REVIEW

JBMR®

Establishing Biomechanical Mechanisms in Mouse Models: Practical Guidelines for Systematically Evaluating Phenotypic Changes in the Diaphyses of Long Bones

Karl J Jepsen,[1] Matthew J Silva,[2] Deepak Vashishth,[3] X Edward Guo,[4] and Marjolein CH van der Meulen[5]

[1]Department of Orthopaedic Surgery, University of Michigan, Ann Arbor, MI, USA
[2]Department of Orthopaedic Surgery, Washington University, St. Louis, MO, USA
[3]Department of Biomedical Engineering, Center for Biotechnology and Interdisciplinary Studies, Rensselaer Polytechnic Institute, Troy, NY, USA
[4]Department of Biomedical Engineering, Columbia University, New York, NY, USA
[5]Department of Biomedical Engineering and Sibley School of Mechanical & Aerospace Engineering, Cornell University, Ithaca, NY, USA

ABSTRACT

Mice are widely used in studies of skeletal biology, and assessment of their bones by mechanical testing is a critical step when evaluating the functional effects of an experimental perturbation. For example, a gene knockout may target a pathway important in bone formation and result in a "low bone mass" phenotype. But how well does the skeleton bear functional loads; eg, how much do bones deform during loading and how resistant are bones to fracture? By systematic evaluation of bone morphological, densitometric, and mechanical properties, investigators can establish the "biomechanical mechanisms" whereby an experimental perturbation alters whole-bone mechanical function. The goal of this review is to clarify these biomechanical mechanisms and to make recommendations for systematically evaluating phenotypic changes in mouse bones, with a focus on long-bone diaphyses and cortical bone. Further, minimum reportable standards for testing conditions and outcome variables are suggested that will improve the comparison of data across studies. Basic biomechanical principles are reviewed, followed by a description of the cross-sectional morphological properties that best inform the net cellular effects of a given experimental perturbation and are most relevant to biomechanical function. Although morphology is critical, whole-bone mechanical properties can only be determined accurately by a mechanical test. The functional importance of stiffness, maximum load, postyield displacement, and work-to-fracture are reviewed. Because bone and body size are often strongly related, strategies to adjust whole-bone properties for body mass are detailed. Finally, a comprehensive framework is presented using real data, and several examples from the literature are reviewed to illustrate how to synthesize morphological, tissue-level, and whole-bone mechanical properties of mouse long bones. © 2015 American Society for Bone and Mineral Research.

KEY WORDS: BIOMECHANICS, BONE; CORTICAL BONE; MOUSE MODELS; FUNCTION; BIOMECHANICAL MECHANISMS

Introduction

The mouse is a remarkable model for identifying the molecular mechanisms underlying common and rare musculoskeletal diseases. Musculoskeletal biology research using mouse models has involved the analysis of inbred, outbred, and hybrid mouse strains to identify quantitative trait loci[1–3] and genes of interest[4]; the analysis of mutated mice to assess gene function[5–7]; and the analysis of adaptive responses to perturbations such as pharmacological treatments,[8] mechanical loading,[9–11] and environmental toxins.[12]

Targeted genetic approaches have been developed to strategically turn on, turn off, or functionally alter nearly any gene of interest and in a tissue-specific and timed manner.[13–17] However, these highly sophisticated genetic approaches are often conducted with nonspecific treatments of the biomechanical data used to establish the effects of the perturbation on bone mechanical function. For example, many studies have concluded that particular genetic perturbations result in a low bone mass phenotype. Because each perturbation may compromise bone mass and fracture resistance through unique biomechanical pathways (Fig. 1), the conclusions of such studies

Received in original form January 14, 2015; revised form April 3, 2015; accepted April 21, 2015. Accepted manuscript online April 27, 2015.
Address correspondence to: Karl J Jepsen, PhD, Glancy Family Scholar, 109 Zina Pitcher Place, Room 2001, A. Alfred Taubman Biomedical Research Science Building, Department of Orthopaedic Surgery, University of Michigan, Ann Arbor, MI 48109-2200 USA.
E-mail: kjepsen@med.umich.edu
Additional Supporting Information may be found in the online version of this article.
Journal of Bone and Mineral Research, Vol. 30, No. 6, June 2015, pp 951–966
DOI: 10.1002/jbmr.2539

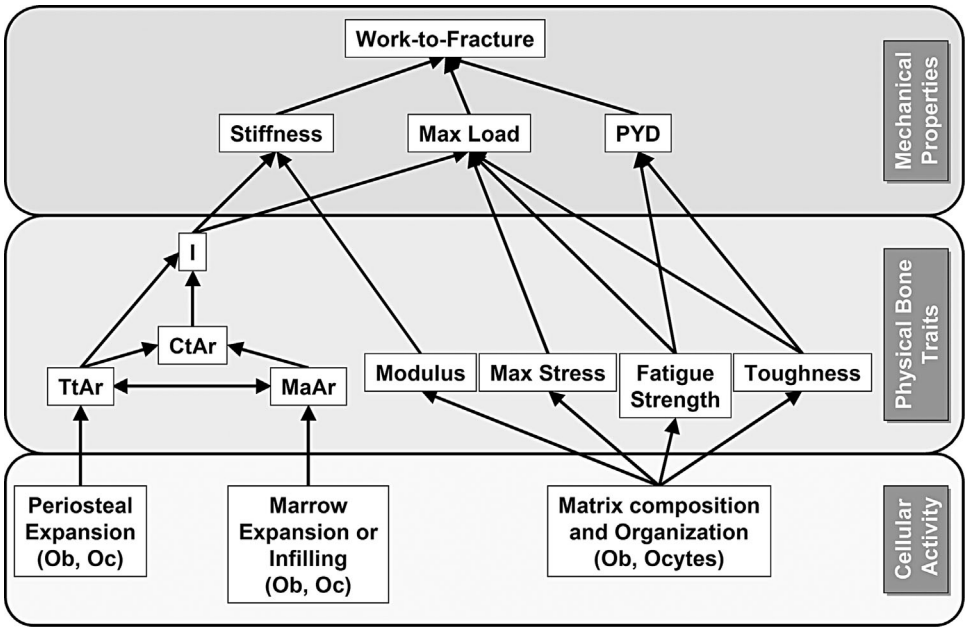

Fig. 1. Flowchart illustrating the collective interactions among morphological traits and tissue-level mechanical properties that arise from specific cellular activities and that together give rise to a unique suite of whole-bone mechanical properties. The changes in specific physical traits and the associated changes in whole-bone mechanical properties together define the unique biomechanical mechanism (pathway) linking genotype with phenotype. PYD = postyield displacement; Ob = osteoblast, Oc = osteoclast, Ocyte = osteocyte; Tt.Ar = total bone area; Ct.Ar = cortical area; Ma. Ar = marrow area.

cannot necessarily be applied directly to interpret bone fragility, a critical function of the skeleton.

Differentiating among biomechanical pathways will provide additional clues to gene function and disease mechanisms,[18] and allow researchers to better relate genetic and molecular mechanisms to the biomechanical mechanisms that define how genetic perturbations affect bone mass and fracture resistance. Similar to how the term *molecular mechanism* refers to the collective interactions among proteins defining the functionality of a molecular pathway, *biomechanical mechanism* refers to the unique collective interactions among morphological, compositional, microstructural, and ultrastructural traits that define whole-bone mechanical function (Fig. 1). Thus, an opportunity remains to match the molecular sophistication afforded by mouse models with more precise and comprehensive biomechanical analyses related to mechanical function.

The goal of this review is to better define biomechanical mechanisms in the skeleton by recommending guidelines to systematically evaluate phenotypic changes in mouse long bones. Further, minimum reportable testing conditions and outcome variables are recommended to facilitate the comparison of data across studies. This review focuses on whole-bone tests of long-bone diaphyses and cortical bone, the typical starting point for a biomechanical assessment in the mouse. We defer the presentation of corticocancellous structures to a separate follow-up work. Systematically evaluating phenotypic data can be complex, particularly when considering the adaptive nature of the skeletal system. Because mouse diaphyses are relatively simple tubular structures with well-defined adaptive properties,[9,10,19–21] developing guidelines first for long bones allows us to develop a broader appreciation for why establishing biomechanical mechanisms provides for more meaningful conclusions regarding how genetic and environmental perturbations impact skeletal function. Further, this review focuses on the most commonly used phenotypic analyses to provide a comprehensive understanding of these key traits rather than an exhaustive list of all available mechanical tests. These guidelines thus represent a starting point for biomechanical phenotyping. As bioengineers, our goal was to describe a methodology that allows biologists, our target audience, to systematically phenotype their mouse models, either in their laboratory or as well-informed collaborators alongside engineering colleagues who may be less familiar with the nuances of the skeletal system.

Types of Whole-Bone Mechanical Tests

Whole-bone mechanical tests of mouse long bones are most often performed in bending, but can also be performed in tension (pull), compression (push), and torsion (twist). [22] These tests are conducted by subjecting the bone to a single loading rate (ie, monotonic tests) or to multiple load cycles (ie, fatigue tests). The type of loading (ie, loading mode) should be selected for relevance to in vivo loads and reflect clinically or functionally applied loads for the bone of interest. In the long bones, the in vivo loads typically include bending and torsion,[23,24] which is why most studies are conducted using one of these two loading modes.

The magnitudes of mechanical properties differ by loading mode. As such, mechanical properties measured in torsion and bending cannot be directly compared on an absolute basis; however, the relative effect of the perturbation should be comparable across studies because these two loading modes

are related to similar morphological parameters.[25] For example, bending and torsion tests provided similar relative information on the changes in mechanical properties of intact mouse femora during postnatal growth.[20] The majority of whole-bone tests are performed in monotonic bending. When choosing between bending and torsion tests, a major consideration is that intact long bones tend to fail in a brittle manner during torsion tests compared to bending.[26] As such, bending tests are recommended if the perturbation is expected to alter bone ductility.

When designing monotonic failure tests, several factors need to be considered in addition to the relevant in vivo loads.[27] Practical considerations such as fixtures and bone alignment contribute to the test design. Bending tests of intact whole bones are straightforward to design and interpret and can be performed in three-point or four-point bending (Fig. 2*A*). Torsion tests also work well on long bones, avoid applying loads to the center of the bone, and are independent of the cross-sectional orientation. Therefore, torsion tests are often used to assess healing of fractured bones. For torsion tests, the long axis of the bone must be aligned with the rotational axis, and fixtures are needed to hold the bone ends and apply the loads[20,27] (Fig. 2*A*).

