

Методы полногеномной амплификации генетического материала единичных клеток

Твеленёва А.А., Мусатова Е.В., Шилова Н.В.

Федеральное государственное бюджетное научное учреждение «Медико-генетический научный центр», 115478, Москва, ул. Москворечье д.1, тел. 007 (0499) 612-86-07 e-mail: al_bardashevskaya@mail.ru

Анализ генетического материала единичных клеток является весьма актуальной задачей для таких областей науки, как онкология, судебная медицина, преимплантационная генетическая диагностика. Полногеномная амплификация является неотъемлемым этапом в исследовании единичных клеток. Среди методов полногеномной амплификации выделяют методы, основанные на термоциклировании, и изотермические методы. Каждый из методов полногеномной амплификации должен обеспечивать максимально возможную представленность и пропорциональность всех участков исследуемого генома, а также обладать минимальным смещением, то есть наиболее точно представлять первичную последовательность ДНК в продуктах полногеномной амплификации. На вышеперечисленные характеристики могут влиять используемые в методах полногеномной амплификации типы праймеров, полимеразы, а также исходное качество и количество исследуемой ДНК. Поэтому в зависимости от используемого протокола качество амплификации может значительно различаться, что, несомненно, должно учитываться при выборе метода полногеномной амплификации в соответствии с поставленной целью исследования.

Ключевые слова: полногеномная амплификация, единичные клетки, репрезентативность амплификации, предпочтительная амплификация, выпадение аллелей.

Информация о конфликте интересов. Авторы заявляют об отсутствии конфликтов интересов.

Whole genome amplification as a method for analysis of single cells

Tveleneva A.A., Musatova E.V., Shilova N.V.

Research Centre for Medical Genetics, Moscow, 115478, Russia, al_bardashevskaya@mail.ru

Analysis of the genetic material of single cells is a very urgent task for such fields of science as oncology, forensic medicine, preimplantation genetic diagnosis. Whole genome amplification is an indispensable stage in the study of single cells. Among methods of whole genome amplification, methods based on thermocycling and isothermal methods are distinguished. Each of the methods of whole genome amplification should ensure the maximum possible representation and proportionality of all parts of the investigated genome, and also have a minimal bias, that is, it is most accurate to represent the primary DNA sequence in products of whole genome amplification. These characteristics can be influenced by the types of primers, polymerases used in the methods of whole genome amplification, as well as the initial quality and quantity of the DNA being studied. Therefore, depending on the protocol used, the quality of the amplification may vary significantly, which should undoubtedly be taken into account when choosing the method of whole genome amplification in accordance with the intended purpose of the study.

Keywords: whole genome amplification, single cells, representative amplification, preferential amplification, allele dropout.

Введение

Анализ генома единичных клеток является актуальной задачей современных исследований в области онкологии, преимплантационной генетической диагностики, изучения гетерогенности клеточных популяций и неинвазивной пренатальной диагностики с использованием клеток плодной природы. Получение достаточного количества ДНК единичных клеток открывает возможности для последующего анализа их генома, транскриптома, эпигенетических особенностей. Для преодоления проблемы малого стартового количества ДНК используются методы полногеномной амплификации. Полногеномная амплификация (ПГА) обеспечивает многократное увеличение количества ДНК, соответствующей полному геному исследуемых клеток, в отличие от локус-специфичной амплификации, во время кото-

рой происходит увеличение количества определенного участка ДНК. Выделяют два основных типа ПГА, основанных на полимеразной цепной реакции (ПЦР) и изотермической амплификации.

Методы, основанные на ПЦР

Первым методом ПГА, основанным на принципе ПЦР, был IRS-PCR (Interspersed Repetitive Sequence PCR), разработанный Nelson с соавторами в 1989 году [1]. Основным недостатком данного метода явилось то, что в продуктах амплификации представленность последовательностей, соответствующих GTG-негативным бэндам хромосом, оказалась значительно выше [2]. С целью преодоления этого ограничения была разработана ПГА с использованием линкерных адаптеров,

LA-PCR (Linker adapter PCR). Данный метод предполагает предварительную рестрикционную обработку ДНК в специфических сайтах с последующим присоединением к концам полученных фрагментов синтетических олигонуклеотидных адаптеров, содержащих последовательность, комплементарную прямому и обратному праймерам [3].

