

СРАВНИТЕЛЬНЫЙ РЕСТРИКЦИОННЫЙ АНАЛИЗ ХРОМОСОМНОЙ ДНК МЫШИ *IN VITRO* И *IN SILICO*

В.А. ЧЕРНУХИН*, М.А. АБДУРАШИТОВ, В.Н. ТОМИЛОВ,
Д.А. ГОНЧАР, С.Х. ДЕГТЯРЕВ

НПО «СибЭнзим», Новосибирск

На основе ранее предложенного метода рестрикционного анализа геномов млекопитающих *in silico* рассчитаны диаграммы распределения фрагментов хромосомной ДНК мыши при ее расщеплении по 17 нуклеотидным последовательностям, являющимися сайтами узнавания эндонуклеаз рестрикции. Проведен анализ имеющихся в базе данных нуклеотидных последовательностей LINE1 повторов мыши и фрагментов ДНК, образующихся при расщеплении этих повторов. Установлено, что абсолютное большинство пиков на диаграммах расщепления хромосомной ДНК соответствует аналогичным пикам, образуемым при расщеплении множества LINE1 повторов ДНК мыши по тем же нуклеотидным последовательностям. Получены экспериментальные картины гидролиза ДНК мыши соответствующими эндонуклеазами рестрикции. Проведено сравнение теоретически рассчитанных диаграмм распределения фрагментов и полученных картин гидролиза ДНК эндонуклеазами рестрикции. Показано, что при анализе препаратов ДНК с помощью электрофореза в геле визуализируются только продукты расщепления LINE1-повторов и γ -сателлитной ДНК мыши. Проведены эксперименты по гидролизу хромосомной ДНК мыши 5-метилцитозин-зависимыми сайт-специфическими эндонуклеазами BlnI, GlnI и GlnII.

Ключевые слова: ДНК-повторы, рестрикционный анализ, метилзависимые эндонуклеазы, эукариотический геном, *Mus musculus*.

В настоящее время первичная структура эухроматиновой части генома мыши определена более чем на 96% [1], что позволяет анализировать структуру ДНК *in silico*.

Ранее нами был предложен метод проведения рестрикционного анализа ДНК млекопитающих *in silico* путем построения диаграмм распределения фрагментов, получаемых при расщеплении хромосомной ДНК по сайтам узнавания эндонуклеаз рестрикции (ЭР) [2]. Для ряда последовательностей ДНК (5'-CCWGG-3', 5'-GATC-3', 5'-GGCC-3' и 5'-CCGG-3') были получены такие диаграммы расщепления хромосомной ДНК крысы, мыши и человека и проведено сравнение теоретических расчетов с экспериментальными результатами по гидролизу хромосомных ДНК эндонуклеазами рестрикции Bst2UI, Kzo9I, HaeIII и MspI, имеющими эти сайты узнавания. В дальнейшем нами был проведен сравнительный рестрикционный анализ расщепления

геномной ДНК крысы [3] и человека [4] и показано хорошее соответствие картин рестрикции *in vitro* и *in silico*.

Целью настоящей работы явилось получение диаграмм распределения фрагментов при расщеплении хромосомной ДНК мыши по 17 сайтам узнавания, включая 4-, 5- и 6-нуклеотидные последовательности и сравнение полученных данных с результатами гидролиза ДНК соответствующими эндонуклеазами рестрикции.

Материалы и методы

Хромосомная ДНК мыши. В экспериментах использовали мышей-самцов линии А/Не в возрасте 5–6 месяцев разведения вивария ИЦиГ СО РАН. Геномная ДНК из печени животных выделялась, как описано ранее [3].

Гидролиз хромосомной ДНК. В работе использовали следующие сайт-специфические эндонуклеазы производства НПО «СибЭнзим» (в скобках указан сайт узнавания соответствующей эндонуклеазы):

AcsI (5'-RAATTY-3'), AspA2I (5'-CCTAGG-3'), AspS9I (5'-GGNCC-3'), BglII (5'-AGATCT-3'), BlnI (5'-GmCNGmC-3'), Bme18I (5'-GGWCC-3'),

* Автор для переписки

© 2007 г. Чернухин Валерий Алексеевич

НПО «СибЭнзим»

630117 Новосибирск, ул. Акад. Тимакова, 2/12

Тел./факс: 8-383-333-49-91

E-mail: valera@sibenzyme.ru

Bse3DI (5'-GCAATG-3' и 5'-CATTCG-3'), Bsp19I (5'-CCATGG-3'), Bst2UI (5'-CCWGG-3'), BstFNI (5'-CGCG-3'), BstSCI (5'-CCNGG-3'), BstX2I (5'-RGATCY-3'), EcoRI (5'-GAATTC-3'), EcoRV (5'-GATATC-3'), Fsp4HI (5'-GCNGC-3'), GluI (5'-GmCGmC-3'), GluI (5'-GmCNGmC-3'), HpaII (5'-CCGG-3'), HspAI (5'-GCGC-3'), MspI (5'-CCGG-3'), PspN4I (5'-GGNNCC-3'), RsaI (5'-GTAC-3'), Sse9I (5'-AATT-3'), SspI (5'-AATATT-3') и TaqI (5'-TCGA-3').

ДНК в количестве 6 мкг расщепляли в 40 мкл реакционной смеси в рекомендуемых производителем SE-буферах для рестрикции ДНК при оптимальных температурах в течение 3 ч.

Электрофорез. Для выявления фрагментов ДНК длиной от 40 до 500 п.н. использовали электрофорез в 8% полиакриламидном геле (ПААГ) в камере Hoefer SE 600/SE 660 (Amersham USA). На гель наносили 6 мкг гидролизованной ДНК на дорожку. Для выявления фрагментов в диапазоне 200–2000 п.н. использовали 1,5% агарозу «Low melting point» («Sigma», USA). На гель наносили 3 мкг гидролизованной ДНК на дорожку.

