Gene mapping and identification for osteoporosis

H-W. Deng¹,²,³ and R.R. Recker¹

¹ Osteoporosis Research Center, ² Department of Biomedical Sciences, Creighton University, Omaha, NE, USA
³ Laboratory of Molecular and Statistical Genetics, College of Life Sciences Hunan Normal University, ChangSha, Hunan, P.R. China

Abstract

Osteoporosis is a disease characterized by fragile bones and high susceptibility to low trauma fractures. It is a serious health problem, especially in elderly women. Bone mineral density (BMD) has been employed most commonly as the index for defining and studying osteoporosis. In this presentation, we use examples of our studies in both Caucasians and Chinese to illustrate the approaches used and some main results obtained on 1) characterizing the degree and the inheritance mode of genetic determination of a complex trait such as BMD; 2) identifying and mapping genes for osteoporosis. The purpose of the presentation is to introduce to the medical researchers how and what modern genetics can do to disentangle the mist of an array of genetics factors, the major determinants, for BMD.

Keywords: Bone Mineral Density, Gene Mapping, Gene Identification, Genetics, Osteoporosis, Osteoporotic Fractures

Low trauma fractures are a major public health problem among the elderly¹,², especially among women. Osteoporosis is a disease with fragile bones susceptible to low trauma fractures. Osteoporosis results in more than 1.3 million osteoporotic fractures a year in the US³, and more than 40% of postmenopausal women, on average, will suffer at least one osteoporotic fracture¹,².

One major measurable determinant of fracture risk is bone mass⁴,⁵. The World Health Organization (WHO) defines osteoporosis quantitatively⁶ as "Osteoporosis: A value for bone mineral density or bone mineral content that is more than 2.5 SD below the young adult mean value". Bone mass is a complex trait in that multiple factors and sometimes their interactions are involved and no single major factor can dominate the determination of its variation in general populations.

In this presentation, we will using examples of our studies in both Caucasians and Chinese to illustrate the approaches used on 1) characterizing the degree and the inheritance mode of genetic determination of a complex trait such as BMD; 2) identifying and mapping genes for BMD. We will mainly use our own studies on BMD for demonstration, though other phenotypes such as bone size and ideally osteoporotic fractures are important and have been studied by us⁷,⁸.

Identifying genes for osteoporosis will lead to the discovery of functional mutations responsible for differential susceptibility to osteoporosis and new research into the various functional mechanisms of distinct allele products. These discoveries will help focus and direct basic research on the mechanisms underlying differential risk to osteoporosis, and create short-cuts to the development of new tools, markers and therapies for diagnosing and treating osteoporosis. The loci identified may also serve as starting points for launching studies of the interaction between genotype and environmental factors (GxE interaction) and gene by gene interactions (epistasis) that affect differential risk to osteoporosis. Ultimately, the results may be useful in devising genotype-specific interventions to reduce risk to osteoporosis. Although it is important to realize that our ultimate goal of molecular genetic studies of osteoporosis is to find genes relevant to the risk for osteoporotic fractures⁹, it suffices to use BMD to demonstrate the approaches used in genetics studies of osteoporosis and in gene mapping and identification for osteoporosis.

Extensive reviews of the field for gene mapping and identification for osteoporosis have been made by others¹⁰,¹¹ and us recently¹²,¹³. The principle and the weakness and strength of various approaches used has been elaborated in Recker and Deng¹³. Our goal here is more restrictive and is focusing on our own experience and studies performed in the past few years. The purpose is to illustrate how genetics and gene mapping efforts can be performed in osteoporosis...
Characterizing the degree of genetic determination of BMD and osteoporotic fractures

The degree of genetic determination of a complex trait can be characterized by heritability, which is defined to be the proportion of phenotypic variation that is attributable to genetic components. Heritability may be determined using human twin studies by comparing the degree of resemblance between di- vs. mono-zygotic twins. However, it is not always feasible to collect twins. It is much easier to collect data on parent-offspring pairs or nuclear families (parents plus their children) or pedigrees (generally multi-generation families). These data would easily yield information on quantifying heritability and other important information such as the mode of inheritance and the genetic correlation between complex traits. Genetic correlation reflects the degree of shared genetic determination between two phenotypically correlated traits.