In fatigue or cyclic loading, a subfailure load is applied repetitively over time until failure occurs and is often used when a perturbation may affect cortical damage initiation and propagation.[28] Additional, more specialized testing methods include fracture toughness tests using prenotched specimens to monitor crack growth[29,30]; time-dependent tests that examine the creep or stress relaxation response[31]; and viscoelastic tests that characterize the response to different loading rates.[32] Test design and data interpretation are more complex in these tests. These tests are sensitive to changes in the bone matrix and are generally used to assess the impact of the perturbation on bone fracture.

Basic Biomechanical Principles

Mechanics terminology

The skeleton has multiple levels of organization across length scales. Here we consider only two: the whole-bone ("structural") level, eg, the femoral diaphysis; and the tissue ("material") level, namely cortical bone. The terms used to describe whole-bone mechanical properties are different from those used to describe tissue-level mechanical properties (Table 1). Precise use of these mechanical terms is analogous to precise use of molecular biology terms: one should not confuse RNA for cDNA, nor should one confuse maximum load for ultimate stress. The primary difference between mechanical properties at these two length scales is that whole-bone properties are measured using intact bones and depend on bone size, whereas tissue-level properties assess the material and are size-independent. Typically, whole-bone mechanical properties are expressed in terms of load (N) and displacement (mm), whereas tissue-level mechanical properties are expressed in terms of stress (N/mm^2) and strain (mm/mm or %).

Whole-bone *stiffness* characterizes how much the entire bone deforms when loaded. The tissue-level equivalent to stiffness is elastic modulus (or Young's modulus), which assesses the resistance to deformation of bone-tissue when loaded. "Tissue-level stiffness" is an acceptable synonym for elastic modulus.

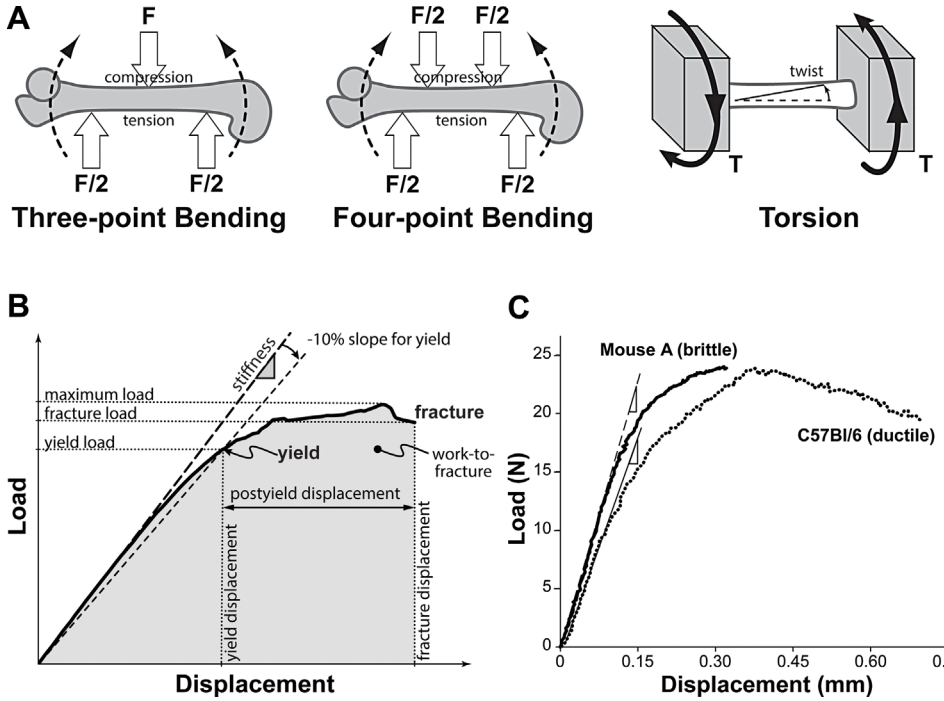

Fig. 2. (*A*) Schematics conveying loading of the mouse femur in three-point bending, four-point bending, and torsion. (*B*) A typical load-displacement curve resulting from a mouse bone loaded to fracture in four-point bending. The key outcome parameters are indicated including the minimum reportable set of whole bone mechanical properties listed in Table 1. Note that a three-point bending test would generate a comparable curve. (*C*) Two load-deflection curves illustrate differences in fracture behavior. Femurs from both mouse strains show similar stiffness and maximum load values, but mouse strain-A femurs fracture in a brittle-manner (low PYD) and B6 femurs fracture in a ductile manner (high PYD). PYD = postyield displacement.

Table 1. Terminology Used for Whole-Bone and Tissue-Level Mechanical Properties

Whole-bone mechanical properties [units]	Tissue-level mechanical properties [units]
Stiffness [N/mm]	Elastic modulus (or tissue-level stiffness) [$N/mm^2 = MPa$]
Maximum load (or whole-bone strength)[a] [N]	Ultimate stress (or tissue-level strength) [$N/mm^2 = MPa$]
Postyield displacement [mm]	Postyield strain [mm/mm (a dimensionless ratio)]
Work-to-fracture [Nmm]	Toughness (or modulus of toughness) [$N/mm^2 = MPa$]

[a]"Load" and "force" can often be used interchangeably; to be precise, "load" is the more general term and may refer to "force" (push or pull) or "moment" (bend or torque).

Likewise, *maximum load* refers to the greatest load (or force) a bone structure withstands before fracturing, whereas ultimate stress is the highest load per unit area that the bone-tissue withstands before fracturing. "Whole-bone strength" is synonymous with maximum load. Postyield displacement (PYD) is a measure of the displacement (or deformation) the bone structure experiences from the start of failure (the yield point) until fracture. Postyield strain is the analogous tissue-level quantity. Work-to-fracture (also called energy-to-fracture) measures the work done by the applied load to deform and fail the bone (also reflects the energy dissipated by the bone structure before breaking apart). At the tissue level, toughness (also called modulus of toughness) designates work per unit of material before fracture. Note that "fracture toughness" is a different tissue-level property, measured only by tests of notched specimens (Supporting Information, Section 3).

Measuring whole-bone mechanical properties

A comprehensive suite of mechanical properties can only be determined by performing mechanical tests. During a monotonic test, load and displacement are measured while the bone is being loaded.[20] For long bones such as the femur, tibia, and humerus, whole-bone tests measure the properties of cortical bone. By convention, the data are represented as a load-displacement curve with load on the *y*-axis and displacement on the *x*-axis (Fig. 2*B*). The load-displacement data are used to calculate whole-bone mechanical properties. We describe four mechanical properties derived from bending tests that are widely used to phenotype mouse long bones. The values of these parameters depend on the dimensions of the test fixtures. Direct comparisons between studies can be made only if equivalent test fixture geometries were used or if the properties were corrected for differences in fixture geometry (Supporting Information, Section 1). For this reason, the test fixture geometry (lower and upper span lengths) must be reported so data can be compared across studies. If the experimental design includes comparisons of groups having bones with lengths that differ by more than 2 mm, then the span lengths should be adjusted for each group and the load-displacement data corrected for test fixture geometry. No mouse long bone achieves the 16:1 diameter to length ratio recommended by ASTM for bending tests (ASTM C1684-08; ASTM, West Conshohocken, PA, USA); although this limitation is unavoidable for mechanical testing of mouse bones. Most studies use a lower span length that is ~50% of the bone length. The experimental methods should report the text fixture geometry in absolute terms and as a percentage of bone length for each test group.

Stiffness

Whole-bone stiffness measures the amount of elastic deformation a structure undergoes when loaded. The term "elastic" refers to the application of low load levels (eg, as experienced during physiological loading) that do not damage the material, and therefore the bone returns to its original state upon unloading. Stiffness is measured as the slope of the initial, linear portion of the load-displacement curve and is in units of load per displacement (N/mm) (Fig. 2*B*). Stiffer structures deform less for a given load than more compliant structures, and would be expected to experience less tissue-level deformation under physiological loads. Stiffness depends on both cortical morphology and bone material properties. Moment of inertia is the morphological trait that should correlate most strongly with bending stiffness, although total and cortical areas should also correlate with stiffness.

Maximum load

Bone fracture occurs after coalescence of microcracks into a macrocrack that propagates through the cortex.[33] Maximum load is simply the greatest load achieved before fracture and is reported in units of Newtons (N). Maximum load depends on both bone morphology and bone material. The moment of inertia, total area, and cortical area should correlate strongly with maximum load.

Postyield displacement

Bone tissue begins to damage when subjected to higher load levels. This damage reduces bone stiffness, causing the bone to show progressively greater displacements as loading continues. Yield is the point at which the load-displacement curve becomes nonlinear and deviates from the linear regression used to calculate stiffness. Unlike the low load levels within the initial elastic region, the mechanical behavior of bone is permanently altered when loaded beyond the yield-point; if unloaded, the bone now returns to a less stiff, permanently deformed state and not to the original, undamaged state. No standard method exists for calculating the yield-point in whole-bone mechanical tests; one simple approach uses a method similar to the calculation of the yield point for tissue-level mechanical tests (ie, the 0.2% offset method). The analogous method in whole-bone tests defines yield as the point at which a regression line that represents a 10% loss in stiffness crosses the load-displacement curve (Fig. 2*B*). Although no single method is recommended for calculating yield, the investigator should report the specific method used to calculate yield, including pertinent quantitative information (eg, 10% loss in stiffness) needed to replicate the method. The displacement (D) that occurs between yielding (D_{yield}) and fracture (D_{fx}) is referred to as postyield displacement ($PYD = D_{fx} - D_{yield}$). *Brittleness* is the absence of or very low PYD; *ductile* is the opposite of brittle and indicates the presence of substantial PYD. Postyield displacement is most often affected by perturbations that disrupt matrix composition or

organization.[26,28,34,35] Thus, a change in PYD often indicates a tissue-level phenotype. Because the fracture process is complex, a study that seeks to assess effects on postyield behavior must be powered to accommodate the higher variance of PYD (see Considerations for Assessing Whole-Bone Mechanical Properties).

Work-to-fracture

Work-to-fracture (also called energy-to-fracture) is a complex, integrative measure of a structure's overall resistance to failure. In whole-bone bending tests, work-to-fracture depends on the combined magnitudes of stiffness, maximum load, and PYD. The work-to-fracture is represented as the area under the load-displacement curve. PYD often has the greatest influence on work-to-fracture. For example, two mouse strains can show similar maximum loads but fracture in a brittle (mouse strain-A) versus a ductile (C57BL/6, B6) manner (Fig. 2*C*). B6 femurs have a significantly greater work-to-fracture despite similar maximum loads, because B6 femurs have increased PYD compared to mouse strain-A femurs. Interpreting differences in work-to-fracture requires considering the relative magnitudes of stiffness, maximum load, and PYD. For example, a perturbation that lowers maximum load but increases PYD could result in no difference in work-to-fracture of the experimental group compared to the control.

