Original Article

Sarcomas in *TP53* Germline Mutation Carriers

A Review of the IARC *TP53* Database

Simona Ognjanovic, PhD[1,2]; Magali Olivier, PhD[3]; Tracy L. Bergemann, PhD[4]; and Pierre Hainaut, PhD[3]

BACKGROUND: Sarcoma is the index diagnosis of Li-Fraumeni syndrome (LFS), a familial predisposition to cancer that also includes brain cancer, breast cancer, and adrenal cortical carcinoma. Germline mutations in the *TP53* gene are detected in approximately 80% of families that fulfill LFS criteria and in 15% to 25% of families that fulfill criteria for Li-Fraumeni-like syndrome (LFS), a group of related syndromes with broader clinical criteria. **METHODS:** The authors of this report used the International Agency for Research on Cancer *TP53* database to analyze the types, age at onset and mutation patterns of sarcoma in *TP53* mutation carriers. Those data were compared with sarcoma types in the general population of Caucasians using data from the National Cancer Institute's Surveillance, Epidemiology, and End Results (SEER) Program. **RESULTS:** Overall, sarcomas represented 25% of tumors in *TP53* mutation carriers, and 95.6% occurred before age 50 years compared with 38.3% before age 50 years in the SEER data set. Sarcomas were more likely to be rhabdomyosarcoma in carriers aged <5 years (odds ratio [OR], 11.6; 95% confidence interval [CI], 6.1-21.9) and osteosarcoma in carriers at any age (aged <20 years: OR, 1.41; 95% CI, 1.02-1.94; age >20 years: OR, 4.61; 95% CI, 2.72-7.83). Early sarcoma (at age <20 years) was associated with missense mutations in exons encoding the DNA-binding domain of p53 protein. Conversely, p53 null mutations (frameshift, splice sites, nonsense) and mutations outside the DNA-binding domain were associated with leiomyosarcoma (OR, 10.1; 95% CI, 3.4-29.9), a type of sarcoma that occurred after age 20 years. **CONCLUSIONS:** The current results further demonstrated genotype-phenotype correlations and age-dependent variations in sarcoma types in carriers of germline *TP53* mutations. ***Cancer*** **2012;118:1387–96.** © *2011 American Cancer Society.*

KEYWORDS: germline *TP53* mutations, sarcoma, Li-Fraumeni syndrome, IARC *TP53* Database, Surveillance, Epidemiology, End Results.

INTRODUCTION

Numerous sarcoma types collectively comprise less than 1% of all cancers and are rare in adults.[1] However, they represent 12.6% of all malignancies in children and adolescents aged <19 years and constitute the third most common solid tumor in this age group.[2] These mesenchymal tumors are morphologically and topologically diverse.[3] They are divided into 2 broad groups, soft tissue sarcoma (STS) and primary bone sarcoma and arise in connective tissues and bones, respectively. Further histopathologic classification distinguishes rhabdomyosarcoma, osteosarcoma, and Ewing sarcoma (together representing >90% of primary sarcoma diagnosed in children and adolescents)[4] as well as malignant fibrous histiocytoma, liposarcoma, leiomyosarcoma, gastrointestinal stromal tumor (GIST), synovial sarcoma, fibrosarcoma, and chondrosarcoma, which represent the most common types of sarcoma in adults.

Genetic susceptibility is a risk factor for several types of sarcoma. In particular, the occurrence of familial sarcoma is the basis of the clinical definition of the Li-Fraumeni syndrome (LFS), an inherited predisposition syndrome characterized by a cluster of early onset cancers.[5,6] To date, *TP53* is the only gene that consistently has been associated with this syndrome.[7] The initial clinical definition of LFS was based on the occurrence of sarcoma in a proband before age 45 years with a first-degree relative with any cancer by age 45 years plus another first-degree or second-degree relative with cancer

Corresponding author: Simona Ognjanovic, PhD, Department of Pediatrics, MMC 715, 420 Delaware Street SE, Minneapolis MN 55455; Fax: (612) 624-7147; ognja001@umn.edu

[1]Division of Pediatric Epidemiology and Clinical Research, University of Minnesota, Minneapolis, Minnesota; [2]Masonic Cancer Center, University of Minnesota, Minneapolis, Minnesota; [3]Molecular Carcinogenesis Group, International Agency for Research on Cancer, Lyon, France; [4]Division of Biostatistics, University of Minnesota, Minneapolis, Minnesota

DOI: 10.1002/cncr.26390, **Received:** February 25, 2011; **Revised:** June 6, 2011; **Accepted:** June 7, 2011, **Published online** August 11, 2011 in Wiley Online Library (wileyonlinelibrary.com)

at age <45 years or sarcoma at any age.[5,6] Over the past 15 years, several modified clinical criteria have been proposed (for review, see Olivier et al[8]). The current definition proposes a set of clinical criteria for referring a patient for *TP53* genetic testing.[9] These criteria (often referred to as "Chompret criteria") identify a set of cancers that constitute the "narrow LFS spectrum" (STS, osteosarcoma, brain tumor, premenopausal breast cancer, adrenocortical carcinoma, leukemia, and bronchoalveolar adenocarcinoma). Among patients who are referred for testing using these criteria, 29% have *TP53* germline mutations identified.[10] Although detailed, population-based data are lacking, it is believed that de novo *TP53* mutations occur in approximately 1 of every 5000 births in the general Western population.[11] A more recent study, however, suggests that the frequency of germline *TP53* mutations in the general population probably is likely 1 in 20,000.[12] In Southern Brazil, a founder *TP53* mutation (a protein variant germline arginine-to-histidine substitution at codon 337 [p.R337H]) was present in approximately 1 in 300 individuals tested,[13] leading to a locally high prevalence of the syndrome.

