Endocrine, *vol. 17, no. 1, 67–75, February 2002* 0969–711X/02/17:67–75/$12.25

A Strategy for Identifying Osteoporosis Risk Genes

David Rowe and Alexander Lichtler

Department of Genetics and Developmental Biology, University of Connecticut Health Center, Farmington, CT

The genetic factors that contribute to the susceptibility of osteoporosis are likely to be extremely heterogeneous, reflecting the complex genetic program that controls the cellular aspects of bone formation and remodeling. Because it is unlikely that current genetic methods will have the power to resolve multiple subtle genetic effects from the environmental contribution, biologic approaches will have to be developed to identify mechanistically similar forms of osteoporosis prior to applying more sophisticated molecular studies. This perspective article proposes a strategy that focuses on those subjects with an inherent impairment of bone formation as the primary cause of diminished bone mass. It views the impairment as a defect within the osteoprogenitor lineage and identifies the stage within the lineage where the progression digresses from normal. It is at the stage of diversion in the lineage that affected cells are isolated for extensive microarray studies with the intent to identify molecular pathways that are underperforming. The technological steps that have to be accomplished for this strategy to be successful rely heavily on promoter green fluorescent protein transgenes that can assess the extent and tempo of lineage progression in primary cells and intact bone, and that can allow for a relatively small subpopulation of cells from a primary bone cell culture to be isolated for molecular analysis. Initially, the strategy has to be validated in murine models with single gene defects affecting the performance of the osteoprogenitor lineage. The experience gained from murine models will allow a similar approach to be applied to humans.

Key Words: Osteoporosis; osteoprogenitor lineage; green fluorescent protein; microarray.

Introduction

There is increasing appreciation that genetic factors play a major role in determining adult bone mass *(1,2)*. Population studies identify a relatively higher risk of osteoporosis in individuals with a northern European and Asian ancestry relative to individuals with an African or Mediterranean background. Family studies have documented mother-daughter pairs with osteoporosis *(3)*, and twin studies also show the importance of genetic determinants of bone mass *(4)*. Inbred mouse strains demonstrate a profound genetic impact on bone mass *(5,6)*. Despite these genetic studies, little progress has been made in identifying genes that are necessary for maintaining normal bone mass, although significant progress has been made in identifying genes that are essential for limb patterning, growth plate regulation, bone cell differentiation, and the structural macromolecules that are essential to cartilage and bone function. This article develops an idealized strategy for identifying genes that are necessary for maintaining skeleton mass—osteoporosis risk genes.

Polygenic Nature of Osteoporosis

Gene knockout studies in mice that produce skeletal defects *(7)* point to transcription factors *(8)*, growth factors *(9)*, signal transduction pathways *(10,11)*, and extracellular matrix molecules that are essential for the patterning of the skeleton, the formation of joints *(12)*, the invasion of osteoblasts *(13)*, the coordinated growth of the metaphysis and cortical bone *(14)*, the production of an osteoid matrix that will mineralize *(15)*, and an ability to remodel existing bone *(16,17)*. With a few exceptions, these traits are expressed in homozygous knockout mice as a severe phenotype such that the mutation has to be maintained in heterozygous mice. In most cases, the carrier mice appear to be normal and do not have obvious bone disease. However, the molecular pathways that are used during embryogenesis to build the initial skeleton are likely to be used again in the acquisition of peak bone mass in the peripubertal years, and in the maintenance of bone mass during adulthood and will be challenged further to meet the increased requirement for bone formation after menopause. If these pathways contribute to skeletal biology in the postnatal period, it would not be surprising that a relative insufficiency rather than a complete knockout of its function could contribute to a less than optimal skeleton in adulthood. Thus, haploid insufficiency (complete loss of one allele), partial insufficiency of one allele, partial insufficiency of both alleles (double heterozygosity), partial insufficiency of multiple genes in the same molecular pathway, or partial insufficiency of genes in different molecular pathways are all potential combinations that could contribute to the inability to maintain a sufficient

Color images of Figs. 1–5 can be downloaded into Photoshop from a web browser at ftp://genetics.uchc.edu/download/rowe lab/endocrine figs

Author to whom all correspondence and reprint requests should be addressed: Dr. David Rowe, Department of Genetics and Developmental Biology, University of Connecticut Health Center, Farmington, CT 06032. E-mail: rowe@neuron.uchc.edu

bone mass. This complexity in molecular pathogenesis will make it difficult to use standard genetic methods to identify osteoporosis risk genes because of the difficulty in identifying a reasonably homogeneous study group.

When a single mutant allele of a gene acts through either a dominant negative or a strong haploid insufficiency mechanism, traditional Mendelian genetics can be used to identify the underlying genetic unit through standard linkage or candidate gene analysis. Examples include osteogenesis imperfecta and vitamin D–resistant rickets. However, in most cases of low bone mass, pedigree analysis is not possible, and, instead, populations of affected individuals are studied by associating a polymorphism within a candidate gene or polymorphismic markers distributed throughout the genome. Although promising studies of the vitamin D receptor *(18)*, estrogen receptor *(19)*, *COL1A1* gene *(20, 21)*, and other genes known to have an important role in bone biology *(22–24)* have shown an increased risk for bone disease with certain genotypes, the association is often weak and not consistent from study to study. Single nucleotide polymorphism (SNP) is a more sensitive approach to associate statistically an allele of a candidate gene to a polygenic disease process *(25)*. This method has the advantage of automated throughput and more sensitive discrimination because it does not depend on a polymorphism altering a restriction site. SNP analysis is receiving broad application through programs and databases initiated by the National Institutes of Health, but the value of this method for studying bone diseases has yet to be demonstrated *(26,27)*.

Quantitative trait analysis is a directed search for candidate genes using polymorphisms distributed throughout the genome *(28)*. It points to regions within a specific chromosome where genetic units exists that are associated with a phenotype. With full definition of the mouse and human genome, this approach will be extremely powerful for identifying genes important to bone biology. However, this approach is best used in a homogeneous genetic background such as the mouse *(6,29)* because its power becomes diluted in the heterogeneous human population *(30)*.

