

**КАНДИДАТНЫЕ SNP-МАРКЕРЫ АТЕРОСКЛЕРОЗА,
КОТОРЫЕ СПОСОБНЫ ДОСТОВЕРНО ИЗМЕНЯТЬ СРОДСТВО
ТАТА-СВЯЗЫВАЮЩЕГО БЕЛКА К ПРОМОТОРАМ ГЕНОВ ЧЕЛОВЕКА**

© 2019 г. М. П. Пономаренко^{1,2, *}, Д. А. Рассказов¹, И. В. Чадаева^{1,2}, Е. Б. Шарыпова¹,
И. А. Драчкова¹, П. М. Пономаренко³, Е. А. Ощепкова¹, Л. К. Савинкова¹, Н. А. Колчанов^{1,2}

¹Федеральный исследовательский центр, Институт цитологии и генетики
Сибирского отделения Российской академии наук, Новосибирск, 630090 Россия

²Новосибирский национальный исследовательский государственный университет, Новосибирск, 630090 Россия

³Университет Ла Верна, Ла Верн 91750, Калифорния, США

*e-mail: pon@bionet.nsc.ru

Поступила в редакцию 30.01.2019 г.

После доработки 18.02.2019 г.

Принята к публикации 09.04.2019 г.

Атеросклероз и связанные с ним такие патологии, как ишемическая болезнь сердца, инфаркт миокарда, стенокардия и инсульт, являются основными причинами смертности в мире. На развитие атеросклероза могут оказывать влияние многие факторы, в частности избыточная масса тела, гипертония, сахарный диабет, гиперлиппротеинемия и другие заболевания в анамнезе, а также генетическая предрасположенность, которая в ряде случаев может быть обусловлена наличием однонуклеотидного полиморфизма (Single Nucleotide Polymorphism, SNP). В этой работе изучались районы промоторов белок-кодирующих генов человека, в которых известны SNP-маркеры изменения сродства ТАТА-связывающего белка (ТВР) к этим промоторам, ассоциированные с сопутствующими атеросклерозу патологиями. В результате мы нашли в базе данных dbSNP неаннотированные SNP, которые могут изменить сродство ТВР/промотор так же, как в случае клинических SNP-маркеров (по прогнозу нашего Web-сервиса SNP_TATA_Z-tester, http://www.mgs.bionet.nsc.ru/cgi-bin/mgs/tatascan_fox/start.pl). Так, SNP-маркер rs35036378 риска опухоли рГ1 снижает сродство ТВР к промотору гена *ESR2* (эстрогеновый рецептор β) и, соответственно, уровень его экспрессии, что является физиологическим маркером кальциноза кровеносных сосудов при атерогенезе. Вблизи этого известного SNP-маркера есть неаннотированный SNP rs766797386, который тоже способен снизить как сродство ТВР к тому же промотору, так и уровень экспрессии гена *ESR2*. Поэтому мы предлагаем rs766797386 как кандидатный SNP-маркер ускорения атерогенеза на стадии кальциноза кровеносных сосудов. Обсуждается возможность использования рекомендованного диетологами рациона с высоким содержанием кальция (Ca) для замедления кальциноза (а также обратный эффект от употребления Ca-обогащенных пищевых добавок) людьми, у которых обнаружены маркеры ускорения кальциноза. Аналогичным образом было предсказано и обсуждено 33 кандидатных SNP-маркера для ускорения или замедления атерогенеза. Полученные данные после клинических проверок могли бы помочь в корректировке рациона и образа жизни в соответствии с данными индивидуальных расшифрованных геномов.

Ключевые слова: атеросклероз, промотор, однонуклеотидный полиморфизм (SNP), ТАТА-связывающий белок (ТВР), сайт ТВР-связывания (ТАТА-бокс).

DOI: 10.1134/S0016675819090121

Атеросклероз – воспалительное заболевание артерий, лежащее в основе развития ряда сердечно-сосудистых патологий, среди которых ишемическая болезнь сердца и инсульт, являющиеся одними из основных причин смертности [1]. Несмотря на то, что атеросклероз считается болезнью стареющих людей, его проявления обнаруживают и у детей, а предвестники атеросклеротических бляшек были обнаружены даже на эмбриональных стадиях развития [2]. Теку-

щее состояние организма индивида в процессе атерогенеза может отражать его генетические особенности, условия и образ жизни, включая перенесенные заболевания, способствующие атеросклерозу. В этой связи можно предполагать, что, корректируя условия и образ своей жизни, становится возможным контролировать течение атерогенеза с помощью информации, полученной в результате расшифровки индивидуального генома.

Согласно общепринятой концепции, развитие атеросклероза начинается с накопления липопротеинов низкой плотности на внутренней стенке сосуда. Затем к образующемуся скоплению присоединяются моноциты, которые поглощают липопротеины и дифференцируются, превращаясь в пенистые клетки (макрофаги). Макрофаги отделяются, унося излишек липопротеинов, либо со временем гибнут, внося вклад в образующееся скопление. Так, шаг за шагом возникают сначала тромбогенные липидные пятна и полосы, затем — их фиброзные утолщения, далее — атеросклеротические бляшки с местными воспалениями и кальцинозом кровеносных сосудов [3, 4]. Согласно клиническим наблюдениям, атерогенез — не монотонный, а импульсный процесс, который развивается преимущественно постпрандиально, т.е. после принятия пищи [5], при острых инфекционных воспалениях [6], при повреждении эндотелия высоким артериальным давлением [4]. Наконец, все это приводит к тромбозу артерии и отмиранию питаемых ею тканей [7]. Таким образом, атерогенез — длительный многоступенчатый процесс, зависящий от таких условий и образа жизни индивида, как пиковые физические нагрузки, инфекционные воспаления, попадание в кровь посторонних веществ при травмах, сбалансированности и регулярности питания, а также от наследственных предрасположенностей индивида, закодированных в геноме.

В рамках крупнейшего мирового научного проекта “1000 геномов” уже было расшифровано 10545 индивидуальных геномов людей [8], итогом сравнения которых стали 100877027 выявленных SNP, документированных в базе данных dbSNP [9] в качестве различий между индивидуальными и референсным геномами человека. Путем сравнения когорт пациентов с заданным заболеванием и условно-здоровых добровольцев исследователи находят достоверные различия между геномами этих групп, выявляя кандидатные SNP-маркеры заболеваний, которые в дальнейшем могут быть использованы в персонализированной медицине [10, 11]. Полученные данные накапливают в базе данных OMIM [12], объем которой достиг 24232 документов. Хотя в белок-кодирующих районах генов встречаются нейтральные SNP (например, синонимические), а в регуляторных — фатальные [12], большинство биомедицинских SNP-маркеров были найдены в белок-кодирующих районах генов, где они повреждают структуры и функции белков, что не поддается никакой коррекции; ничтожно мало — в регуляторных районах генов [13], где SNP влияют лишь на уровень продуктов генов в организме человека, что поддается коррекции как сменой образа жизни, так и медикаментозно. Многие регуляторные клинические SNP-маркеры были идентифицированы вблизи от сайта связывания ТАТА-связывающего белка (ТВР) в про-

моторах генов человека, где они изменяют сродство ТВР/промотор на стадии смещения равновесия от плотной нуклеосомной упаковки хроматина к прединициаторным комплексам транскрипции генов [14]. Это самое первое молекулярное событие инициации транскрипции генов абсолютно необходимо: нокаутные модельные животные *ТВР^{-/-}* останавливаются в своем развитии на стадии бластулы, когда исчерпывают запас материнских ТВР [15]. Такие клинические SNP-маркеры изменяют уровни экспрессии генов пропорционально изменению сродства ТВР к промоторам генов [16].

Ранее нами был создан Web-сервис SNP_TATA_Comparator, оценивающий значимость α различия между сродством ТВР к норме и минорному варианту заданного SNP промотора гена человека [17]. С его помощью нами были предсказаны кандидатные SNP-маркеры для ожирения, агрессивности, хронопатологий, болезни Альцгеймера, аутоиммунных нарушений и репродуктивного потенциала [18–23]. В этой работе предлагается обновление SNP_TATA_Z-tester нашего Web-сервиса [17] и его применение для прогноза кандидатных SNP-маркеров для атеросклероза в качестве продолжения цикла таких прогнозов.

МАТЕРИАЛЫ И МЕТОДЫ

В этой работе изучали проксимальные промоторы, 90 пн, белок-кодирующих генов человека с известными SNP-маркерами для заболеваний, сопутствующих атеросклерозу и ассоциированным с изменениями сродства ТВР к промоторам этих генов. Последовательности ДНК в норме брали из референсного генома человека [8], а минорные получали вручную по описанию в dbSNP [9] и в литературе (например, [24]). В табл. 1 даны обозначения генов в референсном геноме человека [8], идентификаторы SNP в базе данных dbSNP [9] или обозначения SNP в базе данных OMIM [12], последовательности ДНК вокруг SNP, заболевания, сопутствующие атеросклерозу, и литературные ссылки.

