

**АССОЦИАЦИЯ ПОЛИМОРФНЫХ ЛОКУСОВ ГЕНОВ ДЛИННЫХ НЕКОДИРУЮЩИХ РНК (*H19*, *MEG3*, *MALAT1*, *LINC00305*, *LINC00261*, *LINC02227* И *CDKN2B-AS1*) С ХРОНИЧЕСКОЙ ОБСТРУКТИВНОЙ БОЛЕЗНЬЮ ЛЕГКИХ**

© 2024 г. Г. Ф. КORYТИНА<sup>1,2,\*</sup>, Л. З. АХМАДИШИНА<sup>1,3</sup>, В. А. Маркелов<sup>1,2</sup>, Т. Р. Насибуллин<sup>1</sup>, Ю. Г. Азнабаева<sup>2</sup>, О. В. Кочетова<sup>1</sup>, Н. Н. Хуснутдинова<sup>1</sup>, А. П. Ларкина<sup>1</sup>, Н. Ш. Загидуллин<sup>2</sup>, Т. В. Викторова<sup>2</sup>

<sup>1</sup>Институт биохимии и генетики – обособленное структурное подразделение Уфимского федерального исследовательского центра Российской академии наук, Уфа, 450054 Россия

<sup>2</sup>Башкирский государственный медицинский университет, Уфа, 450000 Россия

<sup>3</sup>Уфимский государственный нефтяной технический университет, Уфа, 450006 Россия

\*e-mail: guly\_kory@mail.ru

Поступила в редакцию 21.02.2024 г.

После доработки 09.04.2024 г.

Принята к публикации 17.04.2024 г.

Хроническая обструктивная болезнь легких (ХОБЛ) – это хроническое заболевание, возникающее в результате динамических, кумулятивных ген–средовых взаимодействий, результатом которых является повреждение легочной ткани и изменение ее нормального функционирования, связанное с ускоренным клеточным старением. Длинные некодирующие РНК (днРНК) являются важными эпигенетическими регуляторами различных аспектов клеточного старения. Цель настоящего исследования – выявление ассоциации полиморфных вариантов генов днРНК *H19*, *MEG3*, *MALAT1*, *LINC00305*, *LINC00261*, *CDKN2B-AS1*, *LINC02227* с ХОБЛ. В работе были использованы образцы ДНК больных ХОБЛ ( $N = 703$ ) и здоровых индивидов ( $N = 655$ ), полиморфные локусы анализировали методом ПЦР в реальном времени. Ассоциация с развитием ХОБЛ была установлена для генов *H19* (rs3741219), *MEG3* (rs158663), *LINC02227* (rs2149954), *MALAT1* (rs619586) и *CDKN2B-AS1* (rs4977574). Полигенный анализ позволил выявить информативные ген–генные комбинации, включающие полиморфные варианты днРНК и генов, кодирующих компоненты молекулярных каскадов, связанных с клеточным старением и апоптозом. По результатам множественного регрессионного и ROC-анализа получена прогностическая модель формирования ХОБЛ, в которую вошли ген–генные комбинации и индекс курения ( $P = 4.01 \times 10^{-48}$ , AUC = 0.87).

**Ключевые слова:** хроническая обструктивная болезнь легких, длинные некодирующие РНК, клеточное старение, полигенный анализ, *H19*, *MEG3*, *LINC02227*, *CDKN2B-AS1*.

DOI: 10.31857/S0016675824090094 EDN: ADLQWZ

Хроническая обструктивная болезнь легких (ХОБЛ) – это тяжелое хроническое заболевание, характеризуется стойким и прогрессирующим ограничением воздушного потока в дыхательных путях вследствие развития эмфиземы и обструктивного бронхита и бронхиолита [1]. По данным Всемирной организации здравоохранения (ВОЗ), в настоящее время заболеваемость ХОБЛ превышает 250 млн человек, она является третьей ведущей причиной смерти в мире. В 2060 г. прогнозируемый показатель смертности от ХОБЛ будет составлять более 5.4 млн смертей ежегодно [https://www.who.int/ru/news-room/fact-sheets/detail/chronic-obstructive-pulmonary-disease-(copd)]. В

Российской Федерации наблюдается отчетливый долгосрочный тренд роста заболеваемости данной патологией: так, за 2021-й год было зафиксировано более трех миллионов случаев ХОБЛ [2]. ХОБЛ возникает в результате динамических, кумулятивных и повторяющихся в течение жизни ген–средовых взаимодействий, результатом которых является повреждение легочной ткани и изменение ее нормального функционирования [3]. Развитие ХОБЛ связано с различными факторами, такими как табакокурение, загрязнение воздуха, профессиональное воздействие, генетические и эпигенетические факторы [4]. Длинные некодирующие РНК (днРНК) являются транскриптами, которые

имеют длину более 200 нуклеотидов и не кодируют белок, функционируют как важные регуляторы различных биологических процессов, таких как альтернативный сплайсинг, деградация РНК, ингибирование миРНК, энхансинг и сайленсинг транскрипции, ремоделирование хроматина, посттрансляционная модификация структуры белков [5]. В ряде исследований показано, что в легочной ткани больных ХОБЛ дифференциально экспрессируются днРНК, многие из которых участвуют в регуляции различных аспектов клеточного старения [6]. Развитие и прогрессирование ХОБЛ могут быть связаны как с изменением экспрессии днРНК, так и с нарушением их функционирования вследствие генетического полиморфизма. Вклад полиморфизма днРНК и взаимодействия с генами, белковые продукты которых вовлечены в регуляцию клеточного старения и окислительного стресса и в молекулярный патогенез ХОБЛ, недостаточно изучен. Цель настоящего исследования – выявление ассоциации полиморфных вариантов генов регуляторных днРНК *H19*, *MEG3*, *MALAT1*, *LINC00305*, *LINC00261*, *CDKN2B-AS*, *LINC02227* с развитием ХОБЛ.

## МАТЕРИАЛЫ И МЕТОДЫ

Дизайн исследования – кандидатное исследование по принципу случай – контроль. Использовали образцы ДНК неродственных индивидов, татар по этнической принадлежности, проживающих на территории Республики Башкортостан. Исследование одобрено комитетом по этике ИБГ УНЦ РАН (протокол № 17 от 07.12.2010) и ИБГ УФИЦ РАН (протокол № 19, от 01.11.2022). От всех участников получали информированное добровольное согласие на использование биологического материала в планируемых исследованиях. Все пациенты с ХОБЛ были госпитализированы в отделение пульмонологии Городской клинической больницы № 21 г. Уфы. Диагноз ХОБЛ устанавливали с учетом рекомендаций рабочей группы по “Глобальной стратегии диагностики, лечения и профилактики хронической обструктивной болезни легких” (<http://goldcopd.org>) на основании клинических и лабораторно-инструментальных исследований, включая компьютерную томографию высокого разрешения, спирометрию. В рамках клинико-инструментального обследования у всех участников были оценены показатели внешнего дыхания (жизненная емкость легких (ЖЕЛ), форсированная жизненная емкость легких (ФЖЕЛ), объем форсированного выдоха за первую секунду (ОФВ<sub>1</sub>), соотношение ОФВ<sub>1</sub>/ФЖЕЛ), определена доля курящих пациентов с вычислением индекса курения. Группа больных включала 703 индивида (из них 627 мужчин (86.195%) и 76 женщин (10.81%)), средний возраст составил  $63.04 \pm 12.02$  лет. Среди больных

ХОБЛ курильщиков и бывших курильщиков – 590 человек (83.93%), некурящих 113 (16.08%). Индекс курения у курильщиков составил  $43.08 \pm 25.75$  пачек/лет. В группе больных показатели (в % от нормы) составляли: ОФВ<sub>1</sub> ( $41.99 \pm 19.0$ ), ФЖЕЛ ( $56.65 \pm 22.71$ ), ЖЕЛ ( $58.09 \pm 21.59$ ), ОФВ<sub>1</sub>/ЖЕЛ ( $62.29 \pm 20.98$ ).

