Analysis of Genes Critical for Growth Regulation Identifies *Insulin-like Growth Factor 2 Receptor* Variations with Possible Functional Significance as Risk Factors for Osteosarcoma

Sharon A. Savage,[1,3] Karen Woodson,[2] Elyse Walk,[2] William Modi,[6] Jason Liao,[1] Chester Douglass,[7] Robert N. Hoover,[4] Stephen J. Chanock[1,5] and The National Osteosarcoma Etiology Study Group

[1]Section of Genomic Variation, Pediatric Oncology Branch and [2]Genetics Branch, Center for Cancer Research; [3]Clinical Genetics Branch, [4]Epidemiology and Biostatistics Program, and [5]Core Genotyping Facility, Division of Cancer Epidemiology and Genetics; and [6]Core Genotyping Facility, Science Applications International Corporation-Frederick, Inc., National Cancer Institute, NIH, Bethesda, Maryland; and [7]Harvard School of Dental Medicine, Boston, Massachusetts

Abstract

Background: **Osteosarcoma, the most common malignant primary bone tumor, typically occurs during the adolescent growth spurt. Germ-line genetic variation in genes critical in growth regulation could confer altered risk of osteosarcoma.**
Methods: **Fifty-two common single nucleotide polymorphisms (SNP) in 13 genes were genotyped in a prospective case-control study of osteosarcoma (104 osteosarcoma cases and 74 orthopedic controls). Genotype data analyzed with contingency tables suggested the strongest association with *insulin-like growth factor 2 receptor* (*IGF2R*) SNPs. Additional SNPs were genotyped to capture *IGF2R* common haplotypes and resequencing was done across the *IGF2R* block associated with osteosarcoma risk. Percentage methylation was determined by pyrosequencing of the *IGF2R* variant allele located in a CpG island.**
Results: ***IGF2R* Ex16+88G>A (rs998075) and IVS16+15C>T (rs998074) SNPs were associated with increased risk for osteosarcoma compared with orthopedic controls (haplotype odds ratio, 2.04; 95% confidence interval, 1.29-3.24). Follow-up genotyping showed that *IGF2R* IVS15+213C>T was also associated with increased osteosarcoma risk. Resequence analysis identified two additional SNPs linked to the risk-associated SNPs; linkage disequilibrium was strongest in a 1-kb pair region around them. The Ex16+88G>A SNP is located within a CpG island and alters methylation at that site.**
Conclusion: **This pilot study of germ-line genetic variation in growth pathway genes and osteosarcoma identified a haplotype block in *IGF2R* associated with increased risk of osteosarcoma. The presence of a SNP in this block results in loss of methylation at a CpG island, providing corroborative evidence of a possible functional variant. Our analysis of the *IGF2R* haplotype structure will be applicable to future studies of *IGF2R* and disease risk.** (Cancer Epidemiol Biomarkers Prev 2007;16(8):1667–74)

Introduction

Primary malignancies of the bone, with an annual incidence rate of 8.7 per million children <20 years of age, account for ~6% of childhood cancer (1). Of these, osteosarcoma is the most common with ~400 new cases in the United States per year in individuals <20 years of age (1, 2). The age-specific incidence of osteosarcoma is bimodal, with the first major peak occurring around adolescence and a second smaller peak in adults over age 65 (2, 3). Patients with localized osteosarcoma at presentation have a 60% to 80% long-term survival rate, whereas metastatic disease carries a poorer prognosis (4).

Until recently, studies of genetics and bone cancer risk have focused on familial cancer syndromes in which rare, highly penetrant, germ-line mutations have been identified, such as the Li-Fraumeni syndrome (5). Although the sporadic form of osteosarcoma is more common than the familial type, less is known of the genetics of the former (6). Recent advances in the annotation of common genetic variation, together with increased efficiency in scope and cost for high-throughput genotype analyses, have made it possible to investigate the etiologic contribution of the most common type of variant in the genome, the single nucleotide polymorphism (SNP).

Only a few preliminary studies of genetic variation in primary bone tumors of childhood have been published to date. In osteosarcoma, the Fok1 polymorphism T>C (rs10735810) of the vitamin D receptor gene, which eliminates a transcription start codon, was associated with increased osteosarcoma risk [odds ratio (OR), 1.78] but no association was observed for SNPs in the estrogen receptor or collagen Iα1 (7). In a separate study, an association was reported between a common variant allele of the *TNF* promoter and decreased risk for osteosarcoma but not for Ewing's sarcoma (8). SNPs in the *TP53* gene have recently been evaluated in an osteosarcoma case-control study; no strong association with osteosarcoma risk was found (9). Thus, the contribution of germ-line genetic variants to risk of sporadic osteosarcoma is largely unexplored.

Suggested osteosarcoma risk factors include height and birth weight, but several small studies have yielded conflicting results (10-13), However, because osteosarcoma most commonly occurs in long bones during or shortly after the

Received 3/8/07; revised 5/24/07; accepted 6/6/07.

Grant support: Intramural Research Program of the NIH, National Cancer Institute, Center for Cancer Research, and Division of Cancer Epidemiology and Genetics.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Note: Supplementary data for this article are available at Cancer Epidemiology, Biomarkers & Prevention Online (http://cebp.aacrjournals.org/).

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).

Requests for reprints: Sharon A. Savage, Clinical Genetics Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, 6120 Executive Boulevard, EPS/7018, Rockville, MD 20852. Phone: 301-496-5785; Fax: 301-496-1854. E-mail: savagesh@mail.nih.gov

doi:10.1158/1055-9965.EPI-07-0214

adolescent growth spurt, growth-regulatory pathways may play an important role in osteosarcoma pathogenesis (1-3). The role of growth factors and osteosarcoma is further supported by case reports of osteosarcoma occurring in individuals with acromegaly (14). It is possible that subtle variations in genes that regulate growth, such as those in the growth hormone or the insulin-like growth factor (IGF) pathways, could lead to changes in cellular signaling, altered growth, and increased risk of tumorigenesis. Circulating levels of IGF1 have been associated with the risk of developing prostate, breast, colorectal, and lung cancer (15). A study of prostate cancer cases and controls suggested that inherited variation in the *IGF1* gene could contribute to risk of prostate cancer (16). IGF1 is a potent mitogen for human osteosarcoma cell lines (17, 18). Overexpression of *IGF2* and loss of imprinting occurs in many tumors (15, 19-21). We investigated common genetic variants of key genes (i.e., *IGF1* and *IGF2*), their receptors (*IGF1R* and *IGF2R*), and binding proteins and other genes in the growth hormone pathway as candidate osteosarcoma risk factors.

