UNIVERSIDADE DO ALGARVE

GENETIC AND FUNCTIONAL ANALYSIS OF THE PDB6 LOCUS IN PAGET'S DISEASE OF BONE

Iris Alexandra Lopes da Silva

Tese para a obtenção do grau de Doutor em Ciências Biomédicas

Trabalho efetuado sob a orientação de:

Professora Doutora M. Leonor Cancela

Doutora Laëtitia Michou

2015

UNIVERSIDADE DO ALGARVE

GENETIC AND FUNCTIONAL ANALYSIS OF THE PDB6 LOCUS IN PAGET'S DISEASE OF BONE

Iris Alexandra Lopes da Silva

Tese para a obtenção do grau de Doutor em Ciências Biomédicas

Trabalho efetuado sob a orientação de:

Professora Doutora M. Leonor Cancela

Doutora Laëtitia Michou

2015

Genetic and functional analysis of the PDB6 locus in Paget's disease of bone

Declaração de autoria de trabalho

Declaro ser a autora deste trabalho, que é original e inédito. Autores e trabalhos consultados estão devidamente citados no texto e constam da listagem de referências incluída.

_____

A quem me ajudou a ser quem sou hoje.

“Whatever you are be a good one!”
Abraham Lincoln

AGRADECIMENTOS/ACKNOWLEDGMENTS

Este trabalho não seria possível sem a ajuda de todos os que fizeram parte da minha vida nestes últimos anos. Dos que me ensinaram, dos que me ajudaram, dos que me fizeram crescer ou dos que apenas estiveram ao meu lado para tornarem os meus dias melhores.

Em primeiro lugar quero agradecer à professora Leonor Cancela. Primeiro, por desde sempre ter aberto as portas do seu laboratório para a minha entrada neste mundo fascinante que é o da ciência. Mas principalmente por servir de modelo desde o meu primeiro dia na Universidade do Algarve. Ver como consegue cuidar de todos nós, alunos e membros do seu laboratório, como faz os possíveis e impossíveis para permitir que tenhamos acesso às melhores condições para nos tornarmos os melhores profissionais possíveis. Um muito obrigada por tudo! Espero um dia fazer por outros o que a professora fez por mim.

I would also like to thank to Dr. Laetitia Michou, for being there for me in the most incredible experience in my life. Thank you for making me a much better scientist, for teaching me on how to discuss science and how to think for myself. Your words of encoradgement when I was away from everything that I knew were more important then words can explain. Thank you for everything and I hope that our paths cross again in a near future.

Também gostaria de agradecer à Doutora Natércia Conceição, a minha orientadora de sempre. Não foste no papel desta vez, mas não há outra pessoa que tenha estado tanto ao meu lado como tu. Obrigada por esclareceres as minhas 1001 duvidas, e por me puxares à terra quando eu começo a estravazar. Muito muito obrigada também pelas nossas reuniões no bar ou até no chat do gmail. Foi bom sentir que mesmo do outro lado do oceano te preocupavas comigo. Obrigada por tudo!

Aos meus colegas de laboratório, o melhor laboratório da UAlg (viva os outsiders!!!). Obrigada por tornarem os meus dias mais faceis, seja por me ajudarem a resolver os meus problemas ou por partilharem as minhas queixas. Obrigada ao Vincent, à Sara e ao Gavaia por aturarem a minha queda por 'coisas humanas' num mundo de peixes. À Anabela, à Brigite, ao Ricardo Leite, à Marta Rafael, por apesar de seguirem outros caminhos me ensinarem a dar os primeiros passos neste mundo. Ao gang dos ratos, Márcio, Filipe e António, por manterem sempre o bom ambiente geral do 2.13. Ao Nacho, por ter chegado e revolucionado o BIOSKEL. É bom ter pessoas como tu à nossa volta, espanhol. Obrigadas pelas horas e horas de discussão cientifica, nem sabes o valor que dei! Às meninas da bancada verde, à Cindy por toda a bondade desde o primeiro dia e por me ajudares com problemas às vezes tão faceis de resolver. À Helena, obrigada pelas gargalhadas constantes. À Andreia, a minha princesa que partilhou todos os momentos desde o inicio comigo, aquela que sabia exatamente quando eu entrava em stress ou quando não estava num dia bom. Obrigada por me fazeres sentir como se estivesse em Portugal, mesmo quando estava a milhares de kms de distância. À Joana Rosa e ao Mike, por todos os momentos de diversão neste mundo stessante que é o nosso. Ao Johnny, por me aturar nas minhas conversas (sempre) acima do tom normal. Obrigada pelos momentos de relax que tanto precisei. À minha querida Vânia, por hoje mais do que nunca seres uma das minhas 'pessoas'. Obrigada pelas nossas conversas intermináveis e por me ajudares em tudo! Seja a ver as coisas com mais clareza ou apenas a arrastares me para 'abanar a panela'! És a maior! Finalmente à minha Catxiane! (Ou deverei dizer à minha facada?) Obrigada é pouco para explicar a gratidão que sinto por estares sempre a meu lado em todas as

situações. És muito mais que uma amiga, és familia! Sabes que, hoje ou amanhã, aqui ou em qualquer outro lugar, serás sempre parte de mim.
I would also like to thank to my 'Canadian family'. Thank you to all of you for making me fell like home, even in a strange country with a language that I didn't quite understood. You made all much easier. To Edith for being my company in the lab and for helping me to ajust, and to Sabrina for being the best lab mate ever. Thank you for meeting me on that first day and help me with my first steps to become a real emigrant. To Mariejka for being so fun and always happy. Your laugh is contagious! To Davy for being the most kind and sweet person that I have ever known. You deserve the world! To Nathalie for being always in a good mood and for our French-English talks that were always funny. To Julie and Martin, for being the best, for making my weekends much more fun and for helping me to feel like I was part of your family. I will never forget what you did for me. Finally, to my dear Claudia, my Canadian sister, for all of the talks, the cooking, the laughing, for just being the best friend ever! My life changed since I meet you and it will never be the same. Thank you for being you! Thank you to all for the best experience of my life!! See you very soon!
Às minhas meninas, Ana Teresa, Joana Lopes, Joana Cabrita, Iolanda e Ricardo (sim, fazes parte das 'minhas meninas' get use to it!) por me mostrarem que há coisas que (felizmente!!!) nunca mudam. À Sofia, irmã de coração, por não ser preciso falar para me perceberes, por me entenderes sem ter de me explicar, por estares sempre a meu lado, mesmo quando me afasto. Obrigada soulmate!
À minha Aninhas, que me acompanha desde sempre, que me mostra constantemente que esteja onde eu estiver vamos ser sempre o Woody e o Buzz! Passe o tempo que passar. Obrigada por fazeres parte de mim! Há coisas que podem mudar mas nós seremos sempre nós!
À minha familia por ser a melhor familia do mundo! Por nestes 4 anos me terem apoiado, com o que tinham e o que não tinham, compreendendo ou não o que eu fazia. Eu não seria nada sem vocês. À minha avó pelo amor incondicional, à minha mãe e ao meu pai por me terem ajudado a tornar quem eu sou hoje. Devo-vos tudo a vocês! Obrigada. Um obrigada em especial à minha irmã Irina, que nestes ultimos 4 anos foi a melhor irmã, a melhor amiga, a melhor a fazer-me companhia às 4 da manhã portuguesas enquanto eu estava do outro lado do mundo. Não te digo isto muitas vezes mas espero que saibas que te adoro e te agradeço do fundo do coração por estares sempre do meu lado. És a melhor!!!!!
Finalmente ao meu Nuno. Obrigada pela ajuda, pelo amor, pela amizade e pelo carinho. Pela força e coragem. Contigo sinto-me capaz de tudo! Nestes 4 anos tivémos de ultrapassar grandes obstáculos mas como sempre, juntos, conseguimos isso e muito mais. Sei que não foi fácil para ti e que este mundo da ciência caiu um pouco do céu mas tu mostraste que a minha felicidade era o mais importante e ultrapassaste coisas que poucas pessoas conseguiriam. Hoje sinto que posso ir para onde a vida me levar, desde que esteja contigo tudo correrá bem. Obrigada baby!!!

Como podem ver pela extensão desta secção estou rodeada de muitas pessoas maravilhosas que tornam a minha vida uma experiênicia inesquecível!
Mil obrigadas por tudo!

A produção deste trabalho teve o financiamento da Fundação para a Ciência e Tecnologia, à qual expresso os meus agradecimentos. Referência da bolsa SFRH/BD/77227/2011.

ABSTRACT

Paget's disease of bone (PDB) is the most frequent metabolic bone disease after osteoporosis, and despite having a strong genetic component, *Sequestosome 1* is the only gene directly linked, so far, to this disease. Genome wide linkage and association studies have suggested an association of PDB6 locus with PDB, however no functional studies were performed to try to explain that association. The main goal of this work was to assess the relative contribution of each significant variant identified in the PDB6 locus to the functionality of the candidate genes and their involvement in PDB pathophysiology. In this regard we analyzed the entire PDB6 locus to search for the best candidate genes likely to be involved in bone metabolism. *OPTN*, *CCDC3*, *UCMA*/*GRP*, *CAMK1D*, *PHYH* and *SEPHS1* were selected and screened for genetic variants. Together with the variants found in a previous work from our group (Michou et al 2012) we studied 82 genetic variants, from which we selected rs2234968 and rs3829923 for an association study. The results showed that only rs2234968 was significantly associated with PDB (p-value = 6 x 10^{-3}). In addition, we assessed the effect of a rare variant (RV) found in one PDB patient, RV -9906, and rs1561570, a SNP in *OPTN* strongly associated with PDB. Our functional studies showed that (i) rs3829923 was responsible for an increase of *OPTN* promoter activity due to the appearance of new binding sites for activator nuclear factors E47 and E2F1, while (ii) RV -9906 was responsible for an increase of *OPTN* promoter activity but due to the loss of SP1 binding and consequent loss of its inhibitory effect in gene transcription. Rs2234968 was found to be in *linkage disequilibrium* with two SNPs that changed *OPTN* splicing, thus creating a premature stop codon, and probably resulting in a non-functional protein. In this work we showed that rs1561570, the SNP most associated to PDB, was responsible for the loss of a methylation site that subsequent increase in *OPTN* expression, which in turn increases NF-κB translocation into the nucleus and activation of NF-κB target genes. This will explain the increase in number and activity of the osteoclasts of PDB patients carrying of the rs1561570. In this work we provide new insights about *OPTN* regulation and its role in bone metabolism, with emphasis in osteoclastogenesis, and more importantly the contribution of *OPTN* variants to PDB pathophysiology.

Keywords: Paget's disease of bone, PDB6 locus, optineurin, osteoclastogenesis, rare genetic variants, single nucleotide polymorphism, zebrafish

RESUMO

A doença óssea de Paget (DOP) é a doença óssea metabólica mais frequente a seguir à osteoporose, afectando 3% da população Caucasiana a partir dos 55 anos. Esta doença é caracterizada por um aumento focal e desorganizado do metabolismo ósseo devido ao aumento do número, tamanho e atividade dos osteoclastos, as células responsáveis pela reabsorção óssea. No entanto, posteriormente existe também um aumento no número e atividade dos osteoblastos que, na tentativa de recuperar o osso que foi reabsorvido, dão origem a um osso que não possui uma estrutura normal, tem fraca qualidade e é mais susceptível de sofrer deformações e fraturas. Este tecido ósseo desorganizado é caracterizado por apresentar hipertrofia, fragilidade e hipervascularização, podendo diminuir drasticamente a qualidade de vida dos pacientes. No entanto, a maioria dos casos são assintomáticos. A sua prevalência aumenta com a idade e é mais frequente em países da Europa, como a Inglaterra. As causas da DOP ainda não foram clarificadas. Pensa-se que possa ter uma componente ambiental e que a exposição a agentes como o fumo do tabaco ou arsénio, ou a factores relacionados com uma vida rural como o consumo de leite não pasteurizado e o contacto com animais de quinta, pode contribuir para um aumento da susceptibilidade para desenvolver DOP. A existência de uma infeção viral pelo vírus do sarampo também foi considerada um forte factor de risco, mas essa relação continua ainda por comprovar.

A DOP possui também uma componente genética, com um padrão de hereditariedade autossómico dominante com penetrância incompleta, que vai desde os 17% antes dos 50 anos de idade até aos 80% depois dos 60 anos de idade. Cerca de 15 a 40% dos pacientes apresentam uma história familiar positiva, no entanto pensa-se que este número será muito maior devido ao facto desta ser uma patologia assintomática. O único gene associado à DOP até ao momento é o *Sequestosoma* 1 (*SQSTM1/p62*), que codifica para a proteína p62 que, por sua vez, está envolvida na iniciação da via do NF-κB, regulando a ubiquitinação da TRAF6 e a consequente degradação do inibidor do NF-κB (IκB), o que resulta na libertação do NF-κB para o núcleo. No entanto mais de 60% dos pacientes não têm qualquer mutação no *SQSTM1*, o que sugere que outros genes estarão envolvidos na DOP. Em 2010 estudos de ligação e de associação

sugeriram que existia uma relação entre o *locus* PDB6, nomeadamente o polimorfismo rs1561570 localizado no gene da *Optineurina* (*OPTN*), e a DOP, mas não existem estudos funcionais que expliquem esta relação. Assim, o objectivo principal deste trabalho foi analisar o *locus* PDB6 e tentar explicar a contribuição das variações genéticas encontradas para a DOP. Para isso analisámos toda a região do *locus* PDB6 e procurámos genes que pudessem ter um papel no metabolismo ósseo. Assim, selecionámos os genes *OPTN*, *CCDC3*, *UCMA*/*GRP*, *CAMK1D*, *PHYH* e *SEPHS1* que foram sequenciados utilizando uma amostra de 30 casos familiares e 5 controlos saudáveis. Juntamente com um trabalho anterior elaborado pelo nosso grupo foram encontradas no total 82 variantes genéticas nos seis genes analisados. Destas, duas foram selecionadas devido à sua frequência no grupo de pacientes e devido à sua localização genómica – o rs2234968 (G>A), localizado no primeiro exão codante da *OPTN*, e o rs3829923 (C>T), localizado no promotor partilhado entre os genes *OPTN* e *CCDC3*. Posteriormente, foi feito um estudo de associação com 225 pacientes e 298 controlos saudáveis que mostrou existir uma associação à DOP estatisticamente significativa (*p*-value = 6×10^{-3}) unicamente para o polimorfismo rs2234968.

Os estudos funcionais foram feitos com o objectivo de explicar o efeito (i) do polimorfismo rs1561570 - posteriormente associado à DOP -, (ii) do polimorfismo rs2234968 - descrito como associado à DOP neste trabalho -, e (iii) de duas variantes localizadas no promotor da *OPTN*, o polimorfismo rs3829923 - encontrada neste trabalho - e a variante rara RV -9906 - encontrada num único paciente, num trabalho anterior do nosso grupo. Os resultados obtidos mostraram que o polimorfismo rs3829923 era responsável por um aumento da atividade basal do promotor da *OPTN* devido à criação de um novo local de ligação dos factores de transcrição E47 e E2F1. Também foi demonstrado que a variante rara RV -9906 é responsável por um aumento da atividade do promotor da *OPTN,* mas neste caso devido à perda de ligação e consequente efeito inibitório do factor de transcrição SP1. Em relação ao polimorfismo rs2234968, os resultados mostram que este se encontra em desequilíbrio de ligação com outros dois polimorfismos - o rs10906303 e o rs79529484 - que são responsáveis por uma alteração no padrão de *splicing* da *OPTN*, alterando o quadro de leitura e originando um codão stop prematuro. A proteína resultante perde a maioria dos seus domínios proteicos e por isso é provável que seja degradada. Em relação ao polimorfismo rs1561570 (C>T), através do tratamento do DNA de pacientes e controlos saudáveis com bissulfito, e de técnicas como o qPCR ou *western blot,*

verificámos que o alelo mais comum nos pacientes - o alelo *T* - era responsável pela perda de um local de metilação e pelo aumento da expressão de OPTN. Também foi possível verificar que a presença do alelo *T* aumenta a translocação do NF-κB para o núcleo e a expressão dos seus genes alvo envolvidos na osteoclastogénese (tal como o *NFATc1*). Além disso a presença do alelo *T* também está relacionada com um aumento da taxa de diferenciação de osteoclastos, assim como um aumento no número de núcleos e na atividade destes osteoclastos, explicando assim o fenótipo pagético.

Por fim, foi feita uma análise bioinformática em que se comparou a conservação, ao longo da evolução, dos genes mais relevantes para a DOP - o *SQSTM1* e a *OPTN* -, nomeadamente entre o humano e o peixe-zebra. Através deste estudo verificou-se que ambos os genes são conservados entre humano e peixe-zebra e que aparentam possuir mecanismos de regulação semelhantes, com base na presença de locais putativos de ligação de factores de transcrição em comum nas duas espécies. É de destacar a conservação do resíduo P392 no SQSTM1 (frequentemente mutado em pacientes com DOP) nas duas espécies, o que mostra que o peixe-zebra poderá ser um bom modelo para estudar o efeito desta mutação no metabolismo ósseo. Em relação ao gene da *OPTN* é de realçar o facto deste gene ser expresso nos mesmos tecidos nas duas espécies, entre eles o tecido ósseo, o que indicia que o papel da OPTN possa ser semelhante nestas duas espécies.

Este trabalho permite concluir que o locus PDB6 é relevante no estudo da DOP, dado que possui vários genes com um possível papel no osso, nomeadamente o gene da OPTN. Foi possível decifrar novos mecanismos de regulação do gene *OPTN* desconhecidos até à data e propor um mecanismo que relacione este gene com a osteoclastogénese e com a DOP.

Palavras-passe: Doença Óssea de Paget, locus PDB6, optineurina, osteoclastogénese, variantes genéticas raras, polimorfismo, peixe-zebra

TABLE OF CONTENTS

ABBREVIATIONS & ACRONYMS

ALP	Alkaline Phosphatase
ALS	Amyotrophic Lateral Sclerosis
CAMK1D	Calcium/Calmodulin Dependent Protein Kinase ID
CCDC3	Coiled Coil Domain Containing 3
CHIP	Chromatin Immunoprecipitaion Assay
CSF1	Colony Stimulating Factor 1
CYLD	Cylindromatosis
DC- STAMP	Dendritic Cell–Specific Transmembrane Protein
EMSA	Electrophoretic Mobility Shift Assay
ESH	Expansile Skeletal Hyperphosphatasia
FEO	Familial Expansile Osteolysis
GOLGA6A	Golgin A6 Family, Member A
GWAS	Genome Wide Association Study
IBMPFD	Inclusion Body Myopathy, PDB And Frontotemporal Dementia
IKK	IκB kinase
IL-1	Interleukin-1
IL-1R	Interleukin-1 Receptor
IκB	Inhibitor of NF-κB
JPD	Juvenile PDB
LC3	Light Chain 3
LIR	LC3 Interacting Motif
LZ	Leucine Zipper Domain
MAF	Minor Allele Frequency
MVNP	Measles Virus Nucleocapsid Protein
NEMO	NF- κB Essential Molecule
NF-κB	Nuclear Factor κB
NIK	NF-κB Inducing Kinase
NRP	NEMO Related Protein
NUP205	Nucleoporin 205kDa
ON	Overnight

OPG	Osteoprotegerin
OPTN	Optineurin
PBMCs	Peripheral Blood Mononuclear Cells
PDB	Paget's Disease of Bone
PHYH	Phytanoyl CoA 2 Hydroxylase
POAG	Primary Open Angle Glaucoma
RANKL	Nuclear Factor κB Ligand
RIN3	Ras And Rab Interactor 3
RIP1	Receptor Interacting Protein 1
RT	Room Temperature
SEPHS1	Selenophosphate Synthetase 1
SNP	Single Nucleotide Polymorphism
SQSTM1	Sequestosome 1
TAD	Transcription Activation Domain
TAK1	TGF Beta Activated Kinase 1
TAX1BP1	TAX1 Binding Protein
TBK1	TANK Binding Kinase 1
TF	Transcription Factor
TFBS	Transcriptional Factor Binding Sites
TNFR	Tumor Necrosis Factor α Receptor
TNFRSF11A	Tumor Necrosis Factor Receptor Superfamily, Member 11a
TNFα	Tumor Necrosis Factor α
TRAF6	TNF Receptor Associated Factor 6
TRAP	Tartrate Resistant Acid Phosphatase
UBA	Ubiquitin Binding Associated Domain
UBD	Ubiquitin Binding Domain
UCMA/GRP	Upper zone of growth plate and Cartilage Matrix Associated/Gla Rich Protein
VCP	Valosin Containing Protein
ZnF	Zinc Finger

PREAMBLE

This thesis is divided into six chapters. The first chapter is dedicated to a literature overview about bone metabolism, Paget's disease of bone and OPTN function, with the objective of providing the necessary background to understand the data presented in this work. It is followed by a short description of the main goals.

The second chapter describes the initial part of this work that consisted in the discovery and analysis of genetic variants that might be associated with Paget's disease of bone.

The third chapter describes the functional effect of two variants located in *OPTN* promoter - rs3829923 and RV -9906 - that could affect *OPTN* regulation and is based on a manuscript submitted for publication.

The functional effects of rs2234968, found to be associated to PDB in this work, and rs1561570, the OPTN SNP most associated to Paget's disease of bone, are described in the fourth chapter.

A comparative analysis of *SQSTM1* and *OPTN* genes among several species, namely human and zebrafish, giving a molecular and evolutionary perspective to this work is presented in the fifth chapter, part of it is based on two manuscripts, one published in the Journal of Applied Ichthyology and another submitted to 'Bone'.

Finally, chapter six gathers the main conclusions drawn from the data presented in this thesis, the relevance to this disease of the analyzed SNPs and the perspectives for future works.

CHAPTER 1

GENERAL INTRODUCTION

1. GENERAL INTRODUCTION

1.1 Bone metabolism

Bone comprises the largest proportion of the human body's connective tissue mass. Unlike most other connective tissue, bone matrix is physiologically mineralized and is a unique tissue since it is constantly regenerated throughout life as a consequence of bone turnover. This continuous remodeling is controlled by two distinct processes, bone formation and bone resorption, that are strongly linked and tightly regulated to control and maintain skeletal homeostasis (Raggatt and Partridge, 2010).

There are three main cell types responsible for maintaining the correct balance in the remodeling process (Figure 1.1):

- Osteoblasts are the cells found on the bone surface responsible for bone formation. They are cuboidal cells with a round nucleus, usually organized as a single layer adherent to the surface of bone, and are derived from primitive mesenchymal cells (Mackie, 2003). They are responsible for synthesizing and secreting the collagen fibers, as well as non-collagen proteins such as osteocalcin, which constitute the organic phase of the bone matrix (Berridge, 2012). Although osteoblasts are specialized cells with a determined function, they are not terminally differentiated.
- Osteocytes are osteoblasts that became embedded in the bone matrix. They have a star-shape and are a type of cell that gradually stops secreting osteoid. These cells reside in lacunae and form a network of cytoplasmic extensions inside the canaliculated, dispersed throughout the mineralized bone tissue, and through which they communicate with each other and with surface osteoblasts (Marks and Odgren, 2002).
- Osteoclasts, the multinucleated cells derived from fusion of mononuclear haemopoietic precursors, such as monocytes, are responsible for the resorption of both the organic and inorganic components of bone (Novack, 2011). After their formation the mature osteoclasts can migrate over the bone surface until they find an area

where the bone has to be resorbed. The interaction with the bone surface triggers the osteoclast activation process that converts the mobile cell into a sedentary cell that can begin to resorb bone by acidifying the region located within the resorption pit, which is isolated from the surrounding bone by an actin ring resulting from the rearrangement of the actin cytoskeleton by the osteoclast podosomes (Berridge, 2012).

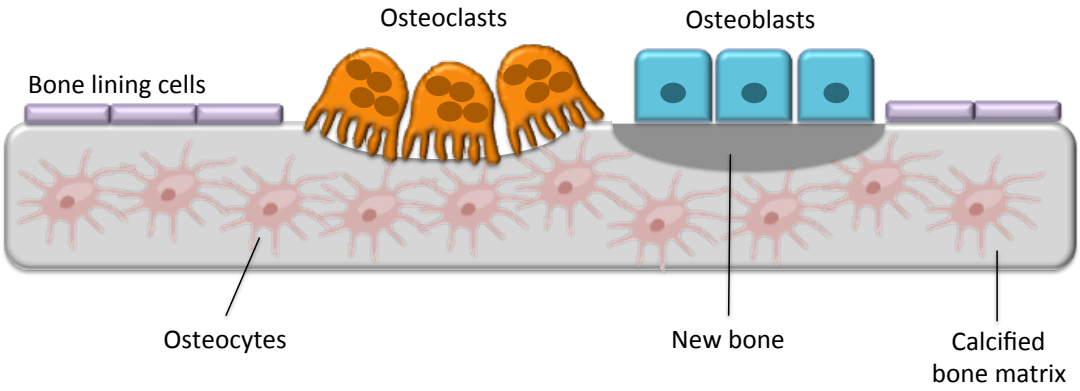


Figure 1.1 - Schematic representation of bone cells involved in bone remodeling. Osteoclasts (orange cells) are the cells responsible for bone resorption due to the action of its enzymes and acidification due to the production of chloridric acid. After bone resorption, osteoblasts (blue cuboid cells) are responsible for the formation of new bone, to maintain bone homeostasis. Osteoblast can also become embedded into the calcified matrix forming the osteocytes (star-shaped pink cells) that are connected through a complex network of cytoplasmic extensions inside lacunae and canaliculi. Adapted from (Matsuo and Irie, 2008).

There is a continuous communication between osteoclasts and osteoblasts in order to maintain a proper bone remodeling process. There are at least three kinds of osteoclast–osteoblast communication. Osteoclasts and osteoblasts can make direct contact, through the binding of membrane-bound ligands and receptors in order to initiate intracellular signaling or by forming gap junctions allowing passage of small molecules between the two types of cells. Communication can also occur through paracrine factors, such as cytokines like interleukins, growth factors, chemokines and other small molecules secreted by either cell type and acting on the other via diffusion. Finally, during bone resorption, osteoclasts may liberate growth factors and other molecules deposited by osteoblasts in the mineralized matrix (Matsuo and Irie, 2008). The osteoblasts exert a profound effect on the process of osteoclastogenesis where they provide the receptor activator of the nuclear factor κB (NF-κB) ligand (RANKL) that will bind to a receptor RANK - present in the osteoclast membrane

(Figure 1.2) (Berridge, 2012). RANK signaling activates a cascade of transcription factors important for osteoclastogenesis, including NF-κB, AP-1 (c-Fos) and NFATc1 (Matsuo and Irie, 2008). Osteoblast also release osteoprotegerin (OPG), which is a soluble RANKL-binding decoy receptor that inhibits the action of RANKL by masking it (Berridge, 2012). The osteoclastogenesis rate is controlled through the ratio between RANKL and OPG, both secreted by the osteoblasts.

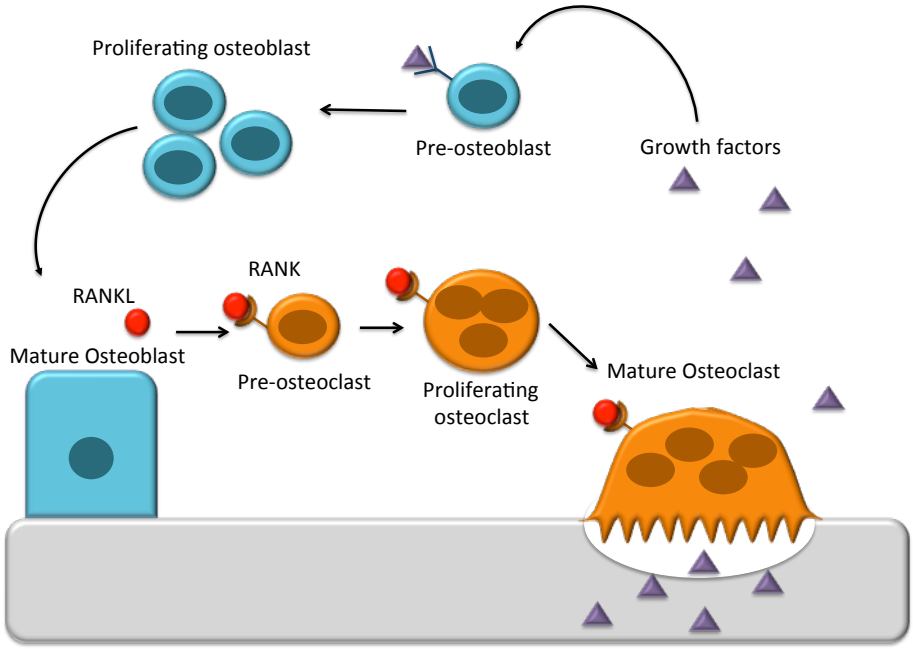


Figure 1.2 – Schematic representation of the crosstalk between osteoblasts and osteoclasts. The osteoblast produces the RANKL that will bind to its receptor in the membrane of the pre-osteoclasts inducing its differentiation. The mature osteoclast starts to resorb bone and releases growth factors that were trapped in the mineralized matrix, inducing the proliferation of pre-osteoblasts. RANK - Receptor activator of the nuclear factor κB, RANKL - Receptor activator of the nuclear factor κB ligand. Adapted from (Lee, 2010).

This process, through which the osteoblasts induce the osteoclasts differentiation, is very well established. However the reverse case is not that well investigated. One hypothesis is that after resorbing the mineralized matrix the osteoclasts release some growth factors that are embedded in the bone, inducing the proliferation of the pre-osteoblasts (Phan et al., 2004). Another theory is that through the interaction between ephrin-ligand and Eph-receptor expressing cells there is a bidirectional signal transduction. Reverse signaling through the ephrinB2 ligand into osteoclasts suppresses differentiation by blocking expression of the transcription factor AP1/c-Fos, while forward signaling through the EphB4 receptor into mesenchymal precursors promotes osteoblast differentiation by inhibition of RhoA activity. Since

ephrinB2 is expressed in mature osteoclasts and EphB4 is expressed by osteoblast precursors, maybe the ephrinB2/EphB4 interaction facilitates the transition from bone resorption to formation (Matsuo and Irie, 2008). However, to this date, no protein or cytokine expressed by the osteoclasts was shown to have a specific paracrine function on osteoblasts.

1.1.1 The role of NF-κB in osteoclasts

An extremely important pathway for osteoclast differentiation is the NF-κB signaling. The pleiotropic NF-κB transcription factors are a family consisting of five members subdivided into two subfamilies (Figure 1.3): The 'Rel' proteins, which includes Rel A (also designated as p65), Rel B, and c-Rel (also known as Rel), and the NF-κB proteins, which include NF-κB1 (p105/p50) and NF-κB2 (p100/p52). Rel A, Rel B and c-Rel contain a transcription activation domain (TAD) and Rel B also has a leucine zipper domain (LZ), which is required for full transcriptional activity (Novack, 2011).

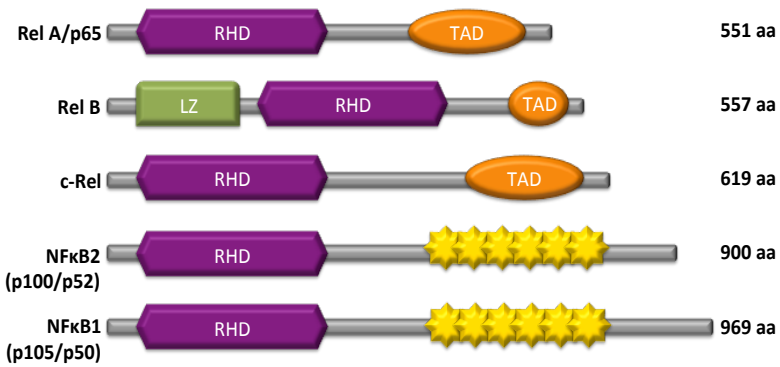


Figure 1.3 – Protein structure of NF-κB family members. NF-κB family is composed by Rel A/p65, Rel B, c-Rel, NF-κB1 (p105/p50) and NF-κB2 (p100/p52). All of these proteins have a Rel-homology domain (RHD). All Rel family members have a transcription activation domain (TAD) and Rel B also has a leucine zipper domain (LZ). NF-κB1 and NF-κB2 also possess several ankyrin repeats (yellow stars). The numbers represent the protein length in amino acids (aa). Adapted from (Soysa and Alles, 2009).

NF-κB1 and NF-κB2 are synthesized as large precursors, p105 and p100, respectively, that become shorter following post-translational modifications into p50 and p52, respectively, with long C-terminal domains that contain multiple ankyrin repeats (Figure 1.3). All of these NF-κB proteins share a Rel-homology domain (RHD) that contains a nuclear localization sequence and is important in dimerization,

DNA binding, and interaction with the inhibitor of NF-κB (IκB) proteins. Whereas the processing of p105 is constitutive upon activation, p100 processing is tightly regulated and highly inducible. NF-κB proteins, p50 and p52 are generally not activators of transcription unless they form dimers with Rel family members. All NF-κB family proteins can form homodimers or heterodimers *in vivo* with the exception of Rel B, which only forms heterodimers. Of the 15 possible dimers, at least 12 are able to bind to 9–10 bp NF-κB DNA binding sites (Hoffmann and Baltimore, 2006).

In most cells, NF-κB dimers are retained in the cytoplasm by inhibitory IκB proteins. These interactions are via the nuclear localization signal present in the RHD domain of the dimers and prevent nuclear localization and DNA binding. IκB proteins include IκBα, IκBβ, and IκBγ and contain ankyrin repeats that mediate binding of the RHD, allowing the binding of these inhibitory proteins with NF-κB1 or NF-κB2. There are two major NF-κB signaling pathways, named canonical (classical) and non-canonical (alternative) (Figure 1.4).

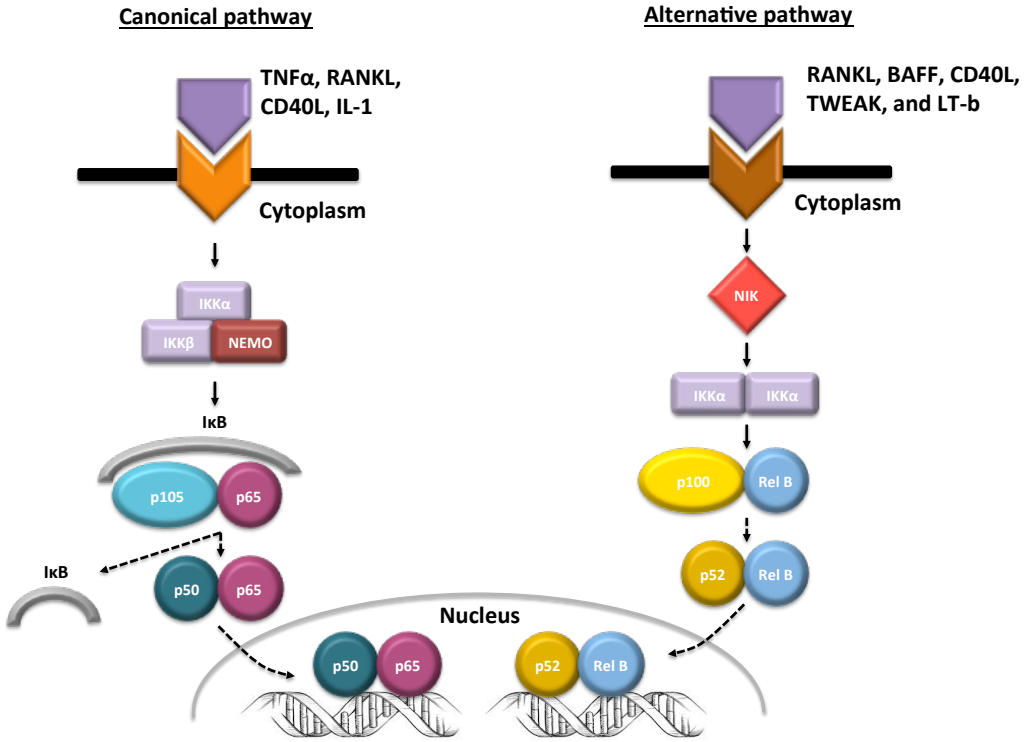


Figure 1.4 – Canonical and non-canonical NF-κB activating pathways. Canonical NF-κB pathway is typically activated through binding of receptor activator of the nuclear factor κB ligand (RANKL), tumor necrosis factor α (TNFα) and interleukin-1 (IL-1) and non-canonical pathway is activated by RANKL, B-cell activating factor (BAFF), CD40L, TWEAK, and lymphotoxin-b (LT-b). When the canonical pathway is activated there is the degradation of inhibitor of NF-κB (IκB) and consequent release of p50/p65 dimers to the nucleus. When the non-canonical pathway is activated there is the processing of p100 and consequent release of p52/Rel B dimers to the nucleus. Adapted from (Madonna et al., 2012).

1.1.1.1 Canonical NF-κB pathway

Canonical NF-κB pathway is typically activated through binding of RANKL, TNFα and IL-1 to their receptors on the cell surface (RANK, TNFR, and IL-1R, respectively), which in turn activates the IKK complex consisting of catalytic subunits IKKα, IKKβ, and the regulatory subunit IKKγ (or NEMO) (Figure 1.5) (Soysa and Alles, 2009). In response to activation of RANK by RANKL or of IL-1R by IL-1, the receptors undergo trimerization and their cytoplasmic domains associate with the adaptor protein TRAF6, which forms a complex with Sequestosome 1 (SQSTM1/p62) and an atypical protein kinase C.

Upon activation by binding to RANK or IL-1R, TRAF6 uses its ubiquitin-ligase activity to conjugate multiple units of ubiquitin to form a Lys63-linked polyubiquitin chain. The downstream TAB2-TAB3 adaptor proteins (in a complex with TAK1), which recognize and bind to this ubiquitin chain, are then phosphorylated and activate a further complex including IKKα, IKKβ and IKKγ. After this step IκB is phosphorylated, subsequently ubiquinated and targeted for proteasomal degradation, releasing the NF-κB dimers (mostly p50/p65) that than enter into the nucleus and selectively activate gene expression and stimulate osteoclast differentiation (Layfield, 2007).

The canonical pathway is rapidly activated by a large number of stimuli and is regulated by auto-regulatory feedback mechanisms to ensure transient activation. This is done by replenishing the pool of IκB proteins via NF-κB activation itself. The newly synthesized IκB proteins enter the nucleus and transport NF-κB dimers back to the cytoplasm. In addition, inactivation of IKK catalytic activity prevents degradation of newly synthesized IκB proteins (Soysa and Alles, 2009).

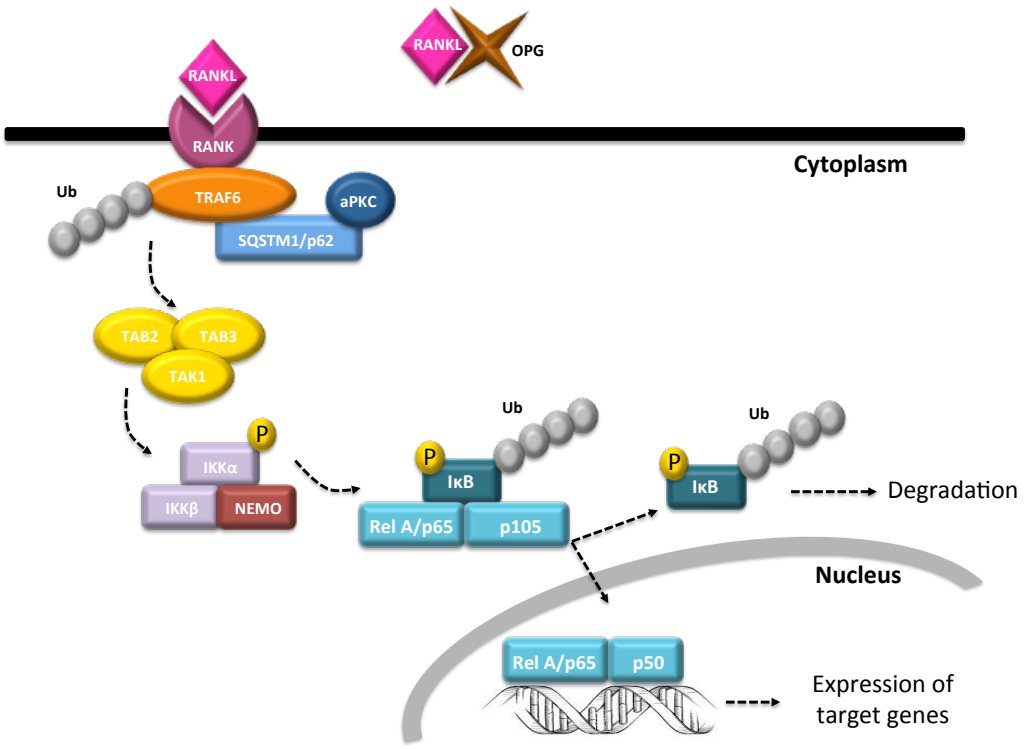


Figure 1.5 - The canonical NF-κB signaling pathway. RANKL is an osteoblast-derived cytokine that stimulates osteoclast activity and its binding to the cell-surface receptor RANK promotes recruitment of TRAF6 to the intracellular tail of RANK. TRAF6 uses its ubiquitin-ligase activity to undergo autoubiquitination. SQSTM1/p62, which complexes with aPKC, uses its ubiquitin-associated domain to regulate this ubiquitination of TRAF6. Assembly of a polyubiquitin chain on TRAF6 leads to activation of TAK1, a kinase that forms a complex with TAB2 and TAB3. TAK1 phosphorylates and activates the complex IKKβ, IKKα and IKKγ/NEMO. This complex in turn phosphorylates IκB, which is subsequently ubiquitinated; in this protein the polyubiquitin chain serves as a signal for the degradation of IκB by the proteasome complex. The NF-κB transcription factor (p65/p50) is then free from its inhibitor IκB and enters the nucleus to stimulate selective gene expression. aPKC - atypical protein kinase C, IκB - inhibitor of NF-κB; IKK - IκB kinase; NEMO - NF-κB essential modulator; NF-κB - nuclear factor κB; OPG - osteoprotegerin; RANK - receptor activator of NF-κB; RANKL - RANK ligand; SQSTM1/p62 - sequestosome 1; TRAF6 - TNF receptor-associated factor 6; Ub – ubiquitin. Adapted from (Layfield, 2007).

1.1.1.2 Alternative NF-κB pathway

In contrast to the canonical pathway, which depends on IKKβ and IKKγ/NEMO, IKKα is the only protein required for activation of the alternative pathway. This pathway is activated by TNF cytokine family members, including RANKL, B-cell activating factor (BAFF), CD40L, TWEAK, and lymphotoxin-b, but not by TNFα itself. Activation of NF-κB-inducing kinase (NIK) by the binding of these ligands to their receptors results in activation of IKKα homodimers and the phosphorylation and proteasome-induced processing of p100 by ubiquitination (Figure 1.6). Ubiquitinated

p100 is not completely degraded; instead it is cleaved to generate an active p52 product. Hence this process is slower than the activation of the classical pathway and leads to delayed activation and nuclear translocation of p52/Rel B under physiological conditions (Dejardin, 2006).

Given the ubiquity of NF-κB in normal cellular metabolism, as well as its association with a wide variety of diseases from autoimmunity to cancer, it is not surprising that the study of NF-κB in bone has now spread beyond to other cell types rather than osteoclasts.

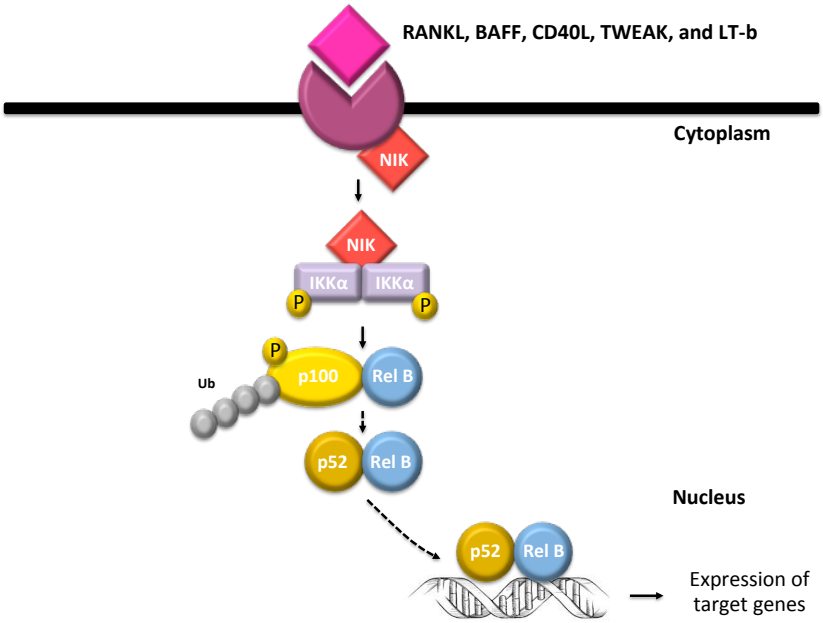


Figure 1.6 - The non-canonical NF-κB signaling pathway. RANKL and other ligands bind to the cell-surface receptor and activate NIK. NIK phosphorylates and activates IKKα homodimers that will label p100 for processing. After conversion into p52, the complex p52/Rel B is translocated to the nucleus and activates target gene expression. IKK - IκB kinase; NIK - NF-κB-inducing kinase; RANK-L - RANK ligand; Ub – ubiquitin. Adapted from (Layfield, 2007).

1.2 Paget's disease of bone

In 1876, Sir James Paget described five patients who had at least two deformed areas of the skeleton (Paget, 1877). These patients had what he called *osteitis deformans*. He believed he was describing a rare inflammatory bone disorder, but after the discovery of X-rays in 1895, the physicians realized that this disease was indeed not that rare, with numerous publications describing similar patients in England, France, and the United States (Leopold-Levi and Londe 1897). By this time, the condition

commonly became known as Paget's disease of bone (PDB).

1.2.1 Epidemiology

PDB is the second most common metabolic bone disorder after osteoporosis, and it affects 3% of Caucasians older than 55 years, and up to 10% of Caucasians older than 80 years (Ralston et al., 2008). It is slightly more common in men than in women and its prevalence increases with age (Figure 1.7A) (Siris and Roodman, 2008). However the prevalence differs amongst various ethnic/geographic groups (Figure 7B).

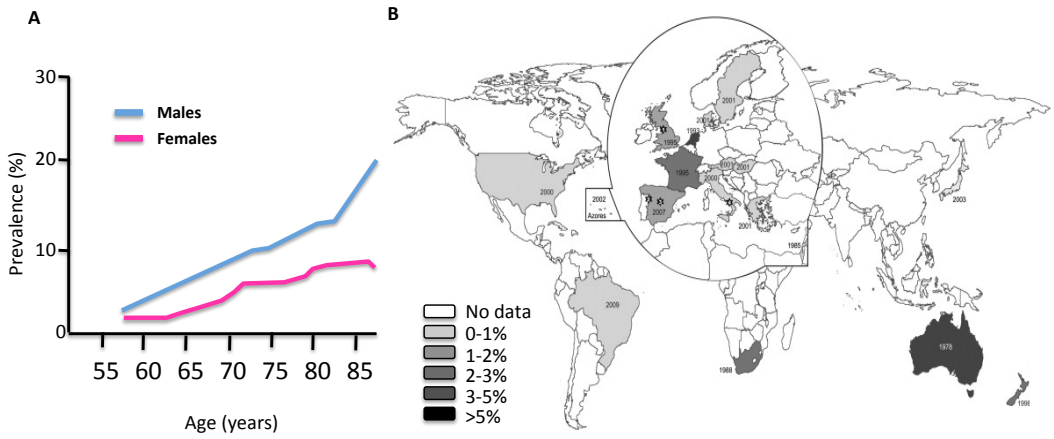


Figure 1.7 – Paget's disease of bone prevalence. A) Demographic data showing an increase in PDB prevalence with age both in males and females. B) World atlas representing the countries with higher PDB prevalence. Adapted from (Corral-Gudino et al., 2013).

This disease is most common in the United Kingdom (UK) and Western Europe but also in former British colonies such as Australia, New Zealand, South Africa, and South America. It is uncommon in Africa, Scandinavia, China, Japan, Southeast Asia, and India (Altman, 2002).

Furthermore, there is evidence of decreasing incidence and severity of PDB in the UK (Cooper et al., 2006) and New Zealand (Doyle et al., 2002) over the past 25-30 years. This decrease in the prevalence of PDB could be due to changes in the influx of migrants from low-prevalence regions, such as India and the Far East. It seems likely that other factors also contribute for this decrease, such as changes in environmental conditions that may trigger an increase in the susceptibility to develop the disease in genetically prone individuals. In fact, in the high prevalence region of Lancashire,

UK, it was reported the presence of arsenic in the water supply, indicating that this could be a risk factor (Lever, 2002). Also, the high prevalence of PDB in Campania, Italy, has been associated with a rural lifestyle, and with non-pasteurized milk consumption and meat that has not undergone health control (Rendina et al., 2006). The contact with domestic animals such as dogs and livestock, also appears to be a risk factor, although controversial (Siris et al., 1990). Cigarette smoke and wood heating were also described as potential risk factors (Guay-Bélanger et al., 2015). These studies show that although PDB prevalence depends on geographical regions, changes in the lifestyle and environmental exposure may contribute to change PDB onset and/or severity.

1.2.2 Clinical features

PDB is characterized by focal abnormal bone remodeling, with increased bone resorption and accelerated, excessive, and disorganized new bone formation. This deregulation of bone metabolism gives rise to a weaker bone, with poor quality, prone to deformities and fracture. It can be classified as monostotic (when only one bone in the skeleton is affected) or polyostotic (when a few bones display pagetic characteristics), the latest producing a more complex clinical setting (Kanis, 1998).

1.2.2.1 Disease progression

Clinically, three phases can be distinguished: first, the initial lytic phase, where there is an increase in osteoclast number. Osteoclasts are believed to be the first type of cells affected, which are increased in number, size and activity, resorbing bone at an incredible rate (Vinod and Reid, 2006). These osteoclasts are also hypersensitive to RANKL, 1,25-$(OH)_2D_3$ (1,25-dihydroxyvitamin D3), and TAFII17 (TATA-binding protein–associated factor [17 kDa], a vitamin D receptor binding protein) and display nuclear and cytoplasmic inclusion bodies (Chung and Van Hul, 2012). Clinical observations suggest that osteoclast-driven bone resorption is an early and progressive event: serial radiographs in several types of bone in the same patient show the progression of a lytic front (Figure 1.8) (Cundy and Bolland, 2008).

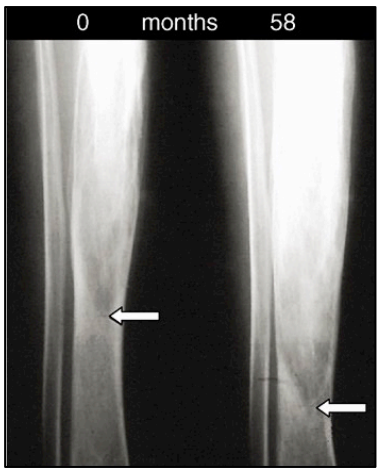


Figure 1.8 – Lytic phase in Paget's disease of bone. Progression of a lytic wedge (white arrows) in a patient's tibia visible on radiographs taken 58 months apart. Adapted from (Cundy and Bolland, 2008).

In the second phase, the mixed phase, and in response to the exaggerated osteoclast activity, bone formation also increases (Vinod and Reid, 2006). The bone formation rate can be exceptionally high: histological studies show that affected bones increase six or seven fold over normal (Kanis, 1998). This phase corresponds to the presence of both abnormally large and active osteoclasts resorbing bone and osteoblasts, which increase in number and try to compensate for the bone loss. Finally, the bone gets into a third phase, the sclerotic phase, in which bone formation is predominant, but that process is so rapid that the new bone is deposited in a chaotic fashion with loss of its normal lamellar pattern (Vinod and Reid, 2006). Therefore, despite affected bones being larger than normal, they are of poor quality – explaining the increase tendency to deform and fracture (Cundy and Bolland, 2008). In this last phase the increased osteoblast activity leaves a highly sclerotic signature in the affected bone, named "woven bone" or "cotton-like bone".

1.2.2.2 Symptoms

PDB patients can have any bone affected, but the axial skeleton, long bones and the skull are usually involved. The bones of the hands and feet are rarely involved. About 5% of patients with radiological PDB experience symptoms. The most common symptom is bone pain but other symptoms include increased skeletal warmth and erythema, due to an increase in the bone vascularity. Complications of PDB such as bowing deformities, due to the week bone structure, and deafness, due to skull enlargement and cranial nerve compression may occur (Figure 1.9) (Vinod and Reid,

2006).

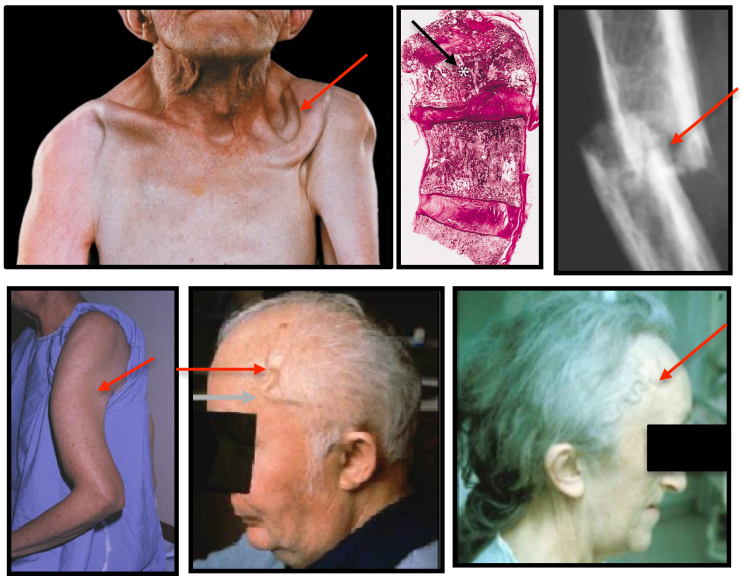


Figure 1.9 – Paget's disease of bone clinical symptoms. Due to a weaker and disorganized bone structure, PDB patients tend to have several clinical manifestations, such as bowing of long bones, skull enlargement or vertebrae collapse. (Pictures kindly provided by Dr. Brown, from the Research center of the CHU de Québec-Université Laval, Québec City, Canada)

Pagetic bone is brittle and may fracture spontaneously, particularly the femur, the humerus, the tibia, and the forearm; these acute fractures, typically transverse, can repair rapidly, although mal union is frequent. Fissure fractures represent a unique feature of PDB that can be distinguished from fractures occurring in other bone diseases (Rousière et al., 2003). However, this disease is mainly asymptomatic. The diagnosis is usually detected as incidental radiological abnormality or by an increased concentration of alkaline phosphatase in serum detected in regular blood analysis.

Some of these clinical features may give rise to secondary complications, associated with PDB. Secondary osteoarthritis is a common complication if a bone is affected next to a joint. If the skull is involved, the patient may develop dental malocclusion or nerve compression that can lead to deafness. Also, hypervascularization in the skull may divert blood from external carotid artery system affecting the brain blood support. This "vascular steal" may provoke drowsiness and apathy. Furthermore, when PDB affects more than 35% of the skeleton, it may cause a high cardiac output failure. This "vascular steal" can also affect the spine causing pain, dysesthesias or paralysis (Kanis, 1998).

The most feared PDB complication is bone malignant degeneration (Figure 1.10), which is estimated to occur in less than 1% of the cases, although the overall risk for

osteosarcoma is 30 times greater in pagetic patients than in the general population (Grimer et al., 2003). There might be an increased concern about the development of sarcoma if a new mass in soft tissue around the bone or a sudden and painful fracture is detected (López et al., 2003).

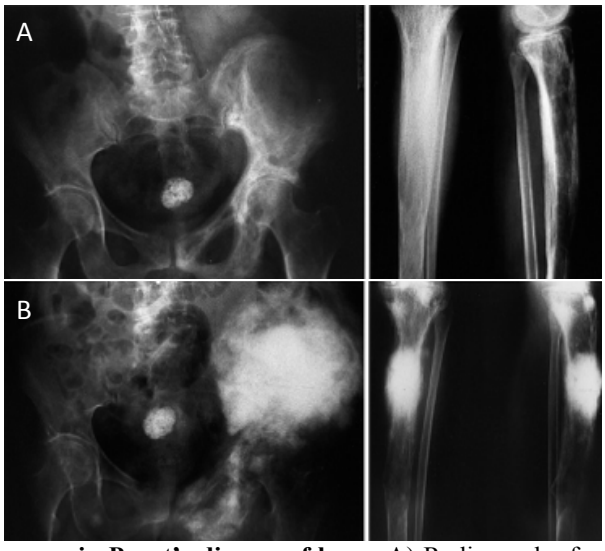


Figure 1.10 – Osteosarcoma in Paget's disease of bone. A) Radiograph of a PDB patient pelvis and left tibia showing the typical changes of uncomplicated PDB. B) Radiograph of the same PDB patient six years later, showing massive sclerotic masses suggestive of malignant transformation. Adapted from (Vuillemin-Bodaghi et al., 2000).

