

SHORT COMMUNICATION

Naoto Keicho · Mitsuru Emi · Mitsuko Kajita
Ikumi Matsushita · Koh Nakata · Arata Azuma
Nobuya Ohishi · Shoji Kudoh

Overestimated frequency of a possible emphysema-susceptibility allele when microsomal epoxide hydrolase is genotyped by the conventional polymerase chain reaction-based method

Received: September 27, 2000 / Accepted: November 10, 2000

Abstract A recent association study suggested that the His113 variant of microsomal epoxide hydrolase (mEPHX) may confer a risk for development of emphysema, presumably by increasing susceptibility to smoking injury. Before considering a possible role of this enzyme in pulmonary disease, we attempted to characterize the genetic polymorphism further. The Tyr/His113 polymorphism within exon 3 of *mEPHX* was initially examined in 62 healthy individuals by conventional methods involving polymerase chain reaction (PCR)-based determination of a restriction fragment length polymorphism (RFLP). Genomic nucleotide sequences, including the polymorphic site and the downstream primer sequence, were further analyzed in 95 unrelated, healthy Japanese volunteers by single-stranded conformation polymorphism (SSCP) analysis and direct sequencing. Genotyping by the first method (PCR-RFLP) revealed that the allelic distribution in our test population apparently deviated from Hardy-Weinberg equilibrium. Sequence analysis showed that a synonymous nucleotide substitution, AAG to AAA (Lys119), was located just within the published primer site. The AAA at codon 119 was present only in alleles with Tyr113, and its frequency reached 0.31 in our panel of 190 Japanese alleles. This substitution potentially hampered PCR amplification because of the nucleotide mismatch, with the result that the fre-

quency of the Tyr113 variation was underestimated. The frequency of His113, a possible emphysema susceptibility allele of the *mEPHX* gene, was thus overestimated when human DNA samples were genotyped in the conventional way. Depending on the population(s) tested, this anomaly could represent a pitfall for PCR-based association studies.

Key words Microsomal epoxide hydrolase · Emphysema · Restriction fragment length polymorphism · Single-stranded conformation polymorphism · Genotyping

Introduction

Xenobiotic metabolizing enzymes are thought to detoxify products of cigarette smoke and other inhaled substances in the lung. Enzyme activity that is altered because of genetic variations might influence susceptibility to diseases such as lung cancer and chronic obstructive pulmonary disease (COPD) (Barnes 1999). Microsomal epoxide hydrolase (mEPHX) is one of the major detoxifying enzymes expressed in the lung. A published in-vitro expression study demonstrated that a His113 substitution for Tyr113 in mEPHX resulted in partial loss of activity (Hassett et al. 1994); moreover, an association study indicated that homozygosity for the His113 variant was significantly more common among patients with emphysema or COPD (Smith and Harrison 1997). Before considering further a possible role of this variation in a variety of lung diseases, we attempted to characterize it genetically.

Subjects and methods

Genomic DNA was extracted from blood samples, donated by 95 healthy and unrelated Japanese individuals, according to the procedures described previously (Emi et al. 1999). Initially we examined 62 of the 95 DNA samples for the Tyr/His113 polymorphism encoded within exon 3 of the

N. Keicho · I. Matsushita · K. Nakata
Department of Respiratory Diseases, Research Institute,
International Medical Center of Japan, Tokyo, Japan

M. Emi (✉) · M. Kajita
Department of Molecular Biology, Institute of Gerontology, Nippon
Medical School, 1-396 Kosugi-cho, Nakahara-ku, Kawasaki 211-8533,
Japan
Tel. +81-44-733-5230; Fax +81-44-733-5192
e-mail: memi@nms.ac.jp

A. Azuma · S. Kudoh
Fourth Department of Internal Medicine, Nippon Medical School,
Tokyo, Japan

N. Ohishi
Department of Respiratory Medicine, University of Tokyo, Tokyo,
Japan

mEPHX gene, using the polymerase chain reaction-restriction fragment length polymorphism (PCR-RFLP) method described in the original report (Smith and Harrison 1997). The PCR amplifications were performed in 25- μ l reaction mixtures containing 100 ng of genomic DNA, 0.5 μ M of each published primer (EPO1; 5'-GATCGATA AGTTCCGTTTCACC-3' and EPO2; 5'-ATCTTAGTCT TGAAGTGAGGAT-3'), 0.2 mM of each dNTP, 1.5 mM of MgCl₂, 10 mM of Tris-HCl (pH 8.3), 50 mM of KCl, and 1 unit of Taq polymerase (Perkin-Elmer, Norwalk, CT, USA) (Tsukamoto et al. 1998; Watanabe et al. 1998). The downstream primer (EPO2) abuts the polymorphic site, and an engineered base change in this primer (shown above in bold type) produces a cleavage site for *EcoRV* endonuclease in the Tyr113 variant only (Smith and Harrison 1997). After initial denaturation at 94°C for 5 min, the reaction mixtures were subjected to 35 cycles of 94°C for 30 s, 56°C for 30 s, and 72°C for 45 s, followed by final extension at 72°C for 5 min. Each 162-bp product was digested with *EcoRV* at 37°C for 2 h, and the fragments were separated on 8% acrylamide gels and visualized by ethidium-bromide staining.

