USE OF THE COLLABORATIVE CROSS GENE MINE MOUSE PHENOTYPE LIBRARY TO IDENTIFY NOVEL GENES REGULATING BONE MASS AND BONE ARCHITECTURE

JINBO YUAN, MBBS, MMedSci

This thesis is presented for the degree of Doctor of Philosophy of the University of Western Australia
School of Biomedical Sciences,
Division of Pathology and Laboratory Medicine,
Queen Elizabeth II Medical Centre, Nedlands, Perth, Australia
2017
THESIS DECLARATION

I, Jinbo Yuan, certify that:

This thesis has been substantially accomplished during enrolment in the degree.

This thesis does not contain material which has been accepted for the award of any other degree or diploma in my name, in any university or other tertiary institution.

No part of this work will, in the future, be used in a submission in my name, for any other degree or diploma in any university or other tertiary institution without the prior approval of The University of Western Australia and where applicable, any partner institution responsible for the joint-award of this degree.

This thesis does not contain any material previously published or written by another person, except where due reference has been made in the text.

The work(s) are not in any way a violation or infringement of any copyright, trademark, patent, or other rights whatsoever of any person.

The research involving human data reported in this thesis was assessed and approved by The University of Western Australia Human Research Ethics Committee. Approval #: HREC2010-253 (Sir Charles Gairdner Hospital).

The research involving animal data reported in this thesis was assessed and approved by The University of Western Australia Animal Ethics Committee. Approval #: RA/3/500/87.

The research involving animals reported in this thesis followed The University of Western Australia and national standards for the care and use of laboratory animals.

The work described in this thesis was funded by Scholarship for International Research Fees and APP1127156 - 'Gene mining for novel molecular determinants of the skeleton'.

Technical assistance was kindly provided by Ms Diana Engineer for the use of microCT, Prof Grant Morahan and Dr Ramesh Ram for gene mapping using GeneMiner, and Mr Kai Chen and Ziyi Wang (The University of Western Australia) for primary cells culture, RNA extraction and performed RT-PCR that is described in Chapter 4, Sections 4.23, 4.25 and 4.3. Correlation with human datasets was performed by Dr Benjamin Mullin (The University of Western Australia) that is described in Chapter 4, Section 4.27.

This thesis does not contain work that I have published, nor work under review for publication.

Signature:

Date: 13/12/2017
ABSTRACT

Osteoporosis is a common systemic skeletal disease associated with low bone mineral density and micro-architectural deterioration of bone tissue leading to an increased risk of fractures. It is a complex disease with contributions from multiple genetic loci and multiple environmental factors and their interactions.

Osteoporosis susceptibility genes are involved in many biological pathways such as RANK/RANKL/OPG, Wnt, endochondral ossification and BMP/TGF-β pathways. Genome-wide association study (GWAS) has successfully identified more than 70 osteoporosis susceptibility genes with genome-wide significance, and another 160 genes that have been suggested to be correlated with BMD or osteoporosis via various genetic approaches. These total 230 genes are estimated to explain up to 30% genetic variation of BMD or osteoporosis. Therefore, the vast majority of heritability of BMD or osteoporosis is missing and requires further understanding of the bone genetics.

The Collaborative Cross (CC) is a large-scale project for generating recombinant inbred strains from eight parental strains that were carefully selected to catch approximately 90% of genetic variation in the mouse genome. Only one study using a small number of CC mice to identify novel genes for bone microarchitecture has been performed previously. In the present study, over 900 CC mice across 70 strains incorporating varying ages and genders were dissected and scanned using microCT. We then analysed parameters including BV/TV, Tb.N, Tb.Sp, Tb.Th, SMI, DA and Ct.Th from reconstructed femur images. QTL haplotype mapping was performed on a total of 848 mice (strains N = 61) to identify candidate genes responsible for bone mass and microarchitecture. Correlation of these genes with variation in human osteoporosis datasets (including GEFOS, UK10K, and TwinsUK) was also performed by Dr Benjamin Mullin.

Over 20 loci that reached genome-wide significance (GWS) were revealed from various traits and more than 80 genes previously associated with bone formation and remodelling from published and unpublished data were pulled out from those loci. In particular, DA in female CC mice was the most promising trait, harvesting a total of six loci that achieved the 95th percentile GWS. It has yielded several strong osteoporosis susceptibility candidate genes such as Setbp1, Mbd1, Nfatc1, Dym, Skor2, Zbtb7c, Sall3, Spp1, Mapk4, Smad2 and Smad7 on chromosome 18, and Col4a2 and Gas6 on chromosome 8. DA in males also identified Wnt3, Wnt9b, Axin2, Gh1, Itga2b, Adam11, Smurf2, Map2k6 and Map3k14 on chromosome 11. Trabecular BV/TV and Tb. N in the older female cohort also identified candidate genes Dnah2 and Hic1 from the locus, which achieved 95th percentile GWS. QTL haplotype mapping of bone volume parameters (BV/TV) in middle aged female mice identified
a locus harbouring genes including *Tnrsf13b, Per1, Alox15b, Aloxe3, Alox12, Alox15* and *Traf4*. QTL scan of BV/TV also pulled out genes *Cdnt* and *Jhy* in male cohort.

Analyses of candidate genes identified in mice within human datasets showed 10 compelling candidates for bone traits including *C11orf63 (JHY), DNAH2, DYM, EN1, HIC1, SETBP1, TNFRSF13B, WNT3, WNT9B* and *ZBTB7C*. These genes either reached gene-wide significance threshold (p ≤ 0.00027778 after Bonferroni correction) or had a top SNP that is above the single-point suggestive threshold (p ≤ 3.00×10⁻⁵). In addition, *NFATC1* reached the gene-wide suggestive threshold (p ≤ 0.001) with p value 0.000678. *EN1, WNT3* and *WNT9B* have previously been associated with BMD at the GWS level. Furthermore, over 50 genes achieved p < 0.05 in the gene-wide scan, these genes were considered “nominally significant associations” with bone traits.

To further characterise the functions of those novel candidate genes associated with both architectural parameters in mice, and human BMD and osteoporosis, we accumulated evidence from web-based resources such as relevant publications, BioGPS, Uniprot, and performed RT-PCR to test the expression profiles of the novel gene *Setbp1* in osteoblasts and osteoclasts. This gene did not show obvious regulation during osteoblastic differentiation, but was downregulated during osteoclastogenesis. We have attempted to build protein-to-protein interaction (PPI) networks for all candidate osteoporosis susceptibility genes identified in the present study. In the future the application of animal models will assist in unveiling the mechanisms underlying the various pathways implicated in this study.

In summary, these results verify the success of our screening program for identification of novel osteoporosis susceptibility genes, and have identified several new genes for further study.
# TABLE OF CONTENTS

**THESIS DECLARATION** ..................................................................................................................II

**ABSTRACT** ..................................................................................................................................III

**ACKNOWLEDGEMENT** ......................................................................................................................X

**PUBLISHED ABSTRACTS AND ARTICLES** .....................................................................................XII

**ABBREVIATIONS** ...............................................................................................................................XIV

**LIST OF FIGURES AND TABLES** ........................................................................................................XXII

**Chapter 1 Bone Biology** ..................................................................................................................1

1.1 Bone Morphology and Organisation ...............................................................................................2

1.1.1 Bone Composition and Organisation ..........................................................................................2

1.1.2 The Functions of Bone ................................................................................................................3

1.2 Bone Cells .....................................................................................................................................4

1.2.1 Osteoblasts ..................................................................................................................................4

1.2.1.1 *Osteoblast Morphology and Function* ....................................................................................4

1.2.1.2 *Osteoblast Formation and Differentiation* ............................................................................5

1.2.1.3 *Regulation of Osteoblast Differentiation* ............................................................................5

1.2.2 Osteocytes ...................................................................................................................................11

1.2.2.1 *Osteocyte Morphology and Functions* ...............................................................................11

1.2.2.2 *Formation and Apoptosis of Osteocytes* ............................................................................12

1.2.2.3 *Regulation of Osteocytes* ....................................................................................................13

1.2.3 Osteoclasts ..................................................................................................................................14

1.2.3.1 *Osteoclast Morphology and Function* ...............................................................................14

1.2.3.2 *Osteoclast Differentiation and Fusion* ..............................................................................15

1.2.3.3 *Regulation of Osteoclast Differentiation and Survival* ....................................................15

1.3 Bone Modelling and Remodelling .................................................................................................18

1.3.1 Skeletal Development and Fracture Healing .............................................................................18

1.3.2 Bone Remodelling ......................................................................................................................19

1.3.3 Cell-to-Cell Communication ......................................................................................................20

1.4 Key Biological Pathways Regulating Bone Homeostasis ...............................................................20
4.1.1 Equipment .............................................................................................................. 67
4.1.2 Chemical Reagents ............................................................................................... 68
4.1.3 Buffers and Solutions ............................................................................................ 69
4.1.4 Tissue Culture Materials and Reagents ................................................................. 71
4.1.5 Oligonucleotide Primers and Housekeeping Genes Primers .................................... 72
4.1.6 Softwares ................................................................................................................ 73
4.1.7 Gene Mapping Links .............................................................................................. 73
4.1.8 Bioinformatics Tools .............................................................................................. 73
4.2 Animals and Methods for Bone Phenotyping and Gene Mapping ................................ 74
4.2.1 Generation of the Collaborative Cross ................................................................... 74
4.2.2 Work Flow of the CC mice Screening ..................................................................... 75
4.2.2.1 MicroCT Analysis of CC Mice ........................................................................... 75
4.2.2.2 Strains Ranking ................................................................................................. 77
4.2.2.3 Haplotype Reconstruction and QTL Mapping of CC Strains .............................. 77
4.2.2.4 Verification of Candidate Genes Using Bioinformatics Approaches .................. 78
4.2.2.5 Correlation with Human Datasets through GWAS ............................................. 80
4.3 Molecular Biology Techniques .................................................................................. 80
4.3.1 Isolation and Culture of Primary Cells .................................................................... 80
4.3.1.1 Isolation and Culture of Bone Marrow Cells ...................................................... 80
4.3.1.2 Isolation and Culture of Osteoblasts ................................................................. 81
4.3.1.3 Osteoclast Culture ............................................................................................ 81
4.3.1.4 Splitting Primary Cells ...................................................................................... 81
4.3.1.5 Cryopreservation of Primary Cells for Long Term Storage ............................... 82
4.3.2 RNA Extraction from Cultured Cells ...................................................................... 82
4.3.3 Real Time Reverse Transcription-Polymerase Chain Reaction (RT-PCR) ............... 82
4.3.3.1 Reverse Transcription and Amplification ......................................................... 82
4.3.3.2 Real Time RT-PCR Setup ................................................................................ 83
4.3.3.3 Real Time-PCR Analysis .................................................................................. 84
7.2.2.3 Novel Genes with Nominally Significant Associations with Bone Traits in Human Populations ........................................................................................................... 194

7.2.2.4 Previously Reported Genes with Nominally Significant Associations with Bone Phenotypes in Human Populations ........................................................................ 204

7.2.2.5 Genes Failed to Correlate with Bone Phenotypes in Human Populations .................. 208

7.2.3 Building PPI Network for Osteoporosis Susceptibility Genes ........................................ 208

7.3 Discussion ............................................................................................................................. 212

Chapter 8 General Overview and Future Directions .................................................................... 222

8.1 Overview of Thesis and General Discussion ........................................................................ 223

8.2 Future Directions .................................................................................................................. 226

Chapter 9 References .................................................................................................................. 229

Appendix ..................................................................................................................................... 305
ACKNOWLEDGEMENTS

I would like to thank to my coordinating supervisor, W/Prof. Jiake Xu whom I have known since the first day when I arrived on the soil of Australia in the year of 2003. Jiake has been so kind to me not only academically but mentally, from providing free accommodation in his house for the first three weeks after arrival, giving support to my previous study at UWA for my master of medical science degree, and to offering opportunity and support to apply scholarships for my PhD degree at this university. It has been a wonderful and rewarding journey in the Cell Signalling group since the first day of enrollment. The joy and excitement gained from this journey is far greater than the intermittent period of confusion, disappointment and fear, mainly because of the persistent guidance and encouragement from Jiake.

I would like to thank to my co-supervisor, Dr Jennifer Tickner for giving me excellent supervision and mentorship. Without Jennifer, I would get loss in analysing the huge amount of microCT data. I have also benefited from exposure to valuable feedback and suggestions on a variety of projects made by Jennifer either during regular laboratory meetings or casual conversations. Language skills have been greatly improved by such long-term exposure.

I would also like to thank to another co-supervisor, Prof. Grant Morahan for providing a large amount of CC mice carcases as research subjects, and guidance on gene mapping using GeneMiner. Grant’s expertise in genetics is the cornerstone of this project.

I would like to acknowledge our collaborator Assoc. Prof Nathan Pavlos for providing insides on Snx27 gene and efforts on NHMRC grant application. Acknowledgement is also given to Dr Benjamin Mullin as our collaborator by associating those candidate genes found in this project with human datasets.

I would like to acknowledge the Australian Government for granting me the SIRF scholarship and UWA for the UIS and top-up scholarships to facilitate my postgraduate studies in Australia.

I would like to extend my appreciation to my fellow students: Baysie Lim, Dian Teguh, Gaurav Jadhav, Heng Qiu, Jacob Kenny, Kai Chen, Mingli Yang, Vincent Kuek, Chao Wang, Colin Wang, Zhou Lin, and many visiting scholars. Kai and Colin helped me a lot in cell cultures and performing RT-PCR.

Special thanks go to CMCA staff assisted me in using microCT to scan samples. The facility manager Dr Kirk Feindel and Ms Diana Engineer gave me grateful support and understanding for scanning more than 1,600 samples that would accelerate the ageing of radiation source of the instrument.

X
Special thanks go to Associate Professor Kristen Nowak who shared her experience of using GeneMiner for haplotype mapping, and commented on the interpretation of QTL mapping results.

Special thanks go to Dr Ramesh Ram who commented on the interpretation of QTL mapping results, and upgraded the threshold settings for the GeneMiner system upon my request.

Special thanks also go to Dr Jingshan Wu who has been a long term mentor academically and spiritually. Her encouragement in seeking both faith and career is unforgettable.

To my wife, Yushan Zhuang, who sacrificed her own career in China and moved to Australia, I deeply appreciate to her effort on carrying and raising two boys. Her unconditional love, just like the daily bread, is consistently pouring power of healing and giving strength to go through the everyday life even the valley of death. To my two little boy, Enoch and Ian, great thanks to their companion, in particular in those hard days. To my parents, sincerely thanks go to their support in any stages and aspects of life.

Last but not least, many thanks to all the brothers and sisters in the Perth Alliance Church, where became my home on the second day of my arrival in Perth. Many have given me either financial or spiritual support or both.

This thesis is dedicated to my wife Yushan and two boys Enoch and Ian, and my parents: you present in my life is such a gift from the God.
ABSTRACTS


XII
ARTICLES


ABBREVIATIONS

µCT  Micro-computed tomography
18S  18s ribosomal RNA
aCML Atypical chronic myeloid leukemia
ACTH Adrenocorticotropic hormone
ACVR2B Activin A receptor type IIB
ADAM A disintegrin and metalloprotease domain
ADAMTS8 ADAM metallopeptidase with thrombospondin type 1 motif 8
ADAMTS15 ADAM metallopeptidase with thrombospondin type 1 motif 15
ADAMTSL4 ADAMTS like 4
ALP Alkaline phosphatase
AML Acute myeloid leukaemia
ANO5 Anoctamin 5
AP-1 Activator protein-1
APOB Apolipoprotein B
APS Ammonium persulfate
ARNT Aryl hydrocarbon receptor nuclear translocator
ARP Actin-related protein
ASTN1 Astrotactin 1
ASTN2 Astrotactin 2
ATF4 Activating transcription factor 4
Baf Bafilomycin A1
BCA Bicinchoninic acid
BFR Bone formation rate
BGLAP Bone gamma-carboxyglutamic acid-containing protein (osteocalcin)
BHLH Basic-Helix-Loop-Helix
BMD Bone mineral density
BMM Bone marrow macrophage/monocyte
BMP Bone morphogenetic protein
BMU Basic multicellular units
BOC Brother of CDON
bp Base pairs
BRINP2 BMP/Retinoic Acid Inducible Neural Specific 2
BS Bone surface
BSA Bovine serum albumin
BSP Bone sialoprotein

XIV
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>BUA</td>
<td>Ultrasound attenuation</td>
</tr>
<tr>
<td>BV</td>
<td>Bone volume</td>
</tr>
<tr>
<td>BV/TV</td>
<td>Bone volume/total volume (bone volume fraction)</td>
</tr>
<tr>
<td>bZIP</td>
<td>Basic leucine zipper</td>
</tr>
<tr>
<td>C2</td>
<td>Conserved region 2</td>
</tr>
<tr>
<td>CC</td>
<td>Collaborative Cross</td>
</tr>
<tr>
<td>cDNA</td>
<td>Complementary Deoxyribonucleic acid</td>
</tr>
<tr>
<td>CED</td>
<td>Cranioectodermal dysplasia</td>
</tr>
<tr>
<td>CEP135</td>
<td>Centrosomal protein 135</td>
</tr>
<tr>
<td>CFU</td>
<td>Colony-forming unit</td>
</tr>
<tr>
<td>CGD-II</td>
<td>Congenital disorder of glycosylation Type II</td>
</tr>
<tr>
<td>Chr</td>
<td>Chromosome</td>
</tr>
<tr>
<td>cKO</td>
<td>Conditional-knockout</td>
</tr>
<tr>
<td>CLOCK</td>
<td>Clock circadian regulator</td>
</tr>
<tr>
<td>cm</td>
<td>Centimeter</td>
</tr>
<tr>
<td>COL27A1</td>
<td>Collagen type XXVII alpha 1 chain</td>
</tr>
<tr>
<td>COL4A1</td>
<td>Collagen type IV alpha 1 chain</td>
</tr>
<tr>
<td>COL4A2</td>
<td>Collagen type IV alpha 2 chain</td>
</tr>
<tr>
<td>Conn.D</td>
<td>Connectivity density</td>
</tr>
<tr>
<td>CREB</td>
<td>cAMP-response element-binding protein</td>
</tr>
<tr>
<td>CSF1</td>
<td>colony stimulating factor 1</td>
</tr>
<tr>
<td>Ct.Ar</td>
<td>Cortical bone area</td>
</tr>
<tr>
<td>Ct.Ar/Tt.Ar</td>
<td>Cortical area fraction</td>
</tr>
<tr>
<td>CTSB</td>
<td>Cathepsin B</td>
</tr>
<tr>
<td>CTSK</td>
<td>Cathepsin K</td>
</tr>
<tr>
<td>CTSS</td>
<td>Cathepsin S</td>
</tr>
<tr>
<td>Ct.Th</td>
<td>Cortical thickness</td>
</tr>
<tr>
<td>CZ</td>
<td>Clear zone</td>
</tr>
<tr>
<td>DA</td>
<td>Degree of anisotropy</td>
</tr>
<tr>
<td>DC-STAMP</td>
<td>Dendritic cell-specific transmembrane protein</td>
</tr>
<tr>
<td>DHD</td>
<td>Dachshund homology domain</td>
</tr>
<tr>
<td>DHH</td>
<td>Desert hedgehog</td>
</tr>
<tr>
<td>DKK1</td>
<td>Dickkopf WNT signalling pathway inhibitor-1</td>
</tr>
<tr>
<td>DKK4</td>
<td>Dickkopf WNT signalling pathway inhibitor-4</td>
</tr>
<tr>
<td>DKKL1</td>
<td>Dickkopf-like protein 1</td>
</tr>
<tr>
<td>DMC</td>
<td>Dyggve-Melchior-Clausen</td>
</tr>
<tr>
<td>DXA</td>
<td>Dual-energy X-ray absorptiometry</td>
</tr>
</tbody>
</table>

XV
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>HH</td>
<td>Hedgehog</td>
</tr>
<tr>
<td>HIF</td>
<td>Hypoxia Inducible Factor</td>
</tr>
<tr>
<td>HLH</td>
<td>Helix-loop-helix</td>
</tr>
<tr>
<td>HPRT</td>
<td>Hypoxanthine-guanine phosphoribosyltransferase</td>
</tr>
<tr>
<td>HMBS</td>
<td>Hydroxymethylbilane synthase</td>
</tr>
<tr>
<td>HOX</td>
<td>Homeobox domain</td>
</tr>
<tr>
<td>HIF</td>
<td>Hypoxanthine guanine phosphoribosyl transferase</td>
</tr>
<tr>
<td>HSC</td>
<td>Haematopoietic stem cell</td>
</tr>
<tr>
<td>HSPG</td>
<td>Heparin sulfate proteoglycan</td>
</tr>
<tr>
<td>IBSP</td>
<td>Integrin-binding sialoprotein</td>
</tr>
<tr>
<td>IGF</td>
<td>Insulin growth factor</td>
</tr>
<tr>
<td>IGFBP</td>
<td>Insulin like growth factor binding protein</td>
</tr>
<tr>
<td>IGFBP7</td>
<td>Insulin like growth factor binding protein 7</td>
</tr>
<tr>
<td>IHH</td>
<td>Indian hedgehog</td>
</tr>
<tr>
<td>IKK</td>
<td>Inhibitor of nuclear factor kappa-B kinase</td>
</tr>
<tr>
<td>IRF-8</td>
<td>Interferon regulatory factor-8</td>
</tr>
<tr>
<td>I-Smad</td>
<td>Inhibitory Smad</td>
</tr>
<tr>
<td>ITAM</td>
<td>Immunoreceptor tyrosine-based activation motif</td>
</tr>
<tr>
<td>ITGA10</td>
<td>Integrin Subunit Alpha 10</td>
</tr>
<tr>
<td>JAG1</td>
<td>Jagged 1</td>
</tr>
<tr>
<td>JNK</td>
<td>c-Jun N-terminal kinase</td>
</tr>
<tr>
<td>kb</td>
<td>Kilo-base</td>
</tr>
<tr>
<td>KCNJ</td>
<td>Potassium voltage-gated channel subfamily J</td>
</tr>
<tr>
<td>kDa</td>
<td>Kilo-dalton</td>
</tr>
<tr>
<td>KDR</td>
<td>Kinase insert domain receptor</td>
</tr>
<tr>
<td>kg</td>
<td>Kilogram</td>
</tr>
<tr>
<td>KIT</td>
<td>KIT proto-oncogene receptor tyrosine kinase</td>
</tr>
<tr>
<td>LAMA1</td>
<td>Laminin subunit alpha 1</td>
</tr>
<tr>
<td>LCS</td>
<td>Lacunocanalicilar system</td>
</tr>
<tr>
<td>LDL-C</td>
<td>Low-density lipoprotein cholesterol</td>
</tr>
<tr>
<td>LEF1</td>
<td>Lymphoid enhancer-binding factor-1</td>
</tr>
<tr>
<td>LGMD2L</td>
<td>Limb girdle muscular dystrophy type 2L</td>
</tr>
<tr>
<td>LGR4</td>
<td>Leucine-rich-repeat-containing G-protein-coupled receptor 4</td>
</tr>
<tr>
<td>LH2</td>
<td>Lipooxygenase homology 2</td>
</tr>
<tr>
<td>LOXHD1</td>
<td>Lipooxygenase homology domains 1</td>
</tr>
<tr>
<td>LRP4</td>
<td>Low-density lipoprotein receptor-related protein-4</td>
</tr>
<tr>
<td>LRP5</td>
<td>Low-density lipoprotein receptor-related protein-5</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Name</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>LRP6</td>
<td>Low-density lipoprotein receptor-related protein-6</td>
</tr>
<tr>
<td>M</td>
<td>Molar</td>
</tr>
<tr>
<td>MAF</td>
<td>Minor allele frequency</td>
</tr>
<tr>
<td>MAP2K6</td>
<td>Mitogen-activated protein kinase kinase 6</td>
</tr>
<tr>
<td>MAP3K14</td>
<td>Mitogen-activated protein kinase kinase 14</td>
</tr>
<tr>
<td>MAPK</td>
<td>Mitogen-activated protein kinase</td>
</tr>
<tr>
<td>MAPK4</td>
<td>Mitogen-activated protein kinase 4</td>
</tr>
<tr>
<td>MATN3</td>
<td>Matrilin-3</td>
</tr>
<tr>
<td>MBD</td>
<td>Methyl-CpG-binding domain protein</td>
</tr>
<tr>
<td>MBD1</td>
<td>Methyl-CpG-binding domain protein 1</td>
</tr>
<tr>
<td>M-CSF</td>
<td>Macrophage colony-stimulating factor</td>
</tr>
<tr>
<td>MITF</td>
<td>Microphthalmia-associated transcription factor</td>
</tr>
<tr>
<td>MECOM</td>
<td>MDS1 and EVI1 complex locus</td>
</tr>
<tr>
<td>mg</td>
<td>Milligram</td>
</tr>
<tr>
<td>mm</td>
<td>Millimeter</td>
</tr>
<tr>
<td>mM</td>
<td>Millimolar</td>
</tr>
<tr>
<td>MMP</td>
<td>Matrix metalloproteinase</td>
</tr>
<tr>
<td>mRNA</td>
<td>Messenger RNA</td>
</tr>
<tr>
<td>MSC</td>
<td>Mesenchymal stem cells</td>
</tr>
<tr>
<td>NBF</td>
<td>Neutral buffered formalin</td>
</tr>
<tr>
<td>NFATC1</td>
<td>Nuclear factor of activated T cells cytoplasmic 1</td>
</tr>
<tr>
<td>NF-κB</td>
<td>Nuclear factor kappa-B</td>
</tr>
<tr>
<td>NGS</td>
<td>Next-generation sequencing</td>
</tr>
<tr>
<td>NICD</td>
<td>Notch intracellular domain</td>
</tr>
<tr>
<td>NIK</td>
<td>NF-κB-inducing kinase</td>
</tr>
<tr>
<td>NMU</td>
<td>Neuromedin U</td>
</tr>
<tr>
<td>NOC</td>
<td>Nocturnin</td>
</tr>
<tr>
<td>OASIS</td>
<td>cAMP responsive element binding protein 3-like 1</td>
</tr>
<tr>
<td>OCN</td>
<td>Osteocalcin</td>
</tr>
<tr>
<td>OPG</td>
<td>Osteoprotegerin</td>
</tr>
<tr>
<td>OPN</td>
<td>Osteopontin</td>
</tr>
<tr>
<td>OPPG</td>
<td>Osteoporosis-pseudoglioma syndrome</td>
</tr>
<tr>
<td>OSCAR</td>
<td>Osteoclast associated receptor</td>
</tr>
<tr>
<td>OSF-1</td>
<td>Osteoblast stimulating factor-1</td>
</tr>
<tr>
<td>OSTM1</td>
<td>Osteopetrosis-associated transmembrane protein 1</td>
</tr>
<tr>
<td>OSX</td>
<td>Osterix</td>
</tr>
<tr>
<td>OVX</td>
<td>Ovariectomised</td>
</tr>
</tbody>
</table>

XVIII
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>PAPPA2</td>
<td>Pregnancy-associated plasma protein A2</td>
</tr>
<tr>
<td>PBS</td>
<td>Phosphate buffered saline</td>
</tr>
<tr>
<td>PC</td>
<td>Proprotein convertase</td>
</tr>
<tr>
<td>PCP</td>
<td>Planar cell polarity pathway</td>
</tr>
<tr>
<td>PDGF</td>
<td>Platelet-derived growth factor</td>
</tr>
<tr>
<td>PDIA3</td>
<td>Protein disulfide isomerase family A member 3</td>
</tr>
<tr>
<td>PER</td>
<td>Period circadian clock 1</td>
</tr>
<tr>
<td>PI3K</td>
<td>Phosphoinositide 3-kinase</td>
</tr>
<tr>
<td>PKC</td>
<td>Protein kinase C</td>
</tr>
<tr>
<td>PLCγ</td>
<td>Phospholipase Cγ</td>
</tr>
<tr>
<td>POSTN</td>
<td>Periostin</td>
</tr>
<tr>
<td>PPAR-γ</td>
<td>Peroxisome proliferator-activated receptor gamma</td>
</tr>
<tr>
<td>PPI</td>
<td>Protein-protein interaction</td>
</tr>
<tr>
<td>PPR</td>
<td>PTH/PTH-related protein receptor</td>
</tr>
<tr>
<td>PRUNE</td>
<td>Prune exopolyphosphatase</td>
</tr>
<tr>
<td>PS1</td>
<td>Presenilin-1</td>
</tr>
<tr>
<td>PTCH1</td>
<td>Patched homologue-1</td>
</tr>
<tr>
<td>PTH</td>
<td>Parathyroid hormone</td>
</tr>
<tr>
<td>PTHrP</td>
<td>PTH related peptide</td>
</tr>
<tr>
<td>PTM</td>
<td>Post-translational modification</td>
</tr>
<tr>
<td>QTL</td>
<td>Quantitative trait locus</td>
</tr>
<tr>
<td>PU.1</td>
<td>Purine-rich binding protein 1</td>
</tr>
<tr>
<td>RA</td>
<td>Rheumatoid arthritis</td>
</tr>
<tr>
<td>RANK</td>
<td>Receptor activator of nuclear factor kappa B</td>
</tr>
<tr>
<td>RANKL</td>
<td>Receptor activator of nuclear factor kappa B ligand</td>
</tr>
<tr>
<td>RB</td>
<td>Ruffled border membrane</td>
</tr>
<tr>
<td>REST</td>
<td>RE1 silencing transcription factor</td>
</tr>
<tr>
<td>RHOB</td>
<td>Ras homolog family member B</td>
</tr>
<tr>
<td>RI</td>
<td>Recombinant inbred</td>
</tr>
<tr>
<td>RNA</td>
<td>Ribonucleic acid</td>
</tr>
<tr>
<td>ROR2</td>
<td>Receptor tyrosine kinase-like orphan receptor 2</td>
</tr>
<tr>
<td>rpm</td>
<td>Revolutions per minute</td>
</tr>
<tr>
<td>RT-PCR</td>
<td>Reverse transcription-polymerase chain reaction</td>
</tr>
<tr>
<td>RUNX2</td>
<td>Runt-related transcription factor 2</td>
</tr>
<tr>
<td>SAA</td>
<td>Serum amyloid A</td>
</tr>
<tr>
<td>SAA1</td>
<td>Serum amyloid A1</td>
</tr>
<tr>
<td>SALL3</td>
<td>Spalt Like Transcription Factor 3</td>
</tr>
</tbody>
</table>
SATB 2  Special AT-rich sequence binding protein 2
SEM  Standard error of the mean
SETBP1  SET Binding Protein 1
SETDB1  SET domain bifurcated 1
SHH  Sonic hedgehog
SKOR2  SKI family transcriptional corepressor 2
SMART  Simple modular architecture research tool
SMC  Smith-McCort
SMI  Structural model index
SMURF  SMAD ubiquitin regulatory factor
SMURF2  SMAD ubiquitin regulatory factor 2
SNP  Single nucleotide polymorphism
SNX  Sorting nexin
SNX19  Sorting nexin 19
SNX27  Sorting nexin 27
SOST  Sclerostin
SOX9  Sex determining region-Y-box 9
SPP1  Secreted phosphoprotein 1 (osteopontin)
SREBP  Sterol regulatory element-binding protein
STAT  Signal transducer and activator of transcription
STRING  Search tool for the retrieval of interacting genes/proteins
SZ  Sealing zone
TACE  TNFα Converting Enzyme
TAE  Tris-acetate EDTA
Tak1  Transforming growth factor-β-activated kinase-1
TAZ  Tafazzin
Tb.N  Trabecular number
Tb.Sp  Trabecular separation
Tb.Th  Trabecular thickness
TBS  Tris-buffered saline
TCF1  T cell factor-1
TEMED  Tetramethylethylenediamine
TGFβ  Transforming growth factor-beta
TGN  Trans-golgi network
TIMP  Tissue inhibitors of metalloproteases
TMEM165  Transmembrane protein 165
TNF  Tumour necrosis factor

XX
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>TNFSF13B</td>
<td>TNF superfamily member 13b</td>
</tr>
<tr>
<td>TNF-α</td>
<td>Tumor necrosis factor-alpha</td>
</tr>
<tr>
<td>TRACP</td>
<td>Tartrate resistant acid phosphatase</td>
</tr>
<tr>
<td>TRAF</td>
<td>TNF receptor-associated factor</td>
</tr>
<tr>
<td>TRAF4</td>
<td>TNF receptor associated factor 4</td>
</tr>
<tr>
<td>TREM2</td>
<td>Triggering receptor expression in myeloid cells-2</td>
</tr>
<tr>
<td>TRIP-1</td>
<td>TGF-β receptor interacting protein</td>
</tr>
<tr>
<td>Trx-1</td>
<td>Thioredoxin-1</td>
</tr>
<tr>
<td>Tt.Ar</td>
<td>Total cortical area</td>
</tr>
<tr>
<td>TV</td>
<td>Tissue volume</td>
</tr>
<tr>
<td>TXNIP</td>
<td>Thioredoxin interacting protein</td>
</tr>
<tr>
<td>v-ATPase</td>
<td>Vacuolar ATPase</td>
</tr>
<tr>
<td>VEGF</td>
<td>Vascular endothelial growth factor</td>
</tr>
<tr>
<td>VKD</td>
<td>Vitamin K-dependent</td>
</tr>
<tr>
<td>VOS</td>
<td>Velocity of sound</td>
</tr>
<tr>
<td>VWFA</td>
<td>Von Willebrand factor A</td>
</tr>
<tr>
<td>WDR35</td>
<td>WD repeat domain 35</td>
</tr>
<tr>
<td>WGS</td>
<td>Whole genome sequencing</td>
</tr>
<tr>
<td>WNT3</td>
<td>Wnt family member 3</td>
</tr>
<tr>
<td>WNT9B</td>
<td>Wnt family member 9B</td>
</tr>
<tr>
<td>WRC</td>
<td>Wnt-responding cell</td>
</tr>
<tr>
<td>β-NGF</td>
<td>β-nerve growth factor</td>
</tr>
<tr>
<td>μg</td>
<td>Microgram</td>
</tr>
<tr>
<td>μl</td>
<td>Microlitre</td>
</tr>
<tr>
<td>μm</td>
<td>Micrometre</td>
</tr>
<tr>
<td>μM</td>
<td>Micromolar</td>
</tr>
</tbody>
</table>
LIST OF FIGURES AND TABLES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Osteoblast differentiation from mesenchymal stem cells (MSCs).</td>
<td>6</td>
</tr>
<tr>
<td>1.2</td>
<td>Osteoclastogenesis.</td>
<td>16</td>
</tr>
<tr>
<td>1.3</td>
<td>Osteoclast-osteoblast crosstalk and bone remodelling.</td>
<td>21</td>
</tr>
<tr>
<td>1.4</td>
<td>RANK/RANKL/OPG signalling pathway.</td>
<td>23</td>
</tr>
<tr>
<td>1.5</td>
<td>Wnt signalling pathway.</td>
<td>25</td>
</tr>
<tr>
<td>1.6</td>
<td>Schematic diagram of Endochondral Ossification.</td>
<td>28</td>
</tr>
<tr>
<td>1.7</td>
<td>TGF-β/BMP signalling Pathway.</td>
<td>31</td>
</tr>
<tr>
<td>2.1</td>
<td>Genetic predisposition and architecture in osteoporosis.</td>
<td>43</td>
</tr>
<tr>
<td>2.2</td>
<td>Workflow of GWAS.</td>
<td>54</td>
</tr>
<tr>
<td>2.3</td>
<td>Important Genes Link to BMD Variations.</td>
<td>57</td>
</tr>
<tr>
<td>2.4</td>
<td>The Collaborative Cross mice (recombinant inbred) breeding scheme.</td>
<td>62</td>
</tr>
<tr>
<td>2.5</td>
<td>Workflow of the Collaborative Cross Mice Screening.</td>
<td>76</td>
</tr>
<tr>
<td>5.1</td>
<td>Trabecular bone volume/tissue volume (BV/TV) (%) in distal femur varies between strains of CC mice and between genders.</td>
<td>90</td>
</tr>
<tr>
<td>5.2</td>
<td>Trabecular number (Tb. N) (1/mm) in distal femur varies between strains of CC mice and between genders.</td>
<td>92</td>
</tr>
<tr>
<td>5.3</td>
<td>Trabecular thickness (Tb. Th) (mm) in distal femur varies between strains of CC mice and between genders.</td>
<td>93</td>
</tr>
<tr>
<td>5.4</td>
<td>Changes in trabecular thickness (Tb. Th) (mm) in distal femur in CC mice in comparison with stratified reference.</td>
<td>94</td>
</tr>
<tr>
<td>5.5</td>
<td>Changes in connectivity density (Conn. D) (1/mm3) in distal femur in CC mice in comparison with C57BL/6J.</td>
<td>95</td>
</tr>
<tr>
<td>5.6</td>
<td>Changes in degree of anisotropy (DA) (α) in distal femur in CC mice in comparison with stratified reference.</td>
<td>96</td>
</tr>
<tr>
<td>5.7</td>
<td>Changes in structural model index (SMI) in distal femur in CC mice in comparison with C57BL/6J.</td>
<td>97</td>
</tr>
<tr>
<td>5.8</td>
<td>Positive correlation between BV/TV and Tb. N, Tb. Th and Conn. D in female.</td>
<td>99</td>
</tr>
<tr>
<td>5.9</td>
<td>Negative correlation between BV/TV and Tb. Sp, SMI and DA in female.</td>
<td>100</td>
</tr>
</tbody>
</table>
Figure 5.10  Positive correlation between BV/TV and Tb. N, Tb. Th and Conn. D in male.

Figure 5.11  Negative correlation between BV/TV and Tb. N, Tb. Th and Conn. D in male.

Figure 5.12  Change of BV/TV over time in female CC mice.

Figure 5.13  Change of BV/TV over time in male CC mice.

Figure 5.14  Changes in cortical thickness (Ct.Th mm) in distal femur in CC mice in comparison with C57BL/6J.

Figure 5.15  Correlation between BV/TV and Ct. Th in CC mice.

Figure 5.16  Change of Ct. Th over time in female CC mice.

Figure 5.17  Change of Ct. Th over time in male CC mice.

Figure 6.1  QTL haplotype mapping of BV/TV in female CC mice.

Figure 6.2  QTL haplotype mapping of BV/TV in three separated female age groups.

Figure 6.3  QTL haplotype mapping of BV/TV in male CC mice.

Figure 6.4  QTL haplotype mapping of BV/TV in old male CC mice.

Figure 6.5  QTL haplotype mapping of osteopetrosis in female CC mice.

Figure 6.6  QTL haplotype mapping of Tb. N in female CC mice.

Figure 6.7  QTL haplotype mapping of Tb. N in male CC mice.

Figure 6.8  QTL haplotype mapping of Tb. Th in female CC mice.

Figure 6.9  QTL haplotype mapping of Tb. Th in middle aged female CC mice.

Figure 6.10  QTL haplotype mapping of Tb. Th in male CC mice.

Figure 6.11  QTL haplotype mapping of Tb. Th in middle aged male CC mice (chr17).

Figure 6.12  QTL haplotype mapping of Tb. Sp in female CC mice.

Figure 6.13  QTL haplotype mapping of Tb. Sp in male CC.

Figure 6.14  QTL haplotype mapping of Tb. Sp in middle aged male CC mice.

Figure 6.15  QLT haplotype mapping of Conn. D in female CC mice.

Figure 6.16  QTL haplotype mapping of Conn. D in male CC mice (chr9).

Figure 6.17  QTL haplotype mapping of Conn. D in male CC.

Figure 6.18  QLT haplotype mapping of SMI in female CC mice.

Figure 6.19  QLT haplotype mapping of SMI in male CC mice.
Figure 6.20  QTL haplotype mapping of SMI in middle aged male CC mice (chr3).

Figure 6.21  QTL haplotype mapping of DA in female CC mice (chr8, peak1).

Figure 6.22  QTL haplotype mapping of DA in female CC mice (excluded CIV2, CAMERON & PUB).

Figure 6.23  QTL haplotype mapping of DA in female CC mice (excluded CIV2, CAMERON, PUB, ROGAN & REV).

Figure 6.24  QTL haplotype mapping of DA in female CC mice (excluded CIV2, CAMERON, PUB and ROGAN).

Figure 6.25  QTL haplotype mapping of DA in young female CC mice (chr18).

Figure 6.26  QTL haplotype mapping of DA in middle aged female CC mice.

Figure 6.27  QTL haplotype mapping of DA in male CC mice.

Figure 6.28  QTL haplotype mapping of DA in young male CC mice.

Figure 6.29  QTL haplotype mapping of DA in middle aged male CC mice (chr4).

Figure 6.30  QTL haplotype mapping of Ct. Th in female CC mice.

Figure 6.31  QTL haplotype mapping of Ct. Th in male CC mice.

Figure 6.32  Summary of identified loci from haplotype mapping of various traits in CC mice.

Table 7.1  Candidate genes achieved significant association with bone phenotypes.

Figure 7.1  Gene expression of murine Setbp1 in osteoclasts and osteoblasts.

Figure 7.2  Visualisation of SETBP1 protein network in various species, generated by STRING programme.

Figure 7.3  Visualisation of TNFRSF13B protein network in various species.

Figure 7.4  Gene expression profile of mouse Tnfrsf13b.

Figure 7.5  Gene expression profile of mouse DYM.

Figure 7.6  Visualisation of DYM protein network in various species.

Figure 7.7  Gene expression profile of mouse Hic1.

Figure 7.8  Visualisation of HIC1 protein network in various species.

Figure 7.9  Gene expression profile of mouse Zbtb7c.

Figure 7.10  Visualisation of ZBTB7C protein network in various species.

Figure 7.11  Gene expression profile of mouse En1.

Figure 7.12  Visualisation of EN1 protein network in various species.
Figure 7.13 Visualisation of WNT3 protein network in various species. 191
Figure 7.14 Visualisation of WNT9B protein network in various species. 192
Table 7.2 Candidate genes achieved nominally significant associations with bone phenotypes. 195
Figure 7.15 Visualisation of SALL3 protein network in various species. 199
Figure 7.16 PPI network for candidate genes identified from CC mice haplotype mapping using STRING programme. 210
Figure 7.17 Protein network of osteoporosis susceptibility genes in homo sapiens (human), built with STRING. 211
Figure 7.18 Summary of OP susceptibility candidate Genes from CC screening. 214
Supplementary Table S2.1 Genes associated with osteoporosis. 298
Supplemental Table S4.1 Summary of CC mice. 314
Supplemental Table S4.2 Cohorts included in BMD association testing. 317
Supplemental Table S4.3 Cohorts included in BUA and VOS association testing. 318
Supplementary Table S5.1 C57BL/6J reference populations. 319
Supplementary Figure S5.1 Trabecular traits in young female CC mice. 320
Supplementary Figure S5.2 Trabecular traits in middle aged female CC mice. 321
Supplementary Figure S5.3 Trabecular traits in old female CC mice. 322
Supplementary Figure S5.4 Traits in young male CC mice. 323
Supplementary Figure S5.5 Traits in middle aged male CC mice. 324
Supplementary Figure S5.6 Traits in old male CC mice. 325
Supplementary Table S5.2
Ranking female CC strains based on BV/TV (% of C57BL/6).

Supplementary Table S5.3
Ranking male CC strains based on BV/TV (% of C57BL/6).

Supplementary Figure S5.7
Positive correlation between BV/TV and Tb. N, Tb. Th and Conn. in young female CC mice.

Supplementary Figure S5.8
Negative correlation between BV/TV and Tb. Sp, SMI and DA in young female CC mice.

Supplementary Figure S5.9
Positive correlation between BV/TV and Tb. N, Tb. Th and Conn. in middle aged female CC mice.

Supplementary Figure S5.10
Negative correlation between BV/TV and Tb. Sp, SMI and DA in middle aged female CC mice.

Supplementary Figure S5.11
Positive correlation between BV/TV and Tb. N, Tb. Th and Conn. in old female CC mice.

Supplementary Figure S5.12
Negative correlation between BV/TV and Tb. Sp, SMI and DA in old female CC mice.

Supplementary Figure S5.13
Positive correlation between BV/TV and Tb. N, Tb. Th and Conn. in young male CC mice.

Supplementary Figure S5.14
Negative correlation between BV/TV and Tb. Sp, SMI and DA in young male CC mice.

Supplementary Figure S5.15
Positive correlation between BV/TV and Tb. N, Tb. Th and Conn. in middle aged male CC mice.

Supplementary Figure S5.16
Negative correlation between BV/TV and Tb. Sp, SMI and DA in middle aged male CC mice.

Supplementary Figure S5.17
Positive correlation between BV/TV and Tb. N, Tb. Th and Conn. in old male CC mice.

Supplementary Figure S5.18
Negative correlation between BV/TV and Tb. Sp, SMI and DA in old male CC mice.

Supplementary Figure S5.19
Ct. Th in female CC mice of various age groups.

Supplementary Figure S5.20
Ct. Th in male CC mice of various age groups.

Supplementary Figure S5.21
Correlation between BV/TV and Ct. Th in female CC mice of various ages.
<table>
<thead>
<tr>
<th>Supplementary Table S7.5</th>
<th>Visualisation of HS1BP3 protein network in various species.</th>
<th>386</th>
</tr>
</thead>
<tbody>
<tr>
<td>Supplementary Table S7.6</td>
<td>Gene expression profile of mouse Cd33.</td>
<td>387</td>
</tr>
<tr>
<td>Supplementary Table S7.7</td>
<td>Visualisation of SMURF2 protein network in various species.</td>
<td>388</td>
</tr>
<tr>
<td>Supplementary Table S7.8</td>
<td>Visualisation of COL27A1 protein network in various species.</td>
<td>389</td>
</tr>
<tr>
<td>Supplementary Table S7.9</td>
<td>Extended protein network of OP susceptibility genes in Mus musculus (mouse).</td>
<td>390</td>
</tr>
<tr>
<td>Supplementary Table S7.10</td>
<td>Extended protein network of OP susceptibility genes in Rattus norvegicus (rat).</td>
<td>391</td>
</tr>
<tr>
<td>Supplementary Table S7.11</td>
<td>Interaction network for canonical Wnt Signalling Pathway.</td>
<td>392</td>
</tr>
<tr>
<td>Supplementary Table S7.12</td>
<td>Interaction network for non-canonical Wnt signalling pathway.</td>
<td>393</td>
</tr>
</tbody>
</table>
Chapter 1
Bone Biology
1.1 Bone Morphology and Organisation

1.1.1 Bone Composition and Organisation

Bone tissue is a type of mineralised connective tissue. It is composed of an organic matrix, minerals and water. Approximately 90% of the organic composition of bone is type I collagen (Young, Kerr, Ibaraki, Heegaard, & Robey, 1992). The rest is made up of noncollagenous proteins which play a crucial role in regulating collagen formation, cell attachment, $\text{Ca}^{2+}$ and hydroxyapatite binding, and mineralisation of bone tissue and microfracture resistance (Young et al., 1992). Bone mineral is composed of poorly crystalline carbonated apatite mineral. Together with collagen fibrils, bone mineral forms the extracellular matrix (ECM) of the bone that provides the structural and biochemical support to the surrounding cells (Gentili & Cancedda, 2009).

Bone can be divided into two types: cortical bone (also called compact or solid bone) and trabecular bone (also called cancellous or spongy bone) (Vashishth, Gibson, Kimura, Schaffler, & Fyhrie, 2002). Cortical bone is made up of fundamental structural and functional units called osteons or Haversian systems (B. Clarke, 2008). Each osteon is composed of concentric layers of hard inorganic matrix (called lamellae) surrounding a central canal, Haversian canal, which contains blood vessels and nerves. Cortical bone is located in the diaphyseal regions of long bones and as the cortex or thin outer shell of most bones. Trabecular bone is a porous structure composed of a lattice of rods and plates, and is primarily present at the end of long bones and in the vertebrae, skull and pelvis. It is highly vascular and frequently filled with bone marrow (B. Clarke, 2008).

Bone tissue is composed of and maintained by four types of cells: osteoblasts, osteocytes, bone lining cells and osteoclasts (Buckwalter, Glimcher, Cooper, & Recker, 1996; Downey & Siegel, 2006). Osteoblasts are derived from mesenchymal stem cells (MSCs) which have the potential to differentiate into mature osteoblasts that form new bone matrix, and these can further differentiate to become osteocytes or cells lining the bone surface (Dennis et al., 1999; Pittenger et al., 1999). They play a role in the initial formation of the skeleton as well as in continuous bone growth and remodelling. Osteocytes are the most abundant cells in mature bone, mainly involved in the maintenance of bone structure and exchange of nutrients and waste between the blood and bone tissues (Dallas, Prideaux, & Bonewald, 2013; Knothe Tate, Adamson, Tami, & Bauer, 2004). Bone lining cells are derived from osteoblasts and cover the surfaces of the bone. The function of lining cells is unclear but they appear to play an important role in bone remodelling (Everts et al., 2002). Osteoclasts are the principal cells responsible for bone resorption and pathological bone loss (Matsuo & Irie, 2008). They are specialised monocyte/macrophage family members that differentiate from hematopoietic precursors (Teitelbaum, 2000a). Terminal differentiation in this lineage is characterised
by the expression of osteoclast marker genes, such as calcitonin receptor and tartrate resistant acid phosphatase (TRAP). Moreover, mature osteoclasts are capable of forming resorption lacunae on bone.

Bone is continuously shaped and changed through remodelling or bone turnover, which relies on the balance between bone formation and resorption (Kular, Tickner, Chim, & Xu, 2012). Bone remodelling is a complex process that occurs due to coordinated actions of the four types of bone cells, which together form the temporary anatomical structure called basic multicellular unit (BMU) (Andersen et al., 2009; Parfitt, 2002; Sims & Martin, 2014). Physiological remodelling is essential for fracture healing, skeletal structural adaptations to mechanical usage and maintenance of calcium homeostasis (Dallas et al., 2013). On the other hand, pathological remodelling characterised by imbalance of bone formation and resorption may result in several bone diseases. For instance, excessive resorption by osteoclasts results in osteoporosis, whereas overwhelming bone formation contributes to osteopetrosis (Sobacchi, Schulz, Coxon, Villa, & Helfrich, 2013; Teitelbaum, 2000b). Overall, the balance between bone formation and resorption is crucial in the achievement and maintenance of healthy bone tissue.

1.1.2 The Functions of Bone

Bone has been classically regarded as a multifunctional organ that plays roles in mechanical support and protection of the internal organs, hematopoiesis, and mineral homeostasis (H. K. Datta, Ng, Walker, Tuck, & Varanasi, 2008; Kular et al., 2012; Parfitt, 2002). More recently, its role as an endocrine organ that help to regulate phosphate and energy metabolism has been recognised (Fukumoto & Martin, 2009; Guntur & Rosen, 2012).

The mechanical functions of bone are predominantly load bearing. Cortical bone takes most of load, and trabecular bone redirects stresses to the stronger cortical shell. Bone also provides a framework for muscles and tendons to attach, promoting movement of the body. Bone also provides protective functions for vital internal organs such as the heart, lungs and brain. Bone can be considered as a hematopoietic (blood-forming) organ due to its bone marrow found in almost any trabecular bone. Bone marrow has been widely recognised as the source of hematopoiesis, producing erythrocytes, lymphocytes and myelocytes. During growth and development, bone marrow appears red indicating active hematopoiesis; during ageing it turns yellow, showing its main composition is fat, which does not have a hematopoiesis function. Abnormal changes in bone marrow architecture such as aplastic anaemia and multiple myeloma leads to disrupted hematopoiesis.
Bone is an important reservoir of bone minerals including calcium and phosphate. The skeleton contains 99% of calcium and 88% of total phosphate. Calcium not only provides bone with its stiffness and strength, but also is essential for many physiological process, exerting a profound effect on muscle contraction, cellular motility, neuronal transmission, cell growth and proliferation (Berridge, Lipp, & Bootman, 2000). Phosphate is another major component of bone and other tissues. It plays a crucial role in the regulation of metabolic processes through phosphorylation (adding phosphate) and dephosphorylation (removing phosphate) (Fischer, 2010).

Bone has been identified as an endocrine organ that is involved in modulating mineral metabolism and metabolic homeostasis by secreting two hormones, fibroblast growth factor 23 (FGF-23), and osteocalcin (Karsenty & Ferron, 2012; Karsenty & Oury, 2012; Shimizu, Fukumoto, & Fujita, 2012). Most of FGF-23, an important factor that regulates serum phosphate and parathyroid hormone (PTH), is synthesised and secreted by osteocytes (Shimizu et al., 2012). Osteocalcin is secreted specifically by osteoblasts and is thought to play a role in metabolic regulation and osteoblast differentiation (Karsenty & Ferron, 2012; N. K. Lee et al., 2007). When undercarboxylated, osteocalcin also helps regulate pancreatic β-cell proliferation and enhance insulin secretion, and increase insulin sensitivity by driving adipocytes to release adiponectin, an important hormone regulating glucose homeostasis as well as fatty acid breakdown (Karsenty & Ferron, 2012; N. K. Lee et al., 2007). In addition, this hormone acts on the Leydig cells of the testis via the pancreas-bone-testis axis to stimulate testosterone production and therefore regulate male fertility (Oury et al., 2011). Therefore bone, through its several hormones, helps regulate mineral and energy metabolism, and male fertility.

1.2 Bone Cells

1.2.1 Osteoblasts

1.2.1.1 Osteoblast Morphology and Function

Osteoblasts are cuboidal mononuclear cells accounting for 4–6% of total resident cells in the bone, the main function of these cells is to synthesise bone matrix (Capulli, Paone, & Rucci, 2014). Osteoblasts initiate the process of bone matrix synthesis by secreting abundant type I collagen and other specific matrix proteins, which form osteoid, followed by subsequent deposit of mineral in the form of hydroxyapatite (Buckwalter et al., 1996; Florencio-Silva, Sasso, Sasso-Cerri, Simoes, & Cerri, 2015). Osteoblasts also secrete high levels of alkaline phosphatase (ALP) and osteocalcin. ALP degrades phosphate-containing compounds and releases phosphate ions inside the matrix vesicles. The hydroxyapatite crystals then form with calcium ion inside the vesicles (Florencio-Silva et al., 2015). Osteocalcin is thought to play a role in metabolic regulation and osteoblast differentiation.
(Guntur & Rosen, 2012). The circulating concentrations of ALP and osteocalcin reflect the rate of bone formation.

Osteoblasts play an essential role in bone modelling, which is a process by which bone is formed by osteoblasts without prior bone resorption (Kular et al., 2012; Parfitt, 2002; Sims & Martin, 2014). This process begins during the foetal stage and stops once longitudinal growth of the skeleton is completed after adolescence, resulting in vigorous changes in bone size and shape during early years of life. Osteoblasts also actively participate in bone remodelling through life, mainly coupled with osteoclasts (P. A. Hill, 1998; Matsuo & Irie, 2008; Sims & Martin, 2014). During this process, bone is first resorbed by osteoclasts, and then formed in the same location by osteoblasts.

### 1.2.1.2 Osteoblast Formation and Differentiation

Osteoblasts are derived from MSCs, which have the potential to differentiate into osteoblasts, chondrocytes, fibroblasts and adipocytes (Ducy, Schinke, & Karsenty, 2000; Pittenger et al., 1999). MSCs were originally identified in the bone marrow and later isolated in many other tissues such as muscle (Jiang et al., 2002), adipose tissue (Gimble & Guilak, 2003) and umbilical cord blood (Karahuseyinoglu, Kocaefe, Balci, Erdemli, & Can, 2008). Osteoblast differentiation initiates with the commitment of osteoprogenitor cells from pluripotential MSCs, these then undergo differentiation into immature osteoblasts, and then mature bone building osteoblasts, and finally ends with formation of osteocytes within the bone matrix, bone lining cells on the bone surface, or programmed cell death (apoptosis) for a fraction of osteoblasts (Harada & Rodan, 2003; Marie, 2008) (Figure 1.1). This process is tightly control by several transcription factors, including Runx2, Osterix, β-Catenin, and Atf4.

### 1.2.1.3 Regulation of Osteoblast Differentiation

Osteoblast differentiation requires the involvement of several transcription factors either in stimulatory or inhibitory ways. The stimulatory transcription factors that upregulate osteoblast differentiation include Runx2, Osterix, β-Catenin, AP1 (activator protein-1), Atf4 (activating transcription factor 4) and Satb2 (special AT-rich sequence binding protein 2) (Fakhry, Hamade, Badran, Buchet, & Magne, 2013). Inhibitory transcription factors include PPAR-γ (peroxisome proliferator-activated receptor gamma), Twist-1 and Twist-2, which downregulate expression of osteoblast genes (Jensen, Gopalakrishnan, & Westendorf, 2010).

