

*На правах рукописи*



СОКОЛОВА ЕКАТЕРИНА АЛЕКСЕЕВНА

**ИССЛЕДОВАНИЕ ГЕНЕТИЧЕСКОЙ  
ПРЕДРАСПОЛОЖЕННОСТИ К РАССЕЯННОМУ  
СКЛЕРОЗУ У ЖИТЕЛЕЙ РФ С ПРИМЕНЕНИЕМ  
СИСТЕМНОГО АНАЛИЗА**

03.01.03 – молекулярная биология

03.01.07 – молекулярная генетика

Автореферат

диссертации на соискание ученой степени

кандидата биологических наук

Новосибирск - 2016

Работа выполнена в Федеральном государственном бюджетном учреждении науки  
Институте химической биологии и фундаментальной медицины СО РАН

Научный руководитель: к.б.н. **Филипенко Максим Леонидович**

Официальные оппоненты: д.б.н., профессор **Гуляева Людмила Федоровна**  
ФГБНУ «Научно-исследовательский институт  
молекулярной биологии и биофизики» СО РАН

к.б.н., н.с. **Бабушкина Надежда Петровна**  
ФГБНУ "Научно-исследовательский институт  
медицинской генетики"

Ведущая организация: ФГБУ "НИИ фундаментальной и клинической  
иммунологии" СО РАН (г. Новосибирск)

Защита состоится «\_\_» \_\_\_\_\_ 201 г. в «\_\_» часов на заседании  
диссертационного совета Д 003.045.01 на базе Федерального государственного  
бюджетного учреждения науки Институт химической биологии и фундаментальной  
медицины СО РАН по адресу 630090, г. Новосибирск, проспект акад. Лаврентьева, 8.

С диссертацией можно ознакомиться в библиотеке Федерального государственного  
бюджетного учреждения науки Институт химической биологии и фундаментальной  
медицины СО РАН и на сайте [www.niboch.nsc.ru](http://www.niboch.nsc.ru).

Автореферат разослан «\_\_» \_\_\_\_\_ 2015 г.

Ученый секретарь диссертационного совета  
к.х.н., доцент



Коваль В. В.

## ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

**Актуальность работы.** Рассеянный склероз (РС) – тяжелое аутоиммунное заболевание нервной системы. Распространенность РС неуклонно растет, более того - наблюдается омоложение заболевания. РС является основной причиной стойкой инвалидизации лиц молодого возраста. К настоящему времени накоплено много информации об ассоциации генетических полиморфных локусов с РС. Большая часть ее разрозненна и противоречива. Ассоциация большинства полиморфных локусов (SNPs) требует верификации путем репликации на репрезентативных выборках разных этнических групп. В то же время для ряда SNPs выполнено много исследований, но на малых выборках, что не позволяет сделать окончательный вывод о роли SNP в патогенезе РС без проведения мета-анализа. Идентификация новых генетических предикторов, ранее не выявленных, требует применения методов системно-биологического анализа путем построения молекулярных сигнальных путей патогенеза РС. Формирование репрезентативной выборки больных РС, проживающих в разных регионах РФ, до настоящего исследования не проводилось. Наличие такой выборки позволило реплицировать ранее выявленные ассоциации SNPs с РС у европеоидов Америки и Европы с высокой мощностью, а также проанализировать наличие различий генетической предрасположенности к РС у людей, проживающих в различных географических регионах РФ.

**Целью** настоящей работы являлось изучение роли полиморфизма отдельных генов в предрасположенности к развитию РС в репрезентативной выборке русской этнической группы с применением современных методов анализа (мета-анализа, анализа молекулярных путей, сетевого анализа).

Для достижения поставленной цели были поставлены следующие **задачи**:

1. Реплицировать ассоциацию с РС полиморфного локуса rs10492972 гена *KIF1B* в русской этнической группе и провести мета-анализ результатов нашего исследования с ранее проведенными;

2. Проанализировать ассоциацию полиморфного локуса -308G>A гена *TNFA* с РС в русской этнической группе, провести мета-анализ результатов нашего исследования с ранее проведенными. Исследовать влияние неравновесия по сцеплению этого полиморфного локуса с HLA-DRB1\*15 аллелем на ассоциацию с РС;

3. Реплицировать ассоциацию полиморфизма гена *CD40* с РС, провести картирование потенциального функционального аллеля и провести теоретический функциональный анализ района гена *CD40*, его содержащего;

4. На основании анализа сигнальных путей, включающих продукт гена *CD40*, и интегральных данных исследований генетики РС провести выбор гена для проверки гипотезы их ассоциативного генетического взаимодействия;

5. Исследовать ассоциацию РС с полиморфным локусом rs2234767 гена *FAS* в русской этнической группе; на основании анализа сигнальных путей, включающих продукт гена *FAS* и интегральных данных исследований генетики РС провести выбор гена для проверки гипотезы их ассоциативного генетического взаимодействия.

**Научная новизна.** Впервые на выборке пациентов с РС русской этнической группы, проживающих в разных регионах России, была реплицирована ассоциация SNPs генов *KIF1B*, *TNFA*, *CD40*, *IL2RA*, *TRAIL*, *TRAIL-R2*; для полиморфного локуса rs2234767 гена *FAS* ассоциация была исследована впервые в мире. С РС у жителей РФ ассоциированы SNPs rs6074022 и rs1883832 гена *CD40*, rs12722489 и rs2104286 гена *IL2RA*, rs2234767 гена *FAS* и rs4894559 гена *TRAIL*.

По результатам мета-анализа аллель rs10492972[C] гена *KIF1B* является статистически значимым протективным фактором развития РС, аллель rs6074022[C] гена *CD40* – фактором риска РС, в то время как SNP -308G>A гена *TNFA* по результатам мета-анализа не ассоциирован с заболеванием.

Впервые на примере русской этнической группы нами было показано, что ассоциация с РС SNP -308G>A гена *TNFA*, идентифицируемая в выборках малого размера, является наведенной и обусловлена неравновесием по сцеплению с аллелем HLA-DRB1\*15.

Показано, что функциональный полиморфный локус гена *CD40* лежит выше старта трансляции гена. Впервые в мире *in silico* выявлено, что сайты связывания транскрипционных факторов VDR, STAT3, IRF-4 – важные регуляторные звенья гена *CD40*, нарушения в которых, вероятно, будут приводить к повышенному риску развития РС. Впервые в мире *in vitro* показано, что аллельные варианты функционального полиморфного локуса в промоторном районе гена *CD40*, который маркируется SNP rs6074022, не приводят к аллельному дисбалансу.

Был проведен комплексный анализ всех статей, посвященных поиску генетических факторов риска РС, на основании которого был составлен список генов, ассоциированных с РС, и проведен анализ молекулярных путей, вовлеченных в патогенез РС.

**Теоретическая и практическая значимость.** Настоящая работа вносит вклад в понимание молекулярных основ патогенеза РС. Подтверждена роль гена *CD40*, проведен поиск месторасположения функционального полиморфного локуса и анализ промоторного региона гена. Полученные данные в будущем могут стать базисом для разработки таргетной терапии РС. Показана первостепенная роль аллеля HLA-DRB1\*15 в локусе HLA-TNF $\alpha$ , что говорит о неперспективности дальнейших ассоциативных исследований гена *TNF $\alpha$*  с РС. Проведенный анализ молекулярных путей, вовлеченных в патогенез РС, выявил кандидатные пути, перспективные для дальнейших экспериментальных исследований как ассоциативных, так и направленных на поиск мишеней для разработки целевого лечения РС.

