

УДК 519.677.57

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ РЕГУЛЯРНЫХ КОНТУРОВ ГЕННЫХ СЕТЕЙ¹⁾

© 2004 г. Г. В. Демиденко*, Н. А. Колчанов**,
В. А. Лихошвай***, Ю. Г. Матушкин**, С. И. Фадеев*

*630090 пр-т Акад. Коптюга, 4, Ин-т матем СО РАН;

**630090 Новосибирск, пр-т Акад. Коптюга, 2,

***628006 Ханты-Мансийск, ул. Мира, 5,

Югорский НИИ информационных технологий)

e-mail:likho@bionet.nsc.ru

Поступила в редакцию 15.12.2003 г.

Рассматриваются гипотетические генные сети (ГГС) – теоретические объекты, конструируемые из элементарных единиц двух типов: генетических элементов и регуляторных механизмов (активация и репрессия). Их функционирование описывается автономными системами дифференциальных уравнений и связанными с ними уравнениями специального вида с запаздывающими аргументами. Формулируются актуальные задачи теории функционирования ГГС, рассматриваемых как математические объекты исследования. Найдены закономерности устройства структурных графов ГГС, позволяющие предсказывать фазовые портреты функционирования ГГС без проведения численных расчетов динамических моделей. Доказана в частном случае эквивалентность замены системы автономных дифференциальных уравнений уравнением с запаздывающим аргументом. Приводятся примеры решения конкретных задач теории функционирования ГГС. Библ. 46. Фиг. 7.

Ключевые слова: математическое моделирование генных сетей, структурные графы, автономные системы обыкновенных дифференциальных уравнений.

1. ВВЕДЕНИЕ

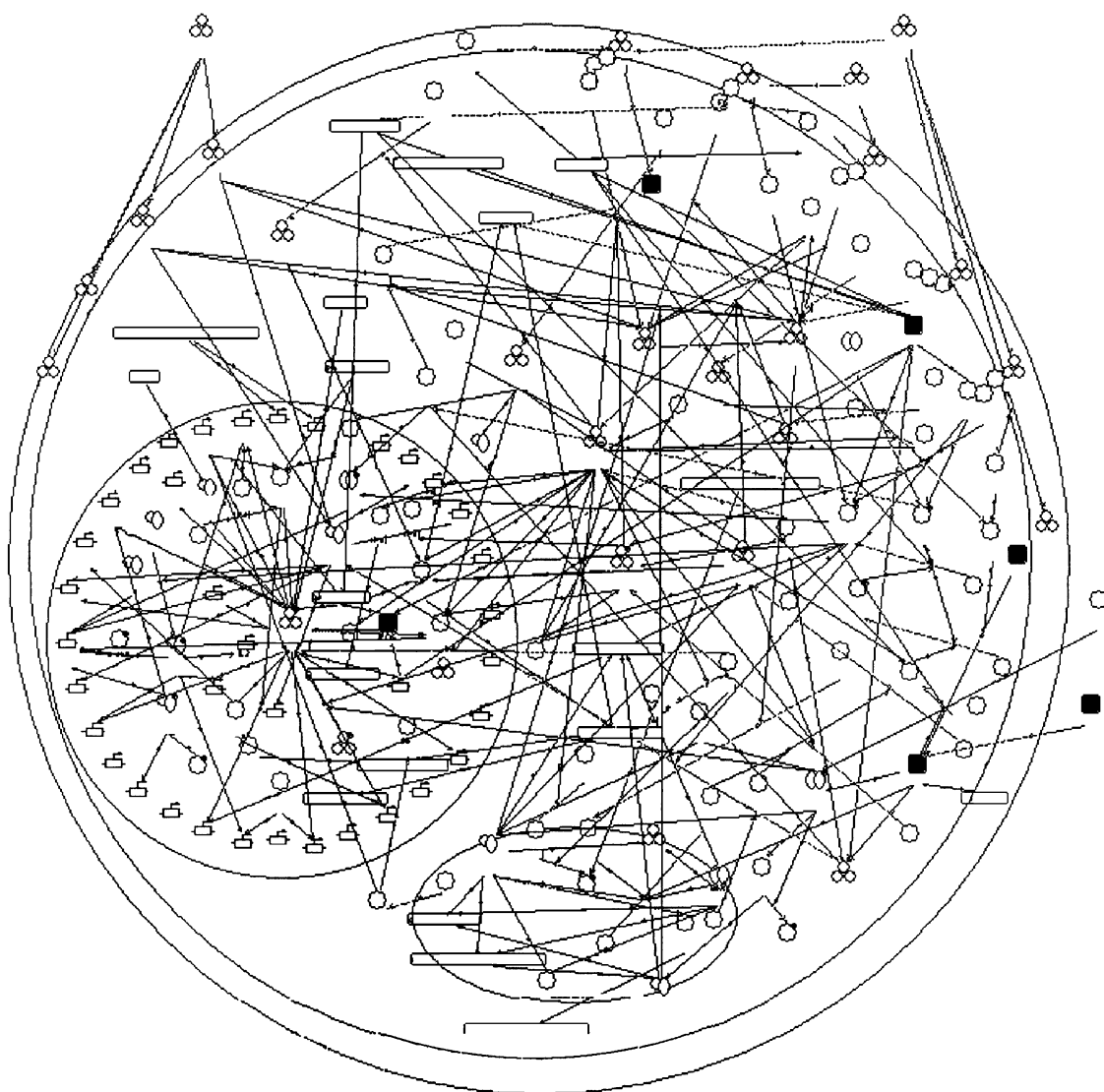
Начавшееся в последнее десятилетие массовое исследование генных сетей (ГС), генетически контролируемых метаболических путей, путей сигнальной трансдукции²⁾ и других сложных молекулярно-генетических систем приводит к исключительно высоким темпам накопления экспериментальной информации, представленной в тысячах публикаций и в компьютерных базах данных и описывающих различные аспекты функционирования ГС (см. [1]). Эти исследования выходят в настоящее время на качественно новый уровень в связи с широким распространением техники ДНК-чипов (microarray-анализа), позволяющей в одном эксперименте изучать функцию многих десятков и тысяч генов (см. [2], [3]). Анализ этих беспрецедентно огромных объемов экспериментальных данных, отражающих сложные процессы функционирования молекулярно-генетических систем, принципиально невозможен без использования современных информационных технологий и эффективных математических методов анализа данных и моделирования биологических систем и процессов. Для обобщения, систематизации и анализа накопленной информации о ГС требуется проведение широкомасштабных теоретических исследований. Это необходимо для понимания принципов их структурной организации, молекулярных механизмов функционирования, закономерностей эволюции ГС, оценки влияния мутаций на функцию ГС, реконструкции ГС на основе экспериментальных данных, для создания искусственных ГС с заданными характеристиками их функционирования (см. [4]). Именно поэтому теоретическое и компьютерное исследование ГС приобретает фундаментальное и первоочередное значение.

¹⁾Работа выполнена при частичной финансовой поддержке РФФИ (коды проектов 02-04-48802, 02-07-90359, 03-01-00328, 03-04-483506, 03-07-96833-з2003югра, 03-04-48829, 04-01-00458), Интеграционного проекта СО РАН (проект 2003-119) и Госконтракта (проект 10002-251/П-25/155-270/2004004-082).

²⁾Трансдукция – перенос бактериофагом генетического материала (участка ДНК) из одной бактериальной клетки в другую, что приводит к изменению наследственных свойств клетки. Перенесенный генетический материал может остаться в составе ДНК фага или включиться в ДНК бактерии.

Генные сети являются структурно сложными пространственными объектами, состоящими из десятков и сотен элементов разной природы и сложности: гены и их регуляторные участки, РНК и белки, кодируемые этими генами, низкомолекулярные соединения, различные комплексы между ферментами и их мишенями и т.д. (см. фиг. 1 где в очень схематичном виде представлена не самая сложная генная сеть). Элементы ГС связываются в единую функциональную систему посредством сложных нелинейных биохимических процессов синтеза и деградации веществ (см. [5]). ГС являются открытыми системами, функционирование которых поддерживается непрерывным поступлением в среду определенных веществ и энергии, а также отводом продуктов деятельности. Функционирование ГС можно характеризовать временными траекториями изменения концентраций некоторой совокупности веществ, принадлежащих ГС.

Важнейшим свойством ГС является способность к изменению состояния (концентрации веществ) в ответ на изменение условий внешней и внутренней среды. Изменение состояния достигается путем изменения уровня экспрессии определенных групп генов посредством веществ-регуляторов. Регуляторные процессы представляют собой последовательности молекулярных событий (часто достаточно сложных и разветвленных), в которых могут быть одновременно задействованы многие вещества, как поступающие извне (внешние сигналы), так и синтезируемые самой генной сетью (внутренние сигналы), а также регуляторные участки генов. Таким об-



Фиг. 1. База данных GENENET: фрагмент генной сети, контролирующей процесс клеточной смерти – апоптоза.

разом, ядром ГС являются гены и кодируемые ими РНК и белки, экспрессия которых подвержена взаимному регулированию (см. [1], [5]). Данные подсети представляют собой регуляторные контуры ГС. Изучение свойств регуляторных контуров является важнейшей задачей теории функционирования ГС, так как именно их наличие обеспечивает ГС уникальную способность адекватно реагировать на изменение внешних условий.

Не менее существенной для описания функционирования ГС является матричность синтеза основных веществ, их составляющих. Ген представляет собой часть нити ДНК, состоящей из мономеров – нуклеотидов (аденин, гуанин, тимин, цитозин), однако количество этих мономеров может достигать десятков тысяч. Хромосома может состоять из нескольких тысяч генов и огромного количества так называемой некодирующей ДНК, которая, однако, часто имеет регуляторное значение. Процесс удвоения ДНК (репликация) при размножении клеток есть процесс синтеза новой матрицы по старой, и подробное описание этого процесса путем последовательного описания присоединения нового мономера к уже частично готовой матрице невозможно по причине гигантской размерности соответствующих систем уравнений. То же можно сказать и о процессах транскрипции (синтез мРНК³⁾ по матрице ДНК) и трансляции (синтез белка по матрице РНК). Таким образом, необходимо найти адекватный способ описания этих процессов, что, по нашему мнению, можно сделать с помощью уравнений с запаздывающим аргументом.

Для моделирования ГС применялись и применяются самые различные подходы: пороговый метод [6], [7], дифференциальные и статистические уравнения [8] – [15], сети Петри [16], [17] и др. Мы используем обобщенный химико-кинетический метод [18], [19], с помощью которого создано уже около десятка полноценных моделей, обладающих прогностическим эффектом [20]–[33].