While a purely genetic approach to identify osteoporosis risk genes is currently the most practical course *(31)*, this strategy may or may not improve our understanding of how a particular genetic unit operates within a tissue as complex as bone. Bone biologists need to develop model systems that can be used in concert with genetic approaches to define function of genes associated with osteoporosis in the context of concepts of bone formation and remodeling. Here the focus will be on bone formation, but a similar model will be required for other aspects of bone biology.

Osteoprogenitor Lineage

The model for understanding the role of osteoporosis risk genes envisions that the osteoprogenitor lineage is continually called on throughout the life of the individual to generate new packets of robust matrix-forming osteoblastic cells *(32,33)*. Initially, the early progenitor cells proliferate to produce a wave of cellular expansion that is necessary to provide sufficient numbers of cells for subsequent differentiation. As this population of cells becomes more restricted to the osteoblast lineage, expression of alkaline phosphatase (AP), type 1 collagen, and parathyroid hormone (PTH) receptor serves to mark these cells as preosteoblasts followed by early osteoblast expressing sialoprotein (BSP). At this stage, the rate of cell division has dropped sharply while the production of extracellular matrix molecules accelerates. With further maturation, the cells become mature osteoblasts and nondividing osteocytes located either as lining cells or within the mineralized bone matrix, some of which can be identified by their production of osteocalcin.

Conceptually, it is convenient to divide the lineage into a proliferating noncommitted phase of cell expansion followed by a commitment and maturing phase of cell differentiation *(34)* (Fig. 1). The cells within these two phases of the lineage have distinctly different cellular properties for which different molecular pathways will play essential controlling roles. Mutations in genes within these molecular pathways will affect the intrinsic ability of the lineage to develop a sufficient number of progenitor cells, or will either impede or accelerate the ability of these cells to acquire full osteoblastic differentiation. In either case, the lineage will be compromised in its ability to provide sufficient numbers of fully differentiated osteoblastic cells resulting in a common phenotype of diminished bone mass. When the genetic deficiency is severe, as in a homozygous knockout mouse, developmental abnormalities are observed because insufficient numbers of differentiated osteoblasts are formed to meet the demand during organogenesis (Fig. 1, row B). When the deficiency in gene function is mild and leads to a subtle alteration in lineage progression, as in a heterozygous knockout, early embryonic development and postnatal bone growth are normal probably because the robustness of cell proliferation and differentiation during this period of life can compensate for mild inadequacies within the lineage (Fig. 1, row C). However, with advancing age, irrespective of any underlying genetic abnormalities, the robustness in proliferation and maturation to full osteoblast is diminished *(35,36)*. In addition, environmental factors such as smoking and physical inactivity contribute to the decreased effectiveness of the lineage to generate mature osteoblastic cells (Fig. 1, row D). With the loss of sex hormones, the lineage is called on to increase its production of mature osteoblasts. It is at this time that a mild genetic deficiency producing intrinsic defects controlling lineage progression combines with environmental factors to act as an osteoporosis risk gene (Fig. 1, row E).

The problem that faces the bone biologist is how to determine where these genetic factors affecting bone mass in mice act on the bone lineage in human subjects. Histologic analysis of bone biopsies can give an impression of cellular

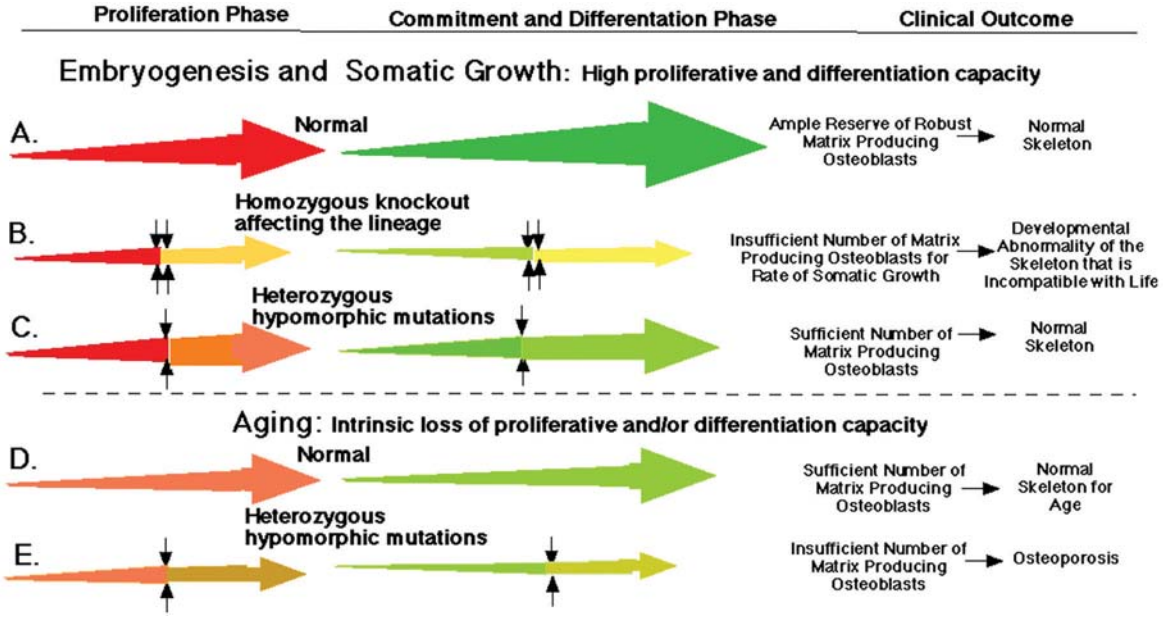

Fig. 1. View of the osteoprogenitor lineage that recognizes two distinct phases (proliferation and commitment/differentiation) that culmulate in formation of mature osteoblastic cells. The inherent tempo and magnitude of each phase is lost with advancing age and can be further compromised by hypomorphic mutation in genes that control each phase. Homozygous mutations have an embryonic phenotype because the lineage cannot keep up with the rapid phase of somatic growth while a heterozygous mutation is usually scored as normal. However, with advancing age, the heterozygous mutation may sufficiently impair the ability of the lineage to maintain bone mass, particularly when demand for increased bone formation associated with estrogen deficiency develops. Each scenario (A–E) is developed in the text.

activity but lacks the discrimination of molecular studies. Microarray analysis of RNA from bone biopsy samples will be difficult to interpret because of the heterogeneity of the cell population within the sample and the inability to appreciate how the lineage is performing in a static sample. Instead, osteoprogenitor cells need to be removed from the affected individual and induced to recapitulate the lineage under standard experimental conditions so that the proliferative and differentiating phases of osteoblast generation can be assessed at the cellular and molecular levels. Initially, development of a murine model system is needed to take advantage of the power of the growing number of gene knockout and chemical mutagenized mice *(37)* that have a skeletal phenotype. The mice will provide the platform to appreciate the cellular and molecular elements of the osteoprogenitor lineage. Once these pathways are understood in the mouse, then applying this strategy to humans becomes a more feasible undertaking.