**Апробация работы и публикации.** По материалам диссертации опубликовано 11 статей в рецензируемых научных журналах. Результаты работы были представлены на конференциях: «Рассеянный склероз: 30 лет спустя» (Новосибирск, Россия, 2011); Международная конференция «Постгеномные технологии для биомедицины» (РВТ-2012) (Новосибирск, Россия, 2012); ECTRIMS 10-13 (Лион, Франция, 2012); 38<sup>th</sup> FEBS Congress “Mechanisms in Biology” (Санкт-Петербург, Россия, 2013).

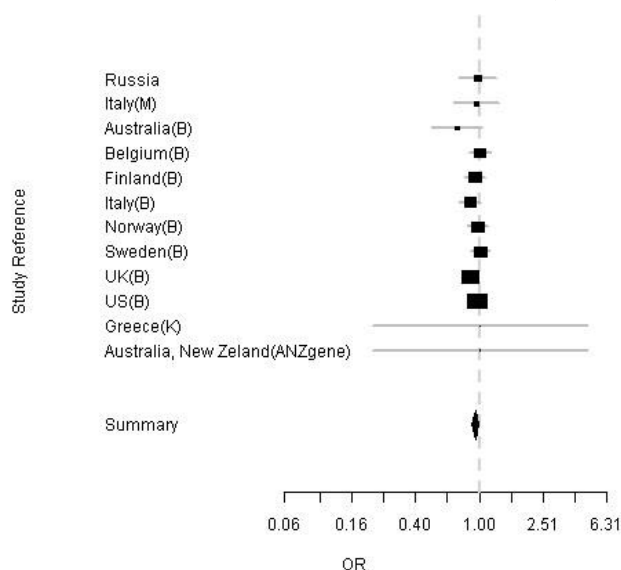
**Структура и объем работы.** Диссертация состоит из введения, обзора литературы обсуждения результатов, экспериментальной части, выводов и списка цитируемой литературы (268 наименований). Работа изложена на 225 страницах, содержит 41 таблицу, 32 рисунка, 27 формул и 1 приложение.

## ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

### 2.1 Ассоциация полиморфного локуса rs10492972 гена *KIF1B* с РС

Нами были определены генотипы SNP rs10492972 гена *KIF1B* в группе РС (n=833) и контрольной группе (n=689). Частоты встречаемости генотипов в контрольной группе статистически значимо не отличались от таковых в группе пациентов с РС (OR=0.99, 95%С.И.=[0.84-1.16], *P* =0.88). Мы провели объединение результатов нашего исследования с ранее опубликованными и выполнили мета-анализ. Тест на гетерогенность значимых различий между исследованиями не выявил (*P*=0.81). По результатам мета-анализа частота аллеля С выше в

контрольной группе (суммарное OR=0.95 (95%С.І.:0.90-0.99) с уровнем статистической значимости  $P=0.02$  (Рис. 1).



**Рис. 1.** Мета-анализ. А-группа из исследования Aulchencko с соавт.; В-группа из исследования Booth с соавт.; М - группа из исследования Martinelli-Boneschi с соавт.; К - группа из исследования Koutsis с соавт.; ANZgene – группа из исследования Booth с соавт.

## 2.2 Природа ассоциации полиморфного локуса -308G->А гена *TNFA* с РС

### 2.2.1 Ассоциация между SNP -308G->А гена *TNFA* и РС у жителей РФ

Для выявления наличия или отсутствия ассоциации SNP -308G->А гена *TNFA* с развитием РС у жителей РФ нами были определены генотипы у 1728 пациентов с РС и 1115 человек контрольной группы. Согласно результатам логистического регрессионного анализа SNP -308G>А не ассоциирован с РС у жителей РФ (OR=1.17, 95%С.І.=[0.99-1.39]  $P=0.07$ ).

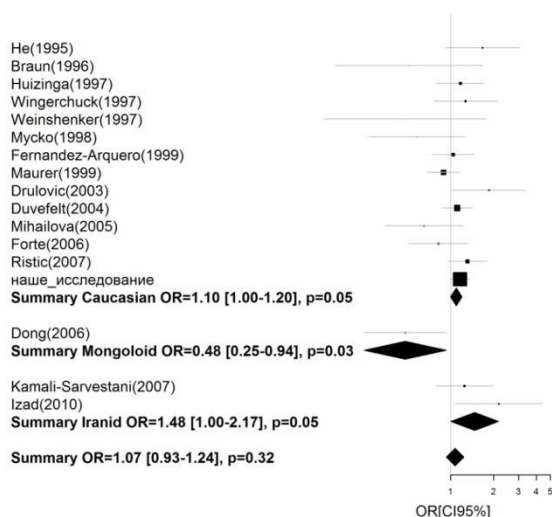
В виду потенциальной стратификации группы РС и контрольная были разбиты на три подгруппы каждая: Центральный федеральный округ (ЦФО), Сибирский федеральный округ (СФО) и Дальневосточный федеральный округ (ДВФО). Нами был проведен аналогичный регрессионный анализ. Ассоциация была выявлена только для жителей СФО (OR=1.36, 95%С.І.=[1.11-1.66]  $P=0.003$ ).

Привлекает внимание, что аллелем риска вопреки ожиданиям стал не редкий аллель А, который согласно исследованиям повышает экспрессию гена *TNFA*, а частый – G. Таким образом, целесообразно суммировать полученные к настоящему времени различными научными коллективами результаты, включая текущее исследование.

### 2.2.2 Мета-анализ ассоциации SNP -308G->А гена *TNFA* с развитием РС

Нами был проведен поиск всех ранее опубликованных ассоциативных исследований SNP -308G->А гена *TNFA* с развитием РС. Из найденных 16 исследований содержало всю необходимую для проведения мета-анализа информацию. Суммарно 4036 пациентов, страдающих РС, и 3879 здоровых волонтера были включены в мета-анализ. Согласно результатам мета-анализа G

аллель не ассоциирован с РС: объединенное OR=1.13 (95%С.И.: 0.95–1.34),  $P=0.18$  (Рис. 2).



**Рис. 2.** «Лесной» график результатов мета-анализа нашего исследования с ранее опубликованными

«Лесной» график не имеет общего тренда в ту или иную сторону от оси  $OR=1$ . Возможны две причины. Первая – малый размер групп сравнения входящих в мета-анализ исследований и как следствие малая статистическая мощность, что привело к широкому разбросу оценки  $OR_G$ . Вторая причина - не смотря на изначальную гипотезу о том, что функциональный полиморфный локус находится в промоторном районе гена *TNFA*, данная замена не является функциональной, но является маркерной для другой замены, находящейся с ней в неравновесии по сцеплению. Ген *TNFA* расположен в области главного комплекса гистосовместимости (МНС). Отличительной чертой МНС является сильное неравновесие по сцеплению между аллелями генов. МНС регион содержит HLA-DRB1 ген, DRB1\*1501 аллель которого был определен как аллель риска развития РС, и, возможно, ассоциация -308G>A локуса гена *TNFA* является наведенной именно этим аллелем.

### 2.2.3 Ассоциация между HLA-DRB1 и РС у жителей РФ

Для выявления является ли ассоциация -308G>A гена *TNFA* наведенной от локуса HLA-DRB1, мы провели репликативное исследование ассоциации с РС аллелей гена HLA-DRB1 на 408 жителях СФО и 627 жителях ЦФО методом секвенирования по Сенгеру (Таблица 1).