В контрольную группу вошли неродственные индивиды, не имевшие хронических заболеваний в анамнезе, в том числе болезней органов дыхания, а также острых респираторных заболеваний на момент сбора биоматериала. Критериями включения в контрольную группу являлись нормальные показатели функции внешнего дыхания (ОФВ<sub>1</sub>/ФЖЕЛ > 70%, ОФВ<sub>1</sub> > 80%) и возраст старше 45 лет. Группа контроля включала 655 индивидов (из них 582 мужчины (88.85%) и 73 женщины (11.15%)), средний возраст составил  $60.67 \pm 11.31$ , курильщики и бывшие курильщики – 552 (84.27%) и некурящие – 103 (15.73%); индекс курения составлял  $39.75 \pm 25.87$  пачек/лет.

**Генотипирование.** ДНК выделяли из лейкоцитов периферической крови с использованием фенольно-хлороформной экстракции. Гены и SNP для анализа были выбраны в соответствии со следующими критериями: имеющие функциональную значимость и/или ранее показаны ассоциацией с другими многофакторными заболеваниями человека, частотой редких аллелей (MAF)  $\geq 5\%$  в популяциях европеоидов по данным базы the National Center for Biotechnology Information (<http://www.ncbi.nlm.nih.gov/projects/SNP/>). Для исследования были выбраны следующие полиморфные локусы генов днРНК: *H19* (rs3741219), *MEG3* (rs7158663), *MALAT1* (rs619586), *LINC00305* (rs2850711), *LINC00261* (rs6048205), *CDKN2B-AS1* (rs4977574), *LINC02227* (rs2149954). Функциональная значимость SNP исследовалась по базам RegulomeDB Version 1.1 (<https://regulomedb.org>), SNPinfo Web Server (<https://snpinfo.niehs.nih.gov>) и HaploReg v3 [7], данные представлены в табл. 1. Полиморфные варианты генов анализировали при помощи полимеразной цепной реакции (ПЦР) в реальном времени коммерческими наборами с флуоресцентной детекцией (<https://www.oligos.ru>, ООО “ДНК-Синтез”, Россия) на приборе BioRad CFX96™ (“BioRad Laboratories”, Inc, USA). Подробно методы анализа описаны нами ранее [8].

**Статистическая обработка результатов.** Описание стандартных методов статистического анализа приведено нами ранее [8]. Анализ отклонения полученных частот генотипов от равновесия Харди–Вайнберга и анализ ассоциаций отдельных SNP с заболеванием проводили с использованием пакета SNPAssoc v 2.0–2 для R [9]. Полиморфный маркер считали ассоциированным с признаком при  $P < 0.05$ , поправку на множественное тестирование проводили с помощью метода оценки доли

**Таблица 1.** Биоинформатический анализ функциональных характеристик, отобранных для исследования полиморфных локусов длинных некодирующих РНК

Ген RefSNP HGVS Names	Хромосомная позиция	Регуляторный		Промоторы гистоновых меток	Энхансеры гистоновых меток	ДНКза тканей	Регуляторные белки	Мотивы	ТФ	Expression QTLs (Haplo Reg, GTEx portal)
		ранг	коэф- фициент							
<i>H19</i> rs3741219 g.7447 T > C	11p15.5	4	0.70497	8 тканей	5 тканей	11 тканей	—	NRSF, YY1	—	19 тканей (в т. ч. легкие)
<i>MEG3</i> rs7158663 g.21819 A > G	14q32.2	1f	0.19549	—	—	Кожа	—	5	да	Кровь, артерии
<i>MALAT1</i> rs619586 g.65266169 A > G	11q13.1	1a	0.99267	23 ткани	Селезенка	48 тканей	4 вида	4	да	14 тканей
<i>LINC00305</i> rs2850711 g.61787038 A > T	18q22.1	4	0.60906	—	—	—	—	—	—	—
<i>LINC00261</i> rs6048205 g.22578963 A > G	20p11.21	4	0.60906	11 тканей	11 тканей	10 тканей	5 видов	14	да	3 ткани
<i>CDKN2B-AS1</i> rs4977574 g.22098574 A > G	9p21.3	2c	0.70567	Жировая ткань	11 тканей	4 ткани	—	Ets, GR	—	Кровь, гипофиз
<i>LINC02227</i> rs2149954 g.157820602 G > A	5q33.3	1f	0.55436	—	4 ткани	Печень, сосуды	4 вида	COMP1	—	—

Примечание. RefSNP согласно базе данных NCBI (<https://www.ncbi.nlm.nih.gov/>), функциональная значимость SNP исследовалась по базам RegulomeDB Version 1.1 (<https://regulomedb.org/>), SNPinfo Web Server (<https://snpinf0.niehs.nih.gov/>), HaploReg v3, GTEx (<https://www.gtexportal.org/>). ДНКза – чувствительный к ДНКазе регион; мотивы – измененные регуляторные мотивы для связывания с регуляторами транскрипции; ТФ – сайты связывания с транскрипционными факторами; регуляторные белки – участки связывания с регуляторными белками, Expression QTLs – expression Quantitative Trait Locus.

**Таблица 2.** Частоты аллелей и генотипов исследованных полиморфных локусов генов днРНК в группах больных ХОБЛ и здоровых индивидов

Ген RefSNP <sup>1</sup>	Редкий аллель	Генотипы, аллели	ХОБЛ n (%) (N = 703)	Контроль n (%) (N = 655)	P	OR (95%CI)
<i>H19</i> 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)
<i>MEG3</i> 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 x 10 <sup>-5</sup>	—
		A/G	932/474 (66.29/33.71)	734/576 (56.03/43.97)	1.14 x 10 <sup>-5</sup>	0.64 (0.55–0.75)
<i>MALAT1</i> 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	1 330/76 (94.59/5.41)	1 261/49 (96.26/3.74)	0.048	1.47 (1.02–2.12)
<i>LINC00305</i> 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	1 078/328 (76.67/23.33)	984/326 (75.11/24.89)	0.366	0.91 (0.77–1.09)
<i>LINC00261</i> 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	1 260/146 (89.62/10.38)	1 181/129 (90.15/9.85)	0.689	1.06 (0.82–1.36)
<i>LINC02227</i> 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)
<i>CDKN2B-AS1</i> 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)

Примечание. P – значимость различий между группами по частотам аллелей и генотипов (тест  $\chi^2$  на гомогенность выборок); OR (95%CI) – показатель отношения шансов для редкого аллеля и 95%-ный доверительный интервал (базовый аллельный тест).

ложноположительных результатов False Discovery Rate (FDR), используя онлайн-ресурс (<https://tools.carbocation.com/FDR>). Поиск ген-генных сочетаний, ассоциированных с заболеванием, проводили с использованием метода Монте-Карло и цепей Маркова с помощью программного обеспечения APSampler (<http://sourceforge.net/projects/apsampler/>) [10]. При построении предиктивных моделей использовали метод логистической регрессии с пошаговым включением переменных, в качестве которых выбирались ген-генные сочетания, полиморфные варианты отдельных генов и клинично-демографические параметры; для оценки эффективности прогностических моделей вычисляли площадь под кривой (AUC – area under the curve); расчеты проводили с помощью программы SPSS v. 22.

## РЕЗУЛЬТАТЫ

В сформированных выборках пациентов с ХОБЛ и контрольной группе проведен анализ семи полиморфных локусов генов днРНК: *H19* (rs3741219), *MEG3* (rs7158663), *MALAT1* (rs619586), *LINC00305* (rs2850711), *LINC00261* (rs6048205), *CDKN2B-AS1* (rs4977574), *LINC02227* (rs2149954). Биоинформатический анализ функциональных характеристик, отобранных для исследования полиморфных локусов днРНК, показал, что большинство SNP оказывали влияние на экспрессию гена или были сцеплены с функциональными локусами гена (табл. 1).

