

Research article

Current limitations of SNP data from the public domain for studies of complex disorders: a test for ten candidate genes for obesity and osteoporosis

Volodymyr Dvornyk†1, Ji-Rong Long†1, Dong-Hai Xiong†1, Peng-Yuan Liu1, Lan-Juan Zhao1, Hui Shen1, Yuan-Yuan Zhang1, Yong-Jun Liu1, Sonia Rocha-Sanchez1, Peng Xiao1, Robert R Recker1 and Hong-Wen Deng*1,2

Address: 1Osteoporosis Research Center and Department of Biomedical Sciences, Creighton University, 601 N. 30th St., Suite 6730, Omaha, NE 68131, USA and 2Laboratory of Molecular and Statistical Genetics, College of Life Sciences, Hunan Normal University, Changsha, Hunan 410081, P. R. China

Email: Volodymyr Dvornyk - dvornyk@creighton.edu; Ji-Rong Long - jrlongyou@creighton.edu; Dong-Hai Xiong - donghai@creighton.edu; Peng-Yuan Liu - pyliu@creighton.edu; Lan-Juan Zhao - zhaolanjuan@creighton.edu; Hui Shen - huishen@creighton.edu; Yuan-Yuan Zhang - zhangyuan2@creighton.edu; Yong-Jun Liu - jun@creighton.edu; Sonia Rocha-Sanchez - saraujo@creighton.edu; Peng Xiao - peng@creighton.edu; Robert R Recker - rrecker@creighton.edu; Hong-Wen Deng* - deng@creighton.edu

* Corresponding author †Equal contributors

Published: 25 February 2004

BMC Genetics 2004, 5:4

Received: 24 November 2003
Accepted: 25 February 2004

This article is available from: http://www.biomedcentral.com/1471-2156/5/4

© 2004 Dvornyk et al; licensee BioMed Central Ltd. This is an Open Access article: verbatim copying and redistribution of this article are permitted in all media for any purpose, provided this notice is preserved along with the article's original URL.

Abstract

Background: Public SNP databases are frequently used to choose SNPs for candidate genes in the association and linkage studies of complex disorders. However, their utility for such studies of diseases with ethnic-dependent background has never been evaluated.

Results: To estimate the accuracy and completeness of SNP public databases, we analyzed the allele frequencies of 41 SNPs in 10 candidate genes for obesity and/or osteoporosis in a large American-Caucasian sample (1,873 individuals from 405 nuclear families) by PCR-invader assay. We compared our results with those from the databases and other published studies. Of the 41 SNPs, 8 were monomorphic in our sample. Twelve were reported for the first time for Caucasians and the other 29 SNPs in our sample essentially confirmed the respective allele frequencies for Caucasians in the databases and previous studies. The comparison of our data with other ethnic groups showed significant differentiation between the three major world ethnic groups at some SNPs (Caucasians and Africans differed at 3 of the 18 shared SNPs, and Caucasians and Asians differed at 13 of the 22 shared SNPs). This genetic differentiation may have an important implication for studying the well-known ethnic differences in the prevalence of obesity and osteoporosis, and complex disorders in general.

Conclusion: A comparative analysis of the SNP data of the candidate genes obtained in the present study, as well as those retrieved from the public domain, suggests that the databases may currently have serious limitations for studying complex disorders with an ethnic-dependent background due to the incomplete and uneven representation of the candidate SNPs in the databases for the major ethnic groups. This conclusion attests to the imperative necessity of large-scale and accurate characterization of these SNPs in different ethnic groups.
Background

A single nucleotide polymorphism (SNP) is generally defined as a stable substitution of a single base with a frequency of more than 0.01 in at least one population [1]. In human genetic studies, SNPs are simply referred to as bi-allelic markers since the other types (tri-allelic and tetra-allelic SNPs) are very rare in the human genome [2]. SNPs have been recognized as an important tool in human genetics and medicine [3,4]. Because SNPs are abundant and scattered throughout the whole human genome, they have been widely used in the genetic association studies of various complex diseases such as obesity, osteoporosis, asthma, hypertension (see, e.g., [5-7]). Various analyses of SNPs across the human genome have been also conducted to determine haplotype patterns in human populations [8-12]. These data are very useful to study the genetic basis of common complex diseases [13].

At present there are several SNP public databases. The largest are dbSNP and HGBase containing together several million SNPs [14]. There also exist other relatively small or specific SNP databases, e.g., TSC [15], JSNP [16], HOWDY [17], GeneSNPs [18]. While the databases have been continuously expanded, the quality and completeness of the deposited SNP data remain to be of particular importance and to be assessed. Recent evaluation of quality and comprehensiveness of about four million candidate SNPs from some public and Celera databases showed that about 6-12% SNPs could not be validated [19]. They represent rare variants, population-specific SNPs and sequencing errors. The other studies reported that only about 50% SNPs are common in any given population. Therefore, measurement of allele frequency and linkage disequilibrium for SNPs in databases should be done to efficiently select a minimal subset of SNPs for population association studies of complex diseases [20,21].

Obesity and osteoporosis are common complex disorders with continuously growing burden and cost for their prevention and treatment. For example, the most recent data for the United States, derived from the third National Health and Nutrition Examination Survey (1988–94), showed ~20% of US men and ~25% of US women are obese and these proportions are increasing [22]. The same tendency has been observed in other human populations [23]. Osteoporosis is another major health problem, particularly in the elderly. More than 40% of postmenopausal women, on average, will suffer at least one osteoporotic fracture [24]. This disorder incurred an estimated direct cost of ~14 billion dollars in the USA alone in 1995 [25].

SNPs of candidate genes for obesity and osteoporosis may be of particular importance in genetic studies of these disorders. The allele frequencies are important in the selection of SNPs for studying complex diseases [26]. For example, association studies of osteoporosis usually generate inconsistent results in different ethnic groups [27]. One important reason is that polymorphisms associated significantly with certain osteoporotic phenotypes in a given ethnic group may be absent or rare in another ethnic group [28-31]. Moreover, racial differences in the prevalence of certain allele could account for certain proportion of disease trait variation between different ethnicities [32]. Finally, a comparison of SNP allele frequencies among different ethnicities can provide valuable information for mapping by admixture linkage disequilibrium (MALD) [33].

Despite the importance of the above data, no attempts have been made to check the quality and completeness of SNPs of candidate genes for obesity and osteoporosis in the public databases. In fact, most allele frequencies in the databases were obtained from studies of relatively small samples (dozens or so), which could yield large sampling errors. In addition to the databases, there are abundant SNP data from many association/linkage studies of these two common complex diseases. These individually published data have not yet been assessed in reference to each other, or to those in the databases. In most association studies, the regular sample sizes used to detect allele frequencies have only been of the order of a hundred or so. However, to obtain reliable estimates of allele frequency distributions, relatively larger sample sizes are needed. For example, for an SNP, to let allele frequency estimate’s error = 0.05, the required sample sizes for frequency distributions (0.5, 0.5), (0.7, 0.3), and (0.9, 0.1) should be at least 1150, 1000, and 400, respectively [34]. Otherwise, the estimates of allele frequencies may be biased to a large extent.

In the present study, we determined allele frequencies of 41 SNPs in 10 candidate genes for obesity and/or osteoporosis in a large American-Caucasian sample (1,873 subjects from 405 nuclear families). We then compared them with the corresponding data for the major ethnic groups obtained from the public databases and various individual studies to check their accuracy and completeness, and to compare the SNP allele frequencies in different ethnic groups.

Results

SNP polymorphism in the studied Caucasian sample

Of the 41 candidate SNPs for obesity and/or osteoporosis identified in the studied Caucasian sample (Table 1), 29 (70.7%) had minor allele frequency ≥0.1. Eight SNPs (19.5%) were monomorphic. Allele frequencies of twelve SNPs in Caucasians were reported for the first time (Table 2). Five of these SNPs were monomorphic and the other seven were polymorphic with minor allele frequencies ranged 0.019–0.398.
Table 1: Summary information about the studied SNPs

