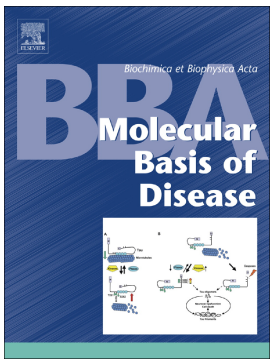
Accepted Manuscript



Effect of genetic variants of OPTN in the pathophysiology of Paget's disease of bone

Iris A.L. Silva, Natércia Conceição, Édith Gagnon, Helena Caiado, Jacques P. Brown, Fernando Gianfrancesco, Laëtitia Michou, M. Leonor Cancela

PII: S0925-4439(17)30363-0
DOI: doi:10.1016/j.bbadis.2017.10.008
Reference: BBADIS 64920

To appear in:

Received date: 23 December 2016
Revised date: 2 October 2017
Accepted date: 4 October 2017

Please cite this article as: Iris A.L. Silva, Natércia Conceição, Édith Gagnon, Helena Caiado, Jacques P. Brown, Fernando Gianfrancesco, Laëtitia Michou, M. Leonor Cancela , Effect of genetic variants of OPTN in the pathophysiology of Paget's disease of bone. The address for the corresponding author was captured as affiliation for all authors. Please check if appropriate. Bbadis(2017), doi:10.1016/j.bbadis.2017.10.008

Effect of genetic variants of OPTN in the pathophysiology of Paget's disease of bone

Iris A.L. Silva, [a, b, 1] Natércia Conceição, [b, c] Édith Gagnon, [d] Helena Caiado, [b,g] Jacques P. Brown, [d, e] Fernando Gianfrancesco, [f] Laëtitia Michou [d, e] * and M. Leonor Cancela [b, c] *

[a]*PhD program in Biomedical Sciences and Medicine, University of Algarve, Faro, 8005-139, Portugal*
[b]*Centre of Marine Sciences (CCMAR), University of Algarve, Faro, 8005-139, Portugal*
[c]*Department of Biomedical Sciences and Medicine, University of Algarve, Faro, 8005-139, Portugal*
[d]*Research centre of the CHU de Québec-Université Laval, Québec City, QC, Canada*
[e]*Division of Rheumatology, Department of Medicine, Université Laval and Department of Rheumatology, CHU de Québec-Université Laval, Québec City, QC, Canada*
[f]*Institute of Genetics and Biophysics "Adriano Buzzati-Traverso," National Research Council of Italy, 80131 Naples, Italy.*
[g]*PhD program in Regenerative Medicine, University of Algarve, Faro, 8005-139, Portugal*

[1]*Present adress: Faculty of Sciences, BioISI - Biosystems & Integrative Sciences Institute, University of Lisboa, Campo Grande, 1749-016, Lisbon, Portugal.*

Correspondence:
M. Leonor Cancela
Department of Biomedical Sciences and Medicine, University of Algarve, 8005 Faro, Portugal
Email: lcancela@ualg.pt
Tel.: + 351289800097

Laetitia Michou
Rhumatologie-R4774, CHU de Québec-Université Laval, 2705 boulevard Laurier, Québec, G1V 4G2, Québec, Canada.
Email: laetitia.michou@crchudequebec.ulaval.ca
Tel: + 1 418 654 2178; Fax: +1 418 654 2142

ABSTRACT

Paget's disease of bone (PDB) is the second most frequent metabolic bone disease after osteoporosis. Genetic factors play an important role in PDB, but to date PDB causing mutations were identified only in the *Sequestosome 1* gene at the *PDB3* locus. OPTN has been recently associated with PDB, however little is known about the effect of genetic variants in this gene in PDB pathophysiology. By sequencing *OPTN* in *SQSTM1* non-carriers PDB patients we found 16 SNPs in regulatory, coding and non-coding regions. One of those was found to be associated with PDB in our cohort - rs2234968. Our results show that rs2238968 effect may be explained by a change in *OPTN* splicing that give rise to a predicted truncated protein. We also performed functional studies on the variants located in *OPTN* promoter – rs3829923 and the rare variant -9906 – to investigate putative regulators of *OPTN*. Our results show that *OPTN* expression seems to be regulated by SP1, RXR, E47, and the E2F family. In conclusion, our work suggests a potential pathophysiological role of SNPs in *OPTN*, giving a new perspective about the regulatory mechanisms of this gene. Ultimately we discovered a new variant associated with PDB in *OPTN*, reinforcing the relevance of this gene for the development of this bone disease.

Keywords: Paget's disease of bone, optineurin, gene regulation, splicing, rs2234968

1. INTRODUCTION

Optineurin was first identified by Li's group [1] using a yeast two-hybrid system and named FIP-2 - interacting protein of Adenovirus E3 14.7-kDa protein. Later, in 2002, mutations in this gene were found to be associated with primary open-angle glaucoma, a disease that causes irreversible bilateral blindness, and the gene was renamed "optineurin" (*OPTN*) [2]. The human *OPTN* gene is located in the short arm of chromosome 10 (13099449-13138308) and spans about 39 Mbp 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. Also, two partial transcripts were also described, but there is no indication of being protein coding. In addition, in 2012 a 'new first exon' was described, upstream from the previously known exon 1, and was labelled as exon 1a [3]. 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 [4]. OPTN is known to undergo posttranslational modifications, after being ubiquitinated it is processed through the ubiquitin-proteasome pathway [5]. It is also phosphorylated at Ser177, which is adjacent to the LC3 interacting region (LIR) site [6]. OPTN is a multifunctional protein involved in several biological functions such as NF-κB regulation, autophagy, membrane trafficking, exocytosis, vesicle transport, transcriptional activation, and reorganizing of actin and microtubules, since it interacts with several proteins, including Rab8, huntingtin, transcription factor IIIA, myosin VI, and TANK binding protein 1

(TBK1). OPTN biological role is not yet fully understood and appears to be complex and involve different mechanisms and pathways (reviewed in [7]). Besides glaucoma, a number of diseases such as neurodegenerative diseases (like Alzheimer's disease, Parkinson's disease [8] and amyotrophic lateral sclerosis [9]) and Paget's disease of bone (PDB) [3,10] have been associated with mutations in *OPTN*. PDB is the second most frequent metabolic bone disorder, after osteoporosis [11], affecting between 1% and 3% of individuals over the age of 55 years [12,13]. This disease is characterized by focal abnormal bone remodelling, with increased bone resorption and accelerated, excessive, and disorganized new bone formation. The pathophysiology of PDB is currently an area of intensive investigation, and this disease seems to have both genetic and non-genetic causes. Fifteen to 40 percent (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 [14–16]. Linkage studies in these families have identified a number of susceptibility loci on chromosomes *6p21* (*PDB1*) [17], *18q21.1-22* (*PDB2*) [18], *5q35* (*PDB3*), *5q31* (*PDB4*) [15], *2q36* (*PDB5*) [19], *10p13* (*PDB6*) [20] and *18q23* (*PDB7*) [19]. The regions identified are typically large and contain several genes that could be possible candidates based on their known functions. Moreover, a genome-wide scan in British families with PDB has shown a linkage to the *10p13* (*PDB6*) locus [20,21]. Recently, reanalysis of data from this genome-wide scan confirmed a genetic association to the *10p13* locus [10], namely to the rs1561570 SNP located in *OPTN* gene, but no PDB causal mutation has been reported to date in this locus. Our group replicated the strong and statistically significant genetic association of rs1561570 (p-value = 5.65×10^{-7}) with PDB in the French-Canadian population [3], and also identified a functional SNP in *UCMA/GRP,* but with a marginal association with PDB, and a rare variant in *OPTN* promoter predicted to alter the putative binding of bone transcription factors [3]. In this work, we have assessed the possible contribution of each

significant variant identified in the *OPTN* gene and have selected rs3829923 and rs2234968 to perform the association study and functional analysis in order to determine the involvement of these variants in the PDB pathophysiology. We also analysed the effect of a rare variant (RV - 9906) found in *OPTN* promoter and described in our previous work [3].

