



HHS Public Access

Author manuscript

J Bone Miner Res. Author manuscript; available in PMC 2023 April 01.

Published in final edited form as:

J Bone Miner Res. 2022 April ; 37(4): 597–607. doi:10.1002/jbmr.4518.

Analysis of the associations between the human fecal microbiome and bone density, structure and strength: the MrOS cohort

Eric S. Orwoll, M.D.¹, Neeta Parimi, M.S.², Jack Wiedrick, M.S.³, Jodi Lapidus, Ph.D.^{3,4}, Nicola Napoli, M.D.⁵, Jeremy E. Wilkinson, Ph.D.⁶, Curtis Huttenhower, Ph.D.⁶, Lisa Langsetmo, Ph.D.⁷, Douglas P. Kiel, M.D., M.P.H.⁸

¹Department of Medicine, Oregon Health & Sciences University, Portland OR

²San Francisco Coordinating Center, San Francisco, CA

³Biostatistics & Design Program, Oregon Health & Science University, Portland, OR

⁴Oregon Health & Science University - Portland State University School of Public Health, Portland, OR

⁵Department of Medicine, Unit of Endocrinology and Diabetes, Campus Bio-Medico University of Rome, Italy; Division of Bone and Mineral Diseases, Washington University in St Louis

⁶Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, MA

⁷School of Public Health, University of Minnesota, Twin Cities, Minneapolis, MN

⁸Hinda and Arthur Marcus Institute for Aging Research, Hebrew Senior Life; Department of Medicine, Beth Israel Deaconess Medical Center & Harvard Medical School, Boston, MA, Broad Institute of MIT & Harvard, Cambridge, MA.

Abstract

Corresponding author Address correspondence to Eric Orwoll, MD, Bone and Mineral Research Unit (CR113), Oregon Health & Science University, 3181 S.W. Sam Jackson Park Rd., Portland OR, 97239. Phone: 503 494 0225. orwoll@ohsu.edu.

Author contributions

E.O. contributed to data collection, study conceptualization, drafting of the analysis plan, data curation, formal analysis, interpretation of results, writing of the original and final draft of the manuscript, and supervision of the study. N.P. contributed to analysis and reviewing of the manuscript. J.W. contributed to analysis, interpretation of results and reviewing of the manuscript. J.L. contributed to data curation, interpretation of results and writing and reviewing of the manuscript. N.N. contributed to interpretation of results and reviewing of the manuscript. J.W. contributed to the design of the analysis plan, interpretation of results, and reviewing of the manuscript. C.H. contributed to the design of the analysis plan, interpretation of results, and reviewing of the manuscript. L.L. contributed to interpretation of results and reviewing of the manuscript. D.K. contributed to data collection, study conceptualization, interpretation of results, reviewing of the manuscript and supervision of the study.

Disclosures

E.O.: Consultation for Amgen, Bayer, BioCon, Radius; research support from Mereo; travel funding from the OI Foundation

N.P.: Nothing to declare

J.W.: Nothing to declare

J.L.: Nothing to declare

N.N.: Nothing to declare

J.E.W.: Nothing to declare

C.H.: Nothing to declare

L.L.: Nothing to declare

D.P.K.: Dr. Kiel has received grants to his institution from Amgen, Inc, and Radius Health. He has received consulting fees from Solarea Bio and Pfizer for serving on Scientific Advisory Boards, and royalties for publication in UpToDate by Wolters Kluwer.

In pre-clinical models, the composition and function of the gut microbiota have been linked to bone growth and homeostasis, but there are few available data from studies of human populations. In a hypothesis generating experiment in a large cohort of community-dwelling older men (N= 831, age range 78-98 years), we explored the associations between fecal microbial profiles and bone density, microarchitecture, and strength measured with total hip DXA and HR-pQCT (distal radius, distal and diaphyseal tibia). Fecal samples were collected and the 16S rRNA gene V4 hypervariable region sequenced. Sequences were bioinformatically processed through the DADA2 pipeline and then taxonomically assigned using SILVA. Generalized linear models as implemented in MaAsLin 2 (Microbiome Multivariable Association with Linear Models) were used to test for associations between skeletal measures and specific microbial genera. The abundances of 4 bacterial genera were weakly associated with bone density, structure, or strength (false discovery rate (FDR) 0.05), and the measured directions of associations of genera were generally consistent across multiple bone measures, supporting a role for microbiota on skeletal homeostasis. However, the associated effect sizes were small (log₂ fold change <±0.35), limiting power to confidently identify these associations even with high resolution skeletal imaging phenotypes, and we assessed the resulting implications for the design of future cohort-based studies. As in analogous examples from genome-wide association studies, we find that larger cohort sizes will likely be needed to confidently identify associations between the fecal microbiota and skeletal health relying on 16S sequencing. Our findings bolster the view that the gut microbiome is associated with clinically important measures of bone health, while also indicating the challenges in the design of cohort-based microbiome studies.

Keywords

Aging; analysis/quantitation of bone; epidemiology; systems biology - bone interactions

Introduction

Complex multicellular organisms exist as a symbiosis between the host and resident microbiota. The largest portion of the human microbiota is in the gut, which is also responsible for the majority of immune, metabolic, and biochemical host-microbial interactions (1). As a result, the composition of the gut microbiota has been linked to a wide variety of normal physiological processes outside the intestine, disturbances have been implicated in the causation of numerous disorders, and there may be many opportunities to therapeutically alter microbiota composition or function (2). Due to the diversity of microbial chemistry and the lifelong nature of association with a personalized gut microbiome, gut microbes could participate in health phenotypes that might be unexpected, such as kidney stone formation, inorganic detoxification, and developmental processes (3, 4).

Skeletal homeostasis appears to be affected by the gut microbiota (5-7). A number of critical experiments in laboratory animals have revealed skeletal effects of changes in gut microbiota composition. Germ-free mice can have altered bone characteristics, including increased bone density (8-11). However, these findings are not consistently present (12-14), and as with many gnotobiotic animal settings, different outcomes are potentially related to

the background strains of the animals used, the protocols for the studies, the age or sex of animals, housing conditions, etc. Similarly, antibiotic treatments have been shown to have a range of effects on bone density; antibiotic-treated mice appear to have higher bone density in some experiments (12), but other reports include adverse (15, 16) or no effects (17). Finally, a number of experiments have demonstrated that probiotic preparations, commonly including one or more *Lactobacillus* species, may positively affect mineral metabolism or bone health (18-21).

Despite the growing series of animal experiments that demonstrate an effect of the gut microbiome on skeletal health, there are very few studies in humans that examine whether there are associations between the gut microbiome and skeletal status. Wang et al. (22) reported the relationship between categories of BMD (primary osteoporosis, osteopenia, normal based on DXA) in a very small number of participants (3 male and 15 female patients), but there are no studies in larger cohorts. Jones et al. recently summarized compelling studies that suggest that the critical adverse skeletal effects of oophorectomy, and the positive effects of some probiotics, are at least in part mediated via alterations the gut microbiota, gut permeability, and immune function (13). The mechanisms that mediate the interaction between the gut microbiota and bone are undoubtedly complex, but may involve combinations of immune function and inflammation (12, 23-25), changes in endocrine factors (26, 27), nutritional effects and prebiotics (28-34), secreted microbial products (35, 36), effects mediated by exercise (37), and others (5, 14, 20, 27, 38, 39).

We thus aimed to test for covariation between the gut microbiome and skeletal biology in humans by studying the associations of the fecal microbiome with multiple measures of bone density, microarchitecture, and strength in a large cohort of community dwelling older men. We obtained stool samples from 831 participants in MrOS - a longitudinal observational study of musculoskeletal health and aging (40, 41). We then performed 16S rRNA amplicon sequencing, and tested for associations between the abundance of microbial clades and human bone measures obtained with high resolution peripheral quantitative computed tomography (HR-pQCT) and dual energy x-ray absorptiometry (DXA). We viewed this experiment as exploratory and hypothesis generating. As the resulting associations were of generally small effect, we utilized this experience to explore the implications for the design of future cohort-based microbiome studies specifically planned to confidently identify microbiome compositions that are associated with skeletal traits.