| SNP     | Genea | Nameb | dbSNP Accession | Mutationd | Domain       | Minor allele frequency | Forward primer | Reverse primer             |
|---------|-------|-------|-----------------|-----------|--------------|------------------------|----------------|--------------------------|
| 1       | APOE  | ss12568587 | G-C | 5’ UTR | 0.357 | TCCCCAGGACCCGCC  | CCCCCAGCGGAGTCAC |
| 2       | APOE  | ss12568609 | G-A | Intron 2 | 0.399 | CTCAGGCTGACCTGGTCC | ACTCTGGGCTAGTGCTTC |
| 3       | APOE  | ss12568607 | T-C | Exon 4 | 0.149 | GGCGGAGGCTGCAAC | CGAGCATGGGCTGCACTTC |
| 4       | APOE  | ss12568612 | C-T | Exon 4 | 0.087 | GCTCGGTAAGGCGCCTCC | GCCGCCCTGAGCTTC |
| 5       | COL1α1 | ss12568606 | G-T | 5’ UTR | 0.154 | GCACCTGGCCTAGAC | CCTATGCGGAGCTGACA |
| 6       | COL1α1 | ss12568597 | G-T | Intron 1 | 0.188 | CAATACGCGGCTTTCACC | CATTCCGAGGGAGGTCCTT |
| 7       | COL1α1 | ss12568598 | G-T | Exon 8 | 0.000 | GAAGAGCTGGAGTAGGAGCA | GGTCGCGAGCTTCAC |
| 8       | COL1α1 | ss12568584 | G-A | Exon 45 | 0.203 | TGCAGTCTTGGCAGCAAC | AGCGAGGATTCAGTCATTCAGAG |
| 9       | ER-α  | ss12568797 | A-G | Exon 1 | 0.484 | TTGAAGCGGGCGACCCACTCA | CGCCGGCTTGGAGCTTC |
| 10      | ER-α  | ss12568796 | T-C | Intron 1 | 0.449 | TGAGCTTCCAGGCTGAC | TGGACGCTAGTACAGTGAGCG |
| 11      | ER-α  | ss12568619 | G-A | Intron 3 | 0.257 | CCAGAGAACAAGTCTGATGCTGACA | TTTACACAAAGGGAGGAGG |
| 12      | ER-α  | ss12568618 | G-C | Exon 4 | 0.218 | ACACGGCGCTGAGCTTTCC | CTGGAGAGGCGAGGCAG |
| 13      | ER-α  | ss12568605 | G-A | Intron 6 | 0.119 | GAGAGGGTTGAGGTG | CTGCTCATGAGTTTTTCAG |
| 14      | ER-α  | ss12568617 | G-A | Exon 10 | 0.380 | CTCAACGTCGAGAGC | CTTGCCCACTTGCTGAAAGAC |
| 15      | IL-6  | ss12568616 | C-G | 5’ UTR | 0.434 | GCCGGAGTGGCCTAGGCA | CAGATGCAAGAGGGTTGCTCAATA |
| 16      | IL-6  | ss12568586 | A-T | Exon 4 | 0.005 | CCTTCCAGGCTGAGC | AGCGAGGATTCAGTCATTCAGAG |
| 18      | LEPR  | ss12568615 | A-G | Exon 14 | 0.290 | AGATTTAGTGTCTGCTGAGCATCAGC | TTAAGGGCACAGTCCGATGGTCTTC |
| 19      | LEPR  | ss12568604 | G-C | Exon 14 | 0.182 | GAGAATTTTGAGAATTAC | GCCTGCCACTGAGTTTGGAG |
| 20      | LEPR  | ss12568614 | G-A | Exon 20 | 0.353 | GCCACGCTGATCAGCAACTTC | GCCTGCCACAGTCATTCGAAAGAC |
| 21      | PTHR1 | ss12568597 | G-A | Intron 1 | 0.001 | GAATTTAGCAGGAGGAGC | CTTGCCCACTTGCGAGAGC |
| 22      | PTHR1 | ss12568592 | A-G | Intron 2 | 0.398 | GAGCCAGGAAGGCTTGCA | CTTGCCCACTTGCGAGAGC |
| 23      | PTHR1 | ss12568591 | C-T | Intron 8 | 0.374 | CCAGGCTTCAATCCAGTGATGACTACA | CTTGCCCACTTGCGAGAGC |
| 24      | PTHR1 | ss12568588 | G-A | Intron 10 | 0.397 | CCTTGAGGCTGGTCTT | CTTGCCCACTTGCGAGAGC |
| 25      | PTHR1 | ss12568580 | G-A | Intron 10 | 0.380 | GCCACGCTGATGCGAC | CTTGCCCACTTGCGAGAGC |
| 26      | TGF-β1 | ss12568613 | C-T | 5’ UTR | 0.311 | GGGCCGCTTGATGCTGCTGAA | GGGGCCGCGTTCCGAG |
| 27      | TGF-β1 | ss12568603 | C-C | Intron 4 | 0.021 | CCAGCCGCTTAAATTCTTCA | CGAGGGAGGCGAGGGAG |
| 28      | TGF-β1 | ss12568593 | G-C | Exon 5 | 0.008 | GAGCTCTAGGCAGGTTAAC | GTTCTACTCCGCAGAAG |
| 29      | TGF-β1 | ss12568602 | C-T | Exon 5 | 0.274 | GGGCTTGTCTTGGATGAAAACAT | TGACTAGCTGCGGAGCC |
| 30      | TNFR2 | ss12568594 | T-C | 5’ UTR | 0.000 | TGCACTGGCTGGTGTTTAGACT | CGTCTACTGCGGAGCC |
| 31      | TNFR2 | ss12568611 | T-G | Exon 6 | 0.198 | AGCACCCGCCAGCCTAC | CTGGGAGGCTGGTGGTTC |
| 32      | UCP3  | ss12568601 | C-T | 5’ UTR | 0.265 | CACTGCCCTGCCGAGC | GGTGCTGGCAGGAGC |
| 33      | UCP3  | ss12568600 | G-A | Exon 2 | 0.000 | GCCCTAAAAGGGAGCCTGCA | CTTGCCCACTTGCGAGAGC |
| 34      | UCP3  | ss12568589 | T-C | Exon 5 | 0.447 | CTTAAGAGGAATTTCCGCGGATTCTC | TCCACGAGGTCTGGTTTCC |
| 35      | UCP3  | ss12568599 | C-T | Exon 7 | 0.000 | GCCACGCTGCTGCAGAGGAGC | CTTGCCCACTTGCGAGAGC |
| 37      | VDR   | ss12568583 | G-A | 5’ UTR | 0.281 | CAGACGGCTGCCTGCTGAC | CAGATAGCTGCGGAGCC |
| 38      | VDR   | ss12568581 | C-T | Exon 2 | 0.373 | TGGGGCTGCCAGTGTTC | GGACGCTGCGGAGCC |
| 39      | VDR   | ss12568582 | C-T | Exon 4 | 0.000 | GACACGCTGGGCTGCA | CTTGCCCACTTGCGAGAGC |
| 40      | VDR   | ss12568610 | G-A | Intron 8 | 0.419 | GTGCCCTCCTACTGCCCTT | CCTCCCCAACTACAGGTTGAGC |
| 41      | VDR   | ss12568608 | T-G | Exon 9 | 0.408 | GGGCCAGGAGCGTGATTC | AGGTGCGCTGAGTTTCAG |

*aDesignation in the present study. bAbbreviations: APOE, apolipoprotein E; COL1α1, collagen type I α1; ER-α, estrogen receptor-α; IL-6, interleukin-6; LEPR, leptin receptor; PTHR1, parathyroid hormone (PTH)/PTH-related peptide receptor type 1; TGF-α1, transforming growth factor-β1; TNFR2, tumor necrosis factor receptor 2; UCP3, uncoupling protein 3; VDR, vitamin D (1,25-dihydroxyvitamin D3) receptor. cCommon name. dMinor alleles are given in bold.
Data on the candidate SNPs for obesity and/or osteoporosis from major ethnic groups: an analysis of literature and databases

Caucasians

In available literature and public SNP databases, we found data about allele frequencies of 29 SNPs (representing 9 genes) in Caucasians, which were shared with the present study. The corresponding data are given in Table 3. The allele frequencies of only 14 SNPs had been previously reported for Caucasians to the public databases (dbSNP and/or HGVbase). No significant differences were found between the allele frequencies in our study and in the public databases, except for SNP31 in Australians ($P = 0.05$, Table 3).

In contrast to the databases, available literature contained data on allele frequencies of the 23 SNPs. We determined significant differences between our and literature data for allele frequencies of four SNPs (SNP16, SNP32, SNP35, and SNP36). We have not found any allele frequency data on the SNPs of the PTHR1 gene (SNP21-SNP25) for Caucasians. Some SNP indicated large discrepancies in allele frequencies between the data from the different databases. For example, the minor allele frequencies of SNP26 for two Caucasian samples reported to dbSNP and HGVbase are 0.43 and 0.26, respectively (Table 3). The differences between the values are not statistically significant due to the small sample sizes ($n = 31$ and $n = 42$, respectively), so it is impossible to conclude what actual allele frequencies are in these populations and whether they are indeed significantly different.

African/African-Americans

Table 4 presents our SNP data in comparison with the corresponding data on African and/or African-American samples obtained from literature and the databases. In the public domain, we found in total only 18 SNPs from seven genes, which are shared with our study. The candidate genes for obesity and/or osteoporosis appear to be underrepresented for Africans/African-Americans in the public SNP databases and literature: only 11 shared SNPs were found in the databases and seven in the literature. Among those, one (SNP36) was monomorphic and four had minor allele frequency <0.1. Significant differences in allele frequencies between Caucasians and Africans/African-Americans were found at three SNPs (SNP16, SNP34 and SNP35). Three more SNPs (SNP32, SNP33, and SNP41) manifested nearly significant differences in allele frequencies between these ethnic groups (Table 4). We found no data on the SNPs of three genes (COL1α1, PTHR1, and TNFR2) for Africans/African-Americans in the public domain, which would be shared with our study.

Similar to the case with Caucasians, there are large inconsistencies between some SNP data for Africans/African-Americans from the different databases. For example, the reported minor allele frequencies of SNP26 from the TSC and dbSNP databases are 0.27 and 0.15, respectively (Table 4). The former is not significantly different ($P = 0.61$) from the corresponding value in Caucasians, while the latter is nearly significantly different ($P = 0.07$). However, it is hard to determine, which value is true because of the small sample sizes ($n = 42$ and $n = 24$, respectively) in the databases.

