Metabolic Polymorphisms and Susceptibility to Cancer
W. Ryder
IARC Scientific Publications No. 148
International Agency for Research on Cancer, Lyon, 1999

Chapter 15. CYP1A1

Kaname Kawajiri

CYP1A1 plays an important role in the metabolism of polycyclic hydrocarbons that occur in the environment and several studies suggest that the genetic polymorphism of the gene may play a role in the predisposition to cancer. In order to evaluate the function of CYP1A1 *in vivo* as a host factor determinant of enviromentally-caused cancers in humans, additional investigations are needed involving not only molecular epidemiological approaches in different ethnic populations but also more direct approaches such as the use of gene-targeted mice as a model system.

The *CYP1A1* gene is of critical importance for the metabolism of polycyclic aromatic hydrocarbons (Omura *et al.*, 1993; Guengerich & Shimada, 1991; Kawajiri & Fujii-Kuriyama, 1991). The gene product, aromatic hydrocarbon hydroxylase (AHH), catalyses the first step in the conversion of many environmental carcinogens, such as benzo(a)pyrene in cigarette smoke, to their ultimate DNA-binding, carcinogenic form. Human CYP1A1 protein is composed of 512 amino acid residues, which is smaller by 12 amino acids than its rodent equivalent. In experimental animals, 1A1 is induced in both the liver and extrahepatic tissues by exogenous polycyclic aromatic hydrocarbons, such as benzo(a)pyrene, 3-methylcholanthrene and TCDD. In contrast, 1A1 in humans is considered to function primarily as an extrahepatic enzyme, because considerable amounts of both mRNA and protein can be detected in lung, lymphocytes, and placenta, in contrast to undetectable levels in most human livers examined.

In animal model systems, genetically regulated levels of AHH activity are associated with susceptibility to chemically developed skin or lung carcinomas (Kouri & Nebert, 1977). Marked differences in the metabolism of polycyclic aromatic hydrocarbons and susceptibility to chemically induced cancer among strains of mice can be designated from a genetic polymorphism of Ah receptor (Ema *et al.*, 1994), a regulatory transcription factor of the 1A1 protein. In humans, reports from numerous laboratories have suggested a relationship between higher levels of AHH inducibility in certain tissues, such as peripheral blood lymphocytes, and the incidence of lung, larynx, renal and ureter cancers (Kellermann *et al.*, 1973b; Emery *et al.*, 1978; Gahmberg *et al.*, 1979; Korsgaad *et al.*, 1984; Kouri *et al.*, 1982). The rationale for such an association was that individuals with high AHH inducibility could readily and efficiently activate polycyclic aromatic hydrocarbons in our environment. However, there are reports from other laboratories suggesting a lack of any relationship between high AHH inducibility in human-derived tissues and susceptibility to cancers (Paigen *et al.*, 1977, 1979; Ward *et al.*, 1978; Karki *et al.*, 1987). At present there is no consensus concerning the role of AHH inducibility in the etiology of cancer susceptibility associated with cigarette smoke. (Caporaso *et al.*, 1991; Nebert *et al.*, 1991).

The *CYP1A1* gene has been localized to chromosome 15 near the MPI locus at 15q22-24 (Hildebrand *et al.*, 1985), and several *CYP1A1* RFLP patterns have been reported (Spurr *et al.*, 1987; Bale *et al.*, 1987; Haugen *et al.*, 1990). Clearly, it might be helpful in predicting the individual risk of cancer if some of the genetic polymorphisms were correlated with cancer susceptibility. Three polymorphisms of the *CYP1A1* gene have been studied extensively in relation to cancer susceptiblity (Fig. 1); two genetically-linked polymorphisms, one producing an *Msp* I recognition site in the 3′ non-coding region and the other a one-base substitution of adenine to guanine in the haem-binding region of exon 7 (*Ile-Val* polymorphism), have been associated with an increased risk of smoking-induced lung cancer in Asians but not in Caucasians (Kawajiri *et al.*, 1993). A novel *Msp* I RFLP in the 3′ non-coding region of the *CYP1A1* gene found only in an African-American population (Crofts *et al.*, 1993) has also been studied. The present chapter briefly reviews the current status of the relationship between metabolic polymorphisms of CYP1A1 and cancer susceptibility in human (Tables 1-3).

Table 1. Distribution of CYP1A1 genotypes (Msp I) in cancer patients and healthy controls

