

Р. Н. Календарь, Ю. М. Сиволап

ПОЛИМЕРАЗНАЯ ЦЕПНАЯ РЕАКЦИЯ С ПРОИЗВОЛЬНЫМИ ПРАЙМЕРАМИ

Полимеразная цепная реакция с произвольными праймерами имеет ряд теоретических и методических особенностей. Для учета некоторых из них предложена математическая модель и описана динамика образования продукта реакции. Приведена формула для вычисления температуры плавления олигонуклеотидных праймеров, которая необходима при расчете оптимальной температуры отжига. На базе идеализированной модели геномной ДНК разработаны алгоритм и компьютерная программа поиска произвольных праймеров. Учтена статистическая закономерность определенных сочетаний нуклеотидов в праймере, указывающая на его специфичность при выявлении генетического полиморфизма. Показана возможность целенаправленного моделирования нуклеотидной последовательности праймеров.

Введение. Полимеразная цепная реакция (ПЦР) является специфическим высокочувствительным методом, который находит применение для решения широкого спектра проблем биологии [1—3]. ПЦР с произвольными праймерами на основе выявляемого генетического полиморфизма используется в качестве быстрого и чувствительного метода исследования эволюционных взаимоотношений, для генного картирования и других целей [4, 5]. Молекулярными маркерами в этой технологии служат участки ДНК, обнаруживающие полиморфные сайты с помощью ПЦР.

Варианты использования амплификации могут быть сгруппированы согласно особенностям самого процесса [6]. Амплификация со специфическими праймерами требует в качестве необходимого условия знания нуклеотидной последовательности амплифицируемого участка для подбора соответствующих праймеров к фланкирующим последовательностям. Амплификация с перемежающимися повторами (Interspersed Repetitive Sequences IRS) Alu-PCR или REP-PCR детерминруется последовательностью ДНК в многочисленных сайтах обеих цепей ДНК [7, 8]. Амплификация с вырожденными праймерами может проходить как недетерминированный процесс — случайнопраймированная амплификация (RPA), праймер-расширенная амплификация (PER) и случайная ПЦР (rPCR) [9—11]. Такой тип обусловлен реассоциацией праймера и матрицы в произвольных локусах, что приводит к стохастической амплификации фрагментов нуклеиновых кислот или даже целого генома со случайного сайта. Случайная амплификация обычно используется при радиоактивном или флуоресцентном мечении ДНК, для увеличения количества ДНК при конструировании библиотек кДНК или при ПЦР-типировании одиночных гаплоидных клеток.

Степень генетического родства индивидуальных организмов или видов можно выявить, сравнивая длину или нуклеотидный состав определенных фрагментов ДНК. Идентификация подобных молекулярных маркеров при анализе RFLP требует информации о нуклеотидном составе последовательности ДНК, клонирования и других процедур, связанных со значительной экспериментальной работой. Некоторое упрощение достигается за счет применения техники многочисленных профилей произвольно амплифицированных ампликонов (МААР, Multiple

Arbitrary Amplicon Profiling) [2], которая подобна методам, выявляющим полиморфизм случайно амплифицированной ДНК, — RAPD (Random Amplified Polymorphic DNA), произвольно праймированной ПЦР (AP-PCR Arbitrarily Primed PCR) и фингерпринт амплифицированной ДНК (DAF — DNA Amplification fingerprinting [4, 5, 13]. Эти независимо разработанные подходы основаны на использовании одного или более произвольных олигонуклеотидных праймеров для амплификации

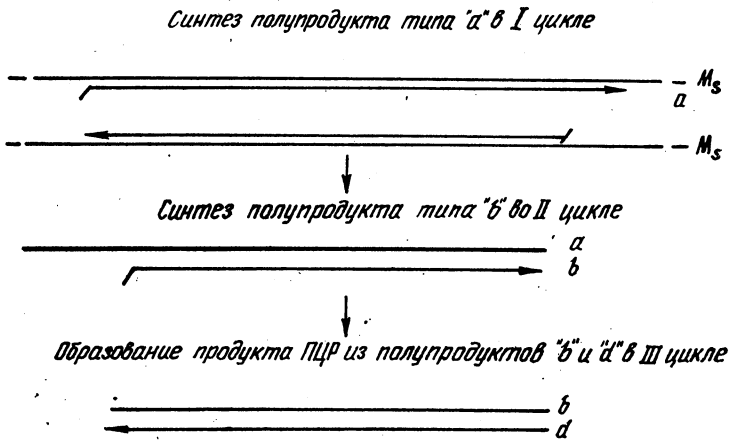


Рис. 1. Схема образования продукта полимеразной цепной реакции

определенных участков ДНК, часто полиморфных, с неизвестной локализацией в геноме. Выявление полиморфизма длин амплифицированных фрагментов (AFLPs — Amplification Fragment Length Polymorphisms) может быть удобным и простым способом идентификации и картирования генов [15, 16].

