Association between TP53 R249S mutation and polymorphisms in TP53 intron 1 in hepatocellular carcinoma

Article *in* Genes Chromosomes and Cancer · October 2013

DOI: 10.1002/gcc.22086 · Source: PubMed

CITATIONS
13

READS
184

12 authors, including:

Sandra Ortiz-Cuaran
Centre Léon Bérard
86 PUBLICATIONS **2,237** CITATIONS
SEE PROFILE

Stéphanie Villar
International Agency for Research on Cancer
52 PUBLICATIONS **1,359** CITATIONS
SEE PROFILE

Marlin D Friesen
Johns Hopkins Bloomberg School of Public Health
123 PUBLICATIONS **6,295** CITATIONS
SEE PROFILE

Geoffroy Durand
Lund University
100 PUBLICATIONS **1,903** CITATIONS
SEE PROFILE

Association Between *TP53* R249S Mutation and Polymorphisms in *TP53* Intron 1 in Hepatocellular Carcinoma

Sandra Ortiz-Cuaran,[1] David Cox,[2] Stéphanie Villar,[1] Marlin D. Friesen,[3] Geoffroy Durand,[1] Amélie Chabrier,[1] Thiravud Khuhaprema,[4] Suleeporn Sangrajrang,[4] Simona Ognjanovic,[5] John D. Groopman,[3] Pierre Hainaut,[1,6]* and Florence Le Calvez-Kelm[1]

[1]International Agency for Research on Cancer, Lyon, France
[2]Cancer Research Center of Lyon, INSERM UI052, Lyon, France
[3]Johns Hopkins University, Bloomberg School of Public Health, Baltimore, MD
[4]National Cancer Institute, Bangkok, Thailand
[5]University of Minnesota, Masonic Cancer Center, Minneapolis, MN
[6]International Prevention Research Institute, Lyon, France

Over 100 single nucleotide polymorphisms (SNP) are validated in the *TP53* tumor suppressor gene. They define haplotypes, which may differ in their activities. Therefore, mutation in cancer may occur at different rates depending upon haplotypes. However, these associations may be masked by differences in mutations types and causes of mutagenesis. We have analyzed the associations between 19 SNPs spanning the *TP53* locus and a single specific aflatoxin-induced *TP53* mutation (R249S) in 85 in hepatocellular carcinoma cases and 132 controls from Thailand. An association with R249S mutation ($P = 0.007$) was observed for a combination of two SNPs (rs17882227 and rs8064946) in a linkage disequilibrium block extending from upstream of exon 1 to the first half of intron 1. This domain contains two coding sequences overlapping with *TP53* (*WRAP53* and *Hp53int1*) suggesting that sequences in *TP53* intron 1 encode transcripts that may modulate R249S mutation rate in HCC. © 2013 Wiley Periodicals, Inc.

INTRODUCTION

The *TP53* tumor suppressor gene (MIM# 191170, tumor protein p53) is frequently altered by missense mutations in human cancers (Petitjean et al., 2007). The *TP53* locus covers 21,451 kb on 17p13.1 (chr17: 7590863–7571720, reference: GRCh37.p5) and contains about 100 validated naturally occurring single nucleotide polymorphisms (SNPs; listed at http://p53.iarc.fr/TP53GeneVariations.aspx). Only a small fraction of these SNPs have been reported to cause measurable perturbations to p53 tumor suppressor functions.

The distribution of several common SNPs show significant geographical and population variations (Whibley et al., 2009; Grochola et al., 2010). The best characterized exon polymorphism is a nonsilent substitution SNP at codon 72 in exon 4 (rs1042522, NC_000017.9:g.7520197G>C, p.R72P) that shows variation across populations and influences tumor suppressor activities of p53. Several studies have shown that the 72P variant has higher capacity than 72R for transactivating *CDKN1A* and induction growth arrest (Pim and Banks, 2004; Salvioli et al., 2005). Alternatively, the 72R variant has increased capacity to relocate in the mitochondria in tumor cell lines (Dumont et al., 2003). Furthermore, the 72R variant is associated with higher expression of the embryonic implantation factor LIF and lower rate of *in vitro* fertilization failure (Kang et al., 2009; Feng et al., 2011). An association has been reported between low winter temperature and the 72R variant in a cohort of 4029 individuals across Eastern Asia, suggesting selection by adaptation to cold climates (Shi et al., 2009). The most studied intron polymorphism is a 16 bp duplication in intron 3 (rs17878362, NM_000546.4:c.97-54delGinsACCTGGAGGGCTGGGG, PIN3). There is evidence that this polymorphism has an effect on p53 mRNA

Additional Supporting Information may be found in the online version of this article.

Supported by: IARC; Ligue contre le Cancer, Département du Rhône, France.

