



Genome-wide search for higher order epistasis as modifiers of treatment effects on bone mineral density in childhood cancer survivors

Cindy Im¹ · Kirsten K. Ness² · Sue C. Kaste^{3,4} · Wassim Chemaitilly^{2,5} · Wonjong Moon² · Yadav Sapkota² · Russell J. Brooke² · Melissa M. Hudson^{2,4} · Leslie L. Robison² · Yutaka Yasui^{1,2} · Carmen L. Wilson²

Received: 24 April 2017 / Revised: 8 November 2017 / Accepted: 23 November 2017 / Published online: 18 January 2018

© European Society of Human Genetics 2018

Abstract

Single-nucleotide polymorphisms (SNPs) contributing to interactions between regulatory elements that modulate gene transcription may explain some of the uncharacterized variation for complex traits. We explored this hypothesis among 856 adult survivors of pediatric cancer exposed to curative treatments that adversely affect bone mineral density (BMD). To restrict our search to interactions among SNPs in regulatory elements, our analysis considered 75523 SNPs mapped to putative promoter or enhancer regions. In anticipation that power to detect higher order epistasis would be low using an exhaustive search and a Bonferroni-corrected threshold for genome-wide significance (e.g., $P < 5.6 \times 10^{-14}$), a novel non-exhaustive statistical algorithm was implemented to detect chromosome-wide three-way regulatory interactions. We used a permutation-based evaluation statistic to identify candidate SNP interactions with stronger associations with BMD than expected. Of the six regulatory 3-SNP interactions identified as candidate interactions ($P < 3.5 \times 10^{-11}$) among cancer survivors exposed to treatments, five were replicated in an independent cohort of survivors ($N = 1428$) as modifiers of treatment effects on BMD ($P < 0.05$). Analyses with publicly available bioinformatics data revealed that SNPs contributing to replicated interactions were enriched for gene expressions ($P = 3.6 \times 10^{-4}$) and enhancer states ($P < 0.05$) in cells relevant for bone biology. For each replicated interaction, implicated SNPs were within or directly adjacent to 100-kb windows of genomic regions that plausibly physically interact in lymphoblastoid cells. Our study demonstrates the utility of a hypothesis-driven approach in revealing epistasis associated with complex traits.

Yutaka Yasui and Carmen L. Wilson contributed equally to this work

Electronic supplementary material The online version of this article (<https://doi.org/10.1038/s41431-017-0050-x>) contains supplementary material, which is available to authorized users.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

✉ Carmen L. Wilson
carmen.wilson@stjude.org

- ¹ School of Public Health, University of Alberta, Edmonton, AB, Canada
- ² Department of Epidemiology and Cancer Control, St. Jude Children's Research Hospital, Memphis, TN, USA
- ³ Department of Radiological Sciences, St. Jude Children's Research Hospital, Memphis, TN, USA
- ⁴ Department of Oncology, St. Jude Children's Research Hospital, Memphis, TN, USA
- ⁵ Division of Endocrinology, Department of Pediatric Medicine, St. Jude Children's Research Hospital, Memphis, TN, USA

Introduction

Survivors of pediatric acute lymphoblastic leukemia (ALL) are at risk for long-term deficits in bone mineral density (BMD) due to childhood cancer treatment exposures, including cranial radiation, antimetabolites (e.g., methotrexate), and glucocorticoids [1–3]. Cranial radiation diminishes BMD through injury to the hypothalamic–pituitary axis, affecting sex and growth hormone secretions that play an important role in bone metabolism [2]. Methotrexate and glucocorticoids decrease BMD by influencing factors that control osteoblast and osteoclast cell activity [2]. Despite common past treatment exposures, pediatric ALL survivors exhibit substantial variation in BMD later in life. An unexplored explanation for some of this uncharacterized variation in BMD is epistasis, where the effect of a locus on a trait is conditional on genotypes observed at other loci.

While studies have investigated pairs of SNPs in select candidate genes with BMD [4, 5], higher order epistasis

involving three or more SNPs is also likely to play a vital role in the genetic architecture of BMD. BMD reflects the cumulative effects of interacting genetic and environmental factors on peak bone mass and bone remodeling [6]. Signaling pathways requiring both spatiotemporal cues and epigenetic modifications of genetic loci guide the differentiation of bone cells from cells of mesenchymal and hematopoietic origin [7]. In a recent genome-wide scan of SNP pair interactions, over half of gene expressions in peripheral blood significantly associated with SNP pairs were influenced by networks involving three SNPs or more [8].

To our knowledge, no studies have explored higher order epistasis and BMD. In general, searches for epistasis are challenged in identifying true interactions between SNPs on a genome-wide scale, largely due to insufficient statistical power. Novel strategies have been applied to increase power and identify reliable interactions. One strategy is to restrict the search for epistasis to SNPs that are likely to contribute to biological interactions, reducing the number of tested interactions [9, 10]. Another strategy is to search for interactions with large effects on phenotypes [8]. Lastly, some epistatic interactions failing to meet conservative genome-wide significance thresholds have been shown to be reliable signals through replication [11].

In this study, we combined all of these strategies to identify higher order epistatic interactions that explain some of the variability of treatment effects on BMD among adult survivors of childhood ALL exposed to BMD-diminishing treatments. We leveraged knowledge that SNPs in interacting enhancer and promoter regions modulate gene expression and thus affect phenotypes [12, 13]. We applied chromatin state annotations [14] to restrict the search for epistasis to SNPs mapped to putative enhancer or promoter regions. To detect interactions between regulatory regions carrying SNPs associated with BMD (hereafter referred to as “SNP interactions”) as potential modifiers of treatment effects, a novel, non-exhaustive statistical algorithm was implemented. Our specific focus was to identify regulatory 3-way SNP interactions associated with BMD in ALL survivors. An independent cohort of cancer survivors was used to replicate candidate regulatory SNP interaction signals as modifiers of treatment effects on BMD. Supplemental bioinformatics analyses were conducted to characterize replicated SNP interactions.

Subjects and methods

Study cohorts

Individuals included in this analysis are participants in the St. Jude Lifetime Cohort Study (SJLIFE) [15]. Eligible

survivors were divided into two cohorts: a discovery cohort of 856 adult survivors of pediatric ALL and a replication cohort consisting of 1428 adult survivors of any non-ALL pediatric cancer (a second cohort of ALL survivors with comparable BMD measurements and genotype data was unavailable). BMD was ascertained using quantitative-computed tomography from the mid-bodies of the first and second lumbar vertebra. A BMD Z-score was computed for each survivor by taking the difference between the average of their two vertebral BMD measurements and the age-matched and sex-matched mean of a reference population, divided by the standard deviation in the reference population. Cumulative doses of cranial radiation (none, >0 to <2400, ≥ 2400 cGy), methotrexate (<5100, ≥ 5100 to <20000, ≥ 20000 mg/m²), and glucocorticoid (<2000, ≥ 2000 to <11000, ≥ 11000 mg/m²) treatment exposures were considered as risk factors for BMD deficiency among ALL survivors [1–3]. We built a multiple linear regression model for BMD Z-scores including sex, categorical treatment exposures, and genetic ancestry estimated using STRUCTURE software [16] (to control for population stratification in our multi-ethnic cohorts) for adjustment in subsequent genetic association analyses. Additional study cohort details are provided in Supplementary Methods.

