DEFINING THE TRANSCRIPTOME OF THE OSTEOCYTE NETWORK

A thesis submitted in fulfillment of the requirements for the degree of Doctor of Philosophy
By

Scott Youlten

Under the supervision of

Professor Peter Croucher

Garvan Institute of Medical Research
St Vincents Clinical School
Faculty of Medicine

UNSW Sydney

Surname/Family Name	:	**Youlten**
Given Name/s	:	**Scott Elliot**
Abbreviation for degree as give in the University calendar	:	**PhD**
Faculty	:	**Faculty of Medicine**
School	:	**St Vincents Clinical School**
Thesis Title	:	**Defining the transcriptome of the osteocyte network**

Abstract 350 words maximum: (PLEASE TYPE)

The skeleton is a multifunctional organ-system, providing structural support to the body and maintaining mineral homeostasis through endocrine interactions with distant organs. Balance between these functions is critical to skeletal health and is regulated by a network of cells distributed throughout bone tissue - the osteocyte network. Little is known of the genes with which the osteocyte network performs this specialised function. As a consequence, understanding of its contribution to skeletal disease is very limited. I hypothesised that defining gene expression in the osteocyte network would reveal genes important to its function and provide insights into skeletal disease.

To test this, I developed techniques to perform transcriptome sequencing on the osteocyte network and analysed how gene expression is influenced by skeletal-site, age and sex. I established experimental and analytical strategies to identify a signature of genes enriched for expression in the osteocyte network and reveal molecular processes enabling its specialised function. These genes were examined for their association with skeletal dysplasia and clinically relevant skeletal traits.

This work revealed that gene expression in osteocytes is highly conserved between skeletal sites, with the exception of a limited number of developmental transcription factors differentially active between adult bone types. Dynamic changes in the osteocyte transcriptome during skeletal maturation were also identified, including the sexually dimorphic regulation of genes associated with perilacunar-remodelling. An osteocyte transcriptome signature was defined - 830 genes enriched for expression within the osteocyte network. Enriched expression in the osteocyte network was the first evidence of skeletal involvement for the majority of signature genes, including novel genes with skeletally-restricted activity alongside known osteocyte markers. This work identified a range of signalling pathways significantly enriched in the osteocyte network, including neuron-like network formation pathways upregulated early in osteocytic differentiation. This osteocyte signature is enriched for gene-orthologs known to cause human skeletal dysplasias and influence bone mineral density.

These discoveries identify the genes and molecular processes that define the osteocyte network and demonstrate that specific expression in the osteocyte network may be a powerful filter to identify genes that cause skeletal disease.

FOR OFFICE USE ONLY Date of completion of requirements for Award:

ORIGINALITY STATEMENT

‘I hereby declare that this submission is my own work and to the best of my knowledge it contains no materials previously published or written by another person, or substantial proportions of material which have been accepted for the award of any other degree or diploma at UNSW or any other educational institution, except where due acknowledgement is made in the thesis. Any contribution made to the research by others, with whom I have worked at UNSW or elsewhere, is explicitly acknowledged in the thesis. I also declare that the intellectual content of this thesis is the product of my own work, except to the extent that assistance from others in the project's design and conception or in style, presentation and linguistic expression is acknowledged.’

Signed .....

Date .....

INCLUSION OF PUBLICATIONS STATEMENT

UNSW is supportive of candidates publishing their research results during their candidature as detailed in the UNSW Thesis Examination Procedure.

Publications can be used in their thesis in lieu of a Chapter if:

- The student contributed greater than 50% of the content in the publication and is the "primary author", ie. the student was responsible primarily for the planning, execution and preparation of the work for publication
- The student has approval to include the publication in their thesis in lieu of a Chapter from their supervisor and Postgraduate Coordinator.
- The publication is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in the thesis

Please indicate whether this thesis contains published material or not.

☐ *This thesis contains no publications, either published or submitted for publication*

☐ *Some of the work described in this thesis has been published and it has been documented in the relevant Chapters with acknowledgement*

☐ *This thesis has publications (either published or submitted for publication) incorporated into it in lieu of a chapter and the details are presented below*

CANDIDATE'S DECLARATION

I declare that:

- I have complied with the Thesis Examination Procedure
- where I have used a publication in lieu of a Chapter, the listed publication(s) below meet(s) the requirements to be included in the thesis.

Name	**Signature**	**Date (dd/mm/yy)**

Abstract

The skeleton is a multifunctional organ-system, providing structural support to the body and maintaining mineral homeostasis through endocrine interactions with distant organs. Balance between these functions is critical to skeletal health and is regulated by a network of cells distributed throughout bone tissue - the osteocyte network. Little is known of the genes with which the osteocyte network performs this specialised function. As a consequence, understanding of its contribution to skeletal disease is very limited. I hypothesised that defining gene expression in the osteocyte network would reveal genes important to its function and provide insights into skeletal disease.

To test this, I developed techniques to perform transcriptome sequencing on the osteocyte network and analysed how gene expression is influenced by skeletal-site, age and sex. I established experimental and analytical strategies to identify a signature of genes enriched for expression in the osteocyte network and reveal molecular processes enabling its specialised function. These genes were examined for their association with skeletal dysplasia and clinically relevant skeletal traits.

This work revealed that gene expression in osteocytes is highly conserved between skeletal sites, with the exception of a limited number of developmental transcription factors differentially active between adult bone types. Dynamic changes in the osteocyte transcriptome during skeletal maturation were also identified, including the sexually dimorphic regulation of genes associated with perilacunar-remodelling. An osteocyte transcriptome signature was defined - 830 genes enriched for expression within the osteocyte network. Enriched expression in the osteocyte network was the first evidence of skeletal involvement for the majority of signature genes, including novel genes with skeletally-restricted activity alongside known osteocyte markers. This work identified a range of signalling pathways significantly enriched in the osteocyte network, including neuron-like network formation pathways upregulated early in osteocytic differentiation. This osteocyte signature is enriched for gene-orthologs known to cause human skeletal dysplasias and influence bone mineral density.

These discoveries identify the genes and molecular processes that define the osteocyte network and demonstrate that specific expression in the osteocyte network may be a powerful filter to identify genes that cause skeletal disease.

Acknowledgements

What an epic adventure. The last four years have been some of the best of my life and the people I've had around me have had a lot to do with that.

The first person I have to thank is Peter. I cannot express how lucky I feel to have had Peter as a supervisor and a mentor. His guidance has always centred on my development, not just as a scientist, but as a person. His patience with me has been astounding and his support has been complete. The unwavering integrity of his character and leadership have meant I have been proud to be part of his team; a sentiment I know I share with all those we work with.

The bone biology division has shown me the diversity of what it means to be a good scientist. The creativity of Pauls ideas, connecting knowledge that for most lies in distinct regions of the brain. Johns critical thinking - he'll tell you if your baby is ugly but never to hurt your feelings. Michelles talent, modesty, determination and friendship. The camaraderie and warmth of Julian, Ryan, Marcelo, Weng-Hua, Nancy, Natalie and Sing. All the Bonies. They say if you're the smartest person in the room you're in the wrong room. I've never had any doubt that I am in the right division.

Being in this institute has been an incredible experience. Every day I have been surrounded by interesting people, passionate about what they do, discussing concepts and hypothesising about science and life. These people have taken time to help me, shared with me this experience and become my friends. These people are why coming here has never felt like going to work.

Lastly, I thank my family. My mother and father for supporting me in every way, my brother Ben and sister Stephanie for laughing at and with me, and my fiancé Mariana for making the future even more exciting.

It wasn't easy but man it's been fun.

Related Publications

Kemp, J. P., Morris, J. A., Medina-Gomez, C., Forgetta, V., Warrington, N. M., Youlten, S. E., ... & Logan, J. G. (2017). Identification of 153 new loci associated with heel bone mineral density and functional involvement of GPC6 in osteoporosis. *Nature genetics*, *49*(10), 1468.

McDonald, M. M., Reagan, M. R., Youlten, S. E., Mohanty, S. T., Seckinger, A., Terry, R. L., ... & Le, L. M. (2017). Inhibiting the osteocyte specific protein sclerostin increases bone mass and fracture resistance in multiple myeloma. *Blood*, blood-2017.

Morris, J. A., Kemp, J. P., Youlten, S. E., Laurent, L., Logan, J. G., Chai, R., ... & Sergio, C. M. (2018). An Atlas of Human and Murine Genetic Influences on Osteoporosis. *bioRxiv*, 338863. (Accepted at Nature Genetics)

Gregson, C. L., Newell, F., Leo, P. J., Clark, G. R., Paternoster, L., Marshall, M., ... & Bassett, J. D. (2018). Genome-wide association study of extreme high bone mass: Contribution of common genetic variation to extreme BMD phenotypes and potential novel BMD-associated genes. *Bone*.

Youlten, S. E. & Baldock, P. A. (Accepted pending corrections). Using mouse genetics to understand human skeletal disease. Submitted to *Bone*.

Youlten, S. E., Kemp, J. P., Logan, J., Leitch, V. D., Chai, R. C., Morris, J. A., ... & Croucher, P. I. (In preparation). An osteocyte transcriptome map identifies the molecular landscape controlling skeletal homeostasis and disease. *To be submitted*.

Table of Contents

Table of Figures

Table of Tables

Abbreviations and Symbols

Abbreviation or symbol	Name and/or definition
ANOVA	Analysis of variance
ATP	Adenosine triphosphate
BIC	Bayesian information criterion
bicor	Bi-weight midcorrelation
BMC	Bone mineral content
BMD	Bone mineral density
BV/TV	Bone volume as a fraction of total volume
CI	Confidence interval
CO_2	Carbon-dioxide
DO	Disease ontology
DXA	Dual-energy X-ray absorptiometry
FACS	Fluorescence-activated cell sorting
FDR	False discovery rate
FPKM	Fragments per kilobase per million mapped reads
GMM	Gaussian mixture model
GO	Gene ontology
GWAS	Genome wide association study
KDE	Kernal density estimate
KEGG	Kyoto encyclopedia of genes and genomes
kV	Kilovolt
LFC	Log_2 fold-change
lncRNA	Long non-coding ribonucleic acid
mA	Milliamp
MGI	Mouse genome informatics
mL	Millilitre
mm	Millimeter
MP	Mammalian phenotype
mRNA	messenger RNA
MSC	Mesenchymal stem cell
n	Number of biological replicates
N_2	Nitrogen
PC	Principal component
PCA	Principal component analysis
poly-A	Poly-adenylated
R	Coefficient of Pearson correlation
R^2	Coefficient of determination
rcf	Relative centrifugal force

Abbreviation or symbol	Name and/or definition
RNA	Ribonucleic acid
RNA-seq	Ribonucleic acid sequencing
rRNA	Ribosomal ribonucleic acid
RT-qPCR	Reverse transcription quantitative polymerase chain reaction
SD	Standard deviation
Tb.N	Trabecular number
Tb.Sp	Trabecular spacing
Tb.Th	Trabecular thickness
TGF-Beta	Transforming growth factor beta
WGCNA	Weighted gene coexpression network analysis
μ	Mean
µCT	Micro computed tomography
µL	Microlitre

1 INTRODUCTION

1.1 Background

1.1.1 The skeleton is controlled by the osteocyte network

The skeleton is alive and constantly changing. From the top of the skull to the tips of the phalanges, the more than 200 bones that make up the healthy human skeletal system are continually optimised to protect and support the rest of the body. Bones in different skeletal sites perform specialised functions, from the levers in the legs used for locomotion, to the protective plates that form the skull. These diverse demands are met by specialised morphologies which are established by distinct developmental processes (Karsenty & Wagner, 2002). Flat bones, such as those that make up the skull, are formed by intramembranous-ossification. In this process bone is formed directly from sheets of osteogenic cells that secrete a collagen template which is then hardened by the incorporation of mineral salts. On the other hand, long bones such as those in the limbs are formed by endochondral-ossification. Bones made this way are first shaped in cartilage which is then replaced with bone tissue and refined by the concerted action of a number of bone-cell types. These bone-cells include the bone-forming osteoblasts, the bone-resorbing osteoclasts and the bone-regulating osteocytes, which coordinate the action of these other cell types to optimise the shape and composition of skeletal tissue once formed (Dallas, Prideaux, & Bonewald, 2013). The mineral from which bone is made also serves as a store of the body's calcium and phosphate, critical chemicals used by a range of other tissues and organs (Civitelli & Ziambaras, 2011). Thus, for the body to be healthy, a balance must be struck between the skeletons' physical requirements and endocrine needs. Our understanding of the genes that are important to maintaining this equilibrium is limited. However, in recent years the osteocyte has been established as the pivotal cell type regulating this balance (Dallas et al., 2013; Schaffler, Cheung, Majeska, & Kennedy, 2014).

Osteocytes arise from osteogenic precursors, mesenchymal stem cells (MSCs) committed to making up the bony tissue of the body (Dallas & Bonewald, 2010; Gegenbaur, 1864). These precursors first become osteoblasts on the bone surface, before some of these cells (~10-30% in humans) become buried in the extracellular collagen-mineral matrix and differentiate into osteocytes (Franz-Odendaal, Hall, & Witten, 2006). Bone-embedded

osteocytes are the most numerous cells in bone, making up >90% of cells, and can live many decades within the hard-mineral tissue of the skeleton (Bloch, Kristensen, & Sørensen, 2012; Frost, 1960). Evidence suggests this process is far from a passive-burial; however, how osteoblasts are "chosen" to become osteocytes and whether this is a cell-intrinsic mechanism or external cue is yet to be established (Dallas & Bonewald, 2010).

Osteocytes form an interconnected sensory cell network distributed throughout bone tissue. As cells transition from osteoblasts to osteocytes they progress through a number of morphologically distinct cell stages, marked by the projection of dendritic processes from the osteocyte cell-surface (Figure 1 A) (Franz-Odendaal et al., 2006). These dendrites permeate the bone tissue, forming connections between osteocytes, cells on the bone surface and in the marrow space (Figure 1 B-D) (Doty, 1981; Palazzini, Palumbo, Ferretti, & Marotti, 1998). These contacts form the vast osteocyte network, 23 trillion connections between ~42 billion osteocytes dispersed throughout the human skeleton (Buenzli & Sims, 2015).

Like other cell networks, such as that in the neurons of the brain and the vascular system, the osteocyte network is an efficiently organised multicellular structure (Kollmannsberger et al., 2017). The non-random arrangement of cell contacts is important to a range of processes, from intercellular communication to nutrient transport and waste removal (Schaffler et al., 2014). As a result, the fidelity of network formation is important for healthy skeletal function, with decreased osteocyte density associated with increased fracture and bone micro damage in humans, and decreased connectivity shown to increase bone-fragility independent of bone-structure changes in mice (Dole et al., 2017; Kerschnitzki et al., 2013; Milovanovic et al., 2015; Qiu, Rao, Palnitkar, & Parfitt, 2003). The interconnectivity also means the function of the osteocyte network is intimately coupled to the *in-situ* environment. Osteocyte differentiation and network formation coinciding with burial in the mineralised bone matrix has presented a challenge in their isolation and study using *ex vivo* systems. Thus, while it is clear that the three-dimensional structure and dendrite contacts are critical to osteocyte network function, the genetic cues and biological mechanisms by which they are is formed are largely unknown.

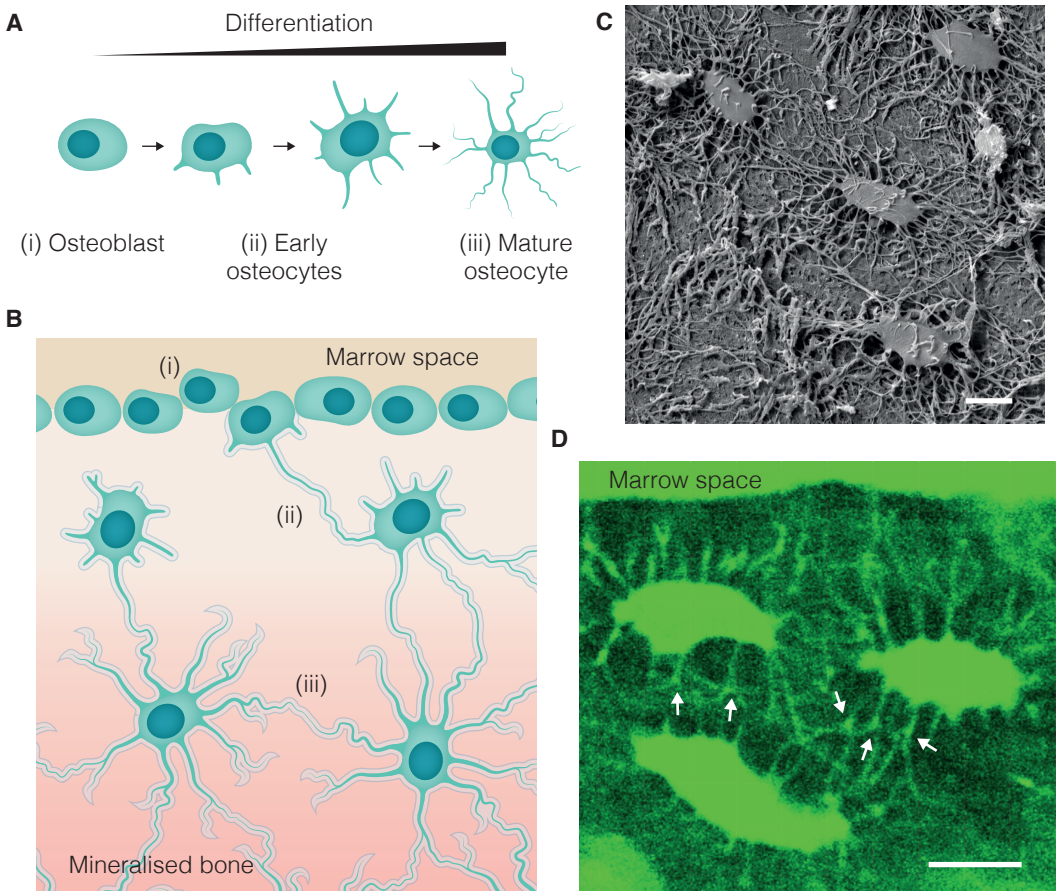


Figure 1 - The interconnected osteocyte network.
(A) Osteocytes differentiate from osteoblasts (i) transitioning through distinct morphological stages and forming dendrites (ii) before reaching maturity (iii). (B) Osteocyte differentiation coincides with network formation and burial in bone tissue. Osteoblasts (i) on the bone surface become early osteocytes (ii) as they form contacts with surrounding cells and are buried in osteoid, maturing as bone tissue becomes mineralised (iii). (C) Acid-etched scanning electron micrograph of the osteocyte lacunar-canalicular system provided by Professor Duncan Bassett. (D) *In-vivo* image of the osteocyte network highlighting contacts between neighboring cells (arrows), provided by Dr Michelle McDonald.

1.1.2 The osteocyte network is the maestro of skeletal maintenance

The structural integrity of the skeleton is critical to its ability to support and protect the body. The skeleton needs to be strong enough that it can withstand the strains of daily life, but also light to enable movement (Seeman & Delmas, 2006). This requires the precise distribution and continual maintenance of bone tissue, achieved through the process of bone-remodelling, where damaged bone is removed, replaced and reinforced (Lanyon, 1993). This process is coordinated by osteocytes as they sense and respond to mechanical

strain (Bellido, 2015). Osteocytes reside within the hard tissue of bone in a fluid filled system of pores, called lacunae, and tiny channels, called canaliculi, containing the cell body and dendrites respectively (James & Steijn-Myagkaya, 1986; Sano et al., 2015). Mechanical loading on the skeleton in the form of movement or weight-bearing compresses bone tissue, deforming the osteocyte network and moving fluid through the lacunae-canalicular system (Y. Han, Cowin, Schaffler, & Weinbaum, 2004; Klein-Nulend, Bakker, Bacabac, Vatsa, & Weinbaum, 2013). Osteocytes sense this deformation and fluid flow with proteoglycan-tethers that connect their cytoskeleton to the extracellular matrix, accentuating their sensitivity to shear stress (Y. Han et al., 2004). These forces alter the expression and activity of important skeletal biochemical pathways, used to transduce signals across the osteocyte network and communicate with effector cell types to regulate bone formation and resorption (Klein-Nulend et al., 2013; Schaffler et al., 2014).

Communication in the osteocyte network occurs through a range of direct and indirect mechanisms. The presence of transmembrane protein channels, or 'gap-junctions', at the contacts between dendrites allows direct chemical communication between osteocytes via the passage of cations, nucleotides and other small molecules to adjacent cells (Doty, 1981). These chemicals are thought to transduce mechanical signals through the osteocyte network, propagating the skeletal response to loading and damage, however much of the data supporting this function is derived from *in-vitro* osteocytic cell systems (Cheng et al., 2001; Yellowley, Li, Zhou, Jacobs, & Donahue, 2000). Gap-junctions have also been observed between osteocytes and a range of other cell types, implying that they too can communicate directly with the osteocyte network, although signals that coordinate the formation and function of these heterotypic juxtacrine interactions are not well understood (Palazzini et al., 1998).

Paracrine signalling is a critical mode of intercellular communication used by the osteocyte network to coordinate other cell types in the process of bone remodelling. In response to bone loading osteocytes secrete a range of bone anabolic factors, increasing bone formation by osteoblasts via small molecules (e.g. nitric oxide), growth factors (e.g. Insulin Growth Factor 1) and hormone-like molecules (e.g. prostaglandins) (Blackwell, Raisz, & Pilbeam, 2010; Lean, Jagger, Chambers, & Chow, 1995; Pitsillides Aa et al., 1995). In addition, mechanical-loading suppresses the expression of a number of osteocyte-

secreted inhibitors of bone anabolic pathways. The Wnt-signalling pathway via Beta-catenin is a powerful driver of osteoblast differentiation and bone formation in osteoblasts (Bonewald & Johnson, 2008). In the absence of loading, osteocytes secrete inhibitors such as Sclerostin (*Sost*) and Dickkopf1 (*Dkk1*), which bind lipoprotein-related receptors to suppress Wnt-signalling (Robling et al., 2008). Both these genes are highly expressed in mature osteocytes relative to other bodily tissues and stages in the osteogenic lineage, and their deletion in osteocytes produces a powerful bone anabolic response in osteoblasts (C. Lin et al., 2009; Poole, 2005). As such, they are among a handful of markers of osteocyte differentiation and demonstrate the orchestration of effector skeletal cell types by the osteocyte network.

Osteocytes also recruit osteoclasts to resorb bone in response to overuse and disuse. Repeated mechanical strain on the skeleton can result in damage to the bone tissue, rupturing osteocyte dendrites and triggering apoptosis (Cardoso et al., 2009; Verborgt, Gibson, & Schaffler, 2000). Similarly, osteocyte apoptosis increases in models of skeletal unloading, although the mechanism by which programmed cell death is triggered is not well understood (Cabahug-Zuckerman et al., 2016). In both cases, osteocyte apoptosis has been shown to preclude osteoclast recruitment to sites of resorption, and inhibition of apoptosis has been shown abrogate this effect (Aguirre et al., 2006; Cabahug-Zuckerman et al., 2016). The most accepted molecular model used to explain this osteocyte-death-driven osteoclast recruitment involves the secretion of the receptor activator of nuclear factor κ-B ligand (*Tnfsf11*/RANKL) (Bellido, 2015). RANKL stimulates the differentiation of osteoclasts by binding to the Receptor activator of nuclear factor κ B (RANK) receptor on the surface of precursors (Boyce & Xing, 2007). Dying cells trigger their network-neighbours to start secreting RANKL, in the process upregulating a number of apoptotic signals including *Hmgb1*, *Bcl2*, *Panx1*, *P2rx7*, *Casp3*, *P2ry2*, and *Cx3cl1* (Cheung et al., 2016; Kennedy, Laudier, Majeska, Sun, & Schaffler, 2014; Wiren, Toombs, Semirale, & Zhang, 2006; Yang et al., 2008). This prompts the cycle of bone remodelling, first removing and then replacing bone at sites of skeletal fatigue. While osteocytic RANKL secretion in this process is very likely to be an important factor, contrasting reports concluding it is both integral and dispensable to the remodelling process suggest more factors may be at play (Cabahug-Zuckerman et al., 2016; Plotkin et al., 2015).

A potentially important difference in these studies examining the osteocytes role in remodelling were the bone types used, one using the femur and the other the vertebrae (Cabahug-Zuckerman et al., 2016; Plotkin et al., 2015). Distinct developmental processes shape the bones that make up the skeletal system which are then maintained by distinct progenitor pools, experience differences in mechanical loading and as a result undergo different rates of remodelling (Leucht et al., 2008; Wand et al., 1992). For osteogenic progenitors and cells lining bone, skeletal site has been shown to be an important factor in bone-healing, a process in which remodelling plays an extensive role (Leucht et al., 2008; Rux et al., 2016; Rux & Wellik, 2017). While the osteocyte network plays a master regulatory role in skeletal remodelling involving communication with these other cell types, how they are affected by skeletal site and bone type is poorly understood.

The osteocyte network can also regulate bone tissue directly in a process termed perilacunar-remodelling (Jähn et al., 2017; Qing et al., 2012). In response to endocrine mineral demand, osteocytes secrete acid into the perilacunar system, dissolving the surrounding bone matrix (Jähn et al., 2017; Qing et al., 2012). Relative to osteoclast-mediated resorption this process provides rapid access to the calcium and phosphate that is locked away in skeletal tissue. This is best known to occur during lactation, liberating the calcium needed for milk production in response to local endocrine signals (Teti & Zallone, 2009). Beyond lactation, osteocytic bone resorption also plays an important role in periods of low mineral consumption, such as hibernation, and contributes to hypo-mineralisation diseases such as Rickets (Bélanger, Jarry, & Uhthoff, 1968; Haller & Zimny, 1977; Marie & Glorieux, 1983). More recently, glucocorticoid suppression of perilacunar-remodelling has been associated with subchondral bone degeneration, and disruption of genes and pathways thought to be involved shown to affect bone quality and strength independent of bone shape (Dole et al., 2017; Fowler et al., 2017; Tang, Herber, Ho, & Alliston, 2012). These studies demonstrate the importance of osteocytic bone resorption to maintaining skeletal homeostasis in response to diseases, genetic perturbations, pharmaceutical agents and contexts of extreme mineral demand. However, the contribution of perilacunar-remodelling to normal skeletal development and mineral balance remains unclear.

1.1.3 The influence of the osteocyte network is not limited to the skeleton

The influence of the osteocyte network extends far beyond the skeleton. In its control of bone composition, the osteocyte network balances local and systemic demands for the minerals from which bone is composed (Dallas et al., 2013). Both calcium and phosphate, the major constituents of bone mineral, are potent biological molecules critical to protein-synthesis, energy dynamics, intra and intercellular signalling in a range of bodily tissues (Civitelli & Ziambaras, 2011). Consequently, the homeostasis of calcium and phosphate available to the body is tightly regulated through the coupled action of the osteocyte network and other organs (Dallas et al., 2013). The osteocyte network is capable of secreting and sensing endocrine factors to liberate these chemicals into the bloodstream or lock them away in bone tissue. Fibroblast growth factor 23 (*Fgf23*) is one such osteocytic endocrine factor, targeting the kidney in the regulation of phosphate metabolism (Lavi-Moshayoff, Wasserman, Meir, Silver, & Naveh-Many, 2010; Liu et al., 2008). Regulating the expression of *Fgf23* are Dentin matrix protein 1 (*Dmp1*) and Matrix extracellular phosphoglycoprotein (*Mepe*), genes with osteocyte restricted patterns of expression (Igarashi, Kamiya, Ito, & Takagi, 2002; Liu, Rowe, Vierthaler, Zhou, & Quarles, 2007). *Dmp1* is commonly used as a marker of early osteocyte differentiation as its expression is markedly upregulated in osteogenic cells at the point they are embedded in bone and integrate into the osteocyte network (Toyosawa et al., 2001). In addition to their influence on systemic mineral homeostasis, *Dmp1* and *Mepe* are also thought to control mineralisation in response to skeletal strain (Gluhak-Heinrich et al., 2003; Harris et al., 2007; Reijnders et al., 2013). Consistent with their dual roles, mutation of *DMP1* causes hypophosphatemic rickets and over expression of *MEPE* is associated with oncogenic hypophosphatemic osteomalacia affecting bone strength in humans (Feng et al., 2006; P. S. N. Rowe et al., 2000). These genes epitomise the complex regulatory role performed by the osteocyte network, integrating mechanical and hormonal signalling to balance conflicting demands of skeletal integrity and nutrient availability.

In addition to mineral regulation, the osteocyte network influences a range of extraosseous tissues and cell types. Genetic ablation of osteocytes results in loss of adipose stores, the deletion of primary lymphoid tissues and a failure to mobilise hematopoietic stem cells, suggesting they have a role in regulating these tissues (Asada et al., 2013; Sato et al., 2013). Moreover, deletion of the G-protein coupled receptor Gs-Alpha

(*Gnas*) specifically in osteocytes induces severe osteopenia and a dramatic expansion of cells of the myeloid lineage, indicating the influence and communication of the osteocytes-network with a variety of cell types in the bone marrow compartment (Fulzele et al., 2013). The osteocyte network also produces a number of myokine factors, capable of stimulating muscle formation and function, however the extent of intercommunication between these two tissues is only now starting to be revealed (Brotto & Bonewald, 2015; Hesse, 2016; Laurent et al., 2016). These examples demonstrate the wide-reaching influence of the osteocyte network, synthesising, secreting and sensing biological signals with consequences for systemic processes and distant tissues. However, beyond these molecules, little is known of the processes and pathways used by osteocytes to perform their highly specialised roles. Unbiased approaches to understanding genes directing the formation and function of the osteocyte network have the potential to reveal novel regulatory mechanisms within the skeleton and beyond.

1.1.4 Transcriptomic studies of the osteocyte network

The last thirty years saw the development of several powerful technologies for unbiased gene discovery. Gene array and next generation sequencing techniques enable the system-scale examination of gene expression which has revealed genes and pathways important in both health and disease (Nagalakshmi et al., 2008; Schena & Shalon, 1995; Wang, Gerstein, & Snyder, 2009). In the exploratory setting, these technologies have numerous advantages over low throughput approaches to measuring gene expression such as reverse transcription quantitative polymerase chain reaction (RT-qPCR). A core strength is that far less needs to be assumed about the expected effect on gene expression prior to experimental manipulation (Morozova, Hirst, & Marra, 2009). Primer-directed candidate gene approaches such as RT-qPCR require expected response genes to be preselected for primer design, guiding, or perhaps limiting, discovery to established pathways and processes. Gene-arrays limit gene selection bias in data collection by measuring the expression of more than 10,000 genes in a single assay through their hybridisation to sequence specific probes (Schena & Shalon, 1995). While the inclusion of probes to which transcripts hybridise to be measured is still an important technological limitation, this scale enables a system-level view of changes in known and novel pathways.

Transcriptome sequencing technology is distinct from the probe hybridisation gene-array approaches. During the sample preparation process, sequence-independent ligation techniques and universal adaptors are used to capture transcript fragments irrespective of their previous identification (Wang et al., 2009). The RNA-base-sequence in these fragments is then determined by sequential fluorescent-nucleotide incorporation and captured as read data, which can be aligned back to the genome to measure gene expression (Conesa et al., 2016). Despite this capacity for unbiased system-scale discovery, just 18 of the 71,032 gene array and RNA-seq experiments catalogued in the ArrayExpress database are returned with the search term 'osteocyte' (https://www.ebi.ac.uk/arrayexpress/). While the lack of raw-data released following publication contributes to this low number, the fact remains that these technologies have been applied sparingly to the osteocyte network.

Many of the transcriptome studies that have been performed on osteocytes have focused on the identification of genes distinguishing these cells from earlier stages in the osteogenic lineage. St John and colleagues performed a transcriptome profile of the osteogenic cell line IDG-SW3 undergoing transition from osteoblast-like morphology to osteocyte-like mineralising and network forming cells (St. John et al., 2014). This cell line recapitulates many aspects of *in-vivo* osteocyte biology more accurately than other *in-vitro* system and has been shown to accelerate bone formation *in-vivo*, demonstrating its capacity to model many aspects of osteogenic differentiation (Woo, Rosser, Dusevich, Kalajzic, & Bonewald, 2011). This work revealed extensive gene expression and epigenetic changes in the process of osteogenic differentiation, demonstrating the temporal expression profiles of known osteocyte markers consistent with the localisation of their *in-vivo* expression (St. John et al., 2014). However, this approach could not distinguish between genes important to osteocytic specialisation from general gene expression changes that occur as part of cellular differentiation. As such, the specific molecular processes involved in osteocyte differentiation and network formation remain to be identified.

Paic and colleagues compared the expression profile of primary osteoblast and osteocyte cells, collagenase digested from neonatal mouse calvarial bone (Paic et al., 2009). These cell suspensions were then FACS sorted based on marker-protein expression,

with the collagen type 1 *Col1a1*-2.3 promoter and the *Dmp1* promoter used to identify osteoblasts and osteocytes respectively. Messenger-RNA expression (mRNA) was then measured by gene chip array. One hundred and forty-three genes were more than 2-fold upregulated in the osteocyte fraction relative to osteoblasts. These included many previously established osteocyte genes supporting the experimental approach and identified many more with unknown skeletal functions. An important consideration in the interpretation of this study was the separation of osteocytes from their extracellular context before examining their transcriptome. While this restricts the contribution of contaminating cell types to the gene expression measurements reported, the function of the osteocyte network is intrinsically linked to its *in-situ* environment. Thus, approaches requiring disaggregation of the osteocyte network are likely to influence the transcriptome of purified cells, potentially explaining the reported absence of the established osteocyte marker Sost in that study.

Alternative approaches to studying the osteocyte transcriptome without the complete disaggregation of the bone cell network have also been reported. Qing and colleagues performed microarray analysis on collagenase-cleaned bone samples to identify genes upregulated in the osteocyte-network associated with perilacunar-remodelling in response to lactation (Qing et al., 2012). Using similar techniques Wasserman et al showed differences in the transcriptome of trabecular osteocytes in response to single or repetitive loading doses (Wasserman et al., 2013). Ayturk et al 2013 and Kelly et al 2016 published methods for removing extra-skeletal tissue from bone and extracting RNA from the osteocyte network, still intact in their mineral environment without the need for collagenase digestion (Ayturk et al., 2013; Kelly, Schimenti, Patrick Ross, & van der Meulen, 2014; Kelly, Schimenti, Ross, & van der Meulen, 2016). Ayturk and colleagues used this approach to assess the gene expression changes occurring due to low density lipoprotein receptor-related 5 (*Lrp5*) mutations, a Wnt/Beta-catenin signalling receptor known to cause skeletal density diseases when expression is altered in humans (Ai, Holmen, Van Hul, Williams, & Warman, 2005; Ayturk et al., 2013). Kelly et al examined how loading affects gene expression differentially between cortical and trabecular bone, showing sustained changes in Wnt-signalling in cortical samples not observed in trabecular samples. All these investigations highlight the power of unbiased transcriptome analysis technologies to

reveal gene expression changes in the osteocyte network associated with their regulation of bone tissue, sensing mechanical strain and responding to endocrine mineral demand.

Importantly, in the development of their technique Ayturk and colleagues also demonstrate significant differences in the transcriptome of osteocytes in intact tissue compared to enzymatically digested and cultured samples, confirming the close coupled nature of osteocyte gene expression and their *in-situ* environment (Ayturk et al., 2013). However, the effect of limb amputation and the time taken to isolate osteocytes once dissociated from the circulatory system using these techniques was not explored. This may be an important consideration given the interaction between the osteocyte network with distant organs and tissues. Ayturk and colleagues also reported an analysis strategy to control for gene expression from contaminating cell populations to improve estimates of inter-replicate variability, demonstrating how in silico approaches can be used to accommodate samples of less than pure cell populations. This approach identified genes enriched in other cell types to be removed from the analysis, however these investigations did not go on to define genes that are preferentially expressed in osteocytes as distinct from housekeeping genes common to other cell populations.

While the studies by Ayturk et al 2013 and Kelly et al 2016 make important technical developments to measuring the transcriptome of the intact osteocyte network, they share a common limitation in that they both use poly-adenylated (poly-A) transcript selection techniques for their transcriptome-sequencing library preparation. Poly-A selection techniques are used in library preparation to limit the amount of read data attributed to un-poly-A ribosomal and mitochondrial RNA transcripts, which are considered uninformative in most experiments yet constitute over 90% of cell transcripts (Conesa et al., 2016). While effective in rRNA depletion, they also bias gene expression data towards processed protein-coding transcripts, potentially losing one of the core advantages of RNA-seq technology (S. Zhao, Zhang, Gamini, Zhang, & Von Schack, 2018). Alternative approaches which deplete rRNA directly without poly-A selection have been documented which enhance the measurement of unprocessed and non-coding transcripts (Adiconis et al., 2013). To the best of our knowledge, no osteocyte transcriptome sequencing experiments capable of measuring protein-coding and non-coding transcripts have been reported, and

as such the expression of the non-coding transcriptome in the osteocyte network *in-vivo* is poorly defined.

It is also important to consider that these early studies of the osteocyte transcriptome were performed in mice. There are important differences in the skeleton humans and mice, including the lack of haversian remodelling in the latter, a fundemental process in human bone remodelling (Bellino, 2000). Moreover, there are differences in the DNA base sequence, gene expression and epigenetic regulation between species (Breschi, Gingeras, & Guigó, 2017; Yue et al., 2014). The understanding of these differences is critical to translate findings in mice to insight into the biology of the human skeleton. Nevertheless, many important morphological, cellular and molecular characteristics are conserved between species, making mice an important model to understand the human skeletal system (Jilka, 2013).

A key advantage of transcriptome sequencing technologies over gene-array approaches is the capacity to detect and measure the expression of unannotated transcripts and genes (Wang et al., 2009). Returning not just transcript expression measurements but also information about the transcribed base sequence, transcriptome-sequencing reads can be used to identify expression at *loci* in the genome that do not overlap with any known genes. Then, by *de-novo* transcriptome assembly, read data can be bioinformatically pieced together based on overlapping sequences, and novel splice junctions, exon structures and transcripts expressed in transcriptome samples can be identified (Grabherr et al., 2011; Pertea et al., 2015). *De-novo* assembly on other tissues has played a critical role in the development of gene models and major transcriptome annotations such as GENCODE and RefSeq, yet these techniques have not been performed on osteocytes (Mudge & Harrow, 2015; Pruitt et al., 2014). Due to this omission, genes and transcripts most specifically expressed in the skeleton may remain to be assembled and discovered. The capacity for transcriptome sequencing technology to unbiasedly measure both protein-coding and non-coding gene expression, as well as novel genes and transcripts, make it a powerful tool to understand the gene expression landscape contributing to osteocyte defining biology.

1.1.5 The osteocyte network in skeletal pathophysiology

The osteocyte network is critical to skeletal health and yet our understanding of its contribution to skeletal genetic disease is limited. Genetic diseases affecting the skeleton can be roughly separated into two categories, the rare monogenic skeletal dysplasias and common genetically complex skeletal diseases. Most rare severe skeletal dysplasias are caused by a damaging mutation in one, or a limited number of *loci* showing clear patterns of inheritance. More than 300 genes causing over 400 rare severe skeletal dysplasia have been identified, many encoding critical components in skeletally important biochemical pathways (Bonafe et al., 2015). Among those are several genes highly enriched for expression in osteocytes, including *SOST* and *DKK1*, mentioned earlier in connection with their Wnt-signalling inhibitory function contributing the skeletal anabolic response to mechanical loading. In humans, mutations affecting the expression of Sclerostin have been shown to cause Sclerosteosis and van-Buchem Syndrome, diseases characterised by extremely high bone mass and strength (Balemans et al., 2002; Brunkow et al., 2001). Similarly, Wnt-receptor mutations that affect *DKK1* binding have been shown to cause bone accrual diseases (Ai et al., 2005). In both cases, the osteocyte restricted expression pattern within the skeleton has led to the development of bone anabolic antibodies with therapeutic potential (Baron & Hesse, 2012; Cosman et al., 2016; Fulciniti et al., 2009; McClung et al., 2014). These examples suggest the discovery of new molecular process controlling the specialised function of the osteocyte network may provide insight into the molecular aetiology of other skeletal genetic disorders and a rational basis for their treatment.

Much more common than the monogenic skeletal dysplasias are the genetically-complex skeletal diseases such as osteoporosis (Montagnani, 2014). Osteoporosis affects more than 25% of people above 50 years of age in Australia, the United States of America and Europe, making it a significant health and economic burden (Australian Institute of Health and Welfare, 2014; Hernlund et al., 2013; Looker & Frenk, 2015). Osteoporosis is a disease of significantly reduced bone mass which leads to bone fragility, increased fracture rates and premature death (Center, Nguyen, Schneider, Sambrook, & Eisman, 1999). A key diagnostic and prognostic metric of osteoporosis is bone mineral density (BMD), measured in the limbs or spine using ultrasound or x-ray technology (Kanis, 2002). BMD variance is shown to be strongly influenced by genetics, with twin studies estimating over 75%

variance is attributable to heritable factors (Pocock et al., 1987). However, in contrast to the rare monogenic dysplasias, diseases such as osteoporosis are caused by the combined effect of many hundreds and potentially thousands of individual sequence variants, the sum of which determine bone density. To identify *loci* contributing to this variance, more than 30 Genome wide association studies (GWAS) for BMD have been performed (https://www.ebi.ac.uk/gwas/search?query=bmd). These have identified hundreds of genes associated with significant *loci*, however together these variants only account for ~12% of BMD variance (Kemp et al., 2017). This suggests many genes that influence BMD and contribute to the development of bone mass diseases have yet to be identified.

While low BMD diseases such as osteoporosis are commonly attributed to excessive bone loss, the failure to reach optimal bone mass during skeletal maturation is another critical determinant of bone health in advancing age (NIH Consensus Development Panel on Osteoporosis Prevention, Diagnosis & Therapy, 2001). During childhood and adolescence, the skeleton undergoes periods of rapid growth in terms of bone length, shape and mineral composition, accruing close to half of the bone mineral in the adult skeleton in the years around sexual maturation (Baxter-Jones, Faulkner, Forwood, Mirwald, & Bailey, 2011). While this period of sexual and skeletal maturation is a critical time for bone mineral acquisition in both males and females, sex -based differences in peak bone mass are well established. Males tend to have higher peak bone density, a difference not explained by nutrition, level of physical activity, body weight or lean mass (Alswat, 2017). These differences early are thought to explain much of the disparity in skeletal disease incidence between sexes in later life, with three times more women diagnosed with osteoporosis than men (Alswat, 2017). While sex hormones are known to play an influential role in both the accrual and loss of bone, these molecules alone do not explain the sex dimorphism in bone mass, suggesting other molecular factors may be involved (Ferrari, Rizzoli, Slosman, & Bonjour, 1998; Sigal, 1984; Wei & Mao, 2007). While the osteocyte network is known to play an important role in the regulation of bone mass, our understanding of its contribution to skeletal maturation and how its function may be influenced by sex is limited.

The osteocyte networks' highly specialised role regulating the skeleton is well established however many of the genes and molecular processes it uses to do this are not

known. This has limited our understanding of its contribution to skeletal genetic diseases. Given the influence of the osteocyte network on bone mass and mineral regulation, the identification of genes important to this function may also help define its contribution to skeletal disease.

1.2 Hypothesis

The critical role of the osteocyte network regulating skeletal mass and mineral dynamics is well established yet, our understanding of the molecular processes with which it performs these specialised functions is very limited. We know little of how the transcriptome may differ between bone type, is regulated during skeletal maturation, or influenced by sex and thus fundamental questions as to the nature of gene expression in the healthy osteocyte network remain to be answered. Critically, these gaps in knowledge also limit understanding of the contribution of the osteocyte network to skeletal disease.

We hypothesise that defining the osteocyte transcriptome will reveal molecular control programs that dictate the formation and function of the osteocyte network, shedding light on the contribution of the osteocyte-network to skeletal genetic diseases.

1.3 Aims

1. Develop methodologies to examine the transcriptome of the osteocyte network in multiple bone types and understand how technical factors may influence gene expression.
2. Identify the transcriptome actively expressed in the osteocyte network across the skeleton, including protein-coding, long non-coding, known-annotated and unknown-unannotated genes and differences between skeletal sites.
3. Examine the transcriptome of the osteocyte network during skeletal maturation in both sexes to reveal genes, pathways and processes that contribute to the sex-related differences in skeletal regulation.
4. Define a signature of genes enriched for expression in the osteocyte network, and with it, the molecular control programs that dictate network formation and function.
5. Determine the contribution of genes enriched for expression in the osteocyte network to rare skeletal diseases and variation in clinically relevant skeletal traits.

2 METHODS

Details of experimental models, reagents, resources and software are listed in the Key Resources Table (Appendix 4).

2.1 Experimental ethics and animal cohort details

2.1.1 Ethical approval

All optimisation, transcriptome sequencing, micro computed tomography (μCT), dual energy X-ray absorptiometry (DXA) and histological experiments were performed on wild-type, immune-competent mice. The Garvan/St Vincent's Animal Ethics Committee (AEC) approved all animal experiments (Protocol ID 16/01 and 12/44). Mice were maintained in a specific pathogen free facility and group housed (2-5 animals per cage) with continuous access to food and water. None of the mice had noticeable health or immune status abnormalities and were not subject to prior procedures.

2.1.2 Sample delay cohort

In-situ isolated osteocyte samples were collected from the left and right humeri of seventeen 14-week-old male wild-type C57BLKaLwRij mice as per section 2.2.1. Limbs were removed immediately after sacrifice and held at room temperature for <10min (n=6, control), 30min (n=4), 45min (n=4), 60min (n=4), 75min (n=4), 90min (n=4), 120min (n=4) or 240min (n=4) post-sacrifice prior to processing. All extra-skeletal tissues were left intact until the specified time point at which point they were processed for transcriptome sequencing as per section 2.2.4. A total of 34 samples were sequenced to an average depth of ~20 million reads per sample.

2.1.3 Bone comparison cohort

In-situ isolated osteocyte samples were collected from left and right tibiae, femora, humeri and calvariae (halved, suture removed) of eight 16-week-old male C57BL6/NTac mice as per section 2.2.1 (n=16 per bone type, 64 samples total). From each mouse, all samples were collected and processed within 20min of animal sacrifice. Histology and μCT analysis were performed on all samples collected from the right side of the body as per section 2.2.3 (n=8 per bone type, 32 samples total). Transcriptome sequencing was performed on all samples collected from the left side of the body as per section 2.2.4 (n=8 per bone type,

32 samples total). Samples were sequenced to an average depth of ~30 million reads per sample (n=8 per bone type).

2.1.4 Skeletal maturation cohort

Left and right humeri were collected from 4, 10, 16 and 26-week-old female and male C57BL6/NTac mice (n=5 per sample type, 80 samples in total). These ages were chosen as they represent key stages in sexual and skeletal development. In C57BL6 mice, puberty begins after 4 weeks of age and a mature skeleton is reached by 16 weeks of age, in which time the skeleton has achieved peak bone mass (Brodt, Ellis, & Silva, 1999; Dutta & Sengupta, 2016; Richman et al., 2001). By collecting bone samples at 4, 10, 16 and 26 weeks, we can examine the transcriptome before and after the growth phase and when the skeleton is mature. Mice breeding was stratified so all ages could be collected within a single 36-hour time period. Samples were collected in groups of 8 mice (one from each time point in each sex) to avoid batch effects. All samples were collected within 15min of sacrifice. Intact bones from the right limbs were used for morphological analysis by DXA and µCT as per section 2.2.3. *In-situ* isolated osteocyte samples were collected from the left limb as per section 2.2.1 and taken for transcriptome sequencing as per section 2.2.4. A total of 40 transcriptome samples were sequenced to an average depth of ~25 million reads per sample (n=5 per sample type).

2.1.5 Osteocyte enrichment cohort

Both humeri were collected from five 10-week-old male C57BL6/NTac mice (n=5 per sample type, 10 samples total). All samples were collected within 20min of sacrifice. Bones from the left limb were processed to obtain *in-situ* isolated osteocytes as per section 2.2.1. Bones from the right side had all soft tissue removed while the marrow was left intact. Transcriptome sequencing was performed on a total of 10 samples as per section 2.2.4 which were sequenced to a depth of ~20 million reads per sample (n=5 per sample type).

2.2 Method details

2.2.1 Sample collection and in-situ osteocyte isolation

Mice were sacrificed by CO_2 asphyxiation and cervical dislocation. To isolate cells within bone tissue, we optimised a methodology that has been reported in a number of

publications (Ayturk et al., 2013; Kelly et al., 2014; Mautner, 1997). This was applied to the tibia, femur and humerus where soft tissue including muscle, ligaments and tendons were removed, before the surface was gently scraped with a scalpel to remove the periosteum. Diaphyseal bone from the tibia was isolated by cutting the tibia at the fibula junction and 1mm distal to the proximal growth plate, and from the femur by cutting the bone immediately proximal to the third trochanter and 1mm proximal to the distal growth plate. The humeri were cut immediately proximal to the deltoid tuberosity and then 1mm proximal to the epicondyles before completely removing the deltoid tuberosity along the bone shaft. Bone marrow from each bone was removed by first flushing with PBS until visibly clean and then centrifuging at 14,000rcf for 15sec. Bones were cut into pieces and snap frozen in liquid N_2 for storage.

Calvarial osteocytes were isolated by dissecting the parietal region of the skull, excluding the sutures and scraping soft tissue from the bone surface. Calvariae were separated into left and right bone fragments. Bone samples were cut into 1mm strips and vortexed in a 1.5mL centrifuge tube with 200µL of PBS and then centrifuged at 14,000rcf for 15sec to remove marrow. Bones were then snap frozen in liquid N_2 for storage.

Sample collections were performed with invaluable assistance from my colleagues Dr Julian Quinn, Dr Sindhu Mohanty, Dr Ryan Chai, Jessica Pettitt, Weng Hua Khoo and Amelia McGlade

2.2.2 RNA extraction optimisation

To compare the quality of RNA obtained from phase separation and column extraction techniques, frozen *in-situ* isolated osteocyte samples were homogenised using a Polytron hand held homogeniser (Kinematica PT 1200 E 11010026) in TRIreagent (Appendix 4). RNA was isolated either according to the TRIreagent manufacturers manual phase separation protocol and cleaned with an additional ethanol-precipitation step or using Direct-zol columns (Appendix 4) as per the manufacturers protocol both with and without the on-column DNase step.

To examine the effect of sample freezing and thawing on RNA quality, *in-situ* isolated osteocyte samples were collected from the tibiae and humeri. RNA was extracted using

the TRIreagent protocol described above from fresh bone samples, snap frozen bone samples, snap frozen samples allowed to thaw in PBS and snap frozen samples allowed to thaw in the dedicated sample thawing agent RNAlater-ICE (Appendix 4). RNA integrity in all relevant experiments was assessed using the Bioanalyser RNA 6000 Nano Kit (Appendix 4) while RNA yield was determined using a Nanodrop (Thermo).

2.2.3 Morphological analysis of bone samples

Samples taken for morphological analysis by histology and μCT were fixed overnight in formalin then transferred to 70% ethanol. μCT was performed using a Skyscan Model 1172 μCT scanner (Bruker) at 50kV, 200mA with a 0.5mm aluminium filter at a pixel size of 4.3μm. Images were captured every 0.4° through 180°, reconstructed and analysed using NRecon and CTAn software (Appendix 4). Three-dimensional rendered scans were created using Drishti-2 software (Appendix 4).

Histological examination was performed following μCT scanning. Samples were decalcified in 0.5m EDTA at 37°C for 24hr and embedded in paraffin. 3μm sections (parasagittal plane) were cut on a Leica RM2265 microtome, mounted on Superfrost plus slides (Appendix 4) and stained with Mayer's hematoxylin and eosin (Appendix 4). Images of each section were captured under 10x and 20x objectives using the Aperio Scanscope slide scanner (Leica) which were processed by Aperio Imagescope (Appendix 4) and Fiji/ImageJ software (Appendix 4).