Materials and Methods

Study Design. Study subjects and controls were drawn from the Bone Disease and Injury Study of Osteosarcoma initiated in 1995 by the Harvard School of Dental Medicine and the National Cancer Institute as described previously (9, 10). Briefly, this prospective case-control study collected blood samples and questionnaire data on individuals seen at orthopedic surgery departments in 10 U.S. medical centers between 1994 and 2000 (Massachusetts General Hospital, Boston, MA; Creighton University/St. Joseph's Hospital and University of Nebraska, Omaha, NE; Children's National Medical Center and Washington Hospital Center, Washington, DC; University of Chicago and Rush Presbyterian St. Luke's, Chicago, IL; University of Florida, Gainesville, FL; University of California, Los Angeles, CA; and Cleveland Clinic, Cleveland OH). Osteosarcoma patients were identified at the time of limb salvage surgery. Orthopedic controls from the Bone Disease and Injury Study of Osteosarcoma were the control group for the current analyses. They were individuals with benign tumors (26%) and other nonneoplastic conditions, such as inflammatory diseases, cysts, and trauma (International Classification of Diseases, Ninth Revision, Clinical Modification codes 289.1, 277.8, 354.0, V711, 682.6, 714.3-756.59, 810.0-885.0, 959.7), excluding those with hip fracture or osteoporosis. Institutional review boards at each of the medical centers approved the study protocol and informed consent was obtained from all study subjects. Approximately 14% of eligible subjects were not enrolled and a further 3% declined to participate in any of the study components (10). Of subjects who were enrolled in the study, 93.5% of cases and 98% of controls completed interviews. Reasons for nonparticipation in the questionnaire component of the study included death, extreme illness following surgery or failure to return for follow-up care, and refusal. Cases (6.3%) and controls (20%) completed interviews but declined to participate in the blood component of the study (10).

The average age of osteosarcoma cases was 26.4 (SD, 16.2; range, 8-77 years) and of orthopedic controls was 24.9 (SD, 14.5; range, 7-68 years; ref. 9). Osteosarcoma cases and orthopedic controls had nearly equal numbers of males and females. There were no identified cases of Paget's disease of the bone in this study.

DNA was isolated by standard methods from blood samples collected from newly diagnosed osteosarcoma cases (n = 124) and orthopedic controls (n = 87). The current analysis was restricted to individuals self-described as Caucasians (104 osteosarcoma cases and 74 orthopedic controls) to reduce potential effects of population stratification.

Genotyping. SNPs were genotyped by Taqman assays (Applied Biosystems) or MGB Eclipse (Epoch Biosciences) at the National Cancer Institute's Core Genotyping Facility.[8] Assay design and conditions are publicly available on the National Cancer Institute's SNP500Cancer Web site[9] (22). Internal, replicated, assay controls for each genotype cluster and a negative control were included on each 384-well assay plate. Fifty-two SNPs were chosen from 13 genes that play critical roles in growth and development as shown in Table 1A and as follows (OMIM abbreviation, Gene ID): *growth hormone 1* (*GH1*, 2688), *growth hormone receptor* (*GHR*, 2690), *dedicator of cytokinesis 5* (*DOCK5*, 80005), *gonadotropin-releasing hormone 1* (*GNRH1*, 2796), *potassium channel tetramerization domain 9* (*KCTD9*, 54793), *gonadotropin-releasing hormone receptor* (*GNRHR*, 2798), *IGF1* (3479), *IGF1R* (3480), *IGF2* (3481), *IGF2R* (3482), *IGF-binding protein 3* (*IGFBP3*, 3486), *IGF-binding protein 5* (*IGFBP5*, 3488), and IGF-binding protein, acid labile subunit (*IGFALS*, 3483). SNPs in *DOCK5* and *KCTD9* were included because haplotype-tagged SNPs for *GNRHR* reside in these adjacent genes, as determined by the Breast and Prostate Cancer Cohort Consortium (23). Genes/SNPs investigated were based on four major criteria: (*a*) genes that play a critical role in growth, (*b*) genes with well-described annotation in the SNP500Cancer and other databases, (*c*) the assay was in used and validated by the National Cancer Institute's Core Genotyping Facility, and (*d*) if possible, nonsynonymous SNPs were included. SNP nomenclature is denoted as recommended by the Human Genome Organization/Human Genetic Variation Society (24).

Sequence Analysis. PCR primers were designed using Primer3[10] appended with M13 forward or reverse tags (sequences available on request). Genomic DNA was amplified by PCR with MJ Research model PTC-225 thermal cyclers (Bio-Rad Laboratories, Inc.) under the following conditions: 10 ng genomic DNA, 0.2 μmol/L of each primer, 200 μmol/L of each deoxynucleotide triphosphate, 2 mmol/L $MgCl_2$, 0.5 unit AmpliTaq Gold DNA polymerase (ABI-Perkin-Elmer), and the manufacturer's buffer. Bidirectional sequencing of amplified DNA was done using the Dye Terminator method (ABI-Perkin-Elmer) and M13 forward and reverse primers and analyzed on ABI-Perkin-Elmer platforms (models 3100 and/or 3700) with Sequence Analysis 3.7 software and Sequencher 4.0.5 software (Gene Codes Corp.). In *IGF2R*, exon 13 to intron 18 (4.7 kb) and coding regions of exons 19 and 20 were resequenced in 30 osteosarcoma cases [6 wild-type (WT), 6 heterozygous, and 18 homozygous variant], 33 orthopedic (23 WT and 10 variant), and 5 osteosarcoma cell lines (HOS, SaOS2, G292, U2, and MG63).

Data Analyses. Each SNP was tested in controls for fitness of Hardy-Weinberg proportion. SNPs with χ^2 P values of <0.05 were excluded from analysis. Osteosarcoma cases were compared with orthopedic controls using contingency tables, additive and dominant (homozygous WT compared with heterozygous plus homozygous variants, i.e., one copy of the variant allele was considered equal to having two copies of the variant) genetic models, and Statistical Analysis System v8.02 software.

Haplotypes were constructed and a case-control permutation test was done using PHASE v2.1 (25, 26). Genotype data from 116 Caucasian individuals in the multiethnic cohort (27)[11] were used to construct haplotypes for common *IGF2R* SNPs (>10% minor allele frequency; Supplementary Table S1). Haploview (v2.05) was used to generate linkage disequilibrium

[8] http://cgf.nci.nih.gov
[9] http://snp500cancer.nci.nih.gov
[10] http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi
[11] http://www.uscnorris.com/MECGenetics/

plots based on absolute D′ values (pairwise comparison of SNPs) and the R-squared statistic (correlation coefficient) for cases and controls separately (25, 26). Haplotype blocks were determined based on PHASE haplotypes using the confidence intervals (CI) set in Haploview (upper CI, 0.98; lower CI, 0.7; ref. 28). Haplotype tagging SNPs were estimated using tagSNPs software (29). HaploStats (30) was used to construct haplotypes and to determine the global score *P* value, haplotype frequencies, and ORs.