Прежде чем приступить к анализу ассоциации полиморфных вариантов генов-кандидатов с развитием ХОБЛ нами были рассчитаны частоты

**Таблица 3.** Статистически значимые результаты анализа ассоциации полиморфных локусов днРНК с ХОБЛ

Ген, SNP	Редкий аллель	N	Генотип, модель	OR <sub>adj</sub> (95%CI)	P <sub>adj</sub>	P <sub>cor-FDR</sub>
<i>H19</i> rs3741219 T > C	C	1358	ТТ ТС+СС, доминантная	1.00 0.74 (0.57–0.96)	0.022	0.022
			ТТ+СС СТ	1.00 0.74 (0.57–0.96)	0.021	0.022
<i>MEG3</i> rs7158663 A > G	G	1358	АА АГ+ГГ, доминантная	1.00 0.54 (0.42–0.70)	2.167 x 10 <sup>-6</sup>	1.733 x 10 <sup>-5</sup>
			лог-аддитивная	0.68 (0.57–0.81)	1.471 x 10 <sup>-5</sup>	5.884 x 10 <sup>-5</sup>
<i>LINC02227</i> rs2149954 G > A	A	1358	ГГ+ГА АА, рецессивная	1.00 0.55 (0.38–0.81)	0.00171	0.0045
			лог-аддитивная	0.81 (0.67–0.96)	0.0165	0.022
<i>CDKN2B-AS1</i> rs4977574 A > G	G	1358	АА АГ+ГГ, доминантная	1.00 0.72 (0.55–0.95)	0.018	0.022
			лог-аддитивная	0.79 (0.66–0.94)	0.0077	0.0154

Примечание. N – количество индивидов, включенных в регрессионный анализ; P<sub>adj</sub> – значимость для теста отношения правдоподобия лог-регрессионной модели с учетом пола, возраста, статуса и индекса курения; OR<sub>adj</sub> – отношение шансов с учетом всех факторов, 95%CI – 95%-ный доверительный интервал для OR; P<sub>cor-FDR</sub> – значимость теста после коррекции (Benjamini-Hochberg FDR Adjusted P-value); лог-аддитивная модель на дозу редкого аллеля – увеличение дозы редкого аллеля в ряду: гомозигота по частому аллелю (0) – гетерозигота (1) – гомозигота по редкому аллелю (2).

аллелей и генотипов в группах и соответствие распределения частот генотипов равновесию Харди–Вайнберга (табл. 2). Выявленные частоты генотипов всех исследованных полиморфных локусов в группе контроля находились в соответствии с равновесием Харди–Вайнберга: *H19* (rs3741219) ( $P_{Х-В} = 0.4715$ ), *MEG3* (rs7158663) ( $P_{Х-В} = 0.2863$ ), *MALAT1* (rs619586) ( $P_{Х-В} = 0.5317$ ), *LINC00305* (rs2850711) ( $P_{Х-В} = 0.3033$ ), *LINC00261* (rs6048205) ( $P_{Х-В} = 0.2288$ ), *CDKN2B-AS1* (rs4977574) ( $P_{Х-В} = 1$ ), *LINC02227* (rs2149954) ( $P_{Х-В} = 0.4817$ ).

Далее проведена оценка статистической значимости различий между группами по распределению частот аллелей и генотипов и рассчитаны показатели отношения шансов для редкого аллеля каждого локуса (базовый аллельный тест). На

следующем этапе методом логистической регрессии анализировали ассоциацию отдельных полиморфных локусов с учетом количественных и бинарных признаков (пол, возраст, статус и индекс курения), вводимых в уравнение регрессии в качестве независимых переменных (табл. 3). Учитывая, что при многофакторных заболеваниях вклад отдельных генов в развитие заболевания может быть небольшим, с использованием программы APSampler проводили поиск информативных ген-генных сочетаний, ассоциированных с ХОБЛ. На заключительном этапе с использованием метода множественного регрессионного анализа с пошаговым включением предикторов и последующего ROC-анализа проводили поиск комплексных клинико-генетических моделей риска развития ХОБЛ (рис. 1).

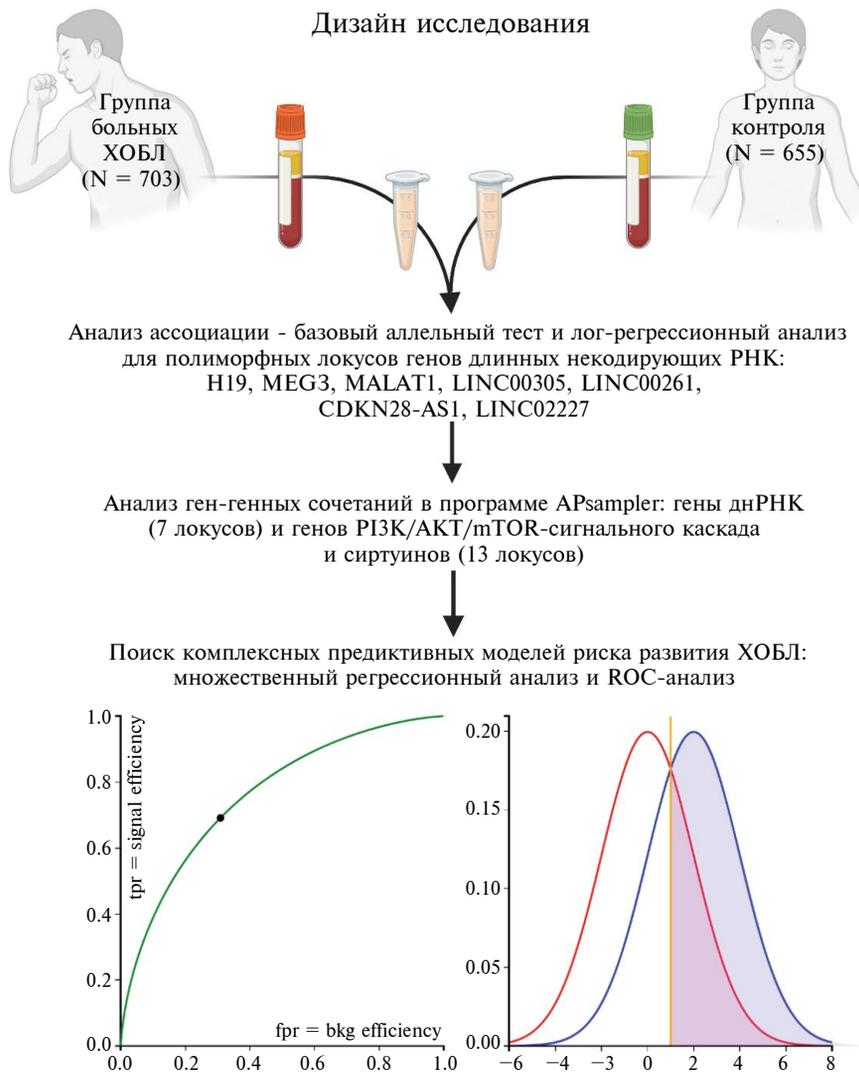


Рис. 1. Дизайн исследования.

*Анализ ассоциации отдельных полиморфных вариантов генов днРНК с развитием ХОБЛ*

Статистически значимые различия по распределению частот аллелей и/или генотипов между группами больных ХОБЛ и контроля были выявлены по генам *H19* (rs3741219), *MEG3* (rs7158663), *MALAT1* (rs619586), *LINC02227* (rs2149954), *CDKN2B-AS1* (rs4977574) (табл. 2).

Ассоциация локуса *H19* (rs3741219) с развитием ХОБЛ была установлена в доминантной модели ( $P_{\text{adj}} = 0.022$ ; OR = 0.74), маркером риска являлся гомозиготный по частому аллелю генотип ТТ ( $P_{\text{adj}} = 0.022$ ; OR = 1.36, 95%CI 1.11–1.69).