Техника МААР различается по количеству продуктов амплификации, длине праймеров, условиям амплификации и способу электрофоретического разделения продуктов. К примеру, в AP-PCR используются праймеры длиной, сравнимой с таковой при обычной ПЦР; в анализе RAPD применяются короткие праймеры около 10 нуклеотидов, в DAF размер праймеров составляет не менее 5 нуклеотидов (обычно 7—8) [4, 5, 13]. Эти и другие различия в конечном результате обуславливают соответствующие условия проведения электрофореза для получения высокоразрешающих профилей продуктов амплификации. Разделение продуктов в RAPD и AP-PCR достигается электрофорезом в агарозном геле с окрашиванием бромистым этидием. Визуализацию продуктов DAF и информацию о ДНК-профилях получают с помощью электрофореза в полиакриламидном геле и окраской серебром для детекции пикограммового уровня ДНК [17].

Моделирование полимеразной цепной реакции. Для начала необходимо определить понятия, которые будут использованы в описании механизма ПЦР с произвольными праймерами.

Основной продукт ПЦР (ампликон) — это двухцепочечная молекула ДНК, концы которой фланкированы последовательностью праймеров и комплементарного им участка ДНК.

Полупродукт ПЦР — это одноцепочечная молекула ДНК, содержащая на конце последовательность праймера. Молекулы полупродуктов участвуют в образовании основного продукта и других полупродуктов реакции (рис. 1). Накопление последних происходит за счет образования новых молекул полупродуктов в каждом цикле.

Цикл ПЦР начинается с денатурации ДНК, затем отжига праймера на ДНК-матрицу и завершается синтезом полупродукта ДНК-полимеразой. ПЦР с произвольными праймерами обычно происходит в присутствии одного праймера, который выполняет функцию прямого

и обратного. Для амплификации участка ДНК праймер образует дуплексы с обеими комплементарными нитями ДНК. 3'-концы праймера ориентированы внутрь амплифицируемого фрагмента (продукта). Другие способы ориентации праймера относительно комплементарных нитей ДНК не приводят к амплификации основного продукта.

Полупродукты полимеразной цепной реакции. Для упрощения моделирования полупродукты амплификации ПЦР были дифференцированы по двум признакам: по уровню их образования и степени комплементарности концевых участков.

Образование полупродуктов зависит от типа матрицы, с которой снята копия данного полупродукта (матрицей может служить как реномная ДНК, так и синтезированные ранее полупродукты).

Каждый тип полупродукта включает комплементарные нити ДНК, синтезированные с комплементарных матриц.

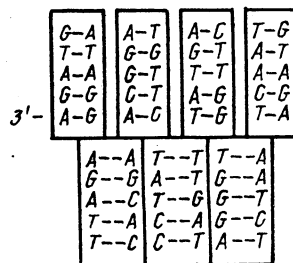


Рис. 2. Формула 8-членного праймера для последовательности ДНК с CG-составом, равным 49,9%. Показаны пять наиболее часто встречающихся пар нуклеотидов в данных положениях праймера

В формулах, описывающих динамику формирования каждого типа полупродуктов ПЦР, учитывается тип матриц, участвующих при их образовании в конкретном цикле n и оценивается вклад других полупродуктов из предшествующих циклов. Тип данного полупродукта не изменяется в течение всей реакции. Формулы состоят из двух частей: одна описывает количество образованных молекул полупродукта в данном цикле, вторая — количество полупродукта, полученного из предыдущих циклов.

Из рис. 1 видно, что в первом цикле образуются молекулы, более длинные, чем продукт, связанный с ДНК-матрицей. Назовем этот тип полупродукта «а». Количество молекул «а», синтезируемых в каждом цикле реакции, одинаково. Сумма молекул комплекса «а» + « M_s » — $S(\langle a \rangle + \langle M_s \rangle)$ (где « M_s » — матричная ДНК в одноцепочечной форме) — равна количеству матричной ДНК $S(\langle M_s \rangle)$. Длина полупродукта типа «а» больше, чем длина последовательности основного продукта, так как терминация синтеза ограничена пределами нити ДНК или ее денатурацией.

Суммарное количество молекул «а» — $S(a_n)$, образованных за n циклов, описывается следующей формулой:

$$S(a_n) = M_s + S(a_{n-1}); \quad (1)$$

$$S(a_n) = 2nM_d. \quad (1.1)$$

Для остальных формул количество матричной ДНК — M_d (M_d в двухцепочечной форме) принято равным единице, в противном случае значенные формулы динамики образования полупродуктов необходимо умножить на соответствующее количество матричной ДНК.

