating these levels as a "quantitative trait" (Bush & Moore, 2012). The genetic variants that influence the levels of these markers will have a clear interpretation. These quantitative traits being easily measurable have made GWAS of blood lipids easily conducted in many cohort studies from different populations. In addition, GWA studies of quantitative traits could be combined to produce a massive meta-analysis

study that is extremely well-powered as done by (Teslovich et al., 2010). However, some diseases don't have a very well-established quantitative measure. Here is when the individuals can be classified as either "disease" or "healthy" which is a binary categorical method. Though, you have to consider the vast difference in measurement error between different studies associated with classifying individuals as either a "case" or a "control" versus precisely measuring a quantitative trait. However, classifying the individuals as cases or controls doesn't necessarily means that the GWAS will be unsuccessful. For instance, multiple sclerosis is a complex clinical phenotype that is often diagnosed over a long period of time by ruling out other possible conditions. GWAS of multiple sclerosis has been enormously successful, implicating more than 10 new genes for the disorder (Habek, Brinar Vv., & Borovecki, 2010).

2.3.6 Quality Control

2.3.6.1 Sex inconsistencies in GWAS sample and chromosomal anomalies

One of the most important steps that must be implemented in any GWA study as a Quality control is to check any potential mismatch in the sample identification, which is typically a result of mishandling errors (S. Turner et al., 2011). The easiest method to identify any sample handling disputes is to check the reported gender of each participant against the predicted gender by the genetic data. The -- check-sex command in PLINK implement the heterozygosity rates of the X-chromosome to determine the sex. Then, the software identifies participants for whom the sex recorded in the ped-file is not compatible with the predicted sex based on genetic data. If there are any discrepancies in the data, the available questionnaires should be

reviewed to make a determination whether there was a sample handling error. In addition, by checking the X chromosome heterozygosity rates, any sex chromosome anomalies such as Kleinfelter syndrome, can be identified (S. Turner et al., 2011).

IID	PEDSEX	SNPSEX	STATUS	F	Explanation
1	1	1	OK	0.98	Male
2	2	2	OK	0.03	Female
3	2	1	PROBLEM	0.99	Recorded female, genetically male
4	1	2	PROBLEM	0.02	Recorded male, genetically female
5	2	0	PROBLEM	0.28	Likely a female with sex chromosome anomaly (e.g. XX/XO mosaic, loss-of-heterozygosity on X)
6	1	0	PROBLEM	0.35	Likely a male with sex chromosome anomaly (e.g. XXY or XX/XY mosaic)

Figure 6 Representative table showing an example of output from command –check sex using PLINK. The Table is retrieved from (Stephen Turner et al., 2011).

2.3.6.2 Sample Relatedness

Another method to examine the sample identity and the pedigree identity simultaneously is by comparing genomic data with the self-reported relationships reported in the questionnaires if available. Using the dense marker genomic data, it is easy to compute kinship estimates and identify related individuals in the study using the command " --genome" in PLINK. In addition to reporting the kinship between the participants, this step is essential for calculating the proportion of loci where two individuals share alleles that are identical by descent (IBD). In this context, individuals who share two alleles that are IBD at every locus in the DNA are either monozygotic twins or one sample was just processed twice. Thus, the individuals who share zero alleles that are IBD at every locus are unrelated. On the other hand, the individuals who share one allele that is IBD at every locus are parent-child samples. In general, siblings can share zero, one, and two alleles. Using the data of kinship, the proportion of loci of individuals who share one allele of IBD can be plotted against the proportion of loci sharing zero alleles IBD (Figure 7). Furthermore, discovering the relatedness in GWAS sample is essential not only for discovering any mishandling errors but to reveals any cryptic relatedness that may be present. As shown in figure 7, individuals who were unrelated represented by the black points or distantly related represented by the blue points line up along the diagonal line in the plot. Individuals along this diagonal line represent up to fifth degree kinship relation. Thus, if those individuals' samples were treated as independent samples in any further analyses in the GWAS, it would result in increased type I and II errors. In this case, mixed model regression analytical method (Aulchenko, de Koning, & Haley, 2007) must be implemented instead of simple linear or logistic regression.

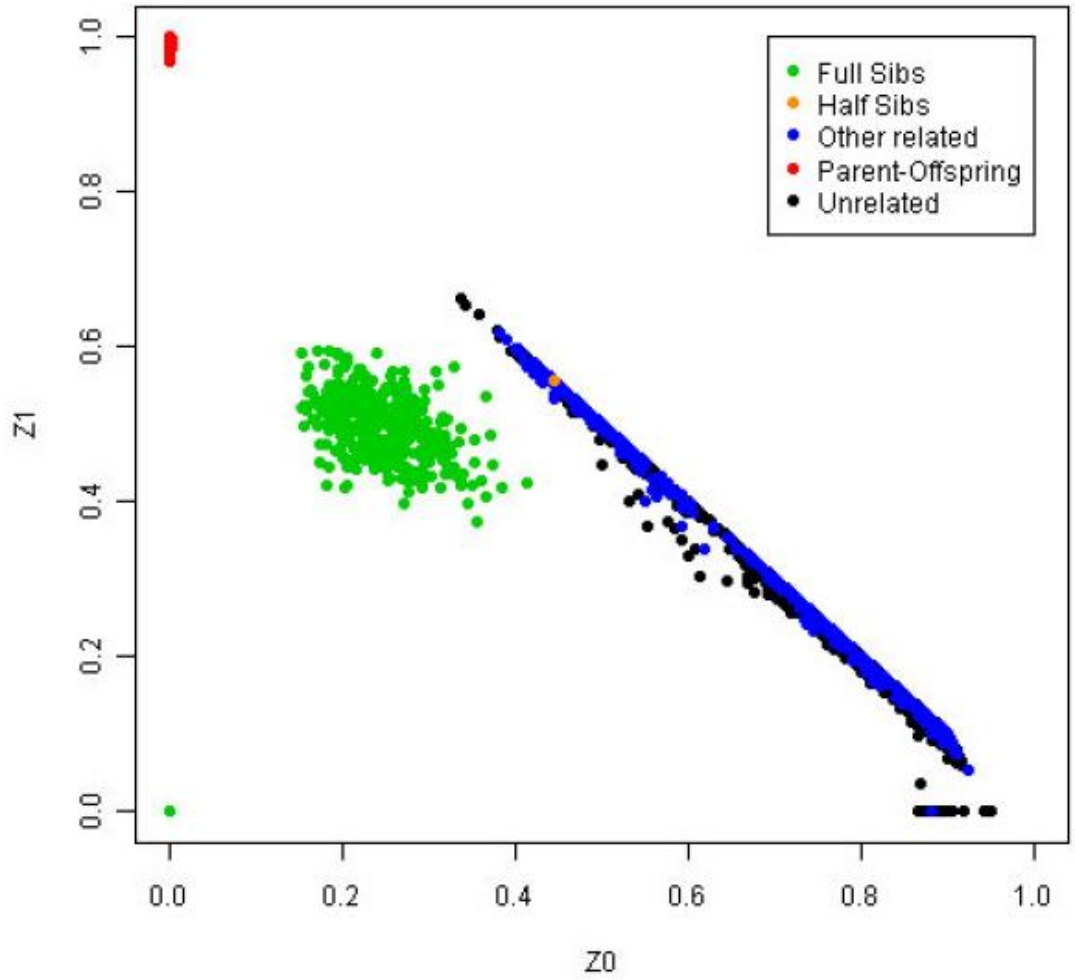

Figure 7 Quality Control: Sample Relatdness Plot.

Representative figure of individuals by their degree of relatedness: the proportion of loci of individuals who share one allele of IBD (Z1) can be plotted against the proportion of loci sharing zero alleles IBD (Z0). The figure is retrieved from (Stephen Turner et al., 2011).

2.3.6.3 Covariate Adjustment

Statistical tests should be implemented to adjust and account for factors that are known to influence the trait under investigation, such as sex, age, study site, and

known clinical covariates (S. Turner et al., 2011). Covariate adjustment is very critical to reduces any spurious associations that results due to any sampling artifacts or biases in study design. However, the covariate adjustment comes at the price of using additional degrees of freedom which may have a negative impact on the statistical power of our results. One of the most important covariates that should be considered in a GWAS is population stratification discussed in section 2.3.6.4.

2.3.6.4 Population Substructure

Population stratification is the presence of systematic differences in the allele frequency between subpopulation groups in the population usually due to different ancestral origins (S. Turner et al., 2011). Population stratification is the prime concern of researchers carrying out association studies(Hellwege et al., 2017). Normally, there are known differences between different ethnic groups in phenotype prevalence, which lead to allele frequencies being highly variable across human subpopulations, meaning that in a sample with multiple ethnicities, ethnic-specific SNPs will likely be associated to the trait due to population stratification. Failure to control the population stratification in GWA study may result in confounding and spurious apparent associations, causing a study to fail for lack of significant results or wasting of resources following false positive signals (Cardon & Palmer, 2003).

Different statistical tools have been developed and implemented into GWAS software to help in detecting population stratification in the GWAS sample and to adjust for it. The genomic control (Devlin & Roeder, 1999; Reich & Goldstein, 2001) helps in controlling the population stratification by estimating an inflation factor. After that, the software helps in adjusting all of the test statistical results downward

by this factor. To account for population stratification, the ancestry of each sample in the dataset is measured using one of two software, either STRUCTURE (Price, Patterson Nj., et al., 2006) or EIGENSTRAT (Hochberg & Benjamini, 1990) methods that compare genome-wide allele frequencies to those of HapMap ethnic groups. To ease the process of quality control of GWAS, another method has been developed to adjust for the population structure with the large sample size and thousands of SNPs. Eigenstrat analysis (Price, Patterson, et al., 2006) uses the Principal Components Analysis to specifically detect and adjust for population stratification in large sample sizes used in GWAS in a computationally effective approach. Researchers preferred this method over the stratified analysis technique as the combined sample often yields more powerful statistical tests, even after adjusting for significant eigenvectors (F. Zhang, Wang, & Deng, 2008). Eigensoft is available open-source found online for free. The software analysis will result in the computation of 10 principal components. If any of these eigenvectors are significantly associated with the phenotype under investigation, these eigenvectors should be adjusted to correct for any bias due to population stratification (S. Turner et al., 2011).

2.3.6.5 The efficiency of genotyping / call rate

The call rate or genotyping efficacy is an indicator of the quality of markers used. The falling of a large number of SNP assays in an individual DNA sample is an indicator of poor-quality DNA that could lead to sporous genotype calling. These samples should be excluded from further analysis. The recommended threshold for genotyping call rate is 98-99 % (S. Turner et al., 2011). However,

the threshold should be applied based on the balance of minimizing the number of samples excluded and maximizing the efficacy of genotyping. Figure 8 represents the proportion of samples (red and blue lines) or markers (SNPs) by (green line) remaining after determining the different threshold. Checking the genotyping efficiency can be done using the command " --missing" in PLINK, which in turn generate a file showing the missingness rate for each individual, which Is basically the proportion of SNPs failing for each individual in the study. The poor quality SNPs are removed based on call rate threshold using the command " –geno" , followed by a threshold for a lower limit of missingness (Ex. "--geno 0.01 "would remove all SNPs showing more than 1% missing). After that, the samples that are below the applied threshold can be excluded from any downstream analyses using the command " –mind" in PLINK (S. Turner et al., 2011).

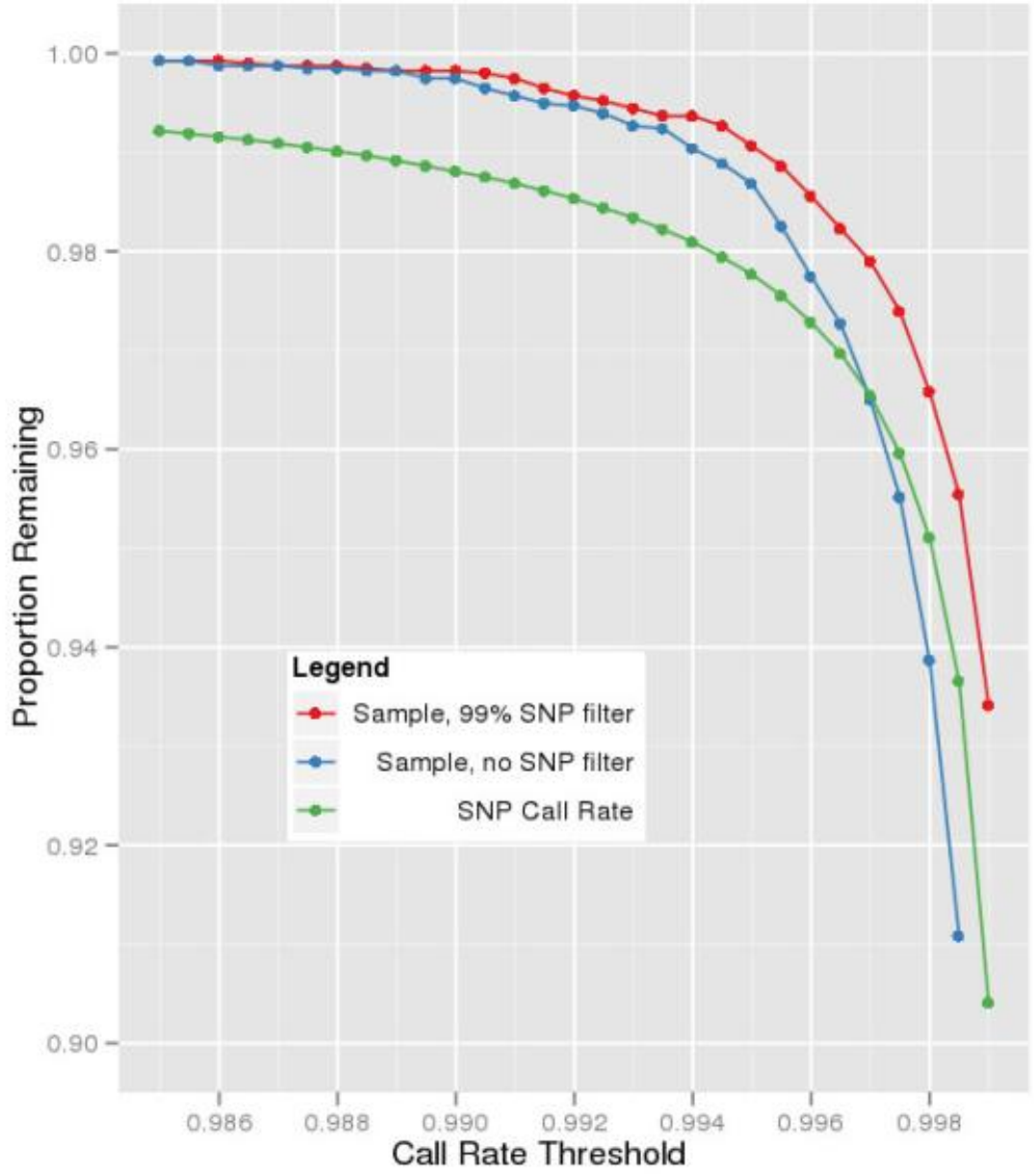

Figure 8 Representative figure of adjusting the call rate efficacy threshold.

The green line represents the SNPs remaining after excluding SNPs below the call rate efficacy threshold. The blue line represents the number of samples remaining after exclusion, whereas the red line represents the number of samples that remains after applying the 99% call rate threshold to exclude poor quality markers. The figure is retrieved from (S. Turner et al., 2011).

2.3.6.6 Minor Allele Frequency (MAF)

One of the most important steps in Quality control and cleaning the GWAS data is to filter SNPs based on the MAF due to the fact that the statistical power is extremely low for rare SNPs. Thus, it is recommended to remove any extremely rare SNPs. Similar to the call rate, the MAF threshold is chosen based on the sample size and the effect sizes that is expected. Power calculation software such as CaTS Power (Skol, Scott, Abecasis, & Boehnke, 2006) can simplify the power calculations for GWAS and inform the investigator of the MAF in which the statistical power of the study becomes extremely low. Using the command "(--freq) in PLINK, MAF can be reported for each SNP, and using the command "--maf", SNPs can be removed from the downstream analysis.

2.3.6.7 Hardy-Weinberg Equilibrium (HWE)

The final step of Quality Control is checking for HWE. By assuming that the sample is under HWE, we can estimate the allele and genotype frequencies from one generation to the next. Any departure from this estimated could be due to an error in the data, population stratification or an actual association with the phenotype under investigation (Wittke-Thompson, Pluzhnikov, & Cox, 2005). HWE can be assessed in the GWAS sample using the command "(--hardy) in PLINK. As mentioned before, any departure from HWE could be due to genotyping error, however, it could be due to a true association with phenotype as well. Thus, SNPs that are extremely out of HWE should not be excluded but flagged for further analysis after GWAS is performed. The quantitative allelic signals at a marker can be utilized to investigate the technical origin behind the deviation from HWE. The null allelic markers can

produce multimodal genotype clusters in the heterozygote and the homozygote clusters as represented by Figure 9 or can produce an unexpected number of samples with no signal as represented by figure 10.

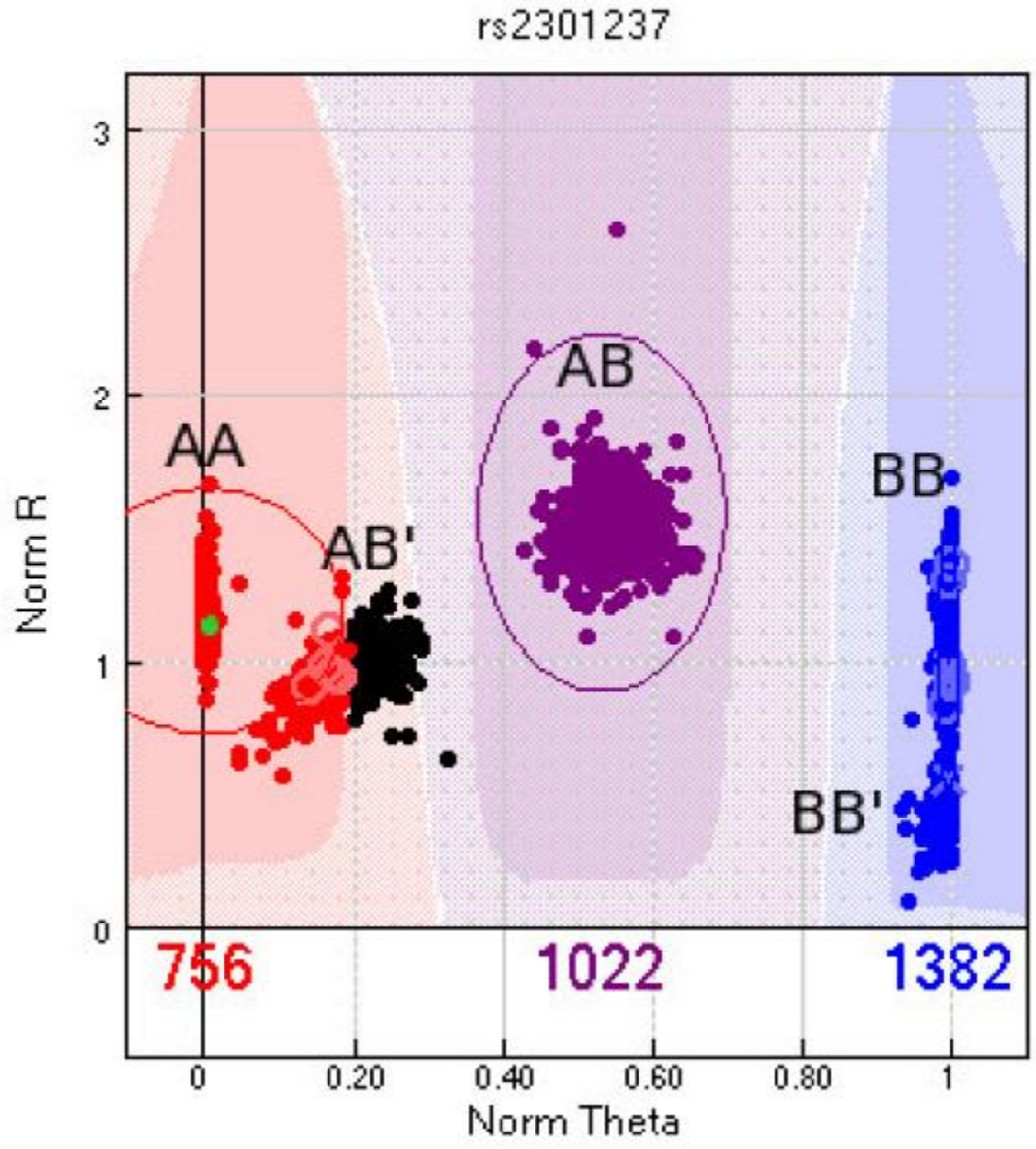

Figure 9 Representative figure of deviation from HWE.

Individuals with AB and BB are divided into sub-clusters AB and AB', BB and B B', while AA cluster are unaffected individuals. The AB/AB' split results in some AB samples miscalled as AA (diagnosed by Mendelian inconsistencies in the genotypes), as well deviation from HWE due to excess homozygosity. The figure is retrieved from (Stephen Turner et al., 2011).

