Searching for perturbed biological pathways and genes through analyzing the expression profile changes in osteoclasts after treatment by bisphosphonates

LONG ZHAO[1], BIN ZHANG[2], FENG WU[3] and XUAN-HUANG CHEN[4]

[1]Department of Orthopedics, The First Hospital of Lanzhou University, Lanzhou, Gansu 730000;
[2]Department of Orthopedics, Chinese Medicine Hospital of Xi'an City, Xi'an, Shanxi 710021;
Departments of [3]Rehabilitation Medicine and [4]Orthopedics, The Affiliated Hospital
of Putian University, Putian, Fujian 351100, P.R. China

Abstract. Criticality pathways and genes related to osteoporosis were identified. We downloaded the expression data of osteoclasts treated with or without bisphosphonates and all human pathways from the public database. Gibbs sampling and Markov chain were performed to identify the disturbed pathways and the hub genes in the disturbed pathways. Pathways and genes with adjusted probability (α_{adj}) ≥0.75 were considered as the disturbed pathways and hub genes. We identified four disturbed pathways (Maturity onset diabetes of the young, Olfactory transduction, Cyanoamino acid metabolism, Taurine and hypotaurine metabolism) and two hub genes (*OR2A4* and *NKX2-2*) with α_{adj} ≥0.75. The expression levels of these disturbed pathways and hub genes were downregulated in bisphosphonates group. In conclusion, four disturbed pathways and two hub genes related to osteoporosis were identified. These results give us a better understanding of the potential mechanism of bisphosphonate treatment and the pathogenesis of osteoporosis.

Introduction

Osteoporosis is a common metabolic bone disease characterized by decreased bone mass, reduced bone density and destruction of the microstructure of the bone, which leads to consequent increase in bone fragility (1). Its occurrence and development is an extremely complicated biological process involves in multiple factors, multi-genes and multi-stage experiences (2-4). Although some osteoporosis-related genes

and pathways have been revealed by many scholars (5,6), the underlying pathogenesis of osteoporosis still remains obscure.

Bisphosphonates (BO), which can inhibit bone resorption, are currently the most commonly used and effective drugs in the treatment of osteoporosis. However, it has been reported to have severe side effects, such as bisphosphonate-related osteonecrosis of the jaw (7). Therefore, understanding the molecular mechanism of osteoporosis, identifying important pathways and genes and taking these signaling pathways as targeting points are important strategies to develop new anti-osteoporosis drugs or prevent osteoporosis. Current advances in high-throughput experimental techniques are of great help in accelerating the identification of the key genes and pathways involved in osteoporosis (8,9).

The Markov chain is a kind of stochastic process with the so-called 'Markov property'. The process has the following characteristics: The next state only depends on the current state, not related to the previous state. Markov chain integral cleverly disassembles analytic integrals into a probabilistic expectation problem and obtains approximate expectations through a large number of samples. Gibbs sampling is a very popular Bayesian method (10). It is assumed that the subjects studied have a certain understanding before sampling, and the prior distribution is often used to describe this understanding. Then, based on the extracted samples, the prior knowledge is corrected to obtain the posterior distribution. The inference of various statistics are based on the posterior distribution (11). According to the probability distributions, researchers can find disturbed pathways and genes in the pathology of complex diseases. In this study, we attempted to predict the criticality pathways and genes related to osteoporosis by Markov methods and Gibbs sampling.

Materials and methods

Data collection. The transcriptomic profiles of osteoclasts treated with bisphosphonates (E-GEOD-63009) were obtained from the ArrayExpress database (http://www.ebi.ac.uk/arrayexpress/). A total of 9 samples were selected for analysis. These samples were divided into two groups: vehicle group (3 samples)

Correspondence to: Dr Xuan-Huang Chen, Department of Orthopedics, The Affiliated Hospital of Putian University, 999 East Dongzhen Road, Licheng, Putian, Fujian 351100, P.R. China
E-mail: whn5156@163.com

and bisphosphonates group (6 simples). After standard prepro-cessing (12,13), we converted the expression profile from probe level to gene symbol level, and removed the duplicated symbols. Finally, a total of 20,514 genes were obtained in each group.