*Correspondence to: Pierre Hainaut, International Prevention Research Institute, 15 Chemin du Saquin, 69130 Ecully, France. E-mail: pierre.hainaut@i-pri.org

Received 29 November 2012; Accepted 5 June 2013

DOI 10.1002/gcc.22086

Published online 1 July 2013 in Wiley Online Library (wileyonlinelibrary.com).

splicing patterns and overall mRNA levels (Gemignani et al., 2004; Marcel et al., 2011). This insertion is associated with increased risk of breast, colorectal, lung, and other cancers (Hu et al., 2010, Sagne et al., 2013). Thus, functional and genetic association studies support that polymorphisms may modulate several aspects of p53 suppressor activity, perhaps in a tissue and context-dependent manner (Whibley et al., 2009). It follows that defined *TP53* alleles might be more prone to undergo somatic inactivation by mutation than others, and the resulting loss of p53 function may depend not only upon the structural and functional effect of each specific mutation but also on the haplotype of the wild-type allele. This hypothesis is supported by results from Mechanic et al. (2007) who analyzed 14 polymorphisms in a case-only study of lung cancer in patients recruited in the greater Baltimore area, Maryland. They found an association with several combinations of SNPs and risk of somatic mutation. The SNPs identified as associated with mutation were distributed in a region encompassing intron 2 to intron 7 and included rs1042522, rs9895829, rs1625895 and rs12951053; haplotype Pro-T-A-G-G). However, SNPs located in the large exon 1 and in the promoter of *TP53* were not analyzed (Mechanic et al., 2007).

To further address whether mutation occurs at different rates depending on the polymorphic nature of the *TP53* gene, we have analyzed the distribution of a selection of 19 common SNPs along the *TP53* locus in subjects recruited in a case-referent study of hepatocellular carcinoma (HCC) conducted in Thailand. In this country, about 30 to 40% of HCC are caused by the association of chronic hepatitis B virus (HBV) infection and aflatoxin (AFB_1), the remaining cases being attributable to a range of etiological factors including alcohol, HCV infections, and metabolic liver syndromes (Galy et al., 2011; Villar et al., 2012; Ortiz et al., 2013). The synergistic effect of AFB_1 and HBV chronic infection causes the formation of a single *TP53* mutation at codon 249 (c.747G>T, R249S, mutant protein p.R249S), resulting from the specific binding of AFB_1 metabolites on the 3rd base of codon 249 (Besaratinia et al., 2009). We have previously shown by sequencing tumor DNA that R249S is detectable in about 30% of HCC cases from Thailand (Galy et al., 2011). Using a quantitative method to assess the presence of this mutation in free circulating plasma DNA, we have detected high concentrations of R249S mutated DNA (>150 copies/ml plasma) in 44% HCC patients with no history of liver cirrhosis before HCC diagnosis and in 20% of patients with HCC developing in a context of pre-existing cirrhosis (Villar et al., 2012). Thus, the presence of high concentrations of mutant DNA in the plasma is strongly associated with HCC status. In contrast, R249S mutant DNA is also detected in a small proportion (3%) of non-cancer (control) subjects from Thailand (Villar et al., 2012). Presence of mutant DNA in the plasma of controls may be a transient phenomenon, likely to reflect mutagenesis by ongoing exposure to AFB_1. In agreement with this hypothesis, we have previously reported R249S DNA in the plasma of non-cancer subjects of West Africa, with quantitative seasonal variations that may result from the seasonal pattern of dietary exposure to AFB_1 (Villar et al., 2011).

We have reasoned that this particular context in which a single mutation type is caused by a specific exposure would be less heterogeneous to analyze the association between mutation and *TP53* SNPs than most other cancers in which *TP53* mutations are structurally and functionally diverse and may occur through different mechanisms at different stages of tumor progression.

MATERIALS AND METHODS

Study Participants

This study was performed using specimens from patients and controls recruited in the context of the International Liver Cancer Study (ILCS), a network of liver cancer studies that uses standardized protocols for sample and data collection (protocols and informed consent forms available on request). ILCS Thailand is a case—referent study in which subjects were recruited from fifty-five provinces across the country from April 2008 to December 2009. We have analyzed 50 cases of HCC without cirrhosis (HCC/no cirrhosis), 35 cases of HCC with cirrhosis (HCC/cirrhosis), and 132 subjects without liver disease, constituting the reference group (Table 1). Diagnosis of HCC was based on concordant clinical examination, abdominal imaging, and plasma levels of alfafoetoprotein >100 ng/ml. Individuals assigned to the reference group presented no clinical evidence of liver disease and were selected among the group of patients that came to the National Cancer Institute, Bangkok, for annual check-up, with similar geographic distribution as for cases. Each participant was interviewed by a trained interviewer and blood and plasma samples were collected for

laboratory examination through oral informed consent. The study was approved by the National Cancer Institute review board (Bangkok, Thailand) and the Ethical Committee of the International Agency for Research of Cancer.

Quantitation of R249S Mutation in Circulating Free DNA

The concentrations of R249S mutant DNA for this study population were reported previously (Villar et al., 2012). Briefly, circulating free DNA (CFDNA) was extracted from 1 ml of plasma using QiAmp circulating nucleic acid Kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions. Purified CFDNA was eluted from the QiAmp Silica column with two volumes (2×50 µl) of water (PCR-grade, Sigma-Aldrich, St Louis, MO). Quantitation of extracted DNA was performed by fluorimetry using PicoGreen (Molecular Probes, Eugene, OR). R249S was detected and quantified against a synthetic, internal standard plasmid by short oligonucleotide mass analysis (SOMA) as described previously (Lleonart et al., 2005; Villar et al., 2011). Plasma concentrations were expressed as copies of R249S-mutated DNA per ml of plasma. The limit of determination for the method was 67 copies of R249S-mutated DNA/ml plasma. Consistent with previous studies, a cutoff level of 150 copies of R249S mutated DNA per ml plasma was used (Villar et al., 2012).

