Recent Natural Selection Identifies a Genetic Variant in a Regulatory Subunit of Protein Phosphatase 2A that Associates with Altered Cancer Risk and Survival

Lukasz F. Grochola,[1] Alexei Vazquez,[2,3,4] Elisabeth E. Bond,[1] Peter Würl,[5] Helge Taubert,[6] Thomas H. Müller,[7] Arnold J. Levine,[2,3] and Gareth L. Bond[1]

Abstract

Purpose: A regulated p53-dependent stress response is crucial in suppressing tumor formation and mediating the response to commonly used cancer therapeutics. However, little is known about the human, inherited genetics of this important signaling pathway.

Experimental Design: Studies of human genetic variants in the *p53* tumor suppressor gene and *MDM2* oncogene have shown that single nucleotide polymorphisms (SNP) can affect p53 signaling, confer cancer risk, and alter outcome, and also suggest that the pathway is under evolutionary selective pressure. Here, we attempt to accelerate the identification of functional p53 pathway SNPs by incorporating these characteristics into an analysis of 142 genes that are known to affect p53 signaling.

Results: We report that a genomic scan for recent natural selection denotes that of the 142 genes studied, the *PPP2R5E* gene that encodes a regulatory subunit of the tumor suppressing protein phosphatase 2A resides in a naturally selected genomic region. We go on to show that a selected SNP in *PPP2R5E* (ε-SNP2) associates with significant allelic differences in the onset (up to 19.2 years; $P = 0.0002$) and risk (odds ratio, up to 8.1; $P = 0.0009$) of soft tissue sarcoma development, as well as overall survival (relative risk, up to 3.04; $P = 0.026$).

Conclusions: The *PPP2R5E* gene is identified as harboring genetic variants that can affect human cancer and are possibly under evolutionary selection pressure. (Clin Cancer Res 2009;15(19):6301–8)

The p53 tumor suppressor pathway is central both in reducing cancer frequency in vertebrates and in mediating the response of commonly used cancer therapies (1, 2). Surprisingly, however, very little is known about the inherited genetic variation of this important signaling pathway (3, 4). The p53 pathway is a cellular stress response pathway that is activated upon stresses such as DNA damage and oncogene activation. Once activated, it initiates cellular responses such as DNA repair, cell cycle arrest, cell death (apoptosis), and senescence (5). A growing body of evidence is emerging in the literature that important genes in this pathway could harbor functional single nucleotide polymorphisms (SNP).

The SNPs in the p53 pathway that are most frequently studied are found in the *p53* and *MDM2* genes (*p53* codon72; ref. 6, rs1042522, C/G; *MDM2* SNP309; ref. 7, rs2279744, T/G). The allelic frequencies for both SNPs vary significantly in different ethnic and racial populations. For example, the G allele of *MDM2* SNP309 was found to be at 10% in African Americans (8), 33% in Northern Europeans (9), 45% in Asians (10), and 50% in Ashkenazi Jewish individuals (9). To explain the significant differences for both polymorphisms, allele-specific selective advantages and disadvantages for evolution and/or reproduction have been proposed (8, 11–13). Indeed, two recent reports have provided evidence of recent natural selection for both polymorphic loci (8, 13). For example, in the first report the *MDM2* haplotype structure was determined in Caucasians (8). It was noted that the unusually long and frequent haplotype that harbors the G-allele of *MDM2* SNP309 deviates significantly from the standard assumptions of selective neutrality using multiple selection tests. These include

Authors' Affiliations: [1]Ludwig Institute for Cancer Research, University of Oxford, Oxford, United Kingdom; [2]The Institute for Advanced Study, Princeton, New Jersey; [3]The Cancer Institute of New Jersey and [4]UMDNJ-Robert Wood Johnson Medical School, New Brunswick, New Jersey; [5]Department of Surgery, Malteser St. Franziskus Hospital, Flensburg, Germany; [6]Department of Oral and Maxillofacial Plastic Surgery, Martin-Luther-University Halle-Wittenberg, Halle, Germany; and [7]German Red Cross Blood Transfusion Service NSTOB, Springe, Germany
Received 3/31/09; revised 6/16/09; accepted 6/17/09; published OnlineFirst 9/22/09.

Grant support: Ludwig Institute for Cancer Research, the Simons Foundation, and the Leon Levy Foundation. Dr. Vazquez's IAS Membership was partially funded by the Helen and Martin Choolijian Founders Circle Member.

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 Clinical Cancer Research Online (http://clincancerres.aacrjournals.org/).

L.F. Grochola and A. Vazquez contributed equally to this article.

Requests for reprints: Gareth L. Bond, Ludwig Institute for Cancer Research, ORCRB, University of Oxford, Old Road Campus, Off Roosevelt Drive, Oxford, OX3 7DQ, United Kingdom. Phone: 44(0)1865-617497; Fax: 44(0)1865-617515; E-mail: gareth.bond@ndm.ox.ac.uk.

doi:10.1158/1078-0432.CCR-09-0797

Translational Relevance

The identification and description of high-frequency genetic variants that alter cancer risk, progression, and outcome is thought to eventually lead both to their use in personalizing prevention and therapeutic strategies, and to the development of novel ones. A large body of evidence strongly suggests that the p53 tumor suppressor pathway is central both in reducing cancer frequency in vertebrates and in mediating the response of commonly used cancer therapies. Here, we describe and utilize an approach that accelerates the identification of functional p53 stress response single nucleotide polymorphisms that affect human cancer. We identify the *PPP2R5E* gene, which encodes a regulatory subunit of the tumor suppressing protein phosphatase 2A, as harboring genetic variants that can affect soft tissue sarcoma and that may be under evolutionary selection pressure. The presented data not only support an important role for *PPP2R5E* in sarcoma development and outcome, but also support its value as a prognostic and predictive factor.

an entropy-based selection test that compares both the frequency and long-range correlations of an allele with a simulated model in which the allele is selectively neutral. Although the precise reason(s) for the selection pressure on these functional p53 pathway SNPs is still under investigation, many reported observations suggest that selection pressure on the p53 pathway may be possible. For example, a well-regulated p53 pathway has been shown to be crucial not only for murine tumor suppression and the response to DNA damaging therapies, but also for proper embryonic development, embryonic implantation, pregnancy rates, and litter sizes (2, 8, 12). Indeed, allelic differences have been noted for both functional p53 pathway SNPs in human fertility (14–17). However, regardless of the precise reasons for the selection pressures, signs of natural selection of p53 variants seem to be a characteristic of functional p53 pathway SNPs as has been further validated by a recent study of the genetic variation of the *MDM4* oncogene (18).