Metabolic pathway enrichment. For the discovery of disturbed pathways associated with osteoclasts, all human pathways were obtained from the Kyoto Encyclopedia of Genes and Genomes (KEGG) database (14), including 287 pathways and 6,894 genes. The expression profiling genes were enriched to the KEGG pathways. The KEGG pathways, with gene intersections ≥5, were screened out and 280 pathways were obtained.

Pathway initial state build. According to the enrichment of gene in each pathway, we calculated the average value of gene expression of each pathway under vehicle condition (state 1) and bisphosphonate condition (state 2), which is regarded as the expression of this pathway. State 1 was considered as the initial state of the pathway and state 2 was considered as a priori value of the path.

According to Markov chain theory, the initial transition probability is obtained from the expression of state 1 and state 2 in the access system, and state three is derived from state 2. State n is sequentially deduced, and state n is only related to the state n-1. The original state of the system has no relation with other Markov processes.

Gibbs sampling. To obtain a Markov chain, Gibbs sampling was performed. An empty set of Gibbs sampling was defined first, and the initial state and priori value data sets were put into the empty set for Gibbs sampling. A k-dimensional vector (k = 280, 280 KEGG pathways), was initialized one of the elements was extracted in turn and the posterior value of the element was calculated. This process was looped k times to generate a new k-dimensional vector, and defined as state 3. After 10,000 times (n=10,000) Gibbs sampling, a k-dimen-sional vector with 10,000 samples is constructed to obtain a Markov chain, and it is considered that the samples generated from 2,000 begin to follow the distribution of the real samples.

Analysis of the disturbed pathways. Through Gibbs sampling, the posterior probability of each pathway was obtained. The probability (α) of occurrence of each pathway was defined as formula 1.

$$\alpha = \frac{\sum_{i=2000}^{10000} P_i}{10000 - 2000 + 1}$$

where P_i represents the posterior value of this pathway in the ith sample. Based on the gene expression values under vehicle condition and bisphosphonate condition, the pathway expres-sion differences between these two conditions were calculated by t-test. Next, the pathways were ranked according to the P-values. Combining the rank and values and α, correction coefficient (*c*) was calculated as formula 2.

$$c = 1 - \frac{rank_i}{n}$$

In formula 2, *n*, represents the number of pathways, $rank_i$ represents the rank of pathway *i*.

Table I. The top 5 genes ranked according to α_{adj}.

Gene	P-value	Rank p	R-value	α	α_{adj}
OR2A4	0.012	2	0.992	0.999	0.992
NKX2-2	0.041	8	0.970	0.876	0.849
CLCA1	0.119	27	0.897	0.829	0.744
PAX6	0.038	7	0.973	0.763	0.743
OR10H2	0.131	31	0.882	0.802	0.707

Subsequently, the adjusted probability (α_{adj}) was calculated. Pathways with $\alpha_{adj} \geq 0.75$ were screened out and considered as the differential pathways.

Identifying key genes. After finding the key pathways, we identified the hub genes from these pathways using Gibbs sampling. The genes in the key pathways were transformed into Markov chains in order to identify key genes by Gibbs sampling. Similar with the identification of key pathways, the posterior probability of each gene in the key pathways was obtained and the probability of each gene was calculated using formula one, where P_i represents the posterior value of this gene in the i$_{th}$ sample. Similarly, gene expression variation was taken into account to adjust the probability. According to the expression level of the genes in the key pathways in different states, the gene expression differences between vehicle group and bisphosphonates group were calculated by Student's t-test. Subsequently, the adjusted probability of the genes in the key pathways was calculated and the genes were ranked according to the adjusted probability. Genes with $\alpha_{adj} \geq 0.75$ were screened out and considered as hub genes.

Results

Data preprocessing and pathway enrichment. The expression data sets of osteoclasts were acquired from the ArrayExpress database, containing 6 samples treated with bisphosphonates and 3 samples not treated by bisphosphonates. After prepro-cessing, a total of 20,514 genes were obtained for the next analysis. Among 287 human pathways obtained from KEGG database, 280 pathways with at least 5 genes were selected for the next analysis.

