

УДК 575.22:577.21:599.9

РЕПЛИКАТИВНОЕ ИССЛЕДОВАНИЕ ПОДВЕРЖЕННОСТИ ШИЗОФРЕНИИ С РАННИМ НАЧАЛОМ У КАЗАХОВ

© 2015 г. В. А. Степанов^{1,2}, А. В. Бочарова², К. З. Садуакасова³, А. В. Марусин²,
Л. А. Конева², К. В. Вагайцева^{1,2}, Г. С. Святова⁴

¹Национальный исследовательский Томский государственный университет, Томск 634050

²Научно-исследовательский институт медицинской генетики
Российской академии медицинских наук, Томск 634050
e-mail: vadim.stepanov@medgenetics.ru

³Казахский национальный медицинский университет им. С.Д. Асфендиярова
Министерства здравоохранения Республики Казахстан, Алматы 050012, Казахстан

⁴Научный центр акушерства, гинекологии и перинатологии Министерства здравоохранения Республики Казахстан,
Алматы 050020, Казахстан

Поступила в редакцию 12.05.2014 г.

В настоящей работе проведен репликативный анализ ассоциаций 15 SNP в области 14 генов, ранее выявленных в широкогеномных исследованиях (GWAS), с шизофренией с ранним началом у казахов. Обнаружена ассоциация ранней шизофрении с маркерами трех генов (*VRK2*, *KCNB2* и *CPVL*). Ассоциация rs2312147 гена *VRK2* с шизофренией ранее была также реплицирована у китайцев, что позволяет рассматривать этот маркер как вероятный расо-специфичный. С помощью методов снижения размерности многофакторных данных были обнаружены две группы из четырех и шести генов, демонстрирующих межгенные эпистатические взаимодействия. Генные онтологии 14 изученных генов сводятся к вариантам одной молекулярной функции (пептидазная активность) и одному биологическому процессу (положительная регуляция процессов биосинтеза). Биоинформационный анализ белок-белковых взаимодействий продуктов изучаемых генов показывает, что продукты шести из 14 генов могут быть вовлечены в единую взаимосвязанную сеть, главным связующим звеном которой является убиквитинилирование их белком UBC.

DOI: 10.7868/S0016675815020149

На протяжении последних лет шизофрения остается одним из наиболее изучаемых психических заболеваний, гетерогенность этиологии и клинических проявлений которого несомненна. По данным ВОЗ это расстройство занимает лидирующее положение в списке патологических состояний, которые приводят к инвалидности и социальной дезадаптации больных. Исследования, направленные на изучение молекулярно-генетических аспектов нейropsychиатрических заболеваний, показали, что шизофрения — типичное многофакторное заболевание (МФЗ), развитие которого является следствием комбинации межгенных и гено-средовых влияний и взаимодействий [1–3].

Одним из основных подходов к выявлению генетической компоненты МФЗ в последние годы являются широкогеномные ассоциативные исследования (GWAS, genome-wide associations studies). Важная составляющая поиска генетической компоненты МФЗ — репликативные исследования. Все маркеры, выявляемые в GWAS или полногеномном анализе сцепления (GWLS), требуют репликации на независимых выборках. На насто-

ящий момент по данным базы опубликованных полногеномных ассоциативных исследований, поддерживаемой Национальным институтом изучения генома человека США (NHGRI), опубликовано 31 широкогеномное исследование шизофрении, проведенное с 2007 по 2013 г. [4] (данные на 10 мая 2014 г.). В этих работах выявлено более 130 однонуклеотидных маркеров (SNP), высокодостоверно ($p < 1 \times 10^{-5}$) ассоциированных с заболеванием.

Одной из проблем интерпретации и применения результатов ассоциативных исследований является популяционная специфичность ассоциаций ген–признак [5]. Проблема заключается в том, что устойчивые сочетания генетических вариантов (как правило, распространенных вариантов полиморфизма), лежащие в основе развития комплексных мультифакторных заболеваний человека, формируются в контексте структуры генофонда и неравновесия по сцеплению, сложившихся в конкретной популяции в ходе действия микроэволюционных факторов [6]. Тем самым сами сочетания маркеров предрасположенности к МФЗ, фиксируемые как ассоциация

маркер—признак, как и сила эффекта отдельных маркеров, могут различаться в популяциях различного этнического происхождения. Подавляющее большинство GWAS по шизофрении, как и в отношении любого другого МФЗ, проведено на европеоидных популяциях. Из 31 опубликованного на сегодняшний день широкогеномного исследования шизофрении 24 проведены только на европеоидах, одна работа — на афроамериканцах в сравнении с белыми американцами [7] и шесть — на популяциях монголоидов Восточной Азии (китайцах и японцах) [8–13]. Таким образом, репликативное исследование генетических маркеров шизофрении, выявленных в GWAS на популяциях различного этнического происхождения, представляется актуальной задачей для поиска маркеров болезни и выявления картины популяционной специфичности ассоциаций.

