



Математическое моделирование динамики генных сетей

A.B. Ратушный

Кафедра информационной биологии ФЕН НГУ



Цели лекции

- Рассмотреть общие понятия:

Что такое модель и для чего она нужна? Что такое моделирование?

Приемы и способы моделирования. Процесс построения модели.

Генные сети как объект моделирования, их специфика и особенности.

- Методы моделирования динамики генных сетей и их классификация

- Обобщенный химико-кинетический метод моделирования

- Формальное описание математических и компьютерных моделей генных сетей

- Автоматическое конструирование математических моделей на основе информации из баз данных

- Проблемы математического моделирования генных сетей

- Верификация параметров математических моделей генных сетей (решение обратной задачи)



Что такое модель?

Модель, в широком смысле любой образ, аналог (мысленный или условный: изображение, описание, схема, чертеж, график, план, карта и т. п.) какого-либо объекта, процесса или явления («оригинала» данной модели), используемый в качестве его «заместителя», «представителя».

Энциклопедия Кирилла и Мефодия



Что такое моделирование?

Моделирование - исследование каких-либо явлений, процессов или систем объектов путем построения и изучения их моделей; использование моделей для определения или уточнения характеристик и рационализации способов построения вновь конструируемых объектов. Моделирование одна из основных категорий теории познания: на идее моделирования по существу базируется любой метод научного исследования как теоретический (при котором используются различного рода знаковые, абстрактные модели), так и экспериментальный (использующий предметные модели).

Энциклопедия Кирилла и Мефодия



Для чего нужна модель?

Для понимания, как устроен изучаемый объект: какова его структура, основные свойства, законы развития и взаимодействия с окружающим миром;

Для того чтобы научится управлять объектом (или процессом) и определить наилучшие способы управления при заданных целях и критериях;

Для того чтобы прогнозировать прямые и косвенные последствия реализации заданных способов и форм воздействия на объект.

Горстко А.Б. Познакомьтесь с математическим моделированием, 1991 г.



Приемы и способы моделирования

Моделирование

Материальное (предметное)

Способы моделирования, при которых исследование ведется на основе модели, воспроизводящей основные геометрические, физические, динамические и функциональные характеристики изучаемого объекта.



Физическое

(реальному объекту противопоставляется его увеличенная или уменьшенная копия)
- макет здания;
- модель самолета;
- планетарий.

Аналоговое

(основано на аналогии процессов и явлений, имеющих различную физическую природу, но одинаково описываемых формально)
- изучение механических колебаний с помощью электрической схемы.

Идеальное

Носит теоретический характер и основано не на материальной аналогии объекта и модели, а на аналогии идеальной, мыслимой.



Интуитивное

(основано на интуитивном представлении объекте исследования)
- жизненный опыт (интуитивная модель окружающего мира).

Знаковое

на (использует в качестве моделей знаковые преобразования об какого-либо вида: схемы, графики, чертежи, формулы, наборы символов и т.д., а также включает совокупность законов, по которым можно оперировать с выбранными знаковыми образованиями и их элементами)



Процесс построения моделей

1. Формируются основные вопросы о поведении системы, ответы на которые мы хотим получить с помощью модели.
3. Из множества законов, управляющих поведением системы, учитываются те, влияние которых существенно при поиске ответов на поставленные вопросы.
5. В дополнение к этим законам, если необходимо, для системы в целом или отдельных ее частей формулируются определенные гипотезы о функционировании. Как правило, эти гипотезы правдоподобны в том смысле, что могут быть приведены некоторые теоретические доводы в пользу их принятия.
7. Гипотезы так же, как и законы, выражаются в форме определенных математических соотношений, которые объединяются в некоторое формальное описание модели.

Горстко А.Б. Познакомьтесь с математическим моделированием, 1991 г.



Генные сети – молекулярная основа жизни

Функционирование живых организмов контролируется информацией, закодированной в их генах. Генетическая информация считывается и преобразуется с помощью фундаментальных генетических процессов: репликации, транскрипции, трансляции и др. Синтезируемые на их основе генетические макромолекулы (ДНК, РНК и белки) взаимодействуют друг с другом, обеспечивая функционирование генных сетей.