Affymetrix Genome-Wide Human SNP Array 6.0 (Affymetrix, Santa Clara, CA) was used to genotype DNA samples in SJLIFE. Genotyping quality control thresholds for exclusion from the analysis in the discovery cohort were as follows: <95% per-sample call rate across markers, <95% SNP call rate across samples, MAF <1%, and Hardy–Weinberg equilibrium $P < 1 \times 10^{-6}$. Imputation of missing measured genotypes was completed using BEAGLE [17]. An allelic R^2 imputation quality score cutoff of >0.5 was applied. For replication, a per-sample call rate of >90% was used to avoid sample exclusion due to the smaller set of SNPs selected for follow-up. Otherwise, genotyping quality control thresholds were identical for discovery and replication cohorts. Imputation was not employed for replication to limit data uncertainty associated with imputation. Per-SNP missingness rates were comparable between cohorts (Supplementary Table 1). Genetic data is available in the European Genome-phenome Archive under study accession number EGAS00001002645 (<https://www.ebi.ac.uk/ega/studies/EGAS00001002645>).

Statistical methodology: discovery analysis

ChromHMM chromatin state annotations [14] were used to map SNPs to putative enhancer or promoter regions. We retained SNPs mapped to “strong enhancer” or “active promoter” elements in any of nine ChromHMM-annotated human cell types, since it was unknown which types would be most relevant. This restriction limited the search for SNP

interactions to 75523 SNPs. Each of these SNPs was dichotomized to create two binary variables, or indicators for carrying at least one non-reference allele or homozygous non-reference alleles. Any binary-encoded SNP variable with frequency <5% was removed to limit evaluations of sparse 3-SNP interactions. A total of 115800 binary SNP variables were retained genome-wide.

Despite this SNP restriction, exhaustively testing enhancer/promoter 3-SNP interactions would entail >260 trillion tests. To decrease the number of effective tests without compromising the search quality, we developed a non-exhaustive, sequential conditioning algorithm based on logic regression [18] to conduct an effective search of the 3-SNP interaction search space. Briefly, logic regression is an adaptive regression methodology that combines generalized linear models with a stochastic search algorithm to identify best-fitting models that include interaction variables (“logic trees”) comprised of binary predictors. Logic regression selects best-fitting models by comparing decrements in model scores. The search for 3-SNP interactions was also restricted by chromosome, reducing the computational burden to a manageable level.

Our algorithm-identified 3-SNP interactions sequentially via logic regression for each chromosome, using forward addition to form a linear predictor that included ten 3-SNP interaction trees per chromosome. Models for each chromosome took the following form:

$$E[Y]=\mu+\delta S+\sum_{j=1}^3\alpha_j A_j+\sum_{k=1}^6\gamma_k T_k+\sum_{p=1}^m\beta_p L_p, \quad (1)$$

where Y is BMD Z-score, S is sex, A_{1-3} are the three STRUCTURE genetic ancestry covariates, T_{1-6} are indicator variables for the three categorical treatment variables (i.e., three 3-level variables), and L_p are the 3-SNP interaction trees ($m = 1, 2, \dots, 10$ trees). By identifying 3-SNP interaction trees conditioned on previously identified trees, the algorithm guides the stochastic search in different directions, yielding 3-SNP interactions that are unlikely to be correlated.

We applied a permutation-based approach to identify candidate 3-SNP interactions for replication follow-up. For each of the algorithm-identified 3-SNP interaction trees, 1000 permutations of BMD Z-scores were used to compute the corresponding empirically derived median for the absolute value of the t -statistic and its median absolute deviation (MAD, a robust measure of variability). Permutations of BMD Z-score values were conditioned on 50 quantiles of the fitted BMD Z-score from the clinical baseline model to approximately preserve relationships between adjustment covariates and BMD Z-score. The search algorithm was applied to these conditionally permuted BMD Z-scores in the exact same manner as the

unpermuted case. To select candidate 3-SNP interactions for replication follow-up, we compared the observed t -statistic of a given tree with the corresponding empirically derived median, similar to the Significance Analysis of Microarray method [19]. Our evaluation statistic is $\frac{t_{\text{obs}} - t_{\text{med}}}{\text{MAD}_{t_{\text{med}}}}$, where t_{obs} is the absolute value of the t -statistic for the p^{th} tree given $(p-1)$ observed trees, and t_{med} and $\text{MAD}_{t_{\text{med}}}$ are the median and MAD, respectively, of the corresponding 1000 conditioned t -statistics from 1000 permutations. If the evaluation statistic was >2 (i.e., t_{obs} was >2 median absolute deviations away from its corresponding permutation-derived median), the 3-SNP interaction tree was selected as a candidate interaction for replication, as the tree explains an unusually large proportion of BMD Z-score variation than expected by chance alone.

After selecting candidate regulatory 3-SNP interaction trees (“original” trees), a “neighborhood” analysis was conducted to identify “proximal” 3-SNP interaction trees with stronger associations with BMD Z-score than original trees. The reasons for conducting this analysis were twofold: (1) our non-exhaustive logic regression-based algorithm may have missed proximal SNP interactions with stronger associations with BMD; and (2) these strongly associated neighborhood SNP interactions may include SNPs that “tag” additional regulatory regions relevant for BMD. Neighborhood trees were constructed with binary-encoded SNP variables (same filtering criteria as the discovery analysis) from SNPs ± 100 kb of SNPs in the original tree, with the same Boolean logic structure as the corresponding original tree. Neighborhood trees selected for follow-up in the replication cohort explained larger proportions of BMD Z-score variation than their corresponding original trees.

Statistical methodology: replication analysis

Since every participant in the discovery cohort received substantial cumulative doses of at least one of the three treatments known to affect BMD, we expected that interaction signals observed in the discovery cohort were potential modifiers of treatment effects on BMD. We therefore defined evidence of replication as significant modification of treatment effects by 3-SNP trees in the replication cohort. We assessed modification of treatment effects using two different approaches: (1) 3-SNP tree interactions with each of the three treatments, and (2) 3-SNP tree main effects among those exposed to each of the three treatments. If the 3-SNP tree had a significant interaction ($P < 0.05$) with at least one of the three treatments or a significant main effect ($P < 0.05$) among those exposed to one of the treatments, we deemed the interaction to be replicated. We further required treatment modification effects in the replication cohort to have the same direction and similar magnitude as the discovery cohort.

Additional explanation of the statistical methodology is given in Supplementary Methods.

Comparison of the proposed method to a benchmark 2-SNP interaction analysis method

We conducted an exhaustive, within-chromosome 2-way SNP interaction analysis among enhancer/promoter SNPs with the linear regression-based epistasis module in PLINK v1.90, a benchmark methodology for epistasis analysis [20]. We also performed a simulation study under three sample size scenarios ($N = 1000, 1500,$ and 2000) to compare the performance, measured by power and positive predictive value (PPV), of our proposed method and the benchmark method's detection of component SNP pairs for replicated 3-SNP interactions. Details for both analyses are provided in Supplementary Methods.

