Romanian Biotechnological Letters *Vol. 23, No. 6, 2018*

ORIGINAL PAPER

Characterization of *TP53* polymorphisms in Romanian colorectal cancer patients

DOI: 10.26327/RBL2017.04

Received for publication, February, 22, 2017
Accepted, March, 2, 2017

DANIELA MURARASU[1,#], LILIANA PUIU[1,#,*], CORINA-ELENA MIHALCEA[1], IOANA MADALINA ALDEA PITICA[2], CRISTINA MAMBET[2], ELENA LACRAMIOARA RADU[2], LILIA MATEI[2], DENISA LAURA DRAGU[2], LAURENTIU SIMION[1], MARIAN AUGUSTIN MARINCAS[1], MIHAELA CHIVU-ECONOMESCU[2], SABIN CINCA[1], LORELEI BRASOVEANU[3], CORALIA BLEOTU[2], CARMEN CRISTINA DIACONU[2,*]

[1]*Prof. Alex. Trestioreanu Institute of Oncology, Bucharest, Romania*
[2]*Cellular and Molecular Department, Stefan S. Nicolau Institute of Virology, Bucharest, Romania*
[3]*Center of Immunology, Stefan S. Nicolau Institute of Virology, Bucharest, Romania*
[#]*These authors contributed equally to this article.*
***Address correspondence to:** Cellular and Molecular Department, Stefan S. Nicolau Institute of Virology, 285 Mihai Bravu, Bucharest 030304, Romania; Carmen C. Diaconu, Liliana Puiu Tel.: +4021 324 25 90; Email: ccdiaconu@yahoo.com; puiu_liliana@yahoo.com*

Abstract

The tumor suppressor TP53 gene is one of the commonest mutated genes in colorectal cancer. Beside somatic mutations, the gene harbors numerous genetic variations that may alter the protein functionality, affecting cancer risk and clinical outcome. Among genetic variations, single nucleotide polymorphisms were investigated in relation to cancer risk, including colorectal cancer susceptibility. The studies revealed that the distribution of polymorphic alleles and their functional relevance varied in a population-dependent manner.

In the present work we analyzed the distribution of eleven TP53 genetic variants in a lot of 63 colorectal cancer patients. The polymorphisms with minor allele frequency between 12%-21% and the resulting haplotypes were examined for association with clinicopathological features of colorectal tumors, TP53 somatic mutations and allelic instability at TP53 locus. The polymorphisms rs1625895, rs1042522, rs1642785 and rs17878362 located in intron 6, exon 4, intron 2 and intron 3, respectively were in strong linkage. The most common haplotypes GGGA1 and ACCA2 were present in 86.4% of the cases. Individual polymorphisms distribution and the two major haplotypes were statistic significantly associated with gender and age of disease onset, the expression of minor alleles and ACCA2 haplotype being prevalent in men and associated with an earlier age of the disease onset.

Keywords: colorectal cancer, *TP53* gene, genetic variations, haplotypes

1. Introduction

Colorectal cancer (CRC) accounts for a significant rate of morbidity and mortality worldwide. It arises as a consequence of the loss of genomic stability that promotes the acquisition of further DNA alterations leading to cell transformation and tumour progression. In about 35% of colorectal cancer risk may be explained by familial or inherited high-penetrance genes. In the remaining sporadic cases, with no identifiable genetic syndrome or family history, the risk of developing CRC during the life time may be explained by genetic variations and gene-environment interaction [1].

The human genome is characterized by a large amount of genetic variations, differences between a typical genome and the reference human genome ranging from 4.1 million to 5.0

million sites. Single nucleotide polymorphisms (SNPs) and short insertion/deletions (indels) account for about 99.9% of human genetic variation [2]. About 93% of the SNPs are found in non-coding DNA regions and a fraction of them is involved in transcriptional regulatory mechanisms [3]. SNPs can also be located in the coding parts of the genes, affecting the function of the protein or the level of gene expression.

Genetic variations exerting a direct effect on gene expression and protein function may act as low-penetrance alleles being associated with increased cancer risk, treatment outcome and prognosis. On the other hand, many cancer-related SNPs have no known biological effect. This discrepancy between a SNP biological function and its association with disease is due to linkage disequilibrium (LD) which is defined by the degree of inheritance within a population of alleles at different loci. Consequently, non-functional SNPs can be associated with an increased risk for developing a disease when they are in high LD with functional SNPs [4]. Therefore, analysis of distribution and extent of LD across the entire human genome can be a more revealing way for the evaluation of disease-associated SNPs than single locus analysis.

Genome-wide association studies (GWAS) identified multiple CRC-related SNPs which are scattered in more than 40 chromosome regions [5-7], being involved in mechanisms underlying genomic instability [8], cell cycle control [9], regulation of gene expression [10] and signal transduction pathways [11]. Although, most of the variants add a small increase in CRC risk, the presence of multiple risk alleles has cumulative effect [12]. Inherited genetic variants may also explain differential side effect rates and response to chemotherapy and radiotherapy of the CRC patients.

p53 tumor suppressor protein, encoded by *TP53* gene, plays a key role in maintenance of genomic stability. In response to cellular stress, p53 has the ability to arrest cell cycle, induce senescence or apoptosis depending on cellular context, this protein playing a significant role in specific pathways such as DNA repair, autophagy, differentiation, and oxidative metabolism.

TP53 is one of the most frequently mutated genes in all cancer types, a mutation rate of 50-60% being reported for CRC tumors IARC mutation database [13]. Besides being frequently mutated, *TP53* is also highly polymorphic. So far 547 SNPs have been mapped in the *TP53* locus along with other sequence variations [14]. Most of the SNPs are located in the non-coding regions of the gene and are thought to have no consequence on cancer. Studies concerning functional characterization, distribution in human populations and association with cancer risk have been focused on the polymorphisms found in the coding region of the *TP53* gene. However, these studies generated contradictions regarding the biological role of *TP53* variations and their relation to cancer predisposition. These discrepancies can be explained by tissue- and age-specific function of *TP53* SNPs, ethnicity, genetic background and environmental exposure [14].

The aim of the present study was to investigate the prevalence of several *TP53* SNPs and to analyze the distribution of haplotypes and their possible associations in 63 CRC patients from Romania.

