Russ. J. Numer. Anal. Math. Modelling, Vol. 24, No. 4, pp. 377–384 (2009)
DOI 10.1515/ RJNAMM.2009.023

On the relative risk concept and TB morbidity in Russia: linking population genetics and epidemiological studies

S. G. RUDNEV*, R. P. SELITSKAYA†, and M. N. BOLDYREVA‡

Abstract — A scheme relating population genetics and TB epidemiology data in Russia is proposed. For this, a genetic susceptibility index is suggested based on relative risk estimates and published data on genetic polymorphisms of various ethnic groups. The index correlates well ($r = 0.746, p < 0.05$) with TB morbidity in genetically and geographically distant populations clearly indicating an important role of the genetic component in determining TB epidemiology in Russia.

1. Introduction

Tuberculosis in Russia and other countries is a significant social problem. According to the WHO data, nearly one third of the Earth's population is infected by tuberculosis mycobacteria and most of infected persons have a latent, inactive form of the infection, which annually turns into the active phase for almost 9 millions people. In 2007, the tuberculosis mortality was about 1.3 millions people and, moreover, 456 thousands death cases were registered among newly detected tuberculosis patients with HIV infection [6]. Russia is among 20 countries most suffering from the burden of tuberculosis.

Debates on the causes of tuberculosis run back to the antiquity. Hippocrates believed that the base is formed by heredity mechanisms. Aristotle and Galenus believed that tuberculosis is transmitted due to a contact with an infected person [23]. Later, with the beginning of industrialization and the growth of cities, tuberculosis was considered to be a disease of low-income groups related to a low quality of life. Nowadays, there is no doubt that the development of this disease is determined by a combination of the properties of a human protection system, mycobacteria, and environmental factors.

The principal DNA region determining the function of human immune system is

*Institute of Numerical Mathematics, Russian Academy of Sciences, Moscow 119333, Russia
e-mail:rudnev@inm.ras.ru

†Research Institute of Phtisiopulmonology, I. M. Sechenov Moscow Medical Academy, Moscow 127994, Russia

‡Institute of Immunology, Russian Federal Agency of Medicine and Biology, Moscow 115478, Russia

The work was supported by the Russian Foundation for Basic Research (07–01–00577).

the major histocompatibility complex, also called human leukocyte antigens (HLA). HLA contains more than 4 thousands base pairs and is located in the short arm of the 6th chromosome. Two main areas are identified in the HLA region, namely HLA-I and HLA-II [27].

At the population level, HLA genes possess a pronounced polymorphism providing antigenic individuality of organisms. HLA contains nearly 20 allele genes. The allele diversity of the major histocompatibility complex is being refined. More than 2700 different alleles of HLA genes have been identified by 2009, i.e., about 135 alleles per allele gene in average [9]. It is believed that the probability of the HLA-identity for two arbitrarily chosen persons is not greater that 10^{-6} [10].

The products of the major histocompatibility complex having antigen specificity (i.e., HLA-antigens) form an HLA-phenotype. The early studies of the HLA region were mainly based on examination of HLA-I polymorphisms using serologic (i.e., cellular mediated) typing. The discovery of the polymerase chain reaction in the middle of 1980-ies permitted one to proceed from characterization of the HLA structure based on expressed gene products to direct analysis of the locus texture by methods of DNA-typing.

Some associations of infectious and other diseases with the presence of particular HLA-antigens were revealed [3], which are due to different participation of HLA-antigens in the immune response. Those associations are not deterministic in nature and have a statistical character. According to one of hypotheses, the formation of the variety of HLA-phenotypes in the evolution process was a mechanism of increasing the species survival under infection pressure [11].

The presence or absence of hereditary susceptibility to diseases can be estimated from the level of the relative risk showing the ratio of the number of disease cases with a given genotypic specificity, i.e, particular allele to the number of cases without that specificity [26]. The relative risk of diseases for the carriers of some alleles of HLA genes may be 1.7 – 90 times higher [15]. The frequency distributions of alleles and the corresponding values of the relative risks may vary on the inter-ethnic and intra-ethnic levels and also depending on the geographic localization of the studied group [10].

Russia differs from other countries in nonuniformity of its ethnic structure and a largely extended territory, taking space of 9000 kilometers from west to east and 5000 kms from north to south. At present, a gradual accumulation of data for HLA-typing of various populations takes place in Russia. In future this can provide a detailed description of ethnoterritorial peculiarities of the genetic polymorphism of the HLA system in norm and disease.