2. MATERIAL AND METHODS

2.1 Study participants and candidate gene sequencing

This study was approved by the CHU de Québec-Université Laval 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 [22–24]. For each individual, peripheral blood was obtained as described in [3]. 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 as described in [3].

2.2 *OPTN* genetic variation screening

The strategies used throughout this study are briefly summarized in Fig. S1. To search for variants within the *OPTN* gene, the exons, their exon–intron boundaries, 5' and 3'-UTRs and the basal promoter were sequenced as previously described [3]. 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 controls and 298 patients, based on their frequency in the discovery group and location in the gene. The allele frequencies were calculated as previously described [3].

2.3 Bioinformatic analysis

To identify the potential effect of the variants found in the coding regions we used the online translate tools Expasy (http://web.expasy.org/translate/), SIFT (http://siftdna.org/www/SIFT_dbSNP.html), Condel (http://bg.upf.edu/fannsdb/), Polyphen (http://genetics.bwh.harvard.edu/pph2/) and Mutation taster software (http://www.mutationtaster.org/). Potential changes in the splice sites were predicted using the Human Splicing Finder tool (http://www.umd.be/HSF/). To identify transcription factor binding sites (TFBSs) for the variants located in promoter regions we used TFsearch (http://diyhpl.us/~bryan/irc/protocol-online/protocol-cache/TFSEARCH.html) and Consite (http://consite.genereg.net/). To analyse 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 (https://www.broadinstitute.org/mpg/snap/).

2.4 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. 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 containing 10% FBS (Life technologies). The cell suspension was added to 6-well plates ($9x10^6$ cells/well) and to Lab-Tek 8 well-slides ($3\text{-}6x10^5$ cells/well). After 24

h, 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).

2.5 Quantitative real-time PCR

In order to test if rs2234968 had an impact on *OPTN* gene expression, we performed qPCR as previously described [3]. We used RNA from PBMCs of six patients and three healthy donors. A quantity of cDNA corresponding to 10 ng of total RNA was used in these analyses. The primers used are displayed in Table S1. For every assay, a negative control was performed in the absence of cDNA template. Fluorescence was measured at the end of each extension cycle and melting profiles of each reaction were performed to check for unspecific product amplification. Relative mRNA expression was calculated using the ΔΔCt method [25]. 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 [26].

2.6 Western blot analyses

Osteoclasts derived from PBMCs from patients and controls were washed once in PBS and lysed using Trizol. The protein concentrations were determined using the Bradford reagent (Bio-Rad). Proteins were separated by an 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 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 1 h, 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.

2.7 *In vitro* functional analysis of *OPTN* promoter SNPs

Four fragments of the *OPTN* promoter were amplified from human genomic DNA using four different primer sets to originate the constructs rs3829923 C/T and RV-9906 G/A. Briefly, a 1.2 kb fragment of the *OPTN* promoter ranging from -315 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 S1) – construct rs3829923 *C*. A point mutation, corresponding to the *T* allele of rs3829923 (position -305), was generated by PCR using the forward primer rs3829923 F2 (Table S1) and the same reverse primer – construct rs3829923 *T*. 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 S1) – construct RV -9906 *G*. 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 S1) and the same primer reverse – construct RV -9906 *A*. The amplified PCR products were cloned into pCRII TOPO (Invitrogen) and the fidelity of the sequence was confirmed by DNA sequencing. The DNA inserts 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.

2.8 Cell transfections and luciferase activity assay

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 nonadherent THP1 cell line was cultured in RPMI 1640 medium supplemented with 10% FBS, 2 mM L-glutamine and 1% 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.

HEK293 cells were cultured on 24 well plates (5 x 10^4 cells/well) and transiently transfected using XtremeGENE HP (Roche), with 250 ng of *OPTN* reporter gene rs3829923 *C* (F1), rs3829923 *T* (F2), RV -9906 *G* (F3), RV -9906 *A* (F4), or empty reporter vector (pGL3-basic). Co-transfections of an additional expression construct (50 ng/µl) carrying each transcription factor were performed using a similar approach.

THP1 cells (5 x 10^4 cells per sample) were transiently transfected with 500 ng of plasmids' DNA using electroporation (Lonza nucleofector technology) in accordance with manufacturer's instructions. Specifically, 100 µl Cell Line Nucleofector® Solution V were used.

After 48 h the cells were lysed and luciferase activity was assayed using a Dual-luciferase Reporter Assay kit (Promega) in accordance with the manufacturer's instructions. All luciferase activities were normalized to the *Renilla* luciferase reporter pRL-Null Luc plasmid (Promega), which was unresponsive to the overexpression of the transcription factors used. All experiments were repeated at least three times. The SP1 expression plasmid and the RXR expression plasmid were a generous gift from 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).

2.9 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 $p<0.05$.

3. RESULTS

3.1 Discovery study – Variants identification in *OPTN*

In this analysis, we have detected 16 genetic variants and their position is shown in Fig. 1 and Table S2. Their respective frequency in the discovery group and *in silico* predicted functions are shown in Table S3. We found two variants with predicted effects on protein sequence, 16 variants predicted to cause loss and/or gain of TFBSs, seven variants predicted to have an effect on splicing, and eight variants predicted to cause loss and/or gain of miRNA binding sites (Table S3).

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

From all these variants, we selected rs3829923 and rs2234968 for genetic association with PDB in our French-Canadian cohort. The two variants were selected due to the high frequency in the discovery group (rs3829923 – found in 13 of the 30 patients and in none of the healthy controls; rs2234968 – found in 15 patients and in three of five controls), due to the minor allele frequency in the European population (25% for rs3829923 and 18% for rs2234968), and due to their localization in the gene (rs3829923 is located in *OPTN* promoter and can affect its expression, and rs2234968 is

located in *OPTN* coding region and can affect either protein-protein interactions, OPTN role or localization in the cell, or its mRNA secondary structure). The genotypes of these variants were in Hardy-Weinberg equilibrium in controls (data not shown). The genetic association study demonstrated that only rs2234968 was associated with PDB (MAF in patients = 33%, MAF in controls = 25%, uncorrected *p*-value = 6 x 10^{-3}, OR = 0.6674; CI 95% = [0.50 ; 0.89]). This association was further replicated in an Italian population composed by 364 controls and 475 PDB patients (MAF in patients = 39%; MAF in controls =32%, uncorrected *p*-value = 4 x 10^{-3}). The association study also showed that rs3829923 was not associated with PDB (MAF in patients = 34%, MAF in controls = 33%, uncorrected *p*-value = 1, OR = 1.0006; CI 95% = [0.76 ; 1.32]).