Analysis of Murine Osteoprogenitor Pathway In Vitro

Primary mouse cell cultures derived from either marrow stromal cells (MSC) or the neonatal calvarial osteoblasts (mCOB) are preferable sources of cells for lineage analysis to immortalized bone cell lines because primary cells have greater phenotypic stability and more easily recognized stages of cell differentiation. In addition, producing primary cultures from mice with defined gene knockouts provides direct linkage between the effect of the mutation on lineage progression in vitro and the consequences of the mutation on intact bone. MSC- and mCOB-derived cultures are distinctly different, each with advantages and disadvantages. The cells that initiate an mCOB culture contain a mixture of osteoblasts, preosteoblasts, and progenitor cells. Within 2 to 3 d after initiating the culture, the majority of cells lose their osteoblastic markers and enter a proliferation phase that continues until confluency is reached. As nodules develop, markers of bone differentiation reappear in a consistent and reproducible matter. What is not clear in this model is the source of the cells that eventually give rise to the bone nodules. Do they arise from early progenitor cells that progress unidirectionally forward or does a more mature cell dedifferentiate, proliferate, and then redifferentiate? Possibly both routes are utilized but the molecular pathways controlling the two processes might be different. Less confusing for lineage studies are MSC cultures in which extremely early progenitor cells initiate the culture and progress forward to produce nodules, some of which become osteoblasts and some of which mineralize. The disadvantage of the MSC culture is its heterogeneous composition of nonosteoblastic nodules and cells that enter other lineages. In either model, the cell population is heterogeneous with cells at different levels of osteoblastic differentiation or within nonosteoblastic lineages.

Molecular analysis of the osteoblast lineage will require methods to produce a relatively homogeneous subpopulation of cells derived from defined but different stages of lineage progression. Antibodies directed to cell-surface proteins have the potential of identifying cells in the early lineage before confluency, and a dense extracellular matrix requires proteolytic steps for cell isolation *(38–40)*. To date, identification of subpopulations of cells by their cell-surface epitopes has been most successfully applied to freshly isolated bone marrow cells rather than to cells that have been maintained in culture. To overcome these cell identification problems, we have chosen to develop transgenes expressing various colors of green fluorescent protein (GFP) that activate at different stages within the lineage. This approach has the advantages of observing lineage progression in real time within the culture system, of isolating intact cells for subsequent culture studies, and of correlating expression of the GFP marker in vitro and in intact bone.

The first step in demonstrating the usefulness of a particular promoter-GFP transgene requires colocalization of GFP expression with endogenous markers of bone cell maturation. Figure 2 depicts our view of lineage progression using endogenous molecular and cellular markers that most investigators would accept as defining a specific stage of cell differentiation *(41)*. The earliest identifiable population is the proliferating multipotential progenitor cell that has not yet become AP positive and does not express readily detectable levels of type I collagen mRNA. These fibroblast-like cells express ALCAM on their cell surface *(42)*, the twist *(43,44)* or Oct4 *(45)* transcription factor, fibroblast growth factor receptors *(46,47)*, and probably fibronectin. The next recognizable stage is the expression of AP within clusters of cells that are surrounded by AP-negative cells. These cells initiate type I collagen synthesis *(41)*, express the PTH receptor *(48)*, and are designated preosteoblastic cells. With time, these clusters become multilayered and the more centrally placed cells become cuboidal in shape and express bone sialoprotein. We have defined cells at this level of development as early osteoblasts. The level of type I collagen synthesis within the preosteoblast stage becomes more prominent in the early osteoblastic cells. The development of a mineralized matrix that encompasses some of the osteoblastic cells is the last visible stage of differentiation. It is at this time that osteocalcin expression is detected, which identifies a mature osteoblast or osteocytes.

Using these endogenous phenotypic markers, we have developed three different Col1a1 promoter fragments that activate at uniquely distinct stages in the lineage *(49)* (Fig. 3). pOBCol3.6GFP is first expressed in the subclass of cells that are AP positive. The transgene continues to remain active as the culture matures, ultimately developing two levels of fluorescent intensity. The low-intensity cells are fibroblastic in appearance and surround the nodule, while the high-intensity cells are located within the nodule and are cuboidal in shape. In intact bone, the transgene identifies elongated cells with low fluorescent intensity in the periosteum as well as more strongly fluorescent cells lining the bone surface. We interpret the period of time of transgene expression in primary culture between early cluster formation and bone cell differentiation as the time when transgene expression identifies preosteoblastic cells. Col2.3 GFP activates concomitantly with bone sialoprotein expression, and it is restricted to the cuboidal cells within the early bone nodule. Within intact bone, Col2.3GFP is not observed in the periosteum but is detected primarily on trabecular and cortical bone surface with some extension within the bone matrix. This marker appears to be sensitive to the transition between a preosteoblast and an osteoblast. The third marker is pOB2.3GFP, which differs from the previous construct by the presence of the collagen's first intron. In culture, this marker activates later than Col2.3GFP and is restricted to cells that are within the mineralized region of the nodule. It becomes active somewhat before the culture acquires osteocalcin expression. In intact bone, it identifies cells lining the bone surface and, to a greater extent than Col2.3GFP, is positive in most of the osteocytes within the bone matrix. We interpret this marker to identify mature osteoblasts and osteocytes. Currently, we are evaluating the efficacy of the fibronectin and osteocalcin promoters as markers that will flank the spectrum of differentiation identified by three Col1a1-derived transgenes.