**Таблица 1.** Частоты аллелей HLA-DRB1 и ассоциативный тест с РС.

DRB1 аллель	СФО	ЦФО	Объединенное значение OR, 95%CI, p
	OR, 95%CI, p	OR, 95%CI, p	
*01	0.36 [0.19-0.71], 0.003	не значимо	0.56 [0.39-0.81], 0.002
*14	не значимо	не значимо	0.30 [0.10-0.90], 0.03
*15	2.34 [1.55-3.53], $5 \times 10^{-5}$	1.65 [1.13-2.40], 0.009	1.93 [1.47-2.55], $3 \times 10^{-6}$

В СФО частота HLA-DRB1\*15 аллеля статистически значимо выше в группе РС, чем в контрольной (29.3% против 15.2%, OR=2.34 (95%С.І.=1.55-3.53),  $P=5 \times 10^{-5}$ ). Частота аллеля HLA-DRB1\*01, напротив, была выше в контрольной группе (6.1% против 14.1%, OR=0.36 (95%С.І.=0.19-0.71),  $P=0.003$ ). Частоты остальных аллелей в СФО статистически значимо не отличались. В ЦФО единственное статистически значимое различие в частотах аллелей между группами РС и контрольной было выявлено для HLA-DRB1\*15 (25.2% против 17.5%, OR=1.65 (95%С.І.=1.13-2.40),  $P=0.009$ ). По результатам мета-анализа в целом по этим двум округам РФ ассоциированы с РС оказались три аллеля: \*01, \*14 и \*15. Однако если HLA-DRB1\*15 аллель определен как фактор риска развития РС (OR=1.93 (95%С.І.= 1.47-2.55),  $P=3 \times 10^{-6}$ , то \*01 и \*14 были определены как протективные факторы (OR=0.56 (95%С.І.=0.39-0.81),  $P=0.002$  и OR=0.30 (95%С.І.= 0.10-0.90),  $P=0.03$ , соответственно).

#### 2.2.4 Неравновесие по сцеплению между rs1800629 (*TNFA* -308G->A) и HLA-DRB1

HLA-DRB1\*15 ассоциирован с развитием РС у жителей РФ. Для проверки гипотезы о том, что ассоциация -308G>A гена *TNFA* у жителей СФО может быть наведенной от HLA-DRB1\*15, мы оценили неравновесие по сцеплению между аллелями локуса HLA-DRB1 и *TNFA* (Таблица 2).

**Таблица 2.** Неравновесие по сцеплению между rs1800629 (*TNFA* -308G->A) и HLA-DRB1

DRB1 аллель	Сибирский федеральный округ				Центральный федеральный округ			
	rs1800629 ( <i>TNFA</i> )				rs1800629 ( <i>TNFA</i> )			
	[A]	[G]	D' <sub>a</sub>	p	[A]	[G]	D' <sub>a</sub>	p
*15	0.002	0.184	0.81( $r^2=0.02$ )	0.0002	0.011	0.228	0.41 ( $r^2=0.007$ )	0.003

Для СФО нами было выявлено статистически более частым носительство гаплотипа HLA-DRB1\*15-rs1800629[G] по сравнению с HLA-DRB1\*15-rs1800629[A] (0.184 против 0.002,  $P=0.0002$ ). Ориентируясь на параметр D' можно заключить, что в СФО аллели HLA-DRB1\*15-rs1800629[G] находятся в неравновесии по сцеплению, в то время как в ЦФО – нет (D'=0.41). Суммарно, ассоциация SNP -308G>A гена *TNFA* может быть наведенной от локуса HLA-DRB1\*15.

### 2.3 Роль SNPs гена CD40 в развитии рассеянного склероза

#### 2.3.1 Ассоциация SNP rs6074022 гена CD40 с РС

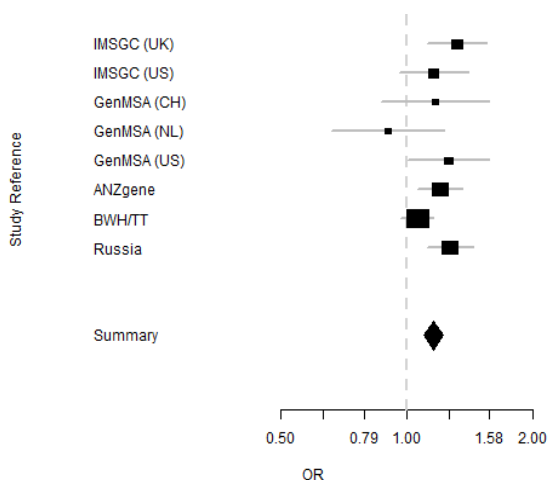
Нами была реплицирована ассоциация с РС rs6074022 на выборке жителей РФ (1679 пациентов с РС и 879 человек контрольной группы). Аллель



C полиморфного локуса rs6074022 ассоциирован с РС (OR=1.27, 95%С.I.=[1.12-1.45]  $P=0.0003$ ).

### 2.3.2 Мета-анализ ассоциации SNP rs6074022 и развития РС

Мы также провели мета-анализ ассоциации SNP rs6074022 и развития РС, включающий результаты нашего исследования и всех ранее опубликованных исследований. Согласно результатам мета-анализа, которые приведены в виде лесного графика на рисунке 3, суммарный OR для всех исследований составил 1.17 (95%С.I.=1.10–1.23) с уровнем статистической значимости  $P=2.24 \times 10^{-12}$ . Тест на гетерогенность не выявил значимых различий между исследованиями ( $\chi^2$  (df=7)=12.16,  $P=0.10$ ).



**Рис. 3** - Мета-анализ ассоциации между аллелем С SNP rs6074022 и РС

Существенно отметить, что полученные нами результаты подтверждают ассоциацию маркерного локуса rs6074022 с развитием РС на полногеномном уровне статистической значимости ( $P=2.24 \times 10^{-12}$ ). Также очень важно, что во всех исследованиях виден тренд того, что именно аллель С SNP rs6074022 является аллелем риска. Это также является показателем того, что функциональный полиморфный локус находится в сильном неравновесии по сцеплению с rs6074022 либо сам маркерный SNP rs6074022 и является функциональным.

### 2.3.3 Картирование функционального полиморфного локуса в гене CD40

На основании литературных данных нами были выбраны еще три SNP с целью картирования функционального полиморфного локуса, отвечающего за ассоциацию с РС: rs1883832, rs1535045, rs11086996. Мы определили их генотипы у 1679 пациентов с РС и 879 человек контрольной группы. Один SNP из трех, rs1883832, статистически значимо ассоциирован с развитием РС (Т аллель риска, OR=1.20, 95%С.I.=1.05-1.38,  $P=7 \times 10^{-3}$ ). Ассоциации двух других SNPs, rs1535045 и rs 11086998, с развитием РС не выявлено. Нами было

оценено неравновесие по сцеплению между исследуемыми SNPs гена CD40. Все исследованные SNPs находятся в неравновесии по сцеплению. С помощью теста отношения правдоподобия (LRT) нами было показано, что маркер rs6074022 находится в большем неравновесии по сцеплению с функциональным полиморфным локусом или им является. Ассоциация же rs1883832 с развитием РС является наведенной, она и слабее по результатам ассоциативного анализа. Скорее всего, функциональный полиморфный локус расположен в промоторе гена CD40.

