

Электронная библиотека GeneNet: описание и моделирование генных сетей животных и растений

Ананько Е.А., Лихошвай В.А., Колпаков Ф.А.,
Подколотный Н.Л., Ратушный А.В., Игнатьева Е.В.,
Подколотная О.А., Степаненко И.Л., Колчанов Н.А.

Институт цитологии и генетики (Сибирское отделение Российской Академии наук),
Новосибирск, 630090, Россия, genenet@bionet.nsc.ru

1 ВВЕДЕНИЕ

Генная сеть - группа координированно экспрессирующихся генов, контролирующая выполнение определенной функции организма. В любой генной сети можно выделить несколько обязательных типов компонентов: ядро генной сети - указанная выше группа генов; белки, кодируемые этими генами; пути передачи сигналов от клеточных мембран в ядра клеток, обеспечивающие активацию или подавление транскрипции генов; отрицательные и положительные обратные связи, стабилизирующие параметры генной сети на определенном уровне или, напротив, отклоняющие их от исходного значения, обеспечивая переход к новому функциональному состоянию; низкомолекулярные компоненты, осуществляющие переключение функции генных сетей в ответ на внешние воздействия (гормоны и другие сигнальные молекулы), энергетические компоненты, различные метаболиты, и т.д. [1].

Важным моментом в функционировании генной сети является ее связь с внешней средой, в том числе и с другими генными сетями. Поэтому в любой генной сети имеются компоненты, обеспечивающие либо восприятие и передачу внешних сигналов, либо способность продуцировать такие сигналы [1, 2].

Характерной особенностью организации генных сетей является их способность к саморегуляции за счет замкнутых регуляторных контуров с отрицательными и положительными обратными связями [3]. Молекулярной основой существования таких регуляторных контуров является наличие сайтов-мишеней в ДНК, РНК и белках, с которыми могут взаимодействовать различные молекулярные компоненты генной сети и внешние регуляторные факторы. Благодаря этим двум типам регуляторных контуров возможно поддержание определенного функционального состояния генной сети или ее переход в другой режим функционирования, в том числе и под влиянием факторов внешней среды.

Регуляция функции генных сетей не ограничивается уровнем транскрипции. Они могут регулироваться на уровне трансляции [4, 5], сплайсинга [5, 6, 7], посттрансляционной деградации белков [8], активного мембранного транспорта [9] и других.

В клетках, тканях, органах и организмах одновременно протекает огромное количество молекулярно-генетических, биохимических и физиологических процессов, контролируемых генными сетями. Для того чтобы понять регуляторные механизмы этих процессов и получить возможность для направленного воздействия на них, необходимо создание эффективных компьютерных методов описания, реконструкции и моделирования генных сетей.

Первые теоретические исследования генных сетей начались в 60-е годы. Они были посвящены рассмотрению общих закономерностей организации молекулярно-генетических систем управления у прокариот [10] и описанию динамики генных сетей в рамках простейших логических схем [11]. Позднее были предложены подходы, основанные на использовании дифференциальных уравнений и стохастических моделей [12 - 14]. До середины 80-х годов развитие теории генных сетей ограничивалось отсутствием достаточного количества экспериментальных данных.

В последнее десятилетие в связи с появлением эффективных методов изучения молекулярных механизмов регуляции экспрессии генов началось стремительное накопление экспериментальных данных по функционированию генных сетей (http://www.witi.cs.uni-magdeburg.de/iti_bm/ibss/courses/). Однако, эти данные разбросаны по множеству статей, каждая из которых содержит результаты исследования лишь отдельных аспектов функционирования какой-либо генной сети. Для накопления информации о генных сетях нами разрабатывается база данных GeneNet [1, 2], описание которой дается в настоящей статье. GeneNet доступна через Интернет (<http://www.mgs.bionet.nsc.ru/systems/mgl/genenet/>) и является частью системы GeneExpress (<http://www.mgs.bionet.nsc.ru/mgs/systems/geneexpress/>)-электронной библиотеки по структуре и функции ДНК, РНК и белков [15, 16].

©Вторая Всероссийская научная конференция
ЭЛЕКТРОННЫЕ БИБЛИОТЕКИ:
ПЕРСПЕКТИВНЫЕ МЕТОДЫ И ТЕХНОЛОГИИ,
ЭЛЕКТРОННЫЕ КОЛЛЕКЦИИ
26-28 сентября 2000г., Протвино

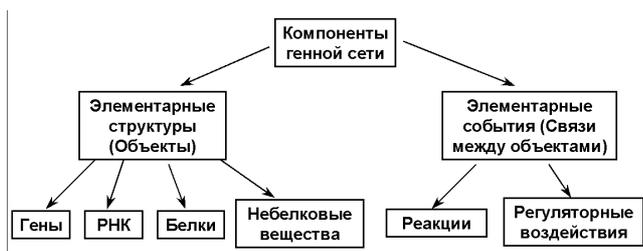


Рис. 1: Классификация компонентов геной сети в системе GeneNet

2 ОБЪЕКТНО-ОРИЕНТИРОВАННОЕ ПРЕДСТАВЛЕНИЕ ИНФОРМАЦИИ.

