



УДК 576.316.3

© 1994 Д. Обрадович, А. М. Бородин, Е. П. Копанцев,
Л. Л. Вагнер, С. В. Волик, О. Д. Ермолаева, Ю. Б. Лебедев,
Г. С. Монастырская, Е. Д. Свердлов

ПОЛУЧЕНИЕ И ХАРАКТЕРИСТИКИ УПОРЯДОЧЕННОЙ БИБЛИОТЕКИ ТРАНСКРИБИРУЕМЫХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ ХРОМОСОМЫ 19 ЧЕЛОВЕКА ИЗ ГИБРИДНЫХ КЛЕТОК ЧЕЛОВЕК-ХОМЯК

*Институт биоорганической химии им. М. М. Шемякина и
Ю. А. Овчинникова РАН, Москва*

Ключевые слова: гя-кДНК, хромосома 19, геном человека, *Alu*-PCR.

Описывается усовершенствованная методика получения хромосом-специфических библиотек транскрибируемых последовательностей генома человека из линий соматических гибридных клеток. Одним из основных преимуществ методики является обогащение материала человекеспецифическими последовательностями гя-РНК на каждой из трех стадий конструирования библиотеки. Разработанный метод был использован для конструирования хромосомспецифической библиотеки из гибридной линии клеток. Полученная ранжированная библиотека охарактеризована. Определена первичная структура 80 произвольно выбранных клонов и проведен ее анализ. Показано, что около 95% клонов являются человекспецифическими и происходят из гя-РНК, для некоторых клонов определена хромосомная локализация. Проанализированы преимущества и недостатки разработанного метода.

В настоящее время наблюдается значительный прогресс в области генетического и физического картирования генома человека [1—5], что создает основу для решения основной, хотя и более отдаленной задачи — составления функциональной карты или, пользуясь термином, предложенным МакКьюзиком [3, 6], функциональной анатомии генома человека. Использование анатомической терминологии, возможно, более обоснованно, чем терминологии географической (картирование), поскольку дает возможность более четко классифицировать различные функции (и соответственно дисфункции) генома: например, клиническая анатомия генома определяет область исследований генома, связанную с анализом генов, вызывающих одно- или многофакторные наследственные заболевания, или генети-

Использованные сокращения: гя-РНК — гетерогенная ядерная РНК, гя-кДНК — гетерогенная ядерная кДНК, STS — sequence tagged site, EST — expressed sequenced tag, cM — сантиморган, PCR — полимеразная цепная реакция, PIC — содержание полиморфной информации (polymorphism information content).

Адрес для переписки: 117871, Москва, ул. Миклухо-Маклая, 16/10.

ческий вклад в широкий спектр заболеваний, включая такие, как диабет, атеросклероз, гипертония. Эволюционная анатомия предполагает сравнительный анализ геномов различных организмов, анатомия развития — исследование генов, координированно включающихся и выключающихся в процессе развития эмбриона и т. д.

Последние годы характеризовались развитием методической базы физико-химической биологии, которая сделала возможным получение карт отдельных хромосом с разрешением, достаточным, чтобы говорить не только о положении таких крупных функциональных единиц генома, как гены, размеры которых порой превышают сотни тысяч и даже миллионы пар оснований, но и о более детализованных картах, которые указывают положение регуляторных элементов, управляющих различными аспектами функционирования генома — репликацией, транскрипцией и т. д. Такие элементы занимают в геноме существенно меньшее пространство, но их значимость для его жизнеспособности трудно переоценить. Наконец, возможно, целесообразно включить в функциональную анатомию генома те его элементы, которые столь часто повторяются в геноме, что трудно ожидать, чтобы они представляли собой просто геномный мусор. Речь идет о многочисленных семействах повторяющихся элементов — SINE, LINE, сателлитах, мини- и микросателлитах, относительно которых высказываются различные гипотезы (см., например, [7]).

Хотя все эти элементы непосредственно связаны с функциональной анатомией генома, первоочередной задачей на сегодня является определение положений генов в хромосомах и установление их функций. Именно этот смысл вкладывает в определение функциональной анатомии МакКьюзик [6]. Большое внимание в настоящее время уделяется проблеме получения транскрипционной карты генома человека [8, 9]. При этом подходы к решению этой задачи можно разделить на две группы: идентификационные (определение транскрибируемости того или иного выделенного фрагмента генома) и выделительные (направленное выделение заведомо транскрибирующихся последовательностей).

В данной работе мы описываем метод создания хромосомспецифических библиотек транскрибирующихся последовательностей человека, относящийся ко второй группе методов, и его применение к созданию библиотеки фрагментов транскрибирующихся последовательностей хромосомы 19. Фрагменты из этой библиотеки могут использоваться в качестве молекулярно-генетических зондов при физическом и генетическом картировании хромосомы 19 и как гибридизационные зонды для получения полноразмерных генов из геномных библиотек.

1. Стратегия метода

Метод основан на принципе, предложенном Корбо и др. [10]. В этой работе для получения хромосомспецифических библиотек транскрибирующихся последовательностей человека использованы гибридные клетки человек-хомяк, содержащие единственную хромосому человека или ее фрагмент. Гетерогенная ядерная РНК (гя-РНК), выделенная из таких клеток, содержит человеческий компонент, отличающийся от гя-РНК хомяка, в частности, присутствием специфических повторов *Alu* (SINE) или *L1* (LINE) (R на рис. 1). С помощью олигонуклеотидных затравок, имеющих последовательности, соответствующие консервативным внутри вида последовательностям этих повторов, можно осуществить дифференциальную обратную транскрипцию человеческих гя-РНК в смеси гя-РНК гибридных клеток.

