A Genome-Wide Scan Identifies Variants in NFIB Associated with Metastasis in Patients with Osteosarcoma

Lisa Mirabello¹, Roelof Koster¹, Branden S. Moriarity²,³,⁴, Logan G. Spector², Paul S. Meltzer⁵, Joy Gary⁶, Mitchell J. Machiela¹, Nathan Pankratz⁶, Orestis A. Panagiotou¹, David Largaespada²,³,⁴,⁵, Zhaoming Wang⁶, Julie M. Gastier-Foster¹⁰, Richard Gorlick¹¹, Chand Khanna⁶, Silvia Regina Caminada de Toledo¹², Antonio S. Petrilli¹², Ana Patiño-Garcia¹³, Luis Sierrasésúmaga¹³, Fernando Lecanda¹³, Irene L. Andreulis¹⁴, Jay S. Wunder¹⁴, Nalan Gokgoz¹⁴, Massimo Serra¹⁵, Claudia Hattinger¹⁵, Piero Picci¹⁵, Katia Scotlandi¹⁵, Adrienne M. Flanagan¹⁶,¹⁷, Roberto Tirabosco¹⁷, Maria Fernanda Amary¹⁷, Dina Halai¹⁷, Mandy L. Ballinger¹⁸, David M. Thomas¹⁹, Sean Davis⁹, Donald A. Barkauskas²⁰, Neyssa Marina²¹, Lee Heiman², George M. Otto³, Kelsie L. Becklin¹, Natalie K. Wolf⁴, Madison T. Weg³, Margaret Tucker¹, Sholom Wacholder¹, Joseph F. Fraumeni Jr¹, Neil E. Caporaso¹, Joseph F. Boland², Belynda D. Hicks³, Aurelie Vogt³, Laurie Burdett³, Meredith Yeager⁹, Robert N. Hoover¹, Stephen J. Chanock¹, and Sharon A. Savage¹
ABSTRACT

Metastasis is the leading cause of death in patients with osteosarcoma, the most common pediatric bone malignancy. We conducted a multistage genome-wide association study of osteosarcoma metastasis at diagnosis in 935 osteosarcoma patients to determine whether germline genetic variation contributes to risk of metastasis. We identified an SNP, rs7034162, in NFIB significantly associated with metastasis in European osteosarcoma cases, as well as in cases of African and Brazilian ancestry (meta-analysis of all cases: \( P = 1.2 \times 10^{-6} \); OR, 2.43; 95% confidence interval, 1.83–3.24). The risk allele was significantly associated with lowered NFIB expression, which led to increased osteosarcoma cell migration, proliferation, and colony formation. In addition, a transposon screen in mice identified a significant proportion of osteosarcomas harboring inactivating insertions in Nfib and with lowered NFIB expression. These data suggest that germline genetic variation at rs7034162 is important in osteosarcoma metastasis and that NFIB is an osteosarcoma metastasis susceptibility gene.

SIGNIFICANCE: Metastasis at diagnosis in osteosarcoma is the leading cause of death in these patients. Here we show data that are supportive for the NFIB locus as associated with metastatic potential in osteosarcoma. Cancer Discov; 5(9): 1–12. © 2015 AACR.

INTRODUCTION

Metastasis to distant secondary sites is the cause of death for the majority of cancer patients; a defining feature of osteosarcoma is its high rate of metastasis. Osteosarcoma is the most common primary malignant bone tumor, with a main incidence peak in adolescence during the pubertal growth spurt and a second, smaller peak in the seventh and eighth decades of life (1–3). Epidemiologic data suggest that growth may contribute to osteosarcoma etiology, and several studies have found common genetic variants associated with risk of osteosarcoma (3–6). However, the disease course for patients with osteosarcoma is variable, and the pathogenesis and prognostic factors that contribute to this variability remain poorly understood.

Patients presenting with metastatic osteosarcoma have a very poor prognosis, with 5-year overall survival rates ranging from 11% to 29% (7–10). Approximately 10% to 25% of patients with osteosarcoma have metastases at the time of diagnosis, and up to 90% of these metastasize to the lungs (11–14). Several factors have been inconsistently proposed to be associated with primary metastasis, including older age, low socioeconomic status, larger tumor size, and tumor location (7, 9, 10, 12, 13). However, the largest international collaborative study of osteosarcoma to date found that none of the clinical or demographic factors evaluated were associated with metastasis; they found that the presence of metastasis at diagnosis increased the risk of a subsequent metastasis 5-fold (14).

To our knowledge, there have been no studies evaluating the role of genetic variation in osteosarcoma metastasis, and there may be an underlying genetic component contributing to the risk of metastasis. A connection between germline genetic variation and susceptibility to metastasis has been recently reported in other cancers (15–17); however, we are unaware of any studies evaluating the association between germline genetic variation and metastasis in pediatric cancers. Here, we assembled clinical outcome data on 935 osteosarcoma cases from our recent genome-wide association study (GWAS; ref. 5) and four independent replication sets to...
determine if common germline genetic variation is associated with risk of metastasis at diagnosis of osteosarcoma. We further used in vitro studies to functionally characterize a novel metastasis-associated locus.