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

Several possible functional effects of rs2234968 were analysed: 1) *OPTN* mRNA secondary structure according to RNA fold program revealed an absence of a predicted effect in mRNA structure (Fig. S2A); 2) *OPTN* expression levels by qPCR using patients and healthy donors samples with all genotypes, showed an absence of correlation between genotype and *OPTN* expression levels (Fig. S2B); and 3) *OPTN* splicing pattern, according to Human Splicing Finder tool, indicated a prediction of a new branch point (73% of probability) and consequently a possible exon 5 skipping, which would change the protein sequence (Fig. S2C). However, by doing a PCR using primers flanking exon 5 and cDNA from patients as template, we showed that rs2234968 did not induce exon 5 skipping (Fig. S2D). 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* introns 2, 3, 5 and 6, respectively (Fig. 2A). Using Human Splicing Finder tool, we determined that rs10906303 and rs79529484 were the most likely to have a functional effect on splicing, since the results show that rs10906303 has a high probability to create a new donor splice site and a new acceptor splice site, and rs79529484 also has a probability to create a new acceptor splice site, while the other SNPs analysed did not have a predicted effect on splicing (Fig. 2B). These effects were tested by PCR using cDNA samples from patients with all genotypes and specific primers (Fig. 2C). Results showed that both rs10906303 and rs79529484 had the potential to alter the splicing sites, causing retention of introns 5 and 6 respectively (Fig. 2C), which would change the amino acid sequence and lead to a predicted premature stop codon (Fig. 2D). By using an antibody against the OPTN C terminus in a western blot analysis, we detected lower OPTN protein levels in the presence of rs2234968 *A* allele (Fig. 2E), but not an additional band corresponding to a truncated protein, suggesting that the *A* allele might be producing a truncated protein lacking the region recognized by the antibody used (C terminus), as predicted by our *in silico* analysis.

3.4 *In silico* prediction of transcription factor binding sites in rs3829923 and RV -9906 located in *OPTN* promoter

In order to explore the functional effect of the variants located in *OPTN* promoter, we searched for TFBSs that could be potentially interrupted by these variants, using the online tools Consite and TFSearch. Rs3829923 causes a change from a *C* allele to a *T* allele (Fig. 3A) and the bioinformatic software predicted the putative binding sites for SAP1, NRF2, E74A and E2F family in the region of the rs3829923 carrying the *C* allele, which are lost in the presence of the *T* allele. On the other hand, in the presence of the *T* allele the software predicted a putative binding site for E47

transcription factor (Fig. 3B). We also assessed the putative effect of *OPTN* RV -9906 G/A (Fig. 3C), a polymorphism previously reported [3], in *OPTN* promoter activity. 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 efficiencies 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 (Fig. 3D).

3.5 Functional analysis of rs3829923 in *OPTN* promoter

To analyze the functionality of the two different alleles of rs3829923, we transfected HEK293 cells and performed luciferase assays with the two constructs, F1 and F2, carrying respectively the *C* or the *T* alleles. The results showed an increase in *OPTN* promoter activity in construct F2 when compared to construct F1, both in HEK293 (Fig. 4A) and in THP1 (Fig. 4B). The cells were then transfected with the F1 construct or the F2 construct together with the transcription factors (TFs) predicted to bind to this region in the *in silico* analysis. The co-transfection of F1 with E47 and E2F1 showed no effect of these TFs when a *C* allele is present (Fig. 5A,B). In contrast, co-transfection of F2 with either E47 or E2F1 showed a significant increase in *OPTN* promoter activity, suggesting a potentiating effect of these TFs in gene transcription when the *T* allele is present (Fig. 5D). No significant effects were observed with SAP1 (Fig. 5E) nor with E2F2, E2F4 and E2F5 (Fig. 5D), while a positive effect on gene transcription was observed for both constructs with E2F3 (p-value < 0.0001) thus indicating that it was independent of the *C/T* variant present (Fig. 5A,D). In order to assess if this positive effect of rs3829923 *T* allele in *OPTN* promoter activity was related to an increase in *OPTN* expression in patients' cells, we performed a western blot analysis using osteoclasts derived from PDB patients 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 (Fig. 6). The overexpression of all transcription factors was assessed by qPCR (Fig. 5C,F).

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

To analyse the functionality of the two different alleles of RV-9906, we transfected HEK293 cells and performed luciferase assays with the two constructs, F3 and F4, carrying respectively the *G* or the *A* alleles. The transient transfections of HEK293 cells with constructs F3 showed a significantly lower activity than with F4 (Fig. 7A), indicating 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 (Fig. 7B,E), indicating that the effect obtained is specific for the *G* allele. After co-transfections of HEK293 cells with either 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) (Fig. 7C,F), which indicates that, in fact, RXR has a binding site in *OPTN* promoter but that it is independent of the G/A variant of RV -9906 present. The overexpression of all transcription factors was assessed by qPCR (Fig. 7D,G). 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 only one PDB patient who was not available for follow up.

4. DISCUSSION

In the present study, we investigated the functional role of three genetic variants within *OPTN*

locus found in the French-Canadian cohort. We detected 16 variants already reported in the NCBI database and we focussed our analysis in three variants: rs3829923 and rs2234968, discovered in our present work and a rare variant found in *OPTN* promoter – RV-9906 – discovered in a previous work from our group [3].

Regarding rs2234968 we found that it was associated with PDB in our population (uncorrected *p*-value = 6 x 10^{-3}) as well as in the Italian population (uncorrected *p*-value = 4 x 10^{-3}), which reinforces the importance and significance of this recently associated variant in PDB pathology. However, our results showed that it has no functional effect on PDB pathogenesis *per se*. Instead we found two SNPs (rs10906303 and rs79529484) in LD with rs2234968 that could explain the association of this variant to PDB. Then, we analysed the effect of rs10906303 and rs79529484 in the splicing pattern of *OPTN* gene and showed that these SNPs caused 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 truncated proteins predicted to lose *OPTN* protein domains, namely the NEMO, UBD and LIR domains which, similarly to the deleterious effect of *SQSMT1* mutations, that impair its main functional domain - UBA - [22], should impair *OPTN* function and thus contribute to the PDB phenotype. Accordingly, an *OPTN* transcript with a retained intron 5 was already described in patients with other diseases like juvenile open-angle glaucoma [27], thus confirming the presence of this abnormal transcript and indicating that it can alter OPTN function or the normal regulation of its protein expression [28]. Also, Obaid et al 2015 had already described a negative role for OPTN in osteoclastogenesis [29], which is in accordance with our results. Nevertheless, 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 analysing if OPTN function is impaired when only the mutant "short" transcript is

expressed, should be highly relevant to further understand its effect in associated pathologies such as PDB. In addition, analysis of the resulting autophagy process or of the expression of osteoclast related genes might also contribute to clarify the effect of these variants in PDB pathophysiology. The fact that this SNP was also found in some healthy controls also indicate that other factors such as, for example, the presence of paramyxovirus, may be required for the development of this disease, and this should also be further investigated. This co-effect of paramyxovirus and genetic mutations was already described by Kurihara's group, who found that p62 knockin mice bred with measles virus nucleocapsid gene (MVNP) mice develop greater numbers of pagetic-like osteoclasts and more dramatic bone lesions than do MVNP mice or p62 knockin mice alone. [30]