Шизофрения является клинически гетерогенным заболеванием. Одна из форм болезни — шизофрения с ранним началом. Неоднозначное отношение ряда психиатрических школ к самому факту существования детской шизофрении требует дальнейших исследований в сравнительно-возрастном аспекте. Как отмечают практически все исследователи, трудности своевременной диагностики шизофрении в значительной мере зависят от того обстоятельства, что ее дебют часто приходится на детский или подростковый возраст — периоды, в которые чрезвычайно трудно различить начальные симптомы болезни и преморбидные особенности. Возрастной патоморфоз детской шизофрении характеризуется, в частности, тем, что формирующийся при ней более сложный по структуре психический дефект сочетается с сопутствующим процессом нарушением развития ребенка (дизонтогенезом), что может затруднять раннюю диагностику этого заболевания. Как и в других возрастных группах, ранняя детская шизофрения протекает непрерывно (злокачественно и вяло) и приступообразно (по степени прогрессивности близко к злокачественному типу, со средней и малой степенью прогрессивности) [14, 15]. В случаях дебюта шизофрении в подростковом возрасте (чаще в период активного пубертата) мы наблюдали параноидную форму шизофрении и приступообразно-прогрессиентную.

Исследования по поиску генетических маркеров ранней формы шизофрении немногочисленны [16–20]. Все они выполнены в рамках стратегии известных кандидатных генов. С точки зрения генетического картирования МФЗ, формы заболевания с ранним началом и более тяжелым течением являются предпочтительными мишенями для исследования ассоциаций, поскольку характеризуются, как правило, более выраженным вкладом генетической компоненты [21].

Задачей настоящей работы был поиск генетических маркеров шизофрении с началом в детском и подростковом возрасте у казахов на основе репликативного анализа маркеров, выявленных в широкогеномных исследованиях шизофрении и ее когнитивных эндотипов.

МАТЕРИАЛЫ И МЕТОДЫ

Были обследованы больные шизофренией дети и подростки, а также взрослые лица (с началом эндогенного процесса в детстве), находившиеся на лечении в клинических отделениях Республиканского научно-практического центра психиатрии, психотерапии и наркологии МЗ РК (клиническая база кафедры психиатрии, психотерапии и наркологии КазНМУ), в отделениях городского Центра психического здоровья г. Алматы, а также дети и подростки с инициальными проявлениями шизофрении группы амбулаторного наблюдения и контингент специализированного учреждения для психохроников г. Алматы.

Выборка больных состояла из 112 человек казахской национальности в возрасте от 4 до 30 лет. Средний возраст в группе больных на момент обследования составил 16.8 лет. За основу формирования групп пробандов для выполнения данного исследования были взяты возраст начала инициальных проявлений и возраст дебюта шизофрении в детском и подростковом возрасте. Диагнозы больных соответствовали рубрикам МКБ-10: ранняя детская шизофрения (код по МКБ-10 — F20.8xx3) — 58 пациентов, параноидная шизофрения с непрерывной формой течения (F20.00) и приступообразно-прогрессиентная шизофрения (F20.01; F20.02) — 54 пациента.

Контрольная группа состояла из 190 индивидов казахской этнической принадлежности в возрасте от 19 до 76 лет, не имеющих в анамнезе психоневрологических заболеваний. Средний возраст в группе контроля составил 31.6 лет. Более старшая возрастная группа была выбрана с целью исключения попадания в нее больных, у которых болезнь еще не проявилась или не была вовремя диагностирована. Национальность больных и членов контрольной группы устанавливалась на основании анкетирования и в исследование включались индивиды, неметисированные в трех поколениях. Исследование было одобрено биоэтическим комитетом ФГБУ НИИМГ ФАНО России. У всех пробандов (их законных представителей) и доноров (лица контрольной группы) было получено информированное согласие на проведение исследований и забор периферической крови.

Для репликативного анализа ассоциаций нами было выбрано 15 однонуклеотидных генетических маркеров, для которых была выявлена высо-

Таблица 1. Маркеры и гены для репликативного анализа ассоциаций с шизофренией

№	SNP	Минорный аллель	Ген
1	rs1502844	C	<i>SLCO6A1</i> , семейство переносчика растворенного органического аниона, элемент 6A1
2	rs9960767	C	<i>TCF4</i> , транскрипционный фактор 4
3	rs2312147	T	<i>VRK2</i> , киназа коровьей оспы 2
4	rs3131296	T	<i>NOTCH4</i> , нейрогенный локус, гомолог 4 белка notch
5	rs12807809	C	<i>NRGN</i> , нейрогранин
6	rs1572299	C	Межгенный регион: <i>TLR4</i> , толл-подобный рецептор 4-го типа и <i>DBC1</i> , белок 1, делегированный при раке мочевого пузыря
7	rs17594526	T	<i>TCF4</i> , транскрипционный фактор 4
8	rs1344706	C	<i>ZNF804A</i> , белок цинкового пальца 804A
9	rs16977195	G	<i>AGBL1</i> , <i>ATP/GTP</i> связывающий белок 1, цитозольная карбоксипептидаза
10	rs7341475	A	<i>RELN</i> , рилин
11	rs8020441	G	<i>ZFP64P1</i> , белок цинкового пальца 64, гомолог псевдогена 1 у мыши
12	rs2247572	T	<i>KCNB2</i> , потенциал-зависимый калиевый канал, Shab-связанное подсемейство, элемент 2
13	rs2616984	G	<i>CSMD1</i> (<i>KIAA1890</i>), CUB и Sushi множественный домен 1
14	rs2229741	T	<i>NRIP1</i> , белок 1, взаимодействующий с ядерным рецептором
15	rs2252521	T	<i>CPVL</i> , вителлогенин-подобная карбоксипептидаза

кодостоверная ассоциация с шизофренией и когнитивными признаками, являющимися эндотипами этого заболевания в недавних GWAS [22–28]. Характеристики генов и генетических маркеров приведены в табл. 1.

ДНК выделяли методом фенольно-хлороформной экстракции из цельной венозной крови. Генотипирование проводили методом ПЦР в реальном режиме времени с помощью TaqMan-проб фирмы “Applied Biosystems” (США) по протоколу производителя на амплификаторе с детекцией ПЦР в реальном времени фирмы Bio-Rad (США).