We estimated the magnitude of genetic determination of the variation and covariation of peak bone mass of the spine and hip (adjusted by age, gender and ethnicity) in 47 independent healthy full-sibling pairs and 27 healthy mother-offspring pairs from our Caucasian population. Roughly speaking, twice the phenotypic correlation between full sib pairs or that between mother-offspring pairs is an estimate for heritability. For the spine and hip, the heritabilities ($h^2$) ($\pm$ SE), were 0.76 (0.34) and 0.84 (0.36) respectively when estimated from full sibs, and 0.86 (0.38) and 0.84 (0.39) respectively when estimated from parent-offspring. Some genetic loci underlying peak bone mass variation at the hip and spine are the same or closely linked, as is reflected by the high genetic correlation of 0.95 (0.05) between them when estimated from full sibs, and 0.57 (0.27) when estimated from parent-offspring respectively. Genetic correlation reflects whether and the extent to which two complex traits (such as BMD at different skeletal sites) may share some loci common or closely linked in the genome. From relative pairs, genetic correlation $r_g$ between two complex traits $X$ and $Y$ denote two

![Figure 1. Distribution of spine and hip $Z_{BMD}$'s (Z scores of BMD) in sib 1 and sib 2 in 47 full sib pairs from different families, and distribution of hip and spine $Z_{BMD}$'s of mother and the mean $Z_{BMD}$'s of her children's. $r$ is the phenotypic correlation. (adopted from Deng et al.)](image)
traits, can then be estimated as:

\[ r_g = \frac{\text{cov}(X_1, Y_2) + \text{cov}(X_2, Y_1)}{2\sqrt{\text{cov}(X_1, X_2)\text{cov}(Y_1, Y_2)}} \]

where "1" and "2" denote for the 1st and 2nd relative in relative pairs. \( \text{cov}(X_i, Y_j) \) is the phenotypic covariance of trait \( X \) in individual "1" and trait \( Y \) in individual "2", and \( \text{cov}(X_i, Y_i) \) is the phenotypic covariance of trait \( X \) in individual "2" and trait \( Y \) in individual "1". \( \text{cov}(X_i, X_j) \) is the phenotypic covariance of trait \( X \) in individuals "1" and "2", and \( \text{cov}(Y_i, Y_j) \) is the phenotypic covariance of trait \( Y \) in individuals "1" and "2". Figure 1 illustrates our phenotypic correlation analyses between relatives in inferring heritability of BMD at spine and hip.

With the aid of some elegant special statistical genetic computer programs, large multi-generation pedigrees may also be used to estimate heritability and genetic correlation between complex traits.

In addition to the above relatively simple and straightforward analyses, more elaborate and complex statistical analyses named as segregation analyses may be applied to data collected from nuclear families and/or pedigrees. Segregation analyses can not only quantify heritability, it may also yield inferences about the genetic architecture of a complex trait under study. In segregation analyses, various models (from the ones without any genetic determination to the ones with genetic determination by polygenes each with minor effects, to the ones with a major gene plus background polygenes etc.) are compared for their compatibility (quantified by maximum likelihoods) with the phenotype data in families and pedigrees. The model most parsimonious and most compatible with the data is determined and chosen. We applied segregation analyses to our Caucasian sample with 941 subjects from 51 pedigrees and concluded that there is a major gene that may explain about 16% of phenotypic variation at the hip and spine. The total heritability is about 66%.