Sections were analysed under 10x objective using OsteoMeasure software (osteometrics). Total cortical area and osteocyte density were assessed for each bone section. In each sample, cells were quantified in the whole length of bone, as standard histological ‘landmarks’ were removed from samples in the osteocyte isolation process. For these calculations, osteocyte lacunae containing visible evidence of nuclei were counted. The viability of osteocytes was assessed morphologically, with osteocytes that occupied a significant fraction of their lacunar space and showed no obvious sign of cell shrinkage, blebbing or other markers of apoptotic cell death pathways considered to be viable. This two-dimensional analysis calculated the density of live osteocytes per μm^2, allowing a comparison between the different bone types. An estimate of three-dimensional osteocyte

density was calculated multiplying the 2D density by the μCT sample volume divided by section thickness.

Histology was used to confirm the efficacy of *in-situ* osteocyte isolation in samples from the bone comparison transcriptome sequencing cohort detailed in section 2.1.3. μCT was used to calculate sample bone volume. Normalised RNA yield was calculated by dividing the RNA yielded from each bone type by the mean bone volume for that sample type. Morphological data collection and analysis was performed by my colleagues Amelia McGlade, Dr Sindhu Mohanty, Alexander Corr and James Smith.

To examine changes in bone structure during skeletal maturation, intact humeri from the skeletal maturation cohort (section 2.1.4) were scanned using DXA. The resulting scans were used to calculate bone length (validated using forceps), bone mineral density (BMD) and bone mineral content (BMC). DXA was performed by Alexander Corr and James Smith. Growth rates between time points were calculated by subtracting the mean of the earlier time point from the later time point and dividing by the number of weeks between time points e.g. BMC growth rate in male samples between 4 and 10 weeks = (mean BMC at 10 weeks - mean BMC at 4 weeks)/(# of weeks age gap i.e. 10-4=6 weeks)). The amount of morphological variance explained and the significance of association with age, sex and the interaction of these two variables was calculated by two-way Analysis of Variance (ANOVA) in GraphPad Prism (FDR ≤ 0.05, Appendix 4).

2.2.4 RNA extraction, transcriptome library preparation and RNA-sequencing

A standardised method was used in the preparation of all transcriptome sequencing experiments. TRIreagent (Appendix 4) was added directly to frozen bone samples and homogenised using a Polytron hand held homogeniser (Kinematica PT 1200 E 11010026). RNA was isolated according to the manufacturers protocol and cleaned with an additional ethanol-precipitation step. RNA yield was determined using a Nanodrop (Thermo) and RNA integrity determined using the Bioanalyser RNA 6000 Nano Kit (Appendix 4).

Total-RNA (250ng) was depleted of ribosomal RNA (rRNA) using RNaseH (Appendix 4) and ribosomal RNA targeting oligonucleotides based on a protocol by Adiconis et al., 2013 (Adiconis et al., 2013). Briefly, total-RNA, spiked with ERCC internal controls (Life

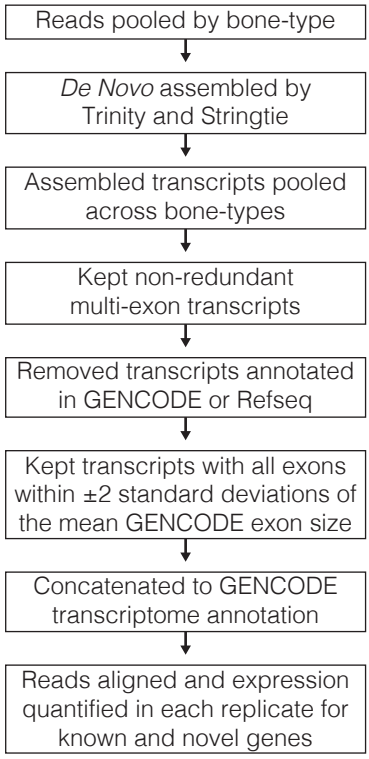


Figure 2 - Filtering approach used to identify novel genes expressed in the osteocyte network.

Technologies 4456739), was incubated with ribosomal-RNA targeting oligos and RNaseH to degrade the rRNAs before the oligos were removed with a DNase treatment (Appendix 4). As an alternative to polyA-capture approaches of rRNA removal, this approach avoids bias towards processed protein-coding transcripts. RNA was repurified using magnetic beads according to the manufacturers protocol (Appendix 4). Total-RNA stranded transcriptome libraries were prepared using the TruSeq Stranded Total RNA LT Sample Prep Kit starting from the fragmentation step in the manufacturers protocol (Appendix 4). 2x125-basepair paired-end sequencing was carried out on the HiSeq 2500 instrument (Illumina) at the Kinghorn Center for Clinical Genomics, Sydney Australia.

2.2.5 De-novo transcriptome assembly and novel gene identification

Transcriptome data was *de-novo* assembled and novel transcripts identified using a bespoke filtering strategy (*Figure 2*). Firstly, read-data was pooled for each bone type in the Bone Comparison Cohort described in section 2.1.3. Next, *de-novo* transcriptome assembly was performed on each bone type using two different assembly strategies: *ab initio*, using Trinity (Appendix 4), and genome-guided, using Stringtie (Appendix 4) (Haas et al., 2013; Pertea et al., 2015). Only multi-exon transcripts assembled by both methods were retained before transcripts assembled in each of the three bones were pooled using Cuffcompare (Appendix 4) to generate a non-redundant union set of assembled transcripts.

Assembled transcripts with splice patterns matching those in RefSeq (Appendix 4) or GENCODE-M5 (Appendix 4) annotations were removed to identify novel transcripts. Remaining transcripts were filtered based on exon length (Mudge & Harrow, 2015; Pruitt, Tatusova, Brown, & Maglott, 2012). Briefly, the mean $\log_2$-exon-length +/- 2 standard

deviation (SD) of GENCODE-M5 annotated exons was calculated. Assembled transcripts containing one or more exons that were longer or shorter than this range were removed. The protein-coding potential of the remaining transcripts was assessed using CPAT software (Appendix 4).

To annotate structures arising from novel *loci* in the genome, those overlapping known annotated transcripts located on the opposite strand were given a "novel_antisense" biotype and given gene IDs beginning with "Obcda", while transcripts located between known genes were given the "novel_intergenic" biotype and assigned gene IDs beginning with "Obcdi". Novel transcripts for both known and novel genes possess transcript IDs begin with "TRINITY". These novel, multi-exon transcripts were then concatenated to the GENCODE-M5 annotation prior to read alignment. This annotation was used for the alignment and quantification of all transcriptome sequencing analyses reported here, including those performed on publicly available data.

Subsequent to this analysis, RefSeq and GENCODE annotations have been updated to include new gene structures, many predicted by computational sequence analysis that have limited evidence as to their expression *in-vivo*. Some of these structures contain splice junctions that overlap novel transcripts reported here, independently supporting their detection in our data. In the case of these overlaps, the GENCODE M13 gene name is contained in brackets beside the unique assigned gene ID.

2.2.6 Transcriptome sample processing

All transcriptome data analysed was processed using the following pipeline. First, transcriptome data were trimmed of low-quality reads and adaptor sequences using Trim Galore (Appendix 4). Next, trimmed data was aligned to the GRCm38.p3 mouse genome guided by the GENCODE-M5 transcriptome annotation (Appendix 4) plus the novel assembled transcripts detailed in section 2.2.5 using STAR software (Appendix 4). Lastly, gene expression was quantified using RSEM software (Appendix 4).

2.2.7 Determination of active genes

Genes actively expressed in a given condition were determined using the following methodology. Sample specific thresholds of gene activity were calculated based on the

bimodal distribution of normalised gene expression, as described in (Hart, Komori, LaMere, Podshivalova, & Salomon, 2013). Briefly, fragments per kilobase per million mapped reads (FPKM) were log_2-normalised (nFPKM), omitting genes with FPKM = 0. The kernel density estimate (KDE) of these values was then calculated (using Scotts rule of thumb for bandwidth) and the maximum KDE value determined. A Gaussian distribution was then fitted, with the mean (μ) at the KDE maximum and the standard deviation (SD) based on normalised expression values greater than μ. The nFPKM values were then transformed to zFPKM using zFPKM = (nFPKM - μ)/SD. Active expression in a sample was defined as those with > -2.6 zFPKM, the conservative range suggested in the original method publication (Hart et al., 2013). This generated sample specific thresholds for gene activity.

Genes were considered actively expressed in a given condition if they were above the sample specific gene activity threshold in all biological replicates. For example, if GeneX was expressed above the sample specific active-expression threshold in 8 or 8 tibia samples in the bone comparison cohort, then GeneX was considered actively expressed in the tibia. Alternatively, if GeneX was expressed above the sample specific threshold in 1-7 tibia samples then GeneX was considered variably expressed in the tibia. If GeneX was not expressed above the activity threshold in any tibia samples then GeneX was considered inactive in the tibia.

Genes were considered to be 'differentially active' if they were actively expressed in one sample type and inactive in another. For example, if GeneX was expressed above the activity threshold in 8 of 8 tibia replicates and 0 of 8 humerus replicates then GeneX was considered to be differentially active between tibia and humeri.

2.2.8 Differential gene expression analysis

Gene-level read counts were subset to genes actively expressed in either condition identified as per section 2.2.7. Differentially expressed genes were identified using the voom-limma analysis pipeline and the topTreat function (Bonferroni adjusted p-value ≤ 0.05, Appendix 4, (Ritchie et al., 2015)).

2.2.9 Principal component analysis

Principal component analysis (PCA) was performed on the gene-level counts of genes actively expressed in any condition (e.g. actively expressed in either tibiae, femora or humeri samples). Counts were scaled, normalised and principal components calculated using the default parameters of the prcomp function (Appendix 4). Samples were clustered based on the first two principal component (PC1 and PC2). The Euclidean distance between groups was calculated based on the mean of the x (PC1) and y (PC2) coordinates for each sample type (sample centroids). The significance of separation between sample types was determined by Hotellings T-test using the ICSNP R-package (Bonferroni adjusted p-value ≤ 0.05, Appendix 4). Fitting of confidence-ellipses and plotting was performed using the ggplot2 R-package (Appendix 4).

2.2.10 Gene Ontology, KEGG and Disease Ontology enrichment analysis

Significantly enriched Gene Ontology (GO) terms were identified using the enrichGO function of the clusterProfiler R-package (Bonferroni adjusted p-value < 0.05, Appendix 4). The gene universe was defined as genes actively expressed in any sample type.

Significantly enriched Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways were identified using the enrichKEGG function of the clusterProfiler R-package (Bonferroni adjusted p-value < 0.05, Appendix 4). The gene universe was defined as with the GO analysis that were also annotated with an ENTREZ id.

Significantly enriched Disease Ontology (DO) terms were identified using the enrichDO function of the DOSE R-package v3.2.0 (Bonferroni adjusted p-value < 0.05, Appendix 4). To do this, human-mouse gene orthologs were identified using the biomaRt R-package (Appendix 4). The gene universe defined as genes actively expressed in any sample type with a human-mouse ortholog annotated with an ENTREZ id.

2.2.11 GO semantic similarity clustering

To identify clusters of semantically similar GO terms, significantly enriched GO 'Biological Processes' results were input into the ReViGO webtool (Appendix 4) which removed redundant terms (similarity > 0.9) and calculated 2D multidimensional scaling (MDS) coordinates based on SimRel semantic similarity algorithm. Distinct clusters of GO terms

were identified based on the MDS coordinates using the mclust function of the mclust R-package (Appendix 4). The optimum number of clusters was selected among models with unequal variance using the Bayesian Information Criterion.

2.2.12 Identification of genes with annotated skeletal effects

To identify genes associated with skeletal biological processes, a list of GO terms directly related to the skeleton was first constructed. To do this, the GO biological process term descriptions was searched using bone-related keywords identifying 116 manually curated terms (Appendix 4). 'Skeletally known' genes were identified as those annotated to any of the skeletal GO terms, as well as to children terms retrieved using "go_parent_term" filter of the biomaRt R-package (Appendix 4).

To identify genes that cause a significant skeletal phenotype when knocked out in mice, a list of mammalian phenotype (MP) terms related to the skeleton was first constructed. To do this, the MP term definitions and descriptions were filtered using bone-related keywords identifying 662 manually curated terms (Appendix 4). Knockout mouse phenotyping data, excluding conditional mutations, in the Mouse Genome Informatics (MGI) database was then screened to identify gene knockout mouse lines associated with skeletal MP terms (Appendix 4).

2.2.13 Gene activity and expression in extra-skeletal organs and tissues

Gene expression in other tissues and organs was assessed using the transcriptome sequencing read data released in association with the study *A circadian gene expression atlas in mammals: Implications for biology and medicine* was obtained and gene expression and activity defined as per the methods section ((R. Zhang, Lahens, Ballance, Hughes, & Hogenesch, 2014), available under the ArrayExpress accession E-GEOD-54652). This high-quality dataset contains 8 replicates of 12 tissues types collected for a single study under controlled conditions. This identified the active genes in 12 extra-skeletal organs and tissues. Data from this study was processed as per section 2.2.6 and actively expressed genes determined as per section 2.2.7.

To identify genes with enriched expression in osteocytes relative to other tissue types, 95% confidence intervals (CI) of the mean normalised expression (FPKM) were calculated for

each tissue type, including osteocytes from each bone type. Genes significantly enriched in osteocytes were identified as those where the lower 95% CI of expression in osteocytes exceeded the upper 95% CI of all other tissues. Principal component analysis comparing the active transcriptome of osteocyte isolated samples and these other tissues was performed as per section 2.2.9.

2.2.14 Transcriptome analysis of Sample Processing Delay Cohort

Transcriptome samples from the sample delay cohort described in section 2.1.2 were processed as per section 2.2.6. Actively expressed genes at each time point were identified as per section 2.2.7. Differentially expressed genes were identified in comparisons between control samples (processed and snap frozen within 10min of animal sacrifice) and delayed samples from each timepoint (30, 45, 60, 75, 90, 120 or 240min post-sacrifice) as per section 2.2.8.

Significantly enriched GO 'Biological Processes' were identified in the differentially expressed between the control and 120min or 240min post sacrifice as per section 2.2.10.

2.2.15 Skeletal Maturation Cohort principal component clustering

Principal component clustering analysis was performed on samples from the Skeletal Maturation Cohort (section 2.1.4) between female samples (any age), male samples (any age) and sexes at each age (4, 10, 16 and 26-weeks-old) as per section 2.2.9. Fitting of 50% CI ellipses and plotting was performed using the ggplot2 R-package (Appendix 4).

2.2.16 Skeletal Maturation Cohort weighted gene co-expression network analysis

To identify clusters of genes with highly correlated patterns of gene expression during skeletal maturation Weighted Gene Co-expression Network Analysis (WGCNA) was performed on the normalised counts of genes actively expressed in either sex at any age (4, 10, 16 or 26-weeks-old) in the Skeletal Maturation Cohort (section 2.1.4) using the WGCNA package (Appendix 4). My colleague Alexander Corr was instrumental in the development of this analysis.

First, the gene-wise 'connectedness' was calculated using the bi-weight midcorrelation function (bicor) across all 40 samples in the skeletal maturation cohort. This bicor measure

of correlation was chosen as it is more robust to outliers compared to mean-based correlation approaches. Next, a soft-thresholding power was chosen based on iterative testing and identification of the power at which gene connectedness resembled a scale-free network, i.e. the scale-free topology model fit R^2>0.9 (power=8). Next, a weighted, signed network adjacency matrix was calculated, raising the gene-wise correlation coefficient to the soft-thresholding power with a 10% outlier threshold (maxPOutliers = 0.1). This weight the strength of connectedness between all genes in the network as opposed to choosing a hard correlation threshold. A topological overlap matrix was constructed based on network adjacency and matrix dissimilarity calculated. This makes the network less sensitive to random noise, using connections with shared neighbours to moderate inter-gene measures of connectedness. Hierarchical clustering was performed on the dissimilarity matrix to group genes based on their connectedness and clusters of highly connected genes identified using the hybrid cutreeDynamic function (Appendix 4). Clusters with correlated patterns of expression were merged (cut-height=0.25) leaving 7 clusters of highly connected genes with distinct patterns of expression during skeletal maturation, with unclustered genes being allocated to an 8th 'Grey' group.

2.2.17 Skeletal maturation cluster characterisation

To identify genes, biological processes and biological pathways with patterns of co-expression influenced by age and sex during skeletal maturation we first identified skeletal maturation clusters significantly associated with these traits. To do this, the overall patterns of gene expression in each WGCNA module were summarised into eigengene values, defined as the first principal component of gene expression variance. The variance of each module eigengene was then tested for significant association with age, sex or the interaction between the two variables by two-way ANOVA using GraphPad Prism software (Appendix 4), identifying the Brown and Purple clusters as those most highly associated with age and the Magenta cluster as that most highly associated with sex. Pearson correlation of Brown and Purple eigengenes with age was calculated for each sex using the cor-function of the base stats package. Significantly enriched GO 'Biological Processes', KEGG pathways and DO terms in selected clusters were identified as per section 2.2.10.

Heatmaps of cluster expression and line plots of selected genes were generated using the mean zscores of normalised gene expression counts calculated across all ages in both

sexes. Heatmaps were produced using the heatmap.2 function of the gplots R-package (Appendix 4), line plots were produced in GraphPad Prism (Appendix 4) and bar plots were produced using ggplot2 R-package (Appendix 4).

2.2.18 Magenta cluster upregulation during lactation

To test the hypothesis that the Magenta cluster identified genes associated with perilacunar-remodelling, we examined their expression in a publicly available expression data from the article *Demonstration of osteocytic perilacunar/canalicular remodelling in mice during lactation* osteocytes ((Qing et al., 2012), available under the ArrayExpress accession E-GEOD-23496). Briefly, this microarray dataset was generated in an investigation of perilacunar-remodelling in lactating mice, identifying several genes involved in this process as significantly upregulated in lactation relative to both virgin mice and mice post-lactation.

To examine skeletal maturation cluster expression in this dataset background expression and control probes were filtered from the data and mean signal intensity calculated in the case of duplicate probes corresponding to a single gene. Competitive gene set testing accounting for inter-gene correlation was performed using the camera function of the limma package (Appendix 4). This compared expression of the Magenta cluster genes between virgin and lactating groups as well as lactating and post-lactation groups. Boxplots of Magenta cluster expression during lactation were generated using the mean zscores of normalised probe intensity calculated across all conditions. Boxplots were produced in using the ggplot2 R-package (Appendix 4).

2.2.19 Skeletal site comparison, differential gene activity and expression analysis

Comparison between *in-situ* isolated osteocytes from the left and right humeri was performed using samples collected before 90min in the Sample Delay Cohort described in section 2.1.2. This timepoint cut-off was chosen as no artefactual changes in gene expression due to samples processing delay were observed in these samples. Differential gene expression analysis between the left and right was performed as per section 2.2.8. The Pearson correlation between the mean voom-normalised counts in left and right samples were calculated using the cor-function of the base stats R-package and the correlation plotted using the ggplot2 R-package (Appendix 4).

Actively expressed genes in the tibiae, femora, humeri and calvariae of the Bone Comparison Cohort (section 2.1.3) were identified as per section 2.2.7. The numbers of active genes in each bone type were classed according to the gene biotype defined in the GENCODE M5 transcriptome annotation. A per sample correlation matrix was calculated based on the voom-normalise counts of genes actively expressed in any bone type, plotted using the ggplot2 R-package (Appendix 4).

Genes that were active only in select sample types were identified as per section 2.2.7. Pairwise differential gene expression analysis was performed as per section 2.2.8 comparing humeri, femora and tibiae samples. Based on power calculations modelled on sample variance, the $\log_2$-fold change (LFC) parameter of the topTreat function was set to 0.5.

Principal component analysis was performed as per section 2.2.9, clustering tibia, femur and humerus samples based on the normalised expression of all active genes and just the active homeobox genes. Homeobox genes were identified using the database described in (Wilming, Boychenko, & Harrow, 2015).

2.2.20 Identification of osteocyte-enriched genes

To identify genes enriched in osteocytes, normalised read count data was compared between the osteocyte isolated and marrow containing samples of the Osteocyte Enrichment cohort described in section 2.1.5. To do this, genes were subset to those actively expressed in either sample type as per section 2.2.7 and then expression values were normalised by library size only. Next, the limma function topTreat (Appendix 4) was used to calculate the gene-wise log2-fold change +/- 95%CI in normalised read count between osteocyte isolated and marrow containing bone samples. No other normalisation was performed as the different cellular composition of the sample types violates assumptions underpinning conventional differential gene expression techniques.

Next, the density distribution of LFC values was calculated using Scotts rule of thumb for bandwidth and plotted, revealing multiple local maxima corresponding with different levels of enrichment in osteocyte isolated samples. We hypothesised that these peaks

correspond to genes with different levels of functional specificity across the cell lineages in the marrow space, with those enriched corresponding to osteocyte specific control of the skeleton, while those showing LFC between samples indicating they were unaffected by changes in cell population performed general housekeeping functions. To group genes identified within these component populations and test this hypothesis a Gaussian Mixture Model (GMM) was fit to the distribution of LFC between conditions. The optimum number of components was first determined using the Bayesian Information Criterion (BIC) via the mclust R-package (Appendix 4). K-means clustering (k = 4) was then used to determine initiation parameters for Expectation-Maximisation fitting of the 4 component GMM using the mixtools R-package (Appendix 4). Significantly enriched GO 'Biological Processes' in each component were then identified by performing GO enrichment analysis as per section 2.2.10 on the top 1000 genes in each component ranked by posterior probability.

Finally, genes significantly enriched by osteocyte isolation were identified. To do this an osteocyte enrichment threshold was then calculated at 2SD above the mean LFC of the second most enriched GMM component (component 2), empirically Osteocyte Enrichment Threshold = 1.63LFC. Individual genes with a lower LFC-95%-CI above the Osteocyte Enrichment Threshold were considered significantly enriched in osteocytes. Scatter plots, GMM diagrams and density plots were visualised using the ggplot2 R-package (Appendix 4).

2.2.21 Gene expression during osteogenic differentiation

To identify genes upregulated in osteocytic cells as distinct from those consistently expressed in their skeletally active precursors, we made use of publicly available transcriptome sequencing data from the publication *The Osteoblast to Osteocyte Transition: Epigenetic Changes and Response to the Vitamin D_3 Hormone* ((St. John et al., 2014), available under the ArrayExpress accession E-GEOD-54783). This transcriptome sequencing data was generated as part of a temporal study of osteocyte differentiation using the osteogenic cell-line IDG-SW3 (Woo et al., 2011). Data from day 3, 14 and 35 of a cell differentiation time-course were chosen as cells at these timepoints phenotypically and molecularly resemble osteoblast, early-osteocyte and mature-osteocyte differentiation stages, respectively.

Raw data files from this experiment were aligned and quantified as per section 2.2.6. Differentially expressed genes were identified in comparisons between osteoblast-like (day 3) vs early-osteocyte-like (day 14), and osteoblast-like (day 3) vs mature-osteocyte-like (day 35) cells as per section 2.2.8. These comparisons were limited to genes with a read count ≥10 in all 3 replicates of either timepoint.

Genes expressed actively expressed in *in-situ* isolated osteocytes from either tibia, femur and humerus that had a read count <10 in all 3 replicates of all timepoints were considered absent *in-vitro*.

2.2.22 Definition of the osteocyte transcriptome signature

The osteocyte transcriptome signature genes were identified based on sequential filtering criteria using *in-vivo* and *in-vitro* (publicly available) data. These criteria can be framed in a series of questions (as per Figure *4*):

- Question 1 - Is the gene actively expressed in all osteocytes? If yes, continue to the next question, if no they were excluded from the signature. There were 13635 genes that are expressed in osteocytes from all bones. Calvarial osteocytes were excluded from this criterion as they were less pure than long bone samples. Active genes were detected as per section 2.2.7.
- Question 2 - Of these 13635, are they enriched in osteocytes (relative to other cell lineages in the marrow space)? If yes continue, if no they were excluded from the signature. These were identified as per section 2.2.20. There were 1439 genes that are expressed in tibiae, femora and humeri and also enriched in osteocytes.
- Question 3 - Of these 1439 are they expressed *in-vitro*? If yes, we asked question 4A, if no, we asked question 4B. These were determined as per section 2.2.21. There were 1267 genes expressed in all bones and enriched in osteocytes that were expressed *in-vitro*, and there were 172 genes expressed in all bones and enriched in osteocytes that were not expressed *in-vitro*.
- Question 4 A - Of these 1267 are they upregulated in osteocyte differentiation? If yes, they were included in the signature, if no, they were excluded. These were determined as per section 2.2.21. There were 780 genes that entered the signature through this route.

- Question 4 B - Of these 172 genes, are they most highly expressed in osteocytes relative to 12 other tissue types? If yes, they were included in the signature, if no, they were excluded. These were determined as per section 2.2.13. This ensured signature inclusion is not limited by *in-vitro* expression. There were 50 genes that entered the signature through this route.

These criteria defined the osteocyte transcriptome signature: 830 genes actively expressed and enriched in osteocytes throughout the skeleton that were either upregulated in osteocyte differentiation *in-vitro* or skeletally restricted *in-vivo* yet absent *in-vitro*.

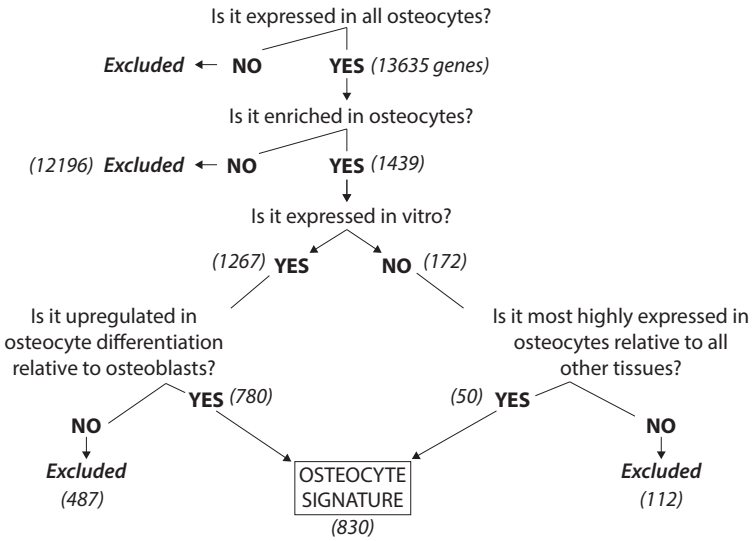


Figure 3 - Osteocyte transcriptome signature inclusion criteria.

(A) A stepwise approach to identifying genes actively expressed in multiple skeletal sites and enriched for expression in the osteocyte network *in-vivo*. Numbers reflect the number of genes remaining at each filtering stage.

2.2.23 Skeletal morphological and mechanical phenotyping in the OBCD pipeline

Knockout mouse phenotyping and analysis was performed by my colleagues Victoria Leitch, John Logan, J. H. Duncan Bassett and Graham Williams at the Imperial College of London.

Skeletal phenotyping was performed by the Origins of Bone and Cartilage Disease (OBCD) program (www.boneandcartilage.com). Samples from 16-week-old female wild-type and

knockout mice lines (*Daam2*$^{-/-}$, *Dact3*$^{-/-}$, *Ldlrad4*$^{-/-}$, n=2 to 6 per mutant genotype) were stored in 70% ethanol. The relative bone mineral content and length of the femur and caudal vertebrae are determined by digital X-ray microradiography (Faxitron MX20, 10μm pixel resolution). μCT (Scanco uCT50, 70kV, 200μA, 0.5mm aluminium filter) was used to determine trabecular bone volume (BV/TV), trabecular number (Tb.N), trabecular thickness (Tb.Th) and trabecular spacing (Tb.Sp). Bone strength and toughness were analysed by destructive 3-point bend testing of the femur (yield load, maximum load, fracture load, % energy dissipated prior to fracture) and compression testing of caudal vertebra 6 and 7 (Instron 5543 load frame, 100N and 500N load cells). Outlier phenotypes were defined by parameters > 2 standard deviations away from the reference mean determined from the 320 age, sex and genetically identical C57BL/6N wild-type controls. Enrichment of outlier skeletal parameters among mice with osteocyte transcriptome signature genes knocked out was determined relative to the frequency of outlier parameters in 626 unselected knockout lines using Fisher's Exact Test in Prism (Appendix 4).

2.2.24 Novel osteocyte signature gene structure and expression

Novel genes were assembled and identified as per section 2.2.5. Novel genes in the osteocyte transcriptome signature were those robust to the signature inclusion criteria outlined in section 2.2.22. Novel osteocyte signature gene structure and read data alignment diagrams were generated using the Gvis R-package (Appendix 4), pooling read data from each bone type. Expression in other organs and tissues was screened using the data detailed in section 2.2.13. Novel gene expression during osteogenic differentiation was examined using the data detailed in section 2.2.21. Bar and line plots were generated using the R-package ggplot2 (Appendix 4).

2.2.25 Osteocyte signature enrichment and GO semantic similarity clustering

Significantly enriched GO 'Biological Processes' and 'Cellular Components' and KEGG pathways in the osteocyte signature were identified as per section 2.2.10. Clusters of semantically similar GO terms were identified as per section 2.2.11. Cluster names were generated by calculating word frequency among GO term descriptions in each cluster using the tm R-package (Appendix 4). Common uninformative words, punctuation and white space were removed. Cluster names were constructed from among the most frequent words associated with each cluster.

Signalling pathway diagrams were constructed using the pathview R-package (Appendix 4), with gene enrichment values calculated as per section 2.2.20. Signalling ligands, modulators and receptors were identified by curating genes associated with significantly enriched KEGG signalling pathways.

2.2.26 Osteocyte signature transition module identification

To identify clusters of osteocyte signature genes with distinct profiles of expression during osteogenic differentiation, four patterns of differential gene expression were defined:

- Transition genes: Genes upregulated between osteoblast-like and early-osteocyte-like cells (FDR <0.05, LFC>0), which were then down regulated between early-osteocyte-like and mature osteocyte-like cells (FDR <0.05, LFC<0) with no significant difference in expression between osteoblast-like and mature-osteocyte-like cells.
- Early activation genes: Genes upregulated between osteoblast-like and early-osteocyte-like cells (FDR <0.05, LFC>0) with no significant difference in expression between early-osteocyte-like and mature-osteocyte-like cells.
- Maturation genes: Genes upregulated between osteoblast-like and early-osteocyte-like cells (FDR <0.05, LFC>0) that were also upregulated between early-osteocyte-like and mature-osteocyte-like cells (FDR <0.05, LFC>0).
- Late activation: Genes that show no significant difference in expression between osteoblast-like and early-osteocyte-like cells that were upregulated in mature-osteocyte-like cells compared to both earlier time points (FDR <0.05, LFC>0).

GO biological processes and KEGG pathways significantly enriched in each cluster were identified as per section 2.2.10. Heatmaps were generated using the heatmap.2-function of the gplots R-package (Appendix 4).

2.2.27 Osteocyte signature association with skeletal disease

To examine the association between osteocyte signature genes and human disease, DO terms significantly enriched in the osteocyte signature were identified as per section 2.2.10.

Osteocyte signature genes orthologs that cause rare skeletal genetic disease in humans were identified among those listed in the Nosology and Classification of Genetic Skeletal

Disorders (Bonafe et al., 2015). Significant over representation of signature orthologs among dysplasia genes was examined under the hypergeometric distribution, using the parameters:

- Population - 12366, the number of genes with human orthologs actively expressed in osteocytes from any bone type.
- Successes in population - 277, the number of skeletal disease-causing genes within the population.
- Sample size - 755, the number of osteocyte signature genes within the population.
- Successes in sample - 54, the number of osteocyte signature genes known to cause skeletal dysplasia within the population.

The enrichment of specific skeletal dysplasia types among osteocyte signature genes was tested using the enricher function of the clusterProfiler R-package (Appendix 4) (Bonferroni $P<0.05$). The gene universe for this analysis was defined as genes actively expressed in osteocytes from any bone type annotated with a human-mouse ortholog. Expression across 12 tissues and organs was analysed using the dataset detailed in section 2.2.13. The heatmap of dysplasia gene expression across tissues was generated using the R-package ggplot2 (Appendix 4).

Osteogenesis Imperfecta causal genes, in addition to those listed in the Nosology and Classification of Genetic Skeletal Diseases, were identified in the Lancet seminar *Osteogenesis Imperfecta* by Antonella Forlino and Joan C Marini with the exception of *SPARC*, identified to cause a recessive form of OI (Forlino & Marini, 2016; Mendoza-Londono et al., 2015). Analysis of OI gene expression during osteocytic differentiation using data detailed in section 2.2.21, with the heatmap generated using the gplots R-package (Appendix 4). Analysis of OI causal gene expression during skeletal maturation and cluster assignment was performed as per section 2.2.16.

2.2.28 Bone mineral density associated gene enrichment analysis

To examine the enrichment of genes associated with bone mineral density (BMD) in the osteocyte signature we analysed two independent GWAS datasets. Dataset 1: 234 genes with human-mouse orthologs identified in the UK biobank GWAS of heel BMD (Kemp et al., 2017). Dataset 2: All genes associated with genome wide significant variants related to

variation in BMD, BMD (spine), BMD (hip), BMD (femoral neck), BMD (wrist) listed in the curated NHGRI-EBI catalogue of published genome-wide association studies (Appendix 4).

The over representation of genes associated with BMD in the osteocyte signature was tested under the hypergeometric distribution, using the parameters:

- Population - the number of genes with human orthologs actively expressed in osteocytes from any bone type.
- Successes in population - the number of BMD associated genes within the population.
- Sample size - the number of osteocyte signature genes within the population.
- Successes in sample - osteocyte signature genes associated with human BMD variance within the population.

Over representation of the osteocyte signature among validated BMD-associated genes was also examined. Validated BMD-associated genes were those annotated with skeletally relevant GO biological processes or skeletal mammalian phenotype (MP) terms when knocked out in mice. Over representation was examined under the hypergeometric distribution for both GO and MP, using the parameters:

- Population - the total number of BMD-associated genes with human orthologs actively expressed in osteocytes.
- Successes in population - the number of genes annotated with a skeletal GO or MP term within the population.
- Sample size - the number of osteocyte signature genes within the population.
- Successes in sample - the number of osteocyte signature genes annotated with a skeletal GO or MP term within the population.

Identification of BMD-associated signature genes most highly expression in osteocytes relative to 12 tissues and organs was performed as per section 2.2.13.

2.3 Quantification and statistical analysis

To avoid systematic batch effects, samples for each transcriptome sequencing cohort were collected in single batches where possible. When not possible sample types were evenly distributed between batches. Statistical methodologies and software used for performing analysis in this work are cited in context in the Methods text.

Analysis using Trinity, Stringtie, Cuffcompare, Trimgalore, STAR and RSEM were performed on a computing cluster running the CentOS 6.8 (Rocks 6.2) Linux operating system. CTAn, NRecon and Drishti were run using a Windows 7 OS. CPAT and ReViGo analyses were run using the web portals (Appendix 4). Two-way ANOVA was performed using GraphPad Prism (Appendix 4). All other statistical analysis was performed in R (Appendix 4).

Error bars reflect mean and standard deviation unless stated otherwise. We employed multiple hypothesis correction wherever significance was evaluated across multiple statistical tests (i.e. differential gene expression analysis, Gene Ontology enrichment, KEGG pathway enrichment, Disease Ontology enrichment).

2.4 Data and software availability

The raw sequencing (fastq), transcriptome assembly data files (GTF files used for alignment and quantification including novel transcripts), read alignment files (BAM), processed gene expression data files (FPKM and read-counts) are deposited at ArrayExpress, accession numbers: E-MTAB-5532 (Bone Comparison Cohort) and E-MTAB-5533 (Osteocyte Enrichment Cohort). The publicly available gene array data from osteocytes in virgin, lactating and post lactation mice was originally released with the publication *Demonstration of osteocytic perilacunar/canalicular remodelling in mice during lactation* and was accessed via ArrayExpress (accession E-GEOD-23496). The publicly available osteoblast to osteocyte differentiation transcriptome data was originally released with the publication *The Osteoblast to Osteocyte Transition: Epigenetic Changes and Response to the Vitamin D_3 Hormone* and was accessed via ArrayExpress (accession E-GEOD-54783). The publicly available organ expression transcriptome data was originally released with the publication *A circadian gene expression atlas in mammals: Implications for biology and medicine* and

was accessed via ArrayExpress (accession E-GEOD-54652). Unless otherwise stated, software used for this analysis is publicly available as listed in the key resources table or in R via the Comprehensive R Archive Network (CRAN) or Bioconductor.

3 SEQUENCING THE OSTEOCYTE NETWORK

Osteocytes are a cell type defined by their environment and yet attempts to study their transcriptome are often preceded by techniques to remove them from their extracellular home. This is because their location within one of the hardest tissues of the body and the morphology of their interconnected dendritic network makes it very challenging to isolate a pure population of primary cells by dissection. Thus, techniques of collagenase digestion, *ex-vivo* cultures, and *in-vitro* models have been used to purify osteocytes for transcriptome analysis (Dallas et al., 2013). However, perhaps unsurprisingly, removing these bone cells from bone tissue using these techniques has been shown to artefactually alter gene expression, an issue pertinent to the sensitivity of transcriptome sequencing (Ayturk et al., 2013).

Whole bone transcriptome analyses have also been performed, in which the osteocyte network is left intact along with the marrow residing within the bone compartment (Rawlinson et al., 2009). This approach, while likely to avoid the introduction of artefactual alterations from purification, faces a different problem - marrow contamination. The distributed organisation of the osteocyte network means the cell density of bone tissue is considerably less than the marrow compartment. In effect, this means most of the cells in bone samples with the marrow intact are not bone cells, and thus the vast majority of RNA and data obtained from these samples pertains to cell types not under investigation. Additionally, the presence of mixed cell populations makes it hard to localise expression measurements as originating from a specific cell type. These challenges illuminate a compromise required for the molecular study of the osteocytes - isolate a pure population of cells from bone tissue and with it change the transcriptome, or, leave the osteocyte network intact within the bone tissue to avoid artefactual transcriptome changes but court the challenges associated with mixed cell populations.

Here we document an approach to perform transcriptome sequencing on the intact osteocyte network, which minimises the influence of other cell types and technical gene expression artefacts. With this technique we examine the effect of sample processing delay on the osteocyte transcriptome and establish a timeframe for collection which minimises artefactual changes in gene expression.

3.1 Methods

3.1.1 Sample details and protocol optimisation

Ethical approval for all animal experiments is detailed in section 2.1.1. The first step in technical development was the optimisation of RNA extraction from osteocyte enriched bone samples detailed in section 2.2.2. The results of these experiments guided the development of the *in-situ* osteocyte isolation and RNA extraction protocols described in sections 2.2.1 and 2.2.4. µCT and histology validation of *in-situ* osteocyte isolation was performed as per section 2.2.3. These samples were taken from the Bone comparison cohort detailed in section 2.1.3.

3.1.2 Transcriptome data generation and analysis

To investigate the effect of sample processing delay and response to amputation on gene expression in the osteocyte network, *in-situ* isolated osteocyte samples were collected as per section 2.2.1 from the Sample Delay Cohort detailed in section 2.1.2. Transcriptome sequencing of these samples was performed as per section 2.2.4, and data processed and active genes determined as per sections 2.2.6 and 2.2.7. Differential gene expression between each timepoint and the control group was performed as per section 2.2.8. GO biological processes significantly enriched among differentially expressed genes were identified as per sections 2.2.10 and 2.2.14.

3.2 Results

3.2.1 In-situ isolation of the osteocyte network

To isolate osteocytes from contaminating cell population without disaggregating the osteocyte network we used a range of physical separation techniques (*Figure 4* A). To isolate cell from the long bones, mixed cell population in the growth plates were removed by dissection and cells from the periosteal bone surface were removed by gentle scraping with a scalpel. Next, the bulk of cells in the marrow space were syringed out with PBS leaving the trabecular bone inside the cortical tube intact. Lastly, cells remaining in the marrow space were removed by centrifugation, this process yielding bone samples with this process yielding bone samples containing both cortical and cancellous bone that were highly enriched for osteocytes intact within their local *in-situ* environment (*Figure 4* B, *Figure*

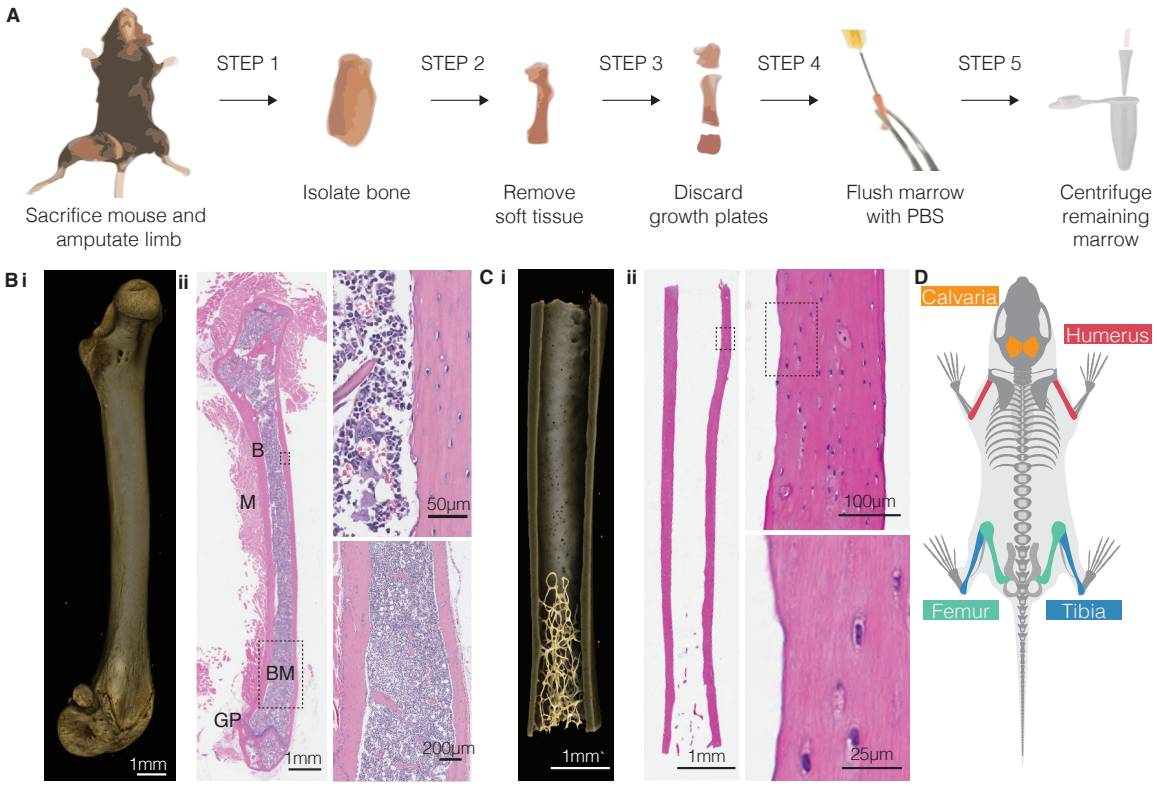


Figure 4 - *In-situ* isolation of the osteocyte network.

(A) Strategy to isolate osteocytes from long bones by removing extra-osseous tissues. (B) Micro-CT (i) and histological (ii) section of unprocessed bone (i-ii). (C) Micro-CT (i) and histological (ii) section of *in-situ* isolated bone sample (i-ii). (D) Schematic of the four bone types from which osteocytes were isolated for transcriptome sequencing.

---

4 C). This technique was optimised for the tibia, femur and humerus, enabling the collection of a consistent volume of bone tissue for all three bone types within 16 minutes of sacrifice (*Figure 4* D, *Figure 5* A and B). To compare the transcriptome of long and flat bones, *in-situ* isolated samples from the calvaria were also collected. Due differences in the plate-like morphology of the of the calvaria and the tube-like long bones, the isolation approach was modified (detailed in method section 2.2.1).

Next, techniques used to extract RNA directly from the mineral bound osteocytes were optimised (*Table 1*). First, we examined the effect on RNA integrity of different RNA extraction approaches, extraction reagents, sample homogenisation techniques, extraction from fresh vs frozen samples, and sample thawing with and without dedicated thawing reagents. We observed that RNA integrity was higher following phase-separation

techniques independent of reagent used compared to column-based RNA extraction methods, particularly if the on-column DNase treatment was followed. Handheld rotor-stator systems for sample homogenisation also slightly outperformed bead-based systems. Snap-freezing samples in liquid nitrogen showed no significant effect on RNA integrity, so long as samples were not allowed to thaw prior to homogenisation, which resulted in complete loss of RNA integrity, irrespective to the addition of dedicated thawing reagents. With this in mind, our approach integrated the snap-freezing of samples after the removal of unwanted extraosseous tissues. This effectively decoupled collections and RNA extraction, increasing the number and efficiency of samples able to be collected in a single batch. A summary of these experiments is contained in *Table 1*.

Table 1 – Optimisation of RNA extraction from osteocyte isolated bone tissue

Test	Control	RNA Integrity	Outcome
Trizol manual extraction	TRIreagent manual extraction	No difference	Continued with TRIreagent manual extraction
TRIreagent column extraction	TRIreagent manual extraction	Slight decrease on column & drastic decrease with on-column DNase	Continued with TRIreagent manual extraction
Bead-based homogenisation	Rotor-stator-based homogenisation	Slight decrease with bead-based system	Continued with rotor-stator homogenisation
Snap frozen processed samples	Fresh processed samples	No difference	Continued with snap frozen samples
Snap frozen samples thawed in PBS	Snap frozen samples extracted from frozen	Complete degradation with thawing	Continued with extraction from frozen samples
Snap frozen samples thawed in RNA-ice	Snap frozen samples extracted from frozen	Complete degradation with thawing	Continued with extraction from frozen samples

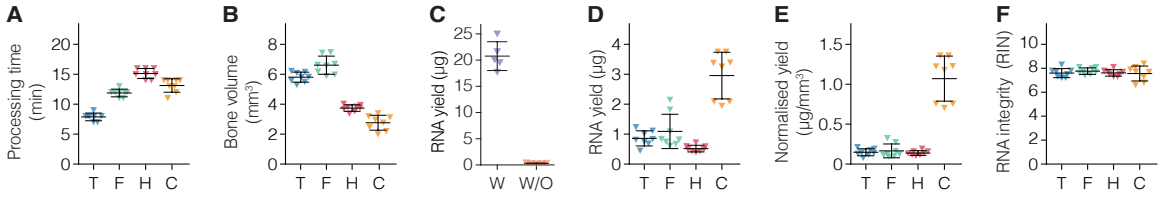


Figure 5 - Consistent extraction of high-quality RNA from *in-situ* isolated osteocyte samples.

(A) Time taken for *in-situ* isolation from animal sacrifice to sample snap freezing. (B) Volume of bone in *in-situ* isolated osteocyte samples. (C) Yield of RNA from bone samples with marrow intact (W) and after *in-situ* isolation without marrow (W/O). (D) Yield of RNA from *in-situ* isolated osteocyte samples. (E) RNA yield normalised by mean bone volume of *in-situ* isolated osteocyte samples. (F) RNA integrity of *in-situ* isolated osteocyte samples (RIN = RNA Integrity Number). T = tibia, F = femur, H = humerus and C = calvaria *in-situ* isolated osteocyte samples. Error bars reflect mean and 95% CI.

In-situ isolation reduced the RNA yielded from long-bone samples by more than 40-fold compared to marrow containing samples, consistent with the fact that most of the RNA in whole bone samples is derived from marrow cells (*Figure 5* C). The yield of RNA was highly consistent within each long-bone type, proportional to the volume of bone tissue in each sample (*Figure 5* D and E). The yield of RNA from calvaria samples was significantly higher than that of the long bones, despite the sample bone volume being slightly lower. This indicated that the purity of calvaria-samples was lower than that of the long bones, with remaining marrow likely accounting for the increased RNA yield. As such, conservative approaches were used in the analysis of samples from the calvaria. For all bone types, RNA integrity measurements were high (RNA Intergity Number > 7) indicating the extraction of high-quality RNA from the *in-situ* isolated osteocyte network (*Figure 5* F). This demonstrated that *in-situ* osteocyte isolation minimised the RNA contribution from undesired cell types in multiple bone types and yielded high-quality RNA suitable for transcriptome sequencing.

3.2.2 The osteocyte network response to amputation and hypoxic stress

Having established an optimised approach to collecting and extracting high integrity RNA from the osteocyte network, we next examined how the transcriptome may be altered by delay from animal sacrifice to sample processing. We hypothesised that limb removal, and thus cessation of blood supply, would induce a hypoxic-stress response in the

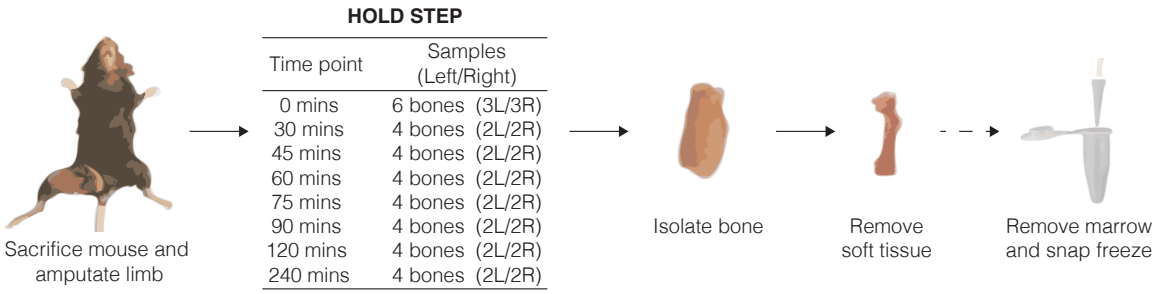


Figure 6 - The response of sample processing delay and amputation in the osteocyte network.

Experimental design used to investigate the effect of sample processing delay and amputation on the osteocyte network.

---

osteocytes, evident in the transcriptome. To examine this hypothesis, we generated and compared transcriptome sequencing data from *in-situ*-isolated osteocyte samples processed immediately following animal sacrifice (within 10mins of sacrifice), to samples processed after being held at room temperature for up to 4 hours after sacrifice and limb removal (*Figure 6*).

Differential gene expression analysis revealed no significantly differentially expressed genes between osteocytes in the control samples and those processed up to 90 minutes post sacrifice (*Figure 7* A). At 2 hours, 164 genes significantly increased expression and 50 were suppressed, a number that grew to 743 up and 819 down regulated by 4 hours post sacrifice. Importantly, we observed that these transcriptional differences could not be explained by differences in RNA integrity, with no significant difference recorded between samples at any time point (*Figure 7* B). The vast majority of the 214 genes differentially expressed at 2 hours were also differentially expressed at the latter time point (190/214, 89%), which indicated the persistence of the gene expression response initiated at 2 hours (*Figure 7* C). Together, these observations indicated that the differences in gene transcript abundance were not due to RNA degradation but rather controlled changes in gene transcription, and that these changes did not occur until more than 1.5 hours after sacrifice.

To identify the biological processes involved in the osteocyte response to limb removal, Gene Ontology enrichment analysis was performed on the genes up and down regulated at each time point. Upregulated at 2 and 4 hours were biological processes

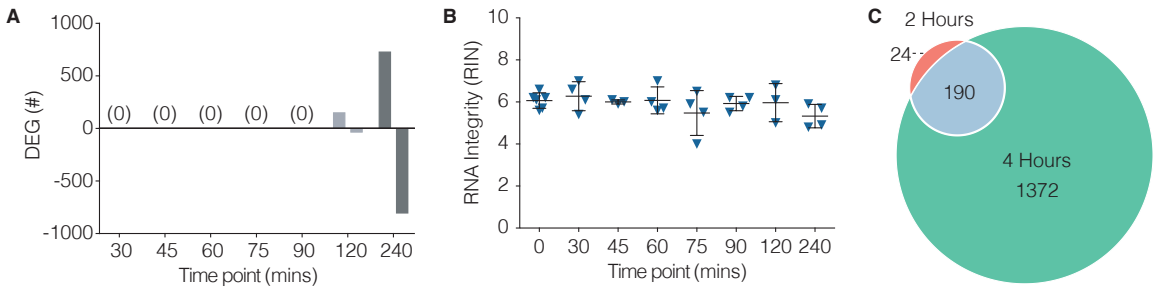


Figure 7 - Differential gene expression induced 90 minutes after amputation.