В 1992 году в публикации Zhang с соавторами была предложена полимеразная цепная реакция с удлинением праймера в циклах преамплификации, PEP-PCR (Primer extension preamplification PCR), при которой используется смесь из 4×10^9 полностью вырожденных олигонуклеотидных праймеров длиной 15 нуклеотидов для достижения универсальной амплификации случайных последовательностей генома. В каждом из 50 циклов ДНК денатурирует при 92°C . Затем праймеры присоединяются к денатурированной ДНК при низкой температуре (37°C), постепенно увеличивающейся до 55°C и удерживающейся на этом уровне в течение четырех минут для элонгации цепи Taq-полимеразой [4].

Был разработан улучшенный протокол PEP-PCR — I-PEP-PCR (Improve PEP-PCR), который отличается от стандартного подходами к лизису клеток, улучшенными условиями термического цикла и добавлением полимеразы с более высокой точностью синтеза. В частности, лизис клеток проводят в специальном буфере, Taq-полимеразу используют совместно с Pwo-полимеразой и вводят дополнительную стадию элонгации цепи при 68°C . Эти изменения протокола позволили со 100%-ной точностью амплифицировать последовательность гена размером более 4 т.п.н. [5].

В 1992 году Telenius с соавторами была предложена ПГА с частично вырожденными олигонуклеотидными праймерами DOP-PCR (Degenerate oligonucleotide primed PCR). Для данного метода характерны две основные особенности. Во-первых, использование частично вырожденного олионуклеотидного праймера длиной 22 нуклеотида. С 3'-конца праймера расположена четко определенная последовательность (ATGTGG), с 5'-конца праймера расположен сайт узнавания рестриктазы XhoI, что позволяет легко клонировать амплифицированные фрагменты ДНК, а между ними находится 6 вырожденных нуклеотидов, обеспечивающих все возможные комбинации последовательности ДНК для осуществления ПГА. Во-вторых, особенностью DOP-PCR является наличие двух типов циклов: низкотемпературных и высокотемпературных. При низкотемпературных циклах отжиг 3'-участка праймера до 5'-конца вырожденного 6-нуклеотидного района праймера происходит при 30°C неспецифически через каждые 4 т.п.н. За время медленного нагрева до 72°C полимеразы производит удлинение присоединенного олигонуклеотида, что позволяет сохранить затравку при высокой температуре элонгации цепи. Таким образом, накапливаются фрагменты, фланкированные DOP-праймерами. При высокотемпературных циклах

происходят специфический отжиг праймеров и амплификация фрагментов, созданных при низкотемпературных циклах [6].

Было определено, что DOP-PCR генерирует относительно короткие продукты ПГА (до 1500 п.н. при использовании высокомолекулярной ДНК). Однако были описаны новые методы, позволяющие получить длинные фрагменты из низкого стартового количества ДНК — метод LL-DOP-PCR, продукты амплификации которого могут варьировать от 0,5 до 10 т.п.н. Это было достигнуто с помощью 3'-5'-экзонуклеазной корректирующей активности Pwo-полимеразы и увеличения времени отжига праймеров. По сравнению со стандартным протоколом DOP-PCR, данный протокол обеспечивает значительно лучшее покрытие для микросателлитных и уникальных последовательностей [7].

Существует еще одна модификация стандартного протокола DOP-PCR — dcDOP-PCR, использующего более вырожденный праймер (10 вырожденных нуклеотидов), 12 неспецифических циклов и Pwo-полимеразу. Эти модификации позволили увеличить число обнаруживаемых коротких tandemных повторов примерно на 45% по сравнению с традиционным методом DOP-PCR [8].

