

ГЕНОМИКА.
ТРАНСКРИПТОМИКА

УДК 575.174:616-056.52

РОЛЬ ЕСТЕСТВЕННОГО ОТБОРА В ФОРМИРОВАНИИ ГЕНЕТИЧЕСКОЙ
СТРУКТУРЫ ПОПУЛЯЦИЙ ПО SNP-МАРКЕРАМ, СВЯЗАННЫМ
С ИНДЕКСОМ МАССЫ ТЕЛА И ОЖИРЕНИЕМ¹

© 2020 г. Е. А. Трифонова^{а, *}, А. А. Попович^а, А. В. Бочарова^а, К. В. Вагайцева^а, В. А. Степанов^а

^аНаучно-исследовательский институт медицинской генетики Томского национально-исследовательского центра
Российской академии наук, Томск, 634050 Россия

*e-mail: ekaterina.trifonova@medgenetics.ru

Поступила в редакцию 08.09.2019 г.

После доработки 23.12.2019 г.

Принята к публикации 24.12.2019 г.

Ожирение становится одной из глобальных проблем — более трети мирового населения имеют избыточный вес. Эта патология существенно влияет на качество жизни и является фактором риска сердечно-сосудистых заболеваний, сахарного диабета, рака и нарушений репродуктивной функции. Нами изучена популяционная вариабельность частот аллелей 26 однонуклеотидных полиморфных маркеров, ассоциированных с ожирением и индексом массы тела по данным полногеномных ассоциативных исследований (GWAS). Генетическая изменчивость проанализирована в популяциях Северной Евразии и популяциях из проекта по изучению разнообразия генома человека (HGDP). Исследованные этнические группы характеризуются высоким уровнем генетического разнообразия, которое коррелирует с климато-географическими параметрами. Результаты тестов на поиск сигналов естественного отбора свидетельствуют о действии отбора на маркеры rs1167827 гена *HIP1*, rs7138803 и rs7164727, расположенные в межгенной области, rs7141420 гена *NRXN3*, rs7498665 гена *SH2B1*, rs7903146 гена *TCF7L2*.

Ключевые слова: популяции человека, генетическое разнообразие, ожирение, индекс массы тела, естественный отбор

DOI: 10.31857/S0026898420030179

Ожирение — одна из актуальных проблем современного мира. Распространенность этой патологии в развитых и развивающихся странах в последнее время резко увеличивается, вызывая значительные социально-экономические последствия. Избыточный вес сокращает продолжительность жизни человека и считается фактором риска сердечно-сосудистых заболеваний, некоторых видов рака, сахарного диабета второго типа, депрессии, заболеваний опорно-двигательного аппарата. Основной патогенетической причиной ожирения и избыточного веса считается нарушение энергетического баланса между расходуемыми и потребляемыми калориями [1]. Наряду с этим показана значимость генетических факторов в индивидуальной предрасположенности к избыточному накоплению жировой ткани. Вклад генетической компоненты в развитие ожирения подтверждается близнецовыми исследованиями, в которых степень влияния наследственности на увеличение индекса массы

тела (ИМТ) составляет 40–70% [2, 3]. Кроме того, в рамках полногеномного анализа ассоциаций (GWAS) обнаружено множество однонуклеотидных полиморфизмов (SNP), связанных с ожирением и с такими параметрами, как ИМТ, объем талии и бедер [4].

Одно из перспективных направлений в изучении наследственных причин заболеваний — эволюционно-генетический подход, способствующий нахождению факторов, влияющих на распространение исследуемого фенотипа в популяциях человека. Этот подход тесно связан с такими представлениями, как гипотезы “экономных генотипов” и “дрейфующих генов”, а также с рядом других эволюционно-генетических концепций [5, 6]. В частности, гипотеза “экономных генотипов” предполагает, что распространенность ожирения в современных популяциях обусловлена направленным отбором аллелей, ассоциированных со способностью “запасать” углеводы, что благоприятствовало выживанию в период голода. Гипотеза “дрейфующих генов” утверждает, что частоты предковых аллелей “дрейфовали” в древних популяциях, а в свя-

¹ Дополнительная информация для этой статьи доступна по doi 10.31857/S0026898420030179 для авторизованных пользователей.

зи с процессами расселения и миграций в условиях новых средовых факторов приобретали новое негативное влияние – приводили к развитию ожирения. Роль межпопуляционных различий в климато-географических факторах как основы различий в процессах метаболизма и обеспечения энергетического баланса подчеркивается и современными эпидемиологическими концепциями [7].

В контексте популяционной дифференциации генетических и средовых факторов ожирения значительный интерес представляют коренные популяции Северной Евразии, в частности Сибири и Дальнего Востока, проживающие на территории с умеренным и арктическим климатом. Эволюционная изменчивость в популяциях Северной Евразии привела, по-видимому, к формированию особой структуры генофонда, специфических частот распространенности заболеваний, включая ожирение, и ассоциированных с ними SNP-маркеров. Однако роль адаптивной эволюции и действия естественного отбора в формировании генетической составляющей популяций человека, проживающих в Северной Евразии, в настоящее время изучена недостаточно. В этой работе в рамках эволюционно-генетического подхода исследованы генетические маркеры, ассоциированные с ожирением и ИМТ в популяциях Северной Евразии и в популяциях из проекта по изучению разнообразия генома человека (HGDP).

ЭКСПЕРИМЕНТАЛЬНАЯ ЧАСТЬ

Материал для исследования представлен популяционными выборками из 14 этнических групп, проживающих на территории России и ближнего зарубежья (русские, карелы, вепсы, удмурты, коми, узбеки, киргизы, ханты, кеты, тувинцы, буряты, якуты, нивхи и коряки), взятыми из биобанка НИИ медицинской генетики Томского НИМЦ. В состав выборок вошли неродственные и неметисированные индивиды в трех поколениях. Также в работу включены данные по 23 популяциям, полученным из проекта HGDP (Human Genome Diversity Project): манденка, мозабиты, пигмеи Биака, друзы, палестинцы, бедуины, сардинцы, баски, французы, синдхи, брахуи, белужди, макраны, калашаи, хазарейцы, бурушо, патаны, майя, пима, каритиана, папуасы, суруи, меланезийцы. Численность общей выборки составила 1672 человека.

Изучены 26 SNP-маркеров, которые в двух и более исследованиях показали ассоциацию с ожирением и показателем ИМТ согласно базе данных “GWAS catalog” [4]. Краткая характеристика изученных полиморфных вариантов представлена в табл. 1. Необходимо отметить, что большинство из SNP расположены в некодирующих участках генома.

Мультиплексное генотипирование полиморфных генетических маркеров проведено методом масс-спектрометрии MALDI-TOF, как описано ранее [25]. Праймеры произведены компанией “Евроген” (Россия). Последовательность праймеров доступна по запросу у авторов.

Соответствие частот аллелей и генотипов равновесию Харди–Вайнберга оценивали с помощью критерия χ^2 . Степень связи ожидаемой гетерозиготности и частот аллелей по изученным SNP с распространенностью ожирения среди взрослого населения и климато-географическими факторами определяли с помощью коэффициента ранговой корреляции Спирмена. Климато-географические показатели (абсолютная широта, абсолютная долгота, среднегодовая температура, температура наиболее холодного месяца, температура наиболее теплого месяца, разброс температур, среднегодовой уровень осадков, средняя относительная влажность) выбраны из базы данных Weatherbase [26]. Данные о распространенности ожирения на территории проживания исследованных популяций получены из статистического сборника 2016 года Министерства здравоохранения Российской Федерации (русские, карелы, вепсы, удмурты, коми, ханты, кеты, тувинцы, буряты, якуты, нивхи и коряки) и данных Всемирной организации здравоохранения (WHO) (популяционные выборки из проекта HGDP, узбеки и киргизы) [27, 28]. Для выявления сигналов естественного отбора по SNP-маркерам использовали тесты F_{DIST} и Юинса–Ваттерсона в программном пакете Arlequin [29]. Функциональную аннотацию генов проводили в онлайн-ресурсах “STRING” и “WebGestalt” [30, 31].

РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

Генетическое разнообразие в популяциях человека по исследованным генетическим маркерам

Значения частот предковых аллелей в популяциях представлены в табл. 1S Приложения. Большинство генетических маркеров оказались полиморфными, кроме локуса rs13107325 в нескольких популяциях. Для этого локуса характерны высокие значения частоты предкового аллеля С в популяционных выборках (от 0.91 у коми до 1.00 у карелов, якутов и нивхов). В целом, популяции Северной Евразии характеризуются существенной вариабельностью по исследованным SNP. Анализ соответствия генотипов равновесию Харди–Вайнберга показал отклонение в 28 случаях (7.69%) из представленных 364. Не обнаружено накопления числа отклонений по отдельным популяциям или маркерам.