Для выявления фрагментов ДНК в диапазоне 500–20000 п.н. применяли электрофорез в 1%-й агарозе «Type I-A, Low EEO» (АГ), («Sigma», USA), на гель наносили 3 мкг гидролизованной ДНК на дорожку.

Для разделения фрагментов в диапазоне от 10 до 50 т.п.н. использовали импульсный электрофорез в 1%-ном растворе агарозного геля [5]. На гель наносили 1 мкг ДНК.

В качестве маркера молекулярных масс ДНК использовали 50 kb ДНК маркер, 1 kb ДНК-маркер и pUC19/MspI ДНК маркер производства НПО «СибЭнзим».

Во всех случаях для электрофореза применяли трис-ацетатный буфер (40 mM Трис-ацетат, pH 8,0; 1 mM ЭДТА). После проведения электрофореза ДНК визуализировали бромистым этидием и фотографировали в УФ-свете.

Данные о нуклеотидной последовательности генома мыши. Последовательность ДНК мыши была получена с ресурса <ftp://ftp.ensembl.org/pub/> (версия от 2 июня 2006 года).

Нуклеотидная последовательность основного мономерного фрагмента γ -сателлитной ДНК мыши взята из [6].

Последовательности LINE1-повторов были извлечены из базы данных геномной ДНК мыши с помощью сервиса Table Browser на сайте (<http://genome.ucsc.edu/cgi-bin/hgTables>) [8].

Полученная выборка LINE1-повторов включала в себя 854172 последовательности с общей длиной ≈ 510 млн. п.н.

Программное обеспечение. Построение диаграмм распределения фрагментов, получаемых при расщеплении ДНК по сайтам узнавания эндонуклеаз рестрикции, в виде зависимости суммарной массы фрагментов фиксированной длины, выраженной в парах оснований, от длины фрагментов (в парах нуклеотидов) проводили по методике, описанной ранее [2].

Результаты и обсуждение

Анализ диаграмм распределения фрагментов ДНК. Ранее нами были получены диаграммы распределения фрагментов при расщеплении хромосомной ДНК мыши по сайтам узнавания рестриктаз HaeIII (5'-GGCC-3'), MspI (5'-CCGG-3'), Kzo9I (5'-GATC-3') и Bst2UI (5'-CCWGG-3') и показано хорошее соответствие между расчетными данными и результатами экспериментов по гидролизу ДНК этими ферментами [2]. В настоящей работе мы провели построение диаграмм распределения фрагментов, получаемых при расщеплении хромосомной ДНК мыши по более широкому спектру сайтов узнавания, включая и шестинуклеотидные последовательности. В соответствии с предложенным ранее методом [2], диаграммы распределения строились с использованием арктангенсоидной шкалы, имитирующей разделение получаемых фрагментов электрофорезом в агарозном геле. Были отобраны только те диаграммы расщепления, которые содержат пики ДНК фрагментов более 5,5 млн. пар оснований. Как было показано в предыдущей работе [2], именно такие пики визуализируются в эксперименте при анализе продуктов расщепления ДНК электрофорезом в агарозном геле. При этом на визуализацию пиковых фрагментов ДНК может оказывать влияние фон, образованный расщеплением основной массы ДНК и способный частично маскировать эти фрагменты [2].

Всего нами было проанализировано расщепление по 18 сайтам узнавания ЭР, приводящее к образованию ДНК фрагментов с высоким значением пиков.

На рисунке 1 приведены диаграммы распределения фрагментов геномной ДНК мыши, полученные для этих последовательностей узнавания.

Максимальные значения высоты пиков (более 10 млн. п.н.) наблюдаются при расщеплении ДНК по сайтам узнавания ферментов Bse3DI (фрагменты длиной 1059 и 2590 п.н.), Bst2UI и BstSCI (фрагмент длиной

1826 п.н.), Fsp4HI (фрагменты длиной 695 и 1613 п.н.), PctI (фрагмент длиной 2859 п.н.) и EcoRI (фрагмент длиной 1373 п.н.). Визуализация последнего фрагмента была описана в литературе ранее [7]. Полосы ДНК, соответствующие этим фрагментам, хорошо видны на электрофореграммах разделения продуктов гидролиза ДНК данными рестриктазами, которые приведены на рисунке 1 над соответствующими диаграммами распределения.

Остальные фрагменты меньшей интенсивности, отмеченные на диаграммах на рисунке 1, также представлены в виде полос на соответствующих фотографиях гелей при гидролизе ДНК ферментами AspA2I (1273, 3343-3344 и 4613 п.н.), BglII (876 п.н.), Bsp19I (3253 п.н.), EcoRV (1883 п.н.), Fsp4HI (583 и 1035) и PspN4I (444, 508, 716 и 1721 п.н.). При гидролизе ДНК ферментом Bme18I на электрофореграмме видны фрагмент 1619 п.н. и дубль фрагментов 809 и 818 п.н. Фрагмент длиной 708 п.н., по-видимому, сливается с фрагментом γ -сателлитной ДНК, имеющим длину около 705 п.н. [6].

Электрофореграммы гидролиза ДНК остальными ферментами имеют значительные участки фоновых фрагментов ДНК, которые маскируют полосы ДНК, соответствующие расчетным данным. В частности, при гидролизе ДНК ферментами Bst2UI и BstSCI маскируются все полосы ДНК, кроме соответствующих фрагментам 1511 и 1826 п.н. При расщеплении ДНК ферментом BstX2I на геле видны полосы, соответствующие фрагментам ДНК длиной 203, 397, 508, 514 и 713 п.о., но не видно фрагмента длиной 1856 п.н., который, по-видимому, маскируется фоном. Аналогичная картина наблюдается с ферментом RsaI: на электрофореграмме в ПААГ не визуализируются пиковые фрагменты длиной больше 503 п.н.

