

УДК 575.174

НЕСТАБИЛЬНОСТЬ МИКРОСАТЕЛЛИТНОГО CAG-ПОВТОРА В ДВУХ ГРУППАХ ГАПЛОТИПОВ ГЕНА МИТОХОНДРИАЛЬНОЙ γ-ДНК-ПОЛИМЕРАЗЫ ЧЕЛОВЕКА

© 2009 г. Б. А. Мальярчук*, М. А. Перкова, М. В. Деренко

Институт биологических проблем Севера Дальневосточного отделения Российской академии наук,
Магадан, 685000

Поступила в редакцию 16.10.2008 г.

Принята к печати 05.02.2009 г.

Исследован полиморфизм двух однонуклеотидных локусов rs2238296 (Т/С) и rs758130 (Т/С) гена митохондриальной ДНК-полимеразы γ (*POLG1*) у индивидов различного этнорасового происхождения (у русских и бурят) с известными генотипами микросателлитного CAG-повтора этого гена. Обнаружено, что аллели с числом повторов, не равным 10, значительно чаще наблюдаются в составе гаплотипа ТТ, чем СС. Результаты филогенетического анализа нуклеотидных последовательностей интрона 2 гена *POLG1* у человека и шимпанзе показали, что более гетерогенный в отношении полиморфизма CAG-повтора гаплотип ТТ является более молодым в сравнении с гаплотипом СС. Полученные данные могут быть использованы для дальнейших исследований связи между полиморфизмом CAG-повтора гена *POLG1* и мужским бесплодием.

Ключевые слова: митохондриальная ДНК-полимераза γ (*POLG1*) человека, микросателлитная нестабильность, анализ сцепления, популяционный полиморфизм.

DIFFERENT INSTABILITY OF MICROSATELLITE CAG REPEAT IN TWO GROUPS OF HAPLOTYPES OF THE HUMAN MITOCHONDRIAL γ-DNA-POLYMERASE GENE, by B. A. Malyarchuk*, M. A. Perkova, M. V. Derenko (Institute of Biological Problems of the North, Far-Eastern Division, Russian Academy of Sciences, Magadan, 685000 Russia; *e-mail: malyarchuk@ibpn.ru). Polymorphism of two single nucleotide loci rs758130 (T/C) and rs2238296 (T/C) of mitochondrial DNA-polymerase gamma (*POLG1*) gene has been studied in individuals of different ethn racial ancestry (in Russians and Buryats) with already known genotypes of microsatellite CAG-repeat of this gene. It was found that alleles with the number of repeats unequal to 10 are more often detected on the background of haplotype TT than CC. Results of phylogenetic analysis of the *POLG1* gene intron 2 nucleotide sequences in humans and chimpanzee have shown that haplotype TT heterogeneous in relation to CAG-repeat polymorphism appears to be evolutionary younger than haplotype CC. The data obtained can be used for further studies of relationships between CAG-repeat polymorphisms and male infertility.

Key words: human mitochondrial γ-DNA-polymerase (*POLG1*), microsatellite instability, linkage disequilibrium, population polymorphism

Репликация митохондриальной ДНК (мтДНК) человека осуществляется с помощью митохондриальной ДНК-полимеразы γ, которая состоит из двух субъединиц – каталитической, кодируемой геном *POLG1* на хромосоме 15q25, и дополнительной, кодируемой геном *POLG2* на хромосоме 17q23-24 [1]. Исследования изменчивости гена *POLG1* проводились, в основном, на выборках пациентов с наследственной митохондриальной патологией, а также в контрольных выборках лиц преимущественно европейского происхождения [2–9]. Это позволило выявить целый ряд мутаций *POLG1*, которые ассоции-

руются с множественными делециями и нуклеотидными заменами в мтДНК скелетных мышц и других тканей у больных (список мутаций приводится, например, в базе данных Human DNA Polymerase Gamma Mutation Database – www.tools.niehs.nih.gov/polg).

