EVALUATION OF GENETIC RISK FACTORS FOR CHILDHOOD ASTHMA USING GENERALIZED MULTIFACTOR DIMENSIONALITY REDUCTION SOFTWARE

Olga Cirstea[1], Liubov Vasilos[1,2], Ala Cojocaru[2], Dorina Savoschin[2]

[1] Department of Pediatrics, State Medical and Pharmaceutical University "Nicolae Testemitanu", Chisinau, Republic of Moldova

[2] Scientific Department of Pediatrics, Institute for Maternal and Child Healthcare, Chisinau, Republic of Moldova

ABSTRACT

Asthma is a multifactorial disease caused by numerous genetic interrelations and also gene-environment interactions. Traditional statistical methods applied for case-control studies in do not offer the possibility for in-depth analysis of the genetic risk because of the "dimensionality" problem. The Generalized Multifactor Dimensionality Reduction software was successfully applied in genetic studies of cancer, hypertension, diabetes, etc.

The aim of our study was to use the GMDR software to analyze the role of genetic interrelations on asthma development in Moldovan children.

Materials and methods. The case-control study included 180 children: 90 children with asthma and 90 healthy controls comparable by age and gender. The GMDR software Beta 0,9 was applied to adjust for discrete quantitative data and covariates. Association of genetic polymorphisms (GSTT1 and GSTM1 gene deletions; GSTP1 313 A > G; NAT2 481 C > T, 590 G > A and 857 G > A; IL-4 -590 C > T, IL-4Rα Arg551Gln; TNFα -308 G > A; (AAT)n repetitions in intron 20 of the NOS1 gene; and CC16 38 G > A) with the risk of developing asthma was studied.

Results. Data analysis using the GMDR method resulted with identification of a combination of four genotypes (GSTT1+, NAT2 *5-*7/*5-*7, NOS1 <12/>12, and IL-4 -590 C/C) that increases by 3.6 folds the risk of childhood asthma development (OR = 3.61; CI 95% 1,45-8,99; p < 0,01).

Conclusions. Gene-gene interactions interfere at different levels of pathophysiological mechanisms of childhood asthma development. Analysis of the interaction models generated by the GMDR software based on study data showed a significant association between asthma risk and polymorphisms in two important groups of genes – xenobiotic-metabolizing genes and the gene responsible for nitric oxide synthesis in airway mucosa.

Keywords: asthma, child, risk factor, genetic polymorphism, GMDR software

INTRODUCTION

Bronchial asthma (AB) is characterized by paroxysmal or persistent shortness of breath, chest tightness, wheezing, sputum production and cough, airflow limitation and a variable degree of bronchial hyperreactivity, and is triggered by endogenous or exogenous stimuli. Genetic predisposition, exposure to allergens and air pollutants, low socioeconomic status, stressful environment and limited access to specialized healthcare services contribute to the high asthma burden worldwide. In Moldova the prevalence of asthma also tends to increase, being diagnosed late or under-diagnosed.

The inflammation that develops at the level of airways mucosa is an important feature of the disease and determines evolution and persistent character of asthma. Still, many underlying mechanisms are not fully understood. A comprehensive approach is needed to reveal various risk factors in-

Corresponding author:
Olga Cirstea, MD, PhD, Clinical Lecturer, Department of Pediatrics, State Medical and Pharmaceutical University "Nicolae Testemitanu", 93 Burebista St., MD-2060, Chisinau, Republic of Moldova
E-mail olga.cirstea@usmf.md

volved in the onset of this disease in childhood. Studies focused on the interaction between genetic and environmental risk factors in asthma are promising in identifying the optimal treatment method and effective preventive measures.

Genetically predisposed individuals being exposed to aggressive and/or triggering factors from environment, such as allergens, viruses, xenobiotics, etc. develop persistent inflammation of the airways and asthma symptoms later on. A number of studies indicate that the disease is largely determined by a polygenic model of inheritance (1). The list of genes involved in the pathogenesis of asthma includes more than 100 loci. Large genome-wide association studies (GWAS) studies identified a number of chromosomal regions related to the pathogenetic mechanisms of asthma (5q23-31, 6p, 11q, 12q, 17q21, etc.) (2), and this list is not ended yet.

