Nucleic Acids Research, 1998, Vol. 26, No. 1 269–270

APC gene: database of germline and somatic mutations in human tumors and cell lines

Pierre Laurent-Puig*, Christophe Béroud[1] and Thierry Soussi[2]

Hôpital Laennec Service de Chirurgie Digestive Générale et Oncologique, 42 Rue de Sevres, 75007 Paris, France, [1]Hôpital Necker Enfants Malades, U383 INSERM, Paris, France and [2]UMR 218 du CNRS, Institut Curie, Pavillon Trouillet Rossignol, 26 rue d'Ulm, 75231 Paris cedex 05, France

Received October 20, 1997; Accepted October 24, 1997

ABSTRACT

A database (http://perso.curie.fr/tsoussi) is described, in which over 1000 mutations in the human APC gene of tumors (colon cancer predominantly) are compiled from the literature. It includes both molecular information about the mutations and clinical data about the patients. Software has been designed to analyse all this information in the database.

INTRODUCTION

Familial adenomatous polyposis (FAP) is an autosomal-dominant precancerous condition characterized by the appearance of hundreds to thousands of adenomatous polyps throughout the entire colorectum. The disease is caused by a germline mutation in the tumor suppressor gene APC localized on chromosome 5q21–22 (1,2). Furthermore, somatic mutation of the APC gene has been identified in sporadic colorectal cancer as well as in some cancer of the stomach, pancreas, thyroid, ovary and other primary sites. As shown below, >98% of APC mutations are either frameshift or nonsense mutations leading to the synthesis of a truncated protein.

Analysis of the APC mutation does not lead to the possibility to perform accurate epidemiological studies as those done with the p53 gene. This is mainly due to the high rate of frameshift mutations that are necessary for the inactivation of APC function. Nevertheless, the analysis of APC mutations indicates that germline and somatic mutations are not alike. The distribution of somatic APC mutations has led to the discovery of a mutation cluster region (MCR) whereas germline mutations are scattered throughout in the first half of the coding region (3). Another interesting observation come from the analysis of the frequency of 1 bp deletion (Table 1). This observation, originally detected by Marshall *et al*. on a small number of patients, is now confirmed with a larger series (4). The origin of this difference is unknown, but it suggests that different mechanisms are involved in these two different cell types.

Table 1. One base deletions are more frequent in tumors

	del 1 bp	del 2 bp
Somatic mutation	112/522 (21.5%)	65/522 (12.45%)
Germline mutation	24/371 (6.5%)	36/371 (7%)
	$P = 0.0001$	$P = 0.20$

CORRELATION BETWEEN GENOTYPE–PHENOTYPE

It is well known that, although all FAP patients have mutations of the same gene and all mutations of the APC gene lead to a C-terminal truncated protein, all FAP patients do not develop uniform disease. It has clearly been shown that patients with identical mutations can develop different clinical features (5). However, some correlations between the location of APC mutation on the gene and clinical manifestations were found. Congenital hypertrophy of the retinal pigment epithelium (called CHRPE) are associated with truncating mutations between codons 463 and 1387 (6,7). An increased of extracolonic manifestations such as desmoid tumors and mandibular lesions were observed with mutations located between codons 1403 and 1578 (8). Some studies have suggested that mutations between codons 1250 and 1464 are associated with an increased number of colorectal tumors (3). On the contrary, an attenuated phenotype of FAP disease was found to be associated with mutations at the 5′ end of the gene. Although only few mutations in the 3′ half of the gene have been published, they all have been associated with a mild phenotype (9–13). No truncated polypeptides were detected in the cells of patients in these families. Van der Luijt *et al*. (1996) speculated that these 3-prime mutations represent null alleles and that the AAPC phenotype is the result of a dosage effect. Patients with such mutations make less than a critical quantity of the normal gene product. The 5′ border of this second region of attenuated FAP could be assigned at codon 1597. It is interesting to note that APC gene mutations in this region are associated with desmoid tumors and absent CHRPE. No correlation has been found between the occurrence of rare manifestations of the disease (i.e. brain tumors, hepatoblastomas and thyroid cancers) and specific APC mutations, suggesting the

To whom correspondence should be addressed. Tel: +33 1 42 86 20 81; Fax: +33 1 42 86 20 72; Email: pierre.laurent–puig@biomedicale.univ–paris5.fr

existence of environmental factors or the existence of modifying genes responsible for the occurrence of these tumors.