PupaSNP[12] (31) and GeneSplicer[13] (32) were used to determine possible intron-exon splice site changes. Evaluation of CpG islands was conducted with the CpG Island Searcher[14] (33).

Methylation Studies. Bisulfite modification of germ-line DNA from cases and orthopedic controls was done using the EZ DNA Methylation kit (Zymo Research) following the manufacturer's instructions. Methylation status in the region of the *IGF2R* Ex16+88G>A polymorphic site was determined using pyrosequencing technology as described (34). The assay to determine methylation status at Ex16+88G>A was designed using PSQ Assay Design software (Biotage). The PCR product for each pyrosequencing reaction was generated in a 25 μL reaction containing 1× AmpliTaq Gold Master Mix, 2 units AmpliGold Taq (Applied Biosystems), 250 μmol/L deoxynucleotide triphosphates, 0.2 μmol/L forward primer (IGF2R-F-5′-CAAAAAACGCGATAAAAGACAA-3′), and 0.2 μmol/L reverse primer (IGF2R-R-5′-biotin-GGGATGATCCAATTGAATTATAGA-3′). The amplifications were carried out at 95°C for 10 min followed by 40 cycles at 95°C for 1 min, 60°C for 1 min, 72°C for 1 min, and a 10-min extension at 72°C. ssDNA from 8 μL of each PCR sample was generated following the PSQ 96 sample preparation guide using a vacuum filtration sample device following the manufacturer's instructions (Biotage). The single-stranded product was annealed to 0.3 μmol/L of sequencing primer IGF2R-S-5′-ATATCTTTCATTATTATAAA-3′, placed at 85°C for 2 min, and cooled to room temperature. Pyrosequencing was done on a PSQ 96 HS system (Biotage) with the Biotage reagent kit. Raw data were analyzed with the allele quantification algorithm of the provided software. Each sample was run in triplicate and mean percentage methylation value was computed for each CpG site. Five osteosarcoma cell lines (HOS, SaOS2, G292, U2, and MG63) with known genotypes (GG, GA, or AA) were used as standards and run on all plates.

Results

Screening of Genetic Variation in Genes Critical for Growth Regulation. Initially, 13 genes (*GH1, GHR, DOCK5, GNRH1, KCTD9, GNRHR, IGF1, IGF1R, IGF2, IGF2R, IGFBP3, IGFBP5*, and *IGFALS*) important in growth and development were evaluated by genotyping 40 SNPs (Table 1A). SNPs were chosen because they represented a broad array of genetic variation across the 13 genes and genotyping assays were available at the start of the study.

Allele frequencies, *P* values, and ORs of all SNPs investigated are shown in Supplementary Table S2. The most notable risk association from the initial analysis was observed between osteosarcoma and the *IGF2R* gene (Table 2A). In the initial screen of seven *IGF2R* SNPs, two SNPs were associated with increased osteosarcoma risk compared with orthopedic controls: a synonymous SNP in exon 16, Ex16+88G>A (T713T), and a nearby SNP in intron 16, IVS16+15C>T (OR, 2.27; 95% CI, 1.12-4.58 and OR, 2.29; 95% CI, 1.14-4.60, respectively). One

Table 1. Genotyped SNPs

Gene	SNP	Location and nucleotide change
A. SNPs genotyped in the initial screen of genetic variation in growth pathway genes*		
GH1	rs6171	−67A>G
	rs6173	Ex1-14T>G
GHR	rs2940927	−143587G>A
	rs1876790	−143121T>C
	rs7735889	IVS-3767A>G
	rs6179	Ex6-61G>A G186G
	rs6182	Ex10+374G>T C440F
	rs6180	Ex10+685A>C I544L
	rs6184	Ex10+790 P579TC>A
DOCK5	rs2709618	Ex20+8G>A P726P
GNRH1	rs6185	Ex1-95G>C W16S
KCTD9	rs1812594	Ex12+93A>G L382L
GNRHR	rs13138607	Ex1+255C>T
	rs4986942	Ex1-70C>T S151S
	rs10031252	IVS1+1241T>A
	rs3822196	IVS1-429T>C
	rs3796718	IVS1-2676A>G
	rs1843593	7,199 bp 3′ of STP A>G
	rs2630488	10,829 bp 3′ of STP C>T
IGF1	rs2162679	IVS1-1682 A>G
	rs1019731	IVS2+4996G>T
	rs2288378	IVS2-16540G>A
	rs6220	Ex4+1830A>G
IGF1R	rs907806	IVS2-89673A>G
	rs9282718	IVS2+61306T>G
	rs2137680	IVS2+61405A>G
	rs2175795	IVS2+61518A>G
	rs3743258	IVS5+289G>A
	rs3743259	IVS5+311A>G
	rs2272037	IVS7-20C>T
	rs2229765	Ex16-58G>A
	rs2016347	3,128 bp 3′ of STP T>G
	rs9282715	3,164 bp 3′ of STP T>C
IGF2	rs3213216	IVS1+1280G>A
	rs3213221	IVS1-285G>C
	rs3213223	IVS1-171C>T
	rs734351	IVS2+384T>C
IGF2R	rs1570070	Ex9+5A>G S350S
	rs894817	Ex12-32G>A G530G
	rs998075	Ex16+88G>A T713T
	rs998074	IVS16+15C>T
	rs629849	Ex34-93G>A R1619G
	rs2282140	IVS34+20C>T
	rs1803989	Ex45+11T>C L2222L
IGFBP3	rs2471551	IVS2-17C>G
	rs9282734	Ex3+70A>C H158P
	rs6413441	IVS4-707A/−
IGFBP5	rs2067039	7,581 bp 3′ of STP T>A
IGFALS	rs9282731	Ex2-359C>T
	rs2230053	Ex2-435G>A
	rs17559	Ex2-615C>T
	rs3751893	Ex2+194C>T
B. Additional *IGF2R* SNPs genotyped that were selected based on the haplotype structure around Ex16+88G>A and IVS16+15C>T		
IGF2R	rs9456497	IVS4-2176A>G
	rs4709390	IVS7-22T>G
	rs4709392	IVS10-1230G>A
	rs435612	IVS11-244G>A
	rs687088	IVS12-492C>T
	rs416572	IVS15+213C>T
	rs648253	IVS17+210A>G
	rs4709393	IVS19+348T>G
	rs3777411	IVS19-511C>T
	rs7746102	IVS22-42T>C
	rs2065396	IVS30+327G>C
	rs3798180	IVS31+651G>A

*The rs number from dbSNP, genomic location, nucleotide change, and amino acid change (if applicable) are noted. SNPs in *DOCK5* and *KCTD9* were genotyped because they comprised the haplotype tagging SNPs for *GNRH1* (see Materials and Methods).