(A) Number of differentially expressed genes (DEG) at timepoints after amputation relative to control samples (immediately processed following sacrifice). (B) RNA integrity of *in-situ* isolated osteocyte samples at timepoints after amputation (RIN = RNA Integrity Number). Error bars reflect mean and 95% CI. (C) Maintenance of differential gene expression beyond 2 hours after amputation.

pertaining to the recruitment of blood supply, such as blood vessel development (GO:0048514, 2hours = 29 associated genes up-regulated, 1.62E-10; 4hours = 87 associated genes up-regulated, 1.43E-22) and angiogenesis (GO:0001570, 2hours = 25 associated genes up-regulated, 3.31E-09; 4hours = 71 associated genes up-regulated, 8.65E-18) (*Table 2* and *Table 3*). This was supported by terms associated with cells and tissues used to make vasculature such as endothelium development (GO:0003158, 2hours = 13 associated genes up-regulated, 2.10E-07; 4hours = 29 associated genes up-regulated, 3.56E-11) and epithelial cell differentiation (GO:0045446, 2hours = 11 associated genes up-regulated, 1.09E-05; 4hours = 25 associated genes up-regulated, 5.94E-09).

A number of genes in signalling pathways known to be important to the skeleton that play an important role in the response to hypoxia were also upregulated (*Table 3*). This included the Notch pathway (GO:0007219, 4hours = 23 genes, 1.77E-03) and Bone Morphogenic Protein (BMP) signalling cascades (GO:0030509, 4hours = 22 genes, 4.92E-04), both important hypoxia response pathways in connective tissues. The 4-hour time point also saw the significant enrichment of multiple terms associated with synaptic organisation (GO:0050808, 4hours = 29 genes, 2.33E-03), synapse signalling (GO:0099536, 4hours = 41 genes, 4.67E-03) and axon guidance (GO:0007411, 4hours = 21 genes, 6.35E-03) in the osteocyte network (*Figure 8* A and B). This revealed that many genes upregulated in osteocytes after limb amputation are involved in the recruitment of blood supply and that

Table 2 – Top 10 significantly enriched GO terms in genes upregulated in osteocytes 2 hours after amputation

GO term description	DEG (#)	P-value
GO:0048514 blood vessel morphogenesis	29	1.62E-10
GO:0001525 angiogenesis	25	3.31E-09
GO:0003158 endothelium development	13	2.10E-07
GO:0045446 endothelial cell differentiation	11	1.09E-05
GO:0001885 endothelial cell development	8	3.61E-04
GO:0001570 vasculogenesis	9	5.37E-04
GO:0035295 tube development	21	1.05E-03
GO:0001936 regulation of endothelial cell proliferation	9	1.21E-03
GO:0003013 circulatory system process	16	1.35E-03
GO:0048729 tissue morphogenesis	20	1.75E-03

P-value is Bonferroni adjusted.
DEG = number of differentially expressed genes associated with term

Table 3 – Top 10 significantly enriched GO terms in genes upregulated in osteocytes 4 hours after amputation

GO term description	DEG (#)	P-value
GO:0048514 blood vessel morphogenesis	87	1.43E-22
GO:0001525 angiogenesis	71	8.65E-18
GO:0003158 endothelium development	29	3.56E-11
GO:1901342 regulation of vasculature development	42	1.47E-09
GO:0035295 tube development	69	1.76E-09
GO:0045446 endothelial cell differentiation	25	5.94E-09
GO:0001667 ameboidal-type cell migration	47	1.81E-08
GO:0090287 cellular response to growth factor stimulus	39	2.36E-08
GO:0045765 regulation of angiogenesis	37	6.19E-08
GO:0071363 cellular response to growth factor stimulus	60	8.51E-08

P-value is Bonferroni adjusted.
DEG = number of differentially expressed genes associated with term

the Notch, BMP and axonal-synaptic signalling pathways may be important in the osteocyte networks' response to prolonged hypoxic stress.

Over represented among the 50 genes down-regulated at 2 hours post sacrifice were 8 biological processes, associated with leukocyte recruitment and cell division (*Table 4*). By 4 hours, a total of 231 biological processes were identified, the top 7 of which related to DNA conformation and nucleosome organization (*Table 5*). These terms revealed the

A

GO Term description	P-value
GO:0050808 synapse organization	2.33E-03
GO:0007268 chemical synaptic transmission	3.98E-03
GO:0098916 anterograde trans-synaptic*	3.98E-03
GO:0099537 trans-synaptic signaling	4.31E-03
GO:0099536 synaptic signaling	4.67E-03
GO:0007411 axon guidance	6.35E-03
GO:0050804 synaptic transmission modulation*	2.34E-02
GO:0007409 axonogenesis	3.85E-02
GO:0061564 axon development	4.48E-02

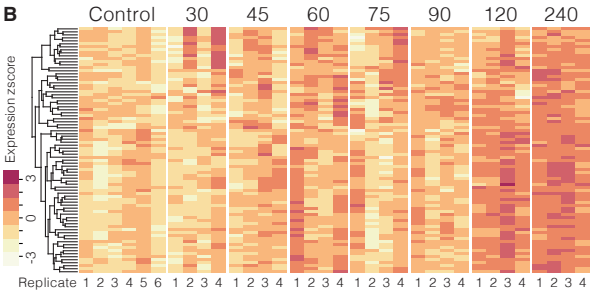


Figure 8 - Axonal-synaptic signaling processes upregulated following amputation.

(A) Significantly enriched biological processes associated with axon guidance and synaptic signaling following amputation (FDR ≤ 0.05). * truncated description. (B) Temporal expression of differentially expressed genes associated with axonal-synaptic signaling following amputation.

consistent down-regulation of genes encoding histone and chromatin components, suggesting the osteocytes response to hypoxia may also involve chromatin level rearrangement in addition to the transcriptional changes observed in our data.

Significantly enriched terms associated with immune and hematopoietic cell recruitment and activation, including leukocyte migration (GO:0050900, 4hours = 56 genes, 1.60E-17), regulation of immune response (GO:0050776, 4hours = 87 genes, 1.72E-17), and lymphocyte activation (GO:0046649, 4hours = 89 genes, 3.49E-17) were also down regulated (*Table 5*). Osteocytes are known to contact and influence different cell types in the marrow space, however the nature, extent and effect of most of these interactions are not yet understood. This data indicated that these processes may be important in the osteocyte response to tissue damage and stress.

Finally, no terms associated with apoptosis or cell death were detected in the analysis of genes up or down regulated at any time point. Further, no significant changes in *RANKL* expression nor other genes associated with osteocyte apoptotic signalling were observed, including *Gja1*/*Cx43*, *Hmgb1*, *Bcl2*, *Panx1*, *P2rx7*, *Casp3* and *P2ry2*. This indicates that upto 4-hours post amputation, without blood circulation, the transcriptome response of the osteocyte network may be one of prolonged cell survival, not programmed cell death.

Table 4 – All significantly enriched GO terms in genes down-regulated in osteocytes 2 hours after amputation

GO term description	DEG (#)	P-value
GO:0045087 innate immune response	10	2.74E-03
GO:0019730 antimicrobial humoral response	4	4.16E-03
GO:0006959 humoral immune response	5	7.11E-03
GO:0002523 leukocyte migration involved in inflammatory response	3	1.05E-02
GO:0000280 nuclear division	8	1.08E-02
GO:0048285 organelle fission	8	2.92E-02
GO:0007059 chromosome segregation	7	3.46E-02
GO:0019731 antibacterial humoral response	3	4.53E-02

P-value is Bonferroni adjusted.
DEG = number of differentially expressed genes associated with term

Table 5 – Top 10 significantly enriched GO terms in genes down-regulated in osteocytes 4 hours after amputation

GO term description	DEG (#)	P-value
GO:0071103 DNA conformation change	59	1.05E-25
GO:0006323 DNA packaging	51	2.19E-23
GO:0006334 nucleosome assembly	41	6.20E-22
GO:0031497 chromatin assembly	42	1.09E-19
GO:0034728 nucleosome organization	41	8.06E-18
GO:0006335 DNA replication-dependent nucleosome assembly	22	9.41E-18
GO:0034723 DNA replication-dependent nucleosome organization	22	9.41E-18
GO:0050900 leukocyte migration	56	1.60E-17
GO:0050776 regulation of immune response	87	1.72E-17
GO:0065004 protein-DNA complex assembly	47	2.08E-17

P-value is Bonferroni adjusted.
DEG = number of differentially expressed genes associated with term

3.3 Discussion

Taken together, these results document a time efficient approach to extract RNA and perform transcriptome sequencing on osteocytes from multiple bone types throughout the skeleton. Previous studies that use similar techniques to isolate osteocytes within bone tissue have demonstrated their advantages relative to collagenase-based approaches (Ayturk et al., 2013; Kelly et al., 2016). However, it was not established how gene expression in the osteocyte network may be artefactually altered by sample collection and separation from the circulatory system.

This data indicates that the osteocyte transcriptome is stable for up to 90-minutes following amputation, providing a window for sample processing before significant changes in gene expression are observed. These experiments informed the optimisation of the protocols detailed in the method sections 2.2.1 and 2.2.4 which were used in subsequent chapters.

Beyond 2 hours, artefactual changes in the osteocyte transcriptome become apparent that are undetectable based on RNA integrity. Interestingly, this timeframe corresponds with reported changes in osteocyte morphology in response to restricted blood supply, observed by electron microscopy (James & Steijn-Myagkaya, 1986). The study by James and Steijn-Myagkaya documented changes in cellular structure associated with cell stress and apoptotic pathway activation such as disorganisation of the cytoskeleton and chromatin condensation. These changes commenced 2-hours after separation from the circulatory system and were well established by 4-hours. While the regulation of genes associated with histone conformation in our data may support the reported chromatin reorganization, there is little indication of a change in cell apoptosis even up to 4-hours post sacrifice. Instead, the transcriptome changes indicate an attempt to re-establish the blood supply, upregulating genes associated with angiogenesis as well as signalling pathways including Notch and axon guidance (Hiyama et al., 2011; Ramasamy, Kusumbe, Wang, & Adams, 2014). A potentially important difference between our study and that of James and Steijn-Myagkaya was the way in which samples were handled post dissection. In contrast to our approach, in which bones were left intact within their surrounding extra-skeletal tissue until processing, James and Steijn-Myagkaya removed all extra-skeletal tissue prior to incubation. Our data suggest that leaving extra-skeletal tissue intact may prolong osteocyte survival following amputation.

Multiple lines of evidence indicate that blood-vessel growth and osteogenesis are somewhat coupled in the skeletal system. During development the forming skeleton regulates limb vascular patterning, while in healing fractures, coupled angiogenesis and osteogenesis is associated with the localisation of osteogenic cells (Eshkar-Oren et al., 2009; Maes et al., 2010). Paracrine control of blood-vessel development by neighbouring tissues is established to occur between the nervous system and in tumour growth, however

whether a similar process is present in skeletal tissue is not known (Weinstein, 2005). Interestingly, axon guidance plays an important role in this cross-talk between the nerves and blood-vessels and is dynamically regulated in our data. While speculative, the patterns of expression in our data may suggest that in response to restricted blood supply the osteocyte network engages transcriptional processes and pathways involved in blood-vessel recruitment. Alternatively, the gene expression profiles observed in our data may also reflect changes in the transcriptome of intracortical blood vessels which are unlikely to be removed by the osteocyte isolation technique used to collect bone samples (Klingberg et al., 2017; Núñez et al., 2018). While these cells are likely to represent only a small fraction of cells in bone tissue relative to osteocytes, their presence may influence gene expression values obtained using our methodology (Schaffler & Kennedy, 2012). Thus, future experiments utilising techniques such immunohistochemistry or fluorescence *in situ* hybridisation, which are capable of localising gene expression signals to specific cell types, would be highly informative to discern the relative contribution of blood vessels and osteocytes to the gene expression patterns observed in our data.

Collectively, this data establishes the stability of the osteocyte transcriptome for up to 90 minutes post amputation, suggesting gene expression in samples collected in this period is minimally affected by technical factors. Beyond this timeframe the osteocyte network responds to amputation by upregulating genes involved in vascular development, with little change in pathways associated with apoptosis. This suggests that the transcriptome response of the osteocyte network favours cell survival up to 4-hours after amputation and may be involved in re-establishing blood supply.

4 THE ACTIVE OSTEOCYTE TRANSCRIPTOME THROUGHOUT THE SKELETON

The osteocyte network maintains the structural integrity of more than 200 bones while balancing endocrine mineral demand. Despite this critical role, we know very little of the genes actively expressed by osteocytes and how this may change throughout the skeletal system. During embryogenesis, bone tissue in different skeletal sites acquire a morphology suited to their array of functions. For example, the flat-bones of the skull such as the parietal calvaria form plate-like structure which protect the brain through the process of intramembranous ossification (Leucht et al., 2008; Rux & Wellik, 2017). In contrast, the long-bones such as the tibia, femur and humerus of the limbs are shaped like rods to provide leverage and support and are formed by endochondral ossification. These processes occur during development and are guided by restricted patterns of transcription factor activity. These patterns define limb identity, ensure the reflected symmetry of limbs from the left and right, and establish proximal-distal bone structures (Hamada, Meno, Watanabe, & Saijoh, 2002; Rux & Wellik, 2017; Wellik, 2007; Zakany & Duboule, 2007). While function of site-specific gene expression during skeletogenesis is well established, far less is known about the role of site-specific gene expression in the adult skeleton.

Site-specific gene expression in osteoprogenitors and cells lining bone has been show to play a role in the fidelity of bone healing within the adult skeleton (Leucht et al., 2008; Rux et al., 2016). Perturbation of these expression patterns alter cell differentiation potential, delaying and reducing the quality of fracture repair. Fractures heal by a process of bone formation and reshaping, known as bone-modelling and bone-remodelling respectively (Kumar & Narayan, 2014). Osteocytes coordinate bone-remodelling, a critical process not just in fracture healing, but also for maintaining the healthy homeostasis of skeletal structure (Dallas et al., 2013). This may suggest osteocytes also play a role in site-specific remodelling processes, yet, the identity or even existence of genes site-specifically expressed in the osteocyte network is poorly defined.

In this chapter, we use transcriptome sequencing to identify all the genes actively expressed the osteocyte network throughout the skeleton, including protein-coding, long

non-coding and novel unannotated genes. We analyse the osteocyte transcriptome across multiple skeletal sites, comparing long-bones from opposing sides of the body, within the same limb, between the fore and hindlimb; and the flat-bones from the skull.

4.1 Methods

4.1.1 Sample details and transcriptome data generation

Ethical approval for all animal experiments is detailed in section 2.1.1 and the Bone Comparison Cohort is described in section 2.1.3. *In-situ* isolated osteocyte samples were collected from the tibia, femur, humerus and calvaria of each mouse in this cohort as per section 2.2.1. Transcriptome data from each bone was generated as per section 2.2.4.

4.1.2 Novel gene assembly and the identification of active genes

De-novo assembly was performed to identify novel transcripts as per section 2.2.5. The expression of known and novel genes was then quantified as per section 2.2.6 and genes actively expressed in the osteocyte network throughout the skeleton determined as per section 2.2.7. Gene activity in the osteocyte network was compared with other organs and tissues as per section 2.2.13.

4.1.3 Skeletal-site-specific gene expression

Gene expression in the osteocyte network was also compared between bone types. First, the correlation of active gene expression in different bone-types was determined as per section 2.2.19. Next, genes that were differentially expressed between skeletal sites were identified as per section 2.2.8. Lastly, bone distinguishing patterns of gene expression were identified by principal component analysis as per section 2.2.9.

4.2 Results

4.2.1 The osteocyte transcriptome is very similar throughout the skeleton

To identify all the genes actively expressed in the osteocyte network throughout the skeleton, we sequenced and assembled the transcriptome of *in-situ*-isolated osteocytes from four bone types, the tibia, femur, humerus and calvarium (*Figure 9* A). Total-RNA transcriptome libraries were used as they enable the measurement of protein-coding and non-coding transcripts. *De-novo* transcriptome assembly was performed on the data from

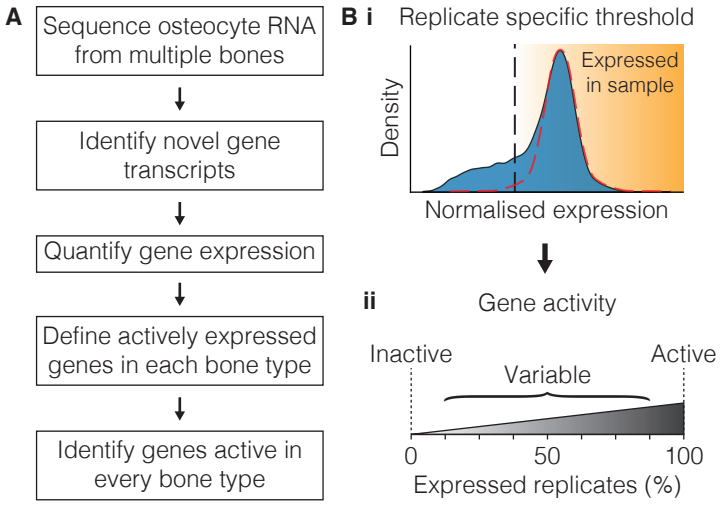


Figure 9 - Identification of genes actively expressed in the osteocyte network.

(A) Strategy used to define genes actively expressed in the osteocyte network. (B) Actively expressed genes defined as those greater than sample specific expression thresholds (i) in all replicates of each bone type (ii).

each bone type to enable the detection of unannotated, novel transcripts potentially restricted to osteocytes. Assembled transcripts were pooled and refined through a bespoke filtering pipeline to remove single-exon transcripts, transcripts with abnormally large or small exons and transcripts with splice-patterns that matched annotated GENCODE or RefSeq annotated gene. The retained multi-exon, unannotated transcripts were then concatenated to the GENCODE transcriptome annotation and gene expression quantified for both the known and novel genes.

Next, we defined all the genes actively expressed in the osteocyte network from each of the four bone types. This approach established sample-specific gene expression thresholds and identified actively expressed genes above this threshold in all 8 replicates of a given bone-type (*Figure 9* B) (described in detail in the method section 2.2.7). An important consideration in this analysis was the use of distinct sample collection methodology for the calvaria-samples. Histological analysis indicated calvaria-samples retained more marrow than those from the long-bones (not shown), also reflected in a significantly higher bone-volume-normalised RNA yield (*Figure 5* E). This reduction in sample purity meant gene expression estimates were distorted due to the contribution of

A

Gene type	Tibiae (#)	Femora (#)	Humeri (#)	All (%)
Protein	12675	12657	12847	95
lncRNA	849	858	866	69
TEC	486	485	496	84
Novel	72	75	75	48
Other	56	55	51	54
Total	14138	14130	14335	92

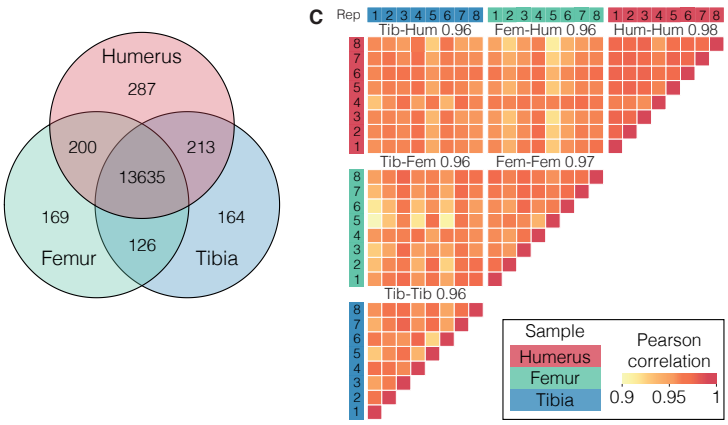


Figure 10 - The osteocyte transcriptome is very similar between bone types.

(A) Number of genes actively expressed in each bone type. The percentages reflect the number of genes actively expressed in all long-bone types. TEC = To be Experimentally Confirmed. (B) Overlap of numbers of genes actively expressed in osteocytes in individual long-bone types. (C) Pearson correlation between the active transcriptome of osteocyte samples from the tibiae, femora and humeri. Numbers represent mean correlation between bone type. Tib = tibia, Fem = femur, Hum = humerus, Rep = biological replicate.

RNA from extra-osseous cell types. To account for this, a conservative approach was taken for the analysis of calvaria samples relative to the long bones from the limbs. The results of this analysis are reported in a separate section 4.2.3 below.

Gene expression in the osteocyte network was very similar between skeletal sites yet distinct from other organs and tissues. Across the tibia, femur and humerus, 14,794 genes were actively expressed in osteocytes, 92% of which were actively expressed in all bone types (*Figure 10* A and B). The gene-type composition of the osteocyte transcriptome in each bone type was very similar, encompassing protein-coding genes (~90%), long non-coding RNAs (lncRNAs, ~6%), a number of genes to be experimentally confirmed (TEC, ~4%) and novel genes for which nothing is known (>50 genes per bone, ~0.4%, explored further in section 6.2.3). The levels of active gene expression were highly correlated between skeletal sites, (R = 0.91-0.98), with the mean-correlation between samples from different bone types as high as between osteocytes from the same bone type (*Figure 10* C). This suggested the pattern of expression across the active osteocyte transcriptome was highly consistent between skeletal sites.

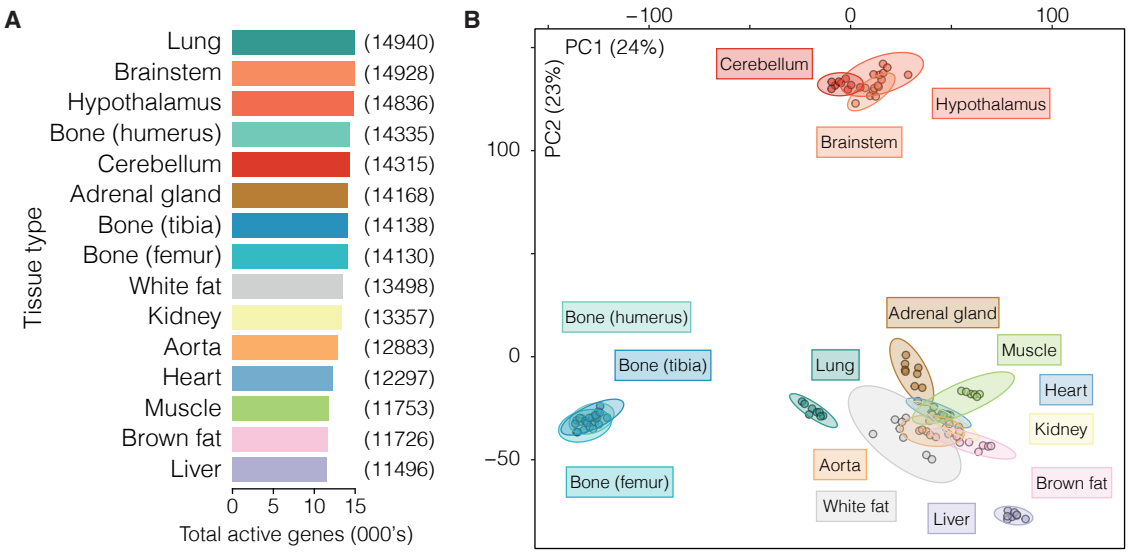


Figure 11 - The active osteocyte transcriptome is distinct from other organs and tissues.

(A) Number of genes actively expressed osteocytes and other organs and tissue types. (B) Sample clustering based on the first two principal components (PC1 and PC2) of active gene expression in osteocytes and 12 other tissue types. Ellipses represent 95% confidence intervals. Percentages reflect variance explained.

The active osteocyte transcriptome in the tibia, femur and humerus was next compared with the actively expressed genes of 12 other organs and tissue types. The number of genes actively expressed in osteocytes was very similar to a number of other tissues, including 8,997 genes common to all tissues sampled (Figure *11* A). Despite the similarity in gene number and overlap in gene activity, the first two principal components of gene expression variation segregated osteocytes from all other tissue types (Figure *11* B). This indicated the pattern of active gene expression in osteocytes is distinct from these other organs.

4.2.2 Differential gene activity in the osteocyte network between skeletal sites

While the overall transcriptome profile of osteocytes was very similar between bone types, we next sought to specifically identify genes with differential expression throughout the skeleton. First, we examined gene expression across the sagittal plane by comparing gene expression in osteocytes from the left humeri, with those collected from the right. No differentially expressed genes were detected despite 10 biological replicates in each group. Moreover, the coefficient of Pearson correlation for gene expression between replicates from left and right humeri was $R > 0.99$ (Figure *12* A). Together, these data indicated that

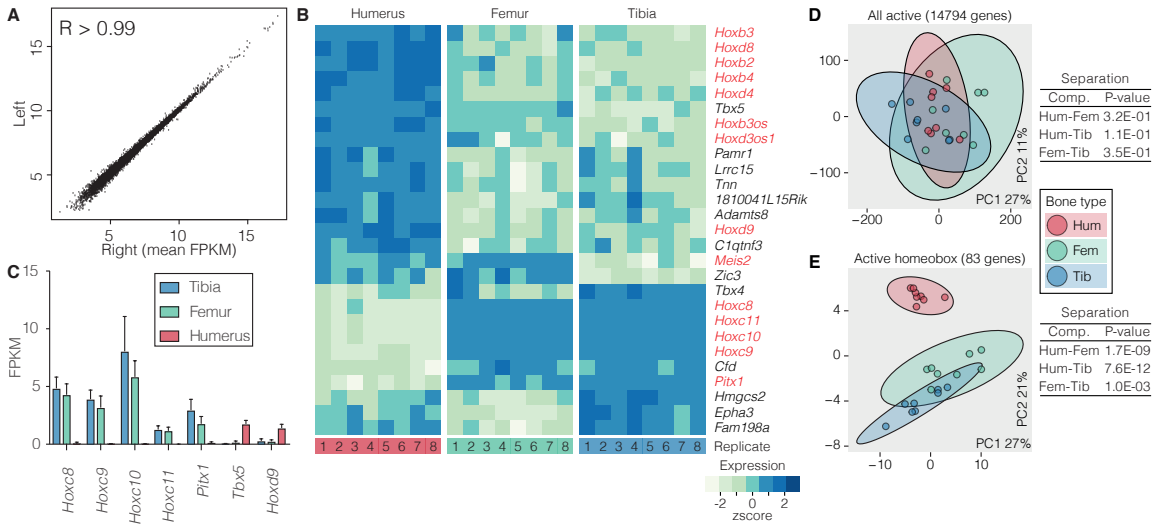


Figure 12 - Skeletal-site-specific gene expression in the osteocyte network.

(A) Pearson correlation between the active osteocyte transcriptome from the humerus of the left and right forelimbs. (B) Significantly differentially expressed genes between osteocytes isolated from different skeletal-sites. Homeobox genes are in red (FDR ≤ 0.05, LFC > 0.5). (C) Normalised expression of differentially active genes ('ON' or 'OFF') between the fore-limb and hindlimb. Error bars reflect mean and SD. (D) Clustering of bone types based on the first two principal components (PC1 and PC2) of the active osteocyte transcriptome and (E) only active homeobox gene expression. Percentages reflect variance explained by individual PCs. Ellipses represent 95% confidence intervals. P-values calculated using on Hotellings T-test. FPKM = Fragments per kilobase per million mapped reads (normalised expression).

gene expression differences that control the formation of mirrored symmetry in the skeleton during development are not evident in the transcriptome of adult osteocytes.

Next, we compared the active osteocyte transcriptome in the tibia, femur and humerus. Twenty-seven genes were differentially expressed between the three bone types (log_2-fold difference >0.5, p<0.05) (Figure *12* B). These included 2 genes that were only active in the humeri and 5 genes restricted to the hindlimbs (Figure *12* C, Appendix 1). The genes expressed only in the humeri were transcription factors T-box 5 (*Tbx5*) and Homeobox-d9 (*Hoxd9*), both known to be pivotal in establishing fore-limb identity during embryogenesis (Agarwal et al., 2003; Takeuchi et al., 1999). Similarly, Homeobox-c8-c11 (*Hoxc8*, *Hoxc9*, *Hoxc10* and *Hoxc11*) and Paired-like homeodomain 1 (*Pitx1*) were expressed only in the hindlimb bones, consistent with their embryonic expression profile. Pitx1 induces expression of *Hoxc10* in a process critical to the specification of hindlimb

identity (DeLaurier, Schweitzer, & Logan, 2006). *Pitx1* is also a causal gene in Liebenberg syndrome, which is characterised by leg-like limb malformation of the arm when *Pitx1* is aberrantly expressed in the forelimb (Spielmann et al., 2012).

Three genes in the homeobox-b (*Hoxb2*, *Hoxb3* and *Hoxb4*), two genes in the homeobox-d (*Hoxd4* and *Hoxd8*) co-linear-gene clusters, as well as two homeobox-associated antisense lncRNAs (*Hoxb3os* and *Hoxd3os1*) were also differentially expressed, all significantly higher in the humeri relative to the hindlimb bones (Figure *12* B). In bone lining cells and osteoprogenitors, site specific *Hoxa11* expression has been shown to be important to the fidelity and rate of fracture healing. Interestingly, no genes from the HoxA cluster were differentially expressed and *Hoxa11* expression was below active levels in all bone types. This supported the efficacy of the *in-situ* osteocyte isolation procedure at removing cells from the bone surface, and indicated that the Hox-genes site-specifically expressed in the osteocyte network are distinct from those reported in other skeletal cell types (Rux & Wellik, 2017). Only two genes were differentially expressed in osteocytes from the tibia compared to the femur, namely the Meis homeobox transcription factor (*Meis2*) and the ZIC family transcription factor *Zic3* both higher in the femur. *Meis2* was also significantly upregulated in osteocytes from the humerus compared to the tibia but not compared to the femur, consistent with their proximal distal expression patterning in the developing limb (Capdevila, Tsukui, Esteban, Zappavigna, & Belmonte, 1999). Collectively, the results indicated that distinct patterns of molecular expression occur in osteocytes from different skeletal sites, distinct from that observed in other skeletal cell types.

Since >50% of the significantly differentially expressed genes were related to the homeobox family of transcription factors, we next examined whether the homeobox-family gene expression alone could distinguish bone types. Supporting previous analysis of osteocyte similarity throughout the skeleton, PCA analysis of all genes actively expressed in osteocytes (n=14,794 genes) was unable to delineate bone type (Figure *12* D). In contrast, the first two principal components of homeobox gene expression (n=83 genes) separated samples by bone type, with significant separation between humeri, femora and tibiae clusters (Figure *12* E). This indicated that site-specific differences in homeobox gene expression distinguish the osteocyte network in different skeletal sites.

4.2.3 Distinct patterns of gene activity in the calvaria

Next, to understand how the osteocyte transcriptome changes between bones with different developmental origins, we compared the active transcriptome of the long bones to that of the calvaria. It is important to note that conventional comparison techniques, such as differential gene expression, are based on assumptions that the tissue composition of samples is highly similar. The inclusion of marrow in the calvaria-samples, not present in samples from the long bones, violate the assumptions underlying these techniques. Thus, long and flat bones were compared based on gene activity (active/inactive) rather than the level of gene expression (high/low).

While the number and identity of active genes in the active transcriptome of the calvarium was very similar to the long bones, a limited number of genes were differentially active between bone types. Ninety-four percent of the 14026 genes in the active transcriptome of the calvarium were also actively expressed in all long bones (Figure *13* A). Comparison between long and flat bones identified 45 genes actively expressed in the tibia, femur and humerus that were inactive in the calvarium (Appendix 1). These included skeletally significant genes such as *Wnt1*, an important stimulator of bone formation associated with severe skeletal disease (Joeng et al., 2017). Twenty percent (9/45) of these genes absent from the calvarium that were expressed in the long bones belonged to the homeobox transcription factor family. Only 1 of the 23 Hox-genes actively expressed in osteocytes from the long-bones were actively expressed in the calvarium (Figure *13* B). Two genes were actively expressed in the calvarium that were inactive in all the other bone types, *Zic2* and *Zic4*, both ZIC-family developmental transcription factors. Together, these results indicated that gene expression in the osteocyte network of the calvarium was somewhat distinct to that of the long bones, with differences centred on the site-specific activity of developmental transcription factors.

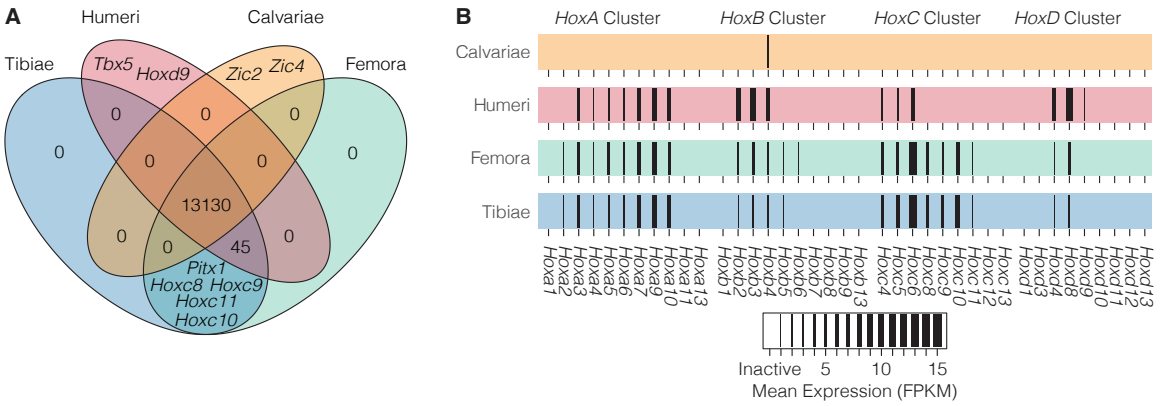


Figure 13 - Distinct patterns of gene activity in osteocytes from the calvaria.

(A) Differentially active genes ('ON' or 'OFF') between bone types, removing genes with variable expression in any bone type. (B) Mean normalised homeobox gene expression in the osteocyte network from different skeletal sites.

4.3 Discussion

These results define the transcriptome of the osteocyte network across a number of skeletal sites. This comprehensive profile of the gene expression in the healthy adult skeleton builds on previous RNA-sequencing studies of the osteocyte network that are limited to comparisons within one bone type (Ayturk et al., 2013; Kelly et al., 2016). These studies also used approaches to data generation that are biased towards protein-coding mRNA and as a consequence very little is known about the lncRNA landscape in the skeleton. Here we identify the active expression of >14,000 genes in the osteocyte network encompassing protein-coding genes, long non-coding RNAs and novel genes not previously described in major transcriptome annotations (discussed in detail in section 8.4).

Our data shows that active gene expression in the osteocyte network is highly correlated between bone types, with the exception of a handful of developmental transcription factors with site specific patterns of activity. While no difference was detected between bone from the left and right limbs, this analysis demonstrated the pattern of homeobox gene expression alone was capable of distinguishing osteocytes from different bone types. These differences Homeobox genes are a class of transcription factors known to play a pivotal role in skeletal patterning *in-utero*, however, their expression in the osteocyte network the adult skeleton is poorly defined (Wellik, 2007; Zakany & Duboule, 2007). A previous report comparing gene expression in whole-bone samples from the adult

hindlimb and the skull documented differences in developmental markers including homeobox genes (Rawlinson et al., 2009). However, this investigation was limited in its ability to localise expression to bone embedded cells as the majority of RNA and thus gene expression data from whole-bone samples is derived from cells in the marrow space. Histological examination and genetic manipulation of osteogenic precursors and cells lining bone have demonstrated that regionally restricted homeobox expression in the adult skeleton are important for the rate and fidelity of fracture healing in a site-specific manner (Leucht et al., 2008; Rux et al., 2016). Interestingly, the genes identified in these studies were not detected in our data. This supports the efficacy with which cells lining bone are removed from *in-situ* isolated osteocyte samples and indicates that site specific homeobox-gene expression within the osteocyte network is distinct from these other skeletal cell types. Consequently, the site-specific patterns of gene expression the osteocyte network may perform a distinct function to that of other osteogenic cell types, potentially related to its highly specialised function within the skeletal system (discussed in detail in section 8.1).

Our analysis also indicated that homeobox gene expression was very limited in calvaria samples from the skull relative to the long bones. Moreover, our data indicated osteocytes from the calvaria express a distinct pattern of developmental transcription factors from the ZIC family, which play an important role in brain development (Elms, Siggers, Napper, Greenfield, & Arkell, 2003; Grinberg et al., 2004). The flat bones of the skull differ from the long bones of the limbs in a number of ways, including the cellular progenitor populations from which they are derived during development and their response to skeletal loading in adulthood (Xiaobing Jiang, Iseki, Maxson, Sucov, & Morriss-Kay, 2002; Rawlinson, Mosley, Suswillo, Pitsillides, & Lanyon, 1995). While none of these genes have an established role in mechanical loading, it is possible that these gene expression patterns reflect differences in load experienced by these bone types. Alternatively, the cell progenitors which go on to form the facial bones of the skull lack homeobox gene expression, similar to what we observe in our parietal samples (Xiaobing Jiang et al., 2002; Leucht et al., 2008; Rice, Rice, & Thesleff, 2003). Leutch *et al* showed that this absence of a ‘Hox-code’ meant bone tissue from the mandible readily integrated with surrounding bone tissue when introduced into defects in the limbs (Leucht et al., 2008). Conversely, bone taken from the limb integrated very poorly into the mandible, affecting the rate and

fidelity of fracture healing. While both comprising the skull, the facial bones such as the mandible and the parietal flat bones which were collected in our samples are derived from distinct progenitor cell types, the former from cranial neural crest cells while the latter is from the paraxial mesoderm (Xiaobing Jiang et al., 2002; Wu, Chen, Tian, & Liu, 2017). Our findings suggest that limited Hox gene expression may be a common feature of these different skull bone types, and distinguish bones from skull from those of the limbs. Future studies may investigate whether this has consequences for fracture healing, analogous to effects observed in the mandible.

5 DYNAMIC GENE EXPRESSION IN THE OSTEOCYTE NETWORK DURING SKELETAL MATURATION

After birth the skeleton continues to mature, with bones fusing, stiffening, lengthening and widening as they take their adult form (Land & Schoenau, 2008; Opperman, 2000). This skeletal maturation proceeds into early adulthood until a plateau in bone accrual is reached (Baxter-Jones et al., 2011; Gordon et al., 2017). At this point the rates of bone accrual and bone loss equilibrate and bone mass remains relatively stable throughout reproductive life (R Rizzoli & Bonjour, 1997). After roughly three decades of skeletal homeostasis this balance shifts in favour of bone resorption and the rate of age-related bone loss increases, particularly in females following menopause (René Rizzoli & Bonjour, 2010; Weaver et al., 2016). While the rate of bone loss is influenced by a range of environmental and hormonal factors, peak bone mass remains a powerful predisposing factor to bone diseases such as osteoporosis in advancing age (Gordon et al., 2017; Hui, Slemenda, & Johnston, 1990). Because of this, the molecular processes that control skeletal maturation and the attainment of peak bone mass in youth may be highly relevant to the process of healthy skeletal aging.

During post-natal maturation the skeleton undergoes periods of rapid growth in terms of bone length, shape and mineral composition. Depending on the skeletal site, up to 46% of the adult BMC is accrued during the adolescent growth period (Baxter-Jones et al., 2011). While diet, exercise and a host of environmental lifestyle factors influence the rate of bone accrual, these explain only 20–40% of the variance in peak bone mass (Weaver et al., 2016). Twin and population studies have shown peak bone mass is a highly heritable trait, indicating genetic factors have a major influence on bone accrual during post-natal growth (Pocock et al., 1987). Despite this, the genes, biological processes and biochemical pathways involved in skeletal maturation are largely unknown.

While the period sexual and skeletal maturation is a critical time for bone mineral acquisition in both males and females, sex -based differences in peak bone mass are well established (Alswat, 2017). While the sex dimorphism in skeletal maturation is commonly attributed the differential activity of androgen and other hormones upregulated in puberty,

other genetic factors independent of sex steroids have also been shown to influence the skeleton (R Rizzoli & Bonjour, 1997; Sigal, 1984; Wei & Mao, 2007). These observations indicate that the molecular control processes governing skeletal maturation may differ between the sexes, and that important genetic factors dictating this difference in skeletal development and disease in may yet to be discovered.

Sex differences also place specific demands on the skeleton accentuated in child-bearing and the calcium demands of lactation. The osteocyte network has been shown play a critical role in meeting these endocrine requirements, upregulating genes capable of resorbing mineral surrounding its cellular network, a process known as perilacunar-remodelling (Jähn et al., 2017; Qing et al., 2012). While this sex -specific setting serves as the architype example for this process, little is known about how perilacunar-remodelling contributes to normal skeletal regulation by the osteocyte network.

The osteocyte network plays a critical role in the regulation of bone mass and yet its specific contribution to postnatal skeletal maturation is poorly defined. Moreover, little is known about how gene expression in the osteocyte network might differ with sex, despite well-established differences in skeletal regulation between the sexes. Here we hypothesise that the morphological differences occurring in the skeleton during maturation are accompanied by changes in transcriptome of the osteocyte network. We define clusters of co-expressed genes dynamically regulated in the osteocyte network during skeletal maturation and identify the pathways and processes with expression influenced by age and sex.

5.1 Methods

5.1.1 Sample details

Ethical approval for animal experiments reported here is detailed in section 2.1.1 and the Skeletal Maturation Cohort is described in method section 2.1.4. Left and right humeri were collected from each animal in this cohort, with the bones from the right side collected intact (removing muscle and ligaments) and used for morphological analysis by DXA and μCT as per section 2.2.3. Bones from the left side were taken for transcriptomic

investigation, with samples processed to obtain *in-situ* isolated osteocytes as per method section 2.2.1.

5.1.2 Transcriptome data generation

Transcriptome sequencing was performed on the *in-situ* isolated samples as per section 2.2.4, with data processed and genes actively expressed in each sample type identified as per sections 2.2.6 and 2.2.7.

5.1.3 Identification and characterisation of co-regulated gene clusters

Samples were clustered based on active gene expression as per section 2.2.15. Clusters of co-expressed genes with differential patterns of regulation during skeletal maturation were identified as per section 2.2.16. The identification of co-expressed clusters with expression patterns significantly influenced by age and/or sex was performed as per section 2.2.17. GO biological processes, KEGG biochemical pathways and Disease Ontology terms significantly overrepresented among genes in each cluster were identified as per section 2.2.10. Methodology used in the examination of Magenta cluster gene expression during lactation is detailed in section 2.2.18.

5.2 Results

5.2.1 Morphological differences during skeletal maturation are accompanied by transcriptome changes in the osteocyte network.

To understand the effects of skeletal maturation on bone composition, we first measured the morphological changes occurring over this time in both sexes (Figure *14* A). In mice, sexual maturity occurs between 6-8 weeks of age and peak bone mass is reached between 16-24 weeks of age (Brodt et al., 1999; Dutta & Sengupta, 2016; Richman et al., 2001). It is important to note that while post-pubertal bone growth is reported in many mouse strains, previous studies have shown no change in the bone length of C57BL/6 mice between 6 and 12 months of age in either sex, indicating that skeletal maturity is achieve by 6 months of age in the mouse strain used for this investigation (Somerville, Aspden, Armour, Armour, & Reid, 2004). Thus, to capture the dynamic changes occurring during skeletal maturation and acquisition of peak bone mass, we assessed the morphology of humeri collected from male and female mice at 4, 10, 16 and 26 weeks of age, measuring

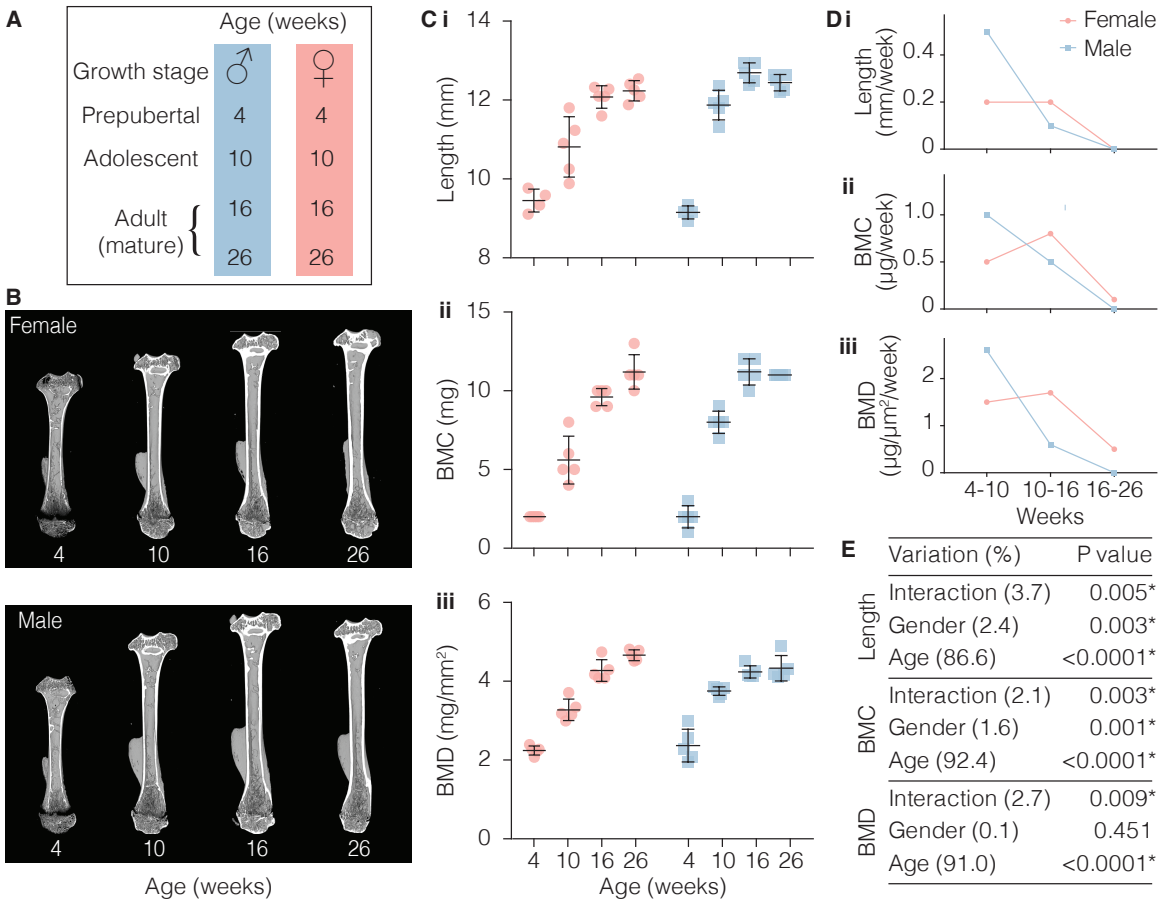


Figure 14 - Bone morphology varies with age and sex during skeletal maturation.

(A) Cohort design used to examine transcriptome changes in the osteocyte network occuring with skeletal maturation. (B) Micro-CT images of humeri collected from female and male mice at different stages of skeletal maturation. (C) Humeri length (i), bone mineral content (BMC) (ii) and bone mineral density (iii) collected from female and male mice at different stages of skeletal maturation. Contralateral humeri were used in the transcriptome investigation of the osteocyte network. (D) Mean humeri growth between time points, measured in terms of length (i), BMC (ii) and BMD (iii). (E) Two-way analysis of the variance (ANOVA) of morphological measurements. Percentages represent the amount of variance in length (i), BMC (ii) and BMD (iii) explained by age, sex and the interaction between age and sex. P-values correspond to the significance of association and are Bonferroni corrected. * significant p-value < 0.05

---

bone-length, bone mineral content (BMC) and bone mineral density (BMD), both cortical and trabecular, by µCT and DXA (Figure *14* B and C). Consistent with this understanding of murine skeletal maturation, the rate of growth in bone length and mineral content was markedly reduced beyond 16 weeks in both sexes, with no difference observed in median bone length in either sex between 16 and 26 weeks (Figure *14* D). Two-way ANOVA showed that while both age and sex have a significant effect on bone morphology individuallyC

these effects are not entirely independent of each other, with the interaction between age and sex significantly associated with morphological variance in all three parameters (Figure *14* E). This indicated that the morphological changes that occur during skeletal maturation are influenced by both age and sex, and that the interaction between the two variables plays a significant role in determining skeletal structure. It is important to note many of the previous reports documenting the effects of age and sex on the skeleton do not specify the sub-strain of C57BL/6 mice used in the study (Brodt et al., 1999; Somerville et al., 2004). Genetic and morphological differences have been reported between the commonly used C57BL/6J strain and the C57BL/6N strain used in our investigations (Fontaine & Davis, 2016; Zurita et al., 2011). However little is known if these differences extend to the skeleton. This may be an important consideration in the extrapolation of these findings to other mouse strains.

Given the master role of the osteocyte network in skeletal regulation, we hypothesised that the changes in skeletal morphology observed in skeletal maturation would be accompanied by changes in the transcriptome of these critical skeletal cells. To test this, transcriptome sequencing was performed on *in-situ* isolated osteocyte samples taken contralateral humeri of animals used for morphological analysis (both sexes at 4, 10, 16 and 26 weeks of age) (Figure *14* A). It is important to note growth plates are excluded from osteocyte-isolated samples mitigating the influence of these regions on gene expression measurements.

First, genes actively expressed in each sample type were determined, revealing subtle differences in the pattern of gene activity in the osteocyte network during maturation (Figure *15* A). In both sexes, the number of genes actively expressed increased with age from ~12500 to ~14000, with more than 80% of genes (12,530) consistently active throughout maturation in both sexes. While similar numbers of lncRNAs and novel genes were active between sexes at each age, only a fraction were active in all sample types (36% and 18% respectively), suggesting their activity is dynamically regulated during skeletal maturation. Principal component analysis of active gene expression showed that both age and sex have a significant influence on gene expression in the osteocyte network. When clustered within sex the first two principal components separated samples by age, with samples arranged from youngest to oldest and the 4-week-old timepoint significantly

separated from the other ages (Figure *15* B). This indicated that the osteocyte transcriptome changes with age in both sexes, transitioning through distinct patterns of expression on the way to skeletal maturity. Comparing sexes across timepoints revealed that gene expression changes associated with sex occur late in skeletal maturation, with significant separation between male and female samples at 16 weeks and beyond (Figure *15* C). Together these analyses demonstrated that the morphological changes occurring during maturation are accompanied by changes in the osteocyte transcriptome, with age playing a significant role early in skeletal maturation and sex-related differences becoming apparent with skeletal maturity.

A

Gene type	Female 4 weeks (#)	Female 10 weeks (#)	Female 16 weeks (#)	Female 26 weeks (#)	Male 4 weeks (#)	Male 10 weeks (#)	Male 16 weeks (#)	Male 26 weeks (#)	Any (#)	All (%)
Protein	12370	12716	12874	12910	12353	12759	12912	12937	13592	86
lncRNA	649	723	724	809	650	714	798	773	1150	36
TEC	414	436	446	484	406	463	475	458	571	55
Novel	57	56	49	66	52	52	71	57	125	18
Other	37	45	44	44	41	39	50	38	60	47
Total	13527	13976	14137	14313	13502	14027	14306	14263	15498	81

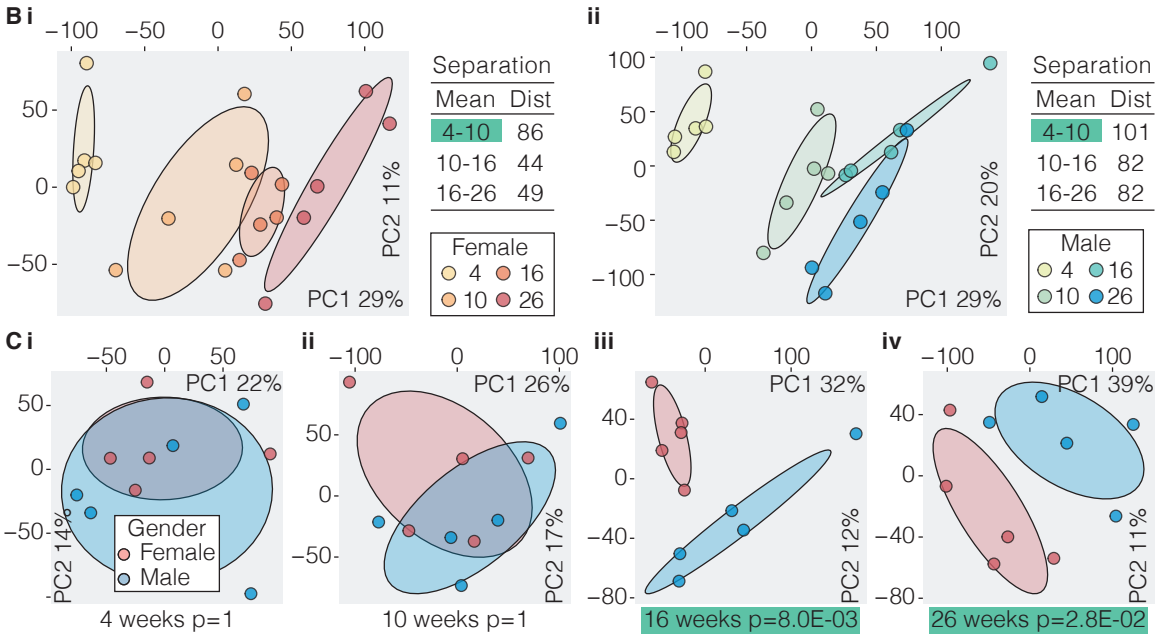


Figure 15 - The transcriptome of the osteocyte network changes during skeletal maturation.