Тестирование равновесия Харди–Вайнберга (РХВ) и расчет ожидаемой гетерозиготности выполняли общепринятыми методами популяционной биометрии. Сравнение частот аллелей и генотипов в группах проводили с помощью критерия максимального правдоподобия χ^2 (МП χ^2). Силу ассоциаций оценивали в значениях показателя соотношения шансов OR и его 95%-ного доверительного интервала (95% CI). Анализ межгенных и ген-средовых взаимодействий проводили методом снижения многомерной размерности в программе MDR (Multifactor Dimensionality Reduction, <http://www.multifactor dimensionalityreduction.org>). Для анализа биологических процессов, связанных с изучаемыми генами, и генных сетей использовали биоинформационные ресурсы DAVID (The Database for Annotation, Visualization and Integrated Discovery), KEGG (Kyoto Encyclopedia of Genes and Genomes) и STRING (Search

Tool for the Retrieval of Interacting Genes/Proteins) [29–31].

РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

Частоты генов и генотипов и ассоциации генетических маркеров с шизофренией с ранним началом

Численности генотипов, частоты минорных аллелей, ожидаемая гетерозиготность и достигнутый уровень значимости соответствия наблюдаемого распределения генотипов ожидаемому при равновесии Харди–Вайнберга у больных и в контрольной группе представлены в табл. 2. Из 30 исследованных распределений генотипов наблюдалось два отклонения от распределения, ожидаемого при РХВ, для однонуклеотидного полиморфизма rs8020441 гена *ZFP64P1* в группе больных шизофренией и для rs2252521 гена *CPVL* в контрольной выборке. В целом частоты аллелей у больных и в контроле близки и находятся в пределах вариаций, наблюдаемых в мировых популяциях по данным проектов НарМар и “1000 геномов” [32, 33].

Сравнение частот аллелей и расчет отношения шансов развития болезни приведены в табл. 3. Для трех из 15 SNP выявлены статистически значимые ассоциации на уровне частот аллелей. Минорный аллель rs2252521 гена *CPVL* (OR = 1.46, $p = 0.037$) и мажорные аллели rs2312147 гена *VRK2* (OR = 1.72, $p = 0.008$) и rs2247572 гена *KCNB2* (OR = 1.54, $p = 0.030$) достоверно чаще встречались среди больных ранней шизофренией казахов

Таблица 2. Распределение генотипов и частоты аллелей у больных и в контрольной группе

№	SNP	MA	Больные						Контроль					
			11	12	22	MAF	H_e	P	11	12	22	MAF	H_e	P
1	rs1502844	C	32	46	22	0.450	0.495	0.480	67	82	41	0.432	0.491	0.097
2	rs9960767	C	94	7	0	0.035	0.067	0.718	180	10	0	0.026	0.051	0.708
3	rs2312147	T	53	34	4	0.231	0.355	0.617	79	90	19	0.340	0.449	0.365
4	rs3131296	T	88	14	0	0.069	0.128	0.458	157	30	2	0.089	0.164	0.689
5	rs12807809	C	51	42	7	0.280	0.403	0.680	108	64	18	0.263	0.388	0.070
6	rs1572299	C	61	37	4	0.221	0.344	0.579	129	56	5	0.174	0.287	0.708
7	rs17594526	T	101	1	0	0.005	0.001	0.960	181	3	0	0.008	0.002	0.913
8	rs1344706	A	26	50	25	0.495	0.500	0.923	50	89	50	0.500	0.500	0.424
9	rs16977195	G	85	13	1	0.076	0.140	0.538	173	16	1	0.047	0.090	0.357
10	rs7341475	A	84	16	2	0.098	0.177	0.254	150	38	1	0.106	0.189	0.390
11	rs8020441	G	55	46	1	0.235	0.359	0.011	97	84	9	0.268	0.393	0.083
12	rs2247572	T	59	34	7	0.240	0.365	0.498	80	89	16	0.327	0.440	0.206
13	rs2616984	G	71	28	1	0.150	0.255	0.327	151	34	5	0.116	0.205	0.082
14	rs2229741	C	40	43	7	0.317	0.433	0.325	82	75	31	0.364	0.463	0.057
15	rs2252521	T	37	47	17	0.401	0.480	0.752	93	69	4	0.232	0.356	0.032

Примечание. MA – минорный (редкий) аллель. Численности генотипов: 11 – гомозиготы по частому аллелю; 12 – гетерозиготы, 22 – гомозиготы по редкому аллелю. MAF – частота минорного аллеля; H_e – ожидаемая гетерозиготность; P – достигнутый уровень значимости соответствия наблюдаемого распределения генотипов ожидаемому при равновесии Харди–Вайнберга. Полужирным шрифтом выделены статистически значимые отклонения от РХВ.