Наиболее изучен в настоящее время полиморфизм микросателлитного повтора в экзоне 2 гена *POLG1* [8–12]. Длина CAG-повторов в аллелях изменяется от 6 до 13 тринуклеотидов. По результатам скрининга CAG-полиморфизма гена *POLG1* в популяциях Евразии (всего исследовано 1330 индивидов из 12 этнических групп) установлено, что аллель, представленный 10 повторами тринуклеотида CAG

* Эл. почта: malyarchuk@ibpn.ru

(10-CAG-аллель), встречается наиболее часто во всех изученных евразийских популяциях: от Польши до Кореи его частота изменяется в диапазоне от 88 до 96% соответственно [12]. Из остальных аллелей, характеризующихся числом повторов, не равным 10 (не-10-CAG-аллели), наиболее высока частота аллеля с 11 повторами (6.7%).

Полиморфизм CAG-повтора может иметь функциональное значение, так как он кодирует полиглутаминовый участок полимеразы, изменение длины которого может повлиять на точность синтеза ДНК и привести соответственно к ошибкам репликации (точечным делециям и инсерциям нуклеотидов и нуклеотидным заменам мтДНК) [13]. Показано, что частота генотипа $(CAG)_{не-10}/(CAG)_{не-10}$ достоверно повышена в выборках мужчин, характеризующихся бесплодием (азооспермией), а также у пациентов с семиномой яичка [8, 11]. Хотя в других исследованиях ассоциация генотипа $(CAG)_{не-10}/(CAG)_{не-10}$ с азооспермией не подтвердилась [9, 14–17], возможность существования такой ассоциации в определенных условиях вполне вероятна.

Частота генотипа $(CAG)_{не-10}/(CAG)_{не-10}$ варьирует на этнорасовом уровне от 0.5% в популяциях Восточной Азии до 2.4% в популяциях Европы [12] и до 9–11% в популяциях Африки [9]. Исследования однонуклеотидного полиморфизма гена *POLG1* показали, что имеются две мажорные группы сцепленных аллелей (гаплотипов) этого гена ([18]; данные проекта HarMap (www.harmap.org)). Так, Буйкиным и соавт. [18] в исследовании полиморфизма локусов rs758130 (T/C) и rs2238296 (T/C) гена *POLG1* в популяциях тувинцев, алтайцев, якутов и русских обнаружено полное (или почти полное у тувинцев) неравновесие по сцеплению аллелей исследованных локусов. Между тем, характер распределения аллелей микросателлитного CAG-повтора относительно двух групп гаплотипов гена *POLG1* до сих пор не изучали, хотя такая информация может быть полезна для выяснения причин ассоциации между вариантами полиморфизма микросателлитного локуса и мужским бесплодием.

Цель настоящей работы – исследование полиморфизма двух локусов гена *POLG1* (rs758130 и rs2238296) у индивидов различного этнорасового происхождения (у русских и бурят) с известными генотипами микросателлитного CAG-повтора, включая анализ распределения его вариантов в составе различных гаплотипов.

ЭКСПЕРИМЕНТАЛЬНАЯ ЧАСТЬ

Исследованы выборки индивидов из русского населения Белгородской области ($n = 50$) и бурят из различных районов Республики Бурятия ($n = 94$). Полиморфизм локусов rs758130 (T/C) и rs2238296 (T/C), расположенных в интроне 2 гена *POLG1* на расстоянии примерно 1500 п.н. друг от друга, ис-

следовали с помощью рестрикционного анализа согласно методике, описанной ранее Буйкиным и соавт. [18]. Для анализа использовали опубликованные нами ранее данные об изменчивости CAG-повтора экзона 2 гена *POLG1* в указанных выше популяциях [12]. Этот участок соответствует номеру rs28567406 в базе данных dbSNP (www.ncbi.nlm.nih.gov/SNP) и расположен примерно в 1500 п.н. в 5'-направлении от локуса rs2238296. Таким образом, порядок расположения локусов следующих (в направлении 5' → 3'): rs28567406 (микросателлитный CAG-повтор) – rs2238296 – rs758130. Используемая в работе нумерация нуклеотидных позиций соответствует таковой в последовательности гена *POLG1*, находящейся в базе данных GenBank (www.ncbi.nlm.nih.gov/Genbank) под номером AC133637.