Для каждой SNP независимо сравнивали последовательности ДНК анцестрального $S_1 = \{s_{1,-90} \dots s_{1,i} \dots s_{1,-1}\}$ и минорного $S_2 = \{s_{2,-90} \dots s_{2,i} \dots s_{2,-1}\}$ аллелей промотора длиной 90 пн с этим SNP перед стартом транскрипции, TSS (Transcription Start Site, $TSS \equiv s_{1,0} \equiv s_{2,0}$; где $s_{\cdot,i} \in \{a, c, g, t\}$) — с помощью алгоритма, подробно представленного в нашей предыдущей статье [69]. При этом сначала независимо оценивали уровень сродства ТВР к анцестральному и минорному аллелям промотора, “ $-\ln(K_D(S_1))$ ” и “ $-\ln(K_D(S_2))$ ”, с использованием следующей формулы:

Таблица 1. Известный и кандидатные SNP-маркеры для атеросклероза (AS), предсказанных в этой работе как способные изменять средство ТВР к моторам генов человека

Ген	dbSNP [9]	5'-фланк $\frac{wt}{mut}$ 3'-фланк $\frac{wt}{mut}$	K _D , нМ				Факт, прогноз (эта работа)	Литературная ссылка
			$\frac{wt}{mut}$	Δ	Z	α		
<i>CETP</i>	-54 Δ 18bp [24]	18 пн cgTgGGGct — gggctccag	$\frac{4}{7}$	↓	7	10 ⁻⁶	Гипералфалипопротеинемия ← AS	[24]
<i>ESR2</i>	rs35036378	cctctggct $\frac{t}{g}$ ttaaagaaga	$\frac{6}{8}$	↓	5	10 ⁻³	Опухоль рГ1	[25]
	rs766797386	ttaaagaaga $\frac{g}{t}$ aagggcctta	$\frac{6}{7}$	↓	3	10 ⁻²	(прогноз) → AS	[26, 27]
<i>IL1B</i>	rs1143627	ttttgaaagc $\frac{c}{t}$ ataaaacag	$\frac{5}{2}$	↑	15	10 ⁻⁶	Рак печени, легких и желудка, гастрит, ожирение (прогноз) → AS	Для обзора см. [17] [28, 27]
<i>INS</i>	rs5505	agatcactgt $\frac{c}{t}$ cttctgccat	$\frac{53}{44}$	↑	4	10 ⁻³	Неонатальный сахарный диабет	[29]
	rs563207167	tcagccctgc $\frac{c}{t}$ tgtctcccag	$\frac{53}{44}$	↑	4	10 ⁻³	(прогноз) → AS	[30, 31]
<i>SOD1</i>	rs7277748	ggctcggcct $\frac{a}{g}$ taaagtagtc	$\frac{2}{7}$	↓	17	10 ⁻⁶	ALS (прогноз) ишемическое осложнение AS передней спинальной артерии	[32], [33]
<i>TPII</i>	rs1800202	gcgctctata $\frac{t}{g}$ aagTgggcag	$\frac{1}{4}$	↓	17	10 ⁻⁶	Гемолитическая анемия и нейромускулярные заболевания как	[35, 36]
	rs781835924	cggcgctc $\frac{t}{c}$ atataagtg	$\frac{1}{2}$	↓	10	10 ⁻⁶	(прогноз) травматическое и ишемическое осложнения AS	[37, 38]
<i>F3</i>	rs563763767	ccctttatag $\frac{c}{t}$ ggcggggca	$\frac{3}{2}$	↑	6	10 ⁻⁶	Инфаркт миокарда, тромбоз (прогноз) → AS	[39], [40]
<i>F7</i>	-33a→c [41]	cccttgaggc $\frac{a}{c}$ gagaactttg	$\frac{53}{62}$	↓	3	0.01	Умеренные кровотечения	[41]
	rs749691733	agaactttgc $\frac{c}{t}$ cgctagctcc	$\frac{53}{66}$	↓	4	10 ⁻³	(прогноз) → AS	[42]

Таблица 1. Продолжение

Ген	dbSNP [9]	5'-фланк $\frac{wt}{mut}$ 3'-фланк $\frac{wt}{mut}$	K_D , нМ				Факт, прогноз (эта работа)	Литературная ссылка	
			$\frac{wt}{mut}$	Δ	Z	α			
<i>HBB</i> и <i>HBD</i> [§]	rs397509430	gggctgggca $\frac{t}{-}$ atacaacagt	$\frac{5}{29}$	↓	34	10^{-6}	Талассемия и устойчивость к малярии [43]		
	rs33980857	gggctgggca $\frac{t}{a,g,c}$ atacaacagt	$\frac{5}{21}$	↓	27	10^{-6}			
	rs34598529	ggctgggcat $\frac{c}{g}$ aaagtcaggg	$\frac{5}{18}$	↓	24	10^{-6}			
	rs33931746	gctgggcata $\frac{a}{g,c}$ aagtcagggc	$\frac{5}{11}$	↓	14	10^{-6}			
	rs33981098	agggctgggc $\frac{a}{g,c}$ taaaagtcag	$\frac{5}{9}$	↓	10	10^{-6}			
	rs34500389	caggctgggg $\frac{c}{a,t,g}$ ataaaagtca	$\frac{5}{6}$	↓	3	0.01			
	rs35518301 [§]	caggaccagc $\frac{a}{g}$ taaaaggcag	$\frac{4}{8}$	↓	11	10^{-6}			
	rs63750953	ctgggcataa $\frac{aa}{-}$ gtcagggcag	$\frac{5}{8}$	↓	9	10^{-6}			
	rs281864525	tgggcataaa $\frac{a}{c}$ gtcagggcag	$\frac{5}{7}$	↓	7	10^{-6}			(прогноз) ← AS [44]
	rs34166473 [§]	aggaccagca $\frac{t}{c}$ aaaaggcagg	$\frac{4}{8}$	↓	18	10^{-6}			
<i>MMP12</i>	rs2276109	gatataact $\frac{a}{g}$ ttagtcactc	$\frac{11}{14}$	↓	3	0.01	Снижение рисков псориаза, астмы и системной склеродермы [45–47]		
	rs572527200	gatgatatca $\frac{a}{g}$ ctatgatca	$\frac{11}{14}$	↓	3	10^{-2}	(прогноз) → AS [48]		

Таблица 1. Окончание

Ген	dbSNP [9]	5'-фланк $\frac{wt}{mut}$ 3'-фланк $\frac{wt}{mut}$	K_D , нМ				Факт, прогноз (эта работа)	Литературная ссылка
			$\frac{wt}{mut}$	Δ	Z	α		
<i>MBL2</i>	rs72661131	tctatttcta $\frac{t}{c}$ atagcctgca	↓	12	10^{-6}	Вариабельный иммунодефицит, преэклампсия, инсульт	[49–51]	
	rs562962093	atctatttct $\frac{a}{g}$ tatagccctgc	↓	15	10^{-6}	(прогноз) → AS	[52, 53]	
<i>NOS2</i>	–51t→c [54]	gtataaatac $\frac{t}{c}$ tcttgctgc	↑	3	0.01	Устойчивость к малярии и эпилепсии (прогноз) → AS	[54–56], [57]	
<i>APOA1</i>	35a→c [58]	ttggctcta $\frac{a}{c}$ atagccctg	↓	5	10^{-6}	Стеатоз печени (прогноз) → AS	[58], [59, 60]	
<i>ACKR1</i>	rs2814778	ttggctcta $\frac{t}{c}$ ctggaagca	↓	4	10^{-3}	Лейкопения и устойчивость к малярии (прогноз) ← AS	[61, 62] [63]	
<i>LEP</i>	rs201381696	tgatcggccc $\frac{a}{g}$ taagaggggc	↓	17	10^{-6}	Ожирение, (прогноз) → AS (прогноз) ← AS при дефиците APOE	[64] [65] [66]	
	rs200487063	tgatcggccc $\frac{g}{a}$ ctataagagg	↑	6	10^{-6}	(прогноз) ← AS, терапия рекомбинантным лептином против AS при диабете I типа	[67]	
<i>GCG</i>	rs183433761	gctggagagt $\frac{a}{g}$ tataaaagca	↓	17	10^{-6}	(прогноз) → AS посттрандиально через рост отношения INS/GCG	[68]	
	rs757035851	tataaaag $\frac{cag}{-}$ tgcgccttgg	↓	3	10^{-3}			

Примечание. wt – норма (референсный геном человека); mut – минорный аллель; –54Δ18bp, 18 пн – делеция “5’-g-71ggggcacaataac_{54-3’”} длиной 18 пн в промоторе гена *SETP*; символ “–” делеция; K_D , константа диссоциации комплекса ТВР/ДНК [34]; Δ – изменение: рост (↑), спад (↓); Z, Z-статистика; α – значимость (α = 1–p, где p, вероятность на рис. 1). Патологии: ALS – боковой амиотрофический склероз; AS – атеросклероз; Атерогенез: ускоренный (→ AS), замедленный (← AS). Гены: *ACKR1* – гликопротеин D; *APOA1* – аполипопротеин А1; *SETP* – белок переноса липидов плазмы; *ESR2* – эстрогеновый рецептор β; *F3* – тромбопластин; *F7* – проконвертин; *GCG* – глюкоган; *HBB* и *HBD* – β- и δ-цепи гемоглобина соответственно; *IL1B* – интерлейкин 1β; *INS* – инсулин; *LEP* – лептин; *MBL2* – манноза-6-фосфат-лигаза; *MMP12* – металлоластаза макрофагов; *NOS2* – индуцибельная синтаза оксида азота; *SOD1* – Cu/Zn-супероксиддисмутаза; *TPI1* – триозофосфатизомераза.

$$-\ln(K_D(S)) = 10.9 - 0.2 \times \{ \ln(K_{SLIDE}(S)) + \ln(K_{STOP}(S)) + \ln(K_{BEND}(S)) \}, \quad (1)$$

здесь 10.9 (натуральных логарифмических единиц) – оценка неспецифического сродства ТВР/ДНК, 10 мМ согласно экспериментальными измерениям [70]; 0.2 – стехиометрический коэффициент для трех шагов связывания ТВР/ДНК, оцененный в работе [71]; K_{STOP} – оценка константы равновесия остановки скольжения ТВР вдоль ДНК на наилучшем потенциальном сайте ТВР-связывания согласно позиционно-весовой матрице $\{w_{j;s}\}$, широко используемой для канонической формы этого сайта, ТАТА-бокса [72]:

$$\ln(K_{STOP}) = \text{MAX}_{(+),(-)\text{DNA цепь}} \left\{ \sum_{j=-1}^{13} w_{j;s_{i+j}} \right\}. \quad (2)$$

В формуле (1) K_{SLIDE} – оценка константы равновесия скольжения ТВР вдоль ДНК вокруг лучшего потенциального сайта ТВР-связывания, который упомянут выше (т.е. ТАТА-боксы ± 5 пн) согласно работе [73]:

$$-\ln(K_{SLIDE}) = \text{MEAN}_{[\text{ТАТА-боксы} \pm 5 \text{ bp}]} \{0.8[\text{ТА}] + 3.4\mu + 35.1\}, \quad (3)$$

здесь $[\text{ТА}]$ – содержание динуклеотида $[\text{ТА}]$ вокруг ТАТА-боксов; μ – ширина малой бороздки в центре ТАТА-боксов в ангстремах [74]; 0.8, 3.4 и 35.1 – регрессионные коэффициенты, оцененные в статье [73].