Локус *MEG3* (rs7158663) значимо ассоциировал с ХОБЛ в доминантной ( $P_{\text{adj}} = 2.167 \times 10^{-6}$ ; OR = 0.54) и лог-аддитивной ( $P_{\text{adj}} = 1.471 \times 10^{-5}$ ; OR = 0.68) моделях; риск развития заболевания связан с генотипом АА ( $P_{\text{adj}} = 2.167 \times 10^{-6}$ ; OR = 1.90, 95%CI 1.52–2.37).

Локус *LINC02227* (rs2149954) ассоциировал с ХОБЛ в рецессивной ( $P_{\text{adj}} = 0.00171$ ; OR = 0.55) и лог-аддитивной ( $P_{\text{adj}} = 0.0165$ ; OR = 0.81) моделях; маркером риска выступал частый аллель G ( $P = 0.015$ ; OR = 1.22, 95%CI 1.04–1.42).

Ассоциация гена *CDKN2B-AS1* (rs4977574) с ХОБЛ была установлена в доминантной ( $P_{\text{adj}} = 0.018$ ; OR = 0.72) и лог-аддитивной ( $P_{\text{adj}} = 0.0077$ ; OR = 0.79) моделях; маркером риска являлся гомозиготный по частому аллелю генотип АА ( $P_{\text{adj}} = 0.018$ ; OR = 1.39, 95%CI 1.09–1.76).

Ассоциация ХОБЛ с локусом *MALAT1* (rs619586) была выявлена только в базовом аллельном тесте, частота редкого аллеля G составила 5.41% в группе больных и 3.74% в контроле ( $P = 0.048$ ; OR = 1.47, 95%CI 1.02–2.12), однако регрессионный анализ не выявил значимых ассоциаций ни в одной из моделей, что связано с низкой частотой аллеля G.

Статистически значимых различий в распределении частот генотипов и аллелей полиморфных локусов генов *LINC00305* (rs2850711) и *LINC00261* (rs6048205) между больными ХОБЛ и контрольной группой выявлено не было.

Таким образом, в результате анализа отдельных полиморфных локусов генов днРНК нами впервые получены данные по ассоциации генов *H19* (rs3741219), *MEG3* (rs7158663), *LINC02227* (rs2149954), *CDKN2B-AS1* (rs4977574) с ХОБЛ.

*Анализ ген–генных сочетаний полиморфных локусов генов днРНК и генов PI3K/АКТ/ mTOR-сигнального каскада и НАД-зависимых деацетилаз семейства сиртуинов*

Выбранные для исследования днРНК выступают в качестве регуляторов различных молекулярных каскадов, связанных с процессами клеточного

старения, окислительного стресса и апоптоза, поэтому целесообразным было рассмотреть комбинированный вклад полиморфных локусов днРНК и ранее исследованных нами генов PI3K/АКТ/mTOR-сигнального каскада и НАД-зависимых деацетилаз семейства сиртуинов [11]. В анализ включали 20 SNP, из них семь – днРНК, исследованных в настоящей работе, и 13 SNP ранее изученных нами генов *PIK3R1*, *AKT1*, *MTOR*, *PTEN*, *SIRT2*, *SIRT1*, *SIRT3*, *SIRT6* [11]. Получено 8192 уникальных паттерна, из которых мы исключили комбинации, не содержащие аллели/генотипы исследованных нами днРНК. Далее критериями отбора выявленных сочетаний были  $P_{\text{FDR}} < 0.01$  и OR < 0.4 (для протективных маркеров) или OR > 2 (для маркеров повышенного риска). Всего было отобрано 33 ген–генных сочетания, которые соответствовали выбранному критерию, из них 15 ассоциировали с риском развития ХОБЛ, 18 были протективными. В табл. 4 представлены 12 наиболее значимых ген–генных комбинаций, ассоциированных с ХОБЛ.

Самые значимые комбинации риска развития ХОБЛ определялись сочетанием генотипа GG гена *SIRT3* (rs536715) с аллелем G локуса *LINC02227* (rs2149954) (OR = 2.31;  $P_{\text{FDR}} = 1.12 \times 10^{-8}$ ) и генотипом АА локуса *MEG3* (rs7158663) (OR = 2.37;  $P_{\text{FDR}} = 1.77 \times 10^{-6}$ ).

Однако большинство выявленных комбинаций риска включали аллель G гена *PIK3R1* (rs3730089) в сочетании с аллелями днРНК – *LINC00305* (rs2850711) – аллель А, *CDKN2B-AS1* (rs4977574) – аллель А и *MALAT1* (rs619586) – аллель G. Наиболее значимая комбинация, которая ассоциирована с пониженным риском развития ХОБЛ, включала аллель А гена *PIK3R1* (rs831125) в сочетании с аллелем А локуса *SIRT3* (rs3782116), генотипом AG локуса *SIRT3* (rs536715) и аллель А локуса *LINC00261* (rs6048205) (OR = 0.16;  $P_{\text{FDR}} = 4.88 \times 10^{-10}$ ). Коровым компонентом наиболее значимых протективных ген–генных паттернов являлась комбинация аллелей А локусов *SIRT3* (rs3782116) и *SIRT3* (rs536715) с аллелем А гена *LINC00261* (rs6048205) и аллелем А *MALAT1* (rs619586).

Следующие три комбинации включали аллель G гена *MEG3* (rs7158663) в сочетании с генотипом ТТ локуса *PIK3R1* (rs10515070) и аллелем А локуса *PIK3R1* (rs3730089) (OR = 0.34;  $P_{\text{FDR}} = 3.01 \times 10^{-7}$ ); с генотипом ТС локуса *SIRT6* (rs107251) (OR = 0.37;  $P_{\text{FDR}} = 3.12 \times 10^{-7}$ ) и одно из сочетаний включало аллель G гена *CDKN2B-AS1* (rs4977574), совместно с аллелями Т генов *MTOR* (rs2295080) и *SIRT6* (rs107251) (OR = 0.28;  $P_{\text{FDR}} = 6.61 \times 10^{-7}$ ).

Итак, мы видим, что полиморфные локусы *MEG3* (rs7158663), *MALAT1* (rs619586), *CDKN2B-AS1* (rs4977574) в ген–генных комбинациях проявляли аллель-специфическое действие, когда одни аллели были частью комбинаций, предрасполагающей

**Таблица 4.** Ген–генные комбинации полиморфных локусов генов днРНК, генов Р1ЗК/АКТ/мTOR–сигнального каскада и НАД–зависимых деацетилаз, наиболее значимо ассоциированные с развитием ХОБЛ

