

МОДЕЛИРОВАНИЕ РЕАКЦИИ АМПЛИФИКАЦИИ ДНК В РАМКАХ  
ТЕОРИИ ВЕТВЯЩИХСЯ ПРОЦЕССОВ С ДВУМЯ ТИПАМИ ЧАСТИЦ© 2010 г. Д. Г. Сочивко, А. А. Фёдоров, Д. А. Варламов,  
В. Е. Курочкин, академик Р. В. Петров

Поступило 27.05.2010 г.

Полимеразная цепная реакция (ПЦР) является сегодня основным молекулярно-генетическим методом, используемым для качественного и количественного анализа специфических последовательностей нуклеиновых кислот. В основе ПЦР лежит разветвленная цепная реакция репликации двухцепочного фрагмента ДНК (амплификация фрагмента). Реакцию проводят путем циклического повторения двух шагов: 1) денатурация – комплементарные цепи амплифицируемого двухцепочного фрагмента расходятся при нагревании реакционной смеси, 2) элонгация – короткие специфические цепи ДНК (праймеры), присутствующие в избытке, связываются с одиночными цепями фрагмента при охлаждении реакционной смеси, и на матрице одиночной цепи происходит синтез второй, комплементарной цепи ДНК. Результатом этих двух шагов является удвоение двухцепочной молекулы исходного амплифицируемого фрагмента. Таким образом, количество копий амплифицируемого фрагмента после  $N$  циклов амплификации можно выразить как  $A_N = A_0 \cdot 2^N$ , где  $A_0$  – исходное количество копий ДНК,  $N$  – количество циклов ПЦР.

Важной чертой ПЦР, отличающей ее от типичных ветвящихся цепных реакций, протекающих в непрерывном времени, является циклический механизм ее проведения, позволяющий синхронизовать пошаговое умножение всех имеющихся в системе молекул. Если в процессе реакции производить количественную регистрацию ее продуктов на каждом цикле, то получим кинетическую кривую ПЦР, имеющую дискретный по времени характер. При этом, задав некоторое пороговое значение количества продукта  $A_t$ , получим, что реакции с раз-

личным исходным количеством фрагмента  $A_0$  достигнут порогового уровня за  $N = \lg_2 \frac{A_t}{A_0}$  циклов амплификации. Этот подход лежит в основе количественного анализа фрагментов ДНК с применением регистрации продуктов ПЦР в реальном времени (ПЦР-РВ).

Ключевым допущением описанного подхода является предположение об удвоении числа амплифицируемых фрагментов ДНК на каждом цикле реакции. Если бы это предположение было истинным, процесс ПЦР протекал бы строго детерминировано, и по количеству продукта на определенном цикле реакции можно было бы однозначно определить исходную концентрацию исследуемой ДНК. Однако в реальности длительность каждого шага реакции ограничена, при этом все процессы протекают с некоторой конечной скоростью. В результате удвоение каждой копии фрагмента на каждом цикле происходит с вероятностью  $p < 1$  и ход ПЦР необходимо описывать в рамках теории случайных ветвящихся процессов с дискретным временем.

Теоретическое исследование ПЦР как ветвящегося процесса проводили в нескольких работах [1, 2]. Общий подход этих работ состоял в том, что вероятность удвоения цепей принималась за постоянную  $p$ , тогда для цикла  $N$  и начального количества копий  $A_{N-1}$  количество копий фрагмента  $A_N$ , синтезированного к концу цикла, будет представлять собой дискретную случайную величину. При этом распределение данной случайной величины определялось с помощью математического аппарата производящих функций для частиц одного типа [3]. Производящую функцию первого цикла можно записать как

$$g_1(s) = A_0 s(q + ps),$$

где  $p$  и  $q$  – вероятности удвоения и не удвоения соответственно,  $A_0$  – исходное количество частиц. То-

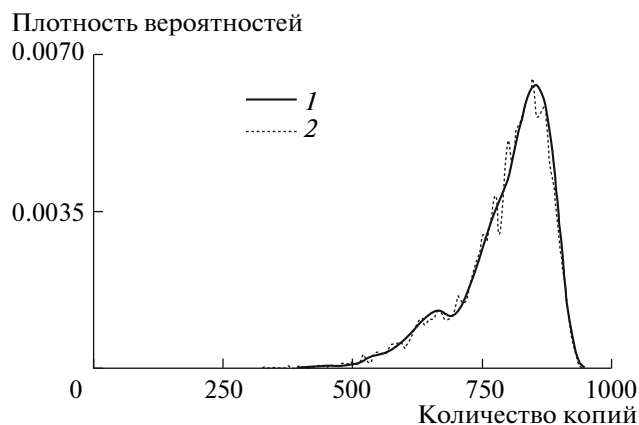
ЗАО “Синтол”, Москва

Институт аналитического приборостроения

Российской Академии наук,

Санкт-Петербург

Институт иммунологии, Москва



**Рис. 1.** Распределения вероятностей количества фрагментов после 10 циклов амплификации, полученные с использованием аналитической (1) и имитационной моделей (2) с двумя типами цепей. Параметры  $P_A = P_B = 0.95$ , исходное количество копий – 1.