Во втором цикле из молекул полупродукта типа «а» предыдущего синтеза $S(a_{n-1})$ образуется новый тип полупродукта — «б». Данный тип полупродукта короче, чем молекулы «а», и в следующем цикле образует продукт реакции.

Суммарное количество молекул типа «б» — $S(b_n)$ (арифметическая прогрессия), образованных за n циклов:

$$S(b_n) = S(a_{n-1}) + S(b_{n-1}); \quad (2)$$

$$S(b_{n-1}) = n^2 - 3n + 2; \quad (2.1)$$

$$S(b_n) = M_d(n^2 - n). \quad (2.2)$$

На 5'-конце праймера возможны некомплементарные матричные нуклеотиды, характерные только для молекул полупродуктов типов «a» и «b».

В третьем цикле из молекул «b» синтезируются молекулы «d», равные по длине молекулам «b». С третьего цикла впервые образуется продукт, который представлен двухцепочечной молекулой ДНК, содержащей цепь молекул «b» и «d».

Суммарное количество синтезированных молекул типа «d» — $S(d_n)$ за n циклов (геометрическая прогрессия):

$$S(d_n) = S(b_{n-1}) + S(d_{n-1}); \quad (3)$$

$$S(d_{n-1}) = 2^n - n^2 + n - 2; \quad (3.1)$$

$$S(d_n) = M_d(2^{n+1} - n^2 - n - 2). \quad (3.2)$$

Молекулы типов «b» и «d» одинаковы по нуклеотидному составу, но происходят от разных матриц (рис. 1, табл. 1).

Особенности молекул этих типов состоят в следующем: во-первых, полупродукт «b» синтезируется на матрице «a» с образованием комплекса «a»+«b» по структуре, не соответствующей характеристике продукта ПЦР (длина молекул типа «a» всегда больше длины продуктов амплификации); во-вторых, матрицей полупродукта «d» является «b», последовательности обеих полностью комплементарны, и комплекс «b»+«d» представляет собой полноценный продукт; синтез «d» происходит в третьем цикле и косвенно связан с полупродуктом «a» через «b». Для удобства описания образования продукта ПЦР из полупродукта типа «b» недостаточно, и поэтому был введен полупродукт типа «d», идентичный ему по нуклеотидной последовательности.

Логические формулы (1—3) объясняют происхождение данного типа полупродукта, а формулы (1.1), (2.2), (3.2) — пригодны для количественных расчетов.

Продукт полимеразной цепной реакции. В образовании основного продукта участвуют молекулы типов «b» и «d», синтезированные в предыдущих циклах, т. е. $S(b_{n-1})$ и $S(d_{n-1})$. Образование продукта ПЦР можно описать как сумму:

$$P_d = S(b_{n-1}) + S(d_{n-1}), \quad (4)$$

где P_d — продукт ПЦР (если количество матрицы $M_d=1$).

Из всех синтезированных в ПЦР молекул только комплексы типов «a+ M_s » и «a»+«b» не являются продуктом реакции и поэтому их количество снижает значение 2^n к значению $2^n - 2n$. Количество ком-

Таблица 1
Динамика образования продукта и полупродуктов ПЦР

Число циклов	Количество полупродуктов*			Продукт*
	«a»	«b»	«d»	
1	2	—	—	—
2	4	2	—	—
3	6	6	2	2
4	8	12	10	8
5	10	20	32	22
6	12	30	84	52
7	14	42	198	114
8	16	56	438	240
20	40	380	2096730	1048536

* $M_d=1$.

Таблица 2
Оптимальная температура отжига

Последовательность (5'—3')	Ta* min—max, °C
CCATGGTACCCGGATCCTCG	55,74—62,81
ATTAACCCCTCACTAAAGGGA	53,00—60,07
GAGCAAGTTCAGCCTGG	52,33—59,41
GACAGACAGACAGACA	48,75—55,82
CGACACGCTG	48,06—55,14
CGACACGC	46,82—53,90
CTGCGTGC	46,87—52,52
ACTTCGTC	45,02—52,10

* Та min — max — оптимальная температура отжига, если продуктом реакции являются молекулы длиной от 300 до 2000 п. н. и (G+T+C)-составом от 45 до 60 %.

плекса $S(\langle a \rangle + \langle M_s \rangle)$ равно таковому $2M_d$, а $S(\langle a \rangle + \langle b \rangle) = 2(n-1)$. Сумма $(\langle a \rangle + \langle M_s \rangle)$ и $S(\langle a \rangle + \langle b \rangle)$ дает $2(M_d + n - 1)$, если $M_d = 1$, то сумма равна $2n$. Образование продукта можно описать следующим образом:

$$P_d = M_d(2^n - 2n) \quad (5)$$

или

$$S(b_{n-1}) + S(d_{n-1}) = 2^n - 2n. \quad (5.1)$$

Данная зависимость объективна для идеальных условий, когда все компоненты реакции (dNTP, Taq-полимеразы, праймеры) в избытке. При недостатке какого-либо компонента эффективность образования продукта ПЦР снижается.

