UPPTEC X 12 017
Examensarbete 30 hp
Juni 2012

Discovery and verification of osteoporosis associated plasma proteins

David Efrem Solomon

UPPSALA UNIVERSITET

Molecular Biotechnology Programme

Uppsala University School of Engineering

UPTEC X 12 017	Date of issue 2012-07

Author

David Efrem Solomon

Title (English)

Discovery and verification of osteoporosis associated plasma proteins

Title (Swedish)

Abstract

Osteoporotic fractures are one of the leading causes for surgery worldwide, especially within the elderly female population. With an aging population, there is an increasing interest in preventive care to avoid higher health care burden. In order to achieve this plasma samples were taken from osteoporotic patients and antibody suspension bead arrays were utilized to discover potential biomarker candidates for osteoporosis. Screening 178 plasma blood samples against 152 protein targets resulted in 3 proteins that show indication of ability to separate between cases and control.

Keywords

Antibody suspension bead array, osteoporosis, post-menopause, fractures, biomarker, screening

Supervisors

Ulrika Qundos & Jochen Schwenk
Biobank Profiling, Science for Life Laboratory, Stockholm

Scientific reviewer

Caroline Kampf
Rudbeck Laboratory

Project name	Sponsors
Language **English**	Security
ISSN 1401-2138	Classification
Supplementary bibliographical information	Pages **25**

Biology Education Centre Biomedical Center Husargatan 3 Uppsala
Box 592 S-75124 Uppsala Tel +46 (0)18 4710000 Fax +46 (0)18 471 4687

Discovery and verification of Osteoporosis associated plasma proteins

DAVID EFREM SOLOMON

POPULÄRVETENSKAPLIG SAMMANFATTNING

Frakturer som orsakas av osteoporos (benskörhet) är en av de vanligaste anledningarna till kirurgiska ingrepp hos äldre kvinnor i världen. Sjukdomen kännetecknas av benvävnad som förändras och minskar i densitet vilket leder till ökad risk för frakturer och orsaken har kopplats till processer på cellulär nivå.

Då osteoporos, åtminstone initialt, inte har synliga symptom och minskningen av benvävnad är dold och inte märks förrän efter första frakturen, så finns det ett intresse av att upptäcka sjukdomen tidigare med andra metoder. Syftet med undersökningen som beskrivs i denna rapport var att försöka hitta en biologisk markör för att identifiera sjukdomen i ett tidigt skede. Detta gjordes genom att nivåerna av olika proteiner undersöktes i blodprov från både patienter som lider av osteoporos och patienter som anses friska, för att sedan försöka hitta en biologisk markör som skiljde proverna emellan.

Den metod (av flera möjliga) som användes för att genomföra undersökningen (Suspension Bead Array, SBA) tillåter till skillnad från många andra metoder undersökning av många prover mot många proteiner samtidigt, i storskala. I detta projekt har 178 blodprover undersökts med avseende på 90 olika proteiner och det har lett till att tre proteiner har hittats som i olika grader kan påvisa skillnad mellan sjuk och frisk.

EXAMENSARBETE, 30 HP
CIVILINGENJÖRSPROGRAMMET MOLEKYLÄR BIOTEKNIK
UPPSALA UNIVERSITET, 2012

ABSTRACT

Osteoporosis is a progressive skeletal disease that is characterized by reduced bone mass and disruption of bone architecture. The bone loss occurs stealthily, which makes it initially asymptomatic and thus requires a high index of suspicion for early detection. This results in Osteoporosis generally being diagnosed very late and often after the first fracture has occurred; signifying the clinical implication of Osteoporosis.

A selection of antibody suspension bead arrays along with antibody targeting proteins were selected from genome-wise association studies and literature searches. The selected arrays were then used to discover and validate potential plasma biomarkers in osteoporotic patients.

Two cohorts containing 178 plasma samples were screened using 152 antibodies targeting 90 unique proteins in total. From the results, two antibodies targeting two different proteins showed indications of separating disease samples from matched control samples in their respective cohorts. One of the proteins shows indication of functioning as an immune-regulator in mouse uteri and is regulated by estradiol. The protein contributes to increased levels of IFN-γ in estrogen deficient mice (simulated postmenopausal state); which is associated to bone loss.

Keywords: Antibody suspension bead array, osteoporosis, fractures, post-menopause, biomarker, screening

TABLE OF CONTENTS

ABBREVIATIONS

BMD	Bone Mineral Density
CV	Coefficient of Variation
DMSO	Dimethyl Sulphoxide
DXA	Dual Energy X-ray Absorptiometry
EDC	1-Ethyl-3-(3-dimethylaminopropyl)-carbodiimide
GWAS	Genome-wide Association Study
HPA	Human Protein Atlas
IFN-Y	Interferon-Gamma
MDC	Multi Disease Cohort
MES	2-[N-Morpholino]-ethanesulfonic acid
MFI	Median Fluorescence Intensity
MS	Mass Spectrometry
NHS	N-Hydroxysuccinimide
OSP	Osteoporosis
PBS	Phosphate Buffered Saline
PBS-T	Phosphate Buffered Saline + Tween20
PIVUS	Prospective Investigation of the Vasculature in Uppsala Seniors cohort
PQN	Probabilistic Quotient Normalization
SAPE	Streptavidin Conjugated R-phycoerythrin
SBA	Suspension Bead Array
SD	Standard Deviation
SMCC	Swedish Mammography Cohort Clinic

1 INTRODUCTION

The human anatomy demands specialized adaptions of the muscles and skeletons to allow rapid mobility and to avoid fractures, by being rigid enough during the early years of human life. At some point in life, especially women in postmenopausal period but also men of older age suffer reduced muscle functions and bones start to deteriorate and with these changes one can expect a risk of fractures, which is the significant clinical implication of Osteoporosis [1]. Using plasma samples from osteoporotic patients, the project utilized antibody suspension bead arrays for the analysis of those plasma samples for discovery and verification of potential biomarkers for osteoporosis.

1.2 OSTEOPOROSIS

Osteoporosis is a progressive skeletal disease that is characterized by reduced bone mass and disruption of bone architecture. Osteoporosis is a disease that was previously thought to be a natural consequence of aging. The general definition of Osteoporosis is based on the terms of "low energy" injury, together with fragility and age. A fall from standing height or any trauma that on a healthy individual would unlikely lead to a fracture is defined as a "low energy" injury [2]. Processes at cellular levels have been linked to the pathogenesis of Osteoporosis, with the smallest functional units being specialized bone cells responsible for bone-forming (osteoblasts) and bone resorption (osteoclasts). The remodeling of bone is a lifelong process that needs to be in balance to uphold plasma calcium homeostasis and prevent bone deficit and is achieved by osteoclasts digesting and osteoblasts forming bone matrix [3].