Perturbed pathways in bisphosphonate-treated osteoclasts. Based on the preprocessed expression data and the screened pathways, the posterior probability of each pathway was obtained by Gibbs sampling. The probability α of occurrence of each pathway was calculated and adjusted. The probability distribution of the 280 pathways is shown in Fig. 1. Under the criterion of $\alpha_{adj} \geq 0.75$, we obtained 4 pathways in total. The 4 pathways are Maturity onset diabetes of the young (hsa04950, $\alpha_{adj}=0.909$), Olfactory transduction (hsa04740, $\alpha_{adj}=0.874$), Cyanoamino acid metabolism (hsa00460, $\alpha_{adj}=0.859$), Taurine and hypotaurine metabolism (hsa00430, $\alpha_{adj}=0.756$), respec-tively. The expression levels of hsa04950, hsa04740, hsa00460 and hsa00430 in vehicle group and bisphosphonates group are shown in Fig. 2. It was obviously observed that the expression
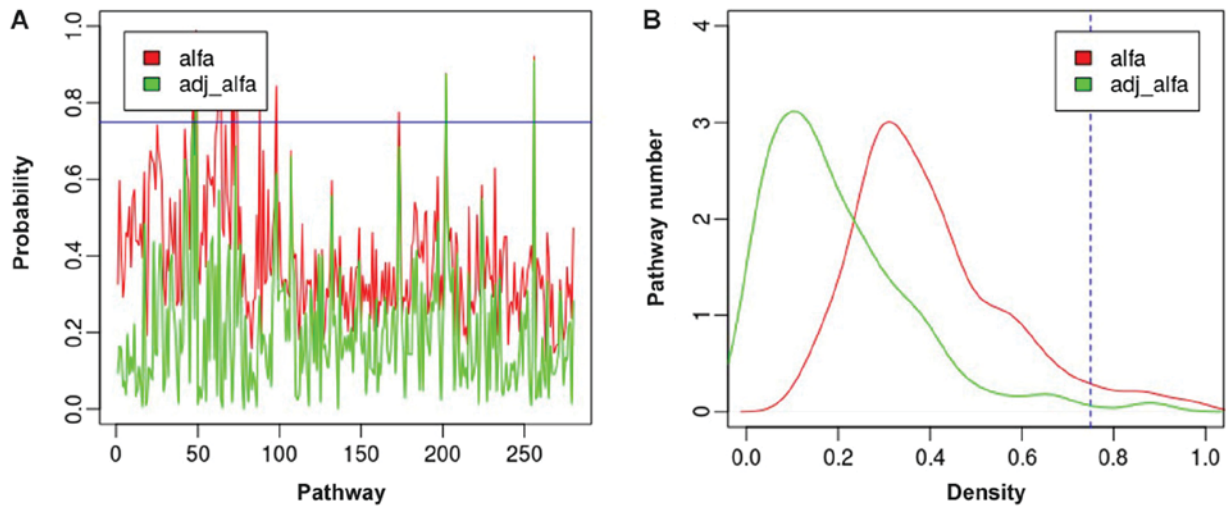
Figure 1. The probabilities and density distribution of pathways. (A) The probability distribution of all human pathways. (B) The density distribution of pathways. The blue line represents the threshold value of posterior value of 0.75, the red line shows the initial probabilities, and the green line shows the adjusted probabilities.
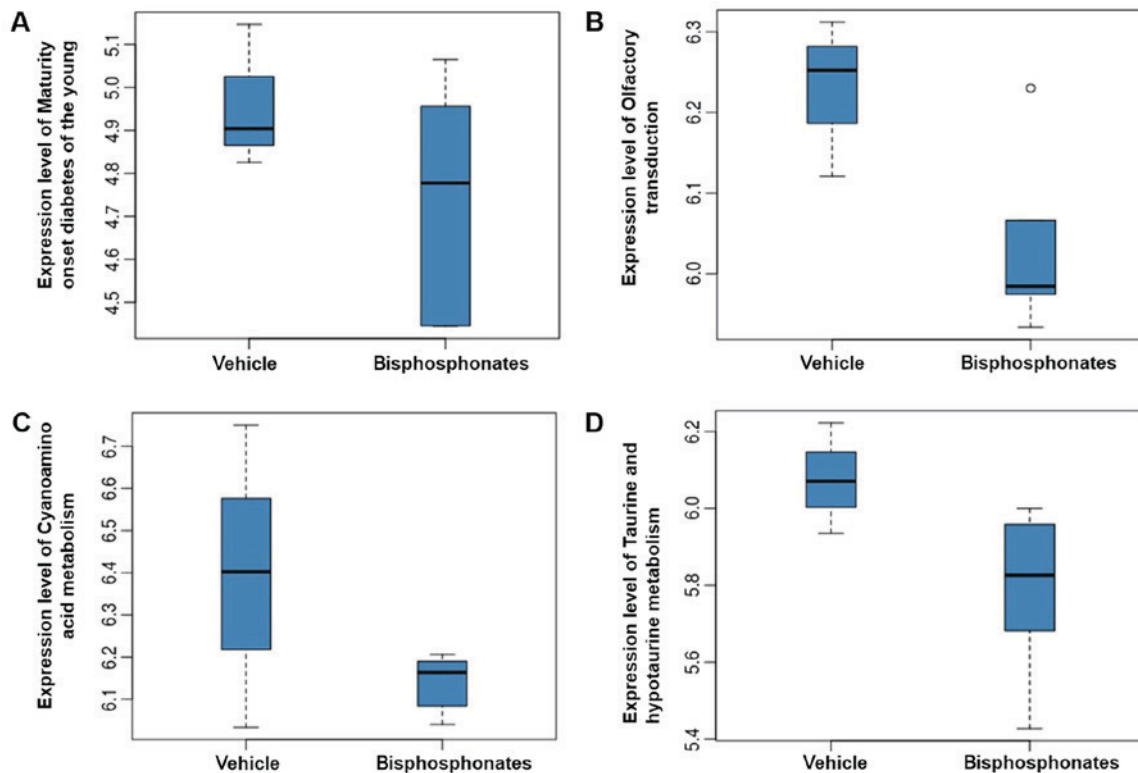

