Genetic polymorphisms in three Iranian populations with different risks of esophageal cancer, an ecologic comparison

Alireza Sepehr\textsuperscript{a}, Farin Kamangar\textsuperscript{a}, Christian C. Abnet\textsuperscript{a}, Saman Fahimi\textsuperscript{b}, Akram Pourshams\textsuperscript{c}, Hossein Poustchi\textsuperscript{c}, Sirous Zeinali\textsuperscript{c}, Masood Sotoudeh\textsuperscript{c}, Farhad Islami\textsuperscript{c}, Dariush Nasrollahzadeh\textsuperscript{c}, Reza Malekzadeh\textsuperscript{c}, Philip R. Taylor\textsuperscript{a}, Sanford M. Dawsey\textsuperscript{a,*}

\textsuperscript{a}The US National Cancer Institute, Bethesda, MD, USA
\textsuperscript{b}The International Agency for Research on Cancer, Lyon, France
\textsuperscript{c}Digestive Disease Research Center, Tehran University of Medical Sciences, Tehran, Iran

Received 10 February 2004; received in revised form 20 May 2004; accepted 20 May 2004

Abstract

The age-standardized incidence of esophageal cancer (EC) varies from 3 to >100/100,000 per year in different provinces of Iran. This striking variation of incidence is associated with differences in ethnic backgrounds, raising the possibility that genetic factors are involved in the pathogenesis of EC. We compared the frequencies of polymorphisms in ten genes that have been hypothesized to have a role in risk of EC (\textit{CYP1A1}, \textit{CYP2A6}, \textit{CYP2E1}, \textit{GSTM1}, \textit{GSTP1}, \textit{GSTT1}, \textit{ADH2}, \textit{ADH3}, \textit{ALDH2}, and \textit{O6-MGMT}) among three Iranian ethnic groups with highly varying rates of EC. These three groups included high-risk Turkomans, medium-risk Turks, and low-risk Zoroastrian Persians. Compared to Zoroastrians, Turkomans had higher frequency of four alleles that are speculated to favor carcinogenesis (\textit{CYP1A1 m1}, \textit{CYP1A1 m2}, \textit{CYP2A6*9}, and \textit{ADH2*1}); these results are consistent with an influence of these allele variants on the population risk of EC. However, none of these four alleles had a high enough prevalence in Turkomans to explain the high rates of EC in this group. Three of these four alleles (\textit{CYP1A1 m1}, \textit{CYP1A1 m2}, \textit{CYP2A6*9}) were less frequent among Turkomans than in some Asian populations with lower risks of EC. We conclude that it is unlikely that variations in these polymorphic genes are major contributors to the high incidence of EC among Turkomans in Iran.

© 2004 Elsevier Ireland Ltd. All rights reserved.

Keywords: Genetic polymorphisms; Esophageal cancer; Iran

1. Background

Esophageal cancer (EC) is the sixth most common cause of cancer mortality worldwide [1]. The incidence of EC is highly variable in different
populations, with more than a 50-fold difference between very high- and very low-risk populations [2].

Such striking differences in EC incidence are observed between different ethnic groups in Iran. Turkomans in northeastern Iran are considered to be a very high-risk group, with age standardized rates (ASR) of above 100/100,000 for both men and women [3]; Turks in Ardabil province, in the northwest of Iran, have a moderately high risk, with ASRs of 15–30/100,000 [3,4]; whereas the rates of EC in pure Zoroastrian Persians in Iran and India are known to be considerably lower, with ASRs of 3–7/100,000 [5].

One possible explanation for the observed differences in EC rates between populations may be differences in their genetic pools. Sunni Muslim Turkomans and Zoroastrian Persians are geographically, ethnically, and religiously distinct. Turkomans are descendants of Oguz Turkic tribes who migrated from the Altai Mountains on the border of China and Mongolia to northern Iran [6], have oriental facial features, and probably have a genetic background similar to East Asians. Zoroastrian Persians are Aryans, and therefore, are considered to be Caucasians. Ardabili Turks are probably a hybrid of Aryans and Turkic tribes who migrated from East Asia. Therefore, the pattern of EC incidence in these groups roughly parallels their ethnic closeness to East Asian ancestors. EC has a very high incidence in reports from parts of Central Asia known as ‘Central Asian Esophageal Cancer Belt’, an area which extends from China and Mongolia to the northern parts of Iran, but it has a low incidence in Caucasians [2]. We hypothesized that a possible contributing factor to the observed differences in EC rates in these three ethnic groups in Iran might be different prevalences of specific genetic polymorphisms in Turkomans, Turks, and Zoroastrians.