[12] http://pupasnp.bioinfo.ochoa.fib.es/
[13] http://www.cbcb.umd.edu/software/GeneSplicer/
[14] http://www.cpgislands.com/

Table 2. Osteosarcoma case-control analysis

A. Osteosarcoma case-control analysis of *IGF2R* screening SNPs*

IGF2R SNP	Genotype	Osteosarcoma cases		Orthopedic controls		Additive	Dominant	
		n = 104	%	*n* = 74	%	χ^2 *P*	OR (95% CI)	*P*
Ex9+5A>G	AA	48\104	46.2	33/74	44.6	0.630		0.837
S350S	AG	42\104	40.4	34/74	45.9			
rs1570070	GG	14\104	13.5	7/74	9.5			
Ex12-32G>A	GG	48\104	46.2	34/74	45.9	0.999		0.978
G530G	GA	45\104	43.3	32/74	43.2			
rs894817	AA	11\104	10.6	8/74	10.8			
Ex16+88G>A	GG	18\103	17.5	24/74	32.4	0.010	2.27 (1.12-4.582)	0.021
T713T	GA	54\103	52.4	40/74	54.1			
rs998075	AA	31\103	30.1	10/74	13.5			
IVS16+15C>T	CC	18\104	17.3	24/74	32.4	0.008	2.29 (1.14-4.60)	0.019
rs998074	CT	53\104	51.0	40/74	54.1			
	TT	33\104	31.7	10/74	13.5			
Ex34-93G>A	GG	75\103	72.8	61/74	82.4	0.326		0.135
R1619G	GA	26\103	25.2	12/74	16.2			
rs629849	AA	2\103	1.9	1/74	1.4			
IVS34+20C>T	CC	78\97	80.4	47/69	68.1	0.148		0.070
rs2282140	CT	17\97	17.5	21/69	30.4			
	TT	2\97	2.1	1/69	1.4			
Ex45+11T>C	TT	97\103	94.2	63/74	85.1	0.044	0.35 (0.13-0.98)	0.044
L2222L	TC	6\103	5.8	11/74	14.9			
rs1803989	CC	0\103	0.0	0/74	0.0			

B. *IGF2R* haplotypes in osteosarcoma cases and orthopedic controls as determined using the screening SNPs †

Haplotype							Osteosarcoma case frequency	Orthopedic frequency	Case-control *P*	OR (95% CI)	Global test *P*
Ex9+5A>G	Ex12-32G>A	Ex16+88G>A	IVS16+15C>T	Ex34-93G>A	IVS34+20C>T	Ex45+11T>C					
A	G	G	C	G	C	C	0.279	0.352	0.115‡	Reference	
A	G	A	T	G	C	C	0.263	0.205	0.168		
G	A	G	C	G	T	C	0.070	0.138	0.166		
G	A	G	C	G	C	T	0.029	0.073	0.206		
G	A	A	T	G	C	C	0.122	0.065	0.052	2.59 (1.00-6.71)	0.08
A	G	A	T	A	C	C	0.059	0.060	0.642		
G	A	A	T	A	C	C	0.065	0.033	0.098		
A	G	G	C	G	T	C	0.020	0.030	0.778		
A	G	A	T	G	T	C	0.037	0.014	0.312		
G	A	A	T	A	T	C	0.015	0.000	NA		
A	G						0.663	0.662	0.980‡	Reference	
G	G						0.014	0.014	0.966		0.41
G	A						0.322	0.311	0.848		
		G	C				0.567	0.405	0.002‡	Reference	0.006
		A	T				0.428	0.595	0.003	2.04 (1.29-3.24)	
				G	C	C	0.694	0.653	0.347‡	Reference	0.142
				G	C	T	0.029	0.724	0.056	0.324 (0.10-1.02)	
				G	T	C	0.135	0.181	0.130		
				A	C	C	0.130	0.077	0.486		

*SNPs were analyzed using an additive genetic model, which compares the three possible genotypes, and a dominant genetic model in which one copy of the variant allele is considered equal to having two copies, and these individuals are compared with the homozygous WT individuals.
† Haplotypes were determined by HaploStats (v1.2.0). The most common haplotype was used as the reference to which other haplotypes were compared. Nonshaded boxes indicate the WT (most frequent in orthopedic controls) allele and gray boxes indicate the variant allele. The case-control *P* value, ORs, and 95% CIs are based on comparison with the most common haplotype. The global test *P* value is derived by HaploStats from a comparison of overall differences in cases and controls.
‡ The *P* value obtained for the reference haplotype (most common) was based on comparison of this haplotype to all other haplotypes.

SNP at the 3′-end of *IGF2R*, Ex45+11T>C (L2222L), was possibly associated with protection from osteosarcoma (OR, 0.35; 95% CI, 0.13-0.98). Similar results were noted in analyses of osteosarcoma cases <30 years of age. SNPs in *IGF2R* seemed to be more strongly associated with osteosarcoma than the other genes in all comparisons; therefore, we chose to focus on characterizing the genetic variation within this gene and its relationship to osteosarcoma risk.

Characterization of Genetic Variation in *IGF2R*

IGF2R Haplotypes Generated by the Seven Screening SNPs. Haplotype analysis revealed potential differences between different regions of the *IGF2R* gene in osteosarcoma cases and orthopedic controls. Haplotypes determined by HaploStats using the seven SNPs initially genotyped in *IGF2R* (Ex9+5A>G, Ex12-32G>A, Ex16+88G>A, IVS16+15C>T,

Ex34-93G>A, IVS34+20C>T, and Ex45+11T>C) in this study suggested an increased risk (OR, 2.59; 95% CI, 1.0-6.7; $P = 0.05$) among individuals with the GAATGCC haplotype (Table 2B). The AT haplotype at Ex16+88G>A and IVS16+15C>T showed an OR of 2.04 (95% CI, 1.29-3.24; $P = 0.003$), suggesting that this region of the gene could be associated with an increased risk of osteosarcoma. The pattern of linkage disequilibrium across the gene (Fig. 1) suggested that SNPs, IGF2R Ex16+88G>A and IGF2R IVS16+15C>T, were part of a small block of linkage disequilibrium. The SNPs genotyped in the 3′ region reside in a distinct block and thus may confer a different risk effect.