Runx2 (also known as Cbfa1), a member of the Runx family, is considered as the master osteoblast transcription factor (Fakhry et al., 2013). It is expressed by MSCs during skeletal development and osteoblast differentiation (Komori, 2005; C. A. Yoshida & Komori, 2005). This transcription factor
Figure 1.1 Osteoblast differentiation from mesenchymal stem cells (MSCs). MSCs can differentiate into osteoblasts, chondrocytes, fibroblasts, adipocytes, and myocytes. BMP, osterix, RUNX2 and PTH are essential factors that drive the mesenchymal osteoblastic lineage. Once the osteoblast has differentiated and completed its cycle of matrix synthesis, it can become a flattened lining cell on the bone surface, be embedded in the bone as an osteocyte, or undergo apoptosis. (Figure modified from Raisz, 1999).
is essential for the commitment of undifferentiated MSCs into the osteoblast lineage and inhibits adipocyte differentiation (Komori, 2005; Y. Zhang, Khan, Delling, & Tobiasch, 2012). Runx2 induces the expression of bone matrix proteins while keeping the osteoblasts at an immature stage (Ducy et al., 1999; Komori, 2005). Mice with global deletion of Runx2 lack osteoblasts, leading to the absent expression of bone matrix protein genes including Spp1 (secreted phosphoprotein 1, also named osteopontin), Ibsp (integrin-binding sialoprotein) and Bglap (bone gamma-carboxyglutamic acid-containing protein, also called osteocalcin) (Adhami et al., 2015; Adhami, Rashid, Chen, & Javed, 2014; Komori et al., 1997; Kormori, 2006). These knockout mice normally die shortly after birth, and their skeletons completely lack bone. In transgenic mice overexpressing Runx2 under the control of 2.3kb Col1a1 promoter osteopenia with multiple fractures was observed as a result of decreased mineralisation and reduced numbers of mature osteoblasts, suggesting that Runx2 negatively controls osteoblast terminal differentiation and maintains osteoblastic cells at an immature stage (Geoffroy, Kneissel, Fournier, Boyde, & Matthias, 2002; Kanatani et al., 2006; W. Liu et al., 2001). Cortical bone of these mice presents a woven bone-like structure, with a severe reduction in cortical but not trabecular bone mass. This is caused by the increased recruitment and activity of osteoclasts in the immature cortical bone, which contains abundant Arg-Gly-Asp (RGD) motifs that promote osteoclast attachment (Geoffroy et al., 2002; W. Liu et al., 2001). In contrast, the extent of mineralisation in the trabecular bone is higher in these mice than in C57BL/6J controls. In more recent studies, these transgenic mice show impaired postnatal bone formation without disrupted embryonic skeletogenesis, implying a crucial role of Runx2 in committed osteoblasts during early postnatal bone formation and remodelling (Adhami et al., 2015; Adhami et al., 2014). Therefore, Runx2 acts differently to control osteoblast differentiation and bone formation; it drives MSCs to the osteoblast lineage to generate an increase in immature osteoblasts, induces the expression of bone matrix proteins genes, whilst inhibiting the formation of fully mature osteoblasts.

Osterix (Osx, also known as Sp7), one of the Sp-family of Krüppel-like zinc finger proteins, is another master transcription factor regulating osteoblast differentiation and function (J. Lu et al., 2016; W. Tang, Li, Osimiri, & Zhang, 2011). It is expressed by osteoblasts of endochondral and membranous bones, and expressed slightly in prehypertrophic chondrocytes, acting as a downstream gene of Runx2 (Baek, de Crombrugghe, & Kim, 2010; K. Nakashima et al., 2002; C. Zhang, 2012). Osx is initially expressed during progression of MSCs into the osteoblast lineage and intensifies during osteoblast differentiation (K. Nakashima et al., 2002). Osx global knockout mice show normal cartilage development but completely defective endochondral and intramembranous bone formation (K. Nakashima et al., 2002). In these mutants, although the expression of Runx2 is normal, preosteoblasts are arrested in their differentiation to become osteoblasts, and mice die within 15 minutes after birth.
due to difficulty in breathing. Conditional Osx knockout mice, generated under the control of 2.3kb Collal-Cre, show osteopenia in adult bone during growth even though they are viable (Baek et al., 2010). In these mice, both bone formation rate (BFR) and mineral apposition rate (MAR) are significantly diminished, suggesting that Osx is essential in regulating osteoblast differentiation and bone formation during early bone development. In human populations, genetic polymorphisms in the Osx locus are associated with low bone mineral density (BMD), indicating an important role for this gene in bone modelling and remodelling (Styrkarsdottir et al., 2008; Timpson et al., 2009).

β-Catenin was first discovered as a part of the cell-to-cell adherens junctions with cadherin and α-catenin, regulating the actin cytoskeleton, and intracellular signalling pathways that control gene transcription (Ozawa, Baribault, & Kemler, 1989; L. Shapiro & Weis, 2009). It was later identified as a key transcriptional coactivator for the expression of Wnt responsive genes (Takemaru & Moon, 2000). More specifically, β-Catenin is a potent transcription factor for osteoblast differentiation (Mbalaviele et al., 2005). It is an essential part of the canonical Wnt pathway (also called Wnt/β-catenin pathway), causing accumulation of unphosphorylated β-catenin molecules in the cytoplasm and translocation of these molecules into the nucleus, and further activating the transcription of downstream genes by binding TCF/LEF transcription factors (Pinzone et al., 2009). Without Wnt signalling, β-catenin would not accumulate in the cytoplasm because it would be degraded by a multiprotein “destruction complex” (Stamos & Weis, 2013). β-Catenin, working with other transcription factors in this pathway, prevents osteoblasts from differentiating into chondrocytes (Day, Guo, Garrett-Beal, & Yang, 2005; T. P. Hill, Spater, Taketo, Birchmeier, & Hartmann, 2005). Its role in chondrocyte maturation and osteoblastogenesis was first revealed in chickens (Hartmann & Tabin, 2000), and later in mice (Day et al., 2005; T. P. Hill et al., 2005). Conditional deletion of the Ctnnb1 gene that encodes β-catenin in Prx1-Cre transgenic mice affects membranous bone formation, suggesting that β-catenin is an essential factor for osteoblast differentiation (T. P. Hill et al., 2005).

AP-1 is not a single protein, rather represents a group of heterodimeric transcription factors composed of various basic leucine zipper domain containing proteins including Jun (c-Jun, JunB, JunD), Fos (c-Fos, Fra-1, Fra-2, FosB) and ATF families (Karin, 1996). It has been linked to a number of cellular processes including transformation, proliferation, differentiation and apoptosis (Ameyar, Wisniewska, & Weitzman, 2003). Specifically, AP-1 acts on osteoblast differentiation and proliferation through its stimulation by TGF-β, PTH, and 1,25-dihydroxy vitamin D3 (Harada & Rodan, 2003). The various members of the AP-1 complex are expressed in different ways during osteoblast differentiation and maturation in vitro (McCabe et al., 1996; E. F. Wagner, 2002). During osteoblast proliferation, all Fos and Jun proteins are highly expressed, followed by a marked decline during the period of ECM production and mineralisation (McCabe et al., 1996). In fully differentiated
osteoblasts, however, Fra-2 and JunD become the major components of the AP-1 complex (McCabe et al., 1996). In c-Fos transgenic mice, uncontrolled osteoblast differentiation results in production of malignant osteoid and osteosarcoma (Grigoriadis, Schellander, Wang, & Wagner, 1993). On the contrary, mice with c-Fos deficiency develop osteopetrosis, a bone disease with excessive bone mass, due to an early block in the differentiation of bone-resorbing osteoclasts (Grigoriadis et al., 1994). Transgenic mice overexpressing Fra-1 in osteoblasts develop a progressive increase in bone mass, developing osteosclerosis of the entire skeleton (Jochum et al., 2000). This phenotype is cell autonomous and is probably caused by enhanced osteoblast differentiation and activity. Moreover, in vitro, Fra-1 transgenic osteoblasts increase differentiation but not proliferation, suggesting that Fra-1 specifically enhances bone formation but not development of osteosarcoma. Mice with knockout of Fra-1 or Fra-2, conversely, develop osteopenia (Bozec et al., 2010; Eferl et al., 2004). Additionally, transgenic mice overexpressing ΔFosB, a splice variant of FosB, develop severe osteosclerosis due to a reduction of inhibitory Smad6 expression (Sabatakos et al., 2008; Sabatakos et al., 2000).

ATF4 is a basic leucine-zipper transcription factor of the ATF/CREB protein family. ATF4 promotes osteoblast differentiation through at least three distinct molecular mechanisms. First, ATF4, through interaction with Runx2, directly activates the expression of osteocalcin, a marker for terminally differentiated osteoblasts (X. Yang & Karsenty, 2004; X. Yang et al., 2004). Atf4 knockout mice have a dramatically reduced bone mass and bone formation rate that persists throughout life, with markedly reduced expression of osteocalcin, suggesting a critical role for ATF4 in osteoblast differentiation (G. Xiao et al., 2005; X. Yang & Karsenty, 2004). ATF4 is also thought to be critical for proliferation and survival of calvarial osteoblasts and primary mouse bone marrow stromal cells (BMSCs) both in vitro and in vivo (X. Zhang et al., 2008). Second, ATF4 acts as an upstream transcriptional activator of the Osx gene and mediates the anabolic actions of PTH in osteoblast differentiation and bone formation in Atf4-deficient mice (S. Yu et al., 2009; S. Yu et al., 2008). Finally, a new role of ATF4 in promotion of osteoblast differentiation via up-regulation of β-catenin protein levels in MSCs and osteoblasts/osteoprogenitors has recently been described (S. Yu et al., 2013).

SATB2 is expressed in branchial arches and osteoblast-lineage cells, and is responsible for preventing craniofacial defects and osteoblast malfunction (J. Zhang et al., 2011). Its gene locates on 2q32-q33, in which a locus responsible for cleft palate was identified (FitzPatrick et al., 2003). SATB2 plays an essential role in facial patterning and bone development (Rosenfeld et al., 2009). In Satb2 knockout mice, in addition to the craniofacial patterning defects, osteoblast differentiation and function are defective, leading to delayed bone formation and mineralisation (Dobreva et al., 2006). Satb2 knockout mice have very similar osteoblast phenotype with Atf4-deficient mice, suggesting that these
two proteins may interact with each other. Evidence reveals that Satb2 interacts not only with ATF4 but also with Runx2 (Dobreva et al., 2006; Gong, Qian, Yang, Wang, & Yu, 2014).

Smads are signal transduction molecules, playing a critical role in the intracellular pathway that transmits TGF-β signals from the cell surface into the nucleus (Massague, Seoane, & Wotton, 2005; Wrana, 2000). In the TGF-β pathway, Smads are directly activated by the membrane receptors of TGF-β and form transcriptional complexes to control target genes (Massague et al., 2005). Smads 1, 5, and 8 work as substrates for the bone morphogenetic protein (BMP) and anti-Muellerian receptors, and Smads 2 and 3 for the TGF-β and activating/nodal receptors (Miyazawa, Shinozaki, Hara, Furuya, & Miyazono, 2002; Rahman, Akhtar, Jamil, Banik, & Asaduzzaman, 2015). Smads 2 and 5, cooperating with Runx2, induce osteoblast differentiation in C2C12 pluripotent mesenchymal precursor cells (K. S. Lee et al., 2000). Without these two BMP-activated Smads, Runx2 alone does not induce osteoblast differentiation (Javed et al., 2009).

**PPAR-γ** (also called PPARG) is a member of the nuclear hormone receptor superfamily of ligand-activated transcription factors that plays an essential role in adipocyte differentiation, insulin signalling, and regulating systemic energy homeostasis (Tontonoz & Spiegelman, 2008). It also controls bone remodelling and regulates bone cell differentiation of both mesenchymal and hematopoietic lineages (Wan, 2010). In the osteoblast lineage, PPAR-γ supresses multiple signalling pathways (Wnt, TGF-β/BMP and IGF-1) and stimulatory transcription factors such as Runx2 and Osterix regulating osteoblast differentiation (Shockley et al., 2009). Recently, PPAR-γ was identified to work with nocturnin (NOC), a circadian protein, in the cytoplasm to enhance PPAR-γ activity (Kawai et al., 2010). *Noc* knockout mice develop reduced numbers of adipocytes and increased bone mass (Kawai et al., 2010). In addition, *Noc* overexpression in cells of the osteoblast lineage enhances adipogenesis and supresses osteoblastogenesis (Kawai et al., 2010). Therefore, PPAR-γ acts as a key negative regulator of osteoblast differentiation in addition to its main role as a crucial positive regulator of adipocyte differentiation.

**Twist-1** (previously called Twist) and **Twist-2** (previously called Dermo-1) are basic helix-loop-helix (HLH) transcription factors that are implicated in cell lineage determination and differentiation (T. D. Howard et al., 1997). **Twist-1** knockout mice die at embryonic day 11.5 due to failure of neural tube closure (Z. F. Chen & Behringer, 1995). **Twist-2** knockout mice are viable, but die a few days after birth due to craniosynostosis, a disease caused by premature osteoblast differentiation in the skull (Sosic, Richardson, Yu, Ornitz, & Olson, 2003). Evidence reveals that premature osteoblast differentiation and bone formation occur in both knockout mice mainly through downregulation of Runx2 (Bialek et al., 2004; Kronenberg, 2004; M. S. Lee, Lowe, Strong, Wergedal, & Glackin, 1999).
Furthermore, overexpression of Twist-1 suppresses osteoblast differentiation without affecting Runx2 expression. Its anti-osteogenic function is mediated by the Twist box, which interacts with the Runx2 DNA binding domain to inhibit its function (Bialek et al., 2004).

1.2.2 Osteocytes

1.2.2.1 Osteocyte Morphology and Functions

Osteocytes are the most abundant cells in mature bone, accounting for over 90% of all bone cells, with there being approximately 10 times more osteocytes than osteoblasts. Osteocytes also have a long lifespan, with an average half-life of 25 years (Bonewald, 2004, 2011). Mature osteocytes are stellate shaped or dendritic cells enclosed within the lacunocanalicular network of bone (Sugawara, Kamioka, Honjo, Tezuka, & Takano-Yamamoto, 2005). The size of cell body varies between different species and even within a single species, from 5-20 μm in diameter, with a cell to cell distance between 20 - 30 μm (Mullender, van der Meer, Huiskes, & Lips, 1996; Sugawara et al., 2005). An osteocyte typically contains a single nucleus and cell processes that radiate towards the mineralising matrix. The typical stellate morphology of the osteocyte is thought to be an intrinsic characteristic of terminal osteocyte differentiation, in addition to the need to build a cellular network (Burger, Klein-Nulend, van der Plas, & Nijweide, 1995). In mature bone, the osteocyte body and its processes are located within lacunae that consist of a number of oblong spaces between the lamellae, and canaliculi, which are microscopic canals between the lacunae of ossified bone. Osteocytes, along with their cytoplasmic processes within the mineralised bone matrix, form the lacunocanalicular system (LCS), a large network intercommunicating the lacunae and canaliculi (Bozal, Sanchez, Mandalunis, & Ubios, 2013; Bozal, Sanchez, & Ubios, 2012; Burger et al., 1995; Knothe Tate et al., 2004). The LCS is a functional syncytium, working as a conduit for metabolic traffic and exchange (Bozal et al., 2012).

Osteocytes have long been considered quiescent bystander cells compared to the osteoblasts and osteoclasts that actively perform bone formation and resorption (Schaffler & Kennedy, 2012). Analysis of osteocyte properties and functions has long been hampered because they are embedded in a mineralised osteoid matrix and difficult to isolate (Bonewald, 2011). Our current understanding of osteocyte biology has been dominated by histomorphometric studies of fixed specimens. Isolation and culture of osteocytes from mice has been technically difficult until recent success in generating protocols to yield an outgrowth of osteocyte-like cells from mature mice (A. R. Stern & Bonewald, 2015; A. R. Stern et al., 2012). The role of osteocytes in skeletal homeostasis is gaining interest, as changes in the osteocyte environment cause the release of ECM proteins, growth factors and
cytokines affecting both osteoblasts and osteoclasts (Aarden, Burger, & Nijweide, 1994; Bellido, 2014; R. B. Martin, 2000; C. H. Turner & Pavalko, 1998). ECM proteins such as osteopontin and dentin matrix protein 1 (DMP1) bind calcium and inhibit mineralisation, acting as regulators of mineral homeostasis (Addison, Masica, Gray, & McKee, 2010; Gericke et al., 2010). Osteocytes also express hormones including insulin-like growth factor 1 (IGF-1), and osteocyte Igf1 conditional knockout mice present with severely impaired longitudinal and periosteal bone growth and thinner calvarial bone (Sheng, Zhou, Bonewald, Baylink, & Lau, 2013). Therefore, osteocyte-derived IGF-1 plays an essential role in regulating bone turnover during developmental growth of intramembranous bone. Recent evidence also shows that osteocytes play a pivotal role in mechanotransduction as the primary mechanosensing cells, therefore they are assumed to be involved in the initiation of bone remodelling (Bonewald, 2011; Feng & McDonald, 2011; Schaffler, Cheung, Majeska, & Kennedy, 2014; Schaffler & Kennedy, 2012). Osteocytes may recognise and respond to mechanical stresses induced by fluid flow in the LCS (Bonewald, 2011; You et al., 2000). A much higher expression of RANKL has been observed in purified osteocytes than in osteoblasts and bone marrow stromal cells, indicating a greater capacity of osteocytes in supporting osteoclastogenesis in vitro (T. Nakashima et al., 2011b). The severe osteopetrotic phenotype observed in Rankl conditional knockout mice in osteocytes also suggests that osteocytes are the predominant producers of RANKL postnatally in vivo. Osteocyte apoptosis may indirectly induce osteoclastogenesis by stimulating stromal/osteoblastic cells to secrete RANKL, therefore initiating the process of bone remodelling (Bellido, 2014). Along with osteoblasts and osteoclasts, osteocytes are also thought to have a role in determining the ECM volume in bone (Noble & Reeve, 2000; Vashishth et al., 2002; Weinstein & Manolagas, 2000).

1.2.2.2 Formation and Apoptosis of Osteocytes

It is well recognised that osteocytes are derived from multipotent MSCs through osteoblast differentiation. The process of osteocyte formation is referred to as osteocytogenesis. The precise mechanisms by which an osteoblast embeds in bone matrix and transform into an osteocyte are still not fully understood. Once a matrix-producing osteoblast stops forming new bone matrix, it may either become an osteocyte, a lining cell, or undergo apoptosis (Figure 1.1) (Bonewald, 2011; Franz-Odendaal, Hall, & Witten, 2006; Manolagas, 2000; L.G. Raisz, 1999). Embedding into the bone matrix drives osteoblasts to transform into osteocytes, and then cease their activity as active bone forming cells (Aarden et al., 1994; Bonewald, 2011). Like osteoblasts and osteoclasts, osteocytes are living cells. Cross-talk among the three main cell types in bone has been documented, mainly either between osteocytes and osteoblasts or osteoblasts and osteoclasts (Burger, Klein-Nulend, & Smit, 2003; Civitelli, 2008; Matsuo & Irie, 2008). Therefore, a functional network of bone cells that
includes osteoclasts, osteoblasts and osteocytes is important for maintaining the integrity and function of bone.

Apoptosis of osteocytes has attracted attention because it may have a major role in bone resorption. Osteocyte apoptosis occurs as a result of immobilisation, withdrawal of oestrogen, glucocorticoid treatment, osteoporosis and osteoarthritis (Dodd, Raleigh, & Gross, 1999; Emerton et al., 2010; Weinstein, Nicholas, & Manolagas, 2000). This process increases bone fragility, which is considered to be closely related to loss of the ability to sense microdamage and/or signal repair (Manolagas, 2000; Noble & Reeve, 2000). Oxygen deprivation elevates hypoxia-induced factor alpha (HIF-α), resulting in apoptosis and induction of the osteoclastogenic factor TNF-α and VEGF (Gross et al., 2001). Loss of oestrogen also induces osteocyte apoptosis (Emerton et al., 2010). In mice receiving bilateral ovariectomy (OVX) increased osteocyte apoptosis was shown in a non-uniform distribution throughout the femoral diaphyses, compared with mice that underwent sham surgery (Emerton et al., 2010). Furthermore, glucocorticoid-induced apoptosis of the cells of the osteoblast lineage could be expected because osteoblasts are highly sensitive to glucocorticoid (Weinstein, Jilka, Parfitt, & Manolagas, 1998; Weinstein et al., 2000). In addition to undergoing apoptosis, osteocytes can undergo a process of self-preservation called autophagy, especially in response to glucocorticoid (Xia et al., 2010).

1.2.2.3 Regulation of Osteocytes

The function of the osteocytes is to maintain the functional matrix, by interacting with osteoclasts and osteoblasts (Capulli et al., 2014; Klein-Nulend, Bacabac, & Mullender, 2005; Thompson, Epari, Bieler, & Duda, 2010). The formation and regulation of osteocytes is regulated by a number of factors that are expressed by preosteoblasts, osteoblasts and osteocytes. Preosteoblasts express many bone-specific markers, including osteonectin, alkaline phosphatase, PTH/PTHrP receptor and integrins (Franz-Odendaal et al., 2006). Preosteoblasts differentiate into active bone matrix-secreting osteoblasts under the control of these markers. During the transition phase, many of the previously expressed bone markers are downregulated or switched off in the osteocyte. Osteocytes may stimulate osteoblast recruitment and differentiation by expressing osteoblast stimulating factor-1 (OSF-1) (Tezuka et al., 1990). The recruited osteoblasts further differentiate into osteocytes after being surrounded by the osteoid matrix that they build, and become another source of OSF-1. Osteocalcin is produced by the late osteoblast and continues to be expressed by the osteocyte (Dallas et al., 2013).

The E11 antibody is used to distinguish mature osteocytes and preosteocytes (Hadjijargyrou, Rightmire, Ando, & Lombardo, 2001; F. Liu, Malaval, & Aubin, 1997). The E11/gp38 is uniquely expressed in osteocytes, not osteoblasts, even though it is also expressed in other tissues (K. Zhang
et al., 2006). It is selectively expressed in the earliest transitional phase, reflecting the activity of osteocyte differentiation. DMP1, Phex, fibroblast growth factor 23 (FGF23) and MEPE, while expressed at low levels in osteoblasts and other tissues, are also found to be highly expressed in osteocytes (Miyagawa et al., 2014; K. Zhang et al., 2006). FGF23, a circulating factor that regulates the renal reabsorption of phosphate and metabolism of vitamin D, is mainly produced by osteocytes (Ubaidus et al., 2009). Furthermore, FGF23 is intensively localised in the endosteal area of cortical bone, which has a high density LCS, suggesting it has a role in bone remodelling.

Sclerostin, a protein encoded by the SOST gene, is highly expressed in osteocytes and directly inhibits bone formation (Poole et al., 2005). Unlike E11/gp38, which is considered as a marker for the early osteocytes, sclerostin is a specific marker for late osteocytes. Both in vitro and in vivo studies of SOST/sclerostin expression are difficult because of the technical difficulties associated with isolation of osteocytes from mammalian bones. Recently, Stern et al. successfully isolated osteocytes from skeletally mature mice through a process of extended collagenase digestions combined with EDTA-based decalcification opening the way for more extensive investigation of osteocyte biology (Stern & Bonewald, 2015; Stern et al., 2012). Inactivating mutations in the SOST gene result in a rare, recessively inherited high bone mass disorder called sclerosteosis (Balemans et al., 2001; X. Li et al., 2008). Based on these data, sclerostin has emerged as a key negative regulator of bone mass, which could be used to combat osteoporosis (X. Li et al., 2009; X. Li et al., 2010).

### 1.2.3 Osteoclasts

#### 1.2.3.1 Osteoclast Morphology and Function

Osteoclasts are the principal bone resorbing cells essential for bone remodelling and calcium homeostasis (Teitelbaum, 2000a). They are multinucleated giant cells, typically having five or more nuclei and are around 150-200 µm in diameter (Jain & Weinstein, 2009). In bone, osteoclasts locate on endosteal surfaces within the Haversian system and on the periosteal surface beneath the periosteum (Z. Li, Kong, & Qi, 2006).

Osteoclasts, existing in the motile and the absorbing phases, are the most important cells in the whole cascade of cellular events needed for bone resorption, called the resorption cycle (Vaanen & Horton, 1995). This process requires a sequence of osteoclast activities. First, motile osteoclasts, which are flattened and non-polarised cells, migrate to the absorption site (Mulari, Vaaraniemi, & Vaananen, 2003). Second, these osteoclasts tightly attach to the bone surface and form four distinct membrane domains including a ruffled border, a sealing zone, a functional secretory domain and a basolateral membrane (Lakkakorpi & Vaananen, 1996; Vaananen, Zhao, Mulari, & Halleen, 2000).
Each of these domains has a unique function. The sealing zone isolates the resorption area from its surroundings. The ruffled border is the resorbing organelle that separates the acidic resorptive environment from the rest of the cells. The basal membrane domain remains outside the sealing zone, facing the bone marrow, and extracellular fluid. The functional secretory domain locates in the center of the basolateral membrane and it removes the degraded bone matrix. Following attachment and establishment of membrane domains, osteoclasts gain resorbing function. They become highly polarised with basolateral and apical membranes, driven by reorganisation of the actin cytoskeleton. The bone matrix is then degraded by acidification and secretion of proteolytic enzymes into the resorption lacunae. Finally, the degraded bone fragments are removed from the resorption lacunae and released into the extracellular space (J. Xu, Cheng, Feng, Pavlos, & Zheng, 2007).

1.2.3.2 Osteoclast Differentiation and Fusion

Osteoclasts are derived from the fusion of mononuclear progenitors of the monocyte-macrophage family in a process termed osteoclastogenesis (Boyle, Simonet, & Lacey, 2003; Teitelbaum, 2000a, 2000b; Yavropoulou & Yovos, 2008). Osteoclastogenesis is a complex process that includes many steps, as shown in Figure 1.2. Circulating monocytic precursor cells can differentiate into macrophages, macrophage giant cells, dendritic cells or osteoclasts depending on the niche of cell differentiation (Udagawa et al., 1990). At the initiation phase of bone remodelling, haematopoietic precursors in the circulation or bone marrow are recruited to the BMU where they undergo osteoclastogenesis. Osteoclast precursors proliferate to form committed preosteoclasts. These committed mononucleated preosteoclasts differentiate and fuse to form multinucleated osteoclasts.

1.2.3.3 Regulation of Osteoclast Differentiation and Survival

Osteoclast differentiation requires the involvement of at least three signalling pathways, which are activated by macrophage colony-stimulating factor (M-CSF), Receptor Activator of NF-κB (RANK) and its ligand, RANKL, and their related transcriptional factors or cytokines (Boyce & Xing, 2008; Boyle et al., 2003; Teitelbaum, 2000a). Among them, M-CSF and RANKL are most important and are thought to be sufficient to induce osteoclastogenesis (Boyle et al., 2003). M-CSF, also called colony stimulating factor 1 (CSF1), is a secreted cytokine that drives haematopoietic stem cells (HSCs) to differentiate into monocytes (Stanley et al., 1997). It is released by osteoblasts and stromal cells following stimulation by PTH, exerting its function on osteoclasts in a paracrine manner.

Osteoclast differentiation is severely suppressed in M-CSF-mutated osteopetrotic op/op mice because these mutants have a serious deficiency of M-CSF (Boyle et al., 2003; Wiktor-Jedrzejczak et al., 1990). Administration of soluble M-CSF to these mutants rescues their osteopetrosis, suggesting the
Figure 1.2 Osteoclastogenesis. Osteoclasts differentiate from haematopoietic precursor cells under the influence of RANKL and M-CSF (CSF-1). These cytokines are both essential to induce expression of genes that commit precursors to the osteoclast lineage, including TRAcP, cathepsin K, calcitonin receptor and β3-integrin. OPG acts as a decoy receptor to block RANKL binding to its cellular receptor RANK and subsequently inhibit osteoclast differentiation (Figure modified from Boyle et al, 2003).
important role of M-CSF in osteoclast biology. In the monocyte/macrophage lineage, M-CSF provides signals necessary for survival and proliferation of early osteoclast precursors through binding to its sole receptor, c-Fms (Hamilton, 1997; Kular et al., 2012; Tanaka et al., 1993). The functional association between M-CSF and c-Fms is established by the evidence that mice lacking Csf1r, the gene coding for c-Fms, develop the same osteopetrotic phenotype as the op/op mutants (X. M. Dai et al., 2002). Binding of M-CSF to c-Fms activates extracellular signal-regulated kinase (ERK) through Grb2 and phosphoinositide 3-kinase (PI3K)/Akt (Pixley & Stanley, 2004; Ross, 2006). The differentiation process of HSCs to osteoclast precursors is induced by transcription factors, such as purine-rich binding protein 1 (PU.1), c-Fos and Mitf. PU.1 stimulates the expression of CSF1R, which further induces expression of RANK by upregulating c-Fos (Anderson et al., 1999; Dekoter, Walsh, & Singh, 1998; Kwon, Lee, Lee, Paik, & Lee, 2005). Knockout mice lacking either c-Fos or PU.1 exhibit an osteopetrotic phenotype, confirming the crucial role of these two transcription factors in osteoclastogenesis (Grigoriadis et al., 1994; Jacenko, 1995).

RANKL, a member of the tumour necrosis factor (TNF) superfamily, is expressed by osteoblasts, stromal cells and other cell lines (Boyle et al., 2003; T. Nakashima et al., 2011a). It is part of the RANK/RANKL/Osteoprotegerin (OPG) signalling pathway which is essential for commitment of the osteoclast precursors to mature osteoclasts. RANKL binds to its receptor RANK, which is present on osteoclasts and their precursors, and stimulates the M-CSF-expanded precursors to commit to the osteoclast phenotype (Lacey et al., 1998). These cells then express key osteoclast markers such as calcitonin receptor and tartrate resistant acid phosphatase (TRAcP) (Boyle et al., 2003). Upon further stimulation with M-CSF and RANKL the preosteoclasts fuse to form multinucleated cells. These cells begin to express more specific osteoclast markers such as CTSK. Although these cells are mature osteoclasts they are not functional as they lack a ruffled border, which is essential for osteoclast bone resorption. Activation of RANK by its ligand results in polarisation of the osteoclast cytoskeleton and formation of the ruffled border (Burgess et al., 1999). Osteoclast differentiation is absent in Rankl-deficient mice, which present a severe osteopetrotic phenotype including increased trabecular bone mass owing to the complete absence of osteoclasts (Boyle et al., 2003; Kong et al., 1999). This osteopetrotic phenotype is caused by the failure of RANKL expression to stimulate osteoclastogenesis (Kong et al., 1999). When Rankl deletion was induced in the osteoblast lineage in 26 week old mice, no detectable change in trabecular bone mass was found, suggesting that the contribution of early osteoblast-derived RANKL to trabecular osteoclastogenesis may be most important in trabecular bone during growth (Xiong et al., 2011). In another study, Rankl knockout mice showed a reduction in cortical thickness, suggesting that osteocytic RANKL may be of importance in endocortical resorption of disuse-associated bone loss (Kartsogiannis et al., 1999).
Rankl-deficient mice also have impaired T and B lymphocyte development and present a complete absence of lymph nodes.

1.3 Bone Modelling and Remodelling

1.3.1 Skeletal Development and Fracture Healing

Skeletal development starts from the first trimester of gestation and continues for years after birth. This process presents into two distinct formats: intramembranous (direct) ossification and endochondral (indirect) ossification (K. W. Ng, Romas, Donnan, & Findlay, 1997). Intramembranous ossification is a process of differentiation of MSCs into bone and is responsible for the formation of most of the skull or craniofacial skeleton. It is most often related to embryonic development, in addition to bone healing (Marsell & Einhorn, 2011; Marsh & Li, 1999; K. W. Ng et al., 1997). Endochondral ossification is a process of MSCs differentiation into a hyaline cartilage template and gradual replacement of this template with mineralised bone tissue (Ortega, Behonick, & Werb, 2004). This process is responsible for the formation of most of the vertebrate appendicular and axial skeleton. Like intramembranous ossification, this process is not limited to embryonic development, but also to fracture healing. The major difference between these two processes is that cartilage is present during endochondral ossification.

During intramembranous ossification in the skull, neural crest-derived mesenchymal cells proliferate and concentrate to produce osteoblasts, which deposit osteoid matrix. These osteoblasts become arrayed along the calcified region of the matrix. Osteoblasts that are trapped within the bone matrix become osteocytes. As calcification proceeds, bony spicules radiate out from the region where ossification began. The entire region of calcified spicules becomes surrounded by compact mesenchymal cells that form the periosteum. The cells on the inner surface of the periosteum also become osteoblasts and deposit osteoid matrix parallel to that of the existing spicules. In this manner, many layers of bone are formed (Marsell & Einhorn, 2011; Marsh & Li, 1999; K. W. Ng et al., 1997).

Endochondral ossification occurs at two distinct sites of ossification in the long bones and vertebrates. One is primary (diaphyseal), and the other is secondary (epiphyseal). Bone development initiates at the primary site, followed by forming a growth plate between the diaphysis and epiphysis. During this process, MSCs differentiate into chondrocytes and form the cartilage template. Vascularisation initiates around these templates, and osteoblasts differentiate around the central area in the bone collar. Chondrocytes in this central area differentiate into hypertrophic chondrocytes and allow vascular invasion. After complete differentiation, they die and ECM remodelling is carried out by osteoclasts and chondroclasts recruited together with the blood vessels. The growth plates are
ultimately responsible for the elongation of the long bones. Later, endochondral ossification occurs at a secondary site, in the epiphysis of the long bones.

Bone modelling is defined as either the formation of bone by osteoblasts or resorption of bone by osteoclasts on a given surface (Migliaccio, Falcone, & Spera, 2004). It can be divided into two types: formation modelling that is induced by osteoblasts, and resorptive modelling that is driven by osteoclasts. The primary function of bone modelling is to alter bone mass and bone shape, or remove damage and therefore maintain bone strength. Bone modelling is an uncoupled process which is characterised by having independent controlled anabolic or catabolic sites. It is an important factor of skeletal growth, functioning as a lifelong optimisation process for the adaptation of bone mass and architecture to meet the functional needs (Bagge, 2000; Jacobs, 2000).

1.3.2 Bone Remodelling

Unlike bone modelling, bone remodelling is a complex process requiring “coupling” between osteoclastic and osteoblastic activities on a given surface. Coordinated function of osteoclasts and osteoblasts ensures that resorption lacunae are filled with new bone produced by osteoblasts so as to maintain bone integrity. Mechanical forces and certain systemic factors, such as steroids or PTH, affect bone remodelling. PTH plays a role in the regulation of the RANKL/OPG ratio by inducing a transient wave of RANKL expression in osteoblasts, but can induce both catabolic and anabolic processes in bones depending on its concentration and release (L.G. Raisz, 1999; Samadfam, Xia, Miao, Hendy, & Goltzman, 2008). Local factors including TRAcP secreted by osteoclasts and IGF and TGF-β, which are released from the bone matrix during bone resorption stimulate bone formation (Hayden, Mohan, & Baylink, 1995; T. J. Martin, 2004; T. J. Martin & Sims, 2005; Mohan & Baylink, 1996).

Bone remodelling occurs within BMUs on the endocortical, trabecular, and intracortical components of the endosteal envelope (Frost, 1991; Parfitt, 1994, 1995). A fully developed BMU comprises of a group of osteoclasts in front forming the cutting cone, a group of osteoblasts behind forming the closing cone, associated connective tissue, blood vessels and nerves filling the cavity (Parfitt, 1994). The BMU maintains its size by excavating and refilling a tunnel through cortical bone or a trench across the surface of trabecular bone. Bone remodelling has been described as a ‘bone remodelling cycle’ consisting of four phases; activation, resorption, reversal and formation (P. A. Hill, 1998). However, in terms of osteoclast-osteoblast communication, it is recently thought to be more convenient to view it as three phases: initiation, transition and termination, as shown in Figure 1.3 (Matsuo & Irie, 2008; Tamma & Zallone, 2012). The initiation phase includes recruitment of osteoclast precursors, differentiation and activation of osteoclasts, and maintenance of bone
resorption. During this phase, initiation of osteoclastogenesis heavily relies on interaction between osteoclast precursors and cells in the osteoblast lineage. Bone lining cells express RANKL and stimulate RANK on osteoclast precursors, while osteoblasts produce M-CSF, which is essential to the survival of cells in the macrophage-osteoclast lineage (H. Yoshida et al., 1990). During the transition phase, osteoclasts stimulate differentiation of osteoblast precursors, activating bone formation in bone resorption lacunae. While bone formation is being activated, bone resorption terminates and osteoclasts undergo apoptosis. Bone formation proceeds gradually and lasts much longer than resorption. Osteocytes produce sclerostin, which suppresses osteoblastic bone formation. During the termination phase, osteoblasts become quiescent and osteoclasts are suppressed, mostly likely through OPG produced by osteoblasts.

1.3.3 Cell-to-Cell Communication

Intercellular communication between osteoclasts, osteoblasts and osteocytes is critically regulated to achieve bone homeostasis (Figure 1.3). Cell-to-cell communication between osteoclasts and osteoblasts is known to occur through at least three methods; direct contact, allowing membrane-bound ligands and receptors to interact and initiate intracellular signalling; via the formation of gap junctions allowing passage of small water-soluble molecules between the two cell types; and via paracrine factors, such as growth factors, cytokines, chemokines and other small molecules (Civitelli, 2008; Nakahama, 2010; Tamma & Zallone, 2012). The mechanisms of cell-to-cell communication between osteocytes and osteoblasts or osteoclasts, however, remain unclear, even though apoptosis in osteocytes is a critical stage in the initiation of resorption (Bellido, 2014; L. Deng et al., 2015; Jilka, Noble, & Weinstein, 2013).

1.4 Key Biological Pathways Regulating Bone Homeostasis

1.4.1 RANKL/RANK/OPG Pathway

RANKL binding to RANK activates an intracellular signalling cascade in osteoclast precursors resulting in activation of NF-κB, NFATc1, Akt/PKB, and JNK, ERK and p38 MAPK pathways (Figure 1.4) (W. Liu, Wang, Wei, Sun, & Feng, 2005; Wong et al., 1999). The RANK receptor lacks an intrinsic capacity to activate protein kinases to mediate signalling and therefore utilises interaction with adaptors or docking proteins, including TNF receptor-associated factors (TRAFs) 2, 3, 5, and 6, Gab2, and Cbl to activate downstream signalling (Darnay, Besse, Poblenz, Lamothe, & Jacoby, 2007; Darnay, Haridas, Ni, Moore, & Aggarwal, 1998; M. C. Walsh & Choi, 2014). Gab2 and Cbl are associated with RANK-mediated activation of c-Src, PI3 kinase (PI3K), and Akt, while TRAFs 2/5 and 6 can activate the TAB1/TAB2/TAK1 complex, resulting in activation of IKKβ and MAPKs
Figure 1.3 Osteoclast-osteoblast crosstalk and bone remodelling. Osteoclast–osteoblast interactions in the BMU. Osteoclasts differentiate from osteoclast precursors under the influence of MCSF and RANKL produced by osteoblast lineage cells including osteocytes. As osteoclasts create a resorption pit, growth factors, including TGF-β and IGF1, are released from the bone matrix. These growth factors may recruit mesenchymal osteoblast progenitors and promote their differentiation into mature cells that secrete osteoid to fill the area of resorbed bone. Some osteoblasts differentiate further into matrix-embedded osteocytes. Cell–cell contact mechanisms may also mediate osteoclast-osteoblast communication. Bidirectional signalling from osteoclast ephrins and osteoblast Eph receptors, and reverse signalling through RANKL on osteoblasts, have both been invoked.
(JNK, ERK and p38). Activation of these pathways promotes translocation and activation of transcription factors such as NFκB, NFATc1, CREB, AP-1 and MITF, which are responsible for osteoclast differentiation. RANKL binding to RANK activates NF-κB through the phosphorylation and inactivation of inhibitory κB kinases (IKKs) and NF-κB inhibitory kinase by TRAF6 (H. Ye et al., 2002); (Cao, Xiong, Takeuchi, Kurama, & Goeddel, 1996; L. Deng et al., 2000; M. C. Walsh, Kim, Maurizio, Molnar, & Choi, 2008). TRAF6, a RING-dependent ubiquitin ligase, belongs to the TRAF family which contains seven members (TRAFs 1, 2, 3, 4, 5, 6, and 7), and mainly mediates signals induced by TNF and pathogen-associated molecular patterns (PAMPs) (Bouwmeester et al., 2004). TRAF6 binds to RANK to activate NF-κB and MAPKs, including Jun N-terminal kinase (JNK) and p38, which play important roles in activation of immune cells and osteoclasts (L. Deng et al., 2000; N. Kobayashi et al., 2001; Lamotte et al., 2007; Wong et al., 1998). TRAF6 knockout mice exhibited severe osteopetrosis, along with phenotypes in other tissues such as thymic atrophy, lymph node deficiency, and splenomegaly; furthermore the RING finger domain of TRAF6 was shown to be essential for the formation of multinucleated osteoclasts (T. Kobayashi et al., 2003). TRAF2 is involved in the TNF-α-mediated activation of MAPK/JNK and NF-κB (Baud et al., 1999; Reinhard, Shamoon, Shyamala, & Williams, 1997). The interaction of TRAF2 with NIK triggers the NF-κB signalling cascade, whereas interaction with apoptosis signal-regulating kinase 1 (ASK1) activates the JNK and p38 MAPK pathways (Ichijo, 1999; Ichijo et al., 1997; Y. K. Lee, Hwang, Kwon, Surh, & Park, 2010). Overexpression of either TRAF2 or TRAF6 is sufficient to activate signalling pathways leading to NF-κB and AP-1 activity in the absence of extracellular stimuli (Aggarwal, 2000; Baud et al., 1999; Song, Regnier, Kirschning, Goeddel, & Rothe, 1997). RANKL additionally activates the Akt/PI3K pathway through a signalling complex involving c-Src and TRAF6 (Wong et al., 1999).

Activated NF-κB translocates into the nucleus and interacts with Nuclear factor of activated T cells cytoplasmic 1 (NFATC1) to trigger the auto-amplification of NFATC1 and the subsequent transcription of osteoclast specific genes, which mediate the completion of the differentiation process (Asagiri et al., 2005; J. H. Kim & Kim, 2014). The binding of TRAF6 to RANK increases intracellular calcium levels, subsequently activates calcineurin and induces NFATC1 expression via forming a complex with c-Fos and c-Jun at the NFATC1 gene promoter (Takayanagi et al., 2002).
Figure 1.4 RANK/RANKL/OPG signalling pathway. Osteoblasts express RANKL, which binds to its receptor, RANK, on the surface of osteoclasts and their precursors. OPG, secreted by osteoblasts and osteogenic stromal stem cells, acts as a decoy receptor to inhibit the effects of RANK by binding to RANKL. The RANK receptor lacks intrinsic enzymatic activity in the intracellular domains and therefore utilises interaction with adaptors or docking proteins, including TRAFs 2, 3, 5, and 6, Gab2, and Cbl to activate downstream signalling. Gab2 and Cbl are associated with RANK-mediated activation of c-Src, PI3 kinase (PI3K), and Akt, while TRAFs 2/5 and 6 can activate the TAB1/TAB2/TAK1 complex, resulting in activation of IKKβ and MAPK. Activation of these pathways promotes translocation and activation of transcription factors such as NFATC1, CREB, NF-κB, AP-1, and MITF, which are responsible for osteoclast differentiation.
1.4.2 Wnt Signalling Pathway

Wnt proteins constitute a large family of cysteine-rich secreted ligands that control development in organisms ranging from nematodes to mammals. Wnt signalling was first identified for its role in carcinogenesis (Nusse & Varmus, 1982), then for its function in embryonic development where it regulates organ development and cellular migration, proliferation, morphology and fate (Rawadi & Roman-Roman, 2005). So far, over 30 Wnt pathway proteins (including 19 known mammalian Wnt proteins) appear to be involved in bone homeostasis (Kikuchi, Yamamoto, & Sato, 2009).

The Wnt signalling pathway is composed of two distinct arms: canonical (Wnt/β-catenin pathway) and noncanonical pathways; the latter can be further divided into two categories: planar cell polarity pathway (PCP) and Wnt/Ca²⁺ pathway (Figure 1.5) (B. T. MacDonald, Tamai, & He, 2009; Rawadi & Roman-Roman, 2005). Wnt/β-catenin pathway is initiated by the binding of Wnt ligands to the frizzled (FZD) co-receptors LRP5 and LRP6. β-catenin subsequently accumulates in the cytoplasm and translocates into the nucleus, where it associates with transcription factors to control target gene transcription (Baron & Kneissel, 2013). The Wnt/β-catenin pathway is thought to affect the entire osteoblastic lineage. The noncanonical Wnt pathways do not involve β-catenin. The PCP is activated via the binding of Wnt to FZD and its co-receptors like PYK or ROR2, while the Wnt/Ca²⁺ pathway helps regulate calcium release from the endoplasmic reticulum in order to control intracellular calcium levels (Komiya & Habas, 2008).

Wnt proteins such as Wnt3, Wnt3a, Wnt4 and Wnt5b have also been found to be associated with bone phenotypes. Mutations in canonical Wnt/β-catenin pathway activators Wnt3 or Wnt3a result in skeletal malformation, indicating that Wnt3 or Wnt3a is required at the earliest stages of limb formation (Galceran, Farinas, Depew, Clevers, & Grosschedl, 1999; Niemann et al., 2004). Wnt4 attenuates bone loss in osteoporosis and skeletal aging mouse models via inhibiting NF-κB activation, mediated by transforming growth factor-β-activated kinase-1 (Tak1) in macrophages and osteoclast precursors independently of β-catenin as part of noncanonical Wnt signalling (Chang et al., 2007; B. Yu et al., 2014). Wnt5b was found to regulate mesenchymal cell aggregation and chondrocyte differentiation through JNK-dependent mechanisms, and cell adhesion through the activation of Src and subsequent cadherin receptor turnover, which is part of the non-canonical PCP (Bradley & Drissi, 2011).

LRP5 and LRP6 have been recognised as key regulators in osteoblastic bone formation. A loss-of-function mutation in the LRP5 gene is associated with low bone mass in osteoporosis-pseudoglioma syndrome (OPPG) and a gain-of-function mutation is related to high bone mass in healthy human subjects (Gong et al., 2001; Little et al., 2002). In animal studies, point mutations in both genes lead
**Figure 1.5 Wnt signalling pathway.** The Wnt/β-catenin pathway is initiated by the binding of Wnt ligands to the FZD co-receptors LRP5 and LRP6. β-catenin subsequently accumulates in the cytoplasm and translocates into the nucleus, where it associates with transcription factors to control target gene transcription via interactions with T-cell factor/lymphoid enhancing factor (TCF/LEF) and additional families of transcription factors. β-catenin can be also phosphorylated in the Axin complex, followed by ubiquitination, leading to proteasomal degradation. This process, in turn, maintains β-catenin at low levels in the cytoplasm and prevents it from translocating to the nucleus, leading to repression of Wnt target genes. The noncanonical Wnt pathways do not involve β-catenin. PCP is activated via the binding of Wnt to FZD and its coreceptors like PYK or ROR2, while the noncanonical Wnt/Ca^{2+} pathway helps regulate calcium release from the endoplasmic reticulum in order to control intracellular calcium levels through Calmodulin kinase (Camk2) and protein kinase C (PKC).
to a low bone mass phenotype (Holmen et al., 2004; Kokubu et al., 2004). Low bone mass in *Lrp5* knockout mice appears to be associated with decreased bone formation while *Lrp6* knockout mice seem to have increased bone resorption. *Lrp6* knockout mice have early developmental problems that are not compatible with life, while mice that carry mutations in both *Lrp5* and *Lrp6* have decreased BMD and limb deformities (Holmen et al., 2004).

Dickkopf1 (Dkk1, encoded by *DKKI*) and sclerostin (encoded by *SOST*) are two key Wnt antagonists. Dkk1 counteracts the Wnt-mediated effects on bone differentiation and adipogenesis through binding competitively with LRP5/LRP6 co-receptors and LRP4 receptors, resulting in inhibition of Wnt signalling. This inhibits chondrogenesis and osteogenesis, therefore preventing endochondral bone formation (Y. Chen et al., 2007; Pinzone et al., 2009). Deletion of *Dkk1* in mice results in severe developmental phenotypes including head defects, limb dysmorphogenesis, and postnatal death (Mukhopadhyay et al., 2001). On the contrary, overexpression of *Dkk1* in osteoblasts causes osteopenia, disruption of the hematopoietic stem cell (HSC) niche, and defects in HSC function (Fleming et al., 2008). Dkk1, together with Wnt proteins (in particular Wnt3), was also found to promote osteolytic metastases, and may facilitate the conversion of osteoblastic metastases to an osteolytic phenotype and vice versa (Pinzone et al., 2009). For instance, Dkk1 blocks prostate cancer-associated osteoblastic metastases without affecting tumour growth, while inhibition of Dkk1 in osteolytic prostate cancer cells altered the nature of bone metastases from osteolytic to osteoblastic (Hall, Bafico, Dai, Aaronson, & Keller, 2005; Hall, Kang, MacDougald, & Keller, 2006).

Sclerostin, a monomeric glycoprotein that is secreted by osteocytes, inhibits not only Wnt signalling but also BMP signalling in osteoblasts (Lewiecki, 2014). Lack of *SOST* expression in bone was found to be linked to rare bone conditions such as the autosomal recessive diseases sclerosteosis and van Buchem disease. These are characterised by high bone mass, suggesting an important role for *SOST* in canonical Wnt signalling (Balemans et al., 2001; Brunkow et al., 2001). Deletion of *Sost* in mice presents a high bone mass phenotype as early as 1 month of age, and high levels of bone formation on trabecular surfaces (C. Lin et al., 2009). In contrast, transgenic mice with human *SOST* develop a low bone mass phenotype (Kramer, Loots, Studer, Keller, & Kneissel, 2010). Due to the important role of *SOST* in the regulation of bone remodelling, this gene has been identified as a pharmaceutical target in osteoporosis and skeletal disorders associated with low bone mass. Investigational monoclonal antibodies to sclerostin such as Romosozumab and Blosozumab have advanced to phase II clinical trials or beyond (Padhi, Jang, Stouch, Fang, & Posvar, 2011; Recknor et al., 2015).
### 1.4.3 Endochondral Ossification Pathway

Endochondral ossification is a complex process by which growing cartilage is gradually replaced by bone to form the growing skeleton. This process mainly occurs at the physis (growth plate) and epiphysis. During this process, multipotent mesenchymal stem cells proliferate and condense, followed by differentiation into chondrocytes, which can further transform into the hypertrophic stage (Figure 1.6). During this stage, the hypertrophic chondrocytes are surrounded by extracellular matrix (ECM) and gradually calcified, of which columns of chondrocytes are formed. The columns are subsequently invaded by metaphyseal vessels and osteoblasts, and form the zone called primary spongiosa, which is composed of calcified cartilage and immature bone. Primary spongiosa gradually remodels under the control of osteoblast-osteoclast and builds up the mature bone of the metaphysis (Mackie et al., 2008; Mackie et al., 2011). Chondrocytes play a key role in endochondral ossification and the sequential changes in chondrocyte behaviour are tightly regulated by both systemic factors and locally secreted factors, including IBSP, PTH related peptide (PTHrP), osteopontin and Sp7, which act on receptors to mediate intracellular signalling and activation of chondrocyte-selective pathways (Clark and Duncan, 2015).

Transcription factors that play critical roles in the regulation of chondrocyte gene expression under the control of these extracellular factors include Runx2, Sox9 and MEF2C (Wuelling & Vortkamp, 2010, 2011). During early chondrogenesis, ECM proteins such as type II collagen alpha 1 (Col2a1), type XI collagen alpha 2 (Col11a2), and aggregan 1 (Agc1) are specifically expressed. On the other hand, during late chondrogenesis expression of these proteins is decreased, while type X collagen alpha 1 (Col10a1), Indian hedgehog (Ihh), and osteopontin expression is increased (Lefebvre & Smits, 2005). Runx2 has largely been associated with chondrocyte hypertrophy and osteoblast commitment. It dominates as the critical transcription factor in late chondrogenesis (Lefebvre & Smits, 2005). Runx2 regulates endochondral ossification through control of chondrocyte proliferation and differentiation (H. Chen et al., 2014; G. H. Tang & Rabie, 2005). Global deletion of Runx2 results in disrupted chondrocyte and osteoblast differentiation and complete loss of skeletal mineralization (Choi et al., 2001; Hinoi et al., 2006). Overexpression of Runx2 in developing chondrocytes by the Col IIa promoter results in accelerated chondrocyte maturation, ectopic ossification of cartilage, and lethality at birth (Ueta et al., 2001). Loss of Runx2 in chondrocytes also impairs OPG/RANKL signalling and chondroclast development (H. Chen et al., 2014). Deletion of Runx2 in chondrocytes causes a dramatic reduction in RANKL and a modest decrease in OPG expression.

SOX9 is a transcription factor of the SRY family, regulating sex determination, chondrocyte differentiation and numerous other developmental events (Lefebvre & de Crombrugghe, 1998). Sox9
Figure 1.6 Schematic diagram of Endochondral Ossification. Endochondral ossification begins with the condensation of multipotent MSCs that subsequently differentiate into chondrocytes and express chondrogenic matrix proteins, including Col2a1, Col11a2, and aggrecan (Agc1). These chondrocytes further differentiate into hypertrophic chondrocytes. Cartilage tissues are subsequently replaced with bone tissues following a series of events, including apoptosis of chondrocytes, degradation of cartilage matrices, and vascular invasion into cartilage. Maturation commences at the mid-point of the long bone shaft and proceeds toward the ends of the long bone. This process requires active bone modelling and remodelling to mature the long bone. Each stage is marked by expression of a particular cohort of genes.
is a master transcription factor for the multiple steps of chondrogenesis, and is predominantly expressed in early chondrogenesis (Akiyama, 2008; Goldring, Tsuchimochi, & Ijiri, 2006; Mori-Akiyama, Akiyama, Rowitch, & de Crombrugghe, 2003). It is a critical regulator of chondrocyte-specific proteins, such as Col2a1, Col11a2, and aggrecan, during cartilage and bone development (Bell et al., 1997; L. J. Ng et al., 1997; Sekiya et al., 2000). Sox9 is highly expressed in chondrocytes of the proliferating and prehypertrophic zone but is completely down-regulated when chondrocytes become hypertrophic and start to express Col10a1, suggesting that Sox9 downregulation in hypertrophic chondrocytes may be a necessary step to initiate cartilage-bone transition in the growth plate (Hattori et al., 2010; Lefebvre, Huang, Harley, Goodfellow, & de Crombrugghe, 1997; Q. Zhao, Eberspaecher, Lefebvre, & De Crombrugghe, 1997). Mutations in one allele of SOX9 in humans result in a rare and severe form of dwarfism termed campomelic dysplasia (CD), a skeletal dysplasia syndrome characterised by sex reversal and skeletal malformations of endochondral bones (T. Wagner et al., 1994). Transgenic mice misexpressing Sox9 in hypertrophic chondrocytes show an almost complete lack of bone marrow in newborns, due to severely retarded vascular invasion into hypertrophic cartilage and impaired cartilage resorption, resulting in delayed endochondral bone formation (Hattori et al., 2010). Sox9 is also necessary for BMP-2-induced chondrogenesis in both mouse embryonic fibroblasts (MEFs) and a multipotent mesenchymal cell line (Jin et al., 2006; Lengner et al., 2004; Mori-Akiyama et al., 2003; Pan et al., 2008). BMP-2 binds to the BMP receptors and stimulates both p38 and Smads phosphorylation. Activation of p38 kinase may increase the binding of NF-Y transcription activator to the proximal promoter of the Sox9 gene, and enhance Sox9 activity to regulate the expression of its target genes (Akiyama et al., 2005; Pan et al., 2008).

MEF2C (5q14.3), a transcription factor that regulates muscle and cardiovascular development, also controls bone development by activating the gene program for chondrocyte hypertrophy (Arnold et al., 2007; Edmondson, Lyons, Martin, & Olson, 1994; Potthoff & Olson, 2007). It is a necessary early regulator of chondrocyte hypertrophy and subsequent growth plate maturation. MEF2C also is a direct regulator of Col10a1 transcription and is required for appropriate temporal and spatial expression of the genetic program of chondrocyte development, including expression of Runx2 and VEGF in endochondral cartilage (Arnold et al., 2007; James, Appleton, Ulici, Underhill, & Beier, 2005). Conditional deletion of Mef2c results in a lack of bone due to impaired chondrocyte hypertrophy and subsequent growth plate vascularisation and endochondral ossification in mice (Arnold et al., 2007). Conversely, a overactivating form of MEF2C causes precocious chondrocyte hypertrophy, ossification of growth plates, and dwarfism (Arnold et al., 2007). Several recent studies have identified MEF2C to be correlated with either BMD or fracture (Velazquez-Cruz et al., 2014; L. Zhang et al., 2014; Zheng et al., 2013).
1.4.4 TGF-β/BMP Pathway

Members of the TGF-β superfamily acting through the TGF-β/BMP signalling pathway are involved in many cellular processes and are critical throughout life, in particular in both embryonic skeletal development and postnatal bone homeostasis (G. Chen, Deng, & Li, 2012; Guo & Wang, 2009). The TGF-β superfamily comprises of TGF-βs (including TGF-β1, TGF-β2, and TGF-β3), BMPs, Nodal, Actin and other related proteins. TGF-βs and BMPs interact with a tetrameric receptor complex comprised of type I and type II receptors at the cell surface that transduce intracellular signals via Smad complex or MAPK cascade. Signal transduction occurs through both the canonical Smad-dependent signalling pathway and the non-canonical-Smad-independent signalling pathway to regulate MSC differentiation during skeletal development, bone formation and bone homeostasis (Figure 1.7) (G. Chen et al., 2012). The canonical signalling requires the TGF-β/BMP ligands, receptors, and Smads, while the non-canonical signalling requires p38 MAPK instead of Smads. In the TGF-β pathway, signals are transduced via phosphorylation of R-Smad (Smad2 or 3) complexes with Smad4, followed by co-translocation into the nucleus or phosphorylation of TAK1 that recruits TAB1 to initiate the MKK-p38 MAPK or MKK–ERK1/2 signalling cascade. Both TGF-β and BMP activity is inhibited by cognate binding proteins, such as Noggin, Grem1, Chordin, CHL, and Fellistatin. Signals are transduced via phosphorylation of Smad1, 5, or 8 or TAK1/TAB1 cascade.