Methods

Participants: The MrOS cohort

The Osteoporotic Fractures in Men Study (MrOS) is a prospective study of 5,994 community-dwelling older men, recruited at six clinical sites in the U.S between 2000 and 2002. The cohort and recruitment methods have been previously described (40, 41). At baseline, participants were at least 65 years of age, able to consent, walked without assistance of another person, and did not have bilateral hip replacement or any condition that in the judgment of the site investigator that would likely impair participation in the study. The Institutional Review Boards at all sites reviewed and approved the study and all

participants provided informed consent. All surviving active MrOS participants were invited to participate in a Year 14 visit occurring between May 2014 and May 2016.

As part of a microbiome study, 1,329 of the men participating in the visit were invited to provide a stool sample starting in 2015. Some men (347) refused or were deemed ineligible to collect specimens because of cognitive or physical limitations that interfered with the sample collection, 982 participants agreed, 951 specimens were obtained, 31 specimens were inadequate, and one did not pass the pipeline validation code after genotyping; thus, 919 had complete microbiome data available (Figure 1). There were a few small but significant differences between those who agreed and those who refused: the men who agreed were slightly younger, more active, stronger, and more often reported good or excellent health than those who declined (42). Participants were also queried about antibiotic use in the 30 days prior to sample collection, and the 66 (7.2%) who reported use were excluded from the analyses. Finally, complete covariate data were missing in 22. Thus, the final analyses included 831 participants.

Microbiome sample collection, 16S rRNA gene sequencing, and sequence data processing

As described previously (42), after collection instructions were provided during the in-person clinic visit, fecal samples were collected by study participants at their homes using the Omnigene Gut collection kit (OMR-200, DNA Genotek, Ottawa, Canada) that includes a stabilizing reagent to preserve stool DNA at ambient temperature for up to 60 days (43, 44). Stool collection was within 1 month after the study visit for 98.3% of participants, and others were within 3 months. Samples were mailed directly to Portland and stored at -80°C until DNA extraction. The specimens were sent to the Alkek Center for Metagenomics and Microbiome Research (CMMR), Baylor College of Medicine in Houston, TX, where the samples were aliquoted and genomic bacterial DNA was extracted using the MO BIO PowerSoil DNA Isolation Kit (MO BIO Laboratories, Inc, Carlsbad, CA). The 16S rRNA gene V4 hypervariable region (hereafter abbreviated 16S) was then amplified using primers 515F and 806R by polymerase chain reaction and sequenced on the MiSeq platform using the 2×250 bp paired-end protocol (Illumina, San Diego, CA). The fecal samples were extracted, amplified, and sequenced in two batches: the first 599 men from whom fecal samples were collected, and those 320 from whom samples were received thereafter. Collection, sample processing, and 16S amplification and sequencing procedures were identical in the two subcohorts. The clinical characteristics of the men in the two batches were highly similar, and microbiome diversity was minimally different between the two batches of samples. Specifically, Shannon alpha diversity (for methods, see below) was minimally, albeit significantly, different (batch 1 median Shannon diversity index 3.66, batch 2 3.76, $p=0.006$). Thus, all analyses were performed using the results from the two batches combined, with adjustment for batch in all analyses. In the analyses, no significant batch effects were seen with HRpQCT measures.

Raw amplicon sequences as FASTQs were then bioinformatically processed through the DADA2 (45) workflow in R which has been wrapped in the reproducible bioBakery workflow (46) with AnADAMA2. Briefly, sequences are demultiplexed, after which DADA2 denoising, filtering, trimming, merge, chimera removal, and taxonomic

classification is executed in R using default parameters for paired-end Illumina sequence data. DADA2 assesses the reads for sequence error rates and corrects on a base by base basis. Chimera removal occurs while simultaneously assigning amplicons to amplicon sequence variants (ASVs). Assigning ASV is based on determining which exact sequences were read and how many times each was read. These data are then combined with an error model for the sequencing run, enabling the comparison of similar reads to determine the probability that a given read at a given frequency is not due to sequencer error. Phylogenetic trees are constructed after alignment of sequences using Clustal Omega (47). ASVs are then taxonomically assigned using the SILVA (48) reference database (version 128).

This yielded a total of 12855 denoised ASVs across all samples, with an average of 25487 reads per sample (ranging from 5429 to 70721 reads). ASVs were excluded if present in <10% of samples or at a relative abundance of <0.0001, leaving 2,788 available for analysis after quality control.

Skeletal measures

HR-pQCT measures were performed using Scanco XtremeCT II machines (Scanco Medical AG, Brüttisellen, Switzerland) as previously described (49). Centrally trained operators acquired scans of the distal radius (9 mm from the articular surface), distal tibia (22 mm from the articular surface), and diaphyseal tibia (centered at 30% of tibial length, as measured externally from the tibial plateau to the tibial malleolus). The radius from the nondominant arm and the tibia from the ipsilateral leg were scanned except in the case of prior fracture, metal shrapnel or implant, or recent non-weight-bearing status persisting >6 weeks; 8% of radius scans and 8% of tibia scans were performed on the dominant limb. Machines were calibrated prior to being used in the present study, and a single cross-calibration density phantom was circulated among the study sites.

Centralized quality assurance and standard analysis of all image data, including micro-finite element analysis (μ FEA), was performed. A central observer read all images for motion artifacts and used an established semiquantitative five-point grading system (1 = superior, 5 = poor) to score image quality; images of grade 4 or 5 were deemed to be of insufficient quality and were excluded from the analytic data set (97% of scans had image grade ≤ 3). The final analyses included 792 radial, 801 distal tibial and 732 proximal tibial images (Figure 1).

A fully automated analysis pipeline was developed to segment the radius and tibia for quantification of bone density and microarchitecture. Volumetric BMD and cross-sectional area of the total, cortical, and trabecular compartments were measured. Cortical porosity and thickness, and trabecular thickness and number were calculated directly. Linear elastic μ FEA of a 1% uniaxial compression was performed using a homogenous elastic modulus of 10 GPa and a Poisson's ratio of 0.3 (Scanco FE Software v1.12; Scanco Medical). The failure load was estimated by calculation of the reaction force at which 7.5% of the elements exceed a local effective strain of 0.7%. The measures were identified as suggested by Whittier et al (50). The precision of HR-pQCT measures have been reported (51, 52). Since associations of bone measures with the gut microbiome may be different depending on the specific site of measurement (e.g. cortical vs trabecular, radial vs tibial)

we examined these associations using several types of bone measures. In that analytical context, the intercorrelations between measurement sites should also be considered; those intercorrelations are shown in Supplemental Figure 1.

Hip DXA scans were performed on Hologic 4500 scanners (Hologic, Waltham, MA, USA) as previously described (53). Centralized quality-control procedures, certification of DXA operators, and standardized procedures for scanning were used to ensure reproducibility of DXA measurements. The coefficient variation of the MrOS DXA scanners estimated using a central phantom ranged from 0.3% to 0.7% for the total hip. Each participant's right hip was scanned unless there was a fracture, implant, hardware, or other problem preventing the right hip from being scanned; in those instances, the left hip was scanned.

Other measurements

Body weight was measured on balance beam or digital scales; height was measured using a Harpenden stadiometer (Holtain, Dyfed, UK). Body mass index (BMI) was calculated as weight (kg)/height (m²). Medical history included self-reported physician diagnosis of diabetes. Alcohol use was assessed as 0-5 or 6 drinks/week. All men completed questionnaires and were interviewed about health status, including an assessment of life space (54). Each self-reported his health as excellent/good, fair, poor, or very poor. The PASE questionnaire (Physical Activity Scale for the Elderly) was used to estimate physical activities (55). All prescription medications were recorded in an electronic medication inventory database and matched to its ingredients based on the Iowa Drug Information Service drug vocabulary (College Pharmacy, University of Iowa, Iowa City, IA, USA) (56). The composite life space score, a comprehensive measure of mobility that captures an individual's movement in their environment, was assessed by questionnaire (54). Diet was assessed using the Block dietary questionnaire, and the response data was reduced using factor analyses to derive two major dietary patterns ("Western" and "prudent") (57, 58) that were used as covariates in these analyses. They represent two dietary patterns that are commonly empirically derived based on the underlying distribution of the dietary habits of participants. The Western dietary pattern is characterized by high intakes of processed meat, fried foods, snacks, and desserts. The prudent dietary pattern is characterized by high intakes of fruits, vegetables, legumes, whole grains, and lean meats like chicken. The scores are a linear combination of foods in that pattern; a higher factor score indicates a higher correlation of foods in that dietary pattern.