Both *p53* codon72 and *MDM2* SNP309 have been shown to associate with differences in the onset and the risk of cancer, as well as survival (7, 19, 20). For example, *p53* mutation carriers (Li-Fraumeni) with the G-allele of *MDM2* SNP309 were shown to be diagnosed with tumors on average 7, 10, and 16 years earlier than carriers that are T/T in genotype (7, 21, 22). Earlier ages of onset associated with individuals with the G-allele, but no known *p53* mutations, were also shown in soft tissue sarcomas (7), lymphoma (9), leukemia (23), and head, neck and oral squamous cell carcinomas (24, 25), and in cancers of the colon (26), breast (9, 27), bladder (28), ovary (29), brain (30), and liver (31). It has been proposed that the high levels of *MDM2* resulting from the G-allele attenuate the *p53* stress response, resulting in a higher mutation rate, poorer DNA repair processes, reduced apoptosis, and senescence, leading to faster and more frequent tumor formation (7). Indeed, G-allele carriers of *MDM2* SNP309 have been shown to harbor an increased risk of developing many cancers (3, 19).

The identification and description of high-frequency SNPs in the p53 stress response pathway that alter cancer risk, progression, and outcome will not only potentially increase our understanding of this tumor suppressor pathway in humans, but could also lead to their use in personalizing prevention and therapeutic strategies. Therefore, in this report we attempt to accelerate the identification of functional genetic variants, by incorporating the above-described characteristics of functional p53 pathway SNPs into an analysis of 142 genes known to be important in mediating and regulating the p53 stress response. Specifically, genetic variants are sought in those genes, which associate with haplotypes that show signs of natural selection, and differences in soft tissue sarcoma risk and survival.

Materials and Methods

Patients. A total of 130 patients (73 females and 57 males; ages 14 to 87 y; mean, 56.3 y) diagnosed with soft tissue sarcoma in the years 1991 to 2001 at the Surgical Clinic 1, University of Leipzig, Germany and at the Institute of Pathology of the Martin-Luther-University Halle, Germany, were included in the study. The mean observation time was 40.5 mo (range, 2-198 mo). Sixty-three patients died from tumor-related causes within the observation time, whereas 67 patients were still alive after this period. All patients underwent surgical treatment and the subsequent mean survival rate was 25.1 mo (range, 2-119 mo). The DNA of 66 patients was extracted from whole blood samples (n = 22) or normal, pathologically confirmed tumor-free tissue, adjacent to the resection specimen (n = 44). The DNA of the remaining 64 patients was obtained from tissue within the tumor confinements. The control cohort consisted of 497 blood donors (Germans with central European origins; 193 females and 304 males, ages 19 to 68 y; mean, 43.9 y) from whom DNA was extracted from whole-blood samples that were obtained at the German Red Cross Blood Transfusion Service NSTOB (Springe, Germany). All persons gave written and informed consent. Approval from the local ethics committee was obtained.

Sequence analysis. DNA from the preserved material was extracted using the Innuprep Blood DNA mini-kit and the Innuprep DNA mini-kit (AJ Innuscreen GmbH). The *PPP2R5E* SNP, rs11158491 (ε-SNP2), was genotyped in the patients and controls using the Applied Biosystems' SNP genotyping assay C___2819093_10 with the AB 7500 Real-Time PCR System (Applied Biosystems).

Statistical analysis. A permutation test was done to determine the statistical significance of the noted increase of the mean age of tumor diagnosis, whereby T/T<C/T<C/C. The statistical significance for the enrichment of T/T carriers in the cases relative to the controls was computed using, first, a Bayesian estimate of the probability density function of the fraction of T/T homozygous in the control population, and second, a binomial test to compute the probability to observe (n)T/T cases in (n) cases. The mean odd ratios (OR) and 95% confidence intervals between the C/C and T/T homozygots were determined using a Bayesian estimate. Case and controls were matched for both gender and age. Survival analysis was done using the Kaplan-Meier and the Cox's multivariate proportional hazards regression model with SPSS 16.0 software (SPSS Inc.). Patients who were alive at the time of the last follow-up were included as censored data into the survival analyses. Values of $P < 0.05$ were considered significant.

Results

To identify genetic variants in *p53* stress response genes that deviate from the standard assumptions of selective neutrality, we utilized a previously described and publicly available map of recent positive selection of the human genome (32).[8] In

[8] Haplotter, http://hg-wen.uchicago.edu/selection/haplotter.htm.

