Polymorphisms in G-quadruplex regions of the TP53 tumour suppressor gene : Impact on cancer susceptibility and expression of p53 N-terminal isoforms

Charlotte Sagne

▶ **To cite this version:**

Charlotte Sagne. Polymorphisms in G-quadruplex regions of the TP53 tumour suppressor gene : Impact on cancer susceptibility and expression of p53 N-terminal isoforms. Agricultural sciences. Université Paris Sud - Paris XI, 2013. English. NNT : 2013PA11T072 . tel-01089806

HAL Id: tel-01089806

https://theses.hal.science/tel-01089806

Submitted on 2 Dec 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Université Paris-Sud

Faculté de Médecine

Ecole doctorale de Cancérologie: Biologie, Médecine, Santé

THÈSE DE DOCTORAT SUR TRAVAUX

LABORATOIRE: INSERM U612 - Institut Curie

DISCIPLINE: Biologie

SPÉCIALITÉ: Génétique et Biologie Moléculaire

Présentée et soutenue publiquement le 27 novembre 2013 par

Charlotte SAGNE

Polymorphisms in G-quadruplex regions of the *TP53* tumour suppressor gene:

Impact on cancer susceptibility and expression of p53 N-terminal isoforms

---

Devant le jury composé de

Dr Joëlle WIELS	Présidente
Dr Sylvie MAZOYER	Rapporteur
Pr Thierry FREBOURG	Rapporteur
Dr Pierre HAINAUT	Examinateur
Dr Jean-Louis MERGNY	Examinateur
Dr Janet HALL	Directrice de thèse

ACKNOWLEDGMENTS

I would like to thank my thesis committee: Pr Thierry Freibourg and Dr Sylvie Mazoyer for their comments on this manuscript and for their questions and discussions, Dr Joëlle Wiels, Dr Jean-Louis Mergny and Dr Pierre Hainaut for their direction and their questions/discussions.

This thesis would not have been possible without the help, support, patience and good advices of my PhD supervisor: Dr Janet Hall. Thanks for all our discussion on p53 polymorphisms and G-quadruplexes and to accept me in your team with my own subject.

I also would like to do a special thank to Dr Pierre Hainaut who help me since 2008 on my project, thank you for your help and your patience and for our scientific discussions over these 4 years. I would like also to tell you that I greatly appreciated the independency you let me during my work in your lab.

I also thank all the Scientifics that I have meet during my work, just to name them Dr Magali Olivier, Dr Maria Isabel Achatz, Dr Patricia Prolla, Dr Marie-Paule Teulade-Fichou, Dr Maria Bota, Dr David Cox. During our discussions, the writing of articles and during meeting, I have developed my reflexion and my scientific approach.

Thanks to Ghyslaine and Agnès for you help during my first experiments in the lab and for your help during my famous WB and for sequencing all the LFS patients.

Thanks also to my colleagues from MOC and U612 teams, to my friends from ESTBB and to the others who have been supportive during the past three years.

I would like to do special thanks for some persons that I met during my licence/master/PhD: Dr Virginie Marcel, Dr Doriane Gouas, Sara Chiker, thanks for our scientific discussions (or not) in or out from the lab, your help and your support during the best and the worst moments.

Thanks also to my family for their support and for their understanding of my choice when I told them that I would like to do a PhD.

My final thanks is for Nadhim who accepted that I do a PhD, half in Lyon, half in Paris, thank you for following me in Paris. Thank you for supporting me and helping me with patience and comprehension during these 3 years.

TABLE OF CONTENTS

Preface *1*

Introduction: The TP53 gene and its modifiers *5*

Part I. The tumour suppressive protein p53 **7**

A. From oncogene to tumour suppressor: A short history of p53 7

B. The *TP53* gene 11

I. Chromosome 17p13.1 organisation 11

II. Role and functions of each p53 domain 12

C. Regulation of p53 23

I. p53 stability 24

II. Subcellular localisation 29

III. Epigenetic silencing of *TP53* promoter 30

IV. Post-transcriptional modifications 29

V. Post-translational modifications 33

VI. Regulation of p53 activity by its isoforms 37

D. p53 functions 38

I. Cell cycle arrest by controlling checkpoints 38

II. Apopotosis 42

III. Other p53 functions 45

IV. The p53 biological repertoire: orchestrating multiple biological functions 50

E. Genetic alterations of *TP53* in cancers 53

I. Somatic mutations in *TP53* gene 53

II. Germline mutations and Li-Fraumeni syndrome 61

III. Polymorphisms in *TP53* 68

Part II. p53 isoforms **64**

Part III. G-quadruplex structures **80**

A. Formation of G-quadruplex structures 88

B. G-quadruplexes chromosomal localisation 89

C. Biological functions of G-quadruplexes 90

IV. Regulation in telomeres 90

V. Role of G-quadruplexes in 5'UTR and promoter region 91

VI. Role of G-quadruplexes in mRNA production and stability 92

VII. Other roles of G-quadruplexes 93

D. G-quadruplexes ligands: biological effects 93

VIII. Biological responses induced by G-quadruplexes ligands 94

IX. Pathways induced by G-quadruplexes ligands 95

Objectives *97*

Results: Role of p53 polymorphisms in p53 expression and cancer susceptibility *101*

Part I. A meta-analysis of cancer risk associated with rs17878362 103

Part II. Effect of rs17878362 on age at cancer onset in germline *TP53* mutation carriers 122

Part III. Role of G4 in intron 3 on p53 mRNA splicing and p53 protein isoform expression 161

Discussion: TP53 polymorphisms, key modulators of p53 cancer susceptibility *216*

Part I. Association between rs17878362 and cancer susceptibility 221

A. Meta-analysis of rs17878362 in relation with risk of sporadic cancer 221

B. Analysis of rs17878362 in relation with risk of familial cancer 225

C. Effects of rs17878362 in sporadic or inherited contexts: an apparent paradox 228

Part II. Towards a functional hypothesis for genetic determinants of p53 regulation 231

Conclusion *236*

References *240*

Annexes *289*

LIST OF FIGURES

Figure 1: p53 history, from its discovery to clinical applications. 10
Figure 2: Schematic representation of the organisation of the human *TP53* gene. 11
Figure 3: Schematic representation of the structural organisation of p53. 13
Figure 4: p53 trans-activation domains functions. 15
Figure 5: Regulation of p53 stability. 24
(Figure 6: p53 regulation by miRNAs. 31
Figure 7: Post-translational modification of p53. 34
Figure 8: Role of p53 in the G1/S transition checkpoint. 40
Figure 9: Role of p53 in G2/M transition checkpoint. 42
Figure 10: Role of p53 in apoptosis. 43
Figure 11: *TP53* mutation distribution depending on cancer types. 54
Figure 12: The distribution of *TP53* somatic mutation types distribution 55
Figure 13: Oncogenic effects of p53 mutants. 58
Figure 14: Codon distribution of *TP53* mutation in LFS. 67
Figure 15: The most frequent polymorphisms found in the *TP53* gene and in its 3'flanking region 68
Figure 16: Schematic representation of human p53 isoforms. 76
Figure 17: Schematic representation of G4 formation and structures in DNA or RNA. 89
Figure 18: Schematic representation for the meta-analysis results for the *TP53* rs17878362 polymorphism 106
Figure 19: Kaplan-Meir disease-free probability estimates in LFS/LFL family members with or without *TP53* mutations. 124
Figure 20: GFP-reporter splicing assay to analyze the role of G4s in alternative splicing of intron 2 172
Figure 21: Effect of the G4 ligand 360A on expression of p53 transcript for A1A1 and A2A2 for rs17878362 lymphoblastoïd cell lines. 173
Figure 22: Schematic representation of the three most frequent polymorphisms between the intron 2 and the exon 4 of the *TP53* gene. 219
Figure 23: Decision tree applied to select publications using PubMed and Web of Science databases. 222
Figure 24: Effect of the rs17878362, rs1042522 and rs1625895 polymorphisms on cancer susceptibility depending on geographic origin and cancer type: a model. 224
Figure 25: Effect of different WT haplotypes in *TP53* mutation carriers: a model. 227
Figure 26: Model of the association of the *TP53* rs17878362 polymorphism in sporadic and germline *TP53* mutation context. 228
Figure 27: Model of the influence of the combination of the rs1642785 and rs17878362 polymorphisms on N-terminal p53 isoforms expression and p53 activity. 233

LIST OF TABLES

Table 1: Summary of phenotypic observation for Knock-in mutated mice carrying mutants at different important phosphorylation site of p53 (Adapted from (Jenkins et al 2012)). **35**

Table 2: Clinical criteria for classic Li-Fraumeni syndrome (LFS) and LFS-like (LFL) criteria, and Chompret criteria. **62**

Table 3: Example of G4 ligands and their pathway. **94**

ABBREVIATIONS

A

A	Adenosine
A	Alanine
aa	Amino acid
ADR	Adrenal cortical carcinoma
ARF	Alternative reading Frame
ASSP	Acid-soluble spore proteins
ATM	Atexia-Telangiectasia Mutated
ATP	Adenosine triphosphate
ATR	ATM and Rad-3-related protein

B

Bax	Bcl-2-associated X protein
Bcl	B-cell lymphoma
BER	Base excision repair
bFGF	Basic fibroblast growth factor
Bid	BH3-interacting domain death agonist
Bp	base pair

C

C	Cytosine
C	Cysteine
CAK	Cyclin-activating kinase
CBP	CREB binding protein
Cdk	Cyclin-dependant kinase
cDNA	Complementary DNA
Chk	Checkpoint kinase
CI	Confidence interval
COP	Constituvely photomorphic
COX-2	Cyclooxygenase-2
CTD	C-terminal basic domain

D

D	Aspartic acid
DBD	DNA binding domain
DNA	Deoxyribonucleic acid
DRAM	Damage-regulated autophagy modulator
DSB	Double strand break

E

EMT	Epithelial-mesemchymal transition
ERK	Extracellular-signal-regulated-kinases

F

F	Phenylalanine
FS	Fully spliced

G

G	Guanine
G	Glycine
G4	G-quadruplex
GADD45	Growth arrest and DNA damage 45

H

H	Histidine
HAUSP	Herpes associated ubiquitin-specific protease
HIF	Hypoxia inducible factor
hnRNP	Heterogeneous nuclear ribonucleoprotein
HR	Homologous recombination

J

JNK	c-Jun N-terminal kinases

K

K	Lysine
kDA	Kilodalton
kb	Kilo base

L

L	Leucine
LFL	Li-Fraumeni like
LFS	Li-Fraumeni Syndrome
LIF	Leukaemia inhibitor factor

M

M	Methionine
M	Mitosis
MAF	Minor allele frequency
MAPK	Mitogen-activated protein kinase
Mdm	Mouse double minute
miRNA	MicroRNA
MRE	Meiotic recombination
mRNA	Messenger RNA

N

n	Nucleotide
NEDD8	Neural precursor cell expressed developmentally down-regulated protein 8
NER	Nucleotide excision repair
NES	Nuclear export signal
NF-Y	Nuclear transcription factor Y
NHEJ	Non-homologous end-joining

NLS	Nuclear localisation signal
NMR	Nuclear magnetic resonance

O

OD	Oligomerisation domain
OR	Odd ratio

P

P	Proline
$p21^{WAF1}$	Cyclin-dependent kinase inhibitor 1
PGC	Peroxisome proliferator-activated receptor γ co-activators
PQS	Putative quadruplex sequence
PUMA	p53 up-regulated modulator of apoptosis
pRB	Retinoblastoma protein
PXXP	Proline-rich domain

Q

Q	Glutamine

R

R	Arginine
RE	Response Element
RING	Really interesting new gene
RLP	Ribosomal protein
RNA	Ribonucleic acid
ROS	Reactive oxygen species
RPA	Replication protein A

S

S	Serine
S phase	Synthesis phase
SH3	Scr homology 3-bind protein
SIRT	Sirtuin
SNP	Single nucleotide polymorphism
SP1	Specific protein 1
STS	Soft tissue sarcoma
SUMO	Small ubiquitin like modifier
SV40	Simian Vacuolating virus 40

T

T	Thymine
T	Threonine
TAD	Transcriptional Domain
TBP	TATA-box binding protein
TCA	Tricarboxylic acid cycle
TCAB	Telomerase Cajal body protein
TERC	Telomerase RNA component
TERT	Telomerase reverse transcriptase

TMPyP4	5, 10, 15, 20-tetra-(*N-methyl-4-pyridyl*)porphyrin chloride

U

USP	Ubiquitin-specific processing protease
UTR	Untranslated region
UV	Ultra violet

V

VEGF	Vascular endothelial growth factor

W

W	Tryptophan
WT	Wild-Type

Y

Y	Tyrosine

Z

Zn	Zinc

PREFACE

The p53 tumour suppressor protein is one of the most studied regulatory proteins in human cancer. This is firstly because it has been recognised since the late eighties, that the *TP53* gene is often mutated and the presence of these mutations is associated with cancer susceptibility and secondly that it carries out a multiplicity of biological functions that are regulated by complex and interlinked pathways. The p53 plays a role in all the functions implicated in the "Hallmarks of Cancer" a concept developed in 2000, and recently updated in 2011 (Hanahan and Weinberg 2000, Hanahan and Weinberg 2011), to understand the molecular mechanisms of carcinogenesis such as proliferation, cell death, growth, senescence, genomic and genetic stability. This multi-action capacity is perhaps the main explanation for the "success" of *TP53* as cancer gene.

Given the multiplicity of its effects and functions, it is of no surprise that the *TP53* gene is highly polymorphic, with significant differences in the frequency of different polymorphic alleles across populations. The *TP53* gene can be expressed as multiple ribonucleic acid (RNA) transcripts, generated by using alternative splicing and alternative promoters, leading to a large diversity at the protein level. The impact of this diversity on individual and population cancer risk however is still poorly understood.

In the nineties, technological developments allowed the tri-dimensional structures of proteins to be determined and opened up the possibility of correlating deoxyribonucleic acid (DNA) and RNA sequences to their formation and their role in transcription/translation regulation. More recently, tri-dimensional structures formed in G-rich sequence in both at the DNA or RNA level called G-quadruplexes (G4) were observed. These structures are involved in different cellular processes such as gene transcription (Bochman et al 2012), genomic stability (Paeschke et al 2013, Ribeyre et al 2009), DNA replication (Cayrou et al 2012, Paeschke et al 2011), messenger RNA (mRNA) splicing (Millevoi et al 2012) and mRNA stability (Millevoi et al 2012).

In this Thesis, I have examined the impact of the complexity of the *TP53* gene and in particular its polymorphic variability on cancer susceptibility in the context of both population and familial settings. I have focused on polymorphisms occurring in G4 structures formed in the p53 pre-mRNA. My results have shown that some polymorphisms located in or close to G4 structures have an impact on the age of cancer onset in subjects carrying a germline *TP53* mutation, as well as on the cancer susceptibility in the general population. Given the possibility that G4s in *TP53* may affect mRNA splicing, stability and p53 isoform expression, these results suggest a possible relation between the genetic diversity of the *TP53* gene and the complexity of p53 regulatory pathways.

INTRODUCTION: THE *TP53* GENE AND ITS MODIFIERS

Part I. The tumour suppressive protein p53

A. From oncogene to tumour suppressor: A short history of p53

In the past thirty years, p53 has emerged as one of the most important molecular factors in human cancer. The term the "p53" protein is used to identify an essentially nuclear phospho-protein with an apparent molecular weight of 53 kDa. The gene encoding p53 is called *TP53*, whereas its mouse homolog is known as *Trp53*. To date, mutations in *TP53* are the most common genetic event in human cancer (p53.iarc.fr). Many of these mutations are missense and are identified by their codon number, preceded by the amino-acid (aa) encoded by the wild-type (WT) codon and followed by the aa encoded by the mutant codon (for example: mutation R175H: substitution of an arginine (R) by a histidine (H) at codon 175). These abbreviations will be widely used throughout this Thesis manuscript.

In the mid 1950's, small DNA tumour viruses such as Simian Vacuolating virus 40 (SV40) and polyoma virus were discovered and were shown to carry the necessary genes, termed viral oncogenes, to lead to a tumorigenic phenotype in infected cells (Eddy et al 1962). Over the following twenty years the question of how these virally encoded proteins can initiate this process and lead to the immortalisation of cells in culture was extensively studied. In 1979, several groups reported an interaction between the SV40 large T-antigen, one of the two main proteins expressed by the SV40 tumour virus, and a 53 kilodalton (kDa) protein, p53. Lane and Crawford and May and collaborators demonstrated that the SV40 large T-antigen co-precipitated with a 53 kDa protein in cells and that these cells contained an equivalent amount of both, suggesting that they two proteins were in a stoichiometric complex in the cell extracts (Lane and Crawford 1979, May et al 1979). Linzer and Levine used antisera from animals carrying SV40 induced tumours to detect both p53 and the viral T-antigen (Linzer and Levine 1979). They also showed that these antibodies immuno-precipitated the p53 protein from transformed cells not infected by SV40, establishing that the p53 protein was a cellular protein and suggesting a possible transforming or oncogenic potential. DeLeo et al reported that animals infected with spontaneously transformed or tumorigenic cells produced antibodies to the p53 protein (**Figure 1**) (DeLeo et al 1979). Taken together these studies suggested that p53 was necessary for the transforming or oncogenic action of the SV40 tumour virus and that p53 is an oncogene.

The next phase of p53 research was to determine the identity of the gene encoding p53, and to characterise its functions (**Figure 1**). This took about 15 years, during which the status of p53 switched from an "oncogene" to a "tumour suppressor gene". The first Complementary DNA (cDNA) and genomic sequences of *TP53* was isolated from immortalized cells. Sequences variations were reported depending on whether tumour or normal cells or tissues were the source of the DNA, leading to confusion as to which of these sequences represented the WT allele. These differences were deciphered by the seminal studies of several groups. Firstly Levine's laboratory showed that the WT p53 protein alone could not transform cells (Reich et al 1983). Moreover, they observed that a mutant p53 protein was often found in cancer cells and could inhibit the transforming activity of oncogenes (Finlay et al 1989, Harvey and Levine 1991, Hinds et al 1989). Secondly the group of Vogelstein reported that *TP53* mutations and loss of alleles were common events in human colon carcinomas (Baker et al 1989, Nigro et al 1989). In the early eighties, it was also shown that p53 was induced in response to DNA damage (Maltzman and Czyzyk 1984) and induced cell death by apoptosis (Yonish-Rouach et al 1991). These properties suggested that p53 was behaving as tumour suppressor protein, in a manner similar akin to the retinoblastoma gene product that had been recently identified (Geiser and Stanbridge 1989). The discovery that germline *TP53* mutations were the genetic defect associated with the Li-Fraumeni syndrome (LFS), characterised by a predisposition to multiple and early onset-cancers (Malkin et al 1990, Srivastava et al 1990), led to the recognition of *TP53* as the "*ultimate tumour suppressor gene*" (Oren 1992).

In 1992, David Lane described p53 functions as those of "the guardian of the genome" (Lane 1992), based on the role of p53 in protecting cells against carcinogenesis DNA damage and abnormal proliferation. Many studies were undertaken to understand the impact of *TP53* mutations in cancers and the molecular mechanisms of p53 effects. The first important role of the p53 protein discovered was its implication in apoptotic death (Yonish-Rouach et al 1991). Next, p53 was described as a transcription factor with a sequence response element that allowed the targeting of a panel of genes (el-Deiry et al 1992, Funk et al 1992). The auto-regulatory loop of p53 stability via Mdm (Mouse Double Minute) 2 (Barak et al 1993), the crystallographic structure of the p53 core domain bound to DNA (Cho et al 1994) and the notion that *TP53* mutations could be understood as "molecular fingerprints" of carcinogens (Hollstein et al 1991), all contributed to the better understanding of the relationships between DNA damage and cell proliferation and establishing the critical role of p53 in cancer.

In 1994, the discovery of the role of p53 in the response to cytotoxic therapeutic treatments suggested that its manipulation could be used to therapeutic benefit (**Figure 1**). For

example, it was shown that the drug PRIMA-1 can restore the suppressive functions of a mutant p53 protein (Bykov et al 2002) and that Nutlin-3 is an antagonist of the interaction between p53 and Mdm2 causing the release of p53 from this complex and by thus controlling its degradation and inducing p53 accumulation in a DNA-damage independent way (Vassilev et al 2004). Although promising, these drugs still have to find their application in cancer treatment. At about the same time, many studies showed that sequence variations in *TP53* (mutations or polymorphisms) could be used as biomarkers for cancer development, treatment response, prognosis and survival (Petitjean et al 2007, Whibley et al 2009). Databases such as the IARC *TP53* database (p53.iarc.fr) and the p53 Web Site (p53.free.fr) were created to compile and categorise these variations. Since the mid-nineties, studies on the p53 regulatory network have allowed a better understanding of p53 functions aided by the discovery of p73 and p63, two proteins encoded by *TP53*-related genes, which are implicated in development, morphogenesis and stress responses (Lane and Levine 2010). In addition, molecular epidemiological studies have more recently led to a new concept: human p53 isoforms, which can also act a large diversity at the protein level as regulators of p53 expression and function (Bourdon et al 2005, Courtois et al 2002, Marcel et al 2011).

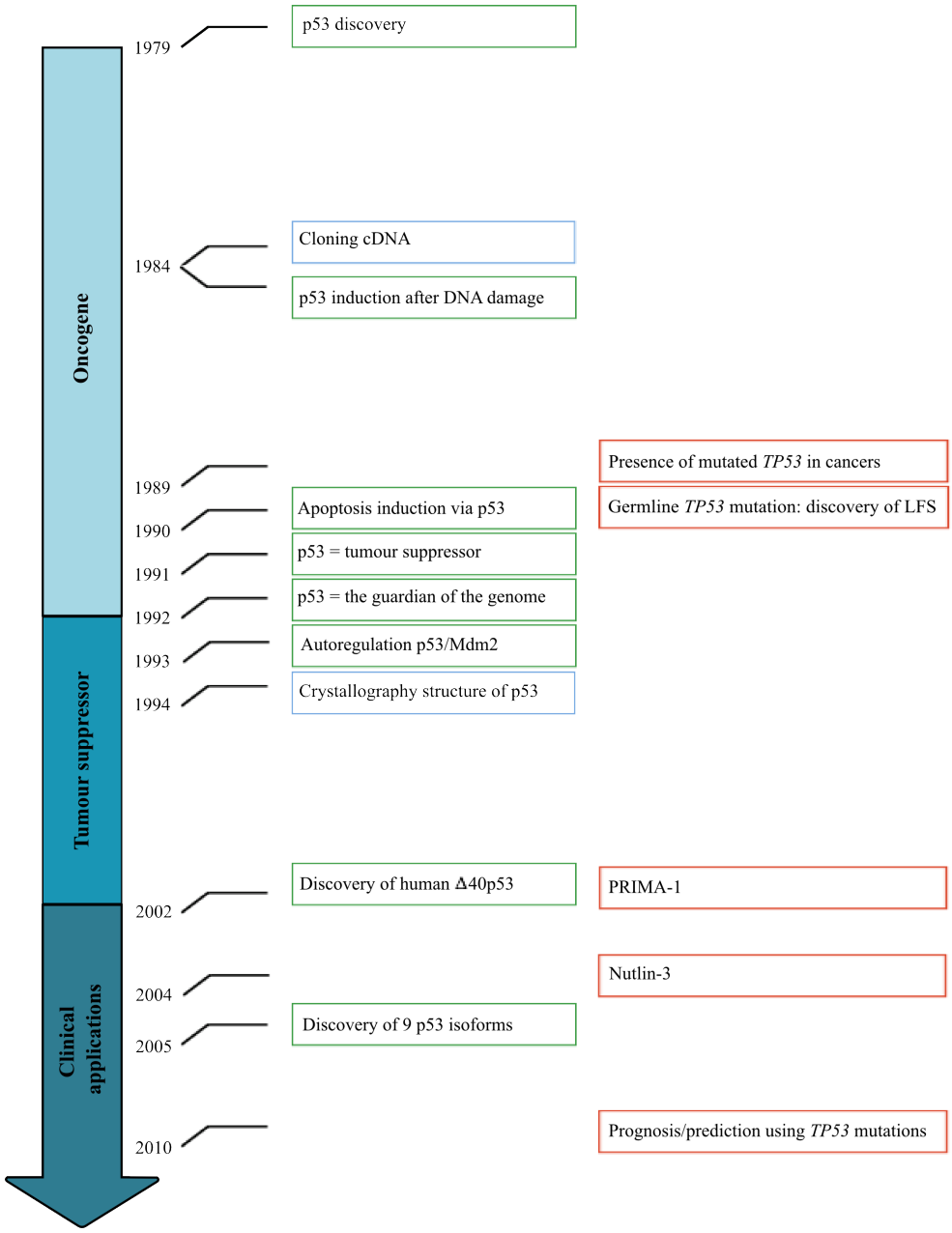


Figure 1: p53 history, from its discovery to clinical applications. Important steps from discovery of p53 in 1979 to clinical applications in 2010. Green boxes: important p53 discoveries; Blue: important steps in p53 methods; Red: current developments from *TP53* mutation studies in cancer therapy using p53 targeting drugs and *TP53* mutation profiles as a prognostic biomarker. Adapted from (Hainaut and Wiman 2009) and Thesis of Virginie Marcel, University Lyon I, 2009.

B. The *TP53* gene

I. Chromosome 17p13.1 organisation

The human *TP53* gene is located on chromosome 17 on the short arm at locus 17p13.1 (Isobe et al 1986, McBride et al 1986). The *TP53* locus is unique on this chromosome arm in having a telomere to centromere orientation, whereas flanking genes are in the reverse orientation. The gene spans about 20 kilobases (kb) and is composed of 11 exons, with a small non-coding first exon and a large conserved intron of about 10 kb (**Figure 2**). The gene lacks a conventional TATA box but contains several sequences with promoter activity, which may regulate p53 expression. The proximal promoter (P1) is located upstream of exon 1 (Lamb and Crawford 1986) and produces the full spliced p53 (FSp53) mRNA and p53I2 mRNA by alternative splicing. The second promoter (P1') is located within the long intron 1 and directs the synthesis of a 1.1 kb mRNA derived from sequences from intron 1 (Hint1p53 transcript). This transcript contains several short open-reading frames but the putative proteins have not been identified so far. Its role and function are unknown (Reisman et al 1988, Reisman et al 1996). Hint1p53 is expressed in a number of human cells and is induced during the terminal differentiation of myeloid leukaemia cells. The third promoter (P2) is located between the end of exon 2 and the beginning of exon 5 and gives rise the p53I4 mRNA that lacks sequences from exon 2 to 5 (Bourdon et al 2005). The organisation of the *TP53* gene is highly conserved through evolution (Soussi et al 1987).

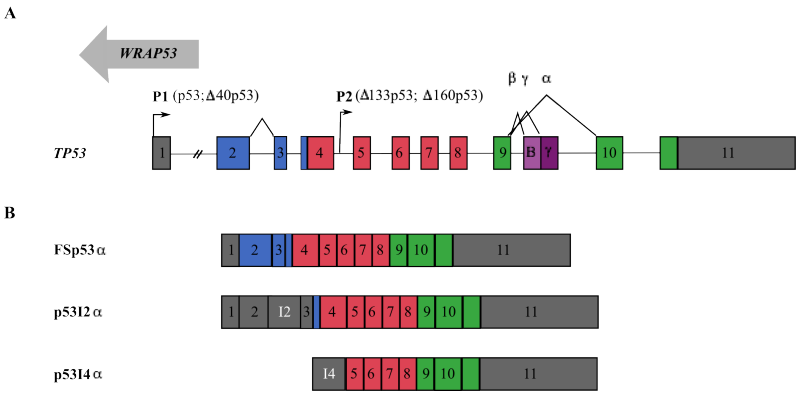


Figure 2: Schematic representation of the organisation of the human *TP53* gene. (**A**) The *TP53* gene structure (Coloured boxes: coding exons; grey boxes: non-coding exons, ↱: alternative promoters; ∧: splicing sites) and the localisation of the *WRAP53* gene, which overlaps exon 1 and intron 1 of the *TP53* gene. The WRAP53 gene is orientated on the opposite direction to *TP53* (**B**) p53 mRNA variants. The proximal promoter P1, located upstream from exon 1, regulates the transcription of FSp53, which encodes both p53 at ATG1 and Δ40p53 at ATG 40 and p52I2 mRNA, which retains the intron 2 containing stop codons by alternative splicing and encodes the Δ40p53. The second promoter P2, located between intron 1 and the exon 4, produces one transcript, the p53I4 mRNA. This mRNA can translate two different products, the Δ133p53 (from ATG 133) and the Δ160p53 (from ATG 160). Adapted from (Bourdon et al 2005, Marcel et al 2011).

The *TP53* gene contains 85 polymorphisms described in the IARC *TP53* Database (p53.iarc.fr) and is considered as a highly polymorphic gene. It has been demonstrated that some polymorphisms such as rs1042522 located in exon 4 at codon 72 (guanine (G) > cytosine (C); R > proline (P)) can modulate p53 functions (Siddique et al 2005, Whibley et al 2009) or cancer susceptibility (Whibley et al 2009). Some of these polymorphisms and their functions are described later in the Introduction, Section E, Part III of the Thesis manuscript.

On this part of chromosome 17p13.1, another gene, *WRAP53*, also called *telomerase Cajal body protein-1* (*TCAB1)* has also been described (Mahmoudi et al 2009). This gene is oriented in the opposite direction from *TP53* and is located immediately upstream of *TP53*, with an overlap between exon 1 and part of intron 1 of *TP53* and exon 1α of *WRAP53*. The transcription level of TP53 is 100-fold higher than that of WRAP53. It was demonstrated that WRAP53 RNA could regulate TP53 RNA translation via an interaction between the sense strand region of WRAP53 and TP53 RNA. This interaction induces Wrap53 and sensitizes cells to p53-dependent apoptosis upon DNA damage (Farnebo 2009, Mahmoudi et al 2009). The interaction between Wrap53 and p53 is also required to maintain a normal level of p53 in cells (Farnebo 2009). Thus, Wrap53 is a new regulator of p53 at the post-transcriptional level.

II. Role and functions of each p53 domain

The p53 protein functions as a tetramer to bind specifically to DNA and acts as a transcription factor for target genes regulated through p53-binding DNA response elements (RE) (**Figure 3**). The p53 protein is divided into 5 distinct structural and functional domains: (1) the N-terminal domain located between aas 1 and 72, (2) the proline domain between aas 73 and 94, (3) the DNA binding domain into aas 102 and 292, (4) the oligomerisation domain at aas 326-355 and (5) the C-terminal basic domain between aas 363 and 393.

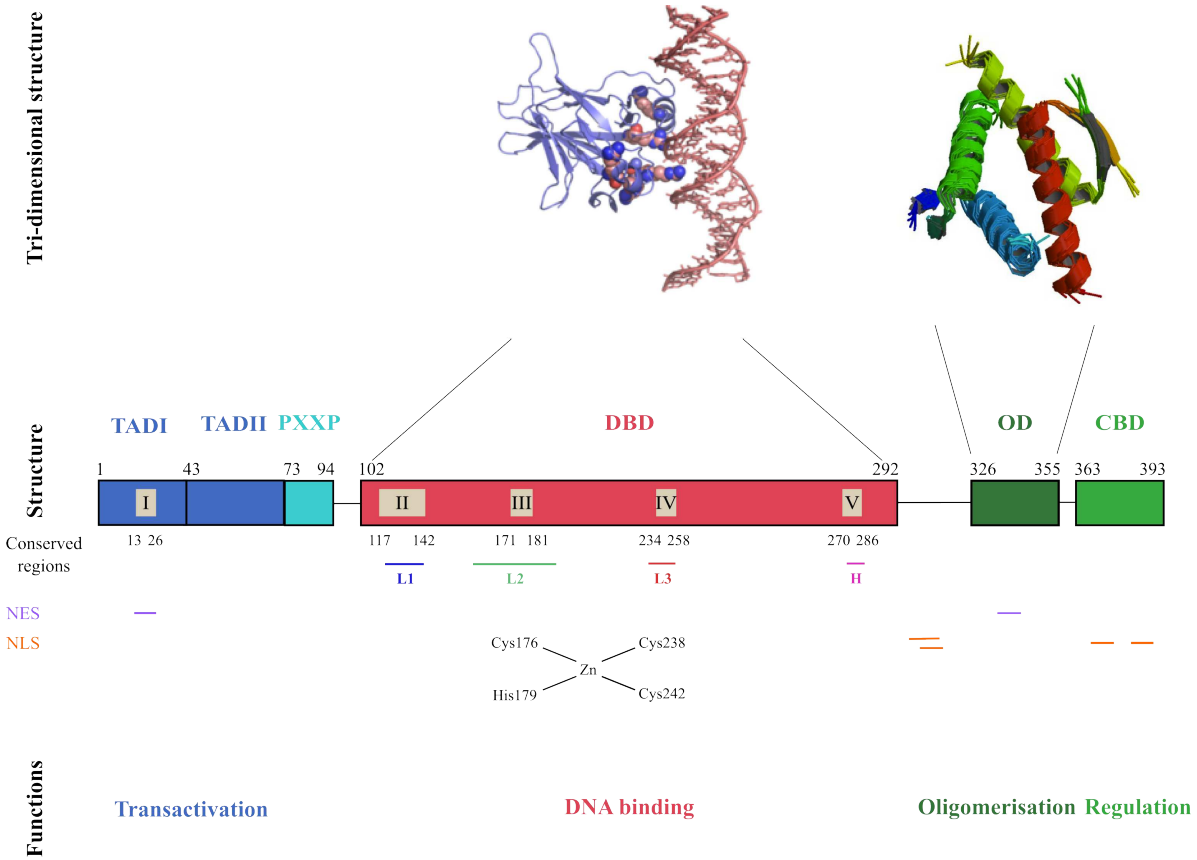


Figure 3: Schematic representation of the structural organisation of p53. The p53 protein has a classical conformation with 4 important domains ("Structure" line) with particular functions ("Functions" line). The tri-dimensional conformations for the DBD and the OD, only, are known ("Tri-dimensional structure" line). The conserved regions, the Nuclear Export Signal (NES) and the Nuclear Localization Signal (NLS) are also represented. The 3 loops (L1 to L3), the hydrophobic region (H) and the tertiary structure of the DBD are shown. TAD: Trans-activation Domain; PXXP; Proline Domain; DBD: DNA Binding Domain; OD: Oligomerisation Domain; CBD: C-terminal Basic Domain. Adapted from (May and May 1999).

1. N-terminal trans-activation domain (TAD, amino acids 1-72)

Structure

The trans-activation domain (TAD) of p53 is located in the N-terminal part of the protein. This region contains a transcriptional activation motif ϕ-x-x-ϕ-ϕ (ϕ = hydrophobic aa and x = any other aa), which is found in many proteins regulating transcription. This region contains a few hydrophobic aas (found at positions 7-9, 16, 17, 26, 58-60), which disfavour the formation of a hydrophobic secondary structure and confer a natively unfolded form to the domain (Dawson et al 2003, Lee et al 2000). This unfolded structure plays an essential role in the regulation of the p53 network by enabling interactions with other proteins. The domain can be divided into two sub-domains: TADI consisting of aas 1-42, which carries the main transcriptional activity, and TADII consisting of aas 43 to 72, which carries a second motif acting essentially as a regulator of TADI. TADI contains an amphipathic α-helix and a nascent turn in the TADII (Lee et al 2000). Through these structural differences, TADI and

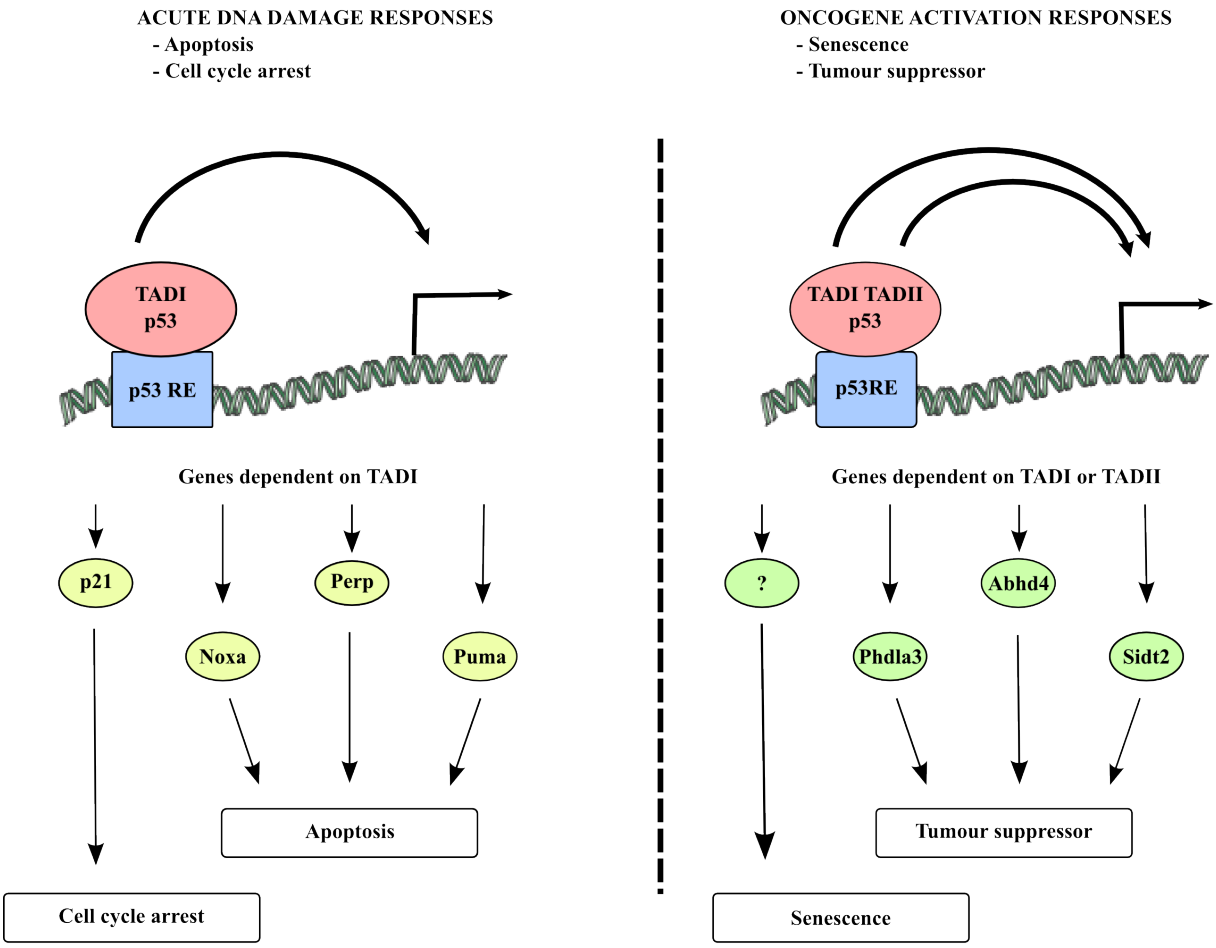
TADII should have distinct activities and also a common activity, which needs both domains. The TADI sub-domain contains a nuclear export signal (NES) between aas 13 and 26 (Zhang and Xiong 2001b).

Functions

Aas 13-26 of TADI represent one of the most conserved regions in the p53 protein through evolution (Soussi et al 1990). The main function of TADI is the interaction with others components of the transcription machinery such as the TATA box binding protein (TBP) and TBP-associated factors components of transcription factor TFIIB (Lu and Levine 1995, Thut et al 1995). It can also interact with viral proteins such as E1B from adenoviruses, which inhibit p53 trans-activation activity (Yew and Berk 1992).

To increase p53 trans-activation activity, cofactors and histone acetyl-transferases (HATs) interact directly with the TADI and TADII (Riley et al 2008). Recently, Attardi and collaborators showed that the TADI and TADII sub-domains, and both TAD domains together interact with different cofactors and HATs (**Figure 4A**) (Bieging and Attardi 2012, Brady et al 2011, Jiang et al 2011). These authors produced transgenic mice with mutations at leucine (L) at codon 25 into glutamine (Q) (L25Q) and at tryptophan (W) at codon 26 into serine (S) (W26S), resulting in inhibition of TADI function or at Phenylalalnine (F) at codon 53 into Q (F53Q) and F54S, inhibiting TADII's role or L25Q, W26S, F53Q and F54S, inactivating both TAD domains (**Figure 4B**) and analysed the differences in p53 protein interactions and in expression of p53-regulated genes. They observed that TADI is involved in the activation of p53 target genes important for the response to acute DNA damage, cell-cycle arrest or apoptosis, but is not capable of tumour suppression in mice. The TADII sub-domain, in contrast to TADI, has no autonomous effects on p53 trans-activation capability or biological activity. However, collectively, the two TADs are essential for p53's tumour suppressor activity. Overall, these results suggest that p53's TAD mediates tumour suppressor activities via TADI-dependent trans-activation of p53- target gene and TADII contributes to tumour suppression by regulating TADI activities via unknown mechanisms. Moreover, each TAD appears to interact with specific cofactors. For example, TADI interacts with TBP, whereas TADII interacts with TFIID and with replication protein A (RPA) (Bochkareva et al 2005, Brady et al 2011). Together, both TADs interact with HATs p300 and CRE Binding protein (CBP) (Ferreon et al 2009, Teufel et al 2007).

A



B

	WT p53	TADI-null p53	TADII-null p53	p53-null
Transactivation	+	+/-*	+	-
Cell cycle arrest	+	-	+	-
DNA damage apoptosis	+	-	+	-
Senescence	+	+	+	-
Tumour suppressor	+	+	+	-

Figure 4: p53 trans-activation domains functions. (**A**) Models for p53 activity depending on the type of stress. For the acute DNA damage response, only the TADI sub-domain is implicated and the p53 response mechanism is described on the left panel. After oncogenic activation responses, both TAD sub-domains are implicated in the p53 response and the target genes implicated are described in the right panel. (**B**) Summary of TAD functions in transgenic mutated mice. TAD: Trans-activation domain; WT p53: entire TAD (TADI + TADII); TADI p53-null: mutation on aas 25 and 26; TADII p53-null: mutation aas 53 and 56; TAD p53-null: mutation on aas 25, 26, 53 and 56; * Minimal trans-activation of most, but not all, p53 target genes. Adapted from (Brady et al 2011).

Role in p53 regulation

The TAD is the most important domain for the regulation of p53 activity involving three pathways. The first is the p53-Mdm2-autoregulatory feedback loop (Barak et al 1993), the second is the post-translational modification of the serines and threonines present in p53 TAD and the third is the formation of a complex between the TAp53 and Δ40p53 isoforms (Hafsi et al 2013).

In 1992, p53 was found to trans-activate *MDM2*, which encodes the Mdm2 protein, a E3 ubiquitin ligase targeting p53 for ubiquitination and proteasome-dependent degradation (Barak et al 1993, Honda et al 1997, Kubbutat and Vousden 1998, Momand et al 1992, Oliner et al 1993, Wu et al 1993). Mdm2 binds to a motif at aas 18-23 (TFS acid aspartic (D) LW motif) located in the TADI sub-domain (Picksley et al 1994). This p53-Mdm2 interaction results in the inhibition of the p53 transcriptional activity by masking the p53 TADI sub-domain. It also promotes the ubiquitination of p53 and induces its ubiquitin-dependent proteosomal degradation (Haupt et al 1997, Honda et al 1997). This regulation constitutively maintains p53 at a low cellular level under normal conditions.

The p53 TAD contains several post-translational modification sites that are phosphorylated by a number of activated kinases and are critical for protein-protein interactions that either modulate the stability and subcellular localization of p53 or affect its function as a transcription factor. These regulatory sites and their upstream kinases are presented in Section I, Part C of this Thesis manuscript.

2. Proline-rich domain (PXXP, amino acids 73-94)

Structure

The p53 proline-rich domain (PXXP) is located between the N-terminal and the DNA binding domain at aas 73-94 (Walker and Levine 1996). The human p53 protein contains 5 repeats of the PXXP motif. This domain contains a frequent exonic polymorphism at codon 72, rs1042522, which consists of a substitution of a guanine (G) to a cytosine (C) corresponding to the substitution of arginine to a proline (Matlashewski et al 1987). In the presence of the arginine variant, the proline-rich domain loses one of its PXXP domains.

Functions

PXXP corresponds to the Scr homology 3-binding protein (SH3) motif suggesting that it may be involved in the physical interaction with elements of signal transduction pathways that contain SH3 domains, including for example inhibitor member of the acid-soluble spore

protein (ASPP) family (iASSP) (Bergamaschi et al 2006) and c-Abelson murine leukemia viral oncogene (Abl), a kinase activated by Ataxia-Telangiectasia Mutated (ATM) (Khanna et al 1998). The deletion of the proline-rich domain does not compromise transcriptional activity or DNA binding activity (Sakamuro et al 1997, Walker and Levine 1996). In addition, in some experimental systems, it was shown that the proline-rich domain was implicated in p53-mediated apoptosis (Sakamuro et al 1997), in suppressing tumour cell growth (Walker and Levine 1996) in the degradation of p53 by the E6 protein of oncogenic human Papilloma Viruses (Li and Coffino 1996). This region of the gene encoding the proline-rich domain is rarely mutated in human cancer.

Role in p53 regulation

To date, few studies have been performed to elucidate the activity of the p53 proline-rich domain in regulating p53 tumour suppressor activity. The mutant T81A p53 failed to induce p53-mediated cell cycle arrest and apoptosis in response to ultra-violet (UV) damage, suggesting that the phosphorylation of threonine 81 by c-Jun N-terminal kinases (JNK) 2 is important for these responses (Buschmann et al 2001). Furthermore, Berger and collaborators observed that deletion of the proline-rich domain increased the sensitivity of p53 to inhibition and degradation by Mdm2, suggesting that the proline rich domain is important for the fine-tuning of p53-Mdm2 interactions (Berger et al 2001). Indeed, they observed, using mutants of the proline-rich domain, that this regulation was mediated by phosphorylation of proline 82 (Berger et al 2005). Phospho-proline 82 is important for p53-Checkpoint Kinase 2 (Chk2) interaction, which allows the phosphorylation of p53 at S20. When S20 is phosphorylated, the interaction p53-Mdm2 is inhibited, preventing p53 degradation via Mdm2 ubiquitination (Craig et al 1999).

3. DNA binding domain (DBD, amino acids 102-292)

Structure

The DNA binding domain (DBD) is located between aas 102 and 292. It is also called the “core” domain and is resistant to proteolysis by enzymes such as thermolysin or proteolytic digestion (Bargonetti et al 1993, Pavletich et al 1993). This domain contains four of the five evolutionary conserved sub-domains of p53: domain II (aas 117-142), domain III (aas 171-181), domain IV (aas 234-258) and domain V (aas 270-286) (Soussi et al 1990). In 1994, Cho and collaborators have determined the crystal structure of the p53 DBD in complex with a DNA target gene (Cho et al 1994). This structure presents a hydrophobic domain

composed by two sheets of β-strands, and a set of loops and α-helices that contains the aas, which form a direct contact with p53 RE DNA. The binding surface of p53 is made of two parts: L2 and L3 loops (two conserved region through evolution) and a short α-helix that bind within the minor groove of p53 RE DNA, with R248 as main DNA interacting aa. L2 and L3 are bridged by a zinc atom, which is linked to cysteine (C) 176, H179 on L2 and C238 and C242 on L3. L1, together with the short S2-S2' β hairpin (aa 124-141) and a large α-helix, forms a motif that occupies the major groove of DNA, with direct aas contacting DNA (including R273, C277, R282 and Lysine (K) 120).

Functions

The majority of the mutations found in the human *TP53* gene are located in the DNA binding domain of p53 (Petitjean et al 2007). A survey of studies using whole genome sequencing suggests that at least 70% of all cancer-related mutations are located in this domain (COSMIC database, http://cancer.sanger.ac.uk/wgs/gene/overview?ln=TP53). This high frequency of mutations is probably due to the alteration of the sequence-specificity of the p53 DNA binding activity by the presence of mutations in the DBD, which can modulate the direct contact with DNA or change the conformation of the DNA-binding structure and thus confer some selective advantage to the "mutated" protein. Indeed, the DBD can bind its target genes only by using a specific RE that allows the recognition of p53 target genes. It is constituted of a repeat of an inverted palindromic 10 base pairs (bp) element matching the consensus 5'-RRRCWW glycine (G) YYY (n=0-13) RRRCWWGYY-3' (R is a purine, W a adenine or thymine base and Y a pyrimidine base) (el-Deiry et al 1992, Funk et al 1992). This sequence orientation increases the specificity of the binding (Funk et al 1992) and reflects the fact that p53 activity depends of the formation of a tetramer consisting of two dimers (Kitayner et al 2006, McLure and Lee 1998) with each p53 monomer in one of the dimers interacting with one half of the 10 bp site (RRRCW). Using reporter assays, Funk and collaborators showed that this palindromic sequence is necessary for p53 DNA binding activity (Funk et al 1992). This sequence is found in between 300 and 1600 potential binding sites in the human genome (Cawley et al 2004, Hoh et al 2002). To date, over 125 protein-coding genes and noncoding RNAs have been shown to be the direct transcriptional target of p53, all of which contain the p53 RE to which p53 binds activating their transcription (Poyurovsky et al 2010, Riley et al 2008). Interestingly, the p53 RE does not perfectly match the consensus sequence in the majority of these cases. The p53 protein thus may have different affinities for some p53 REs, which could explain the large number of p53 target

genes. Another activity of the p53 DBD is the trans-repression of some genes, but the mechanisms involved are less well understood (Oren 2003). For example, the *B-cell lymphoma* (*BCL)-2* promoter contains a p53 RE, which overlaps with the binding site of another more potent activator (Budhram-Mahadeo et al 1999). When p53 binds to the *BCL-2* promoter, the target mRNA level is decreased.

Role in p53 regulation

The p53 DBD is the domain with the least post-translational modification sites known. For instance, no phosphorylation sites have been identified (Riley et al 2008). Recently, some studies have shown that the acetylation of lysine K120 and K164 in this region could increase the trans-activation of *p53 up-regulated modulator of apoptosis (PUMA)*, which influences p53-apoptosis (Sykes et al 2006, Tang et al 2006, Tang et al 2008). The ability to bind DNA is dependent on the DBD's interaction with a Zinc $(Zn)^{2+}$ ion via aas in two clusters around C176-H179 and C238-C242 (Hainaut and Milner 1993). The oxidation of these two particular clusters decreases the ability of the protein to bind to p53 RE and trans-activate target genes *in vivo* (Parks et al 1997). The first cluster contains the aas that interact with a Zn ion and the second is close to the aas implicated in the contact of p53 with its consensus sequence.

Two important ways to influence p53 DNA binding domain activity are well described. The first is the presence of a mutation in the DBD (Olivier et al 2002) and especially ones affecting the bases implicated in the Zn^{2+} ion interaction with the DBD (Kern et al 1991, Kim et al 1997). The second is the structure of the C-terminal basic domain as described later in the text (Hamard et al 2012).

4. Oligomerisation domain (OD, amino acids 326-355)

Structure

The oligomerisation domain (OD) is located between aas 326-355 and is also called the tetramerisation domain (Chene 2001). The structure of this domain has been determined by both X-ray crystallography and nuclear magnetic resonance (NMR) (Clore et al 1994, Jeffrey et al 1995) and is essential for the formation of the tetramer of p53 proteins that consists of a dimer of dimers (Kraiss et al 1988, Lee et al 1994, Mateu et al 1999). In each monomer this region takes up the following conformation: a β-strand (aas 326-333) and an α helix (aas 335-355) linked by a single aa – a G at aa 334. To form a dimer, two monomers of p53 interact by their β-strands to form an antiparallel β-sheet, which allows the association two dimers into a tetramer though their helices. The interface between the helices involved

hydrophobic contact between methionine (M) 340, L344, alanine (A) 347, L348 and L350 and Hydrogen-bond donation between R337 and D352. Mutations at these aas prevent the formation of tetramers (Chene et al 1997, Mateu and Fersht 1998, Waterman et al 1995).

A region containing two important domains for p53 localisation is located between the p53 DBD and the OD. The first, the dominant nuclear localisation signal (NLS)-I also which is upstream of the OD at aas 315 to 325 and a weaker NLS also located upstream of the OD, in a flexible linker region at aas 305-322 (Shaulsky et al 1990b). The second domain also contains the NES (Zhang and Xiong 2001b), which is located at the end of the OD, at aas 339-352.

Functions

The p53 OD plays several roles in p53's DNA binding activity, p53-protein interactions and cellular localisation. First, the formation of p53 oligomers by the OD increases p53 DNA binding activity. Indeed, it has been shown that p53 monomers can interact with p53's REs in a cooperative manner and trans-activate p53 target genes but with a 10-100 fold lower affinity than a p53 tetramer (Balagurumoorthy et al 1995). Secondly, the p53 OD is important for p53's interaction with others proteins. For example, the OD plays an indirect role in p53-Mdm2 interaction (Lomax et al 1998), or with p53-TBP interactions (Liu et al 1993).

Role in p53 regulation

Post-transcriptional modifications can modulate the dynamics of p53 oligomerisation. For example, the phosphorylation of S392 can increase the association constant for oligomerisation by 10-fold *in vitro* (Waterman et al 1996). This domain can interact with proteins such as REGγ, a proteasome activator, which can regulate the cellular distribution of p53 by increasing the mono-ubiquitination of p53 and its subsequent nuclear export and degradation (Liu et al 2010b). These authors also observed that REGγ inhibits p53 tetramerisation and this might enhance p53 cytoplasmic relocalisation, which decreases the amount of active p53 in the nucleus. Zhao and collaborators established mouse lines with or without the NLS-I (Zhao et al 1999). They observed that both trp53-null cells and NLS-I-null cells were deficient for p53-dependent apoptosis after exposure to γ-ionizing radiation or hydrogen peroxide treatment. These results suggest that NLS-I is necessary for p53-mediated apoptosis.

5. C-terminal basic domain (CTD, amino acids 363-393)

Structure

The C-terminal basic domain of p53 (CTD) is located between aas 363-393. Its main characteristic is its natively unfolded form, which confers the capacity to interact with DNA in a sequence nonspecific fashion (Foord et al 1991, Wang et al 1993). It can interact with different DNA structures including insertion/deletion mismatches (Lee et al 1995), γ-irradiated DNA (Miyashita and Reed 1995) or supercoiled DNA (Mazur et al 1999). This domain contains two minor NLS (aas 369-375: NLSII and 379-384: NLSIII) (Dang and Lee 1989, Shaulsky et al 1990b), multiple ubiquitination sites (Michael and Oren 2003), one major site of sumoylation (K386) (Gostissa et al 1999, Rodriguez et al 1999) and several stress-inducible modification sites for phosphorylation, acetylation or glycosylation (Appella and Anderson 2001).

Functions

Early research investigating the functions of this domain was focused on its ability to regulate the sequence-specific DNA binding of p53 (Ahn and Prives 2001). Deletion of the last 30 aas of p53 and the use of the monoclonal C-terminal antibody polyclonal antibody (PAb) 421, a basic peptide which inhibits C-terminal function, resulted in an increase in p53 sequence-specific DNA binding suggesting that the CTD is important for inhibiting the high-affinity p53 DNA binding activity (Hupp et al 1992, Hupp et al 1995). These observations lead to the allosteric hypothesis namely that the C-terminal domain could control the conversion of p53 from a latent to an active form by its interaction with cellular factors resulting in enhanced sequence-specific DNA binding and the transcriptional activity of p53 (Hupp et al 1992, Hupp et al 1995). The p53 latent form was shown to arise from the interaction between the aas 80 and 93 (in the PXXP domain) with its C-terminal domain (Muller-Tiemann et al 1998). More recent studies led to the re-evaluation of this model. Anderson and collaborators found that the nonspecific DNA commonly used in standard *in vitro* DNA-binding assays (for example: Electrophoretic mobility Shift assays (EMSA)) was acting as an inhibitor of p53 DNA binding (Anderson et al 1997). In addition, Ayed and colleagues observed that there was no difference in the NMR structures of the latent and active forms (Ayed et al 2001). Recently, a new role for CTD was identified in DNA binding and trans-activation (Kim et al 2012). The acetylation of K320 and K382 is required for the regulation of the balance between p53 binding to random sequences of DNA or to specific binding p53 RE sequences. In addition, the acetylation of these two K is also required for the

positive regulation of p53 transcription. Thus, the role of the CTD of p53 is essentially repressive but the manner in which this domain causes inhibition of p53 functions remains unclear.

Role in p53 regulation

The CTD of p53 is needed for most p53 biological activities and plays an important role in p53 DNA binding and transcriptional activity (Hamard et al 2012). Hamard and collaborators suggested that the strong interaction of p53 with its targets depends on (1) the sequence of individual p53 RE and their association with p53, (2) post-translational modifications affecting the p53 protein, and more specifically the CTD, (3) p53 protein binding partners and (4) the epigenetic landscape of p53 target genes, regulating the access of p53 to relevant gene regulatory sequences (Hamard et al 2012).

The CTD contains several post-translational modification sites. Lysine aas K370, K372, K373, K381 and K382 can be each acetylated by p300/CREB-binding protein (CBP) (Knights et al 2006). The CTD seems to have a regulatory role on p53 functions by modifying the interaction on the p53 DBD with its target genes or the stability of the p53 tetramer. For example, the CTD is essential for pro-apoptotic p53 target genes such as *PUMA*, *Bcl-2-associated X protein (BAX)* and cell cycle arrest p53 target genes such as *cyclin-dependent kinase inhibitor 1* (*$p21^{WAF1}$*) (Hamard et al 2012). For Mdm2's interaction with p53's CTD, it was observed that both the CTD and TADI were involved and that Mdm2 binding to these domains was not mutually exclusively presupposing that different surfaces of the Mdm2 protein were involved (Poyurovsky et al 2010). The CTD could also interact with the signalling protein 14-3-3σ to stabilise p53 as a tetramer (Schumacher et al 2010).

Liang and collaborators have reported that K305 and K306 were essential for p53 nuclear import (Liang et al 1998, Liang and Clarke 1999a, Liang and Clarke 1999b). Using a mutagenesis approach, they observed that mutated K305A or K306A p53 proteins were deficient for nuclear import of p53. Other punctual mutations between these lysines and NLS-I did not have the same effect. However, deletion of two or more aas abolished p53 nuclear import, suggesting that a part of the basic domain is essential for NLS-I function.

C. Regulation of p53

The p53 expression and activity are regulated by several pathways (**Figure 5**). The first is the modulation of the level of expression in the cell via its interaction with Mdm2 or other proteins interaction such as Mdm4. The second is the regulation of p53's accumulation in the nucleus to promote the transcription of its target genes and the last level is by modulating the level of post-translational modifications (phosphorylation, acetylation), which influences p53's activities. Transcriptional regulation of the p53 promoter does not appear to have a critical role in the induction of p53 in response to many different signals.

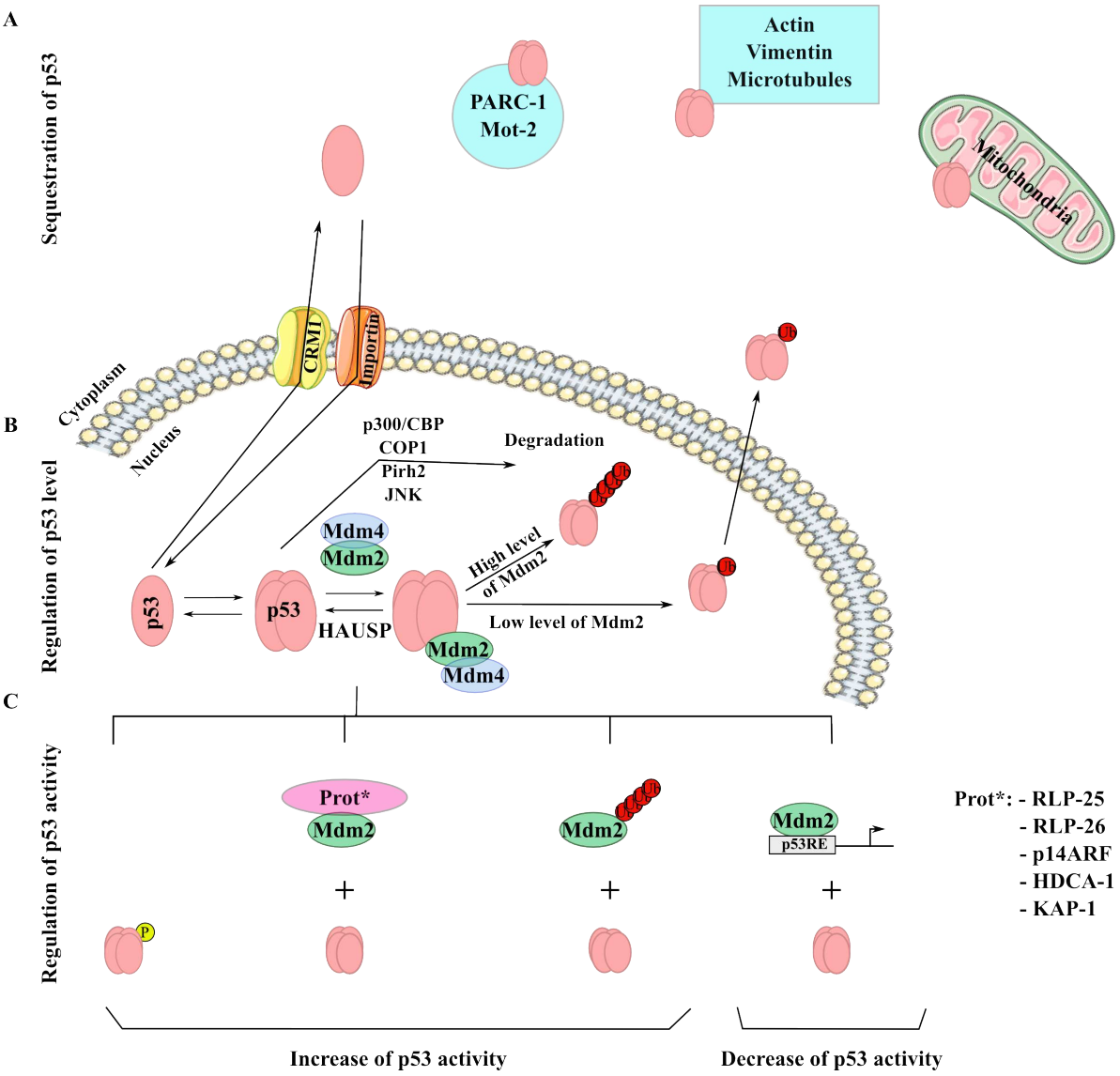


Figure 5: Regulation of p53 stability. (**A**) Sequestration of p53 in the cytoplasm by different mechanisms: PARC-1 or Mot-1 protein-protein interaction, sequestration by actin or vimentin, microtubules or sequestration in the mitochondria. (**B**) Regulation of p53 level. Mdm2 and Mdm4 form a complex, which allows p53 ubiquitination. HAUSP could regulate the formation of this complex by de-ubiquitinating p53 (**C**). Regulation of p53 activity. p53 could be phosphorylated, the absence of Mdm4 decrease p53 degradation mediated by Mdm2, ARF could form a complex with Mdm2 or Mdm2 could poly-ubiquitinate itself and induce its degradation. Adapted from (Haupt et al 1997, Kubbutat et al 1997); Thesis of Virginie Marcel, University Lyon I, 2009.

I. p53 stability

1. The p53/Mdm2 regulatory feedback loop

Mdm2: the main regulator of p53 stability

Mdm2 is a member of the really interesting new gene (RING)-finger-family protein and and an E3 ligase. p53 is one of its substrates and this modification mediates p53's targeted degradation by the proteasome (**Figure 5**) (Jackson and Berberich 2000). Mdm2 can interact with the TADI of p53. Toledo and Wahls have showed that this interaction is

disrupted by phosphorylation of the N-terminal domain of p53 at S9, S15, S20 and S46 and that one or more of these phosphorylation events appeared sufficient to disrupt Mdm2 binding (Toledo and Wahl 2006). However, using mice models with an S to R mutation at S15 or S20, no effect on Mdm2 binding p53 function was noted (Chao et al 2006, Toledo and Wahl 2006). No role of S46 alone was observed on p53 stability either. Thus, threonine (T) 18 phosphorylation seems to be the essential post-translational modification for p53/Mdm2 complex formation (Sakaguchi et al 2000, Schon et al 2002). After p53/Mdm2 complex formation, Mdm2 poly-ubiquitinates p53 in the CTD, providing a signal for p53's export to the cytoplasm and degradation by the 26S proteasome (Kubbutat et al 1999).

After DNA damage, S15 of p53 is phosphorylated, which facilitates the subsequent phosphorylation of S20, S46 and T18 resulting in the disruption of Mdm2 binding to p53 and of the degradation of p53 (**Figure 5C**). The stabilised p53 acts as a transcription factor for its target genes, including Mdm2 resulting in increased Mdm2 levels. This transcriptional regulation by p53 defines a negative regulatory feedback loop between p53/Mdm2 and is important for the cell to be able to restore p53 to its basal level (**Figure 5B**) (Chehab et al 1999, Hirao et al 2000, Stommel and Wahl 2004). The mechanisms, by which p53 escapes Mdm2 binding, depends on the type of stress but seems to be critical in most cell types. In the mouse, depletion of Mdm2 was found to be lethal at early embryonic stages whereas the double depletion of both Mdm2 and trp53 rescued a normal development (Jones et al 1995, Montes de Oca Luna et al 1995). This observation suggests an important role of Mdm2 in the control of p53 levels and activity during development. Indeed the early lethality observed in Mdm2 depleted-mice may be explained by uncontrolled levels of p53 activity, inducing growth arrest and apoptosis. Moreover, over-expression of Mmd2 works as an antagonist that blocks p53-mediated cell cycle arrest and apoptosis (Kruse and Gu 2009, Vousden and Prives 2009). These results show the importance of Mdm2-mediated p53 regulation. To date, several mechanisms modulating Mdm2/p53 interaction have been described (**Figure 5C**). For example, the $p14^{\text{alternate reading frame (ARF)}}$ protein (the alternative product of the *p16* locus) can sequester Mdm2 by binding to its central domain, preventing it from targeting p53 (Llanos et al 2001, Sherr 2006, Weber et al 1999); In addition, Mdm2 is able to self-ubiquitinate and to regulate its own degradation (Chang et al 1998, Linares et al 2003, Stommel and Wahl 2004, Stommel and Wahl 2005) and Mdm2 could also interact with ribosomal proteins (RLP), such as RLP5 and RLP23, which leads to the collapse of p53/Mdm2 complexes.

Mdm2 can also decrease p53 activity by indirect mechanisms. First, Mdm2 can poly-ubiquitinate the RPL26 and induce its degradation by the proteasome (**Figure 5C**) (Ofir-

Rosenfeld et al 2008). It is known that RPL26 binds p53 mRNA, stabilising the transcript and increasing its translation. Thus, the down-regulation of RLP26 by Mdm2 results in a lower level of p53 protein. Secondly, Chromatin Immuno-Precipitation assays (CHIP) have shown that Mdm2 can be found on the promoters of p53 target genes, such as $p21^{WAF1}$ (Arva et al 2005, Tang et al 2008). After DNA damage, p53 is activated and Mdm2 seems to be released from such promoters. Mdm2 could thus negatively regulate some p53 target genes. In addition, Mdm2 interacts with histone modifying enzymes such as histone deacetylase (HDAC) 1 and kinase associated phosphatase (KAP) 1 that regulate the deacetylation of histones at the promoters of target genes and thus could modulate the activation of such promoters (Ito et al 2002, Wang et al 2005).

Mdm4: Regulator of the p53/Mdm2 regulator feedback loop

Mdm4 (also called MdmX) is a RING-finger homologue of Mdm2 and also has an E3 ligase activity. It presents an almost identical p53 binding domain as Mdm2 in its N-terminal region and a RING domain in its C-terminal end (Wang and Jiang 2012). It was first identified as a p53-binding domain partner in 1996 (Shvarts et al 1996). Genetic studies showed that Mdm4 is as essential as Mdm2 for the degradation and the negative regulation of p53. Indeed, *Mdm4* depletion in mice is lethal at an embryonic stage due to the over-expression of p53 (Finch et al 2002, Parant et al 2001). Furthermore, embryonic lethality is also observed in mice carrying mutations in the Mdm4 RING domain due to an accumulation of p53 and an over-expression of p53 target genes such as $p21^{WAF1}$, *BAX* and *MDM2* (Huang et al 2011). Interestingly, when the level of p53 is reduced to a basal level, mice do not show embryonic lethality (Pant et al 2011). These results led to the suggestion that (1) the Mdm4 RING domain interacts with the Mdm2 RING domain and promotes Mdm2-mediated p53 poly-ubiquitination and degradation (Gu et al 2002, Kawai et al 2007, Okamoto et al 2009); (2) the level of basal p53 is important for Mdm2/Mdm4 degradation. Recently, *in vitro*, two main roles were described for the heterodimer of Mdm4/Mdm2: in the presence of p53, Mdm4 is an activator of the E3 ligase activity of Mdm2. In the absence of p53, Mdm4 protects Mdm2 against its self-ubiquitination (Wang et al 2011).

After DNA damage, activation of p53 requires the un-coupling of the Mdm2/Mdm4/p53 complex. Studies show that the kinases ATM/ATM and rad3-related protein (ATR) and Abelson murine leukemia viral oncogene (c-Abl) promotes the translocation of Mdm4 from the nucleus into the cytoplasm via the 14-3-3σ protein (Chen et al 2005, Jin et al 2006, LeBron et al 2006, Li et al 2002a, Okamoto et al 2005, Wang et al

2007, Waning et al 2011). The localisation of Mdm4 to the cytoplasm increases its degradation by the proteasome in an Mdm2-dependent manner (Chen et al 2005, Jin et al 2006, Okamoto et al 2005). This degradation appears to be the key for the inactivation of the Mdm2/Mdm4 E3 ligase activity during the p53 DNA damage response.

2. Others regulators of p53 stability

Other p53 stability regulators have been described in the literature such as the JNK kinase protein, other p53 E3 ligases, the de-ubiquitinating protein Herpes associated ubiquitin-specific protease (HAUSP) ubiquitin-specific protein and the p300/ CREB binding protein (CBP) co-activator proteins.

JNK

JNK appears to have a dual role on p53 stability, either increasing or decreasing p53 stability depending on cellular conditions. Under basal conditions, JNK binds p53 and acts as E3 ligase to target it for ubiquitination by the proteasome (Fuchs et al 1998b). After stress, JNK phosphorylates p53 on T81, inducing a cascade of modifications by acetylation to increase p53 activity (Fuchs et al 1998a, Prives and Hall 1999). As for Mdm2, the phosphorylation of the TAD mediated by JNK decreases p53 degradation and increases its stability.

Others p53 E3 ligases

Pirh2 is a RING domain protein, which is able to modulate p53 stability (Leng et al 2003). Like Pirh2, the human homolog of constitutively photomorphogenic 1 (COP1) has also been described as a p53 interacting RING finger protein (Dornan et al 2004a, Dornan et al 2004b). These two proteins act through a similar mechanism to Mdm2, ubiquitinating and degrading p53 through the proteasome. Interestingly, both are encoded by p53-inducible genes. These results suggest that Pirh2 and COP1 also participate in a p53 negative feedback loop, as for Mdm2. The functional impact of these proteins is similar to Mdm2: over-expression of Pirh2 reduces p53-dependent cell cycle arrest and apoptosis in cancer cell lines (Leng et al 2003), whereas silencing of COP1 by RNA interference increases p53 levels and p53 activation after DNA damage (Dornan et al 2004b); Pirh2 interacts with p53 tetramers though the p53 DBD and/or OD (Leng et al 2003, Sheng et al 2008).

The redundancy of these proteins and the negative regulatory feedback loop could be explained by (1) their expression at different times, for example, after DNA damage, Mdm2 is expressed and not Pirh2 (Duan et al 2006); (2) their synergism: loss of COP1 and Mdm2

expression increases 8-fold p53's half-life compared to the loss of Mdm2 alone (Dornan et al 2004b).

HAUSP

The HAUSP, also known as Ubiquitin-specific processing protease (USP) 7 modulates p53 activity regulation by interacting with both p53 and Mdm2 (Li et al 2002b, Lim et al 2004, Wood 2002). HAUSP directly interacts with p53, de-ubiquitinating it and activating p53 (Li et al 2002b). Studies showed that loss of HAUSP expression increases p53 stabilisation and cell cycle arrest (Cummins et al 2004, Li et al 2004).

p300/CBP: E4 ubiquitin ligases

CBP and p300 are transcription co-activators involved in p53 acetylation, which can also acts as E3/E4 ubiquitin ligases by mediating p53 degradation (Shi et al 2009a). Analysis of p300 or CBP-deficient cells revealed that both were required for endogenous p53 ubiquitination and the degradation of p53 in unstressed cells. Unexpectedly, p300/CBP ubiquitin ligase activities were absent in nuclear extracts and exclusively cytoplasmic. Consistent with the cytoplasmic localization of its E3/E4 activity, CBP deficiency specifically stabilized cytoplasmic, but not nuclear p53. This cytoplasmic localisation is associated with contrary functions of p300/CBP: cytoplasmic p300/CBP ubiquitinates and destabilizes p53, while nuclear p300/CBP activates p53 by acetylation.

$p14^{ARF}$

The $p14^{ARF}$ protein is the product of the alternative reading frame of the p16 locus (Kamijo et al 1997, Sherr 2001). It is able to bind p53, Mdm2 or both when present in a complex (Kamijo et al 1998). Through Mdm2 binding, $p14^{ARF}$ inhibits Mdm2 activity (Honda and Yasuda 1999) and increases Mdm2-mediated degradation of Mdm4 (Pan and Chen 2003). The $p14^{ARF}$ protein is localised in the nucleolus, which co-localises with Mdm2 (Weber et al 1999). These results suggest that $p14^{ARF}$ could sequester Mdm2 in the nucleoplasm and indirectly activate p53. Since $p14^{ARF}$ is a transcriptional target of E2F transcription factors during the S (synthesis) phase of the cell cycle, its expression is increased in cells, which results in the activation of p53 as a safeguard mechanism to control cell growth (Lomazzi et al 2002).

II. Subcellular localisation

p53 is a transcriptional factor and exert its main activity in the nucleus. Export of p53 into the cytoplasm inactivates this function. During cell cycle and after genotoxic stress, p53 appears to shuttle between the nucleus and the cytoplasm (David-Pfeuty et al 1996, Moll et al 1996, Ostermeyer et al 1996, Shaulsky et al 1991a, Shaulsky et al 1991b). NLS and NES sequences are essential for the balance between the rates of import and export of p53 through the cell (Henderson and Eleftheriou 2000, Shaulsky et al 1990a, Shaulsky et al 1990b).

p53 contains two NES, one in the oligomerisation domain (aas 339-352) and the second in the N-terminal domain (aas 11-27) (Stommel et al 1999, Zhang and Xiong 2001a, Zhang and Xiong 2001b). Under basal condition, the NES are exposed when p53 presents its inactive tri-dimensional structure and p53 is exported into the cytoplasm through the export nuclear receptor CRM1 (Kudo et al 1998, Santiago et al 2013). In the cytoplasm, p53 is degraded by the proteasome in a Mdm2-dependent manner (O'Keefe et al 2003). Post-translational modifications such as phosphorylation of S15 and S392, which allows the tetramerisation of p53, interfere with nuclear export (Sakaguchi et al 1997, Zhang and Xiong 2001b). When p53 tetramers are formed, the NES in the C-terminal is not accessible and the p53 tetramers accumulate in the nucleus (Lohrum et al 2001, Stommel et al 1999).

The activity of the NLS is also under post-translational control. As explained in the Introduction, Section B, Part 5, p53 contains several NLS (NLS I: aas 315-325; NLS II: aas 369-375 and NLS III: aas 379-384), the strongest NLS is the NLS I (Shaulsky et al 1990b). This NLS includes a phosphorylation site for cyclin-dependent kinase (Cdk) 2/Cyclin A at S315 and an acetylation site for p300/CBP at K320 (Iyer et al 2004).

p53 crosses the nucleus membrane by fixation to importin α (Kim et al 2000, Liang and Clarke 1999a). Truncation of importin α results in the sequestration of p53 in the cytoplasm. Some additional proteins are implicated in p53 importation into the nucleus such as S100B and Spot-1, which bind to the NLS I motif (Elkind et al 1995, Scotto et al 1999) or Bcl-2, which blocks its re-localisation (Beham et al 1997).

Several proteins can also interact with p53 in the cytoplasm and enable its sequestration including PARC-1 and Mot-2 (Nikolaev et al 2003, Wadhwa et al 1998) or it migration into the cytoplasm, such as actin, vimentin or microtubules (Giannakakou et al 2002, Katsumoto et al 1995, Klotzsche et al 1998).

A fraction of p53 is also found in the mitochondria (Marchenko et al 2000). The mitochondrial p53 can induce the mitochondrial intrinsic apoptotic pathway, which will be developed later in this Thesis (Introduction, Section D, part II-1).

Nuclear import/export and proteasome dependent degradation of p53 are closely related mechanisms. Controversial results have been published on p53 re-localisation into the cytoplasm and its degradation by Mdm2. Some studies showed a degradation of p53 in the cytoplasm only (Freedman and Levine 1998, Roth et al 1998), while others reported that p53 was degraded in the nucleus (Lohrum et al 2001, Stommel and Wahl 2004). Recent studies suggest that the re-localisation of p53 mediated by Mdm2 depends on its ubiquitination status. When Mdm2 levels are low, p53 is only mono-ubiquitinated and is exported to the cytoplasm; while at high level of Mdm2 or in presence of p300 and Mdm2, p53 is poly-ubiquitinated and degraded in the nucleus (Grossman et al 2003, Li et al 2003, Tao and Levine 1999).

III. Epigenetic silencing of *TP53* promoter

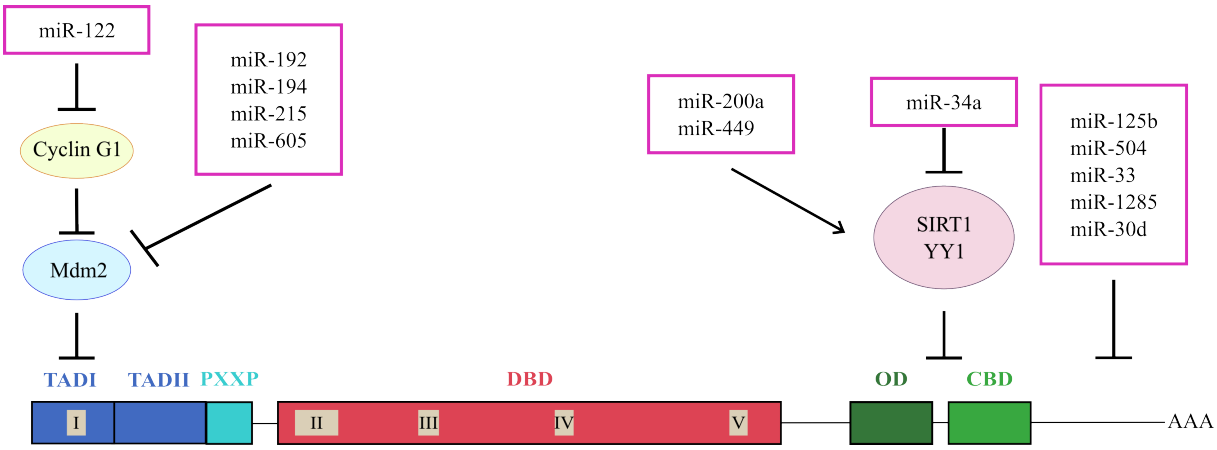
In 1997, Schroeder and collaborators reported a correlation between hyper-methylation of the *TP53* promoter and the decrease of its transcription (Schroeder and Mass 1997). *In vitro* studies, using reporter gene constructs, demonstrated that *TP53* promoter methylation induces a 90% reduction of p53 mRNA expression (Pogribny et al 2000). Since, no reports have been published on naturally occurring DNA methylation of the *TP53* promoter. A number of publications have reported that DNA hyper-methylation of the *TP53* promoter was associated with low levels of mRNA expression in primary hepatocellular carcinoma (Pogribny and James 2002), acute lymphoblastic leukaemia (Agirre et al 2003), chronic lymphocytic leukaemia, breast cancer without *TP53* mutation (Pharoah et al 1999) or glioma (Amatya et al 2005). Based on these results, DNA methylation could be a mechanism used to inactive p53 expression but the role of methylation remains to be fully established. However, it should be noted that there are little data on the levels of DNA methylation of the *TP53* promoter and its variation in non-cancer cells.

IV. Post-transcriptional modifications

1. miRNA

MicroRNAs (miRNAs) are non-coding RNAs of about 22 nucleotides that bind a complementary sequence in the target mRNA and can influence its post-translational regulation, usually by impairing translation or inducing target degradation. About 1,870 human miRNA sequences are currently listed in the miRBase compiled at the University of

Manchester (http://www.mirbase.org/). They are distributed on all chromosomes and miRNA sequences represent about 3-5% of all predicted genes in the human genome. Given the capacity of the same miRNA to target different mRNAs, it is considered that over 50% of all protein coding genes in the genome are regulated by miRNA. Several miRNAs could regulate p53 activity by direct modulation of p53 mRNA and indirect targeting of mRNA encoding p53 regulators. In addition, p53 controls the expression of a network of miRNA with essential functions in cell proliferation, differentiation and survival, such as miR-34, miR-145 or members of the miR-200 family (**Figure 6**) (Hermeking 2012).



(Figure 6: p53 regulation by miRNAs. Certain microRNAs (miR) mediate p53 down-regulation via interaction with the 3'UTR. Others can inhibit the Mdm2-p53 interaction directly or indirectly, such as miR-122. They could also inhibit SIRT1 and YY1 regulation of p53 mRNA in the C-terminal part of the mRNA. TAD: Transcriptional Domain; PXXP: Proline Domain; DBD: DNA Binding Domain; OD: Oligomerisation Domain; CDB: C-terminal Basic Domain. Adapted from (Hermeking 2012).

miRNA-mediated direct p53 inhibition

Several publications have shown that p53 expression is regulated by the interaction between miRNAs and the 3'untranslated region (UTR) of p53 pre-mRNA (**Figure 6**). These miRNA include miR-125b (Le et al 2009), which contains a complementary sequence to the 3'UTR of p53 mRNA. Targeting of p53 mRNA by miR-125b results in decreased p53-mediated apoptosis. Recently, other miRNAs of the miR-125 family were shown to interact with components of the p53 effector network such as *PUMA* (Le et al 2011). Furthermore, elevated expression of miR-125b is associated with an increase of tumour size, invasion, poor prognosis and decreased survival in colorectal cancer (Nishida et al 2011). Taken together, these results suggest that miR-125b an important negative regulator of p53 and of its biological effects. .

Other miRNAs interacting with the 3'UTR of p53 mRNA include miR-504, miR-33, miR-1285 and miR-30d (Herrera-Merchan et al 2010, Hu et al 2010a, Kumar et al 2011, Tian et al 2010). These miRNAs appear to decrease p53 expression and thus to reduce p53's

biological effects such as apoptosis, cell cycle arrest or senescence. Amplification of the gene encoding miR-30d is found in about 30% of solid tumours and is associated with poor clinical survival in ovarian cancer (Li et al 2012a).

Indirect p53 inhibition by miRNAs

Some studies also observed that miRNAs could down-regulate upstream negative regulators of p53 (**Figure 6**). The most studied miRNAs are the miR-34 family. miR-34a can down-regulate the NAD-dependent deacetylase SIRT1, a regulator of metabolic stress responses acting as a tumour suppressor. This down-regulation increases p53 activity and p53 the expression of p53 target genes such as $p21^{WAF1}$ and *PUMA*, which subsequently increases apoptosis (Yamakuchi et al 2008). miR-34a expression is regulated by the p53 protein, creating a positive regulatory loop between p53, miR-34a and SIRT1. miR-200a or miR-449 can increase SIRT1 expression and thus influence this p53 negative loop (Bou Kheir et al 2011, Eades et al 2011). p53 is implicated in a similar feedback loop involving miR-34, since this miRNA can inhibit YY1, a transcriptional suppressor protein, which itself inhibits p53 (Chen et al 2011).

Mdm2 is regulated by several miRNAs, which could thus indirectly modulate p53 expression and activity. miR-122 can down-regulate Cyclin G1, which inhibits the recruitment of Mdm2 by p53, resulting in an increase of p53 levels and activity (Fornari et al 2009). Mdm2 expression can be also modulated by miR-192, miR-194, miR-215 and miR-605, which are all induced by p53 (Braun et al 2008, Pichiorri et al 2010, Xiao et al 2011). These interactions suggest a positive feedback loop involving these miRNA, Mdm2 and p53. Of note, miR-192 appears to be down-regulated in colorectal cancer (Braun et al 2008).

These results show that a complex network of miRNAs can directly or indirectly regulate p53 mRNA levels with the indirect routes creating feedback loops and redundancies and highlight the complexity of p53 regulation.

2. WRAP53

The *WRAP53* gene is located upstream of *TP53*, in the opposite orientation and its coding sequence overlaps with the *TP53* promoter, non-coding exon 1 and the proximal part of intron 1 (**Figure 2A**) (Mahmoudi et al 2009), Initially identified as a natural antisense transcript that up-regulates p53 expression (Yuan et al 2011). Down-regulation of the *WRAP53* gene results in a significant decrease in p53 mRNA and in suppression of p53 induction in response DNA damage. This effect has been attributed to the capacity of WRAP53 mRNA to hybridize with and stabilize p53 mRNA. To analyse the interaction

between WRAP53 transcripts and p53, Mahmoudi and collaborators developed a siRNA against WRAP53 (Mahmoudi et al 2009). They observed, in the absence of WRAP53, a decrease of p53 mRNA level. They also observed a RNA duplex interaction in a head-head manner, which allows the p53's transcription. However, knockdown or over-expression of p53 did not influence WRAP53 expression. These results suggest that WRAP53 is a positive regulator of p53 transcription. In addition, WRAP53 is also known as telomerase Cajal body protein-1 (TCAB1), encoding a protein that interacts with components of active telomerase dyskerin, telomerase reverse transcriptase (TERT) and telomerase RNA component (TERC) and with small Cajal body RNAs, which are involved in splicing regulation (Mahmoudi et al 2010). Depletion of TCAB1 by RNA interference prevents TERC from associating with Cajal bodies, disrupts telomerase-telomere association and prevents telomere elongation. Mutations in *TCAB1* have been associated with Dyskeratosis Congenita (Venteicher and Artandi 2009).

3. G-quadruplexes in 3'UTR

G-quadruplexes (G4s) consist of four-stranded structures occurring in guanine-rich sequences with regulatory effects in DNA or RNA. These structures are widespread in the genome and have been shown to have important regulatory effects on gene transcription (Bochman et al 2012), mRNA splicing (Gomez et al 2004) and mRNA stability (Millevoi et al 2012). G4 structures are presented and discussed in detail in the Introduction, Part III of this Thesis. Decorsière and collaborators observed that after DNA damage generated by UV, the cleavage of p53 pre-mRNA and the addition of the poly(A) tail is deregulated (Decorsiere et al 2011). They showed that this cleavage of the pre-mRNA is regulated by a G4 structure located downstream of the cleavage site. Heterogenenous nuclear ribonucleoproteins (hnRNP) H and F can bind this G4 and stabilize it, enabling p53 expression and p53-induced apoptosis.

V. Post-translational modifications

Under basal conditions, p53 has a short half-life (5 to 30 min) and is maintained at low levels in cells (Jenkins et al 2012). However, under stress condition, p53 is stabilized, activated and accumulates in the nucleus. Stabilisation and activation of p53 are mediated and accompanied by post-translational modifications such as phosphorylation, acetylation, methylation, ubiquitination, neddylation, sumoylation, poly ADP-ribosylation and O-glycosylation. Each type of post-translation modification is clustered in distinct regions of the protein. There are multiple cross talks and redundancies between these post-translational modifications. The precise pattern of modifications may depend upon the particular pathway

of p53 induction such as different forms of stress, the amplitude of inducing signals and of cell responses, and the particular cell and tissue context.

1. Phosphorylation

Phosphorylation is located essentially in the TAD and C-terminal basic domains (Jenkins et al 2012). Overall, phosphorylation of the TAD is critical for p53 stabilisation and activation (Ohki et al 2007) whereas phosphorylation in the CBD regulates p53 oligomerisation (**Figure 7**).

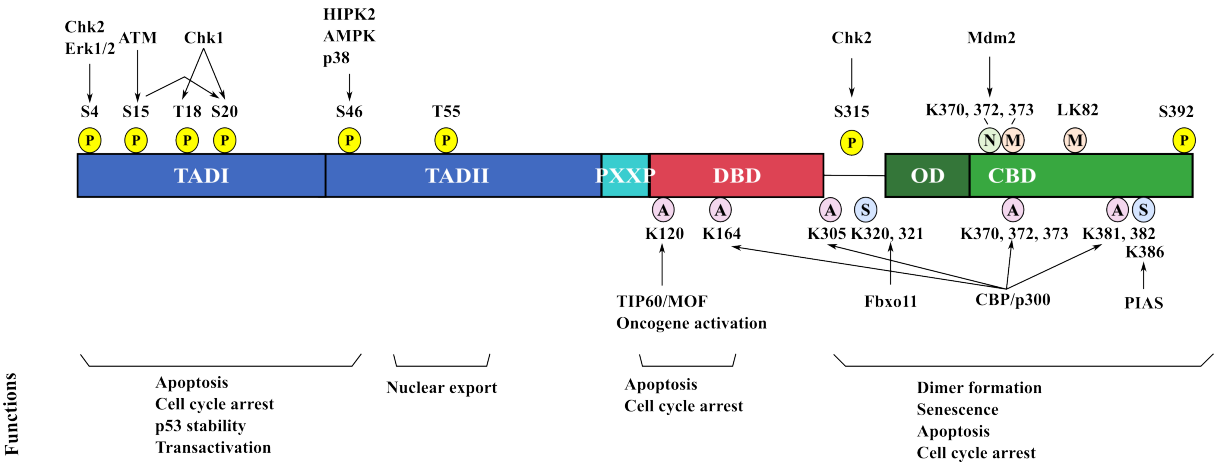


Figure 7: Post-translational modification of p53. Known modifying enzymes (black arrow) are shown above the modifications. Circle: Phosphorylation (yellow); Acetylation (purple); Methylation (orange); Sumoylmation (blue); Neddylation (green). Adapted from (Appella and Anderson 2001, Jenkins et al 2012))

N-terminal trans-activation domain

The p53 TAD phosphorylation network is complex, with some phosphorylations increasing p53 activity while others may decrease it. Recently, Jenkins and collaborators observed that phosphorylation sites in TADI were more important than those in TADII for activating the p53 protein (Jenkins et al 2012). The role of some phosphorylation sites is summarized in **Table 1**.

After DNA damage generating DNA double strand breaks (DSB), one of the first events is the auto-phosphorylation of the ATM kinase and the subsequent phosphorylation of S15 of p53 (Canman et al 1998). This phosphorylation results in the binding to p300/CBP and enables p53 to increase p53-dependent transcription (Dumaz and Meek 1999, Lambert et al 1998). The phosphorylation of S15 also enables the phosphorylation of S20 mediated by Chk1, which increases p53 stabilisation (Shieh et al 2000). These phospho-S are implicated in the p53-dependent G1 cell cycle arrest and apoptosis induction (**Table 1**) (Jenkins et al 2012). S4 is also phosphorylated by Chk2 after DNA damage and is associated with an increase in p53-dependent apoptosis (Shieh et al 2000). In addition, the phosphorylation of S46 by the

kinases HIPK2 (D'Orazi et al 2002), AMPK (Okoshi et al 2008) or p38 (Perfettini et al 2005) has been described as an early step for promoting p53-dependent apoptosis.

In contrast, p53 phosphorylations can also inhibit p53-protein interactions. The presence of the phosphorylated form of T18, carried out by Chk1, reduced by 5- to 25- fold the binding affinity of the TADI-Mdm2 complex (Ferreon et al 2009, Teufel et al 2009). Phosphorylation at T55 also mediates p53 nuclear export (Cai and Liu 2008).

Under conditions of non-genotoxic stress, p53 can also be phosphorylated, for instance the phosphorylation of S15 during hypoxia to stabilize and activate p53 (Koumenis et al 2001) or phosphorylation of S15 via Erk1/2 under conditions of microtubule disorganisation (Sablina et al 2001, Stewart et al 2001). The role of p53 phosphorylation during these stresses is still not well understood.

Table 1: Summary of phenotypic observation for Knock-in mutated mice carrying mutants at different important phosphorylation site of p53 (Adapted from (Jenkins et al 2012)).

Genotype (Mouse)	Correspondence (human)	p53 stability	p53 activity
$p53^{-/-}$	$p53^{-/-}$	NA	None
$p53^{S18A/S18A}$	$p53^{S15A/S15A}$	Normal	Reduced apoptosis and G1 arrest
$p53^{S23A/S23A}$	$p53^{S20A/S20A}$	Modest effect	Reduced apoptosis
$p53^{S18A,S23A/S18A/S23A}$	$p53^{S15A,S20A/S15A,S20A}$	Normal	No apoptosis, impaired G1 arrest
$p53^{L25Q,W26S/L25Q,W26S}$	$p53^{L25Q,W26S/L25Q,W26S}$	Increased	No apoptosis, no G1 arrest
$p53^{T21D,S23D/-}$	$p53^{T18D,S20D/-}$		Increased in basal, no increase after stress
$p53^{F53Q,F54S/F53Q,F54S}$	$p53^{F53Q,F54S/F53Q,F54S}$	Normal	Normal
$p53^{S46A/S46A}$ *	$p53^{S46A/S46A}$	Slight reduction	Reduced apoptosis

NA: not applicable
*: in human p53 transgenic background

C-terminal basic domain

S315 and S392 were the first p53 phosphorylated sites identified in the C-terminal basic domain (Appella and Anderson 2000). After exposure of UV radiation, phosphorylation on S315 mediated by Chk2, has only alittle effect on p53 dimer formation but in combination with phosphorylation at S392, Waterman and collaborators observed an increase in the oligomerisation induced by S392, influencing p53 function (Waterman et al 1995). Phosphorylation at S392 results in the stabilisation of active p53 in its tetrameric state via an intermolecular mechanism: one segment of the C-terminal basic domain bind to the DBD of another subunit in the tetramer (Retzlaff et al 2013).

2. Ubiquitination

As discussed above in detail, ubiquitination of p53 is mediated by two major E3 ligases, Mdm2 and Mdm4. p53 is poly-ubiquitinated in the presence of high levels of Mdm2

and mono-ubiquitinated when Mdm2 is at low levels of Mdm2 (Dai and Gu 2010). Mdm4 specifically contributes to poly-ubiquitination. When p53 is poly-ubiquitinated, it is degraded by the proteasome while mono-ubiquitination of p53 enables the export of p53 into the cytoplasm and induces apoptosis or autophagy (Becker et al 2007, Tasdemir et al 2008a).

p53 can also be modified by the addition of two other ubiquitin-like proteins, Small ubiquitin-like Modifer (SUMO) and Neural precursor cell Expressed Developmentally Down-regulated protein 8 (NEDD8) (Dai and Gu 2010). These proteins are well conserved during evolution and could interact directly with some p53 lysines. SUMO and NEDD8 present the same three-dimensional structure as ubiquitin (Carter and Vousden 2008). The only identified site in p53 that can be sumoylated is K386 by a reaction carried out by the E3 SUMO ligase Protein Inhibitor of Activated Stat (PIAS) (Stehmeier and Muller 2009). The sumoylation of this aa appears to activate a p53 transcriptional program (Melchior and Hengst 2002) and to induce premature senescence (Stehmeier and Muller 2009). Neddylation of p53 is mediated by Mdm2 at lysines, K370, K372 and K373, which are also targeted by ubiquitination (Xirodimas et al 2004). Fbxo11, a F-box protein family that constitutes one of the four subunits of ubiquitin protein ligase complex SKP1-cullin-F-box (SCFs), can neddylate p53 at amino acids K320 and K321 (Abida et al 2007). The neddylation of p53 inhibits p53 transcriptional activity (Abida et al 2007, Xirodimas et al 2004). Recently NEDD8 Ultimate Buster 1 (NUB1), a non-covalent acting protein of NEDD8, was demonstrated to decrease p53 neddylation and increase ubiquitination (Liu and Xirodimas 2010).

3. Acetylation

Acetylation of p53 was shown to be correlated with stabilisation, activation in response to cellular stress and stimulation of sequence-specific DNA-binding (Brooks and Gu 2011). p300/CBP can acetylate different lysines at positions 164, 305, 370, 372, 373, 381 and 382 (Kruse and Gu 2008). These aas could also be ubiquitinated by Mdm2. K120 could also be acetylated TIP60/MOF, two acetyl transferases (Sykes et al 2006, Tang et al 2006). This acetylation of K120 could be activated by oncogene activation (Mellert et al 2007). Mutation of this aa induces activates cell cycle arrest but impairs p53-induced apoptosis (Tang et al 2006).

4. Other post-translational modifications

Methylation of lysines in the C-terminal basic domain also regulates p53 activity. For example, mono-methylation at K370 causes the inhibition of p53 activity (Huang et al 2006), di-methylation at K370 increases p53 activity by promoting interaction with its co-activator

53BP1 (Huang et al 2006) and methylated K372 increases transcription of $p21^{WAF1}$ (Chuikov et al 2004). Mutation of K373 to Q or acetylation of K373 promotes apoptosis by favouring interaction with lower-affinity binding sites like *BAX* and *p53AIP1* promoters (Knights et al 2006, Roy and Tenniswood 2007). The mono-methylated K382 impacts on the trans-activation of some p53 target genes such as $p21^{WAF1}$and *PUMA,* which are inhibited, whereas p53 RE's with weak affinity for p53 are unaffected (Shi et al 2007). Moreover, p53 modified at K382 induces cell cycle arrest but not apoptosis (Roy et al 2005).

p53 could be also modified by O-glycosylation in the DBD domain (Shaw et al 1996), ADP-ribosylation mediated by Poly-ADP Ribose Polymerase (PARP) at K135 (Wesierska-Gadek et al 1996). The biological significance of these modifications remains unclear.

VI. Regulation of p53 activity by its isoforms

The p53 protein must be in a tetrameric form to bind DNA with high affinity. On the other hand, *TP53* encodes at least 12 different protein isoforms with different N- and C-terminal domains, which are differentially regulated by Mdm2. Hetero-oligomerisation between full-length p53 and specific isoforms may therefore regulate p53 stability and activity. A recent study has shown that p53 and Δ40p53, when co-transfected in p53-null cells, could form hetero-oligomers (Hafsi et al 2013). Δ40p53 lacks the TADI, which mediates the standard transcription activation capacity of p53. When expressed in excess as compared to full-length p53, Δ40p53 operates as a negative regulator. However, when expressed at low levels, Δ40p53 may be incorporated in oligomers containing mostly full-length p53, increasing their stability and DNA binding activity by facilitating their escape from Mdm2-dependent degradation (Hafsi et al 2013). In addition, the Δ40p53 can fixe the promoter of some cell cycle p53 target genes such as $p21^{WAF1}$ and altered the p53 transactivation activity (Takahashi et al 2013). These results suggest that the Δ40p53 isoform may exert complex effects on p53 activity, depending upon its level of expression.

D. p53 functions

There has been debate on whether the p53 protein can exert significant biological activities when expressed at baseline, normal levels. Indeed, for a long time, it has been considered that p53 did not play any significant biological role unless activated in response to a wide range of physical, chemical or biological stresses. One of the common denominators of these stresses is the presence of diverse forms of DNA damage. In response to such types of stress, p53 is activated as a high-affinity, sequence specific DNA binding protein and accumulates in the nucleus. This activation is mediated by its dissociation from Mdm2 by post-translational modifications such as phosphorylation at several S or T in the N-terminus (including S15). The p53 protein is implicated in many cellular functions such as cell cycle arrest, apoptosis, DNA repair or senescence. The selection of which among the multiple effector pathways regulated by p53 is activated depends at least in part on its post-translational modifications, themselves reflecting the activation of signalling cascades in response to different forms of stress.

I. Cell cycle arrest by controlling checkpoints

p53 is implicated in cell cycle arrest by controlling the induction of the G1/S and G2/mitosis (M) cell cycle checkpoints. Arrest at these checkpoints may be permanent and lead cells to enter senescence. Alternatively, transient arrest may allow the cell time to repair DNA lesions before resuming proliferation.

1. G1/S transition

The G1/S cell cycle checkpoint is important to prevent cells damaged in G1 from entering the replication phase. This checkpoint is regulated in two different ways. The first is mediated by Cdc25A degradation and the second by p53 activation and the subsequent transcriptional activation of $p21^{WAF1}$. In this section, only the latter mechanism by p53 will be developed (**Figure 8**).

p53 plays a prominent role in the G1/S checkpoint (Kastan et al 1992, Lin et al 1992). After exposure to ionizing radiation (IR) or reactive oxygen species (ROS) formation, ATM and ATR kinases are activated and phosphorylate p53 at S15 (Kastan et al 1992, Siliciano et al 1997). ATM/ATR can also phosphorylate and activate Chk2, which phosphorylates p53 at S20. These phosphorylations of p53 lead to the disruption of its interaction with Mdm2 and

consequently the amount of p53 protein in the cell increases. Once activated and accumulated, p53 up-regulates the transcription of *p21*^{WAF1} (el-Deiry et al 1993, el-Deiry et al 1994, Kastan and Bartek 2004). p21^{WAF1} then binds and inhibits the complex Cyclin E/Cdk2 or Cyclin D/Cdk4 (Harper et al 1993, Stewart and Pietenpol 2001) which are critical for the sequential phosphorylation of the retinoblastoma protein (pRB) at the G1/S transition. Inhibition of pRB phosphorylation, in turn, prevents the release of E2F transcription factors, which directly mediates the expression of many genes essential for progression into S phase. The pRB/E2F protein complex is the main target of the G1/S checkpoint effectors (Sherr 1994). Mouse embryonic fibroblasts lacking *p21*^{WAF1} are deficient in their ability to arrest in G1 in response to DNA damage and nucleotide pool perturbation (Deng et al 1995). The reversibility of the p53-induced G1 arrest appears to be cell-type specific (Agarwal et al 1995, Di Leonardo et al 1994).

In response to an excessive proliferative signalling, such as oncogene activation, the expression of p14^{ARF}, an alternative reading frame product of the *p16* gene, is increased (Lomazzi et al 2002). p14^{ARF} sequesters Mdm2 and prevents its binding with p53, thus allowing the stabilisation of p53 and favouring cell cycle arrest through the activation of p21^{WAF1} and other downstream events.

Finally, p53 can regulate the G1/S transition by modulating miR expression. Indeed, p53 can trans-activate the promoter of the genes of the miR-34 family (He et al 2007) and miR-34 can inhibit Cdk2, Cdk4 and Cyclin E2 activity (Hermeking 2012). Another example is the down-regulation of Cdk4 by miR-145 that is a direct target of p53 (Hermeking 2012). In addition, miR-15a and miR-16-1 target Cyclin E and contribute to the p53-induced cell cycle arrest at the G1/S transition (Ofir et al 2011).

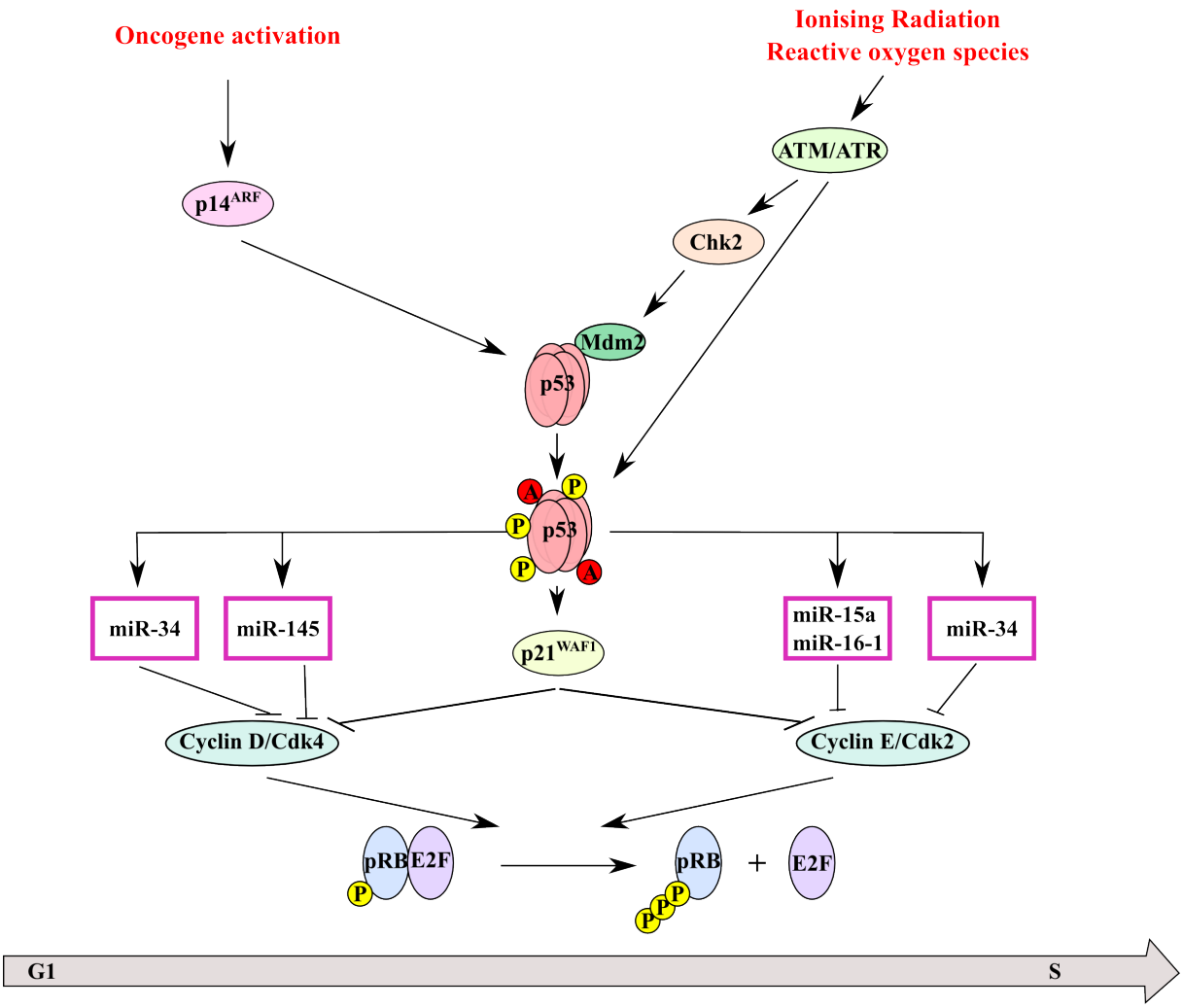


Figure 8: Role of p53 in the G1/S transition checkpoint. The p53 protein can be phosphorylated and activated by two different manners: after oncogene activation via $p14^{ARF}$ or ATM/ATR and Chk2 activation in response to ionising radiation of reactive oxygen species. p53 can thus transactivate genes implicated in G1/S transition such as $p21^{WAF}$ or miRNA (miR-34, miR-145, miR-15a and miR-16-1) which inhibit Cyclin D/Cdk4 or CyclinE/Cdk2 complex and indirectly pRB activation, a major controlling component of the G1/S transition.

2. G2/M transition

G2/M progression is driven by the maturation-promoting factor, a complex of Cyclin B1/Cdk1. The blocking of the activity of this complex leads to the G2 arrest and is induced by multiple signalling pathways, involving the kinases ATM, ATR and their downstream substrates Chk1 and Chk2 (Nyberg et al 2002) and p38 MAPK (**Figure 9**). In addition to the G1/S arrest, p53 has been shown to participate in the G2/M checkpoint through multiple functions, acting mainly in the maintenance and the recovery of this cell cycle arrest. p53 regulates the transition from G2 to M by activating the Cdk inhibitor p21, Growth arrest and DNA damage 45 (GADD45) and 14-3-3σ (Agarwal et al 1995, Guillouf et al 1995, Stewart et al 1995) but also through transcriptional repression of mitotic regulators, including Cyclin B1 (Innocente et al 1999), Cdc25B and PLK1 (Medema and Macurek 2012).

Over-expression of the GADD45 protein has been associated with the G2/M arrest in $p53^{+/+}$ cells but not in $p53^{-/-}$ cells, suggesting a link between p53 and GADD45 in cell cycle arrest. After UV exposure, p53 trans-activates the *GADD45* promoter, which inhibits Cyclin B1/Cdk1 complex formation (Zhan et al 1999). The kinase 14-3-3σ is also induced by p53 and can bind Cdc25b, the phosphatase that controls the de-phosphorylation of tyrosine 15 of Cdk1, located at the active site of the main Cdk controlling G2/M progression. Cdk1 is phosphorylated on tyrosine (Y) 15 during G2 and de-phosphorylation by Cdc25b is the activation signal for entry into mitosis (Hermeking et al 1997, Waterman et al 1998). Binding of 14-3-3σ leads to the sequestration of Cdc25C in the cytoplasm and inhibition of its phosphatase activity (Lopez-Girona et al 2001, Peng et al 1997) and thus its inability to activate the cyclin B1/Cdk1 complex (Peng et al 1997).