RESULTS

Variants at 9p24.1 Associated with Metastasis at Diagnosis

The discovery analysis included 541 cases of European ancestry that passed quality-control metrics and had data on the presence of confirmed metastases at diagnosis (Supplementary Table S1). Metastatic disease was present in 23% of patients with osteosarcoma at diagnosis and was associated with a significantly reduced overall survival \(P = 1.9 \times 10^{-19}\), HR 4.3, 95% confidence intervals (CI), 3.13–5.91; Supplementary Table S2]. Age at diagnosis, gender, tumor subtypes, and tumor location were not significantly different in the patients with metastases at diagnosis compared with patients without metastases at diagnosis (Supplementary Table S2).

An adjusted logistic regression model with log-additive effects identified one locus at 9p24.1 associated with metastasis that approached genome-wide significance in our discovery analysis (Supplementary Table S3; Supplementary Fig. S1). Specifically, six linked intronic SNPs in the nuclear factor 1/B gene (NFIB) on chromosome 9p24.1 were significantly associated with an increased risk of metastasis in cases of African ancestry (\(N = 61\)) and Brazilian (\(N = 107\)) ancestry with GWAS data (Supplementary Table S1); in a meta-analysis combining all cases (European discovery/replication, African, and Brazilian ancestry cases; \(N = 935\)), the rs7034162 association with metastasis gained significance \(P = 1.2 \times 10^{-9}\); OR, 2.60; 95% CI, 1.87–3.62; Table 1 and Fig. 1A). Evaluation of associations with metastasis at diagnosis for the top SNP by inheritance model suggested no departure from the log-additive model (Supplementary Table S4).

We examined variants at 9p24.1 in 85 osteosarcoma cases using targeted genotyping of this region (replication set 1; Supplementary Table S1). The top 100 SNPs from the discovery set were also evaluated in an additional 141 independent cases with GWAS data (replication set 2; Supplementary Table S1). A fixed-effect meta-analysis was applied to combine the results of each replication set (226 total unique cases) and the discovery set (541 unique cases). In the meta-analysis, the 9p24.1 locus marked by SNP rs2890982 was strongly associated with metastasis at diagnosis \(P = 4.9 \times 10^{-8}\); OR, 2.60; 95% CI, 1.87–3.62; Table 1).

To further evaluate this locus, we imputed SNPs across a 1-Mb region centered on the index SNP on 9p24.1. Meta-analyses of the imputed SNPs from both replication sets and the discovery set showed a similar increased risk of metastasis associated with rs7034162 \(P = 3.3 \times 10^{-8}\); OR, 2.62; 95% CI, 1.86–3.68; \(r^2 = 0.8\) with rs2890982; Table 1 and Fig. 1B; Supplementary Table S4). Only the imputed index SNP (rs7034162) additionally showed an association with an increased risk of metastasis in cases of African \(N = 61\) and Brazilian \(N = 107\) ancestry with GWAS data (Supplementary Table S1); in a meta-analysis combining all cases (European discovery/replication, African, and Brazilian ancestry cases; \(N = 935\)), the rs7034162 association with metastasis gained significance \(P = 1.2 \times 10^{-9}\); OR, 2.60; 95% CI, 1.87–3.62; Table 1 and Fig. 1C; Supplementary Table S5).

We additionally evaluated associations for the top SNPs in the European discovery-stage cases compared with the European controls from our previous GWAS \(N = 2,703\); ref. 5; minor allele frequencies (MAF) for the controls are shown in Supplementary Table S6). For the top SNP, rs7034162, the minor allele (A) was significantly associated with an

<p>| Table 1. Summary of the top GWAS and imputed SNP associations with metastasis at diagnosis in the discovery and replication studies |
|----------------------------------|-------------|-----------|-----------|-----------|</p>
<table>
<thead>
<tr>
<th><strong>dbSNP locus</strong></th>
<th><strong>Chr: position</strong></th>
<th><strong>Allele</strong></th>
<th><strong>Stage</strong></th>
<th><strong>Ancestry</strong></th>
<th><strong>Cases with mets. EAF</strong></th>
<th><strong>Cases without mets. EAF</strong></th>
<th><strong>P</strong></th>
<th><strong>OR (95% CI)</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>rs2890982</td>
<td>9p24.1</td>
<td>C/T</td>
<td>Discovery EUR</td>
<td>122</td>
<td>0.25</td>
<td>419</td>
<td>0.11</td>
<td>4.36 × 10^{-7}</td>
</tr>
<tr>
<td>rs7034162</td>
<td>9p24.1</td>
<td>T/A</td>
<td>Discovery EUR</td>
<td>122</td>
<td>0.25</td>
<td>419</td>
<td>0.11</td>
<td>3.14 × 10^{-7}</td>
</tr>
</tbody>
</table>

Abbreviations: AFR, African ancestry; Brazil, Brazilian ancestry; EAF, effect (coded) allele frequency; EUR, European ancestry; mets., metastasis at diagnosis.

aNCBI human genome build 36 position.

bReference/effect allele.

cAncestry based on a STRUCTURE and principal component analysis.
increased risk comparing the cases with metastasis with the controls (OR, 2.04; 95% CI, 1.47–2.82, \( P = 1.84 \times 10^{-5} \)). Conversely, there was a borderline statistically significant inverse association for the minor allele of rs7034162 comparing the cases without metastasis with the controls (OR, 0.77; 95% CI, 0.60–0.99, \( P = 0.039 \)).