TP53 SNPs Genotyping

Genomic DNA was extracted from buffy coat samples and amplified by whole genome amplification (WGA) using GenomiPhi V2 DNA Amplification kit (GE Healthcare Bio-Sciences AB, Uppsala, Sweden). To evaluate the possibility of errors introduced by WGA, we compared genotyping data of a common SNP in the ataxia telangiectasia mutated (*ATM*) gene in a series of independent triplicate WGA and genomic DNAs from 1536 samples of different geographic origin, which showed concordance rates between 98 and 99.8%. In parallel, Illumina Goldengate (Illumina, San Diego, CA) genotyping of 384 common SNPs has been used to compare genomic and WGA DNA from 32 HapMap (Caucasian) samples, with concordance rates over 98%, depending on the SNP considered. Amplified DNA was quantified by PicoGreen and normalized to 6 ng/µl in order to proceed with genetic variations analysis. SNP selection was done according to the tSNPs suggested for the HapMap Asian population and completed with other SNP in order to construct a more complete linkage disequilibrium (LD) block. A set of 19 SNPs was selected for genotyping, including 16 SNPs previously identified for haplotyping *TP53*, which are common to Caucasians and Asians and 3 SNPs, which are specific for Asian populations (Garritano et al., 2010). Of these SNPs, 14 *TP53* Tag SNPs were genotyped as previously described using high resolution melting (HRM) technique, with some modifications (Garritano et al., 2010). Briefly, HRM was performed from a first PCR reaction followed by a nested asymmetric PCR where the same primers were used and an unlabeled-probe designed to perfectly match the most frequent allele of *TP53* SNPs was added. The first PCR reaction (8 µl) contained 30 ng of template DNA, 66.25 µM dNTP (Invitrogen, Carlsbad, CA), 200 nM forward and reverse primers (Sigma Aldrich, St Louis, MO), 0.054 µl of Accustart Taq Polymerase, 1.5 mM Accustart $MgCl_2$, and 10x PCR buffer (Quanta Biosciences, Gaithersburg, MD). The further nested PCR reaction (6 µl) contained 2 µl of a 1/4 dilution of the first PCR product, 66.25 µM dNTP, forward primer and a five-time excess of reverse primer, and unlabeled probe (Sigma Aldrich, St Louis, MO), 1x LC Green dye (Biochem, Salt Lake City, UT), 0.054 µl of Accustart Taq Polymerase, 1.5 mM Accustart $MgCl_2$, and 10x PCR buffer (Quanta Biosciences). Each HRM genotyping experiment included replicates for 96 randomly selected samples. Concordance rate varied between 90.6% and 97.9%, depending on the SNP considered.

SNPs 93, 95, and 96 were genotyped by direct sequencing. Purification/Sequencing/G50 Purification (10 µl of mix) using 0.4 µl (10 U/µl) Exonuclease 1 enzyme, 0.2 µl (1 U/µl) Shrimp Alkaline Phosphatase enzyme (USB Products, Affimetrix, Cleveland, OH), and 9.4 µl TE-4 buffer, sequencing reaction (3 µl of mix) using 0.5 µl water (Eurobio, Courtaboeuf), 1 µl 5X BigDye Buffer, 0.5 µl BigDye V3.1 (Applied Biosystems, Austin, TX), and 1 µl of 5 µM F or R sequencing primer (Sigma Aldrich). After sequence purification on a Multiscreen Filter plate (Millipore, Billerica, MA), sequencing was performed on a Spectrumedix Instruments 96-capillaries sequencer. Primers and probe sequences for HRM and primers for direct sequencing are provided in Supporting Information Table 1. SNPs 3 and 105 were genotyped using a TaqMan SNP genotyping assay (see

TABLE 1. Characteristics of Study Participants

		Status			
		HCC/no cirrhosis *n* (%)	HCC/cirrhosis *n* (%)	Reference *n* (%)	*P* Fisher-Exact Test
Total		50 (100.0)	35 (100.0)	132 (100.0)	
Age	<50	18 (36.0)	7 (20.0)	33 (25.0)	$P = 0.20$
	≥ 50	32 (64.0)	28 (80.0)	99 (75.0)	
Gender	Men	41 (82.0)	27 (77.1)	77 (58.3)	$P = 0.04$
	Women	9 (18.0)	8 (22.9)	55 (41.7)	
HBsAg[a]	Negative	20 (40.0)	16 (47.7)	111 (84.1)	$P < 0.001$
	Positive	30 (60.0)	17 (48.6)	19 (14.4)	
	N/A	0 (0.0)	2 (5.7)	2 (1.5)	
HCV Ab	Negative	43 (86.0)	23 (65.7)	132 (100.0)	–
	Positive	7 (14.0)	11 (31.4)	0 (0.0)	
	N/A	0 (0.0)	1 (2.9)	0 (0.0)	
R249S[b]	<150	28 (56.0)	28 (80.0)	128 (97.0)	$P < 0.001$
	≥150	22 (44.0)	7 (20.0)	4 (3.0)	

[a]Hepatitis B surface antigen; (OR [95% CI]: Reference: 1; HCC/no cirrhosis: 10.93 [4.7–25.2]; HCC/cirrhosis: 11.42 [4.2–31.2]. ORs were adjusted for age and gender.
[b]*TP53* mutation at codon 249 (R249S) copy numbers/ml plasma.

catalog references in Supporting Information Table 1). All amplicons were bidirectionally sequenced, and the electropherograms of both strands were independently double-checked.