1.2.2.3 Diagnosis

PDB diagnosis is often fortuitous, since PDB is asymptomatic in up to 80% of the cases (Selby et al., 2002). PDB hallmarks are found when patients are examined for other reasons. Usually characteristic focal bone lesions with accelerated bone turnover can be detected on bone scan using radionuclide-labeled bisphosphonates or on X-rays. These radiological investigations are important to conform the diagnosis and at determining the extent of the disease (Chung and Van Hul, 2012). When PDB is diagnosed, radiographs typically show a mixed picture of lysis and sclerosis typical of PDB distinctive radiographic appearance (Cundy and Bolland, 2008). Radiographs can also assess complications such as fracture, secondary osteoarthritis and sarcomatous degeneration. Radionuclide bone scintigraphy is also helpful but a little less sensitive than X-rays (Vinod and Reid, 2006). The increased rate of remodeling is also reflected by an increase in bone turnover markers. All bone formation and resorption markers are affected similarly, such as alkaline phosphatase (ALP) and procollagen I N-terminal peptide for bone formation, and urinary N-telopeptide and

alpha-C-telopeptide for bone resorption (Shankar and Hosking, 2006). In the untreated state, the degree to which bone turnover markers are elevated is dependent on both the activity and the extent of the disease. Bone biopsy may be necessary to confirm the diagnosis, particularly if the results of radiographic studies and biochemical markers are inconclusive (Vinod and Reid, 2006).

1.2.2.4 Treatment

It is unclear whether there is value in treating this asymptomatic disease unless there is a clear and imminent risk of fracture. Treatment is indicated in those patients with bone pain, progressive skeletal deformity, compression of the spinal cord or nerve roots, and secondary fractures (Vinod and Reid, 2006).

PDB was an untreatable condition until the mid 1970s, when calcitonin was registered in many countries for the treatment of this disease. However, both salmon and human calcitonin have demonstrated to be ineffective in reducing bone turnover in most cases; moreover, they were shown to be associated with unpleasant side effects. The current first line treatment is bisphosphonates, which act by inhibiting osteoclast function. The greater advantage of these drugs over calcitonin relies in prolonged remissions, which can last for years, and in the healing of the resorption front (Langston and Ralston, 2004).

PDB can be monitored during the treatment with serial evaluations of bone markers. Indications for surgical treatment include secondary osteoarthritis, bone deformity fractures, and compression of nerve roots. After surgery, bone healing may be prolonged and rehabilitation may be delayed. The worst complication of surgery in pagetic patients is bleeding. Because of this, surgery has to be preceded by therapy with bisphosphonates in order to reduce the vascularity of the lesion as much as possible and to improve bone quality. Secondary osteosarcomas require chemotherapy, radiotherapy and/or surgery (Colina et al., 2008). In the end, operative procedures are useful for the relief of pain and for locomotion (Tancioni et al., 2006).

1.2.2.5 PDB-like pathologies

Other rare inherited bone disorders have also been described as having a phenotypic similarity with classical PDB, since they are also characterized by increased bone turnover, bone deformity, bone expansion, and elevated serum ALP concentrations.

These diseases include: familial expansile osteolysis (FEO), expansile skeletal hyperphosphatasia (ESH), early onset familial PDB, juvenile PDB (JPD), and the syndrome of hereditary inclusion-body myopathy, PDB and frontotemporal dementia (IBMPFD).

FEO is a bone dysplasia with many clinical features similar to PDB. The patients present generalized and focal skeletal abnormalities associated with elevated serum ALP and urinary hydroxyproline values, bone pain at affected sites, tooth loss, progressive loss of hearing, and nuclear inclusion bodies similar to those found in PDB patient osteoclasts (Wallace et al., 1989). However, the first symptom in most patients is hearing loss, sometimes in patients with four years old. Bone pain also appears much earlier than in PDB – in the second decade of life - and in some cases is so severe that it is resistant to pain killers and require limb amputation. The disease progression in long bones is almost twice the rate of lesion progression in PDB patients. The most remarkable difference between FEO and PDB is the apparent difference of the rates of osteoblast and osteoclast activity in the late stages of FEO, leading to notable expansion of the medullary cavity and thinning of the cortex, with almost complete replacement of the bone with vascularized adipose tissue (Osterberg et al., 1988).

ESH is a disorder that particularly affects the fingers but it is also characterized by early onset deafness, premature tooth loss, and involvement of the entire skeleton with striking progressive expansion of the long bones. Serum ALP and other markers of bone turnover are considerably elevated in these patients. However, ESH was not considered to be a variant of PDB because the patients also present hypercalcemia and widespread diffuse bone involvement without the presence of focal osteolytic lesions. Also, although an increased number of osteoblasts and osteoclasts were detected on bone biopsy, the osteoclasts were not enlarged to the same extent as seen in PDB (Whyte et al., 2000).

Early onset familial PDB was found in a Japanese family with severe form of PDB, whose symptoms arose in the second or third decade of life. The affected individuals had, for example, 2 to 17 times more serum ALP than normal. This condition had features similar to classic PDB, including axial skeletal involvement, bone expansion, and osteosclerotic lesions, but differed from PDB since they had an earlier age of onset, premature deafness, and premature tooth loss (Nakatsuka et al., 2003).

JPD is a rare autosomal recessive condition with approximately 50 cases reported

worldwide. It is characterized by a high rate of bone turnover; skeletal deformity, bone expansion, bone pain, and an increased risk of pathological fractures, as PDB. However, the first symptoms appear in the early infancy, and its severity increases in the adolescence, therefore this is clearly a more severe condition than PDB (Cundy et al., 2002; Whyte et al., 2002).

IBMPFD is a rare syndrome found in a series of families from the United States where the most prominent feature (occurring in 90% of individuals) is myopathy, present at a mean age of 42 years (range: 3–66 years). The pattern of muscle involvement included proximal and sometimes distal or facial muscles, often in an asymmetric and patchy distribution. Muscle biopsies revealed the presence of vacuolar inclusion bodies in muscle fibers. The PDB component had a penetrance of 43%, and was characterized by typical pagetic lesions involving the skull, spine, and pelvis. Dementia developed at an average age of 54 years and was 37% penetrant. Death typically occurred at a mean age of 58 years because of respiratory and cardiac failure (Kovach et al., 2001).

Despite the similarities with PDB features, all of these conditions have their own particularities; either they are characterized by an earlier diagnosis or by different and/or more severe symptoms. That can be explained by the different genetic mutations present in each type of patient, as described in figure 1.11.

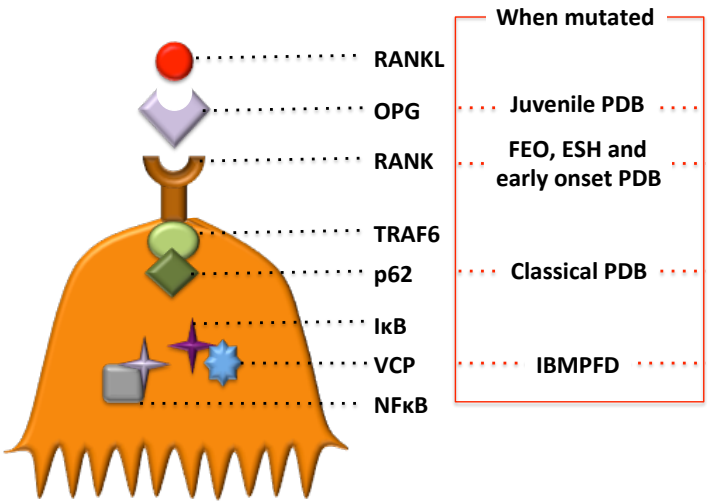


Figure 1.11 – Genes involved in the NF-κB pathway that are important for PDB and PDB-like disorders. Mutations in OPG were found in patients with juvenile PDB, RANK was associated with FEO, ESH and early onset PDB, SQSTM1/p62 with classical PDB and VCP with IBMPFD. RANK - Receptor activator of nuclear factor κB, RANKL – RANK ligand, OPG - Osteoprotegerin, VCP - Valosin containing protein, TRAF 6 - TNF receptor-associated factor 6, NF-κB – Nuclear factor κB, IκB - inhibitor of NF-κB. Adapted from (Lucas et al., 2006a).

By screening FEO families a duplication of 18-bp (84dup18), affecting the RANK signal peptide that segregated with the disease in all affected members, was identified. These results were replicated in families from Spain and the United States (Johnson-Pais et al., 2003; Palenzuela et al., 2002). At the same time, also in the RANK signal peptide, a 75dup27 mutation was found in the Japanese family with early-onset PDB. These mutations were not found in any healthy controls or sporadic PDB cases analyzed. Functional studies confirmed that these mutations are associated with failure of signal peptide cleavage resulting in an increase of RANK protein size. Expression of the 84dup18 and 75dup27 mutants showed increased constitutive activation of NF-κB signaling using promoter-reporter assays (Hughes et al., 2000). These observations led to the search for evidence of RANK mutations in individuals with classical PDB, but none were found, showing that mutation in the RANK gene are not a common cause of the classical PDB (Sparks et al., 2001). Other mutations in this gene were found in early-onset PDB, revealing the importance of RANK in this disease (Ke et al., 2009). Subsequently, after a screening of the RANK gene in two ESH patients there was an association of a 15-bp duplication to the FEO mutation (84dup15), and that was not present in 70 unaffected controls (Whyte and Hughes, 2002). In ESH patients, a 15-bp tandem duplication (84_98dup15) in TNFRSF11A exon 1 was shown to be associated with this disease (Whyte and Hughes, 2002). A mutation screening of TNFRSF11B/OPG was also performed in an Iraqi family in which three of nine siblings had JPD, revealing a homozygous 3-bp deletion in all affected siblings (Cundy et al., 2002). This deletion was predicted to result in the loss of an aspartate residue in OPG protein, which was shown to be crucial for OPG ability to inhibit osteoclastic resorption. Other groups found JPD patients with no OPG mutations, but in these cases their phenotype was much less severe (Chong et al., 2003). The gene associated with IBMPFD was valosin-containing protein (VCP). Thirteen IBMPFD families were screened for mutations in this gene, and six different disease-segregating mutations were identified (Watts et al., 2004). All of those mutations were localized near the VCP CDC48 domain, which is involved in ubiquitin binding. These IBMPFD associated mutations were considered relatively subtle, and only functional in response to oxidative stress and old age. No disease-causing mutations and no evidence of an association between VCP and PDB were found (Lucas et al., 2006b). Therefore, VCP appears to neither cause nor contribute to the pathogenesis of classical, late-onset PDB.

1.2.3 Pathophysiology

The cause of PDB is currently under intensive investigation, and both genetic and non-genetic factors have been implicated in the pathogenesis of this disorder.

1.2.3.1 Environmental Factors

Several observations suggest that environmental factors may contribute to the pathogenesis of PDB. The variable penetrance of this disease within families with a genetic predisposition to PDB, the fact that PDB affects a particular bone or bones rather than the entire skeleton, and the finding that the incidence and severity of the disease has been changing over the last 25 years (Doyle et al., 2002; van Staa et al., 2002) all support the hypothesis that non-genetic factors may be involved in PDB pathogenesis. The first implicated environmental trigger for PDB was a viral infection. That supposition arose from the observation that osteoclasts from PDB patients contained inclusion bodies in their nuclei and cytoplasm that were paramyxoviral-like nucleocapsids (Rebel et al., 1974). Many attempts have been made to detect the expression of paramyxoviral protein and/or DNA in osteoclasts and bone tissue from patients with PDB, but this question remains contradictory.

In a set of reports immunohistochemistry in PDB patients' osteoclast-like cells showed positive staining with antibodies directed against measles virus nucleocapsid protein (MVNP) and respiratory syncytial virus (Mills et al., 1994, 1981) and *in situ* hybridization showed MVNP expression in pagetic bone (Baslé et al., 1986) and in a subset of peripheral blood cells from pagetic patients (Reddy et al., 1996). Other groups have focused on the possibility that canine distemper virus (CDV), rather than MVNP, might be responsible for PDB, based on the observation that patients with PDB were more likely to have owned a dog than non pagetic controls (O'Driscoll and Anderson, 1985), and they detected the CDV mRNA in pagetic bone samples from British patients (Mee et al., 1998). However, none of these studies have demonstrated that a virus may cause PDB. These results suggest that, if involved, a viral infection by itself does not cause PDB but is probably a disease trigger. Transfection of normal human osteoclast precursors with the MVNP gene resulted in the formation of osteoclasts with many of PDB characteristics, including increased rate of osteoclast formation, increase in osteoclast numbers and size, increased bone resorbing capacity,

and hypersensitivity of transfected osteoclast precursors to 1,25-$(OH)_2D_3$. These characteristics are also observed in osteoclasts formed *in vitro* from freshly isolated marrow samples from PDB patients (Kurihara et al., 2000). Additional data was provided by the transgenic mice that express MVNP in the osteoclast lineages (MVNP mice) (Kurihara et al., 2006). Histomorphometric analysis of bones from 17 MVNP and 16 wild-type mice examined between 3 and 14 months of age showed a significant increase in osteoclast numbers and osteoblast activity. These results suggest that expression of MVNP in osteoclasts *in vivo* can induce bone changes that share many of the features of PDB. However, only 30% of the transgenic mice presented this pagetic phenotype (Kurihara et al., 2011, 2000), suggesting that virus may act as a modifier factor rather than a PDB cause.

Still, several groups have failed to detect MVNP or CDV in pagetic bone, peripheral blood, or cultured bone marrow using the same techniques (Birch et al., 2009; Helfrich et al., 2000; Nuovo et al., 1992; Ooi et al., 2000; Ralston et al., 2007). Also, a work from Helfrich and colleagues found no evidence of paramyxoviral protein or nucleic acids in pagetic bone samples or peripheral blood using a wide range of techniques including nested RT/PCR, immunohistochemistry, and *in situ* hybridization (Helfrich et al., 2000). Other recent studies using highly sensitive quantitative PCR-based techniques have also failed to detect evidence of MVNP in peripheral blood cells (Ralston et al., 2007) or in cultured osteoblasts from patients with PDB (Matthews et al., 2008). Another argument against this role of MVNP is the fact that, prior to the era of measles virus vaccination, measles was a ubiquitous infection, while PDB has a distinct geographic and racial distribution, being rare in the Far East but relatively common in the UK, Australia, and New Zealand (Roodman and Windle 2005).

Therefore, several questions remain concerning the involvement of measles in the pathogenesis of PDB. The osteoclast is not a self-renewing cell but is formed by fusion of post-mitotic precursors - thus, other cell types must act as a reservoir for measles virus to persist for long periods of time in patients with PDB (Roodman and Windle, 2005). Also, measles virus infections predominantly occur in children, while PDB is usually diagnosed in patients with more than 55 years old (Roodman and Windle, 2005), suggesting that a viral infection may increase the probability of developing PDB but another trigger is needed later in life.

1.2.3.2 Genetic Factors

Epidemiological studies show a variable prevalence of PDB in different ethnicities, with a higher number of PDB patients in Europe (Altman, 2002); first-degree relatives of patients with PDB have about a sevenfold greater risk for the development of PDB (Siris and Roodman, 2008); if a patient has an early age of diagnosis or deforming bone disease the first-degree relatives also have an increased risk of PDB (Siris et al., 1991); familial PDB also tends to be diagnosed at a younger age and involve more of the skeleton than sporadic disease (Seton et al., 2003). All of these findings suggest that PDB has a clear genetic cause.

Generally, patients with a known PDB family are assigned as familial cases (designated as "familial PDB") and those who are not aware of their positive PDB family history may have the complex or multifactorial form of PDB (also termed "sporadic PDB", although they may have a familial form, but not known by the patient, reason why affected unrelated PDB is probably a better term for PDB without any known family history than sporadic PDB).

There is a positive family history in approximately 15-30% of PDB patients and PDB seems to have an autosomal dominant mode of inheritance (Figure 1.12), with incomplete penetrance (17% before the fifth decade, 80% after the seventh decade of life), however it is hard to determine properly the truly familial cases due to the high percentage of asymptomatic patients (Siris et al., 1991).

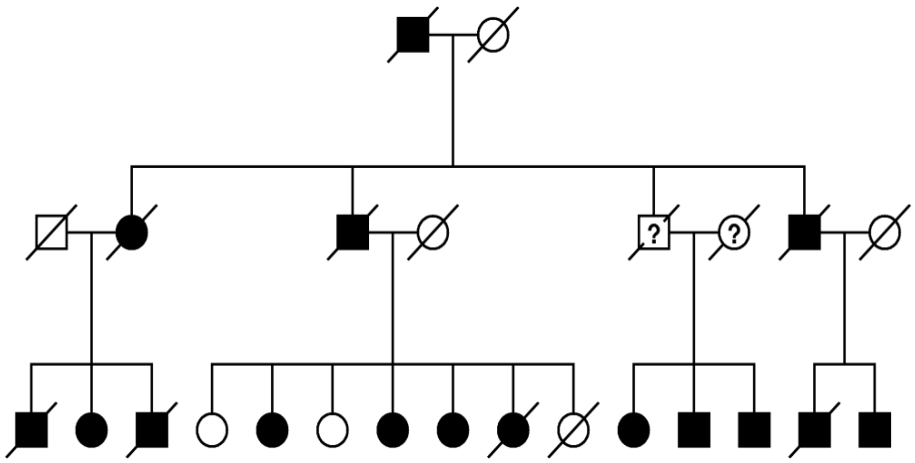


Figure 1.12 – Paget's disease of bone family pedigree. Evaluation of several PDB families shows an autosomal dominant mode of inheritance. Members of family with determined clinical evidence of PDB are represented by black symbols. Open symbols indicate subjects not exhibiting clinical evidence of PDB. Question marks identify individuals whose clinical phenotype is not verifiable. Pedigree developed in this thesis.

The genetic architecture of PDB is not fully understood, but accumulating evidence suggests that it is caused by a combination of rare, high-penetrance variants in several genes, which cause the autosomal dominant inheritance of PDB, and common variants, which individually are not sufficient to cause the disease but act together to increase the risk of developing PDB considerably (Albagha et al., 2011, 2010). These genes are usually key regulators of osteoclast differentiation or function (Figure 1.13).

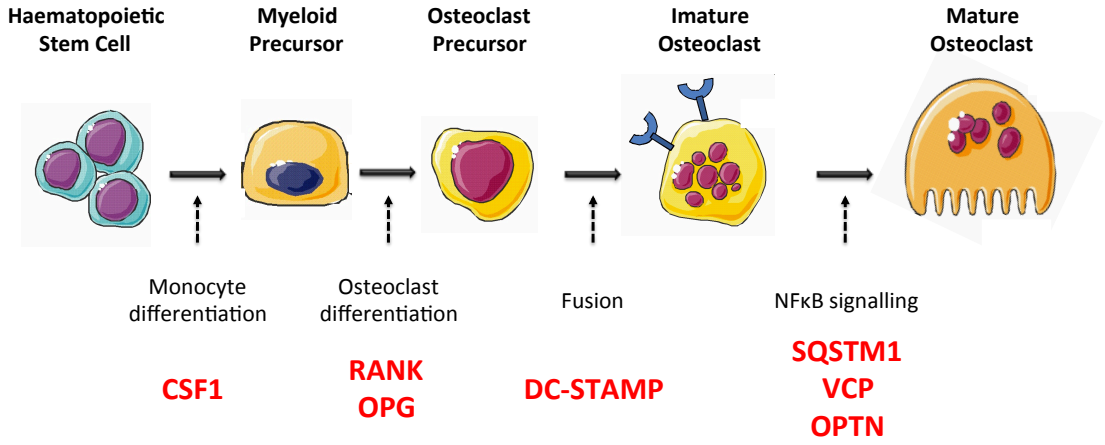


Figure 1.13 – Osteoclast differentiation pathway and genes involved in Paget's disease of bone. Several genes that increase the predisposition to PDB play central roles in the several osteoclast differentiation stages. CSF1 - Macrophage colony-stimulating factor, RANK - Receptor activator of nuclear factor κB, OPG - Osteoprotegerin, DC-STAMP - Dendritic cells specific transmembrane protein, SQSTM1 - Sequestosome 1, VCP - Valosin containing protein, OPTN – Optineurin. Adapted from (Ralston and Layfield, 2012).

Linkage studies of PDB families have identified a number of susceptibility loci on chromosomes 6p21 (PDB1) (Fotino et al., 1977), 18q21.1-22 (PDB2) (Cody et al., 1997), 5q35 (PDB3) (Laurin et al., 2001), 5q31 (PDB4) (Laurin et al., 2001), 2q36 (PDB5) (Good et al., 2002), 10p13 (PDB6) (Hocking et al., 2001) and 18q23 (PDB7) (Laurin et al., 2002). The regions identified are typically large and contain many genes that could be plausible candidate genes for PDB pathogenesis (Helfrich and Hocking, 2008).

1.2.3.2.1 Disease causing genes – *SQSTM1* gene

A strong susceptibility locus for familial PDB on chromosome 5q35 was identified in genome wide linkage scans in French-Canadian and British affected families (Hocking et al., 2001; Laurin et al., 2001). Sequencing of the genes within the selected region identified a mutation that changes a proline to a leucine in the amino

acid 392 (P392L) of the SQSTM1/p62 protein as the cause of the disease in French-Canadian families (Laurin et al., 2002), and subsequently this and other mutations clustering in the UBA domain of SQSTM1 were identified as the cause of the disease in patients of British descent (Hocking et al., 2002). Many missense and truncating *SQSTM1* mutations have now been identified in PDB patients, and all were found in or surrounding the UBA domain of the SQSTM1/p62 protein (Figure 1.14) (Cavey et al., 2006; Morissette et al., 2006).

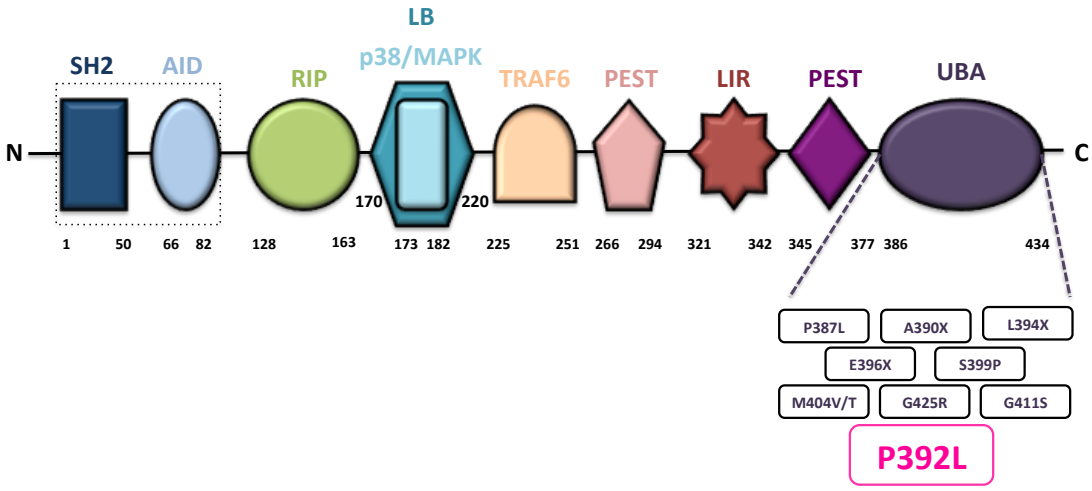


Figure 1.14 – SQSTM1/p62 protein domains and PDB most frequent mutations. SH2 - Scr homology 2, AID - acidic interaction domain, RIP - receptor interaction protein, LB - LIM-containing proteins binding domain, p38/MAPK - p38/mitogen-activated protein kinase, TRAF6 - TNF receptor-associated factor 6-binding domain. PEST - domain rich in proline (P), glutamic acid (E), serine (S) and threonine (T), LIR - LC3-interacting region, UBA - Ubiquitin binding association domain. The numbers represent the domain positions in the protein sequence. The most frequent PDB-related mutations are all localized in the UBA domain, being the P392L the most frequent (present in more than 80% of the cases). Adapted from (Chung and Van Hul, 2012).

SQSTM1/p62, as described in earlier sections, is an adaptor protein that binds ubiquitin and plays an important role in regulating NF-κB signaling (Layfield and Hocking, 2004) as well as autophagy (Ang et al., 2007; Bjørkøy et al., 2009; Pankiv et al., 2007). Mutations in *SQSTM1/p62* are found in 40–50 % of patients with a familial history of PDB and 2.5–10 % of sporadic PDB cases (Hocking et al., 2002; Laurin et al., 2002). These mutations were shown to increase osteoclastogenesis. The osteoclast precursors expressing the P392L mutation are hyper-responsive to osteoclastogenic factors, such as RANKL and TNFα, and SQSTM1/p62 P392L-transfected cells have an enhanced ability to resorb bone (Cavey et al., 2006). The presence of the P392L mutation also leads to a reduction of Cylindromatosis (CYLD) expression, and

therefore to an increase of the NF-κB signaling and *NFATc1* expression in pre-osteoclasts, favoring the progression of osteoclastogenesis and osteoclast activity associated with metabolic bone diseases (Figure 1.15) (Bjørkøy et al., 2005; Mills et al., 1981).

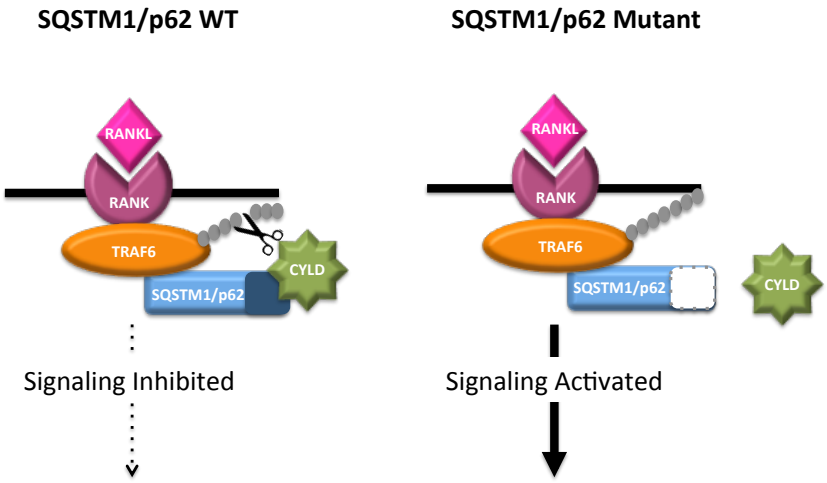


Figure 1.15 – Effect of *SQSTM1/p62* mutations in osteoclastogenesis. Proposed mechanism of the effect of *SQSTM1/p62* mutations in osteoclast activation. Under normal circumstances, the deubiquitinating enzyme CYLD is recruited by the UBA domain of SQSTM1/p62 (dark blue box) to the RANK receptor, where it deubiquitinates TRAF6 (scissors), resulting in inhibition of RANK signaling and osteoclastogenesis. Certain mutations of the SQSTM1/p62 UBA domain (white box) prevent CYLD being recruited, resulting in increased ubiquitination of TRAF6 (gray circles) and activation of RANK signaling. Adapted from (Ralston and Layfield, 2012).

To this date, *SQSTM1/p62* is the only causal gene with mutations in the coding region associated with PDB. However, studies of familial PDB in various populations indicate that *SQSTM1/p62* mutations are found in only about half of the cases analyzed. Positional cloning studies identified other potential susceptibility loci, which may contain other PDB-associated genes to explain the other familial forms of PDB.

1.2.3.2.2 Disease associated variants

Genome wide studies have recently identified other loci with genetic polymorphisms that have been linked to an increase susceptibility to develop PDB and together account for 13% of the familial risk of PDB. These regions contain several genes that have either known or theoretically likely effects on osteoclast functions: *M-CSF/CSF1* - macrophage colony stimulating factor in chromosome 1p13, *TNFRSF11A/RANK* -

18q21.33, *TM7SF4/DC-STAMP* - dendritic-cell-specific transmembrane protein in chromosome 8q22.3, *NUP205* - nucleoporin 205 kDa in chromosome 7q33, *RIN3* - Ras and Rab interactor 3 in chromosome 14q32, *PML* - promyelocytic and *GOLGA6A* - golgin A6 family member A, both in chromosome 15q24, and *OPTN* - optineurin in chromosome 10p13 (Albagha et al., 2011, 2010; Chung et al., 2010; Daroszewska et al., 2004).

- ***CSF1***: Genetic alterations in the *CSF1* gene on chromosome 1p13 were first identified as a PDB predisposing factor by a genome wide association study (GWAS) (Albagha et al., 2010), and the association was reinforced and confirmed in a subsequent study involving seven different populations (Albagha et al., 2011). The higher association observed was with the SNP rs484959, located about 87 kb upstream of *CSF1* transcription start site. The *CSF1* gene encodes M-CSF, an important protein in initiating osteoclast and macrophage differentiation (Figure 1.13) (Tanaka et al., 1993; Yoshida et al., 1990). Loss-of-function mutations of *CSF1* in rats and mice models cause osteopetrosis (Van Wesenbeeck et al., 2002; Yoshida et al., 1990). The mechanisms by which genetic alterations in the *CSF1* gene cause PDB remain unknown, but it seems likely that they upregulate *CSF1*, given the fact that serum M-CSF levels are increased in PDB patients (Neale et al., 2002). That increase in *CSF1* expression will increase the differentiation of haematopoietic stem cells into the monocyte lineage, and those monocytes will later differentiate into osteoclasts, explaining the PDB phenotype (Ralston and Layfield, 2012).
- ***TNFRSF11A/RANK***: The *TNFRSF11A* gene encodes the receptor activator of NF-κB (RANK), which belongs to the TNF receptor family and plays an essential role in osteoclast differentiation and bone resorption (Figure 1.13) (Li et al., 2000). Mutations have not been detected in the exons or intron-exon boundaries of *RANK* in classical PDB (Sparks et al., 2001; Wuyts et al., 2001), but there is evidence from candidate gene studies and GWAS that common variants at the RANK locus could predispose to PDB (Albagha et al., 2011, 2010; Chung et al., 2010). Two common polymorphisms have been described in the protein-coding region of RANK - H141Y (a change from a histidine to a tyrosine at codon 141) and A192V (a change from an alanine to a valine at codon 192). (Sparks et al., 2001)

Functional analyses of these polymorphisms have yielded contradictory results; in one study they performed a reporter assay but no effect of these polymorphisms on RANK-induced NF-κB activation was found (Chung et al., 2010), whereas in the other study the A192V was described as activating RANK signaling in both the presence and the absence of the *SQSTM1/p62* P392L mutation (Gianfrancesco et al., 2012). Further studies will be required to confirm these observations and to determine if the A192V variant is fully responsible for the robust association that has been observed between PDB and common variants at the *TNFRSF11A/RANK* locus (Albagha et al., 2011, 2010; Chung et al., 2010; Wild et al., 2011a).

- ***TM7SF4/DC-STAMP***: The *TM7SF4* gene is localized on chromosome 8q22 and encodes the dendritic cell–specific transmembrane protein (DC-STAMP), an essential cell surface protein that allows the fusion of pre-osteoclasts and formation of multinucleated active osteoclasts (Figure 1.13) (Yagi et al., 2005). The SNP rs2458415 has been strongly associated to PDB in an extended GWAS (741 cases and 2,699 controls), along with six independent replication populations (1,474 cases and 1,671 controls). It seems likely that the genetic variants in this gene that predispose to PDB increase *TM7SF4* expression or cause a gain of function effect at the protein level, increasing the fusion of mononuclear pre-osteoclasts, but further studies will be required to investigate these possibilities (Albagha et al., 2011).
- ***NUP205***: This gene was also associated to PDB by a GWAS (Albagha et al., 2011). The strongest signal was within intron 22 of *NUP205* in the region of the SNP rs4294134. This protein is a component of the nuclear pore complex, which is involved in the regulation of transport of substances between the cytoplasm and nucleus (Grandi et al., 1997). However, this gene has not been so far implicated in bone metabolism. Further studies will be required to investigate how can *NUP205* affect bone cell function.
- ***RIN3***: The *RIN3* gene is localized in chromosome 14q32.12 and encodes the Ras and Rab interactor 3, a protein that interacts with small GTPases such as Ras and Rab, being important in vesicular trafficking (Kajiho et al., 2003; Saito et al., 2002). This gene was tagged by the association of rs10498635 with a large group of PDB patients in a GWAS (Albagha et al.,

2011). The function of *RIN3* in bone is not known, but it could play a role in bone resorption since small GTPases are important in vesicular trafficking and in osteoclast function (Coxon and Rogers, 2003; Van Wesenbeeck et al., 2007). Also, mutations in VCP, a protein also involved in vesicular trafficking, cause a PDB-related pathology - IBMPFD (Watts et al., 2004).

- ***PML*** and ***GOLGA6A***: The 15q24 locus was also identified as associated to PDB by the same GWAS as *NUP205* and *RIN3* (Albagha et al., 2011). The strongest associated SNP (rs5742915) lies within the *PML* gene and causes a change of a phenylalanine to a leucine at codon 645 of the PML protein (P645L). PML is involved in regulating cell growth, apoptosis, and senescence by interacting with various proteins such as p53, Rb, and death-domain associated protein (DAXX) (Salomoni and Pandolfi, 2002). There is evidence that there is a defect in apoptosis in pagetic osteoclasts (Chamoux et al., 2009), providing a potential mechanism through which *PML* could explain the PDB phenotype. The *GOLGA6A* gene encodes a member of the golgin family of coiled-coil proteins associated with the Golgi apparatus, which are important for membrane fusion and as structural supports for the Golgi cisternae. The previously referred SNP (rs5742915) lies near this gene, and therefore the association can be due to the effect of this SNP in *GOLGA6A* and not in *PML*(Albagha et al., 2011). The role of the *GOLGA6A* gene in bone metabolism is unknown at this time but mutations in other members of the golgin family have been shown to cause lethal skeletal dysplasia and a severe form of osteoporosis (Hennies et al., 2008; Smits et al., 2010). Therefore, further functional studies will be required to identify the effect of this genetic variant in these two genes and how they affect bone metabolism.
- ***VCP****:* VCP protein coding mutations have also been associated with IBMPFD but have not been detected in classical PDB (Lucas et al., 2006b). However, two groups have investigated the effect of common genetic variants in the *VCP* gene in this disease. A study in the Belgian population showed evidence of an association between the rs565070 SNP in the *VCP* gene and PDB (Chung et al., 2011). However, in another study in the British population no evidences were found for an association between

common variants at the *VCP* gene and PDB, nor was there evidence of allele sharing at this locus in subjects with familial PDB (Lucas et al., 2006b) Therefore., it remains uncertain whether common variants at the *VCP* locus contribute to the pathogenesis of PDB in the absence of other features of the IBMPFD syndrome.

The association of OPTN and PDB will be addressed further in this thesis.

1.2.3.3 Autophagy

There is increasing interest in the possibility that abnormalities in macroautophagy may contribute to the pathogenesis of PDB (Goode and Layfield, 2010). Autophagy is the biological process that allows the degradation of cytoplasmic proteins, protein aggregates, and damaged organelles. After the encapsulation of the target protein into a double membrane bound structure designed as autophagosome, there is a fusion between the autophagosome and lysosomes or endosomes containing hydrolases (that are responsible for destroying its contents) (Figure 1.16) (Ravikumar et al., 2010). This process involves SQSTM1/p62, OPTN and other ubiquitin-binding proteins, which also play a role in regulating bone cell function (Whitehouse et al., 2010). All of these proteins contain an ubiquitin binding associated (UBA) domain, which allows them to bind to ubiquitinated proteins and organelles that were targeted for degradation (Kirkin et al., 2009a). The SQSTM1/p62 and similar proteins then act as autophagy receptors and bind proteins present in the autophagosomal membranes such as LC3. This allows SQSTM1/p62 and their ubiquitinated cargo to bind to the autophagosome, resulting in degradation of the ubiquitinated cargo and receptor protein.

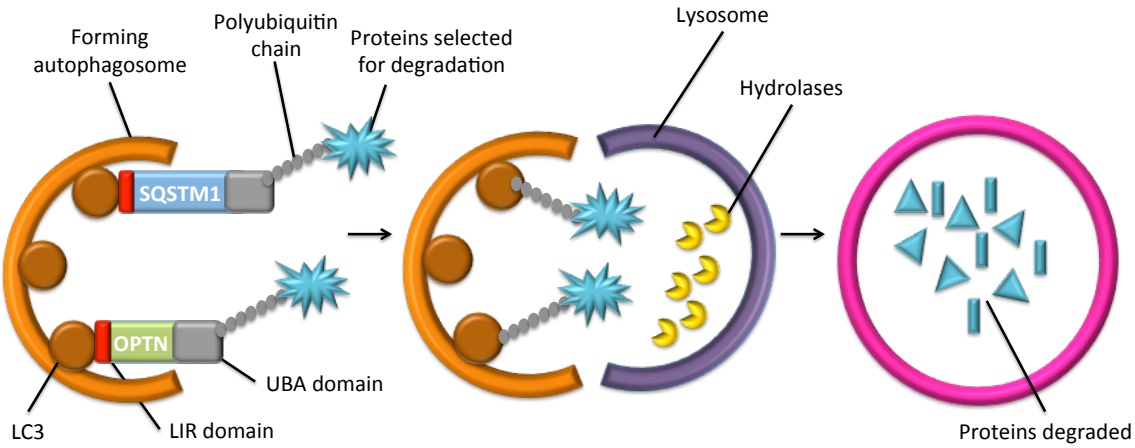


Figure 1.16 - Mechanisms of autophagy. Autophagy receptors, including SQSTM1/p62 or OPTN, interact with polyubiquitinated substrates for autophagy (light blue) through their ubiquitin-binding domains (UBA) (grey box) and with autophagosome-associated proteins such as LC3 (orange circles) through LC3-interacting motifs (red). The substrates are engulfed by the forming autophagosome, which fuses with lysosomes or other endosomes containing hydrolases that degrade the contents of the autophagosome. Adapted from (Ralston and Layfield, 2012)

Several lines of evidence suggest that PDB may be associated with dysregulation of autophagy. First is the fact that genetic variants in PDB patients usually affect genes involved in the process of autophagy, such as *SQSTM1/p62*, *VCP*, and *OPTN*. Indeed, a requirement for SQSTM1/p62's ability to regulate autophagy is the presence of a functional UBA domain, which is the region commonly mutated in PDB (Bolland et al., 2007). Second is the recent observation that mice models harboring the P394L Sqstm1/p62 mutation exhibit a PDB-like skeletal phenotype and have an increased gene expression of *Sqstm1*, as well as several autophagy related genes, such as autophagy-related gene 5 (*atg5*), and light chain 3 gene (*lc3*) in osteoclast precursors, consistent with an dysregulation of the autophagy process. Third is the fact that SQSTM1/p62 protein levels are increased in peripheral blood cells and in osteoclasts from patients with PDB, regardless of the *SQSTM1* mutation status (Chamoux et al., 2009; Collet et al., 2007). Since SQSTM1/p62 is degraded as part of the process of autophagy, raised levels of this protein may indicate a reduced autophagic flux. Fourth, defective autophagy may be a feature of IBMPFD (caused by *VCP* mutations), at least in muscle cells (Ju et al., 2009; Tresse et al., 2010). A final connection between changes in autophagic function and PDB patients' osteoclasts is the presence of nuclear inclusions, which may indicate a lack of protein clearance. Although all of these data indicate that autophagy may be dysregulated or defective in PDB, further work will be required to determine if this is related to abnormalities of

osteoclast activity or to the presence of nuclear inclusions that occur in the disease.

1.3 Hypothesis - Does *OPTN* gene mutations contribute to PDB pathogenesis?

Optineurin (OPTN) has been associated to PDB by an extensive GWAS by Albagha and co-workers, who have found an association between the SNP rs1561570 and PDB in British patients (Albagha et al., 2011). This strong association was also replicated in the French-Canadian population (Michou et al., 2012). However, little is known about the role of this gene in PDB.

1.3.1 Optineurin molecular structure

OPTN is a 67-kDa protein that was first isolated in 1998 in a yeast 2-hybrid screen (Li et al., 1998) and has been shown to have a strong homology to NF-κB essential molecule (NEMO/IKKγ), an important protein in the canonical NF-κB pathway, being named NEMO related protein (NRP) (Schwamborn et al., 2000). In 2002, optineurin or "optic neuropathy inducing" gene was identified as a strong candidate to be one of the leading causes of irreversible bilateral blindness worldwide - primary open-angle glaucoma (POAG) (Rezaie et al., 2002). The human *OPTN* gene is located in chromosome 10 (13099449-13138308) and spans about 39 kb of genomic DNA. It contains three noncoding exons in the 5' untranslated region (UTR) and 13 exons that code for a 577 amino acid protein. Alternative splicing in the 5' UTR generates at least four different transcripts (NM_001008211.1, NM_001008212.1, NM_001008213.1, and NM_001008214.1), but all have the same open reading frame. Alternative splicing in the coding region give rise to three different protein isoforms with 571 (ENST00000378764), 126 (ENST00000424614) and 107 (ENST00000486862) amino acids. In addition, two partial transcripts were also described, but there is no indication of being protein coding. Finally, in 2012 a 'new first exon' was described, upstream from the previously known exon 1, and was labeled as exon 1a (Michou et al., 2012).

The OPTN protein consists of a NEMO-like domain, a leucine zipper motif, multiple coiled-coil motifs, an ubiquitin binding domain (UBD), a microtubule associated

protein 1 light chain 3 (LC3)-interacting motif, and a C-terminal zinc finger (Figure 1.17) (Ying and Yue, 2012).

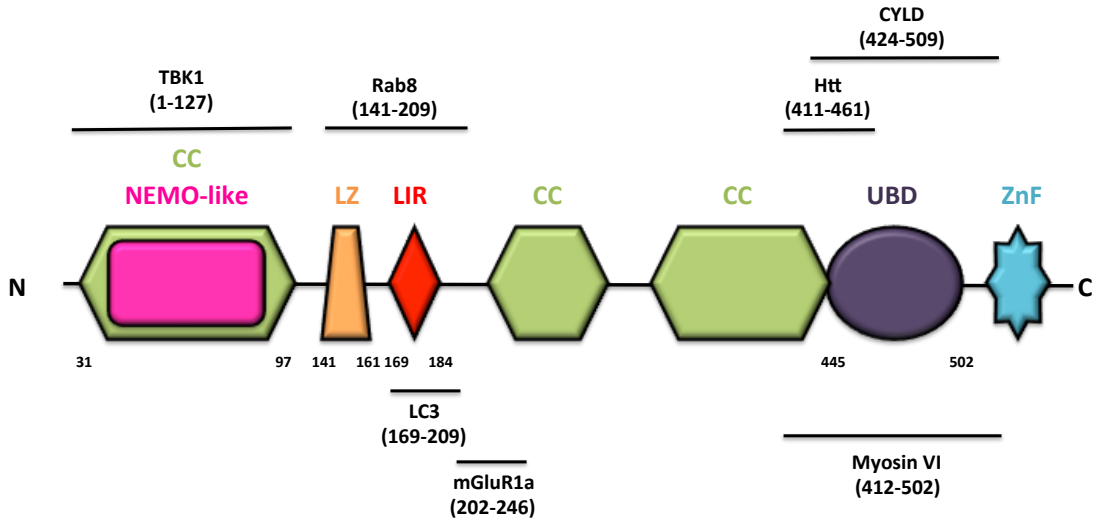


Figure 1.17 - OPTN protein structure and domains. OPTN has several domains such as coiled coil motifs (CC), an ubiquitin binding domain (UBD), a leucine zipper (LZ), a LC3 interacting region (LIR), and a zinc finger (ZnF). The black bars represent the sites of interactions of OPTN with various proteins, including Huntingtin (Htt), Rab8, Light chain 3 (LC3), Cylindromatosis (CYLD), metabotropic glutamate receptor (mGluR), myosin VI, TANK binding kinase 1 (TBK1), and receptor interacting protein 1 (RIP1). The numbers represent the position (in amino acids) in the OPTN protein. Adapted from (Kachaner et al., 2012).

OPTN is a cytosolic protein that is not secreted (Ying et al., 2010). It is ubiquitously expressed in many non-ocular tissues such as heart, brain, placenta, skeletal muscle, kidney, pancreas, adrenal cortex, liver, and the eye (Li et al., 1998; Rezaie and Sarfarazi, 2005; Rezaie et al., 2002). In the eye, OPTN is expressed in the trabecular meshwork, nonpigmented ciliary epithelium, and remarkably, also in retina (Rezaie and Sarfarazi, 2005; Schwamborn et al., 2000; Ying and Yue, 2012). OPTN seems to suffer posttranslational modifications. Since it possesses an ubiquitin binding domain (UBD), after being ubiquitinated it is processed through the ubiquitin-proteasome pathway (Ying et al., 2010). It is also phosphorylated at Ser177, which is adjacent to the LC3 interacting region (LIR) site (Wild et al., 2011b).

1.3.2 Optineurin functions

OPTN is a multifunctional protein involved in regulating several biological functions such as signal transduction, membrane vesicle trafficking, autophagy, NF-κB signaling, and cell survival (Kachaner et al., 2012; Ying and Yue, 2012). These

functions are mediated through interaction with a wide variety of proteins (Figure 1.17).

1.3.2.1 Role of optineurin in vesicular trafficking

Vesicular trafficking is one of the most fundamental biological processes of eukaryotic cells. As the name suggests, it refers to the movement of cargo packaged in the vesicles or cell organelles across the cytosol of the cell. It ensures supply of nutrients and signals to various compartments of the cell, crosstalk between the various organelles inside the cell, secretion and exocytosis (Jahn and Südhof, 1999; Mellman, 1996). Since OPTN interacts with multiple proteins like Rab8, huntingtin, myosin VI, or transferrin receptor (TfR), which are involved in various intra-cellular trafficking pathways, its role in vesicular trafficking is evident (Chalasani et al., 2009; del Toro et al., 2009; Hattula and Peränen, 2000; Sahlender et al., 2005; Vaibhava et al., 2012). But the exact mechanisms by which OPTN performs its functions in trafficking started to be uncovered only recently.

1.3.2.2 Role of optineurin in autophagy

Autophagy receptors are believed to play a crucial role in the selection and recruitment of cargo to autophagosomes by simultaneously binding to LC3 and ubiquitinated cargo (Kirkin et al., 2009b; Kraft et al., 2010; Yang and Klionsky, 2009). OPTN was identified as an autophagy receptor due to its ability to bind LC3 and ubiquitin directly and simultaneously through LIR and UBD domains, as described in section 1.2.3.3 (Figure 1.17) (Wild et al., 2011b). OPTN is involved in clearance of cytosolic Salmonella in macrophages (Wild et al., 2011b). However, so far no specific protein of Salmonella has been identified to bind to OPTN and be targeted to autophagosomes for degradation.

1.3.2.3 Role of optineurin in cell survival and cell death

Recently, using a mouse retinal ganglion cell line, RGC-5, it was shown that knockdown of endogenous *Optn* results in induction of apoptotic cell death due to reduced secretion of neurotrophin 3 (Nt3) and ciliary neurotrophic factor (Cntf) (Sippl et al., 2011). Addition of Nt3 rescued cell death. The levels of *Nt3* or *Cntf* mRNAs were not affected significantly upon knockdown of *Optn*. Rather, knockdown of *Optn*

resulted in collapse of the Golgi structure and accumulation of Nt3 positive vesicles due to a block in vesicle trafficking in the secretory pathway (Sahlender et al., 2005; Sippl et al., 2011). Overexpression of *Optn* sensitizes RGC-5 cells to TNFα induced cell death (Chalasani et al., 2007). Thus it appears that maintenance of optimum level of *Optn* is important for survival of RGCs.

1.3.2.4 Role of optineurin in antiviral signaling

Our body responds to viral infection through innate immune response and produces type I interferons (IFNα/IFNβ). These activate transcription of specific genes to produce an antiviral state in the cells (Müller et al., 1994). A tight regulation of this antiviral signaling is necessary to prevent unwanted tissue damage due to inflammatory response. OPTN has emerged as a negative regulator limiting IFNβ production in response to RNA virus infection (Mankouri et al., 2010). This negative regulation of IFNβ production is mediated by interaction of OPTN with TANK binding kinase 1 (TBK1), a protein kinase involved in the activation of interferon regulatory factor 3/7 (IRF3/7) (Fitzgerald et al., 2003). OPTN inhibits TBK1 mediated phosphorylation of IRF3 induced by Sendai virus or extracellular poly (I:C) (Sakaguchi et al., 2011), controlling the inflammatory response. But another group has suggested that OPTN is an activator of TBK1 and mediates IFNβ production in response to lipopolysaccharide (LPS) or poly (I:C) (Gleason et al., 2011). The UBD of OPTN plays an essential role in this process. It still remains to be clarified the role of OPTN in immune response, and if that depends solely on the immunogenic signal.

1.3.3 OPTN associated pathologies

1.3.3.1 Glaucoma

Glaucoma is the major cause of bilateral blindness and it is characterized by degeneration of the optic nerve, retinal ganglion cell death, and progressive axonal and visual field loss (Kerrigan et al., 1997; Kuehn et al., 2005; Quigley, 2011). Rezaie and coworkers in 2002 studied 54 families that suffered from autosomal dominant inherited adult-onset primary open-angle glaucoma (POAG), leading to the identification of the causative gene on chromosome 10p13, which they designated as optineurin. DNA sequence analyses detected four mutations in patients with POAG:

E50K, M98K, R545Q, and 691_692insAG (2-bp "AG" insertion) (Figure 1.18). These mutations were reported to be responsible for 16.7% of hereditary forms of glaucoma with an additional attributable risk factor of 13.6% in both familial and sporadic cases (Rezaie et al., 2002). In addition to the genetic link, a causal role for E50K in glaucoma was also referred by a recent study of Chi et al. (2010) that observed loss of RGC, reduction of retinal thickness, and excavation of the optic nerve head in E50K-expressing transgenic mice (Chi et al., 2010). Other *OPTN* gene alterations observed in various patient populations include: a missense mutation E322K in *OPTN* exon 10 in Chinese POAG family, H26D alteration in Japanese POAG families, and H486R in POAG and juvenile open-angle glaucoma families (Figure 1.18) (Funayama et al., 2004; Leung et al., 2003; Willoughby et al., 2004; Xiao et al., 2009).

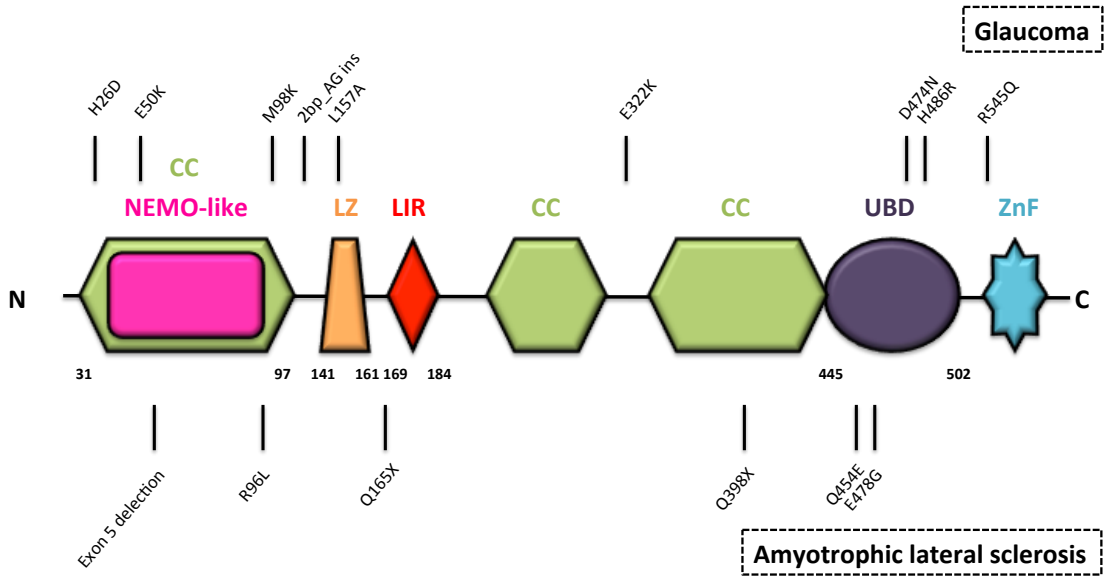


Figure 1.18 – Localization of glaucoma and ALS associated mutations in OPTN protein structure. Various glaucoma causing mutations were identified in the *OPTN* gene. Of these, R545Q and M98K are polymorphisms. Deletion of some of the exons has been found in ALS patients but not in glaucoma patients. CC- coiled coil, UBD- ubiquitin binding domain, LZ- leucine zipper, LIR- LC3 interacting region, ZnF- zinc finger. The numbers represent the position (in amino acids) for some domains in OPTN protein. Adapted from (Kachaner et al., 2012).

1.3.3.2 Amyotrophic lateral sclerosis (ALS)

ALS is a progressive disorder characterized by degeneration of motor neurons of the primary motor cortex, brainstem, and spinal cord (Leigh, 2007). It is a genetically heterogeneous disease. Besides *OPTN*, genes that encode TAR DNA-binding protein

of 43kDa (TDP-43) (Sreedharan et al., 2008), Cu/Zn superoxide dismutase (SOD-1) (Rosen et al., 1993), angiogenin (Greenway et al., 2006), vesicle-associated membrane protein (Nishimura et al., 2004), and VCP (Johnson et al., 2010) are reported to be associated with the classic familial ALS. Maruyama et al. in 2010 reported three mutations in the *OPTN* gene in Japanese familial or sporadic ALS patients: a homozygous deletion of exon 5, a homozygous nonsense mutation Q398X in exon 12, and a heterozygous missense mutation E478G in exon 14. The deletion of exon 5 causes a frame shift leading to the appearance of a premature stop codon, which would be expected to originate a peptide with 58 amino acids in length. Q398X mutation generates a premature stop codon at amino acid 398, truncating the 577 amino acid optineurin protein to one of 397 amino acids in length. This truncation results in a deletion of one of the coiled-coil domain. E478G mutation is located in the UBD domain. The Q398X nonsense mutation and probably the exon 5 deletion mutation as well could cause a decrease in *OPTN* expression resulting from nonsense-mediated mRNA decay of the transcripts carrying the nonsense mutations. The mechanism of recessive mutation causing ALS is expected to be loss of function. On the other hand, the E478G missense mutation increased the immunoreactivity for OPTN in the neural cells. The increased amount and different distribution of the mutated protein might disturb neuronal functions, and accelerate the formation of inclusion bodies in sporadic ALS (Maruyama et al., 2010).

In French familial ALS patients, a heterozygous nonsense 382_383insAG (2-bp AG insertion) mutation and a novel missense mutation R96L were reported (Millecamps et al., 2011). This mutation presumably induces a premature stop codon in exon 6. This will decrease the levels of functional OPTN and therefore is considered to be a loss of function mutation. The missense mutation R96L on the contrary might lead to a gain of function, although no accumulation of OPTN protein was detected in the patients' lymphoblasts. In sporadic ALS patients from a Dutch population, a nonsense Q165X and a missense Q454E mutations were identified. Since the Q165X mutation would probably result in a 72% truncated optineurin protein, binding to Rab8, mGluR1a, TFIIIA, Htt, and myosin VI might be eradicated. The Q454E mutation is located near the area that contains binding sites for Htt and myosin VI and may affect those interactions. However, these mutations were not detected in the Dutch familial ALS patients, suggesting that the genetic background of ALS may be different from one population to another population (van Blitterswijk et al., 2012).

1.3.3.3 Other neurodegenerative diseases

OPTN was shown to be located not only in the skein-like inclusions and round hyaline inclusions in ALS but also in the senile plaques and neurofibrillary tangles in Alzheimer's disease, Lewy bodies and Lewy neuritis in Parkinson's disease, ballooned neurons in Creutzfeldt–Jakob disease, glial cytoplasmic inclusions in multiple system atrophy, and Pick bodies in Pick's disease. This indicates that OPTN is often altered in neurodegenerative conditions (Osawa et al., 2011). The significance of such a finding, however, is unknown. Optineurin aggregates are generally found to be ubiquitin positive, and the OPTN aggregation may be the common process involved in neurodegeneration and cell death. OPTN may be itself an aggregation-prone protein present in the affected neurons and glia. Alternatively, OPTN could also be just secondarily entrapped in the inclusion bodies in various conditions or in ubiquitin (Turturro et al., 2014).

1.3.4 OPTN role in bone metabolism

As stated earlier, the role of OPTN in bone metabolism is not known to this date. The link between OPTN and bone was first suggested by a GWAS performed using a large cohort of PDB patients from the British population (Albagha et al., 2011). A strong association was described between a SNP located in *OPTN* and this disease, showing that when mutated *OPTN* can contribute to bone phenotypes. This association was later reproduced in the French-Canadian population, confirming its relevance. In this study our group showed that in the French-Canadian population the *C* allele was more frequent in healthy controls (present in 52%) while the *T* allele was more frequent in PDB patients (present in 64%) (p-value= 5.65×10^{-7}) (Michou et al., 2012). Since the *T* risk allele is also very common in healthy controls, this gene is more likely to contribute to pathogenesis with other factors than being by itself a disease-causing gene. Several studies have tried to explain the role of OPTN in bone, but the results obtained so far appear to be contradictory.

