International Immunology, Vol. 16, No. 10, pp. 1507–1514
doi:10.1093/intimm/dxh152 Advance Access publication 31 August 2004

Genetic basis of tobacco smoking: strong association of a specific major histocompatibility complex haplotype on chromosome 6 with smoking behavior

George Füst[1,2], Gudmundur J. Arason[3], Judith Kramer[4], Csaba Szalai[5], Jenő Duba[6], Yan Yang[7], Erwin K. Chung[7], Bi Zhou[7], Carol A. Blanchong[7], Marja-Liisa Lokki[8], Sigurdur Bödvarsson[9], Zoltán Prohászka[1,2], István Karádi[1,2], Ágnes Vatay[1], Margit Kovács[1], László Romics[1,2], Gudmundur Thorgeirsson[9] and **C. Yung Yu[7]**

[1]Third Department of Internal Medicine and [2]Research Group of Metabolism and Atherosclerosis, Semmelweis University and Hungarian Academy of Sciences, Budapest, Hungary
[3]Department of Immunology, Institute for Medical Laboratory Sciences, Landspítali University Hospital, Reykjavík, Iceland
[4]Central Laboratory, St John Hospital, [5]Heim Pál Pediatric Hospital and [6]National Institute of Cardiology, Budapest, Hungary
[7]Center for Molecular and Human Genetics, Columbus Children's Research Institute and College of Medicine and Public Health, The Ohio State University, Columbus, Ohio, USA
[8]Tissue Typing Laboratory, Finnish Red Cross Blood Transfusion Service and Transplantation Laboratory, University of Helsinki, Helsinki, Finland
[9]Department of Medicine, Landspitali University Hospital, Reykjavík, Iceland

Keywords: ancestral MHC haplotype, complement component C4A deficiency, odorant receptors, tobacco smoking, tumor necrosis factor α

Abstract

The genetic basis for addiction to tobacco smoking—particularly that of the perception of olfactory stimuli that may be important in reinforcing smoking addiction—is largely unknown. A cluster of genes for olfactory receptors is in close proximity to the MHC region on chromosome 6. Polymorphisms of MHC class III genes (RCCX modules, *TNFA* promoter polymorphisms) were determined in 101 healthy subjects and 232 coronary artery disease (CAD) patients from Hungary with defined tobacco smoking habits. A highly significant association between ever smoking (past + current smokers) and a specific MHC haplotype was observed (odds ratios = 2.14–4.13; *P*-values = 0.012 to <0.001). This haplotype is characterized by the presence of *C4A* null alleles and a solitary short *C4B* gene linked to the *TNF2* allele of the promoter for *TNFA* gene. This haplotype occurred more frequently in the ever smokers than in the never smokers [odds ratio: 4.97 (1.96–12.62); *P* = 0.001], and such associations were stronger in women (odds ratio = 13.6) than in men (odds ratio = 2.79). An independent study of complement C4 protein polymorphism and smoking habits in Icelandic subjects (*n* = 351) yielded similar and confirmative results. Considering the documented link between olfactory stimuli and smoking in females, and the presence of a cluster of odorant receptor genes close to the MHC class I region, our findings implicate a potential role of the MHC-linked olfactory receptor genes in the initiation of smoking.

Introduction

Tobacco smoking has a high impact on the development and progression of serious diseases such as coronary artery disease and lung cancer. Dependence on nicotine is a complex trait influenced by genetic and environmental factors (1,2). Twin studies ranked the heritability estimates for smoking between 46 and 84% (3,4). In a best-fit biometrical model with data from 778 male–male and female–female twin pairs reared together or apart, it was shown that genetic effects accounted for 61% of the variance in liability to regular tobacco use in men; a similar value (63%) was found in women born after 1940 (4). To date, the main candidate genes associated with smoking are cytochrome P450 enzymes implicated in nicotine metabolism

Correspondence to: G. Füst; E-mail: FustGe@kut.sote.hu

Transmitting editor: A. Falus

Received 28 May 2004, *accepted* 28 July 2004

(5) and dopamine and dopamine receptor genes that reinforce the effect of nicotine (6). Using sib-pair linkage analysis, Bergen and colleagues found a significant association between the ever–never smoking trait (EVRNVR) and four genomic regions, among which were two adjacent markers on chromosome 6 (7). These observations prompted us to investigate a possible association between smoking behavior and polymorphisms of MHC genes located on the short arm of chromosome 6. We have examined an association between coronary artery disease (CAD) and polymorphic alleles in the central region of MHC (8–10). In this study we employed healthy subjects and CAD patients from Hungary and Iceland with known smoking habits for genotypic and/or phenotypic analysis of complement components C4A and C4B, constituents of the *RP-C4-CYP21-TNX* (RCCX) modules, and single nucleotide polymorphisms at the promoter region of the tumor necrosis factor α gene (*TNFA*). Compelling evidence was obtained for a strong link between specific MHC haplotype(s) and cigarette smoking.