We applied the correlation analyses in a sample of 44 healthy full-sister pairs and 186 mother-daughter pairs from our Chinese population. For BMD, the narrow-sense heritabilities \( h^2 \) of spine and hip were 0.72 (0.14) and 0.87 (0.14), respectively when estimated by full-sib pairs, and 0.49 (0.07) and 0.77 (0.07), respectively when estimated by mother-daughter pairs. The genetic correlation of BMD between the spine and hip were 0.97 (0.01) and 0.75 (0.005), respectively when estimated by full-sib pairs and mother-daughter pairs. The results are summarized in Table 1 from Jian and colleagues.

In addition, we performed segregation analyses for BMD and bone size in Chinese sample composed of 401 nuclear families with a total of 1,260 individuals. The results indicate a major gene for the hip BMD, whereas there is no evidence of a major gene for the spine BMD. The heritability estimates for the spine and hip BMD are 0.807 and 0.897, respectively. Some programs used for segregation analyses

<table>
<thead>
<tr>
<th>Subjects pairs</th>
<th>( h^2 ) (SE)</th>
<th>( r_g ) (SE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full-sib pairs</td>
<td>0.72 (0.14)</td>
<td>0.87 (0.14)</td>
</tr>
<tr>
<td>Mother-daughter pairs</td>
<td>0.49 (0.07)</td>
<td>0.77 (0.07)</td>
</tr>
</tbody>
</table>

Table 1. The narrow-sense heritability \( h^2 \) and genetic correlation of BMD at the spine and hip (From Jian et al.)

Figure 2. Different approaches of mapping and identification of genes for complex disorders. Linkage analyses generally require pedigrees with multiple affected and nonaffected individuals or affected relative pairs. TDT analyses typically require at least one affected child from nuclear families with both parents (whose phenotypes may be unknown) and at least one parent needs to be heterozygous at the test marker. Association analyses generally require unrelated affected cases and unaffected controls. (revised and adopted from Recker and Deng)

H-W. Deng and R.R. Recker: Genetics of osteoporosis
can be purchased (such as the SAGE program) or can be downloaded from http://linkage.rockefeller.edu/soft/. This web site contains numerous programs for genetics analyses, for gene mapping and identification in humans.

With substantial evidence pointing to the overwhelming importance of genetic factors in BMD determination and with increasing health problem due to osteoporosis, not only in Caucasians but also in Chinese, it is important and imperative to embark on endeavors of molecular genetics studies to map and identify genes for BMD.

**Gene mapping and identification for osteoporosis in humans**

Three approaches have been employed in humans: population association studies, linkage studies, and transmission disequilibrium tests (TDT). Population association studies test whether particular alleles or genotypes are associated with a higher risk or a larger trait value, usually in unrelated population samples. An association usually reflects statistical non-independence (linkage disequilibrium) of a marker allele(s) and a functional mutation(s) underlying the risk or variation of a trait and does not necessarily imply causality. Linkage tests whether there is co-segregation or co-inheritance of alleles with a phenotype under study in pedigrees or affected relative pairs. Linkage refers to close physical locations of genes on one chromosome. TDT tests, in families or discordant relative pairs, both linkage and association of marker alleles with a functional mutation underlying the study trait and is significant if, and only if, both linkage and association exists. Figure 2 revised and adopted from Recker and Deng13 illustrates the principles of the three approaches intuitively.

Population association studies in which candidate genes have been identified a priori based on known biologic functions of gene products are generally statistically much more powerful and the samples are much easier to recruit than linkage studies30-31. However, the association approach

![ANOVA p=0.201](image)