Populations		CYP1A1 genotypes			
		A(m1/m1)	B(m1/m2)	C(m2/m2)	Total
Kawajiri 1998, Kawajiri *et al.* 1993 (Japan)	Control	166 (44.3)	169 (45.1)	40 (10.6)	375
	Lung	130 (39.9)	144 (44.1)	52 (16.0)	326
	Sq	38 (36.1)	43 (41.0)	24 (22.9)	105
	Ad	69 (46.1)	66 (44.0)	15 (10.0)	150
Okada *et al.* 1994 (Japan)	Lung	98 (36.7)	124 (46.4)	45 (16.9)	267
	Sq	33 (38.4)	35 (40.7)	18 (20.9)	86
Kihara *et al.* 1995 (Japan)	Control	81 (43.8)	71 (38.4)	33 (17.8)	185
	Lung	36 (37.1)	45 (46.4)	16 (16.5)	97
Tefre *et al.* 1991 (Norway)	Control	167 (78.8)	43 (20.3)	2 (0.9)	212
	Lung	172 (77.8)	47 (21.3)	2 (0.9)	221
	Sq	59 (73.8)	20 (25.0)	1 (1.2)	80
Hirvonen *et al.* 1992 (Finland)	Control	95 (78.5)	24 (19.8)	2 (1.7)	121
	Lung	65 (74.5)	22 (25.5)	0 (0)	87
	Sq	30 (68.2)	14 (31.8)	0 (0)	44
Drakoulis *et al.* 1994 (Germany)	Control	146 (85.4)	25 (14.6)	0 (0)	171
	Lung	119 (83.8)	22 (15.5)	1 (0.7)	296
	Sq	87 (81.3)	8 (16.8)	2 (1.9)	107
Sugimura *et al.* 1994 (Brazil)	Nonblacks				
	Control	56 (62.2)	27 (30.0)	7 (7.8)	90
	Lung	7 (64.8)	25 (28.4)	6 (6.8)	88
	Blacks				
	Control	14 (63.7)	6 (27.2)	2 (9.1)	22
	Lung	12 (54.5)	7 (31.8)	3 (13.7)	22
Shields *et al.* 1993 (USA)	Control	43 (76.8)	11 (19.6)	2 (3.6)	56
	Lung	33 (68.8)	12 (25.0)	3 (6.3)	48
Kawajiri 1998, Kawajiri *et al.* 1993 (Japan)	Control	166 (44.3)	169 (45.1)	40 (10.6)	375
	Stomach	45 (45.3)	50 (48.0)	9 (8.7)	104
	Colorectal	37 (47.5)	32 (41.0)	9 (11.5)	78
	Breast	15 (48.4)	13 (41.9)	3 (9.7)	31
Okada *et al.* 1994 (Japan)	Pancreas	29 (53.7)	21 (38.9)	4 (7.4)	54
Taioli *et al.* 1995b (USA)	African-Americans				
	Control	51 (60.0)	31 (36.5)	3 (3.5)	85
	Breast	7 (33.4)	10 (47.6)	4 (19.0)	21
	Caucasians				
	Control	146 (79.8)	32 (17.5)	5 (2.7)	183
	Breast	22 (77.3)	8 (26.7)	0 (0)	30
Sivaraman *et al.* 1994 (USA)	Control	23 (49.0)	22 (47.0)	2 (4.0)	47
	Colorectal	23 (53.0)	10 (23.0)	10 (23.0)	43

Sq. = squamous cell carcinoma of the lung; Ad. = adenocarcinoma of the lung

Table 2. Distribution of CYP1A1 genotypes (Ile-Val) in cancer patients and healthy controls

Populations		CYP1A1 genotypes			
		Ile/Ile	Ile/Val	Val/Val	Total
Kawajiri 1998, Kawajiri *et al.* 1993 (Japan)	Control	233 (65.1)	108 (30.2)	17 (4.7)	358
	Lung	188 (57.5)	103 (31.5)	36 (11.0)	327
	Sq	61 (58.7)	30 (28.8)	13 (12.5)	104
	Ad	89 (59.0)	48 (31.8)	14 (9.2)	151
Kihara *et al.* 1995 (Japan)	Control	101 (55.5)	70 (38.5)	11 (6.0)	182
	Lung	59 (61.1)	31 (32.6)	5 (5.3)	95
Drakoulis *et al.* 1994 (Germany)	Control	160 (93.6)	11 (6.4)	0 (0)	171
	Lung	125 (88.0)	15 (10.6)	2 (1.4)	142
Hamada *et al.* 1995 (Brazil)	Control	91 (84.3)	15 (13.9)	2 (1.8)	108
	Lung	70 (70.7)	27 (27.3)	2 (2.0)	99
Alexandrie *et al.* 1994 (Sweden)	Control	306 (93.0)	23 (7.0)	0 (0)	329
	Lung	280 (94.6)	16 (5.4)	0 (0)	296
	Sq	98 (91.6)	9 (8.4)	0 (0)	107
Kawajiri 1998, Kawajiri *et al.* 1993 (Japan)	Control	233 (65.1)	108 (30.2)	17 (4.7)	358
	Stomach	54 (56.8)	37 (39.0)	4 (4.2)	95
	Colorectal	59 (69.4)	21 (24.7)	5 (5.9)	85
	Breast	65 (66.5)	29 (29.6)	4 (4.1)	98
Katoh *et al.* 1995 (Japan)	Control	57 (56.4)	39 (38.6)	5 (5.0)	101
	Urothelial	50 (60.2)	30 (36.2)	3 (3.6)	83
Taioli *et al.* 1995b (USA)	African-Americans				
	Control	78 (94.0)	5 (6.0)	0 (0)	83
	Breast	20 (100)	0 (0)	0 (0)	20
	Caucasians				
	Control	145 (82.9)	28 (16.0)	2 (1.1)	175
	Breast	24 (82.8)	5 (17.2)	0 (0)	29
Ambrosone *et al.* 1995 (USA)	Control	195 (85.5)	31 (13.6)	2 (0.9)	228
	Breast	140 (79.5)	32 (18.2)	4 (2.3)	176
Sivaraman *et al.* 1994 (USA)	Control	33 (70.2)	14 (29.8)	0 (0)	47
	Colorectal:	32 (74.4)	9 (20.9)	2 (4.7)	43

Sq. = squamous cell carcinoma of the lung; Ad. = adenocarcinoma of the lung

CYP1A1 phenotype

Measurement

The CYP1A1-dependent phenotype has been determined through assay of the AHH metabolism of benzo(a)pyrene in human-derived tissues, usually peripheral blood lymphocytes (Nebert, 1978). Basal AHH activity in lymphocytes is so weak that a mitogen, such as phytohaemagglutinin or pokeweed, must be introduced into the medium. Further induction with 3-methylcholanthrene, dibenzanthracene or TCDD may be required. AHH activity is measured using benzo(a)pyrene as a substrate by the fluorometric method, and one unit of AHH activity is defined as the amount of enzyme that catalyses the substrate with the formation of fluorescence equivalent to 1 pmol 3-hydroxy-

Table 3. Distribution of CYP1A1 genotypes (AA) in cancer patients and healthy controls

