Pathway-Based Genome-Wide Association Analysis Identified the Importance of Regulation-of-Autophagy Pathway for Ultradistal Radius BMD

Lishu Zhang,1,2 Yan-Fang Guo,2,3 Yao-Zhong Liu,2 Yong-Jun Liu,2 Dong-Hai Xiong,2 Xiao-Gang Liu,3 Liang Wang,2 Tie-Lin Yang,3 Shu-Feng Lei,2 Yan Guo,2 Han Yan,3 Yu-Fang Pei,3 Feng Zhang,3 Christopher J Papasian,2 Robert R Recker,4 and Hong-Wen Deng1,2,3

1Institute of Bioscience and Biotechnology, School of Science, Beijing Jiaotong University, Beijing, People’s Republic of China
2Departments of Orthopedic Surgery and Basic Medical Sciences, University of Missouri–Kansas City, Kansas City, MO, USA
3The Key Laboratory of Biomedical Information Engineering, Ministry of Education and Institute of Molecular Genetics, School of Life Science and Technology, Xi’an Jiaotong University, Xi’an, People’s Republic of China
4Osteoporosis Research Center, Creighton University, Omaha, NE, USA

ABSTRACT

Wrist fracture is not only one of the most common osteoporotic fractures but also a predictor of future fractures at other sites. Wrist bone mineral density (BMD) is an important determinant of wrist fracture risk, with high heritability. Specific genes underlying wrist BMD variation are largely unknown. Most published genome-wide association studies (GWASs) have focused only on a few top-ranking single-nucleotide polymorphisms (SNPs)/genes and considered each of the identified SNPs/gene sets independently. To identify biologic pathways important to wrist BMD variation, we used a novel pathway-based analysis approach in our GWAS of wrist ultradistal radius (UD) BMD, examining approximately 500,000 SNPs genome-wide from 984 unrelated whites. A total of 963 biologic pathways/gene sets were analyzed. We identified the regulation-of-autophagy (ROA) pathway that achieved the most significant result ($p = 0.005, q_{fdr} = 0.043, p_{fwer} = 0.016$) for association with UD BMD. The ROA pathway also showed significant association with arm BMD in the Framingham Heart Study sample containing 2187 subjects, which further confirmed our findings in the discovery cohort. Earlier studies indicated that during endochondral ossification, autophagy occurs prior to apoptosis of hypertrophic chondrocytes, and it also has been shown that some genes in the ROA pathway (e.g., INFG) may play important roles in osteoblastogenesis or osteoclastogenesis. Our study supports the potential role of the ROA pathway in human wrist BMD variation and osteoporosis. Further functional evaluation of this pathway to determine the mechanism by which it regulates wrist BMD should be pursued to provide new insights into the pathogenesis of wrist osteoporosis. © 2010 American Society for Bone and Mineral Research.

KEY WORDS: OSTEOPOROSIS; BONE MINERAL DENSITY; GENOME-WIDE ASSOCIATION; REGULATION OF AUTOPHAGY; WHITES

Introduction

Osteoporosis (MIM 166710) is a systemic skeletal disease characterized by low bone mass and microarchitectural deterioration of bone tissue with a consequent increase of bone fragility and susceptibility to fractures. Osteoporosis has become one of the most serious public health problems around the world, causing millions of fractures annually. Fractures at the distal radius of the wrist, also known as Colles’ fractures, are the most common fractures in women younger than 75 years of age. Most Colles’ fractures are the combined result of a low-energy fall and loss of wrist bone mass owing to osteoporosis. Bone mineral density (BMD) at the distal forearm is the main determinant of susceptibility to Colles’ fracture.

BMD, the proxy measure for osteoporosis, has a strong genetic determination. The results of twin and family studies indicate that genetics account for 60% to 90% of the variance in BMD at multiple skeletal sites, including the wrist. BMD variation at different skeletal sites has been shown to share some genetic factors. Since wrist fractures were found to be predictive for...
fractures at other skeletal sites,(7–9) the wrist, especially the distal radius, may be a useful site for early assessment of the overall osteoporotic fracture risk.

With the emergence of genome-wide linkage disequilibrium (LD)-based marker panels and the improvements in high-throughput genotyping technology, genome-wide association studies (GWASs) have become feasible. GWASs can systematically survey the whole genome for causal genetic variants for complex traits/diseases and are powerful tools for dissecting the genetic basis for osteoporosis. Several promising candidate genes for osteoporosis or related traits have been identified by GWASs.(10–13)

Most published GWAs consider the effects of genetic markers or haplotypes independently and focus only on genes or markers of top-ranking statistical significance. Genes and/or their products often work together, however, interacting in functional groups or pathways to contribute to phenotypic variation or susceptibility to a disease. Therefore, the approach just outlined, which focuses on individual genes, may not be optimally effective for identifying pathophysiologically significant genetic pathways underlying complex traits or diseases such as osteoporosis. Recently, motivated by the availability of gene set enrichment analysis (GSEA) for gene expression data, Wang and colleagues(14) developed a pathway-based genome-wide association (GWA) analysis method that uses up-to-date knowledge of functional interactions among genes in biologic pathways. With this method, multiple genes in a related pathway can be considered jointly, and association with a group of genes as a whole (pathway) can be detected efficiently, even though some individual genes may have only moderate association evidence. The advantage of this approach was illustrated in its application to two Parkinson disease (PD) GWA data sets and one amyo-trophic lateral sclerosis (ALS) data set.(14)

To identify novel biologic pathways contributing to wrist osteoporosis and Colles’ fractures, we performed a pathway-based GWA analysis on BMD variation at the wrist ultradistal radius (UD) using the method proposed by Wang and colleagues. We identified a significant association, at the pathway-wise level, between UD BMD and the regulation-of-autophagy (ROA) pathway. Our findings support the potential role of the ROA pathway in the pathogenesis of wrist osteoporosis.

