

УДК 577.2:541.1

МАТЕМАТИЧЕСКИЙ АНАЛИЗ КИНЕТИЧЕСКИХ КРИВЫХ ПОЛИМЕРАЗНОЙ ЦЕПНОЙ РЕАКЦИИ

© 2016 г. Д. Г. Сочивко, А. А. Фёдоров, Д. А. Варламов,
В. Е. Курочкин, академик РАН Р. В. Петров

Поступило 15.09.2015 г.

В работе рассмотрены подходы к математическому анализу кинетических кривых полимеразной цепной реакции (ПЦР). Приведены общие принципы математического анализа процесса ПЦР. Описаны возможные подходы к анализу при помощи аппроксимации кривых ПЦР и функций эффективности разными зависимостями. Предложены несколько моделей ПЦР на основе уравнений химической кинетики. Приведены критерии выбора оптимальных функций для описания эффективности ПЦР.

DOI: 10.7868/S0869565216010278

Полимеразная цепная реакция (ПЦР) является сегодня основным молекулярно-генетическим методом 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 основных шагов: 1) денатурации путем нагревания реакционной смеси, при этом комплементарные цепи амплифицируемого двухцепочного фрагмента расходятся; 2) понижения температуры и элонгации, при этом короткие специфические цепи ДНК (праймеры) связываются с одиночными цепями фрагмента, и на матрице одиночной цепи с помощью фермента ДНК-полимеразы происходит синтез второй, комплементарной цепи ДНК.

Если в процессе реакции производить количественную регистрацию ее продуктов на каждом цикле (ПЦР в реальном времени, ПЦР-РВ), то мы получим кинетическую кривую ПЦР, имеющую дискретный по времени характер [1]. Для кривых ПЦР характерна S-образная форма. Количество амплифицируемых фрагментов после N циклов реакции равно $S_N = S_0 (E + 1)^N$, где S_0 – ис-

ходное количество цепей ДНК, E – эффективность реакции, соответствующая вероятности удвоения фрагмента. Если принять E , равную 1, то количественный анализ с помощью ПЦР-РВ становится тривиальной задачей. Однако рассмотрение полученных в эксперименте кривых показывает, что в большинстве случаев E не равно 1 и является переменной величиной, что выражается в замедлении роста кривой ПЦР и выходе на плато. Тогда количество фрагментов после N циклов реакции должно быть выражено как

$$S_N = S_0 \prod_{i=1}^N (1 + E_i), \quad (1)$$

где E_i – эффективность на i -м цикле. На практике тем не менее эффективность на начальных (подпороговых) циклах ПЦР принимают за постоянную величину, которую можно определить при помощи кратных разведений контрольных образцов с известной концентрацией. Ключевым слабым местом такого подхода является необходимое допущение идентичности условий протекания реакции и регистрации её продуктов в разных реакторах для разных образцов. Это допущение нарушается, когда образцы существенно варьируются по характеристикам, прежде всего, по присутствию ингибиторов ПЦР разного рода. Кроме того, условия в разных реакторах могут отличаться в связи с действием случайных факторов (погрешности дозирования реагентов, образование конденсата, пузырьки и т.п.). Результатом являются ошибки анализа двух видов: ошибки в определении порогового цикла и ошибки, связанные с различиями в эффективности реакции.

 ЗАО "Синтол", Москва

 Институт аналитического приборостроения
 Российской Академии наук, Санкт-Петербург
 E-mail: faa@bk.ru

 Всероссийский научно-исследовательский институт
 сельскохозяйственной биотехнологии, Москва

 Государственный научный центр
 "Институт иммунологии"
 ФМБА России, Москва

Чтобы избежать влияния указанных факторов на результаты исследования с помощью ПЦР, необходимо разрабатывать новые эффективные математические методы анализа данных. Работы по математическому анализу кривых ПЦР ведутся с начала 1990-х годов. По уровню анализа эти работы можно разделить на 3 группы: 1) аппроксимация кривых реакции непрерывными функциями; 2) построение функций, описывающих эффективность ПЦР; 3) построение моделей ПЦР, основанных на уравнениях химической кинетики.

