

RESEARCH

Open Access

# Identification of drug candidate for osteoporosis by computational bioinformatics analysis of gene expression profile

Guiyong Yu<sup>1\*</sup>, Litao Wang<sup>1†</sup>, Yazhou Li<sup>2</sup>, Zhihong Ma<sup>3</sup> and Yu Li<sup>1</sup>

## Abstract

**Background:** Osteoporosis is a condition of bones that leads to an increased susceptibility to fracture and consequent painful morbidity. It has become a major issue of life quality worldwide. However, until now, the molecular mechanism of this disease is far from being clear.

**Methods:** In this study, we obtained the gene expression profile of osteoporosis and controls from Gene Expression Omnibus and identified differentially expressed genes (DEGs) using classical t-test method. Then, functional enrichment analyses were performed to identify the dysregulated Gene Ontology categories and dysfunctional pathways in osteoporosis patients compared to controls. Besides, the connectivity map was used to identify compounds that induced inverse gene changes to osteoporosis.

**Results:** A total of 5581 DEGs were identified. We found these DEGs were enriched in 9 pathways by pathway enrichment analysis, including focal adhesion and MAPK signaling pathway. Besides, sanguinarine was identified as a potential therapeutic drug candidate capable of targeting osteoporosis.

**Conclusion:** Although candidate agents identified by our approach may be premature for clinical trials, it is clearly a direction that warrants additional consideration.

**Keywords:** Osteoporosis, Differentially expressed genes, Dysfunctional pathway, Small molecule

## Background

Osteoporosis is a condition of bones that leads to an increased susceptibility to fracture and consequent painful morbidity [1,2]. The prevalence of osteoporosis increases with age due to an imbalance between bone resorption and bone formation during the bone remodeling cycle. Osteoporosis affects up to 30% of women and 12% of men at some point in life and it is a major quality of life issue worldwide [3].

The well accepted pathophysiological mechanisms for osteoporosis include early apoptosis of osteoblasts [4] and osteocytes [5,6], prolongation of the life span of osteoclasts [7] and the imbalance between osteoblastogenesis and adipogenesis of bone marrow mesenchymal stem cells [8]. Many factors influence the risk of osteoporosis, including

predominantly peak bone density along with other factors such as genetic factors, body weight, diet, physical activity, medication use, and coexisting disease [3,9,10]. Besides, lack of estrogen, deficiency of calcium and vitamin D are also important common causes of osteoporosis [11-14]. Various molecular signals were identified to regulate the activation of osteoclasts. Osteoprotegerin (OPG) binds activator for nuclear factor  $\kappa$ B ligand (RANKL), and hence suppresses its ability to increase bone resorption [15]. The role of Wnt signaling pathway is recognized but less well understood. Local production of eicosanoids and interleukins is thought to participate in the regulation of bone turnover, and excess or reduced production of these mediators may underlie the development of osteoporosis [16]. However, until now, the molecular mechanism of this disease is far from being clear.

In the present study, we aim to explore the molecular mechanism of osteoporosis using a computational bioinformatics analysis of gene expression, and to identify

\* Correspondence: yuguiyong123@hotmail.com

†Equal contributors

<sup>1</sup>Department of Orthopedic, The people's Hospital of Hengshui, No.180 Renmin Street, Hengshui 053000, Hebei Province, China  
Full list of author information is available at the end of the article

small molecules for the treatment of osteoporosis. Candidate agents identified by our approach may provide the ground work for a new therapy approach for osteoporosis. However, further evaluations for their potential use are needed.

## Methods

### Affymetrix microarray data

The gene expression profile of GSE 35956 [17] was downloaded from a public functional genomics data repository Gene Expression Omnibus (GEO; <http://www.ncbi.nlm.nih.gov/geo/>), which is based on the Affymetrix GPL570 platform data (Affymetrix Human Genome U133 Plus 2.0 Array). Only 10 genechips were available for further analysis, including 5 genechips from human mesenchymal stem cells (MSCs) of osteoporosis patients and 5 genechips from human MSCs of non-osteoporotic controls. The Human MSCs of elderly patients suffering from osteoporosis were isolated from femoral heads after low-energy fracture of the femoral neck. Control cells were obtained from bone marrow of femoral heads of middle-aged, non-osteoporotic donors after total hip arthroplasty.