In addition, it has been shown that the Reprimo protein induces cell cycle arrest by inhibiting Cdk1 activity and nuclear translocation of the Cdk1/Cyclin B1 complex in a p53-dependent manner (Ohki et al 2000). It has also been observed that the repression of the topoisomerase (Topo) II gene by p53 also helps to block entry into mitosis and strengthens the G2 arrest (Wang et al 1997).

Phosphorylation of Cdk1 by the Cyclin-Activating Kinase (CAK) is one event in the multistep process required for activating CAK. Schneider and collaborators have shown that the activity of CAK is negatively regulated by p53 (Schneider et al 1998).

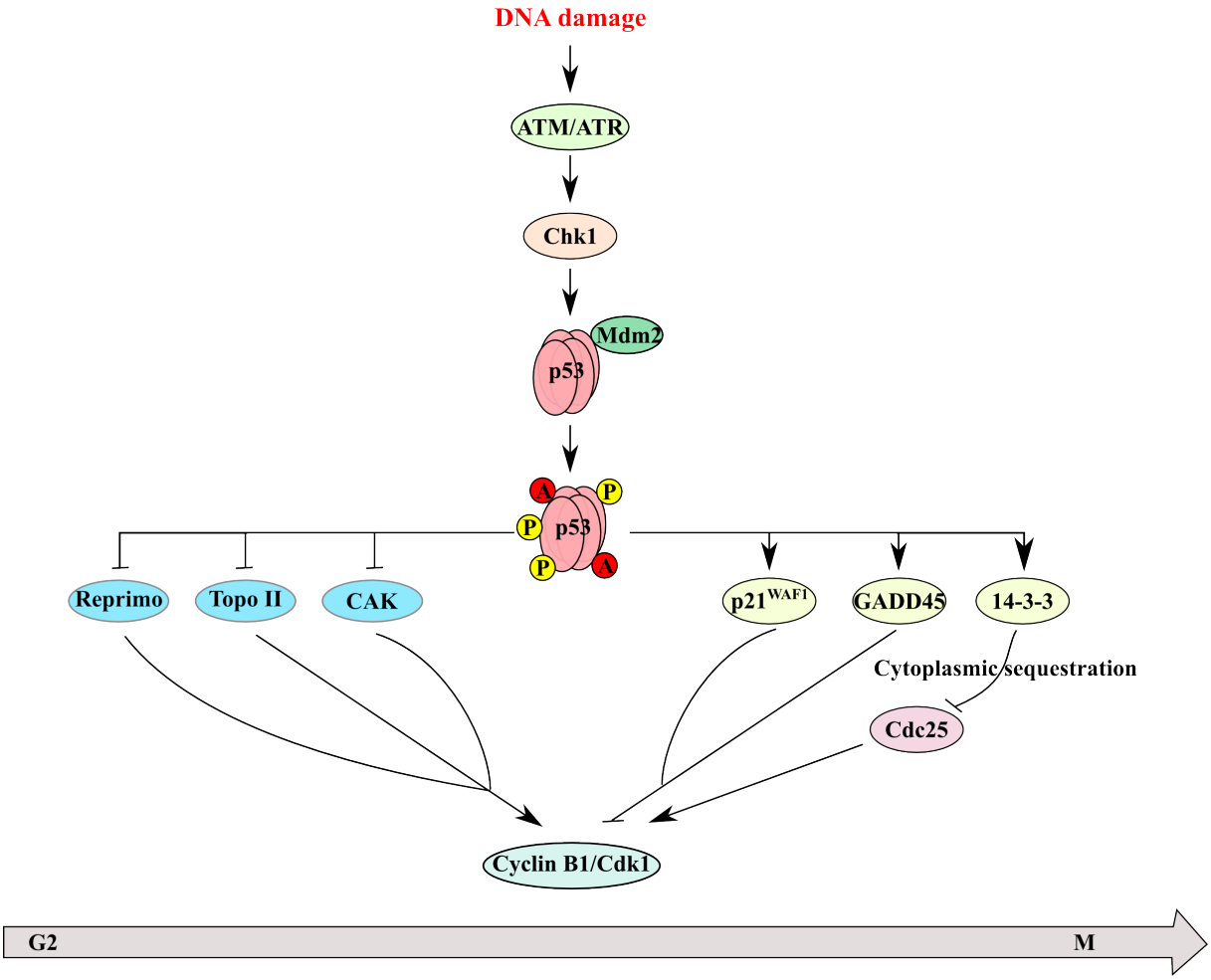


Figure 9: Role of p53 in G2/M transition checkpoint. The p53 protein can be phosphorylated and activated by ATM/ATR and Chk1 activation in response to UV. The phosphorylated p53 can trans-activate genes implicated in the G2/M transition such as *p21*^{WAF}, *GADD45* or *14-3-3σ*, which inhibit Cyclin B1/Cdk1 and G2/M transition. p53 can also inhibit some protein such as Reprimo, Topo II (Topoisomerase II) or CAK (Cyclin-Activating complex): Cdk7, Cyclin H and Mat-1, which activate Cyclin B1/Cdk1 and G2/M transition.

II. Apopotosis

Since the observation of Oren and colleagues that the re-introduction of p53 into p53-deficient myeloid leukaemia cells can induce apoptosis in a manner that could be countered by a pro-survival cytokine (Yonish-Rouach et al 1991), p53 has been shown to be involved in several death pathways such as the mitochondrial pathway, the death receptor pathway and the reactive oxygen species pathway (**Figure 10**).

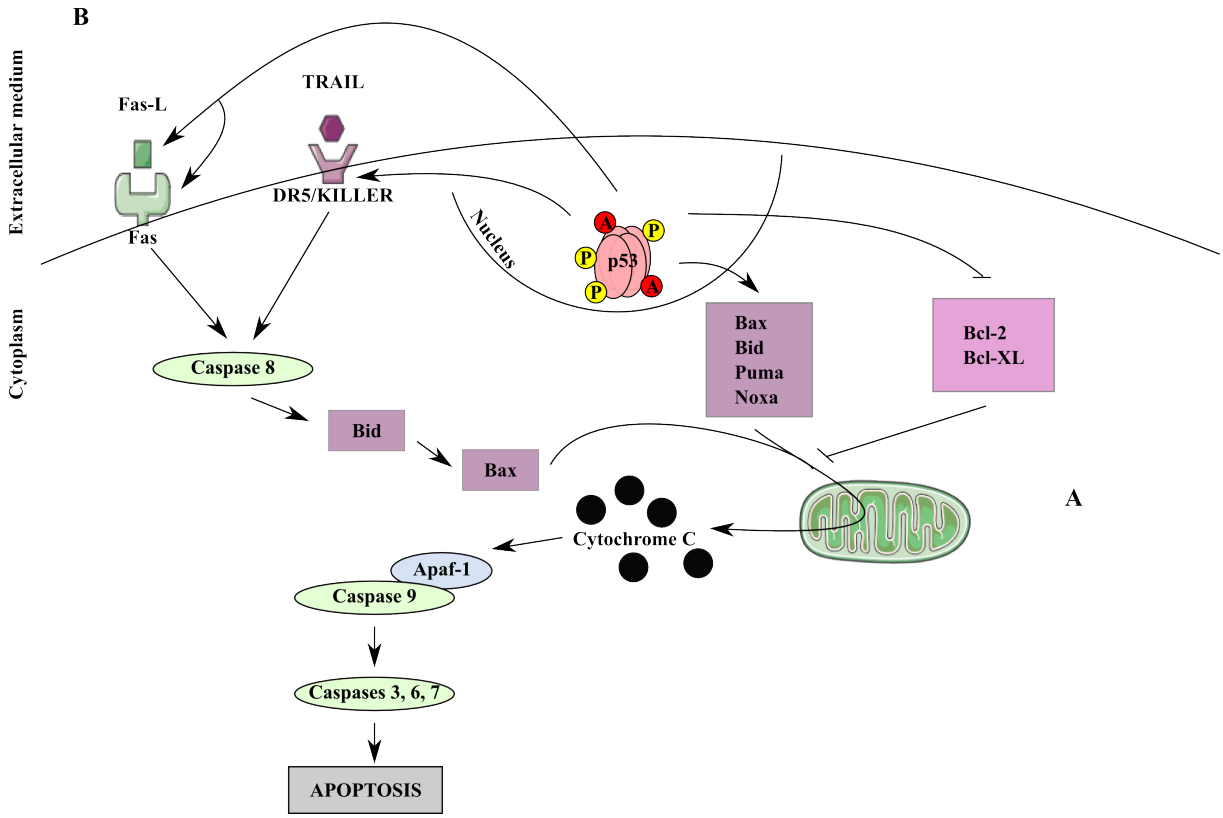


Figure 10: Role of p53 in apoptosis. The p53 protein interacts with different pathways of apoptosis induction, two of which are illustrated here: the mitochondrial intrinsic apoptotic pathway (**A**) and the death receptor mediated extrinsic pathway (**B**).

3. Mitochondrial intrinsic apoptotic pathway

After stress, p53 binds to the p53 RE in the promoters of pro-apoptotic members of the *BAX* (Miyashita et al 1994), *BH3-interacting domain death agonist (BID)* (Sax et al 2002), *PUMA* (Nakano and Vousden 2001) and *NOXA* (Oda et al 2000) (**Figure 10A**). These proteins sequester Bcl-2 and/or Bcl-XL (Yu et al 2001), which are implicated in the control of the mitochondrial membrane permeability. Bcl-2 and Blc-XL are down-regulated by p53 by blocking their transcription (Miyashita et al 1994, Sugars et al 2001). Upon apoptotic stimuli, p53 is induced and inhibits Bcl-2 and /or Bcl-XL, Bax translocates from the cytosol to the mitochondrial outer membrane and deregulates the mitochondrial outer membrane permeability in order to induce the release of cytochrome C and of reactive oxygen species. The released cytochrome C then interacts with Apoptotic protease-activating factor-1 (Apaf1) (Cecconi et al 1998, Moroni et al 2001). Apaf1 could also be directly up-regulated by p53 and it is a key component of apoptosis through its role in caspase 9 activation, initiating the downstream signalling cascade for apoptosis induction through the sequential activation of caspases 3, 6 and 7. The other Bcl-2 pro-apoptotic members such as Puma or Noxa interact with Bcl-2 to release Bax and to promote changes in mitochondrial outer membrane permeability inducing the release of cytochrome C. In addition, a fraction of the cellular p53

translocates into the mitochondria and interacts directly with Bcl-2/Bcl-XL, resulting in the displacement of Bax and the cytochrome C release (Mihara et al 2003). In addition, p53 can also up-regulate miR-34, which inhibits Bcl-2 and cytochrome C release via Bax (Bommer et al 2007). miR-15a and miR-16-1 were shown to induce apoptosis in leukemia cell lines by the down-regulation of Bcl-2 (Cimmino et al 2005).

In addition, p53 can induce PIDD protein expression, implicated in apoptosis induction via caspase 2 activation (Lin et al 2000, Tinel and Tschopp 2004).

4. Death-receptor mediated extrinsic pathway

DNA damage can induce transcriptional up-regulation of some death receptors such as FAS, KILLER/DR5 and DR4 (Takimoto and El-Deiry 2000, Wu et al 1999) through p53-dependent as well as p53-independent mechanisms (El-Deiry 2001) (**Figure 10B**). When p53 is activated in response to an apoptotic stimulus, the Fas Ligand is fixed to Fas. Fas is thus activated and in turn can activate the caspase 8 pathway mediated by the cleavage of the pro-apoptotic Bcl-2 family member Bid resulting in cell death (Bennett et al 1998, Michalak et al 2005, Owen-Schaub et al 1995). Indeed, Bid interacts with some pro-apoptotic Bcl-2 family proteins such as Bax to induce the release of cytochrome C and then the activation of caspase 9. In addition, the caspase 8 pathway can be activated by another mechanism using the tumour-necrosis-factor-related apoptosis inducing ligand (TRAIL), which induces apoptosis through engagement of its receptors, KILLER/DR5 and DR4 (Sheikh et al 1998).

5. Reactive oxygen species pathway

The roles of p53 in regulating levels cellular ROS have been well documented (Hafsi and Hainaut 2011). p53 is associated with ROS generation through the transcription of p53-induced genes (PIGS) that encode redox-dependent enzymes (Johnson et al 1996, Polyak et al 1997). The transcriptional induction of these pro-oxidant genes (Macip et al 2003) results in oxidative stress, followed by ROS release and subsequently apoptosis. These effects synergize with those of genes controlling mitochondrial permeability (see above) in enhancing the leakage of ROS generated by the respiratory chain. In contrast, p53 may also contribute to redox homeostasis by regulating the expression of ROS-detoxifying factors such as *manganese superoxide dismutase* (*MnSOD*) (Dhar et al 2006) thus influencing the ROS balance within cells.

III. Other p53 functions

1. DNA repair

There is accumulating evidence that p53 plays both a direct and an indirect role in DNA repair. For instance, (1) p53 possesses a 3'-5' exonuclease activity; (2) p53 can bind DNA breaks; (3) p53 can interact with the XPB and XPD proteins, involved in Nucleotide Excision Repair (NER) (Wang et al 1996); (4) p53 regulates several factors involved in Base-Excision Repair (BER), including the genes expression of certain glycosylases genes (hOGG1, TDG) (Abedin et al 2013, da Costa et al 2012). It also forms complexes with AP-endonuclease 1 (Ape1/Ref-1), a dual redox/repair enzyme, which is the key endonuclease in BER. The latter interaction appears to have a dual effect, with p53 regulating BER but also APE1/Ref-1 regulating p53, promoting the association of p53 dimers into tetramers and the destacking of higher oligomeric forms into the required tetramer forms which leads to an enhancement of p53 DNA binding activity (Offer et al 2001a, Offer et al 2001b, Zurer et al 2004). In addition, p53 contributes to DNA repair by modulating cell cycle arrest as well as the expression of factors that can impact on the accessibility of the repair machinery, such as GADD45 to DNA (Carrier et al 1999). *GADD45*$^{-/-}$ mice are deficient for NER reparation, suggesting a critical role of this factor in DNA repair (Hollander et al 1993, Smith et al 2000). Finally, p53 up-regulates *p53-inducible ribonucleotide reductase* gene (*p53R2*), which plays a role in supplying deoxyribonucleotides (dNTPs) for DNA repair. In response to DNA damage, *p53R2* is induced in a p53-dependent manner (Tanaka et al 2000).

Recently, using constructs expressing mutant forms of p53 with the S37A or S46A mutations, Serrano and collaborators showed that their phosphorylation is important for homologous recombination (HR) repair (Serrano et al 2013). Indeed, p53 is sequestered by unphosphorylated RPA to maintain a low basal level of p53. After a stress, ATM and ATR induce phosphorylation on S15 of p53 and its DNA repair functions.

2. Senescence

Senescence or biological aging is a tumour suppression mechanism induced by a reduction in telomere length or in response to several forms of stress such as DNA damage or, oncogene activation. Telomere erosion, in particular, induces cells to enter replicative senescence in a p53-dependent manner, thus providing a safeguard system that prevents DNA replication when telomere length is too short to protect chromosomal integrity. Senescence is accompanied and caused by an accumulation of molecular and structural changes that disrupts

metabolism resulting in the arrest of the cell cycle and cell death when DNA repair is not possible (Sahin and DePinho 2012).

p53 is associated with different types of senescence. First, p53 plays a role in replicative senescence. After DNA damage, different kinases such as ATM or ATR are activated and induce the phosphorylation of Chk1, Chk2 and p53 and their activation (Herbig et al 2004). In TERC-deficient mice, lacking the RNA component of telomerase, p53 over-expression is associated with the activation of senescence and reduction of tumour formation (Chin et al 1999). $TERC^{-/-}$ mice show a dwarf phenotype as well as phenotypes associated to premature loss of tissue regeneration, including the skin. Abrogation of p53 in this genetic background rescues both the small size phenotype and the functionality of epidermal stem cells (ESC) of telomerase-deficient mice with dysfunctional telomeres, accompanied by a down-regulation of senescence markers and an increased cell proliferation. Together, these findings indicate the existence of a p53-dependent senescence response acting on stem/progenitor cells with dysfunctional telomeres that is actively limiting their contribution to tissue regeneration (Flores and Blasco 2009).

Secondly, p53 is important in oncogene-induced senescence. In response to Ras oncogene activation, senescence-like G1 arrest is promoted in cells with WT p53 (Serrano et al 1997). Furthermore, Ras regulates p53 through the Mitogen-activated protein kinases (MAPK) pathways involving Raf-1, MEKs (MEK1/2) and MAPKs (Extracellular-signal-regulated-kinases: ERK1/2) (Campbell et al 2007, Lin et al 1998). Ras can also regulate p53 through the activation of DNA damage response via the ATM or ATR pathways (Bartkova et al 2006).

Thirdly, p53 is induced in genotoxic stress-induced senescence. After treatment with chemotherapeutic drugs including doxorubicin, cells enter by senescence and are eventually eliminated in a p53-dependent manner (Wang et al 1998) with an accumulation of $p21^{WAF1}$ protein, which enables a transitory senescence or a p16 over-expression, which enables a non-transitory senescence (Alcorta et al 1996). Other p53 target genes implicated in senescence, such as micro-RNA form miR-34 family, have been also described (He et al 2007).

3. Autophagy

Autophagy serves as a cell survival mechanism via its suppressive role on necrotic cell death. The anti-necrosis function of autophagy has important biological functions in cancer and p53 plays a role in the control of mediators of autophagy (Feng et al 2005, Ghavami et al 2011).

Nuclear p53 can activate several p53 target genes implicated in autophagy, such as damage-regulated autophagy modulator (DRAM), which encodes a lysosomal protein that induces autophagy (Crighton et al 2006). Indeed, in response to DNA damage, p53 directly interacts with DRAM and induces autophagy (Crighton et al 2006, Ghavami et al 2011) or mammalian Target of Rapamycin (mTOR), which is implicated in survival and proliferation (Feng et al 2005). DRAM could also be implicated in apoptosis (Crighton et al 2006). These observations suggest that p53 play a role in the regulation of cell death by controlling both autophagy and apoptosis. In addition, Sestrin2, a protein implicated in the regulation of cell growth and survival, positively regulates autophagy in a p53-dependent manner (Maiuri et al 2009).

In contrast to nuclear p53, cytoplasmic p53 inhibits autophagy. This inhibition is cell cycle-dependent and occurs mostly in G1, less, in S phase but not in the G2/M phase of cell cycle (Tasdemir et al 2008b).

4. Angiogenesis

Angiogenesis is the physiological process by which new blood vessel capillaries are formed. It occurs during normal tissue development but can also occur in adulthood only under certain conditions. For instance, uncontrolled angiogenesis is important in tumour development for delivering both oxygen and nutrients (Folkman et al 1963). It has been demonstrated that p53 plays an important role in the inhibition of angiogenesis by three different pathways (1) inhibition of hypoxia-sensing systems, (2) down-regulation of pro-angiogenesis genes, (3) up-regulation of anti-angiogenesis pathways.

The central regulatory component that responds to oxygen deprivation (hypoxia) is the transcription factor hypoxia inducible factor (HIF) that up-regulates the production of new blood vessels (Pugh and Ratcliffe 2003). It is composed of two subunits, HIF-1α and HIF-1β and can trans-activate several genes implicated in the response to oxygen deprivation such as vascular endothelial growth factor (VEGF), an essential factor for developmental and tumour angiogenesis (Carmeliet et al 1998, Forsythe et al 1996). HIF-1α activity can be inhibited by p53 binding to HIF-1α and allowing the degradation of HIF-1α after p53 activation and accumulation (Ravi et al 2000). The inhibition of HIF-1α by p53 promotes the inhibition of the formation of blood vessel in a hypoxic tumour and does not require p53 transcriptional activity. In addition, miR-107, a p53 target miRNA, targets the HIF-β subunit, which controls hypoxic programming in colon cancer cells (Yamakuchi et al 2010).

Angiogenesis is regulated by a tight network of pro-angiogenic and anti-angiogenic factors, and p53 can influence both aspects of these regulations. For example, p53 can repress pro-angiogenesis factors such as VEGF (Pal et al 2001), basic fibroblast growth factor (bFGF) (Ueba et al 1994), bFGF-binding protein (bFGF-BP) (Sherif et al 2001) or Cyclooxygenase-2 (COX-2), a key player in inflammation (Subbaramaiah et al 1999). p53 inhibits *VEGF* transcription by inhibiting the Specific protein 1 (SP1) transcription factor which activates *VEGF* transcription (Pal et al 2001). p53 is also in competition with TBP (TATA binding protein) for COX-2 promoter binding (Subbaramaiah et al 1999). Finally, it is known that p53 can activate some anti-angiogenesis factors such as Thrombospondin-1 (TSP-1) (Dameron et al 1994), brain-specific angiogenesis inhibitor 1 (BAI1) (Nishimori et al 1997) or ephrin receptors A2 (EPHA2) and ephrin-A1 (Brantley et al 2002, Dohn et al 2001)

5. Metabolism

Control of cellular metabolism is an important key of normal cell behaviour and the deregulation of metabolism could play a role in disease development. p53 is emerging as an important regulator of cell bioenergetics by controlling the pathways by which cells mobilize energy production systems, glucose metabolism, mitochondrial respiration, glutaminolysis and fatty acid oxidation.

Glucose metabolism, (glycolysis) is a major source of energy for many cells, including in particular proliferating cells, through the generation of Adenosine Triphosphatase (ATP) in the absence of oxygen consumption. p53 can negatively regulate glycolysis by controlling the transcription of Hexokinase II, involved in the first step of glycolysis (Mathupala et al 2006) and phospho-glycerate mutase M isoform, important in the 8th step of glycolysis (Ruiz-Lozano et al 1999). In addition, p53 also up-regulates Glucose 6-Phosphate dehydrogenase, the key enzyme redirecting glucose metabolites into the phosphate pentose pathway, a critical pathway for the biosynthesis of reducing equivalents and nucleotides required for proliferation. .

Under aerobic conditions, the pyruvate produced by glycolysis can be fed into the mitochondrial tricarboxylic acid cycle (TCA) for high-yield ATP production via oxidative phosphorylation. The p53 protein influences the TCA cycle for energy generation (Stambolsky et al 2006). For example, p53 regulates the use of aas such as glutamine as a source of carbons for the TCA cycle. This pathway has been shown to be important in cancer cells through the activation of one isoform of the glutaminase (GLS1/KGA), which converts glutamine to glutamate (Wang et al 2010).

Finally, p53 contributes to promote oxidative phosphorylation in the mitochondria by regulating AIP1 and SCO2, two genes involved in the assembly of complex I and complex IV of the respiratory chain, respectively (Kaeser and Iggo 2002, Matoba et al 2006). Thus, on balance, p53 acts as a suppressor of glycolysis and a promoter of oxidative phosphorylation (Kulawiec et al 2009, Lebedeva et al 2009). Loss of p53 function in cancer therefore contributes to the metabolic adaptation of cancer cells to high glycolysis rates even in the presence of oxygen, a phenomenon called "aerobic glycolysis", also widely known as the "Warburg effect" (Warburg 1956).

6. Migration and invasion

Invasion and metastasis are interrelated processes that characterize the malignant state of cancer. p53 can prevent different steps of the metastasis cascade such as the Epithelial-mesenchymal transition (EMT), cellular contraction and propulsion. EMT is a physiological process involved in development and in wound healing, by which epithelial cells can transiently adopt a phenotype of mesenchymal cells, become mobile and contribute to modify the architecture of the stroma to support novel epithelium-mesenchyme interactions. This process has been shown to be critical for the development of the invasion front of epithelial tumors and for creating metastatic niches in distant organ location (Thiery et al 2003). p53 has been implicated in the early stage of metastasis and contributes to the migration of stem cells to distant sites. Studies, using 3D matrices, showed that the loss of p53 increases cell mobility in epithelial cancer cell lines (Xia and Land 2007) and neurons (Qin et al 2009). These results suggest an important role of p53 in migration. p53 can inhibit EMT by different ways such as enhancing Mdm2-mediated degradation of Slug to enhance E-cadherin expression (Wang et al 2009). An important target for p53 in regulating EMT is miR-200c, a member of a family of miRNAs acting as orchestrators of EMT through their capacity to inhibit EMT-activators such as ZEB factors, thereby inducing a reverse process to EMT, the Mesenchymal to Epithelial Transition (MET). MiR-200 family members not only counteract EMT but also suppress stem cells factors, such as Bmi1 (Chang et al 2011). Another important connection between p53, EMT and metastasis involves the Twist1 protein, a regulator of embryogenesis. Twist1 has been shown to induce EMT and is over-expressed in a large fraction of human cancers (Ansieau et al 2008). A common cancer-derived mutant p53 protein, p.R175H, up-regulates Twist1 expression in several cancer cell lines, suggesting that increased Twist1 might occur in cancer cells as the consequence of gain-of-function mutations in *TP53* (Kogan-Sakin et al 2011).

p53 can also contribute to limit cell motility by trans-activate RhoE, which inhibits RhoA activity, a protein implicated in cellular contraction (Gadea et al 2007). Focal Adhesion Kinase (FAK), a protein implicated for cellular tension essential in migration, is also a target gene of p53 (Golubovskaya et al 2008). Additionally, p53 can increase miR-143 expression, which negatively regulates invadopodia that can trigger the degradation of ECM components and basal membranes to allow cells to access the stroma (Quintavalle et al 2010).

7. Regulation of miRNAs machinery

miRNAs are produced from either their own genes or from introns of other genes and can be produced by a specific miRNA pathway. p53 can regulate miRNAs expression in two different ways: by regulating miRNA pathways or miRNA expression directly.

The p53 protein can directly interact with the DEAD-box RNA helicase p68 and enhances its interaction with the DROSHA complex, composed of DROSHA, DGCR8 and p68, with initiates miRNA processing (Suzuki et al 2009). p53 can promote the formation of pri-miRNAs (primary transcript miRNAs), a precursor of pre-miRNAs to increase the level of mature miRNA such as miR-16-1, miR-143 and miR-145. In addition, p53 can influence the selection of miRNA targets by inducing RNA binding proteins such as RBM68, which compete with miRNAs for binding to 3'UTRs of mRNA (Leveille et al 2011). *RBM68* gene is often epigenetically silenced in breast cancers, which can explain miRNA functions for tumour suppression by p53.

p53 regulates gene expression by targeting its specific response element sequence. Recently, it was observed that p53 can also regulate miRNA expression. Similar to the transcription of its target genes, induction of some miRNA requires an intact p53-binding site such as miR-34 (Bommer et al 2007), miR-15/16 (Fabbri et al 2011), miR-107 (Bohlig et al 2011).

IV. The p53 biological repertoire: orchestrating multiple biological functions

The results summarized above demonstrate the multiple functions of p53, which is capable of regulating multiple coordinated, and even sometimes competing, biological responses. The p53 protein has a very large repertoire of targets, being able to bind to more than 1500 genes promoters implicated in several cellular functions such as cell cycle arrest, apoptosis, DNA repair, metabolism or senescence (Wei et al 2006). In addition, p53 also influences many processes by direct binding to proteins. For a long time it has been

considered that p53 could choose between cell cycle arrest and apoptosis, the first response occurring at low levels of DNA damage (protecting cell survival after repair) and the second occurring at high levels of DNA damage (eliminating cells damaged beyond repair). It is now evident that the p53 functions are larger than these basic responses. Factors that may affect the way p53 selects among these many biological effects include cell type and proliferation context, type and intensity of signals inducing p53 activity (leading to a specific panel of post-translational modifications) and presence or absence of a number of p53 partners. These factors define an extended biological context that determines when, how, and to which effect p53 may mediate a set of events that concur to a coordinated growth suppressor response.

At the molecular level, such differential responses may depend on conformational changes of the p53 DBD that impact on the binding affinity to specific p53 REs in target genes. Different factors can modulate the DBD conformation: (1) patterns of post-translation modifications, for example, acetylation of K120 which favours the trans-activation of pro-apoptotic genes (Sykes et al 2006); (2) protein interactions, for example, ASSP binding to p53 DBD, which increases the fixation of p53 to apoptotic target promoters (Qian et al 2002); (3) redox regulation of the conformation of the DNA-binding domain, three aas also involved in the binding of zinc and modulates the tri-dimensional structure of the DBD (Hafsi and Hainaut 2011). Other factors such as p53 complex degradation machineries, or co-expression of p53 isoforms, add further layers of complexity to these mechanisms.

Depending on p53 activation, p53 REs show significant sequence variations depending on the corresponding genes and their biological functions. For example, p53 REs are more conserved in cell cycle arrest target genes than in apoptotic target genes (Horvath et al 2007). The p53 REs of cell cycle genes generally present less than 2 mismatches compared to the optimal consensus sequence, whereas apoptotic genes contain more than 3 mismatches (Qian et al 2002). Diversity also occurs through variation in the number of juxtaposed elements. For example, the *BAX* gene contains 3 juxtaposed elements and is a low-affinity target as compared to $p21^{WAF1}$ contents, which contains only one repeat (Thornborrow and Manfredi 1999).

A biological example of the subtle variations in how p53 selects it biological repertoire was described by Wilson and collaborators (Wilson et al 1998). These authors analysed the activation of p53 in the intestine of mice after whole body irradiation. They observed that the highest accumulation was observed in cells at the bottom of the crypts, which corresponds to the localisation of the progenitor cells. This accumulation was correlated with a predominant apoptotic response. However, in differentiated cells along the

intestinal villosities, p53 accumulated at lower levels, associated with an activation of $p21^{WAF1}$ and cell cycle arrest. Finally, terminally differentiated cells at the top of p53 failed to show any significant accumulation of p53. This example shows that p53 induction and accumulation responses may vary even within a same cell type and tissue, depending of the place of the cells in the differentiation sequence. Overall, these coordinated responses concurred to maintain genetic integrity while at the same time protecting the integrity of the organ.

In recent years, it has emerged that p53 may exert at least part of its suppressor function in a wider biological context than the acute response to various forms of DNA damaging or oncogenic stress. These studies have analysed the physiological role of p53 in coordinating cell cycle arrest, senescence, mitochondrial metabolism and apoptosis in response to the reduction of telomere length. Sahin and collaborators have showed that mice with constitutively dysfunctional telomeres accumulated relatively low levels of persistent DNA damage in a variety of organs including hematopoietic stem cells, heart and liver (Sahin et al 2011). This DNA damage resulted in p53 activation. In such conditions, p53 induced the repression of peroxisome proliferator-activated receptor γ co-activators (PGC)-1α and PGC-1β,. These two factors are the master regulators of mitochondrial physiology and metabolism. p53 induced impaired mitochondrial biogenesis and function, decreased gluconeogenesis, cardiomyopathy and increased production of ROS. These conditions resulted in replicative senescence and in the depletion of pools of stem/progenitor cells responsible for tissue renewal. Ablation of *TP53* substantially restored PGC expression and the functionality of the subsequent redox and energy metabolic network, resituating mitochondrial respiration, cardiac function and gluconeogenesis in liver cells. These observations demonstrate that p53 couples telomere attrition and subsequent DNA damage with metabolic effects, senescence and aging.

E. Genetic alterations of *TP53* in cancers

The expression of the *TP53* gene (OMIM #191117) and the functions of the p53 protein can be modulated by different factors such as the presence of mutations (somatic or germinal) and/or the presence of certain polymorphisms. The presence of both *TP53* mutations and/or polymorphisms can also influence an individual's cancer susceptibility. Indeed, the *TP53* gene is mutated in around 50% of cancers but this proportion depends on the type of cancer. However the correlation between the "genotype" (*TP53* mutation) and its use as a prognostic biomarker or with the "phenotype" of the tumour remains controversial for many cancer types. To identify genetic factors that might impact on this correlation between "genotype" and "phenotype", many labs are examining the roles of *TP53* polymorphisms.

I. Somatic mutations in *TP53* gene

1. Epidemiological data

Somatic *TP53* gene mutations are found in almost every types of cancer at various frequencies (**Figure 11**) (IARC *TP53* database, p53.iarc.fr). The cancers in which *TP53* mutations are most frequently found are colorectal (43.28%) and head and neck (42.51%). Some cancers have an intermediary frequency of *TP53* mutations such as those of the urinary tract (26.25 %), nervous system (25.95%) or a lower frequency of *TP53* mutations (bones cancers: 14.43%).

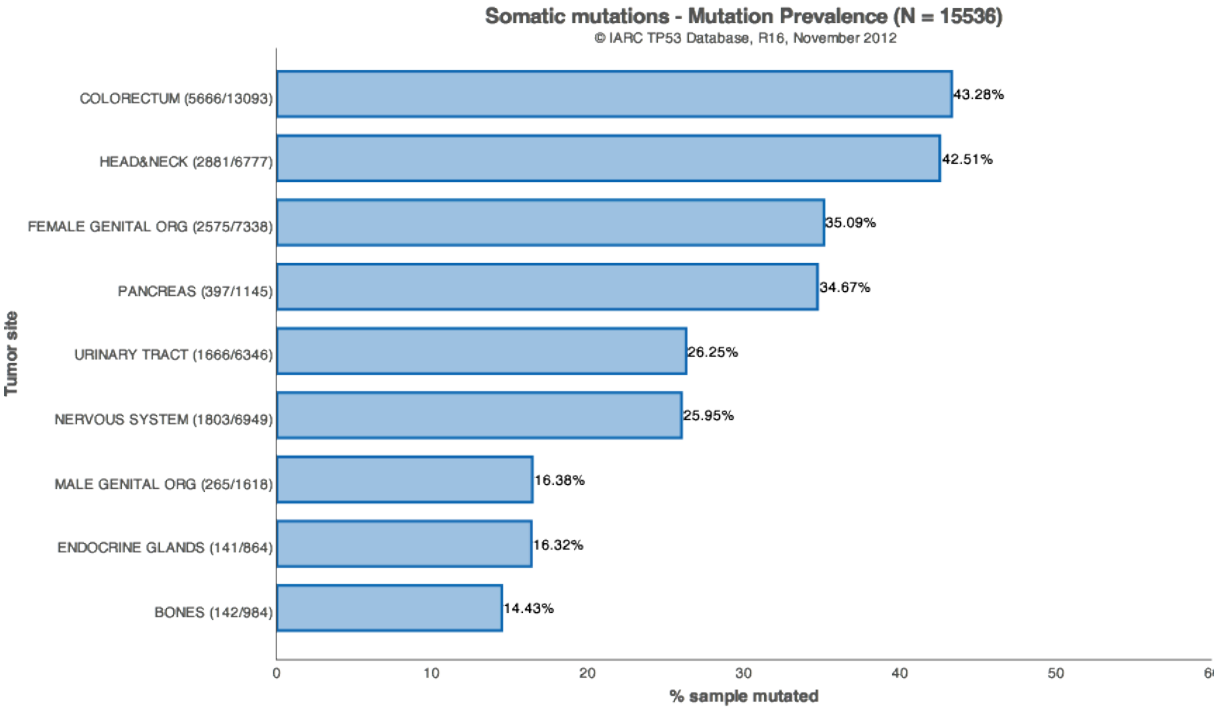


Figure 11: *TP53* mutation distribution depending on cancer types.

The vast majority of the mutations described in *TP53* are small mutations including missense mutations (substitution of an aa by another one); nonsense mutations (formation of a premature stop codon which forms a truncated p53 protein); silent mutations (substitution of a nucleotide by another without any subsequent change in the aa due to the redundancy of the genetic code); insertion or deletion of a small number of nucleotides resulting in a shift in the reading frame (Petitjean et al 2007). In contrast to other tumour suppressor genes that are mainly altered by truncating mutations, the most frequently found mutation type in the *TP53* gene is missense mutations (around 75%) (**Figure 12A**) ((Hussain and Harris 1998, Petitjean et al 2007); p53.iarc.fr). Other alterations include frame shift insertion (9.01%), nonsense mutations (8.14%), silent mutations (3.67%) or large deletions (0.12%) (Petitjean et al 2007). These results suggest that a selection pressure favours the production of an altered p53 protein compared to the loss of the WT protein.

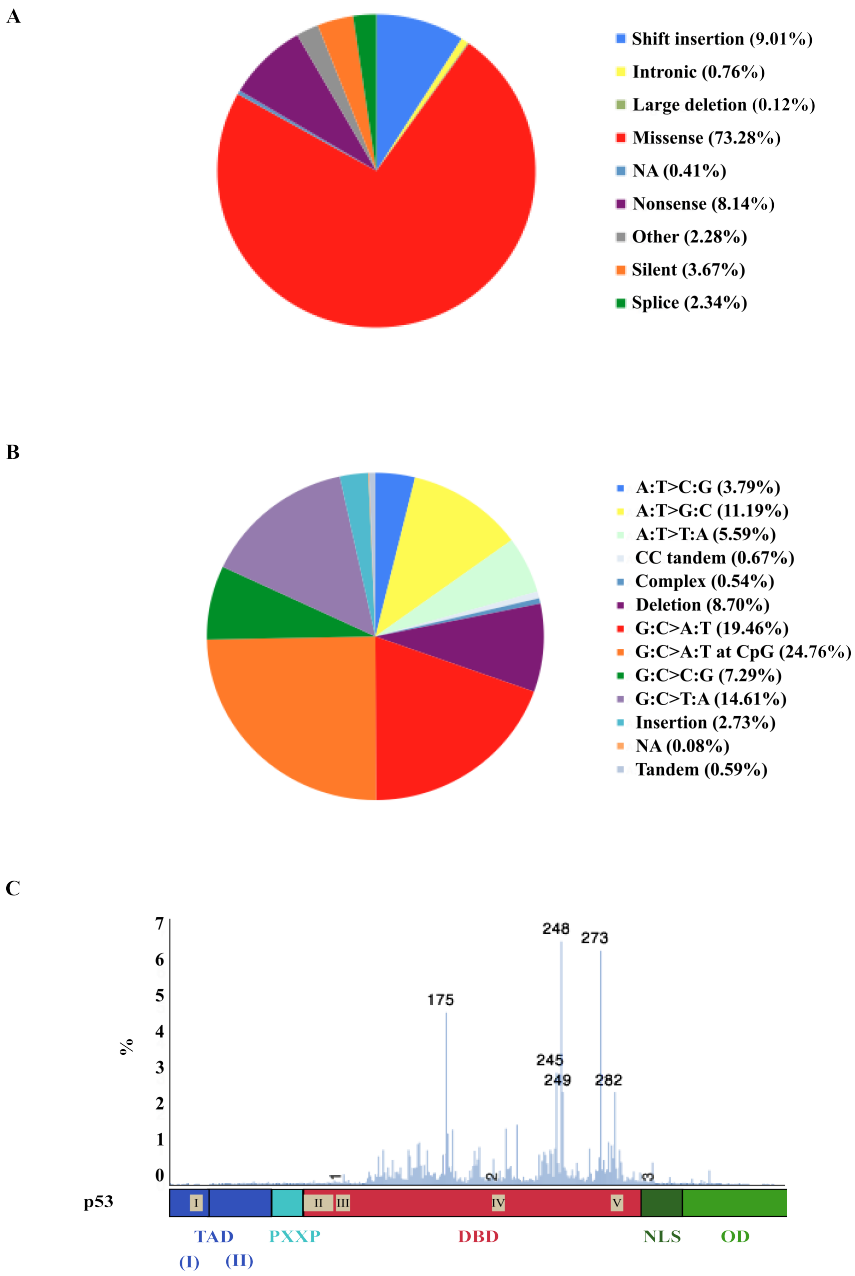


Figure 12: The distribution of *TP53* somatic mutation types distribution (**A**) Percentage of different types of mutations; (**B**) Percentage of the base pairs involved; (**C**) Codon distribution of *TP53* mutations. (n=28,581) (From www.p53-iarc.fr)

The base pairs most frequently mutated are G:C>A:T (44,22%): at CpG sites (24.76%) or at other sites (19.46%) but a large spectrum of mutations is observed in cancers (**Figure 12B**). The *TP53* mutations are mostly located in the DNA binding domain, exons 5 to 8 (around 80%), and one-third of all reported mutations are found in 6 specific codons called "hot spots" (R175H, G245S, R248Q/W, R249S, R273H/C, R282W) (**Figure 12C**) (Hsu et al 2001, Petitjean et al 2007). Two categories of missense mutations are described: conformational mutants, which modulate the tri-dimensional structure of the p53 DBD (for

example: R175H, Y220C, G254S, R249S, R282W or aas located in L2 loop and helix), and contact mutants, which influence the interaction with DNA and the tri-dimensional structure (for example: R248Q/W, R273H/C or aas located in the L3 loop). In addition, *TP53* is sensitive to loss of heterozygosity (LOH), i.e. the loss of one of the two *TP53* alleles in cancers. This was first observed in colorectal carcinomas, which were found to carry a deletion of the WT allele and to maintain the second mutant allele (Baker et al 1989). This finding suggested that the development of a p53-dependent cancer needed two steps: (1) the occurrence of a mutation, which alters the suppressive function of p53 on the first allele, (2) the loss of the WT p53 allele expression on the second allele. Soon after though another study showed that the presence of a *TP53* mutation alone was enough to result in the development of cancer and this hypothesis became obsolete (Nigro et al 1989). A subsequent study showed that half of $p53^{+/-}$ mice, which express a WT p53 protein, develop spontaneous tumours later than $p53^{-/-}$ mice (Venkatachalam et al 2001). These results suggest that the diminution of p53 protein levels due to a haplo-insufficiency is sufficient to increase the risk of cancer development.

2. Association between cancer and specific *TP53* mutation

In several cancers types, mutation patterns bear the hallmarks of chemical damages induced by particular mutagens or by environmental factors leading to the concept of a *TP53* "fingerprint" (Olivier et al 2010). These fingerprints are defined by a higher frequency of a specific type of base change, strand orientation or localisation of a particular base substitution. In some cancers, the factors, which influence the cancer susceptibility and the frequency of *TP53* mutation, are well understood.

In hepatocellular carcinomas, the *TP53* R249S mutation is observed in more than 50% of cases and has not been found in others cancers (Montesano et al 1997). Epidemiological studies have shown that the frequency of this mutation is associated with the presence of the Aflatoxin B1, a mycotoxin produced by *Aspergillus* (Gouas et al 2009). This mycotoxin can form DNA adducts and induce *TP53* mutation formation. These DNA adducts cause the G to thymine (T) transversion at the third base of codon 249 (AGG to AGT; R to S).

In lung cancers, the proportion of G to T transversions is dependent on tobacco exposure. In smokers, 30% of the *TP53* mutations are G to T transversions contrary to the mutation profile seen in non-smokers where only 13% are G to T transversions (Hainaut and Pfeifer 2001). These mutations are mostly found at codons 157, 158, 248 and 273 (Pfeifer et al 2002). These mutation hotspots are associated with adduct positions formed by

benzo(a)pyrene exposure, one of the principal chemical components of tobacco (Denissenko et al 1996).

In non-melanoma skin cancers, the exposure to sunlight and more precisely, UV radiation, increases the presence of tandem CC to TT transitions at codons 177-179 and 278 (Inga et al 1998, Ziegler et al 1994). In skin tumours developed by *Xeroderma pigmentosum* patients, almost 50% of all *TP53* mutations are CC to TT transitions (Daya-Grosjean et al 1995). There is evidence that the preferential accumulation of mutations at these hot spot codons is due to the slower repair of UV-induced lesions at these positions (Tornaletti et al 1993, Tornaletti et al 1994).

3. Biological properties of p53 mutant proteins

To evaluate the properties of the somatic p53 mutant proteins, analyses were performed on their transactivation activity, capacity to induce cell cycle arrest of apoptosis, dominant-negative effect, sensitivity to temperature and gain of function (Brosh and Rotter 2009, Ory et al 1994, Petitjean et al 2007, Weisz et al 2007). Experimental assays in human cells and in yeasts have shown that all p53 mutants were not equivalent and that the variations of their properties separated them into different groups: (1) loss of transactivation function, (2) dominant-negative and (3) gain of function (**Figure 13**) (Kato et al 2003, Petitjean et al 2007).

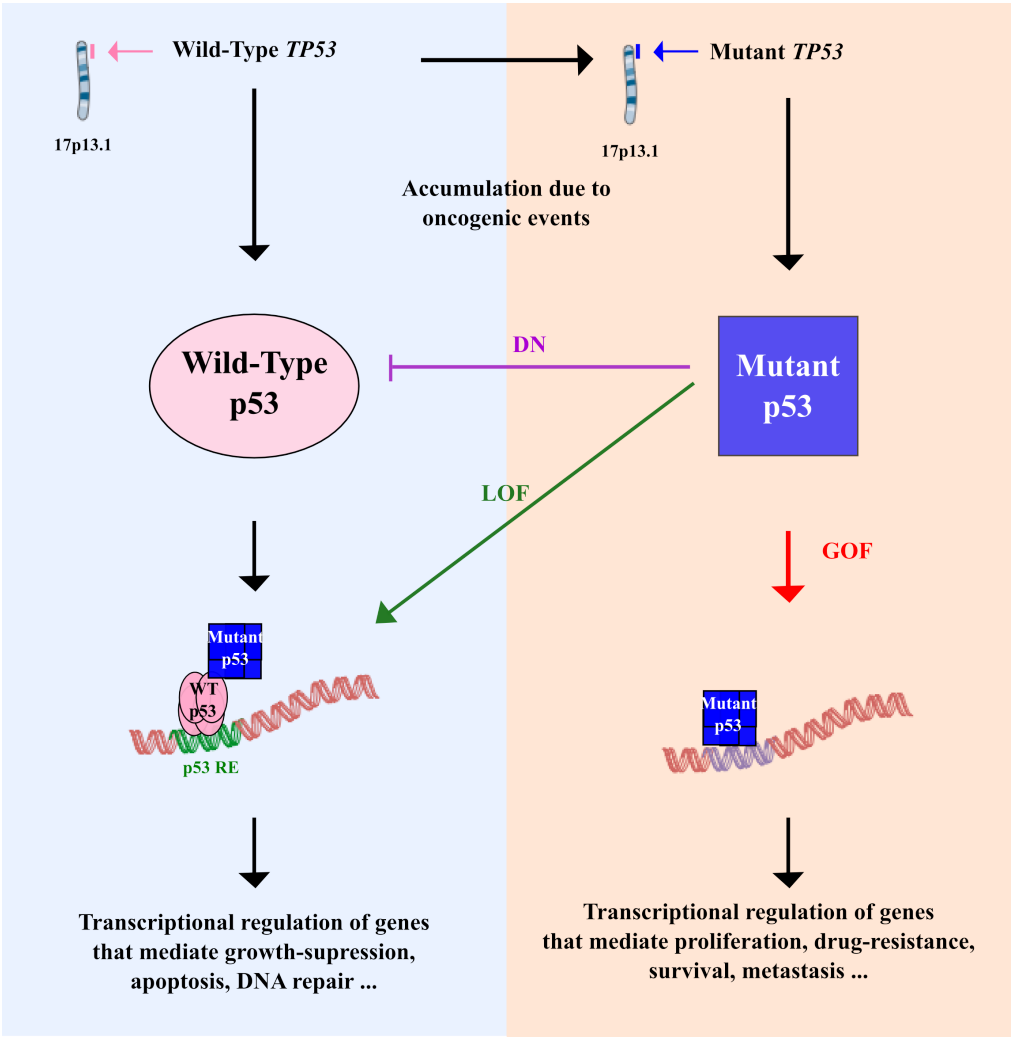


Figure 13: Oncogenic effects of p53 mutants. DN: Dominant-negative effects; LOF: Loss of function or transactivation; GOF: Gain of function (Adapted from (Brosh and Rotter 2009, Guimaraes and Hainaut 2002) and Thesis of Doriane Gouas, University Lyon 1, 2011)

Loss of trans-activation

The main activity of the p53 protein is to regulate gene expression through specific DNA-binding RE located in p53 target genes. For most cancers, the main characteristic is the loss of p53 trans-activation function and the subsequent loss of the p53 target genes' expression (Raycroft et al 1990). In fact, the vast majority of mutations found in cancers are essentially located in the p53 DNA binding domain where the substitution of one nucleotide is associated with the loss of the DNA binding activity (Flaman et al 1994). These non-functional mutants are considerate as "loss of function" mutants or "loss of trans-activation" mutants.

A study using a comprehensive site-directed mutagenesis and a yeast-based functional assay has allowed the transcriptional activity of 2,314 p53 mutants representing all possible amino acid substitutions caused by a point mutation throughout the protein on 8 reporter

genes to be evaluated (Kato et al 2003). This study has shown that only 36% of analysed p53 mutants exhibit complete loss of function and others mutants exhibit a wide diversity of loss of function. Some mutants retain an activity on some p53 REs and show a complete loss of their activity on others, while some mutants retain a partial activity on most REs. Based on the IARC *TP53* database (p53.iarc.fr), these complete "loss of function" mutants represent over 80% of the p53 mutants found in human cancers. Taken together, these results suggest that p53 mutants with loss of function are more selected in cancer development. However, the frequency of the "loss of function" mutants is not the same in all sporadic cancers (Petitjean et al 2007).

Dominant-negative effects

p53 transcriptional activity relies on the formation of tetramers. Since most mutant proteins retain an intact oligomerisation domain mutants may interfere with WT p53 by forming hetero-oligomers less competent for specific DNA binding. The capacity of the mutant proteins to interfere with the WT protein has been studied in yeast and human cell assays. The formation of hetero-oligomers induces (1) the stabilisation of the WT p53 protein by these "dominant-negative" proteins (Shaulian et al 1992); (2) the conversion of the WT p53 tri-dimensional conformation into a mutant tri-dimensional conformation (Milner et al 1991); and (3) the abolition of the DNA binding activity of the WT p53 protein and its transcriptional activity (Farmer et al 1992, Kern et al 1992, Shaulian et al 1992).

Data have been produced on more than 200 mutants and have shown that "dominant-negative" effects may be promoter-, cell type- and mutant type-dependent (Petitjean et al 2007). For example, the R248W mutant decreases the transcriptional activity of the WT p53 while the H179Q mutant completely abolishes it (Unger et al 1993). It seems that the "dominant-negative" effect is more important when the mutant has only a low DNA binding activity (Chene 1998). However, an analysis in the IARC *TP53* database confirmed the correlation between "dominant-negative" effects and frequency of the occurrence in cancer (Petitjean et al 2007). The "dominant-negative" mutations are three times more frequent in breast cancer compared to the non-"dominant-negative" mutations, and twice less frequent in osteosarcomas. "Dominant-negative" effects may thus play a significant role in the selection of mutations due to the fact that they do not need the loss of the second allele to completely abolish the p53 suppressive activity compared to the "loss of function" mutants.

Gain of function

Over 20 years ago, the concept that mutant p53 proteins gain tumour-promoting functions was first proposed by showing that they have oncogene effects in the absence of the WT p53 protein. Many oncogenic functions of mutant p53 have been characterised in mice or cellular models such as the ability to promote tumour development, invasion, migration, angiogenesis survival and proliferation. These properties are referred to as "gain of function". These "gain of function" mutations are mostly located in the DNA-binding domain (hotspots: R175, G245, R248, R249, R273 and R282) suggesting that changes in transcriptional target genes could be the key to the activity of these mutants. Indeed, these mutants activate new genes normally unaffected or repressed by the WT p53 protein, or interfere with other transcription factors (Bossi et al 2006, Muller and Vousden 2013). Studies have observed that "gain of function" mutants are able to regulate the transcription of genes implicated in cell cycle progression such as Cyclin A, Cyclin B1 or Cdk1 (Di Agostino et al 2006) and repress the transcription of genes implicated in apoptosis such as Fas (Zalcenstein et al 2003). Several molecular mechanisms by which p53 mutant may act as by "gain of function" have been proposed: through the interaction of mutant p53 directly using a mutant p53 binding element, or the interaction of mutant p53 with others proteins or co-factors (for example: R175H/ nuclear transcription factor Y (NF-Y) complex recruits p300 and fixes NF-Y REs (Di Agostino et al 2006); DNA topoisomerase 2-binding protein 1 recruits mutant p53 and p300 to target promoters (Liu et al 2011); inactivation of p53 target proteins by their sequestration to inactivate their transcriptional activity (Marin et al 2000, Strano et al 2000); interaction of mutant p53 with proteins to change their function (for example: mutant p53 interacts with Meiotic recombination (MRE) 11, a DNA nuclease required for DNA repair, to prevent the formation of the MRE11-RAD50-NSB1 complex leading to impaired HR (Liu et al 2010a, Song et al 2007).

The frequency of the "gain of function" mutation in cancers could be higher compared to that of the "loss of function" or "dominant-negative" mutants because these mutations have two roles: inactivation of a tumour suppressor and activation of an oncogene. However, the difference in the frequency between "dominant-negative" mutants and "gain of function" is still not clear (Petitjean et al 2007). Most of the "gain of function" mutants also show "dominant-negative" characteristics, so, it is difficult to separate both effects in tumours.

II. Germline mutations and Li-Fraumeni syndrome

1. Characteristics of the Li-Fraumeni syndrome

Definitions

The LFS was identified in 1969 by two American physicians, Frederick Li and Joseph Fraumeni (Li and Fraumeni 1969). This syndrome is characterised by a high incidence of cancers at an early age of onset (**Table 2**) (Li 1995). It is a clinically and genetically rare autosomal dominant syndrome with a complex profile.

Table 2: Clinical criteria for classic Li-Fraumeni syndrome (LFS) and LFS-like (LFL) criteria, and Chompret criteria.

Classic LFS criteria (Li et al 1988)
▪ Proband diagnosed with a sarcoma before age 45; AND ▪ A frist-degree relative with cancer diagnosed before age 45; AND ▪ Another first- or second-degree relative on the same side of the family with cancer diagnosed before age 45 or a sarcoma at any age
Birch criteria (Birch et al 1994)
▪ Proband with any childhood cancer or sarcoma, brain tumour, or adrenocortical carcinoma diagnosed before age 45; AND ▪ First- or second-degree relative with a typical LFS cancer (sarcoma, breast cancer, leukaemia, or adrenocortical carcinoma) diagnosed at any age; AND ▪ A first- or second-degree relative on the same side of the family with any cancer diagnosed under age 60
Eeles criteria (Eeles 1995)
▪ Two first- or second-degree relatives with LFS-related malignancies (sarcoma, breast cancer, brain tumour, leukaemia, adrenocortical tumour, melanoma, prostate cancer, pancreatic cancer) at any age
Chompret criteria (Chompret et al 2001)
▪ Proband diagnosed by a LFS tumour spectrum (sarcoma, breast cancer, brain tumour, adrenocortical carcinoma) before the age of 36; AND ▪ A first- or second-degree relatives with related malignancies (sarcoma, breast cancer (other than breast cancer if the proband is affected by breast cancer), brain tumour, leukaemia, adrenocortical tumour, melanoma, prostate cancer, pancreatic cancer) before age 46
Chompret criteria for *TP53* germline mutation screening (Tinat et al 2009)
▪ Proband diagnosed with a tumour belonging to the LFS tumour spectrum (sarcoma, brain tumour, premenopausal breast cancer, adrenocortical carcinoma, leukaemia, lung bronchoalveolar cancer) before age 46, and at least one first- or second-degree relative affected with a LFS tumour (other than breast cancer if the proband is affected by breast cancer) before age 56, or a relative with multiple primary tumours at any age; OR ▪ A proband with multiple primary tumours (except multiple breast tumours), two of which belong to the LFS tumour spectrum and the first of which occurred before age 46, regardless of family history; OR ▪ A proband with adrenocortical carcinoma or choroid plexus tumour, regardless of family history

The clinical heterogeneity of the syndrome causes problems of classification. The specificity of the tumour spectrum and the number of cancers that develop in these patients has resulted in the evolution of the LFS criteria. The definitions of these modified criteria

(Chompret, Birch, Eeles 1 or Eeles 2) have increased the number of diagnosed patients and have defined the Li-Fraumeni like syndrome (LFL) (**Table 2**) (Birch et al 1994, Chompret et al 2001, Eeles 1995, Tinat et al 2009). The Chompret criteria have been re-evaluated and up-dated (Bougeard et al 2008, Tinat et al 2009). Note that the approach behind these modified Chompret criteria is not to develop an exhaustive clinical definition of the syndrome but to propose a set of operational clinical criteria to identify probands to be referred for *TP53* mutation testing.

Tumour spectrum

The most frequent tumours in individuals with LFS are pre-menopausal breast cancers, soft tissue sarcomas (STS), brain cancers, adrenal cortical carcinomas (ADR) and osteosarcomas. These tumours are essential components of the classic definition of LFS (Li et al 1988, Malkin et al 1990). Several studies have observed that leukaemia are not the most frequent cancer type in LFS patients compared to the general population, so it is considered as a non-specific tumour in terms of the characteristic LFS tumour spectrum (Birch et al 2001, Gonzalez et al 2009, Nichols et al 2001, Ruijs et al 2010).

Other tumours types are also diagnosed in LFS families such as lung cancer, lymphomas, colorectal, skin, stomach, ovarian cancers ((Ruijs et al 2010) ;p53-iarc.fr). These tumours are generally diagnosed at an earlier age than the average age of onset in the general population (Gonzalez et al 2009, Palmero et al 2010, Ruijs et al 2010).

Mutation spectrum

Analyses of segregation have demonstrated that the LFS families' segregation pattern is explained by a hereditary cancer predisposition due to dominant autosomal model (Blattner et al 1979, Lynch et al 1978, Strong et al 1989). To date, the only genetic abnormality associated with LFS IS germline mutations of *TP53* (Malkin et al 1990, Srivastava et al 1990). Such mutations are found in 56-73% of LFS patients using the strict definition for LFS (Bougeard et al 2008, Gonzalez et al 2009, Ruijs et al 2010). In contrast, only 20 to 40% of the LFL families carry a *TP53* germinal mutation (Tabori and Malkin 2008). Recent studies on large cohorts of patients have observed that the *TP53* mutation are found in 20 to 36% in families when the Chompret criteria are applied and 21% in families with the new definition of the Chompret criteria (Bougeard et al 2008, Gonzalez et al 2009, Ruijs et al 2010). A study in the USA has found that 95% of patients with a *TP53* germline mutation correspond to the strict LFS criteria or the Chompret definition (Gonzalez et al 2009).

As in sporadic cancers, more than 80% of the *TP53* mutations found in LFS patients are missense or non-sense and are preferentially G>C or A>T transversions (p53.iarc.fr). Some *TP53* germline mutations have not been reported as somatic mutations. For instance, the R337H mutation found essentially in Brazil is not a common mutation in sporadic cancer cases (Ribeiro et al 2001). The impact of the R337H mutation in LFS/LFL and in p53 functions will be described later, in the Introduction, Part I, section 5. In contrast, some mutations commonly found in sporadic cancer have never been identified in the germline of LFS patients. This is the case, for example, for the R249S mutation associated with hepatocellular carcinoma in a context of exposure to aflatoxin.

2. Penetrance and prevalence

LFS is a highly penetrant syndrome. Carriers of a germline *TP53* mutation have, on average, a 50% chance of developing a cancer before the age of 40 years compared with 1% of the general population, and 90% of *TP53* mutation carriers are diagnosed with cancer by the age of 60 (Birch et al 1998, Strong et al 1992). In a retrospective study of 200 *TP53* germline mutation carriers, 15% developed a second cancer, 4% a third and 2 developed a fourth cancer, with the highest risk of additional cancers in survivors of those who had had childhood malignancies (Hisada et al 1998).

The prevalence of the *TP53* mutation carriers was defined for different cancer of the LFS tumour spectrum using sequencing of exon 4 to 8 only. For example, *TP53* germline mutations are found in between 50 and 80% of ADR in children (cohorts of 6, 14 and 21 patients respectively) (Gonzalez et al 2009, Varley et al 1998, Wagner et al 1994) and 3.9% of ADR developed during adulthood (cohort of 103 patients) (Herrmann et al 2012). There is no obvious association between specific mutation and clinical cancer patterns, in contrast to observations for specific somatic mutations such as R249S in liver cancer.

3. Genetic anticipation in LFS families

Genetic anticipation is characterised by a decreasing age at onset and/or increasing severity of symptoms in successive generations within a pedigree (Tabori and Malkin 2008, Trkova et al 2002). This phenomenon has been observed in LFS families. The mechanism of genetic anticipation in LFS is still under debate. Trkova and collaborators have observed a shorter telomere length in *TP53* mutation carriers compared to the general population (34% decrease in children and 19% decrease in adults) (Trkova et al 2007). Tabori and co-workers have also observed that an earlier age at diagnosis is associated with a shorter telomere length and that *TP53* mutation carriers who developed cancer have shorter telomeres than non-*TP53*

mutation carriers and *TP53* mutation carriers without cancer (Tabori et al 2007). Another mechanism that may contribute to anticipation is DNA copy number variations (CNVs). The accumulation of rare CNVs in patients carrying a p53 mutated DBD may contribute to reported anticipation and the severity of the syndrome (Silva et al 2012). A recent study has observed that 4 rare CNVs, which play a role in chromatin remodelling, are present in 4 patients without deleterious *TP53* mutations and which develop brain cancer (Aury-Landas et al 2013). They have analysed the function role of one of these CNVs, a duplication in *SIRT3*, and suggest that this CNV could play a role in tumour development. This study suggests a deleterious role of constitutive alterations in chromatin remodelling genes in the genetic predisposition of cancer and in particular in brain tumours.

4. Diagnosis and treatment

The detection of tumours in LFS patients is difficult due to the large tumour spectrum associated with the syndrome and the many localisations of these tumours, especially for soft tissue sarcomas. To date, no method allows predicting the cancer type and the age of onset. The analysis of different parameters such as types of mutations, modifier factors or genetic instability, could be used to stratify cancer risks and propose an individual screening strategy (Tabori and Malkin 2008).

In *TP53* mutation carriers, cancers are generally treated by the standard-of-care therapies used for each tumours type. However, these patients present an abnormal radio-sensitivity increasing the risk of the development of a second tumour in the irradiated zones (Chompret et al 2000, Limacher et al 2001, Salmon et al 2007) and as such radiotherapy should be avoided (Kleinerman 2009). For example, a bilateral mastectomy is proposed to women who developed breast cancer to avoid radiotherapy (Heymann et al 2010). In addition, some *TP53* missense mutations are associated with the induction of chemotherapy or radiotherapy resistant tumours (Camplejohn et al 1995, Delia et al 1997, Pepper et al 2003). Thus one of the main objectives for treatment of these individuals has been to develop new drugs to target directly p53 or the p53 pathway. However, after 30 years of research, only 15 molecules have been developed for use in clinical research (Stegh 2012). Five categories of drugs have been developed: (1) genetic therapy using adenovirus to introduce a WT *TP53* in p53-null cells; (2) Mdm2 inhibition using anti-sense oligonucleotides; (3) Mdm2 inhibition using small molecules; (4) vaccination and (5) small molecules targeting mutant p53 variant to restore WT conformation and its trans-activation activity.

5. A special case: Li-Fraumeni syndrome in Brazil and the R337H mutation

In 2001, Ribeiro and collaborators observed the R337H mutation in 35 of 36 ADR Brazilian patients (Ribeiro et al 2001). This mutation is located in the oligomerisation domain of p53 and is found essentially in Southern Brazil with a high frequency: between 93 and 95% of ADR Brazilian patients carry the R337H mutation, 63 to 69% of choroid plexus carcinomas, 7.3% of osteosarcomas; 5.4% of breast cancer and 0.5% of breast cancer before the age of 40 (Custodio et al 2011, Custodio et al 2013, Giacomazzi et al 2013, Gomes et al 2012, Seidinger et al 2011). The tumour spectrum of the LFS/LFL in Brazilian families associated with the R337H mutation is a large spectrum with different types of cancer compared to the "classical" LFS/LFL tumour spectrum: breast cancer (30.4% of R337H carriers develop a breast cancer); STS (10.7%); Brain cancer (10.7%) and ADR (8.9%) (Achatz et al 2007). However, it was noted that there was not always an association between the presence of the R337H mutation and the occurrence of ADR (Sandrini et al 2005). Recently, Giacomazzi and co-workers have identified a R337H homozygotes carrier (Giacomazzi et al 2013). This mutation is found at a frequency 10-20 times higher than other *TP53* mutation in Southern Brazil (Palmero et al 2008) with a frequency of 0.3% in the Brazilian general population and has also been reported in Germany (**Figure 14**) (Achatz et al 2009, Herrmann et al 2012, Palmero et al 2008, Pinto et al 2004). Indeed, the low penetrance of this mutation may have contributed to the maintenance of this mutation in a large population (Garritano et al 2010, Palmero et al 2008). It is considered as a founder mutation and its distribution in Brazil appears to follow a road axis historically used as the main route by merchants of Portuguese origin in the XVIII and XIX century (Garritano et al 2010). Another hypothesis for its maintenance in Brazilian population has been proposed. Silva and collaborators have observed a different magnitude of rare CNV accumulation in the R337H patients compared to patients carrying others *TP53* mutations (Silva et al 2012).

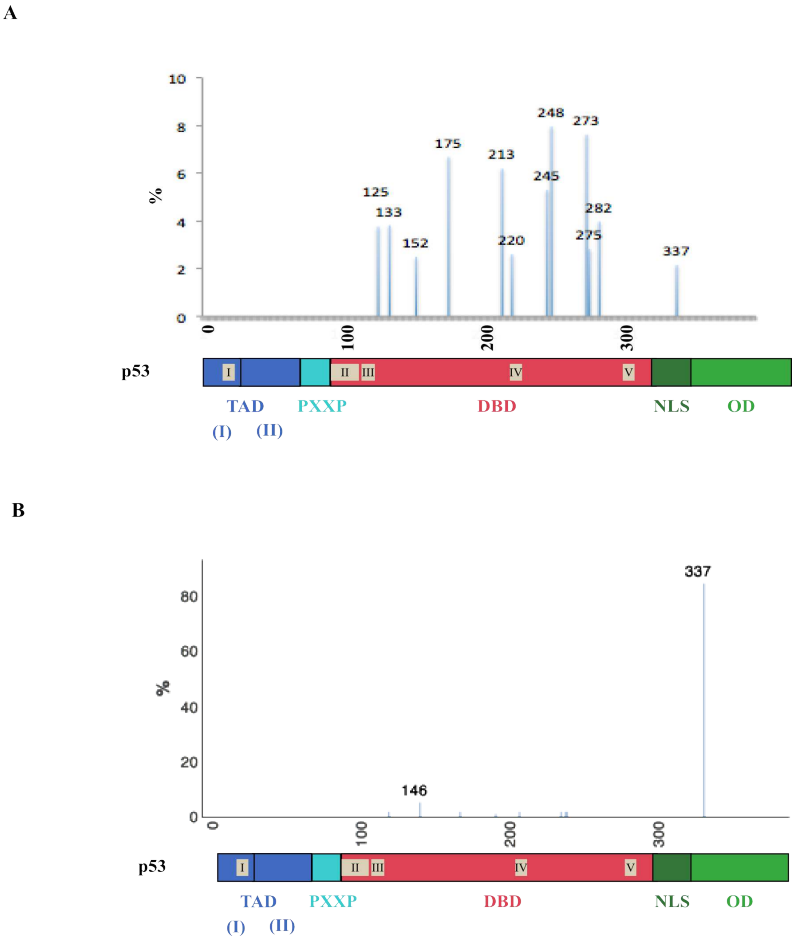


Figure 14: Codon distribution of *TP53* mutation in LFS. (**A**) In LFS patients described around the world except in Brazil. (**B**) In Brazilian LFS patients. Adapted from the IARC *TP53* database (p53.iarc.fr).

The impact of the R337H mutation on p53 suppressive function is due to its localisation in the oligomerisation domain of the protein in an aa important for the formation of the p53 dimer (Galea et al 2005). The mutant oligomerisation domain adopts a native-like structure, which is less stable than the corresponding WT domain (DiGiammarino et al 2002). The difference of stability of the variant protein is due to a high sensitivity to the cellular pH due to the presence of the histidine aa. The pH-dependent p53 dysfunction has been suggested to be the molecular basis for ADR in Brazilian children (DiGiammarino et al 2002). In addition, the presence of this mutation is associated with increased levels of oxidative damages irrespectively of the individual's previous cancer history while no oxidative damage was observed in subjects in a reference group (Macedo et al 2012). This observation suggests that the presence of oxidative damage in R337H patients is related to the carcinogenic process and is not a consequence of malignancy.

III. Polymorphisms in *TP53*

In the *TP53* gene, 85 polymorphisms have been described, including single nucleotide polymorphisms (SNPs) and deletions/insertions of several base pairs and are documented in the IARC *TP53* database (**Figure 15**) (p53.iarc.fr; (Petitjean et al 2007)). Many of these polymorphisms show variations in frequency depending on geographic and population factors (Garritano et al 2010, Whibley et al 2009). More than 90% of the *TP53* polymorphisms are located in non-coding sequences and it has long been presumed that they may have little impact on cancer susceptibility. So far, studies on *TP53* polymorphisms in relation with cancer risk have shown inconsistent results across studies (Garritano et al 2010, Stacey et al 2011, Whibley et al 2009). Moreover, at the functional level, few of the polymorphisms have been related to altered biological and/or biochemical functions of the p53 protein.

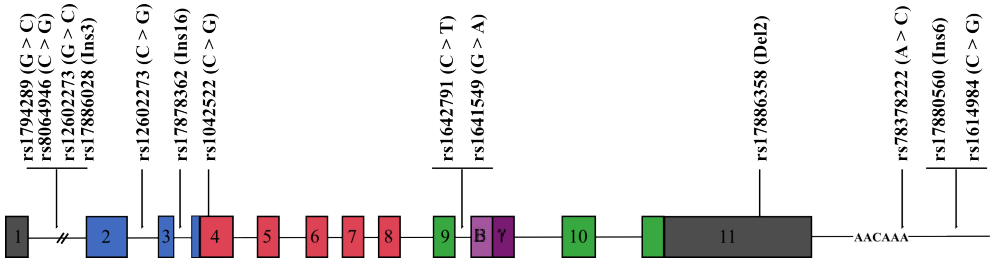


Figure 15: The most frequent polymorphisms found in the *TP53* gene and in its 3'flanking region. The nucleotide variation and the rs number are indicated on a schematic representation of the *TP53* gene. Adapted from p53-iarc.fr and Thesis of Virginie Marcel, University Lyon I, 2009.

1. Polymorphisms located in *TP53* non-coding sequences

rs1642785 (*TP53* PIN2)

Rs1642785, also called *TP53* PIN2, is located in the intron 2 and is the substitution of a G into a C (**Figure 15**). Based on the NCBI dbSNP database (www.ncbi.nlm.nih.gov/SNP/), the C allele frequency varies depending on the population: in Caucasians the frequency of the C allele is around 22% while in the African population the frequency is 55% (Garritano et al 2010, Marcel et al 2009).

Five studies have been performed on the association between the rs1642785 genotype and cancer susceptibility. Two studies have shown no association; no association was found between the risk of cervical cancers and carriage of the rs1642785 polymorphism in Brazilian patients (Fernandes et al 2008) and on the age at first cancer diagnosis in a Brazilian LFS/LFL cohort (Marcel et al 2009). Jha and collaborators have observed that the G allele frequency is higher in gliomas than the C allele frequency (0.55 and 0.36 respectively) in Indian patients but this association was not associated with an increase of cancer risk (Jha et al 2011). The

two last studies have shown that rs1642785 is associated with an increased risk of cancer: associations were reported between the G allele and ovarian cancer in the Danish population (Hogdall et al 2003) and osteosarcomas in Caucasians with a 6.7 fold increase of cancer risk for patients carrying a C allele (Savage et al 2007). The association of rs1642785 and cancer susceptibility could confound by its linkage equilibrium with the rs1042522, a polymorphism located in the coding sequence of *TP53* gene, which will be detailed later (Garritano et al 2010).

rs17878362 (*TP53* PIN3)

The most studied intronic variation in *TP53* is a 16-base pair insertion/duplication in intron 3 called rs17878362 or *TP53* PIN3, consisting of one copy (A1 allele) or two copies (A2 allele) of the sequence ACCTGGAGGGCTGGGG (**Figure 15**) (Lazar et al 1993). The minor allele frequency (MAF) in the Caucasian population is around 10% while it is around 27% in the African population and 2% in Asians (Garritano et al 2010, Marcel et al 2009).

Several case-control studies have reported an increased risk of various cancer types associated with the rs17878362 A2 allele. For example, rs17878362 is associated with an increase of breast cancer (Costa et al 2008, Wang-Gohrke et al 2002) or colorectal cancers (Gemignani et al 2004, Perfumo et al 2006). A recent meta-analysis identified a significant increase in overall cancer risk of 1.14 (95% Confidence Interval (CI): 1.02-1.27) in A2A2 carriers (Hu et al 2010b). However, this conclusion was questioned due to apparent discrepancies between data selected for meta-analysis and the original publications (Lu et al 2011). In addition, the A2 allele is associated with a worse prognosis for individuals with lung cancers (Boldrini et al 2008) and is correlated with poor histological grade in breast cancers (Powell et al 2002).

The A2 allele of rs17878362 is associated with an increase of cancer susceptibility after X-rays exposure, suggesting a role of this polymorphism in p53 function (Hung et al 2006). At the mechanistic level, there is some evidences that rs17878362 has an impact on the p53 mRNA level: the A2 allele expresses less p53 mRNA compared to A1 allele, in a dose-dependent manner (cancer risk associated with the A2A2 genotype compared to the A1A1 is higher than the cancer risk associated with the A1A2 genotype compared to the A1A1 genotype) (Gemignani et al 2004). However, the precise mechanisms underlying an increased cancer risk associated with the rs17878362 A2 allele are not clearly understood.

Polymorphisms located in the 3' flanking region of the *TP53* gene

In the 3' flanking region, 15 polymorphisms are described in the NCBI dbSNP database such as rs17880560, which is an insertion/deletion of 6 bp (A1: one copy of the 6 bp (CACGGC); A2: two copies of these 6 bp) or rs1614984, a C to T transition (**Figure 15**). The most studied polymorphism in the 3'flanking region is rs78378222. This polymorphism consists of a transversion of an A to a C and occurs in the sole poly-adenylation signal of *TP53* (AATAAA>AATACA).

Five studies have recently been performed addressing the association between rs78378222 and cancer risk. Guan and collaborators have observed that rs78378222 is not associated with an increase of melanoma or lung cancer susceptibility (Guan et al 2013). Several studies have observed an increase of cancer risk with the C allele of rs78378222: in basal cell carcinoma, prostate cancer, glioma and colorectal adenoma (Stacey et al 2011); in glioma (Egan et al 2012, Enciso-Mora et al 2013); oesophageal squamous cell carcinoma (Zhou et al 2012) or squamous cell carcinoma of head and neck (Guan et al 2013). To understand the role of the C variant of rs78378222 cancer susceptibility, Stacey and co-workers have extracted pre-mRNA from blood and adipose tissue and observed that heterozygotes expressed less p53 pre-mRNA compared to the AA homozygotes (Stacey et al 2011). In heterozygotes, 73% of the mRNAs contain the WT A allele and 27% of the C variant. Taken together, these results suggest that the C variant impairs proper termination and poly-adenylation of the p53 pre-mRNAs. Recently, Li and collaborators have shown, using gene expression cassettes carrying the human *TP53* cDNA with the 3'flanking region, that the CC genotype presented a decrease of both p53 mRNA and protein compared to the AA genotype (Li et al 2013). In addition, they have also observed a decrease of cellular apoptosis with the CC gene expression cassette. These results support the hypothesis of different biological effect with different genotypes of rs17880560, with possible impact on cancer susceptibility.

2. Polymorphisms located in *TP53* coding regions

Synonymous polymorphisms

Of the 19 polymorphisms described in the *TP53* coding sequence, 8 are synonymous, i.e. the substitution of a nucleotide giving rise to no alteration in the encoded amino acid due to the redundancy of the genetic code (Whibley et al 2009). These polymorphisms thus do not have any direct effects on the p53 protein, but can have an impact of p53 mRNA expression by modulating (1) alternative splicing of pre-mRNA, (2) mRNA stability due to the

modification of the DNA binding sequence of some miRNAs and (3) Mdm2 specific binding of p53 mRNA. Indeed, the Mdm2 protein, which regulates p53 protein degradation, has been implicated in the translation of p53 mRNA (Candeias et al 2008). At codon 36, the G to thymine (T) transition (CCG to CCT) reduces p53-dependent apoptosis by decreasing the binding of Mdm2 with p53 mRNA, consequently reducing p53 levels.

Non synonymous polymorphisms: rs1042522 (*TP53* PEX4)

Probably the most studied *TP53* polymorphism is a SNP in exon 4 encoding an R or a P at aa 72 (rs1042522) also called *TP53* PEX4, G>C) (**Figure 15**) (Matlashewski et al 1987). The minor allele corresponds to the P form of the protein and its frequency in Caucasians is 0.21, 0.45 in Asians and 0.67 in Africans (Garritano et al 2010, Marcel et al 2009). A latitude gradient of the allele frequency was observed depending of the distance with the equator (Niger: 0.63), which decreases in the Arctic Circle (Sweden: 0.17) (Beckman et al 1994, Sjalander et al 1996). This distribution indicates the loss of the pressure of selection on the P allele in the Northern immigrate population. It has been suggested that these variations might be attributable to a role of this polymorphism in the p53 suppressive to UV damages by solar exposure. A recent study has proposed another explanation, linking the R variant of rs1042522 to low winter temperature and to increased fertility through up-regulation of Leukemia Inbibitory Factor (LIF), a key regulator of embryo implantation (Feng et al 2011, Jeong et al 2010, Kang et al 2009, Paskulin et al 2012).

Rs1042522 is located in the p53 proline domain and there is *in vitro* evidence that the R and P p53 protein variants differ in their biological activities. First, the role of this polymorphism in apoptosis was analyzed using p53-null cell lines transfected with constructs that encode either the R or P variant. The different studies have observed an increase of the expression of pro-apoptotic target genes and mitochondrial apoptosis genes, such as *NOXA*, *PUMA* or *PIG-3* suggesting that R is a greater inducer of apoptosis (Cattelani et al 2012, Jeong et al 2010, Zhu et al 2010) with an higher amount of mitochondria DNA damage upon rotenone stress (Altilia et al 2012). These results were confirmed in a mouse model that showed an increase of apoptosis in R mice but no difference in acute responses to UV radiation (Zhu et al 2010). In addition, the association between the R allele and better induction of apoptosis, compared to the P allele, has been confirmed by several observations: (1) a higher proportion of R protein than P protein, tends to relocate in the mitochondria (Dumont et al 2003) and (2) the iASPP protein, which inhibits p53-dependent apoptosis, fixes preferentially the P protein and reduces its role in apoptosis (Bergamaschi et al 2006). An *in*

vitro analysis has shown that the P variant protein enhanced senescence by promoting $p21^{WAF1}$ expression (Cattelani et al 2012). After a stress, cells expressing the P variant protein stop in the G1 phase of cell cycle arrest while apoptosis in favored in cells with the R variant protein (Sullivan et al 2004). The R variant protein is associated with a high frequency of aberrant cells and chromatid breaks in cells after stress, also suggesting a better cell cycle arrest in P cell lines (Litviakov et al 2010).

However, results from systematic studies and meta-analyses have failed to identify a consistent association with breast, gastric or lung cancer risk (Matakidou et al 2003, Zhou et al 2007). The results from individual case-controls studies showed, in general, no association between rs1042522 and cancer risk, however meta-analyses, which increase the number of individuals and thus the statistical power, showed that P variant is associated with an increased cancer risk. For example, the 5 meta-analyses on breast cancer and rs1042522 found an increase of cancer susceptibility associated with the heterozygous carriage of the rs1042522 variant allele. But some studies have shown no difference in cancer risk between individuals carrying the RR corresponding genotype compared to individuals with PP corresponding genotype and other studies an increase of cancer risk associated with the heterozygotes (Dai et al 2009, Li et al 2009, Matakidou et al 2003, Wang et al 2013a). These results suggest that the role of rs1042522 in cancer risk seems to be dependent of cancer type and maybe population type.

In the Li-Fraumeni syndrome, the results concerning the influence of the rs1042522 polymorphism in cancer susceptibility and age of onset are controversial. A first study has shown that the patients with a germinal *TP53* mutation with the R variant develop cancer 12 years before patients with the P variant (Bougeard et al 2006). These results suggest that this SNP can modulate the suppressive functions of p53, which are attenuated by the presence of the mutation. This is supported by the fact that mutant R proteins present a better interaction with the p73 protein compared to the mutant P proteins, which induces a better inhibition of p73, implicated in apoptosis induction after stress (Marin et al 2000). A more recent study did not observed a role of the rs1042522 polymorphism on the age of onset in *TP53* mutant cells (Marcel et al 2009).

3. Polymorphisms in the p53 pathway

Mdm2 SNP309 polymorphism (rs2279744)

Mdm2 is overexpressed in around 10% of cancers and no mutations were found in the *Mdm2* gene in different cancers (Jones et al 1998). However, the *Mdm2* promoter contains a

polymorphism, rs2279744, also called SNP309, which is a single substitution of a T to a G (Bond et al 2004). The T allele frequency was reported to be 0.58 and G 0.42. The polymorphism is located in a RE of the SP-1 transcription factor, which has a higher affinity for the G allele compared to the T allele. In addition, the oestrogen receptor also binds the *Mdm2* promoter in the region of the SNP309 (Hu et al 2009). It was found that the oestrogen receptor can induce *Mdm2* transcription and preferentially in cells with the SNP309 GG genotype. This increase of Mdm2 mRNA level might lead to a higher risk of hormone-related cancers.

Several studies have shown an earlier age of cancer onset or an increase of cancer susceptibility to breast cancer in premenopausal women with the Mdm2 SNP309G variant (Bond et al 2004, Bond et al 2006, Lum et al 2008). However, the role of this polymorphism in the increase of cancer risk is still controversial. An analysis on more than 8,000 cases/controls with breast, lung and colorectal cancers has shown that the SNP309 G variant is not associated with an increase of cancer risk but on the age of cancer onset (Wilkening et al 2007). The G variant was found to be associated with lung cancer risk in some studies (Lind et al 2006) but not in others (Hu et al 2006), however, a recent meta-analysis has shown that the SNP309G variant allele is associated with a significant 1.14 fold increase of lung cancer risk (He et al 2012). Two meta-analyses on gastric cancers have also observed an increase of cancer risk associated with the SNP309G variant allele(Ma et al 2013, Tian et al 2013).

In LFS/LFL patients, several studies have shown that the SNP309G variant is associated with a decrease of age of cancer onset in *TP53* mutation carriers without any difference in non-*TP53* mutation carriers (Bond et al 2004, Bougeard et al 2006, Ruijs et al 2007). However, three studies have contradicted these results because they have not observed any significant statistical difference on age of cancer onset between the different genotypes of SNP309 in *TP53* mutation carriers (Marcel et al 2009, Pinto et al 2009, Renaux-Petel et al 2013). A recent study has observed that the haplotype of *MDM2* SNP285G, a modulator of SP1 binding to Mdm2 promoter and an antagonist of the SNP309G variant, and SNP309G is the higher risk haplotype (5 years earlier onset of cancer) in *TP53* mutation carriers (Renaux-Petel et al 2013). This result suggests that the SNP309G variant could be deleterious when it is not associated with the SNP285C variant.

Polymorphic p53 binding sites

Polymorphisms have been described in the promoters of p53 target genes in p53 REs and can modulate the p53's transactivation activity. Using an *in silico* analysis, Bandele and co-workers have identified 6,538 polymorphisms in the NCBI dbSNP database that were located in putative p53 REs and were predicted to alter p53 DNA interactions (Bandele et al 2011). They also identified 33 polymorphic sequences in experimentally validated p53 binding sites. They confirmed the influence of 32 polymorphisms using a ChIP-seq analysis after Doxorubicin exposure. These SNPs alter gene expression and cellular responses and increase disease susceptibility with cell type specificity. They have found that the C and G residues in the p53 RE are the most conserved and important nucleotides for the p53 DNA binding activity.

Part II. p53 isoforms

This part of the manuscript is more detailed in a review entitled "Biological functions of p53 isoforms through evolution: lessons from animal and cellular models" published in Cell Death and Differentiation (Cell Death Differ. 2011 Dec;18(12):1815-24).

Generation of Human p53 isoforms by Diverse Regulatory Mechanisms

The *TP53* gene is expressed as several protein isoforms with variations in the N- and C- terminal domains (**Figure 16**). *TP53* can encode four isoforms with different truncations in the N-terminal domain: the TAp53 (without deletion of aas), the Δ40p53 (deleted of the first 39 aas, called ΔNp53 or Δ40p53 in the literature), the Δ133p53 (deletion of the first 132 aas) and the Δ160p53 (deleted of the first 159 aas). The production of these N-terminal isoforms is based on different transcriptional and translational mechanisms. The *TP53* gene can also encode also three isoforms with different splicing of the C-terminal domain with partial retention of intron 9 and addition of new aas: the α forms (without deletion of aas), the β forms (addition of 10 new aas) and the γ forms (addition of 15 new aas). The controls of the alternative splicing of intron 9 in still unknown. The TAp53α isoform corresponds to the canonical p53 protein.

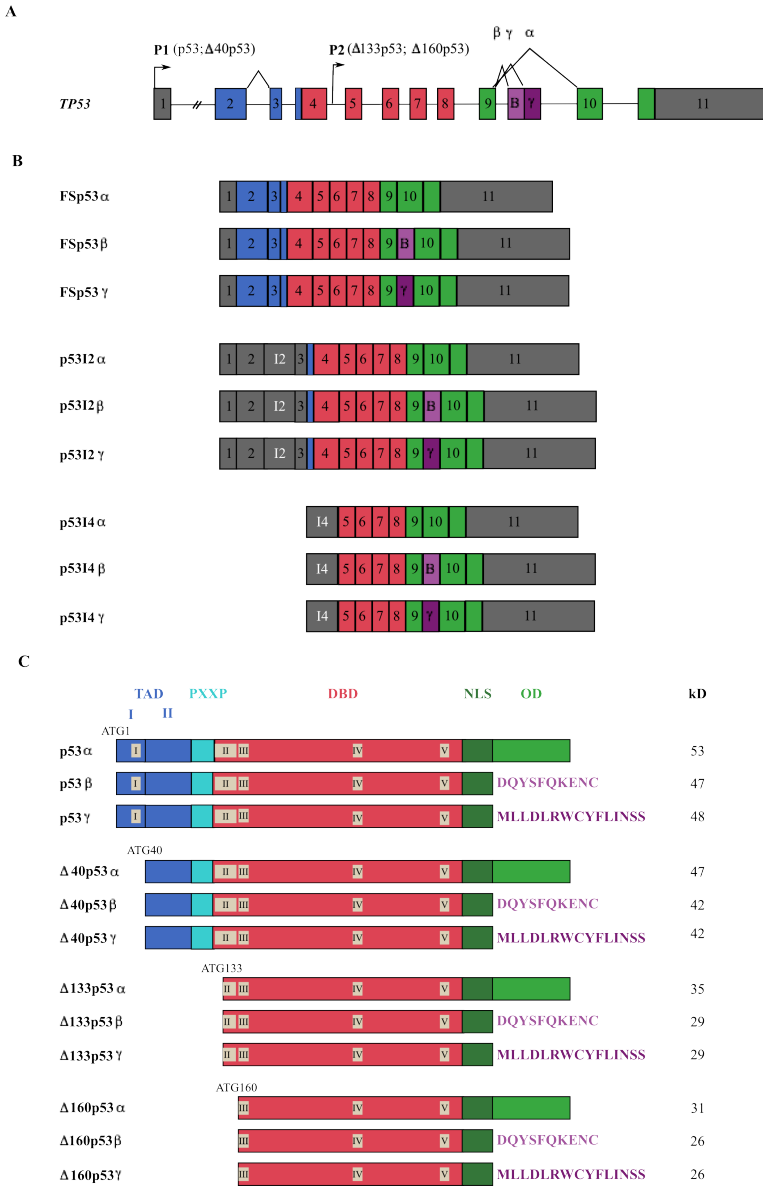


Figure 16: Schematic representation of human p53 isoforms. (**A**) Human *TP53* gene structure. The *TP53* gene, which consists in 11 exons (coloured boxes: coding exons; grey boxes: non-coding exons), is expressed as several p53 isoforms through the usage of alternative promoters (↱), splicing sites (^) or translational initiation sites (|). The proximal promoter P1, located upstream exon 1, regulates the transcription of two transcripts: the fully-spliced p53 mRNA (FSp53: all exons correctly spliced), which encodes both p53 (from ATG1) and the Δ40p53 isoforms (from ATG40); and the p53I2 (or p53EII) mRNA, retaining the entire intron 2 by alternative splicing, that generates Δ40p53 isoform from ATG40, due to the presence of stop codons in the reading frame issued from ATG1. The internal P2 promoter described as encompassing region from intron 1 to exon 5, produces p53I4 mRNA, initiated in intron 4 and encoding the N-terminal Δ133p53 (from ATG133) and Δ160p53 forms (from ATG160). Three different C-terminal p53 forms have been described due to alternative splicing of intron 9: the α forms resulting from the excision of the entire intron 9; the β and γ forms produced by retention of two small parts of intron 9. Some *cis*- and *trans*-regulators driving p53 isoforms expression have been described (purple boxes). (**B**) Human p53 protein isoforms. The classical p53 protein contains a trans-activation domain (TAD, dark blue), a proline-rich domain (PXXP, light blue), a DNA-binding domain (DBD, red) and an oligomerisation domain (OD, dark green) that encompass a nuclear localisation domain (NLS, light green) and five regions conserved through evolution (I to V in grey boxes). Compared to p53, Δ40p53 forms lack the first trans-activation domain, while Δ133p53 and Δ160p53 isoforms lack the entire trans-activation domain and parts of the DNA-binding domain. In C-terminal, the α peptide corresponds to the oligomerisation domain that is replace by new aas, the β and γ peptides (brown). Theoretical molecular weight of p53 protein isoforms are indicated on the right as well as their different names used in the literature.

Biological Functions of human p53 Isoforms

Compared to p53, its isoforms conserve the DNA binding activity but the N-terminal isoforms have lost the transcriptional and conserved the oligomerisation activity and the C-terminal isoforms have lost the oligomerisation and conserved the transcriptional activity. The N-truncated forms act as inhibitors of p53 transcriptional and suppressive activity. For example, the Δ133p53 can play a role in cell cycle arrest and apoptosis by modulating the gene expression of some p53 target genes such as $p21^{WAF1}$, *MDM2* or *BCL-2*. Studies on p53β have shown that this isoform enhances the p53 transcriptional activity on $p21^{WAF1}$ promoter.

Animal Models: clues to the Physiological Significance of Isoforms

The formation of isoforms from the *TP53* gene is conserved though evolution from *Drosophila melanogaster* to *mouse*. The p53 isoforms are analysed in some animal models to understand their functions in cells. In animal models, the p53 isoforms expressions and activity seem to be well conserved throughout evolution. For example, in *Drosophila melanogaster*, the DΔNp53 isoform, equivalent to the human Δ133p53, controls apoptosis or in *Danio renio*, the ZΔ133p53 isoform (ie human Δ133p53) can modulate the transcription of some p53 target genes in cell cycle arrest such as *Cyclin G1* or $p21^{WAF1}$ or in apoptosis such as *BAX* or *REPIMO*.

p53 isoforms and Human Cancers

The first analyses on the impact of p53 isoforms on cancers risk have shown that some polymorphisms located between exon 3 to intron 4 can modulate the basal activity of the internal *TP53* promoter. These results suggest that these polymorphisms may influence the basal level of Δ133p53 and may modulate the cancer risk. In addition, some studies have shown the presence of mutations that affect the production of isoforms. An over-expression of p53 isoforms can also modulate the cancer risk such as *p53β* mRNA over-expression is associated with poorly differentiated ovarian cancers.

Cell Death and Differentiation (2011) 18, 1815–1824
© 2011 Macmillan Publishers Limited All rights reserved 1350-9047/11
www.nature.com/cdd

Review

Biological functions of p53 isoforms through evolution: lessons from animal and cellular models

V Marcel[1], M-L Dichtel-Danjoy[2], C Sagne[3,4,5], H Hafsi[3], D Ma[2], S Ortiz-Cuaran[3], M Olivier[3], J Hall[4,5], B Mollereau[2], P Hainaut[3] and J-C Bourdon*[,1]

The *TP53* tumour-suppressor gene is expressed as several protein isoforms generated by different mechanisms, including use of alternative promoters, splicing sites and translational initiation sites, that are conserved through evolution and within the *TP53* homologues, *TP63* and *TP73*. Although first described in the eighties, the importance of p53 isoforms in regulating the suppressive functions of p53 has only become evident in the last 10 years, by analogy with observations that p63 and p73 isoforms appeared indispensable to fully understand the biological functions of *TP63* and *TP73*. This review summarizes recent advances in the field of 'p53 isoforms', including new data on p63 and p73 isoforms. Details of the alternative mechanisms that produce p53 isoforms and *cis*- and *trans*-regulators identified are provided. The main focus is on their biological functions (apoptosis, cell cycle, aging and so on) in cellular and animal models, including mouse, zebrafish and *Drosophila*. Finally, the deregulation of p53 isoform expression in human cancers is reviewed. Based on these latest results, several developments are expected in the future: the identification of drugs modulating p53 isoform expression; the generation of animal models and the evaluation of the use of p53 isoform as biomarkers in human cancers.
Cell Death and Differentiation (2011) **18,** 1815–1824; doi:10.1038/cdd.2011.120; published online 23 September 2011

Thirty years of research have shown that the p53 tumour-suppressor protein, encoded by *TP53* gene (OMIM 191170), integrates endogenous and exogenous signals to modulate cell fate to stress and cellular environments.[1] In addition, it has emerged that p53 is more than just a 'stress response' factor as it regulates embryo implementation.[2] The ability of p53 to integrate signals implies the existence of multiple and subtle levels of regulation. Over the years, transcriptional, translational and post-translational regulatory mechanisms have been uncovered.[3,4] This biochemical diversity echoes the genetic diversity of the *TP53* locus, which contains multiple genetic polymorphisms defining over 100 distinct *TP53* haplotypes.[5] Recently, an additional layer of regulatory mechanism has emerged through identification of p53 isoforms, which are physiological proteins expressed in normal cells from the *TP53* gene owing to the use of alternative promoters, splicing sites and/or translational initiation sites.[6,7]

The p53 isoforms were first identified in early studies investigating p53 expression patterns. In 1984, Matlashewski *et al.* cloned an N-terminal variant of the human p53 mRNA, whereas in 1985, Rotter and co-workers detected an alternatively spliced C-terminal variant of mouse p53, latter isolated in human cells.[8–10] However, the 'p53 isoform' field has only really emerged in the past 10 years, when it became clear that *TP53* retained the elaborate patterns of isoform expression that characterizes its homologues, *TP63* and *TP73*. The rapid accumulation of descriptive, functional and clinical data on p53 isoforms has led to the emergence of a research community, which held its First International Meeting at the International Agency for Research on Cancer in Lyon, France, in September 2010.[11] This review provides a brief survey of the 'p53 isoforms' field at a time when it is emerging at the forefront of p53 research.

TP63 and *TP73*: the Isoform Paradigm

The two p53-related proteins, p63 and p73, share strong structural, biochemical and biological homologies.[12,13] In particular, they bind specifically to DNA onto conserved p53 response elements (p53REs) by using their DNA-binding domain. In the late nineties, the cloning of *TP63* and *TP73* revealed an elaborate pattern of mRNA expression resulting in several protein isoforms.[14,15] Several N-terminal forms, produced by the use of alternative promoters and/or alternative splicing (i.e., TA forms, which contains the transactivation domain (TAD), *versus* ΔN forms, produced from an

[1]Centre for Oncology and Molecular Medicine, INSERM-European Associated Laboratory, University of Dundee, Ninewells Hospital, Dundee, Scotland DD1 9SY, UK; [2]Apoptosis and Neurogenetics Group, Ecole Normale Supérieure de Lyon, Laboratory of Molecular Biology of the Cell, CNRS UMR5239, University of Lyon, UMS3444 Biosciences Lyon Gerland, 46 allée d'Italie, Lyon 69007, France; [3]Molecular Carcinogenesis Group, International Agency for Research on Cancer, Lyon Cedex 08, France; [4]INSERM U612, Bât. 110-112, Centre Universitaire, Orsay 91405, France and [5]Institut Curie Centre de Recherche, Bât. 110-112, Centre Universitaire, Orsay 91405, France

*Corresponding author: J-C Bourdon, Centre for Oncology and Molecular Medicine, INSERM-European Associated Laboratory, University of Dundee (Mailbox 4), Ninewells Hospital, Medical School, Dundee, Scotland DD1 9SY, UK. Tel: +44 (0) 138 249 6401; Fax: +44 (0) 138 249 6363; E-mail: j.bourdon@dundee.ac.uk
Keywords: tumour suppressor; p53; p63; p73; p53 family; isoforms
Abbreviations: MEFs, mouse embryonic fibroblasts; TAD, transactivation domain; p53REs, p53 response elements
Received 19.5.11; revised 04.8.11; accepted 12.8.11; Edited by G Melino; published online 23.9.11

internal promoter resulting in the presence of a different TAD), were found combined with several C-terminal forms generated by alternative splicing (five for p63: α to ε; seven for p73: α to η).[13] Far from being of minor component, the ΔNp63 and ΔNp73 forms are the major forms expressed in certain cell types.[14,16] In mice, knockout of the entire *TP63* or *TP73* loci (targeting all isoforms) revealed the roles of p63 and p73 in epithelial differentiation and neuronal development, respectively,[15,17] whereas no impact was observed in stem cell commitment.[18] However, with the generation of isoform-specific knockout mice, a subtle interplay between the N-terminal isoforms has recently emerged, with the dynamic expression of N-terminal p63 or p73 isoforms appearing critical for maintaining the normal sequence of cell development (from stem to committed progenitors and then differentiated cells).[19,20] Building on this idea, Aberdam and co-workers analysed the impact of ΔNp63 and TAp63 isoforms in cellular commitment. In murine embryonic stem cells, ΔNp63, but not TAp63, is highly expressed during epidermal commitment and is critical for the expression of the cytokeratins K14 and K5, two markers of keratinocyte proliferation, indicating that only ΔNp63 is required for the commitment of ectodermal into epidermal cells.[21,22] Mills *et al*. observed that ΔNp63α overexpression in mouse embryonic fibroblasts (MEFs) resulted in the bypass of Ras-mediated senescence and enhanced carcinoma development in mice, suggesting that ΔNp63α inhibits senescence and therefore acts as an oncogene.[23] By contrast, overexpression of TAp63 forms in $p53^{-/-}$ MEFs increased senescence and reduced tumour development *in vivo*, consistent with a p53-independent effect.[24] Studying mice deficient for specific p73 protein isoforms, Mak, Melino and co-workers revealed that, like p63, the different p73 isoforms had dual cellular roles. In particular, $TAp73^{-/-}$ mice are spontaneously tumour-prone whereas $\Delta Np73^{-/-}$ cells show impaired tumour formation in nude mice.[25,26] In addition to these cancer-related effects, knockout of either TAp73 or ΔNp73 isoforms resulted in isoform-specific defects in neurogenesis or neurodegeneration, respectively. Overall, these studies show that each N-terminal p63/p73 form has specific roles in regulating distinct cell differentiation pathways. Moreover, they demonstrate that the ΔN forms, by preventing senescence and maintaining progenitor cell status, may act as oncogenes, whereas TA forms, through their capacity to switch cells into a post-mitotic state, may act as tumour suppressors.

Generation of Human p53 Isoforms by Diverse Regulatory Mechanisms

Like *TP63* and *TP73*, the human *TP53* gene encodes several p53 protein isoforms through conserved mechanisms.[6,7] The main and most abundant p53 isoform is the canonical p53 protein, also termed TAp53α, as it contains an entire TAD and the longest C-terminal domain (Figure 1). In addition to the TA forms, three ΔN forms have been identified that differ by their translation initiation site. This is used to designate them as Δ40p53, Δ133p53 and Δ160p53. The four N-terminal p53 forms can be combined with three different C-terminal domains (α, β, γ). Recently, some *cis*- and *trans*-regulators have been identified as specific modulators of p53 isoform expression (Figure 1).

Δ40p53 expression: one form, several mechanisms. Compared with p53, the human N-terminal Δ40p53 forms lack the first 39 amino acids corresponding to the main TAD (Figure 1c). Matlashewski *et al.*[27] identified a p53 mRNA species retaining the entire intron-2 (p53I2), indicating that an alternative splicing event leading to the retention of the *TP53* intron-2 can occur (Figure 1b). They later observed that stop codons in intron-2 of the p53I2 mRNA prevent p53 expression from the first AUG.[28] However, p53I2-transfected cells were found to express a 45-kDa protein, undetectable using antibodies recognizing the p53-TAD epitopes (DO1 or DO7), that corresponds to Δ40p53 initiated at a second AUG at codon-40, encompassed within a strong Kozak consensus.[29] Alternative splicing of intron-2 can be regulated through *G-quadruplex* structures located in intron-3 of the p53 pre-mRNA.[30] Using reporter assays and RNA–*G-quadruplex* ligands, it appears that *G-quadruplex* structures promote the correct splice-out of intron-2, leading to the fully spliced p53 (FSp53) mRNA encoding the full-length p53 protein; *G-quadruplex* disruption however favours the retention of intron-2 and thus p53I2 mRNA expression. This observation is the first clue that the *TP53* sequence itself can modulate its own isoforms' expression through regulation of alternative splicing.

In addition to alternative splicing, Δ40p53 forms can be encoded from the FSp53 mRNA through an internal initiation of translation at codon-40.[29] Two internal ribosomal entry sequences (IRES) have been identified that regulate the translation of either p53 or Δ40p53 (Figure 1b).[31,32] However, the relative contribution, *in vivo*, of each of these mechanisms, which are producing Δ40p53, remains to be fully established.

Production of Δ133p53 and Δ160p53 forms from the internal P2 promoter. As for *TP63* and *TP73*, *TP53* contains an internal promoter that controls the expression of two N-terminal forms (Figure 1a).[33] In addition to the proximal P1 promoter regulating p53 and Δ40p53 expression, an internal P2 promoter located between intron-1 and exon-5 regulates the transcription of p53 mRNAs initiated in intron-4 (p53I4) (Figure 1b).[33] This internal P2 promoter is different from the P* promoter in *TP53* intron-1 identified by Reisman *et al.*[34,35] that regulates the expression of an unrelated p53 transcript encoded by *TP53* intron-1. Site-directed mutagenesis and siRNA methods revealed that translation of p53I4 mRNAs can be initiated at two distinct codons, AUG133 and AUG160, leading to the expression of the Δ133p53 and Δ160p53 proteins, respectively, that lack the TAD and part of the DNA-binding domain (Figure 1c).[36] Surprisingly, Δ160p53 forms are expressed in K562 cells, which are considered as 'p53-null' cells. The *TP53* mutation in K562 cells results in a premature stop codon between AUG133 and AUG160, thus preventing the expression of the TA, Δ40 and Δ133 forms without compromising the Δ160p53 reading frame. Thus, it should be kept in mind that some cells or tumours considered as 'p53-null' because of the presence of frameshift or nonsense mutations, may retain the capacity to express one or several p53 isoforms.[29,36]

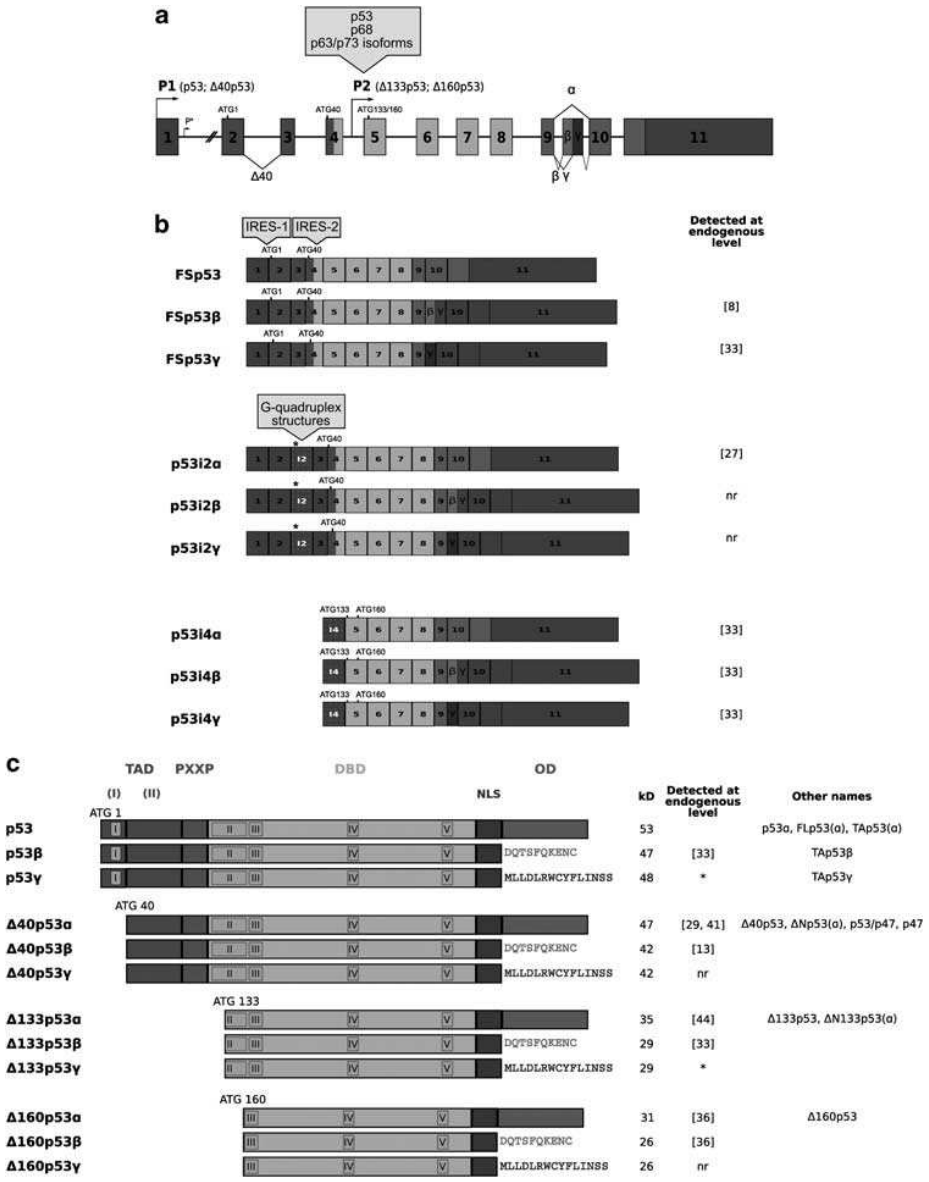


Figure 1 A schematic representation of human p53 isoforms. (**a**) The human *TP53* gene structure. The *TP53* gene, which consists of 11 exons (coloured boxes, coding exons; grey boxes, non-coding exons), expresses several p53 isoforms owing to usage of alternative promoters (┌→), splicing sites (^) or translational initiation sites (I). (**b**) Human p53 mRNA variants. The proximal promoter P1, located upstream from exon-1, regulates the transcription of two transcripts: the fully spliced p53 mRNA (FSp53), which encodes both p53 (from ATG1) and Δ40p53 forms (from ATG40), and the p53I2 mRNA, retaining the entire intron-2 by alternative splicing, which generates Δ40p53 forms from ATG40, owing to the presence of stop codons (*) in the reading frame starting from ATG1. The internal P2 promoter, described as encompassing the region from intron-1 to exon-5, produces p53I4 mRNA, initiated in intron-4 and encoding the N-terminal Δ133p53 (from ATG133) and Δ160p53 forms (from ATG160). Three different C-terminal p53 forms have been described owing to alternative splicing of intron-9: the α-forms resulting from the excision of the entire intron-9, and the β- and γ-forms produced by retention of two small parts of intron-9. Some *cis*- and *trans*-regulators driving p53 isoform expression have been described (purple boxes). Endogenous expression of most of the p53 mRNA variants in human cells has been reported (references shown in parentheses). Grey box, non-coding sequence; NR, not yet reported. (**c**) Human p53 protein isoforms. The canonical p53 protein contains a TAD (blue), a proline-rich domain (PXXP, purple), a DNA-binding domain (DBD, orange) and an OD (green) that encompasses a nuclear localization domain (NLS, green) and five regions conserved through evolution (I–V in grey boxes). Compared with p53, the Δ40p53 forms lack the first TAD, whereas the Δ133p53 and Δ160p53 isoforms lack the entire TAD and parts of the DBD. At the C-terminal, the α-peptide corresponds to the OD that is replace by new residues, the β- and γ-peptides (brown). On the right is indicated the theoretical molecular weight, the detection at endogenous levels (reference in brackets) as well as the different names of the isoforms used in the literature. The color reproduction of this figure is available at the *Cell Death and Differentiation* journal online

Experimental studies showed a p53-dependent regulation of Δ133p53 expression.[37,38] The internal P2 promoter contains p53REs located at the junction of exon-4/intron-4. Promoter deletion, site-directed mutagenesis and chromatin IP experiments demonstrate that the direct binding of p53 onto the p53REs results in an increased expression of both Δ133p53 mRNAs and the Δ133p53α protein. Some experimental evidence suggests that protein expression of

Δ133p53α can be induced through p53-independent mechanisms. In particular, the accumulation of the Δ133p53α protein was observed in response to the knockdown of p68, a DEAD Box helicase involved in multiple transcriptional regulatory processes, and more recently in response to the expression of some p63 or p73 isoforms.[39]

The C-terminal spliced p53 forms, β and γ. The alternative splicing of intron-9 of human *TP53* produces three different C-terminal p53 forms (α, β and γ) (Figure 1a).[8,33] Complete excision of intron-9 results in the expression of the α-forms corresponding to the classical p53 C-terminal domain (oligomerization domain, OD) (Figures 1b and c). On the other hand, partial retention of intron-9 generates the β- or γ-forms, in which the OD is replaced by 10 or 15 new amino acids, respectively. However, the mechanisms that control the alternative splicing of intron-9 are unknown.

