Genome-Wide Association Studies of Skeletal Phenotypes: What We Have Learned and Where We Are Headed

Yi-Hsiang Hsu and Douglas P. Kiel

Hebrew SeniorLife Institute for Aging Research (Y.-H.H., D.P.K.), Department of Medicine, Division of Gerontology, Beth Israel Deaconess Medical and Harvard Medical School, Boston, Massachusetts 02131; and Molecular and Integrative Physiological Sciences Program (Y.-H.H.), Harvard School of Public Health, Boston, Massachusetts 02115

Context: The primary goals of genome-wide association studies (GWAS) are to discover new molecular and biological pathways involved in the regulation of bone metabolism that can be leveraged for drug development. In addition, the identified genetic determinants may be used to enhance current risk factor profiles.

Evidence Acquisition: There have been more than 40 published GWAS on skeletal phenotypes, predominantly focused on dual-energy x-ray absorptiometry-derived bone mineral density (BMD) of the hip and spine.

Evidence Synthesis: Sixty-six BMD loci have been replicated across all the published GWAS, confirming the highly polygenic nature of BMD variation. Only seven of the 66 previously reported genes (*LRP5*, *SOST*, *ESR1*, *TNFRSF11B*, *TNFRSF11A*, *TNFSF11*, *PTH*) from candidate gene association studies have been confirmed by GWAS. Among 59 novel BMD GWAS loci that have not been reported by previous candidate gene association studies, some have been shown to be involved in key biological pathways involving the skeleton, particularly Wnt signaling (*AXIN1*, *LRP5*, *CTNNB1*, *DKK1*, *FOXC2*, *HOXC6*, *LRP4*, *MEF2C*, *PTHLH*, *RSPO3*, *SFRP4*, *TGFBR3*, *WLS*, *WNT3*, *WNT4*, *WNT5B*, *WNT16*), bone development: ossification (*CLCN7*, *CSF1*, *MEF2C*, *MEPE*, *PKDCC*, *PTHLH*, *RUNX2*, *SOX6*, *SOX9*, *SPP1*, *SP7*), mesenchymal-stem-cell differentiation (*FAM3C*, *MEF2C*, *RUNX2*, *SOX4*, *SOX9*, *SP7*), osteoclast differentiation (*JAG1*, *RUNX2*), and TGF-signaling (*FOXL1*, *SPTBN1*, *TGFBR3*). There are still 30 BMD GWAS loci without prior molecular or biological evidence of their involvement in skeletal phenotypes. Other skeletal phenotypes that either have been or are being studied include hip geometry, bone ultrasound, quantitative computed tomography, high-resolution peripheral quantitative computed tomography, biochemical markers, and fractures such as vertebral, nonvertebral, hip, and forearm.

Conclusions: Although several challenges lie ahead as GWAS moves into the next generation, there are prospects of new discoveries in skeletal biology. This review integrates findings from previous GWAS and provides a roadmap for future directions building on current GWAS successes. **(*J Clin Endocrinol Metab* 97: E1958–E1977, 2012)**

ISSN Print 0021-972X ISSN Online 1945-7197
Printed in U.S.A.

doi: 10.1210/jc.2012-1890 Received April 5, 2012. Accepted July 9, 2012.
First Published Online September 10, 2012

Abbreviations: AUC, Area under the curve; BMD, bone mineral density; BMI, body mass index; BR, buckling ratio; CDCV, common disease/common variants; CI, confidence interval; CNV, copy number variant; DXA, dual-energy x-ray absorptiometry; EU, European descent; FN, femoral neck; GEFOS, The Genetic Factors for Osteoporosis Consortium; GWAS, genome-wide association study; GXG, gene-by-gene; HSA, hip structure analysis; LD, linkage disequilibrium; LS, lumbar spine; MAF, minor allele frequency; MHC, major histocompatibility complex; NSA, neck-shaft angle; NW, narrowest width; pQCT, peripheral quantitative computed tomography; SNP, single nucleotide polymorphism; WB, whole body.

In the last 5 yr, the study of genetics of complex diseases has seen a meteoric rise in scope. In contrast to Mendelian diseases caused by a mutation in a single gene, multiple genes and environmental factors contribute to these complex phenotypes. Identifying the relevant genes associated with complex diseases has been difficult, in part because each causal gene only makes a small contribution to overall heritability (1). The study of genetic variants associated with complex phenotypes using a genome-wide association study (GWAS) approach has certain advantages over other approaches, because GWAS: 1) has greater statistical power than using linkage studies to identify common, low-penetrance, disease-susceptible variants (2); 2) has higher resolution mapping with millions of single nucleotide polymorphisms (SNPs) genotyped across the genome to narrow down the disease-susceptible locus into a single gene or even a single sequence variant instead of a linkage 'locus' that includes 10 to 100 genes; 3) uses an "agnostic" approach (3) not requiring prior knowledge of the molecular involvement of candidate genes in the pathophysiology of diseases, which opens up opportunities for novel gene discovery; and 4) offers the prospect of shortening the time and effort required to discover new genetic determinants for complex diseases. There are several milestones that made GWAS efforts feasible: 1) the phenomenon of linkage disequilibrium (LD) (4, 5) was established in the human genome; 2) evidence for the concept of the "common disease/common variants" (CDCV) hypothesis (6); 3) completion of the International HapMap Project (7, 8) (for common sequence variants with minor allele frequency $> 1\%$) and more recently the 1000 Genomes Project (9) [for a deeper resolution of both common and rare sequence variants (10, 11) as well as structural variants (12)] created a catalog of genetic sequence variants across the human genome; 4) high-throughput technologies with accurate genotyping calling algorithms (13, 14) for genotyping hundreds of thousands to millions of SNPs in parallel were developed at a cost that made it possible to genotype large numbers of individuals who have been phenotypically well-characterized; and 5) statistical analysis methods have been developed to impute millions of SNP not actually genotyped based on LD structure of reference populations (15, 16) and to test variant-phenotype associations efficiently (17). GWAS have proliferated so dramatically that a web site hosted by the National Human Genome Research Institute has been created to catalog the findings (http://genome.gov/GWAStudies) (18).

GWAS of Skeletal Phenotypes

Predominance of the bone mineral density (BMD) phenotypes

Since the first GWAS on skeletal phenotypes was published in 2007 (19), there have been more than 40 published GWAS on skeletal phenotypes. The study of skeletal phenotypes over the past 5 yr has been predominantly focused on dual-energy x-ray absorptiometry (DXA)-derived BMD of the hip and spine. DXA-derived BMD is one of the single most important and the strongest predictors of subsequent osteoporotic fracture in both men and women. The ultimate phenotype of the aging skeleton is the low-trauma osteoporotic fracture; however, there are considerable challenges to studying this phenotype given the etiological heterogeneity, the quality of fracture assessment, and the difficulty in identifying gene-environment interactions. On the other hand, DXA-derived BMD measurements are relatively straightforward. The scanning equipment is widely available, and there are only two primary manufacturers of DXA equipment, which is a key consideration when replicating the genetic association findings in independent samples. In addition, for complex traits with polygenic effects, a GWAS approach usually requires considerably large sample sizes (due to the stringent genome-wide significant α level and moderate effect size) to be able to identify SNP-trait associations. BMD has been the most logical choice for GWAS because DXA-derived BMD has been measured in a large number of existing epidemiological studies with DNA samples.

In this review, we primarily focused on GWAS of BMD at different skeletal sites, such as the lumbar spine (LS) (19–30), femoral neck (FN) (19, 23–31), total hip (20–22, 28), whole body (WB) (32), wrist (33), radius (34), tibia (34), and cortical volumetric BMD of the tibia by peripheral quantitative computed tomography (pQCT) (35) as well as hip structure analysis (HSA) (19, 26, 36, 37), hip fracture (38), and Paget's disease of bone (39, 40). The characteristics of selected GWAS are described in Tables 1 and 2.

Study design

To avoid reporting false-positive findings that are commonly found in candidate gene association studies, for this review, we only included GWAS that performed replication analyses. The design of a typical GWAS involves multiple stages (Fig. 1). The first stage is usually a genome-wide discovery effort that relies on association analyses of hundreds of thousands of genotyped SNPs or millions of imputed SNPs. The second stage is a replication stage to replicate top associated SNP in an equal or larger independent sample or samples. This standard study design provides robust association results and avoids potential false-positive findings, although nonreplication may represent false-negative findings resulting from heterogeneity between discovery studies and replication studies.

Except for a handful of studies in East Asian populations (25, 29, 34, 38), most of the GWAS have involved

Caucasian samples of European descent (EU). Most of the EU studies genotyped SNPs using DNA chips with approximately 300K to 550K SNPs. A set of genotyping quality-control processes is typically employed to exclude low-quality SNPs, such as those with a low genotyping call rate, a significant deviation from Hardy-Weinberg equilibrium, SNPs with excessive autosomal heterozygosity, *etc*. Inadequate quality control procedures may introduce false-positive findings (41), as shown in a recent GWAS (42). Additional efforts are typically made to avoid confounding due to the presence of population substructure, which may cause false-positive results due to the difference of allele frequencies in subethnic groups of study samples (43). Possible population substructure is handled by either excluding ethnic outliers (44) or adjusting individuals' ancestral genetic background using methods, such as principal component analysis (PCA) of genotypes (45), in the association analyses. Some GWAS also apply genome control to adjust the λ_{GC} inflation factor for potential population stratification (46, 47).

With the concept of indirect association and the CDCV hypothesis, GWAS benefit greatly from LD (4, 5), which is the correlation between the genotyped SNPs on each array and other untyped nearby causal alleles. Compared with the HapMap Phase II reference panel [Utah residents with Northern and Western European Ancestry from the Centre d'Etude du Polymorphism Humain (CEPH) collection] (7), the coverage, which is the percentage of information of the common variation [minor allele frequency (MAF) > 5%] in the human genome that is captured by genotyped SNP on the DNA chips, is approximately 65 to 85% (at LD index, $r^2 \geq 0.8$) (3, 48). Thus, commercial chip designs potentially capture the majority of common genetic variation in EU populations. It was also recognized that African-derived populations have greater genetic diversity and lower levels of LD, requiring a greater density of SNP to provide genome-wide coverage of common variation.

To be able to directly compare association results of the same SNP across studies with different DNA chips, of large-scale GWAS (27, 28, 30, 34, 35, 40) also have imputed approximately 2 million additional untyped common SNPs (MAF ≥ 5%) based on the LD structure in EU populations. This gain in study efficiency and power through pooling association results of the same imputed SNPs is achieved without additional genotyping and thus permits more comprehensive association studies with current products at no extra cost.

The majority of GWAS of skeletal phenotypes have studied samples with mixed populations of premenopausal women, postmenopausal women, and men with a wide age range (18–99 yr), although the average age is usually 50 yr or greater. There is a general belief that it is likely to be sex-specific genetic variants conferring susceptibility to osteoporosis, and the genetic regulation of bone growth at younger ages and bone loss after peak bone

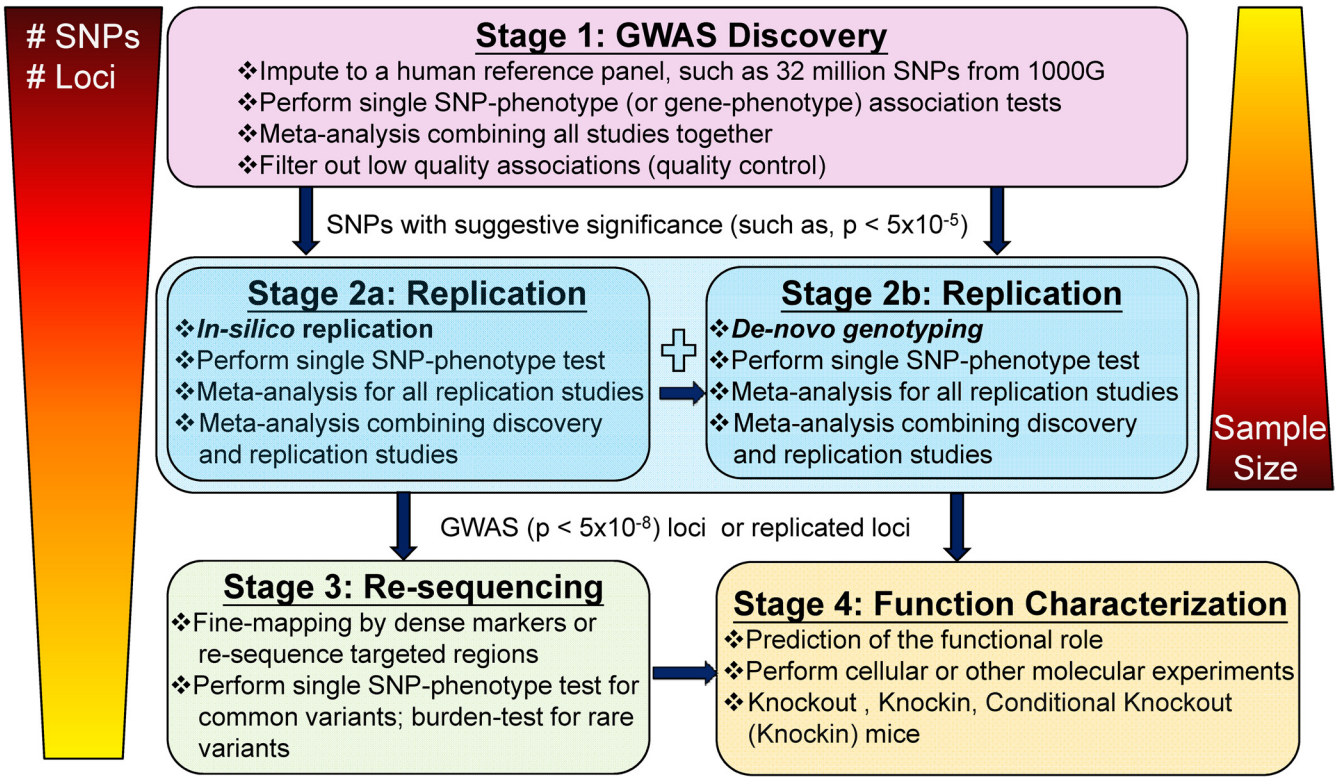

FIG. 1. GWAS design.

