
**ОБЗОРНЫЕ
И ТЕОРЕТИЧЕСКИЕ СТАТЬИ**

УДК 577.2.08

ПРИМЕНЕНИЕ САЙТ-СПЕЦИФИЧЕСКИХ ДНК-ЭНДОНУКЛЕАЗ В ПОЛНОГЕНОМНОМ АНАЛИЗЕ ДНК ЧЕЛОВЕКА

© 2017 г. М. А. Абдурашитов¹, С. Х. Дегтярев^{1, 2, *}

¹ООО “СибЭнзим”, Новосибирск 630117

²Государственный научный центр вирусологии и биотехнологии “Вектор”,
пос. Кольцово Новосибирской области 630559

*e-mail: degt@sibenzyme.com

Поступила в редакцию 02.02.2016 г.

На протяжении последних десятилетий сайт-специфические ДНК-эндонуклеазы служат одним из ключевых инструментов для изучения первичной структуры ДНК и генно-инженерных манипуляций. В обзоре рассмотрены примеры использования этих ферментов при полногеномном анализе ДНК человека, включающие применение эндонуклеаз рестрикции в пробоподготовке при секвенировании на NGS-машинах, а также визуализацию расщепления ДНК-повторов рестриктазами. Особый интерес представляют первые работы по применению ДНК-эндонуклеаз в быстро развивающейся области эпигенетического анализа геномов, чему способствует недавнее открытие нового класса ферментов — 5-метилцитозин-зависимых сайт-специфических ДНК-эндонуклеаз.

Ключевые слова: сайт-специфические эндонуклеазы, геномика, эпигеномика, метилирование ДНК, секвенирование.

DOI: 10.7868/S0016675817010027

Сайт-специфические ДНК-эндонуклеазы — это ферменты, расщепляющие ДНК по определенным последовательностям, называемым сайтами узнавания. Наиболее известными среди этих ферментов являются эндонуклеазы рестрикции II типа, первая из которых (*HindII*) была обнаружена 45 лет назад [1] и которые сыграли большую роль в развитии генной инженерии [2]. Использование эндонуклеаз рестрикции позволило проводить фрагментацию протяженных геномных ДНК и клонировать интересующие фрагменты генома в различные векторные конструкции. Несмотря на то что развитие ПЦР-технологий и методов полногеномного секвенирования несколько снизило интерес к использованию ферментов рестрикции, их значимость для научных исследований остается высокой. В последние годы наряду с эндонуклеазами рестрикции успешное применение в исследованиях и практическом использовании нашли и другие сайт-специфические ДНК-эндонуклеазы, такие как ДНК-никазы [3], интрон-кодируемые нуклеазы [4], ферменты TALEN [5], CRISPR-системы [6] и, наконец, метилзависимые сайт-специфические ДНК-эндонуклеазы [7]. Однако пока только эндонуклеазы рестрикции и метилзависимые сайт-специфические ДНК-эндонуклеазы нашли свое применение в полногеномном анализе ДНК.

ГИДРОЛИЗ ГЕНОМНОЙ ДНК ЧЕЛОВЕКА ЭНДОНУКЛЕАЗАМИ РЕСТРИКЦИИ

Рестрикционный анализ ДНК in vitro u in silico

Применение эндонуклеаз рестрикции для анализа геномных ДНК началось сразу же после открытия этих ферментов [8]. Как правило, геномная ДНК эукариот имеет существенные размеры, в частности геном человека составляет ~3.2 млрд пар оснований (пн) [9, 10]. Как показало первоначальные исследования свойств ДНК, в частности изучение кинетики реассоциации ДНК, геном млекопитающих и человека содержит повторяющиеся последовательности ДНК, или так называемые ДНК-повторы [11, 12]. Разработанные методы секвенирования позволили определить нуклеотидный состав этих фрагментов и установить консенсусные структуры основных ДНК-повторов в геноме человека, таких как Alu, LINE1, разновидностей tandemных повторов сателлитной ДНК и др. [13]. Из-за наличия множества копий этих фрагментов в геноме при расщеплении ДНК человека рестриктазами, имеющими сайты узнавания в тех или иных ДНК-повторах, картина электрофореза продуктов ферментативного гидролиза в агарозном или акриламидном геле должна будет содержать характерные полосы.

С появлением первых результатов секвенирования ДНК-повторов в геноме человека стало