Table 1. Twenty-nine HapMap *PPP2R5E* SNPs with minor allele frequencies >10% having extreme iHS scores

Region	Contig position	SNP	Alleles	Derived allele	Derived allele frequency	iHS
Intron 2	44978228	rs1255641	C/T	C	0.885	2.55
Intron 2	44965697	rs1255627	A/G	A	0.839	2.80
Intron 2	44949747	rs17101239	C/G	C	0.780	2.74
Intron 2	44948642	rs10145422	C/T	T	0.828	2.61
Intron 2	44945051	rs1951213	C/G	G	0.828	2.61
Intron 2	44944616	rs11158492	A/G	G	0.828	2.61
Intron 2	44942312	rs7153764	C/T	C	0.828	2.71
Intron 2	44941827	rs1957045	A/G	G	0.828	2.71
Intron 2	44941305	rs4391977	A/G	G	0.828	2.89
Intron 2	44936970	rs8018107	A/T	T	0.828	2.79
Intron 2	44934705	rs1033745	C/T	C	0.839	2.57
Intron 2	44933964	rs1033744	C/T	C	0.833	2.78
Intron 2	44931618	rs10141179	C/T	C	0.828	2.75
Intron 2	44926433	rs7143816	C/T	T	0.828	2.75
Intron 2	44923835	rs12323479	A/G	G	0.828	2.76
Intron 2	44922047	rs10151527	A/C	C	0.722	-3.00
Intron 2	44922013	rs4902226	C/T	T	0.822	2.88
Intron 3	44916716	rs10152067	A/G	A	0.828	2.60
Intron 3	44914073	rs6573518	C/T	T	0.828	2.61
Intron 3	44910858	rs11158491*	C/T	C	0.722	-2.87
Intron 3	44909091	rs7146069	A/G	G	0.806	2.84
Intron 3	44907997	rs8003647	C/T	T	0.828	2.65
Intron 3	44907693	rs8018307	C/G	C	0.828	2.65
Intron 3	44901047	rs10140679	C/T	C	0.744	-2.65
Intron 4	44888075	rs10145042	C/T	T	0.911	2.68
Intron 5	44880199	rs7141605	C/G	C	0.828	2.94
Intron 5	44879791	rs7140704	A/G	A	0.844	2.76
Intron 5	44876923	rs11158490	G/T	T	0.833	2.85
Intron 11	44853005	rs7146445	A/C	A	0.906	2.53

*PPP2R5E ε-SNP2.

Haplotter, recent positive selection is determined using a haplotype-based approach that looks for enrichment of the classic signal for strong directional selection using the phase II data of the HapMap. The classic signal for strong directional selection is an extended haplotype with very low diversity, or high homozygosity. This is thought to occur when a selected allele increases in frequency so quickly that insufficient time remains to allow recombination to reduce its length. Haplotter utilizes a test

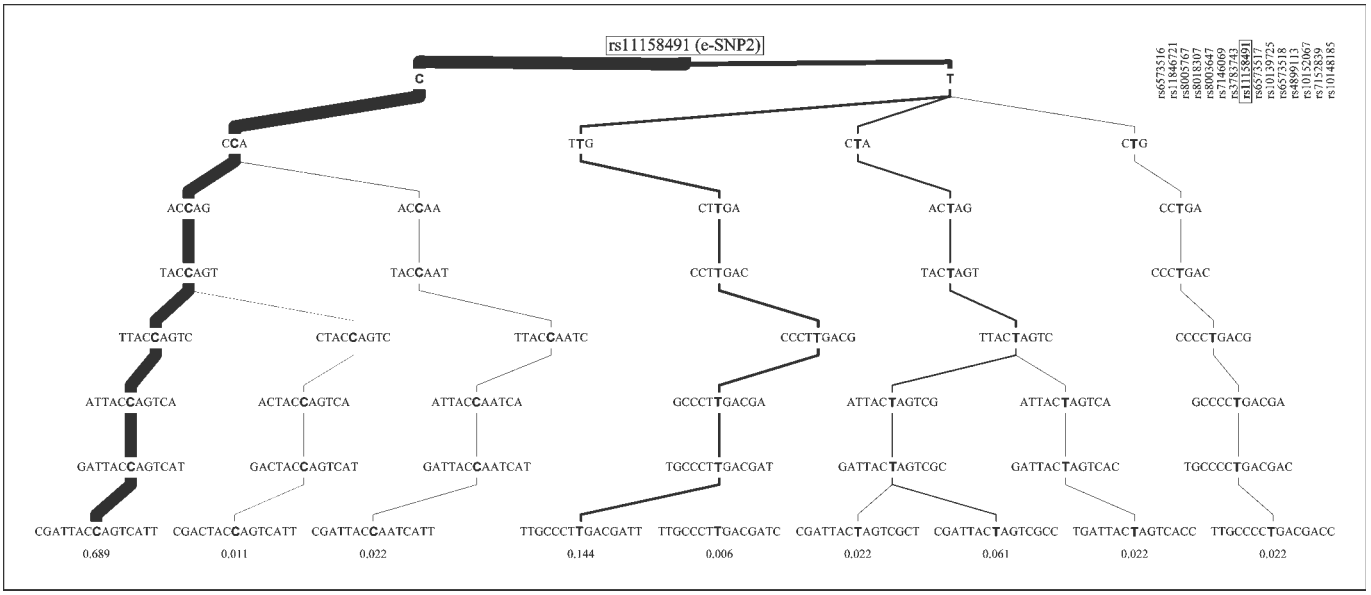

Fig. 1. A depiction of the haplotype structure of the *PPP2R5E* gene in Central Europeans centered on ε-SNP2 (rs11158491). The C-allele is the derived allele of *PPP2R5E* ε-SNP2 and the T allele is the ancestral allele. The numbers at the bottom of each haplotype denote the inferred frequencies. In the upper right corner, the order of the SNPs is reported using the annotation of dbSNP (National Center for Biotechnology Information, rs -numbers).

Table 2. Clinical and histopathologic characteristics of 130 sporadic soft tissue sarcoma patients

	Number of patients
Stage	
I	14
II	52
III	34
IV	30
Grading	
G_1	18
G_2	72
G_3	39
Not known	1
Tumor resection	
Radical (R_0)	96
Not radical (R_1)	18
Not radical (R_2)	14
Not known	2
Tumor localization	
Head	2
Neck	2
Upper extremity	12
Shoulder	1
Thoracic wall	11
Intrathoracal	0
Intraabdominal	12
Abdominal wall	0
Retroperitoneum	15
Hip/pelvis/gluteal	17
Lower extremity	55
Multiple localizations	3
Tumor type	
Liposarcoma	30
Fibrosarcoma	8
Malignant fibrous sarcoma	33
Neurogenic sarcoma	11
Rhabdomyosarcoma	8
Leiomyosarcoma	23
Synovial sarcoma	9
Others	8

statistic called the integrated haplotype score (iHS), which is a measure that includes the degree of haplotype homozygosity around a given SNP (33). Highly positive and negative iHS scores (>2.5 or <-2.5) denote SNPs that harbor higher haplotype homozygosity, compared with other SNPs with similar allele frequencies in the genome. A large negative score denotes an unusually long haplotype carrying the derived (nonancestral) allele; a large positive score denotes a long haplotype carrying the ancestral allele.

