

PERSPECTIVES

Using Global Gene Expression to Dissect the Genetics of Osteoporosis

Charles R. Farber

Center for Public Health Genomics, Departments of Medicine (Division of Cardiovascular Medicine) and Biochemistry and Molecular Genetics, University of Virginia, Charlottesville, Virginia, USA

Abstract

Over the last decade, much effort has been devoted to unraveling the genetics of osteoporosis-related traits. Although progress has been slow, a number of significant advances have been made recently through the use of genome-wide association (GWA) approaches. Despite these successes, however, our understanding is still very incomplete. Global gene expression data are starting to be used to improve gene discovery and elucidate the mechanisms underlying genetic associations. This *Perspective* focuses on these “systems genetics” approaches and how they are being used to complement and enhance genetic analyses and improve our understanding of osteoporosis. *IBMS BoneKEy*. 2010 October;7(10):353-363.
©2010 International Bone & Mineral Society

Introduction

Osteoporosis is a disease of weak and fracture-prone bones that affects approximately 200 million people worldwide (1). It is characterized by low bone mass, coupled with the microarchitectural deterioration of bone, which results in an increased risk of fracture (2). Fractures are a significant public health burden in large part because they are associated with increases in morbidity and mortality (3). Many intrinsic characteristics of the skeleton contribute to bone strength including bone mass, size, morphology and material properties (4). Although such traits are influenced by both the environment and genetics, heritability estimates generally exceed 50%, suggesting that osteoporosis is primarily a genetic disorder (5). As an example, the heritability of BMD at various anatomical sites is commonly estimated at upwards of 80% (5). These data indicate that a thorough understanding of osteoporosis will require the genetic dissection of its component traits.

Genome-wide Association Studies

Over the last three years, osteoporosis genetics has entered a new era – the age of genome-wide association (GWA) (6). GWA studies are performed by genotyping hundreds of thousands of single nucleotide polymorphisms (SNPs) in thousands of unrelated individuals (7). Significant associations are identified by comparing SNP allele frequencies in cases (e.g., individuals with fractures) versus controls (e.g., individuals without fractures) or associating SNP genotypes with a change in a quantitative trait. To date, ~40 unique loci have been identified for areal BMD, bone size and geometry using GWA (6). These findings have confirmed the role of genes such as the estrogen receptor (*ESR1*), tumor necrosis factor receptor superfamily, member 11a, NF κ B activator (*TNFRSF11A*; RANK), tumor necrosis factor (ligand) superfamily, member 11 (*TNFSF11*; RANKL), osterix (*SP7*), among many others, that were first determined to affect bone through functional and candidate gene studies (6). Importantly, many of the GWA loci implicate novel genes that have not been associated previously with

osteoporosis-related traits and their validation will likely reveal novel biological processes that impact bone. Thus, GWA studies have and will continue to revolutionize the genetic analysis of osteoporosis-related traits.

Although GWA studies have resulted in a treasure trove of novel high-resolution genetic associations, the initial studies have also highlighted drawbacks of GWA studies and all strict “genotype to phenotype” approaches. One of the most significant drawbacks is that the GWA variants identified for bone traits to date account only for a small portion of the estimated heritability (8). A recent meta-analysis of five BMD GWA studies comprised of nearly 20,000 individuals identified 20 significantly associated SNPs (9). However, these variants in aggregate only explained 2-3% of the total variance in hip or spine BMD, suggesting that bone traits are far more genetically complex than originally thought (10). It is possible that hundreds, maybe even thousands, of common and rare alleles affect fracture risk, each accounting for a tiny fraction of the variation in BMD at the population level. A second, equally important limitation of GWA studies is that they do not provide information on the function of associated variants. The use of approaches that both improve variant/gene discovery and provide physiologically relevant information on their function will enhance basic bone biology and aid in converting genetic discoveries into new therapies.

Systems Genetics

Cellular systems are comprised of a series of components (11). Examples of components include the genome, transcriptome (all transcribed sequences), proteome (all proteins) and metabolome (all metabolites). Deficiencies in bone cell function are the result of genetic and environmental perturbations that disrupt individual component function or component-component interactions. Thus, qualitative and quantitative information on these various entities can be used to determine how DNA variation alters cellular

function. A few components besides the genome, most notably the transcriptome (due to the widespread application of DNA microarrays), can be assayed in a global high-throughput manner. The evaluation of global gene expression data in the context of a genetics study is generally referred to as systems genetics. This approach is particularly important given that regulatory polymorphisms, and transcriptional perturbations that arise from structural mutations, play key roles in modulating complex diseases (12). Moreover, a number of bone-related GWA associations map to regions devoid of protein-coding genes and are presumed to be due to regulatory polymorphisms (13). There are a number of analytical approaches that can be used to analyze systems genetics data, such as expression SNP (eSNP) identification, causality modeling and network analysis (14).

Gene expression traits derived from DNA microarrays can be mapped by GWA or linkage analysis in a way that is identical to any other quantitative physiological trait (15). Thus, one can identify genetic loci or associations that regulate gene expression on a genome-wide scale. In the context of a GWA study, these associations are referred to as eSNPs. In a linkage analysis, they are referred to as expression quantitative trait loci (eQTLs). To avoid confusion, this *Perspective* uses the term eSNPs to refer to both.