Biological characterization of replicated interactions

We evaluated whether there was an excess of significant gene expressions (expression quantitative trait loci or eQTLs) for SNPs in replicated 3-SNP interactions in bone-related cells/tissues using *cis*-eQTLs achieving study-wide significance from the Genotype-Tissue Expression (GTEx) Project [21] and GHS-Express monocyte transcriptome [22] databases. Using the BMD biology literature, we defined 16 cell or tissue groups to be related to bone out of 45 available cell/tissue groups. Counts of significant eQTLs in bone-related cells/tissues for SNPs of interest were compared to all other SNPs genome-wide with at least one significant eQTL in these databases (~2.6 million SNPs with ~26.4 million eQTLs) using a 2-sided Fisher's exact test.

To investigate the cell- and tissue-specificity of enhancer and promoter states for SNPs contributing to replicated interactions, we conducted enrichment analyses using the 15-state chromatin state annotation data for 127 consolidated human cell types from the Roadmap Epigenomics Mapping Consortium (REMC) [23]. For each cell type, we compared the set of SNPs in replicated interactions with the set of non-overlapping SNPs originally mapped to enhancers/promoters. Frequencies of overlap between SNPs in each set and REMC enhancer or promoter regions were counted in each cell type. Strength of evidence for enrichments was evaluated using a 2-sided Fisher's exact test.

Assays based on chromosome conformation capture (3C) enable study of physical interactions between chromatin regions [12, 13]. We evaluated the likelihood of physical interaction between SNP regions participating in replicated 3-SNP interactions using a publicly available Hi-C data library generated in lymphoblastoid cells [24], visualized with the WashU EpiGenome Browser resource [25].

Details for bioinformatics analyses are available in Supplementary Methods.

Results

The discovery cohort included 856 adult survivors of pediatric ALL. Cohort clinical characteristics are provided in Table 1. Every ALL survivor was exposed to cranial radiation therapy (CRT), methotrexate, and/or

Table 1 Participant characteristics

	Discovery cohort ^a ($N = 856$)	Replication cohort ^b ($N = 1428$)
Characteristic	n (%)	n (%)
Age at BMD measurement (years)		
Median (range)	31.3 (18.4–59.7)	31.6 (18.5–65.9)
Age at diagnosis (years)		
Median (range)	5.0 (0.2–19.5)	9.2 (0–24.8)
Sex		
Male	427 (49.9)	767 (53.7)
Female	429 (50.1)	661 (46.3)
Treatment profile ^c		
Cranial radiation (cGy)		
Median cumulative dose (range)	1800 (0–5100)	0 (0–10600)
None	348 (41.2)	1235 (86.9)
>0 to <2400	215 (25.4)	14 (1.0)
≥2400	282 (33.4)	172 (12.1)
Methotrexate (mg/m^2)		
Median cumulative dose (range)	5462 (85–83350)	0 (0–211900)
<5100	340 (39.9)	1306 (91.5)
≥5100 to <20,000	327 (38.3)	22 (1.5)
≥20,000	186 (21.8)	99 (6.9)
Glucocorticoids (mg/m^2)		
Median cumulative dose (range)	9520 (0–27360)	0 (0–14460)
<2000	328 (38.6)	1239 (86.8)
≥2000 to <11,000	355 (41.8)	160 (11.2)
≥11,000	166 (19.6)	28 (2.0)
BMD Z-score (expressed in SD)		
Median (range)	−0.4 (−3.5, 5.4)	−0.2 (−5.5, 6.0)
≤−1	256 (29.9)	349 (24.4)
≥1	104 (12.1)	249 (17.4)

^aAdult survivors of pediatric acute lymphoblastic leukemia (ALL)

^bAdult survivors of pediatric non-ALL cancers

^cDiscovery cohort: missing cranial radiation, methotrexate, and glucocorticoid cumulative dosage information for 11, 3, and 7 participants, respectively. Replication cohort: Missing cranial radiation, methotrexate, and glucocorticoid cumulative dosage information for 7, 1, and 1 participant(s), respectively

Table 2 Replicated 3-SNP interactions associated with BMD Z-score identified by the novel logic regression-based algorithm

Chr	Regulatory 3-SNP interaction ^a	Single-SNP rsID (hg19 HGVS identifier)	Marginal effect ^b SNP β (P)	Component 2-SNP interaction	Pair effect ^b pair β (P)	Discovery cohort SNP interaction β^c (95% CI; P-value; tree frequency)	Permutation-based evaluation statistic ^d P-value; tree frequency)	Replication cohort ^e replicated β_{int} (P-value; modified treatment effect)
2	rs901466 = {CC,CG} or (rs7569568 = {GG} or rs921319 = {CC})	rs901466 chr2: g.114724122 C>G	-0.457 (0.002)	rs901466 or rs7569568	-0.832 (5.5 × 10 ⁻⁵)	-1.304 (95% CI [-1.668, -0.939]; P = 4.7 × 10 ⁻¹² ; 810)	2.063	-1.769 (P = 0.004; Methotrexate)
		rs7569568 chr2: g.225264598 G>A	-0.139 (0.091)	rs901466 or rs921319	-0.437 (0.004)			
		rs921319 chr2: g.62367720 T>C	0.088 (0.405)	rs7569568 or rs921319	-0.181 (0.036)			
12	rs1020745 = {AG,GG} and (rs2110167 = {GA,AA} and rs10444471 = {GG})	rs1020745 chr12: g.53692955 G>A	0.589 (0.001)	rs1020745 and rs2110167	1.215 (2.8 × 10 ⁻⁸)	1.719 (95% CI [1.265, 2.174]; P = 2.9 × 10 ⁻¹³ ; 22)	2.775	1.402 (P = 0.013; Methotrexate)
		rs2110167 chr12: g.5734319 A>G	0.055 (0.461)	rs1020745 and rs10444471	0.790 (3.5 × 10 ⁻⁵)			
		rs10444471 chr12: g.4677211 G>T	0.081 (0.538)	rs2110167 and rs10444471	0.091 (0.220)			
12	(rs1894331 = {TT} or rs10773093 = {TC,CC}) and rs4768783 = {TT,TC}	rs1894331 chr12: g.11930889G>T	-0.137 (0.091)	rs1894331 or rs10773093	-0.409 (1.7 × 10 ⁻⁵)	-0.508 (95% CI [-0.649, -0.367]; P = 3.1 × 10 ⁻¹² ; 522)	2.514	-0.514 (P = 0.042; Methotrexate)
		rs10773093 chr12: g.125046036T>C	-0.208 (0.011)	rs1894331 and rs4768783	-0.236 (0.005)			
		rs4768783 chr12: g.47592945C>T	-0.263 (0.005)	rs10773093 and rs4768783	-0.294 (1.0 × 10 ⁻⁴)			
13	rs7321815 = {CC} and (rs9315069 = {TC,CC} and rs913071 = {TT,TC})	rs7321815 chr13: g.101701427C>A	0.095 (0.211)	rs7321815 and rs9315069	1.403 (2.5 × 10 ⁻⁹)	1.774 (95% CI [1.283, 2.265]; P = 2.9 × 10 ⁻¹² ; 20)	2.115	1.414 (P = 0.046; Methotrexate)
		rs9315069 chr13: g.31371544T>C	0.767 (2.4 × 10 ⁻⁴)	rs7321815 and rs913071	0.135 (0.076)			
		rs913071 chr13: g.36553105C>T	0.051 (0.666)	rs9315069 and rs913071	0.958 (2.0 × 10 ⁻⁵)			
14	rs887890 chr14: g.75699438T>G	rs887890 chr14: g.75699438T>G	0.199 (0.014)	rs887890 or rs7142110	0.340 (1.1 × 10 ⁻⁵)		2.144	0.568 (P = 0.008; Cranial radiation)