Для анализа изменчивости нуклеотидной последовательности интрона 2 гена *POLG1* исследовали участок ДНК длиной 716 п.н., расположенный между локусами rs28567406 (CAG-повтор) и rs2238296. Этот участок ДНК амплифицировали, используя праймеры 5'-ACAACCTGGACCAGCACTTC-3' и 5'-AAAGGCTGGGGATGCTAAAT-3', подобранные с помощью программы Primer3 [19]. 35 циклов ПЦР проводили в следующем температурном режиме: 94°C – 30 с, 50°C – 60 с и 72°C – 60 с. Амплифицированные участки ДНК секвенировали с использованием набора для циклического секвенирования ДНК Big Dye Terminator (“Applied Biosystems”, v. 3.1) и генетического анализатора ABI Prism 3130 (“Applied Biosystems”, США). Для выравнивания и анализа нуклеотидных последовательностей использовали программы пакета MEGA 3.1 [20].

Соответствие распределения генотипов равновесию Харди-Вайнберга, неравновесие по сцеплению и структуру гаплотипов анализировали с помощью пакета программ Arlequin 3.01 [21]. Для определения частоты гаплотипов использовали алгоритм EM, основанный на методе максимального правдоподобия [22]. Наиболее вероятные структуры гаплотипов реконструировали из мультилокусных генотипов с помощью алгоритма ELB, основанного на псевдобайесовском подходе [21]. Степень дивергенции между нуклеотидными последовательностями интрона 2 гена *POLG1* рассчитывали, основываясь на r -дистанциях, с помощью пакета программ MEGA 3.1 [20].

РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

Исследование полиморфизма локусов rs758130 и rs2238296 гена *POLG1* у русских и бурят показало, что генотипы распределены в популяциях сходным образом (табл. 1). Частота аллеля rs758130*С у бурят составила 42%, у русских – 38%; частота аллеля rs2238296*С у бурят составила 36%, у русских – 35%. В исследованных выборках распределение генотипов по полиморфизму отдельных локусов не откло-

Таблица 1. Распределение генотипов по трем локусам гена *POLG1* у русских и бурят

Локусы			Частота в популяциях	
CAG-повтор, число	rs2238296	rs758130	русские (n = 50)	буряты (n = 94)
10/10	T/T	T/T	0.22 (11)	0.29 (27)
10/10	T/C	T/C	0.34 (17)	0.43 (40)
10/10	C/C	C/C	0.14 (7)	0.11 (10)
10/10	T/T	T/C	0	0.02 (2)
10/10	T/C	C/C	0.06 (3)	0.04 (4)
10/10	T/T	C/C	0	0.02 (2)
10/не-10	T/T	T/T	0.16 (8)	0.05 (5)
10/не-10	T/C	T/C	0.02 (1)	0.04 (4)
10/не-10	T/T	T/C	0.02 (1)	0
не-10/не-10	T/T	T/T	0.02 (1)	0
не-10/не-10	T/C	T/C	0.02 (1)	0

Таблица 2. Распределение гаплотипов по трем локусам гена *POLG1* у русских и бурят

Локусы			Частота в популяциях	
CAG-повтор, число	rs2238296	Rs758130	русские (2n = 100)	буряты (2n = 188)
10	T	T	0.47 (47)	0.54 (101)
10	C	C	0.35 (35)	0.36 (68)
10	T	C	0.04 (4)	0.05 (10)
не-10	T	T	0.13 (13)	0.05 (9)
не-10	C	C	0.01 (1)	0

няется от равновесия Харди-Вайнберга. Полученные данные свидетельствуют о существовании практически полного неравновесия по сцеплению между локусами rs758130 и rs2238296 ($r^2 = 0.844$ у русских, $r^2 = 0.799$ у бурят, $D' = 1$ и $p = 0$ для обеих популяций). Это согласуется с результатами работы Буйкина и соавт. [18], которые исследовали распределение генотипов и гаплотипов гена *POLG1* по локусам rs758130 и rs2238296 у населения Сибири (у тувинцев, алтайцев, якутов и русских).

