



© И. В. Тарковская¹,
О. С. Глозов^{1,2}, Е. Ю. Диткина¹,
Е. С. Вашукова^{1,2}, А. С. Глозов^{1,2},
Р. В. Курилов¹, И. В. Пугачева¹,
О. Л. Белоног¹, И. А. Махрова¹,
М. В. Асеев^{1,2}, Т. Э. Иващенко^{1,2},
В. С. Баранов²

АНАЛИЗ АССОЦИАЦИИ ПОЛИМОРФИЗМА ГЕНОВ МЕТАБОЛИЗМА ЛИПИДОВ С ИНДЕКСОМ МАССЫ ТЕЛА, ОБХВАТОМ ТАЛИИ И ПАРАМЕТРАМИ ЛИПИДОГРАММЫ КРОВИ У ЖЕНЩИН

ВВЕДЕНИЕ

Идентификация генов и аллелей, контролируемых сложные признаки, является одной из важных задач генетики человека. Анализ конкретных аллелей, контролируемых подобные признаки, широко применяется в генетике человека, в медицине и криминалистике (Баранов и др., 2009; Аульченко, 2010).

По данным Всемирной организации здравоохранения (ВОЗ), избыточная масса тела и ожирение являются главными причинами развития таких болезней как сахарный диабет 2-го типа (СД2), сердечно-сосудистые заболевания (ССЗ), онкологические заболевания и прочих. То, что распространенность ожирения выше в женской популяции, чем в мужской, хорошо известно и подтверждается большим числом исследований (Седлецкий, 2007; Дедов, Петеркова, 2006; Подзолкова, 2006; Чазова, 2010). Данный факт обусловлен особенностями эндокринной функции женского организма. Более того у женщин ожирение может происходить из-за некоторых физиологических состояний, таких как беременность, лактация и климакс.

С 1995 г. по рекомендации Международной группы по ожирению ВОЗ (IOTFWHO) оценка ожирения у взрослых пациентов производится путем подсчета индекса массы тела (ИМТ). В исследованиях Diabetes Prevention Program (DPP) и Diabetes Prevention Study (DPS) показано, что снижение массы тела на 7% от исходной уменьшает частоту развития СД2 на 58%. Другим критерием ожирения является объем талии (ОТ). ОТ ≥ 90-го перцентилля указывает на ожирение центрального типа, которое является ведущим фактором риска развития метаболического синдрома (МС) и его осложнений: СД2 и атеросклероза (Рахимова, 2009; Поляков, 2009).

Также фактором риска развития этих заболеваний является нарушение метаболизма липидов. Повышенные уровни общего холестерина (ХС), холестерина липопротеинов низкой (ХСЛПНП) и очень низкой плотности (ХСЛПОНП) могут приводить к таким заболеваниям как ИБС, заболевания коронарных артерий или инсульт. Результаты крупных клинических и генетических исследований показали, что снижение плазменных уровней общего ХС и ХСЛПНП приводит к достоверному снижению риска возникновения ИБС и общей смертности (Prevention of Coronary Heart Disease in Clinical Practice, 1998).

На сегодняшний день роль наследственности в генезисе ожирения не вызывает сомнений (Седлецкий, 2007). Еще в 1962 году американский ученый Neel J. V. в своей теории «экономного генотипа» показал генетическую предрасположенность к ожирению (Neel, 1962).

Ожирение относится к группе мультифакториальных заболеваний (МФЗ), в развитии которых важную роль отводят наследственному фактору (Ройтберг, 2007). Изучение данного типа заболеваний включает в себя комплексный подход — это и выявление генов-кандидатов, и геномный поиск с использованием множества полиморфных маркеров, исследования аллельных вариантов в генных сетях (Дедов, Шестакова, 2006; Баранов и др., 2009).

¹ООО «БиоГлот»,
Санкт-Петербург;

²ФГБУ НИИАГ Д. О. Отта СЗО
РАМН, Санкт-Петербург,

✿ **Методом ПДРФ-анализа изучен полиморфизм 36 генов, вовлеченных в метаболизм липидов, у 212 женщин в возрасте от 18 до 77 лет, проживающих на территории Северо-Западного региона России (города Санкт-Петербурга). Показана ассоциация полиморфизма некоторых изученных генов с индексом массы тела, обхватом талии, уровнями общего холестерина, холестерина липидов низкой и очень низкой плотности. На основании метода логистической регрессии разработана модель, позволяющая проводить первичную оценку изученных параметров у женщин (индекс массы тела, обхват талии, уровни общего холестерина, холестерина липидов низкой и очень низкой плотности) с помощью тестирования соответствующих генетических маркеров.**

✿ **Ключевые слова:** полиморфизм генов; метаболизм липидов; ХС ЛПНП; ХС ЛПОНП; гены-кандидаты; ИМТ; ОТ; корреляция; регрессия.

Поступила в редакцию 29.10.2012
Принята к публикации 16.11.2012

Таблица 1

Клиническая характеристика исследуемой группы

Признак	Min значение признака	Max значение признака	Среднее значение признака
ИМТ	16,7	47,3	25,3
Холестерин	2,5	9,3	5,4
ЛПНП	0,7	7,1	3,2
ЛПОНП	0,1	1,5	0,6
Обхват талии	57	120	79,7

В настоящий момент уже открыто множество генов, участвующих в развитии ожирения. Возможность выявления соответствующих генов, вероятно, позволит системой мероприятий предотвратить развитие тяжелых форм ожирения и связанных с ним осложнений у людей с наследственной предрасположенностью к МС на догоспитальном этапе, а также позволит выделить различные подтипы ожирения, связанные с СД, ССЗ (Walley, 2006).

Генетические и физиологические основы липидного метаболизма хорошо изучены как на модельных объектах, так и на примере моногенных заболеваний. Не будет преувеличением сказать, что уровень липидов в крови человека — один из наиболее хорошо генетически изученных сложных количественных признаков человека (Friedlander et al., 1997; Pillia et al., 2006). Более того, в отличие от большинства сложных количественных признаков человека для уровней липидов известен ряд генов, вариация которых объясняет существенную долю дисперсии признака в популяции, например аллели $\epsilon 2/3/4$ гена *APOE* (Singand Davignon, 1985; Аульченко, 2010).

