LETTER TO THE EDITORS

TP53 mutation spectrum in lung cancers and mutagenic signature of components of tobacco smoke: lessons from the IARC TP53 mutation database

Pierre Hainaut2, Magali Olivier and Gerd P. Pfeifer1

1 Department of Biology, Beckman Research Institute of the City of Hope, Duarte, CA 91010, USA
2 To whom correspondence should be addressed. Tel: +33 4 72738532; Fax: +33 4 72738322; Email: hainaut@iarc.fr

A database of all published TP53 mutations in human cancer is maintained at the International Agency for Research on Cancer (IARC). In lung cancers, TP53 mutation patterns show an exceptionally high prevalence of G→T transversions, mostly occurring at codons demonstrated to be sites of adduction of metabolites of polycyclic aromatic hydrocarbons such as benzo[a]pyrene, one of the major carcinogens of tobacco smoke. These observations have been challenged in a recent ‘Discussion Forum’ by T. Paschke, who claimed that a large number of discrepancies existed in the classification of smoking status between successive releases of the IARC TP53 mutation database and that no statistically significant differences could be found in G→T transversion frequencies between smoking and non-smoking lung cancer patients. In the present Letter we question the methods and the conclusions of the analysis presented by Paschke. Based on an assessment of all published data, we confirm the existence of a highly significant difference in the prevalence of G→T transversions between smoking and non-smoking lung cancer patients.

Missense TP53 mutations are common in human cancers. These mutations are scattered over many different codons and differ by their position and their chemical nature. This variability has allowed one to draw tumour-specific mutation patterns which, in some instances, are consistent with mutagenic mechanisms thought to be involved in the aetiology of cancer (for a recent review see Hainaut and Hollstein, 2000). In 1991 a database of published TP53 mutations was established by C.C. Harris and M. Hollstein in order to facilitate the retrieval and analysis of TP53 mutations. Since 1994 this database has been maintained at the International Agency for Research on Cancer (IARC) and is made freely available as a service to the scientific community (http://www.iarc.fr/p53/Index.html).

In lung cancers a large fraction of TP53 mutations (>30%) are G→T transversions, a type of mutation which is infrequent in tumours other than lung cancers (<12%, aside from hepatocellular carcinoma linked with exposure to aflatoxins) as well as in lung cancers from non-smoking patients (~10%) (Greenblatt et al., 1994; Bennett et al., 1999; Hussain et al., 1999). Previous studies have shown that mutations occur at bases known to be sites of formation of polycyclic aromatic hydrocarbon (PAH) adducts in the coding sequence of TP53. Codons 157, 248 and 273, the three strongest binding codons for benzo[a]pyrene diol epoxide adducts in vitro, contain 21% of all G→T mutations in lung cancers, versus <10% in all other cancers (Denissenko et al., 1996, 1998; Hernandez-Boussard and Hainaut, 1998; Smith et al., 2000). These observations support the notion that patterns of G→T transversions in lung cancers reflect the primary mutagenic signature of DNA damage inflicted by components of tobacco smoke.

The origin of TP53 mutations in lung cancers has recently been questioned in two publications claiming that there were no significant differences in the prevalence of G→T transversions in lung cancers from smoking and non-smoking patients. The first of these reports was published in Proceedings of the National Academy of Sciences (Rodin and Rodin, 2000) and prompted us to reassess the evidence available on TP53 mutations in smokers and non-smokers. This reassessment, which has been published recently, has fully confirmed our previous conclusions (Hainaut and Pfeifer, 2001). The second of these reports, by Paschke, appeared as a ‘Discussion Forum’ in Mutagenesis (Paschke, 2000). Paschke has compared three successive versions of the IARC TP53 database released between 1997 and 1999 (versions R1, R2 and R3) and has identified a large number of discrepancies in the classification of smoking and non-smoking status. In addition, he did not find any statistically significant difference between smoking and non-smoking lung cancer patients with respect to G→T transversion mutation frequencies or in mutational hotspots at codons 157, 248 and 273 using the R3 version (1999) of the database. In conclusion, he raised the hypothesis that previously reported differences might result from the influence of confounding factors such as ethnicity on TP53 mutation spectra. We strongly disagree with these conclusions and with the way Paschke has used the contents of the database.