Populations		CYP1A1 genotypes AA	Aa	aa	Total
Taioli *et al.* 1995a (USA)	Control	103 (83.7)	20 (16.3)	0 (0)	123
	Lung	63 (82.9)	12 (14.0)	1 (1.1)	86
	Sq	24 (92.3)	2 (7.7)	0 (0)	26
	Ad	20 (66.7)	9 (30.0)	1 (3.3)	30
Kelsey *et al.* 1994 (USA)	Control	74 (76.3)	21 (21.6)	2 (2.1)	97
	Lung	60 (83.3)	11 (15.3)	1 (1.4)	72
Taioli *et al.* 1995b (USA)	African-Americans				
	Control	72 (83.7)	14 (16.2)	0 (0)	86
	Breast	17 (81.0)	4 (19.0)	0 (0)	21
	Caucasian				
	Control	183 (100)	0 (0)	0 (0)	183
	Breast	30 (100)	0 (0)	0 (0)	30

Sq. = squamous cell carcinoma of the lung; Ad. = adenocarcinoma of the lung

benzo(a)pyrene in 45 min. In lymphocytes, AHH inducibility has been expressed as the ratio of AHH activity in induced cultures to that in uninduced mitogen-stimulated cultures. The viability of lymphocytes can be determined by the assay of NADH-cytochrome c reductase.

Experimental studies of the role of CYP1A1 phenotype in carcinogenesis

The Ah receptor (Ahr) is a ligand-dependent transcription factor that positively regulates inducible expression of AHH activity (CYP1A1 phenotype). In mice the single Ah locus has been found to govern the genetic differences in the inducibility of AHH for metabolic activation of aromatic hydrocarbons among various inbred strains and, therefore, to be related to the susceptibility of the experimental animals to chemical carcinogenesis (Nebert, 1978). It was reported that responder mice, such as C57BL/6, were at high risk of carcinogenesis induced by these carcinogens, while non-responder mice, such as DBA/2J, were relatively resistant. The Ah receptor exhibits considerable functional and structural variability among inbred strains of mice, and this polymorphism is known to arise from multiple genetic alleles at the Ah locus. Non-responsive strain DBA/2J mice possessed Ahr with an affinity for agonist that was 15-20 times lower than that of the receptor in the responsive C57BL/6 strain. Ema *et al.* (1994) showed that two polymorphisms were involved in different ligand affinity between the two strains of mice; 43 extra amino acids at the C-terminal originated by a codon change from Opal (C57BL) to Arg (DBA) and one amino acid replacement at 375 from Ala (C57BL) to Val (DBA) led to low ligand affinity of the Ahr protein in the DBA (K^d = 1.66 nM for TCDD) compared with that of C57BL mice (K^d = 0.27 nM for TCDD).

Epidemiological evidence of the role of CYP1A1 phenotype in carcinogenesis

Kellermann *et al.* (1973a) reported that inducibility of AHH showed a trimodal distribution in cultured, mitogen-stimulated lymphocytes, thereby suggesting a genetic basis for this variation. This was supported by family pedigree analysis, which revealed that the different AHH activities were under genet-

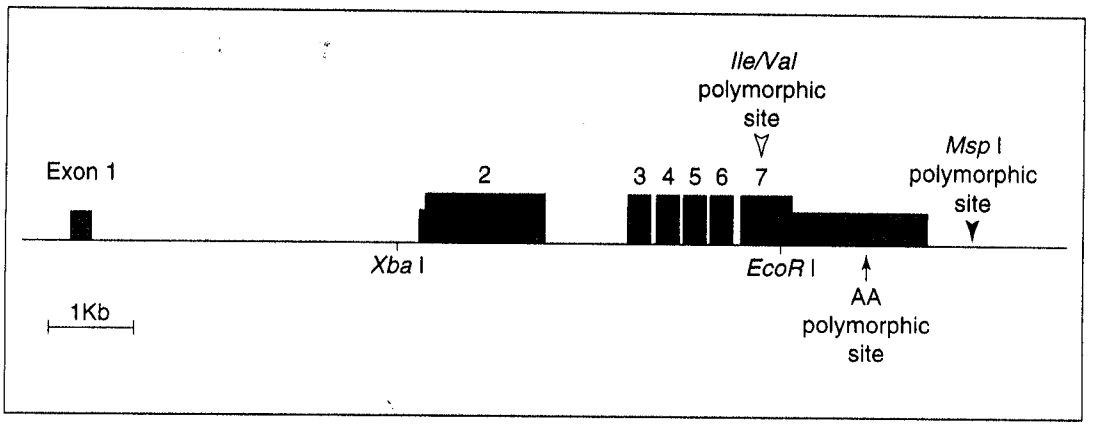

Figure 1. Structural organization and polymorphic sites of the human CYP1A1 gene.

ic control. Reproducibility and trimodal distribution of AHH inducibility in lymphocytes among humans were reconfirmed recently by a strictly developed analysis (Kouri *et al.*, 1982; Trell *et al.*, 1985). On the basis of these results a model was proposed in which the groups with low and high AHH inducibilities are homozygous for two different alleles, and the group with intermediate AHH activity is heterozygous for two alleles. These population studies provided a basis for the hypothesis that susceptibility to smoking-induced cancer may be associated with genetically determined metabolic polymorphism of AHH inducibility.

