TP53 Mutations in Human Cancers: Origins, Consequences, and Clinical Use

Magali Olivier¹, Monica Hollstein², and Pierre Hainaut¹

¹Group of Molecular Carcinogenesis, International Agency for Research on Cancer, 150 Cours Albert Thomas, 69372 Lyon Cedex 08 France
²LIGHT Laboratories, University of Leeds, United Kingdom
Correspondence: hainaut@iarc.fr

Somatic mutations in the TP53 gene are one of the most frequent alterations in human cancers, and germline mutations are the underlying cause of Li-Fraumeni syndrome, which predisposes to a wide spectrum of early-onset cancers. Most mutations are single-base substitutions distributed throughout the coding sequence. Their diverse types and positions may inform on the nature of mutagenic mechanisms involved in cancer etiology. TP53 mutations are also potential prognostic and predictive markers, as well as targets for pharmacological intervention. All mutations found in human cancers are compiled in the IARC TP53 Database (http:\/\/www-p53.iarc.fr/). A human TP53 knockin mouse model (Hupki mouse) provides an experimental model to study mutagenesis in the context of a human TP53 sequence. Here, we summarize current knowledge on TP53 gene variations observed in human cancers and populations, and current clinical applications derived from this knowledge.

Genetic variations in the tumor suppressor gene TP53 (OMIM #191117) contribute to human cancers in different ways. First, somatic mutations are frequent in most cancers (Hollstein et al. 1991). The antiproliferative role of p53 protein in response to various stresses and during physiological processes such as senescence makes it a primary target for inactivation in cancer (Levine 1997). The main modes of TP53 inactivation are single-base substitution and loss of alleles, with inactivation by viral or cellular proteins playing a major role in specific cancers (Tommasino et al. 2003). Second, inheritance of a TP53 mutation causes predisposition to early-onset cancers including breast carcinomas, sarcomas, brain tumors, and adrenal cortical carcinomas, defining the Li-Fraumeni (LFS) and Li-Fraumeni-like (LFL) syndromes (Li et al. 1988; Olivier et al. 2003). Third, TP53 is highly polymorphic in coding and noncoding regions and some of these polymorphisms have been shown to increase cancer susceptibility and to modify cancer phenotypes in TP53 mutation carriers (Whibley et al. 2009).

Whereas tumor suppressors are commonly inactivated by frameshift or nonsense mutations, most TP53 mutations are missense and cause single amino-acid changes at many different positions. Mutations are thus diverse in their type, sequence context, position, and structural...
impact, making it possible to identify mutation patterns in relation with cancer type and etiology. The occurrence of special mutation patterns may inform on the nature of the mutagens that have caused them, making TP53 an interesting gene to analyze in the realm of molecular epidemiology.

Data on mutation prevalence in human cancer can be conveniently accessed through the IARC TP53 database (http://www-p53.iarc.fr/), a resource that compiles all TP53 gene variations reported in human cancers with annotations on tumor phenotype, patient characteristics, and structural and functional impact of mutations (Petitjean et al. 2007b). Recently, it has become possible to confront these observations with experimental data generated in a novel mouse model, the HupKi mouse, that contains a human TP53 sequence at the mouse TP53 locus and recapitulates the effects of environmental mutagens in a human sequence context (Luo et al. 2001). In this article, we review the current knowledge on the origin, causes, and consequences of TP53 variations and mutations in cancer and we discuss their significance as biomarkers in epidemiology and in the clinics.

**TP53 VARIATION LANDSCAPES IN HUMAN CANCERS AND POPULATIONS**

**Somatic Mutations**

Somatic TP53 mutations occur in almost every type of cancer at rates from 38%–50% in ovarian, esophageal, colorectal, head and neck, larynx, and lung cancers to about 5% in primary leukemia, sarcoma, testicular cancer, malignant melanoma, and cervical cancer (Fig. 1). Mutations are more frequent in advanced stage or in cancer subtypes with aggressive behavior (such as triple negative or HER2-amplified breast cancers) (Wang et al. 2004a; Wang et al. 2004b; Langerod et al. 2007). In cancers with low mutation rates, p53 is often inactivated by alternative mechanisms. This is

![Figure 1. TP53 mutations prevalence in sporadic cancers. The proportion of tumors with somatic TP53 mutations is indicated. Data from IARC TP53 Database (R13, November 2008)(Petitjean et al. 2007b).](http://cshperspectives.cshlp.org/)
the case for cervical cancer in which p53 is targeted for degradation by HPV E6 (Tommasino et al. 2003) or for sarcoma that overexpress amplified HDM2.

Based on studies that examined the whole coding sequence, 86% of mutations cluster between codons 125 and 300, corresponding mainly to the DNA binding domain (Fig. 2). Most mutations in this region are missense (87.9%). In contrast, outside this region, missense mutations represent only about 40%, the majority of mutations being nonsense or frameshift. Among single-base substitutions, about 25% are C:G>T:A substitutions at CpG sites. CpG dinucleotides mutate at a rate 10 times higher than other nucleotides, generating transitions (Jones et al. 1992). About 3%–5% of cytosines in the human genome are methylated at position 5’ by a postreplicative mechanism that is restricted to CpG dinucleotides and is catalyzed by DNA methyltransferases. The 5’-methylcytosine (5mC) is less stable than cytosine and undergoes spontaneous deamination into thymine at a rate five times higher than the unmethylated base. This process is enhanced by oxygen and nitrogen radicals, leading to a higher load of CpG transitions in cancers arising from inflammatory precursors such as Barrett’s mucosa or ulcerative colitis (Schmutte et al. 1996; Ambbs et al. 1999; Vannetti et al. 2008). Among the 22 CpG of the DNA-binding domain (DBD), three hotspot codons (175, 248, and 273) represent 60% of CpG mutations and another five residues