Анализ генетических взаимоотношений 14 популяций Северной Евразии и 23 популяций из проекта HGDP методом главных компонент по-

Таблица 1. Характеристика использованных полиморфных генетических маркеров

№	Полиморфный маркер	Ген	Локализация в гене	Аллели*	Фенотип	Источник
1	rs10182181	<i>ADCY3, DNAJC27</i>	Межгенный	A/G	Ожирение ИМТ	8–12
2	rs1167827	<i>HIP1</i>	3'-UTR	A/G	ИМТ	9, 11–13
3	rs12429545	<i>ZNF646P1, LOC105370210</i>	Межгенный	A/G	ИМТ	9–14
4	rs12446632	<i>LOC105371116</i>	Инtron	A/G	Ожирение ИМТ	8–11, 13
5	rs12940622	<i>RPTOR</i>	Инtron	A/G	ИМТ	9, 11
6	rs13107325	<i>SLC39A8</i>	Экзон	A/C/T	ИМТ	9–11, 13, 15
7	rs1514175	<i>LRRC53, FPGT-TNNI3K, TNNI3K</i>	Инtron	C/T	ИМТ	15, 16
8	rs2033529	<i>TDRG1, LRFN2</i>	Межгенный	C/G/T	ИМТ	9, 12, 13
9	rs2531995	<i>ADCY9</i>	3'-UTR	A/G	Ожирение ИМТ	8, 13, 17
10	rs2568958	<i>LOC105378797</i>	Инtron	A/C/G	ИМТ	8, 18
11	rs3101336	<i>LOC105378797</i>	Инtron	A/G	Ожирение ИМТ	8–12, 19
12	rs571312	<i>RNU4-17P, LOC342784</i>	Межгенный	G/T	ИМТ	13, 15, 19
13	rs633715	<i>LINC01741, SEC16B</i>	Межгенный	C/T	ИМТ	8, 12, 13
14	rs6804842	<i>RARB</i>	Инtron	A/G	ИМТ	9, 11, 13
15	rs6864049	<i>LOC105379158</i>	Инtron	A/C/G/T	ИМТ	9, 12, 13
16	rs7138803	<i>BCDIN3D, RPL35AP28</i>	Межгенный	A/G/T	ИМТ	8–13, 15, 18
17	rs7141420	<i>NRXN3</i>	Инtron	C/T	Ожирение ИМТ	8, 9, 11–13
18	rs7164727	<i>ADPGK-AS1, LOC107984801</i>	Межгенный	C/G/T	ИМТ	9, 11, 13
19	rs7195386	<i>RBBP6</i>	Инtron	A/C/T	Ожирение ИМТ	12, 20
20	rs7498665	<i>SH2B1</i>	Экзон	A/G/T	Ожирение ИМТ	8, 18, 21
21	rs7531118	<i>LOC105378797</i>	Инtron	C/T	Ожирение ИМТ	8, 13
22	rs7647305	<i>ETV5, LOC105374257</i>	Межгенный	C/T	ИМТ	13, 18, 22
23	rs7903146	<i>TCF7L2</i>	Инtron	C/G/T	ИМТ	9, 11–13
24	rs8050136	<i>FTO</i>	Инtron	A/C	ИМТ	13, 18, 23
25	rs987237	<i>TFAP2B</i>	Инtron	A/G	ИМТ	8, 10, 13, 15
26	rs9941349	<i>FTO</i>	Инtron	C/T	Ожирение ИМТ	13, 24

Примечание. ИМТ – индекс массы тела.

* Предковый аллель выделен жирным шрифтом.

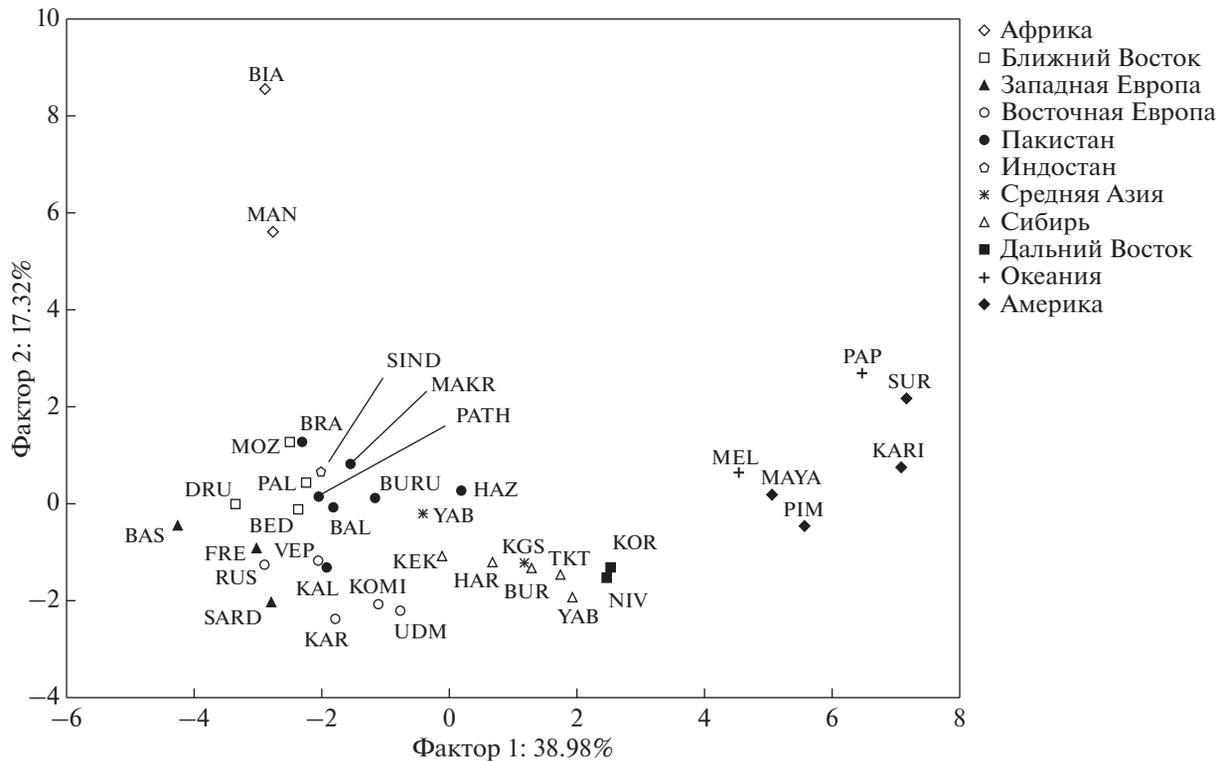


Рис. 1. Положение 37 популяций человека в пространстве главных компонент по частотам аллелей 26 SNP. Примечание: KAR – карелы, VEP – вепсы, UDM – удмурты, RUS – русские, KOMI – коми, UZB – узбеки, KGS – киргизы, HAR – ханты, KEK – кеты, TKT – тувинцы, BUR – буряты, YAB – якуты, NIV – нивхи, KOR – коряки, BAL – белуджи, BED – бедуины, BIA – пигмеи Биака, BRA – брахуи, BURU – бурушо, DRU – друзы, FRE – французы, BAS – баски, HAZ – хазарейцы, KAL – калашы, KARI – каритиана, MAKR – макраны, MAN – манденка, MAYA – майя, MOZ – мозабиты, MEL – меланезийцы, PAL – палестинцы, PAP – папуасы, PATH – патаны, PIM – пима, SARD – сардинцы, SIND – синдхи, SUR – суруи.

казал, что первая и вторая главные компоненты объясняют 56.3% общей изменчивости аллельных частот изученных SNP (рис. 1). Расположение популяций в пространстве главных компонент отражает их географическую локализацию: популяции Африки дистанцированы от населения других регионов мира, как и популяции Океании и Америки, которые расположены в районе наиболее высоких значений первой главной компоненты.

Обнаружена положительная статистически значимая корреляция первой главной компоненты аллельных частот с абсолютной долготой ($R = 0.916$, $p = 0.000$). Вторая главная компонента значимо коррелирует с абсолютной широтой ($R = -0.852$, $p = 0.000$) и несколькими климатическими параметрами: разброс температур ($R = -0.731$, $p = 0.000$), температура наиболее теплого ($R = 0.627$, $p = 0.000$) и холодного месяцев ($R = -0.810$, $p = 0.000$) и среднегодовая температура ($R = 0.772$, $p = 0.000$). Таким образом, первая и вторая главные компоненты могут отражать вклад нескольких климато-географических переменных в распределение частот аллелей в популяциях.

Корреляции частот аллелей и средней ожидаемой гетерозиготности по исследованным SNP с географическими и климатическими параметрам

Корреляционный анализ выявил статистически значимую связь частот предковых аллелей нескольких полиморфных вариантов с абсолютной широтой (10 SNP), абсолютной долготой (22 SNP), среднегодовой температурой (9 SNP), температурой наиболее холодного (11 SNP) и теплого месяца (8 SNP), разбросом температур (9 SNP), среднегодовым уровнем осадков (11 SNP), средней относительной влажностью (9 SNP), с распространенностью ожирения (7 SNP) (табл. 2S Приложения). Следует отметить, что из 26 SNP только локус rs1167827 гена *HIP1* показал корреляцию со всеми исследованными климатическими и географическими параметрами.