The classical p53 C-terminal domain contains the main post-translational modification sites regulating p53 stability, such as the lysines residues ubiquitinated by Mdm2, an E3-ubiquitin ligase regulating p53 stability and activity.[3] Their absence in β- and γ-forms led to the investigation as to whether the stability of the p53β and p53γ proteins is regulated by the Mdm2–ubiquitin–proteasome pathway. To date, results remain controversial. For instance, whereas constitutively overexpressed FLAG-tagged p53β or p53γ do not appear to interact with Mdm2 or to be degraded in a proteasome-dependent manner,[40] Bourdon and co-workers observed that the endogenous p53β protein can be degraded by the proteasome in an Mdm2-dependent manner.

Biological Functions of Human p53 Isoforms

Based on early studies detailed below, the N-terminal isoforms lacking the TAD (i.e., Δ40p53, Δ133p53 and Δ160p53) were expected to only act as dominant-negative regulators of p53 activity. For instance, ectopic Δ40p53α expression downregulated the p53-induced transactivation on reporter genes and counteracted p53-dependent growth suppression in colony formation assays.[28,29] In addition, in human diploid fibroblast WI38 cells, expression of endogenous Δ40p53α increased during the G_1/S transition, in parallel with decreased expression of p21.[29] Furthermore, as Δ40p53α lacks the Mdm2-binding site, it escapes Mdm2-mediated degradation and does not accumulate in response to DNA damage, its expression persisting at low but stable amounts in many cell types.[29,41] These data support the notion that Δ40p53α inhibits basal p53 activities during cell-cycle progression. However, whether Δ40p53α can exert p53-independent effects is still unknown. Hainaut and co-workers observed that Δ40p53α contains an intact DNA-binding domain able to bind to p53REs *in vitro*. Thus, Δ40p53α may also exert intrinsic regulatory effects by competing with p53 for p53RE binding and thereby modulating their accessibility for other transcription factors, regulating the cell fate outcome depending upon cell type and cell context. In addition, Δ40p53α lacks the first TAD, which has been shown to be dispensable for p53 transcriptional activity, and retains the second TAD, which can regulate gene expression.[42,43] Therefore, with our current knowledge, one should be cautious in considering Δ40p53 as a simple dominant-negative inhibitor of p53.

Compared with Δ40p53α, the available data have clearly revealed that Δ133p53 controls p53 activity. In reporter assays, Δ133p53α can also behave as a dominant-negative inhibitor of p53.[13,33] Δ133p53α does not bind to consensus p53REs *in vitro*, consistent with its partial lack of the DNA-binding domain, and thus can also behave as a dominant mutant p53 protein.[38] Experimental studies suggested that, instead of being a strict dominant inhibitor, Δ133p53α is instead a fine modulator of p53's suppressive activity as its expression determines cell fate in response to stress. The knockdown of Δ133p53α expression promotes p53-mediated apoptosis and G_1 cell-cycle arrest in response to doxorubicin treatment, without altering the p53-dependent G_2 cell-cycle arrest.[37] These effects may be due to the ability of Δ133p53α to modulate gene expression in a promoter-dependent manner, as observed for $p21^{WAF1}$, *Mdm2* and *Bcl-2*.[37] Interestingly, Δ133p53 silencing has also been associated with replicative, but not oncogene-induced, senescence in normal human fibroblasts through transcriptional regulation of p53-target genes, including $p21^{WAF1}$ and *mir-34a*.[44] Overall, these results are consistent with an oncogenic capacity of Δ133p53.

The biological functions of the C-terminal p53 isoforms (i.e., p53β and p53γ) remain poorly described and controversial. Bourdon *et al*.[33] showed that, in the absence of stress, endogenous p53β bound to the *Bax* and $p21^{WAF1}$ promoters, but only weakly to that of *Mdm2*. Moreover, in luciferase reporter assays in the absence of stress, the co-expression of p53β and p53 enhanced the p53 transcriptional activity on the $p21^{WAF1}$ promoter but not on the *Bax* promoter, suggesting a promoter-dependent effect. These observations are consistent with the demonstration that p53β cooperates with p53 to accelerate senescence in human fibroblasts.[44] By contrast, experimental studies failed to observe binding of FLAG-tagged p53β or p53γ onto p53RE consensus and to show a role of FLAG-tagged p53β or p53γ in p53-dependent apoptosis or senescence in cells constitutively overexpressing FLAG-tagged p53β or p53γ, and selected to grow in presence of neomycin.[40] Thus, there is still debate on whether p53β or p53γ exert their activities in an autonomous manner or through an interaction with p53. Furthermore, there is no evidence of distinct biological activities for p53β or p53γ.

Overall, current experimental data on the biological roles of p53 isoforms are fragmented. Given that p53 isoforms differ from each other in the three functional domains (TAD, DNA-binding and OD), their potential to modulate p53-dependent responses is expected to be diverse and cell type-dependent. Further insight into which of these functions are of physiological or pathological relevance may come from animal model studies.

Animal Models: clues to the Physiological Significance of Isoforms

The simplest animal model to study p53 isoforms, *Drosophila melanogaster*. The diversification of the *p53* gene family into three members occurred in vertebrates.

Thus, invertebrates such as *Drosophila* contain a single p53-related gene, which encodes three protein isoforms (Figure 2): Dp53, corresponding to the human full-length p53; DΔNp53, a general counterpart of the human N-terminal p53 forms as it is encoded by an mRNA transcribed from an internal promoter (i.e., human Δ133) and contains a

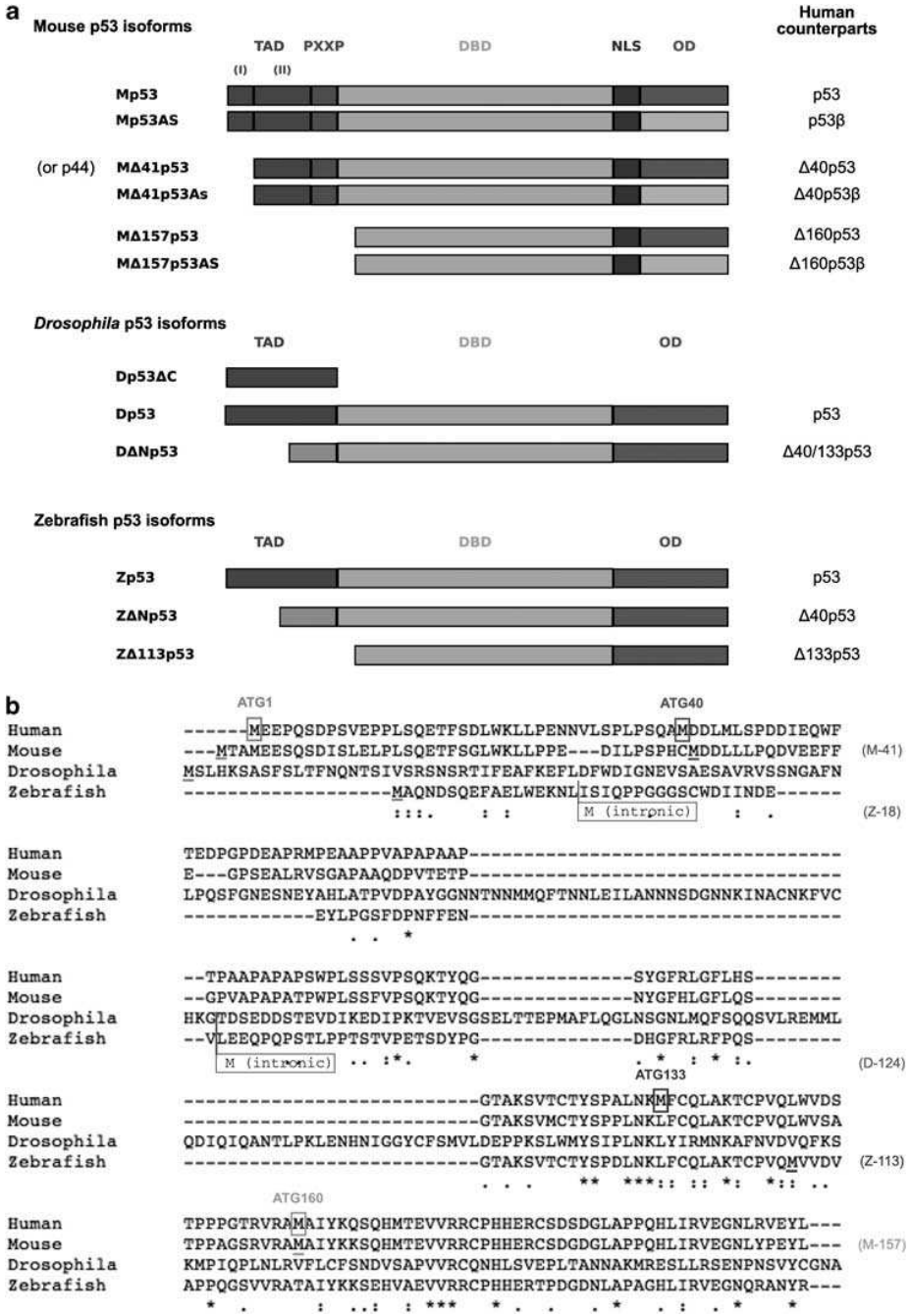


Figure 2 p53 isoforms in animal models. (**a**) Structural organization of p53 isoforms through evolution. Like humans, mouse, *Drosophila* and zebrafish express a full-length p53 protein, which conserves a TAD (blue), a DNA-binding domain (DBD, orange) and an OD (green). Only the mouse Mp53 protein presents a proline-rich domain (PXXP, purple) and a nuclear localization signal (NLS, green). In addition, all these animals express some p53 isoforms that have the same structural organization as the human p53 isoforms owing to the use of alternative promoters and splicing sites. M, mouse protein; D, *Drosophila* protein; Z, zebrafish protein; N-terminal p53 isoform identification, Δ forms; N-terminal p53 isoform denomination, codon number, when initiated ATG occurs in the coding sequence, or N, when initiation occurs in a non-coding sequence; C-terminal p53 isoform identification, AS (alternative splicing, green boxes); grey box, different residues compared with the full-length p53 protein. (**b**) Localization of translation initiation sites in animal p53 sequences. Red, ATG1 generating the full-length p53 protein; green, the methionine used to produce the homologues to the human Δ40p53 forms; blue, the methionine used to produce the homologues to the human Δ133p53 forms; orange, the methionine used to produce the homologues to the human Δ160p53 forms. The color reproduction of this figure is available at the *Cell Death and Differentiation* journal online

Table 1 Available p53 animal models

Species	Name	Modulation of p53 isoforms	Ref.
Drosophila		Overexpression of DΔNp53 in cells	45
		Overexpression of transgenic DΔNp53 under the control of tissue-specific promoters	47,49–53
		Overexpression of transgenic dominant-negative mutant DΔNp53	53,54
		Inactivation of *Dp53* locus	49,53
Zebrafish	Tg(Δ113p53:gfp)	Overexpression of transgenic ZΔ113p53-GFP morpholino (Δ113p53-MO)	61
		Overexpression of ectopic ZΔNp53	59
Mouse	MΔ41p53/MΔ41p53 wt/MΔ41p53	Overexpression of transgenic MΔ41p53	65
	MΔ122p53/MΔ122p53 wt/MΔ122p53 m/m wt/m	Overexpression of transgenic MΔ122p53 'm' = deletion of exon 1–6	68 69

truncated TAD followed by a complete DBD and OD (i.e., human Δ40); and Dp53ΔC, encoded by a short transcript leading to a putative isoform bearing only the TAD.[33,45–47] The *Drosophila p53* gene is activated by irradiation and exerts broad suppressive effects recapitulating those of the p53-family members, including regulation of apoptosis, aging, autophagy, differentiation and growth.[48] Historically, DΔNp53 was the first form identified and previously termed Dp53; hence most of the functional studies to date have focused on the role of this particular isoform (Table 1).

Studies on the morphogenesis of imaginal discs have highlighted the role of DΔNp53 isoforms in the control of cell death.[45,47,49,50] In this system, DΔNp53 controls apoptosis through the Reaper–Hid–Grim (RHG) cascade. Indeed, irradiated imaginal discs from flies with a mutant *Dp53* gene show reduced apoptosis but normal cell-cycle arrest, suggesting a specific regulatory role of the *Dp53* gene product in apoptosis.[45,49] So far, the main isoform implicated in apoptosis appears to be DΔNp53, which directly regulates *reaper* (*rpr*) expression.[45] However, DΔNp53 also exerts effects through other pathways. In particular, DΔNp53 may activate either apoptotic or non-apoptotic responses in photoreceptor cells depending on the cell differentiation status.[50–52] Moreover, DΔNp53 appears to inhibit cellular differentiation in the retina independently of its apoptotic function.[50]

The *Dp53* gene product also exerts important roles in controlling lifespan in a sex- and stage-dependent manner. When overexpressed in adult flies (Table 1), DΔNp53 limited lifespan in females and extended it in males. By contrast, when overexpressed during development, DΔNp53 exerted a similar dose-dependent effect on longevity in both sexes.[53] Conversely, inactivation of the *Dp53* locus increased lifespan in females but had only minor effects in males. Similar phenotypes were observed in *Drosophila* expressing dominant-negative Dp53 mutant transgenes or overexpressing DΔNp53, suggesting that DΔNp53 interferes with Dp53 activity.[54] The role of *Drosophila* p53 has also been investigated in detail in other cellular functions such as DNA repair or compensatory proliferation.[48,55,56] However, the specific roles of each *Drosophila* p53 isoform, Dp53 and DΔNp53, have not been studied in sufficient detail to understand the exact contributions of each isoform to development, stress responses or longevity.

***Danio rerio*, the historical model to study p53 isoforms.** The zebrafish p53 protein, Zp53, recapitulates the suppressive and pro-apoptotic functions of human p53 upon genotoxic stress.[57] So far, only N-terminal Zp53 isoforms have been identified (Figure 2): Zp53, corresponding to the human p53 protein;[58] ZΔNp53, produced through an alternative splicing of intron-2 and thus similar to the human Δ40p53 forms;[59] and ZΔ113p53, produced by an internal promoter located within the zebrafish *Zp53* gene that is regulated by Zp53 itself, and thus equivalent to the human Δ133p53 isoform.[37,38,60] In contrast to the human Δ40p53, AUG1 of ZΔNp53 is located within the partial intronic sequence retained by alternative splicing.[59] Thus in ZΔNp53, the 38 N-terminal residues containing the TAD are replaced by 33 residues derived from the intron-2 sequence.

There is evidence that the ZΔNp53 transcript accumulates in response to γ-ray irradiation.[59] In addition, ectopic expression of ZΔNp53 resulted in a strong developmental phenotype with hypoplasia and malformation of the head, eyes and somites (Table 1). This phenotype is dependent upon the presence of Zp53, the two isoforms forming a protein complex through their ODs. Overexpression of an OD-mutant ZΔNp53 or wild-type ZΔNp53 in a mutant Zp53 background is phenotypically ineffective. These observations suggest that ZΔNp53 exerts its effects by modulating the activity of Zp53 during zebrafish development.

ZΔ113p53 was discovered in a different context (Table 1). Cheng and co-workers found that in zebrafish embryos, loss of the *Def* gene (Digestive organ Expansion Factor) led to defects in the morphogenesis of digestive organs. In a genome-wide screen, they identified a shorter form of Zp53 whose expression was upregulated in $def^{-/-}$ embryos.[60] Upregulation of ZΔ113p53 correlated with increased expression of p53-target genes involved in cell-cycle progression such as *cyclin-G1* and $p21^{WAF1}$, whereas pro-apoptotic genes such as *Bax* and *Reprimo* were not activated. Furthermore, ZΔ113p53 was found to selectively upregulate the *Bcl-2L* anti-apoptotic gene.[61] This pattern of effects is consistent with the notion that ZΔ113p53, like Δ133p53, operates as a modulator of p53 in selectively activating defined target genes.

Mouse, the next generation of animal models. Six p53 isoforms have been described in mice, resulting from combination of three N-terminal p53 isoforms with two different C-terminal isoforms (Figure 2). In addition to the full-length mouse Mp53, Rotter and co-workers identified an alternative p53 mRNA retaining part of intron-10 that encodes a shorter isoform with new residues in place of

the usual OD (Mp53AS), homologous to the human p53β.[10,62] Later, Mowat *et al.*[63] isolated MΔ41p53, the mouse counterpart of human Δ40p53 forms. In addition, Khoury *et al.* isolated a shorter N-terminal form produced by an internal promoter within the mouse *p53* gene, MΔ157p53, equivalent to the human Δ160p53 form. Khoury *et al.* also showed that MΔ41p53 and MΔ157p53 can be expressed as a C-terminal AS variant. As for the human p53 isoforms, Rotter and co-workers reported that Mp53AS modulates Mp53-mediated apoptosis and Mp53 transcriptional activity in luciferase reporter assays.[64]

In 2004, Maier *et al.*[65] described a transgenic mouse overexpressing MΔ41p53 (Table 1). When expressed in a p53-null background, this isoform did not induce any particular phenotype. However, when expressed in a p53-competent background, an increased dosage of MΔ41p53 led to reduced size, accelerated aging and a shorter lifespan associated with hypo-insulinemia and glucose insufficiency.[65–67] These effects were attributed to the hyper-activation of the insulin-like growth factor (IGF)-signalling axis by MΔ41p53, setting in motion a cascade that clamps unimpeded growth through p21.[65] Furthermore, these MΔ41p53-overexpressing mice show cognitive decline and synaptic impairment early in life, also attributable to the hyper-activation of the IGF-1-signalling pathway.[67] These observations are consistent with studies performed *in vitro* and in the zebrafish model, indicating that Δ40p53 isoforms may regulate growth suppression through modulation of p53 activity.

A premature aging phenotype was also reported in a non-physiological knock-in p53 mouse model expressing a 'MΔ122p53' form, truncated for the first 122 residues (Table 1).[68] Although MΔ122p53 has no physiological equivalent in mouse, it can be considered as an 'intermediate' between the MΔ41p53 and MΔ157p53 isoforms, as it lacks the TAD and part of the DNA-binding domain. During adulthood, transgenic $p53^{\Delta122p53/\Delta122p53}$ mice showed premature aging symptoms, such as balding and arthritis, similar to that observed in the $p53^{+/m}$ mice (deletion exon 1–6).[68,69] In addition, earlier tumour onset and shortened lifespan were observed in $p53^{+/\Delta122p53}$ mice as compared with $p53^{+/-}$ mice. Overall, these results support the hypothesis that mouse N-terminal p53 forms may operate as dominant oncogenes to promote cell proliferation and inflammation.

The studies summarized above highlight common themes in the 'p53 isoform' field. First, overall patterns of isoform expression are well-conserved throughout evolution. Second, N-terminal isoforms have a major role as regulators of physiological processes related to development, aging, life-span and, possibly, carcinogenesis. In this respect, two key mechanisms are emerging. The Δ40p53 form exerts regulatory effects on signalling cascades controlled by p53, perhaps through direct interaction between the two isoforms. Conversely, isoforms corresponding to Δ133p53α modulate cell response by regulating gene expression in a p53-dependent and -independent manner. It should, however, be remembered that our current view of isoform activities remains fragmentary and that further studies are needed to better understand their roles and underlying mechanisms.

p53 Isoforms and Human Cancers

Genetic polymorphisms: effects on p53 isoform expression. The *TP53* gene is highly polymorphic, with over 80% of known single-nucleotide polymorphisms (SNPs) located within introns or non-coding 5′ and 3′ sequences.[70] Hainaut and co-workers showed that *G-quadruplex* structures, formed in intron-3 and regulating intron-2 splicing (and thus Δ40 form expression), overlap a common polymorphism, *TP53* PIN3, which consists of a 16-bp duplication (A1, non-duplicated allele; A2, duplicated allele).[30,71] This polymorphism may therefore modulate the structure and/or the stability of the *G-quadruplexes*, and affect Δ40p53 expression (Table 2). Consistent with a functional effect, *TP53* PIN3 has recently been identified as a strong genetic modifier of germline *TP53* mutations. Indeed, carriers bearing A1A1 genotypes developed their first cancer on average 20 years earlier than carriers with an A1A2 genotype.[72] This effect was detected in a Brazilian cohort, in which the p.R337H mutation of partial penetrance is very common owing to a widespread founder effect.[5] It remains to be demonstrated whether a similar effect is observed in carriers of other *TP53* mutation types.

Landi and co-workers have identified 11 different haplotypes defined by eight SNPs in a region from exon-3 to intron-4, overlapping part of the internal P2 promoter regulating Δ133p53 and Δ160p53 expression.[73] Using these different haplotypes as promoters to drive *luciferase* expression, they found significant differences in basal promoter activity that were confirmed by analysing endogenous

Table 2 Role of p53 isoforms in human cancers

Cancer	Isoform	Observation	Ref.
Breast cancer	Δ133p53α	Overexpression	33
	p53β	Loss of expression	33,78
	p53γ	Loss of expression	33,79
		Prognosis of mutp53/p53γ as good as wtp53	
Acute myeloid leukaemia	p53β/γ	Increased expression in response to chemotherapy	78
Ovarian cancer	p53β	Serous and poorly differentiated tumours/worse recurrence-free and overall survival	77
	Δ40p53	Improved recurrence-free survival	74
Li–Fraumeni syndrome	Δ40p53	*G-quadruplex* structure overlapping *TP53* PIN3/early age at first cancer diagnosis in patients carrying A1A1 compared with patients carrying A1A2	72

Δ133p53 expression in samples of normal colonic mucosa. Furthermore, *in silico* and *in vitro* DNA binding analyses suggested that an SNP in intron-4 (rs179287; C>T) affects proteins binding within the P2 promoter.[73] These results suggest that genetic polymorphisms may modulate basal Δ133p53 expression. Whether these effects influence cancer risk remains to be demonstrated.

Mutations affecting the production of isoforms. Given the role of N-terminal p53 forms as inhibitors of p53 transactivation, it can be expected that overexpression of certain isoforms may represent an alternative to a mutation in *TP53* for inactivating p53 in cancers, as observed in some small clinical studies.[33,74,75] Analysis of the IARC *TP53* database identified 1019 somatic cancer mutations that are predicted to disrupt the p53 coding sequence in the N-terminal region while leaving intact the sequence encoding at least Δ160p53. This represents 3.65% of all somatic mutations reported to date (version R15; www-p53.iarc.fr).[70] Therefore, rare and/or silent mutations in *TP53* may differentially affect the expression of p53 isoforms, as observed recently for mutations present in the IRES regulating the relative expression of p53 and Δ40p53 forms encoded by the full-length FSp53 mRNA.[76] With respect to C-terminal isoforms, Hofstetter *et al.*[77] have used an yeast-base functional assay and RT-PCR to show that mutations at splice sites of *TP53* intron-9 can lead to aberrant expression of p53β mRNA in primary cultures of ovarian cancer cells. By contrast, somatic mutations at splice sites of introns 6 and 9 were found to generate spliced mutant p53 mRNAs (p53ζ, p53δ and p53ε), distinct from the physiological isoforms. However, the role of such mutant isoforms in carcinogenesis is not known.

Isoform expression and cancer outcomes. There is emerging evidence that p53 isoform expression is deregulated in human cancers (Table 2). In a study of 245 primary ovarian cancers, Hofstetter *et al.*[77] observed that expression of p53β was associated with serous and poorly differentiated cancers, and, when expressed together with the functional p53 protein, it was correlated with poor recurrence-free and overall survival. By contrast, tumours expressing functional p53 and Δ40p53 showed improved recurrence-free survival of patients compared with tumours expressing no Δ40p53. Studies of patients with acute myeloid leukemia have shown that elevated expression of p53β and/or p53γ in blood cells was correlated with improved responses to chemotherapy,[78] which may predict decreased chemoresistance and improved overall survival. An effect of isoform expression has also been observed on breast cancer prognosis. In a cohort of 127 breast cancer patients, Thompson and co-workers reported that patients whose tumours expressed both mutant p53 and p53γ mRNAs had a prognosis as good as patients whose tumours expressed a wild-type p53, suggesting that the expression of p53γ may abrogate the poor prognosis commonly associated with *TP53* mutations.[79]

Conclusion

The field of 'p53 isoforms' is still in its infancy, but the increasing number of genetic, biochemical and clinical studies have clearly established that p53 isoforms are fundamental and important components of the p53 pathway. Data obtained from animal and cellular models indicate that p53 isoforms regulate the cell fate in response to developmental defects and cell damages by differentially regulating gene expression, both in a p53-dependent and -independent manner. Furthermore, the current data suggest that p53 isoforms have roles in all biological activities regulated by p53 (Figure 3). Therefore, one can reasonably expect that the characterization of the biochemical and biological activities of p53 isoforms will impact on the fields of cancer, embryo development and aging.

Future experiments will be needed to gain further insight into how p53 isoforms modulate the different biological

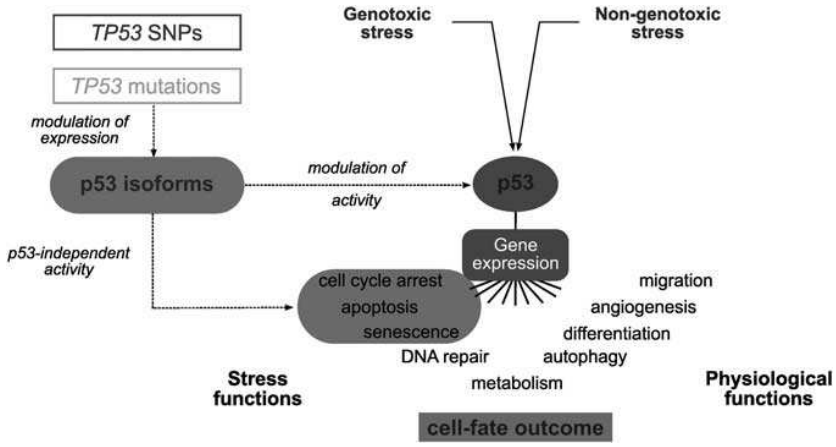


Figure 3 p53 isoforms in the p53 network. p53 integrates the different stress signals to adapt cell fate to the intensity and the nature of stress by regulating several biological functions to maintain genomic and cellular integrity. In addition, p53 controls physiological functions under basal conditions. Recent data suggest that p53 isoforms modulate p53-mediated cell fate outcome and may thus be key components of the p53-mediated decision not only in response to stress but also under basal conditions. It has also been reported that p53 isoforms have p53-independent activities and directly regulate cell-cycle arrest and apoptosis. In addition, genetic alterations, such as *TP53* SNPs and *TP53* mutations, affect the expression of p53 isoforms, which may result in tumorigenesis

activities. However, based on our current understanding of the p63, p73 and p53 isoforms, we have come to realize that the p53 pathway should no longer be considered as regulated by only the p53 protein, but by a set of p63, p73 and p53 isoforms that interplay with each other in regulating physiological functions. Further progress on this front will require the development of robust and standardized tools for the identification and quantification of p53, p63 and p73 protein isoform expression in experimental systems as well as in human tissues.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgements. We thank all the speakers for their participation in the 1st International p53 Isoforms Meeting, and Michelle Wrisez for logistic help. The meeting was funded by IARC, the College of Medicine of the University of Dundee and the Laboratory of Molecular Biology of the Cell (UMR5239, CNRS at ENS de Lyon). Links: IARC TP53 database, http://www-p53.iarc.fr/; abstracts of the International p53 Isoforms Meeting, http://www-p53.iarc.fr/Download/Book_p53Isoforms2010_Final.pdf. VM is supported by a fellowship from the Breast Cancer Campaign; M-LD-D and BM by CNRS ATIP, Fondation pour la Recherche Médicale (FRM) and La Ligue contre le cancer; CS by a PhD grant from the French Ministry of Research; JH by INSERM and Institut Curie; PH and JH by l'Insititut National contre le Cancer (INCa, France; Projet Libre 2009, 2009-192); and J-CB by Cancer Research UK.

Facts

- At least nine different mRNAs are expressed from the *TP53* gene encoding 12 distinct protein isoforms through the use of alternative promoters, splicing sites and internal initiation sites of translation.
- p53 isoforms regulate cell fate in response to developmental defects and cell damage by differentially regulating gene expression, both in a p53-dependent and -independent manner.
- The N-terminal p53 isoforms, Δ40 and Δ133, are not strict dominant-negative inhibitors of p53's suppressive function as they have p53-independent activities.
- The expression patterns and biological activities of p53 isoform are conserved through evolution (human, mouse, zebrafish and *Drosophila*).
- *TP53* SNPs and mutations alter the pattern of p53 isoform expression.
- p53 isoform expression is deregulated in human cancers and is associated with clinical prognosis.

Open Questions

- Development of standardized and robust tools is needed to systematically identify and quantify p53 isoforms in human and animal models.
- Characterization of the biochemical activities of p53 isoforms as well as their molecular targets.
- Is each p53 isoform specialized in a particular biological response? Have they redundant activities?
- Evaluation of the interplay between p53-, p63- and p73-family isoforms.
- Study the potential use of p53 isoform expression in human cancers as biomarkers and as therapeutic targets.

1. Levine AJ, Oren M. The first 30 years of p53: growing ever more complex. *Nat Rev Cancer* 2009; **9**: 749–758.
2. Levine AJ, Tomasini R, McKeon FD, Mak TW, Melino G. The p53 family: guardians of maternal reproduction. *Nat Rev Mol Cell Biol* 2011; **12**: 259–265.
3. Meek DW, Anderson CW. Posttranslational modification of p53: cooperative integrators of function. *Cold Spring Harb Perspect Biol* 2009; **1**: a000950.
4. Vilborg A, Wilhelm MT, Wiman KG. Regulation of tumor suppressor p53 at the RNA level. *J Mol Med* 2010; **88**: 645–652.
5. Garritano S, Gemignani F, Palmero EI, Olivier M, Martel-Planche G, Le Calvez-Kelm F *et al.* 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* 2010; **31**: 143–150.
6. Khoury MP, Bourdon JC. The isoforms of the p53 protein. *Cold Spring Harb Perspect Biol* 2010; **2**: a000927.
7. Marcel V, Hainaut P. p53 isoforms – a conspiracy to kidnap p53 tumor suppressor activity? *Cell Mol Life Sci* 2009; **66**: 391–406.
8. Flaman JM, Waridel F, Estreicher A, Vannier A, Limacher JM, Gilbert D *et al.* The human tumour suppressor gene p53 is alternatively spliced in normal cells. *Oncogene* 1996; **12**: 813–818.
9. Matlashewski G, Lamb P, Pim D, Peacock J, Crawford L, Benchimol S. Isolation and characterization of a human p53 cDNA clone: expression of the human p53 gene. *EMBO J* 1984; **3**: 3257–3262.
10. Wolf D, Harris N, Goldfinger N, Rotter V. Isolation of a full-length mouse cDNA clone coding for an immunologically distinct p53 molecule. *Mol Cell Biol* 1985; **5**: 127–132.
11. Marcel V, Olivier M, Mollereau B, Hainaut P, Bourdon JC. First International p53 Isoforms Meeting: 'p53 isoforms through evolution: from identification to biological function'. *Cell Death Differ* 2011; **18**: 563–564.
12. Melino G. p63 is a suppressor of tumorigenesis and metastasis interacting with mutant p53. *Cell Death Differ* 2011; **18**: 1487–1499.
13. Murray-Zmijewski F, Lane DP, Bourdon JC. p53/p63/p73 isoforms: an orchestra of isoforms to harmonise cell differentiation and response to stress. *Cell Death Differ* 2006; **13**: 962–972.
14. Kaghad M, Bonnet H, Yang A, Creancier L, Biscan JC, Valent A *et al.* Monoallelically expressed gene related to p53 at 1p36, a region frequently deleted in neuroblastoma and other human cancers. *Cell* 1997; **90**: 809–819.
15. Yang A, Walker N, Bronson R, Kaghad M, Oosterwegel M, Bonnin J *et al.* p73-deficient mice have neurological, pheromonal and inflammatory defects but lack spontaneous tumours. *Nature* 2000; **404**: 99–103.
16. Yang A, Kaghad M, Wang Y, Gillett E, Fleming MD, Dotsch V *et al.* p63, a p53 homolog at 3q27-29, encodes multiple products with transactivating, death-inducing, and dominant-negative activities. *Mol Cell* 1998; **2**: 305–316.
17. Mills AA, Zheng B, Wang XJ, Vogel H, Roop DR, Bradley A. p63 is a p53 homologue required for limb and epidermal morphogenesis. *Nature* 1999; **398**: 708–713.
18. Senoo M, Pinto F, Crum CP, McKeon F. p63 is essential for the proliferative potential of stem cells in stratified epithelia. *Cell* 2007; **129**: 523–536.
19. Su X, Chakravarti D, Cho MS, Liu L, Gi YJ, Lin YL *et al.* TAp63 suppresses metastasis through coordinate regulation of Dicer and miRNAs. *Nature* 2010; **467**: 986–990.
20. Su X, Paris M, Gi YJ, Tsai KY, Cho MS, Lin YL *et al.* TAp63 prevents premature aging by promoting adult stem cell maintenance. *Cell Stem Cell* 2009; **5**: 64–75.
21. Medawar A, Virolle T, Rostagno P, de la Forest-Divonne S, Gambaro K, Rouleau M *et al.* DeltaNp63 is essential for epidermal commitment of embryonic stem cells. *PLoS One* 2008; **3**: e3441.
22. Shalom-Feuerstein R, Lena AM, Zhou H, De La Forest Divonne S, Van Bokhoven H, Candi E *et al.* DeltaNp63 is an ectodermal gatekeeper of epidermal morphogenesis. *Cell Death Differ* 2011; **18**: 887–896.
23. Keyes WM, Pecoraro M, Aranda V, Vernersson-Lindahl E, Li W, Vogel H *et al.* DeltaNp63alpha is an oncogene that targets chromatin remodeler Lsh to drive skin stem cell proliferation and tumorigenesis. *Cell Stem Cell* 2011; **8**: 164–176.
24. Guo X, Keyes WM, Papazoglu C, Zuber J, Li W, Lowe SW *et al.* TAp63 induces senescence and suppresses tumorigenesis *in vivo*. *Nat Cell Biol* 2009; **11**: 1451–1457.
25. Wilhelm MT, Rufini A, Wetzel MK, Tsuchihara K, Inoue S, Tomasini R *et al.* Isoform-specific p73 knockout mice reveal a novel role for delta Np73 in the DNA damage response pathway. *Genes Dev* 2010; **24**: 549–560.
26. Tomasini R, Tsuchihara K, Wilhelm M, Fujitani M, Rufini A, Cheung CC *et al.* TAp73 knockout shows genomic instability with infertility and tumor suppressor functions. *Genes Dev* 2008; **22**: 2677–2691.
27. Matlashewski G, Pim D, Banks L, Crawford L. Alternative splicing of human p53 transcripts. *Oncogene Res* 1987; **1**: 77–85.
28. Ghosh A, Stewart D, Matlashewski G. Regulation of human p53 activity and cell localization by alternative splicing. *Mol Cell Biol* 2004; **24**: 7987–7997.
29. Courtois S, Verhaegh G, North S, Luciani MG, Lassus P, Hibner U *et al.* DeltaN-p53, a natural isoform of p53 lacking the first transactivation domain, counteracts growth suppression by wild-type p53. *Oncogene* 2002; **21**: 6722–6728.
30. Marcel V, Tran PL, Sagne C, Martel-Planche G, Vaslin L, Teulade-Fichou MP *et al.* G-quadruplex structures in TP53 intron 3: role in alternative splicing and in production of p53 mRNA isoforms. *Carcinogenesis* 2011; **32**: 271–278.

31. Candeias MM, Powell DJ, Roubalova E, Apcher S, Bourougaa K, Vojtesek B *et al.* Expression of p53 and p53/47 are controlled by alternative mechanisms of messenger RNA translation initiation. *Oncogene* 2006; **25**: 6936–6947.
32. Ray PS, Grover R, Das S. Two internal ribosome entry sites mediate the translation of p53 isoforms. *EMBO Rep* 2006; **7**: 404–410.
33. Bourdon JC, Fernandes K, Murray-Zmijewski F, Liu G, Diot A, Xirodimas DP *et al.* p53 isoforms can regulate p53 transcriptional activity. *Genes Dev* 2005; **19**: 2122–2137.
34. Reisman D, Balint E, Loging WT, Rotter V, Almon E. 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* 1996; **38**: 364–370.
35. Reisman D, Greenberg M, Rotter V. Human p53 oncogene contains one promoter upstream of exon 1 and a second, stronger promoter within intron 1. *Proc Natl Acad Sci USA* 1988; **85**: 5146–5150.
36. Marcel V, Perrier S, Aoubala M, Ageorges S, Groves MJ, Diot A *et al.* Delta160p53 is a novel N-terminal p53 isoform encoded by Delta133p53 transcript. *FEBS Lett* 2010; **584**: 4463–4468.
37. Aoubala M, Murray-Zmijewski F, Khoury MP, Fernandes K, Perrier S, Bernard H *et al.* p53 directly transactivates Delta133p53alpha, regulating cell fate outcome in response to DNA damage. *Cell Death Differ* 2011; **18**: 248–258.
38. Marcel V, Vijayakumar V, Fernandez-Cuesta L, Hafsi H, Sagne C, Hautefeuille A *et al.* p53 regulates the transcription of its Delta133p53 isoform through specific response elements contained within the TP53 P2 internal promoter. *Oncogene* 2010; **29**: 2691–2700.
39. Fuller-Pace FV, Ali S. The DEAD box RNA helicases p68 (Ddx5) and p72 (Ddx17): novel transcriptional co-regulators. *Biochem Soc Trans* 2008; **36** (Part 4): 609–612.
40. Graupner V, Schulze-Osthoff K, Essmann F, Janicke RU. Functional characterization of p53beta and p53gamma, two isoforms of the tumor suppressor p53. *Cell Cycle* 2009; **8**: 1238–1248.
41. Yin Y, Stephen CW, Luciani MG, Fahraeus R. p53 stability and activity is regulated by Mdm2-mediated induction of alternative p53 translation products. *Nat Cell Biol* 2002; **4**: 462–467.
42. Brady CA, Jiang D, Mello SS, Johnson TM, Jarvis LA, Kozak MM *et al.* Distinct p53 transcriptional programs dictate acute DNA-damage responses and tumor suppression. *Cell* 2011; **145**: 571–583.
43. Ohki R, Kawase T, Ohta T, Ichikawa H, Taya Y. Dissecting functional roles of p53 N-terminal transactivation domains by microarray expression analysis. *Cancer Sci* 2007; **98**: 189–200.
44. Fujita K, Mondal AM, Horikawa I, Nguyen GH, Kumamoto K, Sohn JJ *et al.* p53 isoforms Delta133p53 and p53beta are endogenous regulators of replicative cellular senescence. *Nat Cell Biol* 2009; **11**: 1135–1142.
45. Brodsky MH, Nordstrom W, Tsang G, Kwan E, Rubin GM, Abrams JM. *Drosophila* p53 binds a damage response element at the reaper locus. *Cell* 2000; **101**: 103–113.
46. Jin S, Martinek S, Joo WS, Wortman JR, Mirkovic N, Sali A *et al.* Identification and characterization of a p53 homologue in *Drosophila melanogaster*. *Proc Natl Acad Sci USA* 2000; **97**: 7301–7306.
47. Ollmann M, Young LM, Di Como CJ, Karim F, Belvin M, Robertson S *et al. Drosophila* p53 is a structural and functional homolog of the tumor suppressor p53. *Cell* 2000; **101**: 91–101.
48. Rutkowski R, Hofmann K, Gartner A. Phylogeny and function of the invertebrate p53 superfamily. *Cold Spring Harb Perspect Biol* 2010; **2**: a001131.
49. Brodsky MH, Weinert BT, Tsang G, Rong YS, McGinnis NM, Golic KG *et al. Drosophila melanogaster* MNK/Chk2 and p53 regulate multiple DNA repair and apoptotic pathways following DNA damage. *Mol Cell Biol* 2004; **24**: 1219–1231.
50. Fan Y, Lee TV, Xu D, Chen Z, Lamblin AF, Steller H *et al.* Dual roles of *Drosophila* p53 in cell death and cell differentiation. *Cell Death Differ* 2010; **17**: 912–921.
51. Mendes CS, Levet C, Chatelain G, Dourlen P, Fouillet A, Dichtel-Danjoy ML *et al.* ER stress protects from retinal degeneration. *EMBO J* 2009; **28**: 1296–1307.
52. Peterson C, Carney GE, Taylor BJ, White K. reaper is required for neuroblast apoptosis during *Drosophila* development. *Development* 2002; **129**: 1467–1476.
53. Waskar M, Landis GN, Shen J, Curtis C, Tozer K, Abdueva D *et al. Drosophila melanogaster* p53 has developmental stage-specific and sex-specific effects on adult life span indicative of sexual antagonistic pleiotropy. *Aging (Albany NY)* 2009; **1**: 903–936.
54. Bauer JH, Poon PC, Glatt-Deeley H, Abrams JM, Helfand SL. Neuronal expression of p53 dominant-negative proteins in adult *Drosophila melanogaster* extends life span. *Curr Biol* 2005; **15**: 2063–2068.
55. Lu B. Recent advances in using *Drosophila* to model neurodegenerative diseases. *Apoptosis* 2009; **14**: 1008–1020.
56. Martin FA, Perez-Garijo A, Morata G. Apoptosis in *Drosophila*: compensatory proliferation and undead cells. *Int J Dev Biol* 2009; **53**: 1341–1347.
57. Storer NY, Zon LI. Zebrafish models of p53 functions. *Cold Spring Harb Perspect Biol* 2010; **2**: a001123.
58. Cheng R, Ford BL, O'Neal PE, Mathews CZ, Bradford CS, Thongtan T *et al.* Zebrafish (*Danio rerio*) p53 tumor suppressor gene: cDNA sequence and expression during embryogenesis. *Mol Mar Biol Biotechnol* 1997; **6**: 88–97.
59. Davidson WR, Kari C, Ren Q, Daroczi B, Dicker AP, Rodeck U. Differential regulation of p53 function by the N-terminal DeltaNp53 and Delta113p53 isoforms in zebrafish embryos. *BMC Dev Biol* 2010; **10**: 102.
60. Chen J, Ruan H, Ng SM, Gao C, Soo HM, Wu W *et al.* Loss of function of def selectively upregulates Delta113p53 expression to arrest expansion growth of digestive organs in zebrafish. *Genes Dev* 2005; **19**: 2900–2911.
61. Chen J, Ng SM, Chang C, Zhang Z, Bourdon JC, Lane DP *et al.* p53 isoform delta113p53 is a p53 target gene that antagonizes p53 apoptotic activity via BclxL activation in zebrafish. *Genes Dev* 2009; **23**: 278–290.
62. Arai N, Nomura D, Yokota K, Wolf D, Brill E, Shohat O *et al.* Immunologically distinct p53 molecules generated by alternative splicing. *Mol Cell Biol* 1986; **6**: 3232–3239.
63. Mowat M, Cheng A, Kimura N, Bernstein A, Benchimol S. Rearrangements of the cellular p53 gene in erythroleukaemic cells transformed by Friend virus. *Nature* 1985; **314**: 633–636.
64. Almog N, Goldfinger N, Rotter V. p53-dependent apoptosis is regulated by a C-terminally alternatively spliced form of murine p53. *Oncogene* 2000; **19**: 3395–3403.
65. Maier B, Gluba W, Bernier B, Turner T, Mohammad K, Guise T *et al.* Modulation of mammalian life span by the short isoform of p53. *Genes Dev* 2004; **18**: 306–319.
66. Hinault C, Kawamori D, Liew CW, Maier B, Hu J, Keller SR *et al.* {Delta}40 isoform of p53 controls {beta}-cell proliferation and glucose homeostasis in mice. *Diabetes* 2011; **60**: 1210–1222.
67. Pehar M, O'Riordan KJ, Burns-Cusato M, Andrzejewski ME, del Alcazar CG, Burger C *et al.* Altered longevity-assurance activity of p53:p44 in the mouse causes memory loss, neurodegeneration and premature death. *Aging Cell* 2010; **9**: 174–190.
68. Slatter TL, Hung N, Campbell H, Rubio C, Mehta R, Renshaw P *et al.* Hyperproliferation, cancer, and inflammation in mice expressing a {Delta}133p53-like isoform. *Blood* 2011; **117**: 5166–5177.
69. Tyner SD, Venkatachalam S, Choi J, Jones S, Ghebranious N, Igelmann H *et al.* p53 mutant mice that display early ageing-associated phenotypes. *Nature* 2002; **415**: 45–53.
70. Petitjean A, Mathe E, Kato S, Ishioka C, Tavtigian SV, Hainaut P *et al.* Impact of mutant p53 functional properties on TP53 mutation patterns and tumor phenotype: lessons from recent developments in the IARC TP53 database. *Hum Mutat* 2007; **28**: 622–629.
71. Lazar V, Hazard F, Bertin F, Janin N, Bellet D, Bressac B. Simple sequence repeat polymorphism within the p53 gene. *Oncogene* 1993; **8**: 1703–1705.
72. Marcel V, Palmero EI, Falagan-Lotsch P, Martel-Planche G, Ashton-Prolla P, Olivier M *et al.* TP53 PIN3 and MDM2 SNP309 polymorphisms as genetic modifiers in the Li–Fraumeni syndrome: impact on age at first diagnosis. *J Med Genet* 2009; **46**: 766–772.
73. Bellini I, Pitto L, Marini MG, Porcu L, Moi P, Garritano S *et al.* DeltaN133p53 expression levels in relation to haplotypes of the TP53 internal promoter region. *Hum Mutat* 2010; **31**: 456–465.
74. Avery-Kiejda KA, Zhang XD, Adams LJ, Scott RJ, Vojtesek B, Lane DP *et al.* Small molecular weight variants of p53 are expressed in human melanoma cells and are induced by the DNA-damaging agent cisplatin. *Clin Cancer Res* 2008; **14**: 1659–1668.
75. Machado-Silva A, Perrier S, Bourdon JC. p53 family members in cancer diagnosis and treatment. *Semin Cancer Biol* 2010; **20**: 57–62.
76. Grover R, Sharathchandra A, Ponnuswamy A, Khan D, Das S. Effect of mutations on the p53 IRES RNA structure: implications for de-regulation of the synthesis of p53 isoforms. *RNA Biol* 2011; **8**: 132–142.
77. Hofstetter G, Berger A, Fiegl H, Slade N, Zoric A, Holzer B *et al.* Alternative splicing of p53 and p73: the novel p53 splice variant p53delta is an independent prognostic marker in ovarian cancer. *Oncogene* 2010; **29**: 1997–2004.
78. Anensen N, Oyan AM, Bourdon JC, Kalland KH, Bruserud O, Gjertsen BT. A distinct p53 protein isoform signature reflects the onset of induction chemotherapy for acute myeloid leukemia. *Clin Cancer Res* 2006; **12**: 3985–3992.
79. Bourdon JC, Khoury MP, Diot A, Baker L, Fernandes K, Aoubala M *et al.* p53 mutant breast cancer patients expressing p53gamma have as good a prognosis as wild-type p53 breast cancer patients. *Breast Cancer Res* 2011; **13**: R7.

Part III.G-quadruplex structures

A. Formation of G-quadruplex structures

In 1910, the first formation of G-rich structure was reported by Bang. This author observed the formation of a gel with a high concentration of G in aqueous solution, suggesting that G-rich sequences in DNA are able to form higher-order structures. In 1962, Gellert and collaborators observed using X-rays that G could form a tetrameric structure (Gellert et al 1962). In these tetramers, four G molecules form a square planar arrangement in which each guanine is bound to the second adjacent G by a Hoogsteen bond (G1:N-OH-O:G2) (**Figure 17**) (Simonsson 2001). This arrangement is called a G-quartet. The superposition of at least three G-quartets forms a G4 structure. The G4 is stabilized by the presence of monovalent cations that occupy the central cavities between the G-quartets (Williamson et al 1989). The nucleotides present in the sequence, which are not implicated in G-quartet formation, make up participate in the loop. Usually, the loops are small (1-7 nucleotides) and the smaller are the loops, the more stable will be the G4, which is also influenced by the number of guanines implicated in the G4 (3-5 G) (Huppert 2010). Recently, it was observed that G4 appears to be polymorphic in their form depending of the manner in which the structure was allowed to form (**Figure 17**) (Phan et al 2006, Ying et al 2003).

The discovery in 1989 that G4s are formed in eukaryotic telomeres re-awakened interest in these structures and it is now well accepted that G4s are formed in many regions of the genome and are implicated in different biological functions (Sundquist and Klug 1989).

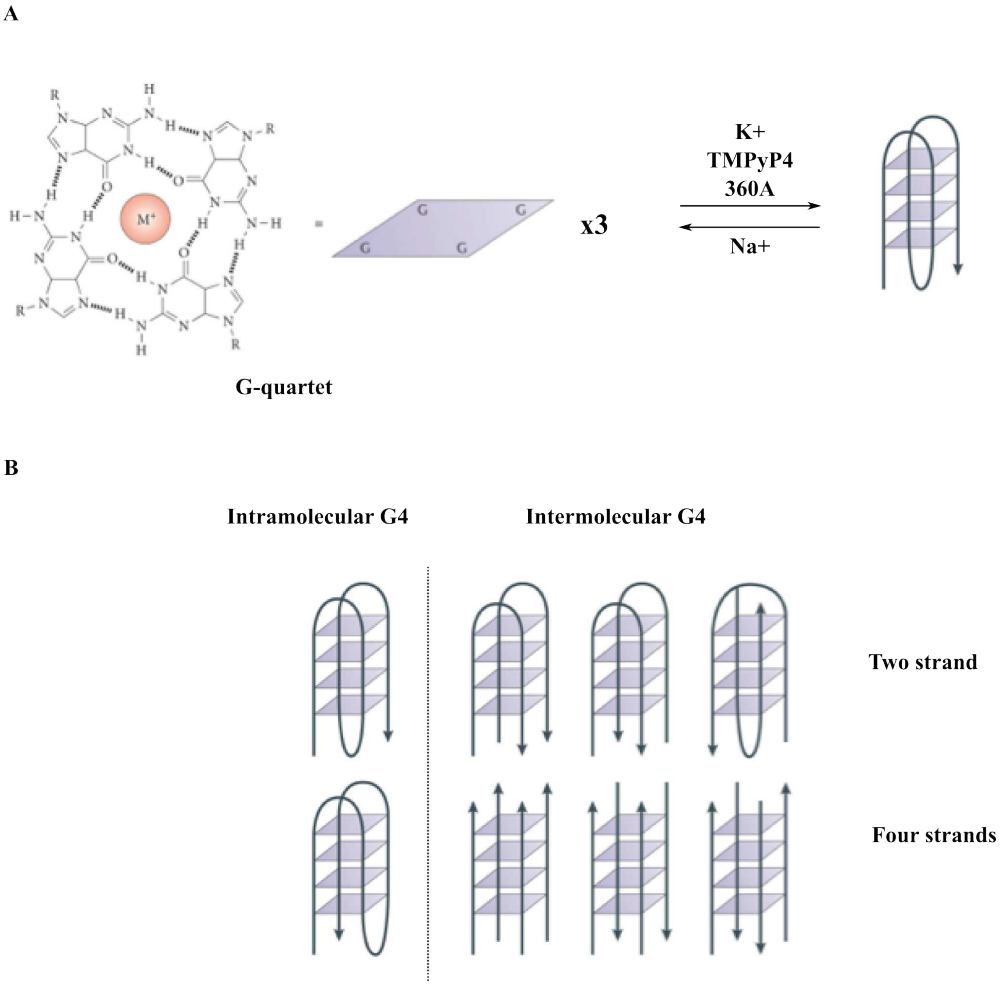


Figure 17: Schematic representation of G4 formation and structures in DNA or RNA. (**A**) Representation of interactions in G-quartet and in G4 structure formation. This quartet is represented as a square. M^+ denotes a monovalent cation. (**B**) Schematic diagrams of intramolecular (left) or intermolecular (right) G4 structures. The arrowheads indicate the direction of the nucleic acid strands. The intermolecular structures shown have two (upper) or four (lower strandsTMPyP4 and 360A are two drugs that modulate G4 structures and the presence of Na+ destabilizes G4 and K+ stabilized the structures. (Adapted from (Bochman et al 2012)).

B. G-quadruplexes chromosomal localisation

The localisation of G4 structures, in DNA or RNA, can be predicted using computational analyses. These identified regions are called Putative Quadruplex Sequences (PQS) (Huppert and Balasubramanian 2005, Huppert 2008a). Different algorithms have been developed based on the number of nucleotides implicated in the structure ($30 \leq n \leq 45$) and on a repeat sequence with at least four runs of G-quartets, in which each G-tract contains at least two guanines. The reference sequence is $G_xX_{y1}G_xX_{y2}G_xX_{y3}G_x$, where X denotes the numbers of guanine implicated in G-quartet ($x \geq 2$) and y1, y2 and y3 the number of nucleotides implicated in the loop ($1 \leq y1, y2, y3 \leq 7$).

To date, a total on around 379,000 PQS was described as potentially being able to form a G4 structure in the human genome (Huppert and Balasubramanian 2005). This number is a third less than would be expected by chance and would suggest an evolutionary pressure against G4 formation (Huppert 2008b). The genomic localisation of G4 structures is preferentially in non-coding regions, with a higher density in telomeres (Huppert et al 2008). The number of PQS is 6.4 times higher in promoter region as compared to the average throughout the whole genome, 8.6 times higher in the nuclease hypersensitive cluster and 230 times higher in regions that are both promoters and nuclease hypersensitive clusters (Huppert and Balasubramanian 2007). G4 structures are also formed in 5'UTR and 3'UTR regions (Huppert et al 2008). Furthermore, G4 structures are preferentially found in the coding strand, suggesting that G4 may play a role in post-transcriptional mechanisms (Rankin et al 2005).

C. Biological functions of G-quadruplexes

As described above, G4 structures can be found in telomeres and in different parts of a gene: 5'UTR, promoter or 3'UTR. This observation suggests that G4s could have different biological functions depending of their localisation, and that they may exert effects through their presence in DNA as well as in RNA.

I. Regulation in telomeres

Telomeres are G-rich repeat sequences (TTAGGG) located in the end of the chromosomes. Telomere DNA contains double-stranded tandem repeats of this sequence followed by a terminal 3' G-rich single-stranded overhang. Due to the enriched G content, the single-stranded telomere sequence can form G4s (Sundquist and Klug 1989). In humans, telomeres exist in a slow equilibrium between the stable G4 structure formed by double-stranded DNA and the unfolded, single-stranded form (Bochman et al 2012). Telomeric G4 structures could be present in polymorphic forms: parallel or antiparallel conformation (**Figure 17B**). The mechanisms of formation of G4 in telomeres are still unclear. The RecQ helicase WNR, causing the Werner syndrome characterised by premature aging and increased risk of cancer, and the RecQ helicase BLM, implicated in Bloom's syndrome associated with higher cancer predisposition, act on telomere structures and can unwind G4 structure *in vitro* (Paeschke et al 2010). Changes in G4 conformation induce a modification of telomerase activity. Antiparallel G4 structures can block telomerase activity, whereas parallel G4s enable

the extension by telomerase (Oganesian and Bryan 2007, Zahler et al 1991). In conclusion, telomeric G4 structures are implicated in different biological functions such as telomere protection and inhibition of telomerase-dependent telomere extension (Lipps and Rhodes 2009).

II. Role of G-quadruplexes in 5'UTR and promoter region

Whole genome studies showed that 47% of promoters contain sequence that could form at least one G4 (Huppert and Balasubramanian 2007, Huppert et al 2008). *In silico* analyses showed that oncogene promoters are richer in G4 structure (69%) compared to tumour suppressor genes (31%), whereas they are under-represented in housekeeping genes. This observation suggests that G4 structures play a role in promoter activity and in gene expression regulation. In humans, G4 structures have been described in a number of genes including *HIFA*α, *BCL-2*, *c-MYC* (Dai et al 2006, De Armond et al 2005, Simonsson et al 1998).

To time, the most studied G4 structure is the one located in the *c-MYC* promoter region (Huppert 2010). Luciferase assays showed that mutations in the G-rich sequence induce a disruption of the G4 and an increase in promoter activity (Siddiqui-Jain et al 2002). Treatment of cells with 5, 10, 15, 20-tetra-(*N-methyl-4-pyridyl*)porphyrin chloride (TMPyP4), a drug which stabilises G4 structures formed in DNA or RNA, induces a decrease of the transcriptional activity of the gene by inhibiting the fixation of the complex of transcriptional proteins. The roles of G4s in other promoter regions are still unclear but may be similar to that seen for *c-Myc*. In addition, supercoiling in or near the promoter region has both positive and negative effects on transcription (Kouzine et al 2008). Similar to stress-induced supercoiling, G4 structures could be formed after stress during transcription (Sun and Hurley 2009). Thus, on the one hand, G4s could inhibit transcription if present in the template DNA strand by blocking the transcription machinery (Bochman et al 2012), however on the other hand, G4s could enhance transcription if present in the non-template DNA strand by maintaining the template strand in a single-stranded conformation. Furthermore, transcription could be affected by the preferential binding of proteins to G4 structures, for example transcriptional enhancers *versus* repressors.

III. Role of G-quadruplexes in mRNA production and stability

1. Role in 3'UTR and poly-adenylation

G4 structures are over-represented in the 3'end of genes and play a role in mRNA poly-adenylation processing (Huppert et al 2008). mRNA 3'end formation or poly-adenylation is a nuclear process, which results in the cleavage of primary transcripts and the acquisition of a poly(A) tail. The poly(A) tail of mature mRNA is essential for its stability, translocation to the cytoplasm and translation. A G-rich region is frequently associated with the poly-adenylation signal (Hu et al 2005) and some of these regions were predicted to fold into G4 structures (Zarudnaya et al 2003). The hnRNP H/F (heterogenous nuclear ribonucleoproteins H/F) can recognise sequences associated with the poly(A) signal and influence splicing and poly-adenylation by increasing the formation and the stability of G4 structures (Bagga et al 1995, Dalziel et al 2007).

Recently, it was described that the pre-mRNA of *TP53* presents G4 structures in its 3' flanking region, upstream the poly-adenylation site (Decorsiere et al 2011). This structure influences the maturation of p53 mRNA by modulating the cleavage of the poly-adenylation site. The formation of this secondary structure is regulated by hnRNP H/F after UV damage.

2. Role in splicing of pre-mRNA

The majority of genes are estimated to undergo alternative splicing (Xiao and Lee 2010). The presence of c*is*-regulatory elements nearby splice sites, such as miRNAs fixation sites or G4 structures can influence spliceosome assembly.

To date, few G4 structures located near splicing sites, in intron or exon regions, have been described. First, a G4 located in intron 5 of the *hTERT* gene is proposed to control the splicing efficiency of *hTERT* by acting as an intronic splicing silencer (Gomez et al 2004). Secondly, a G4 is located in exon 2 of Bcl-X pre-mRNA, a Bcl-2 family member critical for cell survival and apoptosis. This mRNA generates two antagonistic isoforms, Bcl-XL and Bcl-xS (Hai et al 2008). The presence of the G4 decreases the expression of the Bcl-XL and increases that of the Bcl-xS isoform. Thirdly, two independent G4s located in exon 15 of *Fragile mental retardation 1 (FMR1)*, were described as exonic splicing enhancers controlling *FMR1* pre-mRNA alternative splicing (Didiot et al 2008). In addition, several proteins of the *AFF* (AF4/FMR2) gene family are involved in alternative splicing regulation through an interaction with G4s (Melko and Bardoni 2010).

IV. Other roles of G-quadruplexes

Additional roles of G4 structures are emerging, such as regulation of the epigenome and regulation of replication or meiosis (Bochman et al 2012).

A study has shed light on the role of REV1, a DNA repair protein implicated in translesion bypass, in maintaining the patterns and the accurate propagation of histone marks during chromosomal replication. This study showed that, in avian DT40 cells, REV1 was required to maintain repressive chromatin histone marks in the vicinity of regions forming G4. In cells lacking REV1, the inclusion of a G4 into a silent locus led to its depression caused by loss of chromatin histone repressive marks. These observations suggest that G4 structures play a critical role in maintaining repressive epigenetic patterns (Sarkies et al 2010). Another publication confirmed these results and observed, by microarray-based gene expression analysis, that the lack of REV1 in DT40 cells causes genome-wide deregulation of G4-dependent transcription (Sarkies et al 2012).

Concerning the role of G4s at the site of the origin replication, a genome-wide analysis of replication origins indicated that most origins overlap with G4 structures (Besnard et al 2012). The authors proposed that G4 motifs might promote the binding of protein complexes to replication origins and influence their activation.

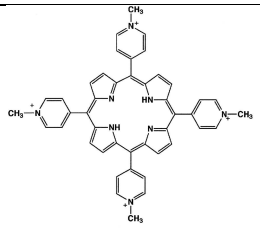
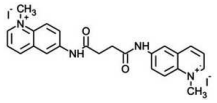
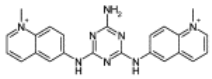
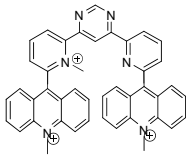
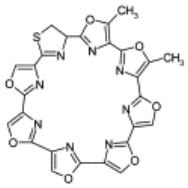
Indirect evidence also suggests that G4 structures are implicated in meiosis. A computational analysis showed, in yeast, an overlap between G4 motifs and preferred meiotic DBS sites (Capra et al 2010). It was also shown that Spo11, the enzyme that generates the DSBs, does not cleave within G4 structures (Pan et al 2011). However, the MRX complex composed by Mre11, Rad50 and Xrs, which act during meiotic DSB formation, has a high affinity for G4 structure *in vitro* (Ghosal and Muniyappa 2005, Ghosal and Muniyappa 2007). These results suggest that G4s may thus participate to the recruitment of MRX to generate DSB formation in the vicinity of G4 regions.

D. G-quadruplexes ligands: biological effects

In cancer, telomerase is over-expressed in 85-90% of the cases (Kim et al 1994). The role of G4 structures in telomeres and their influence in telomerase activity has led to major research efforts towards the development of drugs targeting G4 in telomeres. Different types and generations of ligands have been developed (**Table 3**). Cells treated with G4 ligand show different biological responses depending of cell phenotype, ligand type and concentration

used such as reduction of telomeres size, cell cycle arrest, senescence and/or apoptosis (De Cian et al 2008b).

Table 3: Example of G4 ligands and their pathway.

Name	Structure	Pathway	Ref
TMPyP4		Inducing DNA damage response	(Shi et al 2001)
Bisquinolinium 360A		Telomere aberrations after NHEJ or HR	(Pennarun et al 2005)
12459		Telomere shortening, senescence-like growth arrest and apoptosis	(Gomez et al 2003a)
TAC		Cell cycle arrest and increasing hTERT expression	(Merle et al 2011)
Telomestatin		Cell cycle arrest	(Kim et al 2002)

I. Biological responses induced by G-quadruplexes ligands

In vivo, myeloma cells treated for 7 days with 10 µM of the TMPyP4 ligand present a reduction in telomerase activity of more than 90% (Shammas et al 2003). After 10 days, cells stop to proliferate and after 4 weeks of treatment, the size of telomeres is reduced by 40% compared to non-treated cells. These results suggest that TMPyP4 plays an anti-proliferative function, attributable the reduction of telomeres length due to the telomerase activity inhibition. These results are however confounded by the fact that TMPyP4, one of the first identified synthetic ligands of G4, has broad effects and binds multiple nucleic acid secondary structures other than G4. Using a more specific G4 ligand, 12459, Gomez and collaborators

did not observe a significant reduction of proliferation in the A549 lung cancer cell line over-expressing hTERT (Gomez et al 2003a). Another study has shown that glioma cell lines treated during 10 days with a synthetic G4 ligand showed an large increase in telomeric instability, with degradation of telomeric strands rich in guanines (decrease of 27% in 10 days) and aberrant fusions between telomeres (anaphase bridges), associated with cell cycle arrest (Pennarun et al 2005). Thus, G4 ligands exert direct effects on telomeres, inducing telomere instability and influencing the decrease of telomere size due to inhibition of telomerase activity.

II. Pathways induced by G-quadruplexes ligands

In A549 lung cancer cell lines, low doses of 12459 induce cell cycle arrest only after a very long latency (53 days of treatment) (Riou et al 2002). At this stage, cells show a senescent phenotype with an increase of senescence markers such as β-galactosidase activity. In contrasr, a shorter treatment at a higher drug dose induced massive apoptosis: after 5-12 days of 12459 treatment, surviving cells presented an increase in the expression of Bcl-2, a pro-apoptotic protein (Douarre et al 2005). Apoptosis thus occurred after short-term treatments with concentrations >4 μM and a senescence-like delayed growth arrest after long-term treatments with concentrations <1 μM (Gomez et al 2003b, Riou et al 2002). In addition, after 15 days of a high dose of TMPyP4, over 70% of the cells expressed a high quantity of Annexin V, a marker of apoptosis (Shammas et al 2003).

The decrease of telomere size and their dysfunction induced by G4 ligands lead to the activation of DNA damage signalling. Different activation pathways are implicated depending of the type of ligand and the concentration used. In senescence condition, the 12459 ligand induces a DNA damage pathway via the phosphorylation of Chk1 and p53 S15 (De Cian et al 2008a). In apoptotic conditions, this ligand activates the mitochondrial apoptotic pathway with the phosphorylation of p53 S392. Finally, treatment with another G4 ligand, telomestatin, is associated with increase of ATM and Chk2 phosphorylation with the activation of cell cycle checkpoint proteins $p21^{WAF1}$ and p27 (Tauchi et al 2003). Some resistance to 12459-induced senescence have been described. Resistant 12459-induced senescence cells expressed a high amount of hTERT (Gomez et al 2004). However, inhibition of hTERT did not restore senescence in these cells. Another G4 ligand, 360A, can induce telomere aberrations occurring during or after replication using non-homogous end joining (NHEJ) and HR (Gauthier et al 2012, Pennarun et al 2008). Rad51-dependent HR contributes to specific chromatid-type aberrations such as telomere losses and doublets and DNA-PKcs-dependent

NHEJ induces sister telomere fusions associated with alteration in metaphase-anaphase transition and anaphase bridges (Gauthier et al 2012).

To conclude, G4 ligands induce significant pharmacological responses, which sometimes are counteracted by the development of resistance. Currently, the main question is whether it is possible to develop ligands that preferentially induce apoptosis or senescence. Recently, a new G4 ligand, TAC, was developed (Merle et al 2011). This ligand does not induce a reduction of telomere size. It induces cell cycle arrest in human glioma cell lines and increases hTERT expression. Interestingly, this ligand, associated with a radiation treatment, increases radiation-induced killing compared to only X-ray irradiated cells by increasing DNA DSBs induced by radiation. This observation could lead to new approaches for sensitizing tumour cells to radiation therapy.

OBJECTIVES

Data from genomic databases such as dbSNP suggests that polymorphisms occur in the human genome at an average rate of 1 per 1,910 bases (www.ncbi.nlm.nih.gov/SNP/; (Sachidanandam et al 2001)). In *TP53*, 85 polymorphisms have been validated in the approximately 20 kb of the *TP53* gene and its 3'flanking region, suggesting a rate of one polymorphism per 235 bases, about 8 times higher than in the human genome (p53.iarc.fr; (Petitjean et al 2007)). Thus, *TP53* appears to be highly polymorphic. This may appear counter-intuitive based on the many biological and cellular processes that the p53 protein has been implicated in but may allow a means of expressing alternative RNA transcripts and/or protein isoforms that could provide a means of controlling p53 functions or a selective advantage under some stress or environmental conditions or in certain genetic backgrounds. However, only a few of these polymorphisms have been systematically studied and their possible functional impact is still unclear.

To date, in genome-wide association studies, genetic variations in *TP53* have not been identified as cancer risk factors. The absence of such a strong association does not preclude the possibility that associations may exist in some contexts. One of these contexts is subjects and families that carry a germline *TP53* mutation predisposing to cancer. Several studies (summarized in Introduction, Part I, section II) have provided evidence that age at first cancer onset in *TP53* mutation carriers may vary according to the polymorphic status of *TP53* alleles (Bond et al 2004, Bougeard et al 2006, Marcel et al 2009). However, the biological basis of these effects is not known.

Outside the familial cancer setting, a large number of small-scale studies have addressed the possible effect of specific *TP53* polymorphism on the risk of developing cancers. Meta-analyses of these results have been published for two common SNPs, rs1042522 (Dai et al 2009, Li et al 2009, Matakidou et al 2003, Wang et al 2013b, Zhou et al 2013) and rs17878362 (He et al 2011b, Hu et al 2010b, Hu et al 2010c, Wu et al 2013). These meta-analyses have shown that rare alleles of these polymorphisms are associated with a marginally increased risk of cancer.

In this Thesis, I have focused on one particular polymorphism, rs17878362. This polymorphism consists of a repeat of a 16 bp G-rich sequence located in *TP53* intron 3. The interest in this SNP is motivated by several observations:

(1) It is one of the most common polymorphism in *TP53*, at least in Caucasian (MAF=0.10 in HapMap) and it shows significant differences in its distribution in relation with populations/ethnicity (Garritano et al 2010, Marcel et al 2009);

(2) Previous studies by Janet Hall and colleagues had shown that, in lymphoblastoid cell lines, this polymorphism is associated with a small but reproducible change in the overall level of p53 mRNA (Gemignani et al 2004), and thus could impact of the p53 suppressor activity;

(3) It is located in a structurally complex region of the *TP53* gene, made of small exons (exons 2 and 3) interspaced by very small introns and composed of G-rich strand that can form a G4s structural motif. Intron 3 is 93 bp long and the 16 bp polymorphism thus covers about one-sixth of its total length.

The objectives of this Thesis were:

1) To analyse the association between rs17878362 and risk of cancer by performing a meta-analysis of case-control studies on cancer (Article 1: "A meta-analysis of cancer risk associated with the *TP53* intron 3 duplication polymorphism (rs17878362): geographic and tumor-specific effects" published in Cell Death and Disease; 14:4:e492. doi: 10.1038/cddis.2013.24);
2) To examine the impact of rs17878362 on the age at cancer onset in germline *TP53* mutation carriers using a cohort of 402 Brazilian members (Article 2: Age at cancer onset in germline *TP53* mutation carriers: association with polymorphisms in predicted G4 structures" submitted in Carcinogenesis);

To determine whether the G-rich region of intron 3 can form detectable G4 structure and to investigate how this structure may affect the alternative splicing of p53 mRNA to generate transcripts encoding p53 protein isoforms (Article 3: G-quadruplex structures in *TP53* intron 3: role in alternative splicing and in production of p53 mRNA isoforms" published in Carcinogenesis (32(3):271-8); Article 4: Impact of G-quadruplex structures and the polymorphisms rs17878362 and rs1642785 on the expression of *TP53* transcripts (included in draft form)) .

RESULTS: ROLE OF P53 POLYMORPHISMS IN P53 EXPRESSION AND CANCER SUSCEPTIBILITY

Part I. A meta-analysis of cancer risk associated with rs17878362

This section summarizes the data presented in the annexed manuscript "A meta-analysis of cancer risk associated with the TP53 intron 3 duplication polymorphism (rs17878362): geographic and tumor-specific effects" published in Cell Death and Disease (14:4:e492. doi: 10.1038/cddis.2013.24). Figure and Table numbers correspond to those from the publication.

Background

In *TP53*, 85 polymorphisms are described and many present geographic and population frequency variations (Garritano et al 2010). However, their association with cancer risk is inconsistent, with important variations according to study design, population or type of cancer. The most studied polymorphism is an exonic polymorphism located in the exon 4 at codon 72, rs1042522 (G>C, R>P) (Matlashewski et al 1987). *In vitro*, the two variants R and P present differences in their biological functions (Suspitsin et al 2003), but the role of this polymorphism in susceptibility to cancer risk is still controversial.

The most common intronic polymorphism in *TP53* is a deletion/insertion of 16 bp in intron 3 (rs17878362). The common allele A1 carries one copy of the 16bp motif, whereas the rare allele A2 carries a tandem repeat (two copies) of the 16 bp sequence (Lazar et al 1993). Several individual studies have suggested an association between rs17878362 and the risk of breast, ovarian and lung cancers. However, negative studies have also been reported and it was unclear whether these apparent associations could be affected by publication biases. A meta-analysis had identified a small but significant increase of cancer risk associated with the A2A2 carriers (odd ratio (OR)=1.14; 95% CI=[1.02-1.27]) (Hu et al 2010b) and another one had observed the same result in breast cancer (OR=1.70, 95% CI=[1.20-2.37]) (Hu et al 2010c). However, this conclusion was controversial due to differences between the selected data and the original data (Lu et al 2011).

To understand the influence of the A2 allele of rs17878362 on cancer susceptibility, we performed a detailed meta-analysis on a total of 25 publications regrouping 10,786 cancer cases and 11,377 controls. First we analysed the overall risk associated with the A2 allele.

Secondly, we performed sub-group analyses to determine the risk associated with A2 allele with in relation with ethnicity, geographical origin of the patients or the cancer type considered. Data for rs1042522 and rs1625895 (intron 6, G>A) polymorphisms that were reported in the same publications were also analysed.

Results

First, we have observed that in each control population the rs17878362 allele distribution is in Hardy-Weinberg equilibrium (**Table 1**). Among the 25 publications, 9 studies have observed a significant increase of cancer risk associated with the A2A2 variant compared to the A1A1 variant and 16 have not reported any association between either variant and cancer susceptibility. Overall, our meta-analysis showed that the rs17878362 A2A2 genotype was associated with a statistically significant increase in the overall risk of cancer risk compared to the A1A1 genotype (OR=1.45, 95%CI=[1.22-1.74]) (**Table 2**). This association was confirmed using a cumulative inclusion over time analysis. There was no significant association for A1A2 genotype as compared to A1A1. These effects were not accountable to a publication bias.

Secondly, we have analysed the variation of rs17878362 cancer susceptibility in relation with population and geographical diversity (**Figure 18A of this Thesis**). The overall group of cases and controls was sub-divided into 4 sub-groups of at least 1,000 cases and 1,000 controls (**Table 2**). In the United States, no association was observed in a series of 4,125 cases and 4,716 controls. In Northern Europe and in India, the A2A2 genotype was associated with an increased risk of cancer (OR=1.70; 95%CI=[1.26-2.31] and OR=1.63, 95%CI=[1.10-2.42] respectively) but no effect was seen for the A1A2 genotype compared to A1A1 genotype. In Mediterranean countries, both the A1A2 and the A2A2 genotypes were associated with a significant increase of the risk compared to the A1A1 genotype (OR=1.25, 95%CI=[1.03-1.51] and OR=2.54, 95%CI=[1.53-4.54], respectively), suggesting an allelic A2 dose-dependent cancer susceptibility in this population.

Thirdly, we have exanimated the role of the rs17878362 in cancer risk depending upon cancer site/type (**Figure 18B of this Thesis**). From the data set, three pathology groups could be distinguished, each accounting for at least 1,000 cases and 1,000 controls. No association was observed between lung cancer and rs17878362 genotypes, despite a total of 4,054 cases and 4,045 controls (**Table 3**). An increase of breast cancer risk was observed with the A1A2 genotype (OR=1.18, 95%CI=[1.02-1.37] but not with the A2A2 genotype, compared to the A1A1 genotype. For colorectal cancer, the A2A2 genotype was associated with an increase of

risk compared to the A1A1 genotype (OR=1.67, 95%CI=[1.02-2.74]) but not to the A1A2 genotype.

Fourthly, we have performed a meta-analysis for the rs1042522 and the rs1625895 polymorphisms based on data reported in the set of papers being considered. This gave us data sets containing a total of 8,517 cases and 9,311 controls for rs1042522 and 5,011 cases and 5,100 controls for rs1625895. We found that heterozygote genotypes of rs1042522 or rs1625895 were associated with an increase of cancer risk (rs10425222, RP *versus* RR OR=1.16, 95%CI=[1.05-1.18] and rs1625895 GA *versus* GG OR=1.19, 95%CI=[1.02-1.40]).

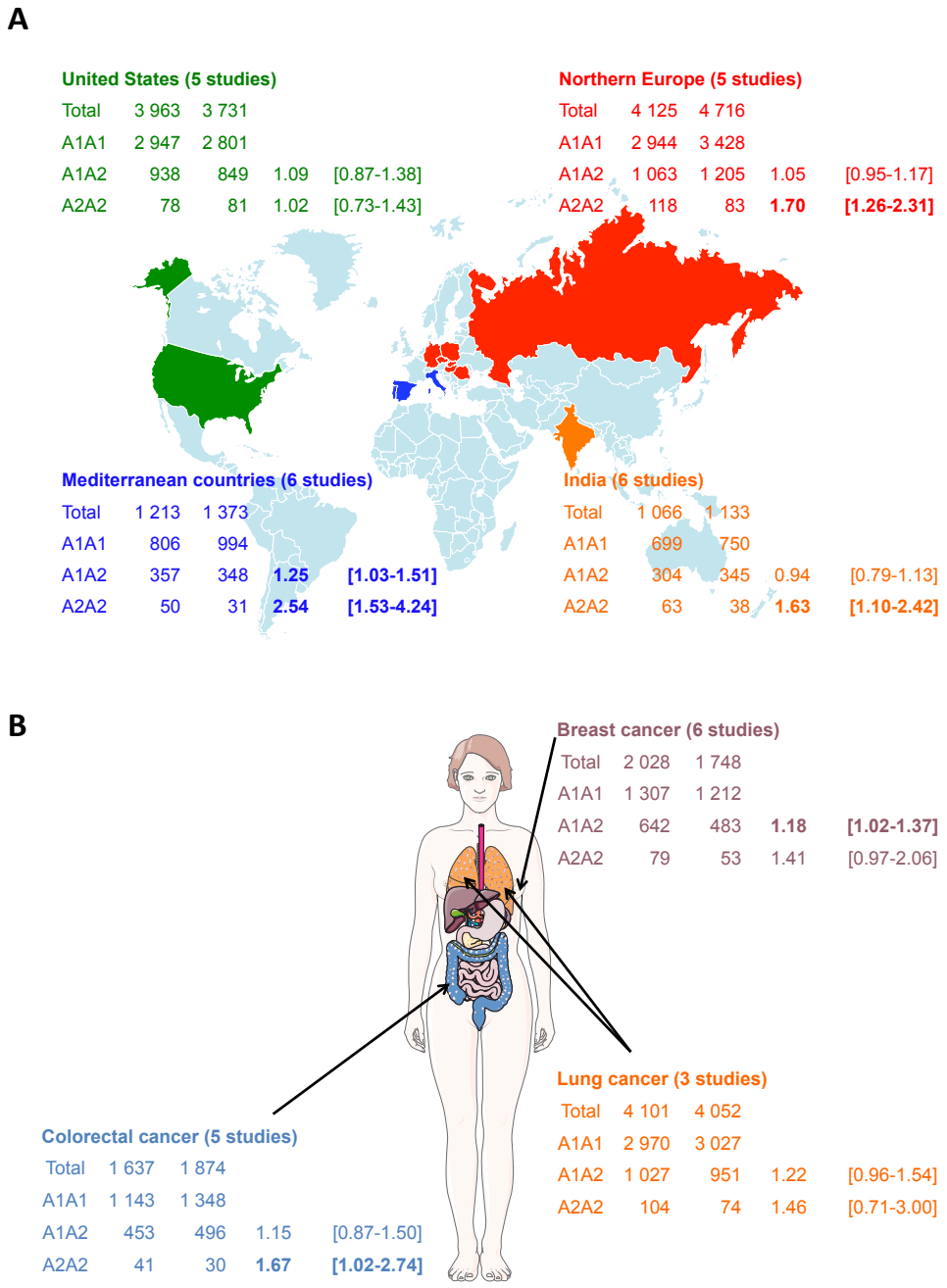


Figure 18: Schematic representation for the meta-analysis results for the *TP53* rs17878362 polymorphism (**A**) by geographical origin and (**B**) by cancer type. In bold: statistical significant results.

Conclusions

The A2A2 rs17878362 genotype is associated with a small but significant increase of cancer susceptibility, with variations in the magnitude and strength of the effect depending upon ethnicity, geographical area and type of cancer. Overall, rs17878362 seems to be a better marker of cancer susceptibility than rs1042522 (for rs17878362, OR=1.45, 95%CI=[1.22-1.74] and for rs1042522, OR=1.16, 95%CI=[1.05-1.18]).

Citation: Cell Death and Disease (2013) 4, e492; doi:10.1038/cddis.2013.24
© 2013 Macmillan Publishers Limited All rights reserved 2041-4889/13
www.nature.com/cddis

A meta-analysis of cancer risk associated with the *TP53* intron 3 duplication polymorphism (rs17878362): geographic and tumor-specific effects

C Sagne[1,2,3], V Marcel[1,2,4], A Amadou[3], P Hainaut[3,5], M Olivier[3] and J Hall*[,1,2]

We have performed a meta-analysis of cancer risk associated with the rs17878362 polymorphism of the *TP53* suppressor gene (*PIN3*, (polymorphism in intron 3), 16 bp sequence insertion/duplication in intron 3), using a compilation of a total of 25 published studies with 10 786 cases and 11 760 controls. Homozygote carriers of the duplicated allele (A2A2) had a significantly increased cancer risk compared with A1A1 carriers (aggregated odds ratio (OR) = 1.45, 95% confidence interval (CI) = 1.22–1.74). However, there was no significant effect for the A1A2 heterozygotes (A1A2 *versus* A1A1 aggregated OR = 1.08, 95% CI = 0.99–1.18). No significant heterogeneity or publication bias was detected in the data set analysed. When comparing populations groups, increased cancer risk was associated with A2A2 carriage in Indian, Mediterranean and Northern Europe populations but not in the Caucasian population of the United States. Analysis by cancer site showed an increased risk for A2A2 carriers for breast and colorectal, but not for lung cancers. These results support that the A2A2 genotype of rs17878362 is associated with increased cancer risk, with population and tumour-specific effects.

Cell Death and Disease (2013) **4,** e492; doi:10.1038/cddis.2013.24; published online 14 February 2013
Subject Category: Cancer

The *TP53* gene (OMIM 191170), encoding the p53 protein, is frequently inactivated in sporadic human tumours, disabling a wide range of anti-proliferative responses regulating cell cycle progression, apoptosis, autophagy, differentiation, senescence, DNA repair and oxidative metabolism.[1–4] The activity of p53 is regulated by multiple transcriptional, post-transcriptional, translational and post-translational mechanisms in response to a wide range of physical and biological stresses, endowing this protein with a pivotal role in preventing DNA replication and cell division in conditions that threaten genetic integrity.[1,5–7] Among these mechanisms, the expression of p53 as multiple protein isoforms with different N- and/or C-terminal domains has recently emerged as a form of regulation that may participate in the diversity of the repertoire of biological effects mediated by p53 (reviewed in Marcel *et al*[8]).

Close to 100 genetic polymorphisms have been identified in TP53 (listed at http://p53.iarc.fr),[9] many of which show geographic and population frequency variations. However, their effects on cancer risk appear to be inconsistent across studies.[10,11] The most studied polymorphism is a single-nucleotide polymorphism (SNP) in exon 4 encoding an arginine (R) or a proline (P) at codon 72 (rs1042522, G > C, R > P at codon 72, PEX4 (polymorphism in exon 4)).[12] There is *in vitro* evidence that the rs1042522 R72 and P72 p53 protein variants differ by their biological activities.[13,14] However, results from systematic studies and meta-analyses have failed to identify a consistent association with cancer risk.[15–19]

The most common intronic variation in *TP53* is a 16-base pair (bp)[11] insertion/duplication in intron 3 (rs17878362, consisting of one copy (A1 allele) or two copies (A2 allele) of the sequence ACCTGGAGGGCTGGGG, PIN3 (polymorphism in intron 3 (rs17878362))).[20] Several case–control studies have reported an increased risk of various cancer types associated with the rs17878362 A2 allele in Caucasians, with the most consistent association reported for breast,[21,22] and colorectal cancers.[23,24] A recent meta-analysis identified a small but significant increase in overall cancer risk of 14% (95% confidence interval (CI) = 1.02–1.27) in homozygote carriers of the A2 allele.[25] However, this conclusion was questioned because of apparent discrepancies between data selected for meta-analysis and the original publications.[26] At the mechanistic level, there is some evidence that rs17878362 may have an impact on the levels[23] and alternative splicing of the *TP53* mRNA, and thus on the ratios of p53 protein isoforms.[8] However, the precise mechanisms underlying an increased cancer risk associated with the rs17878362 A2 allele are not clearly understood.

[1]INSERM U612, Bât 110-112, Centre Universitaire, Orsay, France; [2]Institut Curie, Centre de Recherche, Bât 110-112, Centre Universitaire, Orsay, France and [3]Molecular Carcinogenesis Group, International Agency for Research on Cancer, 150 cours Albert Thomas, Lyon, France
*Corresponding author: Dr J Hall, Institut Curie, Centre de Recherche, INSERM U612, Bât 110-112, Centre Universitaire, Orsay 91405, France.
Tel: + 33 (0)1 69 86 30 61; Fax: + 33 (0)1 69 07 53 27; E-mail: janet.hall@curie.fr
[4]Present address: Cancer Research Centre of Lyon, UMR INSERM 1052 CNRS 5286, Centre Léon Bérard, FNCLCC, Nuclear Domains and Pathologies, 28 rue Laennec, 69373 Lyon, France.
[5]Present address: International Prevention Research Institute, 95 cours Lafayette, 69006 Lyon, France.
Keywords: TP53; rs17878362; intron3; PIN3
Abbreviations: Bp, base pair; CI, confidence interval; MAF, minor allele frequency; OR, odds ratio; PEX4, polymorphism in exon 4 (rs1042522); PIN3, polymorphism in intron 3 (rs17878362); PIN6, polymorphism in intron 6 (rs1625895); SNP, single-nucleotide polymorphism
Received 08.10.12; revised 18.12.12; accepted 21.12.12; Edited by A Stephanou

To assess whether the rs17878362 polymorphism may represent a potentially important and relevant genetic marker contributing to cancer susceptibility, we have performed an independent, two-stage meta-analysis on a total of 10 786 cancer cases and 11 377 controls from 25 published case–control studies. First, we have analysed the overall cancer risk associated with the A2 allele and second we have performed sub-group analyses to examine this association in different populations and for specific cancer types. Data for the rs1042522 and rs1625895 (rs1625895, intron 6, G > A, PIN6 (polymorphism in intron 6)) variant alleles in relation to cancer risk was also compiled and analysed from the same publication set to assess their potential confounding effect.

Results

Characteristics of selected publications. A total of 25 publications out of the 299 identified met the necessary inclusion criteria for the meta-analysis that required the reporting of odds ratio (OR) data and information on the frequency of each allele, which has been verified to be in Hardy–Weinberg equilibrium in each control population (Table 1 and Supplementary Table 1). Two studies[24,27] used the same control populations and they were included only once to avoid over-representation. Overall, nine individual studies reported a significant increase in cancer risk associated with the rs17878362 A2 allele compared with the A1 allele, 16 showed no statistical association between either allele and cancer susceptibility and no study reported an association between the A2 allele and decreased cancer risk (Table 1).

The A2A2 genotype of rs17878362 polymorphism increases cancer risk. On the basis of the results of the heterogeneity testing, a random model was used for the meta-analysis to assess the overall cancer risk in A2 allele carriers (A1A2 or A2A2) (Table 2).[28] The rs17878362 minor allele frequency (MAF) was inferior to 0.17 in control subjects in the different sub-groups and allele ratios were compatible with Hardy–Weinberg equilibrium (data not presented). No significant association with cancer risk was found in the heterozygous A1A2 carriers compared with the homozygous A1A1 carriers (A1A2 *versus* A1A1 aggregated OR = 1.08, 95% CI = 0.99–1.18), however, a significantly increased risk was found for the A2A2 carriers (A2A2 *versus* A1A1 aggregated OR = 1.45, 95% CI = 1.22–1.74). Leave-one-out analyses showed that the aggregated OR for the A1A2 *versus* A1A1 genotypes varied between 1.06 and 1.10 (95% CI between 0.97 and 1.20) and for the A2A2 *versus* A1A1 genotypes between 1.37 and 1.55 (95% CI between 1.15 and 1.91) (Supplementary Table 2). The Egger's bias coefficient was determined to assess a possible bias introduced by any single study. The ORs for Egger's bias coefficient were 0.07, (95% CI = 1.32–1.46) for the A1A2 genotype, and 0.79 (95% CI = 0.62–2.19) for the A2A2 genotype, suggesting no significant publication bias.

To assess the possibility that the overall result might be biased by initial publications reporting a large effect, a cumulative inclusion over time analysis was conducted. For the A1A2 genotype, the first set of studies (four reports published before 2006) had the highest ORs for the association between the A1A2 genotype and cancer risk (Supplementary Table 3). Lower values were reported in the following 2 years, after which the overall result remained stable (aggregated OR 1.08 for 2010 and 2011). For the A2A2 allele, the time trend for the aggregated OR showed little variation, with ORs between 1.37 and 1.45 being reported since 2007, in support of the robustness of this association.

rs17878362-related cancer risk is dependent on ethnicity and geographical origin. To investigate whether rs17878362 related cancer susceptibility varies between populations and geographical regions, the data from the 25 studies were divided into four geographical sub-groups (India, Northern Europe, North America and the Mediterranean area) each containing at least 1000 cases and 1000 controls from a minimum of five independent case–control studies (Table 2, see Table 1 and Supplementary Table 1 for population details). Differences in genotype distribution were noted with that of the Indian controls being statistically different from the three other control sub-groups (India *versus* Mediterranean countries: χ^2 *P*-value 0.01, India *versus* Northern Europe or United States: χ^2 *P*-values < 0.01). The genotype distribution found in the United States' controls (reported as a Caucasian population in the original publications) was also different from that of the Northern Europe controls (χ^2 *P*-value 0.01). No difference in genotype distribution was observed between controls from the Mediterranean and from Northern Europe or United States (Mediterranean countries *versus* Northern Europe: χ^2 *P*-value 0.49, Mediterranean countries *versus* United States: χ^2 *P*-value 0.14).

In this geographical sub-group analysis, the homozygous A2A2 genotype was associated with an increased cancer risk in Indian (A2A2 *versus* A1A1 aggregated OR = 1.63, 95% CI = 1.10–2.42) and Northern Europe populations (A2A2 *versus* A1A1 aggregated OR = 1.70, 95% CI = 1.26–2.31) compared with the homozygous A1A1 genotype. For the Mediterranean population, both the A1A2 and A2A2 genotypes were associated with increased cancer susceptibility in an A2 allelic dose-dependent manner (A1A2 *versus* A1A1 aggregated OR = 1.25, 95% CI = 1.03–1.51; A2A2 *versus* A1A1 aggregated OR = 2.54, 95% CI = 1.53–4.24, *P*-trend < 0.01). In contrast, in the United States' sub-group (3,963 cases and 3,731 controls), no increased cancer susceptibility was associated with carriage of the rs17878362 A2 allele (A1A2 *versus* A1A1 aggregated OR = 1.09, 95% CI = 0.87–1.38; A2A2 *versus* A1A1 aggregated OR = 1.02, 95% CI = 0.73–1.43).

rs17878362-related cancer risk is dependent on cancer type. The risk of developing cancer was assessed for three cancer types: lung, colon and breast, with over 1600 cases and controls included in the analysis (Table 3). For colorectal cancer, homozygous A2A2 carriage was associated with increased susceptibility compared with homozygous A1A1 carriage (A2A2 *versus* A1A1 aggregated

Table 1 Characteristics of the 25 case–control studies selected for TP53 rs17878362 (PIN3) polymorphism meta-analysis

Study numbers and study		Cancer type	Cases	Controls	Population	Minor allele frequency in controls (MAF)			Hardy–Weinberg equilibrium *P*-value for controls		
						rs17878362 (A2)	rs1042522 (P72)	rs1625895 (A)	rs17878362 (A2)	rs1042522 (P72)	rs1625895 (A)
1	Jha *et al.*[40 a]	Glial tissue	84	76	India	0.18	0.55	NA	0.23	0.01[b]	NA
2	Umar *et al.*[41 a]	Oesophagus	255	255	India	0.19	NA	NA	0.33	NA	NA
3	Alawadi *et al.*[42 a]	Breast	229	133	NC	0.31	0.44	NA	0.58	0.01[b]	NA
4	Mittal *et al.*[43 a]	Prostate	177	265	India	0.15	0.24	0.21	0.12	0.28	0.11
5	Malik *et al.*[27 c]	Oesophagus	135	[d]195	India	0.21	NA	NA	0.08	NA	NA
6	Malik *et al.*[27 c]	Gastric	108	[d]195	India	0.21	NA	NA	0.08	NA	NA
7	Naccarati *et al.*[44 a]	Pancreas	240	743	Northern Europe	0.16	0.29	NA	0.10	0.40	NA
8	Polakova *et al.*[45 a]	Colon	612	613	Northern Europe	0.14	0.27	NA	0.15	0.52	NA
9	Ashton *et al.*[30 a]	Endometrial	190	291	NC	0.14	0.24	0.11	0.81	0.97	0.12
10	de Feo *et al.*[46 a]	Gastric	114	295	Mediterranean	0.16	0.25	0.20	0.35	0.13	0.15
11	Hrstka *et al.*[47 a]	Breast	117	108	Northern Europe	0.14	0.45	0.13	0.46	0.00[b]	0.78
12	Gaudet *et al.*[48 a]	Breast	578	390	United States	0.16	0.74	0.85	0.85	0.08	0.93
13	Costa *et al.*[21 c]	Breast	191	216	Mediterranean	0.17	0.17	NA	0.29	0.29	NA
14	Ye *et al.*[49 a]	Bladder	636	618	United States	0.15	0.22	0.15	0.29	0.00[b]	0.13
15	de Vecchi *et al.*[50 a]	Breast	350	352	Mediterranean	0.15	0.23	NA	0.62	0.23	NA
16	Chen *et al.*[51 a]	Head and neck	821	818	United States	0.14	0.27	0.12	0.75	0.07	0.67
17	Tan *et al.*[52 a]	Colon	467	563	Northern Europe	0.17	0.22	NA	0.23	0.98	NA
18	Wang *et al.*[53 a]	Lung	1412	1363	United States	0.13	0.26	0.12	0.45	0.54	0.14
19	Hung *et al.*[54 c]	Lung	2126	2140	Northern Europe	0.13	0.27	NA	0.50	0.74	NA
20	Perfumo *et al.*[24 a]	Colon	60	188[e]	Mediterranean	0.15	0.20	NA	0.21	0.81	NA
21	Perfumo *et al.*[24 c]	Colon	124	188[e]	Mediterranean	0.15	0.20	NA	0.21	0.81	NA
22	Mitra *et al.*[55 a]	Oral cancer	307	342	India	0.19	0.48	NA	0.56	0.20	NA
23	Gemignani *et al.*[23 c]	Colon	374	322	Mediterranean	0.12	0.21	NA	0.60	0.09	NA
24	Wang-Gohrke *et al.*[22 c]	Breast	563	549	Northern Europe	0.16	0.26	0.15	0.92	0.49	0.60
25	Wu *et al.*[56 c]	Lung	516	542	United States	0.10	0.20	0.12	0.05	0.01[b]	0.18

Abbreviations: NA, not available; NC, not classified
[a]no significant increase in cancer risk associated with rs17878362 (TP53 PIN3)
[b]*P*-value <0.05 indicates a Hardy–Weinberg disequilibrium: study exclusion
[c]Significant increase in cancer risk associated with rs17878362 (TP53 PIN3)
[d]Same control population
[e]Same control population

OR = 1.67, 95% CI = 1.02–2.74) (Table 3). A slight but significant increased breast cancer risk was observed in the heterozygous A1A2 carriers compared with the A1A1 carriers (A1A2 *versus* A1A1 aggregated OR = 1.18, 95% CI = 1.02–1.37). However, no altered breast cancer risk was seen in the A2A2 carriers (A2A2 *versus* A1A1 aggregated OR = 1.41, 95% CI = 0.97–2.06), although a significant trend towards increased cancer risk was noted as the number of A2 alleles carried was increased (*P*-trend <0.01). No increased risk of lung cancer was observed for any genotype despite the inclusion of 4101 cases and 4052 controls in the analysis (A2A2 *versus* A1A1 aggregated OR = 1.46, 95% CI = 0.71–3.00).

Association of rs1042522 and rs1625895 genotypes with cancer susceptibility. Among the 25 selected publications, several have analysed cancer risk associated with the rs1042522 and rs1625895 variant alleles (Table 1). For rs1625895, the 10 studies reporting rs1625895-related ORs showed rs1625895 allele ratios compatible with Hardy–Weinberg equilibrium, allowing the pooling of 5011 cancer cases and 5100 controls. For the rs1042522 polymorphism, 8517 cases and 9311 controls were pooled from 17 studies (Supplementary Table 4), while 5 other studies were excluded as the allele ratios for rs1042522 in controls were not compatible with Hardy–Weinberg equilibrium (Table 1). When compared with the respective common homozygous carriers, a small but significant association with cancer risk was observed for heterozygous carriers of the variant allele (rs1042522 R72/P72 *versus* R72/R72 aggregated OR = 1.16, 95% CI = 1.05–1.18; rs1625895 GA *versus* GG aggregated OR = 1.19, 95% CI = 1.02–1.40) (Supplementary Table 4). However, no increased risk was observed in association with the homozygous carriages of the variant alleles at either position.

Discussion

A large number of studies have addressed the association of common *TP53* polymorphisms with cancer risk (reviewed in Whibley *et al.*[10]). Overall, the reported effects are of small amplitude and many studies have reported contradictory results that may result from many causes: small numbers of cases and controls and thus limited statistical power, the selection of specific tumour types, differences between populations and the lack of reliability in SNP genotyping, in particular in earlier studies. Of the *TP53* intronic polymorphisms rs17878362 is the most studied. In this meta-analysis, based on 10 786 cases and 11 377 controls we detected an aggregated OR of 1.45 (95% CI = 1.22–1.74) for increased cancer risk in homozygous carriers of the rare rs17878362 A2 genotype as compared with homozygous carriers of the common A1 genotype. However, no risk was observed when A2A1 carriers were compared with the A1A1 carriers, suggesting that the increased risk associated with rs17878362 follows a recessive model. This result is in

Table 2 Meta-analysis results for the selected case–control studies focused on the TP53 rs17878362 polymorphism

Genotypes	Cases, *n* (%)	Controls, *n* (%)	Heterogeneity, *P*-value	OR	(95% CI)	*P*-trend[a]
Overall (25 studies, MAF = 0.15)						
Total	10 786 (100.0)	11 377 (100.0)				
A1A1	7 639 (70.8)	8 254 (72.5)		1.00	—	<0.01
A1A2	2 823 (26.2)	2 871 (25.2)	0.03[b]	1.08	(0.99–1.18)	
A2A2	324 (3.0)	252 (2.3)	0.06[b]	1.45	(1.22–1.74)	
Geographical origin of studies India (study numbers: 1, 2, 4, 5, 6, 22: MAF = 0.19)						
Total	1 066 (100.0)	1 133 (100.0)				
A1A1	699 (65.6)	750 (66.2)		1.00	–	0.19
A1A2	304 (28.5)	345 (30.5)	0.54[c]	0.94	(0.79–1.13)	
A2A2	63 (5.9)	38 (3.3)	0.07[c]	1.63	(1.10–2.42)	
Mediterranean countries (study numbers: 10, 13, 15, 20, 21, 23: MAF = 0.15)						
Total	1 213 (100.0)	1 373 (100.0)				
A1A1	806 (66.4)	994 (72.4)		1.00	—	<0.01
A1A2	357 (29.4)	348 (25.4)	0.475[c]	1.25	(1.03–1.51)	
A2A2	50 (4.2)	31 (2.2)	0.701[c]	2.54	(1.53–4.24)	
Northern Europe (study numbers: 7, 8, 11, 17, 19, 24; MAF = 0.15)						
Total	4 125 (100.0)	4 716 (100.0)				
A1A1	2 944 (71.4)	3 428 (72.7)		1.00	—	0.03
A1A2	1 063 (25.8)	1 205 (25.5)	0.247[c]	1.05	(0.95–1.17)	
A2A2	118 (2.8)	83 (1.8)	0.795[c]	1.70	(1.26–2.31)	
United States (study numbers: 12, 14, 16, 18, 25; MAF = 0.14)						
Total	3 963 (100.0)	3 731 (100.0)				
A1A1	2 947 (74.3)	2 801 (75.0)		1.00	—	0.65
A1A2	938 (23.7)	849 (22.8)	0.003[b]	1.09	(0.87–1.38)	
A2A2	78 (2.0)	81 (2.2)	0.344[c]	1.02	(0.73–1.43)	

Abbreviations: CI, confidence interval; MAF, minor allele frequency; OR, odds ratio
[a]Fisher's exact test
[b]Heterogeneity *P*-value ≤0.05: performed random model for meta-analysis
[c]Heterogeneity *P*-value >0.05: performed fixed model for meta-analysis

Table 3 Meta-analysis results for the TP53 rs17878362 polymorphism by cancer type

Genotypes	Cases, *n* (%)	Controls, *n* (%)	Heterogeneity *P*-value	OR	(95% CI)	*P*-trend[a]
Breast (Study numbers: 3, 11, 12, 13, 15, 25; MAF = 0.17)						
Total	2 028 (100.0)	1 748 (100.0)				
A1A1	1 307 (64.5)	1 212 (69.3)		1.00	—	<0.01
A1A2	642 (31.7)	483 (27.6)	0.57[b]	1.18	(1.02–1.37)	
A2A2	79 (3.9)	53 (3.0)	0.08[b]	1.41	(0.97–2.06)	
Colon (study numbers: 8, 17, 20, 21, 23; MAF = 0.15)						
Total	1 637 (100.0)	1 686 (100.0)				
A1A1	1 143 (69.8)	1 214 (72.0)		1.00	—	0.08
A1A2	453 (27.7)	444 (26.3)	0.04*	1.15	(0.87–1.50)	
A2A2	41 (2.5)	28 (1.7)	0.33[b]	1.67	(1.02–2.74)	
Lung (study numbers: 18, 19, 25; MAF = 0.13)						
Total	4 054 (100.0)	4 045 (100)				
A1A1	2 977 (73.4)	3 076 (76.0)		1.00	—	<0.01
A1A2	979 (24.2)	898 (22.2)	0.02*	1.22	(0.96–1.54)	
A2A2	98 (2.4)	71 (1.8)	0.03*	1.46	(0.71–3.00)	

Abbreviations: CI, confidence interval; MAF, minor allele frequency; OR, odds ratio
*Heterogeneity *P*-value ≤0.05: performed random model for meta-analysis
[a]Fisher's exact test
[b]Heterogeneity *P*-value >0.05: performed fixed model for meta-analysis

agreement with the recent meta-analysis of Hu and collaborators, despite the fact that the two studies differed in the selection and analysis of data to be included as we used original ORs reported in each publication, which was not the case in the study of Hu *et al*.[25,26] When sub-grouping data according to tumour site, different associations were seen for breast, colon and lung cancer, which were the only three tumour sites for which over 1600 cases and controls was available with the data drawn from at least three different reports. These differences suggest that the contribution of

rs17878362 to susceptibility might be different from one tumour type to the other. In the case of breast cancer, the increased risk was associated only with the heterozygosity status. Tumour type heterogeneity, in term of pathology and molecular profiles including the frequency of *TP53* mutations, may explain these results although this clearly needs further evaluation.[29,30] The lack of significant effect in lung cancer might reflect the overwhelming effect of tobacco smoke as a causative risk factor, masking the much smaller contribution of genetic susceptibility factors such as rs17878362.

Few studies have investigated the impact of rs17878362 on cancer susceptibility with respect to the geographical origin of the cohorts. Here, the observed difference across countries could be due to a different distribution in rs17878362 polymorphism between different ethnic groups. Indeed, Sjalander *et al*.[31] reported a difference in rs17878362 distribution across latitudes, between Swedish, Asian and Mongolian populations, which is independent of rs1042522 distribution. However, in the present meta-analysis, although some differences in the rs17878362 A2 allele frequency were seen between the different geographical regions, no heterogeneity was observed in the overall data set independently of any geographical consideration. Thus, the difference in rs17878362 A2 allele-related cancer susceptibility in the different countries suggests that additional factors, such as environmental factors, lifestyle and other genetic modifiers, may modulate cancer susceptibility associated with this allele.

Several studies have shown that the rs17878362 polymorphism is in linkage disequilibrium with other common TP53 SNPs, including rs1042522.[31,32] In a previous study, we have haplotyped rs17878362 and rs1042522 in a group of mostly Caucasian subjects from Brazil and reported that 71% of the tested population carried the haplotype combining rs17878362 A1 and rs1042522 R72, whereas the haplotype rs17878362 A2/rs1042522 R72 was detected in only 1.5% of the population.[33] In contrast, the A1/P72 and A2/P72 haplotypes were almost equally represented (15 and 12.5% of the population, respectively). This observation suggests that the rs17878362 A2 allele most frequently occurs on a haplotype that also contains rs1042522 P72,[33] raising the possibility that the susceptibility associated with rs17878362 might be driven, or confounded, by other common TP53 SNPs. To evaluate this possibility, we have used the data compiled from the same set of publications to assessed cancer risk associated with rs1042522 and rs1625895 variants in the same data set. The aggregated ORs for the overall analysis showed that the heterozygote carriers of either variant allele had an increased cancer risk, consistent with several previous meta-analyses.[14,25,34] However, the effects observed for rs1042522 and rs1625895 were clearly smaller than for rs17878362 and were observed only in heterozygote carriers of rs1042522 or rs1625895, whereas the effect of rs17878362 appears to follow a recessive model. This would suggest that if rs1042522 and rs1625895 contribute to susceptibility, this effect could occur independently of their association with rs17878362. These results should be interpreted with caution, as no corrections for multiple testing have been performed. Indeed, it is not possible to calculate the number of tests carried in the original papers in order to correct for multiple comparisons. Moreover, it has to be recognized that this analysis was not designed to specifically assess the cancer risk of these two alleles. The linkage disequilibrium between rs17878362 (tagged by rs2909430, which is in linkage disequilibrium with rs17878362, $r2 > 0.9$), rs1042522 and rs1625895 also shows ethnic differences as is reflected in the haplotype frequencies calculated based on published data[9] for three different HapMap populations (Supplementary Table 5). The most frequently found haplotype in the Caucasian and Asian HapMap populations was found to be rs17878362 A1/rs1042522 R72/rs1625895 G (78.13% of the Caucasian and 53.70% of the Asian population), while this only represented 31.67% of the haplotypes seen in the African population. The rs17878362 A1/rs1042522 P72/rs1625895 G haplotype was more frequent in the Asian (43.83%) and African (38.33%) populations than the Caucasian population (11.46%), while the rs17878362 A2/rs1042522 P72/rs1625895 A haplotype was seen in only 1.85% of Asian population compared with 9.37 and 26.11% of the Caucasian and African populations, respectively. Clearly further studies analysing TP53 haplotypes are needed to clarify the specific contribution of each of these common SNPs to cancer susceptibility.

The mechanistic basis of this altered risk associated with the carriage of the rs17878362 A2 allele is still poorly understood. Some evidence links rs17878362 status to differential expression of different p53 isoforms. In lymphoblastoid cell lines established from breast cancer patients the A1A1 genotype was associated with higher constitutive levels of *TP53* mRNA than for the A1A2 and A2A2 alleles.[28] Recently, we have shown that *TP53* intron 3 is involved in the splicing regulation of the *TP53* intron 2, influencing the generation of the fully spliced p53 (FSp53) and the intron-2-retaining p53 (p53I2) mRNA transcripts.[7] These transcripts generate the canonical p53 protein and the N-truncated Δ40p53 isoform, respectively, the latter being a regulator of p53 activity.[8] Using *in silico* algorithms, biophysical measurements and *in vitro* assays we have shown that the RNA sequences present in *TP53* intron 3 pre-RNA can form G-quadruplex structures, whose stability alters the balance of FSp53/p53I2 mRNA species through the modulation of intron 2 splicing.[7] On the basis of the same in silico algorithms, it appears that the rs17878362 duplication may alter the topology of the G4 structures in intron 3 that may impact on the FSp53/p53I2 balance. As the Δ40p53 isoform encoded by the p53I2 mRNA can inhibit p53 transcriptional activity and growth suppressive activity *in vitro* and appears to represent the main form of p53 expressed in mouse embryonic stem cells.[2,35–37] It is possible that the presence of the rs17878362 A2 variant allele could impact on p53 regulatory activity through the modulation of TP53 mRNA transcript patterns, subsequent isoform expression and maintenance of stem cell-like phenotype. Recent evidence suggesting that mRNA encoding Δ40p53 and Δ133p53 isoforms are over-expressed in some forms of ovarian carcinoma is in support of the hypothesis that changes in expression of these isoforms may contribute to carcinogenesis.[38] The mechanism by which the rs17878362 polymorphism modulates cancer risk needs to be fully addressed in appropriate functional genetics studies.

Materials and methods

Literature search and selection criteria. Publications relative to the association between the rs17878362 polymorphism and cancer risk examined in case–control studies were identified using two databases: Pubmed Central (NCBI, NIH) (http://www.nml.ncbi.gov/pubmed) and Web of Science (Thomson Reuters) (http://apps.webofknowledge.com). The publication search was carried out from June 1993, when rs17878362 was first described[20] to December 2011. Several individual search terms, as well as combinations, were used: '*TP53*', 'p53', 'intron3', 'rs17878362', 'polymorphism', 'intron', 'PIN3' and '16bp-Del', as in several publications the major A1 allele is referred to as a deletion of the 16 bp sequence. The publications were reviewed to identify those that met the following inclusion criteria: (i) that the publication reported a formal case–control study analysing cancer susceptibility associated with rs17878362, (ii) results were given as an OR and (iii) the publication was in English.