Путем непосредственного клонирования полученной таким образом гетерогенной ядерной кДНК Корбо и др. смогли получить библиотеку гя-кДНК q24-qter — области X-хромосомы человека, содержащую около 80% человекспецифических клонов. Мы сделали попытку усовершенствовать метод, с тем чтобы увеличить специфичность дискриминации человеческих гя-РНК. С этой целью была разработана схема, представленная на рис. 1 и включающая три стадии, каждая из которых направлена на увеличение содержания человекспецифических клонов в конечной библиотеке. В качестве исходного материала была использована

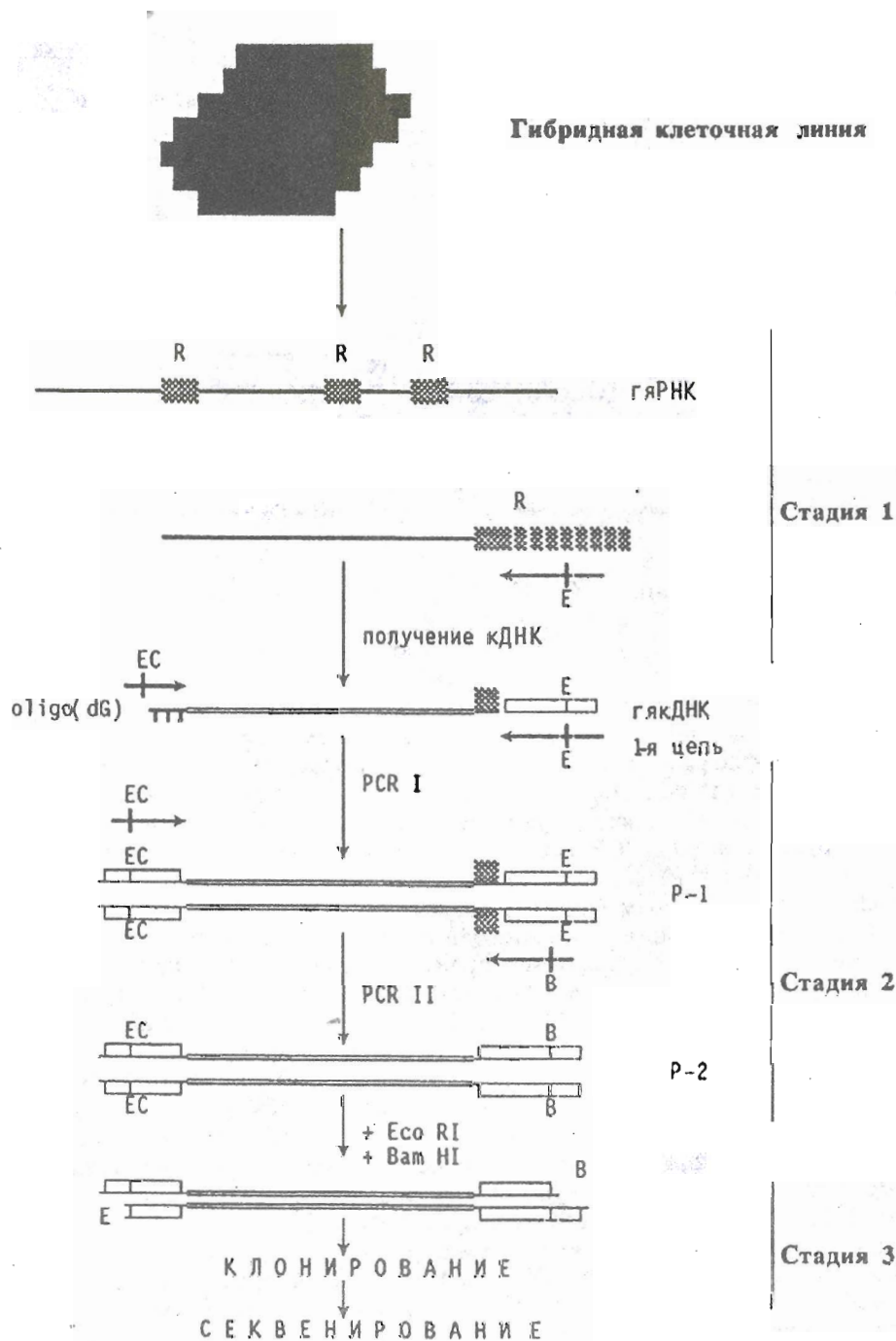


Рис. 1. Схема метода. Повторяющиеся элементы в гγ-РНК обозначены заштрихованными прямоугольниками R. Горизонтальными стрелками отмечены праймеры, вертикальные черточки на стрелках — положения участков узнавания рестриктаз *EcoRI* (E) и *BamHI* (B), EC — oligo(dC)₁₇-праймер с участком узнавания *EcoRI* на 5'-конце. Светлыми прямоугольниками обозначены последовательности, соответствующие праймерам в продуктах PCR P-1 и P-2

Структуры олигонуклеотидных праймеров, использованных в работе*

Праймер	Структура
Е, 517	CAGAATTCGCGATCTCGGCTCACTGCAA
В, ALN3	CAAGGATCCGCTCACTGCAACCTCTGCCTCC
EC	GAGAAATCCCCCCCCCCCCCCCCCCC
16L	GATCTGATGGCCGAGCTTGT
16R	GAATTC AATTGAACTTGGATTATGAGG
42L	CTATGGAGAGGGTTTCAATGA
42R	TGATTCCTCCCTGAATCTCAG
55L	GATCCAGCCCGTGTGGCTTCTCTG
55R	AAAGAGGCAAATAGGCCAAGTCATG

* Жирным шрифтом выделены участки узнавания рестриктаз *Bam*HI и *Eco*RI в праймерах В и Е соответственно.