LD plots were constructed for the reference groups using Haploview 4.2, with the tSNPs that had a genotyping completeness rate of 94% or higher, excluding SNPs which violated Hardy–Weinberg equilibrium (cut-off = 0.0010). Because of the limited power of the study, no attempts were made to examine further the possible association with haplotypes defined by multiple SNPs.

Statistical Analyses

Univariate analysis was conducted using online statistical tools available at VassarStats (http://vassarstats.net/). Adjusted Odds ratios (ORs), X_2, and *P*-values were calculated using STATA 11.1 (College Station, Texas). Odds ratios were adjusted for sex and age. Hardy—Weinberg equilibrium for each SNP was assessed using an online calculator (Court Lab-HW calculator, https://www.tufts.edu) with an alpha of 0.05. The alpha same alpha was taken as significant for all other analyses. Post-hoc power calculations were conducted to qualify better the statistical limitations due to small sample size using calculation tools available at http://homepage.stat.uiowa.edu/~rlenth/Power/index.html). In a case-referent design ($n = 85$ for cases and 132 for controls), OR of association between 2.4 and 3.3 could be detected with power = 80%, alpha = 0.05 (two-sided test) for SNP with minor allele frequencies of 0.1–0.4. In a case–case design (29 R249S-postive and 56 R249S negative cases) power was not sufficient to detect OR of association of haplotypes with R249S of less than 4.4.

RESULTS

Genotyping

A total of 19 SNPs were genotyped in the study population described in Table 1. Results showed that all SNPs except SNP 137 (rs1641549, NC_000017.10:g.7574775C>T) were in Hardy–Weinberg equilibrium (alpha = 0.05). A LD plot constructed with Haploview based on genotyping in controls is shown in Figure 1. Two defined LD blocks were identified. The first one encompassed 4 SNPs from SNP 6 (chr17:7591511) to SNP 23 (chr17:7588006) starting in the 5' untranscribed region upstream of exon 1 and extending into the first half of intron 1. Within this block, SNPs were in very strong pairwise LD, with SNP 6 (rs17551157, NC_000017.10:g.7591511_7591512 insG) in complete LD with SNP 20 (rs17882227, NG_017013.1:g.7304T>C) and SNP 15 (rs8064946, NG_017013.1:g.6553C>G) in complete LD with SNP 23 (rs9909978, NT_010718.16:g.1419963C>T). These two SNP pairs were overlapping, with SNP 15 intercalated between SNP 6 and SNP 20 (Fig. 2). The second LD block, less

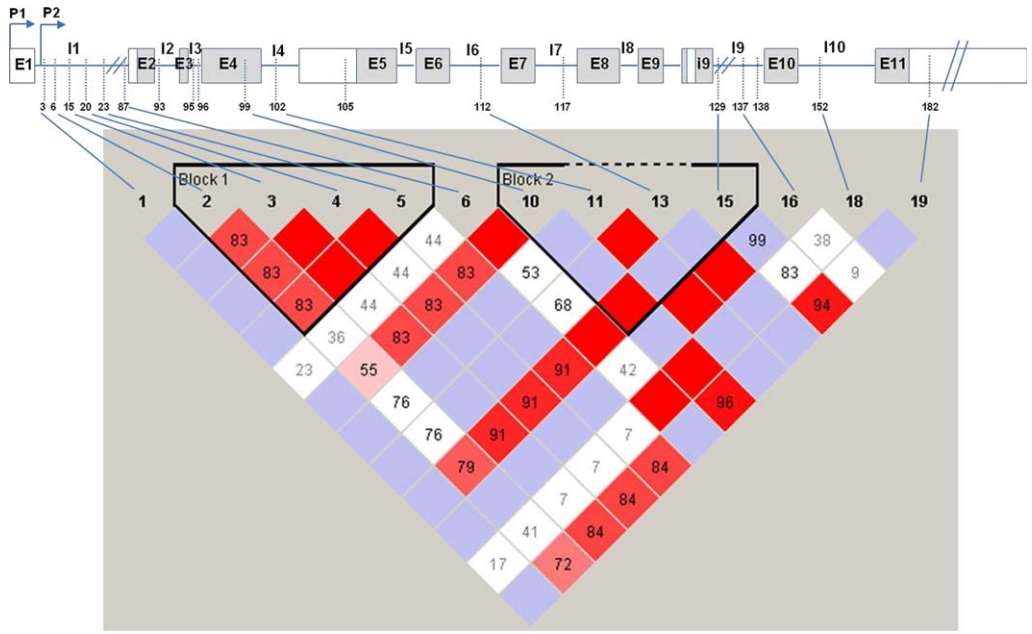

Figure 1. LD plot for 19 *TP53* SNPs. LD plots were constructed for the reference group using [DEFAULT] algorithm of Haploview 4.2 (genotyping percentage $\geq$ 94% and Hardy–Weinberg equilibrium cut-off = 0.0010). The first block encompassed 4 SNPs from SNPs 6 (rs17551157) to SNP 23 (rs9909978) extending from the 5' untranscribed region upstream of exon 1 to the first half of intron 1. The second block encompassed 4 SNPs between SNP99 (rs1042522) and SNP129 (12949655) and corresponded to a region extending from exon 4 to intron 9. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