Earlier detection of Osteoporosis is made difficult by bone loss being an insidious and asymptomatic process and requiring a high amount of suspicion [4]. The repercussion of this is that Osteoporosis is diagnosed very late and often after the first clinical fracture and any patients undergoing therapies at that point of time of the disease progression, is usually done so any further fractures can be prevented. Therefore an earlier detection of Osteoporosis is needed in order to prevent the first fracture, so that a patient does not lose the quality of life because of loss of mobility and autonomy. Additionally there is the increase of morbidity among Osteoporotic patients and depending on where the fractures occur one can also see a correlation with higher mortality rates [5].

The disease which literally translates "*porous bone*" is due to its association with fragility fractures considered as a public health issue which affects people worldwide [6]. Osteoporotic fractures are expected to rise in the future because of an aging population, even though the total population isn't expected to grow markedly in Europe [7]. Monetary costs on health services are associated with the disease due to prolonged and permanent disability with the fractures making it most often impossible for the affected to live by themselves.

Osteoporosis is usually screened for via guidelines of evidence-based methods taking into account risk factors such as age, physical activity and previous fragility fractures, and Dual-energy X-ray Absorptiometry (DXA) in order to predict fracture risks [8, 9]. With imaging techniques, such as DXA, it is possible to measure bone mineral content and therefore assess

bone mass within scanned area of the bone. Due to area of scanned bone a certain level of variance is expected among individuals so bone mineral content is converted to bone mineral density. The score extracted from the imaging technique is an arbitrary value and depends on the measurement technique and skeletal area scanned and does not apply to all sexes or races [10].

Including many environmental factors, the common form of Osteoporosis is considered to be driven by a genetic disorder resulting from the combined action of several genes at once.

1.3 BIOLOGICAL MARKERS

The definition of a biological marker, or biomarker, is an objectively measureable biological feature of a cell, tissue or organism that indicates of biological or pathological process [11]. In medicine one is interested in disease biomarkers as they can identify a certain disease and its state so that a patient can be treated accordingly [12]. Generally, techniques to profile proteins or genes are used together with body fluids and other biological materials to identify biomarkers.

There are many places in the body where biomarkers for subclinical manifestation can occur but the investigation is reserved to body fluids as they make out the least invasive of the specimens that can be extracted. Serum and plasma have been and are important sources of a number of biomarker studies for diseases that are common, such as cancer [13], and among other things, this is due to their accessibility and the minimally invasive sample collection procedure. A growing consensus is that multiple biomarkers in panel will be needed rather than a single biomarker individually in most cases [14].

1.4 PROTEOMICS

In proteomics one tries to identify and understand the protein content of an organism by defining the proteins' abundance, structure, functions and interaction [15]. One major cause of pathologies differences at the cellular levels are altered proteins. To allow a diagnosis of certain diseases, it is necessary to identify which components lead to cellular malfunction.

Proteomic profiling as such is a large-scale analysis of protein expression, post-translational modifications for possible analysis of sample to divulge the protein level differences. A considerable amount of diagnostic tests in clinical laboratories target proteins which makes proteomics especially useful for disease diagnostics. [16]

1.5 PLASMA PROTEOME

In diagnostics one can investigate how normal states differ from disease states by looking at the plasma proteome. Blood circulation plays a very important part of humans and other mammals and it is essential transportation system between all the organs. In total the human plasma can make up about 7% of protein mass [17]. Plasma proteins are in general defined as proteins that function in the circulatory system. Blood plasma is the most accessible sample for medical diagnosis and therefore clinically very important for the ultimate goal of blood test in clinics. Serum and plasma is a better milieu for biomarker discovery as it is easier to obtain compared to tissue samples, but of course there are challenges associated with blood sample as well. Characterizing plasma sample is technically difficult because the abundance of plasma proteins can vary nine orders of magnitude. Serum albumin, which is after hemoglobin is the second most abundant protein in the circulation (~60 mg/ml), is clinically measured when indicating severe liver disease or malnutrition. And on the other end the low-abundant, interleukin-6 (~1 pg/ml), is measured as an indicator of different degrees of inflammation. Both of these proteins can be measured in hospital laboratories and are clinically useful indicators [18]. Considering the previous, one of the most important challenges in biomarker discovery is the discovery of a valid biomarker of a disease by reducing the normal background.

1.6 AFFINITY PROTEOMICS

There are a number of methods and approaches in the search of biomarkers by identifying compounds that can function as biomarkers. One strategy could be the comparison of the protein profile of healthy individuals with the profile of a defined disease state, as protein profiles can serve as powerful diagnostic markers [12].

Mass spectrometry (MS), which measures mass over charge of proteins and in majority their peptides, has had many forensic [19] and toxicological [20] applications during the years. It is an analytical technique well suited for proteome-wide analysis and biomarker discovery of cells but also fluid samples obtained from patients that are either disease-affected or healthy controls. MS-based discoveries have been preferred choice because of the high degree of specificity, but challenges affecting MS persist because of the samples being analyzed, e.g. body fluids, are so complex and therefore components aren't consistently assessed accurately and reproduced easily [12]. One other limitation with MS is the relative few samples that can be analyzed in a short amount of time.

Alternative methods to MS-based discoveries are protein microarrays, found within affinity proteomics, which rely on affinity reagents such as antibodies as capture reagents. The two microarrays technologies of today, planar microarrays and bead-based arrays, are well-suited methods for the screening of a large number of target proteins as well as many patient samples. Both methods have multiplexing possibilities with a number of capture spots, with planar array in the thousands and with bead-based arrays limited by the number of detected bead-identities which currently has been extended to a maximum of 1728 bead identities [21].

Bead-based arrays are suitable instead of planar arrays if antibodies are preferable to be used as capture reagents and if the number of parameters being of interest is fewer. The beads are based on color coded microspheres which can be coated with capture reagents, such as antibodies. The microsphere are analyzed using flow cytometry like instrument with a dual laser system to identify each individual bead-identity as well as fluorophore labeled and captured protein targets (Luminex Corp.).