Statistical analysis. The methodological approach described by Thakkinstian and collaborators was used to carry out our analyses on the association of the rs17878362 polymorphism with cancer risk variant allele with cancer risk and also those on rs1042522 and rs1625895 when data were available in the same panel of selected studies.[28] First, data from both controls and cases were extracted from the selected studies for the TP53 polymorphisms of interest, including the number of subjects, ORs[11] and the corresponding 95% CIs (Table 1 and Supplementary Table 1). Second, the Hardy–Weinberg equilibrium was tested by χ^2 goodness of fit in each study. Third, heterogeneity was determined using the *Q*-test and was considered as present when *Q*-test *P*-value was <0.05. According to the *Q*-test *P*-value, the association between a polymorphism and cancer risk was investigating using either the fixed- or the random-effects models, according to the method of DerSimonian and Laird.[39] Using the same methodology, sub-group analyses were performed by geographic location of the population and cancer type. Sensitivity analyses were conducted to assess the impact of any single study (leave-one-out analysis, cumulative inclusion over time analysis). Publication bias was tested using the Egger test. Statistical analyses were performed using the commercial STATA software (version 11.1, StataCorp LP, College Station, Texas, USA).

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgements. This work was supported by the Institut National Contre le Cancer (INCA), France (2009-192 to JH and PH). Research in Inserm U612 is also supported by funding from Institut Curie and Inserm. CS has a PhD fellowship from the French Ministry of Research and VM's stay in U612 was supported by funding from EU FP7 (Grant Number 249689 for the network of excellence DoReMi (low dose research towards multidisciplinary integration)).

1. Hainaut P, Hollstein M. p53 and human cancer: the first ten thousand mutations. *Adv Cancer Res* 2000; **77**: 81–137.
2. Petitjean A, Mathe E, Kato S, Ishioka C, Tavtigian SV, Hainaut P *et al.* Impact of mutant p53 functional properties on TP53 mutation patterns and tumor phenotype: lessons from recent developments in the IARC TP53 database. *Hum Mutat* 2007; **28**: 622–629.
3. Levine AJ. p53, the cellular gatekeeper for growth and division. *Cell* 1997; **88**: 323–331.
4. Levine AJ, Oren M. The first 30 years of p53: growing ever more complex. *Nat Rev Cancer* 2009; **9**: 749–758.
5. Meek DW, Anderson CW. Posttranslational modification of p53: cooperative integrators of function. *Cold Spring Harb Perspect Biol* 2009; **1**: a000950.
6. Vilborg A, Wilhelm MT, Wiman KG. Regulation of tumor suppressor p53 at the RNA level. *J Mol Med* 2010; **88**: 645–652.
7. Marcel V, Tran PL, Sagne C, Martel-Planche G, Vaslin L, Teulade-Fichou MP *et al.* G-quadruplex structures in TP53 intron 3: role in alternative splicing and in production of p53 mRNA isoforms. *Carcinogenesis* 2011; **32**: 271–278.
8. Marcel V, Dichtel-Danjoy ML, Sagne C, Hafsi H, Ma D, Ortiz-Cuaran S *et al.* Biological functions of p53 isoforms through evolution: lessons from animal and cellular models. *Cell Death Differ* 2011; **18**: 1815–1824.
9. Garritano S, Gemignani F, Palmero EI, Olivier M, Martel-Planche G, Le Calvez-Kelm F *et al.* 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* 2010; **31**: 143–150.
10. Whibley C, Pharoah PD, Hollstein M. p53 polymorphisms: cancer implications. *Nat Rev Cancer* 2009; **9**: 95–107.
11. Stacey SN, Sulem P, Jonasdottir A, Masson G, Gudmundsson J, Gudbjartsson DF *et al.* A germline variant in the TP53 polyadenylation signal confers cancer susceptibility. *Nat Genet* 2011; **43**: 1098–1103.
12. Matlashewski GJ, Tuck S, Pim D, Lamb P, Schneider J, Crawford LV. Primary structure polymorphism at amino acid residue 72 of human p53. *Mol Cell Biol* 1987; **7**: 961–963.
13. Sullivan A, Syed N, Gasco M, Bergamaschi D, Trigiante G, Attard M *et al.* Polymorphism in wild-type p53 modulates response to chemotherapy *in vitro* and *in vivo*. *Oncogene* 2004; **23**: 3328–3337.
14. Suspitsin EN, Buslov KG, Grigoriev MY, Ishutkina JG, Ulibina JM, Gorodinskaya VM *et al.* Evidence against involvement of p53 polymorphism in breast cancer predisposition. *Int J Cancer* 2003; **103**: 431–433.
15. Hiyama T, Yoshihara M, Tanaka S, Chayama K. Genetic polymorphisms and esophageal cancer risk. *Int J Cancer* 2007; **121**: 1643–1658.
16. Koushik A, Platt RW, Franco EL. p53 codon 72 polymorphism and cervical neoplasia: a meta-analysis review. *Cancer Epidemiol Biomar Prev* 2004; **13**: 11–22.
17. Tang NP, Wu YM, Wang B, Ma J. Systematic review and meta-analysis of the association between P53 codon 72 polymorphism and colorectal cancer. *Eur J Surg Oncol* 2010; **36**: 431–438.
18. Matakidou A, Eisen T, Houlston RS. TP53 polymorphisms and lung cancer risk: a systematic review and meta-analysis. *Mutagenesis* 2003; **18**: 377–385.
19. He XF, Su J, Zhang Y, Huang X, Liu Y, Ding DP *et al.* Association between the p53 polymorphisms and breast cancer risk: meta-analysis based on case-control study. *Breast Cancer Res Tr* 2011; **130**: 517–529.
20. Lazar V, Hazard F, Bertin F, Janin N, Bellet D, Bressac B. Simple sequence repeat polymorphism within the p53 gene. *Oncogene* 1993; **8**: 1703–1705.
21. Costa S, Pinto D, Pereira D, Rodrigues H, Cameselle-Teijeiro J, Medeiros R *et al.* Importance of TP53 codon 72 and intron 3 duplication 16bp polymorphisms in prediction of susceptibility on breast cancer. *BMC cancer* 2008; **8**: 32.
22. Wang-Gohrke S, Becher H, Kreienberg R, Runnebaum IB, Chang-Claude J. Intron 3 16 bp duplication polymorphism of p53 is associated with an increased risk for breast cancer by the age of 50 years. *Pharmacogenetics* 2002; **12**: 269–272.
23. Gemignani F, Moreno V, Landi S, Moullan N, Chabrier A, Gutierrez-Enriquez S *et al.* A TP53 polymorphism is associated with increased risk of colorectal cancer and with reduced levels of TP53 mRNA. *Oncogene* 2004; **23**: 1954–1956.
24. Perfumo C, Bonelli L, Menichini P, Inga A, Gismondi V, Ciferri E *et al.* Increased risk of colorectal adenomas in Italian subjects carrying the p53 PIN3 A2-Pro72 haplotype. *Digestion* 2006; **74**: 228–235.
25. Hu Z, Li X, Qu X, He Y, Ring BZ, Song E *et al.* Intron 3 16 bp duplication polymorphism of TP53 contributes to cancer susceptibility: a meta-analysis. *Carcinogenesis* 2010; **31**: 643–647.
26. Lu PH, Wei MX, Li C, Shen W, Chen MB. Need for clarification of data in a recent meta-analysis about TP53 codon 72 polymorphism and cancer susceptibility. *Carcinogenesis* 2011; **32**: 443 author reply 444.
27. Malik MA, Sharma K, Goel S, Zargar SA, Mittal B. Association of TP53 intron 3, 16 bp duplication polymorphism with esophageal and gastric cancer susceptibility in Kashmir Valley. *Oncol Res* 2011; **19**: 165–169.
28. Thakkinstian A, McElduff P, D'Este C, Duffy D, Attia J. A method for meta-analysis of molecular association studies. *Stat Med* 2005; **24**: 1291–1306.
29. Denisov EV, Sukhanovskaya TV, Dultseva TS, Malinovskaya EA, Litviakov NV, Slonimskaya EM *et al.* Coordination of TP53 abnormalities in breast cancer: data from analysis of TP53 polymorphisms, loss of heterozygosity, methylation, and mutations. *Genet Test Mol Biomarkers* 2011; **15**: 901–907.
30. Ashton KA, Proietto A, Otton G, Symonds I, McEvoy M, Attia J *et al.* Polymorphisms in TP53 and MDM2 combined are associated with high grade endometrial cancer. *Gynecol Oncol* 2009; **113**: 109–114.
31. Sjalander A, Birgander R, Kivela A, Beckman G. p53 polymorphisms and haplotypes in different ethnic groups. *Hum Hered* 1995; **45**: 144–149.
32. Beckman G, Birgander R, Sjalander A, Saha N, Holmberg PA, Kivela A *et al.* Is p53 polymorphism maintained by natural selection? *Hum Hered* 1994; **44**: 266–270.
33. Marcel V, Palmero EI, Falagan-Lotsch P, Martel-Planche G, Ashton-Prolla P, Olivier M *et al.* TP53 PIN3 and MDM2 SNP309 polymorphisms as genetic modifiers in the Li-Fraumeni syndrome: impact on age at first diagnosis. *J Med Genet* 2009; **46**: 766–772.
34. Dahabreh IJ, Linardou H, Bouzika P, Varvarigou V, Murray S. TP53 Arg72Pro polymorphism and colorectal cancer risk: a systematic review and meta-analysis. *Cancer Epidemiol, Biomarkers Prev* 2010; **19**: 1840–1847.
35. Courtois S, Verhaegh G, North S, Luciani MG, Lassus P, Hibner U *et al.* DeltaN-p53, a natural isoform of p53 lacking the first transactivation domain, counteracts growth suppression by wild-type p53. *Oncogene* 2002; **21**: 6722–6728.
36. Ghosh A, Stewart D, Matlashewski G. Regulation of human p53 activity and cell localization by alternative splicing. *Mol Cell Biol* 2004; **24**: 7987–7997.
37. Maier B, Gluba W, Bernier B, Turner T, Mohammad K, Guise T *et al.* Modulation of mammalian life span by the short isoform of p53. *Genes Dev* 2004; **18**: 306–319.
38. Hofstetter G, Berger A, Berger R, Zoric A, Braicu EI, Reimer D *et al.* The N-terminally truncated p53 isoform Delta40p53 influences prognosis in mucinous ovarian cancer. *Int J Gynecol Cancer* 2012; **22**: 372–379.
39. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Controlled Clin Trials* 1986; **7**: 177–188.

40. Jha P, Jha P, Pathak P, Chosdol K, Suri V, Sharma MC *et al.* TP53 polymorphisms in gliomas from Indian patients: Study of codon 72 genotype, rs1642785, rs1800370 and 16 base pair insertion in intron-3. *Exp Mol Pathol* 2011; **90**: 167–172.
41. Umar M, Upadhyay R, Khurana R, Kumar S, Ghoshal UC, Mittal B. Role of p53 and p73 genes polymorphisms in susceptibility to esophageal cancer: a case control study in a northern Indian population. *Mol Biol Rep* 2012; **39**: 1153–1162.
42. Alawadi S, Ghabreau L, Alsaleh M, Abdulaziz Z, Rafeek M, Akil N *et al.* P53 gene polymorphisms and breast cancer risk in Arab women. *Med Oncol* 2011; **28**: 709–715.
43. Mittal RD, George GP, Mishra J, Mittal T, Kapoor R. Role of functional polymorphisms of P53 and P73 genes with the risk of prostate cancer in a case-control study from Northern India. *Arch Med Res* 2011; **42**: 122–127.
44. Naccarati A, Pardini B, Polakova V, Smerhovsky Z, Vodickova L, Soucek P *et al.* Genotype and haplotype analysis of TP53 gene and the risk of pancreatic cancer: an association study in the Czech Republic. *Carcinogenesis* 2010; **31**: 666–670.
45. Polakova V, Pardini B, Naccarati A, Landi S, Slyskova J, Novotny J *et al.* Genotype and haplotype analysis of cell cycle genes in sporadic colorectal cancer in the Czech Republic. *Hum Mutat* 2009; **30**: 661–668.
46. De Feo E, Persiani R, La Greca A, Amore R, Arzani D, Rausei S *et al.* A case-control study on the effect of p53 and p73 gene polymorphisms on gastric cancer risk and progression. *Mutat Res* 2009; **675**: 60–65.
47. Hrstka R, Beranek M, Klocova K, Nenutil R, Vojtesek B. Intronic polymorphisms in TP53 indicate lymph node metastasis in breast cancer. *Oncol Rep* 2009; **22**: 1205–1211.
48. Gaudet MM, Gammon MD, Bensen JT, Sagiv SK, Shantakumar S, Teitelbaum SL *et al.* Genetic variation of TP53, polycyclic aromatic hydrocarbon-related exposures, and breast cancer risk among women on Long Island, New York. *Breast Cancer Res Treat* 2008; **108**: 93–99.
49. Ye Y, Yang H, Grossman HB, Dinney C, Wu X, Gu J. Genetic variants in cell cycle control pathway confer susceptibility to bladder cancer. *Cancer* 2008; **112**: 2467–2474.
50. De Vecchi G, Verderio P, Pizzamiglio S, Manoukian S, Bernard L, Pensotti V *et al.* The p53 Arg72Pro and Ins16bp polymorphisms and their haplotypes are not associated with breast cancer risk in BRCA-mutation negative familial cases. *Cancer Detect Prev* 2008; **32**: 140–143.
51. Chen K, Hu Z, Wang LE, Zhang W, El-Naggar AK, Sturgis EM *et al.* Polymorphic TP53BP1 and TP53 gene interactions associated with risk of squamous cell carcinoma of the head and neck. *Clin Cancer Res* 2007; **13**: 4300–4305.
52. Tan XL, Nieters A, Hoffmeister M, Beckmann L, Brenner H, Chang-Claude J. Genetic polymorphisms in TP53, nonsteroidal anti-inflammatory drugs and the risk of colorectal cancer: evidence for gene-environment interaction? *Pharmacogenet Genomics* 2007; **17**: 639–645.
53. Wang W, Spitz MR, Yang H, Lu C, Stewart DJ, Wu X. Genetic variants in cell cycle control pathway confer susceptibility to lung cancer. *Clin Cancer Res* 2007; **13**: 5974–5981.
54. Hung RJ, Boffetta P, Canzian F, Moullan N, Szeszenia-Dabrowska N, Zaridze D *et al.* Sequence variants in cell cycle control pathway, X-ray exposure, and lung cancer risk: a multicenter case-control study in Central Europe. *Cancer Res* 2006; **66**: 8280–8286.
55. Mitra S, Sikdar N, Misra C, Gupta S, Paul RR, Roy B *et al.* Risk assessment of p53 genotypes and haplotypes in tobacco-associated leukoplakia and oral cancer patients from eastern Idia. *Int J Cancer* 2005; **117**: 786–793.
56. Wu X, Zhao H, Amos CI, Shete S, Makan N, Hong WK *et al.* p53 Genotypes and Haplotypes Associated With Lung Cancer Susceptibility and Ethnicity. *J Natl Cancer Inst* 2002; **94**: 681–690.

SOME RIGHTS RESERVED

Cell Death and Disease **is an open-access journal published by** ***Nature Publishing Group*****. This work is licensed under the Creative Commons Attribution-NonCommercial-No Derivative Works 3.0 Unported License. To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-nd/3.0/**

Supplementary Information accompanies the paper on Cell Death and Disease website (http://www.nature.com/cddis)

Supplementary materials

A meta-analysis of cancer risk associated with the *TP53* intron 3 duplication polymorphism (rs17878362): geographic and tumor-specific effects

Charlotte Sagne, Virginie Marcel, Amina Amadou, Pierre Hainaut, Magali Olivier, Janet Hall

Supplementary Table 1. Additional characteristics of the selected case-control studies for rs17878362 (*TP53* PIN3) polymorphism.

Population	Study numbers and Study[×]		Country	Genotyping method	Mean age of cases (SD)	Mean age of controls (SD)	Adjusted method	Ref.
Indian	1	Jha *et al.*, 2011	India	Sequencing	nd[*]	nd[*]	Gender, age, BMI[+], smoking status	(1)
	2	Umar *et al.*, 2011	India	Gel retardation	57.3 (11.9)	55.8 (9.4)	Gender, age	(2)
	4	Mittal *et al.*, 2011	India	Gel retardation	64.4 (5.6)	64.4 (7.1)	Gender, age	(3)
	5	Malik *et al.,* 2011	India	Gel retardation	60.4 (8.4)	58.0 (12.7)	nd[*]	(4)
	6	Malik *et al.,* 2011	India	Gel retardation	55.9 (9.7)	58.0 (12.7)	nd[*]	(4)
	22	Mitra *et al.*, 2005	India	Gel retardation	55.0 (12.0)	50.4 (11.5)	nd[*]	(5)
Mediterranean	10	de Feo *et al.*, 2009	Italy	Gel retardation	66.7 (11.7)	63.5 (13.1)	Age, alcohol consumption, family history of cancer, fruit, vegetable intake, addition of salt to meals	(6)
	13	Costa *et al.*, 2008	Portugal	Gel retardation	53.4 (nd[*])	53.1 (nd[*])	Age	(7)
	15	de Vecchi *et al.*, 2008	Italy	Gel retardation	nd[*]	nd[*]	nd[*]	(8)
	20	Perfumo *et al.*, 2006	Italy	Gel retardation	60.1 (nd[*])	66.7 (nd[*])	nd[*]	(9)
	21	Perfumo *et al.*, 2006	Italy	Gel retardation	68.4 (nd[*])	66.7 (nd[*])	nd[*]	(9)
	23	Gemignani *et al.*, 2004	Spain	Gel retardation	nd[*]	nd[*]	nd[*]	(10)
Northern Europe	7	Naccarati *et al.*, 2010	Czech Republic	Gel retardation	62.2 (10.4)	60.5 (10.7)	Gender, Age, BMI[+], smoking status	(11)
	8	Polakova *et al.*, 2009	Czech Republic	Gel retardation	nd[*]	nd[*]	Age, sex	(12)
	11	Hrstka *et al.*, 2009	Czech Republic	Gel retardation	59.7 (nd[*])	58.86 (nd[*])	nd[*]	(13)
	17	Tan *et al.*, 2007	Germany	Gel retardation	68.2 (nd[*])	66.9 (nd[*])	nd[*]	(14)
	19	Hung *et al.*, 2006	Central Europe	Gel retardation	nd[*]	nd[*]	nd[*]	(15)
	24	Wang-Gohrke *et al.*, 2002	Germany	Gel retardation	nd[*]	nd[*]	nd[*]	(16)
United States	12	Gaudet *et al.*, 2008	USA	Gel retardation	nd[*]	nd[*]	Age	(17)
	14	Ye *et al.*, 2008	USA	Taqman	nd[*]	nd[*]	nd[*]	(18)
	16	Chen *et al.*, 2007	USA	Gel retardation	nd[*]	nd[*]	Age, sex, smoking status, alcohol consumption	(19)
	18	Wang *et al.*, 2007	USA	Gel retardation	61.7 (11.1)	61.4 (9.4)	nd[*]	(20)
	25	Wu *et al.*, 2002	USA	Gel retardation	61.4 (9.7)	60.6 (9.8)	Age, sex, smoking	(21)
Not classified	3	Alawadi *et al.*, 2011	Kuwait-Syria	Gel retardation	48.7 (nd[*])	54.7 (nd[*])	nd[*]	(22)
	9	Ashton *et al.*, 2009	Australia	Gel retardation	nd[*]	nd[*]	Age, BMI[+], HBP[§], diabetes, HRT[#], personal history of cancer, smoking status, alcohol consumption	(23)

[×]: Number corresponds to Table 1, [*]nd: not described, [+]BMI: Body Mass Index, [§]HBP: High Blood Pressure, [#] HRT: Hormone replacement therapy

Supplementary Table 2. Meta-analysis results after removing all studies one by one for A1A2 and A2A2 genotypes of *TP53* rs17878362. The name of the study corresponds to the removed study.

Study Numbers and Study*		A1A2 *vs* A1A1 genotypes			A2A2 *vs* A1A1 genotypes		
		Heterogenity *P* value	OR	95% CI	Heterogenity *P* value	OR	95% CI
1	Jha *et al.*, 2011	0.03	1.08	[0.99-1.19]	0.05	1.47	[1.16-1.86]
2	Umar *et al.*, 2011	0.02	1.08	[0.98-1.18]	0.05	1.45	[1.15-1.85]
3	Alawadi *et al.*, 2011	0.03	1.07	[0.98-1.17]	0.15	1.52	[1.27-1.82]
4	Mittal *et al.*, 2011	0.03	1.09	[0.99-1.19]	0.05	1.49	[1.18-1.88]
5	Malik *et al.,* 2011	0.02	1.07	[0.98-1.18]	0.06	1.41	[1.18-1.70]
6	Malik *et al.,* 2011	0.02	1.08	[0.98-1.18]	0.08	1.40	[1.17-1.68]
7	Naccarati *et al.*, 2010	0.02	1.08	[0.99-1.19]	0.04	1.47	[1.16-1.87]
8	Polakova *et al.*, 2009	0.02	1.08	[0.98-1.18]	0.05	1.46	[1.15-1.86]
9	Ashton *et al.*, 2009	0.03	1.09	[0.99-1.19]	0.05	1.49	[1.17-1.89]
10	de Feo *et al.*, 2009	0.02	1.08	[0.99-1.18]	0.04	1.47	[1.16-1.87]
11	Hrstka *et al.*, 2009	0.02	1.07	[0.98-1.18]	0.04	1.48	[1.17-1.88]
12	Gaudet *et al.*, 2008	0.02	1.09	[0.99-1.19]	0.05	1.50	[1.18-1.91]
13	Costa *et al.*, 2008	0.02	1.08	[0.98-1.18]	0.08	1.42	[1.19-1.70]
14	Ye *et al.*, 2008	**0.11**	**1.10**	**[1.01-1.19]**	0.10	1.51	[1.26-1.82]
15	de Vecchi *et al.*, 2008	0.02	1.07	[0.98-1.18]	0.06	1.42	[1.19-1.71]
16	Chen *et al.*, 2007	0.02	1.07	[0.98-1.18]	0.04	1.49	[1.16-1.89]
17	Tan *et al.*, 2007	**0.05**	**1.09**	**[1.00-1.20]**	0.06	1.48	[1.23-1.77]
18	Wang *et al.*, 2007	0.02	1.07	[0.97-1.18]	0.12	1.55	[1.28-1.86]
19	Hung *et al.*, 2006	0.02	1.08	[0.98-1.19]	0.07	1.37	[1.12-1.66]
20	Perfumo *et al.*, 2006	0.04	1.07	[0.98-1.17]	0.08	1.43	[1.20-1.71]
21	Perfumo *et al.*, 2006	0.02	1.08	[0.99-1.18]	0.05	1.44	[1.20-1.73]
22	Mitra *et al.*, 2005	**0.04**	**1.09**	**[1.00-1.19]**	0.14	1.53	[1.27-1.83]
23	Gemignani *et al.*, 2004	0.05	1.06	[0.99-1.13]	0.04	1.47	[1.16-1.87]
24	Wang-Gohrke *et al.*, 2002	0.03	1.07	[0.97-1.17]	0.05	1.47	[1.15-1.87]
25	Wu *et al.*, 2002	0.12	1.06	[0.99-1.13]	0.05	1.43	[1.20-1.72]

* For study number, number of cases and controls excluded and details of publications see Table I

Supplementary Table 3. The evolution of the meta-analysis results for rs17878362 depending on date of publication of study.

Genotype	Study	Heterogenity *P* value	OR	95% CI
A1A2	2005 (4 studies)	0.01	1.30	[0.94-1.78]
	2006 (7 studies)	0.01	1.23	[0.99-1.52]
	2007 (10 studies)	0.01	1.15	[1.00 1.33]
	2008 (14 studies)	0.01	1.09	[0.96-1.23]
	2009 (18 studies)	0.01	1.08	[0.97-1.20]
	2010 (19 studies)	0.01	1.08	[0.97-1.19]
	2011 (25 studies)	0.03	1.08	[0.99-1.18]
A2A2	2005 (4 studies)	0.13	1.27	[0.82-1.96]
	2006 (7 studies)	0.10	1.67	[1.23-2.27]
	2007 (10 studies)	0.06	1.37	[1.07 1.76]
	2008 (14 studies)	0.03	1.41	[1.03-1.94]
	2009 (18 studies)	0.10	1.40	[1.14-1.71]
	2010 (19 studies)	0.13	1.40	[1.15-1.71]
	2011 (25 studies)	0.06	1.45	[1.22-1.74]

Supplementary Table 4. Meta-analysis results for rs1042522 (*TP53* PEX4) and rs1625895 (*TP53* PIN6) polymorphisms in the selected case-control studies following Hardy-Weinberg equilibrium.

Genotypes	Cases n (%)	Controls n (%)	Heterogenity *P* value	OR	[95% CI]	*P* trend[#]
OVERALL						
rs1042522 (Study numbers: 1, 3, 4, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25; MAF = 0.29)						
Total	8 517 (100.0)	9 311 (100.0)				
R72/R72	4 227 (49.6)	4 833 (51.9)		1.00	-	
R72/P72	3 340 (39.3)	3 586 (38.5)	<0.01[*]	**1.16**	**[1.05-1.18]**	<0.01
P72/P72	950 (11.1)	892 (9.6)	0.02[+]	1.06	[0.91-1.24]	
rs1625895 (Study numbers: 4, 9, 10, 11, 12, 14, 16, 18, 24, 25; MAF = 0.19)						
Total	5 011 (100.0)	5 100 (100.0)				
GG	3 180 (63.5)	3 522 (69.1)		1.00	-	
GA	1 312 (26.2)	1 192 (23.4)	0.01[*]	**1.19**	**[1.02-1.40]**	<0.01
AA	519 (10.3)	386 (7.5)	0.46[*]	1.19	[0.89-1.60]	

Supplementary Table 5: Haplotype distributions for rs17878362*, rs1042522 and rs1625895 in 261 DNA samples from three HapMap populations (African: Nigeria (90 individuals); Asian: China (81 individuals) and Caucasian: Northern or Western European (90 individuals)) based on published data (24).

Polymorphisms			African	Asian	Caucasian
rs17878362*	rs1042522	rs1625895	(%)	(%)	(%)
A1	R72	G	31.67	53.70	78.13
A1	P72	G	38.33	43.83	11.46
A1	R72	A	-	-	0.52
A1	P72	A	3.89	0.62	0.52
A2	R72	G	-	-	-
A2	P72	G	-	-	-
A2	R72	A	-	-	-
A2	P72	A	26.11	1.85	9.37%

***tagged by rs2909430**

REFERENCES

1. Jha P, Jha P, Pathak P, Chosdol K, Suri V, Sharma MC*, et al.* TP53 polymorphisms in gliomas from Indian patients: Study of codon 72 genotype, rs1642785, rs1800370 and 16 base pair insertion in intron-3. *Experimental and molecular pathology* 2011 Apr; **90**(2)**:** 167-172.

2. Umar M, Upadhyay R, Khurana R, Kumar S, Ghoshal UC, Mittal B. Role of p53 and p73 genes polymorphisms in susceptibility to esophageal cancer: a case control study in a northern Indian population. *Molecular biology reports* 2012 Feb; **39**(2)**:** 1153-1162.

3. Mittal RD, George GP, Mishra J, Mittal T, Kapoor R. Role of functional polymorphisms of P53 and P73 genes with the risk of prostate cancer in a case-control study from Northern India. *Archives of medical research* 2011 Feb; **42**(2)**:** 122-127.

4. Malik MA, Sharma K, Goel S, Zargar SA, Mittal B. Association of TP53 intron 3, 16 bp duplication polymorphism with esophageal and gastric cancer susceptibility in Kashmir Valley. *Oncology research* 2011; **19**(3-4)**:** 165-169.

5. Mitra S, Sikdar N, Misra C, Gupta S, Paul RR, Roy B*, et al.* Risk assessment of p53 genotypes and haplotypes in tobacco-associated leukoplakia and oral cancer patients from eastern Idia. *International journal of cancer Journal international du cancer* 2005 Dec 10; **117**(5)**:** 786-793.

6. De Feo E, Persiani R, La Greca A, Amore R, Arzani D, Rausei S*, et al.* A case-control study on the effect of p53 and p73 gene polymorphisms on gastric cancer risk and progression. *Mutation research* 2009 Apr 30; **675**(1-2)**:** 60-65.

7. Costa S, Pinto D, Pereira D, Rodrigues H, Cameselle-Teijeiro J, Medeiros R*, et al.* Importance of TP53 codon 72 and intron 3 duplication 16bp polymorphisms in prediction of susceptibility on breast cancer. *BMC cancer* 2008; **8:** 32.

8. De Vecchi G, Verderio P, Pizzamiglio S, Manoukian S, Bernard L, Pensotti V*, et al.* The p53 Arg72Pro and Ins16bp polymorphisms and their haplotypes are not associated with breast cancer risk in BRCA-mutation negative familial cases. *Cancer detection and prevention* 2008; **32**(2)**:** 140-143.

9. Perfumo C, Bonelli L, Menichini P, Inga A, Gismondi V, Ciferri E*, et al.* Increased risk of colorectal adenomas in Italian subjects carrying the p53 PIN3 A2-Pro72 haplotype. *Digestion* 2006; **74**(3-4)**:** 228-235.

10. Gemignani F, Moreno V, Landi S, Moullan N, Chabrier A, Gutierrez-Enriquez S*, et al.* A TP53 polymorphism is associated with increased risk of colorectal cancer and with reduced levels of TP53 mRNA. *Oncogene* 2004 Mar 11; **23**(10)**:** 1954-1956.

11. Naccarati A, Pardini B, Polakova V, Smerhovsky Z, Vodickova L, Soucek P*, et al.* Genotype and haplotype analysis of TP53 gene and the risk of pancreatic cancer: an association study in the Czech Republic. *Carcinogenesis* 2010 Apr; **31**(4)**:** 666-670.

12. Polakova V, Pardini B, Naccarati A, Landi S, Slyskova J, Novotny J*, et al.* Genotype and haplotype analysis of cell cycle genes in sporadic colorectal cancer in the Czech Republic. *Human mutation* 2009 Apr; **30**(4)**:** 661-668.

13. Hrstka R, Beranek M, Klocova K, Nenutil R, Vojtesek B. Intronic polymorphisms in TP53 indicate lymph node metastasis in breast cancer. *Oncology reports* 2009 Nov; **22**(5)**:** 1205-1211.

14. Tan XL, Nieters A, Hoffmeister M, Beckmann L, Brenner H, Chang-Claude J. Genetic polymorphisms in TP53, nonsteroidal anti-inflammatory drugs and the risk of colorectal cancer: evidence for gene-environment interaction? *Pharmacogenetics and genomics* 2007 Aug; **17**(8)**:** 639-645.

15. Hung RJ, Boffetta P, Canzian F, Moullan N, Szeszenia-Dabrowska N, Zaridze D*, et al.* Sequence variants in cell cycle control pathway, X-ray exposure, and lung cancer risk: a multicenter case-control study in Central Europe. *Cancer research* 2006 Aug 15; **66**(16)**:** 8280-8286.

16. Wang-Gohrke S, Becher H, Kreienberg R, Runnebaum IB, Chang-Claude J. Intron 3 16 bp duplication polymorphism of p53 is associated with an increased risk for breast cancer by the age of 50 years. *Pharmacogenetics* 2002 Apr; **12**(3)**:** 269-272.

17. Gaudet MM, Gammon MD, Bensen JT, Sagiv SK, Shantakumar S, Teitelbaum SL*, et al.* Genetic variation of TP53, polycyclic aromatic hydrocarbon-related exposures, and breast cancer risk among women on Long Island, New York. *Breast cancer research and treatment* 2008 Mar; **108**(1)**:** 93-99.

18. Ye Y, Yang H, Grossman HB, Dinney C, Wu X, Gu J. Genetic variants in cell cycle control pathway confer susceptibility to bladder cancer. *Cancer* 2008 Jun; **112**(11)**:** 2467-2474.

19. Chen K, Hu Z, Wang LE, Zhang W, El-Naggar AK, Sturgis EM*, et al.* Polymorphic TP53BP1 and TP53 gene interactions associated with risk of squamous cell carcinoma of the head and neck. *Clinical cancer research : an official journal of the American Association for Cancer Research* 2007 Jul 15; **13**(14)**:** 4300-4305.

20. Wang W, Spitz MR, Yang H, Lu C, Stewart DJ, Wu X. Genetic variants in cell cycle control pathway confer susceptibility to lung cancer. *Clinical cancer research : an official journal of the American Association for Cancer Research* 2007 Oct 1; **13**(19)**:** 5974-5981.

21. Wu X, Zhao H, Amos CI, Shete S, Makan N, Hong WK*, et al.* p53 Genotypes and Haplotypes Associated With Lung Cancer Susceptibility and Ethnicity. *Journal of the National Cancer Institute* 2002 May 1; **94**(9)**:** 681-690.

22. Alawadi S, Ghabreau L, Alsaleh M, Abdulaziz Z, Rafeek M, Akil N*, et al.* P53 gene polymorphisms and breast cancer risk in Arab women. *Med Oncol* 2011 Sep; **28**(3)**:** 709-715.

23. Ashton KA, Proietto A, Otton G, Symonds I, McEvoy M, Attia J*, et al.* Polymorphisms in TP53 and MDM2 combined are associated with high grade endometrial cancer. *Gynecologic oncology* 2009 Apr; **113**(1)**:** 109-114.

24. Garritano S, Gemignani F, Palmero EI, Olivier M, Martel-Planche G, Le Calvez-Kelm F*, et al.* Detailed haplotype analysis at the TP53 locus in p.R337H mutation carriers in the population of Southern Brazil: evidence for a founder effect. *Human mutation* 2010 Feb; **31**(2)**:** 143-150.

Part II. Effects of age at cancer onset in germline *TP53* mutation carriers

This section summarizes the data presented in the annexed manuscript "Age at cancer onset in germline TP53 mutation carriers: association with polymorphism in predicted G-quadruple structures" in revision to Carcinogenesis. Figure and Table numbers correspond to those from the publication located after this abstract.

Background

The LFS syndrome is associated with a large and heterogeneous spectrum of early age of onset cancers. To date, the only genetic criteria associated with this syndrome is *TP53* germline mutations but the impact on the penetrance of the *TP53* germline mutations are still unclear. This result suggests that several other inheritable traits may predispose is this complex pattern of diseases. Two polymorphisms, one in *TP53* and the second in *MDM2*, a regulator of p53, have been shown to be associated with the severity of LFS/LFL in *TP53* mutation carriers. These SNPs are *TP53* rs1042522 (G>C, A>P at codon 72) and *MDM2* rs2279744 (T>G, SNP309) (Bond et al 2004, Matlashewski et al 1987). However, results remain controversial because of the limited size of the studies and of the multiple possible biases, for example, caused by comparisons among families (Bond et al 2004, Bougeard et al 2006, Marcel et al 2009, Renaux-Petel et al 2013). In a previous study, Marcel and collaborators have used a cohort of LFS/LFL subjects from Brazil to evaluate the effect of rs17878362 on age at cancer onset (Marcel et al 2009). They showed that the A1A2 rs17878362 genotype was associated with a statistically significant increase in the age of first cancer diagnosis in this Brazilian cohort (age difference between A1A1 and A1A2 carriers: 19 years, P value = 0.01).

The Brazilian cohort analysed in this study has added value for the discovery of possible modifier effects. First, in *TP53* R337H carriers, cancer patterns and age at onset are even more diverse than in carriers of other mutations, providing a larger range of phenotypes to compare. Second, assuming that the mutation has only mild effects, it is reasonable to hypothesize that modifiers may have stronger effects than in subjects with highly penetrant

mutations. Third, since *TP53* R337H carriers all share the same mutant allele, it is possible to circumscribe the effects of possible *TP53* modifier polymorphisms to the remaining WT allele.

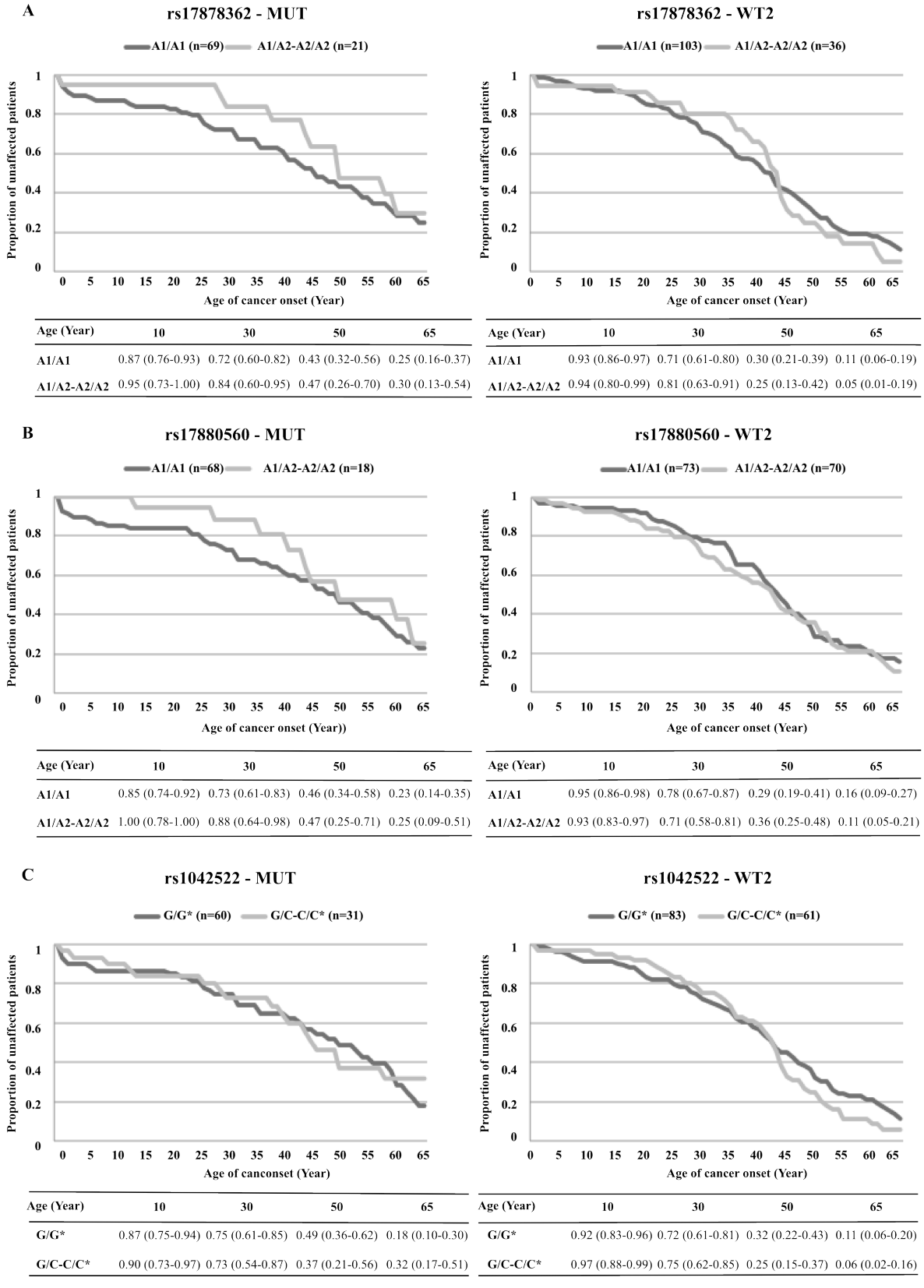
In this study, we have reasoned that polymorphisms in regions containing G4 motifs would have a role as modifiers. We have therefore identified putative or predicted G4 domains in *TP53*, mapped polymorphisms within or around these domains, and used the context of the Brazilian LFS/LFL cohorts to analyse the associations between each of these selected polymorphisms and the age at first cancer onset.

Results

First, we have evaluated the presence of G4 structures in the *TP53* gene using *in silico* programs. This analysis has shown that 5 putative G4 regions were located in the *TP53* gene: in the beginning of intron 1, in the end of intron 1, in intron 3, in intron 6 and in the 3'flanking region (**Figure 1**). Using the IARC *TP53* database and the NCBI dbSNP database, we have identified 85 polymorphisms located within or near these regions.

Secondly, we have assessed the frequency of these polymorphisms in a cohort of 402 Brazilian LFS/LFL patients with or without a germline *TP53* mutation and in a series of 487 Brazilian population controls. We detected only 11 of the 85 polymorphisms in our cohort and only 7 of these 11 polymorphisms had a minor allele frequency (MAF) higher than 0.04 (rs12944939 (G>A, intron 1), rs1642785 (G>C, intron 2), rs17878362, rs10425222, rs1625895 (G>A, intron 6), rs17880560 (6 bp duplication in 3'flanking region, A1: one copy; A2: two copies), rs1614984 (C>T, 3'flanking region) (**Supplementary Table 3**).

Thirdly, focusing on these 7 polymorphisms, we have analysed the impact on age of first cancer diagnosis in LFS/LFL subjects. To perform statistical analyses, we have used a specific model adjusted on the size of the family to calculate the *P value* and thus correct for possible biases due to familial factors. The 402 LFS/LFL subjects represent a total of 145 families, with heterogeneous distribution among families. In these 145 families, 35 had subjects with *TP53* germline mutation. Two groups were considered. The "MUT" group included subjects who carry a germline *TP53* mutation, whereas the WT2 group included subjects from families with no germline *TP53* mutation. None of the 7 SNPs showed any statistically significant association with age at first cancer diagnosis in the WT2 group (**Table 1, Supplementary Table 6 and Figure 2**). In the MUT group, only the rare alleles of rs17878362, located in the described G4 in intron 3, and of rs17880560, located close to the described G4 in the 3'flanking region, were associated with a delay in the age at first cancer, although the effect was of borderline significance (**Figure 19A and B of this Thesis**).



Age (Year)	10	30	50	65
A1/A1	0.87 (0.76-0.93)	0.72 (0.60-0.82)	0.43 (0.32-0.56)	0.25 (0.16-0.37)
A1/A2-A2/A2	0.95 (0.73-1.00)	0.84 (0.60-0.95)	0.47 (0.26-0.70)	0.30 (0.13-0.54)

Age (Year)	10	30	50	65
A1/A1	0.93 (0.86-0.97)	0.71 (0.61-0.80)	0.30 (0.21-0.39)	0.11 (0.06-0.19)
A1/A2-A2/A2	0.94 (0.80-0.99)	0.81 (0.63-0.91)	0.25 (0.13-0.42)	0.05 (0.01-0.19)

Age (Year)	10	30	50	65
A1/A1	0.85 (0.74-0.92)	0.73 (0.61-0.83)	0.46 (0.34-0.58)	0.23 (0.14-0.35)
A1/A2-A2/A2	1.00 (0.78-1.00)	0.88 (0.64-0.98)	0.47 (0.25-0.71)	0.25 (0.09-0.51)

Age (Year)	10	30	50	65
A1/A1	0.95 (0.86-0.98)	0.78 (0.67-0.87)	0.29 (0.19-0.41)	0.16 (0.09-0.27)
A1/A2-A2/A2	0.93 (0.83-0.97)	0.71 (0.58-0.81)	0.36 (0.25-0.48)	0.11 (0.05-0.21)

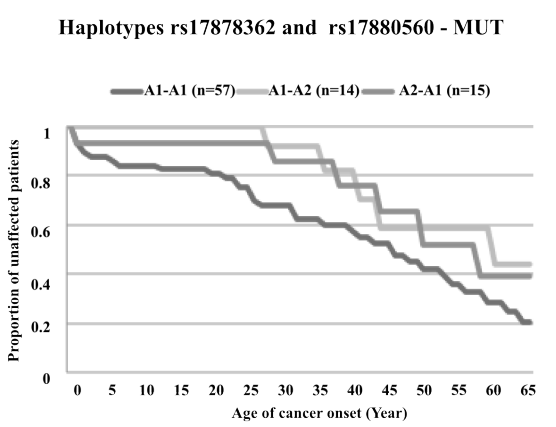
Age (Year)	10	30	50	65
G/G*	0.87 (0.75-0.94)	0.75 (0.61-0.85)	0.49 (0.36-0.62)	0.18 (0.10-0.30)
G/C-C/C*	0.90 (0.73-0.97)	0.73 (0.54-0.87)	0.37 (0.21-0.56)	0.32 (0.17-0.51)

Age (Year)	10	30	50	65
G/G*	0.92 (0.83-0.96)	0.72 (0.61-0.81)	0.32 (0.22-0.43)	0.11 (0.06-0.20)
G/C-C/C*	0.97 (0.88-0.99)	0.75 (0.62-0.85)	0.25 (0.15-0.37)	0.06 (0.02-0.16)

Figure 19: Kaplan-Meir disease-free probability estimates in LFS/LFL family members with or without *TP53* mutations. Kaplan-Meir probabilities are shown for (**A**) rs17878362, (**B**) rs17880560 and (**C**) on the haplotype of the WT allele. In each panel, the left panel corresponds to subjects of the MUT (*TP53* mutation carriers in LFS/LFL family members) and the right panel to subjects of the WT2 group (families with no mutation detected). The tables under the graphs show disease-free probability estimates at different ages (10, 30, 50 and 65 years) according to genotype. Only probabilities up to 65 years are shown.

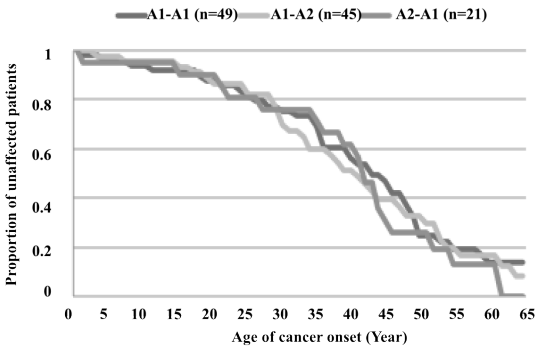
To analyse in detail the impact of the rs17878362 and rs17880560 polymorphisms on age at first diagnosis, we performed haplotype studies associating these 2 polymorphisms (**Figure 19C of this Thesis**). We have observed that the WT2 patients did not show any difference in age of cancer onset (**Table 2 and Figure 3**). In contrast, in the MUT, group, A1A1 (A1 for rs17878362 and A1 for rs17880560) carriers developed cancer on average 25 years before of at least one A2 allele (A1 for rs17878362 and A2 for rs17880560) and (A2-A1: A2 for rs17878362 and A1 for rs17880560). In addition, before the age of 35, only patients carrying the A1-A1 haplotype developed cancer, suggesting that this haplotype identifies a group of subjects who are at high risk for early/childhood cancer (**Figure 20**).

A



Age (Year)	10	30	50	65
A1-A1	0.84 (0.72-0.92)	0.68 (0.54-0.79)	0.42 (0.30-0.56)	0.20 (0.11-0.34)
A1-A2	1.00 (0.73-1.00)	0.92 (0.64-1.00)	0.59 (0.31-0.82)	0.44 (0.20-0.71)
A2-A1	0.93 (0.66-1.00)	0.86 (0.57-0.97)	0.52 (0.26-0.77)	0.39 (0.17-0.66)

B



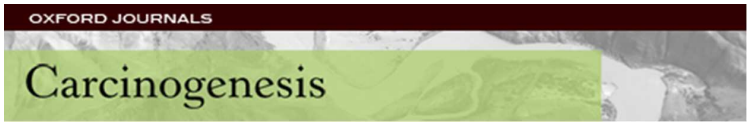
Age (Year)	10	30	50	65
A1-A1	0.94 (0.82-0.98)	0.75 (0.61-0.86)	0.25 (0.14-0.49)	0.14 (0.06-0.27)
A1-A2	0.96 (0.84-0.99)	0.70 (0.54-0.82)	0.33 (0.20-0.49)	0.08 (0.03-0.21)
A2-A1	0.95 (0.74-1.00)	0.76 (0.52-0.91)	0.26 (0.10-0.50)	0.00 (0.00-0.19)

Figure 20: Effect of different wild type haplotypes in *TP53* mutation carriers: a model. *TP53* alleles are represented as rods. (**A**) Mutant allele occurring on a haplotype carrying A1 variants of both rs17878362 and rs17880560 (A1A1). (**B**) Different types of wild type haplotypes. The wild type haplotype defined by A1A1 is considered as a "weak" haplotype (associated with early cancer, indicative low capacity to compensate the loss of p53 function of the mutant allele). The wild type haplotypes defined by A1A2 or A2A1 are considered as "strong haplotype" (associated with later cancer onset, thus providing at least partial compensation for the loss of function of the mutant allele). Of note, our data do not predict the effect of wild type A2A2 haplotypes, or the effects of these haplotypes when the mutation occurs on another haplotype than A1A1.

Conclusion

The A2 alleles of the rs17878362 and the rs17880560 polymorphisms located near or within G4 structures, appear to be protective in patients against an early age at cancer onset. Their combination within defined haplotypes is more effective for protecting against the development of cancer at a young age than any of these polymorphisms alone. These results suggest that genetic heterogeneity affecting the folding and position of G4 may impact on *TP53* suppressor functions.

Carcinogenesis



Age at cancer onset in germline TP53 mutation carriers: association with polymorphisms in predicted G-quadruplex structures

Journal:	*Carcinogenesis*
Manuscript ID:	Draft
Manuscript Type:	Original Manuscript
Date Submitted by the Author:	n/a
Complete List of Authors:	Sagne, Charlotte; Institut Curie, Genotoxicology, Signalisation and Experimental Radiotherapy; Inserm, U612; International Agency for Research in Cancer, Molecular Mechanisms and Biomarkers Group Marcel, Virginie; Institut Curie, Genotoxicology, Signalisation and Experimental Radiotherapy; Inserm, U612 Bota, Maria; International Prevention Research Institute, Research Martel-Planche, Ghyslaine; International Agency for Research in Cancer, Molecular Mechanisms and Biomarkers Group Nogueira, Amanda; A.C. Camargo Cancer Center, Oncogenetics Palmero, Edenir; Barretos Cancer Hospital, Molecular Oncology Research Center Perriaud, Laury; Institut Curie, Genotoxicology, Signalisation and Experimental Radiotherapy; Inserm, U612 Boniol, Mathieu; International Prevention Research Institute, Research Vagner, Stephan; Institut Curie, Genotoxic Stress and Cancer; CNRS, UMR3348 Cox, David; Cancer Research Center of Lyon, INSERM U1052-CNRS UMR5286 Chan, Chang; Cancer Institute of New Jersey, Medicine and Center for Systems Biology Mergny, Jean-Louis; Institut européen de chimie biologie, INSERM U869 Olivier, Magali; International Agency for Research in Cancer, Molecular Mechanisms and Biomarkers Group Prolla, Patricia; Hospital de clinicas de Porto Alegre, Genetics and Molecular Biology Hall, Janet; Institut Curie, Genotoxicology, Signalisation and Experimental Radiotherapy; Inserm, U612 Hainaut, Pierre; International Prevention Research Institute, Research; Achatz, Maria; A.C. Camargo Cancer Center, Oncogenetics
Keywords:	rs17878362, rs1788056, Haplotypes, Li-Fraumeni syndrome

Age at cancer onset in germline *TP53* mutation carriers: association with polymorphisms in predicted G-quadruplex structures

Charlotte Sagne[1,2,3], Virginie Marcel[1,2,§], Maria Bota[4], Ghyslaine Martel-Planche[3], Amanda Nobrega[5], Edenir Inêz Palmero[6], Laury Perriaud[1,2], Mathieu Boniol[4], Stephan Vagner[2,7], David G. Cox[8], Chang S. Chan[9], Jean-Louis Mergny[10,11], Magali Olivier[3], Patricia Prolla[12], Janet Hall[1,2], Pierre Hainaut[4*], Maria Isabel Achatz[5]

Addresses:

[1] INSERM U612, Bât 110-112, Centre Universitaire, Orsay 91405, France

[2] Institut Curie, Centre de Recherche, Bât 110-112, Centre Universitaire, Orsay 91405, France

[3] Molecular Carcinogenesis Group, International Agency for Research on Cancer, 150 cours Albert Thomas, 69372 Lyon Cedex 08, France

[4] International Prevention Research Institute, 15 chemin du Saquin, 69130 Ecully, France

[5] Department of Oncogenetics, A.C. Camargo Cancer Center and National Institute of Science and Technology in Oncogenomics, São Paulo, Brazil

[6] Molecular Oncology Research Center, Barretos Cancer Hospital, Av. Antenor Duarte Vilela, 1331, Barretos, São Paulo, Brazil

[7] CNRS UMR3348 Bât 110-112, Centre Universitaire, Orsay 91405, France

[8] Cancer Research Center of Lyon, INSERM U1052 – CNRS UMR5286, Centre Leon Bérard, 28 rue Laennec, 69373 Lyon, France

For Peer Review

[9] Department of Medicine and Center for Systems Biology, Cancer Institute of New Jersey, University of Dentistry and Medicine of New Jersey, Robert Wood Johnson Medical School, 195 Little Albany St, New Brunswick, NJ, 08901, USA

[10] INSERM U869, Institut Européen de Chimie Biologie, 2 rue Robert Escarpit, 33607 Pessac, France

[11] University of Bordeaux, 33607, Pessac, France

[12] Department of Genetics and Molecular Biology, Federal University of Rio Grande do Sul (UFRGS) and Medical Genetics Service, Hospital de Clínicas de Porto Alegre, and National Institute of Science and Technology in Populational Medical Genetics (INAGEMP), Porto Alegre, Brazil

[§] Present address: Cancer Research Center of Lyon, UMR INSERM 1052 CNRS 5286, University Lyon 1, Centre Léon Bérard, 28 rue Laënnec, 69373 Lyon Cedex 08, France

*Corresponding author:

Pierre HAINAUT, PhD

International Prevention Research Institute

15 chemin du Saquin,

69130 Ecully, France

Phone: +33 4 72 17 11 87

Email: pierre.hainaut@i-pri.org

For Peer Review

Running title: rs17878362 and rs1788056 and cancer risk

Keywords: rs17878362, rs1788056, Haplotypes, Li-Fraumeni syndrome

Abbreviations: MAF: Minor Allele Frequency; OR: Odds Ratio; CI: Confidence Interval; PIN2: Polymorphism in Intron 2 (rs1642785); PIN3: Polymorphism in Intron 3 (rs17878362); PEX4: Polymorphism in Exon 4 (rs1042522); PIN6: Polymorphism in Intron 6 (rs1625895); bp: base pair; Kb: Kilo base; SNP: single nucleotide polymorphism.

For Peer Review

ABSTRACT

Germline *TP53* mutations predispose to multiple cancers defining Fraumeni/Li-Fraumeni-Like Syndrome (LFS/LFL), a disease with large individual disparities. G-quadruplexes (G4) are secondary structural motifs occurring in guanine tracks, with regulatory effects on both DNA and RNA. We analyzed 85 polymorphisms within or near five predicted G4s in *TP53* in search of modifiers of penetrance of LFS/LFL in Brazilian cancer families with (n=34) or without (n=110) *TP53* mutations. Statistical analyses stratified on family structure showed that cancer tended to occur 12.5 to 18.0 years later in mutation carriers who also carried the variant alleles of two polymorphisms within predicted G4-forming regions, rs17878362 (*TP53* PIN3, 16bp duplication in intron 3; p=0.082) and rs17880560 (6bp duplication in 3'flanking region; p =0.067). Haplotype analysis showed that this inverse association was driven by the polymorphic status of the remaining wild type haplotype in mutation carriers. In carriers with wild type haplotype containing at least one variant allele of rs17878362 or rs17880560, cancer occurred about 15 years later than in carriers with other wild type haplotypes (p=0.019). No effect on age of cancer onset was observed in subjects without *TP53* mutation. The G4 in intron 3 has been shown to regulate alternative p53 mRNA splicing, while the biological role of predicted G4s in the 3'flanking region remain to be elucidated. G4-polymorphisms in haplotypes of the wild type *TP53* allele have an impact on LFS/LFL penetrance in germline *TP53* mutation carriers.

For Peer Review

Introduction

Germline mutations in the *TP53* tumor suppressor gene (17p13.1, OMIM #191170) predispose to a range of early-onset cancers that define the Li-Fraumeni Syndrome (1,2). The LFS tumor pattern is dominated by childhood adrenal cortical carcinoma, choroid plexus carcinoma, medulloblastoma and rhabdomyosarcoma, followed by soft-tissue sarcoma, osteosarcoma, pre-menopausal breast cancer and brain tumors in adolescents and young adults. Several definitions of "Li-Fraumeni-like" (LFL) syndrome have been proposed in families showing only partial LFS traits (3-5). Patients from families with LFS/LFL traits often carry germline *TP53* mutations. Currently, no other recurrent germline alteration has been associated with this disease pattern. Depending upon clinical definition, *TP53* mutations are identified in 20 to 70% of the cases (6-9). In current practice, probands with suspected LFS/LFL are referred for *TP53* mutation testing based on a set of individual and familial criteria known as the "modified Chompret criteria" (10). Mutations are detected in about 21-29% of probands matching these criteria (9,11). In *TP53* mutation carriers, the penetrance of the disease is about 50% by age 30-35 and reaches 90% over lifetime. In addition, *TP53* mutation carriers show a tendency to develop a wide range of cancers at an earlier age than in the general population, although there are large individual and familial variations in the age at diagnosis. Thus, it is extremely difficult to develop a protocol for prediction and surveillance of the disease. The current practice is to enroll subjects who tested positive for *TP53* mutations in complex surveillance programs (12) (NCCN guidelines version 2.2013).

The variable individual cancer patterns suggest that other genetic or epigenetic traits may act as modifiers. Studies by Bond *et al.* (13) observed that a single nucleotide polymorphism (SNP) in *MDM2*, encoding a protein that regulates p53 protein stability, was associated with a reduction of 9 years in the age at diagnosis of first cancer. This effect was detected in several independent cohorts of LFS/LFL families (14-17). This SNP (rs2279744, SNP309) has been

For Peer Review

shown to alter the sequence-specific DNA binding of the transcription factor Sp1 to the *MDM2* promoter, thus accounting for differences in levels of Mdm2 protein expression (13). There is evidence that the effect of rs2279744 might be modulated by another SNP in the *MDM2* promoter (rs117039649, SNP285) (18). Borderline effects on age at diagnosis in *TP53* mutation carriers have also been reported for a common non-synonymous SNP in exon 4 of *TP53*, R72P (rs1042522, *TP53* PEX4) (15). In a previous analysis on a cohort of *TP53* mutation carriers from Brazil, we have shown that a 16 bp duplication polymorphism in intron 3 (rs17878362, *TP53* PIN3) was associated with a large difference in the age of cancer diagnosis, with carriers of two alleles without the duplication (A1 allele) developing their first cancer on average 20 years earlier than heterozygote carriers (A1/A2) (15). However, this Brazilian cohort included a large proportion of subjects from apparently unrelated families who carried the same germline *TP53* mutation, R337H, which is common in South-Eastern Brazil due to a widespread founder effect detected in 0.3% of the general population (19). So far, the effect of rs17878362 has not been replicated in an independent LFS/LFL cohort. Thus, it is possible that the effect of the A1 allele of rs17878362 may be due, at least in part, to specific effects on the penetrance of R337H or to other genetic factors in this particular Brazilian background.

G-quadruplexes (G4s) consist of four-stranded structures occurring in guanine-rich sequences. Potential G4-forming structures are widespread in the genome and have been shown to have important regulatory effects on gene transcription (20), genomic stability (21-23) and DNA replication (24), mRNA splicing (25) and mRNA stability (25). Experimentally, drugs that modulate the stability of G4 exert significant effects on the expression (26) and splicing of specific sets of genes (27). Two G4 have been experimentally demonstrated in p53 pre-mRNA. A complex G4 structure is located in intron 3 in which the rs17878362 16bp duplication is located, and regulates the splicing of p53 pre-mRNA into either a fully spliced

For Peer Review

form generating the canonical p53 tumor suppressor protein, or p53I2, a variant mRNA that retains intron 2 (28). This variant mRNA encodes a p53 isoform protein that lacks the 39 N-terminal residues (29) and exerts regulatory effects on p53 suppressor activity (28,30-32). Another G4 has been identified downstream of the cleavage/polyadenylation site of the p53 pre-mRNA. This G4 contributes to p53 pre-mRNA 3'end processing in response to DNA damage (33).

In this study, we have hypothesized that polymorphisms affecting G4 structures may exert modifier effects on age at cancer onset in LFS/LFL subjects. We have identified candidate G4 domains throughout the entire *TP53* locus and analyzed the association between SNPs located within or in the close vicinity to these putative and demonstrated G4s and age at first cancer diagnosis in a series of 402 subjects from Brazilian families with LFS/LFL traits.

MATERIALS AND METHODS

LFS/LFL families

LFS/LFL Brazilian subjects were recruited from families attending the Cancer Risk Evaluation clinics in the Department of Oncogenetics of A.C. Camargo Cancer Center (São Paulo, Brazil) and Hospital de Clinicas, Universidade Federal do Rio Grande do Sul (Porto Alegre) (Ethics statement number 0568/07). A total of 402 subjects from 145 families were included in the study (**Supplementary Table 1**). Families matched at least one LFS/LFL criteria (classic LFS, Birch, Chompret, Eeles criteria), and their familial history was documented over several generations using the Progeny software (Progeny software Inc, Wolfville, Nova Scotia, Canada). Of these families, 15 (10.3%) matched classic LFS criteria while 66 (45.5%) matched Birch or Chompret criteria. Other families matched more relaxed criteria known as Eeles (8). All families were sequenced for *TP53* germline mutations

For Peer Review

throughout the coding sequence (exons 2 to 11 including splice junctions). Mutations were found in 35 families (24.1%), including 19 families with the "Brazilian founder" R337H mutation and 16 families with previously reported DNA-binding domain mutations (34) (http://p53.iarc.fr/TP53GermlineMutations.aspx). The institutional ethics committees of participating institutions approved the study and all patients provided informed consent.

Controls

Four groups of control subjects were selected. Group 1 consisted of 487 cancer-asymptomatic subjects from the general population of Southeast Brazil (age 40- 69 years) (34). Groups 2 to 4 consisted of individuals included as reference samples in the HapMap project. Group 2 (Caucasians) included 30 trios (two parents and a child) of Northern or Western European ancestry, Group 3 (Asians) 90 East Asian subjects, 45 of whom were from the Tokyo and Beijing areas, respectively and Group 4 (Africans) consisted of 30 trios (2 parents, 1 child) from the Yoruba people in Nigeria (35).

Prediction of G-Quadruplex motifs and selection of polymorphisms

The *TP53* sequence NC_000017.10 (ENSG00000141510) containing the entire *TP53* gene sequence plus 2000 bp upstream and downstream of the first and last exons, respectively, was used for *in silico* prediction of G-rich regions likely to form G4 structures using the prediction software Quadfinder (http://bioinformatics.ca/links_directory/tool/10280/quadfinder) and the web-based QGRS Mapper (http://bioinformatics.ramapo.edu/QGRS/index.php) with the following default parameters: G stretch = 3-5; loop size = 1-7; maximum length = 45 (36,37). A total of 5 G4 regions were independently identified by the two softwares. Polymorphisms located in these regions or within 100 bp upstream or downstream of these regions were selected in dbSNP (http://www.ncbi.nlm.nih.gov/snp/). To this set, were added the well-defined *TP53* SNPs rs1642785 (G/C, intron 2) and rs1042522 (G/C, exon 4). Overall, a total of 87 SNPs were selected (**Supplementary Table 2 and Figure 1**).

For Peer Review

Analysis of mutations and polymorphisms

DNA was extracted from peripheral white blood cells obtained by venipuncture with the Qiagen DNA Extraction kit according to manufacturer instructions (QIAamp DNA blood Maxi kit, Qiagen, Hilden, Germany). Purified DNA from HapMap subjects were obtained from the Coriell Institute Biorepository as described earlier (35). *TP53* mutations in coding sequence (exon 2 to 11 including flanking intronic regions containing splice sites) were sequenced according to the protocols of the International Agency for Research on Cancer (http://p53.iarc.fr/Download/TP53_DirectSequencing_IARC.pdf). Genotyping *TP53* polymorphisms was performed by direct sequencing using primers and polymerase chain reaction (PCR) conditions described in **Supplementary Table 2**. Genotyping for *MDM2* rs2279744 was performed using a 5'exonuclease SNP genotyping assay (Proligo, St Louis, Missouri, USA) (probes FAM 5'-cccgcgccgcagc-3' and Hex 5'-cccgcgccgaagc; primers 5'-ttcagggtaaaggtcacggg-3' and 5'-tcaacctgcccactgaacc-3') or by restriction fragment length polymorphism (RFLP) after PCR with primers and condition as described in **Supplementary Table 2**. RFLP was performed using 5 µL of PCR product, digested with 1 unit of MspAI I restriction enzyme (New England Biolabs, Ipswich, UK) during 4 h at 37°C. The results were analyzed on 3% agarose electrophoresis gel. The T allele of rs2279744 generated two RFLP fragments (53 and 158 bp), whereas the G allele generated three fragments (46, 53 and 112 bp). Comparison between 5'exonuclease assay and RFLP showed complete concordance between the two methods.

Haplotype reconstruction

Haplotypes of Brazilian subjects were reconstructed using pooled sequence variant data of individuals from LFS/LFL families and from the Brazilian control group (Group 1). Rare variants (MAF < 0.04) were excluded from the analysis. Haplotypes were constructed based on 5 SNP (rs1642785, rs17878362, rs1042522, rs17880560 and rs1614984) and on *TP53*

For Peer Review

germline mutation status, using R software haplo.stats package (http://mayoresearch.mayo.edu/mayo/research/schaid_lab/software.cfm)

Statistical analysis

Hardy Weinberg equilibrium was assessed for each study group using a web-based calculator (http://www.tufts.edu/~mcourt01/Documents/Court%20lab%20-%20HW%20calculator.xls). Statistical analyses were performed using R software (weights package; http://cran.r-project.org/web/packages/weights/weights.pdf) and running a hierarchical model that takes into account family size to compensate for family bias caused by multiple subjects belonging to the same large pedigree. Briefly, analyses were adjusted for the number of subjects in the family who were tested for *TP53* mutation and for polymorphisms, thus taking into account possible bias due to specific characteristics of large families in the dataset. Comparisons between average ages at first cancer diagnostic were performed by t-test weighted by family size, or by taking into account family sizes either for the whole cohort ("All") or for a particular sub-group analyzed (e.g. only families with defined wild type or mutant *TP53* status; "Group"). Unadjusted Kaplan-Meier disease-free probability estimates were calculated using tools available at http://vassarstats.net/. A "disease" event was assigned as the age at first cancer diagnosis (subsequent diagnoses in the same subjects were not taken into account). Subjects older than 65 years were censored. For each year of age, subjects with no diagnosis were censored at their date of last follow-up.

RESULTS

Distribution of G4s and polymorphisms in *TP53*.

The presence of at least three consecutive guanines along a DNA or RNA strand is (as a first approximation) the classical minimal requirement for intramolecular G4 formation. Longer

For Peer Review

DNA/RNA sequences containing multiples of four G-tracts can, in principle, accommodate higher-order structures defined by multiple G4 blocks, with wide topological and structural variations. The human genome has been predicted to contain up to 376,000 G4s (38). Using *in silico* prediction programs, we have identified 5 predicted G4s in a region of 29.79 Kb encompassing the entire *TP53* sequence (**Figure 1A**). The sequences and positions of these predicted G4s are given in **Figures 1B, 1C and Supplementary Figure 1**. They are located in the proximal and distal part of intron 1, in intron 3, intron 6 and the 3'flanking region, respectively. (**Figure 1B and C**). Of these predicted G4s, the one located in intron 3 (28) and another one (among several predicted domains) in the 3'flank (33), have been previously identified by structural and molecular techniques. A total of 85 polymorphisms reported in dbSNP were located within or near the 5 predicted G4s and their distributions were analyzed in LFS/LFL subjects, Brazilian controls (Group 1) and in the HapMap series (Groups 2-4). Only 11 of the polymorphisms showed allelic variation in these series. Four had a minor allele frequency (MAF) less than 4% including rs78378222 that has been associated with basal cell carcinoma (Prostate cancer, glioma and colorectal adenoma) in an Iceland population (39). The allele distributions of others polymorphisms (polymorphisms A, B, C, E, F and G) and of rs1042522 (*TP53* PEX4, codon R72P) in LFS/LFL families and in the 4 control groups are given in **Supplementary Table 3**. The distribution of these polymorphisms was in Hardy-Weinberg equilibrium (HWE) except for SNP D and E in Group 2 (HapMap Caucasian, $p=0.04$ for both SNPs), SNP G in Group 1 ($p<0.01$) and SNP A and D in LFS/LFL group ($p<0.01$ for both SNPs). When compared to Group 1, the deviation from HWE for SNP D (rs1042522) in LFS/LFL appeared to be due to low numbers of carriers of the C/C (72P/72P) allele combination, whereas numbers of G/G homozygotes and G/C heterozygotes were comparable in both groups. The distribution of polymorphisms was different among the three HapMap series ($p<0.01$, t-test). The Brazilian LFS/LFL and the controls of Group 1

For Peer Review

(Brazilian population) had a SNP profile similar to HapMap Caucasians (p=0.47, Mann-Whitney t test).

Association between polymorphisms in and/or near G4s and age at first diagnosis in LFS/LFL

The summary characteristics of subjects included in the LFS/LFL group are given in **Supplementary Table 1**. Among the 145 families, 35 had a germline *TP53* mutation (MUT group), 19 of which were carriers of the "Brazilian founder" R337H mutation. The main types of cancer diagnosed in carriers of R337H and in carriers of other mutations, and their mean age at diagnosis, are summarized in **Supplementary Table 4**. Overall, R337H carriers developed a range of cancer types which is typical of the LFS spectrum, however with a non-significant tendency for diagnosis at a later age than carriers of other *TP53* mutations. Within families with germline *TP53* mutations, subjects who were not found to carry the familial mutation were identified as the WT1 (Wild type 1) group. Among the 110 families without mutations, 157 cases of cancer were diagnosed among 122 subjects. Subjects from these families without a *TP53* mutation, either with or without cancer, were identified as the WT2 (Wild type 2) group. Types of cancer and ages at diagnosis in this group were compatible with LFS/LFL definitions (**Supplementary table 4**).

We next analyzed the age at diagnosis of a first cancer in relation to the carriage of the *TP53* SNPs A-G and of rs2279744 in the following subgroups: the MUT group (including subjects with any *TP53* germline mutation) and the WT2 group. Mean age at first diagnosis (± standard deviation) was calculated in different subgroups using an adjusted model weighted for family size to take into account a possible familial bias. P-values (t-test) were adjusted on family size either for the whole LFS/LFL cohort ("All") or for the sub-group of patients considered ("Group"). Effects on age at diagnosis were observed only with polymorphism C (rs17878362, 16bp insertion/duplication in intron 3) and polymorphism F (rs17880560, 6bp

For Peer Review

insertion/duplication in 3' flanking region), (**Table 1**; data for other SNPs in **Supplementary Table 5**). For each of the two polymorphisms, alleles were identified as A1 (non-duplicated) and A2 (duplicated), defining three genotypes, A1/A1, A1/A2 and A2/A2 at each polymorphism. No significant difference was found for the age at first cancer diagnosis within the WT2 group in relation with the status of these polymorphisms (for instance the age at first diagnosis ranged between 32.67±18.60 years in subjects with rs17880560 A2/A2 allele status and 39.56±16.43 in subjects with rs17880560 A1/A1 allele status). In contrast, in the MUT group, there was a tendency for a difference of over 12.5 years between subjects with either A1/A1 or A1/A2 genotypes, the A1/A2 mutation carriers having their first diagnosis later than the A1/A1 mutation carriers. The difference in age at diagnosis in relation with genotype was significant in a model adjusted for family size in the whole LFS/LFL cohort ("All"; p=0.044 for rs17878362, p=0.02 for rs17880560) but was at best borderline significant in a model adjusted for the size of families in only the MUT group ("Group"; p=0.081 for rs17878362, p=0.068 for rs17880560). Mean ages at diagnosis in relation to the rs1042522 status, given as comparison in **Table 1**, did not show any significant difference in either WT2 or MUT groups in any of the family weighted models. Results for the *MDM2* rs2279744 show a non-significant tendency for later age at diagnosis in carriers of G/G as compared to T/T genotype, but no effect in T/G heterozygotes (**Supplementary Table 6**).

We repeated these analyses after separating the MUT group in two sub-groups, one with R337H carriers and the second with carriers of other germline *TP53* mutations (**Supplementary Table 7**). Compatible with Supplementary Table 2, patients with R337H had their first diagnosis on average 10-12 years later than patients with other *TP53* mutations. Nevertheless, an effect on age of cancer onset was seen in both mutation groups. In R337H, the difference in age at diagnosis was, on average, of 15.21 years between rs17878362 A1/A1 and A1/A2 carriers, and of 11.14 years between rs17880560 A1/A1 and A1/A2 carriers. In

For Peer Review

carriers of other mutations, these differences were of 9.64 years and 18.73 years, respectively. However, due to small numbers in each of the groups and lack of power, these differences were not statistically significant.

The effects of rs17878362, rs17880560 and rs1042522 genotypes on Kaplan-Meier disease-free probability estimates are shown in **Figure 2**. No effects of any of the SNPs were detected on accrual of cancer until 65 years of age in the WT2 group. In contrast, in the MUT group, the presence of one A2 allele of rs17878362 or of one A2 allele of rs17880560 was associated with a reduced childhood and adolescent cancer risk. Taking as the reference the estimated risk of cancer at age 25 years in A1/A2 carriers, the relative risk for cancer before or at 25 years is 4.0 (95% CI: [2.40-6.40]) in rs17878362 A1/A1 carriers and 3.17 (95% CI: [1.83-5.17]) in A1/A1 rs17880560 carriers. Again, no effect was seen for the rs1042522 genotypes. These observations suggest that variant alleles of both rs17878362 and the rs17880560 polymorphisms are associated with a substantial protection against early life cancer in carriers of germline *TP53* mutations.

Disease-free probability estimates relation with haplotypes defined by rs17878362 and rs17880560

Linkage disequilibrium studies show that rs17878362 and rs17880560 are in strong linkage disequilibrium in all groups analyzed (p=0.024). To determine whether the two alleles may exert effects either alone or in combination, we have reconstructed *TP53* haplotypes in subjects of the MUT and WT2 groups using data for the 5 of the polymorphisms genotyped in this study, plus, in the MUT group, the germline *TP53* mutation (**Supplementary Table 7**). We next used the genotypes at rs17878362 and rs17880560 to define 8 haplotypes (with and without a *TP53* mutation). Of these haplotypes, all were represented in the WT2 or MUT series, namely A1-A1 (carrying non-duplicated forms of rs17878362 and rs17880560), A1-A2 (non-duplicated for rs17878362, duplicated for rs17880560), A2-A1 (duplicated for

For Peer Review

rs17878362 and non-duplicated rs17880560) and A2-A2 (duplicated for both rs17878362 and rs17880560). The "Brazilian founder" R337H mutation was exclusively carried by the A1-A1 haplotype. Of the 16 families with other *TP53* germline mutations, 12 carried the mutation on the same, A1-A1 haplotype. Given the predominance of this haplotype as the carrier of the germline mutation, we reasoned that the modifier effect of *TP53* polymorphisms might be due to variations in the haplotype of the remaining wild type allele, which shows much larger genetic diversity than the mutant haplotype. **Figure 3** shows Kaplan-Meier disease-free probability estimates in MUT and WT2 groups, given one "fixed" A1-A1 *TP53* haplotype (carrying a germline mutation in subjects of the MUT group and a wild type *TP53* in subjects of the WT2 group) and considering the effect of the other *TP53* allele. This analysis showed that, in patients where the haplotype carrying the germline mutation was A1-A1, the presence of either A1-A2 or A2-A1 on the remaining wild type allele was associated with a delay in the age at first cancer diagnosis (**Table 2** and **Figure 3**). A significant difference was observed in mean age at first diagnosis in the MUT group in relation to the haplotype status of the wild type allele after weighting for family size ("Group": p=0.019; "All": p=0.035) (**Table 2**). Importantly, none of the individuals with a wild type A1-A2 haplotype, and only one with a wild type A2-A1 haplotype, had developed cancer by age 25 years among carriers of a germline mutation, whereas 25% (95% CI: 15-38) of subjects with a wild type A1-A1 haplotype had developed a cancer by that age. Thus, the presence of a wild type allele carrying a duplication of either rs17878362 or rs17880560 appears to exert a strong protective effect towards early cancer occurrence in this cohort of patients with germline *TP53* mutation.

DISCUSSION

LFS/LFL is a heterogeneous familial predisposition syndrome with risk of multiple cancers and about 90% penetrance over lifetime in carriers of a germline *TP53* mutation. However,

For Peer Review

there are considerable inter-individual and familial variations in cancer patterns and age at disease onset, raising the hypothesis that a number of other genetic, epigenetic or lifestyle factors may affect the course of the disease. Guanine-rich tracts involved in G4 motifs may represent important targets for genetic polymorphisms with significant functional effects. The enrichment of G4 in key chromosomal regions has suggested a functional role for these motifs in genomic regulation. Data collected from human SNP databases indicated that guanine triplets involved in G4 are more conserved and less polymorphic than their neutral counterparts (40). However, a recent analysis of SNPs occurring within predicted G4 motifs in the promoter of 48 genes has demonstrated a strong correlation between G4 SNPs and expression of the corresponding gene at the individual level (41). These observations support the hypothesis that G4 SNPs may play significant roles in the mechanisms of variations in gene expression among individuals. In this study, we have systematically searched the sequence of the *TP53* locus for areas that may form G4s and we have analyzed polymorphisms in these areas in search of genetic variations that may modulate disease onset in *TP53* mutation carriers and therefore account for at least part of individual and familial variations in LFS/LFL.

Among 5 regions predicted to form G4s in *TP53*, our results identify two polymorphisms, rs17878362 and rs17880560, with a frequency of above 4% that appeared to modulate the age at first cancer onset in *TP53* mutation carriers, but not in subjects with LFS/LFL traits that do not carry a mutation in *TP53*. rs17878362 occurs within a previously demonstrated G4 located in intron 3 (28,33) and rs17880560 overlaps with a predicted G4 in the 3'flanking region of p53 pre-mRNA, the exact structure of which has not yet been studied at the molecular level. Interestingly, the 3'flanking region contains a succession of several G4s. Of these one has been experimentally demonstrated [33] but is distinct from the predicted G4 structure overlapping with rs17880560.

For Peer Review

Both the rs17878362 and rs17880560 polymorphisms consist of duplications introducing additional runs of guanines, which may change the composition, position, structure or stability of the G4s. For both G4 regions, effects were seen in association with the duplication A2 alleles. Compared to the non-duplicated A1 alleles, the presence of one A2 allele of either rs17878362 or rs17880560 retarded by 12.5 to 18.0 years the age at first cancer diagnosis in *TP53* mutation carriers. As a comparison, we used a large group of subjects from families with cancer patterns corresponding to LFS/LFL traits but in whom no familial *TP53* mutation had been identified. Neither of the two polymorphisms had a detectable impact on age at first cancer diagnosis or on cancer-free survival in subjects without mutation. To assess the statistical significance of the effects observed in mutation carriers, we performed a t-test on average age at diagnosis, using a model weighted for family structure and number of members in each family, in order to rule out bias due to family-specific traits that may influence the clinical presentation of LFS/LFL. Although this approach decreases statistical power, it allows clarifying effects that are solely due to the presence of particular allele combinations. Using this method, we detected that each of the polymorphisms had a borderline significant effect on age at first cancer diagnosis (p=0.082 for rs17878362; p=0.067 for rs17880560). However, the magnitude of the effect was large (12.5 to 18.0 years) and separated subjects with cancer diagnosis in childhood or adolescence from subjects with diagnosis in adulthood. Taking as reference the risk of developing cancer before the age 25 years in carriers of one A2 allele of either polymorphism, the relative risk of early cancer in mutation carriers was 4.0 (95% CI: 2.40-6.40) for rs17878362 A1/A1 and 3.17 (95% CI: 1.83-5.17) for rs17880560 A1/A1. By comparison, the common non-synonymous SNP at codon 72 (rs1042522, *TP53* PEX4, R72P) had no effect on age at first cancer diagnosis in this cohort.

For Peer Review

Given that rs17878362 and rs17880560 are in linkage disequilibrium, we reconstructed *TP53* haplotypes to estimate the joint effects of these polymorphisms. We made the assumption that the effects of these polymorphisms were associated with the haplotype of the wild type allele, rather than the mutant allele. This assumption is based on the observation that, in our series of Brazilian families with *TP53* mutation, the mutation is almost systematically located on a haplotype carrying the A1 alleles of both rs17878362 and rs17880560. Thus, the effect of alleles carrying A2 haplotypes, if any, is associated with the properties of the residual wild type allele and with its capacity to compensate the functional defect caused by the mutant allele. Using this framework, we found that presence of haplotypes defined by either A1 rs17878362 - A2 rs17880560 or A2 rs17878362 - A1 rs17880560 had a significant effect on delaying the age at first cancer diagnosis. Among 91 documented subjects carrying a germline mutation, only one of 29 having one of these two haplotypes developed cancer before 25 years of age (3.4%), whereas 15 of 57 having the wild type A1-A1 haplotype developed cancer by this age (26.3%). We therefore conclude that, when the mutation is present on A1-A1 haplotype (the most common haplotype in all populations, in particular Caucasians and Asians), the presence of one A2 variant of either rs17878362 or rs17880560 on the residual wild type allele is protective against the risk of cancer in childhood and adolescence. This effect is compatible with a simple model in which, in the presence of a mutant allele with partial or total loss of function, *TP53* suppressor activity is contributed by the residual wild type allele (**Figure 4**). Thus, wild type haplotypes carrying one A2 allele of either rs17878362 or rs17880560 would define "strong" *TP53* haplotypes, contributing wild type *TP53* activity sufficient to suppress cancer until at least early adulthood, whereas A1-A1 wild type haplotypes would be "weak' *TP53* haplotypes, unable to compensate the loss of function of the mutant haplotype and thus being permissive for early cancer onset.

For Peer Review

The present study differs from previous ones on modifier polymorphisms in *TP53* by several characteristics. First, we have used a weighted model to take into account family size and control for possible bias cause by specific familial traits. Second, we have used as reference a group of LFS/LFL families with no mutation (WT2) rather than a heterogeneous group of "wild type" subjects pooling the members of *TP53* families who have not inherited the mutation (and therefore should be considered as low-risk for cancer) and members of families with no mutations (some of them likely to be at high-risk due to germline mutation in yet unidentified gene(s) other than *TP53*). Third, our series is constituted of LFS/LFL families recruited in Brazil, whereas previous studies were based on Western European or North American cohorts (9,13,14). A common mutant haplotype, R337H, is present in 0.3% of the population of South-Eastern Brazil and represents over half of the germline *TP53* mutations in Brazilian families with LFS/LFL traits (19,35). This mutation is carried by the same haplotype in all families and is considered to result from a widespread founder effect. Since its initial identification in the germline of children with adrenal cortical carcinoma (42), there has been debate as whether R337H should be considered as predisposing to LFS/LFL or only to specific tumor types. Our results on LFS/LFL families have shown that carriers of R337H are prone to develop a spectrum of cancers that covers the whole LFS/LFL spectrum, although the average age at first diagnosis for adolescent and adult cancers tend to be up to 10 years later than in carriers of classic, DNA-binding domain *TP53* mutations commonly associated with LFS (19). In the present study, we can rule out that the modifier effects of rs17878362 and rs17880560 polymorphisms are restricted to this "Brazilian founder" allele because we observe effects of similar type and amplitude in Brazilian carriers of "classic" *TP53* mutations, who have cancer patterns and age at onset identical to Caucasian *TP53* mutation carriers from North America or Western Europe (data not shown).

For Peer Review

The molecular mechanisms responsible for the association of A2 alleles of rs17878362 or rs17880560 with late age at first cancer onset are a matter of conjecture. The G4s containing rs17878362 (intron 3) has been shown to exert a regulatory effect on alternative splicing of intron 2 and on the generation of p53I2, a p53 mRNA that encodes the p53 protein isoform Δ40p53. In animal models, homologs of Δ40p53 down-regulate stem cells and promote aging (in mice) and induce abnormal development (in Zebrafish) exclusively when co-expressed together with full-length, wild type p53 (43,44). The precise structure and biological effect of the predicted polymorphic G4 in the 3'flanking region remain to be demonstrated.

To conclude, this study has identified polymorphisms associated with G4 structures, which modify the penetrance of germline *TP53* mutation in LFS/LFL. They have a general effect on the mean age at first cancer diagnosis but, more importantly, they distinguish patients who are at high risk of developing childhood/adolescent cancers from those who develop mainly, if not exclusively, cancers during adulthood. These findings require evaluation and replication in independent LFS/LFL cohorts but may provide an important marker for developing personalized strategies for surveillance of cancer risk in carriers of germline *TP53* mutations.

SUPPLEMENTARY MATERIAL: Supplementary Figure 1 and Supplementary Tables 1, Supplementary Table 2, Supplementary Table 3, Supplementary Table 4, Supplementary Table 5, Supplementary Table 6 and Supplementary Table 7 can be found at http://carcin.oxfordjournals.org/

ACKNOWLEDGEMENTS

This work was supported by grant INCa 2009-192 "*TP53* intron3" from the French National Cancer Institute to JH, PH and JLM. Research in Inserm U612 is also supported by funding

For Peer Review

from Institut Curie and Inserm. CS has a PhD fellowship from the French Ministry of Research and VM and LP were supported by funding from EU FP7 (Grant Number 249689 for the network of excellence DoReMi (low dose research towards multidisciplinary integration). DGC is support by INSERM. PP was supported by funding from Conselho Nacional de Pesquisa (CNPq grand number 40.0949/2005-9) and Susan G Komen for the Cure (Grant POP 0403033).

CONFLICT OF INTEREST STATEMENT: None declared

REFERENCES

1. Li, F.P., *et al.* (1988) A cancer family syndrome in twenty-four kindreds. *Cancer Res*, **48**, 5358-62.
2. Mai, W.J., *et al.* (2012) Characterization of the tilapia p53 gene and its role in chemical-induced apoptosis. *Biotechnol Lett*, **34**, 1797-805.
3. Birch, J.M., *et al.* (1994) Prevalence and diversity of constitutional mutations in the p53 gene among 21 Li-Fraumeni families. *Cancer Res*, **54**, 1298-304.
4. Eeles, R.A. (1995) Germline mutations in the TP53 gene. *Cancer Surv*, **25**, 101-24.
5. Chompret, A. (2002) The Li-Fraumeni syndrome. *Biochimie*, **84**, 75-82.
6. Malkin, D., *et al.* (1990) Germ line p53 mutations in a familial syndrome of breast cancer, sarcomas, and other neoplasms. *Science*, **250**, 1233-8.
7. Srivastava, S., *et al.* (1990) Germ-line transmission of a mutated p53 gene in a cancer-prone family with Li-Fraumeni syndrome. *Nature*, **348**, 747-9.
8. Varley, J.M., *et al.* (1997) Germ-line mutations of TP53 in Li-Fraumeni families: an extended study of 39 families. *Cancer Res*, **57**, 3245-52.
9. Ruijs, M.W., *et al.* (2010) TP53 germline mutation testing in 180 families suspected of Li-Fraumeni syndrome: mutation detection rate and relative frequency of cancers in different familial phenotypes. *J Med Genet*, **47**, 421-8.
10. Tinat, J., *et al.* (2009) 2009 version of the Chompret criteria for Li Fraumeni syndrome. *J Clin Oncol*, **27**, e108-9; author reply e110.
11. Bougeard, G., *et al.* (2008) Molecular basis of the Li-Fraumeni syndrome: an update from the French LFS families. *J Med Genet*, **45**, 535-8.
12. Villani, A., *et al.* (2011) Biochemical and imaging surveillance in germline TP53 mutation carriers with Li-Fraumeni syndrome: a prospective observational study. *Lancet Oncol*, **12**, 559-67.
13. Bond, G.L., *et al.* (2004) A single nucleotide polymorphism in the MDM2 promoter attenuates the p53 tumor suppressor pathway and accelerates tumor formation in humans. *Cell*, **119**, 591-602.

For Peer Review

14. Bougeard, G., *et al.* (2006) Impact of the MDM2 SNP309 and p53 Arg72Pro polymorphism on age of tumour onset in Li-Fraumeni syndrome. *J Med Genet*, **43**, 531-3.
15. Marcel, V., *et al.* (2009) TP53 PIN3 and MDM2 SNP309 polymorphisms as genetic modifiers in the Li-Fraumeni syndrome: impact on age at first diagnosis. *J Med Genet*, **46**, 766-72.
16. Ruijs, M.W., *et al.* (2007) The single-nucleotide polymorphism 309 in the MDM2 gene contributes to the Li-Fraumeni syndrome and related phenotypes. *Eur J Hum Genet*, **15**, 110-4.
17. Wu, C.C., *et al.* (2011) Joint effects of germ-line TP53 mutation, MDM2 SNP309, and gender on cancer risk in family studies of Li-Fraumeni syndrome. *Hum Genet*, **129**, 663-73.
18. Knappskog, S., *et al.* (2011) Effects of the MDM2 promoter SNP285 and SNP309 on Sp1 transcription factor binding and cancer risk. *Transcription*, **2**, 207-10.
19. Achatz, M.I., *et al.* (2007) The TP53 mutation, R337H, is associated with Li-Fraumeni and Li-Fraumeni-like syndromes in Brazilian families. *Cancer Lett*, **245**, 96-102.
20. Bochman, M.L., *et al.* (2012) DNA secondary structures: stability and function of G-quadruplex structures. *Nat Rev Genet*, **13**, 770-80.
21. Ribeyre, C., *et al.* (2009) The yeast Pif1 helicase prevents genomic instability caused by G-quadruplex-forming CEB1 sequences in vivo. *PLoS Genet*, **5**, e1000475.
22. Paeschke, K., *et al.* (2011) DNA replication through G-quadruplex motifs is promoted by the Saccharomyces cerevisiae Pif1 DNA helicase. *Cell*, **145**, 678-91.
23. Paeschke, K., *et al.* (2013) Pif1 family helicases suppress genome instability at G-quadruplex motifs. *Nature*.
24. Cayrou, C., *et al.* (2012) New insights into replication origin characteristics in metazoans. *Cell Cycle*, **11**.
25. Millevoi, S., *et al.* (2012) G-quadruplexes in RNA biology. *Wiley Interdiscip Rev RNA*, **3**, 495-507.
26. Halder, R., *et al.* (2012) Bisquinolinium compounds induce quadruplex-specific transcriptome changes in HeLa S3 cell lines. *BMC Res Notes*, **5**, 138.
27. Gomez, D., *et al.* (2004) Telomerase downregulation induced by the G-quadruplex ligand 12459 in A549 cells is mediated by hTERT RNA alternative splicing. *Nucleic Acids Res*, **32**, 371-9.
28. Marcel, V., *et al.* (2011) G-quadruplex structures in TP53 intron 3: role in alternative splicing and in production of p53 mRNA isoforms. *Carcinogenesis*, **32**, 271-8.
29. Ghosh, A., *et al.* (2004) Regulation of human p53 activity and cell localization by alternative splicing. *Mol Cell Biol*, **24**, 7987-97.
30. Courtois, S., *et al.* (2002) DeltaN-p53, a natural isoform of p53 lacking the first transactivation domain, counteracts growth suppression by wild-type p53. *Oncogene*, **21**, 6722-8.
31. Yin, Y., *et al.* (2002) p53 Stability and activity is regulated by Mdm2-mediated induction of alternative p53 translation products. *Nat Cell Biol*, **4**, 462-7.
32. Marcel, V., *et al.* (2011) Biological functions of p53 isoforms through evolution: lessons from animal and cellular models. *Cell Death Differ*, **18**, 1815-24.

For Peer Review

33. Decorsiere, A., *et al.* (2011) Essential role for the interaction between hnRNP H/F and a G quadruplex in maintaining p53 pre-mRNA 3'-end processing and function during DNA damage. *Genes Dev*, **25**, 220-5.
34. Petitjean, A., *et al.* (2007) Impact of mutant p53 functional properties on TP53 mutation patterns and tumor phenotype: lessons from recent developments in the IARC TP53 database. *Hum Mutat*, **28**, 622-9.
35. Garritano, S., *et al.* (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**, 143-50.
36. Kikin, O., *et al.* (2006) QGRS Mapper: a web-based server for predicting G-quadruplexes in nucleotide sequences. *Nucleic Acids Res*, **34**, W676-82.
37. Scaria, V., *et al.* (2006) Quadfinder: server for identification and analysis of quadruplex-forming motifs in nucleotide sequences. *Nucleic Acids Res*, **34**, W683-5.
38. Huppert, J.L., *et al.* (2005) Prevalence of quadruplexes in the human genome. *Nucleic Acids Res*, **33**, 2908-16.
39. Stacey, S.N., *et al.* (2011) A germline variant in the TP53 polyadenylation signal confers cancer susceptibility. *Nat Genet*, **43**, 1098-103.
40. Nakken, S., *et al.* (2009) The disruptive positions in human G-quadruplex motifs are less polymorphic and more conserved than their neutral counterparts. *Nucleic Acids Res*, **37**, 5749-56.
41. Baral, A., *et al.* (2012) Quadruplex-single nucleotide polymorphisms (Quad-SNP) influence gene expression difference among individuals. *Nucleic Acids Res*, **40**, 3800-11.
42. Ribeiro, R.C., *et al.* (2001) An inherited p53 mutation that contributes in a tissue-specific manner to pediatric adrenal cortical carcinoma. *Proc Natl Acad Sci U S A*, **98**, 9330-5.
43. Maier, B., *et al.* (2004) Modulation of mammalian life span by the short isoform of p53. *Genes Dev*, **18**, 306-19.
44. Davidson, W.R., *et al.* (2010) Differential regulation of p53 function by the N-terminal DeltaNp53 and Delta113p53 isoforms in zebrafish embryos. *BMC Dev Biol*, **10**, 102.

TABLE AND FIGURES LEGENDS

Fig. 1: Predicted localization of G4 and associated polymorphisms in *TP53*. (**A**) Schematic representation of G4 localization and polymorphisms at the *TP53* locus. G1 to G5: position of predicted G4 motifs. A to G: position of polymorphisms within or near these G4. S1: sequence 1 encompassing G3; S2: Sequence 2 encompassing G5. (**B**) Sequence of S1 region, from intron 2 to intron 3 (small letters), encompassing exon 3 (capital letters). Two G-tracts domains forming putative G4 are shown in bold. The position of polymorphisms B in intron 2 (G/C, rs1642785) and C in intron 3 (16 bp duplication, rs17878362) are underlined

For Peer Review

(**C**) Sequence of S2 region encompassing the 3'flanking region of *TP53* after cleavage site. Bold letters: G-tracts forming putative G4. The position of polymorphisms F (6 bp duplication, rs17880560) and G (C/T, rs1614984) are underlined.

Fig. 2: Genotype dependent Kaplan-Meier disease-free probability estimates in LFL/LFS family members with or without *TP53* mutations. Kaplan-Meir probability is shown for rs17878362 (**A**), rs17880560 (**B**) and rs1042522 (**C**). In each panel, the left panel corresponds to subjects of the MUT group (*TP53* mutation carriers) and the right panel to subjects of the WT2 group (families with no mutation detected). The tables under the graphs show disease-free probability estimates at different ages (10, 30, 50 and 65 years) according to genotype.

Fig. 3: Kaplan-Meier disease-free probability estimates based on the haplotype of the wild-type allele. Kaplan-Meir probabilities in MUT (*TP53* mutation carriers in LFL/LFS family members) (**A**) and WT2 groups (LFL/LFS families with no mutation detected) (**B**) are shown. In the MUT group, the *TP53* mutation is present on a haplotype defined as A1A1 (non-duplicated variant for both polymorphisms rs17878362-rs17880560), and the remaining wild type haplotype is shown. In the WT2 group, in which subjects carry two wild type alleles, one of the alleles has been "fixed" as the A1A1 haplotype and the effect of the other haplotype is shown. Tables under the graphs show disease-free probability estimates at different ages (10, 30, 50 and 65 years) according to haplotypes (A1A1: non-duplicated for both polymorphisms; A1A2: non-duplicated for rs17878362 and duplicated for rs17880560; A2A1: duplicated for rs17878362 and non-duplicated for rs17880560).

For Peer Review

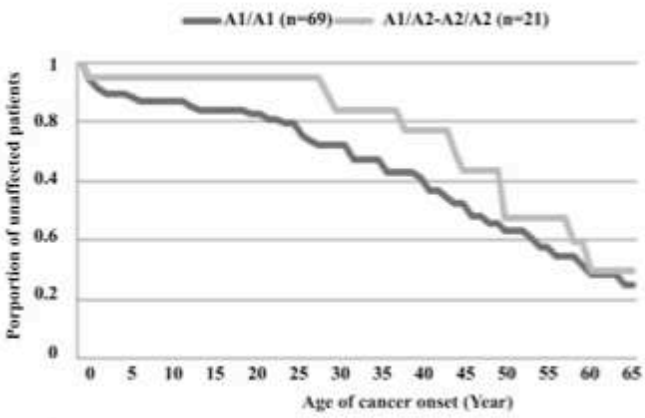
Fig. 4: Effect of different wild type haplotypes in *TP53* mutation carriers: a model. TP53 alleles are represented as rods. Left, mutant allele occurring on a haplotype carrying A1 variants of both rs17878362 and rs17880560 (A1A1). Right, different types of wild type haplotypes. The wild type haplotype defined by A1A1 is considered as a "weak" haplotype (associated with early cancer, indicative low capacity to compensate the loss of p53 function of the mutant allele). The wild type haplotypes defined by A1A2 or A2A1 are considered as "strong haplotype" (associated with later cancer onset, thus providing at least partial compensation for the loss of function of the mutant allele). Of note, our data do not predict the effect of wild type A2A2 haplotypes, or the effects of these haplotypes when the mutation occurs on another haplotype than A1A1.

Supplementary Fig. 1: Schematic representation of the localization of putative G4 and associated SNPs in *TP53*. Sequences of the regions containing G4-1 in the proximal part of intron 1 (**A**), G4-2 in the distal part of intron 1 (**B**) and G4-4 in intron 6 are shown (**C**). G-tracts domains forming putative G4 are shown in bold. The position of polymorphisms is underlined.

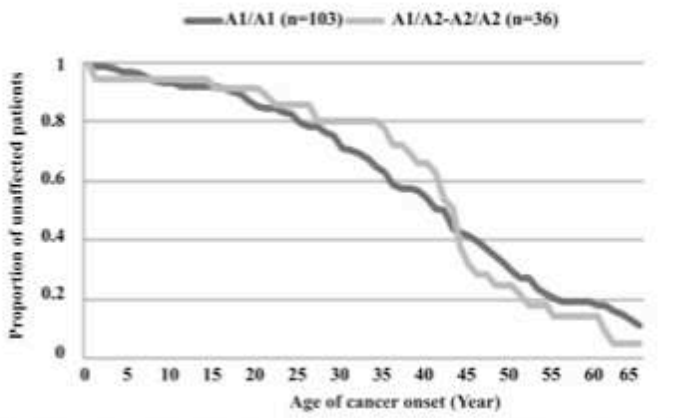
For Peer Review

A
G4-1 G4-2
G4-3
S1
G4-4
G4-5
S2
A B C D E
F; G
B
gtgagtggat ccattggaag ggcaggcccca ccacccccca cccaacccca gccccctagc agagacctgt gggaagcgaa aattccatgg gactgacttt
rs1642785
ctcgtgtctt ctttcagACT TCCTGAAAAC AACGTTCTGG taaggacaagg gttgggctg gggacctgga gggctgggga cctgaaggac tgggggactg
rs17878362
gggggctgag gacctgatcc tctgactgct cttttcaccc atctacag
C
aacgccagtg caggctactg gggtcagcag gtgcaggggt gagtgaggag gtgctgggaa gcagccacct gagtctgcaa tgagtgagg ctgggggacc
cagtgccagg gttccgggag gggaacaaag gctgggact gggtcagtat gcgggctgca tgacaacaag ggagggggtg gctccattca taactcagga
acaaacgtc ccccctccc tccgggcaag gcacgggtgg cacaaggttc tcccccccc ctgcttctag gactgggctg
rs17880560
rs1614984

A

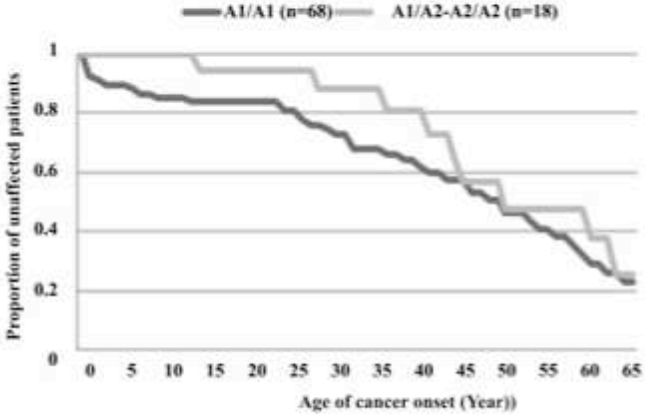


Age (Year)	10	30	50	65
A1/A1	0.87 (0.76-0.93)	0.72 (0.60-0.82)	0.43 (0.32-0.56)	0.25 (0.16-0.37)
A1/A2-A2/A2	0.95 (0.73-1.00)	0.84 (0.60-0.95)	0.47 (0.26-0.70)	0.30 (0.13-0.54)

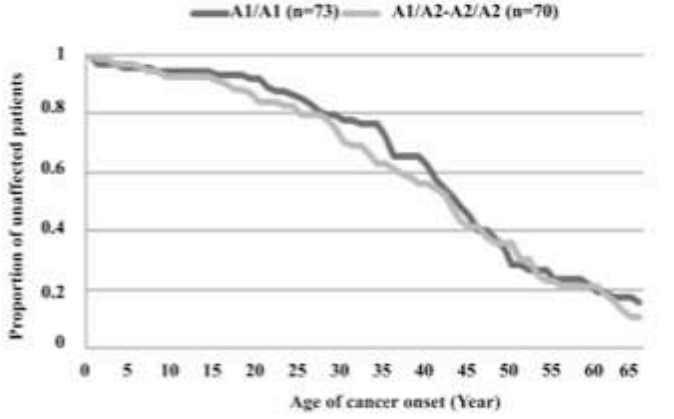


Age (Year)	10	30	50	65
A1/A1	0.93 (0.86-0.97)	0.71 (0.61-0.80)	0.30 (0.21-0.39)	0.11 (0.06-0.19)
A1/A2-A2/A2	0.94 (0.80-0.99)	0.81 (0.63-0.91)	0.25 (0.13-0.42)	0.05 (0.01-0.19)

B

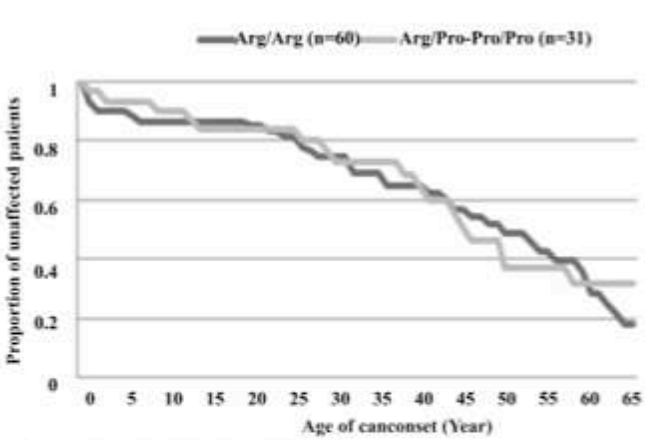


Age (Year)	10	30	50	65
A1/A1	0.85 (0.74-0.92)	0.73 (0.61-0.83)	0.46 (0.34-0.58)	0.23 (0.14-0.35)
A1/A2-A2/A2	1.00 (0.78-1.00)	0.88 (0.64-0.98)	0.47 (0.25-0.71)	0.25 (0.09-0.51)

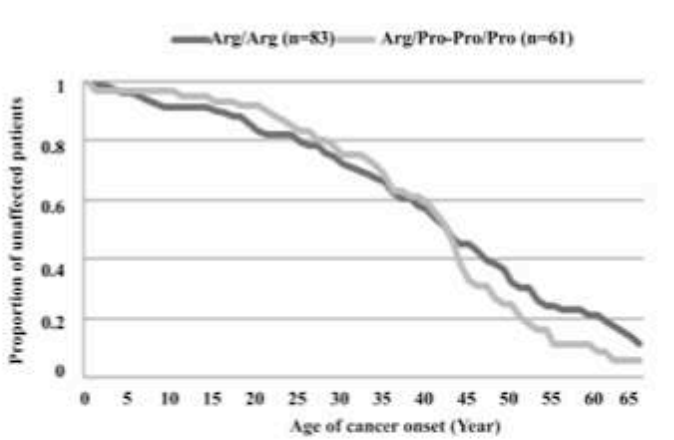


Age (Year)	10	30	50	65
A1/A1	0.95 (0.86-0.98)	0.78 (0.67-0.87)	0.29 (0.19-0.41)	0.16 (0.09-0.27)
A1/A2-A2/A2	0.93 (0.83-0.97)	0.71 (0.58-0.81)	0.36 (0.25-0.48)	0.11 (0.05-0.21)

C

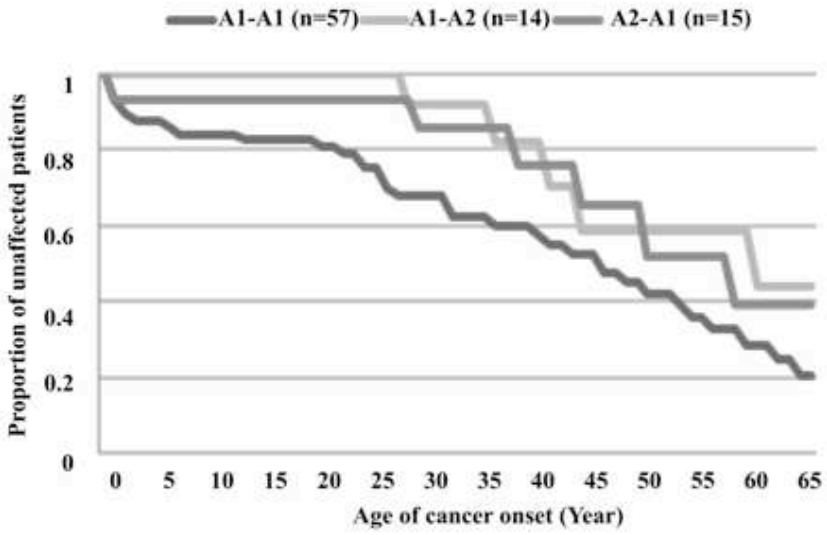


Age (Year)	10	30	50	65
Arg/Arg	0.87 (0.75-0.94)	0.75 (0.61-0.85)	0.49 (0.36-0.62)	0.18 (0.10-0.30)
Arg/Pro-Pro/Pro	0.90 (0.73-0.97)	0.73 (0.54-0.87)	0.37 (0.21-0.56)	0.32 (0.17-0.51)



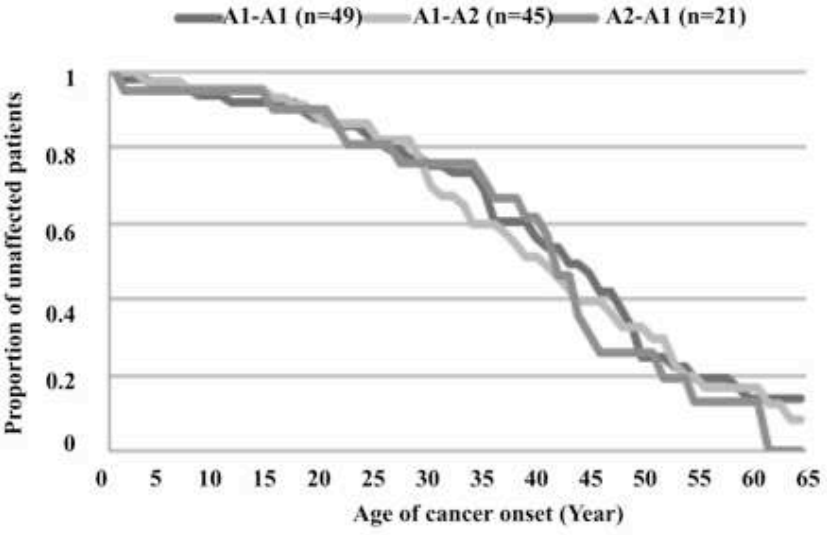
Age (Year)	10	30	50	65
Arg/Arg	0.92 (0.83-0.96)	0.72 (0.61-0.81)	0.32 (0.22-0.43)	0.11 (0.06-0.20)
Arg/Pro-Pro/Pro	0.97 (0.88-0.99)	0.75 (0.62-0.85)	0.25 (0.15-0.37)	0.06 (0.02-0.16)

A



Age (Year)	10	30	50	65
A1-A1	0.84 (0.72-0.92)	0.68 (0.54-0.79)	0.42 (0.30-0.56)	0.20 (0.11-0.34)
A1-A2	1.00 (0.73-1.00)	0.92 (0.64-1.00)	0.59 (0.31-0.82)	0.44 (0.20-0.71)
A2-A1	0.93 (0.66-1.00)	0.86 (0.57-0.97)	0.52 (0.26-0.77)	0.39 (0.17-0.66)

B



Age (Year)	10	30	50	65
A1-A1	0.94 (0.82-0.98)	0.7((0.61-0.86)	0.25 (0.14-0.49)	0.14 (0.06-0.27)
A1-A2	0.96 (0.84-0.99)	0.70 (0.54-0.82)	0.33 (0.20-0.49)	0.08 (0.03-0.21)
A2-A1	0.95 (0.74-1.00)	0.76 (0.52-0.91)	0.26 (0.10-0.50)	0.00 (0.00-0.19)

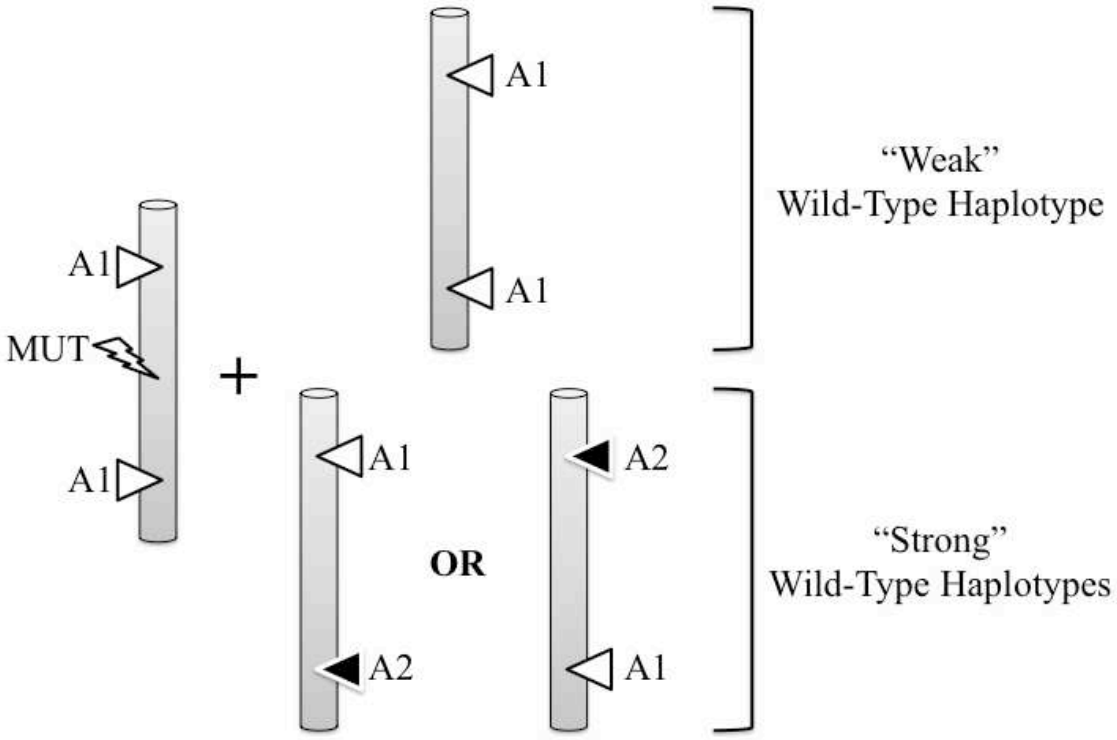

A1
A1
"Weak"
Wild-Type Haplotype
A1
MUT
A1
+
A1
A2
OR
A2
A1
"Strong"
Wild-Type Haplotypes

Table 1: Mean age at 1st cancer diagnosis depending on *TP53* polymorphism genotypes and *TP53* mutation family.

