

## ГЕНЕТИЧЕСКОЕ РАЗНООБРАЗИЕ И НЕРАВНОВЕСИЕ ПО СЦЕПЛЕНИЮ В ЛОКУСЕ МЕТИЛЕНТЕТРАГИДРОФОЛАТРЕДУКТАЗЫ

© 2008 г. Е. А. Трифонова, М. Г. Спиридонова, В. А. Степанов

Научно-исследовательский институт медицинской генетики Томского научного центра  
Сибирского отделения Российской академии медицинских наук,  
Томск 634050;

e-mail: vadim.stepanov@medgenetics.ru

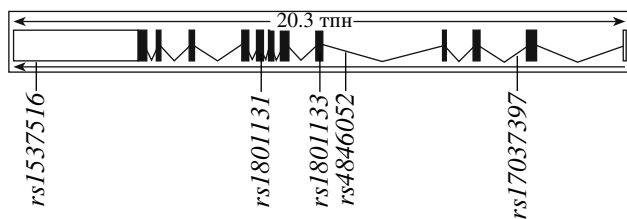
Поступила в редакцию 12.12.2007 г.

Современное и интенсивно исследуемое многими учеными направление в области молекулярной генетики – это изучение блочной архитектуры неравновесия по сцеплению в геноме человека. В представленной работе исследованы генетическая дифференциация и структура неравновесия по сцеплению в гене метилентетрагидрофолатредуктазы (*MTHFR*) в популяциях русских, тувинцев, северных и южных киргизов. Метилентетрагидрофолатредуктаза является ключевым ферментом фолатного цикла, ответственным за восстановление 5,10-метилентетрагидрофолата до 5-метилтетрагидрофолата. Снижение активности данного фермента, часто обусловленное наличием определенных сочетаний аллелей гена *MTHFR*, приводит к увеличению концентрации гомоцистеина в крови. В данной работе определены частоты генотипов и аллелей пяти полиморфизмов гена *MTHFR*: *rs17037397*, *rs4846052*, *rs1801133*, *rs1801131* и *rs1537516* в исследованных группах. Обнаружены значимые генетические различия между популяцией южных киргизов и остальными исследованными группами, а также между русскими и тувинцами. В популяции южных киргизов в гене *MTHFR* выявлен один блок, а в популяциях русских, тувинцев и северных киргизов найдены два блока. Таким образом, структура неравновесия по сцеплению в гене *MTHFR* носит популяционно-специфический характер.

Архитектура неравновесия по сцеплению (**LD**, linkage disequilibrium) в геноме человека является в настоящее время предметом оживленных дискуссий и интенсивных исследований [1–4]. Ряд недавних работ показывает, что в геноме можно выделить блоки сцепленных сайтов, не демонстрирующие свидетельств значительной рекомбинации в истории существования *Homo sapiens*, отделяемые участками с более интенсивным темпом рекомбинации, так называемыми “горячими точками” [2, 3, 5]. Характер неравновесия по сцеплению в современных популяциях человека – это результат комплексного эволюционного процесса, который включает как демографическую историю популяций (изменения эффективной численности, характер подразделенности, миграции), так и гено-специфические факторы, такие как темп мутирования и рекомбинации, давление отбора [6–8]. Обнаружение сильного неравновесия по сцеплению между специфическими мутациями гена и определенными аллелями маркерных локусов имеет важное практическое значение. Часто наблюдается неравновесие по сцеплению между мутантными аллелями гена и одновременно несколькими маркерными локусами. В этом случае анализ маркерных гаплотипов, т. е. наборов аллелей различных локусов, локализованных в одной хромосоме, дает возможность с высокой степенью вероятности оценивать характер мутационно-

го повреждения и проследивать его особенности в популяционных выборках и группах, отягощенных какой-либо патологией. Предполагается, что характеристика LD займет центральное место при картировании генов мультифакториальных заболеваний и позволит реконструировать демографическую историю популяций [2, 3].

Одной из основных причин существования LD в популяции является тесное сцепление. Например, если в момент возникновения мутации, вызывающей заболевание, рядом находился определенный маркерный аллель, то в течение многих поколений этот аллель будет передаваться вместе с мутацией. Рекомбинация постепенно разрушает ассоциацию и происходит это тем быстрее, чем дальше друг от друга расположены локусы. Для тесно сцепленных (1–2 cM) локусов неравновесие по сцеплению сохраняется десятки поколений [9]. Таким образом, неравновесие по сцеплению (LD) является значимой популяционной характеристикой, позволяющей судить о порядке и примерном времени возникновения различных мутаций, а также оценивать возможные механизмы их поддержания в популяции. Результаты, полученные сотрудниками университета в Цинциннати (шт. Огайо, США), демонстрируют, что даже при отсутствии “горячих точек” рекомбинации распределенные случайным образом



**Рис. 1.** Карта гена *MTHFR* и локализация на ней пяти SNPs. Экзоны обозначены прямоугольниками, интроны показаны линиями.

события рекомбинации могут также привести к образованию блоков гаплотипов и что популяционные генетические характеристики, объединенные демографической историей, влияют на блочную структуру. Следовательно, популяции с различной демографической историей будут, вероятно, иметь различные блочные структуры. Это обстоятельство указывает на то, что использование структуры LD при картировании мультифакториальных заболеваний ограничено популяционной специфичностью [7]. Задача данной работы – изучение характера неравновесия по сцеплению в локусе метилентетрагидрофолатредуктазы в различных популяциях.