База данных GeneNet создана в рамках объектно-ориентированного подхода [17]. Компоненты геной сети разделены на два типа: элементарные объекты и элементарные взаимодействия между объектами (рис. 1). Объекты классифицируются как гены, РНК, белки и белковые комплексы, низкомолекулярные небелковые вещества, органы, ткани, клетки, клеточные компартменты и т.д. Взаимодействия между объектами делятся на два класса: (1) реакции - молекулярные события, в результате которых образуются новые объекты (транскрипция, трансляция, мультимеризация, модификация белков и т.д.); (2) регуляторные воздействия - включение, усиление, подавление или полное выключение процесса.

В рамках GeneNet возможно описание любых геной сетей про- и эукариот на различных иерархических уровнях организации: молекулярном, клеточном и организменном [1]. GeneNet позволяет учитывать функциональное состояние компонентов геной сети, их видовую принадлежность, распределенность по различным органам, тканям, клеткам и клеточным компартментам.

3 ВВОД ИНФОРМАЦИИ В БАЗУ ДАННЫХ GENENET.

С помощью специального графического интерфейса [2] пользователь может вводить данные в GeneNet, оперируя естественными понятиями молекулярной биологии, связанными с регуляцией экспрессии генов (рис. 2). Большим достоинством данного интерфейса является то, что при работе с ним нет необходимости знать формат базы данных. Пользователь работает с графическими изображениями объектов и связей между ними и диалоговыми формами для их описания. Программа осуществляет автоматическую трансляцию введенной информации в стандартный формат GeneNet. Каждый класс объектов описывается в отдельной таблице, при этом имеется возможность при необходимости добавлять новые классы объектов (таблицы GeneNet). В настоящее время в базе данных GeneNet имеется 12 взаимосвязанных таблиц, в каждой из которых используется собственный формат данных, учитывающий особенности данного класса объектов.

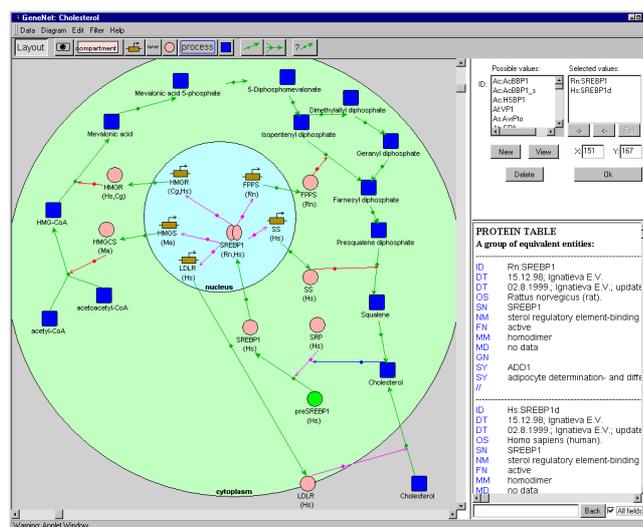


Рис. 2: Интерфейс ввода данных в систему GeneNet. Геной сеть, регулирующая биосинтез холестерина в клетке.

4 ФОРМАЛИЗОВАННОЕ ОПИСАНИЕ КОМПОНЕНТОВ ГЕННЫХ СЕТЕЙ

Формат данных, используемый в GeneNet для описания объектов, основан на использовании нескольких информационных полей и позволяет учитывать особенности каждого класса. Например, при описании белков (таблица PROTEIN, пример входа из которой приведен ниже) указывается полное и краткое название белка (поля SN, NM), синонимы названия (SY), функциональное состояние (active/inactive/no data, поле FN), степень мультимеризации (monomer/homodimer/heterodimer/no data, поле MM), модификация белка (phosphorylated/non phosphorylated, поле MD), ген, кодирующий данный белок (ссылка на таблицу GENE, поле GN), клетки, в которых проводился эксперимент (SO), ссылки на внешние базы данных (DR) и на источник информации (ссылка на таблицу LITER, поле RF).

```
ID Gg:GATA1-p
DT 09.10.97; Ananko E.; created.
DT 19.6.1999.; Podkolodnaya O.A.; updated.
OS Gallus gallus (chicken).
SN GATA1
NM transcription factor GATA1
SY NF-E1
FN active
MM no data
MD phosphorylated
GN Gg:GATA-1
DR TFFACTOR; T00267;
DR SWISSPROT; P17678;
DR EMBL; M26209;
DR PIR; A32993;
SO Gg:Eryth Prog
RF Briegel K, et al. 1996
//
```