Asians

As compared to our SNP data, this ethnic group had corresponding data on 22 SNPs (8 genes) in the public domain (Table 5). The data on 13 SNPs of the 6 genes were available from the public databases, and on 10 SNPs of the 5 genes were reported in the literature. Among those SNPs, two (SNP16 and SNP33) were monomorphic and one (SNP36) had minor allele frequency <0.1. Two more SNPs of the VDR gene (SNP40 and SNP41) have a minor allele frequency about this value; however, this frequency varies in different Asian populations. We observed a significant inconsistency between the values of a minor allele frequency for SNP40 in Asians from the database and literature, respectively. In the database, this value (0.41) is about 3–10 times higher than that reported in the literature (0.043–0.13).
Table 3: Comparison of SNP allele frequencies of the studied Caucasian sample with the data on Caucasians from databases and literature

| Gene | SNP  | Reference population | Database or reference | Data from a database or reference | Caucasians |
|------|------|----------------------|-----------------------|----------------------------------|------------|
|      |      |                      |                       | Frequency | n | Frequency | n |
| APOE | 1    | American Caucasians  | [53]                  | 0.33      | 220 | 0.357     | 1,862 | 0.48 |
|      | 2    | American Caucasians  | dbSNP                 | 0.33      | 24  | 0.399     | 1,839 | 0.64 |
|      |      | Europeans from North Karelia, Finland | dbSNP | 0.4      | 24  | 0.399     | 1,839 | 0.84 |
|      | 3    | European-Americans  | dbSNP                 | 0.12      | 24  | 0.149     | 1,858 | 0.91 |
|      |      | Caucasians          | HGVbase               | 0.16      | 152 | 0.149     | 1,858 | 0.81 |
|      | 4    | European-Americans  | dbSNP                 | 0.23      | 24  | 0.149     | 1,858 | 0.37 |
|      |      | Caucasians          | HGVbase               | 0.08      | 152 | 0.087     | 1,783 | 0.87 |
|      |      | Europeans from North Karelia, Finland | dbSNP | 0.04      | 24  | 0.087     | 1,783 | 0.25 |
| COL1α | 5    | White women with Spanish ancestors | [54] | 0.13      | 256 | 0.154     | 1,856 | 0.36 |
|      | 6    | American Caucasians  | [31]                  | 0.187     | 637 | 0.188     | 1,855 | 0.98 |
| ER-α | 9    | Australian females  | [55]                  | 0.44      | 125 | 0.484     | 1,861 | 0.39 |
|      |      | Greek males         | [56]                  | 0.49      | 50  | 0.484     | 1,861 | 0.95 |
|      | 10   | American females    | [57]                  | 0.45      | 253 | 0.449     | 1,861 | 0.96 |
|      |      | Finnish females     | [58]                  | 0.42      | 322 | 0.449     | 1,861 | 0.38 |
|      | 11   | SANGER 12 DNAs of Caucasian origin | dbSNP | 0.08      | 12  | 0.257     | 1,839 | 0.32 |
|      | 12   | Australian females  | [55]                  | 0.23      | 120 | 0.218     | 1,859 | 0.84 |
|      |      | Slovenian females   | [59]                  | 0.3       | 85  | 0.218     | 1,859 | 0.10 |
|      |      | Caucasians          | dbSNP                 | 0.27      | 31  | 0.218     | 1,859 | 0.59 |
|      | 13   | Caucasians          | HGVbase               | 0.08      | 42  | 0.098     | 1,863 | 1.00 |
|      | 15   | Caucasians          | dbSNP                 | 0.18      | 31  | 0.203     | 1,863 | 0.90 |
| IL-6 | 16   | Caucasians          | dbSNP                 | 0.5       | 31  | 0.434     | 1,860 | 0.58 |
|      |      | English males       | [60]                  | 0.57      | 2,560| 0.434     | 1,860 | <0.01 |
|      |      | Italians            | [61]                  | 0.45      | 183 | 0.434     | 1,860 | 0.74 |
|      |      | Spaniards           | [62]                  | 0.56      | 118 | 0.434     | 1,860 | 0.01 |
|      |      | Finns               | HGVbase               | 0.46      | 400 | 0.434     | 1,860 | 0.37 |
|      |      | Finns               | HGVbase               | 0.48      | 50  | 0.434     | 1,860 | 0.62 |
|      |      | Finns               | HGVbase               | 0.39      | 42  | 0.434     | 1,860 | 0.68 |
|      |      | Finns               | [63]                  | 0.45      | 400 | 0.434     | 1,860 | 0.60 |
| LEPR | 18   | English males       | HGVbase               | 0.27      | 322 | 0.290     | 1,839 | 0.41 |
|      |      | Caucasians          | [64]                  | 0.28      | 56  | 0.290     | 1,839 | 0.98 |
|      |      | Swedish males       | [65]                  | 0.24      | 284 | 0.290     | 1,839 | 0.09 |
|      | 19   | Swedish males       | [65]                  | 0.15      | 284 | 0.182     | 1,776 | 0.21 |
|      | 20   | Caucasians          | [64]                  | 0.38      | 56  | 0.353     | 1,860 | 0.79 |
|      |      | Caucasians          | dbSNP                 | 0.4       | 42  | 0.353     | 1,860 | 0.64 |
|      |      | English males       | HGVbase               | 0.36      | 322 | 0.353     | 1,860 | 0.77 |
| TGF-ß1 | 26   | Caucasians          | dbSNP                 | 0.43      | 31  | 0.311     | 1,840 | 0.21 |
|      |      | Caucasians          | HGVbase               | 0.26      | 42  | 0.311     | 1,840 | 0.59 |
|      |      | Caucasians          | [66]                  | 0.02      | 304 | 0.021     | 1,863 | 0.98 |
|      | 28   | Danes               | [66]                  | 0.027     | 302 | 0.008     | 190   | 0.33 |
|      |      | Caucasians          | female DZ twins       | [67]      | 0.25  | 1,802 | 0.274     | 1,852 | 0.08 |
| TNFR2 | 31   | UK Caucasians       | [68]                  | 0.23      | 192 | 0.198     | 1,842 | 0.39 |
|      |      | Germans             | [69]                  | 0.25      | 94  | 0.198     | 1,842 | 0.33 |
|      |      | Australians         | HGVbase               | 0.26      | 197 | 0.198     | 1,842 | 0.05 |
| UCP3 | 32   | American Caucasians | [70]                  | 0.21      | 56  | 0.265     | 1,767 | 0.49 |
|      |      | Danes               | [71]                  | 0.27      | 857 | 0.265     | 1,767 | 0.85 |
|      |      | French              | [72]                  | 0.22      | 894 | 0.265     | 1,767 | 0.01 |
|      |      | Caucasians          | [73]                  | 0.01      | 165 | 0.000     | 295   | 0.44 |
|      |      | American Caucasians | [70]                  | 0.2       | 58  | 0.258     | 1,862 | 0.38 |
|      | 35   | Caucasians          | [74]                  | 0.5       | 501 | 0.447     | 1,863 | 0.04 |
For gene abbreviations and SNP designations see Table 1. a Present study. b Frequency of the allele corresponding to a minor allele in the present study. c Caucasians with unidentified ethnic background.

| Gene | SNP | Reference population | Database or references | Data from a database or reference | Caucasiansa | P |
|------|-----|----------------------|------------------------|----------------------------------|-------------|---|
| APOE | 2   | African-Americans    | dbSNP                  | 0.48                             | 0.399 1.839 | 0.52 |
|      | 3   | African-Americans    | dbSNP                  | 0.10                             | 0.149 1.858 | 0.56 |
|      | 4   | African-Americans    | dbSNP                  | 0.04                             | 0.087 1.783 | 0.25 |
| ER-α | 11  | SANGER 12 DNAs of African-American origin dbSNP | 0.08 | 12 | 0.257 1.857 | 0.32 |
|      | 12  | African/African-Americans dbSNP | 0.27 | 24 | 0.218 1.859 | 0.61 |
|      | 13  | TSC panel of African-Americans dbSNP | 0.16 | 40 | 0.098 1.863 | 0.27 |
|      | 15  | African/African-Americans dbSNP | 0.28 | 23 | 0.203 1.863 | 0.42 |
| IL-6 | 16  | African/African-Americans dbSNP | 0.96 | 24 | 0.434 1.860 | <0.01 |
| LEPR | 20  | TSC panel of African-Americans dbSNP | 0.45 | 42 | 0.353 1.860 | 0.25 |
| TGF-81 | 26 | TSC panel of African-Americans African/African-Americans dbSNP | 0.27 | 42 | 0.311 1.840 | 0.61 |
|      |     |                      |                        |                                  |             |    |
| UCP3 | 32  | African-American females | [70]                  | 0.15                             | 0.265 1.767 | 0.07 |
|      | 33  | African-Americans      | [73]                  | 0.08                             | 0.000 295   | 0.06 |
|      | 34  | African-American females | [70]                  | 0.45                             | 0.258 1.862 | <0.01 |
|      | 35  | African-American females | [70]                  | 0.247                            | 0.447 1.863 | <0.01 |
|      | 36  | African-Americans      | [70]                  | 0.20                             | 0.447 1.863 | <0.01 |
| VDR  | 38  | African-American males | [77]                  | 0.244                            | 0.373 1.861 | 0.15 |
|      | 40  | African-American males | [77]                  | 0.312                            | 0.419 1.860 | 0.30 |
|      |     | Nigerians              | [82]                  | 0.37                             | 0.419 1.860 | 0.41 |
|      | 41  | African/African-Americans | dbSNP                  | 0.35                             | 0.408 1.859 | 0.71 |
|      |     | African-American males | [77]                  | 0.244                            | 0.408 1.859 | 0.06 |