Важно отметить, что использование традиционных методов статистики для анализа фенотипических признаков человека при изучении большого числа генов не позволяет дать объективную оценку количественных признаков, и существуют серьезные ограничения по статистическим методам обработки данных при их анализе (Реброва, 2003). Поэтому одной из наиболее подходящих для анализа количественных фенотипических показателей является обобщенная линейная регрессионная модель (GLM), которая позволяет работать с непрерывными значениями.

Изучение генетических профилей человека еще далеко от завершения, есть значительный простор для дальнейшего анализа генетического полиморфизма и геномных профилей, влияющих на уровни липидов в сыворотке крови. В связи с этим использование линейной модели для поиска ассоциаций геномных полиморфизмов с уровнем липидов в крови в популяции человека и их прогноза является актуальным.

Поэтому целью данной работы являлся анализ ассоциации полиморфизма генов метаболизма липидов с ИМТ, ОТ и уровнями общего холестерина, холестерина липидов низкой и очень низкой плотности в крови у женщин и разработка общей модели предсказания данных фенотипических признаков.

МАТЕРИАЛЫ И МЕТОДЫ

Характеристика группы

В группу вошли 212 женщин в возрасте от 18 до 77 лет, проживающие на территории Северо-Западного региона России. Средний возраст группы составил $40,9 \pm 13,1$ лет. Все участники исследования подписали информированное согласие на участие в НИР и заполнили анкеты, содержащие следующие данные: индивидуальную информационную карту с личными данными (место рождения, возраст, сведения о родственниках, образовании, работе, физической активности, питании, привычках и др.), антропометрические данные, сведения о распределении жира, болезнях и лекарствах, операциях и др. Кроме того, в рамках исследования был проведен анализ следующих биохимических показателей — уровень глюкозы, общий ХС, ТГ, ХСЛПВП, ХСЛПНП, ХСЛПОНП, инсулин, НОМА-R в сыворотке крови натощак (в утренние часы).

$\text{НОМА-R} = (\text{инсулин плазмы натощак (мкЕд/мл)} \times \text{глюкоза плазмы натощак (ммоль/л)}) / 22,5$.

Клиническая характеристика исследуемой группы представлена в таблице 1.

Молекулярно-генетические исследования: анализ ДНК-полиморфизма

Образцы ДНК выделяли из лейкоцитов периферической крови методом фенол-хлороформной экстракции (Маниатис и др., 1984).

Методом ПЦР-ПДРФ анализа изучены 36 полиморфных ДНК-локусов:

rs3791675 *EFEMP1*,
 rs6763931 *ZBTB38*,
 rs1812175 *HHIP*,
 rs6830062 *LCORL*,
 rs4842838 *ADAMTSL3*,
 rs2554380 *ADAMTSL3*,
 rs1035569 *CDH13*,
 rs849140 *JAZF1*,
 rs5742694 *IGFBP3*,
 rs907806 *IGF1R*,
 rs572169 *GHS-R*,
 rs4800148 *CABLES1*,
 rs1904559 *IFNG*,
 rs731236 *VDR3*,
 rs2568958 *NEGR1*,
 rs7498665 *SH2B1*,

rs4410646 *GHR*,
rs2075356 *GHRL*,
rs1800629 *TNF- α* ,
rs2716609 *LPIN1*,
rs9340799 *ESR1*,
rs10146997 *NRXN3*,
rs545854 *MSRA*,
rs987237 *TFAP2B*,
rs2013751 *CEL*,
rs174570 *FAD s2/s3*,

rs4938303 кластер *APOA1/APOC3/APOA4/APOA5*,
rs5128 *APOA1/APOC3/APOA4/APOA5*,
rs1800588 *LIPC*,
rs3890182 *ABCA1*,
rs16996148 *CILP*,
rs6754295 *APOB*,
rs5888 *SCARB1*,
rs34845087 *LIPE*,
rs10865710 *PPARG*,
rs1801282 *PPARG* (табл. 2).

Таблица 2

Размеры ПЦР-продуктов, ферменты гидролиза, ожидаемые размеры фрагментов ДНК после анализа и соответствующие генотипы

Ген	Номер rs	Замена	Эндонуклеаза	Ожидаемые размеры фрагментов ДНК		
				G/G	G/A	A/A
<i>EFEMP1</i>	rs3791675	G>A	Alu I	198+33нп	231+198+33 нп	231 нп
<i>ZBTB38</i>	rs6763931	G>A	BstDE I	179 нп	162+27 нп	179+162+27 нп
<i>HHIP</i>	rs1812175	A>G	Taq I	250 нп	250+224+26 нп	224+26 нп
<i>LCORL</i>	rs6830062	A>G	BamH I	196 нп	196+138+58 нп	138+58 нп
<i>ADAMTSL3</i>	rs4842838	G>T	Hind III	289 нп	289+267+22 нп	267+22 нп
<i>ADAMTSL3</i>	rs2554380	T>C	BstNS I	205+69 нп	274+205+69 нп	274 нп
<i>CDH13</i>	rs1035569	A>G	Hinf I	210 нп	210+187+29 нп	187+29 нп
<i>JAZF1</i>	rs849140	A>G	Fsp4H I	219 нп	219+140+79 нп	140+79 нп
<i>IGFBP3</i>	rs5742694	T >G	Fok I	200 нп	200+143+57 нп	143+57 нп
<i>IGF1R</i>	rs907806	A>G	Bse4 I	246 нп	246+222+24 нп	222+24 нп
<i>GHS-R</i>	rs572169	G>A	BspAC I	114+96нп	211+114+96 нп	211 нп
<i>CABLES1</i>	rs4800148	G>A	Acl I	185+89нп	274+185+89 нп	274 нп
<i>IFNG</i>	rs1904559	G>A	Ags I	141 нп	141+95+46 нп	95+46 нп
<i>VDR3</i>	rs731236	T>C	Taq I	292 нп	292+190+102нп	190+102 нп
<i>NEGR1</i>	rs2568958	C>T	Bst4C I	140+22нп	162+140+22 нп	162 нп
<i>SH2B1</i>	rs7498665	A>G	BstMW I	154 нп	154+132+22 нп	132+22 нп
<i>GHR</i>	rs4410646	A>C	Bse1 I	150+68нп	218+150+68 нп	218 нп
<i>GHRL</i>	rs2075356	A>G	Acu I	179 нп	179+126+53 нп	126+53 нп
<i>TNF-α</i>	rs1800629	G>A	Bsp19 I	208+20нп	228+208+20 нп	228 нп
<i>LPIN1</i>	rs2716609	T>C	Hae III	247 нп	247+151+96 нп	151+96 нп

