

УДК 575:599.9

ПОЛИМОРФНЫЕ ВАРИАНТЫ ГЕНОВ ДЛИННЫХ НЕКОДИРУЮЩИХ РНК В РАЗВИТИИ САХАРНОГО ДИАБЕТА 2-ГО ТИПА

© 2024 г. О. В. Кочетова^{1, *}, Д. Ш. Авзалетдинова², Т. М. Кочетова²,
Т. В. Викторова², Г. Ф. Корытина¹

¹Уфимский федеральный исследовательский центр Российской академии наук,
институт биохимии и генетики, Уфа, 450054 Россия

²Башкирский государственный медицинский университет, Уфа, 450008 Россия
*e-mail: Olga_MK78@mail.ru

Поступила в редакцию 23.01.2024 г.

После доработки 26.02.2024 г.

Принята к публикации 10.04.2024 г.

Сахарный диабет 2-го типа (СД2) — это заболевание, характеризующееся повышенным содержанием глюкозы в крови, формирующееся в результате нарушения механизмов связывания инсулина с клетками. Использовали образцы ДНК больных ($N = 535$) и здоровых ($N = 475$) индивидов. Нами выявлена ассоциация с риском развития СД2-локусов генов днРНК в аддитивной модели *LINC02227 rs2149954* ($OR = 0.76$, $P = 0.0083$, $P_{FDR} = 0.017$), *LINC00305 rs2850711* ($OR = 1.43$, $P = 0.0017$, $P_{FDR} = 0.004$) и *CDKN2B-AS1 rs4977574* ($OR = 0.70$, $P = 0.0001$, $P_{FDR} = 0.0003$). Для локуса *rs2850711* гена *LINC00305* выявлена ассоциация с уровнем глюкозы натощак ($P = 0.023$), уровнем С-пептида ($P = 0.00001$), для локусов *LINC02227 rs2149954* и *CDKN2B-AS1 rs4977574* установлена ассоциация с гипертензией. Локус *MALAT1 rs619586*, аллель *A* ассоциирован с повышенным уровнем С-пептида ($P = 0.017$), с ЛПНП ($P = 0.012$) и общим холестерином ($P = 0.01$). Для локуса *CDKN2B-AS1 rs4977574* показана ассоциация с уровнем С-пептида ($P = 0.027$). Локус *LINC02227 rs2149954* ($P = 0.0011$, $P_{FDR} = 0.008$) ассоциирован с развитием ожирения. Локус *rs7158663* гена *MEG3* ассоциирован с постинфарктным кардиосклерозом ($P = 0.02$). Проведенный анализ ROC-кривых показал, что исследуемые локусы *MEG3 rs7158663*, *H19 rs3741219*, *MALAT1 rs619586*, *LINC02227 rs2149954*, *LINC00305 rs2850711*, *CDKN2B-AS1 rs4977574* и такие переменные, как пол, возраст обследуемых и уровень ИМТ, включенные в построение модели расчета риска, могут предсказывать развитие СД2 с чувствительностью 98.0% и специфичностью 97.0%, площадь под кривой AUC составила 95.30% (95%CI 93.50–97.40).

Ключевые слова: сахарный диабет 2-го типа, длинные некодирующие РНК, однонуклеотидный полиморфизм.

DOI: 10.31857/S0016675824090083 **EDN:** ADLXNV

Сахарный диабет 2 типа (СД2) является сложным заболеванием, в развитии которого принимают участие как генетические, так и средовые факторы. Выявление предполагаемых генетических вариантов, обуславливающих вариативность генов днРНК, а также определение их функциональной значимости остается слабо изученной проблемой. По результатам широкогеномных ассоциативных исследований (GWAS) были выявлены однонуклеотидные полиморфизмы (ОНП) не только в генах, кодирующих белки, большинство из установленных вариантов определено в некодирующих областях, связанных с дисфункцией поджелудочной железы и развитием диабета [1, 2].

днРНК — это один из наиболее распространенных классов длинных некодирующих РНК, транскрибирующихся у человека. днРНК принято определять как транскрипты длиной более 200

нуклеотидов, выполняющие функции РНК без последующей трансляции в белки [3]. Активное исследование днРНК позволило установить потенциальную роль этих транскриптов в развитии и прогрессировании диабета, ожирения, инсулинорезистентности, а также осложнений, вызываемых диабетом [4]. Известно, что днРНК регулируют метаболизм глюкозы, а также синтез и секрецию инсулина [5].

Полиморфные варианты генов днРНК, расположенные в сайтах связывания с мишенями микроРНК, могут влиять на взаимодействие микроРНК — днРНК и регулировать экспрессию генов, кодирующих белки [6]. Известно, что такие днРНК как *H19*, *MEG3* и *MALAT1* регулируют гомеостаз глюкозы и функцию β -клеток [7, 8]. Ингибитор циклин-зависимой киназы 2В — антисмысловая РНК 1 (*CDKN2B-AS1*), ассоциирован

с развитием диабета типа 2 [9, 10], ишемической болезнью сердца [11], атеросклерозом [12] и онкологическими заболеваниями [13]. По результатам GWAS установлено, что аллель *T* локуса rs2149954 гена *LINC02227* ассоциирован с долголетием [14], также для этого ОНП показана ассоциация с такими показателями, как триглицериды и холестерин [15]. Анализ результатов GWAS выявил ОНП rs2850711 гена *LINC00305*, ассоциированный с развитием воспаления и атеросклерозом [16]. Ассоциации полиморфных вариантов генов днРНК с развитием сахарного диабета типа 2 или сосудистыми осложнениями не вызывают сомнений, однако механизм, посредством которого осуществляется влияние, остается невыясненным. Интеграция знаний о взаимодействии днРНК с другими некодирующими и белок-кодирующими генами будет иметь решающее значение для понимания биологической организации, необходимой для оценки развития многофакторного заболевания. Проведение данного исследования было мотивировано высокой распространенностью СД2 и ожирения в России, так распространенность СД2 в России составляет 5.4%, тогда как распространенность состояния преддиабета достигает 19.3% и ожирения 24.3% [17, 18]. Учитывая данный факт, выявление

потенциальных диагностических маркеров и, в частности, новых потенциальных терапевтических мишеней для лечения диабета и ранней диагностики представляется актуальным.