В 2012 году Zong с соавторами сообщили о создании нового метода амплификации единичных клеток — MALBAC (циклы амплификации, основанные на множественных отжигах и образовании петель — Multiple Annealing and Looping Based Amplification Cycles), позволяющего равномерно амплифицировать до 93% генома единичной клетки. Особенностью данного метода является то, что на первом этапе, этапе преамплификации, с использованием полимеразы, обладающей свойством вытеснения цепи, осуществляется создание перекрывающихся полуампликонов, ограниченных с 3'-конца MALBAC-праймером и охватывающих наибольшую часть генома, по типу «шотган» секвенирования (shotgun sequencing). На втором этапе проводят ПЦР с целью создания полных ампликонов, фланкированных с двух концов специфическими праймерами, которые амплифицируются на третьем этапе реакции [9].

Метод изотермической амплификации

В 2002 году Dean с соавторами описали метод MDA (Multiple displacement amplification), основанный на множественной амплификации путем вытеснения комплементарно-синтезируемой цепи. Принципы этого метода заключаются в следующем. Случайные гексамеры связываются с денатурированной ДНК. Фермент осуществляет синтез комплементарной цепи до тех пор, пока не достигнет вновь синтезированной двухцепочечной ДНК. После этого фермент начинает смещать синтезированную нить и продолжает элонгацию ДНК, в это же время праймеры связываются с вновь синтезированной вытесненной ДНК. Это приводит к тому, что полимери-

зация начинается на новых нитях, образующих гипер-разветвленные структуры [10]. Эту реакцию можно катализировать с помощью Phi29 ДНК-полимеразы или большим фрагментом Bst ДНК-полимеразы. Однако Phi29 ДНК-полимераза более предпочтительна, чем Bst ДНК-полимераза, по причине высокой процессивности и низкой частоты ошибок [11]. Известно, что частота возникновения ошибок при использовании Phi29 ДНК-полимеразы в 100 раз ниже, чем Taq-полимеразы. Такая особенность Phi29 ДНК-полимеразы обусловлена корректирующей 3'-5' экзонуклеазной активностью и добавляет в среднем 20 000 нуклеотидов к праймеру. Помимо этого, Phi29 ДНК-полимераза обладает высокой активностью замещения цепи, таким образом создаются условия для экспоненциального роста амплифицированных продуктов с помощью механизма разветвления, что приводит к высокому выходу продукта реакции. В MDA используются случайные гексамерные праймеры, которые имеют изменение в их 3'-конце в виде тиофосфата для защиты от деградации под воздействием 3'-5' экзонуклеазной активности Phi29 ДНК-полимеразы.

Основные проблемы методов полногеномной амплификации

Проблемами, приводящими к неточной воспроизводимости амплифицируемой последовательности ДНК при ПГА, являются предпочтительная амплификация (РА — preferential amplification) и выпадение аллеля (ADO — allele dropout). ADO представляет собой отсутствие в продуктах амплификации одного из двух гетерозиготных аллелей, тогда как другой аллель успешно амплифицируется. РА отражает процесс амплификации обоих аллелей, при котором в продуктах амплификации один из аллелей представлен в меньшем количестве. Имеется ряд факторов, которые влияют на появление ADO и РА. К РА может приводить неравномерная денатурация, происходящая из-за различий в процентном содержании GC-пар между аллелями. Эффективность отжига праймера также связана с GC-составом. Менее эффективный отжиг праймера на одном аллеле может приводить к РА другого аллеля. При деградации целевой ДНК предпочтительнее амплифицирован может быть более короткий аллель. Помимо этого, РА тесно связана с исходным количеством анализируемых геномов, т.е. количеством анализируемых клеток. Когда их количество очень мало, это создает условия для возникновения РА. Наиболее частой причиной ADO является деградация ДНК и/или её небольшое количество, условия лизиса и ПЦР, а также длина амплифицируемого аллеля. Так, было установлено, что преимущественно выпадает более длинный аллель [12].