Таблица 2 (Продолжение)

Ген	Номер rs	Замена	Эндонуклеаза	Ожидаемые размеры фрагментов ДНК		
<i>ESR1</i>	rs9340799	A>G	XbaI	A/A 228 + 19 нп	A/G 47 + 228 + 19 нп	G/G 247 нп
<i>NRXN3</i>	rs10146997	A>G	Rsa I	A/A 156 + 99 нп	A/G 156 + 99 + 77 + 27 нп	G/G 156 + 77 + 27 нп
<i>MSRA</i>	rs545854	G>C	Bst6 I	G/G 141 нп	G/C 141 + 77 + 67 нп	C/C 77 + 67 нп
<i>TFAP2B</i>	rs987237	A>G	Fae I	A/A 173 + 72 нп	A/G 245 + 173 + 72 нп	G/G 245 нп
<i>CEL</i>	rs2013751	A>G	Acc36 I	A/A 119 + 93 нп	A/G 212 + 119 + 93 нп	G/G 212 нп
<i>FAD s2/s3</i>	rs174570	C>T	Bse8 I	C/C 212 нп	C/T 212 + 109 + 103 нп	T/T 109 + 103 нп
<i>APOA1 APOC3 APOA4 APOA5</i>	rs4938303	G>A	Mbo II	G/G 271 нп	G/A 271 + 149 + 122 нп	A/A 149 + 122 нп
<i>APOA1 APOC3 APOA4 APOA5</i>	rs5128	C>G	EcoICR I	C/C 212 нп	C/G 212 + 135 + 71 нп	G/G 135 + 71 нп
<i>LIPC</i>	rs1800588	C>T	Fae I	C/C 235 нп	C/T 235 + 164 + 71 нп	T/T 164 + 71 нп
<i>ABCA1</i>	rs3890182	G>A	Ags I	G/G 166 нп	G/A 166 + 91 + 75 нп	A/A 91 + 75 нп
<i>CILP</i>	rs16996148	G>T	Fae I	G/G 122 нп	G/T 122 + 97 + 25 нп	T/T 97 + 25 нп
<i>APOB</i>	rs6754295	A>C	Bsc4 I	A/A 219 нп	A/C 219 + 199 + 20 нп	C/C 199 + 20 нп
<i>SCARB1</i>	rs5888	C>T	BstDE I	C/C 143 нп	C/T 143 + 122 + 21 нп	T/T 122 + 21 нп
<i>LIPE</i>	rs34845087	C>G	Fsp4H I	C/C 121 + 50 + 15 нп	C/G 171 + 121 + 50 + 15 нп	G/G 171 + 15 нп
<i>PPARG</i>	rs10865710	C>G	Fae I	C/C 155 + 88 нп	C/G 242 + 155 + 88 нп	G/G 242 нп
<i>PPARG</i>	rs1801282	C>G	BstFN I	C/C 399 нп	C/G 399 + 369 + 30 нп	G/G 369 + 30 нп

Для идентификации всех анализируемых SNP подобраны специфические эндонуклеазы рестрикции и соответствующие олигонуклеотиды для проведения ПЦР. Размеры ПЦР-продуктов, а также использованные для гидролиза эндонуклеазы и размеры ДНК-фрагментов после гидролиза также представлены в таблице 2.

Реакционная смесь (25 мкл) ПЦР содержала 0,4 пмоль каждого праймера (исходная концентрация раствора праймеров 10 pmoI/ul), 67 мМ Трис-НСI (рН 8,6), 166 мМ (NH₄)₂SO₄, 0,01 % Тритон X-100, 1,5 мМ MgCl₂, 0,2 мМ каждого из dNTP («Силекс», Россия) и 2,5 ед. акт. Таq-полимеразы («Силекс», Россия), 100 нг–200 нг ДНК.

Для амплификации исследуемых фрагментов использовали следующий температурный режим ПЦР:

после предварительной денатурации ДНК (5 минут при 95 °С) проводили 35 циклов амплификации в режиме: 94 °С — 30 с, 60 °С — 30 с, 72 °С — 1 мин, далее заключительный этап синтеза при 72 °С — 5 мин.

ПЦР продукты подвергали гидролизу эндонуклеазами рестрикции (табл. 2) в соответствии с рекомендациями фирмы-изготовителя («СибЭнзим», г. Новосибирск).

Продукты амплификации и рестрикции анализировали с помощью электрофоретического разделения в 6%-м неденатурирующем полиакриламидном геле, который далее окрашивали в водном растворе бромистого этидия (0,5 мкг/мл) с последующей визуализацией ДНК на аппарате Praktika (ФРГ) в проходящем ультрафиолетовом свете на трансиллюминаторе Масgrovie (ЛКВ, Великобритания).

Статистическая обработка результатов

Индекс массы тела рассчитывали по формуле:

$$I = \frac{m}{h^2}$$

где: m — масса тела в килограммах, h — рост в метрах.

Для анализа корреляции между генами и признаками применен метод ранговой корреляции по Кендаллу (τ). Данный метод позволяет сравнивать качественные порядковые признаки между собой (генотипы) и качественные порядковые признаки (генотипы) с количественными (ИМТ, уровень ХС и т.д.) (Реброва, 2003).

Для построения модели предсказания количественных признаков (ИМТ, ОБ, ХС, ХС ЛПНП, ХС ЛПОНП) применяли логистическую регрессию. В качестве исходных параметров модели использовались данные об аллелях индивидуумов по рассматриваемым SNP: гомозиготы по аллели, имеющей большую частоту (в исследуемой популяции), получали код «0», гетерозиготы — код «1», гомозиготы по аллели, имеющей меньшую частоту — код «2».

Формула для линейной регрессионной модели в общем виде:

$$Y = \beta_0 + \sum_j \beta_j x_j = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n$$

где Y — это значение предсказываемого признака, β_0 — свободный член, x_j — параметр, соответствующий j -му SNP у индивидуума, может принимать значения {0, 1, 2}, β_j — соответствующий регрессионный коэффициент при параметре x_j (т.е. соответствующий j -му SNP).