Взаимодействия между объектами генной сети описываются в таблице RELATION. Поле **ID** предназначено для описания объектов, вступающих во взаимодействие, и указания их локализации в компартментах клетки или организма. Класс объекта указывается в треугольных скобках (<gene>, <RNA>, <protein>, <substance>). После класса объекта указывается вид организма (Hs - Homo sapiens, Mm - Mus musculus и т.д.). Далее указывается имя объекта и его локализация, разделенные знаком "^". Если во взаимодействие вступает несколько объектов, их описания разделяются запятой. Исходные объекты и продукты взаимодействия разделены стрелкой >. В поле **EF** дается характеристика связи между исходными объектами и объектами, возникшими в результате взаимодействия (прямое или не прямое). Последнее используется в тех случаях, когда неизвестны (или опускаются при описании) промежуточные стадии взаимодействия, приводящие к формированию конечного объекта из исходных. В поле **AT** указывается тип регуляторного влияния ("increase усиление", "decrease ослабление", "switch on включение"). В поле **RF** приведена ссылка на оригинальный источник информации, в поле **DT** - дата создания записи и ее автор. Например, приведенная ниже запись реакции означает, что с гена олигоденилатсинтазы (OAS) человека, локализованного в ядре, в цитоплазме клетки эксперсируется кодируемый белок. Взаимодействие не прямое, поскольку при описании опущены промежуточные стадии транскрипции, процессинга, сплайсинга и т.д.

```
ID <gene>Hs:OAS^ nucleus ->
<protein>Hs:OAS^ cytoplasm
DT 17.5.1999; Ananko E.; created
EF indirect
RF Wathelet M. et al., 1986
//
```

Следующая запись (регуляторного события) означает, что в митохондриях фермент феррохелатаза (FCH) человека катализирует синтез гема (Heme) из предшественников (Fe⁺⁺, Proto IX). Взаимодействие субстратов и фермента прямое.

```
ID <protein>Hs:FCH^ mitochondrion ->>
<substance>Fe++^ mitochondrion, <substance>Proto
IX^ mitochondrion -> <substance>Heme^ mitochondrion
DT 23.6.1999.; Podkolodnaya O.A.; created
AT switch on
EF direct
RF Ponka P., 1997
//
```

5 ВИЗУАЛИЗАЦИЯ СТРУКТУРНО-ФУНКЦИОНАЛЬНОЙ ОРГАНИЗАЦИИ ГЕННЫХ СЕТЕЙ

Большим достоинством созданной технологии является возможность автоматической визуализации генных сетей [1]. Формализованные данные о генной сети, накопленные в базе GeneNet, обрабатываются с помощью специальной программы (GeneNet viewer) и представляются пользователю в виде графической схемы (рис. 2). Каждый ком-

понент генной сети имеет свое изображение, отражающее его особенности. Например, форма изображения белка отражает степень его мультимеризации, цвет - функциональное состояние и т.д. Все изображения на схеме генной сети являются интерактивными, т.е., щелкнув мышкой по выбранному объекту, можно получить его текстовое описание из базы в специальном окошке (рис. 2). По умолчанию схема строится на основе всех имеющихся данных. При этом гомологичные объекты разных организмов изображаются как единственный объект, однако, описание выдается на всю группу. Пользователь имеет возможность отфильтровать данные по виду организма и/или типу клеток. При этом все объекты, не удовлетворяющие условиям фильтра, исчезают со схемы, а оставшиеся не меняют своего положения. Форматированный текст снабжен гипертекстовыми ссылками на другие базы данных (EMBL, SWISS-PROT, TRRD, TRANSFAC, EPD, MEDLINE) и другие таблицы GeneNet.

6 ИНФОРМАЦИОННОЕ СОДЕРЖАНИЕ БАЗЫ ДАННЫХ GENENET

. В настоящее время в базе данных GeneNet описано 23 генных сети, контролирующих процессы липидного метаболизма, кроветворения, противовирусного ответа, функции эндокринной системы, регуляцию развития семени растений, азотфиксацию, тепловой шок и другие процессы. На 8 июня 2000 года в базе описано 310 генов, 529 белков, 132 небелковых вещества, 1132 взаимодействия между объектами 43 организмов. Данные собраны на основе реферирования 429 оригинальных статей.

7 МОДЕЛИРОВАНИЕ ДИНАМИКИ ГЕННЫХ СЕТЕЙ

Для моделирования динамики функционирования генных сетей использован обобщенный химико-кинетический метод моделирования (ОХКММ) [18-20]. ОХКММ ориентирован на формализованное описание закономерностей функционирования произвольных биосистем, а не только генных сетей. Формализация осуществляется на основе блочного принципа, согласно которому каждый процесс выделяется в самостоятельную элементарную подсистему и описывается изолированно от других процессов (рис. 3).

