

ОБЗОРНЫЕ
И ЭКСПЕРИМЕНТАЛЬНЫЕ СТАТЬИ

УДК 575.113:577.21.08:599.9

NotI-STS КАК МАРКЕРЫ ГЕНОВ ХРОМОСОМЫ 3 ЧЕЛОВЕКА

© 2005 г. Г. Е. Сулимова^{1*}, Э. Р. Рахманалиев¹, Е. А. Климов¹, А. А. Компанийцев¹,
И. Г. Удина¹, Е. Р. Забаровский^{2,3}, Л. Л. Киселев²

¹Институт общей генетики им. Н.И. Вавилова Российской академии наук, Москва 119991

²Институт молекулярной биологии им. В.А. Энгельгардта Российской академии наук, Москва 119991

³Microbiology and Tumor Biology Center and Center for Genomics and Bioinformatics, Karolinska Institute,
Stockholm, Sweden

Поступила в редакцию 04.02.2005 г.

На основе нуклеотидных последовательностей ДНК из NotI-“прыжковых” и “связующих” клонотек, специфичных для хромосомы 3 человека, созданы 94 NotI-STS-маркера к 72 индивидуальным NotI-клонам. Порядок расположения NotI-STS-маркеров и их локализация установлены по совокупности данных по радиационному картированию (собственные данные), контиг-картированию, цитогенетическому картированию и картированию *in silico*. Сравнение ДНК NotI-STS с геномной ДНК человека выявило две бреши в областях 3p21.33 (маркер NL1-256) и 3p21.31 (NL3-005) и сегментную дубликацию. Идентичные фрагменты ДНК локализованы в областях 12q и 3p22–21.33 (маркер NL3-007). В области 3q28–q29 (маркер NLM-084) выявлен фрагмент, копии которого локализованы также на хромосомах 1, 2, 15 и 19. Для 69 NotI-STS показаны значимые гомологии с нуклеотидными последовательностями 70 генов и двух кДНК с учетом гомологий к 5'- и 3'-концевым нуклеотидным последовательностям NotI-STS. Ассоциацию NotI-STS с генами подтверждает высокая корреляция между плотностью генов и плотностью NotI-STS-маркеров на карте хромосомы 3 человека. Полученные данные позволяют рассматривать построенную нами NotI-карту как генную карту хромосомы 3 человека и свидетельствуют о перспективности применения NotI-STS в качестве маркеров генов.

Ключевые слова: хромосома 3 человека, NotI-клоны, ПЦР, радиационное (RH)-картирование, NotI-STS-маркеры, физическое картирование, гомологии с генами, геномика.

NotI-STS MARKERS FOR HUMAN CHROMOSOME 3 ARE GENE MARKERS, by G. E. Sulimova^{1*}, E. R. Rakhmanaliev¹, E. A. Klimov¹, A. A. Kompaniytsev¹, I. G. Udina¹, E. R. Zabarovsky^{2,3}, L. L. Kisselev² (¹Vavilov Institute of General Genetics, Russian Academy of Sciences, Moscow, 119991 Russia, *e-mail: galina_sulimova@mail.ru; ²Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, Moscow, 119991 Russia; ³Microbiology and Tumor Biology Center and Center for Genomics and Bioinformatics, Karolinska Institute, Stockholm, Sweden). Ninety four NotI-STS markers to seventy two individual NotI clones were developed basing on DNA nucleotide sequences from NotI-“jumping” and “linking” NotI-libraries of human chromosome 3. The localization of NotI-STS markers and their ordering on chromosome was established by combined data of RH-mapping (our data), contig-mapping, cytogenetic mapping and *in silico* mapping. Performed comparison of NotI-STS DNAs with human genome sequences revealed two gaps in the regions, 3p21.33 (marker NL1-256) and 3p21.31 (NL3-005), and segmental duplication. Identical DNA fragments are localized in the regions 12q and 3p22–21.33 (marker NL3-007). In the region 3q28–q29 (marker NLM-084) a fragment was detected with its identical copies present also on chromosomes 1, 2, 15 and 19. For 69 NotI-STSs, significant homologies with nucleotide sequences of 70 genes and two cDNAs were detected taking in consideration homologies to NotI-STS 5'- and 3'-terminal sequences. Association of NotI-STSs with genes is confirmed by high correlation of gene density distribution with the density of NotI-STS markers on the map of human chromosome 3. Obtained data evidence possibility of NotI-STS marker application as gene markers and allow considering constructed NotI-map as gene map of human chromosome 3.

Принятые сокращения: STS (Sequence Tagged Site) – ДНК-маркеры с известной нуклеотидной последовательностью и единичной локализацией в геноме; EST (Expressed Sequence Tag) – ДНК-маркеры, полученные на основе нуклеотидной последовательности кДНК; RH-гибриды – радиационные гибриды; RH-картирование – радиационное картирование; cR (centi-Ray) – единица расстояния на RH-картах, соответствующая 1% частоты разрывов между двумя маркерами после γ -облучения; cM (сантиморган) – единица расстояния на генетических картах, соответствующая 1% рекомбинации при мейозе; FISH (Fluorescence *In Situ* Hybridization) – флуоресцентная *in situ* гибридизация.

*Эл. почта: galina_sulimova@mail.ru

Нуклеотидная последовательность эухроматиновой части генома человека определена почти полностью (~99%) с ошибкой 1 н. на 10^5 п.н. [1]. Полное секвенирование генома человека, казалось бы, снимает необходимость продолжения многих прежних исследований, однако, создание новых ДНК-маркеров по-прежнему остается актуальной задачей, поскольку их широко используют в диагностике, изучении аллельного полиморфизма генов, в популяционных исследованиях, в сравнительной и функциональной геномике. ДНК-маркеры полезны и при сборке секвенированных фрагментов ДНК человека в непрерывную цепь, а также для независимой оценки достоверности такой реконструкции. Следует напомнить, что на январь 2005 г. в аннотированной последовательности генома человека существует 341 разрыв (брешь) [1]. Предполагают, что большая часть разрывов связана с протяженными (сегментными) дупликациями, составляющими около 5.3% эухроматинового генома [1]. (Сегментными дупликациями принято считать области протяженностью не менее 1 т.п.н. с идентичностью нуклеотидных последовательностей более 90%, которые не являются копиями транспозонных элементов [1].) Ассоциации сегментных дупликаций с разрывами в нуклеотидной последовательности хромосом подробно проанализированы в работе [2]. Использование ДНК-маркеров целесообразно для выявления таких сегментных дупликаций и их картирования.

Особое место среди STS-маркеров разных типов занимают маркеры, созданные на основе нуклеотидных последовательностей, прилегающих к участкам узнавания (GCGGCCGC) рестрикционной эндонуклеазы NotI (NotI-STS) и локализованных, главным образом, в CpG-островках [3–5], ассоциированных с генами [6]. На примере хромосомы 3 человека показано, что 97% из 60 исследованных NotI-STS гомологичны генам или кДНК [7]. Это позволяет рассматривать NotI-STS как универсальные маркеры большинства генов человека и, возможно, генов млекопитающих в целом. Следует отметить, что пока проведена аннотация генов 18 хромосом человека [1], в число которых не входит хромосома 3 (http://vega.sanger.ac.uk/Homo_sapiens/).

В настоящей работе подведен итог исследований по созданию NotI-STS-маркеров, определению их физической локализации и выявлению гомологий с генами и EST-человека. Приведенные данные стали результатом многолетних работ по созданию NotI-STS-карты хромосомы 3 человека. Эти исследования проводили поэтапно, и они включали создание NotI-STS-маркеров [7, 8], их физическое картирование с использованием FISH [9, 10], контиг-картирования [9, 11] и RH-картирования [7, 12] и, наконец, поиск гомологий с генами и EST человека [7, 9, 13].

УСЛОВИЯ ЭКСПЕРИМЕНТА

В работе использованы нуклеотидные последовательности NotI-клонов хромосомы 3 человека [9, 14, 15] и зарегистрированных в GenBank. Препарат ДНК хромосомы 3 человека получен согласно [14]. Олигонуклеотидные праймеры синтезированы в ЗАО “Синтол” (Россия). Подбор олигонуклеотидных праймеров осуществляли с использованием программ “Oligo”, “Primer” и “Oligo-Calculator”.

Локализация NotI-STS-маркеров методом RH-картирования. С целью определения положения NotI-STS-маркеров в геноме человека проводили скрининг панели RH-гибридов GeneBridge 4 RH-Panel (“Research Genetics, Inc.”, США) с помощью полимеразной цепной реакции (ПЦР) на термоциклерах PTC-100™ (“MJ Research Inc.”, США) и Amply-4L (“Биоком”, Россия). Условия амплификации оптимизировали для каждой пары праймеров по температуре отжига и другим параметрам на суммарной ДНК человека и китайского хомячка. Регистрационные номера нуклеотидных последовательностей NotI-клонов и NotI-STS-маркеров (GenBank) приведены в табл. 1. Специфичность полученных праймеров дополнительно тестировали на препарате ДНК хромосомы 3 человека. Тестирование продуктов амплификации проводили после электрофоретического разделения в агарозных или полиакриламидных гелях согласно стандартным методикам [16].