Предложенная математическая модель динамики образования продукта и других молекул реакции амплификации необходима для описания начальных циклов. Накопление молекул полупродуктов типов «a» и «b» незначительно сказывается на общем результате:

$$1/2(S(b_n) + S(d_n)) - P_d = n - 1.$$

Взаимодействие праймера с матрицей. В результате узнавания ДНК-полимеразой комплекса, сформированного при отжиге праймера с сайтом-мишенью, реакция носит специфический характер. Это выражается в предпочтительной амплификации определенных участков ДНК-матрицы и создает воспроизводимую картину полиморфизма [6].

Наличие длинных концевых областей с неполной комплементарностью праймера и матрицы препятствует аффинному взаимодействию фермента с таким комплексом и затрудняет полимеризацию.

Наличие ошибки при связывании праймера с мишенью влияет на воспроизводимость продуктов амплификации и эффективность работы полимеразы, которая узнает и удлиняет «ошибочный» комплекс.

Первые восемь нуклеотидов с 3'-конца праймера включают район с высокой степенью комплементарности к матрице. Единичная замена в этом районе значительно сказывается на результатах амплификации. Поэтому для успешной амплификации на 3'-конце праймера необходимо не менее пяти — шести нуклеотидов, полностью комплементарных матрице [6]. Увеличение праймера на два или три нуклеотида не изменяет количества амплифицированного продукта [6].

Условия проведения полимеразной цепной реакции. На взаимодействие между отдельными нитями, дуплексами, петлями шпилек и другими типами молекул, формируемых в течение МААР, влияет ряд параметров. Использование произвольных праймеров требует определенных условий реакции, которые обуславливают воспроизводимость результатов. Для поддержания необходимого равновесия эти условия должны быть тщательно оптимизированы. Такие параметры, как время и температура отжига, элонгации, денатурации, число циклов, скорость нагрева и охлаждения, концентрация праймеров, матрицы, ионов, ДНК-полимеразы, от которых зависит результат амплификации, по нашему мнению, должны быть стандартизированы для простоты этой технологии. Надо учитывать и качество геномной ДНК.

Особенность рассматриваемой ПЦР, как уже упоминалось, состоит в том, что праймер может быть неполностью комплементарным последовательности ДНК-матрицы при образовании молекул типа «a» на матричной ДНК и типа «b» на матрице «a». Синтез других типов полупродуктов происходит с полностью комплементарной матрицы. Для проведения амплификации необходимо, чтобы в течение первых трех циклов условия отжига были оптимальными для образования неполностью комплементарных дуплексов. После третьего цикла амплификации температуру отжига необходимо увеличить до оптимальной, чтобы в реакцию были вовлечены только молекулы типов «d» и «b» предшествующих циклов. Это позволит уменьшить фоновую амплификацию: синтез молекулы типа «a» и синтез «b» на матрице «a». Боль-

шая часть молекул типа «а» не участвует в амплификации основных продуктов реакции, так как не все варианты амплифицируемых молекул «а» являются матрицами для синтеза молекул типа «b» в связи с отсутствием комплементарных участков для праймера.

Вычисление оптимальной температуры отжига для основных циклов ПЦР производится по формуле, предложенной авторами работы [18]:

$$T_{a}^{\text{опт}} = 0,3 T_{m}^{\text{праймера}} + 0,7 T_{m}^{\text{продукта}} - 14,9. \quad (6)$$

Температуру плавления продукта определяли по стандартной формуле, используемой при гибридизации по Саузерну:

$$T_{m}^{\text{продукта}} = 81,5 + 16,6 \lg [K^{+}] + 0,41 [G + C \%] - 675/L_{p}, \quad (7)$$

где $[K^{+}]$ — концентрация ионов калия в молях; $[G + C \%]$ — количество G и C в продукте ПЦР в процентах; L_{p} — длина продукта в парах нуклеотидов.