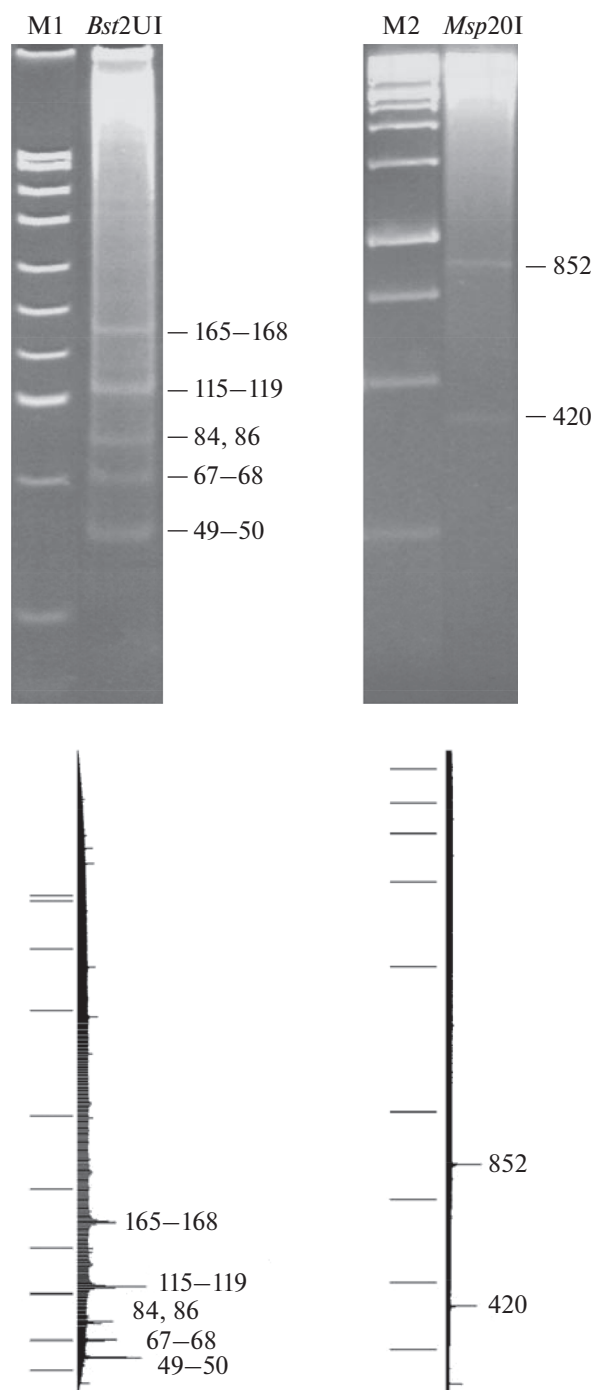


Рис. 1. Сравнение экспериментальных картин гидролиза геномной ДНК человека (вверху) с теоретическими диаграммами распределения суммарных длин фрагментов (внизу) для эндонуклеаз рестрикции *Bst2UI* (сайт узнавания CCWGG) и *Msp20I* (TGGCCA). Маркеры длин фрагментов “pUC19/*MspI*” (M1) и “SE 1kB” (M2). Справа указаны длины видимых фрагментов в пн.

возможным проведение рестрикционного анализа геномной ДНК *in silico*, т.е. получение теоретической картины расщепления ДНК по определенной короткой последовательности ДНК, являю-

щейся сайтом узнавания одной из рестриктаз. Такая работа по установлению сайтов узнавания рестриктаз *MspI*, *HinfI* и *TaqI* на секвенированных к тому времени участках ДНК человека с суммарной длиной 500 млн пн была осуществлена в работе Федоровой и др. [14] и позволила получить первые теоретические картины распределения длин фрагментов геномной ДНК человека.

С завершением проекта “Геном человека” и установлением практически всей структуры геномной ДНК появились возможности проведения полноценного рестрикционного анализа всей геномной ДНК человека *in silico*. Для успешного проведения работы по получению теоретической картины расщепления геномной ДНК на стандартном офисном компьютере нами было разработано специальное программное обеспечение GenomeDivider, позволяющее последовательно анализировать структуру ДНК всех хромосом [15], и была показана применимость этого подхода при анализе геномов мыши и крысы [16, 17]. В работе [18] были получены теоретические картины распределения фрагментов геномной ДНК человека по 15 сайтам узнавания различных эндонуклеаз рестрикции, а также получены экспериментальные картинки гидролиза суммарной лейкоцитарной ДНК человека соответствующими рестриктазами. На рис. 1 показан пример расщепления геномной ДНК человека ферментами *Bst2UI* и *Msp20I* в виде электрофореграммы в сравнении с расчетной диаграммой распределения фрагментов ДНК.

Поскольку именно расщепление ДНК-повторов дает видимый электрофоретический набор полос после гидролиза геномной ДНК, мы предложили проводить рестрикционный анализ *in silico* не всей геномной ДНК человека, а только тех ДНК-повторов, которые представлены в геноме человека в наибольшем количестве копий. Для этого мы использовали полные выборки аннотированных Alu и LINE1 повторов [19] и на рис. 2 показаны сравнительные результаты рестрикционного анализа *in silico* как всей геномной ДНК человека, так и отдельно Alu и LINE1 повторов. Представленные данные подтверждают сделанное ранее заключение, что расщепление именно ДНК-повторов дает основной вклад в выявляемые электрофорезом фрагменты ДНК.

Использование эндонуклеаз рестрикции в пробоподготовке при фрагментации геномной ДНК человека в NGS

При секвенировании геномной ДНК на NGS-машинах необходимо проводить предварительную фрагментацию геномной ДНК. Для этих целей обычно используются ультразвуковая обработка, неспецифические ДНК-эндонуклеазы (“shearase” от ZymoResearch и “fragmentase” от