Kellermann *et al.* (1973a) measured AHH inducibility in mitogen-stimulated lymphocytes from 121 patients with bronchogenic carcinoma and 230 healthy control individuals. The majority of patients were found to be in groups with genetically increased AHH activity (30.6%, 37/121), and only a small proportion had a genetically low AHH activity (5.0%, 6/121). In contrast, high and low AHH groups comprised 10.9% (25/230) and 42.2% (97/230) of the healthy control individuals respectively. In patients with cancer of the larynx the distribution of AHH activity was similar to that of lung cancer patients. It was concluded that individuals with a genetically high AHH activity were at greater risk for developing smoking-associated cancers if they were heavy smokers than were individuals with low AHH activity. Emery *et al.* (1978) showed that the proportion of high AHH inducibility (induced to uninduced AHH activity greater than 4.0) was significantly greater among patients with squamous cell carcinoma (69.4%, 43/62) of the lung than among healthy controls (33.9%, 21/62) matched for age, sex and smoking habits. Gahmberg *et al.* (1979) reported that high induced AHH activity (induced to uninduced AHH activity greater than 1.5) was found in 39% of patients with lung cancer, but in only 15% (61/404) of normal subjects. However, only 17% of the patients with other malignancies had high AHH activity. Korsgaad *et al.* (1984) studied the relationship between AHH inducibility and smoking-associated cancers. In 34 patients with carcinomas of the oral cavity, 41 with larynx cancer and 22 with pulmonary carcinomas there was a highly significant overrepresentation of high AHH inducibility, whereas 30 patients with carcinomas of the renal pelvis and ureter and 46 with urinary bladder carcinomas did not differ significantly in this respect from a control population comprising 92 subjects.

However, an absence of association between AHH inducibility and cancer susceptibility has also been reported. Paigen *et al.* (1977, 1979) conducted two similar studies, using lung (n = 17) and bladder (n = 16) cancer patients. They found that the inducibility of AHH activity in the progeny of bladder cancer patients (n = 53) did not differ from that of their parents and of matched healthy controls (n = 53), whereas half of the lung cancer patients showed a lower AHH inducibility than their progeny (n = 57) and healthy controls (n = 57). Ward *et al.* (1978) reported a hospital-based case-control study of 32 lung cancer patients, 27 larynx cancer patients and 58 controls. The mean AHH inducibility was 3.2 ± 0.02

in patients who had lung cancer, 2.96 ± 0.18 in patients with larynx cancer and 3.29 ± 0.04 for the controls. These values and the distribution of AHH inducibility in pecentiles do not suggest differences between cases and controls. Karki *et al.* (1987) conducted a hospital-based case-control study on AHH inducibility using lymphocytes from 34 patients with pulmonary carcinoma and from 43 non-smoking and 37 smoking controls. The mean inducibility ratio was very similar in all three groups, ranging from 4.5 to 5.5.

A more likely explanation for these contradictory results may relate to the myriad technical difficulties involved in achieving reproducible mitogen activation and subsequent AHH induction in human peripheral blood lymphocytes. Using cryopreserved lymphocytes to improve the assay system, Kouri *et al.* (1982) carried out a case-control study on 51 individuals, 21 of them with lung cancer and 30 with non-malignant pulmonary diseases. All the 14 highest AHH/NADH-cytochrome c reductase levels were in patients with lung cancer. Mean AHH/cytochrome c reductase was 0.89 for lung cancer patients and 0.47 for non-cancer patients. However, whether or not the higher AHH inducibility levels are the cause or the result of primary lung cancer remains to be determined. Although the genetic difference in the susceptibility of mice to chemically induced carcinogenesis is governed by polymorphism of the Ahr, as previously mentioned, no germ line polymorphism of the human Ahr showed a significant association with AHH inducibility or with lung cancer incidence (Kawajiri *et al.*, 1995; Micka *et al.*, 1997).

CYP1A1 genotype

Measurement

(a) Msp I polymorphism in the human CYP1A1 gene

Southern blot analysis

Human lymphocyte DNA (8 μg) is digested with restriction nuclease *Msp* I for 2 hours at 37°C and the products are subjected to electrophoresis in 0.8% agarose for Southern blot analysis (Kawajiri *et al.*, 1990). The DNA fragments are transferred to a nitrocellulose or nylon membrane filter. The filter is hybridized to the ^{32}P-labelled *Xba* I-*Eco* RI fragment of the cloned *CYP1A1* gene in a hybridization solution at 65°C overnight and washed twice with 0.1x SSC containing 0.1% SDS at 65°C for 30 min followed by autoradiography against a Kodak XAR-5 film at -80°C with an intensifying screen. An individual with genotype *A(m1/m1)* is a predominant homozygote, where the *Msp* I site (264th downstream from the poly(A) additional signal) is absent. An individual homozygous for the rare allele is genotype *C (m2/m2)*, derived from a one-base substitution of thymine with cytosine to form the *Msp* I, as confirmed by PCR-direct sequencing. An individual with genotype *B(m1/m2)* is heterozygous for the alleles.

PCR-restriction nuclease digestion analysis

Two synthetic oligonucleotide primers of 21 bases (C47: 5'-CAGTGAAGAGGTGTAGCCGCT-3' and C44: 5'-TAGGAGTCTTGTCTCATGCCT-3' from the 130th to the 150th and from the 449th to the 469th bases respectively, counting from the poly (A) additional signal) are prepared (Hayashi *et al.*, 1991). PCR is carried out with 25 cycles under the following conditions: 1 minute at 95°C for denaturation, 1 minute at 68° C and 1 minute at 72°C for primer annealing and primer extension. The amplified fragments, including the *Msp* I site, are digested with *Msp* I for 2 hours at 37°C and the products subjected to electrophoresis in a 1.8% agarose gel. Genotype *A* is characterized by a 0.34 kb fragment; genotype *B* by 0.14, 0.20 and 0.34 kb; and genotype *C* by 0.14 and 0.20 kb. The genotypes of the *CYP1A1* gene ascribed to the *Msp* I site are identified as restriction fragment length polymorphisms (RFLPs) by the PCR and are in complete agreement with the results of Southern blot analysis.