Обработку данных и получение информации о результатах скрининга RH-панели проводили с помощью программы RHMAPPER через сервер Whitehead Institute/MIT Center for Genome Research (<http://carbon.wi.mit.edu:8000/cgi-bin/contig/rhmapper.pl>), как подробно описано ранее [12]. При локализации NotI-STS-маркеров использовали значение LODscore = 15. Для установления мест локализации картируемых маркеров на физических картах разного типа и совмещения данных с интегрированными картами хромосомы 3 человека использовали серверы: <http://www.ncbi.nlm.nih.gov/>; <http://www.stanford.edu/RH/>; <http://www-genome.wi.mit.edu/>; <http://www.tigr.org/>.

Поиск гомологий NotI-STS-маркеров с нуклеотидными последовательностями, представленными в международных базах данных. Программу анализа нуклеотидных последовательностей BLAST 2.0 (<http://www.ncbi.nlm.nih.gov/BLAST/>) использовали для поиска гомологий исследуемых участков ДНК с нуклеотидными последовательностями, размещенными в базах данных, включая базы данных EMBL и EST. Гомологии с диспергированными повторами убирала с помощью программы Repeat Masker program. Поиск гомологий проводили для NotI-STS, ДНК исходных NotI-клонов, а также NotI-геномных последовательностей, включающих геномные ДНК длиной 1000 п.н.,

Таблица 1. NotI-STS-маркеры хромосомы 3 человека

Маркер	Регистрационный номер		Нуклеотидная последовательность праймеров (5'–3')	Mg ²⁺ /T, mM/°C	Время отжига, мин	Размер продукта, п.н.
	клон	STS				
924-021R	Z22241	924Z021R	F: tgcccctcgtgagatag R: aggggcatggtgactgtaag	3.0/58	0.5	152
924-025D	Z22242	924Z025D	F: ctccaacgcgacagaatc R: atgcacgtcagcgaaactg	3.0/56	0.5	111
924-053D	U49881	924Z053D	F: acaggattcggattcacgtc R: cttctctcgggctttgg	3.0/56	0.5	243
924-059	Z22257	–	F: cggccgtcacgggagagcc R: tgcgtggcggagcgtctggc	2.5/64	0.8	208
R924-059	Z22257	AJ310748	F: tgtaagaaggactgctcagc R: cgattatgggtctagtgtcc	2.0/56	0.5	195
R2924-059D	Z22256	AJ248343	F: cggccgtcacgggagagcc R: gcggtccacttccgcccagc	2.5/64	0.5	174
924-066R	Z22264	AJ310749	F: tcggtgccagagaacctac R: ctccgtagggtaggaggg	2.0/58	0.5	128
924-069D	Z22265	AJ311696	F: taagtgaagctgagctgagcc R: tgtcctgtcctctgcagg	3.0/58	0.5	215
924-082D	U49891	AJ248344	F: ctcggtgtccacggcacctc R: gcctccaagagcctgcgtcc	2.5/67	0.8	198
AP-1R	X87577	AJ248346	F: ctgaaagctctcccagagcc R: cttgctctgacccgggtg	2.5/62	0.5	223
AP-19S	Z22297	AJ248348	F: gggggcagcgctctcagac R: ggaccctgagtgccgggctc	2.5/67	0.8	188
AP-2D	Z22300	AJ248347	F: ctccaacgcgacagaatcg R: gccgtcgagagatgcacgtc	2.5/55	0.8	124
AP-2R	Z22301	AJ409273	F: tattctgtgccggtcttc R: tgatgggcagaagcagaag	3.0/56	0.5	155
AP-20RD	Z22298 Z22299	AJ248349	F: ctccaacgcgacagaatcg R: gccgtcgagagatgcacgtc	2.5/55	0.8	124
AP-20R	Z22299	–	F: tattctgtgccggtcttc R: tgatgggcagaagcagaag	3.0/56	0.5	155
AP-32D	Z22302	AJ248350	F: ctaccacggcctgtgtcaacg R: cgtgtcgggtgtgtcacaacc	2.5/63	0.8	250
AP-34R	Z94756	AJ409272	F: gaggagcgcaactctcacttg R: tgcctcattttagagaccgag	2.5/57	0.8	313
AP-4R	Z22307 Z22306	AP4R	F: ggtgactttgtcagggcc R: atgagataacccatgctcg	3.0/56	0.5	164
AP-40	Z22505	AJ311700	F: cgtagctttcgggcttcc R: tctgcacctagatggctgtg	2.5/58	0.5	280
J31-622D	AJ007299	AJ007299	F: agaagtgcacgtcgcgatct R: cagccgaaacattctacag	2.5/60	0.5	158
J32-032H	Z94761	AJ248386	F: gatggcgggtgaccactttacg R: cttccagacgtgggtccctcc	2.5/63	0.8	175
J32-135H	Z94763	AJ248388	F: ggcccttagccctgccacagc R: cgccagctcccggaaagcgct	2.5/68	0.8	126
J32-135H	Z94763	–	F: gcggttgggctactcgcgggc R: cgccagctcccggaaagcgct	2.5/66	0.8	173
J32-135L	X95829	AJ248387	F: ccttccgctcactcaggcaa R: ccacgggctgcaatgattgg	2.5/68	1.0	350
J32-159L	X95830	AJ248389	F: gcttctggcagctccctcgac R: caggcagtagctgtaccgtgc	2.5/62	0.8	135

Таблица 1. Продолжение

Маркер	Регистрационный номер		Нуклеотидная последовательность праймеров (5'–3')	Mg ²⁺ /T, mM/°C	Время отжига, мин	Размер продукта, п.н.
	клон	STS				
J32-180H	X95831	AJ248390	F: cctccgcagccccagcggca R: gctccccgaccgtcgccattg	2.5/69	0.8	117
J32-180H	X95831	–	F: cctccgcagccccagcggca R: accgtcgccattgcccgggc	2.5/67	0.8	138
J32-184R	AJ311551	AJ310744	F: ggctgtttaccggctgcat R: tcagtcacggctcccattggc	2.5/63	0.8	270
LC-32D	–	–	F: agcccagcgtcccaaatg R: gccgctgagccgtacaaaat	2.5/59	0.8	130
NL1-024R	X87494	AJ248355	F: ctacttcccgggcaagtgcg R: cgtgcccaccctgcagatcg	2.5/65	0.8	133
NL1-024R2	X87494	–	F: ctacttcccgggcaagtgcg R: tggcgtgtatttagcgcgtgc	2.5/65	0.8	150
NL1-065D	Z22324	–	F: gcagctcttacttgctgc R: agcgtctgtctaaaaggagg	2.0/56	0.5	233
NL1-106R	X87479	AJ308711	F: tactcagccgatgactcgg R: agagcaagaccaccacgcg	2.5/56	0.5	187
NL1-155D	Z22411	AJ409270	F: tctacgttccctgttctcgc R: cagatggggatggcagac	2.5/60	0.8	167
NL1-174R	Z22291	–	F: ccaacaagcctagcgagaac R: ccagtgagaagcagggactc	2.0/56	0.5	194
NL1-177R	Z94778	AJ308713	F: ggtatccgcacatcggcagctg R: caggcggcaggatttcggct	2.5/65	0.8	210
NL1-201R	Z22364	AJ311689	F: catacggcgttcccgaagtct R: aatgctttgtgcagcgttgc	2.5/60	0.8	270
NL1-210R	Z22308	AJ248356	F: gccacagtagcgaactgcaag R: cggggtacaagcagaccg	2.5/65	0.8	466
NL1-210R2	Z22308	–	F: gccacagtagcgaactgcaag R: ggaacaaaataggaggcgtg	2.5/57	0.8	102
NL1-211D	Z22342	–	F: agattccaacgtgttccc R: ttcaaaagtgcacaacagtgg	3.0/56	0.5	181
NL1-216D	Z22346	AJ308715	F: acctgtgtcttggtgacc R: ctgctagctgaatggggttc	2.0/58	0.5	156
NL1-229R	Z94780	AJ248357	F: gcgatcacctgaggtcagag R: gcaattctcggtcactgcaa	2.5/63	0.8	187
NL1-229R2	Z224780	AJ308716	F: ctaggctaggcagtcgccac R: ttacttccacgcagcaactcgg	2.5/59	0.8	149
NL1-232R	X87485	AJ248358	F: ggagtccggctcagacacctc R: ggaggtggactctgccgacac	2.5/65	0.8	261
NL1-243R	X87490	AJ248359	F: acgacacgcctgccactcctg R: tgcacttgaggagcctcggg	2.5/66	0.8	137
NL1-243R2	X87489	AJ308717	F: ccacacgacgcctgccac R: ccttctcccgcctatcc	2.5/65	0.8	113
NL1-245D	X87492	AJ409269	F: cagtcaggggaaggcgtgga R: cagttgcccggtttcaagc	2.5/65	0.7	115
NL1-256R	Z94781	AJ311697	F: gccgagaggatggaacagtgc R: atgaaccgaccgccagcaag	2.5/65	0.8	89
NL1-256S	Z94782	AJ248360	F: gggcattagatggaaccacc R: tgtgcatggcaacggttacc	2.5/61	0.5	126
NL1-308D	X87506	AJ248362	F: gccctgctactcccataatc R: caactgaatggccccaaaca	2.5/64	0.8	192

