

RESEARCH ARTICLE

Open Access

# A comprehensive candidate gene approach identifies genetic variation associated with osteosarcoma

Lisa Mirabello<sup>1\*</sup>, Kai Yu<sup>1</sup>, Sonja I Berndt<sup>1</sup>, Laurie Burdett<sup>2</sup>, Zhaoming Wang<sup>2</sup>, Salma Chowdhury<sup>2</sup>, Kedest Teshome<sup>2</sup>, Arinze Uzoka<sup>2</sup>, Amy Hutchinson<sup>2</sup>, Tom Grotmol<sup>3</sup>, Chester Douglass<sup>4</sup>, Richard B Hayes<sup>5</sup>, Robert N Hoover<sup>1</sup>, Sharon A Savage<sup>1</sup> and for the National Osteosarcoma Etiology Study Group

## Abstract

**Background:** Osteosarcoma (OS) is a bone malignancy which occurs primarily in adolescents. Since it occurs during a period of rapid growth, genes important in bone formation and growth are plausible modifiers of risk. Genes involved in DNA repair and ribosomal function may contribute to OS pathogenesis, because they maintain the integrity of critical cellular processes. We evaluated these hypotheses in an OS association study of genes from growth/hormone, bone formation, DNA repair, and ribosomal pathways.

**Methods:** We evaluated 4836 tag-SNPs across 255 candidate genes in 96 OS cases and 1426 controls. Logistic regression models were used to estimate the odds ratios (OR) and 95% confidence intervals (CI).

**Results:** Twelve SNPs in growth or DNA repair genes were significantly associated with OS after Bonferroni correction. Four SNPs in the DNA repair gene *FANCM* (ORs 1.9-2.0,  $P = 0.003-0.004$ ) and 2 SNPs downstream of the growth hormone gene *GH1* (OR 1.6,  $P = 0.002$ ; OR 0.5,  $P = 0.0009$ ) were significantly associated with OS. One SNP in the region of each of the following genes was significant: *MDM2*, *MPG*, *FGF2*, *FGFR3*, *GNRH2*, and *IGF1*.

**Conclusions:** Our results suggest that several SNPs in biologically plausible pathways are associated with OS. Larger studies are required to confirm our findings.

## Background

Osteosarcoma (OS) is the most common primary malignant bone tumor and typically occurs in adolescents and young adults (Damron *et al.*, 2007; Mascarenhas L, 2006; Stiller CA, 2006). OS incidence has a bimodal age distribution; the primary peak occurs during adolescence and a second, much smaller peak is present in the elderly [1,2]. In young patients, OS incidence correlates with puberty and bone growth. The peak incidence of both OS and puberty tend to occur earlier in females. OS incidence is higher in males, who usually grow taller than females, and it typically occurs at sites of rapid bone growth (*e.g.*, the metaphyses of long bones) [1]. The incidence peak in adolescence is followed by a

rapid decline and a plateau when bone growth is complete (after age 24 years) [3]. Several studies have suggested that being taller than average at diagnosis is associated with increased OS risk [4-11]. A recent meta-analysis of height at diagnosis and birth-weight as OS risk factors found that high birth-weight (OR 1.35, 95% CI 1.01-1.79, compared to average birth-weight subjects) and being taller than average were significant OS risk factors (for those  $\geq 90^{\text{th}}$  percentile of height: OR 2.63, 95% CI 1.98-3.49, compared to those  $\leq 50^{\text{th}}$  percentile of height) [12]. In aggregate, these data suggest that growth and development during puberty, and possibly *in utero*, contributes to OS etiology.

OS frequently occurs in several cancer predisposition syndromes, including Li-Fraumeni Syndrome [13], hereditary bilateral retinoblastoma [14], Bloom, Werner, Rothmund Thomson syndromes [15], and Diamond-Blackfan anemia [16]. It also occurs more frequently in

\* Correspondence: mirabellol@mail.nih.gov

<sup>1</sup>Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Rockville, MD 20852, USA

Full list of author information is available at the end of the article

individuals with Paget's disease [17]. However, in the majority of OS cases, there is no known predisposing factor. The data regarding the role of common germline genetic variation in OS risk are sparse. Positive or suggestive associations have been observed for SNPs in the vitamin D receptor (*VDR*; *FokI* polymorphism) [8], tumor necrosis factor- $\alpha$  (*TNF- $\alpha$* ; promoter region -238 SNP) [18], insulin-like growth factor 2 receptor (*IGF2R*) [19], *Fas* [20], transforming growth factor- $\beta$  receptor 1 (*TGFBRI*) [21], and *MDM2* [22] genes, but no associations were observed for the estrogen receptor (*ER*) [8], collagen I[ $\alpha$ ]1 (*COL1A1*) [8], or *TP53* [23] genes.

Peak levels of endogenous sex hormones, growth hormones, and IGF-I levels occur during puberty which also corresponds to peak bone growth rates. It is possible that variation in genes important in bone development, growth, and puberty are modifiers of OS risk. In addition, insulin-like growth factors are known to play critical roles in carcinogenesis [24,25]. Chromosomal aneuploidy in OS cells [26,27] and the increased OS risk observed with genetic syndromes caused by mutations in DNA repair pathways [e.g., *TP53* [13], *WRN*, *BLM*, *RECQL4* [15]] suggests that variants in DNA repair genes may be associated with OS risk. Genes in DNA repair and tumor suppressor pathways may also contribute to OS pathogenesis, because they help maintain the integrity of critical cellular processes and defects in these genes often lead to carcinogenesis. Diamond-Blackfan anemia is associated with an increased frequency of OS and mutations in ribosomal genes (i.e., *RPS19*, *RPS24* and *RPS17*) [16,28]. Thus, it is also feasible that variation in these genes may contribute to OS risk.

