| T | A | N | 1 | 1 | 3 | 3 | B | Dispatch: 21.8.08 | Journal: TAN | CE: M. Mohana Priya |
|---|---|---|---|---|---|---|---|---|---|---|
| Journal Name | | | | Manuscript No. | | | | Author Received: | No. of pages: 4 | ME: Senthil |

Tissue Antigens ISSN 0001-2815

BRIEF COMMUNICATION

# The CR1 C5507G polymorphism is not involved in susceptibility to idiopathic pulmonary fibrosis in two European populations

Z. Kubistova[1], F. Mrazek[1], P. A. Lympany[2], A. L. Lagan[2], A. Arakelyan[1], E. Kriegova[1], K. I. Welsh[2], V. Kolek[1], J. Zatloukal[1], B. Hutyrova[1], R. M. du Bois[2,3] & M. Petrek[1]

1 Palacky University, Olomouc, Czech Republic
2 Royal Brompton Hospital, London, UK
3 National Jewish Medical and Research Center, Denver, CO, USA

**Key words**
Case–control study; complement receptor 1; European population; gene polymorphism; genetics of complex diseases; idiopathic pulmonary fibrosis

**Correspondence**
Martin Petrek, MD
Laboratory of Immunogenomics
Department of Immunology
Palacky University
I. P. Pavlova Str. 6
CZ-775 20 Olomouc
Czech Republic
Tel: +420 58 563 2770
Fax: +420 58 541 5116
e-mail: petrekm@fnol.cz

Received 27 March 2008; revised 16 July 2008; accepted 5 August 2008

doi: 10.1111/j.1399-0039.2008.01133.x

## Abstract

Idiopathic pulmonary fibrosis (IPF), a severe lung disease with unknown aetiology, is thought to have an important genetic component. Single nucleotide polymorphism, C5507G, of the complement receptor 1 (*CR1*) gene, which affects the number of CR1 molecules on erythrocytes, has been associated with susceptibility to IPF in a single European population. To replicate this finding, 53 Czech IPF patients with 203 Czech healthy control subjects and 70 English IPF patients with 149 English controls were investigated. In both populations, there were no significant differences in distribution of CR1 C5507G variants between IPF patients and their appropriate control groups. In conclusion, the association of the CR1 C5507G polymorphism with susceptibility to IPF was not reproducible in Czech and English populations.



Idiopathic pulmonary fibrosis (IPF) is a severe lung disease with unknown cause and unfavourable prognosis, which is characterised by the usual interstitial pneumonia histopathological pattern (1). The current paradigm of IPF pathogenesis is based on the concept of aberrant wound healing in response to, still undefined, injury of the lung (2, 3). It has been suggested that genetic factors may participate in susceptibility to this disease or modify its clinical course. Accordingly, several such genetic variants have been investigated with limited success (4, 5).

Zorzetto et al. have previously reported an association of a functional complement receptor 1 (CR1) single nucleotide polymorphism (SNP), namely CR1 C5507G (rs3811381), with susceptibility to both sarcoidosis (6) and to IPF (7) in Italians. The CR1 C5507G SNP has been shown to have dual character in terms of CR1 allelic structural isoforms: it is located within exon 33 (isoform F) or exon 41 (isoform S) of the *CR1* gene. Indeed, the dual character of CR1 C5507G has emphasised important methodical consequences in a study of a Finnish population (8, 9). Since the original Italian study (7), however, no further confirmatory data on the implication of *CR1* gene variants in the pathogenesis of IPF have been published. In agreement with current rules for replication of genetic association studies (10) we, therefore, investigated if the CR1 C5507G SNP is associated with IPF in two well-characterised Caucasian (Czech and English) groups of IPF patients, together with ethnically matched healthy control subjects.