**Figure 3.** Different outcomes of significant associations of VDR BsmI genotypes with spine BMD. Plotted are the mean and SE bar for different VDR genotypes. P-values of the one-way ANOVA for testing the VDR genotype effects on spine BMD are given, so are the sample sizes (n) for each genotype. Spine BMD in plot a is the raw data. Spine BMD values in plot b are those adjusted for age, height, weight, ERX and ERP polymorphisms (adopted from Deng et al.3)
depends on linkage disequilibrium (statistical association) of markers employed with functional mutations in/near the candidate genes. It is prone to population admixture/stratification in yielding not only false positive results32-35, but also false negative results35. Population stratification/admixture is very difficult to detect, even with large samples and its plaguing effects on association studies are very significant34. Importantly, the association approach tests only association, which may not be relevant to causation, so the association approach per se alone has an inherent difficulty in unambiguously identifying causative genes underlying complex traits. In some studies, such as pharmacogenomic studies, association design may be the only convenient choice36. Linkage studies test only linkage and can search for any genomic region (without any prior knowledge) contributing relatively large variation in complex traits. The linkage approach does not depend on the existence of linkage disequilibrium among genes or markers in adjacent genomic locations and is robust to population admixture/stratification. However, even if significant linkage results are found, extensive fine mapping efforts37,38, which generally depends on linkage disequilibrium, are needed in order to pinpoint a QTL to a small genomic region (~1cM). Only after being fine mapped to a relatively small genomic region, may it be feasible for physical mapping to identify a specific QTL. Traditional linkage results generally reveal large genomic regions (~20cM) that are not feasible to identify causative QTLs. Using the linkage approach to search for genes underlying complex traits generally requires very large samples to be screened and/or genotyped39-41, unless the genetic effect of the locus is large. Linkage studies can be applied to the candidate genes42, or a specific genomic region of interest43, or the whole genome scans44.

The TDT45 and its variants11,40,44,46,47 can be employed to test specific candidate genes for both linkage and association, and is robust to population admixture/stratification48,49. Depending on the distribution of linkage disequilibrium in human genomes and the development of highly automatic SNP (single nucleotide polymorphisms) genotyping, the TDT (as well as association approach) can also be employed for whole-genome scan studies39. The TDT requires existence of linkage disequilibrium of adjacent genes (or markers with functional mutations) and a much denser distribution of genetic markers as compared to traditional linkage studies39 in genomic linkage studies. Since the TDT depends on population linkage disequilibrium and tests for both linkage and association, significant results (if any) can robustly pinpoint a small region containing causative genes with functional mutations50. Therefore, the TDT approach is particularly ideal for testing candidate genes (with functional relevance) or candidate genomic regions that have shown evidence for association or linkage. Judiciously designed sampling schemes in practice can substantially increase the power of linkage and TDT studies51,52.

When testing specific candidate genes, regardless of the approach employed (association, linkage, or TDT), the knowledge obtained from other fields of bone biology regarding the pathophysiology of osteoporosis is utilized. Because genes are only investigated where there is a priori evidence of their potential involvement in the trait concerned, the chance of false positive findings using the candidate gene approach may be minimized. On the other hand, whole-genome scans can identify genomic regions that are important but would otherwise be unknown.

### Association studies of candidate genes

There are a total of 32 candidate genes have been associated with BMD variation31. The list is likely to grow. We selectively tested a few candidate genes and we are still testing more candidate genes in our samples of Caucasians and Chinese. Earlier, we mainly tested a limited few markers for each candidate gene studied. The work is now expanding to a panel of SNP (single nucleotide polymorphisms) about 2kb away from each other and expanding in and around each candidate gene for exhaustive testing of the chosen candidate genes.

In our Caucasian sample, we53 assessed associations of BMD with VDR (vitamin D receptor) BsmI genotypes, and ER-α (estrogen receptor-α) XbaI and PvuII polymorphisms.
(denoted as ERX and ERP respectively) with spine, femoral neck, distal radius BMD, and with total body bone mineral content (tbBMC) in 108 US Mid-western postmenopausal women. We statistically controlled for confounding factors such as height, weight etc. in the analysis. We tested the difference of the mean phenotypic values of different genotypes. No significant association was detected for ER-α genotypes with spine and radius BMD, or for VDR genotypes with femoral neck and radius BMD and tbBMC. No significant interaction between VDR and ER-α genotypes was detected.

