Perspectives

The Future of Mouse Genetics in Osteoporosis Research

Cheryl L. Ackert-Bicknell\textsuperscript{1} and Clifford J. Rosen\textsuperscript{2}

\textsuperscript{1}The Jackson Laboratory, Bar Harbor, Maine, USA
\textsuperscript{2}Maine Medical Center Research Institute, Scarborough, Maine, USA

Abstract

Focused studies examining the genetics of osteoporosis in the mouse began approximately fifteen years ago, but as these studies have progressed it has become apparent that the genetics of osteoporosis is more complicated than originally predicted. Traditional F2 inter-cross mapping in the mouse, while valuable, is going to fall short of the goal of mapping all of the genes underlying this disease. In the past three years there has been considerable development of new techniques and resources for mouse genetics. Herein we describe four such mouse genetic resources and/or techniques that can be readily applied to the study of osteoporosis. First we describe a de novo mouse genetic map that has been developed to alleviate historical marker order problems and a tool that can be used to convert between the genetic map and the physical genome sequence map. We describe haplotype association mapping, a QTL mapping technique that can be used in concert with traditional mapping to identify the underlying genes. We also describe expression QTL mapping and how this technique can be used to augment gene discovery efforts. Lastly, we describe the Collaborative Cross, an ambitious set of recombinant inbred mouse strains currently in production. The application of these and other new tools, techniques and resources will accelerate gene discovery and further our understanding of the genetics of osteoporosis.

Key Words: Mouse genetics; Osteoporosis

Introduction

Mouse genetics has been used successfully for more than five decades to explore mechanisms of disease, identify genes and understand therapeutic pathways. Initially, studies of gene function depended on the identification and genetic mapping of spontaneous mutants and the application of Mendelian genetics to understand heritable diseases. Subsequently, chromosomal mouse markers were identified and studies of complex diseases and polygenic traits began. The establishment of genomic engineering to test function and fine mapping of the mouse genome to identify specific genes further propelled the field.

Studies of osteoporosis in the mouse began in earnest about fifteen years ago, as bone density measurements in animals became standardized. Expectations rose in the late 1990s that mouse genetics could identify not only the genes associated with low bone mass, but also the interactive loci, and the environmental modifiers that determined peak bone density. Furthermore, there was hope that mouse ‘genes’ could accelerate the search for osteoporosis-related genes in humans. However, it soon became apparent that the genetics of this disease was more complicated than anticipated and that models other than just inbred strains would be required. Over the last three years tremendous progress has been made in refining new tools to further our understanding of bone biology. In this Perspective, we outline some of the more notable models and their potential utility in searching for ‘osteoporosis’ genes.

The Genetic Map – the Good, the Bad and the Ugly

A quantitative trait locus (QTL) is a region of the genome that is associated with a given phenotype. QTL mapping in mice has been a successful method for determining the
mechanisms of genetic regulation for a variety of complex traits. To date, over 300 QTLs have been mapped in mice for the phenotypes of bone mineral density (BMD), bone shape and bone strength (1-15). Unlike for other phenotypes, only a handful of the underlying genes have been identified (7;16-19). There are several reasons for this. First, BMD is a complicated phenotype involving more than one cell type, rendering the etiology more difficult to sort out. Second, net BMD is the aggregate of all factors acting upon it such as physiological and environmental factors (i.e., kidney function, stress caused by a dominant animal in the cage, diet, etc.) for which the investigator may or may not be able to control. Third, BMD is likely controlled by many genetic loci with each contributing a small amount to the variance, making the mapping of a single locus difficult. Finally, and most easily fixable, is the fact that there have been problems with the primary mouse genetic map used as the backbone for QTL studies.

The mouse genetic map was derived for the most part from two small discrete mapping crosses (20;21). For each chromosome, a series of Chromosome Committees assembled the information from these crosses and information from other sources, generating the traditional mouse genetic map (22). The traditional mouse map is updated on a regular basis (23), but unfortunately, some historical errors with regard to marker order and relative positions have persisted (24). Like any calculated association, that which is imputed into the equation determines that which is outputted from the equation. Any errors in marker position that are put into the QTL analysis can emerge from the other side as mapping errors.