Метилентетрагидрофолатредуктаза – катализатор единственной внутриклеточной реакции образования 5-метилтетрагидрофолата, который необходим для восстановления гомоцистеина до метионина. Снижение активности данного фермента приводит к накоплению гомоцистеина и развитию умеренной гипергомоцистеинемии. Ген *MTHFR* у человека расположен на коротком плече первой хромосомы (1p36.3). Длина всего кодирующего региона составляет около 20335 пн с расчетной молекулярной массой продукта 74,6 кДа [10]. Ген состоит из 12 экзонов длиной от 102 до 432 пн, 10 интронов длиной от 250 до 1500 пн и одного интрона длиной 4200 пн [11]. На данный момент в гене *MTHFR* описано 109 однонуклеотидных замен (SNP, single nucleotide polymorphism). Наиболее изученным полиморфизмом гена *MTHFR* является вариант, в котором нуклеотид цитозин (С) в позиции 677, относящийся к 5-му экзону, заменен на тимидин (Т), что приводит к замене аминокислотного остатка аланина на валин в области связывания с ко-фактором FAD (флавин-аденин-динуклеотид). При изменении структуры этой области диссоциация фермента с ко-фактором происходит слишком быстро для осуществления полноценной каталитической функции. Показано, что у гомозигот по аллелю *T* наблюдается снижение активности *MTHFR* до 30%, а у гетерозигот – до 65% от обычного уровня. Наличие *TT*-генотипа ведет к повышению уровня ГЦ примерно в 2 раза по сравнению с *CC*-субъектами [12]. Мутантный аллель *T677* гена *MTHFR* распределен в

популяциях мира с высокой гетерогенностью. Его частота у народонаселения нашей планеты варьирует от полного отсутствия в некоторых африканских племенах до 0,55 у испанцев [13].

Другим известным полиморфным вариантом гена *MTHFR* является точечная замена нуклеотида аденина (А) на цитозин (С) в позиции 1298. Это приводит к замене остатка глутамина на остаток аланина в регуляторном домене фермента, что сопровождается небольшим снижением его активности. У лиц, гомозиготных по мутации *A1298C*, активность *MTHFR* составляет приблизительно 60% от нормы. Предполагается, что снижение активности данного фермента связано с изменением его регуляции ингибитором S-аденозилметионином [14].

Многие авторы рассматривают полиморфные варианты *A1298C* и *C677T* гена *MTHFR* в качестве генетических маркеров различных мультифакториальных состояний: сердечно-сосудистых заболеваний, венозных тромбозов, дефекта нервной трубки, мигрени, преэклампсии, сахарного диабета 2-го типа, отслойки плаценты, привычного невынашивания беременности, шизофрении, болезни Альцгеймера, инвазивного цервикального рака, острого лейкоза и других [13, 15–19]. В представленной работе показана структура неравновесия по сцеплению между пятью SNPs (включая *C677T* и *A1298C*) гена *MTHFR*. Данное исследование продолжает цикл работ, посвященных изучению генетического разнообразия и неравновесия по сцеплению в популяциях Северной Евразии [20, 21].

## МАТЕРИАЛЫ И МЕТОДЫ

В качестве маркеров для изучения LD были выбраны следующие пять SNPs гена *MTHFR*: *rs17037397*, *rs4846052*, *rs1801133* (*C677T*), *rs1801131* (*A1298C*) и *rs1537516*. Данные SNPs равномерно распределены в кодирующей части гена *MTHFR* (рис. 1), частота минорного аллеля составляет не менее 5% (по данным базы NSBI). Сайты *C677T* и *A1298C* расположены в 5-м и 8-м экзонах соответственно, *rs17037397* и *rs4846052* локализованы во 2-м и 4-м интронах, а *rs1537516* – в 3' UTR-области [11].

Материал для исследования представлен выборками из популяций, проживающих в различных регионах Северной Евразии: русские (г. Томск), тувинцы (пос. Бай-Тайга, Республика Тыва), южные киргизы (г. Ош, Киргизия) и северные киргизы (г. Бишкек, Киргизия). В исследование включались индивиды, неметисированные по крайней мере в трех поколениях. Общая численность суммарной выборки составила 456 человек. У обследованных индивидов была выделена ДНК, проамплифицированы интересующие полиморфные локусы (методом ПЦР) с последующим рестрикционным анализом специфическими эндонуклеазами. Условия амплификации и ре-

**Таблица 1.** Условия амплификации и ферменты рестрикции для исследованных SNPs

Исследуемый локус	Праймеры	Температура отжига	Фермент рестрикции
<i>rs17037397</i>	5': 5'-AGAACTTAGAACAGTGCCTGGTG-3' 3': 5'-TAGGATCCCCCTAACTCTCAATCA-3'	59°C	<i>Sse9I</i>
<i>rs4846052</i>	5': 5'-GCAGTGGATGCTGAAGACAA-3' 3': 5'-TTATCCCCCACTCCCTTACC-3'	60°C	<i>BstF5I</i>
<i>rs1801133</i>	5': 5'-TGAAGGAGAAGGTGTCTGCGGGA-3' 3': 5'-AGGACGGTGCGGTGAGAGTG-3'	56°C	<i>HinI</i>
<i>rs1801131</i>	5': 5'-ATGTGGGGGGAGGAGGTGAC-3' 3': 5'-GTCTCCCAACTTACCCTTCTCCC-3'	60°C	<i>MboII</i>
<i>rs1537516</i>	5': 5'-GGACCCAGCTACTCTGTCCA-3' 3': 5'-CTCCTTGGGACACGGTCTTA-3'	37°C	<i>AspA2I</i>

стрикции для локусов *C677T* и *A1298C MTHFR* описаны ранее [22, 23]. Для трех SNPs: *rs17037397*, *rs4846052* и *rs1537516* были подобраны комплементарные праймеры, специфические эндонуклеазы, условия ПЦР и рестрикции (табл. 1).