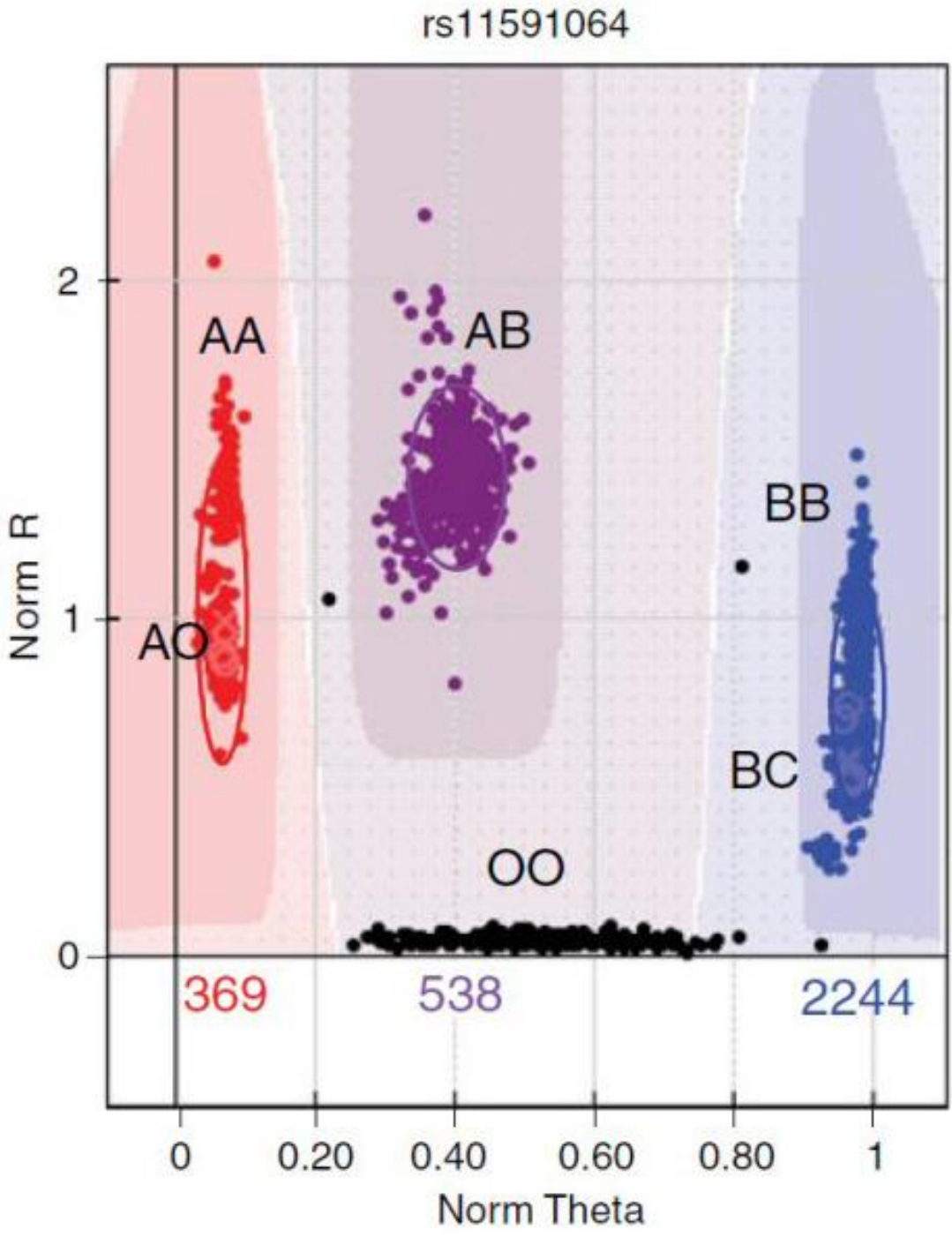

Figure 10 Representative figure of unexpected number of clusters causing diviation from HWE.

Hemizygous individuals cluster at AO and BO. Individuals with homozygous deletions cluster at OO and their genotype calls are missing. The figure is retreived from (Stephen Turner et al., 2011).

2.3.8 Replication and Meta-analysis

2.3.8.1 Statistical Replication

The gold standard for validating any GWAS is the replication of the results in an additional sample that is independent of the first study (Kraft, Zeggini, & Ioannidis, 2009). That said, there are a variety of criteria have been established that aid in defining the "replication" of a GWAS result. This was the subject of a National Human Genome Research Institute working group, which outlined several criteria for establishing a positive replication (Studies et al., 2007). These criteria are discussed in the following paragraphs

Replication studies should have sufficient sample size to detect the effect of the susceptibility allele (Y.-J. Liu, Papasian, Liu, Hamilton, & Deng, 2008). Usually, after the detection of a significant association of a trait with a particular SNPs, estimates of penetrance and allele-frequency parameters for the associated variant facilitate the planning of replication studies (Bush & Moore, 2012). However, when the effects reported in the initial GWAS suffer from the winner's curse, where the detected effect is likely stronger in the GWAS sample than in the general population due to ascertainment bias. Ascertainment bias is a systematic distortion in measuring the true frequency of a phenomenon due to the way in which the data are collected (Nicod & Largiadèr, 2003). To account for this bias, replication studies should include samples that ideally are larger to account for the over-estimation of effect size (Kraft et al., 2009).

One of the most important aspects, in order to have a successful replication study, is that it should be conducted in an independent dataset of samples that are drawn from the same population as the discovery GWAS (Rietveld et al., 2014), in

order to confirm the effect of the allele variation in the original target population. Once the effect is confirmed in the target population, other populations may be sampled to determine if the SNP has an ethnic-specific effect. Replication of a significant result in an additional population is sometimes referred to as *generalization*, meaning the genetic effect is of general relevance to multiple human populations(Bush & Moore, 2012).

2.3.9 Rare variants and unexplained Heritability

The underlying rationale for GWAS is the so-called CD/CV (Common Disease/Common Variation) hypothesis (Schork, Murray, Frazer, & Topol, 2009). Based on improvements in technology, the results from the Human Genome Project and other international consortia efforts, commercial genome-wide SNP microarrays were designed to capture most common variation across the genome. A model for dissecting the constituents of the genetic architecture for complex diseases is illustrated by the bivariate plot of the risk allele frequency versus the genetic effect strength for the genetic variants (Figure 11). In the future, the chance of detecting additional common variants with high risk is small because they should have been found by many existing well-powered scans. On the other hand, uncommon variants with low risk will be difficult to study for most diseases due to the lack of power in many studies, where extremely large sample sizes are necessary to detect them.

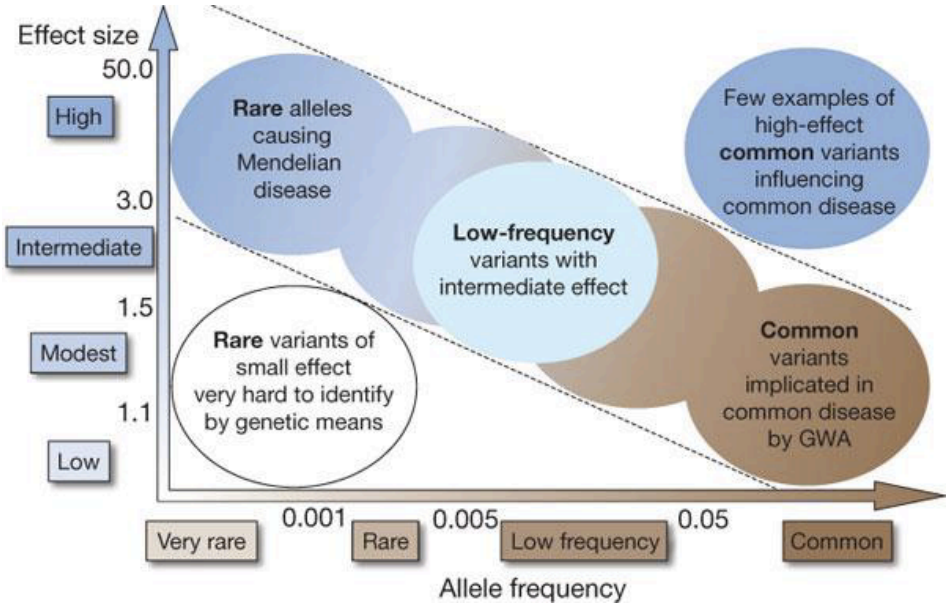

Figure 11 Model of genetic architecture of complex diseases.

The feasibility of identifying genetic variants by risk allele frequency and strength of genetic effect (odds ratio) is presented. The x-axis represents the frequency of the risk allele, and the y-axis represents the effect size. In essence, GWAS mainly detects common variants (MAF> 5%) with moderately low effect sizes (OR <1.5). Retrieved from (Manolio et al., 2009)

Chapter 3: Methodology

3.1 Ethical Statement

All samples used in this study were obtained from Qatar Bio-Bank (QBB). Written informed consent has been already taken from all participants' prior study. Ethical approval was obtained from QBB; QBB IRB MOPH Assurance: MOPH-A-QBB-000222.

3.2 Subjects

Our study cohort started by including 6000 healthy Qatari participants. We excluded 3000 participants who were related. 3000 unrelated healthy Qatari participants aging from 18 to 70 years old remained after exclusion of related participants. Subjects with chronic diseases and conditions that might potentially affect bone mass, structure, or metabolism were excluded. Hence, subjects with chronic disorders involving vital organs (heart, lung, liver, kidney, brain), history of diabetes, high cholesterol level, high blood pressure, kidney diseases, stroke, arthritis, osteoporosis, fractures, Parkinson disease, thyroid disease, hysterectomy, Hodgkin lymphoma, breast, prostate, and lung cancers were excluded from our study. The purpose of these exclusions was to minimize the influence of known environmental and therapeutic factors on bone mineral density and skeletal system.

3.3 Questionnaires

The background questionnaire was self-administered and designed to obtain personal socio-demographic information as well as information regarding the

participants' health history and known risk factors for osteoporosis (e.g., smoking, personal medical history including fracture history, routine consumption of calcium-fortified foods, smoking, alcohol, and caffeine consumption, and general physical activity habits). All participants included in this study were supplied with the main questionnaire from which the demographic variables including age, height, weight, and BMI were obtained. In addition, data regarding the medical history of the participants were collected using a Nurse questionnaire and if the participants are taking any medications or supplements. The nurse questionnaire contained more than 8900 different medication and supplements. Finally, data regarding the participants' diet (smoking status, specific diet, if milk is included in their diet) were collected using the Diet Questionnaire.

3.4 DXA Scan

The BMD data for 3000 individuals were obtained from Qatar Biobank. A trained and certified technician in QBB measured the BMD using the gold standard DXA (GE Lunar Prodigy, Madison, WI). 7 BMD measurements were obtained which are the BMD of the whole body, the lumbar spine (L1-L4), the pelvis bone, the trunk and the femoral neck, femoral troch and femoral ward were all measured. The DXA scan results were reported as absolute values of BMD (g/cm2) and Young adult T-Score**.** The T-scores were used to categorize the BMD values as normal (T-score ≥ - 1), osteopenia (T-score < -1.0 to > -2.5), or osteoporosis (T-score ≤ -2.5) based on the diagnostic criteria of WHO for osteoporosis. The same DXA machine was used for screening of all participants to avoid any variation in BMD measurements.

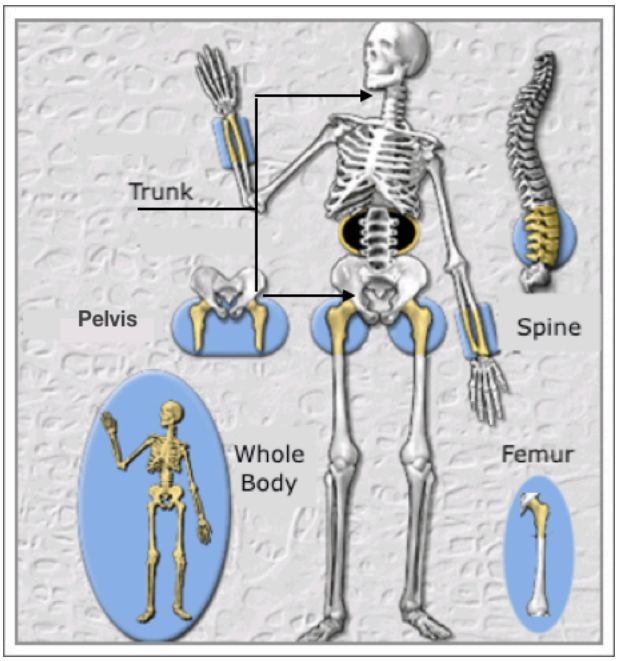

Figure 12 Scheme of body anatomy showing body parts considered in the study. We included 7 BMD measurments; Whole body, Trunk, Pelvis, Spine, Fermoral torch, Femoral Ward and Fermoral Neck.

3.5 Genotyping Method

Blood samples were collected for the 3000 participants, and DNA was extracted using standard Puregene DNA isolation kit by Gentra Systems, Inc., Minneapolis. Samples were aliquoted into barcoded 96-well plates for genotyping. Whole exome sequencing was performed in Qatar Biobank on the blood samples of the 3000 participants.

3.6 Quality Control

Extensive quality control (QC) measures were used to ensure the highest quality data possible (Figure 13). Data from Bead Chips with less than 95% genotyping yields

or heritability error rates greater than 1% were excluded. Samples were excluded when a call rate was less than 95% and a workflow-related problem was suspected. The genomic sample was excluded if the call rate was low (i.e. <95%) due to poor sample quality. SNPs were also excluded when the minor allele frequency (MAF) was less than 1%. Samples with suspected labeling errors (e.g. females with Y chromosome SNP genotypes, duplicate samples with different genotypes, unknown identification numbers) that could not be resolved were also excluded. Variants showing a departure from HWE $P<10^{-6}$ were also excluded (Rafiq et al., 2014; van Leeuwen et al., 2014; Yodsurang et al., 2018). We used genome wide significance P value of 10^{-8} (Fadista, Manning, Florez, & Groop, 2016; Panagiotou & Ioannidis, 2012) ,in which SNPs identified below this regression p value were considered significantly associated with BMD. QC analyses were conducted using PLINK 2.0 toolset (Harvard.edu) (Purcell, Neale, et al., 2007).

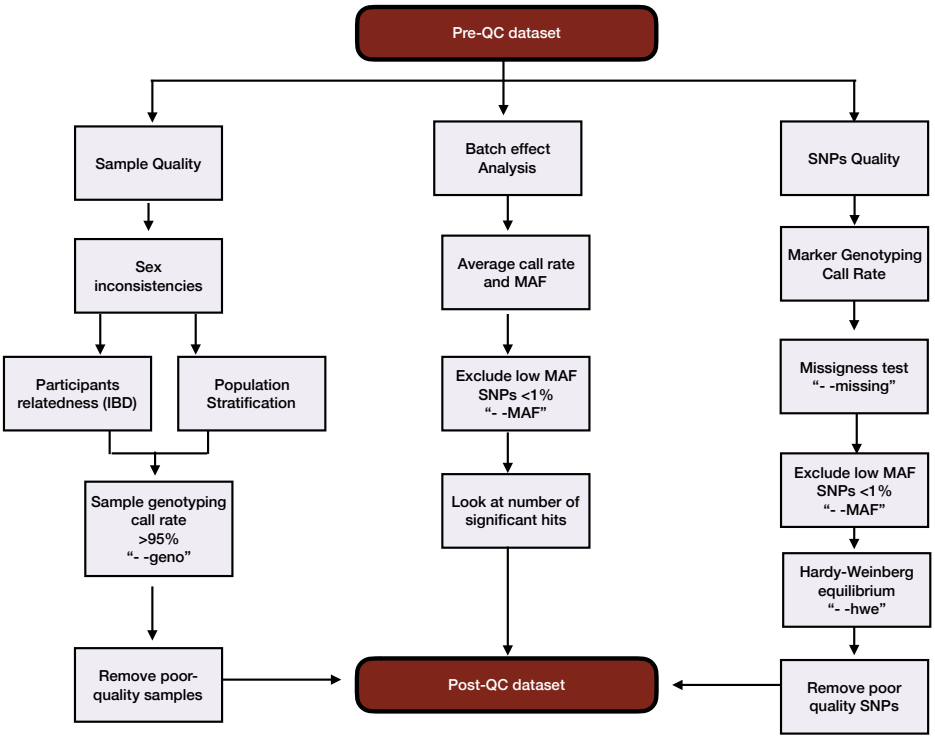

Figure 13 A flow chart overview of the entire GWAS- Quality Control process along with the commands used in PLINK 2.0.

3.7 Statistical Analysis

GraphPad Prism version 7 and IBM SPSS were used for the phenotypic data analysis. The comparison of means for two independent groups was analyzed using the Student's t-test. The biochemical variables were approximately normally distributed, and sample SDs were similar.

Before any of the association analyses, we used principal component analysis implemented in EIGENSTRAT (Price, Patterson Nj., et al., 2006) to correct for any potential population stratification that may lead to spurious association results. The first ten principal components emerging from the EIGENSTRAT analyses, along with

sex, height, weight, and age, were used as covariates to adjust the raw BMD values in each sample.

3.8.1 Linear Model of Regression

The association between a SNP and the polymorphic phenotype was assessed using sex-specific, age-standardized residuals that were analyzed under an additive genetic model. To adjust for the substructure of the Qatari population, we included the four most important principal components (PC), derived from a Multi-Dimensional Scaling analysis of IBS distances using the PLINK (Purcell, Neale B., et al., 2007) software as covariates in the regression analysis. To test for association with a quantitative trait, linear regression was performed by PLINK to obtain the regression coefficient and Wald test asymptotic p-value.

Chapter 4: Results

4.1 The baseline characteristics of participants.

The demographic and biochemical characteristics of the participants are summarized in Table 2 and Table 3, respectively. A total of 3000 participants were included in the present study. We included 1442 (48%) male and 1558 (52%) female in our study. The male participants in our study population had significantly higher weight and height metrics compared to the female participants ($p < 0.001$). Each participant included in our study was supplied with a questionnaire from which the demographic data was collected. Moving on to the biochemical characteristics represented in table 3, there was a significant difference between both genders in all the tested parameters. However, most of the variables were in the normal reference range. Interestingly, the mean of Dihydroxy Vitamin D Total in Qatari men and women was 18.13 ng/ml and 19.16 ng/ml respectively, which is considered below the normal reference range of vitamin D. Detailed vitamin D assessment will be discussed in the below sections.

4.2 The prevalence of vitamin D insufficiency and deficiency.

According to the national institute of health, vitamin D level < 20 ng/ml was considered insufficient and vitamin D level < 12 ng/ml was considered deficient (Harbolic, 2017). 47.2 % and 37.2% of the male and female participants respectively had vitamin D insufficiency. In addition, 22.6% and 24.9% of the male and female participants respectively had vitamin D deficiency. Figure 14 shows that Qatari females (<30 years old) were more prone to vitamin D deficiency compared to older

Qatari women. A similar trend was observed in young Qatari males as well. Thus, we investigated the frequency of taking vitamin D supplements in the 3000 participants.

4.3 The frequency of taking vitamin D supplements.

There was a positive trend observed between taking vitamin D supplements and age (Figure 15). 63% of females in the age group 51-70 were taking vitamin D supplement, while only 24.6% of young females in the age group 18-30 were taking vitamin D supplements (Figure 15). similar results were observed with males, the frequency of taking vitamin D supplements increased with age in Qatari males. However, at a much less frequent compared to females.

4.4 The gender effect on BMD variation

Our inclusion criteria for our study was healthy Qatari participants with age between 18- 70 years old. The total, femoral troch, femoral upper neck and femoral ward for the 3000 participants were assessed using the DXA Scanner. We divided the data from the whole sample population into male and female groups to examine any gender associations with BMD (Table 4). The gender-specific group means were compared using a Student's t-test. The female participants had lower BMD values compared to the male participants ($p<0.01$) as shown in table 4.

4.5 Correlation between Vitamin status and BMD in the Qatari population

We tested the correlation between average t-score of BMD of the total body and femur (upper neck, troch, and ward) with the vitamin D status. Pearson correlation test was used to find the correlation between vitamin D and BMD T- Score for both genders. We divided the data of the 3000 participants into 3 categories; Normal,

vitamin D insufficient and vitamin D deficiency according to WHO criteria discussed earlier. A positive correlation was observed, meaning, individuals who were vitamin D deficient showed lower BMD T-score and vice versa. However, this correlation was not significant ($P>0.05$) for males, but slightly significant for females ($p<0.05$) as shown in figure 16. Thus, we further tested if the BMD results showed any improvement if the individuals were taking vitamin D supplements. No significant difference was observed between both groups in all the tested parameters; whole body, femoral troch, femoral upper neck, and femoral ward BMD as shown in Table 5. For the female whole-body BMD, significant difference was observed between females taking vitamin D supplements and females who are not taking any vitamin D supplements.

4.6 Variants associated with BMD at genome-wide significance level

The genomic data for 3000 participants were obtained from Qatar Biobank. The DNA variants were restricted to those with a minor allele frequency > 0.01. We started with 93,546,361SNPs identified. However, only 1,084,750 SNPs passed the extensive quality control steps as mentioned in the Methodology. We accounted for the population structure in the cohort. In addition, we excluded related individuals from the study as related individuals are also genetically related. As an example, siblings from same biological parents are about 50% genetically related, which can drive the association results and cause bias by creating a subpopulation in the cohort. The genetic association was assessed using PLINK2.0. (Purcell, Neale, et al., 2007). The linear model of regression was conducted in our study including, gender, age, weight, BMI, vitamin D and the leading 10 genomic principal components as covariates

(Kemp et al., 2017; Stuart K. Kim, 2018). A deviation from the random distribution was observed in the Q-Q plot (figure 18) of all tested SNPs, indicating that some SNPs might have an association with BMD. We carried out the linear regression analysis on whole body BMD and 6 BMD measurements obtained from different bones and parts of the bones; the Spine, Pelvis, Trunk, Femur Troch, Femur-Upper neck, Femur-Ward. SNPs were considered as novel variants when no previous reports are found in different genomic databases; Open Target Genetics, UK Biobank cohorts, GEFOS cohorts, GRCh37 ensembl, Clinvar, gnomAD and NCBI data base. In addition, all information regarding the SNPs identified was retrieved from GRCh37 ensembl, NCBI and Clinvar.

In total, 19 autosomal variants were associated with BMD with genome-wide significance ($p<5x10^{-8}$) (Fadista et al., 2016; Panagiotou & Ioannidis, 2012) as shown in table 6. These 19 SNPs are overlapping 20 genes and overlapping 53 transcripts as shown in figure 19. 3 SNPs are present on chromosome 7, 2 SNPs are on chromosome 11, 2 SNPs on chromosome 18, 2 SNPs on chromosome 1, two SNPs on chromosome 2 and one SNP present on each of chromosome 17,21,22,1,9,6, 19,3,4.