Statistical analyses

The analytic sample for the present analysis included those 831 men with 16S sequencing and measures of bone parameters who reported no use of antibiotics within the past month and complete data for the adjustment covariates (Figure 1). The Kruskal–Wallis test was used to determine differences in categorical demographic and clinical covariates such as age, race, diet, body mass index, and clinic site.

Diversity analyses: Alpha diversity (estimated using the Shannon index) is a measure of the richness of a microbial community (total number of taxa and evenness of the taxa present) in a single sample. We tested for alpha diversity associations with HR-pQCT

measures using linear regression controlling for covariates (age, race, clinic site, diet, and BMI) (57, 59). Beta diversity estimates (measures of the differences in microbial abundance between 2 samples) included Bray-Curtis dissimilarity and unweighted and weighted UniFrac distance metrics. Weighted-UniFrac takes into account the relative abundance of species/taxa shared between samples, whereas unweighted-UniFrac only considers presence/absence. Principle coordinates analysis (PCoA) was used to analyze and visualize patterns of beta diversity in sequence count data. To prepare the analytic data set from raw ASV count data we agglomerated ASVs to the genus level, filtered the resulting taxa (requiring at least 10% prevalence and 0.0001 abundance) and used relative abundance to normalize, i.e. control for differences in library sizes. We did not rarefy to equalize counts for any of the analyses. The final dataset included 122 total ASVs. Permutation analysis of variance (PERMANOVA) as implemented by the *adonis* function in R *vegan* was used to test the statistical significance of microbial community dissimilarities across continuous bone measures with 999 permutations.

Associations between microbial abundance and skeletal measures: For association analyses, counts were filtered and then agglomerated (final dataset, 129 ASVs). Generalized linear models as implemented in MaAsLin 2 (Microbiome Multivariable Association with Linear Models) were used to test for associations between skeletal measures and specific microbial genera (60). Bacterial count data were normalized and transformed according to methods available in the package (CSS normalization and log transformation). The default MaAsLin 2 parameters were applied (taxonomic feature prevalent in a minimum of 10% of all samples, minimum percentage relative abundance 0.01%, $P < 0.05$, $q < 0.25$). Models used each genus as the outcome (all ASVs for the same genus were agglomerated) and the standardized phenotypes as predictors of interest; i.e. the association of the relative abundance of the microbe per SD increase in the bone phenotype. False discovery rate (FDR) values for discovery using MaAsLin 2 were calculated with respect to the set of associations for each skeletal measurement site individually (Table 2). All models were additionally adjusted for age, race, clinic site, diet, BMI, batch number, PASE score, self-rated overall health, self-rated diabetes, alcohol use, and number of medications used.

We also evaluated our confidence in the associations of genera with skeletal measures using established approaches for examining local FDR considerations that are similar to those applied to genome association studies (61-63), essentially by examining the overall distribution of z-scores of associations to determine the degree to which effect estimates in the tails of the z-score distribution were expected (followed the theoretical normal null distribution) or unexpected (likely to be non-null). We used these analyses to estimate the attained power of our study design, and finally considered the power needs of similar future studies. In fact, the observed Z-score distribution has somewhat long and asymmetrically dense tails (Supplemental Figure 2), suggesting that at least a proportion of these extreme values (either positive or negative) may represent meaningful associations. Detailed methods for these analyses of the distribution of z-scores are provided in Supplemental Statistical Methods. Briefly, we performed a finite mixture analysis of the association z-scores, assuming a three-component mixture where the central component (centered near zero) is normally distributed and represents the "null" associations, and the outer (extremal right and left) components represent the admixture of "non-null" associations hidden within the

tails of the observed distribution. Based on the empirical spread of all the z-scores (from all bone measures) taken together, we fit a normal distributional curve to the central ("null") component and extrapolated that curve into the tails to yield the distribution of z-scores that would be expected under the null hypothesis (of no association), given the observed data. At each end of the mixture distribution of the observed z-scores, we identified regions in the tails that showed higher density than expected from the normal null distribution and were thus more likely to contain high concentrations of non-null associations. These components of the mixture were considered "enrichments" (of extreme positive or negative values occurring at higher rates than expected for null associations). To clarify the power of our study design and the challenges in interpreting "significant" effects that may or may not represent real associations, we estimated the expected local FDR value for a non-null effect appearing in an enriched area of one of the tails of the z-score distribution. The local FDR differs from the typical FDR adjustments applied to p-values in that the local FDR applies to each z-score individually, and is conditioned on that value; it is interpretable as an empirical Bayes *probability* that the specific z-score of interest is a false finding, i.e. that it comes from the null distribution rather than from a component of enrichment (typical FDR adjustments, e.g. the Benjamini-Hochberg algorithm, only estimate the *average* false discovery rate expected across all "discoveries" with a p-value more extreme than some cutoff; they are not true probabilities and do not apply to individual results). The expected local FDR for a non-null effect is the average local FDR value for z-scores coming strictly from an enrichment, given the estimated locations of the enrichment components. The higher the expected local FDR value, the lower the power to discover true effects using FDR-based cutoff criteria (because the probability that any given finding comes from the null distribution instead is high on average).

Visualization of clusters of phenotypes and taxa of interest: Differences in microbial composition per standard deviation of each HR-pQCT measure were assessed using the multivariable MaAsLin 2 analysis, and the \log_2 fold change and FDR-adjusted p-value for each taxon were used to create a heatmap. We performed a hierarchical clustering analysis to group associations found using MaAsLin 2 across the variety of skeletal measures (HR-pQCT and DXA). We used Cronbach's alpha to assess the coherence of associations within clusters; we probed the volatility of the alpha coefficient to small amounts of shuffling of the components of the clusters.

Analyses were performed using R version 3.6.0 for processing the microbiome data, diversity analysis were performed using the Phyloseq package in R, regression analysis of the ASVs with the bone phenotypes were performed using the MaAsLin 2 package in R, heatmaps were created using the heatmap2 package in R, Stata version 16.1 was used for the finite-mixture analysis of association z-scores and meta-analysis of effect sizes across measurement sites, and SAS version 9.4 was used for creating summary tables.

Results

Participants.

Samples adequate for amplicon sequencing were available from 919 men. A number of men were excluded for use of antibiotics and some outcome measurements were missing

in additional participants (Figure 1); the final analyses included data from 831. The clinical characteristics of these men are provided in Table 1, where we also provide a summary of the diversity measures and the average number of amplicon sequence variants (ASVs) detected. Within this population, none of the skeletal measures were significantly associated with alpha diversity (Shannon diversity; Supplemental Table 1) or beta diversity (examples in Supplemental Figure 3).

Associations between genera and skeletal measures.

After multivariable adjustment, we identified possible associations between skeletal measures and specific bacterial genera (Table 2). At MaAsLin 2-derived FDR = 0.05, the abundances of 4 genera were associated with one or more skeletal measurement sites: greater *Anaerofilum* abundance was associated with lower radial and tibial density, radial and tibial strength, tibial cortical thickness and cortical porosity, and hip DXA BMD; *Methanomassiliicoccus* was associated with greater distal tibial cortical porosity. On the other hand, *Ruminiclostridium 9* was associated with less distal tibial cortical porosity, and *Tyzzarella* with greater tibial density measures. The abundances of 2 additional genera were associated using a slightly less strict criterion (FDR = 0.1): *Lactobacillus* and *Streptococcus* were both associated with worse bone measures at radial and tibial sites. As this was intended to be an exploratory analysis, Table 2 also includes associations with FDR up to 0.25.

Patterns of associations across skeletal phenotypes.

The availability of a number of complementary microarchitectural and density measurements at three skeletal sites allowed us to explore the patterns of their joint associations with microbial taxa (Figure 2). Several overall patterns were apparent. First, there was a group of genera with generally positive associations with measures of bone density, microarchitecture, and failure load (the top cluster in Figure 2), and another with generally negative associations (the bottom cluster in Figure 2).

Second, the direction of associations with measures of cortical porosity tended to be similar, particularly at the two tibial sites (Figure 2, right columns) and were frequently in the opposite direction to the measures of density or structure, as might be expected if perilacunar remodeling was broadly and similarly affected throughout the skeleton, as greater porosity would be expected to lower bone density. Correspondingly, associations between genera and porosity were found to be highly consistent across genera (Cronbach's alpha = 0.75) and were similarly consistent for non-porosity measures (Cronbach's alpha = 0.95). The correlations between measures of cortical porosity and other skeletal measures were low (Supplemental Figure 1) and did not explain the grouping of associations between genera and cortical porosity, and the fact that they are often in the opposite direction from density measures.