гибридная клеточная линия GL-12-42, содержащая хромосому 19 человека и фрагменты хромосом 22 и X, транслоцированные на хромосому хомяка [11]. гя-РНК, выделенная из этих гибридных клеток, была использована в качестве матрицы для синтеза первой цепи кДНК с затравкой, которая представляла собой модификацию известного *Alu*-специфического праймера 517, используемого в *Alu*-PCR-методе анализа человеческих хромосом [12]. Модифицированный праймер 517-Е содержал *Eco*RI-сайт на 5'-конце (табл. 1). К полученной первой цепи кДНК присоединяли oligo(dG) с помощью концевой нуклеотидилтрансферазы и использовали для первого раунда полимеразной цепной реакции (PCR). В качестве праймеров для PCR использовали упомянутый выше праймер 517-Е и (dC)₂₀, содержащий *Eco*RI-сайт на 5'-конце (EC-праймер, табл. 1). Продукт амплификации (P1), полученный в результате 1-й стадии, содержал на обоих концах *Eco*RI-сайты. Он был использован в качестве матрицы для второго раунда амплификации. В этом случае одним праймером служил EC, а другим праймер — ALN3 (В), частично перекрывал первый (517-Е) и не имел на 5'-конце *Bam*HI-сайт. Это частичное перекрывание обеспечивает увеличение специфичности получаемого в результате продукта (P2) [13]. Кроме того, продукт P2 отличается от продукта P1 тем, что он содержит на одном из концов *Eco*RI-сайт за счет EC-праймера, а на другом — *Bam*HI-сайт за счет праймера ALN3. В принципе это позволяет осуществить направленное клонирование продукта P2: после расщепления всей амплификационной смеси рестрикционными эндонуклеазами *Eco*RI и *Bam*HI только продукт P2 будет иметь два разных сайта, *Eco*RI и *Bam*HI, и может быть клонирован в соответствующий вектор, расщепленный этими же двумя рестрикционными эндонуклеазами.

В данной работе вместо *Bam*HI для расщепления продуктов амплификации использовали рестрикционную эндонуклеазу *Sau*3AI, которая расщепляет центральный тетра nukлеотид *Bam*HI специфичной последовательности. Это давало возможность уменьшить размер *Alu*-фрагментов в продуктах амплификации за счет расщепления по существующим в этих последовательностях внутренним *Sau*3AI-сайтам, но исключало дискриминирующую функцию ALN3-праймера в направленном клонировании продукта P2. Несмотря на это, доля продукта P1 в смеси не может быть очень высокой, поскольку в результате 30 циклов его амплификации следует ожидать превышение его концентрации в смеси над продуктом P1 примерно в $8 \cdot 10^6$ раз, если выход амплификации на каждую стадию принять равным 1,7 [14].

2. Получение и анализ библиотеки

Желаемое число независимых клонов в библиотеке оценивали, исходя из следующих соображений. Данная библиотека должна служить источником STS/EST для хромосомы 19 человека. Исходя из общего числа *Alu*-повторов в геноме человека $\sim 5 \cdot 10^5$ [7], в хромосоме 19 человека, занимающей примерно 2% общей длины генома [15] и имеющей длину около $60 \cdot 10^6$ п. о., должно содержаться около 10 000 *Alu*-повторов.

Около 10%, т. е. примерно 1000 таких повторов, может оказаться в гя-РНК, если исходить из того, что в среднем в клетке экспрессируется около 10% генов. Это количество обеспечивает среднее распределение транскрибированных в хромосоме *Alu*-элементов через ~ 60 000 п. о. Задачей современного генетического и/или физического картирования является получение сравнительно равномерно распределенных молекулярно-генетических маркеров через 2—5 сМ и STS/EST через 10^5 п. о. [1—5, 16—19]. Приняв, что в среднем 1 сМ равен $\sim 1 \cdot 10^6$ п. о. и считая, что многие STS/EST-маркеры могут оказаться одновременно молекулярно-генетическими маркерами, можно в качестве первого шага взять число, среднее между 2 сМ и 10^5 п. о., т. е. $\sim 10^6$ п. о., для желаемого расстояния фрагментов библиотеки друг от друга при их распределении вдоль хромосомы. Эта величина определяет нижний предел желаемого числа независимых клонов в библиотеке как 60.

Эти клоны должны быть равномерно распределены вдоль хромосомы. Поэтому, чтобы проводить выбор необходимого числа равномерно распределенных маркеров, нужно получить избыточную библиотеку.

Нами получена библиотека, содержащая 216 клонов, и подготовлена для использования в упорядоченной форме [20].

2.1. Анализ последовательностей случайных клонов из библиотеки

Последовательности 76 клонов из библиотеки определены и проанализированы путем сопоставления с последовательностями, депонированными в банке данных EMBL. Характеристика этих последовательностей приведена в табл. 2. Большинство клонов содержали фрагменты *Alu*-повторов с различной степенью гомологии с консенсусной *Alu*-последовательностью или с различными классами *Alu*-повторов [22]. Их уникальные части не имеют гомологии с последовательностями клонов, депонированными в EMBL-банке.

Наличие высокомонологичных *Alu*-последовательностей может служить доводом в пользу человекспецифичности содержащих их фрагментов ДНК.

Чтобы определить происхождение клонов, не содержащих *Alu*-повторов, был поставлен дополнительный эксперимент.

На рис. 2 приводятся последовательности трех клонов, не содержащих *Alu*-последовательностей из полученной библиотеки, на основании которых были синтезированы три пары праймеров: 16L и 16R; 42L и 42R; 55L и 55R. Эти праймеры были использованы для PCR-амплификации ДНК, полученных из эмбриональных фибробластов человека, из клеток хомяка Ag-17-1, использованных для получения гибридных клеток человек-хомяк GL-12-42, и из самих гибридных клеток.