(A) Number of genes actively expressed in the osteocyte network at different stages of skeletal maturation in each sex. 'Any' corresponds to the number of genes actively expressed in any one age or sex. The percentages reflect the number of genes actively expressed at all stages in both sexes. TEC = To be Experimentally Confirmed. (B) Clustering of different stages of skeletal maturation based on the first two principal components (PC1 and PC2) of the active osteocyte transcriptome in all female (i) and all male (ii) samples. Percentages reflect variance explained by individual PCs. Dist = Euclidean distance between cluster centeroids (mean PC1 and PC2). Clusters with significant separation between centroids are highlighted in green (FDR ≤ 0.05). (C) Clustering of male and female samples based on the first two principal components (PC1 and PC2) of the active osteocyte transcriptome at each stage of skeletal maturation. Percentages reflect variance explained by individual PCs. Clusters with significant separation between centroids are highlighted in green (FDR ≤ 0.05). Ellipses represent 50% confidence intervals. P-values calculated using on Hotellings T-test.

5.2.2 Clusters of co-expressed genes in the osteocyte network are significantly associated with age and sex.

To identify genes whose expression in the osteocyte network is regulated during skeletal maturation, we constructed a weighted co-expression network based on the gene expression in all 40 sequenced samples (*Figure 16* A). Seven clusters of genes with highly-correlated patterns of expression between samples were identified, with genes not displaying an obvious pattern of expression between samples placed in an eighth 'Grey' group (3,680 genes) (*Figure 16* B). The overall pattern of cluster expression across samples was summarised into representative eigengene values, which explain the majority of gene expression variance occurring in each group (explained in detail in method section 2.2.17). The variance of these eigengene values was then analysed across age and sex for each cluster and used to identify coregulated genes significantly associated with these factors during skeletal maturation. Age was seen to significantly influence gene expression in 6 of the 8 clusters, explaining more than ~87% of eigengene expression variance in the purple cluster independent of sex, and ~83% of variance in the brown cluster with a significant yet modest sex contribution of ~5% (*Figure 16* B). Eigengene values indicated expression of these two clusters is divergent in the osteocyte network during skeletal maturation, with purple positively and brown negatively correlated with age in both male and female mice (Purple: Female R=0.86, p=1.09E-06, Male R=0.90, p=9.05E-08; Brown: Female R=-0.86 p=1.18E-06, Male R=-0.90, p=5.75E-08). This analysis also highlighted the magenta cluster, with age and sex each explaining ~35% of eigengene expression variance, and an additional ~7% explained by the interaction of these two variables (*Figure 16* C). Importantly, the Grey cluster of uncorrelated genes showed no significant association with age or sex, consistent with the notion that the function of these genes is not coupled to skeletal maturation. Using this approach, we defined clusters with expression patterns significantly associated with the age and sex, identifying groups of genes co-regulated during skeletal maturation.

5.2.3 Divergent regulation of protein processing and endocrine signalling in the osteocyte network during skeletal maturation.

The two clusters with the strongest association with age were the Brown and Purple clusters, diverging in their expression pattern during skeletal maturation. The 2510 gene Brown cluster was most-highly expressed at 4-weeks and gradually down regulated during

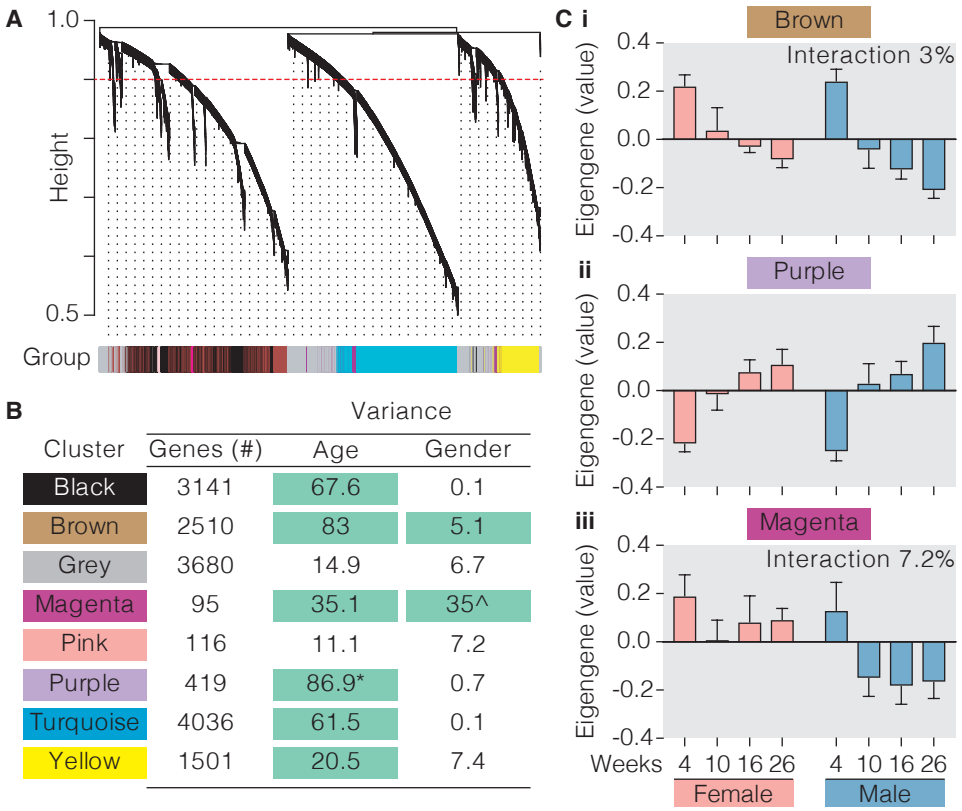


Figure 16 - Clusters of genes co-expressed in the osteocyte network associated with age and sex.

(A) Dendrogram of gene co-expression identifying clusters of genes with highly similar patterns of activity during skeletal maturation. Height refers to the threshold of correlation used to group genes. (B) Seven clusters of genes with distinct patterns of expression during skeletal maturation and one cluster of uncorrelated genes ('grey'). Variance in cluster eigengene expression explained by age and sex. * Cluster most highly associated with age, ^ most highly associated with sex. Significant associations are highlighted in green (FDR ≤ 0.05). (C) The pattern of eigengene expression in the Brown, Purple and Magenta clusters with skeletal maturation across sexes. Interaction = the percentage of variance explained by the interaction of age and sex, significant in the Brown and Magenta clusters (FDR ≤ 0.05).

---

maturation (*Figure 17* A). Gene ontology analysis revealed the significant enrichment of a number of biological processes associated with protein-glycosylation and Golgi-transport, supported by the overrepresentation of genes associated with protein processing, glycan synthesis and export in the KEGG biochemical database (*Figure 17* B and C). Protein glycosylation is critical to normal bone development and homeostasis, with proteoglycans shown to have specific roles in bone cell proliferation, extracellular matrix deposition, mineralisation and remodelling (Nikitovic et al., 2012; Y. Sun et al., 2015). The brown cluster

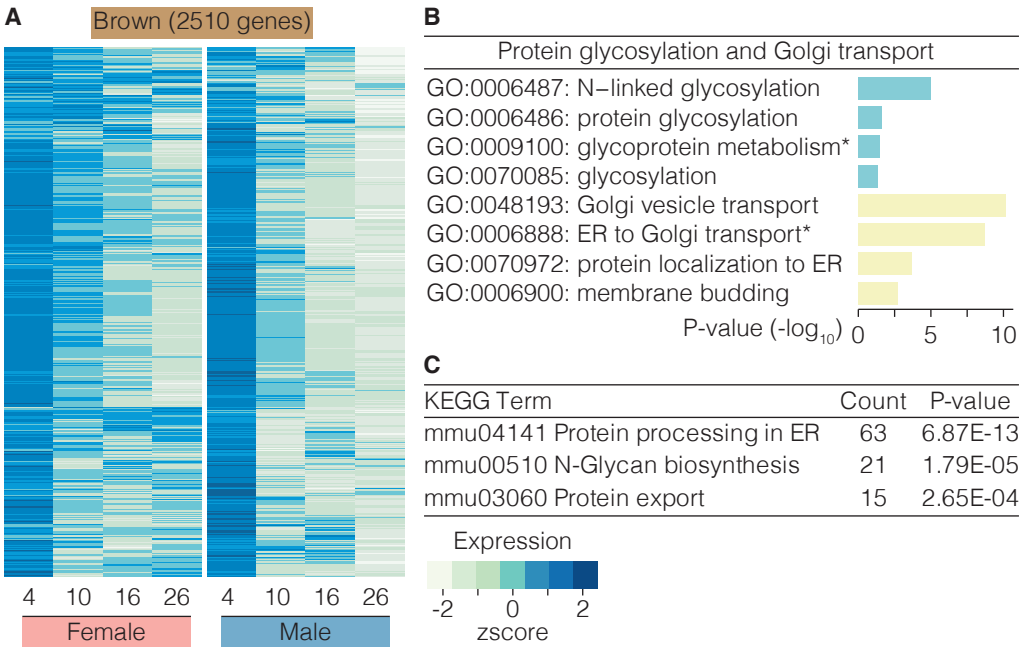


Figure 17 - Down-regulation of protein processing and export in the osteocyte network during skeletal maturation.

(A) Brown cluster expression during skeletal maturation in both sexes. Values reflect mean of z-scores of normalised counts at each time point. (B) Semantically similar clusters of significantly enriched GO biological processes associated with genes in the Brown cluster. * truncated description (FDR ≤ 0.05). (C) KEGG pathways significantly overrepresented in the Brown cluster (FDR ≤ 0.05).

identified a number of proteoglycans such as Biglycan (*Bgn*), Versican (*Vcan*), Brevican (*Bcan*) along with chondroitin sulfate synthase (*Chsy1*) which makes a major proteoglycan constituent of cartilage (Wilson et al., 2012). Additionally, 5 glucosyltransferases were identified in the cluster, including *Galnt1*, critical for the production of at least two major glycoprotein constituents of bone matrix bone sialoprotein and osteopontin (Miwa, Gerken, Jamison, & Tabak, 2010). The pattern of gene expression in the Brown cluster indicate genes associated with protein glycosylation and proteoglycan production are down regulated in the osteocyte network during skeletal maturation and that this process is somewhat influenced by sex.

The Purple cluster identified 419 co-expressed genes upregulated during skeletal maturation (*Figure 18* A). Enrichment analysis of known biological processes and biochemical pathways failed to return any significantly over represented terms, indicating the combined biological function of these genes is poorly defined. One Disease Ontology

term was significantly enriched among the Purple cluster genes, identifying 19 genes associated with endocrine system disease (DOID:28, p=3.20E-03) (*Figure 18* B). Among those were a range of mineralocorticoid and glucocorticoid hormone receptors, including the nuclear receptors *Nr3c1* and *Nr3c2*, as well as the androgen receptor gene *Ar*. In addition to these endocrine receptors, this cluster also highlighted the upregulation of genes associated with skeletally important biological processes during skeletal maturation. These included a number of osteocyte marker genes such as *Sost*, *Mepe* and *Dkk1* , as well as downstream effectors of endocrine signalling processes such as *Bcl2*, *Fas* and *Mecp2*, which have paracrine and autocrine functions (Hockenbery, Zutter, Hickey, Nahm, & Korsmeyer, 1991; Nan et al., 1996; Suda, Takahashi, Golstein, & Nagata, 1993) (*Figure 18* C). Together this indicates the sex independent upregulation of a number of genes associated with endocrine function during skeletal maturation, including sex hormone receptors and downstream signalling molecules, alongside genes known to be important to osteocyte biology.

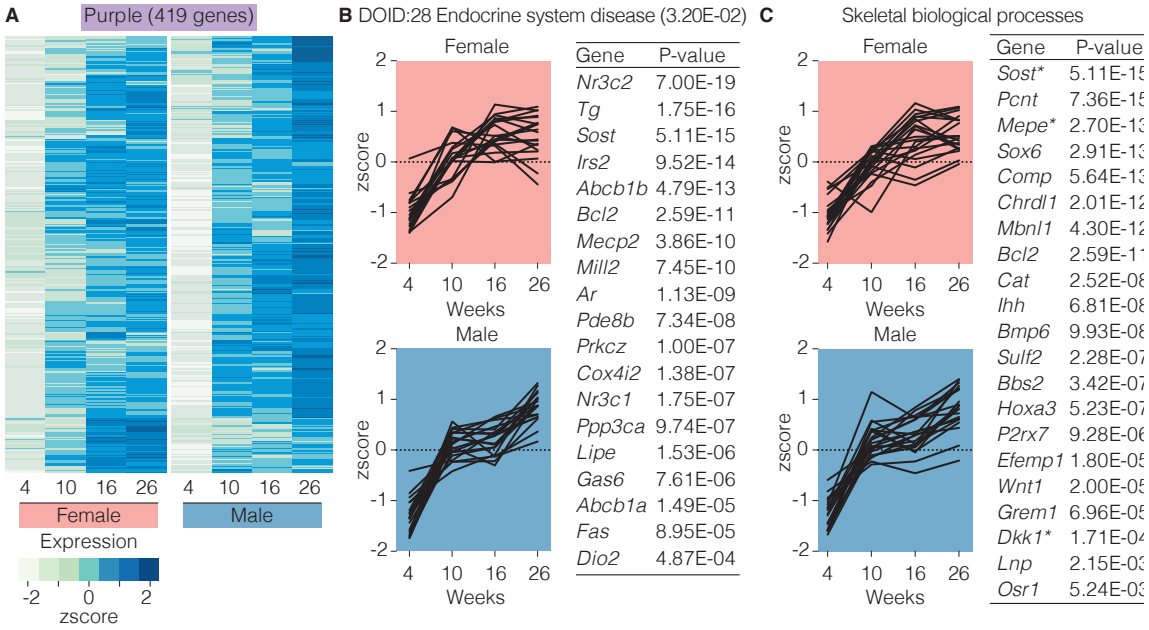
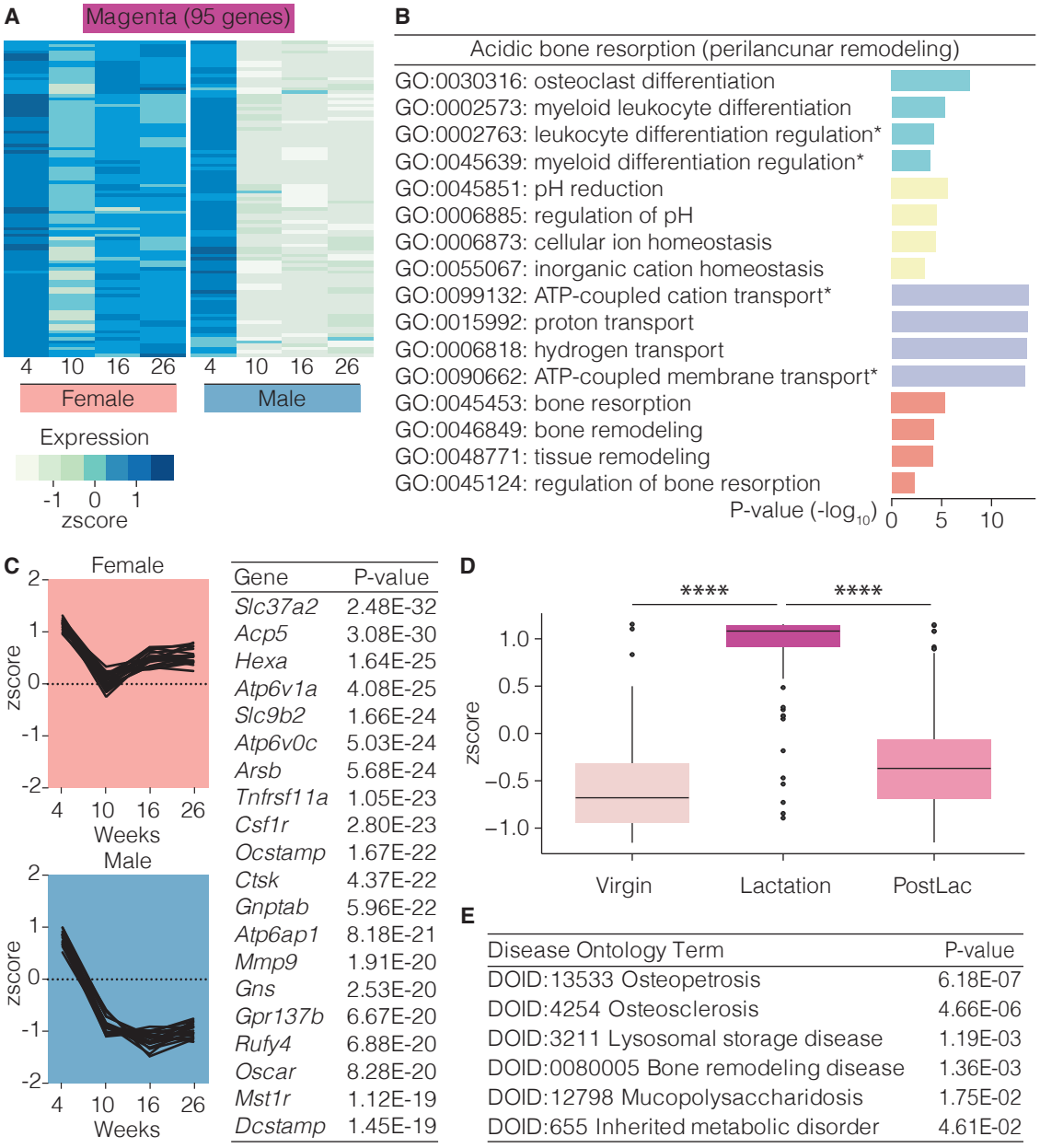


Figure 18 - Genes associated with endocrine diseases and osteocyte markers are upregulated with age in both sexes.

(A) Purple cluster expression during skeletal maturation in both sexes. (B) Expression of 19 genes in the Purple cluster associated with the significantly over-represented Disease Ontology term 'Endocrine System Disease' during skeletal maturation in both sexes (FDR ≤ 0.05). (C) Expression of 21 genes identified in the Purple cluster that are associated with skeletal biological processes annotated in the GO database. * known osteocyte marker genes. Axes reflect mean of z-scores of normalised counts at each time point.

5.2.4 Sexually dimorphic expression of perilacunar-remodelling processes during skeletal maturation.

Identified in the magenta cluster were 95 genes with sex specific patterns of expression during skeletal maturation (*Figure 19* A). At 4 weeks, magenta cluster expression was approximately equal between sexes and collectively higher than all other time points. Between 4-10 weeks magenta cluster expression was downregulated in both sexes, with this reduction greater in male than female samples. The difference between sexes increased with skeletal maturity, as cluster expression was gradually upregulated in female



Gene	P-value
Slc37a2	2.48E-32
Acp5	3.08E-30
Hexa	1.64E-25
Atp6v1a	4.08E-25
Slc9b2	1.66E-24
Atp6v0c	5.03E-24
Arsb	5.68E-24
Tnfrsf11a	1.05E-23
Csf1r	2.80E-23
Ocstamp	1.67E-22
Ctsk	4.37E-22
Gnptab	5.96E-22
Atp6ap1	8.18E-21
Mmp9	1.91E-20
Gns	2.53E-20
Gpr137b	6.67E-20
Rufy4	6.88E-20
Oscar	8.28E-20
Mst1r	1.12E-19
Dcstamp	1.45E-19

Disease Ontology Term	P-value
DOID:13533 Osteopetrosis	6.18E-07
DOID:4254 Osteosclerosis	4.66E-06
DOID:3211 Lysosomal storage disease	1.19E-03
DOID:0080005 Bone remodeling disease	1.36E-03
DOID:12798 Mucopolysaccharidosis	1.75E-02
DOID:655 Inherited metabolic disorder	4.61E-02

Figure 19 - Sexually dimorphic expression of genes associated with perilacunar-remodelling during skeletal maturation.

(A) Magenta cluster expression during skeletal maturation in both sexes. Values reflect mean of z-scores at each time point. Z-scores calculated based on the normalised expression across all 40 male and female samples. (B) Semantically similar clusters of significantly enriched GO biological processes associated with genes in the Magenta cluster (FDR ≤ 0.05). * truncated description. (C) Expression of the top 20 genes associated with the Magenta cluster (ranked by p-value of correlation with module eigengene). Axes reflect mean of z-scores at each time point. (D) Magenta cluster expression in virgin mice, lactating mice and mice post lactation (PostLac). Box plots reflect median and inter-quartile range of z-scores for magenta cluster genes at each time point. Z-scores calculated based on the normalised expression across all replicates of all three conditions (**** $p < 0.0001$). (E) Disease Ontology terms significantly enriched in the Magenta cluster (FDR ≤ 0.05).

samples at 16 and 26 weeks, while expression in male samples remained stably restrained. This pattern of coregulation indicated genes in the magenta cluster are upregulated in both sexes before sexual maturity and preferentially expressed in females as the skeleton matures.

Next, we identified biological processes overrepresented among the co-expressed genes of the magenta cluster. Four groups of semantically similar biological processes were identified among significantly enriched GO terms, associated with bone remodelling, extracellular pH reduction, ATP-coupled proton secretion, and terms related to osteoclast and myeloid cell differentiation (*Figure 19* B). While the process of bone resorption is usually associated with osteoclast function, osteocytes are also capable of directly resorbing their mineral surrounds during perilacunar-remodelling. This process involves the osteocytic-upregulation of many genes conventionally associated with osteoclastic bone resorption, a number of which were among the most highly correlated with magenta module expression (*Figure 19* C). These included Cathepsin K (*Ctsk*), the tartrate resistant acid phosphatase *Acp5* (TRAP) and many components of vacuolar ATPase family. Together, these observations indicated that the genes differentially regulated between sexes during skeletal maturation identified by the magenta cluster may be involved in perilacunar-remodelling.

Challenging this notion were a number of osteoclast marker genes not previously associated with perilacunar-remodelling that were also significantly correlated with the magenta cluster expression, such as the Osteoclast Stimulatory Transmembrane Protein *Ocstamp* and Osteoclast Associated Immunoglobulin-Like Receptor *Oscar* (*Figure 19* C). Their detection raised the possibility that the pattern of magenta cluster expression may be due to the differential retention of bone-resorbing osteoclasts in the *in-vivo* isolated bone samples, and not a product of gene expression in the osteocyte-network. To examine this alternative explanation, we utilised a publicly available microarray dataset which measured the transcriptome osteocytes during lactation (Qing et al., 2012). Much of what is known about perilacunar-remodelling is centred on its role in lactation, with osteocytes acidifying their microenvironment to rapidly liberate bone mineral required for milk production (Jähn et al., 2017; Qing et al., 2012). The handful of genes known to be involved in this process are upregulated in lactating mice compared to virgin animals and then downregulated post-

lactation as lacunae size returns to normal. With our data indicating the magenta cluster was enriched for genes and processes associated with perilacunar-remodelling, we hypothesised genes in the magenta cluster may show a similar pattern of lactation specific upregulation. Importantly, the primary osteocyte samples used to generate this microarray data were collected by 3-step sequential collagenase-digestion of bone fragments, a method demonstrated to be highly effective at removing cells from the bone surface, including osteoclasts (Gu, Nars, Hentunen, Metsikkö, & Väänänen, 2006; Qing et al., 2012). Supporting our hypothesis, a competitive gene set test showed significant upregulation of magenta genes between virgin and lactating animals (p=6.38E-33), which reversed post-lactation (p=1.52E-16) (*Figure 19* D). This lactation-specific upregulation was unique to the magenta cluster, with no significant difference was observed in the other 7 clusters (not shown). Detection of these genes in an independent transcriptome dataset generated by distinct methodologies provided strong evidence that genes commonly used as osteoclast markers are also dynamically expressed in the osteocyte network. Moreover, this indicated that magenta cluster genes, associated with bone resorption that display dimorphic expression patterns between sexes during skeletal maturation, are also upregulated during lactation, consistent with a role in perilacunar-remodelling.

Lastly, we examined the association of magenta cluster genes with diseases listed in the Disease Ontology database (*Figure 19* E). Among the significantly overrepresented disease terms were several disorders of the skeleton, including osteopetrosis (DOID:13533 p=6.18E-07), osteosclerosis (DOID:4254 p=4.66E-06) and bone remodelling diseases (DOID:080005 p=1.36E-03). These terms identify many genes typically associated with bone resorption by osteoclasts that have also been implicated in perilacunar remodelling by osteocytes such as *Acp5* and *Ctsk* (Hayman et al., 1996; Inaoka et al., 1995; Qing et al., 2012). This indicated that genes with sex-specific patterns of expression during skeletal maturation were associated with skeletal disease affecting the accrual and regulation of bone mass.

5.3 Discussion

Here we examine the transcriptome of the osteocyte network during skeletal maturation, examining gene expression in both sexes at key stages of postnatal growth. The role of the osteocyte network regulating the amount and distribution of bone tissue is

well established, yet its association with the skeletal changes that occur during postnatal development is poorly defined. In this work we show that the morphological changes that occur with age and sex are accompanied by changes in gene expression in the osteocyte network, with age-related differences most obvious in young mice and sex -related differences becoming apparent later in skeletal maturation.

Accompanying changes in morphology are a number of biological processes dynamically regulated in the osteocyte network during skeletal maturation. Gene co-expression network analysis revealed the age-related downregulation of processes associated proteoglycan synthesis and the upregulation of a range of endocrine signalling molecules in both sexes. Proteoglycans are important non-collagenous constituents of the bone extracellular matrix important for bone growth and mineralisation, while endocrine signalling is critical in the osteocyte networks management of skeletal-mineral reserves (Dallas et al., 2013; Nikitovic et al., 2012). The divergent regulation of these processes may suggest the osteocyte network plays an important role synthesising of bone matrix components during skeletal growth, however at skeletal maturity endocrine mineral regulation becomes a primary focus. Conversely, there is evidence to suggest that the material, structural and chemical composition of cortical bone changes during postnatal skeletal maturation (Fedarko, Vetter, Weinstein, & Robey, 1992; Miller, Martin, Piez, & Powers, 1967; Torzilli, Takebe, Burstein, Zika, & Heiple, 1982). As gene expression in the osteocyte network is also influenced by factors in the surrounding bone matrix, it is possible the correlations between age and the osteocyte transcriptome reported here result from changes in bone material. Investigations using lineage and age specific genetic engineering technologies are required to discern the cause-effect relationship between the osteocyte network and the changes in bone tissue that occur with skeletal maturation.

This analysis also revealed the sexually dimorphic regulation of perilacunar-remodelling, the process by which the osteocyte network directly remodels its surrounding bone tissue (Qing & Bonewald, 2009). The ability of osteocytes to remove and replace their perilacunar matrix was first reported in patients with Rickets and osteomalacia over 100 years ago and yet still very little is known about the molecular processes involved (v Recklinghausen, 1910). Only a small number of genes involved in perilacunar-remodelling have been identified, many for their role in lactation where they are specifically upregulated

during milk production (Qing et al., 2012). Interestingly, the 95 genes with sexually dimorphic expression patterns in our data are also upregulated during lactation, consistent with their involvement in perilacunar remodelling and suggesting they may too be involved in milk production. Beyond this sex-specific setting, the deletion of the metalloprotease *Mmp13,* inhibition of TGF-Beta signalling and glucocorticoid treatment have all been shown to disrupt perilacunar-remodelling (Dole et al., 2017; Fowler et al., 2017; Tang et al., 2012). These studies associate the disruption of perilacunar-remodelling with an increase in bone fragility and suggest it is an important process for bone quality and strength. Our data suggests that genes associated with perilacunar-remodelling are differentially regulated during skeletal maturation and thus the effects of disruption may be somewhat influenced by age (discussed in detail in section 8.2).

6 THE OSTEOCYTE TRANSCRIPTOME SIGNATURE

As cells differentiate they specialise in their function, and with it their transcriptome (Ng et al., 2008). Genes involved in cell-defining molecular processes are upregulated, often accompanied by changes in cell morphology suited to their refined role. Osteocyte differentiation is marked by distinct morphological and functional stages, transitioning from osteoblasts to early osteocytes before reaching maturity (Franz-Odendaal et al., 2006). A key feature of this specialisation is the formation of dendrites and integration into the osteocyte network, used to communicate and coordinate skeletal regulation. A small number of gene-markers specifically expressed by osteocytes within the skeleton have been discovered, however little is known of the molecular control pathways used in the formation and specialised function of the osteocyte network (Plotkin & Bellido, 2016).

Due to their location, morphology and terminally differentiate state, studying osteocytes is technically challenging. These barriers have meant their omission from all the major transcriptome databases and consequently, the most frequently used transcriptome annotations have been built without gene expression data from the osteocyte network. By omitting these cells, genes and transcripts with expression restricted to the skeleton may yet to be discovered. Without these genes, our understanding of the transcriptome active in the skeleton, and the body, is incomplete.

In this chapter, we document an approach to discovering the genes enriched for expression in the osteocyte network and use it to define an osteocyte transcriptome signature. We identify biological processes and pathways enriched in this osteocyte defining gene list and reveal the molecular control processes upregulated in the differentiation and formation of the osteocyte network.

6.1 Methods

6.1.1 Sample details and transcriptome data generation

Ethical approval for all animal experiments reported here is detailed in section 2.1.1 and the Osteocyte Enrichment Cohort is detailed in section 2.1.5. Briefly, bones from the left-side of the body were processed to obtain *in-situ* isolated osteocytes section 2.2.1,

while bones from the right-side of the body had all soft tissue removed yet the marrow left intact. Transcriptome data from these samples was generated as per section 2.2.4, and data was processed as per section 2.2.6.

6.1.2 Identification of genes enriched for expression in the osteocyte network

Genes enriched for expression in the osteocyte network relative to other cell lineages in the bone marrow were identified as per section 2.2.20. Genes upregulated in osteocytic differentiation and network formation were identified as per section 2.2.21. Genes actively expressed in *in-situ* isolated osteocyte samples that were absent *in-vitro* were as per section 2.2.21. The skeletally enriched expression of genes that were absent in-vitro was determined as per section 2.2.13.

6.1.3 Definition and characterisation of the osteocyte signature

The criteria used to define the osteocyte transcriptome signature are documented in section 2.2.22. Osteocyte signature genes associated with skeletal biological processes in the Gene Ontology database or that produced skeletal phenotypes when knocked out in mice as listed in the Mouse Genome Informatics database were identified as per section 2.2.12. The expression and structure of novel genes identified by *de-novo* transcriptome assembly contained in the osteocyte signature were visualised as per section 2.2.24. Clusters of genes with similar temporal expression patterns during osteocytic-differentiation and network formation were identified in data described in section 2.2.21.

6.2 Results

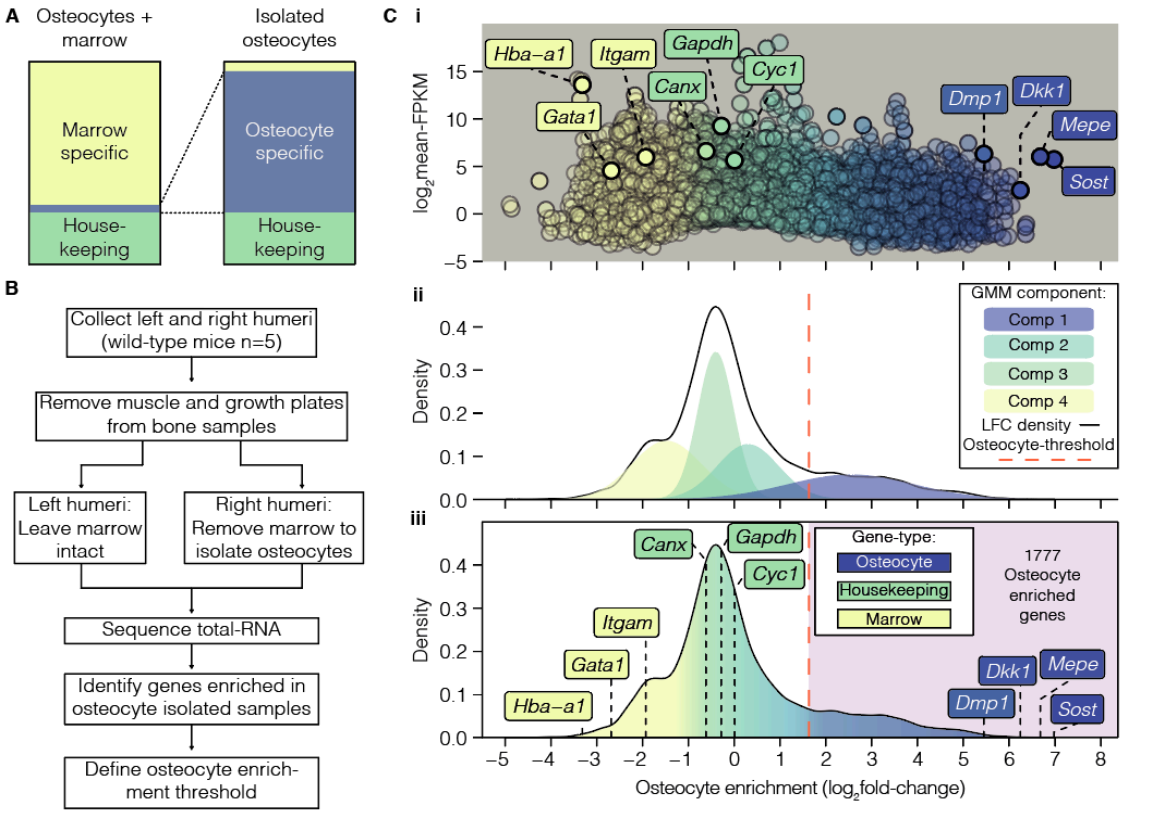
6.2.1 Defining the osteocyte transcriptome signature

To identify the genes that distinguish osteocytes from other cell types, we first identified genes enriched for expression in osteocytes relative to the other cell lineages in the marrow space. We hypothesised that expression of genes contributing to osteocyte-defing functions would be somewhat restricted to osteocytes relative to other cell types, and thus would be enriched for data in samples enriched for osteocytes (Figure *20* A). In contrast, housekeeping genes contributing to common cellular functions are expressed in all cells and thus would be unaffected by sample composition. To test this, we developed a subtractive sequencing approach which compared RNA sequence data of *in-situ* isolated

osteocytes to samples in which bone marrow cells were retained (Figure *20* B). Osteocytes comprised >90% of the cells in the *in-situ* isolated samples but <1% when marrow cells were retained.

Known markers of osteocytes were among the genes most enriched for data by *in-situ* osteocyte isolation (Figure *20* C i). Greater than 100-fold-enrichment (FE) of data was observed for Sclerostin (*Sost*) and Matrix extracellular phosphoglycoprotein (*Mepe*), while Dickkopf-related protein 1 (*Dkk1*) was enriched >70-fold and Dentin matrix acidic phosphoprotein 1 (*Dmp1*) >40-fold. In contrast, the abundance of data for known housekeeping genes was unaffected by changes in the cellular composition, with little change observed for Glyceraldehyde 3-phosphate dehydrogenase (*Gapdh*), Calnexin (*Canx*), and Cytochrome C1 (*Cyc1*). Genes typically found in cell populations known to reside within the marrow space were depleted by marrow removal, such as Hemoglobin subunit alpha 1 (*Hba-a1*), Integrin subunit alpha M1 (*Itgam1*), and GATA-binding factor 1 (*Gata1*). These data indicated the subtractive sequencing approach resulted in the preferential enrichment of genes known to be specifically expressed in the osteocyte network.

Next, we calculated a threshold of enrichment to identify genes significantly enriched in *in-situ* isolated osteocyte samples. We fit a Gaussian Mixture Model to the distribution of gene enrichment between osteocyte-enriched and marrow containing samples (Figure *20* C ii) (a detailed methodological explanation is contained in section 2.2.20). This approach identified 4 component sub-populations with distinct levels of enrichment in osteocyte isolated bone samples. Skeletally related Gene Ontology terms were only identified in the most osteocyte-enriched component (Component 1). Component 2 and component 3 identified genes that were relatively constant between sample types (mean log fold-change ~0). No significant processes were identified in Component 2 while Component 3 contained terms associated with macromolecular synthesis and processing, consistent with general housekeeping processes (Component 3) (Figure *20* D). Component 4 identified genes that were depleted in osteocyte-enriched samples and contained terms associated with cell replication and marrow lineages. This indicated that the GMM model



D

Component 4 GO Terms	Component 3 GO Terms	Component 1 GO Terms
GO:0071103:DNA conformation change	GO:0016071:mRNA metabolic process	GO:0001501:skeletal system development
GO:0007059:chromosome segregation	GO:0034660:ncRNA metabolic process	GO:0001503:ossification
GO:0000280:nuclear division	GO:0032259:methylation	GO:0006936:muscle contraction
GO:0098813:chromosome segregation	GO:0016570:histone modification	GO:0060537:muscle tissue development
GO:0034728:nucleosome organization	GO:0006397:mRNA processing	GO:0007517:muscle organ development
GO:0006333:chromatin assembly	GO:0022613:ribonucleoprotein biogenesis	GO:0003012:muscle system process
GO:0071824:protein-DNA organization	GO:0006399:tRNA metabolic process	GO:0014706:striated muscle development
GO:0051301:cell division	GO:0008213:protein alkylation	GO:0030198:extracellular matrix
GO:0051321:meiotic cell cycle	GO:0043414:macromolecule methylation	GO:0043062:extracellular structure
GO:0051290:protein heterotetramerization	GO:0070585:protein localization to mito*	GO:0007507:heart development
GO:0010564:regulation of cell cycle process	GO:0007005:mitochondrion organization	GO:0051216:cartilage development
GO:0006261:DNA-dependent DNA replication	GO:0071826:ribonucleoprotein organization	GO:0061448:connective tissue development

Figure 20 - Identification of genes enriched for expression in the osteocyte network.

(A) Graphical hypothesis of subtractive-sequencing approach to identify genes enriched for expression in the osteocyte network. (B) Experimental design used to identify genes enriched for expression in osteocytes, comparing data from bone samples with the marrow retained and bone samples with the marrow removed and osteocytes isolated. (C) Gene expression data enrichment in *in-situ* isolated osteocyte samples. Gene enrichment as a function of normalised gene expression (i) highlighting known osteocyte enriched genes, housekeeping genes and markers of marrow cell populations. Osteocyte-enrichment density distribution (ii) overlayed with a 4 component gaussian mixture model used to identify the osteocyte enrichment threshold (red dashed line). Osteocyte-enrichment density distribution (iii) showing highlighting osteocyte-markers among genes significantly enriched in *in-situ* isolated osteocyte samples. (D) Significantly enriched GO terms among the top 1000 genes ranked by posterior probability of association for each GMM component (FDR ≤ 0.05). No significant terms were identified for component 2. * truncated description.

components were associated with distinct biological functions, distinguishing genes associated with skeletal processes from those with housekeeping functions. A threshold of osteocyte enrichment was calculated at the upper 95% confidence interval of Component 2, with genes above this threshold deemed to be significantly enriched by osteocyte isolation and thus enriched for expression in the osteocyte network. In total, this subtractive sequencing methodology identified 1777 genes significantly enriched in osteocytes relative to other cell lineages in the marrow space (Figure *20* C iii). Of those, 1439 were actively expressed in osteocytes from all bone types, the tibia, femur and humerus.

The next step to identifying genes contributing to osteocyte-defining functions was the identification of genes enriched in osteocytes relative to other cells of osteogenic origin. To do this we utilised a transcriptome data from the osteogenic IGD-SW3 cell-line which models osteocytic-differentiation from osteoblast-like cells to mature-osteocyte-like stages (St. John et al., 2014). This step also removed genes which were predominantly expressed in non-osteogeneic cell types, such as blood vessels and nerves, which would be retained in the *in situ* bone tissue samples but would not be present *in vitro* (Klingberg et al., 2017; Mach et al., 2002; Núñez et al., 2018). Differential expression analysis revealed that 5417 of the 13365 genes actively expressed in osteocytes from the tibia, femur and humerus were significantly upregulated in osteocytic cells (day 14 or day 35) relative to the osteoblast-like time-point (day 3). These large-scale transcriptome changes were consistent with the initial reports from this dataset. Importantly, it was also noted that a number of genes highly expressed *in-vivo* were absent in the *in-vitro* system, including *Mepe*, a gene demonstrated to be highly expressed in mature osteocytes by orthogonal techniques (Igarashi et al., 2002). In total, 900 genes actively expressed in the tibia, femur and humerus *in-vivo* were absent *in-vitro*, reinforcing the importance of the osteocyte *in-situ* context in understanding gene expression in the osteocyte network.

Finally, we established criteria to integrate these analyses and identify an osteocyte transcriptome signature – genes whose combined expression defines the osteocyte network (*Figure 21*). These genes were actively expressed in osteocytes from the tibia, femur and humerus and enriched in osteocytes relative to other cells in the marrow space. Of these, 780 were upregulated in osteocytic-differentiation *in-vitro*, and a further 50 genes were most highly expressed in osteocytes relative to 12 other tissues types were not

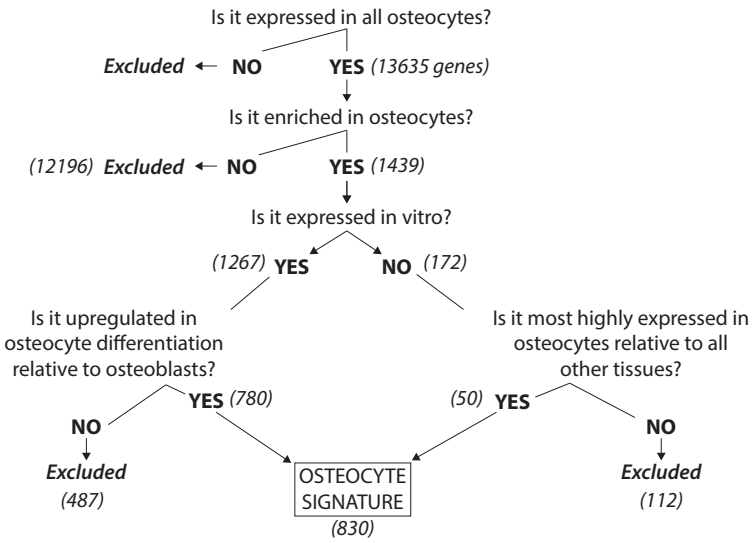


Figure 21 - Osteocyte transcriptome signature inclusion criteria.

(A) A stepwise approach to identifying genes actively expressed in multiple skeletal sites and enriched for expression in the osteocyte network *in-vivo*. Numbers reflect the number of genes remaining at each filtering stage.

detected *in-vitro*. Ninety-five percent of these genes were actively expressed in the osteocyte network throughout skeletal maturation in both sexes, with the remaining 5% active in at least two time points in either sex. Together these criteria defined the 830 gene osteocyte transcriptome signature, genes actively expressed throughout the skeleton and enriched in osteocytes *in-vivo* (Appendix 2).

6.2.2 The osteocyte signature identifies genes enriched in the osteocyte network with no known skeletal function

Given the critical role of osteocytes in skeleton, we hypothesised that the osteocyte transcriptome signature would be enriched for skeletally important genes. Supporting this hypothesis was the significant overrepresentation of genes associated with skeletal biological processes in the Gene Ontology database (GO, 3.7E-35) and skeletal phenotypes in the Mouse Genome Informatics database (MGI, 4.3E-21) (*Figure 22*). However, even with this enrichment of known genes, the vast majority of signature genes (77%) were not identified with a skeletal association, indicating their function in the skeleton is poorly defined. These included many of the genes most enriched by *in-situ* osteocyte isolation, such as the neuronal guanine nucleotide exchange factor *Ngef*, the extracellular signalling

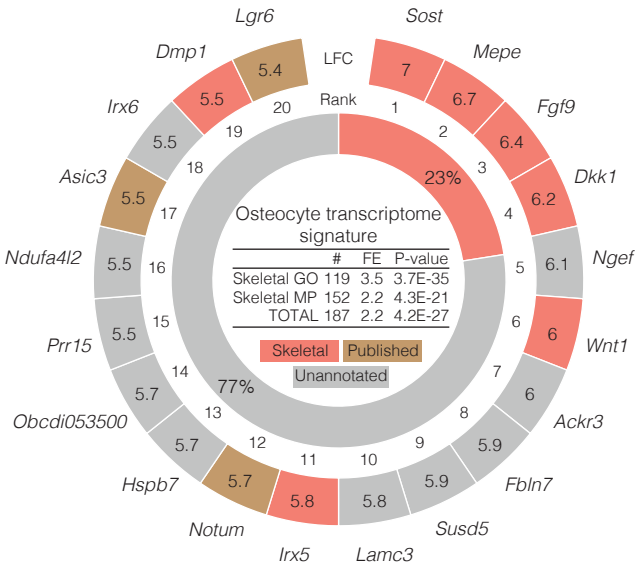


Figure 22 - The osteocyte transcriptome signature contains skeletally known and novel genes.

The proportion of genes annotated with a skeletal gene ontology term (GO annotated) or skeletal phenotype (MGI annotated) in the osteocyte transcriptome signature, highlighting GO annotated, unannotated genes with a published skeletal association and unannotated genes among the top 20 most osteocyte-enriched signature genes. LFC = Log_2 fold-enrichment in gene expression in *in-situ* isolated osteocyte samples.

component Fibulin 7 (*Fbln7*), as well as the atypical chemokine receptor *Ackr3*. Thirty-four lncRNAs were identified in the signature, including *H19* and the maternally expressed gene *Meg3*, known to play important regulatory roles in osteogenic differentiation (Liang et al., 2016; Zhuang et al., 2015). This indicated that while many genes in the osteocyte signature are known to play important roles in the skeleton, for many, enriched expression in the osteocyte network was the first evidence of potential skeletal involvement.

To establish whether osteocyte transcriptome signature genes that had not previously associated with bone biology may have a functional role in the skeleton, we identified osteocyte signature genes that had been screened in the Origins of Bone and Cartilage Disease skeletal phenotyping pipeline (*Table 6*). The OBCD pipeline performs skeletal phenotyping on unselected knockout mice lines generated in the Wellcome Trust Knockout Pipeline, reporting structural and functional assessment of gene knockouts independent of previous association with skeletal biology (Bassett et al., 2012). Fifty of the 626 screened by the OBCD pipeline were identified in the osteocyte transcriptome signature, nearly half of which (23) were identified with significant skeletal phenotype (46%, p=0.09). These included phenotypes affecting bone structure (17 lines), bone strength (11 lines) and both bone structure and strength (6 lines) (*Table 7*). Eighteen of these 23 genes had not previously

Table 6 - OBCD skeletal phenotyping statistics

OBCD Outlier Phenotype	OBCD (Total 626)	Signature (Total 50)	P value	Odds Ratio
Structure and/or Strength	211 (34%)	23 (46%)	0.09	1.675
Structure	144 (23%)	17 (34%)	0.09	1.724
Strength	96 (15%)	11 (22%)	0.22	1.616
Structure and Strength	45 (7%)	6 (12%)	0.26	1.761

Total = total knockout lines screened in OBCD pipeline
Numbers reflect number of knockout lines with skeletal phenotype
% reflects percent of total screened

Table 7 - OBCD knockout mice lines in the osteocyte transcriptome signature

Allele	Genotype	Structure	Strength	Skeletal GO
Amotl1	HOM	TRUE	TRUE	
Arl4d	HOM	TRUE		
Auts2	HET	TRUE		
Bgn	HOM	TRUE		TRUE
Bhlhe40	HOM	TRUE	TRUE	
Cadm1	HOM	TRUE	TRUE	TRUE
Cc2d2a	HET	TRUE		
Cfh	HOM	TRUE		
Daam2	HOM	TRUE	TRUE	
Dact3	HOM	TRUE	TRUE	
Kazn	HOM	TRUE		
Klhl30	HOM	TRUE		
Lamc3	HOM		TRUE	
Ldlrad4	HOM	TRUE		
Ltbp1	HOM	TRUE		
Mamstr	HOM	TRUE		
Matn4	HOM	TRUE		
Mtss1l	HOM		TRUE	
Plxna2	HOM		TRUE	TRUE
Slc9a3r2	HOM	TRUE		
Sparc	HOM	TRUE	TRUE	TRUE
Spns2	HOM	TRUE	TRUE	TRUE
Zcchc14	HET	TRUE		

HOM = Homozygous gene knockout
HET = Heterozygous gene knockout (HOM lethal)
Skeletal GO = Skeletal Gene Ontology biological process

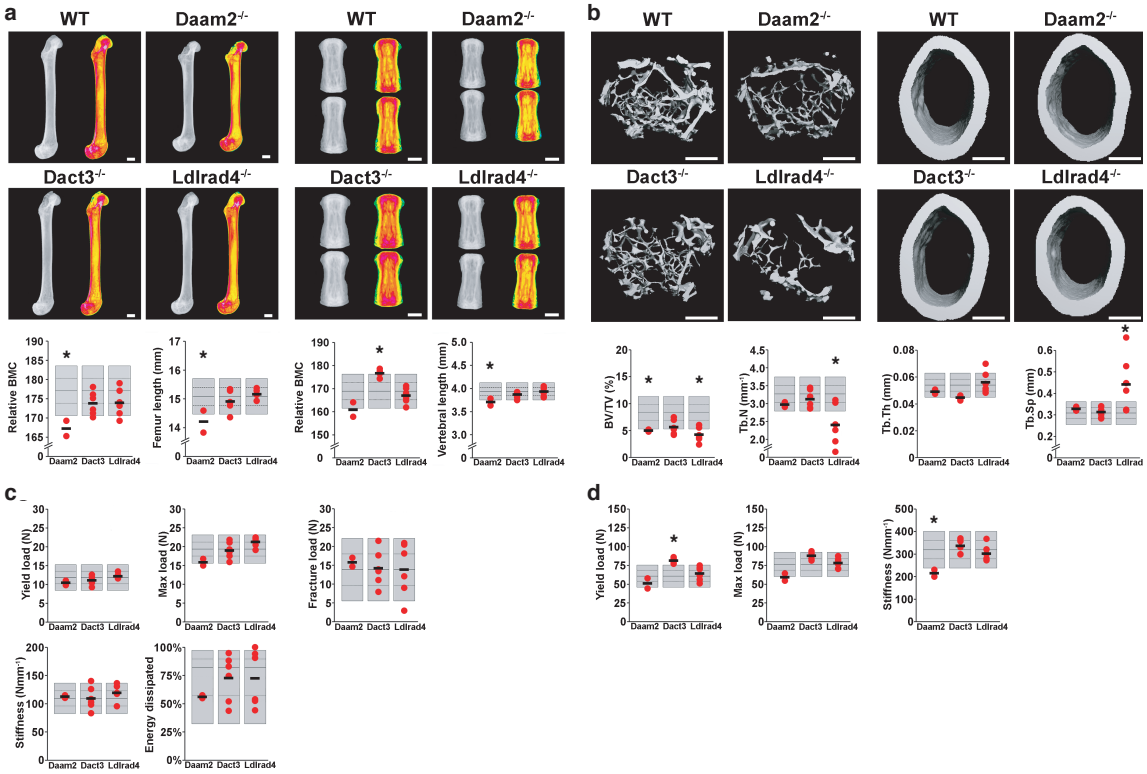


Figure 23 - Skeletally novel osteocyte signature genes produce skeletal phenotypes in knockout mice.

