

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РОССИЙСКОЙ ФЕДЕРАЦИИ

Национальный исследовательский Нижегородский
государственный университет им. Н.И. Лобачевского

Т.В. Лаптева
А.А. Заикин
М.В. Иванченко

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ И ЧИСЛЕННЫЕ МЕТОДЫ ИССЛЕДОВАНИЯ СИСТЕМ ГЕННОЙ РЕГУЛЯЦИИ

Учебно-методическое пособие

Рекомендовано методической комиссией Института
информационных технологий, математики и механики
для студентов ННГУ, обучающихся по направлениям подготовки
01.03.02 и 01.04.02 “Прикладная математика и информатика”,
09.06.01 “Информатика и вычислительная техника”

Нижний Новгород
2019

УДК 519.6 530.182
ББК В 22.19 312.2,
Л-24

Л-24 Лаптева Т.В., Заикин А.А., Иванченко М.В. Математическое моделирование и численные методы исследования систем генной регуляции: Учебно-методическое пособие. – Нижний Новгород: Нижегородский госуниверситет, 2019. – 29 с.

Рецензент: д.ф.-м.н. **О.И. Канаков**

Процессы генной регуляции в живых клетках носят динамический характер и являются крайне сложным объектом исследований как в силу большого числа взаимно регулирующих генов, так и топологии этих взаимодействий. Для понимания этих процессов необходимо прежде всего изучить принципы динамики малых генных сетей, которые, например, могут быть экспериментально реализованы методами синтетической биологии. Цель данного пособия – дать представление о современных математических подходах и численных методах в этой области в рамках теории нелинейной динамики и случайных процессов. Пособие предназначено для студентов и аспирантов Института информационных технологий, математики и механики ННГУ, специализирующихся в области прикладной математики и информатики. Рекомендуется при изучении дисциплин “Математическое моделирование живых систем”, “Математическое моделирование”, “Современные проблемы прикладной математики и информатики”, “Вычислительные задачи прикладной математики”. Для успешного усвоения материала необходимо предварительное изучение дисциплин “Математический анализ”, “Алгебра и геометрия”, “Дифференциальные уравнения”, “Численные методы”.

Ответственный за выпуск: председатель методической комиссии
Института информационных технологий, математики и механики ННГУ
к.ф.-м.н., доцент **А.В. Грезина**

УДК 519.6 530.182
ББК В 22.19 312.2,

© Т.В. Лаптева, А.А. Заикин, М.В. Иванченко 2019

© Нижегородский государственный университет
им. Н.И. Лобачевского, 2019

Содержание

Введение	4
1. Динамические модели транскрипционной регуляции	5
2. Стохастические модели транскрипционной регуляции	9
3. Автоколебания	14
4. Бистабильный элемент	17
5. Последовательная активность генов. Репресселятор	20
6. Коллективная динамика	22
Заключение	25
Литература	26

Введение

Разработка и исследование математических моделей функционирования живых организмов является одним из самых актуальных естественнонаучных направлений. В частности, большое внимание уделяется вопросам регуляции жизнедеятельности клеток за счет динамических процессов взаимной активации и деактивации генов [1]. Такие ансамбли взаимодействующих генов называют генными сетями или цепями [2].

Помимо естественных систем, включающих тысячи генов, методы генной инженерии позволяют создавать малые искусственные генные сети, которые могут быть изучены достаточно детально, как экспериментально, так и теоретически [3]. Синтетические генные сети обычно состоят из нескольких взаимодействующих генов, которые встраиваются в живую клетку и практически не взаимодействуют с собственными генами клетки. Хотя искусственные генные сети, доступные для реализации на текущем уровне развития синтетической биологии, существенно уступают в сложности реальным генным сетям живых клеток, они служат хорошей экспериментальной основой для проверки математических моделей, методов и подходов к описанию регуляторных сетей, в том числе, естественных.

К настоящему времени нелинейная динамика является основным математическим аппаратом исследования динамики сетей генной регуляции. Ее результаты блестяще подтверждаются биологическими экспериментами. Так, на основе синтетических генных сетей реализованы бистабильные переключатели [4], автоколебательные элементы-осцилляторы [5, 6] и счетчики событий [7]. Когда количество генов и молекул в клетке небольшое, и дискретность становится принципиальной, требуется более реалистичное описание в виде уравнений химической кинетики. Соответствующие численные методы математического моделирования также разработаны [11].

Цель данного учебного пособия – дать представление об основных методах математического моделирования и численного исследования систем синтетической генной регуляции [8], привести примеры исследования базовых моделей и их динамических режимов, провести краткий обзор основных результатов и актуальных направлений исследования.

1. Динамические модели транскрипционной регуляции

В основе математических моделей динамики транскрипционной регуляции лежат кинетические уравнения биохимических процессов, которые обеспечивают производство белков в клетке. Непосредственным регулятором экспрессии гена является промоутер, область ДНК, предшествующая кодирующей ген последовательности. При конструировании синтетической геной сети в ее составе предусматриваются транскрипционные факторы – белки, которые могут связываться с используемыми промоутерами и модулировать таким образом активность гена. Различают регулируемые и постоянно активные промоутеры. Регулируемые промоутеры содержат элементы последовательности ДНК, называемые сайтами оператора, связываясь с которыми транскрипционные активаторы усиливают, а транскрипционные репрессоры ослабляют связывание РНК-полимеразы с этим участком ДНК и, как следствие, экспрессию гена в целом (рис. 1).

Рассмотрим метод построения математической модели на примере простейшего промоутера, содержащего единственный сайт оператора [8]. Многие белки, использующиеся в синтетической биологии, приобретают свойства транскрипционных регуляторов только в состоянии олигомеров (комплексов из нескольких молекул белка, обычно двух или четырех), поэтому в нашем примере используем случай димера. Предположим также, что число молекул ДНК и транскрипционного фактора в клетке постоянно. Тогда можно записать следующие химические реакции:



где P и P_2 обозначают одиночные молекулы и димеры белка, O_b и O_u – связанные с димером белка и свободные операторы соответственно, а k_d , k_{-d} , k_b , k_{-b} – кинетические коэффициенты прямых и обратных реакций. Для соответствующих концентраций в равновесном состоянии получаем

$$\begin{aligned} k_d P^2 &= k_{-d} P_2, \\ k_b O_u P_2 &= k_{-b} O_b, \\ O_u + O_b &= N = \text{const}, \end{aligned} \tag{1.2}$$

