Carcinogenesis vol.22 no.3 pp.367–374, 2001

Patterns of p53 G→T transversions in lung cancers reflect the primary mutagenic signature of DNA-damage by tobacco smoke

Pierre Hainaut[1,3] and Gerd P.Pfeifer[2,3]

[1]International Agency for Research on Cancer (WHO), 150 Cours Albert Thomas, 69372 Lyon cedex, France and [2]Department of Biology, Beckman Research Institute of the City of Hope, Duarte, CA 91010, USA

[3]To whom correspondence should be addressed
Email: hainaut@iarc.fr
Email: gpfeifer@coh.org

It is unquestionable that the major cause of lung cancer is cigarette smoking. p53 mutations are common in lung cancers from smokers but less common in non-smokers. A large fraction of the p53 mutations in lung cancers are G→T transversions, a type of mutation that is infrequent in other tumors aside from hepatocellular carcinoma. Previous studies have indicated that there is a good correlation between G→T transversion hotspots in lung cancers and sites of preferential formation of polycyclic aromatic hydrocarbon (PAH) adducts along the p53 gene. The origin of p53 mutations in lung cancer has been questioned by recent reports suggesting that there are no significant differences in p53 mutation spectra between smokers and non-smokers and between lung cancers and non-lung cancers [S.N.Rodin and A.S.Rodin (2000) Human lung cancer and p53: The interplay between mutagenesis and selection. *Proc. Natl Acad. Sci. USA*, 97, 12244–12249]. We have reassessed these issues by using the latest update of the p53 mutation database of the International Agency for Research on Cancer (14 051 entries) as well as recent data from the primary literature on non-smokers. We come to the conclusion that the p53 mutation spectra are different between smokers and non-smokers and that this difference is highly statistically significant (G→T transversions are 30 versus 10%; $P < 0.0001$, χ^2 test). A similar difference is seen between lung cancers and non-lung cancers. At a number of mutational hotspots common to all cancers, a large fraction of the mutations are G→T transversions in lung cancers but are almost exclusively G→A transitions in non-lung cancers. Our data reinforce the notion that p53 mutations in lung cancers can be attributed to direct DNA damage from cigarette smoke carcinogens rather than to selection of pre-existing endogenous mutations.

Introduction

Before the advent of mass production of cigarettes around the time of the first World War, lung cancer was extremely uncommon and represented <1% of all autopsied cancer cases (1). Today, lung cancer is the leading cause of cancer death in the US for both women and men (2). It is estimated that lung cancer kills over one million people each year worldwide. It has long been recognized that the major cause of lung cancer is cigarette smoking (3–5; reviewed in refs 1,6–8). About 90% of the lung cancer deaths in the US are caused by smoking (9). Recent surveys indicate that lung cancer is also rapidly becoming the major cause of cancer death in many developing countries.

Lung cancer arises through a multi-step process in which, according to current understanding, a limited number of crucial genetic alterations need to occur to bring a normal cell onto the path of malignancy (10). Lung cancers are broadly classified into two groups: small cell lung cancer (SCLC) and non-small cell lung cancer (NSCLC), the latter being divided into the major histological subtypes adenocarcinoma, squamous cell carcinoma and large cell carcinoma. More than 50% of all lung cancers carry a mutation in the p53 gene. These mutations are more common in smokers than in non-smokers and there is a dose-dependence of p53 mutation frequency according to the number of cigarettes smoked (11–15). p53 mutations can be found in preneoplastic lesions of the lung (16–20) indicating that they are early events and may be temporally linked to DNA damage from smoking.

It has been noticed that p53 mutations in lung cancer are different from those in other cancers and that an excess of G→T transversions is characteristic for these tumors (21–23). G→T transversions have been likened to a 'molecular signature' of tobacco smoke mutagens in smoking-associated lung cancers for the following reasons: (i) polycyclic aromatic hydrocarbons (PAHs) are prominent carcinogens in tobacco smoke that can produce predominantly this type of mutation (24–28); (ii) PAH adducts are present in DNA extracted from human tissues exposed to tobacco smoke (29–33); (iii) there is an increased frequency of G→T transversions in lung cancers from smokers compared with lung cancers from non-smokers and compared with most other cancers (12–14,22); (iv) a non-transcribed strand bias of G→T transversions can be attributed to preferential repair of adducts on the transcribed strand (34–36); and (v) there is a precise correspondence between mutational hotspots and hotspots of adduct formation by PAHs found in tobacco smoke (37,38).

In a recent issue of the Proceedings of the National Academy of Sciences, Rodin and Rodin argued that p53 mutations in lung cancer may have a different cause, namely selection of pre-existing or endogenous mutations by physiological stresses aggravated by smoking (39). There have been several reports in the medical press about the conclusions drawn in this article (40). In another recent publication (41), Paschke compared mutation data in successive releases of the database of p53 mutations maintained at the International Agency for Research on Cancer (IARC p53 mutation database). Focusing on a number of apparent discrepancies between these successive releases, he concludes that the previously reported differences in the mutation spectrum between smokers and non-smokers are not statistically significant or might be influenced by confounding factors such as ethnicity. The publication by Paschke (41) originated in a private research institution supported by the German Association of Cigarette Manufacturers.

Abbreviations: BPDE, (+/–) anti-benzo[*a*]pyrene-7,8-diol-9,10-epoxide; PAH, polycyclic aromatic hydrocarbon.

Because of the broad implications that these conclusions may have on public health issues, we have re-analyzed all the primary literature currently available on p53 mutations in lung cancers from non-smokers and smokers, as well as in other cancers where tobacco smoke is not a major risk factor.

p53 mutations in smokers and non-smokers

The main claim made by Rodin and Rodin is that 'there are no significant differences between smokers and non-smokers, either in the frequency of different types of mutations or in the frequency of their occurrence along the p53 gene' (39). They argue that a link between smoke-induced mutations and DNA adducts (37) could be fortuitous, and that it is selection of pre-existing endogenous mutations rather than smoking-induced DNA damage that determines the p53 mutational spectrum in lung tumors. Another recent report also claims that there are no differences with respect to G→T transversions between smokers and non-smokers (41).

To re-assess these issues, we have extracted p53 mutation data from the April 2000 release of the IARC p53 mutation database (14 051 entries). It should be stressed that this database is exclusively a repository for mutations described in peer-reviewed articles. It is therefore affected by a number of biases (reviewed in ref. 42). Indiscriminate analysis of the database without selection criteria is very likely to generate biased results, as indeed is the case in the recent analyses published by Rodin and Rodin (39) and Paschke (41). Therefore, we have reassessed each database entry on lung cancers and tobacco consumption from the primary literature. We have excluded samples from known occupational exposures including radon gas-exposed uranium miners, coal miners and individuals exposed to asbestos, mustard gas and high levels of radon gas in their residence. Further, 107 lung cancer mutations from one highly unusual report were excluded (43). These include 59 mutations occurring in non-smokers. In this particular publication, sequence data were obtained after subcloning of PCR products. This experimental design is apparently prone to artifacts. A total of 20 mutations were found in codons of the *k-ras* gene not usually involved in tumor mutations, such as codons 14–32, and only one mutation was found in the commonly mutated codon 12 of *k-ras*. In the p53 gene, one tumor sample harbored no less than 14 different missense mutations and two other samples had 12 p53 mutations each (43). We think it is justified to eliminate these entries. Data obtained with mutant-enriched restriction fragment length polymorphism–polymerase chain reaction (RFLP–PCR) and the yeast functional assay were also not included because of a severe scoring bias inherent to these techniques. The yeast assay picks up mutations that occur in only a small fraction of the tumor cells or mutations that may be derived from DNA damage *in vitro*, is limited to mutations that destroy the transcriptional activation function of p53, and is based on cloning of a single PCR-derived molecule. The RFLP–PCR assay scores mutations only at a few pre-selected restriction sites. In addition, identical mutations, reported two or three times by the same group in different publications, were scored only once in our analysis.