Generally, the "genetic background" is the genotype that includes all related genes involved in the pathogenetic mechanisms which can strongly influence the phenotype or clinical symptoms of the disease. However, genetic studies provide evidence of functional interrelations between genes with regulatory effects that can modify the function of other genes, the so-called phenomenon of ***epistasis***. This is a form of interaction between non-allelic genes, in which one gene masks the phenotypic expression of another gene located on a different locus and on another chromosome. Therefore, the ***epistatic gene* (*suppressor*)** inhibits the expression of another gene, which is dominant but non-allelic with the modifier gene; instead the ***hypostatic gene*** is the allele whose functional expression is suppressed (3). Thus, the complexity of atopic phenotype development is well understood, especially when epistasis interaction involves three or more loci. This phenomenon determines the heterogeneity and variability of clinical manifestations found in different populations with asthma and atopy.

Usually, statistical analysis used to assess the risk of association of genetic risk factors with phenotypic manifestation in multifactorial diseases limits to traditional comparative evaluation of alleles frequency in the group of patients and healthy controls by *chi-square* test and odds ratio (OR), confidence intervals and *p* values. However, we must admit that the most plausible model of inheritance of atopic diseases is the polygenic one, in which the overall effect is determined by the summary effects of individual genetic variants. In asthma studies a relatively modest number of associations between genetic markers and disease manifestations is partly due to the lack of verification for synergistic effects of the studied polymorphisms (4). Hence, estimating the contribution of individual genetic factors to the disease phenotype would represent a complex problem in the view of traditional statistics using parametric methods (5,6).

In order to address this issue multifactor dimensionality reduction method was elaborated, which is an innovative statistical method that offers greater possibilities to adjust qualitative and quantitative parameters of the studied phenotypes and larger flexibility for the proposed design of the investigation (5). This method was successfully used in genetic studies of patients with cancer (7,8), atrial fibrillation (9), autism (10), blood hypertension (11), type II diabetes mellitus (12), and has shown a significant increase of the predicted risk accuracy including highly complex models of the disease.

The purpose of our study was to analyze the role of genetic factors and their interaction in the development of childhood asthma by using computer application GMDR (*Generalized Multifactor Dimensionality Reduction*, software Beta 0.9).

MATERIAL AND METHODS

Data for the complex mathematical analysis was resulted from a case-control study which included a group of 180 children. Ninety children diagnosed with asthma entered the basic group of study (mean age 10.9 ± 0.4 years), all of them hospitalized to the Allergy Unit of the Institute for Maternal and Child Healthcare. The control group included 90 healthy children (mean age 13.5 ± 0.2 years), with no symptoms or complains of allergic diseases. Subjects from both groups were comparable by age and sex. They were selected randomly based on lists and parental informed consent.

Genomic DNA for molecular genetics was extracted from peripheral blood lymphocytes, using special cards intended for transportation of blood samples to the laboratory[1]. Polymerase chain reaction (PCR) using *thermostable* DNA *polymerase* (Thermo Scientific *Taq* DNA *Polymerase,* „Promega", USA) and restriction fragment length polymorphism method (RFLP) were applied to study several *single nucleotide* polymorphisms (SNPs) of the asthma candidate genes: deletion polymorphism of GSTM1 and GSTT1 genes; polymorphisms *313 A* > ***G*** and *341 C* > ***T*** of the GSTP1 gene; several NAT2 SNPs (*481 C* > ***T***,

[1]Molecular genetic investigations were performed in collaboration with the Laboratory of Prenatal Diagnosis of Congenital and Inherited Diseases of the D.O. Ott Research Institute of Obstetrics and Gynecology (Sankt Petersburg, Russia).

*590 G > **A*** and *857 G > **A***); (AAT)n repetitions in intron 20 of the NOS1 gene; SNP *-590 C > **T*** of the IL-4 gene; IL-4Rα Arg551Gln polymorphism; SNP *-308 G > **A*** of the TNF gene; polymorphism *38 G > **A*** of the gene CC16.