Furthermore, the rs7034162 risk allele (A) was significantly associated with worse overall survival in the European cases (discovery stage; \( N = 522 \)) and in all cases (European discovery, African, and Brazilian ancestry cases; \( N = 684 \)) with survival data, with a per-allele HR of 1.34 (95% CI, 1.01–1.77, log-rank \( P = 0.044 \)) and 1.45 (95% CI, 1.16–1.81, log-rank \( P = 3.3 \times 10^{-4} \); Kaplan–Meier survival curve shown in Fig. 1D), respectively. When we restricted our survival analyses to only the cases without metastasis at diagnosis, including all cases from the discovery stage, replication 3, and replication 4 (\( N = 494 \)), the rs7034162 risk allele (A) was associated with worse overall survival, with a per-allele HR of 1.29 (95% CI, 0.92–1.80, log-rank \( P = 0.124 \)), although not statistically significant.

Both rs2890982 (g.14181653 C>T) and rs7034162 (g.14190287 T>A) are located in intron 4 of the NFIB gene (Fig. 1). For rs2890982, the risk allele (T) frequencies vary by population ancestry in the 1000 Genomes Project (phase I genotype data from 1,094 individuals; ref. 18): African 0.70, Asian 0.36, American 0.21, and European 0.14. The risk allele frequencies for rs7034162 (A) show less population variation: European 0.15, American 0.18, African 0.30, and Asian 0.37; in addition, an increased risk of metastasis at diagnosis was associated with the A allele of rs7034162 across all populations studied (Supplementary Table S6).

Sixty-one markers were highly correlated with rs7034162 \( r^2 \geq 0.6 \), 1000 Genomes Project Utah residents with Northern and Western European ancestry from the CEPH collection (CEU) data; CEU haplotype block illustrated in Supplementary Fig. S2) and were associated with metastasis at diagnosis with \( P \leq 1 \times 10^{-5} \). The majority map to potential regulatory elements in the Encyclopedia of DNA Elements (ENCODE; ref. 19) dataset (Supplementary Table S6). Notably, the surrogate SNP rs12377502 maps to a DNAseI hypersensitivity region in osteoblast cells, suggesting this SNP may be within an open chromatin region and have regulatory activity in the relevant cells.

**Figure 1.** Regional plots of the discovery and combined association results, recombination hotspots, and linkage disequilibrium (LD) for the 9p24.1:14,081,842–14,398,982 osteosarcoma metastasis susceptibility locus. Y axes represent the statistical significance (\(-\log_{10}\) transformed \( P \) values) of SNP association results from a trend test (left) and the recombination rate (right). SNPs are color-coded based on pairwise LD (\( r^2 \)) with the most statistically significant SNP, which is labeled and shown in purple. Allelic \( P \) values generated by SNPTEST using score method (two-sided). Physical locations of the SNPs are based on NCBI human genome build 36, and gene annotation was based on the NCBI RefSeq genes from the UCSC Genome Browser. A, European discovery set. B, meta-analyzed European discovery and replication sets (imputed). C, meta-analyzed European discovery/repli­cation sets, plus African and Brazilian ancestry cases (imputed). D, Kaplan–Meier curve of cumulative (cum) survival for patients with osteosarcoma by the rs7034162 genotype.
Decreased NFIB Expression Is Associated with the Risk Allele of rs7034162

We performed expression quantitative trait locus (eQTL)-based analyses using publicly available expression and genotyping data on 17 osteosarcoma cell lines and 29 tumors (20). We evaluated whether top-ranking SNPs were associated with expression of NFIB or other neighboring protein-encoding genes. The risk allele (A) of rs7034162 was significantly associated with a decrease in NFIB expression in osteosarcoma cell lines (N = 17, P = 0.0059) and osteosarcoma tumors (N = 29, P = 0.0211; Fig. 2; Supplementary Fig. S3). There was no association between rs7034162 genotypes and expression of other nearby protein-encoding genes (FREM1, ZDHHC21, and MPDZ; Fig. 2; Supplementary Fig. S4).

NFIB Expression Levels Are Associated with Migration and Growth of Osteosarcoma Cells

The ability of tumor cells to invade and migrate is an important marker of metastatic potential. Therefore, to evaluate the possible involvement of NFIB in osteosarcoma metastatic potential, we analyzed the invasion and migration capacity of three human osteosarcoma cell lines (U2OS, HOS, and OSA) with different expression levels of NFIB. U2OS and HOS cells carry the homozygous non-risk allele (rs7034162: TT), whereas OSA cells carry the homozygous risk allele (rs7034162: AA). U2OS and HOS had significantly higher NFIB expression levels and higher NFIB protein levels than OSA cells (Fig. 3A; Supplementary Figs. S3 and S5A). A Matrigel Transwell invasion and migration assay demonstrated that the invasion and migration rates were inversely correlated with NFIB expression levels in the osteosarcoma cell lines (Fig. 3A and B). siRNA molecules against NFIB were used to deplete NFIB; all three osteosarcoma cell lines showed reduced NFIB mRNA and protein levels compared with control (siNEG)-treated cells (Fig. 3A; Supplementary Fig. S5B). After knockdown of NFIB, there was an increase in invasion and migration in all three osteosarcoma cells compared with the control (Fig. 3B). U2OS and HOS cells, with high endogenous NFIB expression, had a statistically significant increase in invasion and migration after NFIB expression levels were decreased using siRNA.

Figure 2. Relationship between the genotype of the metastasis-associated SNP rs7034162 and expression of NFIB and other nearby protein-encoding genes FREM1, ZDHHC21, and MPDZ in osteosarcoma (OS) cell lines and tumors. Significance is based on linear regression comparing the distribution of NFIB, FREM1, ZDHHC21, and MPDZ expression between the TT homozygous non-risk genotype (N = 34) of rs7034162 genotypes to the heterozygous risk AT (N = 10) and the homozygous risk AA genotypes (N = 2). These analyses were based on the publicly available genotype and expression data from 17 osteosarcoma cell lines and 29 osteosarcoma tumors (20).