Кроме того, показана связь средней ожидаемой гетерозиготности (H_e) с абсолютной долготой ($R = -0.544$; $p = 0.001$), средней относительной влажностью ($R = -0.401$; $p = 0.014$) и среднегодовым уровнем осадков ($R = -0.580$; $p = 0.000$) (рис. 2). Значение гетерозиготности уменьшается с увеличением каждой из этих климато-географи-

ческих переменных. Постепенное снижение H_e26 по мере расселения популяций человека по территории земного шара (корреляция с долготой) можно рассматривать как закономерный результат действия совокупности микроэволюционных факторов в ходе миграции современного человека с доминирующим эффектом нейтральных механизмов — миграция, дрейф генов и изоляция расстоянием. В то же время, ранее на двух других системах генетических маркеров, ассоциированных с иммунозависимыми фенотипами и устойчивостью к холоду, нами была выявлена противоположная тенденция — рост генетического разнообразия популяций по мере отдаления от Экватора, обусловленный, вероятно, адаптивной эволюцией, опосредованной естественным отбором [32–34].

Поиск действия естественного отбора в исследованных популяционных выборках

Для поиска сигналов естественного отбора мы использовали два теста (FDIST и тест Юинса–Ваттерсона), позволяющих оценить отклонения от селективной нейтральности для отдельных маркеров. Тест FDIST основан на оценке генетической дифференциации (F_{ST}) популяционных выборок. Высокий уровень F_{ST} свидетельствует о значительных генетических различиях популяций, что может отражать действие направленного отбора по SNP в исследованных популяциях с разной интенсивностью, низкий уровень F_{ST} — о низком межпопуляционном генетическом разнообразии и влиянии балансирующего отбора. Результаты теста приведены на рис. 3. Тест Юинса–Ваттерсона характеризует действие направленного отбора, если значение наблюдаемой гомозиготности (F_o) статистически значимо выше, чем в пороговой доле случайных распределений (F_e). Если значение F_o меньше F_e , то SNP может находиться под действием балансирующего отбора (табл. 3S и 4S Приложения).

Обнаружено влияние (тест FDIST) направленного естественного отбора на rs1167827 гена *HIP1* ($F_{ST} = 0.252$; $p = 0.000$), маркера rs7164727, находящегося в межгенном участке ($F_{ST} = 0.130$; $p = 0.038$), rs7498665 гена *SH2B1* ($F_{ST} = 0.133$; $p = 0.029$) и rs7903146 гена *TCF7L2* ($F_{ST} = 0.127$; $p = 0.047$), балансирующего отбора — на rs7138803, расположенного в межгенной области ($F_{ST} = 0.046$; $p = 0.047$) и rs7141420 гена *NRXN3* ($F_{ST} = 0.046$; $p = 0.041$). Согласно результатам теста Юинса–Ваттерсона эти же локусы показали отклонение от селективной нейтральности в нескольких изученных популяциях: rs1167827 у франкозюзов, хазарейцев и сардинцев; rs7164727 у коми, киргизов, белуджи, пигмеев Биака и брахуев; rs7498665 у вепсов, удмуртов, киргизов, хантов, коряков и пигмеев Биака;

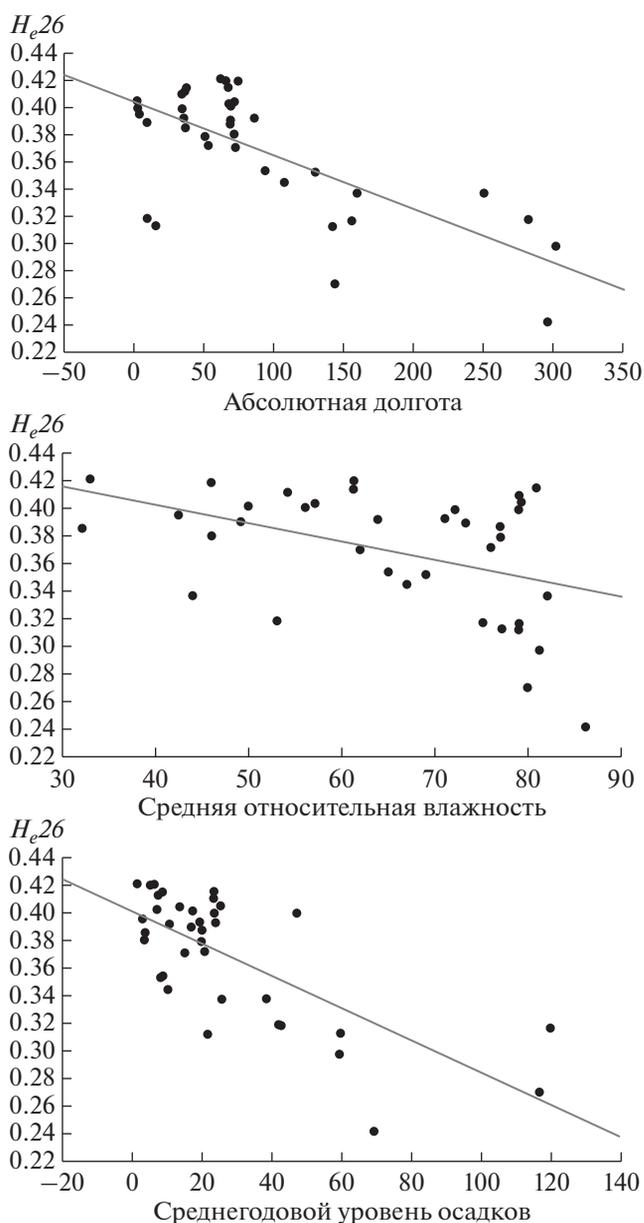


Рис. 2. Связь средней ожидаемой гетерозиготности по 26 SNP с климато-географическими переменными. Ось абсцисс — географический (климатический) параметр. Ось ординат — средняя наблюдаемая гетерозиготность в популяциях.

rs7903146 у брахуев, мозабитов и палестинцев; rs7138803 у удмуртов, киргизов, белуджи, бедуинов, брахуев и синдхов; rs7141420 у карелов, вепсов, коми, узбеков, кетов, тувинцев, бурушо, хазарейцев, палестинцев, сардинцев и синдхов (табл. 3S и 4S Приложения). Кроме того, в соответствии с результатами данного теста, накопление большого числа значимых отклонений от нейтральности обнаруживается в изученных популяциях по локусу rs10182181, локализованному в межгенном регионе (вепсы, удмурты, русские,

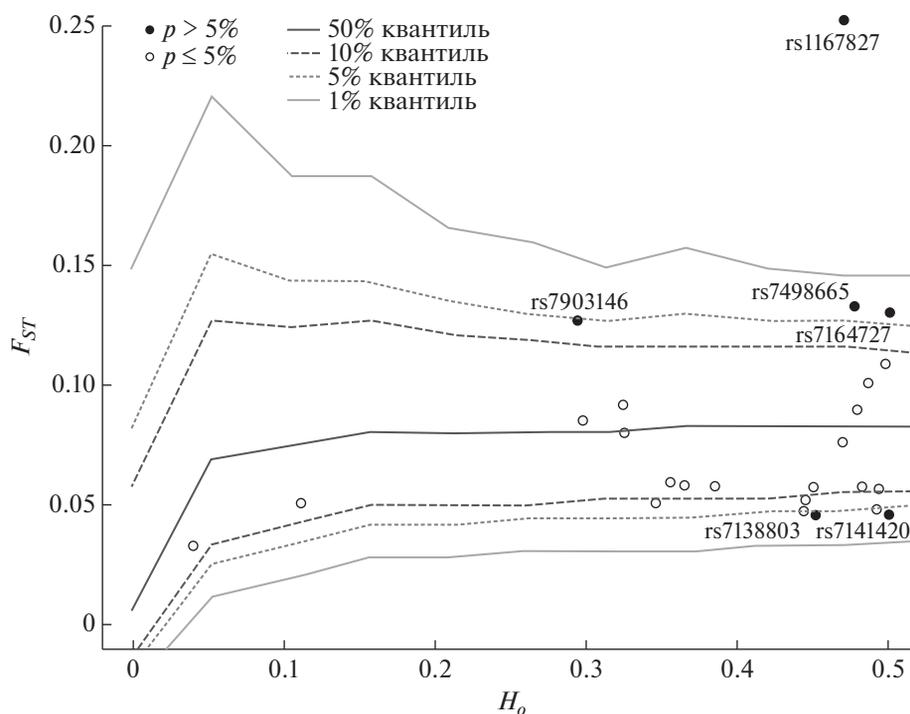


Рис. 3. Зависимость генетической дифференциации популяций от наблюдаемой гетерозиготности (тест FDI_{ST}). Расположение маркеров за пределами 95%-го интервала ожидаемого распределения свидетельствует о значимом действии направленного отбора (при высоких значениях F_{ST}) или балансирующего отбора (при низких значениях F_{ST}). F_{ST} – генетическая дифференциация популяций, H_o – наблюдаемая гетерозиготность.