Результаты PCR-амплификации приведены на рис. 3. В случае клонов рН42/1АВ и рН55/1АВ продукты амплификации ожидаемой длины (237 п. о. для последовательности рН42/1АВ и 198 п. о. для последовательности рН55/1АВ) наблюдаются только при амплификации ДНК из клонов человека и гибридных клеток, но не клеток хомяка. Следовательно, эти последовательности происходят из человеческого компонента гибридных клеток.

При использовании праймеров, специфичных для последовательности рН55/1АВ, в PCR-анализе ДНК человека (но не всех гибридных клеток) наблюдается дополнительная полоса 450 п. о. Специальными экспериментами (будут

Сравнение некоторых клонов библиотеки с консенсусной последовательностью *Alu* [21]

Наименование клона	Иденти- фикатор EMBL	Общая длина (п.о.) ^{1*}	Сравнение с консенсусной последовательностью <i>Alu</i>			
			Компле- ментар- ная ^{2*}	Координаты консенсуса (начало) ^{3*}	Координаты консенсуса (конец)	Степень гомологии с консенсусом, %
pH11/IAB	Z22523	225	Y	1	61	85,2
pH12/IAB	Z22531	172	Y	1	63	85,7
pH13/IAB	Z22530	87	Y	1	76	75
pH14/IAB	Z22529	181	- ^{4*}	-	-	-
pH15/IAB	Z22524	244	-	-	-	-
pH16/IAB	Z22581	162	-	-	-	-
pH17/IAB	Z22578	78	Y	1	44	79,5
pH20/IAB	Z22582	193	Y	134	170	94,5
pH34/IAB	Z22580	195	Y	1	62	87,1
pH37/IAB	Z22583	257	Y	23	85	93,7
pH42/IAB	Z22575	237	-	-	-	-
pH47/IAB	Z22579	172	Y	1	71	81,6
pH55/IAB	Z22584	198	-	-	-	-
pHa1/IDO	Z25794	100	-	-	-	-
pHa260u/IDO	Z24764	295	Y	1	198	82,7
pHa263u/IDO	Z23144	332	Y	2	167	80,7
pHa266u/IDO	Z25410	291	Y	9	199	88,4
pHa272u/DO	Z24760	303	-	-	-	-
pHa273u/IDO	Z24765	404	-	-	-	-
pHa276u/IDO	Z24762	136	Y	2	59	93,1
pHa277u/IDO	Z24761	365	-	-	-	-
pHa280u/IDO	Z24763	150	Y	14	78	83,1
pHa281u/IDO	Z25792	211	Y	1	109	84,4
pHa281u/IDO	Z24792	211	Y	106	181	85,5
pHa282u/IDO	Z24766	177	Y	2	63	82,5
pHa284u/IDO		205	Y	29	177	82,5
pHa285u/IDO	Z25786	337	Y	2	63	85,7
pHa289u/IDO	Z25780	93	Y	2	62	85,4
pHa294u/IDO	Z25785	263	Y	2	63	90,5
pHa299u/IDO	Z25789	291	Y	2	61	83,6
pHa302u/IDO	Z25779	265	-	-	-	-
pHa303u/IDO	Z25784	248	Y	3	196	82,05
pHa306u/IDO	Z25787	330	Y	18	88	86,1
pHa308u/IDO	Z25790	317	Y	15	75	85,2
pHa309u/DO	Z25782	86	Y	11	60	90
pHa310u/IDO	Z25783	115	Y	5	92	85,2
pHa314u/IDO	Z25778	276	-	-	-	-
pHa326u/IDO	Z25412	310	Y	2	63	83,8
pHa327u/IDO	Z23145	97	Y	2	63	88,9
pHa334u/IDO	Z23146	278	Y	2	62	87,1
pHa336u/IDO	Z25411	318	Y	2	63	91,9

Таблица 2 (окончание)

pHa346u/1DO	Z25400	273	-	-	-	-
pHa347u/1DO	Z23148	349	Y	2	61	88,9
pHa384u/1DO	Z25413	275	Y	2	199	81,3
pHa385u/1DO	Z25415	329	Y	8	64	94,7
pHa401u/1DO	Z23147	312	Y	1	231	89,6
pHa402u/1DO	Z23149	322	Y	2	63	80,9
pHa407u/1DO	Z23151	300	-	-	-	-
pHa419u/1DO	Z25417	379	Y	2	63	83,9
pHa420u/1DO	Z25416	309	Y	2	63	83,9
pHa422u/1DO	Z25414	134	Y	4	71	83,8
pHa424u/1DO	Z25401	253	Y	2	59	82,7
pHa425u/1DO	Z25404	347	Y	2	62	91,9
pHa427u/1DO	Z25402	294	Y	2	60	85,2
pHa428u/1DO			-	-	-	-
pHa433u/1DO	Z25403	290	Y	2	63	90,3
pHa438u/1DO	Z25788	242	Y	8	60	92,4
pHa558u/1DO		330	Y	6	88	85,6
pHa561u/1DO		128	Y	6	88	86,7
pHa570u/1DO		262	Y	14	59	85,1
pHa574u/1DO		301	Y	2	63	90,3
pHa577u/1DO		130	Y	2	89	80,7
pHa579u/1DO		301	-	-	-	-
pHa580u/1DO		361	-	-	-	-
pHa586u/1DO		264	Y	83	240	82,3
pHa586u/1DO		264				
pHa588u/1DO		310	Y	6	84	87,3
pHa589u/1DO		124	Y	2	77	63,6
pHa590u/1DO		398	-	-	-	-
pHa593u/1DO		309	Y	2	63	83,9
pHa601u/1DO		310	Y	8	59	94,3
pHa607u/1DO		317	Y	2	62	85,2
pHa612u/1DO		386	Y	2	117	67,4
pHa614u/1DO		262	Y	8	64	90,9
pHa615u/1DO		377	N	219	377	76,6
pHa627u/1DO		336	Y	8	61	84,2
mh50/1AB		206	Y	1	62	85,5
pHa604u/1DO		405	-	-	-	-

1* Дана длина секвенированной части последовательности.