Для аппроксимации кривых ПЦР было предложено использовать непрерывные функции, график которых имеет S-образную форму. В нескольких работах сопоставляли качество аппроксимации разными функциями [2–4]. По результатам сравнения среди функций с четырьмя параметрами логистическая функция обеспечила наилучшую аппроксимацию [2]. Добавление пятого параметра, определяющего степень асимметрии кривой, существенно улучшает качество аппроксимации [3, 4]. Анализ кривых с помощью таких функций носит чисто описательный характер и удобен для простой обработки результатов, поиска уровня фона, нормализации кривых по амплитуде и других подобных операций.

Качество аппроксимации кривых ПЦР непрерывными функциями не очень высоко, поскольку они описывают простые математические модели и не соотносимы с природой реакции. В качестве основы для полноценного анализа кривых ПЦР необходимо исследовать ключевую характеристику ПЦР – функцию эффективности E . Как видно из уравнения (1), определив эту функцию, мы сможем однозначно получить кривую ПЦР при любой начальной концентрации фрагментов S_0 . Для этого функция эффективности должна быть задана как $E(S)$ – зависимость вероятности удвоения фрагмента от концентрации фрагментов на начало цикла ПЦР. Тогда кривая ПЦР будет описываться рекурсивной функцией вида $S_{i+1} = S_i(1 + E(S_i))$.

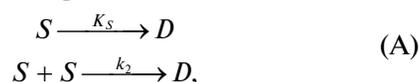
Характер функции эффективности может быть оценен по экспериментальным данным [5]. График функции эффективности представляет собой монотонно убывающую кривую, проходящую от максимального значения, обычно близкого к 1, до нуля. В ряде работ [5–10] исследовали возможность аппроксимации таких графиков различными функциями, такими как линейная, гиперболическая, логистическая, гауссова и другие. Показано, что анализ с использованием аппроксимации функции эффективности обладает существенными преимуществами перед стандартным методом пороговых циклов, а также аппроксимацией кривых ПЦР сигмоидными функциями.

Другой подход к описанию функции эффективности учитывает механизм ПЦР и основан на

анализе уравнений химической кинетики. Для этого создают модель ПЦР той или иной степени детализации, записывают для нее уравнения реакций, выводят из них систему дифференциальных уравнений и далее решают ее, обычно численными методами.

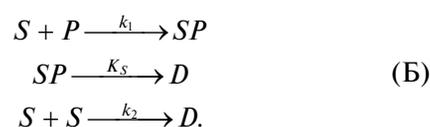
Нами было проведено исследование нескольких вариантов таких моделей, изложение результатов которого составило предмет настоящего сообщения.

Простейшей моделью такого рода можно считать модель, предложенную в работе [6], в которой имеется один параметр и учитывается только работа фермента. Модель с двумя параметрами была предложена в работе [11]. В ней описывали процесс ПЦР как две реакции:

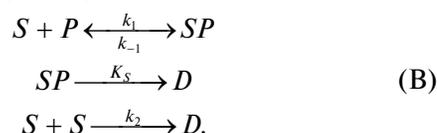


где D – двухцепочный фрагмент, образующийся либо в результате синтеза второй цепи (K_s), либо путем гибридизации двух комплементарных цепей, синтезированных на предыдущих циклах реакции (k_2). Эта модель не рассматривает участие праймеров в реакции, поскольку на начальной фазе ПЦР праймеры находятся в избытке, и их концентрацию можно считать постоянной. Такая модель адекватно описывает участок экспоненциального роста и начало торможения реакции, однако не может описать кривую ПЦР в целом, так как в ней не заложен механизм выхода реакции на плато.