Fifteen SNPs were found to be associated with low Total body BMD. Two of these 15 SNPs we highly significant in whole body, Spine and Trunk BMD, which are rs4727924 (P value = $1.86x10^{-11}$) and rs2536172 (P value= $2.75x10^{-11}$). Both of these SNPs lies within chromosome 7 at q31.31 in *FAM3C* gene which is overlapping with the *WNT16* gene. The third SNP on chromosome 7, which is rs1839588 (P value= $2.00x10^{-08}$) is associated with Pelvis BMD. It is an intronic variant of *SFRP4* gene and is found on band p14.1. 12 of the SNPs are intronic variants on different chromosomes; in *MALAT1* gene on chromosome 11, *CRYBB2P1*gene on

chromosome 22, *FASLG* gene on chromosome 1, *FAM189A2* gene on chromosome 9, *SAG* gene on chromosome 2, *PIGN* gene on chromosome 18, *RP11-15A1.7* gene on chromosome 19, and finally *LSAMP* gene on chromosome 3. The rest four SNPs are intergenic variant, thus, not overlapping any coding genes.

Four SNPs were observed at low BMD of trunk, rs4727924, rs2536172 mentioned earlier, and 2 SNPs on chromosome 4 and 1 which are intronic variant overlapping *BMPR1B* gene and an intergenic variant respectively. Finally, no significant SNPs were observed to be associated with BMD of the Upper neck or Troch BMD of the femur. A summary of all observed significant SNPs (above the significance level of 10^{-08}) are present in Table 6. Figures 20 to 26 represents the Manhattan plots for all SNPs tested for the 7 BMD traits.

4.7 Validation of previous associations with BMD and osteoporosis

The results obtained from our study was compared to previous genome-wide association studies of osteoporosis and BMD. One study performed Meta-analysis for SNPs associated with BMD of the femur upper neck and the lumbar spine. They identified 62 SNPs showing the genome-wide association significance (K. Estrada et al., 2012). From these identified SNPs, 57 SNPs were replicated by a study conducted by the UK biobank. Another study conducted a Meta-analysis on genomic variants associated with BMD and they and identified 9 SNPs with genome-wide significant associations (Moayyeri et al., 2014). All of the previous SNPs were replicated by the UK Biobank study. A third GWA study was conducted on BMD and bone fracture, 13 SNPs were identified (Mullin et al., 2017) and they were replicated by (S. K. Kim, 2018). Finally, Kim et al., identified 899 loci associated with heal BMD, 266 loci are

replicated loci and 613 are new loci with p-values below $1.6x10^{-4}$ (i.e. Bonferroni cutoff) (S. K. Kim, 2018). In our study, we identified 19 SNPs to be associated with spine, pelvis and femur BMD with genome-wide association significance. Six of which those SNPs were replicated by the UK Biobank and GEFOS studies and 13 are new SNPs with no previous data reported.

Table 2 The demographic characteristics of Qatari male and females in our cohort.

Parameter	Male (± SD)	Female (± SD)	P-Value
Age (years)	36.59 (±10.6)	36.19 (±11.6)	>0.329
Height (cm)	172.37 (±6.3)	158.09 (±5.9)	<0.001
Weight (kg)	84.84 (±17.9)	71.18 (±15.9)	<0.001
BMI	28.29 (±5.4)	28.48 (±6.2)	>0.391

The age, height, weight, and BMI of both genders were summarized and compared. N = 3000, Group specific means are compared using Student's t-test.

Table 3 Biochemical Characteristics of 3000 Qatari participant.

Parameter	Male (± SD)	Female (± SD)	P-Value
Sodium (mmol/L)	140.89 (±2.1)	140.03 (±2.1)	P<0.001
Potassium (mmol/L)	4.346 (±0.3)	4.31(±0.3)	P<0.002
Chloride (mmol/L)	101.09 (±2.1)	101.38 (±2.1)	P<0.001
Ca corrected (mmol/L)	2.2754 (±0.8)	2.283 (±0.8)	P<0.001
Phosphorus (mmol/l)	1.1343 (±0.2)	1.1701 (±0.1)	P<0.001
Vitamin D (ng/ml)	18.13 (±11.4)	19.16 (±11.6)	P<0.015

Table 4 Summary of BMD T-score measurements in participants.

	Total Cohort Mean (SD)	Male Mean (SD)	Female Mean (SD)	P-Value
Whole Body	0.39 (±1.15)	0.64 (±1.16)	0.17 (±1.10)	P<0.001
Femoral Torch	-0.37 (±1.177)	-0.59 (±1.25)	-0.67 (±1.01)	P<0.001
Femoral Upper Neck	0.06 (±1.31)	0.37 (±1.427)	-2.25 (±1.11)	P<0.001
Femoral Ward	-0.69 (±1.22)	-0.47 (±1.33)	-0.89 (±1.08)	P<0.001

All BMD measurements are reported in T-score. The female subjects had lower BMD values compared to male subjects. N = 3000, Student's unpaired t-test

Table 5 Vitamin D supplement effect on Bone Mineral Density

	Whole Body	Femoral Torch	Femoral Upper Neck	Femoral Wards
T score for the BMD in Qatari Male				
Vitamin D Supplements	0.67 (±1.2)	-0.11 (±1.2)	0.31 (±1.4)	-0.56 (±1.4)
No vitamin D Supplements	0.63 (±1.2)	-0.05 (±1.2)	0.38 (±1.42)	-0.45 (±1.3)
P Value	>0.6560	>0.6560	>0.8723	>0.2625
T score for the BMD in Qatari Female				
Vitamin D Supplements	0.33 (±1.1)	-0.61 (±0.9)	-0.31 (±1.1)	-0.95 (±1.04)
No vitamin D Supplements	0.12 (±1.0)	-0.69 (±1.0)	-0.19 (±1.1)	-0.87 (±1.0)
P Value	0.0014	>0.2382	>0.9439	>0.2218

Table 6 Genome-Wide Significant SNPs for Whole body, Lumber spine, Pelvis, Trunk and Femoral upper neck, Femoral troch and Femoral ward.

Chr	RefSNP	Position	Band	Ancestor Allele	Effect Allele	P-Value	Gene	Phenotype reported	Study	P-Value	Sample Size
Whole Body BMD											
7	rs4727924	121031879	q31.31	C	T	$1.86x10^{-11}$	FAM3C	BMD	UK Biobank	$4.13x10^{-160}$	194,398
7	rs2536172	120997560	q31.31	A	T	$5.75x10^{-08}$	FAM3C/WNT16 intronic variant	Fractured/broken bones	UK BioBank	$2.09x10^{-180}$	194,398
11	rs202070768	65273453	q13.1	T	C	$1.30x10^{-09}$	MALAT1	No Previous Data	N/A	N/A	N/A
17	rs554808159	61978607	q23.3	C	T	$4.25x10^{-09}$	intergenic variant	No Previous Data	N/A	N/A	N/A
21	rs374876997/rs5533335180	26574354	q21.2	T	deletion	$8.44x10^{-09}$	Splice variant	No Previous Data	N/A	N/A	N/A
22	rs489125	25911056	q12.1	G	A	$2.25x10^{-08}$	CRYBB2P1: intronic variant	Various diseases	UK Biobank	N/A	N/A
1	rs867865671	172626211	q24.3	A	G	$3.03x10^{-08}$	FASLG: intronic Variant	No Previous Data	N/A	N/A	N/A
9	rs73455199	71961260	q21.12	A	G	$3.38x10^{-08}$	FAM189A2: Intron Variant	No Previous Data	N/A	N/A	N/A
6	rs367949909	132861904	q23.2	T	C	$4.94x10^{-08}$	intergenic variant	No Previous Data	N/A	N/A	N/A
18	rs190738498	59831463	q21.33	G	A	$5.71x10^{-08}$	PIGN: Intron Variant	Various disease	UK biobank	N/A	N/A
2	rs1050627711	233310901	q37.1	C	T	$6.08x10^{-08}$	SAG: Intron Variant	No Previous Data	N/A	N/A	N/A
18	rs191429075	59790212	q21.33	C	T	$6.93x10^{-08}$	PIGN: Intron Variant	Various disease	UK biobank	N/A	N/A
1	rs866548296	234651783	q42.2	C	T	$7.77x10^{-08}$	intergenic variant	No Previous Data	N/A	N/A	N/A
19	rs149339318	44503670	q13.31	TA	deletion	$8.35x10^{-08}$	RP11-15A1.7	No Previou	N/A	N/A	N/A

							: intronic variant	s Data			
3	rs142479295	117374777	q13.32	T	dup t	9.68x10^{-08}	LSAMP: intronic variant	No Previous Data	N/A	N/A	N/A
BMD of the Spine											
7	rs4727924	121031879	q31.31	C	T	1.63x10^{-08}	FAM3C/WNT16 intronic variant	BMD, Fracture risk	UK BioBank, GEFOS	5.56x10^{-160}, 4.39x10^{-10}	194,39, 335,587
7	rs2536172	120997560	q31.31	A	T	5.75x10^{-08}	FAM3C/WNT16 intronic variant	Fractured/broken bones	UK BioBank	2.09x10^{-180}	194,398
BMD of the Pelvis											
7	rs1839588	37979888	p14.1	C	T	2.00x10^{-08}	SFRP4/ intronic variant	BMD	UK BioBank	2.70x10^{-24}	194,398
BMD of the Trunk											
7	rs4727924	121031879	rs4727924	C	T	1.63x10^{-08}	FAM3C/WNT16 intronic variant	BMD, Fracture risk	UK BioBank, GEFOS	5.56x10^{-160}, 4.39x10^{-10}	194,39, 335,587
7	rs2536172	120997560	rs2536172	A	T	5.75x10^{-08}	FAM3C/WNT16 intronic variant	Fractured/broken bones	UK BioBank	2.09x10^{-180}	194,398
4	rs1050715238	95757373	q22.3	A	G	4.77x10^{-08}	BMPR1B/ Intronic Variant	No Previous Data	N/A	N/A	N/A
11	rs371319602	96953896	q21	A	C	6.00x10^{-08}	intergenic variant	No Previous Data	N/A	N/A	N/A
BMD of the Femoral Ward											
2	rs62150773	109035597	q12.3	G	C	1.72x10^{-08}	intergenic variant	No Previous Data	N/A	N/A	N/A

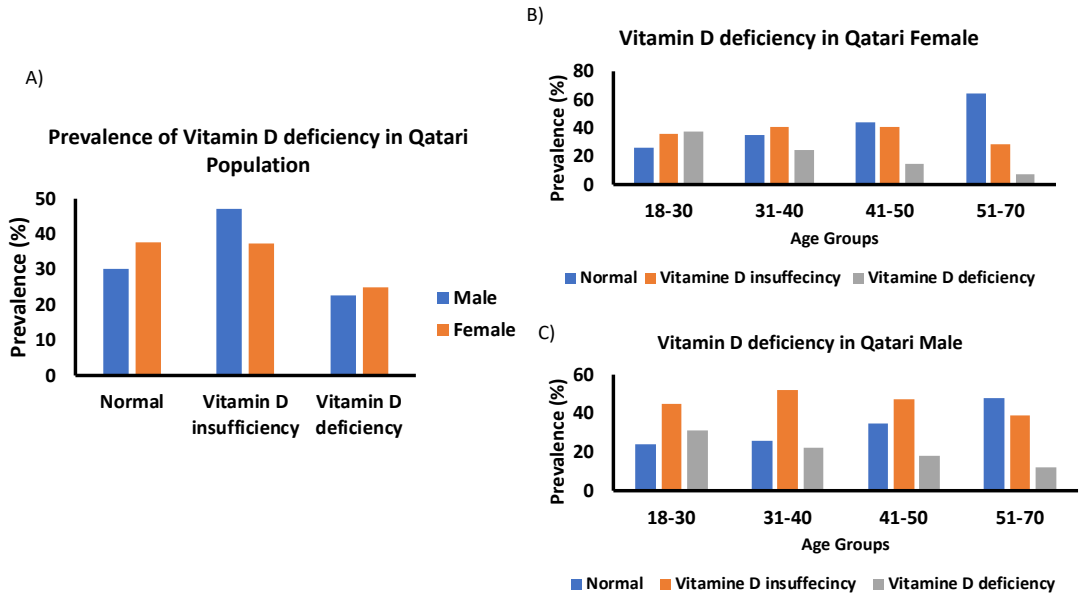

Figure 14 A)The prevelence of Vitmaine D insuffecincy and deficincy in the Qatari population. B) The prevelance of Vitamin D deficiencey in Qatari Females against diffeent age groups C) The prevelance of Vitamin D deficiencey in Qatari against diffeent age groups.

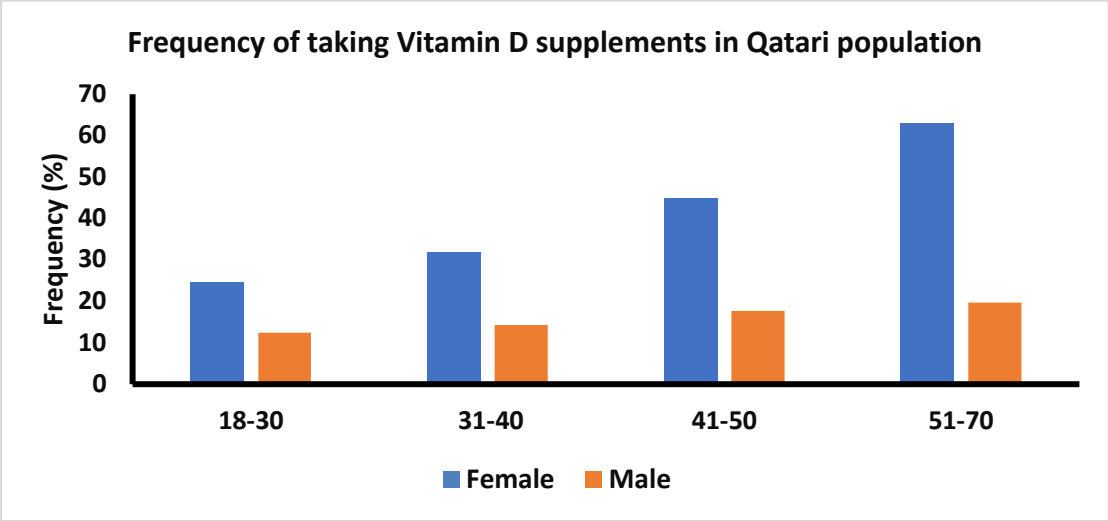

Figure 15 The frequency of taking vitamine D supplements by Qatari females and males. The frequency of taking Vitamin supplements increased with age with both gender.

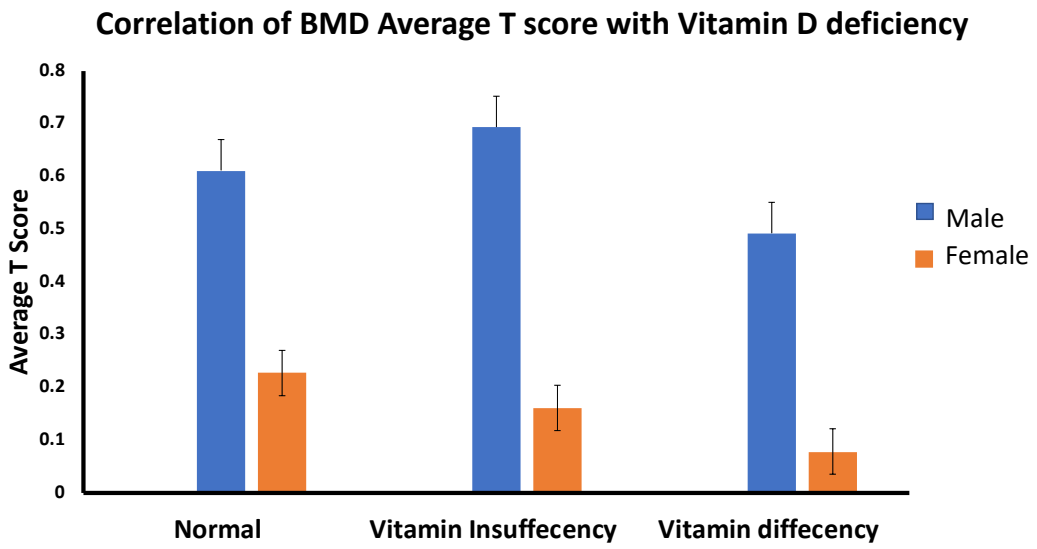

Figure 16 Correlation of vitamin D status with BMD average T score. Pearson correlation was used to find the correlation between vitamin D and Average T score. No correlation was observed between Average T score and Vitamin D in male ($P>0.09$). A positive correlation was observed between Average T score and Vitamin D in females ($p<0.05$).

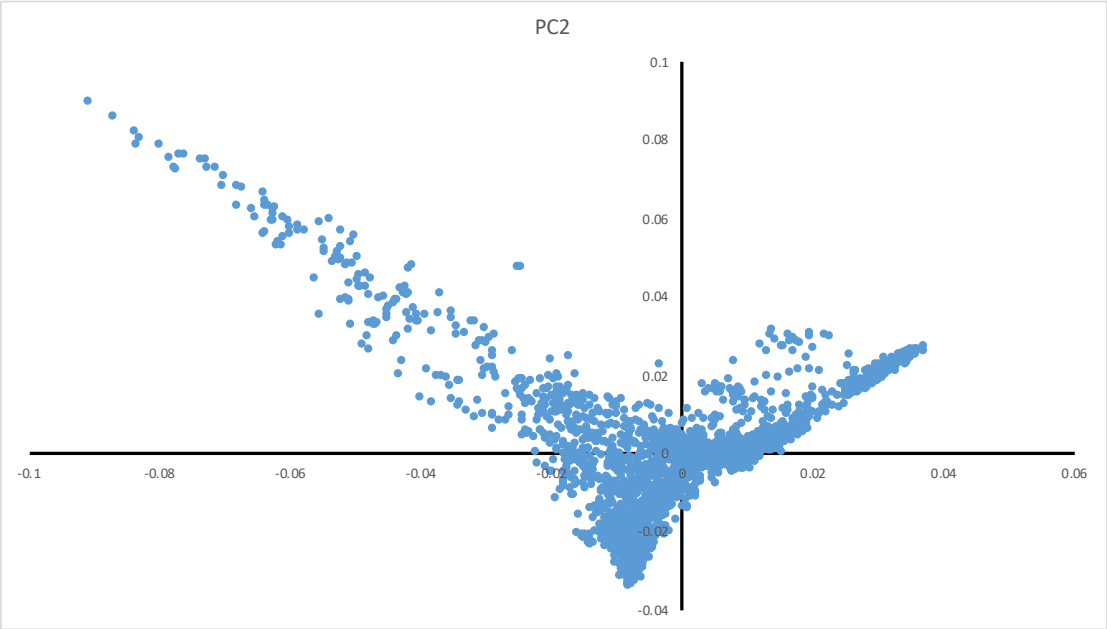

Figure 17 Principle componant analysis of the study cohort. Our cohort shows that the Qatari Population consist of 3 main clusters originating from Arabian origin, Persian origin and african admixture.

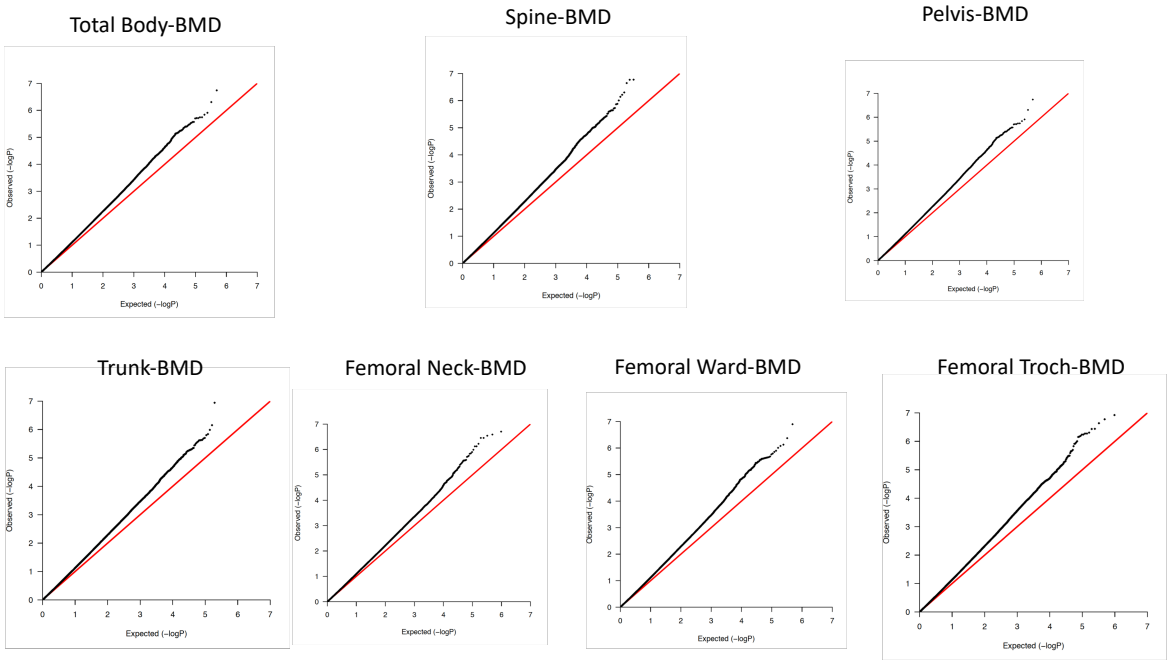

Figure 18 GWAS, quantile-quantile plots (QQ-plots) for the 7 phenotypic traits tested.

Category	Count
Variants processed	19
Variants filtered out	0
Novel / existing variants	0 (0.0) / 19 (100.0)
Overlapped genes	20
Overlapped transcripts	113
Overlapped regulatory features	4

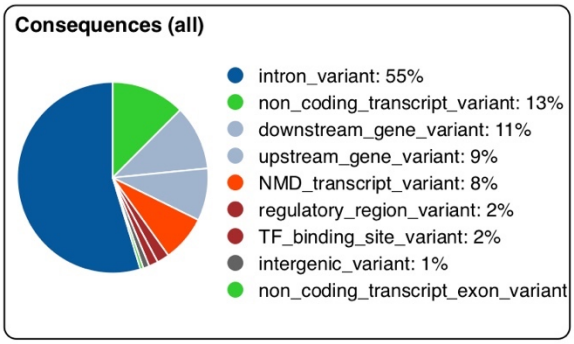

Figure 19 Summary of Significant SNPs obtained from GWAS. The figure was constructed using GRch37, Ensembl. We identified 19 variants overlapping 20 genes.

