

ЗАДАЧИ ТЕОРИИ ФУНКЦИОНИРОВАНИЯ ГЕННЫХ СЕТЕЙ

В.А.Лихошвай* , Ю.Г.Матушкин, С.И.Фадеев¹

Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, 630090.

¹*Институт математики Сибирского отделения Российской академии наук, Новосибирск, 630090.*

- *контактный автор likho@bionet.nsc.ru*

Разработка математической теории функционирования генных сетей является актуальной проблемой информационной биологии. В работе вводятся в рассмотрение гипотетические генные сети (ГГС) – теоретические объекты, конструируемые из элементарных единиц двух типов: матричных элементов и регуляторных механизмов. Их функционирование описывается автономными системами из n уравнений и связанными с ними уравнениями специального вида с запаздывающими аргументами.

В работе приводятся первые практические результаты по исследованию ГГС. Приводятся примеры ГГС, имеющих множество стационарных решений, предельные циклы и аттракторы сложной природы. В соответствии с типами регуляторных связей в рассмотрение вводятся 4 класса ГГС, а также определяются в них канонические представители. Для канонических симметричных ГГС, формулируется (n,k) -критерий, позволяющий предсказывать в определенной области значений параметров предельные свойства решений соответствующих моделей, т.е. предсказывать число и характер имеющихся у них устойчивых предельных состояний. В случае ГГС первого и второго классов формулируется гипотеза об однозначном соответствии устойчивых состояний и 1-баз структурных орграфов этих сетей, а в случае ГГС третьего и четвертого классов - гипотеза об однозначном соответствии устойчивых состояний и $1m$ -баз структурных орграфов.

Вместе с этим авторы хотели бы привлечь внимание к проблемам качественного и численного анализа рассматриваемого круга проблем.

Ключевые слова: гипотетическая генная сеть, математическая модель, компьютерная модель, регуляция, отрицательная связь, положительная связь, стационарные решения, предельные циклы, устойчивость, особые точки, теория генных сетей.

1. Введение

Живые системы являются системами с чрезвычайно сложным внутренним строением. Уникальным свойством таких систем является

способность к самовоспроизведению и к эффективному существованию в изменяющихся условиях внешней среды. Исключительно важная роль в функционировании живых систем принадлежит генным сетям, которые обеспечивают выполнение жизненно важных функций организмов [1,2]. Генные сети контролируют все основные процессы, протекающие в живых системах:

а) поддерживают в организме устойчивые состояния, характеризующиеся постоянством концентраций веществ (гомеостаз, стационарные состояния);

б) обеспечивают периодические незатухающие колебания концентраций определенных групп веществ (осцилляции, циклы);

в) контролируют необратимые процессы: развитие, рост, дифференцировка, апоптоз.

Обычно генные сети состоят из десятков и сотен элементов, которые объединяются между собой посредством сложных нелинейных процессов. Ядро генных сетей представляют группы координировано функционирующих генов. Кроме генов в них входят

1. МРНК и белки, кодируемые этими генами;

2. пути передачи сигналов от клеточной мембраны к ядру клетки, обеспечивающие активацию генов в ответ на внешние регуляторные стимулы;

3. внешние сигналы, гормоны, метаболиты, которые переносят физиологические регуляторные воздействия, обеспечивая изменение состояний генной сети.

В функционировании генных сетей важнейшую, если не определяющую роль, играют отрицательные и положительные регуляторные связи. Именно благодаря их наличию генные сети приобретают способность к саморегуляции и адекватному реагированию на изменение внешних условий.

Несмотря на то, что к настоящему времени накоплен огромный фактический материал о структурных и динамических особенностях функционирования десятков реальных генных сетей, мы пока не имеем сколько-нибудь развитой теории их функционирования, хотя попытки такие предпринимались [3]. Развитие теории актуально как с точки зрения углубления понимания закономерностей функционирования генных сетей, так и с точки зрения выработки эффективных стратегий направленного воздействия на них, в том числе с целью корректировки функции генных сетей, а также для конструирования генных сетей с заранее заданными свойствами. Для успешного решения этих задач необходимо создание эффективных компьютерных методов описания, реконструкции и моделирования генных сетей.

2. Описание гипотетических генных сетей

Природные генные сети состоят из конечного числа элементов двух основных типов. Первый тип представляют вещества - мРНК, белки, их различные формы и промежуточные комплексы, низкомолекулярные соединения и т.д. Ко второму типу относятся процессы, в которых участвуют определенные группы веществ. Процессы связывают вещества между собой в единое функциональное целое – генную сеть. Функционирование генной сети во времени характеризуется изменением концентраций веществ элементов генной сети. Основу функционирования генных сетей составляют биохимические процессы, процессы пассивного и активного переноса веществ и энергии. В том случае, когда среда функционирования генной сети может быть разделена на области, в которых можно предполагать мгновенное перемешивание (однородность пространства), моделирование возможно проводить в терминах обыкновенных дифференциальных уравнений вида

$$\frac{dx_i}{dt} = F_i(X, K) - x_i G_i(X, K), \quad i = 1, 2, \dots, n. \quad (1)$$

Здесь n – количество динамических переменных модели генной сети, $X=(x_1, \dots, x_n)$ – вектор динамических переменных, $K=(k_1, \dots, k_m)$ – вектор параметров, $F_i(X, K)$ и $G_i(X, K)$ – рациональные функции, принимающие неотрицательные значения при всех неотрицательных значениях переменных и параметров, такие что уравнение $F_i(X, K) = x_i G_i(X, K)$ имеет неотрицательное решение относительно x_i при любых заданных значениях параметров $K=(k_1, \dots, k_m)$ и $X_i=(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)$. Таким образом, изучение динамики функционирования генных сетей можно рассматривать как задачу изучения свойств систем вида (1).

Как уже отмечалось, ключевую роль в функционировании генных сетей играют отрицательные и положительные регуляторные связи. Для их изучения введем гипотетические генные сети (ГГС), которые состоят из элементов двух типов: элементов матричного синтеза (коротко, m -элементов) и регуляторных механизмов. Под m -элементом мы будем понимать упорядоченную пару, из которых первый элемент будем называть порождающим, а второй – порождаемым или продуктом. В данных объектах перенос информации и вещества осуществляется в направлении от порождающего элемента к продукту. Физическими прообразами m -элементов являются пары (ген, белок), (ген, РНК), (мРНК, белок) и т.д. В них синтез продукта осуществляется под контролем порождающего элемента, при этом сам порождающий элемент не расходуется в процессе синтеза продукта. Для простоты мы не будем вводить в понятие m -элементов пары, в которых перенос информации осуществляется в обратном направлении: РНК \Rightarrow ДНК, Белок \Rightarrow РНК, хотя формальных трудностей это не составляет. Более того, в дальнейшем мы ограничимся рассмотрением только элементов, прообразами которых являются пары (ген, белок). Такие m -элементы назовем

генетическими (ГмЭ). Делается это единственно с целью облегчения дальнейшего изложения. Все излагаемое ниже может быть легко обобщено и на m -элементы другой природы.

В генетических элементах продуктами являются белки, а порождающими - гены. Обозначать ГмЭ будем в виде упорядоченной пары (g, p) , где g обозначает ген, а p - белок. Два генетических элемента будем считать идентичными, если они имеют одинаковые наименования генов и белков.

Под функционированием ГмЭ будем понимать синтез белка, кодируемого этим элементом. Под активностью функционирования ГмЭ будем понимать скорость синтеза его белка.

Элементами второго типа являются *регуляторные механизмы* (РМ). Под РМ мы будем понимать процесс, посредством которого некоторая совокупность генетических элементов осуществляет регуляцию активности некоторого фиксированного генетического элемента. Вхождение конкретного ГмЭ (g_1, p_1) в РМ другого конкретного ГмЭ (g_2, p_2) будем называть регуляторной связью (РС). В этом случае ГмЭ (g_1, p_1) будем называть регулятором (g_2, p_2) . Если (g_1, p_1) идентичен (g_2, p_2) , то регуляторную связь будем называть авторегуляторной.

Гипотетической геной сетью будем называть конечную совокупность генетических элементов, которые связаны в единую функциональную систему посредством заданных *регуляторных механизмов*.

3. Системы, описывающие функционирование ГГС

3.1. Описание конститутивного синтеза белка в отсутствие регуляции.