Ген–генные комбинации	ХОБЛ (частота)	Контроль (частота)	<i>P</i>	<i>P</i> <sub>FDR</sub>	OR	95%CI
<b>Протективные</b>						
<i>PIK3RI</i> (rs831125) A + <i>SIRT3</i> (rs3782116) A + <i>SIRT3</i> (rs536715) AG + <i>LINC00261</i> (rs6048205) A	4.0	21.0	2.03 x 10 <sup>-12</sup>	4.88 x 10 <sup>-10</sup>	0.16	0.087–0.287
<i>SIRT3</i> (rs3782116) A + <i>SIRT3</i> (rs536715) A + <i>LINC00261</i> (rs6048205) A + <i>MALAT1</i> (rs619586) A	6.80	25.80	2.56 x 10 <sup>-11</sup>	2.52 x 10 <sup>-9</sup>	0.23	0.14–0.37
<i>SIRT3</i> (rs3782116) A + <i>SIRT3</i> (rs536715) A + <i>MALAT1</i> (rs619586) A	8.0	24.23	1.53 x 10 <sup>-10</sup>	8.05 x 10 <sup>-9</sup>	0.26	0.16–0.40
<i>MEG3</i> (rs7158663) G + <i>PIK3RI</i> (rs10515070) TT + <i>PIK3RI</i> (rs3730089) A	9.31	23.08	2.09 x 10 <sup>-8</sup>	3.01 x 10 <sup>-7</sup>	0.34	0.23–0.51
<i>MEG3</i> (rs7158663) G + <i>SIRT6</i> (rs107251) TC	12.09	26.88	2.19 x 10 <sup>-8</sup>	3.12 x 10 <sup>-7</sup>	0.37	0.26–0.53
<i>MEG3</i> (rs7158663) G + <i>MTOR</i> (rs2295080) T + <i>SIRT6</i> (rs107251) T + <i>CDKN2B-AS1</i> (rs4977574) G	6.94	20.78	5.74 x 10 <sup>-8</sup>	6.61x10 <sup>-7</sup>	0.28	0.17–0.46
<b>Рисковые</b>						
<i>SIRT3</i> (rs536715) GG + <i>LINC02227</i> (rs2149954) G	63.00	42.41	2.61 x 10 <sup>-10</sup>	1.12 x 10 <sup>-8</sup>	2.31	1.77–3.01
<i>PIK3RI</i> (rs3730089) G + <i>PTEEN</i> (rs701848) C + <i>LINC00305</i> (rs2850711) A	52.11	33.33	1.42 x 10 <sup>-7</sup>	1.35 x 10 <sup>-6</sup>	2.178	1.62–2.92
<i>MEG3</i> (rs7158663) AA+ <i>SIRT3</i> (rs536715) GG	30.79	15.80	1.99 x 10 <sup>-7</sup>	1.77 x 10 <sup>-6</sup>	2.37	1.69–3.32
<i>PIK3RI</i> (rs3730089) G + <i>SIRT3</i> (rs536715) G + <i>SIRT6</i> (rs107251) C + <i>LINC00305</i> (rs2850711) A	63.42	45.70	2.06 x 10 <sup>-7</sup>	1.82 x 10 <sup>-6</sup>	2.06	1.56–2.72
<i>PIK3RI</i> (rs3730089) G + <i>CDKN2B-AS1</i> (rs4977574) A	20.36	9.65	4.44 x 10 <sup>-6</sup>	2.2 x 10 <sup>-5</sup>	2.39	1.62–3.52
<i>PIK3RI</i> (rs831125) A + <i>PIK3RI</i> (rs3730089) G + <i>MALAT1</i> (rs619586) G	8.06	3.38	0.0008	0.002	2.75	1.47–5.14

к развитию ХОБЛ, а альтернативные аллели этих же полиморфных локусов присутствовали в комбинациях, связанных со сниженным риском заболевания. Необходимо отметить, что для *LINC00261* (rs6048205), *LINC00305* (rs2850711) и *MALAT1* (rs619586) значимые ассоциации проявились только при сочетании с генами *PIK3R1* и сиртуинов (*SIRT3* и *SIRT6*).

Таким образом, анализ ген–генных сочетаний исследованных полиморфных локусов днРНК и генов Р1ЗК/АКТ/мTOR-сигнального каскада и сиртуинов позволил выявить высокоинформативные комбинации, ассоциированные с развитием ХОБЛ, что может указывать на синергизм исследуемых генов.

*Поиск комплексных клиничко-генетических моделей риска развития ХОБЛ с использованием множественного регрессионного анализа*

Проведен поиск комплексных моделей риска ХОБЛ методами множественного регрессионного анализа с пошаговым включением наиболее значимых переменных с последующим ROC-анализом для оценки эффективности полученных прогностических моделей. В качестве предикторов выбирали высокоинформативные ген–генные комбинации, идентифицированные на предыдущем этапе анализа, а также отдельные генотипы или аллели, выявленные при базовом анализе ассоциаций. Далее добавляли клиничко-демографические переменные (пол, возраст, статус и индекс курения) и выбирали наиболее значимые предиктивные модели.

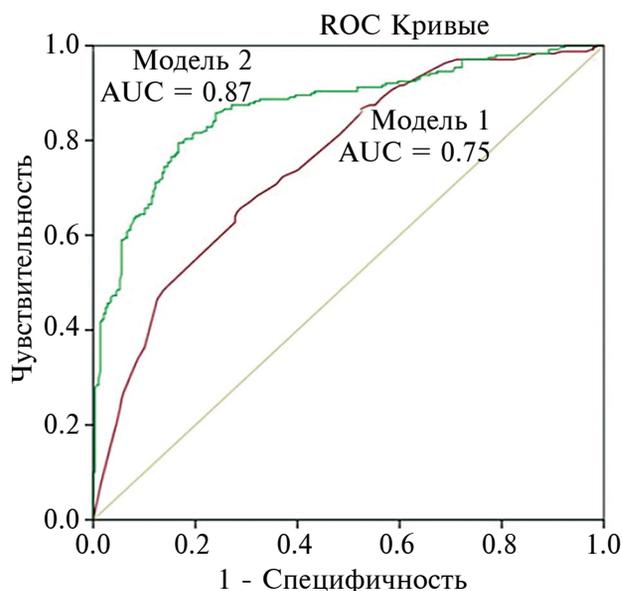
В первую высокоинформативную прогностическую модель риска формирования ХОБЛ вошли ген–генные комбинации и отдельные гены: генотип AA гена *LINC02227* (rs2149954) и аллель G гена *PIK3R1* (rs831125) ( $P = 8.86 \times 10^{-24}$ ) (табл. 5). ROC-анализ полученной модели показал ее умеренную предсказательную способность AUC = 0.75 (95%CI 0.71–0.80, чувствительность – 65.7%, специфичность – 71.2%) (рис. 2).

Вторая значимая прогностическая модель включала, помимо генетических маркеров, индекс курения и характеризовалась высокой предиктивной способностью ( $P = 4.01 \times 10^{-48}$ ) AUC = 0.87 (95%CI 0.84–0.90) (чувствительность – 74.9%, специфичность – 86.3%) (табл. 5, рис. 2), что указывает на возможность данной регрессионной модели эффективно дифференцировать больных с ХОБЛ и здоровых индивидов. В результате множественного регрессионного анализа была оценена прогностическая значимость выявленных маркеров и идентифицированы высокоинформативные комплексные модели риска ХОБЛ, учитывающие как генетические, так и средовые факторы.

## ОБСУЖДЕНИЕ

Проведен анализ ассоциации полиморфных вариантов генов днРНК *H19*, *MEG3*, *MALAT1*, *LINC00305*, *LINC00261*, *CDKN2B-AS*, *LINC02227* с ХОБЛ, изучен комбинированный вклад генов днРНК и ранее исследованных нами генов Р1ЗК/АКТ/мTOR-сигнального каскада и НАД-зависимых деацетилаз семейства сиртуинов в развитии заболевания [11].

Установлена ассоциация локуса rs3741219 гена *H19* с ХОБЛ; маркером риска является генотип ТТ, тогда как у носителей редкого аллеля С риск снижен. Замена Т в положении rs3741219 на С приводит к созданию сайтов связывания для миРНК miR-146b-3p и miR-1539 и снижению уровня экспрессии *H19* у носителей аллеля С [12]. Согласно данным функционального анализа, полиморфный локус rs3741219 расположен на участке ДНК, который связывается с регуляторными белками NRSF и YY1; изменение уровня экспрессии, по данным портала GTEx и HaploReg v3, происходит в том числе и в легочной ткани. Ген *H19* (imprinted maternally expressed transcript) расположен на хромосоме 11p15.5, в уникальном кластере генов, известном как локус H19/IGF-2, который



**Рис. 2.** График площади под кривой (ROC-анализ) для оценки эффективности прогностических регрессионных моделей. AUC – площадь под кривой. Полные характеристики моделей представлены в табл. 5. Модель 1 – AUC = 0.75 (чувствительность – 65.7%, специфичность – 71.2%) включает только генетические маркеры; модель 2 – AUC = 0.87 (чувствительность – 74.9%, специфичность – 86.3%) включает генетические маркеры и индекс курения.