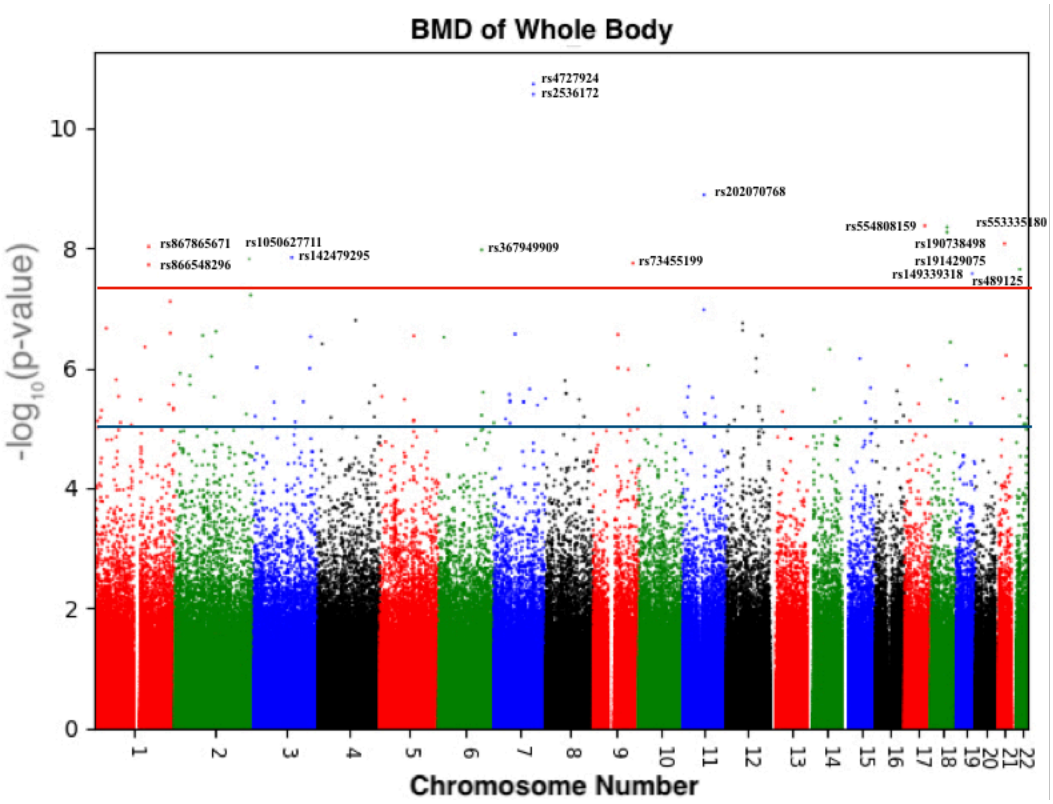

Figure 20 Manhatten plot representing genome-wide association results for whole body BMD of 3000 participents. The red line shows the genome wide significance threshold ($p<5x10^{-8}$). 15 SNPs were identified from Whole body's BMD to be significantly associated with BMD.

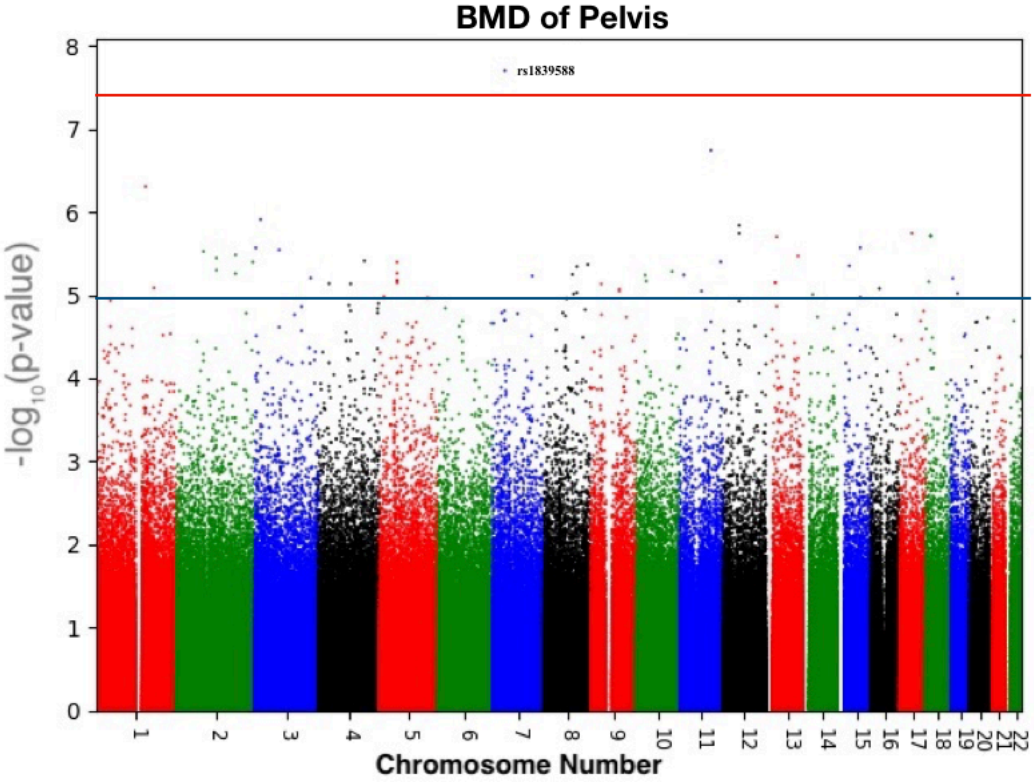

Figure 21 Manhatten plot representing genome genome-wide association results for BMD of Pelvis of 3000 participents. The red line shows the genome wide significance threshold ($p<5x10^{-8}$). 1 SNP was identified from pelvis' BMD to be significantly associated with BMD.

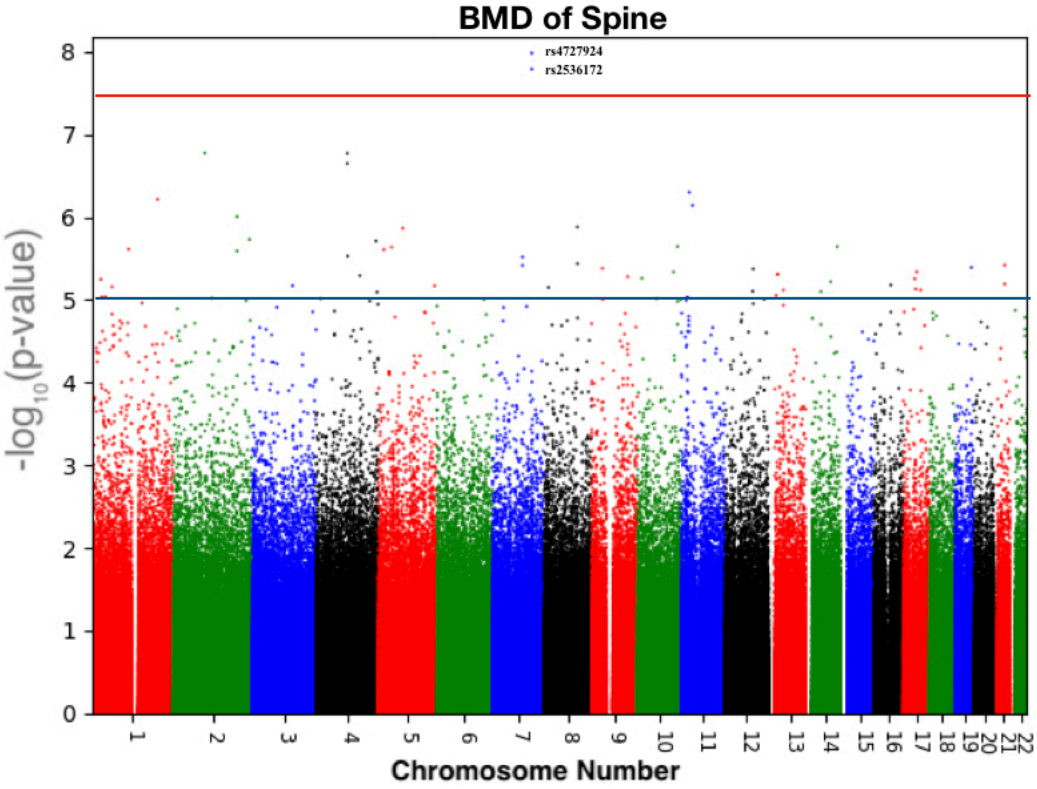

Figure 22 Manhatten plot representing genome-wide association results for BMD of Spine of 3000 participents. The red line shows the genome wide significance threshold ($p<5x10^{-8}$). 2 SNPs were identified from Spine's BMD to be significantly associated with BMD.

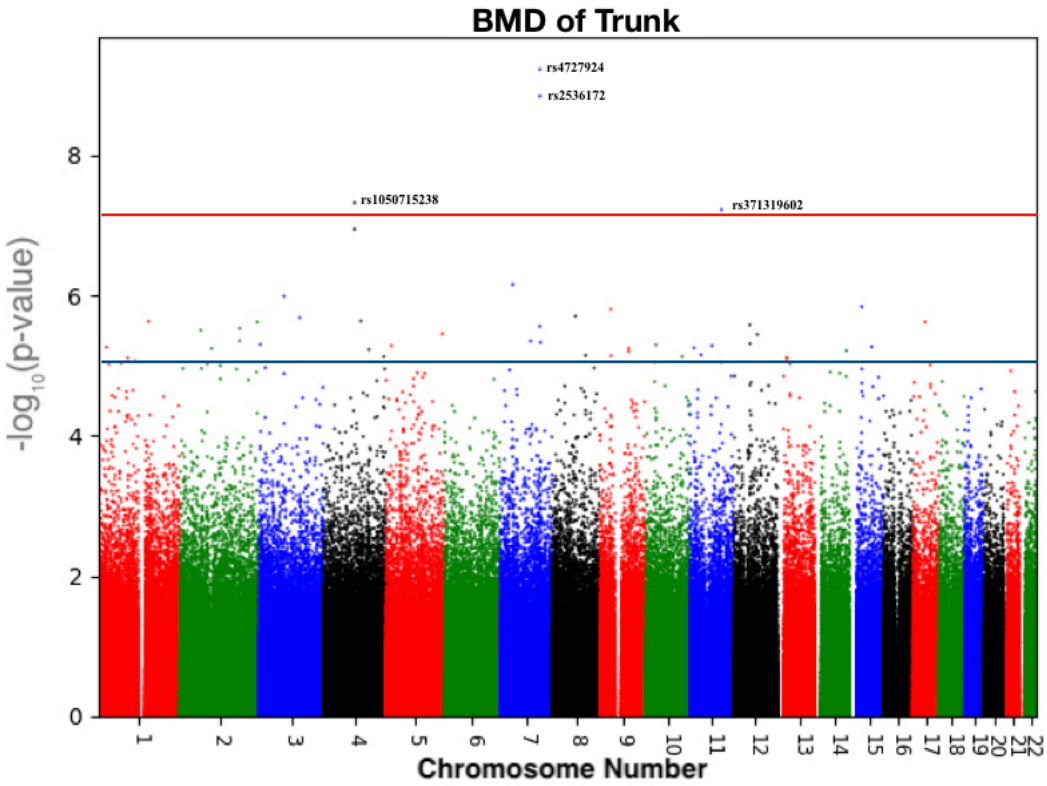

Figure 23 Manhatten plot representing genome-wide association results for BMD of Trunk for 3000 participents. The red line shows the genome wide significance threshold ($p<5x10^{-8}$). 4 SNPs were identified from Trunk's BMD to be significantly associated with BMD.

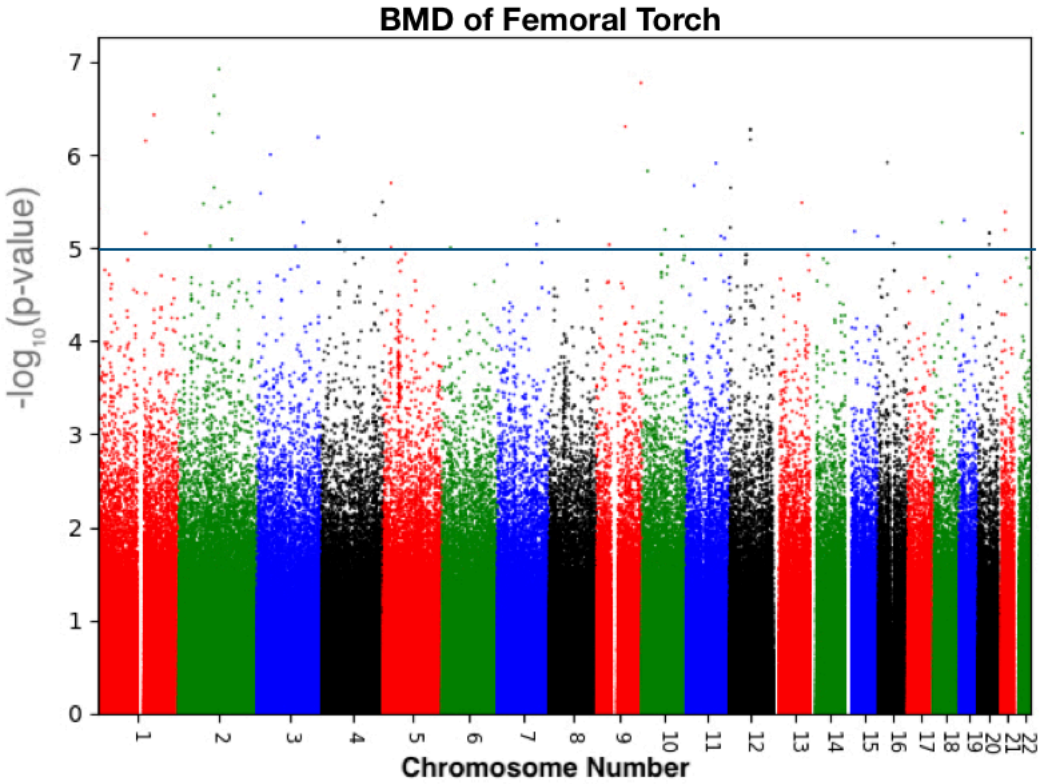

Figure 24 Manhatten plot representing genome-wide association results for BMD of Troch of Femur for 3000 participents. No significant SNPs were detected.

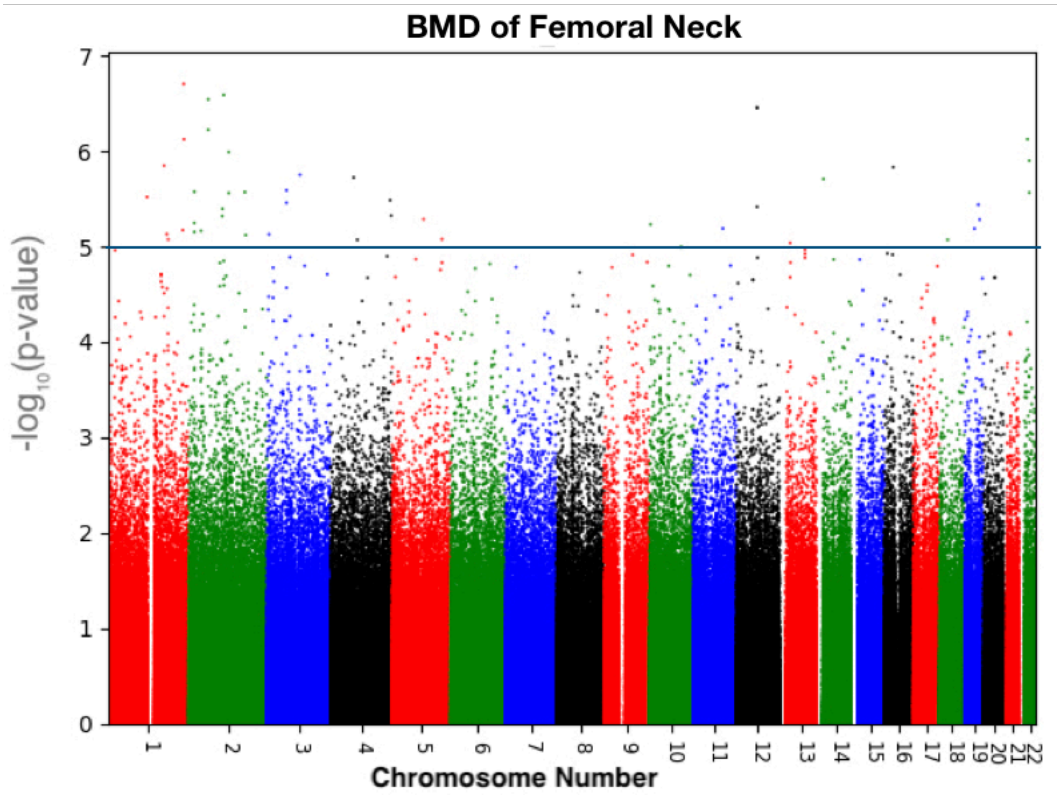

Figure 25 Manhatten plot representing genome-wide association results for BMD of Upper neck of the Femur for 3000 participents. No significant SNPs were detected

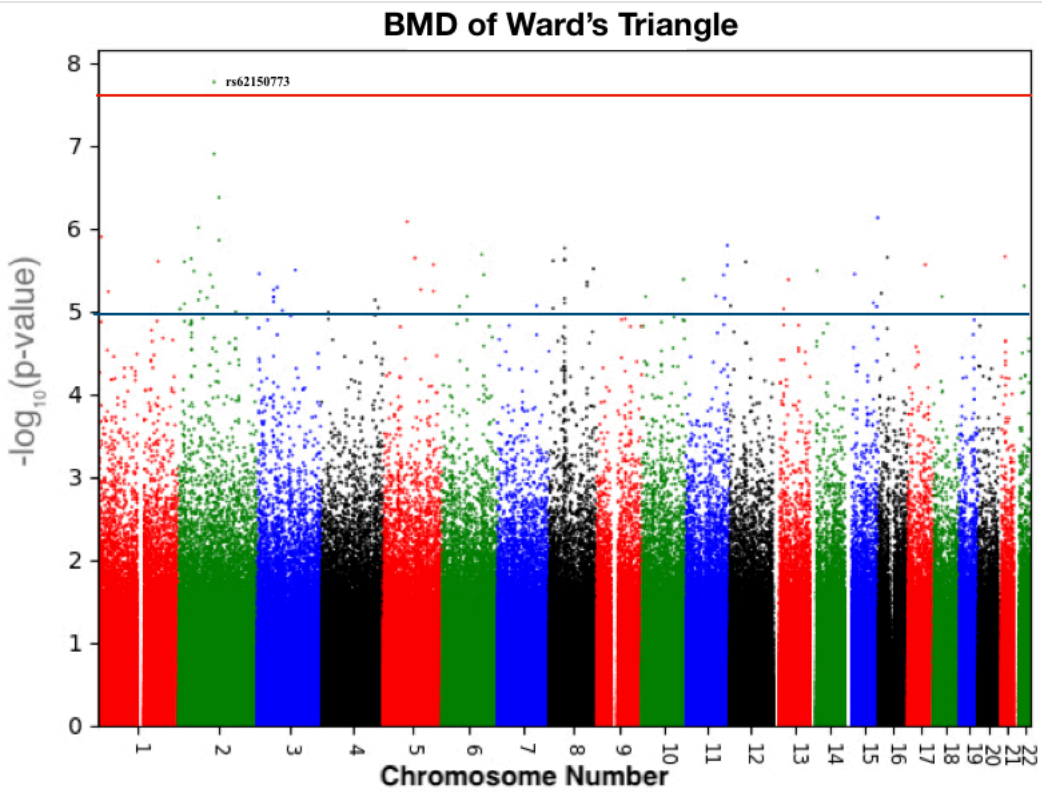

Figure 26 Manhatten plot representing the Genome-wide association results for BMD of Ward of the Femur for 3000 participents. The red line shows the genome wide significance threshold ($p<5x10^{-8}$). 1 SNPs was identified from Wards' triangle BMD to be significantly associated with BMD.

Chapter 5: Discussion

Osteoporosis accounts for more than 8.9 million fracture each year. According to the International Osteoporosis Foundation (IOF), more than 200 million women suffered from osteoporosis globally and 1 in 3 women and 1 in 5 men over the age of 50 experienced an osteoporotic bone fracture.

5.1 Phenotypic variables associated with BMD trait.

Vitamin D plays an essential role in different physiological functions in the human body in addition to its major role in the bone homeostasis process. In this context, several epidemiological studies have revealed the contribution of vitamin D deficiency in wide range of chronic diseases including cardiovascular diseases, metabolic diseases (ex: diabetes mellitus), autoimmune diseases (ex: multiple sclerosis and rheumatoid arthritis) and neoplastic diseases (ex: colon and breast cancer) (Khazai, Judd Se, & Tangpricha, 2008; Kulie et al., 2009; Pittas, Lau J., Hu, & Dawson-Hughes, 2007; Soontrapa & Chailurkit, 2009).