Таблица 3. Анализ ассоциаций генетических маркеров с шизофренией у казахов

№	SNP	Ген	OR	95% CI	P_1	МП χ^2	P_2
1	rs1502844	<i>SLCO6A1</i>	1.08	0.75–1.54	0.671	0.33	0.846
2	rs9960767	<i>TCF4</i>	1.33	0.45–3.85	0.570	0.32	0.568
3	rs2312147	<i>VRK2</i>	0.58	0.38–0.89	0.008	7.58	0.023
4	rs3131296	<i>NOTCH4</i>	0.75	0.37–1.48	0.372	2.01	0.366
5	rs12807809	<i>NRGN</i>	1.09	0.73–1.63	0.664	2.11	0.349
6	rs1572299	<i>TLR4/DBC1</i>	1.35	0.86–2.10	0.168	1.97	0.374
7	rs17594526	<i>TCF4</i>	0.60	0.20–6.48	0.934	0.21	0.645
8	rs1344706	<i>ZNF804A</i>	0.98	0.69–1.40	0.910	0.17	0.920
9	rs16977195	<i>AGBL1</i>	1.65	0.77–3.52	0.163	1.79	0.409
10	rs7341475	<i>RELN</i>	0.92	0.50–1.67	0.768	2.03	0.363
11	rs8020441	<i>ZFP64P1</i>	0.84	0.55–1.27	0.382	3.45	0.178
12	rs2247572	<i>KCNB2</i>	0.65	0.43–0.98	0.030	6.66	0.038
13	rs2616984	<i>CSMD1</i>	1.38	0.81–2.35	0.204	3.80	0.150
14	rs2229741	<i>NRIP1</i>	0.81	0.54–1.20	0.270	4.62	0.099
15	rs2252521	<i>CPVL</i>	1.46	1.01–2.12	0.037	4.77	0.092

Примечание. OR – отношение шансов для минорного аллеля, P_1 – уровень значимости OR; МП χ^2 – критерий максимального правдоподобия χ^2 для сравнения распределения генотипов, P_2 – его уровень значимости. Полужирным шрифтом выделены уровни значимости <0.05.

по сравнению с контрольной группой. Для двух из трех ассоциированных маркеров (SNP генов *VRK2* и *KCNB2*) ассоциация была подтверждена и достоверными различиями в распределении генотипов, оцененными по критерию максимального правдоподобия хи-квадрат. Для rs2252521 гена *CPVL* различия в распределении генотипов носили близкий к статистически достоверному характер ($p = 0.092$). Все три SNP, ассоциированные с ранней шизофренией у казахов, являются некодирующими заменами нуклеотидов, локализованными в интронах.

Анализ межгенных взаимодействий

Для анализа межгенных взаимодействий, оказывающих влияние на фенотипический признак (наличие шизофрении), использовали метод снижения многомерной размерности (MDR) по алгоритму расширенного поиска (exhaustive search algorithm). В результате выявлено восемь возможных сочетаний кумулятивно взаимодействующих локусов. Два из восьми полученных сочетаний показывают высокую кросс-валидационную согласованность (CVC) (10 из 10 непротиворечивых кросс-валидаций), демонстрирующую значимое влияние сочетаний полиморфных маркеров нескольких локусов на фенотип.

Первый набор кумулятивно взаимодействующих вариантов полиморфизма включает четыре SNP (rs1502844, rs2312147, rs8020441, rs2616984). Достигнутый уровень значимости (P) адекватности модели из этих четырех SNP по тесту знаков равен 0.011. Балансовая точность обучающей выборки составила 76.27% при чувствительности 89.15% и специфичности 63.39%. Балансовая точность тестируемых выборок (средняя воспроизводимость по 10 кросс-валидациям) равна 62.29%, чувствительность и специфичность – 70.00 и 56.58% соответственно. Таким образом, первый набор непротиворечивых кросс-валидаций составляют четыре взаимодействующих гена: *SLCO6A1*, *VRK2*, *ZFP64P1* и *CSMD1*, при этом лишь один из них (*VRK2*) демонстрирует статистически значимую ассоциацию с шизофренией у казахов на уровне отдельных SNP.

Второй набор включает шесть однонуклеотидных кумулятивно действующих вариантов полиморфизма и помимо вышеописанных четырех SNP в него входят еще два полиморфных варианта: rs17594526 гена *TCF4* и rs2229741 гена *NR1P1*. Балансовая точность обучающей и тестируемых выборок для этого набора маркеров составила 90.55% (чувствительность 98.95%, специфичность 82.14%) и 53.64% (чувствительность 44.02%, специфичность 65.09%; $P = 0.377$). Для этого сочетания маркеров, несмотря на высокую балансовую точность (>90%) обучающей выборки, наблюдается более слабая классификационная спо-

собность на 10 тестируемых выборках. Вероятнее всего, проблемы небольшого размера выборок, возникающие при множественном тестировании с увеличением числа подгрупп при возрастании количества взаимодействующих генов в обучающей выборке, приводят к снижению чувствительности и потере статистической значимости модели.

ОБСУЖДЕНИЕ

Функции и роль генов, обнаруживших реплицированные ассоциации с шизофренией

В настоящей работе мы выявили ассоциацию шизофрении с ранним началом у казахов Республики Казахстан с маркерами трех генов (*VRK2*, *KCNB2* и *CPVL*), ранее обнаруживших высокодостоверную связь с болезнью или ее эндофенотипами в широкогеномных ассоциативных исследованиях. Ген *VRK2* кодирует белок серин/треонинпротеинкиназу, который принадлежит к группе казеинкиназ I. Эти протеинкиназы фосфорилируют гидроксильную группу в остатках серина или треонина. Они участвуют в контроле цитоплазматических и ядерных процессов, включая репликацию и репарацию ДНК, а также вовлечены в контроль за апоптозом. Ген *VRK2* широко экспрессируется в тканях организма человека в активно делящихся клетках (яичке, лейкоцитах, эмбриональной печени, карциномах). Возможная роль *VRK2* в подверженности к шизофрении, на наш взгляд, может быть связана с тем, что серин/треонинпротеинкиназы связывают белок JIP1, который в нейронных клетках служит фактором антиапоптоза в ответ на стресс и играет роль в развитии аксона [34]. Ассоциация маркеров гена *VRK2* с шизофренией была выявлена в одном из первых GWAS [22] на европеоидах. Позднее ассоциация rs2312147 этого гена с болезнью была реплицирована на китайцах [35]. Независимая репликация ассоциации этого гена у азиатских монголоидов (китайцев и казахов), возможно, демонстрирует расо-специфичный характер ассоциации, связанный, вероятно, с особенностями структуры генофонда или паттернов неравновесия по сцеплению в этом локусе у монголоидов.