Как было показано нами ранее [3], явление клас-теризации близко лежащих фрагментов ДНК позволяет в ряде случаев визуализировать пиковые фрагменты с высотой менее 5,5 млн. п.н. На рисунке 1 приведена картина гидролиза ДНК ферментом SspI, на которой видна полоса ДНК, соответствующая фрагменту длиной 1037 п.н. Однако данный фрагмент в диаграмме расщепления имеет пиковое значение только 3,7 млн. п.н., что ниже вышеуказанной критической высоты пика. Визуализация соответствующей полосы на электрофореграмме связана с наличием сразу нескольких фрагментов, близких по размерам, но существенно меньшей интенсивности, что хорошо видно на более детальной диаграмме распределения фрагментов хромосомной ДНК (рис. 2).

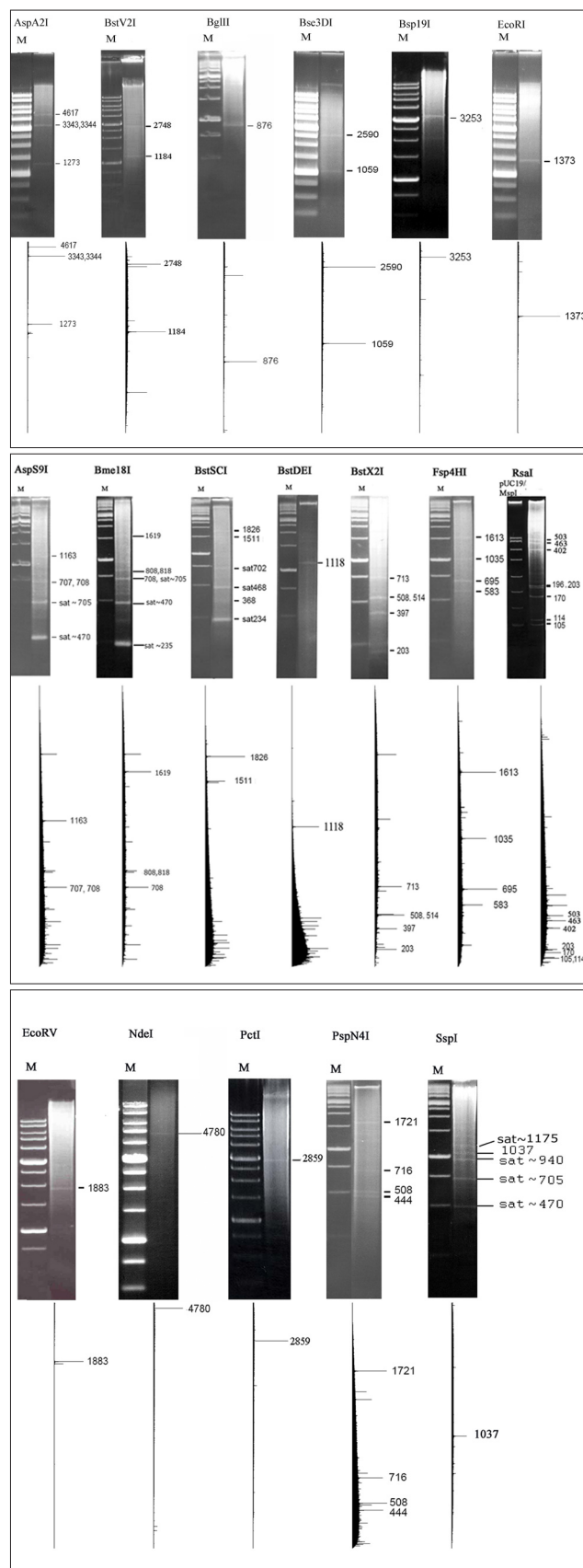


Рис. 1. Сравнение теоретически рассчитанных диаграмм распределения фрагментов геномной ДНК и электрофореграмм, полученных при гидролизе ДНК

соответствующими эндонуклеазами рестрикции. Электрофорез продуктов гидролиза ДНК ферментами AspA2I, BstV2I, BglII, Bse3DI, Bsp19I, EcoRI, EcoRV, NdeI, PstI проводился в 1% агарозном геле, ферментами PspN4I, SspI, AspS9I, Bme18I, BstSCI, BstDEI, BstX2I, Fsp4HI) — в 1,5% агарозном геле и RsaI — в ПААГ. Название фермента указано над фотографией геля. М — 1 kb маркер молекулярных масс ДНК.

На диаграммах числами указаны размеры пиковых фрагментов, образуемых при сайт-специфическом расщеплении хромосомной ДНК мыши (п.н.). Высота пика соответствует количеству нуклеотидов во всех фрагментах данной длины (п.н.)

Как видно из рисунка 1, в случае некоторых ЭР (AspS9I, Bme18I, Bst2UI, BstSCI и SspI) на электрофореграммах хорошо заметны полосы ДНК, которые не соответствуют никаким пиковым фрагментам, представленным на соответствующих диаграммах и, по-видимому, являются продуктами гидролиза γ -сателлитной ДНК мыши [6], не представленной в базе данных геномной ДНК [1].

На рисунке 1 также видно, что яркость этих полос ДНК на фотографии гелей существенно выше свечения фрагментов, получаемых при расщеплении геномной ДНК. Это связано с высоким количеством повторов в сателлитной ДНК и более высокой долей этой ДНК в суммарном препарате ДНК мыши по сравнению с хромосомной ДНК крысы и человека [3, 4]. По некоторым оценкам, доля γ -сателлитной ДНК мыши составляет примерно 10% от общего количества хромосомной ДНК [6].