Variation in genes involved in metabolizing xenobiotics is thought to alter individual risk of cancer [7]. Cytochrome P-450 (CYP) and glutathione-S-transferase (GST) enzymes, which are involved in the activation and detoxification of xenobiotics, alcohol and aldehyde dehydrogenases (ADHs and ALDHs), which metabolize ethanol and retinol, and the DNA repair enzyme O6-methylguanine-DNA methyltransferase (MGMT) have all been suggested to have a role in the etiology of EC. Therefore, we conducted an ecologic study to determine the prevalence of some polymorphic alleles in the genes that code for these enzymes in genomic DNA from Turkomans, Turks, and Zoroastrians and looked for a pattern that might suggest a role for these genes in the etiology of EC.

2. Subjects and methods

The three groups who participated in this ecologic study were Turkomans from Golestan Province, northeastern Iran, Turks from Ardabil Province, northwestern Iran, and Zoroastrian Persians from Tehran. One hundred and ten healthy volunteers from each group, ages 30–60 years, participated in this study. Half of the volunteers in each group were males. The study was approved by the Institutional Review Board of the Digestive Disease Research Center (DDRC), Tehran University of Medical Sciences, Iran, and was conducted in April 2001. The Office of Human Subjects Research of the US National Institutes of Health (NIH) exempted the analysis of the anonymized DNA samples from this study from review.

After signing the informed consent, each participant was interviewed and completed a structured questionnaire to document his or her ethnic background. Ten milliliter of venous blood was drawn from each patient. The blood was immediately frozen and sent on dry ice to the central laboratory of DDRC. Genomic DNA was extracted from the frozen blood and kept at −20°C at DDRC. The samples were anonymized and then were sent to NCI.

DNA was genotyped for polymorphic sites in the following genes: CYP1A1 MspI site (m1 or 3801 T>C), CYP1A1 Ile/Val polymorphism in Exon 7 (m2 or 2455 A>G), CYP1A1 African–American specific polymorphism (m3 or 3205 T>C), CYP2A6 (−48 T>G or *9), CYP2E1 RsaI polymorphism (c2 or −1293 G>C), GSTM1 homozygous deletion (*0/*0 genotype), GSTT1 homozygous deletion (*0/*0 genotype), GSTP1 Codon 104 polymorphism (313 A>G), ADH2 (*2 vs. *1 allele), ADH3 (*2 vs. *1 allele), ALDH2 (*2 vs. *1 allele), and O6-MGMT (290 C>T).

Genetic polymorphism analyses were carried out by Bioserve Biotechnologies Inc. (Laurel, MD) using either Sequenom® (CYP1A1 m1, CYP1A1 m2, CYP1A1 m3, CYP2A6, CYP2E1, GSTP1, ADH3, O6-MGMT), PCR-RFLP (GSTM1, GSTT1), or amplified product length polymorphism (ADH2, ALDH2) assays. Table 1 shows the primers
### Table 1: Primers and different names of the polymorphisms studied