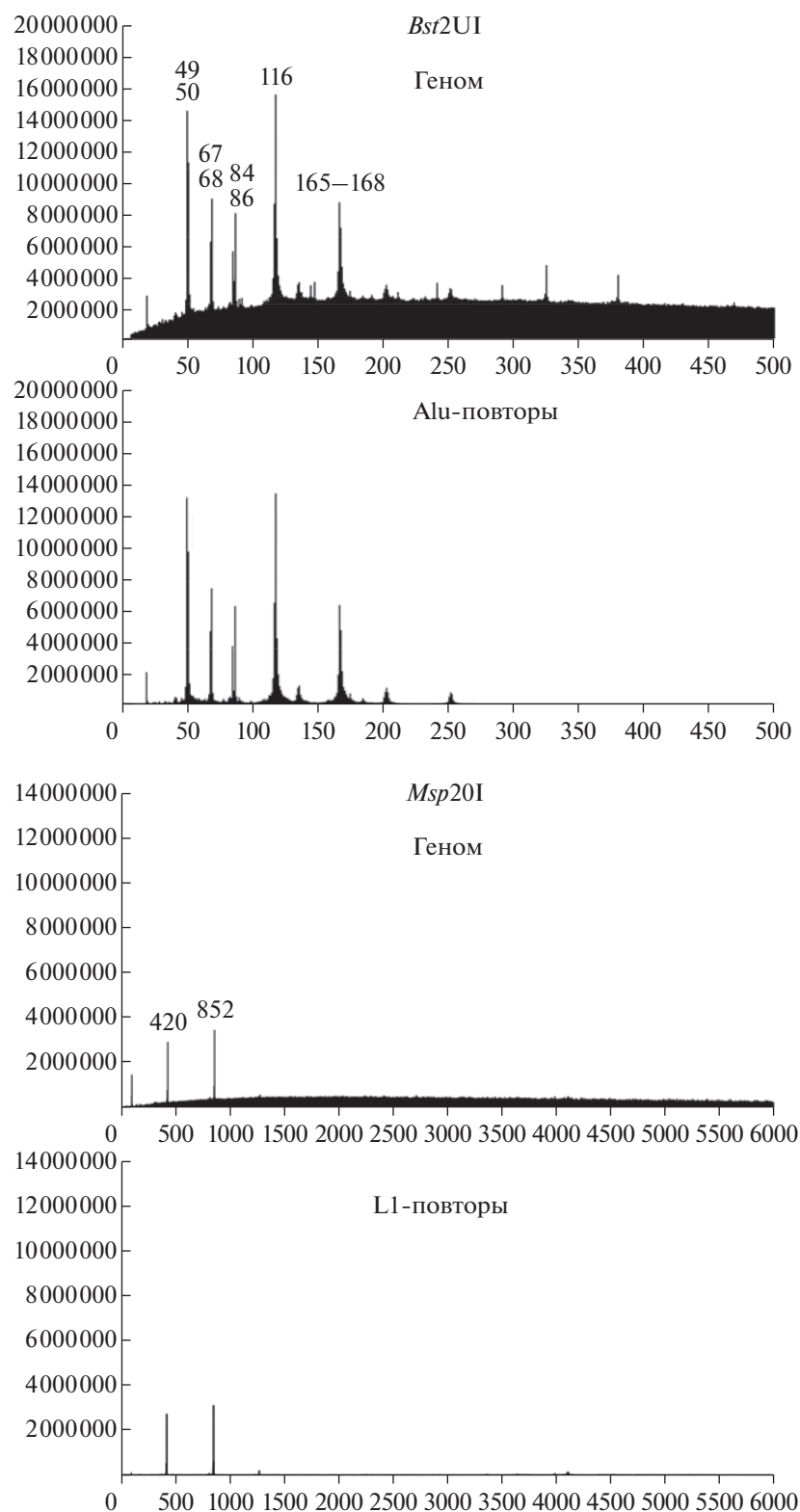


Рис. 2. Диаграммы распределения суммарных длин *Bst2UI*- и *Msp20I*-фрагментов для всего генома и повторов семейств Alu и LINE1. По оси абсцисс — длина геномного фрагмента в пн; по оси ординат — суммарное количество пар нуклеотидов во всех фрагментах данной длины. Указаны длины фрагментов (пн), соответствующие основным пикам.

New England Biolabs) или фрагментация с помощью транспозосом [20]. Однако эндонуклеазы рестрикции также могут успешно применяться для решения ряда задач при пробоподготовке к NGS. Так, использование ферментов рестрикции подтипа IIG позволило создать модифицированный транспозосомный метод для пробоподготовки пикограммных количеств геномной ДНК [21].

Для создания РЕТ-библиотек (paired-end tag) используются эндонуклеазы *MmeI* и *EcoPI5I*, вносящие разрывы в двуцепочечную ДНК на расстояниях 18/20 и 25/27 от сайта узнавания, соответственно. Такие значительные расстояния до точек гидролиза дают возможность точного картирования образуемых концов фрагментов (тэгов) при ресеквенировании геномов. РЕТ-секвенирование успешно применяется для повышения качества *de novo* сборки геномов, для локализации сайтов связывания факторов транскрипции, а также для изучения пространственной организации хроматина [22].

Широкое применение ферменты рестрикции находят в полногеномном поиске молекулярных маркеров для исследований в области популяционной и адаптационной генетики. Геномы разных людей могут различаться по нескольким миллионам нуклеотидов [23], часть из которых входит в сайты узнавания коммерчески доступных ферментов. Сравнение секвенированных рестрикционных фрагментов из различных групп организмов позволяет легко выявлять фрагменты, свойственные лишь одной из исследуемых групп и тем самым обнаружить нуклеотидные различия. Выявленные отличия могут быть в дальнейшем использованы как для разработки методик ПДРФ (полиморфизм длин рестрикционных фрагментов), так и конструирования ДНК-чипов для проведения генотипирования. Такой подход не требует обязательного сравнения со структурой референсного генома, как и проведения сборки секвенируемых геномов, что упрощает проведение биоинформатического анализа. Использование специальных олигонуклеотидных индексов (баркодов) для различения фрагментов из разных геномов дает возможность секвенирования нескольких образцов в ходе одного эксперимента, что существенно удешевляет процесс поиска специфических маркеров [24].

ЭПИГЕНЕТИЧЕСКИЙ АНАЛИЗ ГЕНОМА – АНАЛИЗ МЕТИЛИРОВАНИЯ ГЕНОМНОЙ ДНК С ПОМОЩЬЮ САЙТ- СПЕЦИФИЧЕСКИХ ДНК-ЭНДОНУКЛЕАЗ

*Анализ метилирования геномной ДНК
с помощью эндонуклеаз рестрикции*

Эндонуклеазы рестрикции, во множестве обнаруживаемые у микроорганизмов, представляют

часть бактериальной системы рестрикции–модификации. Другим обязательным компонентом таких систем являются ДНК-метилтрансферазы, которые метилируют хозяйскую ДНК по сайтам узнавания эндонуклеаз рестрикции и тем самым препятствуют ее расщеплению [25]. Метилчувствительность ферментов рестрикции дает возможность их применения в изучении метилирования ДНК. Благодаря зависимости от метилирования ДНК эти ферменты позволяют дискриминировать модифицированные сайты узнавания от немодифицированных, что, видимо, является наиболее простым способом определения статуса метилирования генов и иных функционально значимых участков генома (например, регуляторных участков и CpG-островков) [26].

