A C4887A Polymorphism in Exon 7 of Human CYPIA1: Population Frequency, Mutation Linkages, and Impact on Lung Cancer Susceptibility

Ingolf Cascorbi, Jürgen Brockmöller, and Ivar Roots

Institute of Clinical Pharmacology, University Clinic Charité, Humboldt University of Berlin, Schumannstrasse 2021, D-10098 Berlin, Germany

ABSTRACT

This study reports a C→A transition at position 4887 in exon 7 of cytochrome P4501A1 (CYPIA1), resulting in a threonine→serine exchange in codon 461. The polymorphism is located directly beside the known codon 462-Ile/Val mutation (m2) near the heme binding region. The C4887A mutation leads to the loss of a BsrDI cleavage site, which allows analysis. No linkage of this mutation, termed m4, with other mutations such as m1 (MspI polymorphism in the 3'-flanking region) or m2 was observed on the same DNA strand. Systematic molecular genetic analyses of mutation linkages revealed that mutation m2 is in strict linkage disequilibrium with m1. To distinguish the different CYPIA1 alleles and genotypes, mutation linkages were considered.

Frequency of the m4-containing allele, termed CYPIA1**4, among 880 unrelated Caucasian individuals was 2.95% (95% confidence limits, 2.21%, 3.86%). m2 was found in 7.73%, and m2 in 2.67% of alleles. No case of African black-specific mutation m3 was detected. The allele frequency of CYPIA1**4 among 157 lung cancer patients was 2.87% (95% confidence limits, 1.32%, 5.37%); it was 2.87% (95% confidence limits, 1.71%, 4.49%) in 314 controls matched by age and sex. Thus, the novel m4-mutation may not represent a susceptibility factor for lung cancer.

INTRODUCTION

Cytochrome P-450 1A1 (CYPIA1) is a key enzyme in carcinogen metabolism. By virtue of its polymorphic regulation, CYPIA1 proved a promising genetic biomarker for susceptibility to certain malignancies, particularly lung cancer (1). The mutation detected first (called m1) was a T→C transition 1194 bp downstream of exon 7, creating a new MspI cleavage site (2, 3; Fig. 1). This mutation was found to be overrepresented among lung cancer patients in Japan (4). Reports in Caucasians could not confirm this finding (5–8), perhaps due to the lower allele frequency of 7.3% (7) compared to 33.2% in Japanese (9).

The A→G transition at nucleotide 4889 (m2) is a rare trait in Caucasians but occurs in about 20% of Japanese (10). The mutated enzyme showed enhanced activity (10). Indeed, m2 was overrepresented in Japanese lung cancer patients (9) and also in a German study (7), whereas only a trend was observed in Finland (11). We reevaluated two apparently homozygous carriers of m2 from our own study (7) with the newly available restriction enzyme BsrDI. This led to the discovery that binding of previously applied primers for allele-specific PCR had been partly diminished. DNA sequencing of these samples uncovered a novel mutation (named m4), 2 bases upstream of the m2 mutation site (Fig. 1) in the heme-binding region of the enzyme.

On the basis of an improved PCR/RFLP methodology, the frequency of all known CYPIA1 mutations, including m4 and the African black-specific m3-mutation (12), was determined among 880 unrelated individuals. Mutations were systematically checked for linkages so that we could suggest a classification of CYPIA1 alleles. Moreover, the role of m4 in lung cancer susceptibility has been evaluated.

SUBJECTS AND METHODS

Subjects. A group of 880 individuals [588 males (median age, 57 years) and 292 females (median age, 61 years)] composed of 304 healthy volunteers and 576 patients recruited from the Departments of Internal Medicine (University Clinic Benjamin Franklin, Free University of Berlin, and University Clinic Virchow, Humboldt University of Berlin), Pneumology (Lungenklinik Heckeshorn), and Urology (Krankenhaus Neußölln) in Berlin, Germany. Most patients were hospitalized because of nonmalignant lung diseases, such as chronic obstructions, because of cardiovascular diseases, or with various urological diagnoses. The previously studied 142 lung cancer patients (7) were supplemented by 15 cases [altogether, 124 males (median age, 66 years) and 33 females (median age, 65 years)]. Each lung cancer patient was matched with two reference patients by sex and age (± 2 years). All patients and healthy volunteers were selected in the years 1991–1995; they gave their informed consent, and the study was approved by the Ethics Committee of the University Clinic Charité. To avoid confounding by ethnicity, only subjects of German extraction were included.