### Pathway data

KEGG (Kyoto Encyclopedia of Genes and Genomes) is a collection of online databases dealing with genomes, enzymatic pathways, and biological chemicals [18]. The PATHWAY database records networks of molecular interactions in the cells, and variants of them specific to particular organisms (<http://www.genome.jp/kegg/>).

### Small molecules data

The connectivity map (CMap) can be used to find connections among small molecules sharing a mechanism of action, chemicals and physiological processes, and diseases and drugs [19]. It is the first installment of a reference collection of gene-expression profiles from cultured human cells treated with bioactive small molecules, together with pattern-matching software to mine these data. The CMap dataset comprises genomic profiling data from 6100 treatment-control pairs (instances) involving 1309 bioactive molecules (perturbagens). We downloaded all the profile data for further analysis.

### Differentially expressed genes (DEG) analysis

For GSE 35956, we used R package (v.2.13.0) to analyze the gene expression. The CEL source files from all conditions were processed into expression estimates and performed background correction and quartile data normalization using RMA (Robust Multi-array Average) algorithm [20]. Probe sets were mapped to national center for biotechnology information (NCBI) entrez genes using DAVID [21]. If there are multiple probe sets that

correspond to the same gene, the expression values of those probe sets are averaged. We used the classical t-test to identify differentially expressed genes and defined  $p$ -value  $< 0.05$  to be statistically significant.

### Go ontology analysis

Gene Ontology (GO) analysis has become a commonly used approach for functional studies of large-scale genomic or transcriptomic data [22]. To better understand the functional relevance of the identified DEGs, we performed GO enrichment analysis using goProfiles [23] and searched for over-representation in GO categories in three categories, namely biological process, molecular function and cellular component.

### Pathway enrichment analysis

DAVID (The Database for Annotation, Visualization and Integrated Discovery) consists of an integrated biological knowledgebase and analytic tools aimed at systematically extracting biological meaning from large gene/protein lists [21]. We used the DAVID to identify over-represented KEGG categories in pathways based on the hypergeometric distribution with the count larger than 2 and the FDR less than 0.01.

### Small molecule identification

We first divided the DEGs into two groups: up-regulated group and down-regulated group. Then, we selected the top 500 significant DEGs in each group and performed gene set enrichment analysis compared to the gene profile of a treatment-control pair (instance) in CMap database. The output consisted of a group of chemical perturbagens with an enrichment score ranging from +1 and -1. The score represented the correlation between the query signature profile (osteoporosis) and the gene profile of a treatment-control pair. A high positive connectivity score indicated that the corresponding perturbagen induced the expression of the query signature (osteoporosis), whereas a high negative enrichment score indicated reversal of expression of the query signature (osteoporosis) by the perturbagen. A zero or "null" enrichment score indicated that no effect upon expression of the query signature was recorded.

## Results

### Analysis of DEGs in osteoporosis

Publicly available microarray dataset of human MSCs from osteoporosis and control were obtained from GEO database. We used the classical t-test method to calculate the  $p$ -values of all genes and defined  $p$ -value  $< 0.05$  to be statistically significant. Expressions of 5581 genes were identified differed across osteoporosis and control group.

### Functional annotation of the DEGs

In order to facilitate the functional annotation and analysis of large lists of genes in our result, we identified over-represented GO categories in three ontologies (Figure 1, Figure 2 and Figure 3). In the ontology of cellular component, GO categories of cell and cell part tied for first place in the enrichment (4982 genes). Besides, a total of 3487 genes were enriched in the GO category of organelle. In the ontology of biological process, the most significant GO category was cellular process, and a total of 4311 genes enriched in this category. In the ontology of molecular function, the GO category of binding included 4055 genes which was the most significant category in this ontology. The second significant GO category is catalytic activity (1730 genes).