The tumor suppressor gene *TP53* is the most commonly mutated gene in human cancer, and the prevalence of somatic (mostly missense) *TP53* mutations in sporadic cancer ranging from 10% to 60%.[14] The p53 protein is inducible in response to several forms of genotoxic and nongenotoxic stress. Once activated, it controls biologic processes, including oxidative metabolism, cell cycle, aspects of DNA repair, senescence, and apoptosis. Recent evidence suggests that p53 may regulate the maintenance of pools of stem cells and progenitor cells (for review, see Hafsi and Hainaut[15]). A missense mutation inactivates p53 and, thus, disrupts an integrated set of antiproliferative and growth-suppressive responses. In addition, some mutations can cause "gain of function" whereby mutant p53 protein assumes an oncogenic role and promotes tumor development.[16–18]

To date, a total of 532 families or individuals with germline *TP53* mutations have been reported in the peer-reviewed literature.[19] Information on these mutations, on tumor pathologies and localization, and on age of diagnosis in mutation carriers is compiled into the *TP53* database maintained at the International Agency for Research on Cancer (IARC). For the current review, we analyzed data available in the IARC *TP53* database on sarcoma in carriers of inherited *TP53* mutations. We used data from the Surveillance, Epidemiology, and End Results (SEER) Program to compare the age-related occurrence of different types of sarcoma in *TP53* mutation carriers with a group that represented the general population.

MATERIALS AND METHODS

Germline TP53 *Mutation and Inherited Sarcoma Data Sets*

The data set of germline *TP53* mutation from R14 version of the IARC *TP53* database (http://www-p53.iarc.fr/Germline.html [last accessed October 2010]) was used.[19] This data set was compiled from information on *TP53* germline mutation carriers and on families with LFS and related syndromes affected by cancer published in peer-reviewed articles between 1990 and 2009.[19] The data set included 531 independent families or individuals with germline mutations and compiled data on 2080 cancers that occurred in *TP53* germline mutation carriers. Of those families, 491 (92.5%) were from countries populated mainly by Caucasians, and the remaining families were from Asia (N = 35), Micronesia (N = 1), and India (N = 4). Analyses included only individuals with confirmed germline *TP53* mutation regardless of their familial history of cancer and for whom the age at diagnosis was available. Tumors classified either as "connective, subcutaneous, and other soft tissues"; "bones, joints, and articular cartilage of other and unspecified sites"; or "bones, joints and articular cartilage of limbs," irrespective of tumor localization and histology, were analyzed. All database annotations describing germline mutations, including type, position within the *TP53* gene, and predicted structural and functional effects, were used in the analyses.

Caucasian Population Sarcoma Data Set

The SEER Program data[20] were used to assess the incidence of sarcoma in Caucasians in the United States between 1973 and 2007 for the sarcoma types corresponding to those identified among *TP53* mutation carriers in the IARC database. Caucasian ethnicity was selected, because >90% of the *TP53* carriers analyzed here are expected to be of Caucasian origin (see above). SEER data were taken as representative for the incidence of sarcoma in a Caucasian population, because the United States contributed more sarcoma cases than any other country in the IARC *TP53* database. In addition, counts for different sarcoma types were used to calculate their relative proportions. The SEER Program actively collected information on demographics, tumor site and morphology, stage at diagnosis, treatment, and vital status during this period from 9 registries encompassing 5 states (Connecticut, Hawaii, Iowa, New Mexico, and Utah) and 4 metropolitan areas (Detroit, San Francisco-Oakland, Seattle-Puget Sound, and Atlanta). These 9 SEER

registries represent approximately 9% of the total US population[21] and have an estimated case ascertainment rate of 98%.[22] Primary cases of sarcoma classified by International Classification of Disease for Oncology, 3rd edition[23] and morphology codes present in the International Classification of Childhood Cancer, 3rd edition[24] were included. The following morphologies were analyzed: osteosarcoma (codes 9180/3, 9181/3, 9182/3, 9183/3, 9184/3, 9185/3, 9186/3, 9192/3, 9193/3, 9194/3, 9200/3); rhabdomyosarcoma (codes 8900/3, 8901/3, 8902/3, 8910/3, 8912/3, 8920/3, 8921/3); sarcoma, not otherwise specified (NOS) (code 8800/3); leiomyosarcoma (codes 8890/3 and 8891/3); and liposarcoma (codes 8850/3, 8851/3, 8852/3, 8853/3, 8854/3, 8855/3, 8858/3, 8860/3). In addition, morphologies classified as "other sarcomas" represented a group of sarcoma types rarely observed in carriers. These included spindle cell sarcoma (code 8801/3); fibrosarcoma, NOS (code 8810/3); chondrosarcoma, NOS (code 9220/3); stromal sarcoma (code 8930/3); synovial sarcoma, biphasic (code 9043/3); gliosarcoma (code 9442/3); and dermatofibrosarcoma, NOS (code 8832/3). The proportion of each of these sarcoma types was calculated for children and adolescents (ages birth to 19 years) and for adults (aged $\geq$20 years).

Statistical Methods

For analyses of data with a binary endpoint of *TP53* carriers versus sporadic cancers, we used generalized estimating equations (GEE) and the family units were treated as clusters to account for correlation within *TP53* families.[25] The sporadic cancers were treated as independent units. We fitted equations using 2 different correlation strategies and calculated 2 corresponding *P* values for the risk estimates. The first strategy assumed working independence, which is the default setting for most GEE cluster analyses. The resulting *P* values took into account whether the LFS participants were related. The second strategy assumed an unstructured correlation matrix in which a separate correlation value was estimated for each degree of kinship within a family. When the kinship relationship was unknown for an individual, that individual was treated as an additional proband within the family cluster. Because of the incompleteness of kinship data and very similar results for both sets of *P* values, we present only the *P* values that assumed working independence (which assumed that *TP53* carriers were related but not the degree of their relationships).

We also evaluated possible associations using a general logistic regression method (GLM) in which the degree of relationship between carriers is ignored. Because the GEE models had a heavy computational burden for scenarios evaluating both patients with sporadic sarcoma and *TP53* carriers, we used the GEE model only for associations for which the GLM *P* value was significant. Because GLM *P* values are smaller than GEE *P* values because they ignore correlation within families, it was reasonable to assume that GEE-based tests would not be statistically significant for associations for which GLM-based tests were not significant.