TABLE 1. Published GWAS of BMD phenotypes (up to April 1, 2012)

	GWAS discovery stage					
First author (Ref.)	Traits	Race	Age (yr), mean (range)	No. of samples (% female)	SNP	Covariates
BMD						
Kiel (19)	LS, FN, hip	EU	62.5 (29–86)	1,141 (57%)	100 K	Sex, age, age^2, BMI, height, smoking, physical activity, cohorts, estrogenic status (female)
Styrkarsdottir (20)	LS, hip	EU	63.6 (18–98)	5,861 (87%)	300 K	Sex, age, weight
Styrkarsdottir (21)	LS, hip	EU	64.0 (18–99)	6,865 (87%)	300 K	Sex, age, weight
Richards (23) (GEFOS)	LS, FN	EU	59.9 (18–?)	2,094 (100%)	314 K	Sex, age
Timpson (32)	WB	EU	9.9 (?)	1,518 (51%)	315 K	Sex
Xiong (22)	LS, hip	EU	50.3 (?)	1,000 (50%)	379 K	Sex, age, weight, height
Guo (31)	FN	EU	50.3 (?)	983 (50%)	342 K	Sex, age, weight
Tan (33)	Wrist	EU	50.3 (?)	1,000 (50%)	379 K	Sex, age
Koller (24)	LS, FN	EU	33.2 (25–45)	1,524 (100%)	548 K	Age, weight
Paternoster (35)	Cortical BMD at tibia	EU	17.1 (13–25)	1,934 (28%)#	2.4 M imputed	Sex, age, weight, height
Hsu (26) (GEFOS)	LS, FN	EU	60.8 (29–86)	3,569 (57%)	433 K	Sex, age, age^2, cohorts, estrogenic status (female)
Duncan (28) (GEFOS)	Hip, FN, LS	EU	? (55–85)	Hip or FN BMD Z $<$ -1.5: 900 (100%); Z $>$ 1.5: 1,055 (100%)	2.5 M imputed	Age, age^2, weight, centers
Cho (34)	Radius, tibia	Asian Korean	52.2 (40–69)	8,842 (53%)	2.1 M imputed	Sex, age, area
Kung (25)	LS, FN	Asian Chinese	50.0 (?)	LS or FN BMD, lowest 10%: 424 (100%); highest 10%: 376 (100%)	488 K	Age, weight
Kou (29)	Osteoporosis	Asian Japanese	60.2 (?)	LS or FN BMD, T $<$ -2.5: 157 (100%); T $>$ -2.5: 1,557 (52%)	224 K	N.A.
Rivadeneira (27) (GEFOS)	LS, FN	EU	60.9 (18–96)	19,195 (74%)#	2.5 M imputed	Sex, age, weight
Estrada (30) (GEFOS)	LS, FN	EU	59.6 (?–96)	32,961 (70%)#	2.5 M imputed	Sex, age, weight

(*Continued*)

N.A., Not available; GDPD, Genetic Determinants of Paget's Disease Consortium; BUA, Broadband ultrasound attenuation.

Samples: #, Total number of samples from several independent studies. A meta-analysis was applied to estimate effect size and *P* values of either discovery stage or replication stages. Gene/loci: *, The most significant signal was for a SNP located in the intergenic regions of two nearby genes [gene1–gene2*] or a nearby gene [gene1*]. Phenotype-specific results: We highlighted phenotype-specific results in each study if there was more than one phenotype tested in that particular study. For example, in Styrkarsdottir *et al.*, 2008, two skeletal sites were tested ("Traits" column). *ZBTB40-WNT4* locus was associated with both LS and total hip BMD. However, MHC region was only associated with LS BMD. We listed MHC region as "MHC(LS)" to highight that it was only associated with LS BMD. We listed MHC region as "MHC(LS)" to show it was associated with LS BMD.

TABLE 1. Continued

First author (Ref.)	Replication stage, samples (% female)[#]	Replicated GWAS findings: Previously reported candidate genes	Replicated GWAS findings: Novel GWAS genes/loci	Osteoporotic fractures [gene: OR (95% CI)]
BMD				
Kiel (19)	N.A.			
Strykarsdottir (20)	7,925 (79%)[#]	*ESR1-C6orf97, TNFRSF11B-COLEC10*, TNFSF11-AKAP11**	*ZBTB40-WNT4*, MHC* (LS)	*ZBTB40-WNT4*:* 1.15 (1.07–1.25) *MHC:* 1.09 (1.02–1.16) LRP4: 1.11 (1.03–1.19) *SPTBN1:* 1.11 (1.05–1.17)
Styrkarsdottir (21)	8,510[#] (83%)[#]	*TNFRSF11A* (hip)*, *SOST* (hip)*	*MHC* (LS), *SP7-AAAS* (LS)*, *MARK3* (hip)	*SOST*:* 1.10 (1.04–1.17)
Richards (23) (GEFOS)	6,463[#] (88%)[#]	*LRP5* (LS), *TNFRSF11B-COLEC10**		*LRP5:* 1.3(1.09–1.52)
Timpson (32)	4,178 (49%)		*SP7-AAAS**	
Xiong (22)	4,925 (63%)[#], 2,955 (51%)[#], Chinese 908 (0%) Tobago		*TGFBR3* (spine), *ADAMTS18* (hip)	
Guo (31)	2,557 (55%)	*PTH-FAR1**	*IL21R*	
Tan (33)	1,626 (51%) Chinese		*SOX6**	
Koller (24)	669 (100%) African Am		*CATSPERB* (FN)	
Patermoster (35)	2,803 (52%) 15.5 yr 1,052 (0%) 78.7 yr	*TNFSF11**		
Hsu (26) (GEFOS)	7,721 (72%)[#]	*TNFRSF11B-COLEC10* (LS)*	*WLS**(LS), *SOX6* (LS)	
Duncan (28) (GEFOS)	FN, LS BMD 20,898 (100%)[#]	*TNFRSF11B-COLEC10*, TNFSF11-AKAP11* (LS)*	*WNT4 - ZBTB40* (FN)*, *MEF2C* (FN)*, *SOX6* (FN)*, *FLJ42280* (FN), *GALNT3* (FN), *RSPO3-RPS4XP9* (FN)*	*GALNT3:* vertebral fractures, 0.89 (0.80–0.99); osteoporotic fractures, 0.92 (0.85–0.99)
Cho (34)	Heel BMD 7,861 Korean		*FAM3C* (radius and heel)	
Kung (25)	LS or FN BMD Chinese 456 (100%) lowest 10% 264(100%) highest 10% BMD 3,465 (100%)[#], Chinese 13,913 (81%)[#], EU		*JAG1* (LS)	*JAG1*, 0.7 (0.57–0.93)
Kou (29)	LS or FN BMD T < -2.5 Japanese Cases: 2,092 (100%)[#], Controls: 3,114 (85%)[#]		*FONG*	*FONG*, 1.25 (1.16–1.35)
Rivdeneira (27) (GEFOS)	N.A.	5 loci (bold genes below)	15 loci (bold genes below)	
Estrada (30) (GEFOS)	BMD, 50,933 (77%)[#]; fractures: cases, 31,016 (?); controls, 102,444 (?)	*C6orf97-**ESR1**, **LRP5,** SOST*, **TNFSF11**-AKAP11, **TNFRSF11A,** TNFRSF11B*	*ABCF2, ANAPC1* (FN), ***ARHGAP1-LRP4,*** *AXIN1, C12orf23, C17orf53-**HDAC5**, C18orf19* (FN), *C7orf58* (LS), *CDKAL1-SOX4** (LS), *CPN1, **CRHR1-**MAPT-WNT3, **CTNNB1***, **DCDC5***, *DHH-RHEBL1*, DNM3* (FN), *ERC1-WNT5B*, FAM9B-FAM9A* (LS), *LEKR1, **FOXL1***, *FUBP3, GALNT3, GPATCH1-WDR88-LRP3, HOXC4-HOXC6, IDUA, INSIG2** (LS), *JAG1, KCNMA1** (LS), *KLHDC5-PTHLH** (FN), *LIN7C** (LS), *MARK3, MBL2-DKK1*, **MEF2C*** (FN), **MEPE,** MPP7* (LS), *NTAN1, PKDCC* (FN), *PTX4-CLCN7*, RPS6KA5, RSPO3*, SALL1-CYLD** (FN), *SIDT1-KIAA2018* (FN), ***SLC25A13*,*** *SMG6,* ***SOX6***, ***SOX9**** (FN), ***SP7, SPTBN1***, *SUPT3H-RUNX2, **TXNDC3-STARD3NL***, **WLS**, WNT4*, WNT16-FAM3C, XKR9-LACTB2** (FN), ***ZBTB40****	$P < 5 \times 10^{-8}$ *C18orf19 (FAM210A), DKK1, LRP5, MEPE, SLC25A13, SPTBN1;* $P < 5 \times 10^{-4}$ *C17orf53, CTNNB1, DCDC5, FUBP3, RPS6KA5, SOST, STARD3NL, WNT4, WNT16, ZBTB40*

TABLE 2. Published GWAS of other skeletal phenotypes (up to April 2012)

First author, year (Ref.)	GWAS discovery stage					
	Traits	Race	Age (yr), mean (range)	No. of samples (% female)	SNP	Covariate
Others						
Guo (38)	Hip fracture	East Asian Chinese	69.5 (55–80)	Cases: 350 (65%) Controls: 350 (51%)	281 K	Sex, age, weight, height
Albagha (39)	Paget's disease	EU	?	No SQSTM1 mutations Case: 692 (?) Cotrol: 1001 (?)	294 K	N.A.
Albagha (40), GDPD	Paget's disease	EU	?	No SQSTM1 mutations; cases, 692 (?); controls, 2,699 (?)	2.5 M imputed	Sex
Kiel (19)	HSA, BUA	EU	62.5 (29–86)	1,141 (57%)	100 K	See the first record
Liu (36)	Proximal femur size	EU	50.3 (?)	1,000 (50%)	379 K	Age, age^2, weight, height
Zhao (37)	HSA	EU	50.3 (?)	987 (50%)	379 K	Sex, age, weight, height, age^2*sex
Hsu (26) (GEFOS)	HSA	EU	60.8 (29–86)	3,421 (57%)	433 K	Sex, age, age^2, BMI, height, cohorts, estrogenic status

(*Continued*)

Samples: #, Total number of samples from several independent studies. A meta-analysis was applied to estimate effect size and *P* values of either discovery stage or replication stages. Gene/loci: *, The most significant signal was for a SNP located in the intergenic regions of two nearby genes [gene1–gene2*] or a nearby gene [gene1*]. Phenotype-specific results: We highlighted phenotype-specific results in each study if there was more than one phenotype tested in that particular study. For example, in Styrkarsdottir *et al.*, 2008, two skeletal sites were tested ("Traits" column). *ZBTB40-WNT4* locus was associated with both LS and total hip BMD. However, MHC region was only associated with LS BMD. We listed MHC region as "MHC(LS)" to highight that it was only associated with LS BMD. We listed MHC region as "MHC(LS)" to show it was associated with LS BMD.

mass may be different. Nevertheless, most GWAS have been done with all available participants to maximize sample size, and thus statistical power. A few exceptions have included GWAS focused on children (32), teenagers (35), premenopausal women (24), postmenopausal women (25, 28), and adult women (23) only. However, the relatively smaller sample sizes of these subsets of individuals have limited the statistical power in the discovery stage.

New biological insights of BMD

To date, 66 BMD GWAS loci were replicated. The "replicated" GWAS loci are defined as: 1) SNPs with genome-wide significant associations at the discovery stage and with study-wide statistical significance at the replication stage; 2) SNPs without genome-wide significant associations at the discovery stage, but with genome-wide significant associations when meta-analyzing discovery and replication stages together; or 3) SNPs near genome-wide significant associations at the discovery stage, but achieving study-wide statistical significance at the replication stage after applying multiple testing adjustment. The genome-wide significant cutoff *P* value (α-level) varied among studies ($\alpha = 6 \times 10^{-7}$ to 5×10^{-8}), depending on the number of SNPs genotyped and imputed, and also depending on the methods that were applied for correcting multiple testing. During the replication stage, statistical significance cutoff *P* values were corrected for the number of SNP that were included in the replication analysis. The annotation of the top associated SNPs as being in a gene or a locus is somewhat problematic because, in many instances, the top associated SNPs might be assigned to multiple genes or may lie in intergenic regions. In Tables 1 and 2, we reannotated the SNPs based on human genome reference GRch37.3 and assigned SNP to a locus as follows: 1) top SNPs located within a known gene; 2) top SNPs located in multiple genes (such as *ESR1*-C6orf97); 3) top SNP located in the intergenic regions with two nearby genes existing less than 100 kb [such as *TNFRSF11B-COLEC10** (where the *asterisk* indicates that a SNP is outside the gene)]; 4) top SNPs located in the intergenic regions with only one gene within 100 kb away (such as *SOX6**); or 5) authors specified gene annotation based on a specific reason. For example, Albagha *et al.* (39) assigned the top SNP associated with Paget's disease as being in the *CSF1* locus, despite the fact that it was located far away from the *CSF1* gene and closer to a nearby gene, *EPS8L3*. The annotation to the *CSF1* gene was based on the observation that there was a recombination "hotspot" between the top SNP and the *EPS8l3* gene.

Among the 66 BMD GWAS loci, only seven previously reported genes (*LRP5*, *SOST*, *ESR1*, *TNFRSF11B*,

TABLE 2. Continued

First author (Ref.)	Replication stage, samples (% female)[#]	Replicated GWAS significant findings: Previously reported candidate genes	Replicated GWAS significant findings: Novel GWAS genes/loci	Osteoporotic fractures [gene: OR (95% CI)]
Others				
Guo (38)	Chinese cases: 390 (71%); controls: 516 (64%); hip BMD, 2,955 (51%)[#]; Chinese, 7,007 (57%)[#] EU		*ALDH7A1*	*ALDH7A1:* 2.25 (1.72–2.94)
Albagha (39)	Cases, 481 (?); controls, 520 (?)	*TNFRSF11A**	*CSF1*, OPTN*	Paget's disease, 1.46 (1.30–1.63) ~ 1.82 (1.61–2.04)
Albagha (40), GDPD	Cases, 1,474 (45%)[#]; controls, 1,671 (47%)[#]	TNFRSF11A*	CSF1*, OPTN, PML, RIN3, NUP205, TM7SF4	Paget's disease, 1.34 (1.25–1.45) ~ 1.72 (1.57–1.87)
Kiel (19)	N.A.			
Liu (36)	1,216 (100%)			
Zhao (37)	1,488 (56%), 2,118 (47%) Chinese		RTP3 (BR)	
Hsu (26) (GEFOS)	5,077 (67%)[#]		RAP1A (NW), TBC1D8 (NSA), OSBPL1A (NW)	

TNFRSF11A, *TNFSF11*, and *PTH*) from candidate gene association studies were confirmed by GWAS. It is somewhat surprising that only a handful of candidate genes out of almost 100 previously reported associated candidate genes (from Ref. 49–53 review papers and PubMed searches) were confirmed by GWAS, and candidate genes, such as *VDR*, *MTHFR*, *IGF1*, *IL6*, *CYP19* genes, *etc.*, were not confirmed by GWAS. This phenomenon is not unusual in GWAS of other complex traits or phenotypes such as type 2 diabetes and blood pressure. Several potential explanations include: 1) false-negative findings of GWAS due to the very stringent genome-wide significant cutoff; 2) inadequate statistical power to identify those previously reported associations with modest genetic effects; 3) genetic heterogeneity in subpopulations within the GWAS meta-analysis (*e.g.* different geographic locations and different principal characteristics), making it more difficult to detect genetic variant-phenotype associations; 4) genetic effects with strong gene-gene and gene-environmental interactions; and certainly, 5) false-positive findings of previous candidate gene association studies with small sample size and without appropriate replication.