Methods

Study populations

The Hungarian study populations consisted of two groups. Group A consisted of 101 healthy subjects (29 males, 72 females; 44.0 ± 9.6 years old) (11). In these subjects we determined the complement *C4A* and *C4B* genotypes and analyzed the data together with C4A and C4B phenotypes that included protein polymorphisms in EDTA–plasma, quantitative variations in serum protein concentrations and hemolytic activities, as well as the −308 and −238 G to A promoter polymorphisms of the *TNFA* gene (12). Group B consisted of 232 patients with severe CAD (182 males, 50 females; 58.7 ± 8.2 years old). All CAD patients had symptoms of severe coronary atherosclerosis that included >50% stenosis, clinical signs of stable or unstable angina pectoris, and typical ECG abnormalities. These CAD patients received aorto-coronary by-pass graft by open heart surgery. Main characteristics of these Hungarian subjects including demographic properties, laboratory markers, association of CAD with polymorphism of mannose binding lectin, seropositivity to *Chlamydia pneumoniae*, and high levels of antibodies to the 60 kDa human heat shock protein were reported elsewhere (11–14). Only Caucasians with residence in Hungary were enrolled in the study. The Icelandic study population consisted of 246 healthy subjects (45 males, 201 females, 40.3 ± 10.6 years old) and 105 patients (55 males, 50 females, 54.3 ± 16.8 years old) who were admitted to the Landspitali University Hospital between 1995 and 1998 with chest pain which on further examination was found to be unrelated to cardiovascular diseases. Informed consents were obtained from all participants and the study was approved by the Ethic Committees of the Semmelweis University in Hungary and the Landspitali University Hospital in Iceland.

Registration of the smoking habits

Smoking behavior was registered by a physician at the study entry. Current smokers and those who had quit smoking were together defined as the group of 'ever smokers'. Ever smokers and never smokers were distinguished in both populations according to the criteria described by Bergen *et al.* (7) based on previously determined epidemiological thresholds (15,16). Only cigarette smokers were considered in the analysis. In the subgroup of the 232 Hungarian CAD patients, besides ever/never smoking, other smoking traits such as age at initiation, length of smoking, maximal intensity of cigarette smoking (expressed as number of packets/day) were also available.

Preparation of genomic DNA, Southern blot analysis and DNA probes

Total genomic DNA was extracted from white blood cells using proteinase K digestion, phenol extraction and ethanol precipitation (salt out) method (17). After removal of all personal identifiers, genomic DNA samples from the study populations were sent to the laboratory in Columbus for genotyping experiments, employing a protocol approved by the Institutional Review Board of the Columbus Children's Research Institute.

Southern blot analyses of *Taq*I-digested or *Psh*AI-digested genomic DNAs were performed as described previously (18,19) The *Taq*I genomic blots were hybridized with three specific probes corresponding to the 3′ regions of the *RP* gene that revealed the presence and dosage of *RP1-C4L* (7.0 kb), *RP1-C4S* (6.4 kb), *RP2-C4L* (6.0 kb) and *RP2-C4S* (5.4 kb); the presence and relative dosage of *CYP21B* (3.7 kb) and *CYP21A* (3.2 kb), and the presence and relative dosage of *TNXB* (2.5 kb) and *TNXA* (2.4 kb). The number of RCCX modules was independently confirmed by *Psh*AI RFLP using a RP 3′ probe to determine the relative dosage of *RP1* (7.2 kb) and *RP2* (8.3 kb) genomic fragments. The relative dosages of *C4A* and *C4B* genes were determined by *Psh*AI RFLP using a C4d specific probe corresponding to *C4* gene exons 28–31.

The *TNFA* −238 and −308 promoter polymorphisms were determined by genomic PCR using modified primer sequences and RFLP analyses, after published protocols (20).

C4 phenotyping

EDTA–plasma samples digested by neuraminidase and carboxyl peptidase B were resolved by high voltage agarose gel electrophoresis, subjected to immunofixation using goat anti-human C4 serum, washed to remove diffusible proteins, and stained to reveal the C4A and C4B protein allotypes (21,22).

HLA typing

HLA class I genes *A* and *B* and class II gene *DRB1* were typed by sequence-specific primer (SSP) PCR of genomic DNA, using commercial kits from One Lambda and standardized protocols.

Statistical analysis

The non-parametric Mann–Whitney test was used for group comparisons. Categorical data were compared using the Fisher's exact test. Multiple logistic regression was used to evaluate potential confounders. All tests were two-tailed. Statistical analysis was performed with GraphPad Prism 3.0 (GraphPad Software Inc., San Diego, CA; www.grphpad.com) and SPSS 10.0 (SPSS Inc., Chicago, IL) software.

Results

Relationship between some polymorphism of the central MHC genes and the ever/never smoking trait in Hungarian healthy people and CAD patients

The relationships between smoking behavior and polymorphisms of central MHC genes (*TNFA*, *C4A* and *C4B*; Fig. 1A) were analyzed in a Caucasian population consisting of 101 healthy individuals and 232 CAD patients residing in Hungary (12,24). Results obtained in this population are summarized in Fig. 1 (14). Among different individuals the number of complement component *C4A* genes present in a diploid genome varies from 0 to 5, and the size of a *C4* gene may be long (20.6 kb) or short (14.2 kb) (18,23) (Fig. 1A). There were marked differences in the prevalence of smoking habits among carriers of different numbers of the *C4A* gene (Fig. 1B). When subjects with 0 or 1 *C4A* gene were compared to those with 2 or more *C4A* genes, 44/56 (79%) of the first group and 143/277 (52%) of the second group were ever smokers ($P < 0.001$) (Fig. 1C). The differences were significant in both the healthy subjects ($P = 0.022$) and CAD patients ($P = 0.001$). A higher proportion of the CAD patients (156/232, 67%) than healthy individuals (31/101, 31%) was reported to be ever smokers. We also found significantly more subjects who ever smoked among the carriers than among the non-carriers of the phenotypic C4A*Q0 allele (70% vs 52%, respectively; $P = 0.003$; Fig. 1D).