The power of GFP is the ability to obtain relatively homogeneous subpopulations of cells within the osteoprogenitor lineage using fluorescent activated cell sorting (FACS). For example, approx 50% of the freshly isolated cells from neonatal calvaria contain pOBCol3.6-positive cells while 20% of cells are pOBCol2.3GFP positive. During the first 3–6 d of calvarial culture, GFP expression is gradually lost. As the cells become AP positive, pOBCol3.6GFP can account for 50% of the total cell population while the other transgenes are inactive. Thus, it is during this time that pOBCol3.6 GFP can be used to isolate cells at the preosteoblast stage of differentiation. Col2.3GFP becomes positive 3 to 4 d after pOBCol3.6GFP and can account for approx 20% of the total cell population. Therefore, between the time of activation of Col2.3GFP and pOBCol2.3GFP (approx 3 to 4 d later), a relatively homogeneous population of early osteoblasts can be isolated. Finally, cells that expressed pOBCol2.3 GFP account for 10% of the total cell population and their direct isolation yields a population of very mature osteoblast and osteocytes. Essentially none of the MSC cells that initiate the culture are GFP positive, but as the culture matures, it develops the same pattern and tempo of transgene expression as observed for the mCOB culture.

Our goal is to develop multicolor GFP transgenes that can be distinguished visually and isolated by FACS so that multiple levels of differentiation within the same culture can be subjected to analysis. Using combinatorial color criteria, preosteoblasts would be identified by the presence of color expressed by pOBCol3.6GFP and the absence of

color expressed by Col2.3GFP. Early osteoblastic cells would be identified by the presence of color expressed from Col2.3GFP and the absence the color from pOBCol2.3GFP, while mature osteoblast would be identified by the color expressed from pOBCol2.3GFP. Thus, by producing transgenic mice that are homozygous for various promoter fragments and color versions of GFP, mice harboring multiple color transgenes can be produced from which comprehensive lineage studies can be performed. Our intent is to make these mice available to bone cell investigators as a common platform reagent to facilitate standardized lineage studies.

Identifying Molecular Pathways Controlling Osteoblast Lineage Progression in Mice

Microarray analysis of RNA expression profiles provides the most comprehensive approach to understanding genetic interrelationships in a complex cellular system. The more homogeneous and reductionist the model system, the easier the interpretation of such an analysis. The primary culture system that utilizes GFP markers to isolate subpopulations of cells within the lineages provides a platform to assess the expression profile at specific levels of cell differentiation. From this analysis, it is likely that members of well-defined molecular pathways that have already been defined in simple organisms will be identified, which will aid in understanding its role within the osteoprogenitor pathway.

However, the wide variability in hybridization intensity ratios reported in cDNA microarray studies needs to be controlled to discriminate experimental noise from a meaningful difference in a hybridization signal between control and test samples. Two preliminary steps need to be implemented to overcome some of the uncontrolled variability in a microarray study. The first is the genetic background of the mice used in the study. Using inbred lines of mice has the advantage of a uniform genetic background, but the homozygous nature of the line can emphasize private variations in bone biology and expression that would never be observed in an outbred line *(50)*. While outbred mice will not be subject to strain-specific differences, their nonuniform genetic background will produce unnecessary variability in an expression study that will become particularly problematic when the study requires detecting quantitative differences *(51)*. To overcome this problem, we have chosen to perform our studies in the F1 offsprings of two inbred mice, C57Bl/6 and C3H. These inbred mice have been used for studying the influence of genetic background on bone mass *(6,52)* and MSC cultures *(53,54)* and serve as models for the low and high bone mass, respectively. The F1 offsprings will be uniformly heterozygous at every loci, which should produce a tight phenotype and one that most closely resembles an outbred mouse line. The second requirement is the production of a hybridization slide that is optimized for quantitative measurement of mRNAs present in cell extract *(55, 56)*. Current commercial or public domain expression libraries did not use sources of tissue highly enriched for bone, cartilage, or other connective tissues and thus probably lack many of the genes that might be of interest to the bone biologist. The use of large arrays from diverse tissues is extremely expensive and precludes the use of repeated measures to assess experimental variation.

To overcome these problems, we are developing a connective tissue–focused gene collection from a variety of library sources. To be included in the gene collection, baseline studies will be required to demonstrate that the target gene produces a "yellow" hybridization signal of sufficient strength from a variety of connective tissue extracts that have been divided, labeled separately with cy3 and cy5, and recombined prior to the hybridization step. This step will remove targets with affinities for a particular fluorochrome and segregate targets based on their hybridization intensity, two steps that contribute to variability in a cDNA expression study. In addition, slides will be configured with repetitive gene arrays on multiple slides to allow calculation of the experimental variability of a particular target.

These extensive preliminary steps will offer the best opportunity to measure quantitative changes in expression of genetic pathways in the isolated subpopulation of cells during lineage progression. However, to prove that an observed pathway is playing a deterministic role in lineage progression, it will be necessary to inactivate that gene and observe the consequence of this mutation first on the cellular phenotype and then on the expression profile at each level of cellular differentiation. The source of the cells with a specific inactivation of an endogenous gene will come from existing gene knockout mice. In many cases, gene knockouts exert their phenotype in the homozygous state and heterozygous mice are maintained as normal healthy breeders. However, there is increasing evidence that many of the heterozygous mice in fact have subtle defects that become evident as the animals age *(57)*. If the goal is to identify genes that contribute to osteoporosis in humans, then older heterozygous breeder mice with evidence of diminished bone mass should be a rich source of candidate genes affecting the osteoprogenitor lineage sufficient to compromise maintaining normal bone mass with aging.

