BoneKEy-Osteovision. 2005 February;2(2):14-20
http://www.bonekey-ibms.org/cgi/content/full/ibmske;2/2/14
DOI: 10.1138/20050150

COMMENTARIES

Bigger is Better, But It's Not Just Size That Counts: The Estrogen Receptor Gene and Osteoporosis

Carles Vilarino-Guell and Matthew A. Brown

Institute of Musculoskeletal Sciences, University of Oxford, Oxford, United Kingdom

February 2005

Commentary on: Ioannidis JP, Ralston SH, Bennett ST, Brandi ML, Grinberg D, Karassa FB, Langdahl B, van Meurs JB, Mosekilde L, Scollen S, Albagha OM, Bustamante M, Carey AH, Dunning AM, Enjuanes A, van Leeuwen JP, Mavilia C, Masi L, McGuigan FE, Nogues X, Pols HA, Reid DM, Schuit SC, Sherlock RE, Uitterlinden AG; GENOMOS Study. Differential genetic effects of ESR1 gene polymorphisms on osteoporosis outcomes. *JAMA.* 2004 Nov 3;292(17):2105-14.

Substantial evidence of an important role for estrogen receptor α (*ESR1*) in bone development and osteoporosis risk has been obtained from studies of cases with mutations of the receptor and cell biology experiments. However, whether the receptor plays a role in the heritability of osteoporosis-related phenotypes in the general population has been uncertain, with numerous studies published, mostly small, and none with convincing evidence one way or the other.

With this background in mind, the study by Ioannidis *et al.* (1) of *ESR1* polymorphisms, in the largest case collection yet reported for any osteoporosis study, is a major advance. The paper reports a study of the genetic effect of three previously identified polymorphisms in *ESR1* on osteoporosis. The three polymorphisms are a poly-TA microsatellite (dbSNP rs3138774) in the promoter region and two single nucleotide polymorphisms (SNPs), *Pvu*II (dbSNP rs2234693) and *Xba*I (dbSNP rs9340799), in intron 1 of *ESR1*. The conclusion from this study is a lack of association between any of these markers with BMD at any site, but an association of *Xba*I with all fractures (p = 0.002). This association was particularly significant with regard to vertebral fracture, where there was a 35% reduction in fracture risk in women homozygous for the minor allele (*i.e.*, the XX genotype), with no association detected in men.

This meta-analysis was performed in a cohort of 18,917 subjects from eight European centers. The cohort consisted predominantly of women, with males representing only 23% of the total cohort, still in osteoporosis study terms, a very large sample size. A recent osteoporosis genetics review identified 27 *ESR1* association studies (2). The sample size in these studies ranged from 30-900 cases, with the median study having just 124 cases, highlighting what a substantial jump the current study (1) represents. Only three studies genotyped any marker other than the three investigated by Ioannidis *et al.* (1). A meta-analysis of published data on 5834 individuals genotyped for *Xba*I and *Pvu*II polymorphisms was published in 2002 by Ioannidis *et al.* (3), demonstrating increased hip, lumbar spine, and whole body BMD in homozygotes for the minor allele of the *Xba*I SNP (*i.e.,* the XX genotype) and a reduction in fracture risk.

This paper therefore provides strong confirmation of previous data implicating *ESR1* in fracture risk, but a question remains as to whether the polymorphisms studied influence BMD. How do we interpret this? Possible explanations are that the initial meta-analysis finding may have been

wrong or that the current study has failed to detect a true effect, despite being substantially larger than the initial meta-analysis. At this stage, as discussed below, one cannot determine which of these possibilities is correct.

Certainly, retrospective meta-analyses of candidate gene association studies are prone to incorrect conclusions because of publication/reporting bias and should be interpreted with caution. Many investigators do not publish studies when no association is seen, or only publish in abstract format. Meta-analyses also rarely take into account differences in case ascertainment that may be relevant to the genetic findings. Although methods exist to investigate the presence of publication bias, such methods are relatively insensitive. This may not be the case for osteoporosis linkage studies yet to date no meta-analysis has been reported for these.