Materials and Methods

Subjects

Discovery GWAS sample: This study was approved by the necessary Institutional Review Boards of all the involved institutions. A total of 984 unrelated US whites of European origin were chosen from our established genetic repertoire currently containing more than 10,000 subjects. Signed informed-consent documents were obtained from all study participants before they entered the study.

People with chronic diseases and conditions that potentially might affect bone mass, structure, or metabolism were excluded from the study. These diseases/conditions included chronic disorders involving vital organs (eg, heart, lung, liver, kidney, and brain), serious metabolic diseases (eg, diabetes, hypo- and hyperparathyroidism, hyperthyroidism, etc.), other skeletal diseases (eg, Paget disease, osteogenesis imperfecta, rheumatoid arthritis, etc.), chronic use of drugs affecting bone metabolism (eg, hormone-replacement therapy, corticosteroid therapy, and anticonvulsant drugs), and malnutrition conditions (eg, chronic diarrhea, chronic ulcerative colitis, etc.). In addition, subjects taking anti-bone-resorptive or bone anabolic agents/drugs such as bisphosphonates were excluded from the study.

For all subjects, measurement of anthropometric variables was performed, and a structured questionnaire addressing lifestyles and medical history was administered. Areal BMD (g/cm²) at the UD region of the wrist was measured by dual-energy X-ray absorptiometry (DXA) with Hologic QDR 4500W densitometers (Hologic, Inc., Bedford, MA, USA). The scanners were calibrated daily, and long-term precision was monitored with external phantoms. The coefficient of variation (CV%) of wrist UD BMD measurements obtained on the Hologic scanner was approximately 2.3%.

FHS replication sample: For replication of our GWAS findings, we used a sample from the Framingham Heart Study (FHS). The FHS is a community-based, multigenerational, longitudinal study of cardiovascular disease and its risk factors, with one substudy focusing on osteoporosis. The phenotype and genotype information of the cohort was downloaded from the Framingham SNP Health Association Resource (SHARe), accessed through the database of Genotypes and Phenotypes (dbGaP), developed by the National Center for Biotechnology Information (NCBI) (http://view.ncbi.nlm.nih.gov/dbgap). Appropriate procedures have been taken for the use of the data, which include approval from the University of Missouri–Kansas City Institutional Review Board and signatures on the Data Distribution Agreement by all the UMKC investigators who have access to the data.

The FHS comprises the original cohort, offspring, and generation 3 studies. Subjects used as the replication sample in this analysis are those who participated in exam 23 of the first generation and exams 6 and 7 of the second generation with arm BMD as well as other covariates, including age, weight, height, and sex, available. The arm BMD was scanned by DXA (Lunar DPX-L, Lunar Co., Madison, WI, USA). The detailed protocol is available through the NCBI dbGaP. In total, the sample contains 2187 subjects (817 males and 1370 females) from 477 families.

Genotyping

Discovery GWAS sample: Genomic DNA was extracted from whole human blood using a commercial isolation kit (Genta Systems, Minneapolis, MN, USA) following the protocols detailed in the kit. DNA concentration was assessed by a DUS530 UV/VIS spectrophotometer (Beckman Coulter, Inc., Fullerton, CA, USA). Genotyping with the Affymetrix Mapping 250K Nsp and Affymetrix Mapping 250K Sty arrays was performed at the Vanderbilt Microarray Shared Resource at Vanderbilt University Medical Center (Nashville, TN, USA) using the standard protocol recommended by the manufacturer (Affymetrix, Inc., Santa Clara, CA, USA). Briefly, for each array, 250 ng of genomic DNA was digested with either Nsp1 or Sty1 and ligated to adapters that recognize the cohesive 4-base-pair (bp) overhangs. A generic primer that recognizes the ligated adapter sequence was used to amplify the ligation products in a polymerase chain reaction.

PATHWAY-BASED GWAS FOR WRIST BMD
Bayesian step (B-RLMM) algorithm. According to Affymetrix’s Linear Model with Mahalanobis Distance Classifier with a reached greater than 95%. The B-RLMM algorithm was used for control criterion, and the average call rate for all the subjects with DM call rates of less than 93% were subject to guidelines, DM calls were used for quality control. Specifically,otyping method, please refer to the Framingham SHARe at the Affymetrix 50K supplementary array. For details of the genot- 

generated with the Affymetrix 500K mapping array plus the pathways were retrieved from the BioCarta pathway database, (www.geneontology.org). A total of 260, 190, and 380 annotated gene sets and pathways to be tested in this study, which are the BioCarta pathway database (www.biocarta.com/annotated.html), the Ambion GeneAssist Pathway Atlas (www.ambion.com/tools/pathway), and the Gene Ontology (GO) database (www.geneontology.org). A total of 260, 190, and 380 annotated pathways were retrieved from the BioCarta pathway database, KEGG pathway database, and Ambion GeneAssist Pathway Atlas, respectively. To integrate the GO information into our study, GO annotation files for human genes also were downloaded. We processed the GO annotation files and generated gene sets on the basis of GO level 4 annotations in biologic process and molecular function. Genes whose GO annotations are in level 5 or lower in the GO hierarchy are assigned to their ancestral GO annotations in level 4. In our analysis, we tested only pathways/gene sets in which at least 85% (and in number 10 to 200) of the genes were included in our GWA data set so as to alleviate the multiple-testing problem by avoiding testing too narrowly or too broadly defined functional categories. Overall, 963 pathways/gene sets were analyzed.