Recently, Laken *et al.* (1997) described a novel mechanism for predisposition to colorectal cancer involving a specific polymorphism in the APC gene (14). The change converted the sequence AAATAAAA to (A)8 at codon 1307 (I1307K). Rather than altering the function of the encoded protein, this polymorphism creates a small hypermutable region of the gene leading to various mutational events around this position: insertion or deletion of 1 bp, point mutation leading to a stop codon. They found the polymorphism in 6% of Ashkenazi Jews and ~28% of Ashkenazim with a family history of CRC.

More than 700 somatic mutations have been reported to date in different tumor types. Most of these mutations lead to truncation of the APC protein either by a nonsense mutation (34%) or by frameshift mutation (62%). One exception has to be noted, the nature of the APC somatic mutations observed in hepatoblastoma was unusual, 9 of the 10 mutations were missense, with only 1 case featuring a frameshift mutation due to an insertion (15). Although >90% of the somatic mutations of APC gene were observed in colorectal neoplasm, APC alterations seem to be implicated in tumorigenesis of other cancers such as esophageal (16,17) or gastric (18,19).

DESCRIPTION OF THE DATABASE AND FORMAT

As of October 1997, the database contained >1000 records of mutations—either germline, somatic, or from cell lines. Splice mutations have been omitted. If the same mutation was reported in more than one article, only the first report is taken into account. Most of the mutations described in this database originate either from FAP patients or from colorectal cancer. For germline mutations in FAP patients, each record corresponds to only one patients in one family. Relatives in the family for whom the same mutation has been observed were not recorded. For somatic mutation in colorectal cancer, it has been described by several authors that different mutations could occur in different carcinomas in a single patient (either in different dysplasia, adenomas or carcinomas). For each mutation a single record has been entered corresponding then to a single mutational event. Somatic mutations in the APC gene have been described in other neoplasm and have been also included in the database but their number is too low for statistical analysis. More than 20 polymorphisms have been identified in the APC gene and have been compiled by Nagase *et al.* They have not been taken into account in the database.

AVAILABILITY

The program was developed with the 4th Dimension (4D) package from ACI (version 6.0). This version generates a compiled program that can be used either on a MAC (680xx or PowerMac) or an IBM computer (486 or Pentium). Furthermore, the compilation integrates the runtime of 4th Dimension, alleviating the need for any other software. This stand-alone software has a size of 5 Mo. The APC database has a size of 3 Mo. The software and the databases are freely available. They can be requested from T.Soussi (thierry.soussi@curie.fr) or C.Béroud (béroud@ceylan.necker.fr). Ten formatted floppy disks are necessary for sending the full version of the database and the software. A solid support such as a Syquest cartridge (44 Mo, 88 Mo or 200 Mo), Jaz cartridge (1 Go) or a CD can also be handled.

The APC database is also available on our Web site at the following address: http://perso.curie.fr/tsoussi . Data have been saved as Microsoft Excel files and can be used either with a Mac or an IBM. All databases are also available by mail (two formatted floppy disks are necessary).

ACKNOWLEDGEMENTS

This work was supported by grants from the Association de Recherche sur le Cancer, the Ligue Nationale contre le Cancer (Comité national and Comité de Paris), the Mutuelle Générale de l'Education Nationale and the Fondation pour la Recherche Médicale.