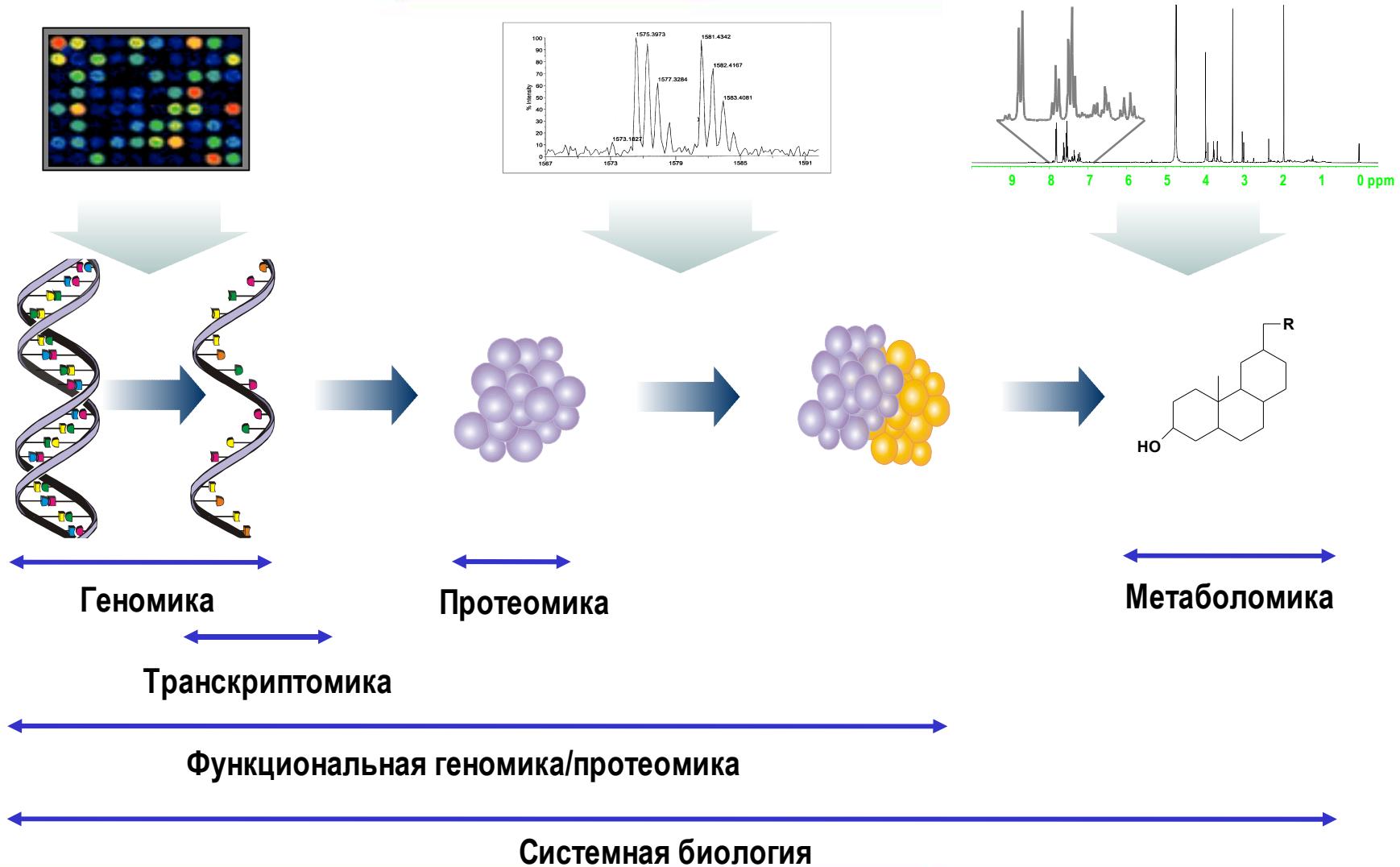


Актуальность проблемы

Одной из центральных проблем постгеномной молекулярной биологии и генетики является исследование принципов организации, функционирования и эволюции генных сетей - сложных молекулярно-генетических систем, обеспечивающих формирование фенотипических (молекулярно-генетических, биохимических, физиологических, морфологических и т.д.) характеристик живых организмов на основе информации, закодированной в их геномах.



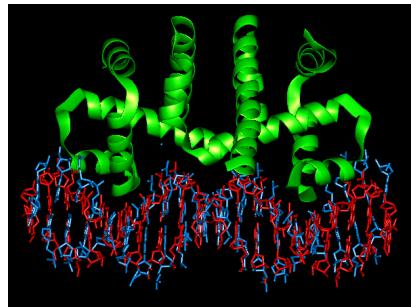
Всестороннее исследование генных сетей



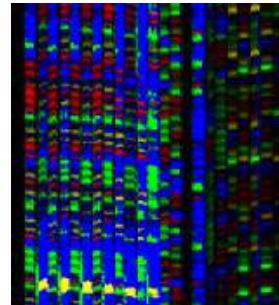


Интеграция данных по генным сетям

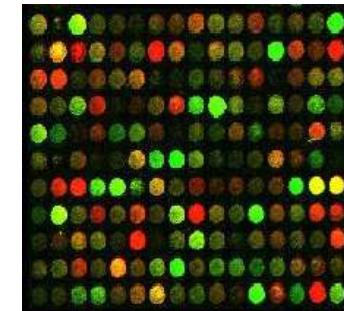
ДНК-белковые
взаимодействия



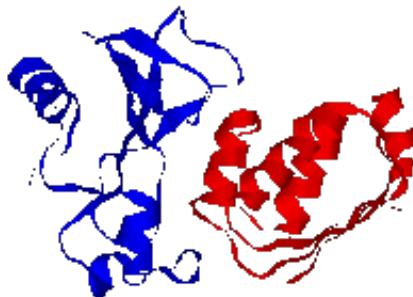
Последовательности геномов



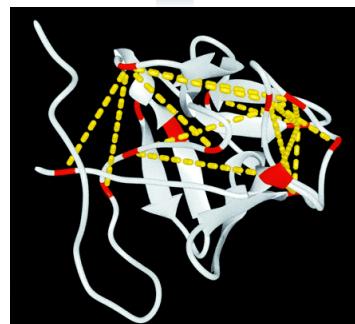
Содержание мРНК



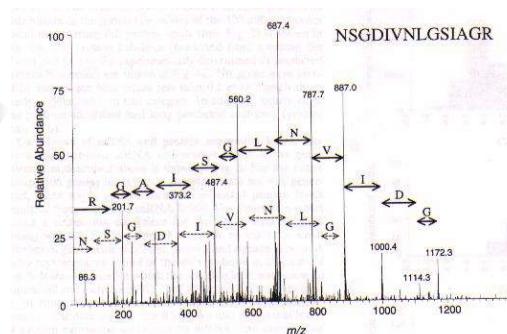
Математические модели
клеток и организмов



