ISSN 1022-7954, Russian Journal of Genetics, 2024, Vol. 60, No. 9, pp. 1233–1246. © Pleiades Publishing, Inc., 2024.

HUMAN GENETICS

# Association of Polymorphic Loci of Long Noncoding RNA Genes (*H19, MEG3, MALAT1, LINC00305, LINC00261, LINC02227,* and *CDKN2B-AS1*) with Chronic Obstructive Pulmonary Disease

**G. F. Korytina[a, b, *], L. Z. Akhmadishina[a, c], V. A. Markelov[a, b], T. R. Nasibullin[a], Y. G. Aznabaeva[b], O. V. Kochetova[a], N. N. Khusnutdinova[a], A. P. Larkina[a], N. Sh. Zagidullin[b], and T. V. Victorova[b]**

[a] *Institute of Biochemistry and Genetics—Subdivision of the Ufa Federal Research Center, Russian Academy of Sciences, Ufa, 450054 Russia*

[b] *Bashkir State Medical University, Ufa, 450008 Russia*

[c] *Ufa State Petroleum Technological University, Ufa, 450064 Russia*

**e-mail: guly_kory@mail.ru*

Received February 21, 2024; revised April 9, 2024; accepted April 17, 2024

**Abstract**—Chronic obstructive pulmonary disease (COPD) is a chronic lung disease resulting from dynamic, cumulative gene-environment interactions that cause lung tissue injury, alteration of its normal function, and acceleration of cellular senescence. Long noncoding RNAs (lncRNAs) function as critical epigenetic regulators of various aspects of cellular senescence. The objective of the present study is to identify the association between polymorphic variants of *H19* (rs3741219), *MEG3* (rs7158663), *MALAT1* (rs619586), *LINC00305* (rs2850711), *LINC00261* (rs6048205), *CDKN2B-AS1* (rs4977574), and *LINC02227* (rs2149954) lncRNAs genes with COPD. DNA samples from COPD patients ($N = 703$) and healthy individuals ($N = 655$) were studied in this study and polymorphic loci were analyzed by real-time PCR. Association with COPD was established with *H19* (rs3741219), *MEG3* (rs7158663), *LINC02227* (rs2149954), *MALAT1* (rs619586) and *CDKN2B-AS1* (rs4977574). Polygenic analysis has made it possible to identify informative gene–gene combinations that include polymorphic variants of the studied lncRNAs genes and genes encoding molecules of signaling cascades associated with cellular senescence and apoptosis. Multiple regression and ROC analysis revealed a COPD risk predictive model, which included gene–gene combinations of lncRNAs genes and smoking index ($P = 4.01 \times 10^{-48}$, AUC = 0.87).

**Keywords:** chronic obstructive pulmonary disease, long noncoding RNAs, cellular senescence, polygenic analysis, *H19*, *MEG3*, *LINC02227*, *CDKN2B-AS1*

**DOI:** 10.1134/S1022795424700777

## INTRODUCTION

Chronic obstructive pulmonary disease (COPD) is a severe chronic disease characterized by persistent and progressive restriction of air flow in the respiratory tract caused by developed emphysema and obstructive bronchitis and bronchiolitis [1]. According to the World Health Organization (WHO), to date, the prevalence of COPD exceeds 250 million people, thus representing the third leading cause of mortality worldwide. By 2060, the expected mortality rate caused by COPD will be more than 5.4 million deaths annually [https://www.who.int/ru/news-room/fact-sheets/detail/chronic-obstructive-pulmonary-disease-(copd)]. In the Russian Federation, a pronounced long-term trend in an increased incidence of this pathology is observed; namely, more than three million COPD cases were registered in 2021 [2]. COPD results from dynamic, cumulative, and repeated gene–environment interactions throughout life, which cause lung tissue damage and changes in its normal functioning [3]. Development of COPD is caused by different factors such as tobacco smoking, air pollution, professional exposure, and genetic and epigenetic factors [4]. Long noncoding RNAs (lncRNAs) represent transcripts 200 nucleotides in length, which do not code proteins but function as important regulators of various biological processes, including alternative splicing, RNA degradation, RNA suppression, transcription enhancing and silencing, chromatin remodeling, and post-translational modification of proteins structure [5]. Several studies demonstrated that lung tissue of COPD patients was characterized by differential expression of lncRNAs, the majority of which are involved in the regulation of various aspects of cellular senescence [6]. COPD development and progression may be caused by both changes in expression of lncRNAs and their impaired functioning owing to genetic polymorphism. The role of lncRNA polymor-

phism and interaction with genes encoding proteins which are responsible for the regulation of cellular senescence, oxidative stress, and COPD remain incompletely studied. The present study aimed to determine the association of polymorphic variants of the genes encoding regulatory lncRNAs including *H19*, *MEG3*, *MALAT1*, *LINC00305*, *LINC00261*, *CDKN2B-AS*, and *LINC02227* with developing COPD.

## MATERIALS AND METHODS

The design of the present study is a candidate case–control study. DNA samples from unrelated individuals, ethnic Tatars, who were residents of the Republic of Bashkortostan, were used in the study. All COPD patients were hospitalized in the Department of Pulmonology at the Ufa City Clinical Hospital no. 21. The diagnosis of COPD was established according to recommendations of the working group on the "Global Strategy for the Diagnosis, Treatment, and Prevention of Chronic Obstructive Pulmonary Disease" (http://goldcopd.org) on the basis of clinical and laboratory instrumental studies, including high-resolution computed tomography and spirometry. As part of a clinical and instrumental examination, all participants were examined for the parameters of external respiration (lungs vital capacity (VC), forced vital capacity (FVC), forced expiratory volume in one second ($FEV_1$), $FEV_1$/FVC ratio); the proportion of smoking patients was identified and their smoking index was calculated. The group of patients included 703 individuals (627 men (86.195%) and 76 women (10.81%)); the mean age was 63.04 ± 12.02 years. Smokers and former smokers included 590 individuals (83.93%), while 113 (16.08%) of COPD patients were nonsmokers. The smoking index in smokers was 43.08 ± 25.75 packs/years. In the patients group, the following values were revealed (% of the normal values): $FEV_1$ (41.99 ± 19.0), FVC (56.65 ± 22.71), VC (58.09 ± 21.59), $FEV_1$/FVC (62.29 ± 20.98).

The control group consisted of unrelated individuals, who reported no chronic diseases in anamnesis, including respiratory diseases and acute respiratory diseases at the time of collection of biological material. The inclusion criteria in the control group were the normal parameters of respiratory function ($FEV_1$/FVC > 70%, $FEV_1$ > 80%) and age over 45 years. The control group included 655 individuals (582 men (88.85%) and 73 women (11.15%)); mean age was 60.67 ± 11.31, of which 552 (84.27%) were smokers and former smokers and 103 (15.73%) were nonsmokers); the smoking index was 39.75 ± 25.87 packs/year.

*Genotyping.* DNA was isolated from the peripheral blood leukocytes using phenol-chloroform extraction. Genes and SNPs for the analysis were selected according to the following criteria: functional significance and/or previously reported association with other multifactorial diseases in humans, minor allele frequency (MAF) ≥5% in Europeans according to the National Center for Biotechnology Information database (http://www.ncbi.nlm.nih.gov/projects/SNP/). The following polymorphic loci in the lncRNA genes were selected for the study: *H19* (rs3741219), *MEG3* (rs7158663), *MALAT1* (rs619586), *LINC00305* (rs2850711), *LINC00261* (rs6048205), *CDKN2B-AS1* (rs4977574), and *LINC02227* (rs2149954). The functional significance of SNPs was assessed according to the RegulomeDB Version 1.1 (https://regulomedb.org), SNPinfo Web Server (https://snpinfo.niehs.nih.gov), and HaploReg v3 databases [7]; the data are presented in Table 1. SNPs were analyzed using real-time polymerase chain reaction (PCR) with commercially available assays with fluorescent detection (https://www.oligos.ru, DNA Synthesis LLC, Russia) using BioRad CFX96™ (Bio-Rad Laboratories, Inc., USA). The methods of analysis have been previously described by our group in detail [8].