Statistical analysis

Analyses in the discovery sample: To detect population stratification that may lead to spurious association results, we used Structure 2.2 software (http://pritch.bsd.uchicago.edu/software.html) and EIGENSOFT 2.0 software (http://genepath.med.harvard.edu/~reich/EIGENSTRAT.html) to investigate the potential substructure/stratification of our sample. The Structure 2.2 program uses a Markov chain Monte Carlo (MCMC) algorithm to cluster individuals into different cryptic subpopulations on the basis of multilocus genotype data. Using the software, we performed independent analyses under three assumed numbers of population strata \((k = 2, 3, \text{and } 4)\) using 200 unlinked, randomly selected markers. To confirm the results achieved through Structure 2.2, we further tested population stratification in our sample using EIGENSOFT 2.0 software, which uses both principal-component analysis and a genomic control approach to correct possible statistical bias caused by ancestral differences in whole-genome association studies. Routinely, GWA statistical analyses on individual SNPs were first conducted by the Wald test implemented in Plink (Version 1.03, http://pngu.mgh.harvard.edu/~purcell/plink) to obtain a statistic for each tested SNP. The parameters, including age, age\(^2\), sex, age/age\(^2\) by-sex interaction, height, and weight, were tested for their associations with UD BMD. The significant \((p < .05)\) terms (age, age\(^2\), sex, and weight) then were included as covariates to adjust the raw UD BMD values for subsequent analyses. The adjusted UD BMD data were normally distributed \((p > .1)\).

Subsequently, pathway-based GWA analysis was performed following the approach developed by Wang and colleagues. Briefly, the main procedures are as follows:

1. Generation of statistic value of gene-phenotype association: Suppose \(N\) genes in the whole genome and each gene with \(x_i\) \((i \leq N)\) SNPs mapped to it. For a specific gene \(G_i\), the highest SNP-phenotype association statistic value among all \(x_i\) SNPs was selected to represent the statistic value of the gene, denoted as \(r_i\). Mostly, SNPs were mapped to their closest gene. And in rare cases, if an SNP was located within shared regions of two overlapping genes, the SNP was mapped to both genes.

2. Ranking statistic values of gene-phenotype association: We ranked all genes by sorting their statistic values from the highest to lowest, denoted as gene list \(L\) \((r_1, r_2, \ldots, r_N)\).

3. Enrichment score calculation: For each given pathway/gene set \(S\), composed of \(N_i\) genes, a statistic enrichment score (ES)
was calculated. ES is a weighted Kolmogorov-Smirnov-like running-sum statistic that reflects the overrepresentation of genes within S at the top of the entire ranking list of genes in the genome. The score is calculated by walking down the gene list L, increasing a running-sum statistic when we encounter a gene in S and decreasing it when we encounter genes not in S. ES is the maximum deviation from zero encountered in the random walk.

$$ES^S = \max_{1 \leq S \leq N_S} \left\{ \sum_{G_i \in S} \left| \frac{r_i} {N_R} \right|^p - \sum_{G_i \notin S} \frac{1} {N - N_S} \right\}$$

where \(N_R = \sum_{G_i \in S} |r_i|^p\), and \(p\) (designated as 1 here) is a parameter that gives higher weight to genes with extreme statistic values. Thus, both the rank of each gene in the entire ranked gene list \(L\) \((N_s)\) and the association statistic value of it \(r\) will contribute to the ES value of a pathway.

4. Permutation and nominal significance assessment: To estimate the significance level of ES, permutation procedures were used to create the null distribution of ES for each pathway/gene set. The phenotype data of sampled subjects were first shuffled. Then the previous three steps were repeated to calculate ES for the pathway/gene set during each permutation. In total, 1000 permutations were done. Finally, a distribution for ES \(ES^S_{null}\) was generated for each pathway/gene set S. The significance of an observed ES \(ES^S\) for a pathway/gene set (nominal \(p\) value) was estimated as the percentage of permutations whose \(ES^S_{null}\) values were greater than the observed \(ES^S\).

5. Multiple-testing adjustments: To compare the significance among pathways with different numbers of genes, first, a normalized ES (NES) was constructed based on the observed ES, mean and standard deviation (SD) of \(ES^S_{null}\).

$$NES^S = \frac{ES^S - \text{mean}(ES^S_{null})}{\text{SD}(ES^S_{null})}$$

Then two measures were used to adjust for multiple-hypothesis testing for more reliable results. False-discovery rate (FDR) is a procedure usually used to control the fraction of expected false-positive findings to stay below a certain threshold. Family-wise error rate (FWER) is another approach to correct for multiple-hypotheses testing that is considered to be a highly conservative procedure seeking to ensure that the list of reported results does not include even a single false-positive gene set.\(^{14,21}\) For a pathway/gene set, let \(NES^S_{observe}\) denote the normalized enrichment score in the observed data. The FDR \(q\) value (denoted as \(q_{fdr}\)) was calculated as the ratio of the fraction of all permutations with \(NES \geq NES^S_{observe}\) to the fraction of observed pathways/gene sets with \(NES \geq NES^S_{observe}\). FWER \(p\) value (denoted as \(p_{fwer}\)) was calculated as the fraction of pathways/gene sets whose greatest NES among all permutations is greater than \(NES^S_{observe}\). In this study, the significance criteria after multiple-testing correction for a pathway is that both FDR \(q\) value and FWER \(p\) value are .05 or less.