Figure 3. NFIB expression correlates with invasion and migration potential of human osteosarcoma cells. A, NFIB expression in OSA, HOS, and U2OS cells was determined 48 hours after transfection with control siRNA (siNEG) or siRNA targeting NFIB (siNFIB). Graphs, relative expression compared with control-treated U2OS cells. The rs7034162 genotype is shown for the osteosarcoma cell lines: U2OS and HOS cells carry the homozygous non-risk allele (TT), and OSA cells carry the homozygous risk allele (AA). B, forty-eight hours after NFIB suppression, cells were evaluated for migratory potential. Cells that invaded through the Matrigel-coated membrane inserts were stained with crystal violet and air-dried, and bright field images were taken (inset photographs depict staining intensity differences). To quantify cell invasion/migration, the inserts were treated with 200 μL of methanol and the absorbance measured in triplicate at 560 nm. Invading cells are depicted relative to control-treated U2OS cells. Results of at least three experiments in triplicate are expressed as mean ± SD. The invasion and migration rate of each cell line was inversely correlated with NFIB expression levels.
NFIB Locus Associated with Osteosarcoma Metastasis

NFIB overexpression of (Fig. 3A; Supplementary Figs. S3 and 5A). In addition, expected because in colony formation with (Fig. 4D). The U2OS cells did not show a significant change of wound healing in HOS and OSA cells (data not shown).

growth factor binding protein 5 (IGFBP5) expression in (Fig. 3A and B). These findings were confirmed in wound-healing cell migration assays and filamentous actin staining assays; there was increased migration of osteosarcoma cells treated with siRNA against NFIB (siNFIB) compared with the siNEG control, and the filamentous actin staining (far right) shows increased podia formation in the cells treated with siNFIB. D, soft-agar colony-formation assay in OSA, HOS, and U2OS cell lines. The HOS and OSA cells showed a significant reduction in colony formation after overexpression of NFIB compared with the Luciferase (Luc) control.

knockdown (P = 0.009 and 0.015, respectively). In the OSA cell line, with low endogenous NFIB expression, there was a small nonsignificant increase in invasion and migration after NFIB suppression (P = 0.151; Fig. 3A and B). These findings were confirmed in wound-healing cell migration assays and filamentous actin staining assays; there was increased migration of osteosarcoma cells treated with siRNA against NFIB compared with the control (Fig. 4A-C). Staining for filamentous actin showed increased podia formation in all three osteosarcoma cell lines after NFIB suppression (shown for U2OS in Fig. 4C), typical of migrating cells. We blindly replicated our findings using a soft-agar colony-formation assay in HOS, OSA, and U2OS cells. Overexpression of NFIB resulted in a significant reduction in colony formation in HOS (P < 0.0001) and OSA (P < 0.0001) cells (Fig. 4D). The U2OS cells did not show a significant change in colony formation with NFIB overexpression. This was expected because NFIB expression is already high in U2OS (Fig. 3A; Supplementary Figs. S3 and 5A). In addition, overexpression of NFIB resulted in a significant reduction of wound healing in HOS and OSA cells (data not shown).

NFIB is a transcription factor that regulates insulin-like growth factor binding protein 5 (IGFBP5) expression in human osteoblasts, and IGFBP5 has been shown to inhibit tumor growth and metastasis of human osteosarcoma cells (21, 22). Therefore, we evaluated if there was a relationship between NFIB and IGFBP5 expression levels in osteosarcoma cell lines and tumors. We found a statistically significant direct correlation between NFIB and IGFBP5 expression levels (P = 1.17 × 10⁻⁸; Supplementary Fig. S6). The AA risk allele of rs7034162 was significantly associated with lower IGFBP5 and NFIB expression in osteosarcoma cell lines and tumors (Supplementary Fig. S3). The U2OS and HOS cells, both carrying the homozygous non-risk allele (rs7034162: TT), had higher NFIB and IGFBP5 expression levels than the OSA cells (carrying the homozygous risk allele, rs7034162: AA; Supplementary Figs. S3 and S7). In addition, NFIB siRNA suppression led to the downregulation of IGFBP5 in HOS and U2OS cells (Supplementary Fig. S7).

Mice with Primary Osteosarcomas and Metastases Harbor Inactivating Transposon Insertions in Nfib

A connection between Nfib and osteosarcoma was also identified in a sleeping beauty (SB) transposon mutagenesis system in mice (23). A significant proportion of primary osteosarcomas and metastases harbored inactivating transposon insertions
in *Nfib* in an analysis of insertions from all chromosomes, including mice on a *Trp53*-deficient and wild-type background (TAPDANCE analysis: \( P = 2.06 \times 10^{-6} \); Fig. 5A). We evaluated tumor tissues from these mice and found a reduced expression of both *Nfib* (\( P = 0.087 \)) and *Igfbp5* (\( P = 0.048 \)) in tumors with inactivating transposon insertions compared with tumors lacking insertions in *Nfib* (Fig. 5B and C). We used IHC staining for NFIB in sections of mouse tumor and adjacent bone to further evaluate tissue-specific changes in NFIB expression. There was a significantly reduced labeling intensity in the
osteosarcoma cells with an Nfib insertion compared with the adjacent normal osteoblasts (P = 0.0018; Fig. 5D and E). There was no NFIB labeling of any osteoclast cells. These mouse data confirm our human cell line findings and further suggest that a decreased expression of NFIB, leading to downregulation of IGFBP5, is important for metastatic ability.