2. ГИПОТЕТИЧЕСКИЕ ГЕННЫЕ СЕТИ – МОДЕЛИ РЕГУЛЯТОРНЫХ КОНТУРОВ РЕАЛЬНЫХ ГЕННЫХ СЕТЕЙ

В ГС действие регуляторных контуров обеспечивает пространственное и динамическое согласование синтеза и деградации веществ между собой и с условиями внешней среды. При длительном сохранении постоянства условий внешней среды ГС в основном приходят в устойчивые режимы функционирования. Выделяются два типа режимов функционирования. Первый тип характеризуется постоянством во времени концентраций веществ – режим стационара. Второй тип характеризуется наличием группы веществ, концентрации которых претерпевают незатухающие изменения во времени, – осциллирующий режим. При изменении условий внешней среды (например, появление в среде некоторого вещества или изменение температуры) ГС выходит из текущего режима функционирования. Эффект реагирования ГС на изменение внешних условий также достигается в силу наличия в них регуляторных контуров. Регуляторные контуры, как правило, оказывают влияние либо на стадию инициации синтеза определенного вещества, либо на стадию его деградации (модификации, диссипации⁴⁾), так что в совокупности устанавливается новый баланс веществ, соответствующий новой ситуации. Если условия внешней среды возвращаются к исходным, то ГС также может вернуться в исходное состояние. Однако она также может перейти в новый режим функционирования, отличный от предыдущего. Если условия внешней среды переходят в некоторое новое состояние и длительное время не меняются, то ГС по прошествии некоторого времени также достигнет некоторого устойчивого состояния. Специфика ГС состоит в том, что обычно точками регуляторного воздействия являются стадии инициации транскрипции⁵⁾ или трансляции или же стадии деградации (модификации⁶⁾, диссипации) веществ. Механизмы регуляции инициации транскрипции или трансляции реализуются через специфические регуляторные участки, расположенные на ДНК и/или РНК. С ними взаимодействуют определенные вещества-регуляторы, как правило комплексы белков. Иногда в регуляторные комплексы могут входить РНК. Непосредственным результатом регуляторного воздействия является изменение скоростей синтеза продуктов, а конечной целью – изменение концентрации конечных продуктов синтеза (ими может быть РНК или белок), которые будут получены только по прошествии некоторого времени, так как синтез РНК и белков осуществляется путем матричного синтеза и представляет собой многостадийный процесс.

³⁾мРНК – матричная РНК, считываемая с ДНК гена.

⁴⁾Диссипация – распад вещества или его удаление из рассматриваемого объема, например из клетки.

⁵⁾Транскрипция – синтез мРНК по матрице ДНК.

⁶⁾Модификация – нерадикальное (химическое) изменение вещества.

Учитывая сказанное, можно записать следующую принципиальную блок-модель регуляторных контуров ГС:

$$\begin{aligned}\frac{dx_{j,i,1}}{dt} &= A_j(x_l \mid l \in B_j) - f_{j,i,1}^+ + f_{j,i,2}^- - d_{j,i,1}, \quad j = \overline{1, m}, \quad i \in P_j, \\ \frac{dx_{j,i,s}}{dt} &= f_{j,i,s-1}^+ - f_{j,i,s}^- - f_{j,i,s}^+ + f_{j,i,s+1}^- - d_{j,i,s}, \\ j &= \overline{1, m}, \quad i \in P_j, \quad s = \overline{2, n_{j,i}-1}, \\ \frac{dx_i}{dt} &= \sum_{j \in G_i} f_{j,i,n_{j,i}}^+ - g_i(x_i, x_l \mid D_i), \quad i = \overline{1, n}.\end{aligned}\tag{2.1}$$

Здесь n – количество конечных продуктов синтеза, m – количество начальных стадий синтеза продуктов (в общем случае $n \neq m$), P_j – множество номеров конечных продуктов, которые синтезируются в результате инициации j -й начальной стадии, $A_j(x_l \mid l \in B_j)$ – механизм регулирования активности инициации j -й начальной стадии, B_j – список номеров конечных продуктов генной сети, регулирующих активность инициации j -й начальной стадии синтеза, $n_{j,i}$ – количество промежуточных стадий в синтезе i -го продукта при инициации j -й начальной стадии, $f_{j,i,s}^+$ – механизм перехода из промежуточной стадии синтеза в следующую промежуточную стадию (прямой процесс), $f_{j,i,s}^-$ – механизм возврата из промежуточной стадии синтеза в предыдущую промежуточную стадию (обратный процесс), $d_{j,i,s}$ – механизм преждевременного прерывания процесса синтеза (сток), $x_{j,i,s}$ – промежуточные стадии синтеза продуктов, x_i есть i -й белок, β_i – константа диссипации i -го белка, G_i – множество номеров начальных стадий, в результате инициации которых синтезируется i -й конечный продукт, g_i – механизм распада конечного продукта.

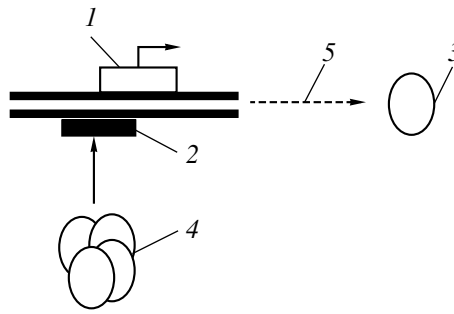
В основе функционирования ГС как молекулярно-генетических систем лежат биохимические процессы и процессы переноса веществ и энергии. Поэтому в принципе для вывода функций A, g, f, d в блок-модели (2.1) можно применить химико-кинетический подход, предполагающий выписывание соответствующих систем биохимических реакций. Ограничением в применении данного подхода для моделирования реальных регуляторных контуров ГС служит невозможность (в силу отсутствия информации) выписывания полных систем биохимических реакций, протекающих в ГС. Поэтому изучение закономерностей функционирования ГС, что равносильно изучению функционирования их регуляторных контуров, эффективнее проводить на гипотетических конструкциях. Это позволяет преодолеть приведенное выше ограничение, поскольку, оттолкнувшись от механизмов реальных ГС, можно выделить в них основные принципы строения и, комбинируя механизмы регулирования синтеза веществ, строить гипотетические регуляторные контуры ГС, которые будем называть *гипотетическими генными сетями* (ГГС) и изучать их свойства.

2.1. Модели механизмов регулирования активностей синтеза конечных продуктов (вид регуляторных функций A)

Формализацию механизма регуляции инициации синтеза РНК/белка будем проводить на основе предположения, что инициация синтеза осуществляется с некоторого участка на ДНК/мРНК. Инициаторный участок, в свою очередь, включает некоторую совокупность регуляторных веществ (комплексов белков, их модифицированных форм и т.д.), которые могут вступать в нековалентные контакты с этими участками и между собой. Будем также считать, что эти контакты приводят к формированию разных альтернативных состояний данного участка (промотора⁷⁾, если речь идет о транскрипции, или сайта инициации трансляции, если речь идет о трансляции), которые выполняют различные функции в системе.

Итак, вид регуляторной функции будем определять из механизма действия веществ-регуляторов на сайты-мишени. Рассмотрим конкретный пример. На фиг. 2 представлена диаграмма репрессируемого промотора; имеется один сайт (оператор) связывания вещества-регулятора (реп-

⁷⁾Промотор – участок ДНК, с которого начинается транскрипция.



Фиг. 2. Определение регуляторных функций: 1 – участок инициации транскрипции (промотор); 2 – сайт регуляции (оператор); 3 – синтезируемый продукт; 4 – репрессор; 5 – многостадийный процесс синтеза продукта.

рессора) и один сайт связывания РНК-полимеразы – фермента, осуществляющего транскрипцию. На промоторе могут происходить следующие события:

(i) репрессор repressor связывается с оператором и блокирует связывание РНК-полимеразы RNA-полимераза связывается с промотором и инициирует транскрипцию.

Всего выделяются три альтернативных состояния: свободный промотор p , промотор в комплексе с репрессором r и промотор в комплексе с РНК-полимеразой R . Выпишем бимолекулярные реакции, описывающие связывание репрессора и РНК-полимеразы с промотором:



и рассмотрим их в равновесии. Получим систему нелинейных уравнений относительно неизвестных p , r , R , repressor и RNA-pol:

$$\begin{aligned} k_1 p \text{ repressor} &= k_2 r, \\ k_3 p \text{ RNA-pol} &= k_4 R, \\ p + r + R &= p_0 + r_0 + R_0, \\ r + \text{repressor} &= r_0 + \text{repressor}_0, \\ R &= \text{RNA-pol} + R_0 + \text{RNA-pol}_0. \end{aligned} \quad (2.3)$$

Здесь нулевой индекс означает начальные значения переменных. Из общей теории обратимых реакций известно, что данная система всегда имеет единственное неотрицательное решение (см. [34]). Допустим, что мы нашли его каким-нибудь способом. Тогда значение R будет представлять долю активных промоторов. Умножив R на константу, имеющую смысл константы скорости инициации транскрипции, получим формулу для выражения регуляторной функции A через переменные repressor₀ и RNA-pol₀.

Описанный метод является универсальным. Для определения вида регуляторной функции записываем регуляторный механизм в виде системы обратимых реакций. Рассматриваем ее в равновесии и выписываем соответствующую систему нелинейных уравнений. Находим ее единственный неотрицательный корень. Выбираем из найденного вектора все компоненты, которые наделены соответствующей активностью, и составляем из них взвешенную сумму. Эта сумма и является необходимой регуляторной функцией. Таким образом, регуляторные функции A определяются как неявные функции от концентраций веществ-регуляторов, которые записываются в виде систем равновесных уравнений, получаемых из систем обратимых реакций. Следует отметить, что относительное неудобство неявной формы определения регуляторных функций проявляется только в том случае, если возникает необходимость явно выписать систему уравнений, так как системы типа (2.3) в общем случае не имеют аналитического решения. Однако численных проблем при этом не возникает, так как существуют эффективные универсальные методы поиска равновесных точек систем обратимых реакций (см. [34]).

Впрочем, существует достаточно простой способ замены неявной формы записи A явной рациональной формой. Пренебрежем изменением концентраций регуляторов в соответствующей системе (например, в системе (2.3) пренебрегаем изменением концентраций репрессора и РНК-полимеразы). Тогда исходная нелинейная система заменяется системой линейных уравнений относительно оставшихся переменных, которые имеют смысл альтернативных состояний регуляторной системы. В рассматриваемом примере система (2.3) заменяется на систему

$$\begin{aligned} k_1 p \text{ repressor} &= k_2 r, & k_3 p \text{ RNA-pol} &= k_4 R, \\ p + r + R &= p_0. \end{aligned} \quad (2.4)$$

Из полученной линейно системы легко находим соответствующие решения, которые и используем в качестве регуляторной функции A . Снова обращаясь к примеру, находим, что $R = p_0 \text{ RNA-pol} / (1 + K_1 \text{ repressor} + K_2 \text{ RNA-pol})$, где p_0 уже имеет смысл суммарной концентрации всех альтернативных состояний регуляторной системы, $K_1 = k_1/k_2$, $K_2 = k_3/k_4$.