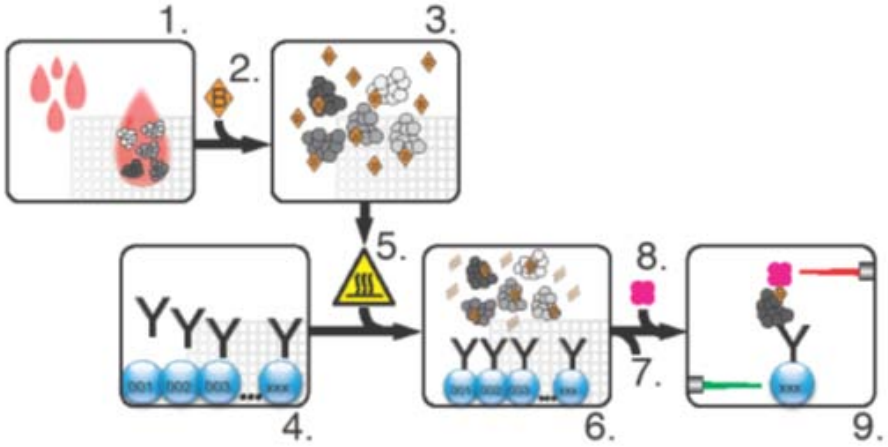

Figure 1. Antibody Suspension [31] bead array workflow, in which patient blood sample is distributed to plate in randomized positions (**1**) and activated biotin is added (**2**) to modify protein content and label samples (**3**). In addition antibodies targeting proteins of interest are coupled to beads with specific bead identities (**4**). Labeled samples are heat-treated to open up epitopes for interaction with antibodies (**5**). Antibody-coupled beads are incubated with labeled samples (**6**) and proteins that have not bound to antibodies are discarded after a wash step (**7**). Fluorescent-labeled streptavidin is added (**8**) to bind with the labeled samples. Using measurement instrument the specific bead-ID and relative antibody intensities can be retrieved.

[Figure 1. Image used with permission from Schwenk et al, MCP 2010]

1.7 HUMAN PROTEIN ATLAS

Limitations of proteomic-based diagnostic tools, which rely on affinity-based analysis, are that validated protein-specific affinity reagents need to be available for targeting any protein of interest. The Human Protein Atlas (HPA) is one of several projects aiming to overcome the lack of validated capture reagents by systematically generating polyclonal antibodies towards all non-redundant protein encoding human genes [22]. The efforts of HPA have resulted in a publicly available database (www.proteinatlas.org) containing already over 15000 antibodies toward human proteins (HPA version 9).

1.8 FROM DISCOVERY TO VERIFICATION OF BIOMARKERS

Going from a discovered biomarker to validated diagnostic marker entails many challenges, such as verification of a finding in additional independent sample collections [23] as wells as with orthogonal methods.

2 MATERIALS AND METHODS

2.1 SAMPLE COLLECTION AND STUDY

The first verification cohort (denoted OSP2) contains 96 samples from equal amount of patients retrieved from the Prospective Investigation of the Vasculature in Uppsala Seniors cohort (PIVUS) collected from men and women in their 70s living in the city of Uppsala, Sweden. The patients further provided medical and drug history and leisure-time physical activity and also underwent dual energy x-ray absorptiometry (DXA, Lunar Prodigy, Lunar corp, Madison, WI, USA). The measured BMD (g/cm^2) of the femoral neck, total body and hip was determined using DXA [24].

The second verification cohort (denoted OSP3) contains 96 samples from just as many patients retrieved from SMCC (Swedish Mammography Cohort Clinic) with the secondary outcome of osteoporosis. The samples were collected between November 2003 and October 2009 from patients living in the city of Uppsala. Beside height and weight measurements, the patients provided blood and urine samples and have also undergone DXA scans. Blood samples provided were control and case samples with case being defined as Osteoporotic patients with T-score at total hip, femoral neck, or spine of <= 2.5 standard deviations (SD) below the mean of a young adult reference range [25].

Clinical information provided with the samples was limited to disease state and age of the patients (see Table 1).

Table 1. Age table for OSP2 and OSP3 sample collection.

	OSP2 (verification cohort 1)		**OSP3 (verification cohort 2)**	
	case	**control**	**case**	**control**
Quantity (patients)	42	47	41	48
Average age, years (min-max)	64,5 (60-69)	63,6 (60-69)	72,5 (71-74)	72,2 (71-74)

2.2 ANTIBODY SELECTION

A total of 177 validated affinity-purified polyclonal antibodies provided by the Human Protein Atlas were used as reagent for to target list of 90 proteins. The target list was retrieved partly from Genome-wide association study performed by the clinical collaborator, Karl Michaelsson (Uppsala University), and the other part was generated by an antibody-centric high-throughput analysis of osteoporosis sample which was a part of a multi disease cohort (MDC), which was performed in the Biobank profiling group at Science for Life Laboratory.

The antibodies were divided and coupled to two 80-plex SBAs including 4 controls for each SBA (MDC-SBA and GWAS-SBA), which resulted in each SBA containing 76 antibodies each with a total of 152 antibodies (see Table 2). Exactly 25 antibodies were therefore excluded from the original antibody list and the remaining antibodies were divided equally to two 80-plex SBAs.

152 out of a total of 177 antibodies targeting 90 unique ENSEMBLE gene-IDs were distributed as follows:

Table 2. Summary of selection strategy for and division of the 177 antibodies targeting 90 unique proteins into two SBAs. SBA; suspension bead arrays.

	MDC-SBA	**GWAS-SBA**
Origin	76 Abs from MDC candidate list	49 Abs from GWAS-list 27 Abs from the MDC list with highest concentration
Controls (anti-HSA, rabbit IgG, anti-human IgG and empty bead)	4	4

2.3 SUSPENSION BEAD ARRAY GENERATION

Antibodies were coupled to beads using 30 µl ($5x10^5$ magnetic beads) per bead-ID (MagPlex microspheres, Luminex Corp.), which were then distributed to a 96-well plate (Greiner BioOne). The beads were then washed and let in solution with activation buffer; 0.1 M, pH 6.0-6.2 NaH_2PO4 (Sodium phosphate monobasic, M=137.99 g/mol), Merck.