Among 59 novel BMD GWAS loci that were not reported by previous candidate gene association studies, several have been shown to be involved in key biological pathways involving the skeleton, particularly Wnt signaling (*AXIN1*, *LRP5*, *CTNNB1*, *DKK1*, *FOXC2*, *HOXC6*, *LRP4*, *MEF2C*, *PTHLH*, *RSPO3*, *SFRP4*, *TGFBR3*, *WLS*, *WNT3*, *WNT4*, *WNT5B*, *WNT16*); bone development: ossification (*CLCN7*, *CSF1*, *MEF2C*, *MEPE*, *PKDCC*, *PTHLH*, *RUNX2*, *SOX6*, *SOX9*, *SPP1*, *SP7*); mesenchymal-stem-cell differentiation (*FAM3C*, *MEF2C*, *RUNX2*, *SOX4*, *SOX9*, *SP7*); osteoclast differentiation (*JAG1*, *RUNX2*); and TGF-signaling (*FOXL1*, *SPTBN1*, *TGFBR3*). We also found knockout mice (Mouse Genome Informatics database) with skeletal phenotypes represented by genes in the top GWAS findings, including *GALNT3*, *AXIN1*, *CLCN7*, *CTNNB1*, *CYLD*, *DKK1*, *FOXL1*, *HOXC6*, *IBSP*, *IDUA*, *LRP4*, *MEPE*, *PKDCC*, *PTHLH*, *RUNX2*, *RSPO3*, *SOST*, *SOX6*, *SOX9*, *SP7*, *SPP1*, *TGFBR3*, and *WLS*. Human monogenic syndromes displaying skeletal phenotypes (OMIM database) were found for *CLCN7*, *GALNT3*, *JAG1*, *LRP4*, *PTHLH*, *SOST*, *SOX9*, *SP7*, and *WNT3*.

There are still 30 BMD GWAS loci without prior molecular or biological knowledge of their involvement in skeletal phenotypes. These novel loci include 1p36.2 (*ZBTB40**), 1q24.3 (*DNM3*), 2p16.2 (*SPTBN1**), 2q13 (*ANAPC1*), 2q33.1 (*FONG*), 3q13.2 (*SIDT1-KIAA2018*), 3q25.31 (*LEKR1*), 7p14.1 (*TXNDC3-STARD3NL**), 7q21.3 (*SLC25A13**), 7q36.1 (*ABCF2*), 8q13.3 (*XKR9-LACTB2*), 9q34.11 (*FUBP3*), 10p11.23 (*MPP7*), 10q22.3 (*KCNMA1**), 10q24.2 (*CPN1*), 11p14.1 (*LIN7C**), 11p14.1 (*DCDC5**), 12q13.2 (*DHH-RHEBL1**), 12q23.3

(C12orf23), 14q32.12 (*RPS6KA5*), 14q32.12 (*CATSPERB*), 14q32.32 (*MARK3*), 16p11 (*IL21R*), 16p13.11 (*NTAN1*), 16q23 (*ADAMTS18*), 17p13.3 (*SMG6*), 17q21.31 (C17orf53-*HDAC5*), 18p11.21 (C18orf19 as *FAM210A*), 19q13.11 (*GPATCH1-LRP3*), and Xp22.31 (*FAM9B*). We performed bioinformatics analyses to characterize the biological functions of these genes and to ascertain whether they may play a role in skeletal biology. Figure 2 shows the biological functional networks for these 66 BMD GWAS loci that were created based on pathways and/or functional groups (Fig. 2, genes grouped in *boxes* except for the *red box*) from KEGG, Biocarta, and IPA Ingenuity (canonical pathway and Gene Ontology only). The edges (*lines*) connecting genes within each box, connecting "boxes" to each other, or connecting a "circle" to a "box" represent the significant relationships and/or gene-gene interactions ($P < 0.05$ after Bonferroni correction) among nodes that were estimated by using GRAIL (54). GRAIL is a tool to examine relationships between genes in different disease-associated loci. Obviously, skeletal biology pathways (*boxes with orange border*) are highly interrelated. We also observed significant interactions between the functional groups that are unrelated to skeletal biology (*blue boxes*, *red box*, and *red circle*) and the functional groups that are related to skeletal biology such as the interactions between "ossification" and "behavior and nervous systems" and interactions between the "Wnt signaling" pathway and "behavior and nervous systems." We also found that several transcription regulators (such as NOTCH1, FUBP3, HOXC4, and SUPT3H) at the nucleus or plasma membranes were significantly interacting with molecules in the Wnt signaling pathway, suggesting their potential involvement in skeletal phenotypes via the Wnt signaling pathway. MAP/microtubule affinity-regulating kinase 3 (protein product of *MAPK3* gene) may also be involved in skeletal phenotypes via the Wnt signaling pathway. BMD GWAS loci were also enriched in "protein kinase signaling," "cancer-genetic disorders," and "reproductive-genetic disorders." However, for these gene-sets, no obvious relations and/or interactions with skeletal biological pathways were found. The genes in the red boxes did not have significant interactions with any other genes among the 66 BMD GWAS loci based on available databases of molecular interactions and biological pathways. Thus, GWAS have provided significant numbers of novel hypotheses and have elucidated the functional implications for skeletal metabolism that will bring new insights into skeletal biology.

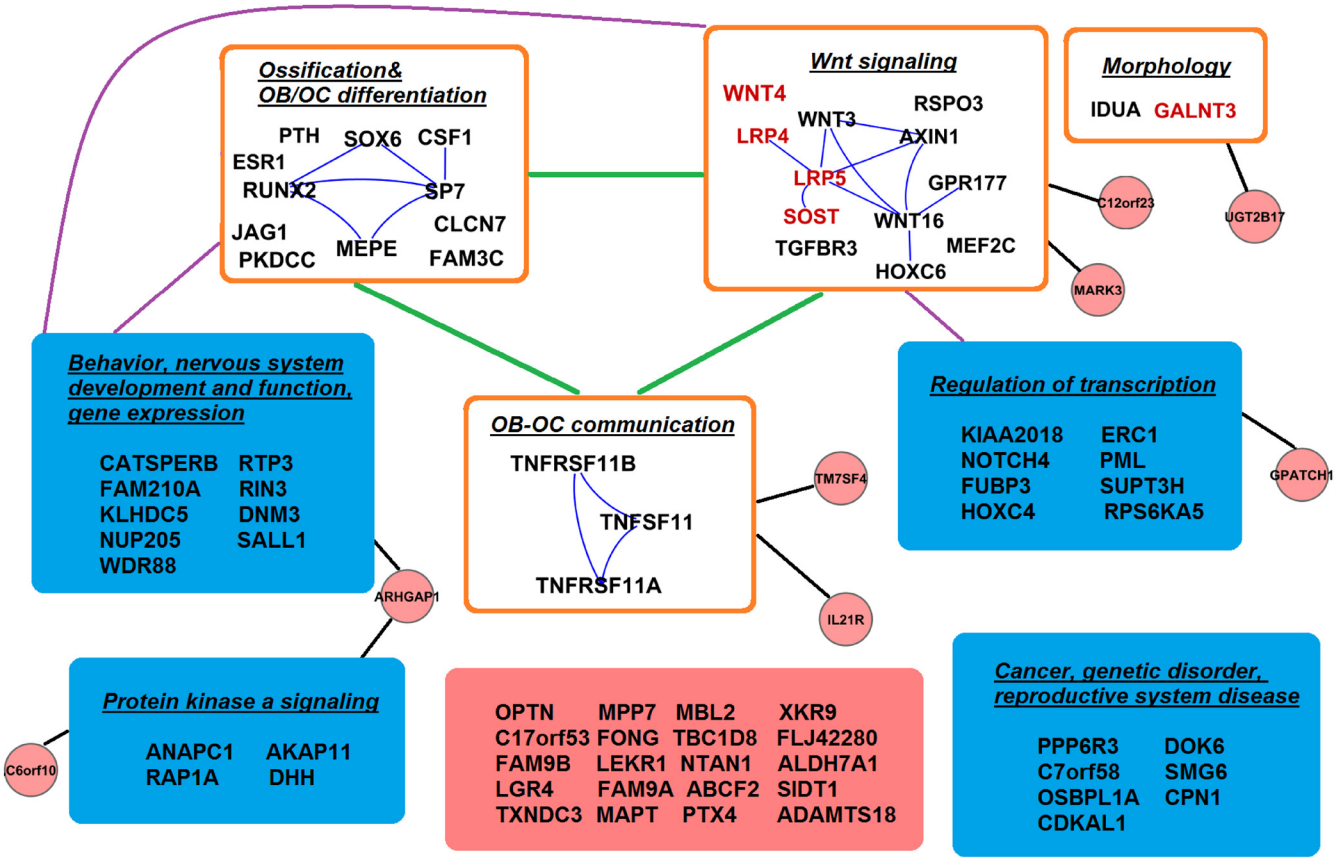

FIG. 2. Biological pathways and functional interaction network analyses for BMD GWAS loci. Molecular pathways and functional gene groups were obtained from KEGG, Biocarta, and the Ingenuity knowledge database (canonical pathways and gene ontology). We performed a gene-set enrichment analysis on 66 BMD GWAS loci. *Boxes with orange borders* are skeletal pathways or skeletal gene groups. *Blue boxes* are other pathways or gene groups. Genes in *red boxes* and *red circles* may be involved in known biological pathways but were not enriched ($P > 0.05$) in any pathways/gene groups among 66 BMD GWAS loci. The edges (*lines*) connecting genes, boxes, or circles are represented as functional interactions among genes with *P* values <0.05 (after multiple testing corrections) from GRAIL analyses.

Age, sex, skeletal sites, or ethnicity-specific BMD GWAS loci

Three GWAS were performed in younger individuals including children (WB BMD) (32), teenagers (cortical volumetric BMD at the tibia) (35), and premenopausal women (LS and FN BMD) (24). SNP rs1021188 located near *TNFSF11* (*RANKL*) was associated with cortical volumetric BMD in teenagers. The association signal of this SNP was independent from SNP in and near this gene that have been reported by other GWAS of LS and FN BMD in adult samples. The *CATSPERB* (24) locus was associated with FN BMD in premenopausal women only. The *CATSPERB* locus is about 500 kb away from the *RPS6KA5* (30) locus, which was associated with both FN and LS BMD in adult men and women. The *FONG* (29) locus was associated with BMD in an East Asian Japanese population only. On the other hand, the *JAG1* (25) and *FAM3C* (34) loci were associated with BMD in both East Asian populations and EU populations. Skeletal site-specific BMD GWAS loci have also been found (Tables 1 and 2). We have highlighted phenotype-specific results in each study if more than one phenotype was tested in that particular study. For example, in Styrkarsdottir *et al.* (20), two skeletal sites, LS and total hip BMD, were tested (Tables 1 and 2, traits column). The ZBTB40-WNT4 locus was associated with both LS and total hip BMD. However, the major histocompatibility complex (MHC) region was only associated with LS BMD. We listed MHC region as "MHC(LS)" to highlight that it was only associated with LS BMD. Most of the published GWAS on skeletal phenotypes did not have adequate power to test sex-specific genetic effects. Estrada *et al.* (30) performed sex-specific association analyses and tested the sex-specific effects using a conservative heterogeneity test. They only found one GWAS locus in Xp22.31 (near FAM9B) to be male-specific in its association with LS BMD. However, the imbalance in sample size between women and men and the conservative heterogeneity test limited the ability of this study to identify sex-specific findings. A formal genome-wide SNP-sex interaction meta-analysis study (55) was recently published (refer to *Gene-by-sex interactions* for details).

Skeletal phenotypes other than DXA BMD

In addition to DXA BMD, one study performed GWAS on cortical volumetric BMD at the tibia by pQCT (35). Devices such as pQCT measure cross-sections of bone (cortical or trabecular bone) and offer the opportunity to examine different skeletal compartments within a bone, which is impossible when using two-dimensional DXA. For example, the SNP rs1021188 in the *TNFSF11* (*RANKL*) gene was associated with cortical volumetric BMD, and the SNP was independent from a previously identified SNP that was associated with DXA BMD (rs9594738) (21) in the same *RANKL* region. This result indicates allelic heterogeneity at the *RANKL* locus. SNP rs1021188 was also associated with increased endosteal circumference, which indicates that *RANKL* may regulate endosteal expansion.

As shown in Table 1, several studies have performed GWAS on hip geometry measures (19, 26, 36, 37), including cortical thickness, buckling ratio (BR), cross-sectional area, femoral neck-shaft angle (NSA), the width of the femoral neck at the narrowest point (NW), femoral neck length, and proximal femur size in EU populations. Genome-wide significant associations with NW were found for SNP located on chromosomes 1p13.2 (*RAP1A*) and 18q11.2 (*OSBPL1A*) (26). Genome-wide significant associations with femoral NSA were found for SNPs located on chromosome 2q11.2 (*TBC1D8*) (26). A polymorphism within the *RTP3* gene was associated with BR in both EU and Chinese populations (37). Hsu *et al.* (26) found that: 1) the *RAP1A* gene was predicted to be causally linked with bone phenotypes in B6xC3H F2 intercross mice; 2) an eSNP (rs494453) located in intron 2 of the *RAP1A* gene was also found to be significantly associated with *RAP1A* gene expression in human primary osteoblasts; and 3) RAP1A expression differed across osteoblast maturation, suggesting that the *RA1PA* gene is a promising gene associated with femoral neck structure. RAP1A is a GTPase that mediates calcium signal transduction and has been found to mediate activities of c-Jun N-terminal kinase. Therefore, variants in the *RAP1A* gene may change the activities of c-Jun N-terminal kinase, which in turn affects osteoblast maturation.