Differences in p53 mutation spectra between smokers and non-smokers have been reported previously (13,14,38). We have confirmed p53 mutations in 99 lung cancer cases of non-smokers from the primary literature (11,44–59). The types of p53 mutations occurring in smokers and non-smokers are shown in Figure 1. Ten percent of the p53 mutations in non-smokers are G→T transversions. This does not agree with the much higher levels (>22%) claimed by Rodin and Rodin (39) and Paschke (41), presumably because of their indiscriminate use of the database sets. The difference between 10% G→T mutations in non-smokers and 29% G→T in smokers is highly statistically significant ($P < 0.0001$; χ^2 test). The frequency of G→T transversions is much higher in lung cancers than it is in any other tumor type except for liver cancers occurring in geographic areas where a contamination of food products with aflatoxins is present (60). In most internal cancers, not directly linked to tobacco consumption, the frequencies of G→T mutations are between 8 and 12% (Figure 1). This is similar to the percentage of G→T mutations in non-smokers. In addition, non-smokers have an increased level of G→A transitions (50%) as opposed to smokers (28%). Again, these differences are statistically highly significant ($P < 0.0001$; χ^2 test) (Figure 1). Interestingly, the frequency of G→T transversions in lung cancers is increased in never-smokers (some of them coal miners) from Silesia, a region with considerable air pollution (61).

It is also of note that, by focusing only on p53 mutations in cases where the smoking status is known, Rodin and Rodin (39) and Paschke (41) removed ~85% of the lung cancer data from their analysis. For these reasons we chose to include categories of both 'designated smokers' (where the smoking status is known from the literature) and 'all lung cancer cases minus non-smokers'. It is reasonable to assume that mutations with no smoking status specified are most probably connected to smoking given the fact that ⩾90% of all lung cancers occur in smokers (9). Indeed, the proportion of G→T transversions in all lung cancers (minus non-smokers) is remarkably similar to that observed in designated smokers when taken separately (Figure 1). When one analyzes the different types of lung cancers, the frequencies of G→T transversions in the April 2000 p53 database are 29% in adenocarcinomas, 31% in squamous cell carcinomas, 28.5% in SCLCs and 43% in large cell carcinomas.

The second claim made by Rodin and Rodin (39) is that any differences between smokers and non-smokers disappear when each DNA strand is analyzed separately. We cannot reproduce these observations. Figure 2 shows that, even when the two DNA strands are analyzed separately (e.g. G→T mutations versus C→A mutations), smokers are clearly different from non-smokers by the much higher involvement of G→T in general ($P < 0.0001$; χ^2 test), a much lower relative frequency of G→A and C→T ($P = 0.001$; χ^2 test), and a more pronounced strand bias for G→T mutations. The latter observation is in fact consistent with transcription-coupled repair of bulky DNA lesions, such as (+/–) anti-benzo[*a*]pyrene-7,8-diol-9,10-epoxide (BPDE) adducts, in the p53 gene (35).

If one examines the distribution of G→T mutations at a higher level of resolution, one finds that the G→T mutations in non-smokers are distributed as follows: one at codon 148, one at codon 158, one at codon 242 and four at codon 249. This distribution does not match with the distribution of G→T transversions in designated smokers and all lung cancer cases minus non-smokers (Figure 3). In these cases, clear mutational hotspots for G→T are seen at codons 157, 158, 245, 248, 249 and 273. All of these, except codon 249, are strong binding sites for BPDE and other PAHs (38). Interestingly, codon 249, which is not a site of strong PAH adduct formation, is over-represented in non-smokers as well. On the other hand, there is only one G→T transversion in non-smokers (1 out of 99