There are two types of eSNPs, local and distant (Fig. 1) (16). Local eSNPs lie in close proximity to the gene they regulate, while distant eSNPs are removed from the structural gene whose expression they control. Examples of local eSNPs would include *cis*-acting polymorphisms in a gene's promoter or intronic regulatory region that alter transcriptional kinetics or exonic mutations that alter mRNA stability. In contrast, distant eSNPs are *trans*-acting transcriptional regulators. An example would be a transcription factor on human Chromosome (Chr.) 2 that affects the expression of both alleles of its target gene on Chr. 10. Examples of local and distant eSNPs for the expression of genes that play

key roles in bone development are provided in Fig. 1A and Fig. 1B. Importantly, eSNPs connect DNA variation with a specific cellular function and, as will be highlighted below, this information can be used to

elucidate the mechanistic underpinnings of genetic associations. How eSNP identification is being used to advance osteoporosis genetics is specifically discussed below.

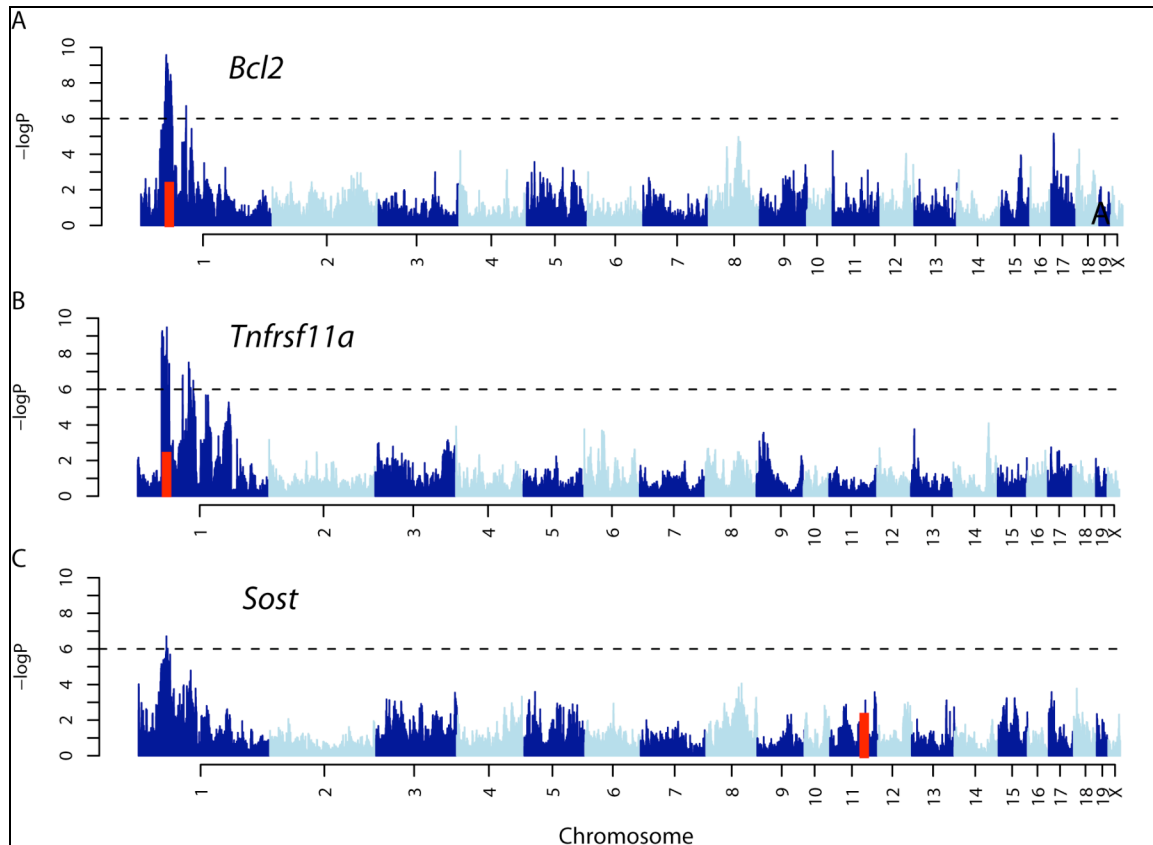


Fig. 1. Local and distant eSNPs affecting the expression of genes that play a key role in bone development. A GWA analysis was performed for *Tnfrsf11a*, *Bcl2* and *Sost* in cortical bone samples (femoral diaphysis free of marrow) from the Hybrid Mouse Diversity Panel (HMDP) (17). The HMDP consists of ~100 inbred strains that have been genotyped at ~135,000 SNPs. Manhattan plots showing that (A) *Bcl2* and (B) *Tnfrsf11a* expression in bone is regulated by local eSNPs. Manhattan plot showing that (C) *Sost* expression in bone is regulated by distant eSNPs. In all three plots the association between SNP genotypes and gene expression is plotted as the $-\log_{10}(P\text{-value})$. The red bars signify the genomic location of each respective gene. The dotted line represents the genome-wide significance ($P < 0.05$) threshold.

Using eSNP Data to Inform Genetic Studies

This section identifies studies that have used eSNP information to determine the most likely causal genes underlying GWA loci and a potential mechanism of action for the associated SNPs. This discussion should be prefaced by mentioning that many of these studies have used data on lymphoblastoid cell lines (LCLs), due to the availability of such data from a large number

of eSNP studies that can be easily obtained through public databases (examples include raw expression data from NCBI's Gene Expression Omnibus database (<http://www.ncbi.nlm.nih.gov/geo/>) and eSNP analysis results from the eQTL browser (<http://eqtl.uchicago.edu/cgi-bin/gbrowse/eqtl/>)). While this may seem like a somewhat irrelevant cell-type for bone it is possible that a gene regulated by a local eSNP in LCLs will also be regulated by the same local eSNP in bone cells or a bone-

relevant tissue/cell type. There are studies that have identified eSNPs in bone-derived cells such as human primary osteoblasts (HObs) (18). While it will be difficult to collect it is likely that systems genetics data on a wide array of tissues (e.g., whole-bone from multiple anatomical locations) and cell types (e.g., osteoblasts, osteoclasts and osteocytes) that are relevant to bone will be much more informative for osteoporosis GWA studies.