Even though all TGF-β subtypes have similar functional structure, only Tgfb2 knockout mice have multiple bone defects, suggesting it is essential for embryonic skeleton development (Sanford et al., 1997). Mutants of either Tgfb1 or Tgfb3 present with a nearly normal skeleton (A. B. Kulkarni et al., 1993; Proetzel et al., 1995). Among the 14 described BMPs, BMP-2, 4, 5, 6, 7, and 9 have high osteogenic capability (Luu et al., 2007). BMP-2 significantly increases osteocalcin expression while BMP-7 induces the expression of osteoblastic differentiation markers, such as ALP and accelerates calcium mineralization (Z. Huang, Ren, Ma, Smith, & Goodman, 2010; B. Shen et al., 2010). Bmp7 knockout mice die shortly after birth and develop skeletal patterning defects restricted to the rib cage, skull, and hindlimbs (Luo et al., 1995), whereas mice with conditional deletion of this gene present with normal skeletogenesis (Tsuji et al., 2010).

1.4.5 Hormone Endocrine Pathway

1.4.5.1 Sex Hormone Endocrine Pathway

Sex hormones (oestrogen and androgen) exert potent influences on the size and shape of the skeleton during growth and contribute to skeletal homeostasis during adulthood (Manolagas, 2000). They act on their target cells by binding to members of the nuclear hormone receptor superfamily; oestrogen
Figure 1.7 TGF-β/BMP signalling Pathway. TGF-βs and BMPs, acting on a tetrameric receptor complex comprised of type I and type II receptors at the cell surface, transduce intracellular signals via Smad complex (canonical) or MAPK cascade (non-canonical). In the TGF-β pathway, signals are transduced via phosphorylation of R-Smad (Smad2 or 3) complexes with Smad4 then co-translocation into the nuclei or phosphorylation of TAK1 that recruits TAB1 to initiate the MKK-p38 MAPK or MKK–ERK1/2 signalling cascade. Both TGF-β and BMP activity is inhibited by specific cognate binding proteins, including Noggin, Grem1, Chordin, CHL, and Fellistatin.
binds to oestrogen receptor ERα (encoded by *ESR1*) or ERβ (encoded by *ESR2*), and androgen binds to the androgen receptor (AR) (Beato & Klug, 2000; Manolagas, O'Brien, & Almeida, 2013). Human bone cells obtained from men and women have similar concentrations of AR and ER suggesting that both sex steroids have an important role in bone mass maintenance (Colvard et al., 1989). The androgen action on bone is more complex in males, because it not only activates the AR but also acts on ERα or ERβ following aromatization (Ohlsson & Vandenput, 2009). In addition, oestrogens and androgens bind to their cognate receptors outside the nucleus. In particular, through binding to sex-steroid receptors in the plasma membrane, oestrogens and androgens can initiate signal transduction by triggering the production of cyclic nucleotides, calcium flux and activation of cytoplasmic kinases. Activation of these kinases, in turn, leads to the phosphorylation of substrate proteins and transcription factors, which mediate some of the gene-regulatory effects of oestrogens (Hammes & Levin, 2007; Kousteni et al., 2003). Ovariectomy (OVX) and orchidectomy (ORX) cause a dramatic increase in osteoblast and osteocyte apoptosis in mice. Oestrogens and androgens suppress osteoblast and osteocyte apoptosis through binding to the ERs or the AR (Kousteni et al., 2001). Oestrogen plays a key role in regulation of bone mass and strength by controlling activity of osteoblasts and osteoclasts (Imai et al., 2009; Syed & Khosla, 2005). The majority of postmenopausal women show a marked decrease in BMD and a high bone turnover rate associated with the rapid loss of oestrogen following the menopause, which is referred to as postmenopausal osteoporosis. Oestrogens inhibit bone turnover by reducing osteoclast-mediated bone resorption and enhancing osteoblast-mediated bone formation via ERα or ERβ (Z. Chen et al., 2008; Khalid & Krum, 2016; T. Nakamura et al., 2007). ERs are classical hormone nuclear receptors and members of the steroid/thyroid hormone superfamily of nuclear receptors. Despite ERα and ERβ acting on various tissues through different pathways, they appear to share similar biological effects, including proliferation, growth, apoptosis, differentiation and angiogenesis (Burns & Korach, 2012). Both ERα and ERβ have an important role in the maintenance of bone homeostasis through effects on osteoblasts, osteoclasts and osteocytes (Cauley, 2015).

ESR1 is expressed in chondrocytes, stromal cells, and osteoblasts, suggesting that both bone and cartilage are regulated by ERα (X. Dai et al., 2014). Interesting, *Esr1* knockout mice only exhibit increased trabecular BMD in males but not in females (Sims et al., 2002). The numbers of osteoclasts and osteoblasts are reduced regardless of gender, leading to a slower bone turnover rate. A possible explanation for this phenomenon can be a failure of the negative feedback loop for estrogen synthesis and a marked increase in testosterone level (Sims et al., 2002). Some authors have also investigated the role of death receptor Fas (CD95) as a possible mediator of bone loss induced by oestrogen withdrawal (Kovacic et al., 2010; Krum et al., 2008). *Fas* knockout mice do not develop bone loss
due to the unchanged number of osteoclasts after OVX, while the number of osteoclasts is elevated in C57BL/6J mice after this procedure (Kovacic et al., 2010). Also, OVX induces greater stimulation of osteoblastogenesis in Fas knockout than in C57BL/6J mice. Fos directly acts on bone through differentiation and apoptosis of osteoblast lineage cells, suggesting its important role in the pathogenesis of postmenopausal osteoporosis. Furthermore, conditional knockout of Fas ligand (FasL) in osteoblasts leads to increased osteoclast numbers and activity, along with reduced bone mass, suggesting that osteoblast-induced osteoclast apoptosis via Fas/FasL signalling is a previously unrecognised mechanism that has an important role in the maintenance of bone mass in both physiological conditions and OVX osteoporosis (L. Wang et al., 2015).

Like ERα, ERβ has been detected by immunohistochemistry in osteoblasts, osteoclasts and osteocytes (Bord, Horner, Beavan, & Compston, 2001; Braidman et al., 2001). Female Esr2 knockout mice exhibit both cortical and trabecular BMD increases by 12 months of age, while the BMD of male counterparts remained unchanged compared to wildtype controls (Windahl et al., 2001). ERβ inhibits ERα -mediated gene transcription in the presence of ERα, whereas, in the absence of ERα, it can partially replace ERα (Lindberg et al., 2003).

Androgen acts on bone could be indirectly mediated by regulation of cytokines and growth factors expressed locally in bone (B. L. Clarke & Khosla, 2009). Androgen upregulates TGF-β and Insulin-like Growth Factors (IGFs), which stimulate bone formation (Kasperk, Wakley, Hierl, & Ziegler, 1997). Interestingly, androgens also stimulates osteoclastogenesis via downregulating IL6 thus promoting bone resorption (Bellido et al., 1995). The AR function is essential for normal bone growth and remodelling in male mice (Kawano, Kawaguchi, & Kato, 2003). In mice completely lacking AR, a reduction in trabecular and cortical bone mass was observed (Notini et al., 2007). The AR has an important role in the homeostasis of the male skeleton, as demonstrated by patients with idiopathic hypogonadotropic hypogonadism or complete androgen insensitivity syndrome due to a loss-of-function mutation in AR having low bone mass (Marcus et al., 2000). The deletion of AR in male mice results in high bone turnover with a decreased trabecular and cortical bone volume (Kawano, Sato, et al., 2003). AR is responsible for the preservation of trabecular bone in male mice (Wiren et al., 2008). Conversely, male transgenic mice overexpressing AR under the control of the 2.3-kb α1(I)-collagen promoter show increased trabecular bone volume but decreased cortical bone (Wiren et al., 2008). Furthermore, enhanced AR signalling in mature osteoblasts leads to reduced expression of osteoblastic markers such as osterix and sclerostin, and osteoclast activation in cortical bone. Men with complete androgen insensitivity syndrome have decreased areal bone density at the lumbar spine and hip compared to age- and sex-matched controls, suggesting that androgens exert direct skeletal effects via the androgen receptor, and not just indirectly via aromatisation to estrogens (Bertelloni et
The role of androgens in maintenance of female skeletal health has not been clearly defined. It is likely that androgens play a significant role in acquisition and maintenance of bone density in women, particularly at puberty (Zborowski, Cauley, Talbott, Guzick, & Winters, 2000). AR gene polymorphisms such as microsatellite or CGA repeat in postmenopausal women may be associated with BMD (H. Y. Chen, Chen, Wu, Tsai, & Tsai, 2003; Retornaz et al., 2006; Zitzmann et al., 2001).

1.4.5.2 Growth Hormone Endocrine Pathway

Growth hormone (GH) is a peptide hormone secreted from the anterior pituitary gland in a pulsatile manner under the control of the regulation of two major hypothalamic hormones, growth hormone releasing hormone (GHRH) that is a positive regulator, and somatostatin that is a negative regulator (Ohlsson, Bengtsson, Isaksson, Andreassen, & Slootweg, 1998; Veldhuis et al., 2001). It has many effects on the body, including anabolic effects in the regulation of longitudinal bone growth (Locatelli & Bianchi, 2014; Ohlsson et al., 1998).

GH increases bone formation in two ways: via a direct interaction with GHRs on osteoblasts and via an induction of endocrine and autocrine/paracrine IGF-1 (Kasukawa, Miyakoshi, & Mohan, 2004; Yakar, Courtland, & Clemmons, 2010). Circulating IGF-1 is mostly synthesised in the liver, but IGF-1 is expressed in all tissues, suggesting that autocrine/paracrine effects of local IGF-1 may be a major mechanism controlling tissue growth (Yakar et al., 1999). In vitro, GH directly acts on human osteoblasts and increases osteoblast differentiation (DiGirolamo et al., 2007; Kassem, Blum, Ristelli, Mosekilde, & Eriksen, 1993; Kassem, Mosekilde, & Eriksen, 1994). IGF-1 can be expressed by osteoblasts and osteoblast-like cells, regulated by GH (Chenu et al., 1990; Langdahl, Kassem, Moller, & Eriksen, 1998). It also increases osteoblast number and stimulates matrix production, acting as an anabolic hormone on bone formation (Langdahl et al., 1998). In humans, it is well-known that GH deficiency causes delay and defects in bone growth, and such defects can be rescued by long term GH replacement (Cowell, Woodhead, & Brody, 2000; Gonc & Kandemir, 2007; Lanes, 2000). Patients homozygous for IGF1 gene deletions exhibit growth retardation and decreased bone volume (Woods et al., 2000; Woods, Camacho-Hubner, Savage, & Clark, 1996). The levels of GH, IGF-1, and some IGF binding proteins (IGFBPs) diminish in the elderly, and in osteoporotic people are associated with greater femoral bone loss, suggesting a consistent effect of the anabolic IGF components on overall bone formation rate (Langlois et al., 1998; Yamaguchi et al., 2006). Interestingly, the age-dependent attenuation of GH, IGF-I, and IGFBP-3 levels among healthy men is not correlated with the reduction of BMD, suggesting that androgens interact with bone metabolism differently (Krassas, Papadopoulou, Koliakos, Konstantinidis, & Kalothetou, 2003).
At the cellular level, the GHR system employs the Janus kinase (JAK) signal transducer and activator of transcription (STAT) signal transduction pathway (Carter-Su & Smit, 1998). The activated GHR is associated with JAK2, a tyrosine kinase that once activated by GH phosphorylates tyrosine residues in STATs-1, -3, -5a, and 5b, with Stat 5b being the predominant target of GH. The STAT proteins translocate to the nucleus where they bind to specific DNA sequences and activate gene transcription. Also, the GH signalling pathway can be mediated by the suppression of cytokine signalling (SOCS) proteins that play an important role in growth and skeletal development as well as inflammation (Ahmed & Farquharson, 2010). The SOCS2 protein, in particular, negatively regulates GH/IGF-1 signalling, bone formation, and endochondral growth (Macrae et al., 2009).

In GH-deficient rats, administration of IGF-1 results in sub-optimal growth (Fielder et al., 1996). Mice with global knockout either Igf-1 or Gh show significant deficits in peak BMD, but that defect is more severe in the Igf-1 nulls (Kasukawa, Baylink, Guo, & Mohan, 2003; Sims et al., 2000; Sjogren et al., 2000). Igf-1 knockout mice have a 24% reduction in cortical bone size and shortened femoral lengths, but an increase in trabecular bone density and connectivity (Bikle et al., 2001). Mice lacking IGF-1 with elevated GH develop a reduction in chondrocyte hypertrophy while in mice lacking GHR both chondrocyte generation and hypertrophy are compromised, producing a compound deficit in long bone growth (J. Wang, Zhou, Cheng, Kopchick, & Bondy, 2004). Such observations support dual roles for GH in promoting longitudinal bone growth: an IGF-1-independent role in growth plate chondrocyte generation and an IGF-1-dependent role in promoting chondrocyte hypertrophy. On the contrary, overexpression of GH in erythroid cells using β-globin promoter increases bone size and BMD in transgenic mice (Saban, Schneider, Bolt, & King, 1996). Igf-1 overexpression in osteoblasts under the osteocalcin promoter (OC-Igf-1 transgenic mice) results in no change in serum IGF-1 levels up to 16 weeks but significant increases in cortical and trabecular BMD as well as trabecular bone volume and trabecular thickness (G. Zhao et al., 2000). Mice with deletion of the IGF-1 receptor resemble Igf-1 null mice in that pups are born smaller than controls and died shortly after birth (J. P. Liu, Baker, Perkins, Robertson, & Efstratiadis, 1993).

Interestingly, GH also has a catabolic effect on bone resorption, probably through direct interaction with osteoclasts or via local IGF-1 secretion by stromal cells such as osteoblasts (Guicheux et al., 1998; Nishiyama et al., 1996). GH and IGF-1 can stimulate production of proinflammatory cytokines such as TNF-α, IL-1beta, and IL-6 in osteoblasts; these cytokines can promote osteoclastogenesis (Renier, Clement, Desfaits, & Lambert, 1996; Uronen-Hansson et al., 2003).

GHRH, produced by the hypothalamus, stimulates GH secretion by stimulating GHRH receptor (GHRHR) on cells in the anterior pituitary. GHRH is released in a pulsatile manner, stimulating
similar pulsatile release of GH. GHRH binding to GHRHR results in increased GH production mainly through pathways like the cAMP-dependent pathway, the phospholipase C pathway and other minor pathways (Mayo, Miller, DeAlmeida, Zheng, & Godfrey, 1996; R. Xu, Zhao, & Chen, 2002). The cAMP-dependent pathway is initiated by the binding of GHRH to its receptor, causing the elevation of cAMP, then stimulating the production of protein kinase A, which phosphorylates and activates the transcription factor cAMP response element binding protein (CREB) (Petersenn & Schulte, 2000). Phosphorylated CREB, together with its coactivators, p300 and CREB-binding protein, stimulates de novo GH production. Autosomal recessive mutations in the GHRHR can result in near total absence of GH, and cause short stature in humans and mice (Carakushansky et al., 2003; S. C. Lin et al., 1993; Salvatori et al., 2001).

Somatostatin, also known as growth hormone–inhibiting hormone (GHIH), is a small but ubiquitous polypeptide that inhibits GH synthesis and release (Eigler & Ben-Shlomo, 2014; Thorner et al., 1990). The somatostatin system, which includes the somatostatin ligand and receptors, regulates anterior pituitary gland function, mainly inhibiting hormone secretion and to some extent pituitary tumor cell growth. This system inhibits pituitary adenylate cyclase/cAMP/PKA signalling, thereby inhibiting pituitary hormone synthesis and cell growth. Somatostatin is also associated with increased protein phosphatase activity in both human GH-secreting pituitary adenoma cells and rat cell lines as well as human non-functioning pituitary tumors (Cervia & Bagnoli, 2007). Somatostatin may contribute to the regulation of bone growth through regulating the proliferation and differentiation of chondroprogenitor and osteoprogenitor cells in vivo, and its ability to counteract the stimulatory effect of insulin (Weiss, Reddi, & Nimni, 1981).

1.4.5.3 Parathyroid Hormone and Calcitonin Endocrine Pathway

PTH, an 84-amino acid peptide hormone synthesised by the parathyroid glands, is essential for the maintenance of calcium homeostasis through its effects on bone, kidney, and the intestine (Silva & Bilezikian, 2015; Silva, Costa, Cusano, Kousteni, & Bilezikian, 2011). The direct actions of PTH are initiated by an interaction with its receptor (PTH1R), a G-protein-coupled receptor expressed in target cells, such as osteoblasts in bone and tubular cells in the kidney (N. S. Datta & Abou-Samra, 2009). The binding of PTH to the PTH1R acts on cell cycle, including stimulation of Gas-mediated activation of adenyl cyclase, which in turn promotes cAMP production and subsequent activation of protein kinase A (PKA) (Kramer, Keller, Leupin, & Kneissel, 2010). The PTH1R is also linked to Gαq-mediated activation of phospholipase and protein kinase C (PKC) (Kousteni & Bilezikian, 2008). PTH has also been shown to downregulate sclerostin, an important regulator of bone formation (Kramer, Keller, et al., 2010; Vignali et al., 2009).
Increased bone resorption is the most recognised catabolic action of PTH. It is one of the essential mechanisms by which PTH maintains calcium homeostasis, particularly in the state of a hypocalcemia. *In vivo*, PTH enhances bone resorption by increasing osteoclastic activity via its action on osteoblasts (Kanzawa, Sugimoto, Kanatani, & Chihara, 2000; S. K. Lee & Lorenzo, 1999). Continuous infusion of PTH increases RANKL and inhibits OPG mRNA expression in osteoblasts (J. C. Huang et al., 2004). *In vivo*, PTH directly increases RANKL expression by activation of cAMP/PKA-CREB pathway, and inhibits OPG expression via a PKA-CREB-AP-1 pathway (Fu, Jilka, Manolagas, & O'Brien, 2002). The catabolic effect of PTH is best represented by primary hyperparathyroidism, a classic disorder of PTH excess. Bone loss and fracture are common in patients with continuously high amounts of circulating PTH (N. Yu et al., 2010).

PTH, however, has also been demonstrated to promote bone formation when it is administrated intermittently (Iwaniec et al., 2007; Jilka et al., 2009; Neer et al., 2001). The anabolic actions of PTH on bone mass depend on its direct action on the osteoblastic lineage. It has been shown that cAMP/PKA signalling is a dominant mechanism by which PTH increases bone anabolism (Silva & Bilezikian, 2015; Silva et al., 2011; D. Yang et al., 2007). The osteoanabolic effect of PTH is achieved through increasing osteoblast number, decreasing osteoblast apoptosis or enhancing osteoblast differentiation (Jilka, 2007; Ogita, Rached, Dworakowski, Bilezikian, & Kousteni, 2008). According to the *in vitro* and *in vivo* studies, PTH increases the expression of genes that typically signal bone formation such as *RUNX2*, *ALP*, *COLIA1* and *OCN* (H. L. Chen et al., 2002; Krishnan et al., 2003; Silva et al., 2011).

PTH-related peptide (PTHrP) has similar structure but different functions compared to PTH (Schluter, 1999). PTH and PTHrP can bind to, and activate, a common G-protein coupled receptor, the PTH/PTHrP receptor or PTH1R (Juppner et al., 1991). Even though PTHrP can mimic many of the functions of PTH, these two peptides serve distinct biological functions. PTH is synthesised by and secreted from the parathyroid glands, while PTHrP is synthesised and expressed by various tissues such as skin, growth plate chondrocytes, bone, kidney and neuronal and glial tissues (Schluter, 1999). PTHrP and its receptor, the PTH/PTHrP receptor are involved in the regulation of chondrocyte differentiation, playing an important role in endochondral bone formation (Amizuka, Warshawsky, Henderson, Goltzman, & Karaplis, 1994; Chung, Lanske, Lee, Li, & Kronenberg, 1998; Juppner, 2000). Overexpression of PTHrP demonstrates delayed chondrocyte differentiation (Schipani et al., 1997). PTHrP knockout mice show accelerated differentiation of chondrocytes and perinatal death with gross skeletal abnormalities (Amizuka et al., 1994; Karaplis & Kronenberg, 1996).
PTH1R is expressed in mature and pre-osteoblasts during rat fetal development and in normal osteoblastic cells (K. Lee, Deeds, & Segre, 1995). The gene encoding the PTH1R is located on chromosome 3 in humans and has a total of 14 exons. Pth1r knockout mice exhibit decreased trabecular bone and increased thickness of cortical bone during foetal development (Lanske et al., 1999). Transgenic mice expressing constitutively active PTH1R in osteoblasts under the control of the Col-1A promoter develop a significant increase in the trabecular bone while a decrease in cortical bone therefore showing both the anabolic and catabolic actions mediated by PTH1R (Calvi et al., 2001). The human PTH type 2 receptor (PTH2R), which has an amino acid sequence similar to PTH1R, is activated by PTH and by tuberoinfundibular peptide of 39 residues (TIP39) (Usdin, 2000). Unlike PTH1R, PTH2R interacts poorly with PTHrP, but like PTH1R, it is strongly coupled to stimulation of cAMP accumulation and, more weakly, in a cell-specific manner, to increases in intracellular calcium concentration. Col2a1-hPTH2R (PTH2R-Tg) transgenic mice are viable and of nearly normal size at birth, but with a decrease in chondrocyte proliferation, suggesting an important role for PTH2R signalling in postnatal growth plate development and subsequent bone mass acquisition (D. Panda, Goltzman, Juppner, & Karaplis, 2009; D. K. Panda, Goltzman, & Karaplis, 2012).

Calcitonin, a 32-amino acid linear polypeptide, is a calcium-regulating hormone produced by parafollicular cells of the thyroid gland, although by different cells than those that produce thyroid hormones, which counters the actions of PTH (Naot & Cornish, 2008). The 'calcitonin family' consists of a group of peptide hormones that share structural similarities with calcitonin, mainly including calcitonin gene-related peptide (CGRP) and amylin. All the peptides of the calcitonin family seem to have a role in bone, although the specific bone effects of the peptides vary. Administration of calcitonin produces rapid lowering of serum calcium levels through binding of calcitonin to a specific receptor. This is predominantly mediated via inhibition of bone resorption induced by morphological changes in osteoclasts (Suzuki et al., 1996). To induce such changes in osteoclasts, calcitonin and its related peptides such as amylin and CGRP cause quiescence through a cAMP-dependent mechanism (Naot & Cornish, 2008; Owan & Ibaraki, 1994). Calcitonin also induces changes in intracellular calcium that lead to osteoclast retraction, an activity which is unique to calcitonin and is not shared by amylin or CGRP (Alam et al., 1993). In vivo, calcitonin increases bone formation via indirect effects on osteoblasts (Wallach, Rousseau, Martin, & Azria, 1999). Amylin and CGRP also stimulate osteoblast differentiation in vitro and in vivo (Villa et al., 2000). Mice with deletions of genes encoding calcitonin/CGRP were predicted to be either unaffected or to exhibit a reduction in bone mass due to enhanced osteoclastic absorption; interestingly, these knockout mice actually have normal serum calcium and exhibit increased trabecular bone mass and bone formation rates at 1 and
3 months of age (Hoff et al., 2002; Lerner, 2006). Calcitonin/CGRP knockout mice do not have significantly increased numbers of osteoclasts or expression of bone resorption markers. Mice with deletion of the amylin gene, however, exhibit bone loss due to enhanced bone resorption (Dacquin et al., 2004). Selective deletion of CGRP also leads to bone loss due to a substantial decrease in bone formation rate, with unaffected numbers of osteoblasts and osteocytes (Schinke et al., 2004). Together, these observations demonstrate calcitonin acts as an inhibitor while CGRP as a stimulator of bone formation.

Receptors for the peptides of the calcitonin family include calcitonin receptor (CTR) and calcitonin receptor-like receptor (Pondel, 2000). CTR works by activating G-proteins found on osteoclasts and its function is mediated through interacting with three receptor activity modifying proteins (RAMPs) (Naot & Cornish, 2008). As Ctr-deficient mice are not viable, two alternative knockout animal models have been developed to study the physiological role of CTR. One is heterozygous Ctr+/− mice that express reduced levels of the gene in osteoclasts compared to the wildtype mice (Dacquin et al., 2004). These mice show increased bone mass due to enhanced bone formation without changes in bone resorption, similar to the phenotype for the calcitonin-deficient animals. Another viable CTR global knockout model is with deletion of more than 94% but less than 100% of the Ctr gene (Davey et al., 2008). These knockout mice exhibit normal serum ultrafiltrable calcium levels and a mild increase in bone formation in males, suggesting that the CTR plays a modest physiological role in the regulation of bone and calcium homeostasis in mice. In human subjects, CTR mutations with a TT genotype have significantly lower lumbar BMD in comparison with the CC genotype, suggesting that the human CTR gene may play a role in BMD (Bandres et al., 2005; Masi et al., 1998).

1.4.5.4 **Vitamin D Endocrine Pathway**

Vitamin D is a group of fat-soluble secosteroids responsible for increasing absorption of calcium, iron, magnesium, phosphorus and zinc. It is a principal factor required for the development and maintenance of bone as well as for maintaining normal calcium and phosphorus homeostasis. Vitamin D has a role in a number of diverse cellular processes, including effects on differentiation and cell proliferation, on hormone secretion, and on the immune system (DeLuca, 2004).

The concept of the existence of the vitamin D endocrine system is firmly established (Bouillon, Okamura, & Norman, 1995; B. L. Riggs, 2003). Vitamin D acts on at least 34 target organs and cells types (including osteoblasts) that are known to contain the nuclear vitamin D receptor (VDR). When the nuclear VDR is occupied by its cognate ligand, 1,25-dihydroxycholecalciferol (also called calcitriol or 1,25-dihydroxyvitamin D₃), this complex is able to modulate genomic events (P. N. MacDonald, Dowd, Zhang, & Gu, 2004; Norman, 2006). Calcitriol is the hormonally active
metabolite of vitamin D₃, and is produced in the kidney and derived from calcifediol (also called 25-hydroxycholecalciferol, or 25-hydroxyvitamin D), a prehormone that is produced in the liver by hydrolysis of cholecalciferol (vitamin D₃) (Norman, 2008). Cholecalciferol is produced from skin following ultraviolet exposure and intake from certain foods. Calcitriol initiates biological responses via regulation both of gene transcription and rapid response via a membrane-receptor complex (Bouillon et al., 1995). Another key participant of the vitamin D endocrine system is the plasma vitamin D-binding protein (DBP), known to have a specific ligand-binding domain for vitamin D-related ligands, which delivers Vitamin D3 and all its metabolites to their target organs and cells (Bishop, Collins, Okamura, & Norman, 1994).

Animal models have identified that VDR is mainly expressed in osteoblasts and hypertrophic chondrocytes, but not in the multinucleated osteoclasts, chondroclasts, and bone marrow stromal cells (Y. Wang, Zhu, & DeLuca, 2014). Conventional Vdr knockout mice fail to identify VDR action in bone because of the animals' systemic defects in calcium metabolism (Eisman & Bouillon, 2014). Osteoblast-specific Vdr knockout mice, however, show a bone mass increase without any dysregulation of mineral metabolism, suggesting that the role of VDR in osteoblasts is as a negative regulator of bone mass (Miso et al., 2003; Yamamoto et al., 2013). Conversely, overexpression of Vdr in osteoblasts and osteocytes in a transgenic mouse model prevents bone loss during vitamin D deficiency (Lam et al., 2014).

The pathways discussed above have been identified as critical in regulating bone homeostasis. Bone disease occurs when the process of bone homeostasis become dysregulated. Understanding the cellular pathways and mechanisms controlling bone mass accrual and maintenance provides us with the tools to understand pathological processes involved in bone diseases.
Chapter 2
Bone Genetics and Diseases
2.1 Bone Genetics

The genetic architecture of most bone diseases, in particular osteoporosis, is typical of a complex disease with contributions from multiple genetic loci and multiple environmental factors and their interactions (M. J. Wade, 2001). The genetic architecture of osteoporosis, describing the number of disease susceptibility loci, the distribution of their allele frequencies, and the size of their effects, remains mostly unknown, even though several genes have been linked to this complex disease (Figure 2.1) (Ralston, 2007; Ralston & Uitterlinden, 2010). Most of these genetic loci have small effects, but a few have large effects on the determination of bone disease development. There are only a few rare bone diseases such as osteogenesis imperfecta and osteopetrosis that follow Mendelian inheritance patterns, which are caused by high-penetration mutations at a single locus (Basel & Steiner, 2009; Rikhotso, Reyneke, & Ferretti, 2008).

Many osteoporosis-related traits such as BMD, bone loss and fracture are highly heritable. BMD is currently the best predictor of osteoporotic fractures and thus a widely used surrogate for osteoporosis. Twin and family studies have shown that between 40% and 90% of the variance in peak BMD is genetically determined (Arden, Baker, Hogg, Baan, & Spector, 1996; Hernandez-de Sosa et al., 2014; G. M. Howard, Nguyen, Harris, Kelly, & Eisman, 1998; Park et al., 2012). Twin studies have also revealed that genes account for 25 - 45% of variation in age-related bone loss (Makovey, Nguyen, Naganathan, Wark, & Sambrook, 2007; B. D. Mitchell & Yerges-Armstrong, 2011; Shaffer et al., 2008; Zhai, Andrew, Kato, Blake, & Spector, 2009). Genetic factors are also likely a minor contributor to the risk of osteoporotic fractures. In the Framingham Heart Study (FHS) with total 4134 cases of vertebral fracture, heritability was estimated up to 43% (C. T. Liu et al., 2012). The heritability of wrist and hip fractures was also estimated at 54% and 68% respectively in perimenopausal women (Andrew, Antioniades, Scurrah, Macgregor, & Spector, 2005; Michaelsson, Melhus, Ferm, Ahlbom, & Pedersen, 2005). So far, approximately 200 genes have been investigated in at least one study for their relationship with BMD or fractures in studies with human populations (see Supplemental Table S2.1).

Genes and genetic polymorphisms influencing osteoporosis-related traits could be different from site to site, female to male, old to young, and even among different ethnic groups. To investigate whether the genes underlying bone acquisition act in a site-specific manner, a genome wide association study (GWAS) assessed the genetic correlation of BMD at axial and appendicular skeletal sites including the upper limbs, lower limbs and skull (Kemp et al., 2014). Meta-analyses of GWAS of BMD were performed in 5,330 children recruited by the Avon Longitudinal Study of Parents and their Children (ALSPAC) and 4,086 children from the Generation R Study. Fifteen independent signals from 13 loci
Figure 2.1 Genetic predisposition and architecture in osteoporosis. Allele frequency is defined as below: common (< 0.5), uncommon (< 0.05), and rare (< 0.005). Mutations are considered as rare variants, mostly with allele frequency less than 0.001, and with extremely large effects. Known alleles that contribute to regulation of BMD include rare variants with large effects (left top, yellow), and common variants with small effects (right bottom, sky blue). Several genes such as COL1a1 and LRP5 include variants that contribute to phenotypes in either dominant or recessive patterns. Uncommon variants with large effects (dark pink) are represented by WNT16 and EN1. However, the most common variants within the genes present small effects, this includes variants in RANK, RANKL, OPG and LRP4. Some genes may have variants in both categories. Common variants with large effects are unlikely to exist, while rare variants with small effects are difficult to identify using current technology. Less common variants with moderate effects are likely to exist, and they may explain the majority of missing inheritance of osteoporosis.
reached genome-wide significance across different skeletal regions. In particular, at 7q31.31 locus, variants at CPED1 were found to exert a larger influence on BMD of the skull and the upper limbs than on BMD of the lower limbs, while variants at Wnt16 influenced the BMD of the upper limbs more than BMD of the skull and lower limbs. This study confirms the association that was observed from a previous GWAS meta-analysis at the 7q31.31 locus, investigating total-body bone mineral density variation in 2,660 children of different ethnic backgrounds (Medina-Gomez et al., 2012). These results suggest that BMD at different skeletal sites is under a mixture of shared and specific genetic and environmental influences. Allelic heterogeneity was found at the Wnt16 locus, showing two independent signals influencing total body and skull BMD variation in children and adults. One of the skull BMD signals mapping to C7orf58 is mostly driven by children, suggesting temporal determination on peak bone mass acquisition. Certain genes may have stronger effects on bone mass at a specific age group.

Most genes influencing bone mass seem to be allelically homogeneous within human populations, correlation of BMD in one ethnic group can be replicated in another group. Some of the most important genes such as LRP5, RANK, OPG and RANKL are linked to BMD variation in a homogeneous manner (Dong et al., 2009). On the other hand, allelic heterogeneity (genetic variations within genes) have also been observed (Jakubowska-Pietkiewicz, Mlynarski, Klich, Fendler, & Chlebna-Sokol, 2012; Ji, Yao, Sun, Zhang, & Han, 2009; L. Shen et al., 2013; Tranah et al., 2008). Gene variants are intrinsic and genuinely exist, most commonly presenting as SNPs. However, many of the reported allelic associations with osteoporosis have not been replicated when tested in other cohorts (Tranah et al., 2008). The inconsistent results may be partly due to extrinsic factors including the lack of statistical power to detect subtle genetic effects and the different approaches for identifying genes and gene variants as well as differences in study design.

2.2 Bone Diseases

2.2.1 Introduction to bone diseases

Most bone-related diseases are complex and attributed to the interaction between many genes, which may affect the phenotypes locally or systemically and have many cell types involved (Ralston & Uitterlinden, 2010; Williams & Spector, 2007). Those genes that have direct effects in osteoblasts, osteoclasts, osteocytes and chondrocytes are closely bone mass related, and changes in their expression or function can result in failures in bone homeostasis and subsequent development of bone diseases.
2.2.2 Osteoporosis

Osteoporosis is a common age-related systemic skeletal disease associated with low BMD and micro-architectural deterioration of bone tissue leading to an increased risk of fractures affecting both women and men (Burr, 2002; Kular et al., 2012; S. A. Riggs, Jr. & Cooney, 1983). Other features of this pathology include low bone mass, altered bone material composition and increased rates of bone turnover. Among all these phenotypes, BMD is the best predictor of bone fracture (Johansson, Kanis, Oden, Johnell, & McCloskey, 2009; Kanis et al., 2001; Kanis et al., 2007). Clinically, osteoporosis is often considered as a silent disease because it is frequently only diagnosed after the first clinical fracture has occurred (Unnanuntana, Gladnick, Donnelly, & Lane, 2010). About 40% of Caucasian postmenopausal women are affected by osteoporosis, and the lifetime fracture risk of a patient with osteoporosis is as high as 40%, with fractures most commonly occur in the spine, the hip, or the wrist (Burge et al., 2007; Unnanuntana et al., 2010).

The measurement of BMD by dual energy x-ray absorptiometry (DXA) is a validated method to diagnose osteoporosis and to predict the risk of fracture (Adler et al., 2000; Rachner, Khosla, & Hofbauer, 2011; Unnanuntana et al., 2010). Osteoporosis is defined by a T-score of -2.5 or more standard deviations below the average of a young adult. The term osteopenia is used to describe the condition when the BMD T-score is above -2.5, but below -1.0, whereas subjects with BMD T-score above -1.0 and below +2.5 are considered to have normal BMD. In recently years, high resolution micro-CT and magnetic resonance imaging (MRI) has become feasible for in vitro or in vivo evaluation of bone architecture (Batiste et al., 2004; Jiang, Zhao, White, & Genant, 2000; Muller & Ruegsegger, 1997; Ritman, Bolander, Fitzpatrick, & Turner, 1998; Thomsen et al., 2005; van der Linden, Waarsing, & Weinans, 2006). Various degrees of correlation of micro-CT and conventional histomorphometry and DXA have been reported, most are satisfactory or high (Chappard, Retailleau-Gaborit, Legrand, Basle, & Audran, 2005; Thomsen et al., 2005). High resolution micro-CT has also demonstrated its advantages in treatment follow-up and the application of bone tissue engineering such as bone scaffolds (Effendy, Khamis, & Shuid, 2013; Peyrin, 2011; J. Yang, Pham, & Crabbe, 2003).

Osteoporosis can be classified as primary osteoporosis, secondary osteoporosis and idiopathic osteoporosis. Primary osteoporosis, the most common form of osteoporosis, can be further divided into categories, including type I osteoporosis, which occurs in hypogonadal women or men; and type II osteoporosis, which is associated with the normal aging process (S. A. Riggs, Jr. & Cooney, 1983). Type I osteoporosis in women is well-known as resulting from a deficiency in estrogen which upregulates circulating levels of cytokines such as IL-1, TNF-α, granulocyte macrophage colony stimulating factor (GM-CSF) and IL-6, which are potent stimuli for osteoclast differentiation and
bone resorption activity (Horowitz, 1993; L. G. Raisz & Rodan, 2003; Russell, Espina, & Hulley, 2006). The increased level of cytokines is mediated by RANKL. The pathogenesis of bone loss in hypogonadal men, in whom circulating androgens predominate but estrogen is minimal, remains uncertain, although it has been suggested that it may be due to direct effects of androgen or estrogen deficiency (B. L. Clarke & Khosla, 2009; Del Fattore, Cappariello, & Teti, 2008; Gennari, Khosla, & Bilezikian, 2008; Kasperk et al., 1997; Moskilde, Vestergaard, & Rejnmark, 2013). Type II osteoporosis occurs in both men and women, usually over the age of 60 (S. A. Riggs, Jr. & Cooney, 1983). A decline in the number and activity of osteoblasts is seen without a decline in the activity of osteoclasts. Thus, a net bone loss is observed due to decreased bone formation (Dempster, 1995). Fractures of cortical bone (e.g., femur and pelvis) occur more commonly in patients with type II osteoporosis.

Secondary osteoporosis refers to osteoporosis caused by underlying diseases (e.g., hyperthyroidism, hyperparathyroidism or Cushing Syndrome) or treatment of such diseases (e.g., with glucocorticoids) (Diab & Watts, 2013). Glucocorticoid-induced osteoporosis has attracted attention because of the extensive use of glucocorticoids for the treatment of chronic inflammatory and autoimmune diseases (Reid, 2000; Weinstein, 2012). Idiopathic juvenile osteoporosis (IJO) is a rare condition of poorly understood aetiology and pathophysiology that affects otherwise healthy children. This condition is characterised clinically by bone pain and vertebral fractures; spontaneous recovery is observed after puberty in the majority of cases (Bacchetta et al., 2013; M. L. Kulkarni & Keshavamurthy, 2004).

Osteoporosis is a multifactorial skeletal pathology with a main genetic component (Delgado-Calle, Garmilla, & Riancho, 2012). The majority of genes associated with this pathology remain unknown because genes identified to date only explain a small portion of inherited bone phenotypes. To find the underlying genes that regulate susceptibility to low bone mass and osteoporosis, different strategies have been used, including linkage studies, studies of candidate genes with a priori hypotheses, hypothesis-free GWAS, and functional studies (Albagha & Ralston, 2006; Karasik & Cohen-Zinder, 2012; Ralston, 2001; Williams & Spector, 2007). Linkage studies, given the multifactorial nature of the common form of osteoporosis, are not the best approach to the study of this pathology because of the low penetrance of the character (Ioannidis et al., 2007). Another approach, study of a candidate gene with a priori hypotheses, has focused mainly on genes regulating bone metabolism and cytokines implicated in bone metabolism. Approximately 200 candidate genes including ESR1, COL1A1, VDR, TGFβ1, IL-6, TNF-α, OPG, and LRP5 have been investigated so far (Ralston & Uitterlinden, 2010). GWAS has been used to analyse the correlation between BMD and chromosomal loci, providing researchers with the ability to work without prior hypotheses (Cheung, Sham, Xiao, Bow, & Kung, 2012; Ioannidis et al., 2007; Richards et al., 2008; Rivadeneira et al.,
This approach, in general, focuses on a few biological pathways such as Wnt/β-catenin signalling, RANKL/RANK/OPG, endochondral ossification and TGF-β/BMP signalling pathways (Estrada et al., 2012; Richards, Zheng, & Spector, 2012). This approach has confirmed several well-established BMD genes, and suggested several novel BMD genes (Cheung et al., 2012; Y. A. Kim et al., 2013). Meta-analyses of GWAS have further added confidence to many proven genes and the approach of GWAS (Rivadeneira et al., 2009; L. Zhang et al., 2014).

Osteoporosis therapies fall into two classes, anti-resorptive drugs, which slow down bone resorption, or anabolic drugs, which stimulate bone formation. Amongst the anti-resorptive drugs, bisphosphonates, with their high affinity for bone, constitute the largest class (Fleisch, 2003; Watts, 1994). Bisphosphonates can be administered either orally or intravenously and are most widely used because they are inexpensive and used across a broad spectrum of osteoporosis types. Other anti-resorptive drugs such as raloxifene, lasofoxifene, strontium ranelate and most recently, denosumab, may represent alternatives for women with postmenopausal osteoporosis (Chesnut et al., 2004; Chitre, Shechter, & Grauer, 2011; Cummings et al., 2010; D'Amelio & Isaia, 2013; D'Haese, Santacruz, & De Broe, 2004; McClung et al., 2001). The anabolic agent PTH produces some reconstruction of the bone skeleton through the deposition of new bone tissue. The overall effect of this is to increase bone strength and reduce fracture risk (Borba & Manas, 2010). More recently, investigational monoclonal antibodies to sclerostin such as Romosozumab and Blosozumab have advanced to phase II clinical trials or beyond (Padhi et al., 2011; Recknor et al., 2015). Targeting the genes that affect BMD via various biological pathways is an effective method for treating osteoporosis and its related traits. Overall, however, the application of genetic knowledge in tackling this disease is still in its infancy.

2.2.3 Osteopetrosis

Osteopetrosis, also known as marble bone disease and Albers-Schönberg disease, refers to a group of rare, heritable disorders of the skeleton characterised by increased bone density and reduction of bone marrow (Balemans, Van Wesenbeeck, & Van Hul, 2005; Filho, de Castro Domingos, de Freitas, & Whaites, 2005; Rikhotso et al., 2008; Stark & Savarirayan, 2009; Tolar, Teitelbaum, & Orchard, 2004). Osteopetrotic conditions vary substantially in their presentation and severity, ranging from neonatal onset with malignant complications such as bone marrow failure, to benign conditions with the incidental radiological findings. Classical manifestations of malignant osteopetrosis consist of fractures, short stature, compression of nervous structures, and pancytopenia. Based on the clinical and radiological presentations, osteopetrosis could be categorised three primary types: infantile, or “malignant,” osteopetrosis, inherited in an autosomal recessive inheritance pattern; “intermediate” autosomal recessive osteopetrosis; and autosomal dominant osteopetrosis (F. Shapiro, 1993).
Recently a new classification system has been developed by the Nosology and Classification of Genetic Skeletal Disorders, which provides an overview of recognised diagnostic entities and groups them by clinical and radiographic features and molecular pathogenesis (Superti-Furga & Unger, 2007; Warman et al., 2011). According to the 2010 revised version, osteopetrosis is classified in Group 23, comprising the osteopetrosis variants and related disorders, which has been expanded following the identification of distinct genetic defects in variants of osteopetrosis (Warman et al., 2011).

Osteopetrosis can be caused by failed osteoclast development or function while bone formation is normal; and at least 10 causative genes in monogenic bone phenotypes have been identified in humans, accounting for 70% of all cases (Cui et al., 2012; Warman et al., 2011). In forms of osteopetrosis with decreased osteoclast counts, an abnormality in the molecular pathways involved in osteoclastogenesis has been suggested (Frattini, Vezzoni, Villa, & Sobacchi, 2007; Sobacchi et al., 2007; Sobacchi et al., 2013). In forms of osteopetrosis with normal or increased osteoclast counts, the bone resorption defect usually results from failed production of a ruffled border (Del Fattore et al., 2008). Most of the genes that are mutated in osteopetrosis encode proteins involved in regulating the intra- and extracellular pH of osteoclasts. Mutations in the gene encoding the A3 subunit of the H+/ATPase proton pump (TCIRG1, also termed ATP6i and OC116) are the most common causes of osteopetrosis and account for 50% of the malignant forms (W. Deng et al., 2001; Frattini et al., 2000; Tolar et al., 2004). About 10% of malignant forms are related to mutations in the chloride channel 7 gene (CLCN7), which encodes a chloride channel found only in osteoclasts (Frattini et al., 2003; Kornak et al., 2001; Tolar et al., 2004). In a clinical study with 94 patients with severe osteopetrosis, fifty-six (60%) were reported having mutations in the TCIRG1 gene while only 12 (13%) patients had mutations in the CLCN7 gene (Frattini et al., 2003). Rarely, mutations in the CA2 gene cause carbonic anhydrase II dysfunction, which accounts for a small proportion of patients with osteopetrosis (Lotan et al., 2006; Shah, Bonapace, Hu, Strisciuglio, & Sly, 2004).

Despite recent advances, in particular in gene identification, our understanding of the genetic basis of this disease is incomplete. However, identification of genes that regulate osteoclast formation and function through assessing osteopetrotic phenotypes has helped identify effective treatments for osteopetrosis, osteoporosis and other bone-related diseases.

### 2.2.4 Paget’s Disease of Bone

Paget’s disease of bone (PDB) is the second most common bone disorder diagnosed after osteoporosis, affecting 2-7% of the population over the age of 55 years in Western countries (Bolland & Cundy, 2013; Roodman & Windle, 2005). It rarely occurs in the population under 40 years of age and has the highest prevalence in the populations of British ancestry, such as the United Kingdom,
the United States, Australia, and New Zealand (Corral-Gudino, Borao-Cengotita-Bengoa, Del Pino-Montes, & Ralston, 2013).

The exact cause of PDB remains unclear; however, it is hypothesised that the primary abnormality is increased osteoclast activity, with an increase in the number and size of osteoclasts causing rapid bone resorption (Bolland & Cundy, 2013; Langston et al., 2007; Ralston, 2008). The excessive osteoclastic activity in PDB causes accelerated bone resorption and results in rapid and excessive formation of disorganised osteoid tissue that is mechanically weaker and more susceptible to deformity and fracture than normal bone. Osteoclast precursors from PDB patients show increased sensitivity to RANKL and calcitriol in comparison with control subjects (Menaa et al., 2000). Furthermore, there is evidence that osteoblastic lineages are also abnormal in PDB (Naot et al., 2007).

Genetic factors play an important role in the pathogenesis of PDB, but current evidence suggests that the disease is genetically heterogeneous and can result from mutations in one of several disease genes (Michou & Brown, 2011; Michou, Collet, Laplanche, Orcel, & Cornelis, 2006; Ralston & Albagha, 2014). Familial PDB is inherited in an autosomal dominant manner with high but incomplete penetrance with increasing age. Linkage analysis has identified potential candidate loci on chromosomes 2p36, 5q31, 5q35, 10p13, and 18q21 (Hocking et al., 2001; Hocking et al., 2002; Laurin, Brown, Morissette, & Raymond, 2002). One gene that predisposes to PDB on 5q35 is SQSTM1 and mutations in this gene are the most common cause of familial PDB (Hocking et al., 2001; Laurin et al., 2001). SQSTM1 mutations are believed to explain about 37% of familial forms and 8% of sporadic forms of PDB, and it is estimated that more than 80% of carriers develop PDB by 80 years old (Morissette, Laurin, & Brown, 2006). The identification of SQSTM1 mutations in PDB established that the SQSTM1/p62 protein played a key role in osteoclast signalling (Chamoux et al., 2009). Osteoclasts from PDB patients were found in higher numbers, contained more nuclei, were more resistant to apoptosis, and exhibited greater bone-resorption capacity than their normal counterparts. However, it has been suggested that other genes remain to be discovered because of a large number of families with autosomal dominant inheritance of classical PDB where involvement of SQSTM1 has been excluded (Hocking et al., 2004; Hocking et al., 2002). Another gene, OPTN, which is located on 10p13, has been implicated as a candidate gene for PDB by linkage analysis in families. This gene seems to account for the development of PDB in the vast majority of families of British descent who do not carry SQSTM1 mutations (Lucas et al., 2008). The TNFRSF11A gene, which encodes RANK, is located on the PDB associated chromosome 18q21 locus. It is well established that this gene plays a critical role in osteoclast differentiation and activity, but the exact mechanism as to how gene variants result in PDB is yet to be fully understood (Sirinian, Papanastasiou, Zarkadis, & Kalofonos, 2013).
2.2.5 Other Bone Disorders

Several rare diseases like osteogenesis imperfecta (OI) and osteoporosis-pseudoglioma syndrome (OPPG) have been identified with profound bone phenotypes affecting bone mass and bone fragility, mostly due to mutations in single genes. These diseases represent the extreme ends in bone mass variability, and the hypothesis is that any gene involved in such significant skeletal pathology has a major effect in the skeleton and that the same genes are likely to contribute to bone mass variability even in the normal population. Therefore, these diseases have provided important insights into the molecular pathways that regulate bone mass, bone quality and bone cell function.

2.2.5.1 Osteogenesis Imperfecta

OI, also known as brittle bone disease or Lobstein syndrome, is a congenital bone disorder characterised by low bone mass and susceptibility to fracture. It is the most common form of bone fragility in children, with estimated at least one per 12,000–15,000 live births (Harrington, Sochet, & Howard, 2014). Clinically, OI is classified as five subgroups, largely based on its severity and different modes of inheritance (F. S. Van Dijk & Sillence, 2014). Among these OI phenotypes, Type I OI includes phenotypes with mild-to-moderate severity, presenting as increased fracture rate and blue sclerae but generally no skeletal deformity or height deficit. Type II OI is lethal in prenatal or neonatal period, whereas type III OI, the progressively deforming form, is the most severe form in patients surviving the neonatal period. Type IV OI, known as ‘common variable OI’, is usually a relatively mild form of OI characterised by the presence of white sclerae.

At present, a total of 17 genetic causes of OI have been described, with COL1A1/2 mutations accounting for a large majority of OI patients, responsible for approximately 90% of all OI cases (F. S. van Dijk et al., 2012). These OI forms are inherited in an autosomal dominant manner. Over the past two decades, autosomal recessive forms of OI have also been discovered. Among these forms, genetic defects interfering with the processing and folding of type I collagen in the endoplasmic reticulum (ER) lead to variable OI phenotypes (Byers & Pyott, 2012). For example, the CRTAP, PP1B and LEPRE1 genes encode three enzymes located in the ER that are involved in the prolyl 3-hydroxylation of type I procollagen. Pathogenic mutations in these genes are responsible for moderate-to-severe syndromic forms of autosomal recessive OI.

The Wnt/β-catenin pathway, which is involved in a wide number of biological processes, has been found to be related to OI (Fahiminiya et al., 2013; Laine et al., 2013). Mutations in the WNT1 gene cause OI, characterized by severe and early-onset skeletal fragility due to loss-of-function in the receptor or ligand of the Wnt signalling pathway (Laine et al., 2013).
2.2.5.2 Osteoporosis-Pseudoglioma Syndrome

OPPG is an extremely rare autosomal recessive condition of congenital blindness and severe childhood osteoporosis with skeletal fragility, caused by loss-of-function mutations in the LRP5 gene, with only approximately 70 cases reported worldwide (Alonso et al., 2015; Laine et al., 2011; Levasseur, Lacombe, & de Vernejoul, 2005). LRP5 is a cell-surface receptor that activates the canonical Wnt/β-catenin pathway, regulating osteoblastic bone formation. The LRP5 protein carries an extracellular domain composed of four β-propellers and EGF motifs and three LDL-receptor domains, a 23–amino acid membrane-spanning segment, and an intracellular domain (Hey et al., 1998). Most loss-of-function mutations contributing to OPPG, including homozygous mutations and compound heterozygotes, have been described in the highly conserved second β-propeller binding domain of the LRP5 protein (Y. Zhang et al., 2004). Activating LRP5 mutations confer the opposing phenotype of high bone mass and fracture resistance (Laine et al., 2011; Van Wesenbeeck et al., 2003).

2.3 Approaches to Identify Genetic Architecture of Bone Diseases

Several methods have been widely used to identify genetic architecture of osteoporosis and its related traits or diseases, of those the principles and their applications underlying these methods are discussed below.

2.3.1 Genome-Wide Linkage Analysis (GWLA)

Genome-wide linkage analysis (GWLA) is the classical approach for gene discovery in an inherited monogenic Mendelian human disease. It involves searching a pattern of inheritance (such as dominant or recessive) for the disease within a family and looking for evidence of segregation of the disease within a family according to that pattern. It typically involves genotyping between 400 and 800 microsatellite markers spread at 5- to 10-cM intervals across the genome. However, higher density panels of SNP markers have become the preferred method for genome-wide linkage scans (Sawcer et al., 2004). For instance, the Affymetrix GeneChip® Human Mapping 10 K array includes 11,560 SNP markers of which all but 337 have unique positions in the genome. These panels significantly increase the accuracy in the prediction of candidate genes in either monogenic or polygenic traits such as osteoporosis. It was successful in identifying the causative gene for many monogenic diseases (Peltonen & McKusick, 2001). Its application in complex traits or diseases, however, remains limited.

Several attempts of using GWLA to detect loci that regulate BMD have failed to achieve consistent results and only very limited replication of linkage of peaks were found in different studies (Devoto et al., 2005; Hsu et al., 2007; Karasik et al., 2002; Ralston et al., 2005; H. Shen et al., 2004). In efforts to improve the statistical power for detection of loci, several meta-analyses of GWLA have been
performed but yielded limited success. For example, a meta-analysis containing 11 (3,097 families with 12,685 individuals) whole-genome scans of BMD or osteoporosis found seven linkage loci across studies but failed to detect evidence of genome-wide significance for any locus (Y. H. Lee, Rho, Choi, Ji, & Song, 2006). Another similar large-scale meta-analysis with 11,842 subjects also did not reach genome-wide statistical significance even though there was evidence for replication of several QTLs identified in individual studies (Ioannidis et al., 2007). This reflects the limited application of linkage analysis in the detection of genes modestly regulating BMD.

2.3.2 Candidate Gene Association Studies
Candidate gene studies have been at the forefront of genetic association studies. The candidate gene approach focuses on assessing the SNPs within genes that have been previously correlated with a specific trait or disease. Candidate gene association studies have been criticised on some aspects, in particular, lack of replication and lack of statistical power due to small samples. Therefore, utilisation of results from candidate gene association studies should be performed with extreme caution.

In recent years, the disadvantages of this approach have been diminished through international collaboration via projects such as the retiring international HapMap project and 1000 genomes project (http://www.internationalgenome.org/). Researchers can use allele frequencies of SNPs and haplotypes to analyse population stratification before pooling data from different populations. The most successful collaborative effort using this approach was using data from 19,195 participants (14,277 women) from five populations of European origin (Richards et al., 2009). 150 candidate genes were identified and 36,016 SNPs in these loci were assessed. SNPs from 9 gene loci (ESR1, LRP4, ITGA1, LRP5, SOST, SPP1, TNFRSF11A, TNFRSF11B, and TNFSF11) were associated with BMD at either femoral neck or lumbar spine, and SNPs from four genes (LRP5, SOST, SPP1, and TNFRSF11A) were significantly associated with fracture risk. The effect size for SNPs (n = 241) that were associated with BMD is small, ranging from 0.04 to 0.18 SD per allele.

2.3.3 Genome-Wide Association Studies (GWAS)
Unlike GWLA, genome-wide association studies (GWAS) are more practical and effective in the identification of genetic loci associated with a disease or trait by scanning multiple markers across the genome. It has greater statistical power than using linkage studies to identify common, low-penetrance and disease-susceptible variants (Hirschhorn & Daly, 2005). GWAS includes several stages (Figure 2.2). The discovery stage focuses on the association of SNPs and traits based on a large cohort with cases and controls. The second stage focuses on the replication of the most promising candidate SNP in an independent cohort. Meta-analysis can be applied to increase the statistical power of individual GWAS during this stage. The final stage focuses on the validation of the detected
association through pathway analyses, determination of mechanism or genetic manipulation in animal models. In this approach thousands SNPs are analysed in cohorts of thousands of unrelated cases and controls. This approach has improved our understanding of the genetic basis of many complex traits or diseases such as type 1 diabetes (Grant & Hakonarson, 2009; Todd et al., 2007), type 2 diabetes (McCarthy & Zeggini, 2009) and breast cancer (Easton et al., 2007; Hunter et al., 2007). The first successful application of this approach using a large numbers of SNPs (n = 92,788) in 94 cases and 658 controls was attempted to identify significant associations between myocardial infarction and two SNPs in LTA (encoding lymphotoxin-alpha) (Ozaki et al., 2002). Another group reported a GWAS of 96 cases and 50 controls for polymorphisms associated with age-related macular degeneration (Klein et al., 2005). In a more recent study of 14,000 cases and 3,000 controls, seven common diseases of major public health importance, including coronary artery disease, hypertension, RA and diabetes were analysed and 58 loci were found to be associated with these diseases (Wellcome Trust Case Control, 2007). These studies thus represent a thorough validation of the GWA approach.

