Available online at www.sciencedirect.com

SciVerse ScienceDirect

journal homepage: http://www.elsevier.com/locate/aob

Gene clustering analysis in human osteoporosis disease and modifications of the jawbone

Paolo Toti [a,1], *Carolina Sbordone* [b], *Ranieri Martuscelli* [c], *Luigi Califano* [b], *Luca Ramaglia* [d], *Ludovico Sbordone* [e,2,*]

[a] *Private Practice, Via Provinciale 87B, 55041 Camaiore (Lucca), Italy*
[b] *Chair of Maxillo-Facial Surgery, School of Medicine, University of Naples "Federico II", Via S. Pansini 5, 80100 Naples, Italy*
[c] *Clinical Department of Odontostomatology and Maxillo-Facial Surgery, School of Medicine, University of Naples "Federico II", Via S. Pansini 5, 80100 Naples, Italy*
[d] *Chair of Oral Surgery, Department of Neurosciences, Reproductive and Odontostomatological Sciences, School of Medicine, University of Naples "Federico II", Via S. Pansini 5, 80131 Naples, Italy*
[e] *Chair of Odontostomatological Diseases, Department of Medicine and Surgery, School of Medicine, University of Salerno, Via S. Allende, 84081 Baronissi (Salerno), Italy*

ARTICLE INFO

Article history:
Accepted 28 February 2013

Keywords:
Computational biology
Genomics
Bone repair
Bone remodelling
Bone resorption
Osteoporosis
Maxilla
Mandible

ABSTRACT

Objective: An analysis of the genes involved in both osteoporosis and modifications of the jawbone, through text mining, using a web search tool, of information regarding gene/protein interaction.
Design: The final set of genes involved in the present phenomenon was obtained by expansion-filtering loop. Using a web-available software (STRING), interactions among all genes were searched for, and a clustering procedure was performed in which only high-confidence predicted associations were considered.
Results: Two hundred forty-two genes potentially involved in osteoporosis and in modifications of the jawbone were recorded. Seven "leader genes" were identified (CTNNB1, IL1B, IL6, JUN, RUNX2, SPP1, TGFB1), while another 10 genes formed the cluster B group (BMP2, BMP7, COL1A1, ICAM1, IGF1, IL10, MMP9, NFKB1, TNFSF11, VEGFA). Ninety-eight genes had no interactions, and were defined as "orphan genes".
Conclusions: The expansion of knowledge regarding the molecular basis causing osteoporotic traits has been brought about with the help of a de novo identification, based on the data mining of genes involved in osteoporosis and in modification of the jawbone. A comparison of the present data, in which no role was verified for 98 genes that had been previously supposed to have a role, with that of the literature, in which another 81 genes, as obtained from GWAS reviews and meta-analyses, appeared to be strongly associated with osteoporosis, probably attests to a lack of information on osteoporotic disease.

* *Corresponding author*. Tel.: +39 329 6332655; fax: +39 089 969642.
E-mail addresses: lsbordone@unisa.it, ludovico.sbordone@med.unipi.it (L. Sbordone).
[1] Formerly at Department of Surgery, University of Pisa, Via Paradisa 2, 56124 Pisa, Italy.
[2] Formerly at Implantology and Periodontology, Department of Surgery, University of Pisa, Pisa, Italy; Complex Operating Unit of Odontostomatology and Implantology, Azienda Ospedaliero-Universitaria Pisana, Pisa, Italy.
0003–9969/$ – see front matter
http://dx.doi.org/10.1016/j.archoralbio.2013.02.013

1. Introduction

Generally, osteoporosis is diagnosed when a decrease in bone mass and bone quality, manifesting itself as bone fragility, may lead to an increased propensity to fractures resulting from minimal trauma.[1]

Even if some forms of osteoporosis have been classified as primary, divided into ''post-menopausal'', or type I, and ''senile'', or type II, and as secondary, manifesting themselves through specific signs and added symptoms stemming from further causes,[2,3] osteoporotic patients have been generally screened and followed through a measurement of their bone mineral density (BMD), which is widely considered as an indicator of the risk factor for fracture (in conjunction with age and gender), and which has oriented preventive therapeutic decisions.[4,5]

The prevention of fractures obtained through an increase in bone formation, together with inhibition of bone resorption, is the ultimate aim of therapy; several strategies are employed, from modification of diet and lifestyle,[6] to the utilization of medications,[7,8] each of which may be quite effective, but also associated with severe side-effects.[9] Moreover, some unrelated BMD factors may play an important role due to the fact that most individuals without BMD-defined osteoporosis develop osteoporotic fractures.[1]

Despite the common genetic variations, probably each of the thousands of genetic variants out of a pool of thousands of genes influences the measured BMD in an infinitesimal way. Experimental explorations of rare genetic variations, helpful in improving our overall understanding of the pathophysiology and pharmacology necessary to prevent osteoporotic fractures, were started and developed using different genetic approaches, such as linkage analysis, candidate and genome-wide association studies (CGAS and GWAS, respectively), genome-wide sequencing and functional studies.[10–12]

The jaw is the secondary target of osteoporosis, even if the basal part of the jaw remains more intact, especially in the male mandible, due to its denser cortical base. So alveolar ridges and maxillary basal parts are mainly exposed to resorption, due to the high percentage of cancellous structure.[13–15] Alveolar ridge resorption can be induced by several metabolic bone diseases, such as vitamin D-resistant rickets,[16] primary and secondary hyperparathyroidism,[17,18] and Paget's disease.[19,20]

For patients who have experienced an intense process of resorption of the edentulous alveolar ridge, an established rehabilitation strategy is that of bone reconstruction with successive insertions of dental implants that support fixed prostheses. Osteoporosis is not an absolute contraindication to prosthetic-implant therapy, although the increased risk of fracture advises for the elimination of the risk factors associated with it.[21] Topic areas regarding osteoporosis, as summarized by Erdogan et al.[21], have been defined, and are listed as follows: changes in alveolar bone morphology in osteoporosis, bone repair and turnover in osteoporosis, alveolar bone regeneration/augmentation associated with low bone density, patients undergoing ridge augmentation procedures.

Through a cluster analysis of the genes involved in the present phenomenon, it was possible to identify certain genes (the so-called ''leader'' genes), which can be safely assumed to play the most essential role in the process as yet performed for other analyses.[22–25] An enormous quantity of data, including tens of thousands of simultaneous transcripts, had to be carefully sifted through in the highly complex analysis required in classical microarray experiments: the information obtained regarding the involved set of genes could then be employed for the design, analysis, and interpretation of microarray experiments that could be specifically targeted, in which the transcripts amounted to only a few hundred.

The primary aim of this study was a de novo identification of relevant genes in osteoporosis, in particular those linked to procedures of bone augmentation. Our secondary aim was to carry out a comparison between resulting primaeval information regarding the obtained final set of genes and the newly putative genes supposed as being associated with osteoporosis, as resulted from several genetic analytical approaches.

2. Materials and methods

2.1. *Genes involved in human osteoporosis*

A bioinformatic/statistical algorithm, less complex than that seen in the pristine analysis,[23] has been employed and utilized

Table 1 – The numbers of the resulting introductory set of genes sorted by different databases used.

Database	Research field	Resulted genes (number)
GenBank	A collection of publicly available annotated nucleotide sequences, including mRNA sequences with coding regions, segments of genomic DNA with a single gene or multiple genes, and ribosomal RNA gene clusters	109
OMIM	A comprehensive, authoritative, and timely compendium of human genes and genetic phenotypes	65
PubMed	Access to all citations from MEDLINE database and additional life sciences journals	188
Genatlas	Contains relevant information with respect to gene mapping and genetic diseases	15
Kegg	Database of biological systems, consisting of genetic building blocks of genes and proteins (KEGG GENES)	5
SA Bio-Sciences	From commercially available DNA-microarrays	155