При гидролизе хромосомной ДНК ферментами AspS9I, Bme18I, BstSCI и Bst2UI происходит выщепление фрагмента, электрофоретическая подвижность которого соответствует длине так называемого основного фрагмента γ -сателлитной ДНК мыши длиной 235 п.н. [6] (рис. 3).

Сайты узнавания этих ферментов представлены в нуклеотидной последовательности основного мономерного фрагмента в позициях 1, 1, 4 и 4, соответственно, и гидролиз tandemных повторов сателлитной ДНК по этим позициям приводит к образованию фрагмента 235 п.н. [6]. Расщепления γ -сателлитной ДНК мыши с образованием такого фрагмента наблюдается также при гидролизе ферментами BstF5I (5'-GGATG-3', позиция 166), FstI (5'-CATG-3', позиция 150), Sse9I (5'-AATT-3', позиция 98), AcsI (5'-RAATTY-3', позиция 97). Кроме того, как было описано в литературе [6], в результате гид-

ролиза сателлитной ДНК образуются также фрагменты кратной длины (≈ 470 п.н., ≈ 705 п.н. и т.д.), состоящие из двух или нескольких мономерных фрагментов.

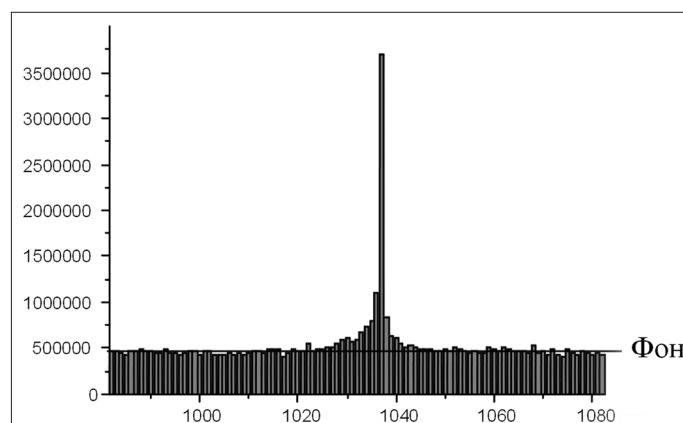


Рис. 2. Диаграмма распределения фрагментов, получаемых при расщеплении геномной ДНК мыши по последовательности 5'-ААТАТТ-3'.

Обозначения на диаграмме:

ось X — длины фрагментов, образуемых при сайт-специфическом расщеплении хромосомной ДНК мыши (п.н.);

ось Y — величины пиков, образуемых при сайт-специфическом расщеплении хромосомной ДНК мыши (п.н.)

На электрофореграмме на рисунке 3 практически на всех дорожках видны такие димеры, тримеры и даже тетрамеры основного фрагмента сателлитной ДНК. Наличие мультимерных фрагментов сателлитной ДНК, очевидно, связано с высокой вариабельностью первичной структуры этой ДНК и изменением нуклеотидной последовательности в сайтах узнавания рестриктаз. Из-за этого не в каждом повторе происходит гидролиз ДНК по соответствующим позициям и, как результат, образуются димеры, тримеры и т.д. При этом для разных ферментов наблюдается различное распределение мультимерных фрагментов по длинам, что, по-видимому, связано с размерами сайта узнавания фермента. В частности, фермент AcsI имеет расширенный сайт узнавания фермента Sse9I и количество тримеров и тетрамеров, наблюдаемое в случае гидролиза ДНК ферментом AcsI, больше, а мономера — меньше, чем в случае расщепления ДНК ферментом Sse9I. На картине гидролиза ДНК ферментом SspI, который узнает шестинуклеотидный невырожденный сайт узнавания, видны фрагменты, соответствующие по длине димеру (470 п.н.), тримеру (705 п.н.), тетрамеру (940 п.н.) и даже пентамеру (1175 п.н.), но почти не заметна полоса, соответствующая основному мономерному

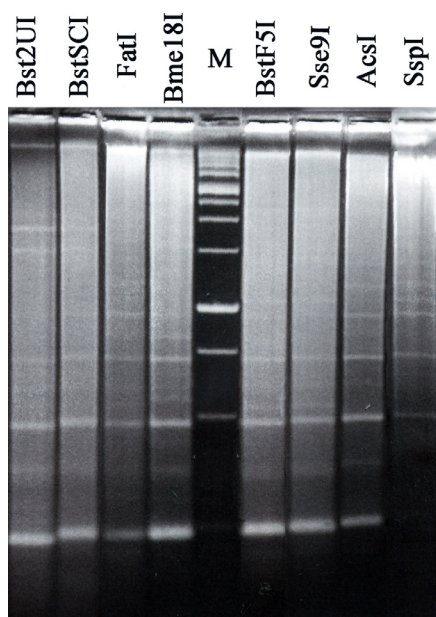


Рис. 3. Гидролиз ДНК мыши эндонуклеазами рестрикции. Название используемых ферментов указано над дорожкой геля. Электрофорез в 1,5% агарозе в ТАЕ-буфере. М — 1 kb маркер молекулярных масс ДНК

фрагменту γ -сателлитной ДНК. Следует также отметить слабый гидролиз как сателлитной, так и геномной ДНК ферментом *FatI*. Видимые на фотографии геля полосы мономерного и полимерных фрагментов сателлитной ДНК, образуемые при обработке ДНК ферментом *FatI*, существенно слабее по сравнению с картиной расщепления ДНК другими мелкоцепляющими ферментами, что говорит о значительной мутационной изменчивости узнаваемой последовательности 5'-CATG-3'.

Таким образом, часть фрагментов ДНК, представленных на электрофореграммах более яркими полосами, образуется в результате расщепления γ -сателлитной ДНК мыши.