GMDR method[2] is intended for the analysis of associations between dichotomous characters and combinations of predictors that can be genetic or exogenous by origin. The software is available to interested researchers free of charge and can be downloaded from the official website www.ssg.uab.edu/gmdr/ (13). The analysis may include an arbitrary number of factors. In our study we aimed to identify the combination of genetic markers with predictive value for the risk of developing asthma in childhood. To achieve this result we introduced the data derived from a case-control study on asthmatics and healthy controls.

The methodology of GMDR software application. The first line of the file is the header line, which includes names of the studied polymorphisms and affiliation of the study subjects to basic or control groups. The group coded with zero ("0") represents the control group and the group coded withy number one ("1") represents asthma patients. Because each marker is represented by a pair of alleles, the file should look like this:

Polymorphism 1	Polymorphism 2	Polymorphism 3	Group
1	2	1	0
0	2	1	0
...			
2	2	1	1

The files with coded results according to this model were processed by the GMDR application, whereafter statistical analysis was performed by dividing the data into two equal subsets, one of which being used as a test set. After checking all possible combinations of genetic variants, the software generated the combination (model) with the highest predictive value, which was selected based on statistical estimation of the reproducibility of all tested models (cross-validation consistency, CVC) and balanced accuracy (Bal. Acc.) that correlated with sensitivity and specificity of the method (5). The "Run Analysis" option starts the multifactor dimensionality reduction analysis in order to identify combinations of genes that increase the risk of the disease. In our study we applied the method of exhaustive search to find the combinations of genotypes from all studied polymorphisms.

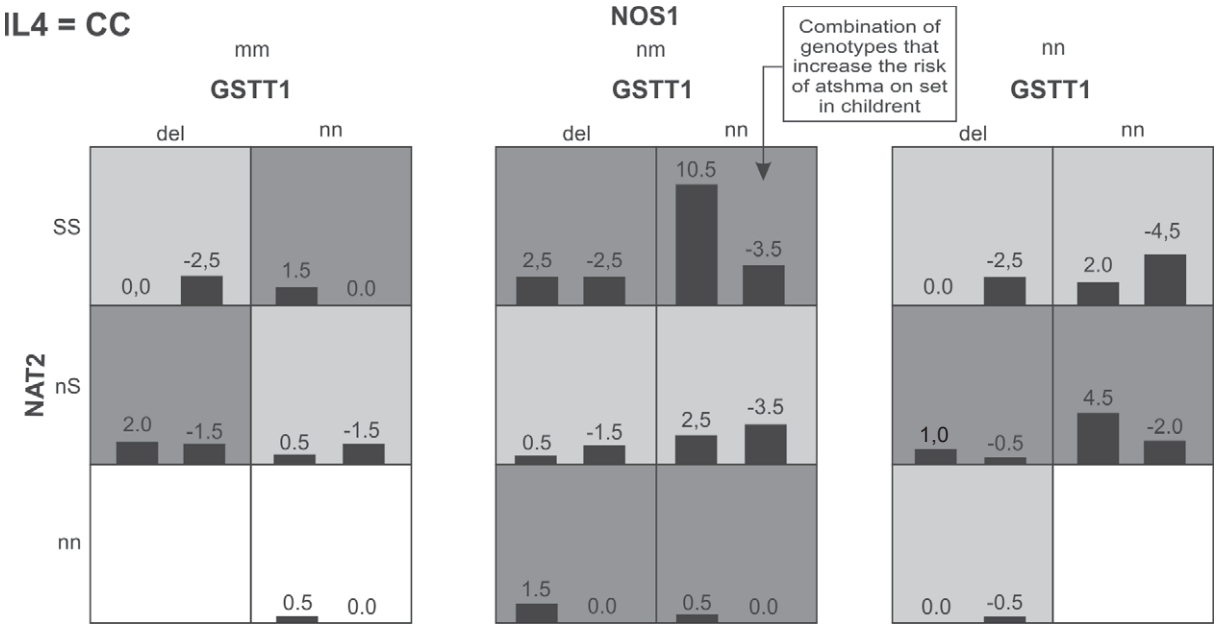

FIGURE 1. *Associations of SPNs genotypes of the GSTT1, NAT2, NOS1 genes with homozygous wild-type genotype of the IL-4-590 C > T polymorphism*

[2] The GMDR software was applyied to analyze the data results within the institutional project „Study of genetic risk factors for development and spread of allergic disorders in children in correlation with the environment" (2009-2010) conducted by the Scientific Department of Pediatrics from the Institute for Maternal and Child Healthcare (Chisinau, Republic of Moldova).