где последнее уравнение вытекает из предположения о постоянстве числа

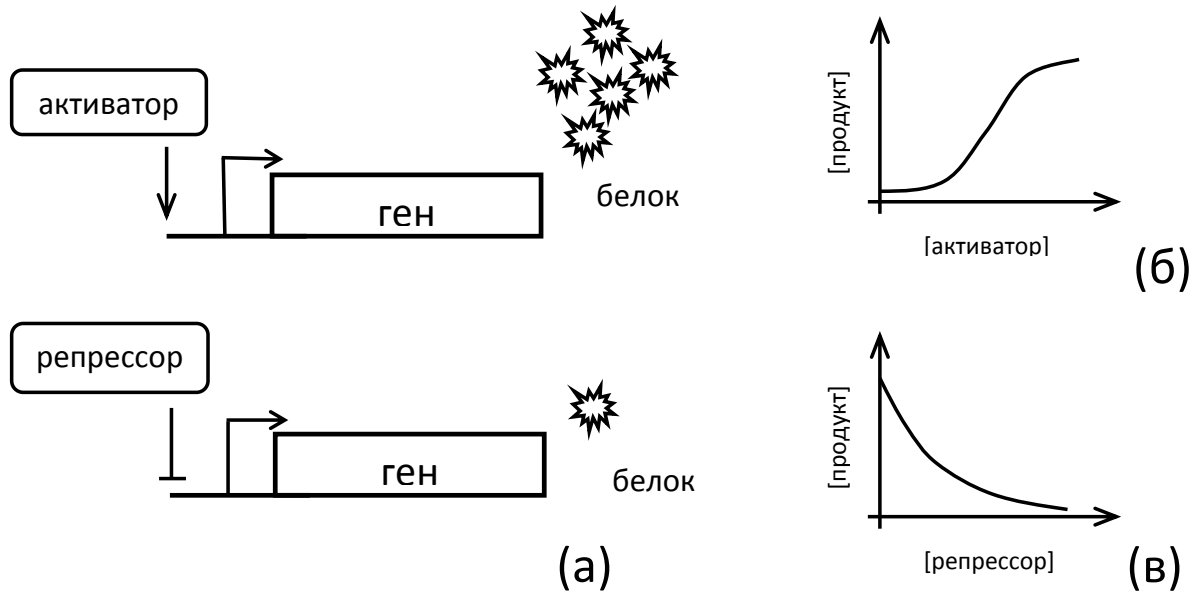


Рис. 1. Два основных типа транскрипционной регуляции – активация и репрессирование генов (а), типичный вид зависимости скорости генерации белка-продукта от концентрации молекул транскрипционного активатора (б) и репрессора (в)

молекул ДНК в клетке, и, следовательно, постоянства полного числа операторов.

Решая эти уравнения, получаем выражения для числа связанных и свободных операторов в виде так называемых функций Хилла [9]:

$$\begin{aligned}
 O_b &= N \frac{(P/K)^2}{1 + (P/K)^2}, \\
 O_u &= N \frac{1}{1 + (P/K)^2}, \\
 K &= \sqrt{\frac{k_{-b}k_{-d}}{k_b k_d}}.
 \end{aligned}
 \tag{1.3}$$

Здесь параметр K носит название константы Хилла и может быть интерпретирован как концентрация транскрипционного фактора, при которой половина операторов связана. Степень при концентрации мономера белка (в данном случае 2) называется коэффициентом Хилла или кооперативностью и соответствует степени олигомеризации транскрипционного фактора (заметим, что в случае сложного взаимодействия с промотором наилучшее совпадение с экспериментально измеренным откликом может дать дробная степень).

Итак, динамика концентрации продукта экспрессии гена, белка X , может быть записана следующим образом:

$$\dot{x} = \alpha_b N \frac{(P/K)^2}{1 + (P/K)^2} + \alpha_u N \frac{1}{1 + (P/K)^2} - r_{deg}(x), \quad (1.4)$$

где α_b, α_u – скорости производства белка геном со связанным и свободным операторами соответственно, а $r_{deg}(x)$ – функция, описывающая зависимость скорости деградации белка X от его концентрации в клетке. Для гена, регулируемого транскрипционным активатором, коэффициенты соотносятся как $\alpha_b \gg \alpha_u$, а для регулируемого транскрипционным репрессором – $\alpha_b \ll \alpha_u$.

В простейшем случае эффект деградации белка может достигаться за счет уменьшения его концентрации в процессе роста и деления клетки. Тогда имеет место так называемая экспоненциальная деградация

$$r_{deg} = \gamma x, \quad (1.5)$$

где γ – некоторый кинетический коэффициент. Во многих случаях, однако, процесс деградации желательно ускорить и (или) сделать контролируемым. Для этого ген модифицируется специальной меткой, которая делает синтетический белок-продукт видимой целью для молекулярных механизмов клеточной деградации. Скорость деградации при этом становится зависимой от числа участвующих в процессе молекул-ферментов (точнее – ограниченной этим числом), и формализм Хилла дает

$$r_{deg} = \gamma \frac{x}{K_m + x}, \quad (1.6)$$

где K_m – некоторая константа [12]. Заметим, что хотя в некотором диапазоне зависимость (1.6) аппроксимируется линейной (1.5), необходимо помнить, что нелинейное ограничение может качественно изменить динамику исследуемой системы.

Важно иметь в виду, что процесс производства белка состоит из нескольких промежуточных этапов: прикрепление РНК-полимеразы к ДНК и создание матричной РНК (транскрипция), транспорт мРНК из ядра (у эукариотов), создание рибосомой белковой аминокислотной последовательности по мРНК (трансляция), олигомеризация белка к конечной функциональной форме (обычно димер или тетрамер) [8]. Моментальное изменение скорости наработки белка при изменении концентрации транскрипционного фактора, заложенное в уравнении (1.4), является не более чем аппроксимацией. Как мы увидим в следующем разделе, реальная динамика может оказаться существенно отличной.

Существует два основных подхода к построению более реалистичных моделей. Первый состоит в более детальном описании процесса. Например, можно выделить этапы транскрипции и трансляции (считая, что результатом последней является функциональный белок). Тогда получим систему дифференциальных уравнений второго порядка:

$$\begin{aligned}\dot{m} &= \alpha_b N \frac{(P/K)^2}{1 + (P/K)^2} + \alpha_u N \frac{1}{1 + (P/K)^2} - r_m(m), \\ \dot{x} &= r_{tl}m - r_x(x),\end{aligned}\tag{1.7}$$

где m – количество молекул мРНК, $r_m(m)$, $r_x(x)$ – функции деградации мРНК и конечного белка, r_{tl} – скорость трансляции. Модель может быть детализирована и далее, например, учетом кинетики олигомеризации белка-продукта. Основная проблема этого подхода состоит в том, что количественные значения параметров промежуточных реакций в подавляющем большинстве случаев неизвестны.

Другим распространенным подходом является феноменологическое описание каскада биохимических реакций как системы с задержкой:

$$\dot{x} = \alpha_b N \frac{(P(t - \tau)/K)^2}{1 + (P(t - \tau)/K)^2} + \alpha_u N \frac{1}{1 + (P(t - \tau)/K)^2} - r_{deg}(x),\tag{1.8}$$

где τ как было показано в [13] – некоторый хорошо определенный параметр.

2. Стохастические модели транскрипционной регуляции

Биохимические процессы в клетке представляют собой набор дискретных молекулярных событий. В отдельной клетке число копий синтетического гена обычно не превосходит нескольких десятков, поэтому эффекты дискретности могут наблюдаться в эксперименте [10]. В связи с этим результаты анализа нелинейных дифференциальных уравнений требуется сопоставлять с результатами численного моделирования соответствующих марковских случайных процессов, которые, вообще говоря, могут достаточно существенно отличаться [11].

Начнем рассмотрение с простейшего процесса случайного распада вещества $S \rightarrow \emptyset$. Пусть вероятность распада одной молекулы за малый промежуток времени равна $P[(x_s(t + \Delta t) = 0 | x_s(t) = 1] = c\Delta t$, где c_0 – кинетический коэффициент реакции. Требование малости Δt вытекает из очевидного условия $c\Delta t < 1$. Тогда для некоторого количества молекул вероятность распада одной из молекул задается выражением

$$P[(x_s(t + \Delta t) = x - 1 | x_s(t) = x] = cx\Delta t. \quad (2.9)$$

При этом критерий малости интервала времени требуется усилить, чтобы вероятность двух и более реакций распада за тот же промежуток была пренебрежимо малой: $\Delta t \ll 1/cx$.