In this work we also analysed the effect of two variants found in *OPTN* promoter: rs3829923 – that was present in 13 patients and not in any of the controls tested - but that after the association study was found not to be associated with PDB (uncorrected *p*-value = 1); and RV -9906 - a rare variant found only in one PDB patient[3]. Although not being associated with PDB, since these variants were located in *OPTN* promoter we hypothesized that they could alter *OPTN* expression by affecting TFs binding 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 inductive effect only in the presence of the *T* allele. E47 is a member of bHLH family and it was described as important for osteoclast maturation and survival [31]. Therefore the gain of an E47 binding site could contribute to explain the onset of a 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 [32], and cell proliferation[33], and this may also contribute to explain the deregulation in the number of osteoblasts (and also osteoclasts) observed in this disease in the presence of the *T* allele. However, the

observed *in vitro* effect due to the presence of the rs3829923 *T* allele, was not reflected as an increase in OPTN protein levels in osteoclasts of patients carrying the rs3829923 *T* allele, which may explain why this variant was not significant in our association study.

A functional analysis of RV-9906 showed that, following co-transfections in HEK293 cells with the *OPTN* promoter carrying either the *G* or the *A* allele, together with SP1 and RXR TFs, we could conclude that RV-9906 *A* allele 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 [34] and SP1 has been described as an important regulator for osteoclast related genes, like RANKL [35] and BMP2 [36]. Our study suggests that SP1 is likely to be an important regulator of *OPTN* since a single alteration in its DNA binding site changed its activity, probably by preventing SP1 binding to the *OPTN* promoter. Unfortunaltely, the protein levels measured in the osteoclasts of the patient carrying RV-9906 were not assessed since the patient did not return for his follow up consultation.

To complement 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. These functional analysis will be a subject of future research. Nevertheless, the results obtained in this work already showed that SP1, E47 and E2F1 could be

new potential regulators of *OPTN* promoter activity, and therefore also new potential therapeutic targets for OPTN related diseases. Furthermore, since these are all bone associated transcription factors, results also suggest a potential role of OPTN in bone metabolism.

5. CONCLUSION

In conclusion, genetic variants of *OPTN* gene seem to be important in PDB patients. We identified two SNPs in *linkage disequilibrium* with rs2234968, a new variant associated with PDB, that are responsible for a change in splicing and possibly for creating a new truncated protein, probably non-functional, providing some insights about the role of this variant in the PDB phenotype of these patients. These data, together with the fact that *OPTN* seems to be regulated by bone related transcription factors such as E2F1, SP1 and E47, reinforce the relevance of this gene in the PDB pathophysiology.

ACKNOWLEDGMENTS

This research was supported in part by the European Regional Development Fund (ERDF) through the COMPETE - Operational Competitiveness Program and national funds through the Portuguese Science and Technology Foundation (FCT), under the project "PEst- CCMAR/Multi/04326/2013 and by Canadian Institutes for Health Research (MOP130457), Canada, Centre Hospitalier Universitaire de Québec (CHUQ) Foundation, Canadian Foundation for Innovation, Fonds de recherche du Québec-santé, Laval University and CHU de Québec Research Centre. The authors would also like to acknowledge the CHU de Québec Research Centre (CHUL) Gene Expression Platform, Quebec, Canada for the quantitative real-time PCR measurements and the Centre of Marine Sciences (CCMAR), Faro, Portugal for the use of the sequencing platform. IS, HC and NC were supported by doctoral and post-doctoral fellowships (Grant number: SFRH/BD/77227/2011, PD/BD/128341/2017 and SFRH/BPD//111898/2015, respectively) from the Portuguese Science and Technology Foundation (FCT). LM is supported by a career award from the Fonds de la recherche en santé du Québec (FRQS), Canada.

Conflict of Interest: The authors declare that they have no conflict of interest.

REFERENCES

1 Li Y, Kang J, Horwitz MS. 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* 1998; **18**: 1601–10.

2 Rezaie T, Child A, Hitchings R, Brice G, Miller L, Coca-Prados M *et al.* Adult-onset primary open-angle glaucoma caused by mutations in optineurin. *Science (80-)* 2002; **295**: 1077–1079.

3 Michou L, Conceição N, Morissette J, Gagnon E, Miltenberger-Miltenyi G, Siris ES *et al.* Genetic association study of UCMA/GRP and OPTN genes (PDB6 locus) with Paget's disease of bone. *Bone* 2012; **51**: 720–8.

4 Ying H, Yue BYJT. Cellular and molecular biology of optineurin. *Int Rev Cell Mol Biol* 2012; **294**: 223–258.

5 Ying H, Shen X, Park B, Yue BYJT. Posttranslational modifications, localization, and protein interactions of optineurin, the product of a glaucoma gene. *PLoS One* 2010; **5**: e9168.

6 Wild P, Farhan H, McEwan DG, Wagner S, Rogov V V, Brady NR *et al.* Phosphorylation of the autophagy receptor optineurin restricts Salmonella growth. *Science* 2011; **333**: 228–33.

7 Slowicka K, Vereecke L, van Loo G. Cellular functions of optineurin in health and disease. *Trends Immunol* 2016; **37**: 621–633.

8 Osawa T, Mizuno Y, Fujita Y, Takatama M, Nakazato Y, Okamoto K. Optineurin in neurodegenerative diseases. *Neuropathology* 2011; **31**: 569–74.

9 Maruyama H, Morino H, Ito H, Izumi Y, Kato H, Watanabe Y *et al.* Mutations of optineurin in amyotrophic lateral sclerosis. *Nature* 2010; **465**: 223–6.

10 Albagha OME, Visconti MR, Alonso N, Langston AL, Cundy T, Dargie R *et al.* Genome-wide association study identifies variants at CSF1, OPTN and TNFRSF11A as genetic risk factors for Paget's disease of bone. *Nat Genet* 2010; **42**: 520–4.

11 Kanis JA. *Pathophysiology and treatment of Paget's disease of bone*. 1998.

12 Altman RD. Epidemiology of Paget's disease of bone. *Clin Rev Bone Min Metab* 2002; **1**: 99–102.

13 Van Staa TP, Selby P, Leufkens HGM, Lyles K, Sprafka JM, Cooper C. Incidence and natural history of Paget's disease of bone in England and Wales. *J Bone Miner Res* 2002; **17**: 465–71.