Меньшая по интенсивности, но значительно более разнообразная часть рестрикционных фрагментов получается при расщеплении эухроматиновой части ДНК мыши, состоящей из рассеянных повторов. Короткие и длинные диспергированные повторы занимают у мыши 39% генома [10]. Наиболее многочисленными повторами у мыши являются длинные диспергированные повторы (LINE), основная часть которых представлена семейством L1-повторов [1], обнаруженных к настоящему времени в большинстве отрядов млекопитающих [9]. Изучаемая нами выборка LINE1-повторов включала в себя более 854 тысяч последовательностей с общей длиной более 500 млн. п.н., что составляет приблизительно 1/5 часть генома мыши [1]. Полноразмерные LINE1-

повторы имеют длину 6–7 т.п.н., однако большинство повторов этого класса представлено укороченными копиями [11] и лишь 5713 имели длину более 6 т.п.н. Ввиду относительно большой длины LINE1 повторов, размеры фрагментов ДНК, образованных при их гидролизе эндонуклеазами рестрикции, существенно различаются.

Отдельные семейства коротких рассеянных повторов (SINE) составляют у мыши в несколько раз меньшую часть генома по сравнению с L1-повторами [1]. В связи с тем, что в выбранных нами условиях анализа продукты расщепления SINE повторов не визуализируются, мы не рассматривали гидролиз SINE повторов в данной работе.

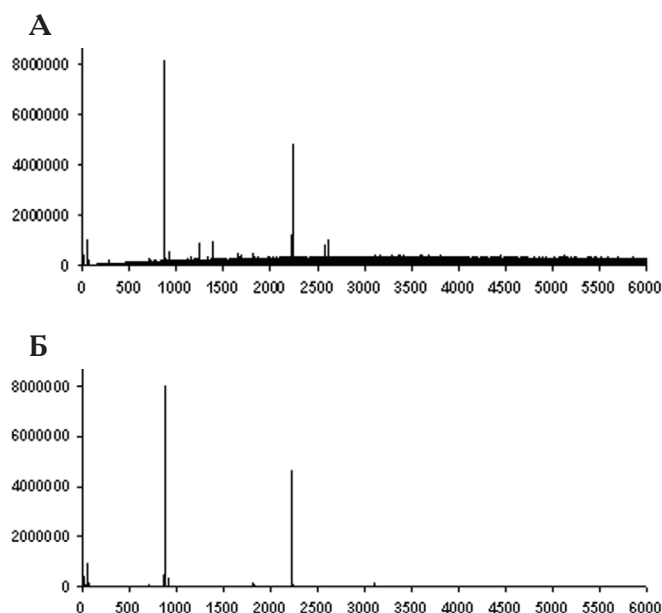
Мы провели сравнение полученных выше диаграмм распределения фрагментов геномной ДНК (см. рис. 1) и диаграмм, рассчитанных для выборки L1-повторов (см. «Материалы и методы»). На рисунке 4 приведены оба типа диаграмм для сайтов узнавания рестриктаз *BglII* и *Asp9I*. Как видно на данном рисунке, все пиковые фрагменты, присутствующие в диаграммах расщепления геномной ДНК, также представлены на диаграммах расщепления L1-повторов.

В большинстве случаев величины пиков фрагментов ДНК, рассчитанные для геномной ДНК, не отличаются от величин пиков, рассчитанных для расщепления выборки L1-повторов (с учетом фона, получаемого при расщеплении геномной ДНК). Так, например, как видно из рисунка 4, высота пикового фрагмента длиной 876 п.н., полученного при расщеплении ДНК по сайту узнавания ЭР *BglII* (5'-AGATCT-3'), практически совпадает в этих двух диаграммах.

Однако в ряде случаев высоты пиковых фрагментов, рассчитанных для L1-повторов, существенно меньше по сравнению со значениями высот пиковых фрагментов на диаграммах, рассчитанных для геномной ДНК. Так, например, при расщеплении по сайту GGCC (сайт узнавания *Asp9I*) фрагмент длиной 708 п.н. имеет значение 8,0 млн. п.н. на диаграмме расщепления геномной ДНК и только 3,6 млн. п.н. — на диаграмме расщепления L1-повторов (рис. 4). Данное расхождение может быть следствием недостаточной представительности выборки L1-повторов.

Влияние метилирования на гидролиз хромосомной ДНК сайт-специфическими эндонуклеазами. Одной из основных причин, влияющих на эффективность гидролиза хромосомной ДНК млекопитающих рядом эндонуклеаз рестрикции, является блокирование расщепления ДНК из-за метилирования цитозина в динуклотидах CG. В соматических клетках млекопитающих уровень

Сайт узнавания BglII (5'-AGATCT-3')



Сайт узнавания AspS9I (5'-GGNCC-3')

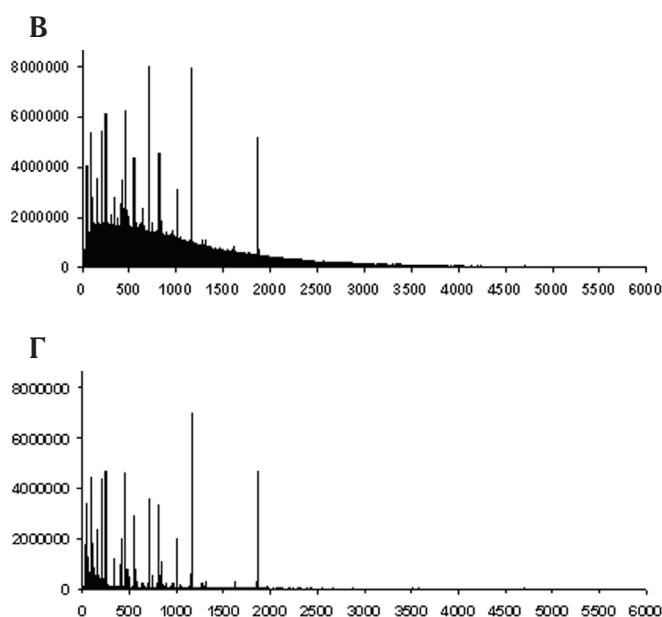


Рис. 4. Сравнение теоретически рассчитанных диаграмм распределения фрагментов, получаемых после расщепления геномной ДНК (А, В) и выборки LINE1-повторов (Б, Г).