Paget's disease of bone is a late-onset metabolic bone disease characterized by focal areas of increased bone remodeling primarily due to increased activity of osteoclasts. Paget's disease of bone affects approximately 8% of older men and approximately 5% women in EU populations. Mutations in the *SQSTM1* gene are the most common genetic causes of classic Paget's disease of bone, accounting for 10 to 50% of cases that run in families and 5 to 30% of cases in which there is no family history of the disease. To understand whether common variants also contribute to Paget's disease, Albagha *et al.* (39, 40) conducted a GWAS on Paget's disease patients without *SQSTM1* mutations. Common SNP located near *TNFRSF11A* (*RANK*) and the *CSF1* gene, and within the *OPTN*, *PML*, *RIN3*, *NUP205*, and *TM7SF4* genes were genome-wide significantly associated with Paget's disease. The risk (per-allele odds ratio) of Paget's disease ranged from 1.34 [95% confidence interval (CI), 1.25–1.45] to 1.72 (95% CI, 1.57–1.87), which is larger than the effect sizes associated with osteoporotic fracture that have been reported by GWAS in EU populations.

CSF1 and *TM7SF4* were found to be involved in osteoclast differentiation.

Effect size and explaining variation in skeletal phenotypes

In most of the GWAS of common phenotypes reported so far, the percentage of variance explained by all GWAS loci is modest. Of all the skeletal phenotypes, the GWAS of Paget's disease yielded the greatest percentage of risk explained [~13% of familial risk of Paget's disease (40)]. The strongest effect of a single SNP was reported by a GWAS in Chinese populations (22). A SNP in the *ADAMTS18* gene explained approximately 3.8% of total variation in hip BMD, and another SNP in the *TGFBR3* gene explained 1.2% of total variance in LS BMD (22). On the other hand, in GWAS in EU populations, most of the SNPs only account for 0.09 to 0.5% of the BMD variance. Rivadeneira *et al.* (27) estimated that all of the 20 BMD GWAS loci together only explained approximately 2.9% of the total genetic variance in LS BMD and approximately 1.9% of the total genetic variance in FN BMD in an EU population. As reported by the largest GWAS meta-analysis of BMD so far, combining all 56 BMD GWAS loci together only explained approximately 5.8% of the total genetic variance in FN BMD (30), which suggests that GWAS findings to date have not accounted for the majority of genetic variance in BMD despite the relatively large sample sizes with necessary power to determine modest genetic effects. This mystery of "missing heritability" is commonly found in many other complex traits and phenotypes (56). Other examples included a finding of 12.5% of the variance in heritability of body height (by180 GWAS loci) (57), 4% of the variance in heritability of body mass index (BMI) (32 GWAS loci) (58), and 16% of the variance in heritability of ulcerative colitis (47 GWAS loci) (59). The challenges of missing heritability are discussed in the section below: Missing low-frequency variants, rare variants, and structural variants.

Attempts have been made to estimate the contribution of genetic variants to the risk for osteoporosis and osteoporotic fracture using the BMD GWAS loci. A simulation study also suggested that genetic profiling could enhance the predictive accuracy of fracture prognosis estimated only by clinical risk factors (such as sex, FN BMD, history of prior fracture, falls during the past 12 months, and age) and help to identify high-risk individuals (60). The Genetic Factors for Osteoporosis consortium (GEFOS) performed a receiver operating characteristics analysis and estimated the area under the curve (AUC) for osteoporosis (T-score < -2.5) and osteoporotic fracture in a prospective study in 2836 postmenopausal Danish women aged 55–86 yr (30). As shown in Fig. 3, age and weight alone predicted osteoporosis with an AUC

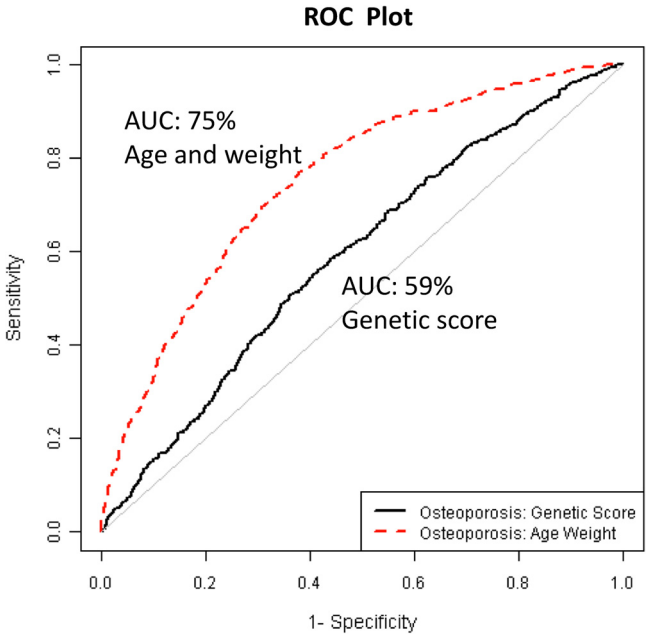

FIG. 3. The area under receiver operating characteristics (ROC) curves of genetic risk scores predicting the risk of osteoporosis (T-score ≤ -2.5) in 2836 postmenopausal EU women (Supplementary Fig. 8, Ref. 30).

of 75% (95% CI, 73–77%). The 56 BMD GWAS loci together showed significant, but modest, predictive ability with an AUC of 59% (95% CI, 56–62%). Adding GWAS loci together with age and weight did not substantially increase the predictive ability with an AUC of 76% (95% CI, 74–78%); however, AUC analyses evaluating the improvement in model fit with the addition of a new risk factor are insensitive to change. A similar pattern was observed for fracture prediction with a smaller AUC. Despite extremely low *P* values ($P < 5 \times 10^{-8}$) and many SNP that were genome-wide significantly associated with BMD, the identified genetic risk loci from GWAS do not seem to be able to predict individuals with low BMD or osteoporotic fracture after taking into account age and weight.

Fracture risks

To date, only one GWAS of fracture in a small case-control sample (discovery stage, 350 hip fracture cases and 350 age-matched controls) of a Chinese elder population has been published (38). SNP rs13182404 within the *ALDH7A1* gene on Chr5q31 was genome-wide significantly associated with hip fracture with a *P* value of 2.1×10^{-9} after combining discovery and replication stages with a total sample size of 740 hip fracture cases and 866 controls. The risk (per-allele odds ratio) of hip fracture was 2.25 (95% CI, 1.72–2.94), which is the largest effect size associated with osteoporotic fracture that has been reported by GWAS of skeletal phenotypes.

While conducting GWAS of BMD, several studies also examined SNP-fracture associations for their top findings (Tables

1 and 2). Among 68 BMD GWAS loci, 21 loci were also associated with fracture risk. They are located on chromosome 1p36 (*ZBTB40-WNT4*), 2q24-q31 (*GALNT3*), 2q33.1 (*FONG*), 3p21 (*CTNNB1*), 4q21.1 (*MEPE*), 6p21.3 (*MHC*), 7p14-p13 (*STARD3NL*), 7q21.3 (*SLC25A13*), 7q31 (*WNT16*), 9q34.11 (*FUBP3*), 10q11.2 (*DKK1*), 11p11.2 (*LRP4*), 11p14.1 (*DCDC5*), 11q13.4 (*LRP5*), 2p21 (*SPTBN1*), 14q31-q32.1 (*RPS6KA5*), 17q11.2 (*SOST*), 17q21.31 (C17orf53), 18p11.21 (C18orf19, recently named *FAM210A*), and 20p12.1-p11.23 (*JAG1*). These loci are statistically significant after Bonferroni correction for multiple testing. The majority of these findings were from our recent publication that attempted to validate GWAS signals for BMD in 31,016 osteoporotic fracture cases and 102,444 controls from more than 40 study cohorts who were part of the GEFOS/GENOMOS consortium framework (30). *JAG1* (25), *FONG* (29), and *ALDH7A1* (38) loci were found to be associated with fracture risk in East Asian populations only. The risk (per-risk allele odds ratio) of osteoporotic fracture ranged from 1.05 (95% CI, 1.02–1.08) for an SNP (risk allele frequency, 67%) located in C17orf53 locus (30) to 1.43 (95% CI, 1.08–1.75) for an SNP (risk allele frequency, 69%) located in the *JAG1* gene (25). To compare the magnitude of effect size across GWAS for each SNP, we flipped the allele and reported the odds ratio of the risk allele, instead of the minor allele in the populations. In this case, both variants in the C17orf53 locus and the *JAG1* gene were actually associated with a decreased risk of osteoporotic fracture because their risk alleles are the "major" alleles (allele frequency $>$ 50%) and the "minor" alleles were associated with an increased risk of osteoporotic fracture. For the remaining loci, most of the per-risk allele odds ratios are less than 1.1, especially in the EU populations, which indicates that the individual effects of these BMD GWAS loci on the risk of osteoporotic fracture is modest. This contrasts with the observation that lower DXA BMD explains 10 to 50% of osteoporotic fracture risk, depending on sex, age, and the population studied. In all cases, the SNP-fracture association showed the same effect direction as the SNP-BMD association (decreasing BMD and increased risk of fracture). One notable finding from these GWAS is that variants in the OPG-RANK-RANKL pathway (*TNFRSF11B*, *TNFRSF11A*, and *TNFSF11* genes) have not been found to be associated with osteoporotic fracture, despite the fact that variants in the OPG-RANK-RANKL pathway (osteoclastogenesis) are the most consistent findings associated with BMD across different studies and populations.

Notwithstanding some of the above inconsistencies relating GWAS findings from BMD and those from osteoporotic fracture, the BMD GWAS top findings that have been associated with osteoporotic fractures seem to predominantly involve bone morphology (*LRP4*, *GALNT3*, *DKK1*) and development, such as mineralization (*WNT4*, *MEPE*) and differentiation of osteoblasts (*CTNNB1*, *DKK1*, *LRP5*, *WNT4*, *SOST*) and osteoclasts (*JAG1*, *CTNNB1*). Several genes are in the Wnt/β-catenin signaling pathway, including *CTNNB1*, *DKK1*, *LRP5*, *SOST*, *WNT4*, and *WNT16*. Some are also involved in diabetic nephropathy or glucose metabolism disorders, including *CTNNB1*, *DKK1*, *LRP5*, *SLC25A13*, *WNT4*, and *WNT16* genes. For the remaining genes, further elucidation of their involvement in biological function and/or processes related to skeletal metabolism is needed.

Structural variation and skeletal phenotypes

In addition to SNP, copy number variants (CNVs) constitute a substantial fraction of genomic variability (61) and affect 20% of the variation in gene expression (62, 63). A CNV, usually larger than 1 kb of DNA length, is a structural variation of genomic sequence that results in the cell having an abnormal number of copies of a DNA segment. Several studies have estimated CNVs from the DNA chips that were used to genotype SNP and performed genome-wide CNV association analyses for hip fracture (64), hip geometry (65, 66), and BMD (65). A common CNV deletion (76.8% of subjects with homozygous deletions) covering the *UGT2B17* gene was associated with hip fracture in a case-control study comprising 350 Chinese hip fracture cases and 350 age- and sex-matched controls as the discovery stage and an independent Chinese sample with 399 hip fracture cases and 400 controls as replication (64). Compared to individuals with deletions of both copies of the *UGT2B17* gene, individuals carrying at least one copy of the *UGT2B17* were at increased risk for hip fracture odds ratio of 1.58 (95% CI, 1.12–2.22). However, this association was not replicated in a subsequent study with 1347 elderly Caucasian women (34.1% of subjects with homozygous deletions) (67), again highlighting the need for robust replication in studies with smaller sample sizes like this.

Outstanding Challenges and New Directions

Within just a 5-yr period, more than 60 novel loci have been identified by GWAS of skeletal phenotypes. The GWAS approach offers the prospect of shortening the time and effort required to discover new genes for osteoporosis compared with previous attempts to use traditional linkage or candidate gene association studies. The newly discovered GWAS genes enhance our current understanding of biological mechanisms underlying the skeletal pheno-

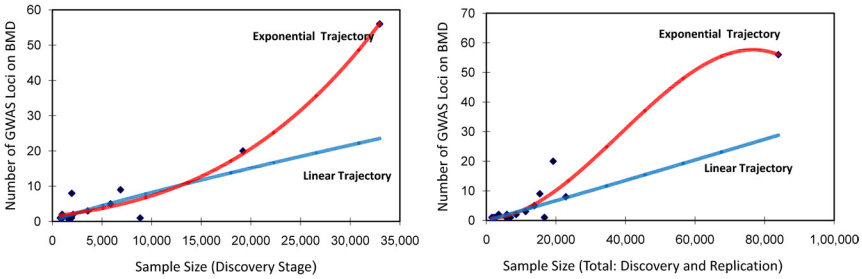

FIG. 4. Relation between sample size and number of GWAS loci identified. The *blue line* is the linear trajectory (linear regression excluded the study with the largest sample size); and the *red line* is the exponential trajectory (exponential regression included the study with largest sample size).

types and can provide "prioritized hypotheses" for biologists to further characterize gene functions and for scientists to develop new treatment, early diagnosis, and even prevention strategies for skeletal disorders. Despite the promise of novel discoveries for genes associated with skeletal phenotypes, the mystery of missing heritability (56) raises doubt about the value of GWAS and dampens the hope of some day using GWAS to develop personalized medicine for treatments. Explanations for this missing heritability have been suggested, including: 1) inadequate sample size to detect variants with smaller effect; 2) common SNPs as surrogate markers poorly linked to causal variants/functional variants; 3) allelic heterogeneity with multiple independent variants at the same locus; 4) rarer variants (minor allele frequency $< 1\%$) that are poorly detected by current genotyping arrays; 5) other structural variations; and 6) inadequate accounting for gene-gene and gene-environment interactions. In the following sections, we discuss challenges that are commonly encountered in GWAS and potential strategies to overcome those challenges. We also present some of the new directions being undertaken.

Sample size does matter

In one of the largest GWAS efforts to date for the height phenotype, 180 loci were discovered. With a sample size of 500,000 with effect sizes equal to or greater than those identified so far, studies estimated that about 42% of the phenotypic variance for height can be explained by all the autosomal common SNP on currently commercial DNA chips (68), and approximately 680 GWAS loci are associated with height (57). According to these estimations, current GWAS remain underpowered to detect statistical signals of all associated common variants. To identify these variants with modest effects, as mentioned previously, GWAS requires large samples of well-phenotyped individuals with available genotyping. Because most common variants have smaller effect sizes, even the largest GWAS of BMD with a sample size close to 33,000 subjects in the discovery stage (30), 56 replicated GWAS loci only explained approximately 5.8% of the total genetic variance in FN BMD. Greater power will require even larger sample sizes. Figure 4 graphically displays this relationship between sample size and the number of replicated GWAS loci. Thus, as the sample sizes of the discovery stage have grown, the number of replicated GWAS loci has grown exponentially (Fig. 4, *left panel*). Studies with smaller sample sizes had reported only a handful of replicated GWAS loci. It is encouraging that less than a doubling of sample size from Rivadeneira *et al.* (27) to Estrada *et al.* (30) more than doubled the number of replicated GWAS loci that were identified (Tables 1 and 2). Nevertheless, it is likely that other sequence variants not achieving genome-wide significance levels may also be true-positive findings. Follow-up of these loci is likely to yield new insights into bone metabolism. However, increasing sample size will not lead to the discovery of all the missing heritability by the current CDCV approach. As an example of height, all of the autosomal common SNP only explain up to half of the heritability of height (68).