The morbidity of pulmonary tuberculosis differs by 2–6 times in various regions of Russia [19]. The general tendency to higher morbidity is observed from west to east. At the same time, the morbidity of tuberculosis in the South Federal District is greater in average than in the Central and North–West Districts. The regional morbidity varies irregularly and may essentially differ for adjacent regions and territories [19]. One may suppose that such distinctions are caused not only by climatic, ecological, social factors, as well as different detection and control effi-

Table 1.
Some HLA-II markers of susceptibility and resistance to tuberculosis.

Alleles and allele groups		Sample size		Country	Reference
Susceptible	Resistant	Ill	Healthy		
DRB1*15	DRB1*11	74	90	China	[25]
DRB1*1501	—	72	36	India	[20]
DRB1*1501; DQB1*0601	DPB1*04	126	87	India	[17]
DRB1*07; DQA1*0101	DQA1*0301; DQA1*0501	40	100	Iran	[1]
DRB1*16	DRB1*13	31	58	Poland	[5]
DRB1*13(6); DRB1*14(6)	—	60	96	Tuva, Russia	[16]
DRB1*04	DRB1*11	147	209	Syria	[8]

ciency, but also by the influence of hereditary susceptibility. A question arises: to what extent do hereditary factors determine the differences in morbidity and other epidemic parameters of tuberculosis infection?

The aim of our study is the development of a method for estimation of the effect of the HLA region genetic polymorphism on tuberculosis morbidity in Russia. The genetic tuberculosis susceptibility index is proposed and a comparison of this index with tuberculosis morbidity data is performed for some regions of Russia.

2. HLA and tuberculosis

Many studies are focused on HLA genes polymorphism under pulmonary tuberculosis. The most interesting is the region HLA-D encoding class II antigens in particular, the gene HLA-DRB1. Probably, HLA-D antigens to a greater extent than other ones determine the trend and intensity of the adaptive immunity in tuberculosis infection and participate in specific recognition of the immunodominant epitope and its presentation to T-lymphocytes. In India and China (these countries are the most affected by tuberculosis) the marker of susceptibility to tuberculosis was the allele group DRB1*15 [17, 20, 25] and the protective ones were allele groups DPB1*04 [17] and DRB1*11 [25]. In Syria, Iran, and Poland, the markers of susceptibility were DRB1*04 [8], DRB1*07 [1], and DRB1*16 [5], respectively. In the examination of representatives of the Tuvinian ethnicity (Russia), susceptibility markers DRB1*13 and DRB1*14 [16] were determined. Interestingly, in the study of the Polish authors, the allele group DRB1*13 was associated with the resistance to tuberculosis [5]. These data are summed up in Table 1. As a whole, the allele groups corresponding to serologic specificities DR2 and DR6 are more often associated with susceptibility to tuberculosis, and DR5 is associated with resistance.

In our recent paper [18], HLA DRB1-markers of hereditary susceptibility to pulmonary tuberculosis are described using computer program IrGene 1.0 [22] based on data for low-resolution DNA-typing. The examination of peripheral blood samples obtained from 300 healthy blood donors from Moscow [4] and from 51 patients with pulmonary tuberculosis treated in the Phtisiopulmonology clinic of

I. M. Sechenov Moscow Medical Academy hospital was performed in the Immunogenetics department of the Immunology Institute of the Federal Medical & Biology Agency (FMBA), Russia. All donors and patients belong to the Russian ethnos. The DNA was extracted by the standard salting-out procedure [13]. The HLA typing of DNA samples was performed using multiprimer PCR [24]. For typing the gene DRB1 of the major histocompatibility complex of class II, the collections of primers HLA-DNA-Tech were used (SPF 'DNA-Technology', Russia).

It was found [18] that increased susceptibility to tuberculosis is significantly associated with four DRB1-genotypes (01/12, 04/15, 04/16, and 11/17). The analysis of the 2×2 contingency table with those genotypes combined into a single group revealed a strong connection between genotypes and TB susceptibility according to the Cheddock scale ($k_a = 0.82$). The values of the relative risk are calculated for allele groups corresponding to serologic specificities. The following allele groups are associated with susceptibility: DRB1*01, DRB1*08, and DRB1*17 (the differences are unreliable). There is a moderate ($k_a = -0.35$), but not reliable negative association of susceptibility to DRB1 homozygosity. We conclude that the allele gene DRB1 relating to an HLA locus of class II actively participates in the pathogenesis of tuberculosis.