#### 2.3.4 Анализ сайтов связывания транскрипционных факторов промоторного региона гена CD40

Мы исследовали, какие антитела, специфичные к ТФ, связываются с промоторным регионом гена CD40. Нами было выявлено три региона в промоторе гена CD40, в которых находится максимальное количество «пиков» антител для ТФ. Вероятнее всего, что первые два региона соответствуют энхансерам, третий регион совпадает со стартом транскрипции гена. Выявленные «пики» к антителам к ТФ были сопоставлены с таблицей SNPs данного региона. Затем выявленные три региона были нами проанализированы на наличие сайтов связывания ТФ *in silico*. Следующим нашим шагом был отбор ТФ, чьи сайты связывания значимо менялись при замене в полиморфном локусе. Из них было выбрано три ТФ перспективных для дальнейшего изучения их влияния на предрасположенность к РС, реализуемую через изменения в промоторном регионе гена CD40: VDR и STAT3, IRF-4.

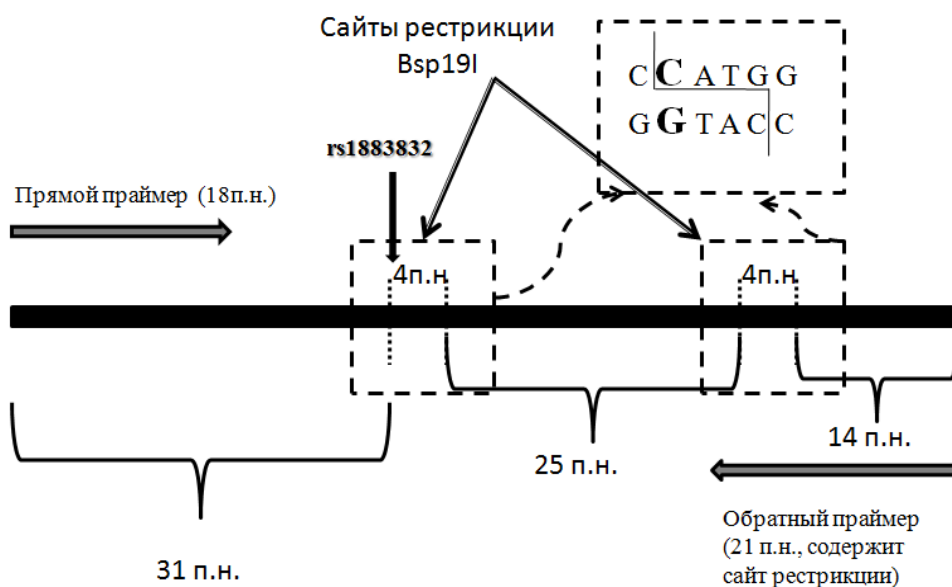
#### 2.3.5 Аллельный дисбаланс гена CD40

Аллельным дисбалансом называют различие в экспрессии двух аллелей одного гена. Так как по результатам нашего исследования функциональный полиморфный локус расположен в промоторной области гена и вероятно приводит к замене в цис-регуляторном элементе, поэтому экспериментальное исследование аллельного дисбаланса включало несколько этапов: 1) выбор SNP в промоторной области гена, который предположительно маркирует функциональный полиморфный локус, изменяющий силу связывания с ТФ, вследствие чего изменяется экспрессия сцепленного с ним аллеля гена; 2) выбор SNP в начале гена, по которому будет оценено количество нарабатываемой мРНК; 3) измерение уровня экспрессии мРНК, нарабатываемой с разных копий гена, в модельных лабораторных образцах (культурах клеток); 4) анализ различий количества мРНК разных аллельных

вариантов гена. В качестве SNP в промоторной области нами был выбран rs6074022, так как нами было убедительно показано, что он маркирует функциональный локус. В качестве SNP, по которому мы оценивали уровень экспрессии мРНК с разных аллельных вариантов, был выбран rs1883832. В качестве модельных лабораторных образцов для измерения уровня экспрессии нами были выбраны мононуклеары крови, полученной от пациентов, страдающих аутоиммунными заболеваниями (n=23), а также пациентов контрольной группы (n=16). Далее мононуклеары делили на пять аликвот, четыре подвергались культивированию с митогеном или без в течение 24 и 48 часов с последующим выделением мРНК и ДНК, из пятой аликвоты мРНК и ДНК была выделена без культивирования.

#### 2.3.5.1 Анализ аллельного дисбаланса методом ПЦР-ПДФ

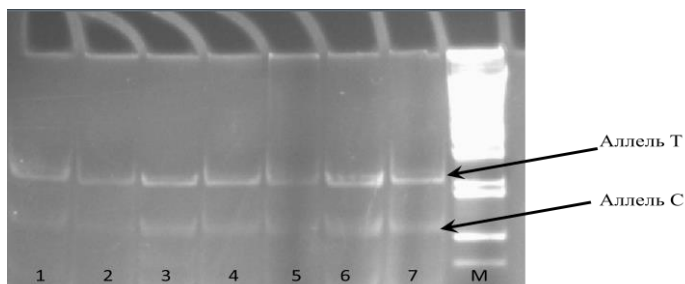
Последовательность вокруг SNP rs1883832 (C->T) в случае отсутствия замены соответствует сайту узнавания эндонуклеазы рестрикции Bsp19I: CCATGG (Рис. 4.)



**Рис. 4.** Продукты ПЦР-ПДФ анализа с указанием длин получаемых фрагментов

Это позволяет оценить соотношение аллелей в кДНК исследуемых образцов, посредством ее амплификации с праймерами, фланкирующими SNP rs1883832, с последующей обработкой эндонуклеазой рестрикции Bsp19I и визуализацией продуктов с помощью электрофореза в акриламидном геле. На рисунке 4 схематически изображены продукты ПЦР-ПДФ анализа с указанием длин получаемых фрагментов в п.н.

Нами были проанализированы все образцы кДНК. Пример электрофореграммы приведен на рисунке 5.



**Рис. 5.** Электрофореграмма продуктов ПЦР-ПДРФ анализа аллельного дисбаланса. М – маркер длин фрагментов, 1-7 – результаты ПЦР-ПДРФ анализа различных образцов.

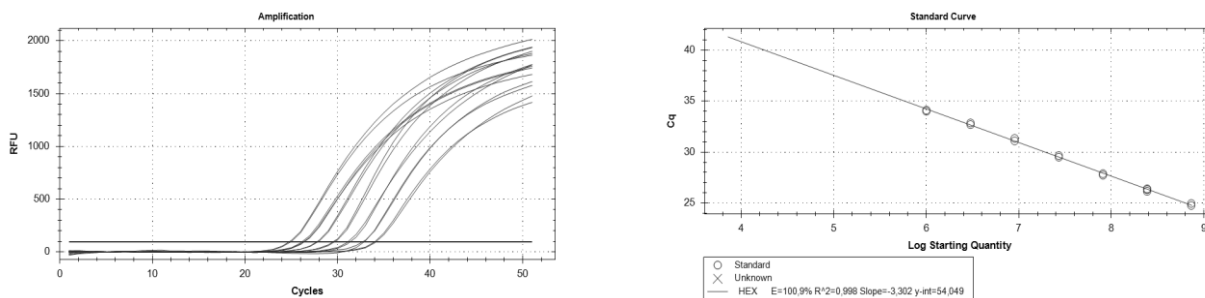
Далее уровень свечения каждого фрагмента гидролиза был оценен с помощью программы GelDoc (BioRad, США) с поправкой на свечение фона. Уровень свечения был поделен на расчетное количество пар нуклеотидов во фрагменте, после чего условные единицы свечения аллеля Т и аллеля С были сравнены для каждого исследованного образца. Для каждой из 5-ти групп аликвот вычисляли среднюю величину аллельного дисбаланса, рассчитанную по формуле 1:

$$\text{аллельный дисбаланс} = \frac{\text{условная единица свечения аллеля Т}}{\text{условная единица свечения аллеля С}} \div 0.57, \quad (1)$$

где 0.57 – это коэффициент несимметричности ПЦР, вычисленный нами эмпирически. Для всех пяти групп данная величина равнялась 1.3. Однако данное отличие в экспрессии аллелей вряд ли имеет биологический эффект.