*Statistical analysis of results.* A description of standard methods of statistical analysis has been previously reported [8]. The analysis of deviation of genotype frequencies from the Hardy–Weinberg equilibrium and association analysis of certain SNPs with a disease were performed using SNPassoc v 2.0–2 package for R [9]. A polymorphic marker was considered to be associated with a phenotype at $P < 0.05$; a correction for multiple comparisons was conducted using the assessment method for the proportion of received false-positive results, False Discovery Rate (FDR), via online tool (https://tools.carbocation.com/FDR). The search for gene–gene interactions associated with a disease was conducted using the Markov chain Monte Carlo method implemented in APSampler software (http://sourceforge.net/projects/apsampler/) [10]. To construct prediction models, a logistic regression with a stepwise forward inclusion of parameters, which comprised gene–gene interactions, SNPs of certain genes, and clinical-demographic parameters, was used. The area under the curve (AUC) was calculated to estimate the efficacy of prognostic models; the calculations were carried out in SPSS v. 22.

## RESULTS

The analysis of seven polymorphic loci of lncRNA genes, including *H19* (rs3741219), *MEG3* (rs7158663), *MALAT1* (rs619586), *LINC00305* (rs2850711), *LINC00261* (rs6048205), *CDKN2B-AS1* (rs4977574), and *LINC02227* (rs2149954), was carried out in the formed groups of COPD patients and control group. Bioinformatic analysis of functional characteristics selected for the study of lncRNA gene polymorphisms demonstrated that the majority of SNPs affected gene expression or were partially linked to the functional gene loci (Table 1).

**Table 1.** Bioinformatic analysis of functional characteristics selected for the study of polymorphic loci of long noncoding RNA genes

| Gene RefSNP HGVS Names | Chromosomal position | Regulatory | | Promoters of histone marks | Enhancers of histone marks | DNase | Regulatory proteins | Motifs | TF | Expression QTLs (Haplo Reg, GTEx portal) |
|---|---|---|---|---|---|---|---|---|---|---|
| | | rank | coefficient | | | | | | | |
| *H19* rs3741219 g.7447 T>C | 11p15.5 | 4 | 0.70497 | 8 tissues | 5 tissues | 11 tissues | – | NRSF, YY1 | | 19 tissues (including lung) |
| *MEG3* rs7158663 g.21819 A>G | 14q32.2 | 1f | 0.19549 | – | – | Skin | – | 5 | Yes | Blood, arteries |
| *MALAT1* rs619586 g.65266169 A>G | 11q13.1 | 1a | 0.99267 | 23 tissues | Spleen | 48 tissues | 4 species | 4 | Yes | 14 tissues |
| *LINC00305* rs2850711 g.61787038 A>T | 18q22.1 | 4 | 0.60906 | – | – | – | – | – | – | – |
| *LINC00261* rs6048205 g.22578963 A>G | 20p11.21 | 4 | 0.60906 | 11 tissues | 11 tissues | 10 tissues | 5 species | 14 | Yes | 3 tissues |
| *CDKN2B-AS1* rs4977574 g.22098574 A>G | 9p21.3 | 2c | 0.70567 | Adipose tissue | 11 tissues | 4 tissues | – | Ets, GR | | Blood, pituitary gland |
| *LINC02227* rs2149954 g.157820602 G>A | 5q33.3 | 1f | 0.55436 | – | 4 tissues | Liver, blood vessels | 4 species | COMP1 | – | – |

RefSNP according to NCBI database (https://www.ncbi.nlm.nih.gov/); functional significance of SNPs was examined on the basis of the RegulomeDB Version 1.1 (https://regulomedb.org), SNPinfo Web Server (https://snpinfo.niehs.nih.gov), HaploReg v3, and GTEx (https://www.gtexportal.org) databases. DNase—a region sensitive to DNase; motifs—modified regulatory motifs for binding with transcription factors; TF—sites of binding with transcription factors; regulatory proteins—regions for binding with regulatory proteins; Expression QTLs—expression Quantitative Trait Locus.

**Table 2.** Allele and genotype frequencies of the examined polymorphic loci of lncRNA genes in the groups of COPD patients and healthy individuals

| Gene RefSNP[1] | Minor allele | Genotypes, alleles | COPD *n* (%) (*N* = 703) | Control *n* (%) (*N* = 655) | *P* | OR (95%CI) |
|---|---|---|---|---|---|---|
| *H19* rs3741219 T>C | C | TT/TC/CC | 282/282/139 (40.11/40.11/19.77) | 216/311/128 (32.98/47.48/19.54) | 0.011 | – |
| | | T/C | 846/560 (60.17/39.83) | 743/567 (56.72/43.28) | 0.074 | 0.86 (0.74–1.01) |
| *MEG*3 rs7158663 A>G | G | AA/AG/GG | 335/262/106 (47.65/37.27/15.08) | 212/310/133 (32.37/47.33/20.31) | $1.24 \times 10^{-5}$ | – |
| | | A/G | 932/474 (66.29/33.71) | 734/576 (56.03/43.97) | $1.14 \times 10^{-5}$ | 0.64 (0.55–0.75) |
| *MALAT1* rs619586 A>G | G | AA/AG/GG | 629/72/2 (89.47/10.24/0.28) | 607/47/1 (92.67/7.18/0.15) | 0.117 | – |
| | | A/G | 1330/76 (94.59/5.41) | 1261/49 (96.26/3.74) | 0.048 | 1.47 (1.02–2.12) |
| *LINC00305* rs2850711 A>T | T | AA/AT/TT | 423/232/48 (60.17/33.00/6.83) | 375/234/46 (57.25/35.73/7.02) | 0.537 | – |
| | | A/T | 1078/328 (76.67/23.33) | 984/326 (75.11/24.89) | 0.366 | 0.91 (0.77–1.09) |
| *LINC00261* rs6048205 A>G | G | AA/AG/GG | 574/112/17 (81.65/15.93/2.42) | 536/109/10 (81.83/16.64/1.53) | 0.482 | – |
| | | A/G | 1260/146 (89.62/10.38) | 1181/129 (90.15/9.85) | 0.689 | 1.06 (0.82–1.36) |
| *LINC02227* rs2149954 G>A | A | GG/GA/AA | 284/361/58 (40.40/51.35/8.25) | 243/320/92 (37.10/48.85/14.05) | 0.003 | – |
| | | G/A | 929/477 (66.07/33.93) | 806/504 (61.53/38.47) | 0.015 | 0.82 (0.70–0.96) |
| *CDKN2B-AS1* rs4977574 A>G | G | AA/AG/GG | 219/344/140 (31.15/48.93/19.91) | 161/329/165 (24.58/50.23/25.19) | 0.008 | – |
| | | A/G | 782/624 (55.62/44.38) | 651/659 (49.69/50.31) | 0.002 | 0.78 (0.67–1.47) |

*P*—significance level of differences in allele and genotype frequencies between the groups ($\chi^2$ test for homogeneity of samples); OR (95%CI)—odds ratio for a minor allele and 95% confidence interval (basic allelic test).

Prior to the association analysis of candidate gene polymorphisms with developing COPD, we calculated allele and genotype frequencies in the groups and the correspondence of genotype frequencies to the Hardy–Weinberg equilibrium (Table 2). Observed genotype frequencies of all examined SNPs in the control group were in accordance with the Hardy–Weinberg test: *H19* (rs3741219) ($P_{H\text{-}W}$ = 0.4715), *MEG3* (rs7158663) ($P_{H\text{-}W}$ = 0.2863), *MALAT1* (rs619586) ($P_{H\text{-}W}$ = 0.5317), *LINC00305* (rs2850711) ($P_{H\text{-}W}$ = 0.3033), *LINC00261* (rs6048205) ($P_{H\text{-}W}$ = 0.2288), *CDKN2B-AS1* (rs4977574) ($P_{H\text{-}W}$ = 1), *LINC02227* (rs2149954) ($P_{H\text{-}W}$ = 0.4817).