So what is the power of the current study, and does it represent a model for future osteoporosis genetic studies?

The study power depends heavily on the proportion of the genetic diversity in *ESR1* that is captured by the genotyping strategy. The markers studied by Ioannidis *et al.* (1) were probably selected because they are the variants most commonly genotyped in the past and certainly not because of any genotypic characteristic, such as the proportion of *ESR1* haplotypes they mark or because they themselves are likely to be functional. The first paper studying *ESR1* polymorphisms in osteoporosis gives no rational for the selection of these markers, and subsequent investigators seem largely to have simply followed this lead (4). Indeed, *Pvu*II and *Xba*I polymorphisms are located in intron 1, far from the intron-exon boundary, and are thus very unlikely to be functional. Furthermore, they lie only 47 bp from one another and are in strong linkage disequilibrium (LD). Thus, the two SNPs are unlikely to mark different haplotypes, and not much is gained by genotyping both markers, rather than just one. To date, there is no compelling evidence that dinucleotide variants have functional effects, and no data suggest that the microsatellite studied here influences *ESR1* function or expression. The decision to pool microsatellite alleles into those with “low” or “high” repeat numbers (the cutoff repeat length is not stated in the paper) is not based on functional data, nor is it likely that it reflects the evolutionary history of microsatellites.

To assess the proportion of genetic diversity identified in this study, we accessed the International HapMap Project (IHMP) database (see www.hapmap.org), a freely available public resource currently containing genotype data on 90 white individuals on 956,730 SNPs. The *ESR1* gene has been extensively studied by the IHMP, and genotype data is currently available (14th release) on 70 polymorphic SNPs, including *Pvu*II. LD analysis using SNPSpD (see http://genepi.qimr.edu.au/general/daleN/SNPSpD/) showed a strong block of LD across the first 37.5 kb of the gene (Figure 1). This means that haplotypes are preserved in this region, but are not commonly carried together with haplotypes from the surrounding region in the general population. This haplotype block contains *Pvu*II and *Xba*I, the two SNPs genotyped by Ioannidis *et al.* (1). It extends nearly 18 kb upstream from the beginning of the gene and includes the area where the analyzed microsatellite is also located. From these data, we calculated that only 20% of all observed *ESR1* haplotypes were “tagged” by the *Pvu*II marker. It is unlikely that *Xba*I and microsatellite genotyping would have tagged many haplotypes from the rest of the gene either, as they all lie within the same haplotype block. The mean pairwise LD between *Pvu*II and other *ESR1* SNPs is D’ = 0.47 (Figure 2). D’ is a measure of LD usually reported in a range from 0 (no LD) to 1 (maximum LD) (5). Thirty percent of SNPs have a D’ of ≤ 0.3, which is considered the reasonable minimum extent of LD for association studies (6).

BoneKEy-Osteovision. 2005 February;2(2):14-20
http://www.bonekey-ibms.org/cgi/content/full/ibmske;2/2/14
DOI: 10.1138/20050150