There are numerous genes that contribute to bone growth and puberty, and DNA repair which could contribute to OS which have not yet been evaluated. We evaluated these hypotheses in an OS association study of candidate genes from the following pathways: growth and hormone metabolism, bone formation, tumor suppressor and DNA repair, and ribosomal. We genotyped 4836 tag-SNPs across 255 candidate genes from these four pathways in 96 OS cases and 1426 cancer-free controls. This approach identified several SNPs in candidate genes from biologically plausible pathways that were associated with OS risk.

## Methods

### Study design and population

OS cases ( $n = 101$ ) were derived from the hospital-based, prospective case-control study, the Bone Disease and Injury Study of Osteosarcoma (BDISO) [29]. Blood samples and questionnaire data on individuals were collected at orthopedic surgery departments in 10 United States medical centers between 1994 and 2000. OS

patients were identified at the time of limb salvage surgery. There were no identified cases of Paget's disease of the bone in this study. Orthopedic controls from BDISO ( $n = 65$ ) were individuals with benign tumors (26%) and other non-neoplastic conditions, such as inflammatory diseases, cysts, and trauma, excluding those with hip fracture or osteoporosis. All individuals were self-described Caucasians. Institutional review boards at each of the medical centers approved the protocol and informed consent was obtained from all study subjects.

An additional 1364 cancer-free control subjects were derived from the Prostate, Lung, Colorectal, and Ovarian (PLCO) Cancer Screening Trial. Individuals aged 55-74 years were enrolled in the screening trial between 1993 and 2001 from 10 different centers in the U.S. All subjects in this study were required to have completed a baseline questionnaire, provided a blood specimen, and consented to participate in etiologic studies of cancer and related diseases. Controls were limited to Caucasians living in the continental U.S. without a diagnosis of adenoma or cancer at baseline. The institutional review boards at the National Cancer Institute and 10 screening centers approved the study.

### Genotyping assays

DNA was isolated from blood specimens using standard methods. Genotyping was conducted on a Custom Infinium<sup>®</sup> BeadChip (iSelect)<sup>™</sup> from Illumina, Inc. The iSelect panel was created by investigators in the Division of Cancer Epidemiology and Genetics, National Cancer Institute (NCI) to target genetic variation in genes potentially important in carcinogenesis and cancer risk. Tag SNPs were identified from the HapMap CEU population assuming a  $r^2$  threshold of 0.80 using the Tagzilla module of the GLU software package (<http://code.google.com/p/glu-genetics/>) across 255 candidate genes, including the region 20 kb upstream and 10 kb downstream from the gene. Additional potentially functional SNPs were forced into the tag-SNP selection for select genes. In this study, a total of 6050 tag-SNPs were genotyped.

The concordance rates between 10 duplicate BDISO and PLCO samples on the iSelect panel were 99.5% and 99.9%, respectively. SNPs were excluded if they had less than a 90% genotyping rate in either study population, were non-variable or had a minor allele frequency (MAF) <1%, or if they failed the Hardy-Weinberg equilibrium test or genotyping validation. Individuals were excluded if they had too much missing genotype data (>10% missing genotypes). A total of 4836 tag-SNPs met these quality control criteria and were included for analysis.

After selecting approximately 500 genetically matched controls, based on 28 k iSelect SNPs, from our 1426

controls and comparing these results to the results using all 1426 controls, we determined that there was no significant difference in the results; and thus, we used all 1426 controls in our analyses to maximize our ability to detect associations with rare alleles. A principal component analysis was performed using a set of 3,843 structure inference SNPs selected from the iSelect BeadChip (27,905 SNPs) to evaluate population substructure among the BDISO individuals and the PLCO controls. There was no evidence of significant population stratification. However, 5 cases, 2 orthopedic controls and 1 PLCO control were considered genetic outliers and excluded from the genotyping analyses for a final sample size of: 96 cases, 63 orthopedic controls, and 1363 PLCO controls.

### Statistical analyses

Logistic regression models were used to estimate the odds ratio (OR) and 95% confidence intervals (CI) for the strength of the association between OS risk independently for each SNP, adjusting for gender. To deal with rare variants, the most appropriate method was chosen in the sequence of logistic regression on the additive trend model, logistic regression on the dominant model, or the Fisher's Exact Test. The most common allele or the homozygote of the common allele was used as the referent category for the additive or dominant model, respectively. Bonferroni corrections ( $P_{adj}$ ) were conducted by gene (for all SNPs in a gene) for correction of multiple statistical tests.

Genes were categorized into biological pathways using the Kegg pathway database (<http://www.genome.jp/kegg/pathway.html>) or by function based on literature review. We conducted gene-level and pathway-level analyses based on Yu *et al* [30]. The gene-level analysis is a global test for the association between the outcome and a subset of SNPs within a given gene or region. The pathway-level analysis is a global test for the association between the outcome and any subset of genes within a given pathway.  $P$ -values for these analysis were estimated with 20,000 permutation steps according to the algorithm given in Yu *et al* [30].

Statistical analyses were performed using SAS software, version 9.1 (SAS Institute, Cary, NC), R language, and PLINK software, version 1.06 (<http://pngu.mgh.harvard.edu/purcell/plink/>). We evaluated the correlation between SNPs [linkage disequilibrium (LD)] across specific gene regions with Haploview version 4.1 [31].