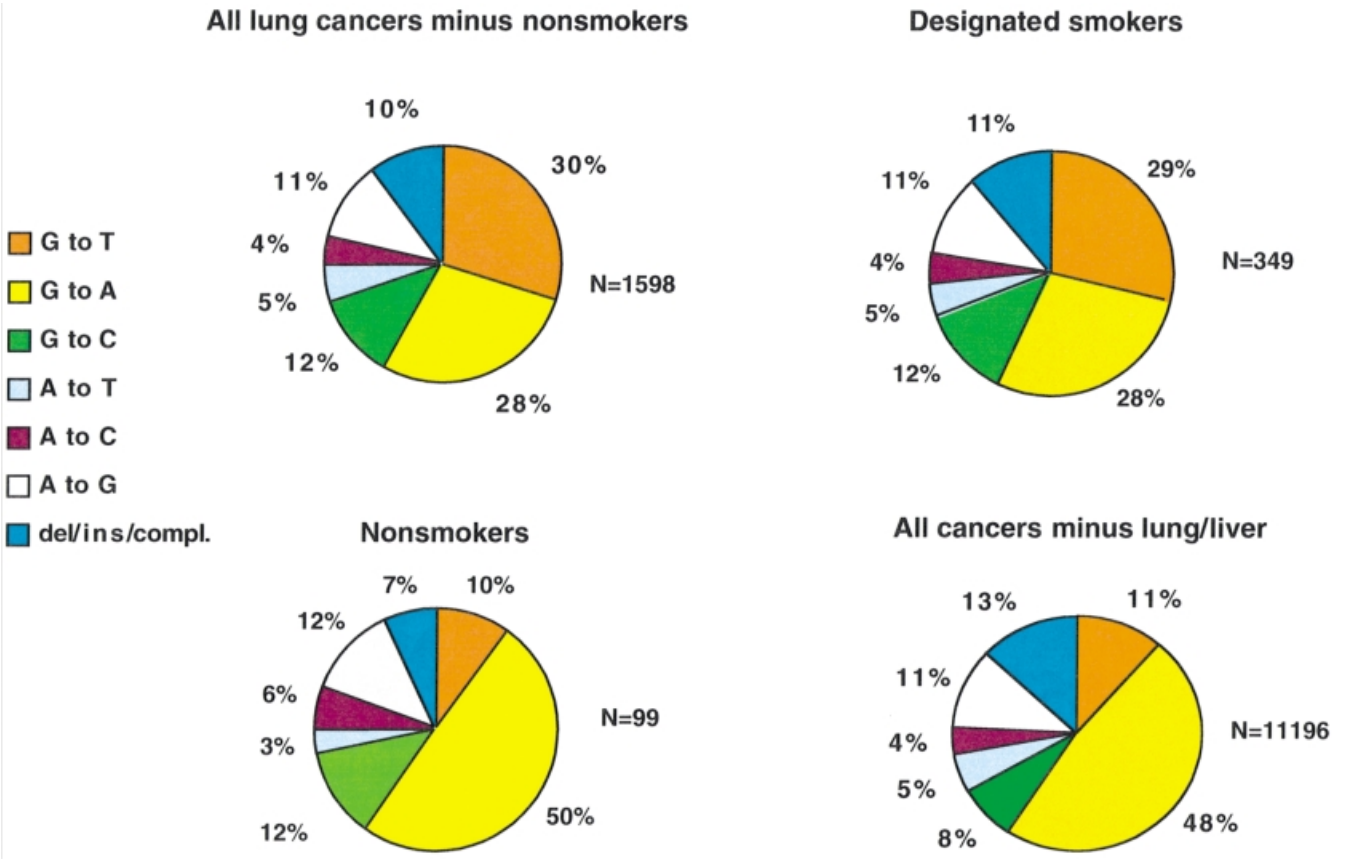

Fig. 1. The types of mutations in the p53 gene of human lung cancers. The numbers were obtained from the April 2000 update of the IARC database with 14 051 entries and from recent publications on p53 mutations in non-smokers (see text for all publications with data on non-smokers). All data on smoking status were confirmed from the primary literature. Cases of radon-, asbestos- and mustard gas-associated p53 mutations were excluded. Mutations reported in studies likely to be affected by laboratory artifacts were also excluded as described in the text. The total number of mutations were: all lung cancers minus non-smokers (n = 1598), non-smokers (n = 99), designated smokers (n = 349), all cancers minus lung and liver cancers (n = 11 196). When smoking-related lung cancers and aflatoxin-related liver cancers are removed, the frequency of G→T transversions in all the remaining cancers is similar to that in lung cancers from non-smokers. Del/ins/compl., deletions, insertions and complex mutations.

total mutations) that occurs at a site, which is a strong target for PAH adduction (codon 158).

The distribution of G→T transversions in lung and other cancers

A comparison of G→T patterns along the p53 gene can be made between lung cancers from smokers versus G→T patterns in selected non-lung tumors. In Figure 3, we have repeated the same analysis as that presented in figure 2 of the paper by Rodin and Rodin (39), using the April 2000 release of the IARC p53 mutation database. It is clear that all major G→T mutational hotspots are much more pronounced in lung cancers than in cancers occurring in tissues considered as 'least accessible to smoke'. In these tissues, codon 176 is the most commonly affected site for G→T transversions and the distribution of these transversions is much broader than in lung cancers. One of the most important differences is observed at codon 158, which represents 12.5% of all G→T transversions in lung cancers from smokers, compared with 1.6% in tumors 'least accessible to smoke' (Figure 3).

Even if there is an accumulation of G→T mutations at some lung cancer hotspots in other tumors, the interpretation of this coincidence does not necessarily need to involve 'selection' of endogenous mutations as proposed (39). The data for non-lung tumors are dominated by a large number of breast, stomach and colon tumors ($n > 2500$). These tissues are either exposed to certain levels of PAHs directly (colon and stomach; refs 62,63) or tumor etiology has been specifically linked to such carcinogens (breast; refs 64–66). Since PAH metabolites target methylated CpG sequences (including those at codons 157, 158, 245, 248 and 273) very specifically (67), the patterns may look somewhat similar to those in lung cancer if one examines only G→T events. What really distinguishes lung cancer from the cancers 'less accessible to smoke' is that a specific G→T pattern (which correlates with the PAH adduct patterns; refs 37,38) is at least three times more frequent in lung cancer than in the other cancers. Rodin and Rodin (39) entirely ignore this crucial difference and their interpretation disguises this fact. In other words, if selection of pre-existing endogenous mutations is the major driving force, as proposed, then why does one see only 9–11% G→T mutations in the other tissues rather than 30%? Of course, selection itself, that is selection from smoking-induced mutations, does play a role, since only mutants that confer tumorigenicity will eventually be detectable in cancer cells. However, selection in p53 is much less dominant than in, for example, the *ras* genes, where only a few codons can be found mutated in tumors. The p53 gene is also different from other tumor suppressor genes in that missense mutations, instead of nonsense or frameshift mutations, are the vast majority of all mutations. A large number of diverse missense mutations are found at high