Таблица 5. Значимые предиктивные регрессионные модели развития ХОБЛ

Предиктор	<i>b</i>	<i>P</i> <sub>Вальда</sub>	OR	95%CI
Модель 1				
<i>PIK3R1</i> (rs3730089) G + <i>SIRT3</i> (rs536715) G + <i>SIRT6</i> (rs107251) C + <i>LINC00305</i> (rs2850711) A	0.635	0.0014	1.887	1.28–2.79
<i>PIK3R1</i> (rs10515070) TT + <i>PIK3R1</i> (rs3730089) A	–0.645	0.0039	0.525	0.34–0.81
<i>PIK3R1</i> (rs831125) G	1.150	4x10 <sup>–8</sup>	3.158	2.09–4.76
<i>LINC02227</i> (rs2149954) AA	–0.587	0.0499	0.556	0.31–1.0
<i>MEG3</i> (rs7158663) AA + <i>SIRT3</i> (rs536715) GG	0.468	0.0471	1.597	1.01–2.53
<i>SIRT3</i> (rs3782116) A + <i>SIRT3</i> (rs536715) A + <i>LINC00261</i> (rs6048205) A	–1.239	0.0002	0.29	0.15–0.55
<i>MEG3</i> (rs7158663) G + <i>MTOR</i> (rs2295080) T + <i>SIRT6</i> (rs107251) T + <i>CDKN2B-AS1</i> (rs4977574) G	–1.298	0.0002	0.273	0.14–0.54
Константа	–0.511	0.001	0.600	
<i>P</i> = 8.86 x 10 <sup>–24</sup> AUC = 0.75 (95%CI 0.71–0.80) (чувствительность – 65.7%, специфичность – 71.2%)				
Модель 2				
<i>PIK3R1</i> (rs3730089) G + <i>SIRT3</i> (rs536715) G + <i>SIRT6</i> (rs107251) C + <i>LINC00305</i> (rs2850711) A	0.521	0.0244	1.684	1.07–2.65
<i>PIK3R1</i> (rs10515070) TT + <i>PIK3R1</i> (rs3730089) A	–0.813	0.0023	0.444	0.26–0.75
<i>PIK3R1</i> (rs831125) G	1.436	6x10 <sup>–9</sup>	4.204	2.59–6.83
<i>LINC02227</i> (rs2149954) AA	–0.814	0.0243	0.443	0.22–0.9
<i>MEG3</i> (rs7158663) AA + <i>SIRT3</i> (rs536715) GG	0.766	0.0067	2.151	1.24–3.74
<i>SIRT3</i> (rs3782116) A + <i>SIRT3</i> (rs536715) A + <i>LINC00261</i> (rs6048205) A	–0.825	0.0224	0.438	0.22–0.89
<i>MEG3</i> (rs7158663) G + <i>MTOR</i> (rs2295080) T + <i>SIRT6</i> (rs107251) T + <i>CDKN2B-AS1</i> (rs4977574) G	–1.184	0.0024	0.306	0.14–0.66
Индекс курения	0.067	4x10 <sup>–21</sup>	1.069	1.07–2.65
Константа	–2.529	7x10 <sup>–16</sup>	0.080	
<i>P</i> = 4.01 x 10 <sup>–48</sup> AUC = 0.87 (95%CI 0.84–0.90) (чувствительность – 74.9%, специфичность – 86.3%)				

Примечание. *b* – бета-коэффициент для переменной; *P*<sub>Вальда</sub> – значимость для статистики Вальда (Wald statistic); OR – exp(*b*) отношение шансов и 95%CI – 95%-ный доверительный интервал для OR; *P* – значение для теста отношения правдоподобия (likelihood ratio (LR) test); AUC – площадь под кривой.

подвергается геномному импринтингу [<https://www.ncbi.nlm.nih.gov/gene/283120/>]. *H19* играет роль ключевого компонента регуляторных сетей, вовлеченных в патогенез некоторых видов рака и фиброза, за счет стимуляции аутофагии, ингибирования апоптоза и усиления эпителиально-мезенхимального перехода (ЭМП), активации сигнальных путей TGF- $\beta$ /SMAD3 и mTOR [13, 14]. Так, показано, что *H19* ингибирует miR-200a и стимулирует PDCD4 (programmed cell death 4), усиливая апоптоз клеток сосудов легких [15]. *H19*, ингибируя miR-19b-3p, стимулирует ферроптоз клеток легких посредством ферритина 1 [16]. *H19* стимулирует фиброзные изменения дыхательных путей, ингибируя компоненты PTEN/АКТ – сигнального каскада [17]. *H19* связывает miR-29a-3p и активирует TNFRSF1A (TNF receptor superfamily member 1A), в результате чего нижележащие регуляторные молекулы стимулируют профибротический и воспалительный фенотипы легочной ткани [18]. Действуя как конкурентная эндогенная РНК (ceRNA), *H19* связывает широкий круг миРНК (miR-200a, miR-107, miR-17, miR-6515-3p, miR-138, miR-203), способствуя пролиферации, инвазии и метастазированию клеток рака легких [19]. Таким образом, повышенная экспрессия *H19* является неблагоприятным фактором развития целого спектра патологических состояний. Полиморфные локусы гена *H19* и, в частности rs3741219, широко исследовались при различных видах рака [20]. Исследования при других заболеваниях малочисленны, так показано отсутствие ассоциации локуса rs3741219 гена *H19* с развитием сахарного диабета второго типа [21].

Наиболее значимые ассоциации с ХОБЛ были получены с локусом *MEG3* (rs7158663); маркерами риска заболевания являлись частый аллель А и генотип AA, носители редкого аллеля G чаще встречались среди здоровых индивидов. Локус *MEG3* (rs7158663) был идентифицирован как наиболее частый компонент ген–генных сочетаний, ассоциированных с развитием ХОБЛ, совместно с полиморфными локусами генов *PIK3R1*, *MTOR*, *SIRT6*, *SIRT3*. Ген *MEG3* (maternally expressed 3) локализован в области 14q32.2, регулирует митохондриальный путь апоптоза [22]. *MEG3* вовлечен в регуляцию TGF- $\beta$ /SMAD3-, Wnt-сигнальных каскадов, активности PI3K/АКТ-пути [23]. Согласно данным функционального анализа, rs7158663 изменяет сайты связывания для нескольких транскрипционных факторов (PAX8, ATF6, PPAR $\gamma$ ) и мотивы для связывания с регуляторами транскрипции (ARNT2, FOXP3, BHLHE40, ELF5). Изменение уровня экспрессии, по данным портала GTeX, происходит в мононуклеарных клетках крови и эндотелии артерий. Согласно базе lncRNASNP2 database (<https://guolab.wchscu.cn/lncRNASNP/>), полиморфизм rs7158663 способен изменять структуру

сворачивания локальной РНК и влиять на взаимодействие с миРНК (miR-4307 и miR-1265) и днРНК, что, в свою очередь, влияет на уровень экспрессии соответствующей миРНК или *MEG3* [24]. В исследовании Gao et al. (2021) было показано, что уровень экспрессии *MEG3* в ткани толстого кишечника у носителей частого аллеля А значимо ниже, чем у гомозигот по редкому аллелю GG [24]. Установлена ассоциация генотипа AA и аллеля А локуса *MEG3* (rs7158663) с развитием сахарного диабета 2-го типа [21]. Показано, что *MEG3* участвует в регуляции воспалительного ответа посредством подавления miR-138, модулируя экспрессию молекул NF- $\kappa$ B-сигнального каскада и провоспалительных цитокинов [25]. *MEG3* связывает miR-181a-5p с последующей стимуляцией сигнального каскада PTEN/pSTAT5/SOCS1 в макрофагах [26], а ингибирование miR-133a-3p приводит к увеличению уровня экспрессии *SIRT1* и снижает степень легочного повреждения [27]. Одной из целевых молекул *MEG3* является miR-181b-3p, которая связана с патологическим ангиогенезом в легких и тем самым формирует основу для развития эмфиземы [28]. *MEG3* может снижать степень клеточного старения эпителиальных клеток легких через ингибирование miR-125a-5p [29]. *MEG3* выступает как ингибитор каскада miR-664a-3p/FHL1, связанного с окислительным стрессом, опосредованным воздействием сигаретного дыма [30].