Subsequently, we estimated the statistical significance of differences in distribution of allele and genotype frequencies between the groups, and the odds ratio values for the minor allele of each locus were calculated (basic allelic test). At the following stage, logistic regression was used to analyze the association of certain SNPs controlling for the quantitative and binary parameters (sex, age, smoking status and index), which were introduced into the regression equation as independent predictors (Table 3). Considering that the impact of single genes in the multifactorial disease development may be rather small, we searched for informative gene–gene interactions associated with COPD using APSampler software. At the

**Table 3.** Statistically significant results of association analysis of lncRNA polymorphic loci with COPD

| Gene, SNP | Minor allele | *N* | Genotype, model | $OR_{adj}$ (95%CI) | $P_{adj}$ | $P_{cor\text{-}FDR}$ |
|---|---|---|---|---|---|---|
| *H19* rs3741219 T>C | C | 1358 | TT<br>TC+CC, dominant | 1.00<br>0.74 (0.57–0.96) | 0.022 | 0.022 |
| | | | TT+CC<br>CT | 1.00<br>0.74 (0.57–0.96) | 0.021 | 0.022 |
| *MEG*3 rs7158663 A>G | G | 1358 | AA<br>AG+GG, dominant | 1.00<br>0.54 (0.42–0.70) | $2.167 \times 10^{-6}$ | $1.733 \times 10^{-5}$ |
| | | | log-additive | 0.68 (0.57–0.81) | $1.471 \times 10^{-5}$ | $5.884 \times 10^{-5}$ |
| *LINC02227* rs2149954 G>A | A | 1358 | GG+GA<br>AA, recessive | 1.00<br>0.55 (0.38–0.81) | 0.00171 | 0.0045 |
| | | | log-additive | 0.81 (0.67–0.96) | 0.0165 | 0.022 |
| *CDKN2B-AS1* rs4977574 A>G | G | 1358 | AA<br>AG+GG, dominant | 1.00<br>0.72 (0.55–0.95) | 0.018 | 0.022 |
| | | | log-additive | 0.79 (0.66–0.94) | 0.0077 | 0.0154 |

*N*—number of individuals included in regression analysis; $P_{adj}$—significance of the likelihood ratio test of the log-regression model controlling for sex, age, and smoking status and index; $OR_{adj}$—odds ratio controlling for all these factors, 95%CI—95% confidence interval for OR; $P_{cor\text{-}FDR}$—test significance after correction (Benjamini-Hochberg FDR Adjusted *P*-value); log-additive model for minor allele dosage—increase in minor allele dosage in a row: major allele homozygote (0), heterozygote (1), minor allele homozygote (2).

final stage, using multiple regression analysis with a stepwise forward inclusion of predictors followed by ROC analysis, a search for complex clinical-genetic models of COPD development was carried out (Fig. 1).

### *Association Analysis of SNPs of lncRNA Genes with COPD Development*

Statistically significant differences in distribution of allele and/or genotype frequencies between the groups of COPD patients and control group were observed for the *H19* (rs3741219), *MEG3* (rs7158663), *MALAT1* (rs619586), *LINC02227* (rs2149954), and *CDKN2B-AS1* (rs4977574) genes (Table 2).

The association of the *H19* (rs3741219) gene locus with COPD development was established for the dominant model ($P_{adj} = 0.022$, OR = 0.74) with homozygous major allele TT genotype as a risk marker ($P_{adj} = 0.022$; OR = 1.36, 95%CI 1.11–1.69).

The *MEG3* (rs7158663) locus was significantly associated with COPD in the dominant ($P_{adj} = 2.167 \times 10^{-6}$, OR = 0.54) and log-additive ($P_{adj} = 1.471 \times 10^{-5}$; OR = 0.68) models; a risk for developing the disease was associated with AA genotype ($P_{adj} = 2.167 \times 10^{-6}$; OR = 1.90, 95%CI 1.52–2.37).

The *LINC02227* (rs2149954) locus was associated with COPD in the recessive ($P_{adj} = 0.00171$, OR = 0.55) and log-additive ($P_{adj} = 0.0165$; OR = 0.81) models; major G allele was a risk marker ($P = 0.015$; OR = 1.22, 95%CI 1.04–1.42).

Association of the *CDKN2B-AS1* gene (rs4977574) with COPD was observed in the dominant ($P_{adj} = 0.018$; OR = 0.72) and log-additive ($P_{adj} = 0.0077$; OR = 0.79) models; homozygous major allele of AA genotype was a risk marker ($P_{adj} = 0.018$; OR =1.39, 95%CI 1.09–1.76).

The association of COPD with the *MALAT1* (rs619586) locus was detected only in the basic allelic test, and the minor allele (G allele) frequency was 5.41% in the patients groups and 3.74% in the control group ($P = 0.048$; OR = 1.47, 95%CI 1.02–2.12). However, a regression analysis revealed no significant associations in any model, which is due to a low frequency of minor G allele.

No statistically significant differences were observed in the distribution of genotype and allele frequencies of *LINC00305* (rs2850711) and *LINC00261* (rs6048205) gene polymorphisms between COPD patients and control group.

Accordingly, as a result of analysis of lncRNA gene polymorphisms, we have primarily revealed the data on the association of the *H19* (rs3741219), *MEG3* (rs7158663), *LINC02227* (rs2149954), and *CDKN2B-AS1* (rs4977574) gene loci and COPD.

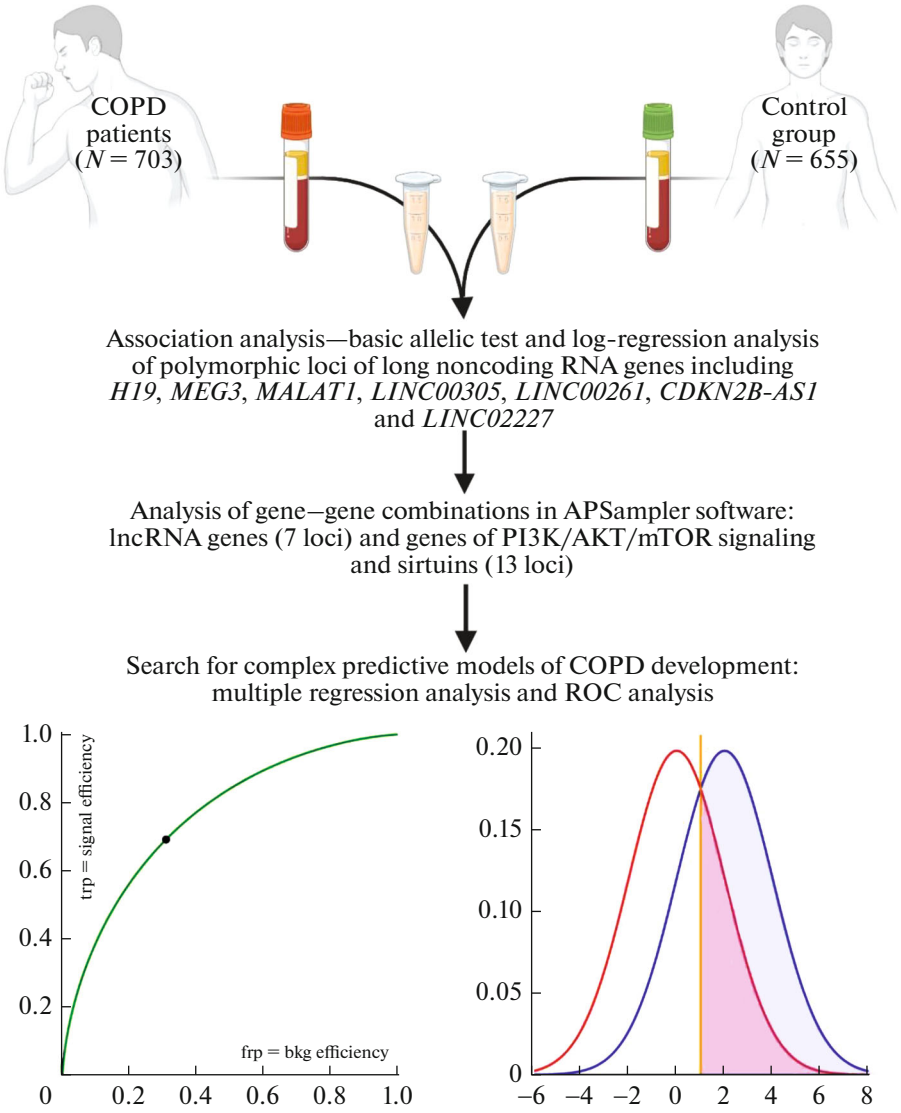


**Fig. 1.** Design of the study.

### *Analysis of Gene–Gene Interactions between Polymorphic Loci of lncRNA Genes and Genes PI3K/AKT/mTOR Signaling and Sirtuin Family of NAD-Dependent Deacetylases*

The lncRNAs selected for the study represent the regulators of various molecular cascades related to the cellular senescence, oxidative stress, and apoptosis; hence, it appears to be reasonable to consider a combined impact of polymorphic loci of lncRNA genes and genes belonging to the PI3K/AKT/mTOR signaling cascade and sirtuin family of NAD-dependent deacetylases [11]. The analysis included 20 SNPs, 13 of which are lncRNAs examined in the present study, and 13 SNPs belong to the *PIK3R1*, *AKT1*, *MTOR*, *PTEN*, *SIRT2*, *SIRT1*, *SIRT3*, and *SIRT6* gene, which were previously examined by our group [11]. We have obtained 8192 unique patterns, which were excluded from the combinations absent of alleles/genotypes of lncRNAs examined by our group. Subsequently, the selection criteria of detected combinations included $P_{FDR} < 0.01$ and OR < 0.4 (for protective markers) or OR > 2 (for risk markers). In total, 33 gene–gene interactions which corresponded to the selected criteria were determined; 15 of them were associated with an increased risk of COPD development, while 18 were protective ones. Table 4 presents the 12 most significant gene–gene interactions associated with COPD.