REFERENCES

1 Polakis,P. (1997) *Biochim. Biophys. Acta.*, **1332**, F127–F147.
2 Kinzler,K.W. and Vogelstein,B. (1996) *Cell*, **87**, 159–170.
3 Nagase,H. and Nakamura,Y. (1993) *Hum. Mutat.*, **2**, 425–434.
4 Marshall,B., Isidro,G., Carvalhas,R. and Boavida,M. (1997) *Hum. Mutat.*, **9**, 286–288.
5 Giardiello,F.M., Krush,A.J., Petersen,G.M., Booker,S.V., Kerr,M., Tong,L.L. and Hamilton,S.R. (1994) *Gastroenterology*, **106**, 1542–1547.
6 Olschwang,S., Tiret,A., Laurent-Puig,P., Muleris,M., Parc,R. and Thomas,G. (1993) *Cell*, **75**, 959–968.
7 Caspari,R., Olschwang,S., Friedl,W., Mandl,M., Boisson,C., Boker,T., Augustin,A., Kadmon,M., Moslein,G., Thomas,G. *et al.* (1995) *Hum. Mol. Genet.*, **4**, 337–340.
8 Davies,D.R., Armstrong,J.G., Thakker,N., Horner,K., Guy,S.P., Clancy,T., Sloan,P., Blair,V., Dodd,C., Warnes,T.W. *et al.* (1995) *Am. J. Hum. Genet.*, **57**, 1151–1158.
9 Spirio,L., Olschwang,S., Groden,J., Robertson,M., Samowitz,W., Joslyn,G., Gelbert,L., Thliveris,A., Carlson,M., Otterud,B. *et al.* (1993) *Cell*, 75, 951–957.
10 Scott,R.J., van der Luijt,R., Spycher,M., Mary,J.L., Muller,A., Hoppeler,T., Haner,M., Muller,H., Martinoli,S., Brazzola,P.L. *et al.* (1995) *Gut*, **36**, 731–736.
11 van der Luijt,R.B., Meera Khan,P., Vasen,H.F., Breukel,C., Tops,C.M., Scott,R.J. and Fodde,R. (1996) *Hum. Genet.*, **98**, 727–734.
12 Friedl,W., Meuschel,S., Caspari,R., Lamberti,C., Krieger,S., Sengteller,M. and Propping,P. (1996) *Hum. Genet.*, **97**, 579–584.
13 Scott,R.J., Froggatt,N.J., Trembath,R.C., Evans,D.G., Hodgson,S.V. and Maher,E.R. (1996) *Hum. Mol. Genet.*, **5**, 1921–1924.
14 Laken,S.J., Petersen,G.M., Gruber,S.B., Oddoux,C., Ostrer,H., Giardiello,F.M., Hamilton,S.R., Hampel,H., Markowitz,A., Klimstra,D., *et al.* (1997) *Nature Genet.*, **17**, 79–83.
15 Oda,H., Imai,Y., Nakatsuru,Y., Hata,J. and Ishikawa,T. (1996) *Cancer Res.*, **56**, 3320–3323.
16 Fein,M. (1997) *Am. J. Gastroenterol.*, **92**, 359–360.
17 Gonzalez,M.V., Artimez,M.L., Rodrigo,L., Lopez-Larrea,C., Menendez,M.J., Alvarez,V., Perez,R., Fresno,M.F., Perez,M.J., Sampedro,A. and Coto,E. (1997) *J. Clin. Pathol.*, **50**, 212–217.
18 Nakatsuru,S., Yanagisawa,A., Ichii,S., Tahara,E., Kato,Y., Nakamura,Y. and Horii,A. (1992) *Hum. Mol. Genet.*, **1**, 559–563.
19 Tamura,G., Maesawa,C., Suzuki,Y., Tamada,H., Satoh,M., Ogasawara,S., Kashiwaba,M. and Satodate,R. (1994) *Cancer Res.* **54**, 1149–1151.