Commentary: Vitamin D receptor polymorphism and bone mineral density: effect size in Caucasians means detection is uncertain in small studies

Nigel Morrison

Vitamin D receptor (VDR) is a nuclear hormone receptor that acts as a transcriptional regulator in response to circulating 1,25 dihydroxyvitamin D$_3$, the active hormonal form of vitamin D. VDR gene polymorphism (VDRGP) have been extensively studied in different diseases, with over 700 primary research articles, although this has focused mainly on the same markers. The VDRGP experience, with its huge literature and appearance of apparently contradictory reports each month, may provide an example of what to expect with other genes in the growing field of analysis of common gene polymorphisms with complex common disorders. Morita et al. provide a typical example of a moderately sized population study of the relationship of VDRGP to bone density and rate of bone loss in Japanese.1 Reviewing the VDRGP literature is beyond the scope of this commentary which will only refer to a limited number of publications. For those interested, Zmuda et al. provide two comprehensive reviews of the literature of VDR related to disease.2,3 Suffice to say that VDRGP have shown positive association to a wide range of divergent diseases, and due to the pleiotropic mode of action of a nuclear-hormone receptor such as the VDR, plausible molecular scenarios of involvement can be constructed for many different diseases. In fact, if functional genetic polymorphism occurs in a transcriptional regulator, one should expect pleiotropism, due to the fact that VDR controls the expression of a large and unknown number of subordinate genes, in both positive and negative senses and in cell-specific manners. The VDR protein is at the centre of the vitamin D endocrine system, a complex physiological system with substantial feedback regulatory mechanisms involved in maintaining serum calcium and 1,25 dihydroxyvitamin D$_3$ within narrow bounds and now known to affect a large number of organs.4 It is possible that the self-regulatory nature of the VDR endocrine system moderates the effect of VDRGP. VDR gene polymorphisms are looking for phenotypes, and judging from the literature, are related to numerous different traits, reflecting pleiotropism. Therefore, although literature has accumulated concerning VDR and bone mineral density (BMD) in particular, this may not necessarily be the most potent effect of genetic variation in VDR.

VDR polymorphisms

The large amount of positive genetic association data in a number of diseases suggests functional consequences of VDR gene polymorphism. The most obvious candidate for a functional change is at the initiation codon where polymorphism truncates the first three amino acids of the amino terminus of the VDR from MEAMAAST(etc) to MAAST(etc). The initiation codon site can be detected using Fok1 restriction enzyme, with the presence of the site designated ‘f’ or M1 and the absence of the site ‘F’ or M4, referring to the alternative initiator methionine. Convincing molecular data shows that the truncation of the VDR to the M4 form results in increased transactivation of target genes by VDR.5 The other VDRGP commonly studied are detected with Bsm1, Apa1 (in the last intron), and Taq1 (in the last exon) within 1090bp of each other, in the order B,A,T. The Taq1 site results in a synonymous isoleucine codon that does not change the amino acid sequence. Since the Taq1 site exists in mRNA, attempts have been made to determine if the mRNA levels are different in heterozygotes, with data suggesting a 35% increase in the t allele over the T allele.6,7 This is paradoxical as the t allele is commonly related to lower BMD but was associated with higher levels of VDR mRNA. Of course, these studies were done using leukocytes and may not reflect events occurring in bone. The effect of the Taq1 marker on relative mRNA levels, and indeed on BMD, may be due to an almost complete disequilibrium between Taq1 and so called long and short alleles of a polyA tract in the 3'UTR of the VDR gene. When the different allelic forms of the 3'UTR were cloned downstream of a reporter gene, differential activity was observed.8 While the B, A and T alleles have a high degree of association, no detectable relationship exists between these and the initiation codon polymorphism, due to the gap of about 34kb. A final polymorphism of functional significance is within a Cdx-2 binding element in the VDR promoter that possibly alters the transcriptional level of the VDR gene.9 While the initiation codon polymorphism is the most convincing candidate for a functional change, the other markers mentioned above may well be in linkage disequilibrium with un-typed polymorphism that alters VDR function in unconventional ways. It is also possible that several functional changes of different magnitude and mechanism may exist within the VDR, each at varying rates of
of linkage disequilibrium in different populations, with the commonly typed VDRGP.

**Multifactorial traits and BMD**

Complex multifactorial diseases are thought to arise as a result of the small incremental effects of many gene loci coupled with environmental risk factors. For any particular age group, BMD has a smooth continuous unimodal normal population distribution. However, there is a very strong age relationship, leading to a decline with age beyond the menopause. Not surprisingly, BMD is a predictor of bone fracture and BMD below a certain age-adjusted numerical value is often used to define osteoporosis. In relationship to BMD, VDR is one of the most studied genetic markers, yet consensus is lacking. Deng et al. suggest failure to detect association could be due to genetic heterogeneity (different alleles contributing to different effects in different populations), study design, and power issues, or other factors such as inadequate controlling of confounding environmental exposures.

