

Определение полиморфизма 5mC/T в позиции Chr1: 245618129 в препаратах ДНК из крови человека методами Glal- и FatI-ПЦР анализа

 epigendx.online/ru/2021/snp2/

А.Г. Акишев^{1*}, Н.А. Нетесова², М.А. Абдурашитов¹, С.Х. Дегтярев¹

¹ НПО «СибЭнзим» Новосибирск

² ООО «Эпигенлаб» Новосибирск

* автор для переписки: Александр Акишев, НПО СибЭнзим, ул. Тимакова, д. 2/12, Новосибирск, 630117, Россия; тел.: +7(383)333-4991; факс: +7(383)333-6853; E-mail: aki@sibenzyme.ru

Методами Glal- и FatI-ПЦР анализа определяли частоты вариантов полиморфизма 5mC/T в положении хр1: 245618129 (по геномной сборке GRCh38.p13) в препаратах ДНК, выделенной из клеток крови 92 человек. Исследование включало (1) выделение лейкоцитарной ДНК из клеток крови, (2) проведение Glal- и FatI-ПЦР анализа фрагмента ДНК хр1: 245617889 — 245618464, (3) определение 5-метилцитозина и тимина в позиции хр1: 245618129 в анализируемых препаратах ДНК и (4) сравнительный анализ полученных результатов.

Показано, что 43 донора (46,74%) имеют гетерозиготный набор С/Т в положении хр1: 245618129, 28 доноров (30,43%) гомозиготны по Т, а 21 донор (22,83%) — по С. Таким образом, принимая во внимание, что клетки крови имеют диплоидный набор хромосом, замена С на Т встречается в 99 из 184 проанализированных вариантов (53,8%). При этом из полученных результатов следует, что цитозин в положении хр1: 245618129 в большей части молекул ДНК находится в метилированной форме (5-метилцитозин).

Ключевые слова: однонуклеотидный полиморфизм 5mC/T, ДНК из крови человека, Glal-ПЦР анализ, FatI-ПЦР анализ

DOI: 10.26213/SE.2019.69.42836

Данные для цитирования:

Акишев А.Г., Нетесова Н.А., Абдурашитов М.А., Дегтярев С.Х. (2021) Определение полиморфизма 5mC/T в позиции Chr1: 245618129 в препаратах ДНК из крови человека методами Glal- и FatI-ПЦР анализа, Эпигенетич ДНК диагност, том 2021(1), DOI: 10.26213/SE.2019.69.42836



Эта статья доступна по лицензии Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International

Full-text-pdf

Введение

Ранее нами было проведено изучение частот аллелей и статуса метилирования цитозинового остатка для полиморфизма С/Т в положении хр16: 75033884 (по геномной сборке GRCh38) в препаратах ДНК 92 доноров [1]. В настоящей работе было проведено изучение другого однонуклеотидного полиморфизма С/Т в положении хр1: 245618129 (rs7516354 по базе данных

dbSNP [2]) в тех же препаратах ДНК. Данный SNP расположен в интроне гена KIF26B, кодирующего кинезин-подобный белок но, в отличие от SNP в позиции chr16: 75033884, не находится в ДНК-повторе.

На рисунке 1 представлена нуклеотидная последовательность участка геномной ДНК, включающая в себя анализируемый полиморфизм. Для проведения контрольной ПЦР в этом случае был подобран фермент рестрикции HaeIII, сайты узнавания которого расположены вне участка амплификации.

Рисунок 1. Нуклеотидная последовательность анализируемого участка геномной ДНК chr1: 245617889 — 245618464.

Фиолетовым фоном выделены сайты узнавания рестриктазы HaeIII (GGCC), окаймляющие амплифицируемый участок. Зоны связывания прямого и обратного праймеров, а также флуоресцентно-меченого TaqMan-зонда выделены желтым, голубым и зеленым фоном, соответственно. Подчеркнут сайт ACGC, который при метилировании расщепляется метилзависимой эндонуклеазой Glai [3]. Цитозиновый остаток в этом сайте, который может замещаться тиминным, выделен красным цветом.

```
TACCTCAGTTTCCCCCTGCTCCTTCCACAGGCCGTGCTGCTCCCGTCTCCTGGCATTCTTCCTCCTGCCCC
TTCTGCCTCTTCTGGGACCTCACCAATTGGGACCCCTCAGTGGTCTTTTCATGCTTATGTTTTTCAGCCTTT
CCTCCCACCAAAGTCAGGGTTACAGAAGGAAGAGAGGAGGCAGGGAGGGAGGGAAAAGGAAAGGAAAGGAA
CCACAATGGCACCTTGCTCTGCCACGCTCCTCTCCGATCGCCTCCTCAACAAAGACGTCAGATTCCCCCTCT
TTGTAATGTCTTCTTTGAATTTACAGCAGTGTGGGGATGTTATGCTTTTGCCTTACTCCGCCTTATATTTTG
GCTGTTTGTGAGCTAGTCCAAGCCCTATTGGACTATAGGTGCCTTGAGAAAGAGTGCCACAGTGCCAGGGCT
GTGTCTGCTGCATGCTGTTTGTGAGCTTGTCCAAGCCCTATTAGACTATAGGTTTCCTTGAGACAGAGTGCCA
CCGTCCTGGGGCTATGTCCGCAGCGTCAGTATCCAAGGCCTATTAGACTATAGGTTTCCTTGAGACAGAGTGC
```