Цель исследования — анализ полиморфных вариантов генов днРНК: *MEG3 rs7158663*, *H19 rs3741219*, *MALAT rs619586*, *LINC02227 rs2149954*, *LINC00305 rs2850711*, *CDKN2B-AS1 rs4977574* у пациентов с СД2, с осложнениями, вызываемыми диабетом и показателями, характеризующими СД2.

МАТЕРИАЛЫ И МЕТОДЫ

В работе были использованы образцы ДНК неродственных индивидов, татар по этнической принадлежности, проживающих на территории Республики Башкортостан. Группа больных СД2 составила 535 человек. Диагноз СД2 ставился на совокупности кодов МКБ-10 [19]. Группа пациентов была сформирована с учетом клинко-инструментального исследования на базе многопрофильного стационара Городская клиническая больница № 21 г. Уфы. Контрольная группа включала 475 практически здоровых индивидов. Характеристика групп пациентов и контроля представлена в табл. 1.

Таблица 1. Характеристика выборок, включенных в исследование

Параметры	Контроль <i>N</i> = 475	СД2 <i>N</i> = 535	<i>P</i>
Возраст, лет, среднее ± Std. dv.	49.64 ± 10.87	53.58 ± 11.87	0.05
Мужчины, <i>n</i> (%)	131 (27.6)	160 (29.9)	0.234
Женщины, <i>n</i> (%)	344 (72.4)	375 (70.1)	
Индекс массы тела (ИМТ) (кг/м ²), среднее ± SD	26.40 ± 1.99	30.94 ± 5.82	< 0.0001
Ожирение и избыточная масса тела, <i>n</i> (%)	105 (22.11)	451 (84.30)	< 0.0001
Длительность СД2, медиана [Q1; Q3]	—	6.22 [2; 17]	—
Артериальная гипертензия, <i>n</i> (%)	—	449 (84.0)	—
Сердечно сосудистые заболевания, <i>n</i> (%)	—	193 (36.1)	—
HbA _{1C} (%), медиана [Q1; Q3]	—	7.5 [7.00; 14.00]	—
Глюкоза натощак (ммоль/л), медиана [Q1; Q3]	4.80 [3.20; 5.90]	7.21 [6.1; 16.00]	< 0.0001
Общий холестерин (ммоль/л), медиана [Q1; Q3]	4.50 [3.30; 6.12]	5.43 [3.20; 10.70]	0.0002
ЛПНП (ммоль/л), медиана [Q1; Q3]	2.70 [0.79; 3.99]	3.06 [0.93; 7.83]	< 0.001
ЛПВП (ммоль/л), медиана [Q1; Q3]	1.10 [0.87; 1.40]	1.20 [0.32; 3.60]	0.05
Триглицериды (ммоль/л), медиана [Q1; Q3]	1.32 [1.10; 2.07]	1.68 [1.15; 3.71]	0.001

Примечание. ЛПНП — липопротеины низкой плотности, ЛПВП — липопротеины высокой плотности, Std. dv. — стандартное отклонение, HbA_{1C} — гликированный гемоглобин.

Исследование одобрено комитетом по этике ИБГ УФИЦ РАН. От всех участников исследования получали информированное добровольное согласие на использование биологического материала в планируемых исследованиях.

ДНК выделяли из венозной крови методом фенольно-хлороформной экстракции. Исследование полиморфных маркеров проводили с использованием полимеразной цепной реакции (ПЦР) с последующим расщеплением соответствующими эндонуклеазами рестрикции *BstF5I* (для ОНП *rs7158663* гена *MEG3*) и *BstH1I* (для *rs3741219* гена *H19*) [5]. Генотипирование локусов генов *MALAT*, *LINC02227* и *LINC00305* осуществляли с помощью аллель-специфической ПЦР. Праймеры для локуса *rs619586* гена *MALAT* – F: 5'- AAAGTCCGCCATTTTGCCAC-3', R: 5'-TGAATGCAAACTACACATGCAGA-3', A-аллеля – 5'-CTTGTTCTCTTGAGGGACACT-3', G-аллеля – 5'-AAGGTGGTAAACTATACCTG-3'; локуса *rs2850711* гена *LINC00305* F: 5'-TGAGAGGCTCTAATGGGACC-3', R: 5'-TCCCATTAATTCCCAACCCACT-3', G: 5'-AAAGATGAGCTTTGCTGTCTG-3', A: 5'-ACTCAAACCTTCCSTAAGCATTACT-3'; локуса *rs2850711* гена *LINC00305* – F: 5'-ACACATGGACAAGAAGCATCAC-3', R: 5'-CCTCTCCCATGTGTTTCCCA-3', A: 5'-AGATCTGAGCAGATCGGTGTTA-3', T: 5'-AAAATCCGGATTGTTCTACCTGA-3'.

Оценку соответствия частоты генотипов и аллелей равновесию Харди–Вайнберга проводили с помощью критерия χ^2 . Анализ ассоциаций с развитием заболевания проводили с использованием программы PLINK v.1.9 [20]. Поправку на множественность сравнения P_{FDR} оценивали с помощью онлайн калькулятора <https://www.sdmproject.com/utilities/?show=FDR>. Ассоциация считалась значимой при уровне P_{FDR} менее 0.05, а 95% CI не пересекал единицу. Анализ ассоциаций, рассчитанных для основной группы, представлен в моделях: кодоминантной и аддитивной, а также в виде аллельного теста. Количественные параметры оценивали с применением критерия Манна–Уитни. Множественный логистический регрессионный анализ и ROQ-анализ проводились с использованием GraphPad Prism 8.4.3.