The most significant combinations of an increased risk of COPD development were determined by a combination of GG genotype of the *SIRT3* (rs536715) gene with G allele of the *LINC02227* locus (rs2149954) (OR = 2.31; $P_{FDR} = 1.12 \times 10^{-8}$) and AA genotype of the *MEG3* (rs7158663) locus (OR = 2.37;

$P_{FDR} = 1.77 \times 10^{-6}$). However, the majority of detected risk combinations included the G allele of the *PIK3R1* gene (rs3730089) in a combination with lncRNA alleles, including *LINC00305* (rs2850711) A allele, *CDKN2B-AS1* (rs4977574) A allele, and *MALAT1* (rs619586) G allele. The most significant combination associated with a diminished risk of COPD development included the *PIK3R1* (rs831125) A allele combined with the *SIRT3* (rs3782116) A allele, *SIRT3* (rs536715) AG genotype, and *LINC00261* (rs6048205) A allele (OR = 0.16, $P_{FDR} = 4.88 \times 10^{-10}$). A core component of the most significant protective gene–gene patterns consisted of combination of *SIRT3* (rs3782116) and *SIRT3* (rs536715) A alleles with *LINC00261* (rs6048205) A allele and *MALAT1* (rs619586) A allele.

The following three combinations consisted of the G allele of the *MEG3* gene (rs7158663) combined with the TT genotype of the *PIK3R1* gene locus (rs10515070) and A allele of the *PIK3R1* gene (rs3730089) (OR = 0.34; $P_{FDR} = 3.01 \times 10^{-7}$); with the TC genotype of the *SIRT6* gene (rs107251) (OR = 0.37; $P_{FDR} = 3.12 \times 10^{-7}$), and one of combinations included *CDKN2B-AS1* (rs4977574) G allele combined with *MTOR* (rs2295080) and *SIRT6* (rs107251) T alleles (OR = 0.28; $P_{FDR} = 6.61 \times 10^{-7}$). Accordingly, polymorphic loci of the *MEG3* (rs7158663), *MALAT1* (rs619586), and *CDKN2B-AS1* (rs4977574) genes in gene–gene interactions demonstrated the allele-specific effect, when certain alleles were the parts of combinations predisposing to COPD, while the alternative alleles of the same polymorphisms were present in the combinations linked to a reduced risk of the disease development. It should be noted that significant associations of the *LINC00261* (rs6048205), *LINC00305* (rs2850711), and *MALAT1* (rs619586) were established only with a combination with *PIK3R1* and sirtuin genes (*SIRT3* and *SIRT6*).

Therefore, the analysis of gene–gene interactions of the studied polymorphisms of lncRNA genes and genes belonging to the PI3K/AKT/mTOR signaling and sirtuins made it possible to detect highly informative combinations associated with COPD, which may indicate synergy between the examined genes.

### *The Search for Complex Clinical-Genetic Risk Models of COPD Development Using Multiple Regression Analysis*

The search for complex risk models of COPD development was carried out via multiple regression analysis with a stepwise forward inclusion of the most significant variables followed by ROC analysis to estimate the efficacy of obtained prognostic models. Highly informative gene–gene interactions identified at the previous stage of the analysis and single genotypes and alleles identified within the basic association analysis were selected as predictors. Subsequently, clinical-demographic variables (sex, age, smoking status and index) were added and the most significant predictive models were selected.

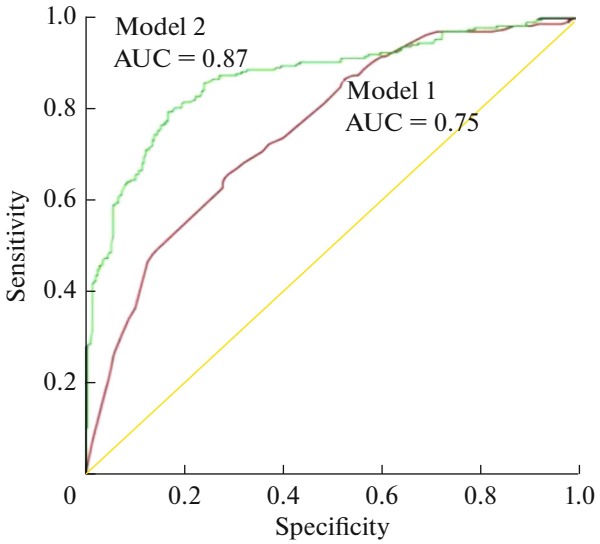


**Fig. 2.** Plot of area under the curve (ROC analysis) for evaluating the efficacy of predictive regression models. AUC—area under the curve. Complete characteristics are shown in Table 5. Model 1 includes only genetic markers—AUC = 0.75 (sensitivity of 65.7%, specificity of 71.2%); model 2 includes genetic markers and smoking index—AUC = 0.87 (sensitivity of 74.9%, specificity of 86.3%).

The first highly informative prognostic model of high risk of COPD development consisted of gene–gene interactions and single genes, including *LINC02227* (rs2149954) AA genotype and *PIK3R1* (rs831125) G allele ($P = 8.86 \times 10^{-24}$) (Table 5). ROC analysis of the obtained model resulted in its moderate level of prognostic ability, i.e., AUC = 0.75 (95%CI 0.71–0.80, sensitivity of 65.7%, specificity of 71.2%) (Fig. 2). The second significant model included the smoking index together with genetic markers and was characterized by high predicting ability ($P = 4.01 \times 10^{-48}$) AUC = 0.87 (95%CI 0.84–0.90) (sensitivity of 74.9%, specificity of 86.3%) (Table 5, Fig. 2), which points to the possibility of this regression model to effectively differentiate COPD patients from the healthy donors. As a result of multiple regression analysis, the prognostic significance of the determined markers was assessed and highly informative complex models of COPD risk accounting for both genetic and environmental factors have been identified.

## DISCUSSION

The association analysis of polymorphic loci of *H19*, *MEG3*, *MALAT1*, *LINC00305*, *LINC00261*, *CDKN2B-AS*, and *LINC02227* lncRNA genes with COPD was carried out, and a combined effect of lncRNA genes and previously examined genes of the