Наконец, в формуле (1) K_{BEND} – оценка константы равновесия стабилизации комплекса ТВР/промотор в силу изгиба оси спирали ДНК в центре ТАТА-боксов:

$$-\ln(K_{BEND}) = \text{MEAN}_{\text{ТАТА-боксы}} \times \{0.9[\text{ТА}, \text{АА}, \text{ТГ}, \text{АГ}] + 2.5[\text{ТА}, \text{ТС}, \text{ТГ}] + 14.4\}, \quad (4)$$

здесь 0.9, 2.5 и 14.4 – коэффициенты регрессии из статьи [73].

Затем независимо оценивали стандартные ошибки оценок сродства ТВР к анцестральному и минорному аллелям промотора, $\delta(S_1)$ и $\delta(S_2)$, в рамках теории нейтральной эволюции [75] с использованием всех возможных замен нуклеотидов “s \rightarrow ξ ” в каждой позиции промотора вокруг лучшего потенциального сайта ТВР-связывания, как это было предложено в работе [76]:

$$\delta(S) = \left[\left(\sum_{1 \leq i \leq 26} \sum_{\xi \in \{a,c,g,t\}} \left[\ln(K_D(\{s_{i-13} \dots \xi \dots s_{i+12}\}) / K_D(\{s_{i-13} \dots s_{i+j} \dots s_{i+12}\})) \right]^2 \right) / 72 \right]^{1/2}. \quad (5)$$

Результатом всех этих расчетов были оценки сродства ТВР к анцестральному и минорному аллелям промотора, охарактеризованные оценками их стандартных ошибок, “ $-\ln(K_D(S_1)) \pm \delta(S_1)$ ” и “ $-\ln(K_D(S_2)) \pm \delta(S_2)$ ”, для сравнения которых вычисляли Z-статистику Фишера следующим образом:

$$Z = \text{abs} \left[\ln(K_D(S_1) / K_D(S_2)) \right] / \left[\delta(S_1)^2 + \delta(S_2)^2 \right]^{1/2}. \quad (6)$$

Наконец, с помощью стандартного пакета “R” [77] по величине Z-статистики оценивали вероятность p гипотезы “ $H_0: \ln(K_D(S_1)) \neq \ln(K_D(S_2))$ ” для принятия при статистической значимости α прогностического решения ($\alpha = 1 - p$):

ЕСЛИ {неравенство “ $-\ln(K_D(S_1)) < -\ln(K_D(S_2))$ ” статистически значимо ($\alpha < 0.05$)},

ТО {прогноз “аллель S_2 увеличит экспрессию гена в сравнении с нормой S_1 ”};

ИНАЧЕ [ЕСЛИ {неравенство “ $-\ln(K_D(S_1)) > -\ln(K_D(S_2))$ ” значимо ($\alpha < 0.05$)},

ТО {прогноз “аллель S_2 уменьшит экспрессию гена в сравнении с нормой S_1 ”};

ИНАЧЕ {прогноз “аллель S_2 незначим в сравнении с нормой S_1 экспрессии гена”}.

Прогнозы этого правила принятия решения были нами ранее многократно верифицированы экспериментально с использованием независимых тест-систем *in vivo* и *ex vivo* (в качестве обзора см. [78]).

В настоящей работе указанный алгоритм применяли, как это показано на рис. 1, с помощью модификации SNP_TATA_Z-tester (http://www-mgs.bionet.nsc.ru/cgi-bin/mgs/tatascan_fox/start.pl), ранее созданного и описанного нами Web-сервиса SNP_TATA_Comparator (<http://www-mgs.bionet.nsc.ru/cgi-bin/mgs/tatascan/start.pl>) [17], который взаимодействует с Web-сервисом UCSC геном браузер [79] (модификация – изменение интерфейса ввода последовательности ДНК анцестрального аллеля промотора без каких-либо изменений алгоритма). В окне “Results” на этом рисунке показан результат анализа: оценки “ $-\ln(K_D(S_1)) \pm \delta(S_1)$ ” и “ $-\ln(K_D(S_2)) \pm \delta(S_2)$ ” (здесь K_D и δ – равновесная константа диссоциации комплекса ТВР/промотор и стандартная ошибка оценки величин “ $-\ln(K_D)$ ”), а также оценка p вероятности принятия гипотезы “ $H_0: K_D(S_1) \neq K_D(S_2)$ ” и соответствующую ей величину Z-статистики Фишера. Колонка “ K_D , нМ” табл. 1 содержит эти результаты (где $\alpha \equiv 1 - p$ является оценкой статистической

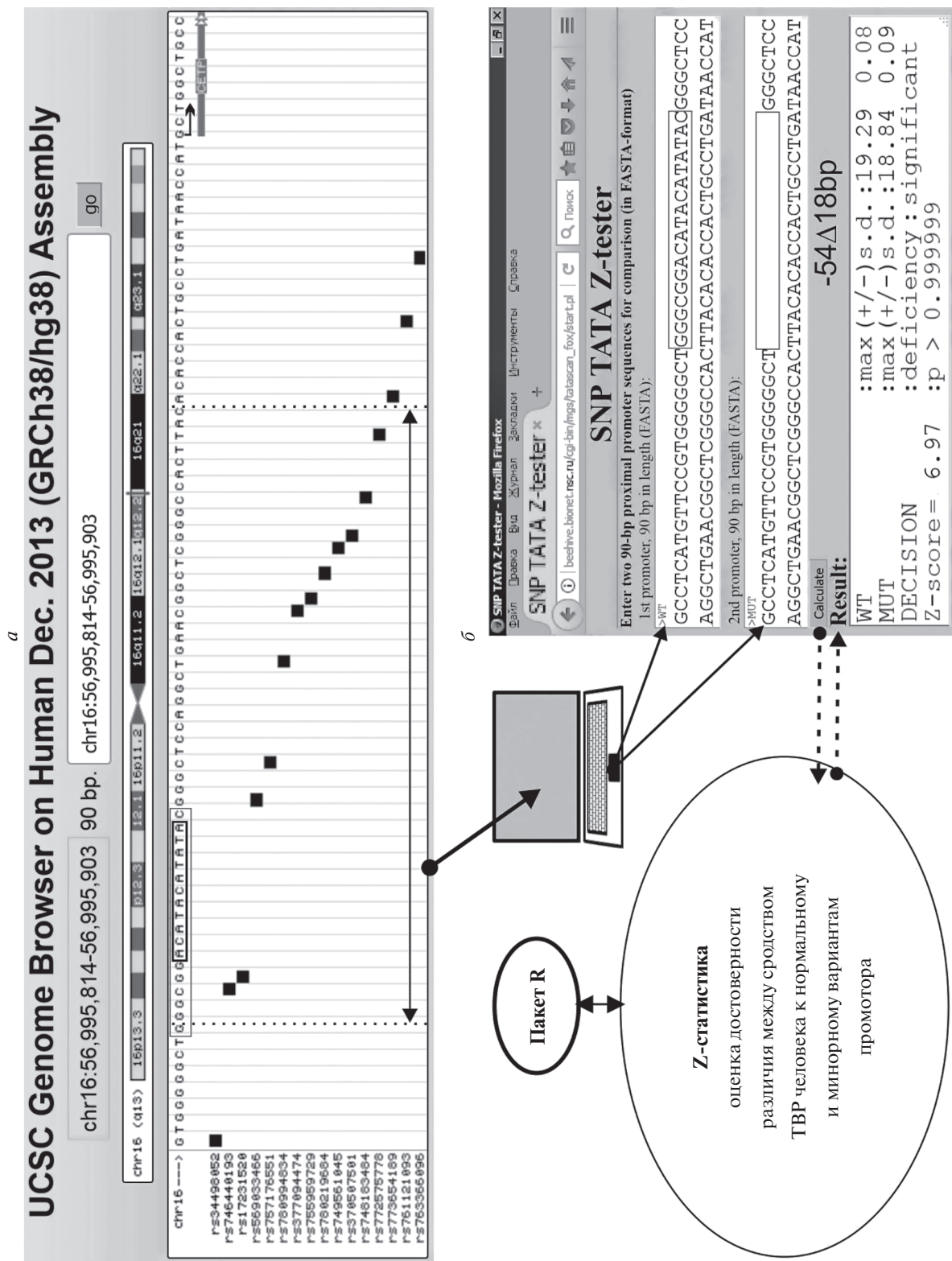


Рис. 1. Применение Web-сервиса SNP_TATA_Z-tester на примере анализа района [—70; —20] (↔) с известным SNP-маркером сниженного риска атеросклероза: делеции g_71ggcggACATACATAc_54 (-54Δ18bp; тонкая рамка) сайта ТВР-связывания (заглавные буквы, жирная рамка) в промоторе гена *CELP2* человека [24]. а — визуализация анализируемого промотора с помощью Web-сервиса “UCSC Genome Browser” [69]; б — ввод данных (→) о норме (wt) и minorном варианте (mut) рассматриваемого промотора, анализ этих данных (пунктирные стрелки) с помощью SNP_TATA_Z-tester и полученный результат в окне “Result”, где в строке “DECISION” можно видеть соответствие этого результата клиническим данным [24].