DISCUSSION

This first multistage GWAS of osteosarcoma metastasis identified a common SNP, rs7034162, in NFIB at 9p24.1 associated with metastasis at diagnosis of osteosarcoma. We replicated the strong association identified in 541 European patients in an additional 394 patients from four case studies, including patients of European, African, and Brazilian ancestry. Our data showed that the risk SNP was associated with a decrease in NFIB expression, and, importantly, there was no association with expression of other nearby protein-encoding genes. Therefore, we focused on characterizing the role of NFIB in osteosarcoma metastatic potential.

NFIB is a member of the NFI gene family, which are site-specific DNA-binding proteins, also known as CAAT box transcription factors; they function in both viral DNA replication and the regulation of genes expressed in almost every organ and tissue (24). NFIB-mediated transcriptional activation or repression of specific gene promoters varies depending on cell type and gene promoter, resulting in the modulation of the expression of more than 100 diverse tissue-specific genes (24). NFIB gene rearrangements and fusions with either HMGA2 or MYB have been reported in lipoma, salivary gland, breast, and head and neck tumors (25–28). Changes in NFIB expression have been associated with specific miRNAs, and NFIB has been reported as a tumor suppressor in some cancers and an oncogene in others (29–33). This implication in both the inhibition and promotion of tumor development and progression in different cancers corroborates the diverse and distinct roles of NFIB in specific tissues and cancer sites. NFIB has not been previously implicated in osteosarcoma etiology or osteosarcoma metastasis. Here, we have shown that lowered NFIB leads to increased osteosarcoma cell line migration, proliferation, and colony formation, which suggests its involvement in osteosarcoma metastasis. The mouse Sb transposon mutagenesis screen independently identified Nfib with close to genome-wide significance to be a tumor-suppressor gene (23), and we further showed that osteosarcomas in mice with Nfib insertions had significantly reduced NFIB compared with the adjacent normal osteoblasts.

Because NFIB had previously been shown to directly regulate IGFBP5 expression in human osteoblasts (34), the relevant cell type, we hypothesized that the association of NFIB with osteosarcoma metastases may be mediated through IGFBP5. Interestingly, IGFBP5 has been previously shown to inhibit tumor growth and metastasis of human osteosarcoma cells (21, 22), growth of mouse osteosarcoma cells (35), and migration of human breast cancer cells (36). IGFBP5 is the most abundant IGFBP stored in bone, and it has been shown to regulate osteoblast proliferation and differentiation (37, 38). Thus, NFIB may modulate osteoblast proliferation and differentiation through the modulation of IGFBP5 expression. IGFBP5 can also both inhibit and enhance insulin-like growth factors (IGF), the critical regulators of bone metabolism (38). The IGF1–IGF1R axis is important for metastasis and migration and has been associated with osteosarcoma risk and survival (20, 39–46). We found a direct correlation between NFIB and IGFBP5 expression in osteosarcoma cell lines and tumors, and the rs7034162 risk allele of NFIB was also associated with reduced IGFBP5 expression. We hypothesize that the NFIB risk allele leads to lowered expression of NFIB and NFIB-mediated lower expression of IGFBP5, which results in less IGFBP5-mediated inhibition of IGF1 (47). This may then lead to an IGF1–IGF1R axis-mediated increase in proliferation, survival, and metastasis of osteosarcoma cells (Supplementary Fig. S8).

Metastatic osteosarcoma at diagnosis signifies a poor prognosis, and, in our study, 23% of patients had metastasis at diagnosis. There are few factors that have been consistently associated with risk of metastasis (14). We have identified a connection between germline genetics and osteosarcoma metastasis at diagnosis and further showed that the risk SNP likely has a role in osteosarcoma metastasis through an effect on NFIB expression levels. Our data suggest that germline genetic variation at rs7034162 in the NFIB gene contributes to susceptibility to metastasis at diagnosis in patients with osteosarcoma.

METHODS

Study Populations

A summary of the participating studies is given in Supplementary Table S1. Five hundred forty-one osteosarcoma cases of European ancestry were included in the discovery set from our previous GWAS, as described (5). Three hundred thirty-three unique osteosarcoma cases for replication of the discovery set findings were from the Genetic Epidemiology of Osteosarcoma study (COG AEPI05N2), the registry component of the Children's Oncology Group (COG), through the University of Minnesota, as previously described (6); the International Sarcoma Kindred Study, Melbourne, VIC (Australia); the Therapeutically Applicable Research to Generate Effective Treatments (TARGET) dataset (48); and from the Instituto de Oncologia Pediátrica GRAACC/UNIFESP and Universidade Federal de Sao Paulo (Brazil). TARGET data were used only for the analysis of the NFIB gene region. Cases from COG AEPI05N2 were restricted to those of European ancestry, as in the discovery set (5). Cases of African ancestry were identified in four studies in the discovery set (Supplementary Table S1) in our previous GWAS ancestry evaluation using a STRUC-TURE analysis and principal component analysis (PCA; ref. 5).