Исходя из формулы (2.9), можно предложить простейший численный метод моделирования процесса распада:

1. Вычислить текущий кинетический коэффициент реакции $a(x) = cx(t)$.
2. Сгенерировать случайное число r , равномерно распределенное на интервале $[0; 1]$.
3. Если $r < a(x)\Delta t$, считать, что на интервале $[t; t + \Delta t]$ произошла реакция распада одной из молекул, и принять $x \leftarrow x - 1$.
4. Увеличить время на величину шага, $t \leftarrow t + \Delta t$.

Легко видеть, что для математического ожидания количества вещества выполняется $\Delta \langle x \rangle = -c \langle x \rangle \Delta t$, а в пределе $\Delta t \rightarrow 0$ получается простейшее дифференциальное уравнение

$$\frac{d}{dt} \langle x \rangle = -c \langle x \rangle, \quad (2.10)$$

решением которого является

$$\langle x(t) \rangle = \langle x(0) \rangle e^{-ct}. \quad (2.11)$$

Конкретная реализация случайного процесса $x(t)$, разумеется, не совпадает с динамикой математического ожидания, однако в пределе большого количества вещества относительная дисперсия случайной величины является малой величиной, и динамическое описание эволюции математического ожидания является вполне достаточным.

Приведенный выше численный алгоритм является, безусловно, корректным, но далеко не самым пригодным для практических целей. Действительно, ввиду требования $a(x)\Delta t \ll 1$, в подавляющем большинстве испытаний неравенство, проверяемое на шаге 3, не будет выполнено. Таким образом, вычисления, производимые на каждом подобном шаге Δt , не несут полезной нагрузки. Особенно остро эта проблема встает, когда кинетический коэффициент становится большим, поскольку при этом шаг метода сильно уменьшается.

В работе [11] предложено элегантное решение этой проблемы. Вместо моделирования процесса Бернулли с близкой к нулю вероятностью события в каждом отдельном испытании, он предложил генерировать случайную последовательность моментов времени, когда события происходят. В самом деле, вероятность того, что событие произойдет через время τ , подчиняется экспоненциальному распределению:

$$P[x_s(t + \tau) = x - 1 | x_s(t') = x, t \leq t' < t + \tau] = a(x)e^{-a(x)\tau}. \quad (2.12)$$

Разрешая формулу для плотности преобразования случайной величины, легко показать, что генерация τ согласно (2.12) может быть произведена по формуле

$$\tau = \frac{1}{a(x)} \ln(1/r), \quad (2.13)$$

где r – случайная величина, равномерно распределенная на интервале $[0; 1]$.

Теперь численный алгоритм моделирования процесса распада выглядит следующим образом:

1. Вычислить текущий кинетический коэффициент реакции $a(x) = cx(t)$.
2. Сгенерировать случайное число r , равномерно распределенное на интервале $[0; 1]$, и получить случайный момент времени следующей реакции распада τ по формуле (2.13).
3. Уменьшить количество молекул $x \leftarrow x - 1$, увеличить время $t \leftarrow t + \tau$.

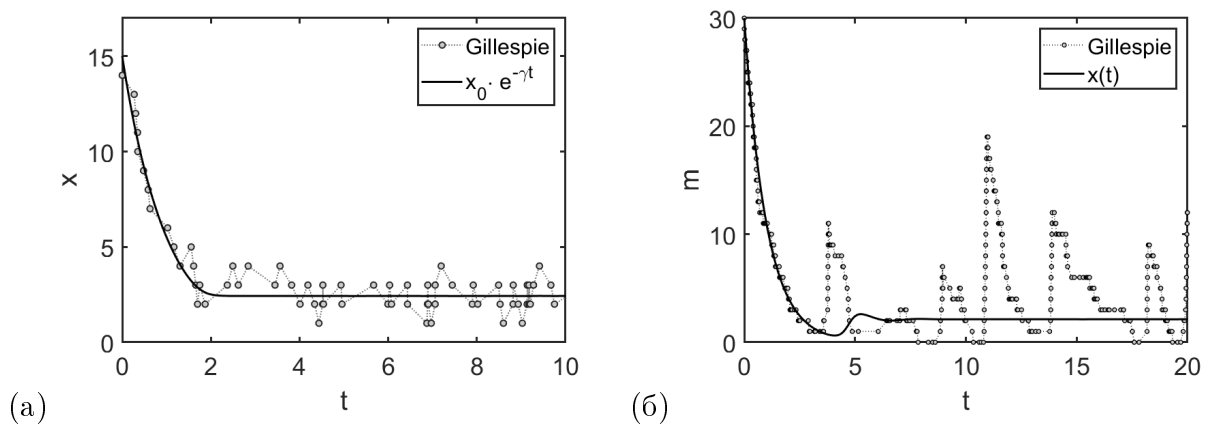


Рис. 2. Сравнение решений динамической модели (а) для процесса распада вещества и модели в виде случайного процесса: точное решение дифференциального уравнения (2.10) $x(t) = x(t)e^{-ct}$ (в общем случае динамическая модель решается численным интегрированием) и реализация стохастического метода [11]; (б) для модели авторепрессора (2.14) с параметрами $K = 200, \beta = \gamma = 1, n = 6$: численное решение дифференциального уравнения сходится к устойчивому состоянию равновесия, в то время как в модели случайного процесса (2.15) наблюдаются нерегулярные колебания.

Пример численной реализации этого алгоритма представлен на рис.2(а) в сравнении с аналитическим решением для математического ожидания этого случайного процесса (2.10) и (2.11).

Представим теперь обобщение алгоритма на случай набора веществ $\{\mathbf{x}_i\}$, реакций $\{\mathbf{R}^{(j)}\}$, каждая из которых приводит к изменению вектора количества молекул $\mathbf{x} \leftarrow \mathbf{x}_i + \nu_i^{(j)}$ с кинетическими коэффициентами $a_j(\mathbf{x})$.

Здесь ключевой идеей является пошаговое рассмотрение события как любой из реакций, а затем конкретизация состоявшейся реакции. В самом деле, имея текущий вектор состояний $\{\mathbf{x}_i\}$, мы можем определить общий кинетический коэффициент реакций, $a_0(\mathbf{x}) = \sum a_j(\mathbf{x})$. Все приведенные выше рассуждения оказываются справедливыми, разумеется, с условием малости $\Delta t \ll 1/a_0(\mathbf{x})^1$. В полной аналогии, время следующего случайного события – любой реакции, которая произошла первой, определяется согласно (2.13). Далее требуется конкретизировать событие – выбрать реакцию из допустимого набора, согласно относительной вероятности каждой: $p_j = a_j(\mathbf{x})/a_0(\mathbf{x})$; отметим, что $\sum p_j = 1$.

Таким образом, численный алгоритм Gillespie в его общей форме имеет следующий вид:

1. Вычислить текущие кинетические коэффициенты реакций $a_j(\mathbf{x})$, суммар-

¹Заметим, что в случае набора веществ и реакций алгоритм Gillespie становится незаменим, поскольку ограничения на временной шаг еще более ужесточаются.