Table 2 (continued)

Chr	Regulatory 3-SNP interaction ^a	Single-SNP rsID (hg19 HGVS identifier)	Marginal effect ^b SNP β (P)	Component 2-SNP interaction	Pair effect ^b pair β (P)	Discovery cohort SNP interaction β^c (95% CI; P-value; tree frequency)	Permutation-based evaluation statistic ^d (P-value; modified treatment effect)	Replication cohort replicated β_{int} (P-value; modified treatment effect)
	(rs887890 = {TG,GG} or rs7142110 = {AA}) or rs1884632 = {GG}							
		rs7142110 <i>chr14:g.51808403G>A</i>	0.191 (0.016)	rs887890 or rs1884632	0.259 (0.001)	0.498 (95% CI [0.352, 0.644]; P = 3.5×10^{-11} ; 604)		
		rs1884632 <i>chr14:g.69418173C>G</i>	0.198 (0.013)	rs7142110 or rs1884632	0.303 (5.5×10^{-5})			

^aShows the single best original or neighborhood interaction detected for each replicated 3-SNP interaction (based on interaction term replication P-values)

^bThe marginal SNP and SNP pair linear regression models include the same set of adjustment covariates as the main SNP and SNP pair genotype classes are consistent with genetic effect codings observed in 3-SNP interactions

^cEstimated mean changes in BMD Z-scores and 95% confidence intervals for 3-SNP interactions in the discovery cohort of ALL survivors, conditioned on previously identified interactions for a given autosome (N = 835; participants with missing treatment values were excluded from discovery analysis)

^dPermutation-based evaluation statistics > 2 were interpreted as explaining an unusually large proportion of BMD Z-score variation

^eAny 3-SNP interaction that modified treatment effects (P < 0.05) in the replication cohort was considered to be replicated. Methotrexate modification effects were detected among non-ALL diagnosis groups exposed to methotrexate (N = 804). CRT modification effects were detected among non-ALL diagnosis groups exposed to CRT (N = 1195)

glucocorticoids during childhood. Our linear regression model with sex, ancestry, and treatment covariates demonstrated that decreases in adjusted mean BMD Z-scores were significantly associated with increasing cumulative dosages for each of these treatments (Supplementary Table 2).

Using the proposed logic regression-based algorithm, we identified 220 3-SNP interactions (10 interactions per chromosome) associated with BMD Z-score. Consistent with previous observations of regulatory complexes involving enhancer–promoter, enhancer–enhancer, or promoter–promoter interactions [26], no restrictions were made on the composition of 3-SNP interactions. Six distinct (uncorrelated) 3-SNP interactions were selected as candidate interactions for replication follow-up using our permutation-based evaluation statistic threshold (values > 2). We considered each of these six distinct 3-SNP interactions separately as genomic “interaction neighborhoods” associated with BMD and looked for other 3-SNP interactions in these “neighborhoods” that were more strongly associated with BMD than the original 3-SNP interactions in the discovery cohort. All 3-way SNP interactions using any SNP located within 100-kb of regulatory loci contributing to the originally selected 3-SNP interactions were assessed. We identified ten additional “neighborhood” 3-SNP interactions that explained larger proportions of BMD Z-score variation than their corresponding original interactions for four of the six selected 3-SNP interactions: this yielded a total of 16 candidate 3-SNP interactions for replication follow-up.

The replication cohort of SJLIFE participants (N = 1428) with a range of non-ALL pediatric cancer diagnoses (Supplementary Table 3) was comparable to the discovery cohort with respect to age, sex, and ancestry distributions (Table 1). Participants in the replication cohort exposed to either CRT or methotrexate received, on average, higher cumulative doses of these treatments compared to the discovery cohort (Supplementary Table 4). Applying our replication definition, 12 of the 16 3-SNP interactions were replicated as modifiers of treatment effects (Supplementary Tables 5–7). Considering the six originally selected 3-SNP interactions, each reflecting a distinct interaction neighborhood, at least one original or neighborhood 3-SNP interaction candidate was replicated for five of the six selected 3-SNP interaction neighborhoods.

Table 2 shows the best replicated original or neighborhood 3-SNP interaction (defined by replication P-value) detected among the five genomic neighborhoods with replicated interactions. Adjusted changes in mean BMD Z-scores for these five best replicated 3-SNP interactions in the discovery cohort ranged from –1.30 to +1.77 SD, with regression coefficient *t* test-based (naive) P-values ranging from 2.9×10^{-13} to 3.5×10^{-11} . Four of these 3-SNP