Patients were diagnosed in the individual hospitals and were prospectively followed up as part of their osteosarcoma treatment. Study centers provided data on patient and clinical variables that were harmonized between the studies, including age at diagnosis, gender, survival, relapse, follow-up, tumor subtype, tumor location, and the presence of histologically confirmed metastatic or nonmetastatic disease at diagnosis. All tumors were histologically confirmed. Not all patients had all variable data. In this study, we focused on the presence or absence of metastatic disease at diagnosis and did not exclude cases that experienced a later relapse. Participating subjects provided informed consent; studies were conducted in accordance with the Declaration of Helsinki and approved by local Institutional Review Boards.

Genotyping

Germline genomic DNA was extracted from either blood or buccal cells drawn from osteosarcoma case series using standard methods. Genotyping of all cases with GWAS data was conducted using the Illumina OmniExpress SNP microarray. Quality-control filtering
was performed as previously described (5). We included autosomal SNPs with a MAF of 25% in the discovery stage, SNPs with a >20% completion rate, and SNPs with no evidence of deviation from the Hardy–Weinberg proportion ($P > 1 \times 10^{-7}$). SNPs were excluded if they had abnormal heterozygosity values of <20% or >31%, expected duplicates, and abnormal X-chromosome heterozygosity. There were 29 duplicated cases with concordance rates of 99.96%. Genotypes for all subject pairs were also computed for close relationships (first and second degree) using the GLU qci.bds module (49) with an IBD0 threshold of 0.70; no first-degree relatives were identified in the cases scanned. SNPs (447,040) passed quality-control metrics for the 541 discovery set cases for final analysis. Only cases of >80% European ancestry were included in the discovery GWAS and the replication COG AEPI05N2 analysis based on STRUCTURE analysis (50) and PCA (51), as previously determined (5). PCA (52) results for individuals in the discovery stage are shown in Supplementary Fig. S9; there were no significant differences between populations ($P > 0.1$).

The top NFIB SNPs associated with metastasis at diagnosis from the discovery set were genotyped using Taqman assays in the International Sarcoma Kindred Study cases, and data from TARGET for osteosarcoma cases were used to examine NFIB SNPs (replication set 1). The top 100 SNPs associated with metastasis at diagnosis were chosen for follow-up in the COG AEPI05N2 cases (replication set 2) and in cases of African and Brazilian ancestry with GWAS data.

We validated the top imputed SNP, rs7034162, to confirm the imputed genotype using TaqMan assays in 89 cases with GWAS data from the discovery stage (16% of the total 541 cases). TaqMan assay (Life Technologies) validation and SNP genotyping were performed at the Cancer Genomics Research Facility. The imputation genotype error rate was 1.12% (1 case had the wrong genotype).

**Genomic Annotation**

For all imputed and genotyped SNPs with a $P$ value of less than $1 \times 10^{-5}$ ($N = 75$) in the NFIB locus, genomic annotation of SNP markers was conducted using the ENCODE (19) tool HaploReg (53), RefSeq (54), and LDlink (55).

**eQTL Analysis**

We performed eQTL-based analyses using publicly available genotype and expression data from 17 osteosarcoma cell lines (GSE36004) and 29 tumors (GSE33383; ref. 20). U2OS, HOS, and SJAS-1 (OSA) osteosarcoma cell lines were included as part of the 17 osteosarcoma cell lines with data and are highlighted in Supplementary Figs. S3 and S4.

**Cell Lines**

The human osteosarcoma cell lines U2OS, HOS, and SJAS-1 (OSA) were obtained from the ATCC for our in vitro studies (March 2014). Cell lines were tested for authentication (February 2015) with a panel of short tandem repeats (STR) using the Identifier Kit (Life Technologies) and compared with the ATCC STR Profile Databases. Osteosarcoma cells were maintained at 37°C in a humidified atmosphere with 5% CO₂ in RPMI-1640 (Cellgro) supplemented with 10% fetal calf serum (FCS; Gibco; Life Technologies). PB/SB-CAG-GFP-PGK-PURO-based transgene expression vectors were generated using the LR Clonase Gateway cloning system (Invitrogen) with cDNAs in Entry vectors with the previously described PB/SB-CAG-DEST-GFP-PGK-PURO, following the manufacturer’s instructions (56). NFIB was purchased from Open Biosystems in Gateway ready pENTR221 vector. NFIB and Luciferase overexpression cell lines were generated by electroporation using 2 μg PB7 transposase and 2 μg transposon plasmid with the NEON transfection system (Invitrogen), following the manufacturer’s protocol. After 2 days, cells were selected with 1 μg/mL puromycin to obtain stable cell lines.

**RNA Interference**

A mix of three siRNA molecules against NFIB and a negative control siRNA (ON-TARGETplus, D-001810-01-20; Thermo Scientific) were purchased from Dharmacon. Subconfluent osteosarcoma cells were transfected in 6-well plates using Lipofectamine RNAiMAX Transfection Reagent according to the manufacturer’s instructions (Life Technologies). Efficacy of the transfection protocol was tested using siGLO (D-001630-02-20) according to the manufacturer’s instructions (Thermo Scientific). NFIB sequences used were (i) AGGAUACU CUGAAAGACCUAUUUU, (ii) GCAAAAGACCCAAAACUAUUU, and (iii) ACUGAAGAAGCCUGUAAAUUU.