Таблица 1. Продолжение

Маркер	Регистрационный номер		Нуклеотидная последовательность праймеров (5'–3')	Mg ²⁺ /T, mM/°C	Время отжига, мин	Размер продукта, п.н.
	клон	STS				
NL1-308R	X87507	AJ248363	F: ccaagaacacagacagctgcgc R: agacataggaggaggagcgtg	2.5/58	0.8	100
NL1-313S	Z94784	AJ248364	F: ggcttagaccagctctccaa R: ccagcacagggtctgccaca	2.5/61	0.8	187
NL1-320R	X87508	AJ248365	F: cgagagattgccgccgaacgc R: cgagccagccaggagtcgaa	2.5/64	0.8	282
NL1-358R	X98703	AJ248366	F: ggaaccgcgcgacctggaac R: atccggctgagcccaggaga	2.5/69	0.8	172
NL1-401RD	AF061621 AF061620	AJ309132	F: aagaagcctgttagtgacgg R: cacaagctctgtaccactgg	2.5/59	0.5	600
NL2-007 D	Z22358	AJ248367	F: ggaacagaccggcaatcacg R: gtcagcgcgtcatcaacctcg	2.5/66	0.8	444
NL2-007R	Z22359	AJ308719	F: agagctggagttgacttagg R: ctacagctcctaggactaga	3.0/56	0.5	180
NL2-008RD	Z22392 Z22391	AJ248368	F: ctacggctcaacttgctgcg R: ggcgacaggcagaagctgag	2.5/64	0.8	381
NL2-008R	Z22391	AJ248369	F: atcggaggttgcggatgcc R: tcaactgcctgcgagcggc	2.5/63	0.8	188
NL2-008R2	Z22391	–	F: gagacgtgcagggccttc R: ggaggcgcagagtcttc	2.5/58	0.5	103
NL2-252D	Z22396	AJ248370	F: ccagccagccccgcaggatg R: tctcagagcactcgcgagcg	2.5/68	0.8	222
NL2-252R	Z22395	AJ409275	F: gatatcccagcctctgaggg R: cacctcgcgagcgtactgtg	2.5/60	0.8	118
NL2-278D	Z22414	–	F: ccgcaggattagttcagtc R: aggtggcttaatgcaggg	3.0/56	0.5	100
NL2-316D	Z22410	AJ308723	F: tctacgttccctgttctgcc R: cagacacgatgaacgcgtttc	2.5/60	0.8	154
NL3-001R	Z94790	AJ309128	F: agccaggcagggcagagcacg R: gctcgcgaaaggctggcctc	2.5/71	1.0	100
NL3-003U	X87527	AJ248371	F: aggaggctcccaccgggag R: gccaggaatcaggccccgag	2.5/62	0.5	323
NL3-004R	X87529	AJ309129	F: cggaaccagccggcgacatg R: gggcgacgcctcccgtctgc	2.5/70	0.7	165
NL3-005R	Z94792	AJ308724	F: tcaggccaagcatgccgatgg R: ggcactcagagccgcgtcgag	2.5/66	0.7	116
NL3-006R	Z94796	AJ308725	F: ggaagttgcggctcactg R: ttagttggaggatgagcaggg	2.5/60	0.7	113
NL3-007R	Z94798	AJ248372	F: ggctcgtcgggaatgattg R: ccttgccctcagctagtcc	2.5/64	0.8	122
NL3-007S	Z94799	AJ306567	F: ctggcgtaaatgtcaccgg R: tggcgctcaatctcgggaca	2.5/59	0.8	203
NL3-008R	Z95354	AJ306566	F: atcggaggttgcggatgcc R: tcaactgcctgcgagcggc	2.5/63	0.8	188
NL3-009D	Z95357	AJ308726	F: gggattgcggcacctcggaa' R: gaaccctttgtgaggccgg	2.5/66	0.7	170
NL3-019RD	AJ296027 AJ296028	–	F: gacagagcagaagtgagt R: ggatccttaaatgcataag	2.5/58	0.6	1182
NLJ-003R	Z22423	AJ248373	F: ttggtccagtctctcggcagg R: tctacacaaccaaggcctc	2.5/66	1.0	341

Таблица 1. Окончание

Маркер	Регистрационный номер		Нуклеотидная последовательность праймеров (5'-3')	Mg ²⁺ /T, mM/°C	Время отжига, мин	Размер продукта, п.н.
	клон	STS				
NLJ-003D	Z22424	AJ308727	F: gtgtagcggacggcagag R: ttccccttcaaccscaac	3.0/56	0.5	153
NLM-007S	Z22474	AJ248374	F: aaatgtgcaccggaaccgag R: ggtcatgtttgcgctgagca	2.5/59	0.5	397
NLM-007D	Z22473	AJ311699	F: tgcaccggaaccgagaaggc R: atattaccgaacgcgcgctgg	2.5/59	0.8	284
NLM-067R	X87545	AJ248375	F: caggccgcaaccctaactc R: cggggacaaaatctcaggcag	2.5/64	0.8	272
NLM-067D	X87544	AJ248376	F: ccgtgtgttggtggcttgca R: cactggccagccctccttctc	2.5/59	0.8	147
NLM-084R	Z95360	AJ248377	F: ggaacaccagcttgccgag R: ggagcccgagcggactgtg	2.5/65	0.7	237
NLM-187R	X87534	AJ248378	F: cccggtccctccataggaa R: gagagagcgcggctctgag	2.5/58	0.8	264
NLM-187R2	X87534	—	F: ttggatgactgaactcgca R: ggatgctgccctagagtgtg	2.5/55	0.8	190
NLM-216R	Z95368	AJ306697	F: tgagtccagatagccagagc R: ttgaaggcaatgctttagagc	2.5/58	1.0	266
NLM-223D	Z22452	AJ248379	F: ctgctcacagcgttaaaccc R: tggaaactgtgtgctctctc	2.5/67	0.8	235
NLM-246D	Z22453	AJ248381	F: cggcgtctctccgttacttc R: ctccgggaatcaacacagctc	2.5/59	0.8	273
NR1-BK20C	AJ326278	AJ248351	F: gggtcagagaccgagtcagag R: cgcagcccagcagttccagag	2.5/62	0.8	138
NRL-082D	Y10918	AJ248382	F: cgacatctgtctggcgcaga R: cccaaaccgggactgggca	2.5/63	0.8	169
NRL-082D2	Y10918	—	F: tggcgcagaaagttatggttc R: gctaccgccacctcacttcc	2.5/56	0.8	109
NRL-090R	X87562	AJ248383	F: cgacgcccgacctcgacac R: atggctctgtgcaacggagac	2.5/62	0.5	254
NRL-097D	X95832	AJ309130	F: accacgagtcggagtcgggg R: acatctgttgagttgacccgcg	2.5/64	0.7	94
NRL-143D	X95834	AJ309131	F: gccgccaccaccactcagc R: ggccccggaagcggaaagtctc	2.5/71	0.7	127
NRL-402R	AJ248384	AJ248384	F: ctactcgggagctgagacag R: ctctcatccctcatctga	2.5/57	0.8	235
NRL1-1D	X87570	AJ308731	F: ctccgcagaagtgtccgggg R: cgggctccaggacaggtaa	2.5/66	0.7	106

примыкающие к 5'- и 3'-концам NotI-участка исследуемого клона, как описано ранее [7].

В работе использовали компьютерные базы данных и Internet-ресурсы, размещенные на сервере NCBI (<http://www.ncbi.nlm.nih.gov/>): GenBank, UniGene, Human Genome Resource и LocusLink, а также базы данных: GDB (Genome Data Base (<http://www.gdb.org/>)), EMBL (<http://www.ebi.ac.uk/embl/>), CEPH-Genethon (<http://www.cephb.fr/bio/cephgenethon-map.html>), Whitehead Institute (<http://www-genome.wi.mit.edu>), RHDB (RHdb) Radiation Hybrid

DataBase (<http://corba.ebi.ac.uk/RHdb/>) и TIGR (<http://www.tigr.org/>).

РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

Создание STS к NotI-клонам хромосомы 3 человека и их локализация

На основе нуклеотидных последовательностей ДНК из NotI-прыжковых и связующих клоно-тек, специфичных для хромосомы 3 человека, созданы 94 NotI-STС-маркера к 72 индивидуальным

NotI-клонам (табл. 1). Праймеры подбирали к нуклеотидным последовательностям, прилегающим либо к 5'-, либо к 3'-концу NotI-участка (за исключением клонов AP-20 и NL1-401, для которых прямой праймер подбирали к 3', а обратный – к 5'-концевой нуклеотидной последовательности), что позволило бы избежать артефактов, если бы клонотека содержала “химерные” клоны. Условия амплификации оптимизировали по концентрации ионов магния, температуре и времени отжига (табл. 1).