Продукты амплификации и рестрикции анализировали с помощью электрофореза в 2%-ном (для *rs17037397*, *rs4846052* и *rs1537516*) или 3%-ном (для *rs1801133* и *rs1801131*) агарозном геле, окрашенном бромистым этидием. Искомые бенды визуализировали в ультрафиолетовом трансиллюминаторе.

Статистическую обработку результатов исследования проводили с помощью пакетов статистических программ ARLEQUIN и Harpview 4.0. Частоты гаплотипов определяли с помощью EM-алгоритма. LD между парами SNP оценивалось с помощью коэффициента  $D'$ , предложенного Левонтином, и коэффициента корреляции  $r^2$  Пирсона. Блочная структура определялась посредством алгоритма “Правило четырех гамет” [7], предусмотренного программным обеспечением Harpview 4.0. Для каждой пары маркеров рассчитывали частоты всех четырех возможных гаплотипов. Наличие рекомбинации допускали, если все четыре встречались с частотой более 0.02. Группу маркеров, в которых для любых двух соседних SNP встречалось не более трех типов гамет, трактовали как блок LD. Уровень генетического разнообразия и межпопуляционной дифференциации вычисляли методом анализа молекулярной вариабельности (AMOVA).

## РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

### Генетическое разнообразие в локусе *MTHFR*

Частоты генотипов и аллелей, гетерозиготность и уровень значимости соответствия равновесию Харди–Вайнберга по пяти изученным SNP представлены в табл. 2. Практически по всем локусам во всех популяциях распределение частот аллелей соответствует равновесию Харди–Вайнберга (исключение составляет лишь локус *A1298C* в популяционной выборке тувинцев). В большинстве случаев наблюдается высокая гетерозиготность, кроме *rs17037397*, что согласуется с мировыми данными.

В исследованных популяциях обнаружено 19 гаплотипов из 32 возможных (табл. 3). В популяционной выборке русских выявлено 16 гаплотипов, тувинцев – 12, у южных и северных киргизов – 10 и 11 соответственно. Все исследованные популяции характеризуются высоким уровнем гаплотипического разнообразия. Известно, что увеличение числа гаплотипов и гаплотипического разнообразия указывает на более древний возраст популяции. Гаплотипы, встречающиеся с частотой более 5%, обозначены как основные гаплотипы. Всего найдено восемь основных гаплотипов, три из них (№ 1, 9, и 12) встречаются с частотой более 10% во всех четырех популяциях и в сумме составляют более 78% хромосом общей выборки. Наибольшее число основных гаплотипов (6) обнаружено у северных киргизов, у русских этот показатель составляет 5, а у тувинцев и южных киргизов – 4. Популяции южных киргизов и тувинцев имеют идентичные основные гаплотипы с приблизительно равными частотами. Значительное генетическое разнообразие в этих популяциях выявлено и по другим генетическим маркерам,

**Таблица 2.** Распределение генотипов и аллелей изученных полиморфных вариантов гена *MTHFR* в исследованных выборках

Исследуемые SNPs		Исследуемые выборки				
		Тувинцы ( <i>N</i> = 134)	Русские ( <i>N</i> = 126)	Южные киргизы ( <i>N</i> = 111)	Северные киргизы ( <i>N</i> = 85)	
rs17037397	Частоты генотипов	CC	0.88	0.89	0.89	0.92
		AC	0.12	0.11	0.11	0.08
		AA	–	–	–	–
	Частота аллеля	A	0.06	0.06	0.05	0.04
	Наблюдаемая гетерозиготность		0.12	0.11	0.11	0.08
	Ожидаемая гетерозиготность		0.25	0.11	0.29	0.11
	Уровень значимости*		1.00	1.00	1.00	1.00
rs4846052	Частоты генотипов	CC	0.46	0.30	0.55	0.40
		CT	0.41	0.53	0.42	0.52
		TT	0.13	0.17	0.03	0.08
	Частота аллеля	T	0.54	0.43	0.24	0.34
	Наблюдаемая гетерозиготность		0.41	0.53	0.42	0.52
	Ожидаемая гетерозиготность		0.45	0.51	0.51	0.53
	Уровень значимости		0.34	0.46	0.12	0.23
rs1801133	Частоты генотипов	CC	0.67	0.50	0.53	0.54
		CT	0.28	0.37	0.37	0.44
		TT	0.05	0.13	0.10	0.02
	Частота аллеля	T	0.19	0.32	0.28	0.24
	Наблюдаемая гетерозиготность		0.28	0.37	0.37	0.44
	Ожидаемая гетерозиготность		0.43	0.45	0.54	0.39
	Уровень значимости		0.26	0.10	0.35	0.13
rs1801131	Частоты генотипов	AA	0.44	0.40	0.62	0.34
		AC	0.35	0.48	0.35	0.54
		CC	0.21	0.12	0.03	0.12
	Частота аллеля	C	0.38	0.36	0.20	0.39
	Наблюдаемая гетерозиготность		0.35	0.48	0.35	0.54
	Ожидаемая гетерозиготность		0.48	0.48	0.48	0.56
	Уровень значимости		0.01	0.70	0.55	0.25
rs1537516	Частоты генотипов	CC	0.84	0.81	0.84	0.80
		CT	0.16	0.17	0.15	0.20
		TT	–	0.02	0.01	–
	Частота аллеля	T	0.08	0.10	0.08	0.10
	Наблюдаемая гетерозиготность		0.16	0.17	0.14	0.20
	Ожидаемая гетерозиготность		0.28	0.18	0.33	0.18
	Уровень значимости		1.00	0.34	0.52	1.00