A number of osteoporosis GWA studies have used eSNP data to query the function of associated variants. One of the first GWA studies to include eSNP data was by Richards *et al.* in which the authors identified a SNP (rs4355801) located upstream of tumor necrosis factor receptor superfamily, member 11b (*TNFRSF11B*; osteoprotegerin (OPG)) that was associated with BMD and risk of osteoporosis (19). Using microarray data from LCLs from 55 unrelated HapMap individuals, it was determined that rs4355801 was also a local eSNP regulating the expression of *TNFRSF11B*. These data are consistent with the observation that perturbing the expression of *Tnfrsf11b* in knockout and transgenic mice alters bone mass (20;21). Similarly, in a large meta-analysis of five BMD GWA studies, Rivadeneira *et al.* identified 20 loci affecting BMD, of which 13 were newly identified (9). Using microarray data on primary human osteoblasts, SNPs within three of the 13 novel loci regulated the expression of the genes (G protein-coupled receptor 177 (*GPR177*), myocyte enhancer factor 2C (*MEF2C*) and forkhead box C2 (*FOXC2*)) nearest the most significant SNPs. Kung *et al.* used quantitative real-time PCR to identify that the SNP rs2273061, which was strongly associated with BMD and located in the third intron of the *JAG1* gene, regulated the expression of *JAG1* in human-derived bone cells and peripheral mononuclear blood cells (22). Additionally, Hsu *et al.* recently identified seven loci affecting BMD and/or femoral neck geometry (23). Three of the top seven SNPs were also found to be eSNPs that regulated the genes (RAS-related protein-1a (*RAP1A*), TBC1 domain family, member 8 (with GRAM domain)

(*TBC1D8*) and *TNFRSF11B*) nearest the trait-associated SNPs. The identification of SNPs associated with bone traits that are also eSNPs not only suggests a mechanism of action, but it also identifies the individual genes within associations (which typically implicate more than one gene) that are most likely responsible for the associations.

In addition to determining if SNPs associated with bone traits potentially do so through changes in gene expression, it is also possible to inform GWA studies by doing the reverse: identifying eSNPs first and then determining if they affect a clinical bone trait. There have been two recent studies, from the same group, that nicely outline this approach (18;24). Both studies utilized microarray data generated on HObs. In the first study, a total of 95 HObs samples were assayed using DNA microarrays (18). The samples were also genotyped at high-density. Using GWA the authors identified several hundred genes regulated by local eSNPs in HObs. Next they cross-referenced the list of eSNPs with a list of the top SNPs identified in a large BMD GWA study (13). They then made two key observations. First, there was a significant enrichment of HObs eSNPs among those that were also associated with BMD. A parallel analysis using LCLs did not reveal this enrichment, suggesting that it is advantageous to use primary bone cells (or presumably bone tissue) for systems genetic studies of osteoporosis as compared to the more assessable cells or cell lines such as LCLs. Second, of the top 10 local eSNPs also associated with BMD, a variant in the serine racemase (*SRR*) gene was found to replicate in two independent studies, providing strong support for the hypothesis that differences in its expression perturb BMD. A similar approach was used in a second study of 60 HObs samples that were profiled using Affymetrix exon arrays (24). Instead of probing the expression of a gene with one or a small number of probes, exon arrays consist of probes for the majority of characterized exons in the genome. Having data on all exons allows one to identify differentially expressed transcript isoforms. In this particular study, these data were used to identify a novel transcript isoform of

the *FAM118A* gene whose expression was regulated by a local eSNP (rs136564). This eSNP was also found to be associated with BMD in two independent studies.

In addition to human GWA studies, eSNP information can be used in the context of mouse linkage studies to identify genes affecting bone mass. One approach is to compare “expression signatures” that are created by BMD QTLs and to experimentally perturb individual positional and functional candidate genes in the same genomic region. If the experimental single gene perturbation signature significantly overlaps the QTL signature then it is likely to be the causal gene. This approach was used to identify arachidonate 5-lipoxygenase (*Alox5*) as a BMD gene (25). Linkage analysis in a cross between the C57BL6/J and DBA/2J mouse strains identified a BMD QTL on Chr. 6. This locus also regulated in *trans* the expression of nearly 2,000 genes (~10% of the genome) in the liver. Of the 172 positional candidates within the Chr. 6 locus, none appeared to be controlling the trait via genetically regulated differences in expression. *Alox5* was the only gene in the region to harbor a missense mutation between the two strains. By generating an expression signature using DNA microarray profiles of livers from *Alox5* knockout mice (*Alox5*^{-/-}), the authors found that many of the genes perturbed by the loss of *Alox5* were regulated by distant eSNPs that co-localized with the Chr. 6 BMD QTL. This observation was consistent with *Alox5* being the causal gene. This hypothesis was further supported by the observation that *Alox5*^{-/-} mice had reduced BMD. This study highlights the fact that gene expression data can be used for gene discovery, even when the causal variant does not directly affect gene expression.