Метилирование ДНК широко распространено в природе и играет важную роль в функционировании как прокариотических, так и эукариотических клеток [27, 28]. У бактерий известны три типа подобных модификаций: метилирование аденина в положении N6 и метилирование цитозина в положениях N4 и C5 [29]. У высших эукариот основным является метилирование цитозинов в положении C5, причем у млекопитающих 5-метилцитозины обычно входят в состав динуклеотидов CG. У растений, кроме CG-сайта, также часто происходит метилирование сайта CNG. Вместе с другими эпигенетическими процессами метилирование ДНК оказывает влияние на регуляцию генной активности, обеспечивая дифференциацию клеток и поддержание их тканеспецифической функциональности [30]. Существенные изменения в метилировании генома наблюдаются при канцерогенезе и ряде других заболеваний, а также при старении, что приводит к нарушению нормального протеомного баланса и развитию патологических процессов [31–35].

Наиболее полная информация о положении метилированных цитозинов в геномах может быть получена при изучении геномной ДНК, предварительно обработанной бисульфитом натрия (метод WGBS, Whole Genome Bisulphite Sequencing) [36]. Такая обработка приводит к замещению неметилированных цитозиновых остатков на урациловые (бисульфитная конверсия). Секвенирование полученных продуктов и сравнение с референсными последовательностями позволяет определить точные позиции всех метилированных и неметилированных цитозинов в геноме. Недостатками такого подхода являются высокая степень деградации ДНК во время обработки в жестких условиях (до 90%), часто наблюдаемая неполнота конверсии и значительная стоимость эксперимента [37]. С целью упрощения анализа данных и снижения стоимости также используется метод RRBS (Reduced Representation Bisulphite Sequencing), в котором геномная ДНК предварительно расщепляется эндонуклеазой

MspI (или *BglII*). Из гидролизата затем отбирается фракция фрагментов определенного диапазона длин, содержащая значительное число геномных CpG-островков. Это позволяет достичь высокой степени покрытия участков, значимых для генной регуляции, в ходе одного эксперимента [38].

В отличие от бисульфитного секвенирования полногеномные методы анализа, основанные на ферментативном расщеплении ДНК, не дают детальной и полной картины расположения всех метилированных цитозинов, а ограничены специфическими сайтами узнавания эндонуклеаз [39–42]. Тем не менее объем получаемой информации вполне достаточен для определения статуса метилирования большей части генов.

На ранних этапах изучения метилирования геномной ДНК, до тех пор, пока не появились доступные секвенирующие машины NGS, использовались методы, основанные на электрофоретическом разделении рестрикционных фрагментов и сравнении получаемых картин для разных геномов. Так, в методе RLGS (Restriction Landmark Genomic Scanning), разработанном еще до определения первичных структур геномов, проводится расщепление исследуемой и контрольной геномной ДНК метилчувствительной эндонуклеазой рестрикции и продукты гидролиза разделяются двумерным электрофорезом в геле [43]. Полученные картины представляют собой наборы пятен, сравнение которых позволяет выявить фрагменты, присутствующие лишь на одном геле. В более поздних методах MS-AP-PCR (Methylation-Sensitive Arbitrarily Primed Polymerase Chain Reaction) [44] и AIMS (Amplification of Inter-Methylated Sites) [45, 46] вводится дополнительная стадия обработки рестрикционных фрагментов – пришивка адаптерных олигонуклеотидов к их концам. Праймер для последующей ПЦР на 5'-конце в основном соответствует последовательности адаптера, но несколько нуклеотидов на его 3'-конце могут быть выбраны любыми. Это дает возможность ограничения числа амплифицируемых фрагментов, что необходимо для получения отчетливой картины разделения при обычном геле-электрофорезе. Достоинством перечисленных методов, хотя они и несколько устарели, является возможность анализа геномов с еще неустановленной нуклеотидной последовательностью с целью поиска различно метилированных участков. Однако эти способы эпигенетического исследования геномов требуют обязательного использования контрольной ДНК для сравнения, а для установления позиций метилирования необходима также элюция фрагментов из геля и их секвенирование.

В последнее время в эпигенетических исследованиях геномов также активно используются ДНК-чиповые технологии. Хотя коммерчески

доступные варианты микрочипов, такие как Roche NimbleGen и Illumina Infinium Human-Methylation450 BeadChip, позволяющие проводить анализ статуса метилирования сотен тысяч CG-сайтов в одном эксперименте, основаны на бисульфитной конверсии, но возможно и создание подобных систем на основе применения пары эндонуклеаз рестрикции – изошизомеров [47], один из которых является метилчувствительным. Хорошо известные ферменты *HpaII* и *MspI* узнают одну и ту же последовательность CCGG, но *MspI* расщепляет сайт C(5mC)GG, а *HpaII* – нет. Чип, разработанный для анализа *HpaII/MspI*-фрагментов генома, позволил определить статус метилирования промоторных регионов 8091 генов [48] и при этом современные технологии предоставляют возможность значительно увеличить число анализируемых сайтов. К общим недостаткам ДНК-чиповых технологий относятся произвольный и ограниченный выбор анализируемых позиций, трудности в выборе праймеров, специфических лишь для анализируемых участков генома, а также необходимость дополнительного дорогостоящего оборудования, способного считывать данные с микрочипов. Но, с другой стороны, простота и скорость анализа, а также возможность будущего сосредоточения чипов следующих поколений на анализе лишь определенных функционально значимых частей генома являются, несомненно, их важным преимуществом.

Современные NGS-технологии становятся все более доступными и также успешно применяются для проведения эпигенетического анализа геномов. Так, метод HELP, изначально разработанный и использованный для проведения микрочиповых исследований метилирования генома мыши, был адаптирован для проведения массового параллельного секвенирования геномной ДНК лимфобластоидной клеточной линии человека [49]. Для проведения секвенирования использовались *HpaII* и *MspI* гидролизаты одного образца геномной ДНК. Картирование фрагментов на референсный геном и сравнение полученных позиций гидролиза между собой помогли выявить как метилированные, так и неметилированные сайты CCGG в геноме клеточной линии. Сходными, также основанными на использовании пары *HpaII/MspI* для пробоподготовки, методами анализа метилирования являются MSCC [50] и Methyl-seq [51].