In human bone research, GWAS have also provided valuable insights into the genetic architecture of the skeleton, predominantly through the study of DXA-derived BMD data, which is one of the strongest predictors of subsequent osteoporotic fracture in both men and women. The first GWAS on skeletal phenotypes was published in 2007, where 70,987 SNPs were evaluated in 1,141 patients by using 11,200 markers (Kiel et al., 2007). Even though several SNPs in candidate genes for osteoporosis, such as rs1801133 in MTHFR; rs1884052 and rs3778099 in ESR1; rs4988300 in LRP5; rs2189480 in VDR; and rs2075555 in COLIA1, were claimed to be associated with osteoporosis, none of these genes achieved genome-wide significance due to the relatively small sample size and low genetic coverage. Subsequent studies published in 2008 together identified five loci that reached the genome-wide significance with BMD, including OPG, LRP5, ESR1, RANKL and ZBTB40 (Richards et al., 2008; Styrkarsdottir et al., 2008). deCODE Genetics identified another four genome-wide significant loci, including the SOST gene at 17q21, the MARK3 gene at 14q32, the SP7 gene at 12q13 and the TNFRSF11A (RANK) gene at 18q21 (Styrkarsdottir et al., 2009).

Individual GWAS may be underpowered to detect the biggest effects of susceptible genes because of the large number of subjects usually required for this approach. Furthermore, the capture efficiency of GWAS varies due to the heterogeneity among studies. Meta-analysis is applied to analyse combined published or new data based on the assumption that the effect and direction for a given genetic variant are the same. Using this approach in GWAS provides a cost-effective way to increase power for the estimation of the impact of gene-gene and gene-environment interactions, pinpoint SNP selection and obtain in silico replication. For instance, the MARK3 gene, encoding microtubule affinity-regulating kinase 3, surprisingly, failed to achieve genome-wide significance in a large-scale
Figure 2.2 Workflow of GWAS. A typical GWAS includes three stages: discovery, replication and validation. The discovery stage focuses on searching the association of SNPs and traits based on a large cohort with cases and controls. The second stage focuses on the replication of the most promising candidate SNP in an independent cohort. Meta-analysis can be applied to increase the statistical power of individual GWAS during this stage. The final stage focuses on the validation of the detected association through pathway analyses, determination of mechanism or genetic manipulation in animal models.
meta-analysis of GWAS with five of Northern European populations (total number = 19,195) (Rivadeneira et al., 2009). This study, conducted by the GEnetic Factors of Osteoporosis (GEFOS) consortium, however, identified 13 new loci reaching genome-wide significance, including 1p31.3 (GPR177), 2p21 (SPTBN1), 3p22 (CTNNB1), 4q21.1 (MEPE), 5q14 (MEF2C), 7p14 (STARD3NL), 7q21.3 (FLJ42280), 11p11.2 (LRP4; ARHGAP1; F2), 11p14.1 (DCDC5), 11p15 (SOX6), 16q24 (FOX1), 17q21 (HDAC5) and 17q12 (CRHR1), along with confirmation of seven proven genes. In a larger GEFOS-2 analysis, with a total of 83,893 subjects from 17 GWAS and multi-ethnic groups, fifty-six loci including 32 novel loci that reach genome-wide significance with BMD either on lumbar spine or femoral neck or both were identified (Estrada et al., 2012). In addition, fourteen BMD loci were also associated with fracture risk (P < 5×10^{-4}), of which six reached P < 5 × 10^{-8} including: 18p11.21 (C18orf19), 7q21.3 (SLC25A13), LRP5, 4q22.1 (MEPE), 2p16.2 (SPTBN1) and 10q21.1 (DKK1). This is the greatest international effort on identification of genetic effects on both BMD and fracture risk. According to their results, interestingly, SNPs in genes of the RANK/RANKL/OPG pathway were not significantly associated with fracture despite being the strongest-associated BMD loci. This is probably because only a small portion of fracture risk can be explained by BMD (Muka et al., 2015; Schuit et al., 2004).

In a recent meta-analysis of GWAS, 13 previously reported loci at the genome-wide significance level were replicated in a cohort of seven GWAS samples and 11,140 subjects (L. Zhang et al., 2014). Apart from replication of previous identified loci, two novel loci, 14q24.2 (rs2274255, SMOC1) in the combined sample of mixed gender and 21q22.13 (rs170183, CLDN14) in the female-specific subject were identified. These two SNPs collectively are thought only to explain 0.3 - 0.7% of phenotypic variation in BMD. Combined with other 13 SNPs, the cumulative effect of all the genome-wide significant SNPs on BMD in this study ranged from 2.8 to 4.9% depending on population and skeletal site. An alternative approach to analysing BMD and fracture risk on either hip or spine using conventional DXA has been undertaken by the GEFOS/GENOMOS consortium. They have focused on identifying genetic determinants of heel bone properties through quantitative ultrasound, which offers an independent measurement to demonstrate a reasonably high genetic heritability of heel bones. In their recent GWAS meta-analysis, nine SNPs had genome-wide significant associations with heel bone DXA BMD, which including six previously reported genes (including ESR1, SPTBN1, RSPO3, WNT16, DKK1, and GPATCH1) and two novel SNPs, rs4869739 (CCDC170, coiled-coil domain containing 170, previously known as C6orf97, 6q25.1), and rs597319 (TMEM135, 11q14.2) (Moayyeri et al., 2014). TMEM135, a gene recently linked to osteoblastogenesis and longevity, was significantly associated with both broadband ultrasound attenuation and velocity of sound (P < 8.23×10^{-14}). This study demonstrates the potential of
ultrasound in the identification of osteoporosis susceptibility genes in addition to DXA, in particular in extremities or in young subjects.

To date, over 150 BMD loci have been replicated and over 70 BMD loci have achieved genome-wide significance based on the database of GWAS literature. Most are non-site-specific but a few are site-dependent (See Figure 2.3 and Table 2.1, and visit HuGE Navigator). However, all these genome-wide significantly associated genes probably explain less than 10% of genetic variance for this trait. Candidate genes such as VDR, THFR, IGF1, surprisingly, have not been confirmed by any GWAS or GWAS meta-analysis (Chesi et al., 2015; Estrada et al., 2012; Hsu & Kiel, 2012; Rivadeneira et al., 2009). This phenomenon is not uncommon in GWAS of other complex traits or phenotypes because of possible false-negative findings of GWAS due to the very strict genome-wide significant threshold.

2.3.4 Whole Genome Sequencing (WGS)

Whole genome sequencing (WGS) is a sequencing technique that is capable of generating a complete catalogue of all variants present within a given DNA sequence rather than having to rely on markers and patterns of linkage disequilibrium. Conventional sequencing of the whole genome of animals or human subjects is time-consuming and lacks genetic coverage. However, large scale WGS can be achieved by using the next-generation sequencing (NGS) platforms (Behjati & Tarpey, 2013; Mardis, 2013; Quail et al., 2012). These sequencing techniques are currently being used for analysis of selected areas such as candidate loci that have emerged from GWAS.

Unlike GWAS, which rely on testing common variants (MAF ≥ 5%), WGS is aimed to assess the role of low frequency (MAF 1 - 5%) or rare (MAF ≤ 1%) genetic variation. A recent sequencing-based study identified a rare nonsense variant, within the leucine-rich-repeat-containing G-protein-coupled receptor 4 (LGR4) gene (c.376C>T), associated with BMD using 4,931 Icelandic subjects with low BMD and 69,034 population-based controls (Styrkarsdottir et al., 2013). Another WGS identified two novel low-frequency non-coding variants with large effects on BMD near EN1 and WNT16 in a total 53,236 subjects, including using 26,534 cases from a combined UK10K/1000Genomes reference panel (Zheng et al., 2015). The SNP (2q14.2, rs11692564[T], MAF = 1.7%) near EN1 had an effect size 4-fold larger than the mean of previously reported common variants for lumbar spine BMD, and was also associated with a decreased risk of fracture (Estrada et al., 2012). Conditional knockout En1Cre/flox mice showed reduced bone mass, probably due to high bone turnover. Another SNP (rs148771817[T], MAF = 1.1%) near WNT16 also showed significant association with BMD, adding evidence that this gene has large effects on osteoporosis across skeletal sites. Overall, these studies provide rationale for WGS to study the genetic architecture of complex traits and disease in the general population.
Figure 2.3 Important Genes Link to BMD Variations. The key genes identified in GWAS for BMD at various skeletal sites: total hip, femoral neck, lumbar spine, wrist or radius and heel.
<table>
<thead>
<tr>
<th>Gene(s)</th>
<th>Locus</th>
<th>Hip BMD</th>
<th>Spine BMD</th>
<th>Wrist/Heel/Radius BMD</th>
<th>Type of Study</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABCF2</td>
<td>7q36.1</td>
<td>+</td>
<td>-</td>
<td>N/A</td>
<td>GWAS meta-analysis</td>
<td>(Estrada et al., 2012)</td>
</tr>
<tr>
<td>ADAMTS18</td>
<td>16q23.1</td>
<td>+</td>
<td>-</td>
<td>N/A</td>
<td>GWAS; meta-analysis</td>
<td>Xiong et al., 2009</td>
</tr>
<tr>
<td>AKAP11</td>
<td>13q14.11</td>
<td>-</td>
<td>+</td>
<td>N/A</td>
<td>GWAS meta-analysis</td>
<td>Rivadeneira et al., 2009</td>
</tr>
<tr>
<td>ALDH7A1</td>
<td>5q31</td>
<td>+</td>
<td>N/A</td>
<td>N/A</td>
<td>GWAS</td>
<td>Guo et al., 2010</td>
</tr>
<tr>
<td>ANAPC1</td>
<td>2q13</td>
<td>+</td>
<td>-</td>
<td>N/A</td>
<td>GWAS meta-analysis</td>
<td>Estrada et al., 2012</td>
</tr>
<tr>
<td>ARHGAP1</td>
<td>11p11.2</td>
<td>+</td>
<td>-</td>
<td>N/A</td>
<td>GWAS meta-analysis</td>
<td>Rivadeneira et al., 2009</td>
</tr>
<tr>
<td>AXIN1</td>
<td>16p13.3</td>
<td>+</td>
<td>+</td>
<td>N/A</td>
<td>GWAS meta-analysis</td>
<td>Estrada et al., 2012</td>
</tr>
<tr>
<td>C6orf97</td>
<td>6q25.1</td>
<td>+</td>
<td>+</td>
<td>N/A</td>
<td>GWAS meta-analysis</td>
<td>Zhang et al., 2014</td>
</tr>
<tr>
<td>C7orf58</td>
<td>7q31.31</td>
<td>-</td>
<td>+</td>
<td>N/A</td>
<td>GWAS meta-analysis</td>
<td>Estrada et al., 2012</td>
</tr>
<tr>
<td>C12orf23</td>
<td>12q23.3</td>
<td>+</td>
<td>-</td>
<td>N/A</td>
<td>GWAS meta-analysis</td>
<td>Estrada et al., 2012</td>
</tr>
<tr>
<td>C16orf38/CLCN7</td>
<td>16p13.3</td>
<td>+</td>
<td>+</td>
<td>N/A</td>
<td>GWAS meta-analysis</td>
<td>Estrada et al., 2012</td>
</tr>
<tr>
<td>C18orf19</td>
<td>18p11.21</td>
<td>+</td>
<td>-</td>
<td>N/A</td>
<td>GWAS meta-analysis</td>
<td>Estrada et al., 2012</td>
</tr>
<tr>
<td>CDKAL1/SOX4</td>
<td>6p22.3</td>
<td>+</td>
<td>+</td>
<td>N/A</td>
<td>GWAS meta-analysis</td>
<td>Estrada et al., 2012</td>
</tr>
<tr>
<td>CLDN14</td>
<td>21q22.13</td>
<td>+</td>
<td>-</td>
<td>N/A</td>
<td>GWAS meta-analysis</td>
<td>Zhang et al., 2014</td>
</tr>
<tr>
<td>CPED1</td>
<td>7q31.31</td>
<td>+</td>
<td>+</td>
<td>(radius, female)</td>
<td>GWAS, GWAS meta-analysis</td>
<td>Chesi et al., 2015; Estrada et al., 2012</td>
</tr>
<tr>
<td>CPN1</td>
<td>10q24.2</td>
<td>+</td>
<td>-</td>
<td>N/A</td>
<td>GWAS meta-analysis</td>
<td>Estrada et al., 2012</td>
</tr>
<tr>
<td>CRHR1</td>
<td>17q12–q22</td>
<td>-</td>
<td>+</td>
<td>N/A</td>
<td>GWAS meta-analysis</td>
<td>Rivadeneira et al., 2009</td>
</tr>
<tr>
<td>CTNNB1</td>
<td>3p22</td>
<td>+</td>
<td>-</td>
<td>N/A</td>
<td>GWAS meta-analysis</td>
<td>Rivadeneira et al., 2009</td>
</tr>
<tr>
<td>CYLD</td>
<td>16q12.1</td>
<td>-</td>
<td>+</td>
<td>N/A</td>
<td>GWAS meta-analysis</td>
<td>Estrada et al., 2012</td>
</tr>
<tr>
<td>DCDC5</td>
<td>11p14.1</td>
<td>-</td>
<td>+</td>
<td>N/A</td>
<td>GWAS meta-analysis</td>
<td>Rivadeneira et al., 2009</td>
</tr>
<tr>
<td>DHH</td>
<td>12q13.12</td>
<td>-</td>
<td>+</td>
<td>N/A</td>
<td>GWAS meta-analysis</td>
<td>Estrada et al., 2012</td>
</tr>
<tr>
<td>DNM3</td>
<td>1q24.3</td>
<td>+</td>
<td>-</td>
<td>N/A</td>
<td>GWAS meta-analysis</td>
<td>Estrada et al., 2012</td>
</tr>
<tr>
<td>ERC1/Wnt5B</td>
<td>12p13.33</td>
<td>+</td>
<td>+</td>
<td>N/A</td>
<td>GWAS meta-analysis</td>
<td>Estrada et al., 2012</td>
</tr>
<tr>
<td>ESR1</td>
<td>6q25.1</td>
<td>+</td>
<td>+</td>
<td>N/A</td>
<td>GWAS</td>
<td>Moayyeri et al., 2014; Zhang et al., 2014</td>
</tr>
<tr>
<td>Gene</td>
<td>Chromosome</td>
<td>Effect</td>
<td>Association</td>
<td>Study(s)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------</td>
<td>------------</td>
<td>--------</td>
<td>-------------</td>
<td>---------------------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FAM3C</td>
<td>7q31.31</td>
<td>+</td>
<td>+</td>
<td>GWAS, GWAS meta-analysis</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Chesi et al., 2015; Estrada et al., 2012</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FAM9B/ KAL1</td>
<td>Xp22.31</td>
<td>-</td>
<td>+</td>
<td>N/A GWAS meta-analysis</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Estrada et al., 2012</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FGFR1L1</td>
<td>4p16.3</td>
<td>+</td>
<td>N/A</td>
<td>GWAS meta-analysis</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Zhang et al., 2014</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FLJ42280</td>
<td>7q21.3</td>
<td>+</td>
<td>+</td>
<td>N/A GWAS meta-analysis</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Zhang et al., 2014; Rivadeneira et al., 2009</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FOXC2</td>
<td>16q24.3</td>
<td>-</td>
<td>+</td>
<td>N/A GWAS meta-analysis</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Rivadeneira et al., 2009</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FOXL1</td>
<td>16q24.3</td>
<td>-</td>
<td>+</td>
<td>GWAS; GWAS meta-analysis</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Bae et al., 2016; Zhang et al., 2014; Rivadeneira et al., 2009</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FUBP3</td>
<td>9q34.11</td>
<td>+</td>
<td>-</td>
<td>N/A GWAS meta-analysis</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Estrada et al., 2012</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GALNT3</td>
<td>2q24</td>
<td>+</td>
<td>-</td>
<td>N/A GWAS</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Duncan et al., 2011</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GPATCH1</td>
<td>19q13.11</td>
<td>-</td>
<td>+</td>
<td>N/A GWAS meta-analysis</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Estrada et al., 2012</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GPR177</td>
<td>1p31.3</td>
<td>+</td>
<td>+</td>
<td>N/A GWAS meta-analysis</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Rivadeneira et al., 2009</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GREM2</td>
<td>1q43</td>
<td>+</td>
<td>+</td>
<td>N/A GWAS meta-analysis</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Paternoster et al., 2013</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HDAC5</td>
<td>17q21</td>
<td>+</td>
<td>-</td>
<td>N/A GWAS Meta-analysis</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Rivadeneira et al., 2009</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IDUA</td>
<td>4p16.3</td>
<td>-</td>
<td>+</td>
<td>N/A GWAS meta-analysis</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Estrada et al., 2012</td>
<td></td>
<td></td>
</tr>
<tr>
<td>INSIG2</td>
<td>2q14.2</td>
<td>-</td>
<td>+</td>
<td>N/A GWAS meta-analysis</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Estrada et al., 2012</td>
<td></td>
<td></td>
</tr>
<tr>
<td>KCNMA1</td>
<td>10q22.3</td>
<td>-</td>
<td>+</td>
<td>N/A GWAS meta-analysis</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Estrada et al., 2012</td>
<td></td>
<td></td>
</tr>
<tr>
<td>KIAA2018</td>
<td>3q13.2</td>
<td>+</td>
<td>-</td>
<td>N/A GWAS meta-analysis</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Estrada et al., 2012</td>
<td></td>
<td></td>
</tr>
<tr>
<td>KLHDC5/ PTHLH</td>
<td>12p11.22</td>
<td>+</td>
<td>-</td>
<td>N/A GWAS meta-analysis</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Estrada et al., 2012</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LERK1</td>
<td>3q25.31</td>
<td>-</td>
<td>+</td>
<td>N/A GWAS meta-analysis</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Estrada et al., 2012</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LIN7C</td>
<td>11p14.1</td>
<td>-</td>
<td>+</td>
<td>N/A GWAS meta-analysis</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Estrada et al., 2012</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LRP4</td>
<td>11p11.2</td>
<td>+</td>
<td>-</td>
<td>N/A GWAS; GWAS meta-analysis</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Rivadeneira et al., 2009</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LRP5</td>
<td>11q13.4</td>
<td>+</td>
<td>+</td>
<td>N/A GWAS; GWAS Meta-analysis</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Bae et al., 2016; Estrada et al., 2012; Rivadeneira et al., 2009</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MARK3</td>
<td>14q32</td>
<td>+</td>
<td>-</td>
<td>N/A GWAS meta-analysis</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Xiao et al., 2013</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MBL2/ DKK1</td>
<td>10q21.1</td>
<td>+</td>
<td>+</td>
<td>N/A GWAS meta-analysis</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Estrada et al., 2012</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MEF2C</td>
<td>5q14.3</td>
<td>+</td>
<td>-</td>
<td>N/A GWAS; GWAS meta-analysis</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Bae et al., 2016; Rivadeneira et al., 2009</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MEPE</td>
<td>4q21.1</td>
<td>-</td>
<td>+</td>
<td>N/A GWAS meta-analysis</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Zhang et al., 2014; Estrada et al., 2012;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gene</td>
<td>Chromosome Position</td>
<td>Effect</td>
<td>Sex</td>
<td>Study (Reference)</td>
<td>Meta-analysis</td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td>---------------------</td>
<td>--------</td>
<td>-----</td>
<td>------------------</td>
<td>--------------</td>
<td></td>
</tr>
<tr>
<td>MIRR31HG/MTAP</td>
<td>9p21.3</td>
<td>N/A</td>
<td>N/A</td>
<td>+ (radius, male)</td>
<td>GWAS</td>
<td>Chesi et al., 2015</td>
</tr>
<tr>
<td>MPP7</td>
<td>10p11.23</td>
<td>-</td>
<td>+</td>
<td>N/A</td>
<td>GWAS meta-analysis</td>
<td>Estrada et al., 2012</td>
</tr>
<tr>
<td>NTANI</td>
<td>16p13.11</td>
<td>+</td>
<td>+</td>
<td>N/A</td>
<td>GWAS meta-analysis</td>
<td>Estrada et al., 2012</td>
</tr>
<tr>
<td>PKDCC</td>
<td>2p21</td>
<td>+</td>
<td>-</td>
<td>N/A</td>
<td>GWAS meta-analysis</td>
<td>Estrada et al., 2012</td>
</tr>
<tr>
<td>RPS6KA5</td>
<td>14q32.12</td>
<td>+</td>
<td>+</td>
<td>N/A</td>
<td>GWAS meta-analysis</td>
<td>Estrada et al., 2012</td>
</tr>
<tr>
<td>RSPO3</td>
<td>6q22-23</td>
<td>+</td>
<td>-</td>
<td>N/A</td>
<td>GWAS</td>
<td>Duncan et al., 2012</td>
</tr>
<tr>
<td>SALL1/CYLD</td>
<td>16q12.1</td>
<td>+</td>
<td>-</td>
<td>N/A</td>
<td>GWAS meta-analysis</td>
<td>Estrada et al., 2012</td>
</tr>
<tr>
<td>SPT3H/RUNX2</td>
<td>6p21.1</td>
<td>-</td>
<td>+</td>
<td>N/A</td>
<td>GWAS meta-analysis</td>
<td>Estrada et al., 2012</td>
</tr>
<tr>
<td>SMG6</td>
<td>17p13.3</td>
<td>+</td>
<td>+</td>
<td>N/A</td>
<td>GWAS meta-analysis</td>
<td>Estrada et al., 2012</td>
</tr>
<tr>
<td>SMOC1</td>
<td>14q24.2</td>
<td>-</td>
<td>+</td>
<td>N/A</td>
<td>GWAS meta-analysis</td>
<td>Zhang et al., 2014</td>
</tr>
<tr>
<td>SOX6</td>
<td>11p15.3</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>GWAS meta-analysis</td>
<td>Zhang et al., 2014; Rivadeneira et al., 2009</td>
</tr>
<tr>
<td>SOX9</td>
<td>17q24.3</td>
<td>+</td>
<td>-</td>
<td>N/A</td>
<td>GWAS meta-analysis</td>
<td>Estrada et al., 2012</td>
</tr>
<tr>
<td>SP7 (Osterix)</td>
<td>12q13</td>
<td>-</td>
<td>+</td>
<td>N/A</td>
<td>GWAS meta-analysis</td>
<td>Rivadeneira et al., 2009</td>
</tr>
<tr>
<td>SPTBN1</td>
<td>2p16.2</td>
<td>-</td>
<td>+</td>
<td>+ (heel)</td>
<td>GWAS; GWAS meta-analysis</td>
<td>Moayyeri et al., 2014; Estrada et al., 2012; Rivadeneira et al., 2009</td>
</tr>
<tr>
<td>STARD3NL</td>
<td>7p14</td>
<td>-</td>
<td>+</td>
<td>N/A</td>
<td>GWAS meta-analysis</td>
<td>Rivadeneira et al., 2009</td>
</tr>
<tr>
<td>TGFB3</td>
<td>1p22.1</td>
<td>-</td>
<td>+</td>
<td>N/A</td>
<td>GWAS; meta-analysis</td>
<td>Xiong et al., 2009</td>
</tr>
<tr>
<td>TNFRSF11A (RANK)</td>
<td>18q21</td>
<td>+</td>
<td>+</td>
<td>N/A</td>
<td>GWAS; GWAS meta-analysis</td>
<td>Zhang et al., 2014; Rivadeneira et al., 2009</td>
</tr>
<tr>
<td>TNFRSF11B (OPG)</td>
<td>8q24.12</td>
<td>+</td>
<td>+</td>
<td>N/A</td>
<td>GWAS; GWAS meta-analysis</td>
<td>Zhang et al., 2014; Rivadeneira et al., 2009</td>
</tr>
<tr>
<td>TNFSF11</td>
<td>13q14</td>
<td>+</td>
<td>-</td>
<td>N/A</td>
<td>GWAS meta-analysis</td>
<td>Rivadeneira et al., 2009</td>
</tr>
<tr>
<td>WNT3</td>
<td>17q21.32</td>
<td>+</td>
<td>N/A</td>
<td>N/A</td>
<td>GWAS meta-analysis</td>
<td>Estrada et al., 2012</td>
</tr>
<tr>
<td>WNT4</td>
<td>1p36.12</td>
<td>-</td>
<td>+</td>
<td>N/A</td>
<td>GWAS meta-analysis</td>
<td>Estrada et al., 2012</td>
</tr>
<tr>
<td>WNT9B</td>
<td>17q21.32</td>
<td>+</td>
<td>N/A</td>
<td>N/A</td>
<td>GWAS meta-analysis</td>
<td>Estrada et al., 2012</td>
</tr>
<tr>
<td>WNT16</td>
<td>7q31.31</td>
<td>+</td>
<td>+</td>
<td>+ (heel)</td>
<td>GWAS meta-analysis</td>
<td>Moayyeri et al., 2014; Zhang et al., 2014; Estrada et al., 2012</td>
</tr>
<tr>
<td>XKR9/ LACTB2</td>
<td>8q13.3</td>
<td>+</td>
<td>-</td>
<td>N/A</td>
<td>GWAS meta-analysis</td>
<td>Estrada et al., 2012</td>
</tr>
<tr>
<td>ZBTB40</td>
<td>1p36.12</td>
<td>N/A</td>
<td>+</td>
<td>N/A</td>
<td>GWAS meta-analysis</td>
<td>Zhang et al., 2014; Rivadeneira et al., 2009</td>
</tr>
</tbody>
</table>
2.3.5 The Collaborative Cross

The Collaborative Cross (CC) is a large, multiparental, recombinant inbred (RI) strain panel that was motivated to serve as a common mouse library for genetic studies in the light of achieving complete sequencing of the mouse and human genomes (Chesler et al., 2008; Churchill et al., 2004). The concept for the CC was officially proposed at the 15th International Mouse Genome Conference (IMGC) in 2001 (Threadgill, Hunter, & Williams, 2002). The project was formally initiated in 2004 at The Jackson Laboratory in the US and expanded to Australia, in which the project is also known as “The Gene Mine”, along with another two collaborated facilities in Kenya and Israel. The aim of the CC project is to develop a new resource that will enhance quantitative trait locus (QTL) and systems genetic analyses in mice.

The CC fundamentally is a specific form of animal GWAS. It is a large-scale project for generating RI strains from eight genetically diverse inbred strains, including five classical inbred strains (A/J, C57BL/6J, 129S1/SvImJ, NOD/LtJ, NZO/H1LtJ) and three wild-derived subspecies (CAST/EiJ, PWK/PhJ, and WSB/EiJ) (Figures 2.4) (Chesler et al., 2008; Churchill et al., 2004; H. Yang et al., 2011). These founder strains were estimated to capture over 90% of allelic variations in laboratory mice originating from Mus musculus and that captured variation is randomly distributed across the mouse genomes. They are colour-coded and arranged randomly in different positions to generate a unique line by sibling mating F2 intercross mice until fully inbred for each line, which requires 20 generations or more. The overall aim of this project is to generate over 1,000 independent RI lines.

Applications of CC mice over the past decade mainly focused on pre-CC mice which are not fully inbred and a wide range of phenotypes have been found in the CC studies (Aylor et al., 2011; Bottomly et al., 2012; Ferris et al., 2013; Kelada et al., 2012; Phillippi et al., 2014). Simulations of the power and precision of genetic mapping with the CC population demonstrated superior performance to traditional or alternative strategies (Broman, 2005; Valdar, Flint, & Mott, 2006). Many other advances in complex trait analysis have been facilitated by the CC project. Imputed genomes were generated for 100 extant mouse strains resulting in almost one billion new genotype calls with less than 0.4% error genome-wide (Keane et al., 2011).

Because the CC strains are inbred, strategies that use these strains do not need genotyping of the mice as their genotypes are available from public databases. The design of GWAS in CC mice, not surprisingly, is fundamentally different from the design of GWAS in humans. Analysing traits of CC strains and searching the associations between phenotypes and genotypes are key steps in the study, followed by validation of genes of interest. Considering several advantages over human studies, including a greater ability in environmental control, fast breeding, reproducible and easier access to
Figure 2.4 The Collaborative Cross mice (recombinant inbred) breeding scheme. Eight founder strains were carefully selected, in which they can catch near 90% of genetic variance in the mouse genome. RI strains are generated by sibling mating F2 intercross mice until fully inbred for each line, which requires 20 generations or more.
trait-relevant tissues, and most importantly, genetic manipulations through gene knockout and allele swaps, CC mice studies offer another fast and steady avenue to explore genetic architecture of complex traits such as osteoporosis.

CC mice breeding is still in progress and its application in genetic research is in its infancy. So far, there is only one published study investigating genes for bone microarchitecture using a cohort of CC mice (Levy, Mott, Iraqi, & Gabet, 2015). This study utilised 31 lines (160 mice of both sexes in total) to perform genome-wide haplotype mapping across 77,808 SNPs, and analysed four traits including trabecular bone volume fraction (BV/TV), trabecular number (Tb.N), trabecular thickness (Tb.Th), and connectivity density (Conn.D) (Levy et al., 2015). Haplotype association mapping revealed six QTLs, namely, Trl1, Trl2 (for BV/TV), Trl3 (for Tb.N), Trl4 and Trl5 (for Tb.Th), and Trl6 (Conn.D) after gender and age stratification for each line. Several candidate genes have been detected to be associated with these traits under the identified QTLs. For example, Oxt, which encodes a precursor of oxytocin, and Avp, encoding arginine-vasopressin, appear to be promising due to their proteins having a direct skeletal effect in other studies (Colaianni et al., 2011; Colaianni et al., 2012; Sun et al., 2016; Tamma et al., 2009). These findings provide a proof-of-concept that the CC is a powerful tool in the identification of genes that are responsible for complex traits or diseases. However, limitations such as small sample size (average of five mice per line), narrow age range (aged 10 to 13 weeks), limited number of strains (n = 37) and limited parameters/traits (four) are present in this study. Such limitations may lead to a suboptimal estimation of candidate genes responsible for bone microarchitecture. Further studies with a large scale of CC mice, therefore, are required.
Chapter 3
Hypothesis and Aims
The genetic architecture of osteoporosis and its related diseases is extremely complex. Even though many genes have been correlated with BMD or osteoporosis using a variety of genetic approaches such as linkage study, candidate gene association study, GWAS, and WGS, the majority of inheritance is still missing. To identify novel genes that explain the missing inheritance further investigation is necessary. The CC is captures significant genetic diversity and would provide sufficient power and resolution for genetic dissection of polygenic traits and construction of systems genetic networks to help identify osteoporosis susceptibility genes.

Preliminary findings in our laboratory utilising micro-computed tomography (µCT) to analyse bone traits in the distal femur of CC mice revealed five prominent osteopetrotic and several osteoporotic phenotypes in the CC mice. This diversity of bone phenotypes observed within the CC mice provides a strong justification for investigating the genetic underpinnings of bone mass and micro architecture in this population. Our research aims to identify candidate genes responsible for bone formation and regulation via phenotyping and genotyping the CC mice. **We hypothesise that the CC will provide a powerful genetic tool in the discovery of genes which are responsible for osteoporosis and related diseases.**

In order to test the hypothesis, the following aims are proposed:

1. To screen the bone phenotypes of the CC mice using µCT.
2. To identify candidate genes responsible for phenotype via haplotype QTL mapping.
3. To verify the candidate genes involvement in bone modelling and remodelling through *in vitro* gene expression analysis and bioinformatics approaches.
Chapter 4
Materials and Methods
4.1 Materials

4.1.1 Equipment

<table>
<thead>
<tr>
<th>Equipment Name / Model</th>
<th>Manufacturer</th>
</tr>
</thead>
<tbody>
<tr>
<td>A&amp;D® HT500 Compact Digital Scale</td>
<td>A&amp;D Engineering Inc., USA</td>
</tr>
<tr>
<td>ATX124 Analytical Balance</td>
<td>Shimadzu Scientific Instruments Inc., USA</td>
</tr>
<tr>
<td>B-28 Incubator</td>
<td>BINDER Inc., USA</td>
</tr>
<tr>
<td>Biological Safety Cabinet Class II</td>
<td>Gelaire Pty. Ltd., Australia</td>
</tr>
<tr>
<td>Centrifuge 5430R</td>
<td>Eppendorf AG, Germany</td>
</tr>
<tr>
<td>Centrifuge 5810R</td>
<td>Eppendorf AG, Germany</td>
</tr>
<tr>
<td>Easypet® Electronic Pipette</td>
<td>Eppendorf AG, Germany</td>
</tr>
<tr>
<td>Eclipse Ti Fluorescent Microscope with Intenslight C-HGFIE Precentered Fiber Illuminator</td>
<td>Nikon Instruments Inc., USA</td>
</tr>
<tr>
<td>ESCO® PCR Cabinet</td>
<td>Esco Technologies Inc. USA</td>
</tr>
<tr>
<td>Grant-Bio PV-1 benchtop vortex</td>
<td>Thermo Fisher Scientific, USA</td>
</tr>
<tr>
<td>Imagequant LAS 4000</td>
<td>GE Healthcare, USA</td>
</tr>
<tr>
<td>Fume Control System FE2000</td>
<td>Johndec Engineering Plastics Pty. Ltd., Australia</td>
</tr>
<tr>
<td>Mastercycler® Pro Thermal Cycler</td>
<td>Eppendorf AG, Germany</td>
</tr>
<tr>
<td>MaxQ™ 4450 Benchtop Orbital Shaker</td>
<td>Thermo Fisher Scientific, USA</td>
</tr>
<tr>
<td>Micro One TOMY centrifuge</td>
<td>Quantum Scientific, Australia</td>
</tr>
<tr>
<td>Milli-Q Integral Water Purification System</td>
<td>Millipore Corporation, USA</td>
</tr>
<tr>
<td>Minispin® plus centrifuge</td>
<td>Eppendorf AG, Germany</td>
</tr>
<tr>
<td>Model 680 Microplate Reader</td>
<td>Bio-Rad, USA</td>
</tr>
<tr>
<td>PolarStar Optima Spectrofluorometer</td>
<td>BMG labtechnologies Pty Ltd., Australia</td>
</tr>
<tr>
<td>PowerPac™ HC Power Supply</td>
<td>Bio-Rad, USA</td>
</tr>
<tr>
<td>Scanscope XT</td>
<td>Aperio Technologies, Canada</td>
</tr>
<tr>
<td>Senographe DS digital mammography machine</td>
<td>GE Healthcare, United Kingdom</td>
</tr>
<tr>
<td>Skyscan 1176 micro-CT scanner</td>
<td>Skyscan, Belgium</td>
</tr>
</tbody>
</table>
### Sterile-Cycle CO₂ Incubator
- **Manufacturer**: Thermo Fisher Scientific, USA

### Thermomixer® Comfort
- **Manufacturer**: Eppendorf AG, Germany

### Tissue Processor TP1020
- **Manufacturer**: Leica Microsystems, Germany

### Veriti® 96-Well Thermal Cycler
- **Manufacturer**: Applied Biosystems, USA

### ViiA™ 7 Real-Time PCR System
- **Manufacturer**: Applied Biosystems, USA

### Xplorer® Electronic Pipette (5-100 μl)
- **Manufacturer**: Eppendorf AG, Germany

#### 4.1.2 Chemical Reagents

<table>
<thead>
<tr>
<th>Reagent</th>
<th>Manufacturer</th>
</tr>
</thead>
<tbody>
<tr>
<td>2-Mercaptoethanol 14.3M</td>
<td>Sigma-Aldrich, USA</td>
</tr>
<tr>
<td>2-Ethoxyethanol</td>
<td>Sigma-Aldrich, USA</td>
</tr>
<tr>
<td>Agarose Powder</td>
<td>Promega Corporation, USA</td>
</tr>
<tr>
<td>Alcian Blue 8GX</td>
<td>Sigma-Aldrich, USA</td>
</tr>
<tr>
<td>Alizarin Red S Monohydrate</td>
<td>MP Biochemicals Inc., USA</td>
</tr>
<tr>
<td>Aluminium Sulfate</td>
<td>Sigma-Aldrich, USA</td>
</tr>
<tr>
<td>Bromophenol Blue</td>
<td>Sigma-Aldrich, USA</td>
</tr>
<tr>
<td>Calcein</td>
<td>Sigma-Aldrich, USA</td>
</tr>
<tr>
<td>Chloroform</td>
<td>Sigma-Aldrich, USA</td>
</tr>
<tr>
<td>ClearMount™ Mounting solution</td>
<td>Invitrogen, Australia</td>
</tr>
<tr>
<td>DPX</td>
<td>BDH Laboratory Supplies, England</td>
</tr>
<tr>
<td>Di-Sodium Phosphate (Na₂HPO₄)</td>
<td>Sigma-Aldrich, USA</td>
</tr>
<tr>
<td>Eosin-Y Disodium Salt</td>
<td>Sigma-Aldrich, USA</td>
</tr>
<tr>
<td>Ethanol</td>
<td>BDH Laboratory Supplies, England</td>
</tr>
<tr>
<td>Ethylene Glycol</td>
<td>Sigma-Aldrich, USA</td>
</tr>
<tr>
<td>Ethylenediaminetetraacetic Acid (EDTA)</td>
<td>BDH Laboratory Supplies, England</td>
</tr>
<tr>
<td>Fast Red Violet LB salt</td>
<td>Sigma-Aldrich, USA</td>
</tr>
<tr>
<td>Formaldehyde Solution</td>
<td>Sigma-Aldrich, USA</td>
</tr>
<tr>
<td>Glacial Acetic Acid</td>
<td>Sigma-Aldrich, USA</td>
</tr>
<tr>
<td>Glutaraldehyde Solution</td>
<td>Thermo Fisher Scientific, USA</td>
</tr>
<tr>
<td>Glycine</td>
<td>Sigma-Aldrich, USA</td>
</tr>
<tr>
<td>Glycerol</td>
<td>BDH Laboratory Supplies, England</td>
</tr>
<tr>
<td>Haematoxylin</td>
<td>Sigma-Aldrich, USA</td>
</tr>
<tr>
<td>Chemical Name</td>
<td>Supplier</td>
</tr>
<tr>
<td>-------------------------------------------</td>
<td>-----------------------------------</td>
</tr>
<tr>
<td>Hydrochloric Acid</td>
<td>BDH Laboratory Supplies, England</td>
</tr>
<tr>
<td>Methanol</td>
<td>BDH Laboratory Supplies, England</td>
</tr>
<tr>
<td>Mono-Sodium Phosphate (NaH₂PO₄)</td>
<td>Sigma-Aldrich, USA</td>
</tr>
<tr>
<td>Naphthol AS-MS Phosphate</td>
<td>Sigma-Aldrich, USA</td>
</tr>
<tr>
<td>Nonidet® P40 Substitute (NP-40)</td>
<td>Fluka BioChemika, Switzerland</td>
</tr>
<tr>
<td>Nuclease-Free Water</td>
<td>Promega Corporation, USA</td>
</tr>
<tr>
<td>Paraformaldehyde</td>
<td>Sigma-Aldrich, USA</td>
</tr>
<tr>
<td>Phlozin B</td>
<td>Sigma-Aldrich, USA</td>
</tr>
<tr>
<td>Phenylmethanesulfonylfluoride (PMSF)</td>
<td>Roche Diagnostics, Germany</td>
</tr>
<tr>
<td>Prolong Gold Antifade Reagent</td>
<td>Invitrogen, USA</td>
</tr>
<tr>
<td>Propan-2-ol</td>
<td>BDH Laboratory Supplies, England</td>
</tr>
<tr>
<td>Protease Inhibitor Cocktail</td>
<td>Roche Diagnostics, Germany</td>
</tr>
<tr>
<td>Skim Milk Powder</td>
<td>Standard supermarket brand</td>
</tr>
<tr>
<td>Sodium Acetate Trihydrate</td>
<td>Sigma-Aldrich, USA</td>
</tr>
<tr>
<td>Sodium Chloride</td>
<td>BDH Laboratory Supplies, England</td>
</tr>
<tr>
<td>Sodium Deoxycholate</td>
<td>BDH Laboratory Supplies, England</td>
</tr>
<tr>
<td>Sodium Hydroxide</td>
<td>BDH Laboratory Supplies, England</td>
</tr>
<tr>
<td>Sodium Iodate</td>
<td>Sigma-Aldrich, USA</td>
</tr>
<tr>
<td>Sodium Orthovanadate (Na₃VO₄)</td>
<td>Sigma-Aldrich, USA</td>
</tr>
<tr>
<td>Sodium Tartrate Dehydrate</td>
<td>AJAX chemicals, Australia</td>
</tr>
<tr>
<td>Triton X-100</td>
<td>Sigma-Aldrich, USA</td>
</tr>
<tr>
<td>Trizma Base</td>
<td>Sigma-Aldrich, USA</td>
</tr>
<tr>
<td>Trizma Hydrochloride</td>
<td>Sigma-Aldrich, USA</td>
</tr>
<tr>
<td>TRIzol® Reagent</td>
<td>Invitrogen, Australia</td>
</tr>
<tr>
<td>Tween-20</td>
<td>MP Biochemicals, France</td>
</tr>
<tr>
<td>Xylene</td>
<td>Hurst Scientific, Australia</td>
</tr>
</tbody>
</table>

4.1.3 Buffers and Solutions

<table>
<thead>
<tr>
<th>Solution</th>
<th>Composition, Preparation and Storage</th>
</tr>
</thead>
<tbody>
<tr>
<td>APS</td>
<td>10 % (w/v) ammonium persulphate dissolved in ddH₂O. Stored at 4°C.</td>
</tr>
<tr>
<td>Chemical</td>
<td>Description</td>
</tr>
<tr>
<td>-------------------</td>
<td>---------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Bromophenol blue</td>
<td>1 % stock prepared in ddH&lt;sub&gt;2&lt;/sub&gt;O.</td>
</tr>
<tr>
<td>BSA-PBS</td>
<td>0.2 % (w/v) BSA dissolved in 1x PBS. Filter sterilised and stored at 4°C.</td>
</tr>
<tr>
<td>dNTPs (dATP, dCTP, dGTP, dTTP)</td>
<td>100 mM stock solution of each dNTP was prepared. 25 % (v/v) of each dNTP was combined to give a 25 mM stock of combined dNTPs. 5 nM dNTP was prepared by dilution of the 25 mM stock in ddH&lt;sub&gt;2&lt;/sub&gt;O. Aliquots stored at -20°C.</td>
</tr>
<tr>
<td>EDTA</td>
<td>14 % (w/v) EDTA diluted in Milli-Q water. Adjusted to pH 8 to dissolve EDTA and then adjusted to pH 7.4.</td>
</tr>
<tr>
<td>Haematoxylin (Gill’s)</td>
<td>250 ml ethylene glycol dissolved in 750 ml in ddH&lt;sub&gt;2&lt;/sub&gt;O, 6 g haematoxylin, 0.6 g sodium iodate, 80 g aluminium sulfate and 20 ml glacial acetic acid. Diluted 1:5 in ddH&lt;sub&gt;2&lt;/sub&gt;O and filtered before use.</td>
</tr>
<tr>
<td>Neutral Buffered Formalin (NBF)</td>
<td>10 % (v/v) formaldehyde solution, 45 mM Na&lt;sub&gt;2&lt;/sub&gt;HPO&lt;sub&gt;4&lt;/sub&gt; and 33 mM NaH&lt;sub&gt;2&lt;/sub&gt;PO&lt;sub&gt;4&lt;/sub&gt; in Milli-Q water.</td>
</tr>
<tr>
<td>Paraformaldehyde</td>
<td>4 % (w/v) paraformaldehyde in 1x PBS. Stored at -20°C.</td>
</tr>
<tr>
<td>PBS</td>
<td>10x PBS stock solution: 70 mM Ha&lt;sub&gt;2&lt;/sub&gt;HP0&lt;sub&gt;4&lt;/sub&gt;, 30 mM HaH&lt;sub&gt;2&lt;/sub&gt;PO&lt;sub&gt;4&lt;/sub&gt;, 1.3 M NaCl dissolved in Milli-Q water. 1x PBS: 10x PBS stock solution diluted 1:10 in ddH&lt;sub&gt;2&lt;/sub&gt;O and adjusted to pH 7.4.</td>
</tr>
<tr>
<td>PMSF</td>
<td>Prepared as stock 17.4 mg/ml in ethanol. Aliquots stored at -20°C.</td>
</tr>
<tr>
<td>Sodium Chloride</td>
<td>5 M stock solution in ddH&lt;sub&gt;2&lt;/sub&gt;O.</td>
</tr>
<tr>
<td>Sodium Hydroxide</td>
<td>5 M stock solution in ddH&lt;sub&gt;2&lt;/sub&gt;O.</td>
</tr>
</tbody>
</table>
Sodium Vanadate (Na$_3$VO$_4$) 0.25 M stock solution from Na$_3$VO$_4$ powder in ddH$_2$O. Aliquots stored at -20°C.

TAE 50x TAE stock solution: 2 M Trizma base, 5.71 % (v/v) glacial acetic acid and 50 mM EDTA. 1x TAE solution: dilution of 50x stock solution.

Triton X-100/PBS 0.1 % (v/v) Triton X-100 in 1x PBS. Filtered and stored at 4°C.

4.1.4 Tissue Culture Materials and Reagents

<table>
<thead>
<tr>
<th>Material / Reagent</th>
<th>Manufacturer</th>
</tr>
</thead>
<tbody>
<tr>
<td>1α25-dihydroxyvitamin D$_3$</td>
<td>Sigma-Aldrich, USA</td>
</tr>
<tr>
<td>B-glycerophosphate</td>
<td>Sigma-Aldrich, USA</td>
</tr>
<tr>
<td>Cell Dissociation Solution, Non-enzymatic</td>
<td>Sigma-Aldrich, USA</td>
</tr>
<tr>
<td>Cell Scraper</td>
<td>Sarstedt, Germany</td>
</tr>
<tr>
<td>Cryopure Cryogenic Vials (1.6ml)</td>
<td>Sarstedt, Germany</td>
</tr>
<tr>
<td>Dexamethasone</td>
<td>Sigma-Aldrich, USA</td>
</tr>
<tr>
<td>Dimethyl Sulfoxide (DMSO)</td>
<td>Sigma-Aldrich, USA</td>
</tr>
<tr>
<td>Falcon Cell Strainer (100 µm)</td>
<td>BD Biosciences, USA</td>
</tr>
<tr>
<td>Fetal Bovine Serum</td>
<td>Invitrogen, New Zealand</td>
</tr>
<tr>
<td>Gibco® α-MEM</td>
<td>Invitrogen, Australia</td>
</tr>
<tr>
<td>Gibco® DMEM</td>
<td>Invitrogen, Australia</td>
</tr>
<tr>
<td>Glass Bottom Culture dish (10 mm with 3 mm No. 1.5 glass cover slip)</td>
<td>MatTek Corporation, USA</td>
</tr>
<tr>
<td>GlutaMAX™</td>
<td>Invitrogen, Australia</td>
</tr>
<tr>
<td>Hanks’ Balanced Salt Solution (HBSS)</td>
<td>Invitrogen, New Zealand</td>
</tr>
<tr>
<td>------------------------------------</td>
<td>-------------------------</td>
</tr>
<tr>
<td>L-ascorbic Acid</td>
<td>Sigma-Aldrich, USA</td>
</tr>
<tr>
<td>Lysosensor™ DND-189</td>
<td>Molecular Probes, USA</td>
</tr>
<tr>
<td>Penicillin/Streptomycin</td>
<td>Invitrogen, Australia</td>
</tr>
<tr>
<td>Serological Pipettes (10 ml, 25 ml)</td>
<td>Sarstedt, Germany</td>
</tr>
<tr>
<td>Tissue Culture Flask (25 cm²)</td>
<td>Nunc, Denmark</td>
</tr>
<tr>
<td>Tissue Culture Flask (75 cm²)</td>
<td>Nunc, Denmark</td>
</tr>
<tr>
<td>Tissue Culture Plate (384-Wells)</td>
<td>Nunc, Denmark</td>
</tr>
<tr>
<td>TrypLE™ Express</td>
<td>Invitrogen, Australia</td>
</tr>
<tr>
<td>Trypsin-EDTA (0.25%)</td>
<td>Invitrogen, Australia</td>
</tr>
</tbody>
</table>

### 4.1.5 Oligonucleotide Primers and Housekeeping Genes Primers

Oligonucleotide primer sequences for Real-time PCR amplifications were designed using Primer-BLAST (https://www.ncbi.nlm.nih.gov/tools/primer-blast/), and screened through NCBI-BLAST to ensure specificity (J. Ye et al., 2012). The primers were purchased from Geneworks, Australia, and shipped in freeze-dried form. They were then reconstructed in sterile and nuclease-free ddH₂O at a final stock concentration of 100 µM. Aliquots of resuspended oligos were stored at -20 °C and placed on ice during usage.

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence (5’ – 3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acp5 (TRACP) Forward</td>
<td>CAGCAGCCAAGGAGGACTAC</td>
</tr>
<tr>
<td>Acp5 (TRACP) Reverse</td>
<td>ACATAGCCCACACCGTTCTC</td>
</tr>
<tr>
<td>Bglap (Osteocalcin) Forward</td>
<td>GCGCTCTGTCTCTCTGACCT</td>
</tr>
<tr>
<td>Bglap (Osteocalcin) Reverse</td>
<td>ACCTTATTGCCCCCTCGTCTGTT</td>
</tr>
<tr>
<td>Hprt Forward</td>
<td>CAGTCCCAGCGTGCTGATTA</td>
</tr>
<tr>
<td>Hprt Reverse</td>
<td>TGGCCTCCCATCTCTCTTCAT</td>
</tr>
<tr>
<td>Setbp1 Forward</td>
<td>TTTGACAGATGCTCGTGGCT</td>
</tr>
<tr>
<td>Setbp1 Reverse</td>
<td>GAAGGGTCAGATCCATGTCC</td>
</tr>
</tbody>
</table>
### 4.1.6 Softwares

<table>
<thead>
<tr>
<th>Software Description</th>
<th>Manufacturer</th>
</tr>
</thead>
<tbody>
<tr>
<td>CT Analyser (CTAn) (version: 1.15)</td>
<td>Bruker® Skyscan 1176, Belgium</td>
</tr>
<tr>
<td>CTvol (version: 2.2.1.0)</td>
<td>Bruker® Skyscan 1176, Belgium</td>
</tr>
<tr>
<td>Dataviewer (version: 1.5.0)</td>
<td>Bruker® Skyscan 1176, Belgium</td>
</tr>
<tr>
<td>Endnote (Version X8)</td>
<td>Clarivate Analytics, NY, USA</td>
</tr>
<tr>
<td>ImageJ (version: 1.47)</td>
<td>National Institutes of Health, USA</td>
</tr>
<tr>
<td>Microsoft Office 2013</td>
<td>Microsoft, USA</td>
</tr>
<tr>
<td>Graphpad Prism (version 6.01)</td>
<td>GraphPad Software, Inc. USA</td>
</tr>
<tr>
<td>NIS-Elements Basic Research (version: 3.22.14)</td>
<td>Nikon Instruments Inc., USA</td>
</tr>
<tr>
<td>NIS-Elements C</td>
<td>Nikon Instruments Inc., USA</td>
</tr>
<tr>
<td>NRecon (version: 1.6.9.8)</td>
<td>Bruker® Skyscan 1176, Belgium</td>
</tr>
<tr>
<td>ViiA™ 7 software (version 1.2.1)</td>
<td>Applied Biosystems, USA</td>
</tr>
</tbody>
</table>

### 4.1.7 Gene Mapping Links

<table>
<thead>
<tr>
<th>Gene Mapping Links</th>
<th>Website</th>
</tr>
</thead>
<tbody>
<tr>
<td>GeneMiner (Geniad)</td>
<td><a href="http://www.sysgen.org/GeneMiner/">http://www.sysgen.org/GeneMiner/</a></td>
</tr>
<tr>
<td>Mouse Genomes Project (Sanger Institute)</td>
<td><a href="http://www.sanger.ac.uk/science/data/mouse-genomes-project">http://www.sanger.ac.uk/science/data/mouse-genomes-project</a></td>
</tr>
<tr>
<td>The Collaborative Cross Database (UNC Systems Genetics)</td>
<td><a href="http://csbio.unc.edu/CCstatus/index.py">http://csbio.unc.edu/CCstatus/index.py</a></td>
</tr>
<tr>
<td>Ensembl Genome Browser 87</td>
<td><a href="http://www.ensembl.org/">www.ensembl.org/</a></td>
</tr>
</tbody>
</table>

### 4.1.8 Bioinformatics Tools

<table>
<thead>
<tr>
<th>Bioinformatics Links</th>
<th>Website</th>
</tr>
</thead>
<tbody>
<tr>
<td>BioGPS</td>
<td><a href="http://www.biogps.org">www.biogps.org</a></td>
</tr>
<tr>
<td>Uniprot</td>
<td><a href="http://www.uniprot.org/">http://www.uniprot.org/</a></td>
</tr>
</tbody>
</table>
4.2 Animals and Methods for Bone Phenotyping and Gene Mapping

4.2.1 Generation of the Collaborative Cross

The Collaborative Cross strains used in this study were generated and bred by Geniad, and housed in a specific pathogen-free facility at the Animal Resources Centre (Murdoch, WA, Australia). This mouse library is originated from eight founder strains: A/J, C57BL/6J, 129S1SvImJ, NOD/LtJ, NZO/HILtJ, CAST/EiJ, PWK/PhJ, and WSB/EiJ. (For convenience, these strains are referred as A/J, C57BL/6J, 129S1, NOD, NZO, CAST, PWK and WSB.) The Australian Code for the Care and Use of Animals for Scientific Purposes was followed, and the mice were maintained with appropriate ethics approvals.

The Geniad breeding program has generated over 900 lines, with over 100 CC strains currently at inbreeding generation 15 or beyond (Morahan, Balmer, & Monley, 2008). All the founder strains, F1 crosses, and numerous CC strains have been genotyped with the MegaMUGA array, providing dense coverage genome-wide by typing 77,808 SNPs. This platform has been developed and used since 2012 and extended Mouse Universal Genotyping Array (MUGA) from 7.5K to 77.8K markers. More detailed information about this mouse genetics resource is displayed on UNC Systems Genetics website (http://csbio.unc.edu/CCstatus/index.py).

The power of the CC was formally calculated and determined that 500 CC strains provided 67% power to detect a QTL with a 5% additive effect, while power achieved to approximately 100% when the QTL effect size exceeded 10% (Valdar et al., 2006). However, it is unrealistic to have 500 strains for testing. So far, most studies used and validated fewer than 100 strains as reasonable numbers to examine the correlations between genotypes and phenotypes (Abu-Toamih Atamni, Ziner, Mott, Wolf, & Iraqi, 2017; Atamni, Botzman, Mott, Gat-Viks, & Iraqi, 2016; Mao et al., 2015; Scoville et al., 2015; Shusterman et al., 2013). In the current study, a
total number of 887 Geniad mice with 64 strains were dissected and scanned, and 848 CC mice including 278 males in 60 strains and 570 females in 61 strains were analysed after exclusion of strains with less than 3 samples and fractured samples (Supplemental Table S4.1).

4.2.2 Work Flow of the CC mice Screening

The CC study starts with searching of the associations among strains and identification of candidate genes within the identified locus, followed by validation of those candidate genes for the phenotype (Figure 4.1). Correlation with human datasets add confidence for the SNPs/genes. Validation can be done by gene expression, pathway analysis and looking at the functions of selected gene through transgenic or knockout mouse model.

4.2.2.1 MicroCT Analysis of CC Mice

Hindlimbs were collected and cleaned of excess soft tissues prior to fixation in 10% neutral buffered formalin for 24 hours. Hindlimbs were then washed in PBS for 24 hours before being prepared for analysis. Hindlimbs were individually wrapped in tissue to immobilize them in a 1.5 ml microcentrifuge tube. 1×PBS was added in tubes to moisturise the samples. Tubes (maximum of 5) were then inserted within a foam tube and taped into the scanning bed of the µCT machine (Bruker® SkyScan 1176) to prevent any movement of the samples that could blur the images. A region from midshaft of the femur to midshaft of the tibia encompassing the knee joint on each hindlimb was then scanned using the following parameters: 50 kV, 500µA, constant exposure, 0.5mm aluminium filter, voxel size 9 nm, 0.4° angular rotation. 931 tomographic sections were acquired for each scanned hindlimb, covering a maximum length of 18 mm per hindlimb, which generated a 3.93-gigabyte file containing 695 raw images. The training for using the µCT machine and technical support were provided by Ms Diana Engineer (The University of Western Australia).

Images were then reconstructed using the SkyScan NRecon program (Bruker®, Belgium) under standardised conditions. The secondary spongiosa region (0.5mm from the growth plate and 1mm in height) within the distal femur was analysed using SkyScan CTAn software (Bruker®, Belgium) to assess 2D/3D parameters of the trabecular bone. Parameters analysed were bone volume/tissue volume (BV/TV, trabecular bone volume), trabecular number (Tb.N), trabecular separation (Tb.Sp), trabecular thickness (Tb.Th), structure model index (SMI), Connectivity density (Conn.D) and degree of anisotropy (DA). Cortical bone analysis was performed in the mid shaft (4 mm below the growth plate with a height of 1 mm), and cortical thickness (Ct.Th)
Figure 4.1 Workflow of the Collaborative Cross Mice Screening. The CC study starts with searching of the associations among strains and identification of candidate genes within the identified locus, followed by validation of those candidate genes for the phenotype. Correlation with human datasets add confidence for the SNPs/genes. Validation can be done by gene expression, pathway analysis and looking at the functions of selected gene through transgenic or knockout mouse model.
was analysed. 3D images were generated using SkyScan CTvol program (Bruker®, Belgium). Dr Jennifer Tickner kindly offered training and support in image reconstruction and analysis.