2.2. Вид функций диссипации q

Важнейшим свойством ГС является наличие в них постоянного кругооборота макромолекул. Без кругооборота невозможно само их существование. Кругооборот поддерживается непрерывным синтезом макромолекул и доставкой (транспортом) их в компартменты⁸⁾ функционирования и непрерывным выводом (диссипацией) этих макромолекул из компартментов системы. Процессы диссипации, так же как и все другие процессы в живых системах, представляют собой цепочки биохимических процессов и процессов пассивного и активного переноса веществ и энергии. Поэтому для построения функций диссипации полностью применимы рассуждения, приведенные в разд. 1. В настоящей работе мы ограничивая использованием простейшей формы функции диссипации, предполагая, что скорость диссипации продукта пропорциональна его концентрации:

$$g(x) = \beta x. \quad (2.5)$$

2.3. Вид функций синтеза f^+ , f^- и функции стока d

В общем случае функции синтеза и стока являются нелинейными функциями многих переменных, имеющих смысл концентраций различных веществ биологической системы, в которой функционируют ГС. Так, если промежуточными стадиями синтеза являются стадии элонгации транскрипции, то функции перехода зависят от концентраций нуклеотидов, а также от источников энергии, переносчиками которой являются молекулы АТФ (аденозинтрифосфат) и ГТФ (гуанитрифосфат). Если это стадии элонгации трансляции, то функции перехода зависят от концентраций аминокислот, транспортных РНК⁹⁾, источников энергии и т.д. Следуя описанной выше схеме, в данном случае также можно выписать соответствующие системы обратимых реакций и определить функции перехода f в неявном или явном виде. Однако, как правило, все переменные таких функций соответствуют веществам, относящимся к таким фундаментальным процессам, как транскрипция и трансляция, а также к другим системам (сплайсинг¹⁰⁾, активный транспорт¹¹⁾ и т.д.), которые по отношению к рассматриваемой ГС являются элементами внешней среды. Поэтому естественно рассматривать эти вещества в качестве параметров. Принимая также во внимание, что в основе процессов синтеза веществ лежат биохимические реакции, приходим к выводу, что по крайней мере на начальном этапе можно ограничиться рассмотрением линейного вида функций прямого перехода

$$f^+(x) = kx$$

и положить функции возврата и стока равными нулю.

⁸⁾Компартмент – рассматриваемый в модели ограниченный объем, в котором происходят моделируемые процессы, например ядро клетки или ее цитоплазма

⁹⁾Транспортная РНК – это РНК, транспортирующая отдельные аминокислоты при процессе трансляции.

¹⁰⁾Сплайсинг – процесс вырезания не кодирующих белок кусков (интронов) в РНК в клетках, имеющих ядро.

¹¹⁾Активный транспорт – перемещение молекул с помощью других молекул.

Таким образом, мы определили класс моделей (2.1), в которых регуляторные функции являются рациональными функциями от аргументов, функции перехода и диссипации являются линейными функциями одного аргумента.

Еще раз подчеркнем, что ГС, описываемые с помощью данного класса моделей, мы называем гипотетическими генными сетями, подчеркивая тем самым, что они являются теоретическими идеализированными объектами и в таком виде в природе не существуют. Цепочки веществ и процессов, начиная со стадии инициации (возможно, регулируемой) синтеза мРНК продукта на ДНК гена и кончая конечной стадией его синтеза, будем называть *генетическими элементами* (ГЭ). В упрощенном виде ГЭ можно представить в виде “ДНК \rightarrow мРНК \rightarrow белок”, где белок осуществляет регуляцию работы другого ГЭ, связываясь с его ДНК. Таким образом, ГС являются наборами ГЭ, которые связаны друг с другом в функциональную систему посредством сети регуляторных механизмов. Сами системы (2.1) будем называть моделями ГС.

3. ОПИСАНИЕ ФУНКЦИОНИРОВАНИЯ СИСТЕМАМИ ДИФФЕРЕНЦИАЛЬНЫХ УРАВНЕНИЙ С ЗАПАЗДЫВАЮЩИМИ АРГУМЕНТАМИ

Как упоминалось выше, в ГС синтез РНК и белков включает большое количество (сотни и тысячи) промежуточных стадий. Прямое описание этих процессов приводит к тому, что системы (2.1) имеют очень большую размерность. Понятно, что это накладывает определенные ограничения на возможности проведения численного анализа таких систем. Возникает вопрос, насколько критичным для свойств систем (2.1) является наличие в них большого количества уравнений стандартного вида, описывающих стадии синтеза. Иными словами, будет ли система (2.1) сохранять принципиальные свойства при различных значениях параметров $n_{i,j}$. Следующий пример показывает, что в общем случае свойства систем (2.1) существенно зависят от числа промежуточных стадий.

Рассмотрим модель

$$\begin{aligned} \frac{dx}{dt} &= \frac{\alpha}{1+x^\gamma} - \theta x \text{ при } n = 0, \\ \frac{dx_1}{dt} &= \frac{\alpha}{1+x^\gamma} - \frac{n}{\tau} x_1, \quad \frac{dx}{dt} = \frac{n}{\tau} x_1 - \theta x \text{ при } n = 1, \\ \frac{dx_1}{dt} &= \frac{\alpha}{1+x^\gamma} - \frac{n}{\tau} x_1, \quad \frac{dx_i}{dt} = \frac{n}{\tau} (x_{i-1} - x_i), \quad i = \overline{2, n}, \\ \frac{dx}{dt} &= \frac{n}{\tau} x_n - \theta x \text{ при } n > 1, \end{aligned} \quad (3.1)$$

которая описывает саморегулирующийся синтез продукта, а количество учитываемых промежуточных стадий синтеза меняется от нуля и выше. Очевидно, при любых неотрицательных значениях $\alpha, \gamma, \tau, \theta$ и n система (3.1) имеет стационарное решение, которое находится из уравнений

$$\alpha = \theta(1+x^\gamma)x, \quad x_i = \theta \frac{\tau}{n} x.$$

При $n = 0$ (второе уравнение отсутствует) для любых значений параметров α, γ данный стационар является единственным устойчивым пределом, к которому стремятся все траектории модели (3.1). Однако при $n > 0$ при определенных значениях параметров α, γ, θ и τ у системы (3.2) уже может существовать аттрактор, имеющий незатухающие колебания. Такая ситуация, например, реализуется, если при $\alpha = 5$ и $\gamma = 3$, $\theta = 0.022$ и $\tau = 100$ начать постепенно увеличивать значение. Тогда, начиная с критического значения $n^* = 5$, стационарное решение теряет устойчивость, что приводит к возникновению автоколебания.

Таким образом, в общем случае учет промежуточного синтеза является важным элементом регуляторных контуров и его нельзя игнорировать без ущерба для полноты описания. Тем не менее имеется способ перехода от систем (2.1) к новому классу систем

$$\frac{dx_i(t)}{dt} = \sum_{j \in G_i} A_i(x_i(t - \tau_{j,i}) \mid l \in B_j) - g_i(x_i(t), x_i(t) \mid l \in D_i), \quad i = 1, 2, \dots, n, \quad (3.2)$$

Здесь число переменных в точности равно числу различных конечных продуктов и при этом не происходит потери свойств систем (2.1). Все параметры и обозначения, кроме $\tau_{j,i}$, имеют тот же смысл, что и для (1), а $\tau_{j,i}$ имеет смысл времени запаздывания появления конечного продукта после начала его синтеза.

Предлагаемый переход основывается на следующей гипотезе, которую мы назовем *предельной теоремой*.

Рассмотрим систему (2.1), для которой в нулевой момент заданы начальные условия

$$x_i(0) = 0, \quad x_{j,i,1}(0) = C_{j,i}, \quad x_{j,i,s}(0) = 0, \quad i = \overline{1, n}, \quad j = \overline{1, m}, \quad s = \overline{2, n_{j,i}}. \quad (3.3)$$

Также рассмотрим $\sum_{j=\overline{1, m}} |P_j$ систем

$$\begin{aligned} \frac{dz_{j,i,1}}{dt} &= -f_{j,i,1}^+ + f_{j,i,2}^- - d_{j,i,1}, \\ \frac{dz_{j,i,s}}{dt} &= f_{j,i,s-1}^+ - f_{j,i,s}^- - f_{j,i,s}^+ + f_{j,i,s+1}^- - d_{j,i,s}, \\ s &= \overline{2, n_{j,i}-1}, \quad j = \overline{1, m}, \quad i \in P_j. \end{aligned}$$

Назовем их *сопутствующими* системе (2.1). Зададим начальные данные: $z_{j,i,1}(0) = 1$ и $z_{j,i,s}(0) = 0$, $j = 1, 2, \dots, m$, $i \in P_j$, $s = 1, 2, \dots, n_{j,i}$, и решим задачу Коши на интервале $[0, \infty)$. Вычислим следующие величины:

$$\tau_{j,i}^{n_{j,i}} = \int_0^\infty t f_{j,i,n_{j,i}}^+ dt, \quad z_{j,i}^{n_{j,i}} = \max_{s=\overline{1, n_{j,i}}} \sup_{t \geq 0} z_{j,i,s}(t).$$

Предположим, что существуют пределы

$$\lim_{n_{j,i} \rightarrow \infty} \tau_{j,i}^{n_{j,i}} = \tau_{j,i}, \quad \lim_{n_{j,i} \rightarrow \infty} z_{j,i}^{n_{j,i}} = 0.$$

Возьмем теперь систему (3.2), соответствующую системе (2.1). Пусть в ней запаздывания $\tau_{j,i}$ положены равными пределам. Пусть все $\tau_{j,i}$, рассмотренные в совокупности по всем генетическим элементам, имеют k различных значений. Обозначим их через t_r , $r = 1, 2, \dots, k$. Не ограничивая общности, можем считать, что $0 \leq t_1 < \dots < t_k$. Обозначим через N_r множество пар номеров (j, i) таких, что запаздывание $\tau_{j,i}$, входящее в регуляторный механизм A_j в системе (2.1), равно t_r .

По определению, для любых номеров $1 \leq j \leq m$ и $1 \leq i \leq n$ полагаем

$$A_j(x_i(t - \tau_{j,i}) \mid l \in B_j) \equiv 0, \quad x_i(t - \tau_{j,i}) = 0 \text{ при } t - \tau_{j,i} < 0, \quad x_i(0) = 0.$$

Через $y_1(t), \dots, y_n(t)$ обозначим функции, которые получаются следующим образом. Если $t_1 > 0$, то полагаем $y_i(t) = 0$, для всех $0 \leq t < t_1$, $i = 1, 2, \dots, n$. Затем переходим к последовательному решению задачи (2.1), (3.3) на интервалах $(t_1, t_2), \dots, [t_{k-1}, t_k), [t_k, \infty)$. Если $t_1 = 0$, то сразу переходим к последовательному решению задачи (2.1), (3.3) на интервалах $[t_1, t_2), \dots, [t_{k-1}, t_k), [t_k, \infty)$, а именно: задав в момент времени t_1 начальное условие

$$y_i(t_1) = \sum_{j: (j,i) \in N_1} C_j,$$

решаем задачу (3.2), (3.3) на интервале $[t_1, t_2)$, откуда находим

$$\lim_{\varepsilon \rightarrow 0} y_i(t_2 - \varepsilon) = y_i(t_2^-).$$

Затем переходим к решению задачи (2.1), (3.3) на интервале $[t_2, t_3)$ с начальным условием

$$y_i(t_2) = y_i(t_2^-) + \sum_{j: (j,i) \in N_2} C_j,$$

что позволяет определить

$$\lim_{\varepsilon \rightarrow 0} y_i(t_3 - \varepsilon) = y_i(t_3^-).$$

Данную процедуру повторяем для каждого следующего интервала, вплоть до последнего. В результате получаем искомые функции $y_i(t)$, определенные на всей временной оси.