$$q_{fdr} = \frac{\% \text{ of permutations with } NES^S_{null} \geq NES^S_{observe}}{\% \text{ of observed } S \text{ with } NES^S_{null} \geq NES^S_{observe}}$$

$$p_{fwer} = \frac{\% \text{ of } S \text{ with } \max(NES^S_{null}) \geq NES^S_{observe}}{\text{total number of } S}$$

Replication analyses with FHS sample: The statistical analyses for the FHS sample were almost the same as those in the discovery GWAS cohort unless some slight modifications regarding that FHS is a family-based study. In brief, first, the raw arm BMD data were adjusted. The parameters, including age, age\(^2\), sex, age/age\(^2\)-by-sex interaction, height, and weight, were tested for their associations with arm BMD. The significant \((p \leq .05)\) terms (age-by-sex, height, and weight) then were included as covariates to adjust the raw arm BMD values for subsequent analyses. The adjusted UD BMD data were normally distributed \((p > .1)\). Second, the routine GWA statistical analyses on individual SNPs were conducted using FBAT software (http://biosun1.harvard.edu/~fbat/hts.htm)\(^{22}\) to obtain a Z statistic for each tested SNP. Third, a pathway-based GWA analysis was performed with almost the same method as that in the discovery cohort unless that the Z statistical values for all genes other than the phenotype data were shuffled when doing the permutation and nominal significance assessment so that the family structure of the population could be maintained.\(^{14}\) A total of 10,000 permutations were done.

Results

Characteristics of study subjects

In the discovery GWAS cohort, a total of 984 individuals were included in the analyses. The basic characteristics of all subjects in this cohort, stratified by sex, are shown in Table 1. The plot of the UD BMD T-scores (mean = -0.53, SD = 2.01) in this population are given in Supplemental Fig. S1. There are 2187 subjects (817 males and 1370 females) from 477 families in the FHS who were studied to replicate the findings from the discovery cohort. The basic characteristics of all subjects in the FHS sample, stratified by sex, are shown in Table 2.

Pathway-based association analyses in the discovery cohort

No significant population stratification was detected in the discovery cohort (984 subjects) by either Structure 2.2\(^{17}\) or EIGENSOFT 2.0 software.\(^{18}\) When using Structure 2.2 software, all subjects were tightly clustered together, suggesting no population stratification (data not shown). The results from EIGENSOFT 2.0 software\(^{18}\) provided further support for the results from Structure 2.2. Only one principal component was significant in the principal-component analysis \((p < .001)\), suggesting that only one population ancestry existed for our sample. Based on genome-wide SNP information, inflation factor \(\lambda\), a measure of population stratification, also was estimated by the genomic control\(^{19}\) approach implemented in EIGENSOFT.

Table 1. Basic Characteristics of the Discovery Cohorta

| Trait       | Male \((n = 489)\) | Female \((n = 495)\) |
|-------------|-------------------|---------------------|
| Age (years) | 50.60 (18.87)     | 50.39 (17.65)       |
| Height (cm) | 177.89 (6.96)     | 163.83 (6.52)       |
| Weight (kg) | 89.00 (14.96)     | 71.40 (15.97)       |
| UD BMD (g/cm\(^2\)) | 0.507 (0.073) | 0.418 (0.065) |

\(a\)The data are mean (SD) of raw values.
Table 2. Basic Characteristics of FHS Subjects

| Trait            | Male (n = 817) | Female (n = 1370) |
|------------------|---------------|-------------------|
| Age (years)      | 63.84 (11.49) | 63.95 (11.74)     |
| Height (inch)    | 68.31 (2.80)  | 62.94 (2.71)      |
| Weight (pounds)  | 176.85 (22.44)| 146.93 (25.78)    |
| Arm BMD (g/cm²)  | 0.983 (0.097) | 0.784 (0.098)     |

*The data are mean (SD) of raw values.

2.0 software. Ideally, for a homogeneous population with no stratification, the value of the inflation factor parameter λ should be equal to or near 1. In our study, \( \lambda = 1.01 \), indicating that the sample population is homogeneous.

Among all 963 pathways tested in our analysis, the ROA pathway, which contains 25 genes, showed the strongest significant association with wrist UD BMD. The ROA pathway achieved a high ES of 0.557 with a p value of .005 (Figs. 1 and 2).

Most of the genes in the ROA pathway ranked at the top in significance in the gene list containing 14,585 genes genome-wide, with 13 genes ranked among the top 2000 genes (Fig. 2). More important, it is the only pathway, among all 963 pathways tested, that achieved statistically significant FDR and FWER values (NES = 2.50, \( q_{FDR} = 0.043, \rho_{FWER} = 0.016 \)).