The beads were activated using activation solution with 0.5 mg N-hydroxysuccinimide (Pierce) and 0.5 mg 1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide (Pierce) in 100 µl activation buffer. The beads were exposed to a 20 min activation time and subsequent washing in 0.05M 2-[N-Morpholino]-ethanesulfonic acid, pH 5.0. Thereafter 1.6 µg of antibody in 100 µl MES-buffer was added together with the beads and incubated for 2 hours in room temperature (650 rpm rotation in dark). The beads were then washed twice in PBST (PBS and 0.05% Tween20, VWR, Haasrode, Belgium) after the two hour incubation and stored in storage buffer (Blocking Reagent for ELISA, Roche) to quench remaining bead activity together with Proclin (Supelco, Bellefonte, PA) to prevent bacterial growth.

The beads are then pooled to create suspension bead arrays (SBAs) by combining an equal amount of volume and sonicating the beads (Branson Ultrasonic Corp.) in plate for 4 min and 40s. The beads were then dispensed to a bead collector and volume adjusted and subsequently transferred to a tube for storage at 4°C.

To evaluate the coupling efficiency of the antibody coupling of beads a coupling test was performed using anti-rabbit IgG conjugated to R-phycoerythrin (Jackson Immunoresearch).

2.4 PRE-ANALYTICAL SAMPLE PREPARATION

Crude plasma samples were randomly distributed in a 96-well plate together with 6 aliquot replicates of a sample pool (pooled samples, n=24) and one sample blank. The two verification cohorts were handled separately.

2.5 SAMPLES LABELING

Labeling is achieved by using biotin as a tag for detection and were performed by first thawing the frozen plasma samples in 4°C and then spun down (2000 rpm) for 2 minutes. Plasma samples were then diluted 1:10 in phosphate buffered saline (PBS) and labeled with 2 mg NHS-PEG4-biotin (Pierce) dissolved in DMSO (10 mg/ml, Dimethyl Sulphoxide, VWR) per 3 µl sample and then stored at 4°C for 2 hours. The labeling plate was vortexed every 30 minutes during this time and after the 2 hours 12.8 µl Tris-HCl (0.5M, pH 8.0, Sigma) was added to the labeling plate to quench the reaction and placed at 4°C for 30 minutes. The labeled samples were placed at -20°C for longer storage.

2.6 ASSAY PROCEDURE

The 1:10-diluted biotinylated samples were thawed in room temperature and subsequently diluted 1:50 in assay buffer (PVX Casein buffer; 0.1% casein (Sigma # c5890-500G), 0.5% PVA (polyvinyl alcohol, Sigma # 136-250G) and 0.8% PVP (polyvinylpyrrilidone, Sigma # 360-500G) and 10% (v/v) rabbit IgG (Bethyl)) making a total of 1:500 sample dilution. The samples were heat-treated at 56°C for 30 min and 23°C for 15 min and finally 45µl sample was transferred to a 96-well assay plate (Greiner BioOne, flat bottom half-area well plate) containing 5µl bead mixture and incubated overnight (16 hours) on a shaker (650 rpm).

The beads were washed 3x100 µl in PBST using a plate washer (Biotek EL406) and incubated in 50 µl paraformaldehyde solution (0.4%) in PBST for 10 min. Next, the beads were washed prior addition of 50 µl Streptavidin Conjugated R-phycoerythrin (0.5 µg/ml, Invitrogen) for 20 min. The beads were washed and measured in 100 µl PBS-T with LX-200 instrument (Luminex Corp.), at a minimum of 50 events per bead ID and well. Data points were summarized as median fluorescence intensity (MFI) in arbitrary units (AU) per bead ID.

The assay was repeated for both sample sets with both bead arrays (MDC-SBA and GWAS-SBA).

2.7 STATISTICAL ANALYSIS

R, which is an open source software environment for statistical computing and graphics [26], was used for data handling, processing and visualization. Data processing includes outlier removal by PCA and normalization by Probabilistic Quotient Normalization [27].

The univariate Wilcoxon rank sum test was applied to identify antibodies able to discriminate between cases and controls (p-value < 0.05). Antibodies were classified as candidates with a p-value < 0.05 in both replicate experiments.

Inter assay variation was assessed by correlation (Pearson's) of both sample and antibody profiles of data sets from replicate experiments and intra assay variation by the calculation of the coefficient of variation per antibody.

3 RESULTS

3.1 SUSPENSION BEAD ARRAY

Coupling test was performed to confirm the coupling efficiency of antibodies with respective bead-IDs, which resulted in intensity levels for each antibody that could be compared with the other antibodies within a SBA. An antibody was flagged if it was below half of the average intensity for all antibodies. MDC-SBA and GWAS-SBA both had 8 antibodies each that did not fit the arbitrary coupling criteria and which were flagged for the subsequent final analysis.

3.2 TECHNICAL AND BIOLOGICAL VARIANCE

A statistical measurement of the dispersion of antibody intensity data points was performed by calculating the average coefficient variance (CV) of the antibody intensities among the replicate samples. Using the averages of the CVs and then subsequently calculating the mean of those CVs for all antibodies make up the variance within the samples. This final average of the average coefficient of variation for all antibodies within the replicate samples make up the technical CV, as can be seen in table 3.

In order to validate the data, the technical CV is compared with the biological CV, which is the average CV of the antibody intensities among non-replicate samples (see Table 3). For the use of the sample data in subsequent analysis of separation of disease and control; the variance between different samples should be much higher than variance within replicate samples

Table 3. Technical and biological CV of all normalized data sets.

SBA	Cohort	Measurement	Technical CV (%)	Biological CV (%)
MDC	OSP2	1	12.7	43.5
		2	15.0	36.8
	OSP3	1	13.8	45.7
		2	16.7	44.3
GWAS	OSP2	1	8.7	33.3
		2	-	-
	OSP3	1	9.6	41.2
		2	-	-

As seen in table 3 above, it is observed that the technical CV is consistently below 25%, and a biological variance that is always at least twice the size of the technical variance. The technical CV for each experiment on the different data sets range from 8.7-16.7 % for normalized data sets and 14-18% for the raw data.

3.3 CORRELATION IN REPEATED MEASUREMENTS

Antibodies in MDC-SBA between the two measurements each of OSP2 and OSP3 sample collections were correlated in order to compare the similarities of the data sets and their repeatability. Out of the 79 antibodies which make up MDC-SBA as many as 11 antibodies between the OSP3 experiments showed less than 0.5 units in correlation (see fig. 2).