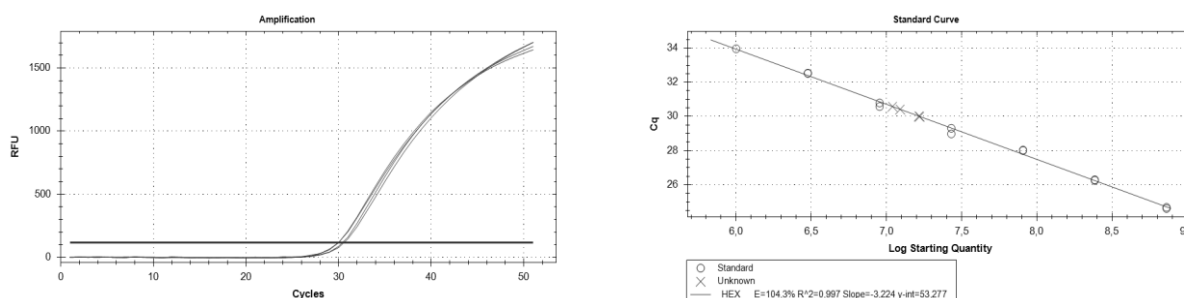
#### *2.3.5.2 Анализ аллельного дисбаланса методом ПЦР в режиме реального времени*

Уровень свечения флюорофора при выходе образца на плато при ПЦР в режиме реального времени коррелирует с исходным количеством детектируемого им аллеля в образце. Следовательно, разница в уровнях свечения может свидетельствовать о различном соотношении аллелей Т и С в образцах кДНК. Для количественной оценки соотношения аллелей в образцах кДНК нами были смоделированы образцы ДНК с заведомо известным соотношением аллелей Т и С SNP rs1883832. Для этого нами были взяты два образца геномной ДНК одинаковой концентрации: один образец – гомозигота СС полиморфного локуса rs1883832, второй – гомозигота ТТ. Концентрация ДНК в образцах была определена методом количественной ПЦР в режиме реального времени с использованием TaqMan-зонда путем построения калибровочной кривой (Рис. 6).



**Рис. 6.** Калибровочная кривая для определения концентрации ДНК в образцах. Последовательные разведения референтного образца в 3 раза.

После определения концентрации ДНК в каждом образце, путем серии последовательных разведений концентрации были выровнены (Рис. 7).



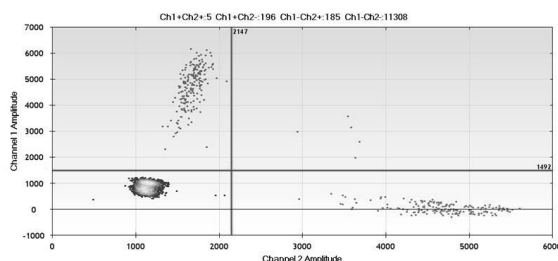
**Рис. 7.** Выравнивание концентрации ДНК образцов, гомозиготных по локусу rs1883832: C/C и T/T.

Далее образцы были смешаны в разных соотношениях так, чтобы в конечных образцах количество аллеля Т равнялось: 0%, 12,5%, 25%, 37,5%, 50%, 62,5%, 75%, 87,5%, 100%. Полученные образцы были использованы в качестве стандарта при измерении количества аллеля Т в исследуемых образцах кДНК. Нами были проанализированы все образцы кДНК. В качестве контроля нами так же были проанализированы образцы геномной ДНК носителей гетерозиготного варианта полиморфного локуса rs1883832. Согласно полученным измерениям в исследованных нами образцах доля аллеля Т была равна 52.6%, и, следовательно, аллельного дисбаланса нет.

### 2.3.5.3 Анализ аллельного дисбаланса методом цифровой ПЦР (Digital PCR)

Третьим методом оценки аллельного дисбаланса была цифровая ПЦР. Принцип метода заключается в том, что исходная ПЦР смесь разбивается на большое количество капель, таким образом, чтобы в каждую каплю попала максимум одна копия исходного образца ДНК. После этого в каждой капле проводится ПЦР, с последующим съемом флюоресцентного сигнала с каждой капли. Далее с помощью программного обеспечения производится подсчет капель с разными флюоресцентными сигналами. Применительно к нашей задаче – оценке аллельного дисбаланса в образце кДНК – соотношение капель с

флуоресцентным сигналом с аллеля Т SNP rs1883832 и аллеля С SNP rs1883832 будет отражать исходное соотношение аллелей в образце кДНК. Нами были проанализированы все образцы кДНК. В качестве контроля, как и в предыдущих двух методах, нами были проанализированы образцы геномной ДНК носителей гетерозиготного варианта SNP rs1883832. Пример соотношения капель с флуоресцентным сигналом от аллелей С и Т приведен на рисунке 8.



**Рис. 8** - Пример соотношения аллелей в анализируемом образце кДНК

Согласно полученным измерениям в исследованных нами образцах аллельного дисбаланса нет. Суммируя результаты всех трех использованных нами методов: ПЦР-ПДРФ, ПЦР в режиме реального времени и цифровой ПЦР – можно заключить, что аллельные варианты функционального полиморфного локуса в промоторном районе гена *CD40*, который маркируется SNP rs6074022, либо не приводят к аллельному дисбалансу, либо аллельный дисбаланс реализуется в узкоспециализированных линиях клеток.

## 2.4 Анализ молекулярных путей (pathways), вовлеченных в патогенез РС

### 2.4.1 Аналитический обзор литературы

Комплексный анализ литературных данных был выполнен на выборке литературных источников в базе PubMed, в которых описаны результаты исследований генетической составляющей РС до мая 2012 года включительно. После персональной аннотации было отобрано 1304 публикации, посвященные генетике РС. Результаты всех публикаций были распределены на публикации со статистически значимыми результатами и незначимыми. Полиморфные локусы были соотнесены с генами, которые они маркируют.

Дополнительно нами были проанализирована информация с двух он-лайн ресурсов, на которых выложены результаты полногеномных исследований ассоциаций: MSGene, A Catalog of Published Genome-Wide Association Studies. С этих двух ресурсов нами была взята информация о SNP, ассоциированных с РС по результатам GWAS. Эти SNP были также соотнесены с генами, которые они маркируют. Все гены (254 гена) были сведены в один список, полученный список генов мы использовали для выявления молекулярных путей (pathways), вовлеченных в патогенез РС.

#### 2.4.2 Молекулярные пути (pathways), вовлеченные в патогенез РС

Выявление молекулярных путей (МП), вовлеченных в патогенез РС, проводили с помощью программной платформы geneXplain. Для получения максимально полной информации мы использовали несколько баз данных (БД): Генная онтология (Gene Ontology), Энциклопедия генов человека и метаболизма (HumanCyc pathways), Reactome, TF classification, TRANSPATH®.