Примечание. \* – уровень значимости соответствия равновесию Харди–Вайнберга. *N* – количество индивидуумов в выборке.

**Таблица 3.** Распределение частот гаплотипов гена *MTHFR* в исследованных выборках

№	Гаплотип	Русские ( <i>N</i> = 126)	Тувинцы ( <i>N</i> = 134)	Южные киргизы ( <i>N</i> = 111)	Северные киргизы ( <i>N</i> = 85)
	5 4 3 2 1*				
1	<b>СССАС</b>	<b>0.314</b>	<b>0.439</b>	<b>0.483</b>	<b>0.328</b>
2	АССАС	0.015	–	–	0.008
3	СССАТ	0.005	–	0.005	–
4	<b>СТСАС</b>	<b>0.059</b>	0.032	0.041	<b>0.069</b>
5	СТСАТ	0.009	–	–	0.009
6	<b>ССССС</b>	–	0.039	0.005	<b>0.082</b>
7	ССССТ	0.004	0.007	–	–
8	АТССС	0.011	–	–	0.011
9	<b>СТССС</b>	<b>0.194</b>	<b>0.233</b>	<b>0.107</b>	<b>0.161</b>
10	<b>АТССТ</b>	0.027	<b>0.054</b>	<b>0.050</b>	0.022
11	<b>СТССТ</b>	0.045	0.006	0.026	<b>0.071</b>
12	<b>ССТАС</b>	<b>0.217</b>	<b>0.141</b>	<b>0.269</b>	<b>0.199</b>
13	СТТАС	0.021	0.004	–	–
14	ССТСС	0.010	0.034	–	0.042
15	АСТСТ	0.003	–	–	–
16	<b>СТТСС</b>	<b>0.061</b>	–	0.011	–
17	СТТСТ	0.006	0.006	–	–
18	АТТСТ	–	0.006	–	–
19	АТТСС	–	–	0.005	–
Число гаплотипов		16	12	10	11
Гаплотипическое разнообразие ( <i>H</i> )		0.807	0.751	0.697	0.781

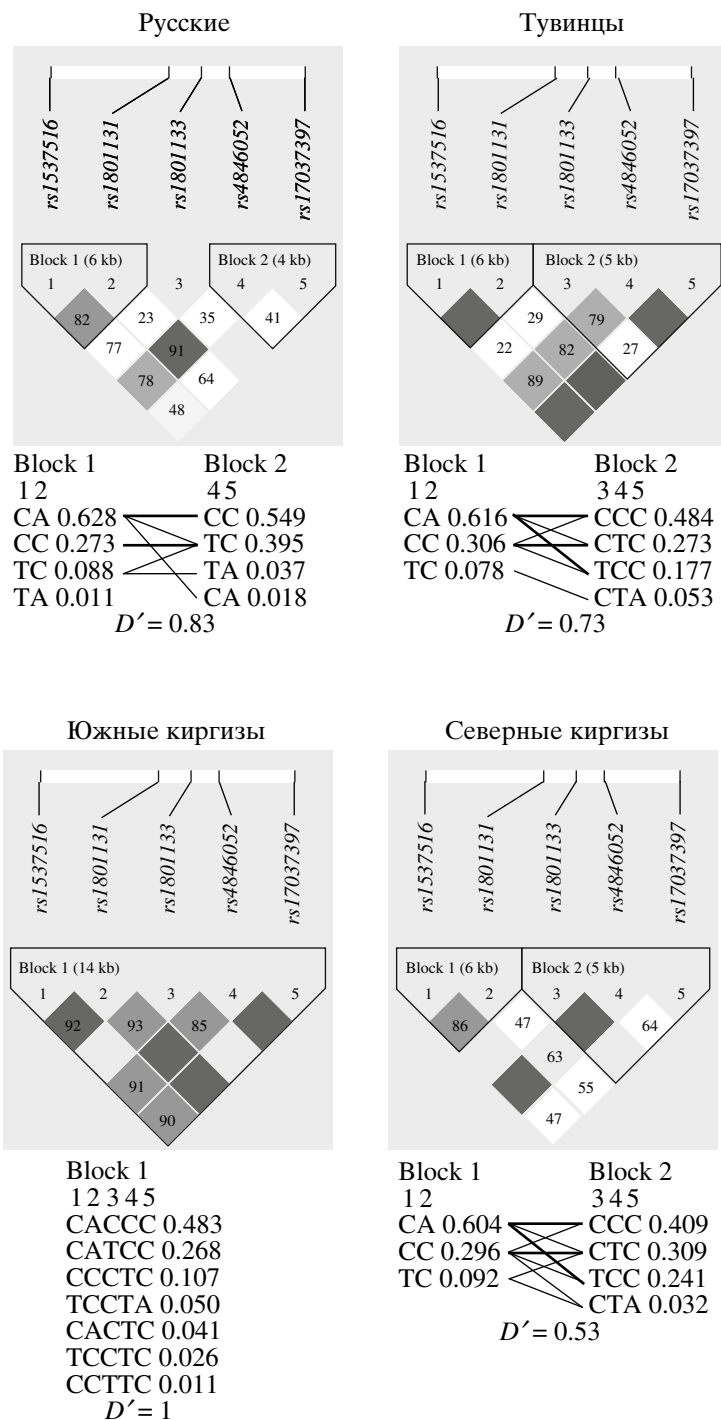
Примечание. \* – нумерация SNPs: 1 – *rs1537516*, 2 – *rs1801131*, 3 – *rs1801133*, 4 – *rs4846052*, 5 – *rs17037397*; жирным шрифтом и серым фоном обозначены основные гаплотипы.