ный коэффициент $a_0(\mathbf{x})$ и относительные вероятности $p_j = a_j(\mathbf{x})/a_0(\mathbf{x})$.

2. Сгенерировать случайное число r_1 , равномерно распределенное на интервале $[0; 1]$, и получить случайный момент времени следующей реакции распада τ по формуле $\tau = \frac{1}{a_0(x)} \ln(1/r_1)$.
3. Сгенерировать случайное число r_2 , равномерно распределенное на интервале $[0; 1]$, и определить соответствующую ему реакцию R_j , найдя минимальный индекс j , для которого выполняется $r_2 < \sum_{j'=1}^j p_{j'}$.
4. Изменить количество молекул в соответствии с выбранной реакцией $\mathbf{x} \leftarrow \mathbf{x} + \nu^{(j)}$, увеличить время $t \leftarrow t + \tau$.

Проиллюстрируем работу алгоритма на системе авторепрессора, математическая модель которого включает две переменные, m – количество молекул мРНК, x – количество молекул белка, и дифференциальные уравнения для математических ожиданий которых выглядят следующим образом (1.7):

$$\begin{aligned}\dot{m} &= \frac{K}{1 + x^n} - m, \\ \dot{x} &= \beta m - \gamma x,\end{aligned}\tag{2.14}$$

где K, β, γ – приведенные безразмерные коэффициенты. Для модели в виде случайных процессов выделяются реакции и их кинетические коэффициенты:

$$\begin{aligned}m &\rightarrow m + 1, \quad a_1 = \frac{K}{1 + x^n}, \\ m &\rightarrow m - 1, \quad a_2 = m, \\ x &\rightarrow x + 1, \quad a_3 = \beta m, \\ x &\rightarrow x - 1, \quad a_4 = \gamma x.\end{aligned}\tag{2.15}$$

Анализируя результаты численного моделирования, интересно отметить, что в динамической и стохастической моделях возможны качественно различные режимы при одних и тех же параметрах. А именно, в то время как во всем пространстве параметров единственным аттрактором динамической системы (2.14) является состояние равновесия, в реализациях случайного процесса при определенных условиях могут наблюдаться незатухающие колебания, рис.2(б). Отметим, что противоречия с требованием соответствия математического ожидания решению динамической системы не возникает: фазы колебаний в различных реализациях случайных процессов не согласованы между собой, и при усреднении получается стационарное значение,

соответствующее координатам состояния равновесия динамической системы.

3. Автоколебания

Рассмотрим, наверное, простейший пример регуляторной генной динамики – модель авторепрессора. Биологически он может быть реализован в виде гена с промотором, ингибируемым белком – продуктом экспрессии гена (рис. 3(а)). С точки зрения нелинейной динамики это система с отрицательной обратной связью, которая при определенных условиях может демонстрировать неустойчивость стационарного режима и возникновение автоколебаний через бифуркацию Андронова-Хопфа [14]. Действительно, в недавнем эксперименте подобный синтетический генный осциллятор был сконструирован, как и более сложный вариант, дополнительно содержащий транскрипционный активатор [6]. В природных генных сетях подобные принципы организации автоколебаний обнаруживаются, например в циркадных осцилляторах, управляющих суточной активностью клеток [15].

Легко видеть, что для любого значения коэффициента Хилла n одномерная модель

$$\dot{x} = \frac{\alpha}{1 + x^n} - x, \quad (3.16)$$

где часть параметров исключена за счет масштабирования переменной и времени, имеет единственное устойчивое состояние равновесия (рис. 3(б)), являющееся корнем полинома

$$x_0^{n+1} + x_0 - \alpha = 0. \quad (3.17)$$

Интересно отметить, что учет динамики мРНК наподобие (1.7) не дает автоколебаний. Состояние равновесия в двумерном фазовом пространстве – либо устойчивый узел, либо фокус. Однако добавление третьего компонента (который может быть интерпретирован как промежуточная, неактивная форма белка, возможно требующая активации неким ферментом или разворачивания) позволяет получить автоколебательный режим, например в модели осциллятора Гудвина [16].

Более биологически ясная математическая модель системы с задержкой

$$\dot{x} = \frac{\alpha}{1 + x^n(t - \tau)} - x, \quad (3.18)$$

также демонстрирует автоколебания. В самом деле, линеаризуя систему (3.18) вблизи состояния равновесия (3.17)

$$x(t) = x_0 + \xi(t), \quad |\xi(t)/x_0| \ll 1, \quad (3.19)$$

после несложных преобразований получаем

$$\dot{\xi}(t) \approx n \left(\frac{x_0}{\alpha} - 1 \right) \xi(t - \tau) - \xi(t). \quad (3.20)$$

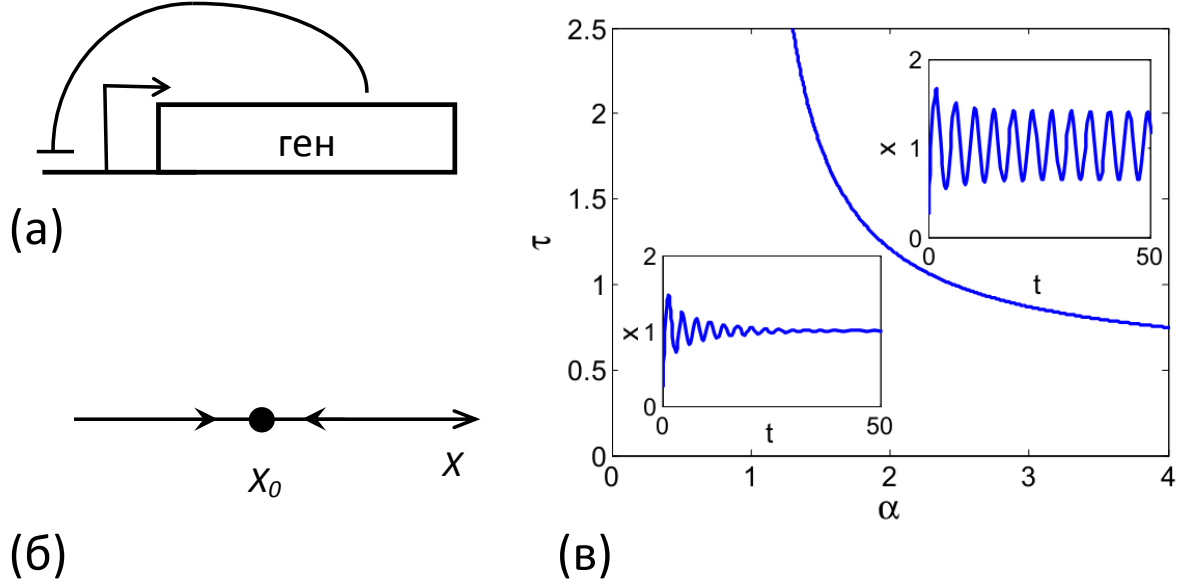


Рис. 3. Принципиальная схема авторепрессора (а), (б) глобально асимптотически устойчивое состояние равновесие в одномерной модели без задержки (3.16); (в) разбиение пространства параметров $\{\alpha, \tau\}$ модели с задержкой (3.18) на области стационарной и автоколебательной динамики и типичные реализации