Vitamin D deficiency has become a global public health problem. Several studies have attracted global attention to the increasing prevalence of vitamin D deficiency around the world. However, most of these studies have focused on the elder population, with only a few studies reporting the prevalence of vitamin D deficiency in the younger population. Recent studies have shown that more than 90% of the non-white population and around 75% of the white population groups in the US suffer from vitamin D insufficiency (25- hydroxyvitamin D < 30 ng/ml) (Adams & Hewison, 2010). On the other hand, the prevalence of vitamin D deficiency varies across various

countries. For instance, vitamin D deficiency is highly prevalent in the UK representing 87% and 60% of the participants in the winter and the summer respectively (25(OH) D < 75 nmol/L) (Gonzalez-Molero et al., 2011). Other countries in Europe showed a moderate risk of vitamin D deficiency, where 25(OH) D< 20 ng/ml) was found in 51%, 50% and 40% in Ireland, Germany and Spain respectively (Gonzalez-Molero et al., 2011; O'Sullivan, Suibhne, Cox, Healy, & O'Morain, 2008). Despite the abundant sunshine in the Middle East and Asia compared to Europe and US, countries in these areas have reported the highest rate of hypovitaminosis D worldwide. For instance, in Thailand, the prevalence of vitamin D deficiency was found to be in 77% of premenopausal women, and it reached up to 90% in India (Goswami, Sk, & Kochupillai, 2008; Soontrapa & Chailurkit, 2009). Similar data was reported from a cross-sectional study of young Jordanian women, where 97% suffered from vitamin D deficiency <50 nmol/L (Gharaibeh & Stoecker, 2009). In Morocco, 91% of healthy adult females (24-77 years) had vitamin D level 25(OH) D<75 nmol/L (Allali et al., 2009), and 72% of adult Lebanese had serum level of vitamin D <12 ng/ml (Gannage-Yared, Chemali R., Yaacoub, & Halaby, 2000). On the other hand, a study done in Tunisia reported a lower prevalence of vitamin D deficiency (47% of the participants) (Meddeb et al., 2005). Shockingly, a study done in 2010, showed that 97% of all health care professionals in Qatar were vitamin D deficient (mean level of 25(OH) D<75 nmol/L) (Mahdy et al., 2010). In fact, Qatar is thought to have one of the highest vitamin D insufficiency and vitamin D deficiency rates (Badawi, Arora, Sadoun, Al-Thani, & Thani, 2012). Sunlight is not the only factor that contributes to vitamin D deficiency, inadequate vitamin D fortification in the dietary products (milk, cereal, and drinks such as orange juice), difference in clothing style between different cultures and seasonal

variations where vitamin D levels are highest in September at the end of summer and lowest in winter (Diehl & Chiu, 2010; O'Sullivan et al., 2008).

In our study, we determined the prevalence of vitamin D insufficiency and deficiency among both genders (18-70 years) of the Qatari population. The Data for 3000 samples were collected from Qatari Biobank. Vitamin D insufficiency (vitamin D level < 20 μg/L) was prevalent in 47.2 % and 37.2 of the male and female participants respectively. Vitamin D deficiency (vitamin D level < 12 μg/L) was prevalent in 22.6 and 24.9 of the male and female participants respectively. In total 69.8 and 62.1 of Qatari males and females have lower than normal vitamin D levels respectively.

Similar data were reported from a study conducted on Saudi females were 79.1% had severe vitamin D deficiency (25(OH) D< 25 nmol/L), while 20.9% showed vitamin D insufficiency or mild to moderate deficiency (serum 25(OH)D between 25–50 nmol/L) with a mean level of 31.3 ± 5.6. The reason why vitamin D deficiency is highly prevalent among the population of that region could be due to the fact that these countries enjoy a sunny climate throughout the year, however, the local population exposure to sunlight is limited due to the high temperature at day time, as noticed in our study.

We have observed the influence of socio-demographic characteristics on vitamin D levels of the Qatari population. Interestingly, younger women and men (<30 years) have more vitamin D deficiency than elder women and men ($P \leq 0.00$). However, other studies reported that vitamin D level decreases with aging (Al-Turki, Sadat-Ali M., Al-Elq, A., & Al-Ali, 2008; Gonzalez-Molero et al., 2011). In our study, we

observed that the educational levels and employment status had no significant impact on vitamin D level, as hypovitaminosis D was highly prevalent among young adults of females and males who are supposedly highly educated and working compared to the elder group. Our findings could be explained by the fact that most highly educated young adults are employed and due to the hot weather of Qatar, most of the Qatari population are employed indoors with less sun exposure, on the other hand, elder people have more free time for sun exposure. In addition, the diet for those in the workplace comprises mainly fast food, which lacks many important vitamins and minerals. Our findings were very similar to a study conducted in Saudi Arabia females (Al-Mogbel, 2012). They have reported that Younger women (<29 years) have more vitamin D deficiency than older women ($P \leq 0.00$). In addition, they have found out that the educational levels and employment status had a large impact on vitamin D level in our study, where hypovitaminosis D was more common among highly educated and working females (P value= 0,014 and 0,000, respectively) (Al-Turki et al., 2008). In addition to the socio-demographic characteristics, the low frequency of taking vitamin D supplements could be a very important contributing factor to the high prevalence of vitamin D deficiency in the younger population. In our study, we found out that around 50% of the Qatari females aged from 51-70 are on vitamin D supplements. However, only 23 % of women at age 18-30 and 34% of women at age 31 to 40 are taking vitamin D supplements. In addition, the difference in the frequency of vitamin D supplements between both genders could explain the reason why vitamin D deficiency is more prevalent among males compared to females. This result might be explained by the fact that usually female are more concern with taking supplements and vitamin for the health of their hair, nails, and skins more than the male gender.

Calcium and vitamin D influence the overall mineralization of the skeleton, bone turnover rate and most importantly, the occurrence of bone fractures. In addition, both calcium and vitamin D play an essential role in the development of peak bone mass and the prevention of age-related bone loss. In the cases of vitamin D deficiency, 1,25(OH)2D will decrease leading to a decrease in calcium absorption. Thus, decreasing bone mineralization. Despite the importance of calcium and vitamin D in the protection of bones, they should not be used solely for the prevention and treatment of osteoporosis.

Elders tend to be at higher risk of vitamin D and calcium deficiency due to decreased dietary intake, usually as a result of decreased overall dietary energy intake and infrequent exposure to sunlight. Even though young adults are at high risk of vitamin D and calcium deficiency as well, very few studies have investigated the determinants of circulating 25-hydroxyvitamin in young adults. A recent cross-sectional study in Europe has shown that almost one-fifth of the youth included in this study suffered from vitamin D deficiency, and more than half had vitamin D insufficiency or worse (Tønnesen, Hovind, Jensen, & Schwarz, 2016).

It is generally known that the uptake of calcium and vitamin D might have a protective influence on bone mineral density (Daly, Brown, Bass, Kukuljan, & Nowson, 2006). Despite all of the studies conducted, the effects of vitamin D_3 as a treatment for osteoporosis in adulthood are controversial. Conflicting reports suggest vitamin D_3 supplementation in adulthood reduces (Bischoff-Ferrari et al., 2005; Tang, Eslick, Nowson, Smith, & Bensoussan, 2007), has no effect (Michaelsson, Melhus, Bellocco, & Wolk, 2003; Smith et al., 2007), or increases the incidence of osteoporotic fractures (Sanders et al., 2010). In our study, a positive correlation between vitamin D

and BMD was observed. However, the correlation was not significant $P<0.05$. To further investigate this correlation, we tested the correlation of vitamin D supplements and BMD t score of different parts of the femur and whole-body BMD among the Qatari population. A positive correlation was observed; however, this correlation was not very significant as well.

In consistent with our finding, a recent study (Trajanoska et al., 2018) analyzed the genetic data of more than 500,000 people in the largest-ever study looking at the genetics of osteoporosis and bone fracture risk. They found that genetic predisposition to low levels of vitamin D and calcium intake – previously thought to be important in determining someone's risk of fracturing their bones – does not affect someone's chances of developing osteoporosis. In addition, a systematic review combining 23 separate studies found that vitamin D supplementation in the people suffering from vitamin D deficiency was not effective in reducing fracture risk (Reid, Bolland, & Grey, 2014). The negative findings have shown that perception that vitamin D works directly on bone cells to promote mineralization is probably incorrect. Thus, the continued widespread use of vitamin D supplements as a treatment or protective measure for osteoporosis in community-dwelling adults without specific risk factors for vitamin D deficiency seems to be inappropriate (Reid et al., 2014).

5.2 Genomic variants associated with BMD trait.

Moving on to the genomic data, a genomic association study was conducted on genomic data of 3000 participants obtained from Qatar Biobank. A total of 19 SNPs were identified in our study with genome-wide significant p-values of $<5x10^{-08}$. 6 SNPs of the 19 SNPs identified were replicated by the UK biobank GWAS and GEFOS.

Inconsistent with previous studies conducted on BMD, our study provides evidence that the genetic basis of BMD phenotype is polygenic, thus, many common genomic variants contribute to the final small effects (K. Estrada et al., 2012; Stuart K. Kim, 2018; Moayyeri et al., 2014; Mullin et al., 2017). Sequencing results have shown that BMD phenotype could be affected by rare genetic variants as well (Zheng et al., 2015). Zheng et al., identified a novel non-coding genetic variant rs148771817 (Pmeta = 1 × 10-11) near *WNT16* gene that showed a large effect on BMD phenotype

5.2.1 Replicated SNPs known to be associated with BMD

Two of the genomic variants, rs4727924 and rs2536172, identified in our study are located in the locus 7q31.31 are associated with whole body, Spine and Femoral Troch BMD. These regions contain three genes *WNT16*, *FAM3C* and *C7orf58*. Interestingly, both SNPs were observed together on BMD of different bones. This could be due to the high LD in this region. *WNT16* gene is the best candidate to affect BMD at this locus based on our current knowledge considering the fact that it belongs to the protein family Wnt. Expectedly, the Wnt signaling pathway is known to play an essential role in bone physiology and specifically for bone formation and remodeling (Krishnan, Bryant, & Macdougald, 2006; Milat & Ng, 2009). Moreover, functional studies conducted on the *WNT16* gene knock out in mice showed a reduction in total BMD at age of 24 weeks. The reduction in total BMD was due to a decrease in bone mineral content and the area of the bone in the mouse (Medina-Gomez et al., 2012). Not only the *WNT16* gene is associated with bones in this region, but also the *FAM3C* gene is widely expressed in different cells including the osteoblasts (Zhu et al., 2002). Up to date, only small information is known about *C7orf58*, the third gene in that area.

However, due to the hypothesis-free nature of GWA study, we cannot exclude the possibility that any of these genes may code for proteins that are essential in BMD phenotype. rs4727924 and rs2536172 were reported previously to be associated with heal BMD trait by (Medina-Gomez et al., 2012), Generation R discovery cohort, UK Biobank and GEFOS at very high genome-wide significance; 4.13e-160 and 2.09e-180, respectively as shown in figure 27 and 28.

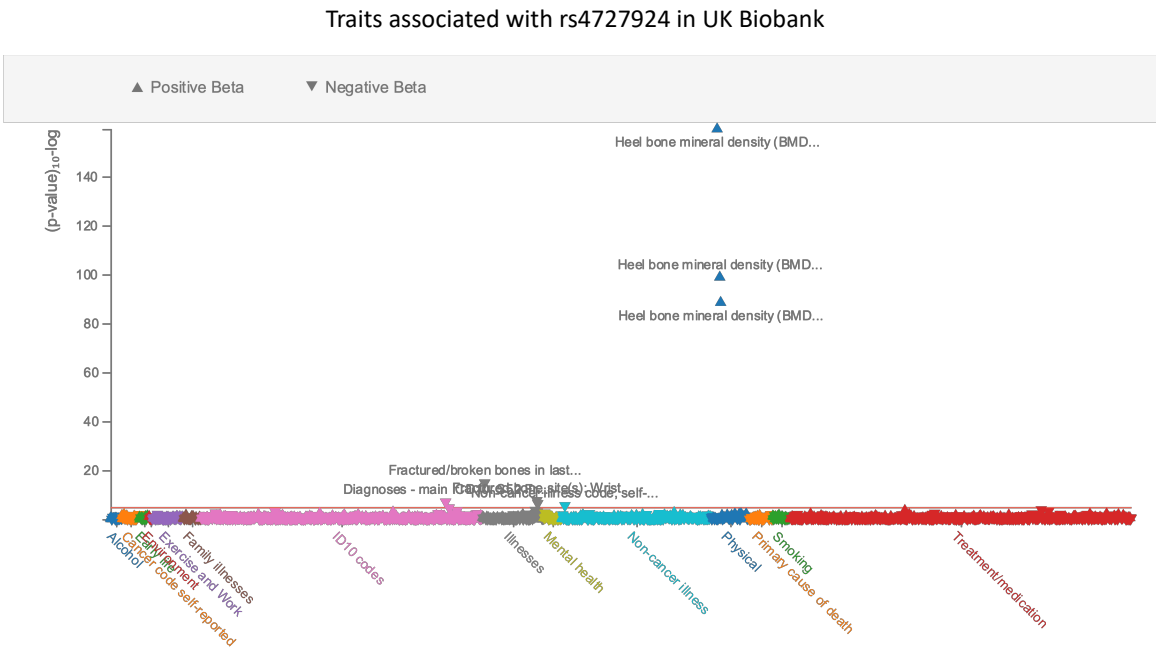

Figure 27 Representative figure from Open Target Genetics representing the traits reported to be associated with rs4727924 by the UK biobank cohort.

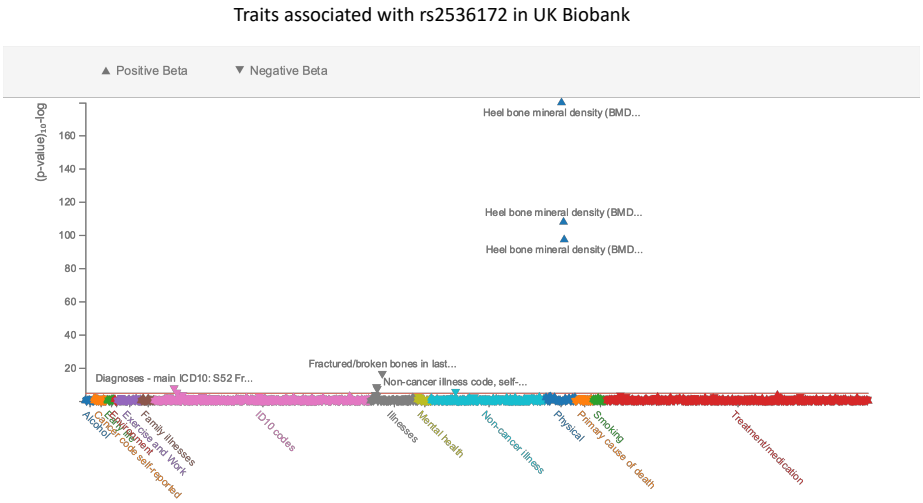

Figure 28 Representative figure from Open Target Genetics representing the traits reported to be associated with rs2536172 by the UK biobank cohort.

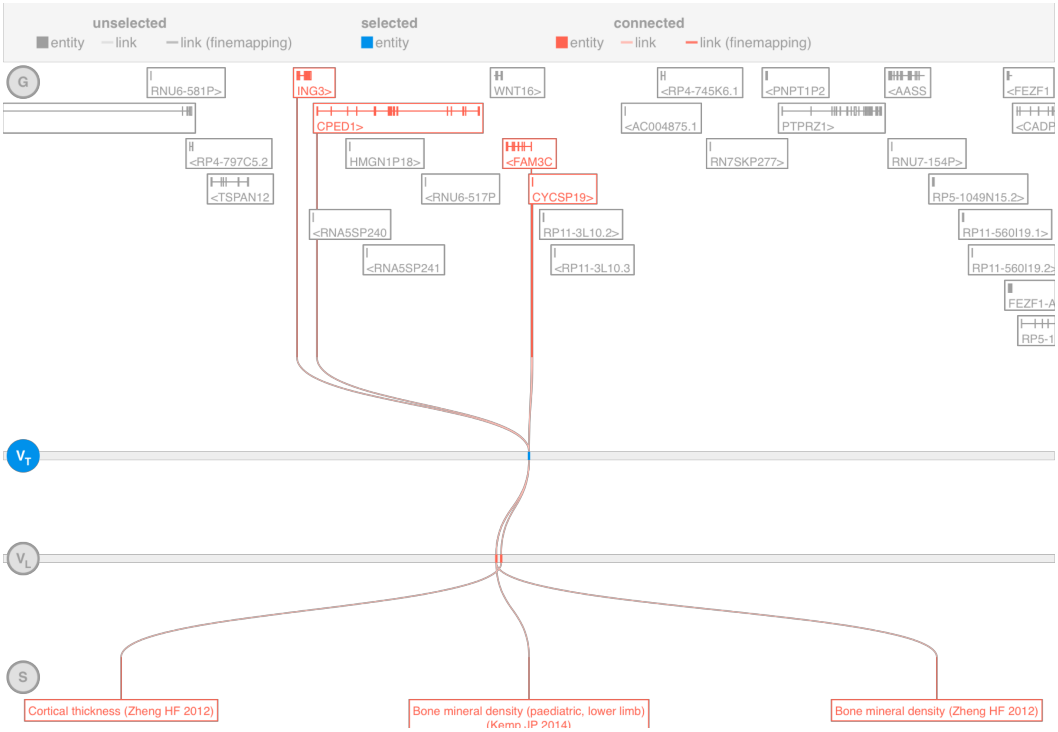

Figure 29 Representative figure of genetic evidence associated with7q31.31 locus. The expansion is shown by LD and fine mapping. The figure is created by Open Target genetics.

The third SNP that we replicated in our study is rs1839588. It was previously reported by UK biobank to be associated with Pelvis BMD as shown in figure 30. It is an intronic variant on locus 7p14.1 overlapping the *SFRP4* gene. It is believed that the *SFRP4* gene act a soluble mediator that bind to WNT ligand and alter the WNT signaling pathway discussed earlier. Studies have shown that SFRP4 protein decreases the bone formation process and inhibits the proliferation of osteoblasts by antagonizing the Wnt signaling. In addition, it was recently shown that sFRP4-dependent Wnt signal plays a critical role in bone remodeling process and age-related bone loss that leads to bone fractures (Haraguchi et al., 2016).

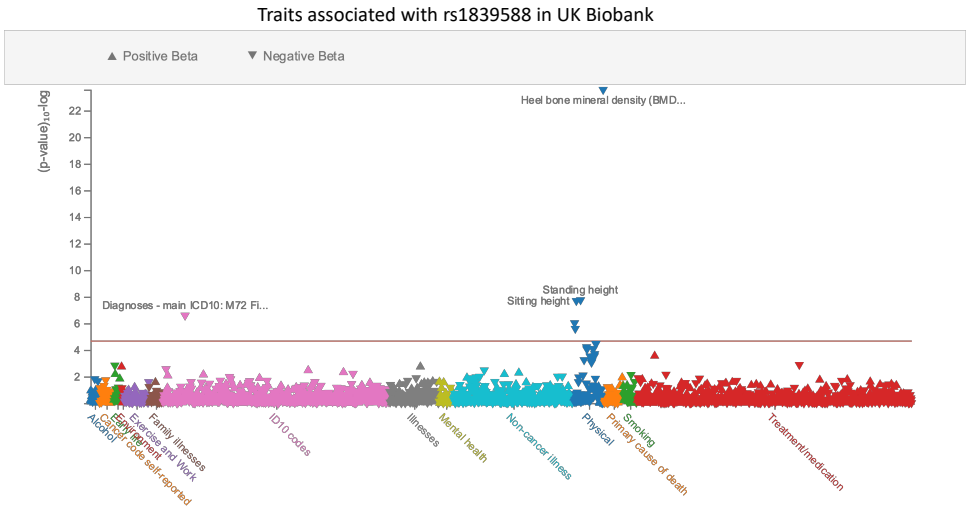

Figure 30 Representative figure from Open Target Genetics representing the traits reported to be associated with rs1839588 by the UK biobank cohort.

Two variants replicated in our study; rs190738498 and rs191429075 were harbored in 18 q21.33 in the *PIGN* gene. Both intronic variants were reported previously by UK Biobank. However, those variants were reported with different phenotypic traits including physical traits but at low significance, as shown in figure 31 & 29. According to open target genetics, rs191429075 variant was found to be linked to *RNF152* gene as well. In our study, these two variants were found to be associated with whole body BMD at a genome-wide significance (10^-8). The *PIGN* gene is one of more than 20 genes that are involved in GPI anchor biosynthesis pathway (Ohba et al., 2014). A recent study has reported the *PIGN* mutation in an Israeli Arab family is

responsible for multiple congenital anomalies including cardiac and skeletal defects (Khayat et al., 2016). However, the exact mechanism is still unknown.

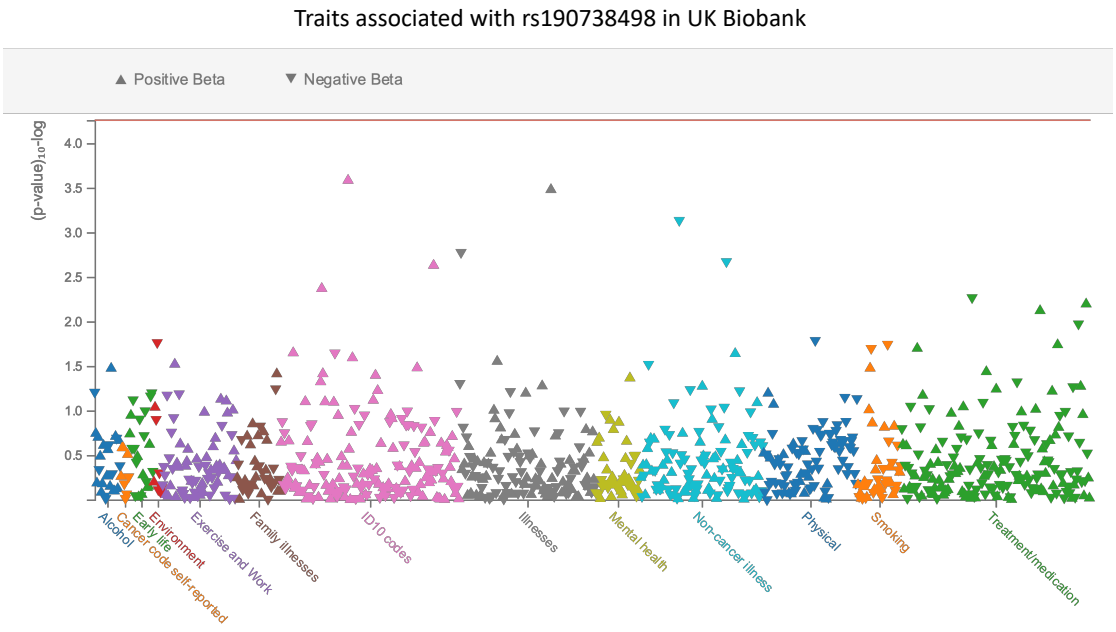

Figure 31 Representative figure from Open Target Genetics representing the traits reported to be associated with rs190738498 by the UK biobank cohort.