С помощью платформы geneXplain нами было проанализировано, какое количество генов из составленного нами списка попадает в каждый аннотированный МП. Полученное количество генов сравнивали с теоретически ожидаемым количеством генов и вычисляли уровень статистической значимости по гипергеометрическому распределению. В БД «Энциклопедия генов человека и метаболизма» и «TF classification» нами не было выявлено МП с уровнем значимости  $P < 0.01$ . В БД Reactome нами было выявлено 2 МП, для которых количество генов из нашего списка статистически значимо ( $P_{adj} < 0.01$ ) превышало теоретически ожидаемое количество генов. В БД TRANSPATH®  $P_{adj}$  был ниже порогового значения для 7 МП. В БД Генная онтология поиск осуществлялся в трех словарях: биологические процессы, клеточные компоненты и молекулярные функции. Для каждого словаря пороговое значение  $P_{adj}$  превысило довольно большое количество МП. Был выявлен 1221 МП в словаре биологические процессы, 30 МП в словаре клеточные компоненты и 50 МП в словаре молекулярные функции.

Суммарно, по результатам анализа подавляющее количество МП, которые были выявлены, можно разделить на четыре группы: пролиферация, дифференциация В и Т клеток; апоптоз; синтез и метаболизм витамина D3; передача сигналов интерлейкинами.

Выявление данных типов МП согласуется с современной концепцией патогенеза РС, в которой важная роль отводится клеткам иммунной системы, В и Т клеткам, так как заболевание аутоиммунной природы. Участие в патогенезе путей, ответственных за синтез и метаболизм витамина D3 согласуется с зависимостью распространенности заболевания от географической широты, а также с регуляцией экспрессии HLA-DRB1\*15 аллеля, главного фактора риска РС. Нарушения в апоптозе Т клеток могут приводить к аккумуляции лимфоцитов в органах лимфоидной системы, а также к нарушению апоптотической элиминации аутореактивных Т клеток, приводящей к развитию аутоиммунных заболеваний. Передача сигналов интерлейкинами является

важным звеном работы иммунной системы. Таким образом, выявленные МП дополнительно подтверждают, что РС - это аутоиммунное заболевание, имеющее воспалительный характер. Более того, детальное исследование генов обнаруженных МП может выявить новые гены, предрасполагающие к развитию РС, что, в свою очередь, может быть одним из методов решения проблемы «потерянной наследственности». Также, в случае верификации участия данных МП в патогенезе РС другими исследовательскими группами, белки и кодирующие их гены могут быть рассмотрены в качестве терапевтических мишеней для таргетной терапии при разработке новых лекарственных препаратов

## **2.5 Анализ ассоциативного генетического взаимодействия**

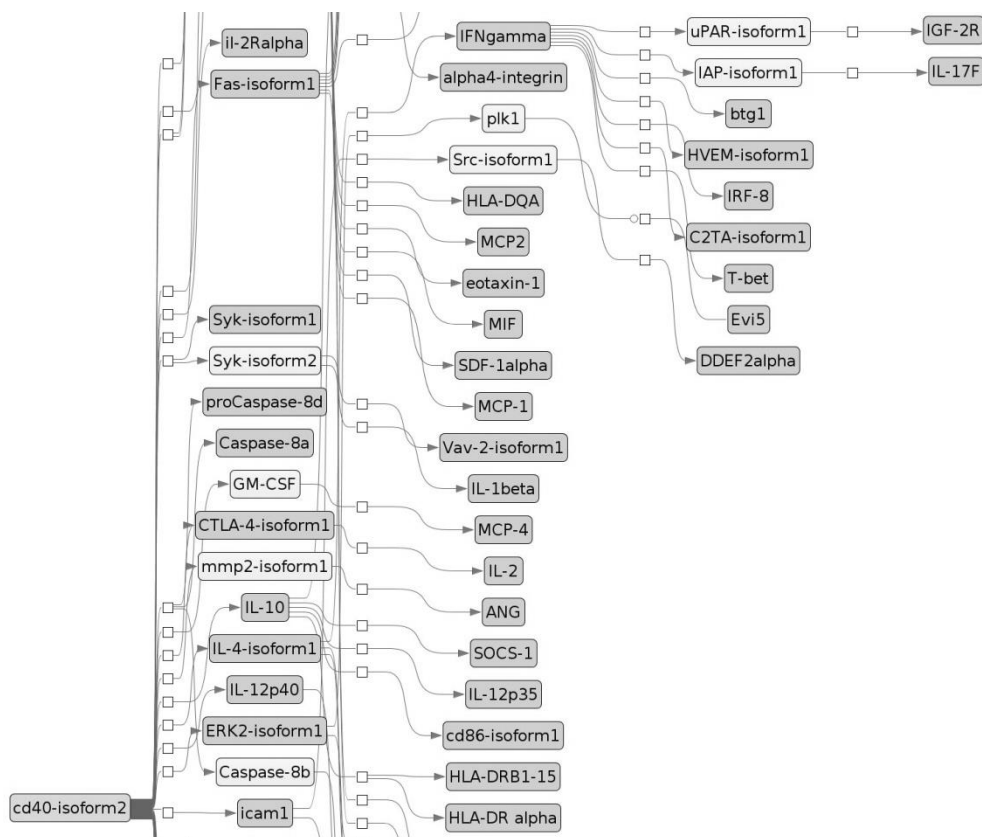
### *2.5.1 CD40*

В нашем исследовании ранее была реплицирована ассоциация SNPs гена *CD40* (rs6074022 и rs1883832). Так как РС – мультифакториальное заболевание, и его развитие обусловлено взаимодействием факторов окружающей среды и множества генов «малых эффектов», то актуальной задачей является выявление, взаимодействие каких именно генов приводит к развитию патологии. Очевидно, что гены, взаимодействие которых, приводит к развитию патогенетических процессов, скорее всего, ассоциированы с РС. В связи с этим целесообразно искать взаимодействие продукта гена *CD40* с продуктами генов, ранее выявленных как ассоциированные с РС.

Сетевой анализ. Нами был проведен сетевой анализ для гена *CD40* с помощью программной платформы geneXplain. Результат анализа представлен в виде рисунка 9, приведена только часть блок-схемы. Выбор второго гена для исследования. Так как по результатам анализа был получен довольно широкий список белков, взаимодействие которых может играть роль в патогенезе РС, выбор гена для анализа мы осуществляли из генов, продукты которых расположены в одном шаге от *CD40*. Из них для исследования взаимодействия с *CD40* нами был выбран ген *IL2RA*, в частности два его SNPs: rs12722489 и rs2104286. Исследование ассоциации с РС SNPs rs12722489 и rs2104286 гена *IL2RA*. Нами были определены генотипы этих двух SNPs у пациентов с РС (n=1266) и контрольной группы (n=1237). Оба SNPs ассоциированы с РС у жителей РФ. Согласно нашим результатам для rs12722489[G] OR равнялся 1.40, 95%CI=1.18-1.68,  $P=2 \times 10^{-4}$ , для rs2104286[A] – OR=1.32, 95%CI=1.14-1.53,  $P=2 \times 10^{-4}$ . Дополнительно мы оценили неравновесие по сцеплению между SNPs



rs12722489 и rs2104286. Данные SNPs находятся в неравновесии по сцеплению в исследуемой нами популяционной группе:  $D' = 0.937$ ,  $R^2 = 0.502$ ,  $\chi^2 = 976.05$ . Так как SNPs находятся в неравновесии по сцеплению, то для исследования ассоциативного взаимодействия генов *IL2RA* и *CD40*, проводимого с помощью метода логистической регрессии необходимо выбрать один SNP из rs12722489 и rs2104286. С помощью LRT для дальнейшего исследования нами был выбран rs12722489.