Recently, a new version of the mouse genetic map has been created. This new map is based on a single large mouse cross (24;25), allowing for the correction of order and spacing errors. The genetic positions (i.e., centimorgan or cM) from the new map can also be linked to physical genomic positions (Mb). The new mouse genetic map is publically available online and a web-based tool has been developed to facilitate going from cM to Mb (see Table 1). A small study comparing mapping in the new versus the traditional genetic map estimates that approximately 20% of the published QTLs may have a peak localization error due to either an issue with the relative distance between the markers around the peak or because of the mis-ordering of the markers in the traditional map (24). For newer crosses in which genotyping was done using single nucleotide polymorphisms (SNPs), updating to the new genetic map may provide little benefit. For older data sets, where gene discovery programs have faltered for certain QTLs, it may well be worth re-mapping the QTLs for those crosses using the new map to ensure that the QTLs really are as expected.

The Collaborative Cross

Two mapping strategies have been widely in mice in the past for QTL mapping. In the first, strain A is crossed to strain B (i.e., the F1 generation) and then the F1 mice are bred back to strain A (i.e., backcross mapping). In the second method, the F1 mice are further intercrossed to make the F2 or second filial generation (i.e., intercross mapping). The backcross/intercross methods are not without their limitations and so mouse genetics has been striving for “a better way” to map QTLs. A single two-inbred strain cross does not capture all of the diversity that exists for a given trait in mice. Multiple crosses are required to identify all of the regions of the genome associated with a given phenotype. Furthermore, the power to detect a QTL is a direct function of the size of the cross and thus power becomes an indirect function of the amount of money and time that can be spent making the cross (26). Resolution is also a major issue. The QTLs detected in a standard F2 mapping cross are usually wide, averaging about 30 cM in breadth, and thus the confidence intervals can incorporate hundreds of genes (27). Lastly, there is a finite limit on the number of phenotypes that can be collected from a single mouse and phenotypes can usually only be measured in
Table 1. Useful web-based resources for mouse genetics.

<table>
<thead>
<tr>
<th>Phenotype Databases</th>
<th>Mouse Phenome Database</th>
<th><a href="http://phenome.jax.org">http://phenome.jax.org</a></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Includes an excellent SNP database that incorporates data from a variety of sources (including Perlagen, Celera, Broad, Roche, etc.); many built-in analysis tools</td>
<td></td>
</tr>
<tr>
<td>Europhenome</td>
<td><a href="http://www.europhenome.org/">http://www.europhenome.org/</a></td>
<td></td>
</tr>
<tr>
<td>Interphenome</td>
<td><a href="http://www.interphenome.org/">http://www.interphenome.org/</a></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Includes links to other phenotype databases</td>
<td></td>
</tr>
<tr>
<td>Knockout Mouse Projects</td>
<td>Knockout Mouse Repository (KOMP)</td>
<td><a href="http://www.komp.org/">http://www.komp.org/</a></td>
</tr>
<tr>
<td>European Conditional Mouse Mutagenesis Project</td>
<td><a href="http://www.eucomm.org/">http://www.eucomm.org/</a></td>
<td></td>
</tr>
<tr>
<td>International Gene Trap Consortium</td>
<td><a href="http://www.genetrap.org/">http://www.genetrap.org/</a></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Includes links to other Gene Trap projects</td>
<td></td>
</tr>
<tr>
<td>Expression Databases and eQTL Mapping</td>
<td>Lusis/Drake Lab Website</td>
<td><a href="http://geneqtl.genetics.ucla.edu/">http://geneqtl.genetics.ucla.edu/</a></td>
</tr>
<tr>
<td></td>
<td>eQTL in B6xDBA (dataset is described in (51)); website includes links to other tools and databases</td>
<td></td>
</tr>
<tr>
<td>The GeneNetwork</td>
<td><a href="http://www.genenetwork.org/">http://www.genenetwork.org/</a></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Includes SNP, phenotype and expression data; excellent tutorials; large focus on BXD Recombinant Inbred strains; includes tools for mapping of eQTL, phenotype QTL, phenotype correlation among other tools; links to other databases and tools</td>
<td></td>
</tr>
<tr>
<td>BioGPS</td>
<td><a href="http://biogps.gnf.org">http://biogps.gnf.org</a></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Expression profiles across tissues in rat, mouse and human; expression in selected tissues across a small number of inbred strains (note: these data are not from crosses)</td>
<td></td>
</tr>
<tr>
<td>Haplotype Association Mapping (HAM)</td>
<td>EMMA Webserver</td>
<td><a href="http://whap.cs.ucla.edu/mpad/">http://whap.cs.ucla.edu/mpad/</a></td>
</tr>
<tr>
<td></td>
<td>HAM QTL using the Efficient Mixed Model Association (EMMA) method; uses phenotype data from the Mouse Phenome Database</td>
<td></td>
</tr>
<tr>
<td>EMMA Source Code</td>
<td><a href="http://mouse.cs.ucla.edu/emma/">http://mouse.cs.ucla.edu/emma/</a></td>
<td></td>
</tr>
<tr>
<td>SNPster</td>
<td><a href="http://snpster.gnf.org">http://snpster.gnf.org</a></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Online tool for Haplotype Association Mapping; is an older program and therefore Mb positions for SNPs are out of date</td>
<td></td>
</tr>
<tr>
<td>Other Relevant Databases Including SNP Database</td>
<td>Mouse Genome Informatics</td>
<td><a href="http://www.informatics.jax.org/">http://www.informatics.jax.org/</a></td>
</tr>
<tr>
<td></td>
<td>Mouse HapMap</td>
<td><a href="http://mouse.cs.ucla.edu/mouseHapMap/">http://mouse.cs.ucla.edu/mouseHapMap/</a></td>
</tr>
<tr>
<td></td>
<td>Imputed SNP database</td>
<td></td>
</tr>
<tr>
<td>Oak Ridge National Laboratory Mutant Mouse Database</td>
<td><a href="http://mouse.oml.gov/mmdb/index.html">http://mouse.oml.gov/mmdb/index.html</a></td>
<td></td>
</tr>
<tr>
<td>Center for Genome Dynamics</td>
<td><a href="http://cgd.jax.org/">http://cgd.jax.org/</a></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Includes several databases and tools for mouse genetics; is the current home of the New Mouse Genetic Map and the Centmorgan/Megabase conversion tool; includes raw data sets from microarrays, QTL mapping crosses, etc., as well as an imputed SNP database</td>
<td></td>
</tr>
<tr>
<td>Pathbase</td>
<td><a href="http://www.pathbase.net/">http://www.pathbase.net/</a></td>
<td></td>
</tr>
<tr>
<td>The Genetic Architecture of Complex Traits in Heterogeneous Stock Mice</td>
<td><a href="http://gscan.well.ox.ac.uk/">http://gscan.well.ox.ac.uk/</a></td>
<td></td>
</tr>
</tbody>
</table>
one experimental condition per mouse (i.e., diet, age, or drug treatment) (28).