With respect to the A allele of −308 *TNFA* promoter polymorphism (*TNF2*), a significantly higher proportion ($P < 0.001$) of the carriers (69/96, 72%) than non-carriers (109/221, 49%) were smokers (Fig. 1E). The difference was detected again in both the healthy subjects ($P = 0.002$) and CAD patients ($P = 0.019$).

There are also polygenic and gene size variations for complement *C4B* genes in the MHC (18,23). However, no significant differences in the smoking behavior were found between the carriers and non-carriers of the *C4B*Q0* allele (Fisher's exact test, $P = 0.169$), subjects with 0–1 versus 2 or more *C4B* genes ($P = 0.859$), number of short *C4* genes (χ^2 test, $P = 0.406$), and homozygous, heterozygous carriers and non-carriers of the −238A *TNFA* promoter allele ($P = 0.591$). A difference of modest significance ($P = 0.040$) was found with the number of long *C4* genes.

Relationship between a specific MHC III haplotype and the ever/never smoking trait in Hungarian healthy people and CAD patients

The absence of a *C4A* gene (*C4A*Q0*) together with the presence of a single short *C4B* gene (mono-S) in the MHC can be conveniently detected by a Southern blot analysis of *Taq*I digested genomic DNA, which appears as a distinct 6.4 kb fragment. In Caucasian populations the strong linkage disequilibrium between the mono-S haplotype with the *TNF2* allele within the MHC class III region on the short arm of chromosome 6 is a characteristic feature of the Caucasian ancestral haplotype (AH) 8.1 (25–27). However, as no factor B and C2 allotyping were done in this study, and HLA class I and class II typing was performed only in a limited number of subjects (see below), an individual with concurrent mono-S-RCCX (C4AQ0, mono-S C4B1) and TNF2 alleles in a diploid genome is not designated as having a probable AH8.1 but as a mono-S-RCCX-TNF2 haplotype carrier. When the frequencies of the ever/never smokers were compared between subjects with and without the *mono-S-RCCX-TNF2* haplotype, we observed remarkable differences (Table 1). *Mono-S-RCCX-TNF2* haplotype carriers were extremely likely to smoke ($P < 0.001$): all except eight (35/43, 81.4%) of the carriers but only 137/268 (51%) of the non-carriers were ever smokers. The differences were highly significant in both the healthy subjects ($P = 0.008$) and CAD patients ($P = 0.005$) (Table 1). In addition, we determined the MHC class I genes *A*, *B* and class II gene *DRB1* polymorphisms in 23 (14 smokers, 9 non-smokers) Hungarian CAD patients. Six smokers had 0 or 1 *C4A* gene. Among them, three had *TNF2* and mono-S, *C4A*Q0 C4B*1*. All of these three patients had at least one HLA-*A*1*, HLA-*B*8* and HLA-*DRB1*03* allele specific to AH8.1. Next, we compared the groups of ever smokers and never smokers for demographic and laboratory data that were registered in both the healthy subjects and CAD patients. Significantly more males than females smoked ($P < 0.001$). The smokers had significantly higher body mass index (BMI) ($P = 0.033$), serum total cholesterol ($P = 0.027$) and triglyceride ($P = 0.008$) levels. Therefore, we recalculated the association between smoking behavior and the tested gene polymorphisms using multiple logistic regression adjusted for these traits. Since the frequencies of smokers were markedly different in the groups of healthy subjects and CAD patients, the data were adjusted to these groups as well (Table 2). Adjusted odds ratios were significant for subjects with 0–1 versus 2 or more *C4A* alleles, carriers versus non-carriers of the *C4A*Q0* allele, carriers versus non-carriers of the *TNF2* allele, and the co-existence of *TNF2*, *C4A*Q0*, *C4B*1*-short. The highest odds ratio (95% CI) of 4.97 (1.958–12.618) ($P = 0.001$) was found between carriers and non-carriers of the *mono-S-RCCX-TNF2* haplotype.

Since we found highly significant interactions ($P = 0.002$ to 0.020) of sex with all genetic traits tested, the data were separately analyzed for males and females (Table 2). The association was stronger in the females than in the males for all traits tested, although the extent of sex-related difference was relatively low with respect to the *TNF2* allele. Again, the difference was especially striking for carriers of the mono-S-RCCX-TNF2 haplotype: female carriers of this haplotype were >13-fold more likely to become smokers than non-carriers, while in males an odds ratio of borderline significance was found.

*Lack of differences between the carriers and non-carriers of the TNF2, C4A*Q0, C4B*1 (short) haplotype in different smoking traits among Hungarian CAD patients*

In the group of CAD patients—where all of these data were available—different traits of smoking behavior were compared between groups of the carriers and non-carriers of the *mono-S-RCCX-TNF2* haplotype (Table 3). Most patients already stopped smoking at the time of by-pass operation. They initiated smoking at 20 years of age on average; used-to-smoke or active smokers had been smoking for an average of 30 years and smoked more than 20 cigarettes/day. Except the differences in the ever/never smoking already depicted