Genotyping Procedure. DNA was extracted from leukocytes by a standard procedure (13). CYPIA1 mutations were characterized by RFLP after PCR. All patients were obtained from TIB Molbiol (Berlin, Germany); PCR reactions were performed with a Perkin-Elmer Applied Biosystems 9600 thermocycler (Weiterstadt, Germany) or with a Biometra Trio thermoblock (Göttingen, Germany). For determination of m1 and m3, an 899-bp DNA fragment was amplified using 1 unit Taq polymerase (Perkin-Elmer Applied Biosystems, Weiterstadt, Germany), 10 μmol/liter of primers M3F 5'-GGCGTGAG-CATTACGACCCTA and P8O 5'-TAGGATCTTGTGCTGACTGC, 0.2 mmol/liter deoxynucleotide triphosphate (Boehringer Mannheim, Mannheim, Germany), and 2.4 mmol/liter MgCl2 in a total volume of 50 μl. PCR conditions were 35 cycles of 0.5 min at 94°C, 1 min at 63°C, and 1 min at 72°C. The PCR product (12.5 μl) was digested with 50 units MspI (New England Biolabs, Schwalbach, Germany), generating smaller fragments in case of the mutation (Table 1). Fragments were evaluated on an ethidium bromide-stained 2.5% agarose gel (Fig. 2). m2 was checked by amplifying a 204-bp fragment with primers M2F 5'-CTGTCCTCCTGGTACAGGAGAC and M2R 5'-TTCCACCCGGTCGACTGATTACGCC (reagent concentrations as above; PCR conditions were 35 cycles of 0.5 min at 94°C, 0.5 min at 63°C, and 0.5 min at 72°C). The product was digested with BsrDI (New England Biolabs; 0.5 units, 65°C overnight). Mutation m4 could be determined from the same 204-bp fragment but using BsaI (New England Biolabs; 2.5 units, 55°C overnight). Both cleavage sites were lost in the case of the mutations and were evaluated on a 3% NuSieve 3:1 agarose gel (Fig. 2).

DNA Sequencing. Two hundred four-bp PCR fragments of exon 7 were processed using a Taq DyeDeoxy terminator sequencing kit (Perkin-Elmer Applied Biosystems). After oligonucleotides were purified from unbound dye terminators by chloroform/phenol extraction, they were analyzed with an Applied Biosystems 373A automated sequencer on a denaturing 6% polyacrylamide gel. Sequences were compared with the published CYPIA1 sequence (14–16) using ABI sequence navigator software.

Mutation Linkage Analysis. Linkage of m1 with m2 was checked after amplification of 2047-bp fragments containing both mutations (Table 2). Cycling conditions were 35 cycles of 0.5 min at 94°C, 1 min at 63°C, and 2 min at 72°C. The PCR product (12.5 μl) was digested first with 0.5 units BsrDI overnight at 65°C and then with 100 units MspI at 37°C, again overnight. Fragments were separated on a 2% agarose gel.

Hypothetical linkage of m1 with m4 was evaluated as follows: 2047-bp fragments were digested with MspI and separated on a 3% agarose gel. The
1841-bp band containing ml was cut out using a sterile scalpel, frozen at \(-20\)°C for 1 h, and thawed at 37°C for 5 min. The gel pieces were transferred to a 0.45-μm filter (UFC3, Millipore, Eschborn, Germany) and eluted as recommended by the manufacturer. The DNA was digested with 5 units BsaI and separated on a 2% agarose gel.

A possible linkage of m2 and m4 was tested as follows: 204-bp PCR fragments were digested with BsrDI followed by ethanol precipitation to remove any enzyme (recognition sites of BsrDI and BsaI overlap by 1 base). After resolubilization, DNA was digested with BsaI (Table 2). Statistics. Expected genotype frequencies were calculated by the Hardy-Weinberg equation \((1 = p^2 + 2pq + q^2)\) from the allele frequencies. Statistical significance of odds ratios was calculated by the exact Fisher’s test. Odds ratios are given with 95% confidence limits.