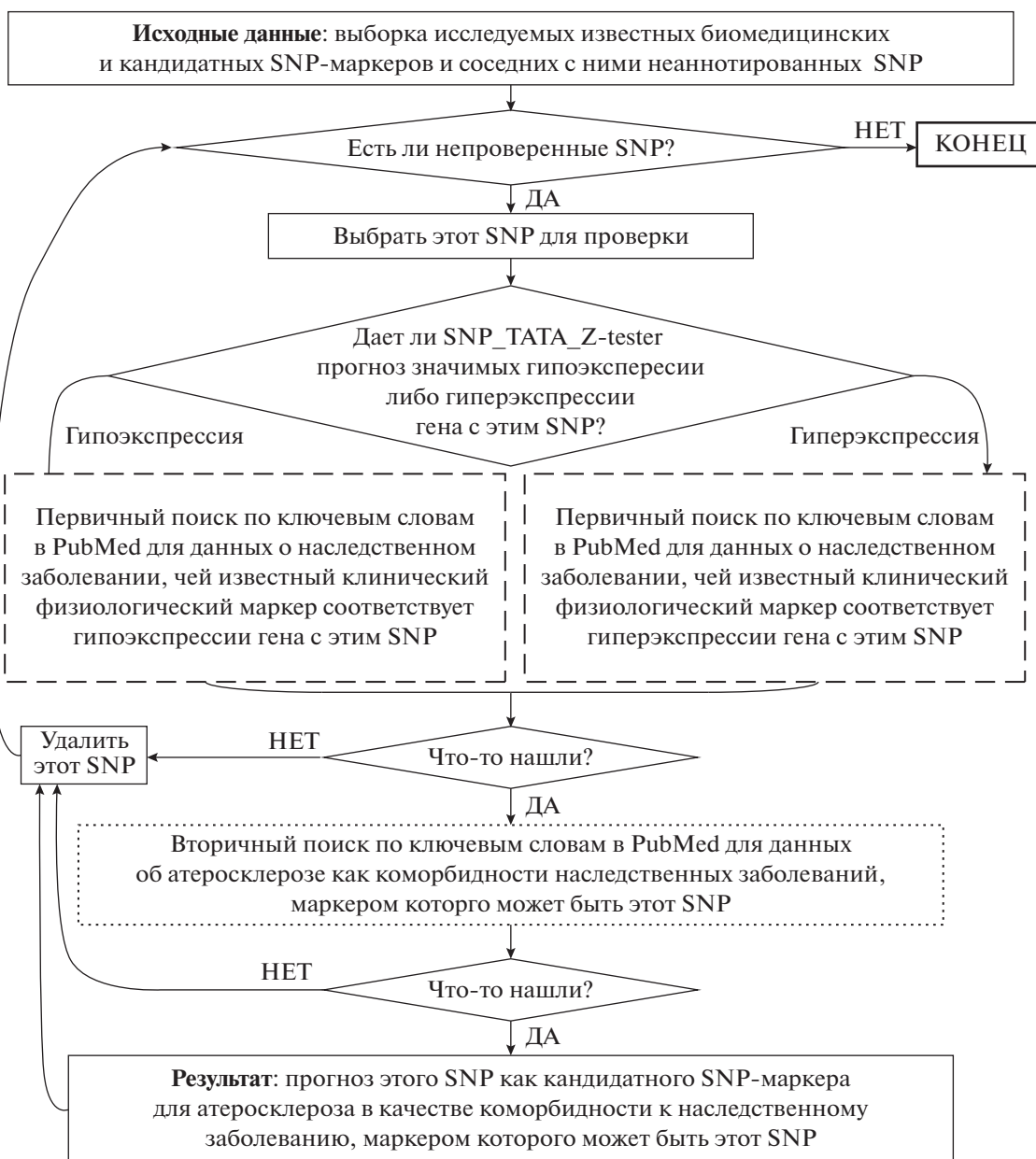


Рис. 2. Блок-схема поиска по ключевым словам для атеросклероза как коморбидности наследственных заболеваний, SNP-маркеры которых могут достоверно изменять уровень сродства TBP к промоторам генов человека.

значимости каждого предсказанного кандидатного SNP-маркера).

Таким образом были исключены все SNP, которые были предсказаны в качестве недостоверно изменяющих сродство TBP/промотор ($\alpha \geq 0.05$, данные не показаны). Затем для каждого оставшегося SNP провели поиск по ключевым словам в базе данных PubMed [80] для клинически доказанных физиологических маркеров атерогенеза, идентичных проявлению этого SNP в изменении уровня экспрессии того гена, промотор которого несет этот SNP, относительно нормы, как это схематически показано на рис. 2.

На рис. 2 два прямоугольника (пунктирные границы) изображают первичный поиск по ключевым словам в базе данных PubMed [80] для наследственных заболеваний, физиологические маркеры которых соответствуют прогнозу достоверного изменения экспрессии гена человека, вызванного исследуемым SNP. Если нет таких данных, этот SNP исключался из дальнейшей проверки, в противном случае проводили вторичный поиск по ключевым словам в PubMed [80] для атеросклероза как коморбидности найденных наследственных заболеваний, чьим маркером может быть проверяемый SNP (прямоугольник с

границами из точек). В зависимости от позитивного или негативного исхода этого дополнительного поиска тестируемый SNP предлагался как кандидатный SNP-маркер атеросклероза (табл. 1) или окончательно отбрасывался (данные не показаны) соответственно.

Полученные в итоге результаты наших прогнозов даны курсивом в двух последних колонках табл. 1, где обозначения “→ AS” и “← AS” соответствуют ускорению и замедлению атерогенеза или усилению и ослаблению атеросклероза (AS) у носителей минорного аллеля SNP в сравнении с нормой.

РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

Было проанализировано 175 SNP промоторов 17 генов человека (*ACKR1*, *APOA1*, *CETP*, *ESR2*, *F3*, *F7*, *GCG*, *HBB*, *HBD*, *IL1B*, *INS*, *LEP*, *MBL2*, *MMP12*, *NOS2*, *SOD1* и *TPI1*), среди которых найдены один известный и 33 кандидатных SNP-маркера для предрасположенности к атеросклерозу, которые были предсказаны в качестве способных достоверно изменять сродство ТВР к промоторам этих генов (табл. 1).

Рассмотрим более подробно пример известного клинического SNP-маркера снижения риска атеросклероза – делеции длиной 18 пн в позиции –54 (здесь и далее: –54Δ18bp) в промоторе гена *CETP* человека [24], с тем, чтобы все остальные 33 кандидатных SNP-маркера, предсказанные в этой работе для атеросклероза, можно было обсудить кратко, по аналогии с этим примером.

Ген CETP человека кодирует белок переноса липидов плазмы, чей дефицит при делеции 5'-g₋₇₁ggcgacatacatatac₋₅₄-3' в промоторе этого гена (рис. 1,а: тонкая рамка) снижает риск атеросклероза [24]. Этот известный биомедицинский SNP-маркер до сих пор не документирован в базе данных dbSNP [9], поскольку он, вероятно, еще не наблюдался в достаточном количестве в 10545 индивидуальных геномах, расшифрованных в рамках проекта “1000 геномов” [8]. Поэтому ожидаемую частоту его встречаемости у человека как биологического вида можно оценить величиной $1/(10454 + 4) < 10^{-5}$, из-за чего ассоциированные с ним патологии соответствуют критерию орфанных (редких) заболеваний [81, 82]. Орфанные заболевания совсем недавно попали в фокус внимания фармакологии в рамках нового направления медицины – предиктивно-превентивной персонализированной медицины [11, 83].

С использованием SNP_TATA_Z-tester был спрогнозирован дефицит *CETP* (рис. 1,б) из-за снижения сродства от 4 нМ ТВР к норме промотора гена *CETP* до 7 нМ к его варианту с этой делецией (табл. 1). Это соответствие между прогнозом *in silico* и клиническими данными в случае из-

вестного SNP-маркера атеросклероза указывает на адекватность результатов SNP_TATA_Z-tester задачам анализа этого заболевания. Кроме того, его прогнозы изменения сродства ТВР/ДНК и уровней экспрессии репортерных генов подтверждены как данными наших экспериментов [34, 84, 85], так и 68 независимых экспериментов других авторов (для обзора см. [86]). Наконец, показано соответствие этих прогнозов клиническим данным (для обзора см. [17, 87]) и фенотипическим проявлениям SNP у лабораторных и хозяйственно ценных животных, растений, микроорганизмов и вирусов (для обзоров см. [69, 76]). Поэтому были аналогично оценены все неаннотированные SNP в районе [–70; –20] этого промотора (рис. 1,а: ↔), в котором проецируются все экспериментально доказанные сайты ТВР-связывания [14]. В итоге среди них не оказалось таких, которые способны снизить сродство ТВР к этому промотору (данные не показаны), как в случае –54Δ18bp. Как можно видеть на рис. 1,а, этот клинический SNP-маркер [24] до сих пор отсутствует в dbSNP [9]. Поэтому мы не указали в табл. 1 ни один из них в качестве кандидатного SNP-маркера атерогенеза. Однако это не означает, что ни один из этих 12 неаннотированных SNP не может быть кандидатным SNP-маркером атерогенеза, поскольку нельзя исключить их влияния на сродство промотора к транскрипционным факторам или к гистоновому октомеру, а также на сайты метилирования промотора, что выходило за рамки исследования влияния SNP на сродство ТВР/промотор. Оценки этих и ряда других молекулярных механизмов влияния SNP на экспрессию генов можно сделать с помощью свободно доступных Web-сервисов (для обзора см. [88, 89]).

Гены ESR2 (эстрогеновый рецептор β) и IL1B (интерлейкин 1β) содержат в промоторах известные SNP-маркеры rs35036378 для ESR2-дефицитной первичной опухоли рГ1 молочной железы [25] и rs1143627 для рака печени, легких и желудка, а также гастрита, язвы желудка, ожирения и ряда психических расстройств вследствие избытка IL1B (для обзора, см. [17]).

В порядке обсуждения следует отметить работы по клиническому исследованию кальциноза, в которых выявили дефицит эстрогенового рецептора β [26] и избыток интерлейкина 1β [28] в качестве физиологических маркеров этой стадии атерогенеза. На основании этих данных предложены кандидатные SNP-маркеры rs35036378 и rs1143627 для ускорения атерогенеза (табл. 1; выделено курсивом в двух последних колонках табл. 1). Прогноз ограничен рамками о клинически найденной корреляции усиления кальциноза с дефицитом эстрогенового рецептора β [26] и с избытком интерлейкина 1β [28] на клинических данных без каких-либо причинно-следственных отношений. Поэтому прогноз кандидатных SNP-маркеров rs35036378 и

rs1143627 лишь эмпирически указывает на предрасположенность к ускоренному кальцинозу без уточнения причин этой особенности протекания атерогенеза.

Вблизи известного SNP-маркера rs35036378 найден неаннотированный SNP rs766797386, который также может вызвать дефицит ESR2 и поэтому также может быть кандидатным SNP-маркером ускорения атерогенеза (табл. 1).

Отметим также данные о том, что повышенное содержание кальция (Ca) в продуктах питания замедляет атерогенез в отличие от Ca-обогащенных пищевых добавок, которые являются фактором риска кальциноза коронарной артерии [27]. Поэтому, в порядке дискуссии, можно ожидать, что у людей с кандидатными SNP-маркерами rs35036378, rs766797386 и rs1143627 ускорения кальциноза эту стадию атерогенеза можно замедлить Ca-богатой диетой, тогда как Ca-обогащенные пищевые добавки могут дополнительно ускорить развитие атеросклероза.