### Results

The characteristics of the study participants are shown in Table 1. The median age of the 96 OS cases was 20.3 years (age range 8 to 80.5). The median age of the 63 orthopedic controls from the BDISO was 18.5 years

**Table 1 Characteristics of cases and controls**

	mean age (SD)	n (%) male	n (%) female	total n
OS Cases	26.6 (16.5)	54 (56.3)	42 (43.7)	96
All Controls	60.9 (9.9)	904 (63.4)	522 (36.6)	1426
Orthopedic Controls	24.7 (15.1)	34 (54.0)	29 (46.0)	63
PLCO Controls	62.6 (5.2)	870 (63.8)	493 (36.2)	1363

$n$  = number of individuals;

SD = standard deviation;

PLCO = Prostate, lung, colon, ovarian cancer cohort.

(age range 7.2-68.5), and the PLCO controls were older with a median age of 62 years (age range 55-75). All participants were self-identified Caucasians and from the continental United States. Additional file 1, Table S1 shows the genes analyzed in each pathway, the number of SNPs and most significant SNP in each gene, and the gene- and pathway-level  $P$  values. There were 161 genes (2835 SNPs) in the DNA repair pathway, 62 genes (1448 SNPs) in the growth/hormone metabolism pathway, 28 genes (534 SNPs) in the bone formation pathway, and 4 genes (19 SNPs) in the ribosomal pathway. We took three approaches to the analyses by evaluating associations with OS at the individual SNP level, gene level, and pathway level. We used the conservative Bonferroni correction to correct for multiple statistical tests.

### Individual SNPs associated with osteosarcoma

Of the 4836 SNPs, 241 (expected is 241.8) were statistically significant ( $P < 0.05$ ) before correction for multiple tests (Additional file 2, Table S2). A SNP downstream of *GHI* (growth hormone 1) was the most significantly associated with OS ( $P_{trend} = 0.0009$ ). Twelve SNPs, all in genes from the DNA repair or growth and hormone pathways, were significantly associated with OS after Bonferroni correction ( $P_{adj} < 0.05$ ) for multiple tests by gene (for all SNPs in a gene; Table 2 and Figure 1). There were 4 SNPs in the DNA repair gene *FANCM* (Fanconi anemia, complementation group M) and 2 SNPs in the growth and hormone gene *GHI* that were significantly associated with OS after correction for multiple tests. The 4 significant SNPs in *FANCM* were not highly correlated using our control data ( $D' = 0.93-1.0$  and  $r^2 = 0.01-0.37$ ) and HapMap CEU population data (Additional file 3, Figure S1). There was high LD among the 2 significant SNPs downstream of *GHI* with our control data ( $D' = 0.96$  and  $r^2 = 0.91$ ) and HapMap CEU population data (Additional file 3, Figure S2). In Figure 1, SNPs in *GHI* clearly appear more significant after correction than SNPs in any other genes. Five SNPs in the upstream or downstream region of candidate genes were significant after correction: downstream of *MDM2* [Mdm2 p53 binding protein homolog

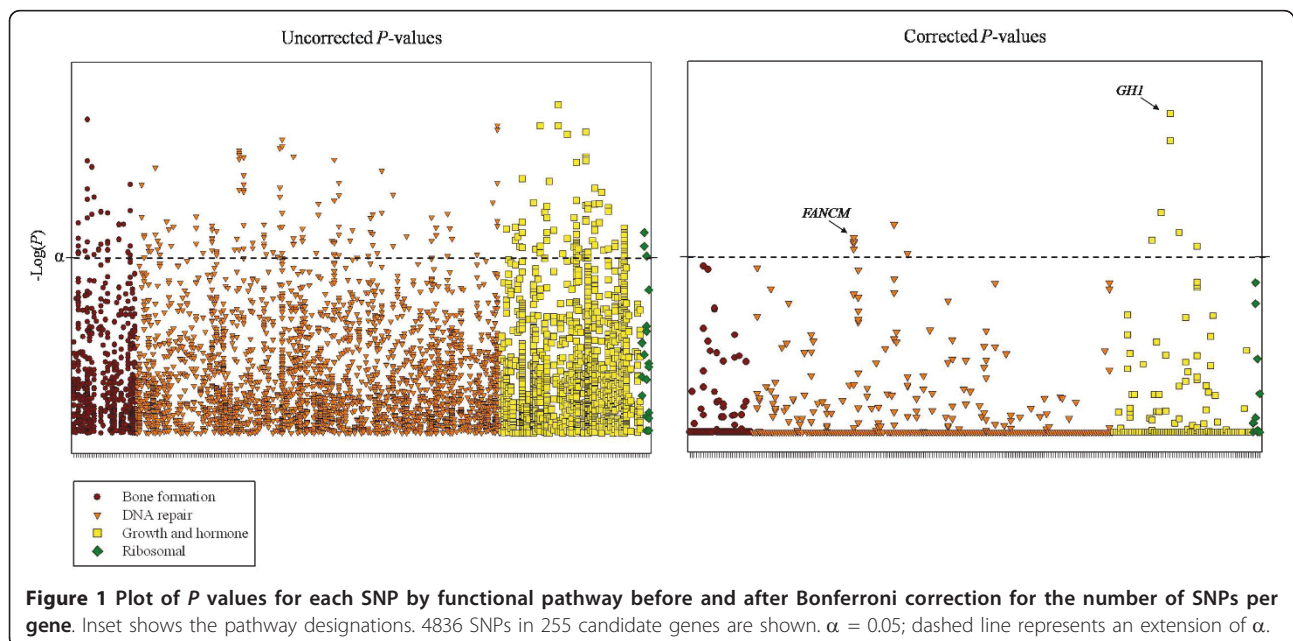
**Table 2 Significant SNPs after Bonferroni correction by gene ( $P_{adj}$ )**