RESULTS

Types of Sarcoma in TP53 Mutation Carriers

In total, 406 sarcomas were retrieved from the IARC *TP53* database, representing 17.4% of all cancers in the database and 36.8% of all cancers (220 of 598 cancers) in patients aged <20 years. This proportion is consistent with the finding that early onset sarcoma is the basis of LFS clinical definition and that >40% of tumors in the analyzed data set are from families that match the strict LFS definition. The main histologic definitions were osteosarcoma (164 tumors; 40.4%), sarcoma NOS (69 tumors; 17%), rhabdomyosarcoma (67 tumors; 16.5%), leiomyosarcoma (37 tumors; 9.1%), liposarcoma (20 tumors; 4.9%), and histiosarcoma (20 tumors; 4.9%). The remaining 29 tumors (7.1%) accounted for 16 different histologic descriptions. Further analysis included only patients who had a confirmed *TP53* mutation and known age at diagnosis. These included 236 sarcomas (Table 1), 18 of which were grouped together into an "other sarcomas" group that included up to 5 tumors per histology. Although there were more females with sarcoma among *TP53* carriers (56.4%) than in the SEER data set (45.4%), this difference was not significant ($P = .20$), nor was there any difference in sex distribution by age ($P = .73$). However, female *TP53* carriers had a marginally increased proportion of osteosarcoma compared with the SEER data set (odds ratio [OR], 1.69; 95% confidence interval [CI], 1.01-2.80; $P = .04$). This sex difference was confined to children/adolescents (Table 1). It is noteworthy that GISTs (the most common type of sarcoma in Caucasians), desmoids tumors, Ewing sarcomas, and angiosarcomas were not represented among the types of sarcomas detected in *TP53* mutation carriers.

Age Distribution of Sarcoma

The distribution of ages at diagnosis for different types of sarcoma in *TP53* mutation carriers and in the SEER data

Table 1. Comparison of Sarcoma Characteristics Reported for Tumor Protein p53 Germline Mutation Carriers in the IARC *TP53* Database With Those Reported for Caucasians in the Surveillance, Epidemiology, and End Results Database (Representing Sporadic Sarcomas)

	Incidence[a]	Mean±SD Age of Onset, y		No. of All Sarcomas (%)		No. of Males (%)[b]	
Tumor Type	Sporadic	*TP53* Carriers, N = 236	Sporadic, N = 34,671	*TP53* Carriers, N = 236	Sporadic, N = 34,671	*TP53* Carriers, N = 103	Sporadic, N = 16,921
Age ≤19 y							
Rhabdomyosarcoma	4.5	4.3±4.4	8.0±2.9	53 (22.4)	1362 (3.9)	28 (57.1)	794 (58.3)
Osteosarcoma	4.4	13.1±3.7	13.3±4.4	73 (30.9)	1401 (4)	32 (45.7)	785 (56)
Sarcoma, NOS	0.7[c]	8.9±6.8	10.9±3.6	17 (7.2)	173 (0.5)	8 (53.3)	81 (46.8)
Liposarcoma	0.2[c]	8.7±4.3	14.8±4.2	4 (1.7)[d]	79 (0.2)	1 (33.3)[d]	36 (45.6)
Leiomyosarcoma	0.3[c]	—	13.7±4.5	0	81 (0.2)	0	39 (48.1)
Other sarcomas		8.4±6.6	13.1±4.4	11 (4.7)	598 (1.7)	4 (36.4)[d]	332 (55.5)
Total				158 (67)	3694 (11.9)	73 (46.2)	2067 (56)
Age ≥20 y							
Rhabdomyosarcoma	1.4	27	53.5±12.9	1 (0.4)[d]	1024 (3)	1 (100)[d]	570 (55.7)
Osteosarcoma	2.5	27.7±6.8	49.4±22.0	18 (7.6)	1891 (5.5)	9 (56.2)	1033 (54.6)
Sarcoma, NOS	5.8	38.3±13.7	64.1±15.7	23 (9.7)	4235 (12.2)	9 (39.1)	2099 (49.6)
Liposarcoma	8.9	31.1±7.7	60.1±14.6	11 (4.7)	6813 (19.7)	4 (36.4)[d]	4058 (59.6)
Leiomyosarcoma	16	44.2±14.0	60.6±14.7	18 (7.6)	10,947 (31.6)	3 (17.6)[d]	4014 (36.7)
Other sarcomas		33.1±10.3	53.1±12.8	7 (3)	6067 (17.5)	3 (42.8)[d]	3080 (50.8)
Total				78 (33)	30,977 (88.1)	30 (38.5)	14,854 (48)

Abbreviations: NOS, not otherwise specified; SD, standard deviation; *TP53*, the tumor protein p53 gene.
[a] Cases per million per year. Fields with <1 represent unstable estimates.
[b] The percentage of females can be calculated as: 100% − % of males.
[c] Because of to low counts, these estimates are unstable.
[d] Fields with <5 cases.

set is provided Table 1. In germline *TP53* mutation carriers, 67% of sarcomas occurred in children and adolescents (aged <20 years), compared with 11.9% of sarcomas in the SEER data set. Thus, sarcoma in adults was less likely to be inherited through *TP53* mutation than as sporadic sarcoma (OR, 0.06; 95% CI, 0.04-0.08; $P < .00001$). When the age distribution was examined for individual sarcoma types, the tumors in *TP53* mutation carriers were significantly more likely to be osteosarcoma than tumors in the SEER data set ($P = .0002$). Compared with other types of sarcoma, the odds that a diagnosed sarcoma would be an osteosarcoma was higher in both children and adolescents (OR, 1.41; 95% CI, 1.02-1.94; $P = .04$) and in adults (OR, 4.61; 95%CI, 2.72-7.83; $P = .0002$) with germline *TP53* mutations. In contrast, an increased risk of rhabdomyosarcoma compared with all other sarcoma types was confined only to mutation carriers aged <5 years (OR, 11.6; 95% CI, 6.1-21.9; $P < .00001$), and an even larger risk was observed for carriers aged <3 years (OR, 16.7; 95% CI, 9.4-29.7; $P < .00001$).