Очевидно, что методы, основанные на сайт-специфическом расщеплении эндонуклеазами рестрикции, так же как и микрочиповый анализ, относятся к методам “низкого разрешения”, не позволяющим охватить все CG-сайты генома. Для более полного изучения метилирования геномных ДНК предложен метод Methyl-MAPS [52]. Выбранная авторами стратегия пробоподго-

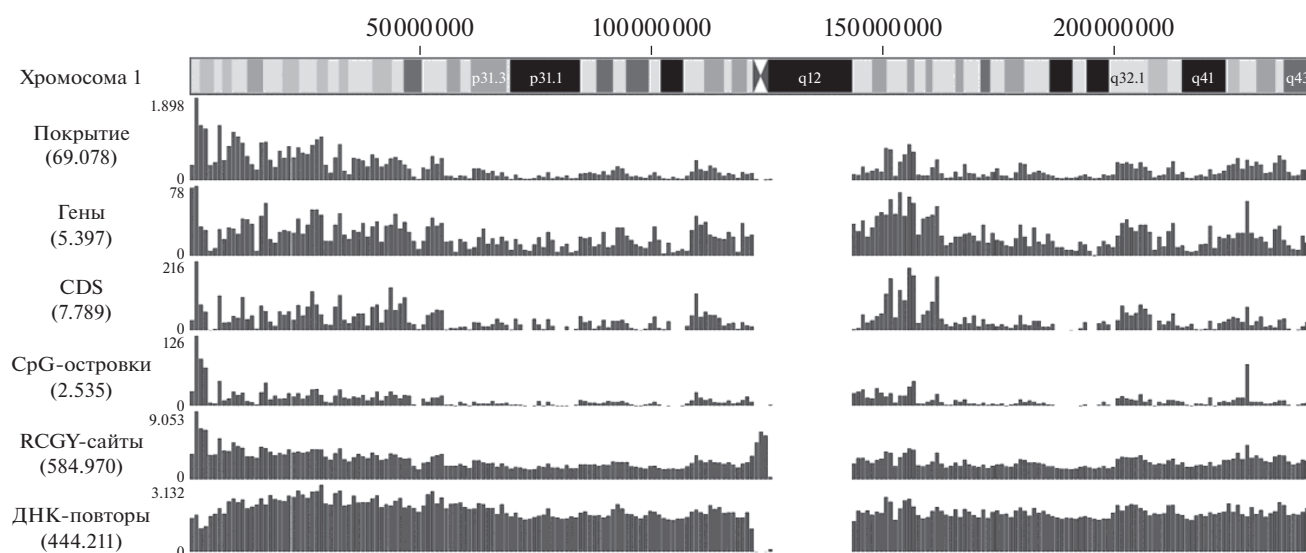


Рис. 3. Диаграммы распределения количеств достоверно покрытых участков и элементов геномной ДНК для хромосомы 1 (по стандартной аннотации к референсному геному). Вверху показана цитологическая карта хромосомы 1. В скобках приведено общее количество участков или элементов генома для данной хромосомы.

товки для секвенирования заключалась в использовании пяти метилчувствительных эндонуклеаз рестрикции, а также метилзависимой нуклеазы MsrBC для селекции неметилированных участков генома. Такой подход, хотя и является достаточно трудоемким и дорогостоящим, позволил оценить статус метилирования 81.4% из ~28 млн сайтов CG в геноме человека.

Анализ метилирования геномной ДНК с помощью сайт-специфических метилзависимых ДНК-эндонуклеаз

5-Метилцитозин-зависимые сайт-специфические ДНК-эндонуклеазы, которые узнают и расщепляют только метилированную ДНК и, как выяснилось в последние годы, широко распространены у бактерий, могут существенно расширить возможности эпигенетических исследований [7, 53]. Это касается как локус-специфического определения статуса метилирования регуляторных участков генов [54, 55], так и полного анализа метилирования.

Одним из первых открытых ферментов этой группы является MsrBC [56], узнающий последовательность R(5mC)N₄₀₋₃₀₀₀R(5mC). Однако вариативность расстояния между двумя частями сайта узнавания и отсутствие определенности для позиции расщепления не дают возможности точной локализации метилированных сайтов с помощью секвенирования MsrBC-фрагментов.

Одну из подгрупп метилзависимых эндонуклеаз составляют ферменты, способные вносить двуцепочечные разрывы в стороне от сайта

узнавания. В частности, к ним относится фермент *MspII*, узнающий последовательность (5mC)NNRN(9/13). Поскольку в ДНК млекопитающих метилированный цитозин входит в состав динуклеотида CG, в котором метилируется также цитозин в составе CG-динуклеотида на нижней цепи, то при расщеплении геномных сайтов YN(5mC)GNR происходит образование фрагментов с преобладающей длиной 32 пн. Такие короткие фрагменты могут быть отсеквенированы с помощью NGS-машины ABI SOLiD и центрально расположенный в них CG-сайт укажет на присутствие метильной метки в нем. Пробное секвенирование такого рода, проведенное в работе [57], позволило картировать ~10.4 млн позиций метилированных CG-сайтов, однако более глубокий анализ для определения статуса метилирования генов не был проведен. Возможно, это связано с тем, что среднее покрытие секвенированных участков оказалось небольшим (5×) и наиболее хорошо покрытые участки относились к повторяющимся элементам ДНК.