2. Materials and methods

Patients and tumor samples

A total of 63 fresh tumor samples and 60 paired normal colon tissue samples were obtained from patients with sporadic colorectal cancer who underwent curative surgical resection at the *Prof. Dr Alex. Trestioreanu* Institute of Oncology, Bucharest, after written informed consents were obtained. All patients were examined by genealogical inquiry and were of Romanian descent. Tissue samples were fixed in 10% neutralized formalin and embedded in paraffin for histological processing or snap-frozen in liquid nitrogen for

molecular biology studies. Routine hematoxylin- and eosin-stained sections were reviewed by a pathologist. In sections assigned as "normal" no tumor cells were present while tumor sections had at least 80% tumoral cells. The patient group included 23 women and 40 men with ages ranging from 39 to 85 years (median 64 years).

The tumor was located in the proximal colon in 14 cases (22.2%), in the distal colon in 17 cases (27.0%) and in the rectum in 32 cases (50.8%). 43 tumors (68.2%) were N0, 11 (17.5%) were N1, 7 (11.1%) were N2 and 2 (3.2%) were Nx. According to tumor depth invasion, 23 (36.6%) tumors were classified as T2, 36 (57.1%) as T3 and 4 (6.3%) as T4. Of the 63 patients 49 (77.8%) had moderately differentiated tumours, 11 (17.5%) presented poorly differentiated tumours and 3 cases (4.8%) had undetermined grade. At the time of initial diagnosis, 48 (76.2%) of patients did not show distant metastasis, while in 14 (22.2%) cases at least one site of metastasis was present. Presence of distant metastasis could not be assessed in one case. Based on Duke's staging criteria, 16 (25.4%) tumors were stage A, 21 (33.3%) were stage B, 12 (19.1%) were stage C and 14 (22.2%) were stage D.

The study was approved by the institutional ethic committee and review board.

DNA extraction

Genomic DNA was isolated using the commercial QIAamp DNA Mini Kit (Qiagen,Hilden, Germany) following the manufacturer's recommendations and was quantified using a NanoDrop ND1000 Spectrophotometer (Thermo Scientific, USA). The extracted DNA was stored at -20°C.

DNA direct sequencing

The entire coding sequence of the *TP53* gene (exons 2–11) and the flanking regions were analyzed using DNA direct sequencing. Standard M13 tails were added to primers in order to facilitate subsequent sequencing. Primer sequences, PCR amplification and sequencing protocols were detailed in a previous study [16]. *TP53* polymorphisms were checked against the NCBI SNP database [17].

Statistical analysis

All data were analyzed by using GraphPad Prism 5 Software Version 5.04 (GraphPad Sofware Inc). Values of $p< 0.05$ were considered statistically significant.

The Haploview 4.2 software package [18] was used to estimate pair-wise LD, detect deviation from the Hardy–Weinberg equilibrium, construct haplotypes, and calculate haplotype frequencies.

3. Results and Discussion

Eleven genetic variants were detected by sequencing all the exons of the *TP53* gene and their flanking regions. Ten of these were affecting a single base within either introns or exons of the *TP53* gene and one was an insertion/deletion type. Four single nucleotide variants (SNVs) were found in the coding regions. In exon 4, the non-synonymous SNV p.R72P (rs1042522), leading to a proline (Pro) to arginine (Arg) substitution, was detected in almost all samples, either in the homozygous or heterozygous form. A silent substitution in exon 4, (p.P36P, rs1800370) which corresponds to a Pro coding codon, located in the transactivation domain, was detected in one sample. A second synonymous variant was detected in exon 6 (p.R213R, rs1800372) in two cases. In exon 8 a homozygous C→T transition resulting in Arg to cysteine (Cys) substitution at amino acid position 283 of the p53 protein (p.R283C, c.847C>T) was detected in one sample. This variation (rs149633775) was also present in the matched normal tissues and is described in literature as a germline mutation.

Six SNVs were located in different intronic regions of the *TP53* gene. In intron 2 the most frequent variant was the rs1642785 polymorphism (c.74+38C>G) present in 98.4% of the samples (62/63), while the rs17883323 polymorphism (c.97-29C>A) was identified in only 5 samples.

Within the sequence of intron 6 the polymorphism rs17880604 (c.673-36G>C) was found in only 4 cases in contrast to the polymorphism rs1625895 (c.672+62A>G), which was present in all but one samples. In intron 10 we detected two rare heterozygous substitutions: rs17880847 (c.1100+30A>T) in one sample and rs17881850 (c.1101-49C>T) in 3 cases. In addition, a common 16-base pair variation in intron 3 (rs17878362) consisting of one copy (A1 allele) or of two copies (A2 allele) of the sequence ACCTGGAGGGCTGGGG was detected at high frequency. The allele and genotype frequencies are summarized in table 1. With the exception of rs149633775, allele ratios of the SNPs analyzed were in agreement with Hardy-Weinberg equilibrium.

Among these 11 variants the rs1625895, rs1042522, rs1642785 and rs17878362 had minor allele frequency (MAF) between 21%-12%. Representative chromatograms for each of these polymorphisms are shown in figure 1.

We examined the association between the 4 variants with clinicopathological features. However, no significant differences were found between individual variants and tumour localization, histological grade, lymph nodes, tumor infiltration, distant metastasis or Duke's stage.

When allele distribution was correlated with gender we found that male expressed minor alleles at a higher level comparing to females. (Men:Women allele count - rs1625895 A 14:66, 1:45, $\chi2$ =6.541, p=0.0105; rs1042522 C 18:62, 3:43, $\chi2$ =5.369, p=0.0205; rs1642785 C 23:57, 3:43, $\chi2$ =8.812, p=0.003; rs17878362 A2 14:66, 2:44, $\chi2$ =4.557, p=0.0328).

Table 1. *TP53* allele and genotype frequencies.