Физическое картирование NotI-STS-маркеров проводили методом RH-картирования с использованием ПЦР-скрининга RH-панели GeneBridge 4. Метод основан на технике гибридизации соматических клеток. В отличие от классического метода, при получении RH-гибридов соматических клеток перед слиянием клетки донора облучают летальными дозами рентгеновских лучей, приводящими к множественным разрывам хромосом. При слиянии клеток фрагменты хромосом донора встраиваются случайным образом в хромосомы реципиента, одна гибридная клетка может содержать до 32% чужеродного генома. Стандартные коммерческие RH-панели содержат минимальное число клонов, в которых представлен весь геном человека с 6–7-кратным перекрытием (RH-панель Gene Bridge 4 содержит 91 клон). Для формирования RH-панели соматических клеток отбирается минимальное количество клонов, в которых представлены все фрагменты генома донора (например, человека) с достаточной степенью перекрытия. Затем с помощью ПЦР проводят скрининг RH-панели с использованием маркеров с уже известной локализацией, и на их основе формируется так называемая рабочая рамка маркеров. Локализация новых детектируемых маркеров устанавливается относительно маркеров рабочей рамки. Метод RH-картирования относится к методам физического картирования с высоким разрешением.

ДНК, выделенная из панели независимых RH-клонов, обеспечивает картирование новых STS-маркеров. Частоту вызванных радиацией разрывов между двумя маркерами используют как меру расстояния, и положение маркера определяют так же, как и при построении карт сцепления. Положение маркеров на карте хромосомы выражают в единицах расстояния 1 cR (centiRay), что соответствует 1% частоты разрывов между двумя маркерами после γ -облучения и зависит от дозы облучения. Для каждой хромосомы соотношение между величинами cR и рекомбинационной единицей расстояния cM варьирует. Например, для хромосомы 3 человека это соотношение в среднем составляет 4 cR/cM. Принцип метода и его достоинства подробно описаны ранее [12, 17]. В 2001 г. построены RH-карты генома человека с разрешением 0.5 млн.п.н. [18]. Эти карты совмещены с

физическими картами, полученными другими методами, а также с цитогенетическими картами, контиг- и STS-картами.

Скрининг RH-панели проводится с помощью ПЦР. Однако применение ПЦР для картирования NotI-STS связано с определенными трудностями, обусловленными обилием GC-пар, повторов и высокой консервативностью фрагментов ДНК, прилегающих к NotI-участкам. В ряде случаев это приводит к появлению положительных сигналов с ДНК китайского хомячка, культуру клеток которого использовали в качестве реципиента при создании RH-панели GeneBridge 4. В связи с этим, чтобы получить достоверные результаты для некоторых клонов, разработаны два или три варианта NotI-STS (табл. 1).

Результаты RH-картирования приведены в табл. 2. Картированные NotI-STS-маркеры зарегистрированы как RH-маркеры в базе данных RHdb, поддерживаемой EMBL (EBI, <http://corba.ebi.ac.uk/RHdb>); регистрационные номера картированных NotI-STS приведены в скобках в табл. 2. Для удобства сопоставления собственных данных с данными по локализации генов, представленными в базах данных, полученные нами результаты картирования NotI-STS на RH-карте Whitehead Institute (WI-RH-Map) приводили в соответствие с локализацией на RH-карте GeneMap'99-GB4 (GM99'-GB4) (табл. 2) на основе локализации маркеров рабочей рамки на обеих RH-картах. Физическое положение контигов определяли на основе RH-картирования одного или двух NotI-STS-маркеров, входящих в их состав. Положение остальных NotI-STS, входящих в контиги, рассчитывали, исходя из размера контигов и физического расстояния между фрагментами ДНК, составляющими контиги [11].

Цитогенетическая локализация NotI-STS-маркеров установлена ранее с использованием FISH [5, 9, 10]. Результаты локализации NotI-STS на цитогенетической карте, полученные с помощью FISH- и RH-картирования, как правило, хорошо совпадали.

Проведен поиск гомологий созданных NotI-STS с нуклеотидными последовательностями генома человека, представленными в базе данных <http://www.genome.gov/10000923>, и проверено соответствие локализации анализируемых STS с позицией в геноме соответствующих им контигов (табл. 2). Использование RH-картирования позволило при уточнении локализации NotI-STS-маркеров сопоставить результаты, полученные разными методами, и установить локализацию NotI-STS-маркеров по совокупности данных по RH-, контиг- [9, 11], цитогенетическому картированию [5, 10] и позиции геномных фрагментов, идентичных нуклеотидным последовательностям анализируемых NotI-STS (картирование *in silico*).

Таблица 2. Локализация NotI-STS-маркеров на разных хромосомных картах и гомологии с нуклеотидными последовательностями из базы данных генома человека

Маркер	Регистрационный номер		Гомолог (ген человека или EST)	Локализация на карте:				Позиция в геноме	
	STS (RHdb)	клон		WT-RH, сR	GM99'-GB4, сR	цитогенетическая	геномный контиг	позиция, млн.п.н.	
NL1-229	AJ308716 (RH135806)	Z94780	<i>LOC442074</i> (предполагаемый ген, сходный с геном фосфатазы 5 с двойной специфичностью)	49.71	46.28	3p25.3	NT_022517	9.635	
NL1-106	AJ308711	X87479	<i>IL17RC</i> (рецептор С интерлейкина 17)	53.72	46.90	3p25.3	NT_022517	9.967	
J32-159L	AJ248389	X95830	ДНК между генами <i>CREDLI</i> (цистеин-богатый белок с EGF-подобными доменами) и <i>FLJ33674</i> (предполагаемый белок FLJ33674)	54.40	49.50	3p25.3	NT_022517	9.980	
NL1-177	AJ308713	Z94778	<i>SH3BP5</i> (BTK-ассоциированный SH3-домен-связывающий белок 5)	72.95	67.42	3p25.1	NT_022517	15.366	
AP-1	AJ248346 (RH135807)	X87577	<i>EAF1</i> (ELL-ассоциированный фактор 1)	73.50	67.84	3p25.1	NT_022517	15.442	
NRL-082	AJ308730	Y10918	<i>SIMP</i> (иммунодоминантный пептид, ассоциированный с МНС)	117.99	112.06	3p22.3	NT_022517	31.545	
NL1-024	AJ311656	X87494 X87493	<i>OSBPL10</i> (ген, подобный гену оксистерол-связывающего белка 10)	130.00	123.50 (Контиг 1)	3p22.3	NT_022517	31.869	
AP-40	AJ311700	Z22505	<i>LOC389108</i> (предполагаемый белок LOC389108)	135.29	126.60 (Контиг 1)	3p22.1	NT_022517	36.799	
NL1-401	AJ309132	AF061621	<i>ITGA9</i> (интегрин $\alpha 9$)	135.96	130.34 (Контиг 1)	3p22.1	NT_022517	37.454	
NRL-402	AJ248384	AJ248384	<i>ITGA9</i> (интегрин $\alpha 9$)	118.00	112.00 (Контиг 1)	3p22.1	NT_022517	37.791	
NLJ-003	AJ308727	Z22424	<i>CTDSPL</i> (полипептид А С-концевого домена РНК-полимеразы II)	136.62	130.03 (Контиг 1)	3p22.1	NT_022517	37.863	
NL3-019	AJ296027 AJ296028	AJ296027 AJ296028	кДНК, 3'-концевой клон (AW614138)	147.38	133.96 (Контиг 1)	3p22.1	NT_022517	38.284	
AP-20	AJ248349 (RH135808)	Z22299	<i>SCN5A</i> (α -пептид натриевого канала, тип V)	146.26	133.27 (Контиг 1)	3p21.33	NT_022517	38.504	
NL3-003	AJ248371	X87526 X87527	<i>GORASP1</i> (стэкинг-белок 1 комплекса Гольджи)	135.78	130.34 (Контиг 1)	3p21.33	NT_022517	38.961	
NL1-308	AJ248363	X87506 X87507	<i>MOBP</i> (основной миелин-ассоциированный белок олигодендроцитов)	146.50	135.41 (Контиг 1)	3p21.33	NT_022517	39.356	
NL1-256	AJ311697 (RH135814)	Z94781 Z94782	(-)	166.7	142.5	3p22-21.33	(-)	(-)	
NL1-232	AJ248358 (RH135809)	X87485	<i>ABHD5</i> (домен-содержащая α/β -гидролаза 5)	167.50	143.17 (Контиг 3)	3p21.32	NT_022517	43.551	
NL3-009	AJ308726	Z95357	<i>SNRK</i> (белок, родственный SNF-1 киназе)	170.68	143.71 (Контиг 3)	3p21.32	NT_022517	43.289	