бедуины, брахуи, бурушо, друзы, французы, хазарейцы, калашы, макраны, патаны и синдхи), маркерам rs2531995 гена *ADCY9* (коми, киргизы, ханты, кеты, буряты, белуджи, друзы, баски, макраны и мозабиты) и rs6864049 гена *LOC105379158* (карелы, вепсы, удмурты, русские, коми, белуджи, бедуины, брахуи, друзы, французы, калашы и макраны).

Следует отметить, что направленный отбор приводит к высокому межпопуляционному генетическому разнообразию и гомозиготности внутри популяции. Естественный отбор, действующий против распространенного аллеля, способствует уменьшению гомозиготности и дифференциации популяций. Следовательно, действие направленного отбора против наиболее распространенного аллеля в популяции может проявляться в виде сигнала балансирующего отбора.

В целом, более 80% изученных SNP (21 из 26) отклоняются от гипотезы селективной нейтральности согласно результатам теста Юинса–Ватерсона (рис. 4). Интересно, что в представленной работе показана корреляция частот аллелей этих генетических маркеров как минимум с одной из исследованных переменных – климато-географических факторами и/или показателем распространенности ожирения на территории

проживания популяций (см. табл. 2S Приложения). С помощью онлайн-ресурсов “STRING” и “WebGestalt” проведена функциональная аннотация генов, в которых локализованы данные 21 SNP. В случае маркера, расположенного в межгенном участке, выбирали два гена, между которыми находится исследуемый полиморфный вариант. Таким образом, в анализ функциональной аннотации вошли 24 гена (*ADCY3*, *ADCY9*, *ADPGK-AS1*, *BCDIN3D*, *DNAJC27*, *FPGT-TNNI3K*, *FTO*, *HIP1*, *LOC105370210*, *LOC105378797*, *LOC105379158*, *LOC107984801*, *LOC342784*, *LRRC53*, *NRXN3*, *RARB*, *RBBP6*, *RNU4-17P*, *RPL35AP28*, *RPTOR*, *SH2B1*, *TCF7L2*, *TNNI3K*, *ZNF646P1*).

Анализ с помощью веб-ресурса “STRING” выявил кластер из 9 функционально взаимодействующих белков (рис. 5), связи между которыми основаны на опубликованных данных (характерно для всех белков, представленных в кластере), коэкспрессии (*ADCY9–ADCY3–DNAJC27–NRXN3–TFO–DNAJC27*, *NRXN3–TNNI3K–ADCY3* и *SH2B1–BCDIN3D*), результатах экспериментальных работ (*ADCY3–DNAJC27–TNNI3K–SH2B1*), гомологии белков (*ADCY9–ADCY3*) и анализе баз данных (*ADCY9–ADCY3*).

Анализ с помощью онлайн-ресурса “WebGestalt” выявил 17 биологических процессов и моле-

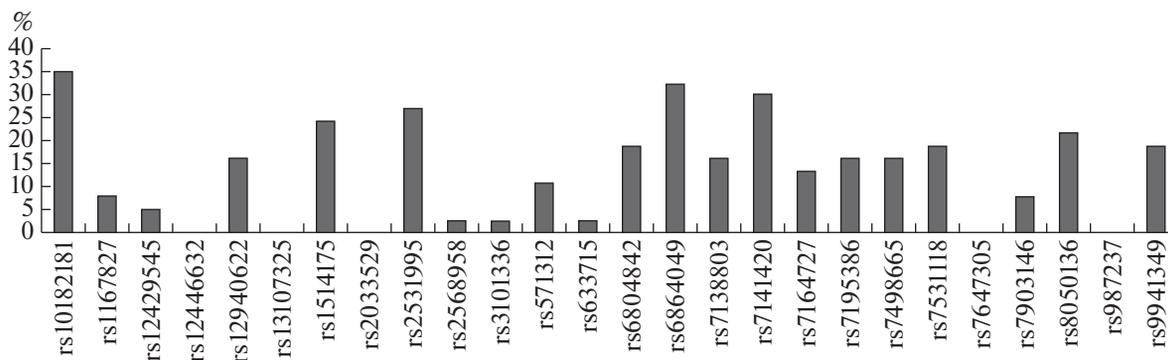


Рис. 4. Доля изученных популяций, в которых зафиксировано отклонение от селективной нейтральности (согласно данным теста Юинса–Ваттерсона).

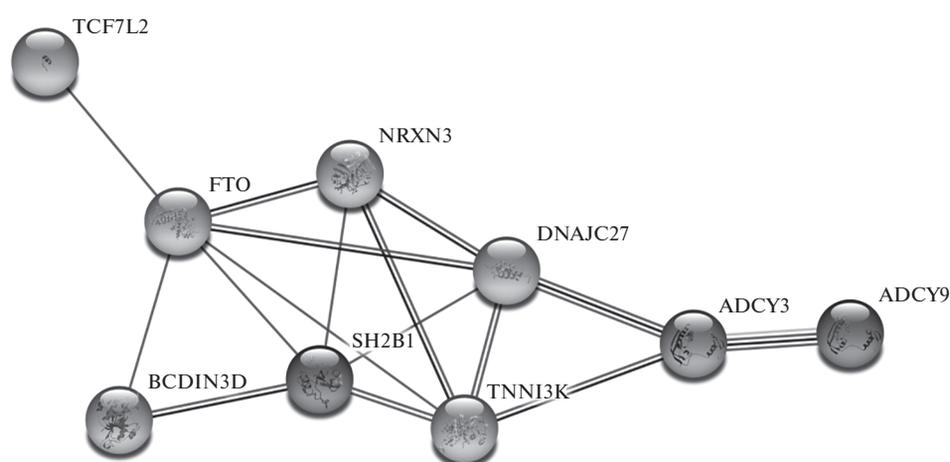


Рис. 5. Белок-белковые взаимодействия продуктов исследуемых генов.

кулярных функций из проекта “Gene Ontology”, в которые вовлечены не менее двух исследуемых генов (табл. 2). Необходимо отметить, что гены *ADCY3* и *ADCY9*, продукты которых участвуют в наиболее сильных белок–белковых взаимодействиях, согласно базе данных “STRING” вовлечены в большинство биологических процессов, ключевые из которых – биосинтез и метаболизм нуклеотидов, водный гомеостаз, ответ на глюкагон. Вероятно, это связано с тем, что белковые продукты этих генов, обладают сходной пространственной структурой, поскольку являются изформами аденилатциклазы, катализирующей превращение АТФ в сАМР [35]. Наибольшее число генов (*HIP1*, *RPTOR*, *SH2B1*, *TCF7L2*) включено в такие категории, как связывание протеинкиназ и киназ (GO: 0019901, GO: 0019900).

Изменения в функционировании описанных биологических процессов (биосинтез и метаболизм нуклеотидов, водный гомеостаз, ответ на глюкагон, связывание протеинкиназ и киназ) могут приводить к различным патологическим состояниям, а также влиять на адаптацию популя-

ций к условиям среды обитания. В частности, различные модификации нуклеотидной последовательности изменяют структуру кодируемых ими белков, что, в свою очередь, может приводить к вариабельности их функциональной активности и физико-химических свойств. Нарушение биосинтеза циклических нуклеотидов – ключевых внутриклеточных медиаторов, может быть связано с молекулярными механизмами адаптации к гипоксии и обеспечением реализации комплекса явлений, называемых “постгипоксической активацией” [36]. Водный гомеостаз, поддержание которого является важнейшей функцией лимфатической системы, необходим для жизнедеятельности организма, определяя его способность адаптироваться к различным средам [37]. Инсулинорезистентность, вызванная глюкагоном, одним из основных гормонов человека, отвечающих за резистентность к инсулину, может провоцировать развитие сахарного диабета второго типа и сердечно-сосудистых заболеваний [38]. Кроме того, известно, что глюкагон участвует в стресс-реакции, в ходе которой секреция глюкагона усиливается под

Таблица 2. Функциональная аннотация исследуемых генов, SNP которых показали статистически значимые результаты