Гипотеза. Для любого $t \geq 0$ имеем

$$\lim_{n_{1,i} \rightarrow \infty, \dots, n_{m,i} \rightarrow \infty} x_i(t) = y_i(t).$$

В настоящее время предельная теорема доказана для частного случая линейных функций перехода вида (2.5) и $k_{j,i,s}^{n_{j,i}} = n_{j,i}/\tau_{j,i}$, $s = \overline{1, n_{j,i}}$ (см. [35]).

4. ИЗУЧЕНИЕ ПРЕДЕЛЬНЫХ СВОЙСТВ ГИПОТЕТИЧЕСКИХ ГЕННЫХ СЕТЕЙ

Системы (3.2) представляют собой очень широкий класс теоретических объектов, так как в качестве регуляторных функций могут выступать практически любые неотрицательные рациональные функции. Поэтому разнообразие режимов функционирования, которые можно у них отыскать, поистине неисчерпаемо. Они могут иметь множественные точки покоя и колебательные циклы практически в любых комбинациях (см. [36] – [39]). Не исключено также, что некоторые конструкции ГГС могут иметь странные аттракторы. На это косвенно указывают исследования поведения булевых кусочно-линейных моделей ГС, которые являются предельным случаем систем (3.2), где запаздывания приняты равными нулю. В связи с этим трудно ожидать, что для систем (3.2) общего вида можно выявить сколько-нибудь универсальные закономерности. Скорее всего их следует искать для некоторых более узких классов моделей. В настоящей работе мы выделим четыре важных с биологической точки зрения класса моделей. Оказывается, что в каждом классе предельные режимы функционирования уже достаточно тесно связаны с определенными характеристиками соответствующих структурных графов.

4.1. Результаты численного исследования классов ГГС.

Классы ГГС определяются путем ввода для каждого из них различных ограничений на разнообразие используемых регуляторных механизмов некоторые общие ограничения.

1. Из рассмотрения исключаются механизмы активации процессов инициации синтеза веществ, т.е. в ГГС будем рассматривать исключительно репрессивные механизмы – отрицательные регуляторные связи.

2. Если ГГС состоит более чем из одного генетического элемента, то продуктам запрещается прямо регулировать (положительно или отрицательно) активность инициации собственного синтеза. Исключение сделаем только для простейшего случая генной сети, состоящей из одного генетического элемента.

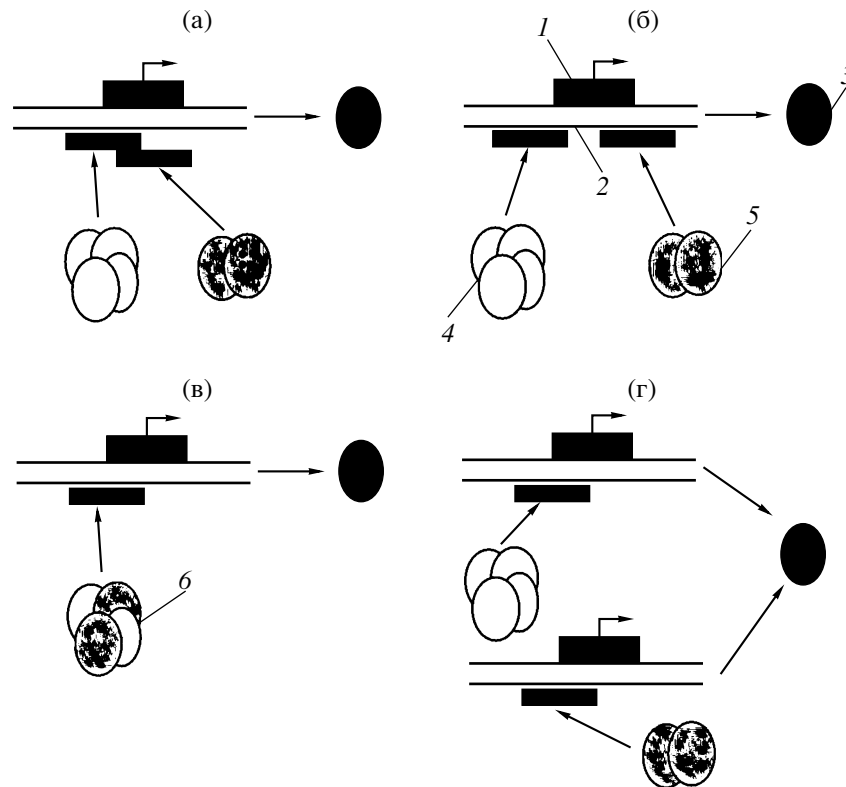
3. Многообразие механизмов диссипации всех продуктов ограничивается простейшим случаем: считаем, что скорость убывания концентрации любого продукта пропорциональна его концентрации (см. п. 2.2).

4. Рассматриваются только четыре различных типа регуляторных механизмов.

К типу 1, схематически представленному на фиг. 3а, отнесем регуляторные механизмы, которые удовлетворяют перечисленным общим ограничениям, а также следующим дополнительным специальным условиям: 1) регуляторами являются *гомомультимеры* продуктов; 2) ингибирование является конкурентным.

Класс 1. Его составляют ГГС, в которых каждый продукт кодируется единственным ГЭ, а все регуляторные механизмы относятся к типу 1. Математическими моделями ГГС этого класса являются системы вида

$$\frac{dx_i(t)}{dt} = \frac{\alpha_i}{1 + \sum_{k \in D_i} \delta_{i,k} x_k^{\gamma_k}(t - \tau_{i,i})} - \beta_i x_i(t), \quad i = 1, 2, \dots, n. \quad (4.1)$$



Фиг. 3. Четыре типа регулирования эффективности экспрессии генетических элементов: 1 – участок инициации транскрипции (промотор); 2 – сайт регуляции (оператор) 3 – синтезируемый продукт; 4, 5 – репрессоры, состоящие из одинаковых субъединиц; 6 – репрессор, состоящий из разных субъединиц.

Здесь и ниже D_i в формулах (4.1) – (4.4) обозначает множество номеров продуктов в некоторой заранее фиксированной нумерации, формирующих гомомультимеры-регуляторы активности синтеза i -го продукта.

К *типу 2* (схема представлена на фиг. 3б) отнесем механизмы, удовлетворяющие следующим дополнительным условиям: 1) регуляторы являются гомомультимерами продуктов, 2) ингибирование не является конкурентным.

Класс 2. К этому классу будем относить только такие ГГС, в которых каждый продукт кодируется единственным ГЭ, а механизмы относятся к типу 2. регулирования. ГГС класса 2 моделируются системами вида

$$\frac{dx_i(t)}{dt} = \frac{\alpha_i}{\prod_{k \in D_i} [1 + \delta_{i,k} x_k^{\gamma_{i,k}} (t - \tau_{i,i})]} - \beta_i x_i(t), \quad i = 1, 2, \dots, n. \quad (4.2)$$

Класс 3. ГГС этого класса будем строить только с использованием регуляторных механизмов *типа 3*. К ним отнесем механизмы, удовлетворяющие следующим дополнительным условиям: регуляторами являются *гетеромультимеры* продуктов; для каждого ГЭ имеется не больше одного регулятора его активности (фиг. 3в). По-прежнему будем считать, что в ГГС класса 3 каждый продукт синтезируется единственным ГЭ. ГГС такого класса моделируются системами вида

$$\frac{dx_i(t)}{dt} = \frac{\alpha_i}{1 + \delta_i \prod_{k \in D_i} x_k^{\gamma_{i,k}} (t - \tau_{i,i})} - \beta_i x_i(t), \quad i = 1, 2, \dots, n. \quad (4.3)$$

Класс 4. Этому классу принадлежат ГГС, в которых регуляторные механизмы относятся к *типу 4* и действия, т.е. регуляторами являются гомомультимеры белков и для каждого ГЭ имеется не больше одного регулятора его активности. При конструировании ГГС класса 4 допуска-

ется, что продукт может синтезироваться более чем одним ГЭ (фиг. 3г). ГГС класса 4 моделируются системами вида

$$\frac{dx_i(t)}{dt} = \sum_{k \in D_i} \frac{\alpha_{i,k}}{1 + \delta_{i,k} x_k^{\gamma_{i,k}}(t - \tau_{k,i})} - \beta_i x_i(t), \quad i = 1, 2, \dots, n. \quad (4.4)$$

Введенные типы являются составными элементами реально существующих в природе регуляторных механизмов. Например, типичным механизмом репрессии транскрипции в прокариотах является взаимодействие репрессора(ов) с определенным(и) сайтом(ами), что приводит к маскировке на ДНК соответствующего сайта посадки РНК-полимеразы. При этом могут встречаться одиночные и множественные сайты связывания репрессоров (фиг. 3, сайты могут стерически взаимопересекаться (фиг. 3а) или не иметь пересечений (фиг. 3в), репрессоры могут быть гомомультимерами (фиг. 3а, б, г) или гетеромультимерами (фиг. 3в), могут существовать множественные сайты инициации транскрипции (фиг. 3г). Многочисленные конкретные примеры можно найти, например, в базе EcoCyc (<http://biocyc.org>).

Хотя введенные типы регуляции в чистом виде в природе обычно не встречаются, представляется не только полезным, но и необходимым провести изучение свойств ГГС классов 1–4. Знание их свойств позволит перейти к предметному изучению свойств конструкций ГГС более общего вида.

4.2. Представление ГГС в виде графов

Пусть фиксирована некоторая типотетическая ГС произвольного вида, а не обязательно класса 1–4. Построим ориентированный граф, в котором вершинам сопоставим продукты ГЭ, а дугами отметим вхождение одних продуктов (дуги исходят из соответствующих вершин) в механизмы регуляции инициации синтеза других продуктов (дуги входят в соответствующие вершины). Назовем данный граф структурным графом ГГС. В общем случае структурный граф отражает лишь часть информации о строении ГГС. Для ее определения необходимо также задать вид регуляторных механизмов и параметры этих механизмов. Таким образом, при определении ГГС естественно выделяются три иерархических информационных уровня. Первый уровень структурный. На нем в виде ориентированного графа фиксируется сеть регуляторных отношений в ГГС. Второй уровень функциональный. На нем в рамках заданной структурной организации задаются конкретные функции, моделирующие регуляторные механизмы. третий уровень параметрический. На нем задаются конкретные значения параметров функционально верифицированных механизмов регулирования процессов инициации синтеза веществ и их деградаций.