Causality Modeling – Orienting Genes and Traits

In GWA studies, the identification of eSNPs associated with bone traits is strong evidence of a functional connection between the gene and clinical trait; however, it is possible that the gene’s expression is “reacting” to the change in phenotype or the

two may be completely independent. Furthermore, in linkage studies, many genes regulated by local eSNPs may reside within bone trait QTLs that span several megabase pairs (Mbp). Approaches that provide a way to prioritize such eSNPs would have the potential to be powerful tools for gene discovery.

Causality modeling algorithms have been developed that allow one to “orient” the relationships between variation in DNA, gene expression traits and clinical traits (26-28). To illustrate how causality modeling works, consider the simple example of a SNP that is associated with both differences in a gene’s expression and BMD. We know that the flow of information has to begin with the SNP (*i.e.*, genetic variation can alter a gene’s expression and/or BMD, but changes in expression or BMD do not alter primary DNA sequences); therefore, the possible relationships can be modeled as: 1) causal (SNP→Gene expression→BMD), 2) reactive (SNP→BMD→Gene expression), or 3) independent (Gene expression←SNP→BMD) (Fig. 2A). Probabilities for each model can be calculated using likelihood-based approaches or structural equation models (27;28). Hypotheses can then be drawn based on relative model probabilities. Typically, the causal model is the one we are most interested in because it links a gene’s expression to a change in a clinical trait; however, the approach can also be used to identify “reactive” genes downstream of other genes, such as key transcriptional regulators. Fig. 2B illustrates this latter point by determining (from the overlapping local and distant eSNP data presented in Fig. 1) that B-cell leukemia/lymphoma 2 (*Bcl2*) expression, and not the expression of *Tnfrsf11a*, is predicted to be responsible for the alterations in sclerostin (*Sost*) expression in mouse cortical bone.

It is important to remember that causality modeling is a statistical prediction based on the given data and unknown hidden confounders can influence the results; therefore, the role of such genes should always be validated. We have demonstrated recently that 88% (8 of 9) of genes predicted

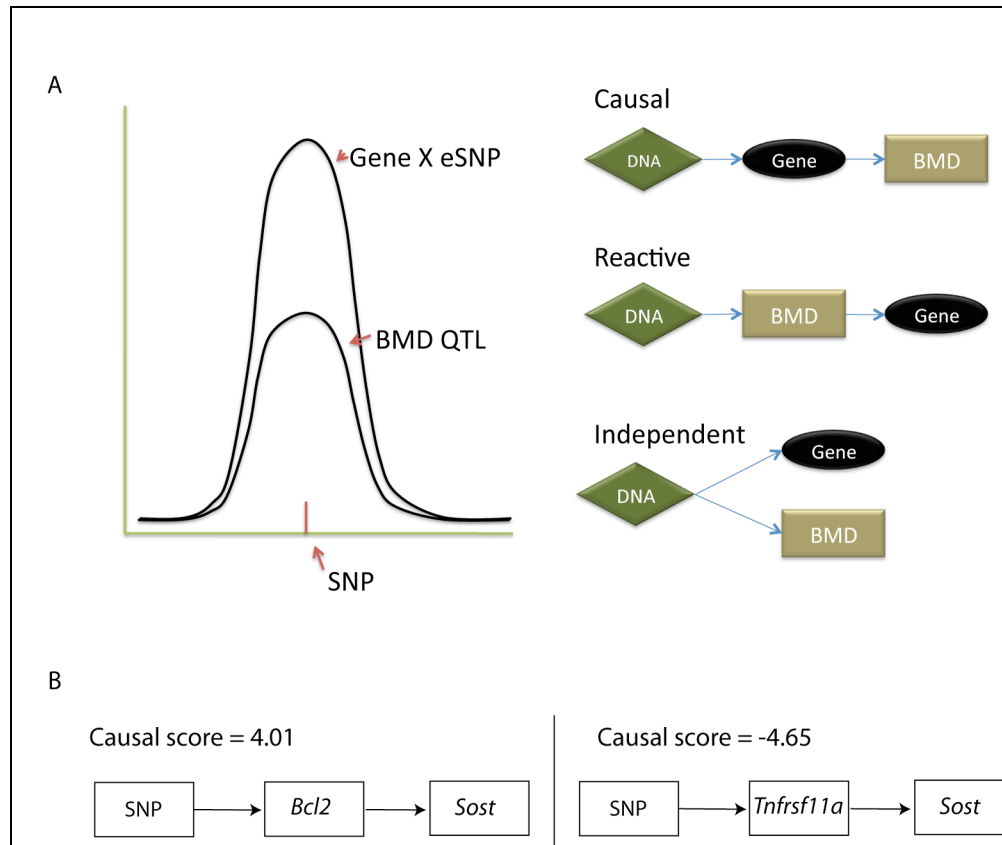


Fig 2. Causality modeling can “orient” the relationships between correlated traits. (A) Example of a SNP pleiotropically controlling both the expression of a gene and BMD. This example illustrates coincidence between an eSNP and BMD QTL identified via linkage (the same concepts are true for eSNPs and BMD associations identified in a GWA study). The coincidence between the eSNP and the BMD QTL suggests there may be a functional connection between the gene’s expression and BMD. Causality modeling can be used to define this relationship by determining the relative likelihoods for three models: 1) the “causal” model that predicts that the gene’s expression is causing a change in BMD, 2) the “reactive” model that predicts that the gene’s expression is reacting to a difference in BMD, and 3) the “independent” model where there is no functional relationship between the gene’s expression and BMD. (B) Causality modeling applied to the eSNP information from Fig. 1 predicted that *Bcl2* and not *Tnfrsf11a* expression was regulating the expression of *Sost* in cortical bone. The causal score is the $-\log_{10}$ of the ratio of causal model probability to the probability of the best competing model. A positive causal score of 4.01 for *Bcl2* suggests it is regulating *Sost* expression, whereas the negative score for *Tnfrsf11a* indicates that it is not causal for *Sost* expression. In the case of *Tnfrsf11a*, the independent model, not the causal model, fits the data the best, hence the negative causal score.

as causally linked to obesity resulted in differences in fat accumulation in transgenic or knockout mice (29).