Категория “Gene Ontology”	Процесс/функция	Ген	<i>p</i>
Биологические процессы			
GO: 0006171	Реакции, связанные с активностью аденилатциклазы	<i>ADCY3, ADCY9</i>	0.0000
GO:0009190	Биосинтез циклических нуклеотидов	<i>ADCY3, ADCY9</i>	0.0002
GO: 0046058	Метаболизм сАМР	<i>ADCY3, ADCY9</i>	0.0002
GO: 0052652	Реакции с участием циклических пуриновых нуклеотидов	<i>ADCY3, ADCY9</i>	0.0002
GO: 0035264	Рост многоклеточного организма	<i>FTO, RARB, RBBP6</i>	0.0002
GO: 0071377	Клеточный ответ на стимуляцию глюкагоном	<i>ADCY3, ADCY9</i>	0.0002
GO: 0003091	Процессы в тканях почек, участвующие в поддержании водного гомеостаза организма	<i>ADCY3, ADCY9</i>	0.0003
GO: 0033762	Процесс, связанный с изменением состояния или активности клетки и/или организма, в результате стимуляции глюкагоном	<i>ADCY3, ADCY9</i>	0.0003
GO: 0046825	Регуляция экспорта белка из ядра	<i>DNAJC27, TCF7L2</i>	0.0004
Молекулярные функции			
GO: 0004016	Активность аденилатциклазы	<i>ADCY3, ADCY9</i>	0.0000
GO: 0009975	Активность циклазы	<i>ADCY3, ADCY9</i>	0.0001
GO: 0016849	Активность фосфор-кислород-лиазы	<i>ADCY3, ADCY9</i>	0.0001
GO: 0030674	Белок-белковые взаимодействия	<i>HIP1, RPTOR, SH2B1</i>	0.0003
GO: 0060090	Активность молекулярного адаптера	<i>HIP1, RPTOR, SH2B1</i>	0.0004
GO: 0019901	Связывание протеинкиназ	<i>HIP1, RPTOR, SH2B1, TCF7L2</i>	0.0010
GO: 0019900	Связывание киназ	<i>HIP1, RPTOR, SH2B1, TCF7L2</i>	0.0017

влиянием катехоламинов, способствуя мобилизации необходимых энергетических ресурсов и снижению содержания инсулина [39]. Следует сказать, что стресс может быть одной из важных составляющих процесса эволюции, поскольку большинство изменений в ходе эволюции возникают под воздействием новых условий окружающей среды, провоцирующих в том числе и стресс [40].

В контексте полученной ассоциации с такими GO-категориями, как “биосинтез и метаболический процесс сАМР”, интересной представляется связь глюкагона с сАМР. Глюкагон вместе с другими гормонами, участвующими в повышении уровня глюкозы в крови, влияет на соответствующие рецепторы печени, генерируя сАМР, который активирует гликогенфосфорилазу печени, превращая гликоген в глюкозу [41]. Известно так-

же обратное влияние сАМР на секрецию глюкагона через альфа-клетки [42].

Естественный отбор способствует адаптации популяций к изменяющимся условиям среды в процессе расселения человека по территории земного шара, накладывая отпечаток на генофонд популяций. Анализ роли действия отбора на формирование структуры отдельных участков генома представляет одну из актуальных задач современной генетики. В связи с чем, в данной работе изучена генетическая структура популяций человека и проведен поиск сигналов действия отбора по кластеру полиморфных генетических маркеров, связанных с ожирением и ИМТ. По совокупности полученных данных наиболее достоверные свидетельства действия естественного отбора показаны для следующих SNP: rs1167827 гена *HIP1*,

rs7138803 и rs7164727, расположенных в межгенной области, rs7141420 гена *NRXN3*, rs7498665 гена *SH2B1*, rs7903146 гена *TCF7L2*. Сигналы адаптации этих генетических маркеров обнаружены как по результатам теста *FDIST*, так и теста Юинса–Ваттерсона. Кроме того, эти же SNP показали связь с климато-географическими факторами и/или распространенностью ожирения. В частности, вариабельность частоты предкового аллеля маркера rs1167827 статистически значимо коррелирует с изменением восьми факторов, полиморфизма rs7138803 – пяти, аллельных вариантов rs7164727 и rs7141420 – трех, rs7498665 – четырех и rs7903146 – шести из восьми факторов, исследованных в данной работе.

Интересно, что ряд изученных локусов связан не только с ИМТ и ожирением, но и, согласно опубликованным данным, ассоциирован также с другими фенотипами. Так, генетический маркер rs1167827 связан с математическими способностями, rs7138803 – с возрастом менархе, rs7498665 и rs7903146 – с сахарным диабетом второго типа [43–47]. Возможная селективная значимость этих SNP обусловлена, по-видимому, функциональной значимостью генов и их продуктов.

Продуктом гена *HIP1* является белок, взаимодействующий с белком гентингином. Этот белок необходим для поддержания гомеостаза как на клеточном, так и на организменном уровнях, а его дефицит приводит к потере веса [48]. Уровень экспрессии *HIP1* коррелирует с повышенным уровнем экспрессии рецептора эпидермального фактора роста в некоторых видах опухолей [49]. Продукт гена *HIP1* играет роль в онкогенезе, обеспечивая выживание опухолевых клеток [50]. Основываясь на тесной связи между некоторыми видами рака и ожирением, можно заключить, что белок *HIP1* участвует, вероятно, в метаболических процессах, влияющих как на ИМТ, так и на онкогенез [51, 52].

Известно, что продукт гена *NRXN3* (нейрексин-3) входит в семейство молекул клеточной адгезии, необходимых для функционирования синапсов [53]. Генетическая изменчивость в этом локусе ассоциирована с такими поведенческими фенотипами, как алкогольная и наркотическая зависимость [54–56]. При этом наркотическая зависимость и ожирение имеют общие неврологические механизмы, посредством которых продукт гена может быть вовлечен в развитие ожирения [57].

Продукт гена *SH2B1* – адаптерный белок, участвующий в передаче сигналов лептина и инсулина, способных регулировать метаболизм глюкозы и липидов. Предполагается, что нарушение молекулярных механизмов действия лептина и инсулина является одним из факторов развития ожирения [58, 59]. Делеция 16p11.2, включающая данный ген, ассоциирована с ранним ожирением, а также

с аномальным развитием почек, кишечника и с нарушениями когнитивных функций (аутизм и задержка развития) [60–62].

Ген *TCF7L2* кодирует фактор транскрипции 7, который входит в состав *Wnt*-сигнального пути и влияет на экспрессию проглюкагона, а, следовательно, на регуляцию уровня глюкозы в крови [63]. Полиморфные варианты этого гена связаны с ИМТ, общим объемом жировых отложений и количественным содержанием висцерального и подкожного жира [64].

Интересным представляется действие естественного отбора на межгенные участки генома, включающие rs7138803 и rs7164727. Так, согласно базе данных “miRDB” (<http://www.mirdb.org>), генетический маркер rs7164727 локализован в сайте связывания следующих микроРНК: hsa-miR-1202, hsa-miR-12121, hsa-miR-4480 и hsa-miR-192-3p, играющих существенную роль в развитии депрессии и различных онкологических заболеваний [65–68]. Кроме того, по данным ресурса “UCSC Genome Browser” аллельный вариант rs7164727 расположен вблизи энхансера (около 2 т.п.н.). В соответствии с существующими представлениями о структуре неравновесия по сцеплению и размере гаплотипических блоков в геноме человека можно предположить, что функциональный эффект этого маркера обусловлен сцеплением с энхансерным регионом [69]. В этом контексте особую значимость приобретают результаты, свидетельствующие о действии недавнего положительного отбора на ряд энхансеров, участвующих в регуляции процессов в иммунной системе человека и обуславливающих тканеспецифичность экспрессии генов [70]. Функциональная значимость маркера rs7138803, вероятно, связана с его регуляторными характеристиками. Показано, что в популяции европеоидов данный SNP является eQTL для локуса *LMBR1L*, кодирующего интегральный мембранный белок, который ранее ассоциировали с участием в механизмах развития конечностей у млекопитающих. Недавно обнаружено, что этот белок играет ключевую роль в регуляции важнейшего внутриклеточного сигнального пути *Wnt*/β-катенин, который вовлечен в процессы гемопоэза, апоптоза, эмбриогенеза, дифференцировку клеток и развитие злокачественных опухолей [71].

В настоящем исследовании отклонение от селективной нейтральности полиморфных маркеров локуса *FTO*, признанного генетического фактора риска развития ожирения и повышенного ИМТ, выявлено только в европеоидных популяциях (удмурты, русские, коми, французы, сардинцы и калашы), что согласуется с результатами работ, в которых сигналы действия естественного отбора на данный локус обнаружены еще в одной европеоидной популяции – у сорбов, проживающих в Германии [72], а также на популяционную

дифференциацию частот аллелей и гаплотипов гена *FTO* [73, 74].

Следует отметить, что с точки зрения эволюционной медицины болезнь можно рассматривать как результат конфликта между произошедшей в прошлом адаптивной селекцией генетических вариантов и современными средовыми условиями, т.е. как несоответствие генетического разнообразия популяций человека настоящему образу жизни. В представленной работе исследованы полиморфные генетические варианты, ассоциированные с ожирением и ИМТ, в популяциях человека. Полученные результаты свидетельствуют о статистически значимой корреляции генетического разнообразия и климато-географических условий среды обитания популяций человека. Нами выявлено действие естественного отбора на ряд локусов, связанных с ожирением и избыточной массой тела. Известно, что развитие этих патологических состояний связано с дисбалансом между потребляемой и расходуемой энергией, в основе которого лежат нарушения метаболизма, генетическая предрасположенность, нарушение поведенческих реакций и влияние внешних факторов [75]. В 50% стран мира распространенность ожирения среди взрослых превышает 20%. [76]. Кроме того, анализ распределения частоты ожирения выявил гендерные и этнические особенности распространенности этого показателя: более низкий уровень среди взрослых зафиксирован в странах Азии и Африки, а более высокие значения обнаружены в Европе, Северной и Южной Америке [76, 77].