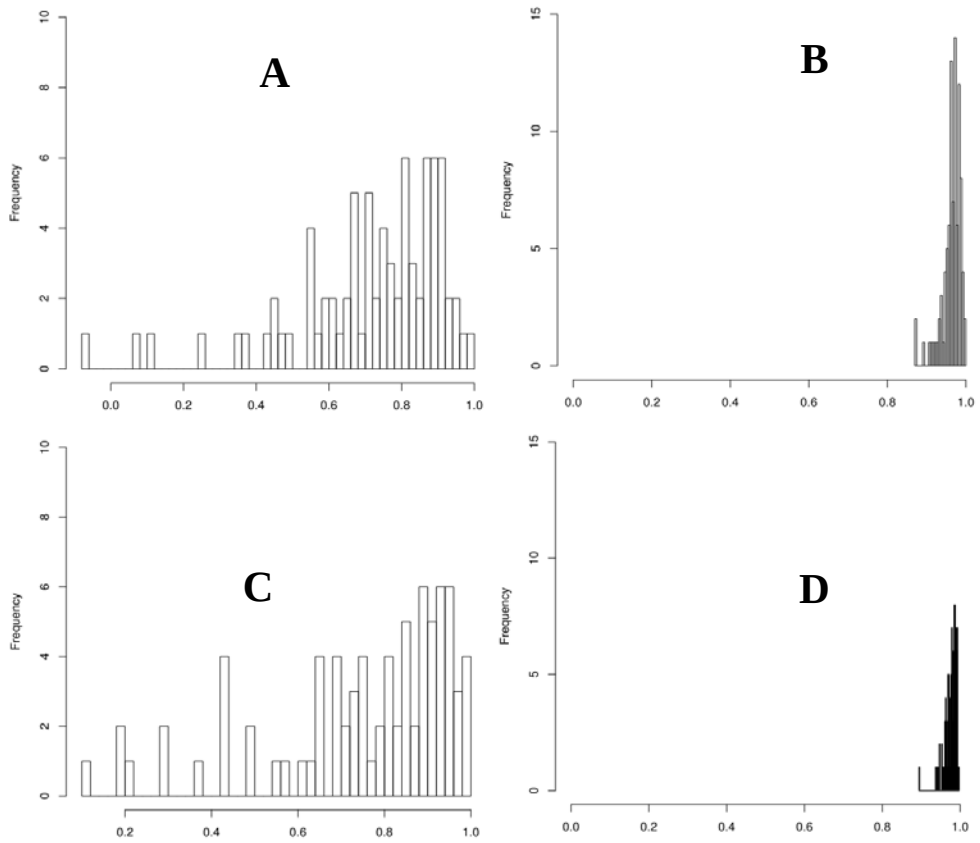

Figure 2. Histogram of antibody intensity correlation, between repeated experiments of OSP2 (**A**) and the patient sample correlation (**B**) within the same cohort experiments. In addition the antibody intensity (**C**) and patient sample (**D**) correlation for the repeated experiments of OSP3.

3.4 ANTIBODY LIST WITH SIGNIFICANT P-VALUES FOR MDC-SBA

The following list of antibodies (see Table 4) have after a Wilcoxon rank sum test on two different cohorts and two different SBAs returned a significant p-value (threshold set at < 0.05) and therefore shown to be able to separate between disease and control. The antibodies that show up in the list after repeated measurements have been classified as interesting, which is the case for HPA032040 (OSP2) and HPA038344 (OSP3) on the two verification cohorts. Antibodies that did not have significance difference between disease and control samples after repeated measurements were classified as false positive results and therefore discarded.

Table 4. Table of antibodies with listed significant p-values with threshold set below 0.05.

Antibody	**MDC-SBA**			
	OSP21	**OSP22**	**OSP31**	**OSP32**
HPA032040	0.0018	0.00048		
HPA026845	0.0200			
HPA028220	0.0290			
HPA038275	0.0320			
HPA037360	0.0390			
HPA037486	0.0400			
HPA029019	0.0470			
HPA028671		0.0025		
HPA026991		0.0040		
HPA027289		0.0140		
HPA031719		0.0270		
HPA027801		0.0280		
HPA038013		0.0480		
HPA038344			0.0017	0.0028
HPA027802			0.0046	
HPA047562			0.0084	
HPA031720				0.0160
HPA027863				0.0460

For table 4 above, the set threshold of p-value 0.05 equals to significant separation between disease and control samples. All antibodies are from the MDC-list, with HPA032040 and HPA038344 (bolded) having significant p-values in repeated measurements in respective cohorts.

3.5 SEPARATION OF DISEASE AND CONTROL WITH MDC-SBA

P-value retrieved from Wilcoxon rank sum test analysis resulted in, HPA38344, showing a separation between disease and control with significant p-value in OSP3 using MDC-SBA (see fig. 3). The results were consistent between repeated measurements and the boxplots showing separation of disease state and healthy state show the same trend.