Figure 2. Type of somatic TP53 mutations in human cancers. (A) Pie charts showing the proportion of the different types of TP53 somatic mutations found in all human cancers. (B) Histogram displaying the position of somatic point mutations in the coding sequence of the TP53 gene. Data from the IARC TP53 Database (R13, November 2008)(Petitjean et al. 2007b).
(196, 213, 245, 282, and 306) account for 26% of these mutations. The lack of mutations at other CpG sites reflects the fact that substitution at these residues does not generate a dysfunctional protein. Although the same CpG hotspot mutations occur in many cancer types, other types of mutations tend to show differences from one cancer to the other. Some of these differences have been linked to the effect of specific mutagens. Geographic differences have also been reported in relation with environmental exposures. These aspects have been extensively discussed in other reviews (Hainaut et al. 2000; Olivier et al. 2004) and some examples are briefly discussed later in this article.

**Germline Mutations: Li-Fraumeni Syndrome**

T53 germline mutations are the underlying cause of LFS, a familial clustering of early onset tumors including sarcomas, breast cancers, brain tumors, and adrenal cortical carcinomas (Li et al. 1988; Malkin et al. 1990). Over the past 20 years, TP53 germline mutations have been detected in about 500 families or individuals with complete or partial LFS features (the latter defined as Li-Fraumeni-like, LFL) (Olivier et al. 2003). LFS/LFL has been generally considered as a rare syndrome (Eeles 1995). However, screening for TP53 germline mutation in patients with early onset breast cancer and unselected for familial history has shown TP53 mutations in 2%–3% of the cases (Lalloo et al. 2006), whereas screening of 525 patients with any kind of cancer family history has identified 91 (17.3%) TP53 mutations (Gonzalez et al. 2009). Based on these results, TP53 mutation may contribute to up to 17% of all familial cancer cases. Studies in southern Brazil have identified many families with a founder mutation (R337H) (Fig. 3). Thus, TP53 germline mutations may be more common than previously recognized, occurring in about 1 in 5,000 to 1 in 20,000 births (Lalloo et al. 2006) (Gonzalez et al. 2009).

Breast cancer and soft tissue and bone sarcoma account for over 50% of tumors in TP53 mutation carriers, followed by adrenocortical carcinomas and brain tumors (Fig. 4). Other cancers include hematological malignancies, gastric, colorectal, and ovarian cancers, occurring at earlier ages than in the general population (Olivier et al. 2003). Rarer cancers associated with TP53 germline mutation are choroid plexus carcinoma or papilloma before the age of 15, Wilms’ tumor, and malignant phyllodes tumors (Birch et al. 2001; Gonzalez et al. 2009).

The distribution of germline mutations is similar to somatic mutations, with mostly missense mutations (77%) located at the same hotspots. The proportion of CpG mutations (54% vs. 25% in somatic mutations) may reflect the spontaneous nature of germline mutations. Genotype–phenotype correlations suggest that the most significant defect is loss of function because large deletions encompassing the whole TP53 gene have been found in LFS families with aggressive features (Bougeard et al. 2003).

**Figure 3.** Geographic distribution of germline TP53 mutations. Number of TP53 germline mutation carrier families in each world region. Data from the IARC TP53 Database (R13, November 2008) (Petitjean et al. 2007b).
TP53 Polymorphisms

Over 80 TP53 polymorphisms have been identified and validated in human populations (IARC TP53 Database, R13). The majority (90%) are located in introns, outside splice sites, or in noncoding exons. Few of them have been tested in functional assays or studied for effects on cancer risk. Among 18 exonic SNPs (Table 1), five are silent and seven are located after the stop codon in exon 11. Four exonic polymorphisms alter the protein sequence and have only subtle effects on transactivation capacity as measured in yeast-based assays (Kato et al. 2003).

V217M is the only nonsilent polymorphism in the DBD. In yeast functional assays, the rare variant shows increased transactivation of some p53 response elements (CDKN1A, BAX, and PMAIP1) and may thus be protective against cancer. G360A, located next to the tetramerization domain (TET) showed slightly reduced activity that may result in an increased cancer risk. However, the impact of these SNP has not been investigated in clinical or epidemiological studies.

P47S has been reported in African populations only, at a frequency of approximately 5% (Felley-Bosco et al. 1993). This polymorphism may affect phosphorylation at S46, which enhances p53-mediated apoptosis (Pistritto et al. 2007; Oda et al. 2000). Phosphorylation of S46 by p38 and homeodomain-interacting protein kinase 2 (HIPK2) requires a proline adjacent to S46. Presence of a serine may thus alter kinase recognition. However, functional data in different experimental systems have proved inconsistent. In mice, blocking S46 phosphorylation has only modest phenotypical consequences (Feng et al. 2006; Toledo et al. 2006). It is possible that this SNP may impact on p53 function only under specific stress conditions. Its
impact on cancer susceptibility remains to be elucidated.

The P72R polymorphism is the most extensively studied both in experimental and population studies. Sharp ethnic differences have been observed, the P72 allele showing a north–south gradient from 0.17 in Swedish Saamis to 0.63 in Africans (Yorubas). A recent report has shown an association of this SNP with latitude within Asian populations (Shi et al. 2009) and has proposed that allele frequency may be an adaptation to differences in winter temperature. R72 is more active than P72 in transactivating the expression of leukaemia-inhibitory factor (Kang et al. 2009), which is critical for blastocyst implantation, leading to the conjecture that R72 could thus provide a selective advantage by reducing the risk of implantation failure in colder climates.