<table>
<thead>
<tr>
<th>dbSNP Number</th>
<th>Assay</th>
<th>Forward primer</th>
<th>Reverse primer</th>
<th>Mass extension primer</th>
</tr>
</thead>
<tbody>
<tr>
<td>CYP1A1 (3801 T&gt;C) (m1)</td>
<td>rs4646903 MALDI-TOF mass array</td>
<td>ACGTTGGATGACTAC CAGAGCTGAGGGTG</td>
<td>ACGTTGGATGACTAC CAGAGCTGAGGGTG</td>
<td>CACTGTAACCTCCACCTCC</td>
</tr>
<tr>
<td>CYP1A1 (1506 A&gt;G) (m2)</td>
<td>rs1048943 MALDI-TOF mass array</td>
<td>ACGTTGGATGATGACTAC TCAGAGCTGAGGGTG</td>
<td>ACGTTGGATGATGACTAC TCAGAGCTGAGGGTG</td>
<td>AGACCTCCAGCGGGCCAA</td>
</tr>
<tr>
<td>CYP1A1 (3205 T&gt;C) (m3)</td>
<td>rs498684 MALDI-TOF mass array</td>
<td>ACGTTGGATGATGCTTCAAC TCACTGCACCTTC</td>
<td>ACGTTGGATGACTCTTGCCCTTCAAC TCACTGCACCTTC</td>
<td>GACCTCCAGCGGGCCAA</td>
</tr>
<tr>
<td>CYP2A6*9 (−48 T&gt;G)</td>
<td>– MALDI-TOF mass array</td>
<td>ACGTTGGATGAGGTGTA GGAATCTGAGCTGACGATACG</td>
<td>ACGTTGGATGAGGTGTA GGAATCTGAGCTGACGATACG</td>
<td>CACTGTAACCTCCACCTCC</td>
</tr>
<tr>
<td>CYP2E1 c2 (−1293 G&gt;C) (RafI)</td>
<td>rs381386 MALDI-TOF mass array</td>
<td>ACGTTGGATGATGACTAC TCAGAGCTGAGGGTG</td>
<td>ACGTTGGATGACTAC TCAGAGCTGAGGGTG</td>
<td>AGACCTCCAGCGGGCCAA</td>
</tr>
<tr>
<td>GSTM1*0/*0 (Homo deletion)</td>
<td>rs1065411 PCR-RFLP</td>
<td>GAACCTCCCGTAAATACGTGTA TAAAGCGTTCCCTTCTTGGTTCAGGAGAA</td>
<td>GAACCTCCCGTAAATACGTGTA TAAAGCGTTCCCTTCTTGGTTCAGGAGAA</td>
<td>CACTGTAACCTCCACCTCC</td>
</tr>
<tr>
<td>GSTT1*0/*0 (Homo deletion)</td>
<td>rs4630 PCR-RFLP</td>
<td>TCACCAGATCAGCCAGCCACA TAAAGCGTTCCCTTCTTGGTTCAGGAGAA</td>
<td>TCACCAGATCAGCCAGCCACA TAAAGCGTTCCCTTCTTGGTTCAGGAGAA</td>
<td>AGACCTCCAGCGGGCCAA</td>
</tr>
<tr>
<td>GSTP1 (313 A&gt;G)</td>
<td>rs947894 MALDI-TOF mass array</td>
<td>ACGTTGGATGACACCTCCTAATAC TAAAGCGTTCCCTTCTTGGTTCAGGAGAA</td>
<td>ACGTTGGATGACACCTCCTAATAC TAAAGCGTTCCCTTCTTGGTTCAGGAGAA</td>
<td>CACTGTAACCTCCACCTCC</td>
</tr>
<tr>
<td>ADH2*2</td>
<td>rs1229984 APLP*</td>
<td>ACGTTGGATGACACCTCCTAATAC TAAAGCGTTCCCTTCTTGGTTCAGGAGAA</td>
<td>ACGTTGGATGACACCTCCTAATAC TAAAGCGTTCCCTTCTTGGTTCAGGAGAA</td>
<td>CACTGTAACCTCCACCTCC</td>
</tr>
<tr>
<td>ADH3*2 (350 A&gt;G)</td>
<td>rs698 MALDI-TOF mass array</td>
<td>ACGTTGGATGACACCTCCTAATAC TAAAGCGTTCCCTTCTTGGTTCAGGAGAA</td>
<td>ACGTTGGATGACACCTCCTAATAC TAAAGCGTTCCCTTCTTGGTTCAGGAGAA</td>
<td>CACTGTAACCTCCACCTCC</td>
</tr>
<tr>
<td>O6-MGMT (290 C&gt;T) (Codon 84)</td>
<td>rs12917 MALDI-TOF mass array</td>
<td>ACGTTGGATGACACCTCCTAATAC TAAAGCGTTCCCTTCTTGGTTCAGGAGAA</td>
<td>ACGTTGGATGACACCTCCTAATAC TAAAGCGTTCCCTTCTTGGTTCAGGAGAA</td>
<td>CACTGTAACCTCCACCTCC</td>
</tr>
</tbody>
</table>

*APLP, amplified product length polymorphism.
and various names of these polymorphism, including their dbSNP numbers. Genotyping was performed without knowledge of the ethnicity of the participants. Genotypes for all of the polymorphisms were successfully determined in at least 90% of the participants in each group, except for GSTP1 and ALDH2: we were able to genotype GSTP1 in only 82% of the Turk and 82% of the Zoroastrian participants, and we were able to genotype ALDH2 in only 84% of the Turk and 72% of the Zoroastrian participants.

We used a $\chi^2$-test to compare the observed genotype frequencies with Hardy-Weinberg proportions and to compare the prevalence of different polymorphisms in the three ethnic groups.

3. Results

Sufficient quantities of DNA samples were obtained from 107 Turkoman, 108 Turk, and 106 Zoroastrian participants. Hardy-Weinberg equilibrium was observed for all of the genes within each ethnic subgroup, except for ADH2. The frequency of ADH2*2/*1 was higher than expected in all of the three ethnic groups, suggesting an advantage for the heterozygote form of this gene.