14 Hocking L, Slee F, Haslam S., Cundy T, Nicholson G, van Hul W *et al.* Familial Paget's disease of bone: patterns of inheritance and frequency of linkage to chromosome 18q. *Bone* 2000; **26**: 577–580.

15 Laurin N, Brown JP, Lemainque a, Duchesne a, Huot D, Lacourcière Y *et al.* Paget disease of bone: mapping of two loci at 5q35-qter and 5q31. *Am J Hum Genet* 2001; **69**: 528–43.

16 Morales Piga AA, Rey Rey JS, Corres Gonzales J, Garcia Sagredo JM, Lopez Abente G. Frequency and characteristics of familial aggregation of Paget ' s disease of bone. *J Bone Miner Res* 1995; **10**: 663–670.

17 Fotino M, Haymovits A, Falk CT. Evidence for linkage between HLA and Paget's disease. *Transplant Proc* 1977; **9**: 1867–8.

18 Haslam SI, Van Hul W, Morales-Piga a, Balemans W, San-Millan JL, Nakatsuka K *et al.* Paget's disease of bone: evidence for a susceptibility locus on chromosome 18q and for genetic heterogeneity. *J Bone Miner Res* 1998; **13**: 911–7.

19 Good D, Busfield F, Fletcher BH, Duffy DL, Kesting JB, Andersen J *et al.* Linkage of Paget disease of bone to a novel region on human chromosome 18q23. *Am J Hum Genet* 2002; **70**: 517–525.

20 Hocking LJ, Herbert C a, Nicholls RK, Williams F, Bennett ST, Cundy T *et al.* 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* 2001; **69**: 1055–61.

21 Lucas GJA, Riches PL, Hocking LJ, Cundy T, Nicholson GC, Walsh JP *et al.* Identification of a major locus for Paget's disease on chromosome 10p13 in families of british descent. *J Bone Miner Res* 2008; **23**: 58–63.

22 Laurin N, Brown JP, Morissette J, Raymond V. Recurrent mutation of the gene encoding sequestosome 1 (SQSTM1/p62) in Paget disease of bone. *Am J Hum Genet* 2002; **70**: 1582–8.

23 Michou L, Morissette J, Gagnon ER, Marquis A, Dellabadia M, Brown JP *et al.* Novel SQSTM1 mutations in patients with Paget's disease of bone in an unrelated multiethnic American population. *Bone* 2011; **48**: 456–60.

24 Morissette J, Laurin N, Brown JP. Sequestosome 1: mutation frequencies, haplotypes, and phenotypes in familial Paget's disease of bone. *J Bone Miner Res* 2006; **21 Suppl 2**: P38–44.

25 Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2^(-ΔΔCT) method. *Methods* 2001; **25**: 402–408.

26 Pachot A, Blond J-L, Mougin B, Miossec P. Peptidylpropyl isomerase B (PPIB): a suitable reference gene for mRNA quantification in peripheral whole blood. *J Biotechnol* 2004; **114**: 121–4.

27 Willoughby CE, Chan LLY, Herd S, Billingsley G, Noordeh N, Levin A V *et al.* Defining the pathogenicity of optineurin in juvenile open-angle glaucoma. *Invest Ophthalmol Vis Sci* 2004; **45**: 3122–30.

28 Lewis BP, Green RE, Brenner SE. Evidence for the widespread coupling of alternative splicing and nonsense-mediated mRNA decay in humans. *Proc Natl Acad Sci U S A* 2003; **100**: 189–92.

29 Obaid R, Wani SE, Azfer A, Hurd T, Jones R, Cohen P *et al.* Optineurin negatively regulates osteoclast differentiation by modulating NF-κB and interferon signaling: implications for Paget's disease. *Cell Rep* 2015; : 1–7.

30 Kurihara N, Hiruma Y, Yamana K, Michou L, Rousseau C, Morissette J *et al.* Contributions of the measles virus nucleocapsid gene and the SQSTM1/p62P392L mutation to Paget's disease. *Cell Metab* 2011; **13**: 23–34.

31 Dey S, Curtis DJ, Jane SM, Brandt SJ. The TAL1/SCL transcription factor regulates cell cycle progression and proliferation in differentiating murine bone marrow monocyte precursors. *Mol Cell Biol* 2010; **30**: 2181–92.

32 Yu S, Yerges-Armstrong LM, Chu Y, Zmuda JM, Zhang Y. E2F1 effects on osteoblast differentiation and mineralization are mediated through up-regulation of frizzled-1. *Bone* 2013; **56**: 234–41.

33 Berman SD, Yuan TL, Miller ES, Lee EY, Caron A, Lees JA. The retinoblastoma protein tumor suppressor is important for appropriate osteoblast differentiation and bone development. *Mol Cancer Res* 2008; **6**: 1440–51.

34 Grant SF, Reid DM, Blake G, Herd R, Fogelman I, Ralston SH. Reduced bone density and osteoporosis associated with a polymorphic Sp1 binding site in the collagen type I alpha 1 gene. *Nat Genet* 1996; **14**: 203–205.

35 Liu J, Yang H, Liu W, Cao X, Feng X. 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* 2005; **96**: 716–27.

36 Xu J, Rogers MB. Modulation of Bone Morphogenetic Protein (BMP) 2 gene expression by Sp1 transcription factors. *Gene* 2007; **392**: 221–229.

TITLES AND LEGENDS TO FIGURES

Fig. 1 – Localization of the genetic variants found in *OPTN*

Schematic representation of the *OPTN* gene. The arrows indicate the transcription start sites and orientation in the chromosome. The light grey boxes represent non-coding exons, and the dark grey represent coding exons. The numbers below the boxes represent the exons length in base pairs. The bars represent the SNPs located in the introns and the arrowheads represent the SNPs located in the exons.

Fig. 2 – SNPs in Linkage Disequilibrium with rs2234968

(A) Localization of the SNPs in *linkage disequilibrium* with rs2234968. The numbers above the boxes (exons) represent the exons length in base pairs. The numbers below the lines (introns) represent the introns length in base pairs; (B) *In silico* prediction of the effects of the SNPs in *linkage disequilibrium* with rs2234968, in *OPTN* splicing; (C) Localization of the primers (F1, R1, F2, R2) used to amplify by PCR the different transcripts and electrophoresis results showing that with the presence of one mutated allele there is 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 retention 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 heterozygeous (*GA*) for rs2234968, which was also

heterozygous for the two intronic SNPs in LD. The results are relative to three independent western blot analyses for each sample.

Fig. 3 – *In silico* prediction of the effect of rs3829923 and RV -9906

(A) rs3829923 localization in *OPTN* gene (Chr10:13099144) 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 *T* allele. (C) RV -9906 localization in *OPTN* gene (Chr10:13099216) showing that this variant is present in *OPTN* promoter and it is responsible for a change from a *G* to an *A*. (D) 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 (Chr10:13099449) and TSS2 represents the second transcription start site (Chr10:13099959).