Pathway	Gene	SNP	Genomic position		Minor Allele	MAF (%) Controls	MAF (%) Cases	OR <sup>†</sup>	95% CI	P	$P_{adj}$	Model or Test <sup>§</sup>	Gene P
DNA repair	<i>FANCM</i>	rs1367580	Chr14: 44714339	Ex14+316 (V878L)	T	10.4	16.7	1.97	(1.26-3.08)	0.0031	0.034	Dominant	0.019
		rs11845507	Chr14: 44721715	IVS16-1012	A	10.5	16.7	1.96	(1.25-3.06)	0.0033	0.037	Dominant	
		rs4900664	Chr14: 44732264	IVS20-2861	T	9.3	15.3	2.00	(1.25-3.17)	0.0035	0.038	Dominant	
		rs7141145	Chr14: 44733578	IVS20-1547	A	10.3	16.5	1.95	(1.24-3.07)	0.0039	0.043	Dominant	
	<i>MDM2</i>	rs1690916	Chr12: 67521673	downstream, no gene	A	42.5	31.3	0.62	(0.45-0.85)	0.0029	0.026	Additive	0.016
	<i>MPG</i>	rs216614	Chr16: 60334	upstream, in <i>RHBDF1</i>	A	0.7	3.1	4.80	-	0.0036	0.047	Fisher's	0.027
Growth and hormone	<i>FGF2</i>	rs11737764	Chr4: 124046230	downstream, in <i>NUDT6</i>	T	8.1	14.1	2.12	(1.33-3.39)	0.0016	0.036	Dominant	0.020
	<i>FGFR3</i>	rs6599400	Chr4: 1754823	upstream, no gene	A	33.0	42.6	1.51	(1.12-2.03)	0.0069	0.021	Additive	0.017
	<i>GHI</i>	rs11079515	Chr17: 59359377	downstream, no gene	G	37.9	49.5	1.61	(1.20-2.16)	0.0016	0.005	Additive	0.002
		rs7921	Chr17: 59359991	downstream, in <i>CD79B</i>	A	27.2	16.1	0.52	(0.35-0.76)	0.0009	0.003	Additive	
	<i>GNRH2</i>	rs3761243	Chr20: 2971022	upstream, no gene	C	29.7	40.5	1.60	(1.19-2.16)	0.0020	0.031	Additive	0.021
	<i>IGF1</i>	rs7956547	Chr12: 101382946	IVS2+10605	C	26.3	15.8	0.53	(0.36-0.79)	0.0019	0.040	Additive	0.021

MAF = minor allele frequency;

<sup>†</sup> Odds ratios were estimated using logistic regression models with the most common allele or genotype as the referent, adjusted for gender;

<sup>§</sup> Results are shown for the model or test chosen to deal with rare variants in the sequence of logistic regression on the additive trend model, logistic regression on the dominant model, or the Fisher's Exact Test.



**Figure 1** Plot of  $P$  values for each SNP by functional pathway before and after Bonferroni correction for the number of SNPs per gene. Inset shows the pathway designations. 4836 SNPs in 255 candidate genes are shown.  $\alpha = 0.05$ ; dashed line represents an extension of  $\alpha$ .

(mouse)], and *FGF2* [fibroblast growth factor 2 (basic)], and upstream of *MPG* (N-methylpurine-DNA glycosylase), *FGFR3* (fibroblast growth factor receptor 3), and *GNRH2* (gonadotropin-releasing hormone 2). A SNP in intron 2 (IVS2+10605) of *IGF1* [insulin-like growth factor 1 (somatomedin C)] was significantly associated with a decreased risk of OS after correction.

#### Genes and pathways associated with osteosarcoma

We evaluated 255 candidate genes from four functional pathways (Additional file 1, Table S1). Fourteen genes were significantly associated with OS (Gene *P*-values <0.05; Additional file 1, Table S1 and Additional file 2, Table S2). The most significantly related genes were *GHI* (Gene *P* = 0.002), *MDM2* (Gene *P* = 0.016), *FGFR3* (Gene *P* = 0.017), and *FANCM* (Gene *P* = 0.019). However, if we correct for multiple tests (255 genes), none remain significant. None of the four pathways were significantly associated with OS (Additional file 1, Table S1).

#### Discussion

The biology of OS pathogenesis is complex and there are limited data on risk factors in the more common sporadic form of OS. Epidemiologic studies of OS suggest that growth and development play a role in etiology [2,32,33]. It occurs primarily in adolescents during puberty [1] when bone growth is rapid and endogenous sex hormones and growth hormones are at their highest, so variation in a gene involved in regulating sex hormones is biologically plausible. Rapidly growing tissue, such as bone during puberty, is known to be highly susceptible to carcinogenesis, possibly due to rapidly proliferating osteogenic cells being more vulnerable to DNA repair errors [4,34]. In addition, chromosomal aneuploidy is extensive in somatic OS cells, which suggests the presence of chromosomal instability [26,27]. The increased frequency of OS in genetic predisposition syndromes [13,15] characterized by mutations in DNA repair pathways suggests that variants in genes involved in DNA repair are also reasonable candidates.