### Identification of dysfunctional pathways in osteoporosis

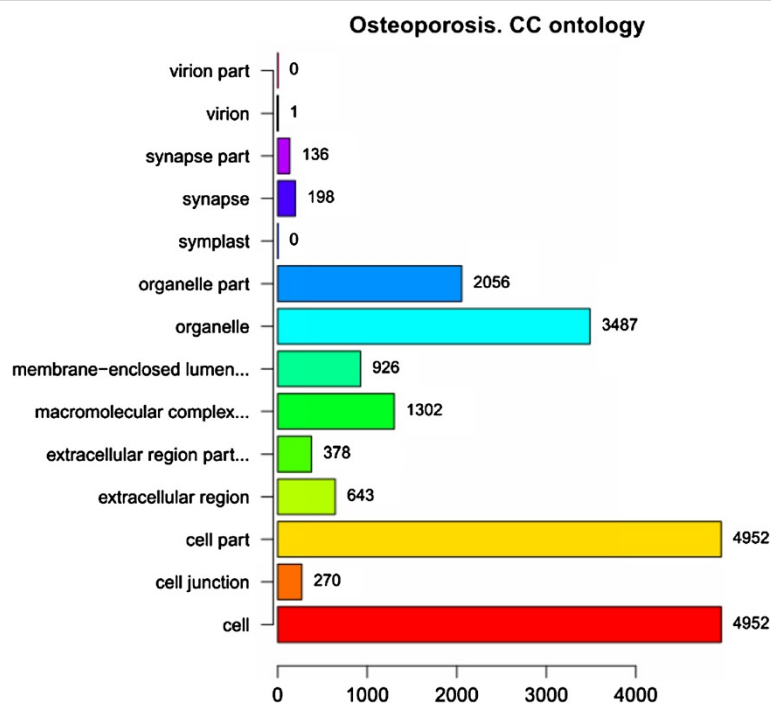
We performed pathway enrichment analysis using the online biological classification tool DAVID. A total of 9 dysfunctional pathways with p-value less than 0.01 were enriched (Table 1). The most significant dysfunctional pathway was focal adhesion with p-value = 5.04E-04. The other significant pathways included MAPK signaling pathway (p-value = 0.001953), allograft rejection (p-value = 0.003975) and DNA replication (p-value= 0.003975).

### Identification of candidate small molecules

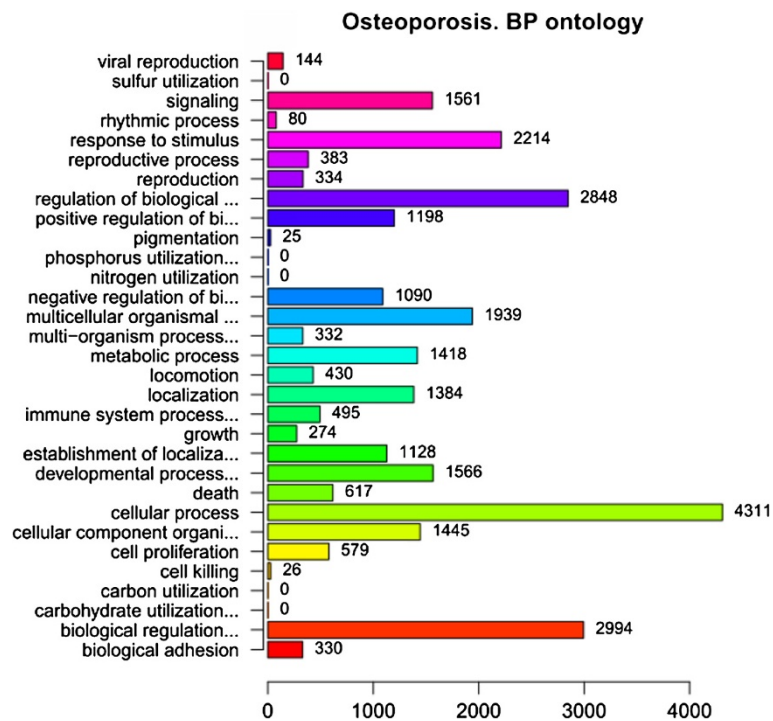
In order to identify candidate small molecules capable to reverse the gene changes of osteoporosis, we performed computational bioinformatics analysis of the derived gene signature using the Connectivity Map. The perturbagens from the CMap were analyzed according to their permuted results, p-values, and enrichment scores. A search against 6100 treatment-control pairs (instances) representing 1309 bioactive small molecules identified large amount small molecules which exhibited positive or negative correlation to the query signature. The top 20 significant small molecules were listed in Table 2. In Table 2, the small molecule of sanguinarine (enrichment score = -0.968) was associated with highly significant negative scores and the small molecule of isoflupredone was associated with highly significant positive score (enrichment score = 0.95).