Элементарные подсистемы описываются в рамках ОХКММ на основе формальных блоков. В ОХКММ под формальным блоком понимается выделенная в отдельную единицу конструкция, которая однозначно характеризуется упорядоченным списком формальных динамических переменных, упорядоченным списком формальных параметров и законом преобразования информации. На рис. 4 приводится 5 формальных блоков, используемых при моделировании. Список формальных блоков может пополняться при необходимости, и они могут иметь произвольные законы преобразования информации - непрерывные, дискретные, стохастические. В ОХКММ собственно моделью является совокупность описанных элементарных

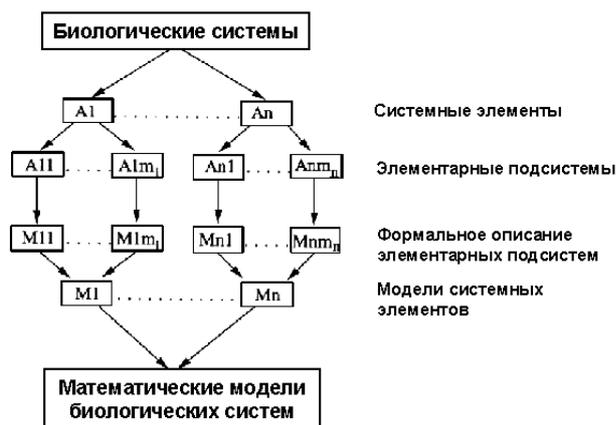


Рис. 3: Блочный принцип моделирования.

Формальный блок 1. Взаимодействующая обратимая реакция: $x_1 + x_2 \xrightleftharpoons[k_2]{k_1} x_3$
 $X = (x_1, x_2, x_3), \bar{P} = (k_1, k_2), F: \frac{dx_1}{dt} = k_2 \cdot x_3 - k_1 \cdot x_1 \cdot x_2, \frac{dx_2}{dt} = \frac{dx_3}{dt} = -\frac{dx_1}{dt}$

Формальный блок 2. Моноспекулярная необратимая реакция: $x \xrightarrow{k} y_1 + y_2 + \dots + y_n$
 $X = (y_1, y_2, \dots, y_n), \bar{P} = k, F: \frac{dx}{dt} = -\frac{dy_1}{dt} = -\frac{dy_2}{dt} = \dots = -\frac{dy_n}{dt} = -k \cdot x, n \geq 0$

Формальный блок 3. Конститутивный синтез: $\xrightarrow{k} x_1 + \dots + x_m$
 $X = (x_1, \dots, x_m), \bar{P} = (k), F: \frac{dx_i}{dt} = k, i = 1, \dots, m, n \geq 1$

Формальный блок 4. Образование n продуктов из одновременно реагирующих субстратов
 $X = (x_1, x_2, \dots, x_m, y_1, y_2, \dots, y_n), \bar{P} = (k_1, k_2, a_1, \dots, a_m, b_1, \dots, b_n)$
 $F: \frac{dx_j}{dt} = -a_j \cdot k_j \cdot x_j, j = 1, \dots, m, \frac{dy_l}{dt} = b_l \cdot k_l \cdot z, l = 1, \dots, n$
 $Z = \frac{k_2 x_1 \dots x_m}{(k_2 + x_1) \dots (k_2 + x_m) - x_1 \dots x_m}, m \geq 2, n \geq 0$

Формальный блок 5. "реакция" $x_1 + \dots + x_m \rightarrow y_1 + \dots + y_n$
 $X = (x_1, \dots, x_m, y_1, \dots, y_n), \bar{P} = (k)$
 $F: \frac{dy_i}{dt} = k \cdot x_1 \dots x_m, m \geq 0, n \geq 1$

Рис. 4: Формальные блоки ОХКММ. \bar{X} - упорядоченный список формальных переменных, \bar{P} - упорядоченный список формальных параметров, F - закон преобразования информации.

процессов, а модели конкретной биосистемы или ее модификаций являются ее конкретными реализациями, которые мы далее будем называть частными моделями.

7.1 Процедура построения моделей в рамках ОХКММ.

Для наглядности обратимся к простейшему примеру. Пусть имеется три "гена" a1, a2 и a3, на которых синтезируются индивидуальные мРНК. Пусть на этих мРНК синтезируются белки, пусть мономерный продукт гена a1 ингибирует транскрипцию с генов a2 и a3, мономерный продукт гена a2 ингибирует транскрипцию с генов a1 и a3, а мономерный продукт гена a3 ингибирует транскрипцию с генов a1 и a2. Пусть эти мономеры являются активаторами транскрипции собственных генов, а димеры белков - ингибиторами.