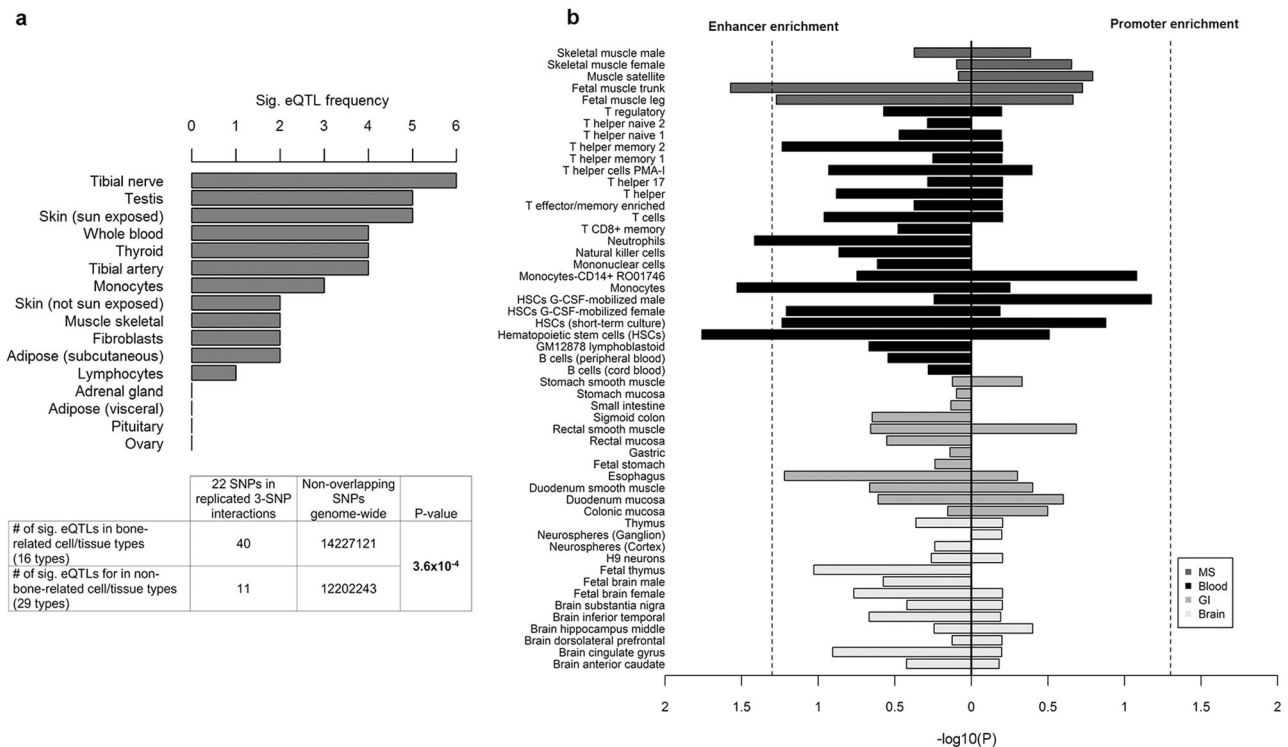


Fig. 1 Biological plausibility of association between identified SNPs in replicated SNP interactions with BMD **a** Counts of significant gene expressions (eQTLs) for the 22 unique SNPs in replicated 3-SNP interactions, grouped by the 16 cell or tissue types related to bone (above); the corresponding enrichment analysis result using ~2.6 million non-overlapping genome-wide SNPs with ~26.4 million eQTLs for comparison (below). **b** Plot of Fisher's exact test P -values

($\log_{10}(P)$) from enhancer (left) and promoter (right) enrichment state analyses for the 22 unique SNPs in replicated 3-SNP interactions, using Roadmap Epigenomics Mapping Consortium chromatin state annotations for two BMD-related human cell categories (musculoskeletal (MS), blood) and two comparison categories (gastrointestinal (GI), brain). Dashed lines correspond to $P < 0.05$

interactions included at least one SNP that was not nominally significant. No component SNP pair fully recovered the entire magnitude of association of its respective 3-SNP interaction. In the discovery cohort, the breakdown of the proportions of variance in BMD Z-score explained by the non-genetic covariates (14.5%) and the five best replicated 3-SNP interactions (14.1%) were comparable (Supplementary Table 8).

To compare the performance of our proposed algorithm to a benchmark SNP interaction association testing method, we conducted an exhaustive, within-chromosome pairwise SNP interaction analysis using the 75523 SNPs mapped to putative regulatory regions. Of the nearly 158 million SNP pair combinations considered, seven pairs achieved genome-wide significance (Bonferroni-adjusted $P < 3.2 \times 10^{-10}$). None were contributing pairs to any of the 220 3-SNP interactions detected with our search algorithm. Considering all SNP pair results with $P < 1.0 \times 10^{-9}$ and the SNP pairs formed by their LD proxy SNPs, none of the 967 original or LD proxy SNP pairs were contributing pairs for any of the 220 3-SNP interactions (Supplementary Table 9). To further distinguish differences in performance between our novel method and the benchmark SNP pair testing

method, we conducted a simulation study. Assuming effect sizes observed in our discovery analysis (Table 2), our proposed method has 18–60% power and 17–49% PPV to detect “true” (replicated) 3-SNP interactions in smaller samples ($N = 1000$), with marked improvements in both statistics with modest increments in sample size (Supplementary Table 10). In comparison, the benchmark SNP pair method is appreciably less powerful and has low PPV for detecting component 2-SNP interactions in underlying true 3-SNP interactions, even with larger sample sizes and under a liberal P -value threshold ($P < 1 \times 10^{-5}$) to select top SNP pairs (Supplementary Table 11).

The overall biological plausibility of association with BMD was assessed for the set of 22 unique SNPs contributing to the 12 replicated original and neighborhood interactions. First, we examined gene expression data, specifically eQTL associations achieving study-wide significance in GTEx Project [21] and GHS-Express [22] databases. Our 22-SNP set had a total of 51 significant eQTLs in 17 cells/tissues, of which 40 were observed among 16 cell/tissue types related to bone (enrichment $P = 3.6 \times 10^{-4}$, relative to the set of non-overlapping SNPs genome-wide with at least one significant eQTL in any of

Table 3 Annotations of replicated regulatory 3-SNP interaction trees

Tree	dbSNP ID	Allele freq (alt) ^a	Chr position, hg19 (locus)	SNP context	Closest genes (±50 kb)	Significant eQTLs ^b	Enhancer in 127 cells/ tissues ^c (bone ratio)	Promoter in 127 cells/ tissues ^c (bone ratio)	Osteoblasts (E129) ^d	Monocytes (E029) ^d	Lymphoblastoid cells (E116) ^d	Fetal muscle trunk (E089) ^d
1	rs901466: C>G	0.26 (G)	114724122 (2q14.1)	Intergenic	<i>LINC01191</i> ; <i>ACTR3</i>	<i>ACTR3</i> ⁴ , <i>SLC35F5</i> ⁴ , <i>RPL23AP7</i> ^{1,4} , <i>DDX11L2</i> ^{2,4} , <i>AC024704.2</i> ^{2,4} , <i>AC010982.1</i> ^{2,4} , <i>AC104653.1</i> ⁴	30 (2:3)	0 (NA)	Enhancer	Enhancer	Weak transcription	Weak Transcription
	rs7569568: G>A	0.16 (A)	225264598 (2q36.2)	Intron	<i>FAM124B</i>	<i>FAM124B</i> ¹	6 (1:5)	20 (7:13)	Weak repressed PC	Enhancer	Weak repressed PC	Active TSS
	rs921319: T>C	0.37 (C)	62367720 (2p15)	Intergenic	<i>COMMD1</i>	<i>FAM161A</i> ⁴	13 (3:10)	2 (1:1)	Quiescent/low	Quiescent/low	Flanking active TSS	Weak Transcription
2	rs1020745: G>A	0.04 (G)	53692955 (12q13.13)	Intron (<i>PFDN5</i>); promoter (<i>C12orf10</i>)	<i>PFDN5</i> ; <i>C12orf10</i> ; <i>MFSD5</i> ; <i>ESPL1</i> ; <i>AAAS</i> ; <i>SP7</i>	<i>AAAS</i> ³ , <i>C12orf10</i> ⁴	56 (23:33)	7 (5:2)	Flanking active TSS	Flanking active TSS	Genic enhancers	Genic Enhancers
	rs2110167: A>G	0.32 (A)	5734319 (12p13.31)	Intron	<i>ANO2</i>	NA	6 (1:1)	0 (NA)	Quiescent/low	Quiescent/low	Quiescent/low	Enhancer
	rs10444471: G>T	0.05 (T)	4677211 (12p13.32)	Coding (<i>DYRK4</i>)	<i>DYRK4</i> ; <i>AKAP3</i> ; <i>RAD51AP1</i> ; <i>C12orf4</i>	NA	13 (7:6)	2 (1:1)	Quiescent/low	Enhancer	Enhancer	Weak Repressed PC
3	rs1894331: G>T	0.44 (G)	11930889 (12p13.2)	Intron	<i>ETV6</i>	NA	49 (19:30)	0 (NA)	Enhancer	Genic enhancers	Genic enhancers	Enhancer
	rs10773093: T>C	0.43 (C)	125046036 (12q24.31)	Intron	<i>NCOR2</i>	NA	68 (19:49)	1 (1:0)	Weak transcription	Enhancer	Flanking active TSS	Enhancer
	rs4768783: C>T	0.43 (C)	47592945 (12q13.11)	Intron (<i>PCED1B</i>)	<i>PCED1B</i> ; <i>PCED1B-AS1</i>	<i>RPL1-493LL12.4</i> ⁴	7 (1:0)	0 (NA)	Quiescent/low	Weak transcription	Quiescent/low	Quiescent/Low
4	rs7321815: C>A	0.35 (A)	101701427 (13q33.1)	Intron (<i>NALCN-AS1</i>)	<i>NALCN-AS1</i> ; <i>NALCN</i>	NA	3 (0:1)	0 (NA)	Quiescent/low	Heterochr.	Quiescent/low	Quiescent/Low
	rs9315069: T>C	0.03 (C)	31371544 (13q12.3)	Intergenic	<i>ALOX5AP</i> ; <i>LINC00398</i>	NA	38 (5:14)	11 (2:9)	Flanking active TSS	Weak repressed PC	Enhancer	Enhancer
	rs913071: C>T	0.36 (C)	36553105 (13q13.3)	Intron (<i>DCLK1</i>)	<i>DCLK1</i> ; <i>MIRS48F5</i>	<i>MAB2L1</i> ³	36 (1:5)	7 (5:2)	Flanking active TSS	Quiescent/low	Quiescent/low	Weak Transcription