РЕЗУЛЬТАТЫ

Распределение генотипов и аллелей ОНП генов днРНК *MEG3 rs7158663*, *H19 rs3741219*, *MALAT rs619586*, *LINC02227 rs2149954*, *LINC00305 rs2850711*, *CDKN2B-AS1 rs4977574* в контрольной группе соответствовало равновесию Харди–Вайнберга (табл. 2). Первоначально мы провели анализ распределения генотипов и аллелей полиморфных локусов генов днРНК у пациентов с СД2 и в

контрольной группе. Статистически значимые различия при сравнении групп пациентов и контроля были получены для локусов *LINC02227 rs2149954*, *LINC00305 rs2850711*, *CDKN2B-AS1 rs4977574* в кодоминантной и аддитивной моделях, а также в аллельном тесте.

Ассоциация с развитием СД2 для локуса *rs2149954* гена *LINC02227* в аддитивной модели: OR = 0.76 (CI95% 0.62–0.93), $P = 0.0083$, $P_{FDR} = 0.017$. Рисковыми являются генотипы CC–CT OR = 2.14 (CI95% 1.34–3.42), $P = 0.001$ ($P_{FDR} = 0.004$) (табл. 2). Ассоциация с развитием СД2 показана для локуса *rs2850711* гена *LINC00305* в аддитивной модели OR = 1.43 (CI95% 1.14–1.80), $P = 0.0017$, $P_{FDR} = 0.004$. Рисковым является аллель T, в аллельном тесте показатель OR составил 1.52 (CI95% 1.20–1.93), $P = 0.0001$, $P_{FDR} = 0.0003$. Ассоциация показана для локуса *rs4977574* гена *CDKN2B-AS1* OR = 0.43 (CI95% 0.30–0.61), $P = 0.0001$. Анализ распределения генотипов и аллелей локуса *rs4977574* гена *CDKN2B-AS1* выявил снижение частоты генотипа AA среди пациентов, в данном случае генотипами риска являются варианты GG–AG, показатель OR для вариантов риска составил 1.89 (CI95% 1.36–2.59), $P = 0.0001$, $P_{FDR} = 0.0003$.

Далее мы провели анализ в зависимости от наличия ожирения. Анализ проводился в группе пациентов. Данные представлены в табл. 3. Ассоциация с ожирением была показана для локусов *rs7158663* гена *MEG3* в аддитивной модели: аллель риска G ($P = 0.04$, $P_{FDR} = 0.07$), а также *rs619586* гена *MALAT* в кодоминантной и аддитивной моделях: аллель риска G ($P = 0.029$, $P_{FDR} = 0.07$ и $P = 0.018$, $P_{FDR} = 0.05$ соответственно). Также ожирение ассоциировано с *rs2149954* гена *LINC02227*, генотипами риска являются CC–CT варианты ($P = 0.0011$, $P_{FDR} = 0.008$), а для генотипа TT и аллеля T установлено протективное значение в кодоминантной и аддитивной моделях ($P = 0.0011$, $P_{FDR} = 0.008$ и $P = 0.0014$, $P_{FDR} = 0.008$ соответственно). Для локуса *rs2850711* и *LINC00305* ассоциация определена в аддитивной модели: аллель риска T ($P = 0.03$, $P_{FDR} = 0.07$). Однако не все полученные различия достигали статистической значимости при поправке на множественность.

Для выявления ассоциации с тяжестью, осложнениями, вызванными СД2, нами проведен анализ данных локусов в зависимости от количественных характеристик СД2 и сопутствующей заболеваемости. С повышенным уровнем С-пептида ассоциированы: генотип AG локуса *rs619586* гена *MALAT* ($P = 0.017$), генотип TT локуса *rs2850711* гена *LINC00305* ($P = 0.00001$), генотип AA локуса *rs4977574* гена *CDKN2B-AS1* ($P = 0.027$) (табл. 4). Повышенный уровень общего холестерина и ЛПНП был установлен для носителей генотипа AA локуса *rs619586* гена *MALAT* ($P = 0.01$ и $P = 0.012$). Повышенный уровень глюкозы, определяемой

Таблица 2. Распределение генотипов и аллелей генов днРНК в группе пациентов СД2 и в контроле

Ген, поли-морфизм	Генотип, аллель	Контроль, (%)	СД2, (%)	OR (95% CI)	<i>P</i>	<i>P</i> _{H-W}	<i>P</i> _{FDR}
MEG3 rs7158663	Кодоминантная AA AG GG	177 (37.3) 210 (44.2) 88 (18.5)	208 (38.9) 239 (44.7) 88 (16.4)	1.00 0.97 (0.74–1.27) 0.85 (0.60–1.22)	0.67	0.071	0.79
	Аллельный тест A G	564 (59.0) 386 (41.0)	655 (61.0) 415 (39.0)	1.08 (0.91–1.29)	0.42		0.61
	Аддитивный	—	—	0.94 (0.78–1.12)	0.47		0.63
H19 rs3741219	Кодоминантная TT CT CC	176 (37) 211 (44.4) 88 (18.5)	199 (37.2) 228 (42.6) 108 (20.2)	1.00 0.96 (0.73–1.26) 1.09 (0.77–1.54)	0.76	0.087	0.80
	Аллельный тест T C	563 (59.0) 387 (41.0)	626 (59.0) 444 (41.0)	0.97 (0.81–1.16)	0.76		0.80
	Аддитивный	—	—	1.04 (0.88–1.24)	0.63		0.79
MALAT1 rs619586	Кодоминантная AA AG GG	429 (90.3) 44 (9.3) 2 (0.4)	461 (86.8) 69 (13) 1 (0.2)	1.00 1.46 (0.98–2.18) 0.47 (0.04–5.15)	0.14	0.34	0.24
	Аллельный тест A G	902 (95.0) 48 (5.0)	991 (93.0) 71 (7.0)	0.74 (0.51–1.08)	0.14		0.24
	Аддитивный	—	—	1.25 (0.85–1.85)	0.25		0.39
LINCO02227 rs2149954	Кодоминантная CC CT TT	172 (36.2) 219 (46.1) 84 (17.7)	234 (44.1) 265 (50) 31 (5.8)	1.00 0.89 (0.68–1.16) 0.27 (0.17–0.43)	0.0001	0.34	0.0003
	Аллельный тест C T	563 (59.0) 387 (41.0)	733 (69.0) 327 (31.0)	0.65 (0.54–0.78)	0.00001		0.00006
	Аддитивный	—	—	0.76 (0.62–0.93)	0.0083		0.017