Пусть мРНК и белки имеют ограниченные време-

```

Исотеми! элементарный процесс 1: Транскрипция
Исотеми! формальный блок 3
gene{ } RNA{pol} => mRNA{ }
Исотеми! список имен "генов" с константами
{a1},{a2},{a3}: k1{ }=10^4 M^-1 sec^-1,

Исотеми! элементарный процесс 2: Трансляция
Исотеми! формальный блок 5
mRNA{ } + ribo{some} => protein{ }
Исотеми! список имен "генов" с константами
{a1}: k2{ }=10^7 M^-1 sec^-1,
{a2}: k2{ }=5-10^7 M^-1 sec^-1
{a3}: k2{ }=10^9 M^-1 sec^-1

Исотеми! элем. процесс 3: репрессия генов a2 и a3
Исотеми! формальный блок 1
gene{ } + protein{a1} <-> r gene{ }
Исотеми! список имен "генов" с константами
{a2},{a3}: k31{ }=10^9 M^-1 sec^-1, k32{ }=0.1 sec^-1

Исотеми! элем. процесс 4: репрессия генов a1 и a3
Исотеми! формальный блок 1
gene{ } + protein{a2} <-> r gene{ }
Исотеми! список имен "генов" с константами
{a1},{a3}: k41{ }=10^9 M^-1 sec^-1, k42{ }=0.1 sec^-1

Исотеми! элем. процесс 5: репрессия генов a1 и a2
Исотеми! формальный блок 1
gene{ } + protein{a3} <-> r gene{ }
Исотеми! список имен "генов" с константами
{a1},{a2}: k51{ }=10^9 M^-1 sec^-1, k52{ }=0.1 sec^-1

Исотеми! элем. процесс 6: формир. димеров белков
Исотеми! формальный блок 1
protein{ } + protein{ } <-> pr otain{ }
Исотеми! список имен "генов" с константами
{a1},{a2},{a2}: k61{ }=10^9 M^-1 sec^-1, k62{ }=0.1 sec^-1

Исотеми! элем. процесс 7: активация транскрипции
Исотеми! форм. блок 3
protein{ } + gene{ } + rRNA{pol} => mRNA{ }
Исотеми! список имен "генов" с константами
{a1},{a2},{a3}: k7{ }=10^9 M^-1 sec^-1

Исотеми! элем. процесс 8: репрессия транскрипции генов димерами белков
Исотеми! формальный блок 1
gene{ } + protein{a1} <-> r gene{ }
Исотеми! список имен "генов" с константами
{a1},{a2},{a3}: k81{ }=10^9 M^-1 sec^-1, k82{ }=0.1 sec^-1

Исотеми! элем. процесс 9: деградация мРНК
Исотеми! формальный блок 3
mRNA{ } =>
Исотеми! список имен "генов" с константами
{a1},{a2},{a3}: k9{ }=10^-4 sec^-1

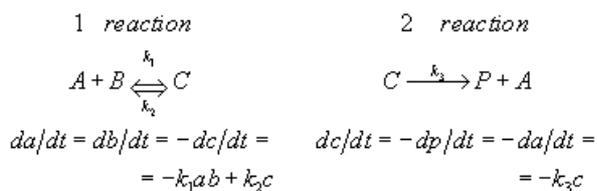
Исотеми! элем. процесс 10: деградация белков
Исотеми! формальный блок 3
protein{ } =>
Исотеми! список имен "генов" с константами
{a1},{a2},{a3}: k10{ }=10^-4 sec^-1

```

Рис. 5: Демонстрационная модель

на жизни. Дадим следующее описание системы в терминах элементарных процессов (рис. 5): здесь **gene**, **RNA**, **mRNA**, **ribo**, **protein**, **protein2**, **rgene**, **rgene2**, **rmRNA** - имена продуктов; $k1, k2, k3, \dots$ - имена констант, а $\{\}, \{\text{pol}\}, \{\text{some}\}, \{\text{a1}\}, \{\text{a2}\}, \{\text{a3}\}$ - их псевдорасширения. Псевдорасширение $\{\}$ считается пустым. Сами фигурные скобки имеют смысл ограничителей и в имена не включаются. Псевдорасширения несут о продукте дополнительную информацию, которая используется при построении конкретных частных моделей. В первом элементарном процессе продукт `gene{ }` обозначает активный промотор "вообще", что задается отсутствием символов между фигурными скобками, **mRNA{ }** - молекулам РНК, синтезируемая с данного промотора. **RNA{pol}** - обозначает конкретную РНК-полимеразу, осуществляющую синтез всех этих мРНК со всех однотипных промоторов. Таким образом, элементарный процесс 1 описывает синтез мРНК на произвольном "гене". Конкретный перечень промоторов задаем списком имен "генов" ($\{\text{a1}\}, \{\text{a2}\}, \{\text{a3}\}$). С каждым именем "гена" всегда связывается определенный список значений параметров. Элементарный процесс 2 описывает процесс трансляции мРНК, синтезируемых с "генов" a1, a2 и a3. Здесь рибосомы обозначены через **ribo{some}**, а синтезируемый белок - через **protein{ }**. Аналогично расшифровываются элементарные процессы 3-10. Для облегчения понимания, мы снабдили каждый элементарный процесс кратким комментарием. В первой строке разъясняется краткая суть процесса и дается его порядковый номер, действительный в рамках данного примера. Во второй - приводится номер формального блока, принятый в рамках рис. 4. По этому номеру можно однозначно восстановить формализм, принятый для данного блока в рамках ОХКММ.