Белок-белковые
взаимодействия



Белковые структуры



Содержание
белков и метаболитов



Ресурсы и базы данных

Regulatory Pathways

GeneNet: Gene Networks Database

<http://wwwmgs.bionet.nsc.ru/mgs/gnw/genenet/>

KEGG: Regulatory Pathways

<http://www.genome.ad.jp/kegg/regulation.html>

SPAD: Signal Transduction

<http://www.grt.kyushu-u.ac.jp/spad/>

CSNDB: Cell Signaling Networks

<http://geo.nihs.go.jp/csndb/>

MIPS: Yeast Pathways

<http://www.mips.biochem.mpg.de/proj/yeast/pathways/index.html>

Interactive Fly: Drosophila Genes

<http://sdb.bio.purdue.edu/fly/aimain/1aahome.htm>

GeNet: Gene Networks Database

http://www.csa.ru/Inst/gorb_dep/inbios/genet/genet.htm

HOX-Pro: Homeobox Genes Database

http://www.iephb.nw.ru/labs/lab38/spirov/hox_pro/hox-pro00.html

Wnt Signaling Pathway

<http://www.stanford.edu/~rnusse/wntwindow.html>

TRANSPATH: Gene Regulatory Pathways

<http://transpath.gbf.de/>

Metabolic Pathways

KEGG: Metabolic Pathways

<http://www.genome.ad.jp/kegg/metabolism.html>

EMP: Enzymes and Metabolic Pathways

<http://emp.mcs.anl.gov/>

WT: Metabolic Reconstruction

<http://wit.mcs.anl.gov/WT2/>

UM-BBD: Microbial Biocatalysis/Biodegradatation

<http://umbbd.ahc.umn.edu/>

BioCyc: Prokaryotic Genes and Metabolism

<http://biocyc.org/>

EcoCyc: E. coli Genes and Metabolism

<http://ecocyc.org/>

HumanCyc: Homo sapiens Genes and Metabolism

<http://humancyc.org/>

MetaCyc: Metabolism from 150 species

<http://ecocyc.org/>

SoyBase: Soybean Metabolism

<http://cgsc.biology.yale.edu/metab.htm>

Metalgen: Genes and Metabolism

<http://indigo.genetique.uvsq.fr/>

Boehringer Mannheim: Biochemical Pathways

<http://www.expasy.ch/cgi-bin/search-biochem-index>

IUBMB-Nicholson Minimaps

<http://www.tcd.ie/Biochemistry/IUBMB-Nicholson/>



Ресурсы и базы данных

Enzymes, Compounds and Reactions

LIGAND: Biochemical Compounds and Reactions

<http://www.genome.ad.jp/ligand/>

ENZYME: Enzymes

<http://www.expasy.ch/enzyme/>

BRENDA: Comprehensive Enzyme Information System

<http://www.brenda.uni-koeln.de/>

Worthington Enzyme Manual

<http://www.worthington-biochem.com/manual/manIndex.html>

Klotho: Biochemical Compounds

<http://www.biocheminfo.org/klotho/>

ChemFinder: Searching Chemicals

<http://chemfinder.camsoft.com/>

ChemIDplus at NLM

<http://chem.sis.nlm.nih.gov/chemidplus/>

PROMISE: Prosthetic Groups and Metal Ions

<http://metallo.scripps.edu/PROMISE/>

GlycoSuiteDB: Glycan Structure Database

<http://www.glycosuite.com/>

CarbBank: Complex Carbohydrate Structure Database

<http://bssv01.lancs.ac.uk/gig/pages/gag/carbbank.htm>

LIPIDBANK: for Web - Lipids

<http://lipidbank.jp/>

WebElements: Periodic Table

<http://www.webelements.com/>

PDB: Protein Data Bank

<http://www.rcsb.org/pdb/>

Protein-Protein Interactions

BRITE: Database for Biomolecular Relations

<http://www.genome.ad.jp/brite/>

DIP: Database of Interacting Proteins

<http://dip.doe-mbi.ucla.edu/>

BIND: Biomolecular Interaction Network Database

<http://www.binddb.org/>

Transcription Factors

TRRD: Transcription Regulatory Regions Database

<http://wwwmgs.bionet.nsc.ru/mgs/gnw/trrd/>

TRANSFAC: Transcription Factor Database

<http://transfac.gbf.de/TRANSFAC/index.html>

RegulonDB: E. coli Transcriptional Regulation

http://www.cifn.unam.mx/Computational_Genomics/regulondb

DBTBS: B. subtilis Transcription Factors

<http://elmo.ims.u-tokyo.ac.jp/dbtbs/>

SCPD: S. cerevisiae Promoter Database

<http://cgsigma.cshl.org/jian/>

DPIInteract: DNA binding proteins

<http://arep.med.harvard.edu/dpinteract/>

Gene Expression Pattern

Axel×db: Xenopus laevis

<http://www.dkfz-heidelberg.de/abt0135/axel×db.htm>

NEXTDB: Caenorhabditis elegans

<http://nematode.lab.nig.ac.jp/>

MAGEST: Halocynthia roretzi

<http://www.genome.ad.jp/mages/>



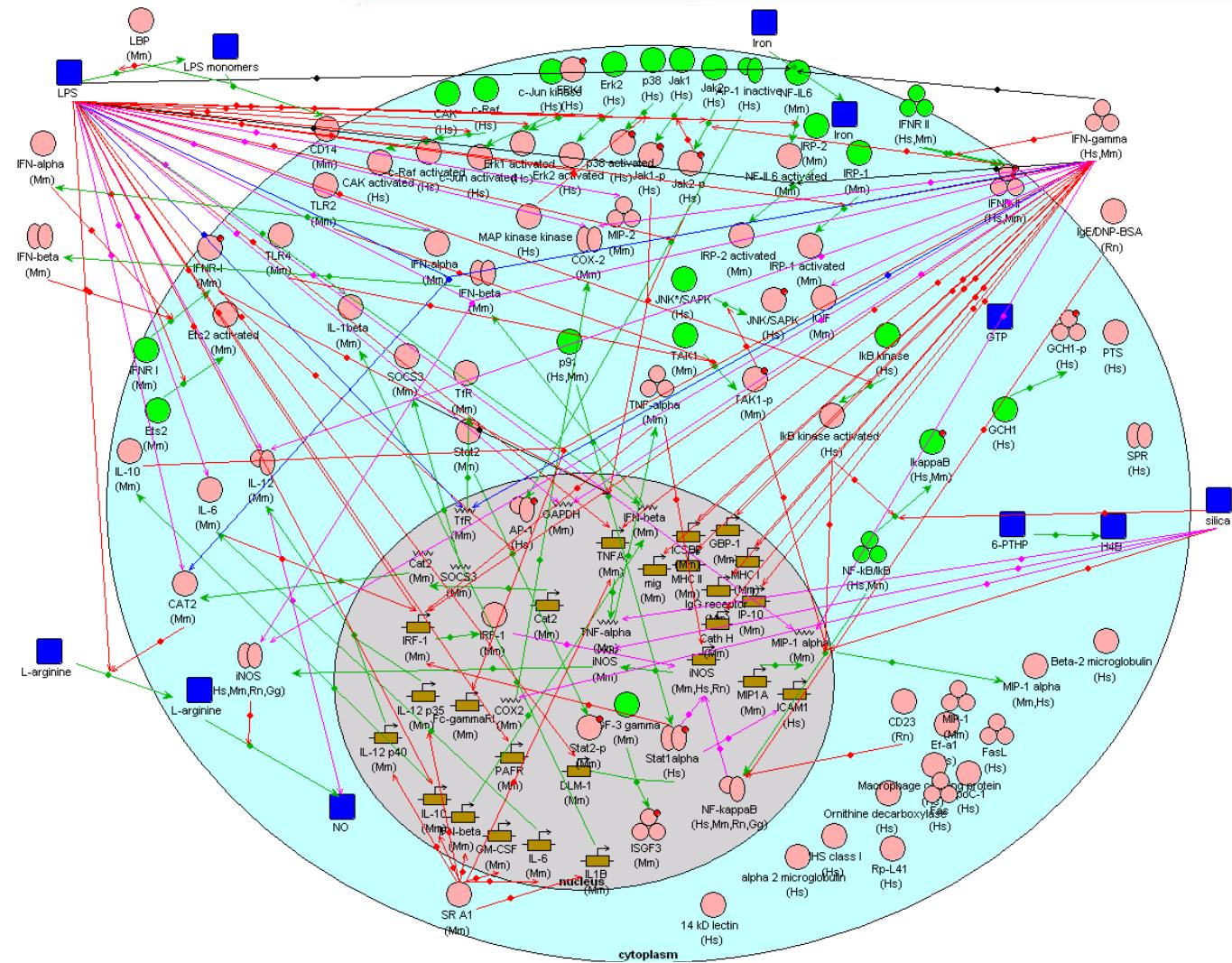
Генные сети как объект исследования

Генные сети – сверхбольшие системы со сверхсложной структурно-функциональной организацией и динамическими свойствами.