Таблица 2. Окончание

Ген, полиморфизм	Генотип, аллель	Контроль, (%)	СД2, (%)	OR (95% CI)	<i>P</i>	<i>P</i> _{H-W}	<i>P</i> _{FDR}
<i>LINC00305</i> rs2850711	Кодоминантная <i>AA</i> <i>AT</i> <i>TT</i>	355 (74.7) 106 (22.3) 14 (3)	362 (67.7) 132 (24.7) 41 (7.7)	1.00 1.22 (0.91–1.64) 2.87 (1.54–5.36)	0.0014	0.088	0.003
	Аллельный тест <i>A</i> <i>T</i>	816 (86.0) 134 (14.0)	856 (80.0) 214 (20.0)	1.52 (1.20–1.93)	0.0001		0.0003
	Аддитивный	—	—	1.43 (1.14–1.80)	0.0017		0.004
<i>CDKN2B-AS1</i> rs4977574	Кодоминантная <i>GG</i> <i>AG</i> <i>AA</i>	124 (26.1) 225 (47.4) 126 (26.5)	188 (35.1) 265 (49.5) 82 (15.3)	1.00 0.78 (0.58–1.04) 0.43 (0.30–0.61)	0.0001	0.27	0.0003
	Аллельный тест <i>G</i> <i>A</i>	473 (50.0) 477 (50.0)	641 (60.0) 429 (40.0)	0.66 (0.55–0.79)	0.00001		0.00006
	Аддитивный	—	—	0.70 (0.58–0.84)	0.0001		0.0003

Примечание. *P*_{H-W} – уровень значимости при определении равновесия Харди–Вайнберга для контрольной группы, *P*_{FDR} – уровень значимости с учетом множественных сравнений. Жирным отмечены статистически значимые различия.

натошак, выявлен среди носителей генотипа *AA* локуса *rs2850711* гена *LINC00305* (*P* = 0.023). Осложнение постинфарктный кардиосклероз чаще встречался у пациентов с генотипами *GG* локуса *rs7158663* гена *MEG3*: OR = 4.91 (CI95% 1.34 – 17.94), (*P* = 0.02) и у пациентов с генотипами *TT* локуса *rs2850711* гена *LINC00305* (*P* = 0.005). Гипертония чаще встречалась у пациентов с генотипами *GG–AG* гена *CDKN2B-AS1* *rs4977574* (*P* = 0.03) и среди носителей вариантов *CT–TT* локуса *rs2149954* гена *LINC02227* (*P* = 0.05).

При анализе ROC-кривых для оценки прогностической значимости выявленных рисков значений в развитии СД2 были построены две модели, для первой модели были учтены только полиморфные локусы: *MEG3* *rs7158663*, *H19* *rs3741219*, *MALAT* *rs619586*, *LINC02227* *rs2149954*, *LINC00305* *rs2850711*, *CDKN2B-AS1* *rs4977574*, во вторую модель были включены эти же локусы, а также такие переменные, как пол, возраст обследуемых и ИМТ.

Эффективность прогноза измеряли с использованием площади под кривой (AUC). ROC-анализ показал AUC, равную 66.5% (95%CI 63.90–69.40) для модели, включающей только исследуемые полиморфные варианты. Для второй модели AUC составила 95.30% (95%CI 93.50–97.40) с чувствительностью 98.0% и специфичностью 97.0%, что указывает на высокую способность показателей, включенных в анализ, правильно классифицировать индивидов с СД2 и здоровых.

ОБСУЖДЕНИЕ

В последние годы большое внимание уделяется исследованию роли некодирующих частей генома человека в развитии заболеваний, особенно это касается онкологических заболеваний [21]. В настоящее время идет активное исследование днРНК с риском развития метаболических расстройств, таких как ожирение, инсулинорезистентность,

Таблица 3. Ассоциация исследованных локусов генов днРНК с ожирением

Ген, полиморфизм	Модель	Генотип	Контроль	Ожирение	OR (95% CI)	<i>P</i>	<i>P</i> _{FDR}
<i>MEG3</i> rs7158663	Кодоминантная	<i>AA</i> <i>AG</i> <i>GG</i>	36.3 44.7 19.9	42.6 44.8 13.8	1.00 1.21 (0.86–1.71) 1.68 (1.03–2.72)	0.18	0.27
	Аддитивная	—	—	—	1.36 (1.01–1.82)	0.04	0.12
<i>H19</i> rs3741219	Кодоминантная	<i>TT</i> <i>CT</i> <i>CC</i>	38.8 42.6 18.7	36.5 46.7 16.8	1.00 0.96 (0.61–1.5) 1.16 (0.82–1.66)	0.58	0.63
	Аддитивная	—	—	—	1.01 (0.81–1.26)	0.94	0.94
<i>MALAT</i> rs619586	Кодоминантная	<i>AA</i> <i>AG</i> <i>GG</i>	89.3 10.3 0.4	86.9 12.8 0.2	1.00 1.81 (1.06–3.10) NA	0.029	0.07
	Аддитивная	—	—	—	2.33 (1.11– 4.90)	0.018	0.05
<i>LINC02227</i> rs2149954	Кодоминантная	<i>CC</i> <i>CT</i> <i>TT</i>	36.4 49.3 14.3	44.4 49.5 6.1	1.00 0.82 (0.59–1.16) 0.35 (0.20–0.60)	0.0011	0.008
	Аддитивная	—	—	—	0.66 (0.51–0.85)	0.0014	0.008
<i>LINC00305</i> rs2850711	Кодоминантная	<i>AA</i> <i>AT</i> <i>TT</i>	78.5 19.6 1.9	72.3 22.7 5	1.00 1.26 (0.85–1.87) 2.84 (0.98–8.23)	0.062	0.124
	Аддитивная	—	—	—	1.41 (1.02–1.94)	0.03	0.07
<i>CDKN2B-AS1</i> rs4977574	Кодоминантная	<i>GG</i> <i>AG</i> <i>AA</i>	27.8 49.8 22.5	32.6 47.4 20	1.00 0.81 (0.56–1.18) 0.76 (0.48–1.19)	0.41	0.49
	Аддитивная	—	—	—	0.87 (0.69–1.08)	0.21	0.28

Примечание. См. табл. 2.