**Рис. 9.** Сетевой анализ гена *CD40*.

Анализ ассоциативного генетического взаимодействия гена *CD40* (rs6074022) и гена *IL2RA* (rs12722489). Исследование взаимодействия SNPs мы проводили с помощью логистического регрессионного анализа путем построения четырех моделей. Первая модель включала только SNP rs6074022 гена *CD40*. Вторая модель включала только SNP rs12722489 гена *IL2RA*. Третья модель включала оба SNP. Четвертая модель включала оба SNP, а также учитывала их взаимодействие. С помощью теста отношения правдоподобия мы сравнили третью модель с двумя вложенными: первой ( $P=0.01$ ) и второй ( $P=0.008$ ). Согласно полученным результатам можно заключить, что третья модель лучше любой из вложенных, и исключение любого из SNP статистически значительно ухудшает ее. Однако сравнение методом LRT третьей модели и четвертой модели, учитывающей взаимодействие SNP, не выявило различий ( $P=0.31$ ). Из полученных результатов видно, что знание генотипов

SNP обоих генов, *CD40* и *IL2RA*, помогает лучше описать ассоциацию с РС, однако выявить взаимодействие этих двух генов в нашем исследовании не удалось.

## 2.5.2 FAS

Выбор первого гена для исследования. Согласно литературным данным Fas система участвует в патогенезе РС. Еще 1996 году Dowling с соавт. были зарегистрированы глиальные клетки, подвергающиеся клеточной гибели<sup>1</sup>. Апоптозный МП был выявлен нами в ходе анализа молекулярных путей в базе данных TRANSPATH®, а также среди МП в БД Генная онтология. Помимо непосредственной гибели олигодендроцитов в очаге воспаления, предполагают, что нарушения в апоптотическом пути может играть роль при апоптотической элиминации аутореактивных Т клеток, отклонения в которой приводят к развитию аутоиммунных заболеваний. Выбор маркерного SNP гена FAS. Ассоциативных исследований полиморфных локусов гена *FAS* с предрасположенностью к РС было проведено немного, и полученные данные разнятся. Участие гена *FAS* активно исследовалось в онкологических процессах. Недавно был проведен крупный мета-анализ ассоциаций полиморфного локуса *FAS*-1377 G->A (rs2234767), по результатам которого была выявлена его ассоциация с развитием рака. Так как связь гена *FAS* с развитием рака предположительно реализуется за счет нарушения апоптоза, как и для РС, и для rs2234767 получены более однородные результаты в ассоциативных анализах, то целесообразно исследовать ассоциацию rs2234767 гена *FAS* с РС. Исследование ассоциации SNP rs2234767 гена FAS с РС у жителей РФ. Исследование ассоциации SNP rs2234767 гена *FAS* с РС мы провели на выборке общей численностью 2354 человека, из них 1341 пациент с РС и 1013 человек контрольной группы. SNP rs2234767 гена *FAS* ассоциирован с РС:  $OR_A=0.81$ , 95% C.I.=0.68-0.96,  $P=0.02$ . Так как нами была выявлена ассоциация SNP rs2234767 гена *FAS* с РС у жителей РФ, то целесообразно провести сетевой анализ для данного гена с целью выбора генов-кандидатов для анализа ассоциативного генетического взаимодействия гена *FAS*. Сетевой анализ гена FAS. Сетевой анализ гена *FAS* проводили аналогичным образом, как и для гена *CD40*. Выбор второго гена для исследования. Сетевой анализ выявил широкий спектр белков, с которыми

---

<sup>1</sup> Dowling P. et al. Involvement of the CD95 (APO-1/Fas) receptor/ligand system in multiple sclerosis brain. // J. Exp. Med. 1996. Vol. 184, № 4. P. 1513–1518.

взаимодействует Fas. Среди белков, расстояние до которых не превышает 4 шага от Fas, есть Apo2L-isoform1 и DR5-L. Белок Apo2L кодируется геном *TRAIL*. Белок DR5-L кодируется геном *TRAIL-R2*. Как и Fas система, сигнальный путь *TRAIL/TRAIL-R2* запускает внешний путь апоптоза посредством активации каспазы 8 и каспазы 10. Это позволяет предположить, что одновременное нарушение взаимодействия Fas/Fas-L и *TRAIL/TRAIL-R2* может приводить к изменениям в воспалительных процессах и тем самым влиять на предрасположенность к таким аутоиммунным заболеваниям, как РС. Для исследования взаимодействия были выбраны SNPs: rs4894559 (*TRAIL*), и rs1001793 (*TRAIL-R2*). Исследование ассоциации с РС SNPs rs4894559 (*TRAIL*) и rs1001793 (*TRAIL-R2*). Нами были определены генотипы SNPs rs4894559 (*TRAIL*) и rs1001793 (*TRAIL-R2*) у пациентов с РС и в контрольной группе. Ассоциация rs4894559 (*TRAIL*) с РС была близка к пограничному уровню статистической значимости:  $OR=0.87$ ,  $95\%CI=[0.75-1.01]$ ,  $P=0.06$ . В нашем исследовании не было выявлено ассоциации SNP rs1001793 (*TRAIL-R2*) с РС. Таким образом, исследование взаимодействия SNP rs2234767 гена *FAS* на наших данных мы можем провести только для SNP rs4894559 гена *TRAIL*. Анализ ассоциативного генетического взаимодействия гена *TRAIL* и гена *FAS*. Для анализа их взаимодействия аналогичным образом, как и для *CD40*, нами были построены четыре модели и сравнены с помощью LRT. Из полученных результатов можно сделать следующие выводы. Добавление в модель, где учтен только один генетический предиктор - rs2234767(*FAS*), дополнительного генетического предиктора не улучшает ее ( $P=0.33$ ). Более того, если из модели, учитывающей оба SNPs убрать предиктор rs2234767(*FAS*), то модель становится статистически значимо хуже ( $P=0.004$ ). Таким образом, согласно нашим результатам rs2234767(*FAS*) является более значимым предиктором развития РС, чем rs4894559(*TRAIL*). Добавление в модель в качестве предиктора взаимодействие этих двух SNPs не улучшало ее прогностических характеристик ( $P=0.16$ ).

Суммарно, нам не удалось выявить ассоциативного взаимодействия генов каспазного сигнального пути, *TRAIL* и *FAS*, в формировании предрасположенности к развитию РС.

## ВЫВОДЫ

1. Исследование, проведенное на выборках ДНК от пациентов и условно здоровых контролей русской этнической группы, проживающих на территории

Западной Сибири и Якутии не выявило ассоциации с РС полиморфного локуса rs10492972 гена *KIF1B*, в то время как по результатам мета-анализа, включающего 10054 пациентов с РС и 9193 участника контрольной группы, показано, что аллель С полиморфного локуса rs10492972 гена *KIF1B* является статистически значимым протективным фактором развития РС.