We evaluated the association between OS and 255 candidate genes, including 4836 SNPs, from four functional pathways (growth and hormone metabolism, bone formation, DNA repair, and ribosomal). We used 3 approaches to comprehensively evaluate these biologically plausible pathways: analyses were performed at the individual SNP level, gene level and pathway level with conservative statistical corrections for multiple testing. While no genes or pathways were significantly associated with OS after correction for multiple tests, the SNP based approach identified some potentially important candidates. A total of twelve SNPs in genes from the growth and hormone metabolism, and DNA repair

pathways were significantly associated with OS risk after correction for multiple tests. Two genes had multiple significant SNPs associated with risk, *FANCM* and *GHI*, after correction for multiple tests.

*FANCM* contained 4 SNPs significantly associated with a similar 2-fold increased risk of OS using a dominant inheritance model, the most in any gene studied. These SNPs were not correlated in our controls. One SNP is located in exon 14 and the minor or risk allele results in a nonsynonymous change from valine to leucine (Ex14+316, Val878Leu). The minor allele of this SNP is the ancestral allele and is highly conserved among other mammalian species. The three other SNPs in *FANCM* were intronic. *FANCM* has DNA-dependent ATPase activity, promotes the dissociation of DNA triplexes, and with other Fanconi anemia-associated proteins, may repair DNA at stalled replication forks [35,36]. DNA repair must be accurate to preserve genome stability for long-term cellular viability; genetic instability is characteristic of cancer cells, and may be due, at least in part, to mutations or variation in genes that function to ensure DNA integrity [37].

Two significant SNPs were located downstream of *GHI*. They appear to be highly correlated in our controls, although one SNP was associated with an increased risk of OS and the other was protective. The variant in *IGF1* (rs7956547, IVS2+10605) associated with a decreased risk of OS was another interesting candidate involved in growth. The insulin-like growth factor signaling system is important in the formation and homeostasis of bone, and differential expression of IGF1 has been observed in osteosarcomas [38-40]. IGF1 expression is stimulated by growth hormone, and OS incidence peaks during puberty with the release of growth hormone. OS cells have been shown to be IGF1-dependent for growth, and inhibiting growth hormone release in mice decreased IGF1 serum levels and inhibited tumor growth [41-43]. In addition, animal model data from dogs, which develop OS similar to human patients (similar sites, histology and treatment response) and large breeds have a 185-fold increased risk of OS compared with small dog breeds [44,45], suggest that a SNP in *IGF1* is a main determinant of small size in dogs and is virtually absent in giant breeds [46]. The data suggest that *GHI* and *IGF1* may play a role in the etiology of OS.

We also specifically evaluated the genes found to be associated with OS in other studies. We previously identified a SNP in *IGF2R* (rs998075; Ex16+88G >A) to be significantly associated with OS risk in study of variation in genes critical in growth regulation [19]. This SNP was also included in our dataset and the significant association replicated in an analysis limited to our BDISO cases and controls (*P* = 0.01), as expected because it is

the same study population. However, the association did not replicate with the addition of our 1363 PLCO controls ( $P = 0.12$ ), which suggests that the original study may have been limited by the number of controls, or possibly these older PLCO controls have a different ethnic mix within whites related to this polymorphism. Others have found significant associations between OS and polymorphisms in *VDR* [8], *TGFBR1* [21], and *MDM2* [22], which were also included in our dataset. We found no significant associations with individual SNPs within *VDR* or *TGFBR1* before or after correction for multiple tests, although our dataset did not include the *VDR* *FokI* polymorphism [8] or *TGFBR1*\*6A variant [21], or at the gene level. One SNP downstream of the DNA repair gene *MDM2*, rs1690916, was significantly associated with a decreased risk of OS after correction (OR 0.62, 95% CI 0.45-0.85,  $P_{\text{adj}} = 0.026$ ), and 3 intronic SNPs were significantly associated with an increased risk of OS before correction (ORs 1.6-1.8,  $P = 0.008-0.02$ ). The previously identified [22]*MDM2* T309G (rs2279744) polymorphism was marginally non-significantly associated with an increased risk of OS (OR 1.31, 95% CI 0.97-1.77,  $P = 0.07$ ). At the gene level, *MDM2* was found to be significantly associated with OS (Gene  $P = 0.016$ ), but not after correction for multiple tests. As others have shown [8,23], the current study did not confirm associations between polymorphisms in *ESR1/2* (including previously analyzed *PvuII* and *XbaI* polymorphisms), *COL1A1*, or *TP53* and OS after correction, or at the gene level.

We also investigated many of the genes with germline mutations in cancer predisposition syndromes associated with an increased frequency of OS [15], including the DNA repair genes *TP53* (mutated in Li-Fraumeni Syndrome), *WRN* (mutated in Werner syndrome), *BLM* (mutated in Bloom syndrome), *RECQL4* (mutated in Rothmund-Thomson Syndrome), and the ribosomal genes *RPS19* and *RPS24* (mutated in Diamond-Blackfan anemia) [28,47,48]. OS also occurs more commonly in older individuals with Paget's disease [17]. Two genes involved in bone formation, *SQSTM1* and *TNFRSF11A*, have been shown to cause Paget's disease [49]. We found that no common SNPs within these genes that predispose to cancer predisposition syndromes or Paget's disease were significantly associated with OS after correction for multiple tests.