В отношении маркеров двух других генов (*KCNB2* и *CPVL*), ассоциацию с которыми мы реплицировали у казахов, механизмы их возможного вовлечения в подверженность к болезни еще менее ясны. Маркеры этих генов были выявлены в недавних GWAS [26, 27] как ассоциированные с когнитивными способностями, являющимися эндофенотипами шизофрении. Репликативные исследования этих маркеров не проводились. Ген *KCNB2* кодирует белок, который является участником потенциал-зависимого калиевого канала. Ген экспрессируется в клетках гладких мышц желудочно-кишечного тракта. Белок, кодируемый

Таблица 4. Генные онтологии 14 генов, в области которых локализованы изучаемые маркеры

Код генной онтологии по Gene Ontology	Молекулярная функция (MF) или биологический процесс (BP)	Вовлеченные гены	<i>P</i>
GO:0004180	Карбоксипептидазная активность (MF)	<i>AGBL1, CPVL</i>	0.028
GO:0008238	Экзопептидазная активность (MF)	<i>AGBL1, CPVL</i>	0.058
GO:0070011	Пептидазная активность, действующая на L-аминокислотные пептиды (MF)	<i>AGBL1, RELN, CPVL</i>	0.064
GO:0010557	Положительная регуляция процессов биосинтеза макромолекул (BP)	<i>TLR4, TCF4, NRIP1</i>	0.067
GO:0008233	Пептидазная активность (MF)	<i>AGBL1, RELN, CPVL</i>	0.069
GO:0031328	Положительная регуляция клеточных процессов биосинтеза (BP)	<i>TLR4, TCF4, NRIP1</i>	0.073
GO:0009891	Положительная регуляция процессов биосинтеза (BP)	<i>TLR4, TCF4, NRIP1</i>	0.075

Примечание. *P* – уровень значимости модифицированного точного теста Фишера, показывающий неслучайность обогащенности данной онтологии изучаемыми генами.

геном *CPVL*, является карбоксипептидазой и имеет сильное сходство по последовательности с карбоксипептидазами серинового типа. Возможно, ассоциация с болезнью маркеров в локусе *CPVL* связана не с этим геном, а с расположенным вблизи локусом *LOC100506497*, который кодирует *miscRNA* с неизвестной функцией.

Биологические процессы, связанные с изучаемыми генами, и генные сети

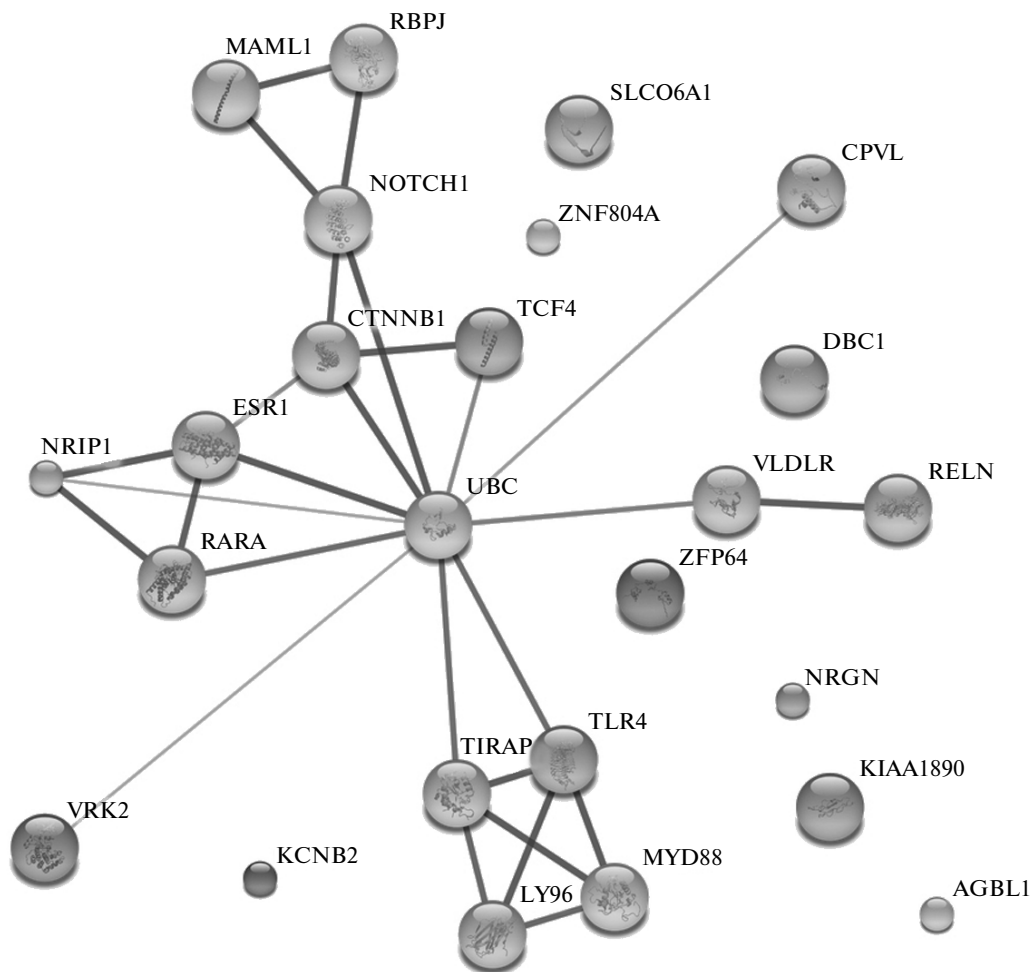
Широкогеномные ассоциативные исследования проводятся в рамках идеологии позиционного картирования и не подразумевают функциональной взаимосвязи ассоциированного маркера с биологическими процессами, лежащими в основе заболевания. Тем не менее в регионе, маркируемом ассоциированным полиморфным участком, должны находиться гены или регуляторные структуры, вовлеченные в метаболические, биохимические или гомеостатические системы, нарушения которых приводят к развитию болезни.