The Collaborative Cross is an ambitious multi-national project that, when completed, will have generated a set of 1000 Recombinant Inbred (RI) lines derived from 8 strains of inbred mice (A/J, C57BL/6J, 129S1/SvImJ, NOD/LtJ, NZO, CAST/Ei, PWK/Ph, and WSB/Ei) (29-31). These carefully constructed strains have been designed to maximize genetic diversity and to allow for improved mapping of complex traits (32). Using computer simulation, it is predicted that a QTL that accounts for 5% of the phenotypic variance among the founder strains could be mapped with a peak location error (i.e., the difference between the actual and predicted QTL peak) of only 1.8 cM using 200 of the Collaborative Cross RI lines (5 replicates per RI line = 1000 mice). In the same simulation, the peak location error for a similar QTL in 1000 F2-intercross mice would be closer to 3.7 cM (33). Furthermore, the unrecombined haplotype block length is predicted to be much shorter than for F2 mice. This suggests that QTL peaks that are detected will be much sharper when using the Collaborative Cross (33). In mice, there are on average 10 genes per Mb and 2 Mb per cM (34). For the average QTL, there are about 600 genes within the 95% confidence interval. It has been predicted that QTL confidence intervals may be as narrow as 0.1 Mb, or a single gene, when mapped with the full panel of the Collaborative Cross (35), making the Collaborative Cross an extremely exciting prospect for genetic mapping.

The final inbred RI mice from the Collaborative Cross will not be completed until 2012 (29), but data is available from the so-called pre-CC mice (mice from distinct lines that are at the mid-stages of inbreeding). Two bones from the Collaborative Cross are presented in Fig. 1, demonstrating the incredible diversity in size and shape in these mice. The bone on the right has a smaller periosteal circumference than the endosteal circumference of the bone on the left and, theoretically, the midshaft of the right-most bone could fit inside the marrow space of the bone on the left.

Fig. 1. µCT images of the femoral midshaft from two mice from the Collaborative Cross. Both of these samples are from male animals, but from separate RI lines. The periosteal circumferences are 7.35 mm (A) and 4.25 mm (B), respectively. Bones were kindly provided by Dr. E. Chesler and Ms. D. Miller of the Oak Ridge National Laboratory.
Haplotype Association Mapping

Another method developed to map QTLs is called “haplotype association mapping” (HAM) or, as previously called, “in silico” mapping. In a traditional F2 mapping experiment, diversity in genotype is generated from crossing two or more strains but this method does not take advantage of the diversity already present in the inbred strains themselves (28). The common laboratory strains of mice arose for the most part from a single small founder population (36). Analysis of the data from the high-density genotyping projects has demonstrated that because of this breeding bottleneck, the genome structure for inbred strains is organized in discrete haplotype blocks (37-39). HAM takes advantage of this haplotype genome structure to discover regions of association between genotype and phenotype in the inbred strains (40). The input data consists of high-density genotype and phenotype data from a large panel of inbred strains.