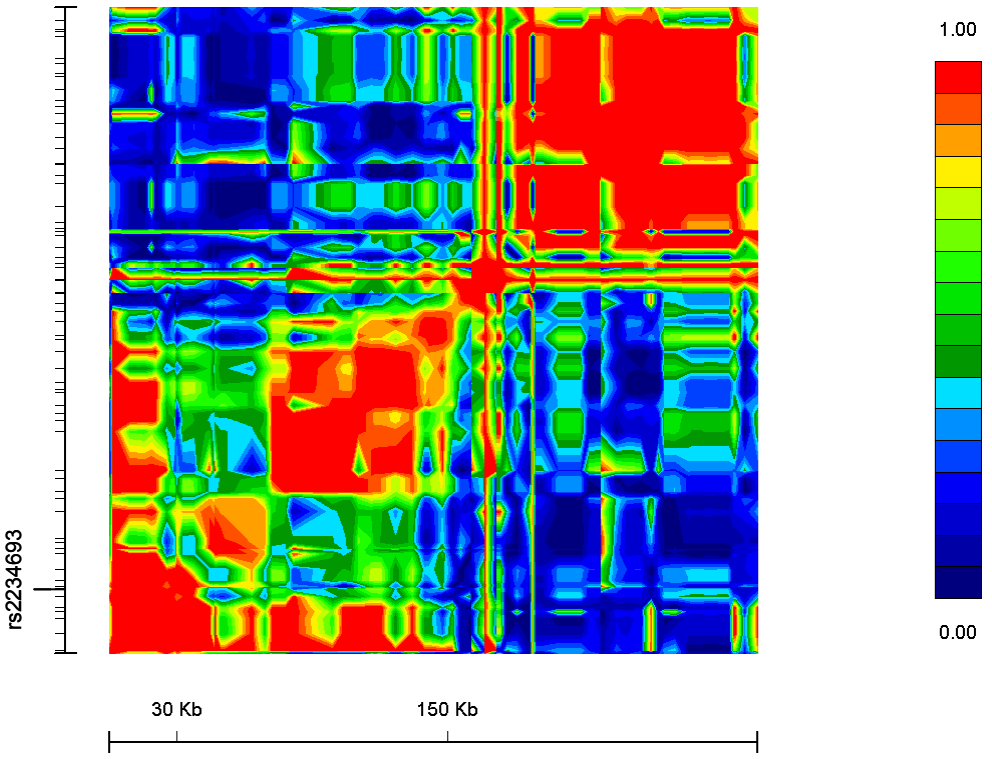

Figure 1: Linkage disequilibrium plot across the *ESR1* gene (measured as D') from the International HapMap Project (see www.hapmap.org). rs2234693 is the *Pvu*II single nucleotide polymorphism (SNP).

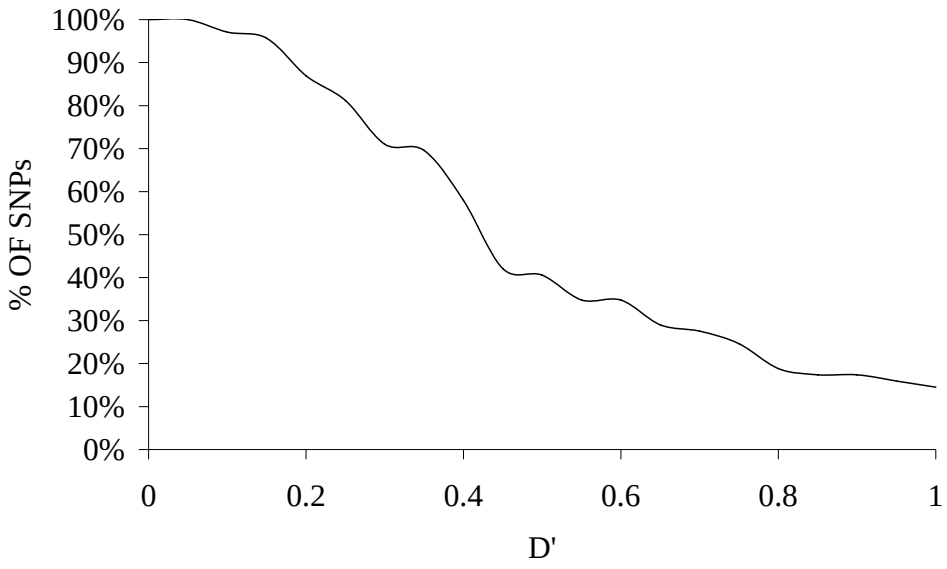

Figure 2: Cumulative pairwise linkage disequilibrium (LD) distribution of *ESR1* single nucleotide polymorphisms (SNPs) with rs2234693. The mean pairwise LD is D'= 0.47.