BMD is strongly related to anthropometric variables such as age, height, and weight. This fact complicates simple analysis of genetic markers in relation to BMD, especially if the effect of a genetic locus varies with age. It becomes difficult to model all the possible regression equations including anthropometric variables with genotype as additive, dominant, or recessive, without a degree of arbitrariness in the analysis. Authors tend to use linear or higher than linear quadratic adjustment for the age-related effect on BMD. Of particular relevance to BMD changes in Asians, Liao et al. analysed 2702 Chinese aged from 5 to 96 years and found that cubic relationships between age and BMD fit best for all skeletal sites tested. The cubic relationship implies that BMD is continually changing with age; a reasonable proposition. Genotype effects may be modelled in ANCOVA with a cubic age relationship, but the possibility that adolescent rates of bone gain and rates of bone loss after the menopause may be affected themselves by VDR genotype severely complicates the analysis of the age effect with genotype. As cross-sectional population-based studies become larger, a more-attractive analytical alternative might be to use a regression spline approach with age knot selection based on important life events, such as skeletal maturity and the onset of menopause. Similar issues of non-linearity exist with the height and weight variables. To complicate matters further, BMD itself is not a true density, but an areal projection with units of grams per square centimetre and is influenced by bone size. More advanced measures of bone mineral content and micro-architecture may provide better target phenotypes for genetic studies.

**The standard quantitative model**

In the quantitative model, the effect of a locus is given by the additive and dominance effects of alleles. The phenotypic variance comprises environmental sources, genetic and gene–environment interaction. The proportion of the variance attributable to the locus is given by the equations below where \( \alpha \) is half the difference between the alternative homozygotes and \( \delta \) is the dominance deviation expressed as proportions of standard deviations:

\[
V_A = 2pq\alpha^2 \quad V_D = 2pq\delta^2
\]

\[
V_G = V_A + V_D = 2pq(\alpha^2 + \delta^2)
\]

\( V_A \) and \( V_D \) are the additive and dominance contributions to the genetic variance at the locus. As BMD is a smooth unimodal normally distributed trait it seems unlikely that high abundance alleles of high effect size exist within populations examined to date: common large effect alleles would result in a platykurtic or bimodal distribution. Realistically, we can only expect a small contribution to the overall phenotypic variance to be attributable to any particular individual gene locus in a large unselected population. Furthermore, any particular gene that influences a trait such as BMD may have multiple different alleles or haplotypes that affect the trait, each with a different genetic effect size (\( \alpha \)) and a different population frequency.

We should expect the distribution of effect sizes at a locus to follow an exponentially decreasing model, where large effect alleles are rare.

**What is the effect size of VDRGP?**

A recent meta-analysis of VDRGP effects on BMD in Caucasians concludes that the \( B \) allele is related to lower BMD at the spine in postmenopausal females and that it acts in relation to BMD such that \( BB = Bb < bb \). In order to estimate the effect size (\( \alpha \)) and the dominance deviation (\( \delta \)) at the VDR locus, I used the figures provided for postmenopausal females by the studies used in meta-analysis. One study (Keil et al.) had a rather extreme violation of Hardy Weinberg equilibrium for the Bsm1 marker; that study was deleted and 14 others accepted. I reconstructed the ANOVA table (Table 1) that would result if the 14 studies had been conducted simultaneously, modelling just genotype effects, including the effects of the studies themselves, and then modelling a genotype–study interaction. In each analysis spine BMD was significantly related to genotype. Incorporating the variation attributable to the study and the interaction only slightly improved the discrimination of the genotype effect although it reduced the estimate of the standard deviation (from 0.17 to 0.14) thereby slightly increasing the genetic effect as a percent of the residual standard deviation. The significant interaction (Table 1) indicates simply that the magnitude of the gene effect varied across different studies: an unsurprising conclusion. Removing the interaction term had little effect on the main effect of genotype (the F ratio changed from 5.86 to 4.65) while the variation attributable to the difference in studies was important, reducing the F ratio to 3.9 (P-value change from 0.003 to 0.02).