(b) Ile-Val polymorphism in the human CYP1A1 gene

Difference in one base at position 4889 in the 7th exon of the *CYP1A1* gene was found by PCR direct sequencing (Hayashi *et al.*, 1991). This novel point mutation resulted in the replacement of Ile by Val at residue 462 in the HR2 region, which was well conserved among the P450 families (Gotoh *et al.*, 1983). The RFLP method cannot be used to detect this polymorphism because there is no suitable restriction site. For screening purposes, therefore, allele-specific PCR amplification (Hayashi *et al.*, 1991) and SSCP (Kawajiri *et al.*, 1996b) analyses have been adopted.

Allele-specific PCR amplification method

Two oligonucleotides of 20-mer (1A1A; 5'-GAAGTGTATCGGTGAGACCA-3' and 1A1G; 5'-GAAGTGTATCGGTGAGACCG-3'), both of which contain the polymorphic site at the 3' end, are synthesized and each is used as a primer for allele-specific PCR amplification together with another strand of 21-mer primer (C53; 5'-GTAGACA-GAGTCTAGGCCTCA-3') located about 190 bp downstream of a polymorphic site detected by sequencing (Hayashi *et al.*, 1991). PCR is performed with 30 cycles under the following conditions: 1 minute at 95°C for denaturation, annealing at 65°C for 1 minute, and extension at 72°C for 1 minute. The PCR products are then subjected to electrophoresis in a 1.8% agarose gel.

Single-strand conformational polymorphism (SSCP)

A pair of primers (5'-GAACTGCCACTTCAGCT-GTCT-3' and 5'-GTAGACAGAGTCTAGGCCTCA-3') are used for screening the *Ile-Val* polymorphism by SSCP analysis (Kawajiri *et al.*, 1996). Genomic DNA (50 ng) is used in a 5-µl PCR reaction mixture containing 10 mM Tris-HCl (pH 8.3), 1.5 mM $MgCl_2$, 50 mM KCl, 0.01% gelatin (w/v), 1.25 mM each of four deoxynucleotide triphosphates except for dCTP, which has a concentration of 0.125 mM, 1µM of each primer, and 0.2 µl of [α-^{32}P]dCTP (3000Ci mmole^{-1}) and Taq DNA polymerase. The PCR is programmed as follows: initial denaturation, 1 minute at 95°C; amplification for 20 seconds at 95°C, 2 minutes at 60°C for 30 cycles; elongation for 1 minute at 72°C. After completion of the PCR the product is diluted 1:100 in loading buffer (95% formamide, 20 mM EDTA, 0.05% bromphenol blue and 0.05% xylene cyanol). The DNA fragments are subjected to electrophoresis at 35 W for approximately 3 hours in 5% non-denaturing polyacrylamide gel with 5% glycerol at room temperature. Upon complete migration the gels are dried and subjected to autoradiography against a Kodak XAR-5 film at -80°C with an intensifying screen. The results given by these two methods are fully consistent with those obtained by direct sequencing.

(c) Race-specific AA RFLP in the CYP1A1 gene

A new *Msp* I RFLP in the *CYP1A1* gene has been found in African-Americans but not in Caucasians or Asians (Crofts *et al.*, 1993). The polymorphism results from a single A-T to G-C transition in the 3' non-coding region about 300 bp upstream from the polyadenylation site. For analysis of the AA RFLP the following two 20-mer primers are used: 5'-GGCTGAGCAATCTGACCC-TA-3' and 5'-ATACCCCCCCCTCACTCC-3'. PCR is performed using initial denaturation at 95°C for 4 minutes followed by 40 cycles of 94°C for 20 seconds, 60°C for 20 seconds and 70°C for 20 seconds, with a final extension at 72°C for 10 minutes. This generates a 221-bp fragment, which is then subjected to digestion with *Msp* I. The digested product is visualized on a 2.5% agarose gel; variant alleles are digested into separate fragments of 125 bp and 96 bp. Genotypes AA, Aa and aa represent the predominant homozygous, heterozygous and rare homozygote respectively (London *et al.*, 1995).

Epidemiological evidence of the association between the CYP1A1 genotypes and cancer susceptibility

(a) Relationship between the Msp I or Ile-Val polymorphisms and cancer

Lung cancer

The genetic association between the *Msp* I and *Ile-Val* polymorphisms of the human *CYP1A1* gene was investigated in Japanese (Hayashi *et al.*, 1991), German (Drakoulis *et al.*, 1994) and Finnish (Hirvonen *et al.*, 1992) general populations revealing that these two loci were very closely associated. A recent population study on the *CYP1A1* genotypes was conducted mainly by PCR-nuclease digestion (*Msp* I polymorphism) or allele-specific PCR amplification (*Ile-Val* polymorphism) analyses.

Since first investigating an association of the *Msp* I polymorphism of the *CYP1A1* gene with predisposition to lung cancer (Kawajiri *et al.*, 1990), especially to smoking-associated squamous cell carcinoma, using Southern blot analysis, we have greatly increased the numbers of cases and healthy controls. In our study, lymphocyte DNAs of 2500 Japanese were isolated from a cohort of a general population and used as controls (Nakachi *et al.*, 1991). The genotypes of 375 randomly selected subjects were determined. Genotypes *A*, *B* and *C* were found in 166 (44%), 169 (45%) and 40 (11%) individuals respectively among the healthy con-

trols. This result showed a good fit with Hardy-Weinberg equilibrium, in which the relative frequencies p^2, $2pq$, and q^2 of genotypes estimated from the gene frequencies p and q must be equal to the observed, with a gene frequency of 0.67 for *m1* and 0.33 for *m2* (Kawajiri *et al.*, 1993). However, when 105 patients with squamous cell carcinoma of the lung were analysed the *C* genotype was found in 24 patients (23%), about twice the frequency among the controls. We also looked at the distribution of three genotypes of the *Ile-Val* polymorphism in cancer patients and healthy controls. The genotypes of *Ile/Ile*, *Ile/Val* and *Val/Val* were found in 233 (65%), 108 (30%) and 17 (5%) individuals respectively among 358 controls (Kawajiri *et al.*, 1993). In 104 patients with squamous cell carcinoma of the lung, genotype *Val/Val* was found in 13 patients (13%). In a population of 86 patients with squamous cell carcinoma of the lung, Okada *et al.* (1994) reported that genotype *C* appeared in 21% (18/86), which was twice as high as in the control subjects.