При вычислении температуры плавления праймера применяли теорию перекрывающихся соседних пар нуклеотидов [19]. Каждой паре нуклеотидов соответствуют свои термодинамические параметры, сумма которых влияет на свойства праймера. Для вычислений использовали свободную энергию Гиббса (dG) для каждой из 16 п. н. [19]:

$$T_{m}^{\text{праймера}} = -k \frac{dG}{L^{3,3/13}}, \quad (8)$$

где $k = 2,9136826$, если концентрация ионов калия в растворе равна 50 мМ; L — длина праймера (в нуклеотидах). В общем же случае формула выглядит следующим образом:

$$T_{m}^{\text{праймера}} = -3,9 \frac{dG}{L^{3,3/13}} + 16,6 \lg [K^{+}]. \quad (8.1)$$

Для стандартной амплификации точно вычисленная оптимальная температура отжига определяет чувствительность амплификации и ее специфичность. Ошибка в три градуса существенно сказывается на результате амплификации, что выражается в уменьшении выхода основного и появлении неспецифического продуктов [14, 18, 20]. Оптимальная температура отжига в AP-PCR различается для коротких продуктов ПЦР и для продуктов с максимальной длиной при прочих равных условиях. Максимальная разница может составить 6 °C (табл. 2). Отжиг можно проводить только в диапазоне температуры, оптимальном для конкретного праймера. Для праймеров равной длины существует общий диапазон температуры отжига, в пределах которого происходит их отжиг с матрицей.

Условием успешной амплификации с произвольными праймерами является высокая концентрация ионов магния (при стандартной амплификации она составляет 1,5 мМ, для МААР — 4 мМ). Такая концентрация способствует инициации синтеза продукта ПЦР с неполностью комплементарного комплекса праймера и мишени при температуре отжига, ниже оптимальной. Отжиг первых трех циклов AP-PCR и RAPD следует проводить при температуре, ниже оптимальной на 10 °C.

Для синтеза продукта амплификации необходима высокая концентрация праймера. Для AP-PCR она составляет 0,2—0,3 мкМ, для DAF — до 3 мкМ [6, 20]. Низкая концентрация праймера приводит к невозможности воспроизведения результатов.

На воспроизводимость результатов влияет также качество используемых препаратов ДНК. Наличие плохо очищенной или гидролизованной ДНК-матрицы вызывает нестабильность воспроизводства продуктов реакции или непредсказуемость результатов [20].

Моделирование праймеров. Применяемые в технике МААР праймеры могут быть представлены нуклеотидной последовательностью с

GC-составом от 50 до 70 % и длиной не менее пяти нуклеотидов. Воспроизводимость результатов может быть связана с GC-составом праймера, но данная зависимость не распространяется на виды с различным GC-составом геномной ДНК [21, 22].

Нуклеотидная последовательность праймеров определяет степень выявляемого меж- и внутривидового полиморфизма. Одни праймеры выявляют межвидовой полиморфизм, но неприменимы на внутривидовом уровне; другие же выявляют внутривидовой полиморфизм и мало пригодны для межвидового исследования в связи с перекрывающимся полиморфизмом. Праймеры можно условно разделить на видовоспецифичные, внутривидоспецифичные и, возможно, специфичные для более высоких таксонов. Вероятно, существует определенная зависимость специфичности праймера от его нуклеотидной последовательности. Нами предложена гипотеза о наличии вышеуказанной специфичности произвольных праймеров и, следовательно, возможности их целенаправленного подбора на основе свойств пар соседних нуклеотидов. Суть ее состоит в следующем: нуклеотидная последовательность праймера с 3'-конца разбивается на пары нуклеотидов, смещенные на один нуклеотид к 5'-концу. Например, праймер 3'-ATTGCCAGTT образует пары:

1-я пара — AT; 2-я — TT; 3-я — TG; 4-я — GC; 5-я — CC; 6-я — CA;
7-я — AG; 8-я — GT; 9-я — TT.

Всего вариантов пар четырех нуклеотидов 16 (2⁴). Если проанализировать все праймеры, выявляющие внутривидовую специфичность (так как они представляют наибольший интерес для исследований), и частоту встречаемости определенных пар нуклеотидов в данном положении праймера, то полученные данные можно использовать для синтеза произвольных праймеров с экстраполированной последовательностью. Такие праймеры с большой вероятностью могут применяться при исследовании внутривидового полиморфизма. Подобную информацию можно получить, анализируя последовательность геномной ДНК, или из банков данных. Для этой цели разработана компьютерная программа, производящая поиск произвольных праймеров в данной последовательности и определяющая их нуклеотидный состав. Поиск произвольных праймеров происходит с учетом заданного числа ошибок комплементарности (для 8-нуклеотидного праймера максимальное число мисметчей равно 3, для 10 нуклеотидов соответственно до 5 и так далее) при необходимом условии полной комплементарности шести нуклеотидов на 3'-конце праймера.