Рассмотрим сначала ГмЭ (g, p) , который имеет некоторый конститутивный уровень функционирования и не подвержен регуляции другими генетическими элементами. Его функционирование выражается в синтезе белка p . А эффективность функционирования – в скорости синтеза белка p . Хорошо известно, что синтез белка в живых клетках складывается из значительного количества последовательно протекающих стадий. Общее количество промежуточных стадий может быть достаточно большим. Для разных генов оно является различным, так как они кодируют РНК и белки различной длины. Для гена, с которого синтезируется РНК длиной 1000 нуклеотидов и белок длиной 300 аминокислот минимальное количество промежуточных стадий, только этапов элонгации транскрипции и трансляции, будет составлять 1900, так что реальное количество их значительно больше.

Будем считать, что количество промежуточных стадий равно n_g . Обозначим каждую промежуточную стадию через $x_i, i = 1, 2, \dots, n_g$. Конститутивный синтез белка p (переменная x) будем описывать системой дифференциальных уравнений:

$$\begin{aligned} \frac{dx_1}{dt} &= \alpha - k_1 x_1, \\ \frac{dx_i}{dt} &= k_{i-1} x_{i-1} - k_i x_i, \quad i = 2, 3, \dots, n_g, \\ \frac{dx}{dt} &= k_{n_g} x_{n_g} - \beta x, \end{aligned} \quad (2)$$

где α - константа скорости конститутивного синтеза, k_i - константы скорости перехода из i -го промежуточного состояния в $(i+1)$ -ое промежуточное состояние, β - константа деградации белка p . Как видим, промежуточные стадии (переменные x_i) синтеза белка в отсутствие регуляции мы описываем линейными уравнениями. В реальности (в живых системах) промежуточные стадии протекают по более сложным нелинейным механизмам. Однако наш опыт моделирования показывает, что при некоторых естественных ограничениях, которые обычно выполняются для биохимических систем, они могут быть хорошо приближены линейными системами (2). (Более подробно об этом говорится в разделе 4.2.)

3.2. Описание регуляции генетического элемента

Будем полагать, что в ГГС регуляция активности генетического элемента направлена на изменение (в сторону уменьшения или увеличения) активности функционирования его порождающего элемента. Будем полагать, что генетические элементы-регуляторы осуществляют функцию регулирования посредством кодируемых ими белков. На формальном уровне наличие регуляции выражается в том, что в системе (2) параметр α становится функцией, зависящей от концентраций белков, кодируемых генетическими элементами-регуляторами данного генетического элемента.

Описывать регуляцию активности ГмЭ будем следующим образом. Считаем, что порождающий элемент g может пребывать в нескольких альтернативных состояниях. Обозначим эти состояния через $u_i, i=1, \dots, n_u$. Считаем, что нам известна схема перехода состояний друг в друга. Эту схему будем задавать ориентированным графом, в котором вершины представляют альтернативные состояния, а дуги - переходы между ними. Каждый переход осуществляется с некоторой скоростью $v_{i,j}$, которая зависит от концентраций ряда белков-регуляторов, что выражается в конкретном виде зависимости $v_{i,j}$ от концентраций белков-регуляторов. По смыслу $v_{i,j} \geq 0$ при любом наборе неотрицательных переменных. Припишем каждой существующей дуге (u_i, u_j)

в качестве веса выражение $v_{i,j}$, а каждой i -й вершине припишем переменную u_i , которая будет иметь смысл вероятности пребывания регуляторного участка в i -м состоянии. Составим следующую алгебраическую систему из n_u уравнений на неизвестные u_i :

$$\sum_{u_j \in D_i^-} v_{i,j} u_j = \sum_{u_j \in D_i^+} v_{j,i}, \quad i = 1, 2, \dots, n_u. \quad (3)$$

Здесь D_i^- - множество вершин, из которых дуги входят в i -ю вершину, D_i^+ - множество вершин, в которые дуги входят из i -ой вершины. Так как сумма по всем вероятностям равняется 1, то любое уравнение линейной системы (3) может быть заменено на уравнение

$$\sum_{i=1}^{n_u} u_i = 1. \quad (4)$$

Исходя из биохимической природы процессов, протекающих в живых клетках, для задания $v_{i,j}$ мы ограничимся использованием полиномиальных выражений от концентраций продуктов генетических элементов-регуляторов:

$$v_{i,j} = \sum k_{i_1 \dots i_m} \prod x_{r_s}^{\gamma_{r_s}}, \quad (5)$$

где все x_{r_s} имеют смысл концентраций продуктов, степени γ_{r_s} задают уровень нелинейности вхождения регуляторного продукта в соответствующий член полинома, $k_{i_1 \dots i_m}$ имеют смысл констант скоростей протекания соответствующих процессов и являются неотрицательными числами. Суммирование в (5) ведется по всем индексам.

Система (3) с уравнением баланса (4) является линейной относительно неизвестных. Она всегда имеет единственное решение (u_1, \dots, u_{n_u}) , $u_i \geq 0$, которое при необходимости может быть выписано в аналитическом виде.

Предположим теперь, что некоторые состояния u_i являются активными формами порождающего элемента. Тогда его суммарную активность будем выражать в виде взвешенной суммы этих состояний, где вес k_i каждого активного состояния задает его относительный вклад в общую активность регуляторного участка:

$$A = \sum_{i=1}^{n_u} k_i u_i \quad (6)$$

Выражение A будем называть моделью регуляции генетического элемента. Заменим в системе (2) α на A . В результате мы приходим к системе, которая описывает *функционирование регулируемого генетического элемента*.

4. Построение моделей функционирования ГГС

4.1. Описание функционирования ГГС автономными системами.

Возьмем теперь некоторое количество ГмЭ (g_j, p_i) , $j=1, \dots, m$, $i=1, \dots, n$. Очевидно, что по смыслу $n \leq m$. Так как генетический элемент полностью определяется порождающим элементом g_j , то договоримся, что ГмЭ имеет тот же номер j . Считаем, что для каждого g_j задан механизм A_j и совокупность белков, участвующих в регуляции активности функционирования генетического элемента (g_j, p_i) является подмножеством всех белков, синтезируемых рассматриваемыми генетическими элементами. Чтобы записать систему дифференциальных уравнений, описывающую функционирование ГГС, состоящей из выбранной совокупности генетических элементов и заданных регуляторных механизмов, для каждого генетического элемента построим системы (2) с учетом соответствующего регуляторного механизма A_j .

Теперь договоримся, что если два генетических элемента кодируют один белок, то последние переменные соответствующих систем (переменные x в системе (2)) являются идентичными. Для всех таких переменных составим балансные уравнения вида:

“скорость наработки белка x ” = “сумма скоростей наработки белка x по всем генетическим элементам, кодирующим белок x ” – “скорость распада белка x ”.

К этим уравнениям добавим все уравнения из систем регуляции всех генетических элементов, исключая из них все последние уравнения. В результате приходим к системе обыкновенных дифференциальных уравнений следующего вида:

$$\begin{cases} \frac{dx_{j,i,1}}{dt} = A_j(x_r | r \in R_j) - k_{j,i,1}^{(n_{j,i})} x_{j,i,1}, & j = \overline{1, m}, \quad i \in P_j, \\ \frac{dx_{j,i,s}}{dt} = k_{j,r,s-1}^{(n_{j,i})} x_{j,i,s-1} - k_{j,i,s}^{(n_{j,i})} x_{j,i,s}, & j = \overline{1, m}, \quad i \in P_j, \quad s = \overline{1, n_{j,i}}, \\ \frac{dx_i}{dt} = \sum_{j \in G_i} k_{j,i,n_{j,i}}^{(n_{j,i})} x_{j,i,n_{j,i}} - \beta_i x_i, & i = \overline{1, n} \end{cases} \quad (7)$$

которую будем называть моделью ГГС. Здесь n – количество различных белков, m – количество различных порождающих элементов, $n_{j,i}$ – количество промежуточных стадий в синтезе i -го белка j -м порождающим элементом, $k_{j,i,s}^{(n_{j,i})}$ – константа скорости перехода s -й промежуточной стадии в $(s+1)$ -ю промежуточную стадию при синтезе i -го белка j -м порождающим элементом, β_i – константа распада i -го белка, P_j – множество номеров белков, которые кодируются j -м порождающим элементом, R_j – множество номеров белков, которые регулируют активность функционирования j -го порождающего

элемента, G_i - множество номеров порождающих элементов, которые кодируют i -й белок. Класс этих систем будем обозначать через МГГС.