Можно расширить данную модель путем введения в реакцию праймеров P , образующих с цепями комплекс SP :



При этом реакция ограничена исчерпанием праймеров, а также появляется третий параметр модели, константа гибридизации праймеров k_1 . Далее можно учесть в этой модели обратимость гибридизации праймеров:



В таком виде модель содержит четыре параметра. Сравнение трёх вариантов модифицированной простейшей модели (A)–(B) проводили путем численного решения соответствующей системы дифференциальных уравнений. Кривые ПЦР приведены на рис. 1, значения параметров указаны в подписи к рисунку. Можно видеть, что протекание ПЦР на начальных циклах практически идентично для всех трех моделей. Введение обратимости гибридизации праймеров несколько

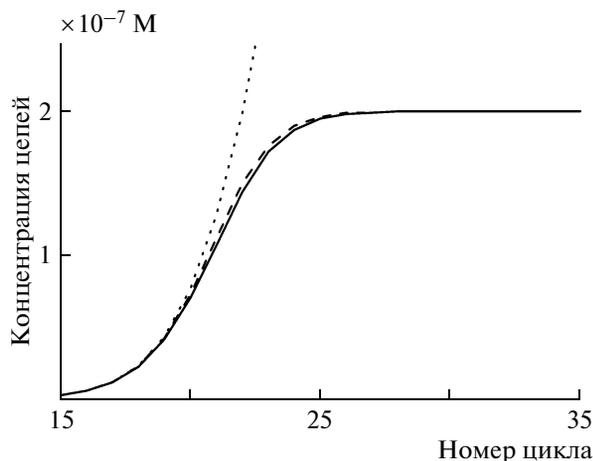


Рис. 1. Кинетические кривые ПЦР, соответствующие моделям (А) – пунктирная линия, (Б) – штриховая линия, (В) – сплошная линия. Параметры моделей: $k_1 - 1 \cdot 10^6 \text{ M}^{-1} \cdot \text{c}^{-1}$, $k_{-1} - 1 \cdot 10^{-4} \text{ c}^{-1}$, $k_2 - 3 \cdot 10^6 \text{ M}^{-1} \cdot \text{c}^{-1}$, $K_s - 2.5 \text{ c}^{-1}$, $S_0 - 1 \cdot 10^{-13} \text{ M}$, $P_0 - 1 \cdot 10^{-7} \text{ M}$.

снижает эффективность реакции на фазе замедления.

Похожая простая модель предложена в работе [12], в ней рассмотрены обратимые гибридизации одиночных цепей с праймерами и зондом, а также необратимая гибридизация с комплементарными цепями. При этом синтез второй цепи на комплексе SP считается мгновенным и полным, и общий процесс последовательной гибридизации и синтеза можно считать необратимым. В результате полученная модель имеет только два кинетических параметра.

Указанные модели исходят из того, что константа реакции синтеза второй цепи является постоянной и не зависит от концентрации реагентов. Это является серьезным упрощением, поскольку данная реакция катализируется ферментом ДНК-полимеразой, а значит, ее необходимо рассматривать как минимум в рамках кинетики Михаэлиса–Ментен. Это может быть сделано, если принять, что в приведенных выше моделях константа реакции $SP \rightarrow D$ описывается уравнением

$$K_S = V_{\max} / (K_m + [S]), \quad (2)$$

где V_{\max} – максимальная скорость реакции, K_m – константа Михаэлиса. Сходный вариант модели был рассмотрен в работе [13]. Можно расширить модель кинетики фермента путем учета возможности неспецифического связывания его с одиночными и двойными цепями ДНК, т.е. конкурентного ингибирования субстратом и продуктами реакции. В этом случае константа скорости реакции задается уравнением