Location	SNPs	Genotype count (frequency)	Allele	Allele count (frequency)	Hardy-Weinberg p Value
Intron 2	**rs1642785**				
	CC	1 (1.6)	C	26 (0.21)	0.4031
	CG	24 (38.1)	G	100 (0.79)	
	GG	38 (60.3)			
	rs17883323				
	CC	58 (92.1)	C	121 (0.96)	1.0
	CA	5 (7.9)	A	5 (0.04)	
	AA	0 (0.0)			
Intron 3	**rs17878362**				
	A1A1	48 (76.2)	A1	110 (0.87)	1.0
	A1A2	14 (22.2)	A2	16 (0.13)	
	A2A2	1 (1.6)			
Exon 4	**rs1800370 (p.P36P)**				
	GG	62 (98.4)	G	125 (0.99)	1.0
	GA	1 (1.6)	A	1 (0.01)	
	AA	0 (0.0)			
	rs1042522 (p.P72R)				
	CC	1 (1.6)	C	21 (0.17)	0.9286
	CG	19 (30.2)	G	105 (0.83)	
	GG	43 (68.2)			
Exon 6	**rs1800372 (p.R213R)**				
	AA	61 (96.8)	A	124 (0.98)	1.0
	AG	2 (3.2)	G	2 (0.02)	
	GG	0 (0.0)			
Intron 6	**rs1625895**				
	AA	1 (1.6)	A	15 (0.12)	1.0
	AG	13 (20.6)	G	111 (0.88)	
	GG	49 (77.8)			
	rs17880604				
	GG	59 (93.6)	G	121 (0.96)	0.1580
	GC	3 (4.8)	C	5 (0.04)	
	CC	1 (1.6)			
Exon 8	**rs149633775**				
	CC	62 (98.4)	C	124 (0.98)	0.0160
	CT	0 (0.0)	T	2 (0.02)	
	TT	1 (1.6)			
Intron 10	**rs17880847**				
	AA	62 (98.4)	A	125 (0.99)	1.0
	AT	1 (1.6)	T	1 (0.01)	
	TT	0 (0.0)			
	rs17881850				
	CC	60 (95.2)	C	123 (0.98)	1.0
	CT	3 (4.8)	T	3 (0.02)	
	TT	0 (0.0)			

We also examined the effect of individual variants on age of cancer onset and we observed a significant decrease of major alleles in patients younger than 50 years of age. (>50:<50 allele count - rs1625895 G 105:11, 6:4, $\chi 2$ =8.175, p=0.0042; rs1042522 G 100:16, 5:5, $\chi 2$ =8.69, p=0.0032; rs1642785 G 95:21, 5:5, $\chi 2$ =5.719, p=0.0168; rs17878362 A1 104:12, 6:4, $\chi 2$ =7.303, p=0.0069).

According to our previous work 46.03% of the analyzed CRC patients carried *TP53* somatic mutations and 50% of the informative cases showed allelic instability at *TP53* locus [16].

The correlation study did not reveal association between individual polymorphisms and increased odds of somatic mutations or allelic instability.

The pattern of pair-wise LD was observed across rs1625895, rs1042522, rs1642785 and rs17878362 (all D' > 0.8; r^2 > 0.3). The most common haplotypes were GGGA1 and ACCA2 accounting for 86.4% of the haplotypes observed in all individuals studied (figure 2).

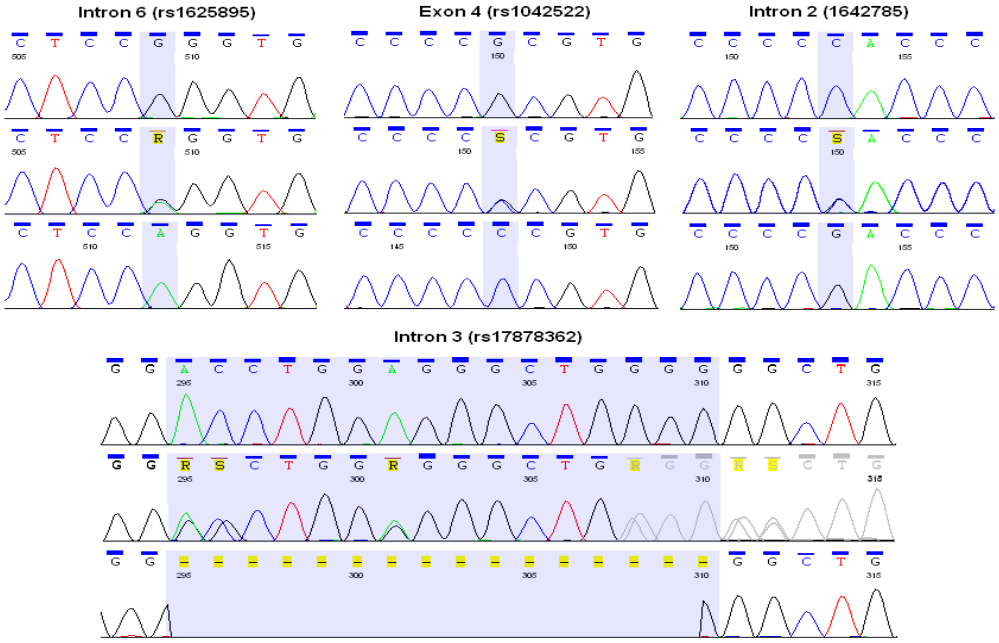

Figure 1. Chromatograms of the most common *TP53* polymorphisms in CRC patients.

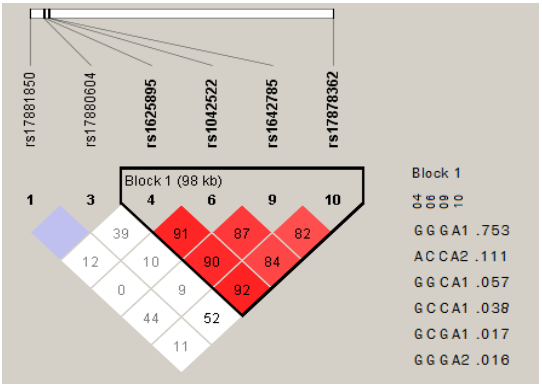

Figure 2. LD plot and block structure of six *TP53* polymorphisms in 63 colorectal tumors. The haplotype block is based on confidence intervals D'. Each diamond represents the pairwise magnitude of LD, with red indicating strong LD (D' > 0.8) and logarithm of odds score (LOD) ≥ 2.0.

The analysis of haplotype distribution for TP53 revealed statistically significant differences between gender and age groups. In particular, the haplotype GGGA1 was prevalent in women (Men:Women allele frequency - 0.685, 0.870, $\chi 2$ =5.318, p=0.0211), while the haplotype ACCA2 was associated with men (Men:Women allele frequency - 0.163, 0.022, $\chi 2$ =5.859, p=0.0155). After the stratification for age of disease onset, the haplotype GGGA1 was associated with older patients (>50:<50 allele frequency - 0.775, 0.497, $\chi 2$ =3.822, p=0.0506) and the haplotype ACCA2 was more frequent in younger patients (>50:<50 allele frequency - 0.086, 0.400, $\chi 2$ =9.178, p=0.0024). We did not observe any association between the 4 polymorphisms haplotype and clinicopathological or molecular features of colorectal tumors.