First, on the issue of G→T transversions in lung cancers, the literature referenced in PubMed (National Library of Medicine, Bethesda, MA) demonstrates a highly significant difference between smokers and non-smokers. The current version of the IARC TP53 mutation database reflects the literature published until December 2000 and contains a total of 1697 mutations in primary lung cancers, including 349 in ascertained smokers and 99 in ascertained non-smokers. The prevalence of G→T transversions in smokers is 29%, compared with 10% in non-smokers (P < 0.0001, χ2 test). In tumours not related to tobacco smoke the prevalence of G→T transversions varies between 8 and 12% (Hainaut and Pfeifer, 2001). As far as ‘hotspot’ mutations are concerned, Paschke (2000) fails to mention that the specificity of mutational hotspots in lung cancers is not at the codon level, but at the base level (Denissenko et al., 1996; Hernandez-Boussard and Hainaut,
1998; Hainaut and Pfeifer, 2001). This is particularly important when analysing mutations at codons 248 and 273, since these codons are mutation hotspots in many tumours other than lung cancers. In lung cancers a large proportion of these mutations are G→T transversions, contrasting with C→T transitions in cancers not related to tobacco smoke (Figure 1). This fact is overlooked by Paschke (2000). Overall, the data currently available fully confirm and even extend our previous observations, supporting the notion that G→T transversions in lung cancers of smokers reflect the primary signature of DNA damage by PAH from tobacco smoke.

Second, it should be noted that Paschke has used the data compiled in the IARC TP53 mutation database in a manner which is against our published recommendations on the nature of the database and the limitations to its exploitation (Hernandez-Boussard et al., 1999). The objective of the database is to compile and compare all mutations reported in the peer reviewed literature, in lung cancers as well as in all other types of cancers. This means that we include all primary descriptions of mutations, irrespective of our opinion on the quality or significance of the publication. In addition to mutation description, we provide a number of annotations derived from the indications given in the publications (as for example smoking status, alcohol consumption, exposure to environmental carcinogens or identified genetic predispositions). These annotations are given solely to allow database users to retrieve specific subsets of data. Any detailed analysis should be based on a critical re-assessment of the primary publications. In his Discussion Forum Paschke does not provide the references for the publications he took into account to compile his data on non-smokers. This is in contrast to our own recent review published in Carcinogenesis (Hainaut and Pfeifer, 2001), which contains an exhaustive list of all relevant papers published before December 2000.

In the case of lung cancer and tobacco smoking, the dataset available in the literature is affected by two important biases. First, a paper by Gao et al. (1997) contains 107 mutations, including 59 mutations in non-smokers, out of a total of 18 tumours from China, with up to 14 distinct point mutations in the TP53 gene in a single tumour. This finding is highly unusual and is most likely to be due to laboratory artifacts. Second, three papers report a total of 114 mutations in tumours from Japan (Huang et al., 1998a; Miyake et al., 1999; Konishi et al., 2000) which had already been published previously by the same group of authors (Huang et al., 1998b). As the subsequent reports do not indicate that the mutations were published before, they meet our (published) inclusion criteria and were therefore included in the IARC TP53 database. Since we do not know which references have been used by Paschke, indiscriminate inclusion of mutations in his dataset may partially explain what he sees as ‘discrepancies’.

Acknowledgements

The IARC TP53 mutation database is partially supported by a grant from the European Community (QLG-1999-00273). G.P. was supported by a grant from the National Cancer Institute (CA84469).

References


Hussain,S.P. and Harris,C.C. (1999) p53 mutation spectrum and load: the


Received on January 18, 2001; accepted on August 17, 2001