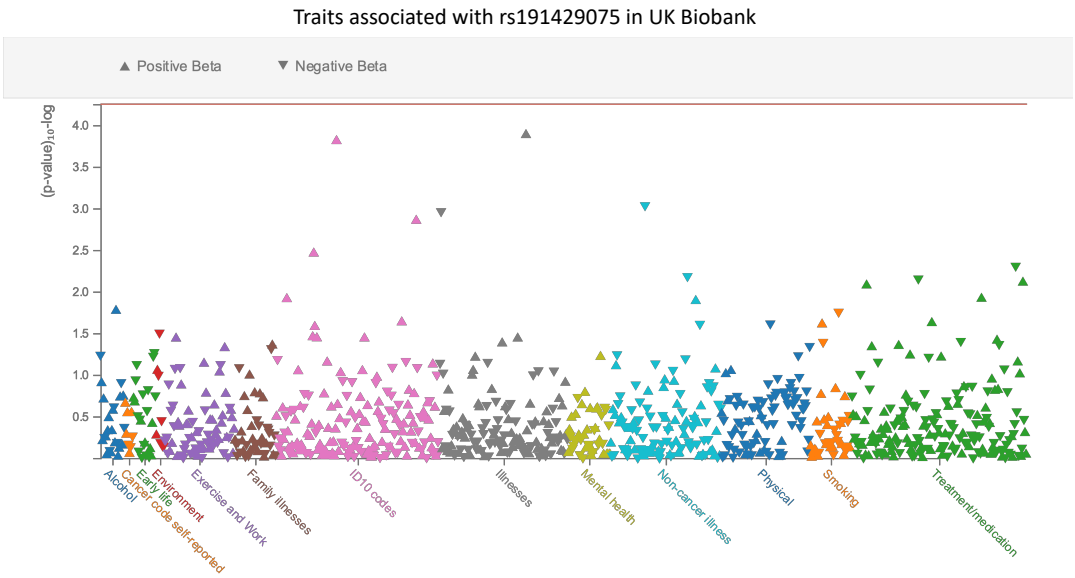

Figure 32 Representative figure from Open Target Genetics representing the traits reported to be associated with rs191429075 by the UK biobank cohort.

The final replicated variant in our study is rs489125 intronic variant. it is found on 22q12.1 locus overlapping *CRYBB2P1* gene. This variant was reported by the UK biobank as well with varies diseases. But most significant with physical health and maternal diseases as shown in figure 33. In our study, it was reported with whole body BMD. *CRYBB2P1* is a pseudogene that belongs to the β-crystallin family. It is transcribed in most tissues except the eye tissue. Till now, there are no proteins associated with that gene (Messina-Baas, Gonzalez-Garay, González-Huerta, Toral-López, & Cuevas-Covarrubias, 2016). In addition, there is no known mechanism regarding how it affects physical phenotype.

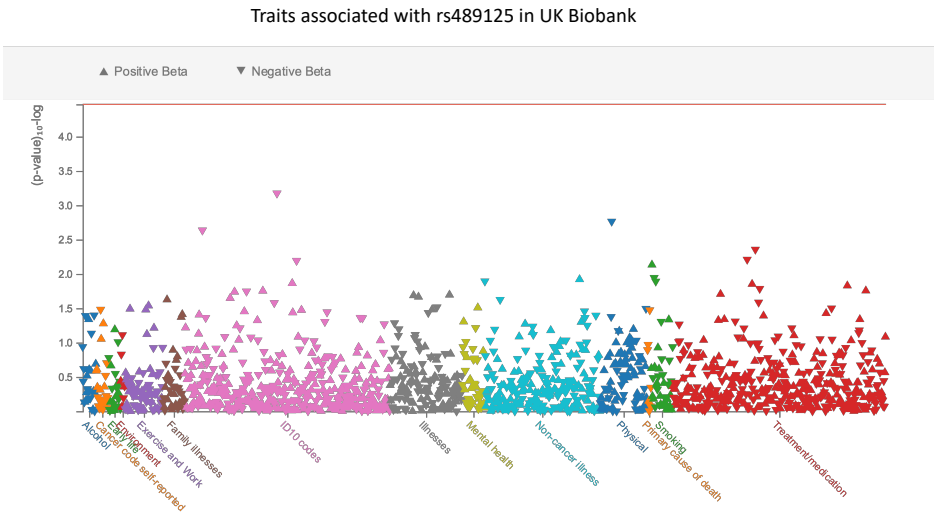

Figure 33 Representative figure from Open Target Genetics representing the traits reported to be associated with rs489125 by the UK biobank cohort.

5.2.2 Novel SNPs identified in our study associated with BMD phenotype.

We identified 13 SNPs associated with whole body, trunk, femoral ward BMD with genome-wide significant p-values of <6.12E-08 in the cohort of 3000 Qatari participants. 7 of these variants were intronic variants harbored in 6 gene loci; *MALAT1, SAG FASLG, FAM189A2, RP11-15A1.7, LSAMP*, and *BMPR1B*. One variant is a splice variant and the remaining 5 variants were intergenic variants.

The rs202070768 intronic variant is found on the 11q13.1 locus. It is found to be overlapping the *MALAT1* gene. A recent study conducted in 2018 demonstrated that *MALAT1* gene play an essential role in the onset of osteolysis (Yang, Zhang, Li, & Wen, 2018). In addition, X yang et al., concluded that *MALAT1* is a potential

therapeutic target for treating osteolysis especially in patients suffering from osteolysis associated with knee replacement (Yang et al., 2018).

The rs374876997/rs553335180 is a splice variant at 21q21.2 locus. The closest gene to this variant is the *MRPL39* gene. Variants in this gene have been associated with BMD trait since 2005. Studies have shown differences in the expression of that gene between individuals with low and high BMD (Y. Z. Liu et al., 2005).

The rs1050627711 is an intronic variant at the q37.1 locus overlapping the *SAG* gene. This variant was detected to be associated with whole body BMD at a genome wide significance of $6.08x10^{-08.}$ The visual/*β*-arrestins, a small family of proteins, encoded by the *SAG* gene, originally described for their role in the desensitization and intracellular trafficking of G protein–coupled receptors (GPCRs), have emerged as key regulators of multiple signaling pathways. Recent studies have shown that *β*-arrestins could play an essential role in bone metabolism and remodeling (Peterson & Luttrell, 2017).

The rs867865671 is an intronic variant at 1q24.3 that was detected in our study at a genome-wide significance of 3.03E-08 in whole body BMD. No previous reports are available regarding this intronic variant association with BMD. However, this variant is found to be overlapping the *FASLG* gene, which plays an essential role in bone formation and altering bone mineral density. Estrogen deficiency in post-menopausal women is known to be a major cause of osteoporosis and low BMD. In this context, estrogen plays a key role in maintaining the ratio between bone formation (osteoblasts) and bone resorption (osteoclasts) by inducing apoptosis of osteoclasts.

Studies have shown that *FASLG* gene plays an important role in bone formation. Basically, estrogen hormone induces *FASL* in osteoblasts, which in turns results in osteoclasts' apoptosis by autocrine mechanism leading to increasing of BMD (Garcia et al., 2013; Jones, 2015; Krum et al., 2008).

The rs142479295 intronic variant is found on the 3q13.32. This variant was identified in our study with genomic-wide significance of 9.68E-08. This variant is overlapping the *LSAMP* gene. A recent study has shown that a frequent deletion in this locus is associated with osteosarcomas, which is the most frequent primary malignant bone tumors (Baroy et al., 2014).

The rs1050715238 intronic variant is found on 4q22.3 locus. It was identified in our study with a genome-wide significance of 4.77E-08 in BMD of the trunk. It is found to be overlapping the *bone morphogenetic protein receptor type 1B* gene *(BMPR1B).* These proteins are essential for inducing the formation of bones and cartilage. This gene is essential for encoding bone morphogenetic protein receptors. The disruption of *BMPR1B* gene led to osteopenia in mice model (Shi et al., 2016). No previous reports on the traits associated with that variant.

The rs73455199 intronic variant is found on the 9q21.12 locus. This variant was identified in our study with a genomic wide significance of 3.38E-08. It is found to be overlapping the *FAM189A2* gene. The rs149339318 intronic variant is found on 19q13.3 locus. This variant was identified in our study with genome-wide significance of 8.35E-08. It is overlapping the *RP11-15A1.7* gene. Until now, there is no information regarding the function of those protein in literature.

5 SNPs identified in our study are intergenic variant; rs554808159 in 17 q23.3, rs367949909 in 6 q23.2, rs866548296 in 1 q42.2, rs371319602 in 11 q21 and rs62150773 in 2 q12.3 locus. In the past 50 years, studies have shown that a large proportion of the DNA is transcribed but not translated into proteins. A recent study has shown that 75% of the human DNA can be transcribed, however, only 1 to 2 % get translated into proteins (Bergmann & Spector, 2014). Thousands of diseases associated SNPs are found in intergenic regions, which make it difficult for researchers to understand their association with diseases. Even though evidence has shown that the non-coding SNPs are frequently located near regulatory elements (Maurano et al., 2012), so far, most of the studies focus on disease-associated SNPs found in the coding regions only. Yet, 93% of the SNPs identified so far through GWA studies are located in non-coding regions including the intergenic region (Maurano et al., 2012). Thus, presenting major challenges to researchers to interpret their association with a particular trait.

Chapter 5: Conclusion

Osteopetrosis is a devastating disease that is characterized by compromised bone strength leading to increased risk of fracture. It is defined as BMD that lies 2.5 SD or more below the average value for young healthy women, as measured with DXA scan. BMD is the most important predictor of fracture risk. According to the executive summary of the IOF audit report, osteoporosis is a neglected disease, not being integrated in to medical curricula of most countries, and the level of awareness about osteoporosis is estimated as poor to medium in Arab countries (IOF, 2011).

It has been known for a long time that vitamin D supplements could improve BMD status and help in decreasing the fracture risk. A recent study (Trajanoska et al., 2018) analyzed the genetic data of more than 500,000 people in the largest-ever study looking at the genetics of osteoporosis and bone fracture risk. They found that genetic predisposition to low levels of vitamin D and calcium intake – previously thought to be important in determining someone's risk of fracturing their bones – does not affect someone's chances of developing osteoporosis. Our study supported their results as we didn't observe any significant correlation vitamin D and BMD. The negative findings have shown that perception that vitamin D works directly on bone cells to promote mineralization is probably incorrect. Thus, the continued widespread use of vitamin D supplements as a treatment or protective measure for osteoporosis in community-dwelling adults without specific risk factors for vitamin D deficiency seems to be inappropriate (Reid et al., 2014).

Moving on to the genomic data, in the past few years, the bone field has witnessed great advances in genome-wide association studies (GWASs) of osteoporosis, with a number of promising genes identified. Many of the identified proteins in GWAS of osteoporosis have clear and relevant mechanisms of action for osteoporosis pathophysiology. In Qatar, we performed the first GWA study of Osteoporosis in the Arab countries to uncover the genetic risk factors associated with BMD and osteoporosis in the Qatari population. In addition, we conducted the first GWA study to include 7 BMD measurements from different parts; whole body, spine, pelvis, trunk, femoral parts in which the best and most accurate measurement of BMD is at the whole body BMD, Spine and pelvis. We included 3000 healthy Qatari participants aging from 18-70 years old in our study. DXA scan was conducted on all of the 3000 participants and BMD was measured for different parts of the body; whole body, spine, pelvis, trunk, femoral upper neck, femoral troch and femoral ward. Our study is the first study to include this wide range of different BMD measurements in a GWA study of osteoporosis.

Our study identified 19 common variants associated with BMD at a genome-wide significance $P<5\times10^{-8}$. 6 SNPs were previously reported by GWAS of UK BioBank cohort and GEFOS cohort. 2 of these SNPs are located in 7q31.31 locus overlapping *WNT16 and FAM3C* genes. WNT signaling pathway plays a critical role in bone formation and remolding. The third SNP is found on chromosome 7 as well in the p14.1 locus, overlapping the *SFRP4* gene. SFRP protein act as a soluble mediator in the WNT signaling, thus, critical for bone biological functions. The last three SNPs that were replicated in our study; 2 SNPs were located in chromosome 18 in locus

q21.33 overlapping the PIGN gene and one SNP in locus 21q12.1, overlapping the CRYBB2P1. However, those three SNPs were reported by the UK biobank to be associated with several traits including the physical phenotype.

We identified 13 novel SNPs to be associated with BMD of Whole body, Spine, Pelvis, Trunk, and Femur (Torch, Ward and Neck) at a genome-wide significance $P<5\times10^{-8}$ that weren't reported previously. 8 of these variants were intronic variants harbored in 8 gene loci; *MALAT1, SAG, MRPL39, FASLG, FAM189A2, RP11-15A1.7, LSAMP*, and *BMPR1B*. One variant is a splice variant and the remaining 4 variants were intergenic variants.

One limitation of our study is that we included participants from age 18- 70 years old. Estrogen has a large effect on BMD and fracture in females. Having female participants after menopause should be excluded to avoid any confounding factors in our study. Our results should be validated in a cohort with younger age groups suffering from low BMD.

In conclusion, our findings highlight the highly polygenic and complex nature underlying BMD variation, shedding light on the pathophysiological mechanisms underlying fracture susceptibility and harboring potential for the future identification of drug targets for the treatment of osteoporosis. In addition, the risk alleles we have identified in our GWA study justify the need for further clinical and biological investigations. Proteins identified and prioritized by our study have identified signaling pathways that represent new drug targets for the prevention and treatment of osteoporosis- a major health care priority. These SNPs alone are unlikely to change

current clinical practice, but as has been shown for other diseases (Maller et al., 2006) extended panels of several SNP markers could be used in the future, in addition to traditional risk factors, to better identify populations who are at high risk for osteoporotic fractures.

The findings of new genetic variants that are not previously described in GWAS studies to be related to BMD variants is of major importance and open horizons for new studies to investigate the molecular mechanisms of a potential relation between MALT1, FASL, and MRPL39 and bone remodeling and development. For future studies, we will conduct a replication study using an independent cohort and will do fine mapping of the association signals in the identified LD regions. In addition, we will conduct in in Silico Analysis of the functional and structural consequences of the novel identified SNPs to find out its molecular mechanism in the bone remodeling and osteogenesis process.

References

Abecasis, G. R., Altshuler, D., Auton, A., Brooks, L. D., Durbin, R. M., Gibbs, R. A., . . . McVean, G. A. (2010). A map of human genome variation from population-scale sequencing. *Nature, 467*(7319), 1061-1073. doi:10.1038/nature09534

Abu, E. O., Bord S., Horner, A., Chatterjee, V. K., & Compston, J. E. (1997). The expression of thyroid hormone receptors in human bone. (8756-3282 (Print)).

Adachi, J. D., Ioannidis G., Olszynski Wp., Brown Jp., Hanley Da., Sebaldt Rj., . . . Goldsmith, C. H. (2002). The impact of incident vertebral and non-vertebral fractures on health related quality of life in postmenopausal women. (1471-2474 (Print)).

Adams, J. S., & Hewison, M. (2010). Update in vitamin D. (1945-7197 (Electronic)).

Agerbaek, M. O., Eriksen Ef., Kragstrup, J., Mosekilde, L., & Melsen, F. (1991). A reconstruction of the remodelling cycle in normal human cortical iliac bone. (0169-6009 (Print)).

Al-Mogbel, E. S. (2012). Vitamin D status among Adult Saudi Females visiting Primary Health Care Clinics. *International journal of health sciences, 6*(2), 116-126. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/23580892.

Al-Turki, H. A., Sadat-Ali M., Al-Elq, A. H., A., F., & Al-Ali, A. K. (2008). 25-Hydoxyvitamin D levels among healthy Saudi Arabian women. (0379-5284 (Print)).

Allali, F., El Aichaoui S., Khazani, H., Benyahia, B., Saoud, B., Saoud B., . . . Hajjaj-Hassouni, N. (2009). High prevalence of hypovitaminosis D in Morocco:

relationship to lifestyle, physical performance, bone markers, and bone mineral density. (1532-866X (Electronic)).

Arden, N. K., Baker J., Hogg, C., Baan, K., & Spector, T. D. (1996). The heritability of bone mineral density, ultrasound of the calcaneus and hip axis length: a study of postmenopausal twins. (0884-0431 (Print)).

Aulchenko, Y. S., de Koning, D. J., & Haley, C. (2007). Genomewide rapid association using mixed model and regression: a fast and simple method for genomewide pedigree-based quantitative trait loci association analysis. *Genetics, 177*(1), 577-585. doi:10.1534/genetics.107.075614

Autier, P., Haentjens P., Bentin J., Baillon Jm., Grivegnee Ar., Closon Mc., & Boonen, S. (2000). Costs induced by hip fractures: a prospective controlled study in Belgium. Belgian Hip Fracture Study Group. (0937-941X (Print)).

Badawi, A., Arora, P., Sadoun, E., Al-Thani, A.-A., & Thani, M. H. A. (2012). Prevalence of vitamin d insufficiency in qatar: a systematic review. *Journal of public health research, 1*(3), 229-235. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/25170469. doi:10.4081/jphr.2012.e36

Barnard, R., Ng Kw., Martin, T. J., & Waters, M. J. (1991). Growth hormone (GH) receptors in clonal osteoblast-like cells mediate a mitogenic response to GH. (0013-7227 (Print)).

Baroy, T., Kresse, S. H., Skarn, M., Stabell, M., Castro, R., Lauvrak, S., . . . Meza-Zepeda, L. A. (2014). Reexpression of LSAMP inhibits tumor growth in a preclinical osteosarcoma model. *Mol Cancer, 13*, 93. doi:10.1186/1476-4598-13-93

Barton, N. H. (2010). Genetic linkage and natural selection. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences, 365*(1552), 2559-2569. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/20643746. doi:10.1098/rstb.2010.0106

Bener, A., Hammoudeh, M., & Zirie, M. (2007). Prevalence and predictors of osteoporosis and the impact of life style factors on bone mineral density. *APLAR Journal of Rheumatology, 10*(3), 227-233. doi:10.1111/j.1479-8077.2007.00294.x

Bergmann, J. H., & Spector, D. L. (2014). Long non-coding RNAs: modulators of nuclear structure and function. *Curr Opin Cell Biol, 26*, 10-18. doi:10.1016/j.ceb.2013.08.005

Bianco, P., Riminucci M., Gronthos, S., & Robey, P. G. (2001). Bone marrow stromal stem cells: nature, biology, and potential applications. (1066-5099 (Print)).

Biobank, Q. (2017). QATAR BIOBANK REPORT 2016-2017.

Bischoff-Ferrari, H. A., Willett, W. C., Wong, J. B., Giovannucci, E., Dietrich, T., & Dawson-Hughes, B. (2005). Fracture prevention with vitamin D supplementation: a meta-analysis of randomized controlled trials. *Jama, 293*(18), 2257-2264. doi:10.1001/jama.293.18.2257

Bonnick, S. L. (2010). Bone Densitometry in Clinical Practice.

Burge, R., Dawson-Hughes., Solomon, D. H., Wong, J. B., King, A., & Tosteson, A. (2005). Incidence and economic burden of osteoporosis-related fractures in the United States, 2005-2025. (0884-0431 (Print)).

Bush, W. S., & Moore, J. H. (2012). Chapter 11: Genome-wide association studies. *PLoS Comput Biol, 8*(12), e1002822. doi:10.1371/journal.pcbi.1002822

Canalis, E., J, P., Gabbitas, B., Rydziel, S., & Varghese, S. (1993). Growth factors regulate the synthesis of insulin-like growth factor-I in bone cell cultures. (0013-7227 (Print)).

Canalis, E., J, P., & Varghese, S. (1993). Skeletal growth factors. (1045-4403 (Print)).

Cardon, L. R., & Palmer, L. J. (2003). Population stratification and spurious allelic association. *Lancet, 361*(9357), 598-604. doi:10.1016/s0140-6736(03)12520-2

Carter Dr Fau - Hayes, W. C., & Hayes, W. C. (1977). Compact bone fatigue damage: a microscopic examination. (0009-921X (Print)). doi:D - NASA: 78023067 EDAT- 1977/01/01 00:00 MHDA- 1977/01/01 00:01 CRDT- 1977/01/01 00:00 PHST- 1977/01/01 00:00 [pubmed] PHST- 1977/01/01 00:01 [medline] PHST- 1977/01/01 00:00 [entrez] PST - ppublish

Carter Dr., & Hayes, W. C. (1977). Compact bone fatigue damage: a microscopic examination. (0009-921X (Print)). doi:D - NASA: 78023067 EDAT- 1977/01/01 00:00 MHDA- 1977/01/01 00:01 CRDT- 1977/01/01 00:00 PHST- 1977/01/01 00:00 [pubmed] PHST- 1977/01/01 00:01 [medline] PHST- 1977/01/01 00:00 [entrez] PST - ppublish

Chen, D., Zhao M., -., & Mundy, G. R. (2004). Bone morphogenetic proteins. (0897-7194 (Print)).

Clarke, B. (2008). Normal bone anatomy and physiology. *Clinical journal of the American Society of Nephrology : CJASN, 3 Suppl 3*(Suppl 3), S131-S139. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/18988698. doi:10.2215/CJN.04151206

Clement-Lacroix, P., Lepescheux, L., Ammann, P., Damotte, D., Goffin, V., Bouchard, B., . . . Kelly, P. A. (1999). Osteoblasts are a new target for

prolactin: analysis of bone formation in prolactin receptor knockout mice. (0013-7227 (Print)).

Cohen, A. J., & Roe, F. J. (2000). Review of risk factors for osteoporosis with particular reference to a possible aetiological role of dietary salt. (0278-6915 (Print)).

Colvard, D. S., Keeting, P. E., Wilson, E. M., Lubahn, D. B., French, F. S., Riggs, B. L., & Spelsberg, T. C. (1989). Identification of androgen receptors in normal human osteoblast-like cells. (0027-8424 (Print)).

Cosman, F., de Beur, S. J., LeBoff, M. S., Lewiecki, E. M., Tanner, B., Randall, S., . . . National Osteoporosis, F. (2014). Clinician's Guide to Prevention and Treatment of Osteoporosis. *Osteoporosis international : a journal established as result of cooperation between the European Foundation for Osteoporosis and the National Osteoporosis Foundation of the USA, 25*(10), 2359-2381. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/25182228. doi:10.1007/s00198-014-2794-2

Cree, M., Soskolne Cl., Belseck, E., Hornig, J., McElhaney, J. E., Brant R., & Suarez-Almazor, M. (2000). Mortality and institutionalization following hip fracture. (0002-8614 (Print)).