Figure 1 compares proportions by age group in patients with sarcoma among *TP53* mutation carriers and in the SEER data set. This comparison indicates the tendency for each sarcoma type to occur at a younger age in *TP53* mutation carriers than in the SEER data set. It is striking that only 4.4% of sarcomas diagnosed in *TP53* mutation carriers occurred in patients aged >50 years in *TP53* carriers in contrast to 62.7% in the SEER data set. In *TP53* carriers, rhabdomyosarcoma was diagnosed almost exclusively among children and adolescents (28 of 29 occurred in patients aged <20 years; 96.6%). In contrast, in the SEER data set, rhabdomyosarcoma was diagnosed in all age groups (1362 of 2386 occurred in patients aged <20 years; 57.1%). Both in carriers and in the SEER data set, leiomyosarcoma and liposarcoma were diagnosed almost exclusively in adults. However, these tumors made for a smaller proportion of all sarcoma types among *TP53* mutation carriers than in the SEER data set (12.3% vs 51.7%, respectively).

Mutation Types and Patterns

Table 2 describes the types, patterns, and hotspot codon distribution of germline *TP53* mutations associated with sarcoma. The types of mutations (missense mutations, 72.8%; frameshift and splice mutations, 16.1%) and the patterns of single base substitutions were similar in carriers

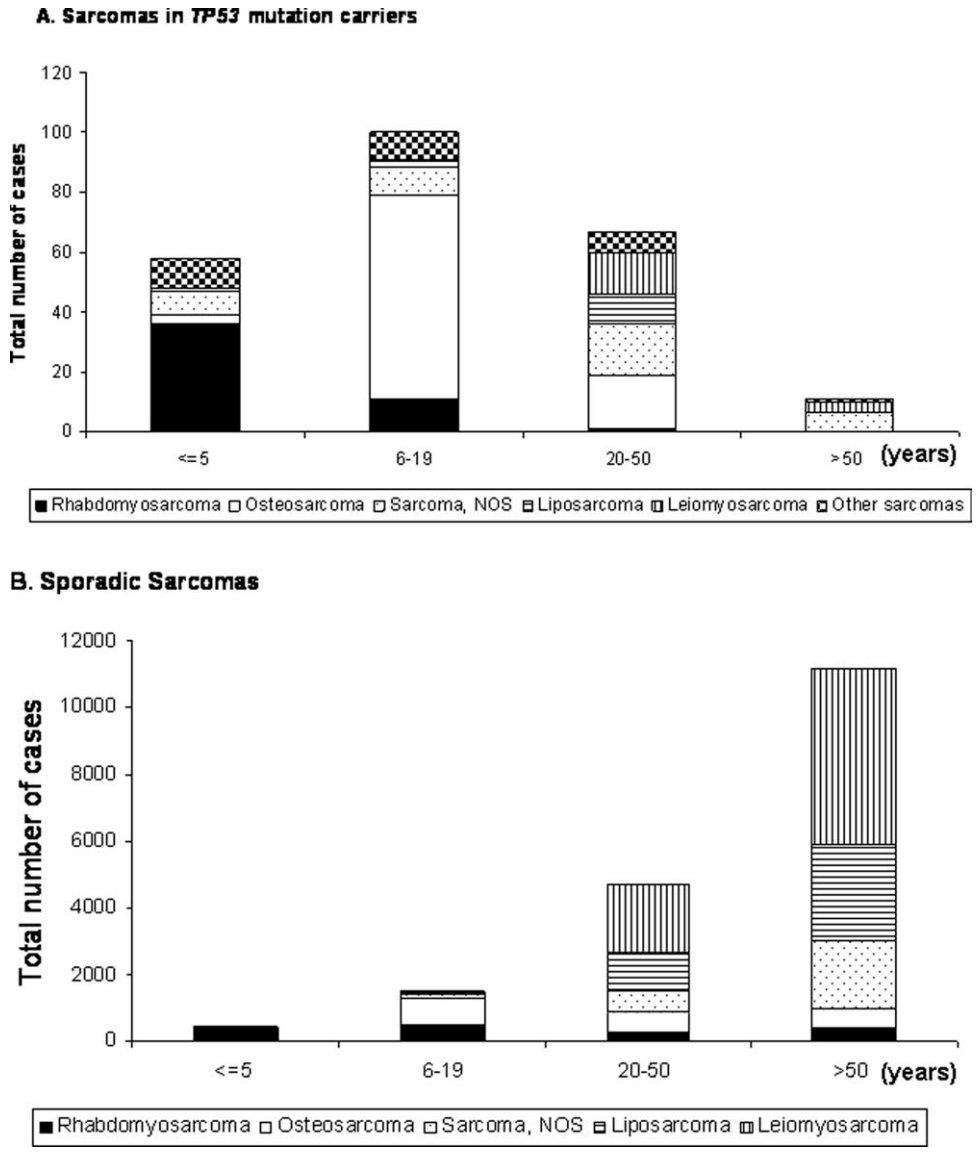

Figure 1. Distribution of the different types of sarcoma is illustrated by age group in (A) tumor protein p53 (*TP53*) germline mutation carriers compared with (B) patients with sporadic sarcomas from the Surveillance, Epidemiology, and End Results database.

with sarcoma and with other cancers (approximately half were guanine:cytosine-to-adenine:thymine substitutions at CpG sites). There was no difference in the age at diagnosis according to mutation type or patterns. In terms of codon distribution (Table 2, Fig. 2), the each of the common "mutation hotspot" codons 273, 248, 282, 175, and 220 was mutated in >5% of carriers. Mutations at codons 175 and 220 were associated with the youngest age of onset (ages 10.9 ± 10.7 years and 11.6 ± 6.6 years, respectively). Codon 273 represented the most common hotspot overall and was particularly frequent in patients with rhabdomyosarcoma. Greater than 20% of all mutations in patients with rhabdomyosarcoma were at codon 273, and mutation at this codon tended toward an association with an increased risk of rhabdomyosarcoma compared with other sarcoma types (OR, 2.5; 95% CI, 0.9-7.2). Figure 2 depicts the codon distribution of mutations associated with 4 different types of sarcoma and of nonsarcoma tumors in *TP53* mutation carriers. Liposarcoma shared with osteosarcoma a high prevalence of codon 245 and 282 mutations. Mutations at codon 337 or 344, which affect residues located outside the DNA binding domain of p53, were associated with a significantly increased risk of leiomyosarcoma (OR, 10.1; 95% CI,

Table 2. Comparison Between Germline Tumor Protein p53 Mutations Associated With Sarcoma and Those Associated With Other Cancers in Carriers Reported in the IARC *TP53* Database