RESULTS AND DISCUSSION

The study of gene-gene interactions in asthma included polymorphisms of the *xenobiotic-metabolizing genes*, genes coding immune mediators of the inflammatory process and genes regulating the inflammatory response in the airways mucosa. Mathematical analysis using the GMDR software identified 34 combinations of the studied genotypes. Some of them are associated with increased risk of developing asthma (marked in gray intense), others – associated with a lower risk (light gray color) (Fig. 1 and 2).

Notably, the model of genetic interaction which reached statistical significance ($p < 0.05$) includes only four SNPs of asthma candidate genes (Fig. 1). Balanced accuracy of the model was 0.67, the sensitivity (Se) – 0.68, specificity (Sp) – 0.65, reproducibility of the results (CV Consistency) – 10/10. This combination of genotypes increased the risk for developing asthma in children 3.6 fold (OR = 3.61, CI

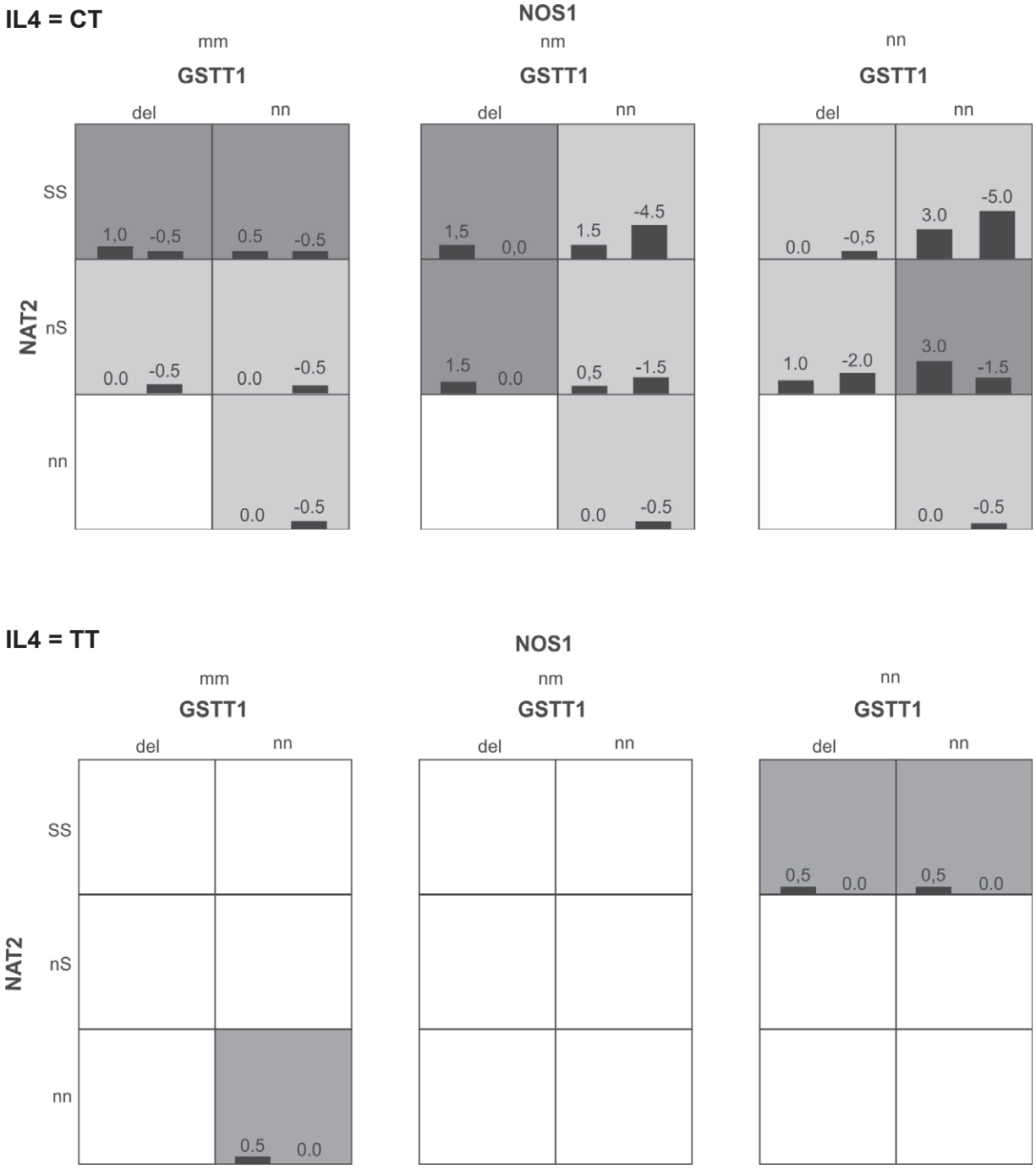

FIGURE 2. *Associations of SPNs genotypes of the GSTT1, NAT2, NOS1 genes with heterozygous and homozygous genotypes of the IL-4 -590 C > T mutation*

1.45 to 8.99, $p < 0.01$) and includes the following genotypes:

- **GSTT1** ($GSTT1^+$, consists of two *functional* alleles);
- **NAT2** (NAT2 *5-*7/*5-*7, homozygous genotypes with both alleles being functionally impaired, so-called "slow acetilators");
- **NOS1** (NOS1 <12/>12, heterozygous genotype) and
- **IL-4** (IL-4 *-590* C/C, homozygous genotype with wild-type alleles).

It must be mentioned that only two of these SNPs genotypes – polymorphisms of the GSTT1 and IL-4 genes, are functionally competent comparing to the NAT2 and NOS1 mutations, which variants include alleles with decreased functionality.

These results are important with regard to the role of these genes in asthma mechanisms. Thus, liver enzyme arylamine N acetyltransferase 2 is involved in the detoxification of drugs and xenobiotics such as arylamines, including diisocyanates and acids used to produce paints, varnishes, adhesives, laminates, etc. Some studies showed that acetylation process can influence the mechanisms of inactivation of the excess of organic amines including histamine, which is responsible for the symptoms of allergic reactions (14).