14 генов, в области которых локализованы изучаемые маркеры (см. табл. 1), не связаны какой-то общей функцией или единым биологическим процессом. Функциональная аннотация этих генов с помощью биоинформационного ресурса DAVID выявляет семь генных онтологий (молекулярных функций и биологических процессов), в которые вовлечены как минимум два из исследуемых генов (табл. 4). Фактически эти онтологии сводятся к вариантам одной молекулярной функции (пептидазная активность) и одному биологическому процессу (положительная регуляция процессов биосинтеза). Пептидазная активность характерна для продуктов трех генов (*AGBL1, RELN, CPVL*), а в регуляцию биосинтеза вовлечены белки, кодируемые тремя другими генами (*TLR4, TCF4, NRIP1*). Восемь других генов

не демонстрируют свидетельств общности биологических функций.

Анализ белок-белковых взаимодействий продуктов 14 изучаемых генов, проведенный с помощью ресурса STRING, показывает, что продукты шести из 14 генов могут быть вовлечены в единую взаимосвязанную сеть, включающую также ряд других (не включенных в настоящую работу) генов (рисунок). Белки *VRK2, CPVL, RELN, TLR4, NRIP1* и *TCF4* включены в единую сеть, центральным звеном которой является белок убиквитин С (*UBC*). Сеть построена на основе анализа экспериментальных данных, коэкспрессии, баз данных и других доказательств без использования инструмента “Text mining”, анализирующего соупоминание генов или их продуктов в одном абзаце. Все шесть указанных белков, равно как и *NOTCH1*, продукт гомолога гена *NOTCH4*, способны связываться убиквитином С. Доказательства связи *UBC* с этими белками получены в экспериментальных работах с помощью различных подходов белковой химии [36–40]. Примечательно, что продукты двух из трех генов, ассоциации которых с шизофренией были реплицированы в настоящей работе (*VRK2* и *CPVL*), также входят в сеть убиквитин-связываемых белков.

Убиквитин С в мономерной или полимерной форме связывается с широким спектром белков и пептидов посредством различных лизиновых остатков полиубиквитиновых цепей. Последние, связываясь с целевыми белками, участвуют в выполнении различных функций, включая регуляцию клеточного цикла, деградацию лизосом, разрушение белков, модификацию киназ, эндоцитоз, активацию транскрипционных факторов. Свободный убиквитин также участвует в активации протеинкиназ и передаче сигнала. Возможно, функции *VRK2* как протеинкиназы и *CPVL*



Сеть белок-белковых взаимодействий, включающая продукты 15 исследуемых генов и 10 их генов-партнеров. Размер круга соответствует размеру белка. Толщина линии между белками отражает доверительный уровень (confidence) связи.

как карбоксипептидазы модифицируются или регулируются посредством связи белков с убиквитином С, и именно это обеспечивает их сетевые взаимодействия в патогенезе шизофрении. Интересно также, что продукты трех генов, демонстрирующих эффекты эпистаза при анализе методами снижения размерности (MDR), — *VRK2*, *TCF4* и *NRIP1* — также входят в сеть убиквитин-связываемых белков.

Таким образом, в настоящей работе впервые проведен репликативный анализ 15 однонуклеотидных маркеров, ассоциированных с шизофренией и ее эндофенотипами по данным GWAS, у казахов. Это исследование является первой работой по поиску генетических маркеров шизофрении с ранним началом в монголоидных популяциях. Несмотря на относительно небольшую по численности выборку больных, что связано с редкостью изучаемой формы шизофрении, для трех генетических маркеров была выявлена статистически достоверная ассоциация с болезнью (SNP в генах *VRK2*, *KCNB2* и *CPVL*). Ассоциация

rs2312147 гена *VRK2* с шизофренией ранее была также реплицирована у китайцев, что позволяет рассматривать этот маркер как вероятный расоспецифичный. С помощью методов снижения размерности многофакторных данных были обнаружены две группы из четырех и шести генов, демонстрирующих межгенные эпистатические взаимодействия. Анализ функций изучаемых генов, биологических процессов и генных сетей позволил предположить, что возможная вовлеченность белка *VRK2* в молекулярные механизмы шизофрении может быть связана с его протеинкиназной активностью в тканях нервной системы, а общим процессом, связывающим ряд изучаемых генов в единую сеть, может являться убиквитинилирование их продуктов белком UBC. Представляет интерес дальнейшее изучение ассоциации исследованных генов с разными формами шизофрении, так как особенности психопатологии, течения и динамики формирования исходных состояний этого процесса дают основа-

ния для предположения о неоднородности его этиопатогенеза и генетической компоненты.