Смоделирована гипотетическая последовательность геномной ДНК с максимальным разнообразием ее нуклеотидной последовательности. Последнее необходимо для увеличения разнообразия последовательности и уменьшения повторов. Частые повторы не участвуют в образовании продуктов AP-PCR, так как структура амплифицированных продуктов имеет петлеобразную форму, нехарактерную для высокоповторяющихся последовательностей. Поэтому при моделировании этой последовательности придерживались принципа максимального разнообразия нуклеотидного состава. Нуклеотидная последовательность «синтезировалась» генератором случайных чисел, где каждому варианту из трех нуклеотидных сочетаний соответствовало множество чисел, генерирующихся компьютером. Образование нуклеотидной цепи происходило за счет добавления сочетаний, «вызываемых» генератором случайных чисел. Например, «вызов» триплета AAA происходил только в случае, если генерировались числа — 1, 65, 129, 193, 321, 385, 449 и 513; для триплета AAT — 2, 66, 130, 194, 258, 322, 386, 450 и 514. Размерность множества генерируемых чисел составляла 576. Длина «синтезируемой» нуклеотидной последовательности была равна 7500 нуклеотидов. Ее было достаточно, чтобы программа отобрала 450 произвольных праймеров длиной 8 нуклеотидов без «ошибок» комплементарности.

тарности праймера с генерированной последовательностью. GC-состав «синтезируемых» нуклеотидных последовательностей составлял от 45 до 60 %.

Сравнение праймеров длиной 8 нуклеотидов, полученных из разных по GC-составу «синтезированных» последовательностей (различие в 5 %) показало, что величина сходства наиболее часто встречаемых пар сочетаний нуклеотидов (первые 9 пар из 16) приближается к половине, но для различных последовательностей с разным GC-составом эта величина ближе к 60 %. Данные приведены в табл. 4 (полужирным шрифтом выделены сходные пары). Таким образом, для произвольных праймеров, происходящих из контрастных по GC-составу исследуемых последовательностей, характерны разные сочетания соседних нуклеотидов. Наблюдается увеличение частоты встречаемости пар нуклеотидов, содержащих С или G, для последовательностей с GC-составом, выше 50 %, и А или Т для последовательностей с GC-составом, ниже 50 %. У произвольных праймеров из последовательностей с близким GC-составом (не более 5 % разницы) частота определенных пар соседних нуклеотидов приблизительно сходна (табл. 3, 4).

Для подбора праймеров необходимо учитывать GC-состав геномной ДНК исследуемого организма. Вероятность амплификации участков ДНК с контрастным содержанием GC-пар представляется низкой. Богатые А — Т последовательности не участвуют в амплификации, так как комплекс праймера с такими участками нестабилен и поэтому меньше вероятность, что они послужат затравкой для полимеразы. В среднем же можно ориентироваться на 50 %-й GC-состав геномной ДНК.

Таблица 3

Зависимость частоты сочетаний соседних пар нуклеотидов от их положения в 8-нуклеотидном праймере (GC-состав исследуемой модельной ДНК равен 49,9 %)

№ пары	Частота встречаемости пары нуклеотидов в праймере* (max→min)									Праймер (3'-)
	1	2	3	4	5	6	7	8	9	
1	ag	gg	aa	tt	ga	at	ta	gt	tc	AGGTTGAA
2	aa	gg	ac	ta	tc	ct	tt	tg	ga	GGATGTGT
3	ac	ct	gt	gg	at	aa	ta	ca	tt	GGA CTGAT
4	tt	at	tg	ca	ct	ag	ta	ga	tc	AACTGTAG
5	tg	ag	tt	gt	ac	gc	ta	ga	at	TTCTGTAG
6	ta	ga	gt	gc	at	ag	tg	cc	ca	GACTTGAG
7	ta	cg	aa	at	tg	ag	gt	ac	gg	GACAGTGT

* Число «ошибок» спаривания равно 0.

Таблица 4

Зависимость частоты сочетаний соседних пар нуклеотидов от их положения в 8-нуклеотидном праймере (GC-состав исследуемой модельной ДНК равен 50,89 %)

№ пары	Частота встречаемости пары нуклеотидов в праймере* (max→min)									Праймер (3'-)
	1	2	3	4	5	6	7	8	9	
1	gt	ag	aa	tt	ct	cg	ac	ca	gc	GTACCCTC
2	ta	gc	ag	ac	tt	cg	ca	gt	tc	GTTCGGAC
3	ca	ac	ta	gc	aa	ag	gt	cc	gg	GTACGGAC
4	ac	aa	cc	cg	gg	ag	tt	ct	gc	CTCAGGTC
5	cc	ta	cg	gg	ac	aa	gt	tc	ga	CAGTTCCA
6	ac	cc	ct	ga	gc	aa	tt	cg	tg	AGCACCTC
7	cc	ac	aa	ca	cg	ga	ct	gc	tc	AGTACCCA

* Число «ошибок» спаривания равно 0.

Отмечено, что GC-состав всех произвольных праймеров выше такового, характерного для модельной последовательности ДНК, примерно на 4 %.