In contrast, the association of adenocarcinoma of the lung with cigarette smoking is less clear than that of squamous cell carcinoma, and the frequencies of genotype *C* and *Val/Val* have not been significantly distinguished from those in healthy populations (Kawajiri *et al.*, 1993). This may be partly ascribed to the heterogeneity of adenocarcinoma of the lung on the basis of the clinical and histopathological characteristics, and the etiologies may be complicated and different among the adenocarcinomas (Suzuki *et al.*, 1990). From this point of view we classified adenocarcinoma of the lung by differentiation grades and examined the involvement of *CYP1A1* polymorphisms (Nakachi *et al.*, 1995). The highest proportion of current or ex-smokers and their cigarette consumption was among the poorly differentiated cases. The two polymorphisms of *CYP1A1* were examined among current or ex-smokers with three differentiated grades, revealing that poorly differentiated adenocarcinoma included significantly elevated frequencies of genotype *C* (26%; 12/46) and *Val/Val* (11%; 5/46), which have been found to be 'susceptible' in squamous cell carcinoma of the lung.

In a study on 221 lung cancer patients and 212 healthy controls in a Norwegian population, Tefre *et al.* (1991) demonstrated the absence of a correlation between the *Msp* I polymorphism and an increased risk of lung cancer. They also found an absence of associations with particular histological types of lung cancer, cigarette smoking history or occupational exposure to asbestos. Hirvonen *et al.* (1992) reported a lack of an association between *Msp* I RFLP and lung cancer risk using 87 lung cancer patients and 121 healthy controls in a Finnish population. A similar lack of association between *Msp* I RFLP and lung cancer incidence was reported by Shields *et al.* (1993), who studied the DNA of 78 subjects who were African-Americans and Caucasians. Drakoulis *et al.* (1994) conducted a case-control study on association between the *Msp* I or *Ile-Val* polymorphism and lung cancer risk in an ethnically homogeneous German population. Although no statistically significant difference was found in the distribution of *Msp* I RFLP between all cell types of lung cancer (n = 142) and controls (n = 171), a trend to overrepresentation of the *m2* allele of the *Msp* I RFLP was observed among 52 squamous cell carcinoma patients. In contrast, the frequency of *Val*-coded allele in lung cancer patients was 2-fold higher than in the control group (OR = 2.16; 95% CI; 0.96-5.11, P = 0.033). They also found that there was a close genetic linkage of the two polymorphisms of the *CYP1A1* gene in the controls, but no linkage was observed among lung cancer patients. In a Brazilian population the *Msp* I polymorphism was not associated with lung cancer susceptibility (110 cases; 112 controls) (Sugimura *et al.*, 1994), while the *Ile-Val* polymorphism was found to be associated with lung cancer (99 cases; 108 controls; OR = 2.26; 95% CI, 1.14-4.47) (Hamada *et al.*, 1995). These results indicate that the site responsible for cancer susceptibility is the *Ile-Val* polymorphism in the catalytic site of CYP1A1.

Other cancer sites

In Japanese people the frequency distributions of the *Msp* I or *Ile-Val* genotypes in patients with other cancer sites, such as stomach (n = 95), colon (n = 85) and breast (n = 98) were the same as in healthy controls (Kawajiri *et al.*, 1993). No association was found with pancreatic cancer (n = 54) (Okada *et al.*, 1994). Taioli *et al.* (1995b) reported an association of the *Msp* I polymorphism with

an increased risk of breast cancer in African-Americans (21 cases; 85 controls; OR = 9.7; 95% CI, 2.0-47.9), but no association in Caucasians (30 cases; 183 controls). Sivaraman *et al.* (1995) reported a positive correlation between the *Msp* I polymorphism and colorectal cancer in Japanese (P = 0.008) and Hawaiian/part-Hawaiian subjects (P<0.001), although the study lacked the capacity to detect a similar association in Caucasians.

(b) Relationship between race-specific AA RFLP and cancer

Lung cancer

In a study on 76 lung cancer patients and 123 healthy controls in an African-American population, Taioli *et al.* (1995a) demonstrated the absence of a correlation between race-specific AA RFLP and an increased risk of all cell types of lung cancer. However, analysis by histological type showed an association between adenocarcinoma (n = 30) of the lung and the AA RFLP genotypes Aa+aa with an odds ratio of 2.6 (95% CI, 1.1-6.3). The mean numbers of packs of cigarettes-year in adenocarcinoma patients with and without at least one variant allele were 5.0 ± 2.5 and 37.2 ± 6.5 (P<0.05) respectively. A lower dose of cigarette smoking is sufficient to exert a carcinogenic effect on the incidence of adenocarcinoma patients carrying the susceptible allele. The absence of an association between AA RFLP and all cell types of lung cancer was also reported by Kelsey *et al.* (1995) (72 cases; 97 controls) and London *et al.* (1995) (144 cases; 230 controls) in an African-American population. Furthermore, London *et al.* did not confirm (in 51 cases) that AA RFLP may be an important risk factor for adenocarcinoma of the lung in the African-American population reported by Taioli *et al.* (1995a).

Other cancer sites

Taioli *et al.* (1995) also examined the role of AA RFLP in susceptibility to breast cancer in African-Americans (21 cases; 86 controls) and Caucasians (30 cases; 86 controls) and found no association in either group.