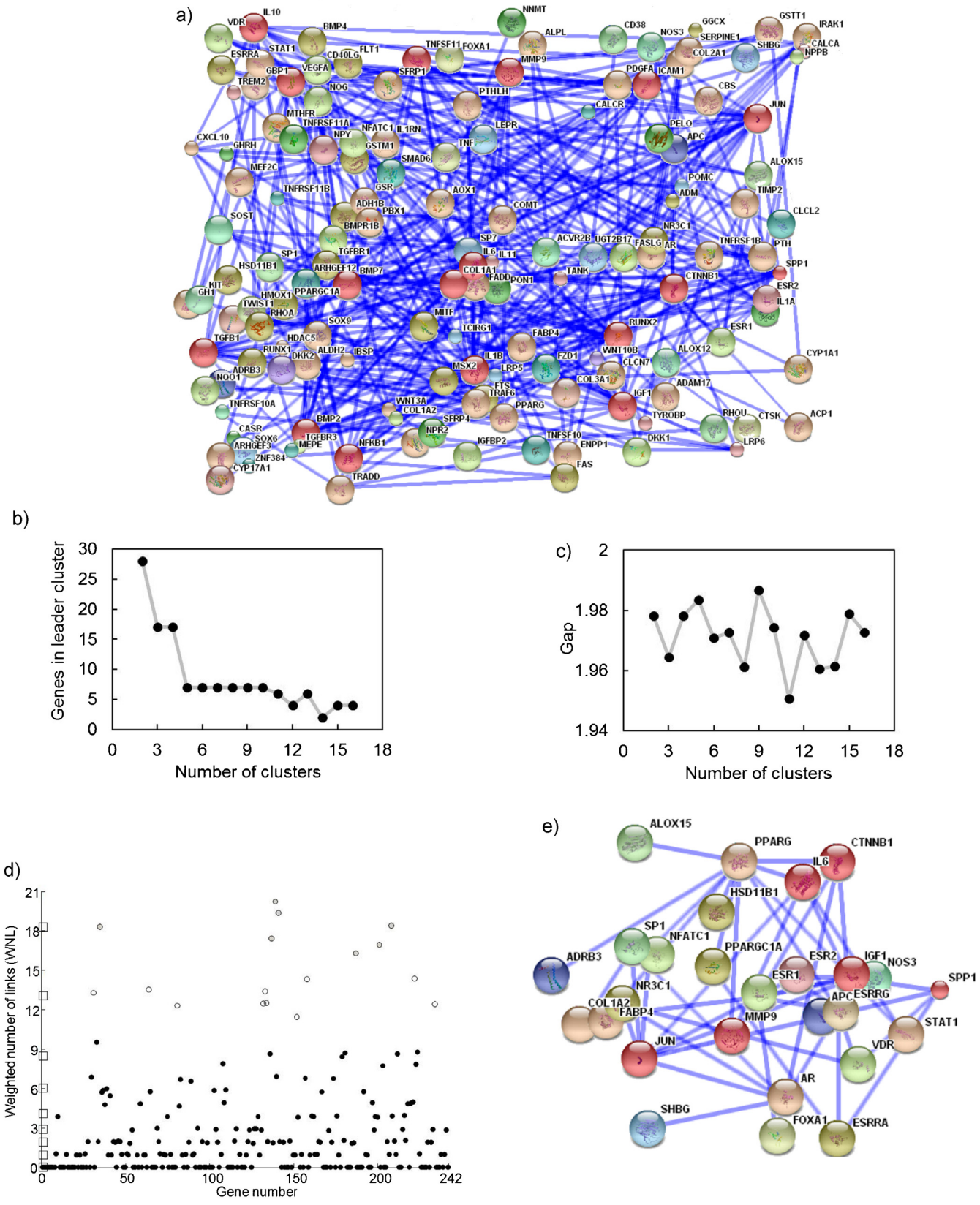

Fig. 1 – Data analysis for osteoporosis process: (a) Final map of interactions of 242 genes involved in the present phenomenon according to STRING. Leader genes and cluster B genes are red. The lines that connect single genes represent predicted functional associations among proteins, with a high degree of confidence; (b) genes belonging to the leader cluster in different *k*-mean clustering experiments with an increasing number of clusters; (c) plot of the gap statistic method for estimating the number of clusters $\hat{k}$; (d) WNL for genes involved in the process. Open circles are cluster B genes. Grey

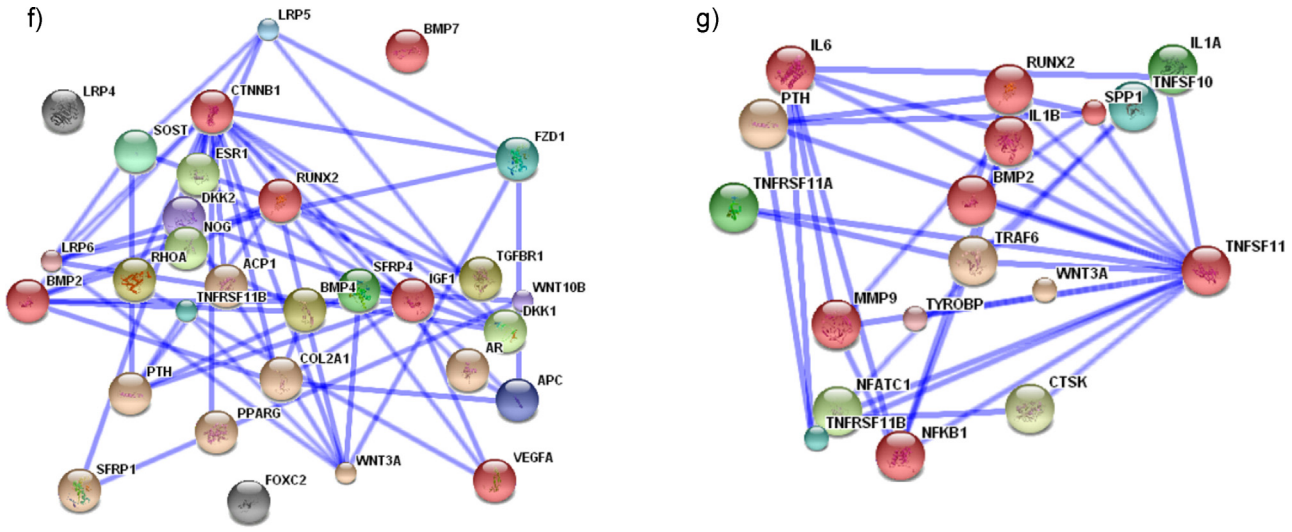

Fig. 1. (Continued).

in the present study.[24,25] This approximate clustering analysis was made up of several steps, and for each of them a different software program was used.

The first step was the establishment of an introductory set of genes by using the web-available engine Entrez (www.ncbi.nlm.nih.gov/sites/gquery, that was linked to the following list of integrated experiment cross-databases: PubMed, Genbank, Kegg, Omim, Genatlas), and via scanning of the commercially available DNA-microarray gene lists, as reported in Table 1.

The search strategy included the following pertinent key words, obtained from a review paper dealing with osteoporosis and modifications of the jawbone,[21] combined using proper boolean logic operators (AND, OR, NOT) for each search in the above-mentioned databases:

(1) Gene, human, osteoporosis;
(2) Bone changes;
(3) Bone alveolar regeneration;
(4) Bone alveolar augmentation;
(5) Ridge regeneration procedure;
(6) Ridge augmentation procedure;
(7) 2 OR 3 OR 4 OR 5 OR 6
(8) repair;
(9) turnover;
(10) resorption;
(11) remodelling;
(12) low density;
(13) 8 OR 9 OR 10 OR 11 OR 12
(14) exclusion search strategy: NOT cancer
(15) 1 AND 7 AND 13 AND 14

Once a new gene was obtained, its name was verified by means of the official Human Genome Organization (HUGO) Gene Nomenclature Committee, or HGNC (available at http://www.genenames.org/), and the approved gene symbol was applied erasing previous symbols or aliases. In the second step, the list of genes was completed; as already described,[25] the new gene dataset was expanded with web-available software STRING (Search Tool for the Retrieval of Interacting Genes/Proteins: version 8.3 available at http://string-db.org/), then false positives were filtered by a further search with PubMed. Consecutive expansion-filtering loops were stopped at the achievement of the convergence, that is when no new gene, potentially involved in the present process, was identified. Before further analysis, the name of each gene of the final convergence set involved in osteoporosis processes was newly changed, when necessary, applying that used in the STRING software.

2.2. *Identification of leader genes*

For each gene (1, 2, 3, ...) of the overall final convergence set, STRING presented all Predicted Functional Partners (α, β, χ, ...); the STRING Scores represented the level of confidence of the combined predicted associations for couples 1α, 1β, 1χ, ... For each gene of the final convergence set, only couples for which there was a high level of confidence (results with a score ≥ 0.9) were considered, and the scores were extracted, allowing the obtaining of a numerical variable called "Weighted Number of Links" (WNL), that is, the arithmetical sum of the high level combined predicted association scores.

circles are cluster leader genes. Squares are the centroids of the cluster groups; Map of interaction for genes involved in osteoporosis, sorted by different pathways according to STRING, so that leader genes and/or cluster B genes (both in red) were exposed. The lines that connect single genes represent predicted functional associations among proteins, with a high degree of confidence: (e) oestrogen endocrine- and vitamin D endocrine-pathway; (f) canonical Wnt/β-catenin signalling pathway; and (g) RANKL/RANK/OPG pathway. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

All gene-related data were entered into a matrix laboratory,[3] allowing calculations to be performed automatically. A k-mean algorithm was applied to the input variable WNL, or ''Weighted Number of Links'', and a partitioning of the overall dataset of genes into mutually-exclusive clusters was automatically performed.

As was already described, a ''gap statistic'' method was applied to the set of data for estimating the optimal number of clusters $\hat{k}$.[25] A null reference distribution, required for an accurate application of the gap statistic method, was generated, i.e. a uniform distribution throughout the data dimension, as described by Tibshirani et al.[26]: reference datasets were obtained by drawing m samples (m = number of genes), repeated 5 times.[26] Comparisons were performed between the logarithmic value of the pooled within-cluster sum of squares around cluster means [$\log(W_k)$] and its logarithmic expectation according to the null reference distribution of the data for the clusters from 2 to 16. The estimate of the number of clusters was given by the value of $\hat{k}$ for which $\log(W_k)$ fell the farthest from the reference curve.[26] Significant differences among WNLs of cluster groups obtained by the gap statistic method, were found by Kruskal–Wallis test (statistical significance at a level of $\alpha = 0.01$), verifying the accurate estimate of the number of clusters.

The first cluster, constituted by the genes with the highest WNL, was named ''leader'', or A: so ''leader genes'' were assumed to play a leading role in each analyzed process. According to their prominence, all other ranked gene classes were named as B, C, D, and so on, until the last gene class, in which genes that had no identified predicted associations (WNL = 0) were called ''orphans''.

Gene score citations (gsc) were also evaluated, as described in a previous paper,[25] by searching for citations in PubMed using Boolean logic on pertinent key words in all introduced combinations, as above described, with sequential scoring strategies: linear weight approach was applied:

- no impact-factor value restriction[27] was applied to the search for citations;
- no discrimination was applied between in vitro or in vivo data;
- increment of the gene score citation (gsc) by 1 for first citation in an article, obtained by the search for citations regarding the analyzed process and published in a journal cited in the Journal Citation Reports;
- increment of the gene score citation (gsc) by 0.5 in case of the multiple references, after the initial reference, that increments gsc by 1, and published by the same group of researchers (at least two of the same researchers) on a specific gene.

For each of the upper cluster genes (Leader and B class) a Specificity Score (SS) was assigned, computed as the ratio between the number of high level predicted interactions of the present analysis and the total number of predicted interactions with a higher level of confidence (results with a score ≥ 0.9) in the whole STRING database. The research concluded on December 8, 2011.

[3] Bioinformatic and Statistics Toolboxes, MatLab 7.0.1, The MathWorks, Natick, MA.

3. Results

After an initial canvassing using the presented key words, 199 genes potentially involved in osteoporosis disease were numbered. The searches carried out among the several databases employed are shown in Table 1. Repetitions of the same genes found in different databases were erased, leading, at the moment of the convergence (four expansion-filtering loops), to 242 genes which constituted the overall final convergence set.