Our group has used this approach to identify candidate causal genes for BMD (30). Using an intercross between the C57BL6/J and C3H/HeJ inbred mouse strains, we identified a total of nine QTLs affecting femoral BMD using DNA microarray data from adipose tissue. A total of 148 genes, located within one of the nine BMD loci, were identified

that were regulated by local eSNPs and their expression was significantly correlated with BMD. Using the Network Edge Orienting (NEO) causality modeling algorithm (27) we determined that 18 of the 148 genes were predicted to be causally linked to changes in BMD. Many of the genes were highly expressed in osteoblasts, suggesting that the expression of these genes in bone, and not adipose tissue, was actually causal and they were detected because their expression in adipose tissue and bone was

highly correlated. Several of the candidate causal genes such as *Twist2*, *Mmp14* and *Wnt9a* are known to be involved in bone development (30).

Causality modeling in mice can also be used to prioritize candidate genes from GWA studies. In a recent large GWA of multiple osteoporosis-related traits (23), we used the likelihood-based causal model selection (LCMS) algorithm (28) to prioritize genes within 109 GWA loci that were suggestively ($P \leq 5.0 \times 10^{-5}$) associated with various bone traits. Using the same B6 x C3H F2 intercross described above, we applied LCMS to the expression of mouse homologs within each association, using microarray expression data from various tissues and a number of bone traits. In total, 12 genes were predicted as causal for at least one bone trait. This approach allowed the authors to prioritize these 12 loci as the most important for future follow-up genetic and functional studies. To date, causality modeling for osteoporosis-related traits has only been applied to mouse data; however, causality modeling can be directly applied to human data, and with a growing number of GWA datasets that include DNA microarray profiles it has the potential to help explain a portion of the missing heritability for bone traits. This approach could be used to "mine" GWA data by identifying the subset of SNPs that, due to a lack of power in the GWA analysis, failed to reach genome-wide significance (e.g., SNPs with P-values less than 0.001), but truly alter osteoporosis by perturbing a gene's expression. An example of such an eSNP would be the IL-6 -174 promoter polymorphism that has been linked to subtle changes in IL-6 expression, IL-6 plasma levels and BMD.

Future Directions

The future of eSNP discovery and systems genetics will be driven by technology. New technologies such as next-generation sequencing (NGS) are already making an impact. NGS can provide a comprehensive digital readout of gene expression (RNA-seq) as well as elucidate qualitative differences, such as alternative splice isoforms and chimeric transcripts, for entire

transcriptomes (31). RNA-seq has been used recently for population-based whole-transcriptome studies to comprehensively identify coding SNPs, eSNPs and SNPs affecting alternative splicing (32;33). In the very near future, as new NGS technologies emerge and costs decrease, it will be possible to use RNA-seq to characterize transcriptomes on a large-scale basis. Such data combined with whole-genome DNA sequences (also generated using NGS) will substantially increase our ability to comprehensively catalog polymorphisms in the population that alter the transcriptome and risk of osteoporosis.

In addition, the collection of data from other cellular components will improve. Array-based and to some extent NGS approaches are already being used to identify genetic variation that affects DNA methylation and its connection with transcriptional alterations (34). The incorporation of proteomic (35;36) and metabolomic data (37) into systems genetics studies is also increasing and this will continue with improvements in technologies for assaying these components. Of course one of the consequences of increases in the volume and types of data for systems genetic studies will be the need for improved analytical and computing resources.

One of the obstacles facing human eSNP and systems genetics studies of osteoporosis is the difficulty of obtaining bone tissue (or other non-osseous tissues that impact bone, e.g., adipose and neuronal tissues) or bone-derived cells from a large population of "normal" subjects for genomic profiling. It is also difficult to characterize large human cohorts for the wide array of bone phenotypes (e.g., bone mass, morphology, mechanical and material properties) that contribute to the risk of fracture. Therefore, it is likely that the mouse, which is amenable to the collection of vast amounts of diverse data and biological samples and shares many molecular similarities in bone development with humans, will play a major role in osteoporosis systems genetics. High-resolution mapping populations such as the HMDP (17;38) and the Collaborative Cross

(CC) (39;40) will make such endeavors much more productive and efficient relative to the designed crosses that have typically been used for such experiments (30).

Conclusion

GWA is rapidly improving our understanding of the genetics of osteoporosis. Systems genetics is an approach that has the potential to complement and enhance GWA through techniques such as eSNP identification and causality modeling. Importantly, systems genetic studies promise to improve our ability to dissect the genetic basis of osteoporosis.

Conflict of Interest: None reported.

Peer Review: This article has been peer-reviewed.