Существующие на сегодняшний день эволюционные концепции, объясняющие распространенность ожирения в современных популяциях человека, в частности, гипотезы “дрейфующих генов” и “экономных генотипов” не раскрывают в полной мере причин влияния глобализации и изменения демографической ситуации на генетическую основу ожирения и этническую вариабельность распространенности этой патологии в мире.

Данные, полученные в нашей работе, свидетельствуют, что наблюдаемая в промышленно развитых странах пандемия ожирения может быть связана не только с изменением типа питания, но и со сменой климатических условий и дифференцированным воздействием факторов окружающей среды на предков современных людей в период миграционных процессов, начавшихся около 70 000 лет назад [78, 79]. В контексте полученных результатов актуальной представляется также концепция “экономных фенотипов”, предполагающая ведущее участие окружающей среды в формировании определенного метаболического профиля плода в эмбриональный период, который на последующих этапах онтогенеза в условиях современного типа питания способ-

ствует развитию инсулинорезистентности и ожирения в сочетании с сердечно-сосудистыми заболеваниями и сахарным диабетом [80, 81]. Эти результаты расширяют представления о структуре и возможных механизмах генетической дифференциации наследственной компоненты ожирения и подтверждают роль естественного отбора в формировании генетической структуры популяций человека по SNP-маркерам, ассоциированным с ИМТ и ожирением.

Работа выполнена на базе Центра коллективного пользования “Медицинская геномика” (НИИ медицинской генетики Томского НИМЦ).

Исследование выполнено при поддержке Российского фонда фундаментальных исследований (грант № 18-04-00758).

Все процедуры, выполненные в исследовании с участием людей, соответствуют этическим стандартам институционального и/или национального комитета по исследовательской этике и Хельсинкской декларации 1964 года и ее последующим изменениям или сопоставимым нормам этики. От каждого участника исследования получено информированное добровольное согласие.

Авторы заявляют об отсутствии конфликта интересов.

СПИСОК ЛИТЕРАТУРЫ

1. Boden G. (2008) Obesity and free fatty acids. *Endocrinol. Metab. Clin. North Am.* **37**, 635–646.
2. Turula M., Kaprio J., Rissanen A., Koskenvuo M. (1990) Body weight in the Finnish Twin Cohort. *Diabetes Res. Clin. Pract.* **10**, S33–36.
3. Wardle J., Carnell S., Haworth C.M., Plomin R. (2008) Evidence for a strong genetic influence on childhood adiposity despite the force of the obesogenic environment. *Am. J. Clin. Nutr.* **87**, 398–404.
4. GWAS Catalog <https://www.ebi.ac.uk/gwas/> Accessed September 20, 2018.
5. Neel J.V. (1999) The “thrifty genotype” in 1998. *Nutr. Rev.* **57**, S2–S9.
6. Speakman J.R. (2008) Thrifty genes for obesity, an attractive but flawed idea, and an alternative perspective: the ‘drifty gene’ hypothesis. *Int. J. Obes. (Lond.)*. **32**, 1611–1617.
7. Sellayah D., Cagampang F.R., Cox R.D. (2014) On the evolutionary origins of obesity: a new hypothesis. *Endocrinology*. **155**, 1573–1588.
8. Berndt S.I., Gustafsson S., Mägi R.; GIANT Consortium (2013) Genome-wide meta-analysis identifies 11 new loci for anthropometric traits and provides insights into genetic architecture. *Nat. Genet.* **45**, 501–512.
9. Locke A.E., Kahali B., Berndt S.I.; GIANT Consortium (2015) Genetic studies of body mass index yield new insights for obesity biology. *Nature*. **518**, 197–206.

10. Felix J.F., Bradfield J.P., Monnereau C.; GIANT Consortium (2016) Genome-wide association analysis identifies three new susceptibility loci for childhood body mass index. *Hum. Mol. Genet.* **25**, 389–403.
11. Justice A.E., Winkler T.W., Feitosa M.F.; GIANT Consortium (2017) Genome-wide meta-analysis of 241258 adults accounting for smoking behaviour identifies novel loci for obesity traits. *Nat. Commun.* **8**, 14977.
12. Akiyama M., Okada Y., Kanai M.; GIANT Consortium (2017) Genome-wide association study identifies 112 new loci for body mass index in the Japanese population. *Nat. Genet.* **49**, 1458–1467.
13. Graff M., Scott R.A., Justice A.E.; GIANT Consortium (2017) Genome-wide physical activity interactions in adiposity – A meta-analysis of 200452 adults. *PLoS Genet.* **13**, e1006528.
14. Warrington N.M., Howe L.D., Paternoster L.; GIANT Consortium (2015) A genome-wide association study of body mass index across early life and childhood. *Int. J. Epidemiol.* **44**, 700–712.
15. Speliotes E.K., Willer C.J., Berndt S.I.; GIANT Consortium (2010) Association analyses of 249796 individuals reveal 18 new loci associated with body mass index. *Nat. Genet.* **42**, 937–948.
16. Graff M., Ngwa J.S., Workalemahu T., Homuth G., Schipf S., Teumer A., Völzke H., Wallaschofski H., Abecasis G.R., Edward L., Francesco C., Sanna S., Scheet P., Schlessinger D., Sidore C., Xiao X., Wang Z., Chanock S.J., Jacobs K.B., Hayes R.B., Hu F., Van Dam R.M.; GIANT Consortium, Crout R.J., Marazita M.L., Shaffer J.R., Atwood L.D., Fox C.S., Heard-Costa N.L., White C., Choh A.C., Czerwinski S.A., Demerath E.W., Dyer T.D., Towne B., Amin N., Oostra B.A., Van Duijn C.M., Zillikens M.C., Esko T., Nelis M., Nikopensius T., Metspalu A., Strachan D.P., Monda K., Qi L., North K.E., Cupples L.A., Gordon-Larsen P., Berndt S.I. (2013) Genome-wide analysis of BMI in adolescents and young adults reveals additional insight into the effects of genetic loci over the life course. *Hum. Mol. Genet.* **22**, 3597–3607.
17. Wen W., Zheng W., Okada Y.; GIANT Consortium (2014) Meta-analysis of genome-wide association studies in East Asian-ancestry populations identifies four new loci for body mass index. *Hum. Mol. Genet.* **23**, 5492–5504.
18. Thorleifsson G., Walters G.B., Gudbjartsson D.F., Steinthorsdottir V., Sulem P., Helgadóttir A., Styrkarsdóttir U., Gretarsdóttir S., Thorlacius S., Jonsdóttir I., Jonsdóttir T., Olafsdóttir E.J., Olafsdóttir G.H., Jonsson T., Jonsson F., Borch-Johnsen K., Hansen T., Andersen G., Jorgensen T., Lauritzen T., Aben K.K., Verbeek A.L., Roeleveld N., Kampman E., Yanek L.R., Becker L.C., Tryggvadóttir L., Rafnar T., Becker D.M., Gulcher J., Kiemeny L.A., Pedersen O., Kong A., Thorsteinsdóttir U., Stefansson K. (2009) Genome-wide association yields new sequence variants at seven loci that associate with measures of obesity. *Nat. Genet.* **41**, 18–24.
19. Wheeler E., Huang N., Bochukova E.G., Keogh J.M., Lindsay S., Garg S., Henning E., Blackburn H., Loos R.J., Wareham N.J., O’Rahilly S., Hurles M.E., Barroso I., Farooqi I.S. (2013) Genome-wide SNP and CNV analysis identifies common and low-frequency variants associated with severe early-onset obesity. *Nat. Genet.* **45**, 513–517.
20. Comuzzie A.G., Cole S.A., Laston S.L., Voruganti V.S., Haack K., Gibbs R.A., Butte N.F. (2012) Novel genetic loci identified for the pathophysiology of childhood obesity in the Hispanic population. *PLoS One.* **7**, e51954.
21. Willer C.J., Speliotes E.K., Loos R.J.; GIANT Consortium (2009) Six new loci associated with body mass index highlight a neuronal influence on body weight regulation. *Nat. Genet.* **41**(1), 25–34.
22. Ng MCY., Graff M., Lu Y.; GIANT Consortium (2017) Discovery and fine-mapping of adiposity loci using high density imputation of genome-wide association studies in individuals of African ancestry: African Ancestry Anthropometry Genetics Consortium. *PLoS Genet.* **13**, e1006719.
23. Namjou B., Keddache M., Marsolo K., Wagner M., Lingren T., Cobb B., Perry C., Kennebeck S., Holm I.A., Li R., Crimmins N.A., Martin L., Solti I., Kohane I.S., Harley J.B. (2013) EMR-linked GWAS study: investigation of variation landscape of loci for body mass index in children. *Front. Genet.* **4**, 268.
24. Cotsapas C., Speliotes E.K., Hatoum I.J., Greenawald D.M., Dobrin R., Lum P.Y., Suver C., Chudin E., Kemp D., Reitman M., Voight B.F., Neale B.M., Schadt E.E., Hirschhorn J.N., Kaplan L.M., Daly M.J.; GIANT Consortium (2009) Common body mass index-associated variants confer risk of extreme obesity. *Hum. Mol. Genet.* **18**, 3502–3507.
25. Степанов В.А., Трифонова Е.А. (2013) Мультиплексное гентипирование однонуклеотидных полиморфных маркеров методом MALDI-TOF масс-спектрометрии: частоты 56 SNP в генах иммунного ответа в популяциях человека. *Молекуляр. биология.* **47**, 976–986.
26. Weatherbase <https://www.weatherbase.com/> Accessed December 2, 2018.
27. Министерство здравоохранения Российской Федерации <https://www.rosminzdrav.ru/ministry/61/22/stranitsa-979/statisticheskie-informatsionnye-materialy/statisticheskiy-sbornik-2017-god> Accessed November 20, 2018.
28. World Health Organization <https://www.who.int/ru> Accessed November 17, 2018.
29. Arlequin ver 3.5.2.2 <http://cmpg.unibe.ch/software/arlequin35/> Accessed November 2, 2018.
30. String <https://string-db.org/> Accessed July 1, 2019.
31. WEB-based Gene Set Analysis Toolkit <http://www.webgestalt.org/> Accessed July 1, 2019.
32. Степанов В.А. (2016) Эволюция генетического разнообразия и болезни человека. *Генетика.* **52**, 852–864.
33. Степанов В.А., Харьков В.Н., Вагайцева К.В., Бочарова А.В., Казанцев А.Ю., Попович А.А., Хитринская И.Ю. (2017) Поиск генетических маркеров адаптации к климату у населения Северной Евразии. *Генетика.* **53**, 1254–1266