Finally, the associations of genera with measures of cortical thickness, cortical density and total density were similar at the distal radius, distal tibia, and diaphyseal tibia, and clustered together. Failure load at the diaphyseal tibia, determined primarily by cortical bone characteristics, is also in this cluster. Associations with measures of trabecular density and

number, and failure load at the distal radius and tibia, tended to be similar, and to represent a separate cluster. The associations of genera abundance with total hip BMD assessed by DXA were generally comparable to those of trabecular measures and failure load assessed by HR-pQCT. These findings are possibly consistent with the correlations between the skeletal measures.

Power and sample size considerations.

We used our results to estimate the attained power of the current study to identify associations of microbiota with skeletal traits. From the overall distribution of the MaAsLin 2-derived z-scores of all associations tested (Figure 3) we identified areas of the tails of the distribution that contained a greater enrichment of effects than would be expected from a normal null distribution comprising all the z-scores considered together (Figure 3). Based on the finite mixture analysis of this distribution (see Statistical Analyses), the expected (i.e. mixture-model-based average) local FDR value across all the enriched portions of the tails (red and green areas in Figure 3) was very large at 0.76. This number represents the average probability of false discovery that would apply to a z-score selected at random from an enrichment area. Since most of the z-scores in these areas are in regions of high null density (the height of the smooth black line shown in Figure 3), they lack salience as true "discoveries". Moreover, the average local FDR value for even the most extreme members of the distribution, for example the three most extreme negative z-scores and the three most extreme positive z-scores, was 0.33, indicating that even in the furthest reaches of the tails, null density is still sufficiently large that the z-scores found there could very plausibly have come from the null distribution instead of from an enrichment. Note that this somewhat large local FDR estimate for the most extreme z-scores contrasts with the lowest FDR values presented in Table 2. That difference is because the FDR values in Table 2 are tail-averaged (i.e. not local) FDR values calculated by MaAsLin 2 for associations at each skeletal measurement site individually, while the distribution-based estimates here use the full set of associations across all measurement sites together, and also consider local FDR (i.e. at each z-score) rather than average tail FDR (for details about local FDR see Statistical Analyses).

In studies of the association of microbiome composition with skeletal traits, low expected false discovery rate (e.g. $FDR < 0.1$) for promisingly large associations would be desirable. The z-scores observed among the most robust associations in the current studies had estimated \log_2 fold changes of approximately 0.25 per standard deviation of bone measure (Table 2), with standard errors on the order of 0.1. An optimistic power calculation assuming enrichment fraction and effect sizes at similar levels to what was observed in the current study estimates that a sample size of no less than $n = 6,000$ would be required, and upwards of $n = 10,000$ or more may be required if the true effect sizes are smaller, the enrichment fraction is lower, or both.

Discussion

The gut microbiome has been linked to bone homeostasis in a variety of animal experiments, and probiotic supplements have been reported to have beneficial effects on bone density

in postmenopausal women (20), but there are few data concerning what elements of the microbiome might be related to skeletal health in humans. This is the first large study to test for associations between the microbiome and skeletal measures in humans. We found suggestive evidence that the abundances of several bacterial genera were associated with skeletal bone density, microarchitecture, and strength. Four genera were associated with at least one skeletal measure (*Anaerofilum*, *Methanomassiliicoccus*, *Ruminiclostridium 9*, and *Tyzzereella*). Two more (*Lactobacillus* and *Streptococcus*) were associated using somewhat less stringent significance criteria. Most associations were negative, but others suggested higher genera abundance was associated with indices of better skeletal health. This is concordant with dysbiotic shifts in other health conditions, which are frequently not attributable to single organisms, but which can instead modify any one of several "more" or "less beneficial" microbes among different individuals. Moreover, these relationships were generally consistent across a variety of skeletal measurement sites. While the study must be considered exploratory and we cannot be highly confident of the associations we identified, our results should be useful in highlighting specific bacterial genera and pathways that may be of particular interest for additional confirmatory research. The genera we report as having the most robust associations have not, to our knowledge, been previously reported to affect skeletal traits and the mechanisms that might mediate those effects are unknown. More detailed studies using metagenomic data generated by shotgun sequencing (i.e. at the species or functional level) would be better suited to identify biological functions that might be important for bone.

The patterns of the associations of the microbiome with skeletal measures may also be informative. The associations between genera abundance and bone were generally consistent across multiple measurement sites and bone compartments. For instance, genera most strongly associated with lower trabecular BMD were also likely to be associated with lower cortical thickness and higher cortical porosity, and hence lower integral BMD and failure load. This pattern is expected if mediated by a broad-based effect of the microbiome on remodeling balance (either positive or negative) that would drive density and porosity measures in opposite directions. This also suggests that putative influences of microbial composition on bone were not limited to a particular bone compartment but involve widespread remodeling-based mechanisms. To some extent, these consistent associations across measures of several anatomical sites are expected based on underlying correlations among the skeletal measures (Supplemental Figure 1). However, these are sometimes quite weak in isolation (e.g. $r \sim 0.4$), making it difficult to assume they completely explain the consistent associations of microbes with multiple skeletal measures. The associations with distal radial cortical porosity were somewhat less consistent, but the accuracy of cortical porosity measures at the distal radius is questionable. Overall, the consistency and biological plausibility of the patterns of associations argues against statistical artifact, and suggests an underlying microbiome-bone interaction.

A notable finding given the nontrivial size of our sample of participants ($n > 800$) was that there were few strong associations between taxa and skeletal measures. This is again in some ways to be expected, as the routes by which the gut microbiome can affect systemic physiology are relatively narrow. Molecular mechanisms to transmit chemical products or immune activity from the gut throughout the body for bone development and/or maintenance

must of necessity act gradually at best, and in extreme cases the latter will be confounded with overt inflammatory disease. A further challenge in this cohort of older men is that variation in the HRpQCT features also reflects genetic variation and the influence of lifetime exposure to environmental effects. Thus, some of the effects of the microbiome would likely be proportionately small. Our experience with these data suggests that studies similar to this one would have to be large (> 6,000 participants) to provide adequate power to confidently detect meaningful associations. Cohort studies that utilize metagenomic data with resolution exceeding that of our 16S rRNA gene amplicon sequencing might need even greater population sizes to provide sufficient power. These challenges are similar to studies using genetic markers, in which effect sizes of common variants with phenotypes are of extremely small magnitude (64-66). Conversely, studies of more focused patient populations might require fewer numbers. For instance, studies of participants with specific clinical characteristics or studies that exclude variables that might confound the relationships (67), could improve power.

The study had several strengths. MrOS is a well-phenotyped cohort with extensive information about the participants' clinical characteristics. We excluded men with recent exposure to antibiotics, and adjusted for potentially confounding factors. Fecal samples were collected using an approach known to preserve nucleic acids and well-established methods were used for bacterial genotyping and analyses. Standardized bone measures were performed at several skeletal sites using state-of-the-art approaches.

We also identified several important limitations. Most importantly, our strongest associations between bone phenotypes and the gut microbiome were relatively small, and it is clear that even in >800 men these results must be considered exploratory. On the other hand, the genera that we have identified may be very useful as specific targets for future replication studies, especially if adequately powered, and eventually to identify pathways or to design targeted research using mouse models. While the homogeneity of the MrOS aids in analysis, this was as a result a cross sectional evaluation that included only elderly males; we could not evaluate the potential influence of age, sex, or racial/ethnic differences, longitudinal measures of the gut microbiome and the skeleton, and their inter-relationships. On the other hand, although the human microbiome can vary in response to a variety of stimuli, it is generally relatively self-stable over long periods (68, 69). Also, as is often necessary in human populations, the study was observational, and causation cannot be determined from our results. Finally, while we examined multivariable models that included a number of factors that might influence bone, the microbiome or their interactions, there may be additional confounding variables that were not considered.