Figure 4. Multipoint linkage analysis results for the chromosomes that had a maximum LOD score close to or greater than 2.0. The solid line is for spine BMD, the dashed line is for hip BMD, and the dot dashed line is for wrist BMD (adopted from Deng et al. 44)
in our sample. However, the VDR genotypes are significantly (p=0.004) associated with ~5.8% spine BMD variation. Both ERX and ERP genotypes are significantly (p=0.02) associated with ~3.5% femoral neck BMD variation. ERX genotypes are significantly (p=0.03) associated with ~2.4% tbBMC variation. However, if the data were analyzed by simple ANOVA as in some previous studies, without adjusting statistically for confounding factors, all the significant results we found here would have gone undetected. Our findings suggest that: 1) VDR and ER-α genotypes may have different effects on BMD at different sites and on tbBMC; and 2) If significant factors influencing bone are not appropriately controlled, true significant associations can be easily missed. Figure 3 adopted from Deng et al.42 illustrates our results for the analyses of spine BMD and the VDR genotype.

In our Chinese sample, we44 evaluated the relationship between estrogen receptor α (ER-α) and vitamin D receptor (VDR) gene with BMD in 649 pre- and postmenopausal healthy Chinese women. The polymorphisms of the ER-α PvuII, XbaI and VDR ApaI were detected by PCR-RFLP. The VDR ApaI is significantly or nearly significantly associated with the spine BMD in both of the pre- and postmenopausal groups (p=0.023 and 0.079, respectively), with AA subjects showing a lower BMD than others. Either the individual polymorphism of ER-α PvuII or XbaI was not associated with BMD in either group. However, we detected significant association between ER-α haplotype with BMD variation in the premenopausal group: px was significantly correlated to lower BMD at the spine (p=0.035); pX was nearly significantly associated with higher BMD at the spine (p=0.066) and the hip (p=0.082). In postmenopausal women, significant interaction was found between ER-α haplotype and VDR gene. With AA genotype (or A allele) at the VDR ApaI locus, subjects with pX had a 8.3% higher BMD at the spine than those without it (p=0.022), and PX carriers had a 11.3% lower BMD at the hip than those without it (p=0.041). Hence, in our study sample of Chinese women, VDR- ApaI is associated with BMD variation at the spine. The haplotype of ER-α gene may serve as a good biomarker for BMD prediction by itself in premenopausal women, or by interaction with VDR gene in postmenopausal women. In addition, we55 found that, VDR-ApaI has a weak association with BMD variation at the spine and the ER-α haplotypes may be associated with BMD variation at some skeletal sites in Chinese males.

Transmission disequilibrium tests (TDT) analysis

The TDT has not been applied often in bone genetic studies, even when compared with the expensive whole genome linkage scans. In the bone field, association studies of candidate genes have been reported but the results are largely inconsistent, and the few available linkage studies have failed to establish linkage of these candidate genes with BMD variation. In addition, whole genome scans in humans and QTL mapping in mice have started to suggest genomic regions and new candidate genes for confirmation and replication. Therefore, we have an ideal situation in which to apply the TDT to genetic studies of bone mass. It will be built upon our extensive work in population association, linkage and QTL mapping studies, to identify and test genes underlying osteoporosis using various approaches.

Deng et al.42 simultaneously tested linkage and/or association of the genes for VDR, osteocalcin (also known as bone Gla protein, BGP), and parathyroid hormone (PTH) with BMD in 630 subjects from 53 human pedigrees. Each of these pedigrees was ascertained through a proband with an extreme BMD value at the hip or spine (Z-score < -1.28). This sample contained more than 11,000 informative relative pairs for linkage tests, including 1,249 sib pairs. The investigators performed tests for linkage alone, association alone, then for both linkage and association by the TDT. We found evidence for association and/or linkage for spine BMD at the VDR gene. Significant results were also found for association and/or linkage for the BGP gene with hip BMD. Our data support the VDR gene as a QTL underlying spine BMD variation and the BGP gene as a QTL underlying hip BMD variation. However, our data do not support the PTH gene as a QTL underlying hip or spine BMD variation. Table 2 summarizes our major results from Deng et al.42. The program QTDT used for analyses can be downloaded from http://www.sph.umich.edu/csg/abecasis/QTDT/.