Clearly, much would be gained by studying a greater proportion of the genetic diversity of the gene. No matter how large the sample size, to detect a disease-causing genetic variant, either the variant itself or markers in linkage disequilibrium with it must be studied. Although more labor intensive, the current study could also be improved by using a selection of different markers. Our entropy analysis identified seven specific SNPs that "tag" more than 90% of the information on the genetic variability of *ESR1*. Only one of the most informative markers for *ESR1* is located in the area covered by variable number tandem repeats and the two SNPs analyzed in this study. Failure to ensure that a high proportion of the genetic diversity of a gene has been captured is not isolated to this study, however -- it is the general experience in osteoporosis genetics, with few exceptions.

Although there are other issues with regard to this paper that could have increased its power, we would like to discuss the following three: the use of fracture as a genetic phenotype, application of the appropriate level of significance, and the assessment and impact of population stratification.

There has been a recent trend to use fracture as a phenotype in genetic studies, based on the argument that it is the clinically relevant endpoint. Although successes in mapping genes for "osteoporosis" on the basis of BMD measures have been reported, an ideal measurement for genetic studies would obviously provide more information about the underlying biology/pathology. It is beyond the scope of this review to discuss this topic at length, but characteristics that we see as essential for such a phenotype would include that trait measurement be noninvasive, cheap and precise, correlated with fracture risk, and heritable.

Four studies have now investigated the heritability of fracture. Deng *et al.* (7) studied the heritability of hip, spine, and wrist fracture in families and showed marginal heritability of hip fracture ($p_{uncorrected} = 0.048$), but not of spine or wrist fracture. A second study by Deng *et al.* (8) suggested possible (but low) heritability of Colles' fracture (narrow-sense heritability = 0.28; statistical significance not stated). A recent twin study also supports heritability of Colles' fracture (9). The interpretation of the findings of a large study of fracture heritability in Scandinavian twins has been debated, with the authors interpreting the findings as not supporting heritability of osteoporotic fracture; however, a reanalysis reaches the opposite conclusion (10,11). Thus, the current evidence suggests that Colles' and hip fracture may be heritable, but the magnitude of the genetic effect is not so large as to be robust and universally detected. No heritability has been demonstrated for vertebral fracture, the phenotype associated with the *Xba*I polymorphism here.

The power and reliability of these studies depend on both accurate diagnosis of fracture and exclusion of fracture in unaffected patients. In many fracture studies, fracture data are self-reported, and no screening is performed to exclude fracture in those reporting no fracture history. Although this is probably reasonable for wrist and hip fracture, it is not adequate for vertebral fracture, which often causes no more than minor symptoms and where a self-reported fracture history is known not to be particularly accurate. It is our opinion that fracture should be used in genetic studies with caution, until there is convincing evidence of heritability, and that objective evaluation should be required to confirm the presence or absence of vertebral fracture, if it is to be studied.

The sample size required for fracture studies is also very difficult to obtain. In this very large study, there were 1072 cases of vertebral fracture and 4952 cases of other fracture types. The association reported with vertebral fracture is for a recessive model, with the minor allele homozygote genotype carrying an association with fracture with an odds ratio of 0.65. The power of the study to identify this association was 100% (calculated using statistics at http://statgen.iop.kcl.ac.uk), assuming *Xba*I is the disease-causing SNP. However, it is more likely that *Xba*I is a marker of a

haplotype carrying another polymorphism, which causes the genetic dysfunction that leads to the effect on fracture risk. In such a case, the power to have detected an effect on fracture is lower, depending on the degree of LD with the true disease-causing allele and the frequency of that allele in the general population. Assuming a minor allele frequency (MAF) of 0.5, for a significance threshold of $p < 0.05$, the study has only 80% power to detect the strength of association reported, if D' > 0.69. A true disease-causing allele may well have a much lower MAF. Assuming a more realistic SNP MAF of 0.1, the study then has only 5% power, even if D' = 1. This further highlights the need to study a substantial proportion of the genetic diversity of any candidate gene and the large sample size required. We feel that negative studies should report their power to detect true genetic associations, using likely genetic models.