4.2. Описание функционирования ГГС системами дифференциальных уравнений с запаздывающими аргументами

Как упоминалось выше, в реальных генных сетях синтез белков складывается из большого количества промежуточных стадий, суммарное время протекания которых ограничено, а каждая промежуточная стадия протекает достаточно быстро. Для описания промежуточных стадий мы используем линейные уравнения вида (2), а для описания механизмов регулирования – нелинейные выражения (6). Последние, с учетом (5), представляют рациональные выражения от концентраций соответствующих белков-регуляторов. Также легко видеть, что значительное количество уравнений (7) являются дифференциальными уравнениями с простейшей по виду линейной правой частью. Благодаря их наличию системы из МГГС, как правило, имеют очень большую размерность. Понятно, что большая размерность существенно затрудняет проведение анализа таких систем. Возникает вопрос, а нельзя ли как-нибудь понизить b [размерность? Оказывается, стандартный вид линейных уравнений, входящих в системы (7), действительно позволяет сделать важный шаг по пути упрощения их и перейти к новому классу моделей, в которых число переменных в точности равно числу различных белков.

Наряду с системой (7) из МГГС будем рассматривать в качестве модели ее функционирования систему уравнений с запаздывающими аргументами следующего вида:

$$\frac{dx_i}{dt} = \sum_{j \in G_i} A_j(x_r(t - \tau_{j,i}) | r \in R_j) - \beta_i x_i, i = 1, 2, \dots, n. \quad (8)$$

Здесь все параметры и обозначения, кроме $\tau_{j,i}$, имеют тот же смысл, что и для (7), а

$$\tau_{j,i} = \sum_{s=1}^{n_{j,i}} 1/k_{j,i,s}^{(n_{j,i})}$$

имеет смысл запаздывающего аргумента, равного времени протекания всех промежуточных стадий синтеза i -го белка с j -го порождающего элемента.

Отметим, что системы (8) являются в некотором роде минимальными системами, которые удовлетворяют требованию адекватности описания функционирования ГГС. Действительно, нижняя граница сложности вида такой системы вытекает из необходимости учета в ней всех регуляторных механизмов. Поэтому в системах, описывающих функционирование ГГС, должны присутствовать динамические переменные, имеющие смысл концентраций белков-регуляторов и механизмы регулирования их скоростей

синтеза. С другой стороны, в таких системах должны быть минимизированы описания второстепенных деталей функционирования ГГС. Системы (8) полностью удовлетворяют перечисленным условиям, так как содержат в качестве динамических переменных только концентрации белков-регуляторов и не содержат никаких иных деталей описания кроме механизмов и запаздываний. Но запаздывания являются важным элементом описания функционирования ГГС, поскольку изменение величины запаздывания может приводить к кардинальному изменению поведения ГГС в целом.

В качестве примера рассмотрим уравнение

$$\frac{dx}{dt} = \frac{\alpha}{1 + x^\gamma(t - \tau)} - x, \quad (9)$$

которое описывает автоингибирование генетического элемента. Очевидно, при любых неотрицательных значениях α и γ , как при $\tau=0$, так и при $\tau > 0$, уравнение (9) имеет только одно стационарное решение, которое находится из уравнения

$$\alpha = (1 + x^\gamma)x. \quad (10)$$

Однако, при других значениях уравнение (9) может иметь в качестве устойчивого режима функционирования предельный цикл. Такая ситуация, например, реализуется, если при $\alpha=5$ и $\gamma=3$ мы начнем постепенно увеличивать значение τ . Тогда, начиная с критического значения $\tau^* = 0.9643$, стационарное решение теряет устойчивость, что приводит к возникновению автоколебания. Можно показать, что критические значения τ^* для уравнения (9) вычисляются по формуле:

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

Обозначим класс систем (8) через МГГСЗ. Между системами из МГГС и МГГСЗ существует тесная взаимосвязь. Она выражается в том, что системы из МГГСЗ являются предельными случаями систем из МГГС. Исходя из опыта численного анализа систем из МГГС и МГГСЗ мы высказываем следующие утверждение, требующее доказательства.

Рассмотрим систему (7) из МГГС, для которой в нулевой момент заданы начальные условия:

$$x_i(0) = 0, \quad x_{j,i,1}(0) = C_{j,i}, \quad x_{j,i,s}(0) = 0, \quad i = \overline{1, n}, \quad j = \overline{1, m}, \quad s = \overline{2, n_{j,i}}. \quad (11)$$

Обозначим через $x_i^{(n_{1,i}, \dots, n_{m,i})}(t)$ - концентрации i -х белков в момент времени t , найденные как решение задачи Коши (7), (11) при фиксированных $n_{j,i}$ и $k_{j,i,s}^{(n_{j,i})}$. Здесь вектор $(n_{1,i}, \dots, n_{m,i})$ вынесен в верхний индекс i -го белка, чтобы подчеркнуть, что решение найдено при фиксированных по всем генетическим элементам количествах промежуточных стадий синтеза белков.

Возьмем теперь систему (8) из МГГСЗ, которая соответствует системе (7) из МГГС. Пусть в ней запаздывания τ , рассмотренные в совокупности по всем генетическим элементам, имеют k различных значений. Обозначим их через $t_r, r=1, \dots, k$. Не ограничивая общности, мы можем считать, что $t_1 < \dots < t_k$. Обозначим через N_r множество генетических элементов (g_j, p_i) , имеющих запаздывание t_r . По определению для любой пары (j, i) из (8) полагаем

$$A_j(x_s(t - \tau_{j,i}) | s \in R_j) \equiv 0, \quad x_i(t - \tau_{j,i}) = 0 \text{ при } t - \tau_{j,i} < 0, \quad x_i(0) = 0.$$

Через $x_1(t), \dots, x_n(t)$ обозначим функции, которые получаются следующим образом. Если $t_1 > 0$, то полагаем $x_i(t) = 0$ для всех $0 \leq t < t_1, i=1, \dots, n$. Затем переходим к последовательному решению задач (7), (12) 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 (7), (12) 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

$$x_i(t_1) = \sum_{j:(g_j, p_i) \in N_1} C_j,$$

решаем задачу (7),(12) на интервале $[t_1, t_2)$, откуда находим

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

Затем переходим к решению задачи (7),(12) на интервале $[t_2, t_3)$ с начальным условием

$$x_i(t_2) = x_i(t_2^-) + \sum_{j:(g_j, p_i) \in N_2} C_j,$$

что позволяет определить

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

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

Сформулируем теперь утверждение. Предположим, что $n_{j,i}$, для всех j и i , таких что пара (g_j, p_i) является генетическим элементом, устремляются к бесконечности так, что выполняются следующие условия:

$$\lim_{n_{j,i} \rightarrow \infty} \sum_{s=1, n_{j,i}} 1/k_{j,i,s}^{(n_{j,i})} = \tau_{j,i} , \quad \lim_{n_{j,i} \rightarrow \infty} \max_{s=1, n_{j,i}} 1/k_{j,i,s}^{(n_{j,i})} = 0 .$$

Тогда для любого $t \geq 0$ имеем:

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

5. Отображение ГГС в виде диаграмм и ориентированных орграфов

5.1. Построение структурно-функциональных диаграмм ГГС.

Пусть имеется некоторая гипотетическая генная сеть. Важными ее характеристиками являются

1. количество генетических элементов, из которых она состоит;
2. количество различных белков, синтезируемых генетическими элементами (напомним, что некоторые генетические элементы могут кодировать идентичные белки);
3. структура регуляторных связей, т.е. перечень для каждого генетического элемента тех генетических элементов, которые являются его регуляторами.

Перечисленные характеристики дают нам важное представление о строении ГГС, которое, однако, не является полным, так как не отражает информацию о механизмах регулирования. Удобным средством наглядной визуализации данной информации является представление ее в виде структурно-функциональной диаграммы. На такой диаграмме порождающие элементы (гены) будем представлять квадратами, порождаемые элементы (белки) - кружками, а промежуточные стадии синтеза белков будем представлять пунктирной дугой, идущей от квадрата к соответствующему кружку. Регуляторные связи на диаграмме будем представлять сплошной дугой, идущей от соответствующего белка к соответствующему гену. Два примера структурно-функциональных диаграмм приводятся на [рис. 1а и 1б](#). На [рис. 1а](#) все генетические элементы кодируют разные белки. На [рис.1б](#) показан пример диаграммы, на которой два генетических элемента кодируют идентичные белки.