Еще одной проблемой полногеномной амплификации является низкая эффективность и точность ампли-

фикации повторяющихся последовательностей ДНК, в частности коротких tandemных повторов, в методах, основанных на ПЦР. Это, вероятно, связано с низкой температурой отжига, характерной для протоколов ПГА, основанных на ПЦР. Согласно докладу Cheung и Nelson, такая проблема главным образом возникает, когда доступно лишь небольшое количество образца [13]. Было установлено, что ПГА на основе MDA лишена этого недостатка [14]. Возможно, это связано с тем, что при изотермической амплификации используется высокопроцессивная полимеразы с 3'-5' экзонуклеазной активностью, что позволяет получать длинные продукты амплификации, охватывающие большинство STR-маркеров.

Исходный материал для амплификации всего генома часто имеет разную степень деградации, поэтому низкий уровень качества ДНК представляет ещё одну проблему для ПГА. Как правило, фрагментация исходного образца является более сложным препятствием для методов ПГА на основе MDA, поэтому методами выбора для ПГА образцов с низким качеством исходной матрицы являются методы, основанные на ПЦР [15]. Так, Pirker с соавторами обнаружили, что для полногеномной амплификации фрагментированной ДНК методом выбора является LA-PCR. Помимо этого, было установлено, что DOP-PCR на таких образцах приводит к ложноположительным результатам [16].

Сравнительная характеристика методов полногеномной амплификации

При выборе протокола ПГА для амплификации генетического материала единичных клеток важным является соотношение задачи исследования с возможностями различных методик. Ряд исследований посвящен сравнительной характеристике имеющихся коммерческих наборов ПГА, основанных на том или ином принципе амплификации. Согласно последним исследованиям методами выбора ПГА для последующего секвенирования генома единичных клеток являются MALBAC и LA-PCR, так как они демонстрируют наилучшую представленность последовательностей геномной ДНК (60% и 50% последовательности генома амплифицируются этими методами соответственно) [17]. Для точного определения CNV (Copy Number Variation — вариации количества копий) DOP-PCR и MALBAC обеспечивают максимальную воспроизводимость результатов и самый низкий коэффициент вариации, характеризующий равномерность в распределении продуктов ПГА по геному. Методы MALBAC и MDA хорошо подходят для одновременного определения CNV и SNV (Single Nucleotide Variation — однонуклеотидные вариации), так как обладают наименьшим уровнем ADO (для данных методов в 21% и 33% аллелей соответственно) [18].

Заключение

При работе с единичными клетками методы ПГА позволяют получить достаточное количество геномной ДНК для последующих молекулярно-генетических и молекулярно-цитогенетических исследований. Основной задачей данных методов является обеспечение наилучшей репрезентативности и пропорциональности амплификации, что достигается путем использования различных праймеров, полимераз, принципа и условий амплификации. При выборе метода ПГА для исследования единичных клеток важно учитывать качество ДНК и соотносить характеристики методов с поставленной задачей.

Список литературы

1. Nelson D.L., Ledbetter S.A., Corbo L., et al. Alu polymerase chain reaction: a method for rapid isolation of human-specific sequences from complex DNA sources // *Proc Natl Acad Sci U S A*, Vol. 86, No. 17, 1989. pp. 6686-90.
2. Korenberg J.R., Rykowski M.C. Human genome organization: Alu, lines, and the molecular structure of metaphase chromosome bands // *Cell*, Vol. 53, No. 3, 1998. pp. 391-400.
3. Ludecke H.J., Senger G., Claussen U., Horsthemke B. Cloning defined regions of the human genome by microdissection of banded chromosomes and enzymatic amplification // *Nature*, Vol. 338, No. 6213, 1989. pp. 348-350.
4. Zhang L., Cui X., Schmitt K., Hubert R., Navidi W., Arnheim N. 85. Zhang L, Cui X, SWhole genome amplification from a single cell: implications for genetic analysis // *Proc Natl Acad Sci USA*, Vol. 89, No. 13, 1992. pp. 5847-5851.
5. Dietmaier W., Hartmann A., Wallinger S., Heinmoller E., Kerner T., Endl E., Jauch K.W., Hofstadter F., Ruschoff J. Multiple mutation analyses in single tumor cells with improved whole genome amplification // *Am J Patho*, Vol. 154, No. 1, 1999. pp. 83-95.
6. Telenius H., Carter N., Bebb C.E., Nordenskjold M., Ponder B.A., Tunnacliffe A. Degenerate oligonucleotide-primed PCR: general amplification of target DNA by a single degenerate primer // *Genomics*, Vol. 13, No. 3. pp. 718-725.
7. Kittler R., Stoneking M., Kayser M. A whole genome amplification method to generate long fragments from low quantities of genomic DNA // *Anal. Biochem*, Vol. 300, No. 2, 2002. pp. 237-244.