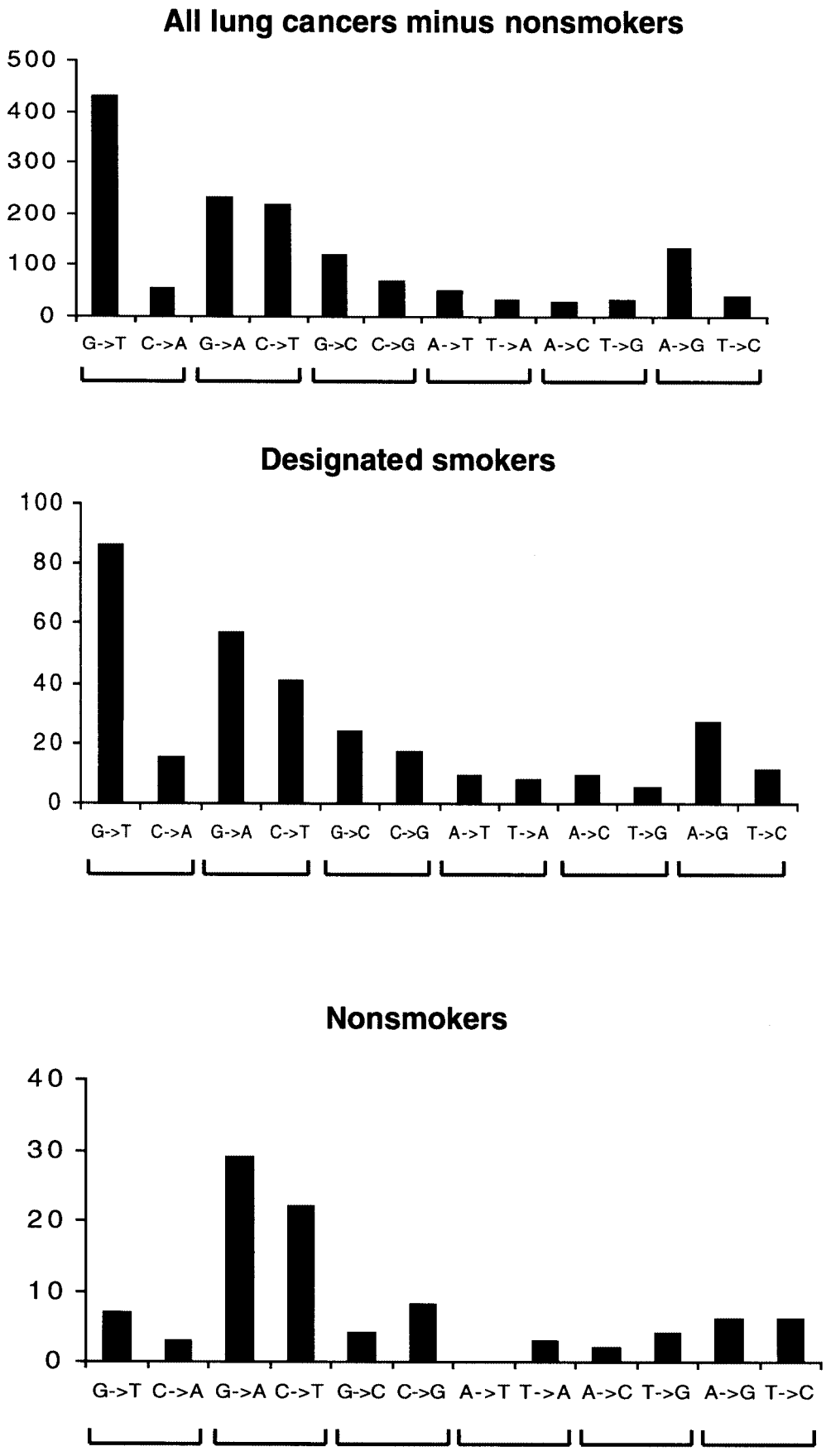

Fig. 2. Strand-specific p53 mutational patterns in lung cancers. The numbers were obtained using the same criteria as described in Figure 1.

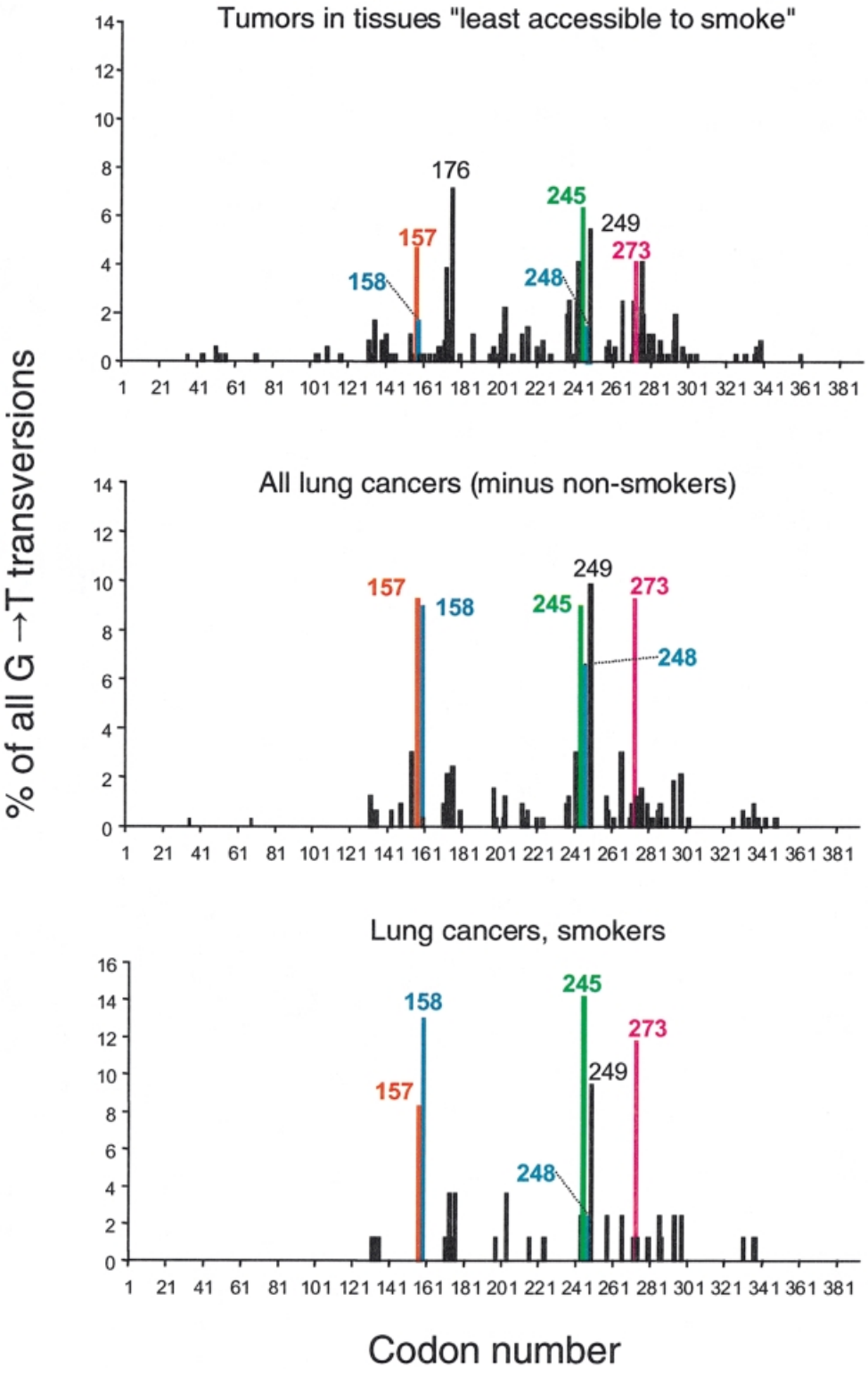

Fig. 3. Comparison of p53 mutational spectra of G→T transversions in lung cancer and spectra of cancers considered to arise in tissues 'least accessible to smoke'. Smokers designated as such in the literature (n = 85) and all lung cancers minus non-smokers (n = 336) are shown separately. Cancers from tissues considered least accessible to smoke include brain, breast, colon, rectum, skin, hematopoietic and reticuloendothelial systems, testes, prostate, stomach and lymph nodes (n = 367). The spectrum of G→T transversions in non-smokers is not drawn because there were only seven data points (one at codon 148, one at codon 158, one at codon 242 and four at codon 249). Hotspots are indicated.

frequency in the p53 database. This makes the p53 gene an excellent tool to analyze possible etiological factors leading to these mutations.

The major lung cancer hotspots 158, 245, 248 and 273 are commonly G→T transversions in lung cancer but are other mutation types (mostly G→A) in other tumors. This does not

Table I. Frequencies of different types of p53 mutations occurring at four different codons in human lung, brain, breast and colon cancers