To have sufficient power to determine genome-wide significant associations for BMD phenotypes, ongoing efforts are attempting to bring all the major cohort studies with genotyping and phenotyping together for the discovery phases of GWAS meta-analyses of skeletal phenotypes, as well as assembling a group of cohorts with phenotyping data and DNA that could be used for *de novo* genotyping of the top findings of the GWAS meta-analyses. Thus, GEFOS (www.gefos.org) and the Genetic Markers for Osteoporosis (GENOMOS) (www.genomos.eu) consortia were established to maximize the samples available for large GWAS meta-analyses with replication. Cohorts from around the world can participate in this joint collaboration to find novel genes contributing to the development of osteoporosis.

Bayesian approach

Compared with the usual test statistics (frequentist), the Bayesian approach may have advantages in GWAS. Because interpretation of the usual association *P* values estimated from frequentist approaches crucially depends on sample size and MAF, using *P* values requires a significant threshold that should be more stringent with increased sample size, which is contrary to common practice. The Bayesian approach, on the other hand, provides an alternative to the *P* value for assessing the consistency of a set of data with a null hypothesis (69). In addition, a Bayesian approach can easily incorporate prior knowl-

edge such as functional and biological information into the models because SNP with functional implications are more likely to be associated with disease status. A Bayes factor, defined as the ratio of the probability of the data under the null and alternative hypotheses, is commonly used as an indicator of the observed dataset that is more likely under alternative hypothesis. A handful of GWAS have been performed using Bayesian approaches (16, 70, 71). However, due to its computational intensity and difficulty in specifying prior distributions for all of the unknown parameters in the model, the Bayesian approach is not widely used (72). Several Bayes factor approximations to *P* values have been proposed such that a Bayesian approach may be an alternative approach of existing frequentist methods in the future (72).

Heterogeneity across study cohorts

One potential limitation of growing sample sizes of GWAS meta-analyses is the possibility of heterogeneity across the large number of studies. Meta-analysis has become a routine part of GWAS, and yet meta-analysis only provides optimum power to find effects that are homogeneous across cohorts. The heterogeneity of information (ancestral genetic background, covariates, genotypes, and phenotypes) not only affects statistical power, but also increases the potential for false-positive findings (73). This makes it imperative to harmonize information across cohorts to the largest extent possible.

Statistical methods exist to retrospectively examine potential heterogeneity between studies or between different phenotypic measurement assays or devices (74, 75). A "Forest plot" can be used to visualize the heterogeneity of effects (76). When heterogeneity exists, appropriate statistical methods, such as random-effects or mixed-effects meta-analysis should be applied (73, 77). Fixed-effects meta-analysis is commonly used in the GWAS setting. However, fixed-effects meta-analysis assumes that the genetic effects are the same across the different studies. Although, fixed-effects meta-analysis provides narrower confidence intervals and significantly lower *P* values for the variants than random-effects meta-analysis, when heterogeneity is present, fixed-effects meta-analysis inflates the type-I error (73). On the other hand, the random-effects (77) or mixed-effects meta-analysis assumes that the mean effect (of each SNP) in each study is different, and the means are usually assumed to be chosen from a Gaussian distribution. The variance of that Gaussian distribution (the amount of between-study heterogeneity) is estimated by the model.

In addition, heterogeneity not only affects statistical power and produces potentially false-positive findings in GWAS discovery, but also reduces the prediction accuracy of diseases (such as fracture risk) when applying genetic risk scores to a population that is in a heterogeneous environment (gene-by-environmental interactions) and/or heterogeneous genetic background. When planning studies, researchers should work together to create standardized procedures that can apply to many cohorts to collect medical information and to measure covariates and phenotypes. Toward this end, the PhenX project has provided a "toolbox" of phenotypes that have been well-validated (https://www.phenx.org).

Missing low-frequent variants, rare variants, and structural variants

With the limitation of current genome-wide genotyping density, GWAS efforts have focused on common SNPs (MAF $\geq$ 5%). Most of the identified common SNPs are not likely to be the causal or functional SNPs. The missing heritability suggests that a small proportion of a large number of common causal variants and a larger proportion of a small number of rare causal variants will contribute to the health of a human individual. Causal variants that are not in LD with the genotyped markers (such as the majority of the common SNP in a DNA chip) are likely to be rare (<1%) and uncommon (<5%) variants (78, 79). There is still considerable debate over three hypotheses that potentially explain this missing heritability: 1) synthetic associations (associations of multiple variants with a phenotype) with only multiple rare causal variants represent a significant proportion of associations detected in GWAS (80, 81); 2) a combination of both common and rare causal variants that are not in high LD with genotyped SNP on the current generation SNP arrays represent a significant proportion of the associations detected in GWAS; or 3) the CDCV hypothesis is still valid, but sample sizes are still too small (82).

Rare variants are predicted to exhibit stronger effect sizes than common variants, consistent with the view that functional allelic variants are subject to purifying selection pressure (83). Recent studies have identified clusters of rare variants related to complex traits (84–86) such as serum lipids (87), type 1 diabetes, sporadic epilepsy syndromes (88), blood pressure (89), hearing loss, sporadic autism (90), and cancers. These studies suggest that complex traits may be due to both common polymorphisms and multiple rare deleterious alleles in protein coding genes. Although the original signals detected in GWAS are often common variants, the discovery of highly penetrant rare variants in the same region might be crucial.

Deep resequencing approaches provide an opportunity to test for associations with all sequence variants (common and rare variants as well as structural variants) that

are, to some extent, much more detailed than the current GWAS with genotyping arrays (78, 79). This resolution is something that GWAS are unable to achieve. Deep resequencing in the promising targeted regions of individuals with extreme phenotypes may be the most promising strategy for accessing rare variants and has the potential to discover rare variants and structural variants neither genotyped on dense SNP arrays nor reliably imputed. The testing of associations between rare variants, structural variants, and skeletal phenotypes is beginning to be explored. The Framingham Study, The United Kingdom Twins Study, the Canadian Multicenter Osteoporosis Study, the Cardiovascular Heart Study, the Women's Health Study, and the Rotterdam Study have all begun to perform next-generation whole-genome sequencing, whole-exome sequencing, and/or targeted sequencing of loci found in previous GWAS meta-analyses. The first results of these efforts should be forthcoming.

Increasing coverage toward "next-generation GWAS"

There are conflicting opinions regarding the use of publicly available random samples as reference panels [such as the International HapMap project and the 1000 Genomes Project (9–11)] to impute disease-related causal variants in specific study populations. This is especially true for those populations not represented in the projects because the disease-related causal variants may not be captured by the individuals in the publicly available reference panels (80, 91, 92). In addition, although long-range phasing and haplotype imputation may be able to capture rare variants well, current imputation methods do not adequately predict rare variants.

With advances in technology for massively parallel genotyping of SNP, the capacity of commercial arrays has evolved to deliver very high density SNP capabilities, thus enabling highly powered "next-generation" GWAS with up to 5 million SNP in a DNA chip from Affymetrix (93) and Illumina. In addition, the recent emergence of the human exome-array will also permit multiple cohorts to genotype both common and rare functional variants in exons that have been identified by previous sequencing studies. This array is much more affordable than next-generation whole-exome sequencing and is being performed in dozens of studies with skeletal phenotypes worldwide.

Gene-by-gene (GXG) interactions

A recent study estimated that GXG interactions may account for a significant portion of the heritability of complex phenotypes (94). To date, few GWAS have incorporated GXG interaction testing due to the fact that the focus has first been on single-locus testing. In addition, testing statistical interaction in relation to a linear model (such as linear regression model) may not directly imply biological or functional interactions in genetics. An exhaustive search of all pair-wise two-locus interactions from a GWAS has been proposed (95). Although it is computationally feasible, it is very time consuming and does not deal with three-way, four-way, or even higher-order interactions. Therefore, sophisticated statistical tools still need to be developed.

Gene-by-sex interactions

The study of the genetic basis of complex phenotypes involves the consideration of genetic and environmental factors and how these interact with each other (GXE). One of the key challenges for current GWAS GXE interactions is the limited statistical power due to small sample size and heterogeneity of the measurement methods and distribution across studies.

Recent studies suggest that sex-specific genetic architecture influences many human phenotypes (96), including reproductive, physiological, and complex disease traits. Some of the underlying mechanisms might be attributed to differential gene regulation in males and females. Strong sexual dimorphism has been observed for BMD. One explanation for this sex-specific predisposition to osteoporosis and fracture risk is the possibility that the differences between men and women are driven by genetic effects determining bone fragility (97). Our recent study performed a GWAS gene-by-sex interaction on LS and FN BMD in approximately 25,000 individuals from seven cohorts and replicated top findings in an additional approximately 24,000 subjects (55). Despite the large collaborative effort involved, no significant gene-by-sex interaction was found. To have adequate power (80%) to detect gene-by-sex interaction signals (explaining 0.08% of total variance), at least 50,000 subjects will be needed. These results suggest that the interaction effects are smaller than the main effects and require even larger sample sizes than the sample sizes required for the main effects.

Pleiotropic effect

Most complex diseases with a heterogeneous pathophysiology cannot be precisely defined by only one phenotype (98, 99). Often the related endophenotypes for diseases may be associated with the same sets of genes (100). Testing related endophenotypes simultaneously is a powerful approach for identifying susceptibility genes and for measuring the pleiotropic effects of genes directly influencing multiple traits. By studying multiple skeletal phenotypes, including BMD of the hip, spine, heel ultrasound, and hip geometric indices in the Framingham Osteoporosis Study, the number of shared GWAS top hits between these phenotypes was increased (101) when the genetic correlation between two skel-

etal phenotypes was higher. It was concluded that most of the similarity between the quantitative bone phenotypes may be attributed to pleiotropic effects of genes.

Another approach to searching for pleiotropy is to leverage large GWAS consortia results and to look up the overlapping GWAS top SNPs between two phenotypes. When this was done with BMD and other phenotypes, several BMD loci (LS or FN) were also found to be associated with other nonskeletal phenotypes, such as *CLDN14* gene (kidney stone) (102), *ITGA1* (fasting glucose) (103), *WNT4-ZBTB40* (ulcerative colitis) (59), *TNFSF11* (Crohn's disease) (104), *MEF2C* and *FUBP3* (height) (57), and *C6orf97-ESR1* (breast cancer) (105). These observations suggest that pleiotropic genetic effects may exist between BMD and other phenotypes that traditionally have not been considered to be related to bone.

To use GWAS information, we recently (106) developed a multivariate genome-wide association approach (empirical-weighted linear-combined test statistics) to model multiple phenotypes simultaneously across the whole genome. Empirical-weighted linear-combined test statistics is a method to directly combine test statistics (aggregated test statistics from GWAS meta-analyses) of correlated phenotypes using a weighted sum of univariate test statistics to maximize the effect size of the overall association tests. Several benefits of using this approach include the following: 1) only aggregated test statistics from GWAS analyses are needed and not individual level data; and 2) the approach is not affected by individuals missing one of the multiple phenotypes. This method has been applied to study potential pleiotropic effects for BMD and reproductive health (such as age at menarche and age at natural menopause) (107) and for BMD and lean body mass (108) in large GWAS meta-analyses.

Integrative genomics: a step beyond statistical signals

Although GWAS provides an unbiased hypothesis-free approach to screen the genetic determinants of traits across the whole genome, the simple statistical signals do not provide the much-needed functional implication to predict the underlying biological processes involved in disease pathophysiology. More than half of the SNPs with genome-wide significant associations are located in intergenic regions between two or multiple genes. Due to the limited understanding of the structure of the human genome, it becomes a challenge to annotate the intergenic SNPs to nearby genes or to characterize the biological consequence of a polymorphism in intergenic regions. Kou *et al.* (29) discovered a GWAS locus associated with osteoporotic fracture in a Japanese population. They applied a protein motif analysis to predict an unknown gene in a chromosome 2q33.1 locus where the genome-wide significant SNP were located. This unknown gene encoded a protein containing a signal peptide and a formiminotransferase domain in its N terminal (FTCD_N domain). They were able to characterize the structure of this novel gene and named this gene "FONG." The molecular involvement of FONG in skeletal biology will need to be further studied.

To overcome this challenge, we proposed a systems genomics approach using likelihood-based causality network modeling to construct regulatory networks (26). This was done by integrating gene expression profiling experiments (whole genome transcripts) from human tissues (expression QTL and expression SNPs), as well as animal and cellular whole genome experimental models (such as PTH-stimulated osteoclastogenesis and osteoblastogenesis of embryonic stem cells) into genome-wide association analysis to prioritize GWAS genes for future functional validations in cellular and animal studies. In addition to genotypes, one can incorporate gene expression profiling from cellular, animal models and human tissues as well as from epigenetics, transcriptomics, and proteomics into a network model. This approach may lead to improving the detection of association signals and to a better understanding of genetic association signals while providing a systems view of the biological processes underlying disease susceptibility.

Beyond BMD phenotypes

In the field of skeletal genetics, the initial success with studies of DXA-derived BMD of LS and FN led to the expansion of phenotypes studied. Additional GWAS efforts are under way to study skeletal phenotypes such as proximal hip geometry (109), bone ultrasound, forearm BMD, pQCT of the LS, high-resolution pQCT of the radius and tibia, biochemical markers such as serum osteocalcin and osteoprotegerin (110), vertebral fracture, hip fracture (111), and nonvertebral fracture (112).