Праймеры длиной 8 нуклеотидов подобны праймерам из 9 и более нуклеотидов, если они не содержат мистметчей. Этим праймерам характерны общие пары нуклеотидов, количество 8-нуклеотидных праймеров больше, чем праймеров длиной 9 нуклеотидов. Ситуация изменяется, если в 9-нуклеотидном праймере допускается хотя бы один мистметч на последних трех нуклеотидах 5'-конца. Тогда 9-нуклеотидных праймеров больше, и в них входит множество 8-нуклеотидных праймеров (табл. 3, 5). Например, последовательность длиной 5823 нуклеотида с GC-составом 49,9 % содержит 285 произвольных праймеров длиной 8 нуклеотидов, для этой же последовательности число праймеров длиной 9 нуклеотидов составит 600. При одной и той же температуре отжига (максимально жесткой для 8-нуклеотидного и близкой к строгой для 9-нуклеотидного праймеров) в первых циклах предпочтительно применение праймеров длиннее 8 нуклеотидов при исследовании полиморфизма внутри вида или на уровне особи. При проведении отжига в максимально мягких условиях предпочтительнее короткие праймеры (8—10 нуклеотидов), так как комплекс праймера с мишенью более стабилен. У длинных праймеров (например, 12 и более нуклеотидов) в условиях, допускающих комплементарность только 6 нуклеотидам на 3'-конце праймера, значение dG увеличивается и, по-видимому, такой комплекс не будет образовываться.

Таблица 5

Зависимость частоты сочетаний соседних пар нуклеотидов от их положения в 9-нуклеотидном праймере (GC-состав исследуемой модельной ДНК равен 49,9 %)

№ пары	Частота встречаемости пары нуклеотидов в праймере* (max→min)									Праймер (3'-)
	1	2	3	4	5	6	7	8	9	
1	aa	tt	ga	ta	gt	tc	gg	ag	at	AACCTGTAT
2	aa	gg	ia	ac	tt	tc	ct	tg	at	AACCCAGTAG
3	ac	at	gg	ta	gt	ag	cc	tt	tc	TTCCTGTAT
4	at	tt	ct	ca	tg	gc	ta	cc	ga	GAAGCGTAC
5	tg	tt	ta	cg	ag	gt	gc	cc	ac	GACCTGTAG
6	ta	gt	ga	cc	at	cg	ag	ac	tg	GAACTGTAG
7	aa	ta	cg	ag	at	gg	tc	tg	cc	GACAGCCCC
8	cc	gt	ta	gg	at	ag	ga	ac	cg	GTTTGTAGG

* Число «ошибок» спаривания равно 1.

В табл. 3—5 приведены произвольные праймеры, полученные в результате моделирования наиболее встречаемых пар нуклеотидов в соответствующих положениях праймера. Эти и подобные им праймеры, вероятнее всего, могут выявлять внутривидовой полиморфизм.

Заключение. Техника МААР является относительно простым и многосторонним диагностическим приемом, подходящим для решения большого числа биологических задач. Амплификация ДНК с использованием произвольных праймеров перспективна при идентификации генов, анализе популяций и потомства и геном картировании. При помощи этого подхода возможно изучение геномов про- и эукариот без предварительной информации относительно их последовательности. МААР представляет новый молекулярно-генетический подход для успешного и быстрого получения ДНК-маркеров. Произвольные праймеры приобретают статус предсказуемых с предопределенной специфичностью. Не вызывает сомнений, что техника МААР быстро заменит менее эффективные методы молекулярно-генетического маркирования.

ПОЛИМЕРАЗНА ЛАНЦЮГОВА РЕАКЦІЯ З ДОВІЛЬНИМИ ПРАЙМЕРАМИ

Резюме

Полимеразна ланцюгова реакція з довільними праймерами має ряд теоретичних та методичних особливостей. Для врахування деяких з них було запропоновано її математичну модель і показано динаміку утворення продукта реакції та інших молекул, які беруть участь у цьому процесі. Наведено формулу для обчислення температури плавлення олігонуклеотидних праймерів, що необхідно при розрахунку оптимальної температури реасоціації. На базі ідеалізованої моделі геномної ДНК було розроблено алгоритм та комп'ютерну програму пошуку довільних праймерів. Враховано статистичну закономірність певних поєднань нуклеотидів у праймері, що свідчить про його специфічність при виявленні генетичного поліморфізму. Показано можливість цілеспрямованого моделювання праймерів.

R. N. Calendar, Yu. M. Syvolap

POLYMERASE CHAIN REACTION WITH ARBITRARY OLIGONUCLEOTIDE PRIMERS

Summary

Polymerase chain reaction with arbitrary oligonucleotide primers has some theoretical and methodics peculiarities. For resolution of some from them was proposed the mathematical model of PCR reaction and dynamics of formation of the reaction product and other molecules taking part in this process. The formula for calculation temperature of oligonucleotide primers melting was proposed to calculate of the optimal temperature of annealing. The idealize model of the genome DNA and computer program of screening of all arbitrary were created. The statistical regularity of appointed combinations of nucleotides in primers showed; it is specific for reveal of the polymorphism. Hence in follows possibility of direct designing primers was showed.