A limitation of the current study was the small number of cases; however, the inclusion of 14 controls per case improved the statistical power to detect SNPs with strong effects. For the additive model and our 96 cases and 1426 controls, we had greater than 80% statistical power to detect an OR of 1.82 for MAFs of 0.1 and an OR of 2.15 for MAFs of 0.05 (with a baseline population

risk of 0.0000001 and type 1 error of 0.05). Another potential limitation of our study was the use of controls from PLCO that were older than the BDISO case and control population. However, the PLCO controls had no history of any cancer, including osteosarcoma, they were limited to Caucasians from the continental US, as were the BDISO cases and controls, and we found no evidence of population stratification in between groups.

Strengths of the current study include the detailed genotyping of biologically plausible pathways which give a higher a priori likelihood. We used three statistical approaches to comprehensively evaluate associations with OS (at the SNP level, gene level, and pathway level). In addition, we used a stringent Bonferroni correction and conservatively interpreted our results to reduce the probability of a Type 1 error.

## Conclusions

We have conducted the largest study of genetic variation in candidate genes associated with OS to date. We identified the presence of genetic variants in candidate genes from biologically plausible pathways important in growth and hormone metabolism, and DNA repair associated with OS, even with conservative statistical corrections. The strongest candidates are *FANCM*, *GH1* and *IGF1*. The potential functional implications of the variation in these genes are currently unknown. However, since OS occurs during a period of rapid bone growth, genes important in growth, puberty, and DNA repair are biologically plausible contributors to OS pathogenesis because of their function in critical cellular processes. Larger studies of common genetic variation in OS and functional studies of the SNPs identified here are required to confirm the significance of our findings.

## Additional material

**Additional file 1: Table S1.** SNPs from 255 candidate genes by functional pathway. This table shows the genes analyzed in each pathway, the number of SNPs in each gene, the  $P$  value for the most significant SNP in each gene, and the gene- and pathway-level  $P$  values.

**Additional file 2: Table S2.** Significant SNPs ( $P < 0.05$ ) in each pathway associated with osteosarcoma before correction for multiple tests. This table lists the statistics (minor allele, MAF, OR, and  $P$  values) for each of the 241 SNPs significantly ( $P < 0.05$ ) associated with osteosarcoma before correction for multiple tests by gene and pathway.

**Additional file 3: Supplementary Figures S1 and S2.** Figure S1. Linkage disequilibrium across *FANCM* using the HapMap Caucasian (CEU) population data (A), and our control data (B) determined using Haploview. This figure illustrates the linkage disequilibrium across *FANCM* and highlights the significant SNPs associated with osteosarcoma after correction for multiple tests. Figure S2. Linkage disequilibrium across *GH1* using the HapMap Caucasian (CEU) population data (A), and our control data (B) determined using Haploview. This figure illustrates the linkage disequilibrium across *GH1* and highlights the significant SNPs associated with osteosarcoma after correction for multiple tests

### Acknowledgements

We are grateful to the BDISO and PLCO participants for their valuable contributions. The National Osteosarcoma Etiology Study Group is represented by Michael A. Simon (University of Chicago), Marc C. Gebhardt (Massachusetts General Hospital), Mark T. Scarborough (Shands Medical Center, University of Florida), Steven Gitelis (Rush Presbyterian and St. Lukes Medical Center), Jeffrey J. Eckardt (University of California at Los Angeles School of Medicine), James R. Neff (Nebraska Health System), Michael J. Joyce (Cleveland Clinic Foundation), Martin Malawer (Washington Cancer Institute), Michael McGuire, (Creighton University), and H. Clarke Anderson (University of Kansas Medical Center). This work was supported in part by the intramural research program of the National Institutes of Health and the National Cancer Institute. This project has been funded in whole or in part with federal funds from the National Cancer Institute, National Institutes of Health, under Contract No. HHSN261200800001E. The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the U. S. Government.

### Author details

<sup>1</sup>Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Rockville, MD 20852, USA. <sup>2</sup>Core Genotyping Facility, National Cancer Institute, SAIC-Frederick, Inc., Gaithersburg, MD, USA. <sup>3</sup>Cancer Registry of Norway, PO Box 5313 Majorstuen, NO-0304 Oslo, Norway. <sup>4</sup>Harvard School of Dental Medicine, Boston, MA, USA. <sup>5</sup>Division of Epidemiology, Department of Environmental Medicine, New York University, New York, NY, USA.

### Authors' contributions

LM helped with the study design, performed the statistical analyses, and drafted the manuscript. KY conducted the global gene and pathway analyses. SIB and RBH designed the genotyping component of the Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial and provided genotype data for these participants. LB, ZW, SC, KT, AU, and AH conducted all of the genotyping assays. TG participated in manuscript preparation. CD and RNH designed the Bone Disease and Injury Study of Osteosarcoma. SAS designed the study, provided input to the analysis strategy, and participated in manuscript preparation. All authors read and approved the final manuscript.

### Competing interests

The authors declare that they have no competing interests.