For gene abbreviations and SNP designations see Table 1. a Present study. b Frequency of the allele corresponding to a minor allele in the present study.
Table 5: Comparison of SNP allele frequencies of the studied Caucasian sample with the data on Asian populations from databases and literature

| Gene | SNP | Reference population | Database or reference | Data from a database or reference | Caucasiansa | P     |
|------|-----|----------------------|-----------------------|-----------------------------------|-------------|-------|
|      |     |                      |                       | Frequencyb n                      | Frequency n |       |
|      |     |                      |                       |                                   |             |       |
| ER-α | 9   | Japanese             | [83]                  | 0.412 200                         | 0.484 1.861 | 0.06  |
|      |     | Thai females         | [84]                  | 0.337 129                         | 0.484 1.861 | <0.01 |
|      | 10  | Korean males         | [85]                  | 0.39 219                         | 0.449 1.861 | 0.11  |
|      |     | Korean females       | [86]                  | 0.415 248                         | 0.449 1.861 | 0.35  |
|      |     | Japanese children    | [87]                  | 0.431 102                         | 0.449 1.861 | 0.80  |
|      |     | Taiwanese            | [88]                  | 0.39 246                         | 0.449 1.861 | 0.09  |
|      |     | Thai females         | [89]                  | 0.414 134                         | 0.449 1.861 | 0.49  |
|      | 11  | SANGER 12 DNAs of Asian origin dbSNP | | 0.36 12 | 0.257 1.857 | 0.49 |
|      |     | Japanese             | [83]                  | 0.483 200                         | 0.218 1.859 | <0.01 |
|      |     | Japanese females     | [90]                  | 0.481 306                         | 0.218 1.859 | <0.01 |
|      | 13  | Japanese             | JSNP                  | 0.34 744                         | 0.098 1.863 | <0.01 |
|      |     | TSC panel of unrelated Asians | TSC | 0.27 32 | 0.098 1.863 | 0.01 |
|      | 14  | Japanese             | JSNP                  | 0.296 752                         | 0.119 1.861 | <0.01 |
|      | 15  | Japanese             | JSNP                  | 0.18 741                         | 0.203 1.863 | 0.21  |
|      |     | Japanese             | [83]                  | 0.21 200                         | 0.203 1.863 | 0.88  |
| IL-6 | 16  | Chinese              | [31]                  | 0.997 147                         | 0.434 1.860 | <0.01 |
|      |     | Chinese males        | [31]                  | 0.998 259                         | 0.434 1.860 | <0.01 |
|      |     | Japanese             | [31]                  | 1.000 388                         | 0.434 1.860 | <0.01 |
| LEPR | 20  | TSC panel of unrelated Asians | dbSNP | 0.13 42 | 0.353 1.860 | <0.01 |
|      |     | Japanese             | JSNP                  | 0.15 748                         | 0.353 1.860 | <0.01 |
| PTHR1 | 22  | Japanese             | JSNP                  | 0.43 747                         | 0.398 1.861 | 0.14  |
|      |     | Japanese             | JSNP                  | 0.43 748                         | 0.374 1.863 | 0.01  |
|      |     | Japanese             | JSNP                  | 0.44 750                         | 0.397 1.855 | 0.05  |
|      |     | Japanese             | JSNP                  | 0.43 746                         | 0.380 1.851 | 0.02  |
| TGF-ß1 | 26  | TSC panel of unrelated Asians | dbSNP | 0.47 36 | 0.311 1.840 | 0.06 |
| TNFR2 | 31  | Japanese             | [91]                  | 0.113 265                         | 0.198 1.842 | <0.01 |
|      |     | Thai                 | [92]                  | 0.127 201                         | 0.198 1.842 | 0.02  |
| UCP3 | 33  | Asiansc              | [73]                  | 0.00 11                          | 0.000 295   | 10.00 |
|      | 35  | Japanese             | JSNP                  | 0.45 750                         | 0.447 1.863 | 0.92  |
|      | 36  | Asiansc              | [73]                  | 0.04 11                          | 0.000 295   | 0.54  |
| VDR  | 37  | Japanese             | JSNP                  | 0.23 750                         | 0.281 1.861 | 0.01  |
|      | 38  | Japanese             | [93]                  | 0.37 195                         | 0.373 1.861 | 0.99  |
|      | 40  | SANGER 12 DNAs of Asian origin dbSNP | | 0.41 12 | 0.419 1.860 | 0.77 |
|      |     | Chinese females (Han nationality) | [94] | 0.043 162 | 0.419 1.860 | <0.01 |
|      |     | Taiwanese            | [95]                  | 0.083 90                          | 0.419 1.860 | <0.01 |
|      |     | Japanese             | [93]                  | 0.13 195                         | 0.419 1.860 | 0.02  |
|      |     | Koreans              | [96]                  | 0.055 211                         | 0.419 1.860 | <0.01 |
|      |     | Thai females         | [97]                  | 0.107 84                          | 0.419 1.860 | 0.01  |
|      | 41  | Chinese (Han nationality) | [98] | 0.05 223 | 0.408 1.859 | <0.01 |
|      |     | Japanese females     | [99]                  | 0.135 119                         | 0.408 1.859 | <0.01 |
|      |     | Koreans              | [96]                  | 0.055 120                         | 0.408 1.859 | <0.01 |

For gene abbreviations and SNP designations see Table 1. a Present study. b Frequency of the allele corresponding to a minor allele in the present study. c Asians with unidentified ethnic background.
A comparison of our Caucasian sample and the Asian populations revealed significant differences in allele frequencies of 13 SNPs located in the 6 genes (Table 5). Another SNP representing the TGF-β1 gene (SNP26) indicated the nearly significant difference ($P = 0.06$). No such differences were observed only in the UCP3 gene.

**Hispanic and Pacific Rim populations**

These populations had the poorest data on the SNPs from the databases and literature as compared to the corresponding data from the present study. We found the data on only seven SNPs of the five candidate genes for Hispanics (Table 6) and five SNPs of the four genes for Pacific Rim populations (Table 7). SNP33 and SNP36 for the UCP3 gene were virtually monomorphic in Hispanics, similar to that in our Caucasian sample, while only one significant difference in allele frequencies was observed at SNP16 between these ethnic populations among the seven compared SNPs (Table 6).

| Gene       | SNP | Database or references | Data from a database or reference | Caucasians$^a$ | $P$  |
|------------|-----|------------------------|-----------------------------------|----------------|-----|
| ER-α       | 12  | dbSNP                  | Frequency$^b$ 23: 0.239            | Frequency 23: 0.218 | 1.859: 1.00  |
|            | 15  | dbSNP                  | Frequency$^b$ 23: 0.196            | Frequency 23: 0.203 | 1.863: 1.00  |
| IL-6       | 16  | dbSNP                  | Frequency$^b$ 22: 0.795            | Frequency 22: 0.434 | 1.860: <0.01 |
| TGF-β1     | 26  | dbSNP                  | Frequency$^b$ 22: 0.413            | Frequency 22: 0.311 | 1.840: 0.36  |
| UCP3       | 33  | [73]                   | Frequency$^b$ 27: 0.00             | Frequency 27: 0.000 | 295: 1.00   |
|            | 36  | [73]                   | Frequency$^b$ 27: 0.02             | Frequency 27: 0.000 | 295: 1.00   |
| VDR        | 41  | dbSNP                  | Frequency$^b$ 23: 0.348            | Frequency 23: 0.408 | 1.859: 0.83  |

For gene abbreviations and SNP designations see Table 1. $^a$Present study. $^b$Frequency of the allele corresponding to a minor allele in the present study.

| Gene       | SNP | Database | Data from a database or reference | Caucasians$^a$ | $P$  |
|------------|-----|----------|-----------------------------------|----------------|-----|
| ER-α       | 12  | dbSNP    | Frequency$^b$ 23: 0.478            | Frequency 23: 0.218 | 1.859: 0.02  |
|            | 15  | dbSNP    | Frequency$^b$ 23: 0.326            | Frequency 23: 0.203 | 1.863: 0.19  |
| IL-6       | 16  | dbSNP    | Frequency$^b$ 23: 0.957            | Frequency 23: 0.434 | 1.860: <0.01 |
| TGF-β1     | 26  | dbSNP    | Frequency$^b$ 24: 0.458            | Frequency 24: 0.311 | 1.840: 0.19  |
| VDR        | 41  | dbSNP    | Frequency$^b$ 23: 0.13             | Frequency 23: 0.408 | 1.859: <0.01 |

For gene abbreviations and SNP designations see Table 1. $^a$Present study. $^b$Frequency of the allele corresponding to a minor allele in the present study.
No data on the SNPs of the 10 candidate genes for obesity and/or osteoporosis were found for Pacific Rim populations in the available literature. All the data presented in Table 7 were obtained from the dbSNP database. Three SNPs out of the five differed significantly in the allele frequencies between Pacific Rim populations and the Caucasian sample in our study.