### 4.2.2.2 Strains Ranking

Strains were ranked based on their various trabecular or cortical parameters/traits in two ways. One is looking at the rankings of strains across various ages. This was done based on the percentage changes either in comparison of various matched aged C57BL/6J mice or average values of parameters of CC mice within the age groups, including young (2-7 months), middle aged (7.1-11.9 months) and old (12 months and older). The former approach was applied in BV/TV, Tb.N, Tb.Th, SMI and Ct.Th. The latter approach was applied in Tb.Sp and DA due to insufficient data from the C57B/6J control. Strains were also ranked based on the actual average values of parameters in three separate age groups.

### 4.2.2.3 Haplotype Reconstruction and QTL Mapping of CC Strains

The genotypes for the eight founders (eight replicates each) on the MegaMUGA genotyping platform from the University of North Carolina CC website. These genotypes were separated into two sets of genotypes per strain, namely homozygous genotypes of allele 1 and homozygous genotypes of allele 2 for the genome to be treated as haploid (inbred). Prof. Grant Morahan’s team input these data in HAPPY software to estimate the founder haplotype having the maximum-likelihood probability for genotype sets of allele 1 and 2 separately (Ram, Mehta, Balmer, Gatti, & Morahan, 2014). They verified these haplotype data and generated the system, GeneMiner (http://www.sysgen.org/GeneMiner/), which is an online software dedicated for gene mapping using CC mice. A detailed description of the mapping pipeline has been provided in this paper (Ram et al., 2014).

To start with the gene mapping processes, ranked strains with numbers were input in the updated GeneMiner version (19/05/2016) and run in both appropriate GWAS and QTL options (“as is” and normalised). Gene mapping was performed under the QTL options run for around 12 minutes then generated QTL haplotypes mapping for each parameter/trait. Haplotype QTL plots showed correlation between SNPs and parameters/traits in a lines, and may yield peak(s) at chromosome(s) which are above thresholds for GWS. These thresholds were computed from 1,000 additive model permutations and set by Prof. Grant Morahan’s team. It was built under the approach recommended by Durrant et al. (Durrant et al., 2011). These thresholds were represented by three lines: red line presenting 95th percentile of confidence which is equivalent to $p < 0.05$, blue line representing 90th percentile confidence which is equivalent to $p < 0.1$, and
orange line representing 37th percentile confidence which is equivalent to p < 0.63 in the QTL plots. Prof Grant Morahan and Dr Ramesh Ram (The University of Western Australia) were kindly offered support in using GeneMiner for gene mapping.

In cases that QTL plots showed high LOD scores but did not achieve any significance, normalised QTL scans were also attempted when applicable. Loci were determined on QTL plots. Genes within loci achieved 95th percentile of confidence were strongly linked to the traits, whereas genes within loci achieved 90th percentile confidence were less confident to determine the correlations. In addition, genes within the loci only achieved 37th percentile of confidence should be dealt with caution, because they are unlikely to contain candidate genes that are associated with traits even though they are reportable.

To identify candidate genes underlying the mapped QTLs, genomic intervals were identified by selecting the QTL peak at each chromosome. Then the intervals were input into the Encode/Coding CC-SNP Finder to filter out all coding SNPs (e.g., missence) using ENCODE database. Genes with deleterious SNP(s) were considered to have a high probability to contribute the trait/phenotype. Extensive searching using FANTOM5 database which is built in the GeneMiner was needed when there was no conclusive candidate variant by missense/regulatory effects because many QTLs are also attributed to structural variants like indels (INsertion/DELetions).

The founder strain(s) contributing to each trait were calculated by GeneMiner. It was determined by deriving coefficients (log odds ratio) of the fit from the logistic/multinominal regression model and using plotting tools implemented in the DOQTL R package (Gatti et al., 2014). We weighed the founder effects based on the results by the system. Founder strain with most prominent distribution which was normally represented by highest P values in founder coefficient plot was considered first because it would contribute the trait/phenotype more. SNPs/genes derived from the key founder strain were assumed to be the key candidates for the trait/phenotype at each QTL. In case of founder strains evenly distributed, genes derived from them were considered to be unlikely strong candidate for trait/phenotype contributors.

4.2.2.4 Verification of Candidate Genes Using Bioinformatics Approaches

Additional information regarding to SNPs and indels of candidate genes was obtained on the database of Mouse Genomes Project (Sanger Institute). Details of significant SNPs were obtained dbsNP Database run by National Institutes of Health, or Ensembl Genome Browser 87. Gene expression data were extracted from BioGPS, and figures for expression profiling were generated
with Graphpad® Prism (version 6.01). Another public gene annotation portal, Entrez Gene, run by National Institutes of Health, was also used in searching critical gene annotation information, including primary sequence data, genome position, associated Gene Ontology terms, gene structure, and genetic variation.

To find out whether those candidates are linked to osteoporosis or related diseases, searching literature was done by typing terms which combined the official symbol of the gene or its known aliases with key words such as osteoporosis, osteoblast, osteoclast and osteocyte using web-based tools such as Pubmed, Google (included Google Scholar), and Endnote X8. For those loci achieved 90th percentile or above confidence, genes identified within them and with literature evidence or positive correlation with BMD in human datasets were considered as strong candidate genes for bone-related genes. Genes within those loci achieved 37th percentile were dealt with caution. They were considered positive if strong literature evidence and/or positive correlation with BMD were present.

Sequence alignment among various species or family members were performed using Uniprot. Dendrogram was also generated with Uniprot. Structures of proteins were built with SMART and/or Prosite MyDomain. Functional domains were highlighted to check whether they are responsible for bone phenotype. Pathways were obtained from KEGG and PathCards. Protein-protein interaction (PPI) mapping of each protein was performed using STRING programme (version 10.5). Network nodes represent proteins; large node indicates a 3D structure is known or predicted, while small nodes represent a protein of unknown 3D structure. In addition, colour nodes indicate query proteins and first shell or interactions, and while nodes indicate second shell of interactions. Edges represent PPIs which are specific and meaningful, and can be divided into three categories, including known interactions, predicted functional PPIs, and others (visit www.string-db.org) (Szklarczyk et al., 2015). The known interactions are derived from curated databases or experiments. Predicted functional interactions include gene neighbourhood, gene fusion and gene cooccurrence. Other PPIs include text-mining, coexpression and homologue.

Building a PPI network for all 86 osteoporosis susceptibility genes identified in our analyses was attempted. Furthermore, to link these 86 candidate genes with 115 known or proven osteoporosis susceptibility genes in homo sapiens (human) was also practiced. In addition, alternative PPI maps were also built in mouse and rat. Besides, integrating key signalling pathways and PPIs was also excised. Sources and tools for such practices included STRING, Cytoscape (http://www.cytoscape.org/), PathPPI (http://proteomeview.hupo.org.cn/PathPPI/PathPPI.html), and PathCards. The latter was mainly used in this study.
4.2.2.5 Correlation with Human Datasets through GWAS

To verify whether any of the candidate genes have a role in bone metabolism in humans, correlation with human bone phenotype data was also attempted, which was performed by Dr Benjamin Mullin, who was one of our collaborators, and has become a colleague at the time of my thesis writing. Human homologues of the mouse candidate genes were identified and interrogated for association with BMD and quantitative ultrasound (QUS) phenotypes in two large genome-wide association study (GWAS) meta-analysis datasets. Testing for association with BMD was performed using the publicly available GEnetic Factors for OSteoporosis (GEFOS) Consortium datasets (2015 release, PMID 26367794), which contain GWAS meta-analysis results for lumbar spine (n = 28,498), femoral neck (n = 32,735) and forearm BMD (n = 8,143) phenotypes (Supplementary Table S4.2) (Zheng et al., 2015). The QUS dataset was generated by Dr Mullin’s group for the phenotypes broadband ultrasound attenuation (BUA, n = 16,433) and velocity of sound (VOS, n = 16,419) using a combination of whole-genome sequence and deeply imputed GWAS genotype data (Supplementary Table S4.3) (Mullin et al., 2017). A gene-based (+/-10Kb) test of association was performed for each gene region in each human dataset using the VEGAS2 software (PMID 25518859), which assigns variants to genes and calculates gene-based empirical association p-values while accounting for the linkage disequilibrium (LD) structure within the gene (Mishra & Macgregor, 2015). Correction for multiple testing in the gene-based and single-point analyses was performed using the Bonferroni method.

4.3 Molecular Biology Techniques

To investigate the RNA expression profile of candidate gene Setbp1, real time RT-PCR experiments were performed with the assistance of Mr Kai Chen and Ziyi Wang (The University of Western Australia). They also helped in isolation and culture of primary cells, RNA extraction, and running the RT-PCR experiments using the ViiA™ 7 Real-Time PCR System (Applied Biosystems).

4.3.1 Isolation and Culture of Primary Cells

4.3.1.1 Isolation and Culture of Bone Marrow Cells

Bone marrow cells were isolated from the long bones of three-month old C57BL/6J mice. The long bones were dissected out from the mice and placed in serum-free α-MEM. The bone marrow was flushed out from the bone marrow cavity with complete α-MEM (α-MEM supplemented with 10% heat inactivated FBS, 2 mM L-glutamine, 100 U/ml Penicillin and 100 µg/ml
Streptomycin) using a 23G needle and filtered through a 100 µm cell strainer. The bone marrow cells were then centrifuged at 1500 rpm/448 × g for 5 minutes, the supernatant was then discarded and the cell pellet was resuspended in complete α-MEM and cultured in 75 cm² tissue culture flasks. All cells were cultured in a 37°C tissue culture incubator with 5 % CO₂.

4.3.1.2 Isolation and Culture of Osteoblasts

Primary osteoblasts are isolated from the long bones of three-month old C57BL/6J mice according to the recommended protocol (Bakker & Klein-Nulend, 2012). The bone marrow was flushed out with 1 × PBS and chopped into small pieces (~ 1 mm³). These pieces were then incubated in collagenase II solution at 37°C with shaking at 200 rpm for 2 hours. An equal amount of complete Dulbecco's Modified Eagle's Medium (DMEM) (DMEM supplemented with 10% heat inactivated FBS, 2 mM L-glutamine, 100 U/ml Penicillin and 100 µg/ml Streptomycin) was added to inhibit further collagenase digestion. The bone pieces were rinsed three times with complete DMEM, then transferred into a 25 cm² tissue culture flask and cultured in complete DMEM. Mature osteoblastic bone cells were observed migrating from the bone pieces after 3 - 5 days. The culture medium was replaced every three days. Bone pieces were maintained for 10 -14 days until migrating cells reached around 70% confluency, when the cells were trypsinised and used for downstream experiments. Bone pieces were discarded.

4.3.1.3 Osteoclast Culture

To culture bone marrow macrophages (BMM), bone marrow cells were cultured in complete α-MEM supplemented with 10 ng/ml M-CSF. Once confluent, the cells were trypsinised and plated in appropriate tissue culture plates for downstream experiments or cryo-preserved for long-term storage.

4.3.1.4 Splitting Primary Cells

To detach cells, the culture media were removed from the flasks and the cell monolayer was washed once with 1 × PBS and incubated with 2 ml of Tryple Express at 37°C for 5 - 20 minutes. A cell scraper was used to further detach cells if necessary. Tryple Express was deactivated by adding complete media and the cell suspension was centrifuged at 1500 rpm/448 × g for 5 minutes to pellet the cells. The supernatant was discarded and the cell pellet was resuspended in fresh complete media. One millilitre of cell suspension was transferred into a new T75 flask containing 9 ml of complete media and incubated in a 37°C CO₂ incubator. The media were completely replaced the following day and the cells were cultured to the appropriate confluence for downstream experiments.
4.3.1.5 Cryopreservation of Primary Cells for Long Term Storage

The cells were detached and pelleted as described above. The cell pellet was resuspended in 92% FBS and 8% DMSO and aliquotted 1 ml per cryovial. The cryovials were stored at -80°C in isopropanol equilibrated cryo-freezing containers for up to a week and transferred to liquid nitrogen for long term storage.

4.3.2 RNA Extraction from Cultured Cells

Total RNA extraction from cultured cells were performed using TRIzol method according to the manufacturer’s guidelines. Briefly, culture media were removed from the cell monolayer and washed once with sterile 1 × PBS. Subsequently, the cells was homogenised by adding TRIzol reagent (1 ml and 0.5 ml in 6 and 12-well plate respectively) directly to the cells and the homogenate was pipetted into 1.5 ml microcentrifuge tubes and ready for RNA extraction. Homogenized samples that were not used immediately for RNA extraction were stored at -80°C until future use.

The homogenised samples were incubated in the TRIzol reagent for 5 minutes at room temperature to allow complete dissociation of nucleoprotein complex. Then chloroform (0.2 ml/l ml TRIzol) reagent was added to each tube and mixed vigorously for 15 seconds. The mixture was then incubated for 2-3 minutes at room temperature before centrifugation at 12,000 g for 15 minutes at 4°C. The transparent upper phase containing RNA was separated from the lower red phenol-chloroform phase. The transparent phase was then carefully pipetted into a nuclease-free 1.5 ml microcentrifuge tube and 0.5 ml of 100% isopropanol was added per 1 ml of TRIzol reagent. The tube with mixture was then incubated at room temperature for 10 minutes, followed by centrifugation at 12,000 g for 10 minutes at 4°C to pellet the RNA. The supernatant was discarded and the RNA pellet was air-dried at room temperature for 10 minutes. Once dried, the pellet was resuspended in 20 µl of nuclease-free water and incubated on a heat block at 55°C for 10 minutes to dissolve the RNA. The RNA was then ready for use or stored at -80°C until future use.

4.3.3 Real Time Reverse Transcription-Polymerase Chain Reaction (RT-PCR)

4.3.3.1 Reverse Transcription and Amplification

The extracted RNA (as described above) was first reverse transcribed to cDNA according to the following protocol:
Reagent | Volume Per Reaction (µl)  
--- | ---  
OligodT (100 µM) | 0.25  
Nuclease-Free water | 5.25  
1 µg RNA diluted in Nuclease-Free water | 10  
Total volume | 15.5

The samples were heated for 3 minutes at 75 °C to denature the RNA and allow the binding of OligodT primers to single stranded RNA. Following which, the following reagents were added to the reaction mixture:

<table>
<thead>
<tr>
<th>Reagent</th>
<th>Volume Per Reaction (µl)</th>
</tr>
</thead>
</table>
dNTP (5 mM) | 2.5 |
5x M-MLV RT buffer | 5 |
M-MLV RT | 1 |
RNasin | 1 |
Total volume | 9.5 |

To amplify the cDNA, the samples were heated at 42 °C for 1 hour followed by heating at 92 °C for 10 minutes. The resulting cDNA was then stored at -20 °C until required.

4.3.3.2 Real Time RT-PCR Setup

Real time-PCR was performed to investigate the gene expression profiles of Setbp1 on osteoblasts and osteoclasts at various time points, including Day 0, 7, 14 and 21 for osteoblasts and Day 0, 1, 3, 5 for osteoclasts. cDNA generated for RT-PCR was diluted 1:6 in nuclease-free water to obtain sufficient quantity for RT-PCR analysis. Each RT-PCR reaction was prepared as follows:

<table>
<thead>
<tr>
<th>Reagent</th>
<th>Volume/Reaction/Well (µl)</th>
</tr>
</thead>
</table>
SYBR Green qPCR Master Mix | 6 |
Forward Primer (10 µM) | 0.3 |
Reverse Primer (10 µM) | 0.3 |
Diluted cDNA in Nuclease-Free water | 5.4 |
Total volume | 12 |
A master mix of the SYBR Green and primers was first prepared. The master mix was mixed thoroughly by vortexing and then loaded into an Optical 384-Well Reaction Plate. Subsequently, the diluted cDNA was loaded into the appropriate wells on top of the SYBR Green-primer master mix. The 384-well plate was then sealed and centrifuged briefly before being placed in the ViiA™ 7 Real-Time PCR System (Applied Biosystems). The real-time-PCR conditions were carried out at 40 cycles of denaturation at 95°C for 15 seconds and annealing at 60°C for 60 seconds. Continuous heating steps of 95°C for 15 seconds, 60°C for 60 sec, and 95°C for 15 seconds were conducted to analyse the melt curve of PCR products. The real time PCR system was programmed as follows:

<table>
<thead>
<tr>
<th>Holding Stage</th>
<th>PCR Stage (×40 cycles)</th>
<th>Melt Curve Stage (Continuous)</th>
</tr>
</thead>
<tbody>
<tr>
<td>50°C for 120 sec, 95°C for 600 sec</td>
<td>95°C for 15 sec, 60°C for 60 sec</td>
<td>95°C for 15 sec, 60°C for 60 sec, 95°C for 15 sec</td>
</tr>
</tbody>
</table>

### 4.3.3.3 Real Time-PCR Analysis

The real time-PCR data (C\textsubscript{T} values) was collected from the ViiA™ 7 software (Applied Biosystems). Melting curve for the PCR products were analysed to verify the specificity of the PCR primers. The normalised fold expression (Fold change) of genes was obtained relative to the housekeeping gene *Hprt* and other genes like *Ocn* and *Trap*. The method for fold change was applied the Livak’s ΔΔC\textsubscript{T} equation (Livak & Schmittgen, 2001). The adapted equation is showed as below:

\[
\text{Fold change} = 2^{-\Delta\Delta C_T}
\]

where \( C_T \) = Mean of \( C_T \) triplicate

\[\Delta C_T = C_T \text{ of Setbp1 of Sample A (Test)} - C_T \text{ Hprt}\]

\[\Delta\Delta C_T = \Delta C_T \text{ of Setbp1 of Sample A (Test)} - \Delta C_T \text{ of Setbp1 of Sample B (Control)}\]

The mean Ct value of *Setbp1* in the experimental group was normalised to the Ct value of *Hprt* to give a ΔCt value, which was further normalised to control samples to obtain ΔΔCt. Three independent cultures were carried out, and osteogenic experiments were performed in triplicate, while experiments of osteoclastogenesis were performed twice.
4.4 Statistical Analysis

Data were entered into Microsoft Excel and statistical analyses conducted using Excel. Figures were generated either with Excel or Graphpad® Prism (version 6.01). Pearson’s correlation plots were applied to analyse the correlations between BV/TV and other traits, including Tb. N, Tb. Th, Tb. Sp, Conn. D, SMI, DA and Ct. Th.

HAPPY software was used to estimate the founder haplotype having the maximum-likelihood probability for genotype sets of allele 1 and 2 separately (Ram et al., 2014). Thresholds for haplotype mapping were computed from 1,000 additive model permutations under the approach recommended by Durrant et al. (Durrant et al., 2011). The founder strain(s) contributing to each trait were determined by deriving coefficients (log odds ratio) of the fit from the logistic/multinomial regression model and using plotting tools implemented in the DOQTL R package (Gatti et al., 2014).

The gene-based (+/- 10Kb) test of association was performed for each gene region in each human dataset using the VEGAS2 software (Mishra & Macgregor, 2015). Significance thresholds used in human populations included Bonferroni-corrected significance threshold for the gene-wide testing is 0.05/188 = 2.66 × 10^{-4} which was based on 188 testing genes. The gene-wide suggestive threshold was set as 0.001. The single-point suggestive threshold was set as 3.00 × 10^{-5}.

For real time RT-PCR experiments, all values are presented as the mean ± SD of the values obtained from two or three experiments. Statistical significance was determined by Student's t test. A p value of less than 0.05 was considered to be significant.
Chapter 5
Characterisation of the Bone Phenotype of CC Mice using Micro Computed Tomography (Micro-CT)
5.1 Introduction

The CC mice are derived from eight carefully selected genetically diverse inbred strains, including five classical inbred strains (A/J, C57BL/6J, 129S1/SvImJ, NOD/LtJ, NZO/H1LtJ) and three wild-derived subspecies (CAST/EiJ, PWK/PhJ, and WSB/EiJ), in order to achieve high levels of variation distributed uniformly throughout the mouse genomes (Churchill et al., 2004). The CC gives each line a unique name and these lines have genetic compositions randomly inherited from six to eight founder strains. Investigating the correlation between phenotype and genotype has been proven to be a powerful way to identify genes responsible for traits like bone mass or bone microarchitecture (Levy et al., 2015).

The CC mice present with a wide range of phenotypes that can be characterised and then the genetic components affecting the traits can be analysed. As mentioned before, there is only one study that utilised a cohort of CC mice to identify genes regulating bone microarchitecture published during the CC mice screening process of the current study (Levy et al., 2015). The published study was the first study in the field of osteoporosis and identified several genes via analysing traits including BV/TV, Tb. N, Tb. Th and Conn. D. It provides a proof-of-concept that the CC is a powerful tool in the identification of genes that are responsible for complex traits or diseases. However, limitations such as small sample size and narrow age range resulted in suboptimal identification of candidate genes responsible for bone microarchitecture. Our study was planned to look at the bone microarchitecture in a larger population of CC mice.

In order to characterise microarchitectural traits, µCT scanning on hindlimbs of CC mice was performed, and analysis of the bone phenotype was completed. The traits that were analysed included essential parameters such as BV/TV, Tb. N, Tb. Th and Tb. Sp for bone mass and additional parameters including Conn. D, SMI and DA for bone microarchitecture in trabecular bone, and a single parameter Ct. Th for cortical bone (Bouxsein et al., 2010). In the current study, totally 64 strains were included; each of these strains had at least 3 mice (see Chapter 4, Supplemental Table S4.1). This cohort consisted of 848 CC mice, including 278 males in 60 strains, and 570 females in 61 strains.

Analyses of various parameters/traits were performed in each CC mouse under the guidelines for assessment of bone mass and bone microarchitecture using microCT (Bouxsein et al., 2010). To find out the phenotypic distribution across lines, we ranked CC lines in descending order. We categorised these lines into four groups, overall (percentage changes with comparison of C57BL/6J or stratification with various age intervals, young (2-7 months), middle aged (7.1-11.9 months), and old (12 months and above). The overall group included all strains with various ages and sample size.
The reference values for age and sex matched C57BL/6J mice (2, 3, 4, 6, 12, and 20 month of age; n = 6–9 per age and per sex) were extracted from previously published data, and calibrated with our data from 3 month old and 6 month old C57BL/6J control mice (Glatt, Canalis, Stadmeyer, & Bouxsein, 2007; Halloran et al., 2002) (Supplementary Table S5.1). References using C57BL/6J as control were applied in traits including BV/TV, Tb. N, Tb. Th, Conn. D, SMI and Ct.Th. For those traits where age matched references were not available strains were stratified by mean values at three different age groups. This “self-reference” approach was used in Tb. Sp and DA.

5.2 μCT results

5.2.1 Analyses of Trabecular Traits of the CC population Using μCT

Trabecular phenotypes were approximately normally distributed

Our analyses showed that all trabecular phenotypes were approximately normally distributed in both sexes. This normal (Gaussian) distribution pattern reflects the large phenotypic variations across the CC population, from overall groups, to young, middle aged and the older cohorts. BV/TV and Tb. N appear to be more normally distributed than other traits in three separate age groups.

Trabecular traits varied across the CC population

Our analyses demonstrated large variations in the seven trabecular traits analysed in this CC cohort. BV/TV, the most important indicator for bone mass was first analysed, and results showed large variations in both sexes (Figure 5.1). We defined those percentage changes within two standard deviations of the C57BL/6J as normal, while those above as osteopetrotic and below as osteoporotic (see Supplementary Table S5.2 & S5.3). This definition was used to classify this CC cohort into three groups: osteopetrotic, normal, and osteoporotic. Most strains fell within the normal category. Several strains showed a consistent high BV/TV in both sexes, including strains VIT, DONNELL, GIG and SOLDIER, while other strains such as PER2, CIV2, GALASUPREME, WAB2 and FIM showed a consistent low BV/TV. However, several strains presented significant differences between sexes. For example, JAFFA was shown to be osteopetrotic (1,236% change vs C57BL/6J) in female mice, while it presented an osteoporotic phenotype (17% change vs C57BL/6J) in males. GAV is another example; it was classified as osteopetrotic (948% change vs C57BL/6J) in female, but its BV/TV was within normal range (80% change vs C57BL/6J) in male. The overall difference in BV/TV within strains and between sexes is shown in Figure 5.1. As we can see from the 3D reconstructions of μCT acquisitions, large variations existed in bone volume within CC strains at a single time point (6 months). DONNELL and CIV2 were selected to represent osteopetrotic and osteoporotic phenotype, respectively. These 3D images represent not only the large variations in bone mass, but also in bone microarchitecture.
Furthermore, large variations also existed in other trabecular traits including Tb. N (Figure 5.2), whereas Tb. Th (Figure 5.3) and Tb. Sp (Figure 5.4) were less variable. When looking at architectural parameters Conn. D (Figure 5.5), and SMI (Figure 5.6) showed large variability, whilst DA values (Figure 5.7) were far less diverse. Interestingly, some outlier strains existed within specific traits. For instance, female mice within the JAFFA strain had an extremely high Conn. D; they also showed a large degree of variability as demonstrated by the extremely large standard deviation, making it difficult to assess this strain for genetic contributions to Conn. D. Exclusion of such outliers was performed to enhance mapping of the genes underpinning the trait.

Trabecular traits also presented variations in separate age groups. In this study four traits, BV/TV, Conn. D, SMI and DA, showed similar distribution patterns across the age groupings as they did when all age groups were analysed together (Supplementary Figures S5.1-5.6). When considering sex variations, large differences were observed in BV/TV, Conn. D and DA, whereas SMI demonstrated as a relatively constant trait across ages in both sexes. Furthermore, CIV2 appeared to be the most consistent outlier across ages and sexes in our analysis. JAFFA, which had a limited sample size in the old female cohort, also demonstrated higher values in BV/TV and Conn. D than other strains (Supplementary Figure S5.3). In particular, JAFFA had a high Conn. D score (298), which is nearly four times greater than the nearest strain, SOLDIER, which had a Conn. D score of 71.
Figure 5.1 Trabecular bone volume/tissue volume (BV/TV) (%) in distal femur varies between strains of CC mice and between sexes. BV/TV was compared to age and sex matched control C57BL/6 values and a % of control value was generated for each strain. This allows comparison of mouse strains across differing age ranges and sexes. (A) Percentage change in BV/TV of CC strains (C57BL/6J strain = 100%, red line) in female mice. (B) Percentage change in BV/TV of CC strains (C57BL/6J strain = 100%) in male mice. (C) Representative 3D reconstructions of the trabecular bone compartment of 6 month old male and female osteoporotic (CIV2), osteopetrotic (DONNELL), and C57BL/6J (normal) mice showing the large variations in bone volume within CC strains.
Figure 5.2 Trabecular number (Tb. N) (1/mm) in distal femur varies between strains of CC mice and between sexes.
Figure 5.3 Trabecular thickness (Tb. Th) (mm) in distal femur varies between strains of CC mice and between sexes.
Figure 5.4 Changes in trabecular Separation (Tb. Sp) (mm) in distal femur in CC mice in comparison with stratified reference.
Figure 5.5 Changes in connectivity density (Conn. D) (1/mm3) in distal femur in CC mice in comparison with C57BL/6J.
Figure 5.6 Changes in structural model index (SMI) in distal femur in CC mice in comparison with stratified reference.
Figure 5.7 Changes in degree of anisotropy (DA) ($\alpha$) in distal femur in CC mice in comparison with C57BL/6J.
Correlations between trabecular traits

Analyses in the current study also demonstrated correlations between trabecular traits. Pearson’s correlation ($R^2$) plots showed BV/TV in females was positively correlated with Tb. N, Tb. Th and Conn. D with coefficient ($r$) of 0.89, 0.15 and 0.58, respectively (Figure 5.8), while it was negatively correlated with Tb. Sp, SMI and DA with coefficient of 0.48, 0.33 and 0.11, respectively (Figure 5.9). In males, Pearson’s correlation plots also showed similar correlations, in which BV/TV was positively correlated with Tb. N, Tb. Th and Conn. D with coefficient of 0.89, 0.20, and 0.9, respectively (Figure 5.10); and negatively correlated with Tb. Sp, SMI and DA with coefficient of 0.56, 0.68 and 0.06, respectively (Figure 5.11). Tb. N, therefore appears to be the most consistent correlative trait with BV/TV in both sexes. Interestingly, Conn. D in both sexes also presented a high positive coefficient with BV/TV (Figure 5.8). Moreover, Tb. Sp and SMI demonstrated higher negative coefficient with BV/TV in males than in females (0.56 vs 0.48, 0.68 vs 0.33, respectively). The remaining traits, including Tb. Th and DA, appear to be less correlated with BV/TV in both sexes, showing coefficients ranged from 0.06 to 0.2. The correlation between traits in individual age groups was also determined and is presented in supplementary figures S5.7-S5.18), these showed similar trends to the combined age groups. The correlation between DA and BV/TV was low in all analyses, suggesting that DA, as the key indicator for bone microarchitecture does not highly correlate with bone mass.
Figure 5.8 Positive correlation between BV/TV and Tb. N, Tb. Th and Conn. D in females. (A) Pearson’s correlation ($R^2$) plots showed coefficient ($r$) of 0.89 for BV/TV vs. Tb.N. (B) Pearson’s correlation plots showed coefficient ($r$) of 0.15 for BV/TV vs. Tb. Th. (C) Pearson’s correlation plots showed coefficient ($r$) of 0.58 for BV/TV vs. Conn. D. JAFFA presented as a outlier in Conn. D with significant higher score in % change of C57BL/6J.
Figure 5.9 Negative correlation between BV/TV and Tb. Sp, SMI and DA in females. (A) Pearson’s correlation plots showed coefficient \( r \) of 0.48 for BV/TV vs. Tb. Th. (B) Pearson’s correlation plots showed coefficient \( r \) of 0.33 for BV/TV vs. SMI. (C) Pearson’s correlation plots showed coefficient \( r \) of 0.11 for BV/TV vs. DA. CIV2 presented as an outlier for SMI and DA. CIV2 presented as an outlier for SMI and DA, and ROGAN was for DA only.
Figure 5.10 Positive correlation between BV/TV and Tb. N, Tb. Th and Conn. D in males. (A) Pearson’s correlation plots showed coefficient ($r$) of 0.89 for BV/TV vs. Tb.N. (B) Pearson’s correlation plots showed coefficient ($r$) of 0.20 for BV/TV vs. Tb. Th. (C) Pearson’s correlation plots showed coefficient ($r$) of 0.9 for BV/TV vs. Conn. D.
Figure 5.11 Negative correlation between BV/TV and Tb. N, Tb. Th and Conn. D in males. (A) Pearson’s correlation plots showed coefficient (r) of 0.56 for BV/TV vs. Tb. Th. (B) Pearson’s correlation plots showed coefficient (r) of 0.68 for BV/TV vs. SMI. (C) Pearson’s correlation plots showed coefficient (r) of 0.06 for BV/TV vs. DA. CIV2 presented an outlier in both SMI and DA.
Age dependence of traits

Bone mass and bone microarchitecture change over time. In our analyses, diverse patterns of change were observed, although most strains demonstrated bone loss over time. We compared the young (2-7 months) and old (12 months and over) female CC mice (n = 23 strains), and showed bone loss over time in all strains except JAFFA and REV (Figure 5.12). Most strains showed bone loss over 12 months old, consistent with what is observed in the reference C57BL/6J strain; however the rate of bone loss was highly variable. For example, MOP had the most significant direct reduction, a drop of 22.5% in measured BV/TV, representing a relative loss of 83.33% of BV/TV over time. Several strains including BEW, ROGAN and CIV2 showed stable bone volume across ages. Interestingly, JAFFA and REV presented as two divergent strains which gain bone even after 12 months. In comparison of young and old animals, JAFFA had a 15.15% increase in measured BV/TV, representing a 119.3% change in relative BV/TV over time. In the REV strain a small direct gain of bone was observed (1.45%), against a background of very low BV/TV, resulting in a large relative change (263.6%). The directly measured increase/decrease in BV/TV appears to be more accurate in determining the change of bone mass because the percentage change may be prominent if the baseline BV/TV is minimal.

Comparison of young and old male CC mice (n = 20 strains) was also performed. As shown in Figure 5.13, several strains, including WAD, CIV2, PEF2 and POH, gained bone over time, but all with minimal increased levels in measured BV/TV. For instance, WAD showed the highest net increase (3.1%), which results in a relative change of 100% in BV/TV. Bone loss was predominately observed across most strains. Strains including CIS, POT, TOFU, ROGAN, CIS2, JUNIOR, JAFFA, WAB2, GALASUPREME and MOP showed at least 50% bone loss over 12 months. For example, JUNIOR lost 16% in measured BV/TV representing a relative loss of 80% of bone volume over the course of a 12 month period. These strains presented more abrupt bone loss than male C57BL/6J mice which showed an relative 22% change over nearly 20 months (Glatt et al., 2007).

In summary, both female and male CC mice presented various bone change patterns. Unlike C57BL/6J mice which have been reported to lose bone during the ageing process, some CC strains were found to gain bone over time. Furthermore, while most strains had bone loss, some showed much greater rates of change, indicating the diversity of phenotypes captured in this screen.
Figure 5.12 **Change of BV/TV over time in female CC mice.** Comparison of young and older CC mice (n = 23). (A) Change in BV/TV (%) across strains. Positive number refers to bone gain, and negative number indicates bone loss. JAFFA and REV demonstrated bone gain, other strains (e.g. MOP, JUNIOR) presented bone loss. (B) % change in BV/TV across strains. REV presented much more change than JAFFA in bone mass.
Figure 5.13 Change of BV/TV over time in male CC mice. Comparison of young and old CC mice (n = 20 strains). (A) Change in BV/TV (%) across strains. WAD, PEF2, CIV2 and POH demonstrated bone gain, other strains (e.g MOP, JUNIOR) presented bone loss. (B) % change in BV/TV across strains. MOP presented most significant percentage change, closely followed by GALASUPREME and WAB2.
5.2.2 Cortical Bone Diversity in CC Mice

Cortical bone thickness (Ct.Th) was also assessed in the present study. Based on Figure 5.14, cortical bone variation was less pronounced than that observed in the trabecular parameters for both male and female cohorts, showing percentage values ranging from 90-140 versus C57BL/6J (set as 100%). In separate age groups, small variations were also observed (Supplementary Figures S5.19, S5.20). No outlier for Ct. Th was identified in these cohorts.

Correlation between BV/TV and Ct. Th was assessed, showing 0.05 in female and 0.003 in male (Figure 5.15). Pearson’s plots in separate age groups also showed nearly zero correlation between these two traits (Supplementary Figure S5.21 & S5.21). These low correlation scores indicate these two traits are poorly correlated, indicating differential regulation of these traits.

Ct. Th change during the ageing process was also examined. Totally 23 and 19 strains with young and older mice were included in female and male cohorts, respectively (Figure 5.16 & 5.17). Strains with a bone loss phenotype slightly outnumbered those with bone gain phenotype in both sexes. The percentage changes in female and male cohort ranged from 0-35% and 0-17%, respectively. These results show that CC mice have a smaller degree of change in Ct. Th than in BV/TV in both sexes. In comparison with the published data, which reported an 8% change in female mice and 22% change in male C57BL/6J mice over their lifetime (Glatt et al., 2007), female CC mice presented significantly larger variations in cortical bone change, while interestingly male animals presented as slightly less variable in cortical phenotype than C57BL/6J counterparts.
Figure 5.14 Changes in cortical thickness (Ct.Th mm) in distal femur in CC mice in comparison with C57BL/6J. (A) Percentage change in Ct.Th of CC strains versus C57BL/6J strain (100%, red line) in female mice. (B) Percentage change in Ct.Th of CC strains versus C57BL/6J strain (100%) in male mice. (C) Representative 3D reconstructions of the cortical bone compartment of 6 month old male and female CIV2, DONNELL, and C57BL/6J normal mice showing the variations in cortical bone within CC strains.
Figure 5.15 Correlation between BV/TV and Ct. Th in CC mice. (A) Pearson’s plot shows a correlation score of 0.05 in female cohort. (B) Pearson’s plot shows a correlation score of 0.003 in male cohort.
Figure 5.16 Change of Ct. Th over time in female CC mice. Comparison of young and older CC mice (n = 23 strains). (A) Change in Ct. Th (mm) across strains. (B) % change in Ct. Th across strains.
Figure 5.17 Change of Ct. Th over time in male CC mice. Comparison of young and older CC mice (n = 19 strains). (A) Change in Ct. Th (mm) across strains. (B) % change in Ct. Th across strains.
5.3 Discussion

We have characterised both trabecular and cortical bone phenotypes of CC mice using µCT. Seven trabecular traits and one cortical trait were included. The ages of this cohort ranged from 2.5 to 23.5 weeks, which is similar to the spectrum of C57BL/6J mice analysed by other researchers (Glatt et al., 2007; Halloran et al., 2002), but is much wider than those CC mice analysed by Levy et al. (Levy et al., 2015). Larger age variations allow us to assess various bone traits comprehensively and bone changes in the ageing process. This study has many advantages over the first published CC study in bone in many aspects, including larger sample size (848 vs 160), more strains (64 vs 31), and age range (2.5-23.5 months vs 10-13 weeks). The narrow age range is insufficient to represent the natural population which includes not only the middle aged, but also the young and aged subjects. Analysis on a particular age group will loss the insights of age-dependence of traits. However, sample size in our study remains relative small if we compare most GWAS studies in human populations. Theoretically, the bigger sample size (more lines and bigger animal numbers) we have, the more detailed phenotypic distribution will have, and further more precise prediction on the identification of genes responsible for those phenotypes will be achieved (Valdar et al., 2006). In other words, to recruit more mice and lines in the current cohort will not only enrich characterisation of CC populations, but also provide a better prediction in the process of haplotype gene mapping.

Our results showed large variations among traits in this CC cohort, in particular in BV/TV, Tb. N, Conn. D and DA. Distribution of CC strains in various traits followed the approximate Gaussian pattern, indicating a natural distribution across strains with existence of outliers such as CIV2 and ROGAN. This pattern is in agreement with the pattern found in 31 lines in Levy et al’s study (Levy et al., 2015). Given several strains presented consistent either osteopetrotic or osteoporotic phenotype across sexes, a few strains such as JAFFA and GAV showed differences between female and male. We speculate that such difference may be determined by sex and genetic factors (Zanotti, Kalajzic, Aguila, & Canalis, 2014). For example, sex hormones such as androgen and oestrogen may play key roles in determining bone mass in these specific strains (De Oliveira, Fighera, Bianchet, Kulak, & Kulak, 2012; Manolagas et al., 2013). Measuring the levels of these hormones in live CC mice may help reveal the correlation between bone mass and hormone levels. Moreover, tracking the hormone levels in specific strains like JAFFA and GAV along with the ageing process may be valuable in the determination of the roles of oestrogen/androgen in age-related bone remodelling.

Our analyses also revealed various correlations between BV/TV and other traits. Significant positive correlation was found in Tb. N and Conn. D, while obvious inverse correlation was shown in Tb. Sp and SMI. Other traits such as Tb. Th, DA and Ct. Th were found to be poorly correlated with BV/TV. BV/TV, as the major indicator for bone mass, has been correlated with Tb. N, Conn. D and Tb. Th.
Negative correlation between BV/TV and Tb. Sp or SMI provides another advantage over the published study which did not include those traits with inverse correlation.

The close positive correlation between BV/TV and Tb. N was demonstrated in the current study, showing Pearson’s correlation coefficient ranged from 0.89 in both sexes in the overall group (Figure 5.8 & 5.10), to striking 0.99 in old female cohort (Supplementary Figure S5.11). The discrepancy between overall group and separate age groups because of the different analysis approaches. In this case, analyses using direct measured numbers of both BV/TV and Tb. N appears to be more accurate due to the consistent high coefficient numbers in separate age groups. However, our results are close to those published studied showing coefficient over 0.9 in mouse models (Levy et al., 2015; Parkinson & Fazzalari, 2003). BV/TV was also shown positive correlation with Conn. D, showing consistent high coefficient numbers in four male groups (ranged from 0.9 - 0.93), and various coefficient in female cohorts (ranged from 0.58 - 0.91). The larger variations in females than males may be explained by the existence of larger variations of these two traits in female rather than male CC mice. In comparison with another CC study, in which showing coefficient $r = 0.79$ in combined cohort with both sexes, our results also present similar correlation pattern between BV/TV and Conn. D.

Obvious negative correlation was found between BV/TV and Tb. Sp in the current study. Our results are not striking as that identified in C57BL/6J cohort ($r = -0.95$), indicating greater variations of CC over C57BL/6J mice (Parkinson & Fazzalari, 2003). Unfortunately, comparison within CC populations cannot be performed due to the missing analyses in the publication (Levy et al., 2015). In addition, our analyses also demonstrated the negative correlation between BV/TV and SMI, showing coefficient ranged from 0.33 - 0.7 with the exception in old female cohort (Supplementary Figure S5.12). Again, no comparison was made thanks to the under-report of this trait in other studies (Halloran et al., 2002; Levy et al., 2015; Parkinson & Fazzalari, 2003).

Interestingly, poor correlation between BV/TV and Ct. Th implies no strong correlation between trabecular and cortical phenotypes. However, contrary to this general correlation, several strains, such as PEF2 and MOP, did present osteoporotic trabecular phenotypes with thick cortical bones. Such observation provides partial support for the hypothesis that cortical thickness may increase in a circumstance with relatively lower trabecular bone mass via undefined adaption mechanisms (Ruimerman, Hilbers, van Rietbergen, & Huiskes, 2005). Increasing the number of strains and samples analysed will allow further clarification of this correlation.

We analysed the percentage change and found both trabecular and cortical bones change over time. Our results suggest that most CC strains are basically agreed with other researchers’ work, in which bone mass has been decreased with age in both mice and humans (H. Chen, Zhou, Fujita, Onozuka,
In trabecular bone, CC mice demonstrated larger variations during the ageing of bone than C57BL/6J mice. This phenomenon is reasonable because of the large number of strains, and genetic diversity of the founder strains (Churchill et al., 2004; Iraqi, Churchill, & Mott, 2008). Furthermore, unlike C57BL/6J mice which showed a decline at 12 months in both sexes (Glatt et al., 2007), some CC strains were found to increase their bone volume after 12 months. We speculate that these strains may show delays in achieving peak bone mass, which could be part of the constitutional delay of pubertal maturation (Gilsanz et al., 2011; Moreira-Andres, Canizo, de la Cruz, Gomez-de la Camara, & Hawkins, 2000; Yap et al., 2004). To investigate whether phenotypes of delayed puberty occur may provide insights of the correlation between constitutional growth and bone formation within those strains. We may have an opportunity for identifying genes that are responsible for constitutional and bone growth, and understanding how the timing and progression of puberty are controlled (Kumar & Boehm, 2013; Mayer et al., 2010).

In the cortical bone, CC mice also presented different age-dependent phenotypes and the percentage changes in females are greater than those in males. In comparison with C57BL/6J mice, of which Ct. Th has been shown percentage change over life by 19% in distal femur and 22% in lumbar spine in males, whereas -4% in distal femur and 8% in lumbar spine in females, CC mice presented greater changes in Ct. Th (Glatt et al., 2007). Bone gain in cortical area, in particular in males may be partially explained that bone loss in trabecular bone could be compensated by the increase of cortical bone thickness to accommodate weight bearing, and perhaps the timing of “peak” bone mass acquisition differs for trabecular and cortical compartments (Glatt et al., 2007). Therefore, in summary, female CC mice vary to a higher extent than C57BL/6J mice in age-related changes in trabecular and cortical bone. Furthermore, sex difference of bone changes within CC mice was observed, which may be caused by a variety of factors such as genetic, oestrogen/androgen level, activity (Amin & Khosla, 2012; Callewaert, Sinnesael, Gielen, Boonen, & Vanderschueren, 2010; Zanotti et al., 2014). These factors may contribute to bone change phenotypes through modifying the balance of osteoclast and osteoblast activities (Roshan-Ghias, Lambers, Gholam-Rezaee, Muller, & Pioletti, 2011; Zanotti et al., 2014). Our results indicate that trabecular bone volume declines more rapidly in many female than in male strains possibly due to higher bone resorption in female mice than in male littermates, of which similar change pattern has been shown in C57BL/6J mice (Glatt et al., 2007; Zanotti et al., 2014). Oestrogen and/or androgen levels may contribute such difference because both sex hormones are potent regulators for bone mass and strength (Manolagas et al., 2013; Wiren, 2005). Oestrogen is involved in osteoblast differentiation by upregulating BMP4 signalling (Matsumoto et al., 2013), and inhibits bone resorption by directly inducing apoptosis of the bone-resorbing osteoclasts by an oestrogen receptor-mediated mechanism (Kameda et al., 1997). Androgen increases bone mass
through enhancing osteoblast activity but inhibiting that of osteoclasts (Sinnesael, Boonen, Claessens, Gielen, & Vanderschueren, 2011; Wiren, 2005). Therefore, female animals are more vulnerable than males towards to abrupt changes in oestrogen levels (Migliaccio, Newbold, Mclachlan, & Korach, 1995; Parikka et al., 2005). Also, the higher androgen levels in males may provide partial protection from age-related decline in trabecular bone (De Oliveira et al., 2012; Miyaura et al., 2001; Wiren, 2005).

There are both advantages and limitations in the current study. In comparison with the published CC paper, we have analysed a bigger sample size and wider age range. However, limitations are also obvious. Firstly, the sample size remains relatively small compared to most GWAS studies in human populations, which could have thousands even up to 100k subjects (Estrada et al., 2012; Mullin et al., 2016; Rivadeneira et al., 2009). The average values of those strains with a small sample size (e.g., RAE2, ZAQ), in particular, are less reliable than those with a larger number (e.g, GALASUPREME, GIG). Secondly, many strains lacked samples across all ages. Therefore, it was impossible to compare bone change over time within those strains. Due to limitations in mouse availability the age categories were quite broad, changes in bone mass within the young category particularly may have resulted in the greater variability in this age group. Thirdly, we did not investigate the correlation between bone mass and other traits such as body mass and body length due to the limited access of fresh or live animals. We may miss many important aspects of this CC cohort. For instance, body mass has been positively associated with BMD, and low body mass is considered as a risk factor for osteoporosis (Jiang et al., 2015; S. J. Kim, Yang, Cho, & Park, 2012; Langsetmo et al., 2012; Zhu, Hunter, James, Lim, & Walsh, 2015). Investigating the relation between bone mass and body mass may identify genes regulating both traits correlatively. To overcome such limitations, more strains and mice with various ages are required, which not only allows us to divide the CC mice (in particular those younger than 7 months old) more precisely in different age groups based on their growth patterns, but also increase the statistical power to analyse the difference between age groups. Furthermore, to obtain corresponding data such as body mass and body length will be beneficial in finding genes that regulate bone and other phenotypes, but this direction is out of the scope of this project.

Given sufficient strains and big enough sample size may be achieved, a critical question may emerge: are the CC mice a suitable population for investigating osteoporosis in human? To answer this fundamental question, several lines of evidence could be addressed, and probably provide a definite answer. It has been well-recognised that most mammals exhibit skeletal deterioration with advancing age (Jilka, 2013), and aged-related bone loss has been well-documented in mouse (in particular C57BL/6J) (Almeida et al., 2007; Ferguson, Ayers, Bateman, & Simske, 2003; Glatt et al., 2007; Halloran et al., 2002). C57BL/6J, as one of the laboratory mouse (Mus musculus), is the most
commonly used model for studying human physiology and disease at molecular, cellular and genetic levels due to few key aspects, including fast breeding (3-week gestation period), high reproductive capacity (~5 - 6 animals/litter), similarity to human physiology and genome, easy controlled housing environment, and most importantly ability for genetic manipulation (Jilka, 2013). C57BL/6J mice exhibit many similar physiological aspects with humans, of which such as loss of trabecular bone, thinning of cortical bone, and increased cortical porosity with advancing age (H. Chen et al., 2013; H. Chen, Zhou, Shoumura, Emura, & Bunai, 2010; L. Wang et al., 2017). More importantly, the ability for gene manipulation has offered a platform to address fundamental questions underlying the regulation of bone cells formation, function and interplay (Davey, MacLean, McManus, Findlay, & Zajac, 2004). CC strains, which inherited genetic information from five laboratory mouse (including C57BL/6J), could share many similar advantages with mouse over human or other mammals in the study of bone biology (Percival et al., 2015; Threadgill, Miller, Churchill, & de Villena, 2011). Furthermore, the success of studies utilised CC mice has added confidence over using this platform in complex diseases such as osteoporosis and diabetes (Abu-Toamih Atamni et al., 2017; Atamni et al., 2016; Levy et al., 2015; Mao et al., 2015).

In particular, CC library may have strains which resemble age-related bone changes in humans better than C57BL/6J mice. This assumption is based on the fact that C57BL/6J mouse may not perfectly reflect the age-related bone loss in human due to its skeleton continues to grow gradually after puberty (6-9 weeks) (Glatt et al., 2007; Halloran et al., 2002), while BMD in humans peaks between the ages of 10 and 19 years and starts to decline with further ageing (Bonjour, Theintz, Buchs, Slosman, & Rizzoli, 1991; H. Chen et al., 2013; H. Chen et al., 2010; L. Wang et al., 2017; W. Yu et al., 1999). The difference between C57BL/6J mice and humans in age-related bone loss may leave the C57BL/6J mouse models less desirable given they have been heavily utilised in bone study. Therefore, applying murine models with similar trend with humans in both sexes in the study of bone loss during ageing process is warranted. In the present study, CC mice present large variations in age-related bone degradation. Unfortunately, we did not find any strain that perfectly match the trend of bone change in humans. With more strains and mice are included in future analyses, such strains may be pulled out.

In summary, CC mice could be a powerful platform in identifying genes responsible for bone mass and bone microarchitecture based on the advantages inherited from mouse models and large variations in bone phenotypes. Besides, specific strains which resemble the age-related degradation of bone in humans may be generated and applied in future study.
Chapter 6
Quantitative Trait Locus Haplotype Mapping of Various Traits in the Collaborative Cross Mice Screening
6.1 Introduction

The CC mice demonstrate a wide range of variation in attributes such as coat colour, body weight, body length, behaviour, and as demonstrated in Chapter 5, bone microarchitecture (Collaborative Cross, 2012). Over 38 million SNPs and indels have been identified among the founder strains, giving diverse genetic coverage within the CC (Munger et al., 2014). Many studies have demonstrated proof-of-principle for applying the CC resource for identification of genes responsible for traits of interest, initially using pre-CC mice (Bottomly et al., 2012; Durrant et al., 2011; Thaisz et al., 2012), and recently using fully genotyped CC strains (Levy et al., 2015; Ram et al., 2014; Rogala et al., 2014; Vered, Durrant, Mott, & Iraqi, 2014). Application of pre-CC and CC mice varied, from coat colour traits, host susceptibility to pathogen infection, to genetic analysis of albuminuria. During the period of our study, there was only one pilot study demonstrating the association between genes and microarchitecture of trabecular bone (Levy et al., 2015).

In the present study, CC mice (n = 848; strain n = 61) were categorised into four groups; overall (percentage changes with comparison of C57BL/6J or stratification with various age intervals), young (2-7 months), middle aged (7.1-11.9 months), and old (12 months and above). For these groups genetic associations for 8 different traits were analysed; BV/TV, Tb. N, Tb. Th, Tb. Sp, Conn.D, SMI, DA, and Ct. Th.

SNP haplotype mapping was initially applied as a supplementary approach to QTL haplotype mapping. We applied the updated version of GeneMiner (updated on 19/05/2016) that included many GigaMuga array SNPs (120K SNPs) to perform QTL scans. However, the databases for SNP based scan was not yet updated at the time of haplotype mapping. Most strains were still tested with the MegaMuga array (77,000 SNPs), with only a few strains typed in GigaMuga. The inequality of these two databases, therefore, provided many inconsistent mapping results. To avoid this problem we focused on haplotype mapping using QTL haplotype scanning alone.

Unlike GWAS studies, where a fixed P-value threshold (such as $5 \times 10^{-8}$) is usually applied to decide whether the variant is replicated or “meaningful”, the thresholds for QTL scans vary, mainly dependent on the sample size. We utilised three thresholds, which were set by Prof Grant Morahan’s team based on the approach recommended by Durrant et al. (Durrant et al., 2011). They were represented by different colours; the red line represents the 95th percentile that is equivalent to a P-value < 0.05 for GWS, the blue line represents the 90th percentile (p < 0.1), and the orange line represents GWS above the 37th percentile (p < 0.63). In other words, loci with peaks above the red line are considered most reliable. The most reliable QTL would have a low P-value (< 0.05), high LOD score, narrow genomic interval and with dominant founder strain(s). Loci only achieving GWS
above the 37th percentile are relatively unreliable (Ram et al., 2014). The genes harboured at loci below the red line are suggestive for significance, but should be interpreted with great caution.

SNPs and genes identified for each trait varied. Details of them are available on the GeneMiner® database via inputting chromosomal interval and clicking the founder strain. Most important information, however, will be presented in the texts and figures in this chapter.

6.2 Results

6.2.1 Trabecular Traits Vary Across the CC Population

As described in Chapter 5, for each CC mouse we analysed the trabecular bone compartment of the distal femoral metaphysis and measured BV/TV, Tb. N, Tb. Th, Tb. Sp, Conn. D, SMI and DA. The first four traits represent trabecular bone mass, and the remainder provide indices relating to the microstructure of trabecular bone. There was a high phenotypic heterogeneity between strains in both females and males.

6.2.1.1 QTL Haplotype Mapping of BV/TV

BV/TV is a strong trait for bone mass

BV/TV is the most important indicator for trabecular bone mass in the µCT analysis. Haplotype QTL mapping was performed over four categories, including overall (percentage changes in comparison with C57BL/6J), young (2-7 months), middle aged (7.1-11.9 months), and old (12 months and above), and in separate sexes.

Loci derived from BV/TV were mostly located on chr4, chr5 and chr11 (Figure 6.1 & 6.2). There was no peak that achieved GWS with 95th percentile of confidence except in the old female cohort (Figure 6.2). However, no gene was pulled out either from the ENCODE or FANTOM5 database. Instead, several genes such as Dnah2, Pfas, Nlrp1a and Nlrp1b were identified from the lower peak within chr11: 68.46-76.73 Mbp. It is also worthwhile to mention the locus within chr5: 74.55-82.03 Mbp, in which several genes such as Lnx1, Kdr, Clock, Cep135, Nmu, Thegl, Noa1, Adgrl3 and Rest were identified, even though this peak only reached the line for least significance (p < 0.63). The CAST founder strain was the major contributor to this trait.

There was no consistent locus across the chromosomes for BV/TV identified in the male cohorts. In the general population, QTLs scan showed two peaks (two LOD drop) at chr9: 34.55 Mbp (LOD = 7.7) and chr9: 41.66 (LOD = 8.7) Mbp, respectively (Figure 6.3). PWK was the major contributor for these peaks. Within the genomic interval 34.55-42.62 Mbp, Olfir family members dominated, harbouring 22 members of this family, on top of other reportable genes such as Ccdc15, Robo4, Robo3.
and Tbrg1. A QTL found at chr3: 50.69-53.11 Mbp also contained several WSB-derived candidate
genes including Noct, Maml3 and Cog6 (Supplementary Figure S6.1).

The QTL at chr12:7.98-9.57 Mbp appeared to contain genes that strongly associated with BV/TV in
male CC mice. This peak achieved 90\textsuperscript{th} percentile of confidence, with LOD (10.1) which was higher
than the threshold (9.0) (Figure 6.4). PWK presented as the sole phenotype contributor within this
genomic interval. ENCODE filtered out 14 SNPs, three of them were deleterious. Several genes were
suggested, including Apob, Gdf7, Slc7a15, Hs1bp3, Laptm4a, Wdr35 and Matn3.