Figure 2. The expression levels of disturbed pathways. The expression levels of (A) Maturity onset diabetes of the young, (B) Olfactory transduction, (C) Cyanoamino acid metabolism, and (D) Taurine and hypotaurine metabolism in vehicle and bisphosphonates groups.

levels of these four pathways were decreased in bisphosphonate groups relative to the vehicle groups.

Perturbed genes in bisphosphonate-treated osteoclasts. In order to identify the hub genes in the key pathways we obtained previously, Gibbs sampling was performed again. The genes in the key pathways were analyzed and a total of 263 genes were obtained for the next analysis. After Gibbs sampling, the probability (α) of occurrence of each gene in these 4 pathways was calculated and adjusted as described in the methods. The probability distribution of the 263 genes is shown in Fig. 3. The top 5 genes ranked according to α_{adj} are listed in Table I.

The genes with $\alpha_{adj} \geq 0.75$ were selected as hub genes. Finally, two hub genes related to osteoporosis were screened out, including *OR2A4* ($\alpha_{adj}=0.992$) and *NKX2-2* ($\alpha_{adj}=0.849$). The expression levels of these two genes in vehicle group and bisphosphonates group are shown in Fig. 4A and B. We found that both the expression of *OR2A4* and *NKX2-2* decreased in bisphosphonate group compared to vehicle group. The
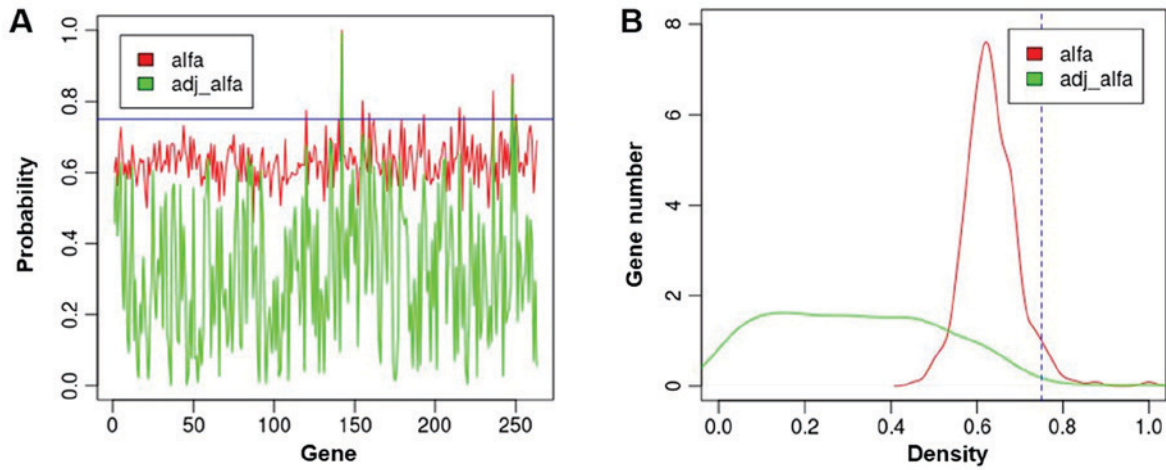
Figure 3. The probabilities and density distribution of genes in the disturbed pathways. (A) The probability distribution of genes in the disturbed pathways. (B) The density distribution of genes in the disturbed pathways. The blue line represents the threshold value of posterior value of 0.75, the red line shows the initial probabilities, and the green line shows the adjusted probabilities.
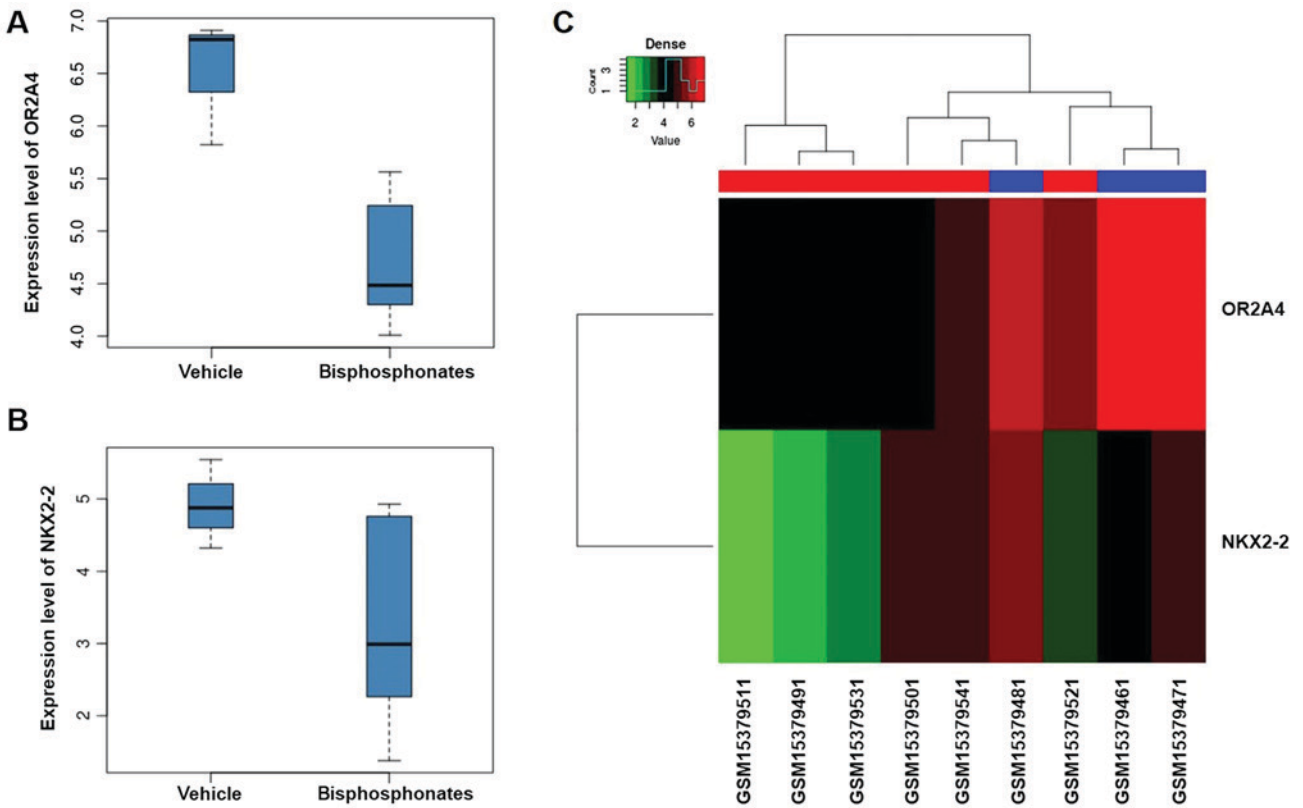