The STRING map of interactions is represented in Fig. 1A for the present phenomenon. As the number of clusters increased, the k-mean clustering gave a number of genes belonging to the leader cluster ≤ 7 (Fig. 1B). The $\hat{k}$ number of clusters estimated by the ''Gap statistic'' method was 9, as shown in Fig. 1C: excluding the leader- and the orphan-clusters, the other clusters were named B, C, and so on to H.

For the analyzed process, seven genes belonged to the cluster of ''leader genes'' (Fig. 1D and Table 2), and results were validated by Kruskal–Wallis test, which computed statistically-significant differences among different clusters: leader genes, B, C, and so on. Another 10 genes were identified in the class B group. The ''leader genes'' were CTNNB1, IL1B, IL6, JUN, RUNX2, SPP1, TGFB1, while the class B gene group was constituted by BMP2, BMP7, COL1A1, ICAM1, IGF1, IL10, MMP9, NFKB1, TNFSF11, VEGFA. Table 2 presents acronyms, official names, and introduced variables of the upper cluster genes. The literature was carefully examined in order to determine which genes showed higher gene score citations (gsc) (Table 3). Among genes belonging to the two highest clusters (leader genes and cluster B genes), only 5 were well known to be linked to osteoporosis and modifications of the jawbone (with a value of gsc > 100). Among upper cluster genes, all belonging to class B, only 4 presented relatively low values of gene score citations (with a value close to 12): BMP7(11), IL10(10.5), ICAM1(10), and MMP9(5). The specificity scores for each leader and class B gene are shown in Table 2. In brief, the mean specificity scores were 20.4% and 21.1% for leader and class B genes, respectively.

In Table 4 is presented the overall set of remaining lower cluster genes (up to H cluster), with the respective gene number, acronyms, official names, and the primary function as reported by STRING. The 98 genes that had no interactions, defined as ''orphan genes'', are ranked in Table 5, in which there are reported a further 81 genes robustly associated with a reduced BMD value (resulting from several meta-analysis and reviews), that had no high level combined predicted association scores related to any gene from the overall final convergence set.

4. Discussion

The measurement of BMD is still the best predictor of osteoporotic fractures, although not all individuals experience loss of BMD with age; the aetiology of osteoporosis is seen to be multifactorial and in combination with clinical risk factors, while the rate of bone loss seems to be genetically determined,

Table 2 – Leader genes and class B genes according to clustering experiments with gene score citations in PubMed, number of predicted interactions, and Specificity Score for osteoporotic phenomenon.

Gene name	Official name	Gene number	Gene score citations (*n*)	Interactions in the analyzed phenomenon	Interactions in whole STRING database (*n*)	Specificity score (%)
Leader genes						
IL1B	Interleukin-1 beta precursor	35	31.5	18	115	15.7
IL6	Interleukin-6 precursor	137	340.5	21	142	14.8
JUN	Transcription factor AP-1 (activator protein 1)	139	37.5	22	229	9.6
RUNX2	Runt-related transcription factor 2	141	78.5	18	55	32.7
SPP1	Osteopontin precursor	187	87.5	19	45	42.2
TGFB1 (CED)	Transforming growth factor beta-1 precursor	201	146.5	19	110	17.3
'CTNNB1'	Catenin beta-1	208	38	19	181	10.5
Class B genes						
BMP2	Bone morphogenetic protein 2 precursor	31	108.5	17	75	22.6
BMP7	Bone morphogenetic protein 7 precursor	64	11	14	47	29.8
COL1A1	Collagen alpha-1 (I) chain precursor	81	85	14	58	24.1
ICAM1	Intercellular adhesion molecule 1 precursor	132	10	14	79	17.7
IGF1	Insulin-like growth factor IA precursor	133	229	17	125	13.6
IL10	Interleukin-10 precursor	134	10.5	13	79	16.5
MMP9	Matrix metalloproteinase-9 precursor	152	5	12	67	17.9
NFKB1	Nuclear factor NF-kappa-B p105 subunit	158	28.5	15	130	11.5
RANKL (TNFSF11)	Tumour necrosis factor ligand superfamily member 11	222	413	15	33	45.5
VEGFA	Vascular endothelial growth factor A precursor	234	24	17	144	11.8

Table 3 – Genes classified according to number of citations among 199 genes specifically involved in the osteoporosis phenomenon.

Gene name	Official name	Gene score citations (gsc)	Interaction in the analyzed phenomenon	Class
PTH	Parathyroid hormone precursor	>100	10	C
GH1	Somatotropin precursor	>100	2	G
TNFRSF11B	Tumour necrosis factor receptor superfamily member 11B precursor (osteoprotegerin)	>100	6	E
TNF	Tumour necrosis factor precursor (TNF-α)	>100	3	F
TNFSF11	Tumour necrosis factor ligand superfamily member 11	>100	15	B
CYP19A1	Cytochrome P450 19A1	>100	1	H
IL6	Interleukin-6 precursor	>100	21	Leader
CA10	Carbonic anhydrase-related protein 10	>100	0	Orphan
VDR	Vitamin D3 receptor	>100	4	F
IGF1	Insulin-like growth factor IA precursor	>100	17	B
ESR1	Oestrogen receptor	>100	8	D
TNFRSF11A	Tumour necrosis factor receptor superfamily member 11A precursor	>100	2	G
TGFB1	Transforming growth factor beta-1 precursor	>100	19	Leader
CALCR	Calcitonin receptor precursor	>100	2	G
BMP2	Bone morphogenetic protein 2 precursor	>100	17	B
SHBG	Sex hormone-binding globulin precursor	>100	1	H

Table 4 – All lower cluster genes with respective gene number, cluster group and primary function as reported in STRING.

Gene name	Gene number	Official name	Protein primary function in STRING	Cluster
RHOA	33	Ras homolog gene family, member A	Regulates a signal transduction pathway linking plasma membrane receptors to the assembly of focal adhesions and actin stress fibres	C
FADD	108	Fas (TNFRSF6)-associated via death domain	Apoptotic adaptor molecule that recruits caspase-8 or caspase-10 to the activated Fas (CD95) or TNFR-1 receptors	C
IL1A	136	Interleukin 1, alpha;	Produced by activated macrophages, IL-1 stimulates thymocyte proliferation by inducing IL-2 release, B-cell maturation and proliferation, and fibroblast growth factor activity	C
PPARG	179	Peroxisome proliferator-activated receptor gamma	Receptor that binds peroxisome proliferators such as hypolipidemic drugs and fatty acids	C
PTH	181	Parathyroid hormone	PTH elevates calcium level by dissolving the salts in bone and preventing their renal excretion	C
STAT1	203	Signal transducer and activator of transcription 1	Signal transducer and activator of transcription that mediates signalling by interferons (IFNs)	C
PTHLH	212	Parathyroid hormone-like hormone	Neuroendocrine peptide which is a critical regulator of cellular and organ growth, development, migration, differentiation and survival and of epithelial calcium ion transport	C
TRADD	223	TNFRSF1A-associated via death domain	Adapter molecule for TNFRSF1A/TNFR1 that specifically associates with the cytoplasmic domain of activated TNFRSF1A/TNFR1 mediating its interaction with FADD	C
TRAF6	224	TNF receptor-associated factor 6	Adapter protein and signal transducer that links members of the tumour necrosis factor receptor family to different signalling pathways by association with the receptor cytoplasmic domain and kinases	C
BMPR1B	30	Bone morphogenetic protein receptor, type IB	On ligand binding, forms a receptor complex consisting of two type II and two type I transmembrane serine/threonine kinases	D
WNT3A	36	Wingless-type MMTV integration site family, member 3A	Ligand for members of the frizzled family of seven transmembrane receptors	D
FZD1	37	Frizzled homolog 1	Receptor for Wnt proteins	D
LRP6	39	Low density lipoprotein receptor-related protein 6	Essential for the Wnt/beta catenin signalling pathway, probably by acting as a coreceptor together with Frizzled for Wnt	D
ESRRG	41	Oestrogen-related receptor gamma	Orphan receptor that acts as transcription activator in the absence of bound ligand	D
BMP4	65	Bone morphogenetic protein 4	Induces cartilage and bone formation	D
COL2A1	83	Collagen, type II, alpha 1	Type II collagen is specific for cartilaginous tissues	D
CXCL10	89	Chemokine (C-X-C motif) ligand 10	Chemotactic for monocytes and T-lymphocytes	D
ESR1	103	Oestrogen receptor 1	Nuclear hormone receptor	D
FASLG	110	Fas ligand (TNF superfamily, member 6)	Cytokine that binds to TNFRSF6/FAS, a receptor that transduces the apoptotic signal into cells	D
IRAK1	140	Interleukin-1 receptor-associated kinase 1	Binds to the IL-1 type I receptor following IL-1 engagement, triggering intracellular signalling cascades leading to transcriptional up-regulation and mRNA stabilization	D
NFATC1	157	Nuclear factor of activated T-cells, cytoplasmic, calcineurin-dependent 1;	Plays a role in the inducible expression of cytokine genes in T-cells, especially in the induction of the IL-2 or IL-4 gene transcription	D
NR3C1	167	Nuclear receptor subfamily 3, group C, member 1	Receptor for glucocorticoids (GC)	D
SOX9	174	SRY (sex determining region Y)-box 9;	Plays an important role in the normal skeletal development	D
SERPINE1	192	Serpin peptidase inhibitor, clade E (nexin, plasminogen activator inhibitor type 1), member 1	This inhibitor acts as 'bait' for tissue plasminogen activator, urokinase, and protein C	D
MEF2C	10	Myocyte enhancer factor 2C	Transcription activator which binds specifically to the MEF2 element present in the regulatory regions of many muscle-specific genes	E
DKK2	38	Dickkopf homolog 2	Inhibitor of Wnt signalling pathway (Potential)	E

Table 4 (*Continued*)