Conclusion

To date, GWAS have identified more than 60 novel loci associated with skeletal phenotypes (predominantly for DXA BMD), and, in doing so, GWAS have provided valuable insights into the genetic architecture of the skeleton. With the "agnostic" approach of GWAS, we expect that more discoveries of the genes involved in the regulation of bone metabolism will emerge compared with traditional candidate gene studies. In addition, high resolution of SNP mapping provides greater power than using linkage studies to identify low-penetrance, disease-susceptible variants. Several of the GWAS findings have highlighted key

biological pathways that influence BMD variation, particularly Wnt signaling, bone development (ossification), mesenchymal-stem-cell differentiation, osteoclast differentiation, and TGF signaling. More than half of the BMD loci identified by GWAS have no known biological function related to skeletal health. These novel loci, each with a subtle effect, together explained approximately 5.8% of the total genetic variance in FN BMD. To gather study cohorts for sufficient sample size, GWAS has fostered the growth of international collaborations on a scale previously unheard of in biological science. With growing collaborations and progress in sequencing/genotyping technologies, it is likely that more diverse populations will be studied and more loci will be discovered. Follow-up of the most promising GWAS top findings will require the identification of functional variants by performing sequencing of targeted regions and ultimately sequencing the whole genome. The ultimate goals of GWAS are to discover new molecular and biological pathways involved in the regulation of bone metabolism that can be leveraged for drug development and to enhance current risk factor profiles such as FRAX (113) by incorporating genetic risk scores. Although several challenges lie ahead as GWAS moves into the next generation, there are prospects of new discoveries in skeletal biology.

Acknowledgments

We acknowledge Dr. Chia-Ho Cheng for creation of Fig. 2.

Address all correspondence and requests for reprints to: Douglas P. Kiel, Institute for Aging Research, Hebrew SeniorLife, 1200 Centre Street, Boston, Massachusetts 02131. E-mail: kiel@hsl.harvard.edu.

Contributions to this review by Y.-H.H. and D.P.K. were funded in part from grants from the National Institute of Arthritis Musculoskeletal and Skin Diseases and the National Institute on Aging (R01 AR/AG 41398, R01 AR061162, and R21 AR056405).

Disclosure Summary: Y.-H.H. has nothing to declare. D.P.K. has received grant support from Amgen, Merck, and Eli Lilly and serves as a consultant for Amgen, Merck, and Eli Lilly.

References

1. **Altshuler D, Daly MJ, Lander ES** 2008 Genetic mapping in human disease. Science 322:881–888
2. **Hirschhorn JN, Daly MJ** 2005 Genome-wide association studies for common diseases and complex traits. Nat Rev Genet 6:95–108
3. **Pe'er I, de Bakker PI, Maller J, Yelensky R, Altshuler D, Daly MJ** 2006 Evaluating and improving power in whole-genome association studies using fixed marker sets. Nat Genet 38:663–667
4. **Rannala B, Reeve JP** 2001 High-resolution multipoint linkage-disequilibrium mapping in the context of a human genome sequence. Am J Hum Genet 69:159–178
5. **Reich DE, Cargill M, Bolk S, Ireland J, Sabeti PC, Richter DJ, Lavery T, Kouyoumjian R, Farhadian SF, Ward R, Lander ES** 2001 Linkage disequilibrium in the human genome. Nature 411:199–204
6. **Reich DE, Lander ES** 2001 On the allelic spectrum of human disease. Trends Genet 17:502–510
7. **Frazer KA, Ballinger DG, Cox DR, Hinds DA, Stuve LL, Gibbs RA, Belmont JW, Boudreau A, Hardenbol P, Leal SM, Pasternak S, Wheeler DA, Willis TD, Yu F, Yang H, Zeng C, Gao Y, Hu H, Hu W, Li C, Lin W, Liu S, Pan H, Tang X, Wang J, Wang W, Yu J, Zhang B, Zhang Q, Zhao H, Zhao H, Zhou J, Gabriel SB, Barry R, Blumenstiel B, Camargo A, Defelice M, Faggart M, Goyette M, Gupta S, Moore J, Nguyen H, Onofrio RC, *et al.*** 2007 A second generation human haplotype map of over 3.1 million SNPs. Nature 449:851–861
8. **Altshuler DM, Gibbs RA, Peltonen L, Dermitzakis E, Schaffner SF, Yu F, Peltonen L, Dermitzakis E, Bonnen PE, Altshuler DM, Gibbs RA, de Bakker PI, Deloukas P, Gabriel SB, Gwilliam R, Hunt S, Inouye M, Jia X, Palotie A, Parkin M, Whittaker P, Yu F, Chang K, Hawes A, Lewis LR, Ren Y, Wheeler D, Gibbs RA, Muzny DM, Barnes C, Darvishi K, Hurles M, Korn JM, Kristiansson K, *et al.*** 2010 Integrating common and rare genetic variation in diverse human populations. Nature 467:52–58
9. **1000 Genomes Project Consortium** 2010 A map of human genome variation from population-scale sequencing. Nature 467:1061–1073
10. **Nielsen R** 2010 Genomics: in search of rare human variants. Nature 467:1050–1051
11. **MacArthur DG, Balasubramanian S, Frankish A, Huang N, Morris J, Walter K, Jostins L, Habegger L, Pickrell JK, Montgomery SB, Albers CA, Zhang ZD, Conrad DF, Lunter G, Zheng H, Ayub Q, DePristo MA, Banks E, Hu M, Handsaker RE, Rosenfeld JA, Fromer M, Jin M, Mu XJ, Khurana E, Ye K, Kay M, Saunders GI, Suner MM, Hunt T, Barnes IH, Amid C, Carvalho-Silva DR, Bignell AH, Snow C, Yngvadottir B, Bumpstead S, Cooper DN, Xue Y, Romero IG, Wang J, Li Y, Gibbs RA, McCarroll SA, Dermitzakis ET, Pritchard JK, Barrett JC, Harrow J, Hurles ME, Gerstein MB, Tyler-Smith C** 2012 A systematic survey of loss-of-function variants in human protein-coding genes. Science 335:823–828
12. **Sudmant PH, Kitzman JO, Antonacci F, Alkan C, Malig M, Tsalenko A, Sampas N, Bruhn L, Shendure J, Eichler EE** 2010 Diversity of human copy number variation and multicopy genes. Science 330:641–646
13. **Gunderson KL, Steemers FJ, Lee G, Mendoza LG, Chee MS** 2005 A genome-wide scalable SNP genotyping assay using microarray technology. Nat Genet 37:549–554
14. **Ragoussis J** 2009 Genotyping technologies for genetic research. Annu Rev Genomics Hum Genet 10:117–133
15. **Li Y, Willer CJ, Ding J, Scheet P, Abecasis GR** 2010 MaCH: using sequence and genotype data to estimate haplotypes and unobserved genotypes. Genet Epidemiol 34:816–834
16. **Marchini J, Howie B, Myers S, McVean G, Donnelly P** 2007 A new multipoint method for genome-wide association studies by imputation of genotypes. Nat Genet 39:906–913
17. **de Bakker PI, Yelensky R, Pe'er I, Gabriel SB, Daly MJ, Altshuler D** 2005 Efficiency and power in genetic association studies. Nat Genet 37:1217–1223
18. **Hindorff LA, Sethupathy P, Junkins HA, Ramos EM, Mehta JP, Collins FS, Manolio TA** 2009 Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. Proc Natl Acad Sci USA 106:9362–9367
19. **Kiel DP, Demissie S, Dupuis J, Lunetta KL, Murabito JM, Karasik D** 2007 Genome-wide association with bone mass and geometry in the Framingham Heart Study. BMC Med Genet 8(Suppl 1):S14
20. **Styrkarsdottir U, Halldorsson BV, Gretarsdottir S, Gudbjartsson DF, Walters GB, Ingvarsson T, Jonsdottir T, Saemundsdottir J, Center JR, Nguyen TV, Bagger Y, Gulcher JR, Eisman JA, Christiansen C, Sigurdsson G, Kong A, Thorsteinsdottir U, Stefansson**

K 2008 Multiple genetic loci for bone mineral density and fractures. N Engl J Med 358:2355–2365

21. **Styrkarsdottir U, Halldorsson BV, Gretarsdottir S, Gudbjartsson DF, Walters GB, Ingvarsson T, Jonsdottir T, Saemundsdottir J, Snorradóttir S, Center JR, Nguyen TV, Alexandersen P, Gulcher JR, Eisman JA, Christiansen C, Sigurdsson G, Kong A, Thorsteinsdottir U, Stefansson K** 2009 New sequence variants associated with bone mineral density. Nat Genet 41:15–17
22. **Xiong DH, Liu XG, Guo YF, Tan LJ, Wang L, Sha BY, Tang ZH, Pan F, Yang TL, Chen XD, Lei SF, Yerges LM, Zhu XZ, Wheeler VW, Patrick AL, Bunker CH, Guo Y, Yan H, Pei YF, Zhang YP, Levy S, Papasian CJ, Xiao P, Lundberg YW, Recker RR, Liu YZ, Liu YJ, Zmuda JM, Deng HW** 2009 Genome-wide association and follow-up replication studies identified ADAMTS18 and TGFBR3 as bone mass candidate genes in different ethnic groups. Am J Hum Genet 84:388–398
23. **Richards JB, Rivadeneira F, Inouye M, Pastinen TM, Soranzo N, Wilson SG, Andrew T, Falchi M, Gwilliam R, Ahmadi KR, Valdes AM, Arp P, Whittaker P, Verlaan DJ, Jhamai M, Kumanduri V, Moorhouse M, van Meurs JB, Hofman A, Pols HA, Hart D, Zhai G, Kato BS, Mullin BH, Zhang F, Deloukas P, Uitterlinden AG, Spector TD** 2008 Bone mineral density, osteoporosis, and osteoporotic fractures: a genome-wide association study. Lancet 371: 1505–1512
24. **Koller DL, Ichikawa S, Lai D, Padgett LR, Doheny KF, Pugh E, Paschall J, Hui SL, Edenberg HJ, Xuei X, Peacock M, Econs MJ, Foroud T** 2010 Genome-wide association study of bone mineral density in premenopausal European-American women and replication in African-American women. J Clin Endocrinol Metab 95: 1802–1809
25. **Kung AW, Xiao SM, Cherny S, Li GH, Gao Y, Tso G, Lau KS, Luk KD, Liu JM, Cui B, Zhang MJ, Zhang ZL, He JW, Yue H, Xia WB, Luo LM, He SL, Kiel DP, Karasik D, Hsu YH, Cupples LA, Demissie S, Styrkarsdottir U, Halldorsson BV, Sigurdsson G, Thorsteinsdottir U, Stefansson K, Richards JB, Zhai G, Soranzo N, Valdes A, Spector TD, Sham PC** 2010 Association of JAG1 with bone mineral density and osteoporotic fractures: a genome-wide association study and follow-up replication studies. Am J Hum Genet 86:229–239
26. **Hsu YH, Zillikens MC, Wilson SG, Farber CR, Demissie S, Soranzo N, Bianchi EN, Grundberg E, Liang L, Richards JB, Estrada K, Zhou Y, van Nas A, Moffatt MF, Zhai G, Hofman A, van Meurs JB, Pols HA, Price RI, Nilsson O, Pastinen T, Cupples LA, Lusis AJ, Schadt EE, Ferrari S, Uitterlinden AG, Rivadeneira F, Spector TD, Karasik D, Kiel DP** 2010 An integration of genome-wide association study and gene expression profiling to prioritize the discovery of novel susceptibility loci for osteoporosis-related traits. PLoS Genet 6:e1000977
27. **Rivadeneira F, Styrkársdottir U, Estrada K, Halldórsson BV, Hsu YH, Richards JB, Zillikens MC, Kavvoura FK, Amin N, Aulchenko YS, Cupples LA, Deloukas P, Demissie S, Grundberg E, Hofman A, Kong A, Karasik D, van Meurs JB, Oostra B, Pastinen T, Pols HA, Sigurdsson G, Soranzo N, Thorleifsson G, Thorsteinsdottir U, Williams FM, Wilson SG, Zhou Y, Ralston SH, van Duijn CM, Spector T, Kiel DP, Stefansson K, Ioannidis JP, Uitterlinden AG** 2009 Twenty bone-mineral-density loci identified by large-scale meta-analysis of genome-wide association studies. Nat Genet 41:1199–1206
28. **Duncan EL, Danoy P, Kemp JP, Leo PJ, McCloskey E, Nicholson GC, Eastell R, Prince RL, Eisman JA, Jones G, Sambrook PN, Reid IR, Dennison EM, Wark J, Richards JB, Uitterlinden AG, Spector TD, Esapa C, Cox RD, Brown SD, Thakker RV, Addison KA, Bradbury LA, Center JR, Cooper C, Cremin C, Estrada K, Felsenberg D, Glüer CC, Hadler J, Henry MJ, Hofman A, Kotowicz MA, Makovey J, Nguyen SC, Nguyen TV, Pasco JA, Pryce K, Reid DM, Rivadeneira F, Roux C, Stefansson K, Styrkarsdottir U, Thorleifsson G, Tichawangana R, Evans DM, Brown MA** 2011 Genome-wide association study using extreme truncate selection identifies novel genes affecting bone mineral density and fracture risk. PLoS Genet 7:e1001372
29. **Kou I, Takahashi A, Urano T, Fukui N, Ito H, Ozaki K, Tanaka T, Hosoi T, Shiraki M, Inoue S, Nakamura Y, Kamatani N, Kubo M, Mori S, Ikegawa S** 2011 Common variants in a novel gene, FONG on chromosome 2q33.1 confer risk of osteoporosis in Japanese. PLoS One 6:e19641
30. **Estrada K, Styrkarsdottir U, Evangelou E, Hsu YH, Duncan EL, Ntzani EE, Oei L, Albagha OM, Amin N, Kemp JP, Koller DL, Li G, Liu CT, Minster RL, Moayyeri A, Vandenput L, Willner D, Xiao SM, Yerges-Armstrong LM, Zheng HF, Alonso N, Eriksson J, Kammerer CM, Kaptoge SK, Leo PJ, Thorleifsson G, Wilson SG, Wilson JF, Aalto V, Alen M, Aragaki AK, Aspelund T, Center JR, Dailiana Z, Duggan DJ, Garcia M, *et al.*** 2012 Genome-wide meta-analysis identifies 56 bone mineral density loci and reveals 14 loci associated with risk of fracture. Nat Genet 44:491–501
31. **Guo Y, Zhang LS, Yang TL, Tian Q, Xiong DH, Pei YF, Deng HW** 2010 IL21R and PTH may underlie variation of femoral neck bone mineral density as revealed by a genome-wide association study. J Bone Miner Res 25:1042–1048
32. **Timpson NJ, Tobias JH, Richards JB, Soranzo N, Duncan EL, Sims AM, Whittaker P, Kumanduri V, Zhai G, Glaser B, Eisman J, Jones G, Nicholson G, Prince R, Seeman E, Spector TD, Brown MA, Peltonen L, Smith GD, Deloukas P, Evans DM** 2009 Common variants in the region around Osterix are associated with bone mineral density and growth in childhood. Hum Mol Genet 18: 1510–1517
33. **Tan L, Liu R, Lei S, Pan R, Yang T, Yan H, Pei Y, Yang F, Zhang F, Pan F, Zhang Y, Hu H, Levy S, Deng H** 2010 A genome-wide association analysis implicates SOX6 as a candidate gene for wrist bone mass. Sci China Life Sci 53:1065–1072
34. **Cho YS, Go MJ, Kim YJ, Heo JY, Oh JH, Ban HJ, Yoon D, Lee MH, Kim DJ, Park M, Cha SH, Kim JW, Han BG, Min H, Ahn Y, Park MS, Han HR, Jang HY, Cho EY, Lee JE, Cho NH, Shin C, Park T, Park JW, Lee JK, Cardon L, Clarke G, McCarthy MI, Lee JY, Lee JK, Oh B, Kim HL** 2009 A large-scale genome-wide association study of Asian populations uncovers genetic factors influencing eight quantitative traits. Nat Genet 41:527–534
35. **Paternoster L, Lorentzon M, Vandenput L, Karlsson MK, Ljunggren O, Kindmark A, Mellstrom D, Kemp JP, Jarett CE, Holly JM, Sayers A, St Pourcain B, Timpson NJ, Deloukas P, Davey Smith G, Ring SM, Evans DM, Tobias JH, Ohlsson C** 2010 Genome-wide association meta-analysis of cortical bone mineral density unravels allelic heterogeneity at the RANKL locus and potential pleiotropic effects on bone. PLoS Genet 6:e1001217
36. **Liu YZ, Wilson SG, Wang L, Liu XG, Guo YF, Li J, Yan H, Deloukas P, Soranzo N, Chinappen-Horsley U, Cervino A, Williams FM, Xiong DH, Zhang YP, Jin TB, Levy S, Papasian CJ, Drees BM, Hamilton JJ, Recker RR, Spector TD, Deng HW** 2008 Identification of PLCL1 gene for hip bone size variation in females in a genome-wide association study. PLoS One 3:e3160
37. **Zhao LJ, Liu XG, Liu YZ, Liu YJ, Papasian CJ, Sha BY, Pan F, Guo YF, Wang L, Yan H, Xiong DH, Tang ZH, Yang TL, Chen XD, Guo Y, Li J, Shen H, Zhang F, Lei SF, Recker RR, Deng HW** 2010 Genome-wide association study for femoral neck bone geometry. J Bone Miner Res 25:320–329
38. **Guo Y, Tan LJ, Lei SF, Yang TL, Chen XD, Zhang F, Chen Y, Pan F, Yan H, Liu X, Tian Q, Zhang ZX, Zhou Q, Qiu C, Dong SS, Xu XH, Guo YF, Zhu XZ, Liu SL, Wang XL, Li X, Luo Y, Zhang LS, Li M, Wang JT, Wen T, Drees B, Hamilton J, Papasian CJ, Recker RR, Song XP, Cheng J, Deng HW** 2010 Genome-wide association study identifies ALDH7A1 as a novel susceptibility gene for osteoporosis. PLoS Genet 6:e1000806
39. **Albagha OM, Visconti MR, Alonso N, Langston AL, Cundy T, Dargie R, Dunlop MG, Fraser WD, Hooper MJ, Isaia G, Nicholson GC, del Pino Montes J, Gonzalez-Sarmiento R, di Stefano M, Tenesa A, Walsh JP, Ralston SH** 2010 Genome-wide association study identifies variants at CSF1, OPTN and TNFRSF11A as ge-