Обозначения на диаграммах:

ось X — длины фрагментов, образуемых при сайт-специфическом расщеплении геномной ДНК или L1-повторов мыши (п.н.) ферментами BglII (А, Б) и AspS9I (В, Г);

ось Y — величины пиков, образуемых при сайт-специфическом расщеплении геномной ДНК или L1-повторов мыши (п.н.).

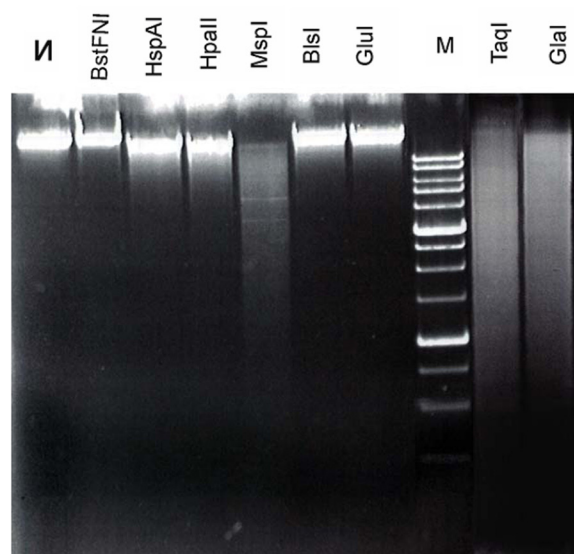
метилирования составляет примерно 70–80% всех CG-пар [12]. Поэтому гидролиз ДНК эндонуклеазами рестрикции, которые содержат в сайте узнавания CG-последовательность, может быть ограничен. Кроме того, как показано для геномов ряда млекопитающих и, в том числе, мыши, доля CG-динуклеотидов в ДНК примерно в 5 раз ниже по сравнению с частотой встречаемости других динуклеотидных последовательностей, состоящих только из нуклеотидов G и C [13, 14]. Поэтому другой причиной, обуславливающей небольшую глубину гидролиза ДНК такими рестриктазами, является более низкая частота встречаемости сайтов узнавания, содержащих CG-динуклеотид [3].

На рисунке 5 показаны результаты расщепления хромосомной ДНК мыши эндонуклеазами рестрикции, у которых в сайте узнавания есть последовательность CG. На рисунке 5А приведена электрофореграмма разделения продуктов гидролиза ДНК, полученная в 1%-ном растворе агарозного геля при постоянном напряжении, а на рисунке 5Б — при переменном напряжении электрического поля (импульсный электрофорез).

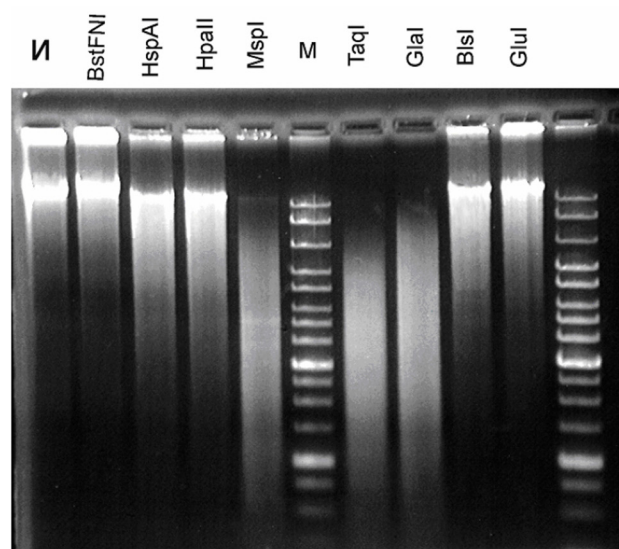
Как видно из рисунка 5А, HpaII совершенно незначительно расщепляет ДНК, хотя MspI (другой фермент с этим же сайтом узнавания CCGG, но не чувствительный к CG-метилированию) эффективно расщепляет ДНК мыши. Также метилированием CG-динуклеотидов в геномной ДНК можно объяснить незначительный гидролиз ДНК ферментами BstFNI (сайт узнавания CCGG) и HspAI (сайт узнавания GCGC).

Использование импульсного гель-электрофореза позволяет разделять фрагменты ДНК в диапазоне до 50000 п.н. [5]. В этом случае, как видно из рисунка 5Б, гидролиз хромосомной ДНК мыши рестриктазами HpaII и HspAI демонстрируется гораздо четче. При расщеплении этими рестриктазами большинство получаемых фрагментов ДНК имеет размеры в диапазоне от 10 до 50 т.п.н.

В случае ЭР BstFNI расщепление хромосомной ДНК даже при использовании импульсного электрофореза практически не визуализируется. Причина этого, по-видимому, заключается в том, что, в отличие от HpaII и HspAI, в сайте узнавания BstFNI содержится сразу два динуклеотида CG, что обуславливает более низкую частоту сайтов узнавания BstFNI по сравнению с последовательностями узнавания HpaII и HspAI. Кроме того, из-за наличия двух CG-динуклеотидов доля сайтов узнавания BstFNI, содержащих метилированный цитозин, существенно выше, по сравнению с сайтами узнавания HpaII и HspAI.