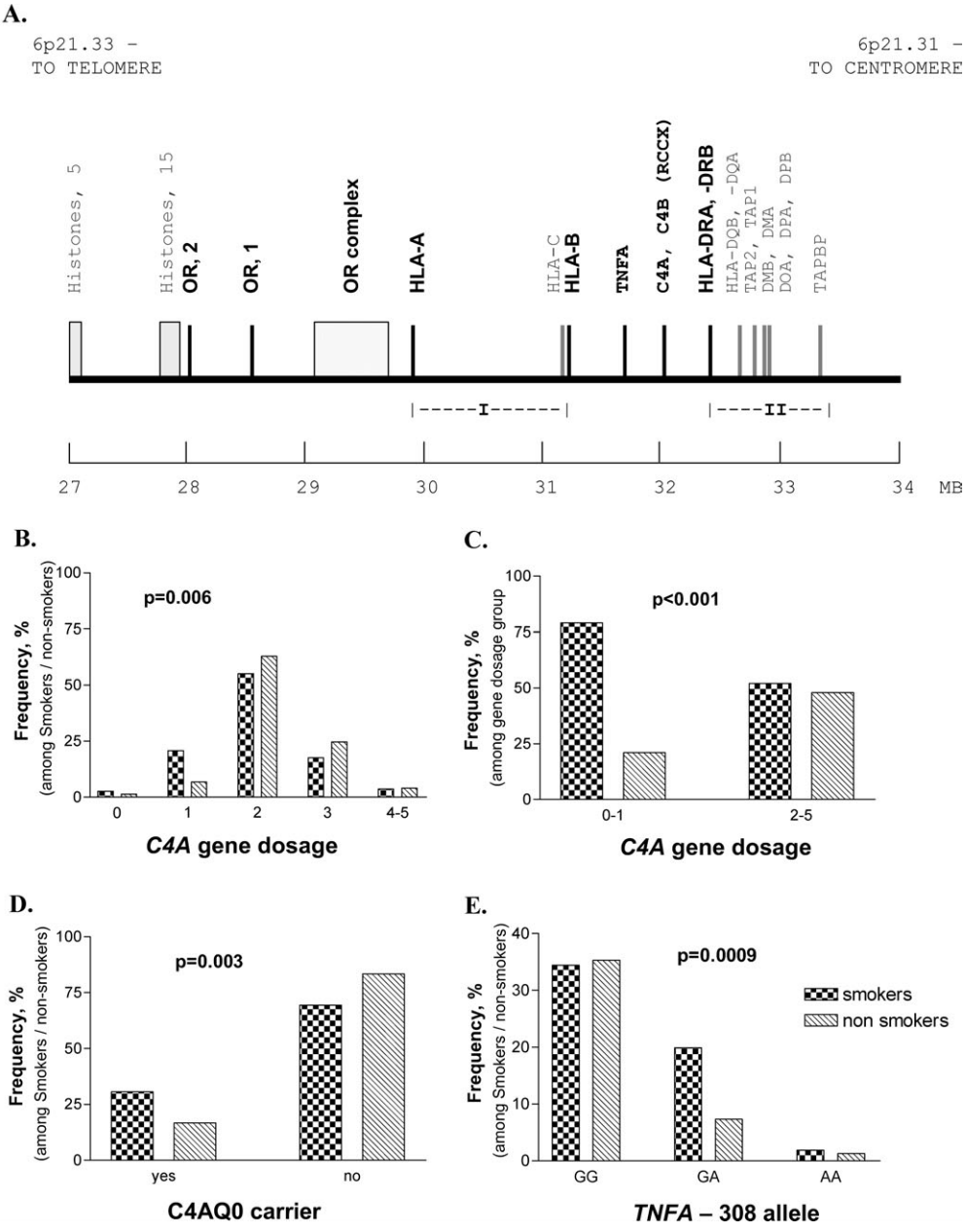

Fig. 1. Strong association of smoking behaviour with markers of the central MHC in healthy individuals and patients with CAD. (A) A simplified genomic map between chromosome 6p21.33 to 6p21.31 with the human MHC and olfactory receptor genes (OR). Numbering and scale were based on data from the NCBI human genome sequence database (http://www.ncbi.nlm.nih.gov/mapview/maps.cgi?ORG=hum&MAPS=ideogr,est,loc&LINKS=ON&VERBOSE=ON&CHR=6) (43). MB, megabase; I and II, MHC class I and class II genes; RCCX, genetic module consisting of serine/threonine nuclear protein kinase RP, complement C4, steroid CYP21 and tenascin TNX (10); numbers after histones and OR represent the numbers of genes present. (B) Frequency of subjects with different gene dosages of complement component *C4A* among the ever smokers or among the never smokers. (C) Frequency of ever smokers and never smokers among the subjects with 0–1 *C4A* genes or among the subjects with 2–5 *C4A* genes. (D) Frequency of subjects with carriers and non-carriers of the phenotypic C4AQ0 allele among the ever smokers or among the never smokers. (E) Frequency of subjects with carriers and non-carriers of *TNF2* (i.e. the A-allele at nucleotide −308 of the *TNFA* promoter region) among the ever smokers or among the never smokers. The healthy subjects and CAD patients were Caucasians from Hungary. The *P*-values for χ^2 test (B, C and E) or Fisher's exact test (D) are indicated.

above, no significant differences between the two groups (i.e. carriers and non-carriers of *mono-S-RCCX-TNF2* haplotype) were found in the proportion of current and past smokers, in the average age of initiation of regular smoking, duration of smoking and number of cigarettes smoked daily (Table 3).