HAM presents several key advantages, not least of which is cost. Data from the large-scale phenotyping projects can be used, obviating the need to raise and phenotype one’s own animals (see Table 1). Similarly, large sets of SNP genotyping data are freely available. But this method is not without its detractors (41-43). The most worrisome problem is the high level of false positives (41), leading to the suggestion that HAM QTLs should be considered true only if validated using another mapping method (43-45). This method has also been criticized for having low power (43;46) and unclear confidence intervals, and concerns have been raised about the use of highly related strains (44). No consensus has been reached about how to define the size of a haplotype or the best method of analysis (28;40;45;47). Regardless of these unresolved issues, this method shows promise for mouse bone genetics as Tang and collaborators recently used this method to identify Cer1 as a candidate gene for BMD, demonstrating the utility of HAM in bone genetics (48). HAM can be used in concert with more traditional, low resolution mapping strategies to aid in candidate gene identification.

Expression QTLs

Studies in numerous species have demonstrated that transcript levels of a gene are genetically regulated (49-52), suggesting that QTLs can be mapped for expression levels (53-55). Microarray is a powerful and widely accessible experimental platform that can be used to acquire expression levels for the majority of the genes in the genome. When expression data are obtained by microarray for a population of genetically diverse individuals, it is then possible to map QTLs for transcript levels of each gene (51;56;57). Such a QTL is called an “expression QTL” or eQTL (specifically reviewed in (58;59)). Expression QTLs can be divided into two types, “cis” acting and “trans” acting (60). Cis, local or proximal eQTLs are defined as QTLs mapped to or extremely near to the gene itself while trans or distant eQTLs are those mapping to anywhere but the location of the gene itself (59).

The experimental design for mapping eQTLs is fairly simple. First, a population of genetically diverse mice is created, such as F2 intercross mice or an out-bred heterogeneous stock or a population of “ready made” genetically diverse mice such as the BXD or BXH RI sets is utilized. Then, tissues from all of the mice are collected, the animals are genotyped, expression by microarray is assessed, and the eQTLs for each gene are mapped (59;61). It is this last analysis step that becomes a little more complicated (62). The obvious stumbling block for this type of experiment for most independent investigators is the cost involved (58). Some studies have suggested that eQTLs with the highest heritability are cis acting (60). Furthermore, cis-acting eQTLs appear consistent across tissues, meaning that if a cis-eQTL is mapped in one tissue in the population, there is high probability that the gene is cis-regulated in other tissues (60;63). This is good news for bone biologists, as to date, there have been no studies published in which eQTLs were mapped using bone-derived tissues. Farber
and colleagues have shown that eQTLs mapped in adipose tissue can be used to narrow traditional bone density QTLs (64). Data sets and analysis tools for eQTLs are now freely available on the web (see Table 1).

**Conclusion**

The hard work and vision of many individuals has placed the future of mouse genetics for osteoporosis research in a new space. An abundance of data, tools and resources is now freely available online. A short list of some of these resources is presented in Table 1. The phenotype databases are expanding frequently. Mouse SNP databases already exist that include high-density genotyping data for selected strains and imputed genotype data for many more strains. In the near future, projects such as the “Mouse Diversity Array” from the Center for Genome Dynamics (http://cgd.jax.org/tools/diversityarray.shtml) will provide a wealth of de novo data including information such as copy number variation. The Collaborative Cross should have the mapping resolution to move the field from Quantitative Trait Loci to Quantitative Trait Genes. The Collaborative Cross will allow for a variety of studies such as environmental perturbation testing on a stable genetic platform, but this set of mice will not be completed until 2012. Commercially available outbred stocks are available now and have been used successfully to map QTLs with high resolution (64;65). The International Mouse Knockout Consortium is planning to make available knockout mice for every protein-coding gene by 2010 (66;67), creating the ability to test candidate genes when they have been identified (35). In sum, the utility of mouse genetics has expanded exponentially with new tools and new databases. These instruments will accelerate gene discovery and the field of functional genomics.

**Conflict of Interest:** None reported.

**Peer Review:** This article has been peer-reviewed.

**References**


and identification of Trps1 as a probable candidate gene. J Bone Miner Res. 2008 Sep;23(9):1529-37.


Copyright 2009 International Bone & Mineral Society