При замене первого нуклеотида С в сайте ACGC на Т возникает сайт узнавания рестриктазы FatI (CATG). Как было показано ранее [1], наличие SNP в сайте CACGC и определение замены 5mC на Т в одной или обеих копиях хромосомы 1 диплоидного набора генома может быть установлено путем проведения Fat- и Glai-ПЦР анализа и определения ΔCq_T и ΔCq_5mC при сравнении с данными контрольного ПЦР. При наличии неметилированного цитозина в позиции SNP в меньшей части молекул ДНК, Glai-ПЦР анализ позволяет оценить долю 5mC.

Таким образом, методами Glai- и FatI-ПЦР анализа мы определяли частоты вариантов полиморфизма 5mC/Т в положении chr1: 245618129 (по геномной сборке GRCh38.p13) в препаратах ДНК, выделенной из клеток крови 92 человек.

Материалы и методы

FatI- и Glai-ПЦР анализ образцов ДНК 92-х условно-здоровых доноров, выделенной из лейкоцитов периферической крови, проводили как описано ранее [1], за исключением замены эндонуклеазы рестрикции для контрольной ПЦР. Для выщепления амплифицируемого фрагмента при проведении контрольной ПЦР использовали рестриктазу HaeIII вместо фермента HinfI [1]. Возраст и пол доноров, а также концентрация полученных препаратов ДНК и их спектрофотометрическая чистота представлены в таблице 1 в работе [1].

Ниже указаны структуры праймеров (1-129d и 1-129r) и флуоресцентно-меченого зонда (1-129z), рассчитанных для проведения ПЦР-анализа участка chr1: 245618040 — 245618159:

1-129z 5' FAM-CCTTTTCCCTCCCTCCCTGCCTCC-BHQ1 3'(24)

1-129d 5' CCAAAGTCAGGGTTACAGAAGGAAGAGA 3'(28)

1-129r 5' CTT TGT TGA GGA GGC GAT CGG AG 3'(23)

ПЦР в реальном времени проводили согласно протоколу производителя [4] в объеме 20 мкл на детектирующем амплификаторе CFX-96 («Bio-Rad», США) по следующей программе:

при 95°C — 3 мин; 5 «слепых» циклов: 95°C — 10 сек, 66°C — 20 сек, 72°C — 15 сек; 40 циклов: 95°C — 10 сек, 66°C — 20 сек (с детекцией флуоресцентного сигнала в канале FAM), 72°C — 15 сек.

По завершении ПЦР, при помощи программного обеспечения амплификатора “Bio-Rad CFX Manager v.2.1”, устанавливали среднее значение Cq и значение среднеквадратичного отклонения для анализируемых образцов N1, N2 и N3 как описано ранее [1] и полученные данные заносили в таблицу.

Результаты и обсуждение

В Таблице 1 показаны значения Cq, полученные при ПЦР-анализе гидролизатов ДНК 92-х доноров крови (столбцы 2-4). В столбцах 5 и 6 указаны значения величин Cq, полученные в FatI- и GluI-ПЦР анализе, за вычетом значения Cq в столбике 2 (контрольный ПЦР на HaeIII-гидролизате), обозначенные как ΔCq_T и ΔCq_5mC , соответственно.

Таблица 1.

Значения Cq в случае гидролиза ДНК по сайтам GGCC (HaeIII), CATG (FatI-ПЦР анализ) и A(5mC)GC (GluI-ПЦР анализ) и значения разницы величины Cq для двух последних реакций относительно первой ($\Delta Cq_T = \Delta Cq_CATG - \Delta Cq_GGCC$ и $\Delta Cq_5mC = \Delta Cq_A(5mC)GC - \Delta Cq_GGCC$), соответственно.