Tissue-specific mutations at codon 157				
	lung	brain	breast	colon
G→T	40 (75)[a]	5 (83)	5 (42)	7 (50)
G→A	5 (9)	0 (0)	2 (17)	0 (0)
G→C	2 (4)	0 (0)	0 (0)	4 (29)
A→T	4 (7)	0 (0)	0 (0)	0 (0)
A→C	1 (2)	0 (0)	2 (17)	0 (0)
A→G	0 (0)	0 (0)	1 (8)	0 (0)
del/ins	1 (2)	1 (17)	2 (17)	3 (21)
total	53 (100)	6 (100)	12 (100)	14 (100)
Tissue-specific mutations at codon 158				
	lung	brain	breast	colon
G→T	37 (62)	2 (11)	2 (33)	1 (10)
G→A	9 (15)	8 (42)	3 (50)	7 (70)
G→C	11 (18)	8 (42)	1 (17)	0 (0)
A→T	0 (0)	0 (0)	0 (0)	0 (0)
A→C	0 (0)	0 (0)	0 (0)	0 (0)
A→G	0 (0)	0 (0)	0 (0)	0 (0)
del/ins	3 (5)	1 (5)	0 (0)	2 (20)
Total	60 (100)	19 (100)	6 (100)	10 (100)
Tissue-specific mutations at codon 248				
	lung	brain	breast	colon
G→T	32 (35)	0 (0)	2 (2)	1 (1)
G→A	50 (55)	60 (97)	75 (96)	162 (98)
G→C	6 (7)	0 (0)	0 (0)	2 (1)
A→T	0 (0)	0 (0)	0 (0)	0 (0)
A→C	0 (0)	0 (0)	0 (0)	0 (0)
A→G	0 (0)	0 (0)	1 (1)	0 (0)
del/ins	2 (2)	2 (3)	0 (0)	0 (0)
Total	90 (100)	62 (100)	78 (100)	165 (100)
Tissue-specific mutations at codon 273				
	lung	brain	breast	colon
G→T	41 (44)	3 (3)	4 (6)	0 (0)
G→A	41 (44)	104 (97)	61 (88)	100 (98)
G→C	7 (8)	0 (0)	2 (3)	0 (0)
A→T	0 (0)	0 (0)	0 (0)	0 (0)
A→C	0 (0)	0 (0)	0 (0)	0 (0)
A→G	3 (3)	0 (0)	1 (1)	0 (0)
del/ins	1 (1)	0 (0)	1 (1)	2 (2)
Total	93 (100)	107 (100)	69 (100)	102 (100)

[a]The total numbers are given with the percentage of the total at each codon in parentheses.

agree with the selection theory put forward by Rodin and Rodin (39). In Table I, we present data for four lung cancer p53 mutational hotspots. The tumor types analyzed are lung, breast, colon and brain. These tumors all have sufficiently large numbers of mutations in the p53 database (between 750 and 1600). Codons 157 and 158 are frequently mutated in lung cancers (mostly as G→T events) but are much less commonly involved in the other tumors (Table I). These two codons are lung cancer-specific mutational hotspots. That codons 248 and 273 are often the targets of G→T mutations in lung cancers is a perfect illustration of the respective roles of mutagens and selection in shaping the mutation pattern of p53. These two codons are those most commonly mutated in the entire p53 database, and this is true in almost every type of human cancer (liver cancer linked with exposure to aflatoxins is the only notable exception). This is so because mutating these residues has a drastic effect on p53 protein function. The residues encoded by codons 248 and 273 mediate contacts between the p53 protein and its target DNA (68,69), and their mutation abrogates the capacity of p53 to act as a transcription factor. However, in lung cancers, the mutations at these codons differ dramatically in their nature from all cancers not associated with tobacco smoke (Table I): they are frequently G→T transversions at the bases identified *in vitro* as sites of adduction of metabolites of benzo[*a*]pyrene and other PAHs (37,38). At codons 248 and 273, 35–44% of the mutations are G→T transversions in lung cancer, but such mutations are virtually non-existent in the other tumors. The non-G→T mutations at these codons in lung cancer may involve carcinogens other than PAHs or, alternatively, they may in fact be caused by PAH metabolites, which can produce mutations other than G→T (70). These observations provide direct evidence that p53 mutations in lung cancer occur by a distinct mechanism, and that this mechanism is consistent with the mutagenic effects of benzo[*a*]pyrene, a major carcinogen present in tobacco smoke.

The issue of silent mutations

To assess the role of tobacco carcinogens in the etiology of p53 mutations, it may be argued that smokers should have more silent mutations than non-smokers and that one should see silent mutations at the same sites, which are strong sites of adduct formation (39). First, we are sceptical about the theories on the origin of silent and multiple mutations in the p53 database (71,72). Silent mutations in p53 occur at a frequency of 4.9% of all mutations in all cancers combined. Somatic mutation frequencies of various genes are at least two orders of magnitude lower than is suggested by these numbers for silent mutations. This has been interpreted as being a consequence of the p53 gene being hypermutable, although a specific mechanism for this presumed hypermutability is not obvious (73). A large fraction of these mutations are probably derived from sequencing problems, errors during scoring of mutations from the sequencing runs, and systematic errors generated during sequencing of PCR products after subcloning into plasmids. In this case, PCR errors from polymerase infidelity will be scored. For example, we found that the frequency of silent mutations is 4.7% (419 of 8925) in the p53 database when the PCR products were sequenced directly, but 9.5% (127 of 1332) when subcloned DNA was used. The silent mutations are found in all tumor types and are scattered over 200 different p53 codons, with no apparent hotspots. The significance of silent mutations is unknown but it is very likely that they represent a certain noise level in the p53 database. This is probably the margin of error we have when we try to draw lessons from p53 mutational patterns. Also, in the analysis of silent mutations, one needs to consider the possibility that a considerable fraction of these mutations may be rare polymorphisms as discussed previously (74). Unfortunately, the surrounding normal tissue of a tumor is rarely analyzed for p53 mutations and a differentiation between rare polymorphism and somatic silent mutation cannot be made. To further illustrate this fact, we found that 22 silent mutations (3.2% of all silent mutations) are reported to occur in cancer at the third base of codon 213 (CGA→CGG). This 'mutation' corresponds to one of the most frequent polymorphisms in the p53 gene (75).

Rodin and Rodin (39) argue that the distribution of silent mutations may provide a more reliable basis for detection of possible carcinogen fingerprints, as these mutations are not subjected to selection. They speculate that, if a carcinogen leaves its mark in the genome, this should be reflected in the number and type of silent mutations. We decided to test this hypothesis by analyzing the number of silent mutations in

hepatocellular carcinoma, a cancer strongly associated with exposure to an exogenous carcinogen inducing G→T transversions, aflatoxin. We found that silent mutations occur in liver cancer at a rate of 5.5% (40 out of 733). Only 17.5% of these silent mutations are GC→TA transversions (7 out of 40) compared with 32% for non-silent mutations. In lung cancer, these numbers are similar. Eight of 60 (13%) of the silent and 475 of 1551 (31%) of the non-silent mutations are G→T mutations. Thirty-nine out of 60 (65%) of the silent lung cancer mutations are G→A transitions compared with 27% G→A mutations for non-silent events. Thus, when a tumor is causally associated with a carcinogen that leaves a fingerprint in the p53 mutation pattern, this fingerprint is not necessarily reflected in the pattern of silent mutations. We believe that this analysis casts serious doubt on the usefulness of the silent mutations in this type of analysis, since many of them may in reality be artifacts or rare polymorphisms not in any way related to tumor etiology. Even if some of the silent G→T mutations were in fact related to carcinogenic exposure, they are so infrequent, relative to non-silent ones (8 versus 475), that they cannot produce a reliable mutation spectrum in lung cancer, and one cannot expect to score any significant number of silent mutation at the site of an adduct hotspot.