Based on the solid spine of LD block definition in Haploview, one haplotype block was identified consisting of four polymorphisms (rs1625895, rs1042522, rs1642785 and rs17878362), which have been extensively studied in cancer genetic epidemiology.

In a series of previous studies, it was demonstrated that the four variants affect either protein functions or mRNA expression level.

The polymorphism rs1625895 displays a G→A transition in intron 6. Although the functional importance of rs1625895 is not confirmed yet, it seems to be responsible for changes in the induction of DNA repair and apoptosis and for the maintenance of genomic stability [19,20]. The minor allele frequency reported in the European population was similar with the frequency found in our study (0.14 vs. 0.12, p= 0.8339).

The SNP rs1042522, a G→C transversion in codon 72 of exon 4, leads to protein variants with different biological properties. The protein containing Arg was shown to be more efficient in inducing apoptosis, a property that correlated with a greater capacity to interact with MDM2 which facilitate nuclear export and mitochondrial localization. In contrast, the Pro variant was found to be more efficient in inducing cell-cycle arrest [21] and DNA repair [15]. Other differences include the ability to suppress transformed cell growth and to activate transcription [16,17]. An important consideration with respect to rs1042522 is the preferential loss of the Pro allele in heterozygous individuals. The retention of the Arg allele, that is preferentially mutated, might be the result of an anti-apoptotic advantage conferred by *TP53* mutations [18]. This polymorphism was found to influence the susceptibility to CRC in some studies, but significant differences in results were obtained in respect to the target population. While certain authors reported that Arg variant increased the CRC risk [19-21], others provided opposite results, favoring Pro variant as being linked to CRC susceptibility [22-24]. On the other hand, in some studies this polymorphism was shown to exert only a borderline/weak effect on cancer risk [25,26]. These contradictory results indicate that 72 codon alleles are unevenly distributed across world populations. In European population the frequency of the minor allele of rs1042522 is 0.29, while in our study was 0.17 (p= 0.064).

The SNP rs17878362, a common intronic variation of the *TP53* gene, presents an insertion/deletion of 16-base pair in intron 3. Rs17878362 is situated in a G-rich region that forms a G-quadruplex structure involved in the alternative splicing of TP53 mRNA. It was demonstrated that this polymorphism changes the relative position of a G-quadruplex structure with respect to *TP53* intron 2 splice junction and alters the balance between the fully spliced p53 transcript (FSp53 mRNA) and an incompletely spliced transcript (p53I2 mRNA). The alternatively spliced transcript retains intron 2 and encodes an isoform (Δ40p53) lacking first 39 N-terminal residues corresponding to the main transactivation domain. This isoform can inhibit suppressive activity of p53. Both rs17878362 variants influence the alternative splicing of the intron 2 with the A2 allele being associated with lower levels of the p53I2 mRNA. Furthermore, several lines of evidence indicated that rs17878362-A2 variant

decreases DNA repair and apoptosis processes [27,28]. The minor allele (A2) frequency in our series of patients was 0.126 while no data are reported in NCBI dbSNP database.

Rs1642785 located in the intron 2 is the G→C substitution. The rs1642785 polymorphism influences the balance between FSp53 and p53I2 mRNAs by modulating the stability of the alternatively spliced mRNA. Perriaud and collaborators [27] demonstrated that the highest FSp53 and p53I2 levels were associated with combined rs1642785-GG/rs17878362-A1A1 alleles, whereas the presence of rs1642785-C with either rs17878362 allele was associated with lower TP53 pre-mRNA, total TP53, FSp53 and p53I2 levels due to the lower stability of transcripts containing rs1642785-C. Considering its location at the center of a region rich in CA repeats, it has been speculated that the destabilization effect of the rs1642785-C allele is a result of long CA repeats in TP53 intronic sequence. Based on the NCBI dbSNP database, the C allele frequency in European population is around 0.29, while in our study the minor allele frequency was 0.21 (p= 0.2529).

Allele distribution of the above mentioned polymorphisms was statistically significant associated with gender and age at diagnosis of colorectal cancer, minor allele being more frequently expressed in men and patients younger than 50 years old.

Previous studies showed that *TP53* rs1625895, rs17878362, rs1042522 and rs1642785 polymorphisms are in strong linkage disequilibrium [29,30]. The most frequently haplotype was rs1625895-G/ rs1042522-G/ rs17878362-A1/ rs1642785-G (75.3%), while the rs1625895-A/ rs1042522-C/ rs17878362-A2/ rs1642785-C haplotype was seen in only 11.1% of colorectal tumors. Sagne and collaborators [31] reported that 78.13% of the Caucasian population carried the haplotype combining rs17878362-A1/rs1042522-G/rs1625895-G, whereas the haplotype rs17878362-A2/rs1042522-C/rs1625895-A was seen in 9.37% of the population. The homozygous carriers of the rare rs17878362-A2 genotype presented an increased CRC susceptibility compared with homozygous carriers of the common A1 genotype (A2A2 vs. A1A1 aggregated OR=1.67, 95% CI=1.02–2.74). In the same data set, the heterozygote carriers of the rs1042522 and rs1625895 SNPs had an increased overall cancer risk, although the effects observed was smaller than for the rs17878362 variants. The results were consistent with several previous meta-analyses and systematic reviews that showed that rs1042522 and rs1625895 are unlikely to be major risk factors for CRC [18,32]. These two polymorphisms belong to the same *TP53* haplotype block as rs17878362. It remains to be determined up to which point their effects can be attributed to the LD with rs17878362.

There are few available data regarding the association between the rs1642785 and cancer risk. One study showed an association between the G allele and ovarian cancer, while a second reported an increased osteosarcoma risk in C allele carriers [33]. To our knowledge there is no study concerning the association of the rs1642785 with CRC susceptibility. In our work the rs1642785-C allele was present on 86.6% of the haplotypes that carry the A2 allele of rs17878362. Given the synergistic effect of these two polymorphisms on the formation of mRNA encoding N-terminal isoforms, it would be interesting to determine whether the presence of allele C can change the impact of rs17878362 on cancer risk.

In our patients the major haplotypes GGGA1 and ACCA2 showed a gender- and age-specific distribution. In particular haplotype ACCA2 was prevalent in men and was associated with earlier age of onset.