The ROA pathway is annotated by the KEGG database. It is the only pathway related to autophagy among all the pathways significant in the gene list containing 14,585 genes genome-wide, with 13 genes ranked among the top 2000 genes (Fig. 2). Among all 963 pathways tested, that achieved statistically significant FDR and FWER values.

Among all 963 pathways tested in our analysis, the ROA pathway, which contains 25 genes, showed the strongest significant association with wrist UD BMD. The ROA pathway achieved a high ES of 0.557 with a p value of .005 (Figs. 1 and 2).

The ROA pathway is annotated by the KEGG database. It is the only pathway related to autophagy among all the pathways tested. As shown in Table 3, the ROA pathway contains 6 genes encoding type I and type II interferons (IFNs), which are pleiotropic cytokines that contribute to regulation of autophagy.

The pathway also contains five genes encoding subunits of kinases, which are autophagy modulators. These kinases include the catalytic subunit and regulatory subunit of AMP-activated protein kinase (AMPK), and unc-51-like kinase (ULK2). The individual gene with the most significant p value among all the genes in this pathway is IFNA14 (\( p = 1.10 \times 10^{-3} \)). Other genes in the ROA pathway that also contribute positively to the ES (ie, the genes that ranked before or at the point of ES, also denoted as leading-edge genes) are ATG12 (\( p = 2.11 \times 10^{-3} \)), ATG21 (\( p = 2.90 \times 10^{-3} \)), IFNA5 (\( p = 3.16 \times 10^{-3} \)), IFNA17 (\( p = 3.49 \times 10^{-3} \)), IFNA8 (\( p = 4.92 \times 10^{-3} \)), IFNA6 (\( p = 5.46 \times 10^{-3} \)), IFNA4 (\( p = 6.02 \times 10^{-3} \)), IFNA7 (\( p = 6.55 \times 10^{-3} \)), ATG7 (\( p = 8.32 \times 10^{-3} \)), IFNA13 (\( p = 1.21 \times 10^{-3} \)), PIK3C3 (\( p = 1.57 \times 10^{-3} \)), and ATG3 (\( p = 1.62 \times 10^{-2} \)) (Table 3).

Information for the SNPs whose \( p \) values were selected to represent the \( p \) values of the 13 leading-edge genes is given in Table 4. All the SNPs are common variants with minor allele frequencies (MAF) of 8% or greater in our sample. We used the FASTSNP program (http://fastsnp.ibms.sinica.edu.tw) to analyze potential functions of these SNPs. The SNP rs7037147 is located in the intron of IFNA14 and is suggested as an intronic enhancer. An A-to-G change at this locus will delete a binding site of the transcription factor SRY. The SNP rs10757219 is located at a transcription factor binding site at the upstream regulatory
region of IFNA8. An A-to-G change at this locus will delete a binding site of the transcription factor c-Myc.

Replication study in the FHS sample

To further confirm our GWAS findings in the discovery cohort, we carried out a replication analysis with a sample from the FHS population, which contains 2187 subjects. Same as that in the discovery cohort, 963 pathways were analyzed. In total, this analysis covered 15,598 genes genome-wide. All 25 genes of the ROA pathway, as shown in Table 3, were included for association testing. For the ROA pathway, the observed ES is 0.24, observed NES = 13.77; based on 10,000 permutation, nominal p value < 10^-4, q_{fdr} = 0.001, and p_{fwer} = 0.007. That is, the ROA pathway showed significant association with arm BMD in this FHS sample even after multiple-testing adjustment. The leading-edge genes of the ROA pathway in the replication FHS cohort are PIK3C3 (p = 1.97 \times 10^{-3}), ATG12 (p = 8.16 \times 10^{-3}), PRKAA2 (p = 8.28 \times 10^{-3})

### Table 3. Genes in Regulation-of-Autophagy (ROA) Pathway

| Gene symbol | Gene ID | Genome location | Full name |
|-------------|---------|----------------|-----------|
| ATG12       | 9140    | 5q21-q22       | ATG12 autophagy-related 12 homologue (S. cerevisiae) |
| ATG3        | 64422   | 3q13.2         | ATG3 autophagy-related 3 homologue (S. cerevisiae) |
| ATG5        | 9474    | 6q21           | ATG5 autophagy-related 5 homologue (S. cerevisiae) |
| ATG7        | 10533   | 3p25.3-p25.2   | ATG7 autophagy-related 7 homologue (S. cerevisiae) |
| BECN1       | 8678    | 17q21          | Beclin 1, autophagy-related. (also known as ATG6) |
| GABARAPL1   | 23710   | 12p13.2        | GABA(A) receptor–associated protein like 1 (ATG8) |
| IFNA1       | 3439    | 9p22           | Interferon, α1 |
| IFNA10      | 3446    | 9p22           | Interferon, α10 |
| IFNA13      | 3447    | 9p22           | Interferon, α13 |
| IFNA14      | 3448    | 9p22           | Interferon, α14 |
| IFNA16      | 3449    | 9p22           | Interferon, α16 |
| IFNA17      | 3451    | 9p22           | Interferon, α17 |
| IFNA2       | 3440    | 9p22           | Interferon, α2 |
| IFNA21      | 3452    | 9p22           | Interferon, α21 |
| IFNA4       | 3441    | 9p22           | Interferon, α4 |
| IFNA5       | 3442    | 9p22           | Interferon, α5 |
| IFNA6       | 3443    | 9p22           | Interferon, α6 |
| IFNA7       | 3444    | 9p22           | Interferon, α7 |
| IFNA8       | 3445    | 9p22           | Interferon, α8 |
| IFNG        | 3458    | 12q14          | Interferon, γ |
| PIK3C3      | 5289    | 18q12.3        | Phosphoinositide-3-kinase, class3 (VPS34) |
| PIK3R4      | 30849   | 3q22.1         | Phosphoinositide-3-kinase, regulatory subunit 4 (VPS15) |
| PRKAA1      | 5562    | 5p12           | AMP-activated protein kinase (AMPK) α1 catalytic subunit |
| PRKAA2      | 5563    | 1p31           | AMP-activated protein kinase (AMPK) α2 catalytic subunit |
| ULK2        | 9706    | 17p11.2        | Unc-51-like kinase |