34. Степанов В.А., Канделария П., Кхо С., Жанг Г., Трифонова Е.А., Симонова К.В., Бочарова А.В., Лайнг И.А., Хайден К., Баккер В., Бьеррегаард П., Линч Н., Хагель И., Маск А.У., Битлз А., Голдблатт Дж., Ле Сеф П. (2013) Деканализация иммунного ответа при расселении современного человека: связь генетического разнообразия в генах иммунной системы с климато-географическими факторами. *Мед. генетика*. **12**, 8–18.
35. Nanoune J., Defer N. (2001) Regulation and role of adenylyl cyclase isoforms. *Annu. Rev. Pharmacol. Toxicol.* **41**, 145–174.
36. Лукьянова Л.Д. (2004) Роль биоэнергетических нарушений в патогенезе гипоксии. *Патол. физиология и эксп. терапия*. **2**, 2–11.
37. Левин Ю.М. (2006) *Прорыв в эндоэкологическую медицину. Новый уровень врачебного мышления и эффективной терапии*. М.: Щербинская типография.
38. Adeva-Andany M.M., Funcasta-Calderón R., Fernández-Fernández C., Castro-Quintela E., Carneiro-Freire N. (2018) Metabolic effects of glucagon in humans. *J. Clin. Transl. Endocrinol.* **15**, 45–53.
39. Чеснокова Н.П., Михайлова А.В. (2006) *Инфекционный процесс*. М.: Акад. естествознания.
40. Parsons P.A. (2005) Environments and evolution: interactions between stress, resource inadequacy and energetic efficiency. *Biol. Rev. Camb. Philos. Soc.* **80**, 589–610.
41. Butcher R.W., Robison G.A., Hardman J.G., Sutherland E.W. (1968) The role of cyclic AMP in hormone actions. *Adv. Enzyme Regul.* **6**, 357–389.
42. Johansson H., Gylfe E., Hellman B. (1989) Cyclic AMP raises cytoplasmic calcium in pancreatic alpha 2-cells by mobilizing calcium incorporated in response to glucose. *Cell Calcium*. **10**, 205–211.
43. Lee J.J., Wedow R., Okbay A.; GIANT Consortium (2018) Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. *Nat. Genet.* **50**, 1112–1121.
44. Sladek R., Rocheleau G., Rung J., Dina C., Shen L., Serre D., Boutin P., Vincent D., Belisle A., Hadjadj S., Balkau B., Heude B., Charpentier G., Hudson T.J., Montpetit A., Pshezhetsky A.V., Prentki M., Posner B.I., Balding D.J., Meyre D., Polychronakos C., Froguel P. (2007) A genome-wide association study identifies novel risk loci for type 2 diabetes. *Nature*. **445**, 881–885.
45. Timpson N.J., Lindgren C.M., Weedon M.N., Randall J., Ouwehand W.H., Strachan D.P., Rayner N.W., Walker M., Hitman G.A., Doney A.S., Palmer C.N., Morris A.D., Hattersley A.T., Zeggini E., Frayling T.M., McCarthy M.I. (2009) Adiposity-related heterogeneity in patterns of type 2 diabetes susceptibility observed in genome-wide association data. *Diabetes*. **58**, 505–510.
46. Elks C.E., Perry J.R., Sulem P.; GIANT Consortium (2010) Thirty new loci for age at menarche identified by a meta-analysis of genome-wide association studies. *Nat. Genet.* **42**, 1077–1085.
47. Xi B., Takeuchi F., Meirhaeghe A.; Kato N., Chambers J.C., Morris A.P., Cho Y.S., Zhang W., Mohlke K.L., Kooner J.S., Shu X.O., Pan H., Tai E.S., Pan H., Wu J.Y., Zhou D., Chandak G.R.; DIAGRAM Consortium; AGEN-T2D Consortium; SAT2D Consortium. (2014) Associations of genetic variants in/near body mass index-associated genes with type 2 diabetes: a systematic meta-analysis. *Clin. Endocrinol. (Oxf)*. **81**, 702–710.
48. Bradley S.V., Hyun T.S., Oravec-Wilson K.I., Li L., Waldorff E.I., Ermilov A.N., Goldstein S.A., Zhang C.X., Drubin D.G., Varela K., Parlow A., Dlugosz A.A., Ross T.S. (2007) Degenerative phenotypes caused by the combined deficiency of murine HIP1 and HIP1r are rescued by human HIP1. *Hum. Mol. Genet.* **16**, 1279–1292.
49. Bradley S.V., Holland E.C., Liu G.Y., Thomas D., Hyun T.S., Ross T.S. (2007) Huntingtin interacting protein 1 is a novel brain tumor marker that associates with epidermal growth factor receptor. *Cancer Res.* **67**, 3609–3615.
50. Rao D.S., Hyun T.S., Kumar P.D., Mizukami I.F., Rubin M.A., Lucas P.C., Sanda M.G., Ross T.S. (2002) Huntingtin-interacting protein 1 is overexpressed in prostate and colon cancer and is critical for cellular survival. *J. Clin. Invest.* **110**, 351–360.
51. Lv W.Q., Zhang X., Zhang Q., He J.Y., Liu H.M., Xia X., Fan K., Zhao Q., Shi X.Z., Zhang W.D., Sun C.Q., Deng H.W. (2017) Novel common variants associated with body mass index and coronary artery disease detected using a pleiotropic cFDR method. *J. Mol. Cell Cardiol.* **112**, 1–7.
52. Haggerty A.F., Sarwer D.B., Schmitz K.H., Ko E.M., Allison K.C., Chu C.S. (2017) Obesity and endometrial cancer: a lack of knowledge but opportunity for intervention. *Nutr. Cancer*. **69**, 990–995.
53. Rowen L., Young J., Birditt B., Kaur A., Madan A., Philipps D.L., Qin S., Minx P., Wilson R.K., Hood L., Graveley B.R. (2002) Analysis of the human neurexin genes: alternative splicing and the generation of protein diversity. *Genomics*. **79**, 587–597.
54. Lachman H.M., Fann C.S., Bartzis M., Evgrafov O.V., Rosenthal R.N., Nunes E.V., Miner C., Santana M., Gaffney J., Riddick A., Hsu C.L., Knowles J.A. (2007) Genomewide suggestive linkage of opioid dependence to chromosome 14q. *Hum. Mol. Genet.* **16**, 1327–1334.
55. Kelai S., Maussion G., Noble F., Boni C., Ramoz N., Moalic J.M., Peuchmaur M., Gorwood P., Simonneau M. (2008) Nr3x3 upregulation in the globus pallidus of mice developing cocaine addiction. *Neuroreport*. **19**, 751–755.
56. Clay S.W., Allen J., Parran T. (2008) A review of addiction. *Postgrad. Med.* **120**, E01–E07.
57. Rapaka R., Schnur P., Shurtleff D. (2008) Obesity and addiction: common neurological mechanisms and drug development. *Physiol. Behav.* **95**, 2–9.
58. O'Brien K.B., O'Shea J.J., Carter-Su C. (2002) SH2-B family members differentially regulate JAK family tyrosine kinases. *J. Biol. Chem.* **277**, 8673–8681.
59. Zhou Y., Rui L. (2013) Leptin signaling and leptin resistance. *Front. Med.* **7**, 207–222.