Genomic sequences that included the downstream primer site were further genotyped in 95 Japanese individuals by PCR-based single-stranded conformation polymorphism (SSCP) analysis (Hirayama et al. 1998) and direct sequencing. Amplification under the same cycling conditions as before, this time using EPO1 and a newly designed downstream primer (EPOL1; 5'-CACATTGTGGAAG AAGGCTG-3'), generated a 233-bp PCR product from each sample.

After 1 μ l of each PCR product was mixed with 7 μ l of deionized formamide and denatured at 94°C for 4 min, 1 μ l of each mixture was electrophoresed at 13°C on a nondenaturing 10% polyacrylamide gel containing 10% glycerol, with an acrylamide:bisacrylamide ratio of 49:1 in 45 mM Tris-borate and 1 mM ethylene diamine tetraacetic acid (EDTA). Single-stranded DNA fragments on the gel were visualized using a silver staining kit (Daiichi Pure Chemicals, Tokyo, Japan). Nucleotide sequences of identified variants were determined by direct sequencing on an ABI PRISM 377 DNA Sequencer with a dRhodamine Terminator Cycle Sequencing Ready Reaction Kit (PE Applied Biosystems, Foster City, CA, USA).

χ^2 analysis based on observed and expected genotype frequencies was performed to determine whether the genotypic distribution was in Hardy-Weinberg equilibrium (Terwilliger and Ott 1994). *P* values less than 0.05 were considered significant in demonstrating deviation from equilibrium.

Results

Our preliminary experiments involved using the conventional PCR-RFLP method to genotype 62 healthy volunteers for polymorphism in exon 3 of the *mEPHX* gene. Six of these samples were barely amplified with the published

primer set; in the others we observed genotypic frequencies for Tyr113/Tyr113, Tyr113/His113, and His113/His113 of 0.43, 0.27, and 0.30, respectively. In particular, the frequency of heterozygotes (0.27) was lower than the 0.49 one would expect on the assumption of Hardy-Weinberg equilibrium; the distribution of these alleles deviated statistically from equilibrium (*P* < 0.001). On the basis of this finding, we speculated that, because it appeared that some alleles of *mEPHX* were not amplified efficiently, some heterozygotes probably had been genotyped as homozygotes.

To test this possibility, we searched the GenBank database and discovered a possible synonymous (silent) single-nucleotide substitution just within the downstream primer (GenBank accession numbers, L25878 and L25879). Therefore, we used EPO1 and a newly designed downstream primer, EPOL1, to amplify individual genomic regions that included codon 113 and the original downstream primer site in 95 healthy volunteers. We analyzed the products by the SSCP method (Fig. 1) and determined the nucleotide sequences. As summarized in Table 1, we identified three different alleles, all encoding lysine at residue 119: Tyr113-AAG119, Tyr113-AAA119, and His113-AAG119. No individual in our Japanese population panel showed the His113-AAA119 homozygous genotype. As to double heterozygotes for the two polymorphisms (*n* = 27), we segregated each haplotype and determined the phase by subcloning the polymorphic region in plasmid vector from each individual. We found that each of them was heterozygous for His113-AAG119 and Thr113-AAA119 (data not shown). The frequencies of the Tyr113 and His113 variations among the 190 Japanese alleles were 0.57 (i.e., 0.26 +

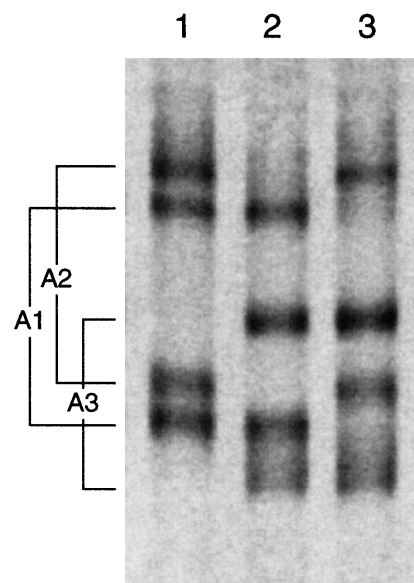


Fig. 1. Genotyping of the *mEPHX* exon 3 polymorphism by polymerase chain reaction-single-stranded conformation polymorphism (PCR-SSCP) analysis. *Lane 1*, A heterozygote for A1 and A2; *lane 2*, a heterozygote for A1 and A3; *lane 3*, a heterozygote for A2 and A3. Sequence analysis showed that the A1, A2, and A3 pairs of bands were single-stranded DNA fragments containing Tyr113-AAG119, Tyr113-AAA119, and His113-AAG119, respectively