5.2. Построение ориентированных графов ГГС

Рассмотрим теперь некоторую диаграмму ГГС. Возьмем вершину, соответствующую какому-нибудь порождающему элементу. Удалим ее и исходящую из нее пунктирную дугу, а все сплошные дуги, которые шли в удаленную вершину, направим в вершину, которая соответствует продукту, порождаемому соответствующим элементом. Прделаем эту процедуру со

всеми порождающими элементами. В результате мы получим оргграф, в котором каждая вершина будет соответствовать какому-нибудь белку, и общее число вершин будет равно числу различных белков. Дугами в таком графе будут дуги, соответствующие регуляторным связям. Полученный оргграф назовем структурным графом (графом структуры, графом регуляторных связей, оргграфом ГГС). Структурные графы отображают еще меньше информации по сравнению с диаграммами. Они показывают нам, какие белки участвуют в регуляции синтеза каждого белка ГГС, но уже не несут информацию о количестве генетических элементов. На [рис.1в и 1г](#) показаны ориентированные оргграфы, полученные из диаграмм [рис.1а и 1б](#).

Структурно-функциональные диаграммы и ориентированные графы являются не только удобным графическим средством отображения различных деталей строения ГГС. При принятии определенных соглашений относительно регуляторных механизмов используемых при построении ГГС, они позволяют однозначно выписывать соответствующие системы из МГГС и МГГСЗ. Некоторые такие соглашения мы рассматриваем в следующем разделе. Однако роль диаграмм и оргграфов этим не ограничивается.

Одна из важных задач состоит в выявлении характеристик оргграфов и диаграмм, которые играют определяющую роль в формировании законов функционирования ГГС. Ниже мы приводим некоторые гипотезы, обобщающие наши численные исследования, которые связывают устойчивые состояния функционирования ГГС с определенными характеристиками структуры соответствующего оргграфа.

6. Выделение классов ГГС

Выше мы описали процедуру задания механизмов регулирования активности генетических элементов. В данном разделе мы опишем альтернативную процедуру, которая не связана с выписыванием соответствующего оргграфа альтернативных состояний порождающего элемента.

Механизм регулирования активности функционирования генетического элемента будем составлять комбинируя четыре формальных выражения.

Первое выражение имеет вид

$$\prod_{u \in P_{j,v,k}} p_u^{h_{i,j,k,u}},$$

Его будем применять для описания действия активного регулятора на соответствующий ген пропорционально произведению некоторых степеней концентраций, входящих в него белков. Здесь и ниже i – номер белка, j – номер генетического элемента, v – номер сайта, u – номер белка, входящего в k -ый активный регулятор, $h_{i,j,k,u}$, $\gamma_{i,j,k,u}$ – степени мультимеризации, $\alpha_{i,j,v,k}$, $\delta_{i,j,v,k}$ – неотрицательные константы,

$P_{j,v,k}$ – множество белков, входящих в активный регулятор.

Простейшей биологической интерпретацией активного регулятора являются природные мультимеры, например λ -репрессор cI и антирепрессор cro фага λ [4,5] и т.д. Однако в более общем случае, такой однозначной интерпретации может и не быть. В общем случае мультипликативная зависимость является приближенным описанием более сложного механизма и в степени могут не быть целыми числами.

Второе формальное выражение имеет вид

$$S_{i,j,v} = \frac{\alpha_{i,j,v,0} + \sum_k \alpha_{i,j,v,k} \prod_{u \in P_{j,v,k}} p_u^{h_{i,j,k,u}}}{\delta_{i,j,v,0} + \sum_k \delta_{i,j,v,k} \prod_{u \in P_{j,v,k}} p_u^{\gamma_{i,j,k,u}}} .$$

Оно применяется для описания независимого воздействия нескольких активных регуляторов на генетический элемент. Простейшим примером в данном случае могут служить промоторы, которые пересекаются с оператором репрессии и с последним могут взаимодействовать несколько различных репрессоров. Например, в системах регуляции активности промоторов P_R , P_L , P_{rm} фага λ λ -репрессор и антирепрессор могут независимо связываться с сайтом O_{R1} (а также с O_{R2} , O_{R3} , O_{L1} , O_{L2} , O_{L3}) [4,5].

Третье формальное выражение

$$A_{i,j} = \prod_v S_{i,j,v} .$$

Оно учитывает совместное воздействие нескольких неконкурирующих независимых регуляторных событий. Пример такой регуляции мы снова находим у промоторов P_{rm} , P_R , P_L , фага λ . Так, участки ДНК, обозначаемые O_{R1} , O_{R2} , O_{R3} (O_{L1} , O_{L2} , O_{L3} , соответственно) составляют опереторные участки, входящие в систему регуляции промоторов P_{rm} , P_R (P_L , соответственно) фага λ и участки не пересекаются в том смысле, что с каждым из них λ -репрессор или антирепрессор могут связываться независимо от того, заняты другие сайты репрессорами или нет [4,5].

Четвертое выражение применяется для вычисления суммарного эффекта, которое может возникать как при суммировании влияния нескольких альтернативных форм, в которых могут существовать генетические элементы, так и при суммировании активностей нескольких генетических элементов:

$$M_i = \sum_j A_{i,j} .$$

Дополнительно будем считать, что для j -го генетического элемента, кодирующего i -й белок, имеется запаздывание в его синтезе, вызванное

многостадийностью процесса. Запаздывание будем обозначать через τ_j . Наличие запаздывания выражается в том, что в механизме регулирования i -го белка в момент времени t выражение $A_{i,j}$ берется в момент времени $t-\tau_j$

Введем теперь следующий перечень ограничений на конструируемые ГГС.

1. Все запаздывания приняты равными 0.
2. Рассматриваются только отрицательные регуляторные связи.
3. Все генетические элементы кодируют различные белки.
4. Генетический элемент имеет только один регуляторный сайт.
5. Активные регуляторы являются гомомультимерами белков одного вида.
6. Для каждого генетического элемента имеется только по одному активному регулятору. Активные регуляторы являются гетеромультимерами белков разного вида.
7. Все активные регуляторы действуют только на один сайт генетического элемента.
8. Каждый активный регулятор действует через самостоятельный непересекающийся с другими сайт генетического элемента.

Выделим четыре класса ГГС, исследование свойств которых представляется важным этапом изучения свойств ГГС произвольных конструкций.

Класс 1. ГГС строятся с ограничениями 1, 2, 3, 4, 5, 7. Их функционирование описывается системами следующего вида:

$$\frac{dx_i}{dt} = \frac{\alpha_i}{1 + \sum_{k \in D_i} \delta_{i,k} x_k^{\gamma_{i,k}}} - \beta_i x_i. \quad (13)$$

Класс 2. ГГС строятся с ограничениями 1, 2, 3, 5, 8. Соответствующие системы имеют следующий вид:

$$\frac{dx_i}{dt} = \frac{\alpha_i}{\prod_{k \in D_i} (1 + \delta_{i,k} x_k^{\gamma_{i,k}})} - \beta_i x_i. \quad (14)$$

Класс 3. ГГС строятся с ограничениями 1, 2, 4, 6 и описываются системами следующего вида:

$$\frac{dx_i}{dt} = \frac{\alpha_i}{1 + \delta_i \prod_{k \in D_i} x_k^{\gamma_{i,k}}} - \beta_i x_i. \quad (15)$$

Класс 4. ГГС строятся с ограничениями 1, 2, 3, 4, 6. Они описываются системами следующего вида

$$\frac{dx_i}{dt} = \sum_{k \in D_i} \frac{\alpha_{i,k}}{1 + \delta_{i,k} x_k^{\gamma_{i,k}}} - \beta_i x_i. \quad (16)$$

Численные расчеты показывают, что ГГС всех классов могут иметь самые различные режимы устойчивого функционирования – точки покоя, циклы, а также более сложные режимы функционирования. Для иллюстрации приведем несколько примеров.

На **рис. 2а** приводится простейшая структурно-функциональная диаграмма ГГС, а на **рис. 2б** – ее структурный граф. В соответствии с определениями, данными выше, по заданному структурному графу мы можем построить ГГС четырех классов. Однако, в силу простоты примера, все они совпадают и приводят к одной и той же системе уравнений :

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

Здесь x_1 и x_2 – концентрации продуктов синтеза, $\alpha \geq 0$ и $\gamma > 1$ – параметры модели. Численный анализ выявил наличие двух параметрических областей. В первой области имеется единственное устойчивое стационарное решение, к которому стремятся все решения, независимо от начальных данных. Во второй области модель имеет два устойчивых стационарных решения. Таким образом, уже самые простые модели ГГС могут иметь нетривиальные режимы поведения.