Работа выполнена при финансовой поддержке РФФИ (грант № 12-04-00595), 7-й рамочной программы Евросоюза (проект 242257 ADAMS) и Научного фонда Томского государственного университета им. Д.И. Менделеева.

СПИСОК ЛИТЕРАТУРЫ

- Riley B., Kendler K.S. Molecular genetic studies of schizophrenia // *Eur. J. Hum. Genet.* 2006. V. 14. P. 669–680.
- Голумбет В.Е. Генетика шизофрении // *Журнал неврологии и психиатрии им. С.С. Корсакова.* 2003. Т. 103. № 3. С. 58–67.
- Хоменко Н.В. Генетические и средовые факторы в развитии шизофрении // *Мед. журнал.* 2012. № 2. С. 15–18.
- Welter D., MacArthur J., Morales J. et al. The NHGRI GWAS Catalog, a curated resource of SNP-trait associations // *Nucl. Acids Res.* 2014. V. 42. (Database issue). P. D1001-D1006.
- Ioannides J.P., Ntzani E.E., Trikalinos T.A. “Racial” differences in genetic effects for complex diseases // *Nat. Genet.* 2004. V. 36. № 12. P. 1312–1318.
- Степанов В.А. Геномы, популяции, болезни: этническая геномика и персонализированная медицина // *Acta Naturae.* 2010. Т. 2. № 4. С. 18–34.
- Shi Y., Li Z., Xu Q., et al. Common variants on chromosome 6p22.1 are associated with schizophrenia // *Nature.* 2009. V. 460. № 7256. P. 753–757.
- Ikeda M., Aleksic B., Kinoshita Y. et al. Genome-wide association study of schizophrenia in a Japanese population // *Biol. Psychiatry.* 2011. V. 69. № 5. P. 472–478.
- Yamada K., Iwayama Y., Hattori E. et al. Genome-wide association study of schizophrenia in Japanese population // *PLoS One.* 2011. V. 6. № 6. e20468.
- Ma X., Deng W., Liu X. et al. A genome-wide association study for quantitative traits in schizophrenia in China // *Genes Brain Behav.* 2011. V. 10. № 7. P. 734–739.
- Yue W.H., Wang H.F., Sun L.D. et al. Genome-wide association study identifies a susceptibility locus for schizophrenia in Han Chinese at 11p11.2 // *Nat. Genet.* 2011. V. 43. № 12. P. 1228–1231.
- Shi Y., Li Z., Xu Q. et al. Common variants on 8p12 and 1q24.2 confer risk of schizophrenia // *Nat. Genet.* 2011. V. 43. № 12. P. 1224–1227.
- Wong E.H., So H.C., Li M. et al. Common Variants on Xq28 Confer Risk of Schizophrenia in Han Chinese // *Schizophr. Bull.* 2013. Sep. 16. [Epub ahead of print]
- Садуакасова К.З. Психический дизонтогенез и ранняя детская шизофрения в аспекте возрастной периодизации психического развития // *Современные проблемы теоретической и клинической медицины.* Алматы, 2004. С. 203–204.
- Юрьева О.П. О типах дизонтогенеза у детей, больных шизофренией // *Невропатологии и психиатрии.* 1970. Т. 70. Вып. 8. С. 1229–1235.
- Addington A.M., Gornick M., Sporn A.L. et al. Polymorphisms in the 13q33.2 gene G72/G30 are associated with childhood-onset schizophrenia and psychosis not otherwise specified // *Biol. Psychiat.* 2004. V. 55. № 10. P. 976–980.
- Gornick M.C., Addington A.M., Sporn A. et al. Dysbindin (DTNBP1, 6p22.3) is associated with childhood-onset psychosis and endophenotypes measured by the Premorbid Adjustment Scale (PAS) // *J. Autism Dev. Disord.* 2005. V. 35. № 6. P. 831–838.
- Maziade M., Martinez M., Rodrigue C. et al. Childhood/early adolescence onset and adult-onset schizophrenia. Heterogeneity at the dopamine D3 receptor gene // *Br. J. Psychiat.* 1997. V. 170. P. 27–30.
- Sekizawa T., Iwata Y., Nakamura K. et al. Childhood-onset schizophrenia and tryptophan hydroxylase gene polymorphism // *Am. J. Med. Genet. B Neuropsychiatr. Genet.* 2005. V. 136B. № 1. P. 106.
- Пахомова С.А., Коровайцева Г.И., Мончаковская М.Ю. и др. Молекулярно-генетическое исследование шизофрении с ранним началом // *Журнал неврологии и психиатрии им. С.С. Корсакова.* 2010. Т. 110. № 2. С. 66–69.
- Пузырев В.П., Степанов В.А. Патологическая анатомия генома человека. Томск: СГТ, 1997. 242 с.
- Stefansson H., Ophof R.A., Steinberg S. et al. Common variants conferring risk of schizophrenia // *Nature.* 2009. V. 460. № 7256. P. 744–747.
- Purcell S.M., Wray N.R., Stone J.L. et al. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder // *Nature.* 2009. V. 460. № 7256. P. 748–752.
- O’Donovan M.C., Craddock N., Norton N. et al. Identification of loci associated with schizophrenia by genome-wide association and follow-up // *Nat. Genet.* 2008. V. 40. № 9. P. 1053–1055.
- Sullivan P.F., Lin D., Tzeng J.Y. et al. Genomewide association for schizophrenia in the CATIE study: results of stage 1 // *Mol. Psychiatry.* 2008. V. 13. № 6. P. 570–584.
- Shifman S., Johannesson M., Bronstein M. et al. Genome-wide association identifies a common variant in the reelin gene that increases the risk of schizophrenia only in women // *PLoS Genet.* 2008. V. 4. № 2. P. e28.
- Cirulli E.T., Kasperaviciute D., Attix D.K. et al. Common genetic variation and performance on standardized cognitive tests // *Eur. J. Hum. Genet.* 2010. V. 18. № 7. P. 815–820.
- Need A.C., Attix D.K., McEvoy J.M. et al. A genome-wide study of common SNPs and CNVs in cognitive performance in the CANTAB // *Hum. Mol. Genet.* 2009. V. 18. № 23. P. 4650–4661.
- Huang D.W., Sherman B.T., Lempicki R.A. Systematic and integrative analysis of large gene lists using DAVID Bioinformatics Resources // *Nature Protoc.* 2009. V. 4. № 1. P. 44–57.
- Kanehisa M., Goto S., Sato Y. et al. KEGG for integration and interpretation of large-scale molecular datasets // *Nucl. Acids Res.* 2012. V. 40. P. D109–D114.
- Franceschini A., Szklarczyk D., Frankild S. et al. STRING v. 9.1: protein-protein interaction networks,