References

1. Reginster JY, Burlet N. Osteoporosis: a still increasing prevalence. *Bone*. 2006 Feb;38(2 Suppl 1):S4-9.
2. Consensus development conference: diagnosis, prophylaxis, and treatment of osteoporosis. *Am J Med*. 1993 Jun;94(6):646-50.
3. Dennison E, Mohamed MA, Cooper C. Epidemiology of osteoporosis. *Rheum Dis Clin North Am*. 2006 Nov;32(4):617-29.
4. Seeman E, Delmas PD. Bone quality--the material and structural basis of bone strength and fragility. *N Engl J Med*. 2006 May 25;354(21):2250-61.
5. Peacock M, Turner CH, Econs MJ, Foroud T. Genetics of osteoporosis. *Endocr Rev*. 2002 Jun; 23(3):303-26.
6. Xu XH, Dong SS, Guo Y, Yang TL, Lei SF, Papasian CJ, Zhao M, Deng HW. Molecular genetic studies of gene identification for osteoporosis: the 2009 update. *Endocr Rev*. 2010 Aug;31(4):447-505.
7. Altshuler D, Daly MJ, Lander ES. Genetic mapping in human disease. *Science*. 2008 Nov 7;322(5903):881-8.
8. Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorf LA, Hunter DJ, McCarthy MI, Ramos EM, Cardon LR, Chakravarti A, Cho JH, Guttmacher AE, Kong A, Kruglyak L, Mardis E, Rotimi CN, Slatkin M, Valle D, Whittemore AS, Boehnke M, Clark AG, Eichler EE, Gibson G, Haines JL, Mackay TF, McCarroll SA, Visscher PM. Finding the missing heritability of complex diseases. *Nature*. 2009 Oct 8;461(7265):747-53.
9. Rivadeneira F, Styrkarsdottir U, Estrada K, Halldórsson BV, Hsu YH, Richards JB, Zillikens MC, Kavvoura FK, Amin N, Aulchenko YS, Cupples LA, Deloukas P, Demissie S, Grundberg E, Hofman A, Kong A, Karasik D, van Meurs JB, Oostra B, Pastinen T, Pols HA, Sigurdsson G, Soranzo N, Thorleifsson G, Thorsteinsdottir U, Williams FM, Wilson SG, Zhou Y, Ralston SH, van Duijn CM, Spector T, Kiel DP, Stefansson K, Ioannidis JP, Uitterlinden AG; Genetic Factors for Osteoporosis (GEFOS) Consortium. Twenty bone-mineral-density loci identified by large-scale meta-analysis of genome-wide association studies. *Nat Genet*. 2009 Nov;41(11):1199-206.
10. Yang J, Benyamin B, McEvoy BP, Gordon S, Henders AK, Nyholt DR, Madden PA, Heath AC, Martin NG, Montgomery GW, Goddard ME, Visscher PM. Common SNPs explain a large proportion of the heritability for human height. *Nat Genet*. 2010 Jul;42(7):565-9.
11. Farber CR, Luskis AJ. Integrating global gene expression analysis and genetics. *Adv Genet*. 2008;60:571-601.
12. Kleinjan DA, van Heyningen V. Long-range control of gene expression: emerging mechanisms and disruption in disease. *Am J Hum Genet*. 2005 Jan;76(1):8-32.

13. Styrkarsdottir U, Halldorsson BV, Gretarsdottir S, Gudbjartsson DF, Walters GB, Ingvarsson T, Jonsdottir T, Saemundsdottir J, Center JR, Nguyen TV, Bagger Y, Gulcher JR, Eisman JA, Christiansen C, Sigurdsson G, Kong A, Thorsteinsdottir U, Stefansson K. Multiple genetic loci for bone mineral density and fractures. *N Engl J Med*. 2008 May 29;358(22):2355-65.
14. Farber CR, Lusic AJ. Future of osteoporosis genetics: enhancing genome-wide association studies. *J Bone Miner Res*. 2009 Dec;24(12):1937-42.
15. Schadt EE, Monks SA, Drake TA, Lusic AJ, Che N, Colinayo V, Ruff TG, Milligan SB, Lamb JR, Cavet G, Linsley PS, Mao M, Stoughton RB, Friend SH. Genetics of gene expression surveyed in maize, mouse and man. *Nature*. 2003 Mar 20;422(6929):297-302.
16. Rockman MV, Kruglyak L. Genetics of global gene expression. *Nat Rev Genet*. 2006 Nov;7(11):862-72.
17. Farber CR, Bennett BB, Orozco L, Kang HM, Qi H, Kostem E, Ghazalpour A, Wen P, Farber EA, Berberyan A, Suwanwela J, Drake TA, Eskin E, Lusic AJ. A novel systems genetics resource for bone biology. *J Bone Miner Res*. 2009 Sep;24(Suppl 1). Available [here](#).
18. Grundberg E, Kwan T, Ge B, Lam KC, Koka V, Kindmark A, Mallmin H, Dias J, Verlaan DJ, Ouimet M, Sinnott D, Rivadeneira F, Estrada K, Hofman A, van Meurs JM, Uitterlinden A, Beaulieu P, Graziani A, Harmsen E, Ljunggren O, Ohlsson C, Mellström D, Karlsson MK, Nilsson O, Pastinen T. Population genomics in a disease targeted primary cell model. *Genome Res*. 2009 Nov;19(11):1942-52.
19. Richards JB, Rivadeneira F, Inouye M, Pastinen TM, Soranzo N, Wilson SG, Andrew T, Falchi M, Gwilliam R, Ahmadi KR, Valdes AM, Arp P, Whittaker P, Verlaan DJ, Jhamai M, Kumanduri V, Moorhouse M, van Meurs JB, Hofman A, Pols HA, Hart D, Zhai G, Kato BS, Mullin BH, Zhang F, Deloukas P, Uitterlinden AG, Spector TD. Bone mineral density, osteoporosis, and osteoporotic fractures: a genome-wide association study. *Lancet*. 2008 May 3;371(9623):1505-12.
20. Simonet WS, Lacey DL, Dunstan CR, Kelley M, Chang MS, Lüthy R, Nguyen HQ, Wooden S, Bennett L, Boone T, Shimamoto G, DeRose M, Elliott R, Colombero A, Tan HL, Trail G, Sullivan J, Davy E, Bucay N, Renshaw-Gegg L, Hughes TM, Hill D, Pattison W, Campbell P, Sander S, Van G, Tarpley J, Derby P, Lee R, Boyle WJ. Osteoprotegerin: a novel secreted protein involved in the regulation of bone density. *Cell*. 1997 Apr 18;89(2):309-19.
21. Kong YY, Yoshida H, Sarosi I, Tan HL, Timms E, Capparelli C, Morony S, Oliveira-dos-Santos AJ, Van G, Itie A, Khoo W, Wakeham A, Dunstan CR, Lacey DL, Mak TW, Boyle WJ, Penninger JM. OPGL is a key regulator of osteoclastogenesis, lymphocyte development and lymph-node organogenesis. *Nature*. 1999 Jan 28;397(6717):315-23.
22. Kung AW, Xiao SM, Cherny S, Li GH, Gao Y, Tso G, Lau KS, Luk KD, Liu JM, Cui B, Zhang MJ, Zhang ZL, He JW, Yue H, Xia WB, Luo LM, He SL, Kiel DP, Karasik D, Hsu YH, Cupples LA, Demissie S, Styrkarsdottir U, Halldorsson BV, Sigurdsson G, Thorsteinsdottir U, Stefansson K, Richards JB, Zhai G, Soranzo N, Valdes A, Spector TD, Sham PC. Association of JAG1 with bone mineral density and osteoporotic fractures: a genome-wide association study and follow-up replication studies. *Am J Hum Genet*. 2010 Feb 12;86(2):229-39.
23. Hsu YH, Zillikens MC, Wilson SG, Farber CR, Demissie S, Soranzo N, Bianchi EN, Grundberg E, Liang L, Richards JB, Estrada K, Zhou Y, van