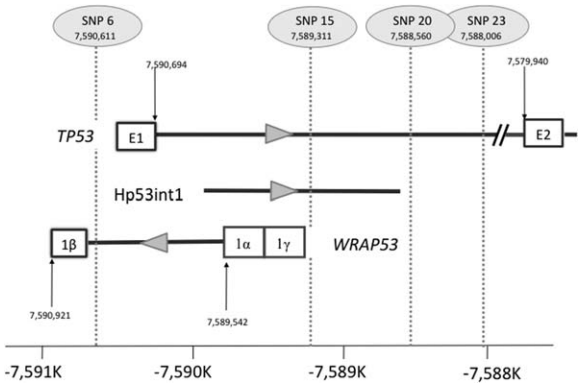

Figure 2. Schematic map of *TP53* intron 1: overlap with *WRAP53*, *Hp53int1*, and position of SNPs associated with R249S mutation. SNP 6 (rs17551157) falls in the alternative intron 2 next to exon1beta of *WRAP53*. SNP 15 (rs8064946) is located 78 bp upstream of exon1-gamma of *WRAP53* and falls upstream of the largest ORF of *Hp53int1*. Finally, the SNP 20 (rs17882227) falls in the 3' region 45 bp downstream of the end of *Hp53int1*.

defined, encompassed 4 SNPs from SNP 99 (exon 4; chr17: 7579472; NC_000017.9:g.7520197G>C, p.R72P) to SNP 129 (intron 9; chr17: 7575733; rs12949655, NG_017013.1:g.20131T>C).

Association with HCC

The characteristics of population for case-referent analysis are described in Table 1. Patients with HCC were grouped in two categories, depending on presence or absence of cirrhosis before liver cancer. Patients who developed HCC without prior evidence of cirrhosis were classified as "HCC/no cirrhosis." The median age was 52 (33–71) for HCC/no cirrhosis and 55 years (range 42–74) for HCC/cirrhosis. Noteworthy, 38% of HCC/no cirrhosis patients were under 50 years of age compared with 18% for patients with HCC/cirrhosis ($P = 0.045$). Male to female ratio was 4:1 for HCC/no cirrhosis and 3:1 for HCC/cirrhosis ($P = 0.04$ when compared with reference group). As expected, a strong association was observed between case status and positivity for Hepatitis B surface Antigen (HBsAg; Adjusted Odd Ratios (OR; [95% CI]): Reference: 1; HCC/no cirrhosis: 10.93 [4.7–25.2]; HCC/cirrhosis: 11.42 [4.2–31.2]) ($P < 0.001$). The prevalence of HCV was higher in HCC/cirrhosis than in HCC/no cirrhosis ($P = 0.044$). Of note, none of the subjects in the reference group was positive for HCV. *TP53* R249S mutation detection in the plasma in these cases and controls has been reported in a previous study (Villar et al., 2012). In agreement with this previous study, we have used a cut-off level of $\geq$150 copies/ml to distinguish HCC that were significantly associated with the R249S mutation from those considered as not associated with mutations. Levels of R249S of 150 copies/ml of plasma or above were detected in 34.1% of all HCC cases

TABLE 2. Allele Distribution of SNP15 and SNP20 According to *TP53* R249S Status

	Cases R249S	HCC (all) <150	HCC (all) ≥150	P^a (Fisher-Exact)	HCC (no cirrhosis) <150	HCC (no cirrhosis) ≥150	P^a (Fisher-Exact)
	Genotype	$n = 82$			$n = 50$		
	GG	20	11	$P = 0.21$	12	9	$P = 0.05$
	GC	22	15		8	12	
	CC	12	2		8	1	
SNP15 (dominant)	GG+GC	42	26	$P = 0.08$	20	21	$P = 0.03$
	CC	12	2		8	1	
SNP20	Genotype	$n = 76$			$n = 46$		
	TT	24	9	$P = 0.12$	11	8	$P = 0.13$
	CT	17	8		9	6	
	CC	8	10		3	9	
SNP20 (dominant)	TT+CT	41	17	$P = 0.04$	20	14	$P = 0.05$
	CC	8	10		3	9	

[a]1-tailed test.

TABLE 3. Association of R249S Status with "at Risk" Combinations Defined by SNP15 and SNP20

	Numbers at risk/not at risk[a]			OR for association with R249S ≥ 150 copies/ml plasma (95%CI)
HCC (all)	7/11	1 (Ref)	10/1	15.71 (1.63 − 151.11)
HCC (no cirrhosis)	3/7	1 (Ref)	9/1	21 (1.78 − 248.11)

[a]"at risk" combination: SNP15 GG or GC and SNP20 CC; "not at risk" combination: SNP15 CC and SNP20 TT or CT

(44% in HCC/no cirrhosis, 20% in HCC/cirrhosis) and in 3% of controls (P for distribution across the 3 groups: <0.001).