Таблица 2. Продолжение

Маркер	Регистрационный номер		Гомолог (ген человека или EST)	Локализация на карте:				Позиция в геноме	
	STS (RHdb)	клон		WI-RH, cR	GM99-GB4, cR	цитогенетическая	геномный контиг	позиция, млн.п.н.	
NL2-007	AJ248367	Z22358 Z22359	<i>EXOSC7</i> (компонент 7 экзосомы) <i>ZDHHC3</i> (белок 3, содержащий ДННС-домен цинковых пальцев)	177.64	147.53	3p21.32	NT_022517	44.978	
NL1-320	AJ248365	X87509	<i>SACMIL</i> (белок, подобный белку <i>SAC1</i> – супрессору мутаций актина у дрожжей)	181.95	152.53	3p21.32	NT_022517	45.692	
NL3-007	AJ248372	Z94798 Z94799	<i>LOC51054</i> (предполагаемый белок переноса гликолипидов)	182.3	148.2 (Контиг 2)	3p21.31	NT_029419 (12q)	(-)	
NLM-067	AJ248375 (RH 125635)	X87545	<i>TMEM16F</i> (трансмембранный белок 16F)	174.70	150.90 (Контиг 2)	3p21.31	NT_022517	46.706	
NL3-006	AJ308725	Z94796	<i>KIAA0540</i> (белок KIAA0540)	180.12	150.01 (Контиг 2)	3p21.31	NT_022517	47.010	
J32-135L	AJ248387	X95829	<i>ARNA</i> (белок – гомолог А из семейства RAS)	189.00	158.10	3p21.31	NT_022517	49.408	
J32-135H	AJ248388 (RH 135827)	Z94763	<i>hmt31467</i> (ген человека, предсказанный <i>in silico</i> программой Gnomon)	185.62	158.15	3p21.31	NT_022517	49.536	
NL3-005	AJ308724	Z94792	(-)	(-)	159.0 (Контиг 4)	3p21.32- p21.2	(-)	(-)	
NRL1-1	AJ308731	X87570 X87571	<i>SEMA3F</i> (семафорин 3F)	217.96	161.61 (Контиг 4)	3p21.31	NT_022517	50.047	
NL3-001	AJ242846	AJ242846 Z94790	<i>GNAI2</i> (нуклеотид-связывающий белок, G- α -2-ингибирующий полипептид)	209.0	160.47 (Контиг 4)	3p21.31	NT_022517	50.151	
NL1-210	AJ409276	Z22308 Z22309	<i>MARKAPK3</i> (митоген-активированная протеинкиназа – активированная протеинкиназа 3)	212.73	162.09 (Контиг 4)	3p21.31	NT_022517	50.613	
NL1-216	AJ308715 (RH 135813)	AJ242843	ДНК между генами <i>ARMET</i> (аргинин-богатый белок, мутирующий на ранних стадиях опухолей) и <i>RBM15B</i> (белок 15B, связывающий мотив РНК) <i>ARMET</i>	217.12	160.99 (Контиг 5)	3p21.31	NT_022517	51.279	
NRL-097	AJ309130	Z22346 X95832	<i>ARMET</i> (аргинин-богатый белок, мутирующий на ранних стадиях развития опухолей)	216.00	160.84 (Контиг 5)	3p21.31	NT_022517	51.381	
NL1-243	AJ248359 (RH 135818)	X87489 X87490	<i>DKFZF564O243</i> (ген предполагаемого белка <i>DKFZF564O243</i>) <i>ACY1</i> (аминоацилаза 1)	218.00	165.50 (Контиг 5)	3p21.31	NT_022517	51.873	
NL1-245	AJ409269	X87492 X87491	<i>ACY1</i> (аминоацилаза 1)	216.41	165.38 (Контиг 5)	3p21.31	NT_022517	51.975	
NL3-008	AJ306566	Z95354 Z95355	ДНК между <i>PRO2730</i> (ген предполагаемого белка <i>PRO2730</i>) и <i>GLYCTK</i> (ген белка, сходного с белком дрозифиллы CG9886)	222.7	171.7 (Контиг 6)	3p21.2	NT_022517	52.229	

Таблица 2. Продолжение

Маркер	Регистрационный номер		Гомолог (ген человека или EST)	Локализация на карте:			Позиция в геноме	
	STS (RHdb)	клон		WI-RH, cR	GM99'-GB4, cR	цитогенетическая	геномный контиг	позиция, млн.п.н.
NL3-004	AJ309129	X87529	<i>PP2C-ETA</i> (протеинфосфатаза 2С, eta-изоформа)	214.21	162.76 (Контиг 6)	3p21.2	NT_022517	52.238
NL2-008	AJ248369 (RH 125637)	Z22392	<i>hmm14181</i> (ген человека, предсказанный <i>in silico</i> программой Spomom)	214.23	162.84 (Контиг 6)	3p21.2	NT_022517	52.271
AP-32	AJ248350 (RH 135811)	Z22302 Z22303	<i>PHF7</i> (белок 7, содержащий пальчиковый PHD-домен)	213.40	161.89 (Контиг 6)	3p21.2	NT_022517	52.300
J32-032H	AJ248386 (RH 135810)	Z94761	<i>NS</i> (нуклеостемин)	214.28	163.99	3p21.2	NT_022517	52.677
NRL-143	AJ309131	X95834 X95835	<i>SPC12</i> (сигнальная пептидаза 12 кДа)	215.40	164.37	3p21.31	NT_022517	52.697
AP-2	AJ409273 (RH 126916)	Z22301 Z22300	<i>AD-017</i> (гликозилтрансфераза)	233.00	198.02	3p21.2	NT_022517	57.437
924-025	AJ310678	Z22242	<i>ARF4</i> (ADP-рибозилирующий фактор 4)	232.33	198.02	3p21.2	NT_022517	57.540
NLM-223	—	Z22452 Z22451	<i>ABHD6</i> (домен-содержащая α/β -гидролаза 6)	229.47	193.81	3p21.2	NT_022517	58.078
J32-184R	AJ310744	AJ311551	<i>NIF3L1BP1</i> (белок 1, сходный с Ngg1-интерактивным фактором 3)	219.34	191.50	3p21.1	NT_022517	63.806
J32-180H	AJ248390	X95831	<i>SCA7</i> (ген спинозомжечковой атаксии 7)	238.03	188.12	3p21.1	NT_022517	63.855
NL1-358	AJ248366	X98703	<i>SCA7</i> (ген спинозомжечковой атаксии 7)	235.80	184.10	3p21.1	NT_022517	63.812
NL1-211	—	Z22342	<i>BAIAP1</i> (BAI1-ассоциированный белок 1)	(—)	(—)	3p14.1	NT_086638	65.999
924-082	AJ223142	U49891	<i>PDZRN3</i> (PDZ-домен, содержащий RING- ² -палец ³)	255.33	213.79	3p14.1	NT_022459	73.482
924-021	AJ310677	Z22241	<i>LOC285296</i> (гомолог B1, сходный с unc-93 <i>C. elegans</i>)	271.08	225.16	3p14.1	NT_022459	75.590
NL1-065	AJ308710	Z22324	<i>FLJ10134</i> (предполагаемый белок FLJ10134)	423.20	357.15	3q12.3	NT_005612	101.532
NL1-174	—	Z22291	<i>TFG</i> (химерный онкоген, слитый с тирозинкиназой TRK)	421.99	352.18	3q12.3	NT_005612	101.749
NLM-187	AJ248378 (RH 126917)	X87534	<i>BOC</i> (клеточный поверхностный рецептор Ig, родственный CDO(BOC))	500.78	406.86	3q13.2	NT_005612	114.251
J31-622D	AJ007299	—	<i>hmm31641</i> (ген человека, предсказанный <i>in silico</i> программой Spomom)	551.82	454.70	3q21.3	NT_005612	127.719
924-059	AJ310748 AJ248343	Z22257	<i>ABTBI</i> (фактор 1, содержащий анкириновый повтор и VTB(POZ)-домен)	541.14	445.50	3q21.3	NT_005612	128.713
AP-4	AJ409271	Z22307 Z22306	<i>hmm31654</i> (ген человека, предсказанный <i>in silico</i> программой Spomom)	551.60	452.72	3q21.3	NT_005612	129.648
NLM-007	AJ248374 AJ311699	Z22474 Z22473	<i>RAB7</i> (ген из RAS-семейства онкогенов)	558.10	447.92	3q21.3	NT_005612	129.765