Bone Morphology (Cross-Sectional Geometric Properties)

Because whole-bone mechanical properties are highly dependent on bone size, a description of mechanical properties is of little value without a corresponding description of morphological properties.[20,25] Herein, we briefly review the morphological properties most relevant to long bone mechanical properties including a minimum reportable set (Table 6) and considerations for completing the Animal Research: Reporting of In Vivo Experiments (ARRIVE) guidelines checklist (Supporting Information, Section 9). Morphological properties are typically determined by micro-CT (μCT) analysis of the bone at the mid-diaphysis; ie, near the site where bone fails during mechanical testing (Fig. 2*A*). Guidelines for evaluating cortical bone morphology using μCT have been presented by Bouxsein and colleagues.[36] Similar to Bouxsein and colleagues,[36] we prescribe a scan region of interest (ROI) that is at least as long as the cortical thickness (~200 μm) and suggest an ROI between 0.5 and 1.0 mm in length. Furthermore, average cross-sectional morphology should be reported as area-values, not volumes. Volume-based measures depend on the ROI size and thus are not readily comparable across studies. If the μCT software does not report cross-sectional area values, then volume measures can be adjusted by the length of the ROI; eg, Ct.Ar = Ct.V/(ROI length), where ROI length = number of slices × slice thickness.

In the cortex, total bone area (Tt.Ar) is the area enclosed by the periosteal surface and marrow area (Ma.Ar) is the area enclosed by the endosteum (Fig. 3). These measures are primary traits that should be reported because each reflects cellular activity on a single surface. Because osteoblasts and osteoclasts may be active on both surfaces, changes in total area and marrow area over time or as a result of a perturbation reflect the net (summed) effect of bone deposition and resorption. The relationship between Tt.Ar and Ma.Ar and cellular activity can be confirmed by dynamic histomorphometry.[37] Cortical area (Ct.Ar) and cortical thickness (Ct.Th) are also important to report and indicate the net amount of bone at the diaphysis. However, Ct.Ar and Ct.Th reflect the net effects of cell activity on two surfaces and thus, are difficult to interpret mechanistically unless presented as part of a dataset that includes Tt.Ar and Ma.Ar. For example, a perturbation that reduces Ct.Th (or Ct.Ar) but increases Tt.Ar would be expected to increase whole-bone stiffness compared to its control. Thus, knowing outer bone size contextualizes the relationship between changes in Ct.Ar or Ct.Th and whole-bone stiffness and maximum load.

The morphological parameter of greatest relevance to the mechanical test depends on the type of loading applied to the whole bone. For loading in compression or tension, engineering theory dictates that Ct.Ar is the most relevant morphological parameter. For loading in bending or torsion, the morphological parameter of most relevance is the moment of inertia (I, J). The moment of inertia reflects both the amount of bone and the spatial distribution of this tissue. For example, two bones can have the same Ct.Ar but the bone with the large periosteal diameter has a larger moment of inertia than a bone with a small periosteal diameter.[25] Moment of inertia describes the geometric contribution of the bone to resisting bending and torsional loading. Unlike simple area measurements, moment of inertia values depend on a reference axis, which must be specified when reporting moment of inertia. In practice, the orientation of the bone cross-section can be challenging to control during μCT scanning. Fortunately, every bone cross-section has a single axis about which the bone is least stiff (ie, easiest to bend). The moment of inertia calculated relative to this axis is the minimum moment of inertia (I_{MIN}). Conversely, the axis about which the bone is most stiff corresponds to the maximum moment of inertia (I_{MAX}). We advocate reporting both I_{MIN} and I_{MAX} to describe moment of inertia. Reporting the moments of inertia relative to the medial-lateral (I_{ML}) and anterior-posterior (I_{AP}) axes is an acceptable alternative. Additional details on calculating moment of inertia from digitized cross-sections are included in the Supporting Information, Section 2.

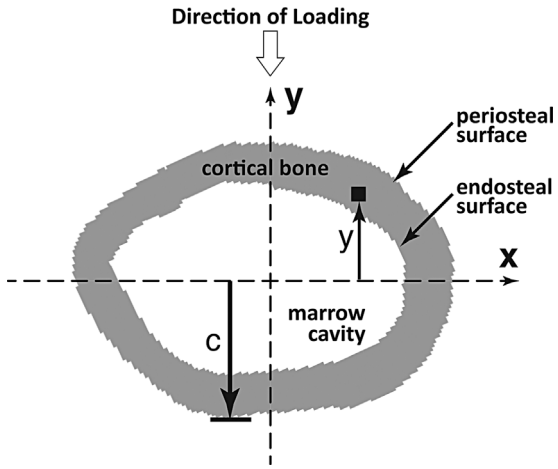

Fig. 3. Idealized cross-section of a mouse mid-shaft illustrating the outer (periosteal) surface, inner(endocortical) surface, marrowcavity, andsolid cortical bone. To quantify morphology, each bone cross-section is discretized into pixels (square), each with a known distance from the bending axis (eg, the distance *y* shown beneath the square).

When considering torsional loading, the relevant moment of inertia is the polar moment of inertia, J, which represents the geometric resistance to twisting a bone about its longitudinal axis. The value of J is independent of the reference axis (*x* or *y*) and calculated as: $J = I_x + I_y = I_{ML} + I_{AP} = I_{MIN} + I_{MAX}$.

Porosity

The relevance of cortical porosity for human bone has become of special interest recently because high-resolution CT scans reveal striking amounts of porosity in human metabolic bone diseases such as diabetes.[38] As a result, interest has increased in quantifying porosity within mouse cortical bone. For mouse bone, porosity can be measured using various technologies, such as histological sections, SEM images, μCT, or nanoCT. Pores within mouse cortical bone range in size across two to three orders of magnitude in dimensions and include canaliculi (0.1 μm to 1 μm in diameter),[39] lacunae (5 μm to 20 μm dimensions),[40] vascular pores (~10 μm in diameter),[41] and macropores that are visible in μCT images (20 μm to 50 μm in diameter).[34,41] Mutations may simply alter the normal porosity[42,43] or result in a "cancellization" of the cortex for more extreme phenotypes.[44] To facilitate the comparison of cortical porosity data across studies, define the type of pore being measured, the size range of the pores measured, and the voxel (or pixel) size of the technology used to quantify porosity.[36]

Assessing Tissue-Level Mechanical Properties

Material properties such as tissue modulus and ultimate stress can be estimated from whole-bone mechanical tests using equations from engineering beam theory. For bending tests, the equations are presented in Table 2. However, because bones do not have standardized geometries, these equations only provide rough estimates of material properties. The limitations of using beam theory equations to estimate the material properties of whole mouse long bones have been detailed elsewhere.[45–48] Several key points should be noted:

- Values of elastic modulus (tissue-level stiffness) are underestimated by a factor of 2-6;
- The aspect ratio (span length:bone width) may be as low as 5:1 for a mouse femur bending test, which is lower than beam theory assumes (>16:1);
- The mouse radius is more slender than the femur and the estimate of modulus will be more accurate;
- Cortical thickness strongly influences the estimates of modulus, with thinner cortices introducing more measurement error than thicker cortices;
- Both absolute values and relative differences in modulus between experimental groups may be inaccurate;
- Yield stress and ultimate stress calculations do not require data on specimen displacement and thus are less variable and more accurate than modulus values.

Given these limitations, why do we present these equations? They are useful as a first-order screening tool when deciding whether additional, more advanced tests of material properties should be undertaken. In no case should material properties estimated from beam theory be the only data used to support a conclusion about changes in material properties. If the perturbation alters tissue-level mechanical properties, then any number of confirmatory materials tests can be conducted. Such tests are becoming more common and offer the advantage of making mechanical observations at the scale of cellular activities. For example, the effects of knocking out a particular gene or set of genes[29] or the effects of mutations of specific genes[30,42] on bone microdamage and fracture could be traced to loss or modifications of key proteins. Several advanced materials tests and considerations for test selection are described in the Supporting Information, Section 3.

Establishing Biomechanical Mechanisms

Whole-bone mechanical properties—adjusting for body size

We begin the biomechanical evaluation by presenting an analysis of whole-bone mechanical properties followed by a systematic analysis of how changes in the whole-bone mechanical properties can be explained by changes in morphology and tissue mineral density (TMD). Assessing whole-bone mechanical properties first followed by the physical properties helps to establish biomechanical mechanisms by systematically moving through the structural hierarchy, ultimately interfacing with the cellular level. Molecular studies use a similar systematic approach when investigating differences in gene expression by confirming protein-level changes followed by testing for changes in downstream molecular targets. This top-down, hierarchically-based order also incorporates the adaptive nature of bone into the biomechanical mechanism. For example, if the perturbation is expected to alter whole-bone function but does not, then these systematic analyses may reveal how the system coordinately adjusted multiple

Table 2. Beam Theory Equations for Estimating Bone Material Properties From Bending Tests

Material property [units]	Three-point bending	Four-point bending
Elastic modulus, E [$N/mm^2 = MPa$]	$E = K L^3/(48 I)$	$E = Ka^2(3L-4a)/12I$
Stress, s [$N/mm^2 = MPa$]	$s = F L c/(4 I)$	$\sigma = F a c/(2 I)$
Toughness [$N/mm^2 = MPa$]	Not recommended[a]	Not recommended[a]

These equations are inaccurate for mouse bones, but may be useful for a "first pass" analysis. See text for discussion of limitations. Parameters are as defined in Fig. 2 and Supporting Fig. 1: K = stiffness, L = span length, a = distance from support to loading point for four-point bending, I = moment of inertia, F = force, c = distance from centroid of cross-section to outermost point on the cross-section, which can be approximated as bone width/2. The moment of inertia used in these calculations should be appropriate to the plane about which the bone is being bent; for a typical mouse femur test in the anterior-to-posterior (or posterior-to-anterior) direction (Supporting Fig. 1) this plane would be I_{ML} and similar to I_{MIN}. The value of force used to calculate stress is either the yield force or the ultimate force, corresponding to the yield stress or the ultimate stress, respectively.

[a]Not recommended because beam theory approximations only apply to the linear region, and toughness calculations involve substantial nonlinear postyield behavior.

morphological and/or tissue-level mechanical properties to maintain function.[6,21,49,50] This adaptation is similar to what may occur when deletion of one gene produces no cellular phenotype because altered expression of another gene(s) in the pathway compensates for the deletion.