На **рис. 3** приводится более сложный пример: ГГС состоит из трех генетических элементов и трех регуляторных связей. На **рис. 3а** приводится структурно-функциональная диаграмма, а на **рис. 3б** – соответствующий структурный граф. Важно отметить, что регуляторные связи расположены по циклу. Снова все ГГС четырех классов совпадают и приводят к системе уравнений с параметрами α и γ :

$$\frac{dx_1}{dt} = \frac{\alpha}{1 + x_3^\gamma} - x_1, \quad \frac{dx_2}{dt} = \frac{\alpha}{1 + x_1^\gamma} - x_2, \quad \frac{dx_3}{dt} = \frac{\alpha}{1 + x_2^\gamma} - x_3.$$

Соответствующий численный анализ выявил существование параметрической области, в которой все решения задачи Коши при задании любых асимметричных начальных условий (т.е условий, для которых не выполняются соотношения $x_i = x_j$ для любых $i, j = 1, \dots, n$) стремятся к единственному устойчивому предельному циклу(**рис. 3г**).

Подчеркнем, что описанные режимы функционирования рассмотренных ГГС найдены в генных сетях аналогичных конструкций, полученных искусственно [6,7]. Это является важным свидетельством адекватности предложенного математического аппарата.

На **рис. 4а** приводится пример структурного графа, в котором семь вершин и восемнадцать дуг. Построим на нем ГГС первого класса. Соответствующая система уравнений имеет вид :

$$\begin{aligned} \frac{dx_1}{dt} &= \frac{\alpha}{1+x_3^\gamma+x_4^\gamma} - x_1, & \frac{dx_2}{dt} &= \frac{\alpha}{1+x_1^\gamma+x_4^\gamma} - x_2, & \frac{dx_3}{dt} &= \frac{\alpha}{1+x_2^\gamma+x_4^\gamma} - x_3, \\ \frac{dx_4}{dt} &= \frac{\alpha}{1+x_1^\gamma+x_2^\gamma+x_3^\gamma+x_5^\gamma+x_6^\gamma+x_7^\gamma} - x_4, \\ \frac{dx_5}{dt} &= \frac{\alpha}{1+x_7^\gamma+x_4^\gamma} - x_5, & \frac{dx_6}{dt} &= \frac{\alpha}{1+x_5^\gamma+x_4^\gamma} - x_6, & \frac{dx_7}{dt} &= \frac{\alpha}{1+x_6^\gamma+x_4^\gamma} - x_7, \end{aligned}$$

В результате численного анализа найдена параметрическая область, $\gamma > 3$, $\alpha > 3$, в которой модель имеет одно устойчивое стационарное решение, три устойчивых предельных циклов и пять глобально неустойчивых предельных циклов. Каждый неустойчивый предельный цикл является устойчивым в некотором собственном гиперпространстве размерности меньше 7 (**рис 4в**). При этом выполняются соотношения: для первого цикла – $x_1(t) = x_5(t+\tau) = x_3(t+2\tau) = x_7(t+3\tau) = x_2(t+4\tau) = x_6(t+5\tau)$, для второго цикла – $x_1(t) = x_6(t+\tau) = x_3(t+2\tau) = x_5(t+3\tau) = x_2(t+4\tau) = x_7(t+5\tau)$, для третьего цикла – $x_1(t) = x_7(t+\tau) = x_3(t+2\tau) = x_6(t+3\tau) = x_2(t+4\tau) = x_5(t+5\tau)$.

Наконец, рассмотрим еще один пример. На **рис.5а** приведен структурный граф, содержащий девять вершин и двадцать четыре дуги. Соответствующая для ГГС первого класса система уравнений имеет вид:

$$\begin{aligned} \frac{dx_1}{dt} &= \frac{\alpha}{1+x_3^\gamma+x_4^\gamma} - x_1, & \frac{dx_2}{dt} &= \frac{\alpha}{1+x_1^\gamma+x_4^\gamma} - x_2, & \frac{dx_3}{dt} &= \frac{\alpha}{1+x_2^\gamma+x_4^\gamma} - x_3, \\ \frac{dx_4}{dt} &= \frac{\alpha}{1+x_1^\gamma+x_2^\gamma+x_3^\gamma+x_5^\gamma+x_6^\gamma+x_7^\gamma+x_8^\gamma+x_9^\gamma} - x_4, \\ \frac{dx_5}{dt} &= \frac{\alpha}{1+x_9^\gamma+x_4^\gamma} - x_5, & \frac{dx_6}{dt} &= \frac{\alpha}{1+x_5^\gamma+x_4^\gamma} - x_6, & \frac{dx_7}{dt} &= \frac{\alpha}{1+x_6^\gamma+x_4^\gamma} - x_7, \\ \frac{dx_8}{dt} &= \frac{\alpha}{1+x_7^\gamma+x_4^\gamma} - x_8, & \frac{dx_9}{dt} &= \frac{\alpha}{1+x_8^\gamma+x_4^\gamma} - x_9. \end{aligned}$$

Численный анализ позволил выявить параметрическую область $\gamma > 3$, $\alpha > 3$, в которой имеется одно устойчивое стационарное решение, один устойчивый предельный цикл и два неустойчивых предельных цикла. При этом есть

основания полагать, что устойчивые колебания при выбранных параметрах являются квазипериодическими.

7. Канонические ГГС и гипотезы об их поведении

Системы уравнений (13)-(16) описывают функционирование ГГС четырех классов. Как видно из структуры правых частей, данные системы содержат большое количество параметров. Это сильно затрудняет непосредственное проведение качественного анализа их поведения во всем параметрическом многообразии, в частности затрудняет выявление связи структуры гипотетической геномной сети с ее динамическими свойствами. Поэтому естественно первоначальное исследование ГГС проводить, ограничившись минимально возможным набором параметров. С этой целью введем в рассмотрение класс канонических гипотетических геномных сетей.

Канонической ГГС называется геномная сеть, описываемая одной из систем (13)-(16) при следующих ограничениях на параметры: $\alpha_i = \alpha_{i,k} = \alpha$, $\delta_i = \delta_{i,k} = 1$, $\gamma_{i,k} = \gamma$, $\beta_i = 1$. Таким образом, в канонических геномных сетях оставлено всего два независимых параметра: α и γ . Тем не менее, этого вполне достаточно с биологической точки зрения, так как α имеет смысл конститутивной скорости синтеза белков, а γ имеет смысл степени нелинейности регуляции. Можно ожидать, что именно в канонических ГГС должно максимально проявляться влияние структурной организации на формирование возможных устойчивых режимов их функционирования. Это ожидание вполне подтверждается численными исследованиями, которые мы обобщаем ниже в виде гипотез-критериев.

При формулировании критериев будем пользоваться терминологией из теории графов, излагаемой в [8]. Пусть имеется ориентированный граф $G(V,W)$, где V – множество вершин, а W – множество дуг. В силу важности приведем определение понятия 1-базы, изменив только обозначения (см.[8], стр.237):

1-базой называется минимальный набор U таких попарно несмежных вершин, что любая вершина орграфа $G(V,W)$ или принадлежит U , или смежна из некоторой вершины U .

Введем понятие 1m-базы. Множество U вершин орграфа назовем 1m-базой, если и только если все вершины из множества $V \setminus U$ попарно несмежны, все вершины, имеющие нулевые полустепени заходов, лежат в U и любая вершина из U , имеющая ненулевую полустепень захода, смежна хотя бы из одной вершины, содержащейся в $V \setminus U$.

7.1 Критерий о числе устойчивых состояний канонических ГГС классов 1-4.

Пусть дан структурный оргграф G . Пусть $M_i(G)$ обозначает модель канонической ГГС i -го класса, построенную на оргграфе G . Тогда существуют такие γ_0 и α_0 , что при любых $\gamma > \gamma_0$ и $\alpha > \alpha_0$

1. для канонических ГГС первого и второго классов количество имеющих у них устойчивых состояний равняется числу 1-баз оргграфа G ;

2. для канонических ГГС третьего и четвертого классов количество имеющих у них устойчивых состояний равняется числу $1m$ - баз оргграфа G .

Далее рассматриваются только первый класс.

3. если каждая вершина G принадлежит хотя бы одной 1-базе, то $M_1(G)$ не имеет других устойчивых состояний;

4. если каноническая гипотетическая генная сеть 1-го класса, построенная на ограничении структурного оргграфа G на совокупность вершин, которые не содержатся ни в одной его 1-базе, не имеет устойчивых колебательных режимов функционирования ни при каких значениях параметров γ и α , то и исходная каноническая ГГС $S_1(G)$ не имеет устойчивых колебательных режимов функционирования.