3. Genetic susceptibility index and TB morbidity

We construct the genetic index of susceptibility to tuberculosis based on data which characterize the frequency of occurrence and the relative risk of tuberculosis development for various genotypic specificities in the ethnoses populating the regions of Russia and compare it with morbidity data.

Let M^i be the TB morbidity in the ith region of Russia. The value M^i can be represented as the weighted sum of partial morbidities (M^i_j):

$$M^i = \sum_{j=1}^{s_i} \alpha^i_j M^i_j \tag{3.1}$$

where j is the ordinal number of the ethnos, α^i_j is the frequency of occurrence of the jth ethnos in the ith region, s_i is the number of ethnoses populating the given region.

The partial morbidity is determined as

$$M^i_j = \sum_{k=1}^{n} \omega^i_{jk} M^i_{jk} \tag{3.2}$$

where M^i_{jk} is the morbidity for the region i, the ethnos j, and the genotypic specificity k, n is the total number of the specificities (for example, the alleles, genotypes, or haplotypes), and ω^i_{jk} are the corresponding frequencies of the genotypic speci-

ficities, so for any i and j we have the relation

$$\sum_{k=1}^{n} \omega_{jk}^{i} = 1. \tag{3.3}$$

The value M_{jk}^{i} is proportional to the *absolute* risk value for the specificity k of the corresponding contingency table for the samples of healthy people and TB patients representing the jth ethnos in the ith region:

$$M_{jk}^{i} = m_{j}^{i} A_{jk}^{i}. \tag{3.4}$$

The factor m_{j}^{i} does not depend on k and is inversely proportional to the value indicating the ratio of the quantities of the considered samples for TB patients and healthy individuals in the jth ethnos in the ith region and for TB patients and healthy people in the given group.

Unfortunately, formula (3.4) cannot be used in practical calculations because the parameters m_{j}^{i} are unknown. As a correlate for the value M_{jk}^{i}, we consider the ratio of the *relative* risk and the frequency of DRB1 gene alleles for representatives of the Russian ethnicity residing in Moscow:

$$S_{jk}^{i} = \frac{\tilde{R}_k}{\tilde{\omega}_k}. \tag{3.5}$$

Moving in the reverse direction and substituting S_{jk}^{i} instead of M_{jk}^{i} in (3.2), and then substituting the obtained expression into (3.1), we get

$$S^{i} = \sum_{j=1}^{s_i} \alpha_{j}^{i} S_{j}^{i} \tag{3.6}$$

where, by the analogy with M_{j}^{i}, it is convenient to call the index S_{j}^{i} the partial susceptibility index; and the parameters α_{j}^{i} characterize the ethnic structure of the ith region. The value S^{i} is called the genetic index of susceptibility to tuberculosis.

Note that in the following, for convenience of comparison of genetic susceptibility indices constructed for various sets of genotypic specificities, it makes sense to consider also the value S^{i} normalized with respect to the total number of specificities n (do not confuse with the sample size!):

$$\hat{S}^{i} = \frac{S^{i}}{n}. \tag{3.7}$$

For example, in case of low-resolution DNA-typing for the gene DRB1 mentioned above, the total number of specificities (allele groups) is 13 (the 14th allele group, DRB1*18, is very rare and usually does not appear). For details of HLA genes allele diversity, see [9].

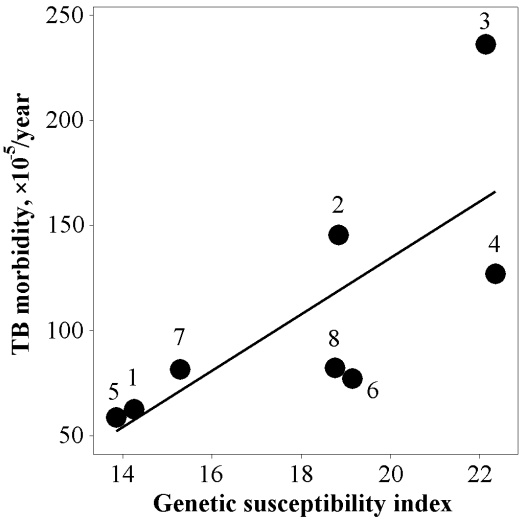

Figure 1. Scatter plot for tuberculosis morbidity in 2007 [19] and genetic susceptibility index for various regions of Russia: 1, Moscow; 2, Buryatia; 3, Tuva; 4, Kalmykia; 5, Tatarstan; 6, Chuvashia; 7, Udmurtia; 8, Mari El.