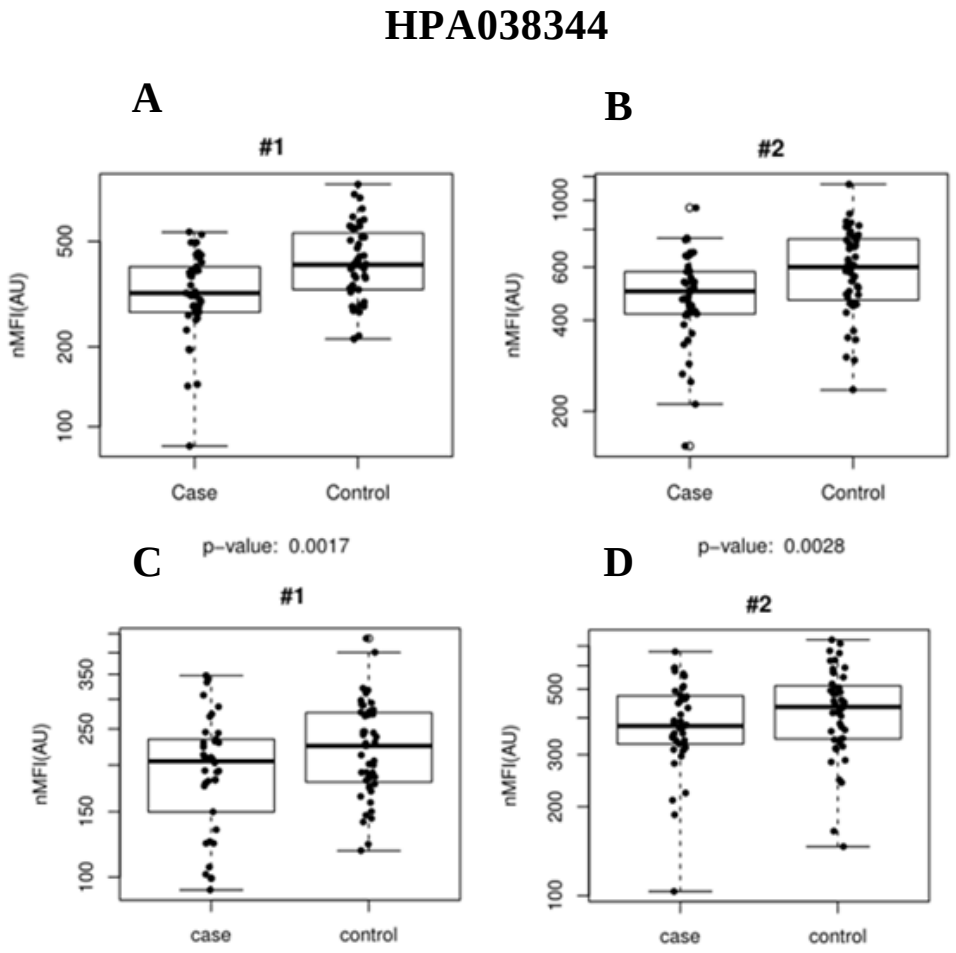

Figure 3. Boxplots of separation of disease and control for repeated measurement of HPA038344 for two verification cohorts. All boxplots show the same trend with the separation for OSP3 being significant on first (**A**) and repeated (**B**) measurements. The retrieved p-value from Wilcoxon rank sum test for OSP2 was not significant but followed the same trend as OSP3 in the first measurement (**C**) and repeated measurement (**D**).

HPA038344, developed similar trends as OSP3 results and significant p-value with OSP2 sample collection using MDC-SBA.

A second antibody, HPA032040, also shows separation between disease and control with significant p-value in first and repeated measurements of OSP2 with MDC-SBA (see fig. 4). The antibody did not show significance in OSP3 and neither did the boxplots reveal a general trend that could be compared with fig. 4.

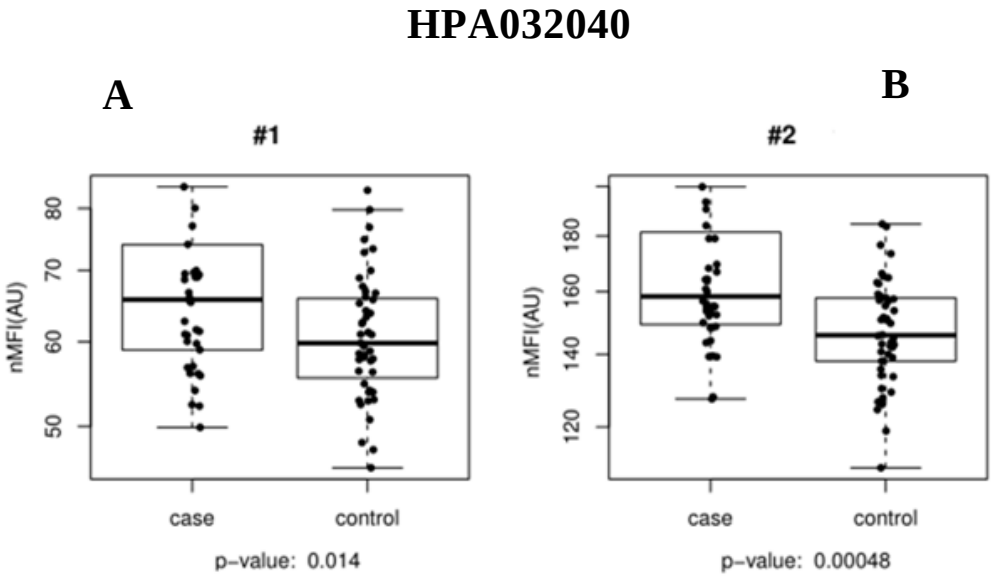

Figure 4. Boxplots of separation of disease and control samples for two measurements of HPA032040 with OSP2. The p-value was significant for the first measurement (**A**) as well as the repeated measurement (**B**) and in addition the trends are similar within this cohort.

3.6 ANTIBODY LIST WITH SIGNIFICANT P-VALUES FOR MDC-SBA

Antibodies listed below (see Table 5) returned after a Wilcoxon rank sum test on different cohorts and different SBAs a significant p-value (threshold set at < 0.05) and therefore shown to be of interest.

Table 5. Table of antibodies with listed significant p-values below the threshold of 0.05

Antibody	**List**	**GWAS-SBA**	
		OSP21	**OSP31**
HPA000450	GWAS	0.012	
HPA027420	MDC	0.03	0.0092
HPA037598	GWAS	0.0056	
HPA024371	GWAS		0.006

From the target list of 4 antibodies targeting respective proteins; one returned a significant p-value on both the OSP2 and OSP3 measurements. The target is supported by re-occurrence in two sample collections but having not been repeated the results are not fully valid following the previous criteria. HPA027420 has significant p-value in two different cohorts but the measurements have not been repeated.

3.7 SEPARATION OF DISEASE AND CONTROL WITH GWAS-SBA

HPA027420, is an antibody that shows a separation between disease and control samples in two experiments; OSP2 and OSP3 samples collections (see fig. 5).