Gene name	Gene number	Official name	Protein primary function in STRING	Cluster
AR	59	Androgen receptor	Steroid hormone receptors are ligand-activated transcription factors that regulate eukaryotic gene expression and affect cellular proliferation and differentiation in target tissues	E
CD40LG	75	CD40 ligand	Mediates B-cell proliferation in the absence of co-stimulus as well as IgE production in the presence of IL-4	E
COL1A2	82	Collagen, type I, alpha 2	Type I collagen is a member of group I collagen (fibrillar forming collagen)	E
FAS	109	Fas (TNF receptor superfamily, member 6)	Receptor for TNFSF6/FASLG	E
HMOX1	129	Heme oxygenase (decycling) 1;	Heme oxygenase cleaves the heme ring at the alpha methene bridge to form biliverdin	E
MSX2	153	Msh homeobox 2;	Probable morphogenetic role	E
NOG	161	Noggin;	Essential for cartilage morphogenesis and joint formation	E
NOS3	162	Nitric oxide synthase 3 (endothelial cell);	Produces nitric oxide (NO) which is implicated in vascular smooth muscle relaxation through a cGMP-mediated signal transduction pathway	E
POMC	176	Proopiomelanocortin	ACTH stimulates the adrenal glands to release cortisol	E
SMAD6	197	SMAD family member 6	Antagonist of signalling by TGF-beta (transforming growth factor) type 1 receptor superfamily members	E
SP1	199	Sp1 transcription factor	Transcription factor that can activate or repress transcription in response to physiological and pathological stimuli	E
SP7	200	Sp7 transcription factor	Transcriptional activator essential for osteoblast differentiation	E
TGFBR1	209	Transforming growth factor, beta receptor 1	On ligand binding, forms a receptor complex consisting of two type II and two type I transmembrane serine/threonine kinases	E
TNFRSF10A	216	Tumour necrosis factor receptor superfamily, member 10a	Receptor for the cytotoxic ligand TNFSF10/TRAIL	E
TNFRSF11B	218	Tumour necrosis factor receptor superfamily, member 11b	Acts as decoy receptor for RANKL and thereby neutralizes its function in osteoclastogenesis	E
TNFRSF1B	220	Tumour necrosis factor receptor superfamily, member 1B	Receptor with high affinity for TNFSF2/TNF-alpha and approximately 5-fold lower affinity for homotrimeric TNFSF1/lymphotoxin-alpha	E
TNFSF10	221	Tumour necrosis factor (ligand) superfamily, member 10	Cytokine that binds to TNFRSF10A/TRAILR1, TNFRSF10B/TRAILR2, TNFRSF10C/TRAILR3, TNFRSF10D/TRAILR4 and possibly also to TNFRSF11B/OPG. Induces apoptosis	E
ALPL	54	Alkaline phosphatase, liver/bone/kidney	This isozyme may play a role in skeletal mineralization	F
CALCA	69	Calcitonin-related polypeptide alpha	CGRP induces vasodilation	F
CYP19A1	93	Cytochrome P450 19, subfamily A1, polypeptide 1	Cytochromes P450 are a group of heme-thiolate monooxygenases	F
ESR2	104	Oestrogen receptor 2 (ER beta)	Nuclear hormone receptor	F
FST	116	Follistatin	Binds directly to activin and functions as an activin antagonist	F
GHRH	124	Growth hormone releasing hormone	GRF is released by the hypothalamus and acts on the adenohypophyse to stimulate the secretion of growth hormone	F
GSTM1	127	Glutathione S-transferase mu 1	Conjugation of reduced glutathione to a wide number of exogenous and endogenous hydrophobic electrophiles	F
GSTT1	128	ENSG00000239562-GSTT1 protein	Unknown	F
IBSP	131	Integrin-binding sialoprotein	Binds tightly to hydroxyapatite	F
PDGFA	171	Platelet-derived growth factor alpha polypeptide	Platelet-derived growth factor is a potent mitogen for cells of mesenchymal origin	F
SOST	198	Sclerosteosis	Negative regulator of bone growth. Interferes with the Wnt signalling pathway	F

Table 4 (*Continued*)

Gene name	Gene number	Official name	Protein primary function in STRING	Cluster
TNF	215	Tumour necrosis factor (TNF superfamily, member 2)	Cytokine that binds to TNFRSF1A/TNFR1 and TNFRSF1B/TNFBR. It is mainly secreted by macrophages and can induce cell death of certain tumour cell lines	F
VDR	233	Vitamin D (1,25-dihydroxyvitamin D3) receptor	Nuclear hormone receptor	F
SFRP4	241	Secreted frizzled-related protein 4	Soluble frizzled-related proteins (sFRPS) function as modulators of Wnt signalling through direct interaction with Wnts	F
FLT1	28	Fms-related tyrosine kinase 1 (vascular endothelial growth factor/ vascular permeability factor receptor)	Receptor for VEGF, VEGFB and PGF	G
APC	34	Adenomatous polyposis coli	Tumour suppressor. Promotes rapid degradation of CTNNB1 and participates in Wnt signalling as a negative regulator	G
ACVR2B	43	Activin A receptor, type IIB	On ligand binding, forms a receptor complex consisting of two type II and two type I transmembrane serine/threonine kinases	G
ADAM 17	44	ADAM metallopeptidase domain 17	Cleaves the membrane-bound precursor of TNF-alpha to its mature soluble form	G
ADH1B	46	Alcohol dehydrogenase 1B (class I), beta polypeptide		G
ADM	47	Adrenomedullin	AM and PAMP are potent hypotensive and vasodilatator agents	G
ALOX15	53	Arachidonate 15-lipoxygenase	Converts arachidonic acid to 15S-hydroperoxyeicosatetraenoic acid	G
CALCR	70	Calcitonin receptor	This is a receptor for calcitonin	G
CTSK	88	Cathepsin K	Closely involved in osteoclastic bone resorption and may participate partially in the disorder of bone remodelling	G
DKK1	96	Dickkopf homolog 1 (Xenopus laevis)	Inhibitor of Wnt signalling pathway	G
ENPP1	99	Ectonucleotide pyrophosphatase/ phosphodiesterase 1	Involved primarily in ATP hydrolysis at the plasma membrane	G
ESRRA	105	Oestrogen-related receptor alpha	Binds to an ERR-alpha response element (ERRE) containing a single consensus half-site, 5′-TNAAGGTCA-3′	G
FOXA1	115	Forkhead box A1	Transcription activator for a number of liver genes such as AFP, albumin, tyrosine aminotransferase, PEPCK, etc.	G
GH1	123	Growth hormone 1	Plays an important role in growth control	G
GSR	126	Glutathione reductase	Maintains high levels of reduced glutathione in the cytosol	G
IL11	135	Interleukin 11	Directly stimulates the proliferation of hematopoietic stem cells and megakaryocyte progenitor cells and induces megakaryocyte maturation resulting in increased platelet production	G
IL1RN	138	Interleukin 1 receptor antagonist	Inhibits the activity of IL-1 by binding to its receptor. Has no IL-1 like activity	G
KIT	142	v-Kit Hardy-Zuckerman 4 feline sarcoma viral oncogene homolog	This is the receptor for stem cell factor (mast cell growth factor). It has a tyrosine-protein kinase activity	G
LEPR	144	Leptin receptor	Receptor for obesity factor (leptin). On ligand binding, mediates signalling through JAK2/STAT3	G
LRP5	148	Low density lipoprotein receptor-related protein 5	Involved in the Wnt/beta catenin signalling pathway, probably by acting as a coreceptor together with Frizzled for Wnt	G
NPY	165	Neuropeptide Y;	NPY is implicated in the control of feeding and in secretion of gonadotrophin-release hormone	G
NQO1	166	NAD(P)H dehydrogenase, quinone 1	The enzyme apparently serves as a quinone reductase in connection with conjugation reactions of hydroquinons involved in detoxification pathways	G
PELO	173	Pelota homolog	Required for normal chromosome segregation during cell division and genomic stability (By similarity)	G

Table 4 (*Continued*)

Gene name	Gene number	Official name	Protein primary function in STRING	Cluster
PPARGC1A	180	Peroxisome proliferator-activated receptor gamma, coactivator 1 alpha	Transcriptional coactivator for steroid receptors and nuclear receptors. Greatly increases the transcriptional activity of PPARG and thyroid hormone receptor on the uncoupling protein promoter	G
RHOU	185	Ras homolog gene family, member U	Acts upstream of PAK1 to regulate the actin cytoskeleton, adhesion turnover and increase cell migration	G
RUNX1	186	Runt-related transcription factor 1	CBF binds to the core site, 5′-PYGPYGGT-3′, of a number of enhancers and promoters	G
TANK	206	TRAF family member-associated NFKB activator	Acts as a regulator of TRAF function by maintaining them in a latent state	G
TGFBR3	210	Transforming growth factor, beta receptor III	Binds to TGF-beta. Could be involved in capturing and retaining TGF-beta for presentation to the signalling receptors	G
TNFRSF11A	217	Tumour necrosis factor receptor superfamily, member 11a, NFKB activator	Receptor for TNFSF11/RANKL/TRANCE/OPGL; essential for RANKL-mediated osteoclastogenesis	G
TWIST1	230	Twist homolog 1	Probable transcription factor, which seems to be involved in the negative regulation of cellular determination and in the differentiation of several lineages including myogenesis, osteogenesis, and neurogenesis	G
TYROBP	232	TYRO protein tyrosine kinase binding protein	Non-covalently associates with activating receptors of the CD300 family	G
HDAC5	9	Histone deacetylase 5	Responsible for the deacetylation of lysine residues on the N-terminal part of the core histones (H2A, H2B, H3 and H4)	H
SOX6	15	SRY (sex determining region Y)-box 6	Transcriptional activator. Binds specifically to the DNA sequence 5′-AACAAT-3′. Plays a key role in several developmental processes, including neurogenesis and skeleton formation	H
PBX1	21	Pre-B-cell leukaemia homeobox 1	Binds the sequence 5′-ATCAATCAA-3′. Acts as a transcriptional activator of PF4 in complex with MEIS1.	H
ITGA1	23	Integrin, alpha 1	Integrin alpha-1/beta-1 is a receptor for laminin and collagen	H
IGFBP2	26	Insulin-like growth factor binding protein 2	IGF-binding proteins prolong the half-life of the IGFs and have been shown to either inhibit or stimulate the growth promoting effects of the IGFs on cell culture	H
UGT2B17	40	UDP glucuronosyltransferase 2 family, polypeptide B17	UDPGTs are of major importance in the conjugation and subsequent elimination of potentially toxic xenobiotics and endogenous compounds	H
ACP1	42	Acid phosphatase 1, soluble	Acts on tyrosine phosphorylated proteins, low-MW aryl phosphates and natural and synthetic acyl phosphates	H
ADRB3	48	Adrenergic, beta-3-, receptor	Beta-adrenergic receptors mediate the catecholamine-induced activation of adenylate cyclase through the action of G proteins	H
ALDH2	51	ALDH2 – aldehyde dehydrogenase 2 family	Mitochondrial	H
ALOX12	52	Arachidonate 12-lipoxygenase	Oxygenase and 14,15-leukotriene A4 synthase activity	H
AOX1	57	Aldehyde oxidase 1		H
ARHGEF12	61	Rho guanine nucleotide exchange factor (GEF) 12	May play a role in the regulation of RhoA GTPase by guanine nucleotide-binding alpha-12 (GNA12) and alpha-13 (GNA13)	H
ARHGEF3	62	Rho guanine nucleotide exchange factor (GEF) 3	Acts as guanine nucleotide exchange factor (GEF) for RhoA and RhoB GTPases	H
CASR	71	Calcium-sensing receptor;	Senses changes in the extracellular concentration of calcium ions	H
CBS	73	Cystathionine-beta-synthase		H
CD38	74	CD38 molecule	Synthesizes cyclic ADP-ribose, a second messenger for glucose-induced insulin secretion	H