In summary, we examined associations of genus-level gut microbial components with measures of bone density, microarchitecture, and strength in a cohort of older men. We found several taxa that were nominally associated with bone measures, and a compelling pattern of genus-bone relationships that suggest the influences of microbial composition on bone are not limited to a particular bone compartment but involve more generalized mechanisms. The magnitudes of the associations we detected were not large, but therapeutically targeting the microbiome over many years could still translate into preservation or enhancement of skeletal integrity. Clearly additional studies should involve

larger cohorts of men and women over wider age ranges and/or more causally incisive methods.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We acknowledge the essential contributions of all the MrOS participants, investigators and staff whose dedication, commitment, and contributions make the MrOS study possible. The Osteoporotic Fractures in Men (MrOS) Study was supported by the National Institute on Aging; the National Institute of Arthritis and Musculoskeletal and Skin Diseases; the National Center for Advancing Translational Sciences; the National Institutes of Health Roadmap for Medical Research (grant numbers U01 AG027810, U01 AG042124, U01 AG042139, U01 AG042140, U01 AG042143, U01 AG042145, U01 AG042168, U01 AR066160, and UL1 TR000128). Dr. Kiel's effort and a portion of gut microbiome genotyping were supported by the National Institute of Arthritis and Musculoskeletal and Skin Diseases R01 AR061445

Data availability

The data that support the findings of this study are available upon request. In addition, MrOS data, including the microbiome genotypes, are available at MrOS Online (<https://mrosonline.ucsf.edu/>).

References

1. Qin J, Li R, Raes J, Arumugam M, Burgdorf KS, Manichanh C, et al. A human gut microbial gene catalogue established by metagenomic sequencing. *Nature*. 2010;464(7285):59–65. [PubMed: 20203603]
2. Gilbert JA, Blaser MJ, Caporaso JG, Jansson JK, Lynch SV, Knight R. Current understanding of the human microbiome. *Nat Med*. 2018;24(4):392–400. [PubMed: 29634682]
3. Claus SP, Guillou H, Ellero-Simatos S. The gut microbiota: a major player in the toxicity of environmental pollutants? *NPJ Biofilms Microbiomes*. 2016;2:16003. [PubMed: 28721242]
4. Stanford J, Charlton K, Stefoska-Needham A, Ibrahim R, Lambert K. The gut microbiota profile of adults with kidney disease and kidney stones: a systematic review of the literature. *BMC Nephrol*. 2020;21(1):215. [PubMed: 32503496]
5. Chen YC, Greenbaum J, Shen H, Deng HW. Association Between Gut Microbiota and Bone Health: Potential Mechanisms and Prospective. *J Clin Endocrinol Metab*. 2017;102(10):3635–46. [PubMed: 28973392]
6. Ohlsson C, Sjögren K. Effects of the gut microbiota on bone mass. *Trends Endocrinol Metab*. 2015;26(2):69–74. [PubMed: 25497348]
7. Yan J, Charles JF. Gut Microbiome and Bone: to Build, Destroy, or Both? *Curr Osteoporos Rep*. 2017;15(4):376–84. [PubMed: 28620867]
8. Sjögren K, Engdahl C, Henning P, Lerner UH, Tremaroli V, Lagerquist MK, et al. The gut microbiota regulates bone mass in mice. *J Bone Miner Res*. 2012;27(6):1357–67. [PubMed: 22407806]
9. Li JY, Chassaing B, Tyagi AM, Vaccaro C, Luo T, Adams J, et al. Sex steroid deficiency-associated bone loss is microbiota dependent and prevented by probiotics. *J Clin Invest*. 2016;126(6):2049–63. [PubMed: 27111232]
10. Schwarzer M, Strigini M, Leulier F. Gut Microbiota and Host Juvenile Growth. *Calcif Tissue Int*. 2018;102(4):387–405. [PubMed: 29214457]
11. Yan J, Herzog JW, Tsang K, Brennan CA, Bower MA, Garrett WS, et al. Gut microbiota induce IGF-1 and promote bone formation and growth. *Proc Natl Acad Sci U S A*. 2016;113(47):E7554–e63. [PubMed: 27821775]

12. D'Amelio P, Sassi F. Gut Microbiota, Immune System, and Bone. *Calcif Tissue Int.* 2018;102(4):415–25. [PubMed: 28965190]
13. Jones RM, Mulle JG, Pacifici R. Osteomicrobiology: The influence of gut microbiota on bone in health and disease. *Bone.* 2018;115:59–67. [PubMed: 28433758]
14. Hernandez CJ. The Microbiome and Bone and Joint Disease. *Curr Rheumatol Rep.* 2017;19(12):77. [PubMed: 29101487]
15. Guss JD, Horsfield MW, Fontenele FF, Sandoval TN, Luna M, Apoorva F, et al. Alterations to the Gut Microbiome Impair Bone Strength and Tissue Material Properties. *J Bone Miner Res.* 2017;32(6):1343–53. [PubMed: 28244143]
16. Hathaway-Schrader JD, Steinkamp HM, Chavez MB, Poulides NA, Kirkpatrick JE, Chew ME, et al. Antibiotic Perturbation of Gut Microbiota Dysregulates Osteoimmune Cross Talk in Postpubertal Skeletal Development. *Am J Pathol.* 2019;189(2):370–90. [PubMed: 30660331]
17. Mikkelsen KH, Vilsbøll T, Holst JJ, Hartmann B, Knop FK, Frost M. No changes in levels of bone formation and resorption markers following a broad-spectrum antibiotic course. *BMC Endocr Disord.* 2018;18(1):60. [PubMed: 30180841]
18. Collins FL, Rios-Arce ND, Schepper JD, Parameswaran N, McCabe LR. The Potential of Probiotics as a Therapy for Osteoporosis. *Microbiol Spectr.* 2017;5(4).
19. McCabe LR, Parameswaran N. Advances in Probiotic Regulation of Bone and Mineral Metabolism. *Calcif Tissue Int.* 2018;102(4):480–8. [PubMed: 29453726]
20. Ohlsson C, Sjögren K. Osteomicrobiology: A New Cross-Disciplinary Research Field. *Calcif Tissue Int.* 2018;102(4):426–32. [PubMed: 29079994]
21. Britton RA, Irwin R, Quach D, Schaefer L, Zhang J, Lee T, et al. Probiotic *L. reuteri* treatment prevents bone loss in a menopausal ovariectomized mouse model. *J Cell Physiol.* 2014;229(11):1822–30. [PubMed: 24677054]
22. Wang J, Wang Y, Gao W, Wang B, Zhao H, Zeng Y, et al. Diversity analysis of gut microbiota in osteoporosis and osteopenia patients. *PeerJ.* 2017;5:e3450. [PubMed: 28630804]
23. Collins FL, Schepper JD, Rios-Arce ND, Steury MD, Kang HJ, Mallin H, et al. Immunology of Gut-Bone Signaling. *Adv Exp Med Biol.* 2017;1033:59–94. [PubMed: 29101652]
24. Belkaid Y, Hand TW. Role of the microbiota in immunity and inflammation. *Cell.* 2014;157(1):121–41. [PubMed: 24679531]
25. Hsu E, Pacifici R. From Osteoimmunology to Osteomicrobiology: How the Microbiota and the Immune System Regulate Bone. *Calcif Tissue Int.* 2018;102(5):512–21. [PubMed: 29018933]
26. Markle JG, Frank DN, Mortin-Toth S, Robertson CE, Feazel LM, Rolle-Kampczyk U, et al. Sex differences in the gut microbiome drive hormone-dependent regulation of autoimmunity. *Science.* 2013;339(6123):1084–8. [PubMed: 23328391]
27. Yan J, Takakura A, Zandi-Nejad K, Charles JF. Mechanisms of gut microbiota-mediated bone remodeling. *Gut Microbes.* 2018;9(1):84–92. [PubMed: 28961041]
28. Weaver CM. Diet, gut microbiome, and bone health. *Curr Osteoporos Rep.* 2015;13(2):125–30. [PubMed: 25616772]
29. Thomas RL, Jiang L, Adams JS, Xu ZZ, Shen J, Janssen S, et al. Vitamin D metabolites and the gut microbiome in older men. *Nat Commun.* 2020;11(1):5997. [PubMed: 33244003]
30. Abrams SA, Griffin IJ, Hawthorne KM, Liang L, Gunn SK, Darlington G, et al. A combination of prebiotic short- and long-chain inulin-type fructans enhances calcium absorption and bone mineralization in young adolescents. *Am J Clin Nutr.* 2005;82(2):471–6. [PubMed: 16087995]
31. Weaver CM, Martin BR, Nakatsu CH, Armstrong AP, Clavijo A, McCabe LD, et al. Galactooligosaccharides improve mineral absorption and bone properties in growing rats through gut fermentation. *J Agric Food Chem.* 2011;59(12):6501–10. [PubMed: 21553845]
32. Roberfroid M, Gibson GR, Hoyles L, McCartney AL, Rastall R, Rowland I, et al. Prebiotic effects: metabolic and health benefits. *Br J Nutr.* 2010;104 Suppl 2:S1–63.
33. Pacifici R. Bone Remodeling and the Microbiome. *Cold Spring Harb Perspect Med.* 2018;8(4).
34. Rizzoli R. Nutritional influence on bone: role of gut microbiota. *Aging Clin Exp Res.* 2019;31(6):743–51. [PubMed: 30710248]