В порядке дискуссии, следует также отметить, что база данных dbSNP [9] содержит частоты встречаемости минорных аллелей некоторых SNP в этнически однородных субпопуляциях, являющихся коренными (аборигенными) для ряда регионов (например, шотландцы в Шотландии), а также основанные на них оценки ожидаемых частот этих аллелей в глобальной популяции людей. Это – миграционные SNP-маркеры для мониторинга в центрах урбанизации населения этнически-специфических демографических и биомедицинских факторов качества жизни людей. Так, спрогнозированный кандидатный SNP-маркер rs35036378 для повышенного риска кальциноза получил оценку популяционной частоты ниже критического порога 2%, тогда как два других – rs766797386 и rs1143627 – не были охарактеризованы такой оценкой. Представляется, что накопление в какой-либо свободно доступной базе данных сведений о популяционных оценках частот как аннотированных, так и неаннотированных SNP могло бы сфокусировать поиск на наиболее адекватных кандидатных SNP-маркерах для решения некоторых задач предиктивно-превентивной персонализированной медицины [11].

Ген *INS* (инсулин) содержит известный SNP-маркер rs5505 для неонатального сахарного диабета с повышенным риском диабета II типа из-за гиперинсулинемии [29], которая в свою очередь является также физиологическим маркером атерогенеза [30]. Вблизи rs5505 найден неаннотированный SNP rs563207167, способный вызвать гиперинсулинемию, и поэтому он может быть кандидатным SNP-маркером ускорения атерогенеза (табл. 1).

Согласно данным, представленным в ретроспективном клиническом обзоре [31], о том, что

физические упражнения и диета против ожирения при сахарном диабете способны также замедлить атерогенез, можно ожидать, что этими методами можно снизить риски сахарного диабета и атеросклероза у людей с кандидатными SNP-маркерами rs5505 и rs563207167 обоих этих заболеваний. Согласно dbSNP [9], популяционные частоты этих кандидатных SNP-маркеров не превышают порог 2%.

Гены *SOD1* (Cu/Zn-супероксиддисмутаза), *TP11* (триозофосфатизомераза), *F3* (тромбопластин) и *F7* (проконвертин) имеют в промоторах известные SNP-маркеры: rs7277748 для бокового амиотрофического склероза, ALS [32], rs1800202 для гемолитической анемии [35] и нейромускулярных заболеваний [36], а также rs563763767 для инфаркта миокарда и тромбоза [39], замену A → C в позиции –31 относительно старта транскрипции (здесь и далее: –31A → C) для умеренных кровотечений [41], являющихся травматическими и ишемическими осложнениями атерогенеза [33, 37, 38, 40, 42]. Вблизи этих биомедицинских SNP-маркеров есть два неаннотированных SNP – rs781835924 и rs749691733, которые способны изменять экспрессию тех же генов также, как это делают известные SNP-маркеры. Поэтому они могут быть кандидатными SNP-маркерами для тех же патологий, включая осложнения атеросклероза (табл. 1).

Отметим, что эти кандидатные SNP-маркеры могли бы помочь интерпретировать результаты совместного таргетного секвенирования геномов лиц с гиперхолестеринемией и их кровных родственников [10], чтобы проследить вклад этих SNP в ускорение атерогенеза при данной патологии. Оценки популяционных частот SNP rs781835924 и rs749691733 в dbSNP [9] не превышают порог 0.01% орфанных заболеваний [71, 72].

Гены *HBB* и *HBD* кодируют β- и δ-цепи гемоглобина. В их промоторах было клинически найдено наибольшее количество SNP-маркеров – rs397509430, rs33980857, rs34598529, rs33931746, rs33981098, rs34500389 и rs35518301 для талассемии и устойчивости к малярии вследствие снижения сродства ТВР к этим промоторам [43]. Ретроспективный обзор [44] сообщает об атеропротекторных свойствах талассемии. Вблизи этих известных биомедицинских SNP-маркеров есть три неаннотированных SNP (rs63750953, rs281864525 и rs34166473), способных повысить риск талассемии и соответственно быть кандидатными SNP-маркерами для замедления атерогенеза (табл. 1).

Ген *MMP12* содержит клинический SNP-маркер rs2276109 снижения риска псориаза [45], астмы [46] и системной склеродермы [47] вследствие сниженного сродства ТВР/промотор, что вызывает дефицит металлоэластазы макрофагов, кодируемой этим геном. Согласно клиническим

данным [48], дефицит MMP12 снижает деградацию фиброгена в крови и следовательно ускоряет атерогенез (табл. 1). Рядом с этим известным SNP-маркером rs2276109 имеется неаннотированный SNP rs572527200, который может снизить сродство ТВР/промотор, вызвать дефицит MMH12 и поэтому быть кандидатным SNP-маркером ускорения атерогенеза.

Ген *MBL2* (маннозосвязывающий лектин) несет в промоторе известный SNP-маркер rs72661131 варибельного иммунодефицита [49], инсульта [51] и преэклампсии [50] из-за дефицита MBL2. Около известного клинического SNP-маркера rs72661131 есть неаннотированный SNP rs562962093, способный также вызывать дефицит MBL2 за счет снижения сродства ТВР к промотору этого гена.

По данным ретроспективного клинического обзора [52] дефицит MBL2 ускоряет атерогенез при тромбозе, формировании и дестабилизации атеросклеротической бляшки. В этом контексте полиморфные варианты rs72661131 и rs562962093 могут быть кандидатными SNP-маркерами ускорения атерогенеза. Кроме того, в статье диетологов [53] описывается механизм снижения рисков инсульта, преэклампсии и атеросклероза при использовании овощно-фруктовых противовоспалительных антиоксидантных диет. Поэтому можно ожидать, что люди с минорными аллелями предсказанных кандидатных SNP-маркеров rs72661131 и rs562962093 ускорения атерогенеза могли бы замедлять развитие атеросклероза в своем организме с использованием такой специально подобранной сбалансированной диеты.

Ген *NOS2* имеет известный SNP-маркер $-51T \rightarrow C$ для эпилепсии [56] и устойчивости к малярии [54, 55] из-за избытка индуцибельной синтазы оксида азота. Этот же маркер является также известным физиологическим маркером ускорения атерогенеза вследствие усиления процессов окисления липопротеинов и активации воспалительного ответа иммунной системы организма [57].

Ген *APOA1* (аполипопротеин А1) несет клинический маркер $-35A \rightarrow C$ стеатоза печени [58] из-за снижения сродства ТВР/промотор и, в итоге, дефицита APOA1 как физиологического маркера ускорения атерогенеза в силу снижения доли липопротеинов высокой плотности [59].

В работе Паролини и соавт. [60] была показана возможность снижения рисков атерогенеза и стеатоза печени при диете на основе антарктического криля. В границах применимости этой находки диетологов, люди с вариантом $-31C$ этого SNP-маркера — известного для стеатоза печени и кандидатного для атеросклероза, могут снизить риски обоих этих заболеваний с использованием диеты на основе антарктического криля [60].

Ген *ACKR1* содержит известный SNP-маркер rs2814778 для лейкопении [61] и устойчивости к малярии [62] из-за низкого сродства ТВР/промотор и соответственно дефицита гликопротеина D, кодируемого этим геном. Отметим данные работы Ван и соавт. [63] об опыте на мышах, где было показано, что дефицит гликопротеина D замедляет атерогенез при недостаточности по аполипопротеину E (табл. 1).

Гены *LEP* (гормон “сытости”, лептин) и *GCG* (гормон “голода”, глюкагон) не имеют известных клинических SNP-маркеров, изменяющих сродство ТВР к промоторам этих генов [9]. Найдены два неаннотированных SNP, rs200487063 и rs34104384, которые увеличивают сродство ТВР к промотору гена *LEP* и могут быть кандидатными SNP-маркерами процессов замедления атерогенеза, поскольку экзогенный рекомбинантный лептин применяют как атеропротекторный препарат [67]. Также найден неаннотированный SNP rs201381696, который может вызвать дефицит лептина и соответственно ожирение [64], известное как фактор риска атеросклероза [65], а также замедления атерогенеза при дефиците APOE [66].

В случае с геном *GCG* были найдены неаннотированные SNP rs183433761 и rs757035851, которые могут снизить сродство ТВР к промотору этого гена и вызвать дефицит глюкагона. В клиническом обзоре Хуббарда и соавт. [68] сообщается о том, что повышенное значение отношения концентраций инсулина и глюкагона в крови после принятия пищи является физиологическим маркером постпрандиального ускорения атерогенеза. Поэтому rs183433761 и rs757035851 могут быть предложены в качестве кандидатных SNP-маркеров для этого пути ускоренного развития атеросклероза (табл. 1).

ЗАКЛЮЧЕНИЕ

Проанализированы 175 SNP промоторов 17 генов человека (*ACKR1*, *APOA1*, *CETP*, *ESR2*, *F3*, *F7*, *GCG*, *HBB*, *HBD*, *IL1B*, *INS*, *LEP*, *MBL2*, *MMP12*, *NOS2*, *SOD1* и *TP11*), абсолютное большинство (142 SNP) из которых были исключены как не влияющие на сродство ТВР к промоторам этих генов. В результате были отобраны остальные 33 кандидатных SNP-маркера атеросклероза, которые способны влиять на сродство ТВР к промоторам 16 генов человека. Важно подчеркнуть, что прежде чем приведенные в этой работе данные о кандидатных SNP-маркерах смогли бы помочь людям, заботящимся о состоянии кровеносных сосудов своего организма, с учетом данных по своему расшифрованному геному, должно быть проведено много сложных, дорогостоящих и трудоемких экспериментальных и клинических исследований. Эти дополнительные исследования по клиническим стандартам и протоколам с ис-

пользованием представительных когорт больных и условно-здоровых добровольцев необходимы для оценки таких характеристик кандидатных SNP-маркеров, как относительный риск RR (Relative Risk) и отношение шансов OR (Odds Ratio), которые существенно уточняют биоинформатическую оценку уровня статистической значимости α , указанную в табл. 1. Эти оценки многократно удешевляются при использовании компьютерных прогнозов самых обоснованных кандидатных SNP-маркеров из всех 8.58 миллиардов потенциально возможных SNP в геноме человека [8, 90]. Поскольку абсолютное большинство SNP являются нейтральными согласно дилемме Холдейна [91] и теории нейтральной эволюции [75], то биоинформатическими методами их можно исключить до начала экспериментальных и клинических исследований.