In total, 475 unrelated Caucasian subjects were enrolled into this case–control study – 53 Czech IPF patients [age, median (range): 53 (22–72); males/females: 36/17] with 203 Czech healthy control subjects [adapted from the study by

Mrazek et al. (11)] and 70 English IPF patients [age, median (range): 63 (47–77); males/females: 46/24; current smoker/former smoker/never smoker/unknown: 5/14/47/4; lung functional parameters, FEV (average % predicted ± SD): 81.5 ± 20, FVC: 81.3 ± 21, $DL_{CO}$: 47.2 ± 17] with 149 3 English controls. The diagnosis of IPF was based on typical clinical and radiological features together with the histopathological appearance of UIP on surgical lung biopsy and was compatible with the ATS/ERS consensus criteria for the diagnosis of IPF (1). Patients were recruited and diagnosed in a single tertiary/quaternary referral centre for each population (Department of Respiratory Medicine, Faculty Hospital, Olomouc, Czech Republic and Interstitial Lung Disease Unit, Royal Brompton Hospital, London, UK). The control populations consisted of gender-matched blood donors or healthy participants of the bone marrow donor registry, recruited from the same regions as patients to ensure ethnic homogeneity. Absence of lung disease in the control subjects was checked by health questionnaire and interview emphasising family history and symptoms of respiratory disease. All subjects agreed to the anonymous use of their biological samples for the purposes of the study, which was performed with the approval of the ethical committees of the Medical Faculty in Olomouc and Royal Brompton Hospital in London. The data on the Czech control group including the frequencies of CR1 C5507G variants have already been published elsewhere (11).

In all study subjects, genotyping was performed by polymerase chain reaction with sequence-specific primers as described previously (11). The sequences of specific primers for allele discrimination were as follows: 5′ATCCGCTGCACAAGTGACCC with 5′CCCTACTAAATCTGGACCTCATC for standard allele and 5′ATCCGCTGCACAAGTGACCG with 5′ CCCTACTAAATCTGGACCTCATC for mutant allele.

The proportion of CR1 C5507G genotypes was determined by direct counting. Allele frequencies and carriage rates (phenotype frequencies) were calculated for each investigated group. Distribution of genotypes for CR1 C5507G SNP was tested for compliance with Hardy–Weinberg equilibrium using chi-squared test. Genotype, allele and phenotype frequencies were compared by chi-squared test between the groups of IPF patients and control population and within the IPF subgroups. Statistical power of the study was calculated according to Lalouel and Rohrwasser (12). The relationship between the CR1 C5507G genotypes and the continuous variables (age and pulmonary functions) was performed by regression analysis (SPSS software v. 15.0).

To determine the distribution of the CR1 C5507G SNP in patients with IPF by comparison with the control population, Czech and English Caucasian IPF patients with ethnically matched healthy control subjects were genotyped. The distribution of CR1 C5507G genotypes complied with Hardy–Weinberg equilibrium in all investigated groups. The statistical power of the present study to replicate the main observation of the original (7) report (difference between the frequency of CR1 5507 GG homozygotes in IPF patients and healthy control subjects) reached 92.3% for the Czech population and 94.8% for the English population.

Although Zorzetto et al. (7) observed a clear overrepresentation of CR1 5507 GG homozygotes among the IPF patients (16%) by comparison with healthy control subjects (3%) in an Italian population, CR1 5507 GG homozygotes were similarly represented in all of our four groups (Czech/English IPF patients and Czech/English controls) and the proportion was no higher than 3% in any group (Table 1). Furthermore, there was no difference between the frequency of the CR1 5507*G allele in IPF patients and healthy controls in both Czech and English populations (CR1 5507*G allelic frequencies; Czech IPF patients: 21%; Czech controls: 18%; English IPF patients: 14%; English controls: 15%). In addition, the proportion of CR1 5507*G allele carriers did not differ between IPF patients and healthy controls (Table 1). Additional comparison revealed significant difference in distribution of CR1 C5507G genotypes between our groups of Czech and English IPF patients ($P = 0.03$, 2 d.f., including 'zero' cell). Nevertheless, this difference is contributed mostly by the moderately different population background of our two IPF groups – overall frequency of CR1 seemed to be slightly lower in the whole English population (including control group) by comparison with Czech population.