35. Lucas S, Omata Y, Hofmann J, Böttcher M, Iljazovic A, Sarter K, et al. Short-chain fatty acids regulate systemic bone mass and protect from pathological bone loss. *Nat Commun.* 2018;9(1):55. [PubMed: 29302038]
36. Tyagi AM, Yu M, Darby TM, Vaccaro C, Li JY, Owens JA, et al. The Microbial Metabolite Butyrate Stimulates Bone Formation via T Regulatory Cell-Mediated Regulation of WNT10B Expression. *Immunity.* 2018;49(6):1116–31.e7. [PubMed: 30446387]
37. McCabe LR, Irwin R, Tekalur A, Evans C, Schepper JD, Parameswaran N, et al. Exercise prevents high fat diet-induced bone loss, marrow adiposity and dysbiosis in male mice. *Bone.* 2019;118:20–31. [PubMed: 29604350]
38. Charles JF, Ermann J, Aliprantis AO. The intestinal microbiome and skeletal fitness: Connecting bugs and bones. *Clin Immunol.* 2015;159(2):163–9. [PubMed: 25840106]
39. Collins FL, Irwin R, Bierhalter H, Schepper J, Britton RA, Parameswaran N, et al. *Lactobacillus reuteri* 6475 Increases Bone Density in Intact Females Only under an Inflammatory Setting. *PLoS One.* 2016;11(4):e0153180. [PubMed: 27058036]
40. Orwoll E, Blank JB, Barrett-Connor E, Cauley J, Cummings S, Ensrud K, et al. Design and baseline characteristics of the osteoporotic fractures in men (MrOS) study--a large observational study of the determinants of fracture in older men. *Contemp Clin Trials.* 2005;26(5):569–85. [PubMed: 16084776]
41. Blank JB, Cawthon PM, Carrion-Petersen ML, Harper L, Johnson JP, Mitson E, et al. Overview of recruitment for the osteoporotic fractures in men study (MrOS). *Contemp Clin Trials.* 2005;26(5):557–68. [PubMed: 16085466]
42. Abrahamson M, Hooker E, Ajami NJ, Petrosino JF, Orwoll ES. Successful collection of stool samples for microbiome analyses from a large community-based population of elderly men. *Contemp Clin Trials.* 2017;7:158–62. [PubMed: 29250605]
43. Anderson EL, Li W, Klitgord N, Highlander SK, Dayrit M, Seguritan V, et al. A robust ambient temperature collection and stabilization strategy: Enabling worldwide functional studies of the human microbiome. *Sci Rep.* 2016;6:31731. [PubMed: 27558918]
44. Song SJ, Amir A, Metcalf JL, Amato KR, Xu ZZ, Humphrey G, et al. Preservation Methods Differ in Fecal Microbiome Stability, Affecting Suitability for Field Studies. *mSystems.* 2016;1(3).
45. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJ, Holmes SP. DADA2: High-resolution sample inference from Illumina amplicon data. *Nat Methods.* 2016;13(7):581–3. [PubMed: 27214047]
46. McIver LJ, Abu-Ali G, Franzosa EA, Schwager R, Morgan XC, Waldron L, et al. bioBakery: a meta-omic analysis environment. *Bioinformatics.* 2018;34(7):1235–7. [PubMed: 29194469]
47. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol.* 2011;7:539. [PubMed: 21988835]
48. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* 2013;41(Database issue):D590–6. [PubMed: 23193283]
49. Langsetmo L, Peters KW, Burghardt AJ, Ensrud KE, Fink HA, Cawthon PM, et al. Volumetric Bone Mineral Density and Failure Load of Distal Limbs Predict Incident Clinical Fracture Independent HR-pQCT BMD and Failure Load Predicts Incident Clinical Fracture of FRAX and Clinical Risk Factors Among Older Men. *J Bone Miner Res.* 2018;33(7):1302–11. [PubMed: 29624722]
50. Whittier DE, Boyd SK, Burghardt AJ, Paccou J, Ghasem-Zadeh A, Chapurlat R, et al. Guidelines for the assessment of bone density and microarchitecture in vivo using high-resolution peripheral quantitative computed tomography. *Osteoporos Int.* 2020;31(9):1607–27. [PubMed: 32458029]
51. Chiba K, Okazaki N, Kurogi A, Isobe Y, Yonekura A, Tomita M, et al. Precision of Second-Generation High-Resolution Peripheral Quantitative Computed Tomography: Intra- and Intertester Reproducibilities and Factors Involved in the Reproducibility of Cortical Porosity. *J Clin Densitom.* 2018;21(2):295–302. [PubMed: 28256308]

52. Manske SL, Zhu Y, Sandino C, Boyd SK. Human trabecular bone microarchitecture can be assessed independently of density with second generation HR-pQCT. *Bone*. 2015;79:213–21. [PubMed: 26079995]
53. Cawthon PM, Ewing SK, Mackey DC, Fink HA, Cummings SR, Ensrud KE, et al. Change in hip bone mineral density and risk of subsequent fractures in older men. *J Bone Miner Res*. 2012;27(10):2179–88. [PubMed: 22648990]
54. Mackey DC, Cauley JA, Barrett-Connor E, Schousboe JT, Cawthon PM, Cummings SR. Life-space mobility and mortality in older men: a prospective cohort study. *J Am Geriatr Soc*. 2014;62(7):1288–96. [PubMed: 24934163]
55. Washburn RA, Smith KW, Jette AM, Janney CA. The Physical Activity Scale for the Elderly (PASE): development and evaluation. *J Clin Epidemiol*. 1993;46(2):153–62. [PubMed: 8437031]
56. Pahor M, Chrischilles EA, Guralnik JM, Brown SL, Wallace RB, Carbonin P. Drug data coding and analysis in epidemiologic studies. *Eur J Epidemiol*. 1994;10(4):405–11. [PubMed: 7843344]
57. Shikany JM, Demmer RT, Johnson AJ, Fino NF, Meyer K, Ensrud KE, et al. Association of dietary patterns with the gut microbiota in older, community-dwelling men. *Am J Clin Nutr*. 2019;110(4):1003–14. [PubMed: 31504105]
58. Rogers TS, Harrison S, Judd S, Orwoll ES, Marshall LM, Shannon J, et al. Dietary patterns and longitudinal change in hip bone mineral density among older men. *Osteoporos Int*. 2018;29(5):1135–45. [PubMed: 29450584]
59. Shardell M, Parimi N, Langsetmo L, Tanaka T, Jiang L, Orwoll E, et al. Comparing Analytical Methods for the Gut Microbiome and Aging: Gut Microbial Communities and Body Weight in the Osteoporotic Fractures in Men (MrOS) Study. *J Gerontol A Biol Sci Med Sci*. 2020;75(7):1267–75. [PubMed: 32025711]
60. Morgan XC, Tickle TL, Sokol H, Gevers D, Devaney KL, Ward DV, et al. Dysfunction of the intestinal microbiome in inflammatory bowel disease and treatment. *Genome Biol*. 2012;13(9):R79. [PubMed: 23013615]
61. Pawitan Y, Michiels S, Koscielny S, Gusnanto A, Ploner A. False discovery rate, sensitivity and sample size for microarray studies. *Bioinformatics*. 2005;21(13):3017–24. [PubMed: 15840707]
62. Chakraborty A, Jiang G, Boustani M, Liu Y, Skaar T, Li L. Simultaneous inferences based on empirical Bayes methods and false discovery rates in eQTL data analysis. *BMC Genomics*. 2013;14 Suppl 8(Suppl 8):S8.
63. Jiang W, Yu W. Controlling the joint local false discovery rate is more powerful than meta-analysis methods in joint analysis of summary statistics from multiple genome-wide association studies. *Bioinformatics*. 2017;33(4):500–7. [PubMed: 28011772]
64. Nishino J, Ochi H, Kochi Y, Tsunoda T, Matsui S. Sample Size for Successful Genome-Wide Association Study of Major Depressive Disorder. *Front Genet*. 2018;9:227. [PubMed: 30002671]
65. Holland D, Wang Y, Thompson WK, Schork A, Chen CH, Lo MT, et al. Estimating Effect Sizes and Expected Replication Probabilities from GWAS Summary Statistics. *Front Genet*. 2016;7:15. [PubMed: 26909100]
66. Hong EP, Park JW. Sample size and statistical power calculation in genetic association studies. *Genomics Inform*. 2012;10(2):117–22. [PubMed: 23105939]
67. Zablocki RW, Schork AJ, Levine RA, Andreassen OA, Dale AM, Thompson WK. Covariate-modulated local false discovery rate for genome-wide association studies. *Bioinformatics*. 2014;30(15):2098–104. [PubMed: 24711653]
68. Sommer F, Anderson JM, Bharti R, Raes J, Rosenstiel P. The resilience of the intestinal microbiota influences health and disease. *Nat Rev Microbiol*. 2017;15(10):630–8. [PubMed: 28626231]
69. Mehta RS, Abu-Ali GS, Drew DA, Lloyd-Price J, Subramanian A, Lochhead P, et al. Stability of the human faecal microbiome in a cohort of adult men. *Nat Microbiol*. 2018;3(3):347–55. [PubMed: 29335554]

