

**А. Л. Буляница, А. Н. Тупик, Г. Е. Рудницкая, А. А. Евстапов** (Санкт-Петербург, ИАП РАН). **Применение классических комбинаторных схем при количественном генетическом анализе методом молекулярных колоний.**

Метод молекулярных колоний является одной из реализаций цифровой полимеразной цепной реакции (цПЦР) [1]. Основное его отличие от классической цПЦР заключается в том, что реакция проводится не в жидкой, а в иммобилизованной гелеобразной среде. Пористая структура геля затрудняет диффузию макромолекул, но не влияет на диффузию компонентов реакции, позволяя осуществить ПЦР. В процессе амплификации копии ДНК-мишени скапливаются вокруг исходной молекулы ДНК, образуя так называемые молекулярные колонии. Так как практически каждая колония происходит из одной молекулы ДНК-мишени, то регистрация молекулярных колоний позволяет обнаруживать, подсчитывать и идентифицировать одиночные молекулы ДНК после амплификации в анализируемой пробе [2]. Тем самым, задача исследователя состоит в корректном подсчете числа колоний в реакционной камере.

Однако, если исходные макромолекулы расположены близко друг к другу, то объединенное наложенное изображение пары (тройки и т. д.) колоний будет ошибочно рассматриваться, как один объект, и количественный анализ (счет колоний) произойдет с погрешностью.

Цель данной работы — обоснование применимости классической комбинаторной схемы для моделирования результатов анализа, а также использование соответствующей модели для выбора условий реализации метода молекулярных колоний при обнаружении ДНК-мишени (например, число циклов амплификации, коэффициента диффузии реагентов и т. п.).

*Исходные положения* для математической модели:

**А.** Равномерность распределения исходных молекул по реакционной камере (круг радиусом  $R = 6-7$  мм), а также отсутствие влияния эффектов, связанных с развитием колонии вблизи ее границ.

**Б.** Формируемая согласно схеме «реакция-диффузия» колония представляет собой круг радиуса  $r_0$  от десятков до сотен мкм.

**В.** В случае наложения друг на друга двух соседних колоний, расстояние между центрами которых менее  $kr_0$ , где  $k = 0,4-0,6$ , они интерпретируются, как один объект. Следовательно, круг соответствующего радиуса можно трактовать, как своеобразный «элемент» или «разряд». Отношение площади реакционной камеры к площади «разряда»  $(R_0/kr_0)^2 = N$  дает общее число «разрядов».

В результате вероятностная модель счета молекулярных колоний аналогична классической комбинаторной задаче последовательного выбора без возвращением из группы  $N$  элементов с числом повторений  $m$  равным числу колоний. Эта задача подробно обсуждена, например, в примерах 14.125-14.127 [3].

*Базовые формулы:*

$P[0] = (N-1)(N-2) \cdots (N-m+1)/N^{m-1}$  — вероятность того, что ни одна из колоний не пересекается с соседними, и тогда погрешность количественного анализа отсутствует;

$P[1 : 2] = C_m^2(N-1) \cdots (N-m+2)/N^{m-1}$  — вероятность наличия одной пары пересекающихся колоний, что приводит к занижению числа колоний на 1;

$P[2 : 2] = \frac{1}{2!} C_m^2 C_{m-2}^2 (N-1) \cdots (N-m+3)/N^{m-1}$  — вероятность наличия двух пар,  $P[1 : 3] = C_m^3 (N-1) \cdots (N-m+3)/N^{m-1}$  — вероятность одной тройки пересекающихся колоний, что приводит к ошибке счета на 2 и т. д. Здесь  $C_m^k$  — число сочетаний из  $m$  по  $k$ .

Аналогично рассчитываются вероятности более сложных схем пересечения колоний и, как следствие, вероятности соответствующей ошибки количественного анализа. Задавая  $m$  и  $N$ , можно найти также и математическое ожидание соответствующей ошибки оценивания.

**Пример 1.** В реакционной камере 14 мм выявлены 15 колоний, средний радиус которых 0,72 мм. Задав  $k = 0, 6$ , получим  $N = 270$ . Тогда вероятность безошибочного счета  $P[0]$  составит 67,3%, вероятность 1 пары пересекающихся колоний — 27,6%. Тем самым, примерно в 95% случаев относительная погрешность количественного анализа не превысит — 1/15 или 6,7%. Математическое ожидание числа «недосчитанных» молекулярных колоний примерно  $(-0,38)$ . Для сравнения, при  $m = 9$  колоний, вероятность безошибочного счета 87,4%, вероятность ошибки в 1 колонию 12,0%. Тем самым, лишь в 0,6% случаев ошибка счета может оказаться большей. Математическое ожидание «недосчета» около  $(-0,13)$ .

**Пример 2.** Известно, что анализируемое число молекулярных колоний в эксперименте не превосходит 100. Требуется оценить, каковы должны быть средние размеры колонии, чтобы математическое ожидание числа «недосчитанных» колоний не превосходило 2, а вероятность того, что число недосчитанных колоний будет 4 и более не превышала бы 10%. Методом перебора получаем, что  $N$  около 2600. Тогда вероятность безошибочного счета составит 14,5%, вероятности ошибок счета 1, 2 и 3 колонии равны 28,8, 27,7 и 17,3%. Тем самым, погрешность 4 и более колоний реализуется с вероятностью около 11,7%, что примерно соответствует требуемым вероятностным характеристикам. В свою очередь,  $N = 2600$  означает  $kr_0 = 134,5$  мкм или  $r_0 = 224,5$  мкм. Тем самым, возможно подобрать условия, в первую очередь, вязкость и пористость гелевой среды при известных длинах фрагментов ДНК-мишени и обеспечить ограничение расширения границ молекулярной колонии указанными пределами.

Используемая классическая комбинаторная схема широко известна, но на ее основе можно решить две важных прикладных задачи количественного генетического анализа: i) моделирование закона распределения и оценивание погрешности счета числа молекулярных колоний и ii) выбор условий анализа, позволяющих формировать молекулярные колонии с требуемыми геометрическими характеристиками.

Работа выполнена в рамках Федеральной целевой программы «Национальная система химической и биологической безопасности Российской Федерации» (2009–2014 годы), НИР «Анализ-13» по теме «Разработка методик регистрации результатов цифровой полимеразной цепной реакции и технологий создания микрочиповых устройств для ее постановки».

#### СПИСОК ЛИТЕРАТУРЫ

1. Vogelstein B., Kinzler K. W. Digital PCR. — Proc. Nat. Acad. Sci. USA, 1999, v. 96, № 16, p. 9236–9241.
2. Четверина Е. В., Четверин А. Б. Нанокolonии: обнаружение, клонирование и анализ индивидуальных молекул. — Успехи биол. химии, 2008, т. 48, с. 3–64.
3. Свешников А. А. Сборник задач по теории вероятностей. М: Наука, 1965.