СД2 [22]. В данном исследовании мы впервые выявили ассоциацию полиморфных локусов генов *LINC02227* rs2149954 и *LINC00305* rs2850711 с развитием СД2, ожирения, а также с осложнениями, вызванными СД2. *LINC02227* rs2149954 – данный полиморфный вариант ассоциирован с долголетием по результатам GWAS [23], также данный локус сцеплен с ОНП rs9313772, ассоциированным с гипертонией и ишемической болезнью сердца (ИБС). *LINC02227* – это длинная межгенная некодирующая РНК 2227. Согласно данным базы <https://regulomedb.org/>, полиморфизму rs2149954 соответствует ранг 1f, указывающий на функциональную значимость данного ОНП (eQTL/caQTL + TF binding / chromatin accessibility peak). Для аллеля *T* установлено повышение уровня экспрессии в жировой ткани, селезенке, крови по данным GTEx проекта (<https://www.gtexportal.org/>). Также показано, что данный участок является значимым при

взаимодействии с такими факторами транскрипции как FOXA1, HNF4A. Известно, что данные транскрипционные факторы вовлечены в регуляцию экспрессии гена *GLUT2*, переносчика глюкозы [24].
LINC00305 является днРНК, относится к классу межгенных, кодирует РНК 305, а также участвует в активации моноцитов, индуцирует продукцию провоспалительных цитокинов, участвует в регуляции апоптоза. Данный ген входит в генную сеть репрессора арилгидрокарбонового рецептора (AHR), участника NF-κB-сигналинга. Полиморфный вариант rs2850711 передается сцепленным в одном кластере, состоящем в общей сложности из 16 SNP, один из сцепленных вариантов (rs2850687) является связывающим регионом для USF2 (Upstream Transcription Factor 2). Продукт гена *USF2* участвует в метаболизме липидов, усиливает активность фактора транскрипции SREBP1c [25].

Таблица 4. Анализ полиморфных вариантов генов днРНК с клиническими характеристиками пациентов

Маркер	Полиморфный локус	Медиана (Q1; Q3)	<i>P</i>
С-пептид	<i>MALAT rs619586</i> <i>AG</i> <i>AA</i>	4.49 (6.32; 2.92) 2.19 (2.28; 1.9)	0.017
	<i>LINC00305 rs2850711</i> <i>AA-AT</i> <i>TT</i>	2.28 (2.11; 2.34) 10.3 (4.87; 15.28)	0.00001
	<i>CDKN-AS rs4977574</i> <i>GG-AG</i> <i>AA</i>	2.27 (2.37; 2.47) 4.57 (2.27; 6.91)	0.027
Общий холестерин	<i>MALAT rs619586</i> <i>AA</i> <i>AG</i>	5.54 (4.92; 5.63) 5.02 (4.88; 5.17)	0.01
ЛПНП	<i>MALAT rs619586</i> <i>AA</i> <i>AG</i>	3.2 (2.98; 3.34) 2.57 (2.22; 2.89)	0.012
Глюкоза натощак	<i>LINC00305 rs2850711</i> <i>AA</i> <i>AT-TT</i>	7.38 (7.11; 7.68) 6.64 (6.45; 6.89)	0.023

Локус *rs2850711* гена *LINC00305* ассоциирован с развитием атеросклероза [26]. Уровень экспрессии гена *LINC00305* выше у носителей аллеля *T* для генотипов *AT* и *TT* по сравнению с носителями аллеля *A* в гомозиготном состоянии [27]. Повышенная экспрессия гена *LINC00305* была выявлена у пациентов с атеросклерозом и ревматоидным артритом, так же как и ассоциация аллеля *T* с риском развития этих заболеваний [27, 28]. Ранее было установлено, что повышенная экспрессия *LINC00305* модулирует (повышает) экспрессию генов воспаления [29].

Для локуса *rs4977574* гена днРНК *CDKN2B-AS1* (*ANRIL*) показана ассоциация как с сердечно-сосудистыми осложнениями при СД2, так и с риском развития СД2 [30]. Авторы в мета-анализе показали, что аллель *G* локуса *rs4977574* обуславливает увеличение уровней ЛПНП у европейцев и, соответственно, предрасположенность к ИБС [31], так же как и среди популяций Азии носители аллеля *G* имеют повышенный риск развития ишемической болезни сердца [32]. Вопрос остается дискуссионным по поводу функциональной значимости полиморфного локуса *rs4977574* гена *CDKN2B-AS1*. Рядом авторов выдвинуто предположение, что аллель *G* приводит к образованию линейных изоформ молекулы *ANRIL* наряду со снижением экспрессии кольцевых транскриптов *ANRIL*. Линейные изоформы *ANRIL* вызывают репрессию генов опухолевых супрессоров — *CDKN2A* и *CDKN2B* [33]. В конечном итоге это приводит к ингибированию апоптоза и чрезмерной пролиферации проатерогенных

клеток. Вместо этого *A*-аллель может способствовать усиленному образованию кольцевых изоформ *ANRIL*. Транскрипты гена *ANRIL* такого типа ингибируют активность комплекса *ReBoW*, необходимого для созревания рРНК. Это, в свою очередь, приводит к дефициту рРНК, ядрышковому стрессу и активации белка *p53*, что приводит к ингибированию деления клеток и активации апоптоза. С другой стороны, показано, что аллель *G* локуса *rs4977574* снижает риск развития рака простаты и ИБС среди иранцев [34], программа *HaploReg v4* предсказывает, что данный локус изменяет сайты связывания с транскрипционными факторами *ETS* и глюкокортикоидного рецептора (*GR*) [35].