Текст статьи был написан М.П. Пономаренко при поддержке Минобрнауки РФ по Программе повышения конкурентоспособности ведущих российских университетов среди ведущих мировых научно-образовательных центров (Проект 5-100); идея и план исследования (Н.А.К.) – Интеграционного проекта СО РАН № 0324-2018-0021; обновление Web-сервиса (Д.А.Р., И.В.Ч. и Е.А.О.) и анализ данных (Е.Б.Ш., И.А.Д. и Л.К.С.) – бюджетных проектов № 0324-2019-0040 и 0324-2019-0042 соответственно.

Авторы заявляют об отсутствии конфликта интересов.

Настоящая статья не содержит каких-либо исследований с использованием в качестве объекта животных.

Настоящая статья не содержит каких-либо исследований с участием в качестве объекта людей.

СПИСОК ЛИТЕРАТУРЫ

1. *Barquera S., Pedroza-Tobías A., Medina C. et al.* Global overview of the epidemiology of atherosclerotic cardiovascular disease // *Arch. Med. Res.* 2015. V. 46. № 5. P. 328–338. <https://doi.org/10.1016/j.arcmed.2015.06.006>
2. *Napoli C., D'Armiento F.P., Mancini F.P. et al.* Fatty streak formation occurs in human fetal aortas and is greatly enhanced by maternal hypercholesterolemia. Intimal accumulation of low density lipoprotein and its oxidation precede monocyte recruitment into early atherosclerotic lesions // *J. Clin. Invest.* 1997. V. 100. № 11. P. 2680–2690. <https://doi.org/10.1172/JCI119813>
3. *Li A.C., Glass C.K.* The macrophage foam cell as a target for therapeutic intervention // *Nat. Med.* 2002. V. 8. № 11. P. 1235–1242. <https://doi.org/10.1038/nm1102-1235>
4. *Lusis A.J.* Atherosclerosis // *Nature.* 2000. V. 407. № 6801. P. 233–241. <https://doi.org/10.1038/35025203>
5. *Hirayama S., Soda S., Ito Y. et al.* Circadian change of serum concentration of small dense LDL-cholesterol in type 2 diabetic patients // *Clin. Chim. Acta.* 2010. V. 411. № 3–4. P. 253–257. <https://doi.org/10.1016/j.cca.2009.11.017>
6. *Lathe R., Sapronova A., Kotelevtsev Y.* Atherosclerosis and Alzheimer-diseases with a common cause? Inflammation, oxysterols, vasculature // *BMC Geriatr.* 2014. V. 14. P. 36. <https://doi.org/10.1186/1471-2318-14-36>
7. *Glass C.K., Witztum J.L.* Atherosclerosis: the road ahead // *Cell.* 2001. V. 104(4). P. 503–516. [https://doi.org/10.1016/S0092-8674\(01\)00238-0](https://doi.org/10.1016/S0092-8674(01)00238-0)
8. *Telenti A., Pierce L.C., Biggs W.H. et al.* Deep sequencing of 10,000 human genomes // *Proc. Natl Acad. Sci. USA.* 2016. V. 113. № 42. P. 11901–11906. <https://doi.org/10.1073/pnas.1613365113>
9. *Sherry S.T., Ward M.H., Kholodov M. et al.* dbSNP: the NCBI database of genetic variation // *Nucl. Acids Res.* 2001. V. 29. № 1. P. 308–311. <https://doi.org/10.1093/nar/29.1.308>
10. *Varzari A., Deyneko I.V., Tudor E., Turcan S.* Polymorphisms of glutathione S-transferase and methylenetetrahydrofolate reductase genes in Moldavian patients with ulcerative colitis: Genotype-phenotype correlation // *Meta Gene.* 2016. V. 7. P. 76–82. <https://doi.org/10.1016/j.mgene.2015.12.002>
11. *Trovato G.M.* Sustainable medical research by effective and comprehensive medical skills: overcoming the frontiers by predictive, preventive and personalized medicine // *EPMA J.* 2014. V. 5. № 1. P. 14. <https://doi.org/10.1186/1878-5085-5-14>
12. *Amberger J.S., Bocchini C.A., Schiettecatte F. et al.* OMIM.org: Online Mendelian Inheritance in Man (OMIM®), an online catalog of human genes and genetic disorders // *Nucleic Acids Res.* 2015. V. 43. Database issue. P. D789–D798. <https://doi.org/10.1093/nar/gku1205>
13. *Zerbino D.R., Wilder S.P., Johnson N. et al.* The Ensembl regulatory build // *Genome Biol.* 2015. V. 16. P. 56. <https://doi.org/10.1186/s13059-015-0621-5>
14. *Ponomarenko M., Mironova V., Gunbin K., Savinkova L.* Hogness box // *Brenner's Encyclopedia of Genetics.* San Diego: Academic Press, 2013. V. 3. P. 491–494. <https://doi.org/10.1016/B978-0-12-374984-0.00720-8>
15. *Martianov I., Viville S., Davidson I.* RNA polymerase II transcription in murine cells lacking the TATA binding protein // *Science.* 2002. V. 298. № 5595. P. 1036–1039. <https://doi.org/10.1126/science.1076327>
16. *Mogno I., Vallania F., Mitra R.D., Cohen B.A.* TATA is a modular component of synthetic promoters // *Genome Res.* 2010. V. 20. № 10. P. 1391–1397. <https://doi.org/10.1101/gr.106732.110>
17. *Ponomarenko M.P., Rasskazov D.A., Arkova O.V. et al.* How to use SNP_TATA_Comparator to find a significant change in gene expression caused by the regulatory SNP of this gene's promoter via a change in affinity of the TATA-binding protein for this promoter // *Biomed. Res. Int.* 2015. V. 2015. P. 359835. <https://doi.org/10.1155/2015/359835>
18. *Arkova O.V., Ponomarenko M.P., Rasskazov D.A. et al.* Obesity-related known and candidate SNP markers

- can significantly change affinity of TATA-binding protein for human gene promoters // *BMC Genomics*. 2015. V. 16. Suppl. 13. P. S5.
<https://doi.org/10.1186/1471-2164-16-S13-S5>
19. Chadaeva I.V., Ponomarenko M.P., Rasskazov D.A. et al. Candidate SNP markers of aggressiveness-related complications and comorbidities of genetic diseases are predicted by a significant change in the affinity of TATA-binding protein for human gene promoters // *BMC Genomics*. 2016. V. 17. Suppl. 14. P. 995.
<https://doi.org/10.1186/s12864-016-3353-3>
 20. Chadaeva I.V., Ponomarenko P.M., Rasskazov D.A. et al. Candidate SNP markers of reproductive potential are predicted by a significant change in the affinity of TATA-binding protein for human gene promoters // *BMC Genomics*. 2018. V. 19. Suppl. 3. P. 0.
<https://doi.org/10.1186/s12864-018-4478-3>
 21. Ponomarenko M., Arkova O., Rasskazov D. et al. Candidate SNP markers of gender-biased autoimmune complications of monogenic diseases are predicted by a significant change in the affinity of TATA-binding protein for human gene promoters // *Front. Immunol.* 2016. V. 7. P. 130.
<https://doi.org/10.3389/fimmu.2016.00130>
 22. Ponomarenko P., Rasskazov D., Suslov V. et al. Candidate SNP markers of chronopathologies are predicted by a significant change in the affinity of TATA-binding protein for human gene promoters // *BioMed. Res. Int.* 2016. V. 2016. P. 8642703.
<https://doi.org/10.1155/2016/8642703>
 23. Ponomarenko P., Chadaeva I., Rasskazov D. et al. Candidate SNP markers of familial and sporadic Alzheimer's diseases are predicted by a significant change in the affinity of TATA-binding protein for human gene promoters // *Front. Aging Neurosci.* 2017. V. 9. P. 231.
<https://doi.org/10.3389/fnagi.2017.00231>
 24. Plengpanich W., Le Goff W., Poolsuk S. et al. CETP deficiency due to a novel mutation in the CETP gene promoter and its effect on cholesterol efflux and selective uptake into hepatocytes // *Atherosclerosis*. 2011. V. 216. № 2. P. 370–373.
<https://doi.org/10.1016/j.atherosclerosis.2011.01.051>
 25. Philips S., Richter A., Oesterreich S. et al. Functional characterization of a genetic polymorphism in the promoter of the ESR2 gene // *Horm. Cancer*. 2012. V. 3. № 1–2. P. 37–43.
<https://doi.org/10.1007/s12672-011-0086-2>
 26. McRobb L.S., McGrath K.C.Y., Tsatralis T. et al. Estrogen receptor control of atherosclerotic calcification and smooth muscle cell osteogenic differentiation // *Arterioscler. Thromb. Vasc. Biol.* 2017. V. 37. № 6. P. 1127–1137.
<https://doi.org/10.1161/atvbaha.117.309054>
 27. Anderson J.J., Kruszka B., Delaney J.A. et al. Calcium intake from diet and supplements and the risk of coronary artery calcification and its progression among older adults: 10-year follow-up of the Multi-Ethnic Study of Atherosclerosis (MESA) // *J. Am. Heart. Assoc.* 2016. V. 5. № 10. P. e003815.
<https://doi.org/10.1161/jaha.116.003815>
 28. Ceneri N., Zhao L., Young B.D. et al. Rac2 modulates atherosclerotic calcification by regulating macrophage interleukin-1 β production // *Arterioscler. Thromb. Vasc. Biol.* 2017. V. 37. № 2. P. 328–340.
<https://doi.org/10.1161/atvbaha.116.308507>
 29. Landrum M.J., Lee J.M., Riley G.R. et al. ClinVar: public archive of relationships among sequence variation and human phenotype // *Nucl. Acids Res.* 2014. V. 42. Database issue. P. D980–D985.
<https://doi.org/10.1093/nar/gkt1113>
 30. Cao W., Ning J., Yang X., Liu Z. Excess exposure to insulin is the primary cause of insulin resistance and its associated atherosclerosis // *Curr. Mol. Pharmacol.* 2011. V. 4. № 3. P. 154–166.
<https://doi.org/10.2174/1874467211104030154>
 31. Liu Z., Zhou Z., Huang G. et al. Long-term effects intensive medical therapy on the development and progression of subclinical atherosclerosis and the metabolic syndrome in Chinese patients with type 2 diabetes mellitus // *Medicine (Baltimore)*. 2016. V. 95. № 46. P. e5201.
<https://doi.org/10.1097/md.0000000000005201>
 32. Niemann S., Broom W.J., Brown R.H.Jr. Analysis of a genetic defect in the TATA box of the SOD1 gene in a patient with familial amyotrophic lateral sclerosis // *Muscle Nerve*. 2007. V. 36. № 5. P. 704–707.
<https://doi.org/10.1002/mus.20855>
 33. Rafael H., David J.O., Vilca A.S. Etiology and treatment of amyotrophic lateral sclerosis // *Am. J. Neurodegener. Dis.* 2017. V. 6. № 1. P. 1–8.
 34. Savinkova L., Drachkova I., Arshinova T. et al. An experimental verification of the predicted effects of promoter TATA-box polymorphisms associated with human diseases on interactions between the TATA boxes and TATA-binding protein // *PLoS One*. 2013. V. 8. № 2. P. e54626.
<https://doi.org/10.1371/journal.pone.0054626>
 35. Watanabe M., Zingg B.C., Mohrenweiser H.W. Molecular analysis of a series of alleles in humans with reduced activity at the triosephosphate isomerase locus // *Am. J. Hum. Genet.* 1996. V. 58. № 2. P. 308–316.
 36. Vives-Corrons J.L., Rubinson-Skala H., Mateo M. et al. Triosephosphate isomerase deficiency with hemolytic anemia and severe neuromuscular disease: familial and biochemical studies of a case found in Spain // *Hum. Genet.* 1978. V. 42. № 2. P. 171–180.
 37. Balla G., Vercellotti G., Eaton J.W., Jacob H.S. Heme uptake by endothelium synergizes polymorphonuclear granulocyte-mediated damage // *Trans. Assoc. Am. Physicians.* 1990. V. 103. P. 174–179.
 38. Kioumourtzoglou M.A., Seals R.M., Gredal O. et al. Cardiovascular disease and diagnosis of amyotrophic lateral sclerosis: a population based study // *Amyotroph. Lateral Scler. Frontotemporal. Degener.* 2016. V. 17. № 7–8. P. 548–554.
<https://doi.org/10.1080/21678421.2016.1208247>
 39. Arnaud E., Barbalat V., Nicaud V. et al. Polymorphisms in the 5' regulatory region of the tissue factor gene and the risk of myocardial infarction and venous thromboembolism: the ECTIM and PATHROS studies // *Arterioscler. Thromb. Vasc. Biol.* 2000. V. 20. № 3. P. 892–898.
<https://doi.org/10.1161/01.atv.20.3.892>
 40. Hasenstab D., Lea H., Hart C.E. et al. Tissue factor overexpression in rat arterial neointima models thrombosis and progression of advanced atherosclerosis //