**RNA Isolation and Quantitative Real-Time PCR for NFIB and IGFBP5**

Total RNA was isolated using the RNeasy Mini Kit according to the manufacturer’s instructions (Qiagen). cDNA was synthesized from total RNA using the High-Capacity cDNA Reverse Transcription Kit, as described by the manufacturer’s protocol (Applied Biosystems; Life Technologies). Quantitative real-time PCR was performed using TaqMan Gene Expression Master Mix (Life Technologies). The expression of NFIB (Hs01029175_m1) and IGFBP5 (Hs00181213_m1) was normalized to the level of HPRT1 (4326321E) in the same sample, resulting in a ΔΔCT from which the 2$^{-ΔΔCT}$ value was derived and compared with control-treated U2OS cells. Results of at least three experiments in triplicate are expressed as mean ± SD.

**Cell Invasion/Migration by Transwell Matrigel Assay**

Subconfluent osteosarcoma cells were infected with the indicated siRNAs or controls. Forty-eight hours after infection, osteosarcoma cells were seeded in RPMI onto the basement membrane matrix (Corning BioCoat Matrigel Invasion Chambers). The Transwell was placed in a 24-well plate that contained 500 μL RPMI-1640 with 10% FCS as chemo attractant. After 24 hours, the noninvading cells and Matrigel matrix were gently removed with a cotton swab. Invasive cells located on the lower side of the chamber were stained with crystal violet, air-dried, and photographed. To quantify cell invasion/migration, the inserts were treated with 200 μL methanol, and the absorbance was measured in triplicate at 560 nm using the GloMax-Multi Detection System (Promega). Results of at least three experiments in triplicate are expressed as mean ± SD.

**Wound-Healing Cell Migration Assay**

Subconfluent osteosarcoma cells were infected with the indicated siRNAs or controls. Wounds were made on the monolayers using a pipette tip 48 hours after infection. Bright field images were taken at time 0 and 16, 24, and/or 30 hours. After the final image (at 16, 24, or 30 hours), the plates were washed, cells fixed with methanol, and stained with Crystal Violet (C008 Hardy Diagnostics).

**Filamentous Actin Staining**

Cells were seeded on coverslips and treated for 48 hours with siRNA as indicated. Six hours after wounds were made, cells were fixed with 4% paraformaldehyde in PBS for 10 minutes and blocked with 0.1% Triton X-100 and 1% BSA in PBS for 20 minutes. This was followed by incubation with Alexa Fluor 488 Phallodin (Life Technologies) in 0.1% Triton X-100 and 1% BSA in PBS for 20 minutes to stain filamentous actin. Finally, cells were mounted with ProLong Diamond Antifade Mountant with DAPI (Life Technologies). The LSM 700 confocal laser scanning microscope (Zeiss) was used for analysis.

**Soft-Agar Colony-Formation Assay**

Cells (1 × 10⁶) were seeded into soft agar in single wells of 6-well plates and allowed to incubate for 2 to 3 weeks. The resultant colonies were then stained with 0.5% crystal violet in buffered formalin for
1 hour. Each well was divided into 4 quadrants and photographed. Colonies were quantified using ImageJ software using a standard colony quantification macro.

**Osteosarcoma SB Screen**

The results of the complete SB screen have been previously published (23). Briefly, mice were generated to target SB mutations only in osteoblast cells in wild-type or Trp53-deficient mice. Genes found to be significantly mutated by transposon mutagenesis were then identified. gCIS analysis of T2/Onc insertion sites was performed using gene-centric software, as previously described (57). RNA was extracted from SB-mutagenized tumors using the RNA mini-prep kit (Invitrogen) and reverse transcribed into cDNA using the Superscript III reverse transcriptase kit (Invitrogen), following the manufacturer’s instructions. Eleven primary tumors with insertions in NFIB and 11 primary tumors without insertions in NFIB were analyzed. Quantitative RT-PCR was performed using FastStart Universal SYBR Green Master (Roche) mix (Roche) using an Eppendorf Realplex2 Mastercycler EP Gradient S. Primer sequences were as follows: NFIB forward: TCTCCCCCAAGTTGAGCCTTT, NFIB reverse: GGTAGATGCT GGGTCGGAAAAA; IGBFBP forward GTACCTGCAACCTGTGAC, IGBFBP reverse GTCTTCTGGTCATTGTCCA; and normalized to GAPDH, forward GTGTTTCTACCCCTGAATGTT and reverse GAGACAACCTGGTCCTCAGTGT.

**Western Blot and IHC Staining**

Tissues were fixed in 10% buffered formalin and embedded in paraffin blocks. Sections were cut at 5-μm thickness and rehydrated through a series of graded ethanolss. Slides underwent antigen retrieval by boiling for 30 minutes in antigen unmasking solution (Vector Laboratories Inc.). Endogenous peroxidases were quenched with 3% hydrogen peroxide solution for 20 minutes. For antibody staining, 10% goat serum in 1× PBST was used for blocking and antibody incubations. Primary NFIB antibody was used at 1:250 (HPA003956; Sigma-Aldrich). Following a series of washes, slides were incubated with biotinylated goat anti-rabbit secondary antibody (1:250; Vector Laboratories Inc.) slides were washed, incubated with the vectastain ABC Kit (Vector Laboratories Inc.) for 30 minutes at room temperature, washed again, and stained using a peroxidase substrate kit DAB (Vector Laboratories Inc.). Finally, slides were counterstained with hematoxylin, dehydrated, cleared with xylene, and mounted with Permount (Fisher Scientific). Immunoblot analysis was performed on whole-cell lysates harvested using RIPA buffer (Sigma-Aldrich) supplemented with phosphatase inhibitors and protease inhibitors (Sigma-Aldrich). Lysates were separated by SDS-PAGE, transferred to a polyvinylidene difluoride membrane (Invitrogen), incubated in 5% BSA in tris-buffered saline and tween (TBST) to block nonspecific binding, and probed with appropriate antibodies. Proteins were detected using the LI-COR Odyssey Fc instrument using a chemiluminescent horseradish peroxidase substrate (Advansta). NFIB antibody used for IHC above was also utilized for immunoblot analysis.