**Table 4.** Gene–gene combinations of polymorphic loci of lncRNA gene, genes of PI3K/AKT/mTOR signaling, and NAD-dependent deacetylases, which demonstrated the highest significance level of association with COPD

| Gene-gene combinations | COPD (frequency) | Control (frequency) | $P$ | $P_{FDR}$ | OR | 95%CI |
|---|---|---|---|---|---|---|
| Protective | | | | | | |
| *PIK3R1* (rs831125) A + *SIRT3* (rs3782116) A + *SIRT3* (rs536715) AG + *LINC00261* (rs6048205) A | 4.0 | 21.0 | $2.03 \times 10^{-12}$ | $4.88 \times 10^{-10}$ | 0.16 | 0.087–0.287 |
| *SIRT3* (rs3782116) A + *SIRT3* (rs536715) A + *LINC00261* (rs6048205) A + *MALAT1* (rs619586) A | 6.80 | 25.80 | $2.56 \times 10^{-11}$ | $2.52 \times 10^{-9}$ | 0.23 | 0.14–0.37 |
| *SIRT3* (rs3782116) A + *SIRT3* (rs536715) A + *MALAT1* (rs619586) A | 8.0 | 24.23 | $1.53 \times 10^{-10}$ | $8.05 \times 10^{-9}$ | 0.26 | 0.16–0.40 |
| *MEG3* (rs7158663) G + *PIK3R1* (rs10515070) TT + *PIK3R1* (rs3730089) A | 9.31 | 23.08 | $2.09 \times 10^{-8}$ | $3.01 \times 10^{-7}$ | 0.34 | 0.23–0.51 |
| *MEG3* (rs7158663) G + *SIRT6* (rs107251) TC | 12.09 | 26.88 | $2.19 \times 10^{-8}$ | $3.12 \times 10^{-7}$ | 0.37 | 0.26–0.53 |
| *MEG3* (rs7158663) G + *MTOR* (rs2295080) T + *SIRT6* (rs107251) T + *CDKN2B-AS1* (rs4977574) G | 6.94 | 20.78 | $5.74 \times 10^{-8}$ | $6.61 \times 10^{-7}$ | 0.28 | 0.17–0.46 |
| Risk | | | | | | |
| *SIRT3* (rs536715) GG + *LINC02227* (rs2149954) G | 63.00 | 42.41 | $2.61 \times 10^{-10}$ | $1.12 \times 10^{-8}$ | 2.31 | 1.77–3.01 |
| *PIK3R1* (rs3730089) G + *PTEN* (rs701848) C + *LINC00305* (rs2850711) A | 52.11 | 33.33 | $1.42 \times 10^{-7}$ | $1.35 \times 10^{-6}$ | 2.178 | 1.62–2.92 |
| *MEG3* (rs7158663) AA+ *SIRT3* (rs536715) GG | 30.79 | 15.80 | $1.99 \times 10^{-7}$ | $1.77 \times 10^{-6}$ | 2.37 | 1.69–3.32 |
| *PIK3R1* (rs3730089) G + *SIRT3* (rs536715) G + *SIRT6* (rs107251) C + *LINC00305* (rs2850711) A | 63.42 | 45.70 | $2.06 \times 10^{-7}$ | $1.82 \times 10^{-6}$ | 2.06 | 1.56–2.72 |
| *PIK3R1* (rs3730089) G + *CDKN2B-AS1* (rs4977574) A | 20.36 | 9.65 | $4.44 \times 10^{-6}$ | $2.2 \times 10^{-5}$ | 2.39 | 1.62–3.52 |
| *PIK3R1* (rs831125) A + *PIK3R1* (rs3730089) G + *MALAT1* (rs619586) G | 8.06 | 3.38 | 0.0008 | 0.002 | 2.75 | 1.47–5.14 |

**Table 5.** Significant predictive regression models of COPD development

| Predictor | *b* | $P_{Wald}$ | OR | 95%CI |
|---|---|---|---|---|
| Model 1 | | | | |
| *PIK3R1* (rs3730089) G + *SIRT3* (rs536715) G + *SIRT6* (rs107251) C + *LINC00305* (rs2850711) A | 0.635 | 0.0014 | 1.887 | 1.28–2.79 |
| *PIK3R1* (rs10515070) TT + *PIK3R1* (rs3730089) A | –0.645 | 0.0039 | 0.525 | 0.34–0.81 |
| *PIK3R1* (rs831125) G | 1.150 | $4 \times 10^{-8}$ | 3.158 | 2.09–4.76 |
| *LINC02227* (rs2149954) AA | –0.587 | 0.0499 | 0.556 | 0.31–1.0 |
| *MEG3* (rs7158663) AA + *SIRT3* (rs536715) GG | 0.468 | 0.0471 | 1.597 | 1.01–2.53 |
| *SIRT3* (rs3782116) A + *SIRT3* (rs536715) A + *LINC00261* (rs6048205) A | –1.239 | 0.0002 | 0.29 | 0.15–0.55 |
| *MEG*3 (rs7158663) G + *MTOR* (rs2295080) T + *SIRT6* (rs107251) T + *CDKN2B-AS1* (rs4977574) G | –1.298 | 0.0002 | 0.273 | 0.14–0.54 |
| Constant | –0.511 | 0.001 | 0.600 | |
| $P = 8.86 \times 10^{-24}$<br>AUC = 0.75 (95%CI 0.71–0.80) (sensitivity of 65.7%, specificity of 71.2% | | | | |
| Model 2 | | | | |
| *PIK3R1* (rs3730089) G + *SIRT3* (rs536715) G + *SIRT6* (rs107251) C + *LINC00305* (rs2850711) A | 0.521 | 0.0244 | 1.684 | 1.07–2.65 |
| *PIK3R1* (rs10515070) TT + *PIK3R1* (rs3730089) A | –0.813 | 0.0023 | 0.444 | 0.26–0.75 |
| *PIK3R1* (rs831125) G | 1.436 | $6 \times 10^{-9}$ | 4.204 | 2.59–6.83 |
| *LINC02227* (rs2149954) AA | –0.814 | 0.0243 | 0.443 | 0.22–0.9 |
| *MEG*3 (rs7158663) AA + *SIRT3* (rs536715) GG | 0.766 | 0.0067 | 2.151 | 1.24–3.74 |
| *SIRT3* (rs3782116) A + *SIRT3* (rs536715) A + *LINC00261* (rs6048205) A | –0.825 | 0.0224 | 0.438 | 0.22–0.89 |
| *MEG3* (rs7158663) G + *MTOR* (rs2295080) T + *SIRT6* (rs107251) T + *CDKN2B-AS1* (rs4977574) G | –1.184 | 0.0024 | 0.306 | 0.14–0.66 |
| Smoking index | 0.067 | $4 \times 10^{-21}$ | 1.069 | 1.07–2.65 |
| Constant | –2.529 | $7 \times 10^{-16}$ | 0.080 | |
| $P = 4.01 \times 10^{-48}$<br>AUC = 0.87 (95%CI 0.84–0.90) (sensitivity of 74.9%, specificity of 86.3%) | | | | |

*b*—beta coefficient for the variable; $P_{Wald}$ – Wald statistic; OR—exp(*b*) odds ratio and 95%CI—95% confidence interval for OR; *P*—*p*-value for likelihood ratio (LR) test; AUC—area under the curve.

PI3K/AKT/mTOR signaling cascade and sirtuin family of NAD-dependent deacetylases genes in development of disease was examined [11].

We have established the association of *H19* rs3741219 with COPD; the TT genotype was a risk marker, while carriers of the minor C allele had a diminished disease risk. T to C substitution at rs3741219 results in the formation of miR-146b-3p and miR-1539 binding sites and a reduced level of *H19* expression in C allele carriers [12]. According to the functional analysis, rs3741219 polymorphism is located at the DNA region which binds to regulatory proteins NRSF and YY1; the changes in the expression level based on the GTEx portal and HaploReg v3 databases occurs also in lung tissue. The *H19* (imprinted maternally expressed transcript) gene is located on chromosome 11p15.5 in the unique cluster of genes known as H19/IGF-2 locus, which is known to be genetically imprinted [https://www.ncbi.nlm.nih.gov/gene/283120]. *H19* plays a role as a key component of regulatory networks involved in the pathogenesis of several cancers and fibrosis via stimulation of autophagy, inhibiting apoptosis and enhanced epithelial-mesenchymal transition (EMT), and activation of TGF-β/SMAD3 and mTOR signaling [13, 14]. Namely, it was shown that *H19* inhibited miR-200a and stimulated PDCD4 (programmed cell death 4), enhancing apoptosis of lung vessel cells [15]. *H19* stimulates ferroptosis of lung cells by ferritin 1 via inhibiting miR-19b-3p [16]. *H19* stimulates fibrous changes in the respiratory tract via inhibition of the components of PTEN/AKT signaling [17]. *H19* binds miR-29a-3p and activates TNFRSF1A (TNF receptor superfamily member 1A), which causes a stimulation of profibrotic and inflammatory phenotypes of lung tissue by downstream regulatory molecules [18]. Acting as competitive endogenous RNA (ceRNA), *H19* binds a set of microRNAs (miR-200a, miR-107, miR-17, miR-6515-3p, miR-138, and miR-203), thus promoting proliferation, invasion, and metastasis of cancer lung cells [19]. Therefore, an enhanced expression of *H19* represents an unfavorable factor of development of a spectrum of pathological states. The *H19* gene polymorphisms, including rs3741219, have been widely studied in various cancer types [20]. Their analysis in other diseases remain insignificant; for instance, no association of the *H19* rs3741219 locus with type 2 diabetes mellitus was reported [21].