Received: 16 December 2010 Accepted: 29 May 2011

Published: 29 May 2011

### References

1. Mirabello L, Troisi R, Savage S: **Osteosarcoma incidence and survival rates from 1973 to 2004: Data from the Surveillance, Epidemiology, and End Results Program.** *Cancer* 2009, **115**(7):1531-1543.
2. Mirabello L, Troisi R, Savage S: **International osteosarcoma incidence patterns in children and adolescents, middle ages and elderly persons.** *Int J Cancer* 2009, **125**(1):229-234.
3. Fraumeni JJ, Boice J: **Bone.** In *Cancer Epidemiology and Prevention*. Edited by: Schottenfeld D FJJ. Philadelphia: W B Saunders Co; 1982.
4. Fraumeni JJ: **Stature and malignant tumors of the bone in childhood and adolescence.** *Cancer* 1967, **20**:967-973.
5. Scranton P, DeCicco F, Totten R, Yunis E: **Prognostic factors in osteosarcoma. A review of 20 years' experience at the University of Pittsburgh Health Center Hospitals.** *Cancer* 1975, **36**:2179-2191.
6. Gelberg K, Fitzgerald E, Hwang S, Dubrow R: **Growth and development and other risk factors for osteosarcoma in children and young adults.** *Int J Epidemiol* 1997, **26**:272-278.
7. Rytting M, Pearson P, Raymond A, Ayala A, Murray J, Yasko A, Johnson M, Jaffe N: **Osteosarcoma in preadolescent patients.** *Clin Orthop Relat Res* 2000, **373**:39-50.
8. Ruza E, Sotillo E, Sierrasesúmaga L, Azcona C, Patiño-García A: **Analysis of polymorphisms of the vitamin D receptor, estrogen receptor, and collagen Ia1 genes and their relationship with height in children with bone cancer.** *J Pediatr Hematol Oncol* 2003, **25**:780-786.

9. Cotterill S, Wright C, Pearce M, Craft A: **Stature of young people with malignant bone tumors.** *Pediatr Blood Cancer* 2004, **42**:59-63.
10. Longhi A, Pasini A, Cicognani A, Baronio F, Pellacani A, Baldini N, Bacci G: **Height as a risk factor for osteosarcoma.** *J Pediatr Hematol Oncol* 2005, **27**:314-318.
11. Goodman M, McMaster J, Drash A, Diamond P, Kappakas G, Scranton PJ: **Metabolic and endocrine alterations in osteosarcoma patients.** *Cancer* 1978, **42**:603-610.
12. Mirabello L, Pfeiffer R, Murphy G, Daw NC, Patiño-García A, Troisi RJ, Hoover RN, Douglass C, Schüz J, Craft AW, et al: **Meta-analysis of height at diagnosis and birth-weight as osteosarcoma risk factors.**, in submission.
13. Varley J: **Germline TP53 mutations and Li-Fraumeni syndrome.** *Hum Mutat* 2003, **21**:313-320.
14. Chauveinc L, Mosseri V, Quintana E, Desjardins L, Schlienger P, Doz F, Dutrillaux B: **Osteosarcoma following retinoblastoma: age at onset and latency period.** *Ophthalmic Genet* 2001, **22**:77-88.
15. Kansara M, Thomas D: **Molecular pathogenesis of osteosarcoma.** *DNA Cell Biol* 2007, **26**:1-18.
16. Lipton J, Federman N, Khabbaze Y, Schwartz C, Hilliard L, Clark J, Vlachos A: **Osteogenic sarcoma associated with Diamond-Blackfan anemia: a report from the Diamond-Blackfan Anemia Registry.** *J Pediatr Hematol Oncol* 2001, **23**:39-44.
17. Hanson MFSM, Merchant A: **Osteosarcoma in Paget's disease of bone.** *Journal of bone and mineral research* 2006, **21**(S2):P58-P63.
18. Patiño-García A, Sotillo-Pieiro E, Modesto C, Sierrases-Maga L: **Analysis of the human tumour necrosis factor-alpha (TNFalpha) gene promoter polymorphisms in children with bone cancer.** *J Med Genet* 2000, **37**:789-792.
19. Savage SA, Woodson K, Walk E, Modi W, Liao J, Douglass C, Hoover RN, Chanock SJ, The National Osteosarcoma Etiology Study Group: **Analysis of genes critical for growth regulation identifies Insulin-like Growth Factor 2 Receptor variations with possible functional significance as risk factors for osteosarcoma.** *Cancer Epidemiol Biomarkers Prev* 2007, **16**:1667-1674.
20. Koshkina N, Kleinerman E, Li G, Zhao C, Wei Q, Sturgis E: **Exploratory analysis of Fas gene polymorphisms in pediatric osteosarcoma patients.** *J Pediatr Hematol Oncol* 2007, **29**:815-821.
21. Hu Y, Pan Y, Li W, Zhang Y, Li J, Ma B: **Association between TGFBR1\*6A and osteosarcoma: A Chinese case-control study.** *BMC Cancer* 2010, **10**(1):169.
22. Toffoli G, Bion P, Russo A, De Mattia E, Cecchin E, Hattinger C, Pasello M, Alberghini M, Ferrari C, Scotlandi K, et al: **Effect of TP53 Arg72Pro and MDM2 SNP309 polymorphisms on the risk of high-grade osteosarcoma development and survival.** *Clin Cancer Res* 2009, **15**:3550-3556.
23. Savage SA, Burdett L, Troisi R, Douglass C, Hoover RN, Chanock SJ: **Germline genetic variation of TP53 in osteosarcoma.** *Pediatr Blood Cancer* 2007, **49**:28-33.
24. Kappel C, Velez-Yanguas M, Hirschfeld S, Helman L: **Human osteosarcoma cell lines are dependent on insulin-like growth factor I for in vitro growth.** *Cancer Res* 1994, **54**:2803-2807.
25. Pollak M, Schernhammer E, Hankinson S: **Insulin-like growth factors and neoplasia.** *Nat Rev Cancer* 2004, **4**:505-518.
26. Sandberg AA, Bridge JA: **Updates on the cytogenetics and molecular genetics of bone and soft tissue tumors: osteosarcoma and related tumors.** *Cancer Genet Cytogenet* 2003, **145**:1-30.
27. Al-Romaih K, Bayani J, Vorobyova J, Karaskova J, Park PC, Zielenska M, Squire JA: **Chromosomal instability in osteosarcoma and its association with centrosome abnormalities.** *Cancer Genet Cytogenet* 2003, **144**:91-99.
28. Shimamura A, Alter B: **Pathophysiology and Management of Inherited Bone Marrow Failure Syndromes.** *Blood Reviews* 2010.
29. Troisi R, Masters M, Joshipura K, Douglass C, Cole B, Hoover R: **Perinatal factors, growth and development, and osteosarcoma risk.** *Br J Cancer* 2006, **95**:1603-1607.
30. Yu K, Li Q, Bergen A, Pfeiffer R, Rosenberg P, Caporaso N, Kraft P, Chatterjee N: **Pathway analysis by adaptive combination of P-values.** *Genet Epidemiol* 2009, **33**:700-709.
31. Barrett J, Fry B, Maller J, Daly M: **Haploview: analysis and visualization of LD and haplotype maps.** *Bioinformatics* 2005, **21**(2):263-265.
32. Glass AGFJ: **Epidemiology of bone cancer in children.** *J Natl Cancer Inst* 1970, **44**:187-199.
33. Miller R: **Contrasting epidemiology of childhood osteosarcoma, Ewing's sarcoma, and rhabdomyosarcoma.** *Nat Cancer Inst Monogr* 1981, **56**:9-15.