2* Y - последовательность комплементарна обычно приводимой в статьях последовательности.

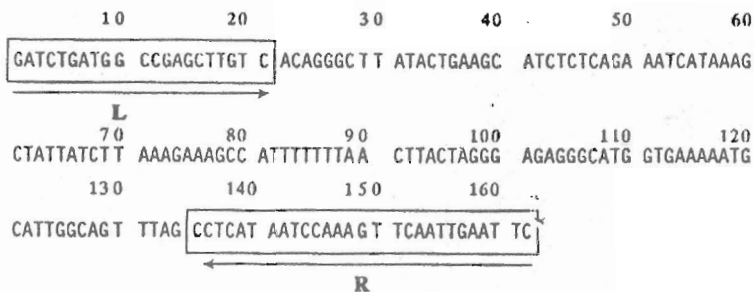
3* Координаты даются относительно начала нашей последовательности.

4* Последовательность не содержит фрагмента, гомологичного *Alu*-последовательности.

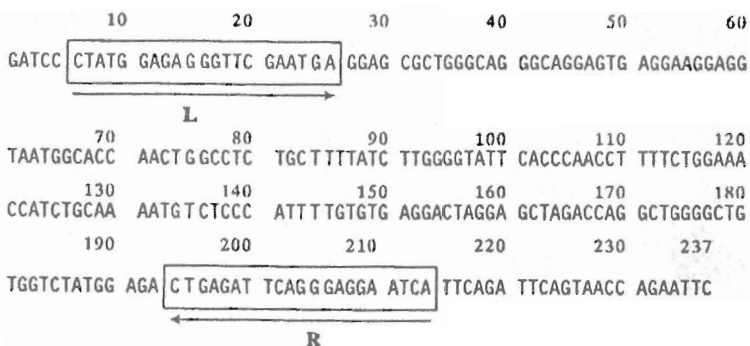
опубликованы отдельно) было показано, что эта полоса появляется в результате амплификации ДНК хромосомы 5 человека. Это означает, что существует семейство генов, содержащих последовательности, гомологичные последовательностям pH55/1AB. Одни из генов этого семейства локализованы на хромосоме 5, другие — на хромосоме 19. Появление слабых полос, соответствующих продуктам амплификации ДНК хомяка, хотя их длины и отличаются от ожидаемых на основании последовательности pH55/1AB, может свидетельствовать о существовании подобного семейства и в геноме хомяка. В случае использования праймеров 16L и 16R продукты амплификации наблюдаются только для ДНК хомяка и гибридных клеток, но не для ДНК человека. Таким образом, этот клон содержит ДНК из генома хомяка.

В целом из 80 проанализированных клонов 67 имеют человеческое происхождение (судя по наличию *Alu*-повторов или по результатам PCR-анализа, подобному выше). Из клонов, не содержащих *Alu*-последовательности, примерно

a) pH16/1AB



b) pH42/1AB



c) pH55/1AB

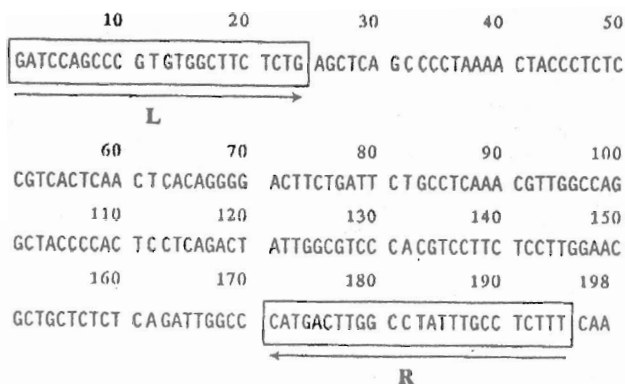


Рис. 2. Структуры некоторых клонов, не содержащих *Alu*-последовательностей. В рамки заключены участки, соответствующие специфическим праймерам (см. табл. 1). Стрелками указано направление синтеза новых цепей в PCR

2/3 также скорее всего являются человеческими (см. PCR-анализ). В целом библиотека содержит примерно 95% клонов человеческого происхождения.

Отдельным вопросом, который требует специального анализа, является доказательство происхождения клонов библиотеки из гя-РНК, а не из возможной примеси ДНК, хотя сам использованный метод выделения гя-РНК [23] сводит подобные примеси к минимуму. Для этой цели были предприняты отдельные эксперименты, гя-РНК подвергали обработке ДНКазой или РНКазой и использовали для синтеза первой цепи кДНК, которую затем подвергали PCR-амплификации с парами праймеров, подобно тому как описано выше для доказательства человеческого происхождения клонов. Предварительная обработка образца РНКазой приводила к полному подавлению PCR-амплификации, тогда как обработка