			WT2			MUT		
			n	**Mean**	**SD**	**n**	**Mean**	**SD**
rs17878362	**A1/A1**		90	37.91	17.06	40	30.65	19.84
	A1/A2		29	39.14	15.22	11	43.36	19.00
	A2/A2		2	35.50	12.02	1	9.00	-
	***P* value**	**Group**	0.959			0.081		
		All	0.747			**0.044**		
		Without	0.918			0.088		
rs17880560	**A1/A1**		59	39.56	16.43	40	32.58	21.36
	A1/A2		43	38.19	16.35	8	45.88	11.68
	A2/A2		12	32.67	18.60	1	14.00	-
	***P* value**	**Group**	0.410			0.068		
		All	0.442			**0.020**		
		Without	0.426			0.152		
rs1042522	**G/G***		68	37.96	18.09	35	33.29	21.11
	G/C*		46	37.76	14.87	16	35.63	19.89
	C/C*		7	42.86	9.63	2	19.50	14.85
	***P* value**	**Group**	0.569			0.440		
		All	0.609			0.545		
		Without	0.742			0.582		

WT2: WT subjects in WT families
MUT: Mutant subjects in mutant families
SD: Standard deviation
Group: *P value* calculated with weight in the analyzed group. In bold, significative *P value*
All: *P value* calculated with weight in the all cohort. In bold, significative *P value*
Without: *P value* without weight. In bold, significative *P value*
*: G corresponds to the Arginine and C to the Proline variants

Table 2: Mean age at 1st cancer diagnosis depending on *TP53* haplotypes and *TP53* mutation family status. One allele, corresponding to the frequent allele (rs17878362-A1 and rs17880560-A1), is fixed and the haplotype frequencies correspond to the second allele.

rs17878362	rs17880560	WT2			MUT		
		n	Mean	SD	n	Mean	SD
A1	**A1**	42	39.36	17.36	35	29.31	20.15
A1	**A2**	35	37.11	15.89	6	46.47	15.05
A2	**A1**	18	37.17	15.05	7	41.71	22.69
A2	**A2**	3	45.33	10.69	1	50.00	-
	Group	0.864			**0.019**		
***P* value**	**All**	0.869			**0.035**		
	Without	0.800			0.138		

WT2: WT subjects in WT families
MUT: Mutant subjects in mutant families
SD: Standard deviation
Group: *P value* calculated with weight in the analyzed group. In bold, significative *P value*
All: *P value* calculated with weight in the all cohort. In bold, significative *P value*
Without: *P value* without weight. In bold, significative *P value*

Supplementary materials

Age at cancer onset in germline *TP53* mutation carriers: association with polymorphism in predicted G-quadruple structures

Charlotte Sagne, Virginie Marcel, Maria Bota, Ghyslaine Martel-Planche, Amanda Nobrega, Edenir Inêz Palmero, Laury Perriaud, Mathieu Boniol, Stephan Vagner, David G. Cox, Chang S. Chan, Jean-Louis Mergny, Magali Olivier, Patricia Prolla, Janet Hall, Pierre Hainaut, Maria Isabel Achatz

TCCCATGTGC TCAAGACTGG CGCTAAAAGT TTTGAGCTTC TCAAAAGTCT AGAGCCACCG TCCAGGGAGC AGGTAGCTGC TGGGCTCCGG GGACACTTTG

CGTTCGGGCT GGGAGCGTGC TTTCCACGAC GGTGACACGC TTCCCTGGAT TGGgtaagct cctgactgaa cttgatgagt cctctctgag tcacgggctc

tcggctccgt gtattttcag ctcgggaaaa tcgct**ggggc tgggggtggg gcagtgggg**a cttagcgagt ttgg**gggtga gtgggatgga agcttgg**cta

gagggatcat cataggagtt gcattgttg**g gagacctggg tgtagatgat ggggg**atgtta ggaccatccg aactcaaagt tgaacgccta **ggcagaggag**

tggagctttg gggaaccttg agccggccta aagcgtactt ctttgcacat ccacccggtg ct

ggagtggttg ttaagagatg agattaaaga agccgagacg ggccattcgt gag**gggtttg taatgcaggg ctgagg**agtg tccgaagaga atgggcaggt

gagcggtgag acagttgttc ttccagaagc tttgcagtga aaggaatcaa agaaatggag ccgtgtatca ggt**ggggaag ggtgggggcc aaggggg**tgt

ccttccccat acagagattg caggctgaga atgactatct ccttgttaac a**ggaggtggg agcagggcac gg**tagctcac acctgtaatc ttggcacttt

a**ggaggctga ggcggg**ccga tcacctgaag taaggagttc gagaccagcc tggccaacat gcaaagccct gtctctacta aaaatacaaa aattagctgg

rs12944939

gtgtggtggt actcgcctgt aatcccagct actcgggaga ctgaggca**gg agaatggctt gaacccggaa ggtagagg**tt gcagtcagct gagatcatgc

cactgtgctc cagcctaggt gacagagaga gactccatca caaaaaaaaa aaaaaaatac aggaagggag ttgggaatag ggtgca

ctggtttgca act**ggggtct ctgggaggag gggttaaggg** tggttgtcag t**ggccctcca ggtgagcagt agggggg**ctt tctcctgctg cttatttgac

rs1625895

ctccctataa ccccatgaga tgtgcaaagt aaatgggttt aactattgca cagttgaaaa aactgaagct tacagaggct aag**ggcctcc cctgcttggc**

tgggcgcagt ggctcatgcc tgtaatccca gcactttggg aggccaagg

Supplementary Table 1: Overall description of the cohort

FAMILIY DESCRIPTION (n=145)		
	n	**(%)**
LFS Criteria*		
Classic	15	(10.34)
Chompret	63	(43.45)
Birch	3	(2.07)
Eeles	64	(44.14)
Mutation status		
WT	110	(75.86)
R337H	19	(13.10)
Others mutations	16	(11.03)
SUBJECT DESCRIPTION (n=402)		
	n	**(%)**
Sex		
Male	141	(36.15)
Female	249	(63.84)
Mean age and cancer		
Mean age (year)	46	
Cancer	196	(48.76)
Mean age first diagnosis of cancer (year)	37	
Mutation Status		
WT in WT family (WT2)	157	(39.05)
WT in MUT family (WT1)	147	(36.57)
R337H	70	(17.41)
Others mutations	28	(6.97)

*: see definitions of criteria at http://p53.iarc.fr/Download/LFSdefs.pdf

For Peer Review

Supplementary Table 2: Primers and PCR conditions used for *TP53* mutation screening, *TP53* and *Mdm2* genotyping. PCR programme A corresponds to 94°C 2 min – 50x (94°C 30 sec – Tm 45sec – 72°C 45 sec) – 72°C 10 min; B to 94°C 2 min – 20x (94°C 30sec – Tm_1 decreased of 0.5°C each 3 cycles 45sec – 72°C 1 min) – 30x (94°C 30 sec – Tm_2 45 sec – 72°C 1 min) – 72°C 10 min; and C to 95°C 15 min – 50x (94°C 30 sec – Tm 30 sec – 72°C 30 sec) – 72°C 10 min. In bold italics are the polymorphisms, which showed allelic variation in the study with a MAF greater than 4% and in bold these presenting a MAF less than 4%..

Gene	Regions	Primer sequences (5'-3') Forward	Reverse	PCR programme	Polymorphisms
TP53	**Intron1 start**	tcccatgtgctcaagactgg	agcaccgggtggatgtgca	A (Tm 62°C)	rs17884410 rs34686922 rs17886250 rs55680224 rs145410568 rs17883908 rs55651612
	Intron1 end	gggagtggttgttaagagatg	tgcaccctattcccaactcc	A (Tm 62°C)	rs1794285 rs183764174 rs113868252 rs12944960 rs1642784 ***rs12944939*** (=rs17880892) rs191229483 rs4380105 rs147856400 rs9894227 rs34576726 (=rs151112663=rs71856623=rs138734480 =rs148946608=rs71919373) rs138932666
	Exons 2 to 4	tctcatgctggatccccact	atacggccaggcattgaagt	A (Tm 61°C)	rs143458271 rs201382018 rs1800369 rs184743157 rs68134313 (=rs68134314=rs68134315) rs2307495 (=rs34067591=rs34967591) ***rs1642785* = *TP53* PIN2** (=rs17880422) rs2307496 (=rs25584095=rs34275090) rs201753350 rs200053580

For Peer Review

rs139100308
rs138066738
rs1502000764
rs66510636
***rs17878362* = *TP53* PIN3**
rs146534833
rs67651209 (=rs67651211=rs67651210)
rs138445402
rs66479507
rs67112752(=rs67112760=rs67112763=rs67112764
=rs67112756=rs67112755=rs67112758=rs67112761
=rs67112757=rs67112762=rs67112759=rs67112754=rs67112753)
rs17883323 (=rs35134940)
rs200989844
rs202217267
rs35117667
rs11575998
rs121912661
rs1800370(=rs11540653=rs4134783=rs17883643=rs4987029)
rs1800371(=rs4134782=rs17883642=rs4987028)
rs201741778
rs144386518
rs142216275
rs66702588
(=rs66702590=rs66702598=rs66702591=rs66702593=rs66702599
=rs66702596=rs66702600=rs66702595=rs66702601=rs66702592
=rs66702597=rs66702602=rs66702594=rs66702589)
***rs1042522* = *TP53* PEX4**
(=rs17844988=rs17882155=rs4134781=rs2229076
=rs60388830=rs17857747=rs3174747)
rs56275308
rs201717599
rs55754907
rs66998338
rs11540654
rs121912658
rs55863639
rs67736424
rs68140816

For Peer Review

					rs67336196
	Exons 5 to 6	tgttcacttgtgccctgact	ttaacccctcctcccagaga	B (Tm_1 63°C; Tm_2 60°C)	
	Intron 6	ctggtttgcaactggggtct	ccttggcctcccaaagtgc	B (Tm_1 65°C; Tm_2 62°C)	rs200372146
					rs35163653
					rs34949160
					rs17884607
					rs1625895
					(=rs17883831)
					rs35177338 (=rs59257978)
					rs8069054 (=rs17884987)
	Exon7	aggcgcactggcctcatctt	tgtgcagggtggcaagtggc	C (Tm 60°C)	-
	Exons 8 to 9	ttgggagtagatggagcct	agtgttagactggaaacttt	B (Tm_1 63°C; Tm_2 60°C)	-
	Exon 10	caattgtaacttgaaccatc	ggatgagaatggaatcctat	B (Tm_1 63°C; Tm_2 60°C)	-
	Exon 11	agaccctctcactcatgtga	tgacgcacacctattgcaag	B (Tm_1 63°C; Tm_2 60°C)	-
	3'UTR and	acattctgcaagcacatctgc	ccgtaatccttggtgagagg	A (Tm 62°C)	rs139306032
	3'Flank				**rs78378222**
					rs17883782
					rs17884947
					rs1794294
					rs55939109
					rs147976274
					rs117562731
					rs199986063
					rs187433842
					rs1794295
					rs151186587
					rs17880560
					(=rs140608273= rs72526905= rs79948390= rs66470553)
					rs1614325
					rs1614984
					(=rs36206383=rs17881740)
	3' Flank	cctctcaccaaggattacgg	tcattcagctacctggcatg	A (Tm 62°C)	-
Mdm2		gagggctttgatgttcctga	gctactagaagttgatggc	A (Tm 63°C)	**rs2279744**

For Peer Review

Supplementary Table 3: Distribution of genotypes and HWE for 7 *TP53* polymorphisms in LFS/LFL and in 4 control groups. The 4 control groups correspond to 3 HapMap groups (Caucasian, Asian and African) and to the Brazilian Control group.

Symbol	rs number	Genotypes	HapMap									Brazilian controls			LFS/LFL (45)		
			African			Asian			Caucasian								
			n	MAF	p HWE	n	MAF	p HWE	n	MAF	p HWE	n	MAF	p HWE	n	MAF	p HWE
A	**rs12944939**	**G/G**	62	0.16	0.86	42	0.32	0.39	76	0.07	**0.02**	256	0.08	0.54	230	0.04	**<0.01**
		G/A	24			35			9			44			13		
		A/A	2			11			2			1			4		
B	**rs1642785**	**G/G**	18	0.55	0.64	18	0.55	0.58	57	0.22	**0.02**	218	0.28	0.69	204	0.26	0.07
		G/C	39			39			23			170			161		
		C/C	26			27			8			30			19		
C	**rs17878362**	**A1/A1**	50	0.27	0.21	87	0.02	0.87	71	0.10	0.32	300	0.16	0.51	255	0.18	**0.04**
		A1/A2	31			3			17			107			123		
		A2/A2	9			0			0			12			6		
D	**rs1042522**	**G/G***	39	0.33	0.71	19	0.55	0.66	57	0.21	**0.04**	208	0.30	0.51	210	0.26	0.12
		G/C*	41			42			22			169			16		
		C/C*	9			28			7			40			20		
E	**rs1625895**	**G/G**	45	0.30	0.31	85	0.02	0.87	75	0.08	**0.04**	247	0.12	0.52	101	0.13	0.83
		G/A	33			3			14			66			31		
		A/A	10			0			0			6			2		
F	**rs17880560**	**A1/A1**	38	0.34	0.67	80	0.06	0.58	47	0.27	0.72	167	0.29	0.19	234	0.22	0.26
		A1/A2	41			10			33			151			119		
		A2/A2	9			0			7			24			21		
G	**rs1614984**	**C/C**	36	0.35	0.74	38	0.37	0.25	30	0.42	0.85	132	0.28	**<0.01**	194	0.29	0.17
		C/T	41			37			42			66			143		
		T/T	10			15			16			31			37		

MAF: Minor allele frequency for Caucasian population
p HWE: *P value* of the Hardy-Weinberg equilibrium chi-square. In bold, significative *P value*
*: G corresponds to the Arginine and C to the Proline variants

For Peer Review

Supplementary Table 4: Comparison between cancer pattern in LFS/LFL cohort and the IARC *TP53* database for germline mutation.

Cancer type	IARC Database*		Brazilian LFS/LFL cohort					
			R337H		Others mutations		WT2	
	n	%	n	%	n	%	n	%
Breast	306	46.65	13	54.17	6	42.86	38	46.34
STS	153	23.32	5	20.83	5	35.71	33	40.24
Brain	131	19.97	2	8.33	1	7.14	8	9.76
ADR	66	10.06	4	16.67	2	14.29	3	3.66

* Version R16

For Peer Review

Supplementary Table 5: Mean age at 1st cancer onset according to genotypes in groups and *TP53* mutations.

			WT			WT1			R337H			Other	
			n	Mean	SD	n	Mean	SD	n	Mean	SD	n	Mean	SD
rs17878362	**A1/A1**		100	38.61	16.90	10	44.90	14.59	26	33.50	20.15	14	25.36	18.79
	A1/A2		34	41.18	15.89	5	47.75	21.06	7	48.71	13.89	4	34.00	25.18
	A2/A2		3	38.67	10.12	1	45.00	-	0	-	-	1	9.00	-
	***P* value**	**Group**	0.888			0.575			0.175			0.518		
		All	0.656			0.109			0.185			0.462		
		Without	0.894			0.957			0.071			0.523		
rs17880560	**A1/A1**		68	40.51	16.80	9	47.50	19.02	28	35.11	20.61	12	26.67	22.81
	A1/A2		46	38.24	15.93	3	39.00	9.85	4	46.25	14.52	4	45.50	10.34
	A2/A2		14	34.14	18.04	2	43.00	15.56	0	-	-	1	14.00	-
	***P* value**	**Group**	0.385			0.870			0.291			0.137		
		All	0.455			0.916			0.138			0.059		
		Without	0.401			0.762			0.308			0.244		
rs1042522	**G/G***		76	38.45	17.78	8	42.63	15.17	23	33.96	20.23	12	32.00	23.56
	G/C*		53	39.15	15.47	7	49.83	17.21	10	43.10	18.25	6	23.17	17.01
	C/C*		8	43.13	8.95	1	45.00	-	0	-	-	2	19.50	14.85
	***P* value**	**Group**	0.670			0.147			0.707			0.668		
		All	0.569			**0.036**			0.493			0.809		
		Without	0.747			0.714			0.229			0.603		

WT1: WT in mutant families
Other: *TP53* mutant carriers with a mutation different to R337H
SD: Standard deviation
Group: *P value* calculated with weight in the analyzed group
All: *P value* calculated with weight in the entire cohort
Without: *P value* without weight
*: G corresponds to the Arginine and C to the Proline variants

For Peer Review

Supplementary Table 6: Mean age at 1st cancer onset according to genotypes in groups and *TP53* mutations.

				WT2			Mut	
			n	Mean	SD	n	Mean	SD
rs12944939	**G/G**		99	37.79	16.76	39	32.10	21.19
	G/A		7	40.57	15.55	1	46.00	-
	A/A		0	-	-	0	-	-
	***P* value**	**Group**	0.675			0.458		
		All	0.511			0.684		
		Without	0.671			0.521		
rs162785	**G/G**		68	37.63	17.73	33	31.21	20.78
	G/C		46	38.24	15.50	16	38.56	19.80
	C/C		7	42.86	9.63	2	19.50	14.85
	***P* value**	**Group**	0.627			0.354		
		All	0.591			0.703		
		Without	0.730			0.321		
rs1625895	**G/G**		72	38.06	16.91	22	26.95	20.24
	G/A		22	41.41	15.61	7	41.86	20.68
	A/A		0	-	-	1	9.00	-
	***P* value**	**Group**	0.410			0.159		
		All	0.236			0.200		
		Without	0.410			0.159		
rs1614984	**C/C**		48	41.00	17.01	34	31.76	21.58
	C/T		52	50.20	17.37	11	43.09	18.21
	T/T		14	43.00	15.56	4	32.50	13.87
	***P* value**	**Group**	0.795			0.578		
		All	0.839			0.303		
		Without	0.799			0.265		
***Mdm2* rs2279744**	**T/T**		51	40.35	16.62	25	37.36	18.94
	T/G		50	33.96	16.80	25	27.04	20.52
	G/G		18	45.11	11.45	3	54.67	11.93
	***P* value**	**Group**	**0.090**			**0.020**		
		All	0.177			**0.030**		
		Without	**0.034**			**0.024**		

SD: Standard deviation
Group: *P value* calculated with weight in the analyzed group
All: *P value* calculated with weight in the entire cohort
Without: *P value* without weight

For Peer Review

Supplementary Table 7: Haplotype reconstruction based on 5 SNPs according to *TP53* mutation status.

		WT2				MUT			
Mutation	**Haplotype[$]***	**n**	**% group**	**% All**	**% without**	**n**	**% group**	**% All**	**% without**
WT	G-A1-G-A1-C	148	47.72	47.72	47.44	48	24.61	23.54	25.81
	G-A1-G-A2-T	75	22.79	22.79	24.04	14	8.52	7.74	7.53
	C-A2-C-A1-T	40	12.81	12.81	12.82	14	7.28	6.66	7.53
	C-A1-C-A1-T	15	4.73	4.73	4.81	5	3.69	4.19	2.69
	C-A1-C-A2-T	11	3.89	3.89	3.53	2	0.53	0.74	1.08
	C-A1-C-A1-C	7	2.91	2.91	2.24	4	2.19	2.70	2.15
	C-A2-C-A2-T	3	1.38	1.38	0.96	0	-	-	-
	C-A1-G-A2-T	2	0.18	0.18	0.64	1	0.05	0.02	0.54
	G-A1-C-A1-C	2	0.61	0.61	0.64	1	1.43	2.14	0.54
	G-A2-G-A1-C	2	0.31	0.31	0.64	1	0.05	0.02	0.54
	C-A1-C-A2-C	1	0.23	0.23	0.32	5	3.69	4.19	2.69
	C-A2-G-A1-T	1	0.46	0.46	0.32	0	-	-	-
	C-A2-C-A1-T	1	0.46	0.46	0.32	1	0.18	0.07	0.54
	C-A2-C-A2-C	0	-	-	-	1	1.43	2.14	0.54
	G-A1-G-A1-T	1	0.46	0.46	0.32	1	0.05	0.02	0.54
	G-A1-C-A1-T	1	0.46	0.46	0.32	0	-	-	-
	G-A1-C-A2-T	1	0.46	0.46	0.32	0	-	-	-
	G-A2-C-A2-T	1	0.15	0.15	0.32	0	-	-	-
MUT	G-A1-G-A1-C	0	-	-	-	86	43.57	45.68	46.24
	C-A2-C-A2-C	0	-	-	-	4	2.86	1.61	2.15
	C-A1-C-A2-T	0	-	-	-	1	1.43	2.14	0.54
	G-A1-G-A1-T	0	-	-	-	1	1.43	0.27	0.54
	C-A2-G-A1-C	0	-	-	-	1	0.71	0.31	0.54

% group: Percentage calculated in each group weight
% All: Percentage calculated in the entire cohort
% without: Percentage calculated without weight
[$]: rs162785, rs17878362, rs1042522, rs17880560 and rs1614984
*: G corresponds to the Arginine and C to the Proline variants

For Peer Review

Part III.Role of G4s in intron 3 on p53 mRNA splicing and p53 protein isoform expression

This section summarizes the data presented in two annexed manuscripts. One is published in Carcinogenesis (G-quadruplex structures in TP53 intron 3: role in alternative splicing and in production of p53 mRNA isoforms), 32(3):271-8. The other is presented as a draft manuscript (Impact of G-quadruplex structures and the polymorphisms rs17878362 and rs1642785 on the expression of TP53 transcripts). Figure and Table numbers correspond to those from the publications.

Background

TP53 gene expression is complex with different transcripts that encode up to 12 protein isoforms with different N- and/or C-terminal domain (Marcel et al 2011). The mRNA from the proximal promoter generates two different N-terminal isoforms: the FSp53 mRNA, which encodes the TAp53 or the canonical p53 protein and the Δ40p53 protein lacking the first 39 aas by an internal translation and the p53I2 mRNA, which encodes the Δ40p53 protein. The internal promoter located between intron 1 and exon 5 can also produce another p53 mRNA, the p53I4, which generates the formation of two N-terminal protein isoforms: the Δ133p53 and the Δ160p53 isoforms.

Several years ago, tri-dimensional RNA structures formed by G4 were identified as new regulators of mRNA splicing. G4 structures are formed in G-rich sequences of DNA or RNA by stacks of "plates" of four guanines linked by Hoogsteen hydrogen bonds (**Manuscript 1, Figure 1**) (Gomez et al 2004). At the RNA level, G4s may play many different functions such as modifying the structures and the functions of the non-coding mRNA (Henn et al 2008), modulating the translation of the mRNA when they are located in the 5' un-translated region of this mRNA (Kumari et al 2007) or modifying the splicing and the expressions pattern of some genes like *hTERT* (Gomez et al 2004) and *Bcl-X* (Hai et al 2008).

Intron 3 of the *TP53* gene contains G-tracts organised in a pattern similar to the G4 structure. Given the very small size of successive introns and exons in that area of the gene, the G-rich tract is located at around 100 bp from the acceptor-splicing site of intron 2, within the canonical distance for the binding of splicing regulatory machinery. Given that alternative splicing of intron 2 leads to a transcript that encodes the Δ40p53 isoforms, we hypothesised that the formation of G4 in intron 3 may modulate the alternative splicing of intron 2 and the formation of the p53I2 mRNA.

Results

The use of the prediction softwares Quadfinder and QGRS Mapper identified that the G-rich tract of intron 3 could form several stable G4 structure. First, we demonstrated the physical existence of these structures in intron 3 using Reverse Transcriptase elongation assay (**Manuscript 1, Figure 3**) and by monitoring the biophysical characteristics of mRNA oligonucleotides encompassing the G-rich region using UV melting and UV thermal difference spectra (**Manuscript 1, Figure 2**). These biochemical and biophysical studies provided direct evidence for G4 structures forming in intron 3 of p53 pre-mRNA, involving in particular two repeats of six guanines.

Secondly, we studied the impact of this structure in the alternative splicing of intron 2 by two different approaches: a biochemical method based on the use of a reporter assay mimicking the alternative splicing of intron 2 in which we have mutated the guanines implicated in the G4 structures to disrupt their formation; and a biological approach involving the treatment of lymphoblastoid cell lines with 360A, a drug which modulates G4 structures formed in RNA (**Manuscript 1, Figure 5A**). Results have shown that G4 structures in intron 3 are implicated in the regulation of the exclusion of intron 2 (**Manuscript 1, Figures 4, 5B and Figure 21 of this Thesis**). The stabilisation of the G4 structures in the intron 3 using A360 increased the exclusion of intron 2 and the formation of the FSp53 mRNA.

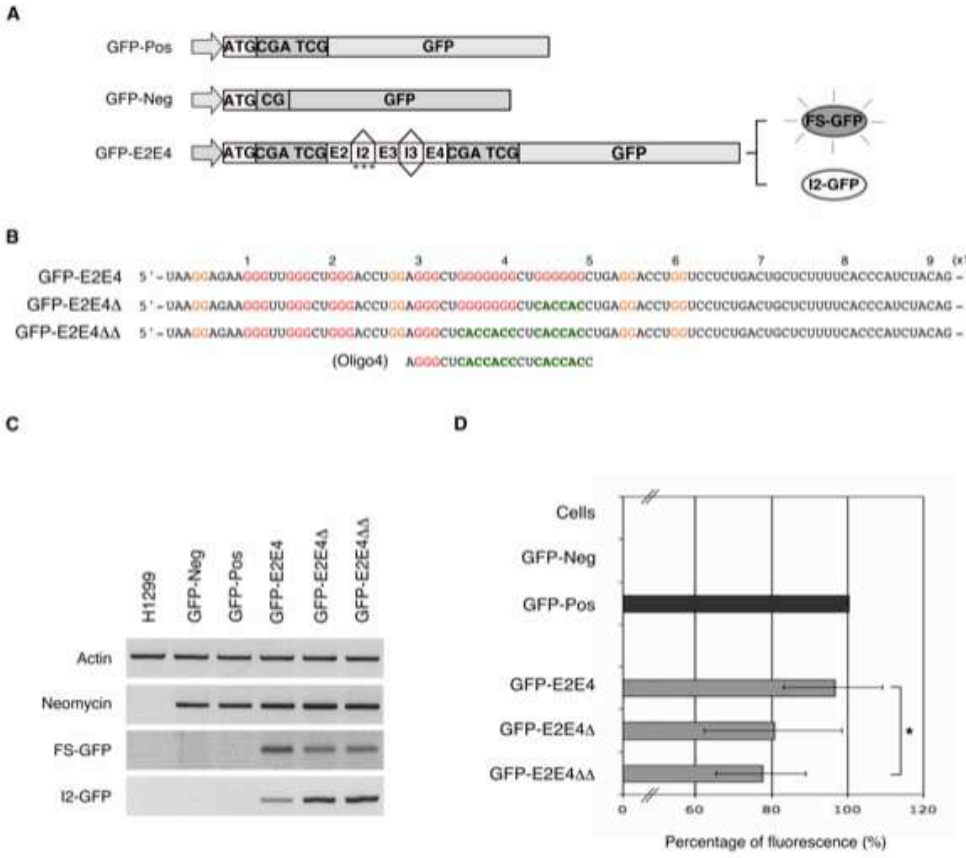


Figure 21: GFP-reporter splicing assay to analyze the role of G4s in alternative splicing of intron 2. (**A**) Schematic representation of constructs used in the GFP-based reporter system. A fragment from exon 2 to exon 4, excluding both ATG 1 and 40, was cloned into the GFP-Pos plasmid (serving as positive control) to construct the GFP-E2E4 reporter. With this reporter, GFP fluorescence is generated only after correct excision of intron 2, since the latter contains stop codons that preclude protein synthesis from ATG1 (stars). Arrow: CMV promoter and Kozak sequence. (**B**) Mutations disrupting *G4s* in GFP-E2E4 reporter. Guanines involved in *G4s* were mutated (green letters). The synthetic RNA Oligo4 was designed to verify by spectroscopic methods that mutation of these guanines prevented the formation of *G4s*. (**C**) mRNA expression analysis of GFP-E2E4, Δ and ΔΔ mutants. Constructs were transfected in H1299 cells and their expression was analyzed by semi-quantitative RT-PCR and agarose gel electrophoresis. As compared to GFP-E2E4, a decrease in fully spliced mRNA (FS-GFP) and an increase of intron 2 retaining mRNA (I2-GFP) were observed. Actin: loading control; neomycin: control of transfection efficiency. (**D**) Analysis of GFP fluorescence of transfected H1299 by flow cytometry. Compared to GFP-Pos GFP-E2E4. *: $p < 0.05$; t-student test based on duplicate of five independent experiments.

In a continuation of this study (**Manuscript 2**), we have determined whether G4s could form in the context of intron 3 sequences carrying the rs17878362 duplication (A2 allele). This was demonstrated using using UV melting, Reverse transcription elongation assay and UV thermal difference spectra (**Manuscript 2, Figures 1 and 2**). Studies comparing the levels of p53 mRNA expression between lymphoblastoid cell lines homozygous for the A1 or A2 alleles suggested that there was an allele-specific difference in the level of expression of p53I2, which was statistically lower the two rs17878362 A2 than in the two rs17878362 A1 homozygous lymphoblastoid cell lines (**Manuscript 2, Figure 3**). After treatment with the RNA binding ligand 360A (**Figure 22 of this Thesis**) or IR, the

FSp53 mRNA levels were increased and p53I2 levels decreased in the A1A1 cell lines, suggesting that under these conditions full splicing is favoured Such a trend was not observed in the A2A2 cells. In addition, carriage of the intron 2 rs164278 (G > C) C allele, which is in strong linkage disequilibrium with the A2 allele, was associated with a lower p53I2 transcript stability. This reduced stability may contribute to the differential expression of p53I2 mRNA levels observed in the A1 and A2 polymorphic sequence context.

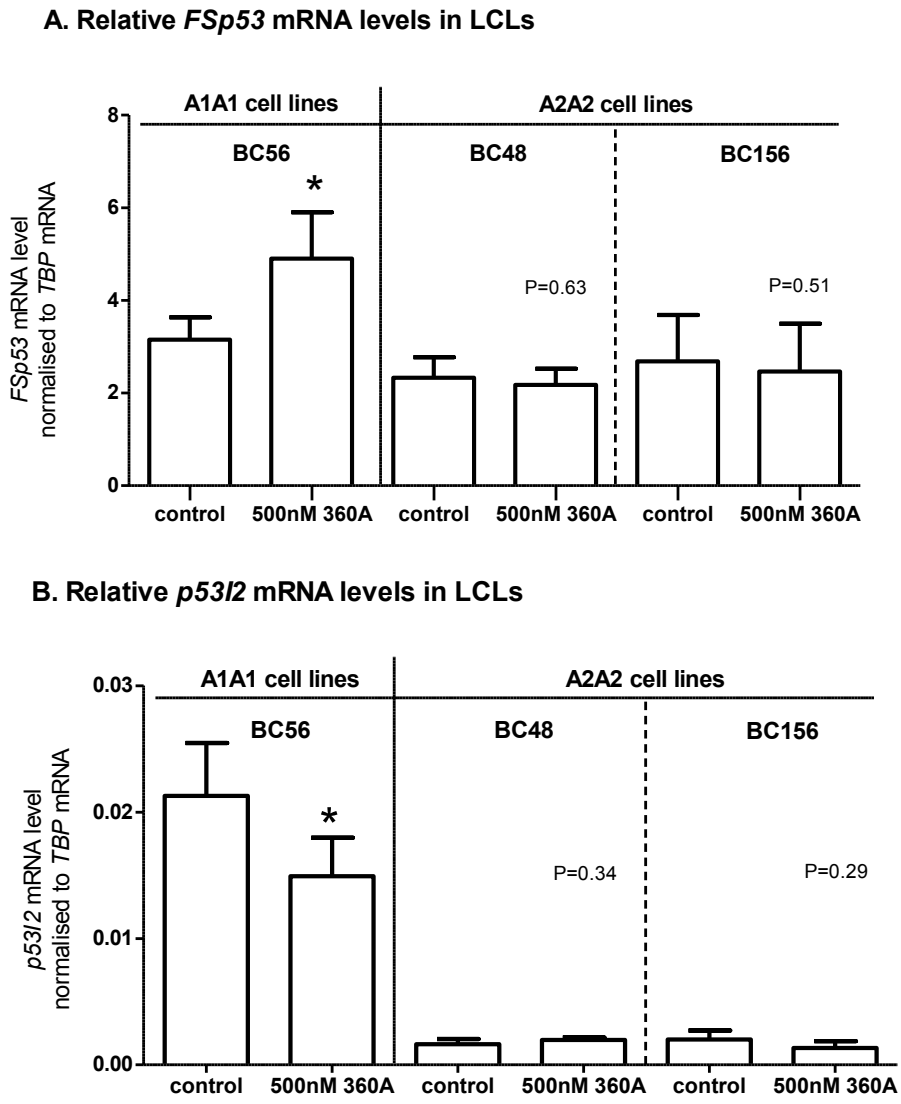


Figure 22: Effect of the G4 ligand 360A on expression of p53 transcript for A1A1 and A2A2 for rs17878362 lymphoblastoïd cell lines. Quantitative reverse transcription-PCR analysis of mRNA isolated from lymphoblastoïd cells using Sybr-Green. Cells were treated for 48h with 500 nM of 360A and levels of (**A**) FSp53 and (**B**) p53I2 were measured. * $P < 0.05$; Student's *t*-test based on triplicate of at least five independent experiments.

Conclusion

G4 structures are present in intron 3 of the *TP53* pre-mRNA and modulate the regulation of the alternative splicing of intron 2 by increasing the exclusion of intron 2 and the expression of FSp53 mRNA. These results are supported by results comparing levels of p53 mRNA in lymphoblastoid cells with A1A1 or A2A2 genotypes. It remains to be demonstrated how the alteration of the balance between FSp53 and p53I2 mRNAs may impact on p53 functions. One possibility is that changes in p53 splicing may modulate the relative levels of expression of TAp53 and of the Δ40p53 protein isoforms.

Carcinogenesis vol.32 no.3 pp.271–278, 2011
doi:10.1093/carcin/bgq253
Advance Access publication November 26, 2010

G-quadruplex structures in *TP53* intron 3: role in alternative splicing and in production of p53 mRNA isoforms

Virginie Marcel[1,7], Phong L.T.Tran[2], Charlotte Sagne[1,3], Ghyslaine Martel-Planche[1], Laurence Vaslin[3], Marie-Paule Teulade-Fichou[4], Janet Hall[3], Jean-Louis Mergny[2,5], Pierre Hainaut[1,*] and Eric Van Dyck[1,6]

[1]Group of Molecular Carcinogenesis, International Agency for Research on Cancer, 150 cours Albert Thomas, 69372, Lyon Cedex 08, France, [2]Muséum National d'Histoire Naturelle, INSERM U565, Centre National de la Recherche Scientifique UMR7196, 75231, Paris Cedex 05, France, [3]INSERM U612, Institut Curie-Recherche, 91405 Orsay, France, [4]Centre National de la Recherche Scientifique UMR176, Institut Curie-Recherche, 91405 Orsay, France, [5]INSERM U869, Institut Européen de Chimie Biologie, Université de Bordeaux, 33607 Pessac, France and [6]Laboratory of Experimental Hemato-Oncology, Public Center for Health (CRP-Santé), L-1526, Luxembourg, UK

[7]Present address: Department of Surgery and Molecular Oncology, INSERM-European Associated Laboratory, Ninewells Hospital and Medical School, Dundee, DD1 9SY, Scotland, UK

*To whom correspondence should be addressed. Tel: +33 4 72 73 84 62; Fax: +33 4 72 73 83 22;
Email: hainaut@iarc.fr

The tumor suppressor gene *TP53*, encoding p53, is expressed as several transcripts. The fully spliced p53 (FSp53) transcript encodes the canonical p53 protein. The alternatively spliced p53I2 transcript retains intron 2 and encodes Δ40p53 (or ΔNp53), an isoform lacking first 39 N-terminal residues corresponding to the main transactivation domain. We demonstrate the formation of *G-quadruplex* structures (G4) in a GC-rich region of intron 3 that modulates the splicing of intron 2. First, we show the formation of G4 in synthetic RNAs encompassing intron 3 sequences by ultraviolet melting, thermal difference spectra and circular dichroism spectroscopy. These observations are confirmed by detection of G4-induced reverse transcriptase elongation stops in synthetic RNA of intron 3. In this region, p53 pre-messenger RNA (mRNA) contains a succession of short exons (exons 2 and 3) and introns (introns 2 and 4) covering a total of 333 bp. Site-directed mutagenesis of G-tracts putatively involved in G4 formation decreased by ∼30% the excision of intron 2 in a green fluorescent protein-reporter splicing assay. Moreover, treatment of lymphoblastoid cells with 360A, a synthetic ligand that binds to single-strand G4 structures, increases the formation of FSp53 mRNA and decreases p53I2 mRNA expression. These results indicate that G4 structures in intron 3 regulate the splicing of intron 2, leading to differential expression of transcripts encoding distinct p53 isoforms.

Introduction

The tumor suppressor p53 protein controls antiproliferative responses to various forms of stress (1). Its function is impaired in >50% of human cancer, mainly by mutation (2). *TP53* gene expression is complex, with different transcripts encoding isoforms carrying distinct N- and C-termini (3,4). To date, 10 isoforms have been identified resulting from the usage of alternative promoters, splice sites and/or translational initiation sites (5). Several of these isoforms differ in their N-terminal region. The N-terminus of p53 contains the main transactivation domain (residues 1–42, transactivation domain I) as well as the binding site of Hdm2, which targets p53 for proteasome degradation and regulates p53 stability (1). Transcription of p53 messenger RNA (mRNA) from the proximal promoter generates two proteins with distinct N-terminal domain. The first corresponds to the canonical p53 protein, assembled from the fully spliced p53 (FSp53) mRNA that retains 11 exons. This protein induces p53-mediated growth suppression in response to stress. The second isoform, Δ40p53 (or ΔNp53), is assembled from an alternatively spliced mRNA retaining intron 2 (p53I2) and lacks the first 39 residues, corresponding to transactivation domain I, as well as Hdm2-binding site (3,6). The use of an internal promoter located in a region between intron 1 and exon 5 generates a third N-terminal isoform, Δ133p53, which lacks the first 132 residues (5).

When expressed in excess to p53, Δ40p53 inhibits transcriptional activity and interferes in the control of cell cycle progression and apoptosis by exerting a negative effect on the expression of p53-target genes (3,6,7). However, the biological circumstances and the molecular mechanisms regulating Δ40p53 expression are still poorly known. Retention of intron 2 in p53I2 mRNA introduces several stop codons in the reading frame of AUG 1, thus preventing the synthesis of a full-length p53 protein. However, p53I2 mRNA can be translated using AUG 40 as initiation site, generating a protein isoform which differs from the canonical p53 by the lack of the first 39 residues. Expression of p53I2 transcript has been reported in cell lines, such as MCF-7, in normal lymphocytes and in primary melanoma isolates (7,8). However, the mechanism that regulates the splicing of p53 pre-mRNA into FSp53 or p53I2 is not understood. Δ40p53 protein isoform can also be produced by internal ribosomal entry site-regulated internal initiation of translation using FSp53 mRNA (9,10).

In recent years, it has been proposed that tridimensional RNA structures such as *G-quadruplexes* may play important roles in regulating splicing (11). These structures result of the propensity of G-rich sequences to fold into four-stranded cation-dependent structures (12). They are formed by the interaction of four guanines organized in a cyclic Hoogsteen hydrogen bonding arrangement termed a G-quartet and by the stacking of several G-quartets (Figure 1A). It is estimated that over 376 000 sequences in the human genome have the potential to adopt *G-quadruplex* structures, most of them located in non-coding regions (13,14). At the RNA level, *G-quadruplexes* may play a number of roles. In non-coding RNAs, they can affect their structures and functions (15,16). In 5′-untranslated region of mRNAs, *G-quadruplexes* have been shown to modulate translation (17,18). When present in introns, *G-quadruplexes* can affect the splicing and expression patterns of genes such as *hTERT* (human telomerase reverse transcriptase), *Bcl-xL* or *FMRP* (Fragile X mental retardation protein) (11,19,20).

The sequence of intron 3 in *TP53* contains tracts of G bases organized in a pattern similar to the one of regions forming *G-quadruplex* structures. Since exon 3 in *TP53* is extremely short (22 bp), we reasoned that motifs located in intron 3 might have an effect on the regulation of the splicing of intron 2. In this study, we provide an evidence for the formation of *G-quadruplex* in *TP53* intron 3 and that these *G-quadruplex* structures may affect the splicing of intron 2, modulating the synthesis of either FSp53 (intron 2 spliced out) or p53I2 (intron 2 retained) mRNAs, which encode different p53 protein isoforms.

Materials and methods

Synthetic RNA oligomers and compounds

RNA oligomers derived from intron 3 of p53 RNA and containing several G-tracts (Figure 1B and C, Table I) were synthesized from IBA (Göttingen, Germany) at 0.2 μmol scale and resuspended in 20–40 μl of bidistilled water (ddH_2O). Concentrations of oligomers were determined by ultraviolet (UV) absorbance. The compound 360A [(2,6-N,N′-methyl-quinolinio-3-yl)-pyridine dicarboxamide] was stored in dimethyl sulfoxide and further diluted in ddH_2O for treating lymphoblastoid cell line for 48 h (21).

Abbreviations: CD, circular dichroism; FSp53, fully spliced p53; GFP, green fluorescent protein; mRNA, messenger RNA; PCR, polymerase chain reaction; RT, reverse transcriptase; TDS, thermal difference spectra; UV, ultraviolet.

© The Author 2010. Published by Oxford University Press. All rights reserved. For Permissions, please email: journals.permissions@oup.com

Downloaded from http://carcin.oxfordjournals.org/ at UniversitÃ© de Limoges. Service commun de la documentation on August 28, 2013

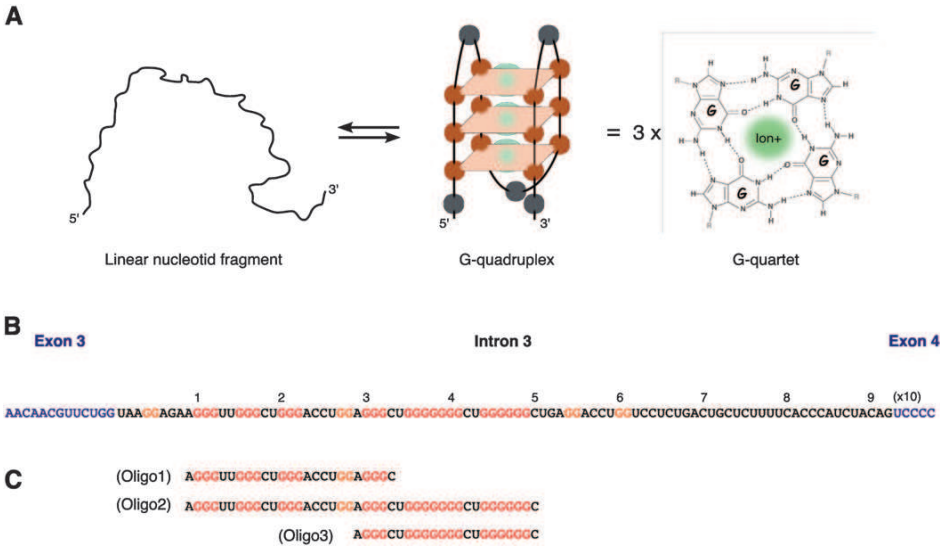


Fig. 1. Schematic representation of *G-quadruplex* and G-rich sequences in intron 3 of *TP53*. (**A**) Formation of *G-quadruplex* structure based on G-quartet element. Middle panel: tridimensional representation of a *G-quadruplex*; right panel: bidimensional representation of a G-quartet, stabilized by cations. (**B**) Partial sequence of *TP53* pre-mRNA from end of exon 3 to start of exon 4. Exonic sequences (blue), intronic sequences (black) and G-repeats (red and orange, three and two guanines, respectively) are shown. (**C**) Location and sequence of synthetic RNA oligomers 1, 2 and 3.

Table I. List of the oligomers used in this study

Oligomer name	Sequence (5′–3′)
Oligo1	AGGGUUGGGCUGGGACCUGGAGGGC
Oligo2	AGGGUUGGGCUGGGACCUGGAGGG-CUGGGGGGGCUGGGGGGC
Oligo3	AGGGCUGGGGGGGCUGGGGGGC
Oligo4	AGGGCUCACCACCCUCACCACC
RF16[a]	AAAAAAAAUUUUUUUU
RK50[a]	CGUAACGUUACG
G4-RNA 22 nts	AGGGUUAGGGUUAGGGUUAGGG
G4-RNA 46 nts	AGGGUUAGGGUUAGGGUUAGGGUUAGG-GUUAGGGUUAGGGUUAGGGG

[a]Self complementary RNA duplexes.

UV melting

The thermal stability of oligonucleotides was characterized in heating/cooling experiments by recording UV absorbance at 295 nm as a function of temperature using an Uvikon XL UV/Vis spectrophotometer (22). UV-melting experiments were conducted as described previously in 10 mM lithium cacodylate buffer, pH 7.2 containing either 100 mM LiCl, 100 mM NaCl or KCl at concentrations of 2 or 10 mM (23). Oligomers were tested at 3 µM strand concentration. The heating and cooling rates were 0.2°C/min. Experiments were performed in 1 cm path length quartz cuvettes.

UV thermal difference spectra

Thermal difference spectra (TDS) were obtained by difference between the absorbance spectra from unfolded (at high temperature, >90°C) and folded oligonucleotides (at low temperature, ~3°C). TDS provide specific signatures demonstrating *G-quadruplex* formation (24). Spectra were recorded between 220 and 330 nm at pH 7.0 in 100 mM NaCl in an Uvikon XL UV/Vis spectrophotometer using quartz cuvettes with an optical path length of 1 cm. RNA strand concentration was 3 M.

Circular dichroism spectroscopy

Circular dichroism (CD) spectra were recorded between 200 and 330 nm on a JASCO-810 spectropolarimeter as described previously (25). Spectra were recorded after heating/cooling experiment and at low temperature (>4°C). Oligomers were tested at 3 M strand concentration.

Fluorescence resonance energy transfer melting assay

The F21T oligonucleotide, mimicking human telomeric motifs and double-labeled with FAM and TAMRA at either end, was synthesized by Eurogentec (Seraing, Belgium) (Table I). Briefly, in the presence of 360A, this oligonucleotide folds into a stable *G-quadruplex*, thus allowing fluorescence transfer to occur. The melting of the F21T probe was recorded by measuring fluorescence emission as a function of temperature. Different unlabeled oligomers were used as competitors, including the *G-quadruplex*-forming Oligos 1, 2 and 3, the non-*G-quadruplex* forming Oligo4 and two negative control oligonucleotides which do not contain G-tracts (Table I) (26,27).

Reverse transcriptase elongation assay

Intron 3 RNA was synthesized by *in vitro* transcription with T7 RNA polymerase (Promega, Madison, WI) of a linearized pBSK plasmids containing intron 3. One microgram of pBSK, denatured in 0.2 N NaOH and 0.2 mM ethylenediaminetetraacetic acid (37°C 30 min), was sequenced using Sequenase T7 (Amersham, Pittsburgh, PA). Primer extensions were performed on 8 µmol of intron 3 RNA and 10^4 c.p.m. of ^{32}P-labeled BSK-R primer. Denaturation (95°C 1 min) and annealing (24°C 10 min) were completed in specific buffer (100 mM NaCl or KCl, 25 mM Tris–HCl pH 8.3, 10 mM $MgCl_2$ and 0.2 mM deoxynucleoside triphosphates) before extension (37°C, 30 min) using 30 U of AMV reverse transcriptase (RT) (Promega). Products were analyzed on 10% polyacrylamide–8 M urea gels following by autoradiography.

Green fluorescent protein-based reporter system

A PvuI restriction site was introduced into the pEGFP-N3 plasmid immediately after the first ATG of green fluorescent protein (GFP) to produce the GFP-Pos plasmid. A negative control plasmid was produced by Klenow fragment digestion at PvuI site. A fragment from exon 2 to exon 4 of *TP53* gene, excluding ATG 1 and 40, was cloned into the PvuI restriction site of GFP-Pos plasmid. Mutant GFP-Δ and GFP-ΔΔ plasmids were produced by QuickChange® site-directed Mutagenesis Kit (Stratagene, Santa Clara, CA) (supplementary Table I available at *Carcinogenesis* Online). In six-well plates, 5×10^5 of H1299 p53-null cells were transfected with 0.2 µg/ml of these plasmids using Fugene (Roche, Basel, Switzerland), RNA expression was analyzed by reverse transcription–polymerase chain reaction (PCR) (supplementary Table IIis available at *Carcinogenesis* Online), whereas GFP protein fluorescence was analyzed by flow cytometry.

Real-time PCR analysis

A lymphoblastoid cell line (BC9) expressing wild-type p53 and derived from breast cancer patient was used (28). Total RNA was extracted using Trizol Reagent (Invitrogen™, Carlsbad, CA). The level of each N-terminal variant p53 mRNA was analyzed using QuantiTect™ Syber® Green PCR Kit (Qiagen, Hamburg, Germany) and specific primers (supplementary Table I available at *Carcinogenesis* Online and Figure 5B) and normalized to that of glycerAldehyde 3-phosphate dehydrogenase. The ΔCt was calculated as an average of three measurements by plate for two quantitative reverse transcription–PCR of at least two independent experiments.

Downloaded from http://carcin.oxfordjournals.org/ at UniversitÃ© de Limoges. Service commun de la documentation on August 28, 2013

Statistical analyses

Analysis of variance and Student's *t*-tests were performed using online tools (http://faculty.vassar.edu/lowry/VassarStats.html).

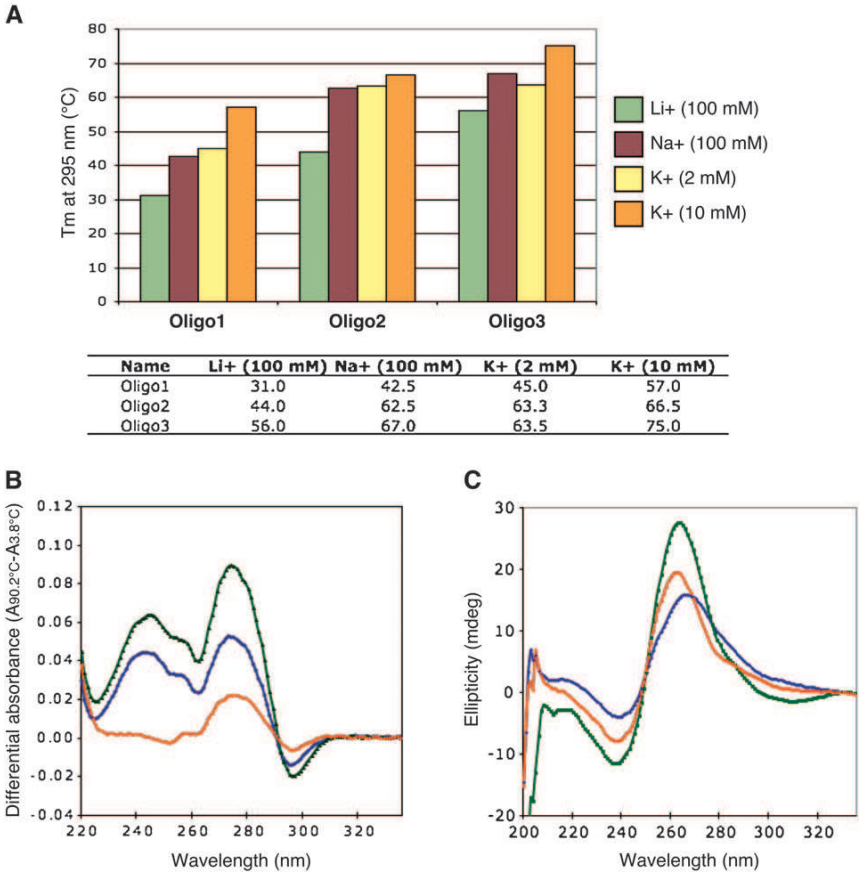
Results

G-quadruplex structures in intron 3 of p53 mRNA

TP53 intron 3 (93 bp) is rich in G bases (41%). To analyze whether this *TP53* region may contain motifs capable of forming *G-quadruplex*, the prediction software Quadfinder server and the web-based server QGRS Mapper were used (29,30). Using default parameters to predict *G-quadruplex* motifs over the whole gene sequence, four candidate regions were identified: two regions in intron 1, one in intron 3 and another in intron 6 (data for intron 3 in supplementary Figure 1A and Bis available at *Carcinogenesis* Online; data for other introns not shown). The predicted *G-quadruplex*-forming motif in intron 3 was located 60 bp downstream of the splice acceptor site of intron 2 (supplementary Figure 1A and Bis available at *Carcinogenesis* Online). This motif is predicted to carry up to 5 G-tracts sequences capable of forming *G-quadruplex* structures (supplementary Figure 1A and Bis available at *Carcinogenesis* Online).

To demonstrate the formation of *G-quadruplex* structures in *TP53* intron 3, we synthesized three RNA oligomers derived from the G-rich region of intron 3 (Figure 1C, Table I) and we analyzed their thermal stability under different ionic conditions by UV melting (Figure 2A). For each of the three oligomers, the highest melting temperatures (T_m) were observed in K^+ (10 mM) buffer (Oligo1: 57°C; Oligo2: 66.5°C and Oligo3: 75°C). T_m was dependent on the nature of the cation, with T_m (Li^+ 100 mM) < T_m (Na^+ 100 mM) < T_m (K^+ 10 mM). This dependence is typical of the effect of these cations on the stability of *G-quadruplexes*. The T_m values of Oligo2 and 3 in K^+ 10 mM (66.5°C and 75°C, respectively) were higher than for Oligo1 (57°C), suggesting that *G-quadruplexes* in Oligo2 and 3 were more stable than in Oligo1. Oligo2 and 3 differ from Oligo1 by the presence of two tracts of six guanines in their 3′ domain (Figure 1C), suggesting that these G-repeats are important for stable G-quadruplex formation.

To further demonstrate the presence of G4, we used two different biophysical methods. First, we analyzed the profiles of the three oligomers using UV TDS (Figure 2B) and CD spectroscopy (Figure 2C). A TDS is obtained by recording the UV absorbance spectra of the unfolded and folded states at temperatures above and below the melting temperature (T_m). The TDS profiles of the three oligomers showed a positive peak at 273 nm and a negative peak at 295 nm (Figure 2B). These two peaks are typical 'signatures' of *G-quadruplex* structures as determined by the analysis of different types of *G-quadruplexes* (24). The variations in TDS profiles are independent of oligomer concentrations but reflect differences in sequence and in stability of G4. In particular, the low amplitude of the peaks observed with Oligo3 may be due to the fact that in this Oligo3, *G-quadruplex* structures are stable and only undergo partial unfolding within the temperature interval. CD spectroscopy also displayed changes in ellipticity, a characteristic of *G-quadruplex* structures, including a negative peak at 240 nm and a positive peak at 260 nm (Figure 2C). Overall, these results



Name	Li+ (100 mM)	Na+ (100 mM)	K+ (2 mM)	K+ (10 mM)
Oligo1	31.0	42.5	45.0	57.0
Oligo2	44.0	62.5	63.3	66.5
Oligo3	56.0	67.0	63.5	75.0

Fig. 2. Detection of *G-quadruplex* structures in synthetic RNA by spectroscopic methods. (**A**) UV melting experiments. Histograms (upper panel) and values (lower panel) of melting temperatures (T_m) ± 1°C at 295 nm for Oligo1–3 in different ionic conditions. (**B**) TDS of Oligos 1–3. Blue: Oligo1; green: Oligo2; Red: Oligo3. Presence of a positive peak at 273 nm and a negative peak at 295 nm are characteristic signatures of *G-quadruplexes* (24). (**C**) CD spectra of Oligo1–3 at 4°C in 10 mM KCl. Blue: Oligo1; green: Oligo2; red: Oligo3.

Downloaded from http://carcin.oxfordjournals.org/ at UniversitÃ© de Limoges. Service commun de la documentation on August 28, 2013

using synthetic RNA confirm *in silico* predictions and provide evidence that *G-quadruplex* structures can be formed in intron 3 of p53 RNA, involving in particular two G-repeats of six guanines.

Detection of G-quadruplexes by RT elongation assay

Next, we used the RT elongation assay to further identify the position of guanines involved in *G-quadruplexes* (Figure 3). This method takes advantage of the differential cation-dependent stability of *G-quadruplexes* to identify structures that interfere with RT elongation (31). K^+ stabilizes *G-quadruplexes*, leading to a K^+-dependent block of complementary DNA synthesis at positions where guanines are involved in a G-quartet. In contrast, in the presence of Na^+, *G-quadruplex* stability is reduced, allowing complementary DNA synthesis to proceed without elongation block. A full-length RT product corresponding to the entire intron 3 RNA was detected in presence of Na^+ (lane 5, '5′-intron 3' arrow in Figure 3). However, in the presence of K^+, stops of RT elongation were observed (lane 6) at positions corresponding to guanines on sequencing analysis, matching those predicted to form the 3′ end of the *G-quadruplexes*. Thus, the detection of several RT-stops is compatible with *in silico* prediction and with spectrometric analyses indicating a critical role of two tracts of six guanines in the formation of several *quadruplex* conformers in intron 3.

Role of G-quadruplex in excision of intron 2 using an in vitro splicing assay

Because alternative splicing of *TP53* involves either retention (p53I2) or skipping (FSp53) of intron 2, we examined whether *G-quadruplex*

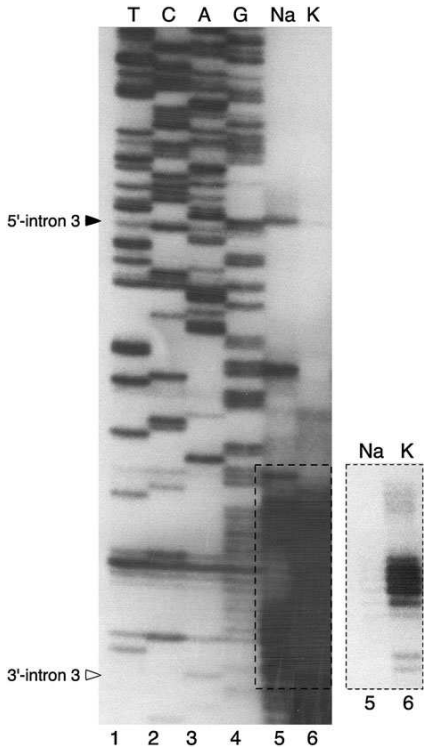


Fig. 3. RT elongation assay identifying RT-stops at guanines involved in *G-quadruplexes*. *G-quadruplexes* were mapped by reverse transcription of intron 3 sequences with a radiolabeled oligonucleotide in the presence of either Na^+ (destabilizing G4) or K^+ (stabilizing G4) and products were separated by denaturing gel electrophoresis (lanes 5–6). Lanes 1–4: dideoxy sequencing of the sequence used for reverse transcription. RT-stops (lanes 5–6) occurred at guanines residues as shown by comparison with the sequencing lanes. Black arrowhead: 5′-intron 3 end; open arrowhead: 3′-intron 3 end; Box on the right: short exposure of the area delimited by dotted lines, identifying the presence of strong stops at defined guanines in lane 6.

in intron 3 were involved in the regulation of this splicing. To analyze the direct effect of *G-quadruplexes* on the splicing of intron 2, we constructed a minigene reporter based on GFP, in which a *TP53* fragment from exons 2 to 4 was cloned between the ATG and the GFP coding sequence (GFP-E2E4, Figure 4A). With this construct, correct splicing of intron 2 (FS–GFP) produces an mRNA that supports GFP synthesis, whereas retention of intron 2 (I2-GFP) introduces stop codons preventing GFP expression. Two mutant versions of the reporter were constructed, in which either one (GFP-E2E4Δ) or both (GFP-E2E4ΔΔ) two tracts of six guanines predicted to be involved in *G-quadruplexes* were substituted by site-directed mutagenesis (Figure 4B). To verify that the mutations introduced in GFP-E2E4ΔΔ abolished *G-quadruplex* formation, we synthesized an RNA oligomer, Oligo4, containing the same mutations (Figure 4B and Table I). When compared with Oligo2 and 3, Oligo4 did not show the typical patterns of *G-quadruplexes* detected by UV-melting curves, TDS and CD spectra (supplementary Figure 1A and Bis available at *Carcinogenesis* Online).

The p53-null cell line H1299 were transfected with the various GFP constructs. All constructs displayed equal transfection efficiency, as determined by reverse transcription–PCR analysis of expression of the *neomycin* gene carried by the plasmids (Figure 4C). When compared with cells expressing the wild-type GFP-E2E4 construct, cells transfected with GFP-E2E4Δ and -E2E4ΔΔ showed an increase in I2-GFP mRNA expression (retaining intron 2) and a decrease of FS–GFP mRNA expression (correct splice-out of intron 2). GFP fluorescence measurements showed that the wild-type GFP-E2E4 construct generated as much fluorescence as the positive GFP control (GFP-pos) (Figure 4). Mutation of guanines resulted in a progressive decrease in fluorescence that depended upon the number of guanines mutated. A significant reduction of fluorescence of ∼25% was observed when comparing GFP-E2E4ΔΔ with GFP-E2E4 ($P < 0.05$). Taken together, these observations support the hypothesis that *G-quadruplexes* in intron 3 modulate the rate at which intron 2 is spliced-out. With wild-type intron 3 sequences presumably capable of forming stable *G-quadruplexes*, splicing of intron 2 appears to be more efficient than when *G-quadruplex* formation is crippled by mutations.

Effects of G-quadruplexes on p53 splicing in lymphoblastoid cells

To determine whether *G-quadruplexes* may modulate the splicing of p53 in intact cells, we analyzed the effect of a specific ligand of *G-quadruplex* structures, 360A, on the expression of FSp53 and p53I2 transcripts in lymphoblastoid cells using real-time PCR (Figure 5). The ligand 360A has high affinity and stabilizing effects for *G-quadruplex* structures in single-stranded nucleic acids (21). To demonstrate that 360A could bind to *G-quadruplexes* in intron 3 of p53 RNA, we used an fluorescence resonance energy transfer melting assay (Figure 4A) (27). A DNA oligonucleotide mimicking human telomeric motifs (F21T) was labeled with FAM at the 5′ end and TAMRA at the 3′ end (26,27). Addition of 360A induced a concentration-dependent stabilization of the *G-quadruplex*, resulting in an increased melting temperature (ΔT_m), at which fluorescence emission ceases due to unfolding of the *G-quadruplex* that suppresses fluorescence transfer. Excess amounts of unlabeled DNA or RNA were used as competitors to determine their capacity to displace 360A from its fluorescent target, thereby reducing its melting temperature. Figure 5A shows that Oligo1, 2 and 3 were capable of acting as competitors, in contrast with mutant Oligo4 and with negative control sequences RF16 and RK50. These results indicate that RNA *quadruplexes* formed in *TP53* intron 3 could bind 360A, further extending the sequence repertoire of G4 structures bound by this ligand.

We next treated human lymphoblastoid cell line with 50 nM 360A for 48 h and analyzed its effects on p53 mRNA expression (Figure 5B and C). Although 360A did not exert significant toxicity at the concentrations used (supplementary Figure 3 available at *Carcinogenesis* Online), it induced opposite, dose-dependent effects on levels of the two p53 transcripts, FSp53 and p53I2 (Figure 5C). Treatment by 360A resulted in a significant 3-fold increase in FSp53 mRNA.

Downloaded from http://carcin.oxfordjournals.org/ at UniversitÃ© de Limoges. Service commun de la documentation on August 28, 2013

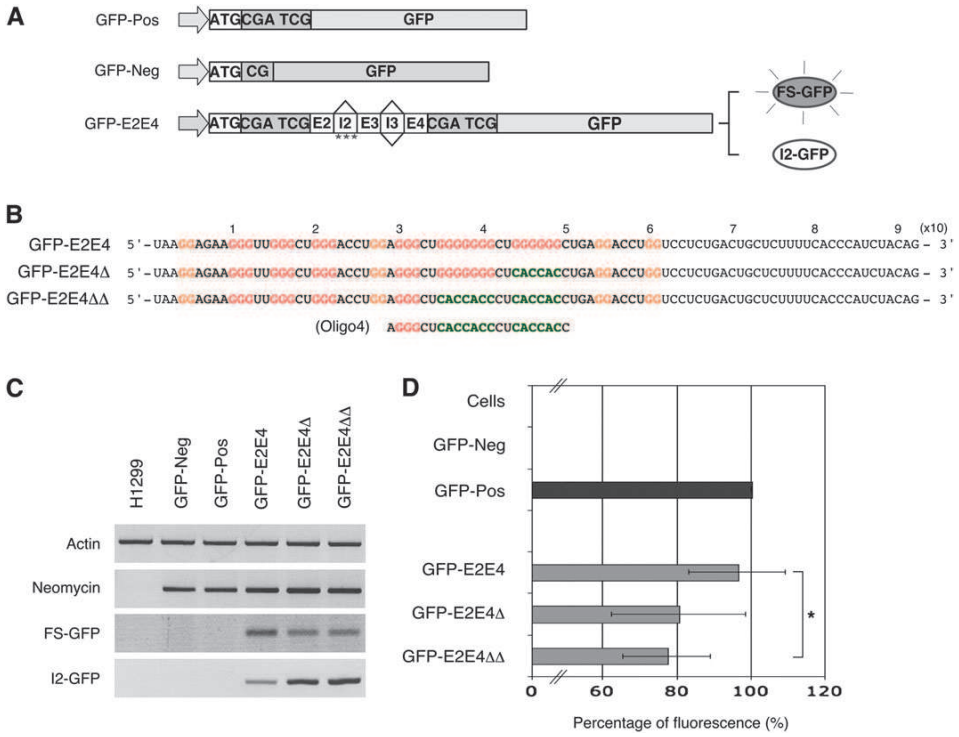


Fig. 4. GFP reporter splicing assay to analyze the role of *G-quadruplexes* in alternative splicing of intron 2. (**A**) Schematic representation of constructs used in the GFP-based reporter system. A fragment from exon 2 to exon 4, excluding both ATG 1 and 40, was cloned into the GFP-Pos plasmid (serving as positive control) to construct the GFP-E2E4 reporter. With this reporter, GFP fluorescence is generated only after correct excision of intron 2 since the latter contains stop codons that preclude protein synthesis from ATG1 (asterisks). Arrow: CMV promoter and Kozak sequence. (**B**) Mutations disrupting *G-quadruplexes* in GFP-E2E4 reporter. Guanines involved in *G-quadruplexes* were mutated (green letters). The synthetic RNA Oligo4 was designed to verify by spectroscopic methods that mutation of these guanines prevented the formation of *G-quadruplexes* (see supplementary Figure 1, available at *Carcinogenesis* Online). (**C**) mRNA expression analysis of GFP-E2E4, Δ and ΔΔ mutants. Constructs were transfected in H1299 cells and their expression was analyzed by semiquantitative reverse transcription–PCR and agarose gel electrophoresis. As compared with GFP-E2E4, a decrease in fully spliced mRNA (FS–GFP) and an increase of intron 2 retaining mRNA (I2-GFP) were observed. Actin: loading control; neomycin: control of transfection efficiency. (**D**) Analysis of GFP fluorescence of transfected H1299 by flow cytometry. Compared with GFP-Pos GFP-E2E4, Δ and ΔΔ mutants exhibited a decrease of GFP fluorescence, indicating an increase of intron 2 retention. $^{*}P < 0.05$; Student's *t*-test based on duplicate of five independent experiments.

In contrast, p53I2 levels were significantly decreased by ~40%. These results support the notion that *G-quadruplexes* are involved in the regulation of intron 2 splicing in intact cells.

Discussion

TP53 is expressed as several isoforms produced by different mechanisms, including alternative splicing (4,7). Differential splicing of intron 2 leads to the generation of FSp53 and p53I2 mRNAs that respectively encode the tumor suppressor p53 and the Δ40p53 isoform, which lacks the transactivation domain and acts as a negative regulator of p53 activity (7,32). Therefore, mechanisms that control the alternative splicing of intron 2 may have an effect on p53 activity by modifying the relative expression levels of p53 and of its Δ40p53 isoform. In this study, we show that a region of intron 3 forming *G-quadruplex* structures in p53 pre-mRNA is involved in the alternative splicing of *TP53* intron 2. First, we used synthetic RNAs containing sequences of intron 3 and we show that these RNA can form *G-quadruplexes* using different spectrometric methods (UV melting, TDS and CD spectroscopy). Next, the presence of *G-quadruplexes* is further demonstrated by RT elongation assays. In these assays, demonstration of a *G-quadruplex* relies on the fact that these structures are very dependent upon the presence of specific cations, being stabilized in the presence of K^+ and destabilized in the presence of Na^+. Furthermore, to demonstrate the involvement of these *G-quadruplex* structures in the splicing of intron 2, we have constructed a GFP reporter system in which fluorescence emission is dependent upon the correct excision of intron 2 sequences. Using this system, we show that mutation of guanines involved in *G-quadruplexes* significantly decreases the splicing-out of intron 2, however, without abolishing it. Finally, to demonstrate that *G-quadruplexes* in intron 3 can impact on p53 splicing in intact cells, we used a novel pharmacological ligand that stabilizes *G-quadruplexes*, 360A, in lymphoblastoid cells and we show that treatment with this ligand induces significant changes in the expression patterns of FSp53 (encoding p53) and p53I2 (encoding Δ40p53) mRNAs. Specifically, 360A increased the levels of FSp53 mRNA and decreased the ones of p53I2 mRNA, consistent with studies using the GFP reporter system showing that presence of stable *G-quadruplexes* is important for efficient splicing-out of intron 2.

The notion that *G-quadruplex* structures can modulate the alternative splicing of several genes is not new: it has been demonstrated in the case of β-tropomyosin, hTERT or Bcl-xL transcripts (11,19,33). Compared with these examples, p53 *G-quadruplexes* show two important features. First, they are located in another intron (intron 3) than the one affected by alternative splicing (intron 2). This particular topography is due to the fact that introns and exons in this region are extremely short and some of the regulatory signals for the splicing of this intron may actually be located in intron 3. The splice acceptor of intron 2 is located just 60 bp upstream of the sequence forming *G-quadruplexes*. Second, the effects we demonstrate here are only partial. They do not induce a complete shift from one splicing pattern to the other, but rather modify the equilibrium between the two

Downloaded from http://carcin.oxfordjournals.org/ at UniversitÃ© de Limoges. Service commun de la documentation on August 28, 2013

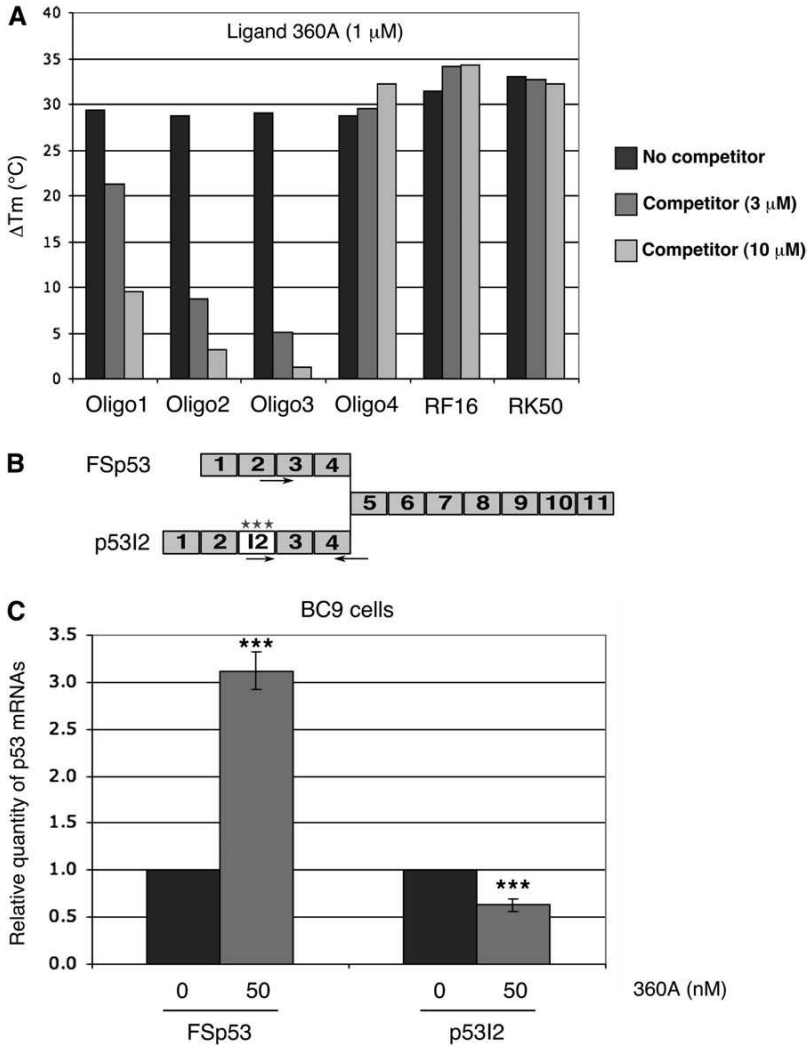


Fig. 5. Effect of the *G-quadruplex* ligand 360A on expression of p53 transcripts. (**A**) fluorescence resonance energy transfer assay demonstrating the capacity of 360A to bind intron 3 *G-quadruplexes*. The ΔT_m values for an oligonucleotide forming a canonical *G-quadruplex* (derived from *hTERT* sequence) were calculated in the presence or in the absence of competing oligonucleotides Oligo1–4 or control double-stranded RNAs (RF16 and RK50). (**B**) Structure of alternative p53 transcripts and positions of primers for quantitative reverse transcription–PCR. The FSp53 mRNA includes exons 1 to 11 and produces p53. The p53I2 mRNA retains intron 2 (I2) but correctly splices out all other introns. Asterisks: stop codons in intron 2; arrows: primers used for real-time PCR. (**C**) Quantitative reverse transcription–PCR analysis of mRNA isolated from BC9 lymphoblastoid cells using Sybr Green. Cells were treated for 48 h with 50 nM of 360A and levels of FSp53 and p53I2 were measured. $^{***}P < 0.001$; Student's *t*-test based on duplicate of five independent experiments.

spliced products in a subtle manner. Using the GFP reporter system, the shift in splicing patterns observed after mutagenesis of candidate guanines is ~25% ($P < 0.05$). In lymphoblastoid cells treated with 360A, we detect significant changes in the levels of both transcripts ($P < 0.001$), with a 3-fold increase for FSp53 and a 40% decrease for p53I2. These observations may be due, at least in part, to the fact that *G-quadruplexes* in intron 3 are polymorphic and may exist as different conformers. It is therefore possible that our site-directed mutagenesis experiments do not completely remove all *G-quadruplex* structures along the sequence of intron 3 but merely displace them or modify their organization. Furthermore studies will be necessary to uncover the full extent of the effect of *G-quadruplexes* on the splicing of intron 2.

The biological consequences of the regulatory effects identified here remain to be determined. Changes in transcript levels are expected to result into changes in the expression of both p53 protein isoforms when *G-quadruplexes* are stabilized, leading to increased p53 expression and decreased Δ40p53 expression. Indeed, we observed that 360A treatment resulted in a slight increased of p53 protein expression but we did not detect any significant impact on Δ40p53 protein levels or on the expression of the p53-target gene $p21^{WAF1}$ (supplementary Figure 4 available at *Carcinogenesis* Online). This suggests that 360A ligand modulates p53 splicing but that this effect does without significant activation of p53. Although p53 has a short half-life due to rapid recognition and proteasome-dependent degradation by Hdm2, Δ40p53 escapes this regulation and is much more stable (3,6). Furthermore, the biological effect of Δ40p53 remains controversial. Biochemical evidence shows that Δ40p53 interacts with p53 protein to form hetero-tetramers and that excess amounts of Δ40p53 downregulate the transactivation capacity of p53. When cotransfected together with wild-type p53 in p53-null cells, Δ40p53 prevents growth suppression by p53 and enhances clonogenicity (3,5,11,34). However, there is no evidence from *in vivo* studies that Δ40p53 operates as a suppressor of p53 function. In mice

Downloaded from http://carcin.oxfordjournals.org/ at UniversitÃ© de Limoges. Service commun de la documentation on August 28, 2013

with wild-type p53, increased dosage of Δ40p53 by expression of a transgene leads to accelerate aging and short lifespan. These mice display cognitive decline and synaptic impairment early in life, attributed to hyperactivation of insulin-like growth factor 1 receptor signaling and altered metabolism of the microtubule-binding protein tau (35,36). In human cells, Δ40p53 has been implicated in G_2 cell cycle arrest in response to endoplasmic reticulum stress (37). Although the molecular mechanisms of these effects are not understood, both mouse and human models suggest that small variations of Δ40p53 are sufficient to modify p53 biological function.

Another important consideration with respect to the potential impact of *G-quadruplex* in intron 3 is that they include guanine residues that are part of a common polymorphism, *TP53* PIN3, consisting of a 16 bp duplication (rs17878362; A1: non-duplicated allele; A2: duplicated allele) (38,39). Several epidemiological case–control studies have identified that the carriers of the A2 allele had a significantly increased risk of several common cancers, including breast, colorectal, lung and ovarian cancers (28,40–42). Recently, we have reported that *TP53* PIN3 is a genetic modifier of germ line *TP53* mutation in the Li-Fraumeni Syndrome, the A1 allele being associated with earlier age at first cancer diagnosis (43). Interestingly, Gemigniani *et al.* (28) described that p53 mRNA levels were lower in lymphoblastoid cell lines with A2 than with A1 alleles, suggesting that this polymorphism may affect the levels of p53 transcripts. *In silico* models predicts that *TP53* PIN3 may alter the topology of *G-quadruplexes* in intron 3 (data not shown). However, the extent of these modifications and their consequences on p53 splicing remains to be investigated.

In conclusion, our results demonstrate that *G-quadruplexes* are present in intron 3 of *TP53* gene and provide support in favor of a role of these structures in the regulation of the alternative splicing of intron 2, thus modulating the patterns of expression of transcripts encoding either p53 or its N-terminally truncated isoform, Δ40p53.

Supplementary material

Supplementary Tables I and II and Figures 1–4 can be found at http://carcin.oxfordjournals.org/

Funding

This work is supported by INSERM, Institut National contre le Cancer (INCa, France, Projet Libre 2009, 2009-192), an Agence Nationale de la Recherche grant (G4-toolbox), the Région Aquitaine (J.-L.M.) and by la Ligue Contre le Cancer (France), including a fellowship from La Ligue Nationale to V.M.

Acknowledgements

The authors thank Dr Hervé Moine for initial help in identifying candidate sequences forming *G-quadruplexes*.

Conflict of Interest Statement: None declared.

References

1. Vousden,K.H. *et al*. (2007) p53 in health and disease. *Nat. Rev. Mol. Cell Biol.*, **8**, 275–283.
2. Petitjean,A. *et al*. (2007) Impact of mutant p53 functional properties on TP53 mutation patterns and tumor phenotype: lessons from recent developments in the IARC TP53 database. *Hum. Mutat.*, **28**, 622–629.
3. Courtois,S. *et al*. (2002) DeltaN-p53, a natural isoform of p53 lacking the first transactivation domain, counteracts growth suppression by wild-type p53. *Oncogene*, **21**, 6722–6728.
4. Bourdon,J.C. *et al*. (2005) p53 isoforms can regulate p53 transcriptional activity. *Genes Dev.*, **19**, 2122–2137.
5. Marcel,V. *et al*. (2009) p53 isoforms—a conspiracy to kidnap p53 tumor suppressor activity? *Cell. Mol. Life Sci.*, **66**, 391–406.
6. Yin,Y. *et al*. (2002) p53 Stability and activity is regulated by Mdm2-mediated induction of alternative p53 translation products. *Nat. Cell Biol.*, **4**, 462–467.
7. Ghosh,A. *et al*. (2004) Regulation of human p53 activity and cell localization by alternative splicing. *Mol. Cell. Biol.*, **24**, 7987–7997.
8. Avery-Kiejda,K.A. *et al*. (2008) Small molecular weight variants of p53 are expressed in human melanoma cells and are induced by the DNA-damaging agent cisplatin. *Clin. Cancer Res.*, **14**, 1659–1668.
9. Candeias,M.M. *et al*. (2006) Expression of p53 and p53/47 are controlled by alternative mechanisms of messenger RNA translation initiation. *Oncogene*, **25**, 6936–6947.
10. Ray,P.S. *et al*. (2006) Two internal ribosome entry sites mediate the translation of p53 isoforms. *EMBO Rep.*, **7**, 404–410.
11. Gomez,D. *et al*. (2004) Telomerase downregulation induced by the G-quadruplex ligand 12459 in A549 cells is mediated by hTERT RNA alternative splicing. *Nucleic Acids Res.*, **32**, 371–379.
12. Lipps,H.J. *et al*. (2009) G-quadruplex structures: *in vivo* evidence and function. *Trends Cell Biol.*, **19**, 414–422.
13. Huppert,J.L. *et al*. (2005) Prevalence of quadruplexes in the human genome. *Nucleic Acids Res.*, **33**, 2908–2916.
14. Todd,A.K. *et al*. (2005) Highly prevalent putative quadruplex sequence motifs in human DNA. *Nucleic Acids Res.*, **33**, 2901–2907.
15. Henn,A. *et al*. (2008) Inhibition of dicing of guanosine-rich shRNAs by quadruplex-binding compounds. *Chembiochem*, **9**, 2722–2729.
16. Gros,J. *et al*. (2008) G-quadruplex formation interferes with P1 helix formation in the RNA component of telomerase hTERC. *Chembiochem*, **9**, 2075–2079.
17. Kumari,S. *et al*. (2007) An RNA G-quadruplex in the 5' UTR of the NRAS proto-oncogene modulates translation. *Nat. Chem. Biol.*, **3**, 218–221.
18. Wieland,M. *et al*. (2007) RNA quadruplex-based modulation of gene expression. *Chem. Biol.*, **14**, 757–763.
19. Didiot,M.C. *et al*. (2008) The G-quartet containing FMRP binding site in FMR1 mRNA is a potent exonic splicing enhancer. *Nucleic Acids Res.*, **36**, 4902–4912.
20. Hai,Y. *et al*. (2008) A G-tract element in apoptotic agents-induced alternative splicing. *Nucleic Acids Res.*, **36**, 3320–3331.
21. Granotier,C. *et al*. (2005) Preferential binding of a G-quadruplex ligand to human chromosome ends. *Nucleic Acids Res.*, **33**, 4182–4190.
22. Mergny,J.L. *et al*. (1998) Following G-quartet formation by UV-spectroscopy. *FEBS Lett.*, **435**, 74–78.
23. Mergny,J.L. *et al*. (2003) Analysis of thermal melting curves. *Oligonucleotides*, **13**, 515–537.
24. Mergny,J.L. *et al*. (2005) Thermal difference spectra: a specific signature for nucleic acid structures. *Nucleic Acids Res.*, **33**, e138.
25. Guedin,A. *et al*. (2008) Sequence effects in single-base loops for quadruplexes. *Biochimie*, **90**, 686–696.
26. Mergny,J.L. *et al*. (2001) Fluorescence resonance energy transfer as a probe for G-quartet formation by a telomeric repeat. *Chembiochem*, **2**, 124–132.
27. De Cian,A. *et al*. (2007) Fluorescence-based melting assays for studying quadruplex ligands. *Methods*, **42**, 183–195.
28. Gemignani,F. *et al*. (2004) A TP53 polymorphism is associated with increased risk of colorectal cancer and with reduced levels of TP53 mRNA. *Oncogene*, **23**, 1954–1956.
29. Kikin,O. *et al*. (2006) QGRS Mapper: a web-based server for predicting G-quadruplexes in nucleotide sequences. *Nucleic Acids Res.*, **34**, W676–W682.
30. Scaria,V. *et al*. (2006) Quadfinder: server for identification and analysis of quadruplex-forming motifs in nucleotide sequences. *Nucleic Acids Res.*, **34**, W683–W685.
31. Weitzmann,M.N. *et al*. (1996) The development and use of a DNA polymerase arrest assay for the evaluation of parameters affecting intrastrand tetraplex formation. *J. Biol. Chem.*, **271**, 20958–20964.
32. Matlashewski,G. *et al*. (1987) Alternative splicing of human p53 transcripts. *Oncogene Res.*, **1**, 77–85.
33. Sirand-Pugnet,P. *et al*. (1995) An intronic (A/U)GGG repeat enhances the splicing of an alternative intron of the chicken beta-tropomyosin pre-mRNA. *Nucleic Acids Res.*, **23**, 3501–3507.
34. Powell,D.J. *et al*. (2008) Stress-dependent changes in the properties of p53 complexes by the alternative translation product p53/47. *Cell Cycle*, **7**, 950–959.
35. Maier,B. *et al*. (2004) Modulation of mammalian life span by the short isoform of p53. *Genes Dev.*, **18**, 306–319.
36. Pehar,M. *et al*. (2010) Altered longevity-assurance activity of p53:p44 in the mouse causes memory loss, neurodegeneration and premature death. *Aging Cell*, **9**, 174–190.
37. Bourougaa,K. *et al*. (2010) Endoplasmic reticulum stress induces G2 cell-cycle arrest via mRNA translation of the p53 isoform p53/47. *Mol. Cell.*, **38**, 78–88.
38. Lazar,V. *et al*. (1993) Simple sequence repeat polymorphism within the p53 gene. *Oncogene*, **8**, 1703–1705.

Downloaded from http://carcin.oxfordjournals.org/ at UniversitÃ© de Limoges. Service commun de la documentation on August 28, 2013

39. Runnebaum,I.B. *et al*. (1994) Multiplex PCR screening detects small p53 deletions and insertions in human ovarian cancer cell lines. *Hum. Genet.*, **93**, 620–624.
40. Hung,R.J. *et al*. (2006) Sequence variants in cell cycle control pathway, X-ray exposure, and lung cancer risk: a multicenter case-control study in Central Europe. *Cancer Res.*, **66**, 8280–8286.
41. Runnebaum,I.B. *et al*. (1995) p53-based blood test for p53PIN3 and risk for sporadic ovarian cancer. *Lancet*, **345**, 994.
42. Wang-Gohrke,S. *et al*. (2002) Intron 3 16 bp duplication polymorphism of p53 is associated with an increased risk for breast cancer by the age of 50 years. *Pharmacogenetics*, **12**, 269–272.
43. Marcel,V. *et al*. (2009) TP53 PIN3 and MDM2 SNP309 polymorphisms as genetic modifiers in the Li-Fraumeni syndrome: impact on age at first diagnosis. *J. Med. Genet.*, **46**, 766–772.

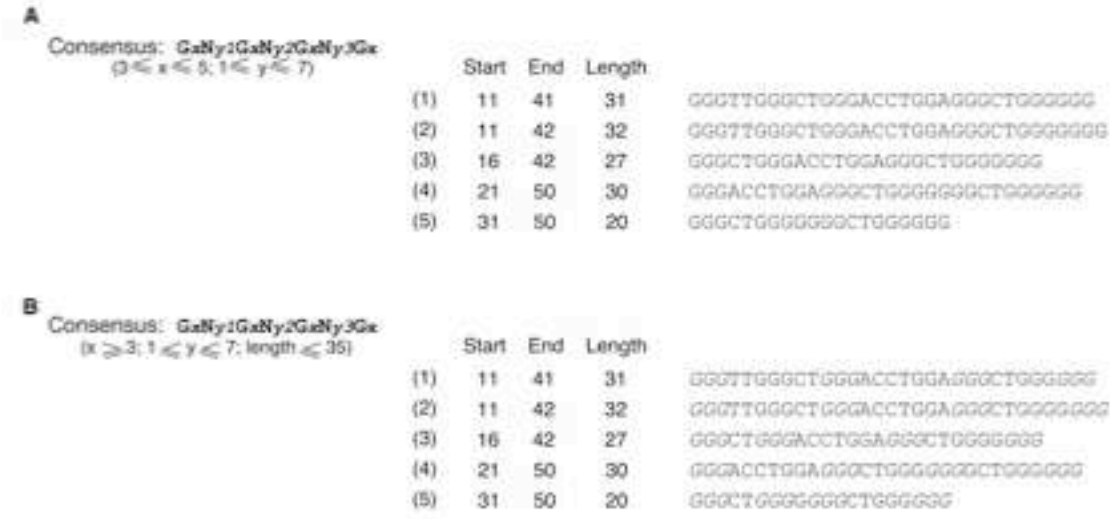
Received August 2, 2010; revised November 5, 2010; accepted November 14, 2010

Downloaded from http://carcin.oxfordjournals.org/ at UniversitÃ© de Limoges. Service commun de la documentation on August 28, 2013

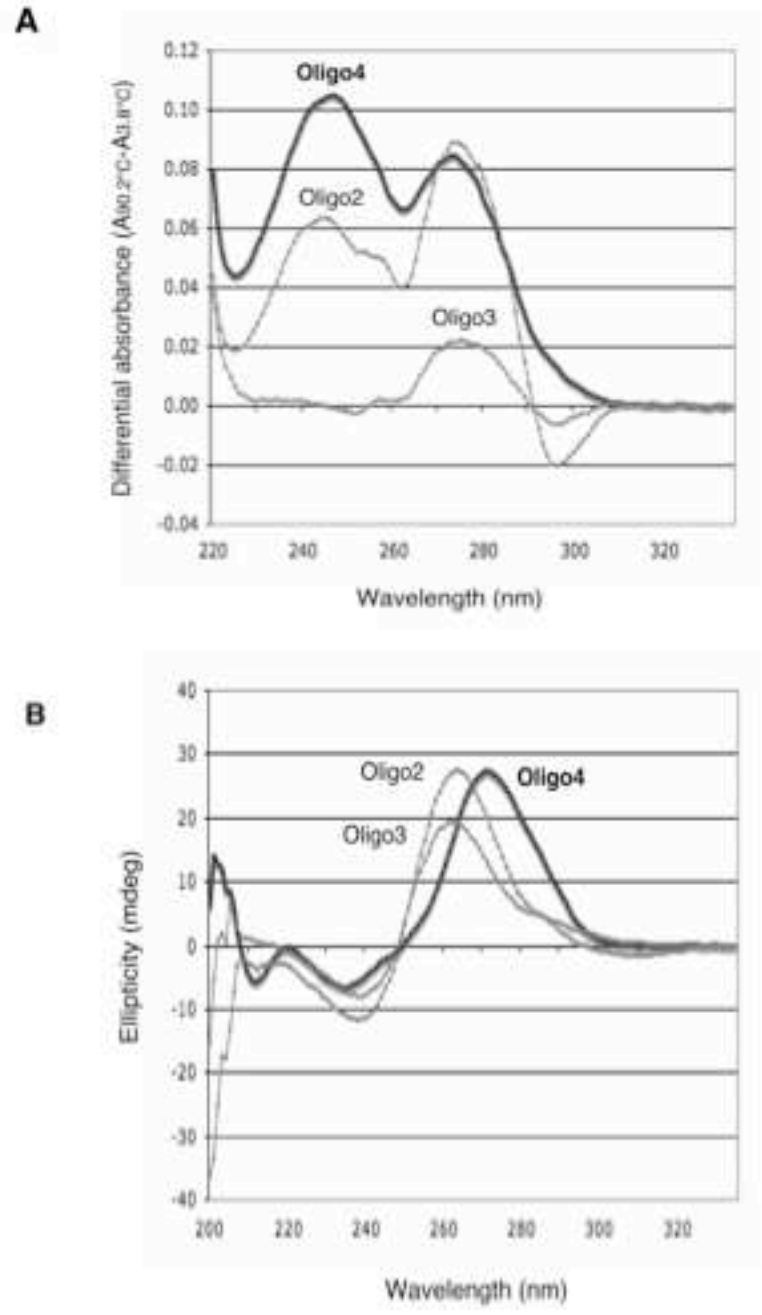
Supplementary materials

G-quadruplex structures in *TP53* intron 3: role in alternative splicing and in production of p53 mRNA isoforms

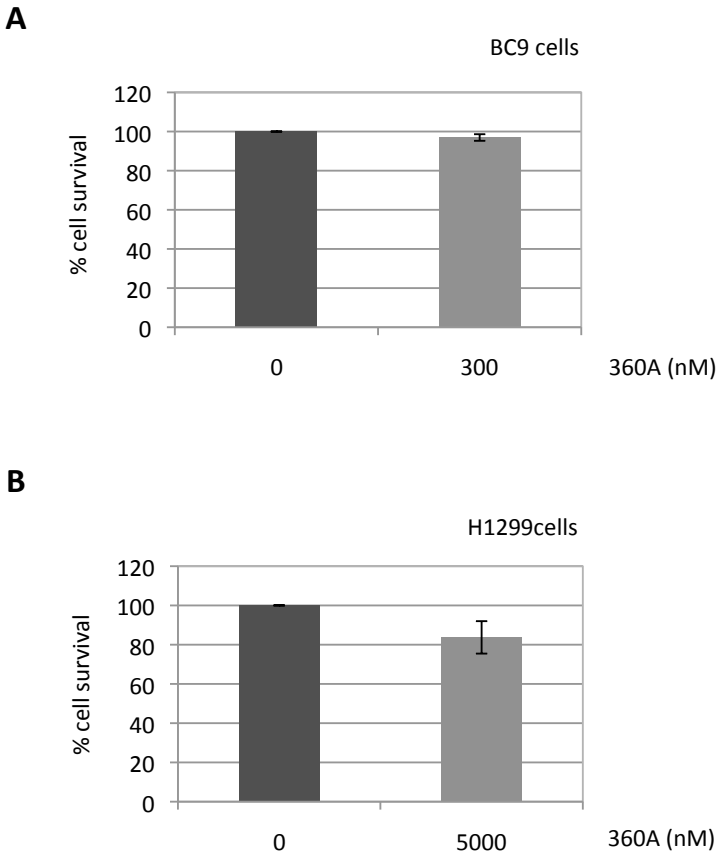
Virginie Marcel, Phong L.T. Tran, Charlotte Sagne, Ghyslaine Martel-Planche, Laurence Vaslin, Marie-Paule Teulade-Fichou, Janet Hall, Jean-Louis Mergny, Pierre Hainaut, Eric Van Dyck



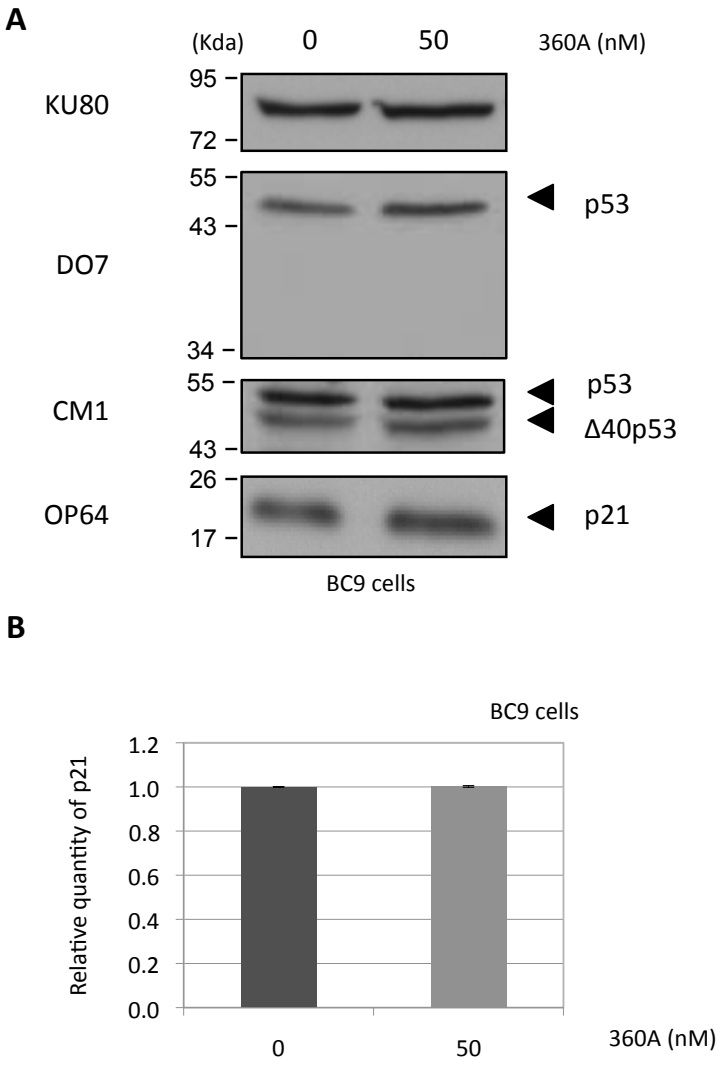
Supplementary Figure 1 – *In silico* prediction of *G-quadruplexes* using Quadfinger server (Scaria *et al.*, 2006) (**A**) and QGRS Mapper (Kikin *et al.*, 2006) (**B**). Italic guanine: guanines involved in G-tetrad formation.



Supplementary Figure 2 – *G-quadruplex* formation in mutated GFP plasmids. (**A**) UV thermal difference spectra of the Oligo4 oligomer. Compared to the TDS profiles of Oligo2 and Oligo3 that present a positive peak at 273 nm and a negative peak at 295 nm, the TDS profile of the Oligo4 has no negative peak, indicating that mutated Oligo4 RNA do not form a G-qua*druplex structure*. (**B**) CD spectroscopy of the Oligo4 oligomer. Both Oligo2 and Oligo3 are associated with a negative peak at 240 nm and a positive peak at 260 nm in CD spectroscopy, while the positive peak of the mutated Oligo4 is present at 270 nm. Overall, the mutated Oligo4 RNA is not able to form a *G-quadruplex* structure, indicating that the mutated GFP plasmids used to analyse the biological effet of the *G-quadruplex* in intron 3 are unable to form such a structure.



Supplementary Figure 3 – Cell viability under 48 hours of 360A ligand treatment. (**A**) Cell viability analysis of lymphoblastoid cells BC9 by trypan blue in response to 48 hours of 300 nM 360A ligand. One volume of cell suspension was diluted in one volume of trypan blue to count dead cells thus coloured in blue. Similar levels of cell survival were observed in non-treated and 360A treated cells (*P*-value = 0.125, t-student test). (**B**) Determination of H1299 cell survival by MTS assays in response to 360A treatment. After 48 hours of treatment, MTS assays were performed to assessed cell survival measured by absorbance using CellTiter 96® Assays (Promega). Similar levels of cell survival were observed in non-treated and 360A treated cells (*P*-value = 0.108). These two analyses indicate that 360A ligand, a RNA-specific *G-quadruplex* ligand, has no cytotoxic effect on BC9 and H1299 cells at high concentration (5 and 0.3 μM).



Supplementary Figure 4 – Impact of 360A drug, a RNA-specific *G-quadruplex* ligand, on p53 expression and transcriptional activity. (**A**) Analysis of p53 protein expression in response to 360A treatment. Protein extracted from BC9 lymophoblastoid cells treated for 48 hours by 50 nM 360A ligand were analysed by western blot. The 360A RNA-specific *G-quadruplex* ligand resulted in a slight increased of p53 protein expression but had no effect on Δ40p53 and p21 expression, suggesting a correlation between p53 mRNA and protein expression in response to 360A ligand that is not associated with an increased of p53 transcriptional activity. DO7: monoclonal p53 antibody, hybridazing the transactivation domain I presents in p53 but absent in Δ40p53; CM1: polyclonal p53 antibody, which detects both p53 and Δ40p53 isoforms; p21: monoclonal antibody specific of p21, a p53-target gene; Ku80: loading control. (**B**) Analysis of p21 expression, a p53-target gene in response to 360A ligand treatment. BC9 lymphoblastoid cells were treated for 48 hours by 50 nM of 360A ligand and RNA expression was analysed by quantitative RT-PCR using Sybr Green. p21 mRNA expression level was not affected by 360A treatment (*P*-value = 0.496, t-student test), suggesting that 360A ligand had no effect on the activation of p53 protein.

Supplementary Table I. Primers used for RT and RT-PCR analysis.

mRNA	Forward primers (5' to 3')		Reverse primers (5' to 3')	
FSp53 **p53I2**	E2/E3 I2	cctatggaaactacttcctg atgggactgactttctgctct	E4/E5	aggggactacgtgcaagt
Gapdh	GapdhF	tctcatggttcacacccatgacgaacatg	GapdhR	aagaagatgcggctgactgtcgagccacat
Actin	ActinF	atcgtggggcgccccaggcacca	ActinR	ctccttaatgtcacgcacgattt
Neomycin	NeoF	agacaatcggctgctctgat	NeoR	caatagcagccagtcccttc
p21		commercial set from Applied Biosystem (Rn00594010_m1*)		
			E4/GFP	gctcaccgatcgtgcttggga
			BSK-R	cccgggctgcaggaattc

Supplementary Table II. Primers used for site-directed mutagenesis.

Forward primers (5' to 3')		Reverse primers (5' to 3')	
PvuI	cgccaccatgcgatcggtgagcaagg	PvuI	ccttgctcaccgatcgcatggtggcg
PvuI/E2	gccgatcggaggagccgca	E4/PvuI	cgcgatcgtgcttgggacgg
GFP-Δ	cctggagggctgggggggctcaccacctgaggacctgg	GFP-Δ	ccaggtcctcaggtggtgagccccccccagccctccagg
GFP-ΔΔ	cctggagggctcaccaccctcaccacctgaggacctggtcct	GFP-ΔΔ	gaggaccaggtcctcaggtggtgagggtgagccctcagg

Impact of G-quadruplex structures and the polymorphisms rs17878362 and rs1642785 on the expression of *TP53* transcripts

Laury Perriaud[1,2], Virginie Marcel[1,2,6], Charlotte Sagne[1,2,], Vincent Favaudon[1,2], Aurore Guédin[3], Marie-Paule Teulade-Fichou[1,4], Pierre Hainaut[5], Jean-Louis Mergny[4], Janet Hall[1,2,*].

[1]Institut Curie, Centre de Recherche, Bât. 110-112, Centre Universitaire, Orsay, F-91405, France.

[2]Inserm, U612, Bât. 110-112, Centre Universitaire, Orsay, F-91405, France.

[3]INSERM, U869, Laboratoire ARNA, Univ. Bordeaux, Institut Européen de Chimie et Biologie, Pessac, F-33607, France.

[4]CNRS, UMR176, Bât. 110-112, Centre Universitaire, Orsay, F-91405, France.

[5]International Prevention Research Institute, Ecully, F-69130, France.

[6]Present address: Centre de Recherche en Cancérologie de Lyon, UMR INSERM 1052 CNRS 5286, Centre Léon Bérard, Lyon, F-69373, France.

* To whom correspondence should be addressed. Tel: +33 (0) 169863066; Fax: +33 (0) 1 69075327; Email: janet.hall@curie.fr

Running title: *TP53* polymorphisms and p53 isoform expression

Keywords: p53 isoforms, alternative splicing, G-quadruplex structures, ionizing radiation, rs17878362, rs1642785

Abstract

G-quadruplex (G4) structures are functionally active tri-dimensional structures formed in DNA and RNA guanine rich regions. We previously reported that G4 structures in intron 3 of the *TP53* pre-RNA modulate *TP53* intron 2 splicing and the balance between FSp53 (fully spliced p53 transcript) and p53I2 (transcript retaining intron 2). The nucleotides involved overlap the polymorphism rs17878362 (A1 wild-type allele, A2: 16 bp insertion). Using biophysical and biochemical approaches, we showed that both alleles form G4 structures that are shifted by 16 bp in relation to the intron 2 splice acceptor site in the A2 allele. Whilst FSp53 transcript levels were similar in rs17878362 A1A1 and A2A2 lymphoblastoid cell lines (LCLs), p53I2 levels was substantially lower and showed allele specific differences with the lowest levels seen in rs17878362 A2A2 LCLs. Treatment of the LCLs with 360A, a G-quadruplex binding ligand, or ionizing radiation increased the FSp53 transcript levels and decreased p53I2 levels in A1A1 cells suggesting that under these conditions full splicing is favored. Such trends were not observed in A2A2 cells. In addition, carriage of the intron 2 rs164278 (G > C) C allele which is in strong linkage disequilibrium with the A2 allele, was associated with a lower p53I2 transcript stability that may contribute to the lower levels of p53I2 transcript seen in the rs17878362 A2 sequence context. These results highlight the impact of *TP53* sequence variation on the formation of G-quadruplex structures, mRNA splicing and stability, and thus on the differential transcript expression of the *TP53* gene.

Introduction

G-quadruplexes (G4) are DNA or RNA tri-dimensional structures formed in guanine rich regions that have the propensity to fold into four-stranded, monovalent cation-dependent structures [1]. Such structures involve the interaction of four guanines in a cyclic Hoogsteen hydrogen bonding arrangement. Using bioinformatic and biochemical approaches G4 structures have been located throughout the genome and their role in gene expression and in genome maintenance revealed. Indeed, these DNA and RNA structures have been involved in the maintenance of genomic stability [2-4], DNA replication [5], gene transcription [6], mRNA splicing [7-8], mRNA synthesis [8-9] and translation [10]. Two G4s have been experimentally demonstrated in the p53 pre-mRNA: one located in intron 3 [7] and another downstream of the cleavage/polyadenylation site [9]. We have previously demonstrated that the G4 structures formed in *TP53* intron 3 regulate the alternative splicing of intron 2 thus altering the balance between expression of the FSp53 and p53I2 transcripts [7]. The FSp53 transcript encodes the full length 53 kDa protein TAp53, corresponding to the canonical p53 protein, and the Δ40p53 protein from an internal initiation of translation [11-12]. The second p53I2 transcript also encodes the Δ40p53 protein isoform [13-14]. The Δ40p53 protein lacks the first 39 residues corresponding to the transactivation domain 1 including the Hdm2-binding site and has been shown to modulate p53 transcriptional activities and p53-mediated growth suppression [11-12, 14-15].