Недавно нами было проведено секвенирование фрагментов генома, образованных при гидролизе ДНК метилзависимой эндонуклеазой *GlaI*, на NGS-приборе Illumina MiSeq [58]. Из всех доступных ферментов этой группы *GlaI* представляется наиболее подходящим для проведения эпигенетических исследований. *GlaI* расщепляет ДНК с образованием тупых концов, а его сайт узнавания R(5mC)GY включает в себя динуклеотид CG и при этом исходный сайт RCGY является оптимальным субстратом для *de novo* метилтрансфераз DNMT3a и DNMT3b человека

[59]. Для того чтобы избежать значительной примеси повторяющейся ДНК в секвенируемых фрагментах, предварительно были проведены теоретические расчеты длин фрагментов, образующихся при расщеплении основных ДНК-повторов (Alu, LINE1, satDNA). Этот теоретический анализ позволил определить, что большая часть *GlaI*-гидролизированных фрагментов из ДНК-повторов имеет размер до 140 пн. Поэтому для секвенирования была отобрана фракция *GlaI*-фрагментов геномной ДНК клеточной линии Raji длиной 140–400 пн. Проведенное картирование секвенированных участков генома показало (рис. 3), что наибольшее покрытие достигнуто для зон расположения функционально значимых частей генома (генов, CpG-островков и белок-кодирующих участков). Сравнение полученных данных с результатами ранее проведенных локус-специфических исследований метилирования регуляторных участков ряда генов-онкосупрессоров [54, 55, 60] выявило их хорошую сходимость. Таким образом, секвенирование *GlaI*-фрагментов также может быть использовано для проведения полногеномного эпигенетического анализа.

ЗАКЛЮЧЕНИЕ

В работах по изучению ДНК человека используются различные сайт-специфические ДНК-эндонуклеазы, однако их применение в основном связано с изучением какого-то конкретного участка генома и таким образом имеет локальный характер. Из всех типов этих ферментов только эндонуклеазы рестрикции и метилзависимые ДНК-эндонуклеазы нашли применение в полногеномном анализе ДНК человека. Однако, на наш взгляд, с развитием NGS-технологий использование рестриктаз в исследованиях всего генома человека, скорее всего, будет отходить на второй план, тогда как роль метилзависимых ДНК-эндонуклеаз в изучении эпигенома человека будет только возрастать. Это связано с тем, что другие имеющиеся инструменты исследования эпигенома имеют существенные недостатки, в частности, использование метилчувствительных рестриктаз имеет ограниченное применение, а бисульфитная конверсия ДНК дает высокий процент ошибок. В то же время расширение списка используемых сайт-специфических метилзависимых ДНК-эндонуклеаз и разработка новых технологий их применения в геномных исследованиях должны привести к увеличению числа работ по изучению эпигенома человека с использованием этих ферментов.

СПИСОК ЛИТЕРАТУРЫ

1. Smith H.O., Wilcox K.W. A restriction enzyme from *Haemophilus influenzae*. I. Purification and general proper-

- ties // J. Mol. Biol. 1970. V. 51. № 2. P. 379–391. doi 10.1016/0022-2836(70)90149-X
2. Loenen W.A., Dryden D.T., Raleigh E.A. et al. Highlights of the DNA cutters: a short history of the restriction enzymes // Nucl. Acids Res. 2014. V. 42. № 1. P. 3–19. doi 10.1093/nar/gkt990
3. Железная Л.А., Качалова Г.С., Артюх Р.И. и др. Никующие эндонуклеазы // Успехи биологической химии. 2009. Т. 49. С. 107–128.
4. Hafez M., Hausner G. Homing endonucleases: DNA scissors on a mission // Genome. 2012. V. 55. № 8. P. 553–569. doi 10.1139/g2012-049
5. Bogdanove A.J., Voytas D.F. TAL effectors: customizable proteins for DNA targeting // Science. 2011. V. 333. № 6051. P. 1843–1846. doi 10.1126/science.1204094
6. Peters J.M., Silvis M.R., Zhao D. et al. Bacterial CRISPR: accomplishments and prospects // Curr. Opin. Microbiol. 2015. V. 27. P. 121–126. doi 10.1016/j.mib.2015.08.007
7. Землянская Е.В., Дегтярев С.Х. Субстратная специфичность и свойства метилзависимых сайт-специфических ДНК-эндонуклеаз // Мол. биология. 2013. Т. 47. № 6. С. 900–913. doi 10.1134/S0026893313060186
8. Roberts R.J. How restriction enzymes became the workhorses of molecular biology // Proc. Natl Acad. Sci. USA. 2005. V. 102. № 17. P. 5905–5908. doi 10.1073/pnas.0500923102
9. International Human Genome Sequencing Consortium. Finishing the euchromatic sequence of the human genome // Nature. 2004. V. 431. № 7011. P. 931–945. doi 10.1038/nature03001
10. Church D.M., Schneider V.A., Graves T. et al. Modernizing reference genome assemblies // PLoS Biol. 2011. V. 9. № 7. P. e1001091. doi 10.1371/journal.pbio.1001091
11. Jelinek W.R., Schmid C.W. Repetitive sequences in eukaryotic DNA and their expression // Annu. Rev. Biochem. 1982. V. 51. P. 813–844. doi 10.1146/annurev.bi.51.070182.004121
12. Biscotti M.A., Olmo E., Heslop-Harrison J.S. Repetitive DNA in eukaryotic genomes // Chromosome Res. 2015. V. 23. № 3. P. 415–420. doi 10.1007/s10577-015-9499-z
13. Hardman N. Structure and function of repetitive DNA in eukaryotes // Biochem. J. 1986. V. 234. № 1. P. 1–11. doi 10.1042/bj2340001
14. Федорова Л.В., Дайзадех И., Федоров А.Н., Рысков А.П. Анализ *in silico* распределения рестриктных фрагментов генома человека по длинам // Генетика. 2001. Т. 37. № 4. С. 456–466.
15. Абдурашитов М.А., Томилов В.Н., Чернухин В.А. и др. Метод рестрикционного анализа геномов млекопитающих *in silico* // Вестник биотехнологии и физико-химической биологии им. Ю.А. Овчинникова. 2006. Т. 2. № 3. С. 29–38.
16. Чернухин В.А., Абдурашитов М.А., Томилов В.Н. и др. Сравнительный рестрикционный анализ хромосомной ДНК мыши *in vitro* и *in silico* // Вестник биотехнологии и физико-химической биологии им. Ю.А. Овчинникова. 2007. Т. 3. № 4. С. 19–27.