Table 4 (*Continued*)

Gene name	Gene number	Official name	Protein primary function in STRING	Cluster
CLCN7	77	Chloride channel 7	Mediates the exchange of chloride ions against protons	H
COL3A1	84	Collagen, type III, alpha 1	Collagen type III occurs in most soft connective tissues along with type I collagen	H
COMT	86	Catechol-O-methyltransferase	Catalyzes the O-methylation, and thereby the inactivation, of catecholamine neurotransmitters and catechol hormones	H
CXCL2	90	Chemokine (C-X-C motif) ligand 2	Produced by activated monocytes and neutrophils and expressed at sites of inflammation	H
CYP17A1	91	Cytochrome P450, family 17, subfamily A, polypeptide 1	Conversion of pregnenolone and progesterone to their 17-alpha-hydroxylated products and subsequently to dehydroepiandrosterone (DHEA) and androstenedione	H
FABP4	107	Fatty acid binding protein 4, adipocyte	Lipid transport protein in adipocytes	H
GBP1	118	Guanylate binding protein 1, interferon-inducible	Binds GTP, GDP and GMP	H
GGCX	122	Gamma-glutamyl carboxylase	Mediates the vitamin K-dependent carboxylation of glutamate residues to calcium binding gamma-carboxyglutamate (Gla) residues with the concomitant conversion of the reduced hydroquinone form of vitamin K to vitamin K epoxide	H
HSD11B1	130	Hydroxysteroid (11-beta) dehydrogenase 1	Catalyzes reversibly the conversion of cortisol to the inactive metabolite cortisone	H
MEPE	150	Matrix extracellular phosphoglycoprotein	Seems to play a role in mineralization	H
MITF	151	Microphthalmia-associated transcription factor	Transcription factor for tyrosinase and tyrosinase-related protein 1	H
MTHFR	155	5,10-Methylenetetrahydrofolate reductase (NADPH)	Catalyzes the conversion of 5,10-methylenetetrahydrofolate to 5-methyltetrahydrofolate, a co-substrate for homocysteine remethylation to methionine	H
NNMT	160	Nicotinamide N-methyltransferase;	Catalyzes the N-methylation of nicotinamide and other pyridines to form pyridinium ions	H
NPPB	163	Natriuretic peptide precursor B;	Acts as a cardiac hormone with a variety of biological actions including natriuresis, diuresis, vasorelaxation, and inhibition of renin and aldosterone secretion	H
NPR2	164	Natriuretic peptide receptor B/guanylate cyclase B	Receptor for atrial natriuretic peptide (ANP), brain natriuretic peptide (BNP), and C-type natriuretic peptide (CNP)	H
PON1	177	Paraoxonase 1	Hydrolyzes the toxic metabolites of a variety of organophosphorus insecticides	H
SFRP1	193	Secreted frizzled-related protein 1	Soluble frizzled-related proteins (sFRPS) function as modulators of Wnt signalling through direct interaction with Wnts	H
SHBG	194	Sex hormone-binding globulin	Functions as an androgen transport protein, but may also be involved in receptor mediated processes	H
TCIRG1	207	T-cell, immune regulator 1, ATPase, H + transporting, lysosomal V0 subunit A3	Part of the proton channel of V-ATPases (By similarity)	H
TIMP2	211	TIMP metallopeptidase inhibitor 2	Complexes with metalloproteinases (such as collagenases) and irreversibly inactivates them	H
TREM2	225	Triggering receptor expressed on myeloid cells 2	May have a role in chronic inflammations and may stimulate production of constitutive rather than inflammatory chemokines and cytokines	H
WNT10B	238	Wingless-type MMTV integration site family, member 10B	Ligand for members of the frizzled family of seven transmembrane receptors. Probable developmental protein	H
ZNF384	240	Zinc finger protein 384	Transcription factor that seems to bind and regulate the promoters of MMP1, MMP3, MMP7 and COL1A1 (By similarity)	H

Table 5 – List of "orphan" genes and new genes strongly associated with osteoporosis, obtained from GWAS reviews and meta-analyses.

Gene name	Gene number	Official name
Orphan genes		
RTP3	1	Receptor (chemosensory) transporter protein 3
PLCL1	2	Phospholipase C-like 1
C17orf53	3	Uncharacterized protein C17orf53
CRHR1	4	Corticotropin releasing hormone receptor 1
DCDC1	5	Doublecortin domain containing 1
DCDC5	6	Doublecortin domain containing 5
ENSG00000197851	7	Putative uncharacterized protein FLJ42280
STARD3NL	8	STARD3 N-terminal like
IL21R	11	Interleukin 21 receptor
MARK3	12	MAP/microtubule affinity-regulating kinase 3
ZBTB40	13	Zinc finger and BTB domain containing 40
FAM3C	14	Family with sequence similarity 3, member C
CCDC27	16	Coiled-coil domain containing 27
SSU72	17	SSU72 RNA polymerase II CTD phosphatase homolog
ISG15	18	ISG15 ubiquitin-like modifier
RERE	19	Arginine-glutamic acid dipeptide (RE) repeats
LTBP2	20	Latent transforming growth factor beta binding protein 2
CLDN14	22	Claudin 14
FLNB	24	Filamin B, beta
HMGA2	25	High mobility group AT-hook 2
TASP1	27	Taspase, threonine aspartase, 1
DMP1	29	Dentine matrix acidic phosphoprotein 1
FOXC2	32	Forkhead box C2 (MFH-1, mesenchyme forkhead 1)
ADAMTS18	45	ADAM metallopeptidase with thrombospondin type 1 motif, 18
ADRBK2	49	Adrenergic, beta, receptor kinase 2
AHSG	50	Alpha-2-HS-glycoprotein
ANGEL1	55	Angel homolog 1
ANKH	56	Ankylosis, progressive homolog
APOE	58	Apolipoprotein E
ARHGEF1	60	Rho guanine nucleotide exchange factor (GEF) 1
B3GAT3	63	Beta-1,3-glucuronyltransferase 3 (glucuronosyltransferase I)
CA10	66	Carbonic anhydrase X; Does not have a catalytic activity
CA8	67	Carbonic anhydrase VIII
CADM1	68	Cell adhesion molecule 1
CAT	72	Catalase
CEACAM1	76	Carcinoembryonic antigen-related cell adhesion molecule 1 (biliary glycoprotein)
CNP	78	2′,3′-Cyclic nucleotide 3′ phosphodiesterase
CNR2	79	Cannabinoid receptor 2 (macrophage)
COL15A1	80	Collagen, type XV, alpha 1
COL5A2	85	Collagen, type V, alpha 2
CP	87	Ceruloplasmin (ferroxidase)
CYP19A1	92	Cytochrome P450, family 19, subfamily A, polypeptide 1
DBP	94	D site of albumin promoter (albumin D-box) binding protein
DCXR	95	Dicarbonyl/L-xylulose reductase
EBP	97	Emopamil binding protein (sterol isomerase)
ENPEP	98	Glutamyl aminopeptidase (aminopeptidase A)
ARHGAP1	100	Rho GTPase activating protein 1
AXIN1	101	Axin 1
CYLD	102	Cylindromatosis (turban tumour syndrome)
EIF6	106	Eukaryotic translation initiation factor 6
ENPP1	111	ectonucleotide pyrophosphatase/phosphodiesterase 1
ERC1	112	ELKS/RAB6-interacting/CAST family member 1
F2	113	Coagulation factor II (thrombin)
FGFR2	114	Fibroblast growth factor receptor 2
GPR177	117	WLS – G protein-coupled receptor 177
HOXA13	119	Homeobox A13
RARG	120	Retinoic acid receptor, gamma
WNT16	121	Wingless-type MMTV integration site family, member 16
WNT4	125	Wingless-type MMTV integration site family, member 4
WNT5B	143	Wingless-type MMTV integration site family, member 5B
LIPC	145	Lipase, hepatic
LOX	146	Lysyl oxidase

Table 5 (*Continued*)