5. если каноническая гипотетическая генная сеть 1-го класса, при достаточно больших значениях параметров γ и α , содержит устойчивые колебательные режимы функционирования, то каноническая гипотетическая генная сеть, построенная на ограничении структурного оргграфа G на совокупность вершин, которые не содержатся ни в одной его 1-базе, также содержит при достаточно больших значениях параметров γ и α устойчивые колебательные режимы функционирования.

Можно показать, что если оргграф $G(V,W)$ не имеет вершин с нулевой полустепенью захода, то некоторое подмножество его вершин U тогда и только тогда является 1-базой, когда $V \setminus U$ является $1m$ -базой. Для таких оргграфов пункты 1 и 2 являются эквивалентными.

Заметим, и это не трудно показать, что все траектории гипотетических генных сетей являются ограниченными. Поэтому из первого пункта данного критерия следует утверждение: если оргграф не имеет 1-баз, то при достаточно больших γ и α канонические ГГС первых двух классов, построенные на данном оргграфе, не имеют устойчивых стационарных решений и имеют незатухающие автоколебательные режимы функционирования. Аналогичное свойство для третьего и четвертого классов следует из пункта 2.

Из критерия также легко выводится следующий вспомогательный критерий. Если последовательное ограничение оргграфа на вершины, не

входящие ни в одну его 1-базу, приводит к пустому орграфу, то канонические генные сети первого и второго классов, построенные на исходном орграфе не имеют устойчивых колебательных режимов функционирования.

Пункт 4 критерия дает только достаточное условие отсутствия устойчивых колебательных режимов. Наличие устойчивых колебаний у канонической ГГС, построенной на ограниченном орграфе еще не гарантирует наличие таких режимов у исходной ГГС.

7.2 Описание глобальных режимов функционирования канонических ГГС, конструируемых на симметричных структурных графах.

Орграф $G(V,W)$ назовем симметричным, если и только если все вершины имеют одинаковые полустепени исходов и существует такое расположение вершин орграфа на окружности, что все достижимые из u вершины, располагаются непосредственно после нее, считая по ходу часовой стрелки. В этом случае множество D_i номеров вершин смежных в вершину u_i в системах (13)-(16) записывается в следующем виде:

$$D_i = \{ \text{mod}_n(j) \mid j = i-1, \dots, i-k-1 \},$$

$$\text{mod}_n(j) = \begin{cases} j, & 1 \leq j \leq n \\ j+n, & j \leq 0 \end{cases}$$

Так как симметричные орграфы полностью характеризуются числом вершин n и полустепенью исхода $(k-1)$, то будем данные орграфы обозначать через $G(n,k)$. На **рис. 26 и 36** приводятся примеры простейших симметричных орграфов. Смысл использования в качестве обозначения не значения полустепени, а значения, увеличенного на единицу, становится ясным из формулируемого ниже (n,k) -критерия.

Обозначим через $M_i(n,k)$ каноническую ГГС i -го класса, построенную на орграфе $G(n,k)$. Имеет место следующий ***(n,k) -критерий существования устойчивых предельных режимов функционирования моделей $M_i(n,k)$.***

Пусть d - наибольший общий делитель n и k . Если $d=k$, то существуют такие γ_0 и α_0 , что при любых $\gamma > \gamma_0$ и $\alpha > \alpha_0$ каждая из моделей $M_i(n,k)$ имеет ровно k устойчивых стационарных решений, а все другие предельные режимы функционирования $M_i(n,k)$ в указанной параметрической области неустойчивы.

Если $d \neq k$, то существуют такие γ_0 и α_0 , что при любых $\gamma > \gamma_0$ и $\alpha > \alpha_0$ модель $M_1(n,k)$ имеет d устойчивых предельных циклов, При этом в указанной параметрической области другие устойчивые режимы функционирования гипотетической генной сети 1 класса невозможны.

8. О некоторых свойствах канонических ГГС первого класса.

Для понимания свойств гипотетических генных сетей важно не только знать о наличии или отсутствии у них устойчивых режимов функционирования, но также важно знать о наличии неустойчивых точек покоя и циклов, а также важно знание отличительных характеристик предельных режимов и умение находить их численно и аналитически. Приведем описание предельных свойств для наиболее изученного нами случая канонической ГГС первого класса.

8.1. Структура точек покоя $M_1(n,k)$

1. При любых γ и α точки покоя удовлетворяют соотношениям $x_i = x_{i+jd}$, $i=1, \dots, d$, $j=1, \dots, n/d-1$.

2. Для любой точки покоя существуют, по крайней мере, два номера $1 \leq i \neq j \leq d$, такие что $x_i = x_j$. Таким образом, модель $M_1(n,k)$ не имеет асимметричных точек покоя, т.е. таких, что для любых $1 \leq i \neq j \leq d$ выполняется неравенство $x_i \neq x_j$.

3. В параметрической области $\gamma > \gamma_0$ и $\alpha > \alpha_0$ для каждого $r < d$ и любой последовательности номеров $1 \leq i_1 < \dots < i_r \leq d$ существует ровно две различные точки покоя, которые удовлетворяют соотношениям:

$$\forall i \neq j, i, j \in \{i_1, \dots, i_r\} \Rightarrow x_i = x_j \text{ и } \forall i \neq j, i, j \in \{1, \dots, d\} \setminus \{i_1, \dots, i_r\} \Rightarrow x_i = x_j.$$

4. Существует ровно одна симметричная точка покоя, которая удовлетворяет соотношению $\forall i \neq j, i, j \in \{1, \dots, d\} \Rightarrow x_i = x_j$. Других точек покоя не имеется.

5. При $d \neq k$ в параметрической области $\gamma > \gamma_0$ и $\alpha > \alpha_0$ все точки покоя являются неустойчивыми.

6. При $d=k$ $\forall 1 \leq r \leq d$ существует устойчивая точка покоя, такая что $\forall i, j \in \{1, \dots, d\} \setminus \{r\}, i \neq j \Rightarrow x_i = x_j$, $x_r = \alpha$, $x_i = 0, \forall i \neq r, i = 1, \dots, n$, принадлежит области притяжения данной точки. Согласно (n,k) -критерию, изложенному выше, всего имеется k таких точек. Все остальные точки покоя являются неустойчивыми.

8.2. Структура предельных циклов $M_1(n,k)$

Если $d \neq k$, то $\forall 1 \leq r \leq d$ точка $x_r = \alpha, x_i = 0, \forall i \neq r, i = 1, \dots, n$, лежит в области притяжения собственного устойчивого предельного цикла, в котором для любого $l=1, \dots, d$ выполняются соотношения $x_{l+id}(t - jTd/n) = x_{l+jd}(t - iTd/n)$, $i, j = 0, \dots, n/d - 1$, где через T обозначен период предельного цикла. Согласно (n,k) -критерию, изложенному выше, всего в модели $M_1(n,k)$ имеется d таких циклов. Все остальные колебательные режимы, если они имеются, являются неустойчивыми.

Следующая процедура дает алгоритм выхода на целую группу предельных режимов функционирования канонических ГГС произвольного класса, в том числе и для $M_1(n,k)$.

Пусть имеется ориентированный оргграф $G(U,D)$, где U – множество вершин, D – множество дуг. Будем считать, что все вершины из U пронумерованы в некотором порядке от 1 до n . Обозначим через D_i множество номеров вершин, из которых дуги входят в вершину u_i : $D_i = \{j \mid (u_j, u_i) \in D\}$. Функцию соответствия вершине ее номера в выбранной нумерации, обозначим через $I(u)$. Далее, разобьем множество вершин на непересекающиеся непустые подмножества: $U = U_1 \cup \dots \cup U_k$, $k \geq 1$, и обозначим его через \mathfrak{K} . Пусть Ω_i обозначает множество пар $\Omega_i = \{(s, |U_s \cap D_i|) \mid U_s \cap D_i \neq \emptyset\}$, где $|U_s \cap D_i|$ – мощность множества $U_s \cap D_i$, $i = 1, \dots, n$, и пусть для разбиения \mathfrak{K} выполняется следующее условие: для всех $j = 1, \dots, k$, и всевозможных двух вершин u и v из U_j имеем $\Omega_{i(u)} \equiv \Omega_{i(v)}$. Возьмем некоторое \mathfrak{K} -разбиение вершин графа $G(U,D)$, которое удовлетворяет введенным выше условиям. По данному разбиению построим оргграф $G(V,W)$, в котором вершинами являются множества U_j , $j = 1, \dots, k$, а дугами входящими в U_j являются дуги, исходящие из вершин, номера которых стоят первом месте пар, входящих в множество Ω_j . Назовем этот граф \mathfrak{K} -порожденным, а операцию порождения обозначим через $P(G(U,D), \mathfrak{K})$.