	Sarcomas		All Cancers	
Mutation Type	**No. of Associated *TP53* Mutations (%)**	**Mean±SD Age at Onset, y**	**No. of Associated *TP53* Mutations (%)**	**Mean±SD Age at Onset, y**
All	236		951	
Missense	172 (72.8)	17.4±15.1	721 (75.8)	24.9±19.6
FS and splice	38 (16.1)	21.1±16.8	132 (13.9)	27.6±15.4
Nonsense	18 (7.6)	17.2±12.7	64 (6.7)	21.4±12.5
Other	8 (3.4)	16.2±14.8	34 (3.5)	22±16.8
Point mutations	202		843	
G:C→A:T at CpG	103 (50.1)	19±17	494 (58.6)	24.3±19.8
A:T→G:C	28 (13.9)	16.9±15.9	88 (10.4)	27.3±16.9
G:C→A:T at non-CpG	28 (13.9)	17.5±14.0	88 (10.4)	27.4±19.3
G:C→T:A	18 (8.9)	16.9±11.8	72 (8.5)	25.7±18.7
G:C→C:G	14 (6.9)	12.8±10.6	41 (4.8)	21.2±16.6
A:T→T:A	10 (4.9)	21.3±3.9	30 (3.5)	27.4±17.8
A:T→C:G	1 (0.5)	11	30 (3.5)	22.8±16.1
Hotspot codons/ missense mutations	172		721	
273	25 (14.5)	15±12.1	71 (9.8)	23.8±16.7
248	23 (13.4)	18.5±17.8	88 (12.2)	22.7±17.1
282	16 (9.3)	13.2±5.7	30 (4.2)	12.2±9.9
175	14 (8.1)	10.9±10.7	47 (6.5)	22.5±15.9
220	10 (5.8)	11.6±6.6	28 (3.9)	21.9±12.8
275	8 (4.6)	17.9±12.4	17 (2.3)	21.1±16.1
245	7 (4)	19.4±12.4	28 (3.9)	30.4±17.7
344	5 (2.9)	38.8±5.5	5 (0.7)	38.8±5.5
133	5 (2.9)	13.8±15.5	33 (4.5)	32.1±13.7
337	5 (2.9)	34±22.8	103 (14.3)	13.9±18.4

Abbreviations: A, adenine; C, cytosine; CpG, cytosine and guanine with a phosphodiester bond; FS, frameshift; G, guanine; SD, standard deviation; T, thymine; *TP53*, the tumor protein p53 gene.

3.4-29.9; $P = .00003$). Comparison between carriers with sarcoma and carriers with other cancers suggested that mutations at codons 220, 275, 282, and 334, together, were represented more in carriers with sarcoma than in carriers with other cancers (OR, 3.1; 95% CI, 1.5-6.4; $P = .001$). These other cancers included breast cancer (N = 252), brain tumors (N = 115), adrenal cortical carcinoma (N = 131), and genitourinary tract cancers (N = 53); whereas lung cancer, colon cancer, and leukemia all amounted to 20 to 30 patients each, and a few other cancer types were represented by smaller numbers (each with N < 10).

Next, we examined whether there were correlations between the expected structural and functional impact of the mutation and specific types of sarcoma (Table 3). Mutations were broadly classified in 2 groups: missense mutations predicting the expression of an altered, mutant protein and other types of mutations (frameshift, splice site, or nonsense mutations) predicting loss of wild-type p53 protein expression. Missense mutations were categorized further as mutations affecting residues located in DNA binding motifs (DBMs), representing specific structural elements within the DNA binding domain of the p53 protein (see Table 3, footnote), or mutations affecting residues located outside these motifs. We compared the actual (observed) distribution of the mutations among these groups in different types of sarcoma with the expected distribution assuming that mutations would be distributed equally among the different types. This assumption is based on the observation of a variety of different tumor types within families that carry the same type of *TP53* mutation. We observed that rhabdomyosarcoma and osteosarcoma had a higher than expected proportion of missense mutations in DBMs and a lower proportion of missense mutations outside DBMs. In contrast, leiomyosarcoma had fewer missense mutations in DBMs than expected. In sarcoma (NOS), liposarcoma, and leiomyosarcoma, frameshift, splice, and nonsense mutations were more frequent than expected. Thus, the mutations that predict the absence of wild-type p53 expression tended to be associated more with late-onset types of