### Table 4. The Most Significant SNPs Mapped to the Leading-Edge Genes of the ROA Pathway

| SNP          | Gene      | p Value   | Role        | Allele | MAF<sup>a</sup> |
|--------------|-----------|-----------|-------------|--------|-----------------|
| rs7037147    | IFNA14    | 1.10 \times 10^{-3} | Intron | A/G   | 0.16            |
| rs6883910    | ATG12     | 2.11 \times 10^{-3} | Upstream | A/C   | 0.39            |
| rs7852323    | IFNA21    | 2.90 \times 10^{-3} | Upstream | A/G   | 0.17            |
| rs1360286    | IFNA5     | 3.16 \times 10^{-3} | Downstream | A/C   | 0.21            |
| rs10964938   | IFNA17    | 3.49 \times 10^{-3} | Upstream | A/C   | 0.18            |
| rs10757219   | IFNA8     | 4.92 \times 10^{-3} | Promoter/regulatory region | A/G   | 0.28            |
| rs2015345    | IFNA16    | 5.46 \times 10^{-3} | Downstream | A/C   | 0.17            |
| rs147749     | IFNA4     | 6.02 \times 10^{-3} | Downstream | A/T   | 0.17            |
| rs7856345    | IFNA7     | 6.55 \times 10^{-3} | Downstream | A/T   | 0.18            |
| rs17536147   | ATG7      | 8.32 \times 10^{-3} | Intron  | C/G   | 0.08            |
| rs637949     | IFNA13    | 1.21 \times 10^{-2} | Upstream | C/G   | 0.17            |
| rs9955346    | PIK3C3    | 1.57 \times 10^{-2} | Upstream | A/G   | 0.17            |
| rs694433     | ATG5      | 1.62 \times 10^{-2} | Upstream | A/G   | 0.21            |

<sup>a</sup>MAF (minor allele frequency) in our sample.
BECN1 (p = 4.45 × 10^{-2}), IFNA13 (p = 4.72 × 10^{-2}), and ATG7 (p = 4.92 × 10^{-2}), five of which (ie, PIK3C3, ATG12, ATGS, IFNA13, and ATG7) are also leading-edge genes in the discovery cohort. Therefore, there is significant overlapping in the genes involved between the discovery and replication cohorts that contribute to the association signals at the ROA pathway level. The information of the most significant SNPs mapped to these leading-edge genes in the replication FHS cohort are presented in Supplemental Table S1.

Discussion

In this study, a genome-wide association analysis was conducted to identify the pathway(s) involved in BMD determination at a clinically important skeletal site—wrist UD. The ROA pathway was shown to be significantly associated with wrist UD BMD (p = 0.05) and was the only pathway, among a total of 963 pathways tested, that achieved statistically significant FDR and FWER values. To further confirm our findings in the discovery cohort, 2187 subjects from the FHS with arm BMD data available were studied as a replication study. The ROA pathway showed significant association with arm BMD in the FHS sample even after the multiple-testing adjustment, which further confirmed our findings in the discovery cohort. This is the first time that the autophagy-related biologic process has been implicated as an underlying factor for wrist BMD variation and hence a risk factor for wrist osteoporosis.

Autophagy is a conserved housekeeping function of eukaryotic cells that permits sequestration of cytoplasmic components in membrane-bound vesicles and delivers them to lysosomes for degradation. Autophagy is subject to suppression or further induction in response to different stresses, starvation, specific hormonal regulation, and other stimuli. The core process of autophagy and the functional interactions among the genes in the ROA pathway are depicted in Fig. 3. The autophagy pathway has numerous proposed physiologic functions. Growing evidence suggests that malfunctions of autophagy contribute to many human diseases, for example, myopathies, liver disease, neurodegeneration, heart disease, and cancer.

Although this article is the first to support the ROA pathway's importance to osteoporosis, the pathway's relevance to bone metabolism has been suggested in previous studies. Autophagy has been recognized as an intermediate stage during the terminal differentiation of chondrocytes. Specifically, during endochondral ossification, induction of autophagy occurs prior to apoptotic hypertrophic maturation in terminally differentiated chondrocytes. These chondrocytes morphologically exhibit autophagic characteristics and express autophagic proteins. The induction of autophagy enhances the survival of chondrocytes in hypoxic microenvironments (eg, low protein, glucose, and O2 concentrations) and facilitates synthesis of the calcified extracellular matrix.