Fig. 4 - rs3829923 effect in *OPTN* basal promoter activity

Transient transfections of (A) HEK293 and (B) THP1 cells showing that the F2 construct (rs3829923 *T* allele) had a significantly higher activity than the construct F1 (rs3829923 *C* allele). Data are the mean ± SD of at least three independent experiments. Significance was determined by one-way Anova. Asterisk *** represents a p-value < 0.001. **** indicates that all the constructs used are significantly different from the promoter less vector pGL3-basic (p-value < 0.0001).

Fig. 5 - **rs3829923 effect in transcription factor binding**

(A) Co-transfection assays showing that E2F3 have an activating effect in F1, while (B) co-transfection with SAP1 revealed that this transcription factor does not have an effect in the presence of rs3829923 *C* allele. (C) qPCR results show that all of the transcription factors were overexpressed after co-transfection with F1 construct. Data are the mean ± SD of at least three independent experiments. Significance was determined by one-way Anova. Asterisk ** represents a *p*-value < 0.01. (D) Co-transfection assays showing that F2 construct (carrying the *T* allele) had a significant increase of luciferase expression, in the presence of E47 and E2F1 not observed with F1 construct (carrying the *C* allele), indicating that the effect of E2F1 and E47 is specific of the *T* allele. E2F2, E2F4 and E2F5 did not have a significant effect on both constructs. (E) Co-transfection with SAP1 revealed that this transcription factor does not have an effect in the presence of rs3829923 *T* allele. 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 that all the constructs used are significantly different from the promoter less vector pGL3-basic (*p*-value < 0.0001). (F) qPCR results show that all of the transcription factor swere overexpressed after co-transfection with F2 construct. Data are the mean ± SD of at least three independent experiments. Significance was determined by one-way Anova. Asterisk ** represents a *p*-value < 0.01.

Fig. 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

Fig. 7 - RV -9906 effect in *OPTN* basal promoter activity

(A) Transient transfections of HEK293 cells showing that the F4 construct (carrying RV -9906 *A* allele) had a significantly higher activity than the F3 construct (carrying the RV -9906 *G* allele). (B) 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 and (C) RXR. (D) qPCR results show that all of the transcription factors were overexpressed after co-transfection with F3 construct. Data are the mean ± SD of at least three independent experiments. Significance was determined by one-way Anova. Asterisk ** represents a *p*-value < 0.01. (E) Transient co-transfections of HEK293 cells showing that SP1 does not have any ihnibitory effect in the presence of the F4 construct (containing the *A* allele). (F) Co-transfections with RXR revealed that this transcription factor also had an inhibitory effect in the presence of F4. Data are mean ± SD of at least three independent experiments. Significance was determined by one-way Anova. Asterisk * indicates a *p*-value < 0.05, *** represents a *p*-value < 0.001, **** indicates that all the construct used are significantly different from the promoter less vector pGL3-basic (*p*-value < 0.0001). (G) qPCR results show that all of the transcription factors were overexpressed after co-transfection with F4 construct. Data are the mean ± SD of at least three independent experiments. Significance was determined by one-way Anova. Asterisk ** represents a *p*-value < 0.01.

Figure 1

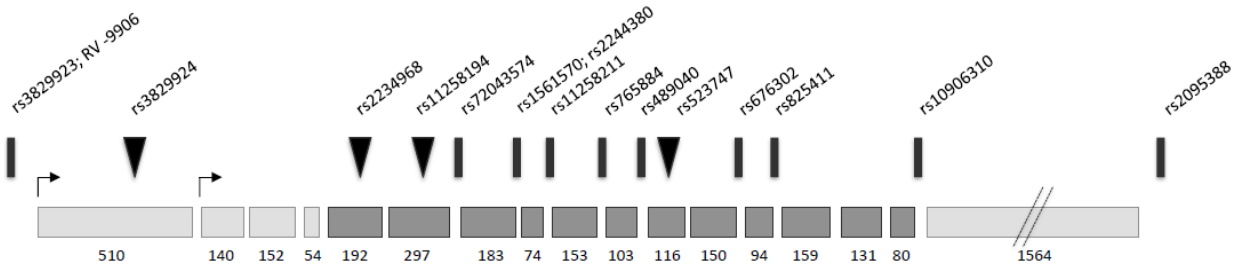
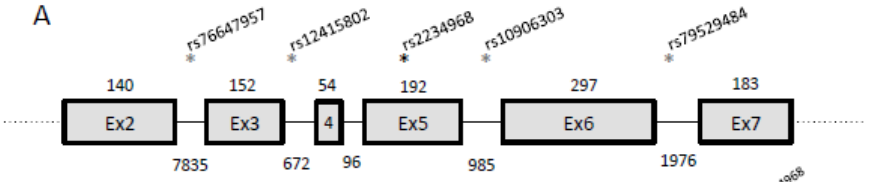


Figure 2

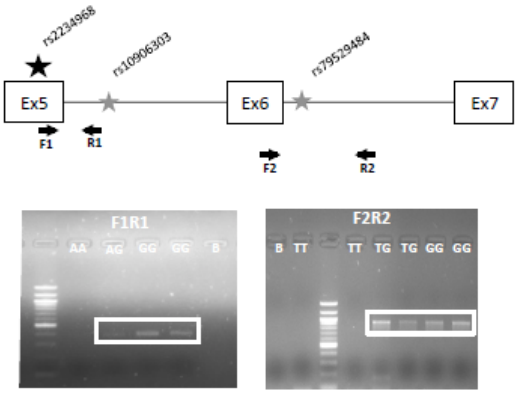
A



B

ID Variant	Splicing DONOR Site	Splicing ACCEPTOR Site
rs76647957	NO	NO
rs12415802	NO	NO
rs10906303	138%	80.9%
rs79529484	NO	0.1%, 3.2%, 2,4%, 6,2%

C



D

>WT
MSHQPLSCLTEKEDSPSESTGNGPPHLAHPNLDTFTP
EELLQQMKELLTENHQLKEAMKLNNQAMKGRFEELSA
WTEKQKEERQFFEIQSKEAKERLMALSHENEKLKEEL
GKLKGKSERSSEDPTDDSRLPRAEAEQEKDQLRT...
>rs10906303
MSHQPLSCLTEKEDSPSESTGNGPPHLAHPNLDTFTP
EELLQQMKELLTENHQLKGEQGWPLCAPFILGLQEMP
SLCAKAWW*
>rs79529484
MSHQPLSCLTEKEDSPSESTGNGPPHLAHPNLDTFTP
EELLQQMKELLTENHQLKEAMKLNNQAMKGRFEELSA
WTEKQKEERQFFEIQSKEAKERLMALSHENEKLKEEL
GKLKGKSERSSEVSRPIHCDVVFFLSLDICSGILRV*