Далее методом пошагового включения/удаления параметров (stepwise algorithm) в случае каждого признака была определена наиболее точная модель и, соответственно, наиболее значимые параметры (SNP). В качестве критерия точности модели использовали информационный критерий Акайке (AIC), который может быть выражен в виде:

$$AIC = 2k + n(\ln(RSS))$$

где k — число параметров модели, n — число индивидуумов, по которым строится модель, а $RSS = \sum_{i=1}^n \varepsilon_i^2$ (residual sum of squares — остаточная сумма квадратов), где ε_i разность реального и предсказанного моделью значения рассматриваемого признака у i -го индивидуума. Средняя ошибка = $1/n \sum_{i=1}^n \varepsilon_i$.

РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

Частоты аллелей по проанализированным 36 генам, вовлеченным в метаболизм липидов, у 212 женщин, проживающих на территории Северо-Западного региона России (города Санкт-Петербурга) представлены в таблице 3.

При сравнении частот аллелей вышеперечисленных генов между группой жителей Санкт-Петербурга и «усредненной» европейской популяцией, представленной в базе NCBI (<http://www.ncbi.nlm.nih.gov/SNP>), популяционные идентификаторы: «HarMap-CEU», «CAUC3», «AFD_EUR_PANEL»), статистические значимые отличия были выявля-

Таблица 3

Частоты генотипов и аллелей по всем исследуемым генам в группе женщин N = 212

Гены	Номер rs	Генотипы и их частоты			Аллели и их частоты	
		A/A	G/A	G/G	A	G
EFEMP1	rs3791675	0,123	0,401	0,476	0,323	0,677
		A/A	G/A	G/G	A	G
ZBTB38	rs6763931 rs1812175	0,156	0,557	0,288	0,434	0,566
		A/A	A/G	G/G	A	G
HHIP	rs1812175	0,033	0,208	0,759	0,137	0,863
		A/A	A/G	G/G	A	G
LCORL	rs6830062	0,759	0,208	0,033	0,863	0,137
		G/G	A/G	G/G	A	G
ADAMTSL3	rs4842838	0,208	0,486	0,307	0,45	0,55
		G/G	G/T	T/T	G	T
ADAMTSL3	rs2554380	0,071	0,311	0,618	0,226	0,774
		C/C	T/C	T/T	C	T
CDH13	rs1035569	0,462	0,415	0,123	0,67	0,33
		A/A	A/G	G/G	A	G
JAZF1	rs849140	0,156	0,533	0,311	0,422	0,578
		A/A	A/G	G/G	A	G
IGFBP3	rs5742694	0,061	0,363	0,575	0,243	0,757
		G/G	T/G	T/T	G	T
IGF1R	rs907806	0,783	0,203	0,014	0,884	0,116
		A/A	A/G	G/G	A	G

Таблица 3 (Продолжение)

Гены	Номер rs	Генотипы и их частоты			Аллели и их частоты	
		A/A	G/A	G/G	A	G
<i>GHS-R</i>	rs572169	A/A	G/A	G/G	A	G
		0,080	0,434	0,486	0,297	0,703
<i>CABLES1</i>	rs4800148	A/A	G/A	G/G	A	G
		0,580	0,377	0,042	0,769	0,231
<i>IFNG</i>	rs1904559	A/A	G/A	G/G	A	G
		0,028	0,231	0,741	0,144	0,856
<i>VDR3</i>	rs731236	C/C	T/C	T/T	C	T
		0,146	0,458	0,396	0,375	0,625
<i>NEGR1</i>	rs2568958	C/C	C/T	T/T	C	T
		0,104	0,467	0,429	0,337	0,663
<i>SH2B1</i>	rs7498665	A/A	A/G	G/G	A	G
		0,396	0,448	0,156	0,62	0,378
<i>GHR</i>	rs4410646	A/A	C/A	C/C	A	C
		0,783	0,212	0,005	0,889	0,111
<i>GHRL</i>	rs2075356	A/A	A/G	G/G	A	G
		0,849	0,137	0,014	0,917	0,083
<i>TNF-α</i>	rs1800629	A/A	G/A	G/G	A	G
		0,019	0,212	0,769	0,125	0,875
<i>LPIN1</i>	rs2716609	C/C	T/C	T/T	C	T
		0,028	0,255	0,717	0,156	0,844
<i>ESR1</i>	rs9340799	A/A	A/G	G/G	A	G
		0,392	0,491	0,118	0,637	0,363
<i>NRXN3</i>	rs10146997	A/A	A/G	G/G	A	G
		0,566	0,401	0,033	0,767	0,233
<i>MSRA</i>	rs545854	C/C	G/C	G/G	C	G
		0,755	0,226	0,019	0,868	0,132
<i>TFAP2B</i>	rs987237	A/A	A/G	G/G	A	G
		0,552	0,349	0,099	0,726	0,274
<i>CEL</i>	rs2013751	A/A	A/G	G/G	A	G
		0,335	0,528	0,137	0,599	0,401
<i>FAD s2/s3</i>	rs174570	C/C	C/T	T/T	C	T
		0,632	0,354	0,014	0,809	0,191
<i>APOA1/APOC3/APOA4/APOA5</i>	rs4938303	A/A	G/A	G/G	A	G
		0,453	0,434	0,113	0,67	0,33
<i>APOA1/APOC3/APOA4/APOA5</i>	rs5128	C/C	C/G	G/G	C	G
		0,792	0,189	0,019	0,887	0,113
<i>LIPC</i>	rs1800588	C/C	C/T	T/T	C	T
		0,571	0,368	0,061	0,755	0,245
<i>ABCA1</i>	rs3890182	A/A	G/A	G/G	A	G
		0,009	0,170	0,821	0,094	0,906
<i>CILP</i>	rs16996148	G/G	G/T	T/T	G	T
		0,882	0,118	0,000	0,941	0,059
<i>APOB</i>	rs6754295	A/A	A/C	C/C	A	C
		0,594	0,349	0,057	0,769	0,231
<i>SCARB1</i>	rs5888	C/C	C/T	T/T	C	T
		0,340	0,476	0,184	0,578	0,422
<i>LIPE</i>	rs34845087	C/C	C/G	G/G	C	G
		0,840	0,151	0,009	0,915	0,085
<i>PPARG</i>	rs10865710	C/C	C/G	G/G	C	G
		0,458	0,458	0,085	0,686	0,314
<i>PPARG</i>	rs1801282	C/C	G/C	G/G	C	G
		0,684	0,302	0,014	0,835	0,165

ны только для маркера rs4842838 гена *ADAMTSL3*. Данные отличия скорее всего обусловлены большим числом проанализированных жителей Европы по сравнению с нашей выборкой (http://www.ncbi.nlm.nih.gov/projects/SNP/snp_viewTable.cgi?pop=1371). Распределение соответствующих генотипов по всем генам в исследуемой группе соответствовало распределению Харди-Вайнберга ($p > 0,05$).