8. Bonnette M.D., Pavlova V.R., Rodier D.N., Thompson L.P., Boone E.L., Brown K.L., Meyer K.M., Trevino M.B., Champagne J.R., Cruz T.D. dcDegenerate oligonucleotide primed-PCR for multilocus, genome-wide analysis from limited quantities of DNA // *Diagn Mol Pathol*, Vol. 18, No. 3, 2009. pp. 165-175.
9. Zong C., Lu S., Chapman A.R., Xie X.S. Genome-wide detection of single-nucleotide and copy-number variations of a single human cell. // *Science*, Vol. 338, No. 6114, 2012. pp. 1622-1626.
10. 1Dean F.B., Hosono S., Fang L., Wu X., Faruqi A.F., Bray-Ward P., Sun Z., Zong Q., Du Y., Du J., Driscoll M., Song W., Kingsmore S.F., Egholm M., Lasken R.S.. Comprehensive human genome amplification using multiple displacement amplification // *Proc Natl Acad Sci U S A*, Vol. 99, No. 8, 2002. pp. 5261-5266.
11. Spits C., Le Caignec C., De Rycke M., Van Haute L., Van Steirteghem A., Liebaers I., Sermon K. Whole-genome multiple displacement amplification from single cells // *Proc Natl Acad Sci USA*, Vol. 1, No. 4, 2006. pp. 1965-1970.
12. Zheng Y.M., Wang N., Li L., Jin F. Whole genome amplification in preimplantation genetic diagnosis // *J Zhejiang Univ Sci B*, Vol. 12, No. 1. pp. 1-11.
13. Cheung V.G., Nelson S.F. Whole genome amplification using a degenerate oligonucleotide primer allows hundreds of genotypes to be performed on less than one nanogram of genomic DNA // *Proc Natl Acad Sci USA*, Vol. 93, No. 25, 1996. pp. 14676-14679.
14. Lledo B., Ten J, Galan FM, Bernabeu R. Preimplantation genetic diagnosis of Marfan syndrome using multiple displacement amplification // *Fertil Steril*, Vol. 86, No. 4, 2006. pp. 945-955.
15. Sabina J., Leamon J.H. Bias in Whole Genome Amplification: Causes and Considerations // *Methods Mol Biol*, Vol. 1347, 2015. pp. 15-41.
16. Pirker C., Raidl M., Steiner E., Elbling L., Holzmann K., Spiegl-Kreinecker S., Aubele M., Grasl-Kraupp B., Marosi C., Miksche M., Berger W. Whole genome amplification for CGH analysis: Linker-adaptor PCR as the method of choice for difficult and limited samples. // *Cytometry A*, Vol. 61, No. 1, 2004. pp. 26-34.
17. Borgstrom E., Paterlini M., Mold J.E., Frisen J., Lundberg J. Comparison of whole genome amplification techniques for human single cell exome sequencing // *PLoS One*, Vol. 12, No. 2, 2017. P. e0171566.
18. Huang L., Ma F., Chapman A., Lu S., Xie X.S. Single-Cell Whole-Genome Amplification and Sequencing: Methodology and Applications. // *Annu Rev Genomics Hum Genet*, No. 16, 2015. pp. 79-102.