Some genes in the ROA pathway also have been considered as important modulating factors for bone development or remodeling. For example, IFN-α can suppress the proliferation of osteoprogenitor cells and modify the expression of a number of important cytokines that are regulators for human osteoprogenitor cell growth and differentiation. IFN-α also can inhibit the differentiation of peripheral blood mononuclear cells (PBMCs) to osteoclasts. Interferon-γ (IFN-γ) is considered to be an antagonist of RANKL and is a well-known potent inhibitor of osteoclast formation and function. IFN-γ also exhibits antiproliferative actions on primary osteoblast cells. Finally, it has been reported that in cultured osteoblasts, AMPK influences the expression of cyclooxygenase 2 (COX-2), which potentially could impact fracture healing because COX-2 is a crucial mediator in mechanically induced bone formation.

In addition to highlighting the potential contribution of the ROA pathway to human wrist BMD deviation, this study also demonstrates the important advantage of pathway-based GWA analysis. None of the genes in the ROA pathway reached a significant level when considered separately; thus the potential impact of the ROA pathway on wrist BMD would not have been detected by individual gene/SNP GWA analyses. By considering critical information about the interaction of a set of functionally related genes and their joint effects, pathway analysis may be a useful paradigm for revealing the polygenic nature of complex diseases. Recently, some candidate pathway association studies have been done, suggesting that the joint actions of common gene variants within pathways may play a major role in predisposing to complex diseases. Examples include the axon guidance pathways for Parkinson disease (PD) and amyotrophic lateral sclerosis (ALS). Differing from the candidate pathway strategies, in the current genome-wide analysis we systematically studied a
large number of pathways annotated by public pathway databases with the aim of identifying pathways associated with UD BMD. This hypothesis-free strategy could make full use of GWA data and would be more useful for discovering new disease-related processes and generating novel hypotheses about pathogenic mechanisms of disease.

It would be interesting to clarify the detailed mechanism of the ROA pathway on wrist BMD. Considering the existing evidence linking some ROA pathway genes with osteoblastogenesis or osteoclastogenesis, it can be imagined that the involvement of autophagy in bone may not be limited to chondrocytes alone. Since autophagy is another mechanism for programmed cell death and shows close interaction with apoptosis, whether autophagy is involved in the survival control of osteoblasts/osteoclasts/osteocytes is an important hypothesis to test. Also, it will be interesting to study whether some lifestyle factors will interact with the ROA pathway or if the ROA pathway would be a candidate target for osteoporosis therapy.

In this study we found that the ROA pathway is associated with human wrist BMD variation. However, we cannot exclude other pathways for their significance in wrist BMD variation, even though these pathways did not pass the significance criteria adopted in this study. Regarding the statistical aspect of this study, it should be noted that some genes inevitably are shared by different pathways. Although the overlap of genes among different pathways will not affect the pathways’ relative ranking in terms of NRES values, interdependence between pathways will lead to a decreased power by affecting FDR and FWER when the causal genes are shared by multiple pathways (owing to the permutation procedure for multiple-testing adjustment). A larger sample size or more specifically annotated pathways will be helpful to improve the power of this pathway-based GWA.

It also should be noted that under the significance criteria adopted in this study, the ROA pathway was not detected to be associated with hip or spine BMD in the discovery cohort (data not shown). The possible reasons for the inconsistent association signals detected at these three skeletal sites may include statistical errors (type I error at the wrist or type II error at the hip/spine) and genetic heterogeneity between different skeletal sites. In particular, type II error at the hip/spine may be the major reason. The type II error may be caused by the relatively low statistical power of this newly developed pathway-based method (as discussed earlier). This low statistical power, when complicated by the possible genetic heterogeneity (ie, smaller effects of the ROA pathway on hip and spine than on wrist), may lead to negative association signals at the hip and spine.

In summary, we applied a novel pathway-based GWA analysis method to systematically screen functional pathways/gene sets associated with wrist UD BMD and thus osteoporosis. The significant enrichment of ROA pathway genes among the top-ranking genes associated with wrist UD BMD, together with the pathway’s functional relevance to bone metabolism, strongly supports an important role of autophagy in human wrist BMD variation. Further detailed and specific functional studies of the ROA pathway will provide new insights into the pathway’s relevance to the physiology of bone and the etiology of wrist osteoporosis.

Acknowledgments

Investigators of this work were supported in part by grants from the NIH (R01 AR050496, R21 AG027110, R01 AG026564, P50 AR055081, and R21 AA015973). The study also benefited from grants from the National Science Foundation of China, the Huo Ying Dong Education Foundation, HuNan Province, Xi’an Jiaotong University, and the Ministry of Education of China. We also thank all contributors to the Framingham Heart Study. The Framingham Heart Study and the Framingham SHARE project are conducted and supported by the National Heart, Lung, and Blood Institute (NHLBI) in collaboration with Boston University. The Framingham SHARE data used for the analyses described in this article were obtained through dbGaP. This article was not prepared in collaboration with investigators of the Framingham Heart Study and does not necessarily reflect the opinions or views of the Framingham Heart Study, Boston University, or the NHLBI.

Disclosures

Dr. Recker has been paid consultant for Merck, Lilly, Wyeth, Procter&Gamble, Amgen, Roche, Glaxo Smith Kline, Novartis, and NPS Allelix; and has received grant/research support from Merck, Lilly, Wyeth, Procter&Gamble, Amgen, Roche, Glaxo Smith Kline, Novartis, NPS Allelix, and Sanofi-Aventis. All other authors state that they have no conflicts of interest.