Bioinformatic Analyses. Based on publicly available genotype data from 116 Caucasians in the multiethnic cohort, we determined the haplotype structure across the *IGF2R* gene using 95 SNPs that had a minor allele frequency of >10% (Fig. 1; Supplementary Table S1; ref. 27).[11] A similar pattern of linkage disequilibrium pattern was observed using the Caucasians from the HapMap (data not shown). Using the CI method to define haplotype blocks, it seemed that the risk-associated SNPs (Ex16+88G>A and IVS16+15C>T) were located in a small haplotype block (Fig. 1, *block 5*). However, it was also possible that blocks 5, 6, and 7 comprised one larger haplotype block in Caucasians. Therefore, 12 additional SNPs were chosen for genotyping in osteosarcoma cases and orthopedic controls: 7 SNPs to further define blocks 5, 6, and 7 (IVS15+213C>T, IVS17+210A>G, IVS19+348T>G, IVS19-511C>T, IVS22-42T>C, IVS30+327G>C, and IVS31+651G>A), 3 SNPs because they resided in a different block (IVS4-2176A>G, IVS7-22T>G, and IVS10-1230G>A), and 2 SNPs that

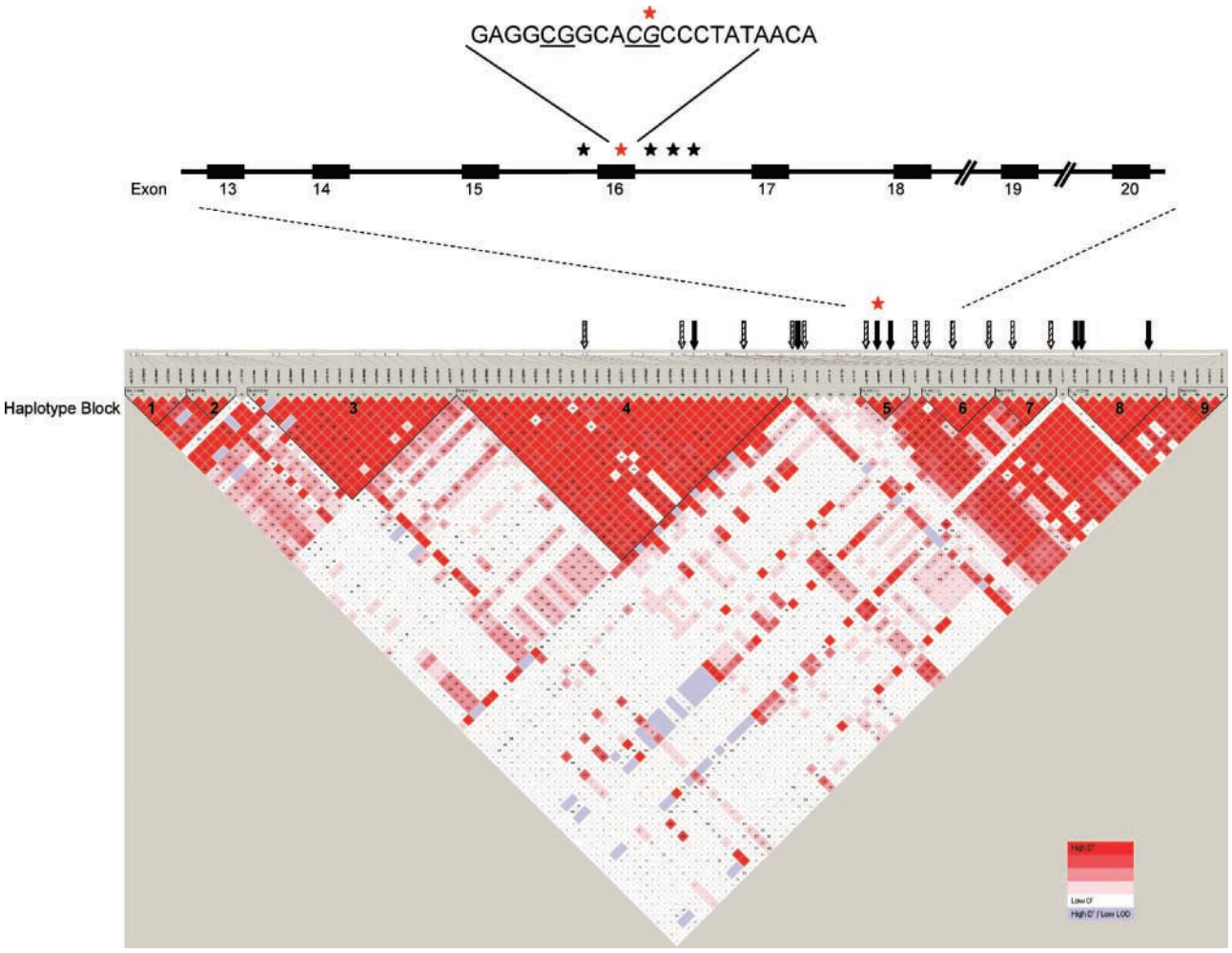

Figure 1. Gene structure of the IGF2R. *IGF2R* (MIM no. 147280) consists of 48 exons spanning 138-kb pairs on chromosome 6q26 with multiple apparent haplotype blocks. These data were used to define a region of *IGF2R* for haplotype-tagged SNP analysis. Haplotypes were determined by PHASE v2 and linkage disequilibrium of SNPs spanning 132-kb pairs (intron 1 to intron 47) with >10% frequency in 116 Caucasian individuals from the multiethnic cohort (listed in Supplementary Table S2) was determined in the PHASE haplotypes with Haploview v3.2. The linkage disequilibrium plot shows absolute D′ values and haplotype blocks as determined by the CI (28) method. D′ values are determined based on the likelihood of linkage disequilibrium between two SNPs. A D′ value of 1 indicates complete linkage disequilibrium and 0 indicates no linkage disequilibrium. Each *square* represents a pairwise comparison between two SNPs. *Red squares,* high degrees of linkage disequilibrium; *white squares,* no linkage disequilibrium; *light purple squares,* high D′ values but lower levels of statistical significance (LOD scores). *Black arrows,* first set of SNPs genotyped in this study; *dashed arrows,* second set of SNPs that were chosen based on haplotype tagging predictions from the multiethnic cohort subjects. SNPs are listed 5′ to 3′ as noted in Table 1. If a SNP was genotyped (in first round) that is not on the figure, the *arrow* is placed between other SNPs to indicate its relative position in the gene. The region sequenced in osteosarcoma cases and orthopedic controls is noted as exon 13 to intron 18 and the coding regions of exons 19 and 20. *Black stars above gene schematic,* SNPs (from *left* to *right*), IVS15+213C>T, Ex16+88G>A, IVS16+15C>T, new1, and IVS22-42T>C; *red star,* Ex16+88G>A. *Underlined bases,* two CpG islands evaluated for methylation.

were located in an area without apparent linkage disequilibrium (IVS11-244G>A and IVS12-492C>T).