netic risk factors for Paget's disease of bone. Nat Genet 42:520–524

40. **Albagha OM, Wani SE, Visconti MR, Alonso N, Goodman K, Brandi ML, Cundy T, Chung PY, Dargie R, Devogelaer JP, Falchetti A, Fraser WD, Gennari L, Gianfrancesco F, Hooper MJ, Van Hul W, Isaia G, Nicholson GC, Nuti R, Papapoulos S, Montes Jdel P, Ratajczak T, Rea SL, Rendina D, Gonzalez-Sarmiento R, Di Stefano M, Ward LC, Walsh JP, Ralston SH** 2011 Genome-wide association identifies three new susceptibility loci for Paget's disease of bone. Nat Genet 43:685–689
41. **Lambert CG, Black LJ** 2012 Learning from our GWAS mistakes: from experimental design to scientific method. Biostatistics 13:195–203
42. **Sebastiani P, Solovieff N, Puca A, Hartley SW, Milista E, Andersen S, Dworkis DA, Wilk JB, Myers RH, Steinberg MH, Montano M, Baldwin CT, Perls TT** 2011 Retraction. Science 330:404
43. **Law B, Buckleton JS, Triggs CM, Weir BS** 2003 Effects of population structure and admixture on exact tests for association between loci. Genetics 164:381–387
44. **Falush D, Stephens M, Pritchard JK** 2003 Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genetics 164:1567–1587
45. **Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D** 2006 Principal components analysis corrects for stratification in genome-wide association studies. Nat Genet 38:904–909
46. **Devlin B, Roeder K** 1999 Genomic control for association studies. Biometrics 55:997–1004
47. **Devlin B, Bacanu SA, Roeder K** 2004 Genomic control to the extreme. Nat Genet 36:1129–1130; author reply 1131
48. **Barrett JC, Cardon LR** 2006 Evaluating coverage of genome-wide association studies. Nat Genet 38:659–662
49. **Liu YZ, Liu YJ, Recker RR, Deng HW** 2003 Molecular studies of identification of genes for osteoporosis: the 2002 update. J Endocrinol 177:147–196
50. **Liu YJ, Shen H, Xiao P, Xiong DH, Li LH, Recker RR, Deng HW** 2006 Molecular genetic studies of gene identification for osteoporosis: a 2004 update. J Bone Miner Res 21:1511–1535
51. **Richards JB, Kavvoura FK, Rivadeneira F, Styrkársdóttir U, Estrada K, Halldórsson BV, Hsu YH, Zillikens MC, Wilson SG, Mullin BH, Amin N, Aulchenko YS, Cupples LA, Deloukas P, Demissie S, Hofman A, Kong A, Karasik D, van Meurs JB, Oostra BA, Pols HA, Sigurdsson G, Thorsteinsdottir U, Soranzo N, Williams FM, Zhou Y, Ralston SH, Thorleifsson G, van Duijn CM, Kiel DP, Stefansson K, Uitterlinden AG, Ioannidis JP, Spector TD** 2009 Collaborative meta-analysis: associations of 150 candidate genes with osteoporosis and osteoporotic fracture. Ann Intern Med 151:528–537
52. **Xu XH, Dong SS, Guo Y, Yang TL, Lei SF, Papasian CJ, Zhao M, Deng HW** 2010 Molecular genetic studies of gene identification for osteoporosis: the 2009 update. Endocr Rev 31:447–505
53. **Sims AM, Shephard N, Carter K, Doan T, Dowling A, Duncan EL, Eisman J, Jones G, Nicholson G, Prince R, Seeman E, Thomas G, Wass JA, Brown MA** 2008 Genetic analyses in a sample of individuals with high or low BMD shows association with multiple Wnt pathway genes. J Bone Miner Res 23:499–506
54. **Raychaudhuri S, Plenge RM, Rossin EJ, Ng AC, Purcell SM, Sklar P, Scolnick EM, Xavier RJ, Altshuler D, Daly MJ** 2009 Identifying relationships among genomic disease regions: predicting genes at pathogenic SNP associations and rare deletions. PLoS Genet 5:e1000534
55. **Liu CT, Estrada K, Yerges-Armstrong LM, Amin N, Evangelou E, Li G, Minster RL, Carless MA, Kammerer CM, Oei L, Zhou Y, Alonso N, Dailiana Z, Eriksson J, García-Giralt N, Giroux S, Husted LB, Khusainova RI, Koromila T, Kung AW, Lewis JR, Masi L, Mencej-Bedrac S, Nogues X, Patel MS, Prezelj J, Richards JB, Sham PC, Spector T, Vandenput L, Xiao SM, Zheng HF, Zhu K, Balcells S, Brandi ML, Frost M, Goltzman D, *et al.*** 12 June 2012 Assessment of gene-by-sex interaction effect on bone mineral density. J Bone Miner Res doi: 10.1002/jbmr.1679
56. **Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, Hunter DJ, McCarthy MI, Ramos EM, Cardon LR, Chakravarti A, Cho JH, Guttmacher AE, Kong A, Kruglyak L, Mardis E, Rotimi CN, Slatkin M, Valle D, Whittemore AS, Boehnke M, Clark AG, Eichler EE, Gibson G, Haines JL, Mackay TF, McCarroll SA, Visscher PM** 2009 Finding the missing heritability of complex diseases. Nature 461:747–753
57. **Lango Allen H, Estrada K, Lettre G, Berndt SI, Weedon MN, Rivadeneira F, Willer CJ, Jackson AU, Vedantam S, Raychaudhuri S, Ferreira T, Wood AR, Weyant RJ, Segrè AV, Speliotes EK, Wheeler E, Soranzo N, Park JH, Yang J, Gudbjartsson D, Heard-Costa NL, Randall JC, Qi L, Vernon Smith A, Mägi R, Pastinen T, Liang L, Heid IM, Luan J, Thorleifsson G, Winkler TW, Goddard ME, Sin Lo K, Palmer C, Workalemahu T, *et al.*** 2010 Hundreds of variants clustered in genomic loci and biological pathways affect human height. Nature 467:832–838
58. **Speliotes EK, Willer CJ, Berndt SI, Monda KL, Thorleifsson G, Jackson AU, Allen HL, Lindgren CM, Luan J, Mägi R, Randall JC, Vedantam S, Winkler TW, Qi L, Workalemahu T, Heid IM, Steinthorsdottir V, Stringham HM, Weedon MN, Wheeler E, Wood AR, Ferreira T, Weyant RJ, Segrè AV, Estrada K, Liang L, Nemesh J, Park JH, Gustafsson S, Kilpeläinen TO, Yang J, Bouatia-Naji N, Esko T, Feitosa MF, Kutalik Z, Mangino M, *et al.*** 2010 Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. Nat Genet 42:937–948
59. **Anderson CA, Boucher G, Lees CW, Franke A, D'Amato M, Taylor KD, Lee JC, Goyette P, Imielinski M, Latiano A, Lagacé C, Scott R, Amininejad L, Bumpstead S, Baidoo L, Baldassano RN, Barclay M, Bayless TM, Brand S, Büning C, Colombel JF, Denson LA, De Vos M, Dubinsky M, Edwards C, Ellinghaus D, Fehrmann RS, Floyd JA, Florin T, Franchimont D, Franke L, Georges M, Glas J, Glazer NL, Guthery SL, Haritunians T, Hayward NK, *et al.*** 2011 Meta-analysis identifies 29 additional ulcerative colitis risk loci, increasing the number of confirmed associations to 47. Nat Genet 43:246–252
60. **Tran BN, Nguyen ND, Nguyen VX, Center JR, Eisman JA, Nguyen TV** 2011 Genetic profiling and individualized prognosis of fracture. J Bone Miner Res 26:414–419
61. **Redon R, Ishikawa S, Fitch KR, Feuk L, Perry GH, Andrews TD, Fiegler H, Shapero MH, Carson AR, Chen W, Cho EK, Dallaire S, Freeman JL, González JR, Gratacòs M, Huang J, Kalaitzopoulos D, Komura D, MacDonald JR, Marshall CR, Mei R, Montgomery L, Nishimura K, Okamura K, Shen F, Somerville MJ, Tchinda J, Valsesia A, Woodwark C, Yang F, Zhang J, Zerjal T, Zhang J, Armengol L, Conrad DF, Estivill X, Tyler-Smith C, Carter NP, Aburatani H, Lee C, Jones KW, Scherer SW, Hurles ME** 2006 Global variation in copy number in the human genome. Nature 444:444–454
62. **Stranger BE, Forrest MS, Dunning M, Ingle CE, Beazley C, Thorne N, Redon R, Bird CP, de Grassi A, Lee C, Tyler-Smith C, Carter N, Scherer SW, Tavaré S, Deloukas P, Hurles ME, Dermitzakis ET** 2007 Relative impact of nucleotide and copy number variation on gene expression phenotypes. Science 315:848–853
63. **Mileyko Y, Joh RI, Weitz JS** 2008 Small-scale copy number variation and large-scale changes in gene expression. Proc Natl Acad Sci USA 105:16659–16664
64. **Yang TL, Chen XD, Guo Y, Lei SF, Wang JT, Zhou Q, Pan F, Chen Y, Zhang ZX, Dong SS, Xu XH, Yan H, Liu X, Qiu C, Zhu XZ, Chen T, Li M, Zhang H, Zhang L, Drees BM, Hamilton JJ, Papasian CJ, Recker RR, Song XP, Cheng J, Deng HW** 2008 Genome-wide copy-number-variation study identified a susceptibility gene, UGT2B17, for osteoporosis. Am J Hum Genet 83:663–674
65. **Deng FY, Zhao LJ, Pei YF, Sha BY, Liu XG, Yan H, Wang L, Yang TL, Recker RR, Papasian CJ, Deng HW** 2010 Genome-wide copy number variation association study suggested VPS13B gene for osteoporosis in Caucasians. Osteoporos Int 21:579–587
66. **Liu SL, Lei SF, Yang F, Li X, Liu R, Nie S, Liu XG, Yang TL, Guo**

Y, Deng FY, Tian Q, Li J, Liu YZ, Liu YJ, Shen H, Deng HW 2011 Copy number variation in CNP267 region may be associated with hip bone size. PLoS One 6:e22035