Естественно встает вопрос о вкладе каждого иерархического информационного уровня описания в формирование свойств ГГС. Наиболее важным представляется изучение условий, в которых задание структурного орграфа уже полностью определяет такие важные свойства, как стационары и/или колебательные режимы. Представляется достаточно очевидным, что в общем случае задание структурного графа еще не позволяет полностью определить функциональный портрет ГГС. Для примера возьмем полный ориентированный граф, имеющий три вершины, и построим на нем следующие модели:

$$\begin{aligned} \frac{dx_1}{dt} &= \frac{\alpha}{1 + x_2^{\gamma_1} + x_3^{\gamma_2}} - x_1, & \frac{dx_2}{dt} &= \frac{\alpha}{1 + x_3^{\gamma_1} + x_1^{\gamma_2}} - x_2, \\ \frac{dx_3}{dt} &= \frac{\alpha}{1 + x_1^{\gamma_1} + x_2^{\gamma_2}} - x_3; \end{aligned} \quad (4.5)$$

$$\begin{aligned} \frac{dx_1}{dt} &= \frac{\alpha}{1 + x_2^{\gamma} x_3^{\gamma}} - x_1, & \frac{dx_2}{dt} &= \frac{\alpha}{1 + x_3^{\gamma} + x_1^{\gamma}} - x_2, \\ \frac{dx_3}{dt} &= \frac{\alpha}{1 + x_1^{\gamma} + x_2^{\gamma}} - x_3. \end{aligned} \quad (4.6)$$

Если положить $\alpha = 10$, $\gamma = \gamma_1 + \gamma_2 = 5$, то система (4.5) будет иметь три устойчивых стационара, а система (4.6) – один. Таким образом, данные ГГС имеют идентичные структурные графы, но разные предельные портреты функционирования. Знание функциональной организации (вто-

рой уровень) ГГС еще не гарантирует знание его предельного портрета. Для того чтобы в этом убедиться, достаточно рассмотреть поведение модели (4.5) с четырьмя разными наборами параметров. Первый – описанный выше, второй $\alpha = 10$, $\gamma_1 = 0$, $\gamma_2 = 0$, третий $\alpha = 10$, $\gamma_1 = 0$, $\gamma_2 = 5$, четвертый $\alpha = 10$, $\gamma_1 = 1.4$, $\gamma_2 = 5$. При первом наборе параметров, как уже говорилось, имеются три стационара, при втором – один стационар, при третьем – один колебательный режим, а при четвертом – три стационара и один колебательный режим функционирования.

Таким образом, чтобы подойти к исследованию взаимоотношений структурного уровня организации ГГС с их динамическими характеристиками, необходимо максимально исключить произвол задания информации на втором (функциональном) и третьем (параметрическом) уровнях. Следовательно, выделение отдельных классов ГГС не только биологически оправдано, но также решает данную методическую задачу, так как по заданному ориентированному структурному графу модели (4.1) – (4.4) определяются однозначно с точностью до параметров.

Однако, как следует из рассмотренного примера, даже в этом случае от конкретных значений параметров предельный портрет зависит достаточно сильно. Чтобы максимально проявить роль структурного уровня организации ГГС в формировании их предельных портретов функционирования, параметрическое многообразие в моделях (4.1) – (4.4) также сведем к минимуму. С этой целью наложим на параметры ограничения:

$$\tau_{k,i} = \tau, \quad \alpha_i = \alpha_{i,k} = \alpha, \quad \delta_i = \delta_{i,k} = 1, \quad \gamma_{i,k} = \gamma, \quad \beta_i = 1.$$

Назовем ГГС классов 1–4 с данным ограничениями *каноническими* (кГГС).

4.3. Исследование канонических ГГС

Продemonстрируем многообразие поведения ГГС на нескольких примерах. Сначала обратимся авторепрессируемого генетического элемента:

$$\frac{dx}{dt} = \frac{\alpha}{1 + x^\gamma(t - \tau)} - x. \quad (4.7)$$

При любых неотрицательных значениях α , γ и $\tau \geq 0$ уравнение (9) имеет стационарное решение, которое находится из уравнения

$$\alpha = (1 + x^\gamma)x.$$

При $\tau = 0$ для любых значений параметров α , γ этот стационар является единственным устойчивым пределом, к которому стремятся все траектории уравнения (4.7). Однако при других значениях параметров α , γ и τ в уравнении (4.7) может также иметься аттрактор с незатухающими колебаниями. Такая ситуация, например, реализуется, если при $\alpha = 5$ и $\gamma = 3$ начать постепенно увеличивать значение τ . Тогда, начиная с критического значения $\tau^* = 0.9643$, стационарное решение теряет устойчивость, что приводит к возникновению автоколебаний. Можно показать, что критические значения τ^* для уравнения (4.7) вычисляются по формуле

$$\tau^* = \frac{\pi - \varphi}{\sqrt{\omega^2 - 1}}, \quad \varphi = \arctg(\sqrt{\omega^2 - 1}), \quad \omega = \gamma - \left(\frac{\gamma}{\alpha}\right)^{\gamma/(\gamma+1)}.$$

Отметим несколько важных моментов. Во-первых, для возникновения автоколебаний в системе (17) необходимым условием является выполнение неравенства $\gamma > 1$, т.е. регуляторная функция должна убывать по x быстрее гиперболической, или, говоря иначе, регуляторная функция должна быть в достаточной степени пороговой. В природных системах высокая пороговость может достигаться разными способами. Наиболее простым является мультимеризация активной молекул репрессора: чем больше количество белковых мономеров, составляющих репрессор, тем выше уровень пороговости. Вторым необходимым условием является требование, чтобы величина параметра α была достаточно большой, т.е. для потенциального наличия в системе колебаний процесс инициации синтеза также должен иметь достаточно высокую базальную эффективность функционирования. Только при одновременном выполнении этих условий существует такое критическое запаздывание τ^* , превышение которого приводит к возникновению колебаний. Открытие цикла в авторегулируемой системе, состоящей из одного генетического элемента, говорит о возможности существования таких систем в природе. В природе действительно имеются автоингибируемые системы. В ГС автоингибирование достаточно широко распространено в системах трансляции. В частности, трансляция некоторых белков *E. Coli* регули-

руется по принципу отрицательной обратной связи [40]–[42]. Такие системы потенциально могут иметь колебательные режимы функционирования. Однако реально такие колебания пока не наблюдались. Одно из возможных объяснений состоит в том, что такие системы в биологической норме не должны иметь колебательные режимы функционирования. Тогда они должны функционировать в параметрической области, исключающей колебания. Так как параметры α и τ в силу специфики процессов могут претерпевать некоторые изменения в сторону увеличения или уменьшения, то следует ожидать, что в таких системах надежное избегание автоколебаний в большинстве случаев должно достигаться за счет того, что активными формами репрессоров являются мономеры белков.

Возьмем теперь полный орграф, состоящий из двух вершин, и построим на нем кГГС. В силу простоты орграфа, все классы порождают одну и ту же модель:

$$\frac{dx_1}{dt} = \frac{\alpha}{1 + x_2^\gamma(t - \tau)} - x_1, \quad \frac{dx_2}{dt} = \frac{\alpha}{1 + x_1^\gamma(t - \tau)} - x_2. \quad (4.8)$$

Данная система моделирует взаимную репрессию активности генетических элементов. Такие генетические системы известны в природе под названием молекулярного триггера. Они выполняют роль переключателей из одного режима функционирования генной сети в другой [6], [40]–[43]. Условия наличия двух стационаров у данного типа ГС исследовались ранее в [43] в математической модели вида

$$\frac{dx_1}{dt} = \frac{\alpha_1}{1 + x_2^{\gamma_1}} - x_1, \quad \frac{dx_2}{dt} = \frac{\alpha_2}{1 + x_1^{\gamma_2}} - x_2,$$

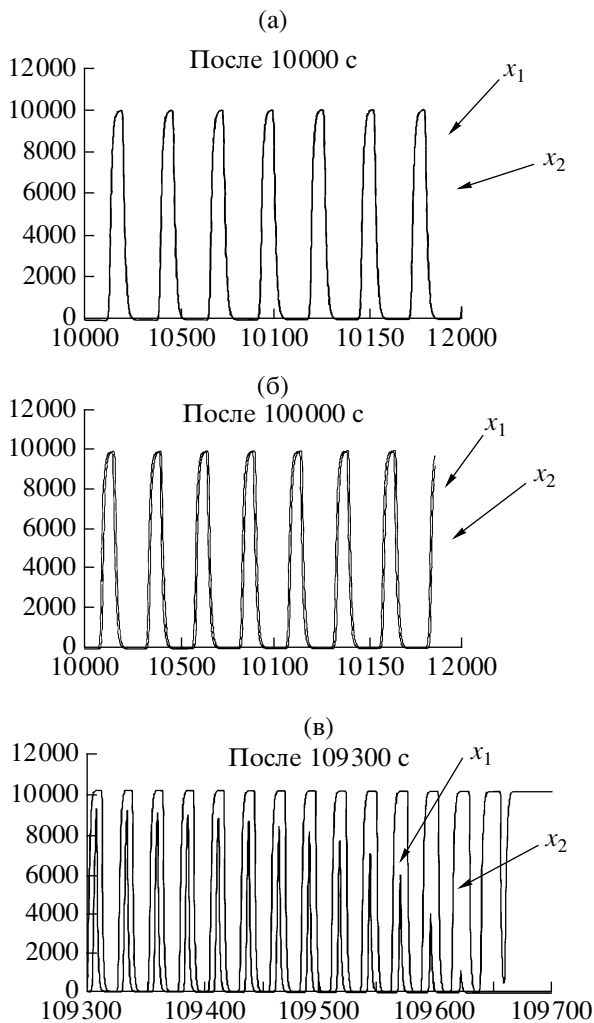
отличающейся от (4.8) наличием разных параметров α и γ для каждого генетического элемента, а также отсутствие параметров запаздывания. В [43] показано, что для того, чтобы двухоперонная, т.е. состоящая из двух взаимно репрессирующих ГЭ, генная сеть могла функционировать как молекулярный триггер, необходимо, чтобы взаимное ингибирование осуществлялось мультимерами порядка выше единицы (т.е. $\gamma_1 > 1$, $\gamma_2 > 1$) и значения параметров, имеющих смысл максимальных скоростей синтеза белков α_1 и α_2 , которые достигаются только в отсутствие репрессоров, были достаточно велики. Отметим, что данный вывод качественно совпадает с аналогичным выводом, сделанным выше. Как будет показано далее, совпадение не является случайным.

Исследование системы (4.8) также выявляет наличие только этих двух типов поведения. Поэтому разнообразие по параметрам α и γ не является существенным фактором формирования триггерных свойств системы. Критичным является их величина. Для придания кГГС свойств молекулярного триггера также требовалось, чтобы параметры α и γ были достаточно велики. Запаздывания также не являются фактором формирования предельных свойств системы. Однако хотя запаздывания не приводят к формированию новых предельных свойств, например устойчивых колебаний, как это происходило в авторегулируемой системе из одного ГЭ, тем не менее роль запаздывания в эволюции траектории системы к одному из предельных состояний оказалась весьма существенной.