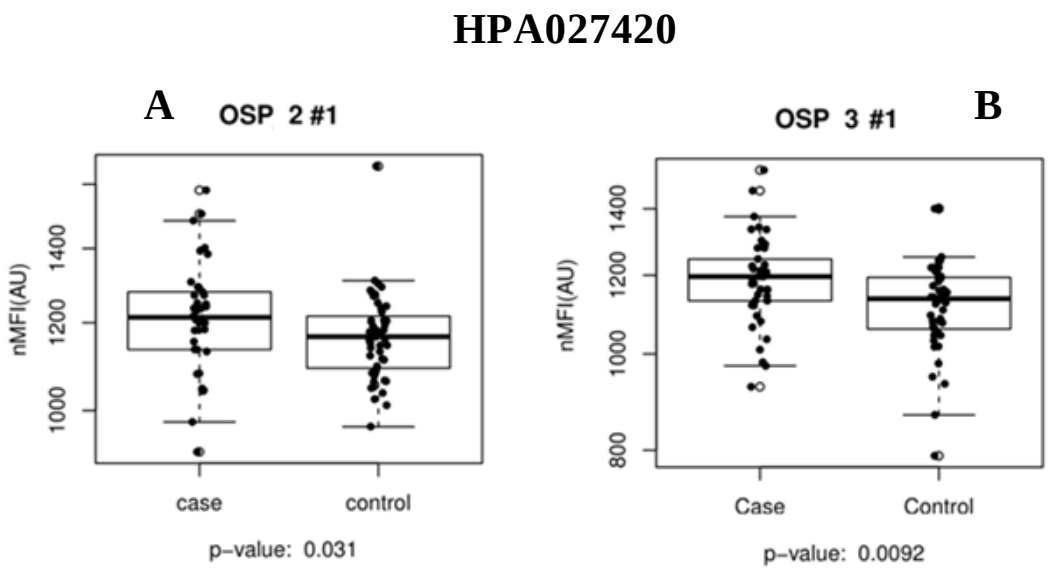

Figure 5. Boxplots of separation of disease against control samples with GAWS-SBA. HPA027420 has significant p-value in both verification cohorts OSP2 (**A**) and OSP3 (**B**) cohort after first experiment.

4 DISCUSSION

4.1 SUMMARY

This project was performed using the antibody suspension bead array method and a collection of antibodies targeting selected proteins of interest. The chosen antibodies were then used to screen samples from osteoporotic and control patients for possible biological markers. Antibodies targeting proteins of interest were selected from a genome-wide association study target list and a discovery phase experiment. The antibodies selected from the discovery phase experiment were previously performed with a smaller sample set related to one of the two cohorts. The sample screening resulted in three protein targets with statistically significant p-values which were either supported by repeated measurements (two targets) or observed in both cohorts (one target).

4.2 DATA ACCURACY

The method used to determine that an assay was performed correctly was by looking at the CV. The desirable quality specification for technical CV is below 25% (with a biological variance that is higher than the technical variance). From the pooled aliquot replicates, the expectation for the technical CV is a non-significant variance in antibody intensities between those samples. Looking at the results, it is fair to assume that replicated samples for each experiment have been very similar and overall consistent (Table 2). The biological variance is consistently at least twice the value of the technical variance, which means we can differentiate between samples.

A correlation was done between the experiments within cohorts to test for data reproducibility and reliability for subsequent data analysis. The repeatability was found to be better on OSP3 over OSP2 if one chooses 0.6 as an arbitrary unit to compare antibody correlation. One aspect in which differences between the experiments can be affected is through heat-treatment; which can show a difference if the temperature were off by a few degrees or the treatment time differed between the experiments. As a result, this would cause a varying degree of interactions among the antibodies and the respective targeted proteins. Another factor that could affect the degree of correlation between repeated experiments is the fluorescent labeled protein SAPE. This is added to bind to biotinylated samples; which could affect the intensity levels between the experiments. Unequal distribution of SAPE could lead to inconsistent results and would be reflected on the lack of similarities between repeated measurements. The patient sample correlation shows very high similarities which is expected since the antibody intensities in each patient sample covers all intensity ranges due to the number of antibodies within each sample.

CANDIDATES

To confirm the hits retrieved, the measurement of the samples was repeated using MDC-SBA. There were 16 hits that were separate cases from the controls that were not reproducible with repeated measurements, and were therefore classified as false positive discoveries. A hit that has been confirmed by repeated measurements strengthens the reliability of the finding, especially if the hits show up between cohorts that differ noticeably. However, this could also be the case of a hit that is age dependent because of the age differences between cohorts.

The two cohorts, OSP2 and OSP3, have an average age that differs by 8.4 years, and also OSP2 has a range that is much broader than OSP3. Though the samples are more similar in age, it would be interesting to look at verifications within a similar age range, since the cohorts differ substantially in that aspect. Age is an important factor when it comes to the different types of osteoporosis: primary type I (postmenopausal) typically occur with age, the age range with type II (senile osteoporosis) patients starts at age 75. Since the verification sample cohorts vary greatly by age, overlapping targets between the cohorts may prove to be substantial; however, perhaps not as expected due to the age difference between samples. This was the case for HPA027420 which targets a protein that separates disease from control samples in both cohorts. This finding will be verified via repeated measurements.

Furthermore, there might be confounding factors not yet identified because of the limited extent of clinical information available at the time of this study.

HPA038344 targets an amino acid sequence that is homologous to both ZNF511 and PRAP1. PRAP1 is a secretory proline-rich protein regulated by 17β-estradiol, and was previously thought of as a pregnancy specific protein as it was overexpressed in the uterus of early and late stage pregnant women [28]. This protein is a sex hormone that is present in both females and males, with males having comparable serum levels to postmenopausal women [29].
In a study of pregnant mice with ovaries removed, to simulate menopause, it was found that the lack of PRAP1 in uteri (which functions as an immunoregulator) increases the expression of Th1 IFN-γ. This goes hand in hand with mice suffering from estrogen deficiency (menopausal), displaying higher levels of IFN-γ which results in greater bone loss [30].
A point which could create a fallacy with the previous reasoning is that the mice which were used on the testing of simulated estrogen deficiency, by removal of ovaries, were only 8-weeks old. For the use of mice as animal models in osteoporosis it is recommended to study them from peak bone mass, which is at the age 4-8 months old [31]. This makes it hard to conclude that IFN-γ is the reason the control samples have close or significantly higher PRAP1 protein compared to the osteoporotic group in the two cohorts after repeated measurements.
Literature search on ZNF511, which is involved in transcriptional regulation, led to no obvious relations to osteoporosis.

CONCLUSION

In conclusion, two antibodies, one for each of the profiled sample cohorts were separated from the disease and control samples, where one of the antibodies (HPA038344)were either in proximity or below the p-value (0.05) in every measurement of both cohorts. This makes the proteins (PRAP1 and ZNF511), which the antibodies target interesting for further research which could be done looking at sibling antibodies and comparing with existing results.

NEXT STEP

Since HPA038344 targets two proteins there is an interest in target identification by assessing the molecular mass detected by this antibody using Western blot. In addition, repeated measurements are required for the plasma samples using GWAS-SBA since only one measurement has been performed. Furthermore, it would be interesting to examine if there would be supportive results using a sandwich assay method with samples from other cohorts than the ones used in this project.