17. Чернухин В.А., Абдурашитов М.А., Томилов В.Н. и др. Сравнительный рестрикционный анализ хромосомной ДНК крысы *in vitro* и *in silico* // Вестник биотехнологии и физико-химической биологии им. Ю.А. Овчинникова. 2006. Т. 2. № 3. С. 39–46.
18. Абдурашитов М.А., Томилов В.Н., Чернухин В.А. и др. Рестрикционный анализ ДНК человека – новые возможности в ДНК-диагностике // Мед. генетика. 2007. Т. 6. № 8. С. 29–36.
19. Karolchik D., Hinrichs A.S., Furey T.S. et al. The UCSC Table Browser data retrieval tool // Nucl. Acids Res. 2004. V. 32 (Suppl. 1). P. D493–496. doi 10.1093/nar/gkh103
20. Head S.R., Komori H.K., LaMere S.A. et al. Library construction for next-generation sequencing: overviews and challenges // Biotechniques. 2014. V. 56. № 2. P. 61–68. doi 10.2144/000114133
21. Parkinson N.J., Maslau S., Ferneyhough B. et al. Preparation of high-quality next-generation sequencing libraries from picogram quantities of target DNA // Genome Res. 2012. V. 22. № 1. P. 125–133. doi 10.1101/gr.124016.111
22. Fullwood M.J., Wei C.L., Liu E.T., Ruan Y. Next-generation DNA sequencing of paired-end tags (PET) for transcriptome and genome analyses // Genome Res. 2009. V. 19. № 4. P. 521–532. doi 10.1101/gr.074906.107
23. Wheeler D.A., Srinivasan M., Egholm M. et al. The complete genome of an individual by massively parallel DNA sequencing // Nature. 2008. V. 452. № 7189. P. 872–876. doi 10.1038/nature06884
24. Davey J.W., Hohenlohe P.A., Etter P.D. et al. Genome-wide genetic marker discovery and genotyping using next-generation sequencing // Nat. Rev. Genet. 2011. V. 12. № 7. P. 499–510. doi 10.1038/nrg3012
25. Pingoud A., Wilson G.G., Wende W. Type II restriction endonucleases – a historical perspective and more // Nucl. Acids Res. 2014. V. 42. № 12. P. 7489–7527. doi 10.1093/nar/gku447
26. Olkhov-Mitsel E., Bapat B. Strategies for discovery and validation of methylated and hydroxymethylated DNA biomarkers // Cancer Med. 2012. V. 1. № 2. P. 237–260. doi 10.1002/cam4.22
27. Jeltsch A. Beyond Watson and Crick: DNA methylation and molecular enzymology of DNA methyltransferases // Chembiochem. 2002. V. 3. № 4. P. 274–293. doi 10.1002/1439-7633(20020402)3:4<274::AID-CBIC274>3.0.CO;2-S
28. Robertson K.D., Jones P.A. DNA methylation: past, present and future directions // Carcinogenesis. 2000. V. 21. № 3. P. 461–467. doi 10.1093/carcin/21.3.461
29. Sánchez-Romero M.A., Cota I., Casadesús J. DNA methylation in bacteria: from the methyl group to the methylome // Curr. Opin. Microbiol. 2015. V. 25. P. 9–16. doi 10.1016/j.mib.2015.03.004
30. He X.J., Chen T., Zhu J.K. Regulation and function of DNA methylation in plants and animals // Cell Res. 2011. V. 21. № 3. P. 442–465. doi 10.1038/cr.2011.23
31. Portela A., Esteller M. Epigenetic modifications and human disease // Nat. Biotechnol. 2010. V. 28. № 10. P. 1057–1068. doi 10.1038/nbt.1685
32. Martín-Subero J.I., Esteller M. Profiling epigenetic alterations in disease // Adv. Exp. Med. Biol. 2011. V. 711. P. 162–177. doi 10.1007/978-1-4419-8216-2_12
33. Brazel A.J., Vernimmen D. The complexity of epigenetic diseases // J. Pathol. 2016. V. 238. № 2. P. 333–344. doi 10.1002/path.4647
34. Ghavifekr F.M., Farshdousti H.M., Shanebandi D., Baradaran B. DNA methylation pattern as important epigenetic criterion in cancer // Genet. Res. Int. 2013. P. 2013:317569. doi 10.1155/2013/317569
35. Jung M., Pfeifer G.P. Aging and DNA methylation // BMC Biol. 2015. V. 13. P. 7. doi 10.1186/s12915-015-0118-4
36. Lister R., Pelizzola M., Dowen R.H. et al. Human DNA methylomes at base resolution show widespread epigenomic differences // Nature. 2009. V. 462. № 7271. P. 315–322. doi 10.1038/nature08514
37. Grunau C., Clark S.J., Rosenthal A. Bisulfite genomic sequencing: systematic investigation of critical experimental parameters // Nucl. Acids Res. 2001. V. 29. № 13. P. e65–5. doi 10.1093/nar/29.13.e65
38. Meissner A., Gnirke A., Bell G.W. et al. Reduced representation bisulfite sequencing for comparative high-resolution DNA methylation analysis // Nucl. Acids Res. 2005. V. 33. № 18. P. 5868–5877. doi 10.1093/nar/gki901
39. Laird P.W. Principles and challenges of genome-wide DNA methylation analysis // Nat. Rev. Genet. 2010. V. 11. № 3. P. 191–203. doi 10.1038/nrg2732
40. Fouse S.D., Nagarajan R.P., Costello J.F. Genome-scale DNA methylation analysis // Epigenomics. 2010. V. 2. № 1. P. 105–117. doi 10.2217/epi.09.35
41. Umer M., Herceg Z. Deciphering the epigenetic code: an overview of DNA methylation analysis methods // Antioxid. Redox Signal. 2013. V. 18. № 15. P. 1972–1986. doi 10.1089/ars.2012.4923
42. Ku C.S., Naidoo N., Wu M., Soong R. Studying the epigenome using next generation sequencing // J. Med. Genet. 2011. V. 48. № 11. P. 721–730. doi 10.1136/jmedgenet-2011-100242
43. Kawai J., Hirotsune S., Hirose K. et al. Methylation profiles of genomic DNA of mouse developmental brain detected by restriction landmark genomic scanning (RLGS) method // Nucl. Acids Res. 1993. V. 21. № 24. P. 5604–5608. doi 10.1093/nar/21.24.5604
44. Liang G., Gonzalgo M.L., Salem C., Jones P.A. Identification of DNA methylation differences during tumorigenesis by methylation-sensitive arbitrarily primed polymerase chain reaction // Methods. 2002. V. 27. № 2. P. 150–155. doi 10.1016/S1046-2023(02)00068-3
45. Frigola J., Ribas M., Risques R.A., Peinado M.A. Methylome profiling of cancer cells by amplification of inter-methylated sites (AIMS) // Nucl. Acids Res. 2002. V. 30. № 7. P. e28. doi 10.1093/nar/30.7.e28
46. Стрельников В.В., Танас А.С., Землякова В.В., Залетаев Д.В. Разработка новых методов диагностики маркеров метилирования ДНК в онкологии // Системы генетических и эпигенетических маркеров в диагностике онкологических новообразований. М.: Медицина, 2009. С. 349–384.