*Differences in the proportion of the ever/never smokers between the Icelandic carriers and non-carriers of phenotypic C4A*Q0*

To validate our observation we investigated the smoking behavior in 351 subjects residing in Iceland (28). This

Table 1. Differences in the frequency of the *TNF2*, *C4A*Q0*, *C4B*1* (short) haplotype between Hungarian ever smokers (ES) and never smokers (NS)

Group	Smoking habit	No. (frequency) of non-carriers for mono-S-RCCX-TNF2 haplotype	No. (frequency) of carriers for mono-S-RCCX-TNF2 haplotype
Healthy individuals	NS	65 (0.765)	5 (0.385)
	ES	20 (0.235)	8 (0.615)
P-value (Fisher's exact test)		0.008	
Patients with CAD	NS	67 (0.361)	3 (0.100)
	ES	116 (0.639)	27 (0.900)
P-value (Fisher's exact test)		0.003	
All subjects	NS	132 (0.493)	8 (0.186)
	ES	136 (0.507)	35 (0.814)
P-value (Fisher's exact test)		<0.001	

Table 2. Logistic regression analysis* of the relationship between smoking habits and polymorphic variants of central MHC genes in 333 Hungarians (healthy individuals + patients with coronary artery disease) stratified according to sex

Polymorphism	Males		Females		All patients	
	Odds ratio (95% CI)	*P*-value	Odds ratio (95% CI)	*P*-value	Odds ratio (95% CI)	*P*-value
*C4A*Q0* carriers vs non-carriers	1.66 (0.81–3.39)	0.168	2.37 (1.01–5.54)	0.046	2.141 (1.18–3.88)	0.012
Carriers of 0–1 vs 2–5 *C4A* genes	2.62 (0.98–7.04)	0.055	6.98 (2.11–23.0)	0.001	4.131 (1.89–9.03)	<0.001
Carriers vs non-carriers of *TNF2* (−308A promoter allele of *TNFA*)	2.74 (1.24–6.06)	0.013	3.36 (1.39–8.14)	0.007	2.225 (1.34–3.69)	0.002
Carriers vs non-carriers of mono-S-RCCX-TNF2 haplotype	2.79 (0.95–8.22)	0.060	13.56 (2.56–71.9)	0.002	4.971 (1.96–12.6)	0.001

*Adjusted to BMI, serum cholesterol and triglyceride value as well as to group (healthy subjects versus patients).

Table 3. Lack of differences between the carriers and non carriers of the mono-S-RCCX-TNF2 haplotype in different smoking traits among Hungarian patients with CAD

	Ever/never smokers	Present/past smokers	Age at start of smoking, mean ± SD	Length of smoking, years, mean ± SD	Maximum cigarettes/day, mean ± SD
Carriers	27/3	3/24	20.2 ± 9.6	31.3 ± 9.8	24.8 ± 11.6
Non-carriers	117/66	14/103	19.0 ± 5.1	29.1 ± 11.9	23.4 ± 11.6
P-value*	0.003	1.000	0.574	0.340	0.600

*Fisher's exact test for contingency tables; two-sample *t*-test for the continuous variable.

population was tested only by C4 phenotyping. The risk to become a smoker was found to be significantly higher for C4A*Q0 carriers than non-carriers and the difference was more pronounced in females than in males (Table 4). As expected, more males (64/100, 64%) than females (115/251, 46%) had ever smoked ($P = 0.003$). The odds ratio of smoking among C4A*Q0 carriers versus non-carriers calculated by age- and sex-adjusted multiple logistic regression was 1.801 (1.057–3.070; $P = 0.031$). The association of the deficient C4A allele with smoking was stronger in females [OR: 1.926 (1.014–3.659), $P = 0.045$] than in males [OR: 1.511 (0.579–3.947); $P = 0.399$]. In contrast, there was no significant relationship ($P = 0.893$) between carriers and non-carriers of the *C4B*Q0* allele with cigarette smoking.

The lower odds ratio of the C4A*Q0 group for the increased risk of smoking in the Icelandic subjects (than those of the two Hungarian study populations) was probably due to the lower resolution of the C4 phenotyping method that sometimes does

Table 4. Differences in frequencies of the apparent C4A*Q0 phenotypes in the male and female Icelandic ever smokers and never smokers.

Group	Smoking habit	No. (frequency) of non-carriers of C4A*Q0	No. (frequency) of carriers of C4A*Q0
Males	Never smoked	36 (0.450)	9 (0.310)
	Ever smoked	44 (0.550)	20 (0.690)
P-value (χ^2 analysis)		0.191	
Females	Never smoked	116 (0.571)	20 (0.417)
	Ever smoked	87 (0.429)	28 (0.583)
P-value (χ^2 analysis)		0.053	
All subjects	Never smoked	152 (0.537)	29 (0.377)
	Ever smoked	131 (0.463)	48 (0.623)
P-value (χ^2 analysis)		0.013	

not yield accurate information on subjects with high *C4* gene dosages (18).

Discussion

Our present findings indicate that in both healthy and CAD study populations, subjects with mono-S-RCCX-TNF2 haplotype had a significant, 5-fold higher risk of becoming smokers compared to those without the haplotype. Strikingly, such association is more pronounced in women, as a >13-fold increased risk was observed in the female cohorts.