	I - Cq_GGCC	II - Cq_CATG	III - Cq_A(5mC)GC	(II - I) ΔCq_T	(III - I) ΔCq_5mC
1172	21.26±0.08	29.73±0.47	21.41±0.15	8.47	0.15
1164	21.69±0.16	22.75±0.07	22.65±0.13	1.06	0.96
1147	22.05±0.12	29.97±0.48	22.25±0.08	7.92	0.2
1133	20.1±0.15	20.89±0.16	21±0.04	0.79	0.9
1187	21.65±0.1	22.9±0.14	22.74±0.03	1.25	1.09
1502	20.99±0.08	22.05±0.08	21.79±0.12	1.06	0.8
1487	21.48±0.06	21.71±0.14	25.83±0.39	0.23	4.35
1304	21.81±0.09	22.66±0.15	22.71±0.09	0.85	0.9
1218	21.41±0.03	30.87±0.41	21.2±0.05	9.46	-0.21
1379	21.23±0.17	22.27±0.03	22.2±0.04	1.04	0.97
1243	22.01±0.04	22.84±0.06	22.88±0.14	0.83	0.87
1332	21.4±0.05	21.43±0.08	24.52±0.1	0.03	3.12
1173	21.35±0.02	21.3±0.06	24.97±0.26	-0.05	3.62
1200	21.71±0.04	22.93±0.14	22.46±0.09	1.22	0.75
1203	22.09±0.09	23.16±0.1	23.18±0.12	1.07	1.09
1209	21.22±0.08	22.34±0.06	22.08±0.19	1.12	0.86
1214	21.44±0.16	29.62±0.4	21.17±0.11	8.18	-0.27
1329	20.25±0.05	21.44±0.08	21.19±0.06	1.19	0.94
1225	21.53±0.1	29.27±0.22	21.46±0.08	7.74	-0.07
1226	21.39±0.08	21.48±0.16	24.26±0.05	0.09	2.87
1221	21.46±0.05	28.9±0.34	21.37±0.05	7.44	-0.09
1134	20.55±0.13	29.84±0.48	20.67±0.01	9.29	0.12
571	22.03±0.1	22.85±0.05	22.87±0.16	0.82	0.84
1161	21.66±0.09	22.45±0.08	22.41±0.14	0.79	0.75
1192	21.14±0.08	26.97±0.06	21.02±0.04	5.83	-0.12
1184	21.45±0.07	21.33±0.13	24.41±0.14	-0.12	2.96
1208	21.36±0.09	21.34±0.09	24.48±0.35	-0.02	3.12
604	22.33±0.07	30.04±0.33	22.39±0.08	7.71	0.06
1086	19.84±0.15	26.87±0.31	19.92±0.09	7.03	0.08
1217	21.11±0.03	20.87±0.11	24.62±0.12	-0.24	3.51
1247	20.88±0.11	21.82±0.16	21.72±0.18	0.94	0.84
1235	21.48±0.12	22.61±0.12	22.44±0.17	1.13	0.96
1345	21.31±0.05	21.23±0.11	25.05±0.02	-0.08	3.74
1094	21.9±0.07	28.98±0.48	21.82±0.06	7.08	-0.08
1219	21.39±0.03	21.5±0.13	25.35±0.09	0.11	3.96
1381	21.15±0.06	22.14±0.05	22.04±0.04	0.99	0.89
1207	21.67±0.08	22.67±0.12	22.49±0.2	1	0.82
1323	21.07±0.07	21.05±0.06	23.19±0.1	-0.02	2.12
1411	21.54±0.08	28.21±0.63	21.62±0.13	6.67	0.08
1145	21.23±0.05	22.29±0.07	22.22±0.16	1.06	0.99
1171	21.6±0.03	22.61±0.06	22.57±0.08	1.01	0.97
1491	20.47±0.13	20.37±0.05	23.4±0.25	-0.1	2.93
603	22.48±0.07	22.49±0.07	25.51±0.07	0.01	3.03
1155	21.29±0.07	22.39±0.11	22.36±0.19	1.1	1.07
1338	21.38±0.15	21.41±0.14	24.53±0.07	0.03	3.15
1099	21.79±0.13	29.36±0.52	21.66±0.2	7.57	-0.13
1111	22.7±0.09	23.67±0.03	23.58±0.1	0.97	0.88