(A) Bone mineral content of femur and caudal vertebrae from wild-type (WT) and knockout mice lines (*Daam2*$^{-/-}$, *Dact3*$^{-/-}$, *Ldlrad4*$^{-/-}$) at postnatal day 112 (P112). Pseudo-colored grey-scale images in which low mineral content is green and high mineral content is pink. Graphs show femur bone mineral content (BMC), femur length, vertebral BMC and vertebral length. (B) Trabecular and cortical bone parameters of femur trabecular bone (left) and mid-diaphysis cortical bone (right) from WT and Ldlrad4-/- mice determined by µCT. Graphs show trabecular bone volume/tissue volume (BV/TV), trabecular number (Tb.N), Trabecular thickness (Tb.Th) and trabecular spacing (Tb.Sp). (C) Femur biomechanical analysis by destructive 3-point bend testing. Graphs show yield load, maximum load, fracture load, stiffness and energy dissipated prior to fracture (Toughness) in WT animals (n>320). (D) Vertebra biomechanical analysis by destructive compression testing. Graphs show yield load, maximum load, stiffness. Reference ranges represent values from female WT mice (n=320) of identical C57Bl6 genetic background to knockout lines with mean (solid line), 1.0SD (dotted line) and 2.0SD (grey box). The individual values for knockout animals are shown with red dots and a thick black line indicating the mean value. * Mean values outside the reference range were considered significant.

been associated with a skeletal biological process nor phenotype, suggesting their role in the skeleton had not previously been identified. Among those were the Disheveled-associated activator of morphogenesis *Daam2*, the dapper antagonist of Beta-catenin *Dact3* and the low-density lipoprotein receptor class A domain-containing gene *Ldlrad4* (Xia Jiang et al., 2008; Lee & Deneen, 2012; Nakano et al., 2014). Bone mineral content (BMC) was significantly affected in either the femur or vertebrae of all three lines, increased in *Dact3*$^{-/-}$ and decreased in *Daam2*$^{-/-}$ and *Ldlrad4*$^{-/-}$ knockouts (*Figure 23*). These changes in BMC were predominantly due to their effects on trabecular bone volume, with little change observed in cortical bone volume in any of the knockout lines. Consistent with these divergent effects on BMC, vertebral functional phenotyping indicated an increased yield load in *Dact3*$^{-/-}$ mice, while *Daam2*$^{-/-}$ showed decreased max load and stiffness with yield load at the lower end of the normal range. The identification of skeletally novel genes with a skeletal phenotype suggested that unknown genes in the osteocyte signature may have an important role in the skeleton.

6.2.3 The osteocyte transcriptome signature identifies novel skeletally restricted genes

Since primary skeletal tissue is typically omitted from major transcriptome analyses, we hypothesised that genes restricted in expression to osteocytes may remain to be discovered. One hundred and three novel assembled genes were actively expressed in in osteocytes from at least one bone type, 49 of which were active in osteocytes from all bones (Figure *24* A and B). Eleven of these were enriched for expression in osteocytes relative to the other cell lineages in the marrow space, which suggested their expression was predominantly in the cells within the mineralised tissue of the skeleton (Figure *24* C). In total, 7 novel assembled genes were identified in the osteocyte signature, which included *Obcdi053500* among the most highly enriched for expression in osteocytes, above the known osteocyte marker *Dmp1* (Figure *24* D).

Multiple isoforms were identified for many of the novel expressed genes, suggesting gene transcripts were post-transcriptional processed, while analysis of sequence coding potential indicated they were non-coding (Figure *24* D). Whilst the function of these novel genes was unclear, the location of *Obcda044440*, antisense to the collagen type 24 alpha 1 chain gene (*Col24a1*) also identified in the osteocytes signature, indicated it may influence the type I collagen fibril component. Six of the 7 genes were most highly enriched in

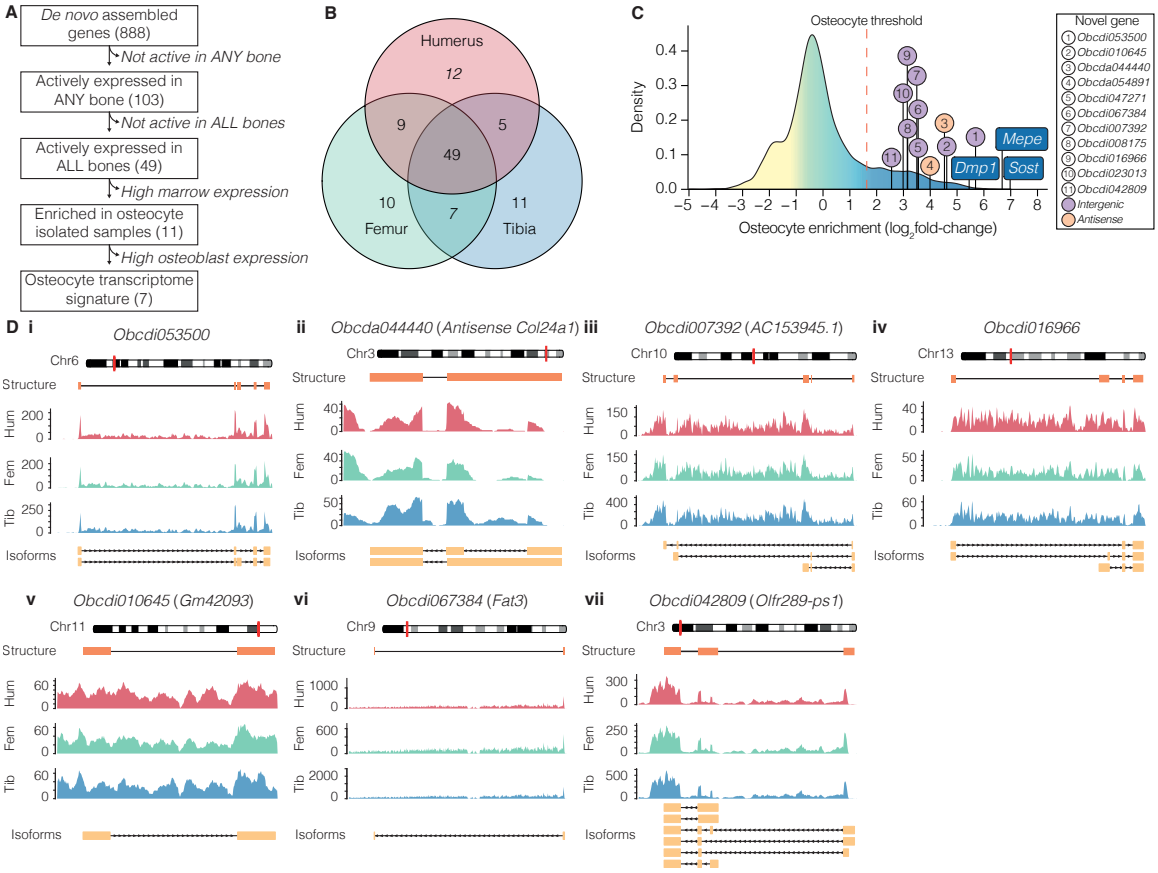


Figure 24 - Novel genes actively expressed in the osteocyte network.

(A) Strategy used to identify novel genes enriched for expression in the osteocyte network. (B) Venn diagram of the novel genes actively expressed in individual bone types including 49 genes in every long-bone type. (C) Eleven novel genes significantly enriched in osteocytes and actively expressed in all long-bones. Intergenic genes are located between annotated genes. Antisense genes are located on the opposite DNA-strand to annotated genes. (D) Novel osteocyte transcriptome signature gene structure diagrams with pooled read data alignment and isoforms.

---

osteocytes relative to 12 other tissue types with the 7th, *Obcdi007392* only detected in the heart outside of the skeleton (*Figure 25* A). Four of the novel signature genes were upregulated in osteocyte differentiation, including *Obcdi007392* and *Obcdi053500* that were restricted in their expression to mature cells (*Figure 25* B). The remaining three were not detected *in-vitro*, despite active expression in all *in-situ* isolated samples from the tibia, femur and humerus. This indicated their expression is somewhat dependent on the *in-vivo* osteocyte context.

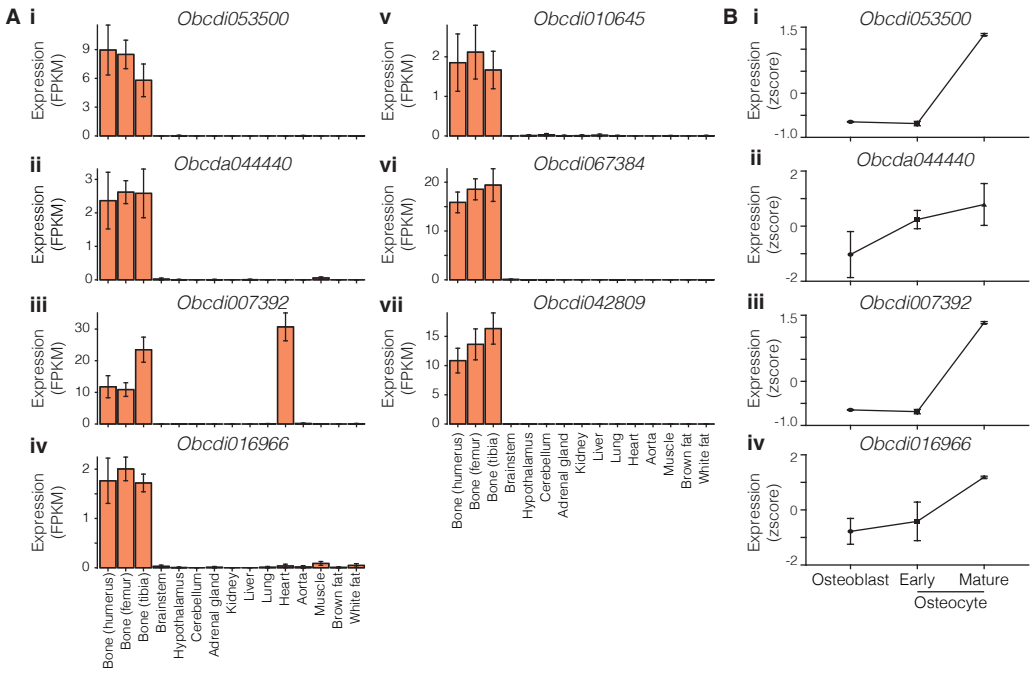


Figure 25 - Skeletally-restricted and maturation-regulated expression of novel signature genes.

(A) Novel osteocyte transcriptome signature gene expression in 12 tissue types relative to osteocytes from three bone types. (B) Temporal expression of novel osteocyte transcriptome signature gene expression in osteocytic differentiation and network formation. Osteoblast = Day 3 of differentiation, Early osteocyte = Day 14 and Mature osteocyte = Day 35). Expression of Obcdi010645, Obcdi067384 and Obcdi042809 was not detected *in-vitro*. Z-score of normalised counts. Error bars reflect mean and SD.

Together these data report novel genes largely restricted in expression to osteocytes, with several undetectable *in-vitro*. These data suggest that osteocytes express genes restricted in expression to the skeleton that had not previously been discovered.

6.2.4 Osteocyte network is enriched for external signalling pathways

To identify the molecular programs that potentially contribute to the highly specialised function of osteocytes within the skeleton, we analysed the signature for over representation of biological processes and biochemical pathways. One hundred and twenty-two biological processes were significantly enriched in the signature (Appendix 2). These included many processes known to be controlled by osteocytes, such as ossification (GO:0001503, p=8.82E-22), skeletal system development (GO:0001501, p=2.43E-17), biomineral tissue development (GO:0031214, p=1.30E-10) and osteoblast differentiation

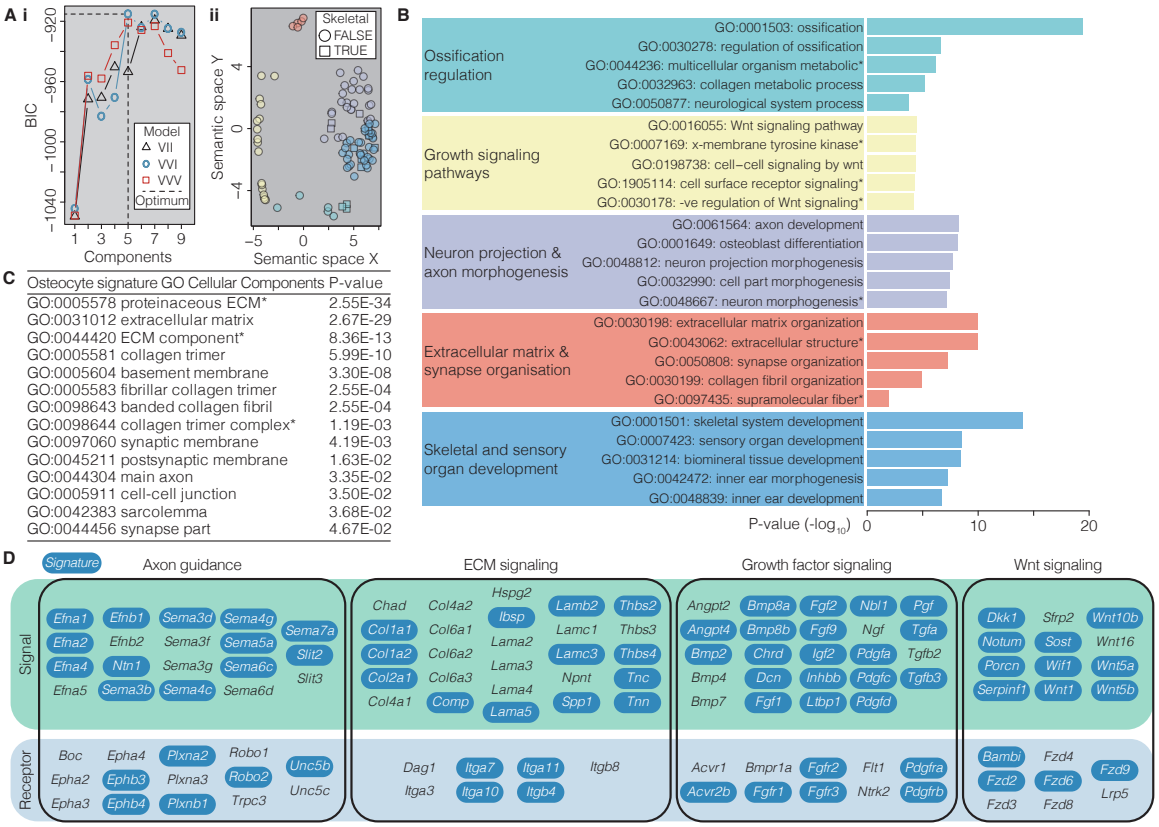


Figure 26 - Extracellular signaling processes are enriched in the osteocyte transcriptome signature.

(A) Semantic-clustering of significantly enriched GO biological processes in the osteocyte transcriptome signature based on Bayesian Information Criterion (BIC) and models of unequal variance (VII = spherical, unequal volume; VVI= diagonal, varying volume and shape; VVV = ellipsoidal, varying volume, shape, and orientation). Square symbols represent 'skeletal' GO terms, directly associated with skeletal biological processes. (B) Clusters of significantly enriched biological processes in the osteocyte transcriptome signature named by word frequency in term descriptions and the top 5 gene ontology terms ranked by p-value in each cluster (FDR ≤ 0.05). (C) Significantly enriched GO Cellular Component terms in the osteocyte transcriptome signature (FDR ≤ 0.05). (D) Signaling and receptor molecules significantly enriched for expression in the osteocyte network and contained in the osteocyte transcriptome signature (highlighted in dark blue).* truncated description

(GO:0001649, p=2.63E-9) among the most enriched. Significant terms were then grouped based on their semantic similarity to identify clusters of functionally similar biological processes (Figure *26* A). Two of the 5 clusters pertained to processes known to be important to skeletal composition, namely *ossification regulation* and *skeletal and sensory organ development* (Figure *26* B). The remaining 3 clusters identified several modes of extracellular signalling and contact organisation, namely *growth factor signalling*,

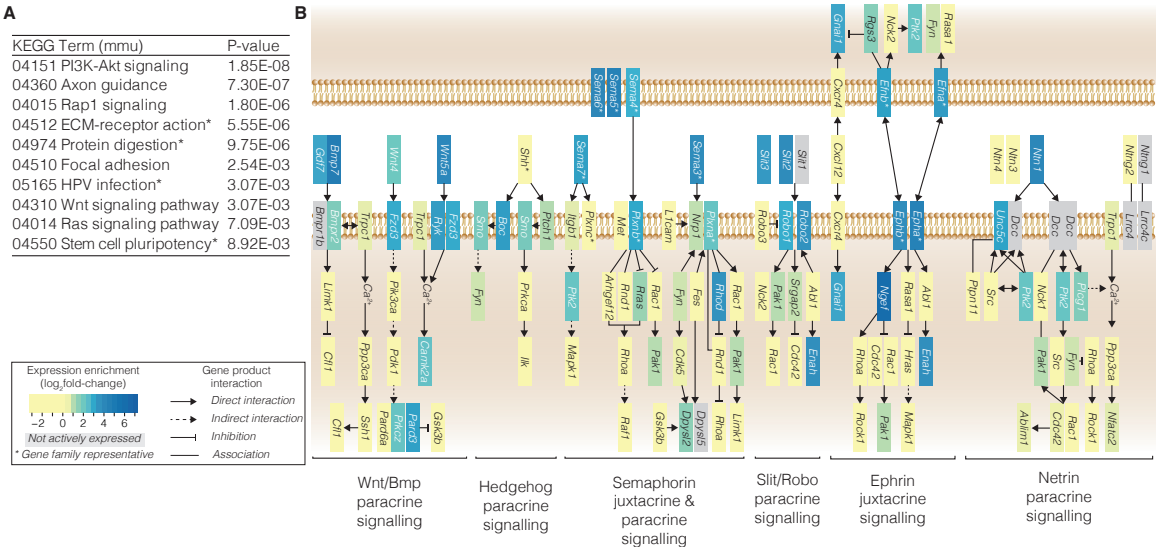


Figure 27 - Enriched expression of axon-guidance molecular machinery in the osteocyte network.

(A) KEGG biochemical pathways significantly enriched in the osteocyte transcriptome signature (Bonferroni adjusted p-value ≤ 0.05). (B) Axon guidance gene expression and enrichment in *in-situ* isolated osteocytes.

extracellular matrix and synapse formation, as well as *neuron projection and axon morphogenesis*. Enrichment analysis of Kyoto Encyclopaedia of Genes and Genomes (KEGG) database identified 'axon guidance' (mmu04360, p=7.30E-07) as one of the most significantly enriched biochemical pathway in the osteocyte transcriptome signature, supported by a significant enrichment of synapse and axon cellular components (Figure *26* C and *Figure 27* A). In total, 9% of the osteocyte transcriptome signature (75 genes) encoded signalling ligands and receptors, including 19 genes from axon guidance pathways (Figure *26* D). The comprehensive enrichment of ligands, receptors and downstream components involved in axon guidance indicated this may be an important control pathway in the osteocyte network (*Figure 27* B).

6.2.5 Neuron-like network formation pathways upregulated in osteocytic differentiation

To understand how the genes and pathways identified in the osteocyte transcriptome signature may contribute to osteocyte network formation, we identified distinct patterns of regulation during osteocytic differentiation (Figure *28* A and B). Four clusters of genes with temporal changes during osteocytic differentiation and network formation were identified:

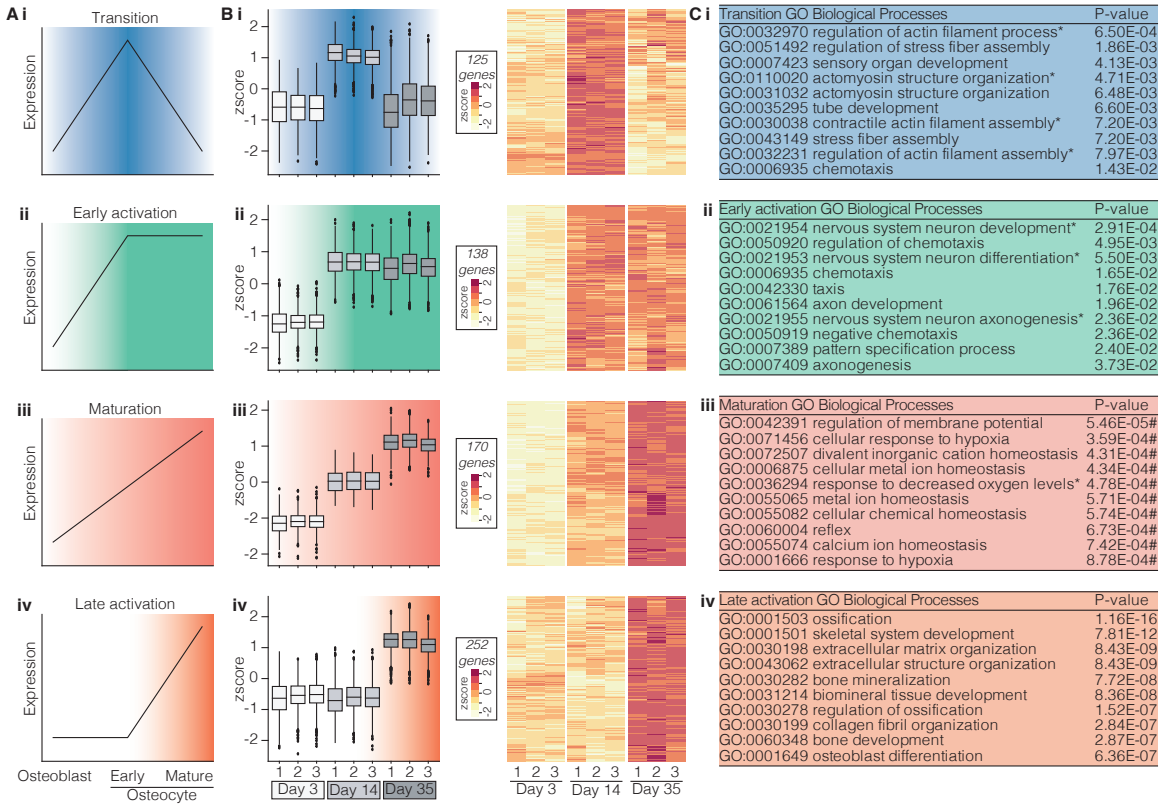


Figure 28 - Signature genes associated with distinct biological processes are temporally regulated during osteocytic differentiation.

(A) Predicted differential gene expression profiles during osteocytic differentiation and network formation. (B) Clusters of significantly differentially expressed osteocyte signature genes temporally regulated during differentiation from osteoblast-like cells (day 3) to osteocyte-like cells (day 14 or day 35). Z-score of normalised counts. (C) Significantly enriched GO biological processes in temporally regulated signature gene clusters. (FDR ≤ 0.05, # = unadjusted p-value, * truncated description).

- Transition cluster - 125 genes that were upregulated early in osteocyte differentiation and network formation which were then down regulated as the osteocyte network matured.
- Early activation cluster - 138 genes that were upregulated early in osteocyte differentiation and network formation which then remained stable through maturation.
- Maturation cluster - 170 genes that were upregulated early in osteocyte differentiation and network formation which continued to increase through maturation.

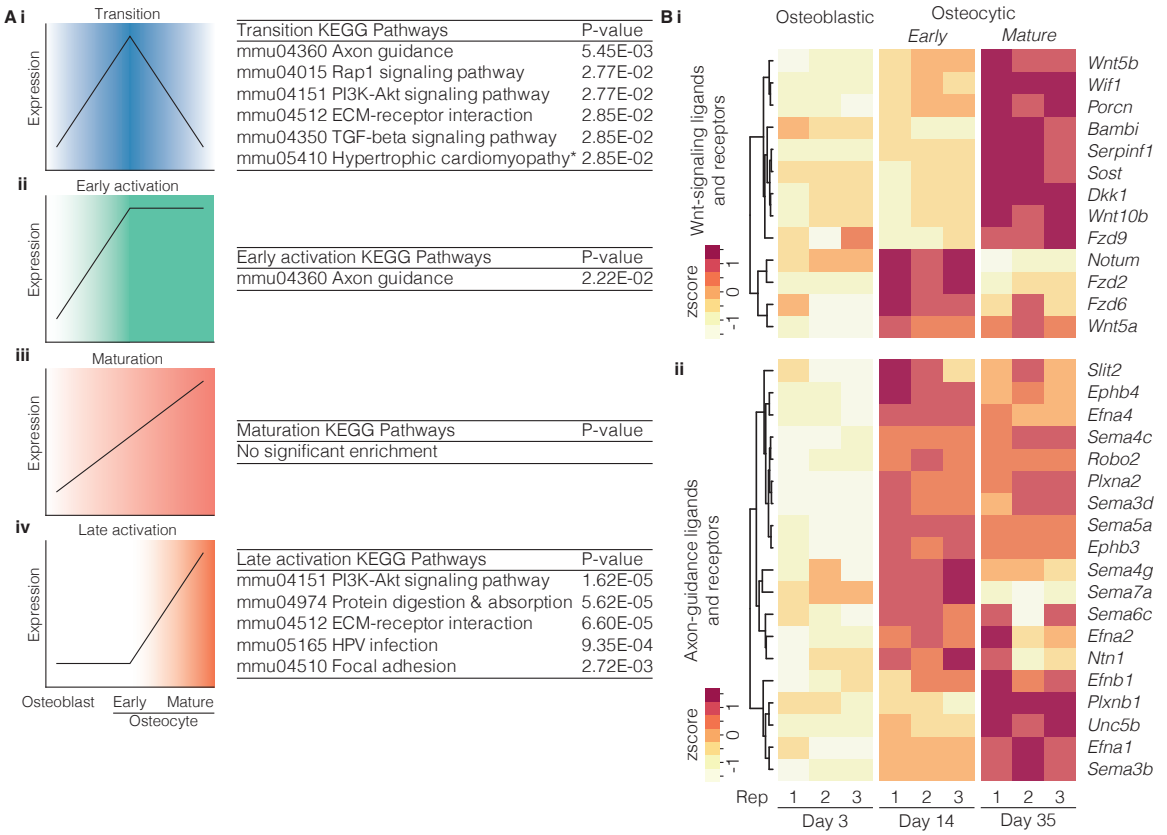


Figure 29 - Genes and pathways associated with network formation are upregulated early in osteocytic differentiation.

(A) KEGG biochemical pathways significantly enriched in Transition (i), Early activation (ii), Maturation (iii, none significant), and Late activation clusters. (FDR ≤ 0.05). (B) Expression of Wnt-signaling (i) and axon guidance (ii) ligands, modulators and receptors during osteocytic differentiation. Z-score of normalised counts. * truncated description

---

- Late activation cluster - 252 genes for which expression did not change between osteoblasts and early osteocytic cells which were then significantly upregulated in the mature network.

This clustering strategy was supported by several osteocyte markers that conformed with established differentiation-stage-specific expression patterns, including Podopalin (*Pdpn*) in the Transition cluster, *Dmp1* in the Maturation cluster and *Sost* in the Late-activation cluster (Dallas et al., 2013; K. Zhang et al., 2006).

GO and KEGG enrichment analysis demonstrated distinct biological processes and pathways upregulated at specific points of osteocyte maturation (Figure *28* C and *Figure*

29 A). Processes such as ossification, skeletal regulation and ion homeostasis, were enriched in *Late activation* and *Maturation* clusters, most highly expressed in the mature network. More than half of the Wnt-signalling ligands, receptors and modulators identified in the osteocyte signature were upregulated late in osteocytic-differentiation, consistent with the role of this pathway in skeletal regulation by the mature osteocyte network (*Figure 29* B).

Enriched in the *Early activation* cluster were a number of biological processes associated with neuron-like axonogenesis and the patterning of cell contacts (Figure *28* C). Further, 'axon guidance' (mmu04360, p=2.22E-02) was the only significantly enriched KEGG pathway in this group, reflecting the upregulation of a range of paracrine and juxtacrine signalling molecules early in osteocyte differentiation (*Figure 29* B). The *Transition* cluster was significantly enriched for stress-fibre assembly processes, structures known to play an important role in mechano-transduction, as well as a number of pathways known to regulate neuron cell-cell contact dynamics, including axon guidance, Rap1 signalling and TGF-Beta signalling (Dodd & Jessell, 1988; Dufort, Paszek, & Weaver, 2011; Fu et al., 2007; Yi, Barnes, Hand, Polleux, & Ehlers, 2010). This indicated that processes and pathways associated with neuron-like axonogenesis, are upregulated early osteocyte differentiation, coinciding with network formation.

6.3 Discussion

Here we document an experimental and analytical approach to identify genes specifically enriched for expression in the osteocyte network *in-vivo* and with it define an osteocyte transcriptome signature. Supporting the efficacy of our approach was the identification of many established osteocyte markers known to play an important role in osteocyte biology. These include *Sost* and *Dkk1*, secreted inhibitors of the Wnt-signalling pathway critical to the osteocytic control of osteoblast function, and *Mepe* and *Dmp1*, key molecules in the regulation of skeletal mineral by the osteocyte network (Schaffler & Kennedy, 2012). Consistent with histological reports demonstrating the restriction of their expression the osteocyte network, these genes were among the most enriched for expression in osteocyte-isolated samples (Gowen et al., 2003; Poole, 2005; Toyosawa et al., 2001). In addition to osteocyte markers, the osteocyte signature was highly enriched for genes associated with skeletal biological process and skeletal phenotypes when

knocked out in mice. It is important to note that the detailed skeletal phenotyping of knockout mice conducted through the OBCD pipelines was performed in female animals, while the signature genes were identified for their expression in males. There is evidence to indicate sex can have a significant influence on skeletal phenotype (D. W. Rowe et al., 2018). Thus, generalisation of these results to imply these genes are phenotypically important in both sexes may require further characterisation of male knockout lines. However, supporting a persistent effect in both sexes was the concordant expression in between male and female mice of all significant genes, with no significant difference in expression observed at any age in our data. Overall, the phenotype identified in the OBCD-screened KO mice indicates that genes that are enriched for expression in the osteocyte network are important for normal skeletal function in at least one gender, potentially due to their role within the osteocyte network. For the majority of osteocyte transcriptome signature genes, their specific expression in the osteocyte network was the first evidence of skeletal involvement. Their enrichment alongside osteocyte markers and many genes known to influence skeletal biology suggests the osteocyte signature may identify skeletally novel genes with a role in the osteocyte network that is yet to be defined (discussed in detail in section 8.3).

A key distinction between this work and previous transcriptome investigations of the osteocyte network was the use of techniques enabling the detection and discovery of non-coding and novel transcripts. Understanding of the non-coding transcriptome in the skeleton is very limited and has largely been based on expression data from effector cell types differentiated *in-vitro* (Hassan, Tye, Stein, & Lian, 2015). Here we identify a significant number of lncRNAs enriched for expression in the osteocyte network, including novel non-coding genes not previously described in major transcriptome annotations. Nearly all of these novel signature genes were only expressed in osteocyte samples relative to 12 other tissue types and several were not detected *in-vitro*. Tissue-specific lncRNAs expression is known to be important to the function of other tissues, however their contribution to the osteocyte network is not well defined (G. Chen et al., 2013; Mercer, Dinger, Sunkin, Mehler, & Mattick, 2008). Detection in the osteocyte signature and lack of expression in other tissues indicates that novel lncRNAs may also contribute to the specialised function of the osteocyte network (discussed in detail in section 8.4).

Along with known and novel genes the signature identified a number of signalling pathways with the potential to influence osteocyte function. Pathways that control neuronal development and axon guidance stood out as programs co-opted by osteocytes to control osteocyte network formation and intercellular communication. Axon guidance is the process by which dendritic cells form specific patterns of cellular contacts (Stoeckli, 2018). First discovered in neurons, axon guidance is now recognised to play a critical role in the formation of cell networks in a range of tissues (Dodd & Jessell, 1988; Hinck, 2004). A number of axon guidance molecules have been shown to affect the skeleton individually including *Sema3a*, which has osteoprotective effects by supressing bone resorption and promoting bone formation, *Sema4d*, which is expressed in osteoclasts and supresses osteoblastic bone formation, as well as a number of Ephrin-Eph receptor pairs which couple bone formation and resorption (Hayashi et al., 2012; Negishi-Koga et al., 2011; C. Zhao et al., 2006). However, the effect of these genes as a function of their combined contribution to the process of axon guidance and osteocyte network formation is not well defined. Network formation is critical to osteocyte differentiation and function and yet the molecular mechanisms involved are poorly understood (Schaffler et al., 2014). The upregulation of axon guidance in early osteocyte differentiation reported here suggests the molecular pathways that coordinate synapse formation between neurons may also direct the formation of the osteocyte network (discussed in detail in section 8.5).

In summary, the osteocyte transcriptome signature orients the expression of hundreds of new genes and a number of signalling pathways within the osteocyte network. These encompass protein-coding, long non-coding, known and novel genes specifically enriched in their expression to the skeleton. While further experimentation is required to deduce their skeletal influence, detection alongside established osteocyte markers and molecular processes known to be important to skeletal function suggests they too may play a role in the function and formation of the osteocyte network. It is important to note that the osteocyte transcriptome signature was defined based on gene expression enrichment within the adult long bones, and that factors such as age, sex and mouse strain may play a role in the genes identified by this approach. Moreover, the transcriptome data presented here represents the average gene expression of both cortical and cancellous bone, as well as anatomical axes across the bone. As such, there may be multiple osteocyte transcriptome signatures expressed in distinct biological contexts. More detailed studies

utilising the experimental and analytical methodologies described here will enable a more complete understanding of the genes that are important for osteocyte function.

7 THE OSTEOCYTE NETWORK IN SKELETAL DISEASE

Genetics has a big influence on skeletal health, contributing to both rare and common genetic diseases (Bonafe et al., 2015; Kemp et al., 2017). In the past, our lack of insight into the genes important to the formation and function of the osteocyte network has limited our understanding of their contribution to such traits and conditions. Despite this, there are several examples of osteocyte restricted genes playing important roles in disease and serving as molecular targets for the development of skeletal therapeutics (Baron & Hesse, 2012). Given the critical role of osteocytes in maintaining skeletal health, we hypothesised that the genes identified in the osteocyte transcriptome signature may be of particular importance to skeletal disease.

In this chapter we investigate the relationship between the osteocyte signature and genes known to cause severe skeletal dysplasia or contribute to clinically significant skeletal traits. We examine the enrichment of disease associated genes orthologs in the signature as a whole, and how genes enriched in the osteocyte network may be associated with particular dysplasia types. We also test the enrichment of signature genes among those associated with human BMD variation by GWAS, distinguishing those that have been supported with functional data.

7.1 Methods

7.1.1 Osteocyte transcriptome signature genes and the association with human skeletal dysplasia

The enrichment of osteocyte signature genes among genes known to cause skeletal dysplasia was calculated as per section 2.2.27. Expression of dysplasia causing genes in other organs and tissues was determined as per section 2.2.13. The enrichment of osteogenesis imperfecta (OI) causal genes in the osteocyte-network relative to other cell lineages in the marrow space was determined as per section 2.2.20. The expression of OI-causal-genes in osteogenic differentiation was analysed as per section 2.2.21. OI-causal gene regulation during skeletal maturation was analysed as per section 2.2.16.

7.1.2 The contribution of the osteocyte network to human BMD variance

The enrichment of signature genes associated with human BMD variance was determined as per section 2.2.28. The previous association of genes with skeletal biological processes or knockout mouse phenotypes was performed as per section 2.2.12. BMD associated genes enriched for expression in osteocyte relative to other organs and tissues were identified as per section 2.2.12.

7.2 Results

7.2.1 Osteocyte signature is enriched for orthologs that cause human skeletal dysplasia

Since osteocytes act as master regulators of the skeleton, we hypothesised that orthologs to genes identified in the osteocyte signature would be associated with human skeletal disease. To examine this hypothesis, we first tested the enrichment of osteocyte signature genes among terms in the Disease Ontology (DO) database, a comprehensive catalogue that annotates human diseases with relevant gene sets. All 9 significantly enriched DO terms were related to diseases of the skeleton, which indicated the osteocyte signature is enriched for genes associated with skeletal disease above all other disease types contained in the DO database (Figure *30* A).

We then compared the osteocyte transcriptome signature genes to the list of causal genes in the nosology of skeletal genetic disorders (Bonafe et al., 2015). Fifty-four of the 277 skeletal dysplasia-causing genes actively expressed in the skeleton were present in the osteocyte signature, related to 93 skeletal diagnoses (Figure *30* B). To determine whether osteocyte signature genes were likely to cause particular types of dysplasia, we ranked dysplasia groups containing more than one causal gene by the proportion of disease-causing genes present in the osteocyte transcriptome signature (Figure *30* C). Many of the top ranked groups pertained to diseases of bone composition and mineralisation, namely *Osteogenesis imperfecta and decreased bone density* group (13/25 casual genes in the signature, P = 0.0001), *Neonatal osteosclerotic dysplasia* (3/4, P = 0.02), *Other sclerosing bone disorders* (5/10, P = 0.03) and the *Abnormal mineralization* group (4/7, P = 0.03). In contrast, only one signature gene of the 16 in the *Osteopetrosis and related disease* group, consistent with current understanding that these are primarily diseases of osteoclastic bone resorption. Twenty-two of the 24 genes known to cause

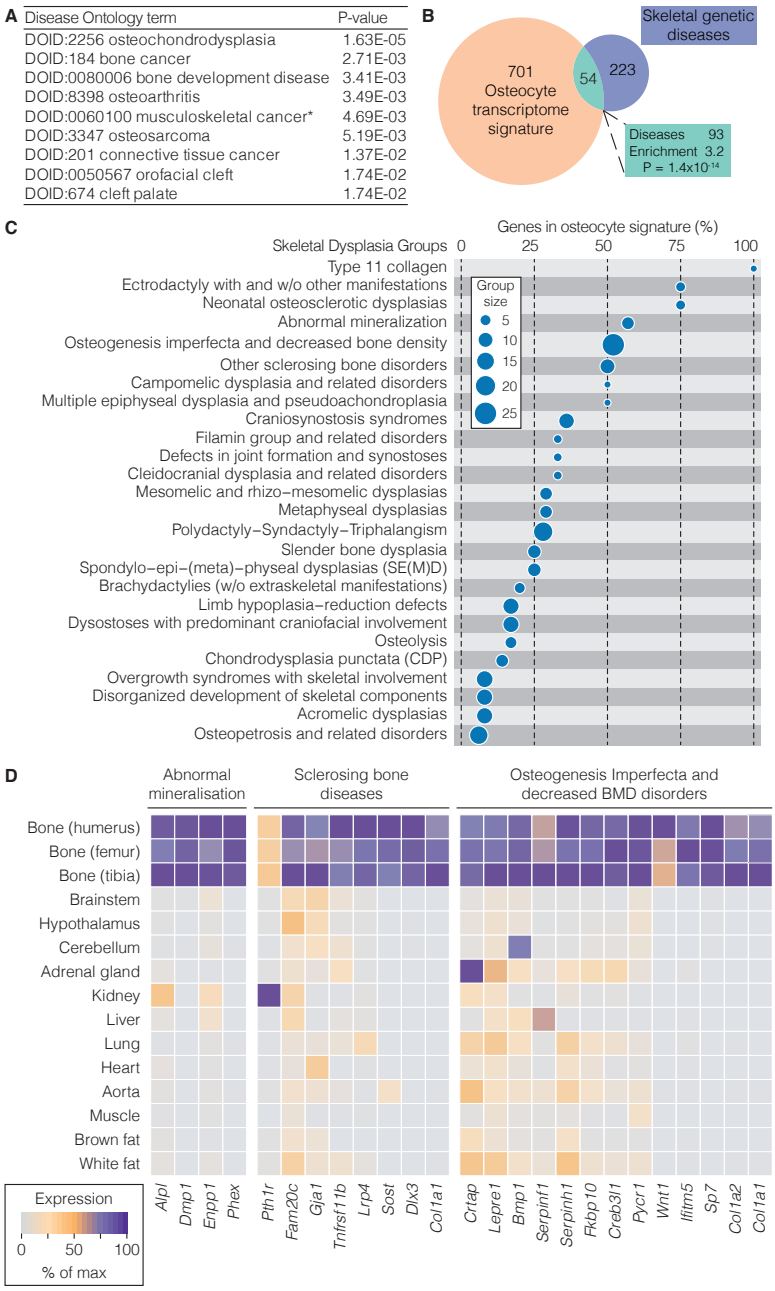


Figure 30 - Gene-orthologs known to cause skeletal dysplasia in humans are over-represented in the osteocyte transcriptome signature.

(A) Significantly enriched Disease Ontology (DO) terms in the osteocyte transcriptome signature (FDR ≤ 0.05, * truncated description). (B) The overlap between human orthologs to osteocyte transcriptome signature genes and the genes that cause rare skeletal genetic disorders. P-value and enrichment are calculated under the hypergeometric distribution. (C) Enrichment of skeletal dysplasia disease groups in the osteocyte transcriptome signature. Group size = number of genes in dysplasia group with mouse orthologs. Percentages reflect the number of osteocyte signature gene orthologs identified in the dysplasia group. (D) Skeletal dysplasia associated osteocyte transcriptome signature gene expression in 12 tissue types compared to osteocytes (% of maximum mean-FPKM).

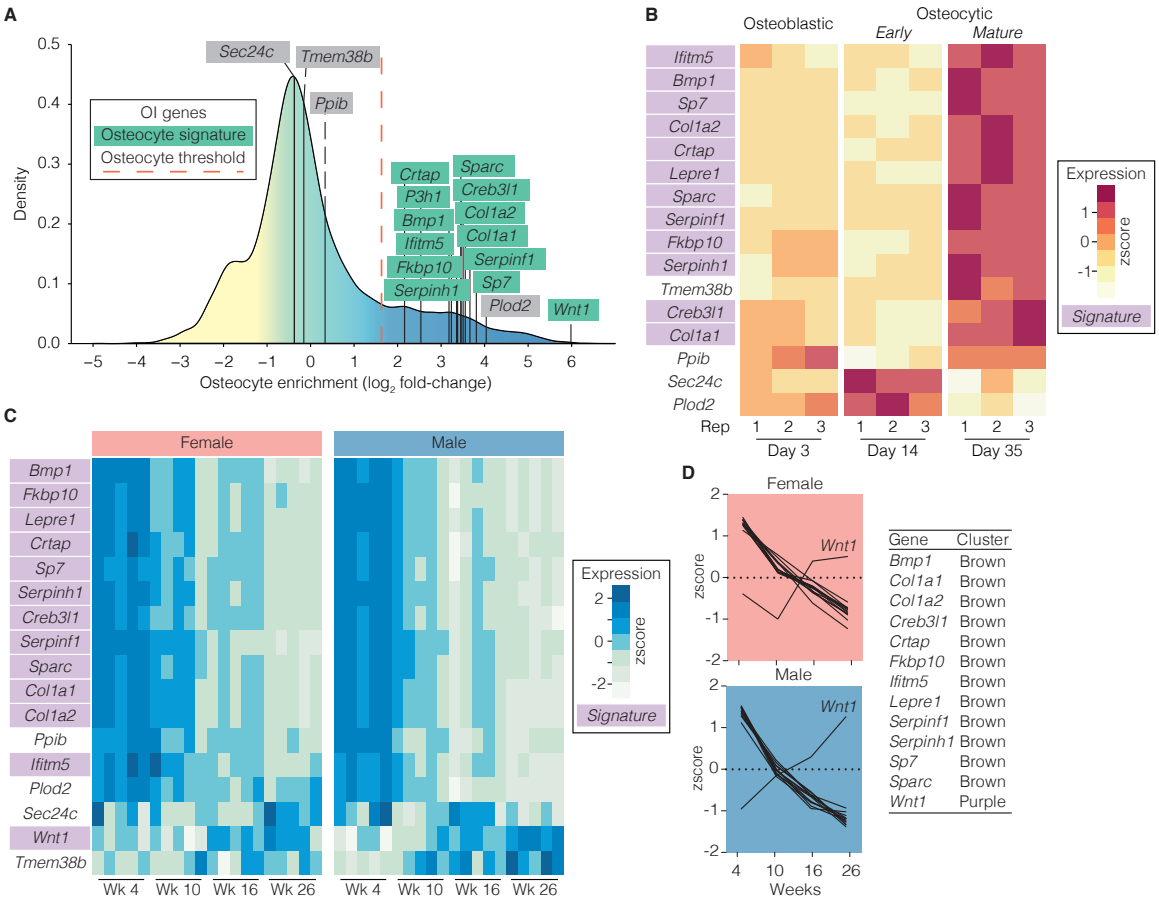


Figure 31 - Genes that cause Osteogenesis Imperfecta (OI) are enriched for expression in the osteocyte network.

(A) Enrichment of OI-causal genes in *in-situ* isolated osteocytes. Osteocyte signature genes are highlighted in green. (B) Expression of OI-causal genes during osteocytic differentiation (z-score of normalised counts). Signature genes are highlighted in purple. (C) OI-causal gene expression during skeletal maturation in both sexes. Signature genes are highlighted in purple. (D) Expression of OI-causal genes identified in the osteocyte transcriptome signature during skeletal maturation.

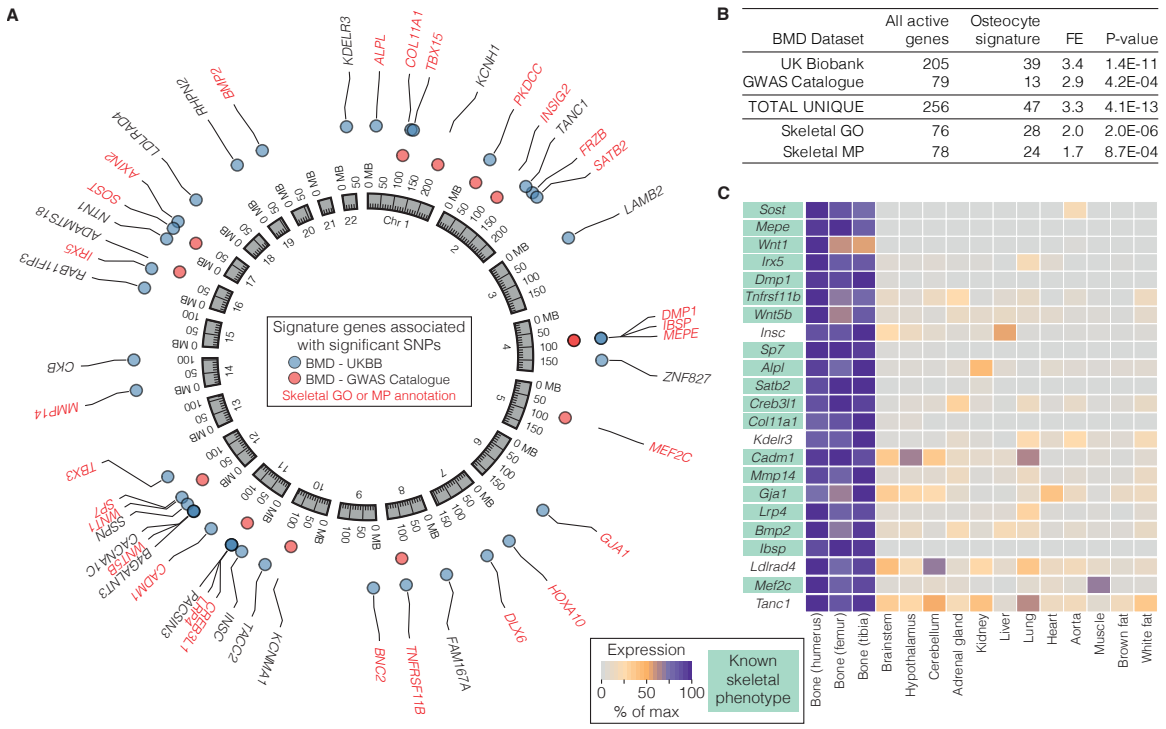
diseases affecting bone-composition and mineralisation were most highly expressed in osteocytes relative to 12 other tissue types (Figure *30* D). These data indicated that many genes known to cause diseases of abnormal bone composition are specifically enriched for expression in the osteocyte network, concurrent with its pivotal role in skeletal regulation.

A large fraction of the signature genes identified in the *Osteogenesis imperfecta and decreased bone density* group were known to cause osteogenesis imperfecta (OI)

specifically, suggesting the osteocyte network may be of particular relevance to this group of genetic disorders. In total, 13 of the 17 genes known to specifically cause OI were in the osteocyte signature with the remaining 4 genes actively expressed in osteocytes but either not enriched relative to cells in the marrow space (*Sec24c*, *Tmem38b* and *Ppib*) or highly expressed in earlier stages in the osteogenic lineage (*Plod2*) (Figure *31* A and B). Interestingly, 12 of the 13 casual osteocytes signature genes were most highly expressed in mature osteocytes relative to early osteocytes with the only exception, *Wnt1*, not detected in the *in-vitro* system (Figure *31* B). Examination of OI-causal gene expression in the osteocyte network during skeletal maturation revealed a pattern of down regulation in both sexes (Figure *31* C). All but one of the osteocyte signature genes known to cause OI were identified in the Brown skeletal maturation cluster, associated with protein production, modification and export (Figure *31* D and *Figure 17*). The only exception was again *Wnt1* which showed a converse pattern of regulation to all other OI-causal genes in the osteocyte signature, increasing in expression with skeletal maturation in both sexes. This suggested that the molecular regulation of *Wnt1* expression within the osteocyte network is distinct from other OI-causal genes. Interesting, OI caused by Wnt1 mutations is reported to be somewhat distinct to other disease subtypes in terms of bone characteristics and response to treatment (Joeng et al., 2017; Palomo et al., 2014). Together, these data indicate that genes enriched for expression in the osteocyte network may play an important role in OI, and that gene regulation within the osteocyte network during skeletal maturation may distinguish between OI subtypes.

7.2.2 The osteocyte network is enriched for genes associated with variance in human bone mineral density

We next examined the relationship between osteocyte signature genes and bone mineral density (BMD). BMD is a highly heritable, genetically complex skeletal trait and a key diagnostic factor for common skeletal diseases such as osteoporosis (Kanis, 2002). To do this, we identified human orthologs to signature genes that had been associated with variance in human BMD using two independent GWAS datasets. These were 1) the National Human Genome Research Institute (NHGRI)- European Bioinformatics Institute (EBI) GWAS catalogue which identified all the significant gene variant associations from 22 GWAS studies of BMD (98 unique genes with mouse orthologs in total), and 2) the most recent UK Biobank BMD GWAS, which identified 234 unique genes with mouse orthologs associated



BMD Dataset	All active genes	Osteocyte signature	FE	P-value
UK Biobank	205	39	3.4	1.4E-11
GWAS Catalogue	79	13	2.9	4.2E-04
TOTAL UNIQUE	256	47	3.3	4.1E-13
Skeletal GO	76	28	2.0	2.0E-06
Skeletal MP	78	24	1.7	8.7E-04

Figure 32 - Genes associated with human BMD variation are enriched for expression in the osteocyte network.

(A) Human orthologs to osteocyte signature genes significantly associated with BMD variance in human GWAS cohorts. (BLUE dots - UK BioBank, RED dots - GWAS catalogue). Genes written in red are annotated with a directly skeletally related GO biological process or skeletal phenotypes in MGI. (B) Osteocyte transcriptome signature GWAS enrichment statistics (P-value and FE = fold-enrichment relative to chance based on hypergeometric distribution). (C) BMD-associated osteocyte transcriptome signature genes expressed most highly in osteocyte relative to 12 tissue types (% of maximum mean-FPKM). Boxed genes have knockout mouse models with a skeletal phenotype.

with significant variants in the data of 142,487 individuals (Kemp et al., 2017; MacArthur et al., 2017). In total 256 of the 301 genes identified by these datasets were actively expressed in osteocytes, 47 of which were contained in the osteocyte signature (Figure *32* A). This represented a 3.27-fold enrichment (FE) for BMD associated genes in the osteocyte gene list than would be expected by chance (p = 4.08E-13) (Figure *32* B). A critical step in GWAS is establishing the plausibility of gene-variant association, usually through the support of biological or phenotypic evidence that the associated gene affects the organ under investigation. Approximately one third (108/301) of all genes associated with BMD identified in these datasets were supported with skeletal biological processes or mouse phenotypes. Of the BMD-associated genes in the osteocyte signature, more than half (28/47) were

annotated with skeletal GO ontology terms or skeletal phenotypes (FE = 1.53, p = 0.001). This indicated that the osteocyte signature was not only enriched for genes associated with BMD, but specifically those that are supported by experimental evidence.