Анализ огромных объемов экспериментальных данных, отражающих сложные процессы функционирования молекулярно-генетических систем, исследование структурной организации, молекулярных механизмов функционирования, закономерностей эволюции генных сетей, оценки влияния мутаций на функцию генных сетей, реконструкции генных сетей на основе экспериментальных данных и т.д., принципиально невозможен без использования современных информационных технологий и эффективных математических методов анализа данных и моделирования биологических систем и процессов.



Пример сложноустроенной генной сети





Основные элементы генной сети

Основные элементы генной сети





Специфика объекта моделирования

Рассматриваемые системы могут состоять из сотен и тысяч компонентов с очень сложной структурно-функциональной организацией, функционирующей на многих иерархических уровнях живых систем

→ **Биохимический уровень** (биохимические процессы и реакции)

→ **Генетический уровень** (расположение и ориентация генов и регуляторных элементов, регуляция экспрессии генов, матричные ориентированные процессы, полигапельность и т.д.)

→ **Субклеточный и клеточный уровни** (комpartmentализация, процессы межкомpartmentного обмена веществами и энергией, активный и пассивный транспорт и т.д.)

→ **Органый, тканевой и организменный уровни** (сверхкомpartmentы, взаимодействие с внешней средой)

→ **Популяционный уровень** (взаимодействие между особями и с внешней средой, возрастная структура популяции, эволюционные процессы)

В большинстве случаев, эти объекты обладают нелинейным поведением в силу наличия в них отрицательных и положительных обратных связей



Специфика объекта моделирования

Биологические системы являются самовоспроизводящимися динамическими системами. Особенностью функционирования биологических систем во времени является то, что у них, как правило, в процессе функционирования не остаются постоянными параметры пространственной организации, а также могут меняться характеристики генетического строения.

Факторы изменения пространственной организации строения биологической системы:
рост организмов, регенерация органов и тканей, деление и гибель клеток, органел, размножение организмов, метаморфоз и т.д.

Факторы изменения генетического строения биологической системы:
репликация, рекомбинация, мутации, делеции, вставки, деление клеток, слияние половых клеток, трансгенез, привнесение новых геномов или частей генома путем внедрение в биологическую систему других биологических систем, как то вирусов, бактерий, паразитов, параллельный перенос генетической информации и т.д.



Классификация методов моделирования

Классификация методов моделирования динамики генных сетей с привязкой ко времени:

- непрерывные методы
- дискретные методы
- гибридные методы



Классификация методов моделирования

Классификация методов моделирования динамики генных сетей с привязкой к биологическому процессу:

- детерминистический подход (подходы, основанные на применении дифференциальных и интегро-дифференциальных уравнений)
- стохастический подход
- логический подход



Классификация методов моделирования

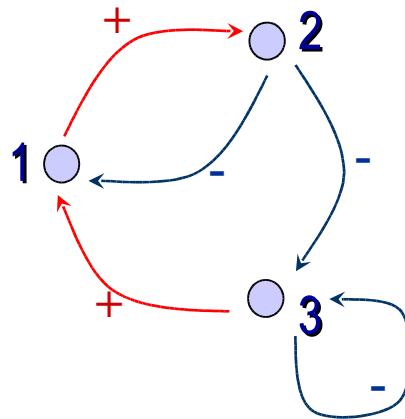
Классификация методов моделирования динамики генных сетей, основанная на типах формализации молекулярно-генетических процессов:

- химико-кинетический подход и его обобщения
- подходы, основанные на дифференциальных уравнениях в частных производных (модели пространственно распределенных систем)
- сети Петри и их обобщения
- Булевы сети и обобщенные логические сети
- Байесовы сети и их обобщения
- пороговый метод и его обобщения (использование кусочно-линейных дифференциальных уравнений и логического подхода)
- стохастическое моделирование и его обобщения
- методы, основанные на системе правил (экспертные системы)



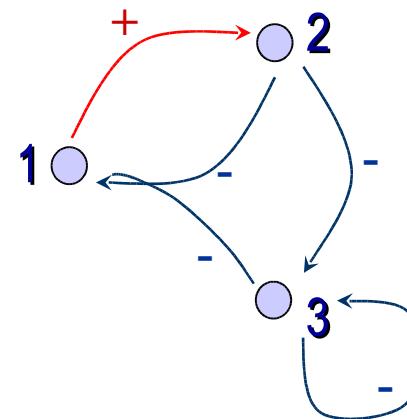
Графическое моделирование генных сетей

Орграф G состоит из конечного множества V вершин и набора E упорядоченных пар различных вершин. Любая такая пара (i, j) называется дугой, или ориентированным ребром, и обычно обозначается ij .



$$V = \{1, 2, 3\}$$

$$E = \{(1, 2, +), (2, 1, -), (2, 3, -), (3, 1, +), (3, 3, -)\}$$



$$V = \{1, 2, 3\}$$

$$E = \{([1], 2, [+]), ([2], 1, [-]), ([2], 3, [-]), ([3], 1, [+]), ([3], 3, [-])\}$$



Построение графических моделей генных сетей

Графические модели генных сетей строятся на основе информации, хранящейся в базах данных и базах знаний, но они также могут быть реконструированы на основе данных по экспрессии генов посредством подходов *inductive* либо *reverse engineering*.

Для выявления паттернов экспрессии (*similar temporal expression patterns*) применяются различные алгоритмы кластеризации экспрессионных данных.

Для дополнительной верификации структуры генной сети и выявления мнимых регуляторных связей применяется ряд различных методик:

- анализ временных запаздываний (*time lags*);
- проведение экспериментов по возмущению системы (*perturbation experiments*);
- картирование коэкспрессирующихся генов на выполнение определенной функции в организме (*mapping of co-expressed genes to pathway functions*);
- скрининг промоторных областей коэкспрессирующихся генов для выявления одинаковых регуляторных районов.