включая гаплотипы Y-хромосомы и некоторые аутосомные маркеры [24].

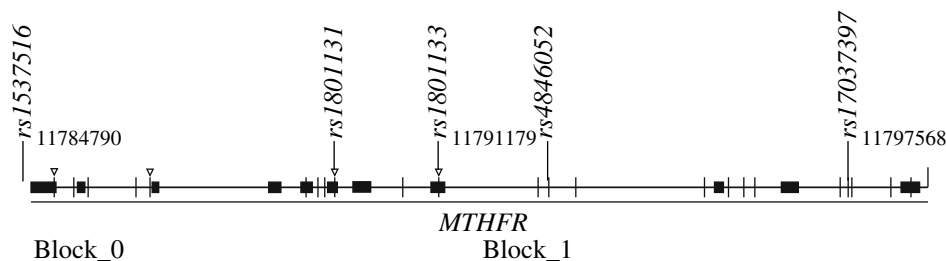
#### Структура LD в гене *MTHFR*

На рис. 2 показана структура LD между исследованными локусами в четырех популяционных выборках. В целом картина неравновесия по сцеплению в гене *MTHFR* имеет популяционно-специфический характер. В популяциях тувинцев и северных киргизов можно выделить два сходных по своей структуре блока (1-й блок включает в себя *rs1801131* и *rs1537516*; 2-й – *rs17037397*, *rs4846052* и

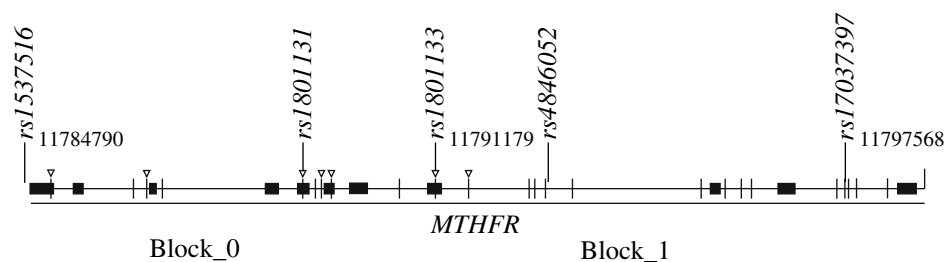
*rs1801133*); у южных киргизов все исследованные SNPs находятся в тесном сцеплении и входят в состав одного блока; в популяции русских выявлены два блока из двух близко расположенных SNPs. Гаплотипические блоки длиной от 4 до 14 тпн представлены несколькими (3–6) основными гаплотипами, которые в сумме составляют более 90% наблюдаемых гаплотипов в выборке, включающей все исследованные популяции. Примечательно, что функционально значимые полиморфизмы *C677T* и *A1298C* во всех популяциях, кроме южных киргизов, не сцеплены.



**Рис. 2.** Структура неравновесия по сцеплению в гене *MTHFR* в популяциях Сибири и Средней Азии. Нумерация соответствует следующему порядку SNPs: 1 – *rs1537516*, 2 – *rs1801131*, 3 – *rs1801133*, 4 – *rs4846052*, 5 – *rs17037397*. Показаны гаплотипы каждого блока и их частоты, мультиаллельный коэффициент  $D'$  Хедрика (Hedrick's multiallelic  $D'$ ). Толстыми и тонкими линиями изображены сочетания гаплотипов в смежных блоках. В ячейках указано значение коэффициента сцепления  $D'$  по Левонгину  $\times 0.01$  (пустая ячейка обозначает  $D' = 1$ ), цветовая гамма отображает силу сцепления между SNPs: темно-серый цвет – сильное сцепление ( $D' = 1$ ,  $LOD > 2$ ), серый и светло-серый – значительное сцепление ( $D' < 1$ ,  $LOD > 2$ ), белый – слабое сцепление ( $D' < 1$ ,  $LOD < 2$ ). Ячейка светло-серого цвета без указанного значения  $D'$  обозначает невозможность расчета неравновесия по сцеплению вследствие низкой частоты минорного аллеля полиморфизма ( $D' = 1$ ,  $LOD < 2$ ).