	<u>I - Cq GGCC</u>	<u>II - Cq CATG</u>	<u>III - Cq A(5mC)GC</u>	<u>(II - I) ΔCq_T</u>	<u>(III - I) ΔCq_5mC</u>
1151	22.08±0.08	29.42±0.3	21.99±0.14	7.34	-0.09
1087	21.47±0.08	27.75±0.11	21.51±0.04	6.28	0.04
1093	21.27±0.02	28.99±0.1	21.38±0.07	7.72	0.11
1129	20.24±0.09	21.14±0.18	21.06±0.12	0.9	0.82
1165	21.39±0.03	21.35±0.02	24.37±0.04	-0.04	2.98
584	21.15±0.06	22.37±0.05	22.16±0.11	1.22	1.01
1159	22.45±0.15	22.33±0.11	25.07±0.09	-0.12	2.62
1358	20.3±0.05	28.55±0.19	20.26±0.14	8.25	-0.04
587	22.23±0.1	23.27±0.08	23.24±0.06	1.04	1.01
561	23.71±0.08	24.56±0.17	24.96±0.18	1.16	0.94
1280	21.8±0.08	22.7±0.02	22.59±0.14	0.9	0.79
1127	20.66±0.15	21.72±0.08	21.51±0.16	1.06	0.85
1135	20.61±0.07	21.57±0.23	21.48±0.09	0.96	0.87
1163	21.62±0.03	22.6±0.07	22.55±0.06	0.98	0.93
1130	19.29±0.05	20.15±0.05	20.19±0.09	0.86	0.9
1137	21.43±0.02	28.25±0.29	21.47±0.1	6.82	0.04
1348	21.35±0.04	29.23±0.11	21.27±0.05	7.88	-0.08
1349	21.66±0.09	22.54±0.03	22.67±0.09	0.88	1.01
1294	21.23±0.11	22.27±0.08	22.08±0.09	1.04	0.85
1369	21.42±0.02	27.83±0.25	21.46±0.12	6.41	0.04
1277	20.75±0.08	20.7±0.04	23.44±0.04	-0.05	2.69
1152	21.77±0.11	29.74±0.19	21.91±0.11	7.97	0.14
1205	22.19±0.05	23.33±0.16	23.12±0.12	1.05	0.93
237	23.12±0.03	24.21±0.11	24.01±0.05	1.09	0.89
559	22.49±0.16	23.61±0.2	23.38±0.05	1.12	0.89
1143	21.34±0.1	21.36±0.06	24.32±0.11	0.02	2.98
1428	21.59±0.13	22.5±0.07	22.47±0.08	0.91	0.88
1365	21.29±0.01	21.24±0.03	23.97±0.1	-0.05	2.68
173	20.4±0.12	21.62±0.15	21.36±0.1	1.22	0.96
1437	21.59±0.03	29.26±0.37	21.71±0.06	7.67	0.12
240	21.69±0.12	27.61±0.29	21.76±0.13	5.59	-0.11
1146	21.64±0.04	22.55±0.08	22.45±0.05	0.91	0.81
1413	19.99±0.05	27.37±0.14	20.05±0.18	7.38	0.06
40	21.25±0.11	30.13±0.22	21.25±0.04	8.88	0
205	20.9±0.1	27.99±0.36	21.02±0.09	7.09	0.12
1115	21.23±0.1	22.24±0.17	22.1±0.02	1.01	0.87
1576	21.95±0.06	23.02±0.11	22.81±0.17	1.07	0.86
178	21.83±0.11	30.24±0.84	21.84±0.1	8.41	0.01
1669	21.45±0.12	22.61±0.24	22.3±0.04	1.16	0.85
121	21.4±0.06	22.25±0.04	22.24±0.09	0.85	0.84
1168	20.44±0.08	20.45±0.04	22.94±0.11	0.01	2.5
1174	21.55±0.04	22.53±0.23	22.47±0.11	0.98	0.92
1160	21.15±0.07	21.21±0.09	24.34±0.05	0.06	3.19
55	23.11±0.11	29.41±0.27	23.05±0.06	6.3	-0.06
56	21.36±0.11	21.25±0.07	22.84±0.12	-0.11	1.48
1202	21.81±0.08	29.27±0.14	21.65±0.06	7.46	-0.16
1162	21.46±0.01	21.36±0.08	24.12±0.17	-0.1	2.66

Анализ данных таблицы 1 показывает, что как и в предыдущей работе [1], выявляются 3 группы образцов ДНК. В первой из них значения ΔCq_T и ΔCq_5mC варьируют от 0.75 до 1.25, что соответствует двукратному уменьшению количества нерасщепленных копий амплифицируемого участка. Таким образом, образцы первой группы представляют собой гетерозиготы по аллелям T:5mC в позиции хр1: 245618129.

Во второй группе значения ΔCq_T варьируют от 5.59 до 9.46, в то время как значения ΔCq_5mC минимальны (от -0.27 до 0.2). Очевидно, что образцы ДНК второй группы практически полностью расщепляются ферментом *FatI* и не содержат сайтов узнавания *GlaI*. Это говорит о том, что в исследуемой позиции SNP находится лишь тиминовый остаток (гомозиготы по Т).

Для третьей группы образцов ДНК характерно повышенное значение ΔCq_5mC (2.12-4.35, за исключением образца №56) при небольших значениях ΔCq_T (от -0.24 до 0.23).