- Circulation. 2000. V. 101. № 22. P. 2651–2657.
<https://doi.org/10.1161/01.cir.101.22.2651>
41. *Kavlie A., Hiltunen L., Rasi V., Prydz H.* Two novel mutations in the human coagulation factor VII promoter // *Thromb. Haemost.* 2003. V. 90. № 2. P. 194–205.
<https://doi.org/10.1160/th02-09-0050>
 42. *Zacharski L.R., Delprete S.A., Kisiel W. et al.* Atherosclerosis and coronary bypass surgery in hereditary factor VII deficiency // *Am. J. Med.* 1988. V. 84. № 5. P. 955–959.
[https://doi.org/10.1016/0002-9343\(88\)90078-2](https://doi.org/10.1016/0002-9343(88)90078-2)
 43. *Martiney J.A., Cerami A., Slater A.F.* Inhibition of hemozoin formation in Plasmodium falciparum trophozoite extracts by heme analogs: possible implication in the resistance to malaria conferred by the beta-thalassaemia trait // *Mol. Med.* 1996. V. 2. № 2. P. 236–246.
 44. *Wang H., Luo W., Wang J. et al.* Paradoxical protection from atherosclerosis and thrombosis in a mouse model of sickle cell disease // *Br. J. Haematol.* 2013. V. 162. № 1. P. 120–129.
<https://doi.org/10.1111/bjh.12342>
 45. *Стародубцева Н.Л., Соболев В.В., Соболева А.Г. и др.* Экспрессия генов металлопротеаз (ММП-1, ММП-2, ММП-9, ММП-12) при псориазе // *Генетика.* 2011. Т. 47. № 9. С. 1254–1261.
 46. *Hunninghake G.M., Cho M.H., Tesfaigzi Y. et al.* MMP12, lung function, and COPD in high-risk populations // *N. Engl. J. Med.* 2009. V. 361. № 27. P. 2599–2608.
<https://doi.org/10.1056/nejmoa0904006>
 47. *Manetti M., Ibba-Manneschi L., Fatini C. et al.* Association of a functional polymorphism in the matrix metalloproteinase-12 promoter region with systemic sclerosis in an Italian population // *J. Rheumatol.* 2010. V. 37. № 9. P. 1852–1857.
<https://doi.org/10.3899/jrheum.100237>
 48. *Motterle A., Xiao Q., Kiechl S. et al.* Influence of matrix metalloproteinase-12 on fibrinogen level // *Atherosclerosis.* 2012. V. 220. № 2. P. 351–354.
<https://doi.org/10.1016/j.atherosclerosis.2011.11.003>
 49. *Boldt A., Culpi L., Tsuneto L. et al.* Diversity of the MBL2 gene in various Brazilian populations and the case of selection at the mannose-binding lectin locus // *Hum. Immunol.* 2006. V. 67. № 9. P. 722–734.
<https://doi.org/10.1016/j.humimm.2006.05.009>
 50. *Sziller I., Babula O., Hupuczi P. et al.* Mannose-binding lectin (MBL) codon 54 gene polymorphism protects against development of pre-eclampsia, HELLP syndrome and pre-eclampsia-associated intrauterine growth restriction // *Mol. Hum. Reprod.* 2007. V. 13. № 4. P. 281–285.
<https://doi.org/10.1093/molehr/gam003>
 51. *Cervera A., Planas A.M., Justicia C. et al.* Genetically-defined deficiency of mannose-binding lectin is associated with protection after experimental stroke in mice and outcome in human stroke // *PLoS One.* 2010. V. 5. № 2. P. e8433.
<https://doi.org/10.1371/journal.pone.0008433>
 52. *Лосин И.Е., Шахнович Р.М., Зыков К.А., Руда М.Я.* Сердечно-сосудистые заболевания и маннозсвязывающий лектин // *Кардиология.* 2014. Т. 54. № 3. С. 64–70.
 53. *Saita E., Kondo K., Momiyama Y.* Anti-inflammatory diet for atherosclerosis and coronary artery disease: antioxidant foods // *Clin. Med. Insights Cardiol.* 2015. V. 8. Suppl. 3. P. 61–65.
<https://doi.org/10.4137/cmc.S17071>
 54. *Burgner D., Rockett K., Ackerman H. et al.* Haplotypic relationship between SNP and microsatellite markers at the NOS2A locus in two populations // *Genes Immun.* 2003. V. 4. № 7. P. 506–514.
<https://doi.org/10.1038/sj.gene.6364022>
 55. *Clark I.A., Rockett K.A., Burgner D.* Genes, nitric oxide and malaria in African children // *Trends Parasitol.* 2003. V. 19. № 8. P. 335–337.
[https://doi.org/10.1016/S1471-4922\(03\)00147-8](https://doi.org/10.1016/S1471-4922(03)00147-8)
 56. *Gonzalez-Martinez J.A., Moddel G., Ying Z. et al.* Neuronal nitric oxide synthase expression in resected epileptic dysplastic neocortex // *J. Neurosurg.* 2009. V. 110. № 2. P. 343–349.
<https://doi.org/10.3171/2008.6.17608>
 57. *Zhao J.F., Shyue S.K., Lin S.J. et al.* Excess nitric oxide impairs LXR(α)-ABCA1-dependent cholesterol efflux in macrophage foam cells // *J. Cell Physiol.* 2014. V. 229. № 1. P. 117–125.
<https://doi.org/10.1002/jcp.24429>
 58. *Matsunaga A., Sasaki J., Han H. et al.* Compound heterozygosity for an apolipoprotein A1 gene promoter mutation and a structural nonsense mutation with apolipoprotein A1 deficiency // *Arterioscler. Thromb. Vasc. Biol.* 1999. V. 19. № 2. P. 348–355.
<https://doi.org/10.1161/01.atv.19.2.348>
 59. *van Capelleveen J.C., Kootte R.S., Hovingh G.K., Bochem A.E.* Myocardial infarction in a 36-year-old man with combined ABCA1 and APOA-1 deficiency // *J. Clin. Lipidol.* 2015. V. 9. № 3. P. 396–399.
<https://doi.org/10.1016/j.jacl.2015.01.006>
 60. *Parolini C., Bjorndal B., Busnelli M. et al.* Effect of dietary components from antarctic krill on atherosclerosis in apoE-deficient mice // *Mol. Nutr. Food Res.* 2017. V. 61. № 12.
<https://doi.org/10.1002/mnfr.201700098>
 61. *Nalls M.A., Wilson J.G., Patterson N.J. et al.* Admixture mapping of white cell count: genetic locus responsible for lower white blood cell count in the Health ABC and Jackson Heart studies // *Am. J. Hum. Genet.* 2008. V. 82. № 1. P. 81–87.
<https://doi.org/10.1016/j.ajhg.2007.09.003>
 62. *Michon P., Woolley I., Wood E.M. et al.* Duffy-null promoter heterozygosity reduces DARC expression and abrogates adhesion of the *P. vivax* ligand required for blood-stage infection // *FEBS Lett.* 2001. V. 495. № 1–2. P. 111–114.
[https://doi.org/10.1016/S0014-5793\(01\)02370-5](https://doi.org/10.1016/S0014-5793(01)02370-5)
 63. *Wan W., Liu Q., Lionakis M.S. et al.* Atypical chemokine receptor 1 deficiency reduces atherogenesis in ApoE-knockout mice // *Cardiovasc. Res.* 2015. V. 106. № 3. P. 478–487.
<https://doi.org/10.1093/cvr/cvv124>
 64. *Zhang Y., Proenca R., Maffei M. et al.* Positional cloning of the mouse obese gene and its human homologue // *Nature.* 1994. V. 372. № 6505. P. 425–432.
<https://doi.org/10.1038/372425a0>
 65. *Skrypnik K., Suliburska J., Skrypnik D. et al.* The genetic basis of obesity complications // *Acta Sci. Pol. Tech-*