**Рис. 3.** Структура неравновесия по сцеплению в гене *MTHFR* в популяциях японцев и китайцев. Прямоугольниками обозначены экзоны, линиями показаны интроны.



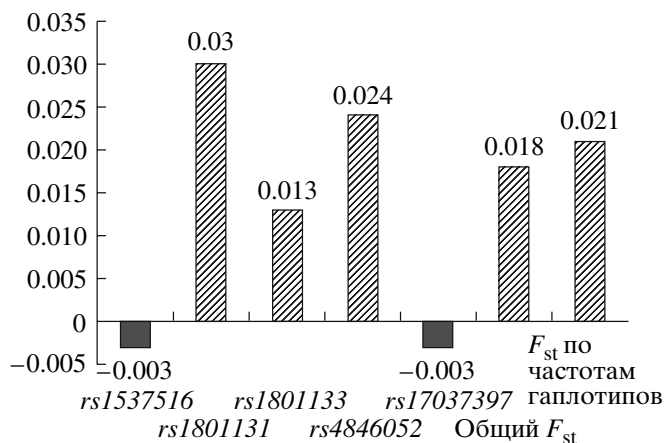
**Рис. 4.** Структура неравновесия по сцеплению в гене *MTHFR* в популяции уроженцев Северной и Западной Европы. Прямоугольниками обозначены экзоны, линиями показаны интроны.

В данном исследовании структура LD определялась с помощью алгоритма “Правило четырех гамет”, с частотой 4-й гаметы, не превышающей 0.02. В целом гаплотипические блоки могут быть определены тремя различными путями [25]. В частности, одним из способов определения блока является включение в него маркеров, между которыми среднее значение  $D'$  (стандартизованный коэффициент LD) больше, чем определенный порог [1, 3]. Второй способ основан на понятии “хромосомного охвата”, блок гаплотипов должен содержать минимальное число SNPs, составляющих большинство общих гаплотипов или обуславливающих сниженный уровень гаплотипического разнообразия [4]. Wang и др. [7] предлагает альтернативный метод для идентификации блоков. Данный метод основан на частоте гамет и не требует определения какого-либо порога [7]. Использование вышеописанных алгоритмов определения структуры LD, или изменение порога  $D'$  в рамках одного исследования, часто приводит к различной картине архитектуры LD в одной и той же популяции. Также одной из причин получения ложноположительных результатов является то, что тесное сцепление генов – не единственная причина возникновения неравновесия по сцеплению. Оно может появиться в выборке из-за подразделенности популяции или кластеризации данных [26]. Популяции человека,

строго говоря, не являются панмиксными. В них всегда присутствует подразделенность, основанная на этнических, религиозных, социальных, культурных особенностях. Кластеризация выражается в том, что при наследственной патологии в группу больных часто попадают близкие родственники. [27]. Другая причина разрушения неравновесия по сцеплению – мутации в маркерных локусах. К счастью, они возникают достаточно редко, особенно если в качестве маркеров используются SNPs [28].

При статистической обработке данных проекта НарМар с помощью программы NarBlock [29, 30], разработанной Zang (данный анализ был проведен посредством iNAP (integrated Haplotype Analysis Pipeline)), получена следующая архитектура неравновесия по сцеплению в гене *MTHFR* в популяциях японцев (Токио), китайцев (Пекин) и жителей Юты с происхождением из Северной и Западной Европы. В популяции японцев выявлено два блока: небольшой блок 0 и протяженный блок 1, охватывающий практически всю длину гена. Аналогичная структура неравновесия по сцеплению наблюдается в популяции китайцев (рис. 3). В европейской популяции также наблюдается два блока. Блок 0 значительно протяженнее, а блок 1 короче подобных блоков у японцев и китайцев (рис. 4).

Как видно из рис. 2, 3 и 4, характер LD в локусе *MTHFR* в популяциях русских, тувинцев и северных



**Рис. 5.** Общая генетическая дифференциация тотальной выборки по исследованным полиморфизмам гена *MTHFR*. Заштрихованные столбцы означают значимые различия ( $p < 0.05$ ).

киргизов соответствует “европейскому” паттерну LD, полученному при анализе данных NapMap (*rs17037397*, *rs4846052* и *rs1801133* входят в состав одного блока; *rs1801131* и *rs1537516* включены в другой). Структура LD у южных киргизов больше напоминает “монголоидный” паттерн (все исследованные SNP, за исключением *rs1537516*, состоят в одном блоке). Возможны два объяснения наблюдаемому сохранению структуры и протяженности блока: недавнее происхождение от общих предков или позитивный отбор, который часто приводит к увеличению размера блока, содержащего полезный аллель [31]. Поскольку у южных киргизов выявлено минимальное число гаплотипов среди всех исследованных популяций, есть вероятность, что в данном случае имел место эффект основателя.

Выявление структуры LD вносит значительный вклад в изучение человеческого генома и эволюции, а также необходимо для дизайна эффективных ассоциативных исследований. Степень гаплотипического разнообразия и протяженность блоков варьируют в разных человеческих популяциях, отражая демографическую историю населения, давление естественного отбора, мутации и рекомбинации [24]. Тем не менее существуют сведения о согласованности в пространственном размещении некоторых гаплотипических блоков в различных популяциях, указывая на возможность существования общего механизма образования данных блоков, в качестве которого большинство авторов рассматривают рекомбинацию [32].