Учитывая широкий спектр функций данной днРНК, вовлеченных в патогенез возраст-ассоциированных заболеваний (регуляция воспалительного каскада, окислительного стресса, апоптоза, PI3K/АКТ-пути), пониженный уровень экспрессии *MEG3*, связанный с полиморфизмом rs7158663, является значимым фактором риска развития ХОБЛ.

Нами установлена ассоциация аллеля А локуса *CDKN2B-AS* (rs4977574) с развитием ХОБЛ, данный вариант также являлся составной частью рискованной ген–генной комбинации в сочетании с полиморфным вариантом гена *PIK3R1* (rs3730089). Аллель G локуса *CDKN2B-AS* (rs4977574) чаще встречался у здоровых индивидов и входил в информативную протективную ген–генную комбинацию в сочетании с аллелями генов *MEG3* (rs7158663), *MTOR* (rs2295080) и *SIRT6* (rs107251). ДнРНК *CDKN2B-AS* (CDKN2B antisense RNA) транскрибируется с антисмысловой цепью кластера генов *CDKN2A/p16<sup>INK4A</sup>*, *CDKN2A/p14<sup>ARF</sup>* и *CDKN2B/p15<sup>INK4B</sup>*, кодирующих ингибиторы циклин-зависимой киназы 4 (CDK4) и (*MTAP*) (methylthioadenosine phosphorylase) на хромосоме 9p21.3, которые играют ключевую роль в контроле клеточной пролиферации, апоптозе, клеточном старении [<https://www.ncbi.nlm.nih.gov/gene/100048912/>]. Молекулярные механизмы действия *CDKN2B-AS1* осуществляются посредством взаимодействия с белками поликомбового репрессивного комплекса 1 и 2

(PRC1 и PRC2), такой комплекс обладает способностью к эпигенетической цис-инактивации генов-мишеней кластера *CDKN2B-CDKN2A* [31]. Исследованный нами rs4977574 влияет на уровень экспрессии *CDKN2B-AS1*; так, по данным портала GTEx и базы HaploReg v3, изменения экспрессии выявлены в крови, rs4977574 изменяет мотивы для регуляторных белков. Наличие аллеля G в локусе rs4977574 *CDKN2B-AS1* приводит к усилению образования линейных изоформ молекулы *CDKN2B-AS1*, способных к связыванию с поликомбовыми белками, наряду со снижением экспрессии кольцевых транскриптов *CDKN2B-AS1* [32]. Результаты полногеномных ассоциативных исследований (GWAS) показали значимую ассоциацию аллеля G локуса rs4977574 с развитием ишемической болезни сердца (ИБС) [33]. Необходимо отметить связь *CDKN2B-AS* с целым спектром возраст-ассоциированных заболеваний [34]. При ХОБЛ ассоциативных исследований полиморфных локусов гена *CDKN2B-AS* не проводилось, однако пониженный уровень экспрессии циркулирующей *CDKN2B-AS* в плазме крови был связан с обострениями ХОБЛ [35].

Ген *LINC02227* (long intergenic non-protein coding RNA 2227 (minus strand)) локализован на хромосоме 5q33.3, rs2149954, был впервые идентифицирован в 2014 г. в результате полногеномных исследований как ассоциированный с долгожительством [36]. Аллель А rs2149954 был связан с низким риском развития сердечно-сосудистых заболеваний и артериальной гипертензии в среднем возрасте и возрасте достижения долголетия [36]. Функциональный анализ показал, что rs2149954 локализован в области сайтов гиперчувствительности к ДНКазе I, участков связывания транскрипционных факторов и энхансеров гистоновых белков. Нами установлена ассоциация локуса *LINC02227* (rs2149954) с развитием ХОБЛ, при этом маркером риска является частый аллель G, тогда как аллель А и генотип АА маркируют устойчивость к развитию заболевания. Аллель G локуса *LINC02227* (rs2149954) был идентифицирован в информативной рискованной ген-генной комбинации с полиморфным вариантом гена *SIRT3* (rs536715), митохондриальной деацетилазой, играющей ключевую роль в регуляции процессов клеточного старения [37].

В результате полигенного анализа нами были получены информативные комбинации, ассоциированные с ХОБЛ, составными компонентами которых были полиморфные варианты генов днРНК *LINC00305* (rs2850711), *LINC00261* (rs6048205) и *MALAT1* (rs619586), не показавшие ассоциацию при базовом анализе отдельных SNP. Аллель А локуса *LINC00305* (rs2850711) был выявлен в двух наиболее значимых рискованных ген-генных комбинациях, связанных с ХОБЛ в сочетании с полиморфными

вариантами генов *PIK3R1*, *PTEN* и сиртуинов (*SIRT6* и *SIRT3*). *LINC00305* (long intergenic non-protein coding RNA 305) расположен в области 18q22.1, является регулятором NF- $\kappa$ B-сигнального каскада, усиливает экспрессию генов, кодирующих провоспалительные цитокины [38]. Аллель А локуса *LINC00261* (rs6048205) был выявлен в двух информативных ген-генных комбинациях, ассоциированных с пониженным риском развития ХОБЛ в сочетании с полиморфными локусами генов *SIRT3*, *PIK3R1* и гена *MALAT1*. Ранее было показано, что минорный аллель Т ассоциировал с развитием ревматоидного артрита, при этом у гомозигот по редкому аллелю ТТ и гетерозигот (АТ) по локусу rs2850711 уровень экспрессии *LINC00305* был значимо выше [39]. *LINC00261* (long intergenic non-protein coding RNA 261 (minus strand)) локализован на участке 20p11.21, rs6048205 был идентифицирован как ассоциированный с уровнем глюкозы при полногеномном исследовании [40]. По данным функционального анализа, rs6048205 находится в 5'-регионе гена и сцеплен с SNP в 3'-нетранслируемой области гена *FOXA2* (forkhead box A2), являющегося регулятором транскрипции и фактором, связанным с регуляцией гомеостаза глюкозы [41]. Rs6048205 изменяет сайты связывания с регуляторными белками (FOXA1, SP1, СЕВРВ, Р300, TCF4), транскрипционными факторами; согласно portalу GTEx, аллель А связан с повышенной экспрессией гена. *LINC00261* участвует в регуляции апоптоза и клеточного гомеостаза, функционируя как отрицательный регулятор Notch- и NF- $\kappa$ B-сигнального каскада, модулирует функцию SMAD3 – ключевого компонента TGF- $\beta$ 1-сигнального пути, подавляя ЭМП [42].

*MALAT1* (metastasis associated lung adenocarcinoma transcript 1) расположен в области 11q13.1, является регулятором экспрессии IL6 и TNF- $\alpha$  [43]. *MALAT1* активно экспрессируется в большинстве тканей организма человека, однако повышенная экспрессия *MALAT1* часто выявляется в различных типах злокачественных новообразований и ассоциирована с неблагоприятным прогнозом и риском метастазирования [44]. Rs619586 локализован в участке связывания с ДНКазой I, с регуляторными белками, транскрипционными факторами и гистоновыми метками. По данным portalа GTEx, изменение экспрессии подтверждено в 14 типах тканей, в том числе крови, артериях и скелетной мускулатуре, при этом аллель G приводит к увеличению уровня экспрессии *MALAT1*. Нами показано, что аллель А *MALAT1* (rs619586) входил в две протективные комбинации, ассоциированные с ХОБЛ, тогда как аллель G, связанный с повышенным уровнем экспрессии *MALAT1*, входил в ген-генное сочетание повышенного риска развития ХОБЛ. Установлено, что уровень экспрессии *MALAT1* положительно коррелирует с тяжестью

заболевания и уровнем провоспалительных цитокинов у больных ХОБЛ, взаимодействует с miR-125b, miR-146a, miR-203, приводя к усилению воспаления, функционируя как регулятор MAPK/NF- $\kappa$ B-сигналинга [45].