The first adjustment that needs to be performed is to account for body size. Because bigger mice tend to have bigger bones, a small change in body size may be associated with significant changes in whole-bone mechanical properties, morphology and, less obviously, tissue-level mechanical properties. No generally agreed upon method exists for adjusting bone traits for body size. Measures of body size can include body mass, body length, bone length, body mass index (BMI), or percent lean mass. The product, body mass $\times$ bone length, is an alternative body size measure that is related to the bending loads applied to long bones.[50,51] The decision as to which body size measure to use for adjustment should be based on available data and scientific objectives. At minimum, we recommend using body mass as a measure of body size. We recommend using the linear regression method that takes into consideration the unique relationship between the trait and body mass for each test group. A more detailed description of this method can be found in the Supporting Information, Section 4. Generalized linear modeling (GLM) can be used when adjusting for multiple confounding variables.

We present two comparisons involving genetic mutations, to convey how conclusions are affected by the adjustment of traits for body size. In the first example, we compare the mechanical properties for the femora of 16-week-old male C57BL/6J (B6) mice and an experimental strain (mouse strain-A) that would often be considered to have a "low bone mass" phenotype. In the second example, we compare the mechanical properties for the femora of 16-week-old male B6 mice and an experimental strain (mouse strain-B) that has a significantly increased body mass (*t* test, $p < 0.0001$) and cortical area and thus may be referred to as a "gain of function" mutation if body size were not considered. B6 mice are included in both comparisons to illustrate how body size–adjusted group means are affected by the comparison mouse strain. We use data from these same examples in the following section demonstrating the full systematic analysis used to establish biomechanical mechanisms. The identities of the experimental mice are hidden to enable the reader to focus on the analysis rather than the specific genetic perturbations. To illustrate the importance of body size effects, we first present conclusions that would be drawn from an analysis of maximum load data. We then present an analysis of the remaining mechanical properties, along with the resolution of a discrepancy in the mechanical data that may arise.

Conclusions based on maximum load data

B6 versus mouse strain-A (Table 3*A*). Maximum load did not differ significantly between B6 and strain-A femurs. Adjusting maximum load for body size did not affect the outcome of this comparison. A graphic display of the data (Fig. 4*A*) conveys the overlap in maximum load values between strain-A and B6 femurs.

B6 versus mouse strain-B (Table 3*B*). When body size effects are ignored, one concludes that the genetic perturbation of strain-B did not affect femoral maximum load. However, when the larger body mass of strain-B was taken into account, the genetic perturbation of strain-B significantly reduced the maximum load relative to body size. A graphic display of the data (Fig. 4*A*) conveys differences in the slope ($p < 0.02$,

Table 3. Comparison of Raw and Adjusted Values for Whole-Bone Mechanical Properties for Femurs Loaded in Four-Point Bending

(*A*) Adult B6 ($n = 40$; body mass $= 27.0 \pm 1.8$ g) and Mouse Strain-A ($n = 10$; body mass $= 25.8 \pm 1.7$ g)				
	Unadjusted		Body mass–adjusted	
	B6	"A"	B6	"A"
Max load (N)	28.0 ± 3.3	27.6 ± 1.6 6%	27.2 ± 2.5	27.7 ± 1.5
	12%	(0.70)	9%	5% (0.53)
Stiffness (N/mm)	197.2 ± 28.6	206.9 ± 22.9	191.5 ± 23.5	**210.9 ± 20.0**
	15%	11% (0.32)	12%	**10% (0.02)**
PYD (mm)	0.33 ± 0.10	**0.13 ± 0.04**	0.33 ± 0.10	**0.13 ± 0.04**
	31%	**34% (0.0001)**	31%	**32% (0.0001)**
Work-to-fracture (Nmm)	10.6 ± 3.2	**7.8 ± 2.4**	10.4 ± 3.1	**7.8 ± 2.4**
	30%	**30% (0.01)**	30%	**30% (0.02)**
(*B*) B6 ($n = 40$; body mass $= 27.0 \pm 1.8$ g) and Mouse Strain-B ($n = 10$; body mass $= 31.6 \pm 2.6$ g)				
	Unadjusted		Body mass–adjusted	
	B6	"B"	B6	"B"
Max load (N)	28.0 ± 3.3	27.3 ± 1.7	30.7 ± 2.5	**26.7 ± 1.6**
	12%	6% (0.53)	8%	**6% (0.0001)**
Stiffness (N/mm)	197.1 ± 28.6	183.4 ± 23.7	218.1 ± 23.5	**180.8 ± 23.5**
	15%	13% (0.19)	11%	**13% (0.0001)**
PYD (mm)	0.33 ± 0.10	0.33 ± 0.06	0.33 ± 0.10	0.34 ± 0.06
	31%	18% (0.9)	31%	18% (0.7)
Work-to-fracture (Nmm)	10.6 ± 3.2	**14.1 ± 3.0**	11.5 ± 3.1	13.6 ± 2.9
	30%	**21% (0.003)**	27%	22% (0.06)

Whole-bone mechanical properties were adjusted for body mass using the linear regression method described in the text. Values are mean $\pm$ SD and coefficient of variation (%); *p* values are shown in parentheses. Bold values indicate significant differences relative to B6, $p < 0.05$.

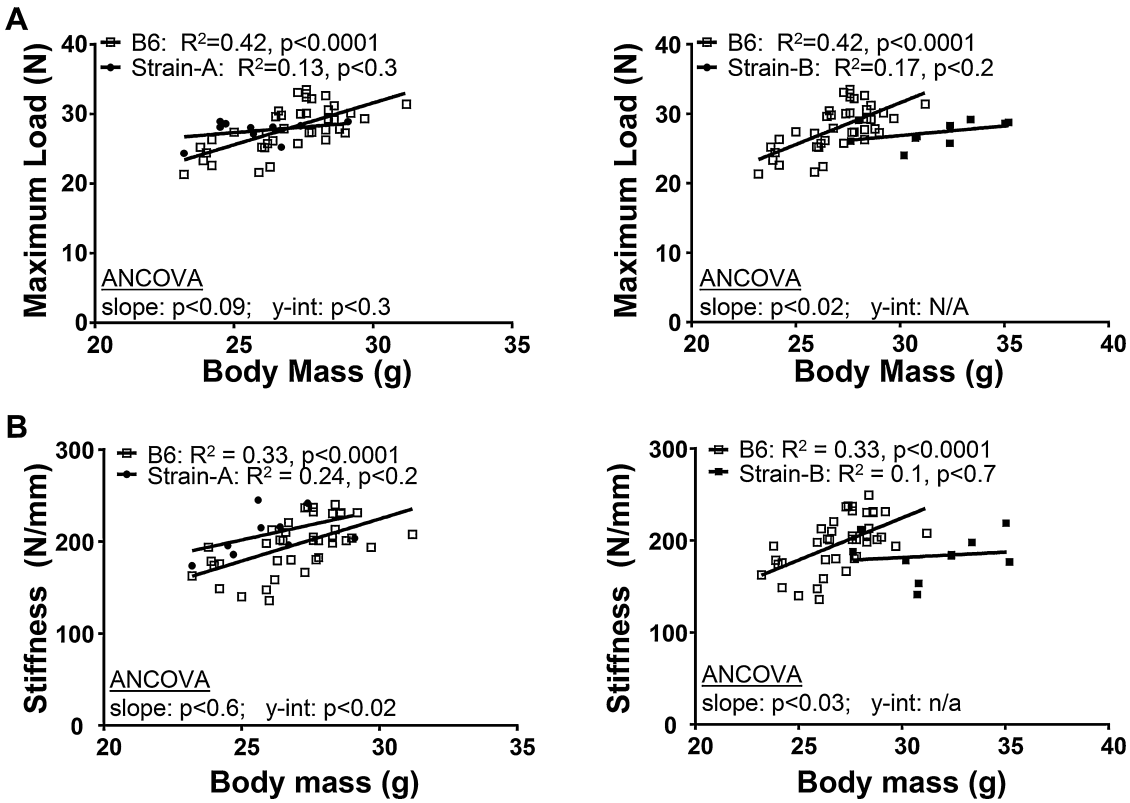

Fig.4. Comparison of linear regressions between (*A*) maximum load to fracture and body mass and (*B*) stiffness and body mass for B6 versus mouse strain-A and B6 versus mouse strain-B.

ANCOVA) of the maximum load-body mass regression between strain-B and B6 femurs. These graphs demonstrate why adjusting maximum load for body size lead to an increase in mean values for B6 femurs and a decrease in mean values for strain-B femurs.

The same analysis methods were applied to whole-bone stiffness, PYD, and work-to-fracture (Table 3). Adjusting stiffness for body mass revealed that strain-A femurs had significantly greater stiffness and strain-B femurs had significantly lower stiffness compared to B6. The differences in stiffness were readily apparent in graphs of stiffness versus body mass (Fig. 4*B*). Note that strain-B femurs showed consistent differences in both maximum load and stiffness relative to B6 femurs, but strain-A femurs showed significant differences in stiffness but not maximum load. In general, perturbations that increase whole-bone stiffness also tend to increase maximum load, because these two mechanical properties depend on similar morphological traits (moment of inertia). The apparent discrepancy between maximum load and stiffness for strain-A femurs may be explained in part by the significant differences in PYD. The lower PYD of strain-A femurs indicates that the genetic perturbation lead to a brittle phenotype. For brittle bones, the damage that initiates during yielding is expected to propagate quickly to a macrocrack resulting in premature failure, thus explaining the lower than expected maximum load values for strain-A femurs. This explanation could be confirmed by conducting additional tissue-level studies such as fracture toughness tests that are independent of bone size (see Supporting Information, Section 3). Table 3 also conveys that mean values for PYD and work-to-fracture, which typically show poor correlations with body mass, are not appreciably altered after adjusting for body mass (although significance is lost in work-to-fracture for strain-B femurs).

Systematically evaluating morphology and tissue mineral density to establish biomechanical mechanisms

Whole-bone mechanical behavior reflects the integrated contributions of bone morphology and tissue-level mechanical properties. Although the whole-bone mechanical properties alone provide insight into skeletal function, more can be learned about the mechanism by systematically examining the contributions of morphological traits and tissue properties to the whole-bone measures. One reason to examine many traits and in a systematic manner is the dependence of function on the combined effects of multiple traits, which means that no single physical property can be used to fully define how a perturbation affected function. For example, Ct.Th is difficult to relate directly to changes in mechanical properties without knowledge of associated changes in outer bone size. One should not selectively report only those properties that are significantly different between test groups. This practice can lead to misleading conclusions and impairs the ability of others to understand and compare results between studies. Like molecular mechanisms, the particular combination of significant and nonsignificant differences in physical traits defines the biomechanical mechanism that is unique to the perturbation under consideration.

Herein, we present a systematic evaluation of specific morphological and compositional traits that we have found useful to identify the biomechanical mechanisms explaining the effect of a perturbation on whole-bone mechanical function. We use an "outside-in" approach, which means we begin by testing for changes in longitudinal growth and periosteal expansion (outside) and end with changes in matrix composition (inside). We present conclusions based on unadjusted data and data adjusted for body mass using the regression method. A comparison of different body size adjustment methods can be found in the Supporting Information, Section 5.