#### Генетическая дифференциация исследованных популяций

Уровень генетической дифференциации изученных популяций по частотам аллелей пяти исследованных SNPs гена *MTHFR* составил 0.018, а по частотам гаплотипов 0.021. Оценка проводилась с помощью коэффициента  $F_{st}$ . Данные о степени генной дифференциации ( $F_{st}$ ) в тотальной выборке по каждому из выбранных маркеров представлены на рис. 5. Дифференциация по локусам *rs1537516* и *rs17037397* оказалась статистически незначимой, возможно это связано с низкой частотой минорного аллеля. Все остальные полиморфные варианты показали достоверную дифференциацию. Анализ межпопуляционной дифференциации по частотам аллелей показал значимые генетические различия ( $p < 0.05$ ) между популяцией южных киргизов и остальными исследованными популяциями: русскими ( $F_{st} = 0.036$ ), тувинцами ( $F_{st} = 0.031$ ) и северными киргизами ( $F_{st} = 0.027$ ), а также выявил достоверные различия между русскими и тувинцами ( $F_{st} = 0.012$ ). Полученные результаты, несомненно, указывают на необходимость дальнейших исследований с привлечением большего количества популяций и маркеров.

В заключение хотелось бы отметить, что частоты аллелей всех изученных полиморфизмов находились в диапазоне мировых. По трем локусам (*rs4846052*, *rs1801133* и *rs1801131*) обнаружена достоверная генная дифференциация. Все четыре популяции характеризовались довольно высоким уровнем гаплотипического разнообразия, а три гаплотипа встречались с частотой более 10% во всех исследованных популяциях. У северных киргизов и тувинцев выявлена сходная структура LD в гене *MTHFR*, а наиболее сильное сцепление отмечалось в популяции южных киргизов. Полученные данные свидетельствуют о том, что архитектура LD в геноме человека носит популяционно-специфический характер и в значительной степени определяется эволюционной историей популяции. Дальнейшее изучение характера генетического разнообразия и неравновесия по сцеплению в геноме в конкретных географических, этнических или популяционных группах позволит реконструировать генетическую историю популяции, выявить следы действия естественного отбора, связанного с адаптивной изменчивостью, идентифицировать функциональные варианты, лежащие в основе предрасположенности к широко распространенным болезням.

Данная работа выполнена при частичной финансовой поддержке гранта РФФИ (№ 07-04-01629-а).