2. Для полиморфного локуса -308G>А гена *TNFA* не реплицирована ассоциация с РС; HLA-DRB1\*15 аллель - фактор риска развития РС, HLA-DRB1\*01 и \*14 - протективные факторы в русской этнической группе. На примере русской этнической группы показано, что ассоциация полиморфного локуса -308G>А гена *TNFA*, ранее идентифицированная при исследовании выборок малого размера, является наведенной от HLA-DRB1\*15.

3. Подтверждена ассоциация РС с полиморфными локусами rs6074022 и rs1883832 гена *CD40* в исследуемой популяции. Показано, что функциональный локус гена *CD40* лежит выше старта трансляции гена. *In silico* показано, что сайты связывания транскрипционных факторов VDR, STAT3, IRF-4 – могут быть вовлечены в регуляцию экспрессии гена *CD40*, изменения которых в силу полиморфизма гена, вероятно, будут приводить к повышенному риску развития РС. *In vitro* показано, отсутствие аллельного дисбаланса для гена *CD40* в качестве потенциального молекулярного механизма реализации ассоциированных аллельных вариантов гена.

4. В результате интегрального анализа опубликованных к настоящему моменту данных о генетике РС выявлено четыре группы молекулярных путей, вовлеченных в патогенез РС: пролиферация, дифференциация В и Т клеток; апоптоз; синтез и метаболизм витамина D3; передача сигналов интерлейкинами.

5. На основании анализа сигнальных путей, включающих продукт гена *CD40*, и сформированного в процессе исследования списка ассоциированных с РС генов выбран наиболее перспективный ген (*IL2RA*) для анализа взаимодействия ассоциаций; верифицирована ассоциация с РС полиморфных локусов rs12722489 и rs2104286 на русской этнической группе. Взаимодействия генов *CD40* и *IL2RA* в нашем исследовании не выявлено.

6. Впервые выявлена ассоциация полиморфного локуса rs2234767 с РС в русской этнической группе и проведен анализ сигнальных путей для гена *FAS*, на основании которого для исследования взаимодействия выбраны гены *TRAIL* и *TRAIL-R2*. Впервые выявлена ассоциация полиморфного локуса *TRAIL*

(rs4894559) с РС в русской этнической группе. Взаимодействия двух генов, *FAS* и *TRAIL*, в нашем исследовании не выявлено.

**Основные результаты диссертации опубликованы в следующих работах:**

1) Ханох Е.В., Рождественский А.С., **Кудрявцева Е.А.**, Какуля А.В., Делов Р.А., Филипенко М.Л. Исследование наследственных факторов предрасположенности к рассеянному склерозу и особенностей его течения в русской этнической группе. // Бюллетень СО РАМН - 2011 - Т.31 - №1 - С.113-118.

2) Смагина И.В., Ельчанинова С.А., Золовкина А.Г., Игнатова Ю.Н., **Кудрявцева Е.А.** Генетические факторы риска рассеянного склероза в популяции Алтайского края. // Журнал неврологии и психиатрии им. Корсакова – 2011 - №5 - С. 42-45.

3) **Kudryavtseva E.A.**, Rozhdestvenskii A.S., Kakulya A.V., Khanokh E.V., Delov R.A., Malkova N.A., Korobko D.S., Platonov F.A., Aref'eva E.G., Zagorskaya N.N., Aliferova V.M., Titova M.A., Babenko S.A., Smagina I.V., El'chaninova S.A., Zolovkina A.G., Lifshits G.I., Puzyrev V.P., Filipenko M.L. Polymorphic locus rs10492972 of the KIF1B gene association with multiple sclerosis in Russia: case control study. // Molecular Genetics and Metabolism -2011 –V.104 - P. 390-394.

4) Ханох Е.В., Рождественский А.С., **Кудрявцева Е.А.**, Какуля А.В., Делов Р.А., Филипенко М.Л. Влияние полиморфных локусов rs1800629 (TNFa), rs6074022 (CD40), rs187238 (IL-18), rs10492972 (KIF1B),rs4149584 (TNFRSF1A) на особенности клинических проявлений рассеянного склероза с учетом гендерной принадлежности в этнической группе русских. // Бюллетень сибирской медицины– 2011 - №2 - С.50-57.

5) Коробко Д.С., **Кудрявцева Е.А.**, Малкова Н.А., Филипенко М.Л. Связь полиморфизмов генов цитокинов со скоростью прогрессирования рассеянного склероза. // Журнал неврологии и психиатрии им. Корсакова – 2012 - №2 - Вып.2 - С.9-15.

6) Smagina I.V., Elchaninova S.A., Zolovkina A.G., Ignatova Yu.N., **Kudryavtseva E.A.** Genetic Risk Factors for Multiple Sclerosis in the Population of the Altai District. // Neuroscience and Behavioral Physiology. –2012 - V. 42 -№8 – P. 876-879.

- 7) Коробко Д.С., Малкова Н.А., Булатова Е.В., Бабенко Л.А., Сазонов Д.В., **Соколова Е.А.**, Филипенко М.Л. Влияние генетических факторов на фенотипическую экспрессию рассеянного склероза. // Журнал неврологии и психиатрии им. Корсакова – 2013 - №2 - Вып.2. - С.10-15.
- 8) **Соколова Е.А.**, Малкова Н.А., Коробко Д.С., Рождественский А.С., Какуля А.В., Ханох Е.В., Делов Р.А., Платонов Ф.А., Попова Т.Е., Арефьева Е.Г., Загорская Н.Н., Алифирова В.М., Титова М.А., Смагина И.В., Ельчанинова С.А., Поповцева А.В., Личенко Ю.Н., Дымова М.А., Вайнер А.С., Пьянкова О.В., Оскорбин И.П., Кечин А.А., Пузырев В.П., Кулакова О.Г., Царева Е.Ю., Фаворова О.О., Щур С.Г., Лащ Н.Ю., Попова Н.Ф., Попова Е.В., Гусев Е.И., Бойко А.Н., Аульченко Ю.С., Филипенко М.Л. Первые результаты объединенного общероссийского исследования по клинической генетике РС. // Журнал неврологии и психиатрии им. Корсакова – 2013 - №2 - Вып.2. - С.6-9.
- 9) **Sokolova E.A.**, Malkova N.A., Korobko D.S., Rozhdestvenskii A.S., Kakulya A.V., Khanokh E.V., Delov R.A., Platonov F.A., Popova T.Y., Aref'eva T.G., Zagorskaya N.N., Alifirova V.M., Titova M.A., Smagina I.V., El'chaninova S.A., Popovtseva A.V., Puzyrev V.P., Kulakova O.G., Tsareva E.Y., Favorova O.O., Shchur S.G., Lashch N.Y., Popova N.F., Popova E.V., Gusev E.I., Boyko A.N., Aulchenko Y.S., Filipenko M.L. Association of SNPs of CD40 Gene with Multiple Sclerosis in Russians. // PLoS ONE - 2013 - V.8 - N.4: e61032. doi:10.1371/journal.pone.0061032.
- 10) Делов Р.А., Рождественский А.С., Маркс Е.А., Смяловский В.Э., Ханох Е.В., Какуля А.В., **Соколова Е.А.**, Агапова О.Ю. Клинико-электрофизиологическая и молекулярно-генетическая характеристика течения ремиттирующего рассеянного склероза. // Журнал неврологии и психиатрии им. Корсакова – 2013, №10, С.55-59.
- 11) **Соколова Е.А.**, Боярских У.А., Аульченко Ю.С., Филипенко М.Л. Генетика рассеянного склероза сегодня. // Успехи современной биологии. 2015. Т.135. № 4. С. 355-369.