Codon 72 is located within a proline-rich region and may affect a putative SH3-binding domain. The current consensus from a large number of studies is that R72 is more effective in inducing apoptosis than P72. However, the relevance of these observations for cancer susceptibility is not understood. An early suggestion that individuals homozygous for R72 may be at greatly higher risk for HPV-related cervical cancers (Storey et al. 1998) has not been substantiated by extensive population-based studies (Klug et al. 2001). Many studies have reported associations between the R72P SNP and risk of different cancers, with inconsistent results. Meta-analyses for breast (Schmidt et al. 2007), lung (Matakidou et al. 2003), and other cancers do not support a significant role for this polymorphism in susceptibility.

### Table 1. Selected polymorphisms in TP53 gene

<table>
<thead>
<tr>
<th>Exon</th>
<th>g_Description</th>
<th>c_Description</th>
<th>AvgHet</th>
<th>SNPlink</th>
<th>Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>4-exon</td>
<td>g.11333C&gt;A</td>
<td>c.102C&gt;A</td>
<td>-</td>
<td>11575998 silent</td>
<td></td>
</tr>
<tr>
<td>4-exon</td>
<td>g.11339G&gt;A</td>
<td>c.108G&gt;A</td>
<td>0.012738</td>
<td>1800370 silent</td>
<td></td>
</tr>
<tr>
<td>4-exon</td>
<td>g.11370C&gt;T</td>
<td>c.139C&gt;T (P47S)</td>
<td>0.029529</td>
<td>1800371 missense</td>
<td></td>
</tr>
<tr>
<td>4-exon</td>
<td>g.11446C&gt;G</td>
<td>c.215C&gt;G (R72P)</td>
<td>0.492248</td>
<td>1042522 missense</td>
<td></td>
</tr>
<tr>
<td>6-exon</td>
<td>g.12708A&gt;G</td>
<td>c.639A&gt;G</td>
<td>0.023526</td>
<td>16956880 silent</td>
<td></td>
</tr>
<tr>
<td>6-exon</td>
<td>g.12718G&gt;A</td>
<td>c.649A&gt;G (V217M)</td>
<td>0.024239</td>
<td>35659787 noncoding</td>
<td></td>
</tr>
<tr>
<td>10-exon</td>
<td>g.16970G&gt;A</td>
<td>c.1079G&gt;C (G360A)</td>
<td>0.019761</td>
<td>35919705 silent</td>
<td></td>
</tr>
<tr>
<td>11-exon</td>
<td>g.18096C&gt;T</td>
<td>c.1182+105C&gt;T</td>
<td>0.019761</td>
<td>35919705 silent</td>
<td></td>
</tr>
<tr>
<td>11-exon</td>
<td>g.18196G&gt;A</td>
<td>c.1182+205G&gt;A</td>
<td>0.019569</td>
<td>34486624 noncoding</td>
<td></td>
</tr>
<tr>
<td>11-exon</td>
<td>g.18305G&gt;A</td>
<td>c.1182+314G&gt;A</td>
<td>0.009892</td>
<td>35659787 noncoding</td>
<td></td>
</tr>
<tr>
<td>11-exon</td>
<td>g.18319G&gt;A</td>
<td>c.1182+328G&gt;A</td>
<td>0.009892</td>
<td>35659787 noncoding</td>
<td></td>
</tr>
<tr>
<td>11-exon</td>
<td>g.18476G&gt;A</td>
<td>c.1182+485G&gt;A</td>
<td>0.091723</td>
<td>4968187 noncoding</td>
<td></td>
</tr>
<tr>
<td>11-exon</td>
<td>g.18560_18561del2</td>
<td>c.96+41_96+56del16</td>
<td>0.436246</td>
<td>17883362 intronic</td>
<td></td>
</tr>
<tr>
<td>2-intron</td>
<td>g.11117C&gt;G</td>
<td>c.74+38C&gt;G</td>
<td>0.468059</td>
<td>1642785 Intronic</td>
<td></td>
</tr>
<tr>
<td>9-intron</td>
<td>g.16143G&gt;A</td>
<td>c.994-742G&gt;A</td>
<td>0.401235</td>
<td>1641549 Intronic</td>
<td></td>
</tr>
<tr>
<td>9-intron</td>
<td>g.14496C&gt;T</td>
<td>c.993+431C&gt;T</td>
<td>0.3432</td>
<td>1642791 Intronic</td>
<td></td>
</tr>
<tr>
<td>7-intron</td>
<td>g.13511T&gt;G</td>
<td>c.782+92T&gt;G</td>
<td>0.310766</td>
<td>12951053 Intronic</td>
<td></td>
</tr>
<tr>
<td>9-intron</td>
<td>g.15185T&gt;C</td>
<td>c.993+1120T&gt;C</td>
<td>0.256705</td>
<td>12949655 Intronic</td>
<td></td>
</tr>
<tr>
<td>7-intron</td>
<td>g.13491C&gt;T</td>
<td>c.782+72C&gt;T</td>
<td>0.246362</td>
<td>12947788 Intronic</td>
<td></td>
</tr>
<tr>
<td>4-intron</td>
<td>g.12273G&gt;A</td>
<td>c.376-91G&gt;A</td>
<td>0.212093</td>
<td>9895829 Intronic</td>
<td></td>
</tr>
<tr>
<td>6-intron</td>
<td>g.12803A&gt;G</td>
<td>c.672+62A&gt;G</td>
<td>0.210458</td>
<td>1625895 Intronic</td>
<td></td>
</tr>
<tr>
<td>4-intron</td>
<td>g.12239T&gt;C</td>
<td>c.376-125C&gt;T</td>
<td>0.157451</td>
<td>9895829 Intronic</td>
<td></td>
</tr>
<tr>
<td>3-intron</td>
<td>g.11299C&gt;A</td>
<td>c.97-29C&gt;A</td>
<td>0.15648</td>
<td>17883323 intronic</td>
<td></td>
</tr>
</tbody>
</table>