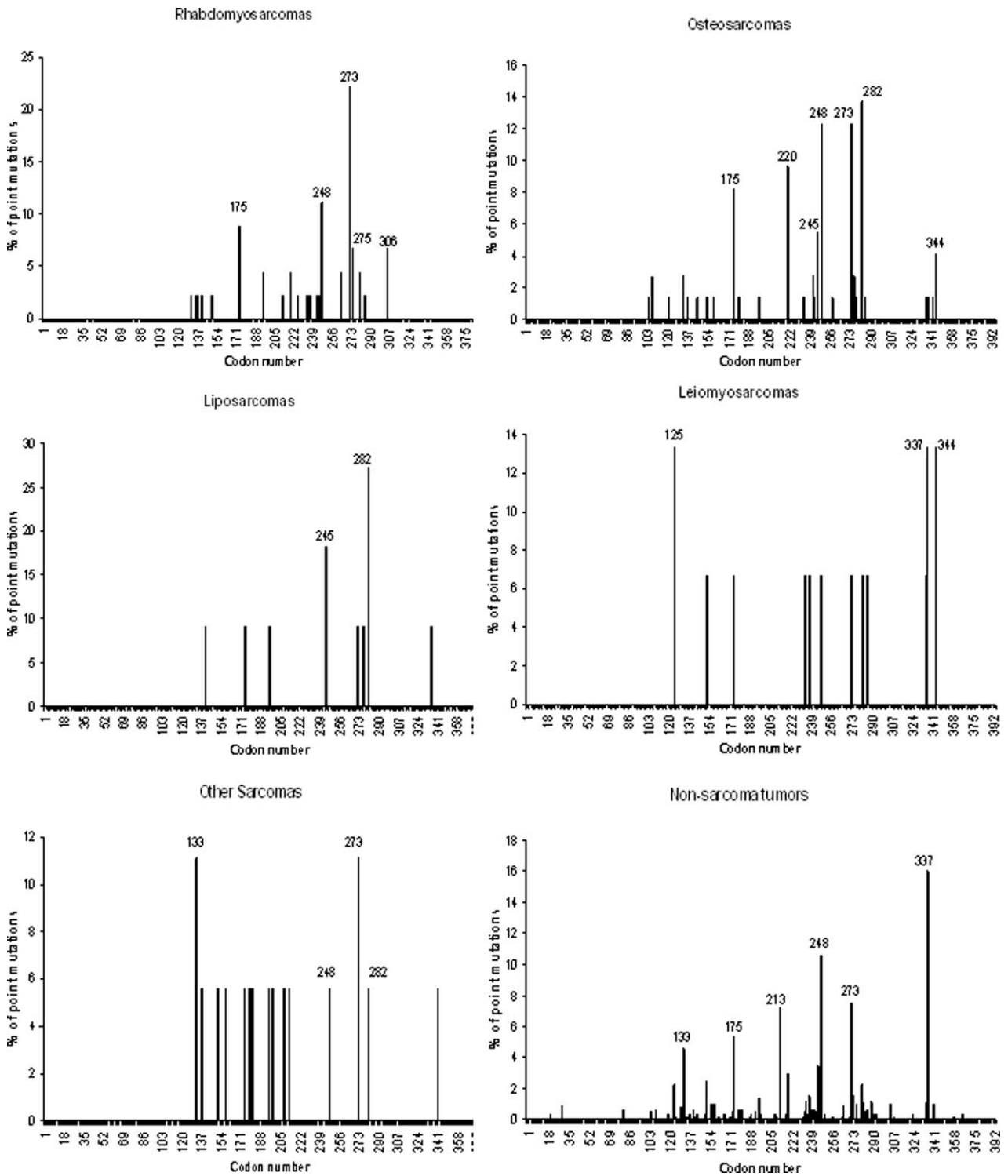

Figure 2. Codon distributions of tumor protein p53 (*TP53*) mutations associated with different types of sarcoma are illustrated in *TP53* germline mutation carriers. Only point mutations in the coding sequence are included.

Table 3. Association Between Structure/Function Groups of *TP53* Mutations and the Type of Sarcoma in Mutation Carriers

Tumor Type	Structural/Functional Characteristics of Mutations				
	Missense in DBM[a]	Missense Outside DBM[a]	FS, Splice, and Nonsense	Other	Total
No. observed					
Rhabdomyosarcoma	29	10	11	4[b]	54
Osteosarcoma	49	20	20	2[b]	91
Sarcoma, NOS	21	8	11	0[b]	40
Liposarcoma	8	1[b]	5	1[b]	15
Leiomyosarcoma	5	6	6	1[b]	18
All tumors, no.	418	303	196	34	951
All tumors, %[c]	43.95	31.86	20.61	3.58	100
No. observed/expected					
Rhabdomyosarcoma	1.22	0.58	0.99	2.07	1
Osteosarcoma	1.23	0.69	1.07	0.61	1
Sarcoma, NOS	1.19	0.63	1.33	0.00	1
Liposarcoma	1.21	0.21	1.62	1.86	1
Leiomyosarcoma	0.63	1.05	1.62	1.55	1

Abbreviations: DBM, DNA-binding motif; FS, frameshift; NOS, not othewise specified.
[a] DBMs located within the DNA binding domain include the following 3-dimensional structures of the p53 protein: loop 2 (L2)/L3 binding to the minor groove of DNA, and L1/sheet/helix 2 motifs binding to the major groove of DNA.
[b] Fields with <5 counts.
[c] This percentage was used to calculate the number of expected tumors.

sarcoma, whereas missense mutations in DBMs that predict stable, accumulating, mutant protein tended to be associated more with early onset forms of sarcoma. However, because of the small numbers in individual categories, the significance of these tendencies could not be assessed.

DISCUSSION

Although, collectively, all sarcoma types are rare in the general population (approximately 15,000 new cases are diagnosed annually in the United States), they are the second most common cancer after breast cancer in *TP53* mutation carriers. This observation was the main clinical clue for identifying the LFS as a distinct genetic entity. Individuals who are members of LFS families are prone to multiple cancers, including breast cancer and, occasionally, colorectal or gastric cancers, which commonly are associated with other cancer predisposition syndromes. Early occurrence of sarcoma played a central role in the clinical definition of LFS,[5,6] which preceded the recognition that germline *TP53* mutations were the underlying genetic defect in these families.[7] To date, however, the distribution, types, age at onset, and possible genotype/phenotype correlations of sarcoma in *TP53* mutation carriers have never been discussed extensively or compared with the distribution of sarcoma types in the general population. In the current study, we have compared data on sarcoma in germline *TP53* mutation carriers compiled in the IARC *TP53* database (irrespective of their detailed family history of cancer) with a data set of histologic types of sarcoma extracted from the SEER database, which is representative of the distribution of sarcoma types in the general, Caucasian population. In total, 236 sarcomas in *TP53* carriers were included in this comparison.

Overall, sarcoma represented 25% of all cancer diagnoses among germline *TP53* carriers in the IARC *TP53* database and 37% in patients aged <20 years, making it the most common germline *TP53*-related cancer in children and adolescents. Approximately 67% of all sarcomas among carriers occurred in individuals aged <20 years, which is a 6-fold higher proportion compared with the general population represented by the SEER data set. Given an overall penetrance of germline *TP53* mutation of approximately 25% by age 20 years (the IARC *TP53* database) and an estimated frequency of germline *TP53* mutation of 1 in 20,000 births[12] to 1 in 5000 births,[11] it is reasonable to assume that sarcoma caused by of germline *TP53* mutation in individuals aged 20 years occurs at the rate of about 0.2 to 0.8 per million population per year, representing approximately 2.5% to 10% of all sarcomas in this age group. This estimate is compatible with the reported detection of a germline *TP53* mutation in

4% to 9% of patients with early onset sarcoma (aged <45 years).[26–28] It follows that the risk of sarcoma before age 20 years may be up to 500 times higher in *TP53* carriers compared with the risk in the general population.