**Рис. 2.** Распределения вероятностей количества фрагментов после 10 циклов амплификации для аналитических моделей с одним типом (1) и двумя типами цепей (2). Параметры  $P_A = P_B = 0.95$ , исходное количество копий – 1.

гда производящую функцию для  $N$ -го цикла можно получить с помощью рекуррентного соотношения

$$g_N(s) = qg_{N-1}(s) + pg_{N-1}(s)^2.$$

Функцию  $g_N(s)$  можно представить в виде

$$g_N(s) = \sum_{j=1}^{2^N} p_j s^j,$$

где  $p_j$  задает вероятность того, что в системе окажется  $j$  частиц.

При исследовании такой модели амплификации было показано, что точность определения начальной концентрации  $A_0$  по кинетической кривой ПЦР с неизбежностью ограничена случайной природой реакции, при этом уровень неопределенности такого анализа зависит от вероятности удвоения и исходного количества фрагмента [1].

Представляется, что авторы этой модели учли не все факторы, определяющие ход реакции. Мы расширили стохастическую модель ПЦР, внося в нее несколько новых элементов, которые, по нашему мнению, могут оказывать принципиальное влияние на результаты моделирования.

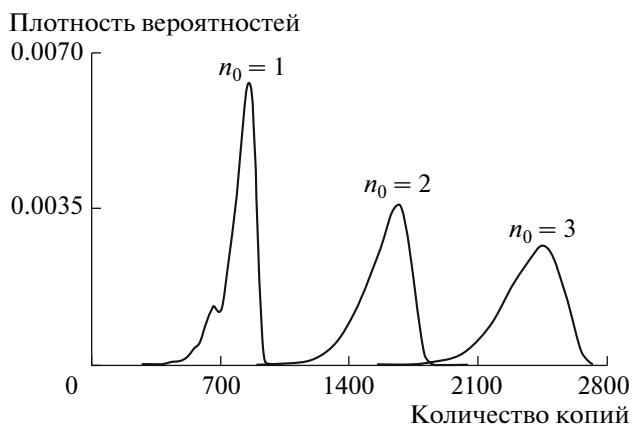
Реакция амплификации протекает на двух комплементарных цепях ДНК, для каждой из которых присутствует свой праймер. Если обозначить цепи ДНК как  $A$  и  $B$ , то на цепи  $A$  синтезируется цепь  $B$ , которая на следующем цикле пополняет пул цепей  $B$ , и аналогично для цепи  $B$ . Таким образом, на каждом цикле происходит независимое удвоение цепей двух типов, имеющих свои значения вероятности  $P_A$  и  $P_B$ . Исходно концентрации цепей  $A$  и  $B$  равны, так как в реакцию вносятся образец двухцепочной ДНК. Далее, по ходу реакции концентрации цепей будут изменяться,

причем амплификация двух пулов цепей будет взаимозависимой.

Описанная схема реакции с двумя типами цепей была исследована при различных параметрах. Расчет распределения вероятностей исходов для процесса амплификации был проведен с помощью двух моделей: аналитической и имитационной.

Для нахождения аналитического решения мы воспользовались матричным представлением производящей функции [3]. Каждый элемент такой матрицы  $p_{ij}$  определяет вероятность того, что на  $N$ -м цикле в системе имеется  $i$  цепей  $A$  и  $j$  цепей  $B$ . Соответственно на следующем шаге размер матрицы удваивается, а ее компоненты определяются суммой вкладов каждого из компонентов матрицы предыдущего цикла. Однако с каждым последующим циклом амплификации объем вычислений экспоненциально увеличивается. Так, уже при  $N > 10$  расчет распределения вероятностей исходов с помощью этой модели требует слишком много машинного времени.

Для сокращения времени расчета была предложена имитационная модель процесса амплификации, позволяющая достаточно быстро и точно оценить распределение вероятностей практически для любого количества циклов амплификации. В основу модели был положен следующий алгоритм. Перед каждым циклом амплификации имеется исходное количество цепей  $A$  и  $B$ . По очереди перебираются все цепи и с помощью датчика случайных чисел и задаваемого значения вероятности успеха ( $P_A$  или  $P_B$  соответственно) определяется, произошел ли синтез комплементарной цепи на данной копии. По итогам цикла определяли результирующее количество цепей  $A$  и  $B$  суммированием исходных и синтезированных. Путем многократного повторения процесса на-



**Рис. 3.** Распределения вероятностей количества фрагментов после 10 циклов амплификации для аналитической модели с двумя типами цепей при различном исходном количестве копий  $n_0$ . Параметры  $P_A = P_B = 0.95$ .