- nol. Aliment. 2017. V. 16. № 1. P. 83–91.
<https://doi.org/10.17306/j.afs.2017.0442>
66. Chiba T., Shinozaki S., Nakazawa T. et al. Leptin deficiency suppresses progression of atherosclerosis in apoE-deficient mice // *Atherosclerosis*. 2008. V. 196. № 1. P. 68–75.
<https://doi.org/10.1016/j.atherosclerosis.2007.01.040>
 67. Jun J.Y., Ma Z., Pyla R., Segar L. Leptin treatment inhibits the progression of atherosclerosis by attenuating hypercholesterolemia in type 1 diabetic Ins2(+)/Akit1(-/-) mice // *Atherosclerosis*. 2012. V. 225. № 2. P. 341–347.
<https://doi.org/10.1016/j.atherosclerosis.2012.10.031>
 68. Hubbard R., Kosch C.L., Sanchez A. et al. Effect of dietary protein on serum insulin and glucagon levels in hyper- and normocholesterolemic men // *Atherosclerosis*. 1989. V. 76. № 1. P. 55–61.
[https://doi.org/10.1016/0021-9150\(89\)90193-7](https://doi.org/10.1016/0021-9150(89)90193-7)
 69. Сулов В.В., Пономаренко П.М., Пономаренко М.П. и др. Полиморфизмы ТАТА-боксов генов хозяйственно важных и лабораторных животных и растений, ассоциированные с их селекционно-ценными признаками // *Генетика*. 2010. Т. 46. № 4. С. 448–457.
 70. Hahn S., Buratowski S., Sharp P.A., Guarente L. Yeast TATA-binding protein TFIIID binds to TATA elements with both consensus and nonconsensus DNA sequences // *Proc. Natl Acad. Sci. USA*. 1989. V. 86. № 15. P. 5718–5722.
<https://doi.org/10.1073/pnas.86.15.5718>
 71. Пономаренко П.М., Савинкова Л.К., Драчкова И.А. и др. Пошаговая модель связывания ТВР/ТАТА-бокс позволяет предсказать наследственное заболевание человека по точечному полиморфизму // *Докл. Акад. наук*. 2008. Т. 419. № 6. С. 828–832.
 72. Bucher P. Weight matrix descriptions of four eukaryotic RNA polymerase II promoter elements derived from 502 unrelated promoter sequences // *J. Mol. Biol.* 1990. V. 212. № 4. P. 563–578.
[https://doi.org/10.1016/0022-2836\(90\)90223-9](https://doi.org/10.1016/0022-2836(90)90223-9)
 73. Ponomarenko M.P., Ponomarenko J.V., Frolov A.S. et al. Identification of sequence-dependent features correlating to activity of DNA sites interacting with proteins // *Bioinformatics*. 1999. V. 15. № 7/8. P. 687–703.
<https://doi.org/10.1093/bioinformatics/15.7.687>
 74. Karas H., Knuettel R., Schulz W. et al. Combining structural analysis of DNA with search routines for the detection of transcription regulatory elements // *Comput. Appl. Biosci.* 1996. V. 12. № 5. P. 441–446.
 75. Kimura M. Evolutionary rate at the molecular level // *Nature*. 1968. V. 217. № 5129. P. 624–626.
<https://doi.org/10.1038/217624a0>
 76. Suslov V.V., Ponomarenko P.M., Efimov V.M. et al. SNPs in the HIV-1 TATA box and the AIDS pandemic // *J. Bioinform. Comput. Biol.* 2010. V. 8. № 3. P. 607–625.
<https://doi.org/10.1142/S0219720010004677>
 77. Waardenberg A.J., Basset S.D., Bouveret R., Harvey R.P. CompGO: an R package for comparing and visualizing Gene Ontology enrichment differences between DNA binding experiments // *BMC Bioinformatics*. 2015. V. 16. P. 275.
<https://doi.org/10.1186/s12859-015-0701-2>
 78. Ponomarenko M., Rasskazov D., Chadaeva I. et al. SNP_TATA_Comparator: genomewide landmarks for preventive personalized medicine // *Front Biosci. (Schol. Ed.)*. 2017. V. 9. № 2. P. 276–306.
<https://doi.org/10.2741/s488>
 79. Haeussler M., Raney B.J., Hinrichs A.S. et al. Navigating protected genomics data with UCSC Genome Browser in a box // *Bioinformatics*. 2015. V. 31. № 5. P. 764–766.
<https://doi.org/10.1093/bioinformatics/btu712>
 80. Lu Z. PubMed and beyond: a survey of web tools for searching biomedical literature // *Database (Oxford)*. 2011. V. 2011. P. baq036.
<https://doi.org/10.1093/database/baq036>
 81. Denis A., Mergaert L., Fostier C. et al. Orphan diseases and orphan medicines: a Belgian and European study // *J. Pharm. Belg.* 2009. V. 86. № 4. P. 131–137.
 82. Rubbert-Roth A. Orphan diseases in rheumatology. Exemplified by polyarteritis nodosa // *Z. Rheumatol.* 2012. V. 71. № 2. P. 119–121.
<https://doi.org/10.1007/s00393-011-0902-7>
 83. Brewer G.J. Drug development for orphan diseases in the context of personalized medicine // *Transl. Res.* 2009. V. 154. № 6. P. 314–322.
<https://doi.org/10.1016/j.trsl.2009.03.008>
 84. Drachkova I., Savinkova L., Arshinova T. et al. The mechanism by which TATA-box polymorphisms associated with human hereditary diseases influence interactions with the TATA-binding protein // *Hum. Mutat.* 2014. V. 35. № 5. P. 601–608.
<https://doi.org/10.1002/humu.22535>
 85. Arkova O., Kuznetsov N., Fedorova O., Savinkova L. A real-time study of the interaction of TBP with a TATA box-containing duplex identical to an ancestral or minor allele of human gene LEP or TPI // *J. Biomol. Struct. Dyn.* 2017. V. 35. № 14. P. 3070–3081.
<https://doi.org/10.1080/07391102.2016.1241190>
 86. Пономаренко П.М., Сулов В.В., Савинкова Л.К. и др. Точное уравнение равновесия четырех шагов связывания ТВР с ТАТА-боксом для прогноза фенотипического проявления мутаций // *Биофизика*. 2010. Т. 55. № 3. С. 400–414.
 87. Пономаренко П.М., Пономаренко М.П., Драчкова И.А. и др. Прогноз изменения аффинности ТАТА-связывающего белка к ТАТА-боксам в результате полиморфизмов ТАТА-боксов промоторов генов человека // *Мол. биол.* 2009. Т. 43. № 3. С. 512–520.
 88. Deplancke B., Alpern D., Gardeux V. The genetics of transcription factor DNA binding variation // *Cell*. 2016. V. 166. № 3. P. 538–554.
<https://doi.org/10.1016/j.cell.2016.07.012>
 89. Gunbin K.V., Ponomarenko M.P., Suslov V.V. et al. Evolution of brain active gene promoters in human lineage towards the increased plasticity of gene regulation // *Mol. Neurobiol.* 2018. V. 55. № 3. P. 1871–1904.
<https://doi.org/10.1007/s12035-017-0427-4>
 90. Wu J., Wu M., Li L. et al. dbWGF: a database and web server of human whole-genome single nucleotide variants and their functional predictions // *Database (Oxford)*. 2016. V. 2016. P. baw024.
<https://doi.org/10.1093/database/baw024>
 91. Haldane J.B.S. The cost of natural selection // *J. Genet.* 1957. V. 55. № 3. P. 511–524.

Candidate SNP-Markers of Atherosclerosis, Which May Significantly Change the Affinity of the TATA-Binding Protein for the Human Gene Promoters

M. P. Ponomarenko^{a, b, *}, D. A. Rasskazov^a, I. V. Chadaeva^{a, b}, E. B. Sharypova^a, I. A. Drachkova^a, P. M. Ponomarenko^c, E. A. Oshchepkova^a, L. K. Savinkova^a, and N. A. Kolchanov^{a, b}

^a*Federal Research Center, Institute of Cytology and Genetics, Siberian Branch, Russian Academy of Sciences, Novosibirsk, 630090 Russia*

^b*Novosibirsk National Research State University, Novosibirsk, 630090 Russia*

^c*University of La Verne, La Verne, 91750 California, USA*

**e-mail: pon@bionet.nsc.ru*

Atherosclerosis (AS) and AS-related pathologies such as coronary heart disease, myocardial infarction, angina pectoris, and stroke are the leading causes of death in the world. The atherogenesis can be influenced by many factors, in particular overweight, hypertension, diabetes, hyperlipoproteinemia and other diseases in anamnesis, as well as genetic predisposition, which may be due to the presence of single nucleotide polymorphism (SNP) in some cases. In this work, we studied only those regions of promoters of the human protein-coding genes, where single-nucleotide (SNP) markers of the changes in the affinity of the TATA-binding protein (TBP) for these promoters have been already associated with the atherosclerosis-related pathologies. As a result, within the dbSNP database we found those unannotated SNPs of these areas that change such affinity as the known biomedical SNP markers do here (according to the predictions made by our Web service SNP_TATA_Z-tester, http://www.mgs.bionet.nsc.ru/cgi-bin/mgs/tatascan_fox/start.pl). For example, the known SNP marker rs35036378 of the high risks of the primary pT1 tumor reduces the TBP affinity for the *ESR2* gene promoter and, thus, the estrogen receptor β abundance in blood that is a known physiological marker of the calcification of blood vessels in atherogenesis. Near this well-known SNP marker, we found an unannotated SNP rs766797386, which can also to reduce the TBP affinity for the same promoter and, thus, to decrease the abundance of the estrogen receptor β in blood. Thus, we propose rs766797386 as a candidate SNP marker for the acceleration of atherogenesis by the calcification of blood vessels. The possibility of using a diet of the natural food of a high abundance of calcium (Ca), which is recommended by nutritionists to slow down calcification in the case of a person having SNP-markers of the accelerated calcification, is discussed in contrast to the use of Ca-enriched nutritional supplements that can cause the opposite effect. By the same way, a total number of 33 candidate SNP markers were predicted and discussed to speed up or slow down atherogenesis. After clinical verification, the candidate SNP markers predicted by this work can help one who would like to slow down his/her atherogenesis by correcting his/her lifestyle using his/her individual genome sequenced.

Keywords: atherosclerosis, promoter, single nucleotide polymorphism (SNP), TATA binding protein (TBP), TBP-binding site (TATA box).