The most significant associations with COPD were determined for the *MEG3* (rs7158663) locus with A allele and AA genotype as risk markers of this disease, while minor G allele carriers were present among healthy individuals. *MEG3* (rs7158663) was identified as the most common component of gene–gene combinations associated with COPD development together with polymorphic loci of *PIK3R1*, *MTOR*, *SIRT6*, and *SIRT3* genes. *MEG3* (maternally expressed 3) gene is located at 14q32.2 and regulates the mitochondrial apoptotic pathway [22]. *MEG3* regulates TGF-β/SMAD3- and Wnt-signaling cascades and PI3K/AKT pathway activity [23]. According to the data of functional analysis, rs7158663 modifies binding sites for several transcription factors (PAX8, ATF6, and PPARG) and motifs for binding with transcription regulators (ARNT2, FOXP3, BHLHE40, and ELF5). According to the GTEx portal, the changes in expression level occur in blood mononuclear cells and arterial endothelium. In accordance with the lncRNASNP2 database (https://guolab.wchscu.cn/lncRNASNP/), rs7158663 enables the changes in the folding structure of local RNA and affects miRNA (miR-4307 and miR-1265) and lncRNA binding, which, in turn, causes changes in expression of corresponding miRNA or *MEG3* [24]. The study by Gao et al. (2021) demonstrated that *MEG3* expression in large intestine tissues was significantly lower in major A allele carriers compared to minor allele homozygotes (GG) [24]. The association of the AA genotype and A allele of *MEG3* (rs7158663) with type 2 diabetes mellitus was established [21]. It was reported that *MEG3* was involved in inflammatory response regulation via suppressing miR-138, modifying expression of molecules of NF-κB-signaling and proinflammatory cytokines [25]. *MEG3* binds miR-181a-5p with subsequent stimulation of the PTEN/pSTAT5/SOCS1 signaling cascade in macrophages [26], while miR-133a-3p inhibition results in exaggerated *SIRT1* expression and reduces the degree of lung impairment [27]. One of the *MEG3* target molecules is miR-181b-3p, which is related to pathological angiogenesis in lungs and, hence, forms the basis for development of emphysema [28]. *MEG3* may reduce the degree of cellular senescence of lung epithelial cells via inhibiting miR-125a-5p [29]. *MEG3* acts as inhibitor of the miR-664a-3p/FHL1 cascade, which is related to oxidative stress mediated by exposure to cigarette smoke [30].

Considering a wide spectrum of functions of this lncRNA, which is involved in the pathogenesis of age-associated disorders (regulation of inflammatory cascade oxidative stress, apoptosis, and PI3K/AKT pathway), a diminished expression of the *MEG3* related to rs7158663 appears to be a significant risk factor for COPD development.

We have established the association of *CDKN2B-AS* (rs4977574) A allele with COPD; this variant also comprised the risk gene–gene combination with *PIK3R1* (rs3730089) gene polymorphism. The *CDKN2B-AS* (rs4977574) G allele was more frequent in healthy individuals and was included in the informative protective gene–gene combination combined with the alleles of the *MEG3* (rs7158663), *MTOR* (rs2295080), and *SIRT6* (rs107251) genes. The *CDKN2B-AS* (CDKN2B antisense RNA) lncRNA is transcribed from the antisense chain of the cluster of genes *CDKN2A/p16*$^{INK4A}$, *CDKN2A/p14*$^{ARF}$, and *CDKN2B/p15*$^{INK4B}$, which encode inhibitors of cyclin-

dependent kinase 4 (CDK4) and (*MTAP*) (methylthioadenosine phosphorylase) at the 9p21.3 region, which play a key role in the control of cellular proliferation, apoptosis, and cellular senescence [https://www.ncbi.nlm.nih.gov/gene/100048912]. Molecular actions of *CDKN2B-AS1* are mediated by its interaction with the proteins of polycomb repressive complex 1 and 2 (PRC1 and PRC2); this complex enables epigenetic *cis*-activation of *CDKN2B-CDKN2A* cluster target genes [31]. The examined rs4977574 affects the expression level of *CDKN2B-AS1*; since according to the GTEx portal and HaploReg v3 database the changes in gene expression were detected in blood, rs4977574 changes the motifs for regulatory proteins. The presence of the G allele at *CDKN2B-AS1* rs4977574 causes an enhanced formation of linear isoforms of the *CDKN2B-AS1* molecule, which enable polycomb proteins binding together with a reduced expression of circular *CDKN2B-AS1* transcripts [32]. The results of genome-wide association studies (GWAS) demonstrated a significant association of rs4977574 G allele with ischemic heart disease (IHD) [33]. It should be noted that *CDKN2B-AS* is related to a spectrum of age-related disorders [34]. Association studies of the *CDKN2B-AS* gene with COPD are absent; however, a reduced expression level of circulating *CDKN2B-AS* in plasma was related to COPD escalation [35].

The *LINC02227* (long intergenic non-protein coding RNA 2227 (minus strand)) gene is located at 5q33.3, and rs2149954 was primarily detected in 2014 as a result of genome-wide studies as associated with longevity [36]. The rs2149954 A allele was linked to a reduced risk for developing cardiovascular disorders and arterial hypertension in middle age and age of longevity [36]. The functional approach demonstrated that rs2149954 was located at sites of DNAase I hypersensitivity and sites of binding of transcription factors and histone enhancers. We have revealed the association of *LINC02227* (rs2149954) with COPD, and a major G allele was a risk marker, while the A allele and AA genotype marked resistance to disease. The *LINC02227* (rs2149954) G allele was identified in the informative gene–gene combination with a polymorphic variant of the *SIRT3* gene (rs536715) and mitochondrial deacetylase, which plays a key role in regulating cellular senescence [37].

As a result of polygenic analysis, we determined the informative combinations associated with COPD, which consisted of polymorphic loci of lncRNA genes, including *LINC00305* (rs2850711), *LINC00261* (rs6048205), and *MALAT1* (rs619586), which showed no association within the basic analysis of single SNPs. The *LINC00305* (rs2850711) A allele was detected in the two most significant risky gene–gene combinations linked to COPD in a combination with polymorphic variants of the *PIK3R1*, *PTEN*, and sirtuin (*SIRT6* and *SIRT3*) genes. *LINC00305* (long intergenic non-protein coding RNA 305) is located at the 18q22.1 region and regulates NF-κB signaling, and enhances the expression of the genes encoding proinflammatory cytokines [38]. The *LINC00261* (rs6048205) A allele was detected in informative gene–gene combinations associated with a reduced risk of COPD development in a combination with polymorphic loci of the *SIRT3, PIK3R1*, and *MALAT1* genes. Previously, it was reported that a minor T allele was associated with rheumatoid arthritis, and minor allele homozygotes (TT) and heterozygotes (AT) of rs2850711 demonstrated an increased level of *LINC00305* expression [39]. *LINC00261* (long intergenic non-protein coding RNA 261 (minus strand)) is located at 20p11.21, and rs6048205 was identified as associated with glucose level at a genome-wide level [40]. On the basis of functional analysis, rs6048205 is located in the 5' region of the gene and is linked to SNP at the 3' untranslated region of the *FOXA2* (forkhead box A2) gene, which is involved in the transcription regulation and represents a factor responsible for regulation of glucose homeostasis [41]. The rs6048205 changes binding sites with regulatory proteins (FOXA1, SP1, CEBPB, P300, and TCF4) and transcription factors. According to the GTEx portal, the A allele is linked to the enhanced gene expression. *LINC00261* is involved in the regulation of apoptosis and cellular homeostasis, functioning as a negative regulator of Notch- and NF-κB signaling, and modulates SMAD3 function, which represents a key component of TGF-β1 signaling via suppressing EMT [42].