Таблица 2. Окончание

Маркер	Регистрационный номер		Гомолог (ген человека или EST)	Локализация на карте:			Позиция в геноме	
	STS (RHdb)	клон		WI-RH, cR	GM99'-GB4, cR	цитогенетическая	геномный контиг	позиция, млн.п.н.
NR1- BK20C	AJ326278	AJ326278	<i>H1FX</i> (белок X из семейства гистонов H1)	552.09	460.16	3q21.3	NT_005612	130.315
LC-32	AJ248351 (RH 135830)	-	<i>H1FX</i> (белок X из семейства гистонов H1)	552.10	460.26	3q21.3	NT_005612	130.355
NL2-278	-	Z22414	<i>MGC3040</i> (предполагаемый белок MGC3040)	593.08	469.96	3q22.3	NT_005612	134.078
NRL-090	AJ248383 (RH 135815)	X87562	<i>GMPS</i> (GMP-синтаза)	703.68	571.81	3q25.31	NT_005612	156.870
NL1-155	AJ409270 (RH 135829)	Z22411	<i>SSR3</i> (рецепторная сигнальная последовательность, транслокон-ассоциированный белок γ)	710.84	576.11	3q25.31	NT_005612	157.593
NL2-316	AJ308723 (RH 135816)	Z22410	<i>SSR3</i> (рецепторная сигнальная последовательность, транслокон-ассоциированный белок γ)	710.85	576.09	3q25.31	NT_005612	157.593
924-053	-	U49882	<i>hmm31848</i> (ген человека, предсказанный <i>in silico</i> программой Gnomon)	(-)	(-)	3q26.33	NT_005612	183.883
NLM-246	AJ248381 (RH 125636)	Z22453 Z22454	<i>ATP11B</i> (АТРаза, класс VI, тип 11B)	833.90	682.99	3q27.3	NT_005612	183.914
924-066	AJ310749	Z22264	<i>LAMP3</i> (ассоциированный с лизосомами мембранный белок 3)	837.80	684.06	3q27.3	NT_005612	184.282
NL1-201	AJ311689 (RH 135819)	Z22364	<i>ABCC5</i> (белок 5 из подсемейства С (CFTR/MRP) АТР-связывающей кассеты)	844.42	689.05	3q27.3	NT_005612	185.056
924-069	AJ311696	Z22265	<i>CLCN2</i> (хлоридный канал 2)	852.83	696.74	3q27.3	NT_005612	185.481
AP-19	AJ248348 (RH 135812)	Z22297 Z22296	<i>ETV5</i> (ets-родственный транскрипционный фактор 5)	853.30	697.30	3q28	NT_005612	187.147
NL1-313	-	Z94783 Z94784	<i>LPP</i> (lim-домен-содержащий белок, вовлеченный в транслокацию t при лимомах)	854.28	697.57	3q28	NT_005612	189.192
NLM-084	AJ248377 (RH 135828)	Z95360 Z95361	<i>LOC440897</i> (белок, сходный с фосфоглюкоматазой) <i>FLJ25222</i> (CXYorf1-родственный белок)	874.9	719.9	(-)	NT_077402 (1p36) NT_022135 (2q14) NT_035325 (15q26) NT_011255 (19p13) NT_090211 (-)	(-)
AP-34	AJ409272 (RH 135826)	Z94756	кДНК, 3'-концевой клон (AW631201)	880.40	730.00	3q29	NT_005535	195.890
NL2-252	AJ409275 (RH 135817)	Z22395	<i>TFRC</i> (рецептор трансферрина (p90, CD71))	880.53	726.38	3q29	NT_029928	197.135
NLM-216	AJ306697	Z95368	<i>FLJ20522</i> (предполагаемый белок FLJ20522)	884.57	730.04	3q29	NT_029928	197.765

Примечание. (-) – отсутствие данных.

Порядок расположения 72 NotI-STS-маркеров на хромосоме 3 человека представлен на рисунке.

Не выявлено гомологий двух NotI-STS (NL1-256 и NL3-005) с геномными ДНК хромосомы 3 человека. ДНК этих маркеров вообще не представлены в геномных контигах, хотя они входят в состав контига, перекрывающего область 3p21–p22 [9, 11], и в международных базах данных зарегистрированы как NotI-клоны, так и NotI-STS, созданные на их основе (табл. 2). Маркеры NL3-007 и NLM-084, согласно данным геномных контигов, локализованы на других хромосомах. Отмечена даже множественная локализация маркера NLM-084 (табл. 2), но не на хромосоме 3. При этом локализация этого маркера на хромосоме 3 доказана нами методом RH-картирования [7]. Клоны NL1-256, NL3-007 и NL3-005 входят в состав контига, перекрывающего область 3p21–p22 [10]. Локализация клона NL1-256 на хромосоме 3 подтверждена также с помощью RH-картирования [7].

Отсутствие фрагментов ДНК, идентичных маркерам NL1-256 и NL3-005, в аннотированной в январе 2005 г. нуклеотидной последовательности генома человека указывает на присутствие брешей в области локализации этих маркеров и позволяет установить точное расположение двух брешей на хромосоме 3. Более того, нуклеотидные последовательности NotI-STS-маркеров и ДНК контига 3p21–p22 [9, 11], которому они принадлежат, могут быть использованы для заполнения брешей и реконструкции непрерывной геномной последовательности хромосомы 3 человека.

Отсутствие на хромосоме 3 человека фрагментов ДНК, идентичных маркерам NL3-007 и NLM-084, по-видимому, объясняется другой причиной. Согласно геномным контигам в базе данных эти маркеры локализованы на других хромосомах, однако их принадлежность хромосоме 3 доказана несколькими независимыми методами [7, 9–11], что указывает на существование сегментных дупликаций в областях локализации маркеров NL3-007 и NLM-084. Фрагменты ДНК, идентичные маркеру NL3-007, присутствуют в геноме на двух хромосомах (3p22–21.33 и 12q). Маркер NLM-084, локализованный нами в области 3q28–q29, имеет множественную локализацию в геноме. В базе данных “Human Genome Resource” фрагменты ДНК, идентичные маркеру NLM-084, представлены в нескольких геномных контигах: NT_077402 (1p36.33), NT_022135 (2q14.1), NT_035235 (15q26.3), NT_011255 (19p13.3), а также в контиге NT_090211, не приписанном ни к одной из хромосом человека. Поскольку полная реконструкция нуклеотидной последовательности хромосомы 3 человека и аннотация генов на этой хромосоме еще не завершены [1], нельзя исключать, что этот контиг локализован в области 3q29.

Таким образом, сравнение ДНК NotI-STS с геномными последовательностями человека выявило две брешы в областях 3p21.33 (маркер NL1-256) и 3p21.31 (NL3-005) и сегментную дупликацию с локализацией идентичных фрагментов ДНК в областях 12q и 3p22–21.33 (маркер NL3-007). В области 3q28–q29 (маркер NLM-084) выявлен фрагмент, идентичные копии которого локализованы также на хромосомах 1, 2, 15 и 19. Полученные данные свидетельствуют о том, что NotI-STS-маркеры могут быть с успехом использованы для выявления брешей и сегментных дупликаций в геномной последовательности ДНК человека.

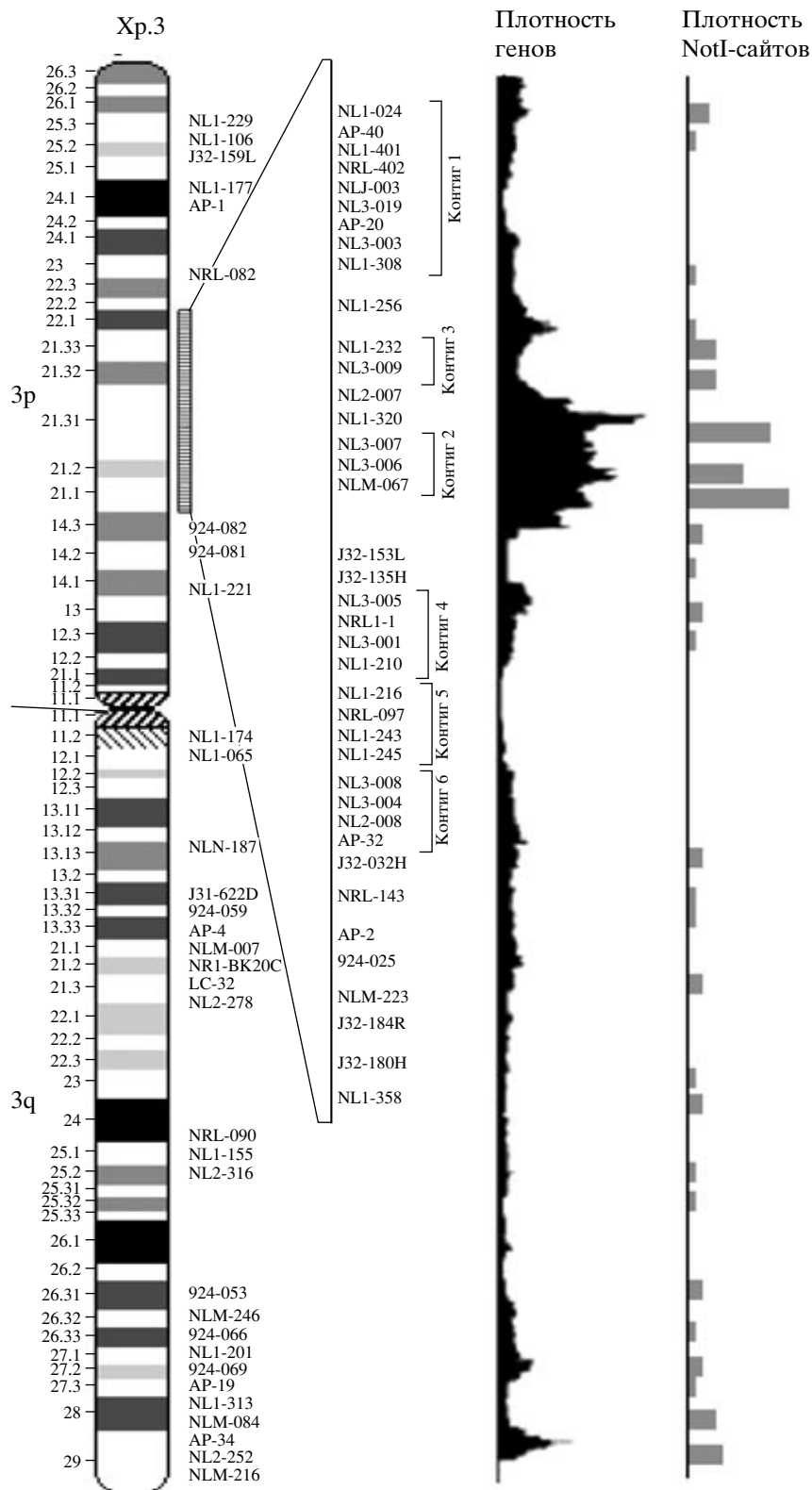
Поиск гомологий NotI-STS-маркеров с генами и EST человека

Поиск гомологий NotI-STS-маркеров с генами и EST человека проводили на основе нуклеотидных последовательностей, представленных в базах данных GenBank, EMBL, TIGR. Схема поиска гомологий описана ранее [7] и в разделе “Условия эксперимента”. Результаты поиска суммированы в табл. 2, в которую включены гомологичные нуклеотидные последовательности 5'- и 3'-участков ДНК каждого индивидуального NotI-клона. Данные по гомологичным участкам ДНК с уровнем идентичности менее 90% не рассматривали.