Presence of "fast" and "slow" acetylators in the population reflects the genetically determined variability of the NAT2 gene. About 50% of Caucasian population show decreased activity of this enzyme (15). However, the role of this gene is much more complex, being also related to impaired immune mechanisms. Batra et al. (16) studied the association of NAT2 SNPs with allergic asthma, serum IgE levels and eosinophilia. The authors have shown that SNP variants of the NAT2 gene are involved in the regulation of the serum IgE levels and eosinophils count, and may therefore influence the asthma evolution. At the same time, N acetyltransferase 2 converts serotonin into melatonin – mediators known for their proinflammatory activity and with important role in asthma pathogenesis (17). Some studies showed that patients with asthma exhibit increased serotonin levels that correlate with the severity of the disease. Also, it decreases the release of Th1 type cytokines and increases the amount of acetylcholine which acts as a bronchoconstrictor agent (18). Melatonin, in turn, induces the synthesis of proinflammatory cytokines (IL-1, IL-2, IL-6, IL-12 and TNFα) and the proliferation of T lymphocytes sensitized to specific antigens together with Th2-type immune response, that finally end with increased IL-4 production (19). The NOS1 gene which was identified in the combination of SPNs in our study plays a separate role in the pathogenesis of asthma. This gene regulates the expression of the NOS2 gene which codes nitric oxide – an important mediator that influences the Th1/Th2 immune balance and triggers the Th2-type cellular response (20).

CONCLUSIONS

1. The described method to analyze data in genetic studies enabled the identification of the combination of genetic polymorphisms that increases the risk for asthma onset in children and represents an important methodological support to understand molecular mechanisms of multifactorial diseases development.

2. Our study results showed that gene interactions interfere at different levels of the pathophysiological mechanisms of asthma development in children. Identification of genetic risk factors for asthma, represented by a combination of mutations in one of the *xenobiotic-metabolizing genes and the* gene regulator of the nitric oxide in the respiratory mucosa, brings evidence that carriers of these genotypes have an increased susceptibility to the negative impact of the environment.