Gene name	Gene number	Official name
LRP4	147	Low density lipoprotein receptor-related protein 4
MAPK8IP2	149	Mitogen-activated protein kinase 8 interacting protein 2
MT1G	154	Metallothionein 1G
MTRR	156	5-Methyltetrahydrofolate-homocysteine methyltransferase reductase
NKRF	159	NFKB repressing factor
OSCAR	168	Osteoclast associated, immunoglobulin-like receptor
P2RX7	169	Purinergic receptor P2X, ligand-gated ion channel, 7
PDE4D	170	phosphodiesterase 4D, cAMP-specific (phosphodiesterase E3 dunce homolog)
PDLIM4	172	PDLIM4 – PDZ and LIM domain 4
PLOD1	175	Procollagen-lysine 1,2-oxoglutarate 5-dioxygenase 1
PON2	178	Paraoxonase 2
PTN	182	Pleiotrophin
QPCT	183	Glutaminyl-peptide cyclotransferase
RAGE	184	Renal tumour antigen
RXFP2	188	Relaxin/insulin-like family peptide receptor 2
SATB2	189	SATB homeobox 2
SBDS	190	Shwachman-Bodian-Diamond syndrome
SCARA3	191	Scavenger receptor class A, member 3
SLC34A1	195	Solute carrier family 34 (sodium phosphate), member 1
SLPI	196	Secretory leucocyte peptidase inhibitor
SPP2	202	Secreted phosphoprotein 2, 24 kDa
STC2	204	Stanniocalcin 2
SULT2B1	205	Sulfotransferase family, cytosolic, 2B, member 1
TMSL1	213	Thymosin-like 4 (pseudogene)
TMSL2	214	Thymosin-like 4 (pseudogene)
TNFRSF18	219	Tumour necrosis factor receptor superfamily, member 18
TRPM7	226	Transient receptor potential cation channel, subfamily M, member 7
TRPV5	227	Transient receptor potential cation channel, subfamily V, member 5
TRPV6	228	Transient receptor potential cation channel, subfamily V, member 6
TSC22D3	229	TSC22 domain family, member 3
TXNRD1	231	Thioredoxin reductase 1
VPS13B	235	Vacuolar protein sorting 13 homolog B
WASF2	236	WAS protein family, member 2
WIPF1	237	WAS/WASL interacting protein family, member 1
WRN	239	Werner syndrome, RecQ helicase-like
SPTBN1	242	Spectrin, beta, non-erythrocytic 1
Genes robustly associated with BMD on GWAS meta-analyses and reviews		
AAAS	–	Achalasia, adrenocortical insufficiency, alacrimia (Allgrove, triple-A)
ABCF2	–	ATP-binding cassette, sub-family F (GCN20), member 2
ADRA1D	–	Adrenergic, alpha-1D-, receptor
AKAP11	–	A kinase (PRKA) anchor protein 11
ALDH7A1	–	Aldehyde dehydrogenase 7 family, member A1
C11orf49	–	UPF0705 protein C11orf49
C12orf10	–	UPF0160 protein MYG1, mitochondrial Precursor
C12ORF23	–	UPF0444 transmembrane protein C12orf23
C16ORF38	–	Neuronal pentraxin-like protein C16orf38 Precursor
C18ORF19	–	Uncharacterized protein C18orf19
C6orf97	–	Coiled-coil domain-containing protein C6orf97
CCDC55	–	CCDC55
CDK5RAP2	–	CDK5 regulatory subunit associated protein 2
CDKAL1	–	CDK5 regulatory subunit associated protein 1-like 1
CKAP5	–	cytoskeleton associated protein 5
CLCN7	–	Chloride channel 7
CPN1	–	Carboxypeptidase N, polypeptide 1
DHH	–	Desert hedgehog homolog
DNM3	–	Dynamin 3
EFCAB5	–	EF-hand calcium binding domain 5
ERLIN1	–	ER lipid raft associated 1
ESPL1	–	extra spindle pole bodies homolog 1
FAM9B	–	Family with sequence similarity 9, member B
FLJ42280	–	Putative uncharacterized protein FLJ42280
FOXE1	–	Forkhead box E1 (thyroid transcription factor 2)
FOXL1	–	Forkhead box L1
FUBP3	–	Far upstream element (FUSE) binding protein 3

Table 5 (*Continued*)

Gene name	Gene number	Official name
GALNT3	–	UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase 3 (GalNAc-T3)
GPATCH1	–	G patch domain containing 1
IDUA	–	Iduronidase, alpha-L-
IFT80	–	Intraflagellar transport 80 Homolog
INSIG2	–	Insulin induced gene 2
JAG1	–	Jagged 1 (Alagille syndrome)
KAL1	–	Kallmann syndrome 1 sequence
KCNMA1	–	Potassium large conductance calcium-activated channel, subfamily M, alpha member 1
KIAA2018	–	KIAA2018
KLHDC5	–	kelch domain containing 5
KPNA4	–	karyopherin alpha 4 (importin alpha 3)
KPRP	–	Keratinocyte proline-rich protein
LCE2A	–	Late cornified envelope 2A
LCE2B	–	Late cornified envelope 2B
LCE2C	–	Late cornified envelope 2C
LCE4A	–	Late cornified envelope 4A
LEKR1	–	Putative uncharacterized protein ENSP00000407071
LIN7C	–	Lin-7 homolog C (C. elegans)
MBL2	–	Mannose-binding lectin (protein C) 2, soluble (opsonic defect)
MFSD5	–	Major facilitator superfamily domain containing 5
MHC	–	ENSG00000233904 – HLA class I histocompatibility antigen, alpha chain E Precursor (MHC class I antigen E)
MPP7	–	Membrane protein, palmitoylated 7 (MAGUK p55 subfamily member 7)
MTHFR	–	5,10-Methylenetetrahydrofolate reductase (NADPH)
NTAN1	–	N-terminal asparagine amidase
OSBPL8	–	Oxysterol binding protein-like 8
PBX1	–	Pre-B-cell leukaemia homeobox 1
PFDN5	–	prefoldin subunit 5
PKDCC	–	Protein kinase domain containing, cytoplasmic homolog
RPS6KA5	–	Ribosomal protein S6 kinase, 90 kDa, polypeptide 5
RSPO3	–	R-spondin 3 homolog
SHFM1	–	split hand/foot malformation (ectrodactyly) type 1
SLC25A13	–	Solute carrier family 25, member 13 (citrin)
SMC4	–	Structural maintenance of chromosomes 4
SMG6	–	Smg-6 homolog, nonsense mediated mRNA decay factor (C. elegans)
SOX4	–	SRY (sex determining region Y)-box 4
SUPT3H	–	Suppressor of Ty 3 homolog
TBC1D1	–	TBC1 (tre-2/USP6, BUB2, cdc16) domain family, member 1
TRIM59	–	Tripartite motif-containing 59
XKR9	–	XK, Kell blood group complex subunit-related family, member 9
ZNF408	–	zinc finger protein 408
ENSG00000163286	–	Alkaline phosphatase, placental (Regan isozyme)
ENSG00000206292	–	Major histocompatibility complex, class II, DO alpha
ERAP1	–	Endoplasmic reticulum aminopeptidase 1
FABP3	–	Fatty acid binding protein 3, muscle and heart (mammary-derived growth inhibitor)
FCGR2A	–	Fc fragment of IgG, low affinity IIa, receptor (CD32)
FKBP2	–	FK506 binding protein 2
FLT3	–	Fms-related tyrosine kinase 3
FOSB	–	FBJ murine osteosarcoma viral oncogene homolog B
FTH1	–	Ferritin, heavy polypeptide 1
GC	–	Group-specific component (vitamin D binding protein)
GCM1	–	Glial cells missing homolog 1
GDF15	–	Growth differentiation factor 15
GRPEL1	–	GrpE-like 1, mitochondrial
KL	–	klotho

even if further inheritable mechanisms, partly independent of BMD, are seen to contribute to an increased incidence of fractures.[28]

From this perspective, osteoporotic clinical phenotyping that now appears anomalous may be explained through the discovery of heretofore unknown genetic information. A great many genes regulating the BMD have been identified by several genetic approaches, such as linkage analysis (discovering inherited monogenic Mendelian human disease) or association studies for identifying susceptibility genes, such

as CGAS (analyzing polymorphic variants in candidate genes and relating haplotype), or GWAS (genotyping large numbers of SNPs across the genome, rather than a few candidate genes), with results that are not perfectly in agreement, and are sometimes conflicting.[11]

Numerous genes of at least three metabolic pathways, these being the estrogenic endocrine pathway, the Wnt/Catenin pathway, and the RANKL/RANK/OPG pathway, have been identified as surely associated to osteoporotic traits, and other genes, instead, have been indicated as being either promising,[29] or robustly associated with BMD,[10,28,30] whereas novel genome-wide-significant loci were supposed to be strongly associated.[1,11,12,31]

The preliminary set counted 199 genes; after four expansion-filtering loops based on both the database-search and microarray gene list scanning, the total number was increased to 242 genes, showing that STRING expansion- and PubMed filtering-process, combined into an iterative loop, were essential to reach the convergence of data and to perform an accurate analysis. STRING was employed for a combination of databases regarding direct physical protein–protein interaction, experimental gene expression, an active role in particular metabolic pathways, and indirect functional associations.

Although previous studies regarding the cell cycle of human T lymphocytes, human periodontal disease, and human bone augmentation/remodelling established that the numbers of genes in the higher clusters were small, our study revealed that 17 genes belonged to leader- or B-cluster; this resulting number of higher cluster genes was in accordance with the increasing set of genes which were either candidates to be associated or were found to be robustly associated with osteoporotic disease.