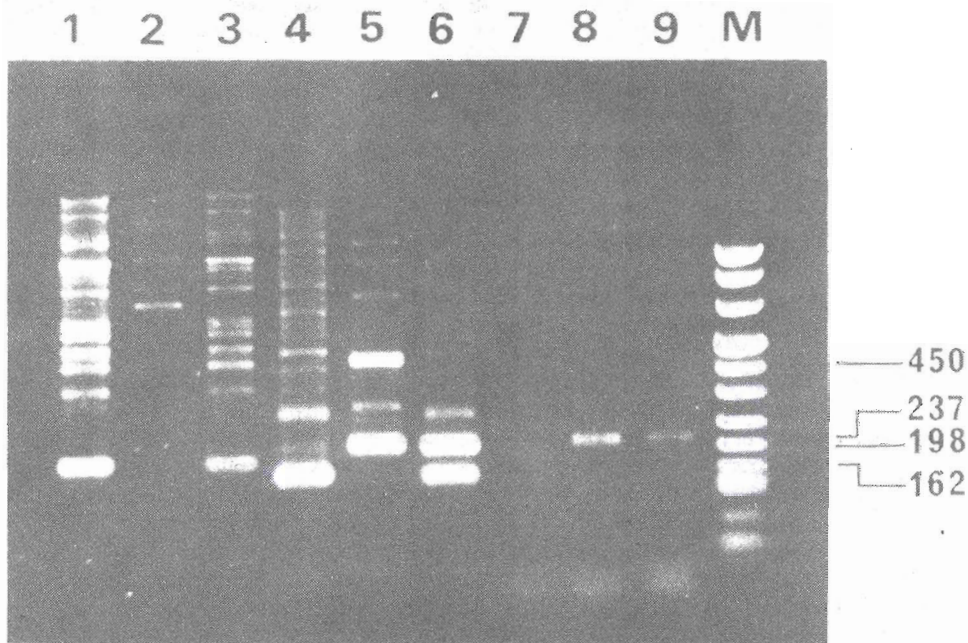


Рис. 3. PCR-анализ некоторых клонов библиотеки. Матрицей служила геномная ДНК хомяка (дорожки 1, 4, 7), человека (дорожки 2, 5, 8) или гибридной клеточной линии (дорожки 3,6,9) в реакции PCR со специфическими праймерами для клонов рН16/1АВ (1—3), рН55/1АВ (4—6) и рН42/1АВ (7—9). Дорожка М — молекулярный маркер Boehringer Mannheim VIII (справа — размер в парах оснований)

ДНКазой мало влияла на выход продуктов (результаты не показаны). Это является свидетельством происхождения клонов библиотеки из гя-РНК.

Таким образом, можно заключить, что полученная библиотека высоко обогащена транскрибируемыми последовательностями, относящимися к человеческим хромосомам, содержащимся в гибридных клетках GL-12-42. Поскольку основной человеческой хромосомой в этих клетках является хромосома 19, можно полагать, что основное количество клонов из этой библиотеки содержит материал из этой хромосомы. Действительно, предварительные данные по отнесению клонов библиотеки к хромосомам человека показывают, что примерно 2/3 полученных клонов относятся к хромосоме 19 (данные будут публиковаться отдельно).

К настоящему времени три группы авторов использовали гя-РНК из гибридных клеток как источник экспрессирующихся последовательностей отдельных хромосом или определенных участков хромосом человека [10, 24—26].

В первой из них [24] авторы создали тотальную библиотеку гя-ДНК из гибридных клеток и затем осуществляли ее скрининг с человеческой ДНК в качестве гибридизационного зонда. Из 800 000 клонов удалось выбрать 275 человекспецифических.

Корбо и др. [10] заменили скрининг тотальных библиотек целенаправленным получением небольших библиотек, содержащих человеческие транскрибируемые последовательности, путем специфической обратной транскрипции, индцируемой *Alu*-специфическими праймерами на гя-РНК из гибридных клеток, и, как уже упоминалось выше, смогли получить библиотеку, содержащую 80% клонов из гя-РНК Xq24-qter-области человека.

Наконец, когда наша работа была завершена [25], появилась статья [26], в которой авторы добавили к методу [10] стадию PCR-амплификации, однако в отличие от нас авторы [26] не использовали систему увеличения специфичности

PCR за счет перекрывающихся праймеров и полученная ими библиотека гя-ДНК из короткого плеча хромосомы 12 содержала около 80% человекспецифических клонов.

Метод, предлагаемый нами, дает возможность достаточно простого получения высокообогащенных библиотек транскрибированных последовательностей отдельных хромосом или их определенных областей из соответствующих гибридных клеток.

3. Потенциальное применение гя-кДНК библиотеки

3.1. Использование в качестве источника STS/EST-маркеров и проб на молекулярно-генетические маркеры (МГМ)

Одной из первоочередных задач программы «Геном человека» является получение физической карты, в которой STS/EST-маркеры [1—5, 19] должны быть равномерно распределены вдоль хромосом с интервалом 10^5 п. о. С другой стороны, планируется сконструировать генетическую карту, в которой молекулярно-генетические маркеры будут распределены вдоль хромосом на расстоянии 2—5 сМ [1—5, 17—18]. STS представляет собой короткую (200—500 п. о.) известную последовательность, обладающую тремя свойствами: а) локализованность в определенном участке хромосомы, б) уникальность для генома, в) специфическая PCR-амплифицируемость соответствующего фрагмента геномной ДНК с праймерами, выведенными на основании данной последовательности. Молекулярно-генетический маркер, направленный на обнаружение сцепленности генов, должен обладать вдобавок к вышеназванным трем свойствам также полиморфизмом. Полиморфность маркера принято определять величиной PIC [1—3, 18] (polymorphism information content), которая варьирует от 0 (неполиморфные последовательности) до величин, близких к 1, для высокополиморфных.