Thus, in order to simplify further analysis and due to the lack of data, we suppose that the variation of partial morbidity for different ethnicities is related *only* to the differences in the distributions of allele frequencies.

In our work we used data from the department of Human Immunogenetics of the Immunology Institute (FMBA, Russia) on the distribution of the gene DRB1 alleles (the parameters ω^i_{jk}) among Russian residents of Moscow ($n = 300$) as well as among the representatives of title nationalities of seven ethnic republics of Russia: Buryats ($n = 87$), Tuvinians ($n = 164$), Kalmyks ($n = 136$), Tatars ($n = 87$), Chuvashs ($n = 78$), Udmurts ($n = 101$), and Mari ($n = 202$) [3]. According to the results of multi-dimensional scaling based on the allelic frequencies of the gene DRB1, these distributions form clusters sufficiently distant from each other and characterize populations with different genetic bases, of the European, Asian, or mixed origin [3]. The data on the demographic structure of Russia's ethnic republics (the parameters α^i_j) were taken from [21].

Figure 1 shows the scatter plot for tuberculosis morbidity in 2007 and the genetic susceptibility index estimates for the considered regions of Russia. The value of correlation coefficient ($r = 0.746$, $p < 0.05$) indicates a significant contribution of the genetic component into morbidity.

4. Conclusion

The results of our study indicate an essential role of hereditary susceptibility as a factor determining differences in pulmonary tuberculosis morbidity in Russia. The existing mathematical models of tuberculosis spread and control in

Russia [2, 12, 14] combine the descriptions of the processes on the demographic and epidemic levels and the parameters of efficiency of the tuberculosis treatment system. These models are adapted for the description of the infection dynamics in genetically homogeneous regions of the Central Federal District of Russia and does not explicitly take into account the contribution of genetic inhomogeneity of the populations to variations in morbidity and other epidemiological parameters.

A high level of correlation of the proposed genetic susceptibility index with tuberculosis morbidity for genetically and geographically distant populations indicates the possibility of the development of models of tuberculosis spread and control in Russia at the national level taking into account the genetic diversity of the population.

Further studies are necessary for implementation of the considered approach. Analysis of statistical properties of the index described above, of basic assumptions and other factors influencing the estimate error is required. Estimation of the corresponding confidence intervals is also necessary. In particular, it is important to establish whether these estimates can be improved based on the relative risk data for the corresponding ethnicities and regions and what is the role of inter-ethnic differences in the frequency distributions of genotypic specificities for different regions of Russia.

The above approach can be used to study correlations between heredity and epidemiology of any other disease of infectious or non-infectious etiology.

References

1. A. A. Amirzargar, A. Yalda, M. Hajabolbaghi, et al., The association of HLA-DRB1, DQA1, DQB1 alleles and haplotype frequency in Iranian patients with pulmonary tuberculosis. *Int. J. Tuberc. Lung Dis.* (2004) **8**, No. 8, 1017–1021.
2. A. A. Avilov and A. A. Romanyukha, Mathematical modelling of tuberculosis propagation and patient detection. *Automation and Remote Control* (2007) **68**, No. 9, 1604–1617.
3. M. N. Boldyreva, *HLA (class II) and Natural Selection. 'Functional' Genotype, Hypothesis of Priority of the 'Functional' Heterozygosis*. Doctoral Diss. in Medical Sciences, Moscow, 2007 (in Russian).
4. M. N. Boldyreva, L. P. Alekseev, R. M. Khaitov, et al., HLA-genetic variety of population of Russia and CIS. I. Russians. *Immunology* (2005), No. 5, 260–267.
5. A. Dubaniewicz, B. Lewko, G. Moszkowska, et al., Molecular subtypes of the HLA-DR antigens in pulmonary tuberculosis. *Int. J. Infect. Dis.* (2000) **4**, No. 3, 129–133.
6. *Global tuberculosis control 2009: epidemiology, strategy, financing: WHO report 2009*. World Health Organization, Geneva, 2009.
7. E. V. Gubler and A. A. Genkin, *Application of Nonparametric Statistical Criteria in Medical-Biological Studies.* Meditsina, Leningrad, 1973 (in Russian).
8. E. I. Harfouch-Hammoud and N. A. Daher, Susceptibility to and severity of tuberculosis is genetically controlled by human leukocyte antigens. *Saudi Med. J.* (2008) **29**, No. 11, 1625–1629.
9. R. Holdsworth, C. K. Hurley, S. G. E. Marsh, et al., The HLA dictionary 2008: a summary of HLA-A, -B, -C, -DRB1/3/4/5, and -DQB1 alleles and their association with serologically defined HLA-A, -B, -C, -DR, and -DQ antigens. *Tissue Antigens* (2009) **73**, No. 2, 95–170.