Most of the discovered ''leader genes'' were the same as those involved in the already examined augmentation/remodelling process[25]: these were divided into *cytokines*, such as interleukins, in particular IL1B, which is produced by activated macrophages that are important mediators of the inflammatory response, whereas IL6 has the capacity to stimulate the development of osteoclasts from their hematopoietic precursors,[32] or such as the growth factor TGFB1, a multifunctional peptide that regulates osteoblast chemotaxis, proliferation and differentiation,[33,34] and *transcription factors*, like RUNX2, that plays an important role in the maturation of osteoblasts as well as in both intramembranous and endochondral ossification, with its ''master switch'' activity,[35–37] and JUN, which is involved in regulating gene expression.[38]

The two adjunctive ''leader genes'' were CTNNB1, an essential molecule of the canonical Wnt/β-catenin signalling pathway, and SPP1, also known as osteopontin, which is bounded tightly to hydroxyapatite, performing interactions between the matrix and cells,[39] and these two genes resulted as being associated with BMD at genome-wide-significant levels in several meta-analyses and reviews.[1,11,31]

The class B gene group is constituted by genes functionally and structurally divided into architectural proteins, non-collagen matrix molecules, and transcription- and cytokines/growth-factors. COL1A1 and ICAM1 belong to the first group. COL1A1 encodes the α-1 chain of type I collagen, the principal component of bone extracellular matrix: the studies performed on the effects on BMD of the three polymorphisms suggested that homozygotes for different haplotypes showed reduced (haplotype2) or increased (haplotype3) BMD.[40] Although COL1A1 has been extensively studied in osteoporosis,[11,12,28] no information has been obtained in connection to ICAM1, which plays an active role in the inflammatory process due to contact between osteoblasts (Obs) and osteoclasts (Ocs), occurring through the Ob membrane complex, in which the membrane protein icam1 binds the counter-receptor expressed at the surface of osteoclasts.[41] Also MMP9, an enzyme involved in degradation of types IV and V collagens, and NFKB1, which has a prominent role in (antibody) clearance, phagocytosis, pathogen recognition, and inflammatory response, have been little examined in more recent meta-analyses and reviews, even if, together with COL1A1, they have been seen to be significantly down-regulated in the bone tissue of osteoporotic women.[42,43]

The analysis of gene-coding cytokines resumed with 5 growth factors, such as BMP2, BMP7, IGF1, VEGFA and TNFSF11, and one interleukin, IL10. Interleukin 10, primarily produced by monocytes and bone morphogenetic proteins (BMPs) belonging to the group of multifunctional peptides (TGFβ superfamily) that control proliferation and differentiation in many cell types, yielded inconclusive results regarding an association with osteoporosis[44,45] as attested to by the absence of these two genes from all the more recent reviews and meta-analyses regarding genome-wide association studies.[1,12,31]

Another two genes, although belonging to upper clusters, appeared not to be associated to osteoporosis after research performed among genome-wide-significant loci of GWAS meta-analyses[1,12,31]: IGF1, encoding a protein similar to insulin in function and structure, which, interacting with some variants of ESR2 (Oestrogen Receptor-β), influences the risk of fracture in postmenopausal women,[46] and the growth factor encoded by VEGFA, that shows several effects on osteoclast differentiation, angiogenesis, and monocyte/macrophage migration.[47]

The last class B gene was TNFSF11, a member of the tumour necrosis factor (TNF) cytokine family, which is a ligand for osteoprotegerin and functions as a key factor for osteoclast differentiation and activation.[48]

The prevention of bone loss in postmenopausal women via oestrogen replacement therapy has shown that an active role is played by the oestrogen endocrine system in regulating bone mass and the occurrence of osteoporosis. The binding between oestrogen and its receptors, such as the most extensively studied oestrogen receptors 1 and 2 (ESR1, ESR2), and oestrogen-related receptors ERR-α (ESRRA) and ERR-γ (ESRRG), resulted in the up-regulation of the expression of many genes. Experiments and link connection analysis showed that oestrogen receptors influenced the occurrence of osteoporosis, both on their own, and by interacting with the higher cluster genes CTNNB1 and IGF1, as well as with the leader SPP1 gene for ESRs and for ESRRs, respectively, as is seen in Fig. 1E.[29]

As regards the vitamin D endocrine pathway and maintenance of calcium homeostasis, effects of vitamin D are mediated through its nuclear transcription factor (VDR), but

the association between VDR polymorphisms (*Cdx2*, *FokI*, *BsmI*, *ApaI*, and *TaqI*) and BMD were uncertain; even if vitamin D binding protein (DBP) binds and transports vitamin D, and mediates bone resorption by directly activating osteoclasts acting as DBP-macrophage activating factor, the presence of other genetic and environmental factors was an indispensable condition for revealing the poor effect of the DBP gene on fracture risk.[49] DBP was an orphan gene, whereas, at third knot, VDR resulted as connected throughout the oestrogen-related receptor gamma (ESRRG) to CTNNB1, SPP1, JUN, IL6, MMP9, and IGF1 (Fig. 1E).

The oestrogen endocrine system and the vitamin D endocrine pathway seem to be interconnected through the ESRRG, which is a candidate gene for osteoporosis, due to consistent associations discovered by recent analysis.[50]

The mechanism of the canonical Wnt/β-catenin signalling pathway has been well known, and its importance in the regulation of bone mass is attested to in the literature.[51] The protein encoded by CTNNB1 (β-catenin) is part of a complex of proteins that is responsible for transmitting signals throughout the nuclear transcriptional activity of β-catenin, when the receptor complex consisting of frizzled family proteins (encoded by frizzled precursors, FZDs) and low-density lipoprotein receptor-related protein 5 (LRP5) or LRP6 co-receptor binds as ligand the Wnt.[52] In Fig. 1F, all osteoporosis-associated genes reviewed by Li and co-workers are presented with all their connections at first knot[29]; the presence of the two leader genes, RUNX2, which is that acting in the proliferation of the osteoprogenitor cell via the canonical Wnt/β-catenin signalling pathway, and CTNNB1, gives evidence that the gene encoding β-catenin, as reported in several GWAS meta-analyses, is strongly associated to osteoporotic disease.[1,11,29,31]

Associations between osteoporosis and the genes belonging to RANKL/RANK/OPG pathway were perfectly established. The three genes of the tumour necrosis factor superfamily, TNFSF11, TNFRSF11A, and TNFRSF11B, encoded ligand rankl (receptor activator of nuclear factor-κB ligand), receptor rank (receptor activator of the nuclear factor-κB), and opg (osteoprotegerin or tumour necrosis factor receptor superfamily member 11B). The formation, activation and survival of multinucleated osteoclasts in normal bone remodelling were promoted by the normal binding between rankl and rank; opg acted as a decoy receptor for rankl, and, thereby, neutralized its function in osteoclastogenesis.[48] In Fig. 1G, rankl, rank and opg are presented with all connections at first knot; most of the links in Fig. 1G were due to rankl, indicating that this ligand may be affected by other pathways.

The present in silico study furnished a list of 242 genes potentially involved in osteoporosis linked to procedures of jaw bone augmentation. Table 5 reports 98 orphan genes along with a further 81 genes, obtained from GWAS reviews and meta-analyses.[1,10–12,28,30,31] The majority of these latter also resulted as ''orphans'', while the remaining ones are members of the lowest class, that is, with, at most, just one significant interaction with present genes belonging to the final convergence set.

The strength of the present analysis, which aimed at the identification of genes playing a relevant role in osteoporosis linked to jawbone reconstruction, was based on a combination of databases regarding direct physical protein–protein interaction, experimental gene expression, an active role in particular metabolic pathways, and indirect functional associations, could also be its weakness, owing to possible ''publication bias''. Some genes of the lower class and, moreover, the ''orphans'', for which information is very scant, could be found to have a significant role in the future, as can be seen from the wide number of genes for which a strong association was indicated, albeit with there being a lack of information regarding their role.

Several microarray experiments were performed in order to identify differentially expressed genes in osteoporotic bone; moreover, genes identified and implicated in osteoporosis pathogenesis were variable in different experimental analyses.[53,54] The set for microarray analysis that is commercially available presents lists specifically addressed to different pathways (SuperArray Bioscience Corporation: 7320 Executive Way Frederick, MD 21704 USA), but the roles of several promising genes (lower class and orphan genes) are not considered. The lack of information concerning osteoporosis was borne out by the role played by both associated genes, and by the numerous old and new ''orphan genes'', not part of the known and well-analyzed pathways; further targeted DNA microarray experiments might help to clarify these still obscure biomolecular mechanisms.

Funding

MIUR Italian Dept. of Education, University and Scientific Research: PRIN 2008 2008K4XXF8_02 (CUP G51J10000030001) and 2008K4XXF8_03 (CUP E61J100000200001).

Competing interest

No conflict of interest.

Ethical approval

No ethical approval is required.

Acknowledgments

The present study has been supported in part through the Ministero dell'Istruzione, Università e Ricerca Scientifica (MIUR Italian Dept. of Education, University and Scientific Research) PRIN 2008 2008K4XXF8_02 (CUP G51J10000030001) to Prof. Ludovico Sbordone, and 2008K4XXF8_03 (CUP E61J100000200001) to Prof. Ranieri Martuscelli.

REFERENCES

1. Richards JB, Zheng HF, Spector TD. Genetics of osteoporosis from genome-wide association studies: advances and challenges. *Nature Reviews Genetics* 2012;**13**(8): 576–88.