В настоящий момент описано более 100 генов-кандидатов, полиморфизм которых показал положительную ассоциацию с признаками ИМТ, ОТ, ХС, ХС ЛПНП и ХС ЛПОНП. Однако сравнительно мало известно о совместном вкладе генов в фенотип этих признаков. Использование традиционных методов статистики (F-критерий, Хи-квадрат, корреляционный анализ и др.) в прогнозировании количественных фенотипических признаков человека на основании исследования большого числа генов не позволяет создавать эффективные модели по их предсказанию. Это связано с тем, что большинство из них не учитывает при множественных сравнениях эпистатические взаимодействия между генами, а также не позволяет оценивать полученные генетические данные в связи с другими параметрами. Проведя анализ имеющихся ме-

тодов обработки данных для исследования количественных фенотипических показателей, нами был выбран метод логистической регрессии, эффективно работающей с непрерывными значениями (количественными признаками) (Xu et al., 2010).

При построении модели предсказания для анализируемых признаков мы использовали анкетные данные индивидов (ИМТ, ОТ), биохимические параметры (ХС, ХС ЛПНП и ХС ЛПОНП) и данные об их генотипах по 36 вышеперечисленным маркерам. Проведен анализ моделей с различными параметрами для определения наиболее точной модели, позволяющей предсказывать ряд параметров человека с учетом его генетических особенностей. В таблице 4 представлены параметры моделей, наиболее четко описывающие ряд признаков у женщин. Формулы для их расчета представлены в разделе «статистическая обработка результатов».

Известно, что эффективность данных моделей можно оценить на основании скорректированного коэффициента детерминации ($\text{adjusted } R^2$), который позволяет сравнивать между собой модели, учитывающие разное число факторов (McCullagh P., Nelder J., 1989). Коэффициент детерминации является одним из основных кри-

Таблица 4

Модели предсказания ИМТ, ОТ, уровней ХС, ХС ЛПНП и ХС ЛПОНП (до 9 генов) на основе исследуемых генетических маркеров

Параметры/ гены	Значение коэффициента	p-value	Параметры/ гены	Значение коэффициента	p-value	Параметры/ гены	Значение коэффициента	p-value
ИМТ			Общий ХС			ХС ЛПНП		
Свободный член уравнения	24,09	$<2e-16$	Свободный член уравнения	4,98	$<2e-16$	Свободный член уравнения	2,82	$<2e-16$
<i>ZBTB38</i>	0,93	0,106	<i>LCORL</i>	-0,23	0,163	<i>JAZF1</i>	0,25	0,024
<i>HNIP</i>	1,32	0,071	<i>JAZF1</i>	0,29	0,022	<i>IGF1R</i>	-0,23	0,154
<i>IGFBP3</i>	1,06	0,089	<i>GHS_R</i>	0,28	0,031	<i>GHS_R</i>	0,21	0,067
<i>IFNG</i>	2,06	0,005	<i>GHR</i>	-0,36	0,071	<i>GHR</i>	-0,46	0,008
<i>NEGR1</i>	-1,00	0,083	<i>TFAP2B</i>	0,24	0,058	<i>TFAP2B</i>	0,17	0,117
<i>NRXN3</i>	1,30	0,055	<i>APOA</i>	0,17	0,164	<i>APOA</i>	0,18	0,106
<i>CEL</i>	-1,18	0,040	<i>ABCA1</i>	-0,37	0,073			
Скорректиро- ванный коэффи- циент детерми- нации (adjusted R -squared)	0,065		Скорректиро- ванный коэффи- циент детерми- нации (adjusted R -squared)	0,065		Скорректиро- ванный коэффи- циент детерми- нации (adjusted R -squared)	0,06	
Средняя ошибка (кг/м ²)	4,1		Средняя ошибка (ммоль/литр)	0,95		Средняя ошибка (ммоль/литр)	0,83	

Таблица 4 (Продолжение)

Параметры / гены	Значение коэффициента	p-value	Параметры / гены	Значение коэффициента	p-value
ХС ЛПОНП			ОТ		
Свободный член уравнения	0,54	< 2e-16	Свободный член уравнения	78,77	< 2e-16
<i>ADAMTSL3</i>	0,06	0,072	<i>ZBTB38</i>	2,31	0,098
<i>ADAMTSL3</i>	-0,06	0,099	<i>HHIP</i>	2,99	0,090
<i>JAZF1</i>	0,05	0,083	<i>IFNG</i>	4,32	0,015
<i>IGFBP3</i>	-0,06	0,065	<i>NEGR1</i>	-2,93	0,035
<i>IGF1R</i>	-0,08	0,056	<i>LPIN1</i>	-2,73	0,116
<i>VDR3</i>	0,08	0,005	<i>NRXN3</i>	2,84	0,081
<i>GHR</i>	-0,07	0,106	<i>CEL</i>	-2,04	0,138
<i>NRXN3</i>	0,05	0,134	—	—	—
<i>TFAP2B</i>	-0,05	0,070	—	—	—
<i>CEL</i>	-0,06	0,055	—	—	—
Скорректированный коэффициент детерминации (adjusted R-squared)	0,101		Скорректированный коэффициент детерминации (adjusted R-squared)	0,059	
Средняя ошибка (ммоль/литр)	0,21		Средняя ошибка (см)	10,33	

териев, по которым проводится оценка эффективности модели. Его значение изменяется в интервале [0;1]. Данный критерий показывает, насколько полно учитываемые факторы и их комбинации, присутствующие в прогностической модели, описывают изменение значений признака. При значении коэффициента 0 — модель считается неэффективной. Основу оценки эффективности модели составляет анализ наблюдаемых и прогнозируемых значений. Расхождение между этими значениями называется ошибкой прогноза. Поскольку значение ошибки прогноза непосредственно зависит от конкретной выборки по которой строится модель (получение списка значимых параметров (из всех рассматриваемых) и значений коэффициентов для них), то и сама ошибка прогноза является случайной величиной. Для анализа степени влияния слагаемых ошибки на ее значение и применяют коэффициент детерминации R^2 , который, по сути, и отображает общий вклад параметров в изучаемый признак.