Обозначим каноническую ГГС i -го класса, построенную на оргграфе $G(U,D)$, через $S_i(G(U,D))$, а систему уравнений, соответствующую $S_i(G(U,D))$ – через $M_i(G(U,D))$ с переменными x_j , $j = 1, \dots, n$. Заменим все x_j в каждом уравнении системы $M_i(G(U,D))$ на соответствующее u_l , где l является номером множества U_l , содержащего x_j . В результате, как легко убедиться, система $M_i(G(U,D))$ преобразуется в систему, в которой имеется по $|U_l|$ идентичных дифференциальных уравнений относительно u_l , $l = 1, \dots, k$. Возьмем по одному такому уравнению и составим новую систему, которую обозначим через $M_i(P(G(U,D), \mathfrak{K}))$. ГГС i -го класса, которая описывается данной системой будем обозначать через $S_i(P(G(U,D), \mathfrak{K}))$ и будем называть \mathfrak{K} -порожденной.

Имеется следующий простой алгоритм построения предельного режима функционирования $S_i(G(U,D))$ по заданному предельному режиму функционирования $S_i(P(G(U,D), \mathfrak{K}))$. Пусть векторная функция (u_1, \dots, u_k) является предельным режимом функционирования $M_i(P(G(U,D), \mathfrak{K}))$. Это может быть точка покоя, цикл или аттрактор более сложной природы. Он может быть устойчивым или неустойчивым по своей природе. Зададим векторную функцию (x_1, \dots, x_n) следующим образом: $x_i = u_l \Leftrightarrow u_i \in U_l$. Тогда вектор-функция (x_1, \dots, x_n) будет представлять предельный режим функционирования того же типа канонической ГГС $S_i(G(U,D))$.

Для ГГС первого класса $S_1(G(U,D))$ мы предполагаем, что вектор-функция (x_1, \dots, x_n) , представляя глобально неустойчивый предельный режим функционирования, будет устойчивым на многообразии $x_i = x_j \Leftrightarrow (\exists l, u_i, u_j \in U_l)$.

Как легко заметить, симметричные структурные графы $G(n,k)$ имеют следующие \mathcal{K} -разбиения. Пусть $n=n_1n_2$, $n_1>k$. Тогда следующее разбиение множества вершин орграфа $G_{n,k}$ на n_1 подмножеств $U_j=\{j,j+n_2,\dots,j+(n_1-1)n_2\}$, $j=1,\dots,n_1$, является \mathcal{K} -разбиением. Это позволяет легко строить неустойчивые режимы функционирования канонических симметричных ГГС сетей. Тем не менее, хотя последнее сильно облегчает проблему описания всех режимов функционирования канонических симметричных ГГС сетей, полностью ее не решает, так как нельзя исключить, что в системах могут быть и другие неустойчивые режимы функционирования, которые нельзя построить с помощью описанной процедуры.

Заключение

В работе описывается класс теоретических объектов - ГГС, исследование свойств которых представляется важным для понимания закономерностей функционирования природных генных сетей, которые обеспечивают выполнение практически всех жизненно важных функций организмов. Предложен математический аппарат для исследования ГГС. Показано, что ГГС могут иметь устойчивые стационарные и циклические режимы функционирования, а также аттракторы более сложной природы. Численные расчеты показывают, что для канонических ГГС при достаточно больших γ и α их предельные свойства определяются только структурно-функциональными связями, которые устанавливаются посредством регулирования активности одних генетических элементов другими. Последнее выразилось в формулировании ряда эмпирических критериев, которые связывают предельные свойства канонических гипотетических генных сетей со свойствами соответствующих структурных графов. Критерии выполняются при некоторых достаточно больших значениях параметров моделей. Применительно к реальным генным сетям это означает, что необходимым условием существования в них необходимого количества стационарных и/или циклических режимов является наличие определенной минимальной сложности (нелинейности) процессов регуляции активности генетических элементов сети. Необходимая сложность может достигаться мультимеризацией белков-репрессоров и/или наличием достаточно большого числа промежуточных стадий, ведущих от гена к белку-регулятору.

Исследование гипотетических генных сетей открывает возможность для синтеза генных сетей с любым наперед заданным количеством стационарных и/или осциллирующих режимов, а также для реализации многих других, возможно более сложных, систем с режимами функционирования, ориентированными на решение определенных практических задач.

Благодарности: Авторы благодарят Н.А. Колчанова, А.В. Ратушного, Г.В. Демиденко, Кагая В.В. за плодотворное обсуждение результатов работы.

Работа получила частичную финансовую поддержку грантов РФФИ 02-04-48802-а, 02-07-90359-в, 01-07-90376-в и Междисциплинарного интеграционного проекта СО РАН "Моделирование фундаментальных генетических процессов и систем".

СПИСОК ЛИТЕРАТУРЫ.

1. Kolchanov N.A. (1997). Регуляция транскрипции генов эукариот: базы данных и компьютерный анализ. Mol. Biol. (Msk.), 31, 581–583.
2. Kolchanov N.A., Anan'ko E.A., Kolpakov F.A., Podkolodnaya O.A., Ignat'eva E.V., Goryachkovskaya T.N., Stepanenko I.L. (2000) [Gene Networks](#) Mol. Biol. (Msk.), 34, 449-460.
3. Thomas R., Thieffry D., Kaufman M. (1995) Dynamical behavior of biological regulatory networks-I. Biological role of feedback loops and practical use of the concept of the loop-characteristic state. Bulletin of Mathematical Biology, 57, 247-276.
4. Johnson AD, Poteete AR, Lauer G, Sauer RT, Ackers GK, Ptashne M. lambda Repressor and cro--components of an efficient molecular switch. Nature 1981 294(5838):217-23. Sugano M., Okamoto H., Ide T. // Agricultural and Biological Chemistry. 1978. V. 42(11). P. 2009-2014.
5. Ptashne. M. A Genetic Switch: Phage Lambda and Higher Organisms. Oxford: Blackwell Scientific, 1992. 217 p.
6. [Gardner TS, Cantor CR, Collins JJ.](#) (2000) Construction of a genetic toggle switch in Escherichia coli. Nature, 403, 339-342.
7. [Elowitz MB, Leibler S.](#) (2000) A synthetic oscillatory network of transcriptional regulators. Nature. 403, 335-338.
8. Харари Ф. Теория графов. М.: Мир, 1973, 300с.

Рисунки к статье.