Булевы сети

Модели булевых сетей (*Boolean networks*) основаны на булевой логике. Предполагается, что каждый ген может находиться в двух состояниях: “экспрессируется”, либо “не экспрессируется”. Для простоты, эти состояния обозначаются “1” и “0”, соответственно. Состояние всей модели описывается списком экспрессирующихся и не экспрессирующихся генов.



Булевы сети

Пример булевой функции для трехгенной системы (XYZ):

$$X(\text{next}) = X(\text{current}) \text{ AND } Y(\text{current})$$

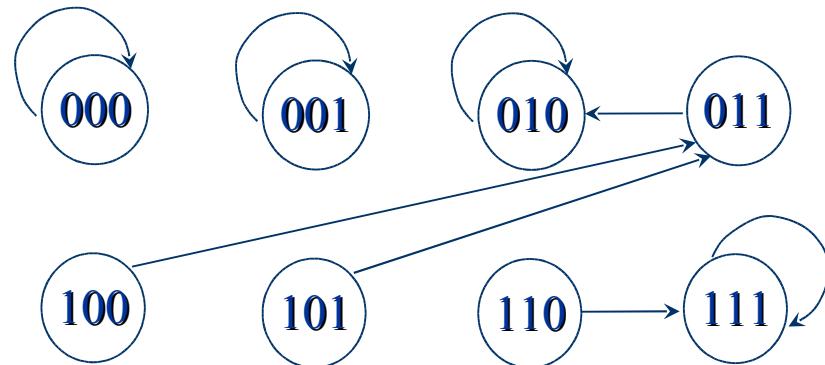
$$Y(\text{next}) = X(\text{current}) \text{ OR } Y(\text{current})$$

$$Z(\text{next}) = X(\text{current}) \text{ OR } \{\text{NOT } Y(\text{current})\} \text{ AND } Z(\text{current})\}$$

Таблица состояний (*truth table*):

<u>Current state</u>	<u>Next state</u>
000	000
001	001
010	010
011	010
100	011
101	011
110	111
111	111

Конечный автомат (*finite-state machine*):





Обобщенный химико-кинетический метод моделирования (ОХКММ)

ОХКММ разработан в ИЦиГ СО РАН и используется нами для моделирования динамики генных сетей;

ОХКММ потенциально содержит все подходы, рассмотренные выше;

ОХКММ ориентирован на формализованное, в первую очередь портретное, описание функционирования произвольных биосистем.



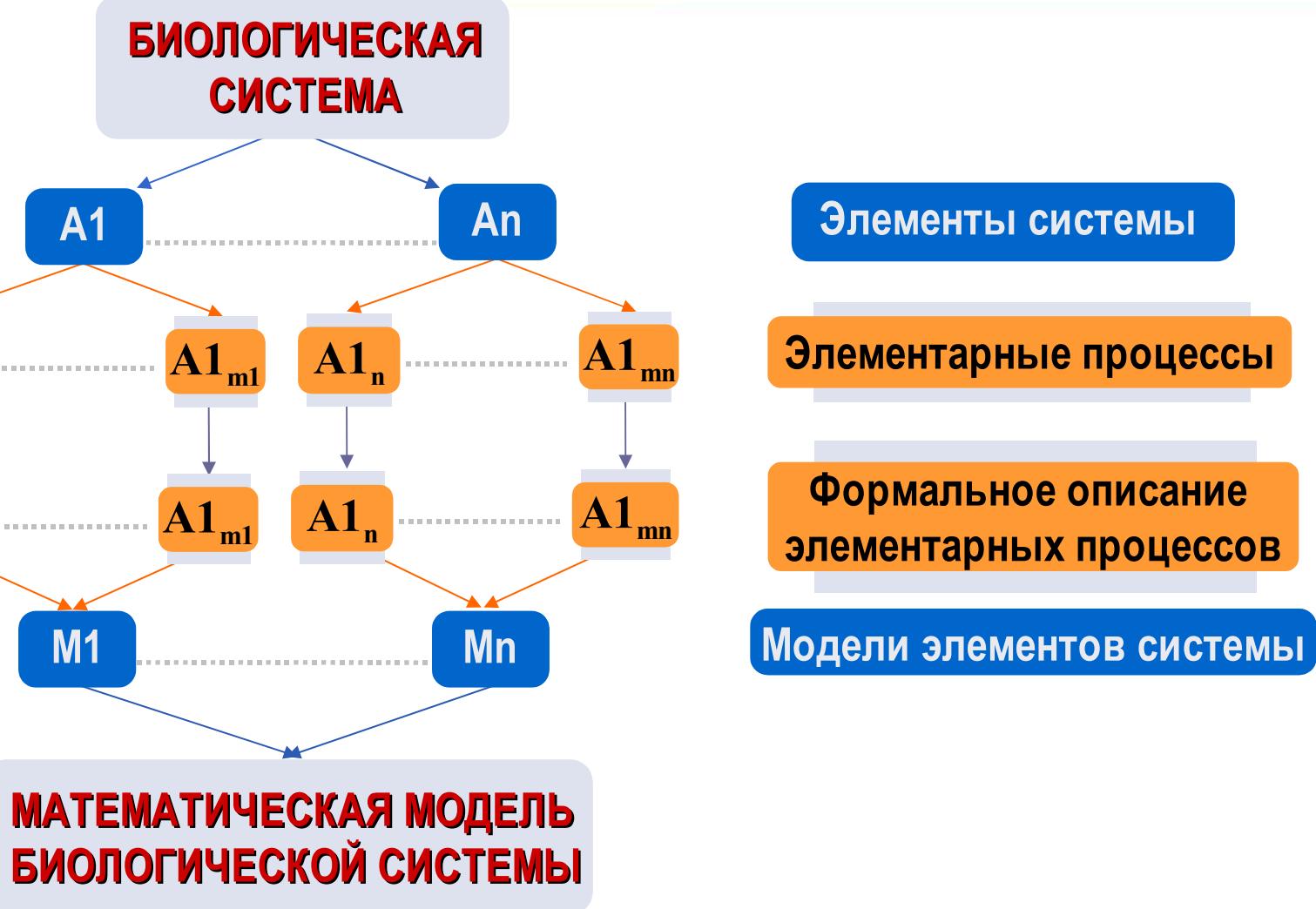
Обобщенный химико-кинетический метод моделирования (ОХКММ)

В рамках ОХКММ развиты средства, позволяющие эффективно моделировать такие структурные и функциональные особенности генных сетей, как:

- Цис- и транс- эффекты;
- Взаимное расположение и ориентацию генов и их регуляторных сайтов;
- Матричность фундаментальных процессов (репликация, транскрипция, трансляция);
- Полиплоидия;
- Генетические перестройки, рекомбинация, кроссинговер;
- Мутации;
- Многокомpartmentость;
- И т.д.