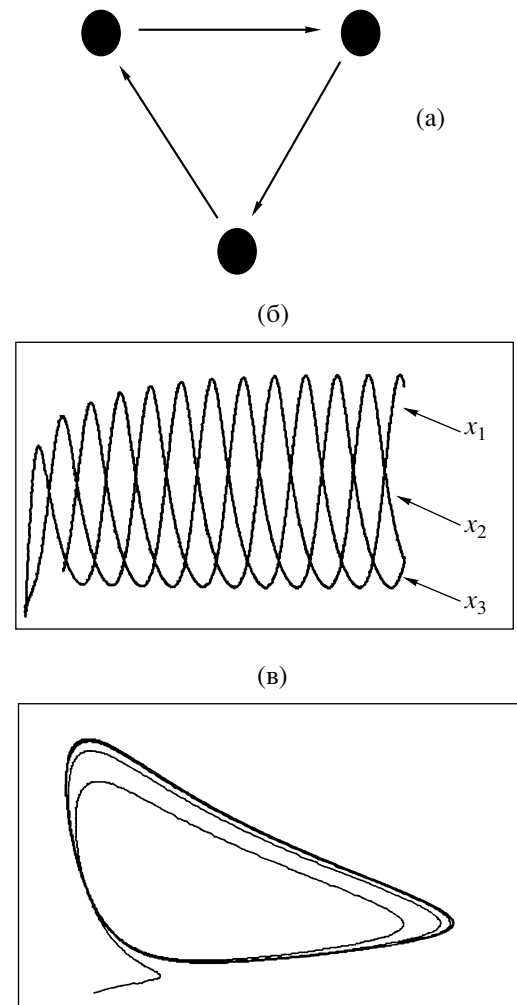
Численно было выявлено, что запаздывание может приводить к возникновению переходных режимов с очень длительной осцилляцией, так что визуально кажется, что система вышла на некоторый предельный автоколебательный режим. Пример такого поведения показан на фиг. 4. Расчет проведен с параметрами $\alpha = 10^4$, $\gamma = 3$, $\tau = 10$. Переходный период длится 109650 с. Причем через 10000 с две кривые настолько близки, что на графике их нельзя различить (фиг. 4а). На данном этапе эволюции системы кажущийся период составляет примерно 27 с. Различия между кривыми становятся заметными только через 100000 с (фиг. 4б). При этом кажущийся период практически не изменяется. Переходный период заканчивается примерно через 109400 с. Причем на переход от осцилляции к полному покою требуется не больше 300 с, что исчезающе мало по сравнению с длительностью переходного периода (фиг. 4в).

На фиг. 5а дан пример эволюции ГГС, построенной на циркулярном орграфе $G_{3,2}$, в области больших значений параметров α и γ к устойчивому прямому циклу (определение циркулярных орграфов данного типа приводится ниже). Построенные на этом орграфе кГГС приводят к системе уравнений (4.5) с параметрами $\gamma_1 = 0$, $\gamma_2 = \gamma$. Для этой системы также выявляются две параметрические области. В одной имеется только одна точка покоя, а в другой точка покоя сменяется автоколебаниями (фиг. 5б, в).

Усложнение орграфа ведет к росту сложности поведения соответствующей кГГС. Рассмотрим кГГС, построенную на орграфе $G_{6,4}$ в области достаточно больших значений параметров α



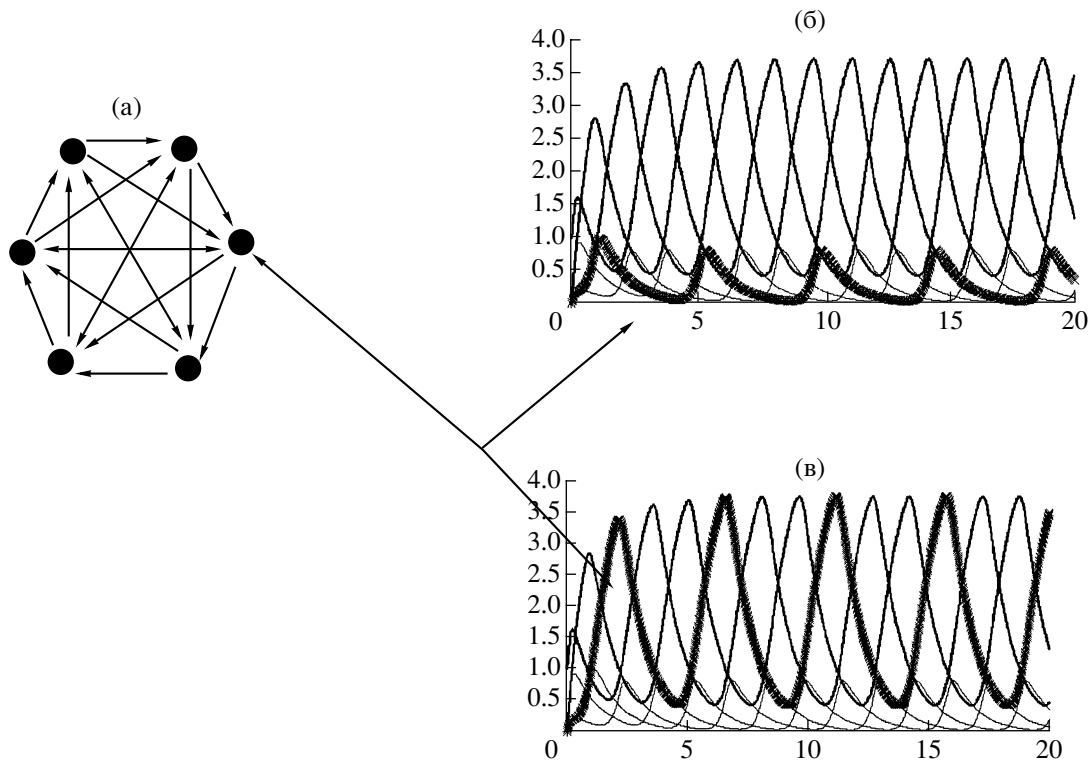
Фиг. 4. Эволюция “молекулярного триггера” из начального состояния $x_1 = 0$, $x_2 = 1.2$ в точку покоя.



Фиг. 5. Эволюция ГГС, построенной на циркулярном графе $G_{3,2}$: (а) граф $G_{3,2}$; (б) эволюция системы во времени; (в) фазовая кривая в плоскости (x_1, x_2) .

и γ к двум автоколебательным аттракторам (фиг. 5). Хотя структура орграфа усложнилась, численный анализ по-прежнему выявляет две крайние параметрические области: область “достаточно малых значений параметров α и γ ”, в которой система может существовать только в одном устойчивом стационаре, и область “достаточно больших значений параметров α и γ ”, в которой система может существовать в двух различных устойчивых состояниях и каждое состояние является автоколебательным (фиг. 6б, в).

Численный анализ других конструкций гипотетических ГС, которые мы не будем рассматривать в этой работе, подтвердил существование в параметрическом многообразии значений параметров α и γ двух областей. Первая область, L , характеризуется относительно малым значением одного или обоих параметров α и γ . В ней существует только один устойчивый стационар. Если значения параметров α и γ не входят в L , то предельные свойства уже существенно зависят от структуры графа. Решающим для выявления свойств становится численный анализ модели кГГС при заданных (α, γ, τ) . При этом для больших размерностей графа (не при всех значениях (α, γ, τ) , взятых вне L) будет наблюдаться один и тот же предельный портрет, т.е. имеется также зависимость от параметров. Тем не менее при численных расчетах было замечено, что можно выделить вторую связную область параметров H – область достаточно больших значений одного или обоих параметров α и γ , в которой формирование предельных режимов функционирования кГГС определяется только структурным графом и не зависит от параметров. В некотором смысле области L и H являются крайними проявлениями единства структуры и функции. В области L



Фиг. 6. Эволюция ГГС класса 1, построенной на циркулярном графе $G_{6,4}$: (а) граф $G_{6,4}$; (б) эволюция системы во времени из точки $(1, 0, 0, 0, 0, 0)$; (в) эволюция системы во времени из точки $(0, 1, 0, 0, 0, 0)$.

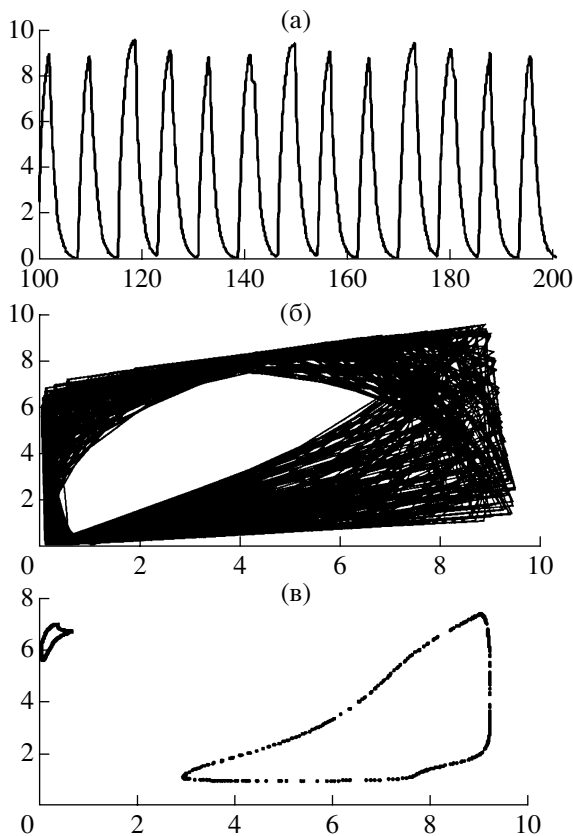
предельный портрет кГГС достаточно простой и не зависит от структурного графа. В области H , напротив, особенности предельного портрета полностью определяются структурным графом.

При малых размерностях орграфа (до пяти вершин), как правило, не удавалось выявить дополнительные области, в которых предельные свойства кГГС отличаются от свойств, наблюдаемых в областях L и H . При больших размерностях орграфа (больше пяти) такие промежуточные области значений параметров уже выявляются. Для примера рассмотрим модель кГГС класса 3, построенной на орграфе $G_{21,5}$. Данная модель при достаточно малых значениях α и γ (например, при $\tau = 0$, $\gamma = 1$ и $\alpha = 10$) имеет только одну устойчивую точку покоя, а при достаточно больших значениях α и γ (например, при $\tau = 0$, $\gamma = 10$ и $\alpha = 10$) модель имеет только один устойчивый предельный цикл.