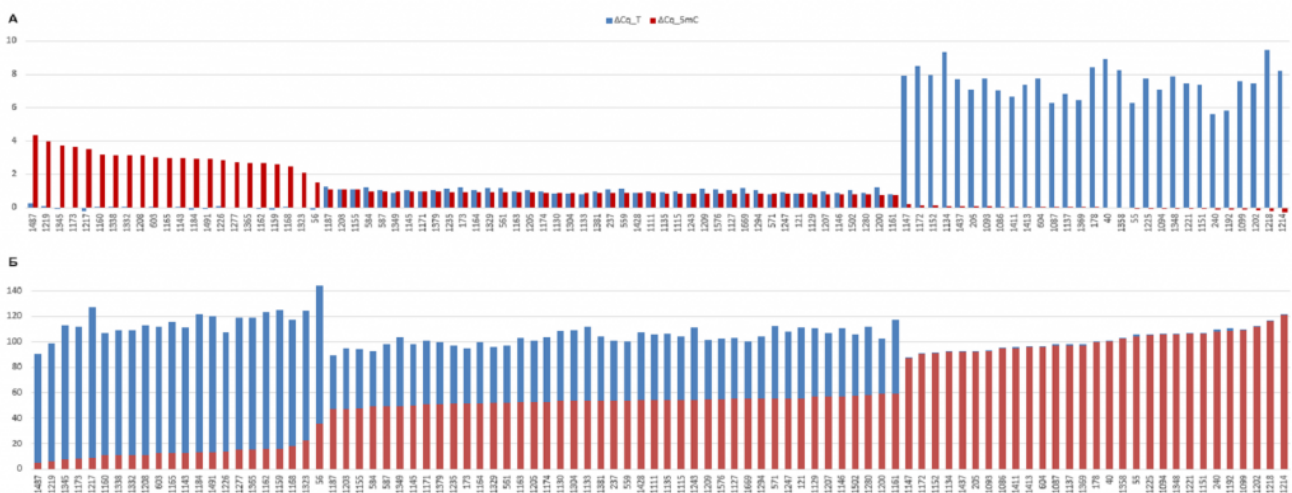
Соответственно, данные образцы можно расценивать как содержащие лишь цитозинный остаток в позиции *chr1: 245618129* (гомозиготы по С), причем этот остаток находится преимущественно в метилированной форме.

Как и в работе [1], значения ΔCq_5mC (от 2.12 до 4.35) существенно меньше величин ΔCq_T (от 5.59 до 9.46), что также, вероятно, связано с наличием неметилированного цитозина, доля которого может достигать почти четверти. Кроме того, также как в работе [1], образец ДНК №56, имеет еще более низкое значение ΔCq_5mC по сравнению с остальными величинами в первой группе. Поскольку *FatI*-ПЦР анализ образца ДНК №56 показывает отсутствие основания Т в позиции *chr1: 245618129*, то, по-видимому, доля неметилированного цитозина в этом положении свыше 30%.

На Рисунке 2А результаты, представленные в таблице 1, отсортированы в порядке уменьшения ΔCq_5mC , а на рисунке 2Б показана доля молекул ДНК, нерасщепленных в ходе *GlaI*- и *FatI*-ПЦР анализа, относительно общего количество молекул ДНК, определенных ПЦР анализом *HaeIII*-гидролизатов (выражена в процентах). Как видно из рисунка 2Б, суммарная доля молекул ДНК, нерасщепленных в ходе *GlaI*- и *FatI*-ПЦР анализа, близка к 100%, при этом на рисунках 2А и 2Б хорошо видны три варианта соответствия значений ΔCq_T и ΔCq_5mC , описанные выше.

Рисунок 2.

Сверху – Значения ΔCq_5mC и ΔCq_T *Cq* для 92-х образцов донорской ДНК лейкоцитов (красным и синим цветом, соответственно). **Снизу** — процент молекул ДНК, нерасщепленных в ходе *GlaI*- и *FatI*-ПЦР анализа, относительно общего количество молекул ДНК, определенных ПЦР анализом *HinfI*-гидролизатов (красным и синим цветом, соответственно).



Таким образом, гетерозиготный набор С:Т имеется в 43 образцах ДНК, 28 образцов представляют собой гомозиготу Т:Т, а 21 образец является гомозиготным по аллелю С. Частота аллеля Т в исследованной выборке составляет 53.8%. Полученные данные показывают, что во всех образцах ДНК цитозинный остаток преимущественно метилирован. Как было показано ранее, метилированная форма CG-динуклеотидов преобладает во всех компартаментах генома

за исключением CpG-островков и первых экзонов, выполняющих функцию регуляции экспрессии генов [5]. Нуклеотид в положении хр1: 245618129 расположен в шестом интроне гена KIF26B и также является в основном 5-метилцитозином.

В Таблице 2 показано число диплоидных сочетаний аллелей и их доля от общего количества проанализированных препаратов ДНК доноров крови, а также эти параметры отдельно для препаратов ДНК крови мужчин и женщин. Анализировались три возможных диплоидных сочетаний аллелей: гомозиготы по аллелям С/С и Т/Т и гетерозигота С/Т для положения хр1: 245618129.

Таблица 2.

Распределение диплоидных наборов аллелей в позиции хр1: 245618129 в общей выборке и отдельно по полам.