Odds ratios of the mono-S-RCCX-TNF2 haplotype carriers to become a regular smoker found in the present study are significantly higher than any other odds ratios for candidate gene association studies with smoking behavior reported so far (1). The highest odds ratio reported previously on a candidate gene associated with smoking was on the *DRD2* polymorphism of the dopaminergic receptor gene; odds ratio 2.71 [(2.05–3.58), $P < 0.001$] (30). The relatively large number of subjects investigated in this study ($n = 684$) allows a statistical power to demonstrate a genetic basis of cigarette smoking in Caucasians with high confidence. By contrast, no associations between these alleles and other traits of smoking behavior, such as the age of starting smoking and the duration and intensity of smoking, were found. This observation is in complete accordance with other studies indicating that genetic influences on different stages of smoking such as initiation, maintenance and cessation may not be identical (1,29).

In Caucasian populations, the presence of a solitary short *C4B* gene coding for a C4B1 allotype accompanied with a phenotypic deficiency of a C4A protein, together with the presence of the *TNF2* allele in the MHC is a characteristic feature of the ancestral haplotype with HLA *A1*, *B8*, *DR3* (AH8.1). This remarkable haplotype has a frequency between 8.6 and 11% in healthy Caucasian populations (14,18), but its frequency is significantly increased in a variety of HLA-associated diseases such as systemic lupus erythematosus (32), type 1 diabetes mellitus, IgA deficiency and common variable immunodeficiency (33–35).

To our knowledge, this is the first molecular genetic study illustrating a very strong association of smoking behavior with any polymorphisms of the MHC. Several genome-wide screening studies that suggest evidence for linkage of nicotine addiction to different chromosome regions have been published (2,4,7,31) The data for linkage studies have been inconclusive so far. However, our present findings are in line with the observations of Bergen *et al.* (7), who found a significant association between the ever smoking trait and, among other regions, with markers on chromosome 6.

No data are available that indicate that heterozygous or complete deficiency of complement C4A, or the presence of *TNF2* alleles *per se* contributes to higher risks of smoking. It can be assumed that carriers of 8.1AH have an increased risk of becoming smokers. On the other hand, since we did not perform factor B and C2 phenotyping and HLA class I and class II typing were performed only in a limited number of individuals, at the present time this assumption can be considered only as a hypothesis. Irrespectively of the haplotype involved, no definite explanation for our present findings can be given at the present time. It is tempting to speculate, however, that the observed association between the mono-S-RCCX-TNF2 haplotype or 8.1AH and ever smoking is related to the presence of a cluster of genes of olfactory receptors in close proximity to the MHC. One of the chemosensory systems is the main olfactory epithelium (MOE), which discriminates odors from the environment (36). There are about 5 million olfactory sensory neurons, each of which expresses one odorant receptor (OR) with a 7-transmembrane domain. The human genome contains about ~1000 *OR* genes and one of the clusters of MOE-expressed *OR* genes is located telomeric to the MHC class I region (Fig. 1A) (39). Alleles from these *OR* genes are in strong linkage disequilibrium with each other and also with specific MHC haplotypes (37). Ehlers *et al.* (40) defined 13 HLA-lined *OR* haplotypes and demonstrated that allelic variation is a general feature of the HLA-linked *OR* genes.

On the other hand, Perkins and coworkers suggested that subjective hedonic and reinforcing effects of smoking behavior were influenced by olfactory/taste stimuli (38). They also noted that the olfactory/taste stimuli-dependent effects of cigarette smoking were stronger in women than in men. By contrast, nicotine itself is less reinforcing on smoking in women than in men (41).Those findings are in line with the sex-related difference found in the present study: we observed a stronger association between the mono-S-RCCX-TNF2 haplotype and smoking habit in women than in man. Therefore, it would be exciting to determine whether specific variants of the MHC-linked *OR* genes (40) contributed to higher risk for initiation of smoking. To this end, it is of interest to note that the MHC has also been implicated for a role in mate choice or sexual attraction (42,43), which is in keeping with the notion that MHC-linked OR genes would be relevant in olfactory-based discrimination (44) and nicotine addiction. Further studies on the genetics and functions of the MHC-linked *OR* genes, and the interaction of OR with other biological and environmental factors, would yield insight into genetic contributions to smoking habits, which can potentially lead to more effective policies or strategies to reduce smoking and to improve public health.

Acknowledgements

We are indebted to the blood donors for their participation in this study. We thank Dr Margrét Leopoldsdóttir, Dr Ragnhildur Magnúsdóttir and Professor Helgi Valdimarsson for their contribution. This work was supported by research grants from the Hungarian Academy of Sciences (2000-107 3,2), Hungarian Ministry of Health (ETT 196/2003), Hungarian National Research Fund (T032661) (GF), the ATHERNET (QLG1-CT-2002-90397) grant of the European Commission, the Icelandic Research Council grant 011330001 and the Science Fund of Landspitali University Hospital, Iceland (GJA), and the US National Institutes of Health grants NIAMS 1R01 AR050078 and NIDDK P01 DK55546 (C.Y.Y.).

Abbreviations

BMI	body mass index
CAD	coronary artery disease
MOE	main olfactory epithelium
OR	odorant receptor
RCCX	four consecutive genes coding for serine/threonine nuclear kinase protein RP, complement component C4, cytochrome P450 CYP21, extracellular matrix protein tenascin TNX
SSP	sequence-specific primer
TNF2	−308A allele of the tumor necrosis factor α gene *TNFA*

References

1 Batra, V., Patkar, A. A., Berrettini, W. H., Weinstein, S. P. and Leone, F. T. 2003. The genetic determinants of smoking. *Chest* 123: 1730.