Conclusions

Overall, the survey of p53 mutation data available clearly indicates that there is an increased frequency of G→T transversions in lung cancer from smokers compared with non-smokers and with most other cancers. In addition, a non-transcribed strand bias of G→T transversions correlates with preferential repair of PAH adducts on the transcribed strand (35). Finally, there is a precise correlation between mutational hotspots and hotspots of adduct formation by PAHs found in tobacco smoke (37,38). We consider that, in the face of this experimental and molecular pathological evidence, the pattern of G→T transversions in lung cancers clearly reflects the primary mutagenic signature of smoke-induced DNA damage rather than selection of pre-existing endogenous mutations.

Acknowledgements

We would like to thank G.Pershagen for specifying the precise location of mutations in a series of lung cancers from non-smokers, M.Olivier for extraction and analysis of data from the IARC p53 database, P.Frankel for help with the statistical analysis and M.F.Denissenko for discussions. The work of the authors was supported by a grant from the National Cancer Institute (CA84469 to G.P.P.). The IARC p53 database is supported by a grant from the European Community (QLG-1999-00273).

References

1. Wynder,E.L. and Hoffmann,D. (1994) Smoking and lung cancer: scientific challenges and opportunities. *Cancer Res.*, **54**, 5284–5295.
2. Wingo,P.A., Ries,L.A., Giovino,G.A., Miller,D.S., Rosenberg,H.M., Shopland,D.R., Thun,M.J. and Edwards,B.K. (1999) Annual report to the nation on the status of cancer, 1973–1996, with a special section on lung cancer and tobacco smoking. *J. Natl Cancer Inst.*, **91**, 675–690.
3. Wynder,E.L. and Graham,E.A. (1950) Tobacco smoking as a possible etiological factor in bronchiogenic carcinoma. A study of six hundred and eighty-four proved cases. *J. Am. Med. Assoc.*, **143**, 329–336.
4. Doll,R. and Hill,A.B. (1950) Smoking and carcinoma of the lung. *Br. Med. J.*, **2**, 739–748.
5. United States Surgeon General (1964) Smoking and Health Pub. No. 1103. Department of Health, Education and Welfare, Washington DC.
6. Loeb,L.A., Ernster,V.L., Warner,K.E., Abbotts,J. and Laszlo,J. (1984) Smoking and lung cancer: an overview. *Cancer Res.*, **44**, 5940–5958.
7. Hecht,S.S. (1999) Tobacco smoke carcinogens and lung cancer. *J. Natl Cancer Inst.*, **91**, 1194–1210.
8. Christiani,D.C. (2000) Smoking and the molecular epidemiology of lung cancer. *Clin. Chest Med.*, **21**, 87–93.
9. Shopland,D.R. (1995) Tobacco use and its contribution to early cancer mortality with a special emphasis on cigarette smoking. *Environ. Health Perspect.*, **103** (Suppl. 8), 131–142.
10. Sekido,Y., Fong,K.M. and Minna,J.D. (1998) Progress in understanding the molecular pathogenesis of human lung cancer. *Biochim. Biophys. Acta*, **1378**, F21–F59.
11. Suzuki,H., Takahashi,T., Kuroishi,T., Suyama,M., Ariyoshi,Y., Takahashi,T. and Ueda,R. (1992) p53 mutations in non-small cell lung cancer in Japan: association between mutations and smoking. *Cancer Res.*, **52**, 734–736.
12. Husgafvel-Pursiainen,K. and Kannio,A. (1996) Cigarette smoking and p53 mutations in lung cancer and bladder cancer. *Environ. Health Perspect.*, **104** (Suppl. 3), 553–556.
13. Hernandez-Boussard,T.M. and Hainaut,P. (1998) A specific spectrum of *p53* mutations in lung cancer from smokers: review of mutations compiled in the IARC *p53* database. *Environ. Health Perspect.*, **106**, 385–391.
14. Bennett,W.P., Hussain,S.P., Vahakangas,K.H., Khan,M.A., Shields,P.G. and Harris,C.C. (1999) Molecular epidemiology of human cancer risk: gene-environment interactions and mutation spectrum in human lung cancer. *J. Pathol.*, **187**, 8–18.
15. Ahrendt,S.A., Chow,J.T., Yang,S.C., Wu,L., Zhang,M.J., Jen,J. and Sidransky,D. (2000) Alcohol consumption and cigarette smoking increase the frequency of p53 mutations in non-small cell lung cancer. *Cancer Res.*, **60**, 3155–3159.
16. Sozzi,G., Miozzo,M., Donghi,R., Pilotti,S., Cariani,C.T., Pastorino,U., Della Porta,G. and Pierotti,M.A. (1992) Deletions of 17p and p53 mutations in preneoplastic lesions of the lung. *Cancer Res.*, **52**, 6079–6082.
17. Sundaresan,V., Ganly,P., Hasleton,P., Rudd,R., Sinha,G., Bleehen,N.M. and Rabbitts,P. (1992) p53 and chromosome 3 abnormalities, characteristic of malignant lung tumours, are detectable in preinvasive lesions of the bronchus. *Oncogene*, **7**, 1989–1997.
18. Wistuba,I.I., Lam,S., Behrens,C., Virmani,A.K., Fong,K.M., LeRiche,J., Samet,J.M., Srivastava,S., Minna,J.D. and Gazdar,A.F. (1997) Molecular damage in the bronchial epithelium of current and former smokers. *J. Natl Cancer Inst.*, **89**, 1366–1373.
19. Mao,L., Lee,J.S., Kurie,J.M. *et al.* (1997) Clonal genetic alterations in the lungs of current and former smokers. *J. Natl Cancer Inst.*, **89**, 857–862.
20. Chung,G.T., Sundaresan,V., Hasleton,P., Rudd,R., Taylor,R. and Rabbitts,P.H. (1995) Sequential molecular genetic changes in lung cancer development. *Oncogene*, **11**, 2591–2598.
21. Hollstein,M., Sidransky,D., Vogelstein,B. and Harris,C.C. (1991) p53 mutations in human cancers. *Science*, **253**, 49–53.
22. Greenblatt,M.S., Bennett,W.P., Hollstein,M. and Harris,C.C. (1994) Mutations in the p53 tumor suppressor gene: clues to cancer etiology and molecular pathogenesis. *Cancer Res.*, **54**, 4855–4878.
23. Hainaut,P. and Hollstein,M. (2000) p53 and human cancer: the first ten thousand mutations. *Adv. Cancer Res.*, **77**, 81–137.
24. Eisenstadt,E., Warren,A.J., Porter,J., Atkins,D. and Miller,J.H. (1982) Carcinogenic epoxides of benzo(a)pyrene and cyclopenta(cd)pyrene induce base substitutions via specific transversions. *Proc. Natl Acad. Sci. USA*, **79**, 1945–1949.
25. Mazur,M. and Glickman,B. (1988) Sequence specificity of mutations induced by benzo(a)pyrene-7,8-diol-9,10-epoxide at endogenous APRT gene in CHO cells. *Somat. Cell Mol. Genet.*, **14**, 393–400.
26. Chen,R.-H., Maher,V.M. and McCormick,J.J. (1990) Effect of excision repair by diploid human fibroblasts on the kinds and locations of mutations induced by (±)-7b,8a-dihydroxy-9a,10a-epoxy-7,8,9,10-tetrahydrobenzo[a] pyrene in the coding region of the *HPRT* gene. *Proc. Natl Acad. Sci. USA*, **87**, 8680–8684.
27. Wei,S.J., Chang,R.L., Bhachech,N., Cui,X.X., Merkler,K.A., Wong,C.Q., Hennig,E., Yagi,H., Jerina,D.M. and Conney,A.H. (1993) Dose-dependent differences in the profile of mutations induced by (+)-7R,8S-dihydroxy-9S,10R-epoxy-7,8,9,10-tetrahydrobenzo(a)pyrene in the coding region of the hypoxanthine (guanine) phosphoribosyltransferase gene in Chinese hamster V-79 cells. *Cancer Res.*, **53**, 3294–3301.
28. Ruggeri,B., DiRado,M., Zhang,S.Y., Bauer,B., Goodrow,T. and Klein-Szanto,A.J.P. (1993) Benzo[a]pyrene-induced murine skin tumors exhibit frequent and characteristic G→T mutations in the p53 gene. *Proc. Natl Acad. Sci. USA*, **90**, 1013–1017.
29. Alexandrov,K., Rojas,M., Geneste,O., Castegnaro,M., Camus,A.M., Petruzzelli,S., Giuntini,C. and Bartsch,H. (1992) An improved fluorometric assay for dosimetry of benzo(a)pyrene diol- epoxide-DNA adducts in smokers' lung: comparisons with total bulky adducts and aryl hydrocarbon hydroxylase activity. *Cancer Res.*, **52**, 6248–6253.
30. Kato,S., Bowman,E.D., Harrington,A.M., Blomeke,B. and Shields,P.G. (1995) Human lung carcinogen-DNA adduct levels mediated by genetic polymorphisms in vivo. *J. Natl Cancer Inst.*, **87**, 902–907.
31. Andreassen,A., Kure,E.H., Nielsen,P.S., Autrup,H. and Haugen,A. (1996)