Построение частных моделей из элементарных процессов осуществляется для конкретных модификаций биосистемы. Информация о конкретном виде модификации биосистемы задается путем ввода списка имен MAP. Формально MAP задается, как линейно-упорядоченный ориентированный список имен. Наиболее близкой аналогией MAP являются генетические карты. В них гены ориентированы и линейно упорядочены. Так как физические носители генетических карт, ДНК или РНК, могут быть кольцевыми, линейными, состоять из нескольких физически не связанных молекул, то MAP, как формальная



Model

$$\begin{cases}
 da/dt = (-k_1ab + k_2c) + (k_3c) \\
 db/dt = (-k_1ab + k_2c) \\
 dc/dt = -(-k_1ab + k_2c) - (k_3c) \\
 dp/dt = (k_3c)
 \end{cases}$$

Рис. 6: Правило суммирования правых частей.

конструкция, наделена в ОХКММ этими же свойствами. Будем называть имена из MAP "генами", а их совокупность "геномом".

Обратимся снова к модели (рис. 5). Зададим, для определенности, "геном"MAP1={a1}, {a2}, содержащий два "гена". Продемонстрируем, как из элементарных процессов 1-10 по "геному"MAP1 строится частная модель. Для краткости обозначим ее через model(MAP1). Отберем все процессы, содержащие имя{a1} в списках имен "генов"этих процессов. Это будут процессы 1, 2, 4-10. В каждом таком процессе заменим псевдорасширения всех продуктов и параметров на расширения, следуя правилам, указанным в табл.1.

Затем отберем из модели все процессы, содержащие имя{a2} - это будут элементарные процессы 1-3, 5-10; и, снова, заменим в них псевдорасширения продуктов и параметров на расширения. Отобранные элементарные процессы являются искомой частной моделью. В ней имеется 10 оригинальных продуктов, отличающихся друг от друга по имени и/или по расширению. Продукты участвуют в 18 элементарных процессах. Если теперь задать "геном"MAP2={a1}, {a2}, {a3}, то model(MAP1) расширится за счетдобавления к ней элементарных процессов 1, 2, 4-10. В частной модели model(MAP2)будет уже 15 динамических переменных, которые будут связаны между собой посредством участия в 27 элементарных процессах. Этим не исчерпывается весь список частных моделей, которые можно построить на основе данных элементарных процессов. Мы можем также построить частные модели для других карт, таких как MAP({a1}), MAP({a2}) и т.д.

Рассмотренный пример, несмотря на свою условность, показывает, что моделирование в рамках обобщенного химико-кинетического метода моделирования, допускает построение, на основе заданного описания элементарных процессов, составляющих исходную биосистему, большого разнообразия частных моделей. Это свойство является особенно важным при описании динамики функционирования генных сетей, так как в природе они существуют в разных генетических вариантах.

Модели, возникающие при описании процессов в терминах химико-кинетических блоков, формально относят-

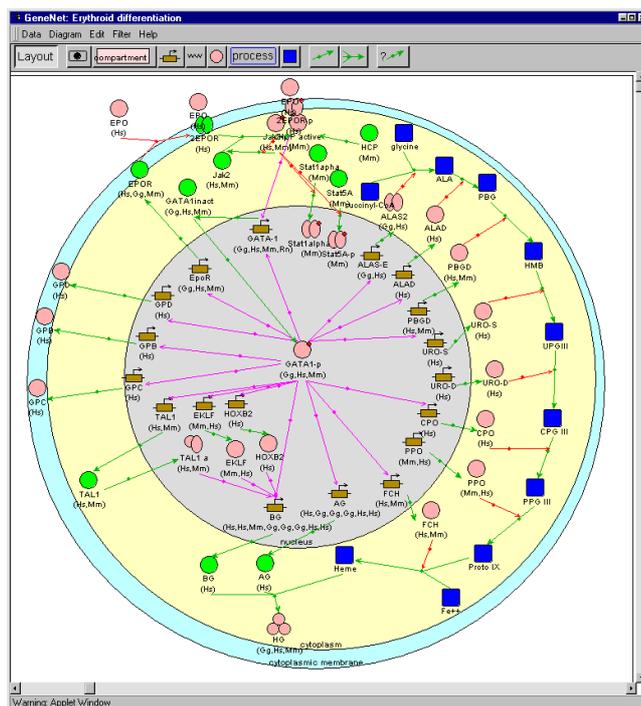


Рис. 7: Генная сеть, регулирующая созревание эритроцита.