2 Sullivan, P. F. and Kendler, K. S. 1999. The genetic epidemiology of smoking. *Nicotine. Tob. Res.* 1 (Suppl. 2) :S51.

3 Heath, A. C., Kirk, K. M., Meyer, J. M. and Martin, N. G. 1999. Genetic and social determinants of initiation and age at onset of smoking in Australian twins. *Behav. Genet.* 29:395.

4 Kendler, K. S., Thornton, L. M. and Pedersen, N. L. 2000. Tobacco consumption in Swedish twins reared apart and reared together. *Arch. Gen. Psychiatry* 57:886.

5 Pianezza, M. L., Sellers, E. M. and Tyndale, R. F. 1998. Nicotine metabolism defect reduces smoking. *Nature* 393:750.

6 Noble, E. P., St Jeor, S. T., Ritchie, T., Syndulko, K., St Jeor, S. C., Fitch, R. J., Brunner, R. L. and Sparkes, R. S. 1994. D2 dopamine receptor gene and cigarette smoking: a reward gene? *Med. Hypotheses* 42:257.

7 Bergen, A. W., Korczak, J. F., Weissbecker, K. A. and Goldstein, A. M. 1999. A genome-wide search for loci contributing to smoking and alcoholism. *Genet. Epidemiol.* 17 (Suppl. 1) :S55.

8 Kumanovics, A., Takada, T. and Lindahl, K. F. 2003. Genomic organization of the mammalian MHC. *Annu. Rev. Immunol.* 21:629.

9 Martinez, O. P., Longman-Jacobsen, N., Davies, R., Chung, E. K., Yang, Y., Gaudieri, S., Dawkins, R. L. and Yu, C. Y. 2001. Genetics of human complement component C4 and evolution the central MHC. *Front Biosci.* 6:D904.

10 Yu, C. Y., Chung, E. K., Yang, Y. *et al.* 2003. Dancing with complement C4 and the RP-C4-CYP21-TNX (RCCX) modules of the major histocompatibility complex. *Prog. Nucleic Acid Res. Mol. Biol.* 75:217.

11 Burian, K., Kis, Z., Virok, D. *et al.* 2001. Independent and joint effects of antibodies to human heat-shock protein 60 and *Chlamydia pneumoniae* infection in the development of coronary atherosclerosis. *Circulation* 103:1503.

12 Vatay, A., Yang, Y., Chung, E. K. *et al.* 2003. Relationship between complement components C4A and C4B diversities and two TNFA promoter polymorphisms in two healthy Caucasian populations. *Hum. Immunol.* 64:543.

13 Rugonfalvi-Kiss, S., Endresz, V., Madsen, H. O. *et al.* 2002. Association of *Chlamydia pneumoniae* with coronary artery disease and its progression is dependent on the modifying effect of mannose-binding lectin. *Circulation* 106:1071.

14 Yang, Y., Chung, E. K., Zhou, B. *et al.* 2003. Diversity in intrinsic strengths of the human complement system: serum C4 protein concentrations correlate with C4 gene size and polygenic variations, hemolytic activities and body mass index. *J. Immunol.* 171:2734.

15 Anda, R. F., Williamson, D. F., Escobedo, L. G., Mast, E. E., Giovino, G. A. and Remington, P. L. 1990. Depression and the dynamics of smoking. A national perspective. *JAMA* 264:1541.

16 Glassman, A. H., Helzer, J. E., Covey, L. S., Cottler, L. B., Stetner, F., Tipp, J. E. and Johnson, J. 1990. Smoking, smoking cessation and major depression. *JAMA* 264:1546.

17 Miller, S. A., Dykes, D. D. and Polesky, H. F. 1988. A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res.* 16:1215.

18 Blanchong, C. A., Zhou, B., Rupert, K. L., Chung, E. K., Jones, K. N., Sotos, J. F., Zipf, W. B., Rennebohm, R. M. and Yung, Y. C. 2000. Deficiencies of human complement component C4A and C4B and heterozygosity in length variants of RP-C4-CYP21-TNX (RCCX) modules in caucasians. The load of RCCX genetic diversity on major histocompatibility complex-associated disease. *J. Exp. Med.* 191:2183.

19 Chung, E. K., Yang, Y., Rupert, K. L., Jones, K. N., Rennebohm, R. M., Blanchong, C. A. and Yu, C. Y. 2002. Determining the one, two, three, or four long and short loci of human complement C4 in a major histocompatibility complex haplotype encoding C4A or C4B proteins. *Am. J. Hum. Genet.* 71:810.

20 Day, C. P., Grove, J., Daly, A. K., Stewart, M. W., Avery, P. J. and Walker, M. 1998. Tumour necrosis factor-alpha gene promoter polymorphism and decreased insulin resistance. *Diabetologia* 41:430.

21 Awdeh, Z. L., Raum, D., Yunis, E. J. and Alper, C. A. 1983. Extended HLA/complement allele haplotypes: evidence for T/t-like complex in man. *Proc. Natl Acad. Sci. USA* 80:259.

22 Sim, E. and Cross, S. J. 1986. Phenotyping of human complement component C4, a class-III HLA antigen. *Biochem. J.* 239:763.

23 Chung, E. K., Yang, Y., Rennebohm, R. M., Lokki, M. L., Higgins, G. C., Jones, K. N., Zhou, B., Blanchong, C. A. and Yu, C. Y. 2002. Genetic sophistication of human complement components C4A and C4B and RP-C4-CYP21-TNX (RCCX) modules in the major histocompatibility complex. *Am. J. Hum. Genet.* 71:823.