	Всего доноров	С/С		С/Т		Т/Т	
		N	%	N	%	N	%
Все доноры	92	21	22.83%	43	46.74%	28	30.43%
Женщины	46	12	26.09%	20	43.48%	14	30.43%
Мужчины	46	9	19.57%	23	50.00%	14	30.43%

Как видно из таблицы 2, общее количество гетерозигот во всей выборке (46,74%) близко к значению 50%, которое и должно быть для такого распределения.

Ранее мы проводили определение SNP 5mC/T в положении хр16: 75033884 в той же выборке препаратов ДНК 92-х доноров [1]. В Таблице 3 представлены данные по распределению гетеро- и гомозигот в положении хр16: 75033884 для каждого из трех аллельных вариантов хр1: 245618129.

Таблица 3

Распределение гомо- и гетерозигот по аллелям С и Т для двух позиций SNP.

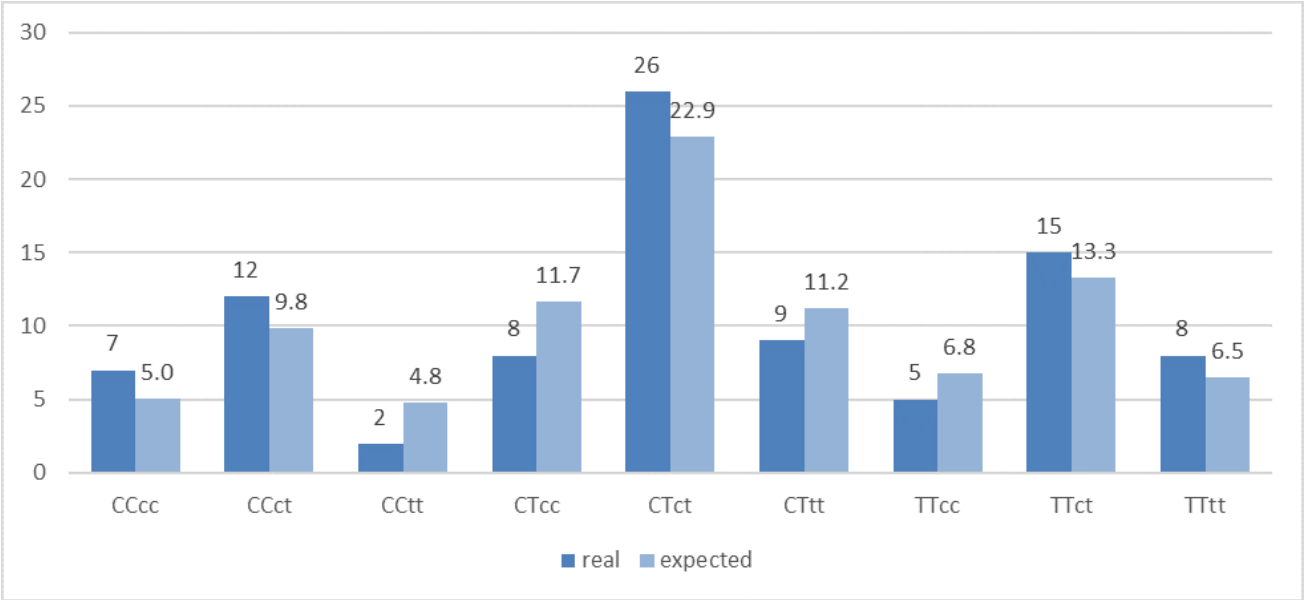
		хр1: 245618129			
		С/С	С/Т	Т/Т	Всего
хр16: 75033884	С/С	7	8	5	20
	С/Т	12	26	15	53
	Т/Т	2	9	8	19
	Всего	21	43	28	92

Как видно из таблицы, гетерозигота 5mC/T в положении хр16: 75033884 распределена приблизительно равномерно по все трем аллельным вариантам в положении хр1: 245618129. В тоже время комбинация гомозиготы по Т в положении хр16: 75033884 и гомозиготы по С в положении хр1: 245618129 встречаются существенно реже, чем другие аллельные варианты.

Мы также сравнили полученные количественные результаты с ожидаемыми по уравнению Харди-Вайнберга [6] для двух переменных позиций, исходя из полученных частот аллелей Т=0.538 для хр1: 245618129 и Т=0.495 для хр16: 75033884 (диаграмма на рисунке 3).

Рисунок 3

Количество различных генотипов по двум позициям SNP, полученных в эксперименте (столбики темно-синего цвета) и по расчетным данным (столбики светло-синего цвета). Заглавными буквами в подписях обозначены диплоидные варианты для позиции chr1: 245618129, прописными — для позиции chr16: 75033884.



Как видно из рисунка 3, в целом, за исключением 3-ей пары столбиков, наблюдается соответствие экспериментальных результатов расчетным данным. Различие в паре столбиков CCtt, вероятно, связано с относительно небольшой изучаемой выборкой.