Выявлены значимые гомологии 70 NotI-STS (более 97% исследованных маркеров) с генами или EST. Участки NotI-связующих клонов, прилегающие к 5'- и 3'-концам NotI-участка, в ряде случаев гомологичны двум разным генам (NL3-003, NL3-007, NL1-243, NRL-143) или участку ДНК, расположенному между двумя генами (J32-159L, NL1-216). В семи случаях обнаружена гомология двух NotI-STS к разным участкам одного гена: NL1-401 и NRL-402 гомологичны гену *ITGA9*, NL1-216 и NRL-097 – гену *ARMET*, NL1-243 и NL1-245 – гену *ACY1*, AP-2 и 924-025 – гену *ARF4*, J32-180H и NL1-358 – гену *SCA7*, NR1-BK20C и LC-32 – гену *HIFX*, NL1-155 и NL2-316 – гену *SSR3* (табл. 2).

В целом, выявлены гомологии созданных нами NotI-STS к 70 генам и двум EST с учетом гомологий к обеим ДНК (5'- и 3'-концам) NotI-связующих клонов. Проведен сравнительный анализ характера распределения генов (по данным [19]) и NotI-участков на хромосоме 3 человека (рисунок). Диаграммы распределения NotI-участков и распределения (плотности) генов на хромосоме 3 человека практически совпадают. Это еще раз подтверждает ассоциацию NotI-участков с генами и позволяет рассматривать построенную нами NotI-карту (рисунок) как генную карту хромосомы 3 человека (табл. 2).

Для более широкого представления о роли генов, обладающих гомологией с локализованными NotI-STS, эти гены классифицировали по функци-



Расположение NotI-STS-маркеров, плотность генов и NotI-участков на хромосоме 3 человека. Порядок NotI-STS-маркеров устанавливали по совокупности данных по RH-картированию, FISH, контиг-картированию и картированию *in silico*. Карта плотности генов построена в соответствии с данными [19], карта плотности NotI-STS-маркеров построена на основе собственных данных.

ям кодируемых ими белков и сопоставили полученные результаты с данными для всего генома человека. Показана высокая корреляция наших данных и данных Вентера и соавт. [20]. Помимо большого количества генов с неизвестной функцией (45.5% – наши данные и 41.7% – данные [20]) выделяются группы генов, кодирующих ферменты (11.4 и 9.2%) и ДНК-связывающие белки (13.2 и 13.5%), к которым можно отнести и факторы транскрипции. Распределение генов разных семейств на хромосоме 3 соответствует, в основном, распределению генов в геноме человека в целом. Однако есть и особенности, характерные для хромосомы 3: высокое содержание протоонкогенов и онкогенов (11.4%), локализованных в основном в области 3p21–p22, по сравнению с геномом в целом (2.9%). Насыщенность хромосомы 3 человека онкогенами [21–23] придает высокую значимость исследованиям по созданию и характеристике маркеров генов, локализованных на этой хромосоме.

Как показано в нашей работе, NotI-STS-маркеры можно рассматривать как универсальные маркеры генов. Они не специфичны по отношению к отдельным группам генов и могут применяться для маркирования генов из разных семейств. Высокий консерватизм нуклеотидных последовательностей, прилегающих к NotI-участкам, позволяет использовать их и для поиска генов у других организмов, вплоть до таких отдаленных групп, как микроорганизмы [24]. NotI-STS-маркеры могут быть с успехом использованы не только в структурной, но и в функциональной геномике. Например, показана эффективность NotI-микрорпанелей для выявления в геноме человека нуклеотидных последовательностей, подвергшихся метилированию (или деметилированию) при различных онкологических заболеваниях [25]. Этот подход, на наш взгляд, может существенно облегчить поиск и локализацию новых онкогенов и генов-супрессоров и может найти широкое применение в будущем.

Работа выполнена при поддержке грантов INTAS (1306-2003), Министерства образования и науки Российской Федерации (программа “Поддержка научных школ”, НШ-827.2003.4 и НШ-2093.2003.4), а также Российского фонда фундаментальных исследований (04-04-08154а).

СПИСОК ЛИТЕРАТУРЫ

1. International Human Genome Sequencing Consortium. 2004. Finishing the euchromatic sequence of the human genome. *Nature*. **431**, 931–945.
2. Eichler E.E., Clark R.A., She X. 2004. An assessment of the sequence gaps: unfinished business in a finished human genome. *Nature Rev. Genet.* **5**, 345–354.
3. Larsen F., Gundersen G., Lopez R., Prydz H. 1992. CpG islands as gene markers in the human genome. *Genomics*. **13**, 1095–1107.
4. Allikmets R.L., Kashuba V.I., Pettersson B., Gizatullin R., Lebedeva T., Kholodnyuk I.D., Bannikov V.M., Petrov N., Zakharyev V.M., Winberg G. 1994. NotI linking clones as a tool for joining physical and genetic mapping of the human genome. *Genomics*. **19**, 303–309.
5. Protopopov A.I., Gizatullin R.Z., Vorobieva N.V., Protopopova M.V., Kiss C., Kashuba V.I., Klein G., Kisselev L.L., Grafodatsky A.S., Zabarovsky E.R. 1996. High resolution FISH mapping of 50 NotI linking clones homologous to genes and cDNAs on human chromosome 3. *Chromosome Res.* **4**, 443–447.
6. Bird A.P. 1987. CpG islands as gene markers in the vertebrate nucleus. *Trends Genet.* **3**, 342–347.
7. Sulimova G.E., Kutsenko A.S., Rakhmanaliev E.R., Udina I.G., Kompaniytsev A.A., Protopopov A.I., Mojsjak E.V., Klimov E.A., Muravenko O.V., Zelenin A.V., Braga E.A., Kashuba V.I., Zabarovsky E.R., Kisselev L.L. 2002. Human chromosome 3: integration of 60 NotI clones into a physical and gene map. *Cytogenet. Genome Res.* **98**, 177–183.
8. Сулимова Г.Е., Удина И.Г., Кунижева С.С., Компанийцев А.А. 1999. Создание NotI-STS-маркеров хромосомы 3 человека. *Молекуляр. биология*. **33**, 791–796.
9. Kashuba V.I., Gizatullin R.Z., Protopopov A.I., Li J., Vorobieva N.V., Fedorova L., Zabarovska V.I., Muravenko O.V., Kost-Alimova M., Domninsky D.A., Kiss C., Allikmets R., Zakharyev V.M., Braga E.A., Sumegi J., Lerman M., Wahlestedt C., Zelenin A.V., Sheer D., Winberg G., Grafodatsky A., Kisselev L.L., Klein G., Zabarovsky E.R. 1999. Analysis of NotI linking clones isolated from human chromosome 3 specific libraries. *Gene*. **239**, 259–271.
10. Protopopov A., Kashuba V., Zabarovska V.I., Muravenko O.V., Lerman M.I., Klein G., Zabarovsky E.R. 2003. An integrated physical and gene map of the 3.5-Mb chromosome 3p21.3 (AP20) region implicated in major human epithelial malignancies. *Cancer Res.* **63**, 404–412.
11. Kashuba V.I., Szeles A., Allikmets R., Nilsson A.S., Bergerheim U.S., Modi W., Grafodatsky A., Dean M., Stanbridge E.J., Winberg G. 1995. A contig of NotI jumping and linking clones cover 2.5 MB in 3p21–p22 region suspected to contain tumor suppressor gene. *Cancer. Genet. Cytogenet.* **81**, 144–150.
12. Сулимова Г.Е., Компанийцев А.А., Мойсяк Е.В., Рахманалиев Э.Р., Климов Е.А., Удина И.Г., Захаров И.А. 2000. Радиационное картирование как один из основных методов построения физических карт геномов человека и животных с высоким уровнем разрешения. *Радиационная биология и радиоэкология*. **40**, 520–528.
13. Kutsenko A.S., Gizatullin R.Z., Al-Amin A.N., Wang F., Kvasha S.M., Podowski R.M., Matushkin Y.G., Gyanchandani A., Muravenko O.V., Levitsky V.G., Kolchanov N.A., Protopopov A.I., Kashuba V.I., Kisselev L.L., Wasserman W., Wahlestedt C., Zabarovsky E.R. 2002. NotI flanking sequences: a tool for gene discovery and verification of the human genome. *Nucleic Acids Res.* **30**, 3163–3170.