Genome-wide scans were also performed to compare those mice with osteopetrotic phenotypes
against those with normal bone mass or osteoporosis in separate sexes, and yielded a peak that
achieved 95\textsuperscript{th} percentile of confidence on chr11 in females (Figure 6.5), and two peaks approaching
the 90\textsuperscript{th} percentile of confidence on chr11 and chr9 in males (Supplementary Figure S6.2). As shown
in Figure 6.6, within the wide genomic interval chr11:67.15-89.73 Mbp, several candidate genes are
suggested, including Gas7, Dnah2, Per1, Alox8, Alox3, Alox12, Alox12e, Alox15, Nlrp1a, Nlrp1b, Itgae, Wdr81, Nos2, Traf4 and Aatf. However, the peak at chr11 in the male cohort revealed no genes;
instead, only two genes (Sorl1 and Tecta) were pulled out from the peak at chr9 (Supplementary
Figure S6.2).
Figure 6.1 QTL haplotype mapping of BV/TV in female CC mice. (A) (Top) Genome-wide scan for % change in BV/TV in comparison with age-matched C57BL/6J. The x-axis shows the chromosomal position and the y-axis shows the $-\log_{10}(P)$ values. The $P$-values were derived from linkage haplotype data. The threshold line (orange) drawn represents 37% (adjusted $P < 0.63$) confidence. (Middle) QTL plot for the chr5 locus. This is a two-LOD support interval. Dashed line is maximum LOD $-2$. (Bottom) The eight coefficients of the QTL model show the effects on the phenotype contributed by each founder haplotype on chr5. The plot of the calculated log-odds ratio of the eight founder alleles over the chromosome where the founders are colour coded. Confidence of founder effects are represented as p values. In this region, CAST (green) contributes most to candidate genes. (B) Suggestion of candidate genes. Enlarged chromosomal region (chr5: 74.55-82.03 Mbp) shows candidate genes such as Lnx1, Kdr, Nmu and Thegl. Underlined genes have at least one predicted deleterious SNP.
Figure 6.2 QTL haplotype mapping of BV/TV in three separated female age groups. (A) QTL plot in young female CC mice. Peaks reached the orange line (37\textsuperscript{th} percentile threshold for GWS at chr5 and chr11. (B) QTL plot in middle aged female CC mice. Peak achieved 90\textsuperscript{th} percentile confidence of GWS at chr5, and another two peaks reached the orange line at chr6 and chr11. (C) QTL plot in old female CC mice. Peak reached the red line (95\textsuperscript{th} percentile threshold for GWS) at chr11, and another two peaks reached the orange line at chr4 and chr18.
Figure 6.3 QTL haplotype mapping of BV/TV in male CC mice. (A) QTL plot revealed one peak achieved 90\textsuperscript{th} percentile of GWS (p < 0.1). PWK was the dominant contributor. (C) Expansion of chromosomal region (chr9: 34.55-43.03 Mbp) contained several genes such as Cdon, Robo4 and Robo3.
Figure 6.4 QTL haplotype mapping of BV/TV in old male CC mice. (A) QTL plot showed one peak (LOD 7.1) achieved the 90th threshold (p < 0.1) at chr12. WSB was the dominant contributor. (B) Expansion of chromosomal region contained genes such as Apob, Laptm4a, Matn3 and Wdr35. There is no human homologue for Slc7a15.
Figure 6.5 QTL haplotype mapping of osteopetrosis in female CC mice. (A) QTL plot for osteopetrosis showed peak reached the red line (P < 0.05) at chr11. (B) Expansion of chromosomal region contained genes such as Per1, Alox8, Alox12, Alox15, Wdr81 and Traf4.
6.2.1.2 QTL Haplotype Scans of Tb. N

Tb. N is highly correlated with BV/TV. Not surprisingly, Tb. N in both genders produced nearly identical peaks on chr5 and chr11 (Figure 6.6 & 6.7). For instance, in the old female cohort, QTL peak was replicated on chr11: 68.46-76.73 Mbp and suggested the exact same set of genes. In the old male cohort, the locus was also replicated, and there were two deleterious SNPs detected on Apob, which was originated from CAST and was missing in the scan for BV/TV. Within the genomic interval chr9:34.41-45.61 Mbp, numerous SNPs and genes were pulled out, including Cdon, Fez1, CCdc15, Robo4, Robo3 and many Olfr family members.
Figure 6.6 QTL haplotype mapping of Tb. N in female CC mice. (A) QTL haplotype scan for the percentage change of Tb. N in female CC mice in comparison of age-matched C57BL/6J. Peaks clearly reached the orange line (p < 0.63) at chr5 and chr11. (B & C) QTL plots also showed peaks at chr5 and chr11 either in young or middle aged group. (D) QTL plot showed peak reached 90th percentile of GWS at chr11.
Figure 6.7 QTL haplotype mapping of Tb. N in male CC mice. (A) QTL haplotype scan for the percentage change of Tb. N in male CC mice in comparison of age-matched C57BL/6J. Peaks reached the orange line (p < 0.63) at chr3 and chr9. (B) QTL scan revealed no significant peak for the young male CC mice. (C) QTL scan of the middle aged male CC mice revealed one peak just reached the orange line at chr3. (D) QTL scan of the middle aged male CC mice showed a peak above the orange line at chr12.
6.2.1.3 QTL Haplotype Mapping of Tb. Th

Tb. Th revealed promising loci and strong candidate genes for bone mass

In females, QTL peaks reached the least threshold (p < 0.63) for all age groups except the older cohort (Figure 6.8 & 6.9). Given there were four peaks that reached or crossed the least threshold line, no promising gene was identified. Interestingly, there was also no gene identified within the narrow genomic interval chr9:121.68-123.11 Mbp, on which significant QTL peak (p < 0.05) was found with LOD score of 14.6, and 129S as the sole contributor. QTL scan in the middle aged cohort, however, contributed several reportable genes within chr9:27.02-34.41 Mbp, PWK-derived genes included Snx19, Adamts15, Adamts8, Prdm10, Tmem45b, Barx2, Arhgap32, Kcnj5 and Kcnj1, and CAST-derived genes included Zbtb44, St14, Aplp2, Nfrkb and Ets1 (Figure 6.9). Shared genes from those two founders included Snx19, Adamts8, Prdm10, Barx2 and Kcnj1. Among all these genes, three (Snx19, Prdm10, Kcnj5) had at least one deleterious SNP.

In males, QTL peaks reached the least threshold for all age groups (Figure 6.10 & 11). In the overall cohort, several CAST-derived genes were pulled out within a narrow genomic space chr4:14.07-14.18 Mbp, including Fbxo42, Clcnka, Zbtb17, Spen, Fblim1, Rsc1a1, Agmat, Casp9, Cela2a, Ctrc and Fhad1. However, no gene was found within another genomic interval chr17: 59.85-63.46 Mbp, which had a higher LOD score (9.7). In the young cohort, QTLs reached the middle threshold (90\textsuperscript{th} percentile, p < 0.1) at chr17:39.79-44.94, with A/J, NZO and WSB as co-founders (Figure 6.10). Five genes including Adgrf5, Mep1a, Enpp4, 9130008F23Rik and Adgrf4 were identified, of which Mep1a had 15 SNPs selected. According to Figure 6.11, within the wide genomic region chr17:59.85-74.06 Mbp, numerous genes such as Fer, Pja2, Tmem232, Txndc2, Ptpm, Lama1, Epb41l3, Gm16519, Myom1, Lpin2, Emilin2, Fam179a and Alk were identified. In addition, QTL at chr9 regenerated genes such as Snx19, Adamts15 and Adamts8 which were shown previously in male CC animals in the same age group, suggesting homogeneity within sexes. On another narrow chromosomal region chr13:98.78-100.86 that reached the least threshold (LOD 8.7), only one gene, Bdp1, was found.
Figure 6.8 QTL haplotype mapping of Tb. Th in female CC mice. (A) QTL haplotype scan for the percentage change of Tb. Th in female CC mice in comparison to age-matched C57BL/6J. Peaks reached the orange line (p < 0.63) at chr6, chr12, chr17 and chr18. (B) QTL plot in the young female mice showed one peak above the red line (p < 0.05) at chr9, and peak at chr3 reached the 37th percentile GWS. No gene was identified within the narrow genomic interval chr9:121.68-123.11 Mbp. (C) QTL plot in the old female group showed no significant peak.
Figure 6.9 QTL haplotype mapping of Tb. Th in middle aged female CC mice. (A) QTL plot showed a peak reached the orange line (37th percentile threshold for GWS) at chr9. PWK and CAST were the major trait contributors.
(B) Expansion of chromosomal region contained few genes such as Snx19, Adams15, Adams8, Kcnj5 and Kcnj1.
Figure 6.10 QTL haplotype mapping of Tb. Th in male CC mice. (A) QTL haplotype scan for the percentage change of Tb. Th in male CC mice in comparison of age-matched C57BL/6J. Peaks clearly reached the orange line ($p < 0.63$) at chr4 and chr17. (B) QTL plot of the Tb. Th in young cohort showed peak at chr17 achieved 90th percentile of GWS (yellow line). (C) QTL plot of the Tb. Th in old cohort showed a peak at chr13 reached the orange line.
Figure 6.11 QTL haplotype mapping of Tb. Th in middle aged male CC mice (chr17). (A) QTL plot clearly showed peaks at chr9, chr16 and chr7 above the orange line. CAST was the major trait contributor at chr17. (B) Expansion of chromosomal region at chr17 contained few genes such as *Fer*, *Pja2*, *Lama1* and *Emilin2*.
6.2.1.4 QTL Haplotype Mapping of Tb. Sp

QTL scans revealed no peak reached the 90\textsuperscript{th} percentile threshold for Tb. Sp in both sexes. While in the general female cohort, QTL showed a peak at chr11:61.79-81.25 Mbp with highest LOD score of 7.0, and reached the least threshold (p < 0.63) (Supplementary Figure S6.3). This peak is overlapping with that for osteopetrosis in female CC cohort, replicating genes such as Per1, Alox8, Alox12, Alox12e, Alox15, Gas7, Nos2 and Itgae. As shown in Figure 6.12, numerous genes were derived from the genomic region chr5:106.46-112.75 from the QTL scan in young female mice, including Cdc7, Pigg, Idua, Zfp605, Thegl, Polr2b, Tmprss11g, Npffr2 and Cxcl5 etc. In the middle aged group, haplotype scan showed three QTL peaks reached the least threshold at chr5, chr6 and chr11. Interestingly, no gene was identified on the wide genomic interval at chr5: 53.13-72.33 Mbp, whereas within chr6:112.6642-125.77012 Mbp, four founder strains including CAST, NOD, 129S and A/J contributed to the trait, and numerous genes were pulled out including Cpne9, Rpusd3, Fancd2, Irak2, Tmem40, Eficab12, Fam21, Olfr213, Zfp637, Zfp248, Ankrd26, Dcp1b, Cecr2, Mical3, A2m, Mug1, Clec4a1, Vmn2r19, Vmn2r25, Vmn2r26, Cdl163, C1rl, Eno2, Spsb2, P3h3, Acrbp, Ncapd2 and Tnfrsf1a. The loci at chr11 also contributed genes previously suggested, and highlighted genes such as Per1, Alox8, Alox12 and Alox15. In the older female group, QTL plot showed a peak at chr7: 26.36-29.23 Mbp, and several reportable genes including Cyp2a12, Sertad1, Fbl, Nlrp9a, Cyp2g1, and Plekhg2.

In males, QTL plots showed no peak above the threshold line with 90\textsuperscript{th} percentile confidence (Figure 6.13). In particular, in the young group, the highest peak failed to even reach the least threshold. In the middle group, even though the peak achieved only 37\textsuperscript{th} percentile of confidence, it is important to highlight it because many genes have been investigated and correlated with bone formation are within the loci (Figure 6.14). Such notable genes included Setdb1, Arnt, Adamtsl4, Ankrd35, Tnrip, Itga10, Gpr89 and Bcl9 derived from ENOCDE, and extra genes such as Snx27, Tmod4, Prune, Ctsk, Ctss, Bola1 and Notch2 derived from FANTOM5.

In the older male group, Apob was pulled out again from the locus within chr12:7.98-10.60 Mbp, adding more confidence in the claim that this gene is associated with bone phenotype (Figure 6.13). Within another locus which was at chr7:30.71-37.54 Mbp, several genes, including Wrtr62, Arhgap33, Igflr1, Cd22, Lgi4, Scgb2b12 and Tdrd12 were suggested but with very low confidence. In addition, the locus at chr10:116.86-122.60 Mbp added more genes including Mdm1, Ilitifb, Ifng, Helb, Irak3, Tbc1d30, Tbk1, Xpot, Tmem5 and Mon2 even though they are statistically less likely to be associated with this trabecular trait.
Figure 6.12 QTL haplotype mapping of Tb. Sp in female CC mice. (A) QTL plot of Tb. Sp in young female CC mice showed a peak reached the orange line at chr5. (B) QTL plot of Tb. Sp in middle aged cohort showed three peaks clearly above the orange line at chr5, chr6 and chr11. (C) QTL plot showed one peak above the orange line at chr7.
Figure 6.13 QTL haplotype mapping of Tb. Sp in male CC. (A) QTL of Tb. Sp in general population (stratified) showed peaks above the orange line at chr6, chr9 and chr16. (B) QTL scan in young cohort revealed no significant peak achieved the least GWS ($p < 0.63$). (C) QTL plot in the old cohort revealed three peaks above the orange line. In particular, Apob was pulled out again within chr12:7.98-10.60 Mbp.
Figure 6.14 QTL haplotype mapping of Tb. Sp in middle aged male CC mice. (A) QTL plot showed the peak at chr3 reached the least GWS. PWK and NZO were two major trait contributors. (B) Expansion of chromosomal region at chr3 contained genes such as Snx27, Setdb1, Arnt, Ctsk, Cts, Txnip, Gpr89 and Notch2.
6.2.1.5 QTL Haplotype Mapping of Conn. D

Conn. D demonstrated strong correlation with other trabecular traits and yielded strong loci and osteoporosis susceptibility genes

In the general population with female CC mice, even though the QTL peak achieved a high LOD score (14.1), it only achieved the least threshold (Figure 6.15). Many genes such as Ccp110, Acsm5, Acsm4, Dnah3 and Otoa were included within the genomic interval chr7:125.24-131.30 Mbp. QLT plots reached no threshold line in young or middle aged cohorts. However, in the older population, QTL peaks above the least threshold were found at chr11 and chr18. At chr11, the locus with a higher peak with LOD score of 20.7, which achieved 90\textsuperscript{th} percentile confidence, demonstrated no SNP or gene derived from the sole significant founder, NOD. The locus at the lower peak (LOD = 18.4) contained only four NOD-derived genes (Dnah2, Nlrp1a, Nlrp1b, 1700016K19Rik) within the wide genomic region (chr11: 68.46-83.61 Mbp). Among these Nlrp1a and Nlrp1b presented with one and two deleterious SNPs, respectively. At chr18, the identified genomic region was even wider (chr18: 5.52-25.47 Mbp), and contained the seven genes Zeb1, Gjd4, Colec12, Rbbp8, Lama3, Ttc39c and Cabyr.

Because the JAFFA strain had significant higher values than other strains (> 4 fold higher than the second closest strain, SOLDIER), a scan without the JAFFA strain was also performed. QTLs at chr4 (LOD = 8.0) and chr5 (LOD = 8.1) above the least threshold were found (Supplementary Figure S6.4). Within Chr4:62:31-63.92 Mbp, PWK, as the sole trait contributor, yielded genes including Zfp618, Kif12, Col27a1, Akna, Tnc and Tlr4. Within chr5:74.97-85.92 Mbp, five genes were detected: Lnx1, Nmu, Cep135, Thegl, Noa1, Adgrl3.

In males, QTL at chr9:40.94-42.62 Mbp was found reaching the red threshold (95\% confidence, p < 0.05) with LOD score of 11.4 (Figure 6.16). Seven SNPs were detected within this narrow genomic region, and three genes with missense SNPs including Sorl1, Tecta and Sc5d were detected. An extensive search of the FANTOM5 database was also performed, adding two more genes including Grik4 and Tbccl. Sorl1 and Grik4 presented motifs related to ESR1. Within chr1: 84.98-135.10 Mbp, numerous PWK-derived genes such as Gpr55, Ash18, Col6a3, Gpr35 and En1 were included within this wide genomic region. Within chr7: 125.24-129.78 Mbp, several genes including Ccp110, Acsm4, Dnah3, Abca15 and Vwa3a were derived from CAST.

In the young male group, QTL plot showed a peak at chr6, which just reached the least threshold (Figure 6.17). Many PWK-derived genes were harboured within the genomic interval chr6:117.31-123:00 Mb. However, these results are unreliable due to the low LOD score and high p value (< 0.63).
QTL scan also revealed peaks which reached the least threshold at chr3 and chr6 in the middle aged group. Within chr3:50.69-52.04 Mbp, only two genes (*Noct* and *Maml3*) were identified from WSB.

QTLs for Conn. D in old male overlapped with Tb. N for the same cohort, but with larger coverage (Figure 6.17). PWK was the sole contributor with a high LOD score (9.5). Within the wide genomic interval chr12:7.985886–26.382286 Mbp, genes such as *Slc7a15, Wdr35, Gen1, Rad51ap2, Nbas* and *Ntsr2* were pulled out.
Figure 6.15 QLT haplotype mapping of Conn. D in female CC mice. (A) QTL haplotype scan for the percentage change of Conn. D in female CC mice in comparison with age-matched C57BL/6J. Peaks reached the orange line (p < 0.63) at chr7. (B & C) QTL plots showed no peak in the young or middle aged cohort. (D) QTL plot showed multiple peaks, only the higher peak at chr11 achieved the 90th percentile GWS. However, no SNP or gene was pulled out at this peak. Instead, genes such as Dnah2, Nlrp1a and Nlrp1b were pulled out from the lower peak (LOD = 18.4).
Figure 6.16 QTL haplotype mapping of Conn. D in male CC mice (chr9). (A) QTL haplotype scan for the percentage change of Conn. D in male CC mice in comparison with age-matched C57BL/6J. QTL plot showed three loci above the orange line, and the locus at chr9 just reached 95th percentile GWS. PWK was the major contributor for the trait. (B) Expansion of chromosomal region at chr9 contained the genes *Sorl1, Sc5d* and *Tecta*. 
Figure 6.17 QTL haplotype mapping of Conn. D in male CC. (A) QTL plot in young cohort showed a peak just reached the orange line. (B) QTL plot in middle aged cohort revealed peaks at chr3 and chr6 above the orange line. (C) QTL plot in the old male CC mice showed peak at chr12 achieved the least GWS.
6.2.1.6 QTL Haplotype Mapping of SMI

SMI as a strong indicator for bone microarchitecture

In the general female population, haplotype scan showed QTL peaks reached GWS (37% confidence) at chr7, chr8 and chrX (Figure 6.18). QTLs within chr7:50.20-59.20 Mbp contained many more genes than genomic intervals at chr8 and chrX. These CAST-derived genes are considered suggestive; these genes included Zfp715, Cldnd, Cd33, Zfp658, Zfp819, Ceacam18, Siglece, Zfp473, Pnkp, Bcl2112, Hrc, Mrgprb4, Mrgprb3, Zdhhc13, Nav2, and Htatip2. QTLs at chr8 had a higher LOD score (9.5). The predominant founder contributor, A/J, had no SNP within chr8:26.04-29.47 Mbp, while the minor contributor, WSB, yielded 10 SNPs and seven genes including Adam9, Tacc1, Fgfr1, Hgsnat, Adgra2, Adrb3 and Chrna6. QTLs at chrX also had a high LOD score (9.6). The sole significant contributor, PWK, yielded genes Cypt15, Cypt14, Tenm1, Gm362, Dcafl211, Prr32, Actrt1, Smarcal and Ocr1.

In the young female group, QTL peak at chr9 just crossed the least threshold, however, yielded no SNP or gene within the chromosomal region chr9: 119.13-123.11 Mbp. In the middle aged group, QTL scan showed a peak at chr11 above the least threshold (Figure 6.18). PWK harvested genes such as Ccdc137, Aspscr1, Dcxr, Ccdc57, Sectm1a, Uts2r and Fn3krp within the narrow genomic interval chr11: 120.16-121.80 Mbp. In the older population, a single significant QTL above the least threshold appeared on chr15:66.05-74.04 Mbp, with LOD of 8.5. Totally five SNPs and four genes including Gm16308, Col22a1, Slec45a4 and Mroh5 were suggested.

In the general male population, three QTL peaks reaching the least threshold were found at chr7, chr9 and chrX (Figure 6.19). QTL at chr9 revealed a similar locus with that found in the scan for Tb. Th in middle aged female cohort. This locus also yielded notable genes such as Snx19, Adams15 and Adams8. QTLs at chr7:25.58-29.23 Mbp also yielded numerous genes from the major founder WSB. Seven of them, namely Pou2f2, Dedd2, Ceacam1, Ceacam2, Nlrp9a, Cyp2g1 and Plekhg2 presented at least one deleterious SNP. QTLs at chrX: 152.18-159.75 Mbp only provided two genes, Scml2 and Ctps2, both were derived from WSB.

QTL haplotype scan in the young male cohort showed no significant locus, while in the middle age group the QTL scan revealed genes such as Noct, Maml3, Cog6, Frem2 and Smad9 (Figure 6.19 & Figure 6.20). In the older population, QTL plots showed peaks reached the least threshold at chr1, chr17 and chr19. Even though QTL presented a highest peak at chr19 (LOD = 7.8), no gene was identified within the narrow genomic space chr19: 59.52-60.22 Mbp. QTLs within chr17: 3.26-7.25 Mbp yielded another two genes, Nox3 and Arid1b.
Figure 6.18 QLT haplotype mapping of SMI in female CC mice. (A) QTL haplotype scan for the percentage change of SMI in female CC mice in comparison with age-matched C57BL/6J. Plot showed three peaks reached GWS with 37\textsuperscript{th} percentile confidence at chr7, chr8 and chrX. (B) QTL plot in young cohort showed peak at chr9 just above the orange line. (C) QTL plot in middle aged cohort demonstrated peak at chr11 achieved the least GWS. (D) QTL plot in the old cohort showed peak at chr15 reached the least GWS.
**Figure 6.19** QLT haplotype mapping of SMI in male CC mice. (A) QTL haplotype scan for the percentage change of SMI in male CC mice in comparison with age-matched C57BL/6J. Plot clearly showed QTLs achieved the least GWS at chr7, chr9 and chrX. (B) Plot in the young male group demonstrated no significant QTL. (C) Plot in the old group showed QTLs achieved GWS with 37th percentile confidence at chr1, chr17 and chr19.
Figure 6.20 QTL haplotype mapping of SMI in middle aged male CC mice (chr3). (A) Plot showed QTLs achieved the least GWS at chr3 and chr18. (B) Expansion of chr3: 50.69–59.42 Mbp included genes such as Cog6, Frem2, Postn and Smad9.
6.2.1.7 QTL Haplotype Mapping of DA
DA in female CC mice presented the most promising and fruitful trabecular-related trait

QTL scans in general females showed peaks reached the least threshold with high LOD scores ranging from 10.5-13.2 at chr7, chr8 and chrX (Figure 6.21). At chr8, two genomic intervals were presented, one was at chr8: 14.57-24.55 Mbp, the other at chr8: 26.04-29.47 Mbp (Supplementary Figure S6.5). Interestingly, A/J, as the major contributor, contributed no SNP within these regions. Instead, all the SNPs and genes were derived from CAST, the minor contributor. Within the first region, 96 SNPs were selected from the ENCODE database, yielding numerous genes such as Mcph1, Xkr5, Defb34, Defb5, Defb9, Fam90a1a, Nek5, Smiml19 and Dkk4. Within the second region, another 18 genes were pulled out. Among them, four genes (Tacc1, Plpp5, Hgsnat and RP23-202K1.1) presented at least one deleterious SNP. QTLs at chr7: 48.22-59.20 Mbp also yielded numerous SNPs and genes from the sole significant founder, CAST. Reportable genes include Vmn2r57, AW146154, Vmn2r59, Vmn2r61, Gm21028, Zfp977, Zfp715, Clndd2, Cd33, Zfp658, Zfp819, Sigeleece, Zfp743, Bcl2l12, Hsd17b14, Fut1, Mrgprb4, Mrgprb3, Htatip2, and Prmt3. QTLs at chrX: 36.25-45.45 Mbp yielded several PWK-derived genes, including Actrt1, Dcaf12l1, Gm362, Ocrl, Prr32, Smarcal and Tenml.

Another QTL peak located at chr18: 76.35-82.88 Mbp with LOD score 8.7 failed to achieve even the least significance. We applied the Gene Mine Strain Selector and filtered out strains CIV2, CAMERON and PUB in the dataset which significantly contributed to the peak at chr18. Interestingly, QTL scan without these strains resulted in the suppression of all the prior peaks and identified a new peak within chr8: 122.51-123.59 Mbp with 95% confidence of GWS (p < 0.05) (Figure 6.22). Two genes were identified, Fam92b and Gm26878. CIV2 was later found to account for this shift because it also contributed the formation of QTL peaks at chr7. In order to suppress the QTL peak at chr8: 122.51-123.59 Mbp, a haplotype scan in the cohort with further exclusion of ROGAN and REV, two of the CAST-derived CC mice attributed to the locus, was attempted, and successfully harvested a QTL above the red line (95th percentile, p < 0.05; LOD = 14.4) at chr1: 159.04-161.38 Mbp (Figure 6.23). Two missense genes (Brinp2, Pappa2) were derived from the sole and significant founder strain, WSB. They were considered as strong candidate genes for trabecular-related phenotypes.

Additional scan was also made through exclusion of CIV2, CAMERON, PUB and ROGAN. QTL peaks at chr11 arose (highest LOD = 10.1) and reached the middle threshold (90th percentile, p < 0.1), while the peak at chr8 was suppressed, representing by LOD drop from 14.4 to 8.7 (Figure 6.24). Attention was paid to NOD, the major contributor (phenotype P value < 0.0001), and identified 42
SNPs and genes such as Ggnbp1, Pi16, Umodl1, H2-Oa, H2-Ob, H2-Eb1, Bttl2, Bttl1, Bttl6, Notch4, Agpat1 and Egfl8.

QTL plot in young females revealed two peaks reached the red line (95% confidence, p < 0.05) with extremely high LOD scores 22.5 and 20.5 at chr7 and chr18, respectively (Figure 6.25). However, within this narrow genomic space, chr7:36.92-37.54 Mbp, only one gene Gm25247 which is a mature miRNA was found (Supplementary Figure S6.6). The region within chr18:73.98-82.88 Mbp was more fruitful; it contained several genes including Skal, Mbd1, Cfap53, Myo5b, Acaα2, Lipp, Dym, Skor2, Hdhd2, Katnal2, Loxdh1, Psstp2, Epg5, Scl14a1, Setbp1, Adnp2, Rbfa, Nfatc1, Gm25071, Sall3 and Zfp236 (Figure 6.25). Five of these genes, namely Mbd1, Dym, Katnal2, Loxdh1 and Adnp2 have at least one deleterious SNP. Interestingly, the well-defined gene, Nfatc1 was found to have two SNPs, both tolerated. Extended searching in the FANTOM5 database revealed a total of 422 SNPs, and several extra genes including Mro, Mapk4, Cxxx1, Cdec11, Rpl17, Smad7, Ctf2, Zbtb7c, Smad2, Katnal2, Pias2, St8 sia5, Rnf165, Atp5a1, Psstp2, Siglec15, Scl14a1, Scl14a2, Gm16286, Hspb111, Pqlc1, Kcng2, Atp9b, Gm17383, Mbp and Gm10524.

In the middle aged female group, standardised haplotype scan did not show any QTL reached the least threshold (37% confidence), even though there were peaks shown at chr7, 8 and 11 with high LOD scores (> 11) (Figure 6.26). Normalised scan, however, showed QTL reached the 90% threshold (P < 0.05) at chr8 with LOD score of 8.3. Within chr8:9.30-14.57 Mbp, CAST contributed 68 SNPs (10 deleterious) and yielded genes including Gm10217, Lig4, Tnfsf13b, Myo16, Col4a2, Naxd, Cars2, Ankrd10, Spaca7, Atpl11a, Mcf21, F10, Tfdp1, Grk1, Gas6, Rasa3, AF366264, Champ1, Erich1 and Gm10699. Aslo, WSB contributed another nine SNPs (two deleterious) and eight genes including Myo16, Irs2, Arhgef7, Tex29, Spaca7, Proz, Gm17023 and Gas6. Extended search at FANTOM5 revealed extra genes including Fam155a, Abhd13, Gm15418, Col4a1, Rab20, Gm5607, Tubgcp3, F7, Proz, Cul4a, Lamp1, Grtp1, Adprhl1, Tnco3, Atp4b, Fam70b, Cdc16 and Dlgap2.

In the older female group, standardised haplotype scan showed multiple QTL peaks across the chromosomes with high LOD scores (from 13-15), but only loci at chr3 clearly crossed the least threshold (Supplementary Figure S6.7). Within chr3: 89.83-93.28 Mbp, numerous PWK-derived SNPs and genes were suggested. Genes with at least one deleterious SNP included Nup210l, Slc27a3, Ivl, Krrp, Crnn, Flg2, Flg, Hmrn and Rptn. Normalised scan revealed QTL reached the red line (p < 0.05) at chr13 with LOD score of 9.0 (Supplementary Figure S6.8). However, no gene was found within this narrow genomic interval (chr13: 45.62-46.02 Mbp). FANTOM5 collected a single gene, Atxn1, which was derived from WSB.
Figure 6.21 QTL haplotype mapping of DA in female CC mice (chr8, peak1). (A) QTL plot of DA in the general population (stratified) showed peaks above the orange line at chr7, chr8 and chrX. There were two peaks at chr8. (B) Expansion of genomic region chr8:14.57-24.55 Mbp contained genes such as Ckap2, Smim19 and Dkk4.
Figure 6.22 QTL haplotype mapping of DA in female CC mice (excluded CIV2, CAMERON & PUB). (A) QTL scan in the general population with exclusion of CIV2, CAMERON and PUB revealed a peak at chr8 achieved the 95th percentile GWS. CAST was the sole but extremely significant trait contributor. (C) Expansion of genomic region chr8:122.51-123.59 Mbp filtered out Fam92b and Gm26878.
Figure 6.23 QTL haplotype mapping of DA in female CC mice (excluded CIV2, CAMERON, PUB, ROGAN & REV). (A) To further suppress the QTL peak at chr8, haplotype scan of the population without key locus contributor ROGAN and REV revealed a peak above the red line. WSB was the sole but extremely significant trait contributor. (B) Two genes, Brinp2 and Pappa2 within the narrow genomic region chr1:159.04 – 161.38 Mbp were pulled out.
Figure 6.24 QTL haplotype mapping of DA in female CC mice (excluded CIV2, CAMERON, PUB and ROGAN). (A) Haplotype scan of the population without CIV2, CAMERON, PUB and ROGAN revealed two QTLs above the orange line, and the locus/loci at chr17 nearly achieved the 95\(^{th}\) percentile GWS. NZO was the major trait contributor within chr17: 24.48 – 34.77 Mbp. (B) Candidate genes such as Notch4 and Egfl8 were included within this region.
Figure 6.25 QTL haplotype mapping of DA in young female CC mice (chr18). (A) Haplotype plot showed QTLs at chr7 and chr18 above the red line. PWK was the sole but major trait contributor within chr18: 73.98 – 82.88 Mbp. (B) Expansion of this genomic region contained genes such as Ska1, Mbd1, Myo5b, Lipg, Dym, Skor2, Hdhd2, Setbp1, Nfatc1 and Sall3.
Figure 6.26 QTL haplotype mapping of DA in middle aged female CC mice. (A) Standardised scan showed no QTL reached the threshold. (B) Normalised scan showed QTL at chr8 reached the red line (90% confidence, P< 0.05). (C) Expansion of genomic region included several genes such as Lig4, Tnfsf13b, Myo16, Col4a1, Col4a2, F10 and Gas6.
**DA in male mice**

In the general group, QTL scan showed a peak at chr7 just reached the least significant threshold, with a high LOD score (25) (Figure 6.27). At first glance, loci within chr7: 48.22-59.20 Mbp contained no genes except *Cd33*. We suspected more significant trabecular-related genes may exist and extended searching at FANTOM5 was also attempted. Several more osteoporosis susceptibility genes such as *Dkk1, Saa1, Saa2, Saa3, Saa4, Sergef, Slc17a6, Abce8, Atf5, Ano5, E2f8* and *Gas2* were suggested.

In the young cohort, haplotype scan showed QTLs at chr1, chr7 and chrX achieved extremely high LOD scores (Figure 6.28). Peaks at chr1 and chr7 just crossed the least threshold with highest LOD scores 36.8 and 37.2 respectively. QTLs at chr7 shared similar genomic region with the overall cohort and suggested genes with low confidence (see Figure 6.27). The locus at chr1: 121.28-122.64 Mbp detected only 13 SNPs from CAST and yielded genes including *Epb41l5, Ptpn4, Cfap221, Dbi* and *Marco*. A normalised scan was also performed, and obtained a peak at chr11 nearly reached the 90th threshold (Figure 6.28). This region contained several genes such as *Itga2b, Adam11, March10, Ace, Gpatch8, Map3k14, Wnt9b, Wnt3, Itgb3*, *Gh, Axin2, Map2k6* and *Smurf2* from the ENCODE and FANTOM5 databases.

In the middle aged group, QTL peaks were observed on chr6, 7, 8 and 11 but failed to achieve significance (Figure 7.29). A normalised scan produced QTL peaks at chr4 with 90% confidence, and chr16 and chr17 close to this threshold (p < 0.1). Within the computed genomic interval chr4: 3.11-27.56 Mbp, founder strains PWK, CAST and NOD contributed numerous genes such as *Otud6b, Necab1, Osgin2, Cngh3, Usp45, Mms22l* and *Tmem64*. Within chr16:41.80-49.24 Mbp, genes such as *Ccde80, Slc35a5, Tmprss7, Phldb2, Plcxd2, Nectin3* and *Morc1* were included. Within chr17:19.75-25.34 mbp, many genes such as *Kremen2, Flywch1, Prss30, Gm5225, Atp6v0c, Rab26, Pkd1, Nih11, Meiob, Eme2, Ift140* and *Unkl* were aslo selected.

In the older group, mutiple QTL peaks was found across chromosomes and achieved 37% confidence at chr11 and 13, with LOD score as high as 20.7 (Supplementary Figure S6.9). Genomic interval chr13:119.38-120.18 Mbp failed to reveal any SNP or gene. QTL at chr11:117.03-121.80 Mbp yielded genes such as *Tha1, Gaa, Slc26a11, Rnf213, Endov, Cep131, Fscn2, Faap100, Oxld1, Ccde137, Rfng, Fasn, Cd7*, and *Zfp750*. 
Figure 6.27 QTL haplotype mapping of DA in male CC mice. (A) QTL plot showed a peak at chr7 achieved the least GWS with a high LOD score (25). CAST contributed most to the trait. (B) Several genes such as *Atf5*, *Cd33*, *Dkk1*, *Saa1* and *Ano5* were identified within chr7: 48.22 – 59.20 Mbp.
Figure 6.28 QTL haplotype mapping of DA in young male CC mice. (A) Standardised QTL scan showed three peaks reached or exceeded the orange line. (B) Normalised scan showed QTL peak at chr11 nearly reached the 90th percentile GWS. (C) Expansion of genomic region (chr11: 102.22 – 112.60 Mbp) contained candidate genes such as *Itga2b, Adam11, March10, Ace, Gpatch8, Map3k14, Wnt9b*, and *Axin2* from the ENCODE database, and extra genes such as *Map3k14, Mapt, Wnt3, Gh* and *Map2k6* from FANTOM5.
Figure 6.29 QTL haplotype mapping of DA in middle aged male CC mice (chr4). (A) Standardised scan showed no QTL reached the threshold. (B) Normalised scan showed QTL at chr4 reached threshold of 90% confidence. (C) Expansion of genomic region included several genes such as Slc26a7, Tmem64 and Osigin2.
6.2.2 Ct. Th in CC Mice

No QTL peaks reached the threshold line with 90\textsuperscript{th} percentile confidence of GWS, presumably due to the small variation in cortical thickness across ages in both sexes (Figure 6.30). Even though QTL plots presented peaks at chr5 and chr9 above the least threshold in the general female population, no SNP or gene was identified within these two regions. In the young female cohort, QTL peaked at chr7 and filtered out several genes including \textit{Trpm5}, \textit{Tssc4}, \textit{Kcnq1}, \textit{Cars}, \textit{Tnfrsf26}, \textit{Tnfrsf22}, \textit{Tnfrsf23}, \textit{Mrnpgr}, \textit{Mrgpre}, \textit{Ctn}, \textit{Ano1}, \textit{Oraov1} and \textit{Cend1}. QTL plot showed no peak reached even the least threshold in the middle aged cohort. In the older cohort, a locus within chr6:8.20-16.93 Mbp harboured genes including \textit{Umad1}, \textit{Thsd7a}, \textit{Vwde} and \textit{Ppp1r3a}, \textit{Rpa3}, \textit{Ica1}, \textit{Foxp2}, \textit{Glcci1}, \textit{Tmem168}, \textit{Foxp2} and \textit{Mdfic}.

In male cohort (general), QTL achieved the least significance at chr11, chr13, chr16 and chr19, with highest peak at chr11 (LOD = 7.0) (Figure 6.31). Potential genes such as \textit{Iba57}, \textit{Zkscan17}, \textit{Atpaf2}, \textit{Aldh3a2}, \textit{Sle5a10}, \textit{Pigl}, \textit{Fbxw10}, \textit{Tekt3} were suggested on chr11. In the young male cohort, four genes including \textit{Mrgpra2a}, \textit{Mrgpra2b}, \textit{Mrgprb5}, \textit{Mrgprb4} and \textit{Mrgprb8} with at least one deleterious SNP on chr7 were observed. In the middle aged group, QTLs achieved the least significance at chr8, 11 and 19. One of these peaks, within chr8:31.12-37.46 Mbp, yielded genes including \textit{Mboat4}, \textit{Ubxn8} and \textit{D8Ertd82e}. Locus/loci within chr11:52.56-67.15 Mbp contained genes such as \textit{Shroom1}, \textit{Sowaha}, \textit{Obscn}, \textit{Iba57}, \textit{Zkscan17}, \textit{Atpaf2}, \textit{Myo15}, \textit{Aldh3a2}, \textit{Sle5a10}, \textit{Pigl}, \textit{Fbxw10}, \textit{Tekt3}, \textit{Elac2}, \textit{Arhgap44}, \textit{Dnah9} and \textit{Myh3}. Within chr19: 55.58-59.04 Mbp, CAST contributed genes like \textit{Nrap}, \textit{Tdrd1} and \textit{Afap1l2}.

In the older male cohort, QTL peaks achieved the least threshold at chr4, chr5 and chr13 (Figure 6.30 & Supplementary Figure S6.10). The locus within chr4:46.15-63.92 Mbp with maximum LOD 7.2 yielded genes such as \textit{Zfp462}, \textit{A1481877}, \textit{Mup21}, \textit{Zfp37}, \textit{Kif12}, and \textit{Col27a1}.
Figure 6.30 QTL haplotype mapping of Ct. Th in female CC mice. (A) QTL scan of the general group showed peaks above the orange line at chr5 and chr9. (B) QTL scan of the young cohort had a peak above the orange line at chr7. (C) QTL scan showed no peak in the middle aged cohort. (D) QTL scan of the older cohort clearly showed a peak above the orange line at chr6.
Figure 6.31 QTL haplotype mapping of Ct. Th in male CC mice. (A) QTL scan of the general group showed peaks above the orange line at chr11, chr13, chr16 and chr19. (B) QTL scan of the young cohort had a peak above the orange line at chr7. (C) QTL scan of the middle aged cohort showed peaks at chr8, chr11 and chr19. (D) QTL scan of the old cohort clearly showed peaks above the orange line at chr8, chr11 and chr19.
6.1 Discussion

This chapter was mainly focused on haplotype mapping of eight selected parameters including seven for trabecular bone and one (Ct.Th) for cortical bone. Because most trabecular parameters differ from male to female, we separated sexes in the present study, and expected “sex-specific” genes to be uncovered. Furthermore, CC mice were categorised into four groups, including the overall, young, middle aged and old. Different mapping results were expected to be observed in different age groups.

The overall group was consolidated via ranking the percentage changes of value in comparison with age-and-gender-matched C57BL/6 mice or stratification with various age intervals. The main purpose for consolidation of CC mice with diverse age was to improve the sample size. This strategy seemed to be feasible, adding few loci that achieved 95\textsuperscript{th} percentile GWS with Conn. D in male CC mice. However, analysis in separate aged groups also worked, yielding some significant loci from BV/TV in old females, and both DA and Tb. Th in young females.

We investigated those genes derived from “reliable” QTLs which met all or most selection criteria such as low P value (< 0.05) and narrow genomic interval first, then dealt with those with less confidence with great caution. Integrating data from other sources such as correlations with human datasets, expression profiles (via BioGPS) and publications may help make a better prediction to decide whether the candidate genes are genuine or false. We have observed loci above the red line that revealed no gene either from ENCODE or FANTOM5 database. One of the explanations for this phenomenon is the significant locus was a “false-positive” result. However, if it is “true-positive”, this result could be explained by two hypotheses. One is that neither the ENCODE or FANTOM5 databases included the responsible SNP which might be a rare variant. The other hypothesis is genes beyond the identified locus regulate the phenotype via “remote” pathways. Because many QTLs are also attributed to structural variants (indels) this needs to be looked at when there is no conclusive candidate variant by missense or regulatory effects.

We also exercised the “normalisation” option for QTL scan using normalised phenotypic data for performing an association mapping analysis. Speculation on whether it was suitable to apply this approach included two aspects: first, to delete those genuine outliers with extreme phenotypes, and second, to apply which approach for phenotype transformation. We excluded two strains (CIV2 and ROGAN) with extremely high values relative to other strains in DA of middle age male group, but this did not change much in the QTL scan. These two strains might not be that significant for the determination of phenotype in this case. But when we applied rank-based tranformation to convert the dataset into a normally-distributed dataset, we sucessfully revealed a locus above the 90\textsuperscript{th} percentile threshold (Figure 7.29). However, none of those genes harboured within the locus appeared to be a strong candidate for the phenotype. This may be not an ideal example for the application of
normalised QTL scan. In another case, a normalised scan was also performed on DA in old female CC mice (Supplementary Figure S7.8). Given a locus above the red line was identified, no gene was pulled out from this locus. Therefore, caution should be paid in the application of normalised QTL scan.

Despite the normalisation option yielding undesirable results, we were inspired by successful application of Gene Mine Strain Selector, which allowed us to pinpoint those strains attributed to the phenotype. Haplotype QTL mapping of DA in the female cohort benefited most from this practice. The standardised (as is) scan revealed peaks above the orange line (p < 0.63), and the peaks covered broad genomic intervals. These peaks were not ideal for the purpose of gene mapping. We then attempted the boost the highest peak at chr8 by excluding strains (CIV2, CAMERON & PUB) that attributed most to the peak at chr18, and successfully lifted the locus on chr8 above the red line (Figure 7.22). Similar practices also boosted up a locus at chr1 that contains two strong candidate genes, namely Brinp2 and Pappa2 (Figure 7.23). This data manipulation is reasonable because we did not change the values of dataset. Rather, this is a promising application of the Selector.

Furthermore, to determine whether amino acid substitutions affect protein function, we applied SIFT which is a sequence homology-based tool that sorts intolerant from tolerant amino acid substitutions and predicts whether an amino acid substitution at a particular position in a protein will have a phenotypic effect (P. C. Ng & Henikoff, 2003). SIFT is thought to predict the phenotype resulting from a substitution more accurately than substitution scoring matrices for three datasets. However, exceptional cases in which a substitution was predicted by SIFT to be neutral or tolerated but experimentally does have a deleterious effect were observed. Nfatc1 could be a perfect example. ENCODE detected no deleterious SNP for this gene. These can be accounted for by query-specific interactions that are not conserved among the protein family members. In addition, to look at the founder effect at each locus may help determine the major-effect genes; the gene derived from the major CC founder may have larger impact on phenotype then those derived from the minor or insignificant founders (Ram et al., 2014). Genes with SNPs retrieved from both ENCODE (mostly missense), and FANTOM5 (enhancers, promoters, motifs) databases may have larger impacts than those with SNPs retrieved from a single database. The speculation is based on the notion that a large gene that encodes a large protein normally has more SNPs, making it a bigger target for mutations than a gene with less SNPs (J. Zhang & Yang, 2015). However, because phenotypes are determined by a variety of coding genes (missense) or regulatory regions (enhancers, promoters, motifs) controlling the expression of these genes, it is very difficult to compare the impacts from missense mutations and from regulatory regions.
This analysis identified more than 300 genes from 15 loci from the eight selected parameters; 86 of them were considered to be associated with bone mass and osteoporosis with preliminary investigation of published or unpublished data (See Figure 6.32 and Supplementary Table S6.1). These loci were spread on 11 chromosomes. Among them, chr3, chr8, chr11 and chr17 contained two loci identified from the present study. In comparison with the similar study published by Levy et al., the present study applied more strains, more mice and a wider range of ages, and yielded more candidate genes for bone mass and osteoporosis (Levy et al., 2015). However, pitfalls have also presented in this analysis, including varying numbers in mice and lack of application of merge analysis that may filter out false-positive results (Atamni et al., 2016).

Accumulating evidence either from published or unpublished studies highly supports that screening CC mice is a powerful way to identify genes correlated to bone phenotypes, mostly via analysing various trabecular-traits such as BV/TV and DA. In particular, DA in female CC mice apparently was the most promising and fruitful trait, harvesting totally six loci achieving the 95th percentile GWS. It has yielded several strong osteoporosis susceptibility candidate genes; some have been well investigated, some are novel. Bone phenotypes and diseases like osteoporosis are extremely complex. To further verify those osteoporosis susceptibility genes, supporting evidence can be collected from relevant publications, BioGPS, gene expression experiments, correlation with human sequences and datasets (e.g. GEFOS, TwinsUK), and animal knockout or knockin models to unveil the mechanisms underlying the various pathways. Also, to build up a protein-protein network (PPI) for all osteoporosis susceptibility genes may help comprehend such mechanisms and uncover part of the vast missing inheritance of BMD or osteoporosis.
Figure 6.32: Summary of identified loci from haplotype mapping of various traits in CC mice.
Chapter 7
Characterise Candidate Genes Responsible for Bone Mass and Microarchitecture
7.1 Introduction

It is well known that many bone phenotypes and diseases like BMD and osteoporosis are extremely complex. In the previous chapters candidate genes responsible for bone mass and osteoporosis have been identified in mice. It is necessary to characterise and prioritise these candidate genes to identify those that are most likely to be associated with bone phenotypes. Unravelling the mechanisms underlying individual or protein family effects will allow us to identify novel or effective therapeutic approaches to tackle bone-related diseases like osteoporosis or metastasis-related bone loss.

Reviewing the literature related to the identified candidate genes is a useful tool to identify genes that have known roles or defined mechanisms in bone biology, for example the identified candidate gene *Nfatc1* is well known for its important role in osteoclast biology giving a high level of confidence that observed variation in this gene may be impacting on the bone phenotype. For the purposes of developing treatments for human bone diseases identification of candidate genes in mice requires correlation of these genes and the genetic associations with human datasets. As described in Chapter 2 there has been a significant number of studies investigating genetic associations with BMD and osteoporosis in human populations. Genuine candidate genes are expected to have significant associations in both human and mouse datasets. Gene expression profiles are also useful tools in identifying whether genes are important in a particular biological system. It is assumed that genes directly expressed and regulated in bone cells will have an impact on the development of bone phenotypes. Also, the temporal expression pattern of genes may give an indication of their role within bone cells if there is little existing functional data for the resultant protein.

Proteins rarely act alone, in many circumstances they group up as molecular machines and act dynamically to undertake biological functions at both cellular and systems levels. A critical step towards unraveling the complex molecular relationships in living systems is the mapping of protein-protein interactions (PPIs). The complete map of protein interactions that can occur in a living organism is called the interactome (Cusick, Klitgord, Vidal, & Hill, 2005; De Las Rivas & Fontanillo, 2010). Interactome mapping has become a major component of current biological research (Lievens, Eyckerman, Lemmens, & Tavernier, 2010; Mehta & Trinkle-Mulcahy, 2016). Integrating PPIs with mechanistic pathways related to proteins may enrich comprehension of their complex functional roles in living systems through interacting with other proteins, and assist in identifying entire pathways that are regulating bone associated phenotypes.

In this chapter we have utilised evidence from a broad range of sources; correlation with the human genome in human datasets, published data derived from the literature or other online resources, investigation of gene expression profiles in bone cells, and mapping of the protein interactions
network, to identify those candidate genes that are most likely to be responsible for the effects seen in our population.

7.2 Results

7.2.1 Correlation with Human Datasets Revealed Novel Candidate Genes Responsible for BMD

Correlation of the genes identified from this present study with human datasets was performed by Dr. Benjamin Mullin. GEFOS cohorts for BMD (2015 release) were applied at three skeletal sites (lumbar spine, femoral neck and forearm) (Zheng et al., 2015), and selective cohorts derived from GEFOS and TwinsUK were applied for BUA and VOS (Mullin et al., 2017; Mullin et al., 2016). Analyses with these populations has showed strong correlations between bone traits and genes by reaching gene-wide significance threshold (p ≤ 0.00027778 after Bonferroni correction) or having a top SNP that is above the single-point suggestive threshold (p ≤ 3.00×10^-5). Those compelling candidates of bone traits included C11orf63 (JHY), DNAH2, DYM, EN1, HIC1, SETBP1, TNFRSF13B, WNT3, WNT9B and ZBTB7C (Table 7.1 & Supplementary Table S7.1). In addition, NFATC1 reached the gene-wide suggestive threshold (p ≤ 0.001) with p = 0.000678 but did not reach the gene-wide significance threshold. Among these 11 genes, six were pulled out from loci that achieved 95th percentile confidence in the CC mouse screen. JHY was derived from a locus reaching the 90th percentile threshold, and EN1, TNFRFSF13B, WNT3 and WNT9B were from loci that reached the 37th percentile threshold. EN1, WNT3 and WNT9B have previously been associated with BMD at the GWS level (Estrada et al., 2012; Karasik, Rivadeneira, & Johnson, 2016; Mo et al., 2015; Zheng et al., 2015). NFATC1, which did not reach the Bonferroni-corrected thresholds for the gene-wide or single-point analyses, has not yet been correlated with vBMD at the GWS level even though its role in osteoclastogenesis has been well established (Yerges et al., 2009). The rest of the genes lack published evidence for a role in bone homeostasis and are considered novel bone modulators. To characterise these seven novel osteoporosis susceptibility genes is one of the main tasks for this study.

7.2.2 Characterisation of Osteoporosis Susceptibility Genes

To characterise osteoporosis genes identified from the current study, supporting evidence was accumulated from our own gene expression analysis and a variety of other online resources, such as Pubmed, BioGPS, Uniprot, KEGG, GeneCardsSuits, ExPASy, and STRING. This section will focus on the characterisation of the candidate genes in Table 7.1, which showed significant or suggestive associations with the human genetic data. Many of these have not previously been associated with human osteoporosis or other bone traits in genetic association studies, whilst several have been previously reported as showing association with bone traits.
### Table 7.1 Candidate genes achieved significant association with bone phenotypes

<table>
<thead>
<tr>
<th>Gene</th>
<th>Bone Phenotype</th>
<th>Gene-wide P</th>
<th>TopSNP</th>
<th>TopSNP P</th>
</tr>
</thead>
<tbody>
<tr>
<td>C11orf63 (JHY)</td>
<td>BUA</td>
<td>0.000319</td>
<td>rs11218892</td>
<td>3.52 ×10^{-6}</td>
</tr>
<tr>
<td></td>
<td>VOA</td>
<td>0.000135</td>
<td>rs7934182</td>
<td>2.58 ×10^{-5}</td>
</tr>
<tr>
<td>DNAH2</td>
<td>FN-BMD</td>
<td>0.0006209999</td>
<td>rs55809496</td>
<td>6.15 ×10^{-6}</td>
</tr>
<tr>
<td>DYM</td>
<td>FN-BMD</td>
<td>0.116883117</td>
<td>rs192369781</td>
<td>5.19 ×10^{-6}</td>
</tr>
<tr>
<td>HIC1</td>
<td>FN-BMD</td>
<td>0.000045</td>
<td>rs9900967</td>
<td>1.1 ×10^{-5}</td>
</tr>
<tr>
<td>SETBP1</td>
<td>LS-BMD</td>
<td>0.02629737</td>
<td>rs4084233</td>
<td>7.19 ×10^{-6}</td>
</tr>
<tr>
<td>TNFRSF13B</td>
<td>FA-BMD</td>
<td>0.000264</td>
<td>rs74577861</td>
<td>3.68 ×10^{-3}</td>
</tr>
<tr>
<td>ZBTB7C</td>
<td>BUA</td>
<td>0.03019698</td>
<td>rs10460075</td>
<td>2.75 ×10^{-3}</td>
</tr>
<tr>
<td>EN1</td>
<td>LS-BMD</td>
<td>0.0007039999</td>
<td>rs147372226</td>
<td>1.09 ×10^{-8}</td>
</tr>
<tr>
<td></td>
<td>BUA</td>
<td>6.49999E-05</td>
<td>rs4144782</td>
<td>9.55 ×10^{-6}</td>
</tr>
<tr>
<td></td>
<td>VOS</td>
<td>7.89999E-05</td>
<td>rs188303909</td>
<td>4.48 ×10^{-7}</td>
</tr>
<tr>
<td>NFATC1</td>
<td>LS-BMD</td>
<td>0.0006779999</td>
<td>rs446628</td>
<td>5.72 ×10^{-5}</td>
</tr>
<tr>
<td>WNT3</td>
<td>LS-BMD</td>
<td>0.000245</td>
<td>rs9912530</td>
<td>6.95 ×10^{-6}</td>
</tr>
<tr>
<td>WNT9B</td>
<td>FA-BMD</td>
<td>0.001639984</td>
<td>rs11869840</td>
<td>1.59 ×10^{-3}</td>
</tr>
</tbody>
</table>

FN-BMD, femoral neck BMD; LS-BMD, lumbar spine BMD; FA-BMD, forearm BMD; BUA, broadband ultrasound attenuation; VOS, velocity of sound.

The novel genes are highlighted in bold.

Significant thresholds uses:
- Gene-wide significance threshold: 0.00027778
- Gene-wide suggestive threshold: 0.001
- Single-point suggestive threshold: 0.00003
7.2.2.1 Novel Osteoporosis Susceptibility Genes Associated with Bone Traits in Human Populations

**SETBP1**

*SETBP1* (SET Binding Protein 1) was the first strong candidate for bone phenotypes pulled out from the present study. It was derived from DA in the young female cohort, and the locus achieved 95th GWS threshold (Figure 6.25). In human populations, we observed an association between *SETBP1* and lumbar spine BMD (gene-wide P = 0.02; top SNP P = 7.19 × 10^{-6}) (Table 7.1). It suggests that this gene may be a strong susceptibility candidate for bone mass and osteoporosis. To better characterise the expression profile of this gene, we carried out real-time quantitative RT-PCR assays to examine the expression of *Setbp1* in murine osteoclasts and osteoblasts. *HPRT* was used as control for both osteogenesis and osteoclastogenesis assays. *Setbp1* was continuously expressed across the different time points. As shown in Figure 7.1, the expression of *Setbp1* shows no significant change during the process of osteogenesis from day 0 to day 21, while the marker gene *OCN* is significantly upregulated during this period. In osteoclastogenesis, *Setbp1* expression is negatively regulated during differentiation from BMMs (day 0), and shows lowest expression levels at day 5. *TRAP*, the marker gene for osteoclastogenesis, shows a consistent upward trend as expected.

*SETBP1* is a protein coding gene that contains several motifs including a ski homology region and a SET-binding region in addition to three nuclear localisation signals (Supplementary Figure S7.1). It contains 1,596 amino acids and multiple alignment of SETBP1 protein sequences from human, mouse and rat showed 94.2% similar identity. Mutations in this gene have been associated with Schinzel-Giedion syndrome (SGS) and atypical chronic myeloid leukemia (aCML) (Ammatuna et al., 2015; Hoischen et al., 2010; Makishima et al., 2013). SGS was first reported in 1978 by Schinzel and Giedion and now is a highly recognisable condition characterised by severe mental retardation, distinctive facial features, susceptibility to malignancies and multiple congenital cardiac and skeletal malformations (Lehman et al., 2008; Schinzel & Giedion, 1978). In particular, patients with SGS are manifested with a wide spectrum skeletal anomalies including a delayed bone maturation, sclerotic skull base, increased cortical density or thickness, broad and dense ribs, and a wide synchondrosis between the exoccipital and supraoccipital bones. The disease phenotype occurs in a sporadic pattern, suggesting heterozygous *de novo* mutations in a single gene as the underlying mechanism, with some exceptions that may be caused by gonadal mosaicism (Hoischen et al., 2010; Suphapeetiporn, Srichomthong, & Shotelersuk, 2011). Conventional genetic approaches have had low efficacy to identify the gene associated with this disease. Recently whole-exome sequencing was applied and has shown its efficacy in the identification of mutations that may either result in a gain-
of-function effect or have a dominant-negative effect (Hoischen et al., 2010). *SETBP1* demonstrates ubiquitous expression, which is consistent with multiple manifestations in various systems in subjects with SGS (Su et al., 2004). The function of SETBP1 remains unknown except in its capacity to bind SET domains. SET is collectively named after three *Drosophila* proteins involved in epigenetic processes, Su(var), E(z) and trithorax (Alvarez-Venegas & Avramova, 2002). Mutations outside of the SET interacting domain do not affect the DNA binding domains (Hoischen et al., 2010). In particular, the keratan sulfates, a kind of glycosaminoglycan, were found to have an important role in bone development, probably through binding the glycosaminoglycan attachment site in the SETBP1 protein (Krueger, Fields, Hildreth, & Schwartz, 1990; Nieduszynski et al., 1990). This phenomenon is consistent with characteristic skeletal malformations in subjects with this disease.