1. Did longitudinal growth differ?

Longitudinal bone growth occurs through the hypertrophy of growth plate chondrocytes followed by matrix calcification. The outcome of this endochondral process is manifested by bone length differences. Testing for changes in bone length (Le or L) is important for understanding the impact of a perturbation on longitudinal growth.

B6 versus mouse strain-A (Table 4 A). After adjusting for differences in body mass, femoral length was not different between B6 and strain-A mice.

B6 versus mouse strain-B (Table 4B). After adjusting for body mass, bone length was not significantly different between B6 and strain-B mice, suggesting that the greater unadjusted length of strain-B femurs was explained by the increased body mass.

2. Did periosteal expansion change?

Changes in periosteal expansion can be attributed to the amount of bone formation relative to resorption occurring on the outer bone surface and can be evaluated by testing for differences in total cross-sectional area (Tt.Ar) between experimental groups. Periosteal circumference can be used as an alternative trait. Moment of inertia does not directly assess changes in periosteal expansion, but rather depends on outer bone diameter and cortical area combined. If bone length differs between experimental groups, then normalizing Tt.Ar by bone length (robustness = Tt.Ar/Le) can be used to test for differences in periosteal expansion.

B6 versus mouse strain-A (Table 4A). Strain-A femurs had a significantly smaller body size–adjusted Tt.Ar compared to B6 femurs, suggesting the genetic perturbation of strain-A suppressed periosteal expansion during growth.

B6 versus mouse strain-B (Table 4B). Strain-B femurs had a significantly smaller Tt.Ar compared to B6 after adjusting for body mass, suggesting the genetic perturbation suppressed periosteal expansion during growth.

3. Did marrow expansion change?

The biological factors defining marrow area (Ma.Ar) are complex, involving net resorption (expansion) and net formation (infilling) at different times during growth.[52] Analyses of Ma.Ar and Tt.Ar are needed to determine whether a change in cortical area (Ct.Ar, see below) arises from biological factors affecting periosteal expansion and/or marrow expansion.

B6 versus mouse strain-A (Table 4A). Strain-A femurs had a significantly smaller Ma.Ar compared to B6, which was expected given that strain-A femurs also had a smaller Tt.Ar. When adjusting Ma.Ar for both body mass and Tt.Ar by GLM, strain-A femurs still had a significantly smaller Ma.Ar compared to B6 but only by 7% ($p < 0.006$, GLM; data not shown). Thus, the genetic perturbation of strain-A significantly suppressed periosteal expansion and had significant but minor effects on marrow expansion.

B6 versus mouse strain-B (Table 4B). After adjusting for differences in body mass, strain-B femurs had a significantly smaller Ma.Ar compared to B6. When adjusting Ma.Ar for both body mass and Tt.Ar, the Ma.Ar for strain-B femurs no longer differed from B6 ($p < 0.32$, GLM; data not shown), suggesting that the genetic perturbation of strain-B affected body mass and Tt.Ar, but not Ma.Ar.

4. Did cortical area change?

A simple measure of mass accumulation (ie, the amount of bone) for diaphyseal bone is cortical area (Ct.Ar). Determining whether the perturbation affected Ct.Ar is complicated by the complex adaptive nature of the skeletal system. Ct.Ar varies with outer bone size, such that narrow (slender) bones have a naturally lower absolute Ct.Ar but a higher relative cortical area (Ct.Ar/Tt.Ar) compared to wide (robust) bones.[21] These naturally varying traits complicate phenotypic analyses; therefore, you should not draw conclusions based on a comparison of Ct.Ar alone. The GLM method can be used to adjust for both body mass and Tt.Ar (or robustness) to test whether the perturbation affected mass accumulation beyond that expected for differences in body mass and outer bone size.[50]

B6 versus mouse strain-A (Table 4A). On an absolute basis, strain-A femurs had a significantly lower Ct.Ar compared to B6 femurs, an expected result because strain-A femurs were more slender (ie, lower Tt.Ar). However, after adjusting for both body mass and Tt.Ar by GLM, strain-A femurs had significantly greater Ct.Ar compared to B6 femurs ($p < 0.0065$; GLM with both body mass and Tt.Ar as covariates). Thus, the genetic perturbation of strain-A promoted greater mass accumulation. This outcome is consistent with the analysis showing that strain-A femurs had a slightly smaller Ma.Ar compared to B6 after adjusting for body mass and Tt.Ar.

B6 versus mouse strain-B (Table 4B). When adjusting Ct.Ar for both body mass and Tt.Ar by GLM, the Ct.Ar of strain-B femurs was not significantly different from B6 femurs ($p < 0.32$ GLM with both body mass and Tt.Ar as covariates). Thus, one concludes that the genetic perturbation of strain-B affected periosteal expansion but not mass accumulation in long bone diaphyses.

5. Did moment of inertia change?

As discussed previously, moment of inertia measures the spatial distribution of bone tissue and defines the morphological resistance of bone to bending and torsional loads. Interpreting changes in moment of inertia can be tricky, because this trait depends heavily on outer bone diameter and thus the effect of the perturbation on periosteal expansion. Importantly, moment of inertia also reflects the amount of bone (Ct.Ar), although to a lesser extent, and thus also depends on changes in mass

Table 4. Comparison of Raw and Body Size–Adjusted Bone Morphology and Tissue Mineral Density Properties for Adult Mouse Femurs

(*A*) B6 Compared to Mouse Strain-A With Lower Body Mass

	Unadjusted		Body mass adjusted	
	B6	"A"	B6	"A"
Le (mm)	15.9 ± 0.2	**15.7 ± 0.2***	15.9 ± 0.2	15.8 ± 0.1
	1.5%	**1.3%**	1.1%	0.9%
Tt.Ar (mm^2)	2.04 ± 0.20	**1.27 ± 0.06***	1.99 ± 0.16	**1.29 ± 0.05***
	10.0%	**4.7%**	8.2%	**3.5%**
Ma.Ar (mm^2)	1.15 ± 0.15	**0.51 ± 0.04***	1.13 ± 0.13	**0.52 ± 0.03***
	13.0%	**7.4%**	11.8%	**6.0%**
Ct.Ar (mm^2)	0.88 ± 0.07	**0.77 ± 0.03***	0.85 ± 0.04	**0.91 ± 0.06****
	8.1%	**3.5%**	4.9%	**6.8%**
I_{MIN} (mm^4)	0.17 ± 0.02	**0.11 ± 0.01***	0.17 ± 0.02	**0.11 ± 0.004***
	13.1%	**5.2%**	11.8%	**4.0%**
I_{MAX} (mm^4)	0.37 ± 0.04	**0.18 ± 0.01***	0.36 ± 0.03	**0.19 ± 0.01***
	11.1%	**7.8%**	9.5%	**6.8%**
J (mm^4)	0.54 ± 0.06	**0.29 ± 0.02***	0.53 ± 0.05	**0.29 ± 0.01***
	11.1%	**5.9%**	9.5%	**4.7%**
TMD (mg HA)	1416 ± 12	**1538 ± 11***	1418 ± 11	**1538 ± 11***
	0.8%	**0.7%**	0.7%	**0.7%**

(*B*) B6 Compared to Mouse Strain-B With Larger Body Mass

	Unadjusted		Body mass adjusted	
	B6	"B"	B6	"B"
Le (mm)	15.9 ± 0.2	**16.4 ± 0.2†**	16.2 ± 0.2	16.2 ± 0.1
	1.5%	**1.2%**	1.1%	0.8%
Tt.Ar (mm^2)	2.04 ± 0.20	2.14 ± 0.14	2.19 ± 0.16	**2.04 ± 0.08†**
	10.0%	6.8%	7.5%	**4.1%**
Ma.Ar (mm^2)	1.15 ± 0.15	1.20 ± 0.12	1.24 ± 0.13	**1.12 ± 0.08†**
	13.0%	9.8%	10.8%	**7.0%**
Ct.Ar (mm^2)	0.88 ± 0.07	**0.94 ± 0.04†**	0.90 ± 0.04	0.88 ± 0.05†††
	8.1%	**4.3%**	4.3%	5.5%
I_{MIN} (mm^4)	0.17 ± 0.02	0.18 ± 0.02	0.19 ± 0.02	0.18 ± 0.01
	13.1%	9.2%	10.7%	7.6%
I_{MAX} (mm^4)	0.37 ± 0.04	0.37 ± 0.03	0.40 ± 0.03	**0.35 ± 0.02†**
	11.1%	8.6%	8.6%	**5.6%**
J (mm^4)	0.54 ± 0.06	0.56 ± 0.04	0.59 ± 0.05	**0.53 ± 0.02†**
	11.1%	7.9%	8.6%	**4.5%**
TMD (mg HA)	1416 ± 12	**1397 ± 8†**	1410 ± 11	**1400 ± 7†**
	0.8%	**0.6%**	0.7%	**0.5%**

Data are presented as mean ± SD, and coefficient of variation (%). Group mean values for Ct.Ar are adjusted for both body mass and Tt.Ar by GLM. Properties were adjusted for body mass using the linear regression method.

*Bold text indicates $p < 0.05$ (B6 versus strain-A; *t* test).

**$p < 0.05$, adjusted for both body mass and Tt.Ar by GLM.

†Bold text indicates $p < 0.05$ (B6 versus strain-B; *t* test).

††$p < 0.05$, adjusted for both body mass and Tt.Ar by GLM.

†††$p > 0.05$, adjusted for both body mass and Tt.Ar by GLM.

accumulation. These combined influences explain why moment of inertia was analyzed after Tt.Ar and Ct.Ar. A further consideration is that bone diaphyses are often elliptically shaped cross-sections, and the geometric resistance to bending loads applied in the anteroposterior (A-P) direction differs from the geometric resistance to bending loads applied in the mediolateral (M-L) direction. Testing for changes in both axes or the ratio of these axes may reveal perturbations that preferentially affect the morphology in one plane over the other plane,[53] such as that found in human bone.[54] Thus, systematically analyzing the moment of inertia may further refine how the genetic perturbation affected bone morphology.

B6 versus mouse strain-A (Table 4A). Strain-A femurs had a significant reduction in the maximum (I_{MAX}, which is close to I_{AP}) and minimum (I_{MIN}, which is close to I_{ML}) moments of inertia compared to B6. Likewise, the polar moment of inertia (J) was significantly reduced in strain-A compared to B6. These outcomes were consistent with the significant reduction in Tt. Ar for strain-A femurs compared to B6 femurs.