The P72R SNP has been studied for its impact on prognosis or response to treatment. Many of these studies must be interpreted cautiously as they lack power and rigorous design. There is, however, evidence of a modifier effect on TP53 mutations, with a bias toward expression of mutant R72 allele in tumors of heterozygous patients with LOH affecting wild-type P72 (Langerod et al. 2002; Bergamaschi et al. 2003; Bonafe et al. 2003; Vikhanskaya et al. 2005; Nelson et al. 2005; Zawlik et al. 2009). Furthermore, experimental studies showed that R72 mutants were more potent in cooperating with EJ-Ras in transforming primary cells and in neutralizing p73 activity (Marin et al. 2000; Bergamaschi et al. 2003). A modifier effect of the R72P SNP has also been reported in germline TP53 mutation carriers, in whom R72 was associated with earlier age at first diagnosis of cancer (Bougeard et al. 2006).

P72R is in linkage disequilibrium with a polymorphism in intron 3 (PIN3), consisting of a duplication of a 16 bp, GC rich sequence (allele frequency: 0.2 in Caucasians). PIN3 has been associated with increased risk of several cancers (Wang-Gohrke et al. 1998; Wang-Gohrke et al. 1999; Gemignani et al. 2004). The duplicated PIN3 allele is associated with reduced expression of TP53 mRNA in lymphoblastoid cell lines (Gemignani et al. 2004). Recently, TP53 germline mutation carriers with duplicated PIN3 were shown to develop cancer, on average, 20 years later than carriers with nonduplicated alleles (Marcel et al. 2009).

Given the complex polymorphic structure of TP53, haplotypes may provide more relevant information than individual polymorphisms. In addition, polymorphisms in gene regulating p53 have a critical impact, as for example MDM2 SNP309 polymorphism (See Grochola et al. 2010).

**STRUCTURAL AND FUNCTIONAL IMPACT OF MUTATIONS**

**Impact on Protein Structure**

The p53 DBD is made of an immunoglobulin-like β-sandwich of two antiparallel β-sheets, providing a scaffold for a flexible DNA-binding surface (Cho et al. 1994). This surface is formed by two large loops (L2 and L3) stabilized by a zinc atom and a loop–sheet–helix motif (loop L1). Zinc binding (coordinated by H179, C176, C238, and C242) is critical for correct folding, and requires reduction of thiol groups on cysteines. Hotspot mutations are at residues involved either in making contacts with DNA or in supporting the structure of the DNA-binding surface. Mutant proteins have thus been classified as “contact” (e.g., R248 and R273) or “structural” (e.g., R175, G245, R249, and R282) (Joerger et al. 2007). The structure of several common mutants has been elucidated by NMR spectroscopy or X-ray crystallography (Joerger et al. 2005; Joerger et al. 2006). DNA-contact mutants retain the overall architecture of the DBD with loss of a critical DNA contact. They may actively prevent DNA binding if a large hydrophobic side chain is introduced (e.g., S241F, R248W, and C277F). Zinc-binding mutants affect the zinc coordination sphere (e.g., C176F, H179R, and C242F). This category includes R175H, the most frequent hotspot mutant, because introduction of histidine residue causes distortions that directly interfere with zinc binding. Substitutions introducing smaller residues at this position have been shown to be less destabilizing with partial retention of function (Bullock et al. 2000). Structural mutants cause distortions that create internal cavities or surface crevices in the protein scaffold, inducing conformational changes in the DNA binding surface. Overall, these studies highlight the structural heterogeneity of mutant proteins, with possible consequences on their biochemical and biological properties.

**Impact on Transcriptional Activities**

p53 regulates transcription through specific binding to response elements in the promoters or introns of target genes (Riley et al. 2008). p53REs are variations of a consensus constituted by two decamers (PuPuPuC(A/T)(T/A) GPyPyPy) separated by a spacer of variable length (El-Deiry et al. 1992). These sequence variations affect p53 binding affinity and may contribute to shape the repertoire of genes
activated by p53 in response to specific signals (Inga et al. 2002). Recent genome-wide analyses of p53 binding sites suggest that hundreds of genes may be up- or down-regulated by p53 (Smeenk et al. 2008). Kato et al. (2003) have used a yeast-based functional assay to analyze the transactivation of 2314 different missense mutants on eight p53REs. All hotspot mutants displayed loss of function on all tested p53REs (Fig. 5). However, several rare mutants retained partial activity on all REs or even full activity on some REs. A small category of mutants even showed increased activity compared with wild-type p53 (so-called ‘supertrans mutants,’ often occurring in loop L1, which is not a frequent target for cancer-related mutations). Using a different type of yeast based assay, Inga et al. (2002) further showed that differential transactivation by wild-type and mutant proteins depends on both protein levels and target sequence.

**Dominant–Negative and Gain-of-Function Effects**

Mutant p53 proteins often accumulate in the nucleus of in situ and metastatic cancer cells, suggesting an oncogenic effect in addition to loss of wild-type suppressor function. Experimentally, several hotspot mutants have been

![Effect distribution of 20999 mutations](image)

**Figure 5.** Functional impact of somatic TP53 mutations in human cancers. (A) Scatter plots showing the frequency (x-axis) of single amino-acid substitutions in relation to their effect (left panel) or functional impacts on transactivation (right panel), and expected substitution rates (y-axis in log). Each point represents a single amino-acid substitution that is shaped and colored according to the mutation effect or functional impact. (B) Pie charts displaying the proportion of all somatic single amino-acid substitutions according to their effect (left panel) or functional impacts on transactivation (right panel). Data from the IARC TP53 Database (R13, November 2008) (Petitjean et al. 2007b).
shown to cooperate with oncogenes for cellular transformation (Hinds et al. 1990), setting the concept of gain-of-function (GOF). Alternatively, mutant p53 proteins may exert dominant–negative effects (DNE) over wild-type p53. Indeed, transcriptional activity relies on the formation of p53 tetramers (dimers of dimers) and incorporation of mutant p53 in these structures may preclude their activity.