The sequence environment in proximity to the G4 in intron 3 is highly polymorphic. Indeed it overlaps the common *TP53* polymorphism rs17878362 also located in intron 3 (*TP53 PIN3*; A1 no-insertion of 16bp, A2: insertion of 16 bp acctggagggctgggg) [16]. The rs17878362 variant allele is found with an allele frequency of between 0.10 – 0.17 in Caucasian populations with higher frequencies (0.15 – 0.21) being found in Indian populations [17]. A recent meta-analysis based on more than 10,000 cases and controls showed that carriage of the rs17878362 A2 allele is significantly associated with an increased cancer risk [17]. In addition, we have reported that rs17878362 is a modifier of the penetrance of germline *TP53* mutations in the Li-Fraumeni Syndrome [18]. *In silico* algorithms predicted that the allelic status of the rs17878362 polymorphism that contains several guanine-tracts may alter the topology of the G4 structures formed in intron 3. However, the functional impact of this polymorphism on the formation of G4 structures and on the regulation of the expression of the FSp53 and p53I2 *TP53* transcripts remains to be fully established.

Making use of biophysical and biochemical approaches, we show that the presence of the *TP53* rs17878362 A2 allele did not affect the formation of the G-quadruplex structure *per se* but modified its relative position in respect to the *TP53* intron/exon boundaries. Using *in vivo* assays we provide evidence that whilst the FSp53 transcript shows no allele specific variation in expression, carriage of the rs17878362 A2 allele substantially lowers the levels of the p53I2 transcript. In contrast G4 stabilisation using ligands or treatment of cells by ionising radiation (IR) results in increased FSp53 and decreased p53I2 transcript levels only in the presence of the rs17878362 A1 allele. In addition we show that the allelic status of rs1642785 (*TP53* PIN2, G > C) located in intron 2 impacts on the stability of the p53I2 transcript suggesting that levels of the FSp53 and p53I2 mRNA levels are strongly dependent upon the *TP53* genetic context.

Materiel and Methods

Synthetic RNA oligomers and compounds

RNA oligomers derived from the *TP53* intron 3 sequence (from A1 allele, 42N: AGGGUUGGGCUGGGGACCUGGAGGGCUGGGGGGCUGGGGGGC, and from A2 allele, 58D: AGGGUUGGGCUGGGGACCUGGAGGGCUGGGGACCUGGAGGGCUGGGGGGCUGGGGGGC) were synthesized by IBA (Göttingen, Germany), resuspended in double distilled water (ddH_2O) and concentrations were determined by ultraviolet (UV) absorbance. The compound 360A [(2,6-N,N'-methyl-quinolinio-3-yl)-pyridine dicarboxamide] was stored at 2mM in dimethyl sulfoxide and further diluted at 100µM in ddH20 for treating lymphoblastoid cell lines. Actinomycin D (Sigma, Saint-Louis, Missouri) was stored at 1mg/ml in dimethyl sulfoxide and used at this concentration to block transcription.

Biophysical methods

G4 formation and stability was assessed using the 42N (3 µM) and 58D (2µM) oligomers in 2.5mM K^+. Thermal difference spectrum (TDS) was measured for the two oligonucleotides at temperatures of 90°C (± 2°C) and 4°C (± 2°C°). Spectra were recorded between 220 and 335 nm in an Uvikon XL UV/Vis spectrophotometer [19] as described previously [7]. Circular dichroism (CD) spectra was recorded using a JASCO-810 spectropolarimeter as described previously [7, 20] after heating at 90°C for 5 minutes following by cooling to 20°C at 1°C/min rate. UV-melting experiments and analysis were

conducted as described previously [7] to characterize the thermal stability of RNA quadruplexes which could be formed by both oligonucleotides. Quadruplex denaturation/renaturation was followed by recording absorbance at 295 nm using the same Uvikon XL UV/Vis spectrophotometer.

Reverse transcriptase (RT) elongation assay

The sequence corresponding to the *TP53* intron 3 rs17878362 A2 allele was introduced into the pBSK plasmid and then used for performing primer extension as described previously [7]. In brief, 1 μg of pBSK containing either the rs17878362 A1- or A2 containing sequence was denatured in 0.2 N NaOH and 0.2 mM ethylenediaminetetraacetic acid (37°C – 30 min) and sequenced using Sequenase T7 (Amersham). Primer extension and sequencing products were analyzed on 10% polyacrylamide-8 M urea gels revealed by autoradiography.

Cell culture

Lymphoblastoid cell lines (LCLs) established from peripheral blood samples from breast cancer (BC) patients [21], were grown in Roswell Park Memorial Institute medium 1640 containing Glutamax (Invitrogen™, Life Technologies, Carlsbad, CA) and supplemented with 10% heat-inactivated fetal bovine serum (Gibco®, Life Technologies) and 1% penicillin/streptomycin. The H1299 cell line, derived from a human lung carcinoma, containing a homozygous partial deletion of the *TP53* gene and lacking expression of all p53 protein isoforms, was cultured in Dulbecco's modified Eagle's medium (DMEM) containing Glutamax (Invitrogen™) and supplemented with 10% fetal bovine serum and 1% penicillin/streptomycin. All cell lines were grown at 37°C in a humidified 5% CO_2 atmosphere.

TP53 genotype determinations

After extraction of genomic DNA from the LCLs using DNeasy Mini kit (Qiagen, Courtaboeuf, France), the region surrounding the rs1642785 and rs17878362 polymorphisms was amplified using the GoTaq® Hot Start Polymerase kit (Promega, Fitchburg, Wisconsin) and primers at a final concentration of 0.5μM (Forward 5' TCTCATGCTGGATCCCCACT 3', Reverse 5' ATACGGCCAGGCATTGAAGT 3' from *TP53* IARC database). The touchdown PCR conditions were 94°C for 2 mins; then 18 cycles of (94°C-30s; 63°C-45s, -0.5°C/3cycles; 72°C-1min) followed by 33 additional cycles of (94°C-30s, 60°C-45s, 72°C-1min) and a final step of 10 min hold at 72°C on a

Veriti 96-well fast thermal cycler (Applied Biosystems®, Life Technologies). PCR products (5µl) were purified using ExoSap (USB) (2µl) following the manufacturer's recommendations and sequenced using 150nM Forward primer, BigDye V1.1 on an Applied Biosystems® Genetic Analyser 3130.

Drug treatment and radiation

The day before treatment or radiation, LCLs were diluted to $0.5x10^6$ cell/ml. Cells were treated for 48 hrs with the G4 binding ligand 360A at a final concentration of 500nM or for 2 to 8 hrs with actinomycin D at a final concentration of 1µg/ml. After each treatment cells were harvested and nuclear RNA extracted.

Cells were irradiated using a 4.5 MeV linear electron accelerator facility built by EuroMeV (Buc, France) operated in a chopped mode as previously described [22]. Real-time measurement of the electron pulse was made through detection of the current generated by electrons trapped in a graphite collimator located close to the electron beam output, and the dose at cells calibrated with the aid of chemical dosimeters. The direction of the electron beam was vertical. The culture flasks were placed in a horizontal position at 540 mm from the electron beam output, in such a way that the thickness of the liquid layer (3 mm) was constant to ensure electronic equilibrium. Cells received a single sub-microsecond pulse with a beam current adjusted to provide the required dose (2.06 ± 0.07 or 0.103 ± 0.013 Gy over the 150 flasks irradiated during these experiments). Cells were harvested 48 hrs after radiation and nuclear RNA extracted.

Green fluorescent protein based splicing reporter system

A *TP53* minigene was used in splicing reporter assays as previously reported [7]. Briefly, the sequences from the end of exon 2 to the beginning of exon 4 of *TP53*, excluding ATG 1 and 40, and carrying either an intron 3 with the rs17878362 A1 allele (pEGFP-E2E4-A1) or an intron 3 sequence corresponding to the rs17878362 A2 allele (pEGFP-E2E4-A2) were introduced between the first ATG and the coding sequence of GFP protein. The day before transfection, p53-null H1299 cells ($0.5x10^5$ cells) were plated in 6-well format. Transient transfection of 50ng of plasmids was performed using Lipofectamine™ 2000 (Invitrogen™) according to the manufacturer's instructions. Cells were harvested 48 hrs after transfection and total RNA extracted.

Total and Nuclear RNA extraction

Total RNA extraction from H1299 transfected cells was performed using the RNeasy Mini kit (Qiagen, Courtaboeuf, France). Nuclear RNA was extracted from LCLs (at least 4.10^7 cells/treatment). Briefly, the cell pellet was resuspended twice in 3 volumes of buffer A (10mM Hepes pH7.9, 1.5mM $MgCl_2$, 10mM KCl, 0.1% Nonidet P-40, 1mM DTT) for 15 min on ice and then centrifuged at 13000rpm for 15min at 4°C. The pellet was resuspended in 2 volumes of buffer B (20mM Hepes pH7.9, 400mM NaCl, 1.5mM $MgCl_2$, 0.2mM EDTA, 0.5mM DTT) and then centrifuged at 13000rpm for 15 min at 4°C followed by a final step of supernatant clean-up to remove salt using the RNeasy Mini kit (Qiagen). Total and nuclear RNA were quantified by spectrophotometry (NanoDrop2000, Thermo Fisher Scientific Inc., Waltham, MA).

Real-time PCR analysis

The levels of nuclear FSp53 and p53I2 transcripts were quantified by reverse transcription (High capacity cDNA reverse transcription, Invitrogen) of 2µg of nuclear RNA followed by quantitative PCR (Fast SYBR® Green master mix, Applied Biosystems®) and using specific primers for each transcript (supplementary Table 1) on a StepOnePlus™ Real-time PCR system (Applied Biosystems®) according to the manufacturer's instructions. Transcript levels were normalized to that of *TBP* (TATA box binding protein). The ΔCt was calculated as an average of the quantitative reverse transcription–PCR of at least four independent experiments. Similar RT-qPCR method was performed to quantify the level of FS-GFP and I2-GFP transcripts as described above except that 1µg of total RNA was used and expression normalized to that of *neomycin* (supplementary Table 1). The ΔCt was calculated with an average of three measurements per plate for at least four experiments.

Results

Formation of G4 structures by the A2-intron 3 TP53 sequence

In a previous study, we reported that the *TP53* intron 3 carrying the non-duplicated version of the rs17878362 A1 allele could form G4 structures that overlap this polymorphic sequence [7]. In this present study we have now investigated whether the duplication of this 16 bp sequence (rs17878362 A2 allele) can affect the formation of G4 structures. Using the online prediction tools Quadfinder and

GQRS Mapper with default parameters [23-24], several putative, overlapping G4 structures were identified in the rs17878362 A2 allele sequence (Supplementary Table 2).

To demonstrate the formation of G4 structures in the rs17878362 A2 containing intron 3, biophysical analyses were performed using two synthetic RNA oligomers derived from this region: a 52-Mer sequence containing the rs17878362 A2 allele 16bp duplication and a 42-mer (derived from the rs17878362 A1 allele and corresponding to oligomer 2 of [7]). First, spectrum profiles of the two intron 3-derived RNA oligomers were determined by TDS and CD. Using these two techniques the G4 signatures were superimposable in the 58D and 42N oligomers (Figure 1A and 1B). Thermal stability was analysed under K^+ conditions by UV melting (Figure 1C). The Tm under these conditions for the 42N oligomer was 52° and for the 58D 52.6° suggesting that the additional bases present in the A2 allele do not have a major impact on the stability of the G4-quadruplex formed. These Tms were obtained at low ionic strength (2.5mM K^+ only), illustrating how stable those RNA quadruplexes can be under near-physiological conditions.

Since no differences in terms of formation and stability were observed using these approaches between the G4 structures of the rs17878362 A1- and A2-intron 3 containing oligomers, we next identified the position of the last guanines involved in the G4 structures formed in the corresponding sequences using a primer extension assay (Figure 2A). This method takes advantage of the differential cation-dependent stability of G4 structures, which stops the reverse transcriptase elongation. While KCl stabilizes G4 structures and promotes reverse transcription (RT) pauses, NaCl reduces the formation of G4 structures and allows complete RT. Using an *in vitro* transcribed A1- and A2-intron 3 RNAs as templates, a band corresponding to the full-length RT product was observed in the presence of NaCl that was shifted by 16bp in the rs17878362 A2 compared to the A1 sequence context. In contrast, several RT pauses were identified in the presence of KCl, confirming that G-quadruplex structures are formed within rs17878362 A1- and A2-intron 3 RNAs. Guanines resulting in RT pauses using the A2-intron 3 corresponded to those following the 16 bp duplicated motif of rs17878362 and were identical to the ones shown to induce RT pauses in the A1-intron 3 RNA (Figure 2B). Taken together these results suggest that although the G4 structures formed in the rs17878362 A1- and A2-intron 3 oligomers exhibit similar biochemical properties, those found in the A2-intron 3 compared to the A1-intron 3 sequence context are shifted by 16 bp away from the splice acceptor site

of intron 2. Thus the sequence context of the polymorphic rs17878362 region differentially affects the location of the G4 structures formed in *TP53* intron 3.

Impact of rs17878362 polymorphism on endogenous TP53 transcript expression

In *order* to investigate the impact of the rs17878362 polymorphism on *TP53* transcript expression we used two approaches. First we took advantage of an *in vitro* splicing assay previously developed [7] that makes use of a mini-gene containing the *TP53* sequence from the end of exon 2 to the beginning of exon 4 upstream of a GFP coding sequence. This allows the splicing of *TP53* intron 2 and thus the expression of the two transcripts of interest to be monitored: FS-GFP resulting from the splicing of intron 2 and I2-GFP resulting from the retention of the entire intron 2. Two polymorphic versions of this mini-gene have been developed that contain either the rs17878362 A1- or the A2-intron 3 sequence of *TP53*. In p53-null transfected H1299 cells the FS-GFP transcript was expressed in similar levels using both the A1 and A2 mini-genes whilst the I2-GFP transcript was expressed at significantly lower levels and showed rs17878362 allele specific differences with transfected cells carrying the rs17878362 A2 mini-gene having a 3-fold lower the level of the I2-GFP transcript than those carrying the rs17878362 A1 mini-gene (Figure 3A).

In order to assess the impact of the rs17878362 polymorphism on the endogenous transcript expression levels under basal conditions *in vivo*, we next quantified FSp53 and p53I2 transcripts in two LCLs homozygous for the rs17878362 A1 allele and two LCLs homozygous for the rs17878362 A2 allele using real-time PCR (Figure 3B). Whilst some variation in both transcript levels was noted between the different independent experiments carried out, the mean level of the FSp53 transcript did not show any allele-dependent variation in agreement with the results from the *in vitro* assays. In contrast the expression levels of the mRNA were significantly lower (up to 100-fold lower) than that of the FSp53 mRNA in the four LCLs and allele specific differences were also noted with the rs17878362 A2A2 LCLs having a 10-fold lower level of p53I2 compared to the rs17878362 A1A1 LCLs. These data suggest that *TP53* transcript expression is differentially affected by the rs17878362 polymorphism status or by a genetic background associated with the *TP53* rs17878362 polymorphism.

Impact of the G4 binding ligand 360A on transcript levels

We next investigated whether the G4 binding ligand 360A which has been shown to stabilise the G4 structures located in intron 3 [7] could impact on the levels of the two *TP53* transcripts. LCLs were used to investigate this as the p53I2 transcript levels were higher than those detectable in the *in vitro* assay. Three LCLs were treated with the ligand 360A (500 nM) for 48 hrs and the transcript levels of FSp53 and p53I2 were then quantified by real-time PCR (Figure 4). In the LCL carrying the homozygous rs17878362 A1 allele (BC56), 360A treatment led to a significant increase of FSp53 RNA levels and decrease of p53I2 RNA levels compared to that seen in untreated cells (Figure 4A and 4B, left panel). These data were in accordance with our previous observations for another cell line homozygous for this rs17878362 allele (BC9 cells) after 360A treatment [7], suggesting that the G4 structures formed in the *TP53* rs17878362 A1 allele can be further stabilized by the 360A ligand favoring the splicing of intron 2 and thus a reduction in p53I2 transcript levels. However in the BC156 and BC48 LCLs homozygous for the rs17878362 A2 allele no major effects on the level of either transcript could be detected after treatment with 360A. Thus whether the G4 structures in this sequence context can be further stabilized remains an open question.

Expression of TP53 transcripts after exposure to IR

In order to assess whether the allele specific expression of TP53 transcripts could be affected by DNA damage, we next exposed the four LCLs to IR using a linear electron accelerator and measured by real-time PCR the effect on FSp53 and p53I2 transcript levels 48 hrs after treatment. In the homozygous LCLs carrying the rs17878362 A1 alleles (BC9 and BC56), treatment with 2Gy resulted in a significant increase of FSp53 mRNA levels associated with a significant decrease of p53I2 transcript level in BC9, with the same trend being observed for these transcripts in BC56 (Figure 5A and B). After exposure to a low radiation dose (0.1Gy) a trend towards an increase in the levels of FSp53 were observed in both LCLs while the level of p53I2 transcript was unchanged at this time point after treatment, suggesting that even after exposure to such low levels of IR FSp53 levels can be modulated (Fig 6C and D). These radiation-induced changes in transcript levels paralleled that seen after the treatment with the G4 binding ligand 360A, in support of our working hypothesis that the stabilization of the G4 structures in intron 3 is associated with the splicing of intron 2 resulting in increased FSp53 levels. In the LCLs homozygous for the rs17878362 A2 allele (BC156 and BC48), radiation exposure, as for treatment with the G4 binding ligand, resulted in no consistent changes in

the profile of transcript levels of FSP53 or p53I2 (with the exception of subtle, but significant fluctuations in BC156 and BC48 after some individual exposure doses).

Impact of the rs1642785 polymorphism on the stability of the p53I2 transcript

Our biophysical, *in vitro* and *in vivo* experimental results support our hypothesis that the sequence surrounding the rs17878362 polymorphism could impact on the splicing of intron 2 and in particular in the context of the A1 allele. However the lack of any detectable stabilization of the G4 structure after treatment with 360A or IR and a concomitant reduction in the level of the p53I2 transcript level in LCLs carrying the rs17878362 A2 allele could be interpreted as evidence for a role of other genetic alterations in the statistically lower levels of the p53I2 transcript seen in this sequence context. Indeed compared to the FSp53 transcript, the entire intron 2 sequence is maintained in the p53I2 transcript and thus any variant alleles that are located within this sequence would also be present uniquely in this transcript. One polymorphism located in intron 2 that is in linkage disequilibrium with rs17878362 is rs1642785 (G>C). Indeed in both the mini-gene *in vitro* assays and the LCLs models presented above carriage of the rs17878362 A1 allele was associated with the rs1642785 G allele and the rs17878362 A2 allele with the rs1642785 C allele. Thus we cannot exclude that the decrease in p53I2 mRNA levels observed in the presence of the rs17878362 A2 allele could be due to an impact of the rs1642785 C allele on either splicing of intron 2 or the transcript's stability. In order to examine this latter possibility we determined the p53I2 mRNA half-life after treatment of the LCLs with actinomycin D. Under these experimental conditions, the half-life of the p53I2 transcript in the two LCLs containing the rs1642785 C and rs17878362 A2 alleles was approx. 2hrs whilst the half-life of the p53I2 transcript in the two LCLs carrying the rs1642785 G and rs17878362 A1 alleles exceeded 8 hrs (Figure 6). Under the same experimental conditions the half-life of the FSp53 transcripts was between 4 and 8 hrs depending on the cell line (data not shown). These results suggested that the allele status of the rs1642785 polymorphism could be a contributory factor to the low levels of the p53I2 transcript seen in the context of the rs17878362 A2 allele.

Discussion

The *TP53* gene expresses several protein isoforms encoded by different transcripts whose expression is finely regulated at the transcriptional and post-transcriptional levels. In particular the FSp53 and

p53I2 transcripts are generated by alternative splicing of the *TP53* pre-RNA and encode two related protein isoforms, the canonical TAp53 protein and the N-truncated Δ40p53 isoform, which modulate p53-mediated growth suppression (reviewed in [25]). We previously reported that the sequence surrounding the most frequently found *TP53* rs17878362 A1 allele, can form G4 structures and that treatment of the BC9 LCL homozygous for the rs17878362 A1 allele with the G4 binding ligand 360A (50nM) for 48 hrs significantly increased the levels of FSp53 transcript while significantly reducing those of the p53I2 transcript. These results suggested that G4 structures in intron 3 regulate the splicing of intron 2 and thus the balance between the FSp53 and p53I2 transcripts [7]. In this present study, we have confirmed and extended these observations in a second, non-related LCL (BC56). A similar inverse expression pattern of changes in FSp53 and p53I2 mRNA levels in these two homozygous rs17878362 A1 LCLs in response to high doses of IR and modulation of FSp53 levels after low IR doses were also observed suggesting that the balance between the FSp53 and p53I2 transcripts is regulated in response to DNA damage.

Next we investigated the potential of synthetic RNA oligomers containing the rs17878362 A2 allele to form G4 structures using different biophysical methods (UV melting, TDS and CD spectroscopy). These approaches and *in vitro* RT elongation assays provided evidence that in this sequence context G4 structures were formed with a shift in location in respect to the splice acceptor site of intron 2. Using a GFP-reporter system containing the rs178783 A2 sequence transfected into *TP53* null cells and LCLs carrying the rs17878362 A2 allele we quantitated the levels of the FSP53 and p53I2 transcript levels *in vivo*. Under these experimental conditions FSp53 transcript levels were similar to those observed in cells transfected with the rs178783 A1 containing plasmids or LCLS carrying the rs17878362 A1 allele. In contrast the basal level of the p53I2 transcript was significantly lower than that associated with the rs17878362 A1 allele in both model systems suggesting that presence of the 16 additional bps and the resulting shift in G4 position in intron 3 could favour the splicing of intron 2. It has to be noted however that despite these differences in the basal levels of the p53I2 transcripts seen between the rs17878362 A1 and A2 LCLs no changes in the level of the p53I2 transcript could be detected after treatment either with the G4 binding ligand or after exposure to IR in the rs17878362 A2 LCLs. This lack of effect of the G4 binding ligand could be interpreted to suggest that the G4 structures formed in the rs17878362A2 sequence context are intrinsically more stable than those found in the rs17878362 A1 context. However the biophysical approaches using the oligomers

corresponding to the two intron 3 sequence did not reveal any differences in G4 stability nor was indicated using the RT-elongation assay. It should also be noted that the absolute level of expression of the p53I2 transcript in the LCLs is lower in the A2 sequence context than after treatment of the rs17878362 A1 cells with 360A and approached the detection limit of the quantitative PCR system which is clearly a technical limitation of the cell model system used. It will clearly be interesting to identify alternative cell model systems where the basal levels of the p53I2 are higher to explore the stability of the rs17878362 A2-derived G4 and to examine whether it can be further modulated by a ligand *in vivo* or following DNA damage. These results could also be interpreted as evidence for a role for other factors either influencing the splicing of intron 2 or the stability of the p53I2 transcript in this sequence context. One candidate would be genetic factors and in particular polymorphisms located in intron 2 such as rs1642785 (G>C) that would be present uniquely in this transcript. The finding of rs1642785 allele specific differences in the half-life of the p53I2 transcript supports the possibility that this SNP may contribute to transcript stability although the molecular basis remains to be established.

G4 structures are known to be structural binding motifs for RNA-binding proteins, including hnRNP proteins, involved in alternative splicing (reviewed in [8]) and it has been shown that the position of hnRNP protein binding is determinant for the inclusion of alternative exons [26]. There is also accumulating evidence for sequence context dependent splicing and it has been reported that the intronic guanine tract density and length correlate with hnRNPH/F enhancer function [27-28] and that the presence of polymorphisms near splicing sites can influence splicing [29]. Clearly the topology of the region implicated in the splicing of intron 2 and the sequence duplicated in the A2 allele in intron 3 will modify certain of these criteria. Indeed the introns and exons in this region are extremely short and some of the regulatory signals for splicing of this intron may actually be located in intron 3. The splice acceptor of intron 2 is located just 60 bp upstream of the sequence forming the G4 in the rs17878362 A1 sequence context. Our findings that exposure to ionizing radiation also results in a decrease in the p53I2 transcript levels in the context of the rs17878362 A1 sequence would suggest that the presence of G4 structure could impact on radiation responses. Indeed the treatment of human glioblastoma cells or bacteria with G4 ligands radiosensitizes cells to X-ray and γ-irradiation exposure, respectively [30-31]. In addition, a recent study reported that stabilization of a G4 structure by hnRNPH/F protein binding in response to UV-B exposure increases p53 mRNA processing thus increasing p53 activity [9].

Overall, our data showed that FSp53 and p53I2 mRNA levels is strongly dependent upon *TP53* genetic context that modulate different transcriptional and post-transcriptional processes regulating p53 isoform expression. Such observations are not restricted to the FSp53 and p53I2 transcripts and the protein isoforms that they encode, other polymorphisms have been involved in the differential expression of p53 isoforms by modulating transcriptional and translational process (reviewed in [25]). We recently showed that rs17878362 and rs1042522 (P>R) polymorphisms modulate promoter activity of the internal *TP53* P2 promoter regulating expression of D133p53 mRNA [32]. In particular, the rs17878362 A1-P72 and rs17878362 A2-R72 alleles exhibit a lower promoter activity than rs17878362 A1-R72 and rs17878362 A2-P72 allele. The impact of *TP53* polymorphisms on internal promoter activity has been extended to 12 polymorphisms in a systematic study also supporting that *TP53* polymorphisms differentially affect D133p53 mRNA expression [33]. In addition, genetic variations in the 5'UTR of FSp53 mRNA have been shown to affect the cap-independent translation of both p53 and D40p53 protein isoforms [34-35]. Expression of p53 isoforms is thus under genetic control resulting in a differential pattern of p53 isoforms found throughout the general population that may directly participate in the increased cancer susceptibility associated with non-coding or synonymous *TP53* polymorphisms.

Our data support growing evidence that non-coding genetic variations modulate expression of p53 isoforms by affecting different transcriptional, post-transcriptional and translational mechanisms. Genetic alterations can modify both nucleotide sequences and structures within p53 RNA that are cis-regulators of p53 expression. Since p53 isoforms have been shown to modulate p53 transcriptional activity and thus the p53-mediated growth suppression, we may expect that in the general population, expression pattern of p53 isoforms is different in an individual-dependent manner. Thus, non-coding polymorphisms by modulating isoforms expression may be as important as coding polymorphism in regulating p53 activity (reviewed in [36]).

Funding

This work was supported by grant INCa 2009-192 "*TP53* intron3" from the French National Cancer Institute to J.H, P.H and J.L.M. Research in Inserm U612 is also supported by funding from Institut Curie and Inserm. C.S has a PhD fellowship from the French Ministry of Research and V.M and L.P were supported by funding from EU FP7 (Grant Number 249689 for the network of excellence DoReMi

(low dose research towards multidisciplinary integration). J.L.M. was supported by grants from ANR (Quarpdiem, TKi-net and Oligoswitch), Conseil Régional d'Aquitaine ("Chaire d'accueil" and Aquitaine-Midi Pyrénées grants) and a “subvention libre” from Fondation ARC.

Conflict of Interest Statement: None declared.

REFERENCES

1. Lipps, H.J. and D. Rhodes, *G-quadruplex structures: in vivo evidence and function.* Trends Cell Biol, 2009. **19**(8): p. 414-22.
2. Ribeyre, C., et al., *The yeast Pif1 helicase prevents genomic instability caused by G-quadruplex-forming CEB1 sequences in vivo.* PLoS Genet, 2009. **5**(5): p. e1000475.
3. Paeschke, K., J.A. Capra, and V.A. Zakian, *DNA replication through G-quadruplex motifs is promoted by the Saccharomyces cerevisiae Pif1 DNA helicase.* Cell, 2011. **145**(5): p. 678-91.
4. Paeschke, K., et al., *Pif1 family helicases suppress genome instability at G-quadruplex motifs.* Nature, 2013. **497**(7450): p. 458-62.
5. Cayrou, C., et al., *Genome-scale identification of active DNA replication origins.* Methods, 2012. **57**(2): p. 158-64.
6. Bochman, M.L., K. Paeschke, and V.A. Zakian, *DNA secondary structures: stability and function of G-quadruplex structures.* Nat Rev Genet, 2012. **13**(11): p. 770-80.
7. Marcel, V., et al., *G-quadruplex structures in TP53 intron 3: role in alternative splicing and in production of p53 mRNA isoforms.* Carcinogenesis, 2011. **32**(3): p. 271-8.
8. Millevoi, S., H. Moine, and S. Vagner, *G-quadruplexes in RNA biology.* Wiley Interdiscip Rev RNA, 2012. **3**(4): p. 495-507.
9. Decorsiere, A., et al., *Essential role for the interaction between hnRNP H/F and a G quadruplex in maintaining p53 pre-mRNA 3'-end processing and function during DNA damage.* Genes Dev, 2011. **25**(3): p. 220-5.
10. Agarwala, P., et al., *The G-quadruplex augments translation in the 5′ untranslated region of transforming growth factor beta2.* Biochemistry, 2013. **52**(9): p. 1528-38.
11. Yin, Y., et al., *p53 Stability and activity is regulated by Mdm2-mediated induction of alternative p53 translation products.* Nat Cell Biol, 2002. **4**(6): p. 462-7.
12. Courtois, S., et al., *DeltaN-p53, a natural isoform of p53 lacking the first transactivation domain, counteracts growth suppression by wild-type p53.* Oncogene, 2002. **21**(44): p. 6722-8.
13. Matlashewski, G., et al., *Alternative splicing of human p53 transcripts.* Oncogene Res, 1987. **1**(1): p. 77-85.
14. Ghosh, A., D. Stewart, and G. Matlashewski, *Regulation of human p53 activity and cell localization by alternative splicing.* Mol Cell Biol, 2004. **24**(18): p. 7987-97.
15. Maier, B., et al., *Modulation of mammalian life span by the short isoform of p53.* Genes Dev, 2004. **18**(3): p. 306-19.
16. Lazar, V., et al., *Simple sequence repeat polymorphism within the p53 gene.* Oncogene, 1993. **8**(6): p. 1703-5.

17. Sagne, C., et al., *A meta-analysis of cancer risk associated with the TP53 intron 3 duplication polymorphism (rs17878362): geographic and tumor-specific effects.* Cell Death Dis, 2013. **4**: p. e492.
18. Marcel, V., et al., *TP53 PIN3 and MDM2 SNP309 polymorphisms as genetic modifiers in the Li-Fraumeni syndrome: impact on age at first diagnosis.* J Med Genet, 2009. **46**(11): p. 766-72.
19. Mergny, J.L., A.T. Phan, and L. Lacroix, *Following G-quartet formation by UV-spectroscopy.* FEBS Lett, 1998. **435**(1): p. 74-8.
20. Guedin, A., et al., *Sequence effects in single-base loops for quadruplexes.* Biochimie, 2008. **90**(5): p. 686-96.
21. Angele, S., et al., *ATM haplotypes and cellular response to DNA damage: association with breast cancer risk and clinical radiosensitivity.* Cancer Res, 2003. **63**(24): p. 8717-25.
22. Favaudon, V., et al., *CO2.- radical induced cleavage of disulfide bonds in proteins. A gamma-ray and pulse radiolysis mechanistic investigation.* Biochemistry, 1990. **29**(49): p. 10978-89.
23. Scaria, V., et al., *Quadfinder: server for identification and analysis of quadruplex-forming motifs in nucleotide sequences.* Nucleic Acids Res, 2006. **34**(Web Server issue): p. W683-5.
24. Kikin, O., L. D'Antonio, and P.S. Bagga, *QGRS Mapper: a web-based server for predicting G-quadruplexes in nucleotide sequences.* Nucleic Acids Res, 2006. **34**(Web Server issue): p. W676-82.
25. Marcel, V., et al., *Biological functions of p53 isoforms through evolution: lessons from animal and cellular models.* Cell Death Differ, 2011. **18**(12): p. 1815-24.
26. Konig, J., et al., *iCLIP reveals the function of hnRNP particles in splicing at individual nucleotide resolution.* Nat Struct Mol Biol, 2010. **17**(7): p. 909-15.
27. Xiao, X., et al., *Splice site strength-dependent activity and genetic buffering by poly-G runs.* Nat Struct Mol Biol, 2009. **16**(10): p. 1094-100.
28. Wang, E., et al., *Global profiling of alternative splicing events and gene expression regulated by hnRNPH/F.* PLoS One, 2012. **7**(12): p. e51266.
29. Lu, Z.X., et al., *Context-dependent robustness to 5' splice site polymorphisms in human populations.* Hum Mol Genet, 2011. **20**(6): p. 1084-96.
30. Merle, P., et al., *Telomere targeting with a new G4 ligand enhances radiation-induced killing of human glioblastoma cells.* Mol Cancer Ther, 2011. **10**(10): p. 1784-95.
31. Beaume, N., et al., *Genome-wide study predicts promoter-G4 DNA motifs regulate selective functions in bacteria: radioresistance of D. radiodurans involves G4 DNA-mediated regulation.* Nucleic Acids Res, 2013. **41**(1): p. 76-89.
32. Marcel, V., et al., *Diverse p63 and p73 isoforms regulate Delta133p53 expression through modulation of the internal TP53 promoter activity.* Cell Death Differ, 2012. **19**(5): p. 816-26.
33. Bellini, I., et al., *DeltaN133p53 expression levels in relation to haplotypes of the TP53 internal promoter region.* Hum Mutat, 2010. **31**(4): p. 456-65.
34. Grover, R., et al., *Effect of mutations on the p53 IRES RNA structure: implications for de-regulation of the synthesis of p53 isoforms.* RNA Biol, 2011. **8**(1): p. 132-42.
35. Khan, D., et al., *Effect of a natural mutation in the 5' untranslated region on the translational control of p53 mRNA.* Oncogene, 2013. **32**(35): p. 4148-59.
36. Whibley, C., P.D. Pharoah, and M. Hollstein, *p53 polymorphisms: cancer implications.* Nat Rev Cancer, 2009. **9**(2): p. 95-107.

FIGURES LEGENDS

Fig 1: TDS, CD and UV melting experiments. (A) Thermal difference Spectra. TDS result from the difference between the absorbance spectrums recorded at 90° ± 2°C and at 4° ± 2°C (in K^+). They were recorded between 220 and 335 nm. (B) Circular dichroism spectra. CD spectra were recorded at 20°C (in K^+) on a JASCO-810 spectropolarimeter using 1 cm path length quartz cuvettes. The oligonucleotides were annealed by eating at 90°C for 5 min, followed by cooling to 20°C at 1°C/min rate. (C) UV melting profiles. Absorbance at 295 nm is plotted as a function of temperature for both sequences in K^+.

Fig 2: RT-elongation assay to identify G4s localization in rs17878362-A1 and -A2 containing sequences. (A) G4s were mapped by reverse transcription of intron 3 sequences in the presence of a radiolabelled oligonucleotide and either K^+ (stabilizing G4) or Na^+ (destabilising G4) and products were separated by denaturing gels electrophoresis. Lanes 1 – 4 and 9 – 12: dideoxy sequencing of the sequence used for reverse transcription. RT-stops (lanes 5 – 8) occurred at guanine residues as shown by comparison with the sequencing lanes. Black arrowhead: 3'-intron 3 end. Open arrow head: 5'-intron 3 end; Brackets show position of rs17878362 sequence. (B) Shorter exposure of area corresponding to the aera delimited by dotted lines identifying the presence of RT-stops. (C) rs17878362-A1 and -A2 sequences single and double underline: PIN3 polymorphic sequence.

Fig 3: Variation in *TP53* transcript levels depending on the rs17878362 allele status. (A) *FS-GFP* (left panel) and *I2-GFP* (right panel) transcript levels in p53-null transfected HT1299 cells were monitored by quantitative RT-PCR with *neomycin* expression used as the reference. The bar represents the mean value ± SEM from at least 4 independent experiments. **P<0.01: Student's paired t-test. (B) FSp53 (left panel) and p53I2 (right panel) transcript levels in LCLs carrying the rs17878362 A1 or A2 allele measured using quantitative RT-PCR with *TBP* expression used as a reference. Each point of scatter plot indicates an independent experiment and for each cell line mean value ± SEM are represented by bars. ***P<0.001: Student's t-test

Fig 4: Effect of G4 ligand on *TP53* transcript levels depending on the status of rs17878362. Forty-eight hours after treatment with 500nM of 360A, the relative FSp53 (A) and relative p53I2 transcript levels were measured (B). Each bar of the histograms indicates mean value ± SEM of at least four experiments performed in each LCL carrying either the rs17878362 A1 or A2 allele. *P<0.05: Student's paired t-test

Fig 5: *TP53* transcript levels after DNA damage induced by IR. FSp53 (A) (C) and p53I2 (B)(D) transcript levels normalized to *TBP* 48 hrs hours after exposure of LCLs to 2 Gy (A and B) or 0.1 Gy (C and D) of IR. The bars of the histograms indicate mean value ± SEM of at least four experiments performed in three LCLs carrying either the rs17878362 A1 or A2 allele. *P<0.05, **P<0.01: Student's paired t-test.

Fig 6: Effect of *rs1642785* on p53I2 mRNA stability. The p53I2 transcripts were quantified by real-time PCR up to 8 hrs after treatment of LCLs with actinomycin D and normalized to *TBP*. Each curve represents relative p53I2 mRNA level in LCLs carrying (A) the rs17878362 A1 and rs1642785 G alleles or (B) the rs17878362 A2 and rs1642785 C alleles. Values are means ± SEM of at least three experiments. The half-life represented the time required to obtain the half of the initial quantity of p53I2 transcript.

SUPPLEMENTARY TABLES

Supplementary Table 1: primers used for Real-Time PCR

Supplementary Table 2. *In silico* prediction of G-quadruplex structures.

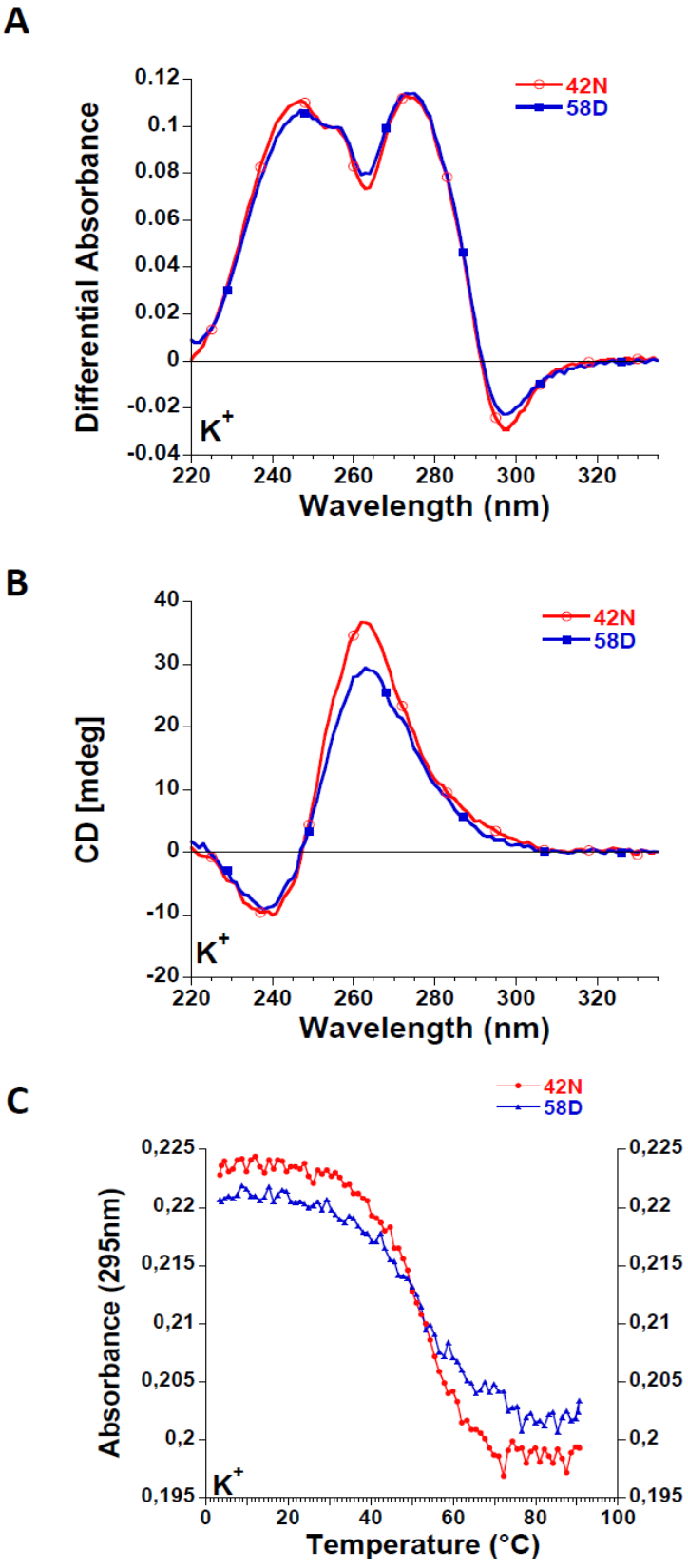
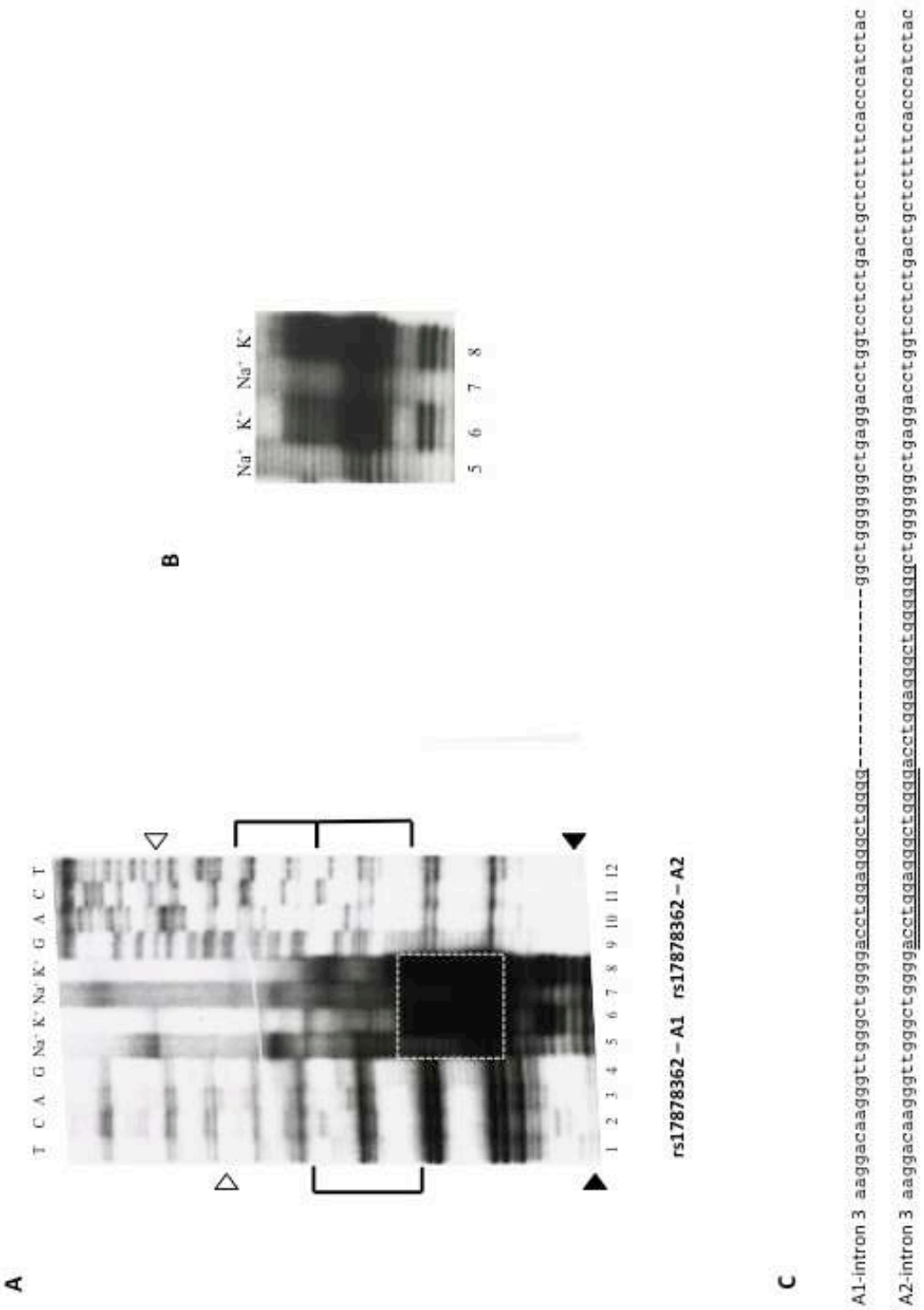

A
Differential Absorbance
0.12
0.1
0.08
0.06
0.04
0.02
0
-0.02
-0.04
220
240
260
280
300
320
Wavelength (nm)
42N
58D
K+
B
CD [mdeg]
40
30
20
10
0
-10
-20
220
240
260
280
300
320
Wavelength (nm)
42N
58D
K+
C
Absorbance (295nm)
0,225
0,22
0,215
0,21
0,205
0,2
0,195
0
20
40
60
80
100
Temperature (°C)
42N
58D
K+

Fig 1

Fig 2


A
T C A G Na+ K+ Na+ K+ G A C T
1 2 3 4 5 6 7 8 9 10 11 12
rs17878362 – A1
rs17878362 – A2
B
Na+ K+ Na+ K+
5 6 7 8
C
A1-intron 3 aagggacaagggttgggctgggacctgaagggctgggg----------ggctgggggggctgaggacctggtcctctgactgctctttcacccatctac
A2-intron 3 aagggacaagggttgggctgggacctgaagggctgggacctgaagggctgggggggctgaggacctggtcctctgactgctctttcacccatctac

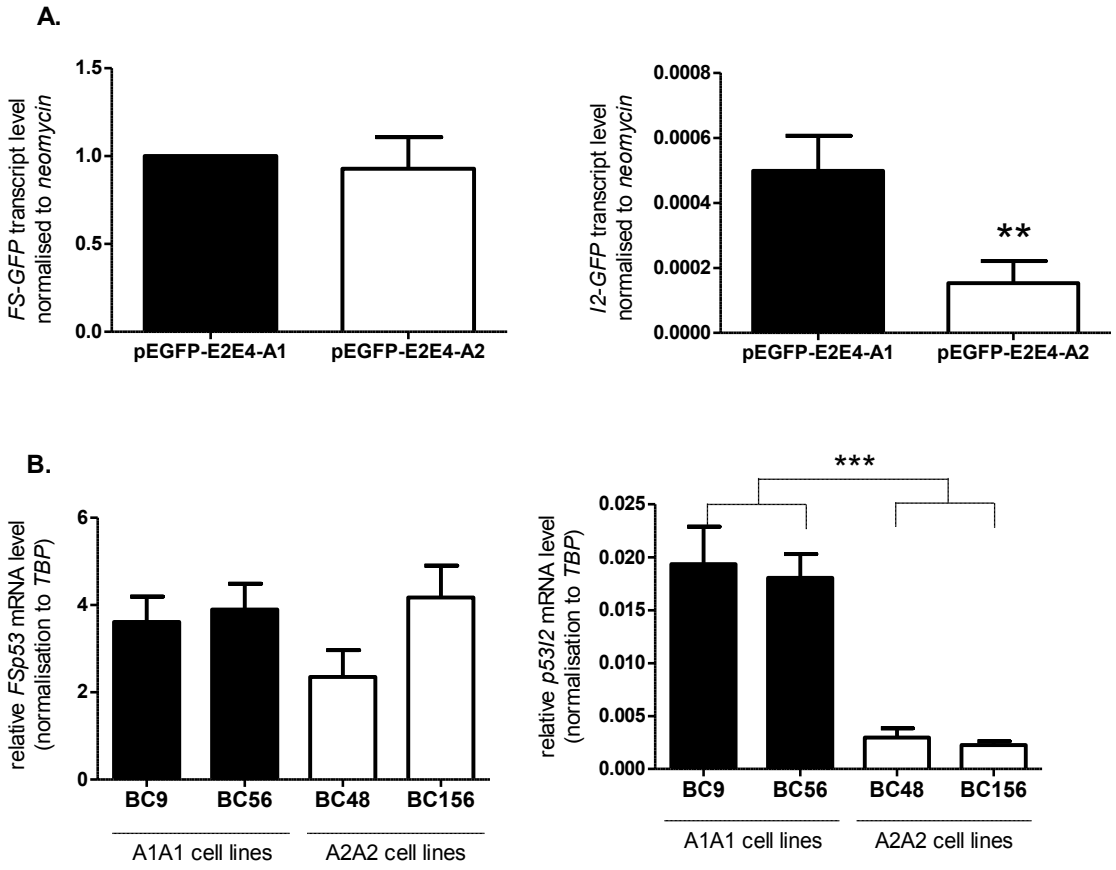

A.
FS-GFP transcript level normalised to neomycin
1.5
1.0
0.5
0.0
pEGFP-E2E4-A1
pEGFP-E2E4-A2
I2-GFP transcript level normalised to neomycin
0.0008
0.0006
0.0004
0.0002
0.0000
**
pEGFP-E2E4-A1
pEGFP-E2E4-A2
B.
relative FSp53 mRNA level (normalisation to TBP)
6
4
2
0
BC9
BC56
BC48
BC156
A1A1 cell lines
A2A2 cell lines
***
relative p53I2 mRNA level (normalisation to TBP)
0.025
0.020
0.015
0.010
0.005
0.000
BC9
BC56
BC48
BC156
A1A1 cell lines
A2A2 cell lines

Fig 3

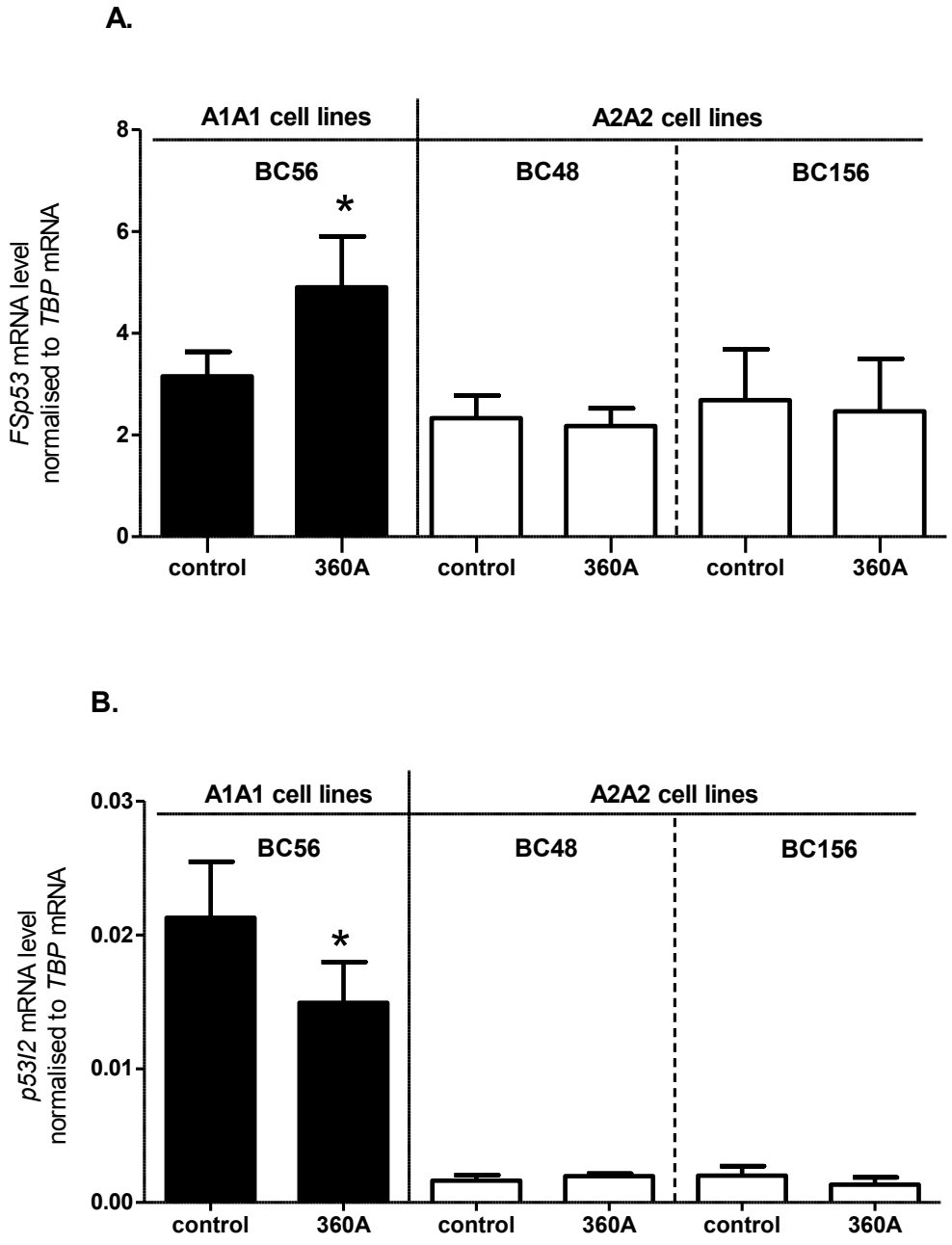

A.
A1A1 cell lines
A2A2 cell lines
BC56
BC48
BC156
*
FSp53 mRNA level
normalised to TBP mRNA
8
6
4
2
0
control
360A
control
360A
control
360A
B.
A1A1 cell lines
A2A2 cell lines
BC56
BC48
BC156
*
p53I2 mRNA level
normalised to TBP mRNA
0.03
0.02
0.01
0.00
control
360A
control
360A
control
360A

Fig 4

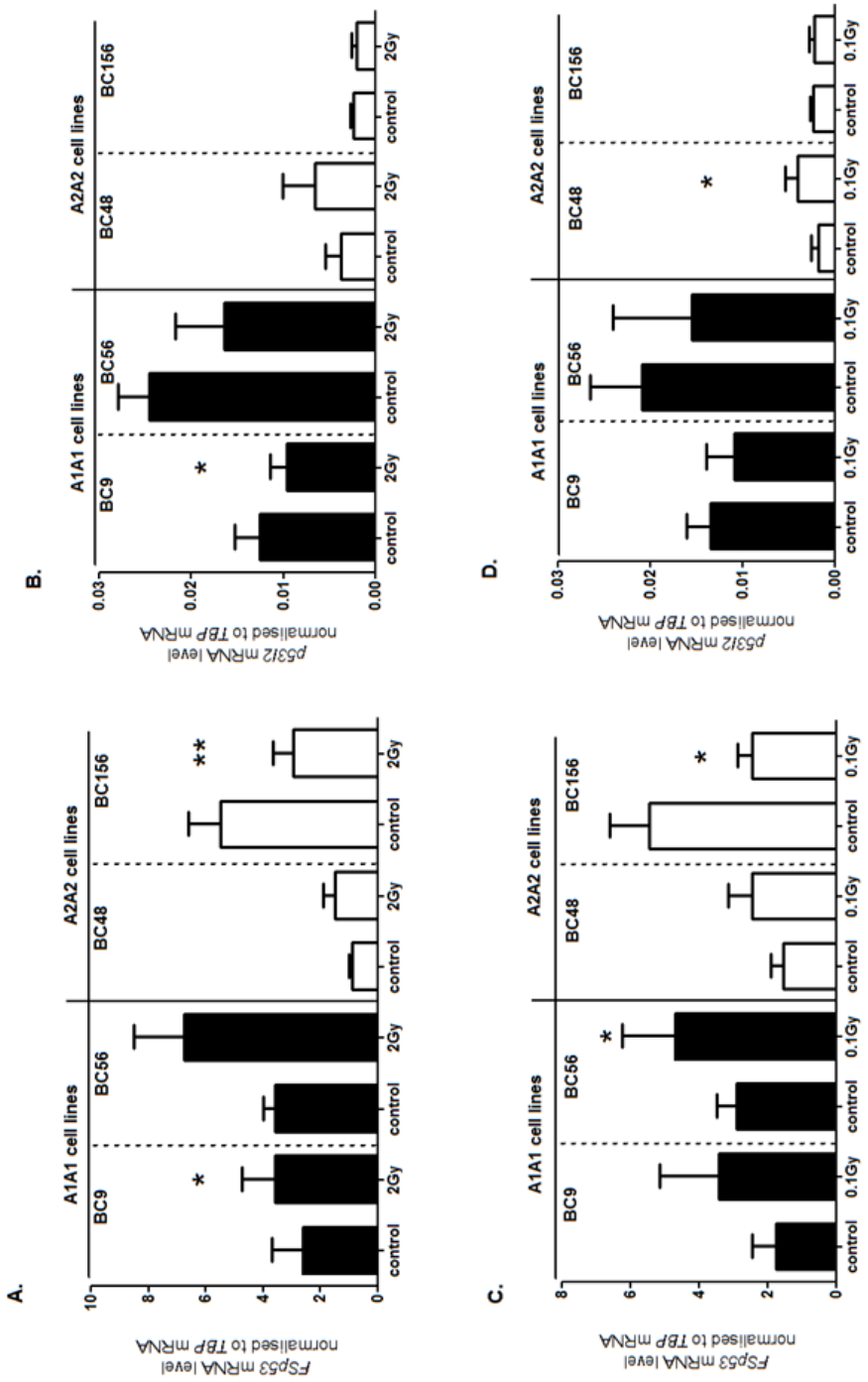

A.
B.
C.
D.
A1A1 cell lines
A2A2 cell lines
BC9
BC56
BC48
BC156
control
2Gy
0.1Gy
FSp53 mRNA level normalised to TBP mRNA
p53I2 mRNA level normalised to TBP mRNA
*
**

Fig 5

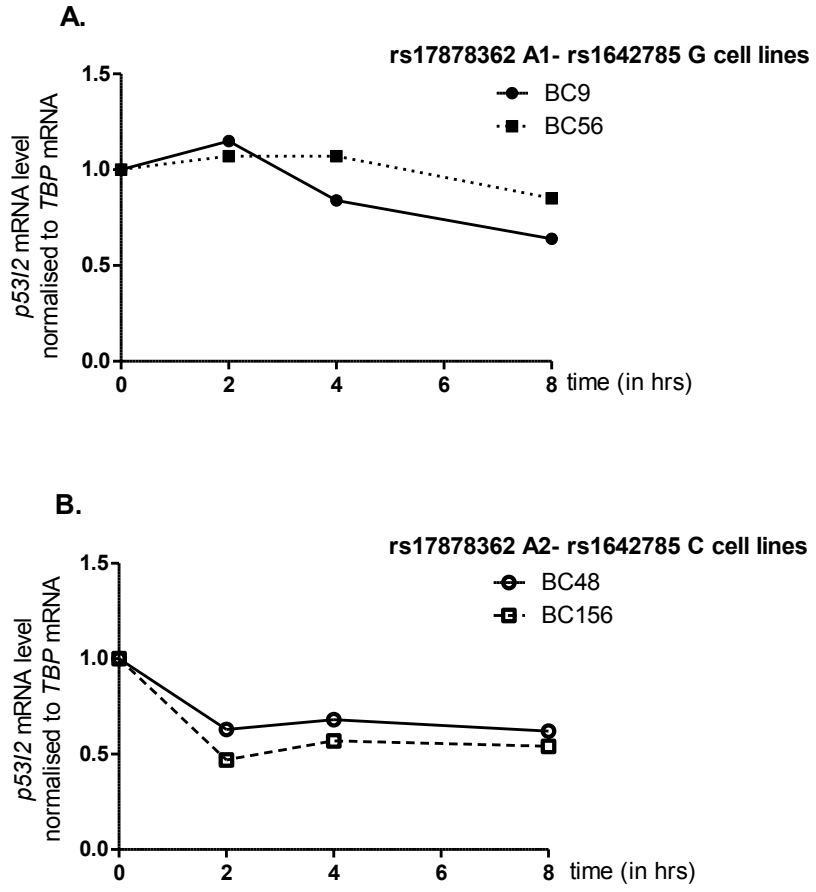

A.
rs17878362 A1- rs1642785 G cell lines
BC9
BC56
p53/2 mRNA level normalised to TBP mRNA
1.5
1.0
0.5
0.0
0
2
4
6
8
time (in hrs)
B.
rs17878362 A2- rs1642785 C cell lines
BC48
BC156
p53/2 mRNA level normalised to TBP mRNA
1.5
1.0
0.5
0.0
0
2
4
6
8
time (in hrs)

Fig 6

Supplementary materials

Impact of G-quadruplex structures and the polymorphisms rs17878362 and rs1642785 on the expression of *TP53* transcripts

Laury Perriaud, Virginie Marcel, Charlotte Sagne, Vincent Favaudon, Aurore Guédin, Marie-Paule Teulade-Fichou, Pierre Hainaut, Jean-Louis Mergny, Janet Hall

Supplementary Table 1

Transcripts	Primers	Sequences (5'-3')	Ref
FSp53	Forward	TGGAAACTACTTCCTGAAAACAACG	New design
	Reverse	GGGAGTACGTGCAAGTCACAGA	
p53I2	Forward	TGACACGCTTCCCTGGAT	New design
	Reverse	TCGCTTCCCACAGGTCTC	
TBP	Forward	CACGAACCACGGCACTGATT	Marcel et al., 2011
	Reverse	TTTTCTTGCTGCCAGTCTGGAC	
FS-GFP	Forward	TGGAAACTACTTCCTGAAAACAACG	New design
	Reverse	GCTCACCGATCGTGCTTGGGAA	Marcel et al. 2011
I2-GFP	Forward	CAGTCAGATCCTAGCGTCGAG	New desgin
	Reverse	AGCAGAAAGTCAGTCCCATGA	
Neomycin	Forward	AGACAATCGGCTGCTCTGAT	Marcel et al., 2011
	Reverse	CAATAGCAGCCAGTCCCTTC	

Supplementary Table 2

			Prediction of TP53 A2-intron 3			
Algorithm	**Consensus**	**Reference**	**Start**	**Length**	**Sequence**	**Score**
Quadfinder	$G_xN_{y1}G_xN_{y2}G_xN_{y3}G_x$ (3 ≤ x ≤ 5) (1 ≤ y ≤ 7)	Scaria et *al*., 2006	11	29	GGGTTGGGCTGGG**ACCTGGAGGGCTGGGG**	
			16	24	GGGCTGGG**ACCTGGAGGGCTGGGG**	
			21	29	GGG**ACCTGGAGGGCTGGGGACCTGGAGGG**	
			31	28	**GGGCTGGGGACCTGGAGGGCTGGGG**GGG	
			36	31	**GGGGACCTGGAGGGCTGGGG**GGGCTGGGGGG	
			37	30	**GGGACCTGGAGGGCTGGGG**GGGCTGGGGGG	
			47	20	**GGGCTGGGG**GGGCTGGGGGG	
GQRS Mapper	$G_xN_{y1}G_xN_{y2}G_xN_{y3}G_x$ (x ≥ 3) (1 ≤ y ≤ 7) (length ≤ 35) (score ≥ 48)	Kikin et *al*., 2006	11	29	GGGTTGGGCTGGG**ACCTGGAGGGCTGGGG**	48
			16	24	GGGCTGGG**ACCTGGAGGGCTGGGG**	47
			21	29	GGG**ACCTGGAGGGCTGGGGACCTGGAGGG**	48
			31	28	**GGGCTGGGGACCTGGAGGGCTGGGG**GGG	48
			37	30	**GGGACCTGGAGGGCTGGGG**GGGCTGGGGGG	50
			47	20	**GGGCTGGGG**GGGCTGGGGGG	48

Bold sequence: 1st single repeat of the 16 bp duplication

Bold underlined sequence: 2nd repeat of the 16 bp duplication

DISCUSSION: *TP53* POLYMORPHISMS, KEY MODULATORS OF P53 CANCER SUSCEPTIBILITY

The human genome contains more than one million sequence variants that are considered as SNPs/polymorphisms, consisting of DNA sequence variations affecting single nucleotides. It is common to include in this definition sequence variations that affect a small number of consecutive base pairs (small insertions, deletions, inversions, tandem substitutions…). SNPs/polymorphisms are heritable as Mendelian traits and represent one of the main genetic bases for the diversity within a species and are thus important targets for evolution.

The human *TP53* gene is highly polymorphic. Some of these polymorphisms may represent variations without functional relevance that occur as markers of genetic distance between human population groups. However, it is likely that several other polymorphisms may have a functional impact and contribute to individual and population differences in the intrinsic level of p53 activity. Such "functional polymorphisms" would be expected to have an effect on both genetic susceptibility and on cancer development. To date, most studies on *TP53* polymorphisms have focused on the small sub-group of polymorphisms that occur in exons and lead to an aa substitution. Indeed, the biological effects of such non-silent polymorphisms are relatively easy to understand, since they may have a very direct impact on protein function. The typical example is rs1042522, a G to C variation at codon 72, specifying either R or P at the corresponding position in the human p53 protein. This polymorphism has been the focus of close to 750 publications indexed in PubMed. Its success as a genetic marker is due to its frequency in human populations (the minor allele, C (encoding P) is present at a frequency of 0.2 to 0.5, depending upon the population) and to the demonstration of functional differences between the R and P p53 proteins *in vitro* and in experimental animals (Whibley et al 2009). Several studies have shown that the 72P variant has a higher capacity than 72R in trans-activating $p21^{WAF1}$ and in inducing growth arrest (Pim and Banks 2004, Salvioli et al 2005). In addition, the 72R variant has an increased capacity to relocate to the mitochondria in tumour cell lines (Dumont et al 2003). Furthermore, the 72R variant is associated with a higher expression of the embryonic implantation factor LIF and lower rate of *in vitro* fertilization failure (Feng et al 2011, Kang et al 2009). An association has also been reported between low winter temperature and the 72R variant in a cohort of 4,029 individuals across Eastern Asia, suggesting selection by adaptation to cold climates (Shi et al 2009b). Compared to this polymorphism, those in introns have been little studied, despite the fact that some of them are also as common in various human populations. Recent advances in the understanding of the role of introns, as well as knowledge on the nature of regulatory motifs affecting RNA transcription, processing and stability in non-coding DNA sequences, is

currently leading the scientific community to re-assess the possible impact of these so far neglected polymorphisms.

The main topic of my Thesis was to understand the biological and pathological impact of rs17878362. This polymorphism consists of a duplication of 16 bp in a G-rich region of *TP53* intron 3. It is located in a portion of the *TP53* gene, which has a complex structure and is highly polymorphic, made of a rapid succession of small exons and introns, and subject to the alternative splicing of intron 2. Within this region, 3 of the most frequent polymorphisms in *TP53* can be found over a distance of 412 bp: rs1642785 in intron 2 (G<C), rs17878362 in intron 3 and rs1042522 in exon 4. These 3 polymorphisms are in partial or near-complete linkage disequilibrium, defining a particular "box" within the *TP53* gene, which may affect the regulation of its expression and function (**Figure 23**) (Garritano et al 2010, Marcel et al 2009).

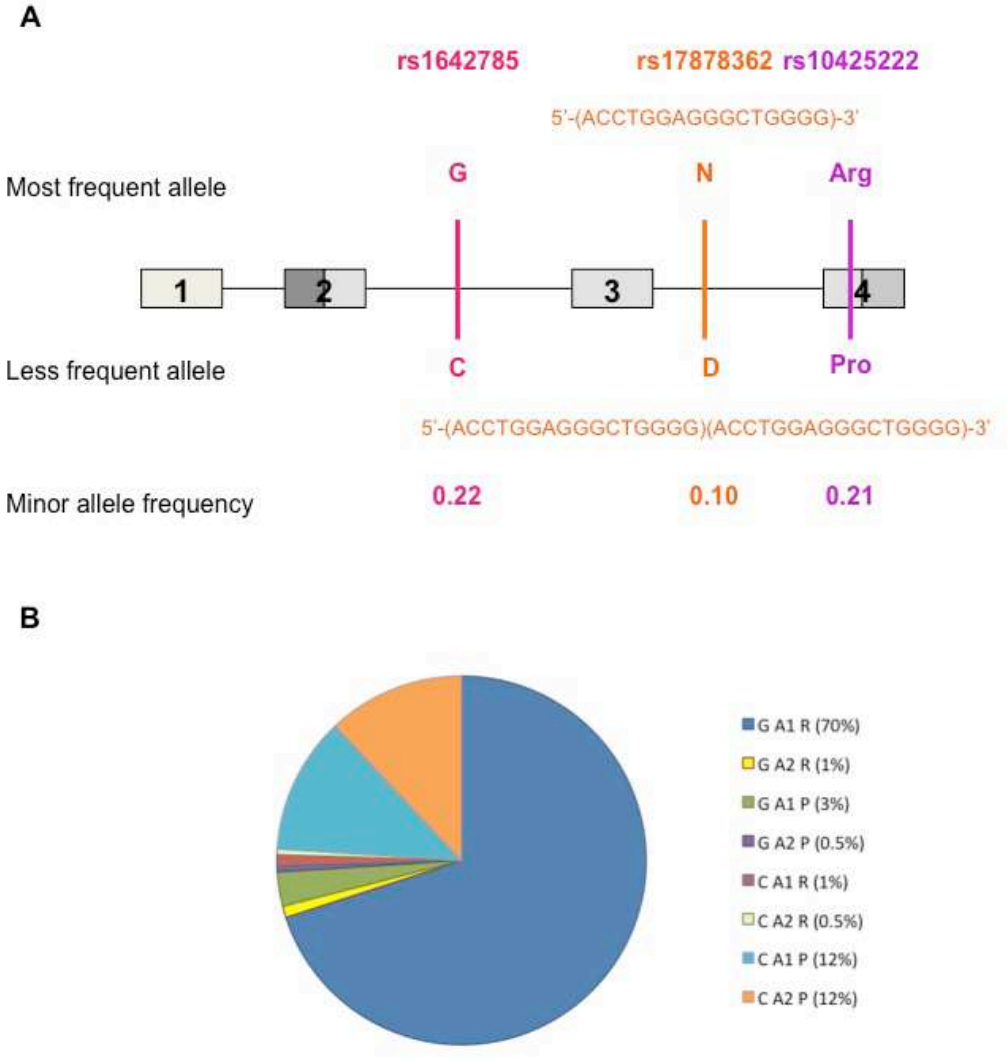


Figure 23: Schematic representation of the three most frequent polymorphisms between the intron 2 and the exon 4 of the *TP53* gene. (A) The localisation of rs1642785, rs17878362 and rs1042522 within *TP53*, their minor allele frequency and (**B**) the distribution of their haplotypes from (Marcel et al 2009).

To address the possible impact of rs17878362, we first performed a meta-analysis of a large dataset of case-controls studies, which analysed the association between rs17878362 and cancer susceptibility. Next, we analysed the effect of this polymorphism in the context of the high cancer risk of subjects carrying a germline *TP53* mutation. Finally, we have integrated structural and functional studies *in vitro* and in cell lines models to demonstrate that rs17878362 has a specific effect on p53 gene expression and regulation by modulating the alternative splicing of intron 2.

Part I. Association between rs17878362 and cancer susceptibility

A. Meta-analysis of rs17878362 in relation with risk of sporadic cancer

Meta-analysis provides a methodology to statistically assess the strength of an association across a series of studies having a similar overall design. One of its main advantages is to take into consideration various sources of bias (such as biases due to small sample size, to publication trends) while taking into account each specific study in the overall assessment. This approach differs from pooled analyses, a method, which consists in assembling a single association study by pooling together the data from different studies. For **Article 1**, we have taken into consideration all studies (1) published in English before January 2012, (2) developed using a strict case-control design, (3) in which rs17878362 has been typed using a reliable methodology (DNA sequencing, allele-specific TaqMan assays, differential gel electrophoresis migration assays), (4) providing their results in the form of an odds ratio with 95% confidence interval (**Figure 24**). Our meta-analysis regrouped 25 such studies, comprising 10,786 cancer cases and 11,377 controls. This is the third meta-analysis of the association between rs17878362 and cancer risk. Previous meta-analyses were published in 2010 ((Hu et al 2010c); 1,823 cases and 2,111 controls focusing on breast cancer, (Hu et al 2010b); 9,801 cases and 10,391 controls) and in 2011 ((He et al 2011a, He et al 2011b); 3,332 cases and 3,400 controls, focusing on breast cancer). However, the 2010 meta-analysis has been the focus of controversy, since some of the data used in this analysis were not in line with those published in the original studies (Lu et al 2011).

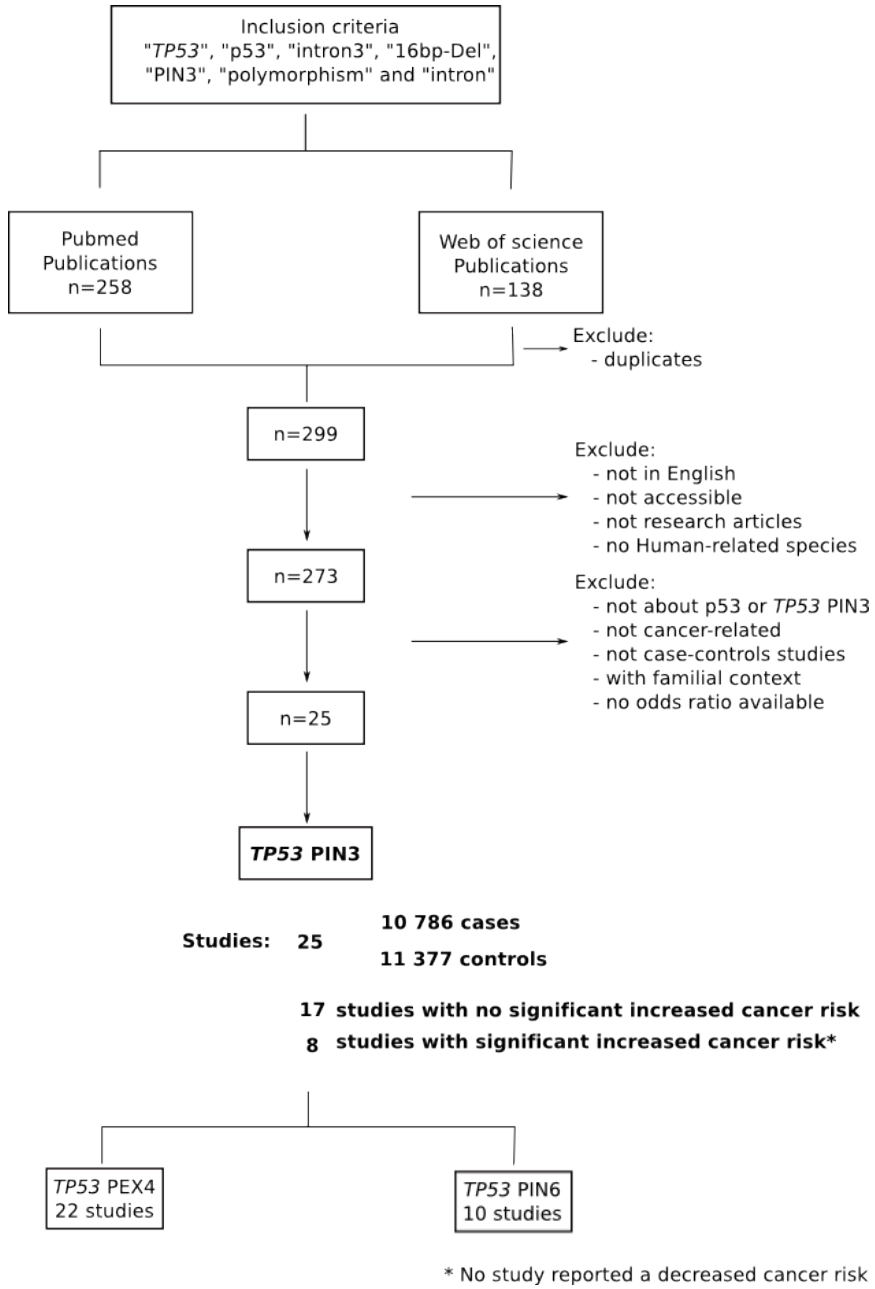


Figure 24: Decision tree applied to select publications using PubMed and Web of Science databases. Using the predefined terms ("*TP53*", "p53", "intron 3", "16bp-Del", "PIN3", "polymorphism", "rs17878362" and "intron"), several publications were found in both PubMed and Web of Science databases. To identify the publications suitable for our meta-analysis, exclusion criteria were applied and are presented on the right panel. Twenty-four studies focused on *TP53* PIN3 polymorphism were included, some of them also presenting data on *TP53* PEX4 and PIN6 polymorphisms. Main characteristics of the selected publications are presented at the bottom panel. n: number of publications, cases and controls.

This controversy led us to perform a rigorous meta-analysis to quantify the association between rs17878362 and cancer susceptibility and to assess this association in different population according to their geographical origin and in different types of cancer.

Our results confirm previous meta-analyses identifying a small but statistically significant association between the A2 allele of rs17878362 and the risk of cancer. Depending upon the cancer pathology considered, the OR in homozygote carriers of the A2 allele varied

between 1.41 and 1.67, suggesting an overall increase in risk of about 50%. Differences were noted between breast, colorectal and lung cancer, which may be due to the heterogeneous nature of these broad cancer types (for example, breast cancer is known to contain several sub-types; lung cancer consists of different histo-pathological types, which also differ in their association with smoking) (**Figure 25**). Furthermore, differences were detected between populations. The strongest association was seen for European Mediterranean countries (A2A2 genotype: OR= 2.52, 95% CI=[1.53-4.24]) whereas no significant association was observed in studies on populations from the US. The reason for such large difference between two populations of essentially Caucasian origin is not known. It may be due to a joint effect of rs17878362 with lifestyle and environment, which may differ among populations and regions. Other genetic modifiers may also play a role, although there is no clear clue on which modifier might significantly differ among these populations of similar origin. Of note, our study did not include Asian populations. Indeed, in this group, the A2 allele of rs17878362 is rare (MAF=2%) and there is not enough structured data to evaluate the impact of this polymorphism in Asian populations (Marcel et al 2009).

Whereas the meta-analysis clearly identifies an increased cancer risk in relation with A2A2 genotype, the strength of the association for individuals with a heterozygous carrier status is not statistically significant (except in Mediterranean countries). When comparing different cancer types, however, there is a small, non-significant increase of the risk in A1A2 carriers, which is compatible with a dosage effect for the impact of the A2 allele for an increased risk of cancer. Further studies on cohorts with better-defined risk factor profiles will be needed to identify and quantify the precise risk associated with A1A2 genotype.

As mentioned above, rs17878362 is in partial LD with rs1042522 (as well as with rs1642785). Using the same case-control studies as for rs17878362, we have meta-analysed the risk of cancer in relation to the rs1042522 genotypes (**Figure 25**). Our analysis detected a significant increase of risk of 16% in heterozygote carriers (G/C). A similar effect was seen with another intronic polymorphism, rs1625895 (G>A; located in intron 6). These two polymorphisms belong to the same *TP53* haplotype block as rs17878362 (Garritano et al 2010). It remains to be determined up to which point their effects (which are weaker than the one of rs17878362) can be attributed to the LD with rs17878362. Current published data on individual polymorphisms do not allow the reconstitution of haplotypes on groups of cases and controls large enough to address this question.

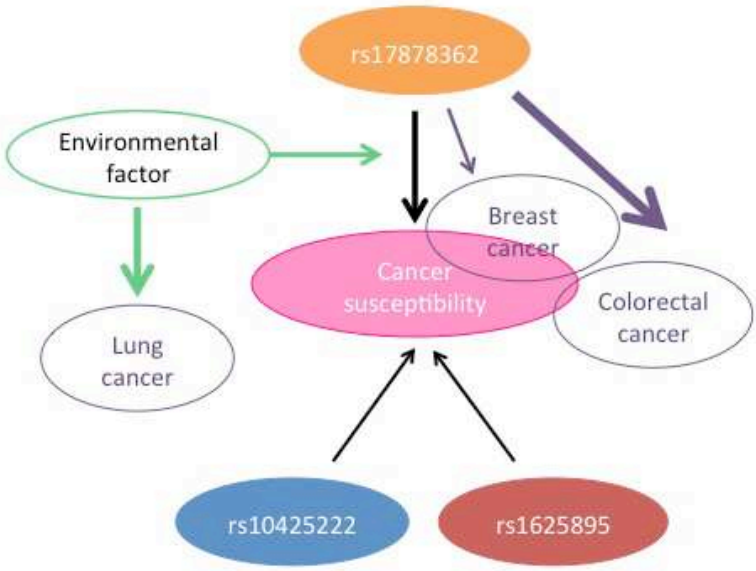


Figure 25: Effect of the rs17878362, rs1042522 and rs1625895 polymorphisms on cancer susceptibility depending on geographic origin and cancer type: a model. The size of the arrow corresponds to the statistical significant association between cancer susceptibility and rs17878362.

Overall, the effect we have detected in this meta-analysis predicts that the A2 genotype is associated with a slightly increased risk of cancer. This would suggest that A2 defines a "weak" *TP53* allele, with less capability to suppress tumorigenesis than alleles carrying the A1 genotype. Another hypothesis, however, can be formulated: subjects with A2 alleles might be more susceptible to somatic inactivation of *TP53* by mutation or loss of heterozygosity, thus leading to an increased risk of cancer in particular in A2A2 carriers. There is currently no data on the association between the possible selective occurrence of *TP53* mutations on defined haplotypes and cancer risk. In a recent study, Ortiz-Cuaran and collaborators have shown that polymorphisms in intron 1, but not rs17878362, were associated with *TP53* mutations at codon 249 in hepatocellular carcinoma (Ortiz-Cuaran et al 2013). In another study, Mechanic and co-workers analysed 14 polymorphisms in lung cancer patients from the greater Baltimore area and have observed an association between the *TP53* polymorphisms rs1042522, rs9895829 (A>G, intron 4), rs1625895 and rs12951053 (T>G, intron 7) and risk of somatic mutation (Mechanic et al 2007). It remains to be determined whether *TP53* alleles carrying the A2 genotype of rs17878362 are more frequently mutated in sporadic cancers than alleles carrying the A1 genotype.

B. Analysis of rs17878362 in relation with risk of familial cancer

The LFS and its variant form LFL are rare autosomal syndromes associated with a specific spectrum of early onset cancer. Currently, the only known genetic basis is carriage of a germline mutation in *TP53*. These mutations are found in 20-30% of the subjects referred for genetic testing based on clinical criteria suggestive of LFS/LFL (modified Chompret criteria) (Tinat et al 2009). The molecular basis of the syndrome in subjects who do not carry a germline *TP53* mutation is not known. Moreover, there is debate as whether all *TP53* mutations are equally penetrant and have the same impact on cancer risk. It is likely that other genetic/epigenetic traits could modify the penetrance of *TP53* mutations to this disease such as *TP53* polymorphisms (such as rs1042522) or *Mdm2* polymorphisms (such as rs2279744). We have addressed the possible impact of rs17878362 on the phenotype of LFS in *TP53* mutation carriers from Brazil, a country where LFS/LFL appears to be more frequent than in other parts of the world, an observation attributable to the high prevalence of a founder *TP53* mutation (R337H) in the population of south Brazil (**Article 2**).

In this study, we have tested our working hypothesis that at least one of the 85 polymorphisms located within or around the 5 regions containing G4 structures of the *TP53* gene may play a role as a genetic modifier of the penetrance of the pattern of cancer incidence in the LFS family members. Therefore, we genotyped these polymorphisms in the LFS subjects and we analysed the association between the carriage of the different alleles with age of cancer onset. For the statistical analysis, we have developed a hierarchic model taking into account the number of tested subjects in each family. We adjusted each sample to the size of its family to suppress the bias induced by large pedigrees.

In LFS/LFL patients, we found that only 11 of the 85 polymorphisms showed allelic variation in these series. Four had a minor allele frequency (MAF) less than 4% including rs78378222 that has been associated with basal cell carcinoma (Prostate cancer, glioma and colorectal adenoma) in an Iceland population . Focusing on these 7 polymorphisms, we analysed their association with the age of first cancer diagnosis. None of these 7 polymorphisms was associated with an increase of cancer risk in the WT2 group. In contrast, in the MUT group, the minor alleles of rs17878362, located in the G4 in the intron 3, and of rs17880560, located in proximity to the G4 in the 3'flanking region, showed a borderline significant tendency for association with a later age at first cancer diagnosis.

To complement this, we reconstituted haplotypes based on the rs17878362 and rs17880560 polymorphism. We did not observe any association of these haplotypes and cancer risk in the WT2 group. In contrast, in the MUT group, A1-A1 (A1 for rs17878362 and A1 for rs17880560) individuals developed cancer on average 25 years earlier than patients with haplotypes carrying at least one A2 allele (A1-A2: A1 for rs17878362 and A2 for rs17880560) and (A2-A1: A2 for rs17878362 and A1 for rs17880560). In addition, only A1-A1 carriers developed cancer before the age of 35 years.

Our results indicate that the A2 alleles of the rs17878362 and rs17880560 polymorphisms appear to be protective against early cancer onset. Their combination is more effective for protecting against early cancer onset than any of these polymorphisms alone. These results suggest that polymorphisms located within or around G4 structures could modulate on *TP53* suppressor functions. These results suggest that *TP53* haplotypes carrying at least one A2 allele of rs17878362 or rs17880560 can exert their suppressor activity even in the face of a mutant A1-A1 haplotype, thus enabling a protection against cancer onset until adult age. These WT haplotypes would therefore qualify as "strong" haplotypes, in contrast to WT A1-A1 haplotypes, which would be less competent to prevent cancer onset, leading to higher risk of childhood cancer (**Figure 26**). These observations, if validated, would be of interest to stratify patients into groups at risk for "early" *versus* "late" diagnosis, each with adapted surveillance programs. Such an approach may be of great interest for the surveillance of cancer risk in Brazilian carriers of R337H, since many of these subjects may have a low risk of early cancer and may therefore not need to undergo the intense surveillance programs proposed to carriers of "classical" *TP53* mutations (Villani et al 2011).

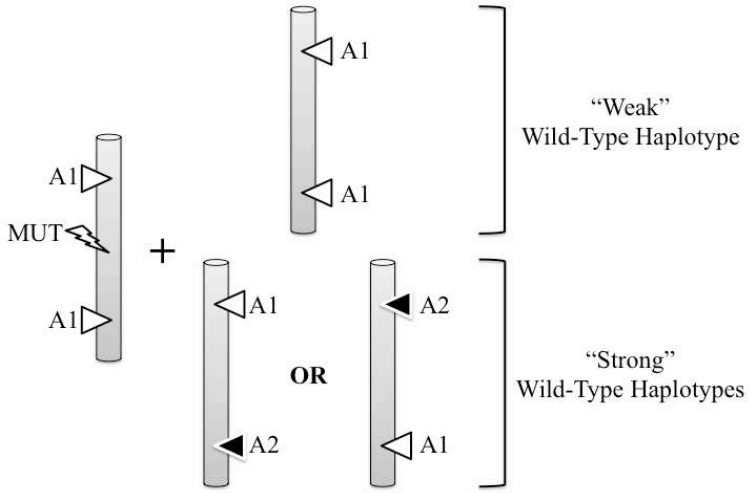


Figure 26: Effect of different WT haplotypes in *TP53* mutation carriers: a model. *TP53* alleles are represented as rods. Left, mutant allele occurring on a haplotype carrying A1 variants of both rs17878362 and rs17880560 (A1A1). Right, different types of WT haplotypes. The WT haplotype defined by A1A1 is considered as a "weak" haplotype (associated with early cancer, indicative low capacity to compensate the loss of p53 function of the mutant allele). The WT haplotypes defined by A1A2 or A2A1 are considered as "strong haplotype" (associated with later cancer onset, thus providing at least partial compensation for the loss of function of the mutant allele). Of note, our data do not predict the effect of WT A2A2 haplotypes, or the effects of these haplotypes when the mutation occurs on another haplotype than A1A1. (Sagne et al., Submitted)

The fact that our series of patients included a large proportion of carriers of the "Brazilian founder mutation" R337H, can be seen as both a constraint and an advantage. It is a constraint since it has to be considered that the effect we have observed might be specific for this particular mutant. It also represents an advantage, since it led us to consider the effects of polymorphisms on the remaining WT allele. Indeed, in R337H carriers, all subjects have exactly the same, mutant haplotype and genetic diversity in *TP53* is due to different polymorphisms occurring on the WT allele. This allele-specific aspect has not been taken into account in studies on *TP53* polymorphism in other LFS/LFL cohorts, in which *TP53* polymorphisms were analysed independently of their occurrence on either the WT or the mutant haplotype. Another consideration is that it is known that R337H carriers have a lower risk of developing early cancer (before age 30) than carriers of "classical", DNA-binding domain *TP53* mutations. Therefore, the age range of disease development in these patients is particularly large, providing a wider "window" for detecting subtle effects of genetic variations. Thus, it will be important to replicate these studies in other, independent LFS/LFL series.

C. Effects of rs17878362 in sporadic or inherited contexts: an apparent paradox

Our conclusions on the role of rs17878362 in sporadic cancers (Meta-analysis: **Article 1**) and in the context of inherited cancers (LFS cohort: **Article 2**) are resulting into an apparent paradox. Whereas, in sporadic cancer, carriers of two A2 alleles appear to have an increased risk of cancer, in inherited cancers, we have identified this allele as "protective" and appear to be associated with the development of cancer at a latter age. It therefore seems that the definition of "weak" and "strong" p53 tumour suppressor alleles is not the same, depending upon whether the context is sporadic or inherited. There is, however, a plausible hypothesis to resolve this apparent paradox (**Figure 26**).

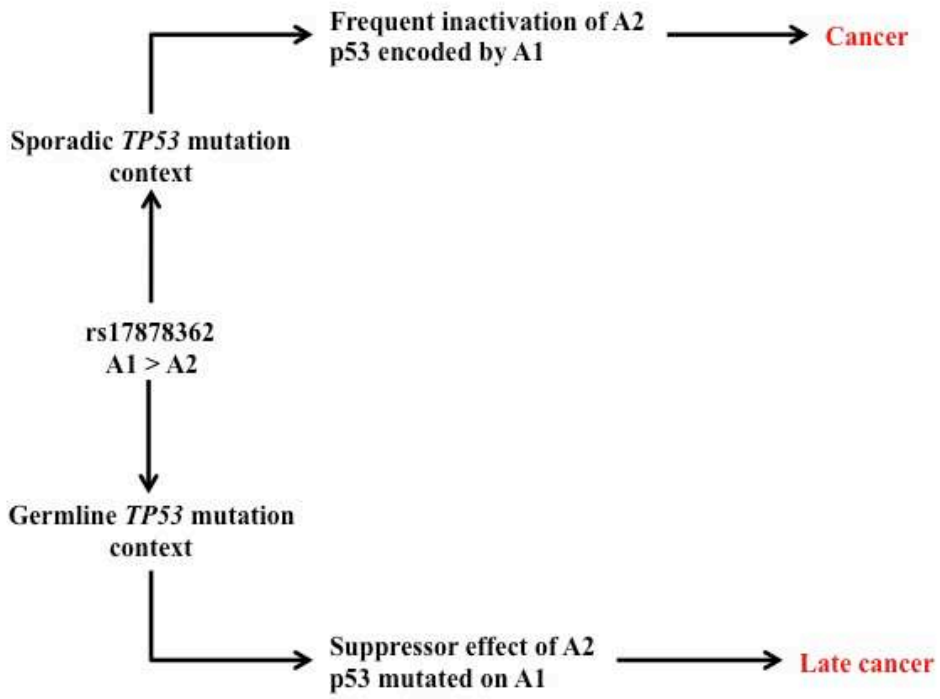


Figure 27: Model of the association of the *TP53* rs17878362 polymorphism in sporadic and germline *TP53* mutation context.

In the context of sporadic cancer, it can be proposed that subjects with A2 alleles of rs17878362 may indeed carry a form of p53 that exerts potent effects as a tumour suppressor. This effect would identify *TP53* in these subjects as a strong limiting factor for the occurrence and development of cancer. Consequently, in these tumours *TP53* would be an important, if not an obligate, target for genetic alterations (mutations or loss of alleles). The notion that *TP53* mutations may preferentially appear on specific alleles has been described in some

studies. In liver cancers, Ortiz-Cuaran and collaborators have identified an association between the presence of the mutation at codon 249 in *TP53* and specific genotypes of 4 polymorphisms located in *TP53* intron 1 (Ortiz-Cuaran et al 2013). In sporadic lung cancers, Mechanic and co-workers have shown that mutations were more frequent on alleles of *TP53* carrying a number of defined polymorphisms (Mechanic et al 2007). However, no study to date has investigated whether different alleles of rs17878362 may be mutated at different rates in sporadic cancers. The hypothesis that haplotypes carrying the A2 allele of rs17878362 may be more "mutable" than alleles carrying the A1 allele may contribute to explain why subjects with A2 alleles may be at higher risk of sporadic cancer.

According to this hypothesis, subjects with inherited mutations on A2 alleles would be expected to have a more severe phenotype than subjects with inherited mutations on A1 alleles. Our dataset of LFS/LFL families is too limited to investigate this question. We could only address it in carriers of other mutations than R337H (which is present on the A1 allele of rs17878362). In families with other mutations, we observed that all but one (15 out of 16) had the germline mutation on the A1 allele of rs17878362. The only subject who carried a *TP53* mutation on an A2 allele also carried a WT A2 allele, and this patient is the only case in our series with early cancer (<25 years) and carrying at least one A2 allele. None of the patients carried a mutant A2 allele and a WT A1 allele. Although we cannot conclude from a single case, we note that, according to our hypothesis, this patient would carry the mutation on a "strong allele" while retaining a "strong" WT allele, a situation in which the intrinsic effects of the two alleles may cancel each other, leading to a phenotype similar to carriers of mutations in an A1A1 background. Further studies in other LFS/LFL series are needed to analyse cancer onset according to the polymorphic status of haplotypes carrying the mutation, and to determine if different *TP53* haplotype are mutated at the same rate in the germline of *TP53* families. In addition, this hypothesis can serve as a basis for experimental strategies aimed at measuring the activity of selected *TP53* haplotypes, for example using a panel of lymphoblastoid cell lines derived from subjects with defined *TP53* genotypes. An interesting experimental model for such studies has been recently described by Zerdoumi and collaborators (Zerdoumi et al 2013). These authors have developed a functional assay of p53 pathway induction after DNA damage in Epstein-Barr virus immortalized lymphocytes, followed by the comparison of gene-expression profiling. In WT cells, they identified 173 genes with an increase of more than two-fold. On these 173 genes, 46 were known as p53 target genes. In LFS cells with "canonical" missense mutations, numbers and levels of induced genes were strongly reduced as compared with controls and LFS cells with null-p53

mutations. Thus, this model seems to be appropriate to score germline missense *TP53* mutations with a dominant negative effect. It would be of great interest to compare the variations in the expression of the 173 gene signatures in lymphoblastoid cell lines with *TP53* mutations in different rs17878362 backgrounds.

Part II. Towards a functional hypothesis for genetic determinants of p53 regulation

G4s are important structural motifs in the regulation of gene expression. Their presence in the pre-mRNA of several genes regulates its stability, splicing and processing. The two polymorphisms identified here as associated with age at cancer onset in LFS/LFL have been associated with differences in p53 mRNA levels. Regarding rs17878362, Gemignani and collaborators have detected that lymphoblastoid cells from subjects homozygous for the A2 allele expressed significantly less total p53 mRNA than subjects homozygous for the A1 genotype (Gemignani et al 2004). Regarding rs17880560, this polymorphism is located in the 3'flanking region that has recently been shown to contain G4 sequences regulating pre-mRNA cleavage at the poly-adenylation site (Decorsiere et al 2011). Although rs17880560 is not precisely located within the main G4 structure identified as responsible for this effect, it affects another putative G4 located closed to this demonstrated G4. Given the possibility that the G4s located over a stretch of bases may be in equilibrium with each other, it is possible that a polymorphism modifying one G4 within a defined region may have a domino effect on the structure of other G4s.

Although located in an intron, rs17878362 is located in an important region for the regulation of the *TP53* gene expression. This region contains the internal promoter P2, responsible for the transcription of the Δ1133p53 and Δ160p53 isoforms and the intron 2, which is alternative spliced, modulating the expression of the Δ40p53 isoform (Marcel et al 2011). The pre-mRNA from the proximal promoter generates two different N-terminal isoforms: FSp53 mRNA, which encode TAp53 (the canonical p53 protein, 393 aas), and p53I2 mRNA, retaining intron 2, which encodes Δ40p53, an isoform of 354 aas produced due to initiation of transcription at AUG40. Retention of intron 2 introduces several stop codons downstream of codon 1 located in exon 2, making AUG40 the first available initiation codon for p53 protein translation from the p53I2 mRNA. Of note, a similar Δ40p53 protein can also be expressed by internal initiation of translation at codon 40 using FSp53. It is currently not possible to distinguish between the Δ40p53 proteins based on whether they have been translated from p53I2 or from internal initiation in FSp53 mRNA.

Examination of the sequence of intron 3 led Eric Van Dick, then a staff scientist at IARC, to speculate that this region might have a structure compatible with formation of a G4. This observation led our interest in identifying this structure and elucidating its possible impact on the regulation of p53 expression. **Article 3** reports that G4 structures are formed in intron 3 of the *TP53* pre mRNA and enable the regulation of alternative splicing of the intron 2 by increasing the exclusion of intron 2 and the expression of FSp53 mRNA. In **Article 4**, we further show that the G4 structures formed in intron 3 can exist in presence of either rs17878362 variants A1 or A2. Both rs17878362 variants influence the alternative splicing of the intron 2 with the A2 allele being associated with lower levels of the p53I2 mRNA. In addition, the balance between FSp53 and p53I2 mRNA is influenced by an additional polymorphism located in intron 2, which is in partial LD with rs17878362. This polymorphism (rs1642785, G>C) is almost systematically present as the C allele on haplotypes that carry the A2 allele of rs17878362 (**Figure 23B**). Presence of this polymorphism has an effect on the stability of the alternatively spliced mRNA, which appears to have a shorter half-life and is thus present at lower levels than in cells with G alleles of rs1642785 combined with A1 alleles of rs17878362. These results suggest that rs17878362, in combination with rs1642785, influences the formation of mRNA encoding N-terminal isoforms by modulating the alternative splicing of the intron 2 at the pre-mRNA level. This synergistic effect of two polymorphisms in different introns may contribute their linkage disequilibrium.

This complex interplay between polymorphisms, G4 and p53 isoform expression provides support to a simple model in which p53 suppressor function is modulated by the level of Δ40p53 protein, acting as a "buffer" to neutralize p53 protein activity. **Figure 28** illustrates this concept in the context of the two most common haplotype combinations of rs17878362 and rs1642785. With the G-A1 haplotype (representing 70% of all haplotypes in Caucasians), both FSp53 and p53I2 are produced. Assuming that both mRNA are equally translated, two distinct p53 proteins, TAp53 and Δ40p53, should be co-expressed in cells. Given that Δ40p53 is deprived of the N-terminal TAD and can inhibit the capacity of TAp53 to trans-activate target genes, this p53 protein combination may correspond to a "weak" p53 suppressor status, where TAp53 transcriptional activity is partially inhibited by the Δ40p53. With the C-A2 haplotype (representing 12.5% of all haplotypes, the second most common combination), TAp53 is expected to be expressed in the presence of a very low quantity of Δ40p53. This protein combination may correspond to a "strong" p53 suppressor status, since

in this situation TAp53 activity may escape the negative control exerted by Δ40p53. Other haplotypes combining rs17878362 and rs1642785 may provide for a whole spectrum of effects between these two extremes. It remains to be determined whether (1) TAp53 and Δ40p53 are actually expressed at a level predicted by the levels of their respective mRNA; (2) whether levels of Δ40p53 modulate the p53-dependent transcriptome in a manner compatible with a reduction of tumour suppression by p53.

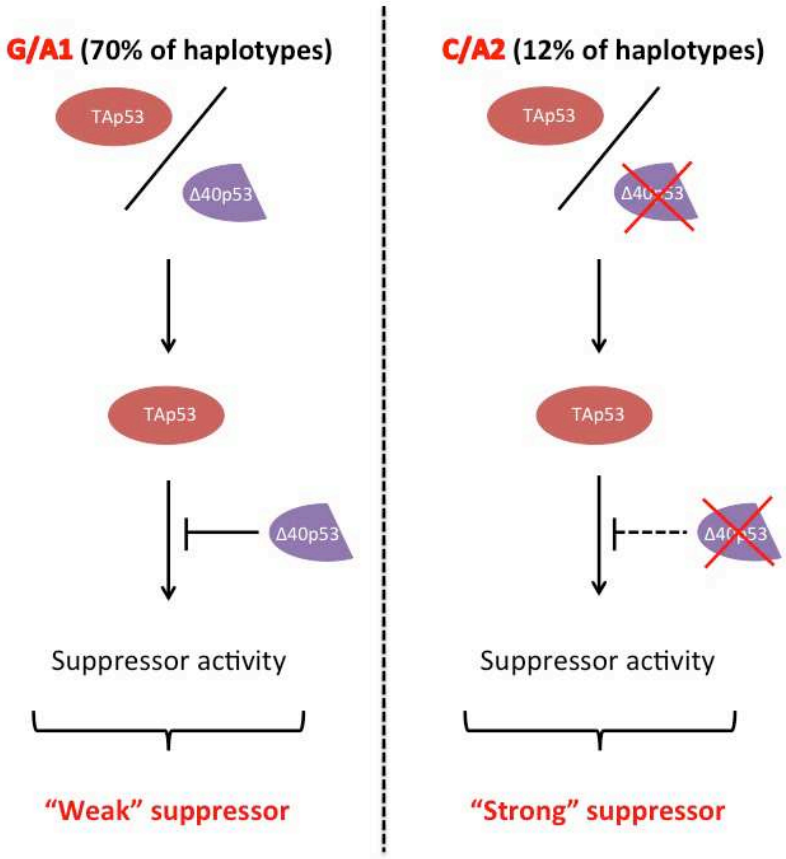


Figure 28: Model of the influence of the combination of the rs1642785 and rs17878362 polymorphisms on N-terminal p53 isoforms expression and p53 activity.

It is acknowledged that the model presented above and in **Figure 28** is extremely speculative, it raises a number of interesting questions. The first is to determine the nature of the biological effects responsible for "weak" and "strong" suppressor status. It is important to consider that Δ40p53, which lacks the Mdm2 binding domain, is not inducible by genotoxic stress in the same way as TAp53. Therefore, it is much more likely to exert its negative effect on p53 function in basal conditions, rather than in conditions were TAp53 is selectively accumulated and activated in response to DNA damage. It follows that the effect of Δ40p53 might be critical for a number of basal p53 functions rather than for the p53-dependent

response to acute DNA damage. Interestingly, a recent study in mice carrying mutants of p53 at acetylation sites has shed a new light on the hypothesis that tumour suppression may be associated with the basal regulation by p53 of a number of genes involved in the control of energy metabolism, rather than with acute induction of cell-cycle arrest or apoptosis (Li et al 2012b). Thus, in future studies, it will be important to examine the basal (non-DNA damage activated) pattern of p53-mediated gene expression in order to identify possible differences in genes involved in the control of energy metabolism.

CONCLUSIONS AND PERSPECTIVES

The aim of this work was to characterise the role of the *TP53* polymorphisms especially the rs17878362 in the regulation of the *TP53* isoforms expression. The hypothesis was that common *TP53* polymorphisms located within introns could have an impact of differential expression of p53 protein isoforms, which could modulate p53 activity. Indeed, results accumulated over the past 10 years have shown that patterns of expression of p53 mRNA and protein were complex, with multiple different mRNA variants and protein isoforms that could influence each other and therefore impact on overall p53 responses.

First, the rs17878362 polymorphism appears to exert different effects on cancer susceptibility depending on whether cancer occurs in a sporadic or familial context. In sporadic cancer, carriage of the A2A2 genotype is associated with an increased risk of cancer compared to the A1A1 genotype, whereas, in inherited cancers, the A2 allele appears to have a "protective" effect, characterized by a later age at first cancer onset in carriers of a germline *TP53* mutation. Of note, in our study, the effect of rs17878362 was due to its presence on the WT allele, since most of the subjects tested had a mutation in an allele carrying the A1 variant. Considering a simple model that assumes that the activity of the mutant allele is lost and that any residual p53 activity in mutation carriers would be driven by the remaining WT allele, these observations suggest that specific *TP53* haplotypes could differ by their capacity to increase the tumour suppression ability. In other terms, haplo-insufficiency in mutation carriers would be dependent upon the haplotype structure, and the biological activity, of the remaining WT allele. This led us to suggest that at least two broad categories of *TP53* haplotypes may exist, "strong" and "weak". It remains unclear how this simple model may account for the fact that the A2 allele of rs17878362 appears to be associated with a higher risk of cancer in a sporadic cancer context. A possible explanation could be that different haplotypes of *TP53* may not be targeted by somatic mutations at the same rate. Haplotypes that specify a "strong" suppressor effect may become more targeted by mutations than haplotypes specifying "weak" suppressor effects, the activity of which may be bypassed by alterations in pathways regulating p53 in the absence of *TP53* mutation. Subjects with A2 alleles may be more "sensitive" to acquire a somatic mutation in *TP53*, and therefore to develop sporadic cancers. There is limited evidence that somatic mutation may not be located at the same rate on different *TP53* haplotypes, but there is no data to support that different haplotypes of rs17878362 are mutated in human cancers. Further studies are needed to examine the associations between rs17878362 haplotypes and *TP53* somatic mutations.

In terms of mechanisms, our work contributes to a growing set of experimental data linking polymorphic G4 structures in *TP53* with regulation of p53 mRNA stability and

processing. Rs17878362 is located in a G-rich region that is demonstrated to form G4s, and that regulates the alternative splicing of p53 mRNA, possibly leading to the differential expression of p53 isoforms. In fact, such effects on p53 mRNA may impact on p53 function just by affecting the levels of mRNA affecting FSp53. In addition to this effect, the production of isoforms capable of interfering with p53 functions may represent a potent mechanism to control and modulate p53 suppressor response, both in response to baseline signals or in response to acute DNA damage.

Secondly, our results show that this polymorphism alone is not the only variations supporting the distinction between "weak" and "strong" alleles. Indeed, it should be considered that these two categories only represent the extreme of a continuous spectrum of *TP53* haplotype defined by several polymorphisms, which may contribute to phenotypic diversity in a coordinated manner. The mechanisms modulated by common polymorphisms in the 412 bp region located between intron 2 and exon 4 are multiple and cover different levels of regulation in a temporal sequence of events. The first polymorphism to exert its effect is rs17878362. This polymorphism affects the processing of p53 pre-mRNA by modulating the selective retention of intron 2 in the final mRNA. Second, rs1642785 may exert an effect at mRNA level by differentially affecting the stability of p53I2 mRNA. Third, another level of differential activity between the products of different *TP53* alleles may be controlled by rs1042522, which specifies either R or P at codon 72 and modulates the activity of the p53 protein. Further functional studies are essential to understand these interactions, therefore shedding new light on the biological mechanisms accounting for the fact that these three polymorphisms have apparently co-evolved in different human populations and lineages.

Finally, it remains to be validated whether the polymorphisms identified in LFS may be of interest as biomarkers to predict individual risk of developing early *versus* late cancers in mutation carriers. This question is critical in the context of the high population frequency of the germline mutation R337H in Brazil. Up to several thousand of subjects may be carriers and it will be impossible to provide them with comprehensive clinical surveillance using standard protocols currently proposed for high penetrance forms of LFS. A reliable marker to "prioritize" for early intensive surveillance, those individuals who have the highest risk to develop childhood cancer, would be extremely useful. Further clinical evaluation will be needed to determine whether screening for rs17878362 genotypes may help to address this important public health concern.

REFERENCES

A

Abedin Z, Louis-Juste M, Stangl M, Field J (2013). The role of base excision repair genes OGG1, APN1 and APN2 in benzo[a]pyrene-7,8-dione induced p53 mutagenesis. *Mutation research* **750:** 121-128.

Abida WM, Nikolaev A, Zhao W, Zhang W, Gu W (2007). FBXO11 promotes the Neddylation of p53 and inhibits its transcriptional activity. *The Journal of biological chemistry* **282:** 1797-1804.

Achatz MI, Olivier M, Le Calvez F, Martel-Planche G, Lopes A, Rossi BM *et al* (2007). The TP53 mutation, R337H, is associated with Li-Fraumeni and Li-Fraumeni-like syndromes in Brazilian families. *Cancer letters* **245:** 96-102.

Achatz MI, Hainaut P, Ashton-Prolla P (2009). Highly prevalent TP53 mutation predisposing to many cancers in the Brazilian population: a case for newborn screening? *The lancet oncology* **10:** 920-925.

Agarwal ML, Agarwal A, Taylor WR, Stark GR (1995). p53 controls both the G2/M and the G1 cell cycle checkpoints and mediates reversible growth arrest in human fibroblasts. *Proceedings of the National Academy of Sciences of the United States of America* **92:** 8493-8497.

Agirre X, Vizmanos JL, Calasanz MJ, Garcia-Delgado M, Larrayoz MJ, Novo FJ (2003). Methylation of CpG dinucleotides and/or CCWGG motifs at the promoter of TP53 correlates with decreased gene expression in a subset of acute lymphoblastic leukemia patients. *Oncogene* **22:** 1070-1072.

Ahn J, Prives C (2001). The C-terminus of p53: the more you learn the less you know. *Nature structural biology* **8:** 730-732.