24 Szalai, C., Fust, G., Duba, J., Kramer, J., Romics, L., Prohaszka, Z. and Csaszar, A. 2002. Association of polymorphisms and allelic combinations in the tumour necrosis factor-alpha-complement MHC region with coronary artery disease. *J. Med. Genet.* 39:46.

25 Awdeh, Z. L. and Alper, C. A. 1980. Inherited structural polymorphism of the fourth component of human complement. *Proc. Natl Acad. Sci. USA* 77:3576.

26 Dawkins, R., Leelayuwat, C., Gaudieri, S., Tay, G., Hui, J., Cattley, S., Martinez, P. and Kulski, J. 1999. Genomics of the major histocompatibility complex: haplotypes, duplication, retroviruses and disease. *Immunol. Rev.* 167:275.

27 Wilson, A. G., de Vries, N., Pociot, F., di Giovine, F. S., van der Putte, L. B. and Duff, G. W.1993. An allelic polymorphism within the human tumor necrosis factor alpha promoter region is strongly associated with HLA A1, B8 and DR3 alleles. *J. Exp. Med.* 177:557.

28 Arason, G. J., Bodvarsson, S., Sigurdarson, S. T., Sigurdsson, G., Thorgeirsson, G., Gudmundsson, S., Kramer, J. and Fust, G. 2003. An age-associated decrease in the frequency of C4B*Q0 indicates that null alleles of complement may affect health or survival. *Ann. N.Y. Acad. Sci.* 1010:496.

29 Kendler, K. S., Neale, M. C., Sullivan, P., Corey, L. A., Gardner, C. O. and Prescott, C. A. 1999. A population-based twin study in women of smoking initiation and nicotine dependence. *Psychol. Med.* 29:299.

30 Comings, D. E., Ferry, L., Bradshaw-Robinson, S., Burchette, R., Chiu, C. and Muhleman, D. 1996. The dopamine D2 receptor (DRD2) gene: a genetic risk factor in smoking. *Pharmacogenetics* 6:73.

31 Straub, R. E., Sullivan, P. F., Ma, Y. *et al.* 1999. Susceptibility genes for nicotine dependence: a genome scan and followup in an independent sample suggest that regions on chromosomes 2, 4, 10, 16, 17 and 18 merit further study. *Mol. Psychiatry* 4:129.

32 Yang, Y., Chung, E. K., Zhou, B., Lhotta, K., Hebert, L. A., Birmingham, D. J., Rovin, B. H. and Yu, C. Y. 2004. The intricate role of complement component C4 in human systemic lupus erythematosus. *Curr. Dir. Autoimmun.* 7:98.

33 Alper, C. A., Marcus-Bagley, D., Awdeh, Z. *et al.* 2000. Prospective analysis suggests susceptibility genes for deficiencies of IgA and several other immunoglobulins on the [HLA-B8, SC01, DR3] conserved extended haplotype. *Tissue Antigens* 56:207.

34 Candore, G., Lio, D., Colonna, R. G. and Caruso, C. 2002. Pathogenesis of autoimmune diseases associated with 8.1 ancestral haplotype: effect of multiple gene interactions. *Autoimmun. Rev.* 1:29.

35 Price, P., Witt, C., Allcock, R., Sayer, D., Garlepp, M., Kok, C. C., French, M., Mallal, S. and Christiansen, F. 1999. The genetic basis for the association of the 8.1 ancestral haplotype (A1, B8, DR3) with multiple immunopathological diseases. *Immunol. Rev.* 167:257.

36 Buck, L. B. 2000. The molecular architecture of odor and pheromone sensing in mammals. *Cell* 100:611.

37 Zhao, H., Ivic, L., Otaki, J. M., Hashimoto, M., Mikoshiba, K. and Firestein, S. 1998. Functional expression of a mammalian odorant receptor. *Science* 279:237.

38 Perkins, K. A., Gerlach, D., Vender, J., Grobe, J., Meeker, J. and Hutchison, S. 2001. Sex differences in the subjective and reinforcing effects of visual and olfactory cigarette smoke stimuli. *Nicotine Tob. Res.* 3:141.

39 Younger, R. M., Amadou, C., Bethel, G. *et al.* 2001. Characterization of clustered MHC-linked olfactory receptor genes in human and mouse. *Genome Res.* 11:519.

40 Ehlers, A., Beck, S., Forbes, S. A., Trowsdale, J., Volz, A., Younger, R. and Ziegler, A. 2000. MHC-linked olfactory receptor loci exhibit polymorphism and contribute to extended HLA/OR-haplotypes. *Genome Res.* 10:1968.

41 Perkins, K. A., Jacobs, L., Sanders, M. and Caggiula, A. R. 2002. Sex differences in the subjective and reinforcing effects of cigarette nicotine dose. *Psychopharmacology (Berl)* 163:194.
42 Jacob, S., McClintock, M. K., Zelano, B. and Ober, C. 2002. Paternally inherited HLA alleles are associated with women's choice of male odor. *Nat. Genet.* 30:175.
43 Yamaguchi, M., Yamazaki, K., Beauchamp, G. K., Bard, J., Thomas, L. and Boyse, E. A. 1981. Distinctive urinary odors governed by the major histocompatibility locus of the mouse. *Proc. Natl Acad. Sci. USA* 78:5817.
44 Doherty, P. C. 2003. On the nose: shared themes for the sensory and immune self. *Nat. Immunol.* 4:1043.