DNE has been experimentally studied in yeast and human cell assays, generating data on the DNE effects for over 200 mutants (Brachmann et al. 1996). Analysis of the IARC TP53 database showed that mutants with DNE toward p53-RE from WAF-1 and RGC promoters were over-represented in cancer (Petitjean et al. 2007b), suggesting that DNE plays a role in selecting for mutations during carcinogenesis.

Only a few mutants have been studied for GOF activities and there is no consensus on the molecular definition of such activities (Petitjean et al. 2007b). Knockin mice with R175H and R273H mutants (Lang et al. 2004; Olive et al. 2004) showed increased metastases and different cancer distribution compared with p53 knockout mice, providing in vivo evidence of GOF. Various molecular mechanisms have been proposed as the basis of GOF, such as interference with the p53-family proteins p63 and p73, interference with the ATM DNA repair pathway (Song et al. 2007), activation of genes normally unaffected or repressed by wild-type p53, interference with other transcription factors, and resistance to drugs (see Oren and Rotter 2010). However, no systematic study has been performed to distinguish mutants with GOF properties. Thus, there is so far no definite evidence that GOF contribute to mutant p53 selection in human cancers.

It should be noted that, when compared with expected numbers, if all mutations were equal, missense mutations in p53 DBD are only slightly overrepresented in cancer (87.1% observed vs. 72.9% expected), whereas nonsense mutations are overrepresented by a factor of two (8% vs. 3.6%) and silent mutations underrepresented by a factor of four (4.8% vs. 22.9%) (Table 2). These observations further support that loss of function is the critical factor for the selection of mutations in cancer. This is particularly striking when considering mutations at hypermutable CpG sites. Of 34 possible missense mutations at CpG sites in the DBD, only seven are frequently observed in cancers. These seven mutants show complete loss of transactivation activities, whereas the 27 rare mutants retain significant transactivation activity on one or more p53RE.

### MUTAGEN-INDUCED MUTATION SPECTRA

#### Human Data

In several types of cancers, mutation patterns bear the hallmarks of chemical damage induced by particular mutagens, leading to the concept of "mutagen fingerprints" (Table 3). These fingerprints are defined by the relative frequencies, types of base change, strand orientation, and location of base substitutions. The factors shaping mutation patterns have been studied in detail in various organisms and experimental models. They include (1) the nature of the environmental agent that chemically modifies a base, (2) the DNA sequence itself (base context), (3) epigenetic modifications at the target sequence (e.g., base methylation), (4) DNA repair and replication processes, (5) transcription activity at the locus, and (6) biological selection (as discussed previously).

Although in many cancers transitions at CpG sites predominate, a complex picture emerges when concentrating on rarer types of mutations. The patterns of these mutations can vary dramatically between two types of cancers, or two patient populations with a given cancer type. In some instances, the mutations reveal

<table>
<thead>
<tr>
<th>Mutation type</th>
<th>Expected mutations</th>
<th>Observed mutations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Missense</td>
<td>1150 (73.4%)</td>
<td>17,191 (87.9%)</td>
</tr>
<tr>
<td>Nonsense</td>
<td>58 (3.7%)</td>
<td>1435 (7.3%)</td>
</tr>
<tr>
<td>Silent</td>
<td>359 (22.9%)</td>
<td>932 (4.8%)</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>1567</strong></td>
<td><strong>19,558</strong></td>
</tr>
</tbody>
</table>

*Somatic mutations reported in the IARC TP53 Database (R13, November 2008).
characteristics expected from the effects of environmental mutagens on DNA. There are four well-documented examples of associations between an etiologic agent and a tumor mutation fingerprint. Three of these have been reviewed in detail: exposure to sunlight and the presence of tandem CC to TT transitions in nonmelanoma skin cancers (Giglia-Mari et al. 2003), tobacco smoking and PAH-induced G to T transversions at specific G:C base pairs oriented with the (premutated) guanine on the nontranscribed strand (codons 157, 158, 245, 248, and 273) (Pfeifer et al. 2002; Pfeifer et al. 2003), and exposure to dietary aflatoxin B1 and G to T transversions at the third base of codon 249 (AGG to AGT) in tumors from HBV carriers (Hussain et al. 2007). The fourth and most recent example links TP53 A to T transversions with crops contaminated with seeds of Aristolochia sp, and is worth mentioning even though the estimated number of cancer patients exposed worldwide to aristolochic acids (AA), the carcinogenic agent, is expected to be small in comparison to the three exposures just cited (Arlt et al. 2007). Pre-mutagenic AA-specific DNA adducts have been detected in individuals suffering from Balkan endemic nephropathy (BEN), a form of nephropathy associated with high risk of urothelial cancer that is historically common in some parts of southeastern Europe. TP53 mutations in urothelial tumors of these patients are predominantly (74%) A to T transversions (Grollman et al. 2007), a form of mutation that is rare in urothelial tumors from non-BEN patients (4%) as well as in human tumors overall (5%). Moreover, BEN tumor mutations show a strong strand bias, with the premutated adenine base residing on the nontranscribed strand in 93% of the instances reported so far. This bias suggests preferential (transcription-coupled) repair of bulky lesions on the DNA strand that serves as template for mRNA synthesis. Reminiscent of the first discoveries of AFB1-associated codon 249 G to T transversions in two small sets of liver tumors (Hsu et al. 1991; Bressac et al. 1991), the AA-associated mutations suggest a mutation fingerprint even though only 19 mutations in this cohort of BEN patients have been reported thus far.