А



Б

Рис. 5. Гидролиз ДНК мышцы эндонуклеазами рестрикции, в сайте узнавания которых есть CG динуклеотид, и метилзависимыми сайт-специфичными эндонуклеазами. Название фермента указано над дорожкой на фотографии геля.

А — электрофорез в АГ при постоянном напряжении электрического поля (М — 1 kb маркер молекулярных масс ДНК).

Б — импульсный электрофорез в АГ (М — 50 kb маркер молекулярных масс ДНК).

И — исходная ДНК

В последние годы появился новый инструмент для анализа статуса метилирования ДНК — сайт-специфичные метилзависимые эндонуклеазы, которые расщепляют ДНК только при наличии в сайте узнавания 5-метилцитозина [15–18].

К таким ферментам относится ClaI, эффективно гидролизующий целый ряд различных метилированных четырехнуклеотидных последовательностей ДНК, содержащих от двух до четырех 5-метилцитозина [16, 19].

Кроме того, описаны метилспецифичные ферменты, которые эффективно гидролизуют последовательность ДНК GCNGC с двумя-четырьмя 5-метилцитозинами (BlnI [17]) или только с четырьмя 5-метилцитозинами (GluI [18]). Результаты расщепления этими тремя ферментами ДНК мышцы приведены на рисунке 4.

На рисунке 4А представлена электрофореграмма анализа продуктов гидролиза, полученная в стандартных условиях разделения фрагментов ДНК. Как видно из этого рисунка, эффективный гидролиз хромосомной ДНК наблюдается только в случае ClaI; при этом глубина ДНК близка к таковой для ферментов MspI и TaqI.

В случае BlnI и GluI гидролиз на приведенной электрофореграмме практически не визуализируется, что связано с низкой частотой встречаемости метилированных сайтов узнавания этих ферментов на хромосомной ДНК. Однако при использовании импульсного электрофореза (рис. 4Б) видно, что BlnI, в отличие от GluI, все-таки гидролизует хромосомную ДНК с образованием высокомолекулярных фрагментов.

Фермент BlnI способен расщеплять метилированные последовательности CGCGGC или GCCGCG, представляющие собой два варианта расширенного на один нуклеотид сайта узнавания (последний подчеркнут). При этом в сайте узнавания появляется два метилированных цитозина. GluI расщепляет последовательность GCNGC только при наличии четырех цитозина и соответственно не способен гидролизовать подобные сайты. Таким образом, в отличие от BlnI, гидролиз ДНК ферментом GluI не возможен из-за недостатка 5-метилцитозина в сайте узнавания.

Заключение

В работе проведен рестрикционный анализ ДНК мышцы *in silico* для широкого спектра последовательностей узнавания, а также получены экспериментальные картины расщепления ДНК соответствующими эндонуклеазами рестрикции. Обнаружено, что экспериментально наблюдаемые фрагменты ДНК являются продуктами расщепления LINE1-повторов и γ -сателлитной ДНК мышцы. Теоретически показано существование более чем пятидесяти пиковых значений фрагментов (или кластеров фрагментов) ДНК, получаемых при расщеплении геном-

ной ДНК мыши по 18 нуклеотидным последовательностям, являющимися сайтами узнавания эндонуклеаз рестрикции.

Установлено, что получаемые диаграммы расщепления геномной ДНК мыши в целом соответствуют диаграммам расщепления выборки соответствующих LINE1-повторов, однако для ряда сайтов наблюдается несоответствие величины пиков, получаемых во втором случае.

Приведенное в работе сравнение рассчитанных диаграмм распределения фрагментов ДНК мыши и полученных картин гидролиза ДНК соответствующими ферментами рестрикции показывает хорошее соответствие теоретических и экспериментальных данных.

В ходе электрофоретического анализа продуктов гидролиза хромосомной ДНК мыши метилзависимыми эндонуклеазами выявлено, что наблюдается существенное расщепление ДНК ферментом *GlaI*: значительно более слабое ферментом *BlsI*, а *GluI* практически не расщепляет ДНК.

Авторы благодарят к.б.н. В.И. Каледина за помощь в работе с животными и к.б.н. Г.В. Васильева за помощь в выделении ДНК.