With respect to the additional 6 SNPs identified within *TP53* sequence, the frequency of minor allele ranges between 0.01 – 0.04. Apart from the rs17880604 polymorphism in intron

6 that showed no association with cancer risk, no similar information could be found for the other 5 SNPs [13,32,34].

The two rare synonymous polymorphisms (rs1800370, p.P36P and rs1800372, p.R213R) located into the coding region of *TP53* gene theoretically do not change the amino acid sequence or structure of the protein. Nevertheless, changes in base sequence could modify protein expression, folding and function, or provoke new splicing events. For example, the polymorphism at codon 36 was shown to affect the ability of p53 to activate apoptosis by modulating the affinity of the TP53 mRNA for MDM2 and, p53 levels consequently [35,36]. In contrast, the R213R polymorphism shows no phenotypic differences [37].

4. Conclusions

Our study found four SNVs (rs1042522, rs1800370, rs1800372, and rs149633775) in the coding regions, six SNVs in intronic regions (rs1642785, rs17883323, rs17880604, rs1625895, rs17880847 and rs17881850) and a 16-base pair duplication in intron 3 (rs17878362) of *TP53* gene, the distribution of allele and genotype frequencies of the SNPs examined being comparable to those of European (or Caucasian) population as reported in the NCBI dbSNP database.

One haplotype block consisting of four polymorphisms (rs1625895, rs1042522, rs1642785 and rs17878362) was identified. The most common haplotypes were GGGA1 and ACCA2 accounting for 86.4% of the haplotypes observed in all individuals studied was associated with gender and earlier age of disease onset and may give a new angle to comprehend under the age of 50 CRC onsets, which seems to have a higher incidence in the last years.

Acknowledgements

We are grateful for the support from P2-2.1-BG-2016-0439, within PNCDI III and PCCA PNII 133/2014-2017.

References

1. J.M. O'BRIEN, Environmental and Heritable Factors in the Causation of Cancer. *Surv. Ophthalmol.*, 45:2, 167-168 (2000).
2. CONSORTIUM GENOMES PROJECT, L.D. BROOKS, R.M. DURBIN, E.P. GARRISON, H.M. KANG, J.O. KORBEL, J.L. MARCHINI, S. MCCARTHY, G.A. MCVEAN, G.R. ABECASIS, A global reference for human genetic variation. *Nature*, 526:7571, 68-74 (2015).
3. M.T. MAURANO, R. HUMBERT, E. RYNES, R.E. THURMAN, E. HAUGEN, H. WANG, A.P. REYNOLDS, R. SANDSTROM, H. QU, J. BRODY, A. SHAFER, F. NERI, K. LEE, T. KUTYAVIN, S. STEHLING-SUN, A.K. JOHNSON, T.K. CANFIELD, E. GISTE, M. DIEGEL, D. BATES, R.S. HANSEN, S. NEPH, P.J. SABO, S. HEIMFELD, A. RAUBITSCHEK, S. ZIEGLER, C. COTSAPAS, N. SOTOODEHNIA, I. GLASS, S.R. SUNYAEV, R. KAUL, J.A. STAMATOYANNOPOULOS, Systematic localization of common disease-associated variation in regulatory DNA. *Science,* 337:6099, 1190-1195 (2012).
4. W.S. BUSH, J.H. MOORE, Chapter 11: Genome-wide association studies. *PLoS Comput. Biol.*, 8 (12), e1002822 (2012).
5. A. TENESA, S.M. FARRINGTON, J.G. PRENDERGAST, M.E. PORTEOUS, M. WALKER, N. HAQ, R.A. BARNETSON, E. THEODORATOU, R. CETNARSKYJ, N. CARTWRIGHT, C. SEMPLE, A.J. CLARK, F.J. REID, L.A. SMITH, K. KAVOUSSANAKIS, T. KOESSLER, P.D. PHAROAH, S. BUCH, C. SCHAFMAYER, J. TEPEL, S. SCHREIBER, H. VOLZKE, C.O. SCHMIDT, J. HAMPE, J. CHANG-CLAUDE, M. HOFFMEISTER, H. BRENNER, S. WILKENING, F. CANZIAN, G. CAPELLA, V. MORENO, I.J. DEARY, J.M. STARR, I.P. TOMLINSON, Z. KEMP, K. HOWARTH, L. CARVAJAL-CARMONA, E. WEBB, P. BRODERICK, J. VIJAYAKRISHNAN, R.S. HOULSTON, G. RENNERT, D. BALLINGER, L. ROZEK, S.B. GRUBER, K. MATSUDA, T. KIDOKORO, Y. NAKAMURA, B.W. ZANKE, C.M. GREENWOOD, J. RANGREJ, R. KUSTRA, A. MONTPETIT, T.J. HUDSON, S.

GALLINGER, H. CAMPBELL, M.G. DUNLOP, Genome-wide association scan identifies a colorectal cancer susceptibility locus on 11q23 and replicates risk loci at 8q24 and 18q21. *Nat. Genet.*, 40:5, 631-637 (2008).