34. Ruddon R: **Cancer Biology**. New York: Oxford University Press, Inc., 4 2007.
35. Gari K, Décaillet C, Stasiak A, Stasiak A, Constantinou A: **The Fanconi anemia protein FANCM can promote branch migration of Holliday junctions and replication forks**. *Mol Cell* 2008, **29**:141-148.
36. Meetei AR, Medhurst AL, Ling C, Xue Y, Singh TR, Bier P, Steltenpool J, Stone S, Dokal I, Mathew CG, *et al*: **A human ortholog of archaeal DNA repair protein Hef is defective in Fanconi anemia complementation group M**. *Nat Genet* 2005, **37**:958-963.
37. Hoijmakers J: **DNA damage, aging, and cancer**. *N Engl J Med* 2009, **361**:1475-1485.
38. Burrow S, Andrusis I, Pollak M, Bell R: **Expression of insulin-like growth factor receptor, IGF-1, and IGF-2 in primary and metastatic osteosarcoma**. *J Surg Oncol* 1998, **69**:21-27.
39. McCarthy T, Centrella M: **Local IGF-I expression and bone formation**. *Growth Horm IGF Res* 2001, **11**:213-219.
40. Conover C: **Insulin-like growth factor-binding proteins and bone metabolism**. *Am J Physiol Endocrinol Metab* 2008, **294**:E10-14.
41. Pollak M, Sem A, Richard M, Tetenes E, Bell R: **Inhibition of metastatic behavior of murine osteosarcoma by hypophysectomy**. *J Natl Cancer Inst* 1992, **84**:966-971.
42. Kappel C, Velez-Yanguas M, Hirschfeld S, Helman L: **Human osteosarcoma cell lines are dependent on insulinlike growth factor I for in vitro growth**. *Cancer Res* 1994, **54**:2803-2807.
43. Pinski J, Schally A, Groot K, Halmos G, Szepeshazi K, Zarandi M, *et al*: **Inhibition of growth of human osteosarcomas by antagonists of growth hormone-releasing hormone**. *J Natl Cancer Inst* 1995, **87**:1787-1794.
44. Withrow S, Powers B, Straw R, RM W: **Comparative aspects of osteosarcoma. Dog versus man**. *Clin Orthop Relat Res* 1991, **270**:159-168.
45. Tjalma R: **Canine bone sarcoma: Estimation of relative risk as a function of body size**. *J Natl Cancer Inst* 1966, **36**:1137-1150.
46. Sutter N, Bustamante C, Chase K, Gray M, Zhao K, Zhu L, Padhukasahasram B, Karlins E, Davis S, Jones P, *et al*: **A single IGF1 allele is a major determinant of small size in dogs**. *Science* 2007, **316**:112-115.
47. Strahm B, Malkin D: **Mini Review: Hereditary cancer predisposition in children: Genetic basis and clinical implications**. *Int J Cancer* 2006, **119**(2001-2006).
48. Lindor N, McMaster M, Lindor C, Greene M, National Cancer Institute DoCP, Community Oncology and Prevention Trials Research Group: **Concise Handbook of Familial Cancer Susceptibility Syndromes - Second Edition**. *JNCI Monographs* 2008, **38**:1-93.
49. Daroszewska A, Ralston S: **Genetics of Paget's disease of bone**. *Clin Sci (Lond)* 2005, **109**:257-263.

#### Pre-publication history

The pre-publication history for this paper can be accessed here:  
<http://www.biomedcentral.com/1471-2407/11/209/prepub>

doi:10.1186/1471-2407-11-209

**Cite this article as:** Mirabello *et al.*: A comprehensive candidate gene approach identifies genetic variation associated with osteosarcoma. *BMC Cancer* 2011 **11**:209.

**Submit your next manuscript to BioMed Central  
and take full advantage of:**

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at  
[www.biomedcentral.com/submit](http://www.biomedcentral.com/submit)