Table 3 (continued)

Tree dbSNP ID	Allele freq (alt) ^a	Chr position, hg19 (locus)	SNP context	Closest genes (± 50 kb)	Significant eQTLs ^b	Enhancer in 127 cells/tissues ^c (bone ratio)	Promoter in 127 cells/tissues ^c (bone ratio)	Osteoblasts (E129) ^d	Monocytes (E029) ^d	Lymphoblastoid cells (E116) ^d	Fetal muscle trunk (E089) ^d
5 rs887890: T>G	0.18 (G)	75699438 (14q24.3)	Intergenic	<i>FOS</i>	<i>EIF2B3</i> , <i>RP11-293M10.1</i> ⁴	28 (17:11)	0 (NA)	Weak transcription	Enhancer	Enhancer	Weak Transcription
rs7142110: G>A	0.43 (G)	51808403 (14q22.1)	Intron	<i>LINC00640</i>	NA	16 (1:7)	1 (0:1)	Quiescent/low	Quiescent/low	Quiescent/low	Quiescent/Low
rs1884632: C>G	0.41 (C)	69418173 (14q24.1)	Intron (<i>ACTN1</i>)	<i>ACTN1</i> ; <i>ACTN1-AS1</i>	<i>ACTN1</i> ⁴	83 (27:56)	11 (3:8)	Genic enhancers	Enhancer	Weak transcription	Enhancer

PC PolyComb, *Heterochir*: heterochromatin, TSS transcription start site

^aAllele frequency for the alternative/minor allele in the discovery cohort of ALL survivors ($N = 856$); sample alternative/minor allele designations were used for binary SNP variable codings

^bSignificant eQTLs in bone-related cell/tissue groups for a given SNP, obtained from GTEx Portal and GHS-Express databases. Transcript superscripts reflect cell/tissue groups: (1) Monocytes, (2) whole blood, (3) muscle skeletal, (4) other endocrinological tissue/pathway

^cFrequency of overlap between SNP and putative enhancer or promoter state in 127 REMC-annotated epigenomes. "Bone ratio" = (frequency of overlap with enhancer/promoter state in a bone-related cell/tissue): (frequency of overlap with enhancer/promoter state in a cell/tissue not related to bone)

^dChromatin state annotation for SNPs in the specified cell/tissue type (E*** = epigenome ID), obtained from REMC

the 45 queried cell/tissue types) (Fig. 1a; Supplementary Tables 12, 13). Second, we used REMC chromatin state annotation data [23] to examine whether our 22-SNP set was enriched in enhancer or promoter states in each of 127 consolidated cell/tissue groups. We observed suggestive enrichment in overlap between SNPs in our 22-SNP set and putative enhancer states in four cell types relevant to bone biology ($P < 0.05$, no Bonferroni adjustment), relative to a background set of 75508 non-overlapping enhancer/promoter SNPs in our original SNP restriction set (Fig. 1b). Consideration of weakly significant enhancer and promoter enrichment analysis results ($P < 0.10$, no Bonferroni adjustment; Supplementary Tables 14, 15) suggests the 22-SNP set is relatively enriched for both regulatory states in monocytes and hematopoietic stem cells, which are related to bone metabolism [7]. For each of the distinct replicated 3-SNP interactions, chromatin contacts between putative regulatory regions containing the three SNPs of interest appeared supported: at least two chromatin contacts connecting the three target loci were observed, each with proximity scores ≥ 2 (Supplementary Table 16; Supplementary Figures 17–20).

The 3-SNP interaction with the strongest evidence of association with BMD was observed between rs1020745 (hg19 chr12:g.53692955G >A; *PFDN5* intronic and *C12orf10* promoter region), rs2110167 (hg19 chr12:g.5734319A >G; *ANO2*, intronic region), and rs10444471 (hg19 chr12:g.4677211G >T; *DYRK4* synonymous coding variant) with an adjusted mean increase in BMD Z-score of 1.72 SD (95% CI: 1.27, 2.17). Both rs10444471 and rs2110167 were more frequently observed in enhancer states in bone-related cell types, whereas rs1020745 overlapped both enhancer and promoter states with relatively high frequencies (Table 3). Hi-C chromatin interaction maps in lymphoblastoid cells connecting the 3-SNP regions showed contact selectivity for the rs1020745 locus, with proximity scores indicating nearly 13-fold interaction enrichment with the rs10444471 locus, and over sixfold interaction enrichment with the rs2110167 locus. Enhancer regions including rs10444471 and rs2110167 may interact distally with a promoter or enhancer region bearing rs1020745, in cell types known to play a role in osteoblast or osteoclast differentiation (Fig. 2). Notably, the rs1020745 locus is known to reside in a region of high linkage disequilibrium [27], implicating several potential gene targets including *SP7*.