47. Roberts R.J., Macelis D. Restriction enzymes and their isoschizomers // Nucl. Acids Res. 1991. V. 19 (Suppl). P. 2077–2109. doi 10.1093/nar/19.suppl.2077
48. Hatada I., Fukasawa M., Kimura M. et al. Genome-wide profiling of promoter methylation in human // Oncogene. 2006. V. 25. № 21. P. 3059–3064. doi 10.1038/sj.onc.1209331
49. Oda M., Glass J.L., Thompson R.F. et al. High-resolution genome-wide cytosine methylation profiling with simultaneous copy number analysis and optimization for limited cell numbers // Nucl. Acids Res. 2009. V. 37. № 12. P. 3829–3839. doi 10.1093/nar/gkp260
50. Ball M.P., Li J.B., Gao Y. et al. Targeted and genome-scale strategies reveal gene-body methylation signatures in human cells // Nat. Biotechnol. 2009. V. 27. № 4. P. 361–368. doi 10.1038/nbt.1533
51. Brunner A.L., Johnson D.S., Kim S.W. et al. Distinct DNA methylation patterns characterize differentiated human embryonic stem cells and developing human fetal liver // Genome Res. 2009. V. 19. № 6. P. 1044–1056. doi 10.1101/gr.088773.108
52. Edwards J.R., O'Donnell A.H., Rollins R.A. et al. Chromatin and sequence features that define the fine and gross structure of genomic methylation patterns // Genome Res. 2010. V. 20. № 7. P. 972–980. doi 10.1101/gr.101535.109
53. Loenen W.A., Raleigh E.A. The other face of restriction: modification-dependent enzymes // Nucl. Acids Res. 2014. V. 42. № 1. P. 56–69. doi 10.1093/nar/gkt747
54. Гончар Д.А., Акишев А.Г., Дегтярев С.Х. BlnI- и GlnI-ПЦР анализ – новый метод исследования метилированных участков ДНК // Вестник биотехнологии и физико-химической биологии им. Ю.А. Овчинникова. 2010. Т. 6. № 1. С. 5–12.
55. Акишев А.Г., Гончар Д.А., Абдурашитов М.А., Дегтярев С.Х. Эпигенетическое типирование маллигнантных клеточных линий человека с помощью BlnI- и GlnI ПЦР анализа // Вестник биотехнологии и физико-химической биологии им. Ю.А. Овчинникова. 2011. Т. 7. № 3. С. 5–16.
56. Raleigh E.A. Organization and function of the mcrBC genes of *Escherichia coli* K-12 // Mol. Microbiol. 1992. V. 6. № 9. P. 1079–1086. doi 10.1111/j.1365-2958.1992.tb01546.x
57. Cohen-Karni D., Xu D., Apone L. et al. The MspJI family of modification-dependent restriction endonucleases for epigenetic studies // Proc. Natl Acad. Sci. USA. 2011. V. 108. № 27. P. 11040–11045. doi 10.1073/pnas.1018448108
58. Abdurashitov M.A., Tomilov V.N., Gonchar D.A. et al. Mapping of R(5mC)GY sites in the genome of human malignant cell line Raji // Biol. Med. (Aligarh). 2015. V. 7. № 4. P. BM-135–15.
59. Handa V., Jeltsch A. Profound sequence preference of Dnmt3a and Dnmt3b mammalian DNA methyltransferases shape the human epigenome // J. Mol. Biol. 2006. V. 348. № 5. P. 1103–1112. doi 10.1016/j.jmb.2005.02.044
60. Абдурашитов М.А., Куксова А.Н., Акишев А.Г. и др. MteI-ПЦР-анализ – новый метод определения статуса метилирования GC-богатых регуляторных участков генов-онкосупрессоров человека // Вестник биотехнологии и физико-химической биологии им. Ю.А. Овчинникова. 2013. Т. 9. № 3. С. 15–23.

Use of Site-Specific DNA Endonucleases in Genome-Wide Studies of Human DNA

M. A. Abdurashitov^a and S. Kh. Degtyarev^{a, b, *}

^aSibEnzyme Ltd., Novosibirsk, 630117 Russia

^bVirology and Biotechnology Research Center Vector, Koltsovo, Novosibirsk oblast, 630559 Russia

*e-mail: degt@sibenzyme.com

During the last decades, site-specific DNA endonucleases have served as a key instrument to study primary structure of DNA and genetic engineering. Here, we describe examples of these enzyme uses in genome-wide analysis of human DNA including restriction endonucleases involvement during sample preparation for sequencing using NGS devices, as well as visualization of cleavage of DNA repeats by endonucleases. The first studies on application of DNA endonucleases in the rapidly developing area of epigenetic analysis of genomes, which is facilitated by the recent discovery of a new class of enzymes, 5-methylcytosine-dependent site-specific DNA endonucleases, are of special interest.

Keywords: site-specific endonucleases, genomics, epigenomics, DNA methylation, sequencing.