Some new candidate SNPs for genetic studies of obesity and osteoporosis
In this study, we identified a number of SNPs, which may be suitable for genetic studies of obesity and osteoporosis. In particular, the six polymorphic SNPs with minor allele frequency ≥0.1, which were firstly determined in Caucasians (Table 2), may be used in population association studies. A number of SNPs with significant variation in allele frequencies in populations of different ethnicity may be appropriate for studying a genetic basis of between-ethnic differences in the rates of obesity and/or osteoporosis. Examples are SNP34, SNP35, and SNP32 for Caucasians and African-Americans (Table 4) and several SNPs located in the ER-α, LEPR, PTHR1, TNFR2, and VDR genes for Caucasians and Asians (Table 5). Some of the SNPs may also be candidates for such studies, if their allele frequencies reported in the databases are validated by using sufficiently large sample sizes (e.g., SNP26 in Asians).

Discussion
Ethnic variation of SNP allele frequencies of the candidate genes for obesity and/or osteoporosis
We observed significant differences in allele frequencies of some studied SNPs for Caucasians (SNP16, SNP35, SNP32, and SNP36, Table 3) between our results and respective data from the databases and literature. This may be due to various factors, such as population admixture and relatively smaller sample size in the other studies [34]. For example, the sample in the present study consisted of the subjects of various ethnic backgrounds (German, French, Dutch, Swedish, some Portuguese and Italian backgrounds) and the relative proportions of these ethnic groups are not ascertained. In the mixed samples from other studies, these proportions may be different that thus influence the allele frequencies to a larger or lesser extent (e.g., SNP35 and SNP36, Table 3). Likewise, monoethnic samples may have allele frequencies significantly different from those in the mixed samples (e.g., SNP16 in English and Spaniards, SNP32 in French, Table 3). Our results showed that major ethnic groups have significantly different allele distribution at some SNP markers of candidate genes for complex disorders. The largest differences were observed between Asian and Caucasian populations, namely at 13 SNPs out of 22 compared (Table 5). Some genes indicated the significant differences at most SNPs compared (e.g., PTHR1 between Japanese and Caucasians, Table 5). These results are consistent with the previously reported data about high differentiation between Asians and Caucasians at candidate genes for bone mass [35,36].

Another important finding is the observed significant differences between Caucasian and Pacific Rim populations in SNP allele frequencies of three important candidate genes for osteoporosis, ER-α, IL-6, and VDR (Table 7). Interestingly, the respective SNP allele frequencies in Asian and Pacific Rim populations have very similar values (Tables 5 and 7). This may suggest that the Pacific Rim sample from the database has mainly Asian ancestry [37] or, perhaps there are some implications for an evolutionary history of these populations [38]. In any event, much more comprehensive data should be obtained to clarify these issues.

There are abundant data about significant between-ethnic differentiation at loci, which may underlie complex diseases, including obesity and osteoporosis (see, e.g., [36,39-41]). Given that there is well-known different incidence of these disorders in various ethnic groups (e.g., [42,43]), such genetic differentiation may have an important implication for studying a genetic basis of ethnicity-specific definitions of obesity and osteoporosis.

Another important application of the SNP markers with high large between-ethnic and small within-ethnic differences is MALD studies of complex diseases in the admixed populations with known parental ethnicities [33,44]. Our results showed that such SNPs exist in the candidate genes for obesity and osteoporosis. However, the large-scale screening of the candidate genes is necessary to identify such markers with sufficient density.

The results of our study suggest that ethnic heterogeneity of large samples may notably affect the observed allele frequencies. It is supported by the fact that several SNP allele frequencies of some sufficiently large monoethnic samples (e.g., SNP16 in English, SNP32 in French) significantly differ from those determined in our sample of mixed Caucasians (Table 3).

SNP databases: current limitations for studying complex diseases
Our analyses showed that the SNP databases in their current status might have some limitations for studies of complex disorders, especially in different ethnic groups, due to incomplete and/or uneven representation of SNPs and/or candidate genes in these groups. As indicated above, of the ten candidate genes examined here, only four have corresponding but incomplete SNP data in the databases for all the major ethnic groups and may be used in the comparative studies of obesity and/or osteoporosis.
in the populations of different ethnicity. The SNP data in the databases for the other six genes (60% examined) need substantial updating. To do this, large-scale studies should be performed for the other major ethnic populations. Given that many complex diseases have different rates in different ethnic groups, the extensive volume of the SNP data needs to be updated and validated. This conclusion essentially corresponds to the recently reported results of the SNP databases evaluation regarding their use for whole-genome association studies in humans [21].

Our study provides an example showing the incompleteness of the SNP data in the current databases for studying complex disorders with an ethnic-dependent background.

Another important problem is inconsistency of the data for some SNP markers between either the different databases or databases and literature. It makes selection of a right SNP for a study rather difficult. For example, a SNP is usually considered to be appropriate for association or linkage studies, if its minor allele frequency is ≥ 0.1 [45,46]. Based on the data of SNP11 from the dbSNP database (Table 3), this SNP is hardly appropriate for the studies in Caucasians, because its minor allele frequency is 0.08. In fact, as determined in the present study, its actual frequency is 0.257 that makes this SNP suitable for the population association and linkage research.

In terms of well-known ethnic differences in incidence of some complex diseases, the discrepancies in the SNP database data may yield wrong conclusion about suitability of particular SNPs for studying genetic basis of these differences. For example, as was mentioned above, SNP26 from the TGF-ß1 gene has the reported minor allele frequency 0.43 (dbSNP) and 0.26 (HGIVbase) for Caucasians, 0.27 (TSC) and 0.15 (dbSNP) for African-Americans, and 0.47 (dbSNP) for Asians (Tables 3,4,5). This gene is a candidate for bone mineral density [47] and, thus, may contribute to the ethnic differences in bone mass and osteoporosis. However, from the above data on the SNP allele frequencies, it is impossible to infer whether this particular SNP may be related to these ethnic differences in bone mass, because none of the values significantly differ from any other due to small sample size and, accordingly, limited statistical power.

Large discrepancies in some allele frequency data between the different databases (e.g., SNP4 for Caucasians) are likely due to small sample sizes and respective large sampling errors of the estimates. The sample sizes in the SNP databases are usually smaller than 50. With such a sample size, a relative sampling error of the allele frequency estimates ranges from 0.12 to 0.67 (fig. 1). In this term, the estimates of the SNP allele frequencies in the present study with the sample size of 1,873 are far more reliable than the respective data from the most databases. Estimation of allele frequencies in small samples is notably affected by heterogeneity of the samples. Given that information about an ethnic background of the samples in the SNP databases is usually scarce, the allele frequency estimates in the databases may be significantly biased due to the ethnic admixture.

The small size of the samples in the databases makes it difficult to powerfully test differences in SNP allele frequencies (Δf) between the data from various sources, especially if absolute values of these differences are not large. Fig. 2 illustrates a statistical power of such estimation. Given two samples of size 50 each, the probability to correctly determine Δf = 0.1 is only 27% (fig. 2A). Having a sample size of 1,800 (similar to ours) increases this probability only up to 40% (fig. 2B). Even a 3-fold difference in the absolute values (e.g., SNP11, Table 4) is not statistically significant due to the insufficient sample size in the databases. These examples suggest that some differences in SNP allele frequency data in the public databases may exist due to improper sampling or too small sample sizes [34]. These inconsistencies may be somewhat reduced by pooling the samples. However, this should be done with caution, because pooling data from small samples with a particular genetic background may introduce a large bias to the new values.

![Figure 1](http://www.biomedcentral.com/1471-2156/5/4)
Furthermore, in many cases, even after pooling, the sample size may remain too small to gain sufficient reliability of the newly obtained allele frequencies.

The SNP data in literature may somewhat supplement those from the databases. They are sometimes obtained with larger samples and, therefore, have smaller sampling error. However, they are not systematized, which makes their collection and use difficult.

In conclusion, our study demonstrated that, although a large volume of SNP data is available in public databases and literature, a great portion of these data needs comprehensive updating and validating in order to be appropriate for genetic studies of complex disorders, such as obesity and/or osteoporosis. Such large-scale studies of the disease-associated SNPs in various ethnicities may provide important insights into the evolutionary history of human populations as well as in etiology of these diseases.

**Methods**

**Subjects**

All the study subjects came from ongoing genetic studies of complex traits that have been approved by the Creighton University Institutional Review Board. All the subjects were Caucasians of western or northern European origin (German, French, Dutch, Swedish), but some had Portuguese and Italian backgrounds. We have recruited 405 nuclear families, each composed of both parents and at least one child. The total sample size was 1,873 subjects, including 840 parents and 1,133 children. All individuals volunteered to participate in the research and signed informed-consent documents before entering the project.

**The study SNPs**

The candidate SNPs and genes for the present study were selected from publicly available sources based on some of the following criteria: 1) functional relevance and importance for obesity and/or osteoporosis; 2) degree of heterozygosity, i.e., allele frequencies, as reported in literature or databases; 3) position in or around the genes; and 4) their use in previous genetic epidemiology studies. After searching public SNP databases (dbSNP, JSNP, HGVbase, and TSC) and literature, we chose 41 SNPs in 10 genes shown to be associated with obesity and/or osteoporosis. The information about the SNPs is given in Table 1. Among them, SNP27 is an insertion/deletion polymorphism of a cytosine (+/- C) and the others are nucleotide substitutions.