67. **Chew S, Mullin BH, Lewis JR, Spector TD, Prince RL, Wilson SG** 2011 Homozygous deletion of the UGT2B17 gene is not associated with osteoporosis risk in elderly Caucasian women. Osteoporos Int 22:1981–1986
68. **Yang J, Manolio TA, Pasquale LR, Boerwinkle E, Caporaso N, Cunningham JM, de Andrade M, Feenstra B, Feingold E, Hayes MG, Hill WG, Landi MT, Alonso A, Lettre G, Lin P, Ling H, Lowe W, Mathias RA, Melbye M, Pugh E, Cornelis MC, Weir BS, Goddard ME, Visscher PM** 2011 Genome partitioning of genetic variation for complex traits using common SNPs. Nat Genet 43:519–525
69. **Stephens M, Balding DJ** 2009 Bayesian statistical methods for genetic association studies. Nat Rev Genet 10:681–690
70. 2007 Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 447:661–678
71. **Stahl EA, Wegmann D, Trynka G, Gutierrez-Achury J, Do R, Voight BF, Kraft P, Chen R, Kallberg HJ, Kurreeman FA, Kathiresan S, Wijmenga C, Gregersen PK, Alfredsson L, Siminovitch KA, Worthington J, de Bakker PI, Raychaudhuri S, Plenge RM** 2012 Bayesian inference analyses of the polygenic architecture of rheumatoid arthritis. Nat Genet 44:483–489
72. **Wakefield J** 2007 A Bayesian measure of the probability of false discovery in genetic epidemiology studies. Am J Hum Genet 81: 208–227
73. **Nakaoka H, Inoue I** 2009 Meta-analysis of genetic association studies: methodologies, between-study heterogeneity and winner's curse. J Hum Genet 54:615–623
74. **Cochran WG** 1954 The combination of estimates from different experiments. Biometrics 10:101–129
75. **Higgins JP, Thompson SG** 2002 Quantifying heterogeneity in a meta-analysis. Stat Med 21:1539–1558
76. **Whitlock MC** 2005 Combining probability from independent tests: the weighted Z-method is superior to Fisher's approach. J Evol Biol 18:1368–1373
77. **DerSimonian R, Laird N** 1986 Meta-analysis in clinical trials. Control Clin Trials 7:177–188
78. **Li Y, Vinckenbosch N, Tian G, Huerta-Sanchez E, Jiang T, Jiang H, Albrechtsen A, Andersen G, Cao H, Korneliussen T, Grarup N, Guo Y, Hellman I, Jin X, Li Q, Liu J, Liu X, Sparsø T, Tang M, Wu H, Wu R, Yu C, Zheng H, Astrup A, Bolund L, Holmkvist J, Jørgensen T, Kristiansen K, Schmitz O, Schwartz TW, Zhang X, Li R, Yang H, Wang J, Hansen T, Pedersen O, Nielsen R, Wang J** 2010 Resequencing of 200 human exomes identifies an excess of low-frequency non-synonymous coding variants. Nat Genet 42:969–972
79. **Gorlov IP, Gorlova OY, Sunyaev SR, Spitz MR, Amos CI** 2008 Shifting paradigm of association studies: value of rare single-nucleotide polymorphisms. Am J Hum Genet 82:100–112
80. **Dickson SP, Wang K, Krantz I, Hakonarson H, Goldstein DB** 2010 Rare variants create synthetic genome-wide associations. PLoS Biol 8:e1000294
81. **Goldstein DB** 2011 The importance of synthetic associations will only be resolved empirically. PLoS Biol 9:e1001008
82. **Anderson CA, Soranzo N, Zeggini E, Barrett JC** 2011 Synthetic associations are unlikely to account for many common disease genome-wide association signals. PLoS Biol 9:e1000580
83. **Kryukov GV, Shpunt A, Stamatoyannopoulos JA, Sunyaev SR** 2009 Power of deep, all-exon resequencing for discovery of human trait genes. Proc Natl Acad Sci USA 106:3871–3876
84. **Cohen JC, Kiss RS, Pertsemlidis A, Marcel YL, McPherson R, Hobbs HH** 2004 Multiple rare alleles contribute to low plasma levels of HDL cholesterol. Science 305:869–872
85. **Nejentsev S, Walker N, Riches D, Egholm M, Todd JA** 2009 Rare variants of IFIH1, a gene implicated in antiviral responses, protect against type 1 diabetes. Science 324:387–389
86. **Wood LD, Parsons DW, Jones S, Lin J, Sjöblom T, Leary RJ, Shen D, Boca SM, Barber T, Ptak J, Silliman N, Szabo S, Dezso Z, Ustyanksky V, Nikolskaya T, Nikolsky Y, Karchin R, Wilson PA, Kaminker JS, Zhang Z, Croshaw R, Willis J, Dawson D, Shipitsin M, Willson JK, Sukumar S, Polyak K, Park BH, Pethiyagoda CL, Pant PV, Ballinger DG, Sparks AB, Hartigan J, Smith DR, Suh E, Papadopoulos N, Buckhaults P, Markowitz SD, Parmigiani G, Kinzler KW, Velculescu VE, Vogelstein B** 2007 The genomic landscapes of human breast and colorectal cancers. Science 318:1108–1113
87. **Romeo S, Pennacchio LA, Fu Y, Boerwinkle E, Tybjaerg-Hansen A, Hobbs HH, Cohen JC** 2007 Population-based resequencing of ANGPTL4 uncovers variations that reduce triglycerides and increase HDL. Nat Genet 39:513–516
88. **Heinzen EL, Radtke RA, Urban TJ, Cavalleri GL, Depondt C, Need AC, Walley NM, Nicoletti P, Ge D, Catarino CB, Duncan JS, Kasperaviciûte D, Tate SK, Caboclo LO, Sander JW, Clayton L, Linney KN, Shianna KV, Gumbs CE, Smith J, Cronin KD, Maia JM, Doherty CP, Pandolfo M, Leppert D, Middleton LT, Gibson RA, Johnson MR, Matthews PM, Hosford D, Kälviäinen R, Eriksson K, Kantanen AM, Dorn T, Hansen J, Krämer G, Steinhoff BJ, Wieser HG, Zumsteg D, Ortega M, Wood NW, Huxley-Jones J, Mikati M, Gallentine WB, Husain AM, Buckley PG, Stallings RL, Podgoreanu MV, Delanty N, Sisodiya SM, Goldstein DB** 2010 Rare deletions at 16p13.11 predispose to a diverse spectrum of sporadic epilepsy syndromes. Am J Hum Genet 86:707–718
89. **Ji W, Foo JN, O'Roak BJ, Zhao H, Larson MG, Simon DB, Newton-Cheh C, State MW, Levy D, Lifton RP** 2008 Rare independent mutations in renal salt handling genes contribute to blood pressure variation. Nat Genet 40:592–599
90. **O'Roak BJ, Deriziotis P, Lee C, Vives L, Schwartz JJ, Girirajan S, Karakoc E, Mackenzie AP, Ng SB, Baker C, Rieder MJ, Nickerson DA, Bernier R, Fisher SE, Shendure J, Eichler EE** 2011 Exome sequencing in sporadic autism spectrum disorders identifies severe de novo mutations. Nat Genet 43:585–589
91. **Morris AP, Zeggini E** 2010 An evaluation of statistical approaches to rare variant analysis in genetic association studies. Genet Epidemiol 34:188–193
92. **Wang K, Dickson SP, Stolle CA, Krantz ID, Goldstein DB, Hakonarson H** 2010 Interpretation of association signals and identification of causal variants from genome-wide association studies. Am J Hum Genet 86:730–742
93. **Hoffmann TJ, Kvale MN, Hesselson SE, Zhan Y, Aquino C, Cao Y, Cawley S, Chung E, Connell S, Eshragh J, Ewing M, Gollub J, Henderson M, Hubbell E, Iribarren C, Kaufman J, Lao RZ, Lu Y, Ludwig D, Mathauda GK, McGuire W, Mei G, Miles S, Purdy MM, Quesenberry C, Ranatunga D, Rowell S, Sadler M, Shapero MH, Shen L, Shenoy TR, Smethurst D, Van den Eeden SK, Walter L, Wan E, Wearley R, Webster T, Wen CC, Weng L, Whitmer RA, Williams A, Wong SC, Zau C, Finn A, Schaefer C, Kwok PY, Risch N** 2011 Next generation genome-wide association tool: design and coverage of a high-throughput European-optimized SNP array. Genomics 98:79–89
94. **Zuk O, Hechter E, Sunyaev SR, Lander ES** 2012 The mystery of missing heritability: genetic interactions create phantom heritability. Proc Natl Acad Sci USA 109:1193–1198
95. **Marchini J, Donnelly P, Cardon LR** 2005 Genome-wide strategies for detecting multiple loci that influence complex diseases. Nat Genet 37:413–417
96. **Ober C, Loisel DA, Gilad Y** 2008 Sex-specific genetic architecture of human disease. Nat Rev Genet 9:911–922
97. **Karasik D, Ferrari SL** 2008 Contribution of gender-specific genetic factors to osteoporosis risk. Ann Hum Genet 72:696–714
98. **Barabási AL** 2007 Network medicine—from obesity to the "diseasome". N Engl J Med 357:404–407
99. **Freimer N, Sabatti C** 2003 The human phenome project. Nat Genet 34:15–21
100. **Goh KI, Cusick ME, Valle D, Childs B, Vidal M, Barabási AL** 2007 The human disease network. Proc Natl Acad Sci USA 104:8685–8690

101. **Karasik D, Hsu YH, Zhou Y, Cupples LA, Kiel DP, Demissie S** 2010 Genome-wide pleiotropy of osteoporosis-related phenotypes: the Framingham Study. J Bone Miner Res 25:1555–1563
102. **Thorleifsson G, Holm H, Edvardsson V, Walters GB, Styrkarsdottir U, Gudbjartsson DF, Sulem P, Halldorsson BV, de Vegt F, d'Ancona FC, den Heijer M, Franzson L, Christiansen C, Alexandersen P, Rafnar T, Kristjansson K, Sigurdsson G, Kiemeney LA, Bodvarsson M, Indridason OS, Palsson R, Kong A, Thorsteinsdottir U, Stefansson K** 2009 Sequence variants in the CLDN14 gene associate with kidney stones and bone mineral density. Nat Genet 41:926–930
103. **Billings L, Hsu YH, Ackerman R, Dupuis J, Voight B, Rasmussen-Torvik LJ, Hercberg S, Lathrop M, Barnes D, Langenberg C, Hui J, Fu M, Bouatia-Naji N, Lecoeur C, An P, Magnusson PK, Surakka I, Ripatti S, Christiansen L, Dalgård C, Folkersen L, Grundberg E, the MAGIC Investigators, the DIAGRAM Consortium, the MuTHER Consortium, *et al.*** 2012 Leveraging genome-wide association data to detect genetic pleiotropy at loci associated with glycemic and skeletal phenotypes. Diabetes 61:2176–2186
104. **Franke A, McGovern DP, Barrett JC, Wang K, Radford-Smith GL, Ahmad T, Lees CW, Balschun T, Lee J, Roberts R, Anderson CA, Bis JC, Bumpstead S, Ellinghaus D, Festen EM, Georges M, Green T, Haritunians T, Jostins L, Latiano A, Mathew CG, Montgomery GW, Prescott NJ, Raychaudhuri S, Rotter JI, Schumm P, Sharma Y, Simms LA, Taylor KD, Whiteman D, Wijmenga C, Baldassano RN, Barclay M, Bayless TM, Brand S, *et al.*** 2010 Genome-wide meta-analysis increases to 71 the number of confirmed Crohn's disease susceptibility loci. Nat Genet 42:1118–1125
105. **Zheng W, Long J, Gao YT, Li C, Zheng Y, Xiang YB, Wen W, Levy S, Deming SL, Haines JL, Gu K, Fair AM, Cai Q, Lu W, Shu XO** 2009 Genome-wide association study identifies a new breast cancer susceptibility locus at 6q25.1. Nat Genet 41:324–328
106. **Hsu YH, Chen X,** A multivariate approach on genome-wide association studies (GWAS) by modeling multiple traits simultaneously to identify pleiotropic genetic effects (invited papers). The Joint Statistical Meetings/American Statistics Association, Miami Beach, FL, 2011 (Abstract #300323)
107. **Hsu YH, Chen X, Elks C, Murabito J, Estrada K, Harris TB, Lunetta KL, Murray A, Ong KK, Perry JR, Rivadeneira FR, Stolk L, Uitterlinden AG, Zillikens CM, Kiel DP, Karasik D,** Genome-wide bivariate association analysis identifies novel candidate genes for ages at menarche and natural menopause and for bone mineral density: the ReproGen and GEFOS Consortia. Program of the 93rd Annual Meeting of The Endocrine Society, Boston, MA, 2011 (Abstract OR17-4)
108. **Hsu YH, Chen X, Zillikens C, Estrada K, Demissie S, Liu C, Zhou Y, Karasik D, Murabito J, Uitterlinden A, Cupples L, Rivadeneira F, Kiel D** 2011 Multi-phenotype genome-wide association meta-analysis on both lean body mass and BMD identified novel pleiotropic genes that affected skeletal muscle and bone metabolism in European descent Caucasian populations. J Bone Miner Res 26:(Suppl 1):S56
109. **Hsu Y, Beck TJ, Brown SJ, Cauley JA, Demissie S, Estrada K, Glazer N, Kaptoge SK, Kiel DP, Richards B, Rivadeneira F, Robbins JA, Soranzo N, Spector TD, Streeten EA, Uitterlinden AG, Wilson SG, Yerges-Armstrong LM, Zillikens C, Karasik D** 2010 Meta-analysis of genome-wide association study (GWAS) identifies several genes for hip bone geometry in Caucasians: the Genetic Factors for Osteoporosis (GEFOS) Consortium. J Bone Miner Res 25(Suppl 1):S448
110. **Cheung CL, Hsu YH, Eriksson J, Kwan J, Karasik D, Benjamin E, Fontes J, Kiel DP, Lorentzon M, Larson M, Kung A, Chee KW, Ohlsson C, Sham PC, Dupuis J** 2011 Meta-analysis of genome-wide association study identifies three loci influencing serum osteoprotegerin levels. J Bone Miner Res 26(Suppl 1):S442
111. **Jackson R, LaCroix A, Aragaki A, Duggan D, Carlson C, Richards B, Dastani Z, Spector T, Eisman J, Nguyen T, Ralston S, McCloskey E, Duncan E, Nicholson G, Leslie W, Goltzman D, Krishnan S, Kooperberg C** 2011 Genome-wide association study of hip fracture suggests loci in biological pathways associated with bone remodeling. J Bone Miner Res 26(Suppl 1):S346
112. **Oei L, Zheng HF, Ntzani E, Estrada K, Ridker P, Garcia ME, Hsu YH, Lehtimäki T, Trompet S, Kaptoge SK, Wilson S, Liu Y, Eriksson J, Amin N, Kung A, Gomez CM, Tsilidis K, Rose K, Vernon-Smith A, Liu CT, Viikari J, Moayyeri A, Minster R, Lorentzon M, Xiao SM, Wareham NJ, Aalto V, Aspelund T, Cummings S, Cupples LA, Evangelou E, Khaw KT, Luben RN, Raitakari O, Siggeirsdottir K, van Duijn CM, Cauley J, Karasik D, Sham PC, Zillikens C, Ohlsson C, Harris TB, Spector TD, Reeve J, Jukema W, Ioannidis J, Kahonen M, Kiel DP, Gudnason V, Uitterlinden AG, Chasman DI, Richards B, Rivadeneira F,** Large-scale meta-analysis of genome-wide association studies for fracture risk: The GEFOS Consortium. Proc 61st Annual Meeting of the American Society of Human Genetics, Montreal, Canada, 2011 (Abstract 847W)
113. **Kanis JA, Johnell O, Oden A, Johansson H, McCloskey E** 2008 FRAX and the assessment of fracture probability in men and women from the UK. Osteoporos Int 19:385–397