Однако при промежуточных значениях (например при $\tau = 0$, $\alpha = 10$, $\gamma = 5$) фазовый портрет модели выглядит иначе: имеется три устойчивых колебательных режима функционирования. Из них два являются циклами, а один не имеет периода. На фиг. 7 приводится график неперiodического поведения, открытого в модели $M_3(21, 5)$. На фиг. 7а показаны временные кривые первой переменной, рассчитанной по модели (2.2), построенной на структурном графе $G_{21,5}$ в области H при $\tau = 0$, $\alpha = 10$, $\gamma = 5$ и начальных данных $x_1(0) = 1$, $x_i(0) = 0$, $i = 2, \dots, 21$. (Заметим, что на периодические режимы системы выходит при задании начальных данных $x_1(0) = 10^{-8}$, $x_i(0) = 0$, $i = 2, \dots, 21$, или $x_1(0) = 10^{+8}$, $x_i(0) = 0$, $i = 2, \dots, 21$, соответственно.)

Апериодичность 21-мерной кривой хорошо видна на фазовой кривой, представленной на фиг. 7б. Анализ отпечатков траектории на плоскости Пуанкаре, проходящей через главную диагональ 21-мерного α -куба $[0 : \alpha, \dots, 0 : \alpha]$ (фиг. 7в)) позволяет предположить, что рассматриваемая кривая является квазипериодической. Приведенный пример примечателен еще и тем, что демонстрирует апериодическое поведение у полностью симметричной ГГС.

Анализ свойств структурных графов в сравнении с динамическими свойствами соответствующих кГГС, выявляемых при “достаточно больших значениях параметров”, позволил связать предельные портреты кГГС в области H с рядом свойств ориентированных графов. Мы полагаем, что выявленные закономерности выполняются для всех конструкций кГГС, а не только для про-



Фиг. 7. Эволюция ГГС класса 3, построенной на циркулярном графе $G_{21,5}$: (а) эволюция системы во времени из точки $(1, 0, \dots, 0)$; (б) фазовая кривая в плоскости (x_1, x_2) ; (в) проекция плоскости Пуанкаре на двумерную грань (x_1, x_2) .

Так как симметричные орграфы полностью характеризуются числом вершин n и полустепенью исхода $(k-1)$, то обозначим данные орграфы через $G_{n,k}$. Очевидно, при фиксированном n существует ровно $n-1$ различных $G_{n,k}$, $k=2, 3, \dots, n$.

Рассмотрим орграфы $G(V_1, W_1), \dots, G(V_m, W_m)$, $m > 0$. Орграф $G(V, W)$, удовлетворяющий условиям

$$V = \bigcup_{i=1}^m V_i,$$

$$W = \bigcup_{i=1}^m (W_i \cup \{(u, v) \mid u \in V_i, v \in V \setminus V_i\}),$$

назовем *сильным объединением орграфов* $G(V_i, W_i)$, $i=1, 2, \dots, m$.

Критерий количества стационаров. Пусть задан орграф $G(V, W)$. Построим на нем кГГС класса 1–4. Тогда для любого заранее фиксированного τ существует такое $\gamma_0(\tau)$, что для любого $\gamma > \gamma_0$ существует такое $\alpha_0(\gamma, \tau)$, что для любого $\alpha > \alpha_0$ число устойчивых стационаров кГГС классов 1, 2 (3, 4) равняется числу $1 - (1m)$ -баз орграфа $G(V, W)$.

Из данного критерия с необходимостью следует, что если структурный граф не имеет $1 - (1m)$ -баз, то соответствующая кГГС для любого τ и достаточно больших γ и α не будет иметь стационаров и, следовательно, будет иметь хотя бы один колебательный режим функционирования. Колебательные свойства кГГС пока не удается связать с какими-либо специфическими характеристиками строения структурных графов общей конструкции. В случае циркулярного структурного графа удастся не только значительно упростить исходную формулировку критерия количества стационаров, но также удастся описать циклические свойства кГГС.

веренных в численном эксперименте. Поэтому выявленные закономерности сформулируем в виде общих критериев (которые имеют статус гипотез, требующих доказательства).

В дальнейшем изложении будут использоваться понятия и термины теории графов. В основном мы будем следовать монографии [44].

Пусть задан некоторый ориентированный граф $G = G(V, W)$, где V – множество вершин, W – множество дуг. Пусть $u, v \in V$, $(u, v) \in W$. Назовем вершину u смежной к v , а v – смежной из u . Полустепенью захода (исхода) вершины v назовем число вершин, смежных из/к v . 1-Базой назовем минимальный набор U таких попарно несмежных вершин, что любая вершина орграфа G или принадлежит U , или смежна из некоторой вершины U .

Введем также понятие $1m$ -базы. Множество U вершин орграфа назовем $1m$ -базой, если и только если выполняется следующее:

- (i) все вершины из множества $\forall U$ попарно несмежны;
- (ii) все вершины, имеющие нулевые полустепени заходов, лежат в U ;
- (iii) любая вершина из U , имеющая ненулевую полустепень захода, смежна хотя бы из одной вершины, содержащейся в $\forall U$.

Орграф назовем *циркулярным*, если и только если все его вершины имеют одинаковые полустепени исходов и существует такое расположение вершин на окружности, что все достижимые из u вершины располагаются непосредственно после нее, считая по ходу часовой стрелки.

Пусть оргграф $G_{n,k}$ является циркулярным и d обозначает наибольший общий делитель n и k . Тогда следующее предложение дает полное описание 1-баз и 1*m*-баз $G_{n,k}$.

Предложение. Если $d = k$, то у $G_{n,k}$ существует ровно до k 1-баз и 1*m*-баз. 1-Базами являются множества вершин $U_i = \{u_{i+dj}, j = 0, 1, \dots, n/d - 1\}$, 1*m*-базами являются множества $V \setminus U_i = \{u_{i+dj}, j \neq 0, 1, \dots, n/d - 1\}$, $i = 1, 2, \dots, k$. При $d \neq k$ не существует ни одной 1 – (1*m*)-базы.

Построим на графе $G_{n,k}$ ГГС классов 1–4, модели которых обозначим для определенности через $M_i(n, k)$, где $i = 1, 2, 3, 4$ обозначает номер класса. Тогда критерий количества стационаров может быть сформулирован в следующей эквивалентной форме (n, k) -критерия.

(n, k) -Критерий количества стационаров. Если $d = k$, то для любого заранее фиксированного τ существует такое $\gamma_0(\tau)$, что для любого $\gamma > \gamma_0$ существует такое $\alpha_0(\gamma, \tau)$, что $M_i(n, k)$, $i = 1, 2, 3, 4$, для любого $\alpha > \alpha_0$ имеет ровно k устойчивых стационаров.

Следующая формулировка позволяет подсчитать количество колебательных режимов у кГГС классов 1, 2, 4, построенных на графе $G_{n,k}$.

(n, k) -Критерий количества циклов. Если $d \neq k$, то для любого заранее фиксированного τ существует такое $\gamma_0(\tau)$, что для любого $\gamma > \gamma_0$ существует такое $\alpha_0(\gamma, \tau)$ что $M_i(n, k)$, $i = 1, 2, 4$, для любого $\alpha > \alpha_0$ имеет d устойчивых предельных циклов.

Для кГГС класса 3 приведенный выше (n, k) -критерий количества циклов не выполняется. Например, в модели $M_3(6, 4)$ при “достаточно больших α и γ ” численно обнаруживается только один устойчивый предельный цикл (расчеты не приводятся). Численные исследования позволяют предположить, что для кГГС класса 3 выполняется (n, k) -критерий в следующей формулировке.

(n, k) -Критерий. Если $d \neq k$, то для любого заранее фиксированного τ существует такое $\gamma_0(\tau)$, что для любого $\gamma > \gamma_0$ существует такое $\alpha_0(\gamma, \tau)$, что для любого $\alpha > \alpha_0$, $M_3(n, k)$ имеет только один устойчивый предельный цикл.

5. ЗАКЛЮЧЕНИЕ

Переход к анализу поведения ГС, в которых регуляция осуществляется механизмами, являющимися комбинациями рассмотренных типов, многократно усложняет проблему поиска подходящих характеристик структурных графов, по которым можно предсказывать стационары и колебательные режимы функционирования.

Сложность проблемы иллюстрируют данные исследований класса булевых (пороговых) моделей ГС (см. [45]), которые являются предельным случаем моделей (2.1). По оценкам авторов, только из пяти генетических элементов можно сконструировать больше 10^{20} ГГС, имеющих различные типы предельных режимов функционирования [38].

Но несмотря на то, что представленные результаты описывают свойства ограниченного многообразия генетических объектов, они закладывают определенную теоретическую базу для решения некоторых задач конструирования ГС. Например, полученные результаты свидетельствуют о принципиальной возможности построения ГС, имеющей любое заданное количество стационаров или предельных циклов. Следующий простой алгоритм, который мы приведем без обоснования, решает задачу построения ГГС с любым наперед заданным количеством стационаров и циклов.

Допустим, необходимо построить ГС, имеющую k стационаров и d циклов. Возьмем оргграфы $G_{n_1k,k}$ и G_{n_2d,n_3d} , где $n_1, n_2, n_3 > 0$, n_2 и n_3 являются взаимно простыми числами. Тогда кГГС класса 1, построенная на их сильном объединении, обладает искомыми свойствами. Напомним, что необходимым условием существования у такой ГС стационарных и/или циклических режимов, прогнозируемых из анализа свойств структурного графа, является наличие определенной нелинейности процессов регуляризации активности генетических элементов ГС (высокое значение параметра γ) и достаточно высокой эффективности экспрессии генов в отсутствие репрессоров (высокое значение параметра α). При конструировании реальных генных сетей необходимая нелинейность должна обеспечиваться подходящей степенью мультимеризации белков-репрессоров, а эффективность экспрессии – использованием “сильных” промоторов и/или эффективных сайтов инициации трансляции [45], [46].

Как упоминалось выше, в ГС синтез РНК и белков включает большое количество (сотни и тысячи) промежуточных стадий. Однако для частного случая доказано, что систему уравнений очень большой размерности можно заменить одним уравнением с запаздывающим аргументом [35]. Полученный результат показывает, что если процесс синтеза имеет достаточно большое количество промежуточных линейных стадий и скорость протекания каждой такой стадии до-

статочна велика, то кинетика наработки конечного продукта практически не зависит от кинетики внутренних стадий синтеза. Все определяется механизмом регуляризации запуска первой стадии синтеза и величиной запаздывания, которая численно равняется среднему суммарному времени протекания всех промежуточных стадий. Таким образом, полученный в работе результат устанавливает отношение между микро- и макроуровнями системы, если под микроуровнем рассматривать отдельные (микро)стадии синтеза, а под макроуровнем – суммарный синтез конечного продукта. Это отношение может выразить следующее

Утверждение. *Отдельная микростадия синтеза тем меньше влияет на кинетику наработки конечного продукта, чем меньшую долю времени она занимает во всей совокупности последовательно протекающих микропроцессов.*

В пределе на макроуровне проявляется только одна характеристика микроуровня – суммарное время протекания процесса синтеза.