2. Riggs BL, Melton LJ. Evidence for two distinct syndromes of involutional osteoporosis. *American Journal of Medicine* 1983;**75**(6):899–901.
3. Downey PA, Siegel MI. Bone biology and the clinical implications for osteoporosis. *Physical Therapy* 2006;**86**(1): 77–91.
4. Kanis JA, Gluer CC. An update on the diagnosis and assessment of osteoporosis with densitometry. Committee of Scientific Advisors, International Osteoporosis Foundation. *Osteoporosis International* 2000;**11**(3):192–202.
5. Lane NE. Epidemiology, etiology, and diagnosis of osteoporosis. *American Journal of Obstetrics and Gynecology* 2006;**194**(2 Suppl.):3–11.
6. Lane JM, Russell L, Khan SN. Osteoporosis. *Clinical Orthopaedics and Related Research* 2000;**372**:139–50.
7. Turner RT, Riggs BL, Spelsberg TC. Skeletal effects of estrogen. *Endocrine Reviews* 1994;**15**(3):275–300.
8. Rodan GA, Martin TJ. Therapeutic approaches to bone diseases. *Science* 2001;**289**(5484):1508–14.
9. Marx RE, Sawatari Y, Fortin M, Broumand V. Bisphosphonateinduced exposed bone (osteonecrosis/ osteopetrosis) of the jaws: risk factors, recognition, prevention, and treatment. *Journal of Oral and Maxillofacial Surgery* 2005;**63**(11):1567–75.
10. Cheung CL, Xiao SM, Kung AWC. Genetic epidemiology of age-related osteoporosis and its clinical applications. *Nature Reviews Rheumatology* 2010;**6**(9):507–17.
11. Ralston SH, Uitterlinden AG. Genetics of osteoporosis. *Endocrine Reviews* 2010;**31**(5):629–62.
12. Cheung CL, Sham PC, Xiao SM, Bow CH, Kung AWC. Meta-analysis of gene-based genome-wide association studies of bone mineral density in Chinese and European subjects. *Osteoporosis International* 2012;**23**(1):131–42.
13. Sbordone C, Toti P, Guidetti F, Califano L, Bufo P, Sbordone L. Volume changes of autogenous bone after sinus lifting and grafting procedures: a 6-year computerized tomographic follow-up. *Journal of Cranio-maxillo-facial Surgery*, in press.
14. Sbordone C, Toti P, Guidetti F, Califano L, Santoro A, Sbordone L. Volume Changes of Iliac Crest Autogenous Bone Grafts after Vertical and Horizontal Alveolar Ridge Augmentation of Atrophic Maxillae and Mandibles: A 6-year Computerized Tomographic Follow-up. *Journal of Oral and maxillofacial Surgery* 2012;**70**(11):2559–65.
15. Sbordone C, Toti P, Guidetti F, Martuscelli R, Califano L, Sbordone L. Healing of Donor Defect After Mandibular Parasymphyseal Block Harvesting: A 6-year Computerized Tomographic follow-up. *Journal of Cranio-Maxillo-Facial Surgery* 2012;**40**(5):421–6.
16. Hillmann G, Geurtsen W. Pathohistology of undecalcified primary teeth in vitamin D-resistant rickets: review and report of two cases. *Oral Surgery Oral Medicine Oral Pathology Oral Radiology and Endodontics* 1996;**82**(2):218–24.
17. Nathan AS, Traiger J, Berman SA. Secondary hyperparathyroidism as a cause of generalized enlargement of the maxilla and mandible. Report of a case. *Oral Surgery Oral Medicine Oral Pathology* 1966;**21**(6):724–31.
18. Frankenthal S, Nakhoul F, Machtei EE, Green J, Ardekian L, Laufer D, et al. The effect of secondary hyperparathyroidism and hemodialysis therapy on alveolar bone and periodontium. *Journal of Clinical Periodontology* 2002;**29**(6):479–83.
19. Hamdy RC. *Paget's disease of bone: assessment and management*. 1st ed. New York: Praeger; 1981.
20. Bender IB. Paget's disease. *Journal of Endodontics* 2003;**29**(11):720–3.
21. Erdogan O, Shafer DM, Taxel P, Freilich MA. A review of the association between osteoporosis and alveolar ridge augmentation. *Oral Surgery Oral Medicine Oral Pathology Oral Radiology and Endodontics* 2007;**104**(6):738.e1–e.
22. Nicolini C, Spera R, Stura E, Fiordoro S, Giacomelli L. Gene expression in the cell cycle of human T-lymphocytes: II. Experimental determination by DNASER technology. *Journal of Cellular Biochemistry* 2006;**97**(5):1151–9.
23. Sivozhelezov V, Giacomelli L, Tripathi S, Nicolini C. Gene expression in the cell cycle of human T lymphocytes: I. Predicted gene and protein networks. *Journal of Cellular Biochemistry* 2006;**97**(5):1137–50.
24. Covani U, Marconcini S, Giacomelli L, Sivozhelevov V, Barone A, Nicolini C. Bioinformatic prediction of leader genes in human periodontitis. *Journal of Periodontology* 2008;**79**(10):1974–83.
25. Sbordone L, Sbordone C, Filice N, Menchini-Fabris GB, Baldoni M, Toti P. Gene clustering analysis in human osseous remodeling. *Journal of Periodontology* 2009;**80**(12):1998–2009.
26. Tibshirani R, Walther G, Hastie T. Estimating the number of clusters in a data set via gap statistic. *Journal of the Royal Statistical Society Series B* 2001;**63**(2):411–23.
27. Fitzpatrick RB. ISI's journal citation reports on the web. *Medical Reference Services Quarterly* 2003;**22**(4):45–56.
28. Mitchell BD, Yerges-Armstrong LM. The genetics of bone loss: challenges and prospects. *Journal of Clinical Endocrinology and Metabolism* 2011;**96**(5):1258–68.
29. Li WF, Hou SX, Yu B, Li MM, Férec C, Chen JM. Genetics of osteoporosis: accelerating pace in gene identification and validation. *Human Genetics* 2010;**127**(3):249–85.
30. Hosoi T. Genetic aspects of osteoporosis. *Journal of Bone and Mineral Metabolism* 2010;**28**(6):601–7.
31. Estrada K, Styrkarsdottir U, Evangelou E, Hsu YH, Duncan EL, Ntzani EE, et al. Genome-wide meta-analysis identifies 56 bone mineral density loci and reveals 14 loci associated with risk of fracture. *Nature Genetics* 2012;**44**(5):491–501.
32. Ihara H, Hirukawa K, Goto S, Togari A. ATP-stimulated interleukin-6 synthesis through P2Y receptors on human osteoblasts. *Biochemical and Biophysical Research Communications* 2005;**326**(2):329–34.
33. Ortiz CO, Chen BK, Bale LK, Overgaard MT, Oxvig C, Conover CA. Transforming growth factor-beta regulation of the insulin-like growth factor binding protein-4 protease system in cultured human osteoblasts. *Journal of Bone and Mineral Research* 2003;**18**(6):1066–72.
34. Wright N, de Lera TL, García-Moruja C, Lillo R, García-Sánchez F, Caruz A, et al. Transforming growth factor-beta1 down-regulates expression of chemokine stromal cell-derived factor-1: functional consequences in cell migration and adhesion. *Blood* 2003;**102**(6):1978–84.
35. Brubaker KD, Vessella RL, Brown LG, Corey E. Prostate cancer expression of runt-domain transcription factor Runx2, a key regulator of osteoblast differentiation and function. *Prostate* 2003;**56**(1):13–22.
36. Schroeder TM, Jensen ED, Westendorf JJ. Runx2: a master organizer of gene transcription in developing and maturing osteoblasts. *Birth Defects Research Part C Embryo Today* 2005;**75**(3):213–25.
37. Komori T. Regulation of bone development and maintenance by Runx2. *Frontiers in Bioscience* 2008;**13**: 898–903.
38. Vesely PW, Staber PB, Hoefler G, Kenner L. Translational regulation mechanisms of AP-1 proteins. *Mutation Research* 2009;**682**(1):7–12.
39. Bonewald LF, Johnson M. Osteocytes, mechanosensing and Wnt signaling. *Bone* 2008;**42**(4):606–15.
40. Stewart TL, Jin H, McGuigan FE, Albagha OM, Garcia-Giralt N, Bassiti A, et al. Haplotypes defined by promoter and intron 1 polymorphisms of the COLIA1 gene regulate bone mineral density in women. *Journal of Clinical Endocrinology and Metabolism* 2006;**91**(9):3575–83.

41. Lavigne P, Benderdour M, Shi Q, Lajeunesse D, Fernandes JC. Involvement of ICAM-1 in bone metabolism: a potential target in the treatment of bone diseases? *Expert Opinion on Biological Therapy* 2005;**5**(3):313–20.
42. Balla B, Kósa JP, Kiss J, Borsy A, Podani J, Takács I, et al. Different gene expression patterns in the bone tissue of aging postmenopausal osteoporotic and non-osteoporotic women. *Calcified Tissue International* 2008;**82**(1):12–26.
43. Balla B, Kósa JP, Kiss J, Podani J, Takács I, Lazáry A, et al. Transcriptional profiling of immune system-related genes in postmenopausal osteoporotic versus non-osteoporotic human bone tissue. *Clinical Immunology* 2009;**131**(2):354–9.
44. Styrkarsdottir U, Cazier JB, Kong A, Rolfsson O, Larsen H, Bjarnadottir E, et al. Linkage of osteoporosis to chromosome 20p12 and association to BMP2. *PLoS Biology* 2003;**1**(3):E69.
45. Freedman BI, Bowden DW, Ziegler JT, Langefeld CD, Lehtinen AB, Rudock ME, et al. Bone morphogenetic protein 7 (BMP7) gene polymorphisms are associated with inverse relationships between vascular calciWcation and BMD: the Diabetes Heart Study. *Journal of Bone and Mineral Research* 2009;**24**(10):1719–27.
46. Rivadeneira F, van Meurs JB, Kant J, Zillikens MC, Stolk L, Beck TJ, et al. Estrogen receptor beta (ESR2) polymorphisms in interaction with estrogen receptor alpha (ESR1) and insulin-like growth factor I (IGF1) variants influence the risk of fracture in postmenopausal women. *Journal of Bone and Mineral Research* 2006;**21**(9):1443–56.
47. Nihrane A, Sezgin G, Dsilva S, Dellorusso P, Yamamoto K, Ellis SR, et al. Depletion of the Shwachman-Diamond syndrome gene product, SBDS, leads to growth inhibition and increased expression of OPG and VEGF-A. *Blood Cells Molecules and Diseases* 2009;**42**(1):85–91.
48. Boyce BF, Xing L. Biology of RANK, RANKL, and osteoprotegerin. *Arthritis Research and Therapy* 2007;**9**(Suppl1):1.
49. Fang Y, van Meurs JB, Arp P, van Leeuwen JP, Hofman A, Pols HA, et al. Vitamin D binding protein genotype and osteoporosis. *Calcified Tissue International* 2009;**85**(2):85–93.
50. Elfassihi L, Giroux S, Bureau A, LaXamme N, Cole DE, Rousseau F. Association with replication between estrogen-related receptor gamma (ESRRG) polymorphisms and bone phenotypes in women of European ancestry. *Journal of Bone and Mineral Research* 2010;**25**(4):901–11.
51. Krishnan V, Bryant HU, Macdougald OA. Regulation of bone mass by Wnt signaling. *Journal of Clinical Investigation* 2006;**116**(5):1202–9.
52. Khosla S, Westendorf JJ, Oursler MJ. Building bone to reverse osteoporosis and repair fractures. *Journal of Clinical Investigation* 2008;**118**(2):421–8.
53. Dvornyk V, Recker RR, Deng HW. Gene expression studies of osteoporosis: implications for microarray research. *Osteoporosis International* 2003;**14**(6):451–61.
54. Hopwood B, Tsykin A, Findlay DM, Fazzalari NL. Gene expression profile of the bone microenvironment in human fragility fracture bone. *Bone* 2009;**44**(1):87–101.