B6 versus mouse strain-B (Table 4B). After adjusting for body size, strain-B femurs had a significant reduction in I_{MAX} and J, but not I_{MIN}. Based on this analysis, one concludes that the reduced Tt.Ar in strain-B results from suppressed periosteal expansion that primarily affected M-L width.

6. Did tissue mineral density or the organic matrix change?

Tissue-mineral density (TMD) measures mineral content by X-ray absorption and is often used as a surrogate measure of tissue-level stiffness.[36,55,56] Note that limitations are associated with TMD values determined using benchtop μCT scanners with polychromatic X-ray sources; comparison of TMD values between studies, in particular, should be made with caution.[36] Similarly, changes in the organic matrix of bone are being considered as these perturbations affect postyield properties.[35] However, much work remains to relate TMD and the organic matrix to the full suite of tissue-level mechanical properties.[57–59] Although the association between TMD and body size is weak, TMD should be compared between experimental groups after adjusting for body size if an association is observed for the mouse strains being examined.

B6 versus mouse strain-A (Table 4A). After adjusting for body mass, strain-A femurs had significantly greater TMD compared to B6 femurs. Thus, one concludes that the suppressed periosteal expansion associated with the genetic perturbation of strain-A was accompanied by a significant increase in TMD.

B6 versus mouse strain-B (Table 4B). Strain-B femurs had significantly reduced TMD compared to B6, even after adjusting for body mass. Thus, one concludes that the suppressed periosteal expansion associated with the genetic perturbation of strain-B not only suppressed periosteal expansion but also impaired TMD beyond that expected for the more slender phenotype.

Summary of biomechanical mechanisms

B6 versus mouse strain-A (Fig. 5A). The genetic perturbation of strain-A significantly impaired periosteal expansion but not whole-bone stiffness and maximum load, because the significantly smaller cross-sectional area of strain-A femurs was associated with coordinated changes in TMD. However, this coordination occurred at the expense of increased bone brittleness (reduced PYD) and decreased work-to-fracture. Further work is needed to confirm that the increased TMD in strain-A femurs reflected an increase in tissue-level stiffness and strength and a reduction in tissue-toughness. This comparison illustrates how the complex adaptive nature of the skeletal system affects biomechanical analyses. Coordinate changes in multiple traits may lead to similar whole-bone mechanical function despite significant differences in the cross-sectional moment of inertia. This analysis also exemplifies the limitations of inferring mechanical function based on engineering beam-theory. According to beam theory, the significant reduction in moment of inertia combined with a modest 7.5% increase in TMD could lead one to incorrectly conclude that the genetic perturbation of strain-A led to a "low bone mass" phenotype and reduced whole bone strength.

B6 versus mouse strain-B (Fig. 5B). The genetic perturbation of strain-B impaired mechanical function by suppressing periosteal expansion and disrupting the coordinate increase in TMD needed to mechanically offset the more slender femoral morphology. The changes in tissue-modulus and strength shown in the biomechanical mechanism would need to be confirmed with independent tissue-level mechanical tests. This analysis illustrates how small changes in body size may confound biomechanical analyses. If body size were not considered, then one incorrectly concludes that the genetic perturbation had no effect on mechanical properties or outer bone size but increased mass accumulation (Ct.Ar).

Considerations for Assessing Whole-Bone Mechanical Properties

What wild-type background strain should I use?

All whole-bone mechanical properties vary among inbred mouse strains.[60] Thus, the control and experimental inbred mouse strains should be of the same genetic background to ensure group comparisons are valid.

Which bone should I test?

This decision depends on the site where you think your perturbation will have the greatest effect. If unclear, the femur is recommended as this bone is long, fairly straight, and considerable data in the literature can be used for comparison. Additional mechanical considerations were described by Schriefer and colleagues.[46]

If I am testing in bending, should I use three-point or four-point bending?

No justification exists to recommend one test over the other, and so the decision of which method to use is often a matter of preference. The primary differences between the two tests include test fixture fabrication and the location of failure. If starting with no experience, three-point bending may be easier to implement because the fixtures are simpler to fabricate. Four-point bending fixtures are more complex to fabricate to ensure that all four loading points contact the bone. Fractures generally occur under the center loading point in three-point bending, which means you decide where the bone fractures based on where the center loading point contacts the bone. In four-point bending, the bone is subjected to a uniform moment (ie, load × distance) between the middle two loading points, and fracture occurs at any location between the middle points, which allows the bone to fail at its weakest location. Outcome variables cannot be directly compared between three-point and four-point bending without adjusting for test fixture geometry, because the magnitude of stiffness, maximum load, PYD, and work-to-fracture depend on the test fixture geometry (Supporting Information, Section 1). For unadjusted data, stiffness and maximum load values will be higher but PYD lower in four-point bending compared to three-point bending. However, stiffness and maximum load, but not PYD, are similar for three-point and four-point bending after adjusting for text fixture geometry. Thus, stiffness and maximum load adjusted for text fixture geometry are comparable across studies.

How fast should I load the bones?

Most bending tests are conducted with the loading points displacing at a rate of 0.5 to 1 mm/s. If the perturbation is

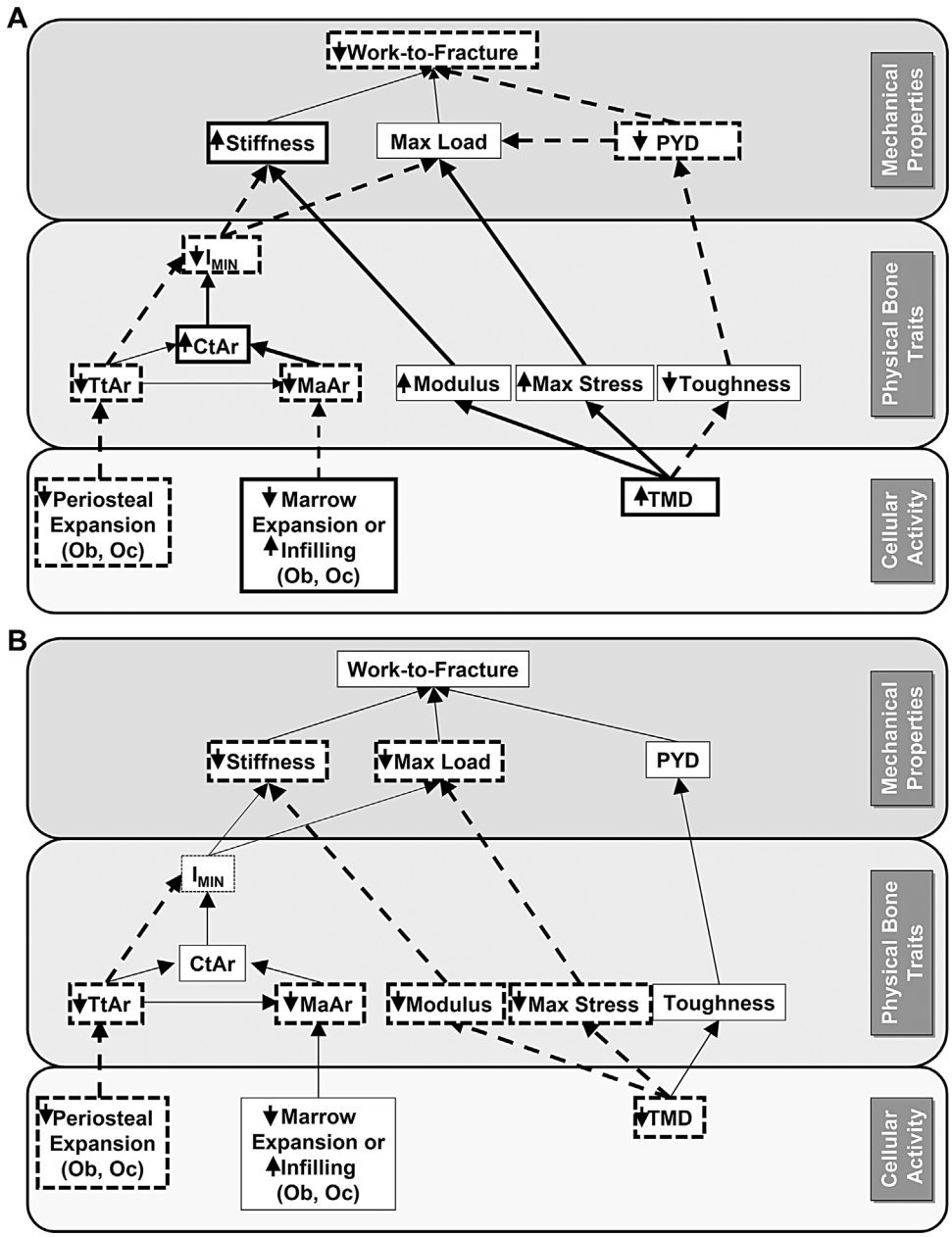

Fig. 5. (*A*) Biomechanical mechanism explaining how the genetic perturbation of mouse strain–A suppressed periosteal expansion but maintained similar whole-bone stiffness and maximum load as B6. Changes in TMD presumably reflect increases in elastic modulus and material strength. How the changes in matrix composition and organization affected PYD remains to be determined for mouse strain-A. (*B*) Biomechanical mechanism explaining how the genetic perturbation of mouse strain-B suppressed periosteal expansion but failed to maintain a similar whole-bone stiffness and maximum load as B6 femurs. Changes in TMD presumably reflect decreased elastic modulus and material strength. PYD did not seem to be affected by the reduced TMD. Bold solid lines indicate positive changes in a trait. Bold dashed lines indicate negative changes in a trait. PYD = postyield displacement; Ob = osteoblast; Oc = osteoclast; Ocyte = osteocyte; TMD = tissue mineral density; I_{MIN} = minimum moment of inertia; Tt.Ar = total bone area; Ct.Ar = cortical area; Ma.Ar = marrow area.

expected to alter postyield behavior, then reducing the loading rate to 0.05 to 0.1 mm/s may improve the potential to detect differences in PYD between experimental groups. A brittle bone will still fracture in a brittle manner, but a more ductile bone will generally show a longer postyield deformation at the slower loading rate. Thus, slower loading rates may increase statistical power by increasing the difference in PYD between experimental groups.

Does orientation of the bone samples matter during bending tests?

In bending, the side being contacted by the middle loading point(s) is undergoing compression, whereas the opposite side is undergoing tension (Fig. 2*A*). Although no study has systematically evaluated this loading condition, pilot studies conducted by the authors suggest that the correspondence

between measurements varies when tests performed in the anterior-to-posterior direction are compared to tests performed in the posterior-to-anterior direction. As such, whole-bone tests should be conducted using consistent loading protocols for reproducibility. Results from one of these pilot studies are described in the Supporting Information, Section 6. Although this pilot study bears replication, the results sufficiently convey that conducting mechanical tests using a consistent loading orientation is central to reproducibility.