Discussion

Previous studies of epistasis have successfully used exhaustive testing methods to assess SNP pair interactions. To detect 3-SNP interactions associated with a complex trait

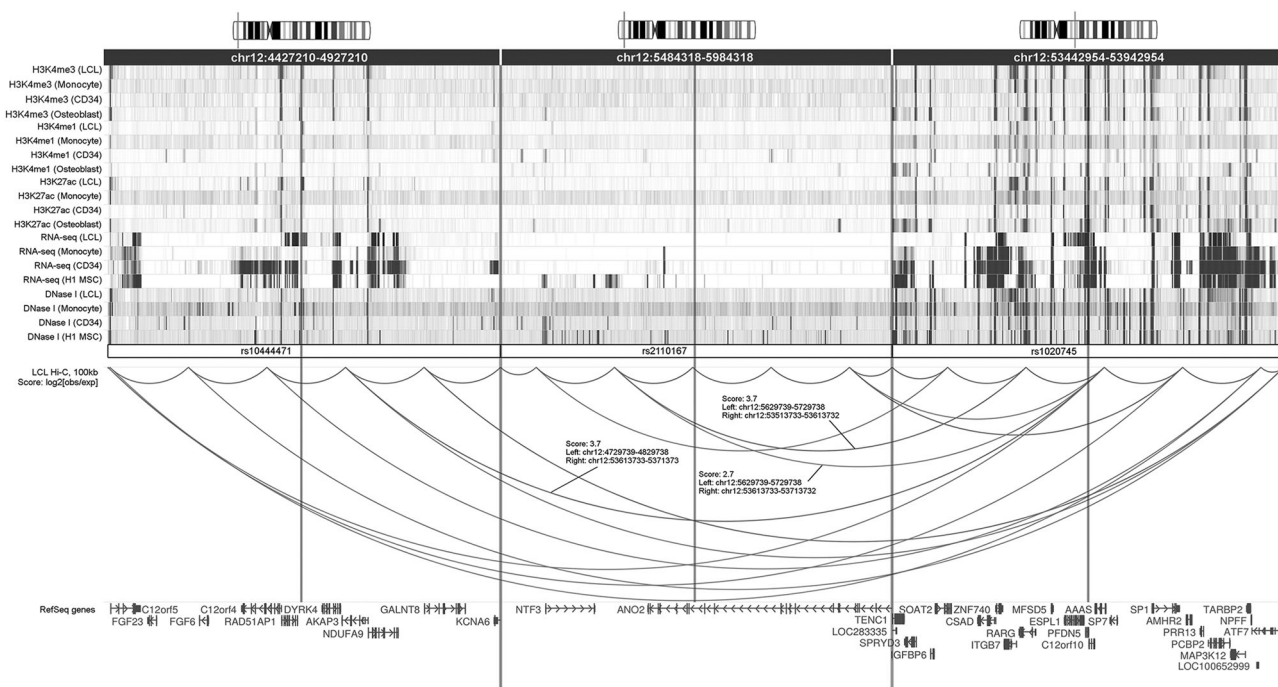


Fig. 2 Chromatin interactions for the chromosome 12 SNP interaction: (rs1020745 = {AG,GG} and (rs2110167 = {GA,AA} and rs10444471 = {GG})). The WashU EpiGenome Browser was used to visualize long-range chromatin interactions within and across three 500-kb windows centered at implicated SNPs. SNP locations are contextualized using ideograms at the top of regional windows and highlighted with vertical lines in the center of each window. Histone modification (H3K4me3, H3K4me1, H3K27ac), RNA-seq, and DNase I hypersensitivity heatmap data tracks were reviewed. Four data tracks per assay for each of four cell/tissue samples are shown:

lymphoblastoid cells (LCLs), peripheral blood mononuclear cells or monocytes, mobilized CD34 cells, and osteoblasts or an osteoblastic precursor proxy (H1 mesenchymal cells). Hi-C data generated with GM06990 LCLs was used to assess evidence for long-range chromatin interactions between SNPs in 3-SNP interaction trees (100-kb bin resolution, $\log_2[\text{observed contact/expected contact}]$ scores). Minimum Hi-C interaction scores were set such that interaction arcs represent chromatin interactions with at least +4-fold observed contact frequency over expected (scores >2)

on a genome-wide scale, we implemented a novel, non-exhaustive logic regression-based algorithm among SNPs mapped to regulatory genomic regions. Specifically, our algorithm: (a) focuses on 3-way interactions that plausibly reflect gene regulation events using SNPs mapped to enhancers or promoters; and (b) considers many epistatic candidates, but only allocates 1 degree-of-freedom for a 3-SNP interaction. The strength of our method is that we use logic regression combined with a conditioning strategy to encourage a multi-directional, stochastic search, bypassing an exhaustive search for 3-way interactions that may miss a true interaction due to lack of statistical power.

Despite known limitations of logic regression (e.g., non-exhaustive searches may miss the “best” interaction solution), we propose our method as a complementary approach to existing exhaustive 2-SNP search methods to detect higher order epistasis. We observed no overlap between top 2-way regulatory SNP interactions identified using a benchmark exhaustive testing method and 3-way regulatory SNP interactions detected with our proposed method. Furthermore, our simulation results revealed that SNP pair

searches are ineffective for detecting 3-SNP interaction patterns associated with variations in BMD, unless component 2-SNP interactions have strong associations with phenotype without the inclusion of an additional SNP. These results suggest exhaustive searches for 2-SNP interactions are not universally effective for detecting higher order epistasis, and novel methods to conduct deliberate searches for higher order epistasis are needed.

To safeguard against the reporting of false positive results, we used a permutation-based evaluation statistic to identify candidate 3-SNP interactions, performed a replication analysis, and conducted additional bioinformatics analyses. We identified six regulatory 3-SNP interactions that potentially modify treatment effects on BMD among adult survivors of pediatric ALL. Five of these 3-SNP interactions were replicated as treatment modification effects in an independent sample. Our bioinformatics analyses indicated that SNPs contributing to replicated interactions had both an excess of gene expressions and an enrichment of enhancer states in cell and tissue types important for bone biology. The plausibility of interactions

between regulatory regions bearing target SNP variants was supported by observations of chromatin contacts that occurred in greater frequencies than expected between regions that overlapped SNPs of interest in lymphoblastoid cells. Although these 3-SNP interactions were not functionally validated, our findings represent viable leads in identifying epistatic interactions with cancer treatment-related effects on BMD.

There are multiple ways to interpret these 3-SNP interactions. Given the long-range chromatin interaction data, it is plausible that epistatic networks consisting of three SNPs embedded in regulatory regions that physically interact jointly affect gene expressions that modify BMD in pediatric cancer survivors exposed to specific cytotoxic treatments. For example, among those exposed to methotrexate, the genomic regulatory region bearing rs1020745 could act as a “hub” for the 3-way chromosome 12 genetic interaction, with rs2110167 and rs10444471 acting as supportive regulatory elements to influence the *SP7* locus (rs1020745). *SP7* has previously been reported as a candidate gene affecting bone biology in both adult and pediatric populations [27, 28], and is known to encode an osteogenic transcription factor, Osterix (*Osx*) [29].