Recent large-scale sequencing studies are having a profound impact on our understanding of mutation load in human cancers (Ding et al. 2008; Stratton et al. 2009). In keeping with the smokers’ lung tumor TP53 mutation data, a screen of 623 genes in 188 lung adenocarcinomas confirmed the higher mutation load and higher fraction of strand-biased G to T transversions in lung tumor of smokers when compared with nonsmokers (Ding et al. 2008). Genome-wide studies highlighted an unexpected and statistically significant elevation in

Table 3. Human carcinogen TP53 signature mutations

<table>
<thead>
<tr>
<th>Carcinogen/(exposure)</th>
<th>Target organ</th>
<th>Metabolic activation</th>
<th>Base substitution</th>
<th>Distinctive</th>
</tr>
</thead>
<tbody>
<tr>
<td>PAH (B[a]P) (smoking)</td>
<td>Lung</td>
<td>Epoxidation</td>
<td>G to T with strand bias</td>
<td>157, 158, 273</td>
</tr>
<tr>
<td>AA (dietary contam.)</td>
<td>Urothelium</td>
<td>Nitroreduction</td>
<td>A to T with strand bias</td>
<td>131, 209, (280)</td>
</tr>
<tr>
<td>Aflatoxin B1 (dietary contam.)</td>
<td>Liver</td>
<td>Epoxidation</td>
<td>G to T</td>
<td>249 (3rd base)</td>
</tr>
<tr>
<td>UV radiation (sunlight)</td>
<td>Skin</td>
<td>n.a.</td>
<td>CC to TT</td>
<td>248, 278</td>
</tr>
<tr>
<td>3-NBA (diesel)</td>
<td>Lung</td>
<td>Nitroreduction</td>
<td>n.i.</td>
<td>G to T</td>
</tr>
<tr>
<td>Vinyl chloride</td>
<td>Liver</td>
<td>Epoxidation</td>
<td>A:T to T:A with strand bias</td>
<td>n.i.</td>
</tr>
</tbody>
</table>
G:C to C:G transversions in breast cancers (28% of all substitutions) compared with colorectal, pancreatic, and brain cancers (7–10%) (Stephens et al. 2005; Sjoblom et al. 2006; Greenman et al. 2007; Jones et al. 2008). Oddly, this is not seen in breast cancer TP53 mutations (9% of which are G:C to C:G transversions) (see Pfeifer and Besaratinia 2009 for discussion of G:C to C:G mutations). Next generation DNA sequencing technologies will offer new opportunities for studying factors that govern mutation load in normal and pre-neoplastic cell populations, including stem cells, which may be especially pertinent to elucidating origins of critical cancer-causing mutations.

Experimental Systems

Experimental mutagenesis studies in lower organisms and mammalian cells have provided evidence linking the four human TP53 tumor mutation signatures discussed above to the suspected carcinogens (Hainaut et al. 2000; vom Brocke J. et al. 2006). However, this approach is difficult to extrapolate to more complex situations involving multiple suspected carcinogens and where etiologic clues are needed. Classical experimental systems to generate mutation spectra are based on mutations in reporter genes that are not cancer related, but that allow efficient selection and recovery of mutants (e.g., the microbial lacI and LacZ genes, or the mammalian TK and HPRT genes). However, a reporter gene will differ from TP53 or any other genes of interest in base sequence context, transcriptional level (thus transcription-coupled repair), and biological selection of mutants, three key factors that shape human TP53 tumor mutation patterns. Until recently, the human TP53 gene was considered a “nonselectable” gene in an experimental setting, which would preclude its use as a reporter of experimentally induced mutations. However, innovative approaches have been developed to map experimental DNA damage along the human TP53 sequence as a first approximation of the sites vulnerable to accumulation of mutations. This strategy was used successfully to show the correlation between human tumor TP53 mutations and the positions of the main DNA photoproducts induced by UV light or those of adducts formed by metabolites of PAH (Denissenko et al. 1996). Another powerful approach was devised to score mutations in human TP53 sequences cloned into a vector by assessing loss of p53 transcriptional transactivation function when the plasmid is then introduced into specially designed yeast strains (Frebourg et al. 1992; Ishioka et al. 1993).