Staining intensity within 16 tumors and adjacent, normal osteoblasts (internal positive control) was assessed by the semiquantitative H-score (58), which was determined by summing the products of the percent of cells with weak (1 × %), moderate (2 × %), and strong (3 × %) nuclear staining, with a maximum possible score of 300.

**Statistical Analyses**

Cases with confirmed metastasis at diagnosis were compared with cases without metastasis at diagnosis. In the discovery and replication stages, risk of metastasis at diagnosis was estimated using logistic regression to calculate the OR and 95% CIs per copy of the minor allele using a multiplicative (log-additive) genetic model with 1 degree of freedom. Logistic regression models were adjusted for study center; age and gender were not significantly associated with metastasis, and thus were removed from the model. Because there were no significant eigenvectors from the PCA of osteosarcoma cases (P > 0.1), they were not included in the model. The estimated inflation factor, λ, of the test statistic for the discovery set was 0.9107 (Supplementary Fig. S10). Logistic regression models for the evaluation of SNP associations comparing the European discovery stage cases with the European controls from our previous GWAS (5) were adjusted for gender and the two significant eigenvectors from the PCA of cases and controls. We had 91.9% power to detect an OR of 2.60 with 122 cases with metastasis and 419 cases without metastasis at an alpha of 0.05 and allele frequency of 10%.

An inverse-variance fixed-effect meta-analysis was applied to combine the results of the discovery set and replication sets. Between-stages heterogeneity was quantified with F (59). We did not observe significant heterogeneity for the top reported SNPs: rs2890982, EUR meta-analysis of discovery and replication sets: F = 0.0, P heterogeneity = 0.39; rs7044162, all ancestry meta-analysis of discovery and replication sets: F = 17.9%, P heterogeneity = 0.30.

We imputed additional SNPs within 1 Mb on either side of the implicated SNPs using IMPUTE2 software and the reference data from both the 1000 Genomes project (March 2012 release; ref. 18) and the DCEG Imputation Reference Set version 1 (60). SNPTEST was used to analyze the posterior SNP dosages from IMPUTE2, adjusting for study center, as described above (61).

A time-to-event analysis for overall survival was assessed in the osteosarcoma cases with these data using Cox proportional hazards regression to estimate HRs and 95% CIs adjusted for study center, ancestry, gender, and age. Not all cases had follow-up and survival data; only 522 European ancestry cases included in the discovery stage, 57 African ancestry cases in replication 3, and 105 Brazilian ancestry cases in replication 4. Overall survival time was estimated as the time from the date of diagnosis until the date of death or the last known alive date; patients were censored at the last date to be alive or if lost to follow-up. Kaplan–Meier survival curves were created, and P values were estimated with a Mantel–Cox log-rank test.

**Disclosure of Potential Conflicts of Interest**

D. Largaespada reports receiving a commercial research grant from Genentech; has received speakers bureau honoraria from Cell Signal­

**Authors’ Contributions**

**Conception and design:** L. Mirabello, L.G. Spector, D. Largaespada, J.S. Wunder, S.J. Chanock, S.A. Savage

**Development of methodology:** L. Mirabello, R. Koster, O.A. Panagiotou, S. Wacholder, J.F. Boland


**Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis):** L. Mirabello, R. Koster, B.S. Moriarty, L.G. Spector, J. Gary, M.J. Machiela, N. Pankratz, O.A. Panagiotou, D. Largaespada, J.S. Wunder, D.M. Thomas, S. Davis, G.M. Otto, S. Wacholder, S.A. Savage

**Writing, review, and/or revision of the manuscript:** L. Mirabello, R. Koster, L.G. Spector, J. Gary, M.J. Machiela, N. Pankratz,
REFERENCES


NFIB Locus Associated with Osteosarcoma Metastasis


A Genome-Wide Scan Identifies Variants in NFIB Associated with Metastasis in Patients with Osteosarcoma

Lisa Mirabello, Roelof Koster, Branden S. Moriarity, et al.

*Cancer Discovery* Published OnlineFirst June 17, 2015.

<table>
<thead>
<tr>
<th>Updated version</th>
<th>Access the most recent version of this article at: doi:10.1158/2159-8290.CD-15-0125</th>
</tr>
</thead>
<tbody>
<tr>
<td>Supplementary Material</td>
<td>Access the most recent supplemental material at: <a href="http://cancerdiscovery.aacrjournals.org/content/suppl/2015/06/17/2159-8290.CD-15-0125.DC1">http://cancerdiscovery.aacrjournals.org/content/suppl/2015/06/17/2159-8290.CD-15-0125.DC1</a></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>E-mail alerts</th>
<th>Sign up to receive free email-alerts related to this article or journal.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reprints and Subscriptions</td>
<td>To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at <a href="mailto:pubs@aacr.org">pubs@aacr.org</a>.</td>
</tr>
<tr>
<td>Permissions</td>
<td>To request permission to re-use all or part of this article, contact the AACR Publications Department at <a href="mailto:permissions@aacr.org">permissions@aacr.org</a>.</td>
</tr>
</tbody>
</table>