Although a second independent cohort of ALL survivors would be desirable for replication analyses, the availability of a replication cohort of non-ALL survivors, which consisted predominantly of survivors of solid tumors or lymphoma, provided the opportunity to assess whether genetic interactions associated with BMD Z-score in the discovery cohort plausibly modified cancer treatment effects on BMD. Our replication results support the discovery findings and underscore the relative importance of treatment exposures, as these epistatic interactions do not appear to be pathological artifacts specific to ALL. To contextualize these treatment effect modifications, consider the chromosome 12 interaction. This putative epistatic interaction may modulate the effects of *SP7* and as a consequence, *Osx* expression levels. Exposure to methotrexate has been linked to decreased *Osx* expression and significant reductions in osteocyte precursor cells and metaphyseal trabecular bone volume in rats [30]. As such, this interaction may counter BMD loss in cancer survivors exposed to methotrexate.

In conclusion, our results demonstrate the feasibility of detecting and replicating higher order interactions between SNPs within putative regulatory regions associated with a complex quantitative trait, using a hypothesis-driven approach. Similar searches can be implemented in other contexts, using known biological interaction mechanisms. Although power to assess larger n^{th} -order interactions decreases as the number of participatory SNPs increases, biologically motivated searches for SNP interaction networks involving more than three SNPs at a time are warranted.

Acknowledgements This project was funded and supported by the St. Jude Lifetime Cohort Study (U01 CA195547), American Lebanese Syrian Associated Charities, Rally Foundation for Childhood Cancer Research, National Institutes of Health Grant R01CA216354, and Alberta Machine Intelligence Institute.

Author contributions Conceived and designed SJLIFE study: KKN, SCK, WC, MMH, LLR, CLW. Conceived and designed analytic methodologies and performed the analysis: CI, YY. Managed data: CI, WM, YS, RJB, CLW. Drafted the paper: CI, YY, CLW. Critical revision and final approval of the paper: All authors.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

- Gurney JG, Kaste SC, Liu W, et al. Bone mineral density among long-term survivors of childhood acute lymphoblastic leukemia: results from the St. Jude Lifetime Cohort Study. *Pediatr Blood Cancer*. 2014;61:1270–6.
- Wasilewski-Masker K, Kaste SC, Hudson MM, Esiashvili N, Mattano LA, Meacham LR. Bone mineral density deficits in survivors of childhood cancer: long-term follow-up guidelines and review of the literature. *Pediatrics*. 2008;121:e705–13.
- Kaste SC, Rai SN, Fleming K, et al. Changes in bone mineral density in survivors of childhood acute lymphoblastic leukemia. *Pediatr Blood Cancer*. 2006;46:77–87.
- Kumar J, Swanberg M, McGuigan F, Callreus M, Gerdhem P, Åkesson K. LRP4 association to bone properties and fracture and interaction with genes in the Wnt-and BMP signaling pathways. *Bone*. 2011;49:343–8.
- Yang TL, Guo Y, Li J, et al. Gene-gene interaction between RBMS3 and ZNF516 influences bone mineral density. *J Bone Miner Res*. 2013;28:828–37.
- Seeman E. Pathogenesis of bone fragility in women and men. *Lancet*. 2002;359:1841–50.
- Takayanagi H. Osteoimmunology: shared mechanisms and crosstalk between the immune and bone systems. *Nat Rev Immunol*. 2007;7:292.
- Hemani G, Shakhbazov K, Westra HJ, et al. Detection and replication of epistasis influencing transcription in humans. *Nature*. 2014;508:249–53.
- Dinu I, Mahasirimongkol S, Liu Q, et al. SNP-SNP interactions discovered by logic regression explain Crohn’s disease genetics. *PLoS ONE*. 2012;7:e43035.
- Emily M, Mailund T, Hein J, Schauer L, Schierup MH. Using biological networks to search for interacting loci in genome-wide association studies. *Eur J Hum Genet*. 2009;17:1231–40.
- Wei W-H, Hemani G, Gyenesei A, et al. Genome-wide analysis of epistasis in body mass index using multiple human populations. *Eur J Hum Genet*. 2012;20:857–62.
- Dekker J, Marti-Renom MA, Mirny LA. Exploring the three-dimensional organization of genomes: interpreting chromatin interaction data. *Nat Rev Genet*. 2013;14:390–403.
- Visel A, Rubin EM, Pennacchio LA. Genomic views of distant-acting enhancers. *Nature*. 2009;461:199–205.
- Ernst J, Kheradpour P, Mikkelsen TS, et al. Mapping and analysis of chromatin state dynamics in nine human cell types. *Nature*. 2011;473:43–9.
- Hudson MM, Ness KK, Nolan VG, et al. Prospective medical assessment of adults surviving childhood cancer: study design,

- cohort characteristics, and feasibility of the St. Jude Lifetime Cohort study. *Pediatr Blood Cancer*. 2011;56:825–36.
16. Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. *Genetics*. 2000;155:945–59.
 17. Browning SR, Browning BL. Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. *Am J Hum Genet*. 2007;81:1084–97.
 18. Ruczinski I, Kooperberg C, LeBlanc M. Logic regression. *J Comp Graph Stat*. 2003;12:475–511.
 19. Tusher VG, Tibshirani R, Chu G. Significance analysis of microarrays applied to the ionizing radiation response. *Proc Natl Acad Sci USA*. 2001;98:5116–21.
 20. Wei W-H, Hemani G, Haley CS. Detecting epistasis in human complex traits. *Nat Rev Genet*. 2014;15:722–33.
 21. GTEx Consortium. Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science*. 2015;348:648–60.
 22. Zeller T, Wild P, Szymczak S, et al. Genetics and beyond—the transcriptome of human monocytes and disease susceptibility. *PLoS ONE*. 2010;5:e10693.
 23. Kundaje A, Meuleman W, Ernst J, et al. Integrative analysis of 111 reference human epigenomes. *Nature*. 2015;518:317–30.
 24. Lieberman-Aiden E, Van Berkum NL, Williams L, et al. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science*. 2009;326:289–93.
 25. Zhou X, Lowdon RF, Li D, et al. Exploring long-range genome interactions using the WashU Epigenome Browser. *Nat Methods*. 2013;10:375–6.
 26. ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome. *Nature*. 2012;489:57–74.
 27. Timpson NJ, Tobias JH, Richards JB, et al. Common variants in the region around Osterix are associated with bone mineral density and growth in childhood. *Hum Mol Genet*. 2009;18:1510–7.
 28. Estrada K, Styrkarsdottir U, Evangelou E, et al. Genome-wide meta-analysis identifies 56 bone mineral density loci and reveals 14 loci associated with risk of fracture. *Nat Genet*. 2012;44:491–501.
 29. Nakashima K, Zhou X, Kunkel G, et al. The novel zinc finger-containing transcription factor osterix is required for osteoblast differentiation and bone formation. *Cell*. 2002;108:17–29.
 30. Georgiou KR, Scherer MA, Fan CM, et al. Methotrexate chemotherapy reduces osteogenesis but increases adipogenic potential in the bone marrow. *J Cell Physiol*. 2012;227:909–18.