The genotype \( bb \) had the highest BMD. Genotypes \( BB \) and \( Bb \) had equivalently low BMD (illustrated in Figure 1). The small difference in mean BMD values between \( BB \) and \( Bb \) was not significant, suggesting that dominance is nearly complete. Since the data were consistent with a dominant effect of the \( B \) allele in conferring low BMD (equivalent to a recessive effect of the \( b \) allele in conferring high BMD), alleles were pooled. In this analysis genotype was a significant predictor of spinal BMD (\( P = 0.005 \)) and the difference in SD was 0.12 (95% CI: 0.04, 0.20), corrected for the effect of study variation and 0.14 (95% CI: 0.06, 0.21) taking into account the interaction.
The differences between the means of the genotypes provide data to solve for the $\alpha$ and $\delta$ components of the quantitative model. The deviations of each genotype derived from a dimorphism are functions of $\alpha$ and $\delta$ and the allele frequencies, $p$ and $q$. Expressed as deviations from the population mean, the genotype means for the VDR Bsm1 genotypes become: $2q(\alpha - p\delta)$ for genotype $bb$ (frequency $p^2$), $\alpha(q-p)+\delta(1-2pq)$ for genotype $Bb$ (frequency $2pq$), and $-2p(\alpha+q\delta)$ for genotype $BB$ (frequency $q^2$). Solving for these values produces standardized coefficients of $\alpha = 0.0585$ SD and $\delta = -0.0541$ SD. The values provide a good fit to the data derived from the combined meta-analysis (Table 2).

Having obtained a reasonable population-based estimate of the additive and dominant quantitative parameters and an idea of the effect size, a simple power exercise can be completed. As a result of the above analysis, it seems reasonable to assume an effect size of between 0.11 and 0.13 SD for the VDR Bsm1 $B$ allele, acting in a dominant manner to confer lower BMD in postmenopausal females. Since $BB$ and $Bb$ genotypes are effectively equivalent, they can be pooled.

The power of a study can be estimated by:

$$Z_{(1-\beta)} = \frac{\Delta \mu}{\delta} \sqrt{N_t}p^2(1-p^2) + Z_\alpha$$

where $Z_\alpha$ is the normal curve cut-off for the desired type one error rate, $p$ and $q$ are the allele frequencies (for $b$ and $B$ respectively), $\Delta \mu$ is the prospective numbers; for 80% power 2680 and 4000

Table 1 Reconstructed ANOVA table modelling the effect of genotype, study, and interaction on spine bone mineral density (BMD) derived from 14 studies on postmenopausal females

<table>
<thead>
<tr>
<th>Genotype</th>
<th>N</th>
<th>Mean BMD</th>
<th>Pooled SD</th>
<th>Difference from mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>$BB$</td>
<td>490</td>
<td>0.898</td>
<td>0.197</td>
<td>0.007</td>
</tr>
<tr>
<td>$Bb$</td>
<td>1412</td>
<td>0.899</td>
<td>0.171</td>
<td>0.006</td>
</tr>
<tr>
<td>$bb$</td>
<td>961</td>
<td>0.918</td>
<td>0.154</td>
<td>0.013</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>2863</td>
<td>0.905</td>
<td>0.171</td>
<td></td>
</tr>
</tbody>
</table>

Table 2 Deviations of genotype means from the standardized population mean (zero) in the quantitative model with dominance deviation

<table>
<thead>
<tr>
<th>Genotypes</th>
<th>$p^2$</th>
<th>$2pq$</th>
<th>$q^2$</th>
<th>$\alpha$</th>
<th>$\delta$</th>
<th>$\alpha$ - $\delta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A1A1 (bb)$</td>
<td></td>
<td></td>
<td></td>
<td>0.0585</td>
<td>0.037</td>
<td>0.021</td>
</tr>
<tr>
<td>$A1A2 (Bb)$</td>
<td></td>
<td></td>
<td></td>
<td>0.042</td>
<td>0.037</td>
<td>0.005</td>
</tr>
<tr>
<td>$A2A2 (BB)$</td>
<td></td>
<td></td>
<td></td>
<td>0.042</td>
<td>0.037</td>
<td>0.005</td>
</tr>
</tbody>
</table>

Estimate of $\alpha = 0.0585$ and estimate of $\delta = -0.0541$, based on data in Table 1. Frequencies are based on Caucasian values; with $b = 0.582$ and $B = 0.418$.

**Post hoc comparisons**

<table>
<thead>
<tr>
<th>Comparisons</th>
<th>Difference</th>
<th>SD difference</th>
<th>$P$-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$BB-bb$</td>
<td>0.020</td>
<td>0.14</td>
<td>0.01</td>
</tr>
<tr>
<td>$Bb-bb$</td>
<td>0.019</td>
<td>0.13</td>
<td>0.001</td>
</tr>
<tr>
<td>$BB-Bb$</td>
<td>0.001</td>
<td>0.01</td>
<td>0.87</td>
</tr>
</tbody>
</table>

* Based on a proportion of the residual SD.