### Discussion

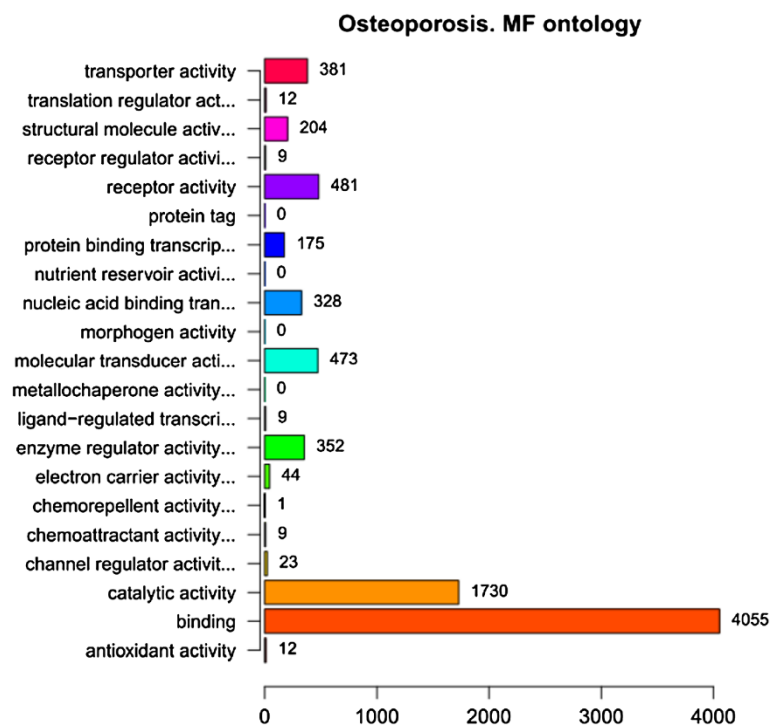
Gene expression profiling in disease reveals the underlying gene activity changes contributing to the disease and enables targets for therapeutic intervention to be identified. In this study, we investigated gene expression profile in human MSCs from patients of osteoporosis and controls, and then identified biologically active small molecules capable to reverse gene changes of osteopo-



**Figure 1** GO Enrichment of DEGs in cellular component ontology. The ordinate indicates the enriched GO categories in cellular component ontology and the abscissa indicates the number of DEGs in each category.



**Figure 2** GO Enrichment of DEGs in biological process ontology. The ordinate indicates the enriched GO categories in biological process ontology and the abscissa indicates the number of DEGs in each category.



**Figure 3** GO Enrichment of DEGs in molecular function ontology. The ordinate indicates the enriched GO categories in molecular function ontology and the abscissa indicates the number of DEGs in each category.

**Table 1 The enriched KEGG pathways**

KEGG-ID	Description	P-Value
hsa04510	Focal adhesion	5.04E-04
hsa04010	MAPK signaling pathway	0.001953
hsa05330	Allograft rejection	0.003975
hsa03030	DNA replication	0.003975
hsa05211	Renal cell carcinoma	0.004548
hsa04912	GnRH signaling pathway	0.005447
hsa04720	Long-term potentiation	0.005452
hsa04110	Cell cycle	0.007658
hsa04120	Ubiquitin mediated proteolysis	0.008467

rosis using computational bioinformatics methods. Results show that a total of 5581 genes were differentially expressed between osteoporosis and controls. In addition, we identified large amount of small molecules which can provide new ideas for the therapeutic studies in osteoporosis.

Up to 5581 genes were identified differentially expressed between osteoporosis and control in our approach. These DEGs may play critical roles in the initiation of osteoporosis, and investigation of them may shed new lights on understanding of the molecular mechanism of osteoporosis. Pathway enrichment analysis of these DEGs indicated a total of 9 pathways were

**Table 2 Small molecules capable to reverse osteoporosis**

Perturbagen	Enrichment score	p-value
viomycin	0.915	0.00004
isoflupredone	0.95	0.00012
adiphenine	0.845	0.00018
medrysone	-0.754	0.00044
luteolin	-0.874	0.00054
thiamphenicol	0.806	0.00062
sulconazole	-0.86	0.00068
apigenin	-0.857	0.00074
cinchonine	0.849	0.00074
thioguanosine	-0.852	0.00092
genistein	0.459	0.00092
roxithromycin	-0.843	0.00105
GW-8510	-0.842	0.00107
nadolol	0.827	0.00135
phthalylsulfathiazole	-0.751	0.00174
rimexolone	-0.825	0.00177
trazodone	-0.899	0.00194
sanguinarine	-0.968	0.00221
camptothecin	-0.895	0.00224
podophyllotoxin	0.807	0.00257

dysregulated in the development of osteoporosis, including focal adhesion and MAPK signaling pathway.