6. I.P. TOMLINSON, E. WEBB, L. CARVAJAL-CARMONA, P. BRODERICK, K. HOWARTH, A.M. PITTMAN, S. SPAIN, S. LUBBE, A. WALTHER, K. SULLIVAN, E. JAEGER, S. FIELDING, A. ROWAN, J. VIJAYAKRISHNAN, E. DOMINGO, I. CHANDLER, Z. KEMP, M. QURESHI, S.M. FARRINGTON, A. TENESA, J.G. PRENDERGAST, R.A. BARNETSON, S. PENEGAR, E. BARCLAY, W. WOOD, L. MARTIN, M. GORMAN, H. THOMAS, J. PETO, D.T. BISHOP, R. GRAY, E.R. MAHER, A. LUCASSEN, D. KERR, D.G. EVANS, C. CONSORTIUM, C. SCHAFMAYER, S. BUCH, H. VOLZKE, J. HAMPE, S. SCHREIBER, U. JOHN, T. KOESSLER, P. PHAROAH, T. VAN WEZEL, H. MORREAU, J.T. WIJNEN, J.L. HOPPER, M.C. SOUTHEY, G.G. GILES, G. SEVERI, S. CASTELLVI-BEL, C. RUIZ-PONTE, A. CARRACEDO, A. CASTELLS, E. CONSORTIUM, A. FORSTI, K. HEMMINKI, P. VODICKA, A. NACCARATI, L. LIPTON, J.W. HO, K.K. CHENG, P.C. SHAM, J. LUK, J.A. AGUNDEZ, J.M. LADERO, M. DE LA HOYA, T. CALDES, I. NIITTYMAKI, S. TUUPANEN, A. KARHU, L. AALTONEN, J.B. CAZIER, H. CAMPBELL, M.G. DUNLOP, R.S. HOULSTON, A genome-wide association study identifies colorectal cancer susceptibility loci on chromosomes 10p14 and 8q23.3. *Nat. Genet.*, 40:5, 623-630 (2008).
7. R.S. HOULSTON, J. CHEADLE, S.E. DOBBINS, A. TENESA, A.M. JONES, K. HOWARTH, S.L. SPAIN, P. BRODERICK, E. DOMINGO, S. FARRINGTON, J.G. PRENDERGAST, A.M. PITTMAN, E. THEODORATOU, C.G. SMITH, B. OLVER, A. WALTHER, R.A. BARNETSON, M. CHURCHMAN, E.E. JAEGER, S. PENEGAR, E. BARCLAY, L. MARTIN, M. GORMAN, R. MAGER, E. JOHNSTONE, R. MIDGLEY, I. NIITTYMAKI, S. TUUPANEN, J. COLLEY, S. IDZIASZCZYK, C. CONSORTIUM, H.J. THOMAS, A.M. LUCASSEN, D.G. EVANS, E.R. MAHER, C. CONSORTIUM, C.C. GROUP, C.C. GROUP, T. MAUGHAN, A. DIMAS, E. DERMITZAKIS, J.B. CAZIER, L.A. AALTONEN, P. PHAROAH, D.J. KERR, L.G. CARVAJAL-CARMONA, H. CAMPBELL, M.G. DUNLOP, I.P. TOMLINSON, Meta-analysis of three genome-wide association studies identifies susceptibility loci for colorectal cancer at 1q41, 3q26.2, 12q13.13 and 20q13.33. *Nat. Genet.*, 42:11, 973-977 (2010).
8. U. PETERS, S. JIAO, F.R. SCHUMACHER, C.M. HUTTER, A.K. ARAGAKI, J.A. BARON, S.I. BERNDT, S. BEZIEAU, H. BRENNER, K. BUTTERBACH, B.J. CAAN, P.T. CAMPBELL, C.S. CARLSON, G. CASEY, A.T. CHAN, J. CHANG-CLAUDE, S.J. CHANOCK, L.S. CHEN, G.A. COETZEE, S.G. COETZEE, D.V. CONTI, K.R. CURTIS, D. DUGGAN, T. EDWARDS, C.S. FUCHS, S. GALLINGER, E.L. GIOVANNUCCI, S.M. GOGARTEN, S.B. GRUBER, R.W. HAILE, T.A. HARRISON, R.B. HAYES, B.E. HENDERSON, M. HOFFMEISTER, J.L. HOPPER, T.J. HUDSON, D.J. HUNTER, R.D. JACKSON, S.H. JEE, M.A. JENKINS, W.H. JIA, L.N. KOLONEL, C. KOOPERBERG, S. KURY, A.Z. LACROIX, C.C. LAURIE, C.A. LAURIE, L. LE MARCHAND, M. LEMIRE, D. LEVINE, N.M. LINDOR, Y. LIU, J. MA, K.W. MAKAR, K. MATSUO, P.A. NEWCOMB, J.D. POTTER, R.L. PRENTICE, C. QU, T. ROHAN, S.A. ROSSE, R.E. SCHOEN, D. SEMINARA, M. SHRUBSOLE, X.O. SHU, M.L. SLATTERY, D. TAVERNA, S.N. THIBODEAU, C.M. ULRICH, E. WHITE, Y. XIANG, B.W. ZANKE, Y.X. ZENG, B. ZHANG, W. ZHENG, L. HSU, R. COLON CANCER FAMILY, G. THE, C. EPIDEMIOLOGY OF COLORECTAL CANCER, Identification of Genetic Susceptibility Loci for Colorectal Tumors in a Genome-Wide Meta-analysis. *Gastroenterol.*, 144:4, 799-807 e724 (2013).
9. W.H. JIA, B. ZHANG, K. MATSUO, A. SHIN, Y.B. XIANG, S.H. JEE, D.H. KIM, Z. REN, Q. CAI, J. LONG, J. SHI, W. WEN, G. YANG, R.J. DELAHANTY, GENETICS, C. EPIDEMIOLOGY OF COLORECTAL CANCER, R. COLON CANCER FAMILY, B.T. JI, Z.Z. PAN, F. MATSUDA, Y.T. GAO, J.H. OH, Y.O. AHN, E.J. PARK, H.L. LI, J.W. PARK, J. JO, J.Y. JEONG, S. HOSONO, G. CASEY, U. PETERS, X.O. SHU, Y.X. ZENG, W. ZHENG, Genome-wide association analyses in East Asians identify new susceptibility loci for colorectal cancer. *Nat. Genet.*, 45:2, 191-196 (2013).
10. UK IBD GENETICS CONSORTIUM, J.C. BARRETT, J.C. LEE, C.W. LEES, N.J. PRESCOTT, C.A. ANDERSON, A. PHILLIPS, E. WESLEY, K. PARNELL, H. ZHANG, H. DRUMMOND, E.R. NIMMO, D. MASSEY, K. BLASZCZYK, T. ELLIOTT, L. COTTERILL, H. DALLAL, A.J. LOBO, C. MOWAT, J.D. SANDERSON, D.P. JEWELL, W.G. NEWMAN, C. EDWARDS, T. AHMAD, J.C. MANSFIELD, J. SATSANGI, M. PARKES, C.G. MATHEW, C. WELLCOME TRUST CASE CONTROL, P. DONNELLY, L. PELTONEN, J.M. BLACKWELL, E. BRAMON, M.A. BROWN, J.P. CASAS, A. CORVIN, N. CRADDOCK, P. DELOUKAS, A. DUNCANSON, J. JANKOWSKI, H.S. MARKUS, C.G. MATHEW, M.I. MCCARTHY, C.N. PALMER, R. PLOMIN, A. RAUTANEN, S.J. SAWCER, N. SAMANI, R.C.