Discussion

Most P450 metabolism results in the detoxification of a wide variety of xenobiotics, although certain chemicals are activated in this process to electrophilic forms that can damage DNA and sometimes produce carcinogenic transformation of the cells (Omura *et al.*, 1993; Guengerich & Shimada, 1991; Kawajiri & Fujii-Kuriyama, 1991). Since this P450-mediated bioactivation is an initial and obligatory step in chemical carcinogenesis, interindividual variation in the metabolic activity of P450s may influence subsequent steps, including detoxification by Phase II enzymes, the formation of DNA-carcinogen adducts and ultimate cancer consequence.

The genetic difference in phenotypic expression and/or structure of P450 genes is worthy of study in order to explain the interindividual or interracial difference in cancer susceptibility (Kawajiri & Fujii-Kuriyama, 1991; Gonzalez, 1995). Most studies in this area compare the frequency of metabolic phenotypes or genotypes of a P450 enzyme between cancer patients and controls. In metabolic phenotype comparison studies it may be necessary to pay more attention to the development of methods of determining phenotypes, and also to the examination of the possible influences of medication on both patients and controls (especially hospital controls), since some P450s can be induced by chemicals (Omura *et al.*, 1993). Even a simple case-control comparison of phenotype or genotype frequency should examine the selection bias of controls, who must be representative of the general population from which patients are drawn. In this sense, hospital controls are, in general, not desirable because they are a more or less biased population. A large DNA library from a general population may be appropriate as a control pool, and a follow-up study using the DNA library can confirm the results of case-control studies.

A close association of smoking-associated lung cancer incidence with the *Msp* I or *Ile-Val* polymorphisms of the human *CYP1A1* gene was found in Japanese people but not in Caucasians. A major reason for this discrepancy was an ethnic difference in the allelic frequency of these polymorphisms. In *Msp* I polymorphism the frequency of the susceptible allele of *m2* was 0.332 in 375 healthy Japanese people (Kawajiri *et al.*, 1993), while frequencies of *m2* in Norwegian (Telfre *et al.*, 1991), Finnish (Hirvonen *et al.*, 1992) and German (Drakoulis *et al.*, 1994) subjects were 0.115, 0.12 and 0.073 respectively. On

the basis of the Hardy-Weinberg equation these data suggest that the frequency of genotype *C* (*m2/m2*) is 110/1000 among Japanese people and about 6-14/1000 among Caucasians, i.e. almost 8 to 18 times more frequent among Japanese subjects than in Caucasians. Clearly, studies involving a large number of subjects will be required to disclose associations of the *Msp* I and *Ile-Val* polymorphisms with lung cancer in different ethnic populations. It is of interest that there is a moderate risk elevation of heterozygous genotype *B* in Finnish (OR = 1.85; P = 0.125) (Hirvonen *et al.*, 1992) or Norwegian (OR = 1.24) (Telfre *et al.*, 1991) study populations. The possibility exists that in Caucasians the heterozygous genotype *B* and *Ile/Val* contribute substantially to the attributable risk of lung cancer.

As for the relationship between *CYP1A1* genotypes and phenotypic expressions, Petersen *et al.* (1991), using family pedigree analysis, demonstrated that a high CYP1A1 inducible phenotype segregated concordantly with the *m2* allele having the *Msp* I site. We constructed and expressed the cDNA for *Ile*- and *Val*-coded CYP1A1 in yeast cells and compared the catalytic activities (Kawajiri *et al.*, 1993). Both AHH activity and mutagenic activity towards benzo(a)pyrene were studied and we found that the Val-type of CYP1A1 showed a higher AHH activity and mutagenicity than the Ile-type, although the enzyme activities were expressed at low levels. Recently, Zhang *et al.* (1996) reported that benzo(a)pyrene-7,8- and 9,10-dihydrodiol formation were comparable when purified CYP1A1-Ile^{462} catalysed the reconstituted reaction compared with catalysis by CYP1A1-Val^{462}. Kiyohara *et al.* (1996) reported that AHH inducibility was correlated with the *Msp* I polymorphism (P<0.0001) but no association was found for the *Ile-Val* polymorphism (P = 0.509). Age-adjusted AHH inducibility in lymphocytes with genotype *A* (n = 38), *B* (n = 37) and *C* (n = 7) was 4.89 ± 0.36, 4.82 ± 0.29 and 13.61 ± 1.44 respectively. They also found that basal AHH activity in lymphocytes with a homozygous mutant *Val/Val* genotype was significantly higher than that of the Ile-homozygote (P<0.05).

Frequency comparison should be followed by a well-designed case-control study. The relative risk is then estimated for phenotypes or genotypes and levels of exposure to carcinogens. The risk estimate of environmental exposures in different phenotypes or genotypes is important not only in the identification of susceptible individuals but also to clarify whether the observed genetic risk elevation results from an interaction of a susceptible genotype with environmental carcinogens or is determined prior to gene-enviroment interaction. Dose-response relationships between phenotypes or genotypes can provide decisive information on this matter. We investigated the difference in susceptibility to squamous cell carcinoma of the lung in terms of *CYP1A1* genotypes, taking the amount of cigarettes consumed into account (Nakachi *et al.*, 1991, 1993, 1995). Patients with a susceptible homozygous genotype of the *Msp* I or *Ile-Val* polymorphisms contracted the carcinoma after smoking fewer cigarettes than those with other genotypes. A case-control study revealed that individuals with the susceptible *Msp* I or *Ile-Val* genotype were at remarkably high risk, with an odds ratio of 6.55 or 8.46 respectively (95% CI, 2.49-17.24 or 2.48-28.85 respectively), at a low dose level of cigarette smoking. On the other hand this relative susceptibility of genotype *C* or *Val/Val* compared to genotype *A* or *Ile/Ile* decreased about 1.5-fold at the higher cigarette dose level. Although the risk of all genotypes increases at higher dose levels the genetic difference in cancer risk tends to reduce at high dose levels where the environmental influence outweighs genetic predispositions. It has also been reported that a lower cigarette dose is sufficient to exert carcinogenic effects on adenocarcinoma carrying the AA polymorphism (Taioli *et al.*, 1995a).