Importantly, a genomic region or gene is only noted to be positively selected when it contains numerous SNPs with extreme iHS scores (32). An empirical *P* value is assigned to each gene as a fraction of genes in the genome with the same number of SNPs with extreme iHS scores, in a window of 50 consecutive SNPs, that is equal to or larger than the given gene. When we analyze the 142 genes important in mediating and regulating the p53 stress response (Supplementary Table S1), only one gene resides in a genomic region that shows significant evidence of recent positive selection. A gene that encodes a regulatory subunit of the protein phosphatase 2A (PP2A) holoenzyme, *PPP2R5E* (protein phosphatase 2, regulatory subunit B′, ε isoform), is found to be in a region of the genome that shows significant signs of natural selection in the Caucasian population of the HapMap (north-western European ancestry, $P = 0.006713$). Specifically, out of the 122 genotyped SNPs in *PPP2R5E*, with minor allele frequencies ≥10%, 29 SNPs have extreme iHS scores (Table 1, 26 SNPs with iHS >2.5 and 3 SNPs with iHS <-2.5).

The subsequent analyses focus on the three SNPs wherein the derived alleles associate with unusually long haplotypes as they might denote a functional nonancestral allele that has undergone a recent selective sweep. The alleles of all three loci are highly linked to each other, to such an extent that these associations are nonrandom and, therefore, in very high linkage disequilibrium (LD). Specifically, rs10151527 in intron 2 (ε-SNP1) is in either complete or almost complete LD with two SNPs in intron 3, rs11158491 (ε-SNP2; $D' = 1$; $R^2 = 1$) and rs10140679 (ε-SNP3; $D' = 1$; $R^2 = 0.87$). Indeed, the derived alleles for all three SNPs are at very high frequencies in the central European populations (>72%), and the low diversity of the derived haplotypes are readily apparent. As seen in Fig. 1, when we analyze the marginal haplotypes of the 15 SNPs centered on ε-SNP2 (rs11158491), it is clear that the high-frequency (73%) derived allele (C) shows significantly lower haplotype diversity than the lower-frequency (27%) ancestral allele (T), and has a dominant haplotype in 69% of the population. This noted high-frequency, low-diversity haplotype is similar to the haplotype containing the functional G-allele of *MDM2* SNP309 (8). We reasoned that the *PPP2R5E* haplotype, denoted by these SNPs, might also harbor one or more functional alleles and, therefore, also show allelic differences in human cancer populations.

The first noted clinical association of the G-allele of *MDM2* SNP309 was its association with an earlier tumor onset in soft tissue sarcoma patients, first in *p53* mutation carriers and subsequently in sporadic cases (7). Soft tissue sarcoma accounts for <1% of all human malignancies with a cumulative incidence of 3 per 100,000 people per year (34). As mentioned above, similar effects have been subsequently observed in a variety of other tumor types (3). To test the hypothesis that the *PPP2R5E* haplotype might also harbor one or more functional SNPs, which show allelic differences in human cancer populations, we genotyped ε-SNP2 in the same cohort of sporadic soft tissue sarcoma patients in which the affect of age-dependent incidence of the *MDM2* SNP309 locus was first described (refs. 7, 9; Table 2).

Similar to *MDM2* SNP309, when the sarcoma patients were separated based on the *PPP2R5E* ε-SNP2 genotype, patients homozygous for the T-allele were found to be diagnosed significantly earlier in life than individuals C/T or C/C in genotype. Specifically, sarcoma patients with a C/C and C/T genotype were diagnosed with a mean age of 58.2 and 57.6 years, respectively, whereas patients homozygous for the T-allele showed an

Table 3. Average (median) age of tumor diagnosis for *PPP2R5E* ε-SNP2 genotypes in years

	n	CC	CT	TT	*P**
Total	130	58.2 (59)	57.6 (62)	46.4 (44)	0.0010
Males	57	59.1 (62)	51.7 (57.5)	39.9 (41.5)	0.0002
Females	73	57.3 (57.5)	61.2 (65)	51.6 (53.5)	1

*Permutation Test.

Table 4. Summary of case-control analyzes of *PPP2R5E* ε-SNP2

	N	C/C *n* (%)	C/T *n* (%)	T/T *n* (%)	*P**	Mean Odds Ratio T/T: C/C (95% CI)†
Controls						
Total	497	266 (53.5)	193 (38.9)	38 (7.6)	—	—
Cases						
Total	130	65 (50)	47 (36.2)	18 (13.8)	0.0250	2.3 (1.05-3.67)
<51 y	43	19 (44.2)	12 (27.9)	12 (27.9)	0.0010	5.3 (1.82-9.58)
≥51 and <66 y	43	23 (53.5)	17 (39.5)	3 (7)	>0.05	
≥66 y	44	23 (52.3)	18 (40.1)	3 (6.8)	>0.05	
Males	57	31 (54.5)	18 (31.5)	8 (14)	>0.05	
<51 y	21	8 (38.1)	6 (28.6)	7 (33.3)	0.004	8.1 (1.57-17.23)
≥51 and <66 y	20	12 (60)	8 (40)	0 (0)	>0.05	
≥66 y	16	11 (68.7)	4 (25.0)	1 (6.3)	>0.05	
Females	73	34 (46.6)	29 (39.7)	10 (13.7)	>0.05	
<51 y	22	11 (50.0)	6 (27.3)	5 (22.7)	0.048	6.1 (0.87-13.65)
≥51 and <66 y	23	11 (47.8)	9 (39.2)	3 (13.0)	>0.05	
≥66 y	28	12 (42.9)	14 (50.0)	2 (7.1)	>0.05	

Abbreviation: 95% CI, 95% confidence interval.
*Probability to observe (*n*) TTs in (*n*) cases given the (*n*) TTs in (*n*) controls. Cases and controls were age and gender matched.
†Bayesian estimate.