Can I combine data for males and females?

Not unless the tests show equivalence among bones from both sexes. On an absolute basis, female mice typically have lower bone properties compared to male mice given their smaller body size. However, after adjusting bone properties for body size, female mice have greater femoral maximum load, a similar robustness, larger cortical area, and higher tissue-mineral density (TMD) compared to male mice.[61] Thus, sex-specific differences in bone properties on an absolute and relative basis preclude combining data in most experiments.

Can I combine data across different ages?

The answer depends on whether bone morphology and/or composition vary across the age range of interest. Bone properties change rapidly during growth, so combining data from different ages up to 12 to 16 weeks of age is not recommended.[20,52,62,63] Although combining data from different ages after 12 to 16 weeks of age is safer, notable exceptions exist. Mouse strains like the C57BL/6J show rapid loss of trabecular bone volume in the distal femur between 2 and 8 months of age,[64] and combining data for multiple ages is not recommended. When an ideal study design (ie, same sex and same age) cannot be achieved, the data should be plotted as a function of age and tested whether the trait or the response to a perturbation correlate significantly with age. Because body mass often increases with age, both age and body mass should be included as covariates in the statistical analysis.

How many mice should I use?

Calculating the number of mice needed to test for significant differences between group means is critical to drawing meaningful conclusions. The coefficient of variation (COV) varies widely among mechanical and physical properties. We have observed that bone length and tissue mineral density (TMD) have COVs in the 0.5% to 1.5% range; cross-sectional morphological traits have COVs in the 5% to 15% range; maximum load has COVs in the 10% to 15% range; stiffness has COVs in the 15% to 20% range; and PYD and work-to-fracture have COVs in the 25% to 50% range.

We conducted a power analysis using data comparing B6 and mouse strain-B femurs to illustrate how the number of mice needed to achieve statistical significance ($p = 0.05$, power = 80%) varies with the primary outcome measure and after adjusting for body size (Table 5). The impact of adjusting traits for body size was based on effects on the variance, not mean values. The COV was reduced by ~20% after adjusting for body size for this example. Thus, adjusting for body size often improves statistical power and reduces the number of mice needed to achieve statistical significance. The greater COVs for postyield properties are expected, because they reflect the intrinsic variability of the fracture process. If the perturbation alters matrix deposition and bone brittleness, the power analysis must use the greater variance of these postyield mechanical properties when calculating sample size. Alternatively, notched tests on whole bones (Supporting Information, Section 3) have a comparatively lower COV compared to tests on intact long bones, and thus may be effective with small samples sizes to evaluate differences in tissue composition and their effect on bone fracture.[65]

Summary

The primary objective of this review was to raise awareness among biologists that systematically evaluating phenotypic data with the intent of establishing biomechanical mechanisms (Fig. 1) may benefit efforts to integrate genetic perturbations with functional outcomes. Unlike engineered materials, living bone is highly adaptive and has a strong biological objective of establishing function.[21,66–70] The adaptive nature of bone can present unique challenges when attempting to define gene function using mouse models. We presented some considerations that are useful in distinguishing phenotypic changes that may arise directly from a genetic perturbation from those that may arise secondarily as an adaptive response to maintain function. Although much work remains to establish a standardized approach for conducting bone biomechanical analyses, our examples clearly demonstrate how adjusting data for body size is critical to drawing accurate conclusions. Thus, we hope the systematic analysis of phenotypic data provides the tools needed to define the biomechanical pathway that is unique to the perturbation under consideration and that may provide additional insight and further refinement of biological mechanisms. Additional examples of systematically evaluating phenotypic data can be found in the Supporting Information, Section 7.

The second objective of this review was to recommend minimum reportable information for experimental testing conditions (Table 6*A*) and outcome variables (Table 6*B*). Our recommendations are based on factors that should improve the comparison of data across studies and facilitate establishing

Table 5. Estimated Sample Sizes for Unadjusted and Body Size–Adjusted (shown in parentheses) Morphological Traits and Whole-Bone Mechanical Properties

		Expected percent difference in group means			
Traits	Average coefficient of variation	5%	10%	25%	50%
Tt.Ar, Ct.Ar, Ma.Ar, I, J	~10%	63 (41)	16 (11)	4 (2)	1 (1)
Stiffness, max load	~15%	141 (91)	36 (23)	9 (4)	2 (1)
PYD, Work-to-fracture	~30%	564 (361)	141 (91)	23 (15)	6 (4)

Sample sizes/group calculated for a two-tailed Student's *t* test. Numbers shown indicate sample sizes for each group. Body size–adjusted values are shown in parentheses.

Table 6.

(*A*) Parameters Required for Comprehensive Description of Whole-Bone Testing Methods in Different Experimental Categories[a]	
Mechanical test hardware	
Testing system	Manufacturer, model number
Sensors	Load cell (range, manufacturer), displacement measurement
Test fixture geometry	Lower (and upper) span length, gauge length[b]
Testing conditions	
Type of test	Axial (tension/compression), bending (three-point or four-point), or torsion
Control mode	Load or displacement control
Loading rate	Rate of applied load or displacement
Data acquisition rate	10× loading rate
Sample positioning	
Orientation of bone	Direction of load relative to anatomical directions
Longitudinal alignment	Location of loading points along bone's long axis
Morphology sampling	
Location measured	Correspondence with gauge length for mechanical testing
ROI location and size	Anatomical markers used to identify location
(*B*) Minimum Measures to Report for Whole-Bone Tests by Categories[c]	
Animal descriptors	
Mouse strain (genotype)	
Age	
Sex	
Body mass	
Bone length	
Whole-bone mechanical testing outcomes	
Maximum load	
Whole-bone stiffness	
Postyield displacement	
Work-to-fracture	
Cross-sectional morphology outcomes	
Total area	
Cortical area	
Marrow area	
Cortical thickness (optional)	
Minimum moment of inertia	
Maximum moment of inertia	

ROI = region of interest.

[a]Descriptors are included with each parameter.

[b]Gauge length refers to the loaded segment of the whole bone tests, which is the spacing of the upper loading point(s) in bending, or the length of the exposed region (ie, not contained within end caps) in torsion.

[c]Comparison of group means adjusted for body size is recommended, unless no body size effect is present.

biomechanical mechanisms. Example text for a Methods section is presented in the Supporting Information, Section 8. Finally, we provide recommendations for completing the ARRIVE guidelines checklist for in vivo experiments that involve biomechanical analyses (Supporting Information, Section 9). This review focused on whole-bone testing and cortical bone; efforts are underway to establish similar guidelines for the biomechanical evaluation of corticocancellous structures, which present their own unique challenges.

Disclosures

All authors state that they have no conflicts of interest.

Acknowledgments

The work reported in this publication was supported by the National Institute of Arthritis and Musculoskeletal and Skin Diseases of the National Institutes of Health under Award Numbers AR44927 (KJJ), AR065424 (KJJ), AR047867 (MJS), AR057235 (MJS), AR058004 (XEG), AR052461 (XEG), AR049635 (DV), and AR064034 (MCHM). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. We thank Lauren Smith and Erin Bigelow for their roles in acquiring the mechanical and morphological data used to illustrate the systematic phenotypic analyses.

Authors' roles: All authors contributed equally to the conceptualization, writing, and editing of this manuscript.

References

1. Beamer WG, Shultz KL, Churchill GA, et al. Quantitative trait loci for bone density in C57BL/6J and CAST/EiJ inbred mice. Mamm Genome. 1999;10(11):1043–9.
2. Li R, Tsaih S-W, Shockley K, et al. Structural model analysis of multiple quantitative traits. PLoS Genet. 2006;2(7):1046–57.