Предполагая $\xi(t) \sim e^{\lambda t}$, приходим к характеристическому уравнению

$$\lambda = n \left(\frac{x_0}{\alpha} - 1 \right) e^{-\tau \lambda} - 1. \quad (3.21)$$

В точке бифуркации Андронова-Хопфа характеристические показатели являются чисто мнимыми $\lambda = \pm i\omega$ [14]. Разделяя действительную и мнимую части (3.21), получаем

$$\begin{aligned} \omega^2 &= n^2 (1 - x_0/\alpha)^2 - 1, \\ \cos \omega \tau &= -\frac{1}{n(1 - x_0/\alpha)}. \end{aligned} \quad (3.22)$$

Принимая во внимание, что $x_0 \leq \alpha$ (3.17), получаем необходимое условие для возникновения автоколебаний $n \geq 2$. Заметим, что авторепрессор *LacI*, использовавшийся в эксперименте [6], становится транскрипционным фактором в состоянии тетрамера, то есть его кооперативность равна 4. Система (3.22) определяет соотношение между мнимой частью характеристического показателя в точке бифуркации (которая может служить оценкой частоты автоколебаний недалеко от бифуркационной границы) и временем задержки:

$$\cos \omega \tau = -\frac{1}{\sqrt{1 + \omega^2}}. \quad (3.23)$$

В частности, для $\omega > 1$ получаем $\omega\tau \sim \pi/2$ и период $T = 2\pi/\omega \sim 4\tau$. Этот результат показывает, что период возникающих колебаний не обязательно близок по величине к времени задержки и может существенно превосходить его.

Бифуркационная кривая может быть получена численным решением уравнений (3.17) и (3.22). Пример такой кривой для $n = 4$ и характерные динамические режимы представлены на рис. 3(в).

4. Бистабильный элемент

Другим крайне важным типом регуляторной динамики является возможность выбора системой одного из нескольких устойчивых состояний в зависимости от начальных условий, бистабильность или мультистабильность. Такое поведение может лежать в основе изменчивости поведения клеток в зависимости от внешней среды, внешних воздействий, от предыстории. Самым известным подобным контуром среди природных является, пожалуй, генный переключатель λ -бактериофага между пассивным и активным состояниями [17].

Синтетический аналог в виде взаимно репрессирующих генов (рис. 4(a)) был впервые экспериментально реализован в двух вариантах и исследован в работе [4]. В обоих контурах одним из репрессоров был выбран *lacI*, вторым – либо температурно-чувствительный λ репрессор, либо *tetR*. Переключение между состояниями с доминирующей экспрессией одного или другого репрессора было реализовано с помощью “импульсов” концентраций химических веществ, понижающих активность избранных репрессоров, либо “импульсов” температуры.

Простейшая математическая модель динамики взаимных репрессоров имеет вид

$$\begin{aligned}\dot{x}_1 &= \frac{K_1}{1 + x_2^{n_2}} - x_1, \\ \dot{x}_2 &= \frac{K_2}{1 + x_1^{n_1}} - x_2,\end{aligned}\tag{4.24}$$

где $x_{1,2}$ – безразмерные концентрации молекул белков-репрессоров, $n_{1,2}$ – их кооперативность, $K_{1,2} > 0$ – скорости производства белков в отсутствие репрессирующего действия оппонента.

Рассмотрим частный случай $K_1 = K_2 = K$, $n_1 = n_2 = n$. В силу симметрии системы уравнений (4.24) состояния равновесия либо лежат на биссектрисе $x_1 = x_2$, либо образуют симметричные пары (x_1, x_2) , (x_2, x_1) . Решения первого типа являются корнями полинома

$$x^{n+1} + x - K = 0.\tag{4.25}$$

Согласно теореме Больцано-Коши уравнение (4.25) всегда имеет ровно один действительный положительный корень [18]. Линеаризация (4.24) в окрестности этого состояния равновесия дает характеристическое уравнение

$$\lambda^2 + 2\lambda + 1 - \frac{n^2}{K^2} x^{2(n+1)} = 0\tag{4.26}$$

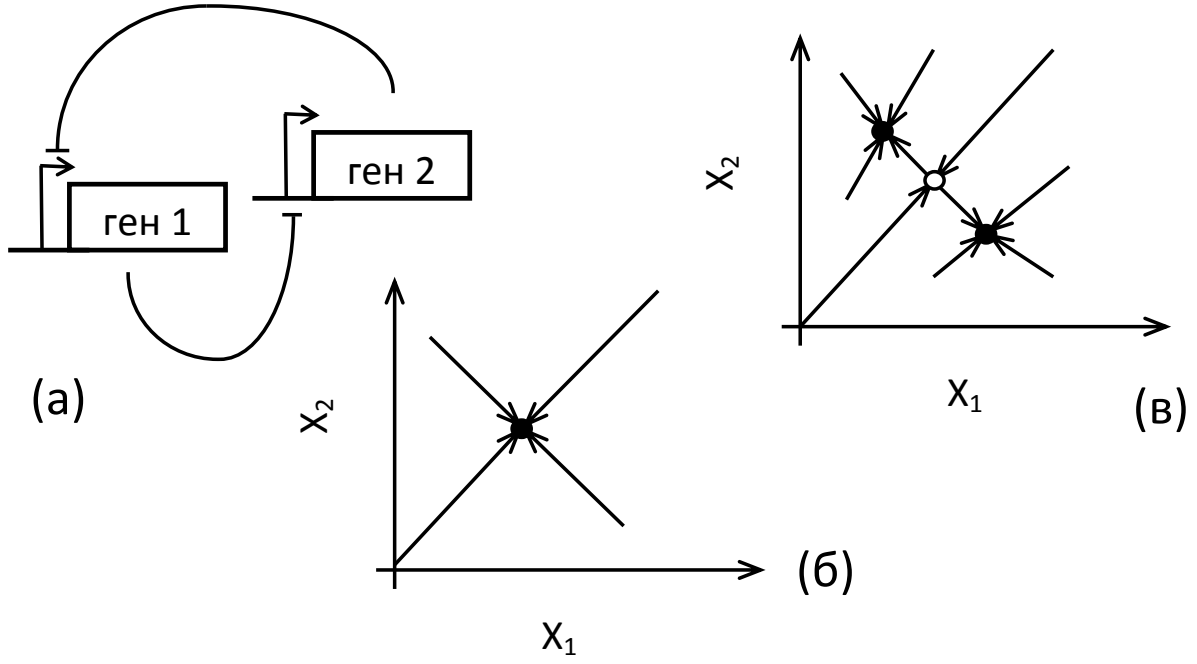


Рис. 4. Принципиальная схема бистабильного элемента, построенного на базе двух взаимных репрессоров (а); качественный вид фазового портрета системы (4.24) в режимах (б) моностабильности и (в) бистабильности

с корнями

$$\lambda_{1,2} = -1 \pm \frac{n}{K} x^{n+1}. \quad (4.27)$$

Используя (4.25) и (4.27), получаем, что симметричное состояние равновесия устойчиво при

$$K < K^* = \frac{n}{n-1} \frac{1}{\sqrt[n]{n-1}} \quad (4.28)$$

и становится седловым при $K > K^*$. Отметим, что потеря устойчивости возможна, только если кооперативность репрессоров больше 1.