11.8-year earlier age of tumor onset at 46.4 years ($P = 0.001$, permutation test; Table 3). Interestingly, a further similarity with the observations made for *MDM2* SNP309 (3, 9, 35, 36) was that these allelic differences show a strong gender-dependence. For ε-SNP2, the greatest allelic difference in the age of onset was found in the male patients. Males T/T in genotype showed a 19.2-year earlier tumor onset when compared with males homozygous for the C-allele ($P = 0.0002$, permutation test; Table 3). A similar, albeit weaker, 5.7-year difference was observed in the female patients (Table 3). However, these differences failed to reach statistical significance.

To expand our analysis of the allelic differences of ε-SNP2 in the risk of developing soft tissue sarcomas, we genotyped ε-SNP2 in 497 blood donors of the same ethnic background. Among them, 38 (T/T, 7.6 %) were homozygous for the ε-SNP2 T-allele, 266 (C/C, 53.5 %) were homozygous for the C-allele, and 193 (T/C, 38.9 %) were heterozygous (Table 4). These ε-SNP2 genotype frequencies differed significantly from those found in the 130 sarcoma patients. Interestingly, an enrichment of individuals homozygous for the T-allele (13.8%) and a decrease in the C/C genotype frequency (50.0%) were noted in the sarcoma patient group, suggesting that T/T individuals harbor a higher risk of sarcoma development (OR, 2.3; $P = 0.025$, Table 4).

The above-described age-of-onset analysis suggests that significant enrichments of particular genotypes in the patient populations, compared with controls, could also be gender- and/or age-dependent. Therefore, incorporating these parameters could identify individuals with an even higher risk of sarcoma development. To do this, we divided either all sarcoma patients, or male patients and female patients, into three groups of almost equal size according to their age at diagnosis: the youngest, intermediately aged, and oldest. Interestingly, upon comparison of the genotype frequencies, the youngest patients in all three groupings (total, males and females) showed a significant enrichment of the T/T genotype when compared with the controls, and a subsequent decrease in the frequency of the C/C genotype (Table 4). Specifically, individuals homozygous for the T-allele made up 27.9% of the youngest patients as opposed to 7% of intermediately aged, 6.8% of oldest cases, and 7.6% of controls (OR, 5.3; $P = 0.0010$; Table 4). Furthermore, in agreement with the previous results, these observations were particularly pronounced in the youngest male patients, whereas men T/T in genotype made up 33.3% of the male cases diagnosed at youngest age, as opposed to 6.3% of oldest male cases, and 7.6 % of blood donors (OR, 8.1; $P = 0.0040$; Table 4). In contrast, men C/C in genotype made up only 38.1% of the male cases diagnosed at youngest age, as opposed to 60% of intermediately aged, 68.7% of oldest male cases, and 53.5% of blood donors. As with the age-of-onset analysis, similar, weaker trends were observed in the female patients, whereas the T/T genotype made up 22.7% of youngest females (OR, 6.1; $P = 0.048$; Table 4) and only 7.1% of oldest women.

Together, the above-described significant allelic differences in age-dependent onset and risk of soft tissue sarcoma development lend strong support to the hypothesis that the *PPP2R5E* haplotypes, denoted by *PPP2R5E* ε-SNP2, might also harbor one or more functional, clinically relevant SNP(s). To further test the perceived clinical relevance of ε-SNP2, we assessed the potential impact of this locus on the overall survival of the sarcoma patients. To begin with, we did a Kaplan-Meier survival analysis that revealed that sarcoma patients C/C in genotype associated with the worst overall survival with a mean survival time of 48.4 months, followed by patients heterozygous for the C-allele (mean survival time of 69.9 months) and subsequently individuals with a T/T genotype (mean survival time of 142.7 months; $P = 0.03$, log rank test; Fig. 2). To exclude potential biases from other independent prognostic factors, we carried out a multivariate Cox's regression survival analysis adjusted to tumor stage, subtype, and localization, as well as type of tumor resection and gender. This analysis further confirmed the initial results, showing a 3.04-fold increased risk for tumor-related death for patients C/C in genotype when compared with individuals homozygous for the T-allele ($P = 0.026$; Fig. 2).

Discussion

In this report, we set out to search for genetic variants in 142 genes known to be important in mediating and regulating the p53 stress response that show signs of natural selection and differences in sarcoma risk and survival. By utilizing a map of recent natural selection (Haplotter), we identified only one gene of these 142 that resides in a genomic region that shows highly significant evidence of selection. This gene (*PPP2R5E*) encodes a regulatory subunit of PP2A. The strongest signs of natural selection that associated with derived (nonancestral) alleles were from three SNPs that were in high linkage disequilibrium with each other. To assess potential differences in sarcoma risk and survival, one of these linked SNPs (*PPP2R5E* ε-SNP2) was genotyped in a cohort of patients diagnosed with soft tissue sarcoma and a control group of the same ethnicity. Interestingly, the subsequent analysis revealed that, in soft tissue sarcoma patients, *PPP2R5E* ε-SNP2 associates strongly with both altered risk and survival. Individuals homozygous for the ancestral allele (T/T) associate with significant allelic differences in the onset and risk of soft tissue sarcoma development in a gender-dependent manner. Specifically, male T/T carriers are diagnosed 19 years earlier than male C/C carriers and have a significantly increased risk of developing early onset disease (OR, 8.1; Table 4). However, once diagnosed with soft tissue sarcoma, C/C carriers associate with a significant and increased risk of tumor-related death (relative risk, 3.04; Fig. 2).