60. Bochukova E.G., Huang N., Keogh J., Henning E., Purmann C., Blaszczyk K., Saeed S., Hamilton-Shield J., Clayton-Smith J., O'Rahilly S., Hurles M.E., Farooqi I.S. (2010) Large, rare chromosomal deletions associated with severe early-onset obesity. *Nature*. **463**, 666–670.
61. Bachmann-Gagescu R., Mefford H.C., Cowan C., Glew G.M., Hing A.V., Wallace S., Bader P.I., Hamati A., Reitnauer P.J., Smith R., Stockton D.W., Muhle H., Helbig I., Eichler E.E., Ballif B.C., Rosenfeld J., Tsuchiya K.D. (2010) Recurrent 200-kb deletions of 16p11.2 that include the SH2B1 gene are associated with developmental delay and obesity. *Genet. Med.* **12**, 641–647.
62. Sampson M.G., Coughlin C.R., Kaplan P., Conlin L.K., Meyers K.E., Zackai E.H., Spinner N.B., Copelovitch L. (2010) Evidence for a recurrent microdeletion at chromosome 16p11.2 associated with congenital anomalies of the kidney and urinary tract (CAKUT) and Hirschsprung disease. *Am. J. Med. Genet. A*. **152A**, 2618–2622.
63. Grant S.F., Thorleifsson G., Reynisdottir I., Benediktsson R., Manolescu A., Sainz J., Helgason A., Stefansson H., Emilsson V., Helgadottir A., Styrkarsdottir U., Magnusson K.P., Walters G.B., Palsdottir E., Jonsdottir T., Gudmundsdottir T., Gylfason A., Saemundsdottir J., Wilensky R.L., Reilly M.P., Rader D.J., Bagger Y., Christiansen C., Gudnason V., Sigurdsson G., Thorsteinsdottir U., Gulcher J.R., Kong A., Stefansson K. (2006) Variant of transcription factor 7-like 2 (TCF7L2) gene confers risk of type 2 diabetes. *Nat. Genet.* **38**, 320–323.
64. Haupt A., Thamer C., Heni M., Ketterer C., Machann J., Schick F., Machicao F., Stefan N., Claussen C.D., Häring H.U., Fritsche A., Staiger H. (2010) Gene variants of TCF7L2 influence weight loss and body composition during lifestyle intervention in a population at risk for type 2 diabetes. *Diabetes*. **59**, 747–750.
65. Lopez J.P., Lim R., Cruceanu C., Crapper L., Fasano C., Labonte B., Maussion G., Yang J.P., Yerko V., Vigneault E., El Mestikawy S., Mechawar N., Pavlidis P., Turecki G. (2014) miR-1202 is a primate-specific and brain-enriched microRNA involved in major depression and antidepressant treatment. *Nat. Med.* **20**, 764–768.
66. Du B., Zhang P., Tan Z., Xu J. (2017) MiR-1202 suppresses hepatocellular carcinoma cells migration and invasion by targeting cyclin dependent kinase 14. *Biomed. Pharmacother.* **96**, 1246–1252.
67. Quan Y., Song Q., Wang J., Zhao L., Lv J., Gong S. (2017) MiR-1202 functions as a tumor suppressor in glioma cells by targeting Rab1A. *Tumour Biol.* **39**(4). <https://doi.org/10.1177/1010428317697565>
68. Botla S.K., Savant S., Jandaghi P., Bauer A.S., Mücke O., Moskalev E.A., Neoptolemos J.P., Costello E., Greenhalf W., Scarpa A., Gaida M.M., Büchler M.W., Strobel O., Hackert T., Giese N.A., Augustin H.G., Hoheisel J.D. (2016) Early epigenetic downregulation of microRNA-192 expression promotes pancreatic cancer progression. *Cancer Res.* **76**, 4149–4159.
69. Ceballos F.C., Joshi P.K., Clark D.W., Ramsay M., Wilson J.F. (2018) Runs of homozygosity: windows into population history and trait architecture. *Nat. Rev. Genet.* **19**, 220–234.
70. Moon J.M., Capra J.A., Abbot P., Rokas A. (2019) Signatures of recent positive selection in enhancers across 41 human tissues. *G3: Genes, Genomes, Genetics.* **9**, 2761–2774.
71. Choi J.H., Zhong X., McAlpine W., Liao T.C., Zhang D., Fang B., Russell J., Ludwig S., Nair-Gill E., Zhang Z., Wang K.W., Misawa T., Zhan X., Choi M., Wang T., Li X., Tang M., Sun Q., Yu L., Murray A.R., Moresco E.M.Y., Beutler B. (2019) LMBR1L regulates lymphopoiesis through Wnt/ β -catenin signaling. *Science*. **364**, eaau0812.
72. Liu X., Weidle K., Schröck K., Tönjes A., Schleinitz D., Breitfeld J., Stumvoll M., Böttcher Y., Schöneberg T., Kovacs P. (2015) Signatures of natural selection at the FTO (fat mass and obesity associated) locus in human populations. *PLoS One*. **10**, e0117093.
73. Mao L., Fang Y., Campbell M., Southerland W.M. (2017) Population differentiation in allele frequencies of obesity-associated SNPs. *BMC Genomics.* **10**, 861.
74. Babenko V., Babenko R., Gamielien J., Markel A. (2019) FTO haplotyping underlines high obesity risk for European populations. *BMC Med. Genomics.* **12**(Suppl 2), 46.
75. Romieu I., Dossus L., Barquera S., Blotière H.M., Franks P.W., Gunter M., Hwalla N., Hursting S.D., Leitzmann M., Margetts B., Nishida C., Potischman N., Seidell J., Stepien M., Wang Y., Westerterp K., Winichagoon P., Wiseman M., Willett W.C.; IARC working group on Energy Balance and Obesity (2017) Energy balance and obesity: what are the main drivers? *Cancer Causes Control.* **28**, 247–258.
76. Souza S.A., Silva A.B., Cavalcante U.M.B., Lima C.M.B.L., Souza T.C. (2018) Adult obesity in different countries: an analysis via beta regression models. *Cad Saude Publica.* **34**(8), e00161417. <https://doi.org/10.1590/0102-311X00161417>
77. Джумагулова А.С., Полупанов А.Г., Халматов А.Н., Алтымышева А.Т., Маматов А.У., Романова Т.А. (2019) Гендерные и этнические особенности распространенности ожирения среди жителей малых городов и сельской местности Кыргызской Республики (по данным исследования “Интерэпид”). *Кардиол. вест.* **14**, 61–66.
78. Rosenberg N.A., Kang J.T. (2015) Genetic diversity and societally important disparities. *Genetics.* **201**, 1–12.
79. López S., van Dorp L., Hellenthal G. (2016) Human dispersal out of Africa: a lasting debate. *Evol. Bioinform. Online.* **11**, 57–68.
80. Hales C.N., Barker D.J., Clark P.M., Cox L.J., Fall C., Osmond C., Winter P.D. (1991) Fetal and in fant growth and impaired glucose tolerance at age 64 years. *Biol. Med. J.* **303**, 1019–1022.
81. Portha B., Grandjean V., Movassat J. (2019) Mother or father: who is in the front Line? Mechanisms underlying the non-genomic transmission of obesity/diabetes via the maternal or the paternal line. *Nutrients.* **11**, E233.

THE ROLE OF NATURAL SELECTION IN THE FORMATION OF THE GENETIC STRUCTURE OF POPULATIONS BY SNP MARKERS IN ASSOCIATION WITH BODY MASS INDEX AND OBESITY

E. A. Trifonova^{1,*}, A. A. Popovich¹, A. V. Bocharova¹, K. V. Vagaitseva¹, and V. A. Stepanov¹

¹*Institute of Medical Genetics, Tomsk National Medical Research Center, Russian Academy of Sciences, Tomsk, 634050 Russia*

**e-mail: ekaterina-trifonova@medgenetics.ru*

Obesity becomes one of the global problems in modern society, provided that more than a third of the world's population suffer from overweight. This phenotype affects the quality of life and is associated with cardiovascular diseases, diabetes, cancer and reproductive disorders. The population variability of allele frequencies of 26 single nucleotide polymorphisms, in association with obesity and body mass index according to the data of genome-wide association studies (GWASs) is discussed in this study. The genetic variability was analyzed in populations of Northern Eurasia and populations from the human genome diversity project (HGDP). The studied population samples are characterized by high genetic diversity that correlates with climatic and geographical parameters. The results of the test for searching natural selection signals in total revealed the selection effect for the rs1167827 of the *HIP1* gene, rs7138803 and rs7164727, located in the intergenic region, rs7141420 of the *NRXN3* gene, rs7498665 of the *SH2B1* gene, rs7903146 of the *TCF7L2* gene.

Keywords: human populations, genetic diversity, obesity, body mass index, natural selection