Table 2 shows the percentages of variant allele frequencies by ethnicity. Frequencies of CYP2E1 c2 and GSTP1 (313 G) alleles and the GSTT1 homozygous deletion genotype did not differ significantly among the three ethnic groups, and the differences in allele frequencies of ADH3*2 and ALDH2*2 were only marginally significant. However, there were significant differences in the frequencies of CYP1A1 m1, CYP1A1 m2, CYP2A6*9, ADH2*2, and O6-MGMT (290T) alleles and in the frequency of the GSTM1*0/*0 genotype among the three groups. The CYP1A1 m3 allele was not seen in any of the blood specimens.

4. Discussion

Differences in the risk of EC between high- and low-risk populations may partly be attributed to host factors, including the genetic make up of the populations. In this study we have investigated differences in the prevalence of polymorphic alleles in genes that are hypothesized to have a role in the etiology of EC among three populations with high, medium, and low incidence of EC. The genes that were investigated in this study were the genes that code for Phase I (mainly CYPs) and Phase II (mainly GSTs) drug metabolizing enzymes, alcohol and aldehyde dehydrogenases, and the DNA repair enzyme O6-MGMT.

CYP1A1 is a Phase I enzyme that metabolizes polycyclic aromatic hydrocarbons (PAHs) and aromatic hydrocarbons (AHs), which are considered to be the primary etiological factor for EC. The enzyme is involved in the metabolism of most of the known carcinogenic PAHs, and the frequency of CYP1A1 m1 allele was found to be significantly higher among the Turkoman population, suggesting an advantage for the heterozygote form of this enzyme.
amines, potential environmental risk factors of EC, and can activate them to ultimate carcinogens [8,9]. Two polymorphisms in CYP1A1 gene, the m1 and m2 alleles, have been extensively studied and both have been shown to increase the catalytic activity of this enzyme and hence to produce more carcinogenic material [8]. Interestingly, the prevalence of these rapid metabolizing alleles correlated very well with the incidence of EC in our study populations. However, the prevalence of m1 and m2 alleles in Turkomans (29 and 11%, respectively) was lower than the mean prevalence reported in a large number of Asian populations (50 and 28%, respectively [10]), many of whom do not have high rates of esophageal cancer. Therefore, although m1 and m2 alleles may contribute to the high risk of EC among Turkomans, they are unlikely to be the main reasons for observing such high EC rates in this group. The CYP1A1*3 (m3) allele was not found in any of these populations; this was consistent with other studies that have found this polymorphism only in African–Americans.

CYP2A6 is a hepatic enzyme that activates several pro-carcinogens including nitrosamines [8,11]. Although reduced activity of CYP2A6 would be expected to decrease the amount of circulating carcinogens and thus be associated with lower rates of tobacco-related cancers, a study examining the association between CYP2A6*4, a variation of the enzyme with completely defective activity, found a higher incidence of lung cancer and EC among non-smokers with the deletion [11]. The authors of this study attributed this finding to the inhibition of first-pass clearance of nitrosamines in the liver. They suggested that when CYP2A6 is active in the liver, carcinogenic material is produced in the liver and is circulated throughout the whole body, exposing each distant organ to only small concentrations of this carcinogen. If the enzyme is inactive in the liver, however, the pre-carcinogenic material may reach the esophagus and other target tissues in higher concentrations and may be converted locally to carcinogenic material by other members of the CYP family, such as CYP2E1. If this speculative rationale is correct, we might expect CYP2A6 alleles with diminished activity to be associated with higher rates of EC. We studied the frequency of CYP2A6*9, a variant allele of the gene that reduces the activity of the enzyme in vitro, in our three study populations. In our study the prevalence of CYP2A6*9 was higher in Turkomans (14%) than in Turks (5%) and Zoroastrians (4%), but slightly lower than in two Japanese populations [12]. The effect of this polymorphism on the incidence of EC has not been studied previously.

The CYP2E1 enzyme activates many carcinogens, including nitrosamines. At least four studies have evaluated the association of the c2 allele of CYP2E1 (previously measured using an RsaI RFLP assay) and the risk of EC. Three of the studies showed no association [13–15], while one showed a protective association between the c2 allele and the risk of EC [16]. Our study did not show any statistical difference in the prevalence of this allele between Turkomans, Turks, or Zoroastrians.

GSTs are Phase II enzymes that catalyze the conjugation of reduced glutathione to a variety of carcinogens, including reactive metabolites of benzo(a)pyrenes and other PAHs, increasing their water solubility and promoting their excretion [7].