Для $n = 2$ ($K^* = 2$) несложно найти все состояния равновесия системы (4.24), воспользовавшись тем, что симметричное решение удовлетворяет уравнению (4.25). После понижения степени полинома, уравнение для асимметричных состояний равновесия будет иметь вид

$$x^2 - Kx + 1 = 0, \quad (4.29)$$

откуда

$$x_{1,2} = \frac{K \pm \sqrt{K^2 - 4}}{2}, \text{ при } K \geq K^* = 2. \quad (4.30)$$

Таким образом, при $K = 2$ имеет место бифуркация трехкратного состояния равновесия, и при $K > 2$ система действительно демонстрирует бистабиль-

ность (рис. 4(б), (в)). В одном из устойчивых состояний доминирует один из репрессоров ($X_1 > X_2$), в другом – другой ($X_1 < X_2$). Выбор состояния зависит от расположения начальных условий на фазовой плоскости относительно сепаратрис седла, или от положения изображающей точки в результате внешнего воздействия на систему.

Результаты исследования более общего случая $K_1 \neq K_2$ и $n_1 \neq n_2$ можно найти в работе [4]. Также интересно, что в работе [19] авторы исследовали влияние скорости принятия клеточных решений на исход процесса. В рассмотренной модели бифуркационные параметры (аналоги $K_{1,2}$ в модели (4.24)) изменялись во времени так, что регуляторная система переходила от моностабильного состояния к бистабильному. Оказалось, что в условиях асимметрии между репрессорами прохождение бифуркационной точки с медленной скоростью приводило к режиму доминирования более сильного репрессора, в то время как быстрое прохождение обеспечивало практически равную вероятность доминирования сильного или слабого. Возможно, что именно эти эффекты играют существенную роль в процессах принятия генно-регулируемых клеточных решений в нестационарных условиях, таких как дифференциация, развитие клеток и онкогенез [20, 21, 22].

5. Последовательная активность генов. Репресселятор

Последовательная, упорядоченная активация генов является еще одним базовым типом генной регуляции. Первая экспериментальная реализация такой синтетической генной динамики – так называемый “репресселятор” – была выполнена в работе [5]. В конструкции контура авторы воплотили принцип игры “камень, ножницы, бумага”: каждый из трех репрессоров (*lacI*, *tetR* и *cI* λ -фага) ингибировал экспрессию “следующего”, как показано на рис. 5(а). В результате в надлежащих условиях наблюдалось возникновение автоколебаний с попеременной экспрессией генов.

Для теоретического анализа авторы использовали следующую математическую модель, детализирующую динамику каждой мРНК m_i и белка-продукта x_i :

$$\begin{aligned}\dot{m}_i &= \frac{\alpha}{1 + x_j^n} - m_i, \\ \dot{x} &= -\beta(x_i - m_i), \\ i &= \{lacI, tetR, cI\}, \\ j &= \{cI, lacI, tetR\}.\end{aligned}\tag{5.31}$$

Координаты состояний равновесия данной системы уравнений определяются из

$$\begin{aligned}m_i &= x_i = x, \\ x^{n+1} + x - \alpha &= 0.\end{aligned}\tag{5.32}$$

Несложно показать, что положительное решение (5.32) всегда существует и единственно. Линеаризуя систему (5.31) вблизи состояния равновесия, получаем характеристическое уравнение

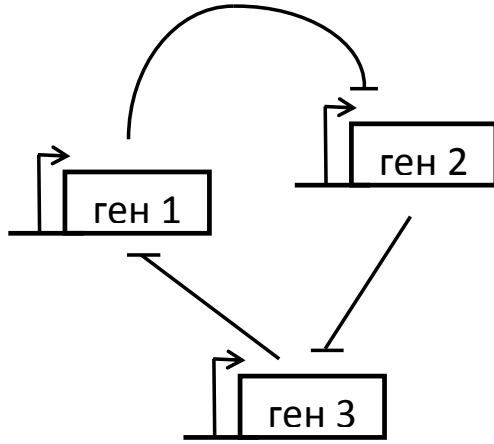
$$\begin{aligned}(1 + \lambda)^3(\beta + \lambda)^3 + \beta^3 d^3 &= 0, \\ d &= \frac{\alpha n x^{n-1}}{(1 + x^n)^2},\end{aligned}\tag{5.33}$$

корни которого удовлетворяют

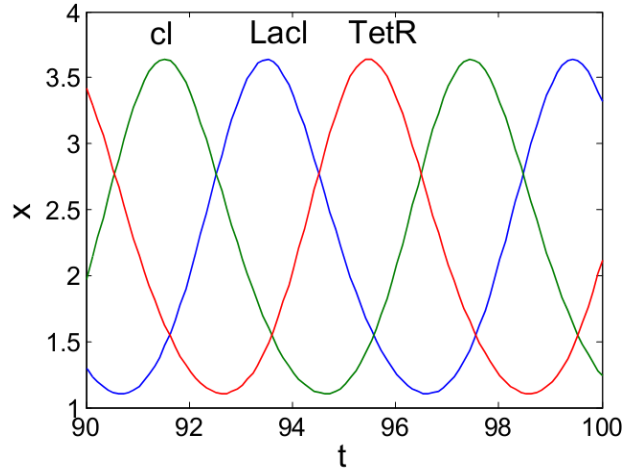
$$\begin{aligned}(1 + \lambda)(\beta + \lambda) + \beta d &= 0, \\ (1 + \lambda)(\beta + \lambda) + \frac{-1 \pm i\sqrt{3}}{2}\beta d &= 0.\end{aligned}\tag{5.34}$$

Корни первого из уравнений всегда отрицательны. Корни второго могут удовлетворять критерию бифуркации Андронова-Хопфа $\Re \lambda = 0, \Im \lambda \neq 0$. Область автоколебаний в пространстве параметров задается неравенством

$$\frac{(\beta + 1)^2}{\beta} < \frac{3d^2}{4 - 2d}.\tag{5.35}$$



(a)



(б)

Рис. 5. Принципиальная схема контура на базе трех репрессоров (а); (б) пример последовательной автоколебательной динамики: показана динамика концентраций белков в клетке при $n = 4$, $\alpha = 10$, $\beta = 5$, над максимумами указаны названия конкретных белков-репрессоров, использовавшихся в работе [5]

В частности для репрессоров-димеров $n = 2$ при $\alpha = 10$ имеем координаты состояния равновесия $x_i = m_i = 2$, параметр $d = 1.6$ и автоколебания при $\beta < \frac{19+4\sqrt{21}}{5}$. Пример последовательной автоколебательной генерации белков-репрессоров для $\beta = 5$ представлен на рис. 5(б). Впоследствии, для систем дифференциальных уравнений с n репрессорами было показано существование не только периодических решений, но и мультистабильности состояний равновесия, а также гетероклинических циклов [23, 24].

6. Коллективная динамика

До сих пор мы рассматривали способы организации динамики генной регуляции внутри одной клетки. Естественным дальнейшим шагом является инженерия межклеточного взаимодействия. Наиболее употребительный способ основан на природном механизме “кворум-сенсинга” (quorum sensing), который используется некоторыми бактериями и микроорганизмами [25]. Механизм основан на синтезе небольших молекул, которые могут легко диффундировать через мембрану клетки и служить лигандами, активирующими транскрипционные факторы (обычно применяются молекулы семейства N-ацил-гомосеринлактонов, англоязычная аббревиатура *AHL*). С помощью межклеточной коммуникации реализованы такие динамические эффекты, как синхронизация колебаний [26], взаимная синхронизация в больших ансамблях [27], синхронизация внешним периодическим сигналом [28], взаимная синхронизация клеточных колоний [29], режим взаимодействия “хищник-жертва” [30], формирование пространственных структур с помощью градиента *AHL* [31], выделение контуров для обработки изображений [32], а также продемонстрирована возможность выполнения логических операций с помощью клеток [33, 34]. Здесь мы рассмотрим одну из задач этого класса.