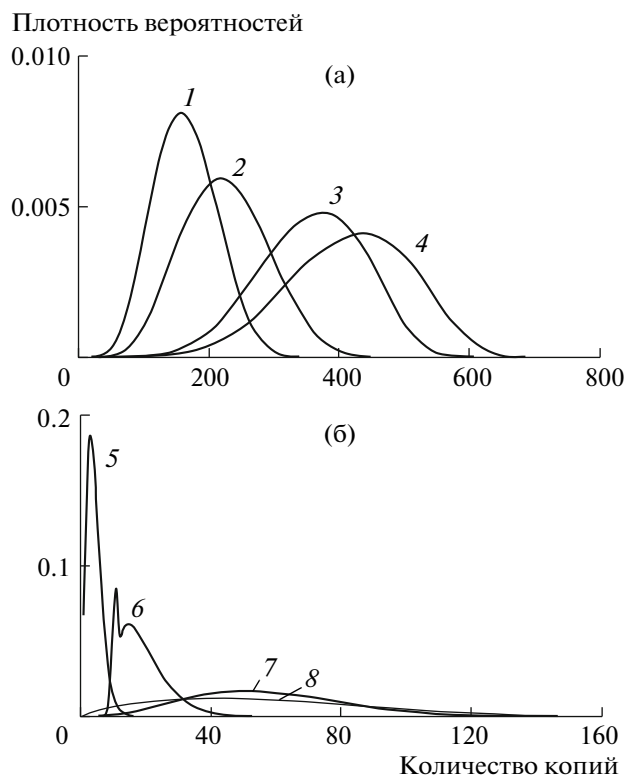
биралась статистика и оценивалось распределение вероятностей для заданного числа циклов.

Распределение вероятностей количества фрагментов после 10 циклов амплификации, полученное с использованием аналитической и имитационной моделей, приведено на рис. 1. Совпадение распределений показывает, что модели построенны корректно.

Для сравнения схемы реакции с цепями одного типа, исследованной ранее [1], и схемы с двумя типами цепей, предложенной в нашей работе, были взяты идентичные параметры моделей (вероятности удвоения для всех цепей 0.95) и для них рассчитаны распределения вероятностей после 10 циклов амплификации (рис. 2). Математические ожидания полученных распределений практически совпадают и составляют 783 фрагмента. Однако в модели с двумя типами цепей распределение вероятностей более компактно (распределение 2), что демонстрирует меньшие погрешности ПЦР по сравнению с ожидаемыми в одноцепочечной модели (распределение 1).

Распределения вероятностей для разного исходного количества частиц при одинаковых параметрах модели приведены на рис. 3. Полученные распределения практически не перекрываются для одной–трех исходных копий. Этот результат важен с точки зрения количественного ПЦР-анализа малых количеств ДНК, поскольку показывает, что стохастическая природа амплификации не препятствует точному определению исходного количества единичных копий ДНК. Таким образом, основным ограничением при таком анализе является погрешность, вносимая прибором для проведения амплификации.

В практическом применении ПЦР обычно сложно добиться одинаковой эффективности



**Рис. 4.** Распределения вероятностей количества фрагментов после 10 циклов амплификации для аналитической модели с двумя типами цепей при различных значениях вероятностей  $P_A$  и  $P_B$ . 1 и 2 – для цепей A и B соответственно при  $P_A = 0.5$  и  $P_B = 0.95$ ; 3 и 4 – для цепей A и B соответственно при  $P_A = 0.7$  и  $P_B = 0.95$ ; 5 и 6 – для цепей A и B соответственно при  $P_A = 0.05$  и  $P_B = 0.95$ ; 7 и 8 – для цепей в моделях с двумя типами и одним типом частиц соответственно для случая  $P_A = P_B = P = 0.5$ .

Исходное количество копий – 1.

синтеза двух комплементарных цепей. Это связано, в частности, с различиями в температурной зависимости гибридизации праймеров, нуклеотидном составе и вторичной структуре амплифицируемых фрагментов. Предложенная схема реакции с двумя типами цепей позволяет в явном виде учесть этот фактор. На рис. 4 приведены примеры распределений для различных вероятностей  $P_A$  и  $P_B$ . Необходимость в использовании такой схемы особенно явно проявляется при значительном различии в вероятностях  $P_A$  и  $P_B$ , например, при рассмотрении аллель-специфичной ПЦР, когда эффективность реакции для одного праймера должна быть максимальной, а для второго – по возможности близкой к нулю (рис. 4б).

Рассмотренная в данной работе модель ПЦР может быть развита в нескольких аспектах. В ней может быть учтен случай, когда эффективности амплификации исходной ДНК и синтезированных в ходе реакции фрагментов различаются.

Также возможно ввести в модель переменные эффективности амплификации цепей, что позволило бы проанализировать статистику для кинетических кривых ПЦР. Кроме того, представляет большой интерес возможность экспериментальной проверки предсказаний модели при анализе единичных копий ДНК.

## СПИСОК ЛИТЕРАТУРЫ

1. *Pescoud J., Jacob C.* // *Biophys. J.* 1996. V. 71. № 7. P. 101–108.
2. *Lalam N., Jacob C., Jagers P.* // *Adv. Appl. Prob.* 2004. V. 36. P. 602–615.
3. *Севастьянов Б.А.* Ветвящиеся процессы. М.: Наука, 1971. 436 с.