Analysis of genotype frequencies of each SNP with HCC status showed an association with HCC status (grouping HCC/no cirrhosis and HCC/cirrhosis) only for the 16 bp insertion polymorphism in intron 3 (PIN3; rs17878362) for heterozygous individuals when compared with the frequent homozygotes (carrying two alleles with one copy of the 16 bp motif) (Adjusted OR: 6.03, 95%CI [1.03–35.08]). All other SNPs had similar allele distributions between cases and controls (Supporting Information Table 2).

Association with High Plasma Concentrations of R249S Mutated DNA

We next conducted a case-only analysis comparing HCC cases without prior diagnosis of cirrhosis that had ≥150 copies/ml or <150 copies/ml for R249S in plasma samples. Among the 19 SNP studied only alleles of SNPs in LD block 1 showed a tendency for association with R249S status (Table 2). A borderline significant association with R249S at ≥150 copies/ml plasma was observed with the G allele of SNP 15 in HCC/no cirrhosis ($P = 0.05$). A trend toward the same effect, however not significant, was observed in all cases [HCC (all), grouping HCC with and without cirrhosis]. No significant effects were seen for allele distributions of SNP 20. The same effects were observed for SNP 6 and SNP 23, which are in complete LD with SNP 20 and SNP 15, respectively. Next, we analyzed allele distributions using a dominant model in which homozygote and heterozygote carriers of the most common allele are grouped as one category. This analysis identified a borderline significant association between R249S status and presence of the major G allele of SNP 15 (GG + GC status, HCC (all): $P = 0.08$; HCC/no cirrhosis: $P = 0.03$). For SNP 20, the inverse association was observed, with a significantly greater proportion of subjects homozygote for the minor C allele being R249S –positive, when compared with carriers of the major T allele (TT+ CT; HCC (all): $P = 0.04$; HCC/no cirrhosis: $P = 0.05$). These observations led us to identify two combinations of alleles defined by SNP 15 and SNP 20. The combination defined by SNP 15 GG or GC and SNP 20 CC represents a combination "at risk" for R249S mutation, whereas the one defined by SNP 15 CC and SNP 20 TT or CT represent the reciprocal, "not at risk" combination. Table 3 shows an estimate of the Odds Ratio (OR) for being positive for R249S mutation (≥150 copies/ml) in relation with these combinations. Taking the "not at risk" combination as reference, OR for positivity for R249S were of 15.71 and 21 for the "at risk" SNP combination for HCC (all) and HCC (no cirrhosis), respectively. Despite the wide 95% confidence

interval, these associations were significant (HCC (all): $P = 0.007$; HCC/no cirrhosis: $P = 0.01$). Importantly, when considering the association with HCC risk in a case-control design, these SNP combinations did not show any association with case status (OR for HCC (all): 1.067; 95%CI [0.67–1.49]; for HCC/no cirrhosis: 1.2; 95%CI [0.48-3.02]). Thus, we conclude that the SNP combination defined by SNP 15 GG or GC and SNP 20 CC is specifically associated with the risk of having levels of R249S $\geq$150 copies/ml in the plasma, but is not associated with the risk of HCC.

DISCUSSION

These results further support that the polymorphic status of *TP53* may influence the occurrence of *TP53* mutation. In a previous study, Mechanic et al. (2007) identified an association between several combination of SNPs located from intron 2 to intron 7 and increased odd ratio for somatic mutation in lung cancer. In this study, we have analyzed HCC cases and controls and we have taken into consideration a set of SNPs that covers the entire *TP53* coding sequence. We have detected an association between presence of high levels of aflatoxin-related R249S mutation and SNPs in strong LD located in a domain that starts immediately upstream of the noncoding exon 1 or *TP53* and extends into the proximal part of intron 1. Considering two sentinel SNPs as representative for the whole domain, we have detected an association between positivity for R249S in HCC and a combination of the major allele of SNP 15 (GG or GC) and the rare allele of SNP 20 (CC). Although the study population is small, the association observed is sufficiently large to reach statistical significance even taking into account the power limitations of the study. The association was seen for HCC (all cases), as well as for a subgroup of HCC with no diagnosis of cirrhosis before cancer. In a previous study, we have shown that high levels of R249S in the plasma were more common in those HCC with no cirrhosis, although such levels could also be observed at a lower frequency in HCC occurring in a context of cirrhosis (Villar et al., 2012). Because of the small size of the study, it has not been possible to reconstitute further the haplotypes potentially associated with R249S. Interestingly, these SNPs were not associated with the risk of HCC although they were, collectively, associated with R249S.