Table 1. Allelic distribution of *mEPHX* exon 3 polymorphism in a normal Japanese population ($n = 95$)

Allele name	Nucleotide sequence	Allele frequency
Tyr-113-AAG (Lys-119)	5'- <u>TACCCTCACTTCAAGACTAAGATT</u> -3'	0.26
Tyr-113-AAA (Lys-119)	5'- <u>TACCCTCACTTCAAGACTAAAATT</u> -3'	0.31
His-113-AAG (Lys-119)	5'- <u>CACCCTCACTTCAAGACTAAGATT</u> -3'	0.43

Nucleotide sequences that matched the original downstream primer (EPO2) for polymerase chain reaction-restriction fragment length polymorphism (PCR-RFLP) analysis are underlined. Substitutions are shown in bold type. TAC and CAC at the 5' end encode Tyr and His, respectively, at residue 113

0.31) and 0.43, and the distribution of these genotypes was in Hardy-Weinberg equilibrium (data not shown). The samples that were barely amplified with the original primer set were found to be homozygous for Tyr113-AAA119. Furthermore, individuals heterozygous for this combination and His113-AAG119 were apparently genotyped as homozygotes for His113 (data not shown).

Discussion

The study reported here revealed that a single-nucleotide substitution can hamper PCR amplification of the *mEPHX* exon 3 polymorphism with the published primer set. For conventional genotyping by the PCR-RFLP method, one primer should contain an engineered base change and abut the polymorphic site so that it can generate an artificial restriction site to discriminate the polymorphism, because no natural restriction sites are available in this region. Unexpectedly, we located another nucleotide substitution close to the polymorphic site and just within the sequence of the published downstream primer. Because of these two mismatches; namely, the engineered base change in the primer and the single-nucleotide substitution within the primer site, the Tyr113-AAA119 allele was not amplified efficiently by PCR and, thus, the frequency of the His113 allele was overestimated.

Because the distribution of the Tyr113/His113 polymorphic alleles determined by the PCR-RFLP method in the work of Smith and Harrison (1997) was in Hardy-Weinberg equilibrium, the allele frequency of Tyr113-AAA119 might have been low or negligible in the population they studied. In fact, the frequency of Tyr113-AAA119 was 0.09 in 76 unrelated Caucasian alleles we tested and relatively lower than that in the Japanese population (unpublished data). Nevertheless, when association studies are conducted in a genetically uncharacterized ethnic population, this allele cannot be neglected, because the potential for overestimation of the frequency of His113 in clinical cases and controls

can lead to false-positive or false-negative results, depending on the PCR conditions. To avoid such a problem, other genotyping methods, such as SSCP and allele-specific hybridization using another PCR primer set, should be performed, as we have shown here. Eliminating the 3' end of the original downstream primer would be an alternative, although the melting temperature of the truncated primer would be less than 50°C and this might make it difficult to set a high annealing temperature for the PCR. We conclude that a possible nucleotide substitution within the downstream primer should be considered a potential pitfall for PCR-based genetic association studies that involve microsomal epoxide hydroxylase and lung diseases.

Acknowledgments This work was supported in part by research grants for Surveys and Research on Specific Diseases from the Ministry of Health and Welfare of Japan, in 1998 and 1999.

References

- Barnes PJ (1999) Molecular genetics of chronic obstructive pulmonary disease. *Thorax* 54:245–252
- Emi M, Keicho N, Tokunaga K, Katsumata H, Souma S, Nakata K, Taguchi K, Ohishi N, Azuma A, Kudoh S (1999) Association of diffuse panbronchiolitis with microsatellite polymorphism at the human interleukin 8 (IL-8) locus. *J Hum Genet* 44:169–172
- Hassett C, Aicher L, Sidhu JS, Omiecinski CJ (1994) Human microsomal epoxide hydrolase: genetic polymorphism and functional expression in vitro of amino acid variants. *Hum Mol Genet* 3:421–428
- Hirayama T, Yamaki E, Hara A, Tsuji M, Hashimoto K, Yamamoto M, Emi M (1998) Five familial hypercholesterolemic kindreds in Japan with novel mutation of the LDL receptor gene. *J Hum Genet* 43:250–254
- Smith CAD, Harrison DJ (1997) Association between polymorphism in the gene for microsomal epoxide hydrolase and susceptibility to emphysema. *Lancet* 350:630–633
- Terwilliger JD, Ott J (1994) Handbook of human genetic linkage. The Johns Hopkins University Press, Baltimore, pp 277–279
- Tsukamoto K, Haruta K, Shiba T, Emi M (1998) Isolation and mapping of a polymorphic CA repeat sequence at the human interleukin 6 locus. *J Hum Genet* 43:71–72
- Watanabe I, Tsukamoto K, Shiba T, Emi M (1998) Isolation and radiation hybrid mapping of dinucleotide repeat polymorphism at the human matrix Gla protein (MGP) locus. *J Hum Genet* 43:75–76