### RESULTS

### Identification and Frequency of m4

DNA sequencing of samples that were apparently homozogously m2 mutant when using allele-specific primers (two cases from a previous publication; Ref. 7) but heterozygous (m2/wild type) when performing a BsrDI digestion revealed a C→A transversion at nucleotide 4887 in codon 461 \([i.e., 2\]

---

**Table 1** Restriction fragment length patterns of CYP1A1 mutations

<table>
<thead>
<tr>
<th>Mutation</th>
<th>Primers for amplification and binding positions (nt)*</th>
<th>Restriction enzyme</th>
<th>Recognition sequence</th>
<th>Cut position (nt)*</th>
<th>Fragment length (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ml 6235 T→C</td>
<td>M3F (5542-5561)/P80 (6420-6440)</td>
<td>MspI</td>
<td>C'CQ_G</td>
<td>Wild type</td>
<td>899</td>
</tr>
<tr>
<td>m2 4889 A→G</td>
<td>M2F (4739-4763)/M2R (4918-4942)</td>
<td>BsrDI</td>
<td>GCAATG_NN'</td>
<td>Wild type</td>
<td>4887</td>
</tr>
<tr>
<td>m3 5639 T→C</td>
<td>M3F/P80</td>
<td>MspI</td>
<td>C'CQ_G</td>
<td>Wild type</td>
<td>5638</td>
</tr>
<tr>
<td>m4 4887 C→A</td>
<td>M2F/M2R</td>
<td>BsaI</td>
<td>GGTCTCN'NNN_</td>
<td>Wild type</td>
<td>4877</td>
</tr>
</tbody>
</table>

* nt, nucleotide. Nucleotide positions refer to the beginning of exon 1 (Fig. 1).

---

**Table 2** Analysis of hypothetical linkages of CYP1A1 mutations

<table>
<thead>
<tr>
<th>Mutation linkage</th>
<th>Amplification with primers</th>
<th>Procedure</th>
<th>Fragment length (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ml and m2</td>
<td>P71/P80</td>
<td>Sequential digestion with MspI and BsrDI</td>
<td>Linked 1600, 1553, 253, 241, 206 Nonlinked 1806, 1347, 253, 241, 206</td>
</tr>
<tr>
<td>ml and m4</td>
<td>P71/P80</td>
<td>MspI digestion excision of the 1841-bp fragment, and subsequent BsaI digestion</td>
<td>Linked 448, 1393 Nonlinked 448, 484, 909</td>
</tr>
<tr>
<td>m2 and m4</td>
<td>M2F/M2R</td>
<td>Sequential digestion with BsrDI and BsaI</td>
<td>Linked 204, 139, 55, 10 Nonlinked 149, 139, 65, 55</td>
</tr>
</tbody>
</table>

* P71, 5'-ATTAAGGTTAGTGGGAGGGGACACG; binding position, nucleotides 4394-4417.
C4887A POLYMORPHISM OF HUMAN CYP1A1

bp upstream of m2 (Fig. 3). This mutation, called m4, led to the loss of a BsaI restriction site, and this was used for determination. The population frequency of m4 was studied in 880 individuals. m4 occurred in 2.95% of the 1760 alleles tested (Table 3). This prevalence was slightly higher than that of m2 (2.67%) but lower than that of m1 (7.73%). The African mutation m1 was not detected in our sample.

**Linkage of Mutations.** Mutation linkage analysis allowed the allocation of the mutations to different alleles. m2 was strictly linked to m1 as tested in 42 reference individuals and 12 lung cancer patients. Among all cases heterozygous for m1 and for m4 (two of the controls and one of the lung cancer patients), the mutations were not linked. Moreover, three individuals presenting with m2 and m4 (one control individual and two lung cancer patients) carried these mutations on different alleles, as could be proven by subsequent digestions with BsrDI and BsaI. Thus, m4 always appeared without any linkage.