Показана роль глюкокортикоидного рецептора в развитии гипергликемии и диабета [36]. Так, повышенная экспрессия глюкокортикоидного рецептора была выявлена у мышей с диабетом, тогда как отсутствие глюкокортикоидного рецептора в гепатоцитах мышей останавливает развитие гипергликемии, это позволяет предположить, что изменение связывания глюкокортикоидного рецептора с участком *9p21* является значимым в развитии диабетической гипергликемии, что было подтверждено настоящим исследованием, тогда как развитие ожирения не ассоциировано с данным локусом. Ассоциация с диабетом для генов *CDKN2A/2B* показана и по результатам проведенных широкогоеномных исследований [37]. Механизмы этого влияния вероятно заключаются в изменении способности к пролиферации бета-клеток поджелудочной железы под действием продуктов

этих генов с возрастом. днРНК *CDKN2B-AS1* способствует усилению пролиферации клеток и снижению апоптоза, а также является участником развития воспалительного ответа [38].

В нашем исследовании локус *rs619589* гена *MALAT1* ассоциирован с уровнем С-пептида, холестерина, ЛПНП и ожирением. *MALAT1* — длинная некодирующая РНК — играет роль в развитии различных видов рака, метаболических и других заболеваний. По данным некоторых авторов, аллель *G* обуславливает снижение экспрессии *MALAT1* в клетках сердца, крови и опухолей [39, 40]. Программа HaploReg v4 предсказывает, что данный локус изменяет сайты связывания для транскрипционных факторов HNF1 и PLZF, принимающих участие в контроле основных биологических процессов, включая энергетический обмен, выработку глюкозы в печени и развитие гипергликемии [41, 42]. Так, аллель *A* связывается с Erythroid-Related Factor (ERF), тогда как аллель *G* не связывается. В настоящем исследовании носители аллеля *G* имели сниженный уровень холестерина, сниженный уровень ЛПНП и сниженный уровень С-пептида, что подтверждает данные Wang G. с соавт. о протективном эффекте аллеля *G* при развитии ишемической болезни сердца [43].

Нами выявлена ассоциация локусов генов днРНК в аддитивной модели — *LINC02227 rs2149954* ($P_{FDR} = 0.017$), *LINC00305 rs2850711* ($P_{FDR} = 0.004$) и *CDKN2B-AS1 rs4977574* ($P_{FDR} = 0.0003$), с риском развития СД2. Для локуса *rs2850711* гена *LINC00305* выявлена ассоциация с уровнем глюкозы натощак ($P = 0.023$), уровнем С-пептида ($P = 0.00001$); для локусов *LINC02227 rs2149954* и *CDKN2B-AS1 rs4977574* установлена ассоциация с гипертонией. Локус *MALAT1 rs619586* ассоциирован с уровнем С-пептида ($P = 0.017$), ЛПНП ($P = 0.012$) и общим холестерином ($P = 0.01$). Для локуса *CDKN2B-AS1 rs4977574* показана ассоциация с уровнем С-пептида ($P = 0.027$). Локус *LINC02227 rs2149954* ($P_{FDR} = 0.008$) ассоциирован с ожирением, локус *rs7158663* гена *MEG3* — с постинфарктным кардиосклерозом ($P = 0.02$). Локус *rs3741219* гена *H19* не продемонстрировал ассоциации ни с показателями инсулинорезистентности, ни с ожирением. Проведена оценка ROC-кривых в развитии СД2. Для модели, включающей шесть локусов генов днРНК, был получен показатель AUC, равный 66.5% (95%CI 63.90–69.40), для второй модели, в которую были включены эти же локусы, а также такие переменные, как пол, возраст обследуемых и ИМТ, AUC составил 95.3% (95%CI 93.5–97.4) с чувствительностью 98.0% и специфичностью 97.0%, что указывает на высокую способность показателей, включенных в анализ, правильно классифицировать индивидов с СД2 и здоровых. Таким образом показано, что на развитие СД2 оказывают влияние как вариабельность генов

днРНК, так и такие «средовые факторы», как пол, возраст, а также наличие ожирения и избыточной массы тела. Для подтверждения выводов необходимо проведение исследования в других группах.

Исследование поддержано Российским научным фондом (№22-25-00010).

Исследование одобрено Этическим комитетом ИБГ УФИЦ РАН, протокол № 8 от 14.03.2012.

Все процедуры, выполненные в исследовании с участием людей, соответствуют этическим стандартам институционального и/или национального комитета по исследовательской этике и Хельсинкской декларации 1964 г. и ее последующим изменениям или сопоставимым нормам этики. От каждого из включенных в исследование участников было получено информированное добровольное согласие. Все обследованные — совершеннолетние.

Авторы декларируют отсутствие явных и потенциальных конфликтов интересов, связанных с публикацией настоящей статьи.

СПИСОК ЛИТЕРАТУРЫ

1. Татосян К.А., Зиневич Л.С., Демин Д.Э., Шварц А.М. Функциональные особенности длинных некодирующих РНК, содержащих последовательности мобильных генетических элементов // Мол. биология. 2020. Т. 54. № 5. С. 718–724. DOI: 10.31857/S0026898420050109
2. Ji E., Kim C., Kim W., Lee E.K. Role of long non-coding RNAs in metabolic control // Biochimica et Biophysica Acta (BBA) — Gene Regulatory Mechanisms. 2020. V. 1863. № 4. P. 194348. <https://doi.org/10.1016/j.bbagrm.2018.12.006>
3. Mahajan A., Taliun D., Thurner M. et al. Fine-mapping type 2 diabetes loci to single-variant resolution using high-density imputation and islet-specific epigenome maps // Nature Genetics. 2018. V. 50. № 11. P. 1505–1513. doi: 10.1038/s41588-018-0241-6
4. Pulit S.L., Stoneman C., Morris A.P. et al. Meta-analysis of genome-wide association studies for body fat distribution in 694 649 individuals of European ancestry // Hum. Mol. Genet. 2019. V. 28. № 1. P. 166–174. doi: 10.1093/hmg/ddy327
5. 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
6. Bak R.O., Mikkelsen J.G. miRNA sponges: Soaking up miRNAs for regulation of gene expression // Wiley Interdisciplinary Reviews: RNA. 2014. V. 5. № 3. P. 317–333. doi: 10.1002/wrna.1213
7. Ding G.L., Wang F.F., Shu J. et al. Transgenerational glucose intolerance with *Igf2/H19* epigenetic alterations in mouse islet induced by intrauterine hyperglycemia