Our analysis indicates that the range of sarcoma types in *TP53* mutation carriers is not covering the full range of sarcoma types that occur in the general population. GISTs (which represent between 25% and 30% of all sarcoma types in the United States), Ewing sarcoma, and desmoid tumors have not been reported to date in *TP53* mutation carriers, consistent with the notion that these sarcoma types may be caused by p53-independent mechanisms.[29] Alternatively, the role of *TP53* mutations may be confined to a specific tumor subgroup, as reported for sporadic patients with Ewing sarcoma, in which mutations are confined to the less frequent, "high-risk" subgroup.[30] The pathologies that constitute the "sarcoma spectrum" in *TP53* mutation carriers include rhabdomyosarcoma, osteosarcoma, liposarcoma, leiomyosarcoma, and sarcoma, NOS.

Sarcoma in *TP53* mutation carriers reveals a very clear age-dependent pattern of occurrence, with first rhabdomyosarcoma (occurring almost exclusively in individuals aged <20 years) followed by osteosarcoma (representing the main type of sarcoma in adolescent carriers), whereas liposarcoma and leiomyosarcoma occur exclusively in adults (aged >20 years). Not surprisingly, the category "sarcoma, NOS" has a more widespread age distribution, consistent with the notion that this category includes several unspecified types of STS. Compared with the SEER data set, the age distribution of sarcoma in *TP53* mutation carriers revealed a statistically significant difference only for osteosarcoma, which represented a greater proportion of sarcoma in both children/adolescents and adults ($P = .0002$).

In the IARC *TP53* database, female carriers appear to have an overall higher risk of cancer, with females representing over 65% of all diagnoses. However, this sex difference is caused by the weight in the data set of female breast cancer. Excluding breast cancer, cancers diagnosed in females represent 51% of total diagnoses, both before and after age 20 years. In sarcoma, we observed a slightly higher proportion of females among *TP53* carriers compared with the SEER data set (56.4% vs 45.4%), although this difference was not statistically significant ($P = .17$). We also observed a small but significantly higher risk of osteosarcoma in childhood or adolescence for female carriers (OR, 1.69; $P = .04$). These observations are consistent with mouse models, which indicated that *TP53* female mutation carrier mice had a significantly higher risk of developing osteosarcoma,[31] as well as with the findings of Hwang and colleagues, who demonstrated a higher risk for female than male carriers for developing cancer in a cohort of patients with childhood sarcoma.[27] However, these sex differences should be confirmed by further investigation.

Comparisons between mutation types, patterns of base changes, and codon positions revealed interesting features in relation to age at onset and type of sarcoma. First, when we examined the structural and functional impact of mutations on the p53 protein, it emerged that mutations associated with early onset sarcoma (rhabdomyosarcoma and osteosarcoma) were more likely to be missense mutations in the DNA-binding domain (particularly in the region of DBMs), whereas mutations that were associated with late-onset sarcoma (liposarcoma and leiomyosarcoma) were more likely to be mutations that deleted the whole p53 protein (nonsense, splice site, or frameshift mutations) or missense mutations outside DBMs. These associations are suggestive of several biologic characteristics of mutations. First, missense mutations in DBMs appear to carry the most severe phenotype (early onset cancer) and to be more deleterious than mutations that occur outside DBMs or mutations that delete the whole p53 protein.[10] These results also are supported in mouse model, including both knock-out[32] and knock-in studies[33] (the latter studies revealed a higher propensity for sarcomas and more frequent metastatic disease). These observations are compatible with the notion that the most severe germline *TP53* mutations are those that induce the production of a mutant p53 protein with dominant-negative functions. It is noteworthy that the p.R337H mutation (guanine-to-adenine substitution at codon 1010) represents a low-penetrance mutation.[34] Thus, it is not surprising that this mutation, along with the p.G334V mutation, which also affects p53 oligomerization, is represented substantially among sarcoma with a later onset (leiomyosarcoma). In contrast, the p.R273H mutation in the DNA binding domain is a high-penetrance mutation: It was observed in close to 25% of carriers with rhabdomyosarcoma, which is the sarcoma type characterized by young age of onset. Further investigations are needed to improve the understanding of genotype-phenotype correlations in LFS-related sarcoma.

In conclusion, the current study emphasizes that individuals who have germline *TP53* mutations have a lifelong risk of developing different sarcoma types, with a pattern of diagnosis that distinguishes between 2 phases in the disease: a childhood/adolescent phase (characterized by the frequent occurrence of rhabdomyosarcoma and osteosarcoma) and an adult phase (characterized by later

onset sarcoma types, such as liposarcoma and leiomyosarcoma). These 2 phases appear to correspond to different types of mutations. The major strengths of this study are that it compiles the largest number of sarcoma cases in *TP53* carriers available in the literature to date and that it exploits population-based cancer registration data to better identify the characteristics of sarcoma associated with germline *TP53* mutation. The main limitation is that the numbers of *TP53* mutation carriers remain small, in particular when patients are subdivided according to different sarcoma types and/or different categories of mutations. In addition, the description of pathologies in the IARC *TP53* database occasionally lacks precision for identifying specific sarcoma subtypes. It will be important to take into account these observations in cohort studies on childhood/adolescent sarcoma types and to improve the standards for collecting and annotating clinical data on germline *TP53* mutation carriers into a large, specialized registry.

FUNDING SOURCES

This study was partially supported by Children's Cancer Research Fund, Minneapolis, Minnesota.

CONFLICT OF INTEREST DISCLOSURES

The authors made no disclosures.