E

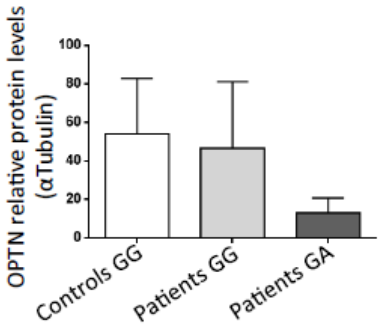
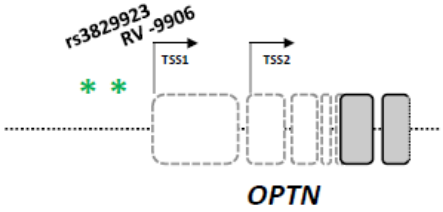


Figure 3



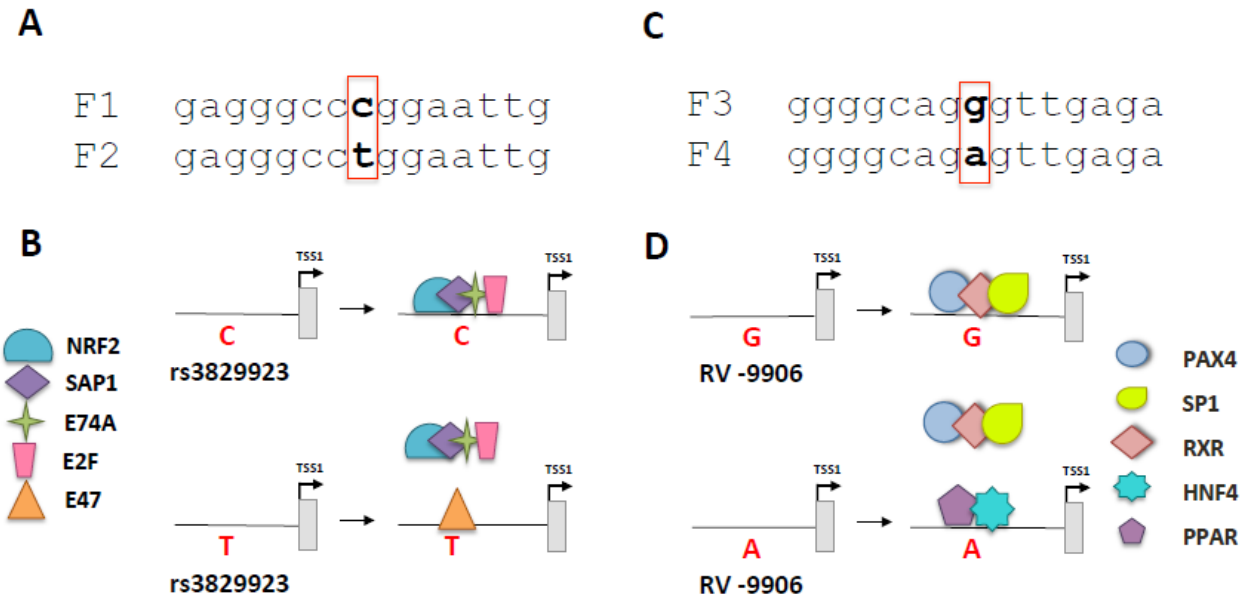


Figure 4

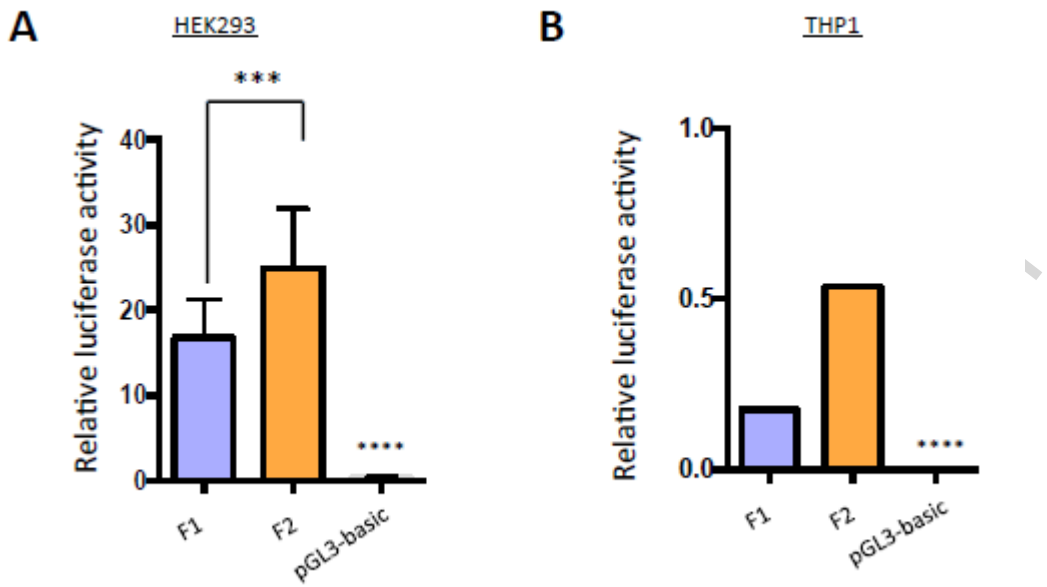


Figure 5

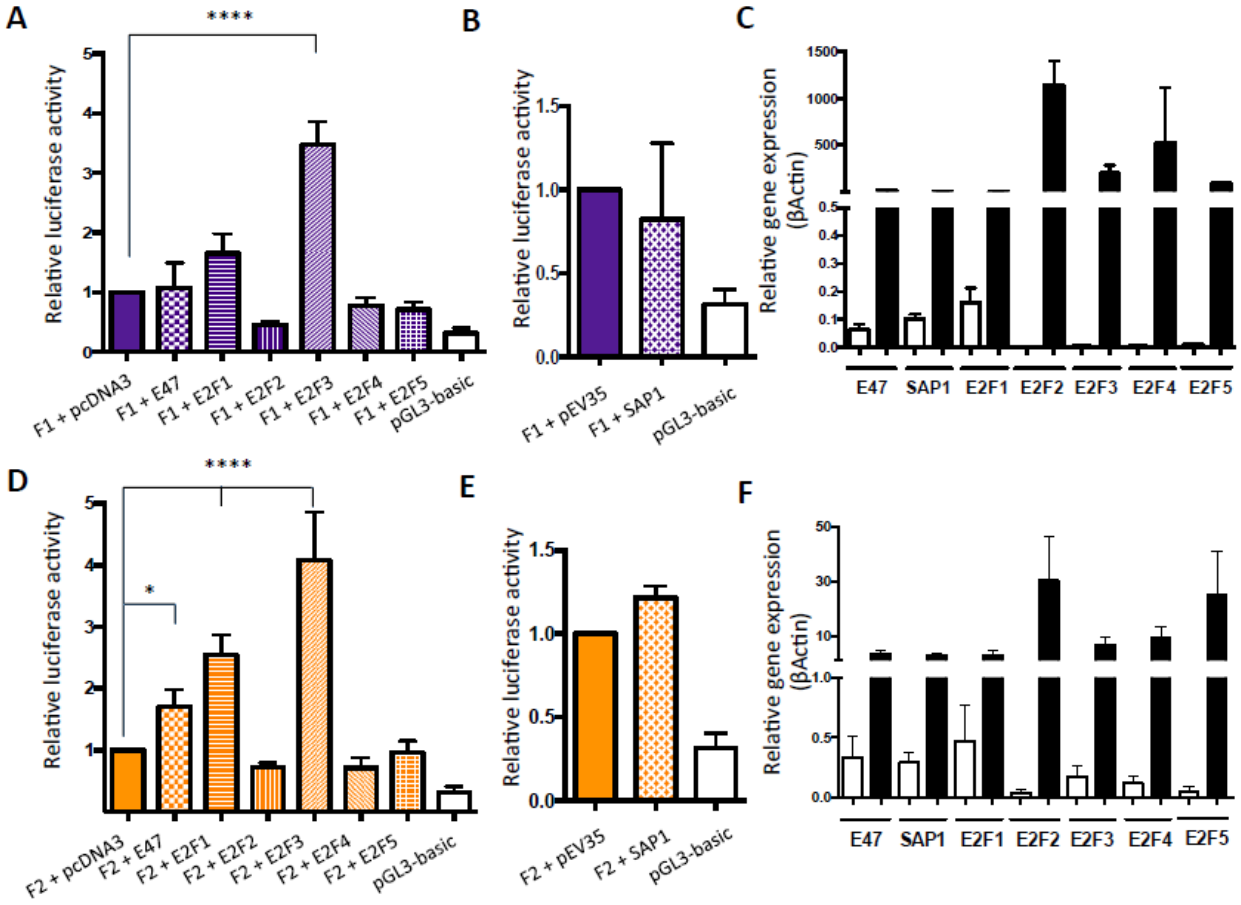


Figure 6

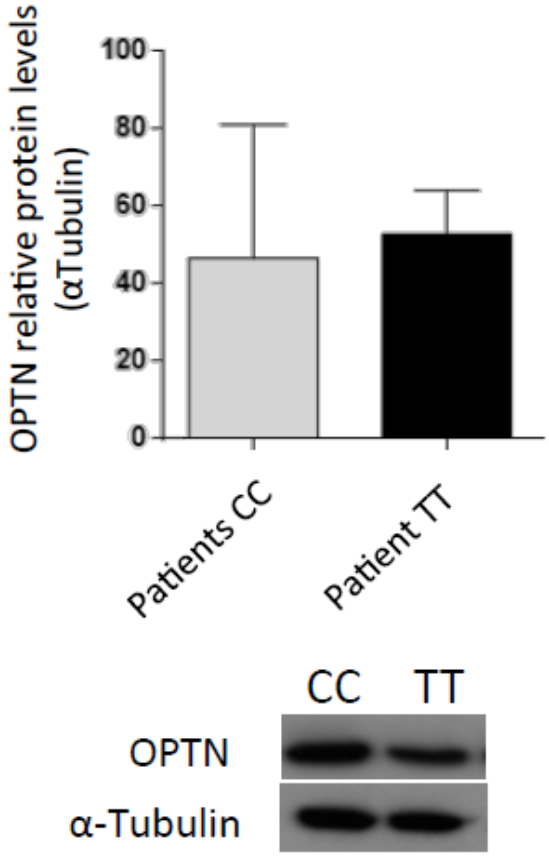
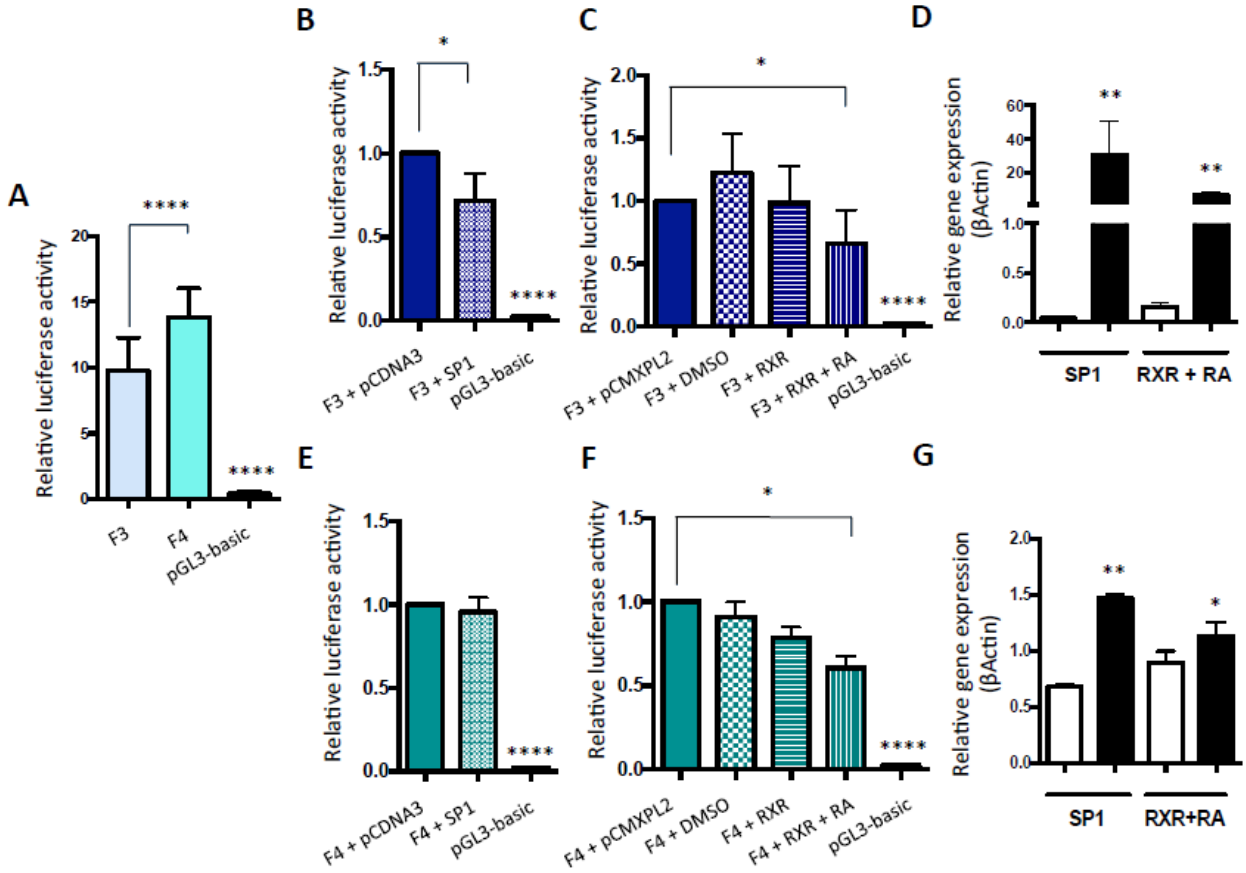


Figure 7



Highlights

- rs2234968 is a new variant associated with Paget's disease of bone.
- rs10906303 and rs79529484 are in linkage disequilibrium with rs2234968.
- rs10906303 and rs79529484 change OPTN splicing and might produce a truncated protein.
- *OPTN* promoter is regulated by bone related transcription factors.