- Nas A, Moffatt MF, Zhai G, Hofman A, van Meurs JB, Pols HA, Price RI, Nilsson O, Pastinen T, Cupples LA, Lulis AJ, Schadt EE, Ferrari S, Uitterlinden AG, Rivadeneira F, Spector TD, Karasik D, Kiel DP. An integration of genome-wide association study and gene expression profiling to prioritize the discovery of novel susceptibility loci for osteoporosis-related traits. *PLoS Genet.* 2010 Jun 10;6(6):e1000977.
24. Kwan T, Grundberg E, Koka V, Ge B, Lam KC, Dias C, Kindmark A, Mallmin H, Ljunggren O, Rivadeneira F, Estrada K, van Meurs JB, Uitterlinden A, Karlsson M, Ohlsson C, Mellström D, Nilsson O, Pastinen T, Majewski J. Tissue effect on genetic control of transcript isoform variation. *PLoS Genet.* 2009 Aug;5(8):e1000608.
25. Mehrabian M, Allayee H, Stockton J, Lum PY, Drake TA, Castellani LW, Suh M, Armour C, Edwards S, Lamb J, Lulis AJ, Schadt EE. Integrating genotypic and expression data in a segregating mouse population to identify 5-lipoxygenase as a susceptibility gene for obesity and bone traits. *Nat Genet.* 2005 Nov;37(11):1224-33.
26. Chaibub Neto E, Ferrara CT, Attie AD, Yandell BS. Inferring causal phenotype networks from segregating populations. *Genetics.* 2008 Jun;179(2):1089-100.
27. Aten JE, Fuller TF, Lulis AJ, Horvath S. Using genetic markers to orient the edges in quantitative trait networks: the NEO software. *BMC Syst Biol.* 2008 Apr 15;2:34.
28. Schadt EE, Lamb J, Yang X, Zhu J, Edwards S, Guhathakurta D, Sieberts SK, Monks S, Reitman M, Zhang C, Lum PY, Leonardson A, Thieringer R, Metzger JM, Yang L, Castle J, Zhu H, Kash SF, Drake TA, Sachs A, Lulis AJ. An integrative genomics approach to infer causal associations between gene expression and disease. *Nat Genet.* 2005 Jul;37(7):710-7.
29. Yang X, Deignan JL, Qi H, Zhu J, Qian S, Zhong J, Torosyan G, Majid S, Falkard B, Kleinhanz RR, Karlsson J, Castellani LW, Mumick S, Wang K, Xie T, Coon M, Zhang C, Estrada-Smith D, Farber CR, Wang SS, van Nas A, Ghazalpour A, Zhang B, Macneil DJ, Lamb JR, Dipple KM, Reitman ML, Mehrabian M, Lum PY, Schadt EE, Lulis AJ, Drake TA. Validation of candidate causal genes for obesity that affect shared metabolic pathways and networks. *Nat Genet.* 2009 Apr;41(4):415-23.
30. Farber CR, van Nas A, Ghazalpour A, Aten JE, Doss S, Sos B, Schadt EE, Ingram-Drake L, Davis RC, Horvath S, Smith DJ, Drake TA, Lulis AJ. An integrative genetics approach to identify candidate genes regulating BMD: combining linkage, gene expression, and association. *J Bone Miner Res.* 2009 Jan;24(1):105-16.
31. Wang Z, Gerstein M, Snyder M. RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet.* 2009 Jan;10(1):57-63.
32. Pickrell JK, Marioni JC, Pai AA, Degner JF, Engelhardt BE, Nkadori E, Veyrieras JB, Stephens M, Gilad Y, Pritchard JK. Understanding mechanisms underlying human gene expression variation with RNA sequencing. *Nature.* 2010 Apr 1;464(7289):768-72.
33. Montgomery SB, Sammeth M, Gutierrez-Arcelus M, Lach RP, Ingle C, Nisbett J, Guigo R, Dermitzakis ET. Transcriptome genetics using second generation sequencing in a Caucasian population. *Nature.* 2010 Apr 1;464(7289):773-7.
34. Gibbs JR, van der Brug MP, Hernandez DG, Traynor BJ, Nalls MA, Lai SL, Arepalli S, Dillman A, Rafferty IP, Troncoso J, Johnson R, Zielke HR, Ferrucci L, Longo DL, Cookson MR, Singleton AB. Abundant quantitative trait loci exist for DNA methylation and gene