5 REFERENCES

1. Pietschmann P, Rauner M, Sipos W, Kerschan-Schindl K. (2009) *Osteoporosis: an age-related and gender-specific disease--a mini-review.* Gerontology. 2009;55(1):3-12. Epub 2008 Oct 24.

2. Lencel P and Magne D (2011) *Inflammaging: The driving force in osteoporosis?* Medical Hypotheses 76:317–321

3. Hadjidakis DJ and Androulakis II. (2006) *Bone remodeling*. Annals of the New York Academy of Sciences. 2006 Dec;1092:385-96.

4. Unnanuntana A, Gladnick BP, Donnelly E, Lane JM (2010) *The assessment of fracture risk.* The Journal of Bone and Joint Surgery. American Volume, 92, pp. 743–753

5. Center JR, Nguyen TV, Schneider D, Sambrook PN, Eisman JA. (1999) *Mortality after all major types of osteoporotic fracture in men and women: an observational study*. Lancet, 353 , pp. 878–882

6. Reginster JY and Burlet N. (2006) *Osteoporosis: a still increasing prevalence.* Bone, 38 (2 Suppl 1), pp. S4–S9

7. Kanis JA and Johnell O. (2004) *Requirements for DXA for the management of osteoporosis in Europe* Osteoporos. International., 16, pp. 229–238

8. Kanis JA (2002) *Diagnosis of osteoporosis and assessment of fracture risk*. Lancet, 359, pp. 1929–1936

9. Cummings SR, Bates D, Black DM. (2002*) Clinical use of bone densitometry: scientific review*. The Journal of the American Medical Association, 288, pp. 1889–1897

10. Peacock M, Turner CH, Econs MJ and Foroud T. (2002*) Genetics of Osteoporosis.* Endocrine Reviews June 1, 2002 vol. 23 no. 3 303-326

11. Naylor S. (2003) Biomarkers: current perspectives and future prospects. Expert Review of Molecular Diagnostics 3: 525–529, 2003

12. Templin MF, Stoll D and Schwenk JM. *Protein microarrays: promising tools for proteomic research*. Proteomics 3(11), 2155–2166 (2003).

13. Patrinos G and Hanashi, S (2010). *Molecular Diagnostics: Application of proteomics to disease diagnostics*. 2nd ed. London: Elsevier Ltd. p285.

14. Ikonomidis I, Michalakeas CA, Lekakis J, Paraskevaidis I and Kremastinos DT. (2009) *Multimarker approach in cardiovascular risk prediction.* Disease Markers 26(5), 273–285.

15. Mishra NC. (2010). *Introduction to Proteomics: Principles and Applications*. John Wiley & Sons. p3.

16. Thongboonkerd V (2007) *Proteomics of Human Body Fluids: Principles, Methods, and Applications.* (Ed.) 2007, XIII, 533 p. 129

17. Anderson NL and Anderson NG. (2002) *The Human Plasma Proteome: History, Character, and Diagnostic Prospects*. November 1, 2002 Molecular & Cellular Proteomics, 1, 845-867.

18. Twyman RM. (2004) *From genomics to proteomics*. Garland Science/BIOS Scientific Publishers, 2004. p.21

19. Dooley KC. (2003*) Tandem mass spectrometry in the clinical chemistry laboratory*. Clinical Biochemistry. 36, 471– 481

20. Saugy M, Cardis C, Robinson N and Schweizer C. (2000) *Test methods: Anabolics*. Bailliere's Best Practice & Research. Clinical Endocrinology & Metabolism. 14, 111– 133

21. Holm A, Wu W and Lund-Johansen F. (2012) *Antibody array analysis of labelled proteomes: how should we control specificity?*, New Biotechnology, Volume 29, Issue 5, 15 June 2012, Pages 578–585

22. Uhlen M, Oksvold P, Fagerberg L, Lundberg E, Jonasson K, Forsberg M, Zwahlen M, Kampf C, Wester K, Hober S, Wernerus H, Björling L and Ponten F. (2010) *Towards a knowledge-based Human Protein Atlas*. Nature Biotechnology. 2010 28(12):1248-50.

23. Drexler DM, Feyen JHM and Sanders M. *Massspectrometry techniques for qualitative and quantitative analysis of biomarkers.* Drug Discovery Today, 1 (2004), pp. 17–23

24. Snellman G, Byberg L, Lind L, Melhus H, Gedeborg R, Mallmin H and Michaëlsson K. *Serum 25-hydroxyvitamin D in relation to BMD and fractures in a Swedish cohort of women and men*. Manuscript

25. Warensjö E, Byberg L, Melhus H, Gedeborg R, Mallmin H, Wolk A and Michaëlsson K (2011) *Dietary calcium intake and risk of fracture and osteoporosis: prospective longitudinal cohort study*. BMJ (Clinical Research edition) 2011; 342 doi: 10.1136/bmj.d1473

26. R Development Core Team (2008). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

27. Dieterle F, Ross A, Schlotterbeck G and Senn H. (2006) *Probabilistic quotient normalization as robust method to account for dilution of complex biological mixtures. Application in 1H NMR metabonomics.* Analytical Chemistry. 2006 Jul 1;78(13):4281-90.

28. Xionga GF, Zhanga YS, Hana BC, Chena W, Yanga Y and Peng JP. (2010) *Estradiol-regulated proline-rich acid protein 1 is repressed by class I histone deacetylase and functions in peri-implantation mouse uterus.* Molecular and Cellular Endocrinology. Volume 331, Issue 1, 1 January 2011, Pages 23–33

29. Stanczyk FZ., Jurow J and Hsing AW. (2010) *Limitations of Direct Immunoassays for Measuring Circulating Estradiol Levels in Postmenopausal Women and Men in Epidemiologic Studies*. Cancer Epidemiology, Biomarkers & Prevention April 2010 19; 903

30. Gao Y, Grassi F, Ryan MR, Terauchi M, Page K, Yang X, Weitzmann MN and Pacifici R. (2007*) IFN-γ stimulates osteoclast formation and bone loss in vivo via antigen-driven T cell activation.* The Journal of Clinical Investigation. 2007 January 2; 117(1): 122–132.

31. Watanabe K and Hishiya A (2005) *Mouse models of senile osteoporosis*. Molecular Aspects of Medicine. Volume 26, Issue 3, June 2005, Pages 221–231