Some mutations in *SETBP1* overlap with the oncogene SKI, suggesting that *SETBP1* might play a role in the regulation of Ski-Ski homodimer and/or Ski-SnoN heterodimer formation, which both result in cellular transformation (Hoischen et al., 2010; Minakuchi et al., 2001). Such transformation are frequently found in myeloid malignancies, showing high levels of *SETBP1* expression in myeloid cancer cell lines (Makishima et al., 2013; Piazza et al., 2013). Some TGF-β target genes are differentially expressed in myeloid cancer cell lines (e.g., atypical chronic myeloid leukaemia cells) with mutated and wild-type *SETBP1*, indicating that Ski and SnoN directly act on TGF-β via its interaction with Smads (X. Liu, Sun, Weinberg, & Lodish, 2001; Piazza et al., 2013). PPI network that includes SETBP1 provides no direct evidence that strongly links SETBP1 and bone; however, its interacting partners, proteins like SET, MECOM or PRDM16, have implied roles in bone via various pathways (Figure 7.2). For instance, SET, as a multitasking protein, is involved in apoptosis, transcription, nucleosome assembly and histone chaperoning, suggesting an indirect role in the regulation of bone cells (Seo et al., 2001). MECOM (MDS1 and EVI1 complex locus) may also play a role in apoptosis through regulation of the JNK and TGF-β signalling, and has been linked to osteoporotic fracture in human populations (Hwang et al., 2013). *Mecom−/−* mice showed early-onset spinal deformity and osteopenia (Juneja et al., 2014). SETBP1 may affect bone formation and remodelling via interacting with these proteins. Further studies will be required to unveil both the physiological role of SETBP1 and its mechanistic role in the leukaemogenesis and bone formation and remodelling.

**TNFRSF13B**

TNF Receptor Superfamily Member 13B (*TNFRSF13B*), also known as *TACI*, is located at 17p11.2 and encodes a protein that is a lymphocyte-specific member of the TNF receptor superfamily. This
Figure 7.1 Gene expression of murine Setbp1 in osteoclasts and osteoblasts. BMMs isolated from C57BL/6J mice were cultured, and RNA was isolated from these cultures on day 0, 7, 14, and 21 for osteoblasts, and day 0, 1, 3 and five for osteoclasts. Real time RT-PCR was carried out to determine the expression levels of (A) OCN, Setbp1, and (B) TRAP, Setbp1 relative to housekeeping gene HPRT. * P < 0.05; ** P < 0.01; *** P < 0.001 when the value of samples in comparison to respective day 0.
Figure 7.2 Visualisation of SETBP1 protein network in various species, generated by STRING programme. (A) SETBP1 is linked to proteins like SET, MECOM and HOXA in humans. (B) Setbp1 is tightly connected with many proteins, including Mecom, Setd7, and Smarca2 in mouse. (C) Rat Setbp1 is also tightly linked to many proteins. Network nodes represent proteins; large node indicates a 3D structure is known or predicted, while small nodes represent a protein of unknown 3D structure. Colour nodes indicate query proteins and first shell or interactions, and while nodes indicate second shell of interactions. Edges represent PPIs which are specific and meaningful, and can be divided into three categories, including known interactions (from curated databases and experimentally determined), predicted functional PPIs (gene neighbourhood, gene fusion and gene cooccurrence), and others (text-mining, coexpression and homologue).
gene was originally discovered because of its ability to induce activation of NFAT via interacting with calcium-modulator and cyclophilin ligand (CAML) (vonBulow & Bram, 1997a, 1997b). TNFRSF13B was later identified a key modulator of humoral immunity by interacting with TNFSF13B (BAFF) and APRIL (Y. M. Wu et al., 2000). Cumulatively, TNFRSF13B can modulate cellular activities through inducing activation of several transcription factors including NFAT, AP-1 and NF-κB. Mutations of TNFRSF13B have been associated with immunodeficiency (Lougaris et al., 2012; Poodt et al., 2009). An association between this gene and bone has not yet been reported.

In the current study, TNFRSF13B was identified from BV/TV in middle aged female CC mice (Figure 6.2), and has been associated with forearm BMD with gene-wide p-value (0.000264) and a top SNP (rs74577861) with p-value (3.68×10^{-5}) that is close the suggestive threshold (Table 7.2). TNFRSF13B binds to APRIL and BAFF, and contains two cysteine-rich domains (CRDs) in its extracellular regions (Hymowitz et al., 2005; R. Wang et al., 2012). Human TNFRSF13B has 293 amino acids, while both mouse and rat have 249 amino acids, showing only 40.5% similarity among these species (Supplementary Figure S7.2). The major structural difference of this protein among species is that human TNFRSF13B has two TACI-CRD2 domains while either mouse or rat has only one. TNFRSF13B could play a role in osteoclastogenesis via binding to other proteins such as TNFSF13B, TRAF2, TRAF5 and TRAF6 (Figure 7.3). Its role in osteoclastogenesis is consistent with the expression profile shown on the BioGPS database, in which murine Tnfrsf13b is highly expressed in osteoclasts, B-cells and mast cells (Figure 7.4). In summary, TNFRSF13B could upregulate osteoclastogenesis through activating osteoclast-related transcription factors including NFAT (in particular Nfatc1), AP-1 and NF-κB (J. H. Kim & Kim, 2014; Soysa & Alles, 2009; E. F. Wagner, 2002).

**DYM**

Dymeclin (*DYM*), also known as Dyggve-Melchior-Clausen Syndrome Protein, is located on 18q21.1. Human DYM contains 669 amino acids, sharing 92% similarity with mouse and rat. Its main functional structure is the dymeclin domain, which is involved in intracellular trafficking (Dimitrov et al., 2009).

*DYM* is highly expressed in macrophages, osteoclasts and osteoblasts (Figure 7.5). Mutations in this gene have been linked to two types of rare recessive osteochondrodysplasia, Dyggve-Melchior-Clausen (DMC) dysplasia and Smith-McCort (SMC) dysplasia, which involve both skeletal defects and mental retardation (El Ghouzzi et al., 2003; Osipovich, Jennings, Lin, Link, & Ruley, 2008; Santos, Fernandes, Nunes, & Almeida, 2009; Thauvin-Robinet et al., 2002). Such mutations result in loss of function of Dymeclin. Mice with *Dym* deficiency have a similar manifestation to that of DMC.
Figure 7.3 Visualisation of TNFRSF13B protein network in various species. (A) Human TNFRSF13B is tightly connected with CAMLG, TRAF2, TRAF5, TNFSF13B, and TRAF6 with database and experimental evidence. (B) Murine Tnfrsf13b has similar connections with human. (C) Rat Tnfrsf13b is also linked to its ligand Tnfsf13b, and other proteins including Traf5 and Traf6.
Figure 7.4 Gene expression profile of mouse Tnrsf13b. This gene is highly expressed in B-cells, mast cells, osteoclasts and immune tissues including spleen, bone marrow and lymph nodes.
and SMC, providing a murine model of the human diseases and demonstrating functional conservation between the mouse and human proteins (Osipovich et al., 2008). Trafficking defects in Dym-deficient cells, combined with proteomic analysis of Dym-interacting proteins, establish dymeclin as a novel protein involved in Golgi organisation and vesicle transport. The PPI is shown in Figure 7.6 and reveals a strong association between DYM and IDUA (α-L-Iduronidase). IDUA has been associated with BMD in GWAS studies (Estrada et al., 2012; Mitchell, et al., 2016; Niu et al., 2016). In our analyses in human populations, although DYM did not achieve GWS (p = 0.11) with BMD at any skeletal site, the SNPs (rs192369781) achieved significance (p = 5.19×10^{-6}) above the single-point suggestive threshold (p < 3.00×10^{-5}) (Table 7.2). DYM could play a role in modulating bone mass via Golgi organisation and vesicle transport, as shown in Dym-deficient mice or humans with osteochondrodysplasia (Osipovich et al., 2008).

**HIC1**

HIC ZBTB transcriptional repressor 1 (HIC1), also named hypermethylated in cancer 1, was first identified as a tumour suppressor gene located at 17p13.3 in a region telomeric to TP53 (Wales et al., 1995). Hypermethylation or deletion in the region of this gene has been associated with many cancers since its discovery (Fujii et al., 1998; Y. Kanai et al., 1998; Koul et al., 2002; Morton et al., 1996). HIC1 has also been implicated in a contiguous-gene syndrome, the Miller-Dieker syndrome, a severe form of lissencephaly accompanied by developmental anomalies (Grimm et al., 1999; Yingling, Toyo-Oka, & Wynshaw-Boris, 2003).

HIC1 is highly conserved between human and mouse, sharing 95.5% amino acid similarity, whilst a much lower similarity (74.11%) was seen between human and rat HIC1 based on the survey of this protein sequences in the Uniprot database. This protein contains a BTB domain and five zinc finger domains. The BTB domain is a versatile PPI motif that is involved in many cellular activities including transcription, cytoskeleton dynamics, ion channel function, and protein ubiquitination (Siggs & Beutler, 2012; Stogios, Downs, Jauhal, Nandra, & Prive, 2005). Zinc finger domains have diverse functions including DNA recognition, RNA packaging, transcriptional activation, regulation of apoptosis, protein folding and assembly, and lipid binding (Laity, Lee, & Wright, 2001).

Murine Hic1 is highly expressed in osteoblasts, mast cells, dendritic cells, and organs such as placenta, uterus and ovary (Figure 7.7). The role of this gene in bone, however, has not yet been investigated. In our analysis with human datasets, HIC1 was found to be associated with femoral neck BMD with gene-wide p-value (0.00045), and a SNP (rs9900967) with p-value 1.1×10^{-5}; both reached suggestive thresholds (Table 7.2). The underlying mechanisms of this gene in bone is totally undetermined.
**Figure 7.5 Gene expression profile of mouse Dym.** This gene is expressed above the average level in HSCs, macrophages, osteoclasts and osteoblasts.
Figure 7.6 Visualisation of DYM protein network in various species. DYM is linked to IDUA, a strong BMD-associated gene.
Figure 7.7 Gene expression profile of mouse *Hic1*. It is highly expressed in osteoblasts, mast cells, and organs including placenta, uterus and ovary.
**HIC1** could be implicated in bone formation and remodelling through hypermethylation, or more likely, manifested as the skeletal phenotypes of the Miller-Dieker syndrome (Carter et al., 2000). *Hic1* deficient mice die perinatally and present various gross developmental defects such as exencephaly, cleft palate, omphalocele and limb abnormalities (Carter et al., 2000). This gene could be a negative regulator of the Wnt signalling pathway involved in digestive tract morphogenesis, dorsal/ventral axis specification, heart development, and potentially bone development (Valenta, Lukas, Doubravska, Fafilek, & Korinek, 2006).

The PPI network of HIC1 did not provide direct evidence that links *HIC1* to bone phenotypes; rather, such links revealed the role of this gene as a tumour suppressor by tight associations with proteins such as TP53, SIRT1, and E2F1 (Figure 7.8). For instance, HIC1 directly down-regulates SIRT1 and thereby is involved in regulation of p53/TP53-dependent apoptotic DNA-damage responses (W. Y. Chen et al., 2005). In addition, *HIC1* shares the locus with *SMG6*, a gene has been significantly associated with femoral neck and lumbar spine BMD, suggesting that it may also act on bone through the adjacent *SMG6* (Estrada et al., 2012). We have also attempted to link HIC1 and SMG6 using STRING but did not show any connection. Based on the accumulated evidence, in particular its association with Miller-Dieker syndrome, mutations of this gene are likely associated with bone phenotypes. This gene warrants further investigation for its role in bone.

**ZBTB7C**

*ZBTB7C*, also called zinc finger and BTB domain containing 7C, *ZBTB36* or *APM-1*, is located on 18q21.1. In our analysis, this gene was derived from DA in young female CC mice, sharing the locus with many other genes such as *NFATC1*, *SETBP1*, *SMAD2*, *SMAD7*, and *MAPK4* (Figure 6.25). ZBTB7C contains 619 amino acids and shares 93.5% similarity among human, mouse and rat. The major domains of this protein include a BTB domain and four zinc finger domains. This protein shares similar structure with above-mentioned protein, HIC1. *ZBTB7C* has been long postulated as a tumour suppressor gene (Buschges et al., 2001; Reuter et al., 1998), a role in bone has not yet been reported.

The expression of murine *Zbtb7c* is high in mature osteoblasts (day 14 & 21) but not osteoclasts, suggesting it may have a role in osteoblast differentiation (Figure 7.9). Also, correlation between BUA and *ZBTB7C* has been demonstrated, showing a top SNP (rs10460775) with p-value ($2.75 \times 10^{-5}$) above the suggestive threshold (Table 7.2). However, the underlying mechanisms that link bone traits and this gene remains a mystery. Based on Figure 7.10, PPI map shows ZBTB7C is tightly connected with proteins such as SREBF1, LRRK2 and PHLPP1. SREBF1 (sterol regulatory element
Figure 7.8 Visualisation of HIC1 protein network in various species. HIC1 is tightly associated with TP53, SIRT1, and E2F1. This gene is highlighted as a tumour suppressor.
Figure 7.9 Gene expression profile of mouse Zbtb7c. It is highly expressed in stomach, large intestine, and mature osteoblasts.
Figure 7.10 Visualisation of ZBTB7C protein network in various species. (A) Human ZBTB7C is tightly connected with proteins such as SREBF1, LRRK2, POTEJ and RIPK4 with database and experimental support. (B) Murine Zbtb7c has connections with Trp53, Ripk4, and Ankfy1. (C) Rat Zbtb7c is also linked to Ripk4 and Ankfy1, and other proteins such as Wdr38 and Wdr90.
binding transcription factor 1) is required for lipid homeostasis, and regulates transcription of the LDL receptor gene as well as the fatty acid and to a lesser degree the cholesterol synthesis pathway (Eberle, Hegarty, Bossard, Ferre, & Foufelle, 2004). LRRK2 (leucine-rich repeat kinase 2) positively regulates autophagy through a calcium-dependent activation of the CaM KK/AMPK signalling pathway, and mutations of LRRK2 have been identified as the most common cause of familial Parkinson's disease (Esteves & Cardoso, 2017; Gomez-Suaga et al., 2012; Orenstein et al., 2013). LRRK2 has been shown to interact with microtubules and mitochondria interfering with their network and dynamics (Esteves & Cardoso, 2017). ZBTB7C may modulate bone via interaction either with lipid metabolism or microtubule modulation or both. This gene appears to be a strong candidate for bone, and deserves further studies to elucidate the mechanisms underlying the association identified in our analysis. Considering the similarity between ZBTB7C and HIC1, mutations of ZBTB7C could be responsible for bone phenotypes, and further investigation on its role in bone is warranted.

**DNAH2**

This gene is also known as dynein axonemal heavy chain 2 or DNAHC2. It is located on 17p13.1 in the human genome. Human DNAH2 has 4,456 amino acids, sharing 90% sequence similarity with mouse. Murine Dnah2 is slightly increased its expression in testis (BioGPS). The expression of Dnah2 in both osteoclasts and osteoblasts is equivalent to the average expression level across all cell and tissue types tested.

DNAH2 is an axonemal inner arm dynein heavy chain, and has been identified as a major component of dyneins, which are microtubule-associated motor protein complexes (Chapelin et al., 1997; Tanenbaum, Vale, & McKenney, 2012, 2013). DNAH2 plays a key role in regulating microtubule sliding in the axonemes of cilia and flagella (Bhabha, Johnson, Schroeder, & Vale, 2016; Estrem, Fees, & Moore, 2017). Immunohistochemical studies have demonstrated the presence of DNAH2 in the flagella of human sperm, suggesting a strong link exists between DNAH2 and spermatogenesis (Jones et al., 2015). This finding is consistent with the expression pattern which has been showed to be highly in testis (BioGPS). Furthermore, a specific disease entity, craniopharyngiomas, have been associated with this gene. It has been demonstrated that the VE1 antibody not only recognises V600E-mutant BRAF proteins, but also DNAH2 (Jones et al., 2015). BRAF V600E mutations have been strongly linked to craniopharyngiomas, probably via the MAPK/ERK pathway (Cantwell-Dorris, O'Leary, & Sheils, 2011).
The role of *DNAH2* in bone remains unknown and its expression in bone cells or bone is unremarkable. However, association between femoral neck BMD and this gene has been established, showing gene-wide significance with p-value (0.000620999) and a top SNP (rs55809496, \( p = 6.15 \times 10^{-6} \)) achieved GWS (Table 7.2). *DNAH2* is therefore a strong candidate for BMD and osteoporosis. We speculate that *DNAH2* may modulate bone phenotype through modifying the microtubules of osteoblasts and osteoclasts. This novel gene requires further investigation.

**C11orf63**

Chromosome 11 open reading frame 63 (*C11orf63*), also referred to as *JHY*, is located on 11q24.1. This gene is poorly characterised, there is no human or rat protein sequence available in Uniprot databases. In addition, neither BioGPS nor STRING showed results, probably due to the lack of investigation of this gene. Interestingly, *C11orf63* has recently been identified as a novel candidate gene for hydrocephalus (Appelbe et al., 2013). In transgenic mice, disruption of this gene has been found to cause abnormal microtubule patterning and juvenile hydrocephalus (Appelbe et al., 2013). This early onset rapidly progressive hydrocephalus is highly associated with the ciliary defects caused by *C11orf63* mutations.

Interestingly, given no significant association has been established with BMD, significant associations have been identified with BUA and VOS; in which BUA analysis presented gene-wide p-value 0.000319, and a SNP (rs11218892) achieved gene-wide significance with \( p = 3.52 \times 10^{-6} \), which is above the single-point suggestive threshold (3.00 \times 10^{-5}); and VOS showed gene-wide p-value (0.000135) that is lower than gene-wide significance threshold (\( p < 0.00027778 \)), and a SNP (rs7934182) with p-value of 2.58 \times 10^{-5} (Table 7.2). Results clearly suggest a strong association between the gene and these bone traits. However, the mechanisms underlying the association between bone phenotypes and mutations/polymorphisms of *C11orf63* is yet to be determined. There has been speculation on whether bone phenotypes can be caused by ciliary defects, which would fit the profile for this protein (Appelbe et al., 2013).

**7.2.2.2 Previously Reported Genes Associated with Bone Traits in Human Populations**

Besides the novel osteoporosis susceptibility genes described above, four genes including *EN1*, *NFATC1*, *WNT3* and *WNT9B* have been previously reported as associated with bone traits in human populations. Excluding *NFATC1*, which only reached the gene-wide suggestive threshold, the remaining three genes reached the single-point suggestive threshold and/or gene-wide significance threshold (Table 7.1 & Supplementary Table S7.1). Such associations with bone traits in our analyses reinforce the confidence in determining the roles of those genes in bone formation and remodelling.
Furthermore, collecting evidence from other sources such as PPIs and pathways could further characterise these genes.

**EN1**

Engrailed Homeobox 1 (*EN1*) is located on 2q14.2 and is one of the homeobox-containing genes, which are thought to have a role in controlling development. It has 82.4% amino acid similarity among human, mouse and rat. The major functional domain of EN1 is the homeobox domain (HOX). Genes containing HOX domains encode homeodomain-containing transcriptional regulators that operate differential genetic programs along the anterior-posterior axis of animal bodies (Ko, Fast, McBride, & Staudt, 1988). EN1 has long been recognised to play a key role in vertebrate limb patterning (Adamska, MacDonald, Sarmast, Oliver, & Meisler, 2004; R. Kimmel, Loomis, Losos, Turnbull, & Joyner, 1999; R. A. Kimmel, Loomis, & Joyner, 1997). Moreover, EN1 has been implicated in the control of pattern formation during development of the CNS (Altieri, Jalabi, Zhao, Romito-DiGiacomo, & Maricich, 2015; dos Santos & Smidt, 2011; Shamim et al., 1999; Zec, Rowitch, Bitgood, & Kinney, 1997).

Mature *En1* is highly expressed in osteoblasts on day 5 and day 21 (Figure 7.11). Mutations in *En1* in mice produced different developmental defects that frequently are lethal, and resembled similar limb abnormalities caused by the deficiency of *Dkk1* (Adamska et al., 2004). *En1* and a Wnt protein, Wnt7a, were found to interact with Dkk1 during limb development in mice. Murine En1 is strongly connected with Dkk1, Wnt7a, Wnt1 and Bmp4 (Figure 7.12B). Furthermore, EN1 has been associated with SHH and IHH in human and rat (Figure 7.12A & C). These connections imply that EN1 has roles not only in Wnt but also in bone-related BMP and HH singalling pathways (G. Chen et al., 2012; R. N. Wang et al., 2014; J. Yang, Andre, Ye, & Yang, 2015). Perinatal *En1* knockout mice exhibitied osteopenia and enhanced skull bone resorption (Deckelbaum, Majithia, Booker, Henderson, & Loomis, 2006). Conditional knockout mice were viable and had lower trabecular BV/TV, Tb. N, and Tb. Th in both L5 vertebrae and femurs in comparison with their littermate controls (Zheng et al., 2015). These observations suggest that EN1 plays a role as a positive regulator of bone formation.

In humans, *EN1* with low-frequency (MAF between 1-5%) non-coding variants has been identified to be a determinant of BMD and fracture via WGS in a large population (Zheng et al., 2015). Based on the WGS study, similar associations between lumbar spine BMD with the rare variants near *EN1* were found in adult males and females. However, *EN1* sex differences have been speculated as a paediatric phenomenon during bone accretion. Mitchell *et al.* focused on children and also
Figure 7.11 Gene expression profile of mouse En1. This gene is highly expressed in osteoblasts on day 5 and day 21.
Figure 7.12 Visualisation of EN1 protein network in various species. (A) EN1 is tightly connected to WNT1, WNT7A, FGF8 and SHH in human. (B) Besides the Wnt1, Wnt7a and Shh, Murine En1 is also strongly linked to Dkk1 and Bmp4. (C) Rat En1 also links to Wnt1, Shh and Ihh.
identified that rare variation near this gene has implications for the lifelong risk of osteoporosis and fracture (J. A. Mitchell, Chesi, et al., 2016). In our analyses, EN1 appears to be the strongest candidate for bone traits, showing significant associations with lumbar spine BMD, BUA and VOS (Table 7.2). A top SNP (rs147372226) was found to be linked to lumbar spine BMD with \( p = 1.09 \times 10^{-8} \), even though its gene-wide \( p \) value (0.000703) failed to reach the significance threshold. In the BUA population, both gene-wide scan \( (p = 6.50 \times 10^{-5}) \) and scan for the top SNP (rs4144782) \( (p = 9.55 \times 10^{-6}) \) achieved significance. Another significant SNP (rs188303909) \( (p = 4.48 \times 10^{-7}) \) was detected in VOS populations, with gene-wide \( p \) value of \( 7.9 \times 10^{-5} \). The SNP rs188303909 has been associated with lumbar spine BMD and fracture, the other two SNPs are considered novel variants which may play key roles in bone. In summary, EN1 may play potential roles in the modulation of multiple pathways including Wnt, BMP and HH. Novel genetic variations identified in our human populations deserve further investigation to clarify their roles in bone formation and remodelling. Because this gene causes a large effect on bone, it is an ideal pharmaceutical target for developing anti-osteoporotic drugs.

**WNT3 and WNT9B**

WNT3 (Wnt family member 3) is clustered with WNT9B (Wnt family member 9B). Both are located in the chromosome 17q21.32 region and are two of 19 currently identified members of the Wnt signalling pathway in mouse and human (Katoh, 2005). WNT3 shares 98% amino acid identity to mouse Wnt3 protein, and 84% to human WNT3A protein, another WNT gene product (Katoh, 2002). WNT9B shares 92.2% amino acid identity to mouse Wnt9b (Kirikoshi, Sekihara, & Katoh, 2001). Given both have a WNT domain, WNT3 appears to be different from WNT9B; they share only 30.4% amino acid similarity. WNT proteins have been implicated in oncogenesis and in several developmental processes, including regulation of cell fate and patterning during embryogenesis (B. T. MacDonald et al., 2009).

Based on the BioGPS database, neither murine Wnt3 nor Wnt9b is highly expressed in osteoclasts or osteoblasts. However, both WNT proteins are strongly associated with other Wnt pathway members such as various FZDs (e.g., FZD1, FZD2, FZD3) and LRP6, according to the PPI mapping (Figure 7.13 & 7.14). Wnt signalling has been shown as an important regulatory pathway in the osteogenic differentiation of mesenchymal stem cells (Monroe, McGee-Lawrence, Oursler, & Westendorf, 2012). WNT3 and WNT9B have been significantly associated with femoral neck BMD in GWAS analyses (Estrada et al., 2012; Karasik et al., 2016; Mo et al., 2015). There is no study to identify whether these two genes are associated with BMD at other skeletal sites. According to Table 7.2, WNT3 was shown to be significantly associated with lumbar spine BMD, showing gene-wide \( p = \)
Figure 7.13 Visualisation of WNT3 protein network in various species. WNT3 is tightly connected with other Wnt pathway members such as FZDs, WLS and LRP6.
Figure 7.14 Visualisation of WNT9B protein network in various species. WNT9b is also strongly associated with other Wnt pathway members such as FZDs and LRP6.
0.000245 (threshold $p < 0.0002778$), and a top SNP (rs9912530) with $p = 6.95 \times 10^{-6}$ which is lower than the single-point suggestive threshold. WNT9B did not achieve gene-wide significance with BMD at the lumbar spine; instead, it was significantly associated with forearm BMD, showing gene-wide $p = 0.00164$, and a top SNP (rs1186984) with $p = 1.59 \times 10^{-5}$. Our findings are different from those previously reported which showed their associations with femoral neck rather than lumbar spine or forearm BMD.

Mechanistically, WNT3 and WNT9B are thought to modulate bone formation and remodelling via the canonical Wnt/$\beta$-catenin pathway (Ring, Neth, Weber, Steffens, & Faussner, 2014). This hypothesis is supported by the observations from previous reports which suggested both genes may regulate midfacial development and lip fusion through the canonical Wnt signalling pathway (Chiquet et al., 2008; Lan et al., 2006; Y. P. Lu et al., 2015; Menezes et al., 2010). However, the investigation on bone phenotypes in either $Wnt3$ or $Wnt9b$ mutated mice is still lacking, and both genes require further investigation.

**NFATC1**

NFATC1 was identified from a strong candidate locus and showed a significant association with lumbar spine BMD [(gene-wide $p = 0.00068$; topSNP (rs446628), $p = 5.72 \times 10^{-5}$)] (Table 7.1). NFATC1 is a member of the family of nuclear factor of activated T-cell (NFAT) transcription factors which were originally identified as key regulators of T-cell development and function (Macian, 2005). NFATC1 has been recognised as a master regulator of osteoclast differentiation (J. H. Kim & Kim, 2014). In particular, this gene regulates several osteoclast specific genes such as $TRAP$, $CTSK$ (cathepsin K), $CTR$ (calcitonin receptor), and $OSCAR$ (osteoclast-associated receptor) through collaboration with $MITF$ (melanogenesis associated transcription factor) and $c-Fos$ (J. H. Kim & Kim, 2014; K. Kim et al., 2005; Takayanagi et al., 2002). The role of NFATC1 in osteoclast differentiation has been investigated through genetically modified mutant mice rather than Nfatc1-deficient mice because they die at E13.5 of cardiac valve defects (de la Pompa et al., 1998). In a viable transgenic model expressing a constitutively nuclear NFATC1 variant ($NFATC1^{nuc}$), an osteopetrotic bone phenotype was observed due to nearly complete absence of osteoclasts and a dramatic increase in osteoblast number (Winslow et al., 2006). Conditional knockout $Nfatc1^{+/-}$ mice developed osteopetrosis, mainly presenting in the long bones and mandible, and inhibition of osteoclastogenesis in vivo and in vitro (Aliprantis et al., 2008). Transcriptional profiling in this study revealed Nfatc1 as a master regulator of the osteoclast transcriptome, promoting the expression of numerous genes required for bone resorption. They observed a dramatic upregulation of $Opg$ mRNA and protein, suggesting that NFATC1 diminishes the induction of OPG that emerges as a consequence of signalling
pathways shared by RANK and other TNF receptor family members. However, ectopic expression of NFATC1 was found to induce osteoclast differentiation in osteoclast precursor cells without RANKL (Takayanagi et al., 2002). The application of NFATC1 inhibitors (e.g., I-BET151) demonstrated an inhibitory effect in the process of osteoclastogenesis, mechanistically via hindering expression of NFATC1 and recruitment of its upstream regulator MYC (Park-Min et al., 2014). These results established that NFATC1 is an indispensable factor for osteoclast differentiation in vitro and in vivo.

Besides its role in osteoclastogenesis, NFATC1 also drives osteoblast proliferation in vivo (Koga et al., 2005; P. H. Stern, 2006; Winslow et al., 2006). Overexpression of NFATC1 stimulates osterix and promote osteoblast differentiation and bone formation (K. Nakashima et al., 2002). NFAT and osterix cooperatively control osteoblastic bone formation through forming a complex that binds to DNA, and this interaction is crucial for the transcriptional activity of osterix (Koga et al., 2005). In the study with NFATC1 nucleic mice, an increase of bone mass and osteoblast number was observed, indicating the role of calcineurin/NFATC1 signalling in osteoblast differentiation (Winslow et al., 2006). To further examine whether NFAT signalling in osteoblasts regulates hematopoietic differentiation, a transgenic mouse model expressing dominant-negative NFAT driven by the 2.3 kb fragment of the collagen-I promoter to disrupt NFAT activity in osteoblasts (dnNFATOB) was generated (Sesler & Zayzafoon, 2013). This model did not present any significant differences in size and weight in comparison with C57BL/6 mice, whilst significantly greater bone formation was observed (Sesler & Zayzafoon, 2013). This study demonstrates that the inhibition of NFAT activity in osteoblasts increases the number of hematopoietic stem/progenitor cells via inhibiting Ca^{2+}-stimulated nuclear translocation of NFAT transcription factors. Even though NFATC1 has been well-established for its role in the maintenance of bone homeostasis, surprisingly, its association with BMD has yet to achieve GWS (Yerges et al., 2009). NFATC1 appears to be a gene of common variants with small effects on bone, and such a notion is consistent with the findings in our analyses.

7.2.2.3 Novel Genes with Nominally Significant Associations with Bone Traits in Human Populations

Based on the correlation analyses in human datasets, there were over 50 genes that achieved p < 0.05 in the gene-wide scan (Table 7.2 & Supplementary Table S7.1). These genes were considered “nominally significantly associated” with bone traits. They may play roles in bone phenotypes in humans even though these genes are less promising than those that reached the gene-wide significance threshold. This section will highlight some of these genes with evidence that links them to bone formation and remodelling.
<table>
<thead>
<tr>
<th>Gene</th>
<th>Bone Phenotype</th>
<th>Gene-wide P</th>
<th>TopSNP</th>
<th>TopSNP P</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACE</td>
<td>BUA</td>
<td>0.02419758</td>
<td>rs1055086</td>
<td>0.001266</td>
</tr>
<tr>
<td>ADAM11</td>
<td>LS-BMD</td>
<td>0.01529847</td>
<td>rs11867421</td>
<td>0.000287</td>
</tr>
<tr>
<td></td>
<td>FN-BMD</td>
<td>0.04389561</td>
<td>rs73307236</td>
<td>0.000514</td>
</tr>
<tr>
<td>ADAMTS15</td>
<td>FA-BMD</td>
<td>0.04039596</td>
<td>rs76389393</td>
<td>0.000933</td>
</tr>
<tr>
<td></td>
<td>VOS</td>
<td>0.019998</td>
<td>rs11222094</td>
<td>0.000143</td>
</tr>
<tr>
<td>ADAMTS8</td>
<td>VOS</td>
<td>0.00889911</td>
<td>rs11222094</td>
<td>0.000143</td>
</tr>
<tr>
<td>ADORA2B</td>
<td>LS-BMD</td>
<td>0.03279672</td>
<td>rs11369177</td>
<td>0.003349</td>
</tr>
<tr>
<td>ALDH3A2</td>
<td>FN-BMD</td>
<td>0.001389986</td>
<td>rs9914784</td>
<td>0.002965</td>
</tr>
<tr>
<td>ALOX15B</td>
<td>FN-BMD</td>
<td>0.001389986</td>
<td>rs9914784</td>
<td>0.002965</td>
</tr>
<tr>
<td>ALOXE3</td>
<td>FN-BMD</td>
<td>0.01679832</td>
<td>rs9914784</td>
<td>0.002965</td>
</tr>
<tr>
<td>ARHGAP27</td>
<td>LS-BMD</td>
<td>0.01239876</td>
<td>rs2139890</td>
<td>0.003664</td>
</tr>
<tr>
<td>AXIN2</td>
<td>FA-BMD</td>
<td>0.02409759</td>
<td>rs4791161</td>
<td>0.001267</td>
</tr>
<tr>
<td></td>
<td>VOS</td>
<td>0.01619838</td>
<td>rs149366495</td>
<td>0.008198</td>
</tr>
<tr>
<td></td>
<td>VOS</td>
<td>0.03779622</td>
<td>rs7219582</td>
<td>0.000303</td>
</tr>
<tr>
<td>BRINP2</td>
<td>BUA</td>
<td>0.01509849</td>
<td>rs66778594</td>
<td>0.000179</td>
</tr>
<tr>
<td>CCDC15</td>
<td>LS-BMD</td>
<td>0.02129787</td>
<td>rs4074187</td>
<td>0.000824</td>
</tr>
<tr>
<td>CHRNA6</td>
<td>FN-BMD</td>
<td>0.00196998</td>
<td>rs7839894</td>
<td>0.000123</td>
</tr>
<tr>
<td>CHRNBP3</td>
<td>FN-BMD</td>
<td>0.005589944</td>
<td>rs56395510</td>
<td>0.000323</td>
</tr>
<tr>
<td>CLOCK</td>
<td>VOS</td>
<td>0.02539746</td>
<td>rs28665088</td>
<td>0.004282</td>
</tr>
<tr>
<td>CNTROB</td>
<td>FN-BMD</td>
<td>0.00897991</td>
<td>rs4792042</td>
<td>0.000247</td>
</tr>
<tr>
<td>COL27A1</td>
<td>FN-BMD</td>
<td>0.02069793</td>
<td>rs72760655</td>
<td>0.000477</td>
</tr>
<tr>
<td>CX3C1 (SPP1)</td>
<td>BUA</td>
<td>0.02009799</td>
<td>rs4939626</td>
<td>0.001821</td>
</tr>
<tr>
<td>ETS1</td>
<td>LS-BMD</td>
<td>0.007879921</td>
<td>rs11221335</td>
<td>0.000103</td>
</tr>
<tr>
<td></td>
<td>FA-BMD</td>
<td>0.02089791</td>
<td>rs3924289</td>
<td>0.00039</td>
</tr>
<tr>
<td>FNTA</td>
<td>BUA</td>
<td>0.04339566</td>
<td>rs145289426</td>
<td>0.019653</td>
</tr>
<tr>
<td>FREM2</td>
<td>VOS</td>
<td>0.04779522</td>
<td>rs139025725</td>
<td>0.00037</td>
</tr>
<tr>
<td>GLP2R</td>
<td>BUA</td>
<td>0.03129687</td>
<td>rs11657503</td>
<td>0.000889</td>
</tr>
<tr>
<td>GPATCH8</td>
<td>FA-BMD</td>
<td>0.02579742</td>
<td>rs138923678</td>
<td>0.00313</td>
</tr>
<tr>
<td>HS1BP3</td>
<td>FN-BMD</td>
<td>0.0189981</td>
<td>rs1377354</td>
<td>0.001554</td>
</tr>
<tr>
<td>HTRA4</td>
<td>FN-BMD</td>
<td>0.0479952</td>
<td>rs75559276</td>
<td>0.00158</td>
</tr>
<tr>
<td>ITGA2B</td>
<td>FA-BMD</td>
<td>0.03509649</td>
<td>rs138923678</td>
<td>0.00313</td>
</tr>
<tr>
<td>Gene</td>
<td>Phenotype</td>
<td>p-value</td>
<td>rsID</td>
<td>MAF</td>
</tr>
<tr>
<td>--------</td>
<td>---------------</td>
<td>----------</td>
<td>------------</td>
<td>-------</td>
</tr>
<tr>
<td>KCNJ1</td>
<td>FN-BMD</td>
<td>0.04579542</td>
<td>rs531106</td>
<td>0.000859</td>
</tr>
<tr>
<td>KIF18B</td>
<td>LS-BMD</td>
<td>0.009999</td>
<td>rs77665779</td>
<td>0.001204</td>
</tr>
<tr>
<td>LOXHD1</td>
<td>VOS</td>
<td>0.01509849</td>
<td>rs74469067</td>
<td>0.000299</td>
</tr>
<tr>
<td>MBD1</td>
<td>BUA</td>
<td>0.02259774</td>
<td>rs4939626</td>
<td>0.001821</td>
</tr>
<tr>
<td>NCOR1</td>
<td>LS-BMD</td>
<td>0.04339566</td>
<td>rs189636761</td>
<td>0.000631</td>
</tr>
<tr>
<td>MCPH1</td>
<td>VOS</td>
<td>0.03019698</td>
<td>rs17077317</td>
<td>0.001446</td>
</tr>
<tr>
<td>NOTCH2</td>
<td>FA-BMD</td>
<td>0.0209979</td>
<td>rs6686155</td>
<td>4.37 × 10^{-5}</td>
</tr>
<tr>
<td>NOTCH4</td>
<td>VOS</td>
<td>0.03819618</td>
<td>rs149448443</td>
<td>0.00257</td>
</tr>
<tr>
<td>PIGL</td>
<td>LS-BMD</td>
<td>0.0319968</td>
<td>rs78495951</td>
<td>0.000114</td>
</tr>
<tr>
<td></td>
<td>FA-BMD</td>
<td>0.04669533</td>
<td>rs77037672</td>
<td>0.000864</td>
</tr>
<tr>
<td>PRDM10</td>
<td>LS-BMD</td>
<td>0.03689631</td>
<td>rs113204914</td>
<td>0.002015</td>
</tr>
<tr>
<td>PROSER1</td>
<td>BUA</td>
<td>0.01159884</td>
<td>rs73458010</td>
<td>0.00046</td>
</tr>
<tr>
<td></td>
<td>VOS</td>
<td>0.02649735</td>
<td>rs73458010</td>
<td>0.002338</td>
</tr>
<tr>
<td>PTPN4</td>
<td>FN-BMD</td>
<td>0.02869713</td>
<td>rs79446483</td>
<td>0.000964</td>
</tr>
<tr>
<td>RASA3</td>
<td>VOS</td>
<td>0.01739826</td>
<td>rs112842768</td>
<td>0.001579</td>
</tr>
<tr>
<td>SALL3</td>
<td>FA-BMD</td>
<td>0.02789721</td>
<td>rs474991</td>
<td>7.64 × 10^{-5}</td>
</tr>
<tr>
<td>SAT2</td>
<td>BUA</td>
<td>0.01369863</td>
<td>rs1050541</td>
<td>3.86 × 10^{-5}</td>
</tr>
<tr>
<td>SKOR2</td>
<td>VOS</td>
<td>0.0179982</td>
<td>rs2048522</td>
<td>0.000771</td>
</tr>
<tr>
<td>SMAD9</td>
<td>FA-BMD</td>
<td>0.01529847</td>
<td>rs181870552</td>
<td>0.000145</td>
</tr>
<tr>
<td>SMIM19</td>
<td>VOS</td>
<td>0.04079592</td>
<td>rs72643264</td>
<td>0.008471</td>
</tr>
<tr>
<td>SMURF2</td>
<td>BUA</td>
<td>0.01949805</td>
<td>rs7208916</td>
<td>0.000256</td>
</tr>
<tr>
<td></td>
<td>VOS</td>
<td>0.03559644</td>
<td>rs9895689</td>
<td>0.001953</td>
</tr>
<tr>
<td>SPACA7</td>
<td>LS-BMD</td>
<td>0.004459955</td>
<td>rs4907529</td>
<td>7.86 × 10^{-5}</td>
</tr>
<tr>
<td>ST8SIA5</td>
<td>FA-BMD</td>
<td>0.008299917</td>
<td>rs608127</td>
<td>0.001993</td>
</tr>
<tr>
<td></td>
<td>VOS</td>
<td>0.02549745</td>
<td>rs74469067</td>
<td>0.000299</td>
</tr>
<tr>
<td>TBRG1</td>
<td>BUA</td>
<td>0.03519648</td>
<td>rs145065870</td>
<td>0.001656</td>
</tr>
<tr>
<td>THSD1</td>
<td>FN-BMD</td>
<td>0.04219578</td>
<td>rs9536035</td>
<td>0.013068</td>
</tr>
<tr>
<td>TMEM33</td>
<td>FA-BMD</td>
<td>0.03869613</td>
<td>rs116501960</td>
<td>0.005053</td>
</tr>
<tr>
<td>TRAF4</td>
<td>VOS</td>
<td>0.02439756</td>
<td>rs113797834</td>
<td>0.00225</td>
</tr>
<tr>
<td>ZDHHC7</td>
<td>FN-BMD</td>
<td>0.002079979</td>
<td>rs62050774</td>
<td>8.98 × 10^{-5}</td>
</tr>
</tbody>
</table>

FN-BMD, femoral neck BMD; LS-BMD, lumbar spine BMD; FA-BMD, forearm BMD; BUA, broadband ultrasound attenuation; VOS, velocity of sound. “Nominally significant associated” is defined as p < 0.05 in gene-wide scan. Genes in bold have been previously reported to be associated with bone phenotypes.
MBD1

In our analysis, MBD1 was derived from the strong candidate locus (18q23) containing genes such as SETBP1 and NFATC1 (Figure 6.25), and associated with BUA (p = 0.02) (Table 7.2). Based on the BioGPS database, this gene is highly expressed in lung, liver and osteoclasts. Methyl-CpG-binding domain protein 1 (MBD1) is encoded by this gene and is a member of a family of nuclear proteins (including MECP2, MBD1, MBD2, MBD3, and MBD4) related by the presence of a methyl-CpG binding domain (MBD) (P. A. Wade, 2001), which repress transcription from methylated gene promoters in vivo and in vitro (Bird, 2002; X. Zhao et al., 2003). Among various species, only 46.3% amino acid identity for MBD1 was identified. MBD1 knockout mice showed increased genomic instability, and reduced neuronal differentiation resulting in decreased neurogenesis and hippocampal malfunction (X. Zhao et al., 2003). However, studies connecting MBD1 and bone phenotype are lacking. Studies on DNA methylation in bone cells showed that demethylation of the proximal promoter region of SOST promoted SOST expression while hypermethylation suppressed its transcription, implying that DNA methylation could effectively modulate bone homeostasis through regulation of DNA binding of bone associated transcription factors (Delgado-Calle, Sanudo, Bolado, et al., 2012). Furthermore, a 170-fold induction of RANKL and a 20-fold induction of OPG mRNA expression were obtained in HEK-293 cells after administration of the DNA demethylating agent 5AzadC, suggesting methylation could play an important role in osteoclast regulation (Delgado-Calle, Sanudo, Fernandez, et al., 2012). The MBD1 PPI network (Supplementary Figure S7.3) shows tight connections with DNA binding and modifying genes such as HDAC3, RARA, SETDB1 and SUV39H1. MBD1 may affect bone cells via repressing gene expression, thereby regulating cell commitment and bone metabolism. MBD1, as an important player in gene silencing, holds some promise in bone modelling and remodelling.

SALL3

SALL3 (Spalt like transcription factor 3) is also located on 18q23 (Figure 6.25). In human populations, SALL3 achieved nominally significant association with forearm BMD [gene-wide p = 0.02; topSNP (rs474991), p = 7.64 × 10^{-5}] (Table 7.2). SALL3 is a member of the Spalt gene family encoding putative transcription factors characterised by seven to nine C2H2 zinc finger motifs (STRING). In vertebrates, the Spalt or Sall genes are expressed in early development and in a variety of embryonic and adult structures (Kuhnlein, Bronner, Taubert, & Schuh, 1997; Nishinakamura et al., 2001; Sakaki-Yumoto et al., 2006). Four murine Sall genes (Sall1 to Sall4, formerly Spalt1 to Spalt4) have been identified; Sall3 is expressed from E7 in the mesoderm and ectoderm, subsequently in the nervous systems, limb buds, palate, heart, and otic vesicles (Ott, Kaestner, Monaghan, & Schutz, 197
1996). Point mutations in SALL1 have been associated with Townes-Brocks syndrome, which is characterised by imperforate anus, pinna malformations, sensorineural hearing loss, polydactyly and kidney impairment (Botzenhart et al., 2005; Kohlhase, Wischermann, Reichenbach, Froster, & Engel, 1998); however, no diseases have been directly linked to SALL3. Mutations of this gene may be involved in 18q deletion syndrome which manifests various defects such as midfacial hypoplasia, delayed growth, and limb and digit abnormalities (Kohlhase et al., 1999; Strathdee et al., 1997). Although SALL3 is likely to be lost in these deletions, it may be not the sole cause for this syndrome because many other genes are located within this region. Sall3 knockout mice are viable at birth but die on the first postnatal day (Parrish et al., 2004). Surprisingly, no major defect was found in the kidneys, heart, skeleton, limbs or spinal cord in the knockout animals. Instead, abnormal feeding behaviour was noticed due to the malformations of the palate, epiglottis and tongue. Because the similarity among Sall family, functional compensation may occur between Sall3 and other members.

Its gene profile on BioGPS that showed no bone related expression. However, according to PPI networks (Figure 7.15), SALL3 is tightly linked to SOX2 and NANOG in humans, and Lef1 (Lymphoid enhancer binding factor 1) and Ctnnb1 (catenin beta 1) in mouse and rat. SOX2 and NANOG are the core regulators of embryonic stem cells and play key roles in tumourigenesis (Allouba, ElGuindy, Krishnamoorthy, Yacoub, & Aguib, 2015; Z. Wang, Oron, Nelson, Razis, & Ivanova, 2012). CTNBB1 has been identified as a key downstream component of the canonical Wnt signalling pathway, and plays a crucial role in controlling bone homeostasis (Mencej-Bedrac, Prezelj, Kocjan, Komadina, & Marc, 2009; Ruiz et al., 2016). Lef1 participates in the Wnt signalling pathway and activates transcription of target genes in the presence of CTNNB1 and EP300, and it has been associated with osteopenia in the absence of this gene (Hoeppner et al., 2009; Noh et al., 2009). SALL3 may have a role in the maintenance of bone homeostasis via modulating the Wnt signalling pathway.

**SKOR2**

SKOR2 (SKI family transcriptional corepressor 2) also termed functional SMAD-suppressing 235 element on chromosome 18 (FUSSEL18) or CORL2 was first described as a novel homologue to the SKI family of transcriptional co-repressors (Arndt, Poser, Schubert, Moser, & Bosserhoff, 2005). It has 1,001 amino acids in human genomes, and shares 89% amino acid identity among human, mouse and rat. SKOR2 has been studied in the differentiation of cerebellar Purkinje cells (PCs) as a PC-selective marker (Minaki, Nakatani, Mizuhara, Inoue, & Ono, 2008; Nakatani, Minaki, Kumai, Nitta, & Ono, 2014), and in the development of head and neck squamous cell carcinomas as a potential
Figure 7.15 Visualisation of SALL3 protein network in various species. (A) SALL3 is tightly connected with proteins such as SOX2, DNMT3A and NANOG in human. (B, C) Sall3 is linked to Lef2 and Ctnnb1 in mouse and rat.
tumour suppressor (Bennett et al., 2010). The role of SKOR2 in bone formation and remodelling is undefined. Our analyses in human populations showed this gene was associated with VOS (gene-wide p = 0.01; top SNP p = 0.000771) (Table 7.2). Its expression level in osteoclasts or osteoblasts is relatively low; however, links between this protein and bone phenotypes have been suggested. Skor2 is equipped with two structural domains in the N-terminal region, a DHD (Dachshund homology domain) (Wilson, Malakhova, Zhang, Joachimiak, & Hegde, 2004), and an adjacent SAND domain that is necessary and sufficient for interaction with Smad3 or Smad4 (Ueki & Hayman, 2003; J. W. Wu et al., 2002). SKOR2 has been associated with SMAD2 based on PPI networks (Supplementary Figure S7.4). Skor2 shares sequence similarity with Ski/Sno, which has been shown to down-regulate TGF-β/BMP signalling pathways through direct interaction with Smad proteins (Deheuninck & Luo, 2009; Ueki & Hayman, 2003). Considering the sequence similarity to Ski/Sno, Skor2 may have an inhibitory effect on TGF-β/BMP signalling pathways, thereby regulating bone formation.

**HS1BP3**

HS1-binding protein 3 (HS1BP3) is located at 2p24.1 in the human genome. This gene was derived from BV/TV and Tb. N in old male CC mice, in which this locus achieved the 90th threshold of confidence (Figure 6.4 & 6.7D). Human HS1BP3 has 392 amino acids, sharing 72.11% similarity with mouse and rat. The major functional domain for this protein is PhoX homologous domain. This domain has been implicated in highly diverse functions such as cell signalling, vesicular trafficking, protein sorting and lipid modification (F. Kanai et al., 2001; T. K. Sato, Overduin, & Emr, 2001; Stahelin, Burian, Bruzik, Murray, & Cho, 2003). There is no publication correlating HS1BP3 with any disease entity. This gene, however, has been implicated in Bell’s palsy and essential tremor based on the GeneCards database. The role of this gene in bone remains unknown. Gene-wide scan showed nominally significant association between femoral neck BMD and this gene (p = 0.018, Table 7.2). Furthermore, murine Hs1bp3 is ubiquitously expressed, mostly in mast cells, macrophages, osteoclasts and osteoblasts (BioGPS). The PPI shows links between HS1BP3 protein and other proteins such as CTTN (cortactin) and SORBS2 (sorbin and SH3 domain containing 2) (Supplementary Figure S7.5), both contributing to the organisation of the actin cytoskeleton and cell shape, critical components controlling osteoclast function (Cestra, Toomre, Chang, & De Camilli, 2005; Cosen-Binker & Kapus, 2006). HS1BP3 could be implicated in the organisation of the actin cytoskeleton and shaping osteoclasts and/or osteoblasts.

**ADAM11**

200
ADAM metallopeptidase domain 11 (ADAM11) is located at 17q21.31, sharing the locus with many other genes such as WNT3, WNT9B, GH1, SMURF2 and AXIN2. ADAM11 is a member of the ADAM (a disintegrin and metalloproteinase) protein family, which comprises a diverse group of multi-domain transmembrane and secreted proteins with various biological functions (Brocker, Vasiliou, & Nebert, 2009). This gene has been recognised as a candidate tumour suppressor gene for human breast cancer based on 17q21.31 microdeletion syndrome (Emi et al., 1993). The role of this gene in bone has yet to be investigated. It is highly expressed in CNS but not in osteoclasts or osteoblasts (BioGPS). In human populations, gene-wide scan showed nominally significant association with lumbar spine BMD (p = 0.015, Table 7.2). Based on PPI network (STRING), ADAM11 is connected with aggrecan, suggesting that this gene may be involved in cartilage development.

**ADAMTS8 and ADAMTS15**

This two genes are clustered and located on 11q24.3 in humans, also known as ADAM metallopeptidase with thrombospondin type 1 motif 8 and ADAM metallopeptidase with thrombospondin type 1 motif 15, respectively. They are members of ADAMTS family that share several distinct protein modules, including a propeptide region, a metalloproteinase domain, a disintegrin-like domain, and a thrombospondin type 1 (TS) motif (Flannery, 2006). Dysregulation and mutations of ADAMTS have been implicated in many diseases such as inflammation, cancer, arthritis, atherosclerosis and multiple sclerosis (Gehan, Alison, Gail, David, & Rowena, 2010; E. A. Lin & Liu, 2010; S. L. Turner, Blair-Zajdel, & Bunning, 2009; W. J. Wang et al., 2015). In particular, the role of ADAMTS as aggrecanases in the degradation of ECM and loss of cartilage in arthritis has been well studied (E. A. Lin & Liu, 2010; Q. Zhang et al., 2015). SOX9 has been identified as a regulator of ADAMTS-induced cartilage degeneration at the early stage of arthritis in humans. Furthermore, inhibition of ADAMTS was markedly increased the production of COL2A1, ACAN and COMP in chondrocytes (Q. Zhang et al., 2015). ADAMTS8 and ADAMTS15 were associated with VOS (p = 0.0089) and forearm BMD (p = 0.04), respectively (Table 7.2). They may play a role in cartilage regulation as aggrecanases.

**CD33**

CD33, also called SIGLEC-3 or sialic acid-binding Ig-ike lectin 3 is located on chromosome 19q13.41 in humans. It is a transmembrane receptor expressed on cells of myeloid lineage and it binds sialic acid (Garnache-Ottou et al., 2005). CD33 expression is down-regulated with development of the myeloid lineage, resulting in low-level expression on peripheral granulocytes and tissue macrophages. It has been considered a marker of acute myeloid leukaemias (AMLs) and its
monoclonal antibodies have been used for the treatment of patients with this disease (Burke et al., 2003; Robertson, Roy, Stone, & Ritz, 1994). Gene-wide scan showed CD33 is associated with VOS (p = 0.018) (Table 7.2). Cd33 is also highly expressed in osteoclasts, bone marrow and bone in mice (Supplementary Figure S7.6). It has been reported that CD33+ blood monocytes were programmed as pre-osteoclasts (Forsyth et al., 2009). Fusion of CD33-expressing pre-osteoclasts formed osteoclasts in giant cell tumour of bone (GCTB) which is characterised by the presence of multinucleated giant cells (osteoclast-like cells) (Forsyth et al., 2009). CD33 may play a role in the modulation of osteoclastogenesis, in particular in primary bone tumours like GCTB and its metastases. It has been applied as a target for treatment of patients with AML (Khan et al., 2017; Olombel et al., 2016).

**KCNJ1**

Potassium voltage-gated channel subfamily J (KCNJ) member 1 (KCNJ1) is harboured at 11p15.1, and belongs to KCNJ family that includes 16 members (D'Avanzo, Cheng, Dong, Nichols, & Doyle, 2010). Like most potassium channels, KCNJ1s are present in most mammalian cells, where they participate in a wide range of physiologic responses. KCNJ1 is an integral membrane protein and inward-rectifier type potassium channel. KCNJ1 is activated by internal ATP, suggesting that it may play an important role in potassium homeostasis (Ho et al., 1993). Loss-of-function mutations in KCNJ1 have been identified as a causal gene to Bartter syndrome, a group of rare autosomal recessive which is characterised by salt wasting, hypokalemic alkalosis, hypercalciuria, and low blood pressure, confirming the dependence of NaCl reabsorption on KCNJ1 - mediated K+ efflux (Sharma & Linshaw, 2011). Kcnj1 knockout mice reproduced the Bartter syndrome, providing further genetic evidence of the transport mechanism (Bailey et al., 2006; Lorenz et al., 2002). Gene-wide scan in human populations showed this gene was associated with femoral neck BMD (p = 0.045) (Table 7.2). The murine Kcnj1 is predominantly expressed in kidney (BioGPS). Mutated KCNJ1 may affect bone homeostasis as a result of changes in serum calcium, which is correlated with levels of potassium and sodium in the circulation.

**LOXHD1**

LOXHD1 (Lipoxygenase homology domains 1) is a protein consisting entirely of LH2 domains, thought to be involved in targeting proteins to the plasma membrane (Bateman & Sandford, 1999). This domain is also found in other membrane or lipid associated proteins, such as ALOX12, ALOX15, LIPG, LPL and PKD1 (Aleem et al., 2008; Bateman & Sandford, 1999; Habermann et al., 2013; Kobe, Neau, Mitchell, Bartlett, & Newcomer, 2014). LOXHD1 has been linked to auditory
defects, which are caused by malfunction of the mechanosensory hair cells in the inner ear (Grillet et al., 2009; Mori et al., 2015; Riazuddin et al., 2012). Besides its function in hair cells, little is known about the role of this gene in other cell lines. Interestingly, we speculate that LOXHD1 plays a role in bone development based on the observation that its analogue ALOX12 is associated with peak BMD in humans (Harslof et al., 2011; Ichikawa et al., 2006; W. J. Xiao et al., 2012). LOXHD1 was associated with VOS (p = 0.01) in our human populations, suggesting a potential role for this gene in bone (Table 7.2). The importance of mechanosensing to bone structural integrity hints at potential roles for proteins involved in this process in bone homeostasis; however, further studies are needed to define any role for this gene in bone modelling and remodelling.

**NOTCH4**

This gene is located at 6p21.32, and is a member of NOTCH family, which includes four Notch receptors (Notch1 to Notch4) and five ligands (Jagged1, Jagged2, Delta-like1 (Dll1), Dll3, and Dll4) with a DSL (Delta, Serrate, and Lag2) domain (Ranganathan, Weaver, & Capobianco, 2011). Human NOTCH4 protein has 2,003 amino acids sharing 76.62% similarity with mouse and rat. NOTCH4 is an important player for Notch signalling, which is involved in many biological processes, including proliferation and survival, angiogenesis, cancer stem cell activity and epithelial-to-mesenchymal transition (Andersson, Sandberg, & Lendahl, 2011; D’Souza, Meloty-Kapella, & Weinmaster, 2010; Q. Zeng et al., 2005). In human populations, gene-wide scan showed association between NOTCH4 and forearm BMD (p = 0.02) (Table 7.2). However, the expression of murine Notch4 in osteoclasts and osteoblasts are at the average levels (BioGPS). In addition, NOTCH4 is tightly connected with SMADs, suggesting that it may play a role in bone via the TGF-β signalling pathway (STRING). It also has a role in vascular formation and morphogenesis during vascular development similar to Notch1 and Notch2, suggesting Notch4 may modulate bone phenotypes partially via vascular formation (Gridley, 2007, 2010; Roca & Adams, 2007).