The role of OPTN in TNFα and NF-κB signaling was long suspected, when it was observed that this protein shares 53% similarity to NEMO (Schwamborn et al., 2000). TNFα was shown to induce *OPTN* expression (Li et al., 1998) and the role of OPTN in NF-κB signaling was shown by Zhu *et al.* (2007) who demonstrated that OPTN

acts as a negative regulator of TNFα induced NF-κB signaling, by binding to polyubiquitinated RIP (Zhu et al., 2007). In the presence of a signal as TNFα, OPTN interacts with CYLD, a deubiquitinase that negatively regulates NF-κB signaling. CYLD removes the ubiquitin from the polyubiquitinated receptor interaction protein (RIP) (Figure 1.19A) (Chalasani et al., 2009; Kovalenko et al., 2003; Sun, 2010; Trompouki et al., 2003). By interacting with CYLD and also with polyubiquitinated RIP, OPTN facilitates deubiquitination of polyubiquitinated RIP by CYLD, inactivating the NF-κB signaling (Nagabhushana et al., 2011). In the absence of OPTN, CYLD is unable to deubiquitinate RIP, leading to the accumulation of polyubiquitinated RIP, phosphorylation and activation of the IKK complex, resulting in enhanced NF-κB activity (Figure 1.19B). Thus, in NF-κB signaling, OPTN seems to act as an adaptor protein that brings together the enzyme CYLD and its substrate - the polyubiquitinated RIP (Figure 1.19) (Nagabhushana et al., 2011).

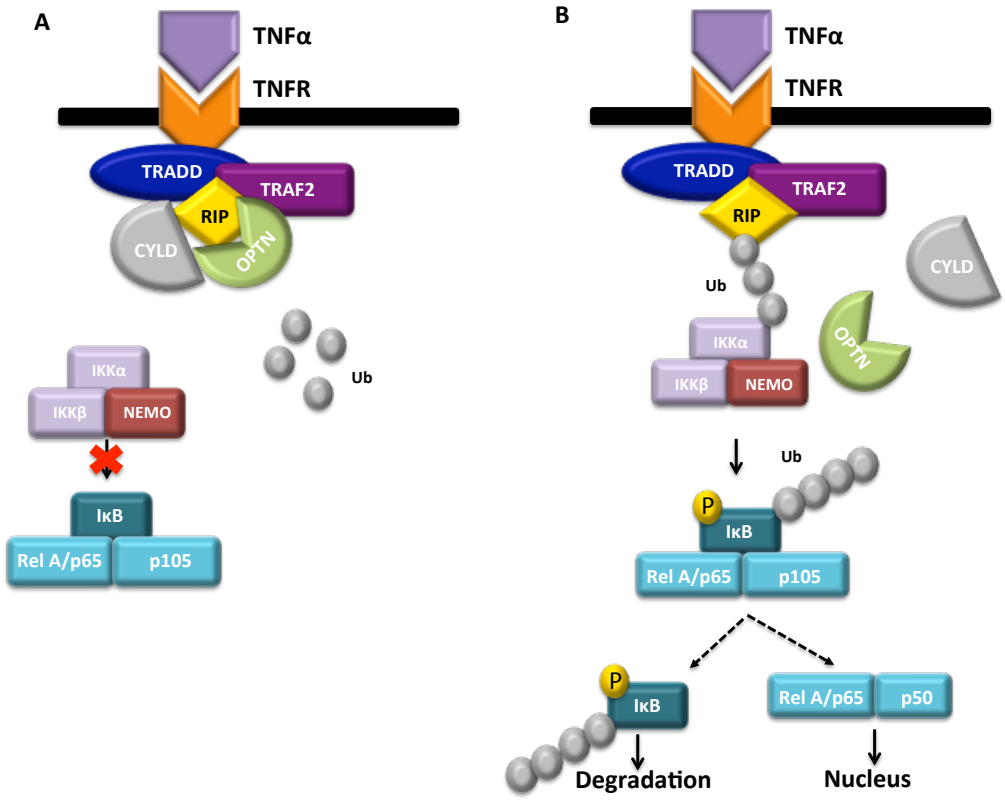


Figure 1.19 - Regulation of TNFα induced NF-κB signaling by OPTN. A) OPTN regulates RIP ubiquitination by acting as a competitive inhibitor of NEMO and binds to ubiquitinated RIP by displacing NEMO. OPTN then recruits CYLD (a deubiquitinase) to the molecular complex facilitating deubiquitination of polyubiquitinated RIP leading to downregulation of downstream pathway. B) In the absence of OPTN, after the binding of TNFα to its receptor and assembly of a multimolecular complex on TNF receptor that binds ubiquitin to RIP, NEMO is recruited, leading to activation of IKK complex. Active IKK phosphorylates IκB, which acts as a trigger for ubiquitination and degradation of IκB. This

leads to the release of p50/p65 complex of NF-κB to the nucleus leading to transcription activation. Adapted from (Zhu et al., 2007).

Concerning PDB, Obaid et al. used *in vitro* knockdown experiments to investigate the expression of *OPTN* during osteoclast formation. *OPTN* expression was not detected after the first day of RANKL treatment but there was a significant increase of its expression from days two to five of osteoclast differentiation. Then, after *OPTN* depletion using lentiviral particles expressing short hairpin RNA targeted against the *OPTN* gene, they found that the number of osteoclasts formed from *OPTN*-depleted cells were significantly higher than those formed from non-targeted cells (number ± SD; 151 ± 21 vs. 69 ± 18; P = 0.002) (Obaid et al., 2012).

Alternatively, functional studies also show that due to its high similarity to NEMO (Figure 1.20), OPTN could be competing for its functions. In terms of structure, the C-terminal NEMO UBD is highly similar to OPTN UBD, and some works showed that the UBD of NEMO can be replaced for the one in OPTN without interfering with its activity, which suggested that they could have similar functions and participate in similar pathways (Laplantine et al., 2009).

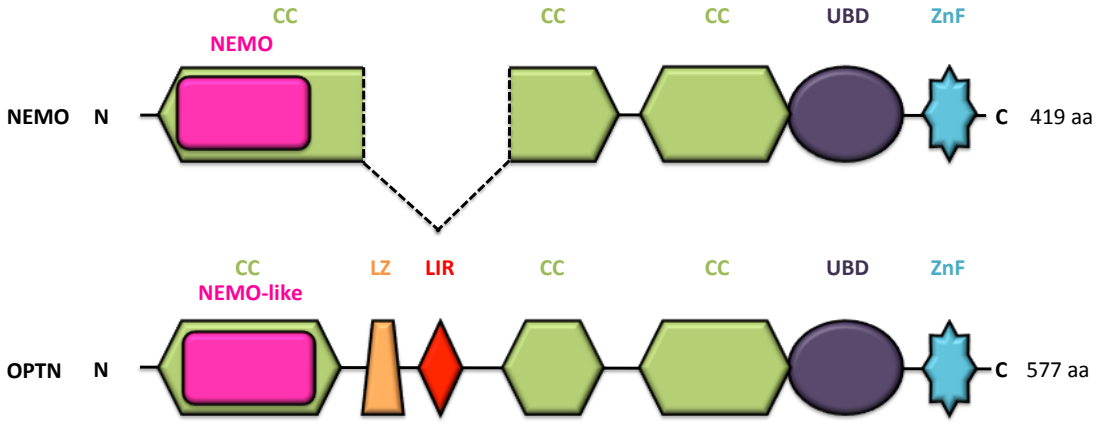


Figure 1.20 - NEMO and OPTN protein structure similarity. Human NEMO and OPTN proteins display strong homology between their domains including coiled-coil regions (CC), zing fingers (ZF), and ubiquitin binding domain (UBD) but differ by a 166 amino acid insert region present in OPTN that is absent from NEMO, and the presence of a leucine zipper (LZ) and a LC3 interacting region (LIR). The protein lengths are represented in amino acids (aa). Adapted from (Kachaner et al., 2012).

OPTN was also shown to cooperate with TAX1 binding protein (TAX1BP1) to potentiate the activation of NF-κB. In an extensive work, human T-lymphotropic virus type 1 (HTLV-1) was shown to encode regulatory and accessory proteins that are involved in viral replication and cell proliferation. Among them, TAX1 plays a

critical role by triggering cell immortalization through various mechanisms (Akagi et al., 1995), including activation of signaling pathways such as NF-κB (Sun and Yamaoka, 2005). This process occurs predominantly in the cytoplasm where TAX1 binds to NEMO and triggers the activation of IKKα and IKKβ (Figure 1.21) (Chu et al., 1999; Harhaj and Sun, 1999; Jin et al., 1999).

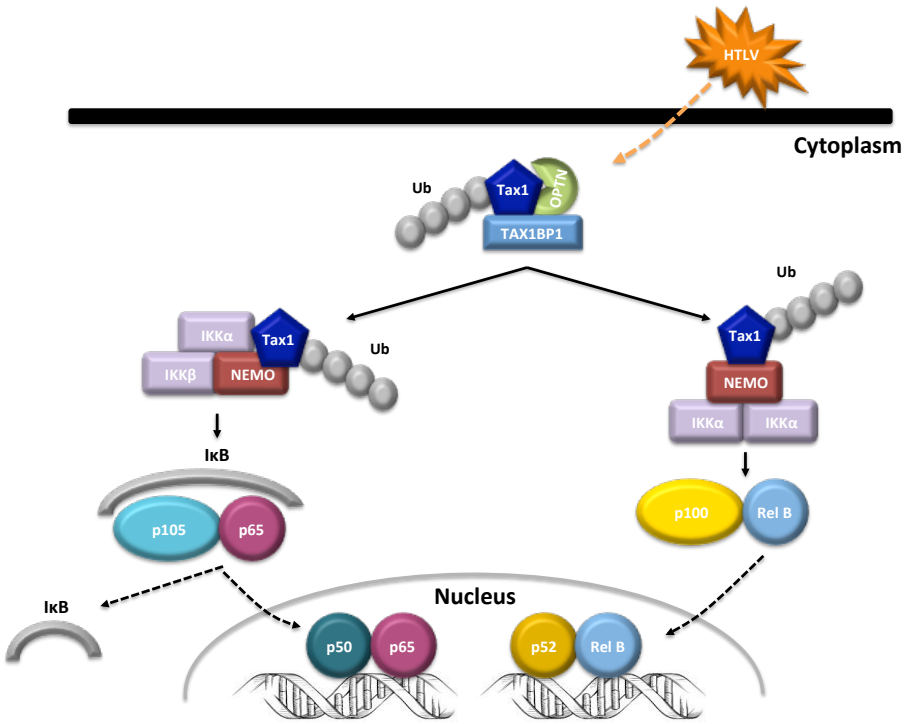


Figure 1.21 - Dysregulation of the NF-κB canonical and non-canonical pathways by TAX1/OPTN interaction. After HTLV-1 infection, dysregulation of the canonical pathway occurs, with the assembly of the TAX1/TAX1BP1/OPTN complex that is responsible for TAX1 ubiquitination. Interaction of ubiquitinated TAX1 with the cytoplasmic IKK complex specifically bound to the IKKγ subunit leads to phosphorylation of IκB, as well as its ubiquitination and subsequent degradation. NF-κB is subsequently activated and translocated into the nucleus. Adapted from (Journo et al., 2009).

To control this process, TAX1 binding protein TAX1BP1 recruits the A20 deubiquitinase and act as a negative control of TNFα, IL-1 and LPS-mediated NF-κB activation, suggesting that TAX1-dependent activation of NF-κB could also be more complex than originally thought (Shembade et al., 2007a). OPTN and TAX1BP1 were shown to form a functional complex with TAX1. Furthermore, this synergistic interaction between TAX1BP1 and OPTN contributes to TAX1 mediated NF-κB activation. OPTN UBD domain binds to TAX1 and facilitates its ubiquitination, allowing its activation and triggering the activation of the IKK system, phosphorylation of IκB and consequently NF-κB signaling. (Shembade et al., 2007b) Therefore this pathway might have a negative regulation mediated by a quaternary

complex containing TAX1BP1, A20, Itch and RNF11 (De Valck et al., 1999; Iha et al., 2008; Shembade et al., 2009, 2008), but TAX1 induced NF-κB activation is mediated by a ternary complex containing TAX1, TAX1BP1 and OPTN (Figure 1.21) (Journo et al., 2009), confirming that the presence of OPTN can also be a positive regulator of NF-κB translocation to the nucleus and activation of its target genes, namely those important for osteoclastogenesis.

1.4 Main Objectives

PDB is the second most frequent metabolic bone disorder after osteoporosis and, despite having a clear genetic background, *SQSTM1* has been so far the only gene which mutations are associated with this disorder. However, mutations in *SQSTM1* only account for around 50% of the reported familial cases, suggesting that other genes may be involved in PDB. After the GWAS performed by Albagha and coworkers (Albagha et al. 2010), a strong association between the PDB6 locus in chromosome 10 and PDB patients has confirmed previous linkage studies (Hocking et al., 2001; Lucas et al., 2008). Data from this GWAS indeed confirmed the previous linkage to PDB6 (10p13) locus, indicating that there is a high probability for the existence of a gene related to PDB in this locus (Albagha et al. 2010). But no PDB causal mutation was reported until the beginning of this thesis in this locus, and therefore PDB6 was considered to be a good 'target locus' to be further investigated. So the general objective of this work was to characterize the PDB6 (10p13) locus in PDB and attempt to identify possible molecular players/targets involved in its pathogenesis.

With this general goal in mind we divided our work plan in four main tasks:

1) Analysis of the PDB6 locus to search for the best candidate genes that could have a role in a bone disease, and screening by Sanger sequencing to search for variants that could have a genetic association with PDB, in a discovery group of 30 familial cases and 5 healthy controls from the French-Canadian population, all without *SQSTM1* mutations;

2) Selection of the most relevant variants and identification of disease-associated variants in a group of 225 unrelated PDB individuals and 298 healthy controls, all without *SQSTM1* mutations;

3) Perform functional analysis of the significant genetic variants using

bioinformatics approaches complemented by functional studies adjusted to each variant.

4) Perform functional analysis of rs1561570, the OPTN SNP most associated to PBD. Since our group had found a genetic association of the SNP rs1561570 in *OPTN* gene with PDB in the French-Canadian population (Frequency in patients – 36%; frequency in controls 52%; *p*-value = 5.65×10^{-7}; OR = 0.53; 95% CI = 0.42–0.69) (Michou et al., 2012), but no functional studies were available that could explain the role of this *OPTN* variant in PDB pathogenesis, we also aimed to analyze and predict how this important SNP may affect *OPTN* transcription or function.

CHAPTER 2

CANDIDATE GENE SELECTION AND SEQUENCING

2. CANDIDATE GENE SELECTION

2.1 PDB6 (10p13) locus candidate gene selection and Sanger sequencing

Iris A.L. Silva
Natércia Conceição
Jacques Brown
Édith Gagnon
Laëtitia Michou *
M. Leonor Cancela *

This chapter is part of a research paper submitted for publication.

Authors' roles: Study design: MLC and LM. Study conduct: IALS and NC. Data collection: IALS, JPB and EG. Data analysis: IALS, EG and NC. Data interpretation: IALS, LM, MLC and NC. Drafting manuscript: IALS. Revising manuscript content: NC, MLC and LM. Approving final version of manuscript: all authors.

2.1.1 Abstract

Paget's disease of bone (PDB) is a common metabolic bone disorder where genetic factors play an important role, but to date PDB causing mutations were only identified in the *Sequestosome* 1 (*SQSTM1*) gene at the PDB3 locus. PDB6 locus has been highly associated with PDB in several European populations and in the French-Canadian population, but no PDB causal mutations were reported in this locus. Therefore, PDB6 was considered a good 'target locus' for further investigation. Relevant candidate genes from PDB6 locus were selected based on their known or possible biological function in bone. For each selected gene, the coding region, splice sites, 5' and 3' untranslated regions and promoter were amplified, using an initial discovery sample of French-Canadian *SQSTM1* non-carriers PDB patients from 30 familial cases and 5 healthy controls. PCR-amplified products were purified and sequenced. Several variants were found and an *in silico* analysis was performed to explain the possible effect of each SNP. We screened *OPTN*, *PHYH*, *UCMA/GRP*,

SEPSH1, *CAMK1D*, and *CCDC3* and sequence analysis of our sample allowed us to detect 59 variants already described and 6 variants never reported to date. The *in silico* analysis showed that the majority of the SNPs could be related to alterations in bone biology possibly resulting in bone related diseases. Two variants were selected for association analysis (in 225 unrelated PDB patients and 298 healthy controls) due to their localization and frequency - rs3829923, rs2234968. The genetic association study demonstrated that only rs2234968 was associated with PDB. In conclusion, PDB6 appears to be a good target to be analyzed and this first step was crucial to identify some genetic variants in these six genes that could explain the association of this locus with PDB.

2.1.2 Introduction

Paget's disease of bone (PDB) is the second most common metabolic bone disorder, after osteoporosis (Kanis, 1998), affecting between 1% to 3% of individuals over the age of 55 years in white populations (Altman et al., 2000; van Staa et al., 2002). The disease is characterized by focal regions of highly exaggerated bone remodeling, with abnormalities in all phases of the remodeling process. The initial phase of PDB is characterized by excessive bone resorption in a focal region. Subsequently, bone formation is also markedly increased, with increased numbers of osteoblasts. The increased population of osteoblasts rapidly deposits new bone in a chaotic fashion so that the bone formed in pagetic lesions has poor quality and is disorganized. The poor quality of pagetic bone accounts for the bowing or even fracture of bones affected by PDB. As rapid bone formation predominates in the more advanced stages of PDB, the lesions become sclerotic, with observed replacement of the bone marrow with vascular and fibrous tissue and thickening of the bone (Hosking, 1981).

The cause of PDB is currently an area of intensive investigation, and both genetic and non-genetic factors have been implicated in the pathogenesis of this disease. The possibility that heredity might play a role in the pathogenesis of PDB was first raised in 1948 (Montagu, 1948). Over the recent years, there has been an increasing interest on the role of genetic factors in the pathogenesis of the disease (Laurin et al., 2001; Michou et al., 2012). Another argument that favors the genetic inheritance of PDB is that 15–40% of affected patients have a first-degree relative with PDB, and numerous studies have described extended families with PDB exhibiting an autosomal dominant

mode of inheritance (Hocking et al., 2000; Laurin et al., 2001; Morales Piga et al., 1995). Linkage studies in these families have identified a number of susceptibility loci on chromosomes *6p21* (PDB1) (Fotino et al., 1977), *18q21.1-22* (PDB2) (Haslam et al., 1998), *5q35* (PDB3), *5q31* (PDB4) (Laurin et al., 2001), *2q36* (PDB5) (Good et al., 2002), *10p13* (PDB6) (Hocking et al., 2001) and *18q23* (PDB7) (Good et al., 2002). Moreover, a genome wide association scan (GWAS) in British PDB families has suggested a linkage to the *10p13* (PDB6) locus (Hocking et al., 2001; Lucas et al., 2008). Recently, reanalysis of data from this genome wide scan confirmed a genetic association to PDB6 locus (Albagha et al., 2010), but no PDB causal mutation was reported to date in this locus, and therefore PDB6 could be a good 'locus target' to be further investigated.

The PDB6 locus has several genes that could have a functional role in PDB, and more interesting, this locus is conserved in all species analyzed, from zebrafish to human, suggesting an evolutionary conservation in their involvement in physiological pathways. In this work, we thus proposed to analyze which could be the most relevant genes in the entire PDB6 locus to the bone biology and assess the possible contributions of each significant variant identified in this locus to the functionality of those genes and their involvement in the pathophysiology of PDB.

2.1.3 Materials and Methods

2.1.3.1 Study participants

This study was approved by the CHU de Québec Ethics Committee and all participants have signed a consent form before inclusion in the study. Phenotype assessment comprised a complete bone evaluation, including total serum alkaline phosphatase, a total body bone scan and skull and pelvis X-rays. We investigated patients with familial form of PDB (one patient per family), unrelated PDB patients and healthy controls, all from the French-Canadian population. Clinical characteristics of these cohorts were previously published (Laurin et al., 2002; Michou et al., 2011; Morissette et al., 2006). For each individual, peripheral blood was obtained by venipuncture and DNA was extracted from blood samples, as already described in Michou et al. (2012). All patients and healthy donors studied here were non-carrier of the *P392L* mutation within the *SQSTM1* gene (PDB3 locus). RNA from total blood

was collected in a subset of patients and controls, on PAXgene tube (Qiagen) and RNA extraction was performed as previously described (Michou et al., 2012).

2.1.3.2 Candidate genes selection and sequencing

After screening the PDB6 locus (6721910 bp-18721437 bp), that spans approximately 26.7 cM between markers D10S189 (19 cM) and D10S548 (45.7 cM) on chromosome *10p13*, and contains over 70 genes, we selected six candidate genes for the sequencing step based on their biological function, genetic mouse models and tissue expression data. The genes selected were genes expressed in bone cells or their precursors and/or those involved in the NF-κB pathway, in the proteasome pathway, in autophagy or apoptosis, in bone cell function and/or survival, in cellular adhesion (*e.g*. integrins), in intercellular communication (cytokines), or in the pathogenesis of other bone metabolic diseases. To search for variants within the PDB6 locus, the exons of the candidate genes, their exon–intron boundaries and the basal promoters were sequenced as previously described (Michou et al., 2012). Thirty samples from patients with PDB and five healthy controls from the French-Canadian population underwent Sanger sequencing, which we refer as the discovery group. For the association study, the SNPs rs3829923 and rs2234968 were selected to be genotyped in a group of 225 patients and 298 controls, based on their frequency in the discovery group and location in the gene. The allele frequencies were calculated as previously described (Michou et al., 2012).

2.1.3.3 Bioinformatics analysis

To identify the potential effect of the variants found in the coding regions we used the online translate tool Expasy (http://web.expasy.org/translate/). To search for the predicted deleterious effect of the variants found in the coding regions we used SIFT (http://siftdna.org/www/Extended_SIFT_chr_coords_submit.html), Condel (http://bg.upf.edu/condel/analysis), Polyphen (http://genetics.bwh.harvard.edu/pph2/) and Mutation taster (http://www.mutationtaster.org/) software. Potential splice sites were predicted for the intronic variants using the Human Splicing Finder tool (http://www.umd.be/HSF/). To identify transcriptional factor binding sites (TFBSs) for the variants located in promoter and intronic regions we used TFsearch (http://www.cbrc.jp/research/db/TFSEARCH.html) and Consite

(http://asp.ii.uib.no:8090/cgi-bin/CONSITE/consite/). To identify miRNA binding sites for the variants located in the 3'UTR we used miRBase (www.mirbase.org). To analyze the effect in the mRNA secondary structure we used RNA fold (http://rna.tbi.univie.ac.at/cgi-bin/RNAfold.cgi). To search for SNPs in *linkage disequilibrium* (LD) we used the SNAP tool (http://www.broadinstitute.org/mpg/snap/).

2.1.4 Results

2.1.4.1 Discovery study – Variants identification in the candidate genes

The six best candidate genes at PDB6 locus were 1) *OPTN* (optineurin), 2) *UCMA/GRP* (upper zone of growth plate and cartilage matrix associated/gla-rich protein), 3) *CAMK1D* (calcium/calmodulin-dependent protein kinase ID), 4) *PHYH* (phytanoyl-CoA 2-hydroxylase), 5) *CCDC3* (coiled-coil domain containing 3) and 6) SEPHS1 (selenophosphate synthetase 1) (Table 2.1). Since *OPTN* and *UCMA/GRP* genes were already screened for variants in a previous work from our group (Michou et al., 2012), only the four later genes were analyzed. In this analysis, we have detected 65 genetic variants and their position is shown in Figure 2.1 and Table S2.1.

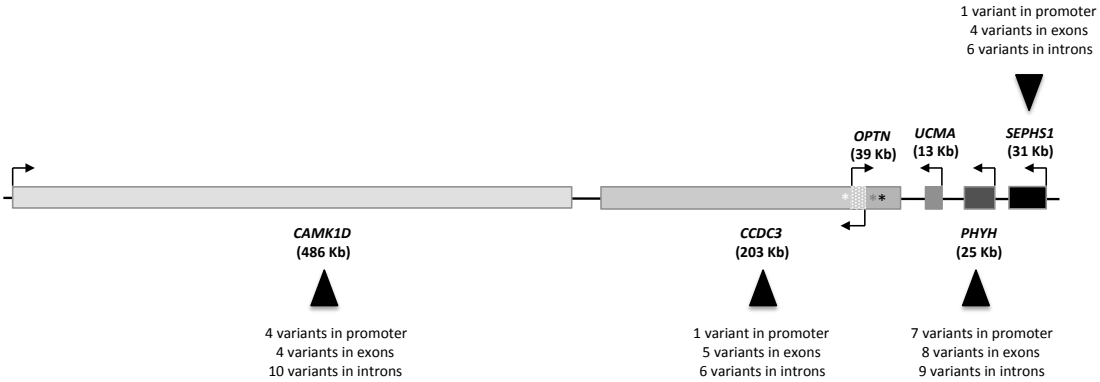


Figure 2.1 - Representation of the candidate genes in the PDB6 locus. Candidate genes in PDB6 locus of chromosome 10, including their names, size and orientation, are represented. The arrows indicate the transcription start site and orientation of each gene in the chromosome. Kb indicates the size in kilobases. The dotted region represents the localization of both overlapped *OPTN* and *CCDC3* first exons. The white * represents rs3829923, the grey * represents rs2234968 and the black * represents rs1561570.

Table 2.1 – Localization and function of the PDB6 genes selected in this study

Gene	Start (bp)	End (bp)	Strand	Ensembl ID	Description	Function	References
CAMK1D	12391481	12877545	+	ENSG00000183049	calcium/ calmodulin-dependent protein kinase 1D	**Activates MAP kinase MAPK3**. The p38/MAPK pathway has been implicated in **osteoclast differentiation or survival.** Phosphorylates **CREB1** pathway, **a known bone regulator.**	(Bergamaschi et al., 2008; Floudas et al., 2014; Li et al., 2002)
CCDC3	12938625	13141652	-	ENSG00000151468	coiled-coil domain containing 3	In mice, is highly **expressed in adipocytes.** It is suppressed by **TNFα, an important cytokine for bone biology.**	(Kobayashi et al., 2010)
OPTN	13141449	13180291	+	ENSG00000123240	optineurin	May function in **cellular morphogenesis and membrane trafficking, vesicle trafficking, and transcription** activation. Probably part of the **TNFα signaling pathway** that can shift the equilibrium toward **induction of cell death.** Its **upregulated by TNFα and NF-κB, two important players in bone biology.**	(Hattula and Peränen, 2000; Sudhakar et al., 2009; Zhu et al., 2007)
UCMA / GRP	13263767	13276331	-	ENSG00000165623	Gla-rich protein/ upper zone of growth plate and cartilage matrix associated	Involved in the **negative control of osteogenic differentiation** of osteochondrogenic precursor cells.	(Cancela et al., 2012)
PHYH	13319796	13344412	-	ENSG00000107537	phytanoyl-CoA 2-hydroxylase	Encodes a **peroxisomal protein that is involved in the α-oxidation of 3-methyl branched fatty acids**. Deficient protein activity has been associated with Zellweger syndrome (**dysmorphic craniofacia**l features) and rhizomelic **chondrodysplasia** punctata.	(Jansen et al., 1996; Wierzbicki et al., 2007)
SEPHS1	13359424	13390297	-	ENSG00000086475	selenophosphate synthetase 1	Encodes an enzyme that synthesizes selenophosphate from selenide and ATP. It seems to be regulated by **STAT3, a known bone regulator.**	("ENCODE Project" 2004; Low et al., 1995)

Their respective frequency in the discovery group and *in silico* predicted functions are shown in Table S2.2 (Appendices section). Together with the previous reported variants in *OPTN* and *UCMA/GRP*, we found 9 variants with predicted effects on protein sequence, 71 variants predicted to cause the loss and/or gain of TFBSs, 35 variants predicted to have an effect on splicing, and 50 variants predicted to cause the loss and/or gain of miRNA binding sites (Table S2.2).

2.1.4.2 Association study – Selection of the most relevant variants found in the discovery study

From the 65 variants found in the discovery study together with those previously described by our group (Michou et al., 2012), we selected rs3829923 (C/T) and rs2234968 (G/A) for genetic association with PDB in our French-Canadian cohort due to 1) the genomic localization; 2) the minor allele frequency (MAF) in the European population (25% and 18%, respectively) available in the dbSNP database and 3) the MAF obtained in our discovery sample. The variant rs3829923 was present in a large number of cases (MAF=43%) but it was not present in any control of our discovery group, which indicates that rs3829923 could be specific to PDB in our population. In addition, it is located in the *CCDC3* first intron (which correspond to the *OPTN* promoter), and *OPTN* has been strongly associated with PDB. The variant rs2234968 was present also in a large number of cases (MAF=53%) whereas the MAF reported in the European population is 18%. Also, this variant is located in the coding region of *OPTN* (exon 5) and could have an impact in OPTN protein. The genotypes of these variants were in Hardy-Weinberg equilibrium in controls (data not shown). The genetic association study of the two selected SNPs identified within the *OPTN* gene, demonstrated that only rs2234968 was associated with PDB (Table 2.2).

Table 2.2- Genetic association analysis of the *OPTN* gene of the two selected variants in 225 unrelated PDB-affected patients and 298 healthy controls from the French-Canadian population.

Variant	Minor allele	Minor allele frequency		Uncorrected *p*-value	Odds Ratio	95% CI
		PDB patients	Healthy controls			
rs2234968 G/A	A	35%	25%	6×10^{-3}	0.6674	[0.50 ; 0.89]
rs3829923 C/T	T	34%	34%	1	1.0006	[0.76 ; 1.32]

CI - Confidence Interval

2.1.5 Discussion

In the present study, we investigated the functional role of numerous genetic variants within the PDB6 locus in PDB, in particular in the bone related genes that can have a role in PDB pathophysiology, such as *OPTN*. We detected 59 variants already reported in the NCBI database and 6 variants never reported to date. After the association study performed using 225 unrelated PDB affected patients and 298 healthy controls only rs2234968 was shown to be associated with PDB. This SNP was selected for further studies regarding its functional effect on this bone disease. Since it is present in *OPTN* coding exon 5, we analyzed the possible effect of this variant in OPTN protein sequence, and the results showed that this is a synonymous variant that do not alter OPTN protein amino acid sequence (Table S2.2, Appendices section). However a more extensive work is necessary to analyze if it affects, for example, the interaction between OPTN and another partner protein. Also, our *in silico* analyses showed that this variant was predicted to cause an exon skipping due to a new branch point (Probability = 73%) (Table S2.2, Appendices section). That could be addressed through a PCR using cDNA from patients with both genotypes. This SNP can also change mRNA secondary structure and affect its stability, giving rise to lower levels of *OPTN* expression. Besides not being associated with PDB, rs3829923 functional effect should also be accessed, since it could give us some evidences of OPTN biological role and promoter activity. In this work this SNP was predicted to cause a change of E47, NRF2, E74A and SAP1 binding site. However transfection assays or functional studies such as electrophoretic mobility shift assay (EMSA) are needed to confirm its effect in *OPTN* promoter activity. In conclusion, the SNPs identified in this study indicate that rs2234968 is associated with PDB patients and that the association of PDB6 locus with PDB could be related with some of these genes that were shown to be involved with bone, namely with *OPTN*.

SUPPLEMENTARY TABLES

Table S2.1 – Identification and localization of the variants found in *PDB6* locus

Gene	ID variant	Localization in genome	Localization in gene	Nucleotide alteration
CCDC3 T1	rs3829923	13141144	exon 2 (5'UTR)	c. -641 G>A
	New variant	13141047	intron 2	c. -581+46 -/A
	New variant	13140511	intron 2	c. -580-14 C>T
	rs7093805	13115942	exon 4 (5'UTR)	c.-344G>T
	rs4750306	13115887	exon 4 (5'UTR)	c.-289A>G
	rs72779542	13115802	intron 4	c.-270+66T>A
	rs61851395	13115776	intron 4	c.-270+92C>G
	rs663177	13091801	intron 5	c. -1 + 127 A>C
CCDC3 T2	New variant	13043228	promoter	g. 13043228 G>T
	New variant	13042597	intron 1	c. 364 + 93 C>T
	rs116856516	12938874	exon 3	c. 797 C>A
	rs10752280	12938929	3' UTR	* 39 A>G
CAMK1D	rs11257730	12390561	promoter	g.12390561G>C
	rs11257731	12390563	promoter	g.12390563G>T
	rs7070092	12390653	promoter	g.12390653C>T
	rs11257733	12390767	promoter	g.12390767C>T
	rs145905575	12391618	exon 1 (5'UTR)	c .-200 G>A
	New variant	12391757	exon 1 (5'UTR)	c. -61 G>A
	rs2895524	12595147	intron 1	c. 93 - 77 C>T
	rs34194224	12595329	exon 2	c. 198 A>G
	rs1077745	12595427	intron 2	c. 225 + 72 G>A
	rs1077744	12595443	intron 2	c. 225 + 88 C>T
	rs139311431	12833088	intron 5	c. 566 - 69 C>G
	rs45497595	12833270	intron 6	c. 641 + 38 G>A
	rs1644417	12858372	intron 8	c. 833 + 45 T>C
	rs1644418	12858409	intron 8	c. 833 + 82 A>G
	rs2482023	12866356	intron 8	c. 834 - 108 A>C
	rs12768271	12866366	intron 8	c. 834 - 98 C>G
	rs1757051	12867598	exon 10	c. 948 C>G
	rs2482075	12870714	intron 10	c. 1039 - 54 A>G
PHYH T1	rs732704	13341889	promoter	g. 13341889 G>C
	rs511959	13341836	promoter	g. 13341836 C>G
	rs608353	13341809	promoter	g. 13341809 A>C
	rs732701	13341731	exon 1 (5'UTR)	c. -102 T>C
	rs1058596	13341579	exon 1 (5'UTR)	c. 50 C>G
	rs3802587	13341493	exon 1 (5'UTR)	c. 135 C>T
	rs3802586	13341378	intron 1	c. 208 + 42 G>A
PHYH T2	rs36143400	13342902	promoter	g. 13342902 G>A
	rs4615920	13342654	promoter	g. 13342654 T>C
	rs72781364	13342538	promoter	g. 13342538 G>A
	rs72781363	13342401	promoter	g.13342401 C>T
	rs76403928	13340375	intron 1	c. 75 - 130 A>C
	rs10796054	13340339	intron 1	c. 75 - 94 C>T

	rs7916926	13340293	intron 1	c. 75 - 48 T>C
	rs28938169	13340236	exon 2	c. 84 C>T
	rs72781358	13337699	intron 2	c. 134 - 93 C>T
	rs1747682	13337588	exon 3	c. 152 C>T
	rs473407	13330402	exon 6	c. 636 A>G
	rs7900830	13330291	intron 6	c. 678 + 69 C>A
	rs78560997	13330194	intron 6	c. 678 + 166 A>G
	rs648173	13325946	intron 6	c.679-107G>A
	rs62619919	13325784	exon 7	c.734G>A
	rs12411977	13320555	intron 8	c. 964 - 101 A>G
	rs11133	13320236	exon 9 (3'UTR)	* 65 G>A
	rs61851599	13390511	promoter	g. 13390511 G>A
	New variant	13387154	intron 1	c. -78 -125 A>G
	ESP_10_13386978	13386978	exon 2 (5'UTR)	c. - 28 G>A
	rs2275128	13378445	intron 3	c. 298-95 T>C
	rs62641683	13378273	exon 4	c. 375 C>G
SEPHS1	rs2275129	13378181	intron 4	c. 405+62 G>C
	rs10752297	13375906	exon 5	c .471 A>G
	rs61851591	13371855	intron 5	c. 561-67 G>T
	rs3740211	13370315	intron 7	c. 751+36 G>A
	rs62641685	13364914	exon 8	c. 885 G>C
	rs10906347	13361421	intron 8	c.965-65G>A

T1 - Transcript 1

T2 - Transcript 2

CHAPTER 3

FUNCTIONAL STUDY OF rs3829923 AND RV -9906

3. FUNCTIONAL STUDY OF rs3829923 AND RV -9906

3.1 Functional study of *OPTN* promoter variants found in Paget's disease of bone patients

Iris A.L. Silva
Natércia Conceição
Laëtitia Michou *
M. Leonor Cancela *

This chapter was submitted as a research paper for publication.

Authors' roles: Study design: MLC, LM and NC. Study conduct: IALS. Data collection: IALS. Data analysis: IALS and NC. Data interpretation: All authors. Drafting manuscript: IALS. Revising manuscript content: NC, MLC and LM. Approving final version of manuscript: all authors.

3.1.1 Abstract

Optineurin (*OPTN*) is a gene located in chromosome 10 that has been associated with several pathologies, including Paget's disease of bone (PDB). OPTN is an important protein for several biological mechanisms such as membrane trafficking, maintenance of the Golgi complex, autophagy, and it seems to be part of the TNFα pathway. However little is known about the role of this gene in bone and how this gene is regulated. Using DNA samples from our cohort of PDB patients we found two SNPs in *OPTN* promoter that could alter *OPTN* expression and give new insights about *OPTN* regulation - rs3829923 (already described in the previous chapter) and RV -9906 (a rare variant found in one PDB patient, described in a previous work from our group (Michou et al. 2012). By using a luciferase assay our results showed that rs3829923 *T* allele increases *OPTN* promoter activity due to a gain of E47 and E2F1 excitatory effect, and the rare variant RV -9906 was responsible for an increase of *OPTN* promoter activity due to a loss of SP1 inhibitory effect. In conclusion, our

work clarified the functional effect of these variants found in *OPTN* gene of PDB patients and how these variants could contribute to PDB pathophysiology.

3.1.2 Introduction

Optineurin (OPTN) is a very important protein for mechanisms such as membrane trafficking, maintenance of the Golgi complex and autophagy. In addition it seems to be part of the TNFα pathway. It is located in chromosome 10 in the positive strand and contains 4 non-coding exons and 13 coding exons, giving rise to 10 different transcripts due to splicing mechanisms (Michou et al., 2012). It encodes a 67-kDa cytoplasmic protein also called FIP-2 (14.7K-interacting protein 2) (Li et al., 1998) or NRP (NEMO-related protein) (Schwamborn et al., 2000) that contains several putative protein domains including a NEMO-like domain, multiple coiled-coil motifs, at least one leucine zipper, an ubiquitin-binding domain (UBD), a microtubule associated protein 1 light chain 3 (LC3)-interacting motif (LIR) (Wild et al., 2011b), and a carboxyl-terminal C2H2 type of zinc finger (Hattula and Peränen, 2000; Li et al., 1998; Schwamborn et al., 2000). *OPTN* mutations have been associated with normal tension glaucoma, a subtype of primary open-angle glaucoma, amyotrophic lateral sclerosis, Alzheimer's disease, Parkinson's disease, Creutzfeldt–Jakob disease, and Pick's disease (Maruyama et al., 2010; Osawa et al., 2011). In addition, *OPTN* mutations were identified as one of the genetic risk factor for Paget's disease of bone (PDB) (Albagha et al., 2010; Chung et al., 2010). PDB is the second most common bone disease after osteoporosis and it is characterized by the disruption of the balance between bone formation and bone resorption, leading to bone deformities (Ankrom and Shapiro, 1998).

The role of OPTN in PDB (or in bone metabolism) is not fully known yet, however some groups suggest that the interaction of OPTN and TAX1 may explain this association. TAX1 has been known to bind NEMO, a cytoplasmic protein involved in NF-κB pathway, triggering the activation of IKKα and IKKβ (Chu et al., 1999; Harhaj and Sun, 1999; Jin et al., 1999) and in turn the activation of NF-κB in the nucleus. Since OPTN is highly similar to NEMO, it could be also capable of activating NF-κB pathway by interacting with TAX1.

Regarding the mechanisms of transcriptional regulation of human *OPTN* gene, the data available indicates that *OPTN* can be induced by tumor necrosis factor α (TNFα), NF-κB and IFNγ (Sudhakar et al., 2013, 2009), however the region analyzed was localized between exon 1 and exon 2. The first non-coding exon of *OPTN* was only described in 2012 (Michou et al., 2012) and therefore there might be unidentified regulators of *OPTN* promoter activity that are upstream of *OPTN* first transcription start site.

In this work we analyze the functional effect of two variants - rs3829923 and RV -9906 -, located in *OPTN* promoter and discovered in a screening using DNA from PDB patients (Silva et al., unpublished results; Michou et al., 2012). By using an *in silico* approach we analyzed the putative transcription factors binding sites (TFBSs) located in those variants and studied their effect in *OPTN* basal promoter activity, unraveling new potential *OPTN* regulators.

3.1.3 Materials and Methods

3.1.3.1 Bioinformatics analysis

To search for putative TFBSs, we used TFsearch (http://www.cbrc.jp/research/db/TFSEARCH.html), Consite (Sandelin et al., 2004) and Genomatix MatInspector (Cartharius et al., 2005), the last with the default settings for the core and matrix similarity of 0.85. Two sequences of 1.2 kb spanning the regions between -309 bp and +846 bp, relative to the transcription initiation, that possesses both variants was used for this analysis. The TFBSs were confirmed using the data available in the ENCODE project (http://www.genome.gov/encode/).

3.1.3.2 Plasmid constructions

Four fragments of the *OPTN* promoter were amplified from human genomic DNA using four different sets of primers. A 1.2 kb fragment of the *OPTN* promoter ranging from -308 to +832 (related to the transcription start site) of the *OPTN* gene was amplified from human genomic DNA, using the primer set rs3829923 F1/R1 (Table S3.1). A point mutation, corresponding to the *T* allele of rs3829923 (position -305), was generated by PCR using the forward primer rs3829923 F2 (Table S3.1) and the same reverse primer. A 1.1 kb DNA fragment of the *OPTN* promoter ranging from -

241 to +832 was also amplified from human genomic DNA using the primer set RV -9906 F1/R1 (Table S3.1). A point mutation corresponding to the *A* allele of RV −9906 (position -232), was generated by PCR using a different primer forward (RV -9906 F2; Table S3.1) and the same primer reverse. The amplified PCR products were cloned into pCRII TOPO (Invitrogen) and the fidelity of the sequence was confirmed by DNA sequencing. The PCR products were then cloned between the *XhoI* and *KpnI* sites of the pGL3-basic luciferase reporter gene vector (Promega) giving rise to construct F1 (corresponding to rs3829923 *C* allele), construct F2 (corresponding to rs3829923 *T* allele), construct F3 (corresponding to RV -9906 *G* allele), and construct F4 (corresponding to RV -9906 *A* allele). All constructs were confirmed by DNA sequencing.

3.1.3.3 Cell culture conditions

The human embryonic kidney HEK293 cell line was grown in Dulbecco's modified eagle medium (DMEM), supplemented with 10% fetal bovine serum (FBS), 2mM L-glutamine and 1% penicillin/streptomycin (P/S). The cells were incubated at 37^{o}C in a humidified atmosphere containing 5% CO_2. The medium, FBS, antibiotics and glutamine were obtained from Invitrogen.

3.1.3.4 Cell transfections

HEK293 cells were plated on 24 well plates (5 x 10^4 cells/well) and then were transiently transfected using XtremeGENE HP (Roche), with 250 ng of *OPTN* reporter gene rs3829923 *C* allele (F1), rs3829923 *T* allele (F2), -9906 *G* allele (F3), -9906 *A* allele (F4), or empty reporter vector (pGL3-basic). A *renilla* luciferase reporter (Promega), 25 ng/well, was used as control for transfection efficiency. Co-transfection of an additional expression construct (50 ng) was performed using a similar approach. 48h after transfections, the cells were lysed and luciferase activity was assayed using a Dual-luciferase Reporter Assay kit in accordance with the manufacturer's instructions (Promega). All luciferase activities were normalized to the *Renilla* luciferase reporter pRL-TK Luc plasmid (Promega). All experiments were repeated at least three times. The SP1 expression plasmid and the RXR expression plasmids was the generous gift of Dr. Roland Schüle (Freiburg, Germany), the E47 expression plasmid was a generous gift from Dr. Xiao-Hong Sun (New York, USA),

the E2F family expression plasmids were a generous gift from Dr. Lieve Verlinden (Leuven, Belgium), and the SAP1 expression plasmid was a generous gift from Dr. Ralf Janknecht (Oklahoma, USA).

3.1.3.5 Preparation of human *in vitro* differentiated mature osteoclasts

Human mature osteoclasts were differentiated *in vitro* using mononuclear cells from blood of patients with PDB. Peripheral blood mononuclear cells (PBMCs) were obtained by density gradient centrifugation using Ficoll-Hypaque. The cells were resuspended ($3x10^6$ cells/mL) in OPTI-MEM (Life Tecnhologies) containing 10% FBS. The cell suspension was added to 6-well plates ($9x10^6$ cells/well) and to Lab-Tek 8 well-slides (Sigma-Aldrich) ($3\text{-}6x10^5$ cells/well). After 24h, the cells were washed thoroughly and lymphocyte-free adherent cells were incubated for three weeks with M-CSF (25 ng/mL, Life technologies) and RANKL (30 ng/mL, Fitzgerald).

3.1.3.6 Western blot analyses

Osteoclasts derived from patients and controls PBMCs were washed once in PBS and lysed using Trizol (Sigma-Aldrich). The protein concentrations were determined using the Bradford reagent (Bio-Rad). Proteins were separated by 8% SDS-polyacrylamide gels and transferred onto PVDF membranes (Life technologies). After transfer, the membranes were blocked in 0.1% TBS/Tween 20 containing 5% nonfat dried milk at room temperature (RT) for 1 h. The membranes were then incubated overnight (ON) at 4°C with anti-OPTN (1:500, #100000 Cayman), or anti-αTubulin (1:2000, #2144 Cell Signaling) in 0.1% TBS/Tween 20, followed by extensive washing using 0,1% TBS/Tween 20 and an incubation with HRP-conjugated secondary antibody (1:2000, Cell signaling) in 0.1% TBS/Tween 20 during 1h, at room temperature. After extensive washing with 0.1% TBS/Tween 20, specific proteins were detected using a chemiluminescence kit (GE Healthcare). The densitometric analysis was performed using ImageJ software.

3.1.3.7 Statistical analyses

All data were expressed as means ± standard deviation of measurements from at least three independent experiments. Comparisons between two groups were made using a two-tailed Student's t-test. For comparisons between multiple groups, one-way ANOVA followed by Tukey's post-hoc test was used. Differences were considered statistically significant when p-value < 0.05.

3.1.4 Results

3.1.4.1 *In silico* prediction of transcription factor binding sites in *OPTN* promoter

In order to explore the functional effect of the *OPTN* promoter SNPs, we searched for TFBSs that could be potentially interrupted by these variants, using the online tools Consite and TFSearch. Rs3829923 is a variant localized in *OPTN* promoter that causes a change from a *C* to a *T* allele (Figure 3.1A) that was described in a previous work (Silva et al., unpublished results). The bioinformatics software predicted the putative binding sites for SAP1, NRF2, E74A, and E2F family in the region of the rs3829923 carrying the *C* allele, and which are lost in the presence of the *T* allele. On the other hand, in the presence of the *T* allele the software predicted the putative binding site for E47 transcription factor (Figure 3.1B).

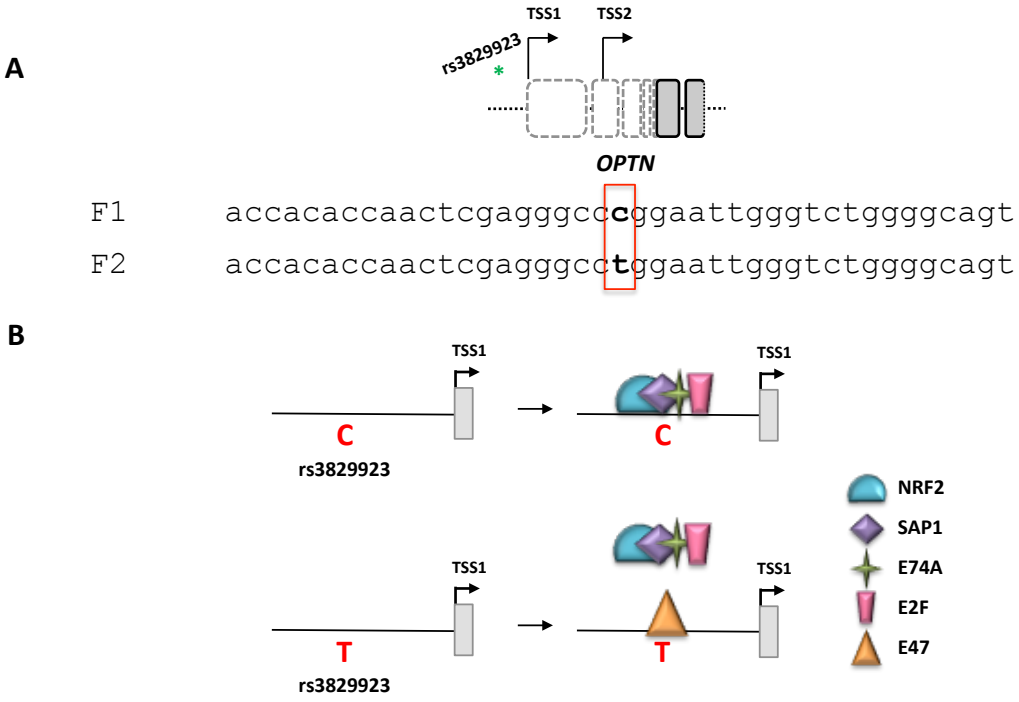


Figure 3.1 – *In silico* prediction of the effect of rs3829923. A) rs3829923 localization in *OPTN* gene showing that this variant is present in *OPTN* promoter and is responsible for a change from a *C* to a *T*. B) Bioinformatics prediction showing putative binding sites for NRF2, SAP1, E74A and E2F in the presence of the *C* allele, and a E47 putative binding site in the presence of rs3829923 *T* allele.

Finally, we assessed the effect of *OPTN* RV -9906 G/A (Figure 3.2A) polymorphism previously reported in *OPTN* promoter activity (Michou et al., 2012). The *in silico* analysis revealed that SP1, PAX4 and RXR binding sites are potentially interrupted by -9906 *A* allele, possibly resulting in loss of their binding efficiency as compared with the -9906 *G* allele. In addition, putative binding sites for HNF4 and PPAR were identified only in the presence of the *A* allele (Figure 3.2B).

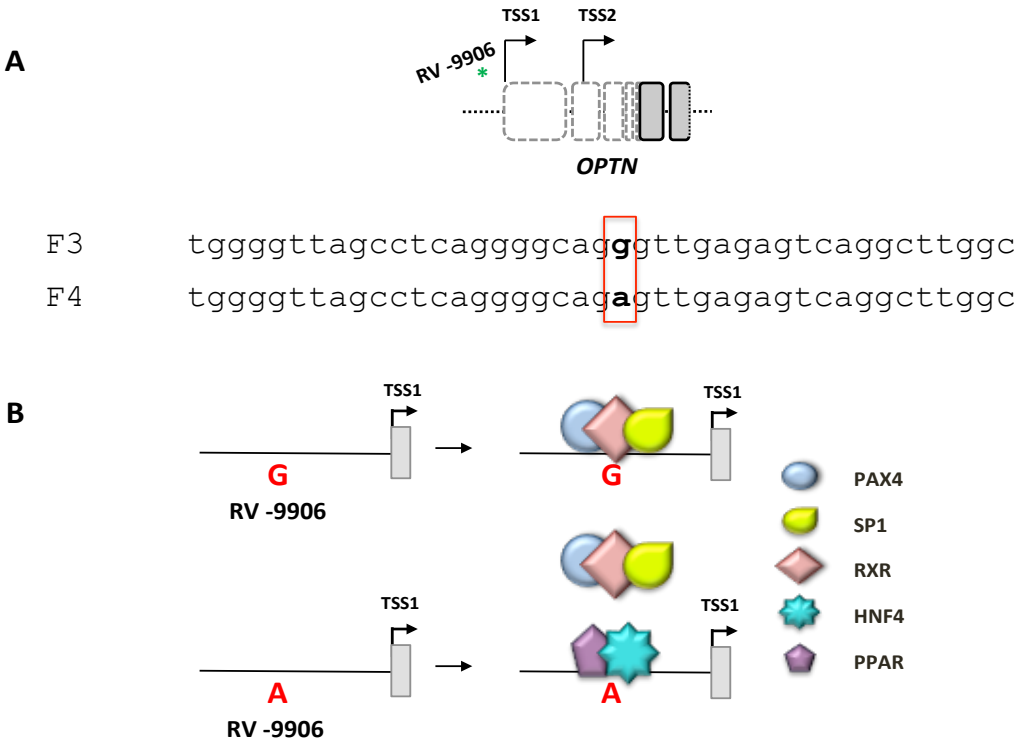


Figure 3.2 – *In silico* prediction of the effect of RV -9906. A) RV -9906 localization in *OPTN* gene showing that this variant is present in *OPTN* promoter and it is responsible for a change from a *G* to a *A*. B) Bioinformatics prediction shows the presence of PAX4, SP1, and RXR putative binding sites in the presence of the RV -9906 *G* allele and the presence of HNF4 and PPAR putative binding sites in the presence of RV -9906 *A* allele. TSS1 represents the first transcription start site and TSS2 represents the second transcription start site.

3.1.4.2 Functional analysis of rs3829923 in *OPTN* promoter

To analyze the functionality of the cloned constructs we performed a luciferase assay in HEK293 cells. The results showed an increase in *OPTN* promoter activity in construct F2 (carrying the *T* allele) when compared to construct F1 (carrying the *C* allele) (Figure 3.3).

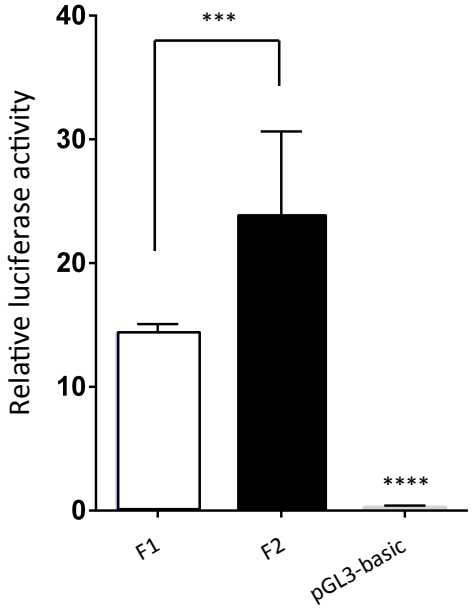


Figure 3.3 - rs3829923 effect in *OPTN* basal promoter activity. Transient transfections of HEK293 cells showing that the F1 construct (rs3829923 *T* allele) had a significantly higher activity than the construct F2 (rs3829923 *C* allele). Data are mean ± SD of at least three independent experiments. Significance was determined by one-way Anova. Asterisk *** represents a *p*-value < 0.001. **** indicates that the two constructs used are significantly different from the promoter less vector pGL3-basic (*p*-value < 0.0001).

The cells were then transfected with the F1 construct (carrying the *C* allele) or the F2 construct (carrying the *T* allele) together with the transcription factors (TFs) predicted in the *in silico* analysis. The co-transfection of F1 construct with E47 showed no effect of this TF when a *C* allele is present. The co-transfection of F2 construct together with E47 showed a significant increase in *OPTN* promoter activity, suggesting that E47 has a potentiating effect only when the *T* allele is present. (Figure 3.4A). These results confirm our *in silico* analysis. SAP1 seemed not to have a significant effect in both constructs used (Figure 3.4B). The presence of E2F3 resulted in a significant increase in reporter activity compared to the transfection conditions without the TF (*p*-value < 0.0001) in both constructs. However, in the presence of E2F1 there was an increase in *OPTN* promoter activity only when using F2 construct (Figure 3.4C). E2F2, E2F4 and E2F5 had no effect in both constructs used (Figure 3.4C).