Using this rationale, heterozygous breeder mice can be crossed with homozygous ColGFP mice to generate litters that all contain the marker gene and half of which carry one dose of the knockout gene. In many cases, the mice are maintained in a partial or complete C57Bl/6 background. To produce a more stable genetic background, the breeders will be crossed with C3H mice homozygous for the GFP transgene to produce a uniform and heterozyous genetic background. When cultures are developed from these mice, the inherent ability of the cells to progress down the lineage can be compared based on standard markers plus expression of GFP. The prediction is that the cultures from affected mice will progress down the lineage normally until a point is reached that is dependent on the defective gene for further

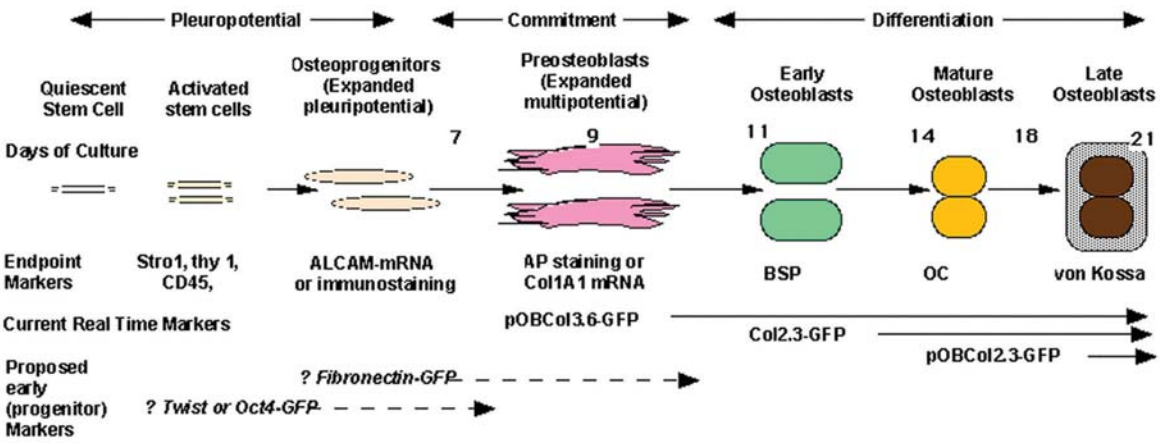

Fig. 2. Working model of osteoprogenitor lineage using histologic, immunologic, and molecular markers. The tempo for progression to full osteoblast differentiation varies with the source of stem cell and culture conditions as well as the genetic backgroud of the mouse used to establish the culture. The promoter GFP transgenes given in italics have yet to be shown to be markers for the earliest stages of the lineage. OC, osteocalcin; BSP, bone sialoprotein.

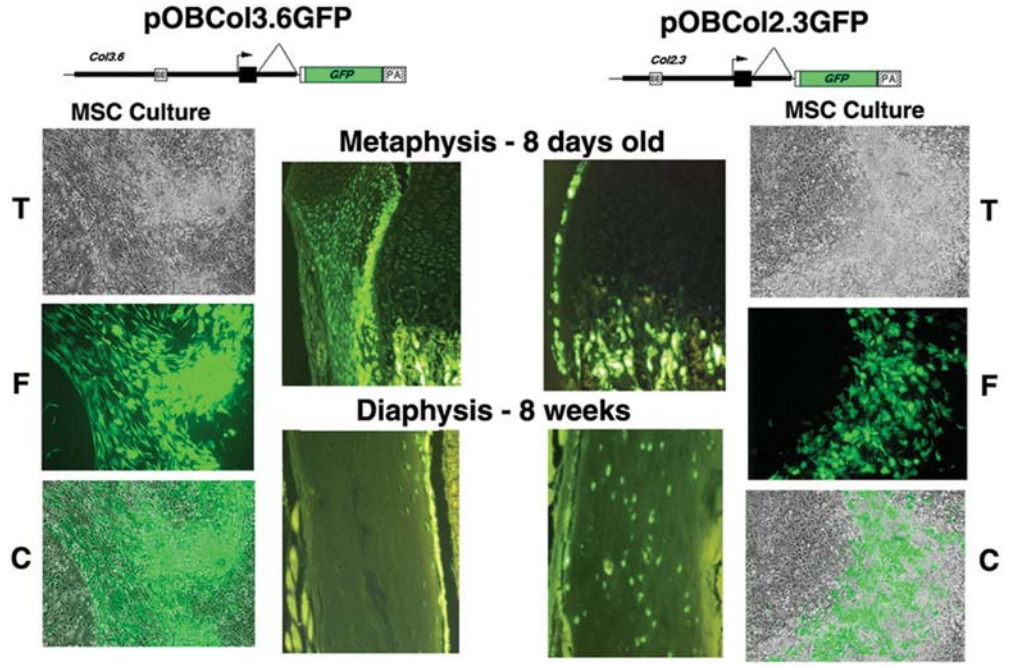

Fig. 3. Expression of collagen-GFP transgenes in primary cell culture and intact mouse bone. The images illustrate the expression of the two most distinctive promoter-GFP transcripts. pOBCol3.6GFP activates at the preosteoblastic stage of the lineage and intensifies in expression as the cells become osteoblastic, while pOBCol2.3GFP activates at full osteoblast differentiation. The relative difference in intensity of expression of pOBCol3.6GFP is illustrated in the MSC culture and in the cells within the collar region of the metaphysis. The uniformity of intensity and restriction of expression of pOBCol2.3GFP to differentiated osteoblasts/osteocytes is seen in the MSC culture (cell restricted to the mineralized nodule) and in the cortical bone within the diaphysis. T, transmitted light; F, fluorescent light; C, composit of the transmitted and fluorescent image.

progression (see Fig. 4B). At that point, subsequent differentiation would be delayed (or accelerated). Microarray analysis will be most informative in the population of cells that first express the delay in progression and to a lesser extent the cells immediately downstream in the differentiation pathway. The analysis should indicate when the hypomorphic gene exerts its effect within the lineage and point to potential candidate genes that are downstream in a potential molecular pathway. Genetic manipulations such as forced expression of a downstream gene may show that the cellular phenotype can be rescued, which will further define the importance of the pathway.