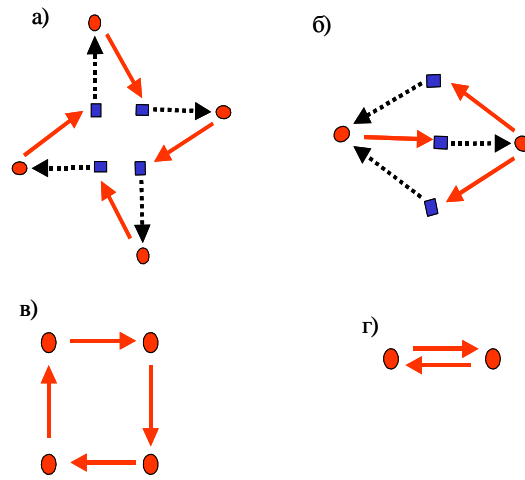
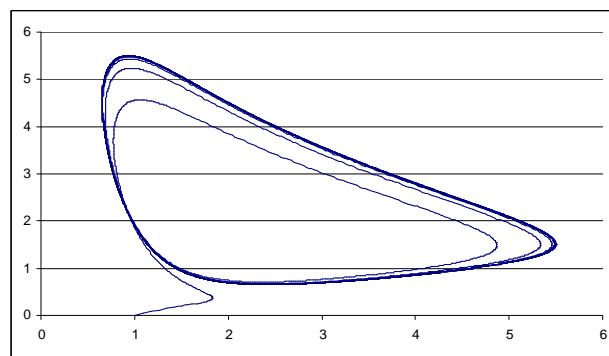
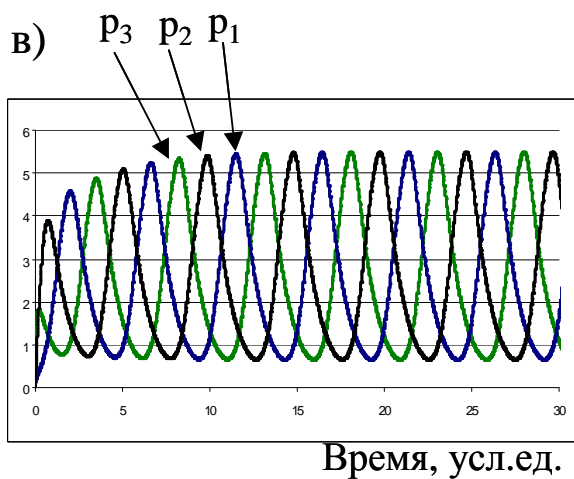
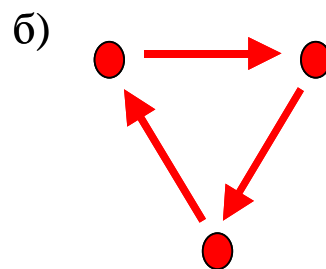
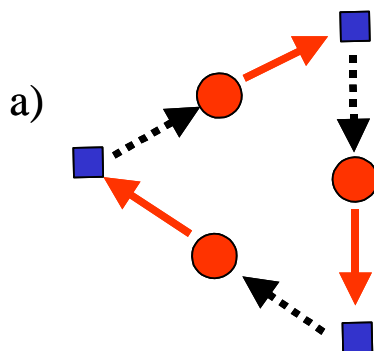


Рис.1 Диаграммы и орграфы гипотетических генных сетей



Рис.2 Простейшая гипотетическая генная сеть.

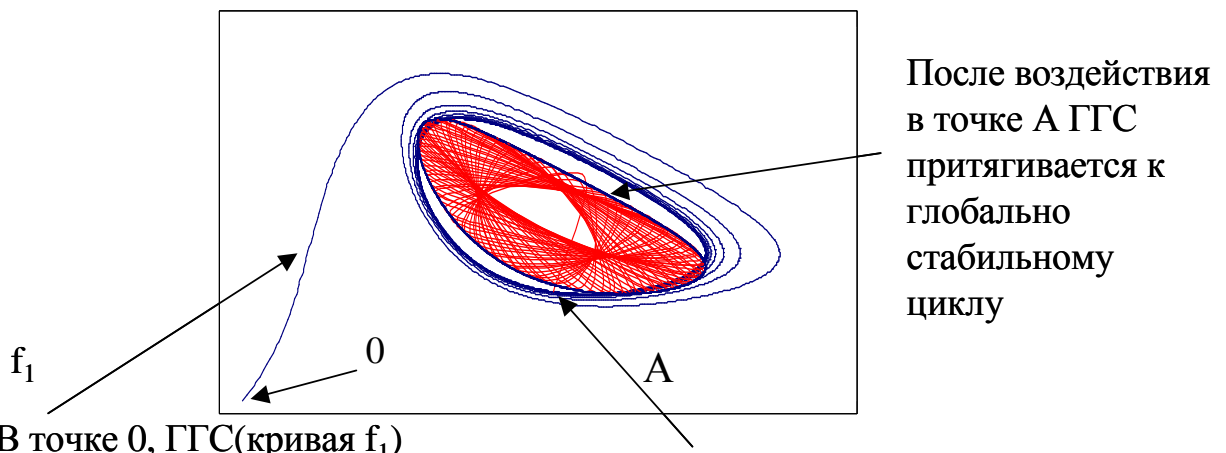
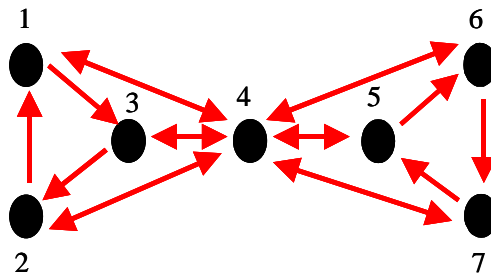


Динамика процесса выхода
на циклический режим функционирования

Фазовый портрет в плоскости (p_1, p_2)

Рис.3 Гипотетическая генная сеть, имеющая циклический режим функционирования

Структурный граф геной сети



После воздействия в точке А ГГС притягивается к глобально стабильному циклу

В точке 0, ГГС(кривая f_1) стартует в плоскости гиперпространства $p_1=p_5$, $p_2=p_6$, $p_3=p_7$ и остается в ней, притягиваясь к собственному циклу

В точке А слабым воздействием функцию вывели из плоскости гиперпространства $p_1=p_5$, $p_2=p_6$, $p_3=p_7$.

Рис.4. Пример геной сети, составленной из семи генетических элементов.

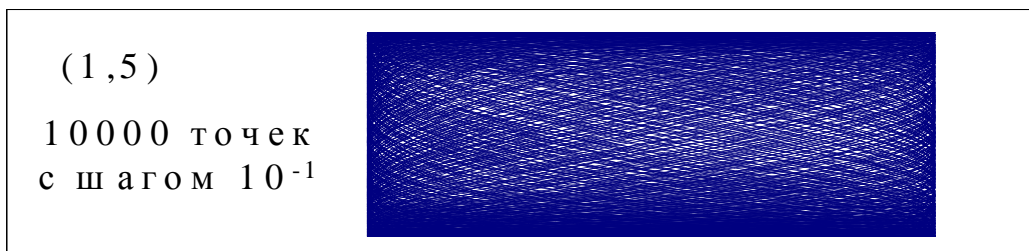
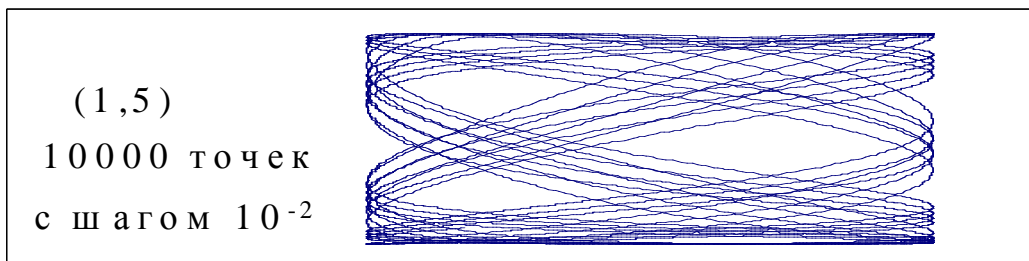
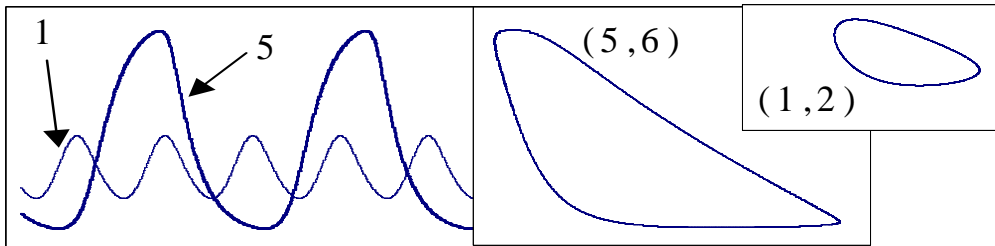
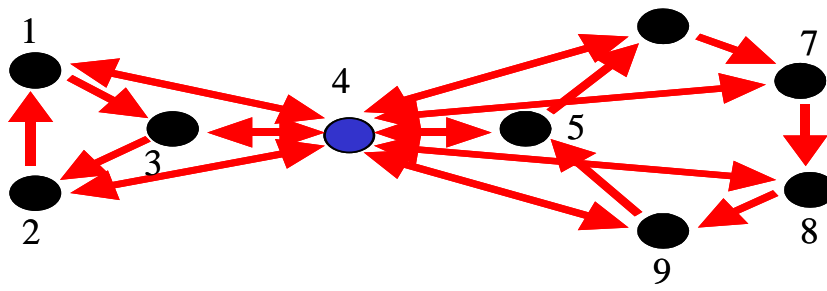


Рис.5. Пример генной сети, составленной из девяти генетических элементов. При $\gamma > 3$, $\alpha > 3$ численно выявляется одна устойчивая точка покоя и один незатухающий режим функционирования, предположительно квазицикл.