$$K_S = V_{\max} / (K_m + [S] + j \cdot [C]), \quad (3)$$

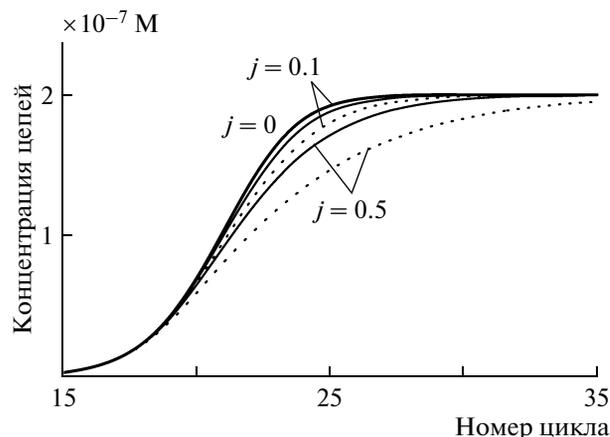


Рис. 2. Кинетические кривые ПЦР, соответствующие модели (В) с учетом ферментативного синтеза цепей, в которой K_S определяется уравнением (3) при значениях $j = 0, 0.1, 0.5$. Сплошные линии – $[C]$ равно концентрации цепей $[S]$ на начало цикла. Пунктирные линии – $[C]$ равно концентрации $[D]$ на текущий момент цикла. Параметры моделей: $k_1 - 1 \cdot 10^6 \text{ M}^{-1} \cdot \text{c}^{-1}$, $k_{-1} - 1 \cdot 10^{-4} \text{ c}^{-1}$, $k_2 - 3 \cdot 10^6 \text{ M}^{-1} \cdot \text{c}^{-1}$, $K_M - 2 \cdot 10^{-9} \text{ M}$, $V_{\max} - 1 \cdot 10^{-8} \text{ M} \cdot \text{c}^{-1}$, $S_0 - 1 \cdot 10^{-13} \text{ M}$, $P_0 - 1 \cdot 10^{-7} \text{ M}$.

где j – коэффициент, соответствующий отношению констант для неспецифического связывания фермента с цепями и для специфического связывания с комплексом SP , $[C]$ – концентрация неспецифически связывающихся цепей. Если принять условие, что фермент в некоторой степени связывается со всеми цепями, то $[C]$ соответствует концентрации одиночных цепей на начало цикла. Если же фермент имеет сродство только к двойным цепям, то $[C]$ представляет собой сумму продуктов синтеза второй цепи и гибридизации комплементарных цепей на текущий момент реакции. Оба эти варианта были нами рассмотрены, и кривые ПЦР для них приведены на рис. 2. При варьировании значений коэффициента j эффект ингибирования был весьма выражен в первом случае и замечен во втором. По данным [14] можно оценить величину j для двойных цепей как не превышающую 0.1, причем для одиночных цепей связывание еще слабее.

Интересно рассмотреть два ключевых параметра, определяющих форму кривой ПЦР в данной модели, а именно максимальную скорость реакции V_{\max} и константу скорости для гибридизации комплементарных цепей k_2 . Согласно кинетике Михаэлиса–Ментен, $V_{\max} = k_{\text{cat}} \cdot [\text{Enz}]$, где k_{cat} – число оборотов фермента при синтезе фрагментов, $[\text{Enz}]$ – полная концентрация фермента в реакции. Число оборотов может в принципе быть оценено на основании общепринятого определения единицы активности ДНК-полимеразы. Если в пробирку с реакционной смесью емкостью 25 мкл внести 1 единицу фермента, то для фраг-