Alcorta DA, Xiong Y, Phelps D, Hannon G, Beach D, Barrett JC (1996). Involvement of the cyclin-dependent kinase inhibitor p16 (INK4a) in replicative senescence of normal human fibroblasts. *Proceedings of the National Academy of Sciences of the United States of America* **93:** 13742-13747.

Altilia S, Santoro A, Malagoli D, Lanzarini C, Ballesteros Alvarez JA, Galazzo G *et al* (2012). TP53 codon 72 polymorphism affects accumulation of mtDNA damage in human cells. *Aging* **4:** 28-39.

Amatya VJ, Naumann U, Weller M, Ohgaki H (2005). TP53 promoter methylation in human gliomas. *Acta neuropathologica* **110:** 178-184.

Anderson ME, Woelker B, Reed M, Wang P, Tegtmeyer P (1997). Reciprocal interference between the sequence-specific core and nonspecific C-terminal DNA binding domains of p53: implications for regulation. *Molecular and cellular biology* **17:** 6255-6264.

Ansieau S, Bastid J, Doreau A, Morel AP, Bouchet BP, Thomas C *et al* (2008). Induction of EMT by twist proteins as a collateral effect of tumor-promoting inactivation of premature senescence. *Cancer cell* **14:** 79-89.

Appella E, Anderson CW (2000). Signaling to p53: breaking the posttranslational modification code. *Pathologie-biologie* **48:** 227-245.

Appella E, Anderson CW (2001). Post-translational modifications and activation of p53 by genotoxic stresses. *European journal of biochemistry / FEBS* **268:** 2764-2772.

Arva NC, Gopen TR, Talbott KE, Campbell LE, Chicas A, White DE *et al* (2005). A chromatin-associated and transcriptionally inactive p53-Mdm2 complex occurs in mdm2 SNP309 homozygous cells. *The Journal of biological chemistry* **280:** 26776-26787.

Aury-Landas J, Bougeard G, Castel H, Hernandez-Vargas H, Drouet A, Latouche JB *et al* (2013). Germline copy number variation of genes involved in chromatin remodelling in families suggestive of Li-Fraumeni syndrome with brain tumours. *European journal of human genetics : EJHG*.

Ayed A, Mulder FA, Yi GS, Lu Y, Kay LE, Arrowsmith CH (2001). Latent and active p53 are identical in conformation. *Nature structural biology* **8:** 756-760.

Bagga PS, Ford LP, Chen F, Wilusz J (1995). The G-rich auxiliary downstream element has distinct sequence and position requirements and mediates efficient 3' end pre-mRNA processing through a trans-acting factor. *Nucleic acids research* **23:** 1625-1631.

Baker SJ, Fearon ER, Nigro JM, Hamilton SR, Preisinger AC, Jessup JM *et al* (1989). Chromosome 17 deletions and p53 gene mutations in colorectal carcinomas. *Science* **244:** 217-221.

Balagurumoorthy P, Sakamoto H, Lewis MS, Zambrano N, Clore GM, Gronenborn AM *et al* (1995). Four p53 DNA-binding domain peptides bind natural p53-response elements and bend the DNA. *Proceedings of the National Academy of Sciences of the United States of America* **92:** 8591-8595.

Bandele OJ, Wang X, Campbell MR, Pittman GS, Bell DA (2011). Human single-nucleotide polymorphisms alter p53 sequence-specific binding at gene regulatory elements. *Nucleic acids research* **39:** 178-189.

Barak Y, Juven T, Haffner R, Oren M (1993). mdm2 expression is induced by wild type p53 activity. *The EMBO journal* **12:** 461-468.

Bargonetti J, Manfredi JJ, Chen X, Marshak DR, Prives C (1993). A proteolytic fragment from the central region of p53 has marked sequence-specific DNA-binding activity when generated from wild-type but not from oncogenic mutant p53 protein. *Genes & development* **7:** 2565-2574.

Bartkova J, Rezaei N, Liontos M, Karakaidos P, Kletsas D, Issaeva N *et al* (2006). Oncogene-induced senescence is part of the tumorigenesis barrier imposed by DNA damage checkpoints. *Nature* **444:** 633-637.

Becker K, Marchenko ND, Maurice M, Moll UM (2007). Hyperubiquitylation of wild-type p53 contributes to cytoplasmic sequestration in neuroblastoma. *Cell death and differentiation* **14:** 1350-1360.

Beckman G, Birgander R, Sjalander A, Saha N, Holmberg PA, Kivela A *et al* (1994). Is p53 polymorphism maintained by natural selection? *Human heredity* **44:** 266-270.

Beham A, Marin MC, Fernandez A, Herrmann J, Brisbay S, Tari AM *et al* (1997). Bcl-2 inhibits p53 nuclear import following DNA damage. *Oncogene* **15:** 2767-2772.

Bennett M, Macdonald K, Chan SW, Luzio JP, Simari R, Weissberg P (1998). Cell surface trafficking of Fas: a rapid mechanism of p53-mediated apoptosis. *Science* **282:** 290-293.

Bergamaschi D, Samuels Y, Sullivan A, Zvelebil M, Breyssens H, Bisso A *et al* (2006). iASPP preferentially binds p53 proline-rich region and modulates apoptotic function of codon 72-polymorphic p53. *Nature genetics* **38:** 1133-1141.

Berger M, Vogt Sionov R, Levine AJ, Haupt Y (2001). A role for the polyproline domain of p53 in its regulation by Mdm2. *The Journal of biological chemistry* **276:** 3785-3790.

Berger M, Stahl N, Del Sal G, Haupt Y (2005). Mutations in proline 82 of p53 impair its activation by Pin1 and Chk2 in response to DNA damage. *Molecular and cellular biology* **25:** 5380-5388.

Besnard E, Babled A, Lapasset L, Milhavet O, Parrinello H, Dantec C *et al* (2012). Unraveling cell type-specific and reprogrammable human replication origin signatures associated with G-quadruplex consensus motifs. *Nature structural & molecular biology* **19:** 837-844.

Bieging KT, Attardi LD (2012). Deconstructing p53 transcriptional networks in tumor suppression. *Trends in cell biology* **22:** 97-106.

Birch JM, Hartley AL, Tricker KJ, Prosser J, Condie A, Kelsey AM *et al* (1994). Prevalence and diversity of constitutional mutations in the p53 gene among 21 Li-Fraumeni families. *Cancer research* **54:** 1298-1304.

Birch JM, Blair V, Kelsey AM, Evans DG, Harris M, Tricker KJ *et al* (1998). Cancer phenotype correlates with constitutional TP53 genotype in families with the Li-Fraumeni syndrome. *Oncogene* **17:** 1061-1068.

Birch JM, Alston RD, McNally RJ, Evans DG, Kelsey AM, Harris M *et al* (2001). Relative frequency and morphology of cancers in carriers of germline TP53 mutations. *Oncogene* **20:** 4621-4628.

Blattner WA, McGuire DB, Mulvihill JJ, Lampkin BC, Hananian J, Fraumeni JF (1979). Genealogy of cancer in a family. *JAMA : the journal of the American Medical Association* **241:** 259-261.

Bochkareva E, Kaustov L, Ayed A, Yi GS, Lu Y, Pineda-Lucena A *et al* (2005). Single-stranded DNA mimicry in the p53 transactivation domain interaction with replication protein A. *Proceedings of the National Academy of Sciences of the United States of America* **102:** 15412-15417.

Bochman ML, Paeschke K, Zakian VA (2012). DNA secondary structures: stability and function of G-quadruplex structures. *Nature reviews Genetics* **13:** 770-780.

Bohlig L, Friedrich M, Engeland K (2011). p53 activates the PANK1/miRNA-107 gene leading to downregulation of CDK6 and p130 cell cycle proteins. *Nucleic acids research* **39:** 440-453.

Boldrini L, Gisfredi S, Ursino S, Lucchi M, Greco G, Mussi A *et al* (2008). Effect of the p53 codon 72 and intron 3 polymorphisms on non-small cell lung cancer (NSCLC) prognosis. *Cancer investigation* **26:** 168-172.

Bommer GT, Gerin I, Feng Y, Kaczorowski AJ, Kuick R, Love RE *et al* (2007). p53-mediated activation of miRNA34 candidate tumor-suppressor genes. *Current biology : CB* **17:** 1298-1307.

Bond GL, Hu W, Bond EE, Robins H, Lutzker SG, Arva NC *et al* (2004). A single nucleotide polymorphism in the MDM2 promoter attenuates the p53 tumor suppressor pathway and accelerates tumor formation in humans. *Cell* **119:** 591-602.

Bond GL, Hirshfield KM, Kirchhoff T, Alexe G, Bond EE, Robins H *et al* (2006). MDM2 SNP309 accelerates tumor formation in a gender-specific and hormone-dependent manner. *Cancer research* **66:** 5104-5110.

Bossi G, Lapi E, Strano S, Rinaldo C, Blandino G, Sacchi A (2006). Mutant p53 gain of function: reduction of tumor malignancy of human cancer cell lines through abrogation of mutant p53 expression. *Oncogene* **25:** 304-309.

Bou Kheir T, Futoma-Kazmierczak E, Jacobsen A, Krogh A, Bardram L, Hother C *et al* (2011). miR-449 inhibits cell proliferation and is down-regulated in gastric cancer. *Molecular cancer* **10:** 29.

Bougeard G, Baert-Desurmont S, Tournier I, Vasseur S, Martin C, Brugieres L *et al* (2006). Impact of the MDM2 SNP309 and p53 Arg72Pro polymorphism on age of tumour onset in Li-Fraumeni syndrome. *Journal of medical genetics* **43:** 531-533.

Bougeard G, Sesboue R, Baert-Desurmont S, Vasseur S, Martin C, Tinat J *et al* (2008). Molecular basis of the Li-Fraumeni syndrome: an update from the French LFS families. *Journal of medical genetics* **45:** 535-538.

Bourdon JC, Fernandes K, Murray-Zmijewski F, Liu G, Diot A, Xirodimas DP *et al* (2005). p53 isoforms can regulate p53 transcriptional activity. *Genes & development* **19:** 2122-2137.

Brady CA, Jiang D, Mello SS, Johnson TM, Jarvis LA, Kozak MM *et al* (2011). Distinct p53 transcriptional programs dictate acute DNA-damage responses and tumor suppression. *Cell* **145:** 571-583.

Brantley MA, Jr., Worley L, Harbour JW (2002). Altered expression of Rb and p53 in uveal melanomas following plaque radiotherapy. *American journal of ophthalmology* **133:** 242-248.

Braun CJ, Zhang X, Savelyeva I, Wolff S, Moll UM, Schepeler T *et al* (2008). p53-Responsive micrornas 192 and 215 are capable of inducing cell cycle arrest. *Cancer research* **68:** 10094-10104.

Brooks CL, Gu W (2011). The impact of acetylation and deacetylation on the p53 pathway. *Protein & cell* **2:** 456-462.

Brosh R, Rotter V (2009). When mutants gain new powers: news from the mutant p53 field. *Nature reviews Cancer* **9:** 701-713.

Budhram-Mahadeo V, Morris PJ, Smith MD, Midgley CA, Boxer LM, Latchman DS (1999). p53 suppresses the activation of the Bcl-2 promoter by the Brn-3a POU family transcription factor. *The Journal of biological chemistry* **274:** 15237-15244.

Buschmann T, Potapova O, Bar-Shira A, Ivanov VN, Fuchs SY, Henderson S *et al* (2001). Jun NH2-terminal kinase phosphorylation of p53 on Thr-81 is important for p53 stabilization and transcriptional activities in response to stress. *Molecular and cellular biology* **21:** 2743-2754.

Bykov VJ, Issaeva N, Selivanova G, Wiman KG (2002). Mutant p53-dependent growth suppression distinguishes PRIMA-1 from known anticancer drugs: a statistical analysis of information in the National Cancer Institute database. *Carcinogenesis* **23:** 2011-2018.

Cai X, Liu X (2008). Inhibition of Thr-55 phosphorylation restores p53 nuclear localization and sensitizes cancer cells to DNA damage. *Proceedings of the National Academy of Sciences of the United States of America* **105:** 16958-16963.

Campbell PM, Groehler AL, Lee KM, Ouellette MM, Khazak V, Der CJ (2007). K-Ras promotes growth transformation and invasion of immortalized human pancreatic cells by Raf and phosphatidylinositol 3-kinase signaling. *Cancer research* **67:** 2098-2106.

Camplejohn RS, Perry P, Hodgson SV, Turner G, Williams A, Upton C *et al* (1995). A possible screening test for inherited p53-related defects based on the apoptotic response of peripheral blood lymphocytes to DNA damage. *British journal of cancer* **72:** 654-662.

Candeias MM, Malbert-Colas L, Powell DJ, Daskalogianni C, Maslon MM, Naski N *et al* (2008). P53 mRNA controls p53 activity by managing Mdm2 functions. *Nature cell biology* **10:** 1098-1105.

Canman CE, Lim DS, Cimprich KA, Taya Y, Tamai K, Sakaguchi K *et al* (1998). Activation of the ATM kinase by ionizing radiation and phosphorylation of p53. *Science* **281:** 1677-1679.

Capra JA, Paeschke K, Singh M, Zakian VA (2010). G-quadruplex DNA sequences are evolutionarily conserved and associated with distinct genomic features in Saccharomyces cerevisiae. *PLoS computational biology* **6:** e1000861.

Carmeliet P, Dor Y, Herbert JM, Fukumura D, Brusselmans K, Dewerchin M *et al* (1998). Role of HIF-1alpha in hypoxia-mediated apoptosis, cell proliferation and tumour angiogenesis. *Nature* **394:** 485-490.

Carrier F, Georgel PT, Pourquier P, Blake M, Kontny HU, Antinore MJ *et al* (1999). Gadd45, a p53-responsive stress protein, modifies DNA accessibility on damaged chromatin. *Molecular and cellular biology* **19:** 1673-1685.

Carter S, Vousden KH (2008). p53-Ubl fusions as models of ubiquitination, sumoylation and neddylation of p53. *Cell Cycle* **7:** 2519-2528.

Cattelani S, Ferrari-Amorotti G, Galavotti S, Defferrari R, Tanno B, Cialfi S *et al* (2012). The p53 codon 72 Pro/Pro genotype identifies poor-prognosis neuroblastoma patients: correlation with reduced apoptosis and enhanced senescence by the p53-72P isoform. *Neoplasia* **14:** 634-643.

Cawley S, Bekiranov S, Ng HH, Kapranov P, Sekinger EA, Kampa D *et al* (2004). Unbiased mapping of transcription factor binding sites along human chromosomes 21 and 22 points to widespread regulation of noncoding RNAs. *Cell* **116:** 499-509.

Cayrou C, Coulombe P, Puy A, Rialle S, Kaplan N, Segal E *et al* (2012). New insights into replication origin characteristics in metazoans. *Cell Cycle* **11**.

Cecconi F, Alvarez-Bolado G, Meyer BI, Roth KA, Gruss P (1998). Apaf1 (CED-4 homolog) regulates programmed cell death in mammalian development. *Cell* **94:** 727-737.

Chang CJ, Chao CH, Xia W, Yang JY, Xiong Y, Li CW *et al* (2011). p53 regulates epithelial-mesenchymal transition and stem cell properties through modulating miRNAs. *Nature cell biology* **13:** 317-323.

Chang YC, Lee YS, Tejima T, Tanaka K, Omura S, Heintz NH *et al* (1998). mdm2 and bax, downstream mediators of the p53 response, are degraded by the ubiquitin-proteasome pathway. *Cell growth & differentiation : the molecular biology journal of the American Association for Cancer Research* **9:** 79-84.

Chao C, Herr D, Chun J, Xu Y (2006). Ser18 and 23 phosphorylation is required for p53-dependent apoptosis and tumor suppression. *The EMBO journal* **25:** 2615-2622.

Chehab NH, Malikzay A, Stavridi ES, Halazonetis TD (1999). Phosphorylation of Ser-20 mediates stabilization of human p53 in response to DNA damage. *Proceedings of the National Academy of Sciences of the United States of America* **96:** 13777-13782.

Chen L, Gilkes DM, Pan Y, Lane WS, Chen J (2005). ATM and Chk2-dependent phosphorylation of MDMX contribute to p53 activation after DNA damage. *The EMBO journal* **24:** 3411-3422.

Chen QR, Yu LR, Tsang P, Wei JS, Song YK, Cheuk A *et al* (2011). Systematic proteome analysis identifies transcription factor YY1 as a direct target of miR-34a. *Journal of proteome research* **10:** 479-487.

Chene P, Mittl P, Grutter M (1997). In vitro structure-function analysis of the beta-strand 326-333 of human p53. *Journal of molecular biology* **273:** 873-881.

Chene P (1998). In vitro analysis of the dominant negative effect of p53 mutants. *Journal of molecular biology* **281:** 205-209.

Chene P (2001). The role of tetramerization in p53 function. *Oncogene* **20:** 2611-2617.

Chin L, Artandi SE, Shen Q, Tam A, Lee SL, Gottlieb GJ *et al* (1999). p53 deficiency rescues the adverse effects of telomere loss and cooperates with telomere dysfunction to accelerate carcinogenesis. *Cell* **97:** 527-538.

Cho Y, Gorina S, Jeffrey PD, Pavletich NP (1994). Crystal structure of a p53 tumor suppressor-DNA complex: understanding tumorigenic mutations. *Science* **265:** 346-355.

Chompret A, Brugieres L, Ronsin M, Gardes M, Dessarps-Freichey F, Abel A *et al* (2000). P53 germline mutations in childhood cancers and cancer risk for carrier individuals. *British journal of cancer* **82:** 1932-1937.

Chompret A, Abel A, Stoppa-Lyonnet D, Brugieres L, Pages S, Feunteun J *et al* (2001). Sensitivity and predictive value of criteria for p53 germline mutation screening. *Journal of medical genetics* **38:** 43-47.

Chuikov S, Kurash JK, Wilson JR, Xiao B, Justin N, Ivanov GS *et al* (2004). Regulation of p53 activity through lysine methylation. *Nature* **432:** 353-360.

Cimmino A, Calin GA, Fabbri M, Iorio MV, Ferracin M, Shimizu M *et al* (2005). miR-15 and miR-16 induce apoptosis by targeting BCL2. *Proceedings of the National Academy of Sciences of the United States of America* **102:** 13944-13949.

Clore GM, Omichinski JG, Sakaguchi K, Zambrano N, Sakamoto H, Appella E *et al* (1994). High-resolution structure of the oligomerization domain of p53 by multidimensional NMR. *Science* **265:** 386-391.

Costa S, Pinto D, Pereira D, Rodrigues H, Cameselle-Teijeiro J, Medeiros R *et al* (2008). Importance of TP53 codon 72 and intron 3 duplication 16bp polymorphisms in prediction of susceptibility on breast cancer. *BMC cancer* **8:** 32.

Courtois S, Verhaegh G, North S, Luciani MG, Lassus P, Hibner U *et al* (2002). DeltaN-p53, a natural isoform of p53 lacking the first transactivation domain, counteracts growth suppression by wild-type p53. *Oncogene* **21:** 6722-6728.

Craig AL, Burch L, Vojtesek B, Mikutowska J, Thompson A, Hupp TR (1999). Novel phosphorylation sites of human tumour suppressor protein p53 at Ser20 and Thr18 that disrupt the binding of mdm2 (mouse double minute 2) protein are modified in human cancers. *The Biochemical journal* **342 (Pt 1):** 133-141.

Crighton D, Wilkinson S, O'Prey J, Syed N, Smith P, Harrison PR *et al* (2006). DRAM, a p53-induced modulator of autophagy, is critical for apoptosis. *Cell* **126:** 121-134.

Cummins JM, Rago C, Kohli M, Kinzler KW, Lengauer C, Vogelstein B (2004). Tumour suppression: disruption of HAUSP gene stabilizes p53. *Nature* **428:** 1 p following 486.

Custodio G, Taques GR, Figueiredo BC, Gugelmin ES, Oliveira Figueiredo MM, Watanabe F *et al* (2011). Increased incidence of choroid plexus carcinoma due to the germline TP53 R337H mutation in southern Brazil. *PloS one* **6:** e18015.

Custodio G, Parise GA, Kiesel Filho N, Komechen H, Sabbaga CC, Rosati R *et al* (2013). Impact of neonatal screening and surveillance for the TP53 R337H mutation on early detection of childhood adrenocortical tumors. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* **31:** 2619-2626.

D'Orazi G, Cecchinelli B, Bruno T, Manni I, Higashimoto Y, Saito S *et al* (2002). Homeodomain-interacting protein kinase-2 phosphorylates p53 at Ser 46 and mediates apoptosis. *Nature cell biology* **4:** 11-19.

da Costa NM, Hautefeuille A, Cros MP, Melendez ME, Waters T, Swann P *et al* (2012). Transcriptional regulation of thymine DNA glycosylase (TDG) by the tumor suppressor protein p53. *Cell Cycle* **11:** 4570-4578.

Dai C, Gu W (2010). p53 post-translational modification: deregulated in tumorigenesis. *Trends in molecular medicine* **16:** 528-536.

Dai J, Dexheimer TS, Chen D, Carver M, Ambrus A, Jones RA *et al* (2006). An intramolecular G-quadruplex structure with mixed parallel/antiparallel G-strands formed in the human BCL-2 promoter region in solution. *Journal of the American Chemical Society* **128:** 1096-1098.

Dai S, Mao C, Jiang L, Wang G, Cheng H (2009). P53 polymorphism and lung cancer susceptibility: a pooled analysis of 32 case-control studies. *Human genetics* **125:** 633-638.

Dalziel M, Nunes NM, Furger A (2007). Two G-rich regulatory elements located adjacent to and 440 nucleotides downstream of the core poly(A) site of the intronless melanocortin receptor 1 gene are critical for efficient 3' end processing. *Molecular and cellular biology* **27:** 1568-1580.

Dameron KM, Volpert OV, Tainsky MA, Bouck N (1994). The p53 tumor suppressor gene inhibits angiogenesis by stimulating the production of thrombospondin. *Cold Spring Harbor symposia on quantitative biology* **59:** 483-489.

Dang CV, Lee WM (1989). Nuclear and nucleolar targeting sequences of c-erb-A, c-myb, N-myc, p53, HSP70, and HIV tat proteins. *The Journal of biological chemistry* **264:** 18019-18023.

David-Pfeuty T, Chakrani F, Ory K, Nouvian-Dooghe Y (1996). Cell cycle-dependent regulation of nuclear p53 traffic occurs in one subclass of human tumor cells and in untransformed cells. *Cell growth & differentiation : the molecular biology journal of the American Association for Cancer Research* **7:** 1211-1225.

Dawson R, Muller L, Dehner A, Klein C, Kessler H, Buchner J (2003). The N-terminal domain of p53 is natively unfolded. *Journal of molecular biology* **332:** 1131-1141.

Daya-Grosjean L, Dumaz N, Sarasin A (1995). The specificity of p53 mutation spectra in sunlight induced human cancers. *Journal of photochemistry and photobiology B, Biology* **28:** 115-124.

De Armond R, Wood S, Sun D, Hurley LH, Ebbinghaus SW (2005). Evidence for the presence of a guanine quadruplex forming region within a polypurine tract of the hypoxia inducible factor 1alpha promoter. *Biochemistry* **44:** 16341-16350.

De Cian A, Gros J, Guedin A, Haddi M, Lyonnais S, Guittat L *et al* (2008a). DNA and RNA quadruplex ligands. *Nucleic Acids Symp Ser (Oxf)***:** 7-8.

De Cian A, Lacroix L, Douarre C, Temime-Smaali N, Trentesaux C, Riou JF *et al* (2008b). Targeting telomeres and telomerase. *Biochimie* **90:** 131-155.

Decorsiere A, Cayrel A, Vagner S, Millevoi S (2011). Essential role for the interaction between hnRNP H/F and a G quadruplex in maintaining p53 pre-mRNA 3'-end processing and function during DNA damage. *Genes & development* **25:** 220-225.

DeLeo AB, Jay G, Appella E, Dubois GC, Law LW, Old LJ (1979). Detection of a transformation-related antigen in chemically induced sarcomas and other transformed cells of the mouse. *Proceedings of the National Academy of Sciences of the United States of America* **76:** 2420-2424.

Delia D, Goi K, Mizutani S, Yamada T, Aiello A, Fontanella E *et al* (1997). Dissociation between cell cycle arrest and apoptosis can occur in Li-Fraumeni cells heterozygous for p53 gene mutations. *Oncogene* **14:** 2137-2147.

Deng C, Zhang P, Harper JW, Elledge SJ, Leder P (1995). Mice lacking p21CIP1/WAF1 undergo normal development, but are defective in G1 checkpoint control. *Cell* **82:** 675-684.

Denissenko MF, Pao A, Tang M, Pfeifer GP (1996). Preferential formation of benzo[a]pyrene adducts at lung cancer mutational hotspots in P53. *Science* **274:** 430-432.

Dhar P, Chattopadhyay K, Bhattacharyya D, Roychoudhury A, Biswas A, Ghosh S (2006). Antioxidative effect of conjugated linolenic acid in diabetic and non-diabetic blood: an in vitro study. *Journal of oleo science* **56:** 19-24.

Di Agostino S, Strano S, Emiliozzi V, Zerbini V, Mottolese M, Sacchi A *et al* (2006). Gain of function of mutant p53: the mutant p53/NF-Y protein complex reveals an aberrant transcriptional mechanism of cell cycle regulation. *Cancer cell* **10:** 191-202.

Di Leonardo A, Linke SP, Clarkin K, Wahl GM (1994). DNA damage triggers a prolonged p53-dependent G1 arrest and long-term induction of Cip1 in normal human fibroblasts. *Genes & development* **8:** 2540-2551.

Didiot MC, Tian Z, Schaeffer C, Subramanian M, Mandel JL, Moine H (2008). The G-quartet containing FMRP binding site in FMR1 mRNA is a potent exonic splicing enhancer. *Nucleic acids research* **36:** 4902-4912.

DiGiammarino EL, Lee AS, Cadwell C, Zhang W, Bothner B, Ribeiro RC *et al* (2002). A novel mechanism of tumorigenesis involving pH-dependent destabilization of a mutant p53 tetramer. *Nature structural biology* **9:** 12-16.

Dohn M, Jiang J, Chen X (2001). Receptor tyrosine kinase EphA2 is regulated by p53-family proteins and induces apoptosis. *Oncogene* **20:** 6503-6515.

Dornan D, Bheddah S, Newton K, Ince W, Frantz GD, Dowd P *et al* (2004a). COP1, the negative regulator of p53, is overexpressed in breast and ovarian adenocarcinomas. *Cancer research* **64:** 7226-7230.

Dornan D, Wertz I, Shimizu H, Arnott D, Frantz GD, Dowd P *et al* (2004b). The ubiquitin ligase COP1 is a critical negative regulator of p53. *Nature* **429:** 86-92.

Douarre C, Gomez D, Morjani H, Zahm JM, O'Donohue M F, Eddabra L *et al* (2005). Overexpression of Bcl-2 is associated with apoptotic resistance to the G-quadruplex ligand 12459 but is not sufficient to confer resistance to long-term senescence. *Nucleic acids research* **33:** 2192-2203.

Duan W, Gao L, Wu X, Zhang Y, Otterson GA, Villalona-Calero MA (2006). Differential response between the p53 ubiquitin-protein ligases Pirh2 and MdM2 following DNA damage in human cancer cells. *Experimental cell research* **312:** 3370-3378.

Dumaz N, Meek DW (1999). Serine15 phosphorylation stimulates p53 transactivation but does not directly influence interaction with HDM2. *The EMBO journal* **18:** 7002-7010.

Dumont P, Leu JI, Della Pietra AC, 3rd, George DL, Murphy M (2003). The codon 72 polymorphic variants of p53 have markedly different apoptotic potential. *Nature genetics* **33:** 357-365.

Eades G, Yao Y, Yang M, Zhang Y, Chumsri S, Zhou Q (2011). miR-200a regulates SIRT1 expression and epithelial to mesenchymal transition (EMT)-like transformation in mammary epithelial cells. *The Journal of biological chemistry* **286:** 25992-26002.

Eddy BE, Borman GS, Grubbs GE, Young RD (1962). Identification of the oncogenic substance in rhesus monkey kidney cell culture as simian virus 40. *Virology* **17:** 65-75.

Eeles RA (1995). Germline mutations in the TP53 gene. *Cancer surveys* **25:** 101-124.

Egan KM, Nabors LB, Olson JJ, Monteiro AN, Browning JE, Madden MH *et al* (2012). Rare TP53 genetic variant associated with glioma risk and outcome. *Journal of medical genetics* **49:** 420-421.

el-Deiry WS, Kern SE, Pietenpol JA, Kinzler KW, Vogelstein B (1992). Definition of a consensus binding site for p53. *Nature genetics* **1:** 45-49.

el-Deiry WS, Tokino T, Velculescu VE, Levy DB, Parsons R, Trent JM *et al* (1993). WAF1, a potential mediator of p53 tumor suppression. *Cell* **75:** 817-825.

el-Deiry WS, Harper JW, O'Connor PM, Velculescu VE, Canman CE, Jackman J *et al* (1994). WAF1/CIP1 is induced in p53-mediated G1 arrest and apoptosis. *Cancer research* **54:** 1169-1174.

El-Deiry WS (2001). Insights into cancer therapeutic design based on p53 and TRAIL receptor signaling. *Cell death and differentiation* **8:** 1066-1075.

Elkind NB, Goldfinger N, Rotter V (1995). Spot-1, a novel NLS-binding protein that interacts with p53 through a domain encoded by p(CA)n repeats. *Oncogene* **11:** 841-851.

Enciso-Mora V, Hosking FJ, Di Stefano AL, Zelenika D, Shete S, Broderick P *et al* (2013). Low penetrance susceptibility to glioma is caused by the TP53 variant rs78378222. *British journal of cancer* **108:** 2178-2185.

Fabbri M, Bottoni A, Shimizu M, Spizzo R, Nicoloso MS, Rossi S *et al* (2011). Association of a microRNA/TP53 feedback circuitry with pathogenesis and outcome of B-cell chronic lymphocytic leukemia. *JAMA : the journal of the American Medical Association* **305:** 59-67.

Farmer G, Bargonetti J, Zhu H, Friedman P, Prywes R, Prives C (1992). Wild-type p53 activates transcription in vitro. *Nature* **358:** 83-86.

Farnebo M (2009). Wrap53, a novel regulator of p53. *Cell Cycle* **8:** 2343-2346.

Feng Z, Zhang H, Levine AJ, Jin S (2005). The coordinate regulation of the p53 and mTOR pathways in cells. *Proceedings of the National Academy of Sciences of the United States of America* **102:** 8204-8209.

Feng Z, Zhang C, Kang HJ, Sun Y, Wang H, Naqvi A *et al* (2011). Regulation of female reproduction by p53 and its family members. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology* **25:** 2245-2255.

Fernandes TA, Lima GL, De Souza FC, Fernandes JV, Meissner RV (2008). Evaluation of the polymorphisms in the exons 2 to 4 of the TP53 in cervical carcinoma patients from a Brazilian population. *Cell Mol Biol (Noisy-le-grand)* **54 Suppl:** OL1025-1031.

Ferreon JC, Lee CW, Arai M, Martinez-Yamout MA, Dyson HJ, Wright PE (2009). Cooperative regulation of p53 by modulation of ternary complex formation with CBP/p300 and HDM2. *Proceedings of the National Academy of Sciences of the United States of America* **106:** 6591-6596.

Finch RA, Donoviel DB, Potter D, Shi M, Fan A, Freed DD *et al* (2002). mdmx is a negative regulator of p53 activity in vivo. *Cancer research* **62:** 3221-3225.

Finlay CA, Hinds PW, Levine AJ (1989). The p53 proto-oncogene can act as a suppressor of transformation. *Cell* **57:** 1083-1093.

Flaman JM, Frebourg T, Moreau V, Charbonnier F, Martin C, Ishioka C *et al* (1994). A rapid PCR fidelity assay. *Nucleic acids research* **22:** 3259-3260.

Flores I, Blasco MA (2009). A p53-dependent response limits epidermal stem cell functionality and organismal size in mice with short telomeres. *PloS one* **4:** e4934.

Folkman J, Long DM, Jr., Becker FF (1963). Growth and metastasis of tumor in organ culture. *Cancer* **16:** 453-467.

Foord OS, Bhattacharya P, Reich Z, Rotter V (1991). A DNA binding domain is contained in the C-terminus of wild type p53 protein. *Nucleic acids research* **19:** 5191-5198.

Fornari F, Gramantieri L, Giovannini C, Veronese A, Ferracin M, Sabbioni S *et al* (2009). MiR-122/cyclin G1 interaction modulates p53 activity and affects doxorubicin sensitivity of human hepatocarcinoma cells. *Cancer research* **69:** 5761-5767.

Forsythe JA, Jiang BH, Iyer NV, Agani F, Leung SW, Koos RD *et al* (1996). Activation of vascular endothelial growth factor gene transcription by hypoxia-inducible factor 1. *Molecular and cellular biology* **16:** 4604-4613.

Freedman DA, Levine AJ (1998). Nuclear export is required for degradation of endogenous p53 by MDM2 and human papillomavirus E6. *Molecular and cellular biology* **18:** 7288-7293.

Fuchs SY, Adler V, Buschmann T, Yin Z, Wu X, Jones SN *et al* (1998a). JNK targets p53 ubiquitination and degradation in nonstressed cells. *Genes & development* **12:** 2658-2663.

Fuchs SY, Adler V, Pincus MR, Ronai Z (1998b). MEKK1/JNK signaling stabilizes and activates p53. *Proceedings of the National Academy of Sciences of the United States of America* **95:** 10541-10546.

Funk WD, Pak DT, Karas RH, Wright WE, Shay JW (1992). A transcriptionally active DNA-binding site for human p53 protein complexes. *Molecular and cellular biology* **12:** 2866-2871.

Gadea G, de Toledo M, Anguille C, Roux P (2007). Loss of p53 promotes RhoA-ROCK-dependent cell migration and invasion in 3D matrices. *The Journal of cell biology* **178:** 23-30.

Galea C, Bowman P, Kriwacki RW (2005). Disruption of an intermonomer salt bridge in the p53 tetramerization domain results in an increased propensity to form amyloid fibrils. *Protein science : a publication of the Protein Society* **14:** 2993-3003.

Garritano S, Gemignani F, Palmero EI, Olivier M, Martel-Planche G, Le Calvez-Kelm F *et al* (2010). Detailed haplotype analysis at the TP53 locus in p.R337H mutation carriers in the population of Southern Brazil: evidence for a founder effect. *Human mutation* **31:** 143-150.

Gauthier LR, Granotier C, Hoffschir F, Etienne O, Ayouaz A, Desmaze C *et al* (2012). Rad51 and DNA-PKcs are involved in the generation of specific telomere aberrations induced by the quadruplex ligand 360A that impair mitotic cell progression and lead to cell death. *Cellular and molecular life sciences : CMLS* **69:** 629-640.

Geiser AG, Stanbridge EJ (1989). A review of the evidence for tumor suppressor genes. *Critical reviews in oncogenesis* **1:** 261-276.

Gellert M, Lipsett MN, Davies DR (1962). Helix formation by guanylic acid. *Proceedings of the National Academy of Sciences of the United States of America* **48:** 2013-2018.

Gemignani F, Moreno V, Landi S, Moullan N, Chabrier A, Gutierrez-Enriquez S *et al* (2004). A TP53 polymorphism is associated with increased risk of colorectal cancer and with reduced levels of TP53 mRNA. *Oncogene* **23:** 1954-1956.

Ghavami S, Mutawe MM, Sharma P, Yeganeh B, McNeill KD, Klonisch T *et al* (2011). Mevalonate cascade regulation of airway mesenchymal cell autophagy and apoptosis: a dual role for p53. *PloS one* **6:** e16523.

Ghosal G, Muniyappa K (2005). Saccharomyces cerevisiae Mre11 is a high-affinity G4 DNA-binding protein and a G-rich DNA-specific endonuclease: implications for replication of telomeric DNA. *Nucleic acids research* **33:** 4692-4703.

Ghosal G, Muniyappa K (2007). The characterization of Saccharomyces cerevisiae Mre11/Rad50/Xrs2 complex reveals that Rad50 negatively regulates Mre11 endonucleolytic but not the exonucleolytic activity. *Journal of molecular biology* **372:** 864-882.

Giacomazzi J, Koehler-Santos P, Palmero EI, Graudenz MS, Rivero LF, Lima E *et al* (2013). A TP53 founder mutation, p.R337H, is associated with phyllodes breast tumors in Brazil. *Virchows Archiv : an international journal of pathology* **463:** 17-22.

Giannakakou P, Nakano M, Nicolaou KC, O'Brate A, Yu J, Blagosklonny MV *et al* (2002). Enhanced microtubule-dependent trafficking and p53 nuclear accumulation by suppression of microtubule dynamics. *Proceedings of the National Academy of Sciences of the United States of America* **99:** 10855-10860.

Golubovskaya VM, Finch R, Kweh F, Massoll NA, Campbell-Thompson M, Wallace MR *et al* (2008). p53 regulates FAK expression in human tumor cells. *Molecular carcinogenesis* **47:** 373-382.

Gomes MC, Kotsopoulos J, de Almeida GL, Costa MM, Vieira R, Filho Fde A *et al* (2012). The R337H mutation in TP53 and breast cancer in Brazil. *Hereditary cancer in clinical practice* **10:** 3.

Gomez D, Aouali N, Londono-Vallejo A, Lacroix L, Megnin-Chanet F, Lemarteleur T *et al* (2003a). Resistance to the short term antiproliferative activity of the G-quadruplex ligand 12459 is associated with telomerase overexpression and telomere capping alteration. *The Journal of biological chemistry* **278:** 50554-50562.

Gomez D, Aouali N, Renaud A, Douarre C, Shin-Ya K, Tazi J *et al* (2003b). Resistance to senescence induction and telomere shortening by a G-quadruplex ligand inhibitor of telomerase. *Cancer research* **63:** 6149-6153.

Gomez D, Lemarteleur T, Lacroix L, Mailliet P, Mergny JL, Riou JF (2004). Telomerase downregulation induced by the G-quadruplex ligand 12459 in A549 cells is mediated by hTERT RNA alternative splicing. *Nucleic acids research* **32:** 371-379.

Gonzalez KD, Noltner KA, Buzin CH, Gu D, Wen-Fong CY, Nguyen VQ *et al* (2009). Beyond Li Fraumeni Syndrome: clinical characteristics of families with p53 germline mutations. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* **27:** 1250-1256.

Gostissa M, Hengstermann A, Fogal V, Sandy P, Schwarz SE, Scheffner M *et al* (1999). Activation of p53 by conjugation to the ubiquitin-like protein SUMO-1. *The EMBO journal* **18:** 6462-6471.

Gouas D, Shi H, Hainaut P (2009). The aflatoxin-induced TP53 mutation at codon 249 (R249S): biomarker of exposure, early detection and target for therapy. *Cancer letters* **286:** 29-37.

Grossman SR, Deato ME, Brignone C, Chan HM, Kung AL, Tagami H *et al* (2003). Polyubiquitination of p53 by a ubiquitin ligase activity of p300. *Science* **300:** 342-344.

Gu J, Kawai H, Nie L, Kitao H, Wiederschain D, Jochemsen AG *et al* (2002). Mutual dependence of MDM2 and MDMX in their functional inactivation of p53. *The Journal of biological chemistry* **277:** 19251-19254.

Guan X, Wang LE, Liu Z, Sturgis EM, Wei Q (2013). Association between a rare novel TP53 variant (rs78378222) and melanoma, squamous cell carcinoma of head and neck and lung cancer susceptibility in non-Hispanic Whites. *Journal of cellular and molecular medicine* **17:** 873-878.

Guillouf C, Rosselli F, Krishnaraju K, Moustacchi E, Hoffman B, Liebermann DA (1995). p53 involvement in control of G2 exit of the cell cycle: role in DNA damage-induced apoptosis. *Oncogene* **10:** 2263-2270.

Guimaraes DP, Hainaut P (2002). TP53: a key gene in human cancer. *Biochimie* **84:** 83-93.

Hafsi H, Hainaut P (2011). Redox control and interplay between p53 isoforms: roles in the regulation of basal p53 levels, cell fate, and senescence. *Antioxidants & redox signaling* **15:** 1655-1667.

Hafsi H, Santos-Silva D, Courtois-Cox S, Hainaut P (2013). Effects of Delta40p53, an isoform of p53 lacking the N-terminus, on transactivation capacity of the tumor suppressor protein p53. *BMC cancer* **13:** 134.

Hai Y, Cao W, Liu G, Hong SP, Elela SA, Klinck R *et al* (2008). A G-tract element in apoptotic agents-induced alternative splicing. *Nucleic acids research* **36:** 3320-3331.

Hainaut P, Milner J (1993). A structural role for metal ions in the "wild-type" conformation of the tumor suppressor protein p53. *Cancer research* **53:** 1739-1742.

Hainaut P, Pfeifer GP (2001). Patterns of p53 G-->T transversions in lung cancers reflect the primary mutagenic signature of DNA-damage by tobacco smoke. *Carcinogenesis* **22:** 367-374.

Hainaut P, Wiman KG (2009). 30 years and a long way into p53 research. *The lancet oncology* **10:** 913-919.

Hamard PJ, Lukin DJ, Manfredi JJ (2012). p53 basic C terminus regulates p53 functions through DNA binding modulation of subset of target genes. *The Journal of biological chemistry* **287:** 22397-22407.

Hanahan D, Weinberg RA (2000). The hallmarks of cancer. *Cell* **100:** 57-70.

Hanahan D, Weinberg RA (2011). Hallmarks of cancer: the next generation. *Cell* **144:** 646-674.

Harper JW, Adami GR, Wei N, Keyomarsi K, Elledge SJ (1993). The p21 Cdk-interacting protein Cip1 is a potent inhibitor of G1 cyclin-dependent kinases. *Cell* **75:** 805-816.

Harvey DM, Levine AJ (1991). p53 alteration is a common event in the spontaneous immortalization of primary BALB/c murine embryo fibroblasts. *Genes & development* **5:** 2375-2385.

Haupt Y, Maya R, Kazaz A, Oren M (1997). Mdm2 promotes the rapid degradation of p53. *Nature* **387:** 296-299.

He L, He X, Lim LP, de Stanchina E, Xuan Z, Liang Y *et al* (2007). A microRNA component of the p53 tumour suppressor network. *Nature* **447:** 1130-1134.

He W, Long J, Xian L, Pang F, Su L, Wei S *et al* (2012). MDM2 SNP309 polymorphism is associated with lung cancer risk in women: A meta-analysis using METAGEN. *Experimental and therapeutic medicine* **4:** 569-576.

He XF, Su J, Wang W (2011a). Need for clarification of data in the recent meta-analysis about p53 codon 72 polymorphism and breast cancer risk. *Breast cancer research and treatment* **129:** 291-292.

He XF, Su J, Zhang Y, Huang X, Liu Y, Ding DP *et al* (2011b). Association between the p53 polymorphisms and breast cancer risk: meta-analysis based on case-control study. *Breast cancer research and treatment* **130:** 517-529.

Henderson BR, Eleftheriou A (2000). A comparison of the activity, sequence specificity, and CRM1-dependence of different nuclear export signals. *Experimental cell research* **256:** 213-224.

Henn A, Joachimi A, Goncalves DP, Monchaud D, Teulade-Fichou MP, Sanders JK *et al* (2008). Inhibition of dicing of guanosine-rich shRNAs by quadruplex-binding compounds. *Chembiochem : a European journal of chemical biology* **9:** 2722-2729.

Herbig U, Jobling WA, Chen BP, Chen DJ, Sedivy JM (2004). Telomere shortening triggers senescence of human cells through a pathway involving ATM, p53, and p21(CIP1), but not p16(INK4a). *Molecular cell* **14:** 501-513.

Hermeking H, Lengauer C, Polyak K, He TC, Zhang L, Thiagalingam S *et al* (1997). 14-3-3 sigma is a p53-regulated inhibitor of G2/M progression. *Molecular cell* **1:** 3-11.

Hermeking H (2012). MicroRNAs in the p53 network: micromanagement of tumour suppression. *Nature reviews Cancer* **12:** 613-626.

Herrera-Merchan A, Cerrato C, Luengo G, Dominguez O, Piris MA, Serrano M *et al* (2010). miR-33-mediated downregulation of p53 controls hematopoietic stem cell self-renewal. *Cell Cycle* **9:** 3277-3285.

Herrmann LJ, Heinze B, Fassnacht M, Willenberg HS, Quinkler M, Reisch N *et al* (2012). TP53 germline mutations in adult patients with adrenocortical carcinoma. *The Journal of clinical endocrinology and metabolism* **97:** E476-485.

Heymann S, Delaloge S, Rahal A, Caron O, Frebourg T, Barreau L *et al* (2010). Radio-induced malignancies after breast cancer postoperative radiotherapy in patients with Li-Fraumeni syndrome. *Radiat Oncol* **5:** 104.

Hinds P, Finlay C, Levine AJ (1989). Mutation is required to activate the p53 gene for cooperation with the ras oncogene and transformation. *Journal of virology* **63:** 739-746.

Hirao A, Kong YY, Matsuoka S, Wakeham A, Ruland J, Yoshida H *et al* (2000). DNA damage-induced activation of p53 by the checkpoint kinase Chk2. *Science* **287:** 1824-1827.

Hisada M, Garber JE, Fung CY, Fraumeni JF, Jr., Li FP (1998). Multiple primary cancers in families with Li-Fraumeni syndrome. *Journal of the National Cancer Institute* **90:** 606-611.

Hogdall EV, Kjaer SK, Glud E, Christensen L, Blaakaer J, Vuust J *et al* (2003). Evaluation of a polymorphism in intron 2 of the p53 gene in ovarian cancer patients. From the Danish "Malova" Ovarian Cancer Study. *Anticancer research* **23:** 3397-3404.

Hoh J, Jin S, Parrado T, Edington J, Levine AJ, Ott J (2002). The p53MH algorithm and its application in detecting p53-responsive genes. *Proceedings of the National Academy of Sciences of the United States of America* **99:** 8467-8472.

Hollander MC, Alamo I, Jackman J, Wang MG, McBride OW, Fornace AJ, Jr. (1993). Analysis of the mammalian gadd45 gene and its response to DNA damage. *The Journal of biological chemistry* **268:** 24385-24393.

Hollstein M, Sidransky D, Vogelstein B, Harris CC (1991). p53 mutations in human cancers. *Science* **253:** 49-53.

Honda R, Tanaka H, Yasuda H (1997). Oncoprotein MDM2 is a ubiquitin ligase E3 for tumor suppressor p53. *FEBS letters* **420:** 25-27.

Honda R, Yasuda H (1999). Association of p19(ARF) with Mdm2 inhibits ubiquitin ligase activity of Mdm2 for tumor suppressor p53. *The EMBO journal* **18:** 22-27.

Horvath MM, Wang X, Resnick MA, Bell DA (2007). Divergent evolution of human p53 binding sites: cell cycle versus apoptosis. *PLoS genetics* **3:** e127.

Hsu CH, Lee SC, Yang YC, Lian ST, Shin SJ, Lin SR (2001). The p53 codon 249 mutant--derived from human functional adrenal tumors--can modify the cell shape of normal adrenocortical transfected cells. *Cancer letters* **170:** 63-71.

Hu J, Lutz CS, Wilusz J, Tian B (2005). Bioinformatic identification of candidate cis-regulatory elements involved in human mRNA polyadenylation. *RNA* **11:** 1485-1493.

Hu W, Feng Z, Levine AJ (2009). The regulation of human reproduction by p53 and its pathway. *Cell Cycle* **8:** 3621-3622.

Hu W, Chan CS, Wu R, Zhang C, Sun Y, Song JS *et al* (2010a). Negative regulation of tumor suppressor p53 by microRNA miR-504. *Molecular cell* **38:** 689-699.

Hu Z, Ma H, Lu D, Qian J, Zhou J, Chen Y *et al* (2006). Genetic variants in the MDM2 promoter and lung cancer risk in a Chinese population. *International journal of cancer Journal international du cancer* **118:** 1275-1278.

Hu Z, Li X, Qu X, He Y, Ring BZ, Song E *et al* (2010b). Intron 3 16 bp duplication polymorphism of TP53 contributes to cancer susceptibility: a meta-analysis. *Carcinogenesis* **31:** 643-647.

Hu Z, Li X, Yuan R, Ring BZ, Su L (2010c). Three common TP53 polymorphisms in susceptibility to breast cancer, evidence from meta-analysis. *Breast cancer research and treatment* **120:** 705-714.

Huang J, Perez-Burgos L, Placek BJ, Sengupta R, Richter M, Dorsey JA *et al* (2006). Repression of p53 activity by Smyd2-mediated methylation. *Nature* **444:** 629-632.

Huang L, Yan Z, Liao X, Li Y, Yang J, Wang ZG *et al* (2011). The p53 inhibitors MDM2/MDMX complex is required for control of p53 activity in vivo. *Proceedings of the National Academy of Sciences of the United States of America* **108:** 12001-12006.

Hung RJ, Boffetta P, Canzian F, Moullan N, Szeszenia-Dabrowska N, Zaridze D *et al* (2006). Sequence variants in cell cycle control pathway, X-ray exposure, and lung cancer risk: a multicenter case-control study in Central Europe. *Cancer research* **66:** 8280-8286.

Hupp TR, Meek DW, Midgley CA, Lane DP (1992). Regulation of the specific DNA binding function of p53. *Cell* **71:** 875-886.

Hupp TR, Sparks A, Lane DP (1995). Small peptides activate the latent sequence-specific DNA binding function of p53. *Cell* **83:** 237-245.

Huppert JL, Balasubramanian S (2005). Prevalence of quadruplexes in the human genome. *Nucleic acids research* **33:** 2908-2916.

Huppert JL, Balasubramanian S (2007). G-quadruplexes in promoters throughout the human genome. *Nucleic acids research* **35:** 406-413.

Huppert JL (2008a). Four-stranded nucleic acids: structure, function and targeting of G-quadruplexes. *Chemical Society reviews* **37:** 1375-1384.

Huppert JL (2008b). Hunting G-quadruplexes. *Biochimie* **90:** 1140-1148.

Huppert JL, Bugaut A, Kumari S, Balasubramanian S (2008). G-quadruplexes: the beginning and end of UTRs. *Nucleic acids research* **36:** 6260-6268.

Huppert JL (2010). Structure, location and interactions of G-quadruplexes. *The FEBS journal* **277:** 3452-3458.

Hussain SP, Harris CC (1998). Molecular epidemiology of human cancer. *Recent results in cancer research Fortschritte der Krebsforschung Progres dans les recherches sur le cancer* **154:** 22-36.

Inga A, Scott G, Monti P, Aprile A, Abbondandolo A, Burns PA *et al* (1998). Ultraviolet-light induced p53 mutational spectrum in yeast is indistinguishable from p53 mutations in human skin cancer. *Carcinogenesis* **19:** 741-746.

Innocente SA, Abrahamson JL, Cogswell JP, Lee JM (1999). p53 regulates a G2 checkpoint through cyclin B1. *Proceedings of the National Academy of Sciences of the United States of America* **96:** 2147-2152.

Isobe M, Emanuel BS, Givol D, Oren M, Croce CM (1986). Localization of gene for human p53 tumour antigen to band 17p13. *Nature* **320:** 84-85.

Ito A, Kawaguchi Y, Lai CH, Kovacs JJ, Higashimoto Y, Appella E *et al* (2002). MDM2-HDAC1-mediated deacetylation of p53 is required for its degradation. *The EMBO journal* **21:** 6236-6245.

Iyer NG, Ozdag H, Caldas C (2004). p300/CBP and cancer. *Oncogene* **23:** 4225-4231.

Jackson MW, Berberich SJ (2000). MdmX protects p53 from Mdm2-mediated degradation. *Molecular and cellular biology* **20:** 1001-1007.

Jeffrey PD, Gorina S, Pavletich NP (1995). Crystal structure of the tetramerization domain of the p53 tumor suppressor at 1.7 angstroms. *Science* **267:** 1498-1502.

Jenkins LM, Durell SR, Mazur SJ, Appella E (2012). p53 N-terminal phosphorylation: a defining layer of complex regulation. *Carcinogenesis* **33:** 1441-1449.

Jeong BS, Hu W, Belyi V, Rabadan R, Levine AJ (2010). Differential levels of transcription of p53-regulated genes by the arginine/proline polymorphism: p53 with arginine at codon 72 favors apoptosis. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology* **24:** 1347-1353.

Jha P, Jha P, Pathak P, Chosdol K, Suri V, Sharma MC *et al* (2011). TP53 polymorphisms in gliomas from Indian patients: Study of codon 72 genotype, rs1642785, rs1800370 and 16 base pair insertion in intron-3. *Experimental and molecular pathology* **90:** 167-172.

Jiang D, Brady CA, Johnson TM, Lee EY, Park EJ, Scott MP *et al* (2011). Full p53 transcriptional activation potential is dispensable for tumor suppression in diverse lineages. *Proceedings of the National Academy of Sciences of the United States of America* **108:** 17123-17128.

Jin Y, Dai MS, Lu SZ, Xu Y, Luo Z, Zhao Y *et al* (2006). 14-3-3gamma binds to MDMX that is phosphorylated by UV-activated Chk1, resulting in p53 activation. *The EMBO journal* **25:** 1207-1218.

Johnson TM, Yu ZX, Ferrans VJ, Lowenstein RA, Finkel T (1996). Reactive oxygen species are downstream mediators of p53-dependent apoptosis. *Proceedings of the National Academy of Sciences of the United States of America* **93:** 11848-11852.

Jones SN, Roe AE, Donehower LA, Bradley A (1995). Rescue of embryonic lethality in Mdm2-deficient mice by absence of p53. *Nature* **378:** 206-208.

Jones SN, Hancock AR, Vogel H, Donehower LA, Bradley A (1998). Overexpression of Mdm2 in mice reveals a p53-independent role for Mdm2 in tumorigenesis. *Proceedings of the National Academy of Sciences of the United States of America* **95:** 15608-15612.

Kaeser MD, Iggo RD (2002). Chromatin immunoprecipitation analysis fails to support the latency model for regulation of p53 DNA binding activity in vivo. *Proceedings of the National Academy of Sciences of the United States of America* **99:** 95-100.

Kamijo T, Zindy F, Roussel MF, Quelle DE, Downing JR, Ashmun RA *et al* (1997). Tumor suppression at the mouse INK4a locus mediated by the alternative reading frame product p19ARF. *Cell* **91:** 649-659.

Kamijo T, Weber JD, Zambetti G, Zindy F, Roussel MF, Sherr CJ (1998). Functional and physical interactions of the ARF tumor suppressor with p53 and Mdm2. *Proceedings of the National Academy of Sciences of the United States of America* **95:** 8292-8297.

Kang HJ, Feng Z, Sun Y, Atwal G, Murphy ME, Rebbeck TR *et al* (2009). Single-nucleotide polymorphisms in the p53 pathway regulate fertility in humans. *Proceedings of the National Academy of Sciences of the United States of America* **106:** 9761-9766.

Kastan MB, Zhan Q, el-Deiry WS, Carrier F, Jacks T, Walsh WV *et al* (1992). A mammalian cell cycle checkpoint pathway utilizing p53 and GADD45 is defective in ataxia-telangiectasia. *Cell* **71:** 587-597.

Kastan MB, Bartek J (2004). Cell-cycle checkpoints and cancer. *Nature* **432:** 316-323.

Kato S, Han SY, Liu W, Otsuka K, Shibata H, Kanamaru R *et al* (2003). Understanding the function-structure and function-mutation relationships of p53 tumor suppressor protein by high-resolution missense mutation analysis. *Proceedings of the National Academy of Sciences of the United States of America* **100:** 8424-8429.

Katsumoto T, Higaki K, Ohno K, Onodera K (1995). Cell-cycle dependent biosynthesis and localization of p53 protein in untransformed human cells. *Biology of the cell / under the auspices of the European Cell Biology Organization* **84:** 167-173.

Kawai H, Lopez-Pajares V, Kim MM, Wiederschain D, Yuan ZM (2007). RING domain-mediated interaction is a requirement for MDM2's E3 ligase activity. *Cancer research* **67:** 6026-6030.

Kern SE, Kinzler KW, Baker SJ, Nigro JM, Rotter V, Levine AJ *et al* (1991). Mutant p53 proteins bind DNA abnormally in vitro. *Oncogene* **6:** 131-136.

Kern SE, Pietenpol JA, Thiagalingam S, Seymour A, Kinzler KW, Vogelstein B (1992). Oncogenic forms of p53 inhibit p53-regulated gene expression. *Science* **256:** 827-830.

Khanna KK, Keating KE, Kozlov S, Scott S, Gatei M, Hobson K *et al* (1998). ATM associates with and phosphorylates p53: mapping the region of interaction. *Nature genetics* **20:** 398-400.

Kim E, Albrechtsen N, Deppert W (1997). DNA-conformation is an important determinant of sequence-specific DNA binding by tumor suppressor p53. *Oncogene* **15:** 857-869.

Kim H, Kim K, Choi J, Heo K, Baek HJ, Roeder RG *et al* (2012). p53 requires an intact C-terminal domain for DNA binding and transactivation. *Journal of molecular biology* **415:** 843-854.

Kim IS, Kim DH, Han SM, Chin MU, Nam HJ, Cho HP *et al* (2000). Truncated form of importin alpha identified in breast cancer cell inhibits nuclear import of p53. *The Journal of biological chemistry* **275:** 23139-23145.

Kim MY, Vankayalapati H, Shin-Ya K, Wierzba K, Hurley LH (2002). Telomestatin, a potent telomerase inhibitor that interacts quite specifically with the human telomeric intramolecular g-quadruplex. *Journal of the American Chemical Society* **124:** 2098-2099.

Kim NW, Piatyszek MA, Prowse KR, Harley CB, West MD, Ho PL *et al* (1994). Specific association of human telomerase activity with immortal cells and cancer. *Science* **266:** 2011-2015.

Kitayner M, Rozenberg H, Kessler N, Rabinovich D, Shaulov L, Haran TE *et al* (2006). Structural basis of DNA recognition by p53 tetramers. *Molecular cell* **22:** 741-753.

Kleinerman RA (2009). Radiation-sensitive genetically susceptible pediatric sub-populations. *Pediatric radiology* **39 Suppl 1:** S27-31.

Klotzsche O, Etzrodt D, Hohenberg H, Bohn W, Deppert W (1998). Cytoplasmic retention of mutant tsp53 is dependent on an intermediate filament protein (vimentin) scaffold. *Oncogene* **16:** 3423-3434.

Knights CD, Catania J, Di Giovanni S, Muratoglu S, Perez R, Swartzbeck A *et al* (2006). Distinct p53 acetylation cassettes differentially influence gene-expression patterns and cell fate. *The Journal of cell biology* **173:** 533-544.

Kogan-Sakin I, Tabach Y, Buganim Y, Molchadsky A, Solomon H, Madar S *et al* (2011). Mutant p53(R175H) upregulates Twist1 expression and promotes epithelial-mesenchymal transition in immortalized prostate cells. *Cell death and differentiation* **18:** 271-281.

Koumenis C, Alarcon R, Hammond E, Sutphin P, Hoffman W, Murphy M *et al* (2001). Regulation of p53 by hypoxia: dissociation of transcriptional repression and apoptosis from p53-dependent transactivation. *Molecular and cellular biology* **21:** 1297-1310.

Kouzine F, Sanford S, Elisha-Feil Z, Levens D (2008). The functional response of upstream DNA to dynamic supercoiling in vivo. *Nature structural & molecular biology* **15:** 146-154.

Kraiss S, Quaiser A, Oren M, Montenarh M (1988). Oligomerization of oncoprotein p53. *Journal of virology* **62:** 4737-4744.

Kruse JP, Gu W (2008). SnapShot: p53 posttranslational modifications. *Cell* **133:** 930-930 e931.

Kruse JP, Gu W (2009). Modes of p53 regulation. *Cell* **137:** 609-622.

Kubbutat MH, Jones SN, Vousden KH (1997). Regulation of p53 stability by Mdm2. *Nature* **387:** 299-303.

Kubbutat MH, Vousden KH (1998). Keeping an old friend under control: regulation of p53 stability. *Molecular medicine today* **4:** 250-256.

Kubbutat MH, Ludwig RL, Levine AJ, Vousden KH (1999). Analysis of the degradation function of Mdm2. *Cell growth & differentiation : the molecular biology journal of the American Association for Cancer Research* **10:** 87-92.

Kudo N, Wolff B, Sekimoto T, Schreiner EP, Yoneda Y, Yanagida M *et al* (1998). Leptomycin B inhibition of signal-mediated nuclear export by direct binding to CRM1. *Experimental cell research* **242:** 540-547.

Kulawiec M, Ayyasamy V, Singh KK (2009). p53 regulates mtDNA copy number and mitocheckpoint pathway. *Journal of carcinogenesis* **8:** 8.

Kumar M, Lu Z, Takwi AA, Chen W, Callander NS, Ramos KS *et al* (2011). Negative regulation of the tumor suppressor p53 gene by microRNAs. *Oncogene* **30:** 843-853.

Kumari S, Bugaut A, Huppert JL, Balasubramanian S (2007). An RNA G-quadruplex in the 5' UTR of the NRAS proto-oncogene modulates translation. *Nature chemical biology* **3:** 218-221.

Lamb P, Crawford L (1986). Characterization of the human p53 gene. *Molecular and cellular biology* **6:** 1379-1385.

Lambert PF, Kashanchi F, Radonovich MF, Shiekhattar R, Brady JN (1998). Phosphorylation of p53 serine 15 increases interaction with CBP. *The Journal of biological chemistry* **273:** 33048-33053.

Lane D, Levine A (2010). p53 Research: the past thirty years and the next thirty years. *Cold Spring Harbor perspectives in biology* **2:** a000893.

Lane DP, Crawford LV (1979). T antigen is bound to a host protein in SV40-transformed cells. *Nature* **278:** 261-263.

Lane DP (1992). Cancer. p53, guardian of the genome. *Nature* **358:** 15-16.

Lazar V, Hazard F, Bertin F, Janin N, Bellet D, Bressac B (1993). Simple sequence repeat polymorphism within the p53 gene. *Oncogene* **8:** 1703-1705.

Le MT, Teh C, Shyh-Chang N, Xie H, Zhou B, Korzh V *et al* (2009). MicroRNA-125b is a novel negative regulator of p53. *Genes & development* **23:** 862-876.

Le MT, Shyh-Chang N, Khaw SL, Chin L, Teh C, Tay J *et al* (2011). Conserved regulation of p53 network dosage by microRNA-125b occurs through evolving miRNA-target gene pairs. *PLoS genetics* **7:** e1002242.

Lebedeva MA, Eaton JS, Shadel GS (2009). Loss of p53 causes mitochondrial DNA depletion and altered mitochondrial reactive oxygen species homeostasis. *Biochimica et biophysica acta* **1787:** 328-334.

LeBron C, Chen L, Gilkes DM, Chen J (2006). Regulation of MDMX nuclear import and degradation by Chk2 and 14-3-3. *The EMBO journal* **25:** 1196-1206.

Lee H, Mok KH, Muhandiram R, Park KH, Suk JE, Kim DH *et al* (2000). Local structural elements in the mostly unstructured transcriptional activation domain of human p53. *The Journal of biological chemistry* **275:** 29426-29432.

Lee S, Elenbaas B, Levine A, Griffith J (1995). p53 and its 14 kDa C-terminal domain recognize primary DNA damage in the form of insertion/deletion mismatches. *Cell* **81:** 1013-1020.

Lee W, Harvey TS, Yin Y, Yau P, Litchfield D, Arrowsmith CH (1994). Solution structure of the tetrameric minimum transforming domain of p53. *Nature structural biology* **1:** 877-890.

Leng RP, Lin Y, Ma W, Wu H, Lemmers B, Chung S *et al* (2003). Pirh2, a p53-induced ubiquitin-protein ligase, promotes p53 degradation. *Cell* **112:** 779-791.

Leveille N, Elkon R, Davalos V, Manoharan V, Hollingworth D, Oude Vrielink J *et al* (2011). Selective inhibition of microRNA accessibility by RBM38 is required for p53 activity. *Nature communications* **2:** 513.

Li C, Chen L, Chen J (2002a). DNA damage induces MDMX nuclear translocation by p53-dependent and -independent mechanisms. *Molecular and cellular biology* **22:** 7562-7571.

Li FP, Fraumeni JF, Jr. (1969). Rhabdomyosarcoma in children: epidemiologic study and identification of a familial cancer syndrome. *Journal of the National Cancer Institute* **43:** 1365-1373.

Li FP, Fraumeni JF, Jr., Mulvihill JJ, Blattner WA, Dreyfus MG, Tucker MA *et al* (1988). A cancer family syndrome in twenty-four kindreds. *Cancer research* **48:** 5358-5362.

Li FP (1995). Identification and management of inherited cancer susceptibility. *Environmental health perspectives* **103 Suppl 8:** 297-300.

Li M, Chen D, Shiloh A, Luo J, Nikolaev AY, Qin J *et al* (2002b). Deubiquitination of p53 by HAUSP is an important pathway for p53 stabilization. *Nature* **416:** 648-653.

Li M, Brooks CL, Wu-Baer F, Chen D, Baer R, Gu W (2003). Mono- versus polyubiquitination: differential control of p53 fate by Mdm2. *Science* **302:** 1972-1975.

Li M, Brooks CL, Kon N, Gu W (2004). A dynamic role of HAUSP in the p53-Mdm2 pathway. *Molecular cell* **13:** 879-886.

Li N, Kaur S, Greshock J, Lassus H, Zhong X, Wang Y *et al* (2012a). A combined array-based comparative genomic hybridization and functional library screening approach identifies mir-30d as an oncomir in cancer. *Cancer research* **72:** 154-164.

Li T, Kon N, Jiang L, Tan M, Ludwig T, Zhao Y *et al* (2012b). Tumor suppression in the absence of p53-mediated cell-cycle arrest, apoptosis, and senescence. *Cell* **149:** 1269-1283.

Li X, Coffino P (1996). High-risk human papillomavirus E6 protein has two distinct binding sites within p53, of which only one determines degradation. *Journal of virology* **70:** 4509-4516.

Li Y, Qiu LX, Shen XK, Lv XJ, Qian XP, Song Y (2009). A meta-analysis of TP53 codon 72 polymorphism and lung cancer risk: evidence from 15,857 subjects. *Lung Cancer* **66:** 15-21.

Li Y, Gordon MW, Xu-Monette ZY, Visco C, Tzankov A, Zou D *et al* (2013). Single nucleotide variation in the TP53 3' untranslated region in diffuse large B-cell lymphoma treated with rituximab-CHOP: a report from the International DLBCL Rituximab-CHOP Consortium Program. *Blood* **121:** 4529-4540.

Liang SH, Hong D, Clarke MF (1998). Cooperation of a single lysine mutation and a C-terminal domain in the cytoplasmic sequestration of the p53 protein. *The Journal of biological chemistry* **273:** 19817-19821.

Liang SH, Clarke MF (1999a). A bipartite nuclear localization signal is required for p53 nuclear import regulated by a carboxyl-terminal domain. *The Journal of biological chemistry* **274:** 32699-32703.

Liang SH, Clarke MF (1999b). The nuclear import of p53 is determined by the presence of a basic domain and its relative position to the nuclear localization signal. *Oncogene* **18:** 2163-2166.

Lim SK, Shin JM, Kim YS, Baek KH (2004). Identification and characterization of murine mHAUSP encoding a deubiquitinating enzyme that regulates the status of p53 ubiquitination. *International journal of oncology* **24:** 357-364.

Limacher JM, Frebourg T, Natarajan-Ame S, Bergerat JP (2001). Two metachronous tumors in the radiotherapy fields of a patient with Li-Fraumeni syndrome. *International journal of cancer Journal international du cancer* **96:** 238-242.

Lin AW, Barradas M, Stone JC, van Aelst L, Serrano M, Lowe SW (1998). Premature senescence involving p53 and p16 is activated in response to constitutive MEK/MAPK mitogenic signaling. *Genes & development* **12:** 3008-3019.

Lin D, Shields MT, Ullrich SJ, Appella E, Mercer WE (1992). Growth arrest induced by wild-type p53 protein blocks cells prior to or near the restriction point in late G1 phase. *Proceedings of the National Academy of Sciences of the United States of America* **89:** 9210-9214.

Lin Y, Ma W, Benchimol S (2000). Pidd, a new death-domain-containing protein, is induced by p53 and promotes apoptosis. *Nature genetics* **26:** 122-127.

Linares LK, Hengstermann A, Ciechanover A, Muller S, Scheffner M (2003). HdmX stimulates Hdm2-mediated ubiquitination and degradation of p53. *Proceedings of the National Academy of Sciences of the United States of America* **100:** 12009-12014.

Lind H, Zienolddiny S, Ekstrom PO, Skaug V, Haugen A (2006). Association of a functional polymorphism in the promoter of the MDM2 gene with risk of nonsmall cell lung cancer. *International journal of cancer Journal international du cancer* **119:** 718-721.

Linzer DI, Levine AJ (1979). Characterization of a 54K dalton cellular SV40 tumor antigen present in SV40-transformed cells and uninfected embryonal carcinoma cells. *Cell* **17:** 43-52.

Lipps HJ, Rhodes D (2009). G-quadruplex structures: in vivo evidence and function. *Trends in cell biology* **19:** 414-422.

Litviakov NV, Denisov EV, Takhauov RM, Karpov AB, Skobel'skaja EV, Vasil'eva EO *et al* (2010). Association between TP53 gene ARG72PRO polymorphism and chromosome aberrations in human cancers. *Molecular carcinogenesis* **49:** 521-524.

Liu DP, Song H, Xu Y (2010a). A common gain of function of p53 cancer mutants in inducing genetic instability. *Oncogene* **29:** 949-956.

Liu G, Xirodimas DP (2010). NUB1 promotes cytoplasmic localization of p53 through cooperation of the NEDD8 and ubiquitin pathways. *Oncogene* **29:** 2252-2261.

Liu J, Yu G, Zhao Y, Zhao D, Wang Y, Wang L *et al* (2010b). REGgamma modulates p53 activity by regulating its cellular localization. *Journal of cell science* **123:** 4076-4084.

Liu K, Ling S, Lin WC (2011). TopBP1 mediates mutant p53 gain of function through NF-Y and p63/p73. *Molecular and cellular biology* **31:** 4464-4481.

Liu X, Miller CW, Koeffler PH, Berk AJ (1993). The p53 activation domain binds the TATA box-binding polypeptide in Holo-TFIID, and a neighboring p53 domain inhibits transcription. *Molecular and cellular biology* **13:** 3291-3300.

Llanos S, Clark PA, Rowe J, Peters G (2001). Stabilization of p53 by p14ARF without relocation of MDM2 to the nucleolus. *Nature cell biology* **3:** 445-452.

Lohrum MA, Woods DB, Ludwig RL, Balint E, Vousden KH (2001). C-terminal ubiquitination of p53 contributes to nuclear export. *Molecular and cellular biology* **21:** 8521-8532.

Lomax ME, Barnes DM, Hupp TR, Picksley SM, Camplejohn RS (1998). Characterization of p53 oligomerization domain mutations isolated from Li-Fraumeni and Li-Fraumeni like family members. *Oncogene* **17:** 643-649.

Lomazzi M, Moroni MC, Jensen MR, Frittoli E, Helin K (2002). Suppression of the p53- or pRB-mediated G1 checkpoint is required for E2F-induced S-phase entry. *Nature genetics* **31:** 190-194.

Lopez-Girona A, Kanoh J, Russell P (2001). Nuclear exclusion of Cdc25 is not required for the DNA damage checkpoint in fission yeast. *Current biology : CB* **11:** 50-54.

Lu H, Levine AJ (1995). Human TAFII31 protein is a transcriptional coactivator of the p53 protein. *Proceedings of the National Academy of Sciences of the United States of America* **92:** 5154-5158.

Lu PH, Wei MX, Li C, Shen W, Chen MB (2011). Need for clarification of data in a recent meta-analysis about TP53 codon 72 polymorphism and cancer susceptibility. *Carcinogenesis* **32:** 443; author reply 444.

Lum SS, Chua HW, Li H, Li WF, Rao N, Wei J *et al* (2008). MDM2 SNP309 G allele increases risk but the T allele is associated with earlier onset age of sporadic breast cancers in the Chinese population. *Carcinogenesis* **29:** 754-761.

Lynch HT, Mulcahy GM, Harris RE, Guirgis HA, Lynch JF (1978). Genetic and pathologic findings in a kindred with hereditary sarcoma, breast cancer, brain tumors, leukemia, lung, laryngeal, and adrenal cortical carcinoma. *Cancer* **41:** 2055-2064.

Ma Y, Bian J, Cao H (2013). MDM2 SNP309 rs2279744 polymorphism and gastric cancer risk: a meta-analysis. *PloS one* **8:** e56918.

Macedo GS, Lisboa da Motta L, Giacomazzi J, Netto CB, Manfredini V, Vanzin CS *et al* (2012). Increased oxidative damage in carriers of the germline TP53 p.R337H mutation. *PloS one* **7:** e47010.

Macip S, Igarashi M, Berggren P, Yu J, Lee SW, Aaronson SA (2003). Influence of induced reactive oxygen species in p53-mediated cell fate decisions. *Molecular and cellular biology* **23:** 8576-8585.

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. *Molecular cell* **33:** 462-471.

Mahmoudi S, Henriksson S, Weibrecht I, Smith S, Soderberg O, Stromblad S *et al* (2010). WRAP53 is essential for Cajal body formation and for targeting the survival of motor neuron complex to Cajal bodies. *PLoS biology* **8:** e1000521.

Maiuri MC, Malik SA, Morselli E, Kepp O, Criollo A, Mouchel PL *et al* (2009). Stimulation of autophagy by the p53 target gene Sestrin2. *Cell Cycle* **8:** 1571-1576.

Malkin D, Li FP, Strong LC, Fraumeni JF, Jr., Nelson CE, Kim DH *et al* (1990). Germ line p53 mutations in a familial syndrome of breast cancer, sarcomas, and other neoplasms. *Science* **250:** 1233-1238.

Maltzman W, Czyzyk L (1984). UV irradiation stimulates levels of p53 cellular tumor antigen in nontransformed mouse cells. *Molecular and cellular biology* **4:** 1689-1694.

Marcel V, Palmero EI, Falagan-Lotsch P, Martel-Planche G, Ashton-Prolla P, Olivier M *et al* (2009). TP53 PIN3 and MDM2 SNP309 polymorphisms as genetic modifiers in the Li-Fraumeni syndrome: impact on age at first diagnosis. *Journal of medical genetics* **46:** 766-772.

Marcel V, Dichtel-Danjoy ML, Sagne C, Hafsi H, Ma D, Ortiz-Cuaran S *et al* (2011). Biological functions of p53 isoforms through evolution: lessons from animal and cellular models. *Cell death and differentiation* **18:** 1815-1824.

Marchenko ND, Zaika A, Moll UM (2000). Death signal-induced localization of p53 protein to mitochondria. A potential role in apoptotic signaling. *The Journal of biological chemistry* **275:** 16202-16212.

Marin MC, Jost CA, Brooks LA, Irwin MS, O'Nions J, Tidy JA *et al* (2000). A common polymorphism acts as an intragenic modifier of mutant p53 behaviour. *Nature genetics* **25:** 47-54.

Matakidou A, Eisen T, Houlston RS (2003). TP53 polymorphisms and lung cancer risk: a systematic review and meta-analysis. *Mutagenesis* **18:** 377-385.

Mateu MG, Fersht AR (1998). Nine hydrophobic side chains are key determinants of the thermodynamic stability and oligomerization status of tumour suppressor p53 tetramerization domain. *The EMBO journal* **17:** 2748-2758.

Mateu MG, Sanchez Del Pino MM, Fersht AR (1999). Mechanism of folding and assembly of a small tetrameric protein domain from tumor suppressor p53. *Nature structural biology* **6:** 191-198.

Mathupala SP, Ko YH, Pedersen PL (2006). Hexokinase II: cancer's double-edged sword acting as both facilitator and gatekeeper of malignancy when bound to mitochondria. *Oncogene* **25:** 4777-4786.

Matlashewski GJ, Tuck S, Pim D, Lamb P, Schneider J, Crawford LV (1987). Primary structure polymorphism at amino acid residue 72 of human p53. *Molecular and cellular biology* **7:** 961-963.

Matoba S, Kang JG, Patino WD, Wragg A, Boehm M, Gavrilova O *et al* (2006). p53 regulates mitochondrial respiration. *Science* **312:** 1650-1653.

May G, Fischer H, Zang KD (1979). SV40-related T-antigen expression in human meningiomas with normal and G-22-monosomic karyotype. *The Journal of general virology* **43:** 697-700.

May P, May E (1999). Twenty years of p53 research: structural and functional aspects of the p53 protein. *Oncogene* **18:** 7621-7636.

Mazur SJ, Sakaguchi K, Appella E, Wang XW, Harris CC, Bohr VA (1999). Preferential binding of tumor suppressor p53 to positively or negatively supercoiled DNA involves the C-terminal domain. *Journal of molecular biology* **292:** 241-249.

McBride OW, Merry D, Givol D (1986). The gene for human p53 cellular tumor antigen is located on chromosome 17 short arm (17p13). *Proceedings of the National Academy of Sciences of the United States of America* **83:** 130-134.

McLure KG, Lee PW (1998). How p53 binds DNA as a tetramer. *The EMBO journal* **17:** 3342-3350.

Mechanic LE, Bowman ED, Welsh JA, Khan MA, Hagiwara N, Enewold L *et al* (2007). Common genetic variation in TP53 is associated with lung cancer risk and prognosis in African Americans and somatic mutations in lung tumors. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology* **16:** 214-222.

Medema RH, Macurek L (2012). Checkpoint control and cancer. *Oncogene* **31:** 2601-2613.

Melchior F, Hengst L (2002). SUMO-1 and p53. *Cell Cycle* **1:** 245-249.

Melko M, Bardoni B (2010). The role of G-quadruplex in RNA metabolism: involvement of FMRP and FMR2P. *Biochimie* **92:** 919-926.

Mellert H, Sykes SM, Murphy ME, McMahon SB (2007). The ARF/oncogene pathway activates p53 acetylation within the DNA binding domain. *Cell Cycle* **6:** 1304-1306.

Merle P, Evrard B, Petitjean A, Lehn JM, Teulade-Fichou MP, Chautard E *et al* (2011). Telomere targeting with a new G4 ligand enhances radiation-induced killing of human glioblastoma cells. *Molecular cancer therapeutics* **10:** 1784-1795.

Michael D, Oren M (2003). The p53-Mdm2 module and the ubiquitin system. *Seminars in cancer biology* **13:** 49-58.

Michalak E, Villunger A, Erlacher M, Strasser A (2005). Death squads enlisted by the tumour suppressor p53. *Biochemical and biophysical research communications* **331:** 786-798.

Mihara M, Shintani S, Nakashiro K, Hamakawa H (2003). Flavopiridol, a cyclin dependent kinase (CDK) inhibitor, induces apoptosis by regulating Bcl-x in oral cancer cells. *Oral oncology* **39:** 49-55.

Millevoi S, Moine H, Vagner S (2012). G-quadruplexes in RNA biology. *Wiley interdisciplinary reviews RNA* **3:** 495-507.

Milner J, Medcalf EA, Cook AC (1991). Tumor suppressor p53: analysis of wild-type and mutant p53 complexes. *Molecular and cellular biology* **11:** 12-19.

Miyashita T, Krajewski S, Krajewska M, Wang HG, Lin HK, Liebermann DA *et al* (1994). Tumor suppressor p53 is a regulator of bcl-2 and bax gene expression in vitro and in vivo. *Oncogene* **9:** 1799-1805.

Miyashita T, Reed JC (1995). Tumor suppressor p53 is a direct transcriptional activator of the human bax gene. *Cell* **80:** 293-299.

Moll UM, Ostermeyer AG, Haladay R, Winkfield B, Frazier M, Zambetti G (1996). Cytoplasmic sequestration of wild-type p53 protein impairs the G1 checkpoint after DNA damage. *Molecular and cellular biology* **16:** 1126-1137.

Momand J, Zambetti GP, Olson DC, George D, Levine AJ (1992). The mdm-2 oncogene product forms a complex with the p53 protein and inhibits p53-mediated transactivation. *Cell* **69:** 1237-1245.

Montes de Oca Luna R, Wagner DS, Lozano G (1995). Rescue of early embryonic lethality in mdm2-deficient mice by deletion of p53. *Nature* **378:** 203-206.

Montesano R, Hainaut P, Wild CP (1997). Hepatocellular carcinoma: from gene to public health. *Journal of the National Cancer Institute* **89:** 1844-1851.

Moroni MC, Hickman ES, Lazzerini Denchi E, Caprara G, Colli E, Cecconi F *et al* (2001). Apaf-1 is a transcriptional target for E2F and p53. *Nature cell biology* **3:** 552-558.

Muller PA, Vousden KH (2013). p53 mutations in cancer. *Nature cell biology* **15:** 2-8.

Muller-Tiemann BF, Halazonetis TD, Elting JJ (1998). Identification of an additional negative regulatory region for p53 sequence-specific DNA binding. *Proceedings of the National Academy of Sciences of the United States of America* **95:** 6079-6084.

Nakano K, Vousden KH (2001). PUMA, a novel proapoptotic gene, is induced by p53. *Molecular cell* **7:** 683-694.

Nichols KE, Malkin D, Garber JE, Fraumeni JF, Jr., Li FP (2001). Germ-line p53 mutations predispose to a wide spectrum of early-onset cancers. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology* **10:** 83-87.

Nigro JM, Baker SJ, Preisinger AC, Jessup JM, Hostetter R, Cleary K *et al* (1989). Mutations in the p53 gene occur in diverse human tumour types. *Nature* **342:** 705-708.

Nikolaev AY, Li M, Puskas N, Qin J, Gu W (2003). Parc: a cytoplasmic anchor for p53. *Cell* **112:** 29-40.

Nishida N, Yokobori T, Mimori K, Sudo T, Tanaka F, Shibata K *et al* (2011). MicroRNA miR-125b is a prognostic marker in human colorectal cancer. *International journal of oncology* **38:** 1437-1443.

Nishimori H, Shiratsuchi T, Urano T, Kimura Y, Kiyono K, Tatsumi K *et al* (1997). A novel brain-specific p53-target gene, BAI1, containing thrombospondin type 1 repeats inhibits experimental angiogenesis. *Oncogene* **15:** 2145-2150.

Nyberg KA, Michelson RJ, Putnam CW, Weinert TA (2002). Toward maintaining the genome: DNA damage and replication checkpoints. *Annual review of genetics* **36:** 617-656.

O'Keefe K, Li H, Zhang Y (2003). Nucleocytoplasmic shuttling of p53 is essential for MDM2-mediated cytoplasmic degradation but not ubiquitination. *Molecular and cellular biology* **23:** 6396-6405.

Oda E, Ohki R, Murasawa H, Nemoto J, Shibue T, Yamashita T *et al* (2000). Noxa, a BH3-only member of the Bcl-2 family and candidate mediator of p53-induced apoptosis. *Science* **288:** 1053-1058.

Offer H, Milyavsky M, Erez N, Matas D, Zurer I, Harris CC *et al* (2001a). Structural and functional involvement of p53 in BER in vitro and in vivo. *Oncogene* **20:** 581-589.

Offer H, Zurer I, Banfalvi G, Reha'k M, Falcovitz A, Milyavsky M *et al* (2001b). p53 modulates base excision repair activity in a cell cycle-specific manner after genotoxic stress. *Cancer research* **61:** 88-96.

Ofir M, Hacohen D, Ginsberg D (2011). MiR-15 and miR-16 are direct transcriptional targets of E2F1 that limit E2F-induced proliferation by targeting cyclin E. *Molecular cancer research : MCR* **9:** 440-447.

Ofir-Rosenfeld Y, Boggs K, Michael D, Kastan MB, Oren M (2008). Mdm2 regulates p53 mRNA translation through inhibitory interactions with ribosomal protein L26. *Molecular cell* **32:** 180-189.

Oganesian L, Bryan TM (2007). Physiological relevance of telomeric G-quadruplex formation: a potential drug target. *BioEssays : news and reviews in molecular, cellular and developmental biology* **29:** 155-165.

Ohki R, Nemoto J, Murasawa H, Oda E, Inazawa J, Tanaka N *et al* (2000). Reprimo, a new candidate mediator of the p53-mediated cell cycle arrest at the G2 phase. *The Journal of biological chemistry* **275:** 22627-22630.

Ohki R, Kawase T, Ohta T, Ichikawa H, Taya Y (2007). Dissecting functional roles of p53 N-terminal transactivation domains by microarray expression analysis. *Cancer science* **98:** 189-200.

Okamoto K, Kashima K, Pereg Y, Ishida M, Yamazaki S, Nota A *et al* (2005). DNA damage-induced phosphorylation of MdmX at serine 367 activates p53 by targeting MdmX for Mdm2-dependent degradation. *Molecular and cellular biology* **25:** 9608-9620.

Okamoto K, Taya Y, Nakagama H (2009). Mdmx enhances p53 ubiquitination by altering the substrate preference of the Mdm2 ubiquitin ligase. *FEBS letters* **583:** 2710-2714.

Okoshi R, Ozaki T, Yamamoto H, Ando K, Koida N, Ono S *et al* (2008). Activation of AMP-activated protein kinase induces p53-dependent apoptotic cell death in response to energetic stress. *The Journal of biological chemistry* **283:** 3979-3987.

Oliner JD, Pietenpol JA, Thiagalingam S, Gyuris J, Kinzler KW, Vogelstein B (1993). Oncoprotein MDM2 conceals the activation domain of tumour suppressor p53. *Nature* **362:** 857-860.

Olivier M, Eeles R, Hollstein M, Khan MA, Harris CC, Hainaut P (2002). The IARC TP53 database: new online mutation analysis and recommendations to users. *Human mutation* **19:** 607-614.

Olivier M, Hollstein M, Hainaut P (2010). TP53 mutations in human cancers: origins, consequences, and clinical use. *Cold Spring Harbor perspectives in biology* **2:** a001008.

Oren M (1992). p53: the ultimate tumor suppressor gene? *FASEB journal : official publication of the Federation of American Societies for Experimental Biology* **6:** 3169-3176.

Oren M (2003). Decision making by p53: life, death and cancer. *Cell death and differentiation* **10:** 431-442.

Ortiz-Cuaran S, Cox D, Villar S, Friesen MD, Durand G, Chabrier A *et al* (2013). Association between TP53 R249S mutation and polymorphisms in TP53 intron 1 in hepatocellular carcinoma. *Genes, chromosomes & cancer* **52:** 912-919.

Ory K, Legros Y, Auguin C, Soussi T (1994). Analysis of the most representative tumour-derived p53 mutants reveals that changes in protein conformation are not correlated with loss of transactivation or inhibition of cell proliferation. *The EMBO journal* **13:** 3496-3504.

Ostermeyer AG, Runko E, Winkfield B, Ahn B, Moll UM (1996). Cytoplasmically sequestered wild-type p53 protein in neuroblastoma is relocated to the nucleus by a C-terminal peptide. *Proceedings of the National Academy of Sciences of the United States of America* **93:** 15190-15194.

Owen-Schaub LB, Angelo LS, Radinsky R, Ware CF, Gesner TG, Bartos DP (1995). Soluble Fas/APO-1 in tumor cells: a potential regulator of apoptosis? *Cancer letters* **94:** 1-8.

Paeschke K, McDonald KR, Zakian VA (2010). Telomeres: structures in need of unwinding. *FEBS letters* **584:** 3760-3772.

Paeschke K, Capra JA, Zakian VA (2011). DNA replication through G-quadruplex motifs is promoted by the Saccharomyces cerevisiae Pif1 DNA helicase. *Cell* **145:** 678-691.

Paeschke K, Bochman ML, Garcia PD, Cejka P, Friedman KL, Kowalczykowski SC *et al* (2013). Pif1 family helicases suppress genome instability at G-quadruplex motifs. *Nature* **497:** 458-462.

Pal S, Datta K, Mukhopadhyay D (2001). Central role of p53 on regulation of vascular permeability factor/vascular endothelial growth factor (VPF/VEGF) expression in mammary carcinoma. *Cancer research* **61:** 6952-6957.

Palmero EI, Schuler-Faccini L, Caleffi M, Achatz MI, Olivier M, Martel-Planche G *et al* (2008). Detection of R337H, a germline TP53 mutation predisposing to multiple cancers, in asymptomatic women participating in a breast cancer screening program in Southern Brazil. *Cancer letters* **261:** 21-25.

Palmero EI, Achatz MI, Ashton-Prolla P, Olivier M, Hainaut P (2010). Tumor protein 53 mutations and inherited cancer: beyond Li-Fraumeni syndrome. *Current opinion in oncology* **22:** 64-69.

Pan J, Sasaki M, Kniewel R, Murakami H, Blitzblau HG, Tischfield SE *et al* (2011). A hierarchical combination of factors shapes the genome-wide topography of yeast meiotic recombination initiation. *Cell* **144:** 719-731.

Pan Y, Chen J (2003). MDM2 promotes ubiquitination and degradation of MDMX. *Molecular and cellular biology* **23:** 5113-5121.

Pant V, Xiong S, Iwakuma T, Quintas-Cardama A, Lozano G (2011). Heterodimerization of Mdm2 and Mdm4 is critical for regulating p53 activity during embryogenesis but dispensable for p53 and Mdm2 stability. *Proceedings of the National Academy of Sciences of the United States of America* **108:** 11995-12000.

Parant J, Chavez-Reyes A, Little NA, Yan W, Reinke V, Jochemsen AG *et al* (2001). Rescue of embryonic lethality in Mdm4-null mice by loss of Trp53 suggests a nonoverlapping pathway with MDM2 to regulate p53. *Nature genetics* **29:** 92-95.

Parks D, Bolinger R, Mann K (1997). Redox state regulates binding of p53 to sequence-specific DNA, but not to non-specific or mismatched DNA. *Nucleic acids research* **25:** 1289-1295.

Paskulin DD, Cunha-Filho JS, Souza CA, Bortolini MC, Hainaut P, Ashton-Prolla P (2012). TP53 PIN3 and PEX4 polymorphisms and infertility associated with endometriosis or with post-in vitro fertilization implantation failure. *Cell death & disease* **3:** e392.

Pavletich NP, Chambers KA, Pabo CO (1993). The DNA-binding domain of p53 contains the four conserved regions and the major mutation hot spots. *Genes & development* **7:** 2556-2564.

Peng CY, Graves PR, Thoma RS, Wu Z, Shaw AS, Piwnica-Worms H (1997). Mitotic and G2 checkpoint control: regulation of 14-3-3 protein binding by phosphorylation of Cdc25C on serine-216. *Science* **277:** 1501-1505.

Pennarun G, Granotier C, Gauthier LR, Gomez D, Hoffschir F, Mandine E *et al* (2005). Apoptosis related to telomere instability and cell cycle alterations in human glioma cells treated by new highly selective G-quadruplex ligands. *Oncogene* **24:** 2917-2928.

Pennarun G, Granotier C, Hoffschir F, Mandine E, Biard D, Gauthier LR *et al* (2008). Role of ATM in the telomere response to the G-quadruplex ligand 360A. *Nucleic acids research* **36:** 1741-1754.

Pepper C, Thomas A, Hoy T, Tighe J, Culligan D, Fegan C *et al* (2003). Leukemic and non-leukemic lymphocytes from patients with Li Fraumeni syndrome demonstrate loss of p53 function, Bcl-2 family dysregulation and intrinsic resistance to conventional chemotherapeutic drugs but not flavopiridol. *Cell Cycle* **2:** 53-58.

Perfettini JL, Castedo M, Nardacci R, Ciccosanti F, Boya P, Roumier T *et al* (2005). Essential role of p53 phosphorylation by p38 MAPK in apoptosis induction by the HIV-1 envelope. *The Journal of experimental medicine* **201:** 279-289.

Perfumo C, Bonelli L, Menichini P, Inga A, Gismondi V, Ciferri E *et al* (2006). Increased risk of colorectal adenomas in Italian subjects carrying the p53 PIN3 A2-Pro72 haplotype. *Digestion* **74:** 228-235.

Petitjean A, Mathe E, Kato S, Ishioka C, Tavtigian SV, Hainaut P *et al* (2007). Impact of mutant p53 functional properties on TP53 mutation patterns and tumor phenotype: lessons from recent developments in the IARC TP53 database. *Human mutation* **28:** 622-629.

Pfeifer GP, Denissenko MF, Olivier M, Tretyakova N, Hecht SS, Hainaut P (2002). Tobacco smoke carcinogens, DNA damage and p53 mutations in smoking-associated cancers. *Oncogene* **21:** 7435-7451.

Phan AT, Luu KN, Patel DJ (2006). Different loop arrangements of intramolecular human telomeric (3+1) G-quadruplexes in K+ solution. *Nucleic acids research* **34:** 5715-5719.

Pharoah PD, Day NE, Caldas C (1999). Somatic mutations in the p53 gene and prognosis in breast cancer: a meta-analysis. *British journal of cancer* **80:** 1968-1973.

Pichiorri F, Suh SS, Rocci A, De Luca L, Taccioli C, Santhanam R *et al* (2010). Downregulation of p53-inducible microRNAs 192, 194, and 215 impairs the p53/MDM2 autoregulatory loop in multiple myeloma development. *Cancer cell* **18:** 367-381.

Picksley SM, Vojtesek B, Sparks A, Lane DP (1994). Immunochemical analysis of the interaction of p53 with MDM2;--fine mapping of the MDM2 binding site on p53 using synthetic peptides. *Oncogene* **9:** 2523-2529.

Pim D, Banks L (2004). p53 polymorphic variants at codon 72 exert different effects on cell cycle progression. *International journal of cancer Journal international du cancer* **108:** 196-199.

Pinto C, Veiga I, Pinheiro M, Peixoto A, Pinto A, Lopes JM *et al* (2009). TP53 germline mutations in Portugal and genetic modifiers of age at cancer onset. *Familial cancer* **8:** 383-390.

Pinto EM, Billerbeck AE, Villares MC, Domenice S, Mendonca BB, Latronico AC (2004). Founder effect for the highly prevalent R337H mutation of tumor suppressor p53 in Brazilian patients with adrenocortical tumors. *Arquivos brasileiros de endocrinologia e metabologia* **48:** 647-650.

Pogribny IP, Pogribna M, Christman JK, James SJ (2000). Single-site methylation within the p53 promoter region reduces gene expression in a reporter gene construct: possible in vivo relevance during tumorigenesis. *Cancer research* **60:** 588-594.

Pogribny IP, James SJ (2002). Reduction of p53 gene expression in human primary hepatocellular carcinoma is associated with promoter region methylation without coding region mutation. *Cancer letters* **176:** 169-174.

Polyak K, Xia Y, Zweier JL, Kinzler KW, Vogelstein B (1997). A model for p53-induced apoptosis. *Nature* **389:** 300-305.

Powell BL, van Staveren IL, Roosken P, Grieu F, Berns EM, Iacopetta B (2002). Associations between common polymorphisms in TP53 and p21WAF1/Cip1 and phenotypic features of breast cancer. *Carcinogenesis* **23:** 311-315.

Poyurovsky MV, Katz C, Laptenko O, Beckerman R, Lokshin M, Ahn J *et al* (2010). The C terminus of p53 binds the N-terminal domain of MDM2. *Nature structural & molecular biology* **17:** 982-989.

Prives C, Hall PA (1999). The p53 pathway. *The Journal of pathology* **187:** 112-126.

Pugh CW, Ratcliffe PJ (2003). Regulation of angiogenesis by hypoxia: role of the HIF system. *Nature medicine* **9:** 677-684.

Qian H, Wang T, Naumovski L, Lopez CD, Brachmann RK (2002). Groups of p53 target genes involved in specific p53 downstream effects cluster into different classes of DNA binding sites. *Oncogene* **21:** 7901-7911.

Qin Q, Baudry M, Liao G, Noniyev A, Galeano J, Bi X (2009). A novel function for p53: regulation of growth cone motility through interaction with Rho kinase. *The Journal of neuroscience : the official journal of the Society for Neuroscience* **29:** 5183-5192.

Quintavalle M, Elia L, Condorelli G, Courtneidge SA (2010). MicroRNA control of podosome formation in vascular smooth muscle cells in vivo and in vitro. *The Journal of cell biology* **189:** 13-22.

Rankin S, Reszka AP, Huppert J, Zloh M, Parkinson GN, Todd AK *et al* (2005). Putative DNA quadruplex formation within the human c-kit oncogene. *Journal of the American Chemical Society* **127:** 10584-10589.

Ravi R, Mookerjee B, Bhujwalla ZM, Sutter CH, Artemov D, Zeng Q *et al* (2000). Regulation of tumor angiogenesis by p53-induced degradation of hypoxia-inducible factor 1alpha. *Genes & development* **14:** 34-44.

Raycroft L, Wu HY, Lozano G (1990). Transcriptional activation by wild-type but not transforming mutants of the p53 anti-oncogene. *Science* **249:** 1049-1051.

Reich NC, Oren M, Levine AJ (1983). Two distinct mechanisms regulate the levels of a cellular tumor antigen, p53. *Molecular and cellular biology* **3:** 2143-2150.

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. *Proceedings of the National Academy of Sciences of the United States of America* **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.

Renaux-Petel M, Sesboue R, Baert-Desurmont S, Vasseur S, Fourneaux S, Bessenay E *et al* (2013). The MDM2 285G-309G haplotype is associated with an earlier age of tumour onset in patients with Li-Fraumeni syndrome. *Familial cancer*.

Retzlaff M, Rohrberg J, Kupper NJ, Lagleder S, Bepperling A, Manzenrieder F *et al* (2013). The regulatory domain stabilizes the p53 tetramer by intersubunit contacts with the DNA binding domain. *Journal of molecular biology* **425:** 144-155.

Ribeiro RC, Sandrini F, Figueiredo B, Zambetti GP, Michalkiewicz E, Lafferty AR *et al* (2001). An inherited p53 mutation that contributes in a tissue-specific manner to pediatric adrenal cortical carcinoma. *Proceedings of the National Academy of Sciences of the United States of America* **98:** 9330-9335.

Ribeyre C, Lopes J, Boule JB, Piazza A, Guedin A, Zakian VA *et al* (2009). The yeast Pif1 helicase prevents genomic instability caused by G-quadruplex-forming CEB1 sequences in vivo. *PLoS genetics* **5:** e1000475.

Riley T, Sontag E, Chen P, Levine A (2008). Transcriptional control of human p53-regulated genes. *Nature reviews Molecular cell biology* **9:** 402-412.

Riou JF, Guittat L, Mailliet P, Laoui A, Renou E, Petitgenet O *et al* (2002). Cell senescence and telomere shortening induced by a new series of specific G-quadruplex DNA ligands. *Proceedings of the National Academy of Sciences of the United States of America* **99:** 2672-2677.

Rodriguez MS, Desterro JM, Lain S, Midgley CA, Lane DP, Hay RT (1999). SUMO-1 modification activates the transcriptional response of p53. *The EMBO journal* **18:** 6455-6461.

Roth J, Dobbelstein M, Freedman DA, Shenk T, Levine AJ (1998). Nucleo-cytoplasmic shuttling of the hdm2 oncoprotein regulates the levels of the p53 protein via a pathway used by the human immunodeficiency virus rev protein. *The EMBO journal* **17:** 554-564.

Roy S, Packman K, Jeffrey R, Tenniswood M (2005). Histone deacetylase inhibitors differentially stabilize acetylated p53 and induce cell cycle arrest or apoptosis in prostate cancer cells. *Cell death and differentiation* **12:** 482-491.

Roy S, Tenniswood M (2007). Site-specific acetylation of p53 directs selective transcription complex assembly. *The Journal of biological chemistry* **282:** 4765-4771.

Ruijs MW, Schmidt MK, Nevanlinna H, Tommiska J, Aittomaki K, Pruntel R *et al* (2007). The single-nucleotide polymorphism 309 in the MDM2 gene contributes to the Li-Fraumeni syndrome and related phenotypes. *European journal of human genetics : EJHG* **15:** 110-114.

Ruijs MW, Verhoef S, Rookus MA, Pruntel R, van der Hout AH, Hogervorst FB *et al* (2010). TP53 germline mutation testing in 180 families suspected of Li-Fraumeni syndrome: mutation detection rate and relative frequency of cancers in different familial phenotypes. *Journal of medical genetics* **47:** 421-428.

Ruiz-Lozano P, Hixon ML, Wagner MW, Flores AI, Ikawa S, Baldwin AS, Jr. *et al* (1999). p53 is a transcriptional activator of the muscle-specific phosphoglycerate mutase gene and contributes in vivo to the control of its cardiac expression. *Cell growth & differentiation : the molecular biology journal of the American Association for Cancer Research* **10:** 295-306.

Sablina AA, Chumakov PM, Levine AJ, Kopnin BP (2001). p53 activation in response to microtubule disruption is mediated by integrin-Erk signaling. *Oncogene* **20:** 899-909.

Sachidanandam R, Weissman D, Schmidt SC, Kakol JM, Stein LD, Marth G *et al* (2001). A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. *Nature* **409:** 928-933.

Sahin E, Colla S, Liesa M, Moslehi J, Muller FL, Guo M *et al* (2011). Telomere dysfunction induces metabolic and mitochondrial compromise. *Nature* **470:** 359-365.

Sahin E, DePinho RA (2012). Axis of ageing: telomeres, p53 and mitochondria. *Nature reviews Molecular cell biology* **13:** 397-404.

Sakaguchi K, Sakamoto H, Lewis MS, Anderson CW, Erickson JW, Appella E *et al* (1997). Phosphorylation of serine 392 stabilizes the tetramer formation of tumor suppressor protein p53. *Biochemistry* **36:** 10117-10124.

Sakaguchi K, Saito S, Higashimoto Y, Roy S, Anderson CW, Appella E (2000). Damage-mediated phosphorylation of human p53 threonine 18 through a cascade mediated by a casein 1-like kinase. Effect on Mdm2 binding. *The Journal of biological chemistry* **275:** 9278-9283.

Sakamuro D, Sabbatini P, White E, Prendergast GC (1997). The polyproline region of p53 is required to activate apoptosis but not growth arrest. *Oncogene* **15:** 887-898.

Salmon A, Amikam D, Sodha N, Davidson S, Basel-Vanagaite L, Eeles RA *et al* (2007). Rapid development of post-radiotherapy sarcoma and breast cancer in a patient with a novel germline 'de-novo' TP53 mutation. *Clin Oncol (R Coll Radiol)* **19:** 490-493.

Salvioli S, Bonafe M, Barbi C, Storci G, Trapassi C, Tocco F *et al* (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.

Sandrini F, Villani DP, Tucci S, Moreira AC, de Castro M, Elias LL (2005). Inheritance of R337H p53 gene mutation in children with sporadic adrenocortical tumor. *Hormone and metabolic research = Hormon- und Stoffwechselforschung = Hormones et metabolisme* **37:** 231-235.

Santiago A, Li D, Zhao LY, Godsey A, Liao D (2013). p53 SUMOylation promotes its nuclear export by facilitating its release from the nuclear export receptor CRM1. *Molecular biology of the cell* **24:** 2739-2752.

Sarkies P, Reams C, Simpson LJ, Sale JE (2010). Epigenetic instability due to defective replication of structured DNA. *Molecular cell* **40:** 703-713.

Sarkies P, Murat P, Phillips LG, Patel KJ, Balasubramanian S, Sale JE (2012). FANCJ coordinates two pathways that maintain epigenetic stability at G-quadruplex DNA. *Nucleic acids research* **40:** 1485-1498.

Savage SA, Burdett L, Troisi R, Douglass C, Hoover RN, Chanock SJ (2007). Germ-line genetic variation of TP53 in osteosarcoma. *Pediatric blood & cancer* **49:** 28-33.

Sax JK, Fei P, Murphy ME, Bernhard E, Korsmeyer SJ, El-Deiry WS (2002). BID regulation by p53 contributes to chemosensitivity. *Nature cell biology* **4:** 842-849.

Schneider E, Montenarh M, Wagner P (1998). Regulation of CAK kinase activity by p53. *Oncogene* **17:** 2733-2741.

Schon O, Friedler A, Bycroft M, Freund SM, Fersht AR (2002). Molecular mechanism of the interaction between MDM2 and p53. *Journal of molecular biology* **323:** 491-501.

Schroeder M, Mass MJ (1997). CpG methylation inactivates the transcriptional activity of the promoter of the human p53 tumor suppressor gene. *Biochemical and biophysical research communications* **235:** 403-406.

Schumacher B, Mondry J, Thiel P, Weyand M, Ottmann C (2010). Structure of the p53 C-terminus bound to 14-3-3: implications for stabilization of the p53 tetramer. *FEBS letters* **584:** 1443-1448.

Scotto C, Delphin C, Deloulme JC, Baudier J (1999). Concerted regulation of wild-type p53 nuclear accumulation and activation by S100B and calcium-dependent protein kinase C. *Molecular and cellular biology* **19:** 7168-7180.

Seidinger AL, Mastellaro MJ, Paschoal Fortes F, Godoy Assumpcao J, Aparecida Cardinalli I, Aparecida Ganazza M *et al* (2011). Association of the highly prevalent TP53 R337H mutation with pediatric choroid plexus carcinoma and osteosarcoma in southeast Brazil. *Cancer* **117:** 2228-2235.

Serrano M, Lin AW, McCurrach ME, Beach D, Lowe SW (1997). Oncogenic ras provokes premature cell senescence associated with accumulation of p53 and p16INK4a. *Cell* **88:** 593-602.

Serrano MA, Li Z, Dangeti M, Musich PR, Patrick S, Roginskaya M *et al* (2013). DNA-PK, ATM and ATR collaboratively regulate p53-RPA interaction to facilitate homologous recombination DNA repair. *Oncogene* **32:** 2452-2462.

Shammas MA, Shmookler Reis RJ, Akiyama M, Koley H, Chauhan D, Hideshima T *et al* (2003). Telomerase inhibition and cell growth arrest by G-quadruplex interactive agent in multiple myeloma. *Molecular cancer therapeutics* **2:** 825-833.

Shaulian E, Zauberman A, Ginsberg D, Oren M (1992). Identification of a minimal transforming domain of p53: negative dominance through abrogation of sequence-specific DNA binding. *Molecular and cellular biology* **12:** 5581-5592.

Shaulsky G, Ben-Ze'ev A, Rotter V (1990a). Subcellular distribution of the p53 protein during the cell cycle of Balb/c 3T3 cells. *Oncogene* **5:** 1707-1711.

Shaulsky G, Goldfinger N, Ben-Ze'ev A, Rotter V (1990b). Nuclear accumulation of p53 protein is mediated by several nuclear localization signals and plays a role in tumorigenesis. *Molecular and cellular biology* **10:** 6565-6577.

Shaulsky G, Goldfinger N, Peled A, Rotter V (1991a). Involvement of wild-type p53 protein in the cell cycle requires nuclear localization. *Cell growth & differentiation : the molecular biology journal of the American Association for Cancer Research* **2:** 661-667.

Shaulsky G, Goldfinger N, Tosky MS, Levine AJ, Rotter V (1991b). Nuclear localization is essential for the activity of p53 protein. *Oncogene* **6:** 2055-2065.

Shaw P, Freeman J, Bovey R, Iggo R (1996). Regulation of specific DNA binding by p53: evidence for a role for O-glycosylation and charged residues at the carboxy-terminus. *Oncogene* **12:** 921-930.

Sheikh MS, Burns TF, Huang Y, Wu GS, Amundson S, Brooks KS *et al* (1998). p53-dependent and -independent regulation of the death receptor KILLER/DR5 gene expression in response to genotoxic stress and tumor necrosis factor alpha. *Cancer research* **58:** 1593-1598.

Sheng Y, Laister RC, Lemak A, Wu B, Tai E, Duan S *et al* (2008). Molecular basis of Pirh2-mediated p53 ubiquitylation. *Nature structural & molecular biology* **15:** 1334-1342.

Sherif ZA, Nakai S, Pirollo KF, Rait A, Chang EH (2001). Downmodulation of bFGF-binding protein expression following restoration of p53 function. *Cancer gene therapy* **8:** 771-782.

Sherr CJ (1994). G1 phase progression: cycling on cue. *Cell* **79:** 551-555.

Sherr CJ (2001). The INK4a/ARF network in tumour suppression. *Nature reviews Molecular cell biology* **2:** 731-737.

Sherr CJ (2006). Divorcing ARF and p53: an unsettled case. *Nature reviews Cancer* **6:** 663-673.

Shi D, Pop MS, Kulikov R, Love IM, Kung AL, Grossman SR (2009a). CBP and p300 are cytoplasmic E4 polyubiquitin ligases for p53. *Proceedings of the National Academy of Sciences of the United States of America* **106:** 16275-16280.

Shi DF, Wheelhouse RT, Sun D, Hurley LH (2001). Quadruplex-interactive agents as telomerase inhibitors: synthesis of porphyrins and structure-activity relationship for the inhibition of telomerase. *Journal of medicinal chemistry* **44:** 4509-4523.

Shi H, Tan SJ, Zhong H, Hu W, Levine A, Xiao CJ *et al* (2009b). Winter temperature and UV are tightly linked to genetic changes in the p53 tumor suppressor pathway in Eastern Asia. *American journal of human genetics* **84:** 534-541.

Shi X, Kachirskaia I, Yamaguchi H, West LE, Wen H, Wang EW *et al* (2007). Modulation of p53 function by SET8-mediated methylation at lysine 382. *Molecular cell* **27:** 636-646.