В благоприятных условиях клетки в колонии делятся, и их плотность возрастает, достигая со временем некоторого предельного значения. Ограничение роста может быть обусловлено рядом факторов, среди которых – ограниченное количество питательных веществ, физического пространства, скопление токсичных продуктов метаболизма. В этих условиях могут активироваться сигнальные регуляторные пути, тормозящие деление клеток. В поле синтетической биологии возникает задача об управлении численностью клеток в популяции с помощью искусственных регуляторных сетей. Помимо фундаментальной стороны вопроса присутствует и практическая. Например, может быть желательным не допустить “перенаселенности” колонии, когда полезная функциональность клеток может оказаться подавленной или ограничить рост одного из штаммов клеток в пользу другого.

Одна из возможных схем была реализована в работе [35]. Синтетический контур в каждой клетке содержит фермент *LuxI*, производящий молекулы *AHL*, которые в комплексе с белком *LuxR* являются транскрипционным активатором для другого фермента, активирующего процесс апоптоза – гибели клетки (рис. 6(а)). Концентрация молекул *AHL* в клетках и межклеточном пространстве пропорциональна плотности клеток, а весь контур работает как система отрицательной обратной связи [14].

В наиболее простом виде динамику генной регуляции можно описать

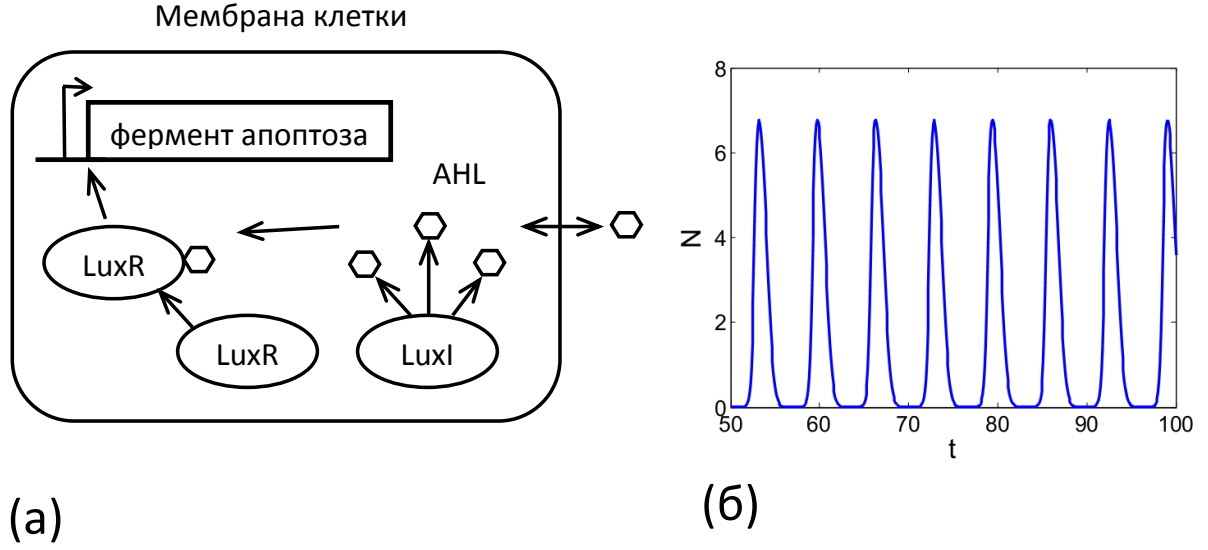


Рис. 6. Принципиальная схема синтетического контура, способного регулировать плотность клеток в популяции (а); (б) пример динамики плотности клеток в автоколебательном режиме $k = 10$, $d = 5$ (6.36)

уравнениями

$$\begin{aligned}
 \dot{N} &= N(k - N) - dEN, \\
 \dot{A} &= N - A, \\
 \dot{R} &= A - R, \\
 \dot{E} &= R - E,
 \end{aligned} \tag{6.36}$$

где N , A , R , E – безразмерные концентрации клеток, AHL , комплекса $AHL-LuxR$ и фермента апоптоза, соответственно, k – максимальная безразмерная концентрация клеток в отсутствие синтетической регуляции, d – коэффициент активации апоптоза ферментом.

Система (6.36) всегда имеет единственное нетривиальное состояние равновесия

$$N^* = A^* = R^* = E^* = \frac{k}{1 + d}. \tag{6.37}$$

Его устойчивость определяется корнями λ характеристического уравнения

$$(N^* + \lambda)(1 + \lambda)^3 + dN^* = 0. \tag{6.38}$$

Несложно показать, что состояние равновесия может терять устойчивость в результате бифуркации Андронова-Хопфа, которой на плоскости парамет-

ров (k, d) отвечает линия

$$\begin{aligned} k^* &= \frac{(3N + 1)(3N^2 + 9N + 8)}{(N + 3)^3}, \\ N^* &= \frac{k^*}{1 + d^*}. \end{aligned} \tag{6.39}$$

При $d > d^*(k)$ система демонстрирует периодические колебания плотности клеток (рис. 6(б)). Колебания численности клеток также удалось получить в эксперименте [35].

Заключение

Исследование механизмов генной регуляции является сегодня одной из наиболее актуальных задач нелинейной динамики. Особую привлекательность этой проблематике придают постоянно возрастающие возможности синтетической биологии, которые позволяют конструировать в хорошей степени изолированные и небольшие регуляторные системы, допускающие контролируемые эксперименты. В данном пособии мы постарались показать на простейших примерах, как методы нелинейной динамики могут быть использованы для анализа и предсказательной инженерии синтетических регуляторных генных сетей. Естественно, что в работах высокого уровня используются не только простейшие математические модели – низкоразмерные системы дифференциальных уравнений, но и многомерные системы, и стохастические модели, весьма детально описывающие молекулярные процессы в клетке и позволяющие достичь не только качественного, но и количественного согласования с экспериментом. Однако исследование таких моделей, как правило, сугубо численное и следует за качественным анализом низкоразмерных систем.