Focal adhesions, which are specialized sites of attachment between cells and the extracellular matrix, play a role in cell motility, cell proliferation, signal transduction and have been proposed to function as mechanosensors [24-26]. Osteoporosis is a result of an imbalance of bone formation and resorption. In osteoporosis, the regenerative capacity of bone is compromised, which may involve altered osteoblast activity. This could be attributed to an inappropriate synthesis and assembly of an extracellular matrix (ECM), altered cell adhesion to the ECM, or be due to inappropriate downstream activation of adhesion-mediated signaling cascades through proteins such as focal adhesion kinase (FAK). Perinpanayagam *et al.* suggested that early adhesion-mediated events, such as cell adhesion, attachment, and FAK signaling may be altered in osteoporotic osteoblast cells [27]. In our result, focal adhesion was the most significant dysfunctional pathways in the initiation of osteoporosis.

MAPK signaling pathways transduces a large variety of external signals, leading to a wide range of cellular responses, including growth, differentiation, inflammation and apoptosis [28]. Several studies have suggested that MAPK signaling pathways contribute greatly to osteoblast differentiation and bone formation via TGF- $\beta$  and bone morphogenic protein (BMP) signaling pathways. Lee *et al.* demonstrated that MAPK pathways converge at the Runx2 gene to control mesenchymal precursor cell differentiation following TGF- $\beta$  induction [29]. Recent study revealed that TGF- $\beta$  signaling promotes osteoprogenitor proliferation, early differentiation, and commitment to the osteoblastic lineage through the selective MAPKs pathways [30]. In addition, MAPK dependent phosphorylation, TGF- $\beta$ /BMP signaling, and Runx2 subnuclear targeting converge to induce the osteogenic phenotype [31].

The identification of a group of small molecules with potential therapeutic efficacy for osteoporosis is an important observation of our work. Data in Table 2 shows that sanguinarine (enrichment score = -0.968) was associated with highly significant negative score, suggesting that this small molecule is capable of targeting osteoporosis. Sanguinarine, a component of sanguinaria extract, has been shown to display antitumor and anti-inflammatory properties in animals [32,33] and to inhibit neutrophil function [34]. Madan *et al.* demonstrate that sanguinarine is a potent suppressor of NF- $\kappa$ B activation that blocks the phosphorylation and degradation of I $\kappa$ B $\alpha$  [35]. Recently, it was discovered that the RANK/RANKL/OPG system is an important signal transduction pathway that regulates osteoclast formation [36]. Targeting of this pathway is a novel therapeutic approach in the management of osteoporosis. Therefore,



sanguinarine may provide promising targets for the future development of novel treatments of osteoporosis. However, further evaluation for their potential use in the treatment of osteoporosis is still needed.

## Conclusion

In conclusion, we have identified a total of 9 dysfunctional pathways in the development of osteoporosis. Among them, focal adhesion and MAPK signaling pathway were the most significant ones. Besides, we have identified that sanguinarine may be a therapeutic drug candidate capable of targeting osteoporosis. Although it may be premature to suggest that this drug might be ready for clinical trials, it is clearly a direction that warrants additional consideration.

## Abbreviations

DEGs: Differentially expressed genes; MSCs: Mesenchymal stem cells; KEGG: Kyoto Encyclopedia of Genes and Genomes; CMap: Connectivity map; GO: Gene Ontology; DAVID: The Database for Annotation, Visualization and Integrated Discovery; ECM: Extracellular matrix; FAK: Focal adhesion kinase; BMP: Bone morphogenetic protein.

## Competing interests

The authors declare that they have no competing interests.