- // Diabetes. 2012. V. 61. № 5. P. 1133–1142.
doi: 10.2337/db11-1314
8. Ding H., Wang F., Shi X. *et al.* LncRNA *MALAT1* induces the dysfunction of β cells via reducing the histone acetylation of the PDX-1 promoter in type 1 diabetes // Experimental and Mol. Pathology. 2020. V. 114. doi: 10.1016/j.yexmp.2020.104432
 9. Zeggini E., Weedon M.N., Lindgren C.M. *et al.* Replication of genome-wide association signals in UK samples reveals risk loci for type 2 diabetes // Science. 2007. V. 316. № 5829. P. 1336–1341. doi: 10.1126/science.1142364
 10. Xiao M., Bai S., Chen J. *et al.* CDKN2B-AS1 participates in high glucose-induced apoptosis and fibrosis via NOTCH2 through functioning as a miR-98-5p decoy in human podocytes and renal tubular cells // Diabetology & Metabolic Syndrome. 2021. V. 13. № 1. P. 1–12. doi: 10.1186/s13098-021-00725-5
 11. Huang K., Zhong J., Li Q. *et al.* Effects of CDKN2B-AS1 polymorphisms on the susceptibility to coronary heart disease // Mol. Genet. & Genomic Med. 2019. V. 7. № 11. doi.org/10.1002/mgg3.955
 12. Li H., Han S., Sun Q. *et al.* Long non-coding RNA CDKN2B-AS1 reduces inflammatory response and promotes cholesterol efflux in atherosclerosis by inhibiting ADAM10 expression // Aging (Albany NY). 2019. V. 11. № 6. P. 1695. doi: 10.18632/aging.101863
 13. Foss-Skiftesvik J., Li S., Rosenbaum A. *et al.* Multi-ancestry genome-wide association study of 4069 children with glioma identifies 9p21.3 risk locus // Neuro-Oncology. 2023. V. 25. № 9. P. 1709–1720. doi: 10.1093/neuonc/noad042
 14. 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
 15. Martin S., Cule M., Basty N. *et al.* Genetic evidence for different adiposity phenotypes and their opposing influences on ectopic fat and risk of cardiometabolic disease // Diabetes. 2021. V. 70. № 8. P. 1843–1856. doi: 10.2337/db21-0129
 16. Zhang D.D., Wang W.T., Xiong J. *et al.* Long noncoding RNA *LINC00305* promotes inflammation by activating the AHRN-NF- κ B pathway in human monocytes // Sci. Reports. 2017. V. 7. № 1. doi: 10.1038/srep46204
 17. Dedov I., Shestakova M., Benedetti M.M. *et al.* Prevalence of type 2 diabetes mellitus (T2DM) in the adult Russian population (NATION study) // Diabetes Res. and Clin. Practice. 2016. V. 115. P. 90–95. doi: 10.1016/j.diabres.2016.02.010
 18. Алфёрова В.И., Мустафина С.В. Распространенность ожирения во взрослой популяции Российской Федерации // Ожирение и метаболизм. 2022. Т. 19. № 1. С. 96–105. <https://doi.org/10.14341/omet12809>
 19. Дедов И.И., Шестакова М.В., Майоров А.Ю. и др. Алгоритмы специализированной медицинской помощи больным сахарным диабетом // Сахарный диабет. 2022. Т. 24. № 1S. С. 1–148. DOI: 10.14341/DM20171S8
 20. Purcell S., Neale B., Todd-Brown K. *et al.* PLINK: A tool set for whole-genome association and population-based linkage analyses // Am. J. Hum. Genet. 2007. V. 81. № 3. P. 559–575. DOI: <https://doi.org/10.1086/519795>
 21. Cao L., Yan G., Yu S. *et al.* Associations of *MALAT1* and its functional single nucleotide polymorphisms with cancer // Pathology-Res. and Practice. 2022. V. 236. doi: 10.1016/j.prp.2022.153988
 22. Yang W., Lyu Y., Xiang R., Yang J. The role of genetic variants in the long non-coding RNA genes *MALAT1* and *H19* in the pathogenesis of childhood obesity, long non-coding RNAs in the pathogenesis of insulin resistance // Int. J. Mol. Sci. 2022. V. 23. № 24. doi: 10.3390/ijms232416054
 23. Wain L.V., Verwoert G.C., O'Reilly P.F. *et al.* Genome-wide association study identifies six new loci influencing pulse pressure and mean arterial pressure // Nat. Genetics. 2011. V. 43. № 10. P. 1005–1011. doi: 10.1038/ng.922
 24. David-Silva A., Freitas H.S., Okamoto M.M. *et al.* Hepatocyte nuclear factors 1 α /4 α and forkhead box A2 regulate the solute carrier 2A2 (Slc2a2) gene expression in the liver and kidney of diabetic rats // Life Sci. 2013. V. 93. № 22. P. 805–813. doi: 10.1016/j.lfs.2013.10.011
 25. Czech M.P., Tencerova M., Pedersen D.J., Aouadi M. Insulin signalling mechanisms for triacylglycerol storage // Diabetologia. 2013. V. 56. P. 949–964. doi: 10.1007/s00125-013-2869-1
 26. Castellanos-Rubio A., Ghosh S. Disease-associated SNPs in inflammation-related lncRNAs // Front. Immunol. 2019. V. 10. DOI: 10.3389/fimmu.2019.00420
 27. 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 // Brit. J. Biomed. Sci. 2020. V. 77. № 3. P. 142–147. doi: 10.1080/09674845.2020.1744942
 28. O'Donnell C.J., Cupples L.A., D'Agostino R.B. *et al.* Genome-wide association study for subclinical atherosclerosis in major arterial territories in the NHLBI's Framingham Heart Study // BMC Med. Genet. 2007. V. 8. P. 1–12. doi: 10.1186/1471-2350-8-S1-S4