Как видно из таблицы 4, наиболее высокие показатели коэффициента детерминации получены для параметра ХС ЛПОНП ($R^2=0.101$). Более того, для этого показателя выявлена наименьшая ошибка прогноза значения ($\pm 0,21$ ммоль/литр) по сравнению с ошибкой уровней общего ХС ($\pm 0,95$ ммоль/литр) и ЛПНП ($\pm 0,83$ ммоль/литр). Полученный нами результат

свидетельствует о большей генетической детерминированности ЛПОНП по сравнению с другими показателями.

Таким образом, используя предложенную нами модель на основании анализа генов: *ZBTB38*, *HHIP*, *IGFBP3*, *IFNG*, *NEGR1*, *NRXN3*, *CEL*, у индивидов женского пола, можно предсказывать ИМТ с ошибкой $\pm 4,1$ кг/м². Модель предсказания значения объема талии включает вышеперечисленные гены для ИМТ и дополнительно ген *LPIN1* (ошибка прогноза $\pm 10,33$ см).

Модели предсказания значения уровня общего ХС строятся на основании анализа генов: *LCORL*, *JAZF1*, *GHS-R*, *GHR*, *TFAP2B*, *APOA*, *ABCA1*, с ошибкой $\pm 0,95$ ммоль/л; значения ЛПНП на основании анализа генов: *JAZF1*, *IGF1R*, *GHS-R*, *GHR*, *TFAP2B*, *APOA*, с ошибкой $\pm 0,83$ ммоль/л; значения ЛПОНП на основании анализа генов: *ADAMTSL3* (оба rs), *JAZF1*, *IGFBP3*, *IGF1R*, *VDR3*, *GHR*, *NRXN3*, *TFAP2B*, *CEL*, с ошибкой $\pm 0,21$ ммоль/л.

Важно отметить, что в моделях фигурируют гены, связь которых с анализируемыми показателями можно объяснить с точки зрения их функции, так и гены, связь которых с изучаемыми параметрами объяснить пока достаточно сложно.

Исходя из полученных результатов, регрессионные модели позволяют проводить первичную оценку ан-

Таблица 5

Значимые результаты корреляционного анализа между генами и параметрами в группе женщин

Ген	Параметр	Kendall		
		Tau	Z	p-value
Общая группа				
<i>JAZF1</i>	ЛПНП	0,104	1,915	0,055
	ЛПОНП	0,112	1,978	0,048
<i>GHS-R</i>	ХС	0,109	1,978	0,048
	ОТ	0,111	2,013	0,044 *
<i>IFNG</i>	ИМТ	0,165	2,965	0,003
	ОТ	0,14	2,494	0,013
<i>VDR3</i>	ЛПОНП	0,146	2,571	0,01
<i>NEGR1</i>	ОТ	-0,121	-2,2	0,028
<i>GHR</i>	ЛПНП	-0,127	-2,223	0,026

* несовпадения с моделью

кетных данные индивидов (ИМТ, ОТ) и биохимических параметров (ХС, ХС ЛПНП и ХС ЛПОНП) как количественных признаков. Разработанные модели возможно применять как для оценки риска развития заболевания, прогноза течения, так и для использования в криминалистике с целью первичной идентификации личности.

Для верификации данных, полученных с помощью регрессионных моделей мы использовали метод корреляционного анализа. Для изучения корреляции между генами и анализируемыми признаками был применен метод ранговой корреляции по Кендаллу (τ). Данный метод позволяет сравнивать качественные порядковые признаки между собой (генотипы) и качественные порядковые признаки (генотипы) с количественными (Реброва, 2003).

Корреляционный анализ по Кендаллу частично подтвердил ассоциацию между генотипами проанализированных генов и ИМТ, ОТ, уровнями ХС, ХС ЛПНП и ХС ЛПОНП, которые были выявлены с помощью регрессионной линейной модели (табл. 5).

Следует так же отметить, что метод линейной регрессии позволил установить ряд новых связей, не выявленных при использовании метода корреляционного анализа.

Таким образом, гены, выявленные с помощью регрессионной модели и верифицированные корреляционным анализом, безусловно участвуют в определении фенотипических и биохимических особенностей индивидуума. Для некоторых показателей (в группе

женщин) скорректированный коэффициент детерминации (adjustedR^2) более 0,1, что свидетельствует о достаточно большой роли генетической компоненты в определении изучаемых признаков. Однако еда, режим питания, физическая активность, стресс, вредные привычки, экология, лекарства — все это может сыграть большую роль в жизни и перекрыть влияние генетических факторов. Полученные результаты могут быть использованы в дальнейших исследованиях, направленных на изучение полиморфизма генов для идентификации личности с помощью ДНК-анализа, и являются важным шагом для развития прикладной генетики в судебной практике.

Работа поддержана государственным контрактом Минобрнауки России № 16.512.11.2035.

ЛИТЕРАТУРА

1. Акульченко Ю.С., 2010. Разработка и применение методов полногеномного анализа генетических ассоциаций сложных признаков: дис... д-ра биол. наук. Новосибирск. 291 с.
2. Баранов В.С., Иващенко Т.Э., Баранова Е.В., 2009. Генетический паспорт — основа индивидуальной и предиктивной медицины. СПб.: Изд-во Н-Л. С. 528.
3. Дедов И.И., Петеркова В.А., 2006. Руководство по детской эндокринологии. М.: Универсум Паблшинг. С. 600.