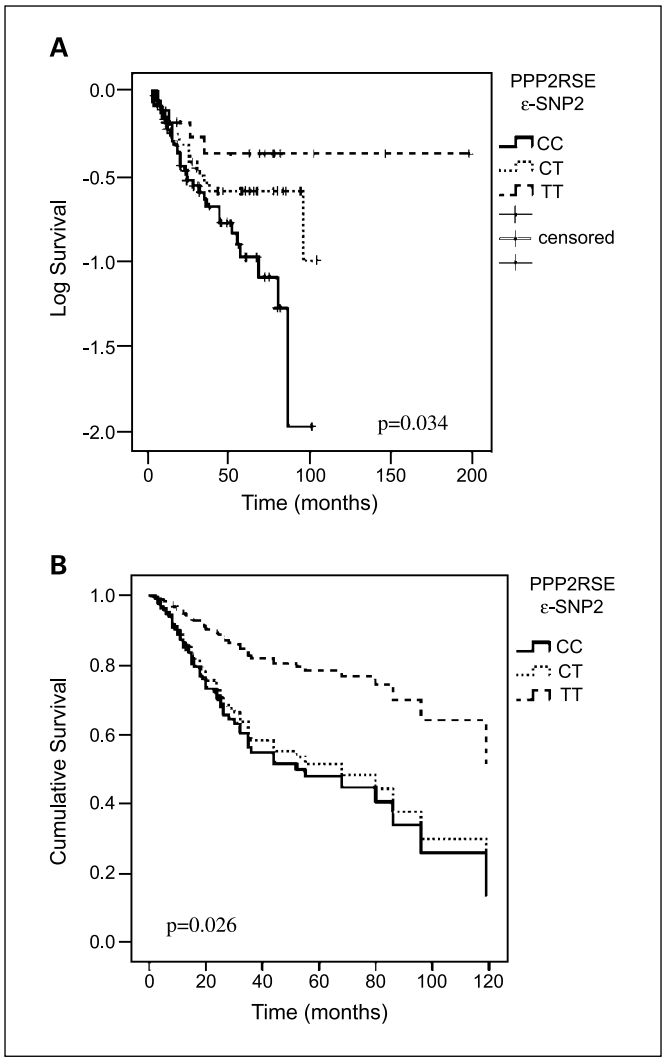

Fig. 2. Survival analysis of *PPP2R5E* ε-SNP2 shows an association of the derived C-allele with an increased risk for tumor-related death. The survival curves of soft tissue sarcoma patients for each genotype are plotted against the survival time in months. Patients who where were alive at the time of follow-up were included as censored data into the analyses. The results of the Kaplan-Meier analysis (*A*) are further supported by the Cox multivariate regression survival analysis (*B*) that was adjusted to tumor stage, subtype, and localization, as well as type of tumor resection and gender. In the Cox analysis, patients with a C/C genotype (*black line*) showed a 3.04-fold increased risk for tumor-related death when compared with patients homozygous for the T-allele (*black dashed line*).

The genomic region noted to be under evolutionary selection pressure also included two other, less-characterized genes, *WDR89* and *SGPP1*. Neither gene has been shown to play a role in the induction or regulation of malignant transformation. In contrast, mounting evidence of the tumor suppressing activity of the PP2A holoenzyme can be found in the literature (37–42). An important mechanism of intracellular signaling is the regulated, reversible, phosphorylation of proteins, and deregulation of these events can play a central role in the malignant transformation of the cell. Recent work has provided compelling evidence for the role of specific protein phosphatases, particularly of PP2A, in carcinogenesis and tumor progression (39, 42). PP2A phosphatase activity has been shown to interact directly with the p53 pathway, causing the dephosphorylation of key residues of *p53* and *MDM2*, resulting in the regulation of *p53* activity levels in cells (43–47). For example, the PP2A-mediated dephosphorylation of *MDM2* has been linked to the ability of *MDM2* to degrade *p53* (47). It will be interesting to determine whether the varying levels of MDM2 protein produced by the different alleles of *MDM2* SNP309 can modify the significant allelic differences of the *PPP2R5E* variant in cancer risk and survival.

The genes that encode the subunits of the PP2A holoenzyme were included in our study due to the above-described role that PP2A plays in regulating p53 signaling (43–47). However, PP2A can affect other, at least partially, p53-independent intracellular signaling pathways that affect human cancer (42). For example, PP2A can antagonize Ras signaling by dephosphorylating c-Myc and RalA and by negatively regulating the PI-3 kinase/Akt signaling pathway (39). Indeed, it will be of future interest to explore whether the significant allelic differences in haplotype structure, cancer risk, and cancer survival noted in this report are due to allelic differences in p53 signaling or other signaling pathways affected by PP2A activity. Further evidence of the importance of PP2A activity in suppressing cellular transformation lies in the wide range of mechanisms that transformed cells have evolved to inhibit its activity. For example, inhibition of PP2A activity has been shown to be mediated by the small tumor antigen of DNA tumor viruses (48) by up-regulation of the c-Myc–specific inhibitor CIP2A (49), by BCR/ABL via SET up-regulation (50), through biallelic mutational inactivation of the Aβ subunit (51), or by decreased expression of the Aα subunit (37).

PP2A is a ubiquitously expressed heterotrimeric protein that accounts for a large fraction of phosphatase activity in eukaryotic cells (39). The various forms of PP2A contain an active core dimer, made up of the catalytic (C-) subunit and a structural scaffold A subunit. The scaffold subunit mediates interaction of the core dimer with a great variety of regulatory (B-) subunits

that determine the localization, specificity, and functions of individual isoforms (42). Four families of B subunits have been described: B/B55/PR55/PPP2R2, B′/B56/PR61/PPP2R5, B″/PR72/PPP2R3, and B‴/PR93/SG2NA/PR110/Striatin (39). Furthermore, five different *PPP2R5* genes have been identified: *PPP2R5A* (α-), *PPP2R5B* (β-), *PPP2R5C* (γ-), *PPP2R5* (δ-) and *PPP2R5E* (ε-isoform; ref. 52). To date, little is known about the specific regulatory role of *PPP2R5E*.