Of the 47 osteocyte transcriptome signature genes associated with BMD, 23 were most highly expressed in osteocytes relative to 12 other tissues sampled (Figure *32* C). Nineteen of these 23 had been demonstrated to play a role in skeletal biological processes or shown to produce a skeletal phenotype when knocked out in mice, more than double what would be anticipated by chance (P = 1.02E-5). Of the remaining 4 genes not previously associated with skeletal biological processes or phenotypes in the MGI, one had been screened through the OBCD skeletal phenotyping pipeline, the low-density lipoprotein receptor class A domain-containing protein *Ldlrad4* (*Figure 23*). Digital x-ray microradiography demonstrated that *Ldlrad4*$^{-/-}$ mice had significantly reduced bone mineral content in the femora compared to *Ldlrad4*$^{+/+}$ mice, with no alteration to bone shape. Vertebral BMC was also seen to be in the lower normal range, with a significant reduction in total bone volume owing to less trabecular bone. This analysis suggested *Ldlrad4* may influence BMD through its function within the osteocyte network and demonstrated how the osteocyte signature may be a useful tool to identify plausible genes candidates in GWAS datasets.

7.3 Discussion

Osteocytes are critical regulators of the skeleton, yet our limited understanding of the genes important to their specialised function has meant that their contribution to skeletal genetic diseases is not well understood. Here we show that the osteocyte signature is highly enriched for genes known to cause severe skeletal dysplasia and contribute to the variance of bone mineral density, indicating that genes specifically enriched in the osteocyte network make a significant contribution to skeletal disease.

While the role of osteocytes in skeletal genetic disease is poorly defined, the identification of disease-causing mutations in genes subsequently localised to the osteocyte network has been instrumental to our understanding of its specialised skeletal function. For example, mutations in SOST that cause sclerosteosis established secreted inhibitors of the Wnt-signalling pathway as a critical molecules used by the osteocyte

network to regulate bone mass (Balemans et al., 2001). Similarly, the identification of DMP1 mutations causing autosomal recessive hypophosphatemic rickets identified a critical role for osteocytes in mineral metabolism (Feng et al., 2006). Here we show that genes with enriched expression in osteocytes are associated with a range of skeletal dysplasias yet this enrichment was not uniform across disease types. Our analysis highlighted the significant enrichment of genes known to cause diseases of bone composition and mineralisation in the osteocyte signature, including more than 80% of genes known to cause OI. OI is caused by defective collagen production, with bone quality compromised by mutations affecting its expression, synthesis and secretion (Forlino & Marini, 2016). As a consequence, OI is typically thought to be a disease of the osteoblasts due to their critical role in bone tissue formation. Our data demonstrates that the genes that cause OI are enriched expression in the osteocyte network, upregulated in late stages of osteocytic differentiation *in-vitro* and dynamically regulated during skeletal growth. This may suggest that the osteocyte network may play a direct role in the synthesis of bone tissue during skeletal maturation, in addition to its regulation of effector cell types. Conversely, gene expression in the osteocyte network is affected by its extracellular environment and thus matrix changes caused by OI could disrupt osteocyte network function. Abnormalities in the osteocyte network accompany a number of OI subtypes however, whether this contributes to disease or is a consequence of abnormal bone formation by osteoblasts has not been addressed (Blouin et al., 2017; Grafe et al., 2014). While decades of research support the integral role osteoblasts play in OI, this work indicates the osteocyte network may too make a critical contribution (discussed in detail in section 8.6).

Bone mineral density (BMD) is a genetically complex skeletal trait used in the diagnosis of common skeletal diseases such as osteoporosis (Feldstein, Elmer, Orwoll, Herson, & Hillier, 2003). Much of what is known about the genetic architecture that influences BMD has come from genome wide association studies (GWAS), detecting DNA variants that cooccur with changes in bone density at the heel, hip or spine (Rivadeneira & Mäkitie, 2016). While a powerful approach to unbiasedly identify *loci* associated with skeletal composition, the prioritisation of genes whose specific function may be contributing to BMD for experimental validation remains a considerable challenge. In other traits, gene expression in trait-related tissues has proven a powerful approach to addressing this challenge (Finucane et al., 2018). Our analysis suggests that the osteocyte

signature is not only enriched for genes associated with variance in human BMD as whole, but also specifically those supported by experimental evidence. Consistent with their established role regulating bone mineral, this suggests genes specifically enriched in the osteocyte network may have a significant influence on BMD.

Together data suggest that the genes, processes and pathways important to the formation and function of the osteocyte network may be of particular relevance to skeletal genetic disease. As a result, the osteocyte transcriptome signature may be a useful resource for the identification of genes that contribute to skeletal pathophysiology, including both rare diseases and complex skeletal traits.

8 DISCUSSION

Here we document an experimental and analytical strategy to determine the genes actively expressed in the osteocyte network *in-vivo*. With this approach we examined the osteocyte transcriptome in multiple bone types, revealing a highly consistent pattern of gene activity throughout the skeleton, with the exception of a handful of developmental transcription factors. We demonstrated the dynamic transcriptome regulation in the osteocyte network during skeletal maturation, defining clusters of co-regulated genes influenced by age and distinct expression patterns between sexes. An osteocyte transcriptome signature is defined, consisting of genes significantly enriched for expression in the osteocyte network *in-vivo*, including novel lncRNAs with skeletally restricted patterns of expression. This also identified molecular processes enriched for expression in the skeleton, including neuron-like network control pathways dynamically regulated during osteocytic differentiation. Lastly, we demonstrated the significant over representation of genes associated with skeletal genetic diseases within the osteocyte signature, including genes that cause rare skeletal dysplasia and those associated with BMD variance. Together these discoveries contextualise gene expression within cell type, skeletal site, sex and maturation to advance our understanding of the molecular processes contributing to the highly specialised function of the osteocyte network.

8.1 Site-specific gene expression in the osteocyte network

The transcriptome analysis of multiple bone types revealed that the vast majority of genes expressed in osteocytes are active throughout the skeleton. Gene expression was indistinguishable between osteocytes from the left or right humeri, indicating that developmental molecular processes that establish parasagittal mirror-symmetry are not maintained in these cells into adulthood. Direct comparison between bone types however revealed the differential activity of a small number of transcription factors, the majority of which were homeobox genes and Hox-associated lncRNAs (*Figure 33*). Hox-genes are arranged in linear clusters throughout the genome and are numbered based on their temporal expression patterns during development. Expression is also spatially restricted, regulated along dorsal-ventral and anterior-posterior axes in the torso and proximal-distally in the limbs. The tight temporal and spatial control of Hox-gene expression is critical to the developmental fidelity of many organ systems. Our analysis resolved patterns of expression

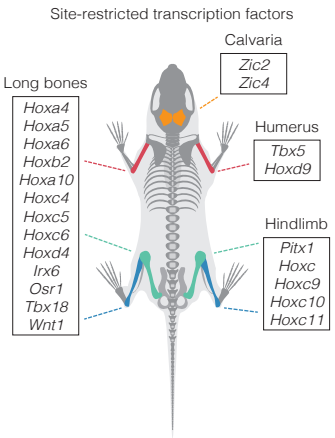


Figure 33 - Site-restricted transcription factor activity in the osteocyte network.

Pattern of site-specific transcription factor expression. Hindlimb = active in tibia and femur, Long bones = active in tibia, femur and humerus.

in adult osteocytes reflecting their developmental localisation, with several genes from the HoxB and HoxC clusters differentially expressed between fore and hindlimbs (Nelson et al., 1996). Interestingly, the specific deletion of HoxB and HoxC clusters has no observable effect on limb identity or development, despite limb restricted patterns in development (Medina-martı & Bradley, 2000; Suemori & Noguchi, 2000). This is in contrast to deletion of HoxA and HoxD genes which produce limb truncations corresponding to their precise site of expression (Kmita et al., 2005). This differential expression between adult bones and dispensable role in development suggests the restricted expression of Hox genes may play a role in the osteocyte network and serve a function beyond skeletal patterning.

Although very little is known about the role of Hox-expression in adult osteocytes, there have been elegant studies examining the role of Hox-gene expression in other osteogenic cell types. Leutch and colleagues showed that the local specificity of Hox-expression impacts the rate and fidelity of bone repair after trauma, a process that recapitulates many of the processes involved in healthy bone remodelling (Leucht et al., 2008). They examined this by transplanting bone progenitors taken from the jaw into defects in the tibia and vice versa, measuring the rate and quality of healing. This design compared *Hoxa11* expressing progenitors from the tibia with Hox-negative progenitor from the jaw, which, along with the calvarium, are derived from neural crest precursors (concordantly, our data indicate Hox-genes are not expressed in osteocytes from the calvarium, with the exception of a low level of *Hoxb4*). The defects in the jaw recovered poorly with progenitors from the tibia, which continued to express *Hoxa11* in contrast to their Hox-negative neighbours. Conversely, the tibial defects recovered normally with the jaw progenitors which began expressing *Hoxa11* in line with the surrounding cells. This role in osteoprogenitors was again examined by Rux et al, who showed that *Hoxa11* deletion differentially affected fracture healing, depending on the proximal distal location of the recovering limb bone (Rux et al., 2016). In contrast to Leutch

et al and in line with our data, Rux et al did not detect *Hoxa11* expression in osteocytes, with immunohistochemistry indicating it is restricted and upregulated in progenitor cells in the marrow space and lining the periosteal bone surface. While this demonstrates that site-restricted Hox-gene activity in skeletal cells have important consequences to site-specific skeletal biology, it also indicates that Hox-gene expression in the osteocyte-network is distinct from that in progenitor populations. Thus, it is possible that the function of Hox-gene expression in the osteocyte network may too be distinct from other osteogenic cell populations.

An important observation in the report from Leucht et al was that Hox-negative cells were able to acquire the Hox-status of their neighbours, with jaw progenitors assuming the *HoxA11* expression in concordance with the cells surrounding tibial defects (Leucht et al., 2008). This suggests the site-specific pattern of Hox-expression in skeletal cells is able to be communicated to cells arriving to bone. Unlike the skeletal progenitors, osteocytes are very long-lived cells, immobilised in the hard-mineral tissue of bone. While speculative, the possibility that stationary osteocytes could communicate site restricted Hox-expression patterns as a mechanism for cell localisation has not escaped our attention. What is clear from our analysis is that the osteocyte transcriptome in the appendicular skeleton contains molecular differences corresponding to their skeletal site of residence. Whether this plays a role in osteocyte skeletal regulation or a broader role in specifying the location of cells within the skeleton and downstream cellular control pathways is to be determined.

8.2 Age and sex specific regulation of the genes associated with perilacunar-remodelling during skeletal maturation

The amount of bone accrued during skeletal maturation is a powerful predictor of skeletal disease in advancing age and can differ between the sexes (Alswat, 2017; Wei & Mao, 2007). While the skeletal regulatory role of the osteocyte network in the adult skeleton is well established, our understanding of its contribution to this period of post-natal skeletal development is far less developed. Here we demonstrate dynamic patterns of gene expression in the osteocyte-network during skeletal maturation and identify a number of co-regulated gene clusters whose regulation is strongly associated with age and sex.

Gene co-expression analysis identified a cluster of genes upregulated in both sexes before puberty and also specifically enriched in females at skeletal maturity. These genes were enriched for processes associated with extracellular-acid-secretion and bone resorption. The osteocyte network has been shown to acidify its extracellular surrounds to liberate bone mineral and directly remodel skeletal tissue surrounding its lacunar-canalicular system in the process of perilacunar-remodelling (Jähn et al., 2017). At the molecular level, much of what is known about perilacunar-remodelling comes from its role in lactation, where genes involved are upregulated within the osteocyte network during milk production and then downregulated post-lactation (Qing et al., 2012). Several genes that play a role in this process were among those most highly correlated with sex -specific expression patterns in our data, including *Acp5*, *Ctsk* and several components of extracellular proton transporters. Moreover, the cluster as a whole was specifically upregulated during lactation consistent with their involvement in perilacunar-remodelling. Collectively, these data indicate that the genes that regulate perilacunar-remodelling during skeletal maturation may be dependent on both age and sex.

While the contribution of perilacunar-remodelling to normal skeletal homeostasis is still being elucidated, it is becoming clear that it is an important process for maintaining bone quality and mass and preventing fragility. TGF-Beta signalling has been identified as a regulatory pathway controlling perilacunar remodelling, with pharmacologic inhibition and osteocyte-specific knockout of TGF-Beta receptors impairing the process and affecting bone quality (Dole et al., 2017). Our data indicate that the TGF-Beta pathway may also be involved in the sex specific activity during skeletal maturation, with the gene encoding the TGF-Beta type 1 receptor *Tgfbr1* among the most strongly correlated with magenta module expression. The suppression of perilacunar-remodelling by glucocorticoids has also been associated with bone degeneration, causing deterioration of the osteocyte lacuna-canalicular network, collagen disorganization, and matrix hyper-mineralization (Fowler et al., 2017). Long-term glucocorticoid use is involved in the management of a range of diseases including nephrotic syndrome, systemic lupus erythematosus, acute lymphoblastic leukemia, and asthma (Kaste et al., 2001; Olenec & Gern, 2009; Rousseau-Nepton, Lang, & Rodd, 2013). Children are thought to be particularly susceptible to the skeletal side effects of these treatments, reducing bone accrual and increasing fragility and fracture risk (Sarinho & Melo, 2017). This data demonstrates perilacunar-remodelling is

upregulated in the osteocyte network of both sexes before puberty, suggesting this process may be of particular importance in childhood.

This data also indicated that perilacunar-remodelling is differentially regulated between sexes at skeletal maturity. While this may indicate a higher rate of direct bone-remodelling by the osteocyte network in female mice, this was not accompanied by a significant difference in BMD, BMC or bone length at mature timepoints. There are a range of factors that could explain this observation. From a technical standpoint, the failure to detect a difference may be due to the technical limitations of DXA to detect fine-scale changes in osteocyte lacunar size, and thus a more detailed examination by μCT or using specialised techniques to measure microstructural changes in lacuna-canalicular architecture may be required (N. E. Lane et al., 2006). Alternatively, if there is no physical evidence of a difference in perilacunar-remodelling, it could be that critical genes involved in pathway activity are not sex -specifically regulated. The genes encoding carbonic anhydrase *Car2* and the matrix metallopeptidase *Mmp13*, thought to be involved in acid production and the breakdown of bone-matrix during perilacunar-remodelling, did not display sex specific expression patterns in our data (Kogawa et al., 2013; Tang et al., 2012). While another matrix metallopeptidase, *Mmp9*, was among the genes most highly correlated with sex specific expression patterns and may perform a similar role as *Mmp13*, no other carbonic anhydrase was identified in place of *Car2*. This may suggest *Car2* expression is critical to acidification of the lacuna-canalicular system and a pivotal molecule in perilacunar-remodelling. Interestingly, Dole and colleagues did not detect a difference in morphological bone phenotype in response to TGF-Beta inhibition, however they did demonstrate microstructural changes in the osteocyte-network and the lacunar environment associated with abnormal perilacunar-remodelling and decreased bone strength (Dole et al., 2017).

Sexual dimorphism in bone strength is well established but typically attributed to the difference in bone geometry between sexes, resulting from differences in sex-hormone and growth-hormone action (Kim et al., 2003). With this in mind, a detailed assessment of the osteocyte-network and lacuna-canalicular system between sexes at skeletal maturity may be required to elucidate the extent to which sexually dimorphic regulation of perilacunar-remodelling contributes to the sex differences in skeleton.

8.3 Skeletally “unknown” genes in the osteocyte transcriptome signature

Examining the transcriptome of the osteocyte network in multiple bone types identified more than 14000 genes actively expressed in osteocytes across the skeleton. While the pattern of expression in osteocytes was distinct from other organs, our analysis of a publicly available dataset suggests that more than >60% of these genes were common to a range of tissues, indicating that many were ‘house-keeping’ genes required for fundamental cellular processes not specific to osteocytes (R. Zhang et al., 2014). Previous transcriptome studies of *in-situ* isolated osteocytes have utilised data from potential contaminating tissues and *in-silico* approaches to control for the contribution of remnant extra-osseous cell types (Ayturk et al., 2013; Kelly et al., 2016). These effectively identified and removed transcripts specifically contributed by these other tissues, but these studies stopped short of defining genes preferentially expressed in the osteocyte network potentially contributing to the specialised biology of these cells. Addressing this, we developed an experimental analytical subtractive-sequencing strategy which identified genes significantly enriched in osteocytes relative to cell lineages in the bone marrow. Skeletal biological processes and known osteocyte markers, shown to be highly specifically expressed in mineral-bound skeletal cells by orthogonal techniques, were among the most enriched genes by this approach. These findings confirmed the capacity for this methodology to distinguish genes enriched in osteocytes. While used here to define a subset of genes enriched in osteocytes, similar experimental designs may be effective in a range of hard to isolate cell populations and those defined by their extracellular context.

The next step in the identification of genes contributing to the specific biology of the osteocyte was the examination of gene expression in an *in-vitro* model of osteocytic differentiation and network formation (St. John et al., 2014). Using publicly available data we identified not only genes upregulated by osteogenic differentiation, but also a significant number of genes absent *in-vitro* that were expressed in all 24 *in-vivo* sequencing samples and enriched in osteocytes relative to marrow lineages and a range of other organs. These "*in-vivo* only" genes demonstrate the interconnection between the extracellular environment the osteocyte transcriptome and thus underscore the importance of conserving this setting in the molecular investigation of this cell type. A strength of *in-vitro* data is the fact that gene expression measurements in this model are unlikely to be affected

by the contribution of other cell lineages or remnant contaminant populations such as blood vessels and nerves. This combination of data sources is a key difference between our approach and that of previous transcriptome investigations. By identifying genes not expressed *in-vitro* yet specifically enriched in the skeleton compared to other tissues *in-vivo*, we leverage the strengths of both model systems in terms of sensitivity and specificity of osteocyte gene expression. Through these analyses we defined the osteocyte transcriptome signature, genes active throughout the skeleton that are preferentially expressed in osteocytes relative to distinct cell lineages in the marrow space and other osteogenic stages.

Early transcriptome investigations of the osteocyte network, such as that by Paic *et al* 2009 or that of Ayturk *et al* 2013, used distinct approaches to collecting samples and controlling for the contribution of contaminating cell types expression data (Ayturk et al., 2013; Paic et al., 2009). Paic *et al* compared the transcriptome of osteoblasts and osteocytes that were disaggregated from the calvaria of neo-natal mice and then FACS sorted using fluorescent markers to isolate the two cell populations. This analysis was performed by microarray, and only a subset of the genes that were differentially expressed between the between osteocytes from osteoblasts were presented. Among those, approximately one quarter (54/194) were identified in the osteocyte signature, with some key markers of osteocyte function absent from the data of Paic *et al*. The limited overlap between these datasets could be due to biological differences in the samples used, such as the age of the mice (neonatal vs adult) or the bone type (skull vs long bones). These differences could also be due to technical factors, caused by the process of collagenase digestion and FACS sorting relative to *in situ* isolation. Moreover, without the raw data for analysis using current methodologies, it is not known whether more genes from the signature were in fact represented among the osteocyte enriched differentially expressed genes that were not reported in the publication. Ayturk *et al* showed that collagenase digestion protocols, like those used by Paic *et al*, induced large artefactual changes in the transcriptome of processed samples. As an alternative, Ayturk and colleagues conducted their investigations using methodologies very similar to those reported here to prepare *in situ* isolated bone samples. They used a different sequencing and bioinformatic strategy from the total-RNA reported here, performing poly-A enriched “mRNA” sequencing. While such a strategy increases data coverage for known genes, it limits the capacity to detect

novel transcripts. Moreover, as with the publication by Paic *et al*, the raw data was not released by Ayturk *et al*,with the publication. As a consequence, a systematic comparison of the data reported in this thesis with these early investigations was not possible.

Analysis of the association between osteocyte signature genes and skeletal significant processes and phenotypes revealed that the vast majority of genes enriched in osteocytes are have no previous annotation with bone biology. These included genes that when knocked out in mice cause significant morphological and functional skeletal consequences, identified by the OBCD phenotyping pipeline. Two of these are known to influence Wnt-signalling, a critical pathway in osteocyte skeletal regulation, namely the Dishevelled associated activator of morphogenesis (*Daam2*) and Dishevelled binding antagonist of Beta-catenin (*Dact3*). Daam2 enhances Beta-catenin signalling by modulating the formation of Wnt-receptor complexes in the nervous system, heart and gut (Ajima et al., 2015; Lee & Deneen, 2012). Dact3 on the other hand is a Wnt-signalling inhibitor, associated with pathway suppression in several cancer types. *Dact3* showed developmentally regulated expression in the dental mesenchyme, a tissue which shares many common molecular and material characteristics with bone (Xia Jiang et al., 2008). Consistent with these opposing influences was a divergence in bone mineral content and bone strength. The demonstration of a significant skeletal phenotype for multiple signature genes not previously associated with bone biology suggests the osteocyte signature may be a useful tool to identify new molecules associated capable of effecting the skeleton.

An important consideration in this analysis is that some skeletal gene associations are not captured in either the MGI phenotype and GO biological process databases. Manual examination of the top twenty signature genes enriched by osteocyte isolation revealed additional genes with published skeletal functions. Among those is *Notum*, the protein product of which is a secreted inhibitor of the Wnt-signalling pathway, a molecular function similar to that of the osteocyte markers *Sost* and *Dkk1* (Kakugawa et al., 2015; Vogel et al., 2016). Like these other molecules, inhibition of Notum stimulates bone formation and thus Notum targeting small molecules have been developed with an eye for skeletal-therapeutic application (Q. Han et al., 2016). Another example is the acid sensing ion channel subunit *Asic3*, used by cells to sense decreases in extracellular pH (S. H. Lin et al., 2016). *Acic3* expression is sensitive to mechanical stimulation in lamprey and mice sensory cells,

functioning as both a mechanoreceptor and chemoreceptor in these cells (Jalalvand, Robertson, Wallén, & Grillner, 2016; S. H. Lin et al., 2016). *Asic3* is expressed in cells in bone and the synovium, thought to play an important role in bone pain caused by arthritis and cancers growing in bone (Hiasa et al., 2017; Jahr, Van Driel, Van Osch, Weinans, & Van Leeuwen, 2005). Acid secretion by the osteocyte network is important in perilacunar-remodelling, however the genes and molecules by which this occurs are largely unknown (Jähn et al., 2017). The detection of *Asic3* among the genes most highly enriched in osteocytes suggests it may be an important molecule to sense mechanical stimuli and respond to differences in extracellular pH within the lacunae-canalicular system. These skeletal associations that are unannotated in the gene ontology database indicate the number of signature genes with known roles in skeletal biology may be higher than estimated. This adds weight to the idea that osteocyte signature genes are of particular importance to skeletal biology, supporting the notion that those genes that are unknown to bone may too play important roles in the skeleton.

Other bone-unknown signature genes are ascribed with functions in other tissues that suggest their activity in osteocytes could have important skeletal consequences. The most highly enriched signature gene without a skeletal association was the Neuronal guanine nucleotide exchange factor gene (*Ngef*, also known as Ephexin), regulated downstream of Ephrin activated EphA receptors and important to cell contact dynamics in neurons (Shamah et al., 2001). A number of Ephrins and Ephrin-receptors were also identified in the osteocyte signature, several known to be indispensable to normal bone formation and remodelling enabling contact driven communication between bone cells (Matsuo & Otaki, 2012). The highly enriched expression of *Ngef in-vivo* implicates the importance of Ephrin signalling control in the osteocyte network. The Atypical chemokine receptor *Ackr3* was also highly enriched in osteocytes. Ackr3 is a sequestering receptor for Cxcl12, an chemokine with chemotactic effects on marrow cells known to play a role in the recruitment of tumour cells and osteoclast-precursors to bone (Benredjem, Girard, Rhainds, St-Onge, & Heveker, 2017; Y. X. Sun et al., 2005; Wright et al., 2005). Ablation of osteocytes has been shown to affect Cxcl12 levels in the marrow space, thought to be important in the regulation of hematopoietic stem and progenitor cells in the marrow space (Asada et al., 2013). Importantly, *Ackr3*, formerly known as *Cxcr7*, has previously been shown to be highly specifically expressed in osteocytes by LacZ reporter vector expression (Gerrits et al.,

2008). Despite this expression pattern, this study of *Ackr3*$^{-/-}$ knockout mouse did not report any significant differences in skeletal phenotype from wild-type in response to known provocation models, perhaps indicating the effects of its activity in the skeleton are realised in extra-skeletal tissues. Osteocytes are known to influence marrow cell dynamics and marrow composition, although this was not examined in the study of *Ackr3*$^{-/-}$ knockout mice (Asada et al., 2013). While this data independently supports the specificity of osteocyte signature gene expression, whether osteocyte derived Ackr3 influences cells in the marrow compartment, potentially via its interaction with Cxcl12, is a hypothesis that is yet to be tested. Fibulin 7 (*Fbln7*, also known as *TM14*) was another bone-unknown gene that was more than 50-fold enriched in osteocytes. Members of the Fibulin family of extracellular matrix proteins are involved in elastic fibre formation, important to the development and function of connective tissues (De Vega, Iwamoto, & Yamada, 2009). *Fbln7* expression has been shown to increase with mineralisation in odontoblasts, dendritic cells that form extracellular matrix components of teeth (Arany, Koyota, & Sugiyama, 2009). Recently, *Fbln7* expression has also been shown to contribute to renal calcification, suggesting it as a local mediator of calcium deposition (Tsunezumi et al., 2018). Given the relevance of such processes to the maintenance of the skeleton, it may follow that Fibulin-7 activity also contributes to osteocytic bone mineral regulation.

Additionally, the biological function is poorly defined for many of the signature genes among those most enriched for expression in the osteocyte network. The sushi-domain containing gene *Susd5* has been associated with hyaluronic acid binding, a major constituent of cartilage, however this annotation was bioinformatically inferred based on domain similarity and is yet to be verified experimentally. The heat shock protein *Hspb7* has been associated with genetic heart disease, however how mutations in this gene may influence disease is yet to be determined. Investigations of public databases show that expression of the proline rich gene *Prr15* is highly tissue restricted and influenced by a number of disease states, however there are yet to be any in depth studies as to how this gene may contribute to cellular function. Lastly, the novel gene *Obcdi053500*, discovered here through *de novo* assembly of the osteocyte transcriptome data, has not been annotated in the major transcriptome annotations or seen to overlap known elements in nucleotide search engines (BLAST etc). While the lack of literature surrounding these genes make it hard to speculate as to how their expression may influence osteocytes, their

detection among the osteocyte signature genes most enriched in bone tissue suggests their function may be important to the specialised biology of this cellular network.

Here we have discussed the plausibility of potential skeletal regulatory roles of a number of genes among the most enriched for expression within skeletal tissue. In addition to these, the osteocyte transcriptome signature contextualises the expression of many hundreds of genes in osteocyte network, increasing by an order of magnitude the number of genes known to be specifically expressed in this skeletal regulatory system. For the vast majority of these genes this is the first evidence of a potential skeletal involvement. An important next step in this work is the validation of osteocyte restricted expression patterns using orthogonal techniques such as immunohistochemistry for signature identified protein-products or fluorescent *in-situ* hybridisation to localised transcript expression. Further investigation is required to elucidate their influence on bone however, the significant enrichment of skeletal biological processes, knockout mouse phenotypes, disease ontology terms and genetic diseases in the osteocyte transcriptome signature suggests these skeletally novel genes may too play important roles regulating the skeleton.

8.4 Non-coding and novel gene expression in the osteocyte network

A key distinction in our experiments to previous osteocyte transcriptome investigations was the use of data generation and bioinformatic methodologies allowing the detection of expression from long non-coding and novel *loci* within the genome. Using this approach ~850 lncRNAs were identified as actively expressed in osteocytes in each bone type. To our knowledge, this is the first investigation into the expression of lncRNAs in primary osteocytes, and as such, the contribution of lncRNAs to osteocyte biology is virtually unknown. A number of lncRNAs with functional roles in other bone cells were identified among osteocyte transcriptome signature genes. Expression of the lncRNA *H19* has been shown to increase with osteogenic MSC differentiation, consistent with its identification in the osteocyte signature (Liang et al., 2016). Retroviral overexpression of this lncRNA accelerated osteoblast differentiation and *in-vivo* bone formation, thought to be due to its effects on Wnt pathway activity. Another signature lncRNA, Maternally expressed gene 3 (*Meg3*), has also been shown to promote osteogenic differentiation in MSCs potentially due to upregulation of Bmp4 expression (Zhuang et al., 2015). Studies in other tissues indicate Meg3 influences TGF-Beta activity, an important regulator of

osteocyte formation and function (Mondal et al., 2015). Given the bone stimulatory activity of these lncRNAs in closely related skeletal cells, examination as to whether they play a similar role in osteocytic regulation of the skeleton may be warranted. In addition to these examples, the signature identified 32 lncRNAs enriched for expression in the osteocyte network. This expression context is the first evidence of potential skeletal effect for many of these genes. As such, these osteocyte signature identified lncRNAs may serve as a solid foundation for future investigation of non-coding regulatory architecture influencing the skeleton.

An important limitation of the transcriptome approaches used in this study is the inability to detect transcripts less than 200bp in length and thus small RNA biotypes, including microRNAs. microRNAs are potent biological molecules in a range of tissues, including other skeletal cell types (Lian et al., 2012). In order to examine their expression, specialised data generation techniques are required optimised for small RNAs. To date, no such methodologies have been performed on primary osteocytes and thus future studies to understand this influence of this layer of regulatory architecture in the osteocyte may be very insightful.

De-novo transcriptome assembly prior to gene quantification enabled the detection of novel genes actively expressed in osteocytes. We hypothesised that genes restricted in expression to osteocytes may remain to be discovered as total-RNA transcriptome data from these cells had not been included in generation of major transcriptome annotations. More than 100 novel genes were found to be actively expressed in osteocytes, including a number identified in the osteocyte transcriptome signature. These novel signature genes were restricted in their expression in other bodily tissues and some selectively expressed only in mature osteocytes, supporting our initial hypothesis. Moreover, the absence from *in-vitro* systems may suggest their role is somewhat governed by the osteocyte cellular environment. In other major tissue systems, including the cardiovascular system, the brain and nervous system and endocrine systems, the analysis of restricted non-coding genes has revealed important roles in tissue-specific biology and identified roles in disease (curated in the LncRNA Disease database, http://www.cuilab.cn/lncrnadisease) (Batista & Chang, 2013; G. Chen et al., 2013). This discovery of osteocyte restricted novel RNAs indicates this may also be true of the skeleton. Given the highly restricted patterns of

expression of these genes, they may also have utility in the development of novel models for investigation of the osteocyte network. Together, these discoveries provide the framework for defining the role of non-coding RNAs, including novel genes, in controlling osteocyte function, the biology of the skeleton and in skeletal disease.

8.5 Neuron-like network control pathways in the osteocyte network

Along with skeletally unknown and unannotated novel genes, the osteocyte signature identified new signalling pathways contributing to the formation and function of these critical skeletal cells. Clusters of significant biological processes, cellular components and biochemical pathways showed a profound enrichment of molecular machinery involved in intercellular communication and interactions with their extracellular environment. A key step in understanding how and when these pathways may contribute to osteocyte biology was the identification of co-expressed clusters in publicly available data from temporal model of osteogenic differentiation. St John and colleagues reported extensive changes in the transcriptome of IDG-SW3 cells as they mature into an interconnected bone-forming network *in-vitro*, mediated by the reshuffling of epigenetic architecture (St. John et al., 2014). These changes indicate a distinction in the molecular programs controlling osteoblasts and osteocytes. A limitation of this initial study was the inability to distinguish between gene expression changes occurring in pathways that might be specific to bone from those that occur in many cell types in response to differentiation. As a result, few skeletal biological processes were detected among differentially expressed genes and the identification of bone-specific changes was guided by prior knowledge of what may be skeletally important (St. John et al., 2014). Our examination of osteocyte signature gene expression patterns in this dataset allowed the identification of both known and unknown pathways enriched for expression in the skeleton *in-vivo*, changing with osteocyte network formation. In contrast to the previous study, this analysis identified the stage specific upregulation of many osteocyte-related regulatory processes involved in collagen production, ossification and osteoblast differentiation.

Examination of signature gene expression revealed a number of processes and pathways associated with the patterning of cell-cell contacts induced at the time of osteocyte network formation *in-vitro*. Among those, Rap1 signalling, though the MAP-PI3K signalling cascade, and TGF-Beta signalling have been shown to be critical to synapse

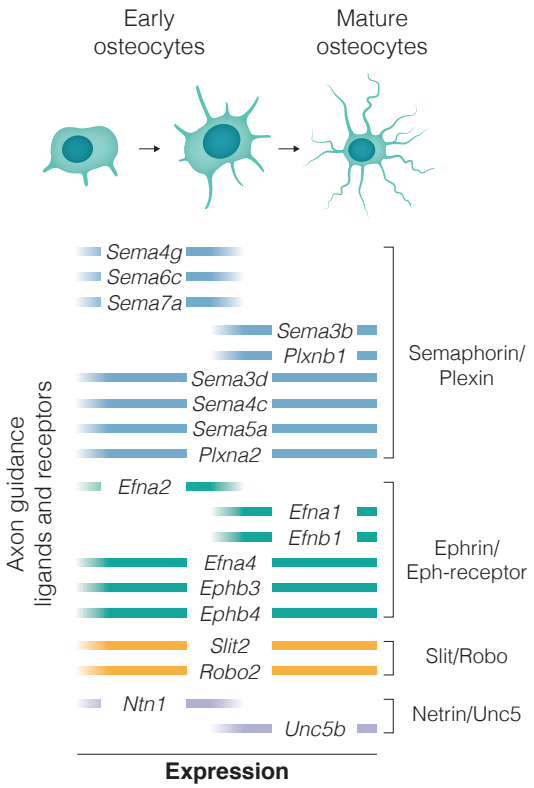


Figure 34 - Axon guidance signaling molecules during osteocyte differentiation and network formation.

Expression of ligands and receptors from the four key axon guidance pathways Semaphorin/Plexin, Ephrin/Eph-receptor, Slit/Robo and Netrin/Unc5 during osteocytic differentiation.

formation, stabilisation and plasticity in neurons (Fu et al., 2007; Ménager, Arimura, Fukata, & Kaibuchi, 2004; Yi et al., 2010). The importance of TGF-Beta signalling to osteocyte function is well established, however its mechanisms of action are still being unravelled (Tang & Alliston, 2013). In a recent investigation by Dole and colleagues, osteocyte specific perturbation of TGF-beta signalling was shown to reduced bone strength and fracture resistance (Dole et al., 2017). The authors of this work attributed this phenotype to a reduction in bone quality, as the bone mass and geometry in knockout mice was not affected. This was posited to be an effect of impaired perilacunar-remodelling, with demonstrated defects in dendrite formation resulting in abnormalities in the osteocyte network. Our analysis indicates these changes in the osteocyte network and downstream effects on bone quality may be due to the disruption of TGF-Beta controlled synapse formation. This would be a new function for the TGF-Beta pathway in osteocytes, akin to its role in the neurons of the brain.

Prominent in this analysis as a novel mode of osteocyte network communication were pathways and processes associated with axon guidance. Axon guidance upregulation coincided with network formation *in-vitro*, identified as the top ranked pathway in the transition cluster and the only significantly enriched KEGG pathway in the early activation cluster. Additionally, evidence that axonal guidance may play a functional role in the established osteocyte network was observed in its upregulation following limb amputation.

First described in neurons, axon guidance is the process by which cell types from many tissues form and maintain specific patterns of dendritic interconnection (Dodd & Jessell, 1988). It is directed by the combined activity of four main signalling pathways, with input from other processes including Wnt and BMP ligands, all seen in our data to be extensively expressed in osteocytes (Figure *34*). These represent both paracrine and juxtacrine modes of communication, with Slit/Robo and Netrin/Unc5 mediated through secreted factors, and members of the Semaphorin/Plexin and the Ephrin/Eph families triggered by cell-cell contact (Stoeckli, 2018). Very little is known about how this process contributes to the dynamics of osteocyte network formation and regulation however, several members of its constituent pathways have important roles in bone cells.

The semaphorin gene family contains both paracrine and juxtacrine ligands, originally identified to control axon pathfinding during the development of the nervous system, but now understood to also affect bone cell homeostasis. Among the paracrine class-3 semaphorins in the osteocyte signature, *Sema3b* expression in osteoblasts has been shown to promote osteoclastogenesis and induces osteopenia in mice (Sutton et al., 2008). Additionally, one of 5 juxtacrine semaphorins in the osteocyte signature, *Sema7a* expression promotes the fusion of osteoclast precursors, and is observed at all stages of *ex-vivo* osteoblast differentiation, consistent with its detection here in osteocytes (Delorme, Saltel, Bonnelye, Jurdic, & Machuca-Gayet, 2005). Variants in *SEMA7A* have been associated with decreased bone mineral density and increased fracture risk in human cohorts, however, whether the cellular and molecular basis for these effects are related to these *in-vitro* observations is not known (Koh et al., 2006). Another of the constituent axon guidance pathways, juxtacrine Ephrin-Eph signalling is critical to bone formation and regulation (Matsuo & Otaki, 2012). Bidirectional Ephrin signalling couples osteoblasts and osteoclasts activity, coordinating their differentiation and activity to form and breakdown bone during remodelling (Allan et al., 2008; Irie et al., 2009; C. Zhao et al., 2006). Ephrin signaling has also been proposed as a mechanism of communication between mature osteoblasts used to coordinate bone remodelling (Allan et al., 2008). In this study, EphrinB2 was seen to increase during osteoblast differentiation and in response parathyroid hormone (PTH) and local PTH-related protein. This study also noted the expression of EphrinB2 in osteocytes, consistent with our data indicating it is actively expressed in osteocytes but not specific to osteocytic stages of the osteogenic lineage. The expression of

complementary signals and receptors in osteocytes and other bone cells indicates a capacity to both send and receive information through Ephrins intercellular connections. While these examples demonstrate the potential for molecules associated with axon guidance to affect the skeleton, our data indicates their integrated activity may be an important factor in the formation and function of the osteocyte network. Recent work comparing the topological properties of the osteocyte network between mice and sheep suggests a universal mechanism may guide its organization during bone formation and mineralization (Kollmannsberger et al., 2017). Axon guidance molecules guide the formation and function of a number of networks in the body, including the nerves and blood vessels likely to be contained within the *in situ* isolated bone samples (Klingberg et al., 2017; Mach et al., 2002; Weinstein, 2005). However, the analysis indicating the upregulation of these pathways during osteocytic differentiation was derived from an *in vitro* system, free from these non-skeletal cell types. Consequently, this data provides strong evidence to suggest the expression of axon guidance molecules and their dynamic regulation during differentiation is occurring in the osteocyte network and not these other cell types. Together, these discoveries suggest that osteocytes utilise molecular machinery of synapse formation used by the brain and other tissues to establish and maintain their skeletal regulatory network.

8.6 The osteocyte transcriptome as a lens to genetic skeletal disease

The osteocyte network plays a pivotal role in skeletal regulation however our understanding of its contribution to skeletal genetic disease has been limited by the identification of genes important to its specialised function. Here we show that genes identified in the osteocyte transcriptome signature are significantly enriched for orthologs known to cause rare monogenic diseases and contribute to clinically significant skeletal traits in humans. The analysis of genes that cause skeletal-dysplasia indicated the osteocyte network is particularly important to diseases of bone composition and quality, characterised by extreme increases and decreases in bone strength. Underscoring this association was the osteocyte-enriched expression of genes that cause Osteogenesis Imperfecta (OI), upregulated in network maturation. While the many subtypes of OI differ in severity and patterns of inheritance, brittle bones and fracture are common disease defining features (Forlino & Marini, 2016). The role of the osteocyte network in OI is not well defined, with understanding of the cellular aetiology of disease centred on the collagen-production

by osteoblasts (Forlino, Cabral, Barnes, & Marini, 2011). However, high osteocyte density is seen in the bone of patients with severe OI subtypes, suggesting the later stages in the osteogenic lineage may also play an important role.

Blouin and colleagues reported a marked increase in osteocyte lacunae density and matrix mineralisation in bone from OI patients with mutations in *IFITM5* (Blouin et al., 2017). The authors of this study suggested these abnormalities were due to exuberant primary bone formation by osteoblasts and altered remodelling. The osteocytic control of osteoblasts is critical to the process of bone remodelling however the possibility that osteocyte-dysfunction may be driving this extreme phenotype was not addressed. Grafe and colleagues also observed increased osteocyte lacunae density in a *Crtap*$^{-/-}$ model of OI (Grafe et al., 2014). In this study, the authors identify excessive TGF-Beta signalling as a common molecular patho-mechanism in recessive (*Crtap*-affecting) and dominant (*Col1a2*-affecting) OI subtypes. They suggest increased TGF-Beta signalling increases the number of osteoblasts in *Crtap*$^{-/-}$ mice and that the increase in osteocyte density occurs due TGF-Beta blocking of osteoblast apoptosis, implying that more cells become embedded in the extracellular matrix. Here, we observe the coincidental upregulation of the TGF-Beta signalling pathway coinciding with osteocytic network formation, suggesting it may play an important role in the formation and integration of cells into the osteocyte network. Suppression of TGF-Beta activity alters osteocyte network formation and function reduces bone strength (Dole et al., 2017). This suggests abnormal TGF-Beta signalling has the potential to directly effect on osteocyte network dynamics and may present an alternative explanation for the abnormalities in osteocyte density observed in severe OI. It may also be possible that the changes in the osteocyte network in OI are due to interactions between osteocytes and their environment, with changes in the extracellular matrix altering expression of these genes in the osteocytes. In such a model, gene dysfunction in osteoblasts would lead to altered bone composition and mechanical properties, sensed by the osteocytes leading to abnormalities in the osteocyte network. Future investigations with cell-lineage specific knockouts may help reveal the cause and/effect nature of OI mutations in the osteocyte network.

Evidence that the osteocyte network contributes to certain subtypes of OI was also reported by Joeng and colleagues (Joeng et al., 2017). Their study showed that osteocyte

specific knockout of *Wnt1* resulted in a low bone mass and spontaneous fractures consistent with the clinical features of OI, which was reversed with *Wnt1* overexpression. While this demonstrated the clinical importance osteocyte gene expression in bone fragility diseases, whether the effects of this expression were realised through the osteocytic-control of osteoblast activity or through a direct cell autonomous mechanism was not elucidated. Our data support the osteocyte-specific expression of *Wnt1* however, our analysis also indicated *Wnt1* was not expressed in the calvarium or *in-vitro* samples, suggesting this may be restricted to the appendicular skeleton and highly dependent on cellular context. Moreover, our data suggests *Wnt1* is unique among known OI causal genes in the osteocyte signature as its expression increases with skeletal maturation. Interestingly, the skeletal characteristics and response to bisphosphonate treatment is distinct in patients carrying *WNT1* mutations as compared to those with defects in *COL1A1/COL1A2* (Palomo et al., 2014). These differences led the authors of this study to suggest that bone diseases caused by mutations in *WNT1* should be considered distinct from other OI subtypes. Our data supports this distinction and suggests that the differences in disease characteristics and drug response may be related to its regulation in the osteocyte network during skeletal maturation. This demonstrates how the insights into the osteocyte transcriptome reported here may help guide investigations into the molecular mechanism and classification of skeletal genetic disease. As such, the osteocyte transcriptome signature might be a powerful biologically-reasoned filter to identify causal genes and guide follow-up experiments in bone-quality conditions that lack molecular diagnosis.

The potential of the osteocyte signature as a filter to identify clinically relevant genes is not limited to rare skeletal diseases. Public health initiatives around the world, like that of the UK Biobank, are producing enormous genotype-phenotype datasets, enabling GWAS of skeletal-BMD among a variety of complex genetic traits (Hagenaars et al., 2016; J. M. Lane et al., 2016; UK Biobank, 2007; Wain et al., 2015). This is shifting the challenge from powering studies capable of identifying *loci*, to the identification of plausible gene associations for follow up investigations (Timpson, Greenwood, Soranzo, Lawson, & Richards, 2018). The use of gene expression data from disease-trait relevant cell types has proven an effective strategy in other diseases-traits, and recent studies also indicate this may also be a powerful strategy for skeletal-BMD (Hertzberg, Katsel, Roussos,

Haroutunian, & Domany, 2015; Sleiman et al., 2014). Our analysis indicates the osteocyte transcriptome signature is enriched not just for BMD associated genes, but specifically for those that are supported by molecular evidence, and phenotypic associations in mice. Intersecting BMD associated genes with the transcriptome data from osteocytes other tissues and organs identified *Ldlrad4*, a regulator of TGF-Beta signalling (Nakano et al., 2014). While not previously annotated with a skeletal effect, screening of *Ldlrad4*$^{-/-}$ mice through the OBCD pipeline showed significantly reduced bone mineral content and trabecular abnormalities. This demonstrates how osteocyte transcriptome data and analyses performed here may be used in the selection of plausible gene associations for further experimental investigation following GWAS.

An important consideration of these studies, particularly in the extrapolation of findings to the interpretation of human disease, is the use of mouse models. The conserved pattern of expression and function of human-mouse genes orthologs in other skeletal cells is thought to be high, however the deficit of primary osteocyte transcriptome data from humans mean estimates for these cells is difficult (Liao, 2005). The conserved expression and skeletal effect of many genes important in human biology in mouse models support the interpretability between the two systems, however there are many instances where gene expression and effect diverge (Thomas, Bourne, Eisman, & Gardiner, 2000). Additionally, there are important differences between the human skeleton and that of the mouse. The lack of osteonal cortical bone remodelling in mice may be the most overt difference between species. While blood vessels do exist on murine cortical bone (Núñez et al., 2018), there is no Haversian system. However, intracortical remodelling does increase cortical porosity in aging mice (Piemontese et al., 2017), an important hallmark of age-related bone loss in humans (Ramchand & Seeman, 2018). While the cellular mechanics of Haversian and surface-based remodelling are similar, whether there are differences in the molecular machinery of each process is yet to be defined. To best power the use of these animal models for investigations into genes affect human skeletal pathophysiology, transcriptome analysis of human osteocytes without removing them from their local *in-vivo* context will be of critical importance.

Collectively, these analyses demonstrate the osteocyte-enriched expression of many genes known to cause skeletal disease and contribute to clinically relevant skeletal traits.

This work suggests the transcriptome of the osteocyte network may be an important lens to understanding the cellular and molecular aetiology of skeletal pathophysiology.

9 CONCLUSIONS

9.1 Key findings

The osteocyte network is a dynamic, multifunctional strain-sensing cellular-control system with a regulatory role that extends beyond the skeleton. The function of the osteocyte network is intimately coupled to its position within one of the hardest tissues of the body. This fact makes its study challenging and has limited the identification of the genes, biological processes and biochemical pathways that enable its specialised role. It was hypothesised that defining the active transcriptome expressed in the osteocyte network would reveal the molecular control-processes key to its formation and specialised function and provide insights into the contribution of the osteocyte network in skeletal disease. To test this, techniques for the system-scale study of gene expression in the osteocyte network *in-situ* were developed, maintaining intercellular contacts and the local extracellular environment. The transcriptome investigations conducted with these methodologies increase the number of genes known to be specifically expressed in the osteocyte-network by an order of magnitude. Supporting this hypothesis, here I report a number of new insights into the molecular machinery involved in the formation and function of the osteocyte network and provide evidence that these genes make a significant contribution to monogenic skeletal disease as well as complex, clinically-relevant skeletal traits. These include:

Technical considerations

- Gene expression in the osteocyte network is stable for up to 90min without a circulatory supply.
- The osteocyte transcriptome is effectively identical across the parasagittal plane and is very similar, but not identical, between bone types.
- Gene expression is dynamically regulated in the osteocyte network during skeletal maturation and is affected by both age and sex.

The dynamic regulation of gene expression in the osteocyte network

- The osteocyte transcriptome is distinct from other organs and tissues.
- The site-specific gene expression patterns in the osteocyte network are distinct from those reported in other osteogenic cell-types.

- Genes encoding structural extracellular matrix proteins are down regulated while endocrine signalling receptors are upregulated in the osteocyte network as the skeleton matures.
- Genes associated with perilacunar-remodelling are dynamically regulated during skeletal maturation and differentially expressed between sexes at skeletal maturity.

Osteocyte-enriched genes and pathways

- The majority of the genes specifically enriched in the osteocyte network have an unknown role in skeletal biology. These include protein-coding genes that cause significant skeletal phenotypes when knocked out in mice and lncRNAs which have been reported to play important regulatory roles in other tissues.
- More than 100 novel genes are actively expressed in the osteocyte network, including several that appear to be skeletally restricted in their expression.
- Biological processes, cellular components and biochemical pathways associated with axon guidance and intercellular network formation are enriched in the osteocyte network and specifically upregulated early in osteocytic differentiation.

Clinically relevant insights

- Genes specifically enriched in the osteocyte-network are significantly associated with a range of human skeletal diseases.
- A large fraction of genes causing monogenic skeletal disorders affecting bone composition and mineral homeostasis are specifically expressed in the osteocyte network.
- More than 80% of the genes that cause osteogenesis imperfecta are identified in the osteocyte transcriptome signature and are down regulated during skeletal maturation.
- Genes associated in GWAS with variation in human bone mineral density are significantly overrepresented in the osteocyte transcriptome signature.
- A significant fraction of osteocyte signature genes associated with BMD-variance are supported by experimental evidence of skeletal effect.

Collectively, these insights help define the osteocyte transcriptome, with skeletal site-specific differences, dynamic regulation during skeletal maturation and a diversity of

external signalling pathways combining to meet mechanical and mineral demands. These findings potentially reveal the molecular machinery of osteocyte network formation, repurposing axon guidance systems used in other organs to establish and maintain patterns of dendritic connections. This work also reinforces the critical role of the osteocyte network in the regulation of bone composition and mineral homeostasis, demonstrating that genes associated with disorders of these processes are highly enriched for expression in osteocytes. The overrepresentation of skeletal-disease associated genes in the osteocyte transcriptome signature suggests specific expression in the osteocyte network may be a powerful filter to identify pathogenic variants and mutations. Together, the methodologies, data and discoveries reported here will provide a platform for future studies of the osteocyte network and their role in human bone diseases.

9.2 Limitations

There are a number of inherent limitations associated with transcriptomic analyses which need to be considered when interpreting this work. The scale of the datasets and gene lists generated by such techniques make it impractical to manually curate the available literature for all genes identified (the osteocyte transcriptome signature alone identifies 830 genes). As an alternative I have made use of curated gene-function databases to provide somewhat heuristic measurements of gene function, such as phenotypic associations in the MGI and annotations to biological processes in the gene ontology database. The use of such techniques to infer potential function from gene lists are well established, with the published evidence supporting each annotation made available. Using these tools, I identify a number of associations, correlations and over representation to provide evidence for the molecular machinery used by the osteocyte network to control the skeleton in health and disease. These revelations should not be taken as definitive statements rather observations that may form the basis of specific hypotheses to be tested in future work.

It is also important to acknowledge that many of the gene expression pattern reported here are yet to be validated using orthogonal approaches, such as fluorescent in situ hybridisation or immunohistochemistry. Candidates for validation are best selected based on specific questions about specific molecules as such techniques used are time consuming, expensive and low throughput. As the nature of this work was intended to be

"hypothesis generating", these approaches were not pursued in this thesis. However, such investigations will be required in future studies to confirm the associations reported here.

9.3 Future directions

The key findings of this thesis generate a number of new hypotheses as to the role of osteocytes within the skeleton and the molecular machinery important to their function. These hypotheses underlie a number of lines of ongoing investigation.