ся к классу систем автономных дифференциальных уравнений. Их правая часть выписывается на основе закона суммирования правых частей (рис. 6). Для численного интегрирования используется метод Гира [21].

7.2 Два примера моделирования динамики функционирования генных сетей

В рамках ОХКММ нами разработаны модели двух генных сетей. Первая модель описывает функционирование генной сети, контролирующей внутриклеточный биосинтез холестерина (рис. 2). Модель состоит из 65 элементарных процессов. Описание генной сети, представленной на рис.2, содержит 40 динамических переменных и 93 константы.

Вторая модель описывает систему регуляции созревания эритроцита (рис. 7) и состоит из 119 элементарных процессов. Модель генной сети, представленной на рис. 7, содержит 68 динамических переменных и 178 констант. Модели адаптированы к экспериментальным данным и позволяют рассчитывать численные характеристики генных сетей в разных условиях ее функционирования. В качестве иллюстрации приведем результаты расчета отечной реакции первой системы на повышение содержания частиц ЛНП в плазме крови в 2 раза (рис. 8).

7.3 Заключение

Для интеграции экспериментальной информации об особенностях организации и механизмах функционирования

Таблица 1: Правила замены псевдорасширений расширениями при построении моделей из элементарных процессов.

	Псевдорасш.	Расширение	Смысл операции
1	0	Натуральное число n	n - порядковый номер "гена" в MAP, имя которого совпадает с именем "гена", в списке имен, приданных элементарному процессу
2	+i	Натуральное число $n+i$	если часть "генома", состоящая из "генов" с номерами $m+1, \dots, N$ является кольцевой и $m_n, n+i > N$ ($N-m$ - число "генов" в кольцевом "геноме"), то в качестве расширения присваивается число $n+i-N+m$
3	-i	Натуральное число $n-i$	если часть "генома", состоящая из "генов" с номерами $m+1, \dots, N$ является кольцевой и $m_n, n-i < m+1$, то в качестве расширения присваивается число $N-n+i+m$.
4	Пустое	Имя гена	Комбинация символов S псевдорасширения становится расширением
5	Другое	S	

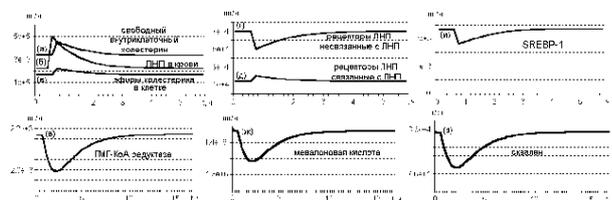


Рис. 8: Кинетики изменения основных компонентов системы регуляции биосинтеза холестерина в клетке (расчет по модели)

генных сетей многоклеточных организмов нами разработана база данных GeneNet, наиболее существенными особенностями которой являются: (а) наличие гибкого языка для формализованного описания всей совокупности элементарных структур, событий и процессов, значимых для функционирования генных сетей; (б) возможность описания в рамках единого подхода как механизмов регуляции экспрессии генов, так и различных метаболических процессов; (в) возможность описания генных сетей как про-, так и эукариот, а также их совместного функционирования; (г) наличие понятной и удобной в работе системы описания генных сетей и ввода информации в базу данных; (д) возможность автоматической визуализации структуры генных сетей на основе информации, накопленной в базе данных. Разрабатываются средства моделирования динамики генных сетей, отражающих фундаментальные молекулярно-генетические процессы, на основе информации, накапливаемой в базе данных GeneNet.

Дальнейшее развитие базы данных GeneNet будет осуществляться по следующим направлениям:

- 1) накопление количественных данных по динамике генных сетей и элементарных процессов;
- 2) совершенствование формата для описания генных сетей с учетом их иерархической организации и пространственной распределенности процессов по различным компартментам организма;
- 3) развитие методов математического моделирования

динамики генных сетей на основе информации, накопленной в базе данных GeneNet;

- 4) расширение круга генных сетей, описываемых в базе данных GeneNet, в первую очередь - генных сетей, обеспечивающих осуществление фундаментальных молекулярно-генетических процессов.