Методология моделирования в ОХКММ





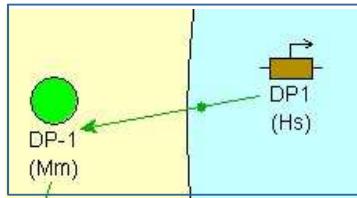
Формализация в ОХКММ

Формализация в рамках ОХКММ осуществляется на основе блочного принципа

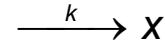
- Описание элементарных подсистем проводится в терминах элементарных процессов
- Основной принцип локальная независимость процессов
- Элементарные процессы описываются на основе формальных блоков
- Формальные блоки однозначно характеризуются (1) упорядоченным списком формальных динамических переменных (X), (2) упорядоченным списком формальных параметров (P) и (3) законом преобразования информации (F)



Примеры формального описания элементарных процессов



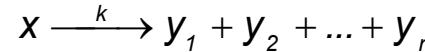
Конститутивный синтез :



$$\bar{X} = (x), \bar{P} = k, F : \frac{dx}{dt} = k \cdot x$$

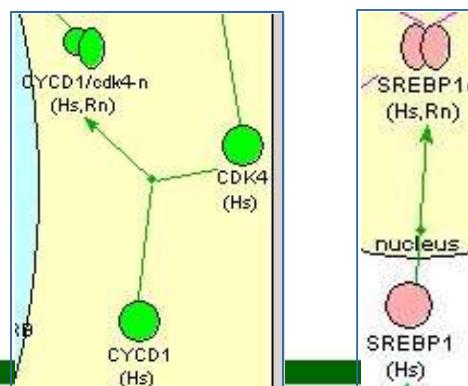


Мономолекулярная необратимая реакция:

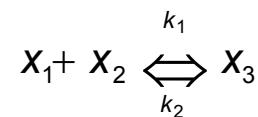


$$\bar{X} = (y_1, y_2, \dots, y_n), \bar{P} = k,$$

$$F : \frac{dx}{dt} = -\frac{dy_1}{dt} = -\frac{dy_2}{dt} = \dots = -\frac{dy_n}{dt} = -k \cdot x, n \geq 0$$



Бимолекулярная обратимая реакция :

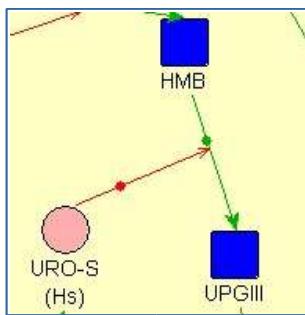
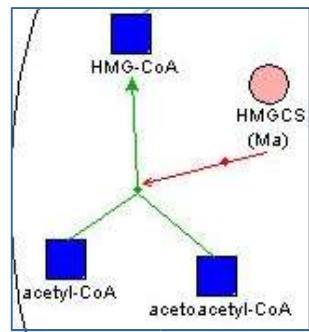


$$\bar{X} = (x_1, x_2, x_3), \bar{P} = (k_1, k_2),$$

$$F : \frac{dx_1}{dt} = \frac{dx_2}{dt} = -\frac{dx_3}{dt} = k_2 \cdot x_3 - k_1 \cdot x_1 \cdot x_2$$



Примеры формального описания элементарных процессов



Обобщенная реакция Михаэлиса – Ментен:

$$\bar{X} = (x_1, x_2, \dots, x_m, y_1, y_2, \dots, y_n),$$

$$\bar{P} = (m, k_i, k_d, a_1, \dots, a_m, b_1, \dots, b_n),$$

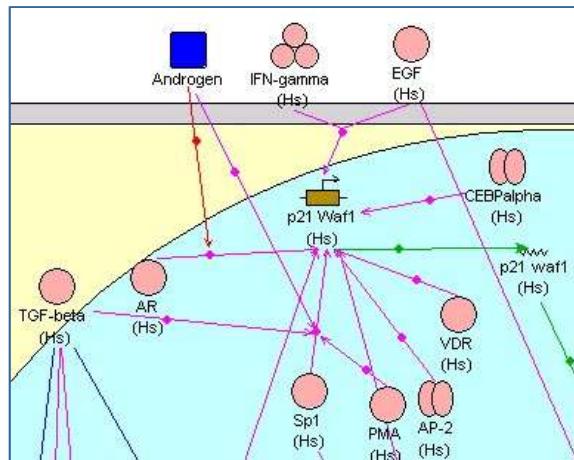
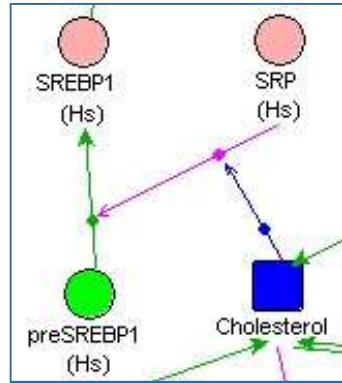
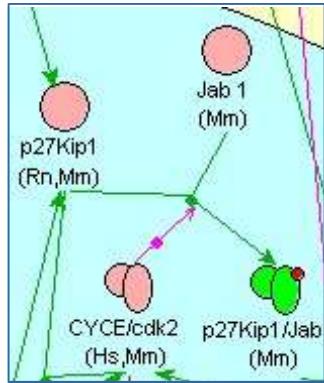
$$F: \frac{dx_j}{dt} = -a_j \cdot k_i \cdot Z, j = 1, \dots, m, m \geq 2,$$

$$\frac{dy_l}{dt} = b_l \cdot k_i \cdot Z, l = 1, \dots, n, n \geq 0$$

$$Z = \frac{k_d \cdot x_1 \cdot \dots \cdot x_m}{(k_d + x_1) \cdot \dots \cdot (k_d + x_m) - x_1 \cdot \dots \cdot x_m}$$