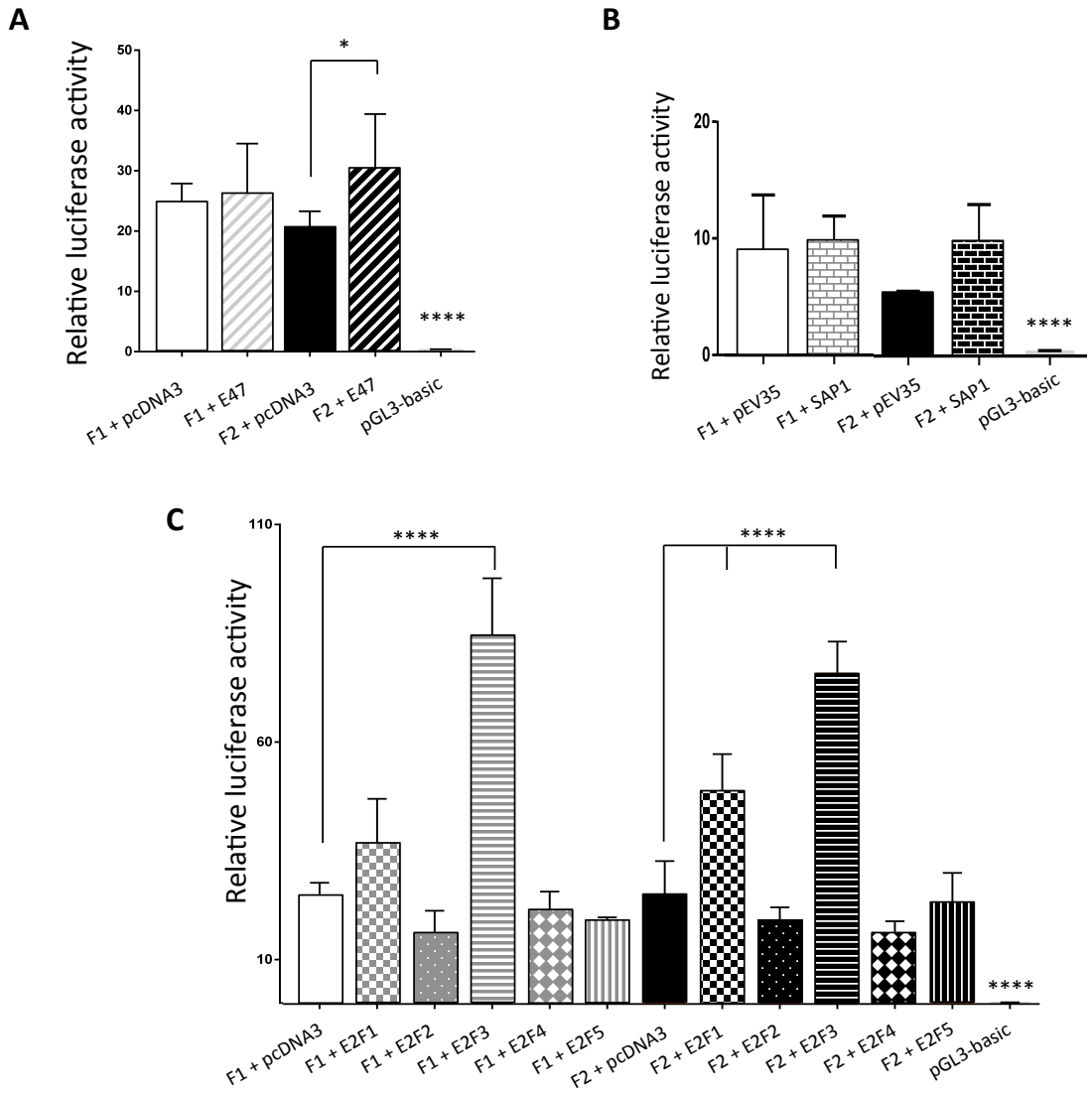


Figure 3.4 – Effect of E47, SAP1 and E2F family in *OPTN* promoter activity. A) Co-transfection assays showing that F2 construct (carrying the *T* allele) had a significant increase of luciferase expression, in the presence of E47 not observed with F1 construct (carrying the *C* allele. B) Co-transfection with SAP1 revealed that this transcription factor does not have an effect in the presence of rs3829923 *C* or *T* alleles. C) Co-transfection assays showing that E2F3 have an activating effect in both F1 and F2 and that F2 (*T* allele) had a significant increase of luciferase expression in the presence of E2F1 that was not observed with F1, indicating that the effect of E2F1 is specific of the *T* allele. EF2, EF4 and EF5 did not have a significant effect on both constructs. Data are mean ± SD of at least three independent experiments. Significance was determined by one-way Anova. Asterisk * indicates a *p*-value = 0.01; **** indicates a *p*-value < 0.0001.

These results indicate that E47 and E2F1 only have an effect in the presence of the *T* allele while E2F3 has an effect in *OPTN* promoter but that effect is independent of rs3829923 *C/T* variant. Actually, E2F family have three additional putative binding sites in *OPTN* gene fragment used in transfection that could also contribute to explain that effect (Figure 3.5).

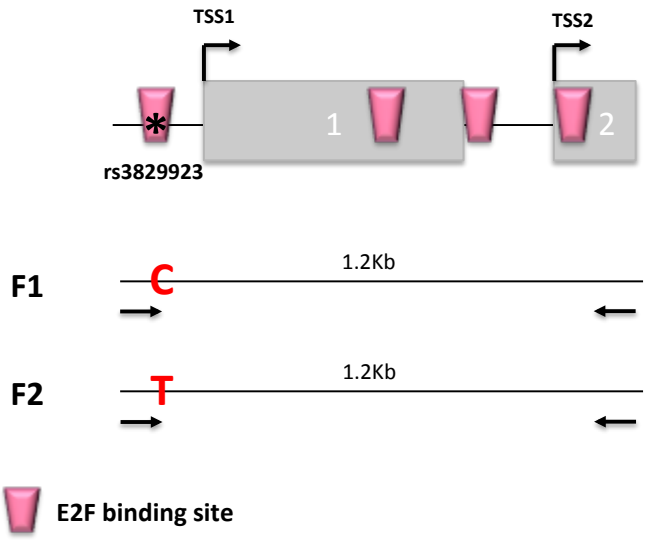


Figure 3.5 – E2F binding sites in F1 and F2 constructs. The *in silico* analysis performed using constructs F1 and F2 shows that E2F family have three more putative binding sites that can explain E2F3 effect in the constructs used. TSS1 represents the first transcription start site. TSS2 represents the second transcription start site. The length of each construct is represented above each construct in kilobases (kb).

In order to access if this positive effect of rs3829923 *T* allele in *OPTN* promoter activity was translated into an increase in *OPTN* expression in patients cells, we performed a western blot analysis using osteoclasts derived from PDB patients peripheral blood mononuclear cells (PBMCs) that were differentiated *in vitro* as described in the material and methods section. However, the *in vitro* effect seen in the transfection assay was not reproduced in patients' osteoclasts, since the presence of rs3829923 *T* allele did not affect the OPTN protein levels (Figure 3.6).

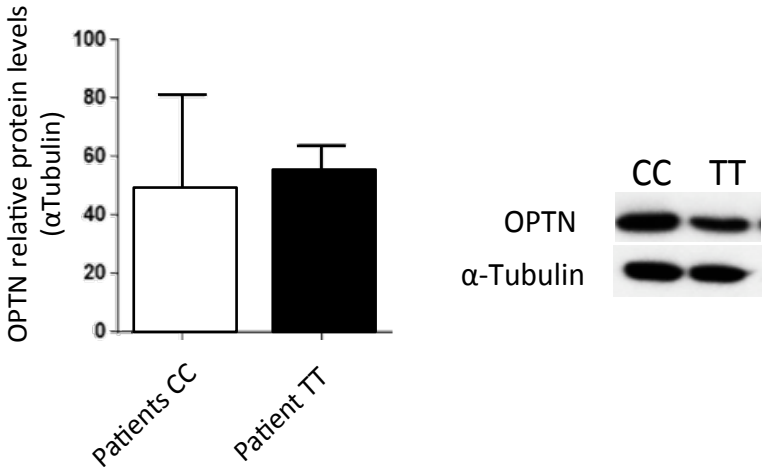


Figure 3.6 - rs3829923 effect in OPTN protein expression. Analysis of OPTN protein expression in PBMCs-derived osteoclasts from non-mutated patients (*CC*) (n=2) and one patient homozygous for

rs3829923 (*TT*). Data are mean ± SD of at least three independent experiments. Significance was determined by a two-tailed Student's t-test.

3.1.4.3 Functional analysis of the rare variant RV -9906 in *OPTN* promoter

The transient transfections of HEK293 cells with constructs F3 and F4 showed that the construct with the *G* allele (construct F3) had a significantly lower activity than the construct with the *A* allele (construct F4) (Figure 3.7).

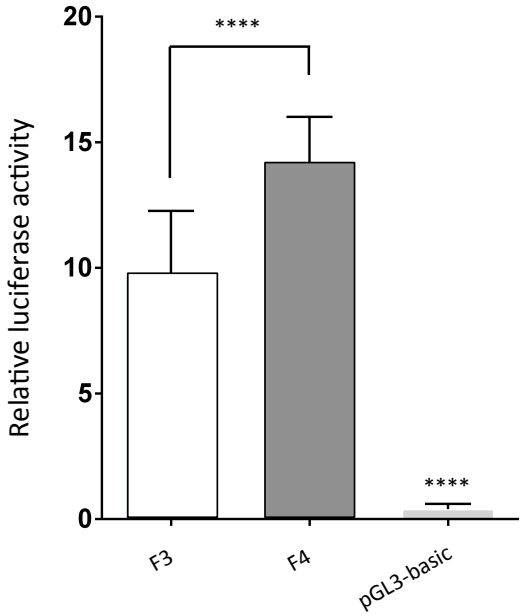


Figure 3.7 - RV -9906 effect in *OPTN* basal promoter activity. Transient transfections of HEK293 cells showing that the F3 construct (carrying RV -9906 *A* allele) had a significantly higher activity than the F4 construct (carrying the RV -9906 *G* allele). Data are mean ± SD of at least three independent experiments. Significance was determined by one-way Anova. Asterisk *** represents a *p*-value < 0.001. **** indicates that both constructs are significantly different from the promoter less vector pGL3-basic (*p*-value < 0.0001).

These results suggest that the presence of the *A* allele (F4 construct) increases the transcriptional activity of the *OPTN* gene. By transient co-transfections we showed that SP1 reduces luciferase expression only in the presence of the F3 construct (*G* allele) (Figure 3.8A), indicating that the effect obtained is specific for the construct containing the *G* allele.

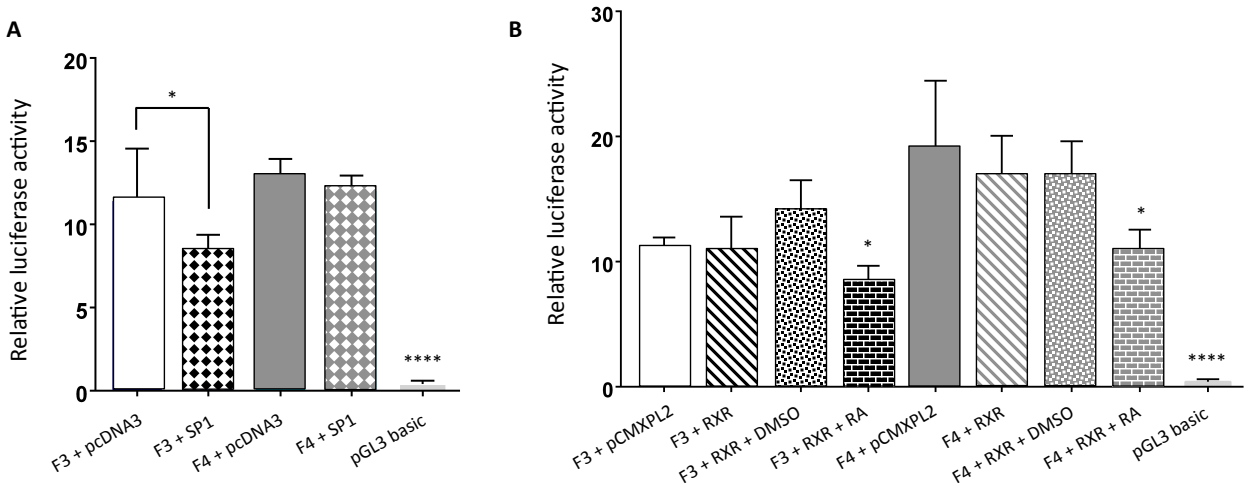


Figure 3.8 – Effect of SP1 and RXR in *OPTN* promoter activity. A) Transient co-transfections of HEK293 cells showing that F3 construct (containing the *G* allele) had a significant inhibition of luciferase expression, in the presence of SP1 not observed with the F4 construct (containing the *A* allele). B) Co-transfections with RXR revealed that this transcription factor had an inhibitory effect in the presence of F3 and F4. Data are mean ± SD of at least three independent experiments. Significance was determined by one-way Anova. Asterisk * indicates that each construct (F3 and F4) transfected with RXR and its ligand (retinoic acid (RA)) is significantly different from their controls (p-value < 0.05). Asterisk **** indicates that all the construct used are significantly different from the promoterless vector pGL3-basic (p-value < 0.0001)

After co-transfections of HEK293 cells with F3 and F4 constructs and with RXR, we observed a decrease in *OPTN* promoter activity with both constructions when the cells were treated with the RXR ligand retinoic acid (RA) (Figure 3.8B), which indicates that, in fact, RXR has a binding site in *OPTN* promoter but that binding is not altered with RV -9906 G/A variant. Actually, RXR has two additional putative binding sites in *OPTN* gene fragment used in these transfection assays that could also contribute to that effect (Figure 3.9).

The effect of RV -9906 in OPTN protein levels was not assessed in the PBMCs from PDB patients because this rare variant was present in one unique PDB patient that was lost to follow up.

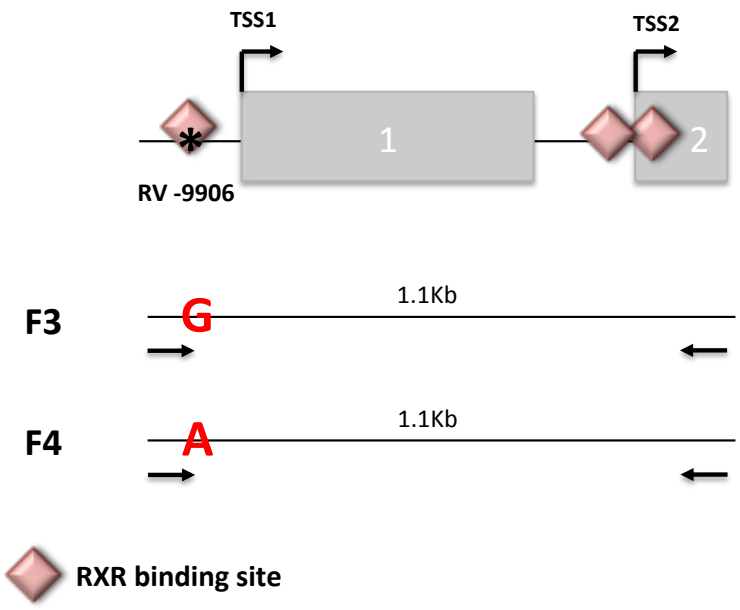


Figure 3.9 – RXR binding sites in F3 and F4 constructs. The *in silico* analysis performed using constructs F3 and F4 sequences shows that RXR have two more putative binding sites that can explain the complex RXR/RA effect in the constructs used. TSS1 represents the first transcription start site. TSS2 represents the second transcription start site. The length of each construct is represented above each construct in kilobases (kb).

3.1.5 Discussion

In the present study we analyzed two variants in *OPTN* promoter that were found in a previous work using a group of 30 familial cases of PDB and 5 healthy controls. One - rs3829923 - was found in 15 cases and not found in any of the controls, the other - RV -9906 - was a rare variant found in one unique PDB patient. Since these variants were located in *OPTN* promoter it was hypothesized that they could alter *OPTN* expression and in that case its functional study would potentially unravel new *OPTN* regulators. Rs3829923 was shown to increase *OPTN* promoter basal activity and following co-transfections, we concluded that E47 and E2F1 have an excitatory effect only in the F1 construct (carrying the *C* allele). E47 is a member of bHLH family and it was described as important for osteoclast maturation and survival (Long et al. 2012). Therefore the gain of a E47 binding site could explain the PDB phenotype since it might be related to an increase in osteoclast survival, giving rise to PDB features. E2F1 has been associated with osteoblast differentiation and mineralization (Yu et al., 2013), and cell proliferation (Berman et al., 2008), which may partially explain the deregulation in the number of osteoblasts (and also osteoclasts) in this

disease. We also analyzed a rare variant found in the *OPTN* promoter (RV -9906) in a unique PDB patient (Michou et al., 2012). Following co-transfections in HEK293 cells with the *OPTN* promoter carrying either the *G* or the *A* allele with SP1 and RXR TFs, we were able to conclude that RV -9906 was in fact responsible for a loss of SP1 inhibitory effect whereas it did not alter the RXR effect. Polymorphisms occurring in SP1 binding sites were previously associated with changes in bone mineral density and osteoporosis (Grant et al., 1996) and SP1 was already described as an important regulator for bone related genes, like NF-κB (Liu et al., 2005) and BMP2 (Xu and Rogers, 2007). SP1 seems to be an important regulator of *OPTN* promoter, since a single alteration in its sequence changed its activity probably due to a loss of SP1 binding site. Taking into account our results on the inhibitory effect of SP1 in RV -9906 *G* allele and the activating effects of E47 and E2F1 in rs3829923 *T* allele - giving rise to higher levels of *OPTN* expression - and the fact that OPTN was described as an important protein for the degradation of NF-κB inhibitor IκB (Journo et al., 2009), our hypothesis is that in normal conditions SP1 binds to *OPTN* promoter inhibiting its activity, giving rise to "normal" levels of *OPTN* expression. With RV -9906 *A* allele this binding site is removed, increasing the *OPTN* promoter activity, which will increase the expression of *OPTN* in this PDB patient. E47 and E2F1 only bind to *OPTN* promoter increasing its activity when the rs3829923 *T* allele is present, giving rise to higher levels of *OPTN* expression. An increase in *OPTN* expression by these mechanisms will give rise to an increase in NF-κB in the nucleus and a consequent increase in osteoclastogenesis-related genes transcription, explaining the PDB phenotype (Figure 3.8).

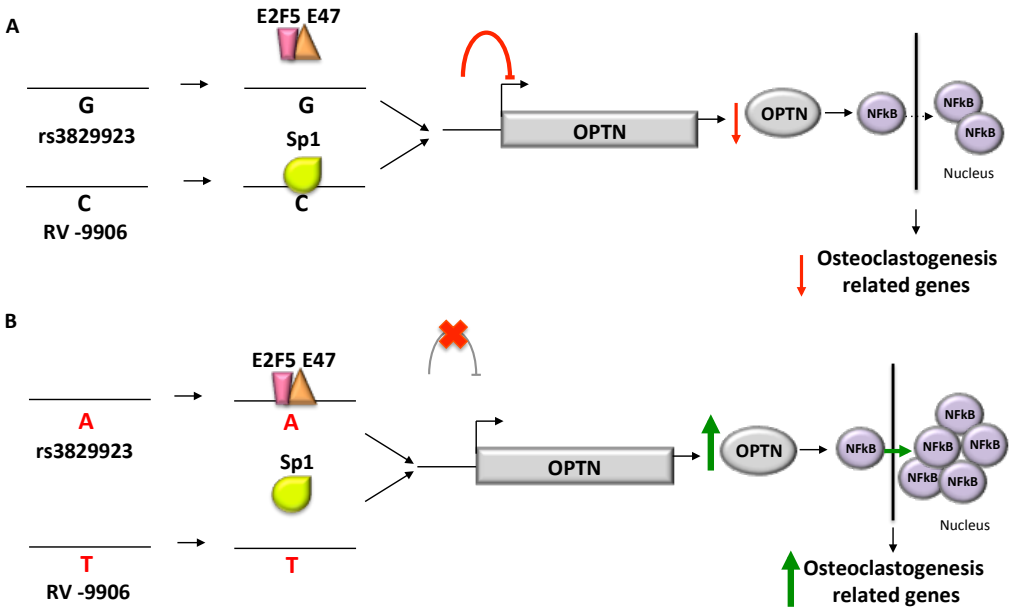


Figure 3.8 - Hypothetical functional effect of rs3829923 and RV -9906 in PDB. A) In the presence of the non-mutated allele in each position there is a decrease in *OPTN* promoter activity due to the inhibitory effect of SP1 in RV -9906. This leads to a decrease in *OPTN* expression levels, and since OPTN is related with the translocation of NF-κB to the nucleus that will give rise to lower levels of osteoclastogenesis related genes transcription. B) In the presence of the *A* allele in RV -9906 the inhibitory effect of SP1 is lost and therefore there is an increase in *OPTN* expression. In the presence of the *T* allele in rs3829923 there is a new binding site for E2F1 and E47, which will also increase *OPTN* expression. An increase in *OPTN* expression will favor the translocation of NF-κB to the nucleus and the transcription of genes important for osteoclastogenesis, explaining the PDB phenotype that is characterized by an increase in the number of osteoclasts.

To extend this analysis of the variants in *OPTN* promoter an electrophoretic mobility shift assay (EMSA) should be performed in order to confirm if in fact SP1, E47 and E2F1 are binding to the cloned promoters. A chromatin immunoprecipitation assay (CHIP) could also be performed aiming at providing additional data to confirm the presence of these *bona fide* TFBSs *in vivo*. In addition, to clarify the role in PDB of the promoter variants identified in this study and its effect in osteoclastogenesis, it would be of interest to do *in vitro* site directed mutagenesis and develop osteoclast precursor isogenic cell lines with and without the mutations, aiming at replicating the expression of the two identified transcript variants and analyzing its effect in the number and activity of resulting osteoclasts.

CHAPTER 4

FUNCTIONAL STUDY OF rs2234968 AND rs1561570

4. FUNCTIONAL STUDY OF rs2234968 AND rs1561570

4.1 Functional study of Paget's disease of bone associated variants found in the PDB6 locus

Iris A.L. Silva
Natércia Conceição
Jacques Brown
Édith Gagnon
Laëtitia Michou *
M. Leonor Cancela *

This chapter has been submitted as a research paper for publication.

Authors' roles: Study design: MLC and LM. Study conduct: IALS and NC. Data collection: IALS, JPB and EG. Data analysis: IALS, EG and NC. Data interpretation: IALS, LM, MLC and NC. Drafting manuscript: IALS. Revising manuscript content: NC, MLC and LM. Approving final version of manuscript: all authors.

4.1.1 Abstract

PDB6 locus was linked to Paget's disease of bone (PDB) in previous works and it contains at least six candidate genes for metabolic bone diseases such as PDB. In our previous work we screened *OPTN*, *PHYH*, *UCMA*, *SEPSH1*, *CAMK1D*, and *CCDC3* and performed an association study to the most important variants found. From the ones selected only rs2234968 was found to be associated with PDB. Also, the common variant rs1561570, within the *OPTN* gene, was previously shown to be strongly associated with PDB in several populations but no functional studies were available to this date. Therefore we aim to unravel the functional effects of these two PDB associated variants in *OPTN* expression and PDB pathophysiology. Rs2238968 functional effect may be explained by rs10906303 and rs79529484, found to be in *linkage disequilibrium* with this SNP. Both SNPs were predicted to create new

acceptor splice sites that may give rise to predicted truncated proteins, probably non-functional. In this study we also showed that the *C* allele of rs1561570 was methylated, and it was also responsible for lower levels of *OPTN* gene expression when compared to the non-methylated *T* allele. In addition, we showed that patients with the *T* allele and cells treated with a demethylating agent have higher rate of translocation of NF-κB to the nucleus and higher expression levels of its target genes (such as *NFATc1* and *IL-6*). The patient carriers of the *T* allele also presented a higher osteoclast differentiation rate, and osteoclasts with a higher number of nuclei and higher bone resorbing capacity. All together, these results show that higher levels of *OPTN* can give rise to higher levels of NF-κB in the nucleus, increasing the expression of its target genes expression, which may explain *OPTN* contribution to PDB. In conclusion, our work clarified the functional effect of rs1561570 and rs2234968 within *OPTN* gene in PDB, giving a new perspective about *OPTN* variants and also indicative of a new role of OPTN in bone biology.

4.1.2 Introduction

Paget's disease of bone (PDB) is one of the most frequent metabolic bone disorders (Kanis, 1998), affecting between 1% and 3% of individuals over the age of 55 years in Caucasians (Altman et al., 2000; van Staa et al., 2002). Both genetic and non-genetic factors have been implicated in this disease. Around 15–40% of affected patients have a first-degree relative with PDB, and numerous studies have described extended families with PDB exhibiting an autosomal dominant mode of inheritance (Hocking et al., 2000; Laurin et al., 2001; Morales Piga et al., 1995). Recently, a genome wide scan in British families with PDB has shown a linkage to the *10p13 (PDB6)* locus (Albagha et al., 2010; Hocking et al., 2001; Lucas et al., 2008), however no PDB causal mutation has been reported to date in this locus. A hot-spot in the GWAS that linked PDB6 locus to PDB was localized in the vicinity of a genetic variant located in the *Optineurin* (*OPTN*) gene (Albagha et al., 2010). *OPTN* gene was previously linked to glaucoma (Rezaie et al., 2002) and neurodegenerative diseases like Alzheimer's disease, Parkinson's disease or amyotrophic lateral sclerosis (Osawa et al., 2011). Besides further replicating the strong and statistically significant genetic association within the *OPTN* gene of rs1561570 (p-value = 5.65×10^{-7}) with PDB in the French-Canadian population (Michou et al., 2012), our group also

identified a functional SNP in OPTN exon 5 with no predicted effect in the protein sequence but highly associated with PDB (*p*-value = 6 x 10^{-3}), already described in chapter 2. OPTN biological role is not yet fully understood but OPTN seems to be involved in NF-κB pathway (Agou et al., 2004). This protein contains several putative domains including a NEMO-like domain, that is present in proteins that are part of IκB kinase (IKK) complex (Ying and Yue, 2012). IKK modifies the inhibitors of NF-κB transcription factor promoting their degradation. The free NF-κB molecules can then enter the nucleus where they activate specific target genes (Agou et al., 2004; Bianchi and Meier, 2009). Also, recently, OPTN was characterized as an autophagy adaptor protein, which regulates selective autophagy of ubiquitin-coated cytosolic *Salmonella enterica*. This function depends on the phosphorylation of its LC3-interacting motif by TANK-binding kinase 1 (TBK1) (Wild et al., 2011b). However, little is known about OPTN role in the autophagy-mediated degradation of misfolded protein inclusions. In this work, we thus proposed to assess the relative contributions of the most associated variants identified in the *PDB6* (*10p13*) locus associated to the functionality of *OPTN* located in this locus and their involvement in the PDB physiopathology.

4.1.3 Materials and Methods

4.1.3.1 Methylation analysis by bisulfite conversion and Sanger sequencing of rs1561570

To analyze if rs1561570 was changing the methylation status we used the MethPrimer (http://www.urogene.org/methprimer/index1.html), an *in silico* tool used to predict CpG island by examining the CG dinucleotides and to design specific primers to analyze those CpG islands. Sodium bisulfite treatment was performed on 1,5 μg of genomic DNA sample using the EpiMark® Bisulfite Conversion Kit (New England Biolabs Inc) following the manufacturers' standard protocol. DNA samples from 130 PDB patients and 225 healthy individuals (with *CT* and *CC* genotypes of rs1561570) were bisulfite converted, with fully-methylated (methylation level of 100%) and unmethylated DNA (0% methylation) samples as controls. The samples with incomplete bisulfite conversion were discarded from this analysis. Bisulfite-PCR amplification was conducted using Taq DNA polymerase and cycling conditions of

45 cycles with an annealing temperature of 56°C. The primers used (Rs1561570 Bisulfite F and Rs1561570 Bisulfite R; Table S3.1) amplify a region of 267 bp around the rs1561570 genomic position. Methylated cytosine was detected by Sanger sequencing of the PCR products.

4.1.3.2 Cell culture conditions

The U937 and HEK293 cell lines were grown in Dulbecco's modified eagle medium (DMEM), supplemented with 10% fetal bovine serum (FBS), 2mM L-glutamine and 1% penicillin/streptomycin (P/S). The T47D cell line was grown in RPMI-1640 medium supplemented with 10% FBS, and 1% P/S. The medium, FBS, antibiotics and glutamine were obtained from Invitrogen. Both cell lines were incubated at 37^{o}C in a humidified atmosphere containing 5% CO_2.

4.1.3.3 Demethylating treatment and immunofluorescence

The localization of NF-κB in T47D and U937 cell lines treated for 72h with 5-Azacitidine (5-Aza) (Sigma) and in osteoclasts derived from controls and patients PBMCs was assessed by immunofluorescence. The cells were washed with phosphate buffered saline (PBS), fixed with 4% paraformaldehyde, and permeabilized in PBS containing 0.1% Triton 100 for 4 min. Non-specific binding was blocked by incubating the cells with PBS supplemented with 0.5% FBS and 0.5% bovine serum albumin (BSA) for 30 min at RT. Cells were then incubated ON at 4^{o}C with a rabbit polyclonal antibody directed against human NF-κB (#8242, Cell Signaling). A rabbit polyclonal antibody against LC3B (#2775, Cell Signaling) was also used in PDB patients' osteoclasts to access effects in autophagy. After rinsing with PBS, the cells were incubated with the secondary antibody Alexa 488 (green) or Alexa 594 (red) goat anti-rabbit (Cell signaling), in 0.5% BSA-PBS for 45 min. The microscope slides were rinsed and mounted with the mounting medium for fluorescence.

4.1.3.4 Preparation of human *in vitro*-differentiated mature osteoclasts

Human mature osteoclasts were differentiated *in vitro* using mononuclear cells from blood of healthy controls and patients with PDB. PBMCs were obtained by density gradient centrifugation using Ficoll-Hypaque. The cells were resuspended ($3x10^6$ cells/mL) in OPTI-MEM (Invitrogen) containing 10% FBS. The cell suspension was added to 6-well plates ($9x10^6$ cells/well) and to Lab-Tek 8 well-slides (Sigma-Aldrich) ($3\text{-}6x10^5$ cells/well). After 24h, the cells were washed thoroughly and lymphocyte-free adherent cells were incubated for three weeks with M-CSF (25 ng/mL, Life Technologies) and RANKL (30 ng/mL, Fitzgerald).

4.1.3.5 Osteoclast morphology assessment and TRAP assay

Mature osteoclast formation was evaluated by quantification of TRAP-positive multinuclear cells using an Acid Phosphatase kit (Sigma-Aldrich), according to the manufacture's recommendations. Briefly, after fixation, the cells were stained for acid phosphatase in the presence of 0.05 M sodium tartrate. The substrate used was napthol AS-BI phosphate. TRAP-positive and TRAP-negative cells were counted in at least 5 random fields using light microscopy.

4.1.3.6 *In vitro* assessment of bone resorption

The bone resorption activity of osteoclasts was tested *in vitro*, using a previously validated model (Itzstein et al., 2000). Briefly, monocytes were seeded onto cortical bone slices and differentiated as described previously. At the end of the culture, cells were removed and bone slices were stained with toluidine blue to score the percentage of bone resorbed area. Bone resorption was evaluated using ImageJ by doing the ratio of the total bone resorbed area versus the total bone slice area.

4.1.3.7 Quantitative real-time PCR

In order to test if the most strongly associated SNPs had an impact on *OPTN* gene expression, we performed qPCR as previously described (Michou et al., 2012). To

analyze rs2234968, we used RNA from PBMCs of six patients and three healthy donors and to analyze rs1561570, we used RNA from PBMCs of 22 patients and 12 healthy donors. Normalization was performed using the reference gene peptidylprolyl isomerase B (*PPIB*), since it was defined as a suitable reference gene for mRNA quantification in peripheral whole blood (Pachot et al., 2004). To test the gene expression of *OPTN* and NF-κB target genes (*NF-κB, IL-6, ELK1,* and *NFATc1*) normalization was performed using the reference gene glyceraldehyde 3-phosphate dehydrogenase (*GAPDH*), since it was defined as a suitable internal control to evaluate gene expression after treatment with 5-Azacitidine (Kobune et al., 2012; Wang et al., 2013). A quantity of cDNA corresponding to 10 ng of total RNA was used in these analyses. The primers used are displayed in Table S3.1.

4.1.3.8 Western blot analyses

Osteoclasts derived from PBMCs from patients and controls were washed once in PBS and lysed using Trizol (Sigma-Aldrich). The protein concentrations were determined using the Bradford reagent (Bio-Rad). Proteins were separated by 8% SDS-polyacrylamide gels and transferred onto PVDF membranes (Life technologies). After transfer, the membranes were blocked in 0.1% TBS/Tween 20 containing 5% nonfat dried milk at room temperature for 1 h. The membranes were then incubated ON at 4°C with anti-OPTN (1:500, #100000 Cayman), anti-LC3B (1:1000, #2775 Cell Signaling) or anti-αTubulin (1:2000, #2144 Cell Signaling) in 0.1% TBS/Tween 20, followed by extensive washing using 0,1% TBS/Tween 20 and an incubation with HRP-conjugated secondary antibody (1:2000, Cell signaling) in 0.1% TBS/Tween 20 during 1h, at room temperature. After extensive washing with 0.1% TBS/Tween 20, specific proteins were detected using a chemiluminescence kit (GE Healthcare). The densitometric analysis was performed using ImageJ software.

4.1.3.9 Statistical analyses

Statistical analyses were performed with GraphPad Prism 5 (GraphPad, La Jolla, CA). Comparisons between two groups were made using a two-tailed Student's t-test. For comparisons between multiple groups, one-way ANOVA followed by Tukey's post-hoc test was used. Differences were considered statistically significant when $p<0.05$.

4.1.4 Results

4.1.4.1 Effect of rs2234968 in *OPTN* splicing and protein sequence

The possible functional effects of rs2234968 in 1) *OPTN* mRNA secondary structure according to RNA fold program revealed an absence of a predicted effect in mRNA structure (Figure 4.1A); 2) *OPTN* expression levels by qPCR using rs2234968qPCR F/R primers (Table S3.1) and patients and healthy donors samples with all genotypes, showed an absence of correlation between genotype and *OPTN* expression levels (Figure 4.1B); and 3) *OPTN* splicing pattern, according to Human Splicing Finder tool, indicated a prediction of a new branch point and consequently a possible exon 5 skipping, which would change the protein sequence (Figure 4.1C). However, by doing a PCR using primers flanking exon 5 (rs2234968 qPCR F/qPCR R, Table S3.1) and cDNA from patients as template, we showed that rs2234968 was not responsible for exon 5 skipping (Figure 4.1D).

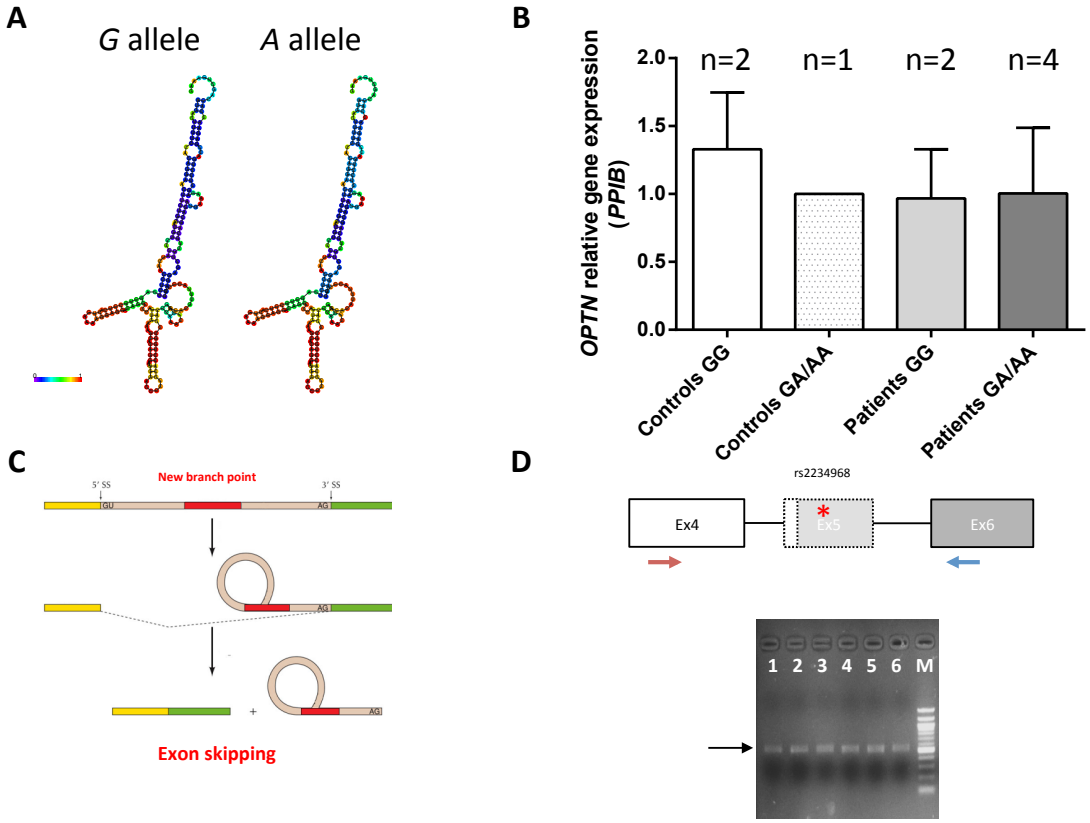


Figure 4.1 - Effect of rs2234968 in mRNA secondary structure, OPTN expression and OPTN splicing. A) *In silico* analysis of the mRNA secondary structure in the region around rs2234968, showing that the *A* allele is not responsible for a change in OPTN mRNA structure. B) Analysis of

OPTN gene expression in PBMCs-derived osteoclasts from non-mutated controls (*GG*) (n=2), mutated controls (n=1), non-mutated patients (*GG*) (n=2), three patients heterozygous (*GA*) and one patients homozygous (n=4) for rs2234968, showing no statistical difference in OPTN gene expression. C) *In silico* prediction of the effect of rs3422968 in OPTN splicing, indicating that rs3422968 could be responsible for exon 5 skipping. D) Localization of the primers (red and blue arrows represent rs2234968 qPCRF1 and rs2234968 qPCR R1, respectively) used to amplify by PCR the region flanking exon 5 and electrophoresis results showing that the A allele is not responsible for a change of OPTN splicing. 1 and 2 represent patients with GG genotype, 3 and 4 patients with GA genotype and 5 and 6 patients with AA genotype. M represents the 1kb marker.

We then searched for SNPs in *linkage disequilibrium* (LD) with rs2234968, which may have a functional effect and explain the association of PDB with rs2234968. We found four SNPs [rs76647957 (r^2 = 0.831; DPrime = 0.912), rs12415802 (r^2 = 0.831; DPrime = 1), rs10906303 (r^2 = 0.958; DPrime = 1) and rs79529484 (r^2 = 0.801; DPrime = 0.954)] in LD with rs2234968 that are located in *OPTN* intron 2, 3, 5 and 6, respectively (Figure 4.2A). Using Human Splicing Finder tool, we determined that rs10906303 and rs79529484 were the most likely to have a functional effect on splicing (Figure 4.1B). This effect was tested by PCR using cDNA samples from patients with all genotypes and specific primers (Figure 4.2C). We found that both rs10906303 and rs79529484 may alter the splicing sites, causing the retention of the intron 5 and 6 respectively (Figure 4.2C), which may change the amino acid sequence and lead to a predicted premature stop codon (Figure 4.2D). By using an antibody against the complete OPTN protein in a western blot analysis, we were also able to see a decrease in OPTN protein levels in the presence of rs2234968 *A* allele (Figure 4.2E), but no band of truncated protein size, suggesting that either the mRNA is degraded by nonsense-mediated mRNA decay or the truncated protein is degraded, since it loses all of its functional domains.

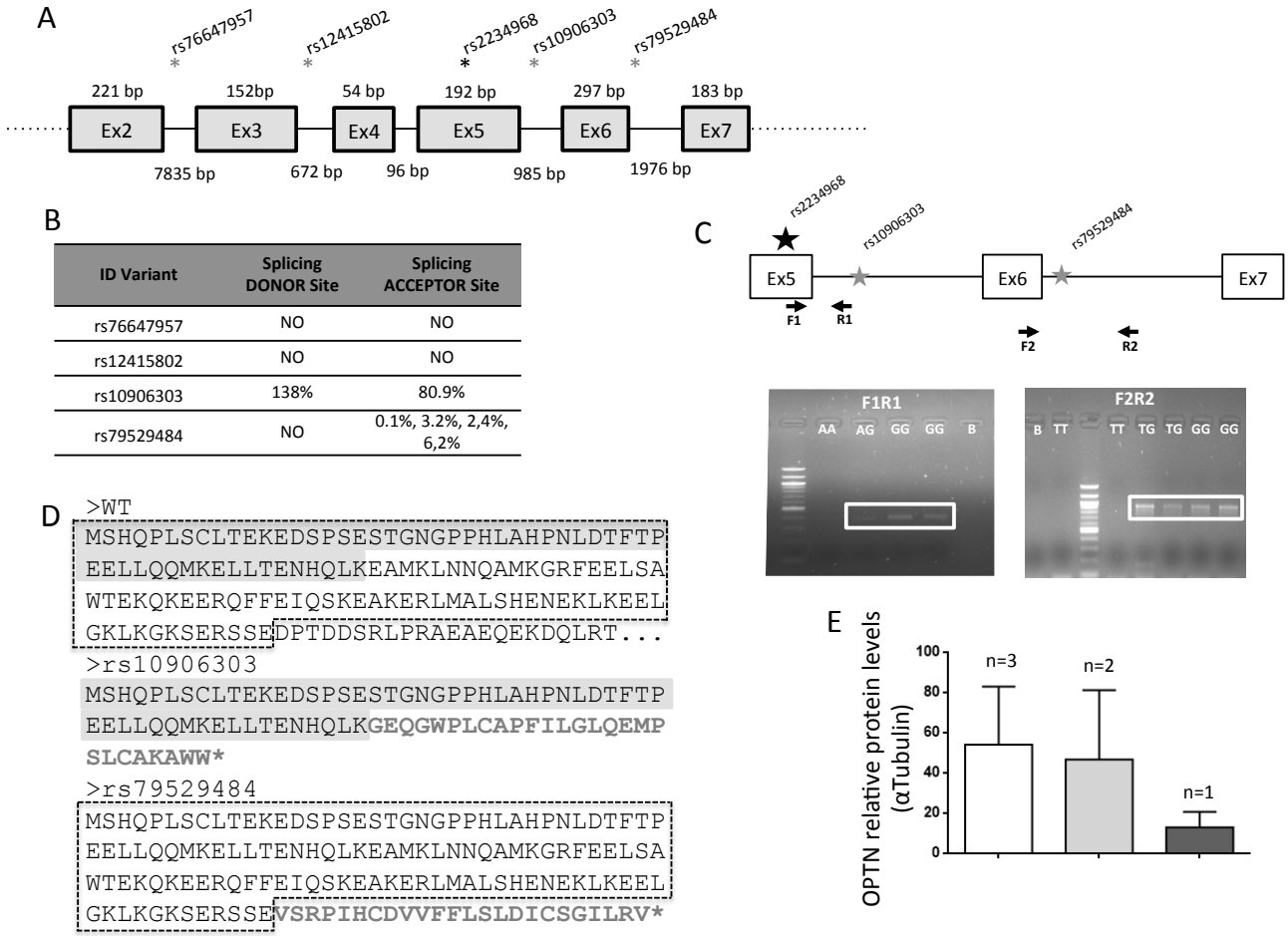


Figure 4.2 - SNPs in Linkage Disequilibrium with rs2234968. A) Localization of the SNPs in linkage disequilibrium with rs2234968; B) *In silico* prediction of the effects of the SNPs in linkage disequilibrium with rs2234968, in *OPTN* splicing; C) Localization of the primers (rs10906303 F1/R1, rs79529484 F2/R2, Table S3.1) used to amplify by PCR the different transcripts and electrophoresis results showing that with the presence of one mutated allele we have a change in *OPTN splicing* pattern. *AA*, *AG*, *GG*, *TT*, and *TG* represent different genotypes from patients samples and B represents the PCR negative control; D) OPTN protein sequence showing a premature stop codon - predicted by the position of the stop codon from the cDNA sequence from patients (represented by an *) - due to the effect of rs10906303 (grey box) and rs79529484 (black dashed box) in splicing. The bold and grey amino acids represent the new protein derived from the skipping of the intron. E) Analysis of OPTN protein expression in PBMCs-derived osteoclasts from non-mutated controls (*GG*) (n=3), non-mutated patients (*GG*) (n=2) and one patient heterozygous (*GA*) for rs2234968.

4.1.4.2 Rs1561570 increases osteoclast differentiation and activity

To better assess the effect of rs1561570 on osteoclasts formation, we performed a TRAP assay. The presence of at least one *T* allele increases significantly osteoclast formation, attested by the number of multinucleated cells yielded from PBCMs cultures, when compared to healthy controls (80% versus 10%, *p*-value = 0.001)

(Figure 4.3A, 4.3C). The osteoclasts generated contained more nuclei than the osteoclasts from healthy controls (*p*-value < 0.001) (Figure 4.3B, 4.3C).

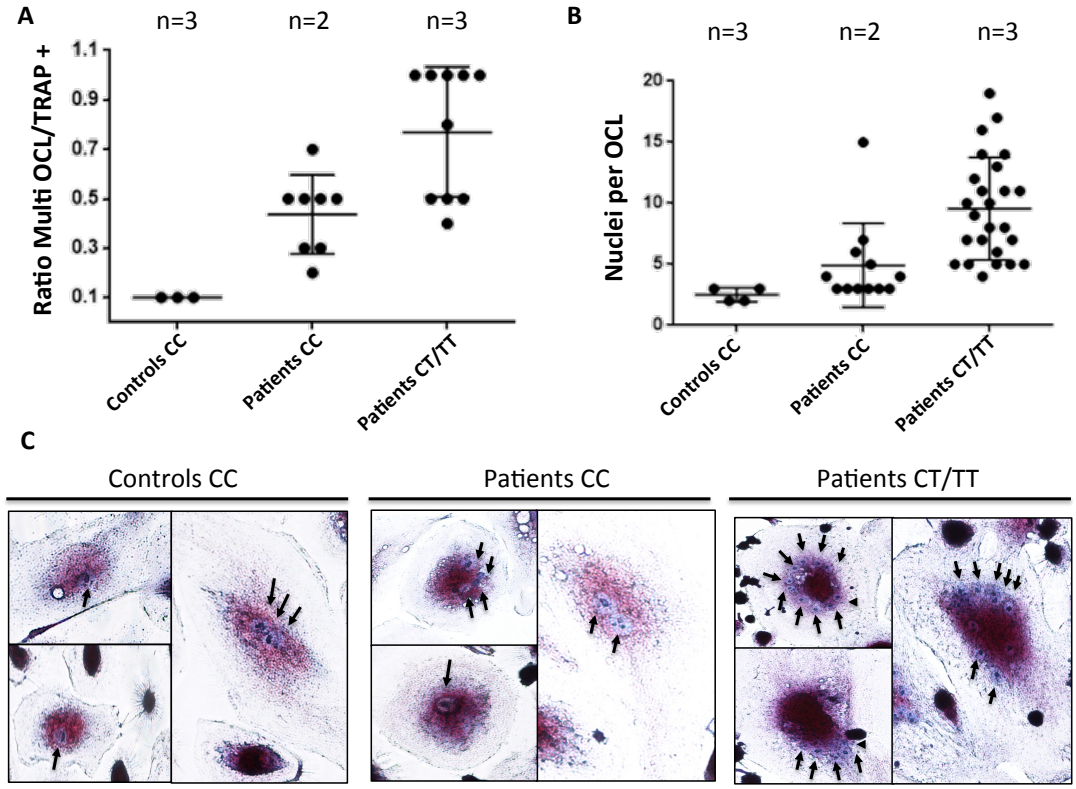


Figure 4.3 - Effect of rs1561570 in osteoclast differentiation. A) The osteoclast differentiation rate (displayed as a ratio between multinucleated osteoclasts and TRAP positive cells) was higher in patients with PDB carrying at least one *T* allele (*p*-value = 0.001). B) Osteoclasts from patients carrying at least one *T* allele contained significantly more nuclei than osteoclasts from healthy controls (*p*-value < 0.001). C) The distribution of the number of nuclei per osteoclast was observed in samples from healthy controls and PDB patients that were stained for TRAP and counterstained with haematoxylin.

Bone resorption was also significantly increased in osteoclast cultures generated from patients with at least one *T* allele, resulting in 3-fold increase in the total resorption area when compared to cultures prepared from healthy controls (*p*-value < 0.001) (Figure 4.4A, 4.4B).

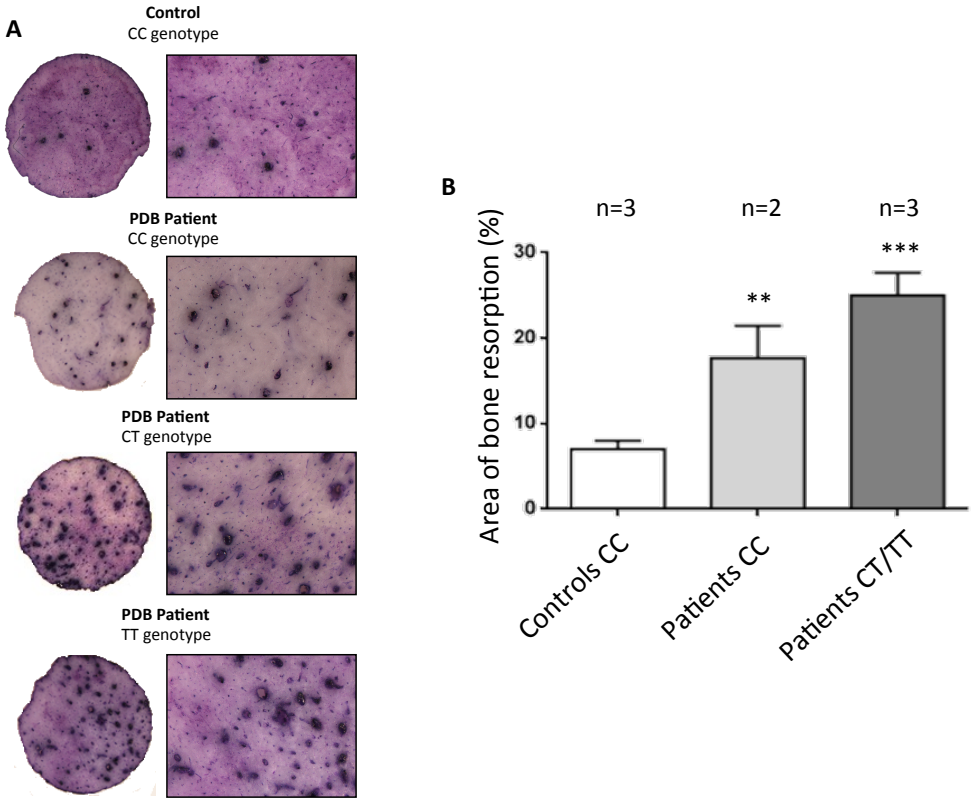


Figure 4.4 - Effect of rs1561570 in osteoclast differentiation. A) Representative images of *in vitro* bone resorption assays (left) with B) quantification of the results (right) by using ImageJ.

4.1.4.3 Rs1561570 does not affect macroautophagy

OPTN has been suggested to be an autophagy adaptor and since impairment in autophagy was reported in PDB, we assessed the autophagy process in the presence of rs1561570 *T* allele. By immunofluorescence and western blot analysis we found that LC3B expression was impaired, however that effect did not seem to be dependent of the presence of a *T* allele (Figure 4.5A, 4.5B).

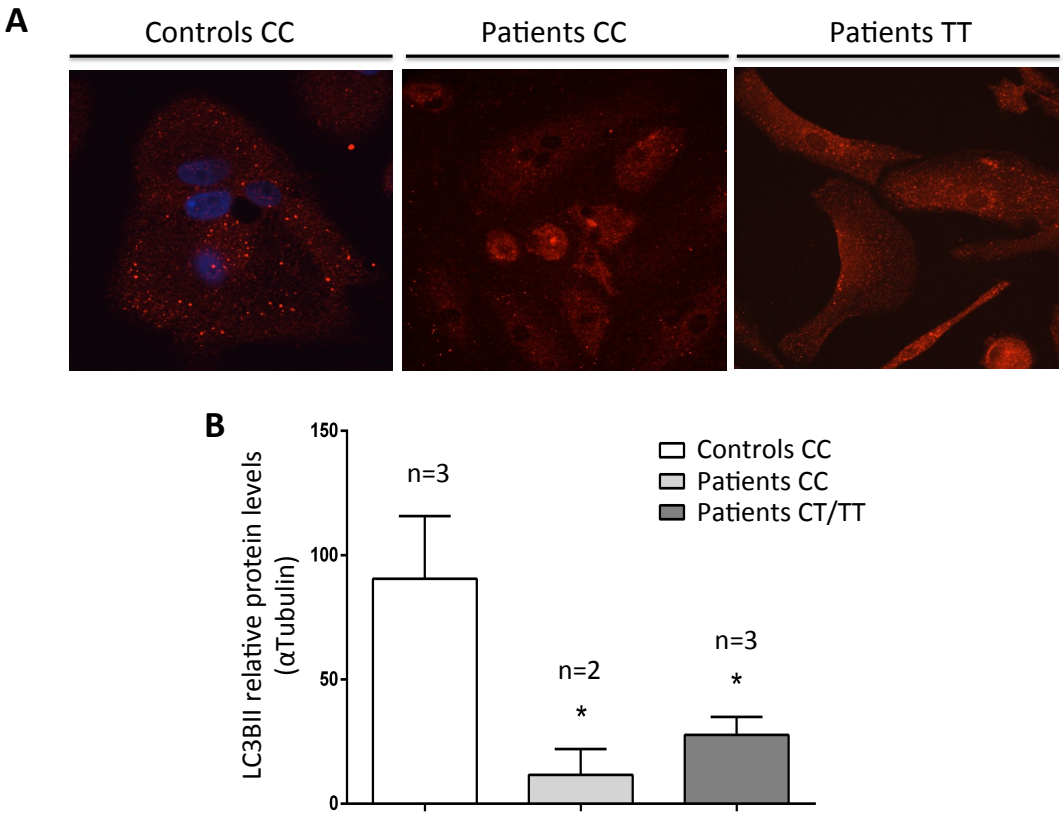


Figure 4.5 - rs1561570 effect in autophagy. A) LC3BII foci (red) in PBMC-derived osteoclasts from controls and PDB patients were analyzed by immunofluorescence. B) The levels of LC3BII protein expression were measured by western blot analysis and related to levels of α-Tubulin (ANOVA, * represents a *p*-value < 0.05).

4.1.4.4 Rs1561570 causes a change in methylation and increases *OPTN* expression

Since our previous study has shown a strong association between rs1561570 (located in OPTN intron 7) and PDB (Michou et al., 2012), we determined the functional effect of rs1561570 on PDB by employing the splice site prediction algorithm, and showed that the allele harboring a *C* was likely to create a potential new acceptor site for splicing and disrupt a potential branch point (Consensus value for mutant sequence = 69.7%). To validate this finding, we performed RT-PCR and qPCR using cDNA samples from patients and controls with the three genotypes. Retention of intronic sequences was not observed in any genotype (Figure 4.6).

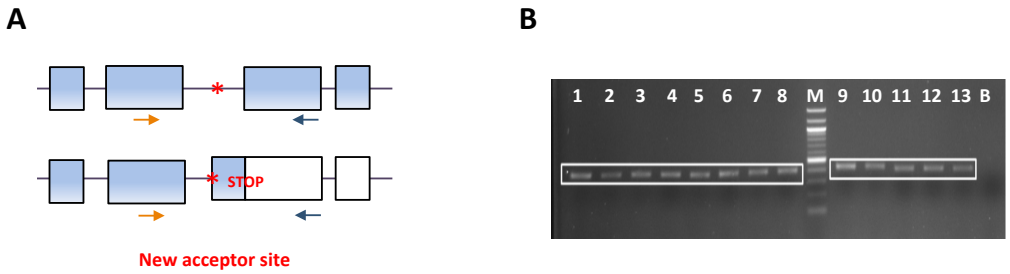


Figure 4.6 - Schematic representation of possible rs1561570 effect in OPTN splicing. A) Bioinformatics prediction from Human Splicing Finder showing that rs1561570 was predicted to cause a new acceptor site. Boxes represent *OPTN* exons. Arrows represent the primers used (red arrow – rs1561570 qPCR F; blue arrow – rs1561570 qPCR R, Table S3.1). The red asterisk represents the SNP rs1561570. B) Electrophoresis results showing that the presence of the mutated allele do not change *OPTN splicing* pattern. 1-3 - PDB patients with *CC* genotype, 4-9 - PDB patients with *CT,* 10 -13 - PDB patients with *TT* genotype, B represents the PCR negative control, and M represents the 1 kb marker.

We then studied the predicted impact of rs1561570 in the *OPTN* methylation status using the Methprimer tool. We found that allele *C* was predicted to cause a CpG dinucleotide, creating a methylation site, which was not present in the presence of the *T* allele (Figure 4.7A). After bisulfite treatment, we confirmed that all samples with at least one *C* allele in their rs1561570 genotype were in fact methylated (results not shown). To correlate the methylation site created by the *C* allele with the *OPTN* transcriptional levels, we performed qPCR and western blot analysis using lymphocytes from patients with the three genotypes. Samples carrying a *T* allele (*CT* or *TT*) showed levels of *OPTN* gene and protein expression higher than samples with *CC* genotype (Figure 4.7B, 4.7C). This result suggested that the methylated *C* allele is associated with a decrease in *OPTN* gene expression levels. The protein expression levels were measured by western blot analysis in PBMC-derived osteoclasts and were in accordance with the results obtained by qPCR and western blot analysis in lymphocytes (Figure 4.7D).