**SNP genotyping**

DNA was extracted from whole blood using a commercial isolation kit (Genta Systems, Minneapolis, MN, USA) following the procedure recommended by the manufacturer. The genotyping involved a polymerase chain reaction (PCR) and invader assay (Third Wave Technology,
FRET mix, and 0.5 µl Cleavase enzyme/MgCl₂ solution diluted PCR product, 1.5 µl probe mix, 1.75 µl Cleavase were carried out in 7.5 µl reaction volume containing 3.75 µl 1:20 in nuclease-free water. The invader reaction was carried out in 7.5 µl reaction volume containing 3.75 µl diluted PCR product, 1.5 µl probe mix, 0.5 µl Cleavase FRET mix, and 0.5 µl Cleavase enzyme/MgCl₂ solution (Third Wave Technology). The reaction mix was overlaid by 15 µl mineral oil and denatured at 95°C for 5 min, and then incubated at 63°C for 20 min on PE9700 Thermal Cycler. After the incubation, the fluorescence intensity for both colors (FAM dye and Red dye) was measured by Cytofluor 4000 (ABI). The data were then loaded to the Invader Analyzer software (Third Wave Technology), and the genotype for every sample was called according to the ratio of the fluorescence intensity of the two dyes.

We initially genotyped all the 41 SNPs in a random sample of 190 to 380 subjects to find SNPs with minor allele frequencies <0.01. Those were then excluded from the further genotyping. Finally, 33 SNPs were genotyped for the whole sample.

Data analysis

PedCheck software [48] was used to verify the accuracy of SNP genotyping in reference to Mendelian inheritance of the alleles within each family. For the 29 polymorphic SNPs the allele frequencies in the nuclear families were estimated by maximum-likelihood method [49] implemented in SOLAR [http://www.sbr.org/sbr/public/soft
ware/solar]. This method uses all available marker information by accounting the dependence between relatives. Differences in allele frequencies of each SNP among various populations were tested using the χ² test or Fisher’s exact two-tailed test as implemented in Advisor software [50]. SNPs with minor allele frequency less than 0.01 were considered to be monomorphic.

Authors’ contributions

VD combined results of the analysis and drafted the manuscript. JRL and DHX performed the literature search, SNP genotyping, and statistical analysis. PYL carried out the statistical analysis. LIZ, HS, YYZ, YIL, SA, and PX performed the SNP genotyping. RRR and HWD conceived of the study, and participated in its design and coordination. All authors read and approved the final manuscript.

Links

Database of Single Nucleotide Polymorphism, dbSNP [http://www.ncbi.nlm.nih.gov/SNP]

Japanese Single Nucleotide Polymorphisms, JSNP [http://snp.ims.u-tokyo.ac.jp]

Human Genome Variation Database, HGVS [http://hvgbase.cgb.ki.se/]

The SNP Consortium, TSC [http://snp.cshl.org/]

Acknowledgments

The investigators were partially supported by grants from Health Future Foundation of the USA, the National Institutes of Health (KO1 AR02170-01, R01 GM60402-01A1), the State of Nebraska Cancer and Smoking Related Disease Research Program, and US Department of Energy (DE-FG03-00ER63000/A00). The study also benefited from support (to HWD) of Hunan Provincial Special Professor Start-up Fund (25000612), Chinese National Science Foundation (CNSF) Outstanding Young Scientist Award (30225025), CNSF Grant (30170504), Seed Fund from the Ministry of Education of P. R. China (25000106), and a key project grant from the Ministry of Education of P. R. China. We thank all the study subjects for volunteering to participate in the study.

References

1. Taylor JG, Choi EH, Foster CB, Chanock SJ: Using genetic variation to study human disease. Trends Mol Med 2001, 7:507-512.
2. Brookes AJ: The essence of SNPs. Gene 1999, 234:177-186.
3. Miller RD, Kwok PY: The birth and death of human single-nucleotide polymorphisms: new experimental evidence and implications for human history and medicine. Hum Mol Genet 2001, 10:2195-2198.
4. Gray IC, Campbell DA, Spurr NK: Single nucleotide polymorphisms as tools in human genetics. Hum Mol Genet 2000, 9:2403-2408.
5. Lin RCY, Wang XL, Dalziel B, Caterson ID, Morris BJ: Association of obesity, but not diabetes or hypertension, with glucocorticoid receptor N363S variant. Obes Res 2003, 11:802-808.
6. Tamura K, Suzuki M, Arakawa H, Tokuyama K, Morikawa A: Linkage and association studies of STAT6 gene polymorphisms and allergic diseases. Int Arch Allergy Immunol 2003, 131:33-38.
7. Yamada Y, Ando F, Niino N, Shimokata H: Association of polymorphisms of interleukin-6, osteocalcin, and vitamin D receptor genes, alone or in combination, with bone mineral density in community-dwelling Japanese women and men. J Clin Endocrinol Metab 2003, 88:3372-3378.
8. Daly MJ, Rioux JD, Schaffner SF, Hudson TJ, Lander ES: High-resolution haplotype structure in the human genome. Nat Genet 2001, 29:229-232.
9. Jeffreys AJ, Kauppi L, Neumann R: Intensely punctate meiotic recombination in the class II region of the major histocompatibility complex. Nat Genet 2001, 29:217-222.
10. Reich DE, Cargill M, Bolk S, Ireland J, Sabeti PC, Richter DJ, Lavery T, Kouyounjian R, Farhadian SF, Ward R, Lander ES: Linkage disequilibrium in the human genome. Nature 2001, 411:199-204.
11. Patil N, Berzo AJ, Hinds DA, Barrett WA, Doshi JM, Hacker CR, Kautzer CR, Lee DH, Marjoribanks C, McDonough DP, Nguyen BT, Norris MC, Sheehan JB, Shen N, Sterin D, Stokowski RP, Thomas DJ, Trulson MO, Vyas KR, Frazer KA, Fodor SP, Cox DR: Blocks of limited haplotype diversity revealed by high-resolution scanning of human chromosome 21. Science 2001, 294:1719-1723.
12. Gabriel SB, Schaffner SF, Nguyen H, Moore JM, Roy J, Blumenstiel B, Higgins J, DeFelice M, Lochner A, Faggar M, Liu-Cordero SN, Rotimi C, Adeyemo A, Cooper R, Ward R, Lander ES, Daly MJ, Altshuler D: The structure of haplotype blocks in the human genome. Science 2002, 296:2225-2229.
13. Phillips MS, Lawrence R, Schidandaman R, Morris AP, Balding DJ, Donaldson MA, Studebaker JF, Ankener WM, Alfisi SV, Kuos F, Childs AL, Pazyrov E, Scott KE, Canlubang DJ, Faith J, Karaki G, Bhuit F, Cui YJ, Dierichsnesian V, Elsousa C, Forman AM, Grecco NM, Hock CR, Kuebler JM, Latipshak JA, Mockler MA, Nachman EF, Res- tine SL, Varde SA, Hojoza MM, Gelfand CA, Broxholme J, Abebasis GR, Boyce-Jacino MT, Cardon LR: Chromosome-wide distribution of haplotype blocks and the role of recombination hot spots. Nat Genet 2003, 33:382-387.

14. Aerts J, Wetzelz Y, Cohen N, Aerssens J: Data mining of public SNP databases for the selection of intragenic SNPs. Hum Mutat 2002, 20:162-173.

15. Tindeloss GA, Stein LD: The SNP Consortium website: past, present and future. Nucleic Acids Res 2003, 31:124-127.

16. Hirakawa M, Tanaka T, Hashimoto Y, Kuroda M, Takagi T, Nakamura Y: SNP: a database of common gene variations in the Japanese population. Nucleic Acids Res 2002, 30:158-162.

17. Marsh S, Kwok P, McLeod HL: SNP databases and pharmacogenetics: great start, but a long way to go. Hum Mutat 2002, 20:174-179.

18. Roos DE, Gabriel SB, Altshuler D: Quality and completeness of SNP databases. Nat Genet 2003, 33:457-458.

19. Marth G, Yeh R, Minton M, Donaldson R, Li Q, Duan S, Davenport R, Miller RD, Kwok PY: Single-nucleotide polymorphisms in the public domain: how useful are they? Nat Genet 2001, 27:371-372.

20. Carlson CS, Eberle MA, Rieder MJ, Smith JD, Kruglyak L, Nickerson DA: Additional SNPs and linkage-disequilibrium analyses are necessary for whole-genome association studies in humans. Nat Genet 2003, 33:518-521.

21. Kuczynski RJ, Flegel KM, Campbell SM, Johnson CL: Increasing prevalence of overweight among US adults. The National Health and Nutrition Examination Surveys, 1960 to 1991. JAMA 1994, 272:205-211.

22. Kopelman PG: Obesity as a medical problem. Nature 2000, 404:635-643.

23. Spencer H, Kramer L: Ethnic, anthropometric, and lifestyle associations with regional variations in body mass. Am J Hum Genet 2002, 70:673-685.