Примеры формального описания элементарных процессов



Регуляторные процессы:

$$-\frac{dX}{dt} = \frac{dY}{dt} = k \cdot \frac{k_1 + \sum_{i=1}^{n_1} \alpha_{1i} A_i^{h_{1i}}}{k_2 + \sum_{i=1}^{m_1} \beta_{i1} I_i^{Y_i} + \sum_{i=1}^{n_1} \alpha_{2i} A_i^{h_{2i}}}.$$

$$\prod_{j=1}^{n_2} \frac{\alpha_{0j} + \alpha_{1j} A_j^{h_{1j}}}{1 + \alpha_{2j} A_j^{h_{2j}}} \cdot \prod_{j=1}^{m_2} \frac{1}{1 + \beta_j I_j^{Y_j}} \cdot X,$$

n_1 – активаторов (A_i) и m_1 – ингибиторов (I_i)
по аддитивному механизму;

n_2 – активаторов (A_j) и m_2 – ингибиторов (I_j)
по мультипликативному механизму



Примеры формального описания элементарных процессов

Суммирование: $x = x_1 + \dots + x_n : \bar{X} = (x, x_1, \dots, x_n), \bar{F} : x = x_1 + \dots + x_n.$

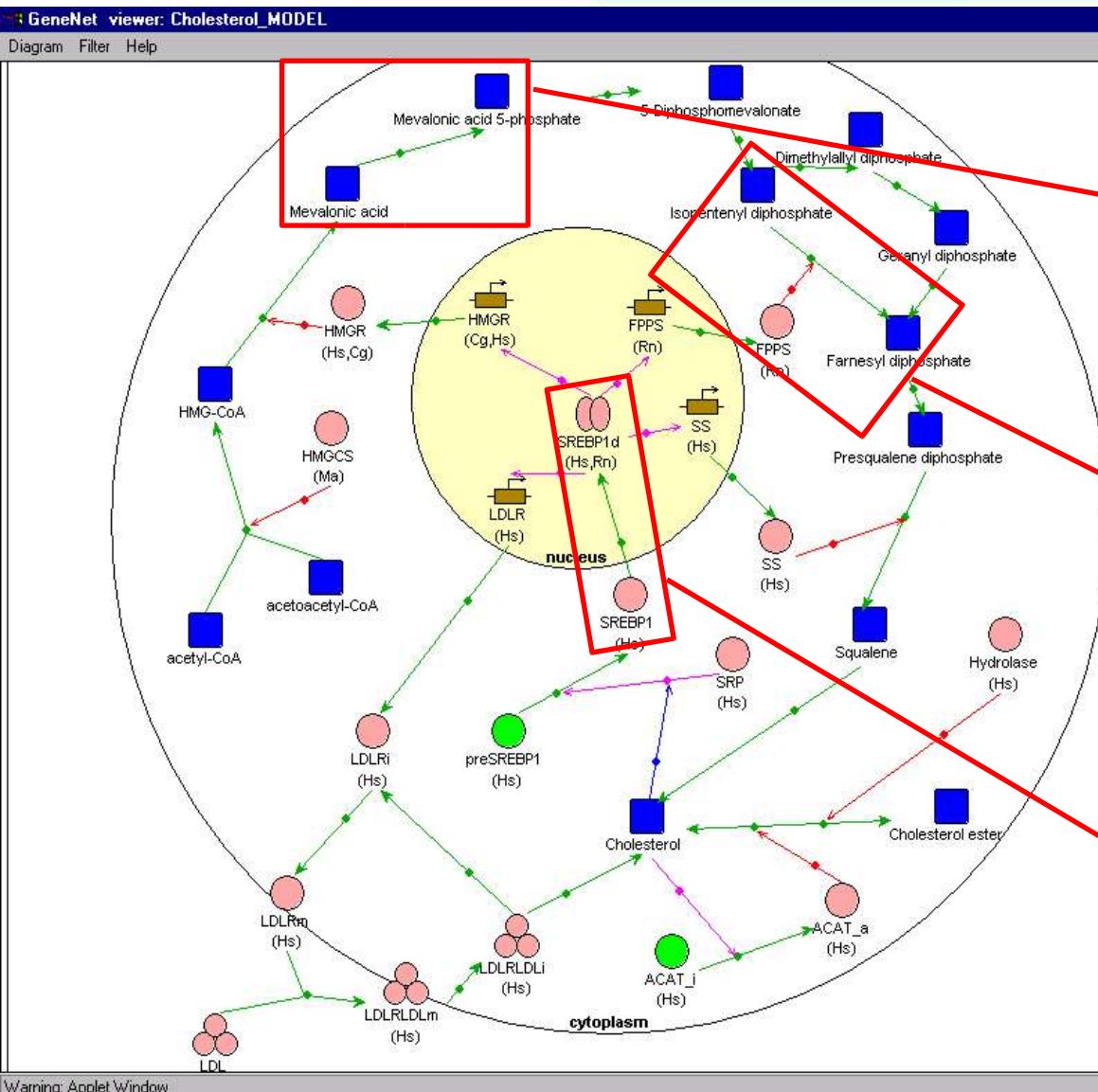
Пороговый автомат: $\bar{X} = (x, x_1, x_2), \bar{F} : x = x_1, \text{ если } x_1 \geq x_2.$

Стochasticкий автомат: $\bar{X} = (x, x_1, y_1, y_2) \bar{F} : x = x_1, \text{ если } y_1 \geq y_2,$
где y_1 – случайная величина, равномерно распределенная на отрезке $[0, 1]$,
 y_2 – число из интервала $[0, 1]$.

В общем случае формальные блоки могут быть любой природы



Примеры формального описания элементарных процессов



Monomolecular reaction:

$$\left. \begin{array}{l} A \xrightarrow{k} B, \quad a = [A], b = [B] \\ -da/db = db/dt = ka \end{array} \right\}$$

Enzymatic reaction:

$$\left. \begin{array}{l} E + S \xrightleftharpoons[k_{-1}]{k_1} ES \xrightarrow{k_2} E + P, \\ e = [E], s = [S], e_a = [ES], p = [P] \\ ds/dt = -dp/dt = -k_3 e_o s / (K_m + s), \\ e_o = e + e_a, \quad K_m = (k_{-1} + k_2)/k_1 \end{array} \right\}$$

Bimolecular reaction:

$$\left. \begin{array}{l} A + B \xrightleftharpoons[k_2]{k_1} C, \quad a = [A], b = [B], c = [C] \\ da/dt = db/dt = -dc/dt = -k_1 ab + k_2 c \end{array} \right\}$$