10. R. M. Khaitov and L. P. Alekseev, Predestination of the immune system: realization of physiological functions providing the genetic identity of an organism. *Physiology and Pathology of Immune System* (2004), No. 8, 3–14.

11. N. A. Mayanskii and A. N. Mayanskii, Nomenclature and functions of the human major histocompatibility complex. *Immunology* (2006), No. 1, 43–46.

12. O. A. Melnichenko and A. A. Romanyukha, A model of tuberculosis epidemiology: estimation of parameters and analysis of factors influencing the dynamics of an epidemic process. *Russ. J. Numer. Anal. Math. Modelling* (2008) **23**, No. 1, 63–75.

13. S. A. Miller, D. Dykes, and H. F. Polesky, A simple salting-out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res.* (1988) **16**, No. 3, 1215.

14. M. I. Perelman, G. I. Marchuk, S. E. Borisov, et al., Tuberculosis epidemiology in Russia: the mathematical model and data analysis. *Russ. J. Numer. Anal. Math. Modelling* (2004) **19**, No. 4, 305–314.

15. R. V. Petrov, *Immunology*. Meditsina, Moscow, 1987 (in Russian).

16. L. E. Pospelov, A. G. Matrakshin, A. F. Malenko, et al., Genetic markers of the HLA region associated with tuberculosis morbidity in Barum-Khemchick region of Tuva republic. *Problems of Tuberculosis and Pulmonary Diseases* (2007), No. 6, 62–64.

17. M. Ravikumar, V. Dheenadhayalan, K. Rajaram, et al., Associations of HLA-DRB1, DQB1 and DPB1 alleles with pulmonary tuberculosis in south India. *Tuber. Lung. Dis.* (1999) **79**, No. 5, 309–317.

18. R. P. Selitskaya, M. N. Boldyreva, S. G. Rudnev, et al., On the polymorphism of the DRB1-locus of HLA region and susceptibility to tuberculosis. *Immunology* (2009) (to appear).

19. M. V. Shilova, *Tuberculosis in Russia in 2007.* Moscow, 2008 (in Russian).

20. U. Sriram, P. Selvaraj, S. M. Kurian, et al., HLA-DR2 subtypes & immune responses in pulmonary tuberculosis. *Indian J. Med. Res.* (2001) **113**, 117–124.

21. *Regions of Russia. Basic characteristics of Subdivisions of Russian Federation, 2007*. Statistical yearbook. Rosstat, Moscow, 2008.

22. E. A. Sytin, IrGene 1.0, interface and code for analysis of population genetics data in immunology. *Collected Articles of Young Scientists of CMC Department*. M. V. Lomonosov Moscow State University, 2009, Issue 6 (in print).

23. H. Takiff, Host genetics and susceptibility. In: *Tuberculosis 2007: From Basic Science to Patient Care* (Eds. J. C. Palomino, S. C. Leao, and V. Ritacco), 2007, pp. 207–262.

24. D. Yu. Trofimov, Development of the multiprimer PCR method for typing HLA genes of class II. *Ph.D. Thesis in Biology*, Moscow, 1996 (in Russian).

25. J. Wang, C. Song, and S. Wang, Association of HLA-DRB1 genes with pulmonary tuberculosis. *Zhonghua Jie He He Hu Xi Za Zhi*. (2001) **24**, No. 5, 302–305.

26. B. Woolf, On estimating the relation between blood group and disease. *Ann. Hum. Genet.* (1955) **19**, 251–253.

27. A. A. Yarilin, *Foundations of Immunology*. Meditsina, Moscow, 1999 (in Russian).