Because this study only examines a small proportion of the genetic diversity of the *ESR1* gene, it does not exclude the presence of genetic variants of *ESR1* associated with BMD, although it does exclude an effect of the microsatellite and two intronic SNPs on BMD. The paper discusses the apparently different findings for fracture and BMD, suggesting that the biological explanation must be the result of some influence of *ESR1* on bone quality, geometry or turnover, cognition, or muscle strength. We feel that this discussion is premature, as no study has yet been performed that adequately screens the genetic diversity of the *ESR1* gene for involvement in any osteoporosis phenotype, such as BMD or fracture.

Although conservative, criteria for reporting results of linkage studies have been established (12,13). These criteria attempt to take into account the problem of inflation of type 1 errors caused by the large number of statistical comparisons made. No such criteria yet exist for association studies, but are needed. Correction needs to account for the number of markers and genes studied, number of different analyses, and number of different phenotypes assessed. As with many osteoporosis genetic studies, Ioannidis *et al.* (1) have not estimated the level of correction that would be appropriate, but it would be substantial. The authors studied three markers using diverse genetic analyses, tested BMD association at two sites, and analyzed fractures as all fractures combined or vertebral fractures alone. The analyses were then repeated in men, women, and a combined dataset and for individual markers and haplotypes separately. Methods have been developed to estimate the level of correction appropriate for the number of markers and haplotypes studied, which could be applied to other correlated factors in the analysis, such as the different phenotypes and subsets being analyzed (14,15).

A further unconsidered element is that a prior aim of the Genetic Markers for Osteoporosis (GENOMOS) consortium (http://www.cgkp.org.uk/topics/cam_genetics/reeve.html) was to investigate many different genes. It is widely accepted that for linkage studies, correction should be made for genome-wide analysis, because such an analysis will eventually be performed. Where it is the *a priori* plan of investigators to study multiple genes, estimating the appropriate level of correction to apply or other methods of confirming the result employed should be considered. If an investigator reports the findings of association tests on several genes in one paper, it would naturally be expected that significance values should be corrected for this. If the investigator chooses to publish separate papers reporting results on individual genes, the same statistical threshold should apply. Clearly, $p_{uncorrected} < 0.05$ cannot be considered sufficient evidence of association to be definitive. As genotyping becomes cheaper and more rapid, this problem will increase. Unless appropriately stringent statistical tests are applied, even more confusion as to the reliability of association study findings is likely.

Lastly, we would like to briefly consider the issue of population stratification. Although it is not mentioned in the article, we assume that this cohort is mainly or exclusively white, although the ethnicities involved range from Scandinavian to Mediterranean. These are clearly genetically diverse groups, and methods have been developed to cope

with this situation (16). Stratification is a potential problem for most populations and has rarely been considered in osteoporosis genetic studies. Although its effects may be less obvious in cohort than case-control studies, they nonetheless exist. Because simple effective methods of assessing and dealing with population stratification are available, they should be employed where the problem is likely to exist.

Does this study represent a model for future osteoporosis studies? We hope not – although the large sample size employed here has clarified much about *ESR1* and osteoporosis. If sample sizes this large are required to identify genetic effects, screening the whole genome will be unfeasible. The GENOMOS collection is a fantastic resource for osteoporosis genetics studies and undoubtedly will make a significant contribution to the field in the future. However, more efficient study designs will have to be employed if hypothesis-free genetic approaches to identify novel disease-causing genetic associations are to be performed.

In summary, this study is a major advance on previous studies of the *ESR1* gene and its effect on BMD and fracture risk. It confirms previous data suggesting that *ESR1* genetic variation is a probable cause of osteoporotic fracture (at least vertebral fracture) and excludes an effect from the specific markers studied on BMD. However, whether other *ESR1* variants influence BMD is as yet unknown, and the genetic variant(s) responsible for the observed association with fracture have yet to be identified with certainty.