A direct method to induce and select TP53 mutations in mammalian cells would be to conduct a rodent cancer assay and sequence the murine TP53 gene in the carcinogen-induced tumors (Zielinski et al. 2002). Although this approach has been used successfully to examine UV-induced mutations in murine skin tumors, it has limited applicability considering the unexplained paucity of TP53 mutations in most mouse tumor types other than skin. Furthermore, murine and human TP53 genes, although highly homologous, differ in DNA sequence at critical hotspot codons, posttranslational modification sites, and intragenic suppressor sites (Hergenhahn et al. 2004). An in vivo carcinogenesis assay in mouse harboring human TP53 gene sequences would bypass this latter difficulty. However, its applicability may still be limited to studies of skin lesions induced by topical application of suspected agents. The HUF Assay (Hupki fibroblast assay) is a novel mutagenesis test in primary cells from “humanized” mice that was designed to circumvent the difficulties discussed above (Liu et al. 2004). HupKi (Human p53-Knockin) mice harbor human TP53 sequences encoding both the polyproline and the DNA binding domain, where most human tumor mutations arise (Luo et al. 2001; Reinbold et al. 2008), replacing the homologous murine gene segments within the mouse TP53 locus. Hupki p53 is functional and regulated normally in mice. The “domain-swap” strategy is necessary, rather than replacement of the entire murine p53 by its human counterpart, because in a complete swap the TP53 gene loses wild-type function in mice, due to aberrant human p53 N-terminal interaction with the murine major negative regulator Mdm2 (Dudgeon et al. 2006).
The HUF assay takes advantage of the fact that p53 inactivation is the critical rate-limiting event in escape of replicative senescence and spontaneous immortalization of mouse embryo fibroblasts (MEF) in culture (Hahn et al. 2002). Fibroblasts are explanted from Hupki embryos, exposed to carcinogens, and cultured until immortalization has occurred. TP53 gene is then sequenced in MEF cell lines in search of mutations induced by carcinogens in a human sequence context. Several human carcinogens have been tested in the HUF assay and shown to induce TP53 mutations that are consistent with signatures identified in human tumors (Pfeifer and Besaratinia 2009; vom Brocke et al. 2006; Reinbold et al. 2008). Of note, aristolochic acid-induced A to T mutations in HUF cells correspond remarkably well with the TP53 tumor mutations in cancer patients with BEN (Nedelko et al. 2009). The HupKi mouse concept is now being extended to a whole family of mice carrying humanized TP53 sequence variants (recapitulating human TP53 polymorphisms, hotspot TP53 mutations, or mutations abrogating critical post-translational modification sites). By crossing Hupki mice with other genetically modified mouse strains (e.g., mice with humanized metabolic or DNA repair enzymes, or knockout strains with deficiencies in specific defense systems), many outstanding questions on the mechanisms of mutagenesis pertinent to human tumorigenesis could be addressed.

EPIGENOLOGICAL, DIAGNOSTIC, AND THERAPEUTIC UTILITY OF TP53 MUTATIONS

Biomarkers in Molecular Epidemiology

As discussed above, specific TP53 mutations are observed in some types of cancer that are likely to have arisen from exposure to known carcinogens. The existence of important variations in mutation patterns between different groups of patients with the same cancer suggests that further mutation fingerprints related to environmental exposures are still to be discovered. This is particularly the case in squamous cell carcinomas of the head and neck and of the esophagus, where variations have been described in relation with geographic origin of the patients, environmental risk factors, and lifestyle. However, in the natural history of human cancers, there is often no evidence for high exposure to genotoxic substances and most mutations may arise through spontaneous replication errors, endogenous pro-mutagenic states (induced by reactive oxygen species), spontaneous deamination, or other mechanisms contributing to the spontaneous decay of DNA (Lindahl 1993). Furthermore, the association between characteristic mutation type and carcinogen needs support from experimental systems and other lines of evidence to gain plausibility. Moreover, any characteristic single-base “fingerprint” mutation induced by an exogenous risk factor may also arise spontaneously, i.e., in the absence of exposure (albeit perhaps rarely), calling for corroborative studies to substantiate links between exposures and cancer mutations. In the realm of molecular epidemiology, where large numbers, control groups, and robust statistics are mandatory, such studies are facilitated by technical advances allowing the identification of mutations in DNA extracted from surrogate biological samples such as plasma, urine, sputum, or exfoliated cells from bronchus, bladder, oral cavity, and esophagus (Kirk et al. 2000). The use of such surrogate materials is compatible with the development of studies that follow consistent epidemiological designs, e.g., case-control and prospective cohort studies.

Biomarkers in the Clinics

Mutations in TP53 are useful markers of tumor clonality to compare, in individual patients, separate clusters of tumor cells from the same lesion, or multiple lesions arising in the same tissue (Dix et al. 1995; Ponten et al. 1997; Franklin et al. 1997). Mutations are also useful for the follow-up of minimal residual disease, for comparison between primary and recurrent tumors and for tracing the origin of distant metastases (Franklin et al. 1997). However, the finding of different TP53 mutations in separate clusters of a tumor does not exclude a clonal
origin with occurrence of TP53 mutations at a late stage. Detection of TP53 mutations may also help to identify early lesions at a high risk of evolution. For example, in esophageal and endometrial cancers, detection of a mutation in low-grade dysplasia should be considered as an indicator of high risk of malignant evolution (Montesano et al. 1998; Jia et al. 2008).

TP53 mutations, but not p53 positive immunohistochemistry (IHC), have been consistently associated with poor prognosis in cancers such as breast, colorectal, head and neck, and leukemia (Petitjean et al. 2007a). In breast cancers, there is now clear evidence that TP53 mutation is an independent marker of poor prognosis, in particular in hormone receptor-positive cases. A large number of studies that have used IHC as a surrogate marker for mutation status have failed to provide such consistent results. IHC leads to an unacceptable number of misclassified cases (false-positive and false-negative) and to a large interstudy variability. TP53 gene status has also been associated with response to specific treatment regimens in breast cancer (Bergh et al. 1995; Aas et al. 1996; Olivier et al. 2006). However, it should be noted that the prognostic and predictive significance of TP53 mutations is extremely variable according to tumor type and/or treatment (Bertheau et al. 2008), and there is therefore no simple, universal clinical message that can be delivered by TP53 mutation analysis.

Mutations may also serve as biomarkers for targeted therapy. Several strategies have been developed to reactivate normal p53 functions in p53 mutated tumors. They include molecules that target broad classes of mutants to reactivate suppressive functions in tumor cells (PRIMA, RTA, scFv) (Caron de Fromentel et al. 1999; Issaeva et al. 2004; Bykov et al. 2005), or compounds that specifically target particular missense mutants to restore wild-type-like structure (Phikan059 targeting R220C)(Boeckler et al. 2008). Interestingly, several of these small pharmacological compounds appear to share a common chemical activity as Michael acceptors and can modify thiols in p53, suggesting that they affect p53 folding through redox regulation (Lambert et al. 2009). Peptides have also been generated that interact specifically with several p53 mutants and block their nonspecific transactivation capacities, inhibiting GOF properties. These approaches complement other approaches such as gene therapy that target tumors that do not express p53 mutants (Senzer et al. 2009).