Cummings, S. R., Nevitt, M. C., Browner, W. S., Stone, K., Fox, K. M., Ensrud, K. E., . . . Vogt, T. M. (1995). Risk factors for hip fracture in white women. Study of Osteoporotic Fractures Research Group. *N Engl J Med, 332*(12), 767-773. doi:10.1056/nejm199503233321202

Daly, R. M., Brown, M., Bass, S., Kukuljan, S., & Nowson, C. (2006). Calcium- and vitamin D3-fortified milk reduces bone loss at clinically relevant skeletal sites

in older men: a 2-year randomized controlled trial. *J Bone Miner Res, 21*(3), 397-405. doi:10.1359/jbmr.051206

Darwish, H. M., & DeLuca, H. F. (1996). Recent advances in the molecular biology of vitamin D action. (0079-6603 (Print)).

Davies, J., Warwick J., Totty, N., Philp, R., Helfrich, M., & Horton, M. (1989). The osteoclast functional antigen, implicated in the regulation of bone resorption, is biochemically related to the vitronectin receptor. (0021-9525 (Print)).

Delaisse, J. M. (2014). The reversal phase of the bone-remodeling cycle: cellular prerequisites for coupling resorption and formation. (2047-6396 (Print)).

Destaing, O., Sanjay A., Itzstein, C., Horne, W. C., Toomre, D., De Camilli, P., & Baron, R. (2008). The tyrosine kinase activity of c-Src regulates actin dynamics and organization of podosomes in osteoclasts. (1939-4586 (Electronic)).

Devlin, B., & Roeder, K. (1999). Genomic control for association studies. *Biometrics, 55*(4), 997-1004.

Diehl, J. W., & Chiu, M. W. (2010). Effects of ambient sunlight and photoprotection on vitamin D status. (1529-8019 (Electronic)).

Dobbs, M. B., Buckwalter, J., & Saltzman, C. (1999). Osteoporosis: the increasing role of the orthopaedist. *The Iowa orthopaedic journal, 19*, 43-52. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/10847516.

Ducy, P., R, Z., Geoffroy, V., V, G., Ridall, A. L., & Karsenty, G. (1997). Osf2/Cbfa1: a transcriptional activator of osteoblast differentiation. (0092-8674 (Print)).

Edwards, A. O., Ritter R 3rd., Abel, K. J., Manning, A., Panhuysen, C., & Farrer, L. A. (2005). Complement factor H polymorphism and age-related macular degeneration. (1095-9203 (Electronic)).

Elmardi, A. S., Katchburian Mv., & Katchburian, E. (1990). Electron microscopy of developing calvaria reveals images that suggest that osteoclasts engulf and destroy osteocytes during bone resorption. (0171-967X (Print)).

Eriksen, E. F., Colvard, D., Berg, N. J., Graham, M. L., Mann, K. G., Spelsberg, T. C., & Riggs, B. L. (1988). Evidence of estrogen receptors in normal human osteoblast-like cells. (0036-8075 (Print)).

Estrada, K., Styrkarsdottir, U., Evangelou, E., Hsu, Y.-H., Duncan, E. L., Ntzani, E. E., . . . Rivadeneira, F. (2012). Genome-wide meta-analysis identifies 56 bone mineral density loci and reveals 14 loci associated with risk of fracture. *Nature genetics, 44*(5), 491-501. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/22504420. doi:10.1038/ng.2249

Estrada, K., Styrkarsdottir, U., Evangelou, E., Hsu, Y. H., Duncan, E. L., Ntzani, E. E., . . . Rivadeneira, F. (2012). Genome-wide meta-analysis identifies 56 bone mineral density loci and reveals 14 loci associated with risk of fracture. *Nat Genet, 44*(5), 491-501. doi:10.1038/ng.2249

Fadista, J., Manning, A. K., Florez, J. C., & Groop, L. (2016). The (in)famous GWAS P-value threshold revisited and updated for low-frequency variants. *European Journal Of Human Genetics, 24*, 1202. Retrieved from https://doi.org/10.1038/ejhg.2015.269. doi:10.1038/ejhg.2015.269

Florencio-Silva, R., Sasso, G. R. d. S., Sasso-Cerri, E., Simões, M. J., & Cerri, P. S. (2015). Biology of Bone Tissue: Structure, Function, and Factors That

Influence Bone Cells. *BioMed research international, 2015*, 421746-421746. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/26247020. doi:10.1155/2015/421746

Frost, H. M. (1990). Skeletal structural adaptations to mechanical usage (SATMU): 1. Redefining Wolff's law: the bone modeling problem. (0003-276X (Print)).

Galante J., Rostoker W., & Ray, R. D. (1970). Physical properties of trabecular bone. (0008-0594 (Print)).

Gannage-Yared, M. H., Chemali R., Yaacoub, N., & Halaby, G. (2000). Hypovitaminosis D in a sunny country: relation to lifestyle and bone markers. (0884-0431 (Print)).

Garcia, A. J., Tom, C., Guemes, M., Polanco, G., Mayorga, M. E., Wend, K., . . . Krum, S. A. (2013). ERalpha signaling regulates MMP3 expression to induce FasL cleavage and osteoclast apoptosis. *J Bone Miner Res, 28*(2), 283-290. doi:10.1002/jbmr.1747

Gharaibeh, M. A., & Stoecker, B. J. (2009). Assessment of serum 25(OH)D concentration in women of childbearing age and their preschool children in Northern Jordan during summer. (1476-5640 (Electronic)).

Globus, R. K., Plouet J., & Gospodarowicz, D. (1989). Cultured bovine bone cells synthesize basic fibroblast growth factor and store it in their extracellular matrix. (0013-7227 (Print)).

Gold, D. T. (2001). The nonskeletal consequences of osteoporotic fractures. Psychologic and social outcomes. (0889-857X (Print)).

Gonzalez-Molero, I., Morcillo S., Valdes, S., Perez-Valero, V., Botas, P., Delgado, E., . . . Soriguer, F. (2011). Vitamin D deficiency in Spain: a population-based cohort study. (1476-5640 (Electronic)).

Goswami, R., Sk, M., & Kochupillai, N. (2008). Prevalence & potential significance of vitamin D deficiency in Asian Indians. (0971-5916 (Print)).

Gullberg, B., Johnell O., & Kanis, J. A. (1997). World-wide projections for hip fracture. (0937-941X (Print)).

Habek, M., Brinar Vv., & Borovecki, F. (2010). Genes associated with multiple sclerosis: 15 and counting. (1744-8352 (Electronic)).

Haines, J. L., Ma., H., Schmidt S., Scott Wk., Olson Lm., P., G., . . . Pericak-Vance, M. A. (2005). Complement factor H variant increases the risk of age-related macular degeneration. (1095-9203 (Electronic)).

Hall, S. E., Criddle Ra., Comito Tl., & Prince, R. L. (1999). A case-control study of quality of life and functional impairment in women with long-standing vertebral osteoporotic fracture. (0937-941X (Print)).

Hammoudeh, M., Al-Khayarin, M., Zirie, M., & Bener, A. (2005). Bone density measured by dual energy X-ray absorptiometry in Qatari women. *Maturitas, 52*(3), 319-327. Retrieved from http://www.sciencedirect.com/science/article/pii/S0378512205001131. doi:https://doi.org/10.1016/j.maturitas.2005.05.011

Haraguchi, R., Kitazawa, R., Mori, K., Tachibana, R., Kiyonari, H., Imai, Y., . . . Kitazawa, S. (2016). sFRP4-dependent Wnt signal modulation is critical for bone remodeling during postnatal development and age-related bone loss.

Scientific Reports, 6, 25198. Retrieved from https://doi.org/10.1038/srep25198. doi:10.1038/srep25198

https://www.nature.com/articles/srep25198#supplementary-information

Harbolic, B. (2017). Vitamin D Deficiency. *MedicineNet*.

Hastbacka, J., A., d. l. C., Kaitila, I., P., S., A., W., & Lander, E. (1992). Linkage disequilibrium mapping in isolated founder populations: diastrophic dysplasia in Finland. (1061-4036 (Print)).

Hellwege, J. N., Keaton, J. M., Giri, A., Gao, X., Velez Edwards, D. R., & Edwards, T. L. (2017). Population Stratification in Genetic Association Studies. *Curr Protoc Hum Genet, 95*, 1.22.21-21.22.23. doi:10.1002/cphg.48

Hernlund, E., A., S., M., I., J., C., C., C., J., S., . . . Kanis, J. A. (2013). Osteoporosis in the European Union: medical management, epidemiology and economic burden. A report prepared in collaboration with the International Osteoporosis Foundation (IOF) and the European Federation of Pharmaceutical Industry Associations (EFPIA). (1862-3514 (Electronic)).

Hochberg, Y., & Benjamini, Y. (1990). More powerful procedures for multiple significance testing. (0277-6715 (Print)).

Hofstetter, W., A., W., Mc., C., R., F., H., F., & Mueller, C. (1992). Detection of transcripts for the receptor for macrophage colony-stimulating factor, c-fms, in murine osteoclasts. (0027-8424 (Print)).

Hou, P., Sato T., Hofstetter, W., & Foged, N. T. (1997). Identification and characterization of the insulin-like growth factor I receptor in mature rabbit osteoclasts. (0884-0431 (Print)).

International HapMap, C. (2005). A haplotype map of the human genome. *Nature, 437*(7063), 1299-1320. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/16255080. doi:10.1038/nature04226

IOF. (2011). The middle east & Africa Regional Audit. Epidemiology, costs & burden of osteoporosis in 2011.

Iolascon, G., Napolano, R., Gioia, M., Moretti, A., Riccio, I., & Gimigliano, F. (2013). The contribution of cortical and trabecular tissues to bone strength: insights from denosumab studies. *Clinical cases in mineral and bone metabolism : the official journal of the Italian Society of Osteoporosis, Mineral Metabolism, and Skeletal Diseases, 10*(1), 47-51. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/23858311. doi:10.11138/ccmbm/2013.10.1.047

Jackson, M., Marks, L., May, G. H. W., & Wilson, J. B. (2018). The genetic basis of disease. *Essays in biochemistry, 62*(5), 643-723. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/30509934. doi:10.1042/EBC20170053

Jensen, E. D., Gopalakrishnan, R., & Westendorf, J. J. (2010). Regulation of gene expression in osteoblasts. *Biofactors, 36*(1), 25-32. doi:10.1002/biof.72

Johnell, O., & Kanis, J. A. (2006). An estimate of the worldwide prevalence and disability associated with osteoporotic fractures. (0937-941X (Print)).

Jones, D. R. (2015). A potential osteoporosis target in the FAS ligand/FAS pathway of osteoblast to osteoclast signaling. *Annals of translational medicine, 3*(14), 189-189. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/26417573. doi:10.3978/j.issn.2305-5839.2015.07.01

Kanis, J. A., O., J., Oden, A., Sembo, I., Redlund-Johnell, I., Dawson, A., . . . Jonsson, B. (2000). Long-term risk of osteoporotic fracture in Malmo. (0937-941X (Print)).

Kanis, J. A., Oden, A., McCloskey, E. V., Johansson, H., Wahl, D. A., & Cooper, C. (2012). A systematic review of hip fracture incidence and probability of fracture worldwide. *Osteoporos Int, 23*(9), 2239-2256. doi:10.1007/s00198-012-1964-3

Keaveny, T. M., & Hayes, W. C. (1993). A 20-year perspective on the mechanical properties of trabecular bone. (0148-0731 (Print)).

Keene, G. S., Parker, M. J., & Pryor, G. A. (1993). Mortality and morbidity after hip fractures. *BMJ (Clinical research ed.), 307*(6914), 1248-1250. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/8166806.

Kelly, T. L., Dm., S., & Neer, R. M. (1989). Calibration and standardization of bone mineral densitometers. (0884-0431 (Print)).

Kemp, J. P., Morris, J. A., Medina-Gomez, C., Forgetta, V., Warrington, N. M., Youlten, S. E., . . . Evans, D. M. (2017). Identification of 153 new loci associated with heel bone mineral density and functional involvement of GPC6 in osteoporosis. *Nat Genet, 49*(10), 1468-1475. doi:10.1038/ng.3949

Kerem, B., Jm., R., Buchanan, J. A., Markiewicz, D., Cox, T. K., Chakravarti, A., . . . Tsui, L. C. (1989). Identification of the cystic fibrosis gene: genetic analysis. (0036-8075 (Print)).

Khayat, M., Tilghman, J. M., Chervinsky, I., Zalman, L., Chakravarti, A., & Shalev, S. A. (2016). A PIGN mutation responsible for multiple congenital anomalies-

hypotonia-seizures syndrome 1 (MCAHS1) in an Israeli-Arab family. *Am J Med Genet A, 170a*(1), 176-182. doi:10.1002/ajmg.a.37375

Khazai, N., Judd Se, & Tangpricha, V. (2008). Calcium and vitamin D: skeletal and extraskeletal health. (1534-6307 (Electronic)).

Kiebzak, G. M., Perser, K., Ambrose Cg., & Heggeness, M. H. (2002). Undertreatment of osteoporosis in men with hip fracture. (0003-9926 (Print)).

Kim, S. K. (2018). Identification of 613 new loci associated with heel bone mineral density and a polygenic risk score for bone mineral density, osteoporosis and fracture. *PloS one, 13*(7), e0200785. doi:10.1371/journal.pone.0200785

Kim, S. K. (2018). Identification of 613 new loci associated with heel bone mineral density and a polygenic risk score for bone mineral density, osteoporosis and fracture. *PloS one, 13*(7), e0200785-e0200785. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/30048462. doi:10.1371/journal.pone.0200785

Kindmark, A., H., T., Johansson, A., Ljunghall, S., & Melhus, H. (1993). Reverse transcription-polymerase chain reaction assay demonstrates that the 9-cis retinoic acid receptor alpha is expressed in human osteoblasts. (0006-291X (Print)).

Klein, R. J., Zeiss., Chew, E. Y., Tsai, J.-Y., Sackler, R. S., Haynes, C., . . . Hoh, J. (2005). Complement factor H polymorphism in age-related macular degeneration. (1095-9203 (Electronic)).

Koller, D. L., Ichikawa, S., Lai, D., Padgett, L. R., Doheny, K. F., Pugh, E., . . . Foroud, T. (2010). Genome-wide association study of bone mineral density in premenopausal European-American women and replication in African-

American women. *The Journal of clinical endocrinology and metabolism, 95*(4), 1802-1809. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/20164292. doi:10.1210/jc.2009-1903

Kraft, P., Zeggini, E., & Ioannidis, J. P. A. (2009). Replication in genome-wide association studies. *Statistical science : a review journal of the Institute of Mathematical Statistics, 24*(4), 561-573. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/20454541. doi:10.1214/09-STS290

Krishnan, V., Bryant, H. U., & Macdougald, O. A. (2006). Regulation of bone mass by Wnt signaling. *J Clin Invest, 116*(5), 1202-1209. doi:10.1172/jci28551

Krum, S. A., Miranda-Carboni, G. A., Hauschka, P. V., Carroll, J. S., Lane, T. F., Freedman, L. P., & Brown, M. (2008). Estrogen protects bone by inducing Fas ligand in osteoblasts to regulate osteoclast survival. *Embo j, 27*(3), 535-545. doi:10.1038/sj.emboj.7601984

Kulie, T., Groff A, Redmer, J., Hounshell, J., Hounshell J, & Schrager, S. (2009). Vitamin D: an evidence-based review. (1558-7118 (Electronic)).

Kuo, T. R., & Chen, C. H. (2017). Bone biomarker for the clinical assessment of osteoporosis: recent developments and future perspectives. (2050-7771 (Print)).

Lanyon, L. E. (1993). Osteocytes, strain detection, bone modeling and remodeling. (0171-967X (Print)).

Laskey, M. A. (1996). Dual-energy X-ray absorptiometry and body composition. (0899-9007 (Print)).

Levy, J., Murray, E., Manolagas, S., & Olefsky, J. M. (1986). Demonstration of insulin receptors and modulation of alkaline phosphatase activity by insulin in rat osteoblastic cells. (0013-7227 (Print)).

Lewontin, R. C., & Kojima, K.-i. (1960). The Evolutionary Dynamics of Complex Polymorphisms. *Evolution, 14*(4), 458-472. Retrieved from http://www.jstor.org/stable/2405995. doi:10.2307/2405995

Liu, Y.-J., Papasian, C. J., Liu, J.-F., Hamilton, J., & Deng, H.-W. (2008). Is replication the gold standard for validating genome-wide association findings? *PloS one, 3*(12), e4037-e4037. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/19112512. doi:10.1371/journal.pone.0004037

Liu, Y. Z., Dvornyk, V., Lu, Y., Shen, H., Lappe, J. M., Recker, R. R., & Deng, H. W. (2005). A novel pathophysiological mechanism for osteoporosis suggested by an in vivo gene expression study of circulating monocytes. *J Biol Chem, 280*(32), 29011-29016. doi:10.1074/jbc.M501164200

Lorincz, C., Manske, S. L., & Zernicke, R. (2009). Bone health: part 1, nutrition. *Sports health, 1*(3), 253-260. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/23015881. doi:10.1177/1941738109334213

Mahdy, S., Al-Emadi Sa., Khanjar, I. A., Hammoudeh, M. M., Sarakbi, H. A., Siam, A. M., & Abdelrahman, M. O. (2010). Vitamin D status in health care professionals in Qatar. (0379-5284 (Print)).

Maller, J., George, S., Purcell, S., Fagerness, J., Altshuler, D., Daly, M. J., & Seddon, J. M. (2006). Common variation in three genes, including a noncoding variant

in CFH, strongly influences risk of age-related macular degeneration. *Nat Genet, 38*(9), 1055-1059. doi:10.1038/ng1873

Manolio, T. A., Collins, F. S., Cox, N. J., Goldstein, D. B., Hindorff, L. A., Hunter, D. J., . . . Visscher, P. M. (2009). Finding the missing heritability of complex diseases. *Nature, 461*(7265), 747-753. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/19812666. doi:10.1038/nature08494

Matsubara, R., Kukita T., Ichigi, Y., Takigawa, I., Qu, P.-F., Funakubo, N., . . . Kukita, A. (2012). Characterization and identification of subpopulations of mononuclear preosteoclasts induced by TNF-alpha in combination with TGF-beta in rats. (1932-6203 (Electronic)).

Maurano, M. T., Humbert, R., Rynes, E., Thurman, R. E., Haugen, E., Wang, H., . . . Stamatoyannopoulos, J. A. (2012). Systematic localization of common disease-associated variation in regulatory DNA. *Science, 337*(6099), 1190-1195. doi:10.1126/science.1222794

Meddeb, N., Sahli H., Chahed, M., Abdelmoula, J., Feki, M., Salah, H., . . . Sellami, S. (2005). Vitamin D deficiency in Tunisia. (0937-941X (Print)).

Medina-Gomez, C., Kemp, J. P., Estrada, K., Eriksson, J., Liu, J., Reppe, S., . . . Rivadeneira, F. (2012). Meta-analysis of genome-wide scans for total body BMD in children and adults reveals allelic heterogeneity and age-specific effects at the WNT16 locus. *PLoS Genet, 8*(7), e1002718. doi:10.1371/journal.pgen.1002718

Melton, L. J., 3rd, Atkinson Ej Fau ., O'Connor, M. K., O'Fallon, W. M., & Riggs, B. L. (1998). Bone density and fracture risk in men. (0884-0431 (Print)).

Melton, L. J., 3rd, Chrischilles Ea Fau., Cooper, C., Lane, A. W., & Riggs, B. L. (1992). Perspective. How many women have osteoporosis? (0884-0431 (Print)).

Messina-Baas, O., Gonzalez-Garay, M. L., González-Huerta, L. M., Toral-López, J., & Cuevas-Covarrubias, S. A. (2016). Whole Exome Sequencing Reveals a Mutation in CRYBB2 in a Large Mexican Family with Autosomal Dominant Pulverulent Cataract. *Molecular syndromology, 7*(2), 87-92. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/27385965. doi:10.1159/000445669

Michaelsson, K., Melhus, H., Bellocco, R., & Wolk, A. (2003). Dietary calcium and vitamin D intake in relation to osteoporotic fracture risk. *Bone, 32*(6), 694-703.

Milat, F., & Ng, K. W. (2009). Is Wnt signalling the final common pathway leading to bone formation? *Mol Cell Endocrinol, 310*(1-2), 52-62. doi:10.1016/j.mce.2009.06.002

Minisola, S. A.-O. h. o. o., Cipriani, C., Occhiuto, M., & Pepe, J. (2017). New anabolic therapies for osteoporosis. (1970-9366 (Electronic)).

Mizrahi, J., Keaveny, T. M., Edwards, W. T., & Hayes, W. C. (1976). Finite-element stress analysis of the normal and osteoporotic lumbar vertebral body. (0362-2436 (Print)).

Mizuno, Y., Hosoi T., Inoue, S., Ikegami, A., Kaneki, M., Akedo, Y., . . . Orimo, H. (1994). Immunocytochemical identification of androgen receptor in mouse osteoclast-like multinucleated cells. (0171-967X (Print)).

Moayyeri, A., Hsu, Y. H., Karasik, D., Estrada, K., Xiao, S. M., Nielson, C., . . . Kaptoge, S. K. (2014). Genetic determinants of heel bone properties: genome-

wide association meta-analysis and replication in the GEFOS/GENOMOS consortium. *Hum Mol Genet, 23*(11), 3054-3068. doi:10.1093/hmg/ddt675

Mullin, B. H., Zhao, J. H., Brown, S. J., Perry, J. R. B., Luan, J., Zheng, H. F., . . . Wilson, S. G. (2017). Genome-wide association study meta-analysis for quantitative ultrasound parameters of bone identifies five novel loci for broadband ultrasound attenuation. *Hum Mol Genet, 26*(14), 2791-2802. doi:10.1093/hmg/ddx174

Munch, S., & Shapiro, S. (2006). The silent thief: osteoporosis and women's health care across the life span. (0360-7283 (Print)).