## Authors' contributions

Design and acquire the data: GY, LW. Analyze and interpret the data: GY, LW, YL, ZM. Draft and revise the manuscript: GY, YL. All authors read and approved the final manuscript.

## Author details

<sup>1</sup>Department of Orthopedic, The people's Hospital of Hengshui, No.180 Renmin Street, Hengshui 053000, Hebei Province, China. <sup>2</sup>Department of Orthopedic, The Third people's Hospital of Hebei Province, Hebei Province, China. <sup>3</sup>Clinical Laboratory, The Second people's Hospital of Tangshan, Hebei Province, China.

Received: 28 July 2012 Accepted: 31 January 2013

Published: 1 March 2013

## References

1. Kanis JA, Melton LJ 3rd, Christiansen C, Johnston CC, Khaltaev N: **The diagnosis of osteoporosis.** *J Bone Miner Res* 1994, **9**:1137-1141.
2. Huybrechts KF, Ishak KJ, Caro JJ: **Assessment of compliance with osteoporosis treatment and its consequences in a managed care population.** *Bone* 2006, **38**:922-928.
3. Ralston SH, de Crombrughe B: **Genetic regulation of bone mass and susceptibility to osteoporosis.** *Genes Dev* 2006, **20**:2492-2506.
4. Manolagas SC, Kousteni S, Jilka RL: **Sex steroids and bone.** *Recent Prog Horm Res* 2002, **57**:385-409.
5. Tomkinson A, Gevers EF, Wit JM, Reeve J, Noble BS: **The role of estrogen in the control of rat osteocyte apoptosis.** *J Bone Miner Res* 1998, **13**:1243-1250.
6. Tomkinson A, Reeve J, Shaw RW, Noble BS: **The death of osteocytes via apoptosis accompanies estrogen withdrawal in human bone.** *J Clin Endocrinol Metab* 1997, **82**:3128-3135.
7. Boyce BF, Hughes DE, Wright KR, Xing L, Dai A: **Recent advances in bone biology provide insight into the pathogenesis of bone diseases.** *Lab Invest* 1999, **79**:83-94.
8. Nuttall ME, Gimble JM: **Controlling the balance between osteoblastogenesis and adipogenesis and the consequent therapeutic implications.** *Curr Opin Pharmacol* 2004, **4**:290-294.
9. Ruiz JC, Mandel C, Garabedian M: **Influence of spontaneous calcium intake and physical exercise on the vertebral and femoral bone mineral density of children and adolescents.** *J Bone Miner Res* 1995, **10**:675-682.
10. Saadi HF, Reed RL, Carter AO, Qazaq HS, Al-Suhaili AR: **Bone density estimates and risk factors for osteoporosis in young women.** *East Mediterr Health J* 2001, **7**:730-737.
11. Marie PJ: **Strontium ranelate: a physiological approach for optimizing bone formation and resorption.** *Bone* 2006, **38**:S10-14.
12. Snieder H, MacGregor AJ, Spector TD: **Genes control the cessation of a woman's reproductive life: a twin study of hysterectomy and age at menopause.** *J Clin Endocrinol Metab* 1998, **83**:1875-1880.
13. Rizzoli R, Boonen S, Brandi ML, Burler N, Delmas P, Reginster JY: **The role of calcium and vitamin D in the management of osteoporosis.** *Bone* 2008, **42**:246-249.
14. Francis RM: **Calcium, vitamin D and involutional osteoporosis.** *Curr Opin Clin Nutr Metab Care* 2006, **9**:13-17.
15. Kostenuik PJ: **Osteoprotegerin and RANKL regulate bone resorption, density, geometry and strength.** *Curr Opin Pharmacol* 2005, **5**:618-625.
16. Raisz LG: **Pathogenesis of osteoporosis: concepts, conflicts, and prospects.** *J Clin Invest* 2005, **115**:3318-3325.
17. Benisch P, Schilling T, Klein-Hitpass L, Frey SP, Seefried L, Raaijmakers N, et al: **The transcriptional profile of mesenchymal stem cell populations in primary osteoporosis is distinct and shows overexpression of osteogenic inhibitors.** *PLoS One* 2012, **7**:e45142.
18. Kanehisa M: **The KEGG database.** *Novartis Found Symp* 2002, **247**:91-101. discussion 101-103, 119-128, 244-152.
19. Lamb J, Crawford ED, Peck D, Modell JW, Blat IC, Wrobel MJ, et al: **The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease.** *Science* 2006, **313**:1929-1935.
20. Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, Scherf U, et al: **Exploration, normalization, and summaries of high density oligonucleotide array probe level data.** *Biostatistics* 2003, **4**:249-264.
21. da Huang W, Sherman BT, Lempicki RA: **Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources.** *Nat Protoc* 2009, **4**:44-57.
22. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al: **Gene ontology: tool for the unification of biology. The Gene Ontology Consortium.** *Nat Genet* 2000, **25**:25-29.
23. Alex S, Jordi O, Miquel S: **goProfiles: goProfiles: an R package for the statistical analysis of functional profiles.** 2010.
24. BurrIDGE K, Fath K: **Focal contacts: transmembrane links between the extracellular matrix and the cytoskeleton.** *Bioessays* 1989, **10**:104-108.
25. Geiger B, Bershadsky A: **Exploring the neighborhood: adhesion-coupled cell mechanosensors.** *Cell* 2002, **110**:139-142.
26. Petit V, Thiery JP: **Focal adhesions: structure and dynamics.** *Biol Cell* 2000, **92**:477-494.
27. Perinpanayagam H, Zaharias R, Stanford C, Brand R, Keller J, Schneider G: **Early cell adhesion events differ between osteoporotic and non-osteoporotic osteoblasts.** *J Orthop Res* 2001, **19**:993-1000.
28. Froy O: **Regulation of mammalian defensin expression by Toll-like receptor-dependent and independent signalling pathways.** *Cell Microbiol* 2005, **7**:1387-1397.
29. Lee KS, Hong SH, Bae SC: **Both the Smad and p38 MAPK pathways play a crucial role in Runx2 expression following induction by transforming growth factor-beta and bone morphogenetic protein.** *Oncogene* 2002, **21**:7156-7163.
30. Matsunobu T, Torigoe K, Ishikawa M, de Vega S, Kulkarni AB, Iwamoto Y, et al: **Critical roles of the TGF-beta type I receptor ALK5 in perichondrial formation and function, cartilage integrity, and osteoblast differentiation during growth plate development.** *Dev Biol* 2009, **332**:325-338.
31. Afzal F, Pratap J, Ito K, Ito Y, Stein JL, van Wijnen AJ, et al: **Smad function and intranuclear targeting share a Runx2 motif required for osteogenic lineage induction and BMP2 responsive transcription.** *J Cell Physiol* 2005, **204**:63-72.
32. Mitscher LA, Park YH, Clark D, Clark GW 3rd, Hammesfahr PD, Wu WN, et al: **Antimicrobial agents from higher plants. An investigation of *Hunnemannia fumariaefolia* pseudoalcoholates of sanguinarine and chelerythrine.** *Lloydia* 1978, **41**:145-150.
33. Lenfeld J, Kroutil M, Marsalek E, Slavik J, Preininger V, Simanek V: **Antiinflammatory activity of quaternary benzophenanthridine alkaloids from *Chelidonium majus*.** *Planta Med* 1981, **43**:161-165.
34. Agarwal S, Reynolds MA, Pou S, Peterson DE, Charon JA, Suzuki JB: **The effect of sanguinarine on human peripheral blood neutrophil viability and functions.** *Oral Microbiol Immunol* 1991, **6**:51-61.

35. Chaturvedi MM, Kumar A, Darnay BG, Chainy GB, Agarwal S, Aggarwal BB: Sanguinarine (pseudocheleerythrine) is a potent inhibitor of NF-kappaB activation, I-kappaBalpha phosphorylation, and degradation. *J Biol Chem* 1997, **272**:30129–30134.
36. Wada T, Nakashima T, Hiroshi N, Penninger JM: RANKL-RANK signaling in osteoclastogenesis and bone disease. *Trends Mol Med* 2006, **12**:17–25.

doi:10.1186/2047-783X-18-5

**Cite this article as:** Yu et al.: Identification of drug candidate for osteoporosis by computational bioinformatics analysis of gene expression profile. *European Journal of Medical Research* 2013 **18**:5.

**Submit your next manuscript to BioMed Central  
and take full advantage of:**

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at  
[www.biomedcentral.com/submit](http://www.biomedcentral.com/submit)