**TRAF4**

TNF receptor associated factor 4 (TRAF4) encodes a member of the TRAF family, which are cytoplasmic scaffold molecules that interact with TNF receptors to mediate the signalling cascade leading to the activation of both the NF-κB and the JNK pathways (Grech, Quinn, Srinivasan, Badoux, & Brink, 2000; Inoue et al., 2000; Wajant, Henkler, & Scheurich, 2001). There are six mammalian TRAFs, sharing a carboxy-terminal TRAF domain that is responsible for oligomerisation of TRAF proteins, binding to receptors, and recruitment of other signal transducers participating in the signalling cascades. Among these TRAFs, TRAF2 and TRAF6 have been well studied with key
roles in the activation of both the NF-κB and JNK pathways (Baud et al., 1999; Reinhard et al., 1997; Ruspi et al., 2014; M. C. Walsh, Lee, & Choi, 2015).

TRAF4 presents several unique properties that distinguish it from the other members of the family (Kedinger & Rio, 2007). These characteristics concern the highly conserved primary sequence of the protein, and a tightly regulated physiological expression during development. TRAF4 has been identified as a key molecule in diverse ontogenic processes, particularly in the nervous system and respiratory system (Blaise et al., 2012; Camilleri-Broet et al., 2007; W. Li et al., 2013; Peng, Li, Chen, & Dong, 2014). Traf4-deficient mice exhibit a high incidence of abnormalities such as spina bifida (impaired neural tube closure), tracheal disruption and axial skeletal malformations (Regnier et al., 2002; Shiels et al., 2000). The skeletal deformities indicate a role for TRAF4 in skeletal development and/or cartilage induction. In human populations, association between VOS and this gene was shown (p = 0.02), indicating a role of TRAF4 in bone.

**SPACA7**

SPACA7 is also called sperm acrosome associated 7 or C13orf28, and its function is unknown. It was associated with lumbar spine BMD with gene-wide p = 0.004, and a topSNP with p = 7.86 × 10^{-5} (Table 7.2). This gene is predominantly expressed in the testis, while its expression in osteoblasts or osteoclasts is average (BioGPS). This gene may modulate bone via androgen secretion (Nguyen, Westmuckett, & Moore, 2014).

**ZDHHC7**

This gene is also called zinc finger DHHC-type containing 7, and is located on 16q24.1. In human populations, gene-wide scan demonstrated association between this gene and femoral neck BMD (p = 0.002), and a topSNP (rs62050774) with p = 8.98 × 10^{-5} (Table 7.2). Murine Zdhhc7 is ubiquitously expressed, with high levels in intestines (BioGPS). Its expression in osteoclasts is high, suggesting this gene may have a role in bone. Furthermore, this gene has been associated with the regulation of cell polarity, a critical function for osteoclast activity, indicating it could play a role in regulating the function of osteoclasts (B. Chen et al., 2016).

7.2.2.4 Previously Reported Genes with Nominally Significant Associations with Bone Phenotypes in Human Populations

Analyses in human populations also showed some previously reported genes with nominally significant associations with bone phenotypes, including AXIN2, SMURF2, COL27A1, NOTCH2, SPP1 and SAT2 (see Table 7.2).
AXIN2

AXIN2 (also called conductin or axil) is located on the same chromosome region with WNT3 and WNT9B. It encodes a axin-related protein, AXIN2, which, along with its homologue AXIN1, is a negative regulator of canonical Wnt signalling through promotion of β-catenin degradation (Jho et al., 2002; Kikuchi, 1999; Leung et al., 2002). This negative feedback loop of Wnt signalling has been associated with colorectal cancer and liver cancer with defective mismatch repair (Koch et al., 2004; Lammi et al., 2004; W. Liu et al., 2000; Lustig et al., 2002; Peterlongo et al., 2005; Taniguchi et al., 2002).

Axin2 shares 44% homology with Axin1 and their functions are similar, but with varying expression patterns (Chia & Costantini, 2005). Axin1 has multiple binding sites for proteins involved in Wnt signal transduction, acting as scaffold for the protein complex involved in β-catenin degradation, and playing as a central component of the canonical Wnt pathway. It has been significantly associated with BMD at femoral neck and lumbar spine (Estrada et al., 2012). Mouse Axin2 also has multiple domains and has been shown to mediate APC-induced β-catenin degradation. Axin1 knockout mice died at Ed9.5 with forebrain truncation, neural tube defects and axis duplications (L. Zeng et al., 1997). Knock-in of the Axin2 gene into the deleted Axin1 gene rescued the phenotype of Axin1-deficient mice, indicating mouse Axin1 and Axin2 are functionally equivalent in vivo (Chia & Costantini, 2005). The homozygous Axin2 knockout mice are viable and fertile, but have craniofacial defects and premature closure of the cranial sutures due to increased β-catenin signalling (H. M. Yu et al., 2005).

Axin2 has been identified as a key regulator of bone remodelling, in particular in adult mice (Yan et al., 2009). In the Axin2 knockout mice, osteoblast proliferation and differentiation were significantly inhibited, whereas osteoclast formation was slightly promoted. Axin2 may regulate osteoblast function through β-catenin signalling in BMS cells, mainly through osteoblast precursor cells. It also inhibited the expression of Bmp2 and Bmp4, and Smad1/5 phosphorylation in osteoblasts. These findings indicate that Axin2 is a key negative regulator in bone remodelling in adult mice and regulates osteoblast differentiation through the β-catenin–BMP signalling pathway in osteoblasts.

Recently, Axin2 has been also found to contribute to healing after skeletal injury (Ransom et al., 2016). This gene was applied to trace Wnt-responding cells (WRCs) and identified a population of long-lived skeletal cells on the periosteum of long bone, and ablation of these WRCs disrupted healing after skeletal injury. These results confirm that Axin2 regulates chondrocyte maturation and axial skeletal development (Dao et al., 2010; Green et al., 2015). In our analyses, this gene was
associated with forearm BMD with gene-wide p = 0.02 (Table 7.2). In short, AXIN2 is implicated in prevention of bone loss, cartilage regeneration and tissue engineering.

**SMURF2**

SMURF2 (SMAD ubiquitination regulatory factor 2) is located on 17q24.1. It has 748 amino acids and shares 84% similarity with mouse and rat. Murine Smurf2 is highly expressed in lung and osteoblasts, in particular in young osteoblasts (BioGPS). Based on our analyses, this gene showed association with BUA with gene-wide p = 0.02 (Table 7.2). Smurf2, along with its homologue Smurf1, plays a critical role in embryogenesis and adult bone homeostasis via regulation of BMP, Wnt, and RhoA signalling pathways (K. Lu et al., 2011). They are close related and members of the Nedd4 family and function in the ubiquitination of both BMP and TGF-β activated Smads for degradation in mammalian cells (Y. Zhang, Chang, Gehling, Hemmati-Brivanlou, & Derynck, 2001). However, Smurf2 specically acts on TGF-β receptor activated Smads for ubiquitination and degradation in primary articular chondrocytes (mainly through Smad7), and stimulates their hypertrophic differentiation and maturation (Kavsak et al., 2000; Q. Wu et al., 2009). Smurfs also function in osteoblast differentiation and regulation by targeting Runx2 and MEKK2 for degradation (Kaneki et al., 2006). Moreover, accelerated endochondral ossification has been observed in ectopic expression of Smurf2 in chondrocytes and perichondrial/periosteal cells by stimulating chondrocyte maturation and osteoblast development through upregulation of β-catenin in Col2a1-Smurf2 embryos (Q. Wu, Chen, Zuscik, O'Keefe, & Rosier, 2008). In a recent study, conditional knockout of Smurf2 in osteoblasts produced the phenotype of germline Smurf2-deficient mice, suggesting that SMURF2 could regulate osteoblast-dependent osteoclast activity rather than directly acting the osteoclast (Z. Xu et al., 2017). Smurf2-deficient mice showed low bone mass, which was opposite to the phenotype of Smurf1-deficient mice, indicating that the functional difference between these two SMURFs in some aspects. The low bone mass phenotype in Smurf2 knockout mice has caused by the increased expression of Rankl as a result of alterations in Smurf2-mediated Smad3 ubiquitination (Z. Xu et al., 2017). Results in this study indicate SMURF2 as a key regulator in the crosstalk between osteoclasts and osteoblasts. In addition, SMURF2 is tightly connected with SMADs in human, mouse and rat (Supplementary Figure S7.7), and murine Smurf2 is also linked to Tgfbr1. These observations greatly suggest a key role of SMURF2 in the maintenance of bone homeostasis.

**COL27A1**

Collagen type XXVII alpha 1 chain (COL27A1) is located on 9q32 in the human genome. This gene encodes a member of the fibrillar collagen family that is highly expressed in developing skeletal
cartilage (Boot-Handford, Tuckwell, Plumb, Rock, & Poulsom, 2003; Pace, Corrado, Missero, & Byers, 2003). It is thought to play a key role during the calcification of cartilage and the transition of cartilage to bone, mainly regulated by the transcription factor SOX9 (Jenkins, Moss, Pace, & Bridgewater, 2005). It has been shown to be highly expressed in mature osteoblasts (BioGPS). In human populations, this gene was found to be associated with femoral neck BMD (p = 0.02, Table 7.2).

Mutations in this gene have recently been identified to be responsible for a novel autosomal recessive disorder, osteochondrodysplasia (also called Steel syndrome), which is characterised by bilateral hip and radial head dislocations, short stature, fusion of carpal bones, scoliosis, and cervical spine anomalies (Flynn et al., 2010; Gariballa et al., 2017; Gonzaga-Jauregui et al., 2015; Kotabagi, Shah, Shukla, & Girisha, 2017). This gene has also been found to be closely linked to other collagens such as COL1A1, COL6A1, COL5A2 etc. (Supplementary Figure S7.8). Based on these observations, COL27A1 is likely to play an important structural role in the pericellular ECM of the growth plate and is required for the organisation of the proliferative zone (Plumb et al., 2011).

**NOTCH2**

NOTCH2 is located on chromosome 1p12, and its genetic product is another member of Notch receptors (Ranganathan et al., 2011). Unlike NOTCH4, whose roles in diseases are less recognised, mutations in NOTCH2 have been reported to cause an autosomal dominant multisystem disorder, Hajdu-Cheney syndrome, which is characterised by severe and progressive bone loss (Narumi et al., 2013; Simpson et al., 2011; W. Zhao et al., 2013). In human populations, this gene was found to be associated with forearm BMD [gene-wide p = 0.02; topSNP (rs6686155), p = 4.37 × 10^{-5}] (Table 7.2).

As a key player in Notch signalling, Notch2 was found to play a crucial role in suppressing differentiation of mesenchymal progenitors in Notch-deficient mice (Hilton et al., 2008). In addition, Notch2 has been recognised as a predominant factor in suppressing osteoblastogenesis via the regulation mediated by RBPjk in which downregulation of RBPjk inhibits Runx2 activity and suppresses Nfatc1 transcription (Hilton et al., 2008; Tu et al., 2012). On the contrary, overexpression of NICD (Notch intracellular domain) under control of the 2.3 kb Col1a1 promoter in committed osteoblastic cells produced a high bone mass phenotype (S. Chen et al., 2013; Tao et al., 2010). Notch2-involved Notch signalling also directly regulates osteoclastogenesis. Deletion of Notch receptors in macrophages enhanced osteoclast differentiation in response to M-CSF and RANKL in vitro (Bai et al., 2008; Colombo et al., 2014; Fukushima et al., 2008; Sekine et al., 2012).
**SPP1**

*SPP1*, also called osteopontin or *OPN*, is located on the same locus with *SETBP1* and *NFATC1*. Its function in bone formation and remodelling has been well established, mainly acting in the regulation of normal mineralisation within the ECM of bones and teeth (Sodek, Ganss, & McKee, 2000), and playing a role in anchoring osteoclasts to the mineralised bone matrix (Denhardt & Noda, 1998). This gene has been associated with BMD but failed to achieve GWS in human (J. H. Chen et al., 2014; Estrada et al., 2012). Our analyses also identified the association between *SPP1* and BUA (p = 0.02, Table 7.2).

**SAT2**

*SAT2*, is officially named spermidine/spermine N1-acetyltransferase family member 2, and located at 17p13.1. This gene was associated with BUA in human populations, showing p = 0.01, and a top SNP (rs1050541) with p = 3.86 × 10^{-5} (Table 7.2). Little is known about this gene in bone. However, this gene is highly expressed in osteoblasts, and its expression is increased during maturation, showing a peak at day 21 (BioGPS). SAT2 has been identified as a key modulator in balancing glutamate, which is the most common excitatory neurotransmitter in the central nervous system (Bleich, Romer, Wiltfang, & Kornhuber, 2003). Beyond the role of glutamate as a neurotransmitter, the importance of glutamate signalling in bone homeostasis has been recognised (Cowan, Seidlitz, & Singh, 2012; Nedergaard, Takano, & Hansen, 2002; Spencer, Hitchcock, & Genever, 2004). Results in our analyses indicate the strong connection between this gene and bone homeostasis, and further suggests the communication mechanisms (crosstalk) between brain and bone (Cowan et al., 2012).

### 7.2.2.5 Genes Failed to Correlate with Bone Phenotypes in Human Populations

This study also identified many genes with strong evidence that links them to bone phenotypes but failed to achieve association with bone phenotypes in human populations. These genes may still modulate bone formation and remodelling. Examples of them include *TXNIP*, *CTSS*, *CTSK*, *MATN3*, *POSTN*, *SNX27*, *GAS6*, *ANOS5*, *PRUNE*, *ARNT*, *SETDB1*, *APOP*, *SMAD2*, *SMAD7*, *ALOX12*, *ALOX15*, and *PAPPA2* (see Supplementary Table S7.1). The details of these genes are included in the Appendix.

### 7.2.3 Building PPI Network for Osteoporosis Susceptibility Genes

Totally 86 candidate genes were identified that had evidence from at least one source linking them to bone phenotypes (such as publications, published databases or correlations with human bone phenotypes in our analyses). Several of these genes have been recognised as proven osteoporosis susceptibility genes such as *NFATC1*, *WNT3*, *WNT9B* and *EN1*, confirming association by both
GWAS analyses and animal models. Most genes have been reported by at least one study but associations with BMD have not yet been established, including *ALOX12, ALOX15, CTSS, GAS6, PAPPA2, POSTN, RHOB, SETDB1*, and *SMURF2*. Seven of them are considered “novel” osteoporosis susceptibility genes because there were no publications that link them to osteoporosis or BMD, of including *TNFRSF13B, C11orf63 (JHY), SETBP1, DNAH2, HIC1, DYM,* and *ZBTB7C*.

PPI networks can provide more information leading to new biological insights for individual genes or clustered genes within various pathways. Hence we have attempted to integrate the individual PPI networks that we presented above into a single PPI map using the 86 candidate genes identified from the present study. Using the STRING programme we built up a loose network with a major cluster that links *WNT3, WNT9B, AXIN2, SMAD2, SMAD7, SMAD9, NOTCH2, NOTCH4, SMURF2,* and *TXNIP* (Figure 7.16). However, most of the candidate genes were isolated, and could not provide much insight amongst the various proteins. Combining these 86 candidate genes with 115 other reported or proven osteoporosis susceptibility genes, however, demonstrated a tight network that links nearly all 201 candidate gene in a single map, with only a few isolated proteins, including *EGFL8, LOXHD1, REST, TMEM165, SNX19* and *PRUNE* (Figure 7.17). PPI mapping was also performed in mouse and rat, showing similarly condensed networks with few isolated proteins in the mouse or rat proteome (Supplementary Figure S7.9 & S7.10). These PPI maps reinforce the notion that osteoporosis proteins rarely work alone; instead, they modulate bone formation and remodelling via collaborating with each other from diverse signalling pathways. Several nodes can be identified from this combined interactome, namely proteins within the Wnt/Notch signalling pathways, TGF signalling pathways, ECM related proteins (collagens), and intracellular signalling (Jak/Stat) pathways.
Figure 7.16 Protein network of osteoporosis susceptibility genes in homo sapiens (human), built with STRING. Inputting 86 candidate genes identified from the present study built up a loose network with a major cluster.
Figure 7.17 Extended protein network of osteoporosis susceptibility genes in Homo sapiens (human). Inputting 201 osteoporosis susceptibility genes including 86 identified from haplotype mapping of CC mice. The PPI network reveals the links between OP proteins present in the human proteome.
7.3 Discussion

This chapter was focused on the characterisation and verification of numerous osteoporosis susceptibility genes identified from haplotype mapping of eight traits in a large CC cohort. More than 80 candidate genes were pulled out from over a dozen significant loci. To characterise these genes, evidence was collected from a variety of sources including publications, correlations with human datasets, protein analysis, expression profiles, PPI analysis and pathway analysis. Therefore, this chapter was heavily based on bioinformatic tools or published literature.

Candidate genes were discussed based on their associations with bone phenotypes and novelty. Basic information for each gene included association with BMD, VOS and BUA (performed by Dr Ben Mullin), sequence alignment (Uniprot), structure prediction (SMART and Prosite MyDomains), expression profiling (BioGPS), PPI network (STRING), in addition to extra information of pathway analysis (PathCards & PEGG) and RNA expression profiles of the earliest identified candidate gene *Setbp1*.

Genes that have been associated with both mouse bone traits and human BMD, VOS and BUA in our analysis are considered strong candidate genes responsible for osteoporosis. Genes without published data to confirm or speculate the association between the gene and any bone-related phenotype are considered novel. Such genes include *SETBP1, DYM, JHY (C11orf63), DNAH2, HIC1, TNFRSF13B* and *ZBTB7C*. Three genes, namely *EN1, WNT3* and *WNT9B*, which showed gene-wide significance in our analyses, have previously showed significant associations with BMD in GWAS studies (Mo et al., 2015; Zheng et al., 2015). Our results not only reinforce the associations between these proven genes with BMD, but also link them with other less reported phenotypes VOS and BUA. Furthermore, novel variants were also identified in our analyses, enriching the knowledge of the missing heritability in osteoporosis. We would highlight the discovery of novel SNPs within *EN1*, which showed strong associations with BMD, VOS and BUA. Interestingly, *NFATC1*, a well-studied gene, did not show association with bone phenotypes as significant as *EN1* or *WNT3/WNT9B* in either GWAS or gene-wide scan in our analyses (Yerges et al., 2009). The results of *NFATC1* between the publication and our analyses are highly consistent, reflecting the close correlation among published studies, gene-wide scan in our human cohorts, and haplotype gene mapping in CC mice.

Some genes that we have identified, but that have yet to be associated with BMD at gene-wide significance levels, also deserve further studies to clarify their roles in bone formation and remodelling. Examples of such genes include *ALOX12, ALOX15, AXIN2, CTSS, GAS6, PAPPA2, POSTN, RHOB, SETDB1, MATN3, SAT2* and *SMURF2*. Inability to link these previously identified genes may reflect the limitations of the human populations applied in our analyses, probably due to
the relatively smaller sample size and less diverse ethnicity compared with other GWAS and meta-GWAS studies (Estrada et al., 2012; Rivadeneira et al., 2009; L. Zhang et al., 2014).

Considering all of these genes, the locus (chr18: 73.98-82.88 Mbp, see Figure 6.25) derived from DA in young female CC cohorts appears to be the strongest and most significant locus for bone mass and bone microarchitecture, which harbours genes including SETBP1, NFATC1, ZBTB7C, DYM, MBD1, SKOR2, LOXHD1, SALL3, SPP1, MAPK4, SMAD2, and SMAD7. This locus has been associated with 18q deletion syndrome which manifests various defects such as midfacial hypoplasia, delayed growth, and limb and digit abnormalities (Kohlhase et al., 1999; Strathdee et al., 1997). Mutations of any gene mentioned above may be associated with this syndrome via various pathways. Another locus (chr11: 102.22 – 112.60 Mbp, see Figure 6.28) derived from DA in the young male cohort, is also a strong candidate locus for bone traits, even though this locus achieved only the lowest significance threshold (p < 0.63) at haplotype mapping of CC mice. This locus contains strong candidate genes such as WNT3, WNT9B, AXIN2, SMURF2, GH1, MAP2K6, MAP3K14, and ITGA2B. BV/TV and Tb. N in old female CC mice attributed the locus within chr11: 68.46-76.73 Mbp (see Figure 6.2) which contributed two significant genes, DNAH2 and HIC1. Ranking these loci based on the accumulative evidence from various sources indicates that two DA-derived loci are the strongest candidates for bone phenotypes in this study.

The candidate genes identified in this study can explain part of missing heritability of BMD or osteoporosis in human populations. They can be categorised into different groups based on their roles in various pathways or networks, such as RANK/RANKL/OPG, Wnt, BMP/TGF-β, fatty acid and energy homeostasis, microtubule and cytoskeletal network etc. (Figure 7.18). For instance, six genes namely Map3k14, Map2k6, Mapk4, Traf4, Ciss and Ctsk are belonged to the RANK/RANKL/OPG, while genes like Alox12, Alox15, Alox12b, Alox15b, Apob and Ano5 can be grouped into the pathways in regulating fatty acid and energy homeostasis. However, apart from these well-defined pathways, bone modulation is affected by many other pathways or genes, with less-known or unknown mechanisms.
Figure 7.18 Summary of OP susceptibility candidate genes from CC screening. Genes are grouped based on their functions or pathways with different ovals. The big red circle in the core covers genes that have been associated with bone phenotypes. Genes (bold) are the strongest candidates for bone mass and bone microarchitecture.
Expression profiling extracted from BioGPS provided a first impression whether the gene is associated with osteoclasts and/or osteoblasts. Genes with high expression levels in the most bone-related cells are treated as strong candidates for bone-related traits or phenotypes. Smurf2, Tmem165, Rhob, and Itga10 exemplifies such claim, by showing high expression in either osteoclasts or osteoblasts or both. However, low expression in either bone cells does not mean that gene is insignificant to the bone phenotype in many occasions. For instance, Rankl, a key activator of osteoclasts shows minimal expression in osteoclasts. Its high expression in osteoblasts is agreed with the tight connection between osteoblasts and osteoclasts (R. B. Martin, 2000). Even given the expression of a gene is low in both bone cells or in bone, we still can not exclude the association between this gene and bone. They may be connected with bone through hormonal or other pathways. Gh (growth hormone) is a perfect example; it is predominately expressed in pituitary gland, cerebral cortex and spinal cord, but is nearly undetectable in bone and bone cells (BioGPS). GH1 is a well established hormone in bone growth (Olney, 2003).

Sequence alignment offers basic information regarding to the similarity between clustered genes such as ALOX12 and ALOX15, WNT3 and WNT9B, or gene among different species (human, mouse and rat). High degrees of similarity from sequence alignment provides evidence that genetic manipulation in animal models largely mirrors the impact of mutations in such genes in humans. Structure analysis presents crucial functional domains that are shared by different species or genes in the same family. Genes containing a domain that has been previously linked to bone phenotypes may have a similar effect on bone. ALOX12 has been consistently associated with peak BMD in various human populations, whereas the role of ALOX15 in the determination of BMD is still lacking confidence because of the existing inconsistent results (Harslof et al., 2011; Ichikawa et al., 2006; Mullin et al., 2007; W. J. Xiao et al., 2011; W. J. Xiao et al., 2012). With the structural evidence that shows both genes share two key domains LH2 and lipoxygenase, ALOX15 is likely to share many similar aspects with ALOX12, including BMD modulation (W. J. Xiao et al., 2012). Interestingly, neither of them showed significant associations with bone phenotypes; instead, both ALOXE3 and ALOX15B presented nominally significant association (p < 0.05) with femoral neck BMD in in our human populations. Discrepancy among studies may be mainly due to the sample size and population. HIC1 and ZBTB7C share similar domains and have been confidently associated with bone phenotypes in the present study. Given the mechanisms underlying the bone formation and/or remodelling remain unclear, they may modulate bone through similar pathways based on their structural similarity. Other clustered genes, WNT3 and WNT9B, demonstrate a better example to support this notion. They share the WNT1 domain, and both have been associated with BMD in large scale GWAS analyses (Estrada et al., 2012; Mo et al., 2015). WNT1 appears to be a key component for all WNT proteins. You could
speculate that other WNT proteins act on bone signalling pathways via interactions with WNT1. They may play slightly different roles in the Wnt pathway due to the existence of other functional domains or regions.

Observations from the loci that contain more than two strong candidate genes suggest that genes within the same locus could work collaboratively in modulating the phenotypes. Tight connections among genes could be represented by sharing a locus or gene fusion. For example, WNT16 and FAM3C share the locus at 7q31.31, and both genes have been identified as strong candidates for osteoporosis (L. Zhang et al., 2014). In the current study, genes like SETBP1, DYM, ZBTB7C and NFATC1 are located at a shared locus. They could also modulate bone in a collaborative manner.

To look at the PPI network using bioinformatic tools such as STRING or PathCards presents a new avenue that links genes and bone phenotype in a holistic manner (Szklarczyk et al., 2015). This practice is quite different from the knockout models that usually target only one or two genes. Because proteins rarely work alone, PPIs provide hints that reflect the behaviours of various genes which are collaboratively working on bone locally or systemically. The PPI network of WNT3 and WNT9B has demonstrated association between these genes and both canonical and non-canonical Wnt signalling pathways (Supplementary Figures S7.11 & 7.12), which is counter to the assumption that they were only involved in canonical Wnt signalling (Bhat, Stauffer, Della Pietra, & Bodine, 2010; Mo et al., 2015). Another example is NOTCH4; the association between NOTCH4 and bone has yet to be established and the assumption has been made on the link between this gene and Wnt or other signalling pathways. However, direct (gene fusion) and indirect (textmining) links between SMADs and NOTCH4 are highly suggestive of a role for NOTCH4 in the TGF-β signalling pathway (STRING). Evidence on the function of EN1, collected from animal and human studies, gives little indication of the mechanisms underlying its role in neurogenesis, bone formation and remodelling (R. A. Kimmel et al., 1997; J. A. Mitchell, Chesi, et al., 2016; Zheng et al., 2015). The PPI network of this gene offers strong evidence indicating its potential roles in the modulation of multiple pathways including Wnt, BMP and HH. Integration of PPI and pathways, therefore, may provide insights which have been ignored or missed by “plain” signalling pathways. Further investigation can be used to verify the predictions made by this practice.

There has been many suggestions that some osteoporosis genes are site-specific (Richards et al., 2012; Tranah et al., 2008). The notion of site-specificity in our study implies two aspects of meaning. One refers to anatomical position, for example lumbar spine, femoral neck, forearm, calcaneus (heel) or calvarium (head). The other refers to trabecular or cortical bone. These two concepts can be applied to “mismatched” results; in which genes identified from one “site” in CC mice may be not equivalent
to the “site” in humans, or the genes pulled out from trabecular bone may have a key role in regulating cortical bone, and vice versa. For instance, \textit{Nfact1} and \textit{Setbp1} were discovered from haplotype mapping the distal femurs of CC mice, but they have been found to be associated with lumbar spine but not femoral neck. Also, our trabecular analysis identified the gene \textit{Postn}; however, this has been previously identified as regulating cortical bone via response to mechanical loading (Bonnet et al., 2009; Gerbaix, Vico, Ferrari, & Bonnet, 2015). Meanwhile, the cortical analysis derived gene \textit{Col27a1} appears to be more important in the development of skeletal cartilage than cortical expansion (Gonzaga-Jauregui et al., 2015; Jenkins et al., 2005). This phenomenon can be explained that those genes are probably non-site-specific. In other words, \textit{Nfact1} and \textit{Setbp1} may not only modulate lumbar spine but also femoral neck or distal femur, and \textit{Postn} and \textit{Col27a1} may play roles in both trabecular or cortical bone.

Even though our analyses were divided into two gender groups, and identified genes from various loci, we have no conclusive findings that indicate whether these genes are “gender-specific”. Furthermore, we have attempted to address whether some genes are “age-specific” by comparing the young and older mice, and revealed a locus on chr2 and pulled out two genes (\textit{Ctcflos} and \textit{Rab22a}). They are unlikely to be genes that determine bone mass in different age groups due to the less reliable P-value (just above the 37th percentile threshold) and no supporting evidence from other sources.

\textit{GPR89}, as a member of GPCRs, may be associated with age-dependent changes in skeletal growth under undefined mechanisms (Hsiao, Boudignon, Halloran, Nissenson, & Conklin, 2010). However, this gene was identified from Tb. Th in middle aged male CC mice. It may be a coincident finding and can not explain the age-dependent changes in bone at this stage. We are expecting age-specific genes could be discovered if we have a larger cohort of CC mice.

Many animal knockout models mimic the phenotype that presents in human with mutations on the same gene. For example, mice with \textit{Dym} deficiency have similar manifestations to that of DMC and SMC, perfectly providing a murine model of the human diseases and presenting functional conservation between the mouse and human proteins (Osipovich et al., 2008). Many genetically modified models are created based on the assumption that the effects presented in mutated animals mirror the manifestations in human with mutations in the same gene. However, caution should be paid when we translate those findings from knockout models to humans. The phenotypic discrepancy between mice and humans of \textit{ANO5} and \textit{WDR35} mutations exemplifies the fact that functions of certain genes may vary among species, even given a high percentage of genetic similarity between humans and animals (Mill et al., 2011; J. Xu et al., 2015). Just taking PPI into account, compensation effects may be induced by the deletion of one gene but other connected genes may fill the deficits
created by the mutated gene and present no obvious defect morphologically (Carelle-Calmels et al., 2009; El-Brolosy & Stainier, 2017).

Several genes that play roles in patterning, neurogenesis and growth seem to have roles in bone growth and remodelling. Those genes include EN1, DYM, SETBP1, SETDB1, TMEM165, WNT3 and WNT9B. Observations from these genes suggest associations between bone growth and neurogenesis. The mechanisms may vary from one gene to the other. For instance, mutations in DYM lead to two types of rare recessive osteochondrodysplasia which are both characterised by skeletal defects and defects in the CNS (Neumann et al., 2006). Loss of function of dymeclin may cause both skeletal and neural defects via Golgi organisation and vesicle transport. SETDB1 is another example in which deletion of Setdb1 causes peri-implantation lethality during embryogenesis, suggesting that this gene is required for peri-implantation development and the survival of embryonic stem cells (Bilodeau, Kagey, Frampton, Rahl, & Young, 2009). This gene is highly expressed in embryonic stem cells and osteoclasts in mice, and in patients with Huntington's disease (Ryu et al., 2006). This gene has also been reported to play a role in bone and cartilage development (Lawson, Yang, & Chansky, 2013). Mutations in SETBP1 present bone and brain phenotypes, suggesting the SET domain of this gene may play dual roles in the development of both organs (Hoischen et al., 2010). Midfacial, CNS and bone phenotypes were observed in mice with mutated Wnt9b also suggesting a dual role for this gene, probably through the Wnt signalling pathway (Lan et al., 2006).

The crosstalk between brain and bone is also indicated by other genes including SAT2, GH1, and possibly BRINP2 and NMU. Indeed, osteoporosis has been considered as a neuroskeletal disease and the discovery that the brain controls bone remodelling has provided a new paradigm for the understanding of bone biology (Rousseaud, Moriceau, Ramos-Brossier, & Oury, 2016; Takeda, 2008, 2009). Accumulating evidence has suggested neurons and neurotransmitters are closely involved in bone remodelling (Takeda, 2009). SAT2, one of the identified candidate genes responsible for bone phenotype, has been recognised as a key enzyme in the maintenance of metabolic glutamine/glutamate balance (Jenstad et al., 2009). As an important neurotransmitter, glutamate has been identified as a key modulator for the brain-bone communication (Cowan et al., 2012). Previous reports have showed that glutamate concentrations within the bone niche are modulated by all bone cell types, indicating a functional crosstalk between the brain and bone (Hinoi et al., 2007; Huggett, Mustafa, O'Neal, & Mason, 2002; Takarada-Iemata et al., 2011; L. Y. Wang, Hinoi, Takemori, Nakamichi, & Yoneda, 2006). In fact, all bone cells are thought to have the capacity to secrete glutamate to some extent in autocrine and/or paracrine manners (Masi, 2012). For example, osteoblasts have been observed to actively secrete glutamate and activate receptors expressed on bone
cell surfaces, providing convincing evidence for an osteoblast-derived intrinsic osteo-glutamatergic signalling mechanism (Genever & Skerry, 2001). GH1 is another example that links the brain and bone. This gene shares a locus with SMURF2, and is close to other genes such as WNT3, WNT9B and AXIN2 (Figure 6.28). Growth hormone plays significant role in the regulation of growth and bone metabolism (Kasukawa et al., 2004; Ohlsson et al., 1998; Olney, 2003). Interestingly, growth hormone could act as a double-edged sword on bone; it stimulates osteoblast differentiation therefore promotes bone formation directly or through IGF-I, or stimulates osteoclast differentiation and activity then promotes bone resorption (Olney, 2003). BRINP2 (BMP/retinoic acid inducible neural specific 2) may be another example linking the brain and bone, which was identified from a locus above the 95th GWS threshold (Figure 6.23) and has been associated with BUA in our human populations (gene-wide p = 0.015, Table 7.1). This gene plays an important role in the differentiation of peripheral sympathetic neurons, indicating a possible role in the central control of bone remodelling (Kawano et al., 2004). Another neuropeptide NMU is also assumed to play a role in central control of bone remodelling based on the fact that the anorexigenic hormone leptin inhibits bone formation through a hypothalamic relay (Appendix) (Ducy, Amling, et al., 2000; Kalra, Dube, & Iwaniec, 2009). Nmu−/− mice show a high bone mass due to increased bone formation, with male mice more severely affected than female mice (S. Sato et al., 2007). In a GWAS study, association between NMU and bone strength regulation has been shown, suggesting the strong link between bone and brain via poorly defined mechanisms (Gianfagna et al., 2013). Along with evidence from other neuropeptides such as CGRP (calcitonin gene-related protein), VIP (vasoactive intestinal peptide) and SP (substance P) which have been associated with bone phenotype, brain and bone definitely have a complex and intimate relationship (Masi, 2012).

With regards to the correlation between bone and lipid homeostasis, numerous genes identified from the present study support the notion that fat and bone metabolism are connected. This relationship has been investigated, however, the underlying mechanisms that connect bone and fat remain equivocal (Adami et al., 2004; Holecki & Wiecek, 2010). It has been assumed that adipose tissue exerts independent effects on bone remodelling by releasing a number of biologically active substances, or through the local interaction between osteoblast and adipocyte differentiation, or via modulating the vascularisation in bone microenvironment (Holecki & Wiecek, 2010). Observations on the lipid and bone phenotypes caused by mutated APOB, for instance, provide clues indicating hyperlipidemia modifies osteoblast or osteoclast activity (Tintut, Morony, & Demer, 2004).

Genes under the endochondral ossification category, including POSTN, GAS6, GAS7, COL4A1, COL4A2 and COL27A1 mainly act on the formation of ECM and chondrogenesis. But some of them
such as *Postn* and *Gas6* may modulate osteoblastogenesis or osteoclastogenesis directly (Gerbaix et al., 2015; Y. S. Nakamura et al., 1998). Furthermore, *Postn* has been associated with bone repair and response to mechanical forces including loading and unloading (Bonnet, 2014; Bonnet, Brun, Rousseau, Duong, & Ferrari, 2017). The role of *Postn* in bone remodelling via response to forces is in agreement with the observation that this gene may modulate both trabecular and cortical bones.

Several genes associated with the microtubule and cytoskeletal network have given us another perspective that links microtubule assembly to bone homeostasis. In fact, studies have shown that microtubule assembly plays an essential role in the maintenance of postnatal bone mass by regulating both osteoblast and osteoclast functions in bone (H. B. Liu et al., 2011; M. Zhao et al., 2009). *WDR35* has been associated with ciliary dysfunction which may lead to the cranioectodermal dysplasia, an autosomal-recessive disease that is characterised by craniosynostosis and ectodermal and skeletal abnormalities (Gilissen et al., 2010). The prominent skeletal abnormalities imply an important association between ciliopathies and bone homeostasis.

Novel genes including *SETBP1*, *HIC1*, *ZBTB7C*, *DNAH2*, *DYM*, *JHY* and *TNFRSF13B* require further verification. Real time RT-PCR is a widely used quantitative approach in bone research (Frank et al., 2002). We applied this approach to assess the expression patterns of murine *Setbp1* in osteoclasts and osteoblasts cultured from C57BL/6 mice. The expression profiles of both marker genes (*Ocn* and *Trap*) in our experiments are in accord with prior studies (W. Huang, Yang, Shao, & Li, 2007; N. C. Walsh et al., 2003). The expression of *Setbp1* was downregulated during osteoclast differentiation, indicating a potential role for this gene in determining cell fate decisions in myeloid cells (Makishima et al., 2013; Piazza et al., 2013). Its expression did not show significant change during osteoblastogenesis, indicating that *Setbp1* may not be regulated during osteoblast differentiation. Rather, this gene may directly or indirectly affect the bone formation at the post-transcriptional level. Western blotting would be the next key step to characterise the *Setbp1* protein (Mahmood & Yang, 2012). Likewise, real time RT-PCR experiments can be performed for other candidate genes to determine the expression of these genes during osteoblastogenesis or osteoclastogenesis.

Application of bioinformatics tools helps understand the correlation between gene and bone. Such tools not only provide clues on how genetic products, in particular from those novel genes, act on bone through direct or indirect pathways, but also offer a view of how those proteins are connected. Such information is valuable because it may guide us to decide the next steps to clarify the roles of novel genes in bone formation and remodelling. The PPI network of EN1, for example, offers hints that link this protein to pathways that is has not previously been associated with (J. A. Mitchell, Chesi,
et al., 2016; Zheng et al., 2015). Furthermore, to apply those bioinformatics tools may help develop anti-osteoporotic drugs by using ‘in silico pharmacology’, in which computer models are developed to model a pharmacologic process (Chakraborty, Doss, Chen, & Zhu, 2014; Sudha, Nussinov, & Srinivasan, 2014).

It is evident that mutations of SETBP1 cause bone phenotypes as part of the constitutional manifestations of Schinzel-Giedion syndrome (Carvalho et al., 2015; Hoischen et al., 2010; Suphapeetiporn et al., 2011). Furthermore, proven links derived from PPIs strongly suggest that SETBP1 is implicated in bone indirectly via various pathways such as JNK and TGF-β signalling (Hwang et al., 2013). SETBP1 may mediate an anti-apoptotic effect of RANKL in osteoclasts via JNK/c-Jun signalling (Ikeda et al., 2008). Taking account of both expression results and PPIs, this gene is probably mediating osteoclast differentiation and apoptosis at the transcriptional level, and mediating osteoblast formation at the post-transcriptional level. Further study is required to delineate the underlying mechanisms/pathways of this gene related to bone metabolism.

In summary, the abundance of genes identified through haplotype mapping various traits in CC mice has demonstrated the potential of the CC in systemic genetics. The CC is a powerful platform for the discovery of novel genes for a complex disease such as osteoporosis. Applying bioinformatics tools enriches the knowledge about potential or proven genes responsible for osteoporosis, from looking at the sequence, structure, interactions and relevant pathways.
Chapter 8
General Overview and Future Directions
8.1 Overview of Thesis and General Discussion

Osteoporosis is a common systemic skeletal disease associated with low BMD and micro-architectural deterioration of bone tissue leading to an increased risk of fractures (Ralston & Uitterlinden, 2010). It is a complex disease with contributions from multiple genetic loci and multiple environmental factors and their interactions.

Osteoporosis susceptibility genes are involved in many biological pathways such as the classic RANK/RANKL/OPG, Wnt, endochondral ossification and BMP/TGF-β pathways (Clark & Duncan, 2015). GWAS has been used successfully to identify more than 70 osteoporosis susceptibility genes with GWS, along with another approximately 160 genes that have been suggested to be correlated with BMD or osteoporosis by at least one study via various genetic approaches including GWAS, GWLA, candidate gene association studies and WGS (Ralston & Uitterlinden, 2010; Richards et al., 2009). These total 230 genes are estimated to explain up to 30% of genetic variances of BMD or osteoporosis. Therefore, the vast majority of heritability of BMD or osteoporosis is missing and requires further understanding of the bone genetics.

The CC is a large-scale project for generating RI strains from eight parental strains which were carefully selected; these founder strains are estimated to represent near 90% of genetic variance in the mouse genome (Threadgill & Churchill, 2012). Considering the advantages of animal models over studies (e.g., GWLA, GWAS) in the human population, such as fast breeding, controlled environment, and potential for genetic manipulations, the CC has inherited these advantages, and provided a phenomenal platform for biomedical research (Elefteriou & Yang, 2011). There was one study published in which CC mice were utilised to identify novel genes for bone microarchitecture during the screening process of our study (Levy et al., 2015).

In the present study, over 900 Geniad mice (CC subpopulation) across 70 strains incorporating varying ages and genders were dissected and scanned using microCT. Strains with less than two mice per gender were excluded in the analysis after generating data on parameters including BV/TV, Tb. N, Tb. Sp, Tb. Th, SMI, DA and Ct. Th from reconstructed femur images. The CC mice presented large variations in trabecular traits including BV/TV, Tb. N, Conn. D and DA. Other trabecular traits including Tb. Sp, Tb. Th and SMI, and the sole cortical parameter Ct. Th appeared to be less varied in the present study. BV/TV was considered the main parameter for bone mass, and DA was used to determine bone microarchitecture.

In total 848 mice (strains N = 61) were analysed, many recently bred or retired strains were not included in the present study because the selection of strains was random, largely based on the stock
we had. If more strains and mice were included, a better prediction of gene mapping would be produced. For the gene mapping purpose, we categorised mice within strains into four groups and ranked them for each parameter. In comparison with Levy et al. paper, our study included larger numbers in strains and mice, and age range (Levy et al., 2015).

QTL haplotype mapping was performed to identify candidate genes responsible for bone mass and microarchitecture. Over 300 genes were pulled out from 20 loci that contained candidate genes for bone traits. We harvested six loci that achieved the 95\textsuperscript{th} percentile GWS from DA in female CC mice. The strongest DA-derived locus at chromosome 18 yielded genes including Setbp1, Mbd1, Nfatc1, Dym, Spp1, Skor2, Zbtb7c, Sall3, Loxhd1, Mapk4, Smad2 and Smad7. A locus at chromosome 1 yielded Brinp2 and Pappa2. DA in the female cohort also pulled out Col4a2 and Gas6 on chromosome 8, and Notch4 and Egfl8 on chromosome 17. In addition, DA in males identified Cd33, Dkk1, Atf5 and Ano5 at chromosome 7, and Wnt3, Wnt9b, Axin2, Gh1, Smurf2, Map2k6 and Map3k14 on chromosome 11. Trabecular BV/TV and Tb. N in the old female cohort identified candidate genes Dnah2 and Hic1 from a locus that achieved 95\textsuperscript{th} percentile GWS. QTL haplotype mapping of female osteopetrosis or BV/TV in middle aged females identified a locus harbouring genes including Tnfrsf13b, Per1, Alox15b, Alox3, Alox12, Alox15 and Traf4. QTL scan of BV/TV also pulled out genes including Clock, Nmu, Rest, Igfbp7, Kdr, Kit and Tmem165 on chromosome 5 in females, and Cdon and Jhy in the male cohort. Furthermore, mapping of BV/TV or Tb. N in old male populations identified a locus containing Apob, Matn3, Wdr35, Hs1bp3, Rhob and Postn. QTL scan of Tb. Sp in middle aged male CC animals contributed genes such as Setdb1, Arnt, Adamts14, Ankrd35, Txnip, Itga10, Snx27, Ctsk, Ctss, Adamts14 and Notch2. QTL plot of SMI in this cohort also pulled out Postn and Smad9. En1 was identified from haplotype mapping of Tb. Sp in the male cohort.

These genes were then correlated with BMD, BUA and VOA in human datasets (GEFOS and BUA and VOA) by Dr. Benjamin Mullin. The number of genes associated with bone formation and remodelling from published and unpublished data were narrowed down to 86. These analyses showed 10 compelling candidates for bone traits, C11orf63 (JHY), DNAH2, DYM, EN1, HIC1, SETBP1, TNFRSF13B, WNT3, WNT9B and ZBTB7C, reaching gene-wide significance threshold (p ≤ 0.00027778 after Bonferroni correction) or having a top SNP that is above the single-point suggestive threshold (p ≤ 3.00×10\textsuperscript{-5}). In addition, NFATC1 reached the gene-wide suggestive threshold (p ≤ 0.001) with p value 0.000678 but did not reach the gene-wide significance threshold. EN1, WNT3 and WNT9B have previously associated with BMD at the GWS level. NFATC1 has not yet been correlated with vBMD at the GWS level even though its role in osteoclastogenesis has been well established (Yerges et al., 2009). Those seven genes without previous reports were considered novel candidates.
for bone phenotypes. Besides these genes there were over 50 genes that achieved \( p < 0.05 \) in the gene-wide scan. These genes were considered “nominally significant associations” with bone traits. Examples of them include \( \text{CTSK, CTSS, MAPK4, SMAD2, SMAD7, SMAD9, SMURF2, PAPPA2, POSTN, BRINP2, MBD1, SALL3, SKOR2, HS1BP3, ADMA11, CD33, AXIN2, COL27A1, NOTCH2, NOTCH4, EGFL8, SPP1, and LOXHD1} \). Taking account of both analyses in CC mice and human populations, loci at chromosomes 18 and 11 are the two strongest candidates for bone phenotypes. In particular, the locus at chromosome 18 has been associated with 18q deletion syndrome which manifests various defects such as midfacial hypoplasia, delayed growth, and limb and digit abnormalities (Kohlhase et al., 1999; Strathdee et al., 1997). Mutations of any gene within this locus may be associated with this syndrome via various pathways. Furthermore, all osteoporosis susceptibility genes were categorised into several groups based on their functions and pathways; which included RANK/RANKL/OPG, Wnt, BMP/TGF-\( \beta \), endochondral ossification, fatty acid and energy homeostasis, microtubule and cytoskeletal network. These genes can explain part of the missing heritability of BMD or osteoporosis in human population.

To characterise these genes, predicted evidence was collected from a variety of sources including publications, protein analysis, expression profiles, PPI analysis and pathway analysis. Bioinformatic tools were heavily used, including sequence alignment (Uniprot), structure prediction (SMART and Prosite MyDomains), expression profiling (BioGPS), PPI network (STRING), in addition to extra information of pathway analysis (PathCards & PEGG). Evidence either from BioGPS or our real time RT-PCR experiments provides comprehensive snapshots that link genes and bone at transcriptional level. PPI analysis extended the scope of protein from an isolated element to a holistic level.

We have no sufficient evidence to pinpoint any gene that is age-specific, gender-specific or site-specific. Given GPR89, as a member of GPCRs, may be associated with age-dependent changes in skeletal growth, it was not derived from our QTL mapping of the comparison of young and older cohorts (Hsiao et al., 2010). Furthermore, trabecular derived gene \( \text{Postn} \) has been considered the strongest candidate gene for cortical bone growth and modulation (Bonnet et al., 2009; Gerbaix et al., 2015), while cortical derived gene \( \text{Col27a1} \) appears to be more importance in the development of skeletal cartilage rather than cortical expansion (Gonzaga-Jauregui et al., 2015; Jenkins et al., 2005). \( \text{Postn} \) and \( \text{Col27a1} \) may be non-site-specific and play roles in both trabecular or cortical bone. Also, \( \text{Nfact1} \) and \( \text{Setbp1} \) were discovered from haplotype mapping the distal femurs of CC mice, but they have been found to be associated with lumbar spine but not femoral neck.

Even though many animal knockout models mimic the phenotype that presents in human with mutations on the same gene, caution should be paid when we translate those findings from knockout
mouse models to humans. Argument has stated that many animal studies often offer poor prediction on human subjects, in particular in drug testing (Bracken, 2009). The phenotypic discrepancy between mice and humans of ANOS5 and WDR35 mutations exemplifies the fact that functions of certain genes may vary among species, given high percentage of genetic similarity between humans and animals (Mill et al., 2011; J. Xu et al., 2015). PPIs provide an explanation that compensation effects may be induced by the deletion of one gene but other connected genes may compensate the deficits caused by the mutated gene and present no obvious defect morphologically (Carelle-Calmels et al., 2009; El-Brolosy & Stainier, 2017).

Most candidate genes were categorised based on their roles within the signalling pathways or networks. We further speculated genes which showed common effects in bone from same pathways or networks. Endochondral ossification genes such as Postn, Gas6, Gas7, Col4a1, Col4a2 and Col27a1 were thought to be mainly act on the formation of ECM and chondrogenesis (Bonnet, Garnero, & Ferrari, 2016; Chao, Hung, & Chao, 2013; Jenkins et al., 2005; Kuo, Labelle-Dumais, & Gould, 2012; Motomura et al., 2007). Observations on the lipid and bone phenotypes caused by mutated APOB, provide clue indicating hyperlipidemia modifies osteoblast or osteoclast activity, and exemplifies the relationship between lipid and bone homeostasis (Yerges-Armstrong et al., 2013). Besides, certain genes from different pathways, such as EN1, DYM, SETBP1, SETDB1, TMEM165, WNT3 and WNT9B, share roles in patterning, neurogenesis and bone growth, indicating a relationship between bone growth and neurogenesis, probably via diverse pathways (Bilodeau et al., 2009; Hoischen et al., 2010; Lan et al., 2006; Neumann et al., 2006). Brain-bone crosstalk has also been suggested in previous reports and exemplified by gene such as SAT2 or GH1 (Masi, 2012; Rousseaud et al., 2016).

The abundance of genes identified through haplotype mapping various traits in CC mice has demonstrated the potential of CC in systemic genetics. CC is a powerful platform for the discovery of novel genes for a complex disease such as osteoporosis. Applying bioinformatics tools enriches the knowledge about potential or proven genes responsible for osteoporosis, through examining the sequence, structure, interactions to relevant pathways of candidate genes.

8.2 Future Directions

This study has opened up many avenues for future work and improvements. The first is to add more CC mice to the analysis. The second is to expand the size of human populations. The third is to further characterise those novel genes using in vitro experiments and in vivo transgenic or knock out mouse models. Theoretically, the more mice are included in the study, the better mapping prediction that will be achieved. To include more mice in those strains with small numbers or include strains that
have not yet been analysed will not only increase the power of analysis for those loci with low confidence but also produce more significant peaks or loci from the analysis. We have analysed the distal femurs of the CC mice, other anatomic locations such as proximal tibia or spine can be included. Moreover, to increase the population size and ethnic diversity in human datasets or replicate the genes in another human populations may identify more genes associated with BMD, BUA or VOA. The populations for BMD analysis cover 14 Caucasian cohorts and were released in 2015 by GEFOS consortium (Zheng et al., 2015). We could expect more cohorts to be included in their next update of the database. It would be beneficial in the identification of osteoporosis susceptibility genes across ethnic groups once cohorts from both Chinese and African populations are included (Lau et al., 2006; Robbins, Hirsch, & Cauley, 2004; S. M. Xiao, Kung, Sham, & Tan, 2013). In our analysis, human datasets for BUA and VOA included three Caucasian populations (Mullin et al., 2017). To integrate more cohorts, in particular from other ethnic groups could generate more precise prediction in searching osteoporosis genes using quantitative ultrasound (Esmaeilzadeh, Cesme, Oral, Yaliman, & Sindel, 2016; Stieglitz, Madimenos, Kaplan, & Gurven, 2016; Yung et al., 2005). It is important to note that the focus of the current study was gene discovery. Furthermore, analyses in our human populations did not identify any association with several previous reported genes including CTSK, SMAD2, MATN3, and GAS6, reflecting the limitation of using relatively small populations, hence our study may have yielded false-negative results. To analyse these genes in future GWAS or meta-GWAS studies with larger and independent populations at replication stage will clarify the association between bone mass and microarchitecture and candidate genes (Thomas et al., 2009).

To investigate specific candidate genes requires expression profiling via real time RT-PCR, and knockout and/or knock-in animal models. In the present study, RT-PCR experiments were performed and repeated on Setbp1, the first identified strong candidate gene. We did not test the other novel osteoporosis susceptibility genes due to time restraints. This technique could be applied to those genes. We utilised HPRT1 (hypoxanthine phosphoribosyltransferase 1) as the sole housekeeping gene in RT-PCR experiments. For a more accurate normalisation of quantitative gene expression data in differentiating osteoblasts or osteoclasts, two set of genes have been recommended (Stephens, Stephens, & Morrison, 2011). In future gene expression experiments, the geometric average of ACTB, HMBS and HPRT1 will be used for osteoblasts, and the average of ACTB, B2M and HPRT1 will be applied in osteoclastogenesis.

To further delineate the role of candidate genes responsible for bone mass and microarchitecture, animals models with genetic modification are necessary (Davey et al., 2004; Elefteriou & Yang, 2011). To elucidate the function a protein (e.g., SETBP1, TNFRSF13B, EN1), modification of its
expression or function in an animal model (particularly mouse) offers an avenue to measure the effect of such alteration on bone structure, function and metabolism (Davey et al., 2004). There are four approaches to genetic modification in animals, which have been heavily used in the identification of gene functions in both bone modelling and remodelling. These approaches include transgenic (overexpression of the target gene), global gene knockout, conditional knockout and gene knock-in. Each of these approaches has its advantages and limitations. Regardless of the difference among these approaches, manipulation of target gene, more precisely target sequences (particular SNPs with highest effect on bone phenotype) would be efficient for observing the functions of gene in bone. However, globally deleting the gene can lead to prenatal death of the animals or early death postnatally, making investigation of the function of the gene in bone remodelling impossible. Examples for such genes include Hic1, En1, Nfact1 and Setdb1 (Carter et al., 2000; de la Pompa et al., 1998; Kang, 2015; Zheng et al., 2015). Creating viable genetically-modified animals, mostly through transgenic approaches would be beneficial in the investigation of the functional role of the target gene; such practice is exemplified by the application of En1\textsuperscript{lacZ/+} knock-in mice (Zheng et al., 2015). In a similar manner, these gene approaches can be applied in the genetic modification of genes identified in our analyses.

Moreover, to determine which gene contains rare variants of large effect or common variants of small effect requires human studies with populations from diverse ethnic groups and a large number of subjects. Those studies commonly need large collaborative groups across different research centres, such as the identification of a low-frequency non-coding variant near EN1 as a determinant of BMD and fracture using WGS (Zheng et al., 2015). GWAS or WGS is helpful in identifying which candidate gene may contain low-frequency non-coding variants. To investigate the phenotypic effect of each allele could help further determine the association between the gene and bone phenotypes like BMD or osteoporotic fracture.

Eventually, candidate genes can be used as potential therapeutic targets for the development of anti-osteoporotic drugs. Translating a gene found in the laboratory into clinical use is a difficult task. The complexity of the PPIs and how changes in this network function in the determination of phenotype make drug discovery more challenging (Scott, Bayly, Abell, & Skidmore, 2016). The use of bioinformatics tools that help clarify the biological structure and functional domains of a protein in a 3D format make the design of drugs easier and more cost-effective via in silico approach (Sudha et al., 2014; Watkins & Arora, 2015). Further understanding the structure and function of those candidate genes (e.g., Setbp1, Dym, Tnfrsf13b) will help to develop novel drugs that are effectively reversing osteoporosis and its related diseases.
Chapter 9

References


systemic osteopenia from inflammation in cherubism. *J Clin Invest, 118*(11), 3775-3789. doi: 10.1172/JCI35711


Cao, Z., Xiong, J., Takeuchi, M., Kurama, T., & Goeddel, D. V. (1996). TRAF6 is a signal transducer for interleukin-1. *Nature*, 383(6599), 443-446. doi: 10.1038/383443a0


239


242


Howard, G. M., Nguyen, T. V., Harris, M., Kelly, P. J., & Eisman, J. A. (1998). Genetic and environmental contributions to the association between quantitative ultrasound and bone mineral


Huang, W., Yang, S., Shao, J., & Li, Y. P. (2007). Signaling and transcriptional regulation in osteoblast commitment and differentiation. *Front Biosci*, 12, 3068-3092.


259


262


Liu, W., Wang, S., Wei, S., Sun, L., & Feng, X. (2005). Receptor activator of NF-kappaB (RANK) cytoplasmic motif, 369PFQEP373, plays a predominant role in osteoclast survival in part by

---

264


10.1016/j.jbspin.2005.05.009


Nedergaard, M., Takano, T., & Hansen, A. J. (2002). Beyond the role of glutamate as a neurotransmitter. *Nature Reviews Neuroscience, 3*(9), 748-755. doi: 10.1038/nrn916


Ogita, M., Rached, M. T., Dworakowski, E., Bilezikian, J. P., & Kousteni, S. (2008). Differentiation and proliferation of periosteal osteoblast progenitors are differentially regulated by estrogens and


273


scale meta-analysis of genome-wide association studies. Nat Genet, 41(11), 1199-1206. doi: 10.1038/ng.446


281


282


Wellcome Trust Case Control, C. (2007). Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature*, 447(7145), 661-678. doi: 10.1038/nature05911


291


increased trabecular bone volume without increased osteoblast proliferation. *Endocrinology, 141*(7), 2674-2682.