The observations reported here remain to be validated in other patient cohorts, and importantly, the regulatory changes associated with *PPP2R5E* ε-SNP2 are yet to be determined. The SNPs that associated with the strongest signs of natural selection with derived (nonancestral) alleles are in high linkage disequilibrium with each other and reside in either intron 2 or intron 3 (Table 1). To our knowledge, there have been no published attempts to characterize the regulatory regions of the *PPP2R5E* gene, and therefore a discussion of whether or not one of these SNPs resides in a potential regulatory region would be premature. Furthermore, more exhaustive searches for other genetic variants closely linked to these SNPs, but not included in the HapMap project, will be necessary to develop a complete list of candidate functional SNPs that merit further experimental investigation into the molecular and cellular mechanisms behind the significant allelic differences in haplotype structure, cancer risk, and cancer survival noted in this report. However, these data strongly suggest that *PPP2R5E* plays a significant regulatory role in the development and progression of soft tissue sarcomas and harbors genetic variants that can affect human cancer and may be under evolutionary selection pressure.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Acknowledgments

We thank Thomas Greither for his invaluable technical assistance, Claire Beveridge and Mark Shipman for their help in the preparation of the manuscript, and William Fairbrother for helpful discussions.

References

1. Haupt S, Haupt Y. Importance of p53 for cancer onset and therapy. Anticancer Drugs 2006;17:725–32.
2. Lozano G, Zambetti GP. What have animal models taught us about the p53 pathway? J Pathol 2005;205:206–20.
3. Bond GL, Levine AJ. A single nucleotide polymorphism in the p53 pathway interacts with gender, environmental stresses and tumor genetics to influence cancer in humans. Oncogene 2007;26:1317–23.
4. Vazquez A, Bond EE, Levine AJ, Bond GL. The genetics of the p53 pathway, apoptosis and cancer therapy. Nat Rev Drug Discov 2008;7:979–87.
5. Murray-Zmijewski F, Slee EA, Lu X. A complex barcode underlies the heterogeneous response of p53 to stress. Nat Rev Mol Cell Biol 2008;9:702–12.
6. Harris N, Brill E, Shohat O, et al. Molecular basis for heterogeneity of the human p53 protein. Mol Cell Biol 1986;6:4650–6.
7. Bond GL, Hu W, Bond EE, et al. A single nucleotide polymorphism in the MDM2 promoter attenuates the p53 tumor suppressor pathway and accelerates tumor formation in humans. Cell 2004;119:591–602.
8. Atwal GS, Bond GL, Metsuyanim S, et al. Haplotype structure and selection of the MDM2 oncogene in humans. Proc Natl Acad Sci U S A 2007;104:4524–9.
9. Bond GL, Hirshfield KM, Kirchhoff T, et al. MDM2 SNP309 accelerates tumor formation in a gender-specific and hormone-dependent manner. Cancer Res 2006;66:5104–10.
10. Hong Y, Miao X, Zhang X, et al. The role of P53 and MDM2 polymorphisms in the risk of esophageal squamous cell carcinoma. Cancer Res 2005;65:9582–7.
11. Beckman G, Birgander R, Sjalander A, et al. Is p53 polymorphism maintained by natural selection? Hum Hered 1994;44:266–70.
12. Hu W, Feng Z, Teresky AK, Levine AJ. p53 regulates maternal reproduction through LIF. Nature 2007;450:721–4.
13. Shi H, Tan SJ, Zhong H, et al. Winter temperature and UV are tightly linked to genetic changes in the p53 tumor suppressor pathway in Eastern Asia. Am J Hum Genet 2009;84:534–41.
14. Kang HJ, Feng Z, Sun Y, et al. Single-nucleotide polymorphisms in the p53 pathway regulate fertility in humans. Proc Natl Acad Sci U S A 2009;106:9761–6.
15. Kay C, Jeyendran RS, Coulam CB. p53 tumour suppressor gene polymorphism is associated with recurrent implantation failure. Reprod Biomed Online 2006;13:492–6.
16. Fang Y, Kong B, Yang Q, Ma D, Qu X. MDM2 309 polymorphism is associated with missed abortion. Hum Reprod 2009;24:1346–9.
17. Firouzabadi RD, Ghasemi N, Rozbahani MA, Tabibnejad N. Association of p53 polymorphism with ICSI/IVF failure and recurrent pregnancy loss. Aust N Z J Obstet Gynaecol 2009;49:216–9.
18. Atwal GS, Kirchhoff T, Bond EE, et al. Altered tumor formation and evolutionary selection of genetic variants in the human MDM4 oncogene. Proc Natl Acad Sci U S A 2009;106:10236–41. Epub 2009 Jun 4.
19. Hu Z, Jin G, Wang L, Chen F, Wang X, Shen H. MDM2 promoter polymorphism SNP309 contributes to tumor susceptibility: evidence from 21 case-control studies. Cancer Epidemiol Biomarkers Prev 2007;16:2717–23.
20. Murphy ME. Polymorphic variants in the p53 pathway. Cell Death Differ 2006;13:916–20.
21. Bougeard G, Baert-Desurmont S, Tournier I, et al. Impact of the MDM2 SNP309 and p53 Arg72Pro polymorphism on age of tumour onset in Li-Fraumeni syndrome. J Med Genet 2006;43:531–3.
22. Ruijs MW, Schmidt MK, Nevanlinna H, et al. The single-nucleotide polymorphism 309 in the MDM2 gene contributes to the Li-Fraumeni syndrome and related phenotypes. Eur J Hum Genet 2007;15:110–4.
23. Swinney RM, Hsu SC, Hirschman BA, Chen TT, Tomlinson GE. MDM2 promoter variation and age of diagnosis of acute lymphoblastic leukemia. Leukemia 2005;19:1996–8.
24. Huang SF, Chen IH, Liao CT, Wang HM, Liou SH, Hsieh LL. Combined effects of MDM2 SNP 309 and p53 mutation on oral squamous cell carcinomas associated with areca quid chewing. Oral Oncol 2009;45:16–22.
25. Nakashima M, Kondo S, Shimizu Y, et al. Impact of MDM2 single nucleotide polymorphism on tumor onset in head and neck squamous cell carcinoma. Acta Otolaryngol 2008;128:808–13.
26. Menin C, Scaini MC, De Salvo GL, et al. Association between MDM2–309 and age at colorectal cancer diagnosis according to p53 mutation status. J Natl Cancer Inst 2006;98:285–8.
27. Wasielewski M, Nagel JH, Brekelmans C, et al. MDM2 SNP309 accelerates familial breast carcinogenesis independently of estrogen signaling. Breast Cancer Res Treat 2006;.
28. Sanchez-Carbayo M, Socci ND, Kirchoff T, et al. A polymorphism in HDM2 (SNP309) associates with early onset in superficial tumors, TP53 mutations, and poor outcome in invasive bladder cancer. Clin Cancer Res 2007;13:3215–20.
29. Bartel F, Jung J, Bohnke A, et al. Both germ line and somatic genetics of the p53 pathway affect ovarian cancer incidence and survival. Clin Cancer Res 2008;14:89–96.
30. Khatri RG, Navaratne K, Weil RJ. The role of a single nucleotide polymorphism of MDM2 in glioblastoma multiforme. J Neurosurg 2008;109:842–8.
31. Yoon YJ, Chang HY, Ahn SH, et al. MDM2 and p53 polymorphisms are associated with the development of hepatocellular carcinoma in patients with chronic hepatitis B virus infection. Carcinogenesis 2008;29:1192–6.
32. Voight BF, Kudaravalli S, Wen X, Pritchard JK. A map of recent positive selection in the human genome. PLoS Biol 2006;4:e72.
33. Sabeti PC, Reich DE, Higgins JM, et al. Detecting recent positive selection in the human genome from haplotype structure. Nature 2002;419:832–7.
34. Toro JR, Travis LB, Wu HJ, Zhu K, Fletcher CD, Devesa SS. Incidence patterns of soft tissue sarcomas, regardless of primary site, in the surveillance, epidemiology and end results program, 1978–2001: An analysis of 26,758 cases. Int J Cancer 2006;119:2922–30.
35. Atwal GS, Rabadan R, Lozano G, et al. An information-theoretic analysis of genetics, gender and age in cancer patients. PLoS ONE 2008;3:e1951.
36. Bond GL, Menin C, Bertorelle R, Alhopuro P, Aaltonen LA, Levine AJ. MDM2 SNP309 accelerates colorectal tumour formation in women. J Med Genet 2006;43:950–2.
37. Chen W, Arroyo JD, Timmons JC, Possemato R, Hahn WC. Cancer-associated PP2A Aα subunits induce functional haploinsufficiency and tumorigenicity. Cancer Res 2005;65:8183–92.
38. Colella S, Ohgaki H, Ruediger R, et al. Reduced