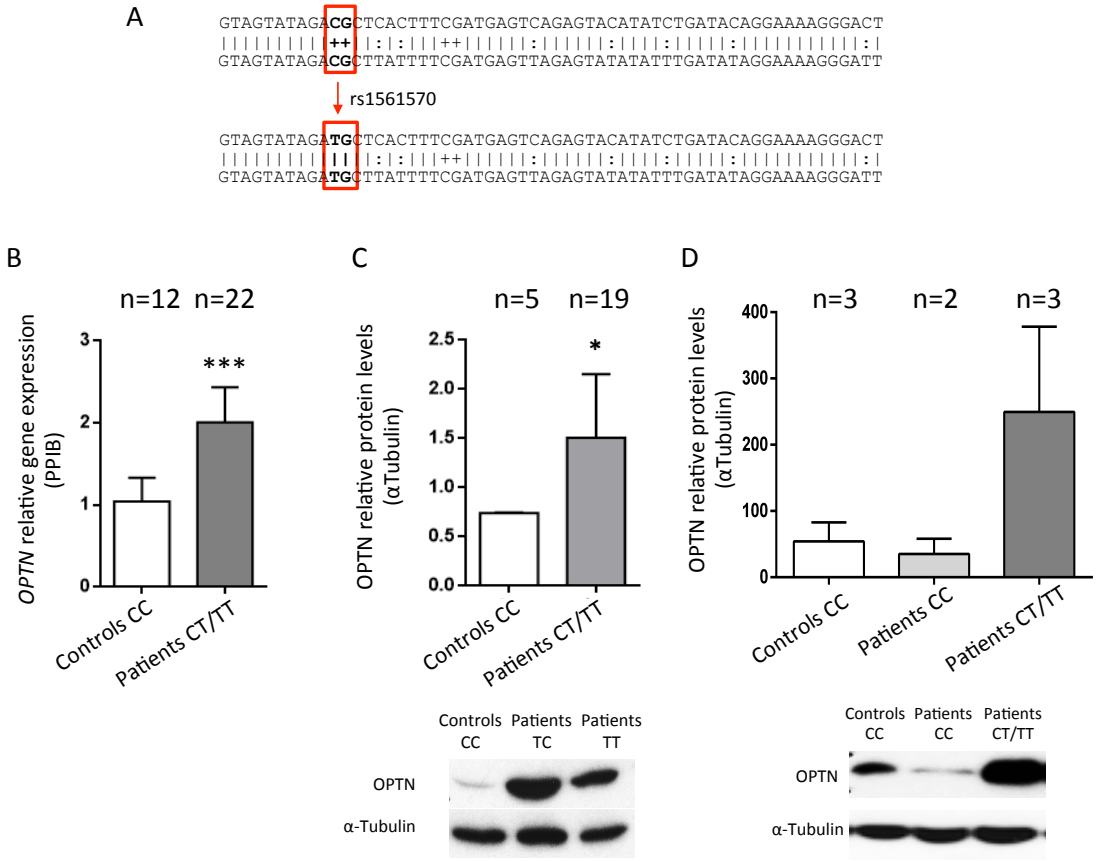


Figure 4.7 - Rs1561570 effect in methylation and *OPTN* expression. A) *In silico* prediction using Methprimer tool showing a methylation site only in the presence of the allele *C* (not the allele *T*). B) Analysis of *OPTN* gene expression in lymphocytes from several controls (n=12) and patients (n=22) with the three genotypes. The levels of *OPTN* gene expression were measured by qPCR normalized to *PPIB* gene expression levels. Values are the mean of at least three independent replicates. C) Analysis of OPTN protein expression in lymphocytes from non-mutated controls (n=5), and patients carrying at least one *T* allele (n=19). D) Analysis of OPTN protein expression in osteoclasts from non-mutated controls (n=3), non-mutated patients (n=2) and patients carrying at least one *T* allele (n=3). The levels of OPTN protein expression were measured by western blot analysis and related to levels of α-Tubulin (t-test and ANOVA, * represents a *p*-value < 0.05, ** represents a *p*-value < 0.005, *** represents a *p*-value < 0.001).

4.1.4.5 Rs1561570 increases NF-κB translocation into the nucleus

To determine the role of rs1561570 on PDB phenotype, we analyzed the NF-κB cellular localization by immunofluorescence following treatment with the demethylating agent, 5-Azacitidine. By using two different cell lines which carry the rs1561570 *C* allele – T47D (*CT* genotype) and U937 (*CC* genotype) - we were able to confirm that before the demethylating treatment NF-κB was localized in the cytoplasm and in the nucleus. However, after *OPTN* demethylation, NF-κB localization is almost exclusive in the nucleus, confirming that a decrease in *OPTN*

methylation (as seen in PDB patients with rs1561570 *T* allele) increases the translocation of NF-κB into the nucleus (Figure 4.8A and 4.8B).

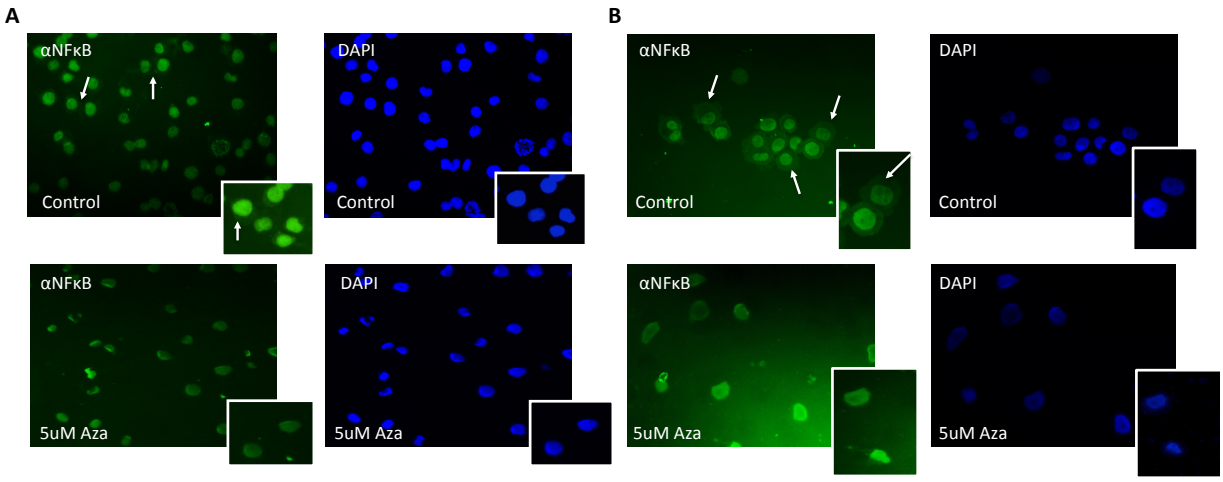


Figure 4.8 - Effect of *OPTN* demethylation in NF-κB localization. NF-κB localization in (A) U937 cells (CC genotype), and in (B) T47D cells (CT genotype), following 5-Azacitidine (Aza) treatment. After *OPTN* demethylation, NF-κB was translocated exclusively to the nucleus.

To confirm the effect of the demethylating agent we measure *OPTN* expression and to confirm that the presence of NF-κB in the nucleus was affecting the expression of its target genes, we measured the expression of *NF-κB, IL-6, ELK1,* and *NFATc1*. Our results showed that there was an increase in *OPTN* expression, which was expected after the demethylation treatment, and in the expression of the NF-κB target genes analyzed. NF-κB levels were maintained as expected (Figure 4.9).

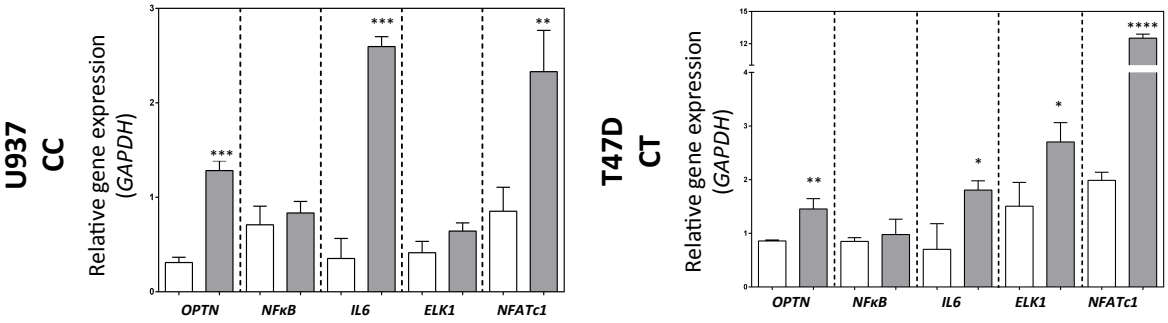


Figure 4.9 - Effect of *OPTN* demethylation in expression of NF-κB target genes. Analysis of NF-κB target genes expression. The levels of *OPTN, NF-κB* and NF-κB target genes (*IL-6, ELK1, NFATc1*) expression were measured by qPCR related to levels of *GAPDH* gene. Values are the mean of at least three independent replicates. (t-test, * represents a p-value < 0.05, ** represents a p-value < 0.005, *** represents a p-value < 0.001, **** represents a p-value < 0.0001).

Since rs1561570 was found to have an *in vitro* effect on the NF-κB translocation into the nucleus, we further analyzed the NF-κB localization in PBMC-derived osteoclasts

from PDB patients carrying different genotypes of rs1561570 and from healthy controls, as well as *NFATc1* expression (since it is the most relevant NF-κB target gene for osteoclast differentiation). Our results showed that, in accordance with our *in vitro* assays, the patients with at least one *T* allele had an increase in NF-κB translocation to the nucleus (showed also as a ratio between co-localization of NF-κB and DAPI) (Figure 4.10A) and an increase in *NFATc1* expression (Figure 4.10B) – another strong evidence that an increase of *OPTN* expression due to the effect of the *T* allele in the *OPTN* methylation status leads to an increase of NF-κB translocation into the nucleus and an increase of its target genes expression.

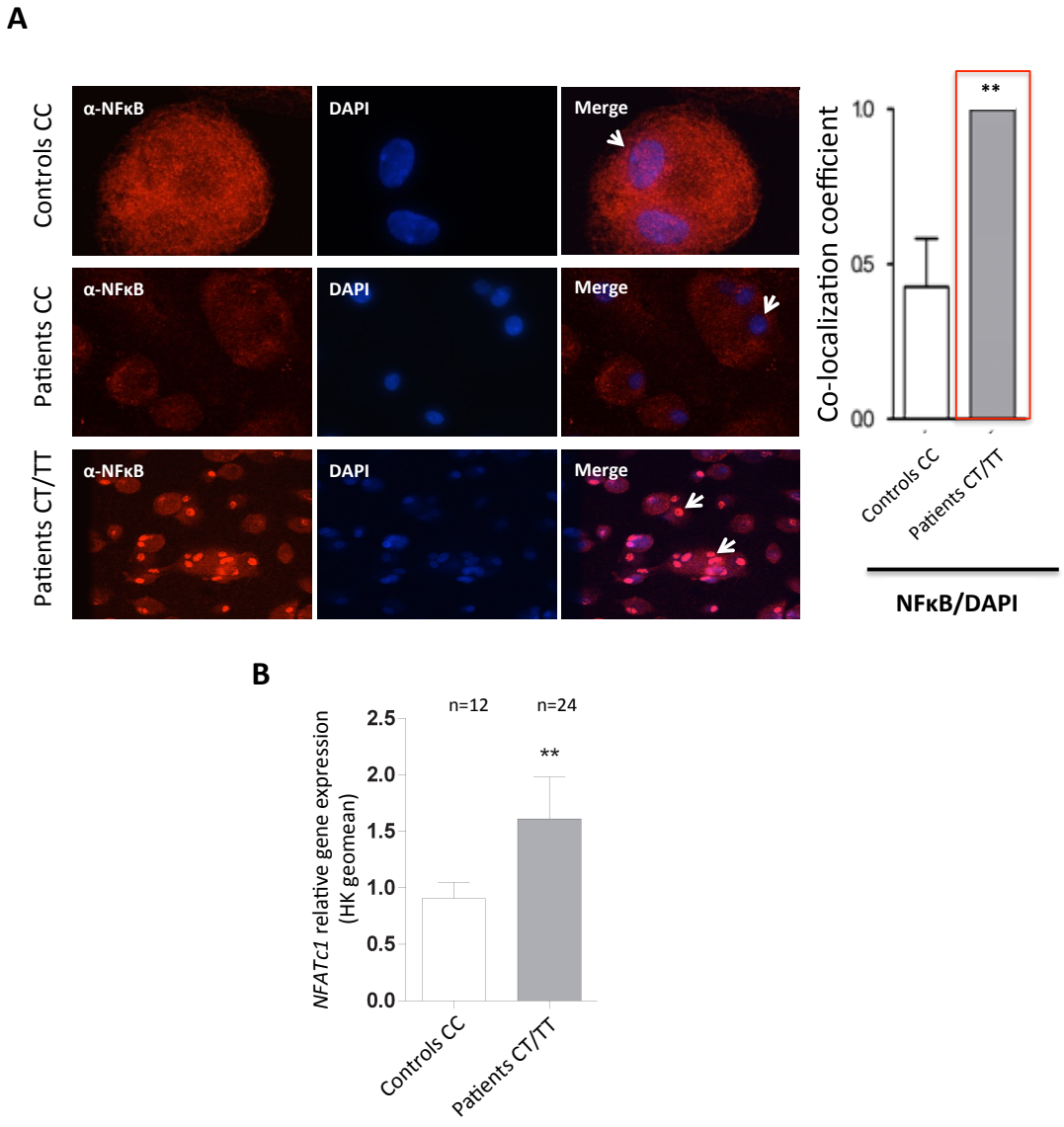


Figure 4.10 - Effect of rs1561570 in NF-κB localization in osteoclasts and in the expression of *NFATc1*. A) NF-κB localization in osteoclasts derived from healthy controls and patients PBCMs showing an increase of the translocation of NF-κB into the nucleus, in the presence of at least one *T* allele. This was also confirmed by co-localization ratio between NF-κB and DAPI – the nuclear

staining. B) Analysis of *NFATc1* gene expression in several controls (n=12) and patients (n=24) with all genotypes. The levels of *NFATc1* expression were measured by qPCR related to levels of the mean of the housekeeping genes (*ATPO5, HPRT1, PPIB, PPIA,* and *18S*). Values are the mean of at least three independent replicates. (t-test, ** represents a *p*-value < 0.005).

4.1.4.6 Rs1561570 may have an effect on PDB severity

By comparing the PDB clinical phenotype of patients carrier of the *CC* genotype of *OPTN* rs1561570 with patients carrying one or two *T* alleles, we can observe a tendency although not significant for an: i) increase of total sALP levels, ii) increase of mean number of affected bones, and iii) increase of disease extension measured by Renier's index – related with the number of *T* alleles in the genotype (Table 4.1). These results suggest that rs1561570 can be a susceptibility variant that could have a dose effect depending on the number of *T* alleles present in the genotype.

4.1.5 Discussion

In the present study, we investigated the functional role of the two most associated variants with PDB within the PDB6 locus. We focused our analysis in the rs2234968, discovered in our previous work and in rs1561570 (C>T), previously described has having a strong association with PDB in the literature (Michou et al., 2012). Regarding rs2234968 we found it was associated with PDB in our population but it has no functional effect on PDB pathogenesis. Instead we found two SNPs (rs10906303 and rs79529484) in LD with rs2234968 that could have a functional effect. Then, we analyzed the effect of rs10906303 and rs79529484 in the splicing pattern of *OPTN* gene, and showed that these SNPs cause the retention of intron 5 and intron 6 respectively, producing new *OPTN* transcripts, both leading to premature stop codons. We hypothesize that these new transcripts encode smaller proteins, which may contribute to the PDB phenotype since those truncated proteins will lose the main functional domains (NEMO and LIR) and therefore are likely to be non-functional. Similarly to the deleterious effect of *SQSMT1* mutations that impair its main functional domain - UBA (Laurin et al., 2002), the predicted loss of *OPTN* protein domains, namely the NEMO, UBD and LIR domains, should impair its function.

Table 4.1. Comparisons of main clinical characteristics between French-Canadian patients with Paget's disease of bone carrier of the rs1561570 *TT* or *TC* genotype to *CC* genotype carrier, considered as the reference.

	Categories of patients			**Comparison of *TC* to *CC* genotype**	**Comparison of *TT* to *CC* genotype**
	CC genotype (n=35)	*TC* genotype (n=93)	*TT* genotype (n=99)	*p*-value	*p*-value
Male sex, n (%)	22 (63)	58 (59)	59 (58)	0.66	0.60
Age at diagnosis (years), mean ± SD	61.9±8.7	63.1±12.0	63.2±11.5	0.60	0.55
Total sALP levels [a]	2.6±1.6	3.1±3.3	3.4±3.5	0.38	0.18
Number of affected bones, mean ± SD	2.3±1.8	2.4±2.4	2.6±2.6	0.82	0.53
Renier's index (%), mean ± SD	9.2±7.4	9.5±8.1	10.5±8.7	0.86	0.44

[a] sALP = serum phosphatase alkaline; for total sALP levels, values are the number of time the Mid point of Normal range ± SD

Accordingly, an *OPTN* transcript with a retained intron 5 was already described in patients with other disease - juvenile open-angle glaucoma (Willoughby et al., 2004), thus indicating that the presence of this abnormal transcript can alter OPTN function or the normal regulation of its protein expression (Lewis et al., 2003). However, further studies will be required to clarify the effect of these transcripts in PDB patients. By quantifying their expression and comparing it with the expression of the normal transcripts, and analyzing if OPTN function is impaired when only the mutant and "short" transcript is expressed should be highly relevant to further understand its effect in associated pathologies such as PDB. In addition, by analyzing the resulting autophagy process or the expression of osteoclast related genes may also contribute to clarify the effect of these variants in PDB pathophysiology.

Regarding rs1561570 (C>T), the studies previously reported showed that the *C* allele was more frequent in controls while the *T* allele was more frequent in PDB patients. Using *in silico* tools, we found that the rs1561570 *T* allele was predicted to cause a loss of a methylation site. With bisulfite treatment, we confirmed that the *C* allele in that position was always methylated. Furthermore, both qPCR and western blot analysis correlated the presence of the *C* allele with a decrease of *OPTN* expression levels. Our immunofluorescence analysis allowed us to confirm that an increase of *OPTN* levels (due to the demethylation treatment in cells and due to the presence of the *T* allele in PDB patients) is related to an increase in NF-κB translocation into the nucleus and to an increased expression of NF-κB target genes related with osteoclastogenesis, such as *NFATc1* – a master transcription factor of osteoclast differentiation (Lorenzo et al., 2007). Journo and its co-workers had already described the mechanism behind the translocation of NF-κB into the nucleus due to OPTN. (Journo et al., 2009) Altogether, our results suggest that, in the presence of two *C* alleles (more frequent in healthy controls), the presence of a methylated site at rs1561570 maintains *OPTN* expression at basal levels, which in turn leads to basal levels of NF-κB translocation to the nucleus, giving rise to normal levels of osteoclastogenesis-related gene expression. However, when patients have at least one *T* allele at rs1561570, that site is no longer methylated, and that correlates with an increase of *OPTN* expression and increase of translocation of NF-κB into the nucleus. This gives rise to an increase in expression of osteoclast related genes, such as the osteoclast major transcription factor, *NFATc1*, probably contributing to the increase in

osteoclast differentiation rate, number of nuclei per cell, and resorbing activity observed in patients with at least one *T* allele (Figure 4.11).

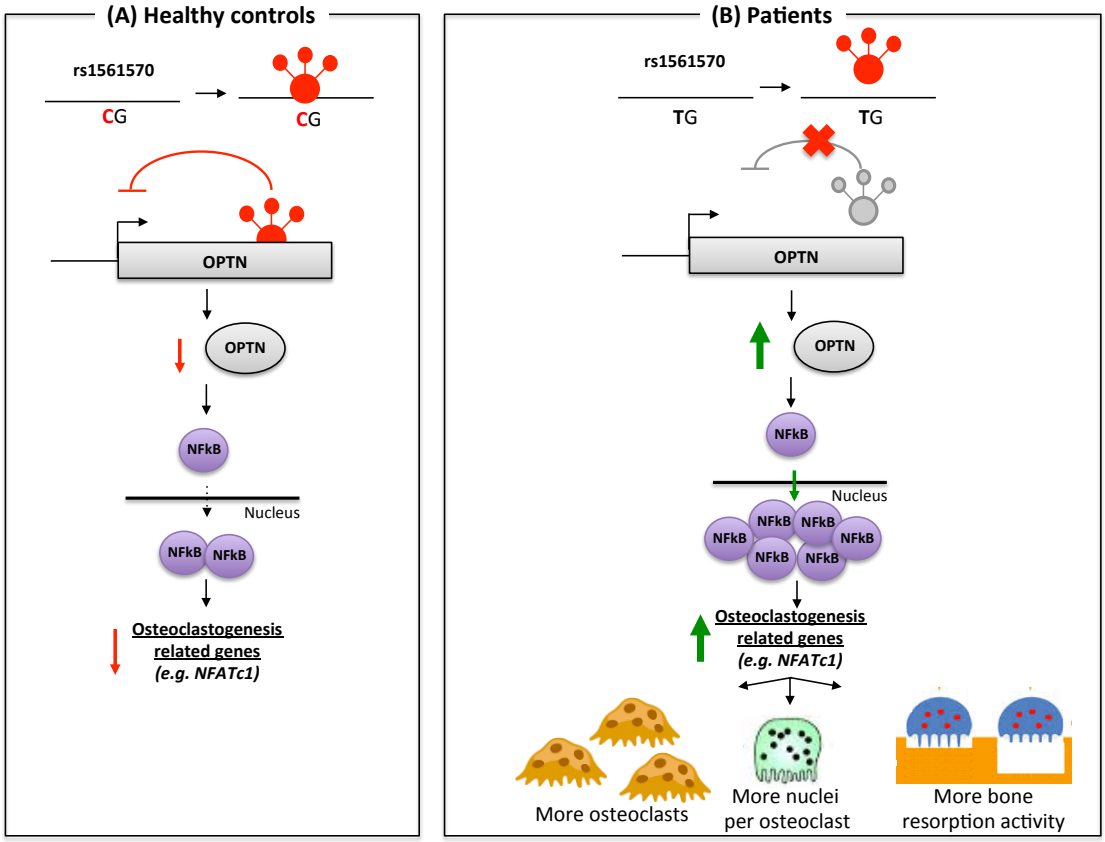


Figure 4.11 - Functional effect of rs1561570 in PDB – a hypothesis. A) In healthy control samples, the most frequent allele is the *C* and the functional data in this work showed that this nucleotide was methylated. Since methylation was associated with transcription repression, this methylated site should maintain *OPTN* expression in basal levels, leading to lower levels of NF-κB in the nucleus thus maintaining basal expression levels of genes related to osteoclastogenesis. B) In PDB patients, the most frequent allele was the *T* that was no longer methylated. The loss of that methylation site will lead to higher levels of *OPTN* transcription, higher levels of NF-κB translocation into the nucleus, and higher levels of expression of genes related to osteoclastogenesis, thus increasing osteoclast number and resulting resorbing activity, therefore contributing to PDB pathogenesis.

A correlation between the hypermethylation status in different genes resulting in a decrease in gene transcription levels affecting related diseases has been previously reported for other genes in the literature (Brenet et al., 2011; Hrašovec et al., 2013; Iwamoto et al., 2011). The change in the *OPTN* expression levels may also be due to the already reported binding of an enhancer - JUND, a member of the AP1 family that is present in osteoclasts (Beranger et al., 2007; David et al., 2001; Inoue et al., 1997) – in the genomic region containing the *T* allele, which is lost in the presence of the *C* allele, as confirmed by the ENCODE project (http://www.genome.gov/encode/). The ENCODE project shows regions of transcription factor binding derived from a large collection of ChIP-seq experiments together with DNA binding motifs identified

within these regions by the ENCODE Factor book repository. The presence of an enhancer might explain why a single change in the DNA sequence of these patients causes an alteration in *OPTN* expression levels. However, our results are not in direct accordance with some previous studies about *OPTN* in PDB. Some groups show that the depletion of OPTN increased the differentiation of osteoclasts (*p*-value = 0.002), showing that OPTN might be a negative regulator of osteoclast differentiation. However the same study also refers that *OPTN* expression significantly increases during days 2 to 5 of *in vitro* osteoclast differentiation (Obaid et al., 2012). Also, another group showed *OPTN* as having a negative effect on osteoclast differentiation but, unlike in our study, that analysis was performed in one patient carrier of a P392L mutation within the *SQSTM1* gene and the effect was mediated by the presence of measles virus nucleocapsid protein (MVNP) (Sun et al., 2014). Our study is unique since it analyses the single effect of rs1561570 in *OPTN* gene and protein expression using several *SQSTM1* negative pagetic patient samples, different cell types (lymphocytes and osteoclasts derived from PBMCs), *in vitro* and *ex vivo* techniques, and provides clear evidences that truly support our hypothesis at several levels.

In conclusion, there are several genes within the *PDB6* locus that could have a functional role contributing to PDB, the *OPTN* being the best candidate. Furthermore, we showed that rs1561570 is responsible for a change in the methylation status of *OPTN* and in its gene expression levels, which may contribute to PDB pathogenesis. Therefore, to this date, *OPTN* is still the most likely gene within this locus to have a functional effect that could contribute to PDB pathophysiology. We have also identified two SNPs in linkage disequilibrium with rs2234968 that are responsible for a change in splicing and possibly for creating a new truncated protein, probably non-functional.

CHAPTER 5

CAN ZEBRAFISH BE A VALID MODEL TO STUDY PAGET'S DISEASE OF BONE?

5. CAN ZEBRAFISH BE A VALID MODEL TO STUDY PAGET'S DISEASE OF BONE?

5.1 Comparative analysis of human and zebrafish SQSTM1: molecular and evolutionary perspectives

Iris A.L. Silva
Natércia Conceição
Laëtitia Michou *
M. Leonor Cancela *

This chapter was published as a research paper in the "Journal of Applied Ichthyology".

Authors' roles: Study design: MLC and LM. Study conduct: IALS. Data collection: IALS. Data analysis: IALS and NC. Data interpretation: All authors. Drafting manuscript: IALS. Revising manuscript content: NC, MLC and LM. Approving final version of manuscript: all authors.

5.1.1 Abstract

Paget's disease of bone (PDB) is the second most frequent metabolic bone disease after osteoporosis and genetic factors play an important role in PDB, but to date the only PDB causal gene identified is the *Sequestosome 1* (*SQSTM1*) gene. Because zebrafish has been validated as model for human genetic diseases our objective was to investigate if the gene structure and chromosomal environment of the *SQSTM1* was similar between zebrafish and human, thus providing a basis for developing a mutant fish capable of modeling PDB. Through a comparative *in silico* analysis, we confirmed that zebrafish *sqstm1* shares not only the same gene structure as its fish and mammalian orthologs, but they also present, within their promoter regions, similar putative binding sites for common transcriptional factors known to affect bone metabolism. The synteny of *SQSTM1* was also determined and results indicate that the cluster of surrounding genes was conserved throughout evolution. The protein comparison revealed a high degree of conservation, with particularly in the functional

domains of the protein. The most common mutation in PDB patients is p.Pro392Leu and the residue Pro392 was found to be 100% conserved in all species analyzed, including zebrafish, confirming its known functional relevance. In conclusion, this study demonstrated that *SQSTM1* is well conserved throughout evolution and therefore fish models, like zebrafish, could be an interesting tool to further investigate the biological role of SQSTM1 in PDB and in bone development.

5.1.2 Introduction

Paget's disease of bone (PDB) is the second most common metabolic bone disorder, after osteoporosis, affecting between 1% and 3% of individuals over the age of 55 years in Caucasian populations (Kanis, 1998). The cause of PDB is currently an area of intensive investigation, and both genetic and non-genetic factors have been implicated in the pathogenesis of this disease. Montagu group first raised the possibility that heredity might play a role in the pathogenesis of PDB (Montagu, 1948), and over recent years there has been an increasing interest towards understanding the role of genetic factors in the pathogenesis of this disease. Genetic factors are clearly an important component of the etiology of PDB, since ethnic differences in the incidence of PDB have been noted, and these persist with emigration to other regions. For example, PDB is common in persons of Anglo-Saxon origin, but the prevalence is low in the Far East and does not change when populations from this region move to areas of higher prevalence, such as the United Kingdom (Cooper et al., 2006). Another argument that favors the genetic inheritance of PDB is that 15–40% of affected patients have a first-degree relative with PDB and numerous studies have described extended families with PDB exhibiting an autosomal dominant mode of inheritance (Leach et al., 2001; Morales Piga et al., 1995).

In 2002, Laurin et al. reported a point mutation (p.Pro392Leu) in *SQSTM1*, in chromosome *5q35*, in two French-Canadian families with PDB and several unrelated patients (Laurin et al., 2002). *SQSTM1* encodes Sequestosome 1, also known as p62, which is an ubiquitin-binding protein involved in the IL-1, TNFα, and RANKL signaling pathways. *SQSTM1* is the gene most frequently linked to PDB and mutations on this gene have been detected in up to 40% of the families with PDB studied (Laurin et al., 2002). The crucial role of SQSTM1 during RANKL-induced

osteoclastogenesis was demonstrated in the *SQSTM1*$^{-/-}$ knockout mice, which displayed altered RANKL-induced osteoclast formation, NF-κB activation, and *NFATc1* expression (Durán et al., 2004). Hiruma et al. (2008) and Daroszewska et al. (2011) also published the presence of some PDB-like phenotypes in Pro394Leu mutant mice, however this model has some clinical and cellular differences from what is observed in humans and therefore does not reproduce the full phenotype of pagetic osteoclasts (Daroszewska et al., 2011; Hiruma et al., 2008). Thus there are still some answers that remain unexplained. In another study, the full PDB phenotype is observed in mice transfected with measles virus nucleocapsid protein (MNVP) alone, but again this does not exactly reproduce what happens in humans, in which PDB phenotype is found in patients without any measles virus infection (Brown and Michou, 2011; Kurihara et al., 2011). Anecdotal reports published symptoms of *osteitis deformans* (referred in humans as PDB) in other animals, like macaque (Hughes and Lang, 1971), rat snake (Hajkova and Knotek, 1998) and python (Preziosi et al., 2007), but once again these models did not develop the full PDB phenotype. More recently, in 2009, Chamoux et al. showed that the p.Pro392Leu mutation actually activates pagetic osteoclasts derived from cultures of peripheral monocytes or bone marrow cells from PDB patients; however they hypothesized that additional events, like environmental or other genetic factors, were eventually needed to lead to the complete phenotype of PDB (Chamoux et al., 2009). Since there are still other questions that remain to be answered, mostly regarding all the factors that are important for the onset of PDB or why do some patients have several bones affected and others have just monostotic lesions, it is imperative to try to develop another model to better understand the more basic aspects of this gene' expression and its contribution to the physiology of PDB disease. Ray-finned fish, particularly zebrafish (*Danio rerio*), are used as experimental models in various research areas such as environmental and developmental biology, genetics, pharmacology and cancer research. Fish relevant features, that demonstrate their suitability for developmental studies, include i) presence of orthologs for many mammalian genes, ii) strong resemblance of biochemical and physiological processes between fish and mammals, and iii) several technical advantages such as external fertilization, transparent embryos and larvae, rapid and external growth and a large progeny (Bolis et al., 2001). In addition, other similarities to mammals, like organ morphology and systems

composition, as well as a complex ossified skeleton, the existence of osteocytic bone and multinucleated osteoclast (Hall and Witten, 2007; Witten and Huysseune, 2009), have turned zebrafish into an unique animal model to study different pathologies, including those affecting bone (Lieschke and Currie, 2007). Although there are many similarities between fish and mammalian bone there are but also differences (Apschner et al., 2011), being the existence of acellular bone (devoided of osteocytes) and mononucleated osteoclasts in most teleost fish species, while mammals have exclusively cellular bone and multinucleated osteoclasts, the major difference. Fish mononucleated osteoclasts are active bone resorbing cells and were shown to contribute to allometric bone growth (Witten et al., 2001). The availability of fish mutants exhibiting features resembling human pathologies, the fact that amino acid sequences of functionally relevant protein domains have been evolutionarily conserved with many similarities to mammalian sequences both at the protein and nucleotide levels (Lieschke and Currie, 2007), among others, has contributed to accept zebrafish as an important and suitable model to study human biological processes, in particular the mechanisms underlying skeleton development and the onset of skeletal disorders. For example, in 2004 Kan group described a transgenic mice overexpressing BMP4 that developed fibrodysplasia ossificans progressiva (FOP)-like phenotype (Kan et al., 2004). FOP is a rare genetic disorder of progressive extraskeletal ossification and results in profoundly decreased mobility of affected individuals. However it was only when Shen's group developed a zebrafish model of FOP that the full molecular mechanism of BMP independent cascade was elucidated (Shen et al., 2009). They showed that the recurrent single nucleotide missense mutation (c.617G→A) in the gene encoding ACVR1, a BMP type I receptor that causes FOP in all classically affected individuals, induced enhanced chondro-osseous differentiation and activate the BMP signaling pathway in a BMP-independent manner, showing that zebrafish is an excellent tool to unravel the molecular mechanisms related to human bone diseases. Only then the mouse model harboring the same mutation and the full FOP phenotype was developed (Chakkalakal et al., 2012), highlighting the important role of zebrafish to assess the functional mutation that caused FOP. Recently, Wiweger et al (2012) have shown that zebrafish harboring mutations in *exostosin* (*ext2*) gene are a model for Multiple Osteochondromas (MO) since ext2−/− fish have skeletal defects that resemble those seen in osteochondromas and have also dental defects associated with the formation and the morphology of

teeth (Wiweger et al., 2012). Their findings from zebrafish model were validated in a dental survey in MO patients showing the presence of the malformed and/or displaced teeth with abnormal enamel. Several studies show how bone remodeling in zebrafish can be compared to the same process in mammals. Zebrafish osteoclasts exhibit a high level of tartrate-resistant acid phosphatase (TRAP) activity (Witten et al., 2001), first inside bone-resorbing cells and then at bone surfaces undergoing resorption. These eroded bone surfaces are then covered with an overlay of new bone, showing a parallelism between fish and human bone remodeling processes (Witten et al., 1997). Besides bone, fish scales can also be a good system to study the cells and mechanisms responsible for regeneration, development and skeletal remodeling (de Vrieze et al., 2011). The availability of zebrafish mutants, like *panther,* that lacks a functional c-fms (receptor for macrophage colony-stimulating factor) gene, crucial for osteoclast proliferation and differentiation, also shows that zebrafish can be a highly relevant model to study phenotypes related to osteoclast dysfunction and altered bone remodeling (Chatani et al., 2011). In addition to bone phenotype, there is more and more evidence that *SQSTM1* mutations may be causal in frontotemporal dementia and amyotrophic lateral sclerosis (ALS). Protein aggregation was found in patients with theses pathologies and some findings provide evidence of a direct genetic role of *SQSTM1* in ALS pathogenesis as it is supposed to be involved in protein-recycling systems by regulating protein degradation pathways (Fecto et al., 2011). However there are still some molecular mechanisms that are not fully understood. Since zebrafish was shown to be a good animal model to analyze protein aggregation in the neural system (Bellipanni et al., 2000; Xi et al., 2011), it may also be a good model to study SQSTM1 role in protein aggregation.

Although no work is available in zebrafish for *sqstm1*, our analysis of the available sequences encoding functional regions of the *sqstm1* gene indicate that they are fully conserved between zebrafish and human, suggesting that its function may also be conserved. This emphasizes the likeliness that phenotypes with some similarities to those observed in mice and human *SQSTM1* mutants may also be developed in fish. For that the main goal of our work was to provide an initial contribution aiming to validate the zebrafish as an animal model to study *SQSTM1*. To achieve this goal we first compared (1) the *SQSTM1* gene structure, (2) the neighbor genes surrounding *SQSTM1* and (3) the corresponding protein sequences, collected from several fish species, with the available human data. Then we focused our analysis only in the

human and zebrafish species and compared (4) the regulatory motifs in the promoters of *SQSTM1* genes, (5) the splicing isoforms and their tissue specificity and (6) the conservation of *SQSTM1* 3D structure between these two species, to provide insights towards the suitability of zebrafish to be used as a model for SQSTM1-related PDB.

5.1.3 Materials and Methods

5.1.3.1 Sequence collection

Databases at GenBank (www.ncbi.nlm.nih.gov) and Ensembl (release v72; www.ensembl.org) were searched for annotated SQSTM1 sequences. A total of 32 species were selected, including fifteen mammals (*Bos taurus, Callithrix jacchus, Cricetulus griseus, Gorilla gorilla, Homo sapiens, Loxodonta africana, Macaca mulatta, Macaca fascicularis, Mus musculus, Oryctolagus cuniculus, Ovies aries, Pan troglodytes, Pongo abelii, Rattus norvegicus* and *Sus scrofa*), eight teleosts (*Danio rerio, Gasterosteus aculeatus, Oreochromis niloticus, Oryzias latipes, Salmo salar, Takifugu rubripes, Tetraodon nigroviridis* and *Xiphophorus maculatus*), three birds (*Gallus gallus, Meleagris gallopavo* and *Taeniopygia guttata*), one reptile (*Anolis carolinensis*), two amphibian (*Xenopus laevis* and *Xenopus tropicalis*), and three insects (*Apis florea, Megachile rotundata* and *Nasonia vitripennis*).

5.1.3.2 Sequence alignment and analysis

The amino acid sequence alignments were created using T-Coffee multiple sequence alignment software (Notredame et al., 2000) with parameters set to default. Pairwise sequence identity values were computed as percent of identical residues over the total number of aligned residues using T-Coffee generated alignments and the Sequence Manipulation Suite (Stothard, 2000).

5.1.3.3 Assessments of conserved synteny

To examine patterns of conserved synteny, chromosomal loci of *SQSTM1* genes in human, mouse, zebrafish, tilapia, platyfish, tetraodon, turkey, lizard and frog were compared by identifying all neighbor genes of *SQSTM1*. The position of each of these

genes was searched in all species using the Ensembl database search function (release v72; http://www.ensembl.org).

5.1.3.4 Identification of alternative spliced isoform and alternative promoter of *SQSTM1* genes

The alternative splicing events in both human and zebrafish were revealed by aligning the cDNAs against the genomic sequences, using the mRNA alignment tool Spidey (ncbi.nlm.nih.gov/spidey/spideyweb.cgi). The presence and features of alternative promoters, tissue expression and protein interactions of the human *SQSTM1* gene were examined with the AceView database (http://www.ncbi.nlm.nih.gov/IEB/Research/Acembly/index.html). The presence and features of spliced isoforms of the zebrafish *sqstm1* gene were examined using the Ensembl database.

5.1.3.5 Prediction of human and zebrafish SQSTM1 three dimensional structures

Human and zebrafish SQSTM1 three-dimensional (3D) structures were predicted with I-TASSER software (http://zhanglab.ccmb.med.umich.edu/I-TASSER) (Roy et al., 2010). Different models were obtained and the best one was chosen according to the DOPE (discrete optimized protein energy) score method (Shen and Sali, 2006). 3D overlapping illustrations were produced using Chimera v1.5 (http://www.cgl.ucsf.edu/chimera) (Pettersen et al., 2004).

5.1.3.6 Analysis of putative transcription factor (TF) binding sites in SQSTM1 regulatory regions

The promoter sequences (1,500 bp upstream of the translation start site) of the *SQSTM1* genes from human and zebrafish were retrieved from Ensembl database, and analyzed for binding sites of 12 TFs (ETS1, MEF2, AP1, TWIST, KRUEPPEL, RXR, LIM, CART1, NF-κB, ERE, GATA, and NFAT) using the web server ConTra (http://bioit.dmbr.ugent.be/contrav2/index.php) (Broos et al., 2011).

5.1.4 Results

5.1.4.1 *SQSTM1* gene structures were conserved throughout evolution

By searching the Ensembl database, we found that the zebrafish *sqstm1* gene is located in chromosome 14 and appears to be a single copy gene, as already suggested by Mostowy et al. (2013) (Mostowy et al., 2013). The *sqstm1* cDNA was aligned with the *sqstm1* genomic sequence, and sites of exon-intron borders were deduced by comparison. The *sqstm1* gene, spanning approximately 5.5 kb in length, is organized into 8 exons and 7 introns and comparison with its orthologs from other metazoan organisms revealed a generally well-conserved exon-intron organization. Indeed, only one gene was identified in all the species analyzed (Figure 5.1).

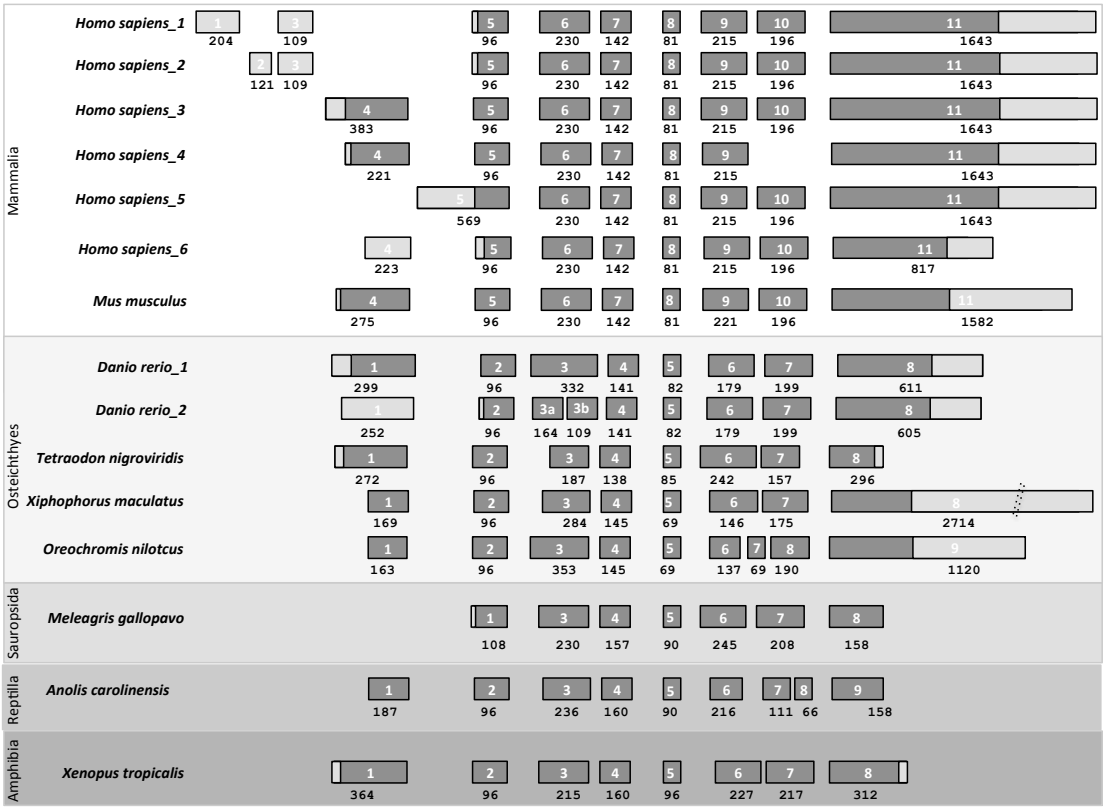


Figure 5.1 - Structural organization of *SQSTM1* gene among species. Schematic diagram of *SQSTM1* genes in humans (*Homo sapiens*), mice (*Mus musculus*), turkey (*Meleagris gallopavo*), zebrafish (*Danio rerio*), platyfish (*Xiphophorus maculatus*), tetraodon (*Tetraodon nigroviridis*), tilapia (*Oreochromis nilotcus*), green anole (*Anolis carolinensis*), and Western clawed frog (*Xenopus tropicalis*). Panels show a *SQSTM1* gene structure comparison among several taxonomic groups (mammalia, osteichthyes, sauropsida, reptilla and amphibia). Six human *SQSTM1* transcripts and two zebrafish *sqstm1* transcripts are also shown. Dark grey boxes represent coding exons, light grey boxes represent non-coding exons. Numbers below the boxes represent their length, in nucleotides. Only exon sizes are to scale. Homo sapiens 1 - ENST00000376929; NM_001142298; uc011dgr.2; Homo sapiens

2 -NM_001142299, uc011dgs.2; Homo sapiens 3 - ENST00000389805; NM_003900; uc003mkw.4; Homo sapiens 4 - ENST00000510187; Homo sapiens 5 -ENST00000360718; uc003mkx.3; Homo sapiens 6 - ENST00000402874. ENST00000514093, ENST00000422245, ENST00000508284, ENST00000504627, ENST00000453046 are protein-coding isoforms (and appear in Table S5.1) but the sequences are not complete.

Zebrafish *sqstm1* exon 1 contains the 5´UTR and the start codon and exon 8 contains the termination codon and the 3´UTR, including two consensus polyadenylation signals (Figure 5.1 and results not shown). We have identified two zebrafish *sqstm1* transcripts coding for protein isoforms Danio rerio_1 (ENSDART00000140061) and Danio rerio_2 (ENSDART00000113012) with 1939 and 1827 bp, with ORFs encoding sequences of 452 and 353 amino acids, respectively (Figure 5.1 and Table S5.1). The transcript Danio rerio_1 seems to have retained the intron 3 when comparing with the genomic structure among the other species, and thus, to allow an easier comparison, we have maintained the same numbering for the zebrafish *sqstm1* exons (Figure 5.1).

The human *SQSTM1* gene is located in chromosome 5 and spans 32 kb in length and at least six transcripts coding for protein isoforms were identified (Figure 5.1) with 2913, 2848, 2986, 1714, 2253 and 2000 bp, with ORFs encoding sequences of 356, 356, 440, 378, 356 and 356 amino acid, respectively (Figure 5.1 and Table S5.1). Exons 1, 2, 3 and 4 are usually skipped by alternative splicing at the mRNA level to produce the following isoforms: Homo sapiens_1 (lacks both exons 2 and 4), Homo sapiens_2 (lacks both exons 1 and 4), Homo sapiens_3 (lacks exons 1-3), Homo sapiens_4 (lacks exons 1-3 and 10), Homo sapiens_5 (lacks exons 1-4 and retains part of intron 4), and Homo sapiens_6 (lacks exons 1-3 and partial 4). We also examined the occurrence of splicing isoforms of human *SQSTM1* gene by using the AceView-NCBI database, a database generated from the experimental data of cDNA sequencing from various human tissues mRNA (Thierry-Mieg and Thierry-Mieg, 2006). As shown in Table S5.1, we found that the human *SQSTM1* encodes multiple open reading frames and has 17 alternatively spliced isoforms (named in alphabetical order, *a* to *q*, based on their protein size) expressed in different tissues. As an example, *SQSTM1* isoform *a* is encoded by 8 exons and results in a 2986 bp mRNA transcript encoding a protein with 440 amino acids, which is expressed in melanoma, skin, lung, and 16 other tissues.

The gene structures for all the species analyzed were constructed according to available data on transcript exons size and number obtained from Ensembl database

after alignment with the genomic sequence using Spidey facilities. We observed that almost every species had a conserved overall gene structure when compared with the human ortholog. It is worthy to mention that although in human there are evidences for the existence of *SQSTM1* transcripts with 5' non coding exons (Figure 5.1) we did not found the same evidences in other species. More important, the coding exons for each species are extremely well conserved which could indicate that this gene encodes a protein with important functions and thus was subjected to an evolutionary pressure responsible for maintaining its molecular structure and sequence.

In order to better understand the relationships between zebrafish and human *SQSTM1*, we next examined the genomic context of *SQSTM1* gene in human chromosome 5, and how this is conserved in other vertebrate species (Figure 5.2). The overall clusters of genes that surround *SQSTM1* were maintained in all the species analyzed showing conservation also at the level of its genomic environment.

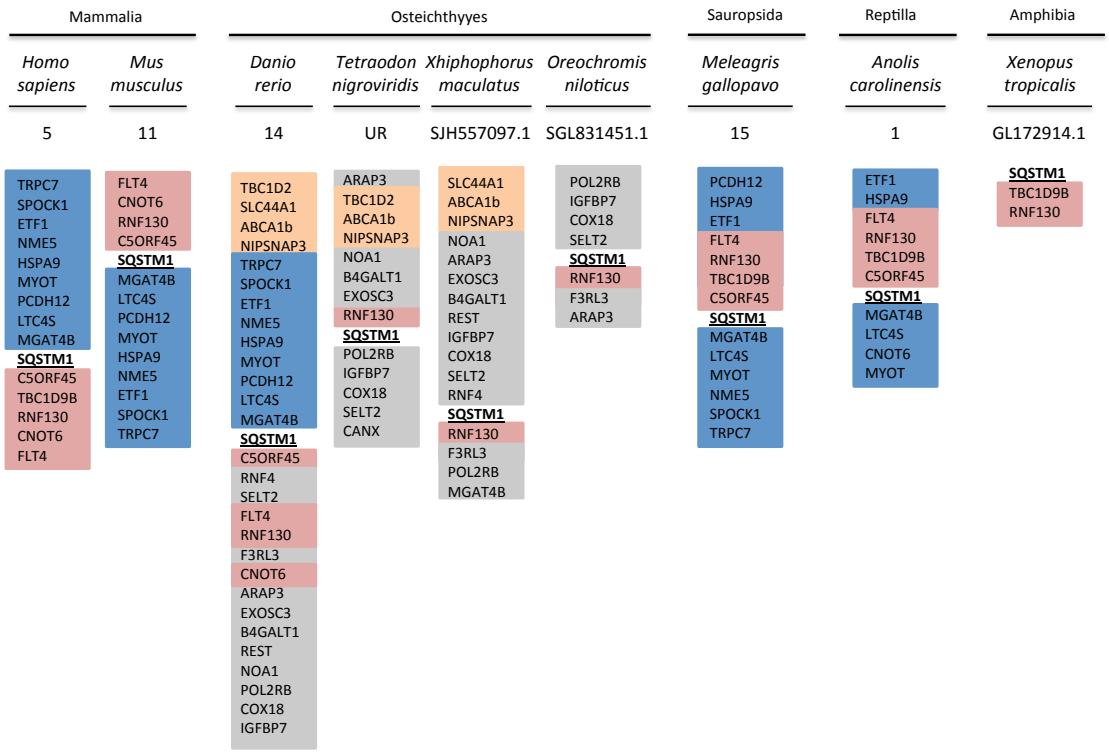


Figure 5.2 - Overall synteny relationships between *SQSTM1* genes. Location of *SQSTM1* genes of two mammals, four teleosts, one bird, one reptile and one amphibian are shown in boxes. Depending on gene clusters, syntenic blocks represented by blue, pink, light grey, and orange boxes. Numbers below bars represent the chromosome or scaffold number in each species. UR in Tetraodon nigroviridis represents UnRandom.

The genes flanking the *SQSTM1* orthologs can be found in two major clusters: one more conserved in mammalian species, which can also be found in zebrafish, including *LTC4S, MGAT4b, RNF130* and *FLT4* (represented in blue and pink in

Figure 5.2), and another one more conserved in fish species, which includes *abca1b*, *arap3, selt2, cox18, polr2b and igfbp7* (represented in grey and orange in Figure 5.2). The arrangement of the genes close to *SQSTM1* was conserved to a large degree in human, mouse, turkey and anole lizard. Genome sequence data in frog was rather limited, and thus synteny analysis was less informative in this vertebrate.

In order to study the evolutionary relationship of zebrafish Sqstm1 with mammalian SQSTM1, we constructed a phylogenetic tree by aligning all 32 collected amino acid sequences of this protein (Figure 5.3). The phylogenetic tree reflects both the taxonomy and the specificity. In agreement with the generally accepted taxonomy of vertebrates, Sqstm1 sequences were clustered into major taxonomic groups (such as insects, reptiles, bony fish, amphibians, birds and mammals).

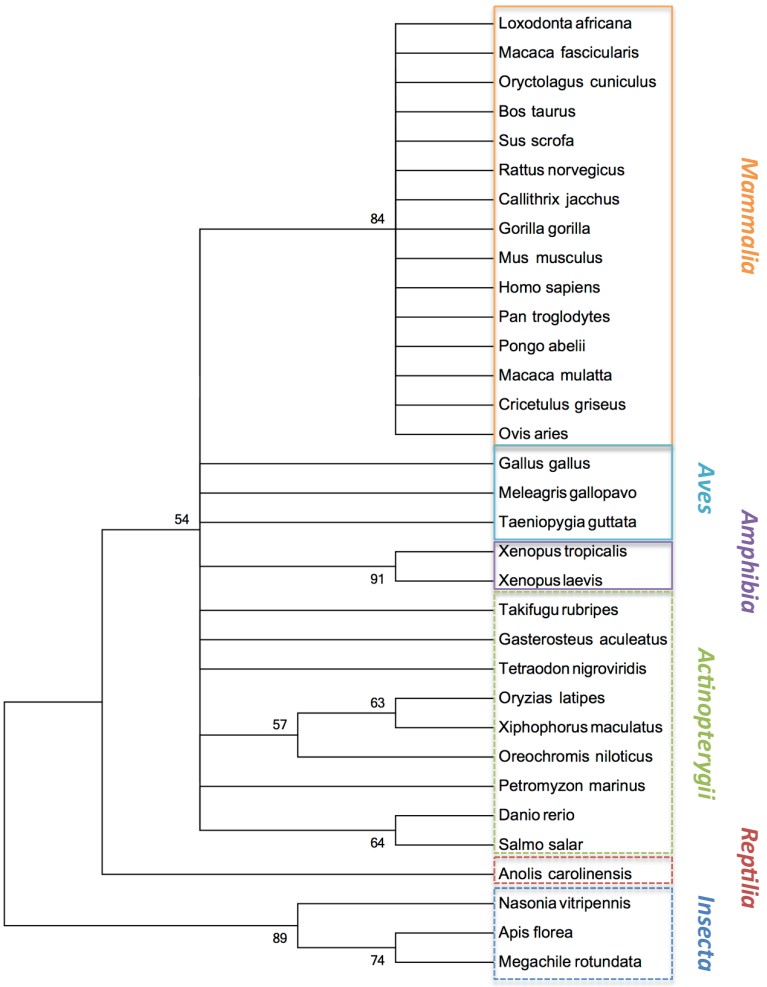


Figure 5.3 - Phylogenetic analysis of SQSTM1 proteins. Phylogenetic relationship of SQSTM1 proteins and evolutionary distances among different species. Neighbor-joining tree was constructed with MEGA facilities using taxonomic data available at NCBI. SQSTM1 proteins are clustered in major taxonomic groups (such as insects, reptiles, amphibians, bony fishes, birds and mammals).

5.1.4.2 SQSMT1 primary and tertiary structures are conserved between human and zebrafish

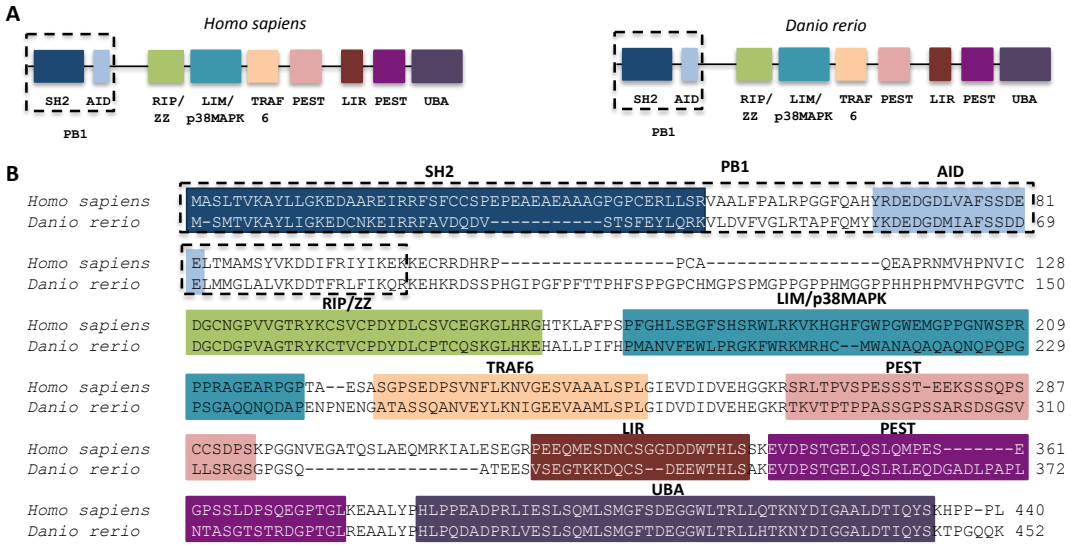
In order to study the protein conservation at the amino acid level, between zebrafish and human, we performed pairwise alignments using species ranging from fish to mammals and the identity percentage acquired with Manipulation Suite facilities revealed an overall sequence identity of 91% between mammalian SQSTM1 proteins and 50% between fish SQSTM1 proteins (Figure 5.4). Particularly, the amino acid sequence of zebrafish Sqstm1 is 43% identical the human SQSTM1 (Figure 5.4).