The strength of this analysis is to view lineage progression from a vantage point of tempo, cell number, and stage of cellular differentiation, all of which determine the intrinsic ability of the lineage to produce ample numbers of mature bone cells. It is likely that distinctly different molecular pathways control all three aspects of the lineage. By exam-

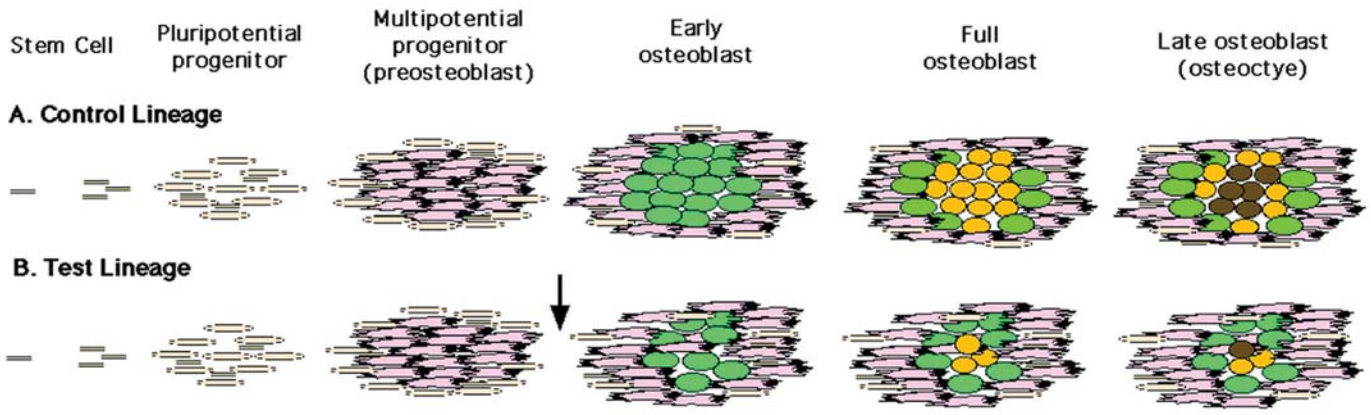

Fig. 4. Identifying cells at defined levels of differentiation within osteoprogenitor lineage using multiplexing of promoter GFP constructs expressing different colors of GFP. Those colors that activate at a particular stage of the lineage while eliminating the color that activates at an earlier stage and continues into later stages are shown. The tempo of cell differentiation of (**A**) a control lineage and (**B**) a test lineage are contrasted. The rate of progression in tempo and number are identical until the effect of the mutation (arrow) acts on further lineage progression. Analysis of cells at this stage of differention will provide the best hope for understanding the role of the mutant gene on lineage progression.

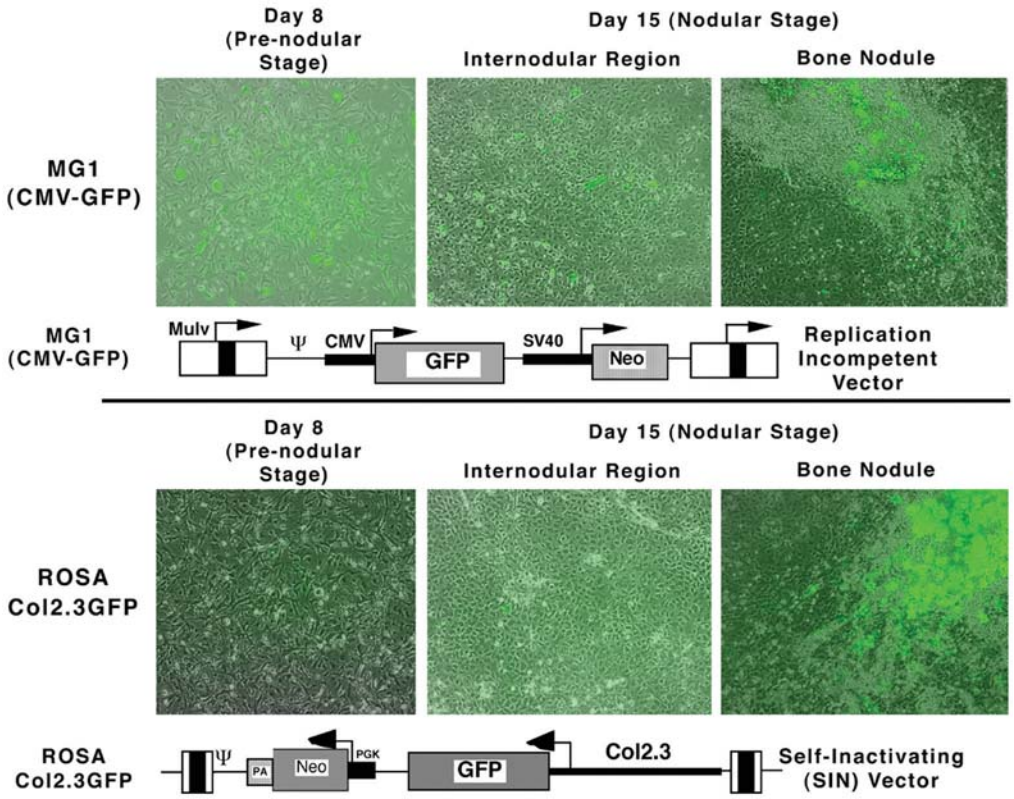

Fig. 5. Comparison of expression pattern of GFP driven from constitutive promoter within replication-incompetent retrovector and Col2.3GFP construct within SIN retrovector in primary murine MSC culture. In the MG1 retrovector construct, GFP is expressed from the cytomegalovirus (CMV) promoter and may also be influenced by the upstream promoter from the Mulv LTR. It generates GFP-positive cells throughout the MSC culture irrespective of the state of cellular differentiation. By contrast, the ROSA retrovector construct, which lacks promoter elements within the LTRs, allows the Col2.3 promoter to express GFP exclusively in the mineralized nodules. Both vectors were pseudotyped in a VSV packaging cell line.

ining a sufficient number of knockout mice, it will be possible to catalog the molecular pathways for each stage within the lineage. It would not be possible to do this analysis by examining the total cell population from a primary culture or from intact bone because of the heterogeneous nature of the tissue. That type of analysis will primarily reflect the failure of the lineage to generate mature cells and will obscure the crucial rate-limiting phase and immediate down-

stream members of a molecular pathway that affects lineage progression.