- expression in human brain. *PLoS Genet.* 2010 May 13;6(5):e1000952.
35. Melzer D, Perry JR, Hernandez D, Corsi AM, Stevens K, Rafferty I, Lauretani F, Murray A, Gibbs JR, Paolisso G, Rafiq S, Simon-Sanchez J, Lango H, Scholz S, Weedon MN, Arepalli S, Rice N, Washecka N, Hurst A, Britton A, Henley W, van de Leemput J, Li R, Newman AB, Tranah G, Harris T, Panicker V, Dayan C, Bennett A, McCarthy MI, Ruokonen A, Jarvelin MR, Guralnik J, Bandinelli S, Frayling TM, Singleton A, Ferrucci L. A genome-wide association study identifies protein quantitative trait loci (pQTLs). *PLoS Genet.* 2008 May 9;4(5):e1000072.
36. Garge N, Pan H, Rowland MD, Cargile BJ, Zhang X, Cooley PC, Page GP, Bunger MK. Identification of quantitative trait loci underlying proteome variation in human lymphoblastoid cells. *Mol Cell Proteomics.* 2010 Jul;9(7):1383-99.
37. Dumas ME, Wilder SP, Bihoreau MT, Barton RH, Fearnside JF, Argoud K, D'Amato L, Wallis RH, Blancher C, Keun HC, Baunsgaard D, Scott J, Sidelmann UG, Nicholson JK, Gauguier D. Direct quantitative trait locus mapping of mammalian metabolic phenotypes in diabetic and normoglycemic rat models. *Nat Genet.* 2007 May;39(5):666-72.
38. Bennett BJ, Farber CR, Orozco L, Kang HM, Ghazalpour A, Siemers N, Neubauer M, Neuhaus I, Yordanova R, Guan B, Truong A, Yang WP, He A, Kayne P, Gargalovic P, Kirchgessner T, Pan C, Castellani LW, Kostem E, Furlotte N, Drake TA, Eskin E, Lusis AJ. A high-resolution association mapping panel for the dissection of complex traits in mice. *Genome Res.* 2010 Feb;20(2):281-90.
39. Chesler EJ, Miller DR, Branstetter LR, Galloway LD, Jackson BL, Philip VM, Voy BH, Culiati CT, Threadgill DW, Williams RW, Churchill GA, Johnson DK, Manly KF. The Collaborative Cross at Oak Ridge National Laboratory: developing a powerful resource for systems genetics. *Mamm Genome.* 2008 Jun;19(6):382-9.
40. Churchill GA, Airey DC, Allayee H, Angel JM, Attie AD, Beatty J, Beavis WD, Belknap JK, Bennett B, Berrettini W, Bleich A, Bogue M, Broman KW, Buck KJ, Buckler E, Burmeister M, Chesler EJ, Cheverud JM, Clapcote S, Cook MN, Cox RD, Crabbe JC, Crusio WE, Darvasi A, Deschepper CF, Doerge RW, Farber CR, Forejt J, Gaile D, Garlow SJ, Geiger H, Gershenfeld H, Gordon T, Gu J, Gu W, de Haan G, Hayes NL, Heller C, Himmelbauer H, Hitzemann R, Hunter K, Hsu HC, Iraqi FA, Ivandic B, Jacob HJ, Jansen RC, Jepsen KJ, Johnson DK, Johnson TE, Kempermann G, Kendzioriski C, Kotb M, Kooy RF, Llamas B, Lammert F, Lassalle JM, Lowenstein PR, Lu L, Lusis A, Manly KF, Marcucio R, Matthews D, Medrano JF, Miller DR, Mittleman G, Mock BA, Mogil JS, Montagutelli X, Morahan G, Morris DG, Mott R, Nadeau JH, Nagase H, Nowakowski RS, O'Hara BF, Osadchuk AV, Page GP, Paigen B, Paigen K, Palmer AA, Pan HJ, Peltonen-Palotie L, Peirce J, Pomp D, Pravenec M, Prows DR, Qi Z, Reeves RH, Roder J, Rosen GD, Schadt EE, Schalkwyk LC, Seltzer Z, Shimomura K, Shou S, Sillanpää MJ, Siracusa LD, Snoeck HW, Spearow JL, Svenson K, Tarantino LM, Threadgill D, Toth LA, Valdar W, de Villena FP, Warden C, Whatley S, Williams RW, Wiltshire T, Yi N, Zhang D, Zhang M, Zou F; Complex Trait Consortium. The Collaborative Cross, a community resource for the genetic analysis of complex traits. *Nat Genet.* 2004 Nov;36(11):1133-7.