Из этих задач вытекает и наиболее прямое применение полученной библиотеки — в качестве источника STS/EST-маркеров, т. е. одновременно известной и транскрибируемой последовательности. Обычно в качестве источника STS/EST рассматривают кДНК (см., например, [1—3, 27, 28]). Однако с точки зрения выбора праймеров для амплификации геномной ДНК последовательности кДНК представляют определенную сложность: выбранные для праймеров последовательности могут оказаться в геномной ДНК разделенными интроном. С этой точки зрения транскрибируемые последовательности, получаемые при анализе гя-кДНК, более безопасны, поскольку гя-РНК содержат интроны. Таким образом, полученная библиотека представляет более надежный источник STS/EST, чем библиотеки кДНК, хотя, конечно, не все клоны из нее будут отвечать всем вышесформулированным требованиям. Клоны рН42/1АВ и рН55/1АВ дают возможность амплификации уникальных последовательностей геномной ДНК и поэтому могут служить STS/EST-маркерами с момента их локализации на определенной области хромосомы (амплификация двух хромосом в случае пробы рН55/1АВ не препятствует ее использованию в качестве STS-маркера, поскольку получаемые фрагменты имеют для разных хромосом разные длины).

Можно полагать, что в качестве источника молекулярно-генетических зондов для построения генетических карт библиотеки гя-кДНК имеют еще большие преимущества по сравнению с библиотеками кДНК с точки зрения возможной их полиморфности. кДНК представляет собой экзонные части генов, которые более консервативны, чем их интроны [29]. Поскольку *Alu*-последовательности сосредоточены в основном в нетранслируемых частях генов [30], прилегающие к *Alu*-участкам уникальные последовательности также по большей части имеют интронное происхождение и поэтому должны обладать большими величинами PIC и быть более полезны в качестве молекулярно-генетических маркеров.

3.2. Использование в качестве гибридационных проб для получения полноразмерных генов

Это применение библиотек также очевидно, и здесь они также имеют существенное преимущество перед библиотеками кДНК.

Многие гены в геноме являются членами семейств генов, которые обладают высокой степенью гомологии, особенно в экзонных частях [31]. Поэтому использование кДНК в качестве гибридационных зондов вызывает серьезные проблемы при идентификации индивидуальных генов семейств в библиотеках генов. Зонды, происходящие из более вариабельных нетранслируемых частей генов, в этом отношении существенно более перспективны, поскольку каждая из них характеризует индивидуального члена семейства.

4. Ограничения метода

Разработанный метод имеет ряд ограничений. Одно из них связано с экспрессией лишь части генов в данной линии гибридных клеток. Этот недостаток является общим ограничением всех подходов, нацеленных на картирование экспрессирующих последовательностей, в том числе кДНК [27]. Клетки мозга, экспрессирующие наибольшее число различных генов, содержат не более 30% потенциально возможных мРНК, в клетках других тканей эта цифра еще меньше. Это ставит серьезные проблемы перед исследователями при получении полных функциональных карт генома. В этой связи ситуация для гя-РНК несколько лучше, особенно если принять во внимание данные об отсутствии абсолютных запретов на первые стадии экспрессии тканеспецифических генов в любом типе клеток [32]. Таким образом, можно надеяться, что проблема получения транскрипционной библиотеки, содержащей полный набор транскрибируемых генов, принципиально решаема и требует лишь технической разработки. Следует, однако, заметить, что эта проблема не стоит на первом этапе построения физических и генетических карт, использующих STS/EST-систему, поскольку расположение этих маркеров через $(1000-100) \cdot 10^3$ п. о. не предполагает выделения всех транскрибируемых последовательностей.

Экспериментальная часть

Выделение гя-РНК из гибридных клеток. Фракцию гя-РНК выделяли из ядер клеточной линии человек — китайский хомячок GL-12-42 в соответствии с описанной методикой [23]. В стандартных экспериментах получали 500 мкг РНК из 10^8 клеток.

Получение гя-кДНК. Синтез первой цепи гя-кДНК проводили в описанных [10] условиях с праймером 517 и при температуре 55°C . Полученную гя-кДНК обрабатывали концевой дезоксирибонуклеотидилтрансферазой (Promega, США) и по окончании реакции материал подвергали 30 циклам PCR с помощью 517 с EC-праймеров. Для второго раунда PCR использовали праймеры EC и внутренний по отношению к 517-му праймер В. Полученную гя-кДНК после обработки ферментами EcoRI и Sau3AI клонировали в плазмидном векторе pGEM-3Z.

Секвенирование ДНК. Определение первичной последовательности гя-кРНК проводили после переклонирования фрагментов в вектор M13mp10 согласно стандартной методике [33].

Гибридационный анализ. Агарозный электрофорез, перенос ДНК-фрагментов на нейлоновую мембрану и гибридизацию проводили по стандартным методикам [34]. Радиоактивное мечение проб осуществляли методом статистической затравки [34].

PCR-амплификация. Амплификацию проводили в термоблоке Perkin — Elmer/Cetus в 50 мкл буфера, содержащего 67 мМ трис-НСI (рН 8,8), 17 мМ $(\text{NH}_4)_2\text{SO}_4$, 1 мМ дитиотреит, 2 ед. акт. Таq-ДНК-полимеразы («Гентест», Москва),

3 mM MgCl₂, 100 нг ДНК-матрицы, 250 мкмоль/л каждого dNTP и 25 пмоль специфического праймера, в следующем циклическом режиме: 94° С — 10 с, 65° С — 10 с и 72° С — 2 мин. В стандартных экспериментах обычно проводили 30 циклов PCR.