However, the mechanisms of genetic predisposition to lung cancer remain largely obscure. Lung carcinogenesis seems to start from a clonal expansion of the cells that gained a selective growth advantage through early genetic change in the cells. Thus, genetic predisposing factors to smoking-associated lung cancer, such as *CYP1A1* polymorphisms, may affect the mutational frequencies of target genes in early genetic alterations. In this respect, we examined *p53* mutations in relation to *CYP1A1* polymorphisms, using surgical specimens of 148 non-small-cell lung cancer (NSCLC) patients who were smokers (Kawajiri *et al.*, 1996a). The frequency of *p53*

mutations among heavy smokers was higher than in patients who had never smoked (P<0.01; OR = 3.74; 95% CI, 1.46-9.56). Smokers with susceptible rare homozygous alleles, such as either the *Msp* I or *Ile-Val* polymorphism of the *CYP1A1* gene, exhibit a 4.5-fold (P<0.005; OR, 4.48; 95% CI, 1.64-12.26) or 5.5-fold (P<0.01: OR, 5.52; 95% CI, 1.55-19.64) higher risk of having a mutation of the *p53* gene than those with non-susceptible predominant homozygous alleles of the gene. A recent report showed that susceptibility to hepatocellular carcinoma induced by aflatoxin B_1 was associated with low activity of the detoxification enzyme epoxide hydrolase and *GSTM1* genotypes, resulting in increased *p53* mutation (McGlynn *et al.*, 1995). It has been reported that *p53* mutations in patients with NSCLC were associated not only with the genesis and progression of lung cancer but also with shortened survival as predictors of poor prognosis (Mitsudomi *et al.*, 1993; Horio *et al.*, 1993). We have recently shown that a germ line *Msp* I polymorphism of the *CYP1A1* gene is associated with various clinical parameters responsible for the poorer prognosis in patients with NSCLC; it is also an independent factor indicative of prognostic significance at the non-resectable advanced stage of the disease (Goto *et al.*, 1996).

References

Alexandrie, A.-K., Sundberg, M.I., Seidegard, J., Tornling, G. & Rannug, A. (1994) Genetic susceptibility to lung cancer with special emphasis on CYP1A1 and GSTM1: a study on host factors in relation to age at onset, gender and histological cancer types. *Carcinogenesis*, 15, 1785-1790

Ambrosone, C.B., Freudenheim, J.L., Graham, S., Marshall, J.R., Vena, J.E., Brasure, J.R., Laughlin, R., Nemoto, T., Michalek, A.M., Harrington, A., Ford, T.D. & Shields, P. Cytochrome (1995) P4501A1 and glutathione S-transferase (M1) genetic polymorphisms and postmenopausal breast cancer risk. *Cancer Res.*, 55, 3483-3485

Bale, A.E., Nebert, D.W. & McBride, O.W. (1987) Subchromosomal localization of the dioxin-inducible P_1-450 locus (CYP1) and description of two RFLPs detected with a 3' P_1-450 cDNA probe. *Cytogenet. Cell. Genet.*, 46, 574-575

Caporaso, N., Landi, M.T. & Vineis, P. (1991) Relevance of metabolic polymorphisms to human carcinogenesis: evaluation of epidemiological evidence. *Pharmacogenetics*, 1, 4-19

Crofts, F., Cosma, G.N., Currie, D., Taioli, E., Toniolo, P. & Garte, S. (1993) A novel *CYP1A1* gene polymorphism in African-Americans. *Carcinogenesis*, 14, 1729-1731

Drakoulis, N.D., Cascorbi, I., Brockmöller, C.R. & Roots, G.I. (1994) Polymorphisms in the human *CYP1A1* gene as susceptibility factors for lung cancer: exon-7 mutation (4889 A to G), and a T to C mutation in the 3'-flanking region. *Clin. Investig.*, 72, 240-248

Ema, M., Ohe, N., Suzuki, M., Mimura, J., Sogawa, K., Ikawa, S. & Fujii-Kuriyama, Y. (1994) Dioxin-binding activities of polymorphic forms of mouse and human Ah receptors. *J. Biol. Chem.*, 269, 27337-27343

Emery, A.E.H., Anand, R., Danford, N., Duncan, W. & Paton, L. (1978) Aryl-hydrocarbon-hydroxylase inducibility in patients with cancer. *Lancet*, 1, 470-472

Gahmberg, C.G., Sekki, A., Kosunen, T.U., Holsti, L.R. & Makela, O. (1979) Induction of aryl hydrocarbon hydroxylase activity and pulmonary carcinoma. *Int. J. Cancer*, 23, 302-305

Gonzalez, F.J. (1995) Genetic polymorphism and cancer susceptibility: Fourteenth Sapporo Cancer Seminar. *Cancer Res.*, 55, 710-715

Goto, I., Yoneda, S., Yamamoto, M. & Kawajiri, K. (1996) Prognostic significance of germ line polymorphisms of the *CYP1A1* and glutathione *S*-transferase genes in patients with non-small-cell lung cancer. *Cancer Res.*, 56, 3725-3730

Gotoh, O., Tagashira, Y., Iizuka, T. & Fujii-Kuriyama, Y. (1983) Structural characteristics of cytochrome P-450. Possible location of the haeme-binding cysteine in determined amino-acid sequence. *J. Biochem.*, 93, 807-817