В таблице 4 приведены данные по частотам аллелей в позициях chr1: 245618129 и chr16: 75033884 для различных популяций по данным консорциума «1000 genomes» [7], из которого видно, что полученные нами данные наиболее соответствуют результатам SNP анализа ДНК представителей европейской популяции. Для европейцев установленные частоты аллелей Т для позиций chr1: 245618129 и chr16: 75033884 составляют 0.588 [8] и 0.543 [9], соответственно, что близко к значениям частот 0.538 и 0.495 для этих позиций, полученных нами ранее [1] и в этой работе.

Таблица 4

Частоты аллелей С и Т для двух позиций SNP по данным проекта «1000 геномов»

		chr1: 245618129		chr16: 75033884	
Популяция	Число аллелей	С	Т	С	Т
Африканцы	1322	0.6346	0.3654	0.9402	0.0598
Европейцы	1006	0.4125	0.5875	0.4573	0.5427
Американцы	694	0.39	0.61	0.578	0.422
Южные азиаты	978	0.298	0.702	0.726	0.274
Восточные азиаты	1008	0.129	0.871	0.8323	0.1677
Всего	5008	0.3886	0.6114	0.7294	0.2706

На рисунке 4 приведено сравнение экспериментальных данных, полученных при секвенировании 503 геномов (1006 аллелей) жителей Европы, с теоретическими расчетами в форме, аналогичной рис.3. Как видно из сравнения рисунков 3 и 4, картины распределения генотипов по позициям $chr1: 245618129$ и $chr16: 75033884$, выявленные у европейцев и полученные в нашей работе, отличаются незначительно. На рисунке 5 эти данные представлены вместе в виде сравнения по частотам SNP. Как видно из рис.5 кратные различия в частотах SNP для двух выборок, полученные в эксперименте, наблюдаются в случае гомозигот по С в обоих сайтах, а также в случае гомозиготы по С для позиции $chr1: 245618129$ и гомозиготы по Т для позиции $chr16: 75033884$. И если второй случай, как было отмечено выше, можно объяснить малым размером выборки в нашей работе, то различие в доле гомозигот по С в обоих положениях является, по-видимому, отражением реальных различий генотипов в изучаемых выборках ДНК жителей Сибири и Европы. Известно, что распределение гомо- и гетерозигот в реально существующих популяциях сильно зависит от миграционных потоков и выделения в обществе более или менее обособленных субпопуляций [10], что характерно в целом для современного мира.

Рисунок 4

Количество различных генотипов по двум позициям SNP, полученных в эксперименте (столбики темно-оранжевого цвета) и по расчетным данным (столбики светло-оранжевого цвета) для 503 геномов европейской популяции из базы данных «1000 геномов»

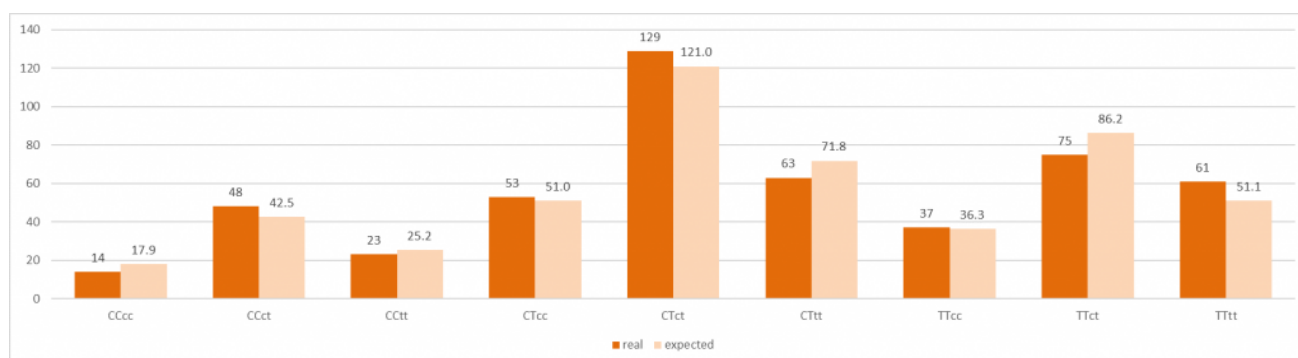
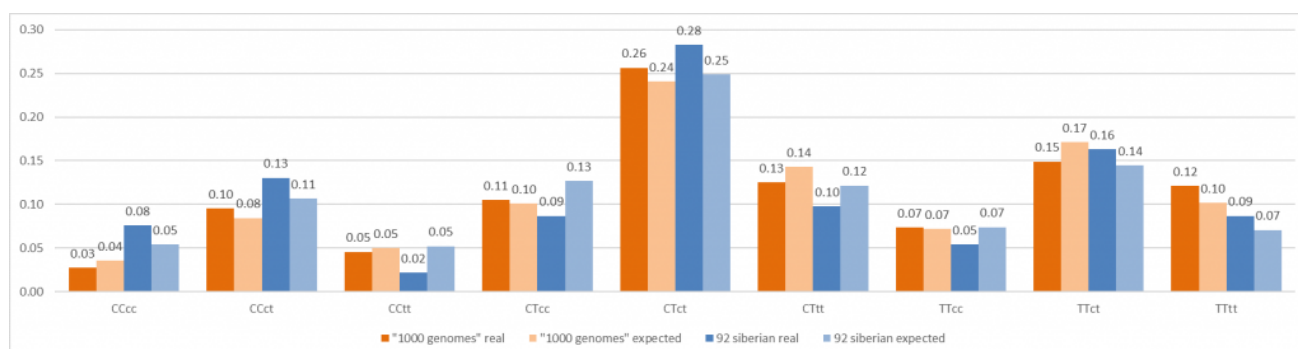