Hypothesis: Osteocyte transcriptome signature genes can be used as sentinels to identify common molecular control networks and pathways important to skeletal function. The osteocyte signature revealed many hundreds of genes enriched in the osteocyte network whose function in the skeleton is poorly defined or completely unknown. Unbiased screening through the Origins of Bone and Cartilage Disease (OBCD) skeletal phenotyping pipeline has revealed many signature genes that have significant skeletal consequences when deleted in mice (Bassett et al., 2012). To understand how these deletions lead to skeletal abnormalities, several of these signature knockouts have been selected for further transcriptome interrogation including *Cadm1*$^{-/-}$, *Daam2*$^{-/-}$ and *Spns2*$^{-/-}$. Moreover, samples collections and sequencing experiments have been conducted trialling the direct integration of the *in-situ* osteocyte isolation methodology in the Sanger knockout pipeline. The success of such an endeavour would make available material for the sequencing of the osteocyte network in hundreds of comprehensively phenotyped knockout mice lines. Comparing the osteocyte transcriptome in a range of skeletally significant lines could prove a powerful approach to discovering common genes and control pathways affecting bone.

Hypothesis: Novel osteocyte-restricted lncRNAs influence osteocyte network function and skeletal phenotype. This work also revealed a number of novel long non-coding genes restricted in expression to the osteocyte network. In other organ systems such as the brain, tissue-restricted lncRNAs have been shown to play important roles in tissues specific biological processes (Barry, 2014). To understand how these genes may influence the skeleton, four novel gene knockout mice lines (Obcdi008175, Obcdi007392, Obcdi042809 and Obcdi053500) have been developed and are being expanded for further molecular and phenotypic characterisation. Additionally, the identification of novel cell restricted genes and promoters can have considerable utility in the development of new

tools for molecular manipulation. Future work, establishing the potential of these novel *loci* as markers of the osteocyte network, may be coupled with the development of new models for the investigation of the osteocyte network.

Hypothesis: The osteocyte transcriptome signature, and specifically those genes within the brown skeletal maturation cluster, can be used to identify novel mutations that cause OI. The significant over representation of orthologs associated with human skeletal dysplasia among genes specifically enriched in the osteocytes supports evidence that the osteocyte network plays a key role in skeletal genetic diseases. These data suggest the osteocyte transcriptome could be a powerful filter to identify causal genes in patients with genetic skeletal diseases that lack molecular diagnosis, particularly in diseases of abnormal bone composition, mineralisation or density. In line with this and in collaboration with Genomics England, the utility of the osteocyte transcriptome signature as a screen for genomic data from a large cohort of OI patients for which the causal gene is unknown is currently being examined. This work may provide valuable insights into the molecular and cellular aetiology of rare skeletal diseases and the role of the osteocyte network in the pathophysiology of skeletal diseases.

Hypothesis: The osteocyte transcriptome signature can identify novel genes that influence bone mineral density variance in human populations. The identification of plausible candidate genes associated with significant *loci* remains a considerable challenge in the GWAS community. The intersection of GWAS *loci* with transcriptome data from trait-relevant tissues is one strategy being used to explore the genetic architecture contributing to clinically significant traits affecting other organ systems. As such, the osteocyte transcriptome data and signature have been integrated into the gene prioritisation pipelines of leading BMD GWAS consortia. As an alternative to GWAS, family studies are a powerful approach to uncovering genes contributing to heritable traits. Similarly, the use of the osteocyte signature to identify variants contributing to bone mass variation between relatives has revealed plausible candidates that are being validated in mouse models. This uptake by the skeletal genetic community suggests the data and discoveries reported here will help guide future investigations into the genetic complexity of skeletal traits and common skeletal diseases.

Hypothesis: The osteocyte network plays a role in the growth of cancers in the skeleton. Cancers that grow in bone disrupt the dynamics of skeletal regulation and compromise skeletal structure. Despite the master regulatory role of the osteocyte network the skeleton little is known about how it is affected by tumour burden. Transcriptome sequencing on *in-situ* isolated osteocyte samples from bones burdened with multiple myeloma has been compared with wild-type samples. Initial analysis revealed the altered expression of sclerostin along with a number of secreted inhibitors of Wnt-signalling, which can be targeted with therapeutic antibodies to build bone lost in disease. Future work will utilise the subtractive sequencing experimental methodology documented here to identify tumour-driven changes occurring specifically in the osteocyte network. These insights have the potential to advance understanding of the skeletal effects of cancers growing in bone and potentially reveal new regulatory pathways important to skeletal health and disease.

9.4 Concluding statement

The important role the skeletal system plays keeping us upright and healthy is often underestimated. If you move beyond the field of bone biology, fellow scientists are often surprised to hear to that their skeleton is almost completely replaced every decade or that the survival following hip fracture is comparable to that following breast cancer. Part of the reason the skeleton is so often overlooked is that it is made from a difficult tissue to study. While we know a lot about the effector skeletal cells that reside on the bone-surface, the key controlling cells, the osteocytes, are buried in the skeleton itself. Despite these difficulties, decades of discovery have defined molecular pathways controlling skeletal homeostasis and demonstrated distant interactions with a diversity of organs. Guided by these insights, this work capitalises on advances in technology to orient the expression of hundreds of skeletally novel genes within the osteocyte network, promoting them into the realm of "known-unknowns", the fuel of further research.

The limited understanding of the genes important to the osteocyte networks control of the skeleton has made it challenging to define its role in skeletal genetic disease. Here we provide evidence that genes enriched for expression in the osteocyte network are associated with a range of skeletal diseases and show that may be characterised by distinct patterns of expression during cell differentiation and skeletal maturation. These findings suggest the osteocyte network may play an important role a range skeletal diseases, and

indicates that the osteocyte transcriptome signature may be a useful tool to aid future discovery of genes important to skeletal health.

More broadly, these data may prove a valuable resource for bringing investigation of the skeleton into the wider scientific community. The scarcity of skeletal sequencing data has meant bone is often overlooked in analyses of gene expression throughout the body. In defining the transcriptome of the osteocyte network, the methodologies, data and discoveries documented here may enable an integrative understanding of the skeleton as one of the bodies major organ systems.

10 APPENDICES

Appendix 1 - Site-specific gene activity in the osteocyte network

Gene symbol	Tib	Fem	Hum	Cal	Activity
Zic4	0.13	0.00	0.01	2.25	Calvaria
Zic2	0.00	0.01	0.02	2.08	Calvaria
Tbx5	0.03	0.18	1.78	0.00	Fore-limb
Hoxd9	0.32	0.28	1.43	0.02	Fore-limb
Hoxc11	1.31	1.19	0.01	0.00	Hindlimb
Hoxc8	4.87	4.32	0.05	0.00	Hindlimb
Pitx1	2.99	1.80	0.13	0.00	Hindlimb
Hoxc10	8.08	5.87	0.01	0.00	Hindlimb
Hoxc9	3.93	3.21	0.02	0.00	Hindlimb
Wnt1	2.95	2.10	3.09	0.05	Long-bones
Sowahb	2.46	2.25	2.58	0.16	Long-bones
Adamtsl2	12.14	13.38	12.84	0.31	Long-bones
Irx6	3.72	4.65	4.49	0.14	Long-bones
Grm4	1.28	1.65	1.21	0.12	Long-bones
Grip2	3.77	3.47	3.97	0.29	Long-bones
9330154J02Rik	2.97	2.62	3.03	0.23	Long-bones
ENSMUSG00000105079	1.70	1.63	2.10	0.27	Long-bones
Tmem151a	1.64	1.30	1.61	0.12	Long-bones
9330159N05Rik	1.47	1.52	1.56	0.07	Long-bones
9130019P16Rik	1.99	2.27	2.79	0.16	Long-bones
Osr1	5.84	6.72	6.34	0.15	Long-bones
Gm13441	1.71	1.79	2.06	0.07	Long-bones
Gm15998	1.56	1.47	1.59	0.15	Long-bones
Obcdi010645	1.67	1.85	2.12	0.13	Long-bones
Hoxc5	5.80	5.11	3.17	0.08	Long-bones
Grik2	3.86	6.10	5.38	0.32	Long-bones
Gm4117	1.34	1.08	1.30	0.23	Long-bones
Colgalt2	1.26	1.22	1.28	0.34	Long-bones
Fhl5	1.80	1.67	2.12	0.19	Long-bones
Adamts18	2.52	2.61	2.96	0.29	Long-bones
Lingo3	2.28	2.29	2.11	0.27	Long-bones
Hoxc4	4.88	4.35	2.67	0.00	Long-bones
Olfr558	2.43	2.30	2.56	0.30	Long-bones

Gene symbol	Tib	Fem	Hum	Cal	Activity
Rnf152	0.97	0.97	1.15	0.30	Long-bones
Obcdi047271	1.79	1.54	1.48	0.18	Long-bones
Kcnk3	1.34	1.13	1.71	0.08	Long-bones
NA	1.61	1.54	5.90	0.29	Long-bones
Hoxc6	13.44	12.66	6.69	0.01	Long-bones
Hoxb2	1.24	2.49	8.05	0.06	Long-bones
Plekhg6	1.51	1.22	1.32	0.09	Long-bones
Bmp8b	1.73	1.61	1.72	0.11	Long-bones
Mab21l2	6.01	4.51	4.94	0.05	Long-bones
Exph5	1.12	1.53	1.78	0.17	Long-bones
Tbx18	3.83	1.89	3.71	0.17	Long-bones
NA	1.40	1.88	1.71	0.12	Long-bones
Hoxa10	6.14	5.14	5.45	0.05	Long-bones
Fetub	4.40	3.32	2.83	0.22	Long-bones
Hoxa6	2.65	2.56	2.03	0.07	Long-bones
Sod3	2.53	1.96	2.73	0.38	Long-bones
Hoxa5	3.29	3.64	3.49	0.28	Long-bones
Hoxa4	1.24	1.30	1.69	0.11	Long-bones
Fam84a	1.56	1.61	0.98	0.16	Long-bones
Thpo	1.18	1.05	1.25	0.22	Long-bones
Macrod2os2	1.24	1.37	1.04	0.06	Long-bones

Appendix 2 – The osteocyte transcriptome signature.

The size of the osteocyte transcriptome signature makes it impractical to be included as a table in print. A digital copy of the osteocyte transcriptome signature can be obtained by contacting Scott Youlten (s.youlten@garvan.org.au) or Professor Peter Croucher (p.croucher@garvan.org.au).

Appendix 3 - Significantly enriched GO biological processes in the osteocyte transcriptome signature.

GO ID	Description	P-value	Ratio
GO:0001503	ossification	8.82E-22	75/322
GO:0001501	skeletal system development	2.43E-17	75/377
GO:0007423	sensory organ development	4.98E-11	63/365
GO:0031214	biomineral tissue development	1.30E-10	30/98
GO:0030198	extracellular matrix organization	1.97E-10	39/164
GO:0043062	extracellular structure organization	1.97E-10	39/164
GO:0001649	osteoblast differentiation	2.63E-09	40/185
GO:0042472	inner ear morphogenesis	8.85E-09	23/67
GO:0061564	axon development	1.22E-08	54/321
GO:0030278	regulation of ossification	1.91E-08	37/171
GO:0060348	bone development	3.57E-08	35/158
GO:0048812	neuron projection morphogenesis	4.35E-08	64/433
GO:0048839	inner ear development	5.99E-08	31/129
GO:0032990	cell part morphogenesis	7.34E-08	67/470
GO:0050808	synapse organization	8.70E-08	38/188
GO:0044236	multicellular organism metabolic process	1.41E-07	24/82
GO:0048667	cell morphogenesis involved in neuron differentiation	1.47E-07	60/403
GO:0090596	sensory organ morphogenesis	1.88E-07	37/184
GO:0030282	bone mineralization	1.88E-07	24/83
GO:0048562	embryonic organ morphogenesis	3.80E-07	39/206
GO:0048705	skeletal system morphogenesis	1.65E-06	35/180
GO:0035295	tube development	2.77E-06	65/488
GO:0043583	ear development	4.15E-06	31/151
GO:0032963	collagen metabolic process	9.83E-06	20/70
GO:0048568	embryonic organ development	1.19E-05	49/331
GO:0030199	collagen fibril organization	1.51E-05	15/40
GO:0061448	connective tissue development	1.83E-05	34/187
GO:0048598	embryonic morphogenesis	1.91E-05	60/455
GO:0016055	Wnt-signalling pathway	2.02E-05	46/305
GO:0060021	palate development	2.20E-05	20/73
GO:0198738	cell-cell signaling by Wnt	2.24E-05	46/306
GO:0007169	transmembrane receptor protein tyrosine kinase signaling pathway	2.77E-05	57/426

GO ID	Description	P-value	Ratio
GO:0071363	cellular response to growth factor stimulus	2.78E-05	56/415
GO:0048589	developmental growth	2.91E-05	63/494
GO:0051216	cartilage development	2.97E-05	28/137
GO:1905114	cell surface receptor signaling pathway involved in cell-cell signaling	3.70E-05	50/353
GO:0060541	respiratory system development	4.52E-05	32/175
GO:0001655	urogenital system development	5.32E-05	39/243
GO:0050877	neurological system process	6.02E-05	60/469
GO:0007411	axon guidance	6.16E-05	25/116
GO:0070848	response to growth factor	7.13E-05	56/426
GO:0097485	neuron projection guidance	7.40E-05	25/117
GO:0030178	negative regulation of Wnt signaling pathway	9.28E-05	24/110
GO:0050919	negative chemotaxis	9.44E-05	12/28
GO:0030324	lung development	9.59E-05	29/153
GO:0030323	respiratory tube development	1.50E-04	29/156
GO:0007389	pattern specification process	1.52E-04	42/283
GO:0006935	chemotaxis	2.08E-04	49/361
GO:0042330	taxis	2.48E-04	49/363
GO:0060349	bone morphogenesis	4.73E-04	18/71
GO:0048514	blood vessel morphogenesis	5.38E-04	55/440
GO:0001525	angiogenesis	5.41E-04	48/361
GO:0060828	regulation of canonical Wnt signaling pathway	8.47E-04	28/159
GO:0001822	kidney development	9.13E-04	32/198
GO:0071772	response to BMP	9.82E-04	23/115
GO:0071773	cellular response to BMP stimulus	9.82E-04	23/115
GO:0010092	specification of animal organ identity	1.10E-03	9/18
GO:0010975	regulation of neuron projection development	1.19E-03	52/416
GO:0042476	odontogenesis	1.32E-03	17/68
GO:0045778	positive regulation of ossification	1.44E-03	18/76
GO:0060322	head development	2.07E-03	53/435
GO:0048729	tissue morphogenesis	2.16E-03	55/459
GO:0055023	positive regulation of cardiac muscle tissue growth	2.45E-03	11/30
GO:0007507	heart development	3.47E-03	54/454
GO:0030900	forebrain development	3.53E-03	35/241
GO:0045667	regulation of osteoblast differentiation	3.84E-03	21/106
GO:0048565	digestive tract development	3.91E-03	17/73

GO ID	Description	P-value	Ratio
GO:0060563	neuroepithelial cell differentiation	4.09E-03	13/44
GO:0007416	synapse assembly	4.24E-03	20/98
GO:0048736	appendage development	4.33E-03	25/143
GO:0051962	positive regulation of nervous system development	4.93E-03	50/412
GO:0045165	cell fate commitment	4.95E-03	25/144
GO:0042692	muscle cell differentiation	5.02E-03	43/332
GO:0043410	positive regulation of MAPK cascade	5.32E-03	44/344
GO:0035108	limb morphogenesis	5.35E-03	23/126
GO:0002009	morphogenesis of an epithelium	5.81E-03	48/391
GO:0035265	organ growth	5.96E-03	24/136
GO:0035239	tube morphogenesis	6.69E-03	40/302
GO:0055123	digestive system development	8.32E-03	18/85
GO:0007420	brain development	8.98E-03	48/397
GO:0042475	odontogenesis of dentin-containing tooth	9.37E-03	13/47
GO:0060485	mesenchyme development	1.00E-02	28/179
GO:0016049	cell growth	1.01E-02	47/387
GO:0021954	central nervous system neuron development	1.44E-02	14/56
GO:0090287	regulation of cellular response to growth factor stimulus	1.53E-02	29/193
GO:0007156	homophilic cell adhesion via plasma membrane adhesion molecules	1.56E-02	13/49
GO:0097435	supramolecular fiber organization	1.77E-02	55/491
GO:0060350	endochondral bone morphogenesis	2.70E-02	12/44
GO:1901863	positive regulation of muscle tissue development	2.79E-02	14/59
GO:0048636	positive regulation of muscle organ development	2.79E-02	14/59
GO:0060562	epithelial tube morphogenesis	2.88E-02	36/275
GO:0007605	sensory perception of sound	3.30E-02	16/76
GO:0055024	regulation of cardiac muscle tissue development	4.18E-02	15/69
GO:0090090	negative regulation of canonical Wnt signaling pathway	4.24E-02	17/86
GO:0048638	regulation of developmental growth	4.47E-02	34/258

Appendix 4 - Key resources table

Details of experimental models, reagents, datasets and software used in this work.

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Chemicals, Peptides, and Recombinant Proteins		
Mayer's hematoxylin and eosin	Sigma-Aldrich	CAT# MHS1
Critical Commercial Assays		
TRIreagent	Sigma-Aldrich	CAT# T9424
Direct-zol columns	Zymo Research	CAT # R2061
RNAlater-ICE	ThermoFisher	CAT# AM7030
RNA 6000 Nano Kit	Agilent Technologies	CAT# 5067-1511
Hybridase Thermostable RNase Hybridase	Epicentre	CAT# H39500
TURBO DNase	Life Technologies	CAT# AM2238
TruSeq Stranded Total RNA LT Sample Prep Kit	Illumina	CAT# RS-122-2201
Agencourt RNAClean XP	Beckman Coulter Genomics	CAT# A63987
ERCC ExFold RNA Spike-in Mixes	Life Technologies	CAT# 4456739
DNase I	New England Biosciences	CAT# M0303S
Tetro cDNA Synthesis Kit	Bioline	CAT# BIO-65043
iTAQ Universal SYBR Green Supermix	Biorad	CAT# 1725121

Data		
Osteocyte transcriptome data – sample delay cohort	This project	E-MTAB-5532
Osteocyte transcriptome data – bone comparison cohort	This project	E-MTAB-5532
Osteocyte transcriptome data – osteocyte enrichment cohort	This project	E-MTAB-5533
Osteocyte transcriptome data – skeletal maturation cohort	This project	TBD
Gene array data from osteocytes in virgin, lactating and post lactation mice	Qing et al., 2012	E-GEOD-23496
Osteoblast to osteocyte transition transcriptome data	St. John et al., 2014	E-GEOD-54783
Transcriptome data from 12 organs and tissues	R. Zhang et al., 2014	E-GEOD-54652
GENCODE M5 (GRCm38.p3) comprehensive gene annotation	Mudge & Harrow, 2015	ftp://ftp.sanger.ac.uk/pub/gencode/Gencode_mouse/release_M5/gencode.vM5.annotation.gtf.gz
RefSeq transcriptome annotation	Pruitt, Tatusova, Brown, & Maglott, 2012	http://genome.ucsc.edu/ via “Table Browser” on the 15 October, 2015.
Gene Ontology (GO) terms and GO IDs	Ashburner et al., 2000; Carbon et al., 2017	http://www.informatics.jax.org/downloads/reports/go_terms.mgi
Mammalian Phenotype terms and descriptions	(Smith & Eppig, 2009)	http://www.informatics.jax.org/downloads/reports/VOC_MammalianPhenotype.rpt
MGI knockout mouse phenotyping data, excluding conditional mutations	Smith, Blake, Kadin, Richardson, & Bult, 2018	http://www.informatics.jax.org/downloads/reports/MGI_GenePheno.rpt
GWAS Catalog	MacArthur et al., 2017	www.ebi.ac.uk/gwas. Accessed 26/09/2017, version 1.0.1
Experimental Models: Organisms/Strains		

Mouse: C57BL/6NTac (NTac)	Wellcome Trust Sanger Institute	RRID:IMSR_TAC:b6
Software and Algorithms		
affy v1.56.0	(Gautier, Cope, Bolstad, & Irizarry, 2004)	https://bioconductor.org
Affycoretools v1.50.6	(MacDonald, 2008)	https://bioconductor.org
Aperio Imagescope	Leica	https://www.leicabiosystems.com/digital-pathology/manage/aperio-imagescope/
clusterProfiler v3.4.4	(Yu, Wang, Han, & He, 2012)	https://bioconductor.org
CPAT	Wang et al., 2013	http://lilab.research.bcm.edu/cpat/index.php
CTAn	Bruker	http://bruker-microct.com/products/downloads.htm
Cufflinks v2.2.1	(Trapnell et al., 2012)	http://cole-trapnell-lab.github.io/cufflinks
DOSE v3.2.0	(Yu, Wang, Yan, & He, 2015)	https://bioconductor.org
Drishti-2 v2.6.1	(Limaye, 2012)	https://github.com/nci/drishti
Fiji	ImageJ	https://fiji.sc/
ggplot2 v2.2.1	(Wickham, 2016)	https://cran.r-project.org
gplots v3.0.1	(Warnes et al., 2009)	https://cran.r-project.org
GraphPad Prism v7	GraphPad Software	https://www.graphpad.com/scientific-software/prism/
Gviz v1.20.1	(Hahne & Ivanek, 2016)	https://bioconductor.org
ICSNP v1.1-1	(Nordhausen, Sirkiä, Oja, & Tyler, 2007)	https://cran.r-project.org

limma v3.32.7	(Law, Chen, Shi, & Smyth, 2014; Ritchie et al., 2015)	https://bioconductor.org
mclust v5.3	(Fraley & Raftery, 2003)	https://cran.r-project.org
mixtools v1.1.0	(Benaglia, Chauveau, Hunter, & Young, 2009)	https://cran.r-project.org
NRecon	Bruker	http://bruker-microct.com/products/downloads.htm
OsteoMeasure	Osteometrics	http://www.osteometrics.com/
Pathview v1.16.5	(Luo & Brouwer, 2013)	https://bioconductor.org
R v3.4.0	(Team, 2013)	https://www.r-project.org
reshape2 v1.4.2	(Wickham, 2007)	https://cran.r-project.org
ReViGo	(Supek, Bošnjak, Škunca, & Šmuc, 2011)	http://revigo.irb.hr
RSEM v1.2.21	(Li & Dewey, 2011)	http://deweylab.github.io/RSEM/
STAR v2.4.1d	(Dobin et al., 2013)	https://github.com/alexdobin/STAR
Stringtie v1.0.4	(Pertea et al., 2015)	https://ccb.jhu.edu/software/stringtie
tm v0.7-1	(Meyer, Hornik, & Feinerer, 2008)	https://cran.r-project.org
TrimGalore v0.3.3	(Krueger, 2015)	http://www.bioinformatics.babraham.ac.uk/projects/trim_galore
Trinity v2.0.6	(Haas et al., 2013)	https://github.com/trinityrnaseq/trinityrnaseq
VennDiagram v1.6.17	(H. Chen & Boutros, 2011)	https://cran.r-project.org

11 BIBLIOGRAPHY

Adiconis, X., Borges-Rivera, D., Satija, R., Deluca, D. S., Busby, M. A., Berlin, A. M., … Levin, J. Z. (2013). Comparative analysis of RNA sequencing methods for degraded or low-input samples. *Nature Methods*, *10*(7), 623–629. https://doi.org/10.1038/nmeth.2483

Agarwal, P., Wylie, J. N., Galceran, J., Arkhitko, O., Li, C., Deng, C., … Bruneau, B. G. (2003). Tbx5 is essential for forelimb bud initiation following patterning of the limb field in the mouse embryo. *Development (Cambridge, England)*, *130*(3), 623–633. https://doi.org/10.1242/dev.00191

Aguirre, J. I., Plotkin, L. I., Stewart, S. A., Weinstein, R. S., Parfitt, A. M., Manolagas, S. C., & Bellido, T. (2006). Osteocyte apoptosis is induced by weightlessness in mice and precedes osteoclast recruitment and bone loss. *Journal of Bone and Mineral Research*, *21*(4), 605–615. https://doi.org/10.1359/jbmr.060107

Ai, M., Holmen, S. L., Van Hul, W., Williams, B. O., & Warman, M. L. (2005). Reduced affinity to and inhibition by DKK1 form a common mechanism by which high bone mass-associated missense mutations in LRP5 affect canonical Wnt signaling. *Molecular and Cellular Biology*, *25*(12), 4946–4955. https://doi.org/10.1128/MCB.25.12.4946-4955.2005

Ajima, R., Bisson, J. A., Helt, J. C., Nakaya, M. A., Habas, R., Tessarollo, L., … Cohen, E. D. (2015). DAAM1 and DAAM2 are co-required for myocardial maturation and sarcomere assembly. *Developmental Biology*, *408*(1), 126–139. https://doi.org/10.1016/j.ydbio.2015.10.003

Allan, E. H., Häusler, K. D., Wei, T., Gooi, J. H., Quinn, J. M. W., Crimeen-Irwin, B., … Martin, T. J. (2008). EphrinB2 regulation by PTH and PTHrP revealed by molecular profiling in differentiating osteoblasts. *Journal of Bone and Mineral Research*, *23*(8), 1170–1181. https://doi.org/10.1359/jbmr.080324

Alswat, K. A. (2017). Gender Disparities in Osteoporosis. *Journal of Clinical Medicine Research*, *9*(5), 382–387. https://doi.org/10.14740/jocmr2970w

Arany, S., Koyota, S., & Sugiyama, T. (2009). Nerve growth factor promotes differentiation of odontoblast-like cells. *Journal of Cellular Biochemistry*, *106*(4), 539–545. https://doi.org/10.1002/jcb.22006

Asada, N., Katayama, Y., Sato, M., Minagawa, K., Wakahashi, K., Kawano, H., … Tanimoto, M. (2013). Matrix-embedded osteocytes regulate mobilization of hematopoietic stem/progenitor cells. *Cell Stem Cell*, *12*(6), 737–747. https://doi.org/10.1016/j.stem.2013.05.001

Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., … Sherlock, G. (2000). Gene ontology: Tool for the unification of biology. *Nature Genetics*, *25*(1), 25–29. https://doi.org/10.1038/75556

Australian Institute of Health and Welfare. (2014). *Estimating the prevalence of osteoporosis in Australia*. Retrieved from http://www.aihw.gov.au/WorkArea/DownloadAsset.aspx?id=60129548481

Ayturk, U. M., Jacobsen, C. M., Christodoulou, D. C., Gorham, J., Seidman, J. G., Seidman, C. E., … Warman, M. L. (2013). An RNA-seq protocol to identify mRNA expression changes in mouse diaphyseal bone: Applications in mice with bone property altering Lrp5 mutations. *Journal of Bone and Mineral Research*, *28*(10), 2081–2093. https://doi.org/10.1002/jbmr.1946

Balemans, W., Ebeling, M., Patel, N., Van Hul, E., Olson, P., Dioszegi, M., … Van Hul, W. (2001). Increased bone density in sclerosteosis is due to the deficiency of a novel

secreted protein (SOST). *Human Molecular Genetics*, *10*(5), 537–543. https://doi.org/Doi 10.1093/Hmg/10.5.537

Balemans, W., Patel, N., Ebeling, M., Van Hul, E., Wuyts, W., Lacza, C., … Van Hul, W. (2002). Identification of a 52 kb deletion downstream of the SOST gene in patients with van Buchem disease. *J Med Genet*, *39*, 91–97.

Baron, R., & Hesse, E. (2012). Update on bone anabolics in osteoporosis treatment: Rationale, current status, and perspectives. *Journal of Clinical Endocrinology and Metabolism*, *97*(2), 311–325. https://doi.org/10.1210/jc.2011-2332

Barry, G. (2014). Integrating the roles of long and small non-coding RNA in brain function and disease. *Molecular Psychiatry*, *19*(4), 410–416. https://doi.org/10.1038/mp.2013.196

Bassett, J. H. D., Gogakos, A., White, J. K., Evans, H., Jacques, R. M., van der Spek, A. H., … Williams, G. R. (2012). Rapid-Throughput Skeletal Phenotyping of 100 Knockout Mice Identifies 9 New Genes That Determine Bone Strength. *PLoS Genetics*, *8*(8). https://doi.org/10.1371/journal.pgen.1002858

Batista, P. J., & Chang, H. Y. (2013). Long noncoding RNAs: Cellular address codes in development and disease. *Cell*, *152*(6), 1298–1307. https://doi.org/10.1016/j.cell.2013.02.012

Baxter-Jones, A. D. G., Faulkner, R. A., Forwood, M. R., Mirwald, R. L., & Bailey, D. A. (2011). Bone mineral accrual from 8 to 30 years of age: An estimation of peak bone mass. *Journal of Bone and Mineral Research*, *26*(8), 1729–1739. https://doi.org/10.1002/jbmr.412

Bélanger, L. F., Jarry, L., & Uhthoff, H. K. (1968). Osteocytic osteolysis in Paget's disease. *Revue Canadienne de Biologie*, *27*(1), 37–44.

Bellido, T. (2015). Osteocyte-Driven Bone Remodeling. *Calcified Tissue International*, *94*(1), 25–34. https://doi.org/10.1007/s00223-013-9774-y.Osteocyte-Driven

Bellino, F. L. (2000). Nonprimate animal models of menopause: Workshop report. *Menopause*. https://doi.org/10.1097/00042192-200007010-00004

Benaglia, T., Chauveau, D., Hunter, D., & Young, D. (2009). mixtools: An R Package for Analyzing Finite Mixture Models. *Journal of Statistical Software*, *32*(6), 1–29. https://doi.org/10.18637/jss.v032.i06

Benredjem, B., Girard, M., Rhainds, D., St-Onge, G., & Heveker, N. (2017). Mutational analysis of atypical chemokine receptor 3 (ACKR3/CXCR7) interaction with its chemokine ligands CXCL11 and CXCL12. *Journal of Biological Chemistry*, *292*(1), 31–42. https://doi.org/10.1074/jbc.M116.762252

Blackwell, K. A., Raisz, L. G., & Pilbeam, C. C. (2010). Prostaglandins in bone: Bad cop, good cop? *Trends in Endocrinology and Metabolism*, *21*(5), 294–301. https://doi.org/10.1016/j.tem.2009.12.004

Bloch, S. L., Kristensen, S. L., & Sørensen, M. S. (2012). The Viability of Perilabyrinthine Osteocytes: A Quantitative Study Using Bulk-Stained Undecalcified Human Temporal Bones. *Anatomical Record*, *295*(7), 1101–1108. https://doi.org/10.1002/ar.22492

Blouin, S., Fratzl-Zelman, N., Glorieux, F. H., Roschger, P., Klaushofer, K., Marini, J. C., & Rauch, F. (2017). Hypermineralization and High Osteocyte Lacunar Density in Osteogenesis Imperfecta Type V Bone Indicate Exuberant Primary Bone Formation. *Journal of Bone and Mineral Research*, *32*(9), 1884–1892. https://doi.org/10.1002/jbmr.3180

Bonafe, L., Cormier-Daire, V., Hall, C., Lachman, R., Mortier, G., Mundlos, S., … Unger, S. (2015). Nosology and classification of genetic skeletal disorders: 2015 revision. *American Journal of Medical Genetics, Part A*, *167*(12), 2869–2892.

https://doi.org/10.1002/ajmg.a.37365
Bonewald, L. F., & Johnson, M. L. (2008). Osteocytes, mechanosensing and Wnt signaling. *Bone*, *42*(4), 606–615. https://doi.org/10.1016/j.bone.2007.12.224
Boyce, B. F., & Xing, L. (2007). The RANKL/RANK/OPG pathway. *Current Osteoporosis Reports*, *5*(3), 98–104. https://doi.org/10.1007/s11914-007-0024-y
Breschi, A., Gingeras, T. R., & Guigó, R. (2017). Comparative transcriptomics in human and mouse. *Nature Reviews Genetics*, *18*(7), 425–440. https://doi.org/10.1038/nrg.2017.19
Brodt, M. D., Ellis, C. B., & Silva, M. J. (1999). Growing C57Bl / 6 Mice Increase Whole Bone Mechanical Properties by Increasing Geometric and Material Properties. *Journal of Bone and Mineral Research*, *14*(12).
Brotto, M., & Bonewald, L. (2015). Bone and muscle: Interactions beyond mechanical. *Bone*, *80*, 109–114. https://doi.org/10.1016/j.bone.2015.02.010
Brunkow, M. E., Gardner, J. C., Van Ness, J., Paeper, B. W., Kovacevich, B. R., Proll, S., … Mulligan, J. (2001). Bone dysplasia sclerosteosis results from loss of the SOST gene product, a novel cystine knot-containing protein. *American Journal of Human Genetics*, *68*(3), 577–589. https://doi.org/10.1086/318811
Buenzli, P. R., & Sims, N. A. (2015). Quantifying the osteocyte network in the human skeleton. *Bone*, *75*, 144–150. https://doi.org/10.1016/j.bone.2015.02.016
Cabahug-Zuckerman, P., Frikha-Benayed, D., Majeska, R. J., Tuthill, A., Yakar, S., Judex, S., & Schaffler, M. B. (2016). Osteocyte Apoptosis Caused by Hindlimb Unloading is Required to Trigger Osteocyte RANKL Production and Subsequent Resorption of Cortical and Trabecular Bone in Mice Femurs. *Journal of Bone and Mineral Research*, *31*(7), 1356–1365. https://doi.org/10.1002/jbmr.2807
Capdevila, J., Tsukui, T., Esteban, C. R., Zappavigna, V., & Belmonte, J. C. I. (1999). Control of Vertebrate Limb Outgrowth by the Proximal Factor Meis2 and Distal Antagonism of BMPs by Gremlin. *Molecular Cell*, *4*(5), 839–849. https://doi.org/10.1016/S1097-2765(00)80393-7
Carbon, S., Dietze, H., Lewis, S. E., Mungall, C. J., Munoz-Torres, M. C., Basu, S., … Westerfield, M. (2017). Expansion of the gene ontology knowledgebase and resources: The gene ontology consortium. *Nucleic Acids Research*, *45*(D1), D331–D338. https://doi.org/10.1093/nar/gkw1108
Cardoso, L., Herman, B. C., Verborgt, O., Laudier, D., Majeska, R. J., & Schaffler, M. B. (2009). Osteocyte Apoptosis Controls Activation of Intracortical Resorption in Response to Bone Fatigue. *Journal of Bone and Mineral Research*, *24*(4), 597–605. https://doi.org/10.1359/jbmr.081210
Center, J. R., Nguyen, T. V, Schneider, D. L., Sambrook, P. N., & Eisman, J. (1999). Mortality after all major types of osteoporotic fracture in men and women : an observational study. *The Lancet*, *353*, 878–882.
Chen, G., Wang, Z., Wang, D., Qiu, C., Liu, M., Chen, X., … Cui, Q. (2013). LncRNADisease: A database for long-non-coding RNA-associated diseases. *Nucleic Acids Research*, *41*(D1), 983–986. https://doi.org/10.1093/nar/gks1099
Chen, H., & Boutros, P. C. (2011). VennDiagram: A package for the generation of highly-customizable Venn and Euler diagrams in R. *BMC Bioinformatics*, *12*. https://doi.org/10.1186/1471-2105-12-35
Cheng, B., Zhao, S., Luo, L. F., Sprague, E., Bonewald, L. F., & Jiang, J. X. (2001). Expression of functional gap junctions and regulation by fluid flow in osteocyte-like MLO-Y4 cells. *Journal of Bone and Mineral Research*, *16*(2), 249–259. https://doi.org/10.1359/jbmr.2001.16.2.249

Cheung, W. Y., Fritton, J. C., Morgan, S. A., Seref-Ferlengez, Z., Basta-Pljakic, J., Thi, M. M., … Schaffler, M. B. (2016). Pannexin-1 and P2X7-Receptor Are Required for Apoptotic Osteocytes in Fatigued Bone to Trigger RANKL Production in Neighboring Bystander Osteocytes. *Journal of Bone and Mineral Research*, *31*(4), 890–899. https://doi.org/10.1002/jbmr.2740

Civitelli, R., & Ziambaras, K. (2011). Calcium and phosphate homeostasis: concerted interplay of new regulators. *Journal of Endocrinological Investigation*, *34*(7 Suppl), 3–7. https://doi.org/10.1080/07853890701689645

Conesa, A., Madrigal, P., Tarazona, S., Gomez-Cabrero, D., Cervera, A., McPherson, A., … Mortazavi, A. (2016). A survey of best practices for RNA-seq data analysis. *Genome Biology*, *17*(1), 1–19. https://doi.org/10.1186/s13059-016-0881-8

Cosman, F., Crittenden, D. B., Adachi, J. D., Binkley, N., Czerwinski, E., Ferrari, S., … Grauer, A. (2016). Romosozumab Treatment in Postmenopausal Women with Osteoporosis. *New England Journal of Medicine*, *375*(16), 1532–1543. https://doi.org/10.1056/NEJMoa1607948

Dallas, S. L., & Bonewald, L. F. (2010). Dynamics of the transition from osteoblast to osteocyte. *Annals of the New York Academy of Sciences*, *1192*, 437–443. https://doi.org/10.1111/j.1749-6632.2009.05246.x

Dallas, S. L., Prideaux, M., & Bonewald, L. F. (2013). The osteocyte: An endocrine cell . . . and more. *Endocrine Reviews*, *34*(5), 658–690. https://doi.org/10.1210/er.2012-1026

De Vega, S., Iwamoto, T., & Yamada, Y. (2009). Fibulins: Multiple roles in matrix structures and tissue functions. *Cellular and Molecular Life Sciences*, *66*(11–12), 1890–1902. https://doi.org/10.1007/s00018-009-8632-6

DeLaurier, A., Schweitzer, R., & Logan, M. (2006). Pitx1 determines the morphology of muscle, tendon, and bones of the hindlimb. *Developmental Biology*, *299*(1), 22–34. https://doi.org/10.1016/j.ydbio.2006.06.055

Delorme, G., Saltel, F., Bonnelye, E., Jurdic, P., & Machuca-Gayet, I. (2005). Expression and function of semaphorin 7A in bone cells. *Biology of the Cell / under the Auspices of the European Cell Biology Organization*, *97*(7), 589–597. https://doi.org/10.1042/BC20040103

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., … Gingeras, T. R. (2013). STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics*, *29*(1), 15–21. https://doi.org/10.1093/bioinformatics/bts635

Dodd, J., & Jessell, T. M. (1988). Axon guidance and the patterning of neuronal projections in vertebrates. *Science*. https://doi.org/10.1126/science.3055291

Dole, N. S., Mazur, C. M., Acevedo, C., Lopez, J. P., Monteiro, D. A., Fowler, T. W., … Alliston, T. (2017). Osteocyte-Intrinsic TGF-β Signaling Regulates Bone Quality through Perilacunar/Canalicular Remodeling. *Cell Reports*, *21*(9), 2585–2596. https://doi.org/10.1016/j.celrep.2017.10.115

Doty, S. B. (1981). Morphological evidence of gap junctions between bone cells. *Calcified Tissue International*, *33*(1), 509–512. https://doi.org/10.1007/BF02409482

Dufort, C. C., Paszek, M. J., & Weaver, V. M. (2011). Balancing forces: Architectural control of mechanotransduction. *Nature Reviews Molecular Cell Biology*, *12*(5), 308–319. https://doi.org/10.1038/nrm3112

Dutta, S., & Sengupta, P. (2016). Men and mice: Relating their ages. *Life Sciences*, *152*(May 2016), 244–248. https://doi.org/10.1016/j.lfs.2015.10.025

Elms, P., Siggers, P., Napper, D., Greenfield, A., & Arkell, R. (2003). Zic2 is required for neural crest formation and hindbrain patterning during mouse development. *Developmental Biology*, *264*(2), 391–406. https://doi.org/10.1016/j.ydbio.2003.09.005

Eshkar-Oren, I., Viukov, S. V., Salameh, S., Krief, S., Oh, C. -d., Akiyama, H., … Zelzer, E. (2009). The forming limb skeleton serves as a signaling center for limb vasculature patterning via regulation of Vegf. *Development*, *136*(8), 1263–1272. https://doi.org/10.1242/dev.034199

Fedarko, N. S., Vetter, U. K., Weinstein, S., & Robey, P. G. (1992). Age-related changes in hyaluronan, proteoglycan, collagen, and osteonectin synthesis by human bone cells. *Journal of Cellular Physiology*, *151*(2), 215–227.

Feldstein, A., Elmer, P. J., Orwoll, E., Herson, M., & Hillier, T. (2003). Bone Mineral Density Measurement and Treatment for Osteoporosis in Older Individuals With Fractures: A Gap in Evidence-Based Practice Guideline Implementation. *Arch Intern Med*, *163*(18), 2165–2172. https://doi.org/10.1001/archinte.163.18.2165

Feng, J. Q., Ward, L. M., Liu, S., Lu, Y., Xie, Y., Yuan, B., … White, K. E. (2006). Loss of DMP1 causes rickets and osteomalacia and identifies a role for osteocytes in mineral metabolism. *Nature Genetics*, *38*(11), 1310–1315. https://doi.org/10.1038/ng1905

Ferrari, S., Rizzoli, R., Slosman, D., & Bonjour, J. P. (1998). Familial resemblance for bone mineral mass is expressed before puberty. *Journal of Clinical Endocrinology and Metabolism*, *83*(2), 358–361. https://doi.org/10.1210/jc.83.2.358

Finucane, H. K., Reshef, Y. A., Anttila, V., Slowikowski, K., Gusev, A., Byrnes, A., … Price, A. L. (2018). Heritability enrichment of specifically expressed genes identifies disease-relevant tissues and cell types. *Nature Genetics*, *50*(4), 621–629. https://doi.org/10.1038/s41588-018-0081-4

Fontaine, D. A., & Davis, D. B. (2016). Attention to background strain is essential for metabolic research: C57BL/6 and the international knockout mouse consortium. *Diabetes*, *65*(1), 25–33. https://doi.org/10.2337/db15-0982

Forlino, A., Cabral, W. A., Barnes, A. M., & Marini, J. C. (2011). New perspectives on osteogenesis imperfecta. *Nature Reviews Endocrinology*, *7*(9), 540–557. https://doi.org/10.1038/nrendo.2011.81

Forlino, A., & Marini, J. C. (2016). Osteogenesis imperfecta. *The Lancet*, *387*(10028), 1657–1671. https://doi.org/10.1016/S0140-6736(15)00728-X

Fowler, T. W., Acevedo, C., Mazur, C. M., Hall-Glenn, F., Fields, A. J., Bale, H. A., … Alliston, T. (2017). Glucocorticoid suppression of osteocyte perilacunar remodeling is associated with subchondral bone degeneration in osteonecrosis. *Scientific Reports*, *7*(November 2016), 1–13. https://doi.org/10.1038/srep44618

Fraley, C., & Raftery, A. E. (2003). Enhanced Model-Based Clustering, Density Estimation, and Discriminant Analysis Software: MCLUST. *Journal of Classification*, *20*(2), 263–286. https://doi.org/10.1007/s00357-003-0015-3

Franz-Odendaal, T. A., Hall, B. K., & Witten, P. E. (2006). Buried alive: How osteoblasts become osteocytes. *Developmental Dynamics*, *235*(1), 176–190. https://doi.org/10.1002/dvdy.20603

Frost, H. M. (1960). In vivo osteocyte death. *Journal of Bone and Joint Surgery.* Boston:

Fu, Z., Lee, S. H., Simonetta, A., Hansen, J., Sheng, M., & Pak, D. T. S. (2007). Differential roles of Rap1 and Rap2 small GTPases in neurite retraction and synapse elimination in hippocampal spiny neurons. *Journal of Neurochemistry*, *100*(1), 118–131. https://doi.org/10.1111/j.1471-4159.2006.04195.x

Fulciniti, M., Tassone, P., Hideshima, T., Vallet, S., Nanjappa, P., Ettenberg, S. A., … Munshi, N. C. (2009). Anti-DKK1 mAb (BHQ880) as a potential therapeutic agent for multiple myeloma. *Blood*, *114*(2), 371–379. https://doi.org/10.1182/blood-2008-11-191577

Fulzele, K., Krause, D. S., Panaroni, C., Saini, V., Barry, K. J., Liu, X. L., … Pajevic, P. D.

(2013). Myelopoiesis is regulated by osteocytes through Gs alpha-dependent signaling. *Blood*, *121*(6), 930–939. https://doi.org/DOI 10.1182/blood-2012-06-437160

Gautier, L., Cope, L., Bolstad, B. M., & Irizarry, R. A. (2004). Affy - Analysis of Affymetrix GeneChip data at the probe level. *Bioinformatics*, *20*(3), 307–315. https://doi.org/10.1093/bioinformatics/btg405

Gegenbaur, C. (1864). *Untersuchungen zur vergleichenden Anatomie der Wirbelthiere* (Vol. 1). W. Engelmann.

Gerrits, H., Van Ingen Schenau, D. S., Bakker, N. E. C., Van Disseldorp, A. J. M., Strik, A., Hermens, L. S., … Gossen, J. A. (2008). Early postnatal lethality and cardiovascular defects in CXCR7-deficient mice. *Genesis*, *46*(5), 235–245. https://doi.org/10.1002/dvg.20387

Gluhak-Heinrich, J., Ye, L., Bonewald, L. F., Feng, J. Q., MacDougall, M., Harris, S. E., & Pavlin, D. (2003). Mechanical Loading Stimulates Dentin Matrix Protein 1 (DMP1) Expression in Osteocytes In Vivo. *Journal of Bone and Mineral Research*. https://doi.org/10.1359/jbmr.2003.18.5.807

Gordon, C. M., Zemel, B. S., Wren, T. A. L., Leonard, M. B., Bachrach, L. K., Rauch, F., … Winer, K. K. (2017). The Determinants of Peak Bone Mass. *Journal of Pediatrics*, *180*, 261–269. https://doi.org/10.1016/j.jpeds.2016.09.056

Gowen, L. C., Petersen, D. N., Mansolf, A. L., Qi, H., Stock, J. L., Tkalcevic, G. T., … Brown, T. A. (2003). Targeted disruption of the osteoblast/osteocyte factor 45 gene (Of45) results in increased bone formation and bone mass. *Journal of Biological Chemistry*, *278*(3), 1998–2007. https://doi.org/10.1074/jbc.M203250200

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., … Regev, A. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology*, *29*(7), 644–652. https://doi.org/10.1038/nbt.1883

Grafe, I., Yang, T., Alexander, S., Homan, E. P., Lietman, C., Jiang, M. M., … Lee, B. (2014). Excessive transforming growth factor-β signaling is a common mechanism in osteogenesis imperfecta. *Nature Medicine*, *20*(6), 670–675. https://doi.org/10.1038/nm.3544

Grinberg, I., Northrup, H., Ardinger, H., Prasad, C., Dobyns, W. B., & Millen, K. J. (2004). Heterozygous deletion of the linked genes ZIC1 and ZIC4 is involved in Dandy-Walker malformation. *Nature Genetics*, *36*(10), 1053–1055. https://doi.org/10.1038/ng1420

Gu, G., Nars, M., Hentunen, T. A., Metsikkö, K., & Väänänen, H. K. (2006). Isolated primary osteocytes express functional gap junctions in vitro. *Cell and Tissue Research*, *323*(2), 263–271. https://doi.org/10.1007/s00441-005-0066-3

Haas, B. J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P. D., Bowden, J., … Regev, A. (2013). De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nature Protocols*, *8*(8), 1494–1512. https://doi.org/10.1038/nprot.2013.084

Hagenaars, S. P., Harris, S. E., Davies, G., Hill, W. D., Liewald, D. C. M., Ritchie, S. J., … Deary, I. J. (2016). Shared genetic aetiology between cognitive functions and physical and mental health in UK Biobank (N=112 151) and 24 GWAS consortia. *Molecular Psychiatry*, *21*(11), 1624–1632. https://doi.org/10.1038/mp.2015.225

Hahne, F., & Ivanek, R. (2016). Visualizing genomic data using Gviz and bioconductor. In *Statistical Genomics* (pp. 335–351). Springer.

Haller, A. C., & Zimny, M. L. (1977). Effects of Hibernation on Interradicular Alveolar Bone. *Journal of Dental Research*, *56*(12), 1552–1557. https://doi.org/10.1177/00220345770560122601

Hamada, H., Meno, C., Watanabe, D., & Saijoh, Y. (2002). Establishment of vertebrate left-right asymmetry. *Nature Reviews Genetics*. https://doi.org/10.1038/nrg732

Han, Q., Pabba, P. K., Barbosa, J., Mabon, R., Healy, J. P., Gardyan, M. W., … Carson, K. G. (2016). 4H-Thieno[3,2-c]chromene based inhibitors of Notum Pectinacetylesterase. *Bioorganic and Medicinal Chemistry Letters*, *26*(4), 1184–1187. https://doi.org/10.1016/j.bmcl.2016.01.038

Han, Y., Cowin, S. C., Schaffler, M. B., & Weinbaum, S. (2004). Mechanotransduction and strain amplification in osteocyte cell processes. *Proceedings of the National Academy of Sciences*, *101*(47), 16689–16694. https://doi.org/10.1073/pnas.0407429101

Harris, S. E., Gluhak-Heinrich, J., Harris, M. A., Yang, W., Bonewald, L. F., Riha, D., … Nicollela, D. (2007). DMP1 and MEPE expression are elevated in osteocytes after mechanical loading in vivo: theoretical role in controlling mineral quality in the perilacunar matrix. *Journal of Musculoskeletal & Neuronal Interactions*. https://doi.org/10.1016/j.biotechadv.2011.08.021.Secreted

Hart, T., Komori, H. K., LaMere, S., Podshivalova, K., & Salomon, D. R. (2013). Finding the active genes in deep RNA-seq gene expression studies. *BMC Genomics*, *14*(1). https://doi.org/10.1186/1471-2164-14-778

Hassan, M. Q., Tye, C. E., Stein, G. S., & Lian, J. B. (2015). Non-coding RNAs : Epigenetic regulators of bone development and homeostasis. *Bone*, *81*, 746–756. https://doi.org/10.1016/j.bone.2015.05.026

Hayashi, M., Nakashima, T., Taniguchi, M., Kodama, T., Kumanogoh, A., & Takayanagi, H. (2012). Osteoprotection by semaphorin 3A. *Nature*, *485*(7396), 69–74. https://doi.org/10.1038/nature11000

Hayman, A. R., Jones, S. J., Boyde, A., Foster, D., Colledge, W. H., Carlton, M. B., … Cox, T. M. (1996). Mice lacking tartrate-resistant acid phosphatase (Acp 5) have disrupted endochondral ossification and mild osteopetrosis. *Development*, *3162*, 3151–3162.

Hernlund, E., Svedbom, A., Ivergård, M., Compston, J., Cooper, C., Stenmark, J., … Kanis, J. A. (2013). Osteoporosis in the European Union: Medical management, epidemiology and economic burden: A report prepared in collaboration with the International Osteoporosis Foundation (IOF) and the European Federation of Pharmaceutical Industry Associations (EFPIA). *Archives of Osteoporosis*, *8*(1–2). https://doi.org/10.1007/s11657-013-0136-1

Hertzberg, L., Katsel, P., Roussos, P., Haroutunian, V., & Domany, E. (2015). Integration of gene expression and GWAS results supports involvement of calcium signaling in Schizophrenia. *Schizophrenia Research*, *164*(1–3), 92–99. https://doi.org/10.1016/j.schres.2015.02.001

Hesse, E. (2016). Muscle and Bone: Combating the Evil Side of the Connection. *Journal of Bone and Mineral Research*, *31*(9), 1647–1651. https://doi.org/10.1002/jbmr.2912

Hiasa, M., Okui, T., Allette, Y. M., Ripsch, M. S., Sun-Wada, G. H., Wakabayashi, H., … Yoneda, T. (2017). Bone pain induced by multiple myeloma is reduced by targeting V-ATPase and ASIC3. *Cancer Research*, *77*(6), 1283–1295. https://doi.org/10.1158/0008-5472.CAN-15-3545

Hinck, L. (2004). The Versatile Roles of “ Axon Guidance ” Cues in Tissue Morphogenesis, *7*, 783–793.