29. Zhang D-D., Wang W-T., Xiong J. et al. Long noncoding RNA *LINC00305* promotes inflammation by activating the AHRR-NF- κ B pathway in human monocytes // Sci. Reports. 2017. V. 7. № 1. <https://doi.org/10.1038/srep46204>
30. Rivera N.V., Carreras-Torres R., Roncarati R. et al. Assessment of the 9p21.3 locus in severity of coronary artery disease in the presence and absence of type 2 diabetes // BMC Med. Genet. 2013. V. 14. P. 1–8. doi: 10.1186/1471-2350-14-11
31. Wei B., Liu Y., Li H. et al. Effect of 9p21.3 (lncRNA and CDKN2A/2B) variant on lipid profile // Front. in Cardiovascular Med. 2022. V. 9. doi: 10.3389/fcvm.2022.946289
32. Altshuler D., Anand S.S., Andersen K. et al. Large-scale association analysis identifies 13 new susceptibility loci for coronary artery disease // Nat. Genet. 2011. V. 43. № 4. P. 333–338. doi: 10.1038/ng.784
33. Xu B., Fang Z., He S. et al. ANRIL polymorphism rs4977574 is associated with increased risk of coronary artery disease in Asian populations: A meta-analysis of 12,005 subjects // Medicine. 2018. V. 97. № 39. doi: 10.1097/MD.00000000000012641
34. Taheri M., Pouresmaeili F., Omrani M.D. et al. Association of *ANRIL* gene polymorphisms with prostate cancer and benign prostatic hyperplasia in an Iranian population // Biomarkers in Medicine. 2017. V. 11. № 5. P. 413–422. <https://doi.org/10.2217/bmm-2016-0378>
35. 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. Ac. Res. 2016. V. 44. № D1. P. D877–D881. doi.org/10.1093/nar/gkv1340
36. Kokkinopoulou I., Diakoumi A., Moutsatsou P. Glucocorticoid receptor signaling in diabetes // Int. J. Mol. Sci. 2021. V. 22. № 20. doi: 10.3390/ijms222011173
37. Kong Y., Sharma R.B., Nwosu B.U., Alonso L.C. Islet biology, the *CDKN2A/B* locus and type 2 diabetes risk // Diabetologia. 2016. V. 59. № 8. P. 1579–1593. doi: 10.1007/s00125-016-3967-7
38. Holdt L.M., Teupser D. Long Noncoding RNA *ANRIL*: Lnc-ing genetic variation at the chromosome 9p21 locus to molecular mechanisms of atherosclerosis // Front. in Cardiovascular Med. 2018. V. 5. doi: 10.3389/fcvm.2018.00145
39. Li Q., Zhu W., Zhang B. et al. The *MALAT1* gene polymorphism and its relationship with the onset of congenital heart disease in Chinese // Biosci. Rep. 2018. V. 38. № 3. doi: 10.1042/BSR20171381
40. Zheng J., Pang C.H., Du W. et al. An allele of rs619586 polymorphism in *MALAT1* alters the invasiveness of meningioma via modulating the expression of collagen type V alpha (COL5A1) // J. Cel. Mol. Med. 2020. V. 24. № 17. P. 10223–10232. doi.org/10.1111/jcmm.15637
41. Chen S., Qian J., Shi X. et al. Control of hepatic gluconeogenesis by the promyelocytic leukemia zinc finger protein // Mol. Endocrin. 2014. V. 28. № 12. P. 1987–1998. doi: 10.1210/me.2014-1164
42. Kavitha B., Ranganathan S., Gopi S. et al. Molecular characterization and re-interpretation of *HNF1A* variants identified in Indian MODY subjects towards precision medicine // Front. in Endocrin. 2023. V. 14. doi: 10.3389/fendo.2023.1177268
43. Wang G., Li Y., Peng Y. et al. Association of polymorphisms in *MALAT1* with risk of coronary atherosclerotic heart disease in a Chinese population // Lipids in Health and Disease. 2018. V. 17. P. 1–7. <https://doi.org/10.1186/s12944-018-0728-2>

Polymorphic Variants of Long Non-Coding RNA Genes in the Development of Type 2 Diabetes Mellitus

O. V. Kochetova^{1, *}, D. Sh. Avzaleydinova², T. M. Kochetova², T. V. Viktorova², G. F. Korytina¹

¹Ufa Scientific Center of the Russian Academy of Sciences, Institute of Biochemistry and Genetics, Ufa, 450054 Russia

²Bashkir State Medical University, Ufa, 450008 Russia

*e-mail: Olga_MK78@mail.ru

Type 2 diabetes mellitus (T2DM) is a disease characterized by increased blood glucose, formed as a result of impaired mechanisms of insulin binding to cells. DNA samples of patients with T2DM ($N = 535$) and healthy individuals ($N = 475$) were used. We identified the association of lncRNA gene loci *LINC02227 rs2149954* (OR = 0.76, $P = 0.0083$, $P_{\text{FDR}} = 0.017$), *LINC00305 rs2850711* (OR = 1.43, $P = 0.0017$, $P_{\text{FDR}} = 0.004$) and *CDKN2B-AS1 rs4977574* (OR = 0.70, $P = 0.0001$, $P_{\text{FDR}} = 0.0003$) in an additive model with T2DM. The *LINC00305* gene locus *rs2850711* showed association with fasting glucose level ($P = 0.023$), C-peptide level ($P = 0.00001$), *LINC02227 rs2149954* and *CDKN2B-AS1 rs4977574* loci showed association with hypertension. *MALAT rs619586* gene polymorphism was associated with C-peptide ($P = 0.017$), LDL ($P = 0.012$) and total cholesterol levels ($P = 0.01$). *CDKN2B-AS1 rs4977574* gene polymorphism showed association with C-peptide levels ($P = 0.027$). *SNP LINC02227 rs2149954* was associated with obesity ($P = 0.0011$, $P_{\text{FDR}} = 0.008$). *MEG3 rs7158663* gene polymorphism was associated with post infarct atherosclerosis ($P = 0.02$). ROC curve analysis showed that the studied loci and variables such as sex, age of subjects and BMI level, included in the construction of the risk calculation model, could predict the development of T2DM with a sensitivity of 98.0% and specificity of 97.0%, the area under the AUC curve was 95.30% (95%CI 93.50–97.40).

Keywords: type 2 diabetes mellitus, long non-coding RNAs, single nucleotide polymorphisms.