7.4 Благодарности

Работа поддержана Российским фондом фундаментальных исследований (гранты 98-04-49479, 98-07-91078, 98-07-90126, 99-07-90203, 00-04-49229, 00-04-49255, 00-07-90337), Всероссийской программой "Геном человека" и Комитетом по науке и технике РФ.

7.5 Список литературы:

1. Kolpakov F.A., Ananko E.A., Kolesov G.B., Kolchanov N.A. (1998) GeneNet: a database for gene networks and its automated visualization. *Bioinformatics*, **14**, 529-537.
2. Kolpakov F.A., Ananko E.A. (1999) Interactive data input into the GeneNet database. *Bioinformatics*, **15**, 713-714.
3. Колчанов Н.А. Регуляция транскрипции генов эукариот: базы данных и компьютерный анализ // *Mol. Biol.* 1997. Т. 31. С. 581-583.
4. Buss W.C. and Stepanek J. (1993) Characterization of the inhibition of renal translation in the Sprague-Dawley rat following in vivo cyclosporin A. *Int. J. Immunopharmacol.*, **15**, 63-76.
5. Pyronnet S., Vagner S., Bouisson M., Prats A.C., Vaysse N. and Pradayrol L. (1996) Relief of ornithine decarboxylase messenger RNA translational repression induced by alternative splicing of its 5' untranslated region. *Cancer Res.*, **56**, 1742-1745.
6. Nandabalan K. and Roeder G.S. (1995) Binding of a cell-type-specific RNA splicing factor to its target regulatory sequence. *Mol. Cell. Biol.*, **15**, 1953-1960.
7. Yao K.S., Godwin A.K., Johnson C. and O'Dwyer P.J. (1996) Alternative splicing and differential expression of DT-diaphorase transcripts in human colon tumors and

in peripheral mononuclear cells in response to mitomycin C treatment. *Cancer Res.*, **56**, 1731-1736.

8. Hochstrasser M. (1996) Protein degradation or regulation: Ub the judge. *Cell*, **84**, 813-815.

9. Weissmuller G. and Bisch P.M. (1993) Autocatalytic cooperativity and self-regulation of ATPase pumps in membrane active transport. *Eur. Biophys. J.*, **22**, 63-70.

10. Ратнер В.А. Генетические управляющие системы. Новосибирск: Наука, 1966. 181 с.

11. Kauffman S. (1969) Metabolic stability and Epigenesis in randomly constructed genetic net. *J. Theoret. Biol.*, **22**, 437-467.

12. Thomas R. (1973) Boolean formalization of genetic control circuits. *J. Theor. Biol.*, **42**, 563-585.

13. Savageau M. (1985) A theory of alternative designs for biochemical control systems. *Biomed. Biochim. Acta.*, **44**, 875-880.

14. McAdams H. and Arkin A., (1997) Stochastic mechanism in gene expression. *Proc. Natl. Acad. USA*, **94**, 814-819.

15. Kolchanov N.A., Ponomarenko M.P., Kel A.E., et al. (1998) GeneExpress: a computer system for description, analysis, and recognition of regulatory sequences of the eukaryotic genome. The Sixth International Conference on Intelligent Systems for Molecular Biology June 28 - July 1 1998, Montreal, Canada, pp.95-104.

16. Колчанов Н.А., Лаврюшев С.В. и др. ГЕНЭКС-

ПРЕСС: электронная библиотека по структурам и функциям ДНК, РНК и белков. Электронные библиотеки: перспективные методы и технологии, электронные коллекции. СПб, 1999, стр. 161-169.

17. Schweigert S., Herde P.V., Sibbald P.R. (1995) Issues in incorporation semantic integrity in molecular biological object-oriented databases. *Comput. Appl. Biosci.*, **11**, 339-347.

18. Likhoshvai V.A., Bazhan S.I., Belova O.E. Mathematical modeling of molecular genetic system regulation interferon induction and antiviral action. (1992) *Modeling and computer methods in molecular biology and genetics* (ed. by V.A.Ratner and N.A.Kolchanov, Nova Science Publishers, Inc., USA), 293- 299.

19. Belova O.E., Likhoshvai V.A., Bazhan S.I. (1994) The Knowledge-Base System for Investigation of Regulation and Functioning of Molecular-Genetic Systems *Proceedings of EWHCI'94* The East-West International Conference on Human- Computer Interaction, St. Petersburg, 2-6 August, 1994. **2**, 115-118.

20. Bazhan S.I., Likhoshvai V.A., Belova O.E. (1995) Theoretical analysis of the regulation of interferon expression during priming and blocking. *J. Theor. Biol.*, **175**, 149-160.

21. Gear C.W. (1971) The automatic integration of ordinary differential equations. *Communs. ACM.*, **14**, 176-190.