СПИСОК ЛИТЕРАТУРЫ

1. Stephens J. C., Cavanaugh M. L., Gradie M. I., Mader M. L., Kidd K. K. // *Science*. 1990. V. 250. P. 237—250.
2. Weissenbach J., Gyapay G., Dib C., Vignul A., Morissette J., Millasseau P., Vaysseix G., Luthrop M. // *Nature*. 1992. V. 359. P. 794—801.
3. McKusick V. A. // *FASEB J*. 1991. V. 5. P. 12—20.
4. Cox D. R., Burmeister M., Price E. R., Kim S., Myers R. M. // *Science*. 1990. V. 250. P. 245—250.
5. Billings P. R., Smith C. L., Cantor C. R. // *FASEB J*. 1991. V. 5. P. 28—34.
6. McKusick V. A. // *J. Hered.* 1990. V. 71. P. 370—391.
7. Vogt P. // *Human Genet.* 1990. V. 84. P. 301—336.
8. Hochgeschwender U., Brennan M. B. // *BioEssays*. 1991. V. 13. P. 139—144.
9. Hochgeschwender U. // *Trend. Genet.* 1992. V. 8. P. 41—44.
10. Corbo L., Maley J. A., Nelson D. L., Caskey C. T. // *Science*. 1990. V. 249. P. 652—655.
11. Копанцев Е. П., Нестерова Т. Б., Бородин А. М., Закиян С. М. // *Генетика*. 1993. Т. 29. С. 40—51.
12. Nelson D. L. // *Methods: a Companion Methods in Enzymology*. 1991. V. 2. P. 60—74.
13. Ehrlich H. A., Gelfand D., Sninsky J. J. // *Science*. 1991. V. 252. P. 1643—1651.
14. Bloch W. // *Biochemistry*. 1991. V. 30. P. 2735—2747.
15. Therman E. // *Human Chromosomes Structure Behavior Effects*. 2th. Ed. N. Y.: Berlin; Heidelberg; Tokyo: Springer-Verlag, 1986.
16. Galau G. A., Klein W. H., Britten R. J., Davidson E. H. // *Arch. Biochem. and Biophys.* 1977. V. 179. P. 584—599.
17. White R., Leppert M., Bishop D. T., Barker D., Berkowitz J., Brown C., Callahau P., Holm T., Jerominski L. // *Nature*. 1985. V. 313. P. 101—105.
18. Botstein D., White R. L., Scolnick M., Davis R. W. // *Amer. J. Hum. Genet.* 1980. V. 32. P. 314—331.
19. Olson M., Hood L., Cantor C., Botstein D. // *Science*. 1989. V. 245. P. 1434—1435.
20. Lennon G. G., Lehrach H. // *Trend. Genet.* 1991. V. 7. P. 314—317.
21. Jurka J., Smith T. // *Proc. Nat. Acad. Sci. USA*. 1988. V. 85. P. 4775—4778.
22. Jurka J., Milosavljevic A. // *J. Mol. Evol.* 1991. V. 32. P. 105—121.
23. Zerivits K., Akusjarvi G. // *Gene Anal. Technol.* 1989. V. 6. P. 101—109.
24. Liu P., Legersky R., Siciliano M. J. // *Science*. 1989. V. 246. P. 813—815.
25. Kopantzev E., Launer G., Lebedev Yu., Wagner L., Borodin A., Sverdlov E. // *Abstract of Genome Mapping and Sequencing Conference*. Cold Spring Harbor. May 12—16. N. Y.: CSH, 1993. P. 134.
26. Baens M., Chaffanet M., Cassiman J.-J., van den Berhe H., Marynen P. // *Genomics*. 1993. V. 16. P. 214—218.
27. Adams M. D., Kelley J. M., Gocayne J. D., Dubnik M., Polymeropoulos M. H., Xiao H., Merril C. R., Wu A., Olde B., Moreno R. F., Kerlavage A. R., McCombie W. R., Venter J. G. // *Science*. 1991. V. 252. P. 1651—1656.
28. Polymeropoulos M. H., Xiao H., Glodec A., Gorski W., Adams M. D., Moreno R. F., Fitzgerald M. G., Venter J. G., Merril C. R. // *Genomics*. 1992. V. 12. P. 492—496.
29. Кумура М. Молекулярная эволюция: теория нейтральности: Пер. с англ. М.: Мир, 1985.
30. Zuckerkandl E., Laffer G., Jurka J. // *J. Mol. Evol.* 1989. V. 29. P. 504—512.
31. Moore G. P. // *Trends Biochem. Sci.* 1983. V. 8. P. 411—414.
32. Sarkar G., Sommer S. S. // *Science*. 1989. V. 244. P. 331—334.
33. Royal A., McGraw D. F. // *Analyt. Biochem.* 1984. V. 143. P. 298—303.
34. Sambrook J., Fritsch E. F., Maniatis T. // *Molecular Cloning*. 2ed. CSH: Cold Spring Harbor Laboratory Press, 1989.

Поступила в редакцию
30.XI.1993

После доработки
29.III.1994

*D. Obradovitch, A. M. Borodin, E. P. Kopantzev, L. L. Wagner,
S. V. Volik, O. D. Ermolaeva, Yu. B. Lebedev,
G. S. Monastyrskaya, E. D. Sverdlov*

**CONSTRUCTION AND CHARACTERISTICS OF ARRAYED LIBRARY
OF TRANSCRIBED HUMAN CHROMOSOME 19 SEQUENCES
FROM THE HUMAN-HAMSTER HYBRID CELL LINE**

*M. M. Shemyakin and Yu. A. Ovchinnikov Institute of Bioorganic Chemistry,
Russian Academy of Sciences, Moscow 117871, ul. Miklukho-Maklaya, 16/10*

Key words: hn-cDNA, chromosome 19, human genome, *Alu*-PCR.

A new technique based on *Alu*-PCR amplification of hn-RNA is described for the extraction of human-specific transcribed sequences from a hybrid cell line. Arrayed library of hn-cDNA was constructed and characterized by sequencing about 80 individual clones. A high enrichment by human-specific sequences (about 95%) was demonstrated.