24. Lack of association between vitamin D receptor genotypes and osteoporosis in Koreans. Lack of an intronic Sp1 binding-site polymorphism at the collagen type I gene in healthy Korean women. Bone 1999, 24:135-137.

25. Lack of an intronic Sp1 binding-site polymorphism at the collagen type I gene in healthy Korean women. Bone 1999, 24:135-137.

26. Davis JW, Novotny R, Wasnich RD, Ross PD: Ethnic, anthropometric, and lifestyle associations with regional variations in peak bone mass. Calcif Tissue Int 1999, 65:100-105.

27. Davis JW, Novotny R, Wasnich RD, Ross PD: Ethnic, anthropometric, and lifestyle associations with regional variations in peak bone mass. Calcif Tissue Int 1999, 65:100-105.

28. Davis JW, Novotny R, Wasnich RD, Ross PD: Ethnic, anthropometric, and lifestyle associations with regional variations in peak bone mass. Calcif Tissue Int 1999, 65:100-105.

29. Davis JW, Novotny R, Wasnich RD, Ross PD: Ethnic, anthropometric, and lifestyle associations with regional variations in peak bone mass. Calcif Tissue Int 1999, 65:100-105.

30. Davis JW, Novotny R, Wasnich RD, Ross PD: Ethnic, anthropometric, and lifestyle associations with regional variations in peak bone mass. Calcif Tissue Int 1999, 65:100-105.

31. Davis JW, Novotny R, Wasnich RD, Ross PD: Ethnic, anthropometric, and lifestyle associations with regional variations in peak bone mass. Calcif Tissue Int 1999, 65:100-105.

32. Davis JW, Novotny R, Wasnich RD, Ross PD: Ethnic, anthropometric, and lifestyle associations with regional variations in peak bone mass. Calcif Tissue Int 1999, 65:100-105.

33. Davis JW, Novotny R, Wasnich RD, Ross PD: Ethnic, anthropometric, and lifestyle associations with regional variations in peak bone mass. Calcif Tissue Int 1999, 65:100-105.

34. Davis JW, Novotny R, Wasnich RD, Ross PD: Ethnic, anthropometric, and lifestyle associations with regional variations in peak bone mass. Calcif Tissue Int 1999, 65:100-105.

35. Davis JW, Novotny R, Wasnich RD, Ross PD: Ethnic, anthropometric, and lifestyle associations with regional variations in peak bone mass. Calcif Tissue Int 1999, 65:100-105.

36. Davis JW, Novotny R, Wasnich RD, Ross PD: Ethnic, anthropometric, and lifestyle associations with regional variations in peak bone mass. Calcif Tissue Int 1999, 65:100-105.

37. Davis JW, Novotny R, Wasnich RD, Ross PD: Ethnic, anthropometric, and lifestyle associations with regional variations in peak bone mass. Calcif Tissue Int 1999, 65:100-105.

38. Davis JW, Novotny R, Wasnich RD, Ross PD: Ethnic, anthropometric, and lifestyle associations with regional variations in peak bone mass. Calcif Tissue Int 1999, 65:100-105.

39. Davis JW, Novotny R, Wasnich RD, Ross PD: Ethnic, anthropometric, and lifestyle associations with regional variations in peak bone mass. Calcif Tissue Int 1999, 65:100-105.

40. Davis JW, Novotny R, Wasnich RD, Ross PD: Ethnic, anthropometric, and lifestyle associations with regional variations in peak bone mass. Calcif Tissue Int 1999, 65:100-105.

41. Davis JW, Novotny R, Wasnich RD, Ross PD: Ethnic, anthropometric, and lifestyle associations with regional variations in peak bone mass. Calcif Tissue Int 1999, 65:100-105.

42. Davis JW, Novotny R, Wasnich RD, Ross PD: Ethnic, anthropometric, and lifestyle associations with regional variations in peak bone mass. Calcif Tissue Int 1999, 65:100-105.

43. Davis JW, Novotny R, Wasnich RD, Ross PD: Ethnic, anthropometric, and lifestyle associations with regional variations in peak bone mass. Calcif Tissue Int 1999, 65:100-105.

44. Davis JW, Novotny R, Wasnich RD, Ross PD: Ethnic, anthropometric, and lifestyle associations with regional variations in peak bone mass. Calcif Tissue Int 1999, 65:100-105.

45. Davis JW, Novotny R, Wasnich RD, Ross PD: Ethnic, anthropometric, and lifestyle associations with regional variations in peak bone mass. Calcif Tissue Int 1999, 65:100-105.

46. Davis JW, Novotny R, Wasnich RD, Ross PD: Ethnic, anthropometric, and lifestyle associations with regional variations in peak bone mass. Calcif Tissue Int 1999, 65:100-105.

47. Davis JW, Novotny R, Wasnich RD, Ross PD: Ethnic, anthropometric, and lifestyle associations with regional variations in peak bone mass. Calcif Tissue Int 1999, 65:100-105.

48. Davis JW, Novotny R, Wasnich RD, Ross PD: Ethnic, anthropometric, and lifestyle associations with regional variations in peak bone mass. Calcif Tissue Int 1999, 65:100-105.

49. Davis JW, Novotny R, Wasnich RD, Ross PD: Ethnic, anthropometric, and lifestyle associations with regional variations in peak bone mass. Calcif Tissue Int 1999, 65:100-105.

50. Davis JW, Novotny R, Wasnich RD, Ross PD: Ethnic, anthropometric, and lifestyle associations with regional variations in peak bone mass. Calcif Tissue Int 1999, 65:100-105.

51. Davis JW, Novotny R, Wasnich RD, Ross PD: Ethnic, anthropometric, and lifestyle associations with regional variations in peak bone mass. Calcif Tissue Int 1999, 65:100-105.

52. Davis JW, Novotny R, Wasnich RD, Ross PD: Ethnic, anthropometric, and lifestyle associations with regional variations in peak bone mass. Calcif Tissue Int 1999, 65:100-105.

53. Davis JW, Novotny R, Wasnich RD, Ross PD: Ethnic, anthropometric, and lifestyle associations with regional variations in peak bone mass. Calcif Tissue Int 1999, 65:100-105.

54. Davis JW, Novotny R, Wasnich RD, Ross PD: Ethnic, anthropometric, and lifestyle associations with regional variations in peak bone mass. Calcif Tissue Int 1999, 65:100-105.

55. Davis JW, Novotny R, Wasnich RD, Ross PD: Ethnic, anthropometric, and lifestyle associations with regional variations in peak bone mass. Calcif Tissue Int 1999, 65:100-105.
56. Evangelopoulos D, Alevizaki M, Lekakis J, Cimponeriu A, Papamichael C, Kominais A, Kolofotis A, Moutsatsou P. Molecular analysis of the estrogen receptor α gene in men with coronary artery disease: association with disease status. Clin Chim Acta 2003, 331:37-44.

57. Willing M, Sowers MR, Aron D, Clark MK, Burns TL, Bunten C, Cruchfield M, D’Agostino D, Jannausch M. Bone mineral density and its changes in white women: estrogen and vitamin D receptor genotypes and their interaction. J Bone Miner Res 1998, 13:695-705.

58. Salmen T, Heikkinen AM, Mahonen A, Kroger H, Komulainen M, Saarikoski S, Honkanen R, Maenpaa PH. Early postmenopausal bone loss is associated with PvuII estrogen receptor gene polymorphism in Finnish women: effect of hormone replacement therapy. J Bone Miner Res 2000, 15:315-321.

59. Jurada S, Marc J, Prezelj J, Kocijancic A, Komel R. Codon 325 sequence polymorphism of the estrogen receptor α gene and bone mineral density in postmenopausal women. J Steroid Biochem Mol Biol 2001, 78:15-20.

60. Humphries SE, Luong LA, Ogg MS, Hawe E, Miller GJ. The interleukin-6 -174 G/C promoter polymorphism is associated with risk of coronary heart disease and systolic blood pressure in healthy men. Eur Heart J 2001, 22:2243-2252.

61. Flex A, Gaetani E, Pola R, Santoliquido A, Alof F, Papale P, Dal Lago A, Pola E, Serricchio M, Tondi P, Pola P. The -174 G/C polymorphism of the interleukin-6 gene promoter is associated with peripheral arterial occlusive disease. Eur J Vasc Endovasc Surg 2002, 24:624-628.

62. Vozarova B, Fernandez-Real JM, Knowler WC, Gallart L, Hanson RL, Gruber JD, Ricart W, Vendrell J, Richart C, Tataranni PA, Wolfoff JD. The interleukin-6 (-174) G/C polymorphism promoter is associated with type-2 diabetes mellitus in Native Americans and Caucasians. Hum Genet 2003, 112:49-53.

63. Hullkonnent J, Fertovaara A, Antonen J, Pasterneck A, Hurme M. Elevated interleukin-6 plasma levels are regulated by the promoter region polymorphism of the IL6 gene in primary Sjogren’s syndrome and correlate with the clinical manifestations of the disease. Rheumatology (Oxford) 2001, 40:626-631.

64. Thompson DB, Ravussin E, Bennett PH, Bogardus C. Structure and sequence variation at the human leptin receptor gene in lean and obese Pima Indians. Hum Mol Genet 1997, 6:875-879.