14. Zabarovsky E.R., Allikmets R., Kholodnyuk I., Zabarovska V.I., Paulsson N., Bannikov V.M., Kashuba V.I., Dean M., Kisselev L.L., Klein G. 1994. Construction of representative NotI linking libraries specific for the total human genome and for human chromosome 3. *Genomics*. **20**, 312–316.
15. Zabarovsky E.R., Kashuba V.I., Pettersson B., Petrov N., Zakharyev V., Gizatullin R., Lebedeva T., Bannikov V., Pokrovskaya E.S., Zabarovska V.I. 1994. Shot-gun sequencing strategy for long-range genome mapping: a pilot study. *Genomics*. **21**, 495–500.
16. Sambrook J., Fritsch E.F., Maniatis T. 1989. *Molecular Cloning: A Laboratory Manual*. N.Y.: Cold Spring Harbour Laboratory Press.
17. Рахманалиев Э.Р., Климов Е.А., Сулимова Г.Е. 2004. Методы картирования геномов млекопитающих. *Успехи современной биологии*. **124**, 286–302.
18. Olivier M., Aggarwal A., Allen J., Almendras A.A., Bajorek E.S., Beasley E.M., Brady S.D., Bushard J.M., Bustos V.I., Chu A., Chung T.R., De Witte A., Denys M.E., Dominguez R., Fang N.Y., Foster B.D., Freudenberg R.W., Hadley D., Hamilton L.R., Jeffrey T.J., Kelly L., Lazzeroni L., Levy M.R., Lewis S.C., Liu X., Lopez F.J., Louie B., Marquis J.P., Martinez R.A., Matsuura M.K., Misherghi N.S., Norton J.A., Olshen A., Perkins S.M., Perou A.J., Piercy C., Piercy M., Qin F., Reif T., Shepard K., Shokoohi V., Smick G.A., Sun W.L., Stewart E.A., Fernando J., Tejada Tran N.M., Trejo T., Vo N.T., Yan S.C., Zierten D.L., Zhao S., Sachidanandam R., Trask B.J., Myers R.M., Cox D.R. 2001. A high-resolution radiation hybrid map of the human genome draft sequence. *Science*. **291**, 1298–1302.
19. Caron H., van Schaik B., van der Mee M., Baas F., Riggins G., van Sluis P., Hermus M.C., van Asperen R., Boon K., Voute P.A., Heisterkamp S., van Kampen A., Versteeg R. 2001. The human transcriptome map: clustering of highly expressed genes in chromosomal domains. *Science*. 2001. **291**, 1289–1292.
20. Venter J.C., Adams M.D., Myers E.W., Li P.W., Mural R.J., Sutton G.G., Smith H.O., Yandell M., Evans C.A., Holt R.A., Gocayne J.D., Amanatides P., Ballew R.M., Huson D.H., Wortman J.R., Zhang Q., Kodira C.D., Zheng X.H., Chen L., Skupski M., Subramanian G., Thomas P.D., Zhang J., Gabor Miklos G.L., Nelson C., Broder S., Clark A.G., Nadeau J., McKusick V.A., Zinder N., Levine A.J., Roberts R.J., Simon M., Slayman C., Hunkapiller M., Bolanos R., Delcher A., Dew I., Fasulo D., Flanigan M., Florea L., Halpern A., Hannenhalli S., Kravitz S., Levy S., Mobarry C., Reinert K., Remington K., Abu-Threideh J., Beasley E., Biddick K., Bonazzi V., Brandon R., Cargill M., Chandramouliswaran I., Charlab R., Chaturvedi K., Deng Z., Di Francesco V., Dunn P., Eilbeck K., Evangelista C., Gabrielian A.E., Gan W., Ge W., Gong F., Gu Z., Guan P., Heiman T.J., Higgins M.E., Ji R.R., Ke Z., Ketchum K.A., Lai Z., Lei Y., Li Z., Li J., Liang Y., Lin X., Lu F., Merkulov G.V., Milshina N., Moore H.M., Naik A.K., Narayan V.A., Neelam B., Nusskern D., Rusch D.B., Salzberg S., Shao W., Shue B., Sun J., Wang Z., Wang A., Wang X., Wang J., Wei M., Wides R., Xiao C., Yan C., Yao A., Ye J., Zhan M., Zhang W., Zhang H., Zhao Q., Zheng L., Zhong F., Zhong W., Zhu S., Zhao S., Gilbert D., Baumhueter S., Spier G., Carter C., Cravchik A., Woodage T., Ali F., An H., Awe A., Baldwin D., Baden H., Barnstead M., Barrow I., Beeson K., Busam D., Carver A., Center A., Cheng M.L., Curry L., Danaher S., Davenport L., Desilets R., Dietz S., Dodson K., Doup L., Ferreira S., Garg N., Gluecksmann A., Hart B., Haynes J., Haynes C., Heiner C., Hladun S., Hostin D., Houck J., Howland T., Ibegwam C., Johnson J., Kalush F., Kline L., Koduru S., Love A., Mann F., May D., McCawley S., McIntosh T., McMullen I., Moy M., Moy L., Murphy B., Nelson K., Pfannkoch C., Pratts E., Puri V., Qureshi H., Reardon M., Rodriguez R., Rogers Y.H., Romblad D., Ruhfel B., Scott R., Sitter C., Smallwood M., Stewart E., Strong R., Suh E., Thomas R., Tint N.N., Tse S., Vech C., Wang G., Wetter J., Williams S., Williams M., Windsor S., Winn-Deen E., Wolfe K., Zaveri J., Zaveri K., Abril J.F., Guigo R., Campbell M.J., Sjolander K.V., Karlak B., Kejariwal A., Mi H., Lazareva B., Hatton T., Narechania A., Diemer K., Muruganujan A., Guo N., Sato S., Bafna V., Istrail S., Lippert R., Schwartz R., Walenz B., Yooshep S., Allen D., Basu A., Baxendale J., Blick L., Caminha M., Carnes-Stine J., Caulk P., Chiang Y.H., Coyne M., Dahlke C., Mays A., Dombroski M., Donnelly M., Ely D., Esparham S., Fosler C., Gire H., Glanowski S., Glasser K., Glodek A., Gorokhov M., Graham K., Gropman B., Harris M., Heil J., Henderson S., Hoover J., Jennings D., Jordan C., Jordan J., Kasha J., Kagan L., Kraft C., Levitsky A., Lewis M., Liu X., Lopez J., Ma D., Majoros W., McDaniel J., Murphy S., Newman M., Nguyen T., Nguyen N., Nodell M., Pan S., Peck J., Peterson M., Rowe W., Sanders R., Scott J., Simpson M., Smith T., Sprague A., Stockwell T., Turner R., Venter E., Wang M., Wen M., Wu D., Wu M., Xia A., Zandieh A., Zhu X. 2001. The sequence of the human genome. *Science*. **291**, 1304–1351.
21. Yang Q., Yoshimura G., Mori I., Sakurai T., Kakudo K. 2002. Chromosome 3p and breast cancer. *J. Hum. Genet.* **47**, 453–459.
22. Брага Э.А., Киселев Л.Л., Забаровский Е.Р. 2004. От идентификации геномного полиморфизма к диагностическим и прогностическим маркерам эпителиальных опухолей человека. *Молекуляр. биология*. **38**, 179–190.
23. Chakraborty S.B., Dasgupta S., Roy A., Sengupta A., Ray B., Roychoudhury S., Panda C.K. 2003. Differential deletions in 3p are associated with the development of head and neck squamous cell carcinoma in Indian patients. *Cancer Genet. Cytogenet.* **146**, 130–138.
24. Zabarovsky E.R., Petrenko L., Protopopov A., Vorontsova O., Kutsenko A.S., Zhao Y., Kilosanidze G., Zabarovska V., Rakhmanaliev E., Pettersson B., Kashuba V.I., Ljungqvist O., Norin E., Midtvedt T., Mollby R., Winberg G., Ernberg I. 2003. Restriction site tagged (RST) microarrays: a novel technique to study the species composition of complex microbial systems. *Nucleic Acids Res.* **15**, e95.
25. Li J., Protopopov A., Wang F., Senchenko V., Petushkov V., Vorontsova O., Petrenko L., Zabarovska V., Muravenko O., Braga E., Kisselev L., Lerman M.I., Kashuba V., Klein G., Ernberg I., Wahlestedt C., Zabarovsky E.R. 2002. NotI subtraction and NotI-specific microarrays to detect copy number and methylation changes in whole genomes. *Proc. Natl. Acad. Sci. USA*. **99**, 10724–10729.