В популяциях распространены, главным образом, два гаплотипа – ТТ и СС (табл. 2). Гаплотип ТТ зарегистрирован у русских и бурят с частотой 60 и 59% соответственно, гаплотип СС – с частотой 36% в обеих популяциях. Гаплотип ТС, имеющий, по-видимому, рекомбинантное происхождение, зарегистрирован с частотой 4% у русских и 5% у бурят. Этот результат, в общем, согласуется с данными Буйкина и соавт. [18], у которых гаплотип ТС у тувинцев обнаруживается с частотой 2%.

При анализе характера распределения генотипов микросателлитного CAG-повтора показано, что имеется сцепление между ним и локусами rs758130 и rs2238296 ($D' > 0.8$, $p < 0.05$). При этом аллели, харак-

теризующиеся числом повторов, не равным 10 (не-10-CAG-аллели), в популяциях русских и бурят обнаруживаются преимущественно в составе гаплотипа ТТ (табл. 2). Всего их пять: $(CAG)_{10}TT$, $(CAG)_{10}CC$, $(CAG)_{10}TC$, $(CAG)_{не-10}TT$ и $(CAG)_{не-10}CC$. Аллель не-10-CAG в составе гаплотипа СС наблюдался только один раз в русской выборке (табл. 2), в то время как весь спектр аллелей (от 6 до 12 повторов) микросателлитного CAG-повтора отмечен у ТТ-гаплотипов.

Чтобы уточнить распределение аллелей CAG-повтора относительно гаплотипов, мы изучили полиморфизм локусов rs758130 и rs2238296 в группе индивидов ($n = 35$), характеризующихся генотипами $(CAG)_{10}/(CAG)_{не-10}$ и $(CAG)_{не-10}/(CAG)_{не-10}$, из общей выборки, насчитывающей 556 человек из различных популяций русского населения европейской части России. Нами выявлено шесть гаплотипов: ранее обнаруженные $(CAG)_{10}TT$, $(CAG)_{10}CC$, $(CAG)_{не-10}TT$, $(CAG)_{не-10}CC$ и новые $(CAG)_{10}CT$, $(CAG)_{не-10}TC$ (табл. 3 и 4). Как видно, не-10-аллели CAG-повтора крайне редко обнаруживаются в составе гаплотипа СС – всего 2 случая (3%) против 35 случаев (50%), зарегистрированных в составе гаплотипа ТТ (табл. 4). При этом необходимо отметить,

Таблица 3. Распределение генотипов по трем локусам гена *POLG1* в выборке лиц из этнических русских, характеризующихся генотипами (CAG)₁₀/(CAG)_{не-10} и (CAG)_{не-10}/(CAG)_{не-10}

Локусы			Частота в выборке ($n = 35$)
СAG-повтор, число	rs2238296	rs758130	
10/не-10	Т/Т	Т/Т	0.29 (10)
10/не-10	Т/С	Т/С	0.46 (16)
10/не-10	С/С	С/С	0.03 (1)
10/не-10	Т/Т	Т/С	0.03 (1)
10/не-10	Т/С	С/С	0.03 (1)
10/не-10	Т/С	Т/Т	0.03 (1)
не-10/не-10	Т/Т	Т/Т	0.09 (3)
не-10/не-10	Т/С	Т/С	0.03 (1)
не-10/не-10	Т/Т	Т/С	0.03 (1)

Таблица 4. Распределение гаплотипов по трем локусам гена *POLG1* в выборке из русских индивидов, характеризующихся генотипами (CAG)₁₀/(CAG)_{не-10} и (CAG)_{не-10}/(CAG)_{не-10}

Локусы			Частота в выборке ($2n = 70$)
СAG-повтор, число	rs2238296	rs758130	
10	Т	Т	0.16 (11)
10	С	С	0.26 (18)
10	С	Т	0.01 (1)
не-10	Т	Т	0.50 (35)
не-10	С	С	0.03 (2)
не-10	Т	С	0.04 (3)

что спектр аллелей СAG-повтора, выявленных в составе гаплотипов ТТ и СС, частично пересекался: весь спектр от 6 до 12 повторов выявлен в составе ТТ-гаплотипа, и часть спектра от 8 до 12 повторов зарегистрирована в составе СС-гаплотипа. В целом же, результаты показывают, что нестабильность микросателлитного СAG-повтора, определяемая частотой аллелей с числом повторов, не равным 10, в значительно большей степени выражена в составе гаплотипа ТТ, чем СС.