Shieh SY, Ahn J, Tamai K, Taya Y, Prives C (2000). The human homologs of checkpoint kinases Chk1 and Cds1 (Chk2) phosphorylate p53 at multiple DNA damage-inducible sites. *Genes & development* **14:** 289-300.

Shvarts A, Steegenga WT, Riteco N, van Laar T, Dekker P, Bazuine M *et al* (1996). MDMX: a novel p53-binding protein with some functional properties of MDM2. *The EMBO journal* **15:** 5349-5357.

Siddique MM, Balram C, Fiszer-Maliszewska L, Aggarwal A, Tan A, Tan P *et al* (2005). Evidence for selective expression of the p53 codon 72 polymorphs: implications in cancer development. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology* **14:** 2245-2252.

Siddiqui-Jain A, Grand CL, Bearss DJ, Hurley LH (2002). Direct evidence for a G-quadruplex in a promoter region and its targeting with a small molecule to repress c-MYC transcription. *Proceedings of the National Academy of Sciences of the United States of America* **99:** 11593-11598.

Siliciano JD, Canman CE, Taya Y, Sakaguchi K, Appella E, Kastan MB (1997). DNA damage induces phosphorylation of the amino terminus of p53. *Genes & development* **11:** 3471-3481.

Silva AG, Achatz IM, Krepischi AC, Pearson PL, Rosenberg C (2012). Number of rare germline CNVs and TP53 mutation types. *Orphanet journal of rare diseases* **7:** 101.

Simonsson T, Pecinka P, Kubista M (1998). DNA tetraplex formation in the control region of c-myc. *Nucleic acids research* **26:** 1167-1172.

Simonsson T (2001). G-quadruplex DNA structures--variations on a theme. *Biological chemistry* **382:** 621-628.

Sjalander A, Birgander R, Saha N, Beckman L, Beckman G (1996). p53 polymorphisms and haplotypes show distinct differences between major ethnic groups. *Human heredity* **46:** 41-48.

Smith ML, Ford JM, Hollander MC, Bortnick RA, Amundson SA, Seo YR *et al* (2000). p53-mediated DNA repair responses to UV radiation: studies of mouse cells lacking p53, p21, and/or gadd45 genes. *Molecular and cellular biology* **20:** 3705-3714.

Song H, Hollstein M, Xu Y (2007). p53 gain-of-function cancer mutants induce genetic instability by inactivating ATM. *Nature cell biology* **9:** 573-580.

Soussi T, Caron de Fromentel C, Mechali M, May P, Kress M (1987). Cloning and characterization of a cDNA from Xenopus laevis coding for a protein homologous to human and murine p53. *Oncogene* **1:** 71-78.

Soussi T, Caron de Fromentel C, May P (1990). Structural aspects of the p53 protein in relation to gene evolution. *Oncogene* **5:** 945-952.

Srivastava S, Zou ZQ, Pirollo K, Blattner W, Chang EH (1990). Germ-line transmission of a mutated p53 gene in a cancer-prone family with Li-Fraumeni syndrome. *Nature* **348:** 747-749.

Stacey SN, Sulem P, Jonasdottir A, Masson G, Gudmundsson J, Gudbjartsson DF *et al* (2011). A germline variant in the TP53 polyadenylation signal confers cancer susceptibility. *Nature genetics* **43:** 1098-1103.

Stambolsky P, Weisz L, Shats I, Klein Y, Goldfinger N, Oren M *et al* (2006). Regulation of AIF expression by p53. *Cell death and differentiation* **13:** 2140-2149.

Stegh AH (2012). Targeting the p53 signaling pathway in cancer therapy - the promises, challenges and perils. *Expert opinion on therapeutic targets* **16:** 67-83.

Stehmeier P, Muller S (2009). Regulation of p53 family members by the ubiquitin-like SUMO system. *DNA repair* **8:** 491-498.

Stewart N, Hicks GG, Paraskevas F, Mowat M (1995). Evidence for a second cell cycle block at G2/M by p53. *Oncogene* **10:** 109-115.

Stewart ZA, Pietenpol JA (2001). p53 Signaling and cell cycle checkpoints. *Chemical research in toxicology* **14:** 243-263.

Stewart ZA, Tang LJ, Pietenpol JA (2001). Increased p53 phosphorylation after microtubule disruption is mediated in a microtubule inhibitor- and cell-specific manner. *Oncogene* **20:** 113-124.

Stommel JM, Marchenko ND, Jimenez GS, Moll UM, Hope TJ, Wahl GM (1999). A leucine-rich nuclear export signal in the p53 tetramerization domain: regulation of subcellular localization and p53 activity by NES masking. *The EMBO journal* **18:** 1660-1672.

Stommel JM, Wahl GM (2004). Accelerated MDM2 auto-degradation induced by DNA-damage kinases is required for p53 activation. *The EMBO journal* **23:** 1547-1556.

Stommel JM, Wahl GM (2005). A new twist in the feedback loop: stress-activated MDM2 destabilization is required for p53 activation. *Cell Cycle* **4:** 411-417.

Strano S, Munarriz E, Rossi M, Cristofanelli B, Shaul Y, Castagnoli L *et al* (2000). Physical and functional interaction between p53 mutants and different isoforms of p73. *The Journal of biological chemistry* **275:** 29503-29512.

Strong LC, Williams WR, Ferrell RE, Tainsky MA (1989). Genetic analysis of childhood sarcoma. *Princess Takamatsu symposia* **20:** 151-157.

Strong LC, Williams WR, Tainsky MA (1992). The Li-Fraumeni syndrome: from clinical epidemiology to molecular genetics. *American journal of epidemiology* **135:** 190-199.

Subbaramaiah K, Altorki N, Chung WJ, Mestre JR, Sampat A, Dannenberg AJ (1999). Inhibition of cyclooxygenase-2 gene expression by p53. *The Journal of biological chemistry* **274:** 10911-10915.

Sugars KL, Budhram-Mahadeo V, Packham G, Latchman DS (2001). A minimal Bcl-x promoter is activated by Brn-3a and repressed by p53. *Nucleic acids research* **29:** 4530-4540.

Sullivan A, Syed N, Gasco M, Bergamaschi D, Trigiante G, Attard M *et al* (2004). Polymorphism in wild-type p53 modulates response to chemotherapy in vitro and in vivo. *Oncogene* **23:** 3328-3337.

Sun D, Hurley LH (2009). The importance of negative superhelicity in inducing the formation of G-quadruplex and i-motif structures in the c-Myc promoter: implications for drug targeting and control of gene expression. *Journal of medicinal chemistry* **52:** 2863-2874.

Sundquist WI, Klug A (1989). Telomeric DNA dimerizes by formation of guanine tetrads between hairpin loops. *Nature* **342:** 825-829.

Suspitsin EN, Buslov KG, Grigoriev MY, Ishutkina JG, Ulibina JM, Gorodinskaya VM *et al* (2003). Evidence against involvement of p53 polymorphism in breast cancer predisposition. *International journal of cancer Journal international du cancer* **103:** 431-433.

Suzuki HI, Yamagata K, Sugimoto K, Iwamoto T, Kato S, Miyazono K (2009). Modulation of microRNA processing by p53. *Nature* **460:** 529-533.

Sykes SM, Mellert HS, Holbert MA, Li K, Marmorstein R, Lane WS *et al* (2006). Acetylation of the p53 DNA-binding domain regulates apoptosis induction. *Molecular cell* **24:** 841-851.

Tabori U, Nanda S, Druker H, Lees J, Malkin D (2007). Younger age of cancer initiation is associated with shorter telomere length in Li-Fraumeni syndrome. *Cancer research* **67:** 1415-1418.

Tabori U, Malkin D (2008). Risk stratification in cancer predisposition syndromes: lessons learned from novel molecular developments in Li-Fraumeni syndrome. *Cancer research* **68:** 2053-2057.

Takahashi R, Markovic S, Scrable H (2013). Dominant Effects of Delta40p53 on p53 Function and Melanoma Cell Fate. *The Journal of investigative dermatology*.

Takimoto R, El-Deiry WS (2000). Wild-type p53 transactivates the KILLER/DR5 gene through an intronic sequence-specific DNA-binding site. *Oncogene* **19:** 1735-1743.

Tanaka H, Arakawa H, Yamaguchi T, Shiraishi K, Fukuda S, Matsui K *et al* (2000). A ribonucleotide reductase gene involved in a p53-dependent cell-cycle checkpoint for DNA damage. *Nature* **404:** 42-49.

Tang Y, Luo J, Zhang W, Gu W (2006). Tip60-dependent acetylation of p53 modulates the decision between cell-cycle arrest and apoptosis. *Molecular cell* **24:** 827-839.

Tang Y, Zhao W, Chen Y, Zhao Y, Gu W (2008). Acetylation is indispensable for p53 activation. *Cell* **133:** 612-626.

Tao W, Levine AJ (1999). Nucleocytoplasmic shuttling of oncoprotein Hdm2 is required for Hdm2-mediated degradation of p53. *Proceedings of the National Academy of Sciences of the United States of America* **96:** 3077-3080.

Tasdemir E, Maiuri MC, Galluzzi L, Vitale I, Djavaheri-Mergny M, D'Amelio M *et al* (2008a). Regulation of autophagy by cytoplasmic p53. *Nature cell biology* **10:** 676-687.

Tasdemir E, Maiuri MC, Orhon I, Kepp O, Morselli E, Criollo A *et al* (2008b). p53 represses autophagy in a cell cycle-dependent fashion. *Cell Cycle* **7:** 3006-3011.

Tauchi T, Shin-Ya K, Sashida G, Sumi M, Nakajima A, Shimamoto T *et al* (2003). Activity of a novel G-quadruplex-interactive telomerase inhibitor, telomestatin (SOT-095), against human leukemia cells: involvement of ATM-dependent DNA damage response pathways. *Oncogene* **22:** 5338-5347.

Teufel DP, Freund SM, Bycroft M, Fersht AR (2007). Four domains of p300 each bind tightly to a sequence spanning both transactivation subdomains of p53. *Proceedings of the National Academy of Sciences of the United States of America* **104:** 7009-7014.

Teufel DP, Bycroft M, Fersht AR (2009). Regulation by phosphorylation of the relative affinities of the N-terminal transactivation domains of p53 for p300 domains and Mdm2. *Oncogene* **28:** 2112-2118.

Thiery J, Dorothee G, Haddada H, Echchakir H, Richon C, Stancou R *et al* (2003). Potentiation of a tumor cell susceptibility to autologous CTL killing by restoration of wild-type p53 function. *J Immunol* **170:** 5919-5926.

Thornborrow EC, Manfredi JJ (1999). One mechanism for cell type-specific regulation of the bax promoter by the tumor suppressor p53 is dictated by the p53 response element. *The Journal of biological chemistry* **274:** 33747-33756.

Thut CJ, Chen JL, Klemm R, Tjian R (1995). p53 transcriptional activation mediated by coactivators TAFII40 and TAFII60. *Science* **267:** 100-104.

Tian S, Huang S, Wu S, Guo W, Li J, He X (2010). MicroRNA-1285 inhibits the expression of p53 by directly targeting its 3' untranslated region. *Biochemical and biophysical research communications* **396:** 435-439.

Tian X, Tian Y, Ma P, Sui CG, Meng FD, Li Y *et al* (2013). Association between MDM2 SNP309 T>G and risk of gastric cancer: a meta-analysis. *Asian Pacific journal of cancer prevention : APJCP* **14:** 1925-1929.

Tinat J, Bougeard G, Baert-Desurmont S, Vasseur S, Martin C, Bouvignies E *et al* (2009). 2009 version of the Chompret criteria for Li Fraumeni syndrome. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* **27:** e108-109; author reply e110.

Tinel A, Tschopp J (2004). The PIDDosome, a protein complex implicated in activation of caspase-2 in response to genotoxic stress. *Science* **304:** 843-846.

Toledo F, Wahl GM (2006). Regulating the p53 pathway: in vitro hypotheses, in vivo veritas. *Nature reviews Cancer* **6:** 909-923.

Tornaletti S, Rozek D, Pfeifer GP (1993). The distribution of UV photoproducts along the human p53 gene and its relation to mutations in skin cancer. *Oncogene* **8:** 2051-2057.

Tornaletti S, Rozek D, Pfeifer GP (1994). Mapping of UV photoproducts along the human P53 gene. *Annals of the New York Academy of Sciences* **726:** 324-326.

Trkova M, Hladikova M, Kasal P, Goetz P, Sedlacek Z (2002). Is there anticipation in the age at onset of cancer in families with Li-Fraumeni syndrome? *Journal of human genetics* **47:** 381-386.

Trkova M, Prochazkova K, Krutilkova V, Sumerauer D, Sedlacek Z (2007). Telomere length in peripheral blood cells of germline TP53 mutation carriers is shorter than that of normal individuals of corresponding age. *Cancer* **110:** 694-702.

Ueba T, Nosaka T, Takahashi JA, Shibata F, Florkiewicz RZ, Vogelstein B *et al* (1994). Transcriptional regulation of basic fibroblast growth factor gene by p53 in human glioblastoma and hepatocellular carcinoma cells. *Proceedings of the National Academy of Sciences of the United States of America* **91:** 9009-9013.

Unger T, Mietz JA, Scheffner M, Yee CL, Howley PM (1993). Functional domains of wild-type and mutant p53 proteins involved in transcriptional regulation, transdominant inhibition, and transformation suppression. *Molecular and cellular biology* **13:** 5186-5194.

Varley JM, Chapman P, McGown G, Thorncroft M, White GR, Greaves MJ *et al* (1998). Genetic and functional studies of a germline TP53 splicing mutation in a Li-Fraumeni-like family. *Oncogene* **16:** 3291-3298.

Vassilev LT, Vu BT, Graves B, Carvajal D, Podlaski F, Filipovic Z *et al* (2004). In vivo activation of the p53 pathway by small-molecule antagonists of MDM2. *Science* **303:** 844-848.

Venkatachalam S, Tyner SD, Pickering CR, Boley S, Recio L, French JE *et al* (2001). Is p53 haploinsufficient for tumor suppression? Implications for the p53+/- mouse model in carcinogenicity testing. *Toxicologic pathology* **29 Suppl:** 147-154.

Venteicher AS, Artandi SE (2009). TCAB1: driving telomerase to Cajal bodies. *Cell Cycle* **8:** 1329-1331.

Villani A, Tabori U, Schiffman J, Shlien A, Beyene J, Druker H *et al* (2011). Biochemical and imaging surveillance in germline TP53 mutation carriers with Li-Fraumeni syndrome: a prospective observational study. *The lancet oncology* **12:** 559-567.

Vousden KH, Prives C (2009). Blinded by the Light: The Growing Complexity of p53. *Cell* **137:** 413-431.

Wadhwa R, Takano S, Robert M, Yoshida A, Nomura H, Reddel RR *et al* (1998). Inactivation of tumor suppressor p53 by mot-2, a hsp70 family member. *The Journal of biological chemistry* **273:** 29586-29591.

Wagner J, Portwine C, Rabin K, Leclerc JM, Narod SA, Malkin D (1994). High frequency of germline p53 mutations in childhood adrenocortical cancer. *Journal of the National Cancer Institute* **86:** 1707-1710.

Walker KK, Levine AJ (1996). Identification of a novel p53 functional domain that is necessary for efficient growth suppression. *Proceedings of the National Academy of Sciences of the United States of America* **93:** 15335-15340.

Wang C, Ivanov A, Chen L, Fredericks WJ, Seto E, Rauscher FJ, 3rd *et al* (2005). MDM2 interaction with nuclear corepressor KAP1 contributes to p53 inactivation. *The EMBO journal* **24:** 3279-3290.

Wang Q, Zambetti GP, Suttle DP (1997). Inhibition of DNA topoisomerase II alpha gene expression by the p53 tumor suppressor. *Molecular and cellular biology* **17:** 389-397.

Wang Q, Liu D, Bai Y (2010). T-cryptantigen (TCA) activation in sever pneumonia complicated with multiple organ failure. *Transfusion and apheresis science : official journal of the World Apheresis Association : official journal of the European Society for Haemapheresis* **43:** 361-364.

Wang S, Chen L, Zhao Q, Rong H, Wang M, Gong W *et al* (2013a). Effect of TP53 codon 72 and MDM2 SNP309 polymorphisms on survival of gastric cancer among patients who receiving 5-fluorouracil-based postoperative adjuvant chemotherapy. *Cancer chemotherapy and pharmacology* **71:** 1073-1082.

Wang S, Lan X, Tan S, Wang S, Li Y (2013b). P53 codon 72 Arg/Pro polymorphism and lung cancer risk in Asians: an updated meta-analysis. *Tumour biology : the journal of the International Society for Oncodevelopmental Biology and Medicine*.

Wang SP, Wang WL, Chang YL, Wu CT, Chao YC, Kao SH *et al* (2009). p53 controls cancer cell invasion by inducing the MDM2-mediated degradation of Slug. *Nature cell biology* **11:** 694-704.

Wang X, Ohnishi K, Takahashi A, Ohnishi T (1998). Poly(ADP-ribosyl)ation is required for p53-dependent signal transduction induced by radiation. *Oncogene* **17:** 2819-2825.

Wang X, Wang J, Jiang X (2011). MdmX protein is essential for Mdm2 protein-mediated p53 polyubiquitination. *The Journal of biological chemistry* **286:** 23725-23734.

Wang X, Jiang X (2012). Mdm2 and MdmX partner to regulate p53. *FEBS letters* **586:** 1390-1396.

Wang XW, Vermeulen W, Coursen JD, Gibson M, Lupold SE, Forrester K *et al* (1996). The XPB and XPD DNA helicases are components of the p53-mediated apoptosis pathway. *Genes & development* **10:** 1219-1232.

Wang Y, Reed M, Wang P, Stenger JE, Mayr G, Anderson ME *et al* (1993). p53 domains: identification and characterization of two autonomous DNA-binding regions. *Genes & development* **7:** 2575-2586.

Wang YV, Wade M, Wong E, Li YC, Rodewald LW, Wahl GM (2007). Quantitative analyses reveal the importance of regulated Hdmx degradation for p53 activation. *Proceedings of the National Academy of Sciences of the United States of America* **104:** 12365-12370.

Wang-Gohrke S, Becher H, Kreienberg R, Runnebaum IB, Chang-Claude J (2002). Intron 3 16 bp duplication polymorphism of p53 is associated with an increased risk for breast cancer by the age of 50 years. *Pharmacogenetics* **12:** 269-272.

Waning DL, Lehman JA, Batuello CN, Mayo LD (2011). c-Abl phosphorylation of Mdm2 facilitates Mdm2-Mdmx complex formation. *The Journal of biological chemistry* **286:** 216-222.

Warburg O (1956). On respiratory impairment in cancer cells. *Science* **124:** 269-270.

Waterman JL, Shenk JL, Halazonetis TD (1995). The dihedral symmetry of the p53 tetramerization domain mandates a conformational switch upon DNA binding. *The EMBO journal* **14:** 512-519.

Waterman MJ, Waterman JL, Halazonetis TD (1996). An engineered four-stranded coiled coil substitutes for the tetramerization domain of wild-type p53 and alleviates transdominant inhibition by tumor-derived p53 mutants. *Cancer research* **56:** 158-163.

Waterman MJ, Stavridi ES, Waterman JL, Halazonetis TD (1998). ATM-dependent activation of p53 involves dephosphorylation and association with 14-3-3 proteins. *Nature genetics* **19:** 175-178.

Weber JD, Taylor LJ, Roussel MF, Sherr CJ, Bar-Sagi D (1999). Nucleolar Arf sequesters Mdm2 and activates p53. *Nature cell biology* **1:** 20-26.

Wei CL, Wu Q, Vega VB, Chiu KP, Ng P, Zhang T *et al* (2006). A global map of p53 transcription-factor binding sites in the human genome. *Cell* **124:** 207-219.

Weisz L, Oren M, Rotter V (2007). Transcription regulation by mutant p53. *Oncogene* **26:** 2202-2211.

Wesierska-Gadek J, Schmid G, Cerni C (1996). ADP-ribosylation of wild-type p53 in vitro: binding of p53 protein to specific p53 consensus sequence prevents its modification. *Biochemical and biophysical research communications* **224:** 96-102.

Whibley C, Pharoah PD, Hollstein M (2009). p53 polymorphisms: cancer implications. *Nature reviews Cancer* **9:** 95-107.

Wilkening S, Bermejo JL, Hemminki K (2007). MDM2 SNP309 and cancer risk: a combined analysis. *Carcinogenesis* **28:** 2262-2267.

Williamson JR, Raghuraman MK, Cech TR (1989). Monovalent cation-induced structure of telomeric DNA: the G-quartet model. *Cell* **59:** 871-880.

Wilson JW, Pritchard DM, Hickman JA, Potten CS (1998). Radiation-induced p53 and p21WAF-1/CIP1 expression in the murine intestinal epithelium: apoptosis and cell cycle arrest. *The American journal of pathology* **153:** 899-909.

Wood SA (2002). Dubble or nothing? Is HAUSP deubiquitylating enzyme the final arbiter of p53 levels? *Science's STKE : signal transduction knowledge environment* **2002:** pe34.

Wu D, Zhang Z, Chu H, Xu M, Xue Y, Zhu H *et al* (2013). Intron 3 sixteen base pairs duplication polymorphism of p53 contributes to breast cancer susceptibility: evidence from meta-analysis. *PloS one* **8:** e61662.

Wu GS, Burns TF, McDonald ER, 3rd, Meng RD, Kao G, Muschel R *et al* (1999). Induction of the TRAIL receptor KILLER/DR5 in p53-dependent apoptosis but not growth arrest. *Oncogene* **18:** 6411-6418.

Wu X, Bayle JH, Olson D, Levine AJ (1993). The p53-mdm-2 autoregulatory feedback loop. *Genes & development* **7:** 1126-1132.

Xia M, Land H (2007). Tumor suppressor p53 restricts Ras stimulation of RhoA and cancer cell motility. *Nature structural & molecular biology* **14:** 215-223.

Xiao J, Lin H, Luo X, Luo X, Wang Z (2011). miR-605 joins p53 network to form a p53:miR-605:Mdm2 positive feedback loop in response to stress. *The EMBO journal* **30:** 524-532.

Xiao X, Lee JH (2010). Systems analysis of alternative splicing and its regulation. *Wiley interdisciplinary reviews Systems biology and medicine* **2:** 550-565.

Xirodimas DP, Saville MK, Bourdon JC, Hay RT, Lane DP (2004). Mdm2-mediated NEDD8 conjugation of p53 inhibits its transcriptional activity. *Cell* **118:** 83-97.

Yamakuchi M, Ferlito M, Lowenstein CJ (2008). miR-34a repression of SIRT1 regulates apoptosis. *Proceedings of the National Academy of Sciences of the United States of America* **105:** 13421-13426.

Yamakuchi M, Lotterman CD, Bao C, Hruban RH, Karim B, Mendell JT *et al* (2010). P53-induced microRNA-107 inhibits HIF-1 and tumor angiogenesis. *Proceedings of the National Academy of Sciences of the United States of America* **107:** 6334-6339.

Yew PR, Berk AJ (1992). Inhibition of p53 transactivation required for transformation by adenovirus early 1B protein. *Nature* **357:** 82-85.

Ying L, Green JJ, Li H, Klenerman D, Balasubramanian S (2003). Studies on the structure and dynamics of the human telomeric G quadruplex by single-molecule fluorescence resonance energy transfer. *Proceedings of the National Academy of Sciences of the United States of America* **100:** 14629-14634.

Yonish-Rouach E, Resnitzky D, Lotem J, Sachs L, Kimchi A, Oren M (1991). Wild-type p53 induces apoptosis of myeloid leukaemic cells that is inhibited by interleukin-6. *Nature* **352:** 345-347.

Yu X, Kubota H, Wang R, Saegusa J, Ogawa Y, Ichihara G *et al* (2001). Involvement of Bcl-2 family genes and Fas signaling system in primary and secondary male germ cell apoptosis induced by 2-bromopropane in rat. *Toxicology and applied pharmacology* **174:** 35-48.

Yuan JM, Li XD, Liu ZY, Hou GQ, Kang JH, Huang DY *et al* (2011). Cisplatin induces apoptosis via upregulating Wrap53 in U-2OS osteosarcoma cells. *Asian Pacific journal of cancer prevention : APJCP* **12:** 3465-3469.

Zahler AM, Williamson JR, Cech TR, Prescott DM (1991). Inhibition of telomerase by G-quartet DNA structures. *Nature* **350:** 718-720.

Zalcenstein A, Stambolsky P, Weisz L, Muller M, Wallach D, Goncharov TM *et al* (2003). Mutant p53 gain of function: repression of CD95(Fas/APO-1) gene expression by tumor-associated p53 mutants. *Oncogene* **22:** 5667-5676.

Zarudnaya MI, Kolomiets IM, Potyahaylo AL, Hovorun DM (2003). Downstream elements of mammalian pre-mRNA polyadenylation signals: primary, secondary and higher-order structures. *Nucleic acids research* **31:** 1375-1386.

Zerdoumi Y, Aury-Landas J, Bonaiti-Pellie C, Derambure C, Sesboue R, Renaux-Petel M *et al* (2013). Drastic effect of germline TP53 missense mutations in Li-Fraumeni patients. *Human mutation* **34:** 453-461.

Zhan Q, Antinore MJ, Wang XW, Carrier F, Smith ML, Harris CC *et al* (1999). Association with Cdc2 and inhibition of Cdc2/Cyclin B1 kinase activity by the p53-regulated protein Gadd45. *Oncogene* **18:** 2892-2900.

Zhang Y, Xiong Y (2001a). Control of p53 ubiquitination and nuclear export by MDM2 and ARF. *Cell growth & differentiation : the molecular biology journal of the American Association for Cancer Research* **12:** 175-186.

Zhang Y, Xiong Y (2001b). A p53 amino-terminal nuclear export signal inhibited by DNA damage-induced phosphorylation. *Science* **292:** 1910-1915.

Zhao W, He C, Rotter V, Merrick BA, Selkirk JK (1999). An intragenic deletion of nuclear localization signal-1 of p53 tumor suppressor gene results in loss of apoptosis in murine fibroblasts. *Cancer letters* **147:** 101-108.

Zhou C, Chen H, Wang A (2013). P53 codon 72 polymorphism and lung cancer risk: evidence from 27,958 subjects. *Tumour biology : the journal of the International Society for Oncodevelopmental Biology and Medicine*.

Zhou L, Yuan Q, Yang M (2012). A functional germline variant in the P53 polyadenylation signal and risk of esophageal squamous cell carcinoma. *Gene* **506:** 295-297.

Zhou Y, Li N, Zhuang W, Liu GJ, Wu TX, Yao X *et al* (2007). P53 codon 72 polymorphism and gastric cancer: a meta-analysis of the literature. *International journal of cancer Journal international du cancer* **121:** 1481-1486.

Zhu F, Dolle ME, Berton TR, Kuiper RV, Capps C, Espejo A *et al* (2010). Mouse models for the p53 R72P polymorphism mimic human phenotypes. *Cancer research* **70:** 5851-5859.

Ziegler A, Jonason AS, Leffell DJ, Simon JA, Sharma HW, Kimmelman J *et al* (1994). Sunburn and p53 in the onset of skin cancer. *Nature* **372:** 773-776.

Zurer I, Hofseth LJ, Cohen Y, Xu-Welliver M, Hussain SP, Harris CC *et al* (2004). The role of p53 in base excision repair following genotoxic stress. *Carcinogenesis* **25:** 11-19.

ANNEXES

Oncogene (2010) 29, 2691–2700
© 2010 Macmillan Publishers Limited All rights reserved 0950-9232/10 $32.00
www.nature.com/onc

ORIGINAL ARTICLE

p53 regulates the transcription of its Δ133p53 isoform through specific response elements contained within the *TP53* P2 internal promoter

V Marcel, V Vijayakumar, L Fernández-Cuesta, H Hafsi, C Sagne, A Hautefeuille, M Olivier and P Hainaut

Molecular Carcinogenesis Group, International Agency for Research on Cancer, Lyon, Cedex, France

The tumor suppressor p53 protein is activated by genotoxic stress and regulates genes involved in senescence, apoptosis and cell-cycle arrest. Nine p53 isoforms have been described that may modulate suppressive functions of the canonical p53 protein. Among them, Δ133p53 lacks the 132 proximal residues and has been shown to modulate p53-induced apoptosis and cell-cycle arrest. Δ133p53 is expressed from a specific mRNA, p53I4, driven by an alternative promoter P2 located between intron 1 and exon 5 of *TP53* gene. Here, we report that the P2 promoter is regulated in a p53-dependent manner. Δ133p53 expression is increased in response to DNA damage by doxorubicin in p53 wild-type cell lines, but not in p53-mutated cells. Chromatin immunoprecipitation and luciferase assays using P2 promoter deletion constructs indicate that p53 binds functional response elements located within the P2 promoter. We also show that Δ133p53 does not bind specifically to p53 consensus DNA sequence *in vitro*, but competes with wild-type p53 in specific DNA-binding assays. Finally, we report that Δ133p53 counteracts p53-dependent growth suppression in clonogenic assays. These observations indicate that Δ133p53 is a novel target of p53 that may participate in a negative feedback loop controlling p53 function.
Oncogene (2010) **29,** 2691–2700; doi:10.1038/onc.2010.26; published online 1 March 2010

Keywords: p53; Δ133p53 isoform; promoter; transcriptional regulation; genotoxic stress

Introduction

The tumor suppressor *TP53* gene is one of the most frequently altered genes in human cancers (Petitjean *et al*., 2007). The expression of its product, the p53 protein, is tightly controlled through protein interaction with the E3-ubiquitin ligase Hdm2 that induces p53 proteasome-dependent degradation (Haupt *et al*., 1997; Kubbutat *et al*., 1997). Under stress, p53 is post-translationally modified, escapes to Hdm2-mediated degradation, accumulates in the nucleus and regulates the transcription of several target genes involved in growth suppressive responses including cell-cycle arrest, senescence and apoptosis. Among cell-cycle arrest genes, $p21^{WAF1/CIP1}$ encodes a cyclin-dependent kinase inhibiting cyclin:CDK complexes at both G1/S and G2/M (el-Deiry *et al*., 1993; Waldman *et al*., 1995). p53-dependent apoptosis is mediated through several distinct pathways involving genes, such as *BAX* or *PUMA*, that induce mitochondrial apoptosis through caspase activation (Miyashita and Reed, 1995; Villunger *et al*., 2003). A complex and critical question is to understand the molecular mechanisms by which p53 induces different types of suppressive responses, depending on the nature and intensity of the stress as well as on the cell and tissue characteristics (Vousden, 2006).

TP53 has been shown to produce nine isoforms (Bourdon *et al*., 2005). These isoforms combine three different N-terminal transactivation domains (TADs) (the full-length N-terminus present in p53, Δ40 and Δ133) with three different C-terminal oligomerization domains (α, β and γ, generated by alternative splicing) (Marcel and Hainaut, 2009). The Δ40p53 and Δ133p53 isoforms lack, respectively, 39 residues, which carry the main TAD, or 132 residues, corresponding to the TAD, the proline-rich region and the proximal part of the DNA-binding domain (Figure 1a) (Courtois *et al*., 2002; Bourdon *et al*., 2005). Δ40p53 isoform binds p53 response elements (p53REs), but has no intrinsic transactivation capacity, resulting in the inhibition of p53 transcriptional and growth suppressive functions (Courtois *et al*., 2002; Ghosh *et al*., 2004). Recent studies have shown that Δ133p53 can also inhibit both p53 transcriptional activity in reporter assays and p53-dependent apoptosis (Bourdon *et al*., 2005; Murray-Zmijewski *et al*., 2006). Thus, p53 isoforms may regulate some aspects of p53 functions. The expression of both p53 and Δ40p53 is controlled by the P1 promoter located upstream of exon 1 (Figure 1a) (Tuck and Crawford, 1989; Ghosh *et al*., 2004). A second promoter, P2, has been identified within *TP53* gene, in a region that spans the distal part of intron 1 to the proximal part of exon 5, covering about 1.5 kb (Bourdon *et al*., 2005). The resulting transcript, p53I4, carries the 3' end of intron 4 followed by all exons from

Correspondence: Dr P Hainaut, Molecular Carcinogenesis and Biomarkers Group, International Agency for Research on Cancer, 150 cours Albert Thomas, Lyon Cedex 08, 69372, France.
E-mail: hainaut@iarc.fr
Received 27 May 2009; revised 13 January 2010; accepted 19 January 2010; published online 1 March 2010

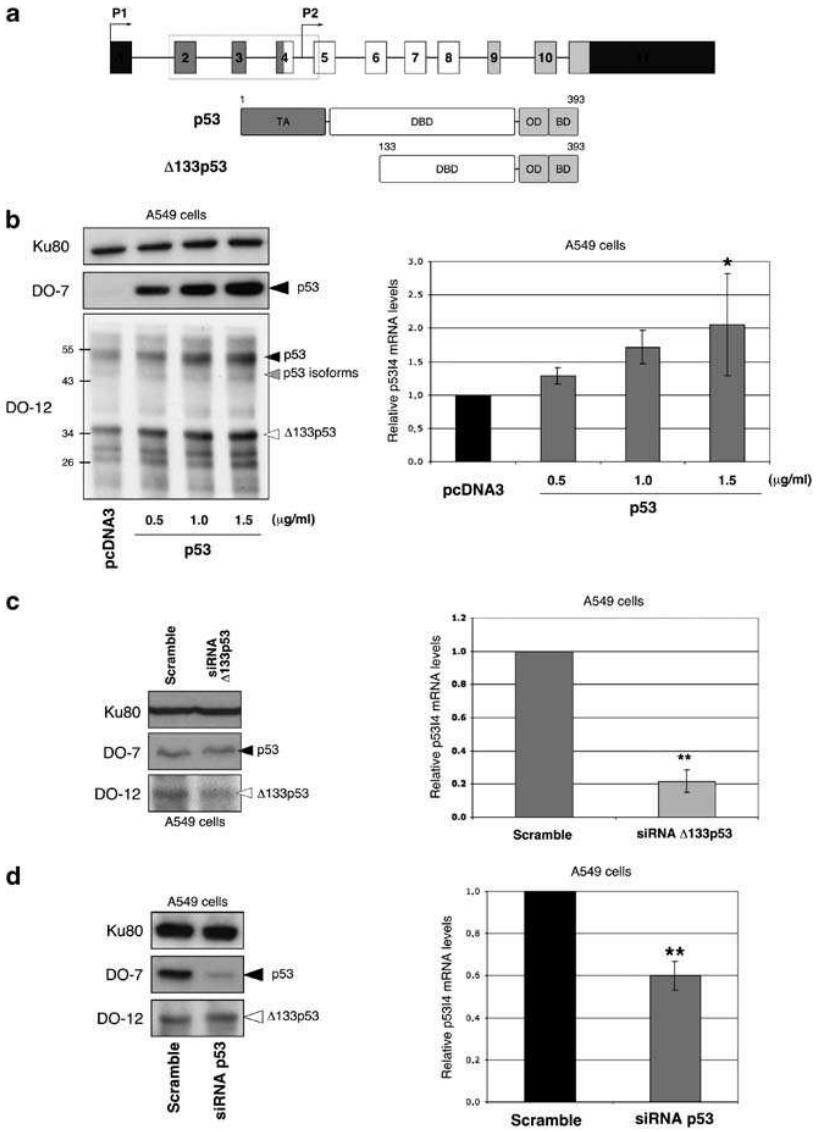


Figure 1 Correlation between p53 and Δ133p53 expression. (**a**) The *TP53* gene contains two promoters. The proximal P1 promoter regulates the expression of the p53 protein, which contains a transactivation domain (TAD), a DNA-binding domain (DBD), an oligomerization domain (OD) and a basic domain (BD). Fragment from intron 1 to exon 5 (box) contains the alternative P2 promoter that regulates the expression of the truncated Δ133p53 isoform. (**b**) Increased amounts of pcDNA3-p53 vector, expressing p53 cDNA, were transfected in A549 cells. p53 protein was detected by western blot 48 h post-transfection using DO-7 antibody (black arrowhead) and Δ133p53 using DO-12 antibody in the same extracts (white arrowhead) (left panel). p53I4 mRNA levels were quantified by real-time PCR using a forward primer hybridizing within intron 4 (left panel). Increased expression of p53 was correlated with an increased expression of Δ133p53 at both protein and mRNA levels. Ku80: loading control; $*P<0.05$; $**P<0.01$. (**c**) siRNA-targeting p53I4 mRNA, which encodes Δ133p53 isoform, was introduced in A549 cells. The expression of p53I4 mRNA (left panel) and expression of p53 and Δ133p53 proteins (right panel) were analyzed, respectively, by real-time PCR or western blot. siRNA treatment reduces the expression levels of both p53I4 mRNA and Δ133p53 protein, whereas it not affects p53 expression levels. (**d**) In A549 cells, p53 expression was specifically knocked-down using a siRNA directed at exon 4 and Δ133p53 expression was analyzed at protein (left panel) and mRNA (right panel) levels. Decrease of p53 expression was correlated with a decreased expression of Δ133p53 at mRNA, but not at protein levels. Ku80: loading control. The DO-7 western blot was exposed for 1 min in (**a**) and for 10 min in (**b**).

5 to 11 and supports Δ133p53 synthesis. However, the regulation of the P2 promoter is still unknown.

TP53 belongs to a family that contains two other members, *TP63* and *TP73*, with strong functional and structural similarities (Kaghad *et al*., 1997; Yang *et al*., 1998). These two genes contain an internal promoter, which regulates the production of N-truncated isoforms (Kaghad *et al*., 1997; Ishimoto *et al*., 2002). The p53 protein has been shown to bind the internal promoter of *TP63* and *TP73* and to act either as a negative or positive transcription factor (Grob *et al*., 2001; Harmes *et al*., 2003). These observations led us to investigate

whether the *TP53* internal P2 promoter may share similar functional characteristics. We show that the P2 promoter of *TP53* contains several p53REs, which can drive the expression of a reporter gene in a p53-dependent manner. Furthermore, we show that p53 binds P2 promoter in intact cells and regulates the expression of Δ133p53 in response to DNA damage. Finally, we provide evidence that Δ133p53 counteracts growth suppression by p53 in a clonogenic assay, suggesting a function of Δ133p53 in p53-regulatory loop.

Results

Δ133p53 expression correlates with p53 protein levels

To determine the function of p53 in Δ133p53 expression, we first overexpressed p53 in p53-expressing A549 cell lines using pcDNA3-p53 vector that contains p53 cDNA (Figure 1b). The p53 protein levels were verified using DO-7 antibody, which recognizes an epitope (18–23) present in the N-terminal domain of p53, but absent in Δ133p53 (Figure 1b, left panel). To detect Δ133p53 protein, we used DO-12 antibody, which recognizes a part of the DNA-binding domain common to both proteins (260–264), but allows to distinguish them on the basis of their different electrophoretic motilities (Supplementary Figure 1a). To verify that the 34 kDa band corresponds to Δ133p53, we used a siRNA that specifically targets p53I4 mRNA, encoding Δ133p53 protein. As shown in Figure 1c and in Supplementary Figure 1b, this siRNA significantly reduced both p53I4 mRNA levels and the intensity of the 34 kDa band, while not affecting the levels of the canonical p53 protein. After p53 overexpression, an increase in Δ133p53 protein was observed that correlated with increased amounts of transfected pcDNA3-p53 vector and of expressed p53 protein (Figure1b; Supplementary Figure 1a). A similar increase in p53I4 mRNA was also detected in these conditions with a significant twofold increase in the presence of 1.5 μg/ml of pcDNA3-p53-transfected vector (Figure 1b, right panel). Comparable effects were observed in another cell line expressing detectable basal levels of wild-type p53, MN1 (data not shown). These observations show that overexpression of p53 induces an increase in Δ133p53 isoform at both mRNA and protein levels.

To further show that p53 regulates p53I4 mRNA levels, we silenced p53 expression in A549 cells. A siRNA-targeting exon 4 of *TP53* gene was used to selectively eliminate full-length, but not p53I4 transcript. At protein level, p53 silencing did not result in a detectable variation of Δ133p53 despite reducing p53 protein levels by $>70\%$ (Figure 1c, left panel). In contrast, at mRNA level, a significant decrease of 40% in p53I4 was observed (Figure 1c, right panel). These data show that removal of p53 in A459 cells decreases basal levels of p53I4 mRNA, but not of Δ133p53 protein. This suggests that p53 silencing affects Δ133p53 transcription, but that within the time course of this experiment, the effect of this decrease is not detectable at the protein level, maybe because of its relatively long half-life.

Δ133p53 expression is modulated in a p53-dependent manner in response to stress

To investigate the possible regulation of Δ133p53 by p53 in response to stress, we assessed p53I4 mRNA expression in response to DNA damage induced by doxorubicin in A459 cells (Figure 2a). Protein accumulation of p53 and Hdm2, a p53-target gene, was verified by western blot in response to two different doses of doxorubicin at 8 h of treatment. We observed a dose-dependent increase in Δ133p53 protein expression correlated with a concordant increase in p53I4 mRNA, reaching a significant threefold increase at the highest dose. However, this dose-dependent increase of Δ133p53 is limited in cells transfected with a siRNA-targeting p53I4 mRNA (Supplementary Figure 1c). These observations indicate that Δ133p53 is induced by p53 in response to DNA damage.

To assess the impact of p53 on Δ133p53 expression at basal levels, we next used the MN1 breast cancer cell line grown in estrogen-free conditions (Figure 2b). This condition has been shown to reduce basal levels of p53 in estrogen-dependent breast cancer cell lines (Hurd *et al*., 1997). At protein level, estrogen deprivation resulted in a marked decrease in p53 and Hdm2 expression after long deprivation times. These variations were paralleled with matched variations in Δ133p53 expression, at both protein and mRNA levels. Using real-time PCR, the decrease after 12 days of deprivation (20% decrease) was statistically significant. Together, these observations indicate that treatments regulating basal levels of p53 protein, either positively or negatively, exert matched effects on Δ133p53 mRNA and protein levels.

Δ133p53 isoform interferes with p53 DNA-binding and growth suppression

It has been shown that Δ133p53 isoform can form hetero-oligomers with p53 protein and can inhibit p53-dependent transcription in luciferase assays (Murray-Zmijewski *et al*., 2006; Chen *et al*., 2009). To investigate whether Δ133p53 isoform may modulate p53 DNA-binding activity, we performed electro-mobility-shift assays (EMSA) using a ^{32}P-labeled p53RE consensus oligonucleotide (Figure 3a). Western blot confirmed the presence of the two proteins at expected levels (Supplementary Figure 2a). In the presence of p53 alone, a shift was observed, which was both super-shifted and stabilized by the addition of the monoclonal PAb421 antibody, as described earlier (Figure 3a) (Verhaegh *et al*., 1997). In the presence of Δ133p53, no shift was observed, whereas in the presence of the two proteins, Δ133p53 seems to significantly decrease the super-shift generated by the formation of DNA:p53:PAb421 complex (Figure 3a; Supplementary Figure 2b). These observations indicate that Δ133p53 does not bind p53RE DNA on its own, which is consistent with the

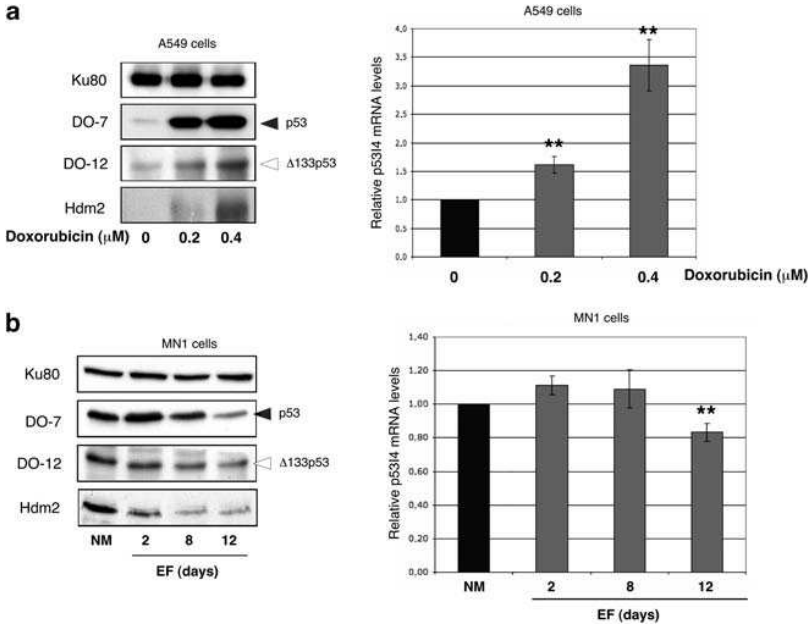


Figure 2 p53-dependent modulation of Δ133p53 expression. (**a**) A549 cells were treated during 8 h with the DNA-damaging drug doxorubicin. The induction of p53, Δ133p53 and Hdm2 was determined by western blot (left panel) and the expression of p53I4 mRNA was analyzed using real-time PCR (right panel). Ku80: loading control; $^{**}P<0.01$. (**b**) Wild-type p53-expressing MN1 cells were cultured for up to 12 days in estrogen-free medium (EF). p53, Hdm2 and Δ133p53 proteins as well as p53I4 mRNA were analyzed as in (**a**). Modulation of p53 by estrogen deprivation correlated with Δ133p53 protein and p53I4 mRNA variations. NM: normal medium.

notion that it lacks L1 loop of the DBD (residues 117–142) (Bourdon *et al*., 2005). Furthermore, it reduces p53 DNA-binding activity, maybe by forming hetero-oligomers unable to bind p53RE DNA. To verify this hypothesis, EMSA were conducted using DO-7 antibody in addition to PAb421 (Supplementary Figure 2b). Indeed, hetero-oligomers containing Δ133p53, which lacks DO-7 epitope, would undergo an incomplete shift in the presence of DO-7 antibody. No such shift was detected, suggesting that p53:Δ133p53 complexes are unable to bind specifically to p53RE DNA. The above results suggest that Δ133p53 isoform may inhibit p53 transcriptional capacity by inhibiting its DNA-binding activity.

To determine whether such an effect may inhibit p53 suppressive function, a colony formation assay was performed using p53-null H1299 cells transfected with vectors expressing p53, Δ133p53 or both (Figure 3b). After 14 days of selection, cells transfected with empty vector were spread over the surface of the well. Conversely, expression of p53 drastically reduced colony formation, in agreement with its growth suppressive activity. When compared with empty vector, cells transfected with Δ133p53 tended to form less colonies; however, Δ133p53 did not exert a suppressive effect similar to the one of p53. When co-transfected at equal amounts, Δ133p53 at least partially prevented the suppressive effect of p53. These observations suggested that Δ133p53 may counteract p53 suppressive function.

p53 transactivates the internal P2 promoter of TP53 gene

Δ133p53 expression at mRNA levels was correlated with the expression levels of p53 protein, suggesting that p53 transactivates the internal P2 promoter. To test this hypothesis, we generated a luciferase reporter driven by the P2 promoter sequence identified by Bourdon *et al*. (2005), consisting in a 1.5 kb fragment from intron 1 to exon 5 of *TP53* gene. This luciferase reporter was co-transfected in p53-null H1299 cells with various amounts of pcDNA3-p53 expression vector (Figure 4a, right panel). The expression of exogenous p53 protein and of its target gene Hdm2 was verified by western blot (Figure 4a, left panel). The basal transcriptional activity of the P2 promoter in the absence of p53 was fivefold higher than the one of the Basic promoterless luciferase vector. In the presence of pcDNA3-p53 vector at concentrations above 0.25 μg/ml, luciferase activity was significantly increased by about sixfold over the basal transcriptional activity of the P2 promoter. The luciferase activity was concordant with p53 expression levels (Figure 4a, left panel) and showed a dose-dependent increase with amounts of transfected pcDNA3-p53 vectors between 25 and 500 ng/ml (Supplementary Figure 3a). These results indicate that the P2 promoter activity is regulated by p53.

It has been reported that Δ40p53 isoform lacks part of the TAD and has no transcriptional activity toward reporter system carrying p53REs (Courtois *et al*., 2002; Ghosh *et al*., 2004). Different amounts of Δ40p53 were transfected in the presence of the luciferase

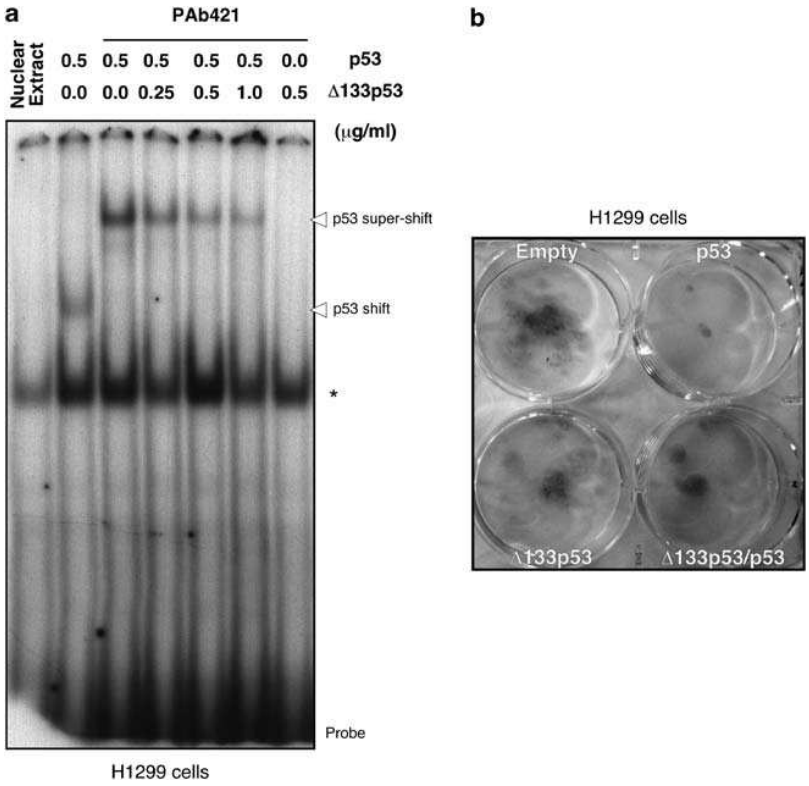


Figure 3 Effects of Δ133p53 on p53 biochemical and biological properties. (**a**) Electro-mobility-shift assays (EMSA) measuring DNA-binding activity. p53-null H1299 cells were co-transfected with a constant amount of pcDNA3-p53 vector (0.5 μg/ml) in the presence of increasing amounts of pcDNA3-Δ133p53 vector (0.25–1 μg/ml). Nuclear extracts were incubated with a ^{32}P-labeled p53RE oligonucleotide, with or without PAb421, which stabilizes and super-shifts p53:DNA complexes. DNA:p53 shifts and DNA:p53:PAb421 super-shift are indicated by arrowheads. *: non-specific band. (**b**) Colony formation assays were performed using p53-null H1299 cells transfected with 0.5 μg/ml of wild-type pcDNA3-p53 vector and/or −Δ133p53 vector during 14 days under 0.5 mg/ml of neomycin selection. As compared with wild-type p53, Δ133p53 did not suppress cell growth (60% of colonies compared with empty conditions) and partially reverted the suppression induced by wild-type p53 (10% in presence of p53 compared with 40% in presence of both p53 and Δ133p53).

reporter to investigate whether p53 TAD is required to modulate the P2 promoter activity (Supplementary Figure 3b). Compared with the significant increase in luciferase activity with pcDNA3-p53 vector, Δ40p53 alone has only minimal, non-significant effects on luciferase activity. These data suggest that p53 transcriptional activity is required to transactivate the P2 promoter.

p53REs are located within the P2 promoter

The p53 protein binds to response elements defined as repeats of a palindromic consensus sequence (5′-RRRCWWGYYY-3′) (el-Deiry *et al*., 1992). We used MatInspector software to identify putative p53RE by analyzing the fragment from intron 1 to exon 5 of *TP53* that corresponds to the P2 promoter (Cartharius *et al*., 2005). This DNA fragment contains four putative p53REs (Figure 4b; Supplementary Figure 4a). The first p53RE, termed as RE-1, is located in the 5′ end of intron 3. RE-2 is contained within exon 4, whereas the RE-3 and -4 are located 9 bp apart in intron 4 just upstream of the initiation start site of transcription. None of these REs shows a perfect match with the p53 consensus, the degree of matching varying from 60% (RE-3) to 85% (RE-2) (Supplementary Figure 4b).

To evaluate the contribution of these p53REs, we generated four deletion mutants of the promoter into the luciferase reporter (Figure 4b). In the absence of p53, these reporters showed differences in their basal transcriptional activity, suggesting that they may contain different regulatory elements in addition to p53 (Supplementary Figure 5a). These differences in basal activities have been normalized to evaluate the activation induced by p53. As shown in Figure 4b, p53 clearly stimulated the activity of the three P2 luciferase constructs (9–10-fold over Basic): P2 (containing RE-1 to RE-4); P2Δ1 (RE-1 to RE-4) and P2Δ2 (RE-2 to RE-4). However, it had only minor effects on P2Δ3 (RE-3/4) and P2Δ4 (no RE) (less than twofold over Basic). These results suggest that, among the four putative REs, RE-2 may have the major effect.

To investigate the function of RE-2 in greater details, we introduced mutations at critical bases of RE-2 by successive rounds of site-directed mutagenesis (Figure 4c). The resulting reporter constructs, termed as P2 Mut1 and Mut2, carry, respectively, 2 or 4 introduced mismatches with the consensus sequences. In the presence of the

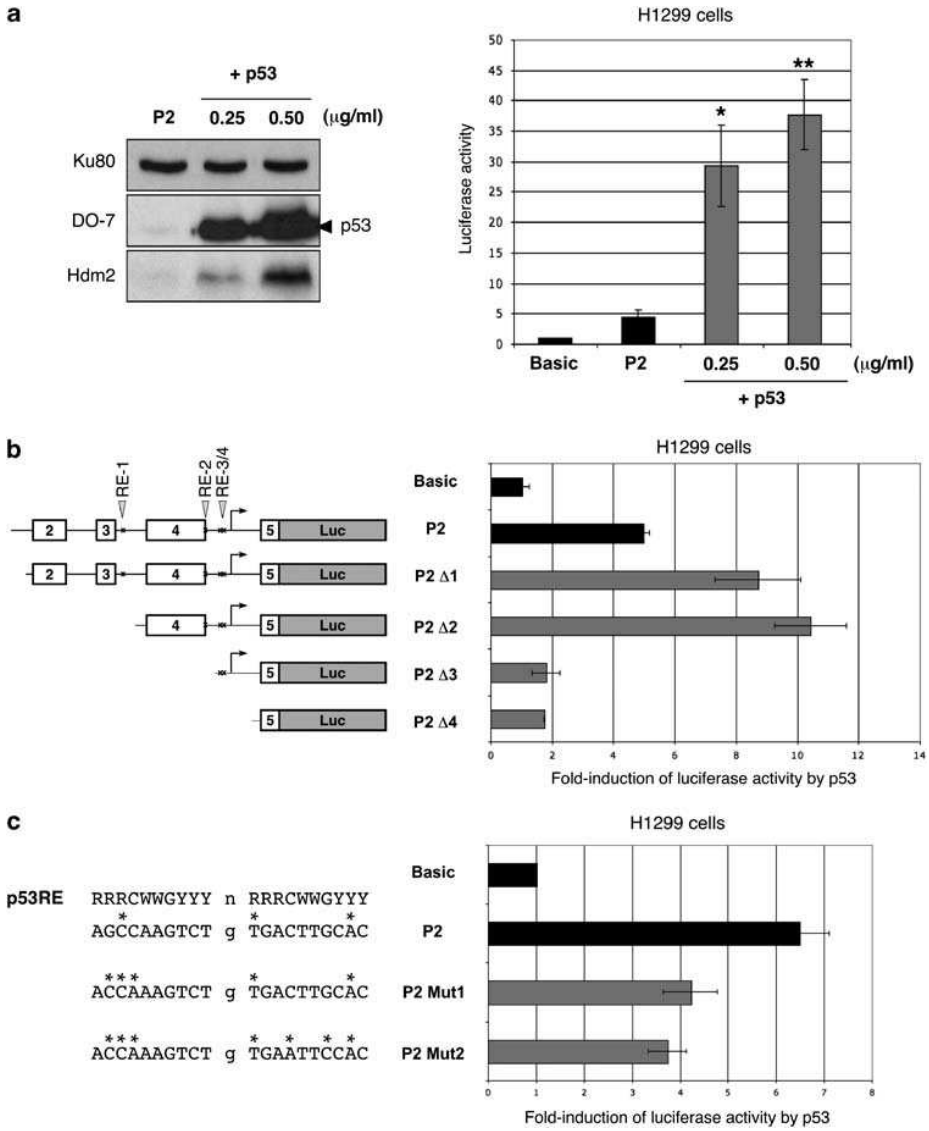


Figure 4 P2 promoter activity is regulated by p53. (**a**) In p53-null H1299 cells, luciferase assays (right panel) were performed using a luciferase reporter under the control of the 1.5 kb segment of *TP53* from intron 1 to exon 5, in the presence of increasing amounts of pcDNA3-p53, the expression of which was verified by western blot (left panel). Increase in luciferase activity was concordant with p53 protein levels. Basic: promoterless luciferase reporter; P2: luciferase reporter driven by *TP53* segment; Ku80: loading control; $*P<0.05$; $**P<0.01$. (**b**) MatInspector software (Cartharius *et al*., 2005) was used to predict four putative p53-binding sites (RE-1 to -4) in the promoter P2. Luciferase reporter deletion constructs were engineered as pictured and tested in the presence of p53. Mean fold induction was calculated as the ratio between luciferase activity of each construct in the presence of pcDNA3-empty or -p53 vector. p53 induced at least a fivefold increase of luciferase activity using P2, Δ1 or Δ2. In contrast, no induction was observed using Δ3 and Δ4. (**c**) Luciferase assays (right panel) were performed in cells co-transfected with pcDNA3-p53 vector and with either one of two plasmids carrying mutations in RE-2 (P2 Mut1 and P2 Mut2; left panel). Mutations in p53RE induced a 40% decrease of luciferase activity. p53RE: consensus sequence of p53RE; star: mismatch between p53RE and RE-2 in the promoter P2.

pcDNA3-empty vector, the three P2 plasmids exhibited the same luciferase activity, indicating no abolition of their basal transcriptional activities (Supplementary Figure 5b). In the presence of pcDNA3-p53 vector, the promoter activities of P2 Mut1 and Mut2 were reduced by about 40% as compared with wild-type P2 (Figure 4c). These observations suggest that RE-2 located in exon 4 participate in the p53-dependent regulation of P2 promoter, but does not account on its own for the full activation of P2 by p53.

p53 binds the internal promoter P2 of TP53 gene

We used chromatin immunoprecipitation (ChIP) to analyze the binding of p53 on different segments of the P2 promoter in A549 cells in basal conditions. The DO-7 antibody was used to immunoprecipitate p53 and non-immunized IgG were used as negative control (Figure 5). After ChIP, five PCR products were analyzed (Figure 5a). As controls, we used primers designed to amplify the $p21^{WAF1/CIP1}$ promoter (positive

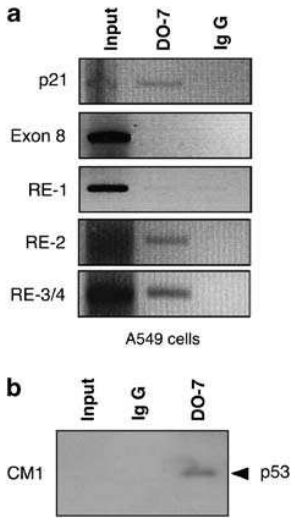


Figure 5 Binding of p53 on P2 promoter. (**a**) Chromatin immunoprecipitation assay (ChIP) for p53 protein onto P2 promoter in wild-type p53-expressing A549 cells in basal conditions. Purified DNA immunoprecipitated by p53 was analyzed by PCR using specific primers to amplify regions located in $p21^{WAF1/CIP1}$ promoter (positive control), in exon 8 of *TP53* gene (negative control) and in three parts of promoter P2 containing the putative p53REs RE-1, RE-2 and RE-3/4). Input: ChIP positive control; IgG: ChIP negative control; DO-7: immunoprecipitation using DO-7 antibody. (**b**) In A549 cells, p53 was immunoprecipitated using DO-7 antibody and immunodetected using CM1. A band of low intensity was detected in DO-7 immunoprecipitated sample.

control) and exon 8 of *TP53* gene (negative control). ChIP with IgG did not generate any amplification. In contrast, ChIP with DO-7 antibody generated an amplification signal for $p21^{WAF1/CIP1}$ promoter, but not for *TP53* exon 8. With P2 promoter fragments, no amplification was observed with primers targeting a DNA fragment containing RE-1 (Figure 5a). In contrast, amplifications were observed for fragments encompassing RE-2 and RE-3/4. These observations suggest that p53 physically binds several p53REs located between exon 4 and intron 4, but not in intron 3.

Mutant p53 does not induce Δ133p53 expression

To further assess the specificity of the wild-type p53 effect on the P2 promoter, we analyzed its regulation in response to DNA damage by doxorubicin in several cell lines expressing p53 with different functional status (Figure 6a). In wild-type p53 cells, doxorubicin treatment induced both p53 and Δ133p53 expression as expected. In contrast, in three cell lines expressing a non-functional p53 protein, treatment with doxorubicin had either no or only minimal effects on the expression of p53 and Δ133p53.

These observations were extended by analyzing the effects of several common 'hotspot' *TP53* mutants on the expression of the P2 luciferase reporter (Figure 6b). In contrast to wild-type p53, none of these mutants succeeded in activating the reporter. These observations support the notion that only wild-type p53 with functional capacities transactivates the P2 promoter leading to Δ133p53 expression.

Discussion

Recently, it has been shown that *TP53* gene is expressed as several isoforms (Bourdon *et al*., 2005). Among them, only Δ133p53 is regulated by an alternative promoter, P2, located within *TP53* (Bourdon *et al*., 2005). Its overexpression has been observed in squamous cell carcinoma of the head and neck as well as in breast cancers, suggesting a function of Δ133p53 in tumorigenesis (Bourdon *et al*., 2005; Boldrup *et al*., 2007). However, the biological regulation and function of Δ133p53 are still unknown. Here, we have analyzed the P2 promoter and we show that its activity is regulated by p53. First, modulation of wild-type p53 expression induces Δ133p53 expression at both mRNA and protein levels. Second, the P2 promoter contains p53REs, which are involved in p53 responsiveness. Third, Δ133p53 prevents the binding of wild-type p53 to DNA segment containing its consensus sequence and inhibits p53 suppressive activity in a clonogenic assay. These data suggest that Δ133p53 may contribute to an auto-regulatory process by which p53 controls its own suppressor function.

At the time of Δ133p53 identification, several evidences supported the presence of an internal promoter involved in the production of Δ133p53 (Bourdon *et al*., 2005). A transcript, p53I4, initiated within intron 4, but exhibiting the same splicing pattern as full-length p53 mRNA downstream of exon 5, was observed in several human cell lines, in Drosophilia and Zebrafish (Bourdon *et al*., 2005; Chen *et al*., 2005). In addition, a 1.5 kb segment of *TP53* spanning from intron 1 (+10810) to exon 5 (+12396) conferred a sixfold induction to a luciferase reporter (Bourdon *et al*., 2005). In this study, we confirm that the same *TP53* sequence possesses a promoter activity (fivefold induction of reporter) and we show that this segment confers responsiveness toward wild-type, but not mutant p53. Furthermore, using MatInspector software, we identified four putative p53REs upstream of the transcription initiation site. These p53RE have a degree of degeneration, which is in the same range as p53REs of many common p53-target genes (Menendez *et al*., 2007). Using luciferase and ChIP experiments, we show that transcriptional activation by p53 is dependent on two contiguous regions containing p53REs: one in the 3'-end of exon 4 that correlates with the main p53 responsiveness and one in intron 4. These observations correlate with the one of Chen *et al*. (2009), who recently reported that the internal P2 promoter of *TP53* Zebrafish was induced by p53 through two close p53REs located in intron 4, just upstream the initiation site of transcription. These authors suggest that the two response elements operate in a cooperative manner. These data suggest that human Δ133p53 isoform and its zebrafish

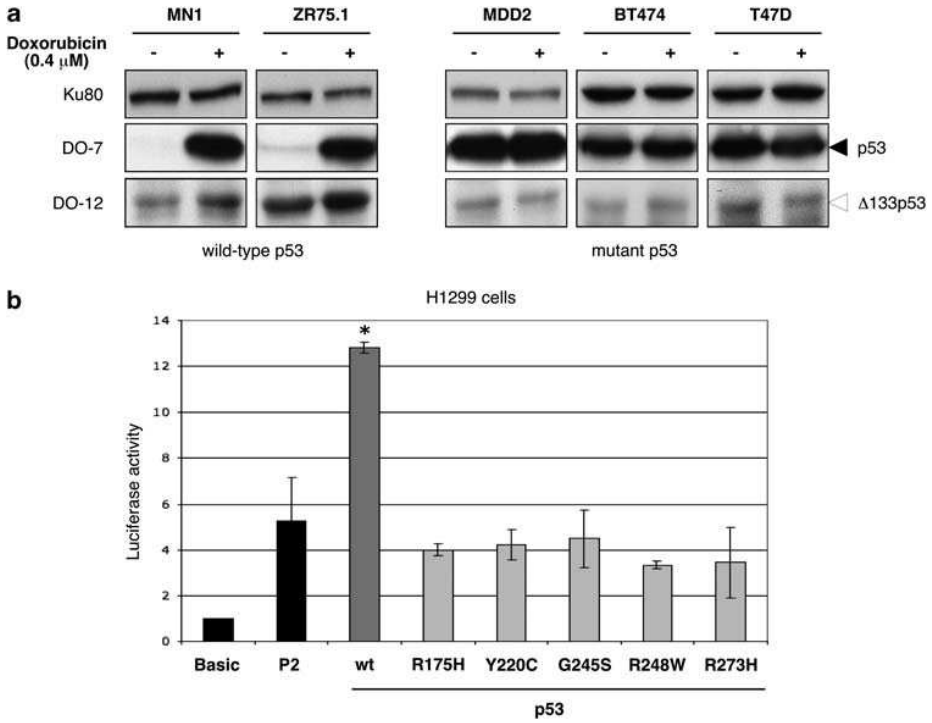


Figure 6 Effect of mutant p53 on Δ133p53 expression. (**a**) Breast cancer cells carrying different *TP53* mutation status were treated for 8 h with doxorubicin at 0.4 μM. Protein extracts were analyzed by western blot to determine the expression of p53 (DO-7) and of Δ133p53 (DO-12). p53 mutant cells overexpress an inactive p53 protein. Ku80: loading control. (**b**) Luciferase assays were performed in the presence of different mutant p53 as indicated. As compared with wild-type p53, mutant p53 had no effect on P2 promoter activity. Wt: wild-type p53; $*P<0.05$.

homolog are regulated by a conserved mechanism involving p53 protein.

In addition to *in vitro* data, we reported p53-dependent regulation of Δ133p53 isoform in cancer cell lines expressing wild-type, but not mutant p53. To detect Δ133p53, we have used the monoclonal antibody DO-12 that recognize a p53 epitope (260–264). This antibody has been described as the most appropriate to detect Δ133p53 among a panel of seven common p53 antibodies. Using siRNA-targeting p53I4 mRNA to specifically knock-down Δ133p53, we verified that the 34 kDa band indeed corresponds to Δ133p53 protein (Supplementary Figure 1b) (Bourdon, 2007). In response to DNA damage induced by doxorubicin, an increase in Δ133p53 expression followed the accumulation of p53 by about 8 h, a time-course compatible with that of products of other p53-regulated genes such as *Hdm2*, *BAX* or $p21^{WAF1/CIP1}$ (Figure 2a and data not shown). In Zebrafish, a similar lag of 6 h has been observed between induction of p53 and Δ113p53 on ribosomal stress (Danilova *et al*., 2008). The p53-dependent regulation of Δ133p53 expression in response to stress was also reported in response to γ-rays exposure (Chen *et al*., 2009). Although several reports have described induction of Δ133p53 expression in response to stress, no data on basal activity of P2 promoter have been reported. Here, using siRNA and estrogen deprivation that both results in a reduction of basal p53 expression, we observed a decrease in Δ133p53 expression at mRNA levels. These observations suggest that p53 has an important function in the maintenance of Δ133p53 expression at basal levels. Although p53 siRNA did not have a detectable effect on Δ133p53 protein expression after 48 h, this may be explained by the fact that Δ133p53 lacking the Hdm2-binding domain may escape Hdm2-mediated degradation, being thus more stable than p53 (Haupt *et al*., 1997; Kubbutat *et al*., 1997).

In contrast to p53, Δ133p53 lacks the first 30 residues of the DNA-binding domain encoding the L1 loop and part of the loop-sheet-helix motif that form the structure binding to the major groove of DNA in p53RE (Cho *et al*., 1994). As expected, we observed that this protein is unable to bind DNA containing the p53 consensus sequence using the same *in vitro* conditions than for p53. However, Δ133p53 could decrease the binding capacity of p53 on p53RE. As it has been described that p53 and Δ133p53 can form hetero-tetramers (Chen *et al*., 2009), our data suggest that Δ133p53 tetramers lack sequence-specific DNA-binding capacity. The inhibition of p53-binding capacities by Δ133p53 is in agreement with the inhibitory effect of Δ133p53 on both p53 transcriptional activity and p53-dependent apoptosis (Bourdon *et al*., 2005; Murray-Zmijewski *et al*., 2006). Here, we provide support for long-term anti-suppressive effects of Δ133p53 by showing that co-expression of Δ133p53

and p53 could partially restore the proliferation capacity of H1299 cells in a clonogenic assay. However, despite the accumulation of evidences suggesting that Δ133p53 may counteract p53 suppressive function, there is still limited information available on the physiological function of Δ133p53. One of the problems is that Δ133p53 seems to be expressed at low levels as compared with p53, in particular after DNA damage. Recent results in Zebrafish shed a light on a possible mechanism of action (Chen *et al*., 2009). They observed that the Zebrafish homologous of Δ133p53 could antagonize p53-induced apoptosis through activating Bcl-2L (homologous of the human Bcl-xL), and that knockdown of Δ113p53 enhanced p53-mediated apoptosis under stress conditions. These results are consistent with the ones reported here and indicate that small changes of Δ133p53 may contribute to a novel feedback pathway that modulates the p53 response.

In this study, we observed that the internal P2 promoter involved in Δ133p53 expression is regulated by p53. This feature is a familial trait also observed in the other *TP53* gene family members, *TP63* and *TP73*, which express N-truncated proteins that are regulated by internal promoters responsive to wild-type p53 (Grob *et al*., 2001; Harmes *et al*., 2003). These N-truncated proteins participate in auto-regulatory feedback loops that regulate p53 suppressive functions. This model seems to be relevant for Δ133p53 isoform as well. Thus, Δ133p53 is another player involved in these auto-regulatory feedback loops, which may selectively interfere with p53 capacity to transactivate target genes. It remains to be determined in which cell type and stress type context this regulation may have a major function.

Materials and methods

Cell culture
Two human cell lines expressing wild-type *TP53* were employed: lung A459 and breast ZR75.1 cells. Human p53-null H1299 cells were used for luciferase assays and EMSA. Two human breast cell lines expressing mutant *TP53*, BT474 (E285K) and T47D (L194F) were used. Two isogenic breast cell lines derived from MCF-7 were employed, MN1 (stably transfected with an empty vector) and MDD2 (expressing a mini-peptide exerting negative effect on the endogenous wild-type p53) (Bacus *et al*., 1996). Cells were treated for 8 h with doxorubicin (0.2 or 0.4 μM). Breast MN1 cells were treated with medium deprived of both Red phenol and estradiol.

Vectors
Exogenous p53 isoforms (p53, Δ40p53 and Δ133p53) were produced using pcDNA3 vectors containing human cDNA of each isoforms (Courtois *et al*., 2002). Five mutant pcDNA3-p53 were used (R175H, Y220C, G245S, R248W and R273H). The pcDNA3-empty vector was used as a negative control. Several luciferase reporter vectors were developed using pGL3-Basic, a promoterless luciferase plasmid (Promega, Fitchburg, WI, USA). The pGL3-P2 (or P2) vector contains the human P2 promoter (from intron 1 to exon 5) upstream of *Luciferase* gene. Four deleted and two mutant vectors were constructed using P2 vector (Supplementary Table 1). The pGL3-Basic was used as negative control to standardize experiments and phRL-null plasmid as control of transfection efficiency to normalized experiments.

Transfection
The pcDNA3 expression vectors (1.5 μg/ml completed with pcDNA3 empty) were transfected using 9 μl of Fugene (Roche, Indianapolis, IN, USA) in A549 and H1299, harvested 48 h latter for overexpression and 24 h latter for EMSA assays. Silencing of p53 mRNA was performed in A549 cells seeded in antibiotic-free medium and directly transfected with 20 nM of either scramble or p53-targeted siRNAs (Eurogentec, Liege, Belgium) (Supplementary Table 1) using 12 μl of Hyperfect (Qiagen, Huntsville, AL, USA). Cells were transfected twice at 24 h interval and harvested after 48 h of treatment. For luciferase assays using Fugene (Roche), 500 ng/ml of pGL3 plasmids were transfected with 10 ng/ml of phRL-null and with 1 μg/ml of pcDNA3 expression vectors. Cells were harvested 48 h post-transfection and luciferase activity was analyzed using Dual Luciferase Reporter Assay System (Promega).

Real-time quantitative RT–PCR
Total RNA was extracted using Nucleospin RNA II kit (Machery-Nalgene, Duren, Germany) and reverse transcription was performed using random primers (Promega) and SuperScript II (Invitrogen, Carlsbad, CA, USA). RNA expression was analyzed by real-time quantitative RT–PCR using Brilliant SYBR Green Master Mix (Qiagen) and 0.4 μM of primers (Supplementary Table 1). The p53I4 mRNA expression level is normalized to that of Gapdh (lung cells) or of 28S (breast cells). The control of the experiment is used as the reference. All samples were analyzed in triplicates and two Q-PCRs were performed for each experiment.

Western blot
Proteins were extracted using RIPA-like buffer (50 mM Tris–HCl pH7.4, 250 mM NaCl, 0.1% SDS, 0.5% NP-40, 2 mM DTT, 1 mg/ml protease inhibitors) and were analyzed on a 10% SDS–polyacrylamide gel using several antibodies: DO-7 (monoclonal anti-p53 antibody specific of TAD; Dako, Cambridgeshire, UK); DO-12 (monoclonal anti-p53 antibody specific of DBD), which has been described as the most appropriate to detect Δ133p53 among a panel of seven common p53 antibodies (Bourdon, 2007); CM1 (polyclonal anti-p53 antibody; Novocastra, Newcastle, UK); Hdm2 (Abcam) and Ku80 (used as a loading control; Abcam). Detection was performed with the ECL kit (Amersham, Cambridge, MA, USA) by autoradiography.

Chromatin immunoprecipitation assay
In 100 mm plates, 3×10^6 A549 cells were seeded. The following day, 1% formaldhehyde was added onto cells for 10 min at room temperature to cross-link proteins to DNA. Sonication was performed (twice: 10 × (5 s on: 5 s off) at 21% amplitude) (Vibra cell 75041) and immunoprecipitation was carried out using 4 μg of DO-7 antibody (anti-p53) and protein G-agarose beads (Invitrogen). DNA:p53 complexes were dissociated by reverse cross-linking using NaCl 200 mM and DNA was extracted by phenol/chloroform. Immunoprecipitated DNA was amplified by PCR (Supplementary Table 1).

Electro-mobility-shift assay
Nuclear cellular proteins of H1299-transfected cells were extracted and incubated with mixture containing ^{32}P-radio-

llabeling oligonucleotide with p53RE consensus as described earlier (Supplementary Table 1) (Verhaegh *et al*., 1997). Two antibodies were used to shift DNA:p53 complex: the monoclonal PAb421 antibody recognizing p53 DBD and known to stabilize DNA:p53 complex; and the monoclonal DO-7 antibody specific of p53 protein.

Statistical analysis

Anova and Student's *t*-tests were performed using on-line tools (http://faculty.vassar.edu/lowry/VassarStats.html). In the histograms, black bars indicate the control used for statistics and *P*-values were indicated by (*) when $P<0.05$ and (**) when $P<0.01$.

Conflict of interest

The authors declare no conflict of interest.

Acknowledgements

VM is supported by la Ligue National Contre le Cancer. LF-C is supported by The Association for International Cancer Research (AICR). This project is funded by la Ligue Régionale du Rhône Contre le Cancer.

References

Bacus SS, Yarden Y, Oren M, Chin DM, Lyass L, Zelnick CR *et al*. (1996). Neu differentiation factor (Heregulin) activates a p53-dependent pathway in cancer cells. *Oncogene* **12**: 2535–2547.

Boldrup L, Bourdon JC, Coates PJ, Sjostrom B, Nylander K. (2007). Expression of p53 isoforms in squamous cell carcinoma of the head and neck. *Eur J Cancer* **43**: 617–623.

Bourdon JC. (2007). p53 and its isoforms in cancer. *Br J Cancer* **97**: 277–282.

Bourdon JC, Fernandes K, Murray-Zmijewski F, Liu G, Diot A, Xirodimas DP *et al*. (2005). p53 isoforms can regulate p53 transcriptional activity. *Genes Dev* **19**: 2122–2137.

Cartharius K, Frech K, Grote K, Klocke B. (2005). MatInspector and beyond: promoter analysis based on transcription factor binding sites. *Bioinformatics* **21**: 2933–2942.

Chen J, Ng SM, Chang C, Zhang Z, Bourdon JC, Lane DP *et al*. (2009). p53 isoform delta113p53 is a p53 target gene that antagonizes p53 apoptotic activity via BclxL activation in zebrafish. *Genes Dev* **23**: 278–290.

Chen J, Ruan H, Ng SM, Gao C, Soo HM, Wu W *et al*. (2005). Loss of function of def selectively up-regulates Delta113p53 expression to arrest expansion growth of digestive organs in zebrafish. *Genes Dev* **19**: 2900–2911.

Cho Y, Gorina S, Jeffrey PD, Pavletich NP. (1994). Crystal structure of a p53 tumor suppressor-DNA complex: understanding tumorigenic mutations. *Science* **265**: 346–355.

Courtois S, Verhaegh G, North S, Luciani MG, Lassus P, Hibner U *et al*. (2002). DeltaN-p53, a natural isoform of p53 lacking the first transactivation domain, counteracts growth suppression by wild-type p53. *Oncogene* **21**: 6722–6728.

Danilova N, Sakamoto KM, Lin S. (2008). Ribosomal protein S19 deficiency in zebrafish leads to developmental abnormalities and defective erythropoiesis through activation of p53 protein family. *Blood* **112**: 5228–5237.

el-Deiry WS, Kern SE, Pietenpol JA, Kinzler KW, Vogelstein B. (1992). Definition of a consensus binding site for p53. *Nat Genet* **1**: 45–49.

el-Deiry WS, Tokino T, Velculescu VE, Levy DB, Parsons R, Trent JM *et al*. (1993). WAF1, a potential mediator of p53 tumor suppression. *Cell* **75**: 817–825.

Ghosh A, Stewart D, Matlashewski G. (2004). Regulation of human p53 activity and cell localization by alternative splicing. *Mol Cell Biol* **24**: 7987–7997.

Grob TJ, Novak U, Maisse C, Barcaroli D, Luthi AU, Pirnia F *et al*. (2001). Human delta Np73 regulates a dominant negative feedback loop for TAp73 and p53. *Cell Death Differ* **8**: 1213–1223.

Harmes DC, Bresnick E, Lubin EA, Watson JK, Heim KE, Curtin JC *et al*. (2003). Positive and negative regulation of deltaN-p63 promoter activity by p53 and deltaN-p63-alpha contributes to differential regulation of p53 target genes. *Oncogene* **22**: 7607–7616.

Haupt Y, Maya R, Kazaz A, Oren M. (1997). Mdm2 promotes the rapid degradation of p53. *Nature* **387**: 296–299.

Hurd C, Khattree N, Dinda S, Alban P, Moudgil VK. (1997). Regulation of tumor suppressor proteins, p53 and retinoblastoma, by estrogen and antiestrogens in breast cancer cells. *Oncogene* **15**: 991–995.

Ishimoto O, Kawahara C, Enjo K, Obinata M, Nukiwa T, Ikawa S. (2002). Possible oncogenic potential of DeltaNp73: a newly identified isoform of human p73. *Cancer Res* **62**: 636–641.

Kaghad M, Bonnet H, Yang A, Creancier L, Biscan JC, Valent A *et al*. (1997). Monoallelically expressed gene related to p53 at 1p36, a region frequently deleted in neuroblastoma and other human cancers. *Cell* **90**: 809–819.

Kubbutat MH, Jones SN, Vousden KH. (1997). Regulation of p53 stability by Mdm2. *Nature* **387**: 299–303.

Marcel V, Hainaut P. (2009). p53 isoforms—a conspiracy to kidnap p53 tumor suppressor activity? *Cell Mol Life Sci* **66**: 391–406.

Menendez D, Inga A, Jordan JJ, Resnick MA. (2007). Changing the p53 master regulatory network: ELEMENTary, my dear Mr Watson. *Oncogene* **26**: 2191–2201.

Miyashita T, Reed JC. (1995). Tumor suppressor p53 is a direct transcriptional activator of the human bax gene. *Cell* **80**: 293–299.

Murray-Zmijewski F, Lane DP, Bourdon JC. (2006). p53/p63/p73 isoforms: an orchestra of isoforms to harmonise cell differentiation and response to stress. *Cell Death Differ* **13**: 962–972.

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.

Tuck SP, Crawford L. (1989). Characterization of the human p53 gene promoter. *Mol Cell Biol* **9**: 2163–2172.

Verhaegh GW, Richard MJ, Hainaut P. (1997). Regulation of p53 by metal ions and by antioxidants: dithiocarbamate down-regulates p53 DNA-binding activity by increasing the intracellular level of copper. *Mol Cell Biol* **17**: 5699–5706.

Villunger A, Michalak EM, Coultas L, Müllauer F, Böck G, Ausserlechner MJ *et al*. (2003). p53- and drug-induced apoptotic responses mediated by BH3-only proteins puma and noxa. *Science* **302**: 1036–1038.

Vousden K. (2006). Outcomes of p53 activation—spoilt for choice. *J Cell Sci* **119**: 5015–5020.

Waldman T, Kinzler KW, Vogelstein B. (1995). p21 is necessary for the p53-mediated G1 arrest in human cancer cells. *Cancer Res* **55**: 5187–5190.

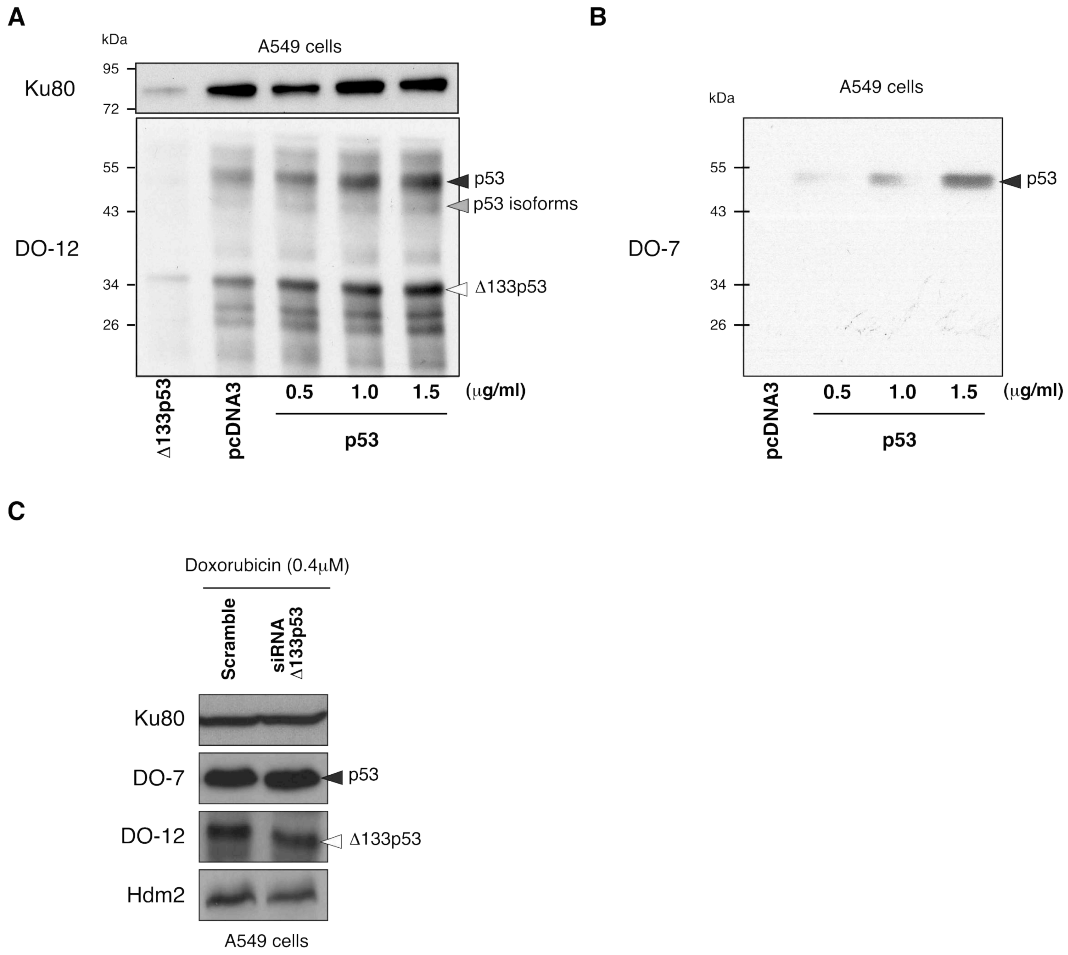
Yang A, Kaghad M, Wang Y, Gillett E, Fleming MD, Dotsch V *et al*. (1998). p63, a p53 homolog at 3q27-29, encodes multiple products with transactivating, death-inducing, and dominant-negative activities. *Mol Cell* **2**: 305–316.

Supplementary Information accompanies the paper on the Oncogene website (http://www.nature.com/onc)

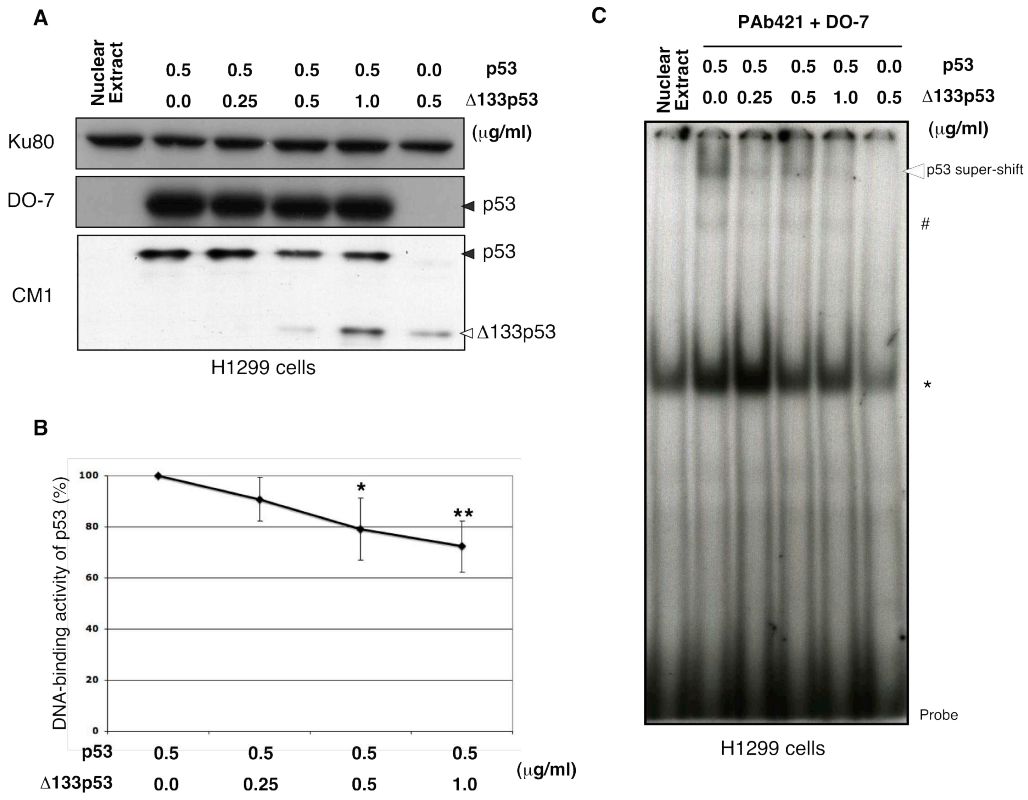
Supplementary materials

p53 regulates the transcription of its Δ133p53 isoform through specific response elements contained within the *TP53* P2 internal promoter

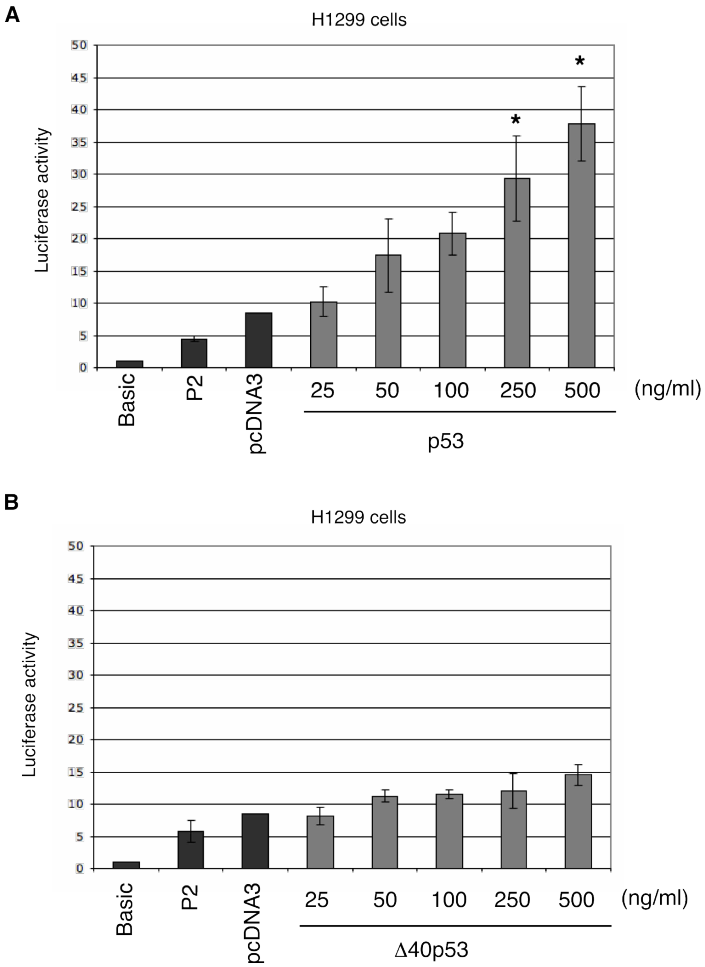
Virginie Marcel, Vineetha Vijayakumar, Lynnette Fernandez-Cuesta, Hind Hafsi, Charlotte Sagne, Agèns Hautefeuille, Magali Olivier, Pierre Hainaut



Supplementary Figure 1. Increased Δ133p53 expression dependent on p53 over-expression. In p53 expressing A549 cells, pcDNA3-p53 vector was transfected at 3 different doses and *TP53* expression levels were analyzed by western blot using polyclonal DO-12 antibody (loading of 50μg) (**A**) or DO-7 antibody (loading of 30μg) (**B**). pcDNA3: negative control; Δ133p53: positive control (Δ133p53 over-expression using pcDNA3-Δ133p53 vector, loading of 10μg); Ku80: loading control. (**C)** A549 cells were transfected with siRNA targetting p53I4 mRNA and, in addition, treated with doxorubicin. No change in p53 and Hdm2 protein levels was observed, while Δ133p53 protein levels decrease in presence of the siRNA.



Supplementary Figure 2. Modulation of p53 DNA-binding properties by Δ133p53 isoform. (**A**) Analysis by western blot of the same extracts as for EMSA (Figure 3a). pcDNA3-p53 and/or -Δ133p53 were co-transfected in p53-null H1299 cells to determine the effect of Δ133p53 isoform on p53 DNA-binding activity by EMSA. In nuclear extract, p53 and Δ133p53 expression were verified using DO-7 or CM1 polyclonal antibody. The expression levels of both p53 and Δ133p53 correspond to the amount of transfected pcDNA3 vectors. Ku80: loading control. (**B**) EMSA assays (Figure 3a) were quantified. *: $P<0.05$; **: $P<0.01$. (**C**) EMSA in the presence of PAb421 and DO-7 antibodies. The PAb421 monoclonal anti-p53 antibody recognizes both p53 and Δ133p53 isoform, while the DO-7 monoclonal antibody recognizes only p53 protein on an N-terminal epitope. In the presence of DO-7 antibody, only one band corresponding to the p53 supershift was observed. These observations indicate that Δ133p53 isoform prevents p53 DNA-binding capacity, most likely through the formation of heterotetramers unable to bind DNA. #: band that corresponds to a residual amount of canonical p53 protein not shifted by DO-7; *: unspecific band.



Supplementary Figure 3. p53-dependent transactivation of internal P2 promoter. (**A**) Activation of P2 promoter by p53. In p53-null H1299 cells, luciferase assays were performed using a reporter containing a fragment of *TP53* gene, which corresponds to the internal P2 promoter (from the end of intron 1 to the beginning of exon5), in the presence of increasing amount of pcDNA3-p53 vector. At 500 ng/ml of transfected p53, the luciferase activity was increased by about 4-fold. (**B**) Effect of Δ40p53 on P2 promoter activity. Similar assays were performed in presence of increasing amount of pcDNA3-Δ40p53 vector. No significant increase of luciferase activity was observed.

A

+10810

aaacattgga agagagaatg tgaagcagcc attcttttcc tgctccacag gaagccgagc tgtctcagac actggcatgg

tgttgggggа ggggcctcct cctctgcagg cccaggtgac ccagggttgg aagcgtctca tgctggatcc ccacttttcc

tcttgcag**ca gccagactgc cttccgggtc actgccatgg aggagccgca gtcagatcct agcgtcgagc cccctctgag**

tcaggaaaca ttttcagacc tatggaaact gtgagtggat ccattggaag ggcaggccac caccccgacc ccaaccccag

ccccctagca gagacctgtg ggaagcgaaa attcatggga ctgactttct gctcttgtct ttcag**acttc ctgaaaacaa**

cgttctggta aggacaaggg ttgggctggg accctggaggg ctgggggggc tgggggctg aggacctggt cctctgactg

ctcttttcac ccatctacag **tccccttgc cgtcccaagc aatggatgat ttgatgctgt ccccggacga tattgaacaa**

tggttcactg aagacccagg tccagatgaa gctcccagaa tgccagaggc tgctccccgc gtggcccctg caccagcagc

tcctacaccg gcggcccctg caccagcccc ctcctggccc ctgtcatctt ctgtcccttc ccagaaaacc taccagggca

gctacggttt ccgtctgggc ttcttgcatt ctgggacagc caagtctgtg acttgcacgg tcagttgccc tgaggggctg

gcttccatga gacttcaatg cctggccgta tccccctgca tttcttttgt ttggaacttt gggattcctct tcaccctta

ggcttcctgt cagtgttttt ttatagttta cccacttaat gtgtgatctc tgactcctgt cccaaagttg aatattcccc

ccttgaattt gggcttttat ccatcccatc acaccctcag catctctcct ggggatgcag aacttttctt tttcttcatc

cacgtgtatt ccttggcttt tgaaaataag ctcctgacca ggcttggtgg ctcacacctg caatcccagc actctcaaag

aggccaaggc aggcagatca cctgagcccc aggagttcaa gaccagcctg ggtaacatga tgaaacctcg tctctacaaa

aaaatacaaa aaattagcca ggcatggtgg tgcacaccta tagtcccagc cactcaggag gctgaggtgg gaagatcact

+1

tgaggccagg agatggaggc tgcagtgagc tgtgatcaca ccactgtgct ccagcctgag tgacagagca agaccctatc

tcaaaaaaaa aaaaaaagaa aagctcctga ggtgtagacg ccaactctct ctagctcgct agtgggttgc aggaggtgct

tacacatgtt tgtttctttg ctgccgtgtt ccagttgctt tatctgttca cttgtgccct gactttcaac tctgtctcct

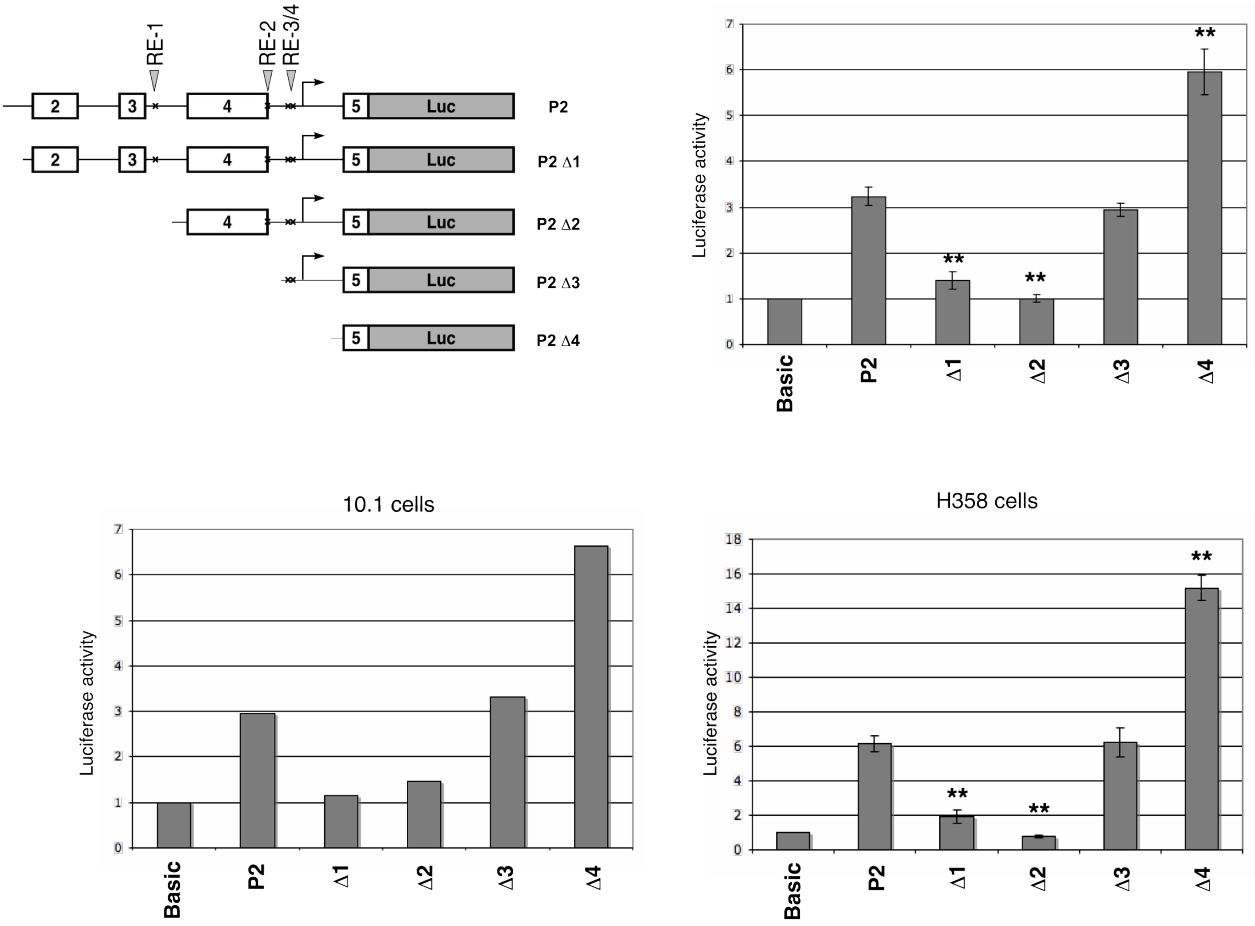
tcctcttcct acag**tactcc cctgccctca acaagATGtt ttgccaa**

B

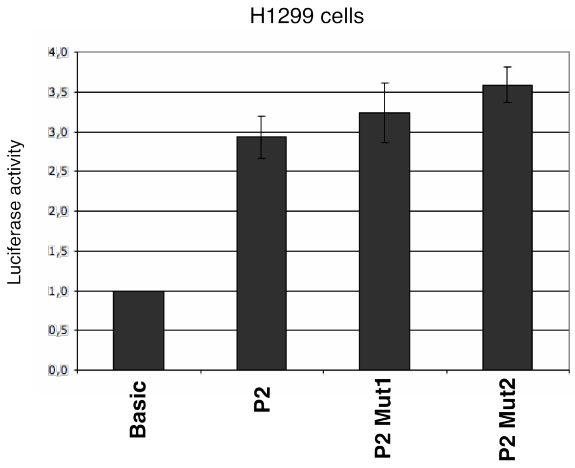
Name of p53RE	Location	Sequence	Percentage of similarity with the consensus (%)
p53 consensus		RRRCWWGYYY n RRRCWWGYYY	
RE-1	intron 3	GGACAAGGGT t GGGCTGGGAC	75
RE-2	exon 4	AGCCAAGTCT g TGACTTGCAC	85
RE-3	intron 4	ATCCACGTGT a TTCCTTGGCT	60
RE-4	intron 4	GCTCCTGACC AGGCTTGGTG	70

Supplementary Figure 4. Predictive location of p53 response elements within the internal P2 promoter. (**A**) Trans-regulators of the P2 promoter. The analysis of the P2 promoter (fragment from the end of intron 1 to the beginning of exon 5) was done using MatInspector software (Cartharius et al., 2005). Normal letter: intronic sequence; bold letter: exonic sequence; capital letter: initiation of translation (ATG133); arrow: initiation site of transcription of p53I4 mRNA; thin line box: putative TATA box; bold line boxes: putative p53 response element. (**B**) Divergence of p53RE according to the p53 consensus sequence. Four putative p53RE predicted by MatInspector software are known. RE-1/4: p53RE located within P2 promoter; R: purine; Y: pyrimidine; W: adenine or thymidine; underlined nucleotide: mismatch compared to the p53 consensus sequence.

A



B



Supplementary Figure 5. Regulatory regions of the internal P2 promoter of *TP53* gene. (**A**) Cis-regulators of the P2 promoter. Luciferase assays were performed using 4 plasmids carrying different deletions of the P2 promoter in p53-null murine cell line 10.1 and in two p53-null human lung cell lines H1299 and H358. In the three cell lines, regions upstream of exon 4 correspond to enhancer cis-regulatory elements and regions downstream of exon 4 correspond to silencer cis-regulatory elements. **: P-value < 0.01 compared to *TP53* (luciferase reporter system co-transfected with empty vector) calculated by Anova. (**B**) Basal transcriptional levels of mutated *TP53* plasmids. To determine the role of p53RE-2 in the transactivation of P2 promoter by p53 protein, two mutated luciferase system were developed. These two plasmids, P2 Mut1 and Mut2, present the same basal transcriptional activity than P2 luciferase system.

Supplementary tables

Supplementary Table 1. List of oligonucleotids used for this study.

Assay	Name	Sequence
real time PCR	Gapdh-F	5′-tctcatggttcacacccatgacgaacatg-3′
	Gapdh-R	5′-aagaagatgcggctgactgtcgagccacat-3′
	28S-F	5′-cgatccatcatccgcaatg-3′
	28S-R	5′-agccaagctcagcgcaac-3′
	p53I4-F (intron 4)	5′-ttcaactctgtctccttcct-3′
	p53I4-R (exons 5/6)	5′-gctgctcagatagcgatggtctggc-3′
EMSA	p53CON	5′-ggacatgcccgggcatgtcc-3′
ChIP	p21-F	5′-cattgttcccagcacttcctctc-3′
	p21-R	5′-agaaagccaatcagaccacag–3′
	Exon8-F	5′-ttccttactgcctcttgctt-3′
	Exon8-R	5′-aggcataactgcacccttgg-3′
	RE1-F	5′-aggcataactgcacccttgg-3′
	RE1-R	5′-ggggactgtagatgggtgaa-3′
	RE2-F	5′-tggaactttgggattcctct-3′
	RE2-R	5′-aacctaccagggcagctacg-3′
	RE3/4-F	5′-ggcttttatccatcccatca-3′
	RE3/4-R	5′-ccttggcctctttgagagtg-3′
siRNA	p53	5′-caaugguucacugaagacc-3′
	Δ133p53	5′-uguucacuugugcccugacuuucaa-3′
Plasmid production	P2(intron1)-F	5′-gcgagctcacattggaagaga-3′
	P2(exon5)-R	5′-gcaagcttggcaaaacatcttgt-3′
	P2Mut1-F	5′-gggacaccaaagtctgtgacttgcacggtcagttgccctgagg-3′
	P2Mut1-R	5′-cctcagggcaactgaccgtgcaagtcacagacaatttggtgtc-3′
	P2Mut2-F	5′-gggacaccaaagtctgtgaattccacggtcagttgccctgagg-3′
	P2Mut2-R	5′-cctcagggcaactgaccgtggaattcacagacaatttggtgtc-3′

ABSTRACT

The *TP53* gene is a highly polymorphic gene with 85 polymorphisms described. Some of these have been associated with an increase of cancer susceptibility, for example rs10425222 that can modulate certain p53 activities. However for others such as rs17878362, the most studied intronic polymorphism, the association with cancer risk is more controversial.

To investigate the influence of rs17878362 on cancer susceptibility, we analysed its role in sporadic and familial contexts. The results are paradoxical with an increase of sporadic cancer associated with the rs17878362 A2A2 genotype whereas the rs17878362 A2 allele is associated with a "protective" effect in the context of Li-Fraumeni patients carrying a *TP53* germline mutation on an A1 haplotype. These observations suggest that specific *TP53* haplotypes could modulate p53's tumour suppression capacities. A possible hypothesis to explain this could be that somatic mutations are carried on different haplotypes of *TP53* present at different allele frequencies in the population.

In addition, *TP53* is expressed as several protein isoforms, such as Δ40p53, which inhibits p53's suppressive activity. Δ40p53 can be produced from an alternative spliced transcript that retains intron 2. We have shown that G-quadruplexes, tri-dimensional structures formed in G-rich sequences, are formed in intron 3 and regulate the retention of intron 2 and the formation of the p53I2 transcript. We also observed that rs1642785 (located in intron 2) could regulate p53I2's stability. These results suggest that the *TP53* polymorphisms located in a 412 bp region located between exon 2 and exon 4 regulate the expression of p53 isoforms in a temporal sequence of events by modulating the pre-mRNA formation (rs17878362), mRNA stability (rs1642785) and protein functions (rs1042522).

p53 isoforms' expression is thus finely regulated by mechanisms involving *TP53* polymorphisms, which are also associated with altered cancer susceptibility.

KEY WORDS

Tumour suppressor protein, p53, isoforms, polymophism, rs17878362, alternative splicing, G-quadruplex structure, Li-Fraumeni syndrome, meta-analysis

RESUME

Le gène *TP53* est extrêmement polymorphique avec 85 polymorphismes décrits. Certains de ces polymorphismes sont associés à une augmentation du risque de cancer, par exemple rs10425222 peut moduler les fonctions de p53. Cependant, pour d'autres, comme le rs17878362 qui est le polymorphisme intronique le plus étudié, leur association avec une augmentation du riques au cancer est controversée.

Pour analyser l'association entre le polymorphisme rs17878362 et la susceptibilité au cancer, nous avons analysé son rôle dans des contextes de cancers sporadiques et familiaux. Les résultats obtenus pour le polymorphisme rs17878362 sont paradoxaux avec une augmentation des cancers sporadiques associée avec le génotype A2A2 alors que l'allèle A2 est associé avec un effet « protectif » chez les patients atteints du syndrome de Li-Fraumeni porteurs d'une mutation germinale de *TP53* situé sur l'haplotype A1. Ces observations suggèrent que des haplotypes spécifiques de *TP53* pourraient moduler les capacités suppressives de p53. Une hypothèse possible est que les différents haplotypes de *TP53* présenteraienrt des mutations somatiques à des fréquences différentes dans la population.

De plus, le gène *TP53* exprime différentes isoformes, comme le Δ40p53, inhibant l'activité suppressive de p53. Le Δ40p53 peut être produite par le maintien de l'intron 2 par épissage alternatif. Nous avons montré que les G-quadruplexes, des structures tridimensionnelles formées dans des régions riches en G, sont formés dans l'intron 3 et régulent la rétention de l'intron 2 et la formation du transcrit p53I2. Nous avons aussi observé que le polymorphisme rs1652785 (localisé dans l'intron 2) semble réguler la stabilité du p53I2. Ces résultats suggèrent que les polymorphismes de *TP53* localisés dans une région de 412 pb située entre l'exon 2 et l'exon 4 régulent l'expression des isoformes de p53 dans une séquence temporelle d'évènements en modulant la formation des pré-ARNm (rs17878362), la stabilité des ARNm (rs1642785) et les fonctions protéiques (rs10425222).

L'expression des isoformes de p53 est donc finement régulée par des mécanismes impliquant les polymorphismes de *TP53* qui sont aussi associés avec une altération dans la susceptibilité au cancer.

MOTS-CLES

Protéine suppresseur de tumeur, p53, isoformes, polymorphismes, rs17878362, épissage alternatif, structure de type G-quadruplexe, syndrome de Li-Fraumeni, méta-analyse