	Dre	Gac	Tru	Tni	Oni	Ssa	Xma	Mmus	Hsa	Ggo	Cja	Rno	Ssc	Bta	Ocu	Laf	Ptr	Pab	Mmu	Oai	Xtr	Xla
Dre	452																					
Gac	48	421																				
Tru	48	60	420																			
Tni	49	61	69	421																		
Oni	47	58	58	63	436																	
Ssa	58	48	48	48	48	463																
Xma	51	61	63	64	66	51	397															
Mmus	43	44	46	43	42	43	46	442														
Hsa	43	44	45	43	41	42	46	91	440													
Ggo	43	44	45	43	41	42	46	91	99	440												
Cja	43	44	45	43	41	42	45	91	96	96	439											
Rno	43	44	46	43	42	43	45	97	91	91	91	439										
Ssc	44	44	45	44	41	43	45	91	93	93	92	91	440									
Bta	43	44	45	44	41	42	45	88	92	92	90	89	94	440								
Ocu	43	45	46	45	42	42	46	90	93	93	91	90	94	91	521							
Laf	41	40	41	43	40	40	46	78	80	81	80	78	81	79	79	445						
Ptr	43	44	43	43	41	42	46	91	100	100	96	91	93	93	93	81	440					
Pab	43	44	44	43	41	42	45	90	99	99	95	90	92	92	92	80	99	436				
Mmu	43	44	45	43	41	42	46	90	98	98	96	91	93	92	93	81	98	97	440			
Oai	43	44	45	44	41	42	46	89	94	94	92	89	94	98	92	80	94	93	93	440		
Xtr	41	41	43	43	39	39	44	54	53	53	53	53	53	52	54	47	53	53	53	53	452	
Xla	41	41	44	43	39	40	44	52	52	52	52	52	52	52	53	47	52	52	52	53	90	454

Figure 5.4 - Pairwise percent identities among SQSTM1 protein sequences. Diagonal values are sequence lengths in amino acids and shaded areas indicate identities within different groups of organisms (purple - amphibian; orange - mammalian; green - teleosts). The identity values were calculated from alignments described in the text. Atlantic salmon (*Salmo salar*, Ssa), African clawed frog (*Xenopus laevis*, Xla), African savanna elephant (*Loxodonta Africana*, Laf), cattle (*Bos taurus*, Bta), chimpanzee (*Pan troglodytes*, Ptr), common marmoset (*Callithrix jacchus,* Cja), house mouse (*Mus musculus*, Mmu), human (*Homo sapiens*, Hsa), Nile tilapia (*Oreochromis niloticus*, Oni), Norway rat (*Rattus norvegicus*, Rno), pig (*Sus scrofa*, Ssc), rabbit (*Oryctolagus cuniculus,* Ocu), rhesus monkey (*Macaca mulatta*, Mm), sheep (*Ovies aries*, Oar), Southern platyfish (*Xiphophorus maculatus*, Xma), spotted green pufferfish (*Tetraodon nigroviridis*, Tni), stickleback (*Gasterosteus aculeatus*, Gac), Sumatran Orangutan (*Pongo abelii*, Pab), torafugu (*Takifugu rubripes*, Tru), Western clawed frog (*Xenopus tropicalis*, Xtr), Western gorilla (*Gorilla gorilla*, Ggo), zebrafish (*Danio rerio*, Dre).

Protein size was also found to be conserved along evolution, with zebrafish presenting a Sqstm1 encoding 452 amino acids, and the human SQSTM1 encoding 440 amino acids (Figure 5.4). Interestingly, based on T-coffee alignment, the amino acid sequences identity in the main domain of the SQSTM1 are high, with 81% identity in ubiquitin-associated (UBA) domain (Figure 5.5), a domain known to be involved in SQSTM1 autophagy (Daroszewska et al., 2011; Ravikumar et al., 2010). On the other

hand the LIM/p38MAPK domain is less conserved, sharing only 17% of the amino acid sequence (Figure 5.5).



C

Protein Domain	*Homo sapiens* lenght (aa)	*Danio rerio* length (aa)	Share identity (%)
SH2	50	40	38
AID	16	16	75
RIP/ZZ	35	35	71
LIM/p38MAPK	50	52	17
TRAF6	26	26	54
PEST	29	30	30
LIR	21	19	48
PEST	33	40	45
UBA	50	50	81

Figure 5.5 - SQSTM1 protein domains conservation between zebrafish and human. A) Schematic representation of SQSTM1 protein domains in human (*Homo sapiens*) and zebrafish (*Danio rerio*) proteins. B) Alignment of SQSTM1 proteins from human and zebrafish. The boxes highlight the protein domains represented in panel A. C) Comparison of protein domains size in both species and pairwise percent identities among these domains. aa – amino acids. . SH2 - Src-homology 2 domain, AID - acidic interaction domain, RIP/ZZ - AMPA receptor interacting protein/zinc finger region, PEST - proline, glutamic acid, serine, and threonine rich domains, PB1 - Phox and Bem1p domain, (p38/MAPK) - p38 mitogen-activated protein kinase, TRAF6 - tumor necrosis factor receptor-associated factor 6 binding domain, LIR - LC3-interacting region, UBA - ubiquitin-associated domain.

Of relevance to the conservation of these domains and possibly their function is the fact that residue Pro392, shown in Figure 5.6, known to be the most frequently mutated in PDB patients (Hocking et al., 2001; Laurin et al., 2002, 2001), is 100% conserved in protein sequences from all species analyzed (Figure 5.6).

```
Bos_taurus                  EGPTGLKEAALYPHLPPEADPRLIESLSQMLSMGFSDEGGWLTRLLQTKNYDIGAALDTIQY
Callithrix_jacchus          EGPTGLKEAALYPHLPPEADPRLIESLSQMLSMGFSDEGGWLTRLLQTKNYDIGAALDTIQY
Gorilla_gorilla             EGPTGLKEAALYPHLPPEADPRLIESLSQMLSMGFSDEGGWLTRLLQTKNYDIGAALDTIQY
Homo_sapiens                EGPTGLKEAALYPHLPPEADPRLIESLSQMLSMGFSDEGGWLTRLLQTKNYDIGAALDTIQY
Loxodonta_africana          EGPTGLKEAALYPHLPPEADPRLIESLSQMLSMGFSDEGGWLTRLLQTKNYDIGAALDTIQY
Macaca_mulatta              EGPTGLKEAALYPHLPPEADPRLIESLSQMLSMGFSDEGGWLTRLLQTKNYDIGAALDTIQY
Mus_musculus                EGPTGLKEAALYPHLPPEADPRLIESLSQMLSMGFSDEGGWLTRLLQTKNYDIGAALDTIQY
Oryctolagus_cuniculus       EGPTGLKEAALYPHLPPEADPRLIESLSQMLSMGFSDEGGWLTRLLQTKNYDIGAALDTIQY
Ovis_aries                  EGPTGLKEAALYPHLPPEADPRLIESLSQMLSMGFSDEGGWLTRLLQTKNYDIGAALDTIQY
Pan_troglodytes             EGPTGLKEAALYPHLPPEADPRLIESLSQMLSMGFSDEGGWLTRLLQTKNYDIGAALDTIQY
Pongo_abelii                EGPTGLKEAALYPHLPPEADPRLIESLSQMLSMGFSDEGGRLTRLLQTKNYDIGAALDTIQY
Rattus_norvegicus           EGPTGLKEAALYPHLPPEADPRLIESLSQMLSMGFSDEGGWLTRLLQTKNYDIGAALDTIQY
Sus_scrofa                  EGPTGLKEAALYPHLPPEADPRLIESLSQMLSMGFSDEGGWLTRLLQTKNYDIGAALDTIQY
Danio_rerio                 DGPTGLREAALYPHLPQDADPRLVESLSQMLSMGFTDEGGWLTRLLHTKNYDIGGALDTIQY
Gasterosteus_aculeatus      LGQTGLKEAALYPHLPQEADPRLVESLSLMLSMGFTDEGGWLTRLLQAKNFDIGAALDAIQY
Oreochromis_niloticus       AGPSGLREAALYPHLPQEADPRLVESLASMLSMGFTDEGGWLTRLLQAKNFDIGAALDAIQY
Salmo_salar                 QKSLTLAEAAAYPHLPQDADPRLVESLSQMLAMGFTDEGGWLTRLLHTKDCDVGAALDTIHY
Takifugu_rubripes           QHPTGLKEAALYPHLPDEADPRLVESLAQMLSMGFTDEGGWLTRLLQAKNYDIGAALDAIQY
Tetraodon_nigroviridis      --PAGLKEAALYPHLPEEADPRLVESLAQMLSMGFTDEGGWLTRLLQAKNFDIGAALDAIQF
Xiphophorus_maculatus       QGPTGLREAALYPHLPQEADPRLVESLAAMLSMGFGDEGGWLTHLLQAKNGDIGAALDAIQY
Xenopus_laevis              HAPTGLREAALYPHLPPEADPRLIETLSQMLSMGFSDEGGWLTRLLEAKQYDIGSALDAMQS
Xenopus_tropicalis          HAPTGLREAALYPHLPPEADPRLIESLSQMLSMGFTDEGGWLTRLLEAKQFDIGSALDAMQS
```

Figure 5.6 – Conservation of Pro392 amino acid. Alignment of SQSTM1 proteins with ClustalW showing the conservation of the Pro392, an important amino acid for SQSTM1 and highly associated with Paget's disease of bone, in all species analyzed, represented by the grey box.

Zebrafish Sqstm1 not only presents protein domains similar to those found in the human protein but they also show remarkable three-dimensional (3D) structure conservation (Figure 5.7), an additional evidence for the existence of an evolutionary pressure to maintain the structure and function of SQSMT1.

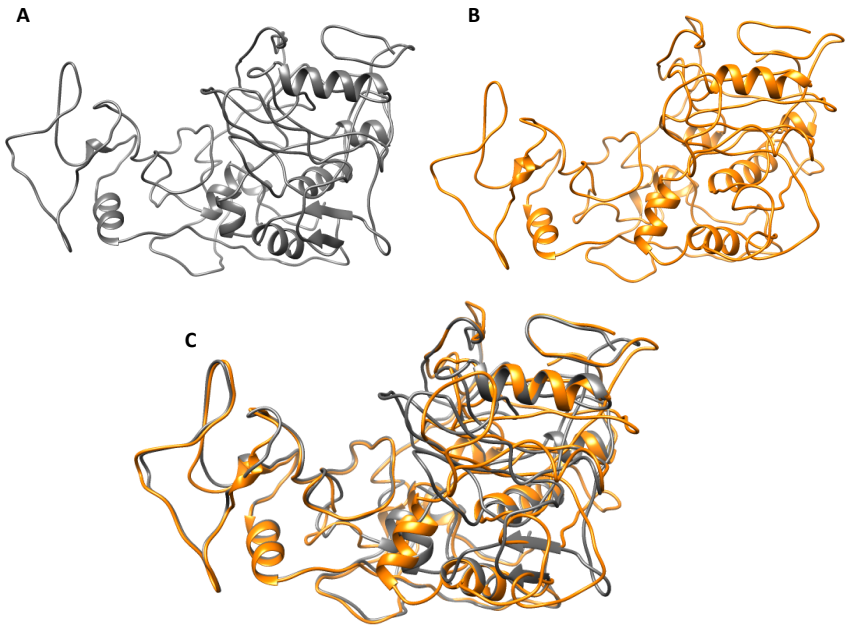


Figure 5.7 – Human and zebrafish SQSTM1 3D structure. Prediction of human and zebrafish SQSTM1 three dimensional (3D) structure. A) 3D structure of Sqstm1 from zebrafish. B) 3D structure of SQSTM1 from human. C) 3D alignment of zebrafish (dark grey) and human (orange) SQSTM1.

Accordingly, and as shown in Figure 5.7, despite differences existing at the amino acid level, the predicted 3D structure of zebrafish Sqstm1 is very similar to that of its human counterpart, with the important structural amino acids being conserved in the same spatial position, further indicating that zebrafish Sqstm1 is likely to have an *in vivo* function similar to that of human SQSMT1, and emphasizing the relevance of zebrafish as a suitable animal model for SQSMT1 related research.

5.1.4.3 *SQSTM1* is regulated by bone related transcription factors both in human and in zebrafish

Transcription factors (TFs) are nuclear proteins that regulate the expression of target genes. We examined the presence of transcription factors binding sites (TFBSs) known from the literature to bind human *SQSTM1*, in the promoters of human and zebrafish *SQSTM1* using Contra facilities. Data in Figure 5.8 also suggests that both human and zebrafish *SQSTM1* appear to be putative targets of TFs known to be important for regulation of bone metabolism like ETS1, NF-κB, AP1, Twist, GATA, ERE, RXR, Krueppel, LIM, Cart-1, Mef2 and NFAT, which could explain its role in bone biology. Interestingly, the transcription factor binding elements Cart-1 and AP1 identified in the human promoter are found to be conserved in the same position in the zebrafish promoter. In addition to TFBSs, both a TATA and a CAAT boxes were also conserved in the zebrafish promoter, which may indicate the existence of similar mechanisms for regulation of expression.

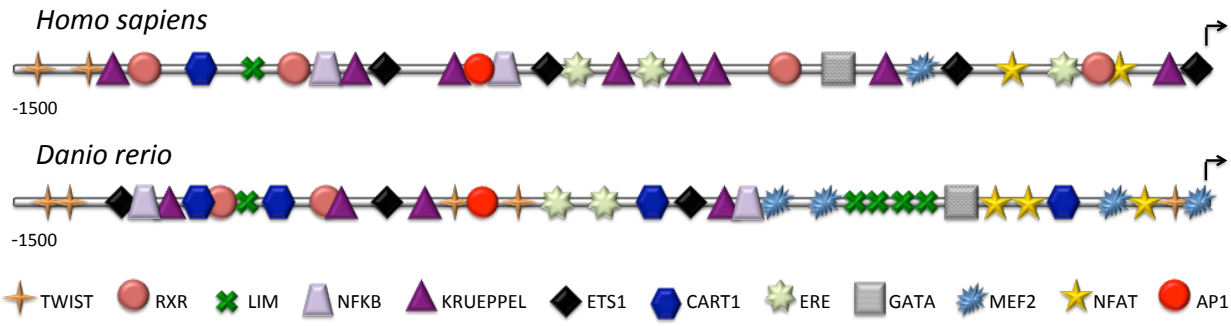


Figure 5.8 – Comparative *SQSTM1* promoter analysis between human and zebrafish. Sequence analysis of *SQSTM1* promoter. Several putative transcription factors binding motifs were identified with Contra analysis. The numbering is related to the translation start site.

5.1.5 Discussion

Fish and humans present large similarities when comparing gene structure and function and their involvement in physiological processes. In the case of zebrafish *sqstm1,* given its conservation in terms of gene structure and protein organization when compared to mammalian genes and its predicted regulation by bone related TFs, it is likely that it also plays an important role in fish bone. In this work we analyzed SQSTM1 gene and protein conservation between human and zebrafish and the results showed that SQSTM1 is well conserved throughout evolution. The overall clusters of genes that surround *SQSTM1* were maintained in all the species analyzed showing conservation also at the level of its genomic environment. Postlethwait et al. (1998) compared the chromosomal locations of zebrafish genes with their mammalian orthologs and found highly conserved syntenic relationships between them (Postlethwait et al., 1998). These comparative analyses could help to predict the positions of some genes from mammalian gene maps and also provide new insights into the corresponding gene function by comparison with the known functions of its human orthologs (Gates et al., 1999). Interestingly, both in human and in zebrafish, the gene closer to *SQSTM1* is *MGAT4b*, a gene encoded in the opposite direction, which in zebrafish is 11,7 kb apart, while in human both genes share the first exons (exons 1-3).

Regarding the protein sequence, the results showed a high amino acid sequences identity in the main domain of the SQSTM1, with 81% identity in ubiquitin-associated (UBA) domain, a domain known to be involved in SQSTM1 autophagy (Daroszewska et al., 2011; Ravikumar et al., 2010). More interesting is the conservation of the residue Pro392, known to be the most frequently mutated in PDB patients (Hocking et al., 2001; Laurin et al., 2002, 2001), is 100% conserved in protein sequences from all species analyzed (Figure 5.6). This finding confirms that this key amino acid in the UBA domain is crucial to SQSMT1 function, providing additional evidence for the maintenance of the autophagy function in this protein. Studies in mouse showed that this mutation is sufficient to cause several aspects of the PDB phenotype but by itself was not sufficient to induce the complete PDB phenotype (Daroszewska et al., 2011; Hiruma et al., 2008). It was hypothesized that further events, like environmental or other genetic factors were eventually needed to

lead to the full development of PDB phenotype (Chamoux et al., 2009; Hiruma et al., 2008; Roodman, 2010). Interestedly, the residue Pro392 is located in the human *SQSMT1* exon 10, an exon that is lost in one of the identified transcript isoforms (Homo sapiens_4; Figure 5.1). This exon codes for the second PEST domain (formed by amino acids 345-377, Figure 5.5) a region rich in proline (P), glutaminic acid (E), serine (S), and threonine (T) (Rechsteiner and Rogers, 1996; Rogers et al., 1986) and containing many S/TP sequences, which are minimum consensus phosphorylation sites for protein kinases. It was hypothesized that disruption of a PEST sequence could result in a mutant SQSTM1 protein that is more stable than the wild type, inducing a more sustained signaling (Beyens et al., 2006).

Using bioinformatics tools we found in zebrafish *sqstm1* promoter some putative TFBSs already known to regulate human *SQSTM1* (Du et al., 2009; Thompson et al., 2003; Vadlamudi and Shin, 1998), like ETS1 and NF-κB (Figure 5.8), and found that one ETS1 position in the zebrafish *sqstm1* promoter is conserved, showing that also *SQSTM1* regulation mechanisms seems to be preserved.

Our work shows that, by using bioinformatics tools, we can predict that zebrafish, which presents unique advantages over other vertebrate models (Bolis et al., 2001), is likely to be a suitable model to study some aspects of the function of SQSTM1 in PDB, providing a convenient and powerful *in vivo* platform in which genetic and physiologic characteristics of PDB can be identified and studied (Lieschke and Currie, 2007; Santoriello and Zon, 2012), with particular emphasis on skeletal defects associated to PDB, as already observed for other genetic diseases affecting the skeleton.

5.2 Comparative analysis of human and zebrafish OPTN: molecular and evolutionary perspectives

Iris A.L. Silva
Natércia Conceição
Laëtitia Michou *
M. Leonor Cancela *

This chapter has been submitted as a research paper for publication.

Authors' roles: Study design: MLC and LM. Study conduct: IALS. Data collection: IALS. Data analysis: IALS and NC. Data interpretation: All authors. Drafting manuscript: IALS. Revising manuscript content: NC, MLC and LM. Approving final version of manuscript: all authors.

5.2.1 Abstract

Optineurin (OPTN) is a protein encoded by the *OPTN* gene. This protein is involved in several cellular mechanisms such as autophagy, NF-κB signaling, cellular morphogenesis, membrane and vesicle trafficking, and transcription activation. Mutations in *OPTN* have been described in glaucoma, amyotrophic lateral sclerosis and other neurological diseases. More recently a polymorphism in this gene was also identified by a genome wide association study (GWAS) to be associated with Paget's disease of bone (PDB), making *OPTN* a strong candidate gene to be involved in PDB; however defining the molecular mechanisms by which variants in this gene may contribute to this bone disease requires further studies. Because zebrafish has been validated as a good model to study bone related diseases, the objective of the present work was to evaluate if zebrafish could be a good system to study the molecular mechanisms through which OPTN may contribute to PDB pathogenesis. Through a comparative analysis we observed that OPTN is encoded by a single copy gene both in human and zebrafish and its genomic structure is also conserved. The neighbor genes and chromosomal localization were also maintained, which strongly suggest that zebrafish *optn* is the ortholog of human *OPTN*. Bioinformatics analysis indicate that zebrafish and human *OPTN* seem to be regulated by common transcriptional

factors related to bone such as TWIST, KRUEPPEL, LIM domain, NFAT, NF-κB and RXR. Furthermore, OPTN protein comparison between several species revealed a high degree of conservation in the functional domains of the protein and in its 3D structure. In conclusion, this study demonstrated that OPTN is well conserved throughout evolution, and therefore fish models, like zebrafish, which has been previously validated as a good model to study bone related pathologies, could be considered to further study the biologic role of OPTN in bone diseases and bone development.

5.2.2 Introduction

Optineurin (*OPTN*) is a gene linked to several conditions that go from neurological diseases, such as open-angle glaucoma and amyotrophic lateral sclerosis (ALS) (Kachaner et al., 2012; Osawa et al., 2011; Ying and Yue, 2012), to bone diseases, such as Paget's disease of the bone (PDB) (Albagha et al., 2010). OPTN was originally identified as a gene responsible for primary open-angle glaucoma, a progressive blinding disease, with the loss of the retinal ganglion cells and damage to their axons that make up the optic nerve (Rezaie et al., 2002). Subsequent groups have implicated many different mutations in *OPTN* with this disease (Chalasani et al., 2009; Funayama et al., 2004; Leung et al., 2003; Swarup et al., 2013). ALS is a progressive debilitating condition where the loss of spinal motor neurons leads to paralysis and death, and researchers have also found mutations in *OPTN* to be associated with various forms of this disease (Belzil et al., 2011; Czell et al., 2013; Iida et al., 2012; Lattante et al., 2012; Millecamps et al., 2011). This gene encodes a cytosolic protein that interacts with a number of proteins and therefore is important in basic cellular functions such as vesicle trafficking, maintenance of the Golgi apparatus, NF-κB pathway, antibacterial and antiviral signaling, cell division control, and autophagy. Mutation or level alteration of *OPTN* expression results in adverse consequences in the cells leading to diseases (Gao et al., 2014; Park et al., 2010; Turturro et al., 2014). The molecular mechanisms, however, are not yet understood.

The human *OPTN* gene encodes a protein that contains multiple coiled-coil domains, at least one leucine zipper and a C-terminal zinc finger (Li et al., 1998). The OPTN protein from different species has high amino acid homology and the amino acid in the position 50, which is often mutated in several diseases, is conserved (Rezaie and

Sarfarazi, 2005). Animal models where the effect of mutations in *OPTN* expression is analyzed are mainly focused in mimicking neurological diseases, such as glaucoma or ALS. Chi et al. (2010) described the phenotypic characteristics of transgenic mice overexpressing wildtype or mutated *OPTN* (E50K, H486R, and a deletion of the leucine zipper) (Chi et al., 2010). After 16 months, histologic abnormalities were exclusively observed in the retina of E50K mutant mice, with loss of retinal ganglion cells and connecting synapses in the peripheral retina, thinning of the nerve fiber layer at the optic nerve head at normal intraocular pressure, and massive apoptosis and degeneration of the entire retina. Introduction of the E50K mutation disrupted the interaction between Optn and Rab8 GTPase, a protein involved in the regulation of vesicle transport from Golgi to plasma membrane. Chi et al. (2010) concluded that alteration of the Optn sequence could initiate significant retinal degeneration in mice (Chi et al., 2010). Two other different knock-in mouse lines have been developed that interfere with OPTN function: the $OPTN^{D477N}$ line interferes with ubiquitin binding (Gleason et al., 2011) and the $OPTN^{470T}$ line causes a C-terminal truncation, which also eliminates the ubiquitin-binding domain (Munitic et al., 2013). Neither of these published works described any obvious phenotype for these mouse lines, with the exception of an increase in embryonic lethality related to the genetic background of the $OPTN^{470T}$ line (Munitic et al., 2013). In addition, another paper showed that transient knockdown of *optn* in zebrafish embryos resulted in abnormal morphology and defects in spinal motor axon guidance, a defect that can be related to problems in directed cell migration (Korac et al., 2013). Another work developed in the attempt to try to explain the *in vivo* effects of *optn* mutation was performed by Paulus and Link (2014). This group assessed the cellular consequences of *optn* knock-down in zebrafish embryos and they showed that the loss of *optn* expression resulted in increased cell death, as well as subtle cell morphology, cell migration and vesicle trafficking defects (Paulus and Link, 2014). However no bone phenotypes were accessed in any of those models.

Zebrafish is used as an experimental model in various research areas and has relevant features, including several technical advantages such as external fertilization, transparent embryos and larvae, rapid and external growth and a large progeny, strong resemblance of biochemical and physiological processes with mammals and presence of orthologs for many mammalian genes. Other similarities to mammals, like a complex ossified skeleton, the existence of osteocytic bone and multinucleated

osteoclast makes zebrafish an unique animal model to study different pathologies, including those affecting bone (Bolis et al., 2001; Lieschke and Currie, 2007). Zebrafish can be a good model to compare the bone remodeling mechanisms since zebrafish osteoclasts exhibit high activity of tartrate-resistant acid phosphatase and the eroded bone surfaces are covered with new bone, just like in human (Hall and Witten, 2007; Witten and Huysseune, 2009), and there are models that lack crucial genes for osteoclast proliferation and differentiation, such as *panther* (Chatani et al., 2011).

5.2.3 Materials and methods

5.2.3.1 Sequence collection

Sequence databases at NCBI GenBank (www.ncbi.nlm.nih.gov) and Ensembl (release v72; www.ensembl.org) were searched for annotated OPTN sequences. A total of 11 species were selected, including two mammals [human (*Homo sapiens*) and mouse (*Mus musculus*)], seven teleosts [zebrafish (*Danio rerio*), stickleback (*Gasterosteus aculeatus*), Nile tilapia (*Oreochromis niloticus*), medaka (*Oryzias latipes*), salmon (*Salmo salar*), fugu (*Takifugu rubripes*), and platyfish (*Xiphophorus maculatus*)], one bird [chicken (*Gallus gallus*)], and one amphibian [frog (*Xenopus tropicalis*)].

5.2.3.2 Sequence alignment and analysis

The amino acid sequence alignments were created using T-Coffee software (Notredame et al., 2000) with parameters set to default. Sequence composition logos were constructed from the resulting alignments using the WebLogo facility (Crooks et al., 2004) - sequence logos are pictorial representations of the protein sequence in which the height of each amino acid letter is proportional to its frequency at a particular site. For each protein alignment, pairwise sequence identity and similarity values (p-distances) were computed as percent of identical and similar residues over the total number of aligned residues using the Sequence Manipulation Suite, respectively (Stothard, 2000).

5.2.3.3 Assessments of gene synteny

To determine the patterns of conserved gene synteny, we annotated the genes upstream and downstream of the *OPTN* gene in the four teleost fish species, plus human, chicken and frog using the Ensembl database search function (release v72; http://www.ensembl.org). The synteny pattern was then confirmed using Synteny Database (Catchen et al., 2009).

5.2.3.4 Prediction of human and zebrafish OPTN three dimensional structures

Human and zebrafish OPTN three-dimensional (3D) structures were predicted with I-TASSER software (http://zhanglab.ccmb.med.umich.edu/I-TASSER) (Roy et al., 2010). Different models were obtained and the best one was chosen according to the DOPE (discrete optimized protein energy) score method (Shen and Sali, 2006). 3D overlapping illustrations were produced using Chimera v1.5 (http://www.cgl.ucsf.edu/chimera) (Pettersen et al., 2004).

5.2.3.5 Analysis of putative transcription factor binding sites (TFBSs) in *OPTN* regulatory regions

The promoter sequences (1,500 bp upstream of the transcription start site) of the *OPTN* gene from human and zebrafish were retrieved from Ensembl database, and masked for repetitive elements by the program RepeatMasker (www.repeatmasker.org) using the default mode. The sequences were then analyzed using the web server ConTra (http://bioit.dmbr.ugent.be/contrav2/index.php) (Broos et al., 2011) to determine binding sites for 14 TFs (E2F, E47, SP1, ETS1, MEF2, TWIST, KRUEPPEL, RXR, LIM, CART1, NF-κB, ERE, GATA, and NFAT).

5.2.3.6 *In silico* EST-based gene expression profiles

To determine if *OPTN* genes were differentially expressed in human and in zebrafish tissues, gene expression profiles were analyzed with NCBI Unigene EST profiles. These analyses determined approximate gene expression patterns inferred from transcript counts and cDNA library sources as reported by sequence submitters to the

Unigene database. The EST and Unigene data of NCBI are available in a standardized format, which makes it possible to extract information from the two databases and obtain sufficient data to estimate gene expression profiles. Gene mRNA levels are described as gene transcript units per million transcripts (TPM).

5.2.4 Results

5.2.4.1 *OPTN* gene structure structures were conserved throughout evolution

Human and zebrafish *OPTN* gene sequences and structures were obtained by searching the Ensembl database, and we found that the zebrafish *optn* gene is located in chromosome 4 and appears to be a single copy gene. The *optn* gene has approximately 11 kb and it is organized into 14 exons and 13 introns and comparison with its human ortholog revealed a generally well-conserved exon-intron organization for the coding exons (Figure 5.9).

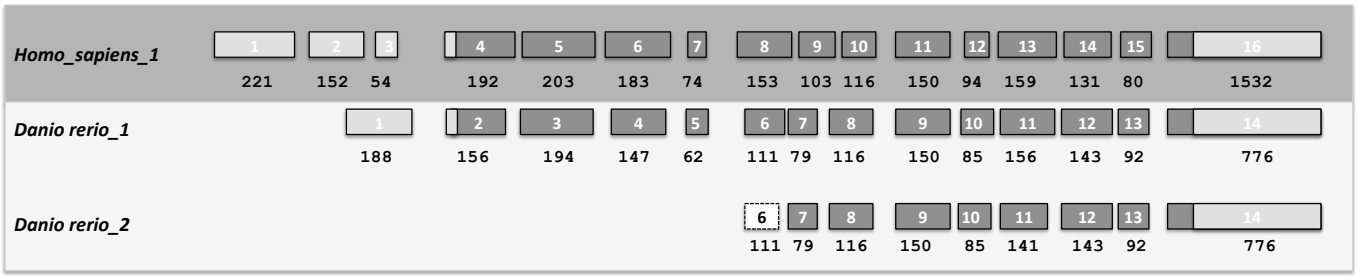


Figure 5.9 - Comparative analysis of human and zebrafish *OPTN* gene structure. The comparison between human and zebrafish overall gene structure revealed a high conservation at genomic level. The gene structure, namely the number of coding exons is conserved, which is the first indication of a strong gene function conservation. Dark grey boxes represent exonic coding regions, light grey boxes represent exonic non-coding regions. Numbers below the boxes indicate exons length in base pairs.

Zebrafish *optn* exon1 and 2 span the 5´UTR and the start codon is present in exon 2. Exon 14 contains the termination codon and the 3´UTR that includes one consensus polyadenylation signal (Figure 5.9). We have identified two zebrafish *optn* transcripts coding for protein isoforms: Danio_rerio_1 (ENSDART00000014036) and Danio rerio_2 (ENSDART00000133616) with 2,455 and 1,614 bp, and open reading frames encoding sequences of 517 and 314 amino acids, respectively. The transcript Danio_rerio_2 seems to have a different cDNA sequence due to alternative splicing in exon 11 as well as an incomplete 5' coding sequence region (Figure 5.9).

The human *OPTN* gene is located in chromosome 10 and spans 39 kb in length and at least eight transcripts coding for protein isoforms were identified. The gene is organized in three non-coding exons that span the 5'UTR and 13 protein coding exons (Figure 5.9). The most interesting feature is the similar coding exon sizes of these two species, suggesting that these genes have sufficiently important functions to result in their sequence being highly conserved throughout vertebrate evolution.

5.2.4.2 Synteny within chromosomal locations of *OPTN* gene

To determine whether *OPTN* occurs within syntenic regions in mammals and fish, we evaluated the positions of an additional six to ten genes with chromosomal proximity to *OPTN* (Figure 5.10). Analysis of synteny within the chromosomal regions that harbor the *OPTN* gene demonstrated that genes flanking the *OPTN* orthologues are conserved in their positions in the mammalian and fish species analyzed. Not only the fish species revealed a highly conserved block but also species such as chicken and frog showed a block of *OPTN* neighbor genes conserved when compared with the human ortholog (Figure 5.10).

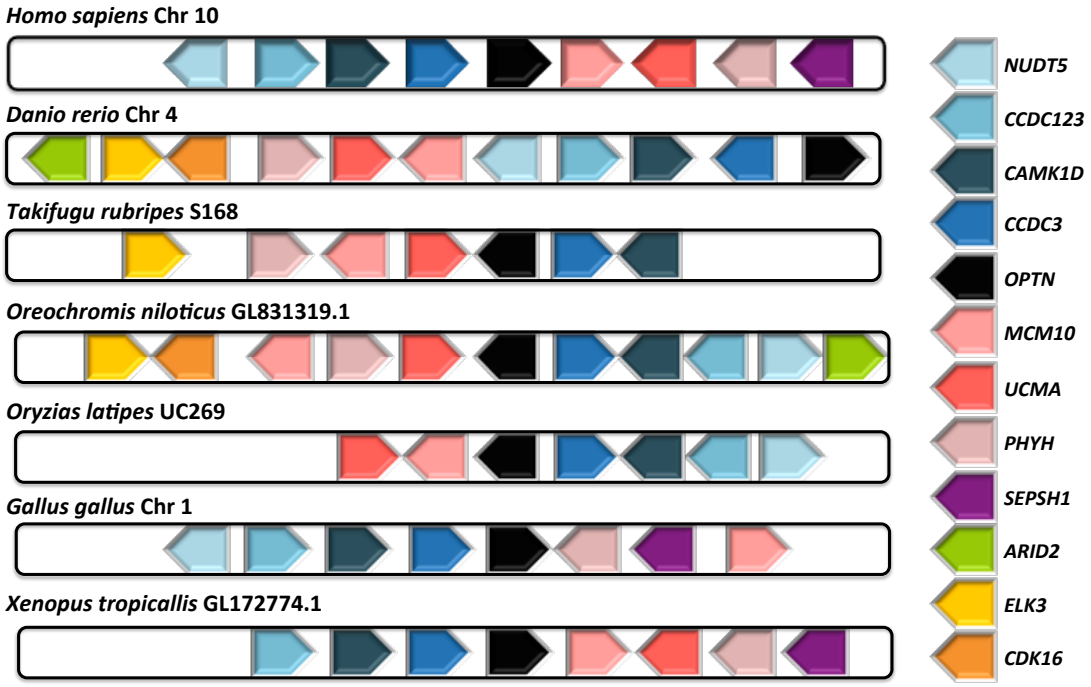


Figure 5.10 - Syntenic relationships between *OPTN* genes. *OPTN* has the same neighbor genes in several species, despite their taxa. The direction of the arrows indicates the direction of the respective gene in the genome. Chr represents chromosome. In *Takifugu rubripes* the S means Scaffold and in

Oryzias latipes the UC means UltraContig, meaning that in these species the genes were not assigned to a specific chromosome.

The conserved synteny drawn using Synteny Database (Catchen et al., 2009), which is specifically designed to identify conserved synteny using Ensembl data, showed that a large region of human chromosome 10 surrounding *OPTN* shares a well-recognizable synteny with the *optn* region on chromosome 4 of zebrafish, providing another solid evidence for a conserved synteny (Table S5.2). A gene-by-gene comparison showed 16 pairs of orthologous genes surrounding the *OPTN* orthologues, confirming our results. In summary, the syntenic relationships of the zebrafish *optn* gene are conserved in the other fish species analyzed, and the same syntenic conservation is maintained with respect to human *OPTN*, showing that not only *OPTN* but that the entire chromosomal context was conserved throughout evolution.

5.2.4.3 *OPTN* is regulated by bone related transcription factors both in human and in zebrafish

Because transcription factor (TFs) can be "switch" genes regulating the expression of other genes, we selected 14 bone related TFs known to regulate several genes and examined the frequency of their binding sites in the human *OPTN* promoter and compared it to its frequency in the corresponding promoter of zebrafish *optn* gene. The results indicate that both human and zebrafish *OPTN* appear to be putative targets of TFs like ETS1, NF-κB, TWIST, GATA, ERE, RXR, KRUEPPEL, LIM, CART1, MEF2 and NFAT, all of them previously associated to skeletal metabolism, and thus could explain the *OPTN* role in bone biology (Figure 5.11). Interestingly, the transcription factor binding elements CART1, GATA, E47, NFAT and MEF2 identified in the human promoter are found to be conserved in the same position in the zebrafish promoter. In addition, the NF-κB binding site is located just upstream of the transcription start site, which may indicate the existence of similar mechanisms for regulation of expression by this known *OPTN* regulator.

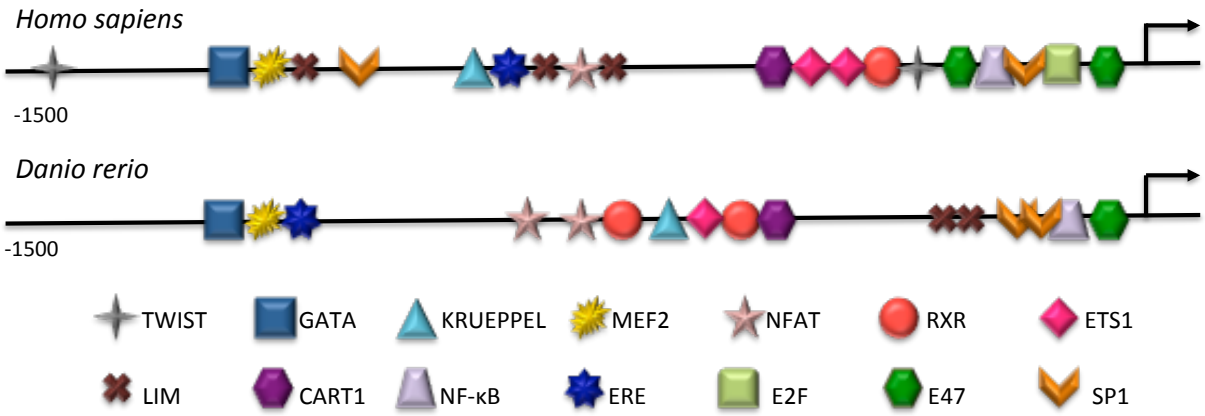


Figure 5.11 - Comparison of 1.5 kb promoter region between human and zebrafish. The analysis of 1.5 kb of *OPTN* regulatory region in these two species shows the conservation of several TFBSs. The arrow represents the transcription start site. The numbers correspond to the length in base pairs of the fragment analyzed.

Also, we analyzed the putative TFBSs studied in a previous work (Silva et al., unpublished results) and that have an effect in human *OPTN* promoter activity. In this analysis the results show that zebrafish *optn* seems to have an E47 binding site in the same region as in the human gene and two SP1 binding sites around the same region as in human *OPTN* promoter. On the contrary, we did not find any E2F family binding sites in the zebrafish *optn* promoter, which suggests that this TF might not be involved in *optn* regulation in zebrafish (Figure 5.11).

5.2.4.4 *OPTN* is expressed in similar tissues in human and in zebrafish

The expression profiles of the *OPTN* genes in human and zebrafish tissues are shown in Figure 5.12. *OPTN* gene expression seems to have a similar pattern in these two species, being expressed in the neural tissues, such as brain, both in human and zebrafish. Also, *OPTN* is known to be highly expressed in the human eye and our results showed that *optn* expression was also found in the zebrafish eye. However, human *OPTN* seems to be highly expressed in the reproductive system while in zebrafish its expression was barely detected. *Optn* expression in the zebrafish skin seems to be relevant and higher than the expression in the eye for example, while in human that is not observed. However the most relevant result is the fact that one of the tissues with the highest expression is bone, in both species (Figure 5.12), indicating that this gene can have an important role in bone development or bone biology, which was conserved in evolution.

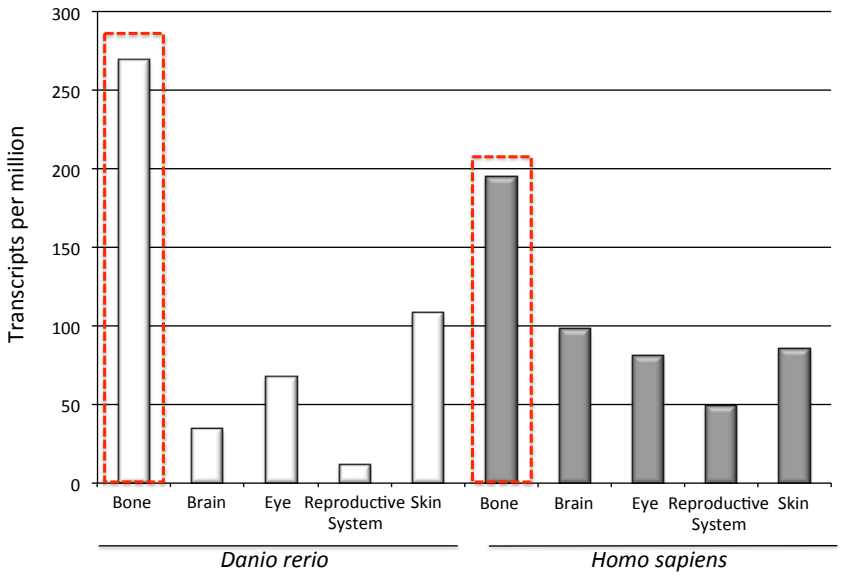


Figure 5.12 - *OPTN* expression profile in human and zebrafish. Analysis of ESTs in some tissues available in the Unigene database (http://www.ncbi.nlm.nih.gov/unigene) reveal that *OPTN* is expressed in the same tissues in zebrafish and in human, particularly it is highly expressed in bone, both in human and in zebrafish, suggesting a major role in this tissue.

5.2.4.5 OPTN primary and tertiary structures are conserved between human and zebrafish

Orthologues to the human *OPTN* gene were identified in the teleost fish genomes of zebrafish, stickleback, medaka, torafugu, platyfish, salmon and tilapia. Pairwise protein alignments and their percentage identity values revealed that the percentage of conservation between mammalians and fish is around 35-40% (Figure 5.13). Protein size was found to be very similar, with medaka's being the smallest protein and mouse's the largest protein. Overall the protein size for OPTN is conserved between human and the teleost fish analyzed (Figure 5.13).

	Dre	Gac	Ola	Oni	Ssa	Tru	Xma	Hsa	Mmu	Gga	Xtr
Dre	524										
Gac	54	558									
Ola	58	70	489								
Oni	57	71	75	521							
Ssa	58	63	61	63	525						
Tru	54	72	71	74	60	535					
Xma	57	68	75	74	61	69	525				
Hsa	37	40	38	39	38	38	37	577			
Mmu	37	39	37	37	35	37	36	77	584		
Gga	38	40	39	39	38	38	38	57	53	556	
Xtr	33	32	32	33	33	32	32	42	41	42	531

Figure 5.13 - Conservation of OPTN proteins. Pairwise alignments and identity percentage reveal that OPTN is well conserved throughout evolution. Dre – *Danio rerio*, Gac – *Gasterosteus aculeatus*,

Ola – *Oryzias latipes*, Oni – *Oreochromis niloticus*, Ssa – *Salmo Salar*, Tru – *Takifugu rubripes*, Xma – *Xiphophorus maculatus*, Hsa – *Homo sapiens*, Mmu – *Mus musculus*, Gga – *Gallus gallus*, Xtr – *Xenopus tropicallis*.

However, more important than an overall protein structure is the conservation of the residues that constitute the main OPTN functional domains. NEMO domain, which confers the same cellular function as the NEMO/IKKγ subunit, is 67% similar between zebrafish and human; the LIR domain, which serves as an interacting domains with LC3 autophagy protein, is 100% similar; UBD, the domain that allows the binding of ubiquitin to OPTN and therefore is important for its function in binding to proteins that were signalized to degradation, is 62% similar; and finally the ZnF domain is 62% similar between these two species (Figure 5.14).

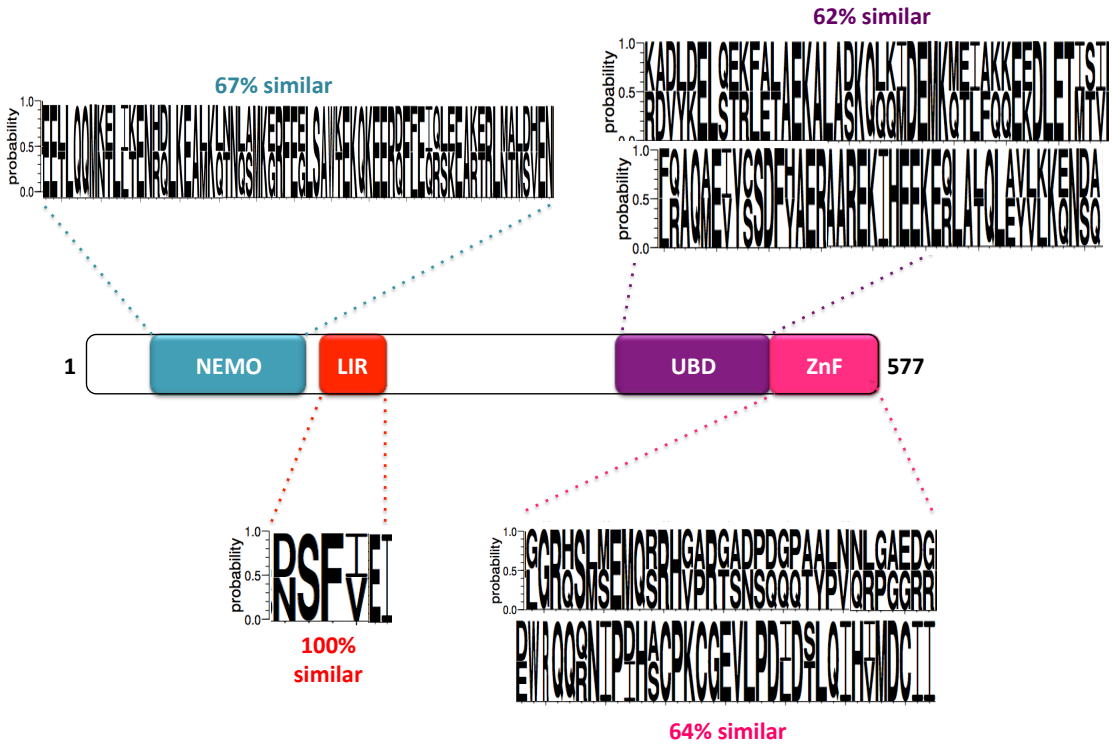


Figure 5.14 - Conservation of OPTN protein domains. Alignment of human and zebrafish OPTN protein sequences shows that the functional domains of OPTN are highly conserved, which presumes a maintenance both in protein function as in expression patterns. UBD represents the ubiquitin binding domain, the LIR represents the LC3 interacting region, and ZnF represents the zinc finger domain.

Also, not only the primary structure of the most important domains is well conserved but also the predicted three-dimensional (3D) structure. To analyze the predicted tertiary structure of OPTN human and zebrafish proteins, we modeled the 3D structure of the human OPTN protein (Figure 5.15A) and zebrafish Optn (Figure 5.15B), since there was no published data about any OPTN 3D structure in the Protein Database. Predicted 3D structures of human and zebrafish OPTN were overlaid to

highlight its conservation (Figure 5.15C), showing just some differences in the N-terminal and C- terminal residues. The high conservation of OPTN domains from primary to tertiary structure and both in zebrafish and mammals strongly suggests that OPTN function has been maintained throughout evolution, thus validating zebrafish as a suitable model for OPTN research.

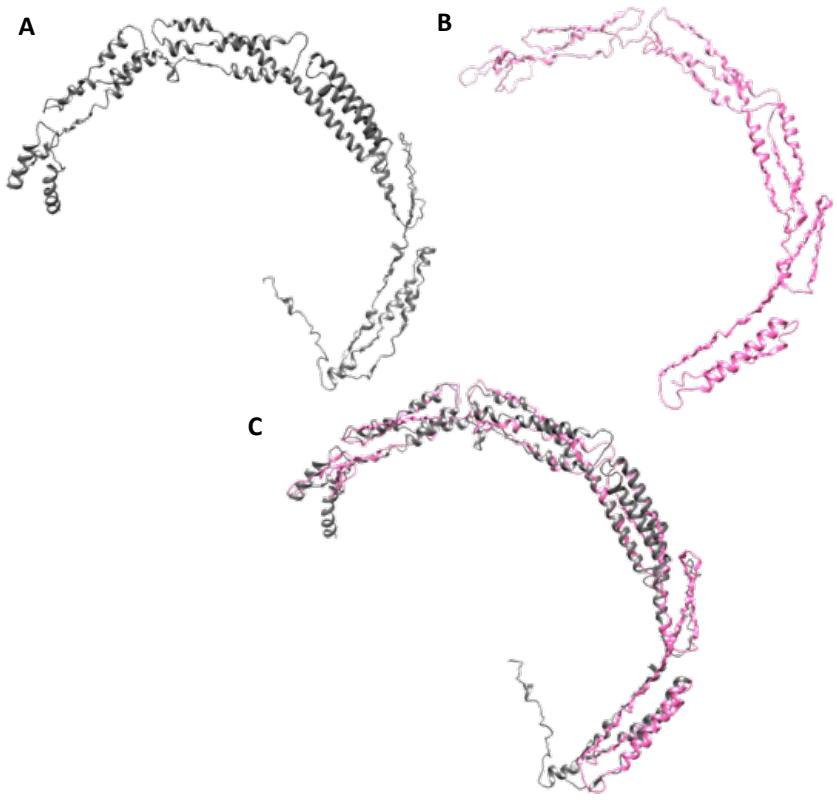


Figure 5.15 - Conservation of the predicted 3D structure of OPTN proteins. OPTN 3D structure is also conserved between human and zebrafish. A) The 3D structures of human (grey) and B) zebrafish (pink) were obtained using I-TASSER software. C) Overlap of the two OPTN proteins highlight the conservation of their structure, and it shows that there are only some differences in the N- and C-terminal regions.

5.2.5 Discussion

The objective of this work was to provide preliminary data towards the use of zebrafish as an alternative model to study *optn* gene expression and Optn role in bone and also complement data available for human and mouse systems. Our strategy included *in silico* analysis to determine (i) the molecular characterization of *optn*, by analyzing its gene structure, (ii) the study of *optn* promoter, by predicting the binding of several transcription factors, (iii) the analysis of *optn* tissue expression by data

available in the Unigene database in NCBI, and (iv) the characterization of protein structural conservation between zebrafish and human.

By sharing with mammals a number of important characteristics (*e.g.* organ systems, gene functions, developmental organization and physiological/biochemical mechanisms, including mechanisms of cartilage and bone formation), fish have been considered a suitable model organism to study mechanisms underlying vertebrate development, including skeletogenesis (Berghmans et al., 2005; Kabashi et al., 2011). Also, the zebrafish is widely accepted as a valuable model for studies of vertebrate development (Driever et al., 1994; Kimmel, 1989) and is being used as a model organism for several human diseases (Zon, 1999), presenting unique advantages over other vertebrate models. In this work, OPTN was shown to be conserved between human and zebrafish, both at gene and protein levels. Zebrafish *optn* was shown to have a gene structure similar to human, to be putatively regulated by the same bone related TFs and to be expressed in many of the same tissues, namely in bone, which suggest that *optn* might have an important role in zebrafish bone.

To identify the location of the *optn* gene in the zebrafish genome we searched the Ensembl database (http://www.ensembl.org/) and found only a single copy of this gene, located in chromosome 4, providing evidence that *optn* evolved to single copy in the zebrafish lineage after the teleost genome duplication (Amores et al., 1998; Postlethwait et al., 2000; Taylor, 2003). This result is in agreement with information collected from databases for other teleosts. Indeed, only one gene was also identified in medaka (*Oryzias latipes*), Southern platyfish (*Xiphophorus maculates*), torafugu (*Takifugu rubripes*), three-spined stickleback (*Gasterosteus aculeatus*) and Atlantic salmon (*Salmo salar*). The zebrafish *optn* gene spans approximately 11 kb of chromosomal DNA and it is organized in 13 coding exons, a structure comparable to those described for its mammalian orthologues. The zebrafish *optn* gene was mapped to chromosome 4 in a region showing extensive conserved synteny with the small arm of human chromosome 10, where *OPTN* is located together with coiled-coil domain containing 3 (CCDC3), upper zone of growth plate and cartilage matrix associated/gla-rich protein (UCMA/GRP), minichromosome maintenance complex component 10 (MCM10) and phytanoyl-CoA 2-hydroxylase (PHYH). Taken together, our results confirmed that the zebrafish sequence reported here is a true ortholog of human *OPTN*. By comparing the chromosomal locations of zebrafish genes and their mammalian orthologues (including human), Postlethwait et al. (1998) discovered

highly conserved syntenic relationships (Postlethwait et al., 1998). Indeed, a growing number of studies have shown extensive conserved synteny between zebrafish and human chromosomes (Amores et al., 1998; Barbazuk et al., 2000; Cancela et al., 2012; Catchen et al., 2009; Fazenda et al., 2012; Gates et al., 1999). The analysis of these syntenic regions demonstrated that the genes flanking *OPTN* are conserved in their positions in all teleost fish genomes analyzed, as well as in human. Unsurprisingly, we also observed a more conserved synteny among the teleost fish analyzed than between teleost and human. Therefore, it is of interest to determine the chromosomal locations of *optn* in other fish species to be able to reconstruct the ancestral gene organization and fully understand their evolution in the fish lineage. In addition, comparative mapping of *optn* in various teleost genomes will contribute to improve our understanding of their evolutionary and functional implications among vertebrates.

Comparative analysis of genomic sequences has become an important tool to identify regulatory regions in genomic DNA and their conservation in non-coding regions has been related to their importance for regulating gene expression (Koop and Hood, 1994). Our *in silico* analysis showed that the transcription factor NF-κB has a putative binding site in *OPTN* promoter, both in zebrafish and in human. It is known that NF-κB regulates human *OPTN* (Sudhakar et al., 2009) and these results show that probably NF-κB is also an important regulator of *optn* in zebrafish. In addition, our latest results showed that E47 and SP1 have an effect in human *OPTN* promoter activity (Silva et al., unpublished results) and our *in silico* analysis show that these TFs might regulate also zebrafish *optn*. These results indicate that human and zebrafish *OPTN* might share the same regulatory mechanisms. The analysis of tissue specific expression available in Unigene database shows that *OPTN* is expressed in many of the same tissues in both species, reinforcing the theory of a common regulatory network.

The more important results are regarding the protein sequence, where we showed an important amino acid sequences similarity in the main domains of the OPTN protein, with 67% similarity in NEMO-like domain, 62% similarity in ubiquitin-associated domain (UBD), a domain known to be involved in OPTN autophagy (Ying and Yue, 2012) and 100% similarity in the LIR domain, which is vital for OPTN interaction with LC3 and OPTN participation in the autophagy process (Wild et al., 2011b). The identification of genetic signatures and domains in coding regions can provide

relevant data for studying the functional and evolutionary characteristics of aligned regions (Nielsen, 2005). The highly conserved nature of the OPTN domains – from fish to human – strongly suggests that OPTN functions have been maintained throughout vertebrate evolution.

Our study provides novel insights about the similarities between zebrafish and human OPTN and contributes to the establishment of zebrafish as a suitable model to study the role of OPTN in bone metabolism, osteoclastogenesis, gene transcription and to investigate functionally OPTN-related human pathologies. Moreover, zebrafish have life-cycle related traits that make them ideal study subjects such as, ease of reproduction, high fecundity and growth rate with juveniles exhibit external development, plus they can be easily manipulated genetically through the use of transgenic overexpression constructs and specific transcription inhibition. Because of these advantages, zebrafish appears to be an ideal model for studying the consequences of functional inactivation or overexpression of OPTN and thus contributing to better understand the genetic events occurring in humans that lead to pathologies such as Paget's disease of bone.

SUPPLEMENTARY TABLES

Table S5.1 – Known mRNA and protein features and corresponding tissue expression of alternatively spliced isoforms encoding human and zebrafish SQSTM1.

Gene	Isoform	Exons	mRNA (bp)	Protein (aa)	Tissue
Hsa *SQSTM1*	a	8	2986	440	skin (39), lung (34), melanotic melanoma (32) and 55 other tissues
	b	7	1714	378	placenta cot 25-normalized (1), thyroid gland (1)
	c	9	2913	356	pancreas (3), placenta (3), pancreatic islet (3) and 20 other tissues
	d	7	2253	356	skin (23), lung (20), eye (15), melanoma (15) and 84 other tissues
	e	8	2000	356	amygdala(3), hippocampus (2), hypothalamus(2) and 14 other tissues
	f	3	602	190	liver (1)
	g	6	842	167	NA
	h	4	628	140	lung (2), leiomyosarcoma (1), leukocyte (1) and 4 other tissues
	i	4	572	73	brain
	j	4	578	69	endometrium, adenocarcinoma cell line from uterus
	k	3	2405	NPP	NA
	l	2	1722	NPP	NA
	m	3	921	NPP	NA
	n	4	808	NPP	NA
	o	2	656	NPP	NA
	p	4	562	NPP	NA
	q	4	532	NPP	NA
Dre *sqstm1*	a	8	1939	452	NA
	b	9	1827	353	NA
	c	5	557	NPP	NA
	d	3	1010	NPP	NA
	e	3	531	NPP	NA

Hsa – Homo sapiens; Dre – Danio rerio; bp – base pairs; aa – amino acid. NA - Information not available; NPP - No protein product.