Comparative synchronous fluorescence spectrophotometry and 32P-postlabeling analysis of PAH-DNA adducts in human lung and the relationship to TP53 mutations. *Mutat. Res.*, **368**, 275–282.
32. Kriek,E., Rojas,M., Alexandrov,K. and Bartsch,H. (1998) Polycyclic aromatic hydrocarbon-DNA adducts in humans: relevance as biomarkers for exposure and cancer risk. *Mutation Res.*, **400**, 215–231.
33. Besarati Nia,A., Van Straaten,H.W., Godschalk,R.W., Van Zandwijk,N., Balm,A.J., Kleinjans,J.C. and Van Schooten,F.J. (2000) Immunoperoxidase detection of polycyclic aromatic hydrocarbon-DNA adducts in mouth floor and buccal mucosa cells of smokers and non-smokers. *Environ. Mol. Mutagen.*, **36**, 127–133.
34. Chen,R.-H., Maher,V.M., Brouwer,J., van de Putte,P. and McCormick,J.J. (1992) Preferential repair and strand-specific repair of benzo(a)pyrene diol epoxide adducts in the *HPRT* gene of diploid human fibroblasts. *Proc. Natl Acad. Sci. USA*, **89**, 5413–5417.
35. Denissenko,M.F., Pao,A., Pfeifer,G.P. and Tang,M. (1998) Slow repair of bulky DNA adducts along the nontranscribed strand of the human *p53* gene may explain the strand bias of transversion mutations in cancers. *Oncogene*, **16**, 1241–1247.
36. Wijnhoven,S.W., Kool,H.J., van Oostrom,C.T., Beems,R.B., Mullenders,L.H., van Zeeland,A.A., van der Horst,G.T., Vrieling,H. and van Steeg,H. (2000) The relationship between benzo[a]pyrene-induced mutagenesis and carcinogenesis in repair-deficient Cockayne syndrome group B mice. *Cancer Res.*, **60**, 5681–5687.
37. Denissenko,M.F., Pao,A., Tang,M. and Pfeifer,G.P. (1996) Preferential formation of benzo[a]pyrene adducts at lung cancer mutational hotspots in *P53*. *Science*, **274**, 430–432.
38. Smith,L.E., Denissenko,M.F., Bennett,W.P., Li,H., Amin,S., Tang,M. and Pfeifer,G.P. (2000) Targeting of lung cancer mutational hotspots by polycyclic aromatic hydrocarbons. *J. Natl Cancer Inst.*, **92**, 803–811.
39. Rodin,S.N. and Rodin,A.S. (2000) Human lung cancer and p53: The interplay between mutagenesis and selection. *Proc. Natl Acad. Sci. USA*, **97**, 12244–12249.
40. Larkin,M. (2000) More to smoking and lung cancer than meets the eye? *Lancet*, **356**, 1499.
41. Paschke,T. (2000) Analysis of different versions of the IARC p53 database with respect to G→T transversion mutation frequencies and mutation hotspots in lung cancer of smokers and non-smokers. *Mutagenesis*, **15**, 457–458.
42. Hernandez-Boussard,T.M., Montesano,R. and Hainaut,P. (1999) Sources of bias in the detection and reporting of p53 mutations in human cancer: analysis of the IARC p53 mutation database. *Genet. Anal.*, **14**, 229–233.
43. Gao,H.G., Chen,J.K., Stewart,J., Song,B., Rayappa,C., Whong,W.Z. and Ong,T. (1997) Distribution of p53 and *K-ras* mutations in human lung cancer tissues. *Carcinogenesis*, **18**, 473–478.
44. Takeshima,Y., Seyama,T., Bennett,W.P., Akiyama,M., Tokuoka,S., Inai,K., Mabuchi,K., Land,C.E. and Harris,C.C. (1993) p53 mutations in lung cancers from non-smoking atomic-bomb survivors. *Lancet*, **342**, 1520–1521.
45. Isobe,T., Hiyama,K., Yoshida,Y., Fujiwara,Y. and Yamakido,M. (1994) Prognostic significance of p53 and ras gene abnormalities in lung adenocarcinoma patients with stage I disease after curative resection. *Jpn J. Cancer Res.*, **85**, 1240–1246.
46. Lee,L.N., Shew,J.Y., Sheu,J.C. *et al.* (1994) Exon 8 mutation of p53 gene associated with nodal metastasis in non-small-cell lung cancer. *Am. J. Respir. Crit. Care Med.*, **150**, 1667–1671.
47. Law,J.C., Whiteside,T.L., Gollin,S.M., Weissfeld,J., El-Ashmawy,L., Srivastava,S., Landreneau,R.J., Johnson,J.T. and Ferrell,R.E. (1995) Variation of p53 mutational spectra between carcinoma of the upper and lower respiratory tract. *Clin. Cancer Res.*, **1**, 763–768.
48. Takagi,Y., Koo,L.C., Osada,H. *et al.* (1995) Distinct mutational spectrum of the p53 gene in lung cancers from Chinese women in Hong Kong. *Cancer Res.*, **55**, 5354–5357.
49. Lung,M.L., Wong,M.P., Skaanild,M.T., Fok,C.L., Lam,W.K. and Yew,W.W. (1996) p53 mutations in non-small cell lung carcinomas in Hong Kong. *Chest*, **109**, 718–726.
50. Takagi,Y., Osada,H., Kuroishi,T. *et al.* (1998) p53 mutations in non-small-cell lung cancers occuring in individuals without a past history of active smoking. *Br. J. Cancer*, **77**, 1568–1572.
51. Hojo,S., Fujita,J., Yamadori,I. *et al.* (1998) Heterogeneous point mutations of the p53 gene in pulmonary fibrosis. *Eur. Respir. J.*, **12**, 1404–1408.
52. Wang,Y.C., Chen,C.Y., Chen,S.K., Cherng,S.H., Ho,W.L. and Lee,H. (1998) High frequency of deletion mutations in p53 gene from squamous cell lung cancer patients in Taiwan. *Cancer Res.*, **58**, 328–333.