**CONCLUDING REMARKS**
The TP53 gene has been the most extensively sequenced tumor gene before the era of large scale sequencing, generating a large amount of data to identify links between carcinogens, mutation fingerprints, and tumorigenesis. Recent data produced by genome-wide sequencing have confirmed that overall mutation patterns in cancer genomes are quite similar to those in TP53 itself. The existence of several in vitro and in vivo mutagenesis assays such as the HUF assay provides tools to experimentally corroborate observations on mutation patterns in human cancers. Thus, studies on TP53 mutations will further contribute to the understanding of gene-environment interactions in cancer, in particular when comparing variations in TP53 mutation patterns in relation to different cohorts of patients.

In terms of clinical applications, TP53 mutations have proven to be extremely complex biomarkers. Despite impressive progress in mechanistic understanding of p53 structure and function, p53 research has not yet generated applications of wide impact on cancer management and therapy. Somehow, the complexity of the p53 field may have acted as a deterrent for clinical applications. The fact that mutations in TP53 are diverse in their biological effects and that they may occur at many different stages before or during tumor development makes it impossible to derive simple messages uniformly applicable to all clinical contexts. Translating TP53 mutation into the clinics will require large, structured clinical trials in which patients with defined p53 status are recruited on the basis of specific inclusion criteria, randomized for treatment according to determined regimens, and followed up for long-term therapeutic and clinical end points. Databases such as the current TP53 mutation databases will

Cite this article as Cold Spring Harb Perspect Biol 2010;2:a001008

13
have a critical role in collecting, structuring, and annotating these data, allowing for the interpretation of TP53 mutation and their educated use in standard molecular pathology practice.

REFERENCES


TP53 Mutation Spectrum in Cancer


Langerod A, Zhao H, Borgan O, Nesland JM, Bukholm IR, Ikdahl T, Karesen R, Borresen-Dale AL, Jeffrey SS. 2007. TP53 mutation status and gene expression profiles are
M. Olivier, M. Hollstein, and P. Hainaut


# TP53 Mutations in Human Cancers: Origins, Consequences, and Clinical Use

Magali Olivier, Monica Hollstein and Pierre Hainaut

*Cold Spring Harb Perspect Biol* 2010; doi: 10.1101/cshperspect.a001008 originally published online November 11, 2009

<table>
<thead>
<tr>
<th>Subject Collection</th>
<th>The p53 Family</th>
</tr>
</thead>
<tbody>
<tr>
<td>The Origins and Evolution of the p53 Family of Genes</td>
<td>The Tumor Suppressor p53: From Structures to Drug Discovery</td>
</tr>
<tr>
<td>Vladimir A. Belyi, Prashanth Ak, Elke Markert, et al.</td>
<td>Andreas C. Joerger and Alan R. Fersht</td>
</tr>
<tr>
<td>Mouse Models of p53 Functions</td>
<td>p53 Regulation of Metabolic Pathways</td>
</tr>
<tr>
<td>Guillermina Lozano</td>
<td>Eyal Gottlieb and Karen H. Vousden</td>
</tr>
<tr>
<td>TP53 Mutations in Human Cancers: Origins, Consequences, and Clinical Use</td>
<td>The Regulation of the p53-mediated Stress Response by MDM2 and MDM4</td>
</tr>
<tr>
<td>Magali Olivier, Monica Hollstein and Pierre Hainaut</td>
<td>Mary Ellen Perry</td>
</tr>
<tr>
<td>p53 Research: The Past Thirty Years and the Next Thirty Years</td>
<td>Zebradfish Models of p53 Functions</td>
</tr>
<tr>
<td>David Lane and Arnold Levine</td>
<td>Narie Y. Storer and Leonard I. Zon</td>
</tr>
<tr>
<td>Transcriptional Regulation by P53</td>
<td>p63 and p73, the Ancestors of p53</td>
</tr>
<tr>
<td>Rachel Beckerman and Carol Prives</td>
<td>V. Dötsch, F. Bernassola, D. Coutandin, et al.</td>
</tr>
<tr>
<td>p53-based Cancer Therapy</td>
<td>Pathologies Associated with the p53 Response</td>
</tr>
<tr>
<td>David P. Lane, Chit Fang Cheok and Sonia Lain</td>
<td>Andrei V. Gudkov and Elena A. Komarova</td>
</tr>
<tr>
<td>Phylogeny and Function of the Invertebrate p53 Superfamily</td>
<td>Single-nucleotide Polymorphisms in the p53 Signaling Pathway</td>
</tr>
<tr>
<td>Rachael Rutkowski, Kay Hofmann and Anton Gartner</td>
<td>Lukasz F. Grochola, Jorge Zeron-Medina, Sophie Mériaux, et al.</td>
</tr>
<tr>
<td>Tied Up in Loops: Positive and Negative Autoregulation of p53</td>
<td>Clinical Outcomes and Correlates of TP53 Mutations and Cancer</td>
</tr>
<tr>
<td>Xin Lu</td>
<td>Ana I. Robles and Curtis C. Harris</td>
</tr>
</tbody>
</table>

For additional articles in this collection, see [http://cshperspectives.cshlp.org/cgi/collection/](http://cshperspectives.cshlp.org/cgi/collection/)

Copyright © 2010 Cold Spring Harbor Laboratory Press; all rights reserved