Рисунок 5

Сравнение частот различных генотипов по двум позициям SNP, полученных в настоящем исследовании с таковыми для 503 геномов европейской популяции из базы данных «1000 геномов»



Таким образом, с помощью методов FatI- и GluI-ПЦР анализа нами были определены генотипы 92-х доноров по позиции $chr1: 245618129$, которая находится в интроне гена KIF26B, кодирующего кинезин-подобный белок. Показано, что 43 донора (46,74%) имеют гетерозиготный набор С/Т в положении $chr1: 245618129$, 28 доноров (30,43%) гомозиготны по Т, а 21 донор (22,83%) – по С. Принимая во внимание, что клетки крови имеют диплоидный набор хромосом, замена С на Т встречается в 99 из 184 проанализированных вариантов (53,8%). При

этом Glal-ПЦР анализ показал, что цитозин в положении chr1: 245618129 в большей части молекул ДНК находится в метилированной форме (5-метилцитозин), однако доля неметилированного цитозина может достигать 1/5 и более.

Методы Fat1- и Glal-ПЦР анализа могут быть использованы в медико-генетических исследованиях для установления предрасположенности к тем или иным заболеваниям и определения частоты нежелательных аллелей в данной географической местности или в относительно изолированных популяциях. Использование Glal-ПЦР анализа позволяет также определить степень метилирования цитозина в случае, если он является одним из вариантов SNP. В настоящее время установлена связь между метилированием определенных участков генома и развитием онкологических и других патологий, что делает актуальным внедрение подобных эпигенетических методов диагностики в практику [11].

Литература

1. А.Г. Акишев, Н.А. Нетесова, М.А. Абдурашитов, Дегтярев (2021) *Определение полиморфизма 5mC/T в повторе AluSx (CHR16: 75033884) в препаратах ДНК из крови человека методами Glal- и Fat1-ПЦР анализа*. Эпигенетич ДНК диагност. том 2021(1), <https://epigendx.online/ru/2021/snp/>
2. Sherry ST, Ward MH, Kholodov M, Baker J, Phan L, Smigielski EM, Sirotkin K. *dbSNP: the NCBI database of genetic variation*. Nucleic Acids Res. 2001;29(1):308-311.
3. Tarasova, G.V., Nayakshina, T.N., Degtyarev, S.K. *Substrate specificity of new methyl-directed DNA endonuclease Glal*. BMC Molecular Biol 9, 7 (2008).
4. *CFX96 and CFX384 Real-Time PCR Detection Systems. Instruction Manual*. <https://www.biorad.com/webroot/web/pdf/lsr/literature/10010424.pdf> Accessed May 01 2021
5. Edwards JR, Yarychivska O, Boulard M, Bestor TH. *DNA methylation and DNA methyltransferases*. Epigenetics Chromatin. 2017;10:23.
6. Stark AE, Seneta E. *A reality check on Hardy-Weinberg*. Twin Res Hum Genet. 2013 16(4):782-789.
7. Sudmant, P., Rausch, T., Gardner, E. et al. *An integrated map of structural variation in 2,504 human genomes*. Nature 526, 75–81 (2015).
8. dbSNP Short Genetic Variations. Reference SNP (rs7516354) Report https://www.ncbi.nlm.nih.gov/snp/rs7516354#frequency_tab Accessed May 01 2021
9. dbSNP Short Genetic Variations. Reference SNP (rs66499664) Report https://www.ncbi.nlm.nih.gov/snp/rs66499664#frequency_tab Accessed May 01 2021
10. Grover VK, Cole DE, Hamilton DC. *Attributing Hardy-Weinberg disequilibrium to population stratification and genetic association in case-control studies*. Ann Hum Genet. 2010;74(1):77-87.
11. Chen Y, Hong T, Wang S, Mo J, Tian T, Zhou X. *Epigenetic modification of nucleic acids: from basic studies to medical applications*. Chem Soc Rev. 2017;46(10):2844-2872.