Table S5.2 – Conserved syntenic block around *OPTN* genes in zebrafish (*Danio rerio*) and human (*Homo sapiens*) performed by Synteny Database

Danio rerio		*Homo sapiens*
ENSDARG00000017549 *(phyh)* Dre: 4 position: 7,008k		ENSG00000107537 *(PHYH)* Hsa: 10 position: 13,320k
ENSDARG00000027799 *(ucma a)* Dre: 4 position: 7,022k		ENSG00000165623 *(UCMA)* Hsa: 10 position: 13,264k
ENSDARG00000045815 *(mcm10)* Dre: 4 position: 7,031k		ENSG00000065328 *(MCM10)* Hsa: 10 position: 13,204k
ENSDARG00000078073 *(nudt5)* Dre: 4 position: 7,051k		ENSG00000165609 *(NUDT5)* Hsa: 10 position: 12,207k
ENSDARG00000075025 *(cdc123)* Dre: 4 position: 7,056k		ENSG00000151465 *(CDC123)* Hsa: 10 position: 12,238k
ENSDARG00000074905 *(camk1d a)* Dre: 4 position: 7,068k		ENSG00000183049 *(CAMK1D)* Hsa: 10 position: 12,391k
ENSDARG00000026052 *(ccdc3)* Dre: 4 position: 7,182k		ENSG00000151468 *(CCDC3)* Hsa: 10 position: 12,939k
ENSDARG00000002663 ***(optn)*** Dre: 4 position: 7,197k		ENSG00000123240 ***(OPTN)*** Hsa: 10 position: 13,141k
ENSDARG00000075501 *(cdnf)* Dre: 4 position: 8,650k		ENSG00000185267 *(CDNF)* Hsa: 10 position: 14,861k
ENSDARG00000058030 *(hspa14)* Dre: 4 position: 8,655k		ENSG00000187522 *(HSPA14)* Hsa: 10 position: 14,880k
ENSDARG00000045704 *(dclre1c)*		ENSG00000152457 *(DCLRE1C)*

Dre: 4 position: 8,666k		Hsa: 10 position: 14,939k
ENSDARG00000045705 *(meig1)* Dre: 4 position: 8,672k		ENSG00000197889 *(MEIG1)* Hsa: 10 position: 15,001k
ENSDARG00000062983 *(fbxo18)* Dre: 4 position: 8,916k		ENSG00000134452 *(FBXO18)* Hsa: 10 position: 5,932k
ENSDARG00000003822 *(ankrd16)* Dre: 4 position: 10,523k		ENSG00000134461 *(ANKRD16)* Hsa: 10 position: 5,904k
ENSDARG00000005451 *(gdi2)* Dre: 4 position: 10,529k		ENSG00000057608 *(GDI2)* Hsa: 10 position: 5,807k
ENSDARG00000079698 *(asb13 a)* Dre: 4 position: 10,544k		ENSG00000196372 *(ASB13)* Hsa: 10 position: 5,681k
ENSDARG00000032765 *(net1)* Dre: 4 position: 10,638k		ENSG00000173848 *(NET1)* Hsa: 10 position: 5,455k

Dre – Danio rerio
Hsa – Homo sapiens
ENSDRAG – assession number of zebrafish genes
ENSG – assession number of human genes
Position – Gene's position in genome

CHAPTER 6

FINAL CONCLUSIONS AND FUTURE PERSPECTIVES

6. FINAL CONCLUSIONS AND FUTURE PERSPECTIVES

Paget's disease of bone (PDB) is the most common metabolic bone disorder after osteoporosis, affecting more than 3% of the Caucasian population. Besides *SQSTM1* gene, no other genes were associated to PDB. However after the GWAS published in 2010 several other candidate genes were suggested to be associated to this disease. One of the most relevant results was the strong association between rs1561570 in PDB6 locus, a SNP located in the intron 7 of *OPTN* gene, and PDB with a *p*-value of 6×10^{-13}. This strong association was replicated in the French-Canadian population but no functional studies had been performed to explain this association.

This work presents new and exciting insights into the role of PDB6 locus and *OPTN* gene in the pathogenesis of PDB and in bone metabolism. In this study, we sought to identify novel genetic variants in the PDB6 locus that could be risk factors for PDB, in patients without *SQSTM1* mutations, using a candidate gene approach to explain the contribution of the PDB6 locus to PDB. After screening the entire PDB6 locus, we selected the best candidate genes to play a role in bone metabolism: *OPTN, UCMA/GRP, CCDC3, SEPSH1, PHYH* and *CAMK1D*. By using a discovery group consisting in 30 familial cases and 5 healthy controls, we identified 82 genetic variants in these genes. Due to their frequency, we selected rs2234968 and rs3829923 to be tested for genetic association in a sample of 225 PDB patients without mutations in *SQSTM1* and 298 healthy controls. This sample has more than 90% power to detect significant association ($p<0.05$) when risk allele frequency differs from at least 0.09 between cases and controls. Rs2234968 was shown to be associated with PDB (*p*-value = 6×10^{-3}) but rs3829923 was not associated.

Functional studies were performed in order to explain rs2234968 association to PDB. Rs2234968 was shown to be in *linkage disequilibrium* with rs10906303 and rs79529484 and we demonstrated that these variants induced a new acceptor site, changing *OPTN* splicing pattern. We predict that this change in *OPTN* splicing alters the open reading frame and creates a premature stop codon, giving rise to a truncated (probably not functional) protein. However there are still some questions that remain to be answered. It would be interesting to prove that *OPTN* splicing mechanisms are

effectively altered due to the presence of rs2234968 *A* allele, by using cDNA from PDB patients with rs2234968 *A* allele and cloning the full *OPTN* open reading frame, to analyze the new cDNA sequence and confirm the existence of a premature stop codon. Also, in the western blot performed in this work, we used an antibody detecting the entire OPTN protein and the results showed the presence of the expected 67kDa band correspondent to the full-length OPTN protein, however with lower levels. Whether the antibody used did not react with the possible short-size band corresponding to the protein produced by the effect of rs2234968 *A* allele or this mRNA isoform was unstable and rapidly degraded is unknown. In a future work, an antibody that recognizes the N- terminal of OPTN protein should be used to assess if there is indeed production of a shorter isoform by revealing the existence of a short-size band in the western blot. Also, the mRNA half-life could be assessed by doing a mRNA decay assay to confirm the existence of a shorter non functional *OPTN* transcript rapidly degraded by the cell, thus explaining the lower OPTN protein levels that we detected in patients harboring rs2234968 *A* allele. In addition, further studies will be required to clarify the effect of these transcripts in PDB patients. By quantifying their expression and comparing it with the expression of the normal transcripts, and analyzing if OPTN function is impaired when only the mutant and "short" transcript is expressed, should be highly relevant to further understand its effect in associated pathologies, such as PDB. Analyzing the autophagy process, by quantifying the LC3B expression in the presence of autophagy drugs such as Bafilomycin A1 or Rapamycin, or the expression of osteoclast related genes in PDB patients with rs2234968 *A* allele, might also contribute to clarify the effect of these variants in PDB pathophysiology.

We also analyzed rs3829923 and RV -9906 effect in *OPTN* promoter activity and the results showed that these SNPs are responsible for an increase in *OPTN* promoter basal activity, probably due to both the activating effect of E47 and E2F1 only in the presence of the rs3829923 *A* allele, and the inhibitory effect of SP1 in RV -9906 *G* allele. The presence of two DNA fragments that only differ in the alleles corresponding to each variant provides the first evidence that rs3829923 *C* allele is needed for E47 and E2F1 binding to *OPTN* promoter and that RV -9906 *A* allele is sufficient to disrupt the SP1 effect. However an electrophoretic mobility shift assay (EMSA) should be performed in order to confirm that, in fact, E47 and E2F1 are binding in the rs3829923 region and that SP1 is binding in the RV -9906 region. A

chromatin immunoprecipitation assay (CHIP) could also be performed to provide additional *in vivo* data, to confirm if these transcription factors are able to bind to *OPTN* promoter in the osteoclasts derived from the PBMCs from PDB patients. In addition, to clarify the role of this variant in osteoclastogenesis, it would be of interest to do an *in vitro* site directed mutagenesis and develop osteoclast precursor isogenic cell lines with and without the variants, aiming at replicating the expression of the two identified variants and analyzing its effect in the number and activity of resulting osteoclasts. To assess the effect of these variants in a more physiological context we could perform co-cultures of osteoclasts derived from PBMCs from PDB patients harboring this variant together with stromal cells or pre-osteoblasts. This would permit to investigate if the stromal cells or pre-osteoblasts from PDB patients harboring these variants produce in fact more RANKL, giving rise to an increase in osteoclastogenesis, or if the osteoclasts of these PDB patients are more sensitive to the same concentration of RANKL.

The most relevant result is, without a doubt, the explanation on how does rs1561570, the most associated SNP to PDB, contributes to this disease, to osteoclastogenesis and to bone metabolism. Using different techniques, such as *in silico* analysis, bisulfite treatment, osteoclast differentiation, demethylating treatment, gene expression analysis by qPCR, and protein analysis by western blot and immunofluorescence, we were able to show that rs1561570 *T* allele is responsible for a loss of a methylation site, an increase in *OPTN* expression, an increase of NF-κB translocation to the nucleus resulting in an increase in expression of its target genes, and thus contributing to promote the increase in number, size and activity of osteoclasts, all possibly contributing to PDB phenotype. Despite some limitations, such as a low number of patients and controls, which may have resulted in lack of statistical power in some of the gene and protein expression analyses performed, the results obtained appear to be quite consistent and have provided evidence towards confirming our hypothesis. Indeed, our study gave several indications, from gene to protein analysis, from *in vitro* to patient samples, from genetic to morphological cellular features, that show that OPTN is important for NF-κB translocation into the nucleus, activation of its target genes and consequent increase in the osteoclastogenesis, providing the first input on how rs1561570 can contribute to the PDB phenotype and/or its severity by acting as a modifier. In fact, together with our results on autophagy impairment in PDB patients, we can hypothesize that in healthy individuals, the autophagy process works normally

both in the presence of the *C* allele, when the *OPTN* expression is controlled at basal levels (Figure 6.1), and in the presence of the *T* allele, when the *OPTN* expression is potentiated due to the loss of a methylation site but the efficient autophagy process signals the excessive OPTN for degradation by ubiquitination, and *OPTN* expression is controlled to a basal level. In both situations this will give rise to normal levels of NF-κB in the nucleus and regular osteoclastogenesis rate (Figure 6.1). In PDB patients, where the autophagy process is defective, with no alteration in the rs1561570 position the PDB phenotype might be explained by the presence of other genetic variants that lead to an increase of osteoclastogenesis (Figure 6.1). The effect of rs1561570 *T* allele in PDB patients is explained by the increase in *OPTN* expression and decreased degradation due to the autophagy defect present in these patients. Therefore, there is an accumulation of OPTN that leads to NF-κB translocation to the nucleus at a much higher rate than normal, giving rise to more osteoclasts, and thus explaining the PDB phenotype in these patients (Figure 6.1).

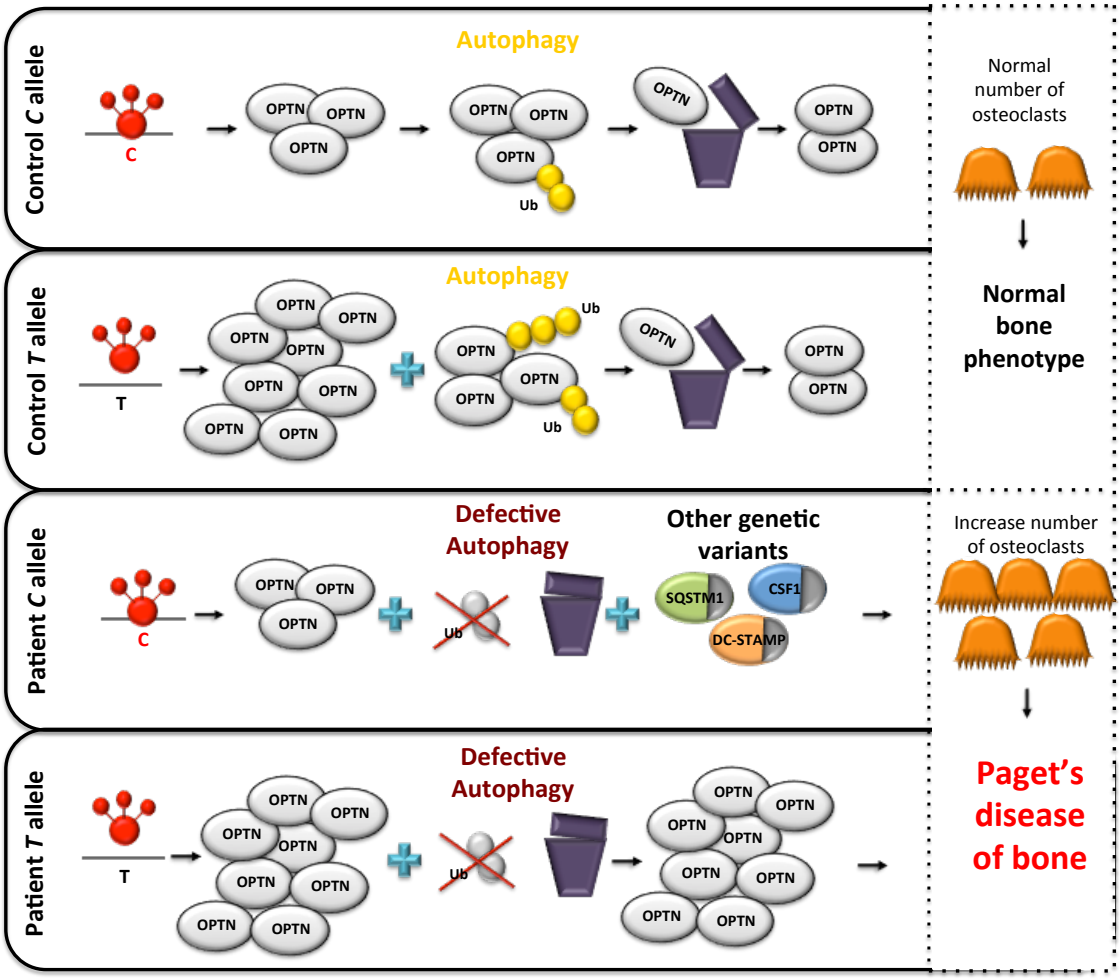


Figure 6.1 - Hypothetical mechanism of rs1561570 effect in PDB. In controls the autophagy process can control the *OPTN* levels potentiated by the presence of the *T* allele in the rs1561570 position. Basal

levels of OPTN give rise to basal levels of NF-κB in the nucleus, basal levels of osteoclastogenesis and the maintenance of the normal bone metabolism. In PDB patients with the *C* allele other genetic variations might explain the PDB phenotype, however in the presence of the *T* allele, there is an accumulation of OPTN due to a defective autophagy process that is unable to control *OPTN* levels. The increase in OPTN will increase NF-κB translocation to the nucleus, NF-κB target gene expression and an increase in osteoclastogenesis, explaining the PDB phenotype. Ub represents the ubiquitin chains.

These results also suggest that the overexpression of *OPTN* could be a good pretherapeutic target in order to prevent the major PDB symptoms. By inhibiting the increase of *OPTN* expression using, for example, a microRNA targeting *OPTN*, the defective autophagy in patients would not represent a problem, and the phenotype would not be as critical. In fact, in our *in silico* analysis of the OPTN variants we detected putative binding sites for several microRNAs related to bone, such as miR-320a, miR-302d-3p, miR-3128, miR-145 and miR-365a (Baglìo et al., 2013; Franceschetti et al., 2014; He et al., 2015; Jia et al., 2013; Peng et al., 2011), which indicates that *OPTN* might be a target for these microRNAs, which could be used to inhibit *OPTN* overexpression caused by genetic variants, such as rs1565170. Also, to complete our model, the cause of the autophagy defect in PDB patients should be analyzed in detail, since this is probably the best therapeutic target to prevent PDB phenotype.

In addition, these results also elucidate the role of *OPTN* overexpression as a risk factor for other diseases, such as glaucoma. Glaucoma is characterized by an increase in the intraocular pressure that give rise to degeneration of the optic nerve, retinal ganglion cell death, and progressive axonal and visual field loss (Kerrigan et al., 1997; Kuehn et al., 2005; Quigley, 2011). The role of *OPTN* as an osteoclastogenesis enhancer can explain the increase of calcium in the blood stream and in the eye. In excess, calcium acts as an oxidant and therefore *OPTN* overexpression will cause an increase in the oxidative stress in the eye and damage the trabecular meshwork. This will lead to an increase of the intraocular pressure due to aqueous humor retention, causing the main symptoms in glaucoma. Also, calcium deposits may obstruct the draining of the aqueous humor by the Schlemm's canal in the eye. Taking this into account, future pre-therapeutical inhibition of *OPTN* expression may be valuable not only for PDB patients but also for glaucoma patients, for example.

In need of an alternative animal model, we questioned if zebrafish could be a good model to study PDB biological mechanisms or potential therapeutics, since zebrafish has been described as an excellent model to study bone related pathologies and can

thus be considered a promising tool for the initial development and validation of new therapeutic approaches (Barrett et al., 2006; Berghmans et al., 2005; Huttner et al., 2013; Jagadeeswaran and Liu, 1997; Kabashi et al., 2011; Sadler et al., 2005; Swanhart et al., 2011). Keeping that in mind, we analyzed the evolutionary conservation between zebrafish and human of the only gene associated with PDB - *SQSTM1* - and of *OPTN*, the other candidate gene. Our results provide evidence supporting the use of zebrafish as a valid model to study both SQSTM1 P392L mutation and OPTN function in bone, since this gene is highly expressed in zebrafish bone. To understand the mechanistic effects of *OPTN* and *SQSTM1* mutations, it would be interesting to generate zebrafish transgenic lines using the CRISPR-Cas9 system, by microinjection of zebrafish zygotes with constructs containing a mutation in *OPTN* exon 5 that led to a premature stop codon, to mimic rs2234968, or containing the mutation P308L in the zebrafish *sqstm1* gene that is equivalent to the P392L observed in the human *SQSTM1*. A combination of the two mutations in these genes to assess the combined effect of *OPTN* and *SQSTM1* mutations could also be performed. In addition, it would be interesting to use the Tol2 system with a TRAP promoter to overexpress *OPTN* in pre-osteoclasts, to reproduce the effect of the most associated SNP in OPTN, rs1561570. A combination of the effect of rs1561570 and Sqstm1 P308L should be also studied.

Although much remains to be done, novel and valuable data have been collected within the scope of this work, which contributed towards a better understanding of OPTN function and regulation but also towards the evolutionary relationship between *SQSTM1* and *OPTN* and the feasibility of zebrafish as a valid animal model for studying PDB. Ultimately this work contributed to further clarify the contribution of PDB6 locus and OPTN in PDB.

REFERENCES

Agou, F., Courtois, G., Chiaravalli, J., Baleux, F., Coïc, Y.M., Traincard, F., Israël, A., Véron, M., 2004. Inhibition of NF-kappa B activation by peptides targeting NF-kappa B essential modulator (nemo) oligomerization. J. Biol. Chem. 279, 54248–54257.

Akagi, T., Ono, H., Shimotohno, K., 1995. Characterization of T cells immortalized by Tax1 of human T-cell leukemia virus type 1. Blood 86, 4243–4249.

Albagha, O.M.E., Visconti, M.R., Alonso, N., Langston, A.L., Cundy, T., Dargie, R., Dunlop, M.G., Fraser, W.D., Hooper, M.J., Isaia, G., Nicholson, G.C., del Pino Montes, J., Gonzalez-Sarmiento, R., di Stefano, M., Tenesa, A., Walsh, J.P., Ralston, S.H., 2010. Genome-wide association study identifies variants at CSF1, OPTN and TNFRSF11A as genetic risk factors for Paget's disease of bone. Nat. Genet. 42, 520–524.

Albagha, O.M.E., Wani, S.E., Visconti, M.R., Alonso, N., Goodman, K., Brandi, M.L., Cundy, T., Chung, P.Y.J., Dargie, R., Devogelaer, J.P., Falchetti, A., Fraser, W.D., Gennari, L., Gianfrancesco, F., Hooper, M.J., Van Hul, W., Isaia, G., Nicholson, G.C., Nuti, R., Papapoulos, S., Montes, J.D.P., Ratajczak, T., Rea, S.L., Rendina, D., Gonzalez-Sarmiento, R., Di Stefano, M., Ward, L.C., Walsh, J.P., Ralston, S.H., 2011. Genome-wide association identifies three new susceptibility loci for Paget's disease of bone. Nat. Genet. 43, 685–689.

Altman, R.D., 2002. Epidemiology of Paget's disease of bone. Clin Rev Bone Min. Metab 1, 99–102.

Altman, R.D., Bloch, D.A., Hochberg, M.C., Murphy, W.A., 2000. Prevalence of pelvic Paget's disease of bone in the United States. J. Bone Miner. Res. 15, 461–465.

Amores, A., Force, A., Yan, Y.L., Joly, L., Amemiya, C., Fritz, A., Ho, R.K., Langeland, J., Prince, V., Wang, Y.L., Westerfield, M., Ekker, M., Postlethwait, J.H., 1998. Zebrafish hox clusters and vertebrate genome evolution. Science 282, 1711–1714.

Ang, E.S.M., Zhang, P., Steer, J.H., Tan, J.W., Yip, K., Zheng, M.H., Joyce, D.A., Xu, J., 2007. Calcium / Calmodulin-Dependent Kinase Activity is Required for Efficient Induction of Osteoclast Differentiation and Bone Resorption by Receptor Activator of Nuclear Factor Kappa B Ligand (RANKL) 787–795.

Ankrom, M.A., Shapiro, J.R., 1998. Paget's disease of bone (osteitis deformans). J. Am. Geriatr. Soc. 46, 1025–1033.

Apschner, A., Schulte-Merker, S., Witten, P.E., 2011. The Zebrafish: Disease Models and Chemical Screens, Methods in cell biology, Methods in Cell Biology. Elsevier.

Baglìo, S.R., Devescovi, V., Granchi, D., Baldini, N., 2013. MicroRNA expression profiling of human bone marrow mesenchymal stem cells during osteogenic differentiation reveals Osterix regulation by miR-31. Gene 527, 321–331.

Barbazuk, W.B., Korf, I., Kadavi, C., Heyen, J., Tate, S., Wun, E., Bedell, J.A., McPherson, J.D., Johnson, S.L., 2000. The syntenic relationship of the zebrafish and human genomes. Genome Res. 10, 1351–1358.

Barrett, R., Chappell, C., Quick, M., Fleming, A., 2006. A rapid, high content,in vivo model of glucocorticoid-induced osteoporosis. Biotechnol. J. 1, 651–655.

Baslé, M.F., Fournier, J.G., Rozenblatt, S., Rebel, A., Bouteille, M., 1986. Measles virus RNA detected in Paget's disease bone tissue by in situ hybridization. J. Gen. Virol. 67 (Pt 5), 907–913.

Bellipanni, G., Murakami, T., Doerre, O.G., Andermann, P., Weinberg, E.S., 2000. Expression of Otx Homeodomain Proteins Induces Cell Aggregation in Developing Zebrafish Embryos. Dev. Biol. 223, 339–353.

Belzil, V. V, Daoud, H., Desjarlais, A., Bouchard, J.P., Dupré, N., Camu, W., Dion, P.A., Rouleau, G.A., 2011. Analysis of OPTN as a causative gene for amyotrophic lateral sclerosis. Neurobiol. Aging 32, 555.e13–4.

Beranger, G.E., Momier, D., Guigonis, J.M., Samson, M., Carle, G.F., Scimeca, J.C., 2007. Differential binding of poly(ADP-Ribose) polymerase-1 and JunD/Fra2 accounts for RANKL-induced Tcirg1 gene expression during osteoclastogenesis. J. Bone Miner. Res. 22, 975–983.

Bergamaschi, A., Kim, Y.H., Kwei, K.A., Choi, Y.L.A, Langerod, A., Han, W., Noh, D., Huntsman, D.G., Jeffrey, S.S., Pollack, J.R., 2008. CAMK1D amplification implicated in epithelial-mesenchymal transition in basal-like breast cancer. Mol. Oncol. 2, 327–339.

Berghmans, S., Jette, C., Langenau, D., Hsu, K., Stewart, R., Look, T., Kanki, J.P., 2005. Making waves in cancer research: new models in the zebrafish. Biotechniques 39, 227–237.

Berman, S.D., Yuan, T.L., Miller, E.S., Lee, E.Y., Caron, A., Lees, J.A., 2008. The retinoblastoma protein tumor suppressor is important for appropriate osteoblast differentiation and bone development. Mol. Cancer Res. 6, 1440–1451.

Berridge, M.J., 2012. Calcium signalling remodelling and disease. Biochem. Soc. Trans. 40, 297–309.

Beyens, G., Wuyts, W., Cleiren, E., de Freitas, F., Tiegs, R., Van Hul, W., 2006. Identification and molecular characterization of a novel splice-site mutation (G1205C) in the SQSTM1 gene causing Paget's disease of bone in an extended American family. Calcif. Tissue Int. 79, 281–288.

Bianchi, K., Meier, P., 2009. A tangled web of ubiquitin chains: breaking news in TNF-R1 signaling. Mol. Cell 36, 736–742.

Birch, M.A., Taylor, W., Fraser, W.D., Ralston, S.H., Hart, C.A., Gallagher, J.A., 2009. Absence of paramyxovirus RNA in cultures of pagetic bone cells and in pagetic bon. J. Bone Miner. Res. 9, 11–16.

Bjørkøy, G., Lamark, T., Brech, A., Outzen, H., Perander, M., Overvatn, A., Stenmark, H., Johansen, T., 2005. p62/SQSTM1 forms protein aggregates degraded by autophagy and has a protective effect on huntingtin-induced cell death. J. Cell Biol. 171, 603–614.

Bjørkøy, G., Lamark, T., Pankiv, S., Øvervatn, A., Brech, A., Johansen, T., 2009. Monitoring autophagic degradation of p62/SQSTM1. Methods Enzymol. 452, 181–197.

Bolis, C.L., Piccolella, M., Dalla Valle, A.Z., Rankin, J.C., 2001. Fish as model in pharmacological and biological research. Pharmacol. Res. 44, 265–280.

Bolland, M.J., Tong, P.C., Naot, D., Callon, K.E., Wattie, D.J., Gamble, G.D., Cundy, T., 2007. Delayed development of Paget's disease in offspring inheriting SQSTM1 mutations. J. Bone Miner. Res. 22, 411–415.

Brenet, F., Moh, M., Funk, P., Feierstein, E., Viale, A.J., Socci, N.D., Scandura, J.M., 2011. DNA methylation of the first exon is tightly linked to transcriptional silencing. PLoS One 6.

Broos, S., Hulpiau, P., Galle, J., Hooghe, B., Van Roy, F., De Bleser, P., 2011. ConTra v2: a tool to identify transcription factor binding sites across species, update 2011. Nucleic Acids Res. 39, 74–78.

Brown, J., Michou, 2011. Emerging strategies and therapies for treatment of Paget’s disease of bone. Drug Des. Devel. Ther. 225.

Cancela, M.L., Conceição, N., Laizé, V., 2012. Gla-Rich Protein, a New Player in Tissue Calcification? Adv. Nutr. 3, 174–181.

Cartharius, K., Frech, K., Grote, K., Klocke, B., Haltmeier, M., Klingenhoff, A., Frisch, M., Bayerlein, M., Werner, T., 2005. MatInspector and beyond: promoter analysis based on transcription factor binding sites. Bioinformatics 21, 2933–2942.

Catchen, J.M., Conery, J.S., Postlethwait, J.H., 2009. Automated identification of conserved synteny after whole-genome duplication. Genome Res. 19, 1497–1505.

Cavey, J.R., Ralston, S.H., Sheppard, P.W., Ciani, B., Gallagher, T.R.A., Long, J.E., Searle, M.S., Layfield, R., 2006. Loss of ubiquitin binding is a unifying mechanism by which mutations of SQSTM1 cause Paget's disease of bone. Calcif. Tissue Int. 78, 271–277.

Chakkalakal, S.A., Zhang, D., Culbert, A.L., Convente, M.R., Caron, R.J., Wright, A.C., Maidment, A.D.A., Kaplan, F.S., Shore, E.M., 2012. An Acvr1 R206H knock-in mouse has fibrodysplasia ossificans progressiva. J. Bone Miner. Res. 27, 1746–1756.

Chalasani, M.L., Radha, V., Gupta, V., Agarwal, N., Balasubramanian, D., Swarup, G., 2007. A glaucoma-associated mutant of optineurin selectively induces death of retinal ganglion cells which is inhibited by antioxidants. Invest. Ophthalmol. Vis. Sci. 48, 1607–1614.

Chalasani, M.L., Swarup, G., Balasubramanian, D., 2009. Optineurin and its mutants: molecules associated with some forms of glaucoma. Ophthalmic Res. 42, 176–184.

Chamoux, E., Couture, J., Bisson, M., Morissette, J., Brown, J.P., Roux, S., 2009. The p62 P392L mutation linked to Paget's disease induces activation of human osteoclasts. Mol. Endocrinol. 23, 1668–1680.

Chatani, M., Takano, Y., Kudo, A., 2011. Osteoclasts in bone modeling, as revealed by in vivo imaging, are essential for organogenesis in fish. Dev. Biol. 360, 96–109.

Chi, Z.L., Akahori, M., Obazawa, M., Minami, M., Noda, T., Nakaya, N., Tomarev, S., Kawase, K., Yamamoto, T., Noda, S., Sasaoka, M., Shimazaki, A., Takada, Y., Iwata, T., 2010. Overexpression of optineurin E50K disrupts Rab8 interaction and leads to a progressive retinal degeneration in mice. Hum. Mol. Genet. 19, 2606–2615.

Chong, B., Hegde, M., Fawkner, M., Simonet, S., Cassinelli, H., Coker, M., Kanis, J., Seidel, J., Tau, C., Tüysüz, B., Yüksel, B., Love, D., 2003. Idiopathic hyperphosphatasia and TNFRSF11B mutations: relationships between phenotype and genotype. J. Bone Miner. Res. 18, 2095–2104.

Chu, Z.L., Shin, Y.A., Yang, J.M., DiDonato, J.A., Ballard, D.W., 1999. IKKgamma mediates the interaction of cellular IkappaB kinases with the tax transforming protein of human T cell leukemia virus type 1. J. Biol. Chem. 274, 15297–15300.

Chung, P.Y.J., Beyens, G., de Freitas, F., Boonen, S., Geusens, P., Vanhoenacker, F., Verbruggen, L., Van Offel, J., Goemaere, S., Zmierczak, H.G., Westhovens, R., Devogelaer, J.P., Van Hul, W., 2011. Indications for a genetic association of a VCP polymorphism with the pathogenesis of sporadic Paget's disease of bone, but not for TNFSF11 (RANKL) and IL-6 polymorphisms. Mol. Genet. Metab. 103, 287–292.

Chung, P.Y.J., Beyens, G., Riches, P.L., Van Wesenbeeck, L., de Freitas, F., Jennes, K., Daroszewska, A., Fransen, E., Boonen, S., Geusens, P., Vanhoenacker, F., Verbruggen, L., Van Offel, J., Goemaere, S., Zmierczak, H.G., Westhovens, R., Karperien, M., Papapoulos, S., Ralston, S.H., Devogelaer, J.P., Van Hul, W., 2010. Genetic variation in the TNFRSF11A gene encoding RANK is associated with susceptibility to Paget's disease of bone. J. Bone Miner. Res. 25, 2592–2605.

Chung, P.Y.J., Van Hul, W., 2012. Paget's disease of bone: evidence for complex pathogenetic interactions. Semin. Arthritis Rheum. 41, 619–641.

Cody, J.D., Singer, F.R., Roodman, G.D., Otterund, B., Lewis, T.B., Leppert, M., Leach, R.J., 1997. Genetic Linkage of Paget Disease of the Bone to Chromosome 18q. Am. J. Hum. Genet. 61, 1117–1122.

Colina, M., La Corte, R., De Leonardis, F., Trotta, F., 2008. Paget's disease of bone: a review. Rheumatol. Int. 28, 1069–1075.

Collet, C., Michou, L., Audran, M., Chasseigneaux, S., Hilliquin, P., Bardin, T., Lemaire, I., Cornélis, F., Launay, J., Orcel, P., Laplanche, J., 2007. Paget's Disease of Bone in the French Population: Novel 22, 310–317.

Cooper, C., Harvey, N.C., Dennison, E.M., van Staa, T.P., 2006. Update on the epidemiology of Paget's disease of bone. J. Bone Miner. Res. 21 Suppl 2, P3–8.

Corral-Gudino, L., Borao-Cengotita-Bengoa, M., Del Pino-Montes, J., Ralston, S., 2013. Epidemiology of Paget's disease of bone: A systematic review and meta-analysis of secular changes. Bone 55, 347–352.

Coxon, F.P., Rogers, M.J., 2003. The role of prenylated small GTP-binding proteins in the regulation of osteoclast function. Calcif. Tissue Int. 72, 80–84.

Crooks, G.E., Hon, G., Chandonia, J.M., Brenner, S.E., 2004. WebLogo: a sequence logo generator. Genome Res. 14, 1188–1190.

Cundy, T., Bolland, M., 2008. Paget disease of bone. Trends Endocrinol. Metab. 19, 246–253.

Cundy, T., Hegde, M., Naot, D., Chong, B., King, A., Wallace, R., Mulley, J., Love, D.R., Seidel, J., Fawkner, M., Banovic, T., Callon, K.E., Grey, A.B., Reid, I.R., Middleton-Hardie, C.A., Cornish, J., 2002. A mutation in the gene TNFRSF11B encoding osteoprotegerin causes an idiopathic hyperphosphatasia phenotype. Hum. Mol. Genet. 11, 2119–2127.

Czell, D., Andersen, P.M., Neuwirth, C., Morita, M., Weber, M., 2013. Progressive aphasia as the presenting symptom in a patient with amyotrophic lateral sclerosis with a novel mutation in the OPTN gene. Amyotroph. Lateral Scler. Frontotemporal Degener. 14, 138–140.

Daroszewska, A., Hocking, L.J., McGuigan, F.E., Langdahl, B., Stone, M.D., Cundy, T., Nicholson, G.C., Fraser, W.D., Ralston, S.H., 2004. Susceptibility to Paget's Disease of Bone Is Influenced by a Common Polymorphic Variant of Osteoprotegerin. J. Bone Miner. Res. 19, 1506–1511.

Daroszewska, A., van 't Hof, R.J., Rojas, J.A., Layfield, R., Landao-Basonga, E., Rose, L., Rose, K., Ralston, S.H., 2011. A point mutation in the ubiquitin-associated domain of SQSMT1 is sufficient to cause a Paget's disease-like disorder in mice. Hum. Mol. Genet. 20, 2734–2744.

David, J.P., Rincon, M., Neff, L., Horne, W.C., Baron, R., 2001. Carbonic anhydrase II is an AP-1 target gene in osteoclasts. J. Cell. Physiol. 188, 89–97.

De Valck, D., Jin, D.Y., Heyninck, K., Van de Craen, M., Contreras, R., Fiers, W., Jeang, K.T., Beyaert, R., 1999. The zinc finger protein A20 interacts with a novel anti-apoptotic protein which is cleaved by specific caspases. Oncogene 18, 4182–4190.

De Vrieze, E., Sharif, F., Metz, J.R., Flik, G., Richardson, M.K., 2011. Matrix metalloproteinases in osteoclasts of ontogenetic and regenerating zebrafish scales. Bone 48, 704–712.

Dejardin, E., 2006. The alternative NF-kappaB pathway from biochemistry to biology: pitfalls and promises for future drug development. Biochem. Pharmacol. 72, 1161–1179.

Del Toro, D., Alberch, J., Lázaro-Diéguez, F., Martín-Ibáñez, R., Xifró, X., Egea, G., Canals, J.M., 2009. Mutant huntingtin impairs post-Golgi trafficking to lysosomes by delocalizing optineurin/Rab8 complex from the Golgi apparatus. Mol. Biol. Cell 20, 1478–1492.

Doyle, T., Gunn, J., Anderson, G., Gill, M., Cundy, T., 2002. Paget's disease in New Zealand: evidence for declining prevalence. Bone 31, 616–619.
Driever, W., Stemple, D., Schier, A., Solnica-Krezel, L., 1994. Zebrafish: genetic tools for studying vertebrate development. Trends Genet. 10, 152–159.
Du, Y., Wooten, M.C., Wooten, M.W., 2009. Oxidative damage to the promoter region of SQSTM1/p62 is common to neurodegenerative disease. Neurobiol. Dis. 35, 302–310.
Durán, A., Serrano, M., Leitges, M., Flores, J.M., Picard, S., Brown, J.P., Moscat, J., Diaz-Meco, M.T., 2004. The atypical PKC-interacting protein p62 is an important mediator of RANK-activated osteoclastogenesis. Dev. Cell 6, 303–309.
ENCODE Project Consortium: The ENCODE (ENCyclopedia Of DNA Elements), 2004. Science 306, 636–640.
Fazenda, C., Silva, I. a. L., Cancela, M.L., Conceição, N., 2012. Molecular characterization of two paralog genes encoding Gla-rich protein (Grp) in zebrafish. J. Appl. Ichthyol. 28, 377–381.
Fecto, F., Yan, J., Vemula, S.P., Liu, E., Yang, Y., Chen, W., Zheng, J.G., Shi, Y., Siddique, N., Arrat, H., Donkervoort, S., Ajroud-Driss, S., Sufit, R.L., Heller, S.L., Deng, H.X., Siddique, T., 2011. SQSTM1 mutations in familial and sporadic amyotrophic lateral sclerosis. Arch. Neurol. 68, 1440–1446.
Fitzgerald, K.A., McWhirter, S.M., Faia, K.L., Rowe, D.C., Latz, E., Golenbock, D.T., Coyle, A.J., Liao, S.M., Maniatis, T., 2003. IKKepsilon and TBK1 are essential components of the IRF3 signaling pathway. Nat. Immunol. 4, 491–496.
Floudas, C.S., Um, N., Kamboh, M.I., Barmada, M.M., Visweswaran, S., 2014. Identifying genetic interactions associated with late-onset Alzheimer's disease. BioData Min. 7.
Fotino, M., Haymovits, A., Falk, C.T., 1977. Evidence for linkage between HLA and Paget's disease. Transplant. Proc. 9, 1867–1868.
Franceschetti, T., Dole, N.S., Kessler, C.B., Lee, S.K., Delany, A.M., 2014. Pathway analysis of microRNA expression profile during murine osteoclastogenesis. PLoS One 9, e107262.
Funayama, T., Ishikawa, K., Ohtake, Y., Tanino, T., Kurosaka, D., Kimura, I., Suzuki, K., Ideta, H., Nakamoto, K., Yasuda, N., Fujimaki, T., Murakami, A., Asaoka, R., Hotta, Y., Tanihara, H., Kanamoto, T., Mishima, H., Fukuchi, T., Abe, H., Iwata, T., Shimada, N., Kudoh, J., Shimizu, N., Mashima, Y., 2004. Variants in optineurin gene and their association with tumor necrosis factor-alpha polymorphisms in Japanese patients with glaucoma. Invest. Ophthalmol. Vis. Sci. 45, 4359–4367.
Gao, J., Ohtsubo, M., Hotta, Y., Minoshima, S., 2014. Oligomerization of optineurin and its oxidative stress- or E50K mutation-driven covalent cross-linking: possible relationship with glaucoma pathology. PLoS One 9, e101206.
Gates, M.A., Kim, L., Egan, E.S., Cardozo, T., Sirotkin, H.I., Dougan, S.T., Lashkari, D., Abagyan, R., Schier, A.F., Talbot, W.S., 1999. A genetic linkage map for zebrafish: comparative analysis and localization of genes and expressed sequences. Genome Res. 9, 334–347.
Gianfrancesco, F., Rendina, D., Di Stefano, M., Mingione, A., Esposito, T., Merlotti, D., Gallone, S., Magliocca, S., Goode, A., Formicola, D., Morello, G., Layfield, R., Frattini, A., De Filippo, G., Nuti, R., Searle, M., Strazzullo, P., Isaia, G., Mossetti, G., Gennari, L., 2012. A nonsynonymous TNFRSF11A variation increases NFκB activity and the severity of Paget's disease. J. Bone Miner. Res. 27, 443–452.
Gleason, C.E., Ordureau, A., Gourlay, R., Arthur, J.S.C., Cohen, P., 2011. Polyubiquitin binding to optineurin is required for optimal activation of TANK-binding kinase 1 and production of interferon β. J. Biol. Chem. 286, 35663–35674.
Good, D., Busfield, F., Fletcher, B.H., Duffy, D.L., Kesting, J.B., Andersen, J., Shaw, J.T.E., 2002. Linkage of Paget disease of bone to a novel region on human chromosome 18q23. Am. J. Hum. Genet. 70, 517–525.
Goode, A., Layfield, R., 2010. Recent advances in understanding the molecular basis of Paget disease of bone. J. Clin. Pathol. 63, 199–203.
Grandi, P., Dang, T., Pané, N., Shevchenko, A., Mann, M., Forbes, D., Hurt, E., 1997. Nup93, a vertebrate homologue of yeast Nic96p, forms a complex with a novel 205-kDa protein and is required for correct nuclear pore assembly. Mol. Biol. Cell 8, 2017–2038.
Grant, S.F., Reid, D.M., Blake, G., Herd, R., Fogelman, I., Ralston, S.H., 1996. Reduced bone density and osteoporosis associated with a polymorphic Sp1 binding site in the collagen type I alpha 1 gene. Nat. Genet. 14, 203–205.
Greenway, M.J., Andersen, P.M., Russ, C., Ennis, S., Cashman, S., Donaghy, C., Patterson, V., Swingler, R., Kieran, D., Prehn, J., Morrison, K.E., Green, A., Acharya, K.R., Brown, R.H.,

Hardiman, O., 2006. ANG mutations segregate with familial and "sporadic" amyotrophic lateral sclerosis. Nat. Genet. 38, 411–413.
Grimer, R.J., Cannon, S.R., Taminiau, A.M., Bielack, S., Kempf-Bielack, B., Windhager, R., Dominkus, M., Saeter, G., Bauer, H., Meller, I., Szendroi, M., Folleras, G., San-Julian, M., van der Eijken, J., 2003. Osteosarcoma over the age of forty. Eur. J. Cancer 39, 157–163.
Guay-Bélanger, S., Cormier, J.G., Michou, L., 2015. La maladie osseuse de Paget, une condition évanescente ? Rev. Rhum. 82, 223–229.
Hajkova, P., Knotek, Z., 1998. Osteitis Deformans in Black Rat Snake. Eur. Assoc. Zoo- Wildl. Vet. 2, 469–472.
Hall, B.K., Witten, P.E., 2007. Plasticity of and transitions between skeletal tissues in vertebrate evolution and development., in: Anderson, J.S., Sues, H.D. (Eds.), Major Transitions in Vertebrate Evolution. Indiana University Press, Bloomington, pp. 13–56.
Harhaj, E.W., Sun, S.C., 1999. IKKgamma serves as a docking subunit of the IkappaB kinase (IKK) and mediates interaction of IKK with the human T-cell leukemia virus Tax protein. J. Biol. Chem. 274, 22911–22914.
Haslam, S.I., Van Hul, W., Morales-Piga, A., Balemans, W., San-Millan, J.L., Nakatsuka, K., Willems, P., Haites, N.E., Ralston, S.H., 1998. Paget's disease of bone: evidence for a susceptibility locus on chromosome 18q and for genetic heterogeneity. J. Bone Miner. Res. 13, 911–917.
Hattula, K., Peränen, J., 2000. FIP-2, a coiled-coil protein, links Huntingtin to Rab8 and modulates, cellular morphogenesis. Curr. Biol. 10, 1603–1606.
He, C., Gao, H., Fan, X., Wang, M., Liu, W., Huang, W., Yang, Y., 2015. Identification of a novel miRNA-target gene regulatory network in osteosarcoma by integrating transcriptome analysis. Int. J. Clin. Exp. Pathol. 8, 8348–8357.
Helfrich, M.H., Hobson, R.P., Grabowski, P.S., Zurbriggen, A., Cosby, S.L., Dickson, G.R., Fraser, W.D., Ooi, C.G., Selby, P.L., Crisp, A.J., Wallace, R.G.H., Kahn, S., Ralston, S.H., 2000. A Negative Search for a Paramyxoviral Etiology of Paget's Disease of Bone: Molecular, Immunological, and Ultrastructural Studies in U.K. Patients. J. Bone Miner. Res. 15, 2315–2329.
Helfrich, M.H., Hocking, L.J., 2008. Genetics and aetiology of Pagetic disorders of bone. Arch. Biochem. Biophys. 473, 172–182.
Hennies, H.C., Kornak, U., Zhang, H., Egerer, J., Zhang, X., Seifert, W., Kühnisch, J., Budde, B., Nätebus, M., Brancati, F., Wilcox, W.R., Müller, D., Kaplan, P.B., Rajab, A., Zampino, G., Fodale, V., Dallapiccola, B., Newman, W., Metcalfe, K., Clayton-Smith, J., Tassabehji, M., Steinmann, B., Barr, F.A., Nürnberg, P., Wieacker, P., Mundlos, S., 2008. Gerodermia osteodysplastica is caused by mutations in SCYL1BP1, a Rab-6 interacting golgin. Nat. Genet. 40, 1410–1412.
Hiruma, Y., Kurihara, N., Subler, M.A., Zhou, H., Boykin, C.S., Zhang, H., Ishizuka, S., Dempster, D.W., Roodman, G.D., Windle, J.J., 2008. A SQSTM1/p62 mutation linked to Paget's disease increases the osteoclastogenic potential of the bone microenvironment. Hum. Mol. Genet. 17, 3708–3719.
Hocking, L.J., Slee, F., Haslam, S., Cundy, T., Nicholson, G., van Hul, W., Ralston, S., 2000. Familial Paget's disease of bone: patterns of inheritance and frequency of linkage to chromosome 18q. Bone 26, 577–580.
Hocking, L.J., Herbert, C.A., Nicholls, R.K., Williams, F., Bennett, S.T., Cundy, T., Nicholson, G.C., Wuyts, W., Van Hul, W., Ralston, S.H., 2001. Genomewide search in familial Paget disease of bone shows evidence of genetic heterogeneity with candidate loci on chromosomes 2q36, 10p13, and 5q35. Am. J. Hum. Genet. 69, 1055–1061.
Hocking, L.J., Lucas, G.J.A., Daroszewska, A., Mangion, J., Olavesen, M., Cundy, T., Nicholson, G.C., Ward, L., Bennett, S.T., Wuyts, W., Van Hul, W., Ralston, S.H., 2002. Domain-specific mutations in sequestosome 1 (SQSTM1) cause familial and sporadic Paget's disease. Hum. Mol. Genet. 11, 2735–2739.
Hoffmann, A., Baltimore, D., 2006. Circuitry of nuclear factor kappaB signaling. Immunol. Rev. 210, 171–186.
Hosking, D.J., 1981. Paget's disease of bone. Br. Med. J. (Clin. Res. Ed). 283, 686–688.
Hrašovec, S., Hauptman, N., Glavač, D., Jelenc, F., Ravnik-Glavač, M., 2013. TMEM25 is a candidate biomarker methylated and down-regulated in colorectal cancer. Dis. Markers 34, 93–104.
Hughes, A.E., Ralston, S.H., Marken, J., Bell, C., MacPherson, H., Wallace, R.G., van Hul, W., Whyte, M.P., Nakatsuka, K., Hovy, L., Anderson, D.M., 2000. Mutations in TNFRSF11A, affecting the signal peptide of RANK, cause familial expansile osteolysis. Nat. Genet. 24, 45–48.
Hughes, H.C., Lang, C.M., 1971. Osteitis deformans in a macaque (macaca mulatta). Vet. Pathol. 8, 414–420.

Huttner, I.G., Trivedi, G., Jacoby, A., Mann, S.A., Vandenberg, J.I., Fatkin, D., 2013. A transgenic zebrafish model of a human cardiac sodium channel mutation exhibits bradycardia, conduction-system abnormalities and early death. J. Mol. Cell. Cardiol. 61, 123–132.

Iha, H., Peloponese, J.M., Verstrepen, L., Zapart, G., Ikeda, F., Smith, C.D., Starost, M.F., Yedavalli, V., Heyninck, K., Dikic, I., Beyaert, R., Jeang, K.T., 2008. Inflammatory cardiac valvulitis in TAX1BP1-deficient mice through selective NF-kappaB activation. EMBO J. 27, 629–641.

Iida, A., Hosono, N., Sano, M., Kamei, T., Oshima, S., Tokuda, T., Kubo, M., Nakamura, Y., Ikegawa, S., 2012. Optineurin mutations in Japanese amyotrophic lateral sclerosis. J. Neurol. Neurosurg. Psychiatry 83, 233–235.

Inoue, D., Santiago, P., Horne, W.C., Baron, R., 1997. Identification of an osteoclast transcription factor that binds to the human T cell leukemia virus type I-long terminal repeat enhancer element. J. Biol. Chem. 272, 25386–25393.

Itzstein, C., Espinosa, L., Delmas, P.D., Chenu, C., 2000. Specific antagonists of NMDA receptors prevent osteoclast sealing zone formation required for bone resorption. Biochem. Biophys. Res. Commun. 268, 201–209.

Iwamoto, K., Bundo, M., Ueda, J., Oldham, M.C., Ukai, W., Hashimoto, E., Saito, T., Geschwind, D.H., Kato, T., 2011. Neurons show distinctive DNA methylation profile and higher interindividual variations compared with non-neurons. Genome Res. 21, 688–696.

Jagadeeswaran, P., Liu, Y.C., 1997. A Hemophilia Model in Zebrafish: Analysis of Hemostasis. Blood Cells, Mol. Dis. 23, 52–57.

Jahn, R., Südhof, T.C., 1999. Membrane fusion and exocytosis. Annu. Rev. Biochem. 68, 863–911.

Jansen, G., Mihalik, S., Watkins, P., Moser, H., Jakobs, C., Denis, S., Wanders, R.J., 1996. Phytanoyl-CoA hydroxylase is present in human liver, located in peroxisomes, and deficient in Zellweger syndrome: direct, unequivocal evidence for the new, revised pathway of phytanic acid alpha-oxidation in humans. Biochem. Biophys. Res. Commun. 229, 205–210.

Jia, J., Tian, Q., Ling, S., Liu, Y., Yang, S., Shao, Z., 2013. miR-145 suppresses osteogenic differentiation by targeting Sp7. FEBS Lett. 587, 3027–3031.

Jin, D.Y., Giordano, V., Kibler, K. V, Nakano, H., Jeang, K.T., 1999. Role of adapter function in oncoprotein-mediated activation of NF-kappaB. Human T-cell leukemia virus type I Tax interacts directly with IkappaB kinase gamma. J. Biol. Chem. 274, 17402–17405.

Johnson, J.O., Mandrioli, J., Benatar, M., Abramzon, Y., Van Deerlin, V.M., Trojanowski, J.Q., Gibbs, J.R., Brunetti, M., Gronka, S., Wuu, J., Ding, J., McCluskey, L., Martinez-Lage, M., Falcone, D., Hernandez, D.G., Arepalli, S., Chong, S., Schymick, J.C., Rothstein, J., Landi, F., Wang, Y.D., Calvo, A., Mora, G., Sabatelli, M., Monsurrò, M.R., Battistini, S., Salvi, F., Spataro, R., Sola, P., Borghero, G., Galassi, G., Scholz, S.W., Taylor, J.P., Restagno, G., Chiò, A., Traynor, B.J., 2010. Exome sequencing reveals VCP mutations as a cause of familial ALS. Neuron 68, 857–864.

Johnson-Pais, T.L., Singer, F.R., Bone, H.G., McMurray, C.T., Hansen, M.F., Leach, R.J., 2003. Identification of a novel tandem duplication in exon 1 of the TNFRSF11A gene in two unrelated patients with familial expansile osteolysis. J. Bone Miner. Res. 18, 376–380.

Journo, C., Filipe, J., About, F., Chevalier, S.A., Afonso, P. V, Brady, J.N., Flynn, D., Tangy, F., Israël, A., Vidalain, P.O., Mahieux, R., Weil, R., 2009. NRP/Optineurin Cooperates with TAX1BP1 to potentiate the activation of NF-kappaB by human T-lymphotropic virus type 1 tax protein. PLoS Pathog. 5, e1000521.

Ju, J.S., Fuentealba, R.A., Miller, S.E., Jackson, E., Piwnica-Worms, D., Baloh, R.H., Weihl, C.C., 2009. Valosin-containing protein (VCP) is required for autophagy and is disrupted in VCP disease. J. Cell Biol. 187, 875–888.

Kabashi, E., Brustein, E., Champagne, N., Drapeau, P., 2011. Zebrafish models for the functional genomics of neurogenetic disorders. Biochim. Biophys. Acta - Mol. Basis Dis. 1812, 335–345.

Kachaner, D., Génin, P., Laplantine, E., Weil, R., 2012. Toward an integrative view of Optineurin functions. Cell Cycle 11, 2808–2818.

Kajiho, H., Saito, K., Tsujita, K., Kontani, K., Araki, Y., Kurosu, H., Katada, T., 2003. RIN3: a novel Rab5 GEF interacting with amphiphysin II involved in the early endocytic pathway. J. Cell Sci. 116, 4159–4168.

Kan, L., Hu, M., Gomes, W.A., Kessler, J.A., 2004. Transgenic Mice Overexpressing BMP4 Develop a Fibrodysplasia Ossificans Progressiva (FOP)-Like Phenotype. Am. J. Pathol. 165, 1107–1115.

Kanis, J.A., 1998. Pathophysiology and treatment of Paget's disease of bone. London: Martin Duntz, 1998, pp 110-138.

Ke, Y., Yue, H., He, J., Liu, Y., Zhang, Z., 2009. Early onset Paget's disease of bone caused by a novel mutation (78dup27) of the TNFRSF11A gene in a Chinese family. Acta Pharmacol. Sin. 30, 1204–1210.

Kerrigan, L.A., Zack, D.J., Quigley, H.A., Smith, S.D., Pease, M.E., 1997. TUNEL-positive ganglion cells in human primary open-angle glaucoma. Arch. Ophthalmol. (Chicago, Ill. 1960) 115, 1031–1035.

Kimmel, C.B., 1989. Genetics and early development of zebrafish. Trends Genet. 5, 283–288.

Kirkin, V., Lamark, T., Johansen, T., Dikic, I., 2009a. NBR1 cooperates with p62 in selective autophagy of ubiquitinated targets. Autophagy 5, 732–733.

Kirkin, V., McEwan, D.G., Novak, I., Dikic, I., 2009b. A role for ubiquitin in selective autophagy. Mol. Cell 34, 259–269.

Kobayashi, S., Fukuhara, A., Taguchi, T., Matsuda, M., Tochino, Y., Otsuki, M., Shimomura, I., 2010. Identification of a new secretory factor, CCDC3/Favine, in adipocytes and endothelial cells. Biochem. Biophys. Res. Commun. 392, 29–35.

Kobune, M., Iyama, S., Kikuchi, S., Horiguchi, H., Sato, T., Murase, K., Kawano, Y., Takada, K., Ono, K., Kamihara, Y., Hayashi, T., Miyanishi, K., Sato, Y., Takimoto, R., Kato, J., 2012. Stromal cells expressing hedgehog-interacting protein regulate the proliferation of myeloid neoplasms. Blood Cancer J. 2:e87.

Koop, B.F., Hood, L., 1994. Striking sequence similarity over almost 100 kilobases of human and mouse T-cell receptor DNA. Nat. Genet. 7, 48–53.

Korac, J., Schaeffer, V., Kovacevic, I., Clement, A.M., Behl, C., Terzic, J., Dikic, I., 2013. Europe PMC Funders Group Ubiquitin-independent function of optineurin in autophagic clearance of protein aggregates 126, 580–592.

Kovach, M.J., Waggoner, B., Leal, S.M., Gelber, D., Khardori, R., Levenstien, M.A., Shanks, C.A., Gregg, G., Al-Lozi, M.T., Miller, T., Rakowicz, W., Lopate, G., Florence, J., Glosser, G., Simmons, Z., Morris, J.C., Whyte, M.P., Pestronk, A., Kimonis, V.E., 2001. Clinical Delineation and Localization to Chromosome 9p13.3–p12 of a Unique Dominant Disorder in Four Families: Hereditary Inclusion Body Myopathy, Paget Disease of Bone, and Frontotemporal Dementia. Mol. Genet. Metab. 74, 458–475.

Kovalenko, A., Chable-Bessia, C., Cantarella, G., Israël, A., Wallach, D., Courtois, G., 2003. The tumour suppressor CYLD negatively regulates NF-kappaB signalling by deubiquitination. Nature 424, 801–805.

Kraft, C., Peter, M., Hofmann, K., 2010. Selective autophagy: ubiquitin-mediated recognition and beyond. Nat. Cell Biol. 12, 836–841.

Kuehn, M.H., Fingert, J.H., Kwon, Y.H., 2005. Retinal ganglion cell death in glaucoma: mechanisms and neuroprotective strategies. Ophthalmol. Clin. North Am. 18, 383–395.

Kurihara, N., Hiruma, Y., Yamana, K., Michou, L., Rousseau, C., Morissette, J., Galson, D.L., Teramachi, J., Zhou, H., Dempster, D.W., Windle, J.J., Brown, J.P., Roodman, G.D., 2011. Contributions of the Measles Virus Nucleocapsid Gene and the SQSTM1/p62P392L Mutation to Paget's Disease. Cell Metab. 13, 23–34.