Литература

- [1] Jacob, F. Genetic regulatory mechanisms in synthesis of proteins / F. Jacob, J. Monod // J. Mol. Biol. — 1961. — Vol. 3. — P. 318–356.
- [2] Hasty, J. Engineered gene circuits / J. Hasty, D. McMillen, J.J. Collins // Nature. — 2002. — Vol. 420. — P. 224–230.
- [3] Nandagopal, N. Synthetic biology: Integrated gene circuits / N. Nandagopal, M.B. Elowitz // Science. — 2011. — Vol. 333. — P. 1244–1248.
- [4] Gardner, T.S. Construction of a genetic toggle switch in escherichia coli / T.S. Gardner, C.R. Cantor, J.J. Collins // Nature. — 2000. — Vol. 403. — P. 339–342.
- [5] Elowitz, M.B. A synthetic oscillatory network of transcriptional regulators / M.B. Elowitz, S. Leibler // Nature. — 2000. — Vol. 403. — P. 335–338.
- [6] A fast, robust and tunable synthetic gene oscillator / J. Stricker, S. Cookson, M.R. Bennett, et al. // Nature. — 2008. — Vol. 456. — P. 516–519.
- [7] Synthetic gene networks that count / A.E. Friedland, T.K. Lu, X. Wang, et al. // Science. — 2009. — Vol. 324. — P. 1199–1202.
- [8] O'Brien, E. Modeling synthetic gene oscillators / E. O'Brien, E. van Itallie, M.R. Bennett // Math. Biosciences. — 2012. — Vol. 236. — P. 1–15.
- [9] Hill, A.V. The possible effects of the aggregation of the molecules of haemoglobin on its dissociation curves / A.V. Hill // J. Physiol. — 1910. — Vol. 40. — P. i–vii.
- [10] Stochastic gene expression in a single cell / M.B. Elowitz, A.J. Levine, E.D. Siggia, et al. // Science. — 2002. — Vol. 297. — P. 1183–1186.
- [11] Gillespie, D.T. Stochastic smulation of chemical kinetics / D.T. Gillespie // Annual Review of Physical Chemistry. — 2007. — Vol. 58. — P. 35–55.
- [12] Johnson, K.A. The original Michaelis constant: Translation of the 1913 Michaelis–Menten paper / K.A. Johnson, R.S. Goody // Biochemistry. — 2011. — Vol. 50 (39). — P. 8264–8269.
- [13] Bel, G. The simplicity of completion time distributions for common complex biochemical processes / G. Bel, M. Munsky, I. Nemenman // Phys. Biol. — 2010. — Vol. 7. — P. 1–9.
- [14] Некоркин, В. И. Лекции по основам теории колебаний: учебное пособие / Некоркин, В.И. — Нижний Новгород: Нижегородский госуниверситет, 2011. — P. 233.
- [15] Expression of a gene cluster kaiabc as a circadian feedback process in cyanobacteria / M. Ishiura, S. Kutsuna, D. Aoki, et al. // Science. — 1998. — Vol. 281. — P. 1519–1523.

- [16] Goodwin, B.C. Oscillatory behavior in enzymatic control processes / B.C. Goodwin // *Adv. Enzyme Regul.* — 1965. — Vol. 3. — P. 425–428.
- [17] Johnson A. D. Poteete, A.R. Lambda repressor and cro-components of an efficient molecular switch / A.R. Johnson, A.D. Poteete, G. Lauer, et al. // *Nature.* — 1981. — Vol. 294. — P. 217–223.
- [18] Фихтенгольц, Г. М. Основы математического анализа: в 2 т. Т.1 / Фихтенгольц, Г.М. — М.: Наука, 1968. — P. 423.
- [19] Nene, N. Speed-dependent cellular decision making in nonequilibrium genetic circuits / N. Nene, J. Garcia-Ojalvo, A. Zaikin // *PLoS ONE.* — 2012. — Vol. 7. — P. 1–7.
- [20] Bifurcation dynamics in lineage-commitment in bipotent progenitor cells / S. Huang, Y.P. Guo, G. May, et al. // *Dev. Biol.* — 2007. — Vol. 305. — P. 695–713.
- [21] Dynamic filopodia transmit intermittent delta-notch signaling to drive pattern refinement during lateral inhibition / M. Cohen, M. Georgiou, N.L. Stevenson, et al. // *Dev. Cell.* — 2010. — Vol. 19. — P. 78–89.
- [22] Epigenetic stem cell signature in cancer / M. Widschwendter, H. Fiegl, D. Egle, et al. // *Nat. Genet.* — 2007. — Vol. 39. — P. 157–158.
- [23] A generalized model of the repressilator / S. Müller, J. Hofbauer, L. Endler, et al. // *J. Math. Biol.* — 2006. — Vol. 53. — P. 905–937.
- [24] Strelkowa, N. Transient dynamics around unstable periodic orbits in the generalized repressilator model / N. Strelkowa, M. Barahona // *Chaos.* — 2011. — Vol. 21. — P. 023104(1)–023104(10).
- [25] Miller, M.B. Quorum sensing in bacteria / M.B. Miller, B.L. Bassler // *Annu. Rev. Microbiol.* — 2001. — Vol. 55. — P. 165–199.
- [26] Synchronizing genetic relaxation oscillators by intercell signaling / D. McMillen, N. Kopell, J. Hasty, et al. // *Proc. Natl. Acad. Sci. USA.* — 2002. — Vol. 99. — P. 679–684.
- [27] Danino, T. A synchronized quorum of genetic clocks / T. Danino, O. Mondragón-Palomino, L. Tsimring // *Nature.* — 2010. — Vol. 463. — P. 326–330.
- [28] Entrainment of a population of synthetic genetic oscillators / O. Mondragon-Palomino, T. Danino, J. Selimkhanov, et al. // *Science.* — 2011. — Vol. 333. — P. 1315–1319.
- [29] Sensing array of radically coupled genetic biopixels / A. Prindle, P. Samayoa, I. Razinkov, et al. // *Nature.* — 2012. — Vol. 481. — P. 39–44.
- [30] A synthetic escherichia coli predator–prey ecosystem / F.K. Balagadde, H. Song, J. Ozaki, et al. // *Mol. Syst. Biol.* — 2008. — Vol. 4. — P. 187–194.
- [31] A synthetic multicellular system for programmed pattern formation / S. Basu, Y. Gerchman, C.H. Collins, et al. // *Nature.* — 2005. — Vol. 434. — P. 1130–1134.
- [32] A synthetic genetic edge detection program / J.J. Tabor, H.M. Salis, Z.B. Simpson, et al. // *Cell.* — 2009. — Vol. 137. — P. 1272–1294.

- [33] Tamsir, A. Robust multicellular computing using genetically encoded nor gates and chemical “wires” / A. Tamsir, J.J. Tabor, C.A. Voigt // Nature.— 2011.— Vol. 469.— P. 212–215.
- [34] Distributed biological computation with multicellular engineered networks / S. Regot, J. Macia, N. Conde, et al. // Nature.— 2011.— Vol. 469.— P. 207–211.
- [35] Long-term monitoring of bacteria undergoing programmed population control in a microchemostat / F.K. Balagadde, L. You, C.L. Hansen, et al. // Science.— 2005.— Vol. 309.— P. 137–140.

Татьяна Владимировна **Лаптева**
Алексей Анатольевич **Заикин**
Михаил Васильевич **Иванченко**

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ И ЧИСЛЕННЫЕ МЕТОДЫ ИССЛЕДОВАНИЯ СИСТЕМ ГЕННОЙ РЕГУЛЯЦИИ

Учебно-методическое пособие

Федеральное государственное автономное образовательное учреждение
высшего образования
«Национальный исследовательский Нижегородский государственный
университет им. Н.И. Лобачевского».
603950, Нижний Новгород, пр. Гагарина, 23.

Подписано в печать . Формат 60×84 1/16.
Бумага офсетная. Печать цифровая. Гарнитура Таймс.
Усл. печ. л. 1,8. Заказ № . Тираж 15 экз.

Редакционно-издательское управление (РИУ)
Нижегородского госуниверситета им. Н.И. Лобачевского.
603950, Нижний Новгород, пр. Гагарина, 23.