References

1. Johnell O, Kanis J. Epidemiology of osteoporotic fractures. Osteoporos Int. 2005;16:53–7.
2. Chao EY, Inoue N, Koo TK, Kim YH. Biomechanical considerations of fracture treatment and bone quality maintenance in elderly patients and patients with osteoporosis. Clin Orthop Relat Res. 2004;425:12–25.
3. Muller ME, Webber CE, Bouxsein ML. Predicting the failure load of the distal radius. Osteoporos Int. 2003;14:345–352.
4. Weichetova M, Stepans JJ, Haas T, Michalska D. The risk of Colles’ fracture is associated with the collagen Iα1 (Isp1) polymorphism and ultrasound transmission velocity in the calcaneus only in heavier postmenopausal women. Calcif Tissue Int. 2005;76:98–106.
5. Tse KY, Macias BR, Meyer RS, Hargens AR. Heritability of bone density: regional and gender differences in monozygotic twins. J Orthop Res. 2009;27:150–154.
6. Duncan EL, Cardon LR, Sinshheimer JS, Wass JA, Brown MA. Site and gender specificity of inheritance of bone mineral density. J Bone Miner Res. 2003;18:1531–1538.
7. Cuddihy MT, Gabriel SE, Crowson CS, O’Fallon WM, Melton LJ 3rd. Forearm fractures as predictors of subsequent osteoporotic fractures. Osteoporos Int. 1999;9:469–475.
8. Gardsell P, Johnell O, Nilsson BE, Gullberg B. Predicting various fragility fractures in women by forearm bone densitometry: a follow-up study. Calcif Tissue Int. 1999;52:348–353.
9. Mallmin H, Ljunghall S, Persson I, Naessens T, Krusemo UB, Bergstrom R. Fracture of the distal forearm as a forecaster of subsequent hip fracture: a population-based cohort study with 24 years of follow-up. Calcif Tissue Int. 1993;52:269–272.
10. Styrkarsdottir U, Hallldorsson BV, Gretarsdottir S, et al. Multiple genetic loci for bone mineral density and fractures. N Engl J Med. 2008;358:2355–2365.
30. Roach HI, Rivadeneira F, Inouye M, et al. Bone mineral density, osteoporosis, and osteoporotic fractures: a genome-wide association study. Lancet. 2008;371:1505–1512.

29. Liu YZ, Wilson SG, Wang L, et al. Identification of PLCL1 gene for hip bone size variation in females in a genome-wide association study. PLoS ONE. 2008;3:e3160.

28. Kiel DP, Demissie S, Dupuis J, Lunetta KL, Murabito JM, Karasik D. Genome-wide association with bone mass and geometry in the Framingham Heart Study. BMC Med Genet. 2007;8 (Suppl 1): S14.

27. Wang K, Li M, Bucan M. Pathway-Based Approaches for Analysis of Genomewide Association Studies. Am J Hum Genet. 2007;81:1278–1283.

26. Di X, Matsuzaki H, Webster TA, et al. Dynamic model based algorithms for screening and genotyping over 100K SNPs on oligonucleotide microarrays. Bioinformatics. 2005;21:1958–1963.

25. Rabbee N, Speed TP. A genotype calling algorithm for affymetrix SNP arrays. Bioinformatics. 2006;22:7–12.

24. Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. Genetics. 2000;155:945–959.

23. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis for stratification in genomewide association studies. Nat Genet. 2006;38:904–909.

22. Devlin B, Roeder K. Genomic control for association studies. Biometrics. 1999;55:997–1004.

21. Purcell S, Neale B, Todd-Brown K, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007;81:559–575.

20. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A. 2005;102:15545–15550.

19. Horvath S, Xu X, Laird NM. The family based association test method: strategies for studying general genotype–phenotype associations. Eur J Hum Genet. 2001;9:301–306.

18. Mizushima N, Levine B, Cuervo AM, Klionsky DJ. Autophagy fights disease through cellular self-digestion. Nature. 2000;408:87–90.

17. Espert L, Codogno P, Biard-Piechaczek M. Involvement of autophagy in viral infections: antiviral function and subversion by viruses. J Mol Med. 2007;85:811–823.

16. Tallozy C, Zang W, Virgin HWt, et al. Regulation of starvation- and virus-induced autophagy by the elf2alpha kinase signaling pathway. Proc Natl Acad Sci U S A. 2002;99:190–195.

15. Pyo JO, Jang MH, Kwon YK, et al. Essential roles of Atg5 and FADD in autophagic cell death: dissection of apoptotic cell death into vacuole formation and cell death. J Biol Chem. 2005;280:20722–20729.

14. Shapira IM, Adams CS, Freeman T, Srinivas V. Fate of the hypertrophic chondrocyte: microenvironmental perspectives on apoptosis and survival in the epiphyseal growth plate. Birth Defects Res C Embryo Today. 2005;75:330–339.

13. Bohensky J, Shapiro IM, Leshinsky S, Watanabe H, Srinivas V. PIM-2 is an independent regulator of chondrocyte survival and autophagy in the epiphyseal growth plate. J Cell Physiol. 2007;213:246–251.

12. Zuschik MJ, Hilton MJ, Zhang X, Chen D, O’Keefe RJ. Regulation of chondrogenesis and chondrocyte differentiation by stress. J Clin Invest. 2008;118:429–438.

11. Roach HI, Aigner T, Kouri JB. Chondroptosis: a variant of apoptotic cell death in chondrocytes? Apoptosis. 2004;9:265–277.