expression of the Aα subunit of protein phosphatase 2A in human gliomas in the absence of mutations in the Aα and Aβ subunit genes. Int J Cancer 2001;93:798–804.

39. Mumby M. PP2A: unveiling a reluctant tumor suppressor. Cell 2007;130:21–4.

40. Ruediger R, Pham HT, Walter G. Alterations in protein phosphatase 2A subunit interaction in human carcinomas of the lung and colon with mutations in the A β subunit gene. Oncogene 2001;20:1892–9.

41. Ruediger R, Pham HT, Walter G. Disruption of protein phosphatase 2A subunit interaction in human cancers with mutations in the A α subunit gene. Oncogene 2001;20:10–5.

42. Westermarck J, Hahn WC. Multiple pathways regulated by the tumor suppressor PP2A in transformation. Trends Mol Med 2008;14:152–60.

43. Dohoney KM, Guillerm C, Whiteford C, et al. Phosphorylation of p53 at serine 37 is important for transcriptional activity and regulation in response to DNA damage. Oncogene 2004; 23:49–57.

44. Kimura SH, Nojima H. Cyclin G1 associates with MDM2 and regulates accumulation and degradation of p53 protein. Genes Cells 2002;7: 869–80.

45. Li HH, Cai X, Shouse GP, Piluso LG, Liu X. A specific PP2A regulatory subunit, B56γ, mediates DNA damage-induced dephosphorylation of p53 at Thr55. EMBO J 2007;26:402–11.

46. Moule MG, Collins CH, McCormick F, Fried M. Role for PP2A in ARF signaling to p53. Proc Natl Acad Sci U S A 2004;101:14063–6.

47. Okamoto K, Li H, Jensen MR, et al. Cyclin G recruits PP2A to dephosphorylate Mdm2. Mol Cell 2002;9:761–71.

48. Arroyo JD, Hahn WC. Involvement of PP2A in viral and cellular transformation. Oncogene 2005;24:7746–55.

49. Junttila MR, Puustinen P, Niemela M, et al. CIP2A inhibits PP2A in human malignancies. Cell 2007;130:51–62.

50. Neviani P, Santhanam R, Trotta R, et al. The tumor suppressor PP2A is functionally inactivated in blast crisis CML through the inhibitory activity of the BCR/ABL-regulated SET protein. Cancer Cell 2005;8:355–68.

51. Sablina AA, Chen W, Arroyo JD, et al. The tumor suppressor PP2A Abeta regulates the RalA GTPase. Cell 2007;129:969–82.

52. Martens E, Stevens I, Janssens V, et al. Genomic organisation, chromosomal localisation tissue distribution and developmental regulation of the PR61/B′ regulatory subunits of protein phosphatase 2A in mice. J Mol Biol 2004;336: 971–86.