Литература

1. Mouse Genome Sequencing Consortium. Initial sequencing and comparative analysis of the mouse genome // *Nature*. — 2002. — Vol. 420. — P. 520–562.
2. Абдурашитов М.А., Томилов В.Н., Чернухин В.А., Гончар Д.А., Дегтярев С.Х. Метод рестрикционного анализа геномов млекопитающих *in silico* // Вестник биотехнологии и физико-химической биологии имени Ю.А. Овчинникова. — 2006. — Т. 2. — № 3. — С. 29–38.
3. Чернухин В.А., Абдурашитов М.А., Томилов В.Н., Гончар Д.А., Дегтярев С.Х. Сравнительный рестрикционный анализ хромосомной ДНК крысы *in vitro* и *in silico* // Вестник биотехнологии и физико-химической биологии имени Ю.А. Овчинникова. — 2006. — Т. 2. — № 3. — С. 39–46.
4. Абдурашитов М.А., Чернухин В.А., Томилов В.Н., Гончар Д.А., Дегтярев С.Х. Рестрикционный анализ повторяющихся последовательностей ДНК человека // Медицинская генетика. — 2007. — Т. 6. — № 8. — С. 29–36.
5. Зернов Ю.П., Перминова Л.В., Пустошилова Н.М. Разделение двуниевых молекул ДНК электрофорезом в агарозе с использованием пульсирующего поля // Известия СО АН СССР, серия биологическая. — 1989. — Вып. 2. — С. 104–107.
6. Horz W., Altenburger W. Nucleotide sequence of mouse satellite DNA // *Nucleic Acids Res.* — 1981. — Vol. 9. — P. 683–696.
7. Cheng S.-M., Schildkraut C.L. A family of moderately repetitive sequences in mouse DNA // *Nucleic Acids Res.* — 1980. — Vol. 8. — P. 4075–4090.
8. Karolchik D., Hinrichs A.S., Furey T.S., Roskin K.M., Sugnet C.W., Haussler D., Kent W.J. The UCSC Table Browser data retrieval tool // *Nucleic Acids Res.* — 2004. — Vol. 32 (Suppl. 1). — D493–D496.
9. Hutchison C.A. III, Hardies S.C., Loeb D.D., Shehee W.R., Edgell M.H. LINEs and related retrotransposons: Long interspersed repeated sequences in the eukaryotic genome / In: *Mobile DNA* (Eds. Berg D.E., Howe M.M.). — American Society for Microbiology, Washington. — 1989. — P. 593–617.
10. Shonbach Ch. From masking repeats to identifying to functional repeats in the mouse transcriptome // *Briefings in bioinformatics*. — 2004. — Vol. 5. — N 2. — P. 107–117.
11. Wincker P., Jubier-Maurin V., Roizes G. Unrelated sequences at 5'-end of the mouse LINE-1 repeated elements define two distinct subfamilies // *Nucleic Acids Res.* — 1987. — Vol. 15. — P. 8593–8606.
12. Ehrlich M., Gama S.M., Huang L.H., Midgett R.M., Kuo K.C., McCune R.A., Gehrke C. Amount and distribution of 5-methylcytosine in human DNA from different types of tissues or cells // *Nucleic Acids Res.* — 1982. — Vol. 10. — P. 2709–2721.
13. Ohno S. Universal rule for coding sequence construction: TA/CG deficiency-TG/CT excess // *Proc. Natl. Acad. Sci. USA*. — 1988. — Vol. 85. — P. 9630–9634.
14. Yomo T., Ohno S. Concordant evolution of coding and noncoding regions of DNA made possible by the universal rule of TA/CG deficiency-TG/CT excess // *Proc. Natl. Acad. Sci. USA*. — 1989. — Vol. 86. — P. 8452–8456.
15. Чернухин В.А., Наякишина Т.Н., Абдурашитов М.А., Томилова Ю.Э., Мезенцева Н.В., Дедков В.С., Михненко Н.А., Гончар Д.А., Дегтярев С.Х. Новая эндонуклеаза рестрикции *GlaI* узнает метилированную последовательность 5'-GCGC-3' // Биотехнология. — 2006. — № 4. — С. 31–35.
16. Томилова Ю.Э., Чернухин В.А., Дегтярев С.Х. Зависимость активности сайт-специфической эндонуклеазы *GlaI* от количества и положения метилированных цитозинов в узнаваемой последовательности 5'-GCGC-3' // Вестник биотехнологии и физико-химической биологии имени Ю.А. Овчинникова. — 2006. — Т. 2. — № 1. — С. 30–39.
17. Чернухин В.А., Томилова Ю.Э., Чмуж Е.В., Соколова О.О., Дедков В.С., Дегтярев С.Х. Сайт-специфическая эндонуклеаза *BlsI* узнает последовательности ДНК 5'-G(5mC)NGC-3' и расщепляет ее с образованием 3'-выступающих концов // Вестник биотехнологии и

- физико-химической биологии имени Ю.А. Овчинникова. — 2007. — Т. 3. — № 1. — С. 28–33.
18. Чернухин В.А., Чмуж Е.В., Томилова Ю.Э., Наякина Т.Н., Гончар Д.А., Дедков В.С., Дегтярев С.Х. Новая сайт-специфическая эндонуклеаза GluI узнает метилированную последовательность ДНК 5'-G(5mC)NG(5mC)-3'/3'-(5mC)GN(5mC)G-5' // Вестник биотехнологии и физико-химической биологии имени Ю.А. Овчинникова. — 2007. — Т. 3. — № 2. — С. 13–17.
19. Tarasova G.V., Nayakshina T.N., Degtyarev S.Kh. Substrate specificity on new methyl-directed DNA endonuclease GluI // BMC Molecular Biology. — 2008. — Vol. 9. — P. 7.

Список сокращений

ЭР — эндонуклеаза рестрикции, рестриктаза;

ПААГ — полиакриламидный гель;

п.н. — пары нуклеотидов,

т.п.н. — тысяча пар нуклеотидов.

COMPARATIVE ANALYSIS OF MOUSE CHROMOSOMAL DNA DIGESTION WITH RESTRICTION ENDONUCLEASES IN VITRO AND IN SILICO

V.A. CHERNUKHIN, M.A. ABDURASHITOV, V.N. TOMILOV,
D.A. GONCHAR, S.Kh. DEGTYAREV

SibEnzyme SPA, Novosibirsk

Theoretical diagrams of mouse chromosomal DNA cleavage at 18 nucleotide sequences 4-6 bp in length, which are the recognition sites of restriction endonucleases, have been plotted based on earlier suggested method of mammalian genomes restriction analysis in silico. Analysis of mouse LINE 1 repeats, presented in database, and products of these repeats cleavage at the same nucleotide sequences has been carried out. In general, the diagrams of mouse chromosomal DNA digestion correspond to diagrams of LINE1-repeats cleavage. Mouse chromosomal DNA hydrolysis with restriction endonucleases, which possess the corresponding recognition sites, has been performed. A comparison of DNA hydrolysis patterns and the plotted diagrams has revealed a good correspondence between the experimental and theoretical data. Only LINE1 repeats and satellite DNA cleavage products are visualized in experiments on chromosomal DNA cleavage with subsequent gel-electrophoresis. Mouse chromosomal DNA cleavage with new methyl-dependent site-specific DNA endonucleases BslI, GluI and GluI has been performed.

Keywords: DNA repeats, restriction analysis, methyl-dependent endonucleases, eukaryotic genome, *Mus musculus*.