REFERENCES

1. Jemal A, Siegel R, Ward E, et al. Cancer statistics, 2006. *CA Cancer J Clin.* 2006;56:106-130.
2. Ries LAG, Melbert D, Krapcho M, et al, eds. SEER Cancer Statistics Review, 1975–2004 [based on the November 2006 SEER data submission, posted to the SEER web site 2007]. Bethesda, MD: National Cancer Institute; 2007.
3. Brennan MF. Soft tissue sarcoma: advances in understanding and management. *Surgeon.* 2005;3:216-223.
4. Gurney JG, Young JL, Roffers SD, Smith MA, Bunin GR. Soft Tissue Sarcomas. SEER Pediatric Monograph. Bethesda, MD: National Cancer Institute; 2005.
5. Li FP, Fraumeni JF Jr. Rhabdomyosarcoma in children: epidemiologic study and identification of a familial cancer syndrome. *J Natl Cancer Inst.* 1969;43:1365-1373.
6. Li FP, Fraumeni JF Jr. Soft-tissue sarcomas, breast cancer, and other neoplasms. A familial syndrome? *Ann Intern Med.* 1969;71:747-752.
7. Malkin D, Li FP, Strong LC, et al. Germ line p53 mutations in a familial syndrome of breast cancer, sarcomas, and other neoplasms. *Science.* 1990;250:1233-1238.
8. Olivier M, Goldgar DE, Sodha N, et al. Li-Fraumeni and related syndromes: correlation between tumor type, family structure, and TP53 genotype. *Cancer Res.* 2003;63:6643-6650.
9. Tinat J, Bougeard G, Baert-Desurmont S, et al. 2009 Version of the Chompret criteria for Li Fraumeni syndrome. *J Clin Oncol.* 2009;27:e108-e109; author reply e110.
10. Bougeard G, Sesboue R, Baert-Desurmont S, et al. Molecular basis of the Li-Fraumeni syndrome: an update from the French LFS families. *J Med Genet.* 2008;45:535-538.
11. Lalloo F, Varley J, Ellis D, et al. Prediction of pathogenic mutations in patients with early onset breast cancer by family history. *Lancet.* 2003;361:1101-1102.
12. Gonzalez KD, Noltner KA, Buzin CH, et al. Beyond Li Fraumeni syndrome: clinical characteristics of families with p53 germline mutations. *J Clin Oncol.* 2009;27:1250-1256.
13. Achatz MI, Hainaut P, Ashton-Prolla P. Highly prevalent TP53 mutation predisposing to many cancers in the Brazilian population: a case for newborn screening? *Lancet Oncol.* 2009;10:920-925.
14. Hainaut P, Hollstein M. p53 and human cancer: the first ten thousand mutations. *Adv Cancer Res.* 2000;77:81-137.
15. Hafsi H, Hainaut P. Redox control and interplay between p53 isoforms: roles in the regulation of basal p53 levels, cell fate and senescence [published online ahead of print May 4, 2011]. *Antioxid Redox Signal.* 2011.
16. Olivier M, Petitjean A, Marcel V, et al. Recent advances in p53 research: an interdisciplinary perspective. *Cancer Gene Ther.* 2009;16:1-12.
17. Scian MJ, Stagliano KE, Anderson MA, et al. Tumor-derived p53 mutants induce NF-kappaB2 gene expression. *Mol Cell Biol.* 2005;25:10097-10110.
18. Strano S, Dell'Orso S, Di Agostino S, Fontemaggi G, Sacchi A, Blandino G.Mutant p53: an oncogenic transcription factor. *Oncogene.* 2007;26:2212-2219.
19. Petitjean A, Mathe E, Kato S, et al. Impact of mutant p53 functional properties on TP53 mutation patterns and tumor phenotype: lessons from recent developments in the IARC TP53 database. *Hum Mutat.* 2007;28:622-629.
20. Surveillance, Epidemiology, and End Results (SEER) Program. SEER Limited-Use Data (1975–2005) [released April 2008; based on the November 2007 submission]. Bethesda, MD: National Cancer Institute, DCCPS, Surveillance Research Program, Cancer Statistics Branch; 2008.
21. Ries LA, Smith MA, Gurney JG, et al, eds. SEER Cancer Statistics Review. Cancer Incidence and Survival Among Children and Adolescents. NIH Publ. No. 99–4649. Bethesda. MD: SEER Program, National Cancer Institute; 1999.
22. Zippin C, Lum D, Hankey BF. Completeness of hospital cancer case reporting from the SEER Program of the National Cancer Institute. *Cancer.* 1995;76:2343-2350.
23. Fritz A, Jack A, Parkin DM, et al. eds. International Classification of Diseases for Oncology. 3rd ed. Geneva, Switzerland: World Health Organization; 2000.
24. Steliarova-Foucher E, Stiller C, Lacour B, Kaatsch P.International Classification of Childhood Cancer, third edition. *Cancer.* 2005;103: 1457-1467.
25. Zeger SL, Liang KY. Longitudinal data analysis for discrete and continuous outcomes. *Biometrics.* 1986;42:121-130.
26. Diller L, Sexsmith E, Gottlieb A, Li FP, Malkin D.Germline p53 mutations are frequently detected in young children with rhabdomyosarcoma. *J Clin Invest.* 1995;95:1606-1611.
27. Hwang SJ, Lozano G, Amos CI, Strong LC. Germline p53 mutations in a cohort with childhood sarcoma: sex differences in cancer risk. *Am J Hum Genet.* 2003;72:975-983.
28. Toguchida J, Yamaguchi T, Dayton SH, et al. Prevalence and spectrum of germline mutations of the p53 gene among patients with sarcoma. *N Engl J Med.* 1992;326:1301-1308.
29. Lopez-Guerrero JA, Machado I, Scotlandi K, et al. Clinicopathological significance of cell cycle regulation markers in a large series of genetically confirmed Ewing's sarcoma family of tumors. *Int J Cancer.* 2011;128:1139-1150.
30. Neilsen PM, Pishas KI, Callen DF, Thomas DM. Targeting the p53 pathway in Ewing sarcoma [serial online]. *Sarcoma.* 2011;2011:746939.
31. Donehower LA, Harvey M, Vogel H, et al. Effects of genetic background on tumorigenesis in p53-deficient mice. *Mol Carcinog.* 1995;14:16-22.
32. Jacks T, Remington L, Williams BO, et al. Tumor spectrum analysis in p53-mutant mice. *Curr Biol.* 1994;4:1-7.
33. Lang GA, Iwakuma T, Suh YA, et al. Gain of function of a p53 hot spot mutation in a mouse model of Li-Fraumeni syndrome. *Cell.* 2004;119:861-872.
34. Figueiredo BC, Sandrini R, Zambetti GP, et al. Penetrance of adrenocortical tumours associated with the germline TP53 R337H mutation. *J Med Genet.* 2006;43:91-96.