GSTM1, one of the major members of the GST family, has a null polymorphism (GSTM1*0/*0 or homozygous deletion of the gene). Individuals with the null polymorphism lack the detoxifying capabilities attributed to GSTM1, and therefore may be more susceptible to EC and a range of other cancers. However, the results of several case-control studies conducted so far are inconsistent [9,15–22]. The observed trend in our data, which showed the lowest prevalence of GSTM1*0/*0 in Turkomans and the highest prevalence in Zoroastrians, shows that GSTM1*0/*0 is not a major carcinogenic factor for EC in the Turkoman population.

GSTT1, another member of the GST family also has a null polymorphism. None of the four case-control studies that have examined the effect of this polymorphism on the risk EC have found any association [16,19,21,23]. Our three populations were very similar regarding the prevalence of this polymorphism.

GSTP1 has a coding region single nucleotide polymorphism which results in replacement of the amino acid Val by Ile at residue 104. Four case-control studies have evaluated the association between this polymorphism and the risk of EC: two have found no association [16,24], one has found a positive association [23], and one has found a negative association [25]. Our study did not find any
significant difference in the prevalence of this polymorphism among our three ethnic populations.

Alcohol is converted to acetaldehyde in the hepatocytes by ADH, and acetaldehyde is in turn converted to acetate via ALDH. Since acetaldehyde is a known carcinogen [26], hyperactive forms of ADH and hypoactive forms of ALDH were originally hypothesized to be associated with a higher risk of EC.

Contrary to initial expectations, case-control studies have shown that the ADH2*2 allele, which encodes a hyperactive form of ADH2, is associated with a reduced EC risk [15,17,27–30]. Several possible explanations have been suggested for this observation [26]. For example, it has been suggested that hyperactive forms of ADH2 may reduce the time that alcohol lingers in the esophageal mucosa and therefore may reduce the risk of esophageal cancer. Our study also showed a lower prevalence of the ADH2*2 allele in high-risk Turkomans (51%) than in low-risk Zoroastrians (68%). For comparison, the prevalence of this allele in the Japanese population is approximately 78% [17].

We found the frequency of the ADH2*1/ADH2*2 genotype to be much higher in all the three ethnic groups than what was expected from Hardy-Weinberg proportions. Whether this represents a true heterozygote advantage is not clear, but this observation deserves further attention.

The ADH2 and ADH3 genes are reported to be in linkage disequilibrium among Asians; the ADH3*1 allele has been shown to be tightly linked to the ADH2*2 allele [26]. A Japanese study showed that after adjusting for ADH2 status, there was no longer an association between ADH3 polymorphisms and the risk of EC [17]. In our study, the prevalence of ADH3*2 allele was lower in Zoroastrians than in the other two groups, but the difference was not statistically significant.

ALDH2*2, a mutant allele that is common in East Asians and rare in other ethnic groups, encodes a variant form of ALDH2 that has very low activity and results in delayed removal of acetaldehyde from blood [26]. This polymorphism has been consistently shown to be positively associated with a higher risk of EC [15,17,27,29,31–35]. Our results show that the frequency of this allele is very low in all three groups; therefore this polymorphism is unlikely to be a reason for the high incidence of EC observed among Turkomans.

MGMT is involved in the repair of alkylated DNA, thereby protecting the cells from neoplastic events. By accepting the alkyl group from the alkylated base, this protein restores the original base [36]. Several polymorphisms have been found in the gene that encodes this protein, including a missense polymorphism in codon 84 which results in substitution of phenylalanine for leucine [37]. The role of this polymorphism in esophageal cancer has not been previously examined. Our results show that this polymorphism was much more common in Turkomans than in the other two ethnic groups.

In summary, this study adds to our knowledge of the allele frequencies of the tested polymorphisms in previously untested ethnic groups. Our data suggests that CYP1A1 m3, GSTM1*0, GSTT1*0, GSTP1 313 G, ADH3*2, and ALDH*2 alleles play a very small role, if any, in the high incidence of EC observed among Turkomans. Compared to Zoroastrians, Turkomans have higher frequency of four alleles that are presumed to favor carcinogenesis (CYP1A1 m1, CYP1A1 m2, CYP2A6*9, and ADH2*1); these results are consistent with an influence of these allele variants on the population risk of EC and thus strengthen the notion that these allele variants may be associated with an increased risk of EC. However, none of these four alleles had extremely high prevalence in Turkomans and three of these four alleles (CYP1A1 m1, CYP1A1 m2, CYP2A6*9) were less frequent among Turkomans than in some Asian populations with lower risks of EC. Therefore, it is unlikely that these alleles are the main reasons for observing a very high incidence of EC among Turkomans.

References