Finally, we performed further subanalysis within the groups of IPF patients to show any possible relationship of the CR1 C5507G SNP with specific IPF phenotypes. In both Czech and English IPF populations, no association of CR1 C5507G variants with gender or with age was observed ($P > 0.05$ for all comparisons). Similarly, no differences in distribution of CR1 C5507G variants were observed when English IPF patients were stratified according to their smoking habit and pulmonary function. The latter subanalyses could not be performed in the Czech IPF group because of the lack of the complete data.

In summary, CR1 C5507G variants were equally represented both in IPF patients and ethnically matched healthy control subjects from two tertiary (Czech and English) centres. Hence, the present data based on two population cohort design do not confirm the findings from Italian population (7) and do not support the hypothesis that CR1 C5507G variant is directly involved in the susceptibility to IPF.

Current biological concepts of IPF do not consider immune complexes (IC) as essential factors in the pathogenesis of this disease (2, 3). Accordingly, the CR1 gene that encodes the IC 'scavenger' molecule may not be regarded as a classical 'candidate' potentially contributing to genetic

**Table 1** Genotype frequencies, allele frequencies and carriage rates of the CR1 C5507G SNP in Czech, English and Italian populations shown in absolute (relative) values

| Population | Czech ($n = 256$) | | English ($n = 219$) | | Italian ($n = 240$) | |
|---|---|---|---|---|---|---|
| CR1 C5507G | Controls | IPF patients | Controls | IPF patients | Controls | IPF patients |
| *n* | 203 | 53 | 149 | 70 | 166 | 74 |
| Genotype frequencies | | | | | | |
| CC | 133 (0.66) | 31 (0.58) | 106 (0.71) | 53 (0.76) | 111 (0.67) | 35 (0.47) |
| CG | 66 (0.32) | 22 (0.42) | 41 (0.28) | 15 (0.21) | 50 (0.30) | 27 (0.37) |
| GG | 4 (0.02) | 0 (0.00) | 2 (0.01) | 2 (0.03) | 5 (0.03) | 12 (0.16) |
| *P* values* | | | | | | |
| GG genotypes | 0.30 | | 0.43 | | 0.0002 | |
| Allele frequencies | | | | | | |
| C | 332 (0.82) | 84 (0.79) | 253 (0.85) | 121 (0.86) | 272 (0.82) | 97 (0.66) |
| G | 74 (0.18) | 22 (0.21) | 45 (0.15) | 19 (0.14) | 60 (0.18) | 89 (0.34) |
| Carriage rates | | | | | | |
| C | 199 (0.98) | 53 (1.00) | 147 (0.99) | 68 (0.97) | 161 (0.97) | 62 (0.84) |
| G | 70 (0.34) | 22 (0.42) | 43 (0.29) | 17 (0.24) | 55 (0.33) | 39 (0.53) |

* *P* values for comparison of the proportion of CR1 5507 GG homozygotes between the groups of IPF patients and ethnically matched healthy control subjects of particular population. The data for an Italian population have been adapted from the article by Zorzetto et al. (7).

susceptibility in IPF. Nevertheless, it is recognised that variants of genes not considered as primary candidates for disease pathogenesis have recently been identified and subsequently verified as true susceptibility markers in 'hypothesis free' whole genome association studies for complex diseases (13).

It has been repeatedly suggested that studies implicating genetic variants in the susceptibility to complex diseases should be independently replicated (10). This recommendation is in line with empirical observations that only a minority of original studies reporting association of gene polymorphism with complex diseases has been verified in subsequent studies. In general, there may be several potential explanations for such discrepancies, including differences in disease phenotype definition and homogeneity, type I error (false-positive data because of random fluctuation), linkage disequilibrium of the marker with the true causative variant limited to particular ethnic populations or stratification in selection of control subjects (14). In this context, we acknowledge that the difference between the median age of the groups of Czech IPF patients and control subjects (53 *vs* 33 years) and a lack of the data on age distribution in the English control group may be considered as a limitation of the present study. Nevertheless, because the incidence of IPF is extremely low in the general population and, based on numerous previous reports showing no association of CR1 C5507G with age (6, 7), the probability that our results have significantly been confounded by the differences in age distribution between the cases and controls is very low.