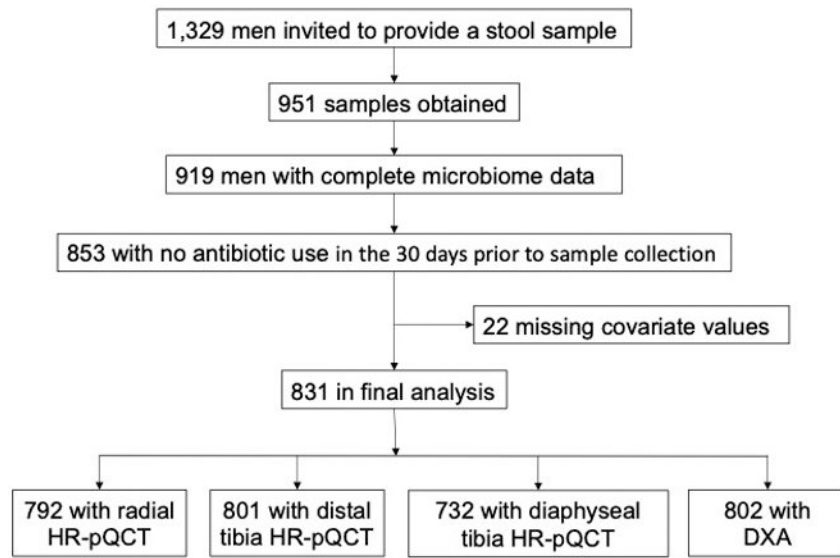


Figure 1. Selection of participants; fully described in Methods.

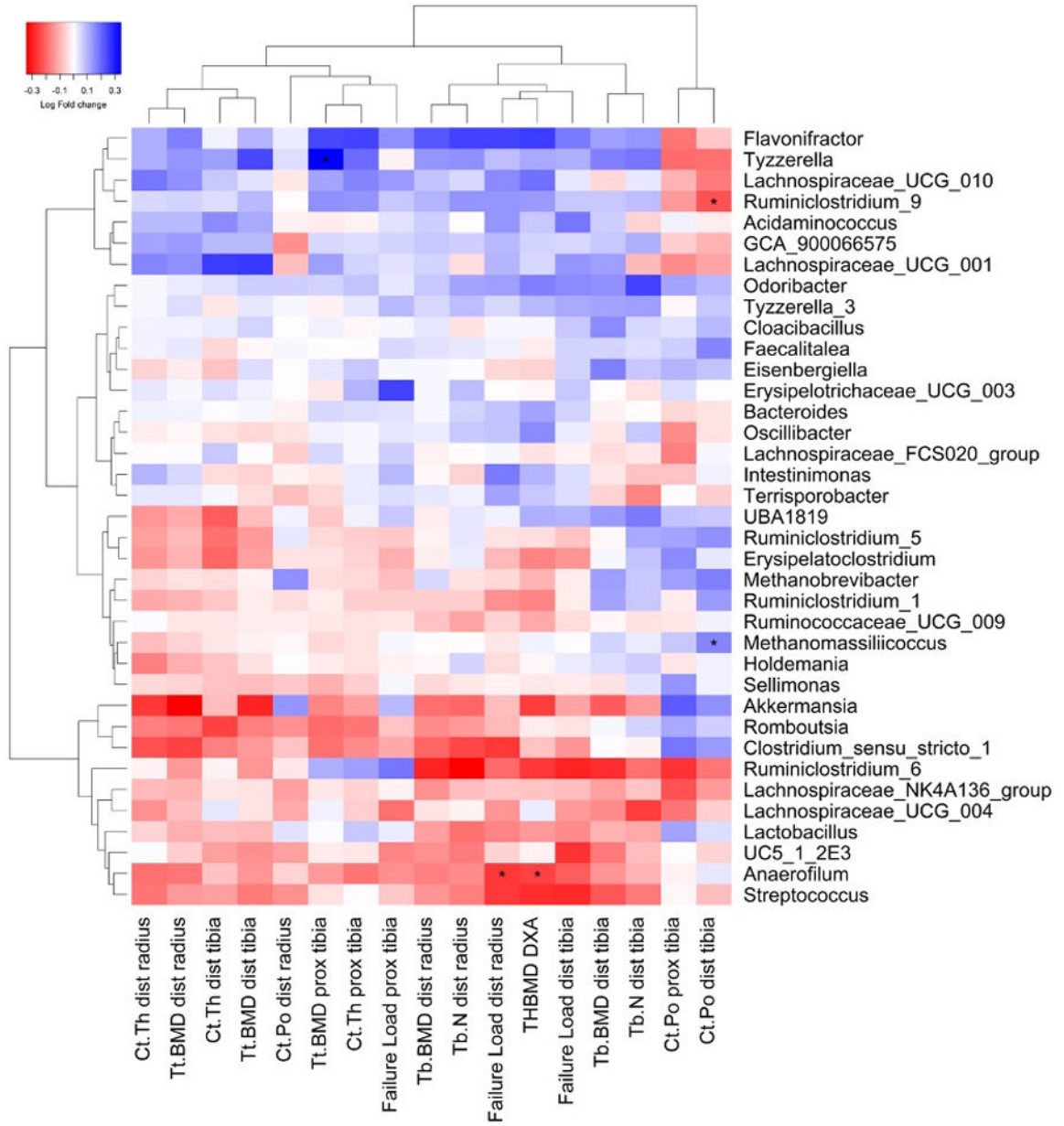


Figure 2. Associations between gut microbial taxa and bone measures. Differences in microbial composition per standard deviation of each bone measure were assessed using generalized linear models in MaAsLin 2, and the \log_2 fold change and FDR-adjusted p-value for each taxon were used to create the heatmap. Only those associations with FDR < 0.25 are included. A hierarchical clustering was performed to group associations across the variety of skeletal measures. *p < 0.05