Hiyama, A., Skubutyte, R., Markova, D., Anderson, D. G., Yadla, S., Sakai, D., … Risbud, M. V. (2011). Hypoxia activates the notch signaling pathway in cells of the intervertebral disc: Implications in degenerative disc disease. *Arthritis and Rheumatism*, *63*(5), 1355–1364. https://doi.org/10.1002/art.30246

Hockenbery, D. M., Zutter, M., Hickey, W., Nahm, M., & Korsmeyer, S. J. (1991). BCL2

protein is topographically restricted in tissues characterized by apoptotic cell death. *Proceedings of the National Academy of Sciences*. https://doi.org/10.1073/pnas.88.16.6961

Hui, S. L., Slemenda, C. W., & Johnston, C. C. (1990). The contribution of bone loss to postmenopausal osteoporosis. *Osteoporosis International*, *1*(1), 30–34. https://doi.org/10.1007/BF01880413

Igarashi, M., Kamiya, N., Ito, K., & Takagi, M. (2002). In situ localization and in vitro expression of osteoblast/osteocyte factor 45 mRNA during bone cell differentiation. *Histochemical Journal*, *34*(5), 255–263. https://doi.org/10.1023/A:1021745614872

Inaoka, T., Bilbe, G., Ishibashi, O., Tezuka, K. I., Kumegawa, M., & Kokubo, T. (1995). Molecular cloning of human cDNA for cathepsin K: Novel cysteine proteinase predominantly expressed in bone. *Biochemical and Biophysical Research Communications*. https://doi.org/10.1006/bbrc.1995.1013

Irie, N., Takada, Y., Watanabe, Y., Matsuzaki, Y., Naruse, C., Asano, M., … Matsuo, K. (2009). Bidirectional signaling through EphrinA2-EphA2 enhances osteoclastogenesis and suppresses osteoblastogenesis. *Journal of Biological Chemistry*, *284*(21), 14637–14644. https://doi.org/10.1074/jbc.M807598200

Jähn, K., Kelkar, S., Zhao, H., Xie, Y., Tiede-Lewis, L. A. M., Dusevich, V., … Bonewald, L. F. (2017). Osteocytes Acidify Their Microenvironment in Response to PTHrP In Vitro and in Lactating Mice In Vivo. *Journal of Bone and Mineral Research*, *32*(8), 1761–1772. https://doi.org/10.1002/jbmr.3167

Jahr, H., Van Driel, M., Van Osch, G. J. V. M., Weinans, H., & Van Leeuwen, J. P. T. M. (2005). Identification of acid-sensing ion channels in bone. *Biochemical and Biophysical Research Communications*, *337*(1), 349–354. https://doi.org/10.1016/j.bbrc.2005.09.054

Jalalvand, E., Robertson, B., Wallén, P., & Grillner, S. (2016). Ciliated neurons lining the central canal sense both fluid movement and pH through ASIC3. *Nature Communications*, *7*. https://doi.org/10.1038/ncomms10002

James, J., & Steijn-Myagkaya, G. L. (1986). Death of osteocytes. Electron microscopy after in vitro ischaemia. *The Journal of Bone and Joint Surgery. British Volume*, *68*(4), 620–624. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/3733842

Jiang, X., Iseki, S., Maxson, R. E., Sucov, H. M., & Morriss-Kay, G. M. (2002). Tissue origins and interactions in the mammalian skull vault. *Developmental Biology*, *241*(1), 106–116. https://doi.org/10.1006/dbio.2001.0487

Jiang, X., Tan, J., Li, J., Kivimäe, S., Yang, X., Zhuang, L., … Yu, Q. (2008). DACT3 Is an Epigenetic Regulator of Wnt/β-Catenin Signaling in Colorectal Cancer and Is a Therapeutic Target of Histone Modifications. *Cancer Cell*, *13*(6), 529–541. https://doi.org/10.1016/j.ccr.2008.04.019

Jilka, R. L. (2013). The relevance of mouse models for investigating age-related bone loss in humans. *Journals of Gerontology - Series A Biological Sciences and Medical Sciences*, *68*(10), 1209–1217. https://doi.org/10.1093/gerona/glt046

Joeng, K. S., Lee, Y. C., Lim, J., Chen, Y., Jiang, M. M., Munivez, E., … Lee, B. H. (2017). Osteocyte-specific WNT1 regulates osteoblast function during bone homeostasis. *Journal of Clinical Investigation*, *127*(7), 2678–2688. https://doi.org/10.1172/JCI92617

Kakugawa, S., Langton, P. F., Zebisch, M., Howell, S. A., Chang, T. H., Liu, Y., … Vincent, J. P. (2015). Notum deacylates Wnt proteins to suppress signalling activity. *Nature*, *519*(7542), 187–192. https://doi.org/10.1038/nature14259

Kanis, J. A. (2002). Osteoporosis III: Diagnosis of osteoporosis and assessment of fracture risk. *Lancet*, *359*(9321), 1929–1936. https://doi.org/10.1016/S0140-6736(02)08761-5

Karsenty, G., & Wagner, E. F. (2002). Reaching a genetic and molecular understanding of skeletal evelopment. *Developmental Cell*, *2*(4), 389–406. https://doi.org/10.1016/S1534-5807(02)00157-0

Kaste, S. C., Jones-Wallace, D., Rose, S. R., Boyett, J. M., Lustig, R. H., Rivera, G. K., … Hudson, M. M. (2001). Bone mineral decrements in survivors of childhood acute lymphoblastic leukemia: Frequency of occurrence and risk factors for their development. *Leukemia*, *15*(5), 728–734. https://doi.org/10.1038/sj.leu.2402078

Kelly, N. H., Schimenti, J. C., Patrick Ross, F., & van der Meulen, M. C. H. (2014). A method for isolating high quality RNA from mouse cortical and cancellous bone. *Bone*, *68*, 1–5. https://doi.org/10.1016/j.bone.2014.07.022

Kelly, N. H., Schimenti, J. C., Ross, F. P., & van der Meulen, M. C. H. (2016). Transcriptional profiling of cortical versus cancellous bone from mechanically-loaded murine tibiae reveals differential gene expression. *Bone*, *86*, 22–29. https://doi.org/10.1016/j.bone.2016.02.007

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

Kennedy, O. D., Laudier, D. M., Majeska, R. J., Sun, H. B., & Schaffler, M. B. (2014). Osteocyte apoptosis is required for production of osteoclastogenic signals following bone fatigue in vivo. *Bone*, *64*, 132–137. https://doi.org/10.1016/j.bone.2014.03.049

Kerschnitzki, M., Kollmannsberger, P., Burghammer, M., Duda, G. N., Weinkamer, R., Wagermaier, W., & Fratzl, P. (2013). Architecture of the osteocyte network correlates with bone material quality. *Journal of Bone and Mineral Research*, *28*(8), 1837–1845. https://doi.org/10.1002/jbmr.1927

Kim, B. T., Mosekilde, L., Duan, Y., Zhang, X. Z., Tornvig, L., Thomsen, J. S., & Seeman, E. (2003). The structural and hormonal basis of sex differences in peak appendicular bone strength in rats. *Journal of Bone and Mineral Research*, *18*(1), 150–155. https://doi.org/10.1359/jbmr.2003.18.1.150

Klein-Nulend, J., Bakker, A. D., Bacabac, R. G., Vatsa, A., & Weinbaum, S. (2013). Mechanosensation and transduction in osteocytes. *Bone*, *54*(2), 182–190. https://doi.org/10.1016/j.bone.2012.10.013

Klingberg, A., Hawwari, I., Bornemann, L., Brenzel, A., Müller, S., Lang, S., … Gunzer, M. (2017). A network of trans-cortical capillaries as basis for blood circulation in long bones. *Manuscript Submitted for Publication (Cell)*. https://doi.org/10.1038/s42255-018-0016-5

Kmita, M., Tarchini, B., Zàkàny, J., Logan, M., Tabin, C. J., & Duboule, D. (2005). Early developmental arrest of mammalian limbs lacking HoxA/HoxD gene function. *Nature*, *435*(7045), 1113–1116. https://doi.org/10.1038/nature03648

Kogawa, M., Wijenayaka, A. R., Ormsby, R. T., Thomas, G. P., Anderson, P. H., Bonewald, L. F., … Atkins, G. J. (2013). Sclerostin regulates release of bone mineral by osteocytes by induction of carbonic anhydrase 2. *Journal of Bone and Mineral Research*. https://doi.org/10.1002/jbmr.2003

Koh, J. M., Oh, B., Lee, J. Y., Lee, J. K., Kimm, K., Kim, G. S., … Kim, S. Y. (2006). Association study of semaphorin 7a (sema7a) polymorphisms with bone mineral density and fracture risk in postmenopausal Korean women. *Journal of Human Genetics*. https://doi.org/10.1007/s10038-005-0331-z

Kollmannsberger, P., Kerschnitzki, M., Repp, F., Wagermaier, W., Weinkamer, R., & Fratzl, P. (2017). The small world of osteocytes: Connectomics of the lacuno-canalicular

network in bone. *New Journal of Physics*, *19*(7). https://doi.org/10.1088/1367-2630/aa764b

Krueger, F. (2015). Trim Galore! trim_galore_v0.4.0. *A Wrapper Tool around Cutadapt and FastQC to Consistently Apply Quality and Adapter Trimming to FastQ Files*. Retrieved from http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/

Kumar, G., & Narayan, B. (2014). The biology of fracture healing in long bones. *Classic Papers in Orthopaedics*, 531–533. https://doi.org/10.1007/978-1-4471-5451-8_139

Land, C., & Schoenau, E. (2008). Fetal and postnatal bone development: reviewing the role of mechanical stimuli and nutrition. *Best Practice and Research in Clinical Endocrinology and Metabolism*, *22*(1), 107–118. https://doi.org/10.1016/j.beem.2007.09.005

Lane, J. M., Vlasac, I., Anderson, S. G., Kyle, S. D., Dixon, W. G., Bechtold, D. A., … Saxena, R. (2016). Genome-wide association analysis identifies novel loci for chronotype in 100,420 individuals from the UK Biobank. *Nature Communications*, *7*, 1–10. https://doi.org/10.1038/ncomms10889

Lane, N. E., Yao, W., Balooch, M., Nalla, R. K., Balooch, G., Habelitz, S., … Bonewald, L. F. (2006). Glucocorticoid-treated mice have localized changes in trabecular bone material properties and osteocyte lacunar size that are not observed in placebo-treated or estrogen-deficient mice. *Journal of Bone and Mineral Research*, *21*(3), 466–476. https://doi.org/10.1359/JBMR.051103

Lanyon, L. E. (1993). Osteocytes, strain detection, bone modeling and remodeling. *Calcified Tissue International*, *53*(S1), S102–S107. https://doi.org/10.1007/BF01673415

Laurent, M. R., Dubois, V., Claessens, F., Verschueren, S. M. P., Vanderschueren, D., Gielen, E., & Jardí, F. (2016). Muscle-bone interactions: From experimental models to the clinic? A critical update. *Molecular and Cellular Endocrinology*, *432*, 14–36. https://doi.org/10.1016/j.mce.2015.10.017

Lavi-Moshayoff, V., Wasserman, G., Meir, T., Silver, J., & Naveh-Many, T. (2010). PTH increases FGF23 gene expression and mediates the high-FGF23 levels of experimental kidney failure: a bone parathyroid feedback loop. *AJP: Renal Physiology*, *299*(4), F882–F889. https://doi.org/10.1152/ajprenal.00360.2010

Law, C. W., Chen, Y., Shi, W., & Smyth, G. K. (2014). Voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biology*, *15*(2), 1–17. https://doi.org/10.1186/gb-2014-15-2-r29

Lean, J. M., Jagger, C. J., Chambers, T. J., & Chow, J. W. (1995). Increased insulin-like growth factor I mRNA expression in rat osteocytes in response to mechanical stimulation. *American Journal of Physiology.* Bethesda, MD: https://doi.org/10.1152/ajpendo.1995.268.2.E318

Lee, H. K., & Deneen, B. (2012). Daam2 Is Required for Dorsal Patterning via Modulation of Canonical Wnt Signaling in the Developing Spinal Cord. *Developmental Cell*, *22*(1), 183–196. https://doi.org/10.1016/j.devcel.2011.10.025

Leucht, P., Kim, J.-B., Amasha, R., James, A. W., Girod, S., & Helms, J. A. (2008). Embryonic origin and Hox status determine progenitor cell fate during adult bone regeneration. *Development*, *135*(17), 2845–2854. https://doi.org/10.1242/dev.023788

Li, B., & Dewey, C. N. (2011). RSEM: Accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics*, *12*. https://doi.org/10.1186/1471-2105-12-323

Lian, J. B., Stein, G. S., Van Wijnen, A. J., Stein, J. L., Hassan, M. Q., Gaur, T., & Zhang, Y. (2012). MicroRNA control of bone formation and homeostasis. *Nature Reviews Endocrinology*, *8*(4), 212–227. https://doi.org/10.1038/nrendo.2011.234

Liang, W. C., Fu, W. M., Wang, Y. B., Sun, Y. X., Xu, L. L., Wong, C. W., … Zhang, J. F. (2016). H19 activates Wnt signaling and promotes osteoblast differentiation by functioning as a competing endogenous RNA. *Scientific Reports*, *6*(February), 1–11. https://doi.org/10.1038/srep20121

Liao, B.-Y. (2005). Evolutionary Conservation of Expression Profiles Between Human and Mouse Orthologous Genes. *Molecular Biology and Evolution*, *23*(3), 530–540. https://doi.org/10.1093/molbev/msj054

Limaye, A. (2012). Drishti: a volume exploration and presentation tool, *8506*, 85060X. https://doi.org/10.1117/12.935640

Lin, C., Jiang, X., Dai, Z., Guo, X., Weng, T., Wang, J., … He, L. (2009). Sclerostin mediates bone response to mechanical unloading through antagonizing Wnt/β-catenin signaling. *Journal of Bone and Mineral Research*, *24*(10), 1651–1661. https://doi.org/10.1359/jbmr.090411

Lin, S. H., Cheng, Y. R., Banks, R. W., Min, M. Y., Bewick, G. S., & Chen, C. C. (2016). Evidence for the involvement of ASIC3 in sensory mechanotransduction in proprioceptors. *Nature Communications*, *7*(May), 1–15. https://doi.org/10.1038/ncomms11460

Liu, S., Rowe, P. S. N., Vierthaler, L., Zhou, J., & Quarles, L. D. (2007). Phosphorylated acidic serine-aspartate-rich MEPE-associated motif peptide from matrix extracellular phosphoglycoprotein inhibits phosphate regulating gene with homologies to endopeptidases on the X-chromosome enzyme activity. *Journal of Endocrinology*, *192*(1), 261–267. https://doi.org/10.1677/joe.1.07059

Liu, S., Zhou, J., Tang, W., Menard, R., Feng, J. Q., & Quarles, L. D. (2008). Pathogenic role of Fgf23 in *Dmp1* -null mice. *American Journal of Physiology-Endocrinology and Metabolism*, *295*(2), E254–E261. https://doi.org/10.1152/ajpendo.90201.2008

Looker, A. C., & Frenk, S. M. (2015). Percentage of Adults Aged 65 and Over With Osteoporosis or Low Bone Mass at the Femur Neck or Lumbar Spine: United States, 2005–2010. *Centers for Disease Control and Prevention*, (August), 2005–2010.

Luo, W., & Brouwer, C. (2013). Pathview: An R/Bioconductor package for pathway-based data integration and visualization. *Bioinformatics*, *29*(14), 1830–1831. https://doi.org/10.1093/bioinformatics/btt285

MacArthur, J., Bowler, E., Cerezo, M., Gil, L., Hall, P., Hastings, E., … Parkinson, H. (2017). The new NHGRI-EBI Catalog of published genome-wide association studies (GWAS Catalog). *Nucleic Acids Research*, *45*(D1), D896–D901. https://doi.org/10.1093/nar/gkw1133

MacDonald, J. W. (2008). affycoretools: Functions useful for those doing repetitive analyses with Affymetrix GeneChips. *R Package Version*, *1*(0).

Mach, D. B., Rogers, S. D., Sabino, M. C., Luger, N. M., Schwei, M. J., Pomonis, J. D., … Mantyh, P. W. (2002). Origins of skeletal pain: Sensory and sympathetic innervation of the mouse femur. *Neuroscience*. https://doi.org/10.1016/S0306-4522(02)00165-3

Maes, C., Kobayashi, T., Selig, M. K., Torrekens, S., Roth, S. I., Mackem, S., … Kronenberg, H. M. (2010). Osteoblast precursors, but not mature osteoblasts, move into developing and fractured bones along with invading blood vessels. *Developmental Cell*, *19*(2), 329–344. https://doi.org/10.1016/j.devcel.2010.07.010

Marie, P. J., & Glorieux, F. H. (1983). Relation between hypomineralized periosteocytic lesions and bone mineralization in vitamin D-resistant rickets. *Calcified Tissue International*, *35*(1), 443–448. https://doi.org/10.1007/BF02405074

Matsuo, K., & Otaki, N. (2012). Bone cell interactions through Eph/ephrin: Bone modeling, remodeling and associated diseases. *Cell Adhesion and Migration*, *6*(2), 148–156.

https://doi.org/10.4161/cam.20888
Mautner, M. N. (1997). Directed panspermia. 3. strategies and motivation for seeding star-forming clouds. *JBIS - Journal of the British Interplanetary Society*, *50*(3), 93–102. https://doi.org/10.1210/me.2009-0085
McClung, M. R., Grauer, A., Boonen, S., Bolognese, M. A., Brown, J. P., Diez-Perez, A., … Bone, H. G. (2014). Romosozumab in Postmenopausal Women with Low Bone Mineral Density. *New England Journal of Medicine*, *370*(5), 412–420. https://doi.org/10.1056/NEJMoa1305224
Medina-martı, O., & Bradley, A. (2000). A Large Targeted Deletion of Hoxb1 – Hoxb9 Produces a Series of Single-Segment Anterior Homeotic Transformations. *Developmental Biology*, *83*, 71–83. https://doi.org/10.1006/dbio.2000.9683
Ménager, C., Arimura, N., Fukata, Y., & Kaibuchi, K. (2004). PIP3is involved in neuronal polarization and axon formation. *Journal of Neurochemistry*, *89*(1), 109–118. https://doi.org/10.1046/j.1471-4159.2004.02302.x
Mendoza-Londono, R., Fahiminiya, S., Majewski, J., Tétreault, M., Nadaf, J., Kannu, P., … Rauch, F. (2015). Recessive Osteogenesis Imperfecta Caused by Missense Mutations in SPARC. *American Journal of Human Genetics*, *96*(6), 979–985. https://doi.org/10.1016/j.ajhg.2015.04.021
Mercer, T. R., Dinger, M. E., Sunkin, S. M., Mehler, M. F., & Mattick, J. S. (2008). Specific expression of long noncoding RNAs in the mouse brain. *Proceedings of the National Academy of Sciences*, *105*(2), 716–721. https://doi.org/10.1073/pnas.0706729105
Meyer, D., Hornik, K., & Feinerer, I. (2008). Text mining infrastructure in R. *Journal of Statistical Software*, *25*(5), 1–54.
Miller, E. J., Martin, G. R., Piez, K. A., & Powers, M. J. (1967). Characterization of chick bone collagen and compositional changes associated with maturation. *Journal of Biological Chemistry*, *242*(23), 5481–5489. https://doi.org/74
Milovanovic, P., Zimmermann, E. A., Riedel, C., Scheidt, A. vom, Herzog, L., Krause, M., … Busse, B. (2015). Multi-level characterization of human femoral cortices and their underlying osteocyte network reveal trends in quality of young, aged, osteoporotic and antiresorptive-treated bone. *Biomaterials*, *45*, 46–55. https://doi.org/10.1016/j.biomaterials.2014.12.024
Miwa, H. E., Gerken, T. A., Jamison, O., & Tabak, L. A. (2010). Isoform-specific O-glycosylation of osteopontin and bone sialoprotein by polypeptide N-acetylgalactosaminyltransferase-1. *Journal of Biological Chemistry*, *285*(2), 1208–1219. https://doi.org/10.1074/jbc.M109.035436
Mondal, T., Subhash, S., Vaid, R., Enroth, S., Uday, S., Reinius, B., … Kanduri, C. (2015). MEG3 long noncoding RNA regulates the TGF-β pathway genes through formation of RNA-DNA triplex structures. *Nature Communications*, *6*. https://doi.org/10.1038/ncomms8743
Montagnani, A. (2014). Bone anabolics in osteoporosis: Actuality and perspectives. *World Journal of Orthopedics*, *5*(3), 247. https://doi.org/10.5312/wjo.v5.i3.247
Morozova, O., Hirst, M., & Marra, M. A. (2009). Applications of New Sequencing Technologies for Transcriptome Analysis. *Annual Review of Genomics and Human Genetics*, *10*(1), 135–151. https://doi.org/10.1146/annurev-genom-082908-145957
Mudge, J. M., & Harrow, J. (2015). Creating reference gene annotation for the mouse C57BL6/J genome assembly. *Mammalian Genome*, *26*(9–10), 366–378. https://doi.org/10.1007/s00335-015-9583-x
Nagalakshmi, U., Wang, Z., Waern, K., Shou, C., Raha, D., Gerstein, M., & Snyder, M. (2008). The transcriptional landscape of the yeast genome defined by RNA

sequencing. *Science*. https://doi.org/10.1126/science.1158441

Nakano, N., Maeyama, K., Sakata, N., Itoh, F., Akatsu, R., Nakata, M., … Itoh, S. (2014). C18 orf1, a novel negative regulator of transforming growth factor-β signaling. *Journal of Biological Chemistry*, *289*(18), 12680–12692. https://doi.org/10.1074/jbc.M114.558981

Nan, X., Tate, P., Li, E., Bird, A., Nan, X., Tate, P., & Li, E. N. (1996). DNA methylation specifies chromosomal localization of MeCP2. *Molecular and Cellular Biology*. https://doi.org/10.1128/MCB.16.1.414

Negishi-Koga, T., Shinohara, M., Komatsu, N., Bito, H., Kodama, T., Friedel, R. H., & Takayanagi, H. (2011). Suppression of bone formation by osteoclastic expression of semaphorin 4D. *Nature Medicine*, *17*(11), 1473–1480. https://doi.org/10.1038/nm.2489

Nelson, C. E., Morgan, B. a, Burke, A. C., Laufer, E., DiMambro, E., Murtaugh, L. C., … Tabin, C. (1996). Analysis of Hox gene expression in the chick limb bud. *Development (Cambridge, England)*, *122*(5), 1449–1466. https://doi.org/10.1038/342767a0

Ng, F., Boucher, S., Koh, S., Sastry, K. S. R., Chase, L., Lakshmipathy, U., … Tanavde, V. (2008). identify markers and signaling pathways important in differentiation of PDGF , TGF- B , and FGF signaling is important for differentiation and growth of mesenchymal stem cells (MSCs): transcriptional profiling can identify markers and signaling pathways, *112*(2), 295–307. https://doi.org/10.1182/blood-2007-07-103697

NIH Consensus Development Panel on Osteoporosis Prevention, Diagnosis, A., & Therapy. (2001). Osteoporosis prevention, diagnosis, and therapy. *Jama*, *Feb 14; 28*(11), 785–795. https://doi.org/10.1001/jama.285.6.785

Nikitovic, D., Aggelidakis, J., Young, M. F., Iozzo, R. V., Karamanos, N. K., & Tzanakakis, G. N. (2012). The biology of small leucine-rich proteoglycans in bone pathophysiology. *Journal of Biological Chemistry*, *287*(41), 33926–33933. https://doi.org/10.1074/jbc.R112.379602

Nordhausen, K., Sirkiä, S., Oja, H., & Tyler, D. E. (2007). ICSNP: Tools for multivariate nonparametrics. *R Package Version*, 0–1.

Núñez, J. A., Goring, A., Javaheri, B., Razi, H., Gomez-Nicola, D., Hesse, E., … Clarkin, C. E. (2018). Regional diversity in the murine cortical vascular network is revealed by synchrotron X-ray tomography and is amplified with age. *European Cells and Materials*, *35*, 281–299. https://doi.org/10.22203/eCM.v035a20

Olenec, J., & Gern, J. E. (2009). Effect of Long-term Corticosteroid Use on Bone Mineral Density in Children: A Prospective Longitudinal Assessment in the Childhood Asthma Management Program (CAMP) Study. *Pediatrics*. Elk Grove Village, IL: https://doi.org/10.1542/peds.2009-1870SSS

Opperman, L. A. (2000). Cranial sutures as intramembranous bone growth sites. *Developmental Dynamics*, *219*(4), 472–485. https://doi.org/10.1002/1097-0177(2000)9999:9999<::AID-DVDY1073>3.0.CO;2-F

Paic, F., Igwe, J. C., Nori, R., Kronenberg, M. S., Franceschetti, T., Harrington, P., … Kalajzic, I. (2009). Identification of differentially expressed genes between osteoblasts and osteocytes. *Bone*, *45*(4), 682–692. https://doi.org/10.1016/j.bone.2009.06.010

Palazzini, S., Palumbo, C., Ferretti, M., & Marotti, G. (1998). Stromal cell structure and relationships in perimedullary spaces of chick embryo shaft bones. *Anatomy and Embryology*. https://doi.org/10.1007/s004290050145

Palomo, T., Al-Jallad, H., Moffatt, P., Glorieux, F. H., Lentle, B., Roschger, P., … Rauch, F. (2014). Skeletal characteristics associated with homozygous and heterozygous WNT1 mutations. *Bone*, *67*, 63–70. https://doi.org/10.1016/j.bone.2014.06.041

Pertea, M., Pertea, G. M., Antonescu, C. M., Chang, T. C., Mendell, J. T., & Salzberg, S. L. (2015). StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nature Biotechnology*, *33*(3), 290–295. https://doi.org/10.1038/nbt.3122

Piemontese, M., Almeida, M., Robling, A. G., Kim, H.-N., Xiong, J., Thostenson, J. D., … Jilka, R. L. (2017). Old age causes de novo intracortical bone remodeling and porosity in mice. *JCI Insight*, *2*(17). https://doi.org/10.1172/jci.insight.93771

Pitsillides Aa, Rawlinson Sc, Suswillo Rf, Bourrin S, Zaman G, & Lanyon Le. (1995). Mechanical strain-induced NO production by bone cells: a possible role in adaptive bone (re)modeling? *FASEB Journal*, *9*(15), 1614–1622.

Plotkin, L. I., & Bellido, T. (2016). Osteocytic signalling pathways as therapeutic targets for bone fragility. *Nature Reviews Endocrinology*, *12*(10), 593–605. https://doi.org/10.1038/nrendo.2016.71

Plotkin, L. I., Gortazar, A. R., Davis, H. M., Condon, K. W., Gabilondo, H., Maycas, M., … Bellido, T. (2015). Inhibition of osteocyte apoptosis prevents the increase in osteocytic receptor activator of nuclear factor κb Ligand (RANKL) but does not stop bone resorption or the loss of bone induced by unloading. *Journal of Biological Chemistry*, *290*(31), 18934–18942. https://doi.org/10.1074/jbc.M115.642090

Pocock, N. A., Eisman, J. A., Hopper, J. L., Yeates, M. G., Sambrook, P. N., & Eberl, S. (1987). Genetic determinants of bone mass in adults. A twin study. *J Clin Invest*, *80*(3), 706–710. https://doi.org/10.1172/jci113125

Poole, K. E. S. (2005). Sclerostin is a delayed secreted product of osteocytes that inhibits bone formation. *The FASEB Journal*, *18*, 1–18. https://doi.org/10.1096/fj.05-4221fje

Pruitt, K. D., Brown, G. R., Hiatt, S. M., Thibaud-Nissen, F., Astashyn, A., Ermolaeva, O., … Ostell, J. M. (2014). RefSeq: An update on mammalian reference sequences. *Nucleic Acids Research*, *42*(D1), 756–763. https://doi.org/10.1093/nar/gkt1114

Pruitt, K. D., Tatusova, T., Brown, G. R., & Maglott, D. R. (2012). NCBI Reference Sequences (RefSeq): Current status, new features and genome annotation policy. *Nucleic Acids Research*, *40*(D1), 130–135. https://doi.org/10.1093/nar/gkr1079

Qing, H., Ardeshirpour, L., Divieti Pajevic, P., Dusevich, V., Jähn, K., Kato, S., … Bonewald, L. F. (2012). Demonstration of osteocytic perilacunar/canalicular remodeling in mice during lactation. *Journal of Bone and Mineral Research*, *27*(5), 1018–1029. https://doi.org/10.1002/jbmr.1567

Qing, H., & Bonewald, L. F. (2009). Osteocyte remodeling of the perilacunar and pericanalicular matrix. *International Journal of Oral Science*, *1*(2), 59–65. https://doi.org/10.4248/ijos.09019

Qiu, S., Rao, D. S., Palnitkar, S., & Parfitt, A. M. (2003). Reduced iliac cancellous osteocyte density in patients with osteoporotic vertebral fracture. *Journal of Bone and Mineral Research*, *18*(9), 1657–1663. https://doi.org/10.1359/jbmr.2003.18.9.1657

Ramasamy, S. K., Kusumbe, A. P., Wang, L., & Adams, R. H. (2014). Endothelial Notch activity promotes angiogenesis and osteogenesis in bone. *Nature*, *507*(7492), 376–380. https://doi.org/10.1038/nature13146

Ramchand, S. K., & Seeman, E. (2018). The Influence of Cortical Porosity on the Strength of Bone During Growth and Advancing Age. *Current Osteoporosis Reports*, *16*(5), 561–572. https://doi.org/10.1007/s11914-018-0478-0

Rawlinson, S. C. F., McKay, I. J., Ghuman, M., Wellmann, C., Ryan, P., Prajaneh, S., … Kingsmill, V. J. (2009). Adult rat bones maintain distinct regionalized expression of markers associated with their development. *PLoS ONE*, *4*(12). https://doi.org/10.1371/journal.pone.0008358

Rawlinson, S. C. F., Mosley, J. R., Suswillo, R. F. L., Pitsillides, A. A., & Lanyon, L. E. (1995).

Calvarial and limb bone cells in organ and monolayer culture do not show the same early responses to dynamic mechanical strain. *Journal of Bone and Mineral Research*. https://doi.org/10.1002/jbmr.5650100813

Reijnders, C. M. A., Van Essen, H. W., Van Rens, B. T. T. M., Van Beek, J. H. G. M., Ylstra, B., Blankenstein, M. A., … Bravenboer, N. (2013). Increased expression of matrix extracellular phosphoglycoprotein (MEPE) in cortical bone of the rat tibia after mechanical loading: Identification by oligonucleotide microarray. *PLoS ONE*, *8*(11), 1–12. https://doi.org/10.1371/journal.pone.0079672

Rice, D. P. C., Rice, R., & Thesleff, I. (2003). Molecular mechanisms in calvarial bone and suture development, and their relation to craniosynostosis. *European Journal of Orthodontics*, *25*(2), 139–148. https://doi.org/10.1093/ejo/25.2.139

Richman, C., Kutilek, S., Miyakoshi, N., Srivastava, A. K., Beamer, W. G., Donahue, L. R., … Mohan, S. (2001). Postnatal and pubertal skeletal changes contribute predominantly to the differences in peak bone density between C3H/HeJ and C57BL/6J mice. *Journal of Bone and Mineral Research*, *16*(2), 386–397. https://doi.org/10.1359/jbmr.2001.16.2.386

Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., & Smyth, G. K. (2015). Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Research*, *43*(7), e47. https://doi.org/10.1093/nar/gkv007

Rivadeneira, F., & Mäkitie, O. (2016). Osteoporosis and Bone Mass Disorders: From Gene Pathways to Treatments. *Trends in Endocrinology and Metabolism*, *27*(5), 262–281. https://doi.org/10.1016/j.tem.2016.03.006

Rizzoli, R., & Bonjour, J.-P. (2010). Determinants of Peak Bone Mass Acquisition. *Osteoporosis*. London : https://doi.org/10.1007/978-1-59745-459-9_1

Rizzoli, R., & Bonjour, J. P. (1997). Hormones and bones. *The Lancet*, *349*, sl20-sl23.

Robling, A. G., Niziolek, P. J., Baldridge, L. A., Condon, K. W., Allen, M. R., Alam, I., … Turner, C. H. (2008). Mechanical stimulation of bone in vivo reduces osteocyte expression of Sost/sclerostin. *Journal of Biological Chemistry*, *283*(9), 5866–5875. https://doi.org/10.1074/jbc.M705092200

Rousseau-Nepton, I., Lang, B., & Rodd, C. (2013). Long-term bone health in glucocorticoid-treated children with rheumatic diseases. *Current Rheumatology Reports*, *15*(3), 6–9. https://doi.org/10.1007/s11926-012-0315-x

Rowe, D. W., Adams, D. J., Hong, S. H., Zhang, C., Shin, D. G., Renata Rydzik, C., … Ackert-Bicknell, C. (2018). Screening Gene Knockout Mice for Variation in Bone Mass: Analysis by μCT and Histomorphometry. *Current Osteoporosis Reports*, *16*(2), 77–94. https://doi.org/10.1007/s11914-018-0421-4

Rowe, P. S. N., De Zoysa, P. A., Dong, R., Wang, H. R., White, K. E., Econs, M. J., & Oudet, C. L. (2000). MEPE, a new gene expressed in bone marrow and tumors causing osteomalacia. *Genomics*, *67*(1), 54–68. https://doi.org/10.1006/geno.2000.6235

Rux, D. R., Song, J. Y., Swinehart, I. T., Pineault, K. M., Schlientz, A. J., Trulik, K. G., … Wellik, D. M. (2016). Regionally Restricted Hox Function in Adult Bone Marrow Multipotent Mesenchymal Stem/Stromal Cells. *Developmental Cell*, *39*(6), 653–666. https://doi.org/10.1016/j.devcel.2016.11.008

Rux, D. R., & Wellik, D. M. (2017). Hox genes in the adult skeleton: Novel functions beyond embryonic development. *Developmental Dynamics*, *246*(4), 310–317. https://doi.org/10.1002/dvdy.24482

Sano, H., Kikuta, J., Furuya, M., Kondo, N., Endo, N., & Ishii, M. (2015). Intravital bone imaging by two-photon excitation microscopy to identify osteocytic osteolysis in vivo. *Bone*, *74*, 134–139. https://doi.org/10.1016/j.bone.2015.01.013

Sarinho, E. S. C., & Melo, V. M. P. P. (2017). Doença Óssea Induzida Pelos Glicocorticoides: Mecanismos E Importância Na Prática Pediátrica. *Revista Paulista de Pediatria*, *35*(2), 207–215. https://doi.org/10.1590/1984-0462;2017/;35;2;00007

Sato, M., Asada, N., Kawano, Y., Wakahashi, K., Minagawa, K., Kawano, H., … Katayama, Y. (2013). Osteocytes regulate primary lymphoid organs and fat metabolism. *Cell Metabolism*, *18*(5), 749–758. https://doi.org/10.1016/j.cmet.2013.09.014

Schaffler, M. B., Cheung, W. Y., Majeska, R., & Kennedy, O. (2014). Osteocytes: Master orchestrators of bone. *Calcified Tissue International*, *94*(1), 5–24. https://doi.org/10.1007/s00223-013-9790-y

Schaffler, M. B., & Kennedy, O. D. (2012). Osteocyte signaling in bone. *Current Osteoporosis Reports*. https://doi.org/10.1007/s11914-012-0105-4

Schena, M., & Shalon, D. (1995). Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science*, *270*(5235), 467–470. https://doi.org/10.1126/science.270.5235.467

Seeman, E., & Delmas, P. D. (2006). Bone quality--the material and structural basis of bone strength and fragility. *The New England Journal of Medicine*, *354*(21), 2250–2261. https://doi.org/10.1056/NEJMra053077

Shamah, S. M., Lin, M. Z., Goldberg, J. L., Estrach, S., Sahin, M., Hu, L., … Greenberg, M. E. (2001). EphA receptors regulate growth cone dynamics through the novel guanine nucleotide exchange factor ephexin. *Cell*, *105*(2), 233–244. https://doi.org/10.1016/S0092-8674(01)00314-2

Sigal, V. L. (1984). On constructing a quantitative theory of open capillary chromatography. *Journal of Liquid Chromatography*, *7*(10), 1961–1968. https://doi.org/10.1677/JOE-10-0209

Sleiman, P. M. A., Wang, M. L., Cianferoni, A., Aceves, S., Gonsalves, N., Nadeau, K., … Hakonarson, H. (2014). GWAS identifies four novel eosinophilic esophagitis loci. *Nature Communications*, *5*, 1–5. https://doi.org/10.1038/ncomms6593

Smith, C. L., Blake, J. A., Kadin, J. A., Richardson, J. E., & Bult, C. J. (2018). Mouse Genome Database (MGD)-2018: Knowledgebase for the laboratory mouse. *Nucleic Acids Research*, *46*(D1), D836–D842. https://doi.org/10.1093/nar/gkx1006

Smith, C. L., & Eppig, J. T. (2009). The mammalian phenotype ontology: Enabling robust annotation and comparative analysis. *Wiley Interdisciplinary Reviews: Systems Biology and Medicine*, *1*(3), 390–399. https://doi.org/10.1002/wsbm.44

Somerville, J. M., Aspden, R. M., Armour, K. E., Armour, K. J., & Reid, D. M. (2004). Growth of C57Bl/6 Mice and the Material and Mechanical Properties of Cortical Bone from the Tibia. *Calcified Tissue International*, *74*(5), 469–475. https://doi.org/10.1007/s00223-003-0101-x

Spielmann, M., Brancati, F., Krawitz, P. M., Robinson, P. N., Ibrahim, D. M., Franke, M., … Mundlos, S. (2012). Homeotic arm-to-leg transformation associated with genomic rearrangements at the PITX1 locus. *American Journal of Human Genetics*, *91*(4), 629–635. https://doi.org/10.1016/j.ajhg.2012.08.014

St. John, H. C., Bishop, K. A., Meyer, M. B., Benkusky, N. A., Leng, N., Kendziorski, C., … Pike, J. W. (2014). The Osteoblast to Osteocyte Transition: Epigenetic Changes and Response to the Vitamin D $_3$ Hormone. *Molecular Endocrinology*, *28*(7), 1150–1165. https://doi.org/10.1210/me.2014-1091

Stoeckli, E. T. (2018). Understanding axon guidance: are we nearly there yet? *Development*, *145*(10), dev151415. https://doi.org/10.1242/dev.151415

Suda, T., Takahashi, T., Golstein, P., & Nagata, S. (1993). Molecular cloning and expression of the fas ligand, a novel member of the tumor necrosis factor family. *Cell*.

https://doi.org/10.1016/0092-8674(93)90326-L
Suemori, H., & Noguchi, S. (2000). Hox C cluster genes are dispensable for overall body plan of mouse embryonic development. *Developmental Biology*, *220*(2), 333–342. https://doi.org/10.1006/dbio.2000.9651
Sun, Y., Weng, Y., Zhang, C., Liu, Y., Kang, C., Liu, Z., … Wang, Z. (2015). Glycosylation of Dentin Matrix Protein 1 is critical for osteogenesis. *Scientific Reports*, *5*, 1–15. https://doi.org/10.1038/srep17518
Sun, Y. X., Schneider, A., Jung, Y., Wang, J., Dai, J., Wang, J., … Taichman, R. S. (2005). Skeletal localization and neutralization of the SDF-1(CXCL12)/CXCR4 axis blocks prostate cancer metastasis and growth in osseous sites in vivo. *Journal of Bone and Mineral Research*, *20*(2), 318–329. https://doi.org/10.1359/JBMR.041109
Supek, F., Bošnjak, M., Škunca, N., & Šmuc, T. (2011). Revigo summarizes and visualizes long lists of gene ontology terms. *PLoS ONE*, *6*(7). https://doi.org/10.1371/journal.pone.0021800
Sutton, A. L. M., Zhang, X., Dowd, D. R., Kharode, Y. P., Komm, B. S., & MacDonald, P. N. (2008). Semaphorin 3B Is a 1,25-Dihydroxyvitamin D $_3$ -Induced Gene in Osteoblasts that Promotes Osteoclastogenesis and Induces Osteopenia in Mice. *Molecular Endocrinology*, *22*(6), 1370–1381. https://doi.org/10.1210/me.2007-0363
Takeuchi, J. K., Koshiba-Takeuchi, K., Matsumoto, K., Vogel-Höpker, A., Naitoh-Matsuo, M., Ogura, K., … Ogura, T. (1999). Tbx5 and Tbx4 genes determine the wing/leg identity of limb buds. *Nature*, *398*(6730), 810–814. https://doi.org/10.1038/19762
Tang, S. Y., & Alliston, T. (2013). Regulation of postnatal bone homeostasis by TGFβ. *BoneKEy Reports*, *2*(October 2012), 1–5. https://doi.org/10.1038/bonekey.2012.255
Tang, S. Y., Herber, R. P., Ho, S. P., & Alliston, T. (2012). Matrix metalloproteinase-13 is required for osteocytic perilacunar remodeling and maintains bone fracture resistance. *Journal of Bone and Mineral Research*, *27*(9), 1936–1950. https://doi.org/10.1002/jbmr.1646
Team, R. C. (2013). R: A language and environment for statistical computing.
Teti, A., & Zallone, A. (2009). Do osteocytes contribute to bone mineral homeostasis? Osteocytic osteolysis revisited. *Bone*, *44*(1), 11–16. https://doi.org/10.1016/j.bone.2008.09.017
Thomas, G. P., Bourne, A., Eisman, J. A., & Gardiner, E. M. (2000). Species-divergent regulation of human and mouse osteocalcin genes by calciotropic hormones. *Experimental Cell Research*, *258*(2), 395–402. https://doi.org/10.1006/excr.2000.4912
Timpson, N. J., Greenwood, C. M. T., Soranzo, N., Lawson, D. J., & Richards, J. B. (2018). Genetic architecture: The shape of the genetic contribution to human traits and disease. *Nature Reviews Genetics*, *19*(2), 110–124. https://doi.org/10.1038/nrg.2017.101
Torzilli, P. A., Takebe, K., Burstein, A. H., Zika, J. M., & Heiple, K. G. (1982). The material properties of immature bone. *Journal of Biomechanical Engineering*, *104*(1), 12–20.
Toyosawa, S., Shintani, S., Fujiwara, T., Ooshima, T., Sato, A., Ijuhin, N., & Komori, T. (2001). Dentin matrix protein 1 is predominantly expressed in chicken and rat osteocytes but not in osteoblasts. *Journal of Bone and Mineral Research*, *16*(11), 2017–2026. https://doi.org/10.1359/jbmr.2001.16.11.2017
Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D. R., … Pachter, L. (2012). Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nature Protocols*, *7*(3), 562–578. https://doi.org/10.1038/nprot.2012.016
Tsunezumi, J., Sugiura, H., Oinam, L., Ali, A., Thang, B. Q., Sada, A., … Yanagisawa, H.

(2018). Fibulin-7, a heparin binding matricellular protein, promotes renal tubular calcification in mice. *Matrix Biology*, (2017). https://doi.org/10.1016/j.matbio.2018.04.014

UK Biobank. (2007). *UKBB ethics and governance framework 2007.pdf* (Vol. 0). Retrieved from https://www.ukbiobank.ac.uk/wp-content/uploads/2011/05/EGF20082.pdf?phpMyAdmin=trmKQlYdjjnQIgJ,fAzikMhEnx6

v Recklinghausen, F. (1910). Untersuchungen über Rachitis und Osteomalazie. In *Fischer Jena*.

Verborgt, O., Gibson, G. J., & Schaffler, M. B. (2000). Loss of osteocyte integrity in association with microdamage and bone remodeling after fatigue in vivo. *Journal of Bone and Mineral Research*, *15*(1), 60–67. https://doi.org/10.1359/jbmr.2000.15.1.60

Vogel, P., Read, R. W., Hansen, G. M., Powell, D. R., Kantaputra, P. N., Zambrowicz, B., & Brommage, R. (2016). Dentin Dysplasia in Notum Knockout Mice. *Veterinary Pathology*, *53*(4), 853–862. https://doi.org/10.1177/0300985815626778

Wain, L. V., Shrine, N., Miller, S., Jackson, V. E., Ntalla, I., Artigas, M. S., … Francks, C. (2015). Novel insights into the genetics of smoking behaviour, lung function, and chronic obstructive pulmonary disease (UK BiLEVE): A genetic association study in UK Biobank. *The Lancet Respiratory Medicine*, *3*(10), 769–781. https://doi.org/10.1016/S2213-2600(15)00283-0

Wand, J. S., Smith, T., Green, J. R., Hesp, R., Bradbeer, J. N., & Reeve, J. (1992). Whole-body and site-specific bone remodelling in patients with previous femoral fractures: relationships between reduced physical activity, reduced bone mass and increased bone resorption. *Clinical Science.* London :

Wang, Z., Gerstein, M., & Snyder, M. (2009). RNA-Seq: A revolutionary tool for transcriptomics. *Nature Reviews Genetics*. https://doi.org/10.1038/nrg2484

Warnes, G. R., Bolker, B., Bonebakker, L., Gentleman, R., Huber, W., Liaw, A., … Moeller, S. (2009). gplots: Various R programming tools for plotting data. *R Package Version*, *2*(4), 1.

Wasserman, E., Webster, D., Kuhn, G., Attar-Namdar, M., Müller, R., & Bab, I. (2013). Differential load-regulated global gene expression in mouse trabecular osteocytes. *Bone*, *53*(1), 14–23. https://doi.org/10.1016/j.bone.2012.11.017

Weaver, C. M., Gordon, C. M., Janz, K. F., Kalkwarf, H. J., Lappe, J. M., Lewis, R., … Zemel, B. S. (2016). *The National Osteoporosis Foundation's position statement on peak bone mass development and lifestyle factors: a systematic review and implementation recommendations*. *Osteoporosis International* (Vol. 27). https://doi.org/10.1007/s00198-015-3440-3

Wei, N., & Mao, Y. J. (2007). Dirichlet eigenvalue problems for p-subelliptic operators consisting of vector fields. *Fangzhi Gaoxiao Jichukexue Xuebao*, *20*(4), 388–391. https://doi.org/10.1359/jbmr.2001.16.11.1962

Weinstein, B. M. (2005). Vessels and nerves: Marching to the same tune. *Cell*. https://doi.org/10.1016/j.cell.2005.01.010

Wellik, D. M. (2007). Hox patterning of the vertebrate axial skeleton. *Developmental Dynamics*, *236*(9), 2454–2463. https://doi.org/10.1002/dvdy.21286

Wickham, H. (2007). Reshaping data with the reshape package. *Journal of Statistical Software*, *21*(12), 1–20.

Wickham, H. (2016). *ggplot2: elegant graphics for data analysis*. Springer.

Wilming, L. G., Boychenko, V., & Harrow, J. L. (2015). Comprehensive comparative homeobox gene annotation in human and mouse. *Database*, *2015*(February), 1–12.

https://doi.org/10.1093/database/bav091
Wilson, D. G., Phamluong, K., Lin, W. Y., Barck, K., Carano, R. A. D., Diehl, L., … Solloway, M. J. (2012). Chondroitin sulfate synthase 1 (Chsy1) is required for bone development and digit patterning. *Developmental Biology*, *363*(2), 413–425. https://doi.org/10.1016/j.ydbio.2012.01.005
Wiren, K. M., Toombs, A. R., Semirale, A. A., & Zhang, X. (2006). Osteoblast and osteocyte apoptosis associated with androgen action in bone: Requirement of increased Bax/Bcl-2 ratio. *Bone*, *38*(5), 637–651. https://doi.org/10.1016/j.bone.2005.10.029
Woo, S. M., Rosser, J., Dusevich, V., Kalajzic, I., & Bonewald, L. F. (2011). Cell line IDG-SW3 replicates osteoblast-to-late-osteocyte differentiation in vitro and accelerates bone formation in vivo. *Journal of Bone and Mineral Research*, *26*(11), 1–23. https://doi.org/10.1016/j.biocel.2009.12.011.Human
Wright, L. M., Maloney, W., Yu, X., Kindle, L., Collin-Osdoby, P., & Osdoby, P. (2005). Stromal cell-derived factor-1 binding to its chemokine receptor CXCR4 on precursor cells promotes the chemotactic recruitment, development and survival of human osteoclasts. *Bone*, *36*(5), 840–853. https://doi.org/10.1016/j.bone.2005.01.021
Wu, T., Chen, G., Tian, F., & Liu, H. X. (2017). Contribution of cranial neural crest cells to mouse skull development. *International Journal of Developmental Biology*, *61*(8–9), 479–493. https://doi.org/10.1387/ijdb.170051gc
Yang, J., Shah, R., Robling, A. G., Templeton, E., Yang, H., Tracey, K. J., & Bidwell, J. P. (2008). HMGB1 is a bone-active cytokine. *Journal of Cellular Physiology*, *214*(3), 730–739. https://doi.org/10.1002/jcp.21268
Yellowley, C. E., Li, Z., Zhou, Z., Jacobs, C. R., & Donahue, H. J. (2000). Functional Gap Junctions Between Osteocytic and Osteoblastic Cells. *Journal of Bone and Mineral Research*, *15*(2), 209–217. https://doi.org/10.1359/jbmr.2000.15.2.209
Yi, J. J., Barnes, A. P., Hand, R., Polleux, F., & Ehlers, M. D. (2010). TGF-β Signaling Specifies Axons during Brain Development. *Cell*, *142*(1), 144–157. https://doi.org/10.1016/j.cell.2010.06.010
Yu, G., Wang, L.-G., Han, Y., & He, Q.-Y. (2012). clusterProfiler: an R Package for Comparing Biological Themes Among Gene Clusters. *OMICS: A Journal of Integrative Biology*, *16*(5), 284–287. https://doi.org/10.1089/omi.2011.0118
Yu, G., Wang, L. G., Yan, G. R., & He, Q. Y. (2015). DOSE: An R/Bioconductor package for disease ontology semantic and enrichment analysis. *Bioinformatics*, *31*(4), 608–609. https://doi.org/10.1093/bioinformatics/btu684
Yue, F., Cheng, Y., Breschi, A., Vierstra, J., Wu, W., Ryba, T., … Ren, B. (2014). A comparative encyclopedia of DNA elements in the mouse genome. *Nature*, *515*(7527), 355–364. https://doi.org/10.1038/nature13992
Zakany, J., & Duboule, D. (2007). The role of Hox genes during vertebrate limb development. *Current Opinion in Genetics and Development*, *17*(4), 359–366. https://doi.org/10.1016/j.gde.2007.05.011
Zhang, K., Barragan-Adjemian, C., Ye, L., Kotha, S., Dallas, M., Lu, Y., … Bonewald, L. F. (2006). E11/gp38 Selective Expression in Osteocytes: Regulation by Mechanical Strain and Role in Dendrite Elongation. *Molecular and Cellular Biology*, *26*(12), 4539–4552. https://doi.org/10.1128/MCB.02120-05
Zhang, R., Lahens, N. F., Ballance, H. I., Hughes, M. E., & Hogenesch, J. B. (2014). A circadian gene expression atlas in mammals: Implications for biology and medicine. *Proceedings of the National Academy of Sciences*, *111*(45), 16219–16224. https://doi.org/10.1073/pnas.1408886111
Zhao, C., Irie, N., Takada, Y., Shimoda, K., Miyamoto, T., Nishiwaki, T., … Matsuo, K.

(2006). Bidirectional ephrinB2-EphB4 signaling controls bone homeostasis. *Cell Metabolism*, *4*(2), 111–121. https://doi.org/10.1016/j.cmet.2006.05.012

Zhao, S., Zhang, Y., Gamini, R., Zhang, B., & Von Schack, D. (2018). Evaluation of two main RNA-seq approaches for gene quantification in clinical RNA sequencing: PolyA+ selection versus rRNA depletion. *Scientific Reports*, *8*(1), 1–12. https://doi.org/10.1038/s41598-018-23226-4

Zhuang, W., Ge, X., Yang, S., Huang, M., Zhuang, W., Chen, P., … Li, B. (2015). Upregulation of lncRNA MEG3 promotes osteogenic differentiation of mesenchymal stem cells from multiple myeloma patients by targeting BMP4 transcription. *Stem Cells*, *33*(6), 1985–1997. https://doi.org/10.1002/stem.1989

Zurita, E., Chagoyen, M., Cantero, M., Alonso, R., González-Neira, A., López-Jiménez, A., … Montoliu, L. (2011). Genetic polymorphisms among C57BL/6 mouse inbred strains. *Transgenic Research*, *20*(3), 481–489. https://doi.org/10.1007/s11248-010-9403-8