*MALAT1* (metastasis associated lung adenocarcinoma transcript 1) is located at the 11q13.1 region and regulates the expression of IL6 and TNF-a [43]. *MALAT1* is actively expressed in the majority of human tissues; however, an enhanced *MALAT1* expression is frequently observed in various types of malignant neoplasms and is associated with unfavorable prognosis and metastasis risk [44]. The rs619586 is located in the region of DNase I with regulatory proteins and transcription factors and histone marks. According to the GTEx portal, changes in expression were confirmed in 14 tissues, including blood, arteries, and skeletal muscles, and the G allele results in exaggerated *MALAT1* expression. We have demonstrated that the *MALAT1* (rs619586) A allele was included in two protective combinations associated with COPD, whereas the G allele related to the increased expression of *MALAT1* gene was included in gene–gene combination of the enhanced risk of COPD development. It was established that the level of *MALAT1* expression positively correlated with the severity of disease and the level of proinflammatory cytokines in COPD patients, bound to miR-125b, miR-146a, and miR-203, provided an increased inflammation, and functioned as a regulator of MAPK/NF-κB signaling [45].

As a result of the conducted study, we have reported for the first time a significant association of polymorphic loci of lncRNA genes including *H19*,

*MEG3*, *CDKN2B-AS*, and *LINC02227* with COPD and gene–gene interactions between the genes of the PI3K/AKT/mTOR signaling and sirtuin family of NAD-dependent deacetylases genes and the examined lncRNAs (*MEG3*, *CDKN2B-AS*, *MALAT1*, *LINC00261*, and *LINC00305*). The data obtained indicate that molecular pathogenesis of COPD may include the mechanisms related to the impaired regulation of stress reactions preventing cellular senescence, which are significantly affected by long noncoding RNAs. Noncoding RNAs that are related to apoptosis, cellular homeostasis, oxidative stress, and cellular senescence as potential biomarkers and therapy targets may become the basis for the development of a novel strategy of COPD diagnostics and treatment.

## FUNDING

This study was supported by the Russian Science Foundation (no. 23-25-00019, https://rscf.ru/project/23-25-00019/) using the equipment from the Center of Collective Use Biomics and USE KODINK (Institute of Biochemistry and Genetics, Ufa Federal Research Center, Russian Academy of Sciences).

## ETHICS APPROVAL AND CONSENT TO PARTICIPATE

The study was approved by the Bioethics Committee at the Institute of Biochemistry and Genetics, Ufa Scientific Center, Russian Academy of Sciences (Protocol no. 17, dated December 7, 2010) and Institute of Biochemistry and Genetics, Ufa Federal Research Center, Russian Academy of Sciences (Protocol no. 19, dated November 1, 2022). An informed voluntary consent for the use of biological material in the study was obtained from all the participants.

## CONFLICT OF INTEREST

The authors of this work declare that they have no conflicts of interest.

## REFERENCES

1. Agusti, A., Celli, B.R., Criner, G.J., et al., Global initiative for chronic obstructive lung disease 2023 report: GOLD executive summary, *Eur. Respir. J.*, 2023, vol. 61, no. 4, p. 2300239. https://doi.org/10.1183/13993003.00239-2023
2. Chuchalin, A.G., Avdeev, S.N., Aisanov, Z.R., et al., Chronic obstructive pulmonary disease: federal clinical guidelines for diagnosis and treatment, *Pul'monologiya*, 2022, vol. 32, no. 3, pp. 356—392. https://doi.org/10.18093/0869-0189-2022-32-3-356-392
3. Agustí, A., Melén, E., DeMeo, D.L., et al., Pathogenesis of chronic obstructive pulmonary disease: understanding the contributions of gene-environment interactions across the lifespan, *Lancet Respir. Med.*, 2022, vol. 10, no. 5, pp. 512—524. https://doi.org/10.1016/S2213-2600(21)00555-5
4. Brandsma, C.A., van den Berge, M., Hackett, T.L., et al., Recent advances in chronic obstructive pulmonary disease pathogenesis: from disease mechanisms to precision medicine, *J. Pathol.*, 2020, vol. 250, no. 5, pp. 624—635. https://doi.org/10.1002/path.5364
5. Bridges, M.C., Daulagala, A.C., and Kourtidis, A., LNCcation: lncRNA localization and function, *J. Cell Biol.*, 2021, vol. 220, no. 2. https://doi.org/10.1083/jcb.202009045
6. Devadoss, D., Long, C., Langley, R.J., et al., Long noncoding transcriptome in chronic obstructive pulmonary disease, *Am. J. Respir. Cell Mol. Biol.*, 2019, vol. 61, no. 6, pp. 678—688. https://doi.org/10.1165/rcmb.2019-0184TR
7. Ward, L.D. and Kellis, M., HaploReg v4: systematic mining of putative causal variants, cell types, regulators and target genes for human complex traits and disease, *Nucleic Acids Res.*, 2016, vol. 44, no. D1, pp. D877—D881. https://doi.org/10.1093/nar/gkv1340
8. Korytina, G.F., Akhmadishina, L.Z., Aznabaeva, Y.G., et al., Associations of the NRF2/KEAP1 pathway and antioxidant defense gene polymorphisms with chronic obstructive pulmonary disease, *Gene*, 2019, vol. 692, pp. 102—112. https://doi.org/10.1016/j.gene.2018.12.061
9. González, J.R., Armengol, L., Solé, X., et al., SNPassoc: an R package to perform whole genome association studies, *Bioinformatics*, 2007, vol. 23, no. 5, pp. 644—645. https://doi.org/10.1093/bioinformatics/btm025
10. Favorov, A.V., Andreewski, T.V., Sudomoina, M.A., et al., A Markov chain Monte Carlo technique for identification of combinations of allelic variants underlying complex diseases in humans, *Genetics*, 2005, vol. 171, no. 4, pp. 2113—2121. https://doi.org/10.1534/genetics.105.048090
11. Korytina, G.F., Akhmadishina, L.Z., Markelov, V.A., et al., Role of PI3K/AKT/mTOR signaling pathway and sirtuin genes in chronic obstructive pulmonary disease development, *Vavilovskii Zh. Genet. Sel.*, 2023, vol. 27, no. 5, pp. 512—521. https://doi.org/10.18699/VJGB-23-62
12. Song, Y., Xing, H., Zhou, L., et al., LncRNA H19 modulated by miR-146b-3p/miR-1539-mediated allelic regulation in transarterial chemoembolization of hepatocellular carcinoma, *Arch. Toxicol.*, 2021, vol. 95, no. 9, pp. 3063—3070. https://doi.org/10.1007/s00204-021-03119-8
13. Lu, Q., Guo, Z., Xie, W., et al., The lncRNA H19 mediates pulmonary fibrosis by regulating the miR-196a/COL1A1 axis, *Inflammation*, 2018, vol. 41, no. 3, pp. 896—903. https://doi.org/10.1007/s10753-018-0744-4
14. Xu, J.L., Hua, T., Ding, J., et al., FOXF2 aggravates the progression of non-small cell lung cancer through targeting lncRNA H19 to downregulate PTEN, *Eur. Rev. Med. Pharmacol. Sci.*, 2019, vol. 23, no. 24, pp. 10796—10802. https://doi.org/10.26355/eurrev_201912_19782