In our Chinese sample of 1,260 subjects from 401 nuclear families, we66 simultaneously test linkage and/or association of the estrogen receptor α gene polymorphism with peak bone mass in 401 Chinese nuclear families (both parents plus their female children) with 1,260 subjects, with the 458 children aged between 20-40. All the subjects were genotyped by PCR-RFLP at polymorphic PvuII and XbaI sites inside the ER-α gene. BMD were measured at lumbar spine (L1-L4) and hip (femoral neck, trochanter, intertrochanteric region). Raw BMD values were adjusted by age, height, and weight as covariates. We detected marginally significant results for within-family association (transmission disequilibrium) (p = 0.054) between the spine BMD variation and the ER-α XbaI genotypes. For the hip BMD variation, significant (p < 0.05) linkage results were generally found for the two intragenic markers. Analyses of the haplotypes defined by the two markers confers further evidence for linkage of the ER-α with the hip peak bone mass variation. In conclusion, this study suggests that the ER-α gene may have minor effects on peak bone mass variation in our Chinese population.

Linkage studies

Linkage studies, particularly those for the whole genome, are rare in the field of bone genetics, largely due to the high cost involved. We57 genotyped 400 markers throughout the whole human genome for 635 people from 50 large Caucasian pedigrees identified via probands with extremely low BMD values. This sample contains more than 11,000 informative relative pairs for linkage tests, including 1,249
sib pairs. Figure 4 from Deng and colleagues presents the results for the chromosomes which are identified as important. The magnitude of the logarithm of odds (LOD) scores in the plots convey the information regarding the magnitude of significance of the genomic regions identified. The genomic regions identified are broad (~20cM) and contain some important candidate genes, for example, the IL-6 on 7p22, the IGF 1 (insulin growth factor) on 12q24. Fine mapping effort is needed to narrow down the regions to manageable size (~1cM). Before that, the genomic regions needed to be confirmed using larger samples with sufficient statistical power to guard against false results. More than 4,000 subjects have been recruited from large pedigrees (one pedigree contains more than 500 subjects) as well as nuclear families that should form a solid basis for us to eventually identify the genes involved for BMD variation.

Conclusion

With the extensive molecular genetic research initiated recently in the bone field, we are making steady progress toward mapping and identification of genes for osteoporosis. This is true as witnessed not only by the interesting, albeit somewhat debatable, results reviewed earlier and here, but also by the problems and challenges surfacing during our progress. By continuously confronting these problems and challenges utilizing the multiple approaches, as well as new approaches such as DNA microarrays for gene expression analyses and proteomics for large-scale comparative assays of functions and quantities of proteins that may complement and confirm each other, we are on our way to unraveling the tangle of the genetic variation in osteoporosis.

Acknowledgements

This project was partially supported by grants from the Health Future Foundation of the USA, from the National Institute of Health (K01 AR02170-01, R01 GM60402-01A1), from the State of Nebraska Cancer and Smoking Related Disease Research Program, the State of Nebraska Tobacco Settlement Fund, and the US department of Energy (DE-FG03-00ER63080/A00). The project has also been supported by the Hunan Province Special Professor Start-up Fund (25000612), Outstanding Young Scientist Award (30025025), a general grant (30170504), and a key project grant (30230210) from the Chinese National Science Foundation (CNSF), a seed grant (25000500) and a key project grant from the Ministry of Education of P. R. China, and a young scientist development grant from the Huo Ying Dong Education Foundation of Hong Kong.

References

35. Deng HW. Population admixture may appear to mask, change or reverse genetic effects of genes underlying complex traits. Genetics 2001; 159:1319-1323.