Our results suggest that elements contained in the LD block 1 of the *TP53* gene may impact on the occurrence of R249S mutation in *TP53*. As this region is located over 10 kb upstream of exon 7 that contains codon 249, it is unlikely to contribute to a sequence context influencing mutagenesis or DNA repair at codon 249. However, there is a basis to suggest that specific elements in LD block 1 may contribute to the regulation of *TP53* function. This region contains two sequences that overlap with *TP53*, *WRAP53* and *Hp53int1* (Fig. 2). Initially identified as a natural antisense transcript that upregulates p53 expression, *WRAP53* (MIM# 612661, gene locus 17p13) is also known as *TCAB1* (telomerase Cajal body protein-1) and is mutated in Dyskeratosis Congenita (Mahmoudi et al., 2009; Batista et al., 2011). The TCAB1 protein interacts with components of active telomerase dyskerin, TERT, and TERC and with small Cajal body RNAs (scaRNAs), which are involved in splicing regulation. Depletion of *TCAB1* by RNA interference prevents TERC from associating with Cajal bodies, disrupts telomerase–telomere association and prevents telomere elongation (Mahmoudi et al., 2010). Aside from this role, downregulation of *WRAP53* results in a significant decrease in p53 mRNA and in suppression of p53 induction in response to DNA damage. This effect has been attributed to the capacity of *WRAP53* mRNA to hybridize with and stabilize p53 mRNA (Mahmoudi et al., 2010). SNP 6 (rs17551157), located upstream of *TP53* noncoding exon 1, falls in the alternative intron 2 that immediately follows exon1beta of *WRAP53*, whereas SNP 15 (rs8064946) is located 78 bp upstream of exon1gamma. The effect, if any, of SNP 6 (rs17551157) and SNP 15 (rs8064946) on *WRAP53* expression and splicing is unknown. *Hp53int1* (D17S2179E, GenBank: U58658.1) has been identified by Reisman et al. (1988, 1996) as a 1125 bp polyadenylated cDNA corresponding to a mRNA transcribed from the p53P2 promoter, located ~1000 bp downstream of the 3'end of the noncoding exon 1 within the 10,738 bp-long intron 1. The *Hp53int1* gene is oriented in the same direction as *TP53* and is expressed in myeloid leukemia cells HL-60 and U937 in a differentiation-dependent manner. The largest open reading frame (on the +1 frame of the cDNA) encodes a putative 118 residues protein with a region of homology to part of the DNA-binding domain of NFkappaB. SNP 15 (rs8064946) falls at position 32 in a sequence of 43 residues located upstream of the largest ORF (Reisman et al., 1996). The base change (CCA to GCA) is potentially nonsilent (proline to alanine) but the corresponding

sequence lacks a conventional initiation codon. SNP 20 (rs17882227) is located 45 bp downstream of the 3' end of the 1125 bp cDNA. To date, there are no data on the functional significance of Hp53int1 and on its potential role as a regulator of p53 expression or activity.

We suggest that these two SNPs (rs8064946 and rs17882227) may modulate the potency of *TP53* alleles to exert wild type, suppressor effect (making its inactivation by mutation required for development of HCC), or mutant, gain-of-function effects (thus acting in a way that only specific mutant haplotypes may exert such oncogenic effects). However, considering the small sample size and the limited power of the study, in particular in the case-case comparison design, these findings will need to be replicated on a larger group of subjects study and, ideally, in other populations where R249S mutation in HCC is common. Further studies are needed to determine if a similar association also exists with *TP53* mutations occurring in other cancer types and/or in other exposure contexts.

ACKNOWLEDGMENTS

The authors acknowledge Amelie Plymoth (Karolinska Institute, Stockholm) and Anant Karalak (National Cancer Institute, Bangkok) for sample collection in Thailand.

REFERENCES

Batista LF, Pech MF, Zhong FL, Nguyen HN, Xie KT, Zaug AJ, Crary SM, Choi J, Sebastiano V, Cherry A, Giri N, Wernig M, Alter BP, Cech TR, Savage SA, Reijo Pera RA, Artandi SE. 2011. Telomere shortening and loss of self-renewal in dyskeratosis congenita induced pluripotent stem cells. Nature 474:399–402.

Besaratinia A, Kim SI, Hainaut P, Pfeifer GP. 2009. In vitro recapitulating of TP53 mutagenesis in hepatocellular carcinoma associated with dietary aflatoxin B1 exposure. Gastroenterology: 137:1127–1137.

Dumont P, Leu JI, Della PA, III, George DL, Murphy M. 2003. The codon 72 polymorphic variants of p53 have markedly different apoptotic potential. Nat Genet 33: 357–365.

Feng Z, Zhang C, Kang HJ, Sun Y, Wang H, Naqvi A, Frank AK, Rosenwaks Z, Murphy ME, Levine AJ, Hu W. 2011. Regulation of female reproduction by p53 and its family members. FASEB J 25:2245–2255.

Galy O, Chemin I, Le RE, Villar S, Le Calvez-Kelm F, Lereau M, Gouas D, Vieco B, Suarez I, Navas MC, Chevallier M, Norder H, Srivatanakul P, Karalak A, Sangrajrang S, Trepo C, Hainaut P. 2011. Mutations in TP53 and CTNNB1 in relation to Hepatitis B and C infections in hepatocellular carcinomas from Thailand. Hepat Res Treat 2011:697162.

Garritano S, Gemignani F, Palmero EI, Olivier M, Martel-Planche G, Le Calvez-Kelm F, Brugieres L, Vargas FR, Brentani RR, shton-Prolla P, Landi S, Tavtigian SV, Hainaut P, Achatz MI. 2010. Detailed haplotype analysis at the TP53 locus in p.R337H mutation carriers in the population of Southern Brazil: evidence for a founder effect. Hum Mutat 31:143–150.

Gemignani F, Moreno V, Landi S, Moullan N, Chabrier A, Gutierrez-Enriquez S, Hall J, Guino E, Peinado MA, Capella G, Canzian F. 2004. A TP53 polymorphism is associated with increased risk of colorectal cancer and with reduced levels of TP53 mRNA. Oncogene 23:1954–1956.