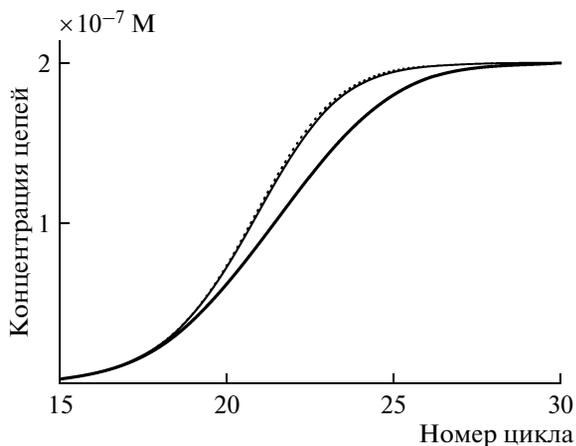


Рис. 3. Кинетические кривые ПЦР, соответствующие модели (В) с учетом ферментативного синтеза цепей, в которой K_s определяется уравнением (2). Пунктирная линия – $V_{\max} = 2 \cdot 10^{-8} \text{ М} \cdot \text{с}^{-1}$; тонкая линия – $V_{\max} = 1 \cdot 10^{-8} \text{ М} \cdot \text{с}^{-1}$; полужирная линия – $V_{\max} = 2 \cdot 10^{-9} \text{ М} \cdot \text{с}^{-1}$. Параметры модели: $k_1 - 1 \cdot 10^6 \text{ М}^{-1} \cdot \text{с}^{-1}$, $k_{-1} - 1 \cdot 10^{-4} \text{ с}^{-1}$, $k_2 - 3 \cdot 10^6 \text{ М}^{-1} \cdot \text{с}^{-1}$, $K_M - 2 \cdot 10^{-9} \text{ М}$, $S_0 - 1 \cdot 10^{-13} \text{ М}$, $P_0 - 1 \cdot 10^{-7} \text{ М}$.

мента длиной 100 п.н. значение V_{\max} составит $2 \cdot 10^{-9} \text{ М} \cdot \text{с}^{-1}$. Поскольку на практике используют от 1 до 5 единиц фермента на реакцию, мы рассчитали кривые ПЦР для соответствующего диапазона значений V_{\max} (рис. 3). Можно видеть, что увеличение концентрации фермента сверх 5 единиц на реакцию не оказывает существенного эффекта на повышение эффективности.

Значение константы гибридизации комплементарных цепей k_2 имеет смысл рассматривать в связи со значением константы образования комплекса одиночных цепей с праймерами k_1 . Оба процесса имеют общую природу, различаясь лишь длиной гибридизующихся участков – порядка сотен пар нуклеотидов для первого и 20 п.н. – для второго. Известно, что константа скорости гибридизации ДНК пропорциональна квадратному корню из длины фрагментов [15]. На рис. 4 приведены кривые ПЦР для разных значений k_2 , относящихся к k_1 как 1, 3 и 10. Увеличение длины амплифицируемого фрагмента приводит к росту значения k_2 , что оказывает существенное влияние на эффективность реакции. Если учесть, что при этом происходит пропорциональное линейное снижение числа оборотов фермента, можно понять, насколько важно при разработке систем для анализа методом ПЦР придерживаться минимально возможной длины ампликонов.

На основе результатов моделирования имеется возможность рационального выбора функции для аппроксимации зависимости $E(S)$. Такая функция должна иметь отрицательную производ-