Although the consensus criteria for the diagnosis of IPF are generally accepted (1), some authors speculate that there is still a possibility that IPF may consist of several disease entities, which leads to the same or similar clinical (histopathological) manifestation (5). For example, the familial form of pulmonary fibrosis may be separated from sporadic IPF based on the basis of the epidemiological definition (4, 5). In this regard, familial lung fibrosis has been reported to be associated with mutations in the surfactant protein C gene (15), but these variants likely do not contribute to the susceptibility to the sporadic form of IPF (16). Applying this concept to our study, it is possible that the Italian group of IPF patients manifested a different pattern of disease phenotype than in our cohorts, which may be one of the most plausible explanations for the discordant results described. Hypothetically, *CR1* gene variants may confer susceptibility to one of the environmental (infectious?) IPF triggers to which the Italian population is exposed. However, because both studies recruited IPF patients (phenotypes) according to the same criteria (1), this speculation cannot be confirmed until further detailed disease stratification according to the phenotype and 'aetiology' (yet unknown) can be applied.

With regard to the possibility of randomly false (positive or negative) findings, the strength of the present study lies in its two-cohort design, which substantially reduces the possibility that false-negative data have been obtained, especially when both our populations had sufficient statistical power to replicate the original findings (7). However, we limited our investigation to the CR1 C5507G SNP, although Zorzetto et al. investigated three but strongly linked, SNPs and performed haplotype analysis. Nevertheless, haplotype analysis from our previous study of CR1 polymorphisms and sarcoidosis in a Czech population (11) revealed almost complete linkage disequilibrium between the less common alleles of the CR1

coding variants (CR1 5507*G–3650*G), which is consistent with the finding in the Italian and other populations (7, 17).

In conclusion, despite adequate statistical power, the present study observed no association between the CR1 C5507G SNP and susceptibility to IPF in two Caucasian (Czech and English) populations. Our investigations, therefore, do not provide further evidence for the concept that CR1 C5507G is involved in the genetic predisposition to IPF.

## Acknowledgments

The authors acknowledge the technical help of J. Onderkova, S. Zachova and A. Vevodova. This work was supported in part by the Internal Grant Agency of the Ministry of Health, Czech Republic (IGA NR9037) and the Czech government (MSYCR-MSM6198959205). The findings of this study were presented in part at the 1st Joint Meeting of European National Society of Immunology and 16th European Congress of Immunology, 6–9 September 2006, Paris, France.

## References

1. American Thoracic Society/European Respiratory Society International Multidisciplinary Consensus Classification of the Idiopathic Interstitial Pneumonias. *Am J Respir Crit Care Med* 2002: **165**: 277–304.
2. Hunninghake GW, Schwarz MI. Does current knowledge explain the pathogenesis of idiopathic pulmonary fibrosis? A perspective. *Proc Am Thorac Soc* 2007: **4**: 449–52.
3. Studer SM, Kaminski N. Towards systems biology of human pulmonary fibrosis. *Proc Am Thorac Soc* 2007: **4**: 85–91.
4. Lawson WE, Loyd JE. The genetic approach in pulmonary fibrosis: can it provide clues to this complex disease? *Proc Am Thorac Soc* 2006: **3**: 345–9.
5. Grutters JC, du Bois RM. Genetics of fibrosing lung diseases. *Eur Respir J* 2005: **25**: 915–27.
6. Zorzetto M, Bombieri C, Ferrarotti I et al. Complement receptor 1 gene polymorphisms in sarcoidosis. *Am J Respir Cell Mol Biol* 2002: **27**: 17–23.
7. Zorzetto M, Ferrarotti I, Trisolini R et al. Complement receptor 1 gene polymorphisms are associated with idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 2003: **168**: 330–4.
8. Hodgson U, Tukiainen P, Laitinen T. The polymorphism C5507G of complement receptor 1 does not explain idiopathic pulmonary fibrosis among the Finns. *Respir Med* 2005: **99**: 265–7.
9. Zorzetto M, Ferrarotti I, Campo I, Luisetti M. CR1 gene polymorphism in Finland. *Respir Med* 2006: **100**: 186–7.
10. Chanock SJ, Manolio T, Boehnke M et al. Replicating genotype-phenotype associations. *Nature* 2007: **447**: 655–60.
11. Mrazek F, Kvezereli M, Garr E et al. Complement receptor 1 single nucleotide polymorphisms in Czech and Dutch patients with sarcoidosis. *Tissue Antigens* 2008: **71**: 77–80.
12. Lalouel JM, Rohrwasser A. Power and replication in case-control studies. *Am J Hypertens* 2002: **15**: 201–5.
13. Samani NJ, Erdmann J, Hall AS et al. Genomewide association analysis of coronary artery disease. *N Engl J Med* 2007: **357**: 443–53.
14. Cordell HJ, Clayton DG. Genetic association studies. *Lancet* 2005: **366**: 1121–31.
15. Nogee LM, Dunbar AE 3rd, Wert SE, Askin F, Hamvas A, Whitsett JA. A mutation in the surfactant protein C gene associated with familial interstitial lung disease. *N Engl J Med* 2001: **344**: 573–9.
16. Lawson WE, Grant SW, Ambrosini V et al. Genetic mutations in surfactant protein C are a rare cause of sporadic cases of IPF. *Thorax* 2004: **59**: 977–80.
17. Xiang L, Rundles JR, Hamilton DR, Wilson JG. Quantitative alleles of CR1: coding sequence analysis and comparison of haplotypes in two ethnic groups. *J Immunol* 1999: **163**: 4939–45.