**CYP1A1** alleles are defined (Table 3) with reference to the different mutation linkages. Alleles with 6235C (m1) were called *2A. m2 was in strict linkage disequilibrium with m1; therefore, this variance of *2A was termed *2B. The allele providing the novel, nonlinked mutation m4 was called *4.

**Frequency of CYP1A1 Genotypes in Controls and in Lung Cancer.** Within the 880 individuals tested, all observed CYP1A1 genotype frequencies matched exactly the expected percentages as calculated from the allele frequencies (Table 4). Because mutated alleles were relatively rare, homozygous carriers occurred only in a few cases. One individual provided a homozygous genotype CYP1A1*4f1@4 (0.11%). CYP1A1*4 was equally distributed among 157 lung cancer patients (2.87%; 95% confidence limits, 1.32%, 5.37%) and the 314 matched reference patients (2.87%; 95% confidence limits, 1.71%, 4.49%; Table 5). Stratification by gender and smoking habits failed to reveal

---

**Table 3. Frequency of four CYP1A1 mutations in 880 unrelated Caucasian individuals and allocation to specified alleles**

<table>
<thead>
<tr>
<th>Allele</th>
<th>Nucleotide at position (nt)</th>
<th>Allele frequency</th>
<th>95% confidence limits</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>4887 (m2)</td>
<td>4889 (m2)</td>
<td>5639 (m1)</td>
</tr>
<tr>
<td>CYP1A1*1 (wild type)</td>
<td>C</td>
<td>A</td>
<td>T</td>
</tr>
<tr>
<td>CYP1A1*2A</td>
<td>C</td>
<td>A</td>
<td>T</td>
</tr>
<tr>
<td>CYP1A1*2B</td>
<td>C</td>
<td>G</td>
<td>T</td>
</tr>
<tr>
<td>CYP1A1*3</td>
<td>C</td>
<td>A</td>
<td>C</td>
</tr>
<tr>
<td>CYP1A1*4</td>
<td>A</td>
<td>A</td>
<td>T</td>
</tr>
</tbody>
</table>

Mutation frequency (n) = 52, 47, 0, 136
% = 2.95, 2.67, 0.00, 7.73
2.5% confidence limit = 2.21, 1.97, 0.00, 6.52
97.5% confidence limit = 3.86, 3.54, 0.21, 9.07

* Referred to 1760 alleles.

**Table 4. Frequency of CYP1A1 genotypes among 880 unrelated Caucasian individuals**

<table>
<thead>
<tr>
<th>Genotype of CYP1A1</th>
<th>n</th>
<th>%</th>
<th>95% confidence limits</th>
<th>Expected*</th>
</tr>
</thead>
<tbody>
<tr>
<td>*1/*1</td>
<td>701</td>
<td>79.67</td>
<td>76.85, 82.27</td>
<td>78.79</td>
</tr>
<tr>
<td>*1/*2A</td>
<td>81</td>
<td>9.20</td>
<td>7.38, 11.13</td>
<td>9.03</td>
</tr>
<tr>
<td>*1/*2B</td>
<td>42</td>
<td>4.77</td>
<td>3.46, 6.40</td>
<td>4.77</td>
</tr>
<tr>
<td>*1/*3</td>
<td>47</td>
<td>5.34</td>
<td>3.95, 7.04</td>
<td>5.28</td>
</tr>
<tr>
<td>*1/*4</td>
<td>1</td>
<td>0.11</td>
<td>0.00, 0.63</td>
<td>0.26</td>
</tr>
<tr>
<td>*2A/*2A</td>
<td>4</td>
<td>0.45</td>
<td>0.12, 1.16</td>
<td>0.27</td>
</tr>
<tr>
<td>*2A/*2B</td>
<td>2</td>
<td>0.23</td>
<td>0.03, 0.82</td>
<td>0.30</td>
</tr>
<tr>
<td>*2B/*2B</td>
<td>1</td>
<td>0.11</td>
<td>0.00, 0.42</td>
<td>0.07</td>
</tr>
<tr>
<td>*2B/*3</td>
<td>1</td>
<td>0.11</td>
<td>0.00, 0.63</td>
<td>0.16</td>
</tr>
<tr>
<td>*2B/*4</td>
<td>1</td>
<td>0.11</td>
<td>0.00, 0.63</td>
<td>0.09</td>
</tr>
<tr>
<td>Total</td>
<td>880</td>
<td>100.00</td>
<td></td>
<td>100.00</td>
</tr>
</tbody>
</table>

*Expected frequencies are calculated from the allele frequencies in Table 3 according to the Hardy-Weinberg law.
a statistically significant difference (Table 5). There was only a trend of increased CYP1A1*4 frequency among the nonsmokers and the moderately smoking patients. Tumor histology had no impact on m4 frequency (data not shown).