The power of a study can be estimated by:

$$Z_{(1-\beta)} = \frac{\Delta \mu}{\delta} \sqrt{N_t}p^2(1-p^2) + Z_\alpha$$

where $Z_\alpha$ is the normal curve cut-off for the desired type one error rate, $p$ and $q$ are the allele frequencies (for $b$ and $B$ respectively), $\Delta \mu$ is the prospective numbers; for 80% power 2680 and 4000

Table 3 shows that large numbers are required to have reasonable power to detect an effect of the magnitude suggested by meta analysis (0.13SD difference between the two groups). Japanese allele frequencies were estimated as $b = 0.885$ and $B = 0.115$. For 80% power to detect the effect at $P = 0.01$, 3046 Caucasians or 4700 Japanese are required. Basing the power calculation expressly on the estimates of $\alpha$ and $\delta$ shown in Table 2, the difference is slightly smaller (0.11 SD) resulting in larger prospective numbers; for 80% power 2680 and 4000 Caucasians (or 3600 and 5400 Japanese) are needed at $P$ of 0.05 and 0.01 respectively. These are sobering numbers, considering
Haplotypes in different ethnic groups

In Caucasians the Bsm1 and Taq1 markers are essentially equivalent with strong linkage disequilibrium between these markers. A major difference in haplotype distributions exists between Caucasians and Chinese, Japanese and Koreans.33 Haplotype frequencies are in the following order of abundance: Caucasians, $baT^T$, $BaT$, then $Bat$; Asians, $bat^T$, $bat^T$, $Bat$ and then $BaT^T$. The second frequent haplotype found in Caucasians ($BaT$ frequency = 0.41) that has been associated with low BMD, is rare in the Asian samples (frequency = 0.06). By typing only Apa1 and Taq1 VDR alleles in Japanese samples, it is not possible to discriminate the most abundant haplotype $bat^T$ (frequency = 0.56) from the third abundant haplotype $Bat$ (frequency = 0.15). Determining the contribution of individual haplotypes to a quantitative trait in a cross-sectional analysis is clearly a difficult task. While the dominance deviation effect described above in Caucasians is attributable largely to haplotype $BaT^T$, the effect of haplotype $bat^T$ is not clear.

In the present study, Morita et al.1 report effect sizes between 0.09 and 0.12SD, with genotype $aaTT$ lower than that of $AaTT$, with the effect at the mid radius but no appreciable effect at the other sites of the spine and the femoral neck. This compares homozgyous haplotype $at^T$ with individuals heterozygous for haplotypes $at^T$ and $AT$. Since Bsm1 was not genotyped, it is not possible to discriminate between haplotypes $bat^T$ and $BaT$ in the $aaTT$ homozygotes, meaning that the $aaTT$ genotypes are admixed.

Conclusions

If an effect cannot be detected in a study of reasonable size, is it important? Morita et al.1 suggest that VDRGP have a negligible effect on BMD and rate of change of BMD in Japanese. That may be so, but it would seem too early to judge given the general lack of exploration for other polymorphism in VDR. The fact that VDRGP have been associated with an increasing array of conditions and diseases unrelated to bone is consistent with pleiotropism, which should be expected for a transcriptional regulator. The power analysis above suggests that it would be uncertain to detect an effect of the same magnitude as observed in Caucasians, with a cross-sectional analysis of Japanese, even one the size of that done by Morita et al.1 with 778 pre- and 604 post-menopausal women. Despite that, Kim et al.34 report that in Koreans, genotyping the $baT$ haplotype system, in conjunction with VDRGP long and short alleles, resulted in a significant association with BMD in a study of 417 postmenopausal females. Kim et al.34 found improved discrimination of the genetic effect by typing all markers. Morita et al.1 and Kim et al.34 represent the VDRGP story in microcosm: two similar studies in similar ethnic groups, one yielding significant relationships and the other not. If the effect size is smaller than originally imagined, then many studies that are currently feasible will fail to detect a significant effect and this will only be resolved by meta-analysis.
Genetic risk factors in cross-sectional association studies are not different in character to any other type of predictor that may be detected sporadically in an epidemiology study due to effect size. False positive signals do not tend to be replicated, especially with the regularity of the VDRGP. It is currently easier to collect quality clinical data, as in Morita et al. than to obtain dense genetic information. Alternatively, different study designs, based on large multigenerational pedigrees may provide more power to detect VDR linkage and association simultaneously.

In part as a response to the concerns regarding power in genetic analysis of complex traits, very large genetic epidemiology studies are contemplated with numbers in the hundreds of thousands. The success of such whole genome association studies will depend on the quality of clinical data, genetic data, and selection of participants. With companies like Perlegen developing million SNP analyses per patient, the dream of whole genome sequence for each participant of a mass epidemiology study is not distant fantasy. One suspects that even these large whole population studies will depend on meta-analysis.

References