4. Дедов И.И., Шестакова М.В., 2006. Сахарный диабет и артериальная гипертензия. М.: Медицинское информационное агентство. 344 с.
5. Маниатис Т., Фрич Э., Сэмбрук Д., 1984. Молекулярное клонирование. М.: Мир. 480 с.
6. Подзолкова Н.М., Кузнецова И.В., Глазкова О.Л., 2006. Ожирение и репродуктивная функция женщины: Учебное пособие. М. 28 с.
7. Поляков В.К., Аверьянов А.П., Болотова Н.В., 2009. Нормативы индекса массы тела и обхвата талии: их роль в диагностике ожирения у детей школьного возраста // Педиатрия. Т. 88, № 6. С. 17–20.
8. Рахимова Г.Н., Азимова Ш.Ш., 2009. Оценка частоты метаболического синдрома среди детей и подростков с ожирением согласно новым критериям Международной Диабетической Ассоциации // Педиатрия. Т. 88, № 6. С. 14–17.
9. Реброва О.Ю., 2002, 2003, 2006. Статистический анализ медицинских данных. Применение пакета программ Statistica. М.: МедиаСфера. С. 312.
10. Ройтберг Г.Е., 2007. Метаболический синдром. М.: МЕДпресс-информ. С. 224.
11. Седлецкий Ю.И., 2007. Современные методы лечения ожирения: Руководство для врачей. СПб.: ЭЛБИ-СПб. С. 416.
12. Чазова И.У., Ильина Е.В., Терещенко С.Н., 2010. Поражение сердечно-сосудистой системы на фоне терапии лекарственными средствами, влияющими на аппетит и массу тела // Системные гипертензии. № 1. С. 47–51.
13. Friedlander Y., Austin M.A, Newman B., Edwards K. et al., 1997. Heritability of longitudinal changes in coronary-heart-disease risk factors in women twins / Am. J. Hum. Genet. N 60. P. 1502–1512.
14. McCullagh P., Nelder J., 1989. Generalized Linear Models. Second Edition. Chapman&Hall. CRC. 512 p.
15. Neel J. V., 1962. Diabetes Mellitus: A “Thrifty” Genotype Rendered Detrimental by “Progress”? // Am. J. Hum. Genet. Vol. 14, N 4. P. 353–362.
16. Pillia G., Chen W., Scuteri A., Orru M. et al., 2006. Heritability of cardiovascular and personality traits in 6,148 Sardinians // PLoS Genet. P. 2.
17. Prevention of Coronary Heart Disease in Clinical Practice. Recommendations of the Second Joint Task Force of the European and other Societies on Coronary Prevention, 1998 // Eur. Heart. J. N 19. P. 1434–1503.
18. Sing C., Davignon J., 1985. Role of the apolipoprotein E polymorphism in determining normal plasma lipid and lipoprotein variation // Am. J. Hum. Genet. Vol. 37. P. 268–285.
19. Walley A.J., Blakemore A.I., Froguel P., 2006. Genetics of obesity and the prediction of risk for health // Hum. Mol. Genet. Vol. 15, N 2. P. 124–130.
20. Xu S., Hu Z., 2010. Generalized Linear Model for Interval Mapping of Quantitative Trait Loci // Theor. Appl. Genet. Vol. 121. N 1. P. 47–63.
21. <http://www.ncbi.nlm.nih.gov/SNP>.
22. http://www.ncbi.nlm.nih.gov/projects/SNP/snp_viewTable.cgi?pop=1371.

ANALYSIS OF ASSOCIATION OF LIPID METABOLISM GENES POLYMORPHISM WITH BMI, WAIST CIRCUMFERENCE AND BLOOD LIPIDOGRAM PARAMETERS IN WOMEN

Tarkovskaya I. V., Glotov O. S., Ditkina E. Y., Vashukova E. S., Glotov A. S., Kurilov R. V., Pugacheva I. V., Belonog O. L., Makhrova I. V., Aseyev M. V., Ivashchenko T. E., Baranov V. S.

☉ **SUMMARY:** Using the PCR-RFLP method we have studied polymorphism of 36 genes involved in lipid metabolism in 212 women, residents of the North-West Region of Russia (St. Petersburg), aged 18 to 77. We found an association of polymorphism in several candidate genes with body mass index, waist circumference, total cholesterol level, low density lipoprotein cholesterol level and very low density lipoprotein cholesterol level. We propose a logistic regression model for a primary assessment of these parameters in women based on corresponding genetic markers tests.

☉ **KEY WORDS:** gene polymorphism; lipid metabolism; LDL; VLDL; candidate genes; BMI; WC; correlation; regression.

☉ Информация об авторах

Тарковская Ирина Васильевна — лаборант. ООО «БиоГлот». 197341, Санкт-Петербург, Коломяжский пр., д. 28, к. 2, пом. 31-Н, лит. А.

Глотов Олег Сергеевич — к. б. н., с. н. с., н. с., Федеральное бюджетное учреждение Российской академии медицинских наук Научно-исследовательский институт акушерства и гинекологии им. Д. О. Отта Северо-Западного отделения РАМН, лаборатория пренатальной диагностики врожденных и наследственных заболеваний. 199034, Санкт-Петербург, Менделеевская линия, д. 3. ООО «БиоГлот». 197341, Санкт-Петербург, Коломяжский пр., д. 28, к. 2, пом. 31-Н, лит. А. E-mail: olgotov@mail.ru.

Tarkovskaya Irina Vasilyevna — laboratory, LLC BioGlot. Kolomyazhskiy Ave., 28, R. 2, of. 31-H, lit. A, Saint-Petersburg, 197341, Russia.

Glotov Oleg Sergeevich — candidate of biological scientist, senior scientist, laboratory of prenatal diagnosis of congenital and inherited diseases, D. O. Ott Research Institute of Obstetrics and Gynecology, RAMS. Mendeleev Line, 3, Saint-Petersburg, 199034, Russia. Scientist of LLC BioGlot. Kolomyazhskiy Ave., 28, R. 2, of. 31-H, lit. A, Saint-Petersburg, 197341, Russia. E-mail: olgotov@mail.ru.