Figure 4. The expression levels of disturbed genes. The expression levels of (A) *OR2A4* and (B) *NKX2-2* in vehicle and bisphosphonate groups. (C) Heatmap results of hub genes. Blue represents the expression of hub genes in vehicle group and red represents the expression of hub genes in bisphosphonate group.

heatmap of two hub genes is shown in Fig. 4C. We found that the distribution of these 2 genes in 9 samples was consistent with the results observed above.

Discussion

In this study, four perturbed biological pathways (Maturity onset diabetes of the young, Olfactory transduction, Cyanoamino acid metabolism, Taurine and hypotaurine metabolism), that may be associated with osteoporosis were

identified by Gibbs sampling. Among these 4 pathways, Maturity onset diabetes of the young (MODY) pathway showed the highest α_{adj} value (0.909). MODY is a monogenic form of type 2 diabetes, with the characteristic of early onset and autosomal dominant inheritance (15). Six genes are known to cause MODY, including *HNF4α* (*MODY1*), *HNF1α* (*MODY3*), *PDX1* (*MODY4*), *HNF1β* (*MODY5*), *NEUROD1* (*MODY6*) and glucokinase catalyzes (16). In bisphosphonates treated osteoclasts, we found that the levels of *HNF4α*, *PDX1* (*MODY4*) and *HNF1β* had almost no change, while the levels

of *HNF1α* and *NEUROD1* were upregulated. It has been reported that diabetes mellitus can be complicated by osteoporosis which called diabetic osteoporosis, but its correlation has not been widely recognized (17). Many researchers have demonstrated that in type 1 diabetes patients, the bone mass was reduced and fracture risk was increased. However, in type 2 diabetes, this risk is controversial (17). Some scholars have suggested that type 2 diabetes patients may possess a reduced risk of osteoporosis with an increased bone mineral density and body weight (18). While, recent studies have demonstrated that fracture risk increased in type 2 diabetes patients, even though the bone mineral density increased or was independent (19-21). Here, we found that MODY pathway was disturbed in osteoclasts treated with bisphosphonates, indicating that this pathway was associated with the activity of osteoclasts and played critical roles in the development of osteoporosis.

Furthermore, we identified two hub genes (*OR2A4* and *NKX2-2*) that may be associated with osteoporosis using Gibbs sampling. Among these two genes, *NKX2-2* was involved in MODY pathway and gained the second highest $α_{adj}$ value 0.849. NKX2-2 (named as NKX2.2 or NKX2B), contains a homeobox domain and is a member of NK2 family (22). This protein has been suggested to be involved in the neuronal development (23) and the differentiation of pancreatic β cells. Sussel *et al* have reported that lacking NKX2.2 caused an incompletely differentiated state of pancreatic β cells and NKX2.2 null mice suffered from diabetes immediately after birth due to lack of insulin (24). Hence NKX2.2 played a critical role in the development of diabetes which have a certain relationship with osteoporosis. Furthermore, Smith *et al* reported that inhibiting the expression of NKX2.2 by RNAi could suppress oncogenesis of Ewing's sarcoma (25). This observation implied that NKX2.2 has the potential to serve as a therapeutic target for Ewing's sarcoma which is a malignant bone tumor composed of small round cells (25). Although the exact relationship between MODY pathway and *NKX2-2* gene and osteoporosis has not been reported yet, the decreased expression of MODY pathway and *NKX2-2* in bisphosphonate treated osteoclasts indicating they may play critical roles in osteoporosis development.