# Author Query Form

## Journal: Tissue Antigens

## Article: tan_1133

Dear Author,

During the copy-editing of your paper, the following queries arose. Please respond to these by marking up your proofs with the necessary changes/additions. Please write your answers on the query sheet if there is insufficient space on the page proofs. Please write clearly and follow the conventions shown on the attached corrections sheet. If returning the proof by fax do not write too close to the paper's edge. Please remember that illegible mark-ups may delay publication.
Many thanks for your assistance.

| Query No. | Query | Reply |
|---|---|---|
| 1 | Please provide department for all three affiliations. | |
| 2 | Please spell out FEV and FVC in the sentence 'In total, 475…'. | |
| 3 | Please spell out 'UIP' and 'ATS/ERS' in the sentence 'The diagnosis of IPF…'. | |
| 4 | Please provide author names for the first affiliation. | |

# MARKED PROOF

# Please correct and return this set

Please use the proof correction marks shown below for all alterations and corrections. If you wish to return your proof by fax you should ensure that all amendments are written clearly in dark ink and are made well within the page margins.

| *Instruction to printer* | *Textual mark* | *Marginal mark* |
|---|---|---|
| Leave unchanged | ··· under matter to remain | New matter followed by or or |
| Insert in text the matter indicated in the margin | | |
| Delete | through single character, rule or underline<br>or<br>through all characters to be deleted | or |
| Substitute character or substitute part of one or more word(s) | through letter or<br>through characters | new character or<br>new characters |
| Change to italics | under matter to be changed | |
| Change to capitals | under matter to be changed | |
| Change to small capitals | under matter to be changed | |
| Change to bold type | under matter to be changed | |
| Change to bold italic | under matter to be changed | |
| Change to lower case | Encircle matter to be changed | |
| Change italic to upright type | (As above) | |
| Change bold to non-bold type | (As above) | |
| Insert 'superior' character | through character or<br>where required | or<br>under character<br>e.g. or |
| Insert 'inferior' character | (As above) | over character<br>e.g. |
| Insert full stop | (As above) | |
| Insert comma | (As above) | |
| Insert single quotation marks | (As above) | or and/or<br>or |
| Insert double quotation marks | (As above) | or and/or<br>or |
| Insert hyphen | (As above) | |
| Start new paragraph | | |
| No new paragraph | | |
| Transpose | | |
| Close up | linking characters | |
| Insert or substitute space between characters or words | through character or<br>where required | |
| Reduce space between characters or words | between characters or words affected | |