Additional IGF2R Genotypes and Haplotypes. These analyses showed that IVS15+213C>T was possibly associated with osteosarcoma risk (OR, 1.85; 95% CI, 1.01-3.39; $P = 0.048$; Supplementary Table S3A). IVS15+213C>T, Ex16+88G>A, and IVS16+15C>T lie within 493 bp of each other. SNPs from the flanking blocks were not associated with osteosarcoma (Table 1; Fig. 1). In addition, three SNPs downstream from the risk-associated SNPs may be associated with protection from osteosarcoma: IVS22-42T>C (OR, 0.54; 95% CI, 0.29-0.99), IVS30+327G>C (OR, 0.50; 95% CI, 0.27-0.93), and IVS31+651G>A (OR, 0.35; 95% CI, 0.18-0.68).

Haplotypes constructed from the 19 genotyped SNPs in *IGFR2* showed that the global test *P* value for differences between osteosarcoma cases and orthopedic controls was 0.065 (Supplementary Table S3B). Osteosarcoma cases were compared with orthopedic controls and global test *P* values and ORs were determined for the SNPs comprising each haplotype block. The addition of IGF2R-15 to the haplotype analysis showed OR for osteosarcoma cases compared with orthopedic controls with the CAT haplotype to be 2.00 (95% CI, 1.15-3.47), whereas the TAT haplotype OR was 2.20 (95% CI, 1.23-3.92). This suggests that the greatest contribution to risk is probably from the Ex16+88G>A and IVS16+15C>T AT haplotype.

IGF2R Sequence Analyses. The 4.7-kb pair region between exon 13 and intron 18 and the coding regions of exons 19 and 20 were resequenced in individuals with and without the Ex16+88G>A and IVS16+15C>T variant alleles, including 30 osteosarcoma cases (6 WT, 6 heterozygous, and 18 variant), 33 orthopedic controls (23 WT and 10 variant), and 5 osteosarcoma cell lines, in an effort to identify rare mutations that could be linked to the more common SNPs/haplotypes in osteosarcoma cases. The observed variation present in the osteosarcoma cell lines was consistent with those seen in the osteosarcoma cases (Supplementary Table S4; Supplementary Fig. S1). A 2-bp deletion (AA) was identified in intron 18 (IVS18+280-281) in two cell lines, HOS and MG63, but no novel coding mutations were observed in study subjects nor osteosarcoma cell lines. Sequence analysis in the cases and controls identified an untyped SNP in intron 17 (IVS17+160A>T). This SNP is in strong linkage disequilibrium with the SNPs already genotyped in the study. A SNP in intron 17 (IVS17+210A>G), which was not strongly associated with osteosarcoma in our case-control study, frequently cosegregated with the risk alleles. A SNP in intron 13 (IVS13+64G>T) was common and was often found on the risk haplotype of Ex16+88G>A and IVS16+15C>T (AT), but it was not in strong linkage disequilibrium with these SNPs (absolute D′ = 0.68 and 0.56; R-squared = 0.21 and 0.29 for osteosarcoma cases and orthopedic controls, respectively). This sequence analysis suggested that linkage disequilibrium was strongest in a 1-kb pair region surrounding SNPs Ex16+88G>A and IVS16+15C>T (intron 15 to intron 16).

Functional Studies of IGF2R Ex16+88G>A. We examined IVS15+213C>T, Ex16+88G>A, and IVS16+15C>T *in silico* using PupaSNP (31) and GeneSplicer (32) and determined that none alters intron-exon splice sites. Because *IGF2R* is imprinted in the mouse models and emerging data suggest the same for humans (35, 36), we investigated the region associated with osteosarcoma risk for CpG islands. CpG Island Searcher (33) predicted that Ex16+88G>A and IVS16+15C>T were located in a CpG island (Fig. 1; data not shown) and that the A allele of Ex16+88G>A removes a CpG site that could potentially change methylation status.

Percentage methylation of genomic DNA was determined at Ex16+88G>A and another, nonvariant CpG site 6 bp upstream of the SNP (Fig. 1), in 15 osteosarcoma cases (5 GG, 5 GA, and 5 AA genotypes at Ex16+88G>A) and 8 orthopedic controls (3 GG, 1 GA, and 4 AA genotypes). Regardless of case-control status, individuals who were homozygous GG (WT) had 85% (±6%) methylation at that site (Fig. 2). Those who were heterozygous for the SNP (GA) had 54% (±12%) methylation

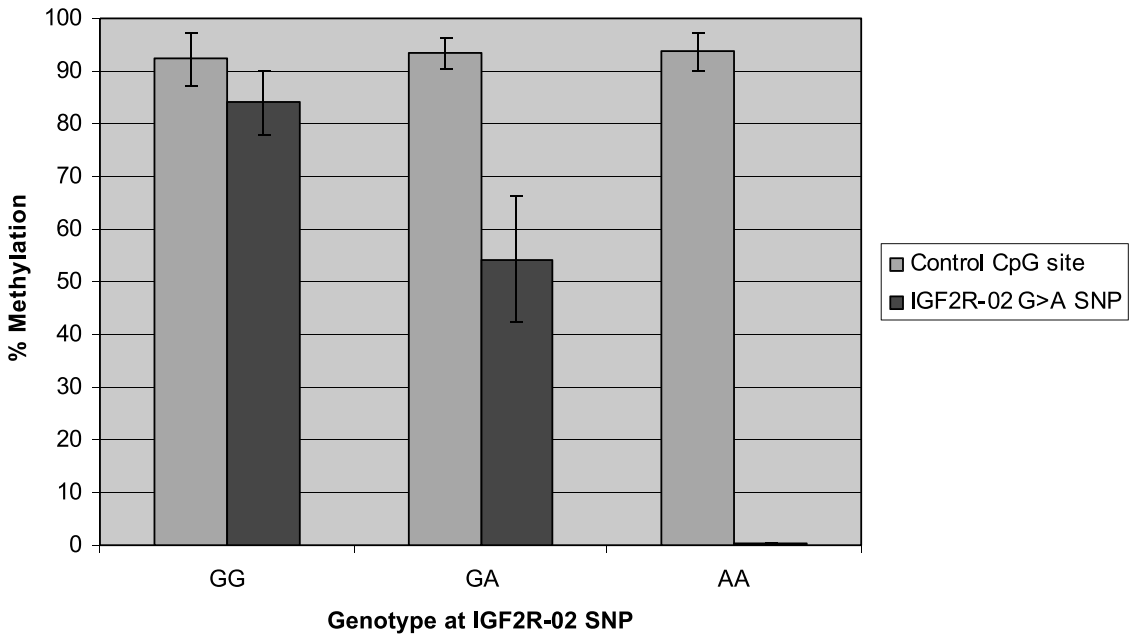

Figure 2. Effect of IGF2R-02 G>A SNP on methylation. Genomic DNA from osteosarcoma cases and orthopedic controls was evaluated for the degree of methylation at this SNP site. The sequence evaluated was GAGG*CG*GCA*C(G>A)*CCCTATAACA. The control CpG site is the *first underlined CG site*. The SNP site is the *second underlined site* with the G>A SNP noted. Percentage methylation for osteosarcoma cases and orthopedic controls at each site is noted for GG, GA, and AA genotypes. *Bars,* SD of percentage methylation. Individuals with the AA genotype had 0% to 2% methylation at IGF2R-02 (average, 0.22; SD, 0.22).

and individuals homozygous for the variant (AA) had 0.2% methylation (±0.2%) at Ex16+88G>A. The nonvariant CpG site showed 92% to 94% average methylation regardless of genotype at the variant Ex16+88G>A site.