В результате проведенного исследования впервые в нашей работе показаны значимая ассоциация полиморфных локусов генов днРНК *H19*, *MEG3*, *CDKN2B-AS* и *LINC02227* с ХОБЛ и наличие межгенных взаимодействий генов PI3K/AKT/mTOR-сигнального каскада и НАД-зависимых деацетилаз семейства сиртуинов и исследованных днРНК (*MEG3*, *CDKN2B-AS*, *MALAT1*, *LINC00261*, *LINC00305*). Полученные нами данные указывают на то, что молекулярный патогенез ХОБЛ может включать механизмы, связанные с нарушением регуляции стрессовых реакций, препятствующих клеточному старению, при которых ключевую роль играет сеть длинных некодирующих РНК. Связанные с апоптозом, клеточным гомеостазом, окислительным стрессом и клеточным старением некодирующие РНК как потенциальные биомаркеры и мишени для терапии могут стать основой для разработки новой стратегии диагностики и лечения ХОБЛ.

Исследование выполнено за счет гранта Российского научного фонда № 23-25-00019, <https://rscf.ru/project/23-25-00019/>. с использованием оборудования ЦКП "Биомика" и УНУ "КОДИНК" (ИБГ УФИЦ РАН).

Все процедуры, выполненные в исследовании с участием людей, соответствуют этическим стандартам институционального комитета по исследовательской этике и Хельсинкской декларации 1964 года и ее последующим изменениям или сопоставимым нормам этики.

От каждого из включенных в исследование участников было получено информированное добровольное согласие.

Авторы заявляют, что у них нет конфликта интересов.

## СПИСОК ЛИТЕРАТУРЫ

1. Agustí 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. V. 61. № 4. P. 2300239. doi: 10.1183/13993003.00239-2023
2. Чучалин А.Г., Авдеев С.Н., Айсанов З.Р. и др. Хроническая обструктивная болезнь легких: федеральные клинические рекомендации по диагностике и лечению // *Пульмонология.* 2022. Т. 32. № 3. С. 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. V. 10. № 5. P. 512–524. doi: 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. V. 250. № 5. P. 624–635. doi: 10.1002/path.5364
5. Bridges M.C., Daulagala A.C., Kourtidis A. LNCcation: lncRNA localization and function // *J. Cell Biol.* 2021. V. 220. № 2. doi: 10.1083/jcb.202009045
6. Devadoss D., Long C., Langley R.J. et al. Long non-coding transcriptome in chronic obstructive pulmonary disease // *Am. J. Respir. Cell Mol. Biol.* 2019. V. 61. № 6. P. 678–688. doi: 10.1165/rcmb.2019-0184TR
7. Ward L.D., Kellis M. HaploReg v4: Systematic mining of putative causal variants, cell types, regulators and target genes for human complex traits and disease // *Nucl. Acids Res.* 2016. V. 44. № D1. P. D877–881. doi: 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. V. 692. P. 102–112. doi: 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. V. 23. № 5. P. 644–645. doi: 10.1093/bioinformatics/btm025
10. Favorov A.V., Andreevski 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. V. 171. № 4. P. 2113–2121. doi: 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 // *Vavil. Zh. Genet. Selektsii.* 2023. V. 27. № 5. P. 512–521. doi: 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. V. 95. № 9. P. 3063–3070. doi: 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. V. 41. № 3. P. 896–903. doi: 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. V. 23. № 24. P. 10796–10802.  
doi: 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. Nucl. Acids.* 2018. V. 13. P. 44–54.  
doi: 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. V. 13. P. 23–36.  
doi: 10.1016/j.bioactmat.2021.11.013
17. *Yu H., Qi N., 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. V. 14. P. 71–80.  
doi: 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. V. 247.  
doi: 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. V. 10. № 4. P. 1351–1366.  
doi: 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. V. 62. № 6. P. 565–575.  
doi: 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. V. 675. P. 265–271.  
doi: 10.1016/j.gene.2018.07.002
22. *Al-Rugeebah A., Alanazi M., Parine N.R.* MEG3: An oncogenic long non-coding RNA in different cancers // *Pathol. Oncol. Res.* 2019. V. 25. № 3. P. 859–874.  
doi: 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. V. 3. № 17.  
doi: 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. V. 11.  
doi: 10.3389/fonc.2021.796774.
25. *Li R., Fang L., Pu Q. et al.* MEG3-4 is a miRNA decoy that regulates IL-1 $\beta$  abundance to initiate and then limit inflammation to prevent sepsis during lung infection // *Sci. Signal.* 2018. V. 11. № 536.  
doi: 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. V. 78. № 6. P. 617–630.  
doi: 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. V. 88.  
doi: 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. V. 9. № 7.  
doi: 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-1 $\alpha$  // *Int. J. Biol. Sci.* 2022. V. 18. № 2. P. 661–674.  
doi: 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. V. 14. P. 2319–2329.  
doi: 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. V. 19. № 7. P. 664–670.  
doi: 10.1038/nsmb.2315
32. *Holdt L.M., Teupser D.* Long noncoding RNA *ANRIL*: Lnc-ing genetic variation at the chromosome 9p21 locus to molecular mechanisms of atherosclerosis // *Front. Cardiovasc. Med.* 2018. V. 5.  
doi: 10.3389/fcvm.2018.00145
33. *Van der Harst P., Verweij N.* Identification of 64 novel genetic loci provides an expanded view on the genetic architecture of coronary artery disease // *Circ. Res.* 2018. V. 122. № 3. P. 433–443.  
doi: 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. V. 6. № 4.  
doi: 10.1371/journal.pgen.1000899
35. *Ge J., Geng S., 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. V. 33. № 2.  
doi: 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. V. 23. № 16. P. 4420–4432.  
doi: 10.1093/hmg/ddu139
37. *Sun W., Liu C., Chen Q. et al.* SIRT3: A new regulator of cardiovascular diseases // *Oxid. Med. Cell Longev.* 2018. V. 2018.  
doi: 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- $\kappa$ B pathway in human monocytes // *Sci. Rep.* 2017. V. 7.  
doi: 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. V. 77. № 3. P. 142–147.  
doi: 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. V. 44. № 6. P. 659–669.  
doi: 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. V. 99. № 8. P. E1580-5.  
doi: 10.1210/jc.2013-4503
42. *Chen Z., Xiang L., Li L. et al.* TGF- $\beta$ 1 induced deficiency of linc00261 promotes epithelial-mesenchymal-transition and stemness of hepatocellular carcinoma via modulating SMAD3 // *J. Transl. Med.* 2022. V. 20. № 1. P. 75.  
doi: 10.1186/s12967-022-03276-z
43. *Zhu R., Liu X., 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. V. 11. № 1. P. 58.  
doi: 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. V. 100. № 12.  
doi: 10.1097/MD.00000000000023716
45. *Wu W., Wang S., Zhang L. et al.* Mechanistic studies of MALAT1 in respiratory diseases // *Front. Mol. Biosci.* 2022. V. 9.  
doi: 10.3389/fmolb.2022.1031861

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

G. F. Korytina<sup>1, 2\*</sup>, L. Z. Akhmadishina<sup>1, 3</sup>, V. A. Markelov<sup>1, 2</sup>, T. R. Nasibullin<sup>1</sup>,  
Y. G. Aznabaeva<sup>2</sup>, O. V. Kochetova<sup>1</sup>, N. N. Khusnutdinova<sup>1</sup>, A. P. Larkina<sup>1</sup>,  
N. Sh. Zagidullin<sup>2</sup>, T. V. Victorova<sup>2</sup>

<sup>1</sup>*Institute of Biochemistry and Genetics - Subdivision of the Ufa Federal Research Centre of the Russian Academy of Sciences, Ufa, 450054 Russia*

<sup>2</sup>*Bashkir State Medical University, Ufa, 450008 Russia*

<sup>3</sup>*Ufa State Petroleum Technological University, Ufa, 450064 Russia*

\*e-mail: guly\_kory@mail.ru

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 non-coding 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 allowed 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 non-coding RNAs, cellular senescence, polygenic analysis, *H19*, *MEG3*, *LINC02227*, *CDKN2B-AS1*.