Исследование изменчивости участка гена *POLG1* длиной 716 п.н., расположенного в интроне 2 между СAG-повтором и локусом rs2238296, у 20 гомозиготных по локусам rs2238296 и rs758130 индивидов показало, что в нем имеются три однонуклеотидные замены, которые известны из базы данных dbSNP под номерами rs2283430, rs2239286 и rs2247233. Оказалось, что гаплотипу СС соответствует гаплотип ТСА для локусов rs2283430, rs2239286 и rs2247233, а

гаплотипу ТТ – два гаплотипа GGG и GGA. Сравнительный анализ нуклеотидных последовательностей этого участка гена *POLG1* у человека и шимпанзе (NW_001225258 в базе данных GenBank) показал, что в филогенетическом отношении гаплотип ТТ более молодой: его последовательности в большей степени отличаются от последовательности гаплотипа шимпанзе (в среднем, 1.1% дивергенции), чем последовательность гаплотипа СС (0.76%).

Из результатов настоящей работы следует также, что идентичные не-10-СAG-аллели формировались независимо в разных группах гаплотипов гена *POLG1*. Этот факт необходимо учитывать при проведении молекулярно-генетических исследований мужчин с бесплодием, поскольку вполне возможно, что проявление азооспермии может ассоциироваться с генотипами (CAG)_{не-10}/(CAG)_{не-10} какой-то одной из разновидностей гена *POLG1*. Возможно, этим и объясняется наблюдаемая в литературе противоречивость результатов исследования ассоциативных связей между мужским бесплодием и полиморфизмом микросателлитного СAG-повтора гена *POLG1*. Для подтверждения этого предположения требуются дополнительные исследования полиморфизма гена *POLG1* в норме и при нарушении сперматогенеза.

Авторы выражают благодарность И.К. Дамбуевой за помощь в проведении данного исследования.

Работа получила финансовую поддержку Дальневосточного отделения Российской академии наук (06-3-А-06-176, 09-3-А-06-221).

СПИСОК ЛИТЕРАТУРЫ

- Walker R.L., Anziano P., Meltzer P.S. 1997. A PAC containing the human mitochondrial DNA polymerase gamma gene (*POLG*) maps to chromosome 15q25. *Genomics*. **40**, 376–378.
- van Goethem G., Dermaut B., Lofgren A., Martin J.J., van Broeckhoven C. 2001. Mutations of *POLG* is associated with progressive external ophthalmoplegia characterized by mtDNA deletions. *Nature Genet.* **28**, 211–212.
- Barthelemy C., de Baulny H.O., Lombes A. 2002. D-loop mutations in mitochondrial DNA: link with mitochondrial DNA depletion? *Hum. Genet.* **110**, 479–487.
- Lamantea E., Tiranti V., Bordoni A., et al. 2002. Mutations of mitochondrial DNA polymerase gamma are a frequent cause of autosomal dominant and recessive progressive external ophthalmoplegia. *Ann. Neurol.* **52**, 211–219.
- Del Bo R., Bordoni A., Sciacco M., Di Fonzo A., Galbati S., Crimi M., Bresolin N., Comi G.P. 2003. Remarkable infidelity of polymerase gamma A associated with mutations in *POLG1* exonuclease domain. *Neurology*. **61**, 903–908.
- Di Fonzo A., Bordoni A., Crimi M., Sara G., Del Bo R., Bresolin N., Comi G.P. 2003. *POLG* mutations in sporadic mitochondrial disorders with multiple mtDNA deletions. *Hum. Mutat.* **22**, 498–499.
- Kollberg G., Jansson M., Perez-Bercoff A., Melberg A., Lindberg C., Holme E., Moslemi A.-R., Oldfors A. 2005. Low frequency of mtDNA point mutations in patients with