3. Saless N, Litscher SJ, Houlihan MJ, et al. Comprehensive skeletal phenotyping and linkage mapping in an intercross of recombinant congenic mouse strains HcB-8 and HcB-23. Cells Tissues Organs. 2011;194(2–4): 244–8.
4. Klein RF, Allard J, Avnur Z, et al. Regulation of bone mass in mice by the lipoxygenase gene Alox15. Science. 2004;303(5655):229–32.
5. Ducy P, Desbois C, Boyce B, et al. Increased bone formation in osteocalcin-deficient mice. Nature. 1996;382(6590):448–52.
6. Maloul A, Rossmeier K, Mikic B, Pogue V, Battaglia T. Geometric and material contributions to whole bone structural behavior in GDF-7-deficient mice. Connect Tissue Res. 2006;47(3):157–62.
7. Yakar S, Canalis E, Sun H, et al. Serum IGF-1 determines skeletal strength by regulating subperiosteal expansion and trait interactions. J Bone Miner Res. 2009;24(8):1481–92.
8. Saini V, Marengi DA, Barry KJ, et al. Parathyroid hormone (PTH)/PTH-related peptide type 1 receptor (PPR) signaling in osteocytes regulates anabolic and catabolic skeletal responses to PTH. J Biol Chem. 2013;288(28):20122–34.
9. Brodt MD, Silva MJ. Aged mice have enhanced endocortical response and normal periosteal response compared with young-adult mice following 1 week of axial tibial compression. J Bone Miner Res. 2010;25(9):2006–15.
10. Fritton JC, Myers ER, Wright TM, van der Meulen MC. Loading induces site-specific increases in mineral content assessed by microcomputed tomography of the mouse tibia. Bone. 2005;36-(6):1030–8.
11. Melville KM, Kelly NH, Surita G, et al. Effects of deletion of ER-Alpha in osteoblast-lineage cells on bone mass and adaptation to mechanical loading differs in female and male mice. J Bone Miner Res. Forthcoming. Epub 2015 Feb 24. DOI:10.1002/jbmr.2488
12. Pelch KE, Carleton SM, Phillips CL, Nagel SC. Developmental exposure to xenoestrogens at low doses alters femur length and tensile strength in adult mice. Biol Reprod. 2012;86(3):69.
13. McCauley LK. Transgenic mouse models of metabolic bone disease. Curr Opin Rheumatol. 2001;13(4):316–25.
14. Murray SA. Mouse resources for craniofacial research. Genesis. 2011;49(4):190–9.
15. Menke DB. Engineering subtle targeted mutations into the mouse genome. Genesis. 2013;51(9):605–18.
16. Bouabe H, Okkenhaug K. Gene targeting in mice: a review. Methods Mol Biol. 2013;1064:315–36.
17. Piret SE, Thakker RV. Mouse models: approaches to generating in vivo models for hereditary disorders of mineral and skeletal homeostasis. In: Thakker RV, Whyte MP, Eisen EJ, Igarashi T, editors. Genetics of bone biology and skeletal diseases. London, UK: Academic Press: 2013. p. 181–204.
18. Blank RD. Breaking down bone strength: a perspective on the future of skeletal genetics. J Bone Miner Res. 2001;16(7):1207–11.
19. Bonadio J, Jepsen KJ, Mansoura MK, Jaenisch R, Kuhn JL, Goldstein SA. A murine skeletal adaptation that significantly increases cortical bone mechanical properties. Implications for human skeletal fragility. J Clin Invest. 1993;92(4):1697–705.
20. Brodt MD, Ellis CB, Silva MJ. Growing C57Bl/6 mice increase whole bone mechanical properties by increasing geometric and material properties. J Bone Miner Res. 1999;14(12):2159–66.
21. Jepsen KJ, Hu B, Tommasini SM, et al. Genetic randomization reveals functional relationships among morphologic and tissue-quality traits that contribute to bone strength and fragility. Mamm Genome. 2007;18(6–7): 492–507.
22. Turner CH, Burr DB. Basic biomechanical measurements of bone: a tutorial. Bone. 1993;14(4):595–608.
23. Burr DB, Milgrom C, Fyhrie D, et al. In vivo measurement of human tibial strains during vigorous activity. Bone. 1996;18(5):405–10.
24. Fritton SP, McLeod KJ, Rubin CT. Quantifying the strain history of bone: spatial uniformity and self-similarity of low-magnitude strains. J Biomech. 2000;33(3):317–25.
25. van der Meulen MC, Jepsen KJ, Mikic B. Understanding bone strength: size isn't everything. Bone. 2001;29(2):101–4.
26. Silva MJ, Brodt MD, Wopenka B, et al. Decreased collagen organization and content are associated with reduced strength of demineralized and intact bone in the SAMP6 mouse. J Bone Miner Res. 2006;21(1):78–88.
27. Burstein AH, Frankel VH. A standard test for laboratory animal bone. J Biomech. 1971;4(2):155–8.
28. Jepsen KJ, Pennington DE, Lee YL, Warman M, Nadeau J. Bone brittleness varies with genetic background in A/J and C57BL/6J inbred mice. J Bone Miner Res. 2001;16(10):1854–62.
29. Poundarik AA, Diab T, Sroga GE, et al. Dilatational band formation in bone. Proc Natl Acad Sci U S A. 2012;109(47):19178–83.
30. Carriero A, Zimmermann EA, Paluszny A, et al. How tough is brittle bone? Investigating osteogenesis imperfecta in mouse bone. J Bone Miner Res. 2014;29(6):1392–401.
31. Lynch JA, Silva MJ. In vivo static creep loading of the rat forelimb reduces ulnar structural properties at time-zero and induces damage-dependent woven bone formation. Bone. 2008;42-(5):942–9.
32. Maruyama N, Shibata Y, Mochizuki A, et al. Bone micro-fragility caused by the mimetic aging processes in alpha-klotho deficient mice: in situ nanoindentation assessment of dilatational bands. Biomaterials. 2015;47:62–71.
33. Vashishth D, Tanner KE, Bonfield W. Contribution, development and morphology of microcracking in cortical bone during crack propagation. J Biomech. 2000;33(9):1169–74.
34. Jepsen KJ, Goldstein SA, Kuhn JL, Schaffler MB, Bonadio J. Type-I collagen mutation compromises the post-yield behavior of Mov13 long bone. J Orthop Res. 1996;14(3):493–9.
35. Tang SY, Allen MR, Phipps R, Burr DB, Vashishth D. Changes in non-enzymatic glycation and its association with altered mechanical properties following 1-year treatment with risedronate or alendronate. Osteoporos Int. 2009;20(6):887–94.
36. Bouxsein ML, Boyd SK, Christiansen BA, Guldberg RE, Jepsen KJ, Muller R. Guidelines for assessment of bone microstructure in rodents using micro-computed tomography. J Bone Miner Res. 2010-25(7):1468–86.
37. Dempster DW, Compston JE, Drezner MK, et al. Standardized nomenclature, symbols, and units for bone histomorphometry: a 2012 update of the report of the ASBMR Histomorphometry Nomenclature Committee. J Bone Miner Res. 2013;28(1):2–17.
38. Patsch JM, Burghardt AJ, Yap SP, et al. Increased cortical porosity in type 2 diabetic postmenopausal women with fragility fractures. J Bone Miner Res. 2013;28(2):313–24.
39. You LD, Weinbaum S, Cowin SC, Schaffler MB. Ultrastructure of the osteocyte process and its pericellular matrix. Anat Rec A Discov Mol Cell Evol Biol. 2004;278(2):505–13.
40. Mader KS, Schneider P, Muller R, Stampanoni M. A quantitative framework for the 3D characterization of the osteocyte lacunar system. Bone. 2013;57(1):142–54.
41. Schneider P, Stauber M, Voide R, Stampanoni M, Donahue LR, Muller R. Ultrastructural properties in cortical bone vary greatly in two inbred strains of mice as assessed by synchrotron light based micro- and nano-CT. J Bone Miner Res. 2007;22(10):1557–70.
42. Carriero A, Doube M, Vogt M, et al. Altered lacunar and vascular porosity in osteogenesis imperfecta mouse bone as revealed by synchrotron tomography contributes to bone fragility. Bone. 2014;61:116–24.
43. Kuhnisch J, Seto J, Lange C, et al. Multiscale, converging defects of macro-porosity, microstructure and matrix mineralization impact long bone fragility in NF1. PLoS One. 2014;9(1):e86115.
44. Lai LP, Lotinun S, Bouxsein ML, Baron R, McMahon AP. Stk11 (Lkb1) deletion in the osteoblast lineage leads to high bone turnover, increased trabecular bone density and cortical porosity. Bone. 2014;69:98–108.
45. Silva MJ, Brodt MD, Fan Z, Rho JY. Nanoindentation and whole-bone bending estimates of material properties in bones from the senescence accelerated mouse SA MP6. J Biomech. 2004;37-(11):1639–46.

46. Schriefer JL, Robling AG, Warden SJ, Fournier AJ, Mason JJ, Turner CH. A comparison of mechanical properties derived from multiple skeletal sites in mice. J Biomech. 2005;38(3):467–75.
47. Voide R, van Lenthe GH, Muller R. Bone morphometry strongly predicts cortical bone stiffness and strength, but not toughness, in inbred mouse models of high and low bone mass. J Bone Miner Res. 2008;23(8):1194–203.
48. van Lenthe GH, Voide R, Boyd SK, Muller R. Tissue modulus calculated from beam theory is biased by bone size and geometry: implications for the use of three-point bending tests to determine bone tissue modulus. Bone. 2008;43(4):717–23.
49. Yershov Y, Baldini TH, Villagomez S, et al. Bone strength and related traits in HcB/Dem recombinant congenic mice. J Bone Miner Res. 2001;16(6):992–1003.
50. Smith LM, Bigelow EM, Nolan BT, Faillace ME, Nadeau JH, Jepsen KJ. Genetic perturbations that impair functional trait interactions lead to reduced bone strength and increased fragility in mice. Bone. 2014;67:130–8.
51. Selker F, Carter DR. Scaling of long bone fracture strength with animal mass. J Biomech. 1989;22(11–12):1175–83.
52. Price CP, Herman BC, Lufkin T, Goldman HM, Jepsen KJ. Genetic variation in bone growth patterns defines adult mouse bone fragility. J Bone Miner Res. 2005;20(11):1983–91.
53. Mikic B, Van der Meulen MC, Kingsley DM, Carter DR. Mechanical and geometric changes in the growing femora of BMP-5 deficient mice. Bone. 1996;18(6):601–7.
54. Macdonald HM, Cooper DM, McKay HA. Anterior-posterior bending strength at the tibial shaft increases with physical activity in boys: evidence for non-uniform geometric adaptation. Osteoporos Int. 2009;20(1):61–70.
55. Wagner DW, Lindsey DP, Beaupre GS. Deriving tissue density and elastic modulus from microCT bone scans. Bone. 2011;49(5):931–8.
56. Bi X, Patil CA, Lynch CC, et al. Raman and mechanical properties correlate at whole bone- and tissue-levels in a genetic mouse model. J Biomech. 2011;44(2):297–303.
57. McBride SH, McKenzie JA, Bedrick BS, et al. Long bone structure and strength depend on BMP2 from osteoblasts and osteocytes, but not vascular endothelial cells. PLoS One. 2014;9(5):e96862.
58. Amend SR, Uluckan O, Hurchla M, et al. Thrombospondin-1 regulates bone homeostasis through effects on bone matrix integrity and nitric oxide signaling in osteoclasts. J Bone Miner Res. 2015;30-(1):106–15.
59. Vashishth D. The role of the collagen matrix in skeletal fragility. Curr Osteoporos Rep. 2007;5(2):62–6.
60. Jepsen KJ, Akkus OJ, Majeska RJ, Nadeau JH. Hierarchical relationship between bone traits and mechanical properties in inbred mice. Mamm Genome. 2003;14(2):97–104.
61. Jepsen KJ, Bigelow EM, Schlecht SH. Women build long bones with less cortical mass relative to body size and bone size compared with men. Clin Orthop Relat Res. Forthcoming. Epub 2015 Feb 18. DOI:10.1007/s11999-015-4184-2.
62. Miller LM, Little W, Schirmer A, Sheik F, Busa B, Judex S. Accretion of bone quantity and quality in the developing mouse skeleton. J Bone Miner Res. 2007;22(7):1037–45.
63. Main RP, Lynch ME, van der Meulen MC. Load-induced changes in bone stiffness and cancellous and cortical bone mass following tibial compression diminish with age in female mice. J Exp Biol. 2014;217 (Pt 10):1775–83.
64. Glatt V, Canalis E, Stadmeyer L, Bouxsein ML. Age-related changes in trabecular architecture differ in female and male C57BL/6J mice. J Bone Miner Res. 2007;22(8):1197–207.
65. Ritchie RO, Koester KJ, Ionova S, Yao W, Lane NE, Ager JW 3rd. Measurement of the toughness of bone: a tutorial with special reference to small animal studies. Bone. 2008;43(5):798–812.
66. Wright S. Correlation and causation. J Agric Res. 1921;20:557–85.
67. Waddington CH. Canalization of development and the inheritance of acquired characters. Nature. 1942;14:563–5.
68. Currey JD. Mechanical properties of bone tissues with greatly differing functions. J Biomech. 1979;12(4):313–9.
69. Frost HM. Bone "mass" and the "mechanostat": a proposal. Anat Rec. 1987;219(1):1–9.
70. Jepsen KJ, Hu B, Tommasini SM, et al. Phenotypic integration of skeletal traits during growth buffers genetic variants affecting the slenderness of femora in inbred mouse strains. Mamm Genome. 2009;20(1):21–33.