53. Huang,C.L., Taki,T., Adachi,M., Konishi,T., Higashiyama,M., Kinoshita,M., Hadama,T. and Miyake,M. (1998) Mutations of p53 and K-ras genes as prognostic factors for non-small cell lung cancer. *Int. J. Oncol.*, **12**, 553–563.
54. Marchetti,A., Pellegrini,S., Sozzi,G. *et al.* (1998) Genetic analysis of lung tumours of non-smoking subjects: p53 gene mutations are constantly associated with loss of heterozygosity at the FHIT locus. *Br. J. Cancer*, **78**, 73–78.
55. Miyake,M., Adachi,M., Huang,C., Higashiyama,M., Kodama,K. and Taki,T. (1999) A novel molecular staging protocol for non-small cell lung cancer. *Oncogene*, **18**, 2397–2404.
56. Gealy,R., Zhang,L., Siegfried,J.M., Luketich,J.D. and Keohavong,P. (1999) Comparison of mutations in the p53 and K-ras genes in lung carcinomas from smoking and nonsmoking women. *Cancer Epidemiol. Biomarkers Prev.*, **8**, 297–302.
57. Yngveson,A., Williams,C., Hjerpe,A., Lundeberg,J., Soderkvist,P. and Pershagen,G. (1999) p53 Mutations in lung cancer associated with residential radon exposure. *Cancer Epidemiol. Biomarkers Prev.*, **8**, 433–438.
58. Tokuchi,Y., Hashimoto,T., Kobayashi,Y. *et al.* (1999) The expression of p73 is increased in lung cancer, independent of p53 gene alteration. *Br. J. Cancer*, **80**, 1623–1629.
59. Husgafvel-Pursiainen,K., Boffetta,P., Kannio,A., Nyberg,F., Pershagen,G., Mukeria,A., Constantinescu,V., Fortes,C. and Benhamou,S. (2000) p53 mutations and exposure to environmental tobacco smoke in a multicenter study on lung cancer. *Cancer Res.*, **60**, 2906–2911.
60. Hussain,S.P. and Harris,C.C. (1999) p53 mutation spectrum and load: the generation of hypotheses linking the exposure of endogenous or exogenous carcinogens to human cancer. *Mutat. Res.*, **428**, 23–32.
61. Rusin,M., Butkiewicz,D., Malusecka,E. *et al.* (1999) Molecular epidemiological study of non-small-cell lung cancer from an environmentally polluted region of Poland. *Br. J. Cancer*, **80**, 1445–1452.
62. Alexandrov,K., Rojas,M., Kadlubar,F.F., Lang,N.P. and Bartsch,H. (1996) Evidence of anti-benzo[a]pyrene diolepoxide-DNA adduct formation in human colon mucosa. *Carcinogenesis*, **17**, 2081–2083.
63. Godschalk,R.W., Moonen,E.J., Schilderman,P.A., Broekmans,W.M., Kleinjans,J.C. and Van Schooten,F.J. (2000) Exposure-route-dependent DNA adduct formation by polycyclic aromatic hydrocarbons. *Carcinogenesis*, **21**, 87–92.
64. Biggs,P.J., Warren,W., Venitt,S. and Stratton,M.R. (1993) Does a genotoxic carcinogen contribute to human breast cancer? The value of mutational spectra in unravelling the aetiology of cancer. *Mutagenesis*, **8**, 275–283.
65. Rundle,A., Tang,D., Hibshoosh,H., Estabrook,A., Schnabel,F., Cao,W., Grumet,S. and Perera,F.P. (2000) The relationship between genetic damage from polycyclic aromatic hydrocarbons in breast tissue and breast cancer. *Carcinogenesis*, **21**, 1281–1289.
66. Williams,J.A. and Phillips,D.H. (2000) Mammary expression of xenobiotic metabolizing enzymes and their potential role in breast cancer. *Cancer Res.*, **60**, 4667–4677.
67. Denissenko,M.F., Chen,J.X., Tang,M.S. and Pfeifer,G.P. (1997) Cytosine methylation determines hot spots of DNA damage in the human P53 gene. *Proc. Natl Acad. Sci. USA*, **94**, 3893–3898.
68. Cho,Y., Gorina,S., Jeffrey,P. and Pavletich,N.P. (1994) Crystal structure of a p53 tumor suppressor-DNA complex: a framework for understanding how mutations inactivate p53. *Science*, **265**, 346–355.
69. Walker,D.R., Bond,J.P., Tarone,R.E., Harris,C.C., Makalowski,W., Boguski,M.S. and Greenblatt,M.S. (1999) Evolutionary conservation and somatic mutation hotspot maps of p53: correlation with p53 protein structural and functional features. *Oncogene*, **18**, 211–218.
70. Shukla,R., Liu,T., Geacintov,N.E. and Loechler,E.L. (1997) The major, N2-dG adduct of (+)-anti-B[a]PDE shows a dramatically different mutagenic specificity (predominantly, G→A) in a 5′-CGT-3′ sequence context. *Biochemistry*, **36**, 10256–10261.
71. Rodin,S.N., Holmquist,G.P. and Rodin,A.S. (1998) CpG transition strand asymmetry and hitch-hiking mutations as measures of tumorigenic selection in shaping the p53 mutation spectrum. *Int. J. Mol. Med.*, **1**, 191–199.
72. Rodin,S.N. and Rodin,A.S. (1998) Strand asymmetry of CpG transitions as indicator of G1 phase-dependent origin of multiple tumorigenic p53 mutations in stem cells. *Proc. Natl Acad. Sci. USA*, **95**, 11927–11932.
73. Strauss,B.S. (1998) Hypermutability in carcinogenesis. *Genetics*, **148**, 1619–1626.
74. Strauss,B.S. (1997) Silent and multiple mutations in p53 and the question of the hypermutability of tumors [published erratum appears in *Carcinogenesis* (1998) **19**, 237]. *Carcinogenesis*, **18**, 1445–1452.
75. Carbone,D., Chiba,I. and Mitsudomi,T. (1991) Polymorphism at codon 213 within the p53 gene. *Oncogene*, **6**, 1691–1692.

Received November 27, 2000; accepted January 2, 2001