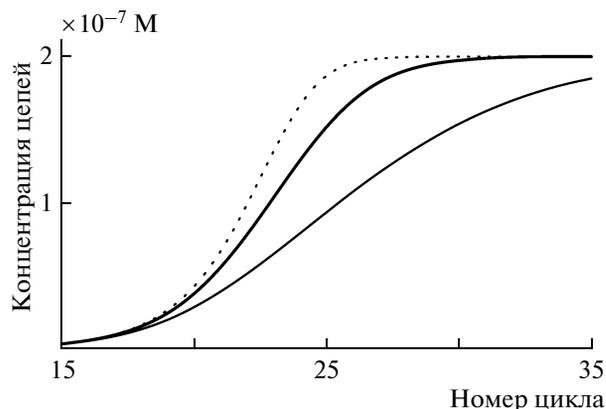


Рис. 4. Кинетические кривые ПЦР, соответствующие модели (В) с учетом ферментативного синтеза цепей, при разном значении отношения k_2/k_1 . Пунктирная линия – $k_2 = 1 \cdot 10^6 \text{ с}^{-1}$, полужирная линия – $k_2 = 3 \cdot 10^6 \text{ с}^{-1}$, тонкая линия – $k_2 = 1 \cdot 10^7 \text{ с}^{-1}$. Параметры модели: $k_1 - 1 \cdot 10^6 \text{ М}^{-1} \cdot \text{с}^{-1}$, $k_{-1} - 1 \cdot 10^{-4} \text{ с}^{-1}$, $V_{\max} - 1 \cdot 10^{-8} \text{ М} \cdot \text{с}^{-1}$, $K_M - 2 \cdot 10^{-9} \text{ М}$, $S_0 - 1 \cdot 10^{-13} \text{ М}$, $P_0 - 1 \cdot 10^{-7} \text{ М}$.

ную при $S = 0$, убывать монотонно и без перегибов, достигая нуля при некотором конечном значении S . Выбор оптимальной функции позволяет реализовать эффективные алгоритмы автоматизированного анализа результатов ПЦР-РВ, в том числе без использования контрольных образцов.

Работа выполнена при финансовой поддержке Министерства образования и науки Российской Федерации (уникальный идентификационный номер ПНИЭР RFMEFI60714X0095) и методическом содействии Центра коллективного пользования “Биотехнология” Всероссийского научно-исследовательского института сельскохозяйственной биотехнологии (проект RFMEFI62114X0003).

СПИСОК ЛИТЕРАТУРЫ

1. Higuchi R., Fockler C., Dollinger G., Watson R. // *Bio-technology*. 1993. V. 9. № 11. P. 1026–1030.
2. Zhao S., Fernald RD. // *J. Comput. Biol.* 2005. V. 12. № 8. P. 1047–1064.
3. Guescini M., Sisti D., Rocchi M.B., Stocchi L., Stocchi V. // *BMC Bioinformatics*. 2008. V. 9. P. 326.
4. Spiess A.N., Feig C., Ritz C. // *BMC Bioinformatics*. 2008. V. 9. P. 221.
5. Сочивко Д.Г., Федоров А.А., Лавров В.В., Курочкин В.Е., Петров Р.В. // *ДАН*. 2011. Т. 439. № 5. С. 696–699.
6. Schnell S., Mendoza C. // *J. Theor. Biol.* 1997. V. 184. № 4. P. 433–440.
7. Lalam N., Jacob C., Jagers P. // *Adv. Appl. Probability*. 2004. V. 36. № 2. P. 602–615.

8. *Alvarez M.J., Vila-Ortiz G.J., Salibe M.C., Podhajcer O.L., Pitossi F.J.* // BMC Bioinformatics. 2007. V. 8. № 1. P. 85.
9. *Batsch A., Noetel A., Fork C., Urban A., Lazic D., Lucas T., Pietsch J., Lazar A., Schömig E., Gründemann D.* // BMC Bioinformatics. 2008. V. 9. № 1. P. 95.
10. *Lievens A., Van Aelst S., Van den Bulcke M., Goetghebeur E.* // Nucl. Acids Res. 2012. V. 40. № 2. P. e10–e10.
11. *Boggy G.J., Woolf P.J.* // PloS One. 2010. V. 5. № 8. P. e12355.
12. *Smith M.V., Miller C.R., Kohn M., Walker N.J., Portier C.J.* // BMC Bioinformatics. 2007. V. 8. № 1. P. 409.
13. *Gevertz J.L., Dunn S.M., Roth C.M.* // Biotechnol. and Bioeng. 2005. V. 92. № 3. P. 346–355.
14. *Langer A., Schräml M., Strasser R., Daub H., Myers T., Heindl D., Rant U.* // Sci. Repts. 2015. V. 5. 12066.
15. *Wetmur J.G., Davidson N.* // J. Mol. Biol. 1968. V. 31. № 3. P. 349–370.