TREMBATH, A.C. VISWANATHAN, N. WOOD, C.C. SPENCER, J.C. BARRETT, C. BELLENGUEZ, D. DAVISON, C. FREEMAN, A. STRANGE, P. DONNELLY, C. LANGFORD, S.E. HUNT, S. EDKINS, R. GWILLIAM, H. BLACKBURN, S.J. BUMPSTEAD, S. DRONOV, M. GILLMAN, E. GRAY, N. HAMMOND, A. JAYAKUMAR, O.T. MCCANN, J. LIDDLE, M.L. PEREZ, S.C. POTTER, R. RAVINDRARAJAH, M. RICKETTS, M. WALLER, P. WESTON, S. WIDAA, P. WHITTAKER, P. DELOUKAS, L. PELTONEN, C.G. MATHEW, J.M. BLACKWELL, M.A. BROWN, A. CORVIN, M.I. MCCARTHY, C.C. SPENCER, A.P. ATTWOOD, J. STEPHENS, J. SAMBROOK, W.H. OUWEHAND, W.L. MCARDLE, S.M. RING, D.P. STRACHAN, Genome-wide association study of ulcerative colitis identifies three new susceptibility loci, including the HNF4A region. *Nat. Genet.*, 41:12, 1330-1334 (2009).

11. P. BRODERICK, L. CARVAJAL-CARMONA, A.M. PITTMAN, E. WEBB, K. HOWARTH, A. ROWAN, S. LUBBE, S. SPAIN, K. SULLIVAN, S. FIELDING, E. JAEGER, J. VIJAYAKRISHNAN, Z. KEMP, M. GORMAN, I. CHANDLER, E. PAPAEMMANUIL, S. PENEGAR, W. WOOD, G. SELLICK, M. QURESHI, A. TEIXEIRA, E. DOMINGO, E. BARCLAY, L. MARTIN, O. SIEBER, C. CONSORTIUM, D. KERR, R. GRAY, J. PETO, J.B. CAZIER, I. TOMLINSON, R.S. HOULSTON, A genome-wide association study shows that common alleles of SMAD7 influence colorectal cancer risk. *Nat. Genet.*, 39:11, 1315-1317 (2007).
12. A. ABULI, A. CASTELLS, L. BUJANDA, J.J. LOZANO, X. BESSA, C. HERNANDEZ, C. ALVAREZ-URTURI, M. PELLISE, C. ESTEBAN-JURADO, E. HIJONA, A. BURON, F. MACIA, J. GRAU, R. GUAYTA, S. CASTELLVI-BEL, M. ANDREU, P.R. GROUP, Genetic Variants Associated with Colorectal Adenoma Susceptibility. *PLoS One*, 11 (4), e0153084 (2016).
13. N. BUYRU, A. TEZOL, N. DALAY, p53 intronic G13964C variant in colon cancer and its association with HPV. *Anticancer Res.*, 25:4, 2767-2769 (2005).
14. E.V. DENISOV, N.V. CHERDYNTSEVA, N.V. LITVIAKOV, E.A. MALINOVSKAYA, N.N. BABYSHKINA, V.A. BELYAVSKAYA, M.I. VOEVODA, TP53 Gene Polymorphisms in Cancer Risk: The Modulating Effect of Ageing, Ethnicity and TP53 Somatic Abnormalities. In: Tumor Suppressor Genes Cheng Y, editor: *InTech* 2012. pp. 79-110.
15. M. SIDDIQUE, K. SABAPATHY, Trp53-dependent DNA-repair is affected by the codon 72 polymorphism. *Oncogene*, 25:25, 3489-3500 (2006).
16. N. VOLODKO, M. SALLA, B. EKSTEEN, R.N. FEDORAK, H.Q. HUYNH, S. BAKSH, TP53 codon 72 Arg/Arg polymorphism is associated with a higher risk for inflammatory bowel disease development. *World J. Gastroenterol.*, 21:36, 10358-10366 (2015).
17. A. PETITJEAN, M.I. ACHATZ, A.L. BORRESEN-DALE, P. HAINAUT, M. OLIVIER, TP53 mutations in human cancers: functional selection and impact on cancer prognosis and outcomes. *Oncogene*, 26:15, 2157-2165 (2007).
18. I.J. DAHABREH, C.H. SCHMID, J. LAU, V. VARVARIGOU, S. MURRAY, T.A. TRIKALINOS, Genotype misclassification in genetic association studies of the rs1042522 TP53 (Arg72Pro) polymorphism: a systematic review of studies of breast, lung, colorectal, ovarian, and endometrial cancer. *Am. J. Epidemiol.*, 177 (12), 1317-1325 (2013).
19. R. SCHNEIDER-STOCK, C. BOLTZE, B. PETERS, R. SZIBOR, O. LANDT, F. MEYER, A. ROESSNER, Selective loss of codon 72 proline p53 and frequent mutational inactivation of the retained arginine allele in colorectal cancer. *Neoplasia*, 6:5, 529-535 (2004).
20. L.O. PEREZ, M.C. ABBA, F.N. DULOUT, C.D. GOLIJOW, Evaluation of p53 codon 72 polymorphism in adenocarcinomas of the colon and rectum in La Plata, Argentina. *World J. Gastroenterol.*, 12:9, 1426-1429 (2006).
21. A. DAKOURAS, N. NIKITEAS, E. PAPADAKIS, M. PERAKIS, D. VALIS, G. RALLIS, N. TZANAKIS, G. PEROS, C. TSIGKRIS, C. KITTAS, P. KARAKITSOS, P53Arg72 homozygosity and its increased incidence in left-sided sporadic colorectal adenocarcinomas, in a Greek-Caucasian population. *Anticancer Res.*, 28:2A, 1039-1043 (2008).
22. Z.Z. ZHU, A.Z. WANG, H.R. JIA, X.X. JIN, X.L. HE, L.F. HOU, G. ZHU, Association of the TP53 codon 72 polymorphism with colorectal cancer in a Chinese population. *Jpn. J. Clin. Oncol*., 37:5, 385-390 (2007).
23. Z. CAO, J.H. SONG, Y.K. PARK, E.J. MAENG, S.W. NAM, J.Y. LEE, W.S. PARK, The p53 codon 72 polymorphism and susceptibility to colorectal cancer in Korean patients. *Neoplasma*, 56:2, 114-118 (2009).
24. A.A. AIZAT, S.N. SHAHPUDIN, M.A. MUSTAPHA, Z. ZAKARIA, A.S. SIDEK, M.R. ABU HASSAN, R. ANKATHIL, Association of Arg72Pro of P53 polymorphism with colorectal cancer susceptibility risk in Malaysian population. *Asian Pac. J. Cancer Prev.*, 12:11, 2909-2913 (2011).
25. F. GEMIGNANI, V. MORENO, S. LANDI, N. MOULLAN, A. CHABRIER, S. GUTIERREZ-ENRIQUEZ, J. HALL, E. GUINO, M.A. PEINADO, G. CAPELLA, F. CANZIAN, A TP53 polymorphism is associated with increased risk of colorectal cancer and with reduced levels of TP53 mRNA. *Oncogene*, 23:10, 1954-1956 (2004).