Discussion

In this study, we report that common SNPs in a defined region of the *IGF2R* gene are associated with risk of osteosarcoma. After screening the subjects of a hospital-based, prospective case-control study of osteosarcoma, we noted that two SNPs, in close proximity, were in a block linkage disequilibrium in *IGF2R*, which conferred an increased risk of osteosarcoma. Additional analyses further defined the linkage disequilibrium and common haplotypes occurring around these two risk SNPs, Ex16+88G>A and IVS16+15C>T. The supplemental analysis of the multiethnic cohort high frequency SNPs (>10% in Caucasians) enabled us to identify an additional SNP, IVS15+213C>T, with slightly less effect but still in partial linkage disequilibrium with Ex16+88G>A and IVS16+15C>T. Based on our genotype results, we conducted directed resequence analysis of cases and controls both with and without the haplotypes associated with risk of osteosarcoma. Although we did not find new or rare mutations, we did identify two additional SNPs of interest, IVS17+160A>T and IVS17+210A>G.

Ex16+88G>A resides in a region predicted to be a CpG island. The presence of the variant A allele of Ex16+88G>A results in loss of methylation at that CpG site. At this site, eight GG (five osteosarcoma cases and three orthopedic controls), six GA (five osteosarcoma cases and one orthopedic control), and nine AA (five osteosarcoma cases and four orthopedic controls) individuals were studied. Although the relative numbers are small, the loss of methylation is significant and consistent with predicted CpG island locations. The G to A change prevents methylation at that site. Although this SNP may not radically affect gene or protein effect function, because it codes for a synonymous amino acid change, it is possible that, due to modified methylation at this site, a subtle alteration in gene expression or splicing could occur and possibly alter cancer risk. Although the biological significance, if any, of the loss of methylation at this site is unclear, these findings suggest genetic variation may facilitate epigenetic changes that could lead to altered gene expression and function.

IGF2R has critical functions in growth and development, and limited studies of these effects in the context of genetic variation are intriguing. For instance, there is increased transmission of the G allele of Ex16+88G>A to children with insulin-dependent diabetes mellitus from mothers compared with those with the A allele (37). This disproportionate transmission of the maternal allele along with recent evidence for imprinting of *IGF2R* in humans (35, 38) suggest that variants in *IGF2R* could be important in disease risk. A study of childhood growth and *IGF2R* Ex34-93G>A did not show an association with birth size, but AA homozygotes had slower growth velocity and remained shorter during the 7 years of childhood studied (39). Individuals who are taller than their peers are at increased risk of osteosarcoma in some, but not all, studies (6, 11, 13, 40). However, in our study, we observed no association between *IGF2R* Ex34-93G>A and osteosarcoma.

Our findings in osteosarcoma could be applicable to other types of cancer. For example, because *IGF2* is overexpressed in many tumor cell lines (15, 19, 21), it is possible that interactions with its receptor, IGF2R, are critical in tumorigenesis and that genetic variation in *IGF2R* may affect this interaction. Specifically, because IGF2R binds only to IGF2 and acts as an antagonist to IGF2, it has been considered a potential tumor suppressor (21, 41). Ongoing studies in the Breast and Prostate Cancer Cohort Consortium are investigating the role of common genetic variants in the IGF pathway for possible associations with breast and prostate cancer (23).

This study is limited by its small sample size relative to the number of tests done. As in all genetic association studies, it is possible that false-positive associations have been identified and/or that false-negative associations have been missed. Despite these limitations, this study has several strengths. The genes were selected based on a high prior probability of significance based on biological relevance to specific growth pathways, which have been implicated in osteosarcoma pathogenesis. By selecting biologically plausible genes, the identification of false positives is less likely (42). SNPs throughout the genome are in varying degrees of linkage disequilibrium and most cannot be considered independent markers of risk (43). The small haplotype block in *IGF2R* that was identified adds evidence that this region of the gene may be important in osteosarcoma risk. We present this as a pilot study and believe we have provided sufficient new detailed genetic and laboratory evidence to warrant follow-up studies in osteosarcoma aimed at clarifying the contribution of common genetic variants in *IGF2R* to osteosarcoma risk.

In conclusion, we have identified a SNP that alters methylation in an *IGF2R* CpG island that seems to be associated with increased risk of osteosarcoma. Our results will inform the design of larger case-control studies, as well as studies of parents and affected children, which will also be required to evaluate the role of differential transmission of the *IGF2R* Ex16+88G>A SNP in the context of osteosarcoma.

Acknowledgments

We thank the patients for participating in this study, the members of the National Osteosarcoma Etiology Study Group for the data collection, and Dr. Rebecca Troisi (National Cancer Institute) for helpful discussions.

References

1. Gurney JG, Swensen AR, Bulterys M. Malignant bone tumors. In: Ries LAG, Smith MA, Gurney JG, et al., editors. Cancer incidence and survival among children and adolescents. Bethesda (MD): National Cancer Institute, NIH Pub. No. 99–4649; 1999. p. 99–110.
2. Herzog CE. Overview of sarcomas in the adolescent and young adult population. J Pediatr Hematol Oncol 2005;27:215–8.
3. Miller RW, Boice JD, Jr., Curtis RE. Bone cancer. In: Schottenfeld D, Fraumeni JF, editors. Cancer epidemiology and prevention. 2nd ed. New York (NY): Oxford University Press; 1996. p. 971–83.
4. Wittig JC, Bickels J, Priebat D, et al. Osteosarcoma: a multidisciplinary approach to diagnosis and treatment. Am Fam Physician 2002;65:1123–32.
5. Varley JM. Germline TP53 mutations and Li-Fraumeni syndrome. Hum Mutat 2003;21:313–20.
6. Fuchs B, Pritchard DJ. Etiology of osteosarcoma. Clin Orthop 2002;397:40–52.
7. Ruza E, Sotillo E, Sierrasesumaga L, Azcona C, Patino-Garcia A. Analysis of polymorphisms of the vitamin D receptor, estrogen receptor, and collagen Iα1 genes and their relationship with height in children with bone cancer. J Pediatr Hematol Oncol 2003;25:780–6.
8. Patio-Garcia A, Sotillo-Pieiro E, Modesto C, Sierrases-Maga L. Analysis of the human tumour necrosis factor-α (TNFα) gene promoter polymorphisms in children with bone cancer. J Med Genet 2000;37:789–92.
9. 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.
10. Troisi R, Masters MN, Joshipura K, Douglass C, Cole BF, Hoover RN. Perinatal factors, growth and development, and osteosarcoma risk. Br J Cancer 2006;95:1603–7.
11. Operskalski EA, Preston-Martin S, Henderson BE, Visscher BR. A case-control study of osteosarcoma in young persons. Am J Epidemiol 1987;126:118–26.
12. Pui CH, Dodge RK, George SL, Green AA. Height at diagnosis of malignancies. Arch Dis Child 1987;62:495–9.
13. Longhi A, Pasini A, Cicognani A, et al. Height as a risk factor for osteosarcoma. J Pediatr Hematol Oncol 2005;27:314–8.
14. Lima GA, Gomes EM, Nunes RC, et al. Osteosarcoma and acromegaly: a case report and review of the literature. J Endocrinol Invest 2006;29:1006–11.
15. Pollak MN, Schernhammer ES, Hankinson SE. Insulin-like growth factors and neoplasia. Nat Rev Cancer 2004;4:505–18.
16. Cheng I, Stram DO, Penney KL, et al. Common genetic variation in IGF1 and