15. Wang, R., Zhou, S., Wu, P., et al., Identifying involvement of H19-miR-675-3p-IGF1R and H19-miR-200a-PDCD4 in treating pulmonary hypertension with melatonin, *Mol. Ther. Nucleic Acids*, 2018, vol. 13, pp. 44—54. https://doi.org/10.1016/j.omtn.2018.08.015
16. Zhang, R., Pan, T., Xiang, Y., et al., Curcumenol triggered ferroptosis in lung cancer cells via lncRNA H19/miR-19b-3p/FTH1 axis, *Bioact. Mater.*, 2021, vol. 13, pp. 23—36. https://doi.org/10.1016/j.bioactmat.2021.11.013
17. Yu, H., Qi, N., and Zhou, Q., LncRNA H19 inhibits proliferation and migration of airway smooth muscle cells induced by PDGF-BB through miR-21/PTEN/Akt axis, *J. Asthma Allergy*, 2021, vol. 14, pp. 71—80. https://doi.org/10.2147/JAA.S291333
18. Bu, N., Gao, Y., Zhao, Y., et al., LncRNA H19 via miR-29a-3p is involved in lung inflammation and pulmonary fibrosis induced by neodymium oxide, *Ecotoxicol. Environ. Saf.*, 2022, vol. 247. https://doi.org/10.1016/j.ecoenv.2022.114173
19. Liao, J., Chen, B., Zhu, Z., et al., Long noncoding RNA (lncRNA) H19: an essential developmental regulator with expanding roles in cancer, stem cell differentiation, and metabolic diseases, *Genes Dis.*, 2023, vol. 10, no. 4, pp. 1351—1366. https://doi.org/10.1016/j.gendis.2023.02.008
20. Li, L., Huang, Q., Yan, F., et al., Association between long non-coding RNA H19 polymorphisms and breast cancer risk: a meta-analysis, *Women Health,* 2022, vol. 62, no. 6, pp. 565—575. https://doi.org/10.1080/03630242.2022.2096748
21. Ghaedi, H., Zare, A., Omrani, M.D., et al., Genetic variants in long noncoding RNA H19 and MEG3 confer risk of type 2 diabetes in an Iranian population, *Gene*, 2018, vol. 675, pp. 265—271. https://doi.org/10.1016/j.gene.2018.07.002
22. Al-Rugeebah, A., Alanazi, M., and Parine, N.R., MEG3: an oncogenic long non-coding RNA in different cancers, *Pathol. Oncol. Res.*, 2019, vol. 25, no. 3, pp. 859—874. https://doi.org/10.1007/s12253-019-00614-3
23. Gokey, J.J., Snowball, J., Sridharan, A., et al., MEG3 is increased in idiopathic pulmonary fibrosis and regulates epithelial cell differentiation, *JCI Insight*, 2018, vol. 3, no. 17. https://doi.org/10.1172/jci.insight.122490
24. Gao, X., Li, X., Zhang, S., et al., The association of MEG3 gene rs7158663 polymorphism with cancer susceptibility, *Front. Oncol.*, 2021, vol. 11. https://doi.org/10.3389/fonc.2021.796774
25. Li, R., Fang, L., Pu, Q., et al., MEG3-4 is a miRNA decoy that regulates IL-1β abundance to initiate and then limit inflammation to prevent sepsis during lung infection, *Sci. Signal*, 2018, vol. 11, no. 536. https://doi.org/10.1126/scisignal.aao2387
26. Su, Y., Silva, J.D., Doherty, D., et al., Mesenchymal stromal cells-derived extracellular vesicles reprogramme macrophages in ARDS models through the miR-181a-5p-PTEN-pSTAT5-SOCS1 axis, *Thorax*, 2023, vol. 78, no. 6, pp. 617—630. https://doi.org/10.1136/thoraxjnl-2021-218194
27. Chen, L., Xie, W., Wang, L., et al., MiRNA-133a aggravates inflammatory responses in sepsis by targeting SIRT1, *Int. Immunopharmacol.*, 2020, vol. 88. https://doi.org/10.1016/j.intimp.2020.106848
28. Green, C.E., Clarke, J., Bicknell, R., et al., Pulmonary microRNA changes alter angiogenesis in chronic obstructive pulmonary disease and lung cancer, *Biomedicines*, 2021, vol. 9, no. 7. https://doi.org/10.3390/biomedicines9070830
29. Wu, H., Ma, H., Wang, L., et al., Regulation of lung epithelial cell senescence in smoking-induced COPD/emphysema by microR-125a-5p via Sp1 mediation of SIRT1/HIF-1a, *Int. J. Biol. Sci.*, 2022, vol. 18, no. 2, pp. 661—674. https://doi.org/10.7150/ijbs.65861
30. Zhong, S., Chen, C., Liu, N., et al., Overexpression of hsa-miR-664a-3p is associated with cigarette smoke-induced chronic obstructive pulmonary disease via targeting FHL1, *Int. J. Chron. Obstruct. Pulmon. Dis.*, 2019, vol. 14, pp. 2319—2329. https://doi.org/10.2147/COPD.S224763
31. Guil, S., Soler, M., Portela, A., et al., Intronic RNAs mediate EZH2 regulation of epigenetic targets, *Nat. Struct. Mol. Biol.*, 2012, vol. 19, no. 7, pp. 664—670. https://doi.org/10.1038/nsmb.2315
32. Holdt, L.M. and Teupser, D., Long noncoding RNA *ANRIL*: lnc-ing genetic variation at the chromosome 9p21 locus to molecular mechanisms of atherosclerosis, *Front. Cardiovasc. Med.*, 2018, vol. 5. https://doi.org/10.3389/fcvm.2018.00145
33. van der Harst, P. and Verweij, N., Identification of 64 novel genetic loci provides an expanded view on the genetic architecture of coronary artery disease, *Circ. Res.*, 2018, vol. 122, no. 3, pp. 433—443. https://doi.org/10.1161/CIRCRESAHA.117.312086
34. Cunnington, M.S., Santibanez Koref, M., Mayosi, B.M., et al., Chromosome 9p21 SNPs associated with multiple disease phenotypes correlate with *ANRIL* expression, *PLoS Genet.*, 2010, vol. 6, no. 4. https://doi.org/10.1371/journal.pgen.1000899
35. Ge, J., Geng, S., and Jiang, H., Long noncoding RNAs antisense noncoding RNA in the INK4 locus (*ANRIL*) correlates with lower acute exacerbation risk, decreased inflammatory cytokines, and mild GOLD stage in patients with chronic obstructive pulmonary disease, *J. Clin. Lab. Anal.*, 2019, vol. 33, no. 2. https://doi.org/10.1002/jcla.22678
36. Deelen, J., Beekman, M., Uh, H.W., et al., Genome-wide association meta-analysis of human longevity identifies a novel locus conferring survival beyond 90 years of age, *Hum. Mol. Genet.*, 2014, vol. 23, no. 16, pp. 4420—4432. https://doi.org/10.1093/hmg/ddu139
37. Sun, W., Liu, C., Chen, Q., et al., SIRT3: a new regulator of cardiovascular diseases, *Oxid. Med. Cell Longevity*, 2018, vol. 2018. https://doi.org/10.1155/2018/7293861
38. Zhang, D.D., Wang, W.T., Xiong, J., et al., Long noncoding RNA LINC00305 promotes inflammation by activating the AHRR-NF-κB pathway in human monocytes, *Sci. Rep.*, 2017, vol. 7. https://doi.org/10.1038/srep46204

39. Wahba, A.S., Ibrahim, M.E., Mesbah, N.M., et al., Serum LINC00305 expression and its genetic variant rs2850711 are associated with clinical and laboratory features of rheumatoid arthritis, *Br. J. Biomed. Sci.*, 2020, vol. 77, no. 3, pp. 142—147. https://doi.org/10.1080/09674845.2020.1744942

40. Manning, A.K., Hivert, M.F., Scott, R.A., et al., A genome-wide approach accounting for body mass index identifies genetic variants influencing fasting glycemic traits and insulin resistance, *Nat. Genet.*, 2012, vol. 44, no. 6, pp. 659—669. https://doi.org/10.1038/ng.2274

41. Johnson, M.E., Schug, J., Wells, A.D., et al., Genome-wide analyses of ChIP-Seq derived FOXA2 DNA occupancy in liver points to genetic networks underpinning multiple complex traits, *J. Clin. Endocrinol. Metab.*, 2014, vol. 99, no. 8, p. E1580-5. https://doi.org/10.1210/jc.2013-4503

42. Chen, Z., Xiang, L., Li, L., et al., TGF-β1 induced deficiency of linc00261 promotes epithelial-mesenchymal transition and stemness of hepatocellular carcinoma via modulating SMAD3, *J. Transl. Med.*, 2022, vol. 20, no. 1, p. 75. https://doi.org/10.1186/s12967-022-03276-z

43. Zhu, R., Liu, X., and He, Z., Long non-coding RNA H19 and MALAT1 gene variants in patients with ischemic stroke in a northern Chinese Han population, *Mol. Brain*, 2018, vol. 11, no. 1, p. 58. https://doi.org/10.1186/s13041-018-0402-7

44. Chen, M., Cai, D., Gu, H., et al., MALAT1 rs619586A/G polymorphisms are associated with decreased risk of lung cancer, *Medicine* (Baltimore), 2021, vol. 100, no. 12. https://doi.org/10.1097/MD.0000000000023716

45. Wu, W., Wang, S., Zhang, L., et al., Mechanistic studies of MALAT1 in respiratory diseases, *Front. Mol. Biosci.*, 2022, vol. 9. https://doi.org/10.3389/fmolb.2022.1031861

*Translated by A. Kazantseva*