- with increased coverage and integration // *Nucl. Acids Res.* 2013. V. 41 (Database issue). P. D808–D815.
32. *The 1000 Genomes Project Consortium.* An integrated map of genetic variation from 1092 human genomes // *Nature.* 2012. V. 491. P. 56–63.
 33. *The International HapMap 3 Consortium.* Integrating common and rare genetic variation in diverse human populations // *Nature.* 2010. V. 467. P. 52–58.
 34. *Steinberg S., de Jong S., Irish Schizophrenia Genomics Consortium, et al.* Common variants at *VRK2* and *TCF4* conferring risk of schizophrenia // *Hum. Mol. Genet.* 2011. V. 20. № 20. P. 4076–4081.
 35. *Li M., Wang Y., Zheng X. et al.* Meta-analysis and brain imaging data support the involvement of *VRK2* (rs2312147) in schizophrenia susceptibility // *Schizophr. Res.* 2012. V. 142. P. 200–205.
 36. *Kuhn M., Simonovic M., Roth A. et al.* *AIP4/Itch* regulates Notch receptor degradation in the absence of ligand // *PLoS One.* 2008. V. 3. № 7. P. e2735.
 37. *Yamada M., Ohnishi J., Ohkawara B. et al.* NARF, an nemo-like kinase (NLK)-associated ring finger protein regulates the ubiquitylation and degradation of T cell factor/lymphoid enhancer factor (TCF/LEF) // *J. Biol. Chem.* 2006. V. 281. № 30. P. 20749–20760.
 38. *Altun M., Kramer H.B., Willems L.I. et al.* Activity-based chemical proteomics accelerates inhibitor development for deubiquitylating enzymes // *Chem. Biol.* 2011. V. 18. № 11. P. 1401–1412.
 39. *Danielsen J.M., Sylvestersen K.B., Bekker-Jensen S. et al.* Mass spectrometric analysis of lysine ubiquitylation reveals promiscuity at site level // *Mol. Cell. Proteomics.* 2011. V. 10. № 3. P. M110.003590.
 40. *Afrazi A., Sodhi C.P., Good M. et al.* Intracellular heat shock protein-70 negatively regulates TLR4 signaling in the newborn intestinal epithelium // *J. Immunol.* 2012. V. 88. № 9. P. 4543–4557.

Replicative Study of Susceptibility to Childhood-Onset Schizophrenia in Kazakhs

V. A. Stepanov^{a,b}, A. V. Bocharova^b, K. Z. Saduakassova^c, A. V. Marusin^b,
L. A. Koneva^b, K. V. Vagaitseva^{a,b}, and G. S. Svyatova^d

^a National Research Tomsk State University, Tomsk, 634050 Russia

^b Research Institute for Medical Genetics, Russian Academy of Medical Sciences, Tomsk, 634050 Russia
e-mail: vadim.stepanov@medgenetics.ru

^c Asfendiyarov Kazakh National Medical University, Ministry of Health of the Republic of Kazakhstan, Almaty, 050012 Kazakhstan

^d Scientific Center for Obstetrics, Gynecology and Perinatology, Ministry of Health of the Republic of Kazakhstan, Almaty, 050020 Kazakhstan

This paper reports the results of replicative analysis of associations of 15 SNPs in a region of 14 genes previously identified in genome-wide association studies (GWAS) with early-onset schizophrenia in Kazakhs. An association of early-onset schizophrenia with genetic markers in three genes (*VRK2*, *KCNB2*, and *CPVL*) was found. An association of rs2312147 in the *VRK2* gene with schizophrenia was also previously reported in the Chinese population, so this marker may be considered as possibly race-specific. Two groups consisting of four and six genes demonstrating intergenic epistatic interactions were revealed by multifactor dimensionality reduction methods. The gene ontologies of 14 studied genes were reduced to variants of one molecular function (peptidase activity) and one biological process (positive regulation of biosynthesis processes). Bioinformatic analysis of the protein–protein interactions of products of the genes under study demonstrates that the products of six out of 14 genes may be involved in a single interrelated network, the major connecting link of which is represented by their ubiquitination by the UBC protein.

English translation of paper is published in “Russian J. Genetics” (2015, vol. 51, no. 2), www.maik.ru.