Application of Lineage/Microarray Strategy to Humans with Osteoporosis

It is the heterogeneity of osteoporosis, including genetic, hormonal, and behavioral factors, that confounds the detection of genes contributing to this bone disease. One step toward reducing this heterogeneity is a strategy to enrich for individuals with an intrinsic failure to produce sufficient numbers of mature osteoblast by removing those in whom hormonal/environmental factors are the major contributor to osteoporosis. In other words, we need to focus on individuals likely to harbor gene defects that act in the cell autonomous manner to control lineage proliferation and differentiation. The types of problems that would act in a cell autonomous manner include a reduced ability to proliferate sufficient numbers of precursor cells, an alteration (either prematurely or delayed) in the rate that these cells acquire a fully differentiated phenotype, or the premature apoptotic loss of cells in either phase of the lineage. Advancing age will only contribute further to an intrinsic problem of cell proliferation and differentiation, underlying the importance of controlling for age. Thus, the causes for the intrinsic defect within the osteoprogenitor lineage will be complex and call for improved clinical characterization to distinguish family pedigrees with osteoporosis likely to have a genetic contribution.

A potential test to distinguish individuals with a cell autonomous defect from those with a cell nonautonomous defect is to induce osteoprogenitor cells derived from the affected individual to form bone in a standardized environment. Heterotopic bone formation appears to provide an environment conducive to bone formation, and intrinsic defects within bone formation are reproduced in this model system *(58,59)*. In this test, fibroblastic cells derived from MSC cultures are embedded in a three-dimensional matrix and implanted in the subcutaneous space of an immunotolerant mouse. Over the course of 2–4 wk, a small bone nodule develops in which the transplanted cells make bone while the marrow that invests it is derived from the host. Thus, the outcome of an initial screening study could stratify individuals whose MSC-derived cells either do or do not yield a heterotopic bone at a particular threshold. In reality, this is a clinical study that will need extensive validation, but such an analysis, when coupled with a thorough clinical evaluation, might well provide a measure of stratification of osteoporosis that could be useful for clinical classification and drug selection.

Once a pedigree has been identified that demonstrates a cell intrinsic defect in bone formation, a more structured assessment of in vitro lineage progression can be initiated that resembles the strategy developed for the mouse. An essential reagent for the analysis is self-inactivating (SIN) retrovectors that can deliver ColGFP constructs in a way that will allow the marker genes to respond to endogenous cellular signals (Fig. 5). While standard retrovectors can efficiently transduce genes into a wide variety of mammalian cells, the promoter/enhancer elements within the long terminal repeat (LTR) elements of the virus overpower an included mammalian promoter. SIN retrovectors are a transducing agent in which the promoter/enhancer elements have been removed so that the included mammalian promoter transgene is free to respond to endogenous transcriptional signals *(60)*. When the SIN retrovectors are produced in a vesicular stromatitis virus (VSV) packaging cell line, extremely high-titer vectors are produced that are capable of infecting the majority of the cells in a culture without the need for antibiotic selection *(61)*. These vectors can deliver the Col-driven GFP constructs that activate in a temporal and spatial manner similar to their expression patterns in primary MSC cultures derived from transgenic mice. Although validation of these virus-delivered marker transgenes for lineage identification still needs rigorous documentation, it should be possible to build reagent retrovectors containing different promoter fragments and colors of GFP. With these reagent vectors, the intrinsic capacity of the osteoprogenitor pathway of MSC-derived cells from individual patients can be assessed in intact bone using the heterotopic bone assay and in real time in primary MSC cultures from which a subpopulation of cells can be isolated for subsequent microarray analysis.

Despite all these steps, it is likely that interpretation of data from a microarray study of lineage progression will be difficult. Unlike the mice in which all of the genes are normal except for the one that is hypomorphic, osteoporotic humans are likely to have a number of underperforming genes and partially dampened molecular pathways. Some will be identifiable from a pedigree analysis, and some may be a consequence of aging or other extrinsically dependent factor. However, the mouse data will have mapped a number of molecular pathways essential to the stages of proliferation and/or differentiation that should be directly transferable to studies of lineage performance and gene expression in humans. From the mouse studies, informatic strategies for pathway analysis will be developed to appreciate essential pathways for lineage progression, and these informatic strategies will be used to interpret the data derived from humans. The prediction is that patterns of expression will point to certain underperforming molecular pathways that are contributing to a diminished lineage proliferation or differentiation such that candidate genes within that pathway can be specifically examined for mutations that account for a lower level of gene activity. In fact, it would not be surprising that more than one gene within one pathway or more than one gene in two or more pathways could have an impact on the lineage. Proof that a particular candidate gene is con-

tributing to this problem could be developed by transduction with a replacement gene construct to determine whether lineage performance improved either in culture or in the heterotopic bone assay and by clinical linkage studies.

Clearly this is not a strategy that is going to be fully realized in the immediate future. It provides a general outline that certainly will need to be modified as experience is gained. The essential elements of this strategy are the promoter GFP constructs in mice and the retrovectors to deliver these constructs to human cells so as to assess lineage potential both in culture and in heterotopic bone implants. Also necessary, but not fully discussed here, is the need for gene array sets that are optimized for studying the osteoprogenitor lineage in both mice and humans. The transgenic, retrovector, and gene array sets are reagents that must be available to basic and clinical investigators as a way to standardize their application to murine models and human subjects with bone disease. However, this strategy will enable a more rational and informative evaluation of patients with bone disease. Identifying genetic risk factors will focus behavioral and therapeutic interventions on individuals prior to the onset of bone disease as well as aid the choice of drugs tailored to a specific pathophysiologic cause for osteoporosis.

Acknowledgment

This work was supported by grants AR43457, AR44545, and PO1-38933 from the US Public Health Service.