65. Evangelopoulos D, Alevizaki M, Lekakis J, Cimponeriu A, Papamichael C, Kominais A, Kolofotis A, Moutsatsou P. Molecular analysis of the estrogen receptor α gene in men with coronary artery disease: association with disease status. Clin Chim Acta 2003, 331:37-44.

66. Langdahl BL, Carstens M, Stenkjær L, Eriksen EF. Meirhaeghe A, Amouyel P, Helbecque N, Cottel D, Otaze S, Froguel P. An uncoupling protein 3 gene polymorphism associated with a lower risk of developing Type II diabetes and with atherogenic lipid profile in a French cohort. Diabetologia 2000, 43:1424-1428.

67. Chung WK, Luke A, Cooper RS, Rotini C, Vidal-Puig A, Rosenbaum M, Chua M, Solanes G, Zheng M, Zhao L, Le Duc C, Esberg A, Chu F, Murphy E, Schreier M, Aronne L, Caprio S, Kahle B, Gordon D, Leal SM, Goldsmith R, Andreu AL, Brung C, D’Amato S, Leibl RL. Genetic and physiologic analysis of the role of uncoupling protein 3 in human energy homeostasis. Diabetes 1999, 48:1890-1895.

68. Vanhoucke CM, Chagnon YC, Rice T, Perusse L, Muzzin P, Giacobini JP, Gagnon J, Wilmore JH, Leon AS, Skinner JS, Rao DC, Bouchard C. Uncoupling protein 3 gene is associated with body composition changes with training in HERITAGE study. J Appl Physiol 2002, 92:1111-1118.

69. Chiu KC, Chuang LM, Yoon C. The vitamin D receptor polymorphism in the translation initiation codon is a risk factor for insulin resistance in glucose tolerant Caucasians. BMC Med Genet 2001, 2,12.

70. Bretherton-Watt D, Given-Wilson R, Mansi JL, Thomas V, Carter N, Cotterill PA. Vitamin D receptor polymorphisms are associated with breast cancer risk in a UK Caucasian population. Br J Cancer 2001, 85:171-175.

71. Bell NH, Morrison NA, Nguyen TV, Elsman J, Hollis BW. Apolipoproteins of the vitamin D receptor predict bone density of the lumbar spine and hip but not of bone density in young men. J Lab Clin Med 2001, 137:133-140.

72. Ferrari SL, Rizzoli R, Slosman DO, Bonjour JP. Do dietary calcium and age explain the controversy surrounding the relationship between bone mineral density and vitamin D receptor gene polymorphism? J Bone Miner Res 1998, 13:1611-1615.

73. Suzuki F, Zehgoud F, Rossignon C, Walrant O, Garabedian M. Association between vitamin D receptor gene polymorphism and sex-dependent growth during the first two years of life. J Clin Endocrinol Metab 1997, 82:2966-2970.

74. Taverna M, Sola A, Goyostegu E, Pacher N, Bruzzo F, Slama G, Reach G, Selam JL. TaqI polymorphism of the vitamin D receptor and receptor of severe diabetic retinopathy. Diabetologia 2002, 45:432-446.

75. Tao C, Yu T, Garnett S, Briody J, Knight J, Woodhead H, Cowell CT. Vitamin D receptor alleles predict growth and bone density in girls. Arch Dis Child 1998, 79:488-493.

76. Fischer PR, Thacher TD, Pettitot JM, Jorde LB, Eccleshall TR, Feldman D. Vitamin D receptor polymorphisms and nutritional rickets in Nigerian children. J Bone Miner Res 2000, 15:2206-2210.

77. Sasaki M, Tanaka Y, Kanetsuya M, Sakuragi N, Dahiya R. Polymorphisms of estrogen receptor α gene in endometrial cancer. Biochem Biophys Res Commun 2002, 297:558-564.

78. Ongphiphadhanakul B, Chanprasertyothin S, Payattikul P, Saetung S, Ongphiphadhanakul N, Payattikul P, Saetung S. Polymorphisms in the estrogen receptor α gene in Thai women. J Clin Endocrinol Metab 2004, 89:173-175.

79. Kikuchi T, Hashimoto N, Kawasaki T, Uchiyama M. Polymorphism of the vitamin D receptor gene polymorphism is associated with skeletal responsiveness to estrogen in postmenopausal women. J Endocrinol Invest 2001, 24:749-755.

80. Koh JM, Kim DJ, Hong JS, Park JY, Lee KU, Kim SY, Kim GS. Estrogen receptor α gene polymorphisms of Prull and Xbal influence association between leptin receptor gene polymorphism (Gil223Arg) and bone mineral density in young men. Eur J Endocrinol 2002, 147:777-783.

81. Han KO, Moon IG, Kang YS, Chung HY, Min HK, Han IK. Nonassociation of estrogen receptor gene genotypes with bone mineral density and estrogen responsiveness to hormone replacement therapy in Korean postmenopausal women. J Clin Endocrinol Metab 1997, 82:991-995.

82. Kikuchi T, Hashimoto N, Kawasaki T, Uchiyama M. Association of serum low-density lipoprotein cholesterol metabolism with oestrogen receptor gene polymorphisms in healthy children. Acta Paediatr 2000, 89:42-45.

83. Iac CI, Liao DL, Bai YM, Lin CC, Yu SC, Chen JY, Wang YC. Association study of the estrogen receptor polymorphisms with tardive dyskinesia in schizophrenia. Neuropsychobiology 2002, 46:173-175.

84. Ongphiphadhanakul B, Rajatanavin R, Chanprasertyothin S, Paseu N, Chailurkit L, Chansiriporn S, Pravalai G, Rajatanavin R. Association of a T262C transition in exon 1 of estrogen-receptorα gene with skeletal responsiveness to estrogen in postmenopausal women. J Endocrinol Invest 2001, 24:749-755.

85. Koh JM, Kim DJ, Hong JS, Park JY, Lee KU, Kim SY, Kim GS. Estrogen receptor α gene polymorphisms of Prull and Xbal influence association between leptin receptor gene polymorphism (Gil223Arg) and bone mineral density in young men. Eur J Endocrinol 2002, 147:777-783.

86. Han KO, Moon IG, Kang YS, Chung HY, Min HK, Han IK. Nonassociation of estrogen receptor gene genotypes with bone mineral density and estrogen responsiveness to hormone replacement therapy in Korean postmenopausal women. J Clin Endocrinol Metab 1997, 82:991-995.

87. Kikuchi T, Hashimoto N, Kawasaki T, Uchiyama M. Association of serum low-density lipoprotein cholesterol metabolism with oestrogen receptor gene polymorphisms in healthy children. Acta Paediatr 2000, 89:42-45.
α gene that is associated with calcium excretion in urine. *J Bone Miner Metab* 2000, 18:153-157.

91. Shibue T, Tsuchiya N, Komata T, Matsushita M, Shiota M, Ohashi J, Waku M, Matsuta K, Tokunaga K: Tumor necrosis factor α 5’-flanking region, tumor necrosis factor receptor II, and HLA-DRB1 polymorphisms in Japanese patients with rheumatoid arthritis. *Arthritis Rheum* 2000, 43:753-757.

92. Hananantachai H, Patarapotikul J, Looareesuwan S, Ohashi J, Naka I, Tokunaga K: Lack of association of -308A/G TNFA promoter and 196R/M TNFR2 polymorphisms with disease severity in Thai adult malaria patients. *Am J Med Genet* 2001, 102:391-392.

93. Ban Y, Taniyama M, Ban Y: Vitamin D receptor gene polymorphism is associated with Graves’ disease in the Japanese population. *J Clin Endocrinol Metab* 2000, 85:4639-4643.

94. Zhang H, Tao G, Wu Q, Liu J, Gao Y, Chen R, Leng X: Vitamin D receptor gene polymorphism in postmenopausal women of the Han and Uygur nationalities in China. *Chin Med J (Engl)* 2000, 113:787-789.

95. Huang CM, Wu MC, Wu JY, Tsai FJ: Association of vitamin D receptor gene BsmI polymorphisms in Chinese patients with systemic lupus erythematosus. *Lupus* 2002, 11:31-34.

96. Lee CK, Hong JS, Cho YS, Yoo B, Kim GS, Moon HB: Lack of relationship between vitamin D receptor: polymorphism and bone erosion in rheumatoid arthritis. *J Korean Med Sci* 2001, 16:188-192.

97. Ongphiphadhanakul B, Rajatanavin R, Chanprasertyothin S, Chailurkit L, Plaseu N, Teerarungsakul K, Sirisriro R, Komindr S, Puavilai G: Vitamin D receptor gene polymorphism is associated with urinary calcium excretion but not with bone mineral density in postmenopausal women. *J Endocrinol Invest* 1997, 20:592-596.

98. Zhao J, Zhou X, Meng X, Liu G, Xing X, Liu H, Xu L: Polymorphisms of vitamin D receptor gene and its association with bone mineral density and osteocalcin in Chinese. *Chin Med J (Engl)* 1997, 110:366-371.

99. Iki M, Saito Y, Dohi Y, Kajita E, Nishino H, Yonemasu K, Kusaka Y: Greater trunk muscle torque reduces postmenopausal bone loss at the spine independently of age, body size, and vitamin D receptor genotype in Japanese women. *Calcif Tissue Int* 2002, 71:300-307.