В связи с приведенной биохимической интерпретацией полученного результата возникает естественный вопрос о критичности условий линейности и обратимости промежуточных стадий для выполнения доказанной предельной теоремы. Действительно, в реальных биологических системах отдельные стадии синтеза ДНК, РНК и белков являются линейными и необратимыми только в первом приближении. В целом же они оказываются нелинейными, так как являются совокупностями биохимическим реакций. По этой же причине стадии синтеза все в большей мере теряют необратимость по мере того, как мы рассматриваем все более мелкие подстадии, постепенно приближаясь к уровню протекания элементарных биохимических событий на уровне взаимодействия атомов. Поэтому изучение предельных переходов в системах типа описывающих многостадийный синтез вещества при различных механизмах протекания промежуточных стадий синтеза подводит точную теоретическую базу под интуитивное понимание того, что при определенных условиях для адекватного моделирования процессов на макроуровне не требуется полного знания механизмов функционирования системы на ее микроуровнях. Этот результат представляется важным этапом в построении теории генных сетей.

Представленные в работе результаты могут найти применение в областях биотехнологии и фармакогенетики, в том числе при конструировании ГС с заранее заданными свойствами, решении задач поиска оптимальных стратегий управления ГС, конструировании “интеллектуальных” генетических систем, способных обеспечивать в организме максимально сбалансированную коррекцию параметров его функционирования (нарушенных, например, вследствие наличия некоего генетического дефекта или болезни) с целью приведения их к индивидуальной норме.

СПИСОК ЛИТЕРАТУРЫ

1. Колчанов Н.А., Ананько Е.А., Колпаков Ф.А. и др. Генные сети // Мол. биология. 2000. Т. 34. № 4. С. 533–544.
2. Fellenberg K., Hauser N.C., Brors B. et al. Correspondence analysis applied to microarray data // Proc. Nat. Acad. Sci. USA. 2001. V. 98/ P. 10781–10786.
3. Sherlock G., Hernandez-Boussard T., Kasarskis A. et al. The Stanford microarray database // Nucl. Acids Res. 2001. V. 29. P. 152–155.
4. Ananko E.A., Podkolodny N.L., Stepanenko I.L. et al. GeneNet: a database on structure and functional organization of gene networks // Nucl. Acids Res. 2002. V. 30. № 1. P. 398–401.
5. Kolchanov N.A., Ananko E.A., Likhoshvai V.A. et al. Gene networks description and modeling in the GeneNet system // Gene Regulation and Metabolism. Cambridge: MIT Press, 2002. Ch. 7. P. 149–180.
6. Патнер В.А. Генетические управляющие системы. Новосибирск: Наука, 1966.
7. Galimzyanov A.V., Tchuraev R.N. An investigation of the structural stability of drosophila control gene subnetwork in computer experiments // Bioinformatics of Genome Regulation and Structure. Boston etc.: Kluwer Acad. Publs, 2004. P. 301–310.
8. Kauffman S.A. Metabolic stability and epigenesis in randomly constructed genetic net // J. Theor. Biol. 1969. V. 22. P. 437–467.
9. Thomas R. Boolean formalization of genetic control circuits // J. Theor. Biol. 1973. V. 42. P. 563–585.
10. Thomas R. Logical analysis of systems comprising feedback loops // J. Theor. Biol. 1978. V. 73. № 4. P. 631–656.
11. Savageau M. A theory of alternative designs for biochemical control systems // Biomed. Biochim. Acta. 1985. V. 44. P. 875–880.
12. Mjolsness E., Sharp D.H., Reinitz J. A connectionist model of development // J. Theor. Biol. 1991. V. 152. № 4. P. 429–454.

13. McAdams H., Arkin A. Stochastic mechanisms in gene expression // Proc. Nat. Acad. Sci. USA. 1997. V. 94. P. 814–819.
14. Marnellos G., Mjolsness E. A gene network approach to modeling early neurogenesis in Drosophila // Pacific Symposium on Biocomputing. Maui, Hawaii, USA: World Scient., 1998. P. 30–41.
15. Savageau M.A. Design principles for elementary gene circuits: Elements, methods, and examples // CHAOS. 2001. V. 11. № 1. P. 142–159.
16. Hofstadt R., Meineke F. Interactive modeling and simulation of biochemical networks // Comput. Biol. Med. 1995. V. 25. P. 321–334.
17. Freier A., Lange M., Hofstaedt R. Integrative analysis of gene networks using dynamic process pattern modeling // Bioinformatics of Genome Regulation and Structure. Boston etc.: Kluwer Acad. Publs. 2004. P. 257–264.
18. Likhoshvai V.A., Matushkin Yu.G., Vatin Yu.N., Bazhan S.I. A generalized chemical kinetic method for simulating complex biological systems. A computer model of λ phage ontogenesis // Comput. Technol. 2000. V. 5. № 2. P. 87–99.
19. Лихошвай В.А., Матушкин Ю.Г., Ратушный А.В. и др. Обобщенный химико-кинетический метод моделирования генных сетей // Мол. биология. 2001. Т. 35. № 6. С. 1972–1980.
20. Ватолин Ю.Н., Вазхан С.И., Лихошвай В.А. Математическое моделирование внутриклеточного развития фагов // I Отраслевая конф. по применению матем. моделирования и АСУ в науч. учреждениях и на предприятиях (1–3 марта 1978 г.). Новосибирск, 1978. С. 22–32.
21. Кравченко В.В., Шамин В.В., Гилева И.П. и др. Механизм регуляции трансляции полицистронных мРНК. Роль взаимного расположения цистронов // Докл. АН СССР. 1987. Т. 295. № 3. С. 745–748.
22. Бачинский А.Г., Лихошвай В.А. Влияние дефектных интерферирующих частиц на репродукцию вируса венозного энцефаломиелита лошадей в клеточных культурах. Кольцово: ВНИИМБ, 1988.
23. Кравченко В.В., Шамин В.В., Гилева И.П. и др. Влияние альтернативных вторичных структур на эффективность трансляции полицистронной мРНК // Докл. АН СССР. 1988. Т. 301. № 2. С. 480–484.
24. Вазхан С.И., Лихошвай В.А. Математическая модель регуляции внутриклеточного онтогенеза вируса гриппа // Тезисы докл. Всес. конф. “Матем. моделирование системы иммунитета и инфекционного процесса”. Кольцово, 26–28 июля 1989. Новосибирск, 1989. С. 9–10.
25. Likhoshvai V.A. Rare codons: fortunity or regularity? // Modeling and Comput. Meth. in Molecular Biol. and Genetics. New York: Nova Sci. Publs, 1992. P. 463–469.
26. Bazhan S.I., Likhoshvai V.A., Belova O.E. Theoretical analysis of the regulation of interferon expression during priming and blocking // J. Theor. Biol. 1995. V. 175. P. 149–160.
27. Bazhan S.I., Likhoshvai V.A., Belova O.E. Theoretical analysis of the regulation of interferon expression during priming and blocking // J. Theor. Biol. 1995. V. 175. P. 149–160.
28. Николенко Г.Н., Кравченко В.В., Сваровская Е.С. и др. Регуляция трансляции дистального гена LacZ полицистронной мРНК потоком рибосом с проксимального гена // Биоорган. химия. 1997. Т. 23. № 3. С. 200–204.
29. Likhoshvai V.A., Matushkin Yu.G. Computer model for analysis of evolutionary drift of synonymous codons along mRNA // Comput. Technol. 2000. V. 5. № 2. P. 57–63.
30. Ratusny A.V., Podkolodnaya O.A., Ananko E.A., Likhoshvai V.A. Mathematical model of erythroid cell differentiation regulation // Proc. I Internat. Conf. on Bioinformatics or Genome Regulation and Structure. Novosibirsk, 2000. V. 1. P. 203–206.
31. Nedosekina E.A., Ananko E.A., Likhoshvai V.A. Construction of mathematical model of the gene network on macrophage activation under the action of IFN- γ and LPS // Proc. II Internat. Conf. on Bioinformatics of Genome Regulation and Structure (BGRS’20002). Novosibirsk, 2002. V. 2. P. 74–76.
32. Ратушный А.В., Лихошвай В.А., Игнатьева Е.В. и др. Компьютерная модель генной сети регуляции биосинтеза холестерина в клетке: анализ влияния мутаций // Докл. РАН. 2003. Т. 389. № 2. С. 90–93.
33. Latypov A.F., Nikulichev Yu.V., Likhoshvai V.A. et al. Problems of control of gene networks in a space of stable states // Proc. IFAC Workshop “Modelling and Analys. of Logic Controlled Dynamic Systems”. Irkutsk, 2003. P. 251–266.
34. Фадеев С.И., Ермакова А., Гудков А.В. Применение метода продолжения по параметру для расчета равновесных состояний, определяемых минимизацией свободной энергией Гиббса: Препринт № 7. Новосибирск: Ин-т матем. СО РАН, 1955. 33 с.
35. Лихошвай В.А., Фадеев С.И., Демиденко Г.В., Матушкин Ю.Г. Моделирование многостадийного синтеза вещества без ветвления уравнением с запаздывающим аргументом // Сибирский. индустр. матем. 2004. Т. 7. № 1(17). С. 73–94.
36. Thomas R., Thieffry D., Kaufman M. Dynamical behavior of biological regulatory networks. II. Biological role of feedback loops and practical use of the concept of the loop-characteristic state // Bull. Math. Biology. 1995. V. 57. P. 247–276.
37. Лихошвай В.А., Матушкин Ю.Г., Фадеев С.И. О связи графа генной сети с качественными режимами ее функционирования // Мол. биология. 2001. Т. 35. № 6. С. 1080–1087.

38. Лихошвай В.А., Матушкин Ю.Г., Фадеев С.И. Задачи теории функционирования генных сетей // Сибирский индустр. матем. 2003. Т. 6. № 2(14). С. 64–80.
39. Лихошвай В.А., Фадеев С.И. О гипотетических генных сетях // Сибирский ж. индустр. матем. 2003. Т. 6. № 3(15). С. 134–153.
40. Johnson A.D., Poteete A.R., Lauer G. et al. Lambda repressor and cro-components of an efficient molecular switch // Nature. 1981. V. 294³5838). P. 217–233.
41. Van Gemen B., Twisk J., Van Knippenberg P.H. Autogenous regulation of the escherichia coli ksgA gene at the level of translation // J. Bacteriology. 1989. V. 171. № 7. P. 4002–4008.
42. Ptashne M. A genetic switch: Phage lambda and higher organisms. Oxford: Blackwell Sci., 1992.
43. Gardner T.S., Cantor C.R., Collins J.J. Construction of a genetic toggle switch in Escherichia coli // Nature. 2000. V. 403. P. 339–342.
44. Харари Ф. Теория графов. М.: Мир, 1973.
45. Elowitz M.B., Leibler S. A synthetic oscillatory network of transcriptional regulators // Nature. 2000. V. 403. P. 335–338.
46. Edwards R., Glass L. Combinatorial explosion in model gene networks // CHAOS. 2000. V. 10. P. 691–704.