✉ Информация об авторах

- Диткина Екатерина Юрьевна** — н.с., ООО «БиоГлот». 197341, Санкт-Петербург, Коломяжский пр., д. 28, к. 2, пом. 31-Н, лит. А.
- Вашукова Елена Сергеевна** — биолог, н.с., Федеральное бюджетное учреждение Российской академии медицинских наук Научно-исследовательский институт акушерства и гинекологии им. Д.О. Отта Северо-Западного отделения РАМН, лаборатория пренатальной диагностики врожденных и наследственных заболеваний. 199034, Санкт-Петербург, Менделеевская линия, д. 3. ООО «БиоГлот». 197341, Санкт-Петербург, Коломяжский пр., д. 28, к. 2, пом. 31-Н, лит. А. E-mail: iagmail@ott.ru.
- Глов Андрей Сергеевич** — к.б.н., с.н.с., н.с., Федеральное бюджетное учреждение Российской академии медицинских наук Научно-исследовательский институт акушерства и гинекологии им. Д.О. Отта Северо-Западного отделения РАМН, лаборатория пренатальной диагностики врожденных и наследственных заболеваний. 199034, СПб, Менделеевская линия, д. 3. ООО «БиоГлот». 197341, СПб, Коломяжский пр., д. 28, к. 2, пом. 31-Н, лит. А. E-mail: iagmail@ott.ru.
- Курилов Роман Владимирович** — н.с., ООО «БиоГлот». 197341, СПб, Коломяжский пр., д. 28, к. 2, пом. 31-Н, лит. А.
- Пугачева Ирина Владимировна** — н.с., ООО «БиоГлот». 197341, СПб, Коломяжский пр., д. 28, к. 2, пом. 31-Н, лит. А.
- Белоног Ольга Львовна** — н.с., ООО «БиоГлот». 197341, СПб, Коломяжский пр., д. 28, к. 2, пом. 31-Н, лит. А.
- Махрова Ирина Александровна** — н.с., к.м.н., ООО «БиоГлот». 197341, СПб, Коломяжский пр., д. 28, к. 2, пом. 31-Н, лит. А.
- Асеев Михаил Владимирович** — к.б.н., с.н.с., н.с., Федеральное бюджетное учреждение Российской академии медицинских наук Научно-исследовательский институт акушерства и гинекологии им. Д.О. Отта Северо-Западного отделения РАМН, лаборатория пренатальной диагностики врожденных и наследственных заболеваний. 199034, СПб, Менделеевская линия, д. 3. ООО «БиоГлот». 197341, СПб, Коломяжский пр., д. 28, к. 2, пом. 31-Н, лит. А. E-mail: iagmail@ott.ru.
- Ивашченко Татьяна Эдуардовна** — д.б.н., в.н.с., н.с., профессор, Федеральное бюджетное учреждение Российской академии медицинских наук Научно-исследовательский институт акушерства и гинекологии им. Д.О. Отта Северо-Западного отделения РАМН, лаборатория пренатальной диагностики врожденных и наследственных заболеваний. 199034, СПб, Менделеевская линия, д. 3. ООО «БиоГлот». 197341, СПб, Коломяжский пр., д. 28, к. 2, пом. 31-Н, лит. А. E-mail: iagmail@ott.ru.
- Баранов Владислав Сергеевич** — д.м.н., член-корр. РАМН, профессор, зав. лаб., Федеральное бюджетное учреждение Российской академии медицинских наук Научно-исследовательский институт акушерства и гинекологии им. Д.О. Отта Северо-Западного отделения РАМН, лаборатория пренатальной диагностики врожденных и наследственных заболеваний. 199034, СПб, Менделеевская линия, д. 3. E-mail: iagmail@ott.ru.
- Ditkina Ekaterina Yuryevna** — research, LLC BioGlot. Kolomyazhskiy Ave., 28, R. 2, of. 31-H, lit. A, Saint-Petersburg, 197341, Russia.
- Vashukova Elena Sergeevna** — biolog, laboratory of prenatal diagnosis of congenital and inherited diseases, D.O. Ott Research Institute of Obstetrics and Gynecology, RAMS. Mendeleev Line, 3, Saint-Petersburg, 199034, Russia. Scientist of LLC BioGlot. Kolomyazhskiy Ave., 28, R. 2, of. 31-H, lit. A, Saint-Petersburg, 197341, Russia. E-mail: iagmail@ott.ru.
- Glotov Andrey Sergeevich** — candidate of biological scientist, senior scientist, laboratory of prenatal diagnosis of congenital and inherited diseases, D.O. Ott Research Institute of Obstetrics and Gynecology, RAMS. Mendeleev Line, 3, Saint-Petersburg, 199034, Russia. Scientist of LLC BioGlot. Kolomyazhskiy Ave., 28, R. 2, of. 31-H, lit. A, Saint-Petersburg, 197341, Russia. E-mail: iagmail@ott.ru.
- Kurilov Roman Vladimirovich** — research, LLC BioGlot. Kolomyazhskiy Ave., 28, R. 2, of. 31-H, lit. A, Saint-Petersburg, 197341, Russia.
- Pugacheva Irina Vladimirovna** — researcher, LLC BioGlot. Kolomyazhskiy Ave., 28, R. 2, of. 31-H, lit. A, Saint-Petersburg, 197341, Russia.
- Belonog Olga Lvovna** — research, LLC BioGlot. Kolomyazhskiy Ave., 28, R. 2, of. 31-H, lit. A, Saint-Petersburg, 197341, Russia.
- Makhrova Irina Aleksandrovna** — Ph.D, research, LLC BioGlot. Kolomyazhskiy Ave., 28, R. 2, of. 31-H, lit. A, Saint-Petersburg, 197341, Russia.
- Aseyev Mikhail Vladimirovich** — candidate of biological scientist, senior scientist, laboratory of prenatal diagnosis of congenital and inherited diseases, D.O. Ott Research Institute of Obstetrics and Gynecology, RAMS. Mendeleev Line, 3, Saint-Petersburg, 199034, Russia. Scientist of LLC BioGlot. Kolomyazhskiy Ave., 28, R. 2, of. 31-H, lit. A, Saint-Petersburg, 197341, Russia. E-mail: iagmail@ott.ru.
- Ivashchenko Tatyana Eduardovna** — professor, doctor of biological sciences, laboratory of prenatal diagnosis of congenital and inherited diseases, D.O. Ott Research Institute of Obstetrics and Gynecology, RAMS. Mendeleev Line, 3, Saint-Petersburg, 199034, Russia. Senior scientist of of LLC BioGlot. Kolomyazhskiy Ave., 28, R. 2, of. 31-H, lit. A, Saint-Petersburg, 197341, Russia. E-mail: iagmail@ott.ru.
- Baranov Vladislav Sergeevich** — professor, doctor of biological sciences, head of the laboratory of prenatal diagnosis of congenital and inherited diseases. D.O. Ott Research Institute of Obstetrics and Gynecology, RAMS. Mendeleev Line, 3, Saint-Petersburg, 199034, Russia. E-mail: iagmail@ott.ru.