OR2A4 (olfactory receptor, family 2, subfamily A, member 4) which gained the highest $α_{adj}$ value 0.992, is a member of olfactory receptors that involves in olfactory transduction pathway. Olfactory receptors are expressed in a series of tissues. Tsai *et al* have demonstrated that *OR2A4* was expressed in human skin cells and affected cytokinesis and cell proliferation (26). To this date, few studies have demonstrated the direct association between olfactory transduction pathway/olfactory receptors and osteoporosis. Besides, the relationships between osteoporosis and the other two pathways we identified in this study (Cyanoamino acid metabolism pathway and Taurine and hypotaurine metabolism pathway) also have limited studies. Further investigation may reveal new mechanisms of osteoporosis and provide new targets for osteoporosis treatment.

In conclusion, we found 4 disturbed pathways (Maturity onset diabetes of the young, Olfactory transduction, Cyanoamino acid metabolism, Taurine and hypotaurine metabolism) and two hub genes (*OR2A4* and *NKX2-2*) in

osteoclasts treated with bisphosphonates using Gibbs sampling. These results give us a better understanding of the potential mechanism of bisphosphonate treatment of osteoporosis and the pathogenesis of osteoporosis.

Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Authors' contributions

LZ conceived the study and drafted the manuscript. BZ and FW acquired the data; LZ, BZ and FW analyzed the data. XHC contributed to the analysis of the data and revised the manuscript. All authors read and approved the final manuscript.

Ethics approval and consent to participate

Not applicable.

Patient consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

1. Maraka S and Kennel KA: Bisphosphonates for the prevention and treatment of osteoporosis. BMJ 351: h3783, 2015.
2. Ralston SH and Uitterlinden AG: Genetics of osteoporosis. Endocr Rev 31: 629-662, 2010.
3. Pietschmann P, Rauner M, Sipos W and Kerschan-Schindl K: Osteoporosis: An age-related and gender-specific disease - a mini-review. Gerontology 55: 3-12, 2009.
4. Seeman E: Bone quality: The material and structural basis of bone strength. J Bone Miner Metab 26: 1-8, 2008.
5. Thouverey C and Caverzasio J: Focus on the p38 MAPK signaling pathway in bone development and maintenance. Bonekey Rep 4: 711, 2015.
6. Abu-Amer Y: NF-kappaB signaling and bone resorption. Osteoporos Int 24: 2377-2386, 2013.
7. Mukudai Y, Kondo S, Koyama T, Li C, Banka S, Kogure A, Yazawa K and Shintani S: Potential anti-osteoporotic effects of herbal extracts on osteoclasts, osteoblasts and chondrocytes in vitro. BMC Complement Altern Med 14: 29, 2014.
8. Zhou Z, Gao M, Liu Q and Tao MDJ: Comprehensive transcriptome analysis of mesenchymal stem cells in elderly patients with osteoporosis. Aging Clin Exp Res 27: 595-601, 2015.
9. van Zoelen EJ, Duarte I, Hendriks JM and van der Woning SP: TGFβ-induced switch from adipogenic to osteogenic differentiation of human mesenchymal stem cells: Identification of drug targets for prevention of fat cell differentiation. Stem Cell Res Ther 7: 123, 2016.