References

1. Bansal, A., et al. (2000). *J. Clin. Densitom.* **3(4),** 333–338.
2. Cardon, L. R., et al. (2000). *J. Bone Miner. Res.* **15(6),** 1132–1137.
3. Jones, G. and Nguyen, T. V. (2000). *J. Bone Miner. Res.* **15(10),** 1998–2004.
4. Stewart, T. L. and Ralston, S. H. (2000). *J. Endocrinol.* **166(2),** 235–245.
5. Beamer, W. G., et al. (1996). *Bone* **18(5),** 397–403.
6. Beamer, W. G., et al. (2001). *J. Bone Miner. Res.* **16(7),** 1195–1206.
7. Gunther, T. and Schinke, T. (2000). *Trends Endocrinol. Metab.* **11(5),** 189–193.
8. Yamaguchi, T. P., et al. (1999). *Genes Dev.* **13(24),** 3185–3190.
9. Sims, N. A., et al. (2000). *J. Clin. Invest.* **106(9),** 1095–1103.
10. Yamaguchi, T. P., et al. (1999). *Development* **126(6),** 1211–1223.
11. Milenkovic, L., et al. (1999). *Development* **126(20),** 4431–4440.
12. Gong, Y., et al. (1999). *Nat. Genet.* **21(3),** 302–304.
13. Takeda, S., et al. (2001). *Genes Dev.* **15(4),** 467–481.
14. Chung, U. I., et al. (2001). *J. Clin. Invest.* **107(3),** 295–304.
15. Wennberg, C., et al. (2000). *J. Bone Miner. Res.* **15(10),** 1879–1888.
16. Boskey, A. L., et al. (1998). *Bone* **23(3),** 187–196.
17. Delany, A. M., et al. (2000). *J. Clin. Invest.* **105(7),** 915–923.
18. Obermayer-Pietsch, B. M., et al. (2000). *J. Bone Miner. Res.* **15(10),** 1950–1955.
19. Albagha, O. M., et al. (2001). *J. Bone Miner. Res.* **16(1),** 128–134.
20. Mann, V., et al. (2001). *J. Clin. Invest.* **107(7),** 899–907.
21. Uitterlinden, A. G., et al. (2001). *J. Bone Miner. Res.* **16(2),** 379–385.
22. Ota, N., et al. (2000). *Genes Immun.* **1(4),** 260–264.
23. Langdahl, B. L., et al. (2000). *J. Bone Miner. Res.* **15(3),** 402–414.
24. Keen, R. W., et al. (2001). *Rheumatology (Oxford)* **40(1),** 48–54.
25. Grupe, A., et al. (2001). *Science* **292(5523),** 1915–1918.
26. Loughlin, J., et al. (2000). *J. Rheumatol.* **27(3),** 779–784.
27. Ogawa, S., et al. (1999). *Biochem. Biophys. Res. Commun.* **260(1),** 122–126.
28. Koller, D. L., et al. (2000). *J. Clin. Endocrinol. Metab.* **85(9),** 3116–3120.
29. Yershov, Y., et al. (2001). *J. Bone Miner. Res.* **16(6),** 992–1003.
30. Nguyen, T. V., Blangero, J., and Eisman, J. A. (2000). *J. Bone Miner. Res.***15(3),** 392–401.
31. Blank, R. D. (2001). *J. Bone Miner. Res.* **16(7),** 1207–1211.
32. Dennis, J. E., et al. (1999). *J. Bone Miner. Res.* **14(5),** 700–709.
33. Pittenger, M. F., et al. (1999). *Science* **284(5411),** 143–147.
34. Aubin, J. E. (1998). *Biochem. Cell. Biol.* **76(6),** 899–910.
35. Nishida, S., et al. (1999). *J. Bone Miner. Metab.* **17(3),** 171–177.
36. Muraglia, A., Cancedda, R., and Quarto, R. (2000). *J. Cell. Sci.* **113(Pt. 7),** 1161–1166.
37. Justice, M. J., et al. (1999). *Hum. Mol. Genet.* **8(10),** 1955–1963.
38. Gronthos, S., et al. (1999). *J. Bone Miner. Res.* **14(1),** 47–56.
39. Chen, X. D., et al. (1999). *J. Bone Miner. Res.* **14(3),** 362–375.
40. Seshi, B., Kumar, S., and Sellers, D. (2000). *Blood Cells Mol. Dis.* **26(3),** 234–246.
41. Dacic, S., et al. (2001). *J. Bone Miner. Res.* **16(7),** 1228–1236.
42. Bruder, S. P., et al. (1998). *J. Bone Miner. Res.* **13(4),** 655–663.
43. Bourgeois, P., et al. (1998). *Hum. Mol. Genet.* **7(6),** 945–957.
44. Fang, M. A., et al. (2001). *J. Cell. Biochem.* **80(4),** 550–559.
45. Pesce, M. and Scholer, H. R. (2001). *Stem Cells* **19(4),** 271–278.
46. Rice, D. P., et al. (2000). *Development* **127(9),** 1845–1855.
47. Walsh, S., et al. (2000). *Bone* **27(2),** 185–195.
48. Wada, Y., et al. (1998). *Bone* **22(5),** 479–485.
49. Kalajzic, I., et al. (2002). *J. Bone Miner. Res.* **17,** 15–25.
50. Silver, L. M. (1995). *Mouse genetics: concepts and applications.* Oxford University Press: New York.
51. Dozmorov, I., Bartke, A., and Miller, R. A. (2001). *J. Gerontol. A Biol. Sci. Med. Sci.* **56(2),** B72–B80.
52. Rosen, C. J., et al. (2000). *Bone* **27(4),** 521–528.
53. Dimai, H. P., et al. (1998). *Bone* **22(3),** 211–216.
54. Linkhart, T. A., et al. (1999). *J. Bone Miner. Res.* **14(1),** 39–46.
55. Kerr, M. K., Martin, M., and Churchill, G. A. (2000). *J. Comput. Biol.* **7(6),** 819–837.
56. Kerr, M. K. and Churchill, G. A. (2001). *Genet. Res.* **77(2),** 123–128.
57. Montero, A., et al. (2000). *J. Clin. Invest.* **105(8),** 1085–1093.
58. Krebsbach, P. H., et al. (1997). *Transplantation* **63(8),** 1059–1069.
59. Bianco, P., et al. (1998). *J. Clin. Invest.* **101(8),** 1737–1744.
60. Stover, M. L., et al. (2001). *Mol. Ther.* **3(4),** 543–550.
61. Kalajzic, I., et al. (2001). *Virology* **284(1),** 37–45.