Kurihara, N., Reddy, S. V., Menaa, C., Anderson, D., Roodman, G.D., 2000. Osteoclasts expressing the measles virus nucleocapsid gene display a pagetic phenotype. J. Clin. Invest. 105, 607–614.

Kurihara, N., Zhou, H., Reddy, S. V, Palacios, V.G., Subler, M.A., Dempster, D.W., Windle, J.J., Roodman, G.D., 2006. Expression of Measles Virus Nucleocapsid Protein in Osteoclasts Induces Paget's Disease-Like Bone Lesions in Mice. J. Bone Miner. Res. 21, 446–455.

Langston, A.L., Ralston, S.H., 2004. Management of Paget's disease of bone. Rheumatology (Oxford). 43, 955–959.

Laplantine, E., Fontan, E., Chiaravalli, J., Lopez, T., Lakisic, G., Véron, M., Agou, F., Israël, A., 2009. NEMO specifically recognizes K63-linked poly-ubiquitin chains through a new bipartite ubiquitin-binding domain. EMBO J. 28, 2885–2895.

Lattante, S., Conte, A., Zollino, M., Luigetti, M., Del Grande, A., Marangi, G., Romano, A., Marcaccio, A., Meleo, E., Bisogni, G., Rossini, P.M., Sabatelli, M., 2012. Contribution of major amyotrophic lateral sclerosis genes to the etiology of sporadic disease. Neurology 79, 66–72.

Laurin, N., Brown, J.P., Lemainque, A., Duchesne, A., Huot, D., Lacourcière, Y., Drapeau, G., Verreault, J., Raymond, V., Morissette, J., 2001. Paget disease of bone: mapping of two loci at 5q35-qter and 5q31. Am. J. Hum. Genet. 69, 528–543.

Laurin, N., Brown, J.P., Morissette, J., Raymond, V., 2002. Recurrent mutation of the gene encoding sequestosome 1 (SQSTM1/p62) in Paget disease of bone. Am. J. Hum. Genet. 70, 1582–1588.

Layfield, R., 2007. The molecular pathogenesis of Paget disease of bone. Expert Rev. Mol. Med. 9, 1–13.

Layfield, R., Hocking, L.J., 2004. SQSTM1 and Paget's disease of bone. Calcif. Tissue Int. 75, 347–357.
Leach, R.J., Singer, F.R., Roodman, G.D., 2001. The genetics of Paget's disease of the bone. J. Clin. Endocrinol. Metab. 86, 24–8.
Lee, N.K., 2010. Molecular Understanding of Osteoclast Differentiation and Physiology. Endocrinol. Metab. 25, 264.
Leigh, P.N., 2007. Chapter 13 Amyotrophic lateral sclerosis. Handb. Clin. Neurol. 82, 249–278.
Leopold¬-Levi, M., Londe, A., 1897. Application des rayons de Roentgen a l'etude de la texture. N. iconog. la Salpetriere, Paris. 198–201.
Leung, Y., Fan, B., Lam, D., Lee, W., Tam, P., Chua, J., 2003. Different optineurin mutation pattern in primary open-angle glaucoma. Invest. Ophthalmol. Vis. Sci. 44, 3880–3884.
Leung, Y.F., Fan, B.J., Lam, D.S.C., Lee, W.S., Tam, P.O.S., Chua, J.K.H., Tham, C.C.Y., Lai, J.S.M., Fan, D.S.P., Pang, C.P., 2003. Different optineurin mutation pattern in primary open-angle glaucoma. Invest. Ophthalmol. Vis. Sci. 44, 3880–3884.
Lever, J.H., 2002. Paget's disease of bone in Lancashire and arsenic pesticide in cotton mill wastewater: a speculative hypothesis. Bone 31, 434–436.
Lewis, B.P., Green, R.E., Brenner, S.E., 2003. Evidence for the widespread coupling of alternative splicing and nonsense-mediated mRNA decay in humans. Proc. Natl. Acad. Sci. U. S. A. 100, 189–192.
Li, J., Sarosi, I., Yan, X.Q., Morony, S., Capparelli, C., Tan, H.L., McCabe, S., Elliott, R., Scully, S., Van, G., Kaufman, S., Juan, S.C., Sun, Y., Tarpley, J., Martin, L., Christensen, K., McCabe, J., Kostenuik, P., Hsu, H., Fletcher, F., Dunstan, C.R., Lacey, D.L., Boyle, W.J., 2000. RANK is the intrinsic hematopoietic cell surface receptor that controls osteoclastogenesis and regulation of bone mass and calcium metabolism. Proc. Natl. Acad. Sci. U. S. A. 97, 1566–1571.
Li, X., Udagawa, N., Itoh, K., Suda, K., Murase, Y., Nishihara, T., Suda, T., Takahashi, N., 2002. p38 MAPK-mediated signals are required for inducing osteoclast differentiation but not for osteoclast function. Endocrinology 143, 3105–3113.
Li, Y., Kang, J., Horwitz, M.S., 1998. Interaction of an adenovirus E3 14.7-kilodalton protein with a novel tumor necrosis factor alpha-inducible cellular protein containing leucine zipper domains. Mol. Cell. Biol. 18, 1601–1610.
Lieschke, G.J., Currie, P.D., 2007. Animal models of human disease: zebrafish swim into view. Nat. Rev. Genet. 8, 353–367.
Liu, J., Yang, H., Liu, W., Cao, X., Feng, X., 2005. Sp1 and Sp3 regulate the basal transcription of receptor activator of nuclear factor kappa B ligand gene in osteoblasts and bone marrow stromal cells. J. Cell. Biochem. 96, 716–727.
Long, C.L., Berry, W.L., Zhao, Y., Sun, X.H., Humphrey, M.B., 2012. E proteins regulate osteoclast maturation and survival. J. Bone Miner. Res. 27, 2476–2489.
López, C., Thomas, D. V, Davies, A.M., 2003. Neoplastic transformation and tumour-like lesions in Paget's disease of bone: a pictorial review. Eur. Radiol. 13 Suppl 4, 151–163.
Lorenzo, J., Horowitz, M., Choi, Y., Schett, G., Takayanagi, H., 2007. Osteoimmunology - Interactions of the Immune and Skeletal Systems. Springer.
Low, S.C., Harney, J.W., Berry, M.J., 1995. Cloning and functional characterization of human selenophosphate synthetase, an essential component of selenoprotein synthesis. J. Biol. Chem. 270, 21659–21664.
Lucas, G.J.A., Daroszewska, A., Ralston, S.H., Al, L.E.T., 2006a. Contribution of Genetic Factors to the Pathogenesis of Paget's Disease of Bone and Related Disorders 21. Suppl 2, 31-37.
Lucas, G.J.A., Mehta, S.G., Hocking, L.J., Stewart, T.L., Cundy, T., Nicholson, G.C., Walsh, J.P., Fraser, W.D., Watts, G.D.J., Ralston, S.H., Kimonis, V.E., 2006b. Evaluation of the role of Valosin-containing protein in the pathogenesis of familial and sporadic Paget's disease of bone. Bone 38, 280–285.
Lucas, G.J.A., Riches, P.L., Hocking, L.J., Cundy, T., Nicholson, G.C., Walsh, J.P., Ralston, S.H., 2008. 10p13 in Families of British Descent 23, 58–63.
Mackie, E.J., 2003. Osteoblasts: novel roles in orchestration of skeletal architecture. Int. J. Biochem. Cell Biol. 35, 1301–1305.
Madonna, G., Ullman, C.D., Gentilcore, G., Palmieri, G., Ascierto, P.A., 2012. NF-κB as potential target in the treatment of melanoma. J. Transl. Med. 10, 53.
Mankouri, J., Fragkoudis, R., Richards, K.H., Wetherill, L.F., Harris, M., Kohl, A., Elliott, R.M., Macdonald, A., 2010. Optineurin negatively regulates the induction of IFNbeta in response to RNA virus infection. PLoS Pathog. 6, e1000778.
Marks, S.C., Odgren, P.R., 2002. Principles of Bone Biology, Principles of Bone Biology. Elsevier.

Maruyama, H., Morino, H., Ito, H., Izumi, Y., Kato, H., Watanabe, Y., Kinoshita, Y., Kamada, M., Nodera, H., Suzuki, H., Komure, O., Matsuura, S., Kobatake, K., Morimoto, N., Abe, K., Suzuki, N., Aoki, M., Kawata, A., Hirai, T., Kato, T., Ogasawara, K., Hirano, A., Takumi, T., Kusaka, H., Hagiwara, K., Kaji, R., Kawakami, H., 2010. Mutations of optineurin in amyotrophic lateral sclerosis. Nature 465, 223–226.

Matsuo, K., Irie, N., 2008. Osteoclast-osteoblast communication. Arch. Biochem. Biophys. 473, 201–209.

Matthews, B.G., Afzal, M.A., Minor, P.D., Bava, U., Callon, K.E., Pitto, R.P., Cundy, T., Cornish, J., Reid, I.R., Naot, D., 2008. Failure to detect measles virus ribonucleic acid in bone cells from patients with Paget's disease. J. Clin. Endocrinol. Metab. 93, 1398–2401.

Mee, A.P., Dixon, J.A., Hoyland, J.A., Davies, M., Selby, P.L., Mawer, E.B., 1998. Detection of canine distemper virus in 100% of Paget's disease samples by in situ-reverse transcriptase-polymerase chain reaction. Bone 23, 171–175.

Mellman, I., 1996. Endocytosis and molecular sorting. Annu. Rev. Cell Dev. Biol. 12, 575–625.

Michou, L., Conceição, N., Morissette, J., Gagnon, E., Miltenberger-Miltenyi, G., Siris, E.S., Brown, J.P., Cancela, M.L., 2012. Genetic association study of UCMA/GRP and OPTN genes (PDB6 locus) with Paget's disease of bone. Bone 51, 720–728.

Michou, L., Morissette, J., Gagnon, E.R., Marquis, A., Dellabadia, M., Brown, J.P., Siris, E.S., 2011. Novel SQSTM1 mutations in patients with Paget's disease of bone in an unrelated multiethnic American population. Bone 48, 456–460.

Millecamps, S., Boillée, S., Chabrol, E., Camu, W., Cazeneuve, C., Salachas, F., Pradat, P.F., Danel-Brunaud, V., Vandenberghe, N., Corcia, P., Le Forestier, N., Lacomblez, L., Bruneteau, G., Seilhean, D., Brice, A., Feingold, J., Meininger, V., LeGuern, E., 2011. Screening of OPTN in French familial amyotrophic lateral sclerosis. Neurobiol. Aging 32, 557.e11–3.

Mills, B.G., Frausto, A., Singer, F.R., Ohsaki, Y., Demulder, A., Roodman, G.D., 1994. Multinucleated cells formed in vitro from Paget's bone marrow express viral antigens. Bone 15, 443–448.

Mills, B.G., Singer, F.R., Weiner, L.P., Holst, P.A., 1981. Immunohistological demonstration of respiratory syncytial virus antigens in Paget disease of bone. Proc. Natl. Acad. Sci. U. S. A. 78, 1209–1213.

Montagu, M.F.A., 1948. Paget ' s Disease (Osteitis deformans) and Heredity 94–95.

Morales Piga, A.A., Rey Rey, J.S., Corres Gonzales, J., Garcia Sagredo, J.M., Lopez Abente, G., 1995. Frequency and Characteristics of Familial Aggregation of Paget ' s Disease of Bone. J. Bone Miner. Res. 10, 663–670.

Morissette, J., Laurin, N., Brown, J.P., 2006. Sequestosome 1: mutation frequencies, haplotypes, and phenotypes in familial Paget's disease of bone. J. Bone Miner. Res. 21 Suppl 2, 38–44.

Mostowy, S., Boucontet, L., Mazon Moya, M.J., Sirianni, A., Boudinot, P., Hollinshead, M., Cossart, P., Herbomel, P., Levraud, J.P., Colucci-Guyon, E., 2013. The zebrafish as a new model for the in vivo study of Shigella flexneri interaction with phagocytes and bacterial autophagy. PLoS Pathog. 9, e1003588.

Müller, U., Steinhoff, U., Reis, L.F., Hemmi, S., Pavlovic, J., Zinkernagel, R.M., Aguet, M., 1994. Functional role of type I and type II interferons in antiviral defense. Science 264, 1918–1921.

Munitic, I., Giardino Torchia, M.L., Meena, N.P., Zhu, G., Li, C.C., Ashwell, J.D., 2013. Optineurin Insufficiency Impairs IRF3 but Not NF- B Activation in Immune Cells. J. Immunol. 191, 6231–6240.

Nagabhushana, A., Bansal, M., Swarup, G., 2011. Optineurin is required for CYLD-dependent inhibition of TNFα-induced NF-κB activation. PLoS One 6, e17477.

Nakatsuka, K., Nishizawa, Y., Ralston, S.H., 2003. Phenotypic characterization of early onset Paget's disease of bone caused by a 27-bp duplication in the TNFRSF11A gene. J. Bone Miner. Res. 18, 1381–1385.

Neale, S.D., Schulze, E., Smith, R., Athanasou, N.A., 2002. The influence of serum cytokines and growth factors on osteoclast formation in Paget's disease. QJM 95, 233–240.

Nielsen, R., 2005. Molecular signatures of natural selection. Annu. Rev. Genet. 39, 197–218.

Nishimura, A.L., Mitne-Neto, M., Silva, H.C.A., Richieri-Costa, A., Middleton, S., Cascio, D., Kok, F., Oliveira, J.R.M., Gillingwater, T., Webb, J., Skehel, P., Zatz, M., 2004. A mutation in the vesicle-trafficking protein VAPB causes late-onset spinal muscular atrophy and amyotrophic lateral sclerosis. Am. J. Hum. Genet. 75, 822–831.

Notredame, C., Higgins, D.G., Heringa, J., 2000. T-Coffee: A novel method for fast and accurate multiple sequence alignment. J. Mol. Biol. 302, 205–217.

Novack, D.V., 2011. Role of NF-κB in the skeleton. Cell Res. 21, 169–182.

Nuovo, M.A., Nuovo, G.J., MacConnell, P., Forde, A., Steiner, G.C., 1992. In situ analysis of Paget's disease of bone for measles-specific PCR-amplified cDNA. Diagn. Mol. Pathol. 1, 256–265.
O'Driscoll, J.B., Anderson, D.C., 1985. Past pets and Paget's disease. Lancet (London, England) 2, 919–921.
Obaid, R., Wani, S., Ralston, S.H., Albagha, O.M., 2012. OPTN negatively regulates osteoclast formation in vitro. Bone 50, S92–S93.
Ooi, C.G., Walsh, C.A., Gallagher, J.A., Fraser, W.D., 2000. Absence of measles virus and canine distemper virus transcripts in long-term bone marrow cultures from patients with Paget's disease of bone. Bone 27, 417–421.
Osawa, T., Mizuno, Y., Fujita, Y., Takatama, M., Nakazato, Y., Okamoto, K., 2011. Optineurin in neurodegenerative diseases. Neuropathology 31, 569–574.
Osterberg, P.H., Wallace, R.G., Adams, D.A., Crone, R.S., Dickson, G.R., Kanis, J.A., Mollan, R.A., Nevin, N.C., Sloan, J., Toner, P.G., 1988. Familial expansile osteolysis. A new dysplasia. J. Bone Joint Surg. Br. 70, 255–260.
Pachot, A., Blond, J.L., Mougin, B., Miossec, P., 2004. Peptidylpropyl isomerase B (PPIB): a suitable reference gene for mRNA quantification in peripheral whole blood. J. Biotechnol. 114, 121–124.
Paget, J., 1877. On a Form of Chronic Inflammation of Bones (Osteitis Deformans). Med. Chir. Trans. 60, 37–64.
Palenzuela, L., Vives-Bauza, C., Fernández-Cadenas, I., Meseguer, A., Font, N., Sarret, E., Schwartz, S., Andreu, A.L., 2002. Familial expansile osteolysis in a large Spanish kindred resulting from an insertion mutation in the TNFRSF11A gene. J. Med. Genet. 39, E67.
Pankiv, S., Clausen, T.H., Lamark, T., Brech, A., Bruun, J.A., Outzen, H., Øvervatn, A., Bjørkøy, G., Johansen, T., 2007. p62/SQSTM1 binds directly to Atg8/LC3 to facilitate degradation of ubiquitinated protein aggregates by autophagy. J. Biol. Chem. 282, 24131–24145.
Park, B., Ying, H., Shen, X., Park, J.S., Qiu, Y., Shyam, R., Yue, B.Y.J.T., 2010. Impairment of protein trafficking upon overexpression and mutation of optineurin. PLoS One 5, e11547.
Paulus, J.D., Link, B.A., 2014. Loss of Optineurin In Vivo Results in Elevated Cell Death and Alters Axonal Trafficking Dynamics. PLoS One 9, e109922.
Peng, X., Guo, W., Liu, T., Wang, X., Tu, X., Xiong, D., Chen, S., Lai, Y., Du, H., Chen, G., Liu, G., Tang, Y., Huang, S., Zou, X., 2011. Identification of miRs-143 and -145 that Is Associated with Bone Metastasis of Prostate Cancer and Involved in the Regulation of EMT. PLoS One 6, e20341.
Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., Ferrin, T.E., 2004. UCSF Chimera-a visualization system for exploratory research and analysis. J. Comput. Chem. 25, 1605–1612.
Phan, T.C., Xu, J., Zheng, M.H., 2004. Interaction between osteoblast and osteoclast: Impact in bone disease. Histol. Histopathol. 19, 1325–1344.
Postlethwait, J.H., Woods, I.G., Ngo-Hazelett, P., Yan, Y.L., Kelly, P.D., Chu, F., Huang, H., Hill-Force, A., Talbot, W.S., 2000. Zebrafish comparative genomics and the origins of vertebrate chromosomes. Genome Res. 10, 1890–1902.
Postlethwait, J.H., Yan, Y.L., Gates, M.A., Horne, S., Amores, A., Brownlie, A., Donovan, A., Egan, E.S., Force, A., Gong, Z., Goutel, C., Fritz, A., Kelsh, R., Knapik, E., Liao, E., Paw, B., Ransom, D., Singer, A., Thomson, M., Abduljabbar, T.S., Yelick, P., Beier, D., Joly, J.S., Larhammar, D., Rosa, F., Westerfield, M., Zon, L.I., Johnson, S.L., Talbot, W.S., 1998. Vertebrate genome evolution and the zebrafish gene map. Nat. Genet. 18, 345–349.
Preziosi, R., Diana, A., Florio, D., Gustinelli, A., Nardini, G., 2007. Osteitis deformans (Paget's disease) in a Burmese python (Python molurus bivittatus) – A case report. Vet. J. 174, 669–672.
Quigley, H.A., 2011. Glaucoma. Lancet 377, 1367–1377.
Raggatt, L.J., Partridge, N.C., 2010. Cellular and molecular mechanisms of bone remodeling. J. Biol. Chem. 285, 25103–25108.
Ralston, S.H., Afzal, M.A., Helfrich, M.H., Fraser, W.D., Gallagher, J.A., Mee, A., Rima, B., 2007. Multicenter blinded analysis of RT-PCR detection methods for paramyxoviruses in relation to Paget's disease of bone. J. Bone Miner. Res. 22, 569–577.
Ralston, S.H., Langston, A.L., Reid, I.R., 2008. Pathogenesis and management of Paget's disease of bone. Lancet 372, 155–163.
Ralston, S.H., Layfield, R., 2012. Pathogenesis of Paget disease of bone. Calcif. Tissue Int. 91, 97–113.
Ravikumar, B., Sarkar, S., Davies, J.E., Futter, M., Garcia-Arencibia, M., Green-Thompson, Z.W., Jimenez-Sanchez, M., Korolchuk, V.I., Lichtenberg, M., Luo, S., Massey, D.C.O., Menzies, F.M., Moreau, K., Narayanan, U., Renna, M., Siddiqi, F.H., Underwood, B.R., Winslow, A.R.,

Rubinsztein, D.C., 2010. Regulation of mammalian autophagy in physiology and pathophysiology. Physiol. Rev. 90, 1383–1435.
Rebel, A., Malkani, K., Basle, M., Bregeon, C., Patezour, A., Filmon, R., 1974. Ultrastructural characteristics of osteoclasts in Paget's disease. Rev. Rhum. Mal. Osteoartic. 41, 767–771.
Rechsteiner, M., Rogers, S.W., 1996. PEST sequences and regulation by proteolysis. Trends Biochem. Sci. 21, 267–271.
Reddy, S. V, Singer, F.R., Mallette, L., Roodman, G.D., 1996. Detection of measles virus nucleocapsid transcripts in circulating blood cells from patients with Paget disease. J. Bone Miner. Res. 11, 1602–1607.
Rendina, D., Gennari, L., De Filippo, G., Merlotti, D., de Campora, E., Fazioli, F., Scarano, G., Nuti, R., Strazzullo, P., Mossetti, G., 2006. Evidence for increased clinical severity of familial and sporadic Paget's disease of bone in Campania, southern Italy. J. Bone Miner. Res. 21, 1828–1835.
Rezaie, T., Child, A., Hitchings, R., Brice, G., Miller, L., Coca-Prados, M., Héon, E., Krupin, T., Ritch, R., Kreutzer, D., Crick, R.P., Sarfarazi, M., 2002. Adult-onset primary open-angle glaucoma caused by mutations in optineurin. Science 295, 1077–1079.
Rezaie, T., Sarfarazi, M., 2005. Molecular cloning, genomic structure, and protein characterization of mouse optineurin. Genomics 85, 131–138.
Rogers, S., Wells, R., Rechsteiner, M., 1986. Amino acid sequences common to rapidly degraded proteins: the PEST hypothesis. Science 234, 364–368.
Roodman, G.D., 2010. Insights into the pathogenesis of Paget's disease. Ann. N. Y. Acad. Sci. 1192, 176–180.
Roodman, G.D., Windle, J.J., 2005. Science in medicine Paget disease of bone 115.
Rosen, D.R., Siddique, T., Patterson, D., Figlewicz, D.A., Sapp, P., Hentati, A., Donaldson, D., Goto, J., O'Regan, J.P., Deng, H.X., 1993. Mutations in Cu/Zn superoxide dismutase gene are associated with familial amyotrophic lateral sclerosis. Nature 362, 59–62.
Rousière, M., Michou, L., Cornélis, F., Orcel, P., 2003. Paget's disease of bone. Best Pract. Res. Clin. Rheumatol. 17, 1019–1041.
Roy, A., Kucukural, A., Zhang, Y., 2010. I-TASSER: a unified platform for automated protein structure and function prediction. Nat. Protoc. 5, 725–738.
Sadler, K.C., Amsterdam, A., Soroka, C., Boyer, J., Hopkins, N., 2005. A genetic screen in zebrafish identifies the mutants vps18, nf2 and foie gras as models of liver disease. Development 132, 3561–3572.
Sahlender, D.A., Roberts, R.C., Arden, S.D., Spudich, G., Taylor, M.J., Luzio, J.P., Kendrick-Jones, J., Buss, F., 2005. Optineurin links myosin VI to the Golgi complex and is involved in Golgi organization and exocytosis. J. Cell Biol. 169, 285–295.
Saito, K., Murai, J., Kajiho, H., Kontani, K., Kurosu, H., Katada, T., 2002. A novel binding protein composed of homophilic tetramer exhibits unique properties for the small GTPase Rab5. J. Biol. Chem. 277, 3412–3418.
Sakaguchi, T., Irie, T., Kawabata, R., Yoshida, A., Maruyama, H., Kawakami, H., 2011. Optineurin with amyotrophic lateral sclerosis-related mutations abrogates inhibition of interferon regulatory factor-3 activation. Neurosci. Lett. 505, 279–281.
Salomoni, P., Pandolfi, P.P., 2002. The role of PML in tumor suppression. Cell 108, 165–170.
Sandelin, A., Wasserman, W.W., Lenhard, B., 2004. ConSite: web-based prediction of regulatory elements using cross-species comparison. Nucleic Acids Res. 32, 249–252.
Santoriello, C., Zon, L.I., 2012. Hooked! Modeling human disease in zebrafish. J. Clin. Invest. 122, 2337–2343.
Schwamborn, K., Weil, R., Courtois, G., Whiteside, S.T., Israël, A., 2000. Phorbol esters and cytokines regulate the expression of the NEMO-related protein, a molecule involved in a NF-kappa B-independent pathway. J. Biol. Chem. 275, 22780–22789.
Selby, P.L., Davie, M.W.J., Ralston, S.H., Stone, M.D., 2002. Guidelines on the management of Paget's disease of bone. Bone 31, 366–373.
Seton, M., Choi, H.K., Hansen, M.F., Sebaldt, R.J., Cooper, C., 2003. Analysis of environmental factors in familial versus sporadic Paget's disease of bone-the New England Registry for Paget's Disease of Bone. J. Bone Miner. Res. 18, 1519–1524.
Shankar, S., Hosking, D.J., 2006. Biochemical assessment of Paget's disease of bone. J. Bone Miner. Res. 21 Suppl 2, 22–27.
Shembade, N., Harhaj, N.S., Liebl, D.J., Harhaj, E.W., 2007a. Essential role for TAX1BP1 in the termination of TNF-α-, IL-1- and LPS-mediated NF-κB and JNK signaling. EMBO J. 26, 3910–3922.

Shembade, N., Harhaj, N.S., Yamamoto, M., Akira, S., Harhaj, E.W., 2007b. The Human T-Cell Leukemia Virus Type 1 Tax Oncoprotein Requires the Ubiquitin-Conjugating Enzyme Ubc13 for NF- B Activation. J. Virol. 81, 13735–13742.

Shembade, N., Harhaj, N.S., Parvatiyar, K., Copeland, N.G., Jenkins, N.A., Matesic, L.E., Harhaj, E.W., 2008. The E3 ligase Itch negatively regulates inflammatory signaling pathways by controlling the function of the ubiquitin-editing enzyme A20. Nat. Immunol. 9, 254–262.

Shembade, N., Parvatiyar, K., Harhaj, N.S., Harhaj, E.W., 2009. The ubiquitin-editing enzyme A20 requires RNF11 to downregulate NF-κB signalling. EMBO J. 28, 513–522.

Shen, M.Y., Sali, A., 2006. Statistical potential for assessment and prediction of protein structures. Protein Sci. 15, 2507–2524.

Shen, Q., Little, S.C., Xu, M., Haupt, J., Ast, C., Katagiri, T., Mundlos, S., Seemann, P., Kaplan, F.S., Mullins, M.C., Shore, E.M., 2009. The fibrodysplasia ossificans progressiva R206H ACVR1 mutation activates BMP-independent chondrogenesis and zebrafish embryo ventralization. J. Clin. Invest. 119, 3462–3472.

Sippl, C., Bosserhoff, A.K., Fischer, D., Tamm, E.R., 2011. Depletion of optineurin in RGC-5 cells derived from retinal neurons causes apoptosis and reduces the secretion of neurotrophins. Exp. Eye Res. 93, 669–680.

Siris, E., Roodman, G.D., 2008. Paget's disease of bone., in: Rosen, C.J., Compston, J.E., Lian, J.B. (Eds.), Primer on the Metabolic Bone Diseases and Disorders of Mineral Metabolism, 7th Edition American Society for Bone and Mineral Research. pp. 335–343.

Siris, E.S., Kelsey, J.L., Flaster, E., Parker, S., 1990. Paget's disease of bone and previous pet ownership in the United States: dogs exonerated. Int. J. Epidemiol. 19, 455–458.

Siris, E.S., Ottman, R., Flaster, E., Kelsey, J.L., 1991. Familial aggregation of Paget's disease of bone. J. Bone Miner. Res. 6, 495–500.

Smits, P., Bolton, A.D., Funari, V., Hong, M., Boyden, E.D., Lu, L., Manning, D.K., Dwyer, N.D., Moran, J.L., Prysak, M., Merriman, B., Nelson, S.F., Bonafé, L., Superti-Furga, A., Ikegawa, S., Krakow, D., Cohn, D.H., Kirchhausen, T., Warman, M.L., Beier, D.R., 2010. Lethal skeletal dysplasia in mice and humans lacking the golgin GMAP-210. N. Engl. J. Med. 362, 206–216.

Soysa, N.S., Alles, N., 2009. NF-kB functions in osteoclasts. Biochem. Biophys. Res. Commun. 378, 1–5.

Sparks, A.B., Peterson, S.N., Bell, C., Loftus, B.J., Hocking, L., Cahill, D.P., Frassica, F.J., Streeten, E.A., Levine, M.A., Fraser, C.M., Adams, M.D., Broder, S., Venter, J.C., Kinzler, K.W., Vogelstein, B., Ralston, S.H., 2001. Mutation screening of the TNFRSF11A gene encoding receptor activator of NF kappa B (RANK) in familial and sporadic Paget's disease of bone and osteosarcoma. Calcif. Tissue Int. 68, 151–155.

Sreedharan, J., Blair, I.P., Tripathi, V.B., Hu, X., Vance, C., Rogelj, B., Ackerley, S., Durnall, J.C., Williams, K.L., Buratti, E., Baralle, F., de Belleroche, J., Mitchell, J.D., Leigh, P.N., Al-Chalabi, A., Miller, C.C., Nicholson, G., Shaw, C.E., 2008. TDP-43 mutations in familial and sporadic amyotrophic lateral sclerosis. Science 319, 1668–1672.

Stothard, P., 2000. The sequence manipulation suite: JavaScript programs for analyzing and formatting protein and DNA sequences. Biotechniques 28, 1102, 1104.

Sudhakar, C., Nagabhushana, A., Jain, N., Swarup, G., 2009. NF-kappaB mediates tumor necrosis factor alpha-induced expression of optineurin, a negative regulator of NF-kappaB. PLoS One 4, e5114.

Sudhakar, C., Vaibhava, V., Swarup, G., 2013. IRF-1-binding site in the first intron mediates interferon-γ-induced optineurin promoter activation. Biochem. Biophys. Res. Commun. 437, 179–184.

Sun, Q., Adamik, J., Windle, J., Roodman, G.D., Galson, D., 2014. Decreased Optineurin Mediates MVNP Effects on Pagetic Osteoclast Formation. ASBMR Abstr. B.

Sun, S.C., 2010. CYLD: a tumor suppressor deubiquitinase regulating NF-kappaB activation and diverse biological processes. Cell Death Differ. 17, 25–34.

Sun, S.C., Yamaoka, S., 2005. Activation of NF-κB by HTLV-I and implications for cell transformation. Oncogene 24, 5952–5964.

Swanhart, L.M., Cosentino, C.C., Diep, C.Q., Davidson, A.J., de Caestecker, M., Hukriede, N.A., 2011. Zebrafish kidney development: basic science to translational research. Birth Defects Res. C. Embryo Today 93, 141–156.

Swarup, G., Vaibhava, V., Nagabhushana, A., 2013. Functional Defects Caused by Glaucoma – Associated Mutations in Optineurin, in: Rumelt, S. (Ed.), Glaucoma - Basic and Clinical Aspects. InTech.

Tanaka, S., Takahashi, N., Udagawa, N., Tamura, T., Akatsu, T., Stanley, E.R., Kurokawa, T., Suda, T., 1993. Macrophage colony-stimulating factor is indispensable for both proliferation and differentiation of osteoclast progenitors. J. Clin. Invest. 91, 257–263.

Tancioni, F., Di Ieva, A., Levi, D., Aimar, E., Debernardi, A., Colombo, P., Gaetani, P., Rodriguez Y Baena, R., 2006. Spinal decompression and vertebroplasty in Paget's disease of the spine. Surg. Neurol. 66, 189–191.

Taylor, J.S., 2003. Genome Duplication, a Trait Shared by 22,000 Species of Ray-Finned Fish. Genome Res. 13, 382–390.

Thierry-Mieg, D., Thierry-Mieg, J., 2006. AceView: a comprehensive cDNA-supported gene and transcripts annotation. Genome Biol. 7 Suppl 1, S12.1–14.

Thompson, H.G.R., Harris, J.W., Wold, B.J., Lin, F., Brody, J.P., 2003. p62 overexpression in breast tumors and regulation by prostate-derived Ets factor in breast cancer cells. Oncogene 22, 2322–2333.

Tresse, E., Salomons, F.A., Vesa, J., Bott, L.C., Kimonis, V., Yao, T.P., Dantuma, N.P., Taylor, J.P., 2010. VCP/p97 is essential for maturation of ubiquitin-containing autophagosomes and this function is impaired by mutations that cause IBMPFD. Autophagy 6, 217–227.

Trompouki, E., Hatzivassiliou, E., Tsichritzis, T., Farmer, H., Ashworth, A., Mosialos, G., 2003. CYLD is a deubiquitinating enzyme that negatively regulates NF-κB activation by TNFR family members. Nature 424, 793–796.

Turturro, S., Shen, X., Shyam, R., Yue, B.Y., Ying, H., 2014. Effects of mutations and deletions in the human optineurin gene. Springerplus 3, 99.

Vadlamudi, R.K., Shin, J., 1998. Genomic structure and promoter analysis of the p62 gene encoding a non-proteasomal multiubiquitin chain binding protein. FEBS Lett. 435, 138–142.

Vaibhava, V., Nagabhushana, A., Chalasani, M.L.S., Sudhakar, C., Kumari, A., Swarup, G., 2012. Optineurin mediates a negative regulation of Rab8 by the GTPase-activating protein TBC1D17. J. Cell Sci. 125, 5026–5039.

Van Blitterswijk, M., van Vught, P.W.J., van Es, M.A., Schelhaas, H.J., van der Kooi, A.J., de Visser, M., Veldink, J.H., van den Berg, L.H., 2012. Novel optineurin mutations in sporadic amyotrophic lateral sclerosis patients. Neurobiol. Aging 33, 1016.e1–1016.e7.

Van Staa, T.P., Selby, P., Leufkens, H.G.M., Lyles, K., Sprafka, J.M., Cooper, C., 2002. Incidence and natural history of Paget's disease of bone in England and Wales. J. Bone Miner. Res. 17, 465–471.

Van Wesenbeeck, L., Odgren, P.R., Coxon, F.P., Frattini, A., Moens, P., Perdu, B., MacKay, C.A., Van Hul, E., Timmermans, J.P., Vanhoenacker, F., Jacobs, R., Peruzzi, B., Teti, A., Helfrich, M.H., Rogers, M.J., Villa, A., Van Hul, W., 2007. Involvement of PLEKHM1 in osteoclastic vesicular transport and osteopetrosis in incisors absent rats and humans. J. Clin. Invest. 117, 919–930.

Van Wesenbeeck, L., Odgren, P.R., MacKay, C.A., D'Angelo, M., Safadi, F.F., Popoff, S.N., Van Hul, W., Marks, S.C., 2002. The osteopetrotic mutation toothless (tl) is a loss-of-function frameshift mutation in the rat Csf1 gene: Evidence of a crucial role for CSF-1 in osteoclastogenesis and endochondral ossification. Proc. Natl. Acad. Sci. U. S. A. 99, 14303–14308.

Vinod, K.A., Reid, D.M., 2006. Paget ' s disease of bone. Surgery 24, 390–391.

Vuillemin-Bodaghi, V., Parlier-Cuau, C., Cywiner-Golenzer, C., Quillard, A., Kaplan, G., Laredo, J.D., 2000. Multifocal osteogenic sarcoma in Paget's disease. Skeletal Radiol. 29, 349–353.

Wallace, R.G., Barr, R.J., Osterberg, P.H., Mollan, R.A., 1989. Familial expansile osteolysis. Clin. Orthop. Relat. Res. 265–277.

Wang, R., Löhr, C. V., Fischer, K., Dashwood, W.M., Greenwood, J.A., Ho, E., Williams, D.E., Ashktorab, H., Dashwood, M.R., Dashwood, R.H., 2013. Epigenetic inactivation of endothelin-2 and endothelin-3 in colon cancer. Int. J. Cancer 132, 1004–1012.

Watts, G.D.J., Wymer, J., Kovach, M.J., Mehta, S.G., Mumm, S., Darvish, D., Pestronk, A., Whyte, M.P., Kimonis, V.E., 2004. Inclusion body myopathy associated with Paget disease of bone and frontotemporal dementia is caused by mutant valosin-containing protein. Nat. Genet. 36, 377–381.

Whitehouse, C.A., Waters, S., Marchbank, K., Horner, A., McGowan, N.W.A., Jovanovic, J. V., Xavier, G.M., Kashima, T.G., Cobourne, M.T., Richards, G.O., Sharpe, P.T., Skerry, T.M., Grigoriadis, A.E., Solomon, E., 2010. Neighbor of Brca1 gene (Nbr1) functions as a negative regulator of postnatal osteoblastic bone formation and p38 MAPK activity. Proc. Natl. Acad. Sci. 107, 12913–12918.

Whyte, M.P., Hughes, A.E., 2002. Expansile skeletal hyperphosphatasia is caused by a 15-base pair tandem duplication in TNFRSF11A encoding RANK and is allelic to familial expansile osteolysis. J. Bone Miner. Res. 17, 26–9.

Whyte, M.P., Mills, B.G., Reinus, W.R., Podgornik, M.N., Roodman, G.D., Gannon, F.H., Eddy, M.C., McAlister, W.H., 2000. Expansile skeletal hyperphosphatasia: a new familial metabolic bone disease. J. Bone Miner. Res. 15, 2330–44.

Whyte, M.P., Obrecht, S.E., Finnegan, P.M., Jones, J.L., Podgornik, M.N., McAlister, W.H., Mumm, S., 2002. Osteoprotegerin Deficiency and Juvenile Paget's Disease. N. Engl. J. Med. 347, 175–184.

Wierzbicki, A., Lloyd, M., Schofield, C.J., Feher, M.D., Gibberd, F.B., 2007. Peroxisomal disorders affecting phytanic acid alpha-oxidation: a review. Biochem. Soc. Trans. 35, 881–886.

Wild, P., Farhan, H., McEwan, D.G., Wagner, S., Rogov, V. V, Brady, N.R., Richter, B., Korac, J., Waidmann, O., Choudhary, C., Dötsch, V., Bumann, D., Dikic, I., 2011a. Phosphorylation of the autophagy receptor optineurin restricts Salmonella growth. Science 333, 228–233.

Wild, P., Farhan, H., McEwan, D.G., Wagner, S., Rogov, V. V, Brady, N.R., Richter, B., Korac, J., Waidmann, O., Choudhary, C., Dötsch, V., Bumann, D., Dikic, I., 2011b. Phosphorylation of the autophagy receptor optineurin restricts Salmonella growth. Science 333, 228–233.

Willoughby, C.E., Chan, L.L.Y., Herd, S., Billingsley, G., Noordeh, N., Levin, A. V, Buys, Y., Trope, G., Sarfarazi, M., Héon, E., 2004. Defining the pathogenicity of optineurin in juvenile open-angle glaucoma. Invest. Ophthalmol. Vis. Sci. 45, 3122–3130.

Witten, P.E., Bendahmane, M., Abou-Haila, A., 1997. Enzyme histochemical characteristics of osteoblasts and mononucleated osteoclasts in a teleost fish with acellular bone (Oreochromis niloticus, Cichlidae). Cell Tissue Res. 287, 591–599.

Witten, P.E., Hansen, A., Hall, B.K., 2001. Features of mono- and multinucleated bone resorbing cells of the zebrafish Danio rerio and their contribution to skeletal development, remodeling, and growth. J. Morphol. 250, 197–207.

Witten, P.E., Huysseune, A., 2009. A comparative view on mechanisms and functions of skeletal remodelling in teleost fish, with special emphasis on osteoclasts and their function. Biol. Rev. Camb. Philos. Soc. 84, 315–346.

Wiweger, M.I., Zhao, Z., van Merkesteyn, R.J.P., Roehl, H.H., Hogendoorn, P.C.W., 2012. HSPG-deficient zebrafish uncovers dental aspect of multiple osteochondromas. PLoS One 7, e29734.

Wuyts, W., Van Wesenbeeck, L., Morales-Piga, A., Ralston, S., Hocking, L., Vanhoenacker, F., Westhovens, R., Verbruggen, L., Anderson, D., Hughes, A., Van Hul, W., 2001. Evaluation of the role of RANK and OPG genes in Paget's disease of bone. Bone 28, 104–107.

Xi, Y., Noble, S., Ekker, M., 2011. Modeling neurodegeneration in zebrafish. Curr. Neurol. Neurosci. Rep. 11, 274–282.

Xiao, Z., Meng, Q., Tsai, J.C., Yuan, H., Xu, N., Li, Y., 2009. A novel optineurin genetic mutation associated with open-angle glaucoma in a Chinese family. Mol. Vis. 15, 1649–1654.

Xu, J., Rogers, M.B., 2007. Modulation of Bone Morphogenetic Protein (BMP) 2 gene expression by Sp1 transcription factors. Gene 392, 221–229.

Yagi, M., Miyamoto, T., Sawatani, Y., Iwamoto, K., Hosogane, N., Fujita, N., Morita, K., Ninomiya, K., Suzuki, T., Miyamoto, K., Oike, Y., Takeya, M., Toyama, Y., Suda, T., 2005. DC-STAMP is essential for cell-cell fusion in osteoclasts and foreign body giant cells. J. Exp. Med. 202, 345–351.

Yang, Z., Klionsky, D.J., 2009. Autophagy in Infection and Immunity, Current topics in microbiology and immunology, Current Topics in Microbiology and Immunology. Springer Berlin Heidelberg, Berlin, Heidelberg.

Ying, H., Shen, X., Park, B., Yue, B.Y.J.T., 2010. Posttranslational modifications, localization, and protein interactions of optineurin, the product of a glaucoma gene. PLoS One 5, e9168.

Ying, H., Yue, B.Y.J.T., 2012. Cellular and Molecular Biology of Optineurin. Int. Rev. Cell Mol. Biol. 294, 223–258.

Yoshida, H., Hayashi, S.I., Kunisada, T., Ogawa, M., Nishikawa, S., Okamura, H., Sudo, T., Shultz, L.D., Nishikawa, S.I., 1990. The murine mutation osteopetrosis is in the coding region of the macrophage colony stimulating factor gene. Nature 345, 442–444.

Yu, S., Yerges-Armstrong, L.M., Chu, Y., Zmuda, J.M., Zhang, Y., 2013. E2F1 effects on osteoblast differentiation and mineralization are mediated through up-regulation of frizzled-1. Bone 56, 234–241.

Zhu, G., Wu, C.J., Zhao, Y., Ashwell, J.D., 2007. Optineurin Negatively Regulates TNFα- Induced NF-κB Activation by Competing with NEMO for Ubiquitinated RIP. Curr. Biol. 17, 1438–1443.

Zon, L.I., 1999. Zebrafish: a new model for human disease. Genome Res. 9, 99–100.

APPENDICES

Table S2.2 – Detection of genetic variants in the candidate genes in *PDB6* locus in the discovery sample and *in silico* analysis

Variant	Minor allele frequency		Functional predicted effect					
	PDB patients	Healthy controls	Protein sequence	TF binding		Splicing	miRNA binding	
				GAIN	LOSS		GAIN	LOSS
rs3829923 G/A	0.23	0.00	-	THING1-E47	NRF-2, E74A, SAP1	-	Hsa-miR-1324, Hsa-miR-4738-3p	-
c. -580 - 14 C>T	0.02	0.00	-	IK-, SOX 17	KR, CF2-II, BROAD-COMPLEX_4	-	Hsa-miR-4729	-
rs7093805 G/T	0.16	0.30	-	-	ADR1, CAP	-	-	-
rs4750306 A/G	0.20	0.20	-	ADR1, STRE, MZF1, HUNCHBACK	HSF	-	Hsa-miR-5010-3p	Hsa-miR-4778-3p
rs72779542 T/A	0.20	0.20	-	OCT1, BR-C Z, BROAD COMPLEX 1/4, HUNCHBACK, HFH2	TBP	-	-	Hsa-miR-4777-5p

rs61851395 C/G	0.20	0.20	-	-	-	-	Hsa-miR-5095	-
rs663177 A/C	0.62	0.50	-	HSF, E74A	CDXA, HSF, TBP, ATHB-1	-	-	Hsa-miR-5680
c. -133-31 G>T	0.02	0.00	-	CDXA	TATA, TBP, AGL3, ARNT	-	-	-
c. 364 + 93 C>T	0.02	0.00	-	-	CAP	-	-	-
rs116856516 C/A	0.02	0.00	p.Pro266His	P300, CDXA, MZF	ADR-1, STRE	New branch point	-	-
rs10752280 A/G	0.98	1.00	-	ATHB5	CRE-BP, THING1-E47, BZIP910	-	-	-
rs11257730 G/C	0.11	0.00	-	-	ADR1, SP1, MZF1	-	-	Hsa-miR-4286, Hsa-miR-3972, Hsa-miR-4433-3p, Hsa-miR-

								4498,v Hsa-miR-4433-5p
rs11257731 G/T	0.53	0.70	-	CP2, THING1-E47	ADR1, SP1, MZF1	-	-	Hsa-miR-3972, Hsa-miR-4433-3p, Hsa-miR-4498, Hsa-miR-4433-5p
rs7070092 C/T	0.11	0.00	-	-	-	-	Hsa-miR-4669,v Hsa-miR-4664-5p	Hsa-miR-1909-5p
rs11257733 C/T	0.11	0.00	-	BZIP910	GATA1, C-REL, P65, E74A	-	Hsa-miR-513a-5p	Hsa-miR-18b-3p, Hsa-miR-3126-5p,
rs145905575 G/A	0.05	0.00	-	-	-	-	-	Hsa-miR-3194-3p
c. -61 G>A	0.02	0.00	-	CAP	SP1, CFI-USP	-	Hsa-miR-675-5p	Hsa-miR-1469, Hsa-miR-3960, Hsa-miR-

								4665-3pv
rs2895524 C/T	0.39	0.40	-	AML-1	-	New acceptor site	Hsa-miR-371a-5p, Hsa-miR-371b-3p	-
rs34194224 A/G	0.02	0.00	p.Ile66Met	BRN-2, STAF	HUNCHBACK	Loss of an acceptor site	Hsa-miR-3658,v Hsa-miR-4716-3p	-
rs1077745 G/A	0.06	0.10	-	-	C-ETS, E74A	-	-	-
rs1077744 C/T	0.42	0.30	-	-	-	New acceptor site	Hsa-miR-5008-3p	-
rs139311431 C/G	0.02	0.00	-	HMG-IY	AML-1A, THING1-E47	New acceptor site	Hsa-miR-637	Hsa-miR-3926
rs45497595 G/A	0.05	0.00	-	-	E74A, NRF-2, TBP	New acceptor site	Hsa-miR-4713-5p	-

rs1644417 T/C	0.02	0.00	-	DELTAE, USF, ARNT, N-MYC	ADR1, GATA 1, GATA 2, HSF, TAL1BETA-E47	New acceptor site	-	-
rs1644418 A/G	0.02	0.00	-	C-MYB, HFH-1, BROAD-COMPLEX_4	CDXA, CAP, C-MYB, SKN-1, TBP, SQUA	Loss of a donor site	-	-
rs2482023 A/C	0.79	0.90	-	ADR1, MZF1, CFI-USP	-	New acceptor site	Hsa-miR-4523	-
rs12768271 C/G	0.05	0.10	-	-	-	New donor site	-	-
rs1757051 C/G	0.74	0.80	p.Val316	C-FOS	CAP, ATF, BZIP910	New donor site	-	-
rs2482075 A/G	0.80	0.90	-	ADR1, SU-H	IK-2	-	-	Hsa-miR-3620, Hsa-miR-4639-3p
rs732704 G/C	0.26	0.10	-	-	GCR1, ADR1	-	Hsa-miR-150-3p, Hsa-miR-501-3p	Hsa-miR-4687-5p

rs511959 C/G	0.97	1.00	-	-	STUAP, CFI-USP	-	-	Hsa-miR-139-3p
rs608353 A/C	0.97	1.00	-	ADR1, STATX, E74A	TTK, SNAIL	-	-	-
rs732701 T/C	0.26	0.10	-	GBP, CREB	STUAP	-	Hsa-miR-663a	-
rs1058596 C/G	0.36	0.40	-	-	-	-	Hsa-miR-4730, Hsa-miR-1292	-
rs3802587 C/T	0.03	0.00	-	MYF	-	New acceptor site	-	Hsa-miR-3187-5p
rs3802586 G/A	0.36	0.40	-	AP-4, CAP, C-FOS	-	-	Hsa-miR-4265, Hsa-miR-146b-3p, Hsa-miR-3160-5p, Hsa-miR-3160-3p	-

rs36143400 G/A	0.02	0.00	-	SNAIL	NRF-2, E74A	-	Hsa-miR-4326	Hsa-miR-4669, Hsa-miR-665
rs4615920 T/C	0.24	0.10	-	NRF-2, E74A, MYF	CAP	-	Hsa-miR-1913	Hsa-miR-22-3p
rs72781364 G/A	0.03	0.00	-	-	-	-	-	-
rs72781363 C/T	0.03	0.00	-	HEN1, CREB, BZIP910	-	-	Hsa-miR-329	-
rs76403928 A/C	0.36	0.40	-	HNF-3BETA	-	-	-	Hsa-miR-5096
rs10796054 C/T	0.36	0.40	-	-	ADR1, C-FOS, SNAIL	-	-	-
rs7916926 T/C	0.39	0.20	-	STUAP, SNAIL	-	-	-	Hsa-miR-3622a-5p, Hsa-miR-3622b-3p

rs28938169 C/T	0.21	0.20	p.Pro29Ser	HSF, TEF-1	-	-	-	-
rs72781358 C/T	0.03	0.00	-	-	-	-	Hsa-miR-25-3p	Hsa-miR-1228-3p, Hsa-miR-3124-5p
rs1747682 C/T	1.00	1.00	p.Asn51	CDXA, SOX 5, ATHB-1	HFH-2, BZIP910	-	-	-
rs473407 A/G	0.27	0.30	p.Thr112	HSF2, HSF1, BSAP	HSH-2	New acceptor site	-	-
rs7900830 C/A	0.18	0.30	-	HSF, SRY, CF2-II, AGL3, TBP, HFH-1	-	New acceptor site	-	-
rs78560997 A/G	0.02	0.10	-	SN, DELTAE, MYOD	NHF-3BETA, FREAC-4, SOX 5	New acceptor site	-	Hsa-miR-5010-5p

rs648173 G/A	0.50	0.70	-	HSF, SRY, HUNCHBACK	-	New acceptor site	Hsa-miR-3611	Hsa-miR-3202, Hsa-miR-519c-5p, Hsa-miR-520a-5p, Hsa-miR-523-5p, Hsa-miR-518e-5p, Hsa-miR-522-5p, Hsa-miR-519a-5p, Hsa-miR-519b-5p
rs62619919 G/A	0.02	0.10	p.Arg245Gln	SN, DELTAE, MYOD, USF, ARNT, N-MYC	-	New donor/acceptor site	Hsa-miR-324-3p	Hsa-miR-4640-5p
rs12411977 A/G	0.52	0.30	-	HSF, CDXA	HUNCHBACK, SOX 5, ATHB-1	New acceptor site	-	-
rs11133 G/A	0.30	0.40	-	HFH-2, HNF-3BETA, SOX 17, ATHB-1, BROAD COMPLEX-4	HSF-2, BROAD COMPLEX-1/4, SOX 5, FREAC-4	-	Hsa-miR-3915	Hsa-miR-664-5p, Hsa-miR-515-5p

rs61851599 G/A	0.12	0.30	-	USF, CREB, AP1, DELTAE, HLF, BZIP910	PACC, MYC-MAX	New donor site	-	-
c. -78 -125 A>G	0.02	0.00	-	-	TST1, CDXA, TBP, AGL3, CF2-II	-	-	-
ESP_10_13386978 G/A	0.02	0.00	-	HSF, CAP	AP4	-	-	Hsa-miR-3147
rs2275128 T/C	0.50	0.70	-	MZF1, DORSAL-1, DORSAL-2, NF-KAPPAB, C-REL, P50	-	-	-	Hsa-miR-1247-3p
rs62641683 C/G	0.02	0.00	p.Leu125	CAP, MYF	SPZ1	New acceptor site	Hsa-miR-378g	Hsa-miR-151a-3p
rs2275129 G/C	0.33	0.40	-	-	-	New donor site	-	Hsa-miR-3922-3p
rs10752297 A/G	0.49	0.70	p.Thr157	BZIP910	MYF, TAL1BETA-E47A, SNAIL	New donor/acceptor site	-	-

rs61851591 G/T	0.12	0.30	-	SRY, BR-C Z, SBF-1, DFD, BROAD COMPLEX 4, SOX 5, HFH-1	EVI-1	New donor site	-	-
rs3740211 G/A	0.36	0.40	-	-	ADR1, STRE, SNAIL	New donor/acceptor site	-	-
rs62641685 G/C	0.03	0.00	p.Val293	TTK69	MYF, SNAIL	New donor/acceptor site	-	Hsa-miR-338-5p, Hsa-miR-3065-3p
rs10906347 G/A	0.35	0.40	-	SRY, SPZ1, AML1	MYF, HEN1	New acceptor site	-	Hsa-miR-3675-5p
RV (-9906) G>A	0.02	0	-	-	CAP, GCBOX	-	-	-
rs3829924 G>A	0.016	0	-	HUNCHBACK, HMG-IY	-	-	-	Hsa-miR-578
rs2234968 G>A	0.30	0.42	-	CF2-II	MAX, USF, ARNT, N-MYC, FREAC-4	New branch point	Hsa-miR-302a-3p, Hsa-miR-302d-3p,	Hsa-miR-619

							Hsa-miR-5188	
rs11258194 T>A	0.03	0.08	p.Met98Lys	-	UBX	-	-	-
rs72043574 delta	0.42	0.33	-	-	HMG-YV, HFH2, BROAD COMPLEX 1/4, HUNCHBACK	-	-	-
rs1561570 T>C			-	-	AGL3, CF2-II, TAL1BETA-E47	New donor/acceptor site	-	-
rs2244380 C>T	0.84	0.92	-	STATX, TATA, HMG IY, SQUA	SU_H	New acceptor site	-	Hsa-miR-145-5p, Hsa-miR-452-5p
rs11258211 G>A	0.02	0	-	GATA1, GATA2, NIT2, SPZ1	THING1-E47	New donor/acceptor site	-	-
rs765884 T>C	0.38	0.58	-	-	SQUA	New acceptor site	-	Hsa-miR-3164

rs489040 A>G	0.58	0.75	-	-	MATALP, CDXA, SKN1	-	-	-
rs523747 A>G	1	1	p.Lys322Glu	HSF	AP4, HEN1	New acceptor site	Hsa-miR-891a	Hsa-miR-19a-3p
rs676302 G>T	0.72	0.75	-	NKX2, HMG IY	ADR1, HUNCHBACK	New acceptor site	Hsa-miR-3128	-
rs825411 A>G			-	V-MYB, GATA2, SPZ1	HFH2, SOX 5	New donor/acceptor site	-	-
rs10906310 C>A	0.22	0.08	-	BROAD COMPLEX 1/4, HFH1, HUNCHBACK, SOX 5	SNAIL	-	-	Hsa-miR-653, Hsa-miR-365a-5p
rs2095388 A>G			-	FTZ, AGL3	S8, BROAD COMPLEX 4	-	-	-

- indicates that the variant does not have a predicted functional effect.

Table S3.1 – Primers list

Name	Sequence (5' - 3')
Sense primers	
Rs1561570 qPCR F	CAGAAAGGTCATCTGAGGACC
Rs2234968 qPCR F	GAGAACCACCAGCTGAAAGAAG
Rs10906303 F1	TGAAAGCACAGGAAATGGAC
Rs79529484 F2	AAGAGCGTCTAATGGCCTTG
Rs3829923 F1	CTCGAGGGCCCGGAATTGGGTCT
Rs3829923 F2	CTCGAGGGCCTGGAATTGGGTCT
RV -9906 F3	CAGGGGCAGGGTTGAGAGTCAGGCT
RV -9906 F4	CAGGGGCAGAGTTGAGAGTCAGGCT
Rs1561570 Bisulfite F	TTGTTATAATTGTGTTTGGGTATAGGG
OPTN qPCR F	CAATGTCCCATCAACCTCTCAGC
NF-κB qPCR F	CAGGAGAGGATGAAGGAGTTGTG
ELK1 qPCR F	TCTCCTCCACAGAAAACACACAGC
IL6 qPCR F	GAAAGCAGCAAAGAGGCACTGG
NFATc1 qPCR F	ATGTCGGAGTTTCTGAGTTTCA
Antisense primers	
Rs1561570 qPCR R	CCTTCTGCTTCTCCTTCAGCC
Rs2234968 qPCR R	GTCATCAGTGGGGTCCTCAGA
Rs10906303 R1	CTTGGTGGTGAGCTCCCTTCT
Rs79529484 R2	TTATAGATGGTGCAGATGGGC
Rs3829923 R1	TGCCGAGGAAACAGTGACCCTGA
RV -9906 R1	TGCCGAGGAAACAGTGACCCTGA
Rs1561570 Bisulfite R	ACTCAAAAAATCCTACCTCCTCAA
OPTN qPCR R	ACGCTCTTTTGCTTCTTTGCTCTG
NF-κB qPCR R	CATAAGTGTTTTGGAAGGAGCAGGA
ELK1 qPCR R	ACTCACATCCAAAAACGCAGACAG
IL6 qPCR R	GGTTGGGTCAGGGGTGGTTATT
NFATc1 qPCR R	TCTCCACGAAAATGACCTTG