СПИСОК ЛІТЕРАТУРИ

1. Mullis K. B., Faloona F. A., Scharf S. et al. Specific enzymatic amplification of DNA *in vitro*: The polymerase chain reaction // Cold Spring Harbor Symp. Quant. Biol.—1986.—51.—P. 263—273.
2. Деґабов В. Г. Локальная амплификация нуклеиновых кислот — новый метод исследования // Молекуляр. биология.—1990.—24.—С. 304—309.
3. Варганетян А. Б. Полимеразная цепная реакция // Там же.—1991.—25.—С. 936—937.
4. Williams J. G. K., Kubelic A. R., Livak K. J. et al. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers // Nucl. Acids Res.—1990.—18.—P. 6531—6535.
5. Welsh J., McClelland M. Fingerprinting genomes using PCR with arbitrary primers // Ibid.—P. 7213—7218.
6. Caetano-Anolles G. Amplifying DNA with arbitrary oligonucleotide primers // PCR Meth. and Applications.—1993.—3.—P. 85—94.
7. Nelson D. L., Ledbetter S. A., Cordo L. et al. Alu polymerase chain reaction: A method for rapid isolation of human-specific sequences from complex DNA sources // Proc. Nat. Acad. Sci. USA.—1989.—86.—P. 6686—6690.
8. Versalovic J., Koeuth T., Lupski J. R. Distribution of repetitive DNA sequences in eubacteria and applications to fingerprinting of bacterial genomes // Nucl. Acids Res.—1991.—19.—P. 6823—6831.
9. Pat. N 5.043.272. USA. Amplification of nucleic acid sequences using oligonucleotides of random sequence as primers / J. L. Hartley, 1991.
10. Zhang L., Cui X., Schmitt K. et al. Whole genome amplification from a single cell: Implications for genetic analysis // Proc. Nat. Acad. Sci. USA.—1992.—89.—P. 5847—5851.
11. Froussard P. rPCR: A powerful tool for random amplification of whole RNA sequences // PCR Meth. and Applications.—1993.—2.—P. 185—190.
12. Caetano-Anolles G., Bassam B. J., Gresshoff P. M. DNA fingerprinting: MAAPing out a RAPD redefinition? // Bio/Technology.—1992.—10.—P. 937.
13. Caetano-Anolles G., Bassam B. J., Gresshoff P. M. DNA amplification fingerprinting using very short arbitrary oligonucleotide primers // Ibid.—1991.—9.—P. 553—557.

14. *Сиволап Ю. М., Календарь Р. Н.* Генетический полиморфизм ячменя, детектируемый ПЦР с произвольными праймерами // *Генетика*.—1994.—**31**.—С. 12—18.
15. *Caetano-Anolles G., Bassam B. J., Gresshoff P. M.* DNA amplification fingerprinting: A strategy for genome analysis // *Plant. Mol. Biol. Rep.*—1991.—**9**.—P. 292—305.
16. *Tingey S. V., gel Tufo J. P.* Genetic analysis with random amplified polymorphic DNA markers // *Plant Physiol.*—1993.—**101**.—P. 349—352.
17. *Bassam B. J., Caetano-Anolles G., Gresshoff P. M.* Fast and sensitive silve staining of DNA in polyacrylamide gels // *Anal. Biochem.*—1991.—**80**.—P. 81—84.
18. *Rychlik W., Soencer W. Y., Rhoads R. F.* Optimization of the annealing temperature for DNA amplification in vitro // *Nucl. Acids Res.*—1990.—**18**.—P. 6400—6414.
19. *Kenneth J. B., Frank R., Blocker H., Marky L.* Predicting DNA duplex stability from the base sequence // *Proc. Nat. Acad. Sci. USA.*—1986.—**83**.—P. 3746—3750.
20. *Сиволап Ю. М., Календарь Р. Н.* Исследование генетического полиморфизма злаковых растений при помощи ПЦР со случайными праймерами // *Цитология и генетика*.—1994.—№ 6.—С. 54—61.
21. *Ferrenberg A. M., Landau D. P., Wong Y. J.* Monte Carlo simulations: Hidden errors from «good» random number generators // *Phys. Rev. Lett.*—1992.—**69**.—P. 3382—3384.
22. *Hayes B.* The wheat of fortune // *Amer. Sci.*—1993.—**81**.—P. 114—118.

Селекц.-генет. ин-т Укр. Академии Аграр. наук, Одесса

Получено 18.10.94