REFERENCES

1. **Carroll W.** Asthma genetics: pitfalls and triumphs. *Paed. Respir. Reviews* 2005; 6: 68-74.
2. **Vercelli D.** Discovering susceptibility genes for asthma and allergy. *Nature Rev. Immunol.* 2008; 8: 169-182.
3. **Erickson D.** Quantitative trait loci: mapping the future of QTL's. *Heredity* 2005; 95: 417-418.
4. **Moore J.H.** The ubiquitous nature of epistasis in determining susceptibility to common human diseases. *Hum. Hered.* 2003; 56: 73-82.
5. **Hahn L.W., Ritchie M.D., Moore J.H.** Multifactor dimensionality reduction software for detecting gene-gene and gene-environment interactions. *Bionformatics* 2003; 19(3): 376-382.
6. **Mei H., Cuccaro M.L., Martin E.R.** Multifactor dimensionality reduction-phenomics: A novel method to capture genetic heterogeneity with use of phenotypic variables. *Am. J. Hum. Gen.* 2007; 81: 1251-1261.
7. **Andrew A.S., Nelson H.H., Kelsey K.T. et al.** Concordance of multiple analytical approaches demonstrates a complex relationship

between DNA repair gene SNPs, smoking and bladder cancer susceptibility. *Carcinogenesis* 2006; 27(5): 1030-1037.
8. **Ritchie M.D., Hahn L.W., Roodi N. et al.** Multifactor-dimensionality reduction reveals high-order interactions among estrogen-metabolism genes in sporadic breast cancer. *Am. J. Hum. Genet.* 2001; 69: 138-147.
9. **Tsai C.T., Lai L.-P., Lin J.-L. et al.** Renin-angiotensin system gene polymorphisms and atrial fibrillation. *Circulation* 2004; 109(13): 1640-1646.
10. **Ma D.Q., Whitehead P.L., Menold M.M. et al.** Identification of significant association and gene-gene interaction of GABA receptor subunit genes in autism. *Am. J. Hum. Genet.* 2005; 77(3): 377-388.
11. **Moore J.H, Williams S.M.** New strategies for identifying gene-gene interactions in hypertension. *Ann. Med.* 2002; 34(2): 88-95.
12. **Cho Y.M., Ritchie M.D., Moore J.H. et al.** Multifactor-dimensionality reduction shows a two-locus interaction associated with Type 2 diabetes mellitus. *Diabetologia* 2004; 47(3): 549-554.
13. **Lou X.Y., Chen G.B., Yan L. et al.** GMDR (Generalized Multifactor Dimensionality Reduction); software available at http://www.ssg.uab.edu/gmdr/(accesat la data de 23 iunie 2012).
14. **Seiler N.** Catabolism of polyamines. *Amino Acids* 2004; 26: 217-233.
15. **Sabbagh A., Langaney A., Darlu P. et al.** Worldwide distribution of NAT2 diversity: implications for NAT2 evolutionary history. *BMC Genet.* 2008; 9: 21.
16. **Batra J., Sharma S.K., Ghosh B.** Arylamine N-acetyltransferase gene polymorphisms: markers for atopic asthma, serum IgE and blood eosinophil counts. *Pharmacogenomics* 2006; 7: 673-682.
17. **Klein D.C.** Arylalkylamine N-acetyltransferase: "the Timenzyme". *J. Biol. Chem.* 2007; 282: 4233-4237.
18. **Ménard G., Turmel V., Bissonnette E.Y.** Serotonin modulates the cytokine network in the lung: involvement of prostaglandin E2. *Clin. Exp. Immunol.* 2007; 150: 340-348.
19. **Sutherland E.R., Ellison M.C., Kraft M., Martin R.J.** Elevated serum melatonin is associated with the nocturnal worsening of asthma. *J. Allergy Clin. Immunol.* 2003; 112: 513-517.
20. **Togashi H., Sasaki M., Frohman E. et al.** Neuronal (type I) nitric oxide synthase regulates nuclear factor kappaB activity and immunologic (type II) nitric oxide synthase expression. *Proc. Natl. Acad. Sci. USA* 1997; 94: 2676-2680.