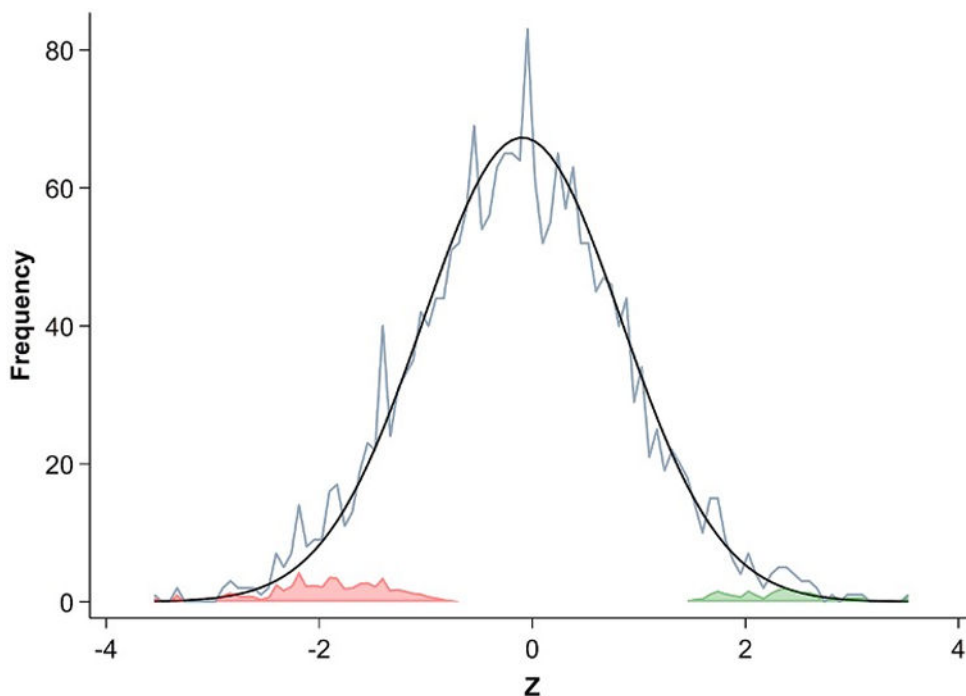


Figure 3. Finite mixture analysis of association z-score distribution. The smooth black curve is an estimate of the hypothetical normal distribution of null associations, based on the observed range and density of z-scores near zero. The jagged light-blue line shows the distribution (in bins of $1/14 \sim 0.0714$ z-score units) of all actual, observed z-scores of the taxa-skeletal measure associations. The colored areas represent the number of taxa-skeletal measure associations not expected to occur within the hypothetical normal distribution of null effects, and thus representing potentially true associations: red the number of potentially true negative effects, and green the potentially true positive effects. The false discovery rates at each location on the z scale can be visually estimated by comparing the height of the red or green areas at that location to the height of the overall histogram (light-blue line) at the same location. See Supplemental Statistical Methods for a fuller description of the finite mixture model.

Table 1.

Cohort characteristics. N= 831. Mean (SD) or N (%).

Age (years)	84.2 (4)
BMI (kg/m ²)	27 (3.7)
PASE score	123 (66)
Self-rated health status	
Excellent/good N (%)	744 (90)
Fair, poor or very poor N (%)	87 (10)
Diabetes N (%)	124 (15)
Alcohol	
0-7 (drinks/week) N (%)	503 (60)
>7 (drinks/week) N (%)	326 (40)
Western diet	-0.03 (-1.0)
Prudent diet	-0.01 (-1.0)
Total number of medications taken	9 (6)
Composite life-space score	85 (22)
Shannon alpha diversity*	3.6 (0.6)
Observed amplicon sequence variants (ASVs)* N (range)	144 (48)
Min Bray-Curtis dissimilarity*	0.6 (0.05)
Distal radial	
Failure load (FL)(kN)	4921 (1344)
Total BMD (Tt.BMD)(mg HA/cm ³)	276 (60)
Trabecular BMD (Tb.BMD)(mg HA/cm ³)	170 (39)
Trabecular number (Tb.N)(1/mm)	1.4 (0.2)
Cortical thickness (Ct.Th)(mm)	1.0 (0.2)
Cortical porosity (Ct.Po)(%)	0.02 (0.01)
Distal tibial	
Failure load (FL)(kN)	4921 (1344)
Total BMD (Tt.BMD)(mg HA/cm ³)	276 (60)
Trabecular BMD (Tb.BMD)(mg HA/cm ³)	170 (39)
Trabecular number (Tb.N)(1/mm)	1.4 (0.2)
Cortical thickness (Ct.Th)(mm)	1.0 (0.2)
Cortical porosity (Ct.Po)(%)	0.02 (0.01)
Total hip BMD; gm/cm ²	0.94 (0.15)

Table 2.

Associations between taxa abundance and bone measures from the MaAsLin multivariate adjusted model. Includes all associations with FDR q-values ≤ 0.25 .

GENUS	Mean relative abundance*	Bone phenotype	Log 2 fold change**	FDR q-value***
Anaerofilum	.00028	FL distal radius	-0.262	0.019
		BMD DXA THD	-0.249	0.027
		FL distal tibia	-0.214	0.075
		Ct.Th distal radius	-0.189	0.136
		Tt.BMD distal radius	-0.184	0.148
		Tb.BMD distal radius	-0.166	0.186
		Ct.Th diaphyseal tibia	-0.186	0.195
		Tb.N distal radius	-0.155	0.247
Ruminiclostridium 9	.00321	Ct.Po distal tibia	-0.232	0.027
Methanomassiliicoccus	.00012	Ct.Po distal tibia	0.165	0.027
Tyzzerella	.00084	Tt.BMD diaphyseal tibia	0.338	0.048
		Tt.BMD distal tibia	0.244	0.165
Lactobacillus	.00035	Tb.N distal radius	-0.191	0.076
		FL distal radius	-0.161	0.159
		FL distal tibia	-0.155	0.162
Streptococcus	.00151	FL distal tibia	-0.285	0.100
		BMD DXA THD	-0.275	0.120
		FL distal radius	-0.262	0.145
		Tb.BMD distal tibia	-0.223	0.230
Ruminiclostridium 5	.00314	Ct.Th distal tibia	-0.191	0.115
		Ct.Po distal tibia	0.147	0.250
Sellimonas	.00009	Ct.Po diaphyseal tibia	0.141	0.117
Romboutsia	.00103	Ct.Th distal tibia	-0.253	0.131
Lachnospiraceae FCS020 group	.00013	Ct.Po diaphyseal tibia	-0.165	0.139
Ruminococcaceae UCG-009	.00008	Tb.N distal radius	-0.124	0.141
	.00008	Total hip BMD DXA	-0.114	0.193
UBA1819	.00233	Ct.Th distal tibia	-0.219	0.148
Cloacibacillus	.00027	Tb.BMD distal tibia	0.154	0.150
Flavonifractor	.00161	Total hip BMD DXA	0.261	0.151
		FL distal radius	0.253	0.164
		Tb.N distal radius	0.253	0.166
		Ct.Th diaphyseal tibia	0.252	0.214
		Tb.BMD distal radius	0.222	0.230
		Tt.BMD diaphyseal tibia	0.245	0.231

GENUS	Mean relative abundance [*]	Bone phenotype	Log 2 fold change ^{**}	FDR q-value ^{***}
Ruminiclostridium 6	.01081	Tb.N distal radius	-0.340	0.166
Erysipelatoclostridium	.00047	Ct.Th distal tibia	-0.205	0.154
Holdemania	.00020	Ct.Th distal radius	-0.171	0.166
UC5-1-2E3	.00161	FL distal tibia	-0.270	0.172
Akkermansia	.02971	Tt.BMD distal radius	-0.344	0.182
Eisenbergiella	.00025	Tb.BMD distal tibia	0.169	0.188
GCA-900066575	.00017	Ct.Po distal radius	-0.151	0.193
Ruminiclostridium 1	.00028	Total hip BMD DXA	-0.164	0.200
Lachnospiraceae UCG-001	.00290	Tt.BMD distal tibia	0.263	0.202
		Ct.Th distal tibia	0.256	0.225
Odoribacter	.00350	Tb.N distal tibia	0.248	0.205
Clostridium sensu stricto 1	.00310	FL distal radius	-0.262	0.210
		Tt.BMD distal radius	-0.254	0.227
Faecalitalea	.00030	Ct.Po distal tibia	0.162	0.212
Tyzzrella 3	.00013	Tb.N distal tibia	0.129	0.221
Intestinimonas	.00191	FL distal radius	0.176	0.221
Methanobrevibacter	.00080	Ct.Po distal tibia	0.171	0.223
Erysipelotrichaceae UCG-003	.00131	FL diaphyseal tibia	0.250	0.225
Lachnospiraceae UCG-004	.00268	Tb.N distal tibia	-0.258	0.226
Acidaminococcus	.00099	FL distal tibia	0.185	0.233
Lachnospiraceae NK4A136 group	.01274	Ct.Po diaphyseal tibia	-0.231	0.241
Lachnospiraceae UCG-010	.00096	Total hip BMD DXA	0.186	0.242
Bacteroides	.25190	Total hip BMD DXA	0.122	0.243
Terrisporobacter	.00034	Tb.N distal tibia	-0.162	0.243
Oscillibacter	.00329	Total hip BMD DXA	0.157	0.246

* Mean relative abundance is the abundance of each genus (from raw counts) as a proportion of all genera.

** Log 2 fold change is the log-ratio of the association of the relative abundance of the microbe per SD increase in the bone phenotype.

*** The q-value is the expected proportion of false positives among all features at least as extreme as the observed one.