## СПИСОК ЛИТЕРАТУРЫ

1. Reich D.E., Cargill M., Bolk S. et al. Linkage disequilibrium in the human genome // *Nature*. 2001. V. 411. P. 199–204.
2. Jeffreys A.J., Kauppi L., Neumann R. Intensely punctuate meiotic recombination in the class II region of the major histocompatibility complex // *Nat. Genet.* 2001. V. 29. P. 217–222.
3. Gabriel S.B., Schaffner S.F., Nguyen H. et al. The structure of haplotype blocks in the human genome // *Science*. 2002. V. 296. P. 2225–2229.
4. Daly M.J., Rioux J.D., Schaffner S.F. et al. High-resolution haplotype structure in the human genome // *Nat. Genet.* 2001. V. 29. P. 229–232.
5. Wall J.D., Pritchard J.K. Haplotype blocks and linkage disequilibrium in the human genome // *Nature*. 2003. V. 4. P. 587–597.
6. Tishkoff S.A., Verrelli B.C. Patterns of human genetic diversity // *Annu. Rev. Genomics Hum. Genet.* 2003. V. 4. P. 293–340.
7. Wang N., Akey J.M., Zhang K. et al. Distribution of recombination crossovers and the origin of haplotype blocks: the interplay of population history, recombination, and mutation // *Am. J. Hum. Genet.* 2002. V. 71. P. 1227–1234.
8. Rosenberg N.A., Pritchard J.K., Weber J.L. et al. Genetic structure of human populations // *Science*. 2002. V. 298. P. 2381–2385.
9. Ott J. *Analysis of Human Genetic Linkage*. Baltimore and London: The Johns Hopkins Univ. Press, 1999. 382 p.
10. Goyette P., Pai A., Milos R. et al. Gene structure of human and mouse methylenetetrahydrofolate reductase // *Mammalian Genome*. 1998. V. 9. P. 652–656.
11. Электронная база данных NSBI. Режим доступа: <http://www.ncbi.nlm.nih.gov/SNP/>; dbSNP.
12. Morita H., Taguchi J., Kurihara H. Genetic polymorphism of 5,10-methylenetetrahydrofolate reductase (MTHFR) as a risk factor for coronary artery disease // *Circulation*. 1997. V. 95. P. 2032–2036.
13. Спиридонова М.Г., Степанов В.А., Пузырев В.П. О роли полиморфных вариантов гена 5,10-метилентетрагидрофолатредуктазы в патогенезе сердечно-сосудистых заболеваний // *Клинич. медицина*. 2001. Т. 2. С. 10–16.
14. Weisberg I., Tran P., Christensen B. et al. A second genetic polymorphism in methylenetetrahydrofolate reductase (MTHFR) associated with decreased enzyme activity // *Mol. Genet. Metab.* 1998. V. 64. P. 169–172.
15. Ray J.G. Meta-analysis of hyperhomocysteinemia as a risk factor for venous thromboembolic disease // *Arch. Intern. Med.* 1998. V. 26. P. 2101.
16. Nelen W.L., Steegers E.A., Eskes T.K., Blom H.J. Genetic risk factor for unexplained recurrent early pregnancy loss // *Lancet*. 1997. V. 35. P. 861.
17. Zetterberg H., Regland B., Palmer M. Increased frequency of combined methylenetetrahydrofolate reductase C677T and A1298C mutated alleles in spontaneously aborted embryos // *Eur. J. Hum. Genet.* 2002. V. 10. P. 113–118.
18. Isotalo P.A., Donnelly J.G. Prevalence of methylenetetrahydrofolate reductase mutations in patients with venous thrombosis // *Mol. Diagn.* 2000. V. 5. P. 59–66.
19. Little J., Sharp L., Duthie S., Narayanan S. Colon cancer and genetic variation in folate metabolism: The clinical bottom line // *J. Nutr.* 2003. V. 133. P. 3758–3766.
20. Марусин А.В., Степанов В.А., Спиридонова М.Г. и др. Полиморфизм генов этанолметаболизирующих ферментов *ADH1B*, *ADH7* и *CYP2E1* и риск развития алкоголизма в русской популяции Западно-Сибирского региона // *Мед. генетика*. 2006. Т. 5. № 7. С. 51–56.
21. Степанов В.А. *Этногеномика населения Северной Евразии*. Томск: Печатная мануфактура, 2002. 243 с.
22. Frosst P., Blom H.J., Milos R. et al. A candidate genetic risk factor for vascular disease: A common mutation methylenetetrahydrofolate reductase // *Nat. Genet.* 1995. № 10. P. 111–113.
23. Friedman G., Goldschmidt N., Friedlander Y. A common mutation A1298C in human methylenetetrahydrofolate reductase gene: Association with plasma total homocysteine and folate concentrations // *J. Nutr.* 1999. № 129. P. 1656–1661.
24. Shifman S., Kuypers J., Kokoris M. et al. Linkage disequilibrium patterns of the human genome across populations // *Hum. Mol. Genet.* 2003. V. 7. P. 771–776.
25. Barrett J.C., Fry B., Maller J., Daly M.J. Haploview: Analysis and visualization of LD and haplotype maps // *Bioinformatics*. 2005. № 21. P. 263–265.
26. Abecasis G.R., Ghosh D., Nichols T.E. Linkage disequilibrium: Ancient history drives the new genetics // *Hum. Hered.* 2005. V. 59. № 2. P. 118–124.
27. Freedman M.L., Reich D., Penney K.L. et al. Assessing the impact of population stratification on genetic association studies // *Nat. Genet.* 2004. V. 36. № 4. P. 388–393.
28. Johnson G.C., Todd J.A. Strategies in complex disease mapping // *Curr. Opin. Genet. Dev.* 2000. V. 10. № 3. P. 330–334.
29. Zhang K., Qin Z., Chen T. et al. HapBlock: Haplotype block partitioning and tag SNP selection software using a set of dynamic programming algorithms // *Bioinformatics*. 2005. V. 21. P. 131–134.
30. Song C.M., Yeo B.H., Tanso E. et al. iHAP – integrated haplotype analysis pipeline for characterizing the haplotype structure of genes // *Bioinformatics*. 2006. № 7. P. 525.
31. Liskova P., Hysi P.G., Williams D. et al. Study of p.N247S KERA mutation in a British family with cornea plana // *Molecular Vision*. 2007. V. 13. P. 13339–13347.
32. Rana N.A., Ebenezer N.D., Webster A.R. et al. Recombination hotspots and block structure of linkage disequilibrium in the human genome exemplified by detailed analysis of *PGM1* on 1p31 // *Hum. Mol. Genet.* 2004. V. 24. P. 3089–3102.

## Genetic Diversity and the Structure of Linkage Disequilibrium in the Methylenetetrahydrofolate Reductase Locus

E. A. Trifonova, M. G. Spiridonova, and V. A. Stepanov

*Institute for Medical Genetics, Russian Academy of Medical Sciences, Tomsk, 634050 Russia;*

*e-mail: vadim.stepanov@medgenetics.ru*

Investigation of linkage disequilibrium block architecture in human genome is modern, intensely investigated field of molecular genetics. In the present study, genetic differentiation and linkage disequilibrium pattern in the methylenetetrahydrofolate reductase (*MTHFR*) locus was examined in the populations of Russians, Tuvinians, and Northern and Southern Kyrgyzes. Methylenetetrahydrofolate reductase is the key enzyme of folate cycle, responsible for reduction of 5,10-methylenetetrahydrofolate to 5-methyltetrahydrofolate. Decreased enzymatic activity of this protein often caused by certain associations of *MTHFR* alleles results in the increased plasma homocysteine levels. In the population groups examined, genotype and allele frequencies at five polymorphic *MTHFR* loci: *rs17037397*, *rs4846052*, *rs1801133*, *rs1801131*, and *rs1537516* were evaluated. Statistically significant genetic differences between the population group of Southern Kyrgyzes and the other groups, as well as between Russians and Tuvinians, were demonstrated. In the *MTHFR* gene from the population of Southern Kyrgyzes one block was revealed; in the populations of Russians, Tuvinians, and Northern Kyrgyzes two blocks were detected. Thus, the structure of linkage disequilibrium in the *MTHFR* locus demonstrated population-specific pattern.