- PEO associated with *POLG1* mutations. *Eur. J. Hum. Genet.* **13**, 463–469.
8. Blomberg Jensen M., Leffers H., Petersen J.H., Daugaard G., Skakkebaek N.E., Rajpert-De Meyts E. 2008. Association of the polymorphism of the CAG repeat in the mitochondrial DNA polymerase gamma gene (*POLG*) with testicular germ-cell cancer. *Ann. Oncol.* **19**, 1910–1914.
 9. Westerveld G.H., Kaaij-Visser L., Tanck M., van der Veen F., Repping S. 2008. CAG repeat length variation in the polymerase gamma (*POLG*) gene: effect on semen quality. *Mol. Hum. Reprod.* **14**, 245–249.
 10. Rovio A.T., Marchington D.R., Donat S., et al. 2001. Mutations at the mitochondrial DNA polymerase (*POLG*) locus associated with male infertility. *Nature Genet.* **29**, 261–262.
 11. Rovio A.T., Abel J., Ahola A.L., et al. 2004. A prevalent *POLG* CAG microsatellite length allele in humans and African great apes. *Mamm. Genome.* **15**, 492–502.
 12. Malyarchuk B.A., Papuga M., Grzybowski T., Wozniak M., Derenko M.V., Rogozin I.B., Rychkov S.Y., Czarny J., Zakharov I.A., Miscicka-Sliwka D. 2005. Low variability of the *POLG* (CAG)_n repeat in North Eurasian populations. *Hum. Biol.* **77**, 355–365.
 13. Copeland W.C., Ponamarev M.V., Nguyen D., Kunkel T.A., Longley M.J. 2003. Mutations in DNA polymerase gamma cause error prone DNA synthesis in human mitochondrial disorders. *Acta Biochim. Polonica.* **50**, 155–167.
 14. Krausz C., Guarducci E., Becherini L., Degl' Innocenti S., Gerace L., Balercia G., Forti G. 2004. The clinical significance of the *POLG* gene polymorphism in male infertility. *J. Clin. Endocrinol. Metab.* **89**, 4292–4297.
 15. Akinin-Seifer I.E., Touraine R.L., Lejeune H., Jimenez C., Chouteau J., Siffroi J.P., McElreavey K., Bienvenu T., Patrat C., Levy R. 2005. Is the CAG repeat of mitochondrial DNA polymerase gamma (*POLG*) associated with male infertility? A multi-centre French study. *Hum. Reprod.* **20**, 736–740.
 16. Brusco A., Michielotto C., Gatta V., et al. 2006. The polymorphic polyglutamine repeat in the mitochondrial DNA polymerase gamma gene is not associated with oligozoospermia. *J. Endocrinol. Invest.* **29**, 1–4.
 17. Harris T.P., Gomas K.P., Weir F., Holyoake A.J., McHugh P., Wu M., Sin Y., Sin I.L., Sin F.Y. 2006. Molecular analysis of polymerase gamma gene and mitochondrial polymorphism in fertile and subfertile men. *Int. J. Androl.* **29**, 421–433.
 18. Буйкин С.В., Голубенко М.В., Погребенкова В.В., Цимбалюк И.В., Пузырев В.П. 2006. Ген митохондриальной γ-полимеразы (*POLG*): Частота и анализ сцепления двух однонуклеотидных замен (SNP) в популяциях народов Сибири. *Молекуляр. биология.* **40**, 1081–1083.
 19. Rozen S., Skaletsky H. 2000. In: *Bioinformatics Methods and Protocols: Methods in Molecular Biology*. Eds Krawetz S., Misener S. Totow: Humana Press, pp. 365–386.
 20. Kumar S., Tamura K., Nei M. 2004. MEGA3: Integrated software for molecular evolutionary genetics analysis and sequence alignment. *Brief. Bioinform.* **5**, 150–163.
 21. Excoffier L., Laval G., Schneider S. 2005. Arlequin ver. 3.0: An integrated software package for population genetics data analysis. *Evol. Bioinform. Online* **1**, 47–50.
 22. Excoffier L., Slatkin M. 1995. Maximum-likelihood estimation of molecular haplotype frequencies in a diploid population. *Mol. Biol. Evol.* **12**, 921–927.