**DISCUSSION**

**Improvement of Genotyping CYP1A1.** The introduction of the restriction enzyme BsrDI for evaluation of nucleotide 4889 (m2) instead of an allele-specific set of primers led indirectly to the discovery of the new C4887A transversion. Whereas the allele-specific amplification revealed homozygous mutants in two lung cancer patients, the BsrDI RFLP showed only heterozygosity. DNA sequencing and subsequent linkage analysis showed clearly that both samples provided one *2B allele and one *4 allele (Fig. 3). The reason for the mistyping using allele-specific primers was probably the diminishment of the annealing capability of the wild type-specific primer caused by the neighboring 4889A mutation. Difficulties with such genotyping techniques are well known and have already been commented on (17). The nomenclature follows the recent rules as provided one *2B allele and one *4 allele (Fig. 3). The reason for the mistyping using allele-specific primers was probably the diminishment of the annealing capability of the wild type-specific primer caused by the neighboring 4889A mutation. Difficulties with such genotyping techniques are well known and have already been communicated for CYP1A1 with a different strategy to avoid these problems (17, 18).

**Characteristics of the Novel C4887A Mutation.** The significance of threonine exchanged to asparagine in codon 461 is still unknown. Both are hydrophilic amino acids. Interestingly, an exchange of the adjacent hydrophobic amino acid residues, isoleucine to valine, results in an increase of enzyme activity (10). m4 is not in linkage disequilibrium with m2. The role of genotypes containing only nonlinked m1 (i.e., the CYP1A1*2A allele) is not clear, at least not in Caucasians. Interestingly, m1 has been associated with a high degree of inducibility (21). The novel mutation m4, however, was equally distributed among lung cancer patients and controls. With our sample size, an odds ratio $\approx 2.8$ could be excluded by power analysis (type-I error 5%, type-II error 20%) for m4 in heterozygous configuration. Thus, the change to asparagine in codon 461 does not appear to modulate lung cancer susceptibility.

In those studies in which allele-specific PCR was used, some of the data on m2 should be rechecked because of the potential disturbance by the m4 mutation to the primer-binding conditions. However, our previous results on the overrepresentation of m2 in lung cancer patients (7) are still valid, despite two lung cancer cases, who were apparently homozygous for m2 but now show heterozygosity at nucleotides 4887 and 4889. The newly calculated odds ratio of m2 in 157 cases and 314 controls is 3.01 (95% confidence limits, 1.29, 7.26; $P_{0.004}$; data not shown). Hence, CYP1A1 allele *2B has been confirmed as an important host factor of lung cancer in Caucasians. Our further work aims at the enzymatic properties, inducibility, and membrane topology of the CYP1A1 allele *4 product.

**ACKNOWLEDGMENTS**

As the DNA source of m3, venous blood samples from United States blacks were kindly donated by Dr. Helmut F. Cascorbi, Department of Anesthesiology, University Hospitals, Cleveland, Ohio. We thank H. Maszynski and P. Pietzsch for excellent technical assistance.

**REFERENCES**

8. Hamada, G. S., Sugimura, H., Suzuki, I., Nagura, K., Kiyokawa, E., Iwase, T.,


A C4887A Polymorphism in Exon 7 of Human CYP1A1: Population Frequency, Mutation Linkages, and Impact on Lung Cancer Susceptibility

Ingolf Cascorbi, Jürgen Brockmöller and Ivar Roots


Updated version Access the most recent version of this article at:
http://cancerres.aacrjournals.org/content/56/21/4965

E-mail alerts Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.