26. J.J. WANG, Y. ZHENG, L. SUN, L. WANG, P.B. YU, J.H. DONG, L. ZHANG, J. XU, W. SHI, Y.C. REN, TP53 codon 72 polymorphism and colorectal cancer susceptibility: a meta-analysis. *Mol. Biol. Rep.*, 38:8, 4847-4853 (2011).
27. L. PERRIAUD, V. MARCEL, C. SAGNE, V. FAVAUDON, A. GUEDIN, A. DE RACHE, C. GUETTA, F. HAMON, M.P. TEULADE-FICHOU, P. HAINAUT, J.L. MERGNY, J. HALL, Impact of G-quadruplex structures and intronic polymorphisms rs17878362 and rs1642785 on basal and ionizing radiation-induced expression of alternative p53 transcripts. *Carcinogenesis*, 35:12, 2706-2715 (2014).
28. V. MARCEL, P.L. TRAN, C. SAGNE, G. MARTEL-PLANCHE, L. VASLIN, M.P. TEULADE-FICHOU, J. HALL, J.L. MERGNY, P. HAINAUT, E. VAN DYCK, G-quadruplex structures in TP53 intron 3: role in alternative splicing and in production of p53 mRNA isoforms. *Carcinogenesis*, 32:3, 271-278 (2011).
29. M. GARCIA-CLOSAS, V. KRISTENSEN, A. LANGEROD, Y. QI, M. YEAGER, L. BURDETT, R. WELCH, J. LISSOWSKA, B. PEPLONSKA, L. BRINTON, D.S. GERHARD, I.T. GRAM, C.M. PEROU, A.L. BORRESEN-DALE, S. CHANOCK, Common genetic variation in TP53 and its flanking genes, WDR79 and ATP1B2, and susceptibility to breast cancer. *Int. J. Cancer*, 121:11, 2532-2538 (2007).
30. J.M. SCHILDKRAUT, E.L. GOODE, M.A. CLYDE, E.S. IVERSEN, P.G. MOORMAN, A. BERCHUCK, J.R. MARKS, J. LISSOWSKA, L. BRINTON, B. PEPLONSKA, J.M. CUNNINGHAM, R.A. VIERKANT, D.N. RIDER, G. CHENEVIX-TRENCH, P.M. WEBB, J. BEESLEY, X. CHEN, C. PHELAN, R. SUTPHEN, T.A. SELLERS, L. PEARCE, A.H. WU, D. VAN DEN BERG, D. CONTI, C.K. ELUND, R. ANDERSON, M.T. GOODMAN, G. LURIE, M.E. CARNEY, P.J. THOMPSON, S.A. GAYTHER, S.J. RAMUS, I. JACOBS, S. KRUGER KJAER, E. HOGDALL, J. BLAAKAER, C. HOGDALL, D.F. EASTON, H. SONG, P.D. PHAROAH, A.S. WHITTEMORE, V. MCGUIRE, L. QUAYE, H. ANTON-CULVER, A. ZIOGAS, K.L. TERRY, D.W. CRAMER, S.E. HANKINSON, S.S. TWOROGER, B. CALINGAERT, S. CHANOCK, M. SHERMAN, M. GARCIA-CLOSAS, G. AUSTRALIAN OVARIAN CANCER STUDY, Single nucleotide polymorphisms in the TP53 region and susceptibility to invasive epithelial ovarian cancer. *Cancer Res.*, 69:6, 2349-2357 (2009).
31. C. SAGNE, V. MARCEL, A. AMADOU, P. HAINAUT, M. OLIVIER, J. HALL, A meta-analysis of cancer risk associated with the TP53 intron 3 duplication polymorphism (rs17878362): geographic and tumor-specific effects. *Cell Death Dis.*, 4, e492 (2013).
32. E. MAMMANO, C. BELLUCO, M. BONAFE, F. OLIVIERI, E. MUGIANESI, C. BARBI, M. MISHTO, M. COSCI, C. FRANCESCHI, M. LISE, D. NITTI, Association of p53 polymorphisms and colorectal cancer: modulation of risk and progression. *Eur. J. Surg. Oncol.*, 35:4, 415-419 (2009).
33. S.A. SAVAGE, L. BURDETT, R. TROISI, C. DOUGLASS, R.N. HOOVER, S.J. CHANOCK, G. NATIONAL OSTEOSARCOMA ETIOLOGY STUDY, Germ-line genetic variation of TP53 in osteosarcoma. *Pediatr. Blood Cancer*, 49:1, 28-33 (2007).
34. A. SJALANDER, R. BIRGANDER, L. ATHLIN, R. STENLING, J. RUTEGARD, L. BECKMAN, G. BECKMAN, P53 germ line haplotypes associated with increased risk for colorectal cancer. *Carcinogenesis,* 16:7, 1461-1464 (1995).
35. M.M. CANDEIAS, L. MALBERT-COLAS, D.J. POWELL, C. DASKALOGIANNI, M.M. MASLON, N. NASKI, K. BOUROUGAA, F. CALVO, R. FAHRAEUS, P53 mRNA controls p53 activity by managing Mdm2 functions. *Nat. Cell Biol.,* 10:9, 1098-1105 (2008).
36. B.S. STRAUSS, Role in tumorigenesis of silent mutations in the TP53 gene. *Mutat. Res.*, 457:1-2, 93-104 (2000).
37. D.A. PILGER, P.L. LOPEZ, F. SEGAL, S. LEISTNER-SEGAL, Analysis of R213R and 13494 g-->a polymorphisms of the p53 gene in individuals with esophagitis, intestinal metaplasia of the cardia and Barrett's Esophagus compared with a control group. *Genomic Med.*, 1:1-2, 57-63 (2007).