Примеры формального описания элементарных процессов

$d(\text{Cholin})/dt =$

$$+ \{ KISBin * \text{Squalene} \} - \{ KISBestr * \text{Cholin} * \text{ACAT} / (KMSBestr + \text{Cholin} + \text{ACAT}) \} + \{ KISBhydr * \text{Cholestr} * \text{Hydrolas} / (KMSBhydr + \text{Cholester} + \text{Hydrolas}) \} + \{ 1785 * KISBfree * \text{LDLRLDLi} \} + \{ -Kin1prot * \text{Cholin} * \text{SRPprot} + Kin2prot * \text{SRP-Chol} \} + \{ Ksrpcdeg * \text{SRP-Chol} \} - \{ Kutichol * \text{Cholin} \}$$

$d(\text{LDL}_\text{R})/dt =$

$$+ \{ -Kclaw1 * \text{LDL}_\text{R} * \text{LDL}_\text{R} + Kclaw2 * \text{LDLRLDLo} \} + \{ Kfre_\text{LDL} * t * \text{LDLRin} \} - \{ KLDLRdeg * \text{LDL}_\text{R} \}$$

$d(\text{LDL})/dt =$

$$+ \{ -Kclaw1 * \text{LDL}_\text{R} * \text{LDL} + Kclaw2 * \text{LDLRLDLo} \} + \{ KsynLDL * \text{PLDL} \} - \{ KutiLDL * \text{LDL} \}$$

$d(\text{LDLRin})/dt =$

$$+ \{ KISBldlr * \text{LDLRgene} * \text{SREBPdim} / (KMSBldlr + \text{LDLRgene} + \text{SREBPdim}) \} - \{ Kfre_\text{LDLRin} * t * \text{LDLRin} \} + \{ Kfree * \text{LDLRLDLi} \} - \{ KdegLRin * \text{LDLRin} \}$$

Константы мономолекулярных
реакций, с⁻¹

$$KISBfree = 1 * 10^{-2}, Ksrpcdeg = 2.5 * 10^{-5}, Kfre_\text{LDL} = 1 * 10^{-2}, KLDLRdeg = 1.93 * 10^{-6}, KsynLDL = 3.32 * 10^{-2}, KutiLDL = 3.21 * 10^{-6}, KdegLRin = 1.93 * 10^{-6}$$

Константы бимолекулярных
реакций

$$Kclaw1 = 3.3 * 10^{-1} \text{ л} * \text{ммоль}^{-1} * \text{с}^{-1} \quad Kclaw2 = 1 * 10^{-3} \text{ с}^{-1}$$

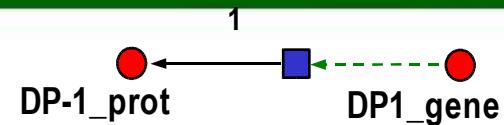
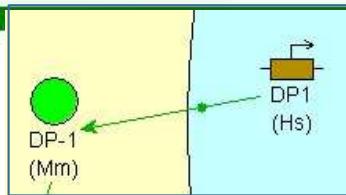
Константы ферментативных
реакций

Константы оборота ферmenta, с ⁻¹	Константы Михаэлиса, ммоль/л
$KISBestr = 1 * 10^{+02}, KISBhydr = 1.9 * 10^{+03}, KISBldlr = 1 * 10^{-1}$	$KMSBestr = 6.67 * 10^{-03}, KMSBhydr = 7.6 * 10^{-01}, KMSBldlr = 3.3 * 10^{-05}$

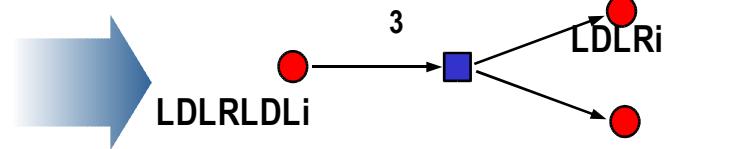
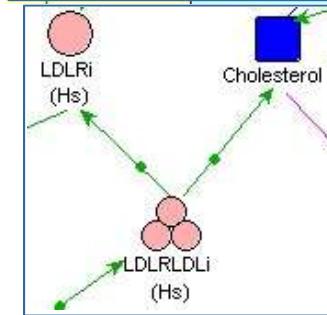


Представление генных сетей в виде двудольных орграфов

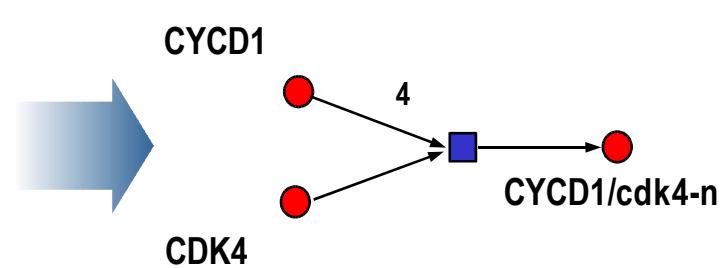
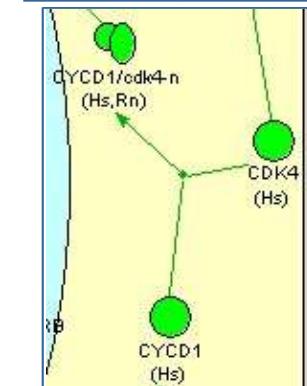
Конститутивный синтез



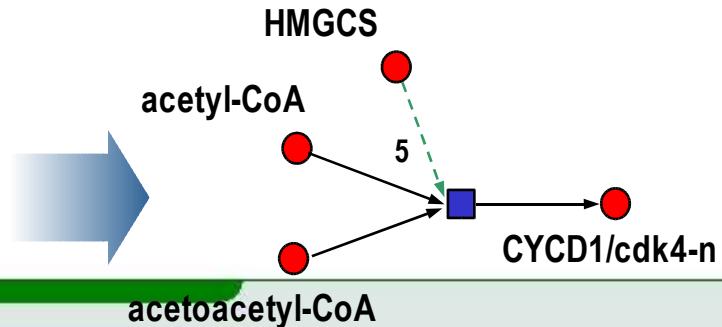
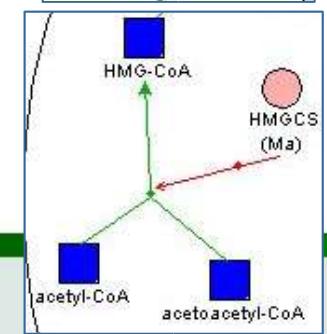
Мономолекулярная реакция



Бимолекулярная реакция