10. Lavine M and West M and Praeger A: A Bayesian method for classification and discrimination. Can J Stat 20: 451-461, 1992.
11. Mclachlan GJ and Krishnan T: The EM algorithm and extensions. 2nd edition. John Wiley & Sons, Hoboken, NJ, 2007.
12. Bolstad BM, Irizarry RA, Astrand M and Speed TP: A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. Bioinformatics 19: 185-193, 2003.
13. Irizarry RA, Bolstad BM, Collin F, Cope LM, Hobbs B and Speed TP: Summaries of affymetrix GeneChip probe level data. Nucleic Acids Res 31: e15, 2003.
14. Aoki-Kinoshita KF and Kanehisa M: Gene annotation and pathway mapping in KEGG. Methods Mol Biol 396: 71-91, 2007.
15. Yamagata K, Furuta H, Oda N, Kaisaki PJ, Menzel S, Cox NJ, Fajans SS, Signorini S, Stoffel M and Bell GI: Mutations in the hepatocyte nuclear factor-4alpha gene in maturity-onset diabetes of the young (MODY1). Nature 384: 458-460, 1996.
16. Winckler W, Weedon MN, Graham RR, McCarroll SA, Purcell S, Almgren P, Tuomi T, Gaudet D, Boström KB, Walker M, *et al*: Evaluation of common variants in the six known maturity-onset diabetes of the young (MODY) genes for association with type 2 diabetes. Diabetes 56: 685-693, 2007.
17. Leidig-Bruckner G, Grobholz S, Bruckner T, Scheidt-Nave C, Nawroth P and Schneider JG: Prevalence and determinants of osteoporosis in patients with type 1 and type 2 diabetes mellitus. BMC Endocr Disord 14: 33, 2014.
18. Heath H III, Melton LJ III and Chu CP: Diabetes mellitus and risk of skeletal fracture. N Engl J Med 303: 567-570, 1980.
19. Janghorbani M, Van Dam RM, Willett WC and Hu FB: Systematic review of type 1 and type 2 diabetes mellitus and risk of fracture. Am J Epidemiol 166: 495-505, 2007.
20. Schwartz AV, Sellmeyer DE, Ensrud KE, Cauley JA, Tabor HK, Schreiner PJ, Jamal SA, Black DM and Cummings SR; Study of Osteoporotic Features Research Group: Older women with diabetes have an increased risk of fracture: A prospective study. J Clin Endocrinol Metab 86: 32-38, 2001.
21. Strotmeyer ES, Cauley JA, Schwartz AV, Nevitt MC, Resnick HE, Zmuda JM, Bauer DC, Tylavsky FA, de Rekeneire N, Harris TB, *et al*: Diabetes is associated independently of body composition with BMD and bone volume in older white and black men and women: The Health, Aging, and Body Composition Study. J Bone Miner Res 19: 1084-1091, 2004.
22. Lessnick SL and Owen LA: NKX2-2 (NK2 homeobox 2). Atlas Genet Cytogenet Oncol Haematol 12: 375-376, 2008.
23. Briscoe J, Sussel L, Serup P, Hartigan-O'Connor D, Jessell TM, Rubenstein JL and Ericson J: Homeobox gene Nkx2.2 and specification of neuronal identity by graded Sonic hedgehog signalling. Nature 398: 622-627, 1999.
24. Sussel L, Kalamaras J, Hartigan-O'Connor DJ, Meneses JJ, Pedersen RA, Rubenstein JL and German MS: Mice lacking the homeodomain transcription factor Nkx2.2 have diabetes due to arrested differentiation of pancreatic beta cells. Development 125: 2213-2221, 1998.
25. Smith R, Owen LA, Trem DJ, Wong JS, Whangbo JS, Golub TR and Lessnick SL: Expression profiling of EWS/FLI identifies NKX2.2 as a critical target gene in Ewing's sarcoma. Cancer Cell 9: 405-416, 2006.
26. Tsai T, Veitinger S, Peek I, Busse D, Eckardt J, Vladimirova D, Jovancevic N, Wojcik S, Gisselmann G, Altmüller J, *et al*: Two olfactory receptors-OR2A4/7 and OR51B5-differentially affect epidermal proliferation and differentiation. Exp Dermatol 26: 58-65, 2017.