prostate cancer risk in the Multiethnic Cohort. J Natl Cancer Inst 2006;98:123–34.
17. Kappel CC, Velez-Yanguas MC, Hirschfeld S, Helman LJ. Human osteosarcoma cell lines are dependent on insulin-like growth factor I for *in vitro* growth. Cancer Res 1994;54:2803–7.
18. Pollak MN, Polychronakos C, Richard M. Insulinlike growth factor I: a potent mitogen for human osteogenic sarcoma. J Natl Cancer Inst 1990;82:301–5.
19. Furstenberger G, Senn HJ. Insulin-like growth factors and cancer. Lancet Oncol 2002;3:298–302.
20. Grimberg A, Cohen P. Role of insulin-like growth factors and their binding proteins in growth control and carcinogenesis. J Cell Physiol 2000;183:1–9.
21. Yu H, Rohan T. Role of the insulin-like growth factor family in cancer development and progression. J Natl Cancer Inst 2000;92:1472–89.
22. Packer BR, Yeager M, Burdett L, et al. SNP500Cancer: a public resource for sequence validation, assay development, and frequency analysis for genetic variation in candidate genes. Nucleic Acids Res 2006;34:D617–21.
23. Hunter DJ, Riboli E, Haiman CA, et al. A candidate gene approach to searching for low-penetrance breast and prostate cancer genes. Nat Rev Cancer 2005;5:977–85.
24. den Dunnen JT, Antonarakis SE. Nomenclature for the description of human sequence variations. Hum Genet 2001;109:121–4.
25. Stephens M, Donnelly P. A comparison of bayesian methods for haplotype reconstruction from population genotype data. Am J Hum Genet 2003;73:1162–9.
26. Stephens M, Smith NJ, Donnelly P. A new statistical method for haplotype reconstruction from population data. Am J Hum Genet 2001;68:978–89.
27. Kolonel LN, Altshuler D, Henderson BE. The multiethnic cohort study: exploring genes, lifestyle and cancer risk. Nat Rev Cancer 2004;4:519–27.
28. Gabriel SB, Schaffner SF, Nguyen H, et al. The structure of haplotype blocks in the human genome. Science 2002;296:2225–9.
29. Stram DO. Tag SNP selection for association studies. Genet Epidemiol 2004;27:365–74.
30. Lake SL, Lyon H, Tantisira K, et al. Estimation and tests of haplotype-environment interaction when linkage phase is ambiguous. Hum Hered 2003;55:56–65.
31. Conde L, Vaquerizas JM, Santoyo J, et al. PupaSNP Finder: a web tool for finding SNPs with putative effect at transcriptional level. Nucleic Acids Res 2004;32:W242–8.
32. Pertea M, Lin X, Salzberg SL. GeneSplicer: a new computational method for splice site prediction. Nucleic Acids Res 2001;29:1185–90.
33. Takai D, Jones PA. The CpG island searcher: a new WWW resource. In Silico Biol 2003;3:235–40.
34. Colella S, Shen L, Baggerly KA, Issa JP, Krahe R. Sensitive and quantitative universal Pyrosequencing methylation analysis of CpG sites. Biotechniques 2003;35:146–50.
35. Vu TH, Li T, Hoffman AR. Promoter-restricted histone code, not the differentially methylated DNA regions or antisense transcripts, marks the imprinting status of IGF2R in human and mouse. Hum Mol Genet 2004;13:2233–45.
36. Wutz A, Smrzka OW, Schweifer N, Schellander K, Wagner EF, Barlow DP. Imprinted expression of the Igf2r gene depends on an intronic CpG island. Nature 1997;389:745–9.
37. McCann JA, Xu YQ, Frechette R, Guazzarotti L, Polychronakos C. The insulin-like growth factor-II receptor gene is associated with type 1 diabetes: evidence of a maternal effect. J Clin Endocrinol Metab 2004;89:5700–6.
38. Smrzka OW, Fae I, Stoger R, et al. Conservation of a maternal-specific methylation signal at the human IGF2R locus. Hum Mol Genet 1995;4:1945–52.
39. Petry CJ, Ong KK, Wingate DL, et al. Genetic variation in the type 2 insulin-like growth factor receptor gene and disparity in childhood height. Growth Horm IGF Res 2005;15:363–8.
40. Fraumeni JF, Jr. Stature and malignant tumors of bone in childhood and adolescence. Cancer 1967;20:967–73.
41. Oates AJ, Schumaker LM, Jenkins SB, et al. The mannose 6-phosphate/insulin-like growth factor 2 receptor (M6P/IGF2R), a putative breast tumor suppressor gene. Breast Cancer Res Treat 1998;47:269–81.
42. Wacholder S, Chanock S, Garcia-Closas M, El Ghormli L, Rothman N. Assessing the probability that a positive report is false: an approach for molecular epidemiology studies. J Natl Cancer Inst 2004;96:434–42.
43. Savage SA, Chanock SJ. Genetic association studies in cancer: good, bad or no longer ugly? Hum Genomics 2006;2:415–21.

Analysis of Genes Critical for Growth Regulation Identifies *Insulin-like Growth Factor 2 Receptor* Variations with Possible Functional Significance as Risk Factors for Osteosarcoma

Sharon A. Savage, Karen Woodson, Elyse Walk, et al.

Cancer Epidemiol Biomarkers Prev 2007;16:1667-1674.