References

1. Ioannidis JP, Ralston SH, Bennett ST, Brandi ML, Grinberg D, Karassa FB, Langdahl B, van Meurs JB, Mosekilde L, Scollen S, Albagha OM, Bustamante M, Carey AH, Dunning AM, Enjuanes A, van Leeuwen JP, Mavilia C, Masi L, McGuigan FE, Nogues X, Pols HA, Reid DM, Schuit SC, Sherlock RE, Uitterlinden AG; GENOMOS Study. Differential genetic effects of ESR1 gene polymorphisms on osteoporosis outcomes. *JAMA.* 2004 Nov 3;292(17):2105-14.

2. Liu YZ, Liu YJ, Recker RR, Deng HW. Molecular studies of identification of genes for osteoporosis: the 2002 update. *J Endocrinol.* 2003 May;177(2):147-96.

3. Ioannidis JP, Stavrou I, Trikalinos TA, Zois C, Brandi ML, Gennari L, Albagha O, Ralston SH, Tsatsoulis A; ER-alpha Genetics Meta-Analysis. Association of polymorphisms of the estrogen receptor alpha gene with bone mineral density and fracture risk in women: a meta-analysis. *J Bone Miner Res.* 2002 Nov;17(11):2048-60.

4. Kobayashi S, Inoue S, Hosoi T, Ouchi Y, Shiraki M, Orimo H. Association of bone mineral density with polymorphism of the estrogen receptor gene. *J Bone Miner Res.* 1996 Mar;11(3):306-11.

5. Lewontin, R.C. The Interaction of Selection and Linkage. I. General considerations; heterotic models. *Genetics.* 1964 Oct;50:49-67.

6. Kruglyak L. What is significant in whole-genome linkage disequilibrium studies? Am J Hum Genet. 1997 Oct;61(4):810-2.

7. Deng HW, Mahaney MC, Williams JT, Li J, Conway T, Davies KM, Li JL, Deng H, Recker RR. Relevance of the genes for bone mass variation to susceptibility to osteoporotic fractures and its implications to gene search for complex human diseases. *Genet Epidemiol.* 2002 Jan;22(1):12-25.

8. Deng HW, Chen WM, Recker S, Stegman MR, Li JL, Davies KM, Zhou Y, Deng H, Heaney R, Recker RR. Genetic determination of Colles' fracture and differential bone mass in women with and without Colles' fracture. *J Bone Miner Res.* 2000 Jul;15(7):1243-52.

9. Andrew T, Antioniades L, Scurrah KJ, Macgregor AJ, Spector TD. Risk of wrist fracture in women is heritable and is influenced by genes that are largely independent of those influencing BMD. *J Bone Miner Res.* 2005 Jan;20(1):67-74.

10. Kannus P, Palvanen M, Kaprio J, Parkkari J, Koskenvuo M. Genetic factors and osteoporotic fractures in elderly people: prospective 25 year follow up of a nationwide cohort of elderly Finnish twins. *BMJ.* 1999 Nov 20;319(7221):1334-7.

11. MacGregor A, Snieder H, Spector TD. Genetic factors and osteoporotic fractures in elderly people. Twin data support genetic contribution to risk of fracture. *BMJ.* 2000 Jun 17;320(7250):1669-70; author reply 1670-1.

12. Lander E, Kruglyak L. Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nat Genet.* 1995 Nov;11(3):241-7.

13. Nyholt DR. All LODs are not created equal. *Am J Hum Genet.* 2000 Aug;67(2):282-8.

14. Nyholt DR. A simple correction for multiple testing for single-nucleotide polymorphisms in linkage disequilibrium with each other. *Am J Hum Genet.* 2004 Apr;74(4):765-9.

15. Becker T, Knapp M. A powerful strategy to account for multiple testing in the context of haplotype analysis. *Am J Hum Genet.* 2004 Oct;75(4):561-70.

16. Pritchard JK, Donnelly P. Case-control studies of association in structured or admixed populations. *Theor Popul Biol.* 2001 Nov;60(3):227-37.