Grochola LF, Zeron-Medina J, Meriaux S, Bond GL. 2010. Single-nucleotide polymorphisms in the p53 signaling pathway. Cold Spring Harb Perspect Biol 2: a001032.

Hu Z, Li X, Qu X, He Y, Ring BZ, Song E, Su L. 2010. Intron 3 16 bp duplication polymorphism of TP53 contributes to cancer susceptibility: A meta-analysis. Carcinogenesis 31:643–647.

Kang HJ, Feng Z, Sun Y, Atwal G, Murphy ME, Rebbeck TR, Rosenwaks Z, Levine AJ, Hu W. 2009. Single-nucleotide polymorphisms in the p53 pathway regulate fertility in humans. Proc Natl Acad Sci USA 106:9761–9766.

Lleonart ME, Kirk GD, Villar S, Lesi OA, Dasgupta A, Goedert JJ, Mendy M, Hollstein MC, Montesano R, Groopman JD, Hainaut P, Friesen MD. 2005. Quantitative analysis of plasma TP53 249Ser-mutated DNA by electrospray ionization mass spectrometry. Cancer Epidemiol Biomarkers Prev 14:2956–2962.

Mahmoudi S, Henriksson S, Corcoran M, Mendez-Vidal C, Wiman KG, Farnebo M. 2009. Wrap53, a natural p53 antisense transcript required for p53 induction upon DNA damage. Mol Cell 33:462–471.

Mahmoudi S, Henriksson S, Weibrecht I, Smith S, Soderberg O, Stromblad S, Wiman KG, Farnebo M. 2010. WRAP53 is essential for Cajal body formation and for targeting the survival of motor neuron complex to Cajal bodies. PLoS Biol 8:e1000521.

Marcel V, Tran PL, Sagne C, Martel-Planche G, Vaslin L, Teulade-Fichou MP, Hall J, Mergny JL, Hainaut P, Van DE. 2011. G-quadruplex structures in TP53 intron 3: role in alternative splicing and in production of p53 mRNA isoforms. Carcinogenesis 32:271–278.

Ortiz-Cuaran S, Villar S, Gouas D, Ferro G, Plymoth A, Khuhaprema T, Kalalak A, Sangranjang S, Friesen M, Groopman JD, Hainaut P. 2013. Association between HBX status, aflatoxin-induced R249S TP53 mutation and risk of hepatocellular carcinoma in a case-control study from Thailand. Cancer Lett 331:46–51.

Petitjean A, Achatz MI, Borresen-Dale AL, Hainaut P, Olivier M. 2007. TP53 mutations in human cancers: Functional selection and impact on cancer prognosis and outcomes. Oncogene 26: 2157–2165.

Pim D, Banks L. 2004. p53 polymorphic variants at codon 72 exert different effects on cell cycle progression. Int J Cancer 108:196–199.

Reisman D, Greenberg M, Rotter V. 1988. Human p53 oncogene contains one promoter upstream of exon 1 and a second, stronger promoter within intron 1. Proc Natl Acad Sci USA 85: 5146–5150.

Reisman D, Balint E, Loging WT, Rotter V, Almon E. 1996. A novel transcript encoded within the 10-kb first intron of the human p53 tumor suppressor gene (D17S2179E) is induced during differentiation of myeloid leukemia cells. Genomics 38: 364–370.

Sagne C, Marcel V, Amadou A, Hainaut P, Olivier M, Hall J. 2013. A meta-analysis of cancer risk associated with the *TP53* intron 3 duplication polymorphism (rs17878362): Geographic and tumor-specific effects. Cell Death Dis 14: e492.

Salvioli S, Bonafe M, Barbi C, Storci G, Trapassi C, Tocco F, Gravina S, Rossi M, Tiberi L, Mondello C, Monti D, Franceschi C. 2005. p53 codon 72 alleles influence the response to anticancer drugs in cells from aged people by regulating the cell cycle inhibitor p21WAF1. Cell Cycle 4:1264–1271.

Shi H, Tan SJ, Zhong H, Hu W, Levine A, Xiao CJ, Peng Y, Qi XB, Shou WH, Ma RL, Li Y, Su B, Lu X. 2009. Winter temperature and UV are tightly linked to genetic changes in the p53 tumor suppressor pathway in Eastern Asia. Am J Hum Genet 84:534–541.

Villar S, Le Roux-Goglin E, Gouas DA, Plymoth A, Ferro G, Boniol M, Lereau M, Bah E, Hall AJ, Wild CP, Mendy M, Norder H, van der SM, Whittle H, Friesen MD, Groopman JD, Hainaut P. 2011. Seasonal variation in *TP53* R249S-mutated serum DNA with aflatoxin exposure and Hepatitis B virus infection. Environ Health Perspect 119:1635–1640.

Villar S, Ortiz-Cuaran S, bedi-Ardekani B, Gouas D, Nogueira da CA, Plymoth A, Khuhaprema T, Kalalak A, Sangrajrang S, Friesen MD, Groopman JD, Hainaut P. 2012. Aflatoxin-induced *TP53* R249S mutation in hepatocellular carcinoma in Thailand: Association with tumors developing in the absence of liver cirrhosis. PLoS One 7: e37707.

Whibley C, Pharoah PD, Hollstein M. 2009. p53 polymorphisms: Cancer implications. Nat Rev Cancer 9:95–107.