National Institutes of Health. (2017). NIH OSTEOPOROSIS AND RELATED BONE DISEASES

National Osteoporosis Foundation. (2012). Bone health basics.

Nevitt, M. C., Ettinger B., Black, D. M., Stone, K., Jamal, S. A., Ensrud, K., . . . Cummings, S. R. (1998). The association of radiographically detected vertebral fractures with back pain and function: a prospective study. (0003-4819 (Print)).

Nicod, J.-C., & Largiadèr, C. R. (2003). SNPs by AFLP (SBA): a rapid SNP isolation strategy for non-model organisms. *Nucleic acids research, 31*(5), e19-e19. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/12595568.

Nordin, M. (2012). Basic Biomechanics of the musculoskeletal system.

O'Neill, T. W., & Roy, D. K. (2005). How many people develop fractures with what outcome? *Best Practice & Research Clinical Rheumatology, 19*(6), 879-895. Retrieved from

http://www.sciencedirect.com/science/article/pii/S1521694205000732. doi:https://doi.org/10.1016/j.berh.2005.06.003

O'Sullivan, M., Suibhne, N., Cox, G., Healy, M., & O'Morain, C. (2008). High prevalence of vitamin D insufficiency in healthy Irish adults. (1863-4362 (Electronic)).

Ohba, C., Okamoto, N., Murakami, Y., Suzuki, Y., Tsurusaki, Y., Nakashima, M., . . . Saitsu, H. (2014). PIGN mutations cause congenital anomalies, developmental delay, hypotonia, epilepsy, and progressive cerebellar atrophy. *Neurogenetics, 15*(2), 85-92. doi:10.1007/s10048-013-0384-7

Olivier, M. (2003). A haplotype map of the human genome. (1531-2267 (Electronic)).

Osterhoff, G., Morgan, E. F., Shefelbine, S. J., Karim, L., McNamara, L. M., & Augat, P. (2016). Bone mechanical properties and changes with osteoporosis. *Injury, 47 Suppl 2*(Suppl 2), S11-S20. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/27338221. doi:10.1016/S0020-1383(16)47003-8

Panagiotou, O. A., & Ioannidis, J. P. (2012). What should the genome-wide significance threshold be? Empirical replication of borderline genetic associations. *Int J Epidemiol, 41*(1), 273-286. doi:10.1093/ije/dyr178

Papaioannou, A., Kennedy Cc Fau - Ioannidis, G., Ioannidis G Fau - Gao, Y., Gao Y Fau - Sawka, A. M., Sawka Am Fau - Goltzman, D., Goltzman D Fau - Tenenhouse, A., . . . Adachi, J. D. (2008). The osteoporosis care gap in men with fragility fractures: the Canadian Multicentre Osteoporosis Study. (0937-941X (Print)).

Peacock, M., Econs, M. J., Turner, C. H., & Foroud, T. (2002). Genetics of Osteoporosis. *Endocrine Reviews, 23*(3), 303-326. Retrieved from https://doi.org/10.1210/edrv.23.3.0464. doi:10.1210/edrv.23.3.0464

Peterson, Y. K., & Luttrell, L. M. (2017). The Diverse Roles of Arrestin Scaffolds in G Protein-Coupled Receptor Signaling. *Pharmacological reviews, 69*(3), 256-297. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/28626043. doi:10.1124/pr.116.013367

Pittas, A. G., Lau J., Hu, F. B., & Dawson-Hughes, B. (2007). The role of vitamin D and calcium in type 2 diabetes. A systematic review and meta-analysis. (0021-972X (Print)).

Price, A. L., Patterson, N. J., Plenge, R. M., Weinblatt, M. E., Shadick, N. A., & Reich, D. (2006). Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet, 38*(8), 904-909. doi:10.1038/ng1847

Price, A. L., Patterson Nj., Plenge, R. M., Weinblatt, M. E., Shadick, N. A., & Reich, D. (2006). Principal components analysis corrects for stratification in genome-wide association studies. (1061-4036 (Print)).

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D., . . . Sham, P. C. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet, 81*(3), 559-575. doi:10.1086/519795

Purcell, S., Neale B., Todd-Brown, K., Todd-Brown K., Thomas, L., Ferreira, M. A. R., . . . Sham, P. C. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. (0002-9297 (Print)).

Rafiq, S., Khan, S., Tapper, W., Collins, A., Upstill-Goddard, R., Gerty, S., . . . Eccles, D. (2014). A genome wide meta-analysis study for identification of common variation associated with breast cancer prognosis. *PloS one, 9*(12), e101488-e101488. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/25526632. doi:10.1371/journal.pone.0101488

Rees A., Shoulders, C. C., Stocks, J., Galton, D. J., & Baralle, F. E. (1983). DNA polymorphism adjacent to human apoprotein A-1 gene: relation to hypertriglyceridaemia. (0140-6736 (Print)).

Reich, D. E., & Goldstein, D. B. (2001). Detecting association in a case-control study while correcting for population stratification. *Genet Epidemiol, 20*(1), 4-16. doi:10.1002/1098-2272(200101)20:1<4::Aid-gepi2>3.0.Co;2-t

Reid, I. R., Bolland, M. J., & Grey, A. (2014). Effects of vitamin D supplements on bone mineral density: a systematic review and meta-analysis. *Lancet, 383*(9912), 146-155. doi:10.1016/s0140-6736(13)61647-5

Reinholt, F. P., Hultenby K Fau - Oldberg, A., Oldberg A Fau - Heinegard, D., & Heinegard, D. (1990). Osteopontin--a possible anchor of osteoclasts to bone. (0027-8424 (Print)).

Rietveld, C. A., Conley, D., Eriksson, N., Esko, T., Medland, S. E., Vinkhuyzen, A. A. E., . . . Social Science Genetics Association, C. (2014). Replicability and robustness of genome-wide-association studies for behavioral traits. *Psychological science, 25*(11), 1975-1986. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/25287667. doi:10.1177/0956797614545132

Riggs Bl Fau., & Melton, L. J. (1983). Evidence for two distinct syndromes of involutional osteoporosis. (0002-9343 (Print)).

Riggs BL., & Melton LJ. (1992). The prevention and treatment of osteoporosis. *N Engl J Med*.

Ritchie, M. D., Denny Jc Fau - Crawford, D. C., Crawford Dc Fau - Ramirez, A. H., Ramirez Ah Fau - Weiner, J. B., Weiner Jb Fau - Pulley, J. M., Pulley Jm Fau - Basford, M. A., . . . Roden, D. M. (2010). Robust replication of genotype-phenotype associations across multiple diseases in an electronic medical record. (1537-6605 (Electronic)).

Rizzoli, R., Poser, J., & Bürgi, U. (1986). Nuclear thyroid hormone receptors in cultured bone cells. *Metabolism, 35*(1), 71-74. Retrieved from http://www.sciencedirect.com/science/article/pii/0026049586900983. doi:https://doi.org/10.1016/0026-0495(86)90098-3

Robbins, J., Hirsch C., Whitmer R., Cauley J., & Harris, T. (2001). The association of bone mineral density and depression in an older population. (0002-8614 (Print)).

Sadat-Ali, M., Al-Habdan, I. M., Al-Turki, H. A., & Azam, M. Q. (2012). An epidemiological analysis of the incidence of osteoporosis and osteoporosis-related fractures among the Saudi Arabian population. *Ann Saudi Med, 32*(6), 637-641. doi:10.5144/0256-4947.2012.637

Sahu, M., & Prasuna, J. G. (2016). Twin Studies: A Unique Epidemiological Tool. *Indian journal of community medicine : official publication of Indian Association of Preventive & Social Medicine, 41*(3), 177-182. Retrieved from

https://www.ncbi.nlm.nih.gov/pubmed/27385869. doi:10.4103/0970-0218.183593

Sanders, K. M., Stuart, A. L., Williamson, E. J., Simpson, J. A., Kotowicz, M. A., Young, D., & Nicholson, G. C. (2010). Annual high-dose oral vitamin D and falls and fractures in older women: a randomized controlled trial. *Jama, 303*(18), 1815-1822. doi:10.1001/jama.2010.594

Schaffler, M. B., & Kennedy, O. D. (2012). Osteocyte signaling in bone. *Current osteoporosis reports, 10*(2), 118-125. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/22552701. doi:10.1007/s11914-012-0105-4

Schork, N. J., Murray, S. S., Frazer, K. A., & Topol, E. J. (2009). Common vs. rare allele hypotheses for complex diseases. *Current opinion in genetics & development, 19*(3), 212-219. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/19481926. doi:10.1016/j.gde.2009.04.010

Seeman, E., & Delmas, P. D. (2006). Bone quality--the material and structural basis of bone strength and fragility. (1533-4406 (Electronic)).

Sharma, S., Tandon, V., Mahajan, A., Kour, A., & Kumar, D. (2006). *Preliminary screening of osteoporosis and osteopenia in urban women from Jammu using calcaneal QUS Indian* (Vol. 60).

Shi, C., Iura, A., Terajima, M., Liu, F., Lyons, K., Pan, H., . . . Sun, H. (2016). Deletion of BMP receptor type IB decreased bone mass in association with compromised osteoblastic differentiation of bone marrow mesenchymal

progenitors. *Scientific reports, 6*, 24256-24256. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/27048979. doi:10.1038/srep24256

Shikada, Y., Yonemitsu Y., Koga, T., Onimaru, M., Nakano, T., Okano, S., . . . Sueishi, K. (2005). Platelet-derived growth factor-AA is an essential and autocrine regulator of vascular endothelial growth factor expression in non-small cell lung carcinomas. (0008-5472 (Print)).

Shilbayeh, S. (2003). Prevalence of osteoporosis and its reproductive risk factors among Jordanian women: a cross-sectional study. *Osteoporos Int, 14*(11), 929-940. doi:10.1007/s00198-003-1458-4

Skol, A. D., Scott, L. J., Abecasis, G. R., & Boehnke, M. (2006). Joint analysis is more efficient than replication-based analysis for two-stage genome-wide association studies. *Nat Genet, 38*(2), 209-213. doi:10.1038/ng1706

Slatkin, M. (2008). Linkage disequilibrium--understanding the evolutionary past and mapping the medical future. *Nature reviews. Genetics, 9*(6), 477-485. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/18427557

https://www.ncbi.nlm.nih.gov/pmc/PMC5124487/. doi:10.1038/nrg2361

Smith, H., Anderson, F., Raphael, H., Maslin, P., Crozier, S., & Cooper, C. (2007). Effect of annual intramuscular vitamin D on fracture risk in elderly men and women--a population-based, randomized, double-blind, placebo-controlled trial. *Rheumatology (Oxford), 46*(12), 1852-1857. doi:10.1093/rheumatology/kem240

Soontrapa, S., & Chailurkit, L.-O. (2009). Hypovitaminosis D in Thailand. (0125-2208 (Print)).

Sözen, T., Özışık, L., & Başaran, N. Ç. (2017). An overview and management of osteoporosis. *European Journal of Rheumatology, 4*(1), 46-56. Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC5335887/. doi:10.5152/eurjrheum.2016.048

Stadler, Z. K., Thom, P., Robson, M. E., Weitzel, J. N., Kauff, N. D., Hurley, K. E., . . . Offit, K. (2010). Genome-wide association studies of cancer. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology, 28*(27), 4255-4267. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/20585100 https://www.ncbi.nlm.nih.gov/pmc/PMC2953976/. doi:10.1200/JCO.2009.25.7816

Stenbeck, G. (2002). Formation and function of the ruffled border in osteoclasts. (1084-9521 (Print)).

Studies, N.-N. W. G. o. R. i. A., Chanock, S. J., Manolio, T., Boehnke, M., Boerwinkle, E., Hunter, D. J., . . . Collins, F. S. (2007). Replicating genotype–phenotype associations. *Nature, 447*, 655. Retrieved from https://doi.org/10.1038/447655a. doi:10.1038/447655a

Sweileh, W. M., Al-Jabi, S. W., Zyoud, S. e. H., Sawalha, A. F., & Ghanim, M. A. (2014). Osteoporosis is a neglected health priority in Arab World: a comparative bibliometric analysis. *SpringerPlus, 3*, 427-427. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/25152852 https://www.ncbi.nlm.nih.gov/pmc/PMC4141069/. doi:10.1186/2193-1801-3-427

Tang, B. M., Eslick, G. D., Nowson, C., Smith, C., & Bensoussan, A. (2007). Use of calcium or calcium in combination with vitamin D supplementation to prevent

fractures and bone loss in people aged 50 years and older: a meta-analysis. *Lancet, 370*(9588), 657-666. doi:10.1016/s0140-6736(07)61342-7

Tarride, J. E., Hopkins, R. B., Leslie, W. D., Morin, S., Adachi, J. D., Papaioannou, A., . . . Goeree, R. (2012). The burden of illness of osteoporosis in Canada. *Osteoporosis international : a journal established as result of cooperation between the European Foundation for Osteoporosis and the National Osteoporosis Foundation of the USA, 23*(11), 2591-2600. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/22398854

https://www.ncbi.nlm.nih.gov/pmc/PMC3483095/. doi:10.1007/s00198-012-1931-z

Teslovich, T. M., Musunuru K., Smith, A. V., Edmondson, A. C., Stylianou, I. M., Koseki, M., . . . Kathiresan, S. (2010). Biological, clinical and population relevance of 95 loci for blood lipids. (1476-4687 (Electronic)).

Teti, A., Rizzoli R Fau - Zambonin Zallone, A., & Zambonin Zallone, A. (1991). Parathyroid hormone binding to cultured avian osteoclasts. (0006-291X (Print)).

Thomas, D. M., Udagawa N., Hards, D. K., Quinn, J. M., Moseley, J. M., Findlay, D. M., & Best, J. D. (1998). Insulin receptor expression in primary and cultured osteoclast-like cells. (8756-3282 (Print)).

Tian, L., Yang, R., Wei, L., Liu, J., Yang, Y., Shao, F., . . . Guo, T. (2017). Prevalence of osteoporosis and related lifestyle and metabolic factors of postmenopausal women and elderly men: A cross-sectional study in Gansu province, Northwestern of China. *Medicine, 96*(43), e8294-e8294. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/29068999

https://www.ncbi.nlm.nih.gov/pmc/PMC5671832/. doi:10.1097/MD.0000000000008294

Tønnesen, R., Hovind, P. H., Jensen, L. T., & Schwarz, P. (2016). Determinants of vitamin D status in young adults: influence of lifestyle, sociodemographic and anthropometric factors. *BMC public health, 16*, 385-385. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/27170258

https://www.ncbi.nlm.nih.gov/pmc/PMC4863340/. doi:10.1186/s12889-016-3042-9

Tosteson, A. N., Gabriel Se., Grove Mr ., Moncur, M. M., Kneeland, T. S., & Melton, L. J., 3rd. (2001). Impact of hip and vertebral fractures on quality-adjusted life years. (0937-941X (Print)).

Trajanoska, K., Morris, J. A., Oei, L., Zheng, H.-F., Evans, D. M., Kiel, D. P., . . . Rivadeneira, F. (2018). Assessment of the genetic and clinical determinants of fracture risk: genome wide association and mendelian randomisation study. *BMJ, 362*, k3225. Retrieved from http://www.bmj.com/content/362/bmj.k3225.abstract. doi:10.1136/bmj.k3225

Turner, S., Armstrong, L. L., Bradford, Y., Carlson, C. S., Crawford, D. C., Crenshaw, A. T., . . . Ritchie, M. D. (2011). Quality control procedures for genome-wide association studies. *Current protocols in human genetics, Chapter 1*, Unit1.19-Unit11.19. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/21234875

https://www.ncbi.nlm.nih.gov/pmc/PMC3066182/. doi:10.1002/0471142905.hg0119s68

Turner, S., Armstrong, L. L., Bradford, Y., Carlson, C. S., Crawford, D. C., Crenshaw, A. T., . . . Ritchie, M. D. (2011). Quality control procedures for

genome-wide association studies. *Curr Protoc Hum Genet, Chapter 1*, Unit1.19. doi:10.1002/0471142905.hg0119s68

Udagawa, N., Takahashi, N., Akatsu, T., Tanaka, H., Sasaki, T., Nishihara, T., . . . Suda, T. (1990). Origin of osteoclasts: mature monocytes and macrophages are capable of differentiating into osteoclasts under a suitable microenvironment prepared by bone marrow-derived stromal cells. *Proceedings of the National Academy of Sciences of the United States of America, 87*(18), 7260-7264. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/2169622 https://www.ncbi.nlm.nih.gov/pmc/PMC54723/.

van Leeuwen, E. M., Smouter, F. A. S., Kam-Thong, T., Karbalai, N., Smith, A. V., Harris, T. B., . . . van Duijn, C. M. (2014). The challenges of genome-wide interaction studies: lessons to learn from the analysis of HDL blood levels. *PloS one, 9*(10), e109290-e109290. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/25329471 https://www.ncbi.nlm.nih.gov/pmc/PMC4203717/. doi:10.1371/journal.pone.0109290

Warshawsky H., Goltzman, D., Rouleau, M. F., & Bergeron, J. J. (1980). Direct in vivo demonstration by radioautography of specific binding sites for calcitonin in skeletal and renal tissues of the rat. (0021-9525 (Print)).

Wei, L. L., Leach Mw., -., Miner, R. S., & Demers, L. M. (1993). Evidence for progesterone receptors in human osteoblast-like cells. (0006-291X (Print)).

Wiktorowicz, M. E., Goeree R Fau - Papaioannou, A., Papaioannou A Fau - Adachi, J. D., Adachi Jd Fau - Papadimitropoulos, E., & Papadimitropoulos, E. (2001).

Economic implications of hip fracture: health service use, institutional care and cost in Canada. (0937-941X (Print)).

Wittke-Thompson, J. K., Pluzhnikov, A., & Cox, N. J. (2005). Rational inferences about departures from Hardy-Weinberg equilibrium. *Am J Hum Genet, 76*(6), 967-986. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/15834813 https://www.ncbi.nlm.nih.gov/pmc/PMC1196455/. doi:10.1086/430507

Xu, L. X., Kukita T Fau - Nakano, Y., Nakano Y Fau - Yu, H., Yu H Fau - Hotokebuchi, T., Hotokebuchi T Fau - Kuratani, T., Kuratani T Fau - Iijima, T., . . . Koga, T. (1996). Osteoclasts in normal and adjuvant arthritis bone tissues express the mRNA for both type I and II interleukin-1 receptors. (0023-6837 (Print)).

Yang, X., Zhang, Y., Li, Y., & Wen, T. (2018). MALAT1 enhanced the proliferation of human osteoblasts treated with ultra-high molecular weight polyethylene by targeting VEGF via miR-22-5p. *International journal of molecular medicine, 41*(3), 1536-1546. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/29328414 https://www.ncbi.nlm.nih.gov/pmc/PMC5819924/. doi:10.3892/ijmm.2018.3363

Yodsurang, V., Tang, Y., Takahashi, Y., Tanikawa, C., Kamatani, Y., Takahashi, A., . . . Matsuda, K. (2018). Genome-wide association study (GWAS) of ovarian cancer in Japanese predicted regulatory variants in 22q13.1. *PLoS One, 13*(12), e0209096. doi:10.1371/journal.pone.0209096

Zhang, F., Wang, Y., & Deng, H.-W. (2008). Comparison of population-based association study methods correcting for population stratification. *PloS one,*

3(10), e3392-e3392. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/18852890 https://www.ncbi.nlm.nih.gov/pmc/PMC2562035/. doi:10.1371/journal.pone.0003392

Zhang, Z., Chen J Fau - Jin, D., & Jin, D. (1998). Platelet-derived growth factor (PDGF)-BB stimulates osteoclastic bone resorption directly: the role of receptor beta. (0006-291X (Print)).

Zheng, H. F., Forgetta, V., Hsu, Y. H., Estrada, K., Rosello-Diez, A., Leo, P. J., . . . Richards, J. B. (2015). Whole-genome sequencing identifies EN1 as a determinant of bone density and fracture. *Nature, 526*(7571), 112-117. doi:10.1038/nature14878

Zhu, Y., Xu, G., Patel, A., McLaughlin, M. M., Silverman, C., Knecht, K., . . . Young, P. R. (2002). Cloning, expression, and initial characterization of a novel cytokine-like gene family. *Genomics, 80*(2), 144-150.

APPENDIX A: ETHICAL APPROVAL

QATAR BIOBANK INSTITUTIONAL REVIEW BOARD

Email: qbbresearch@qf.org.qa

QBB IRB MOPH Registration: MOPH-QBB-IRB-011
QBB IRB MOPH Assurance: MOPH-A-QBB-000222

Date	02/07/2018
Lead Principal Investigator	Dr. Mohammad Haris
Committee Action	**Expedited Approval**
Approval Date:	02/07/2018
IRB Protocol #:	E -2018-QBB- RES-ACC-0112-0054
Study Title	Genome-wide association study (GWAS) to uncover genetic risk factors associated with low bone mineral density and Osteoporosis in Qatari population.
Expiration Date	02/07/2018 -**01/07/2019**

Dear PI,

The above-referenced protocol was approved following expedited review by the QBB Institutional Review Board.

The project involves no more than minimal risk to subjects and qualifies under Expedited Category 4: Research involving materials (data, documents, records, or specimens) that: have been collected; or, Will be collected solely for non-research purposes (such as medical treatment or diagnosis).

It is the Principal Investigator's responsibility to obtain IRB review and continued approval before one month from the expiration date.

You may not continue any research activity beyond the expiration date without approval by the Institutional Review Board.
It is a condition of this approval that you report promptly to the IRB any serious, unanticipated adverse events experienced by subjects in the course of this research, whether or not they are directly related to the study protocol.

Amendments: If you wish to change any aspect of this study, such as the procedures, the consent forms, or the investigators, please communicate your requested changes to the IRB. The new procedure is not to be initiated until the IRB approval has been given.

Please retain a copy of this letter with your approved protocol.

Sincerely yours,

Dr. Khalid Al-Ansari
Chair, Institutional Review Board QBB

Qatar Biobank P.O.Box 5825 Doha Qatar Tel: +974 4 454 8351 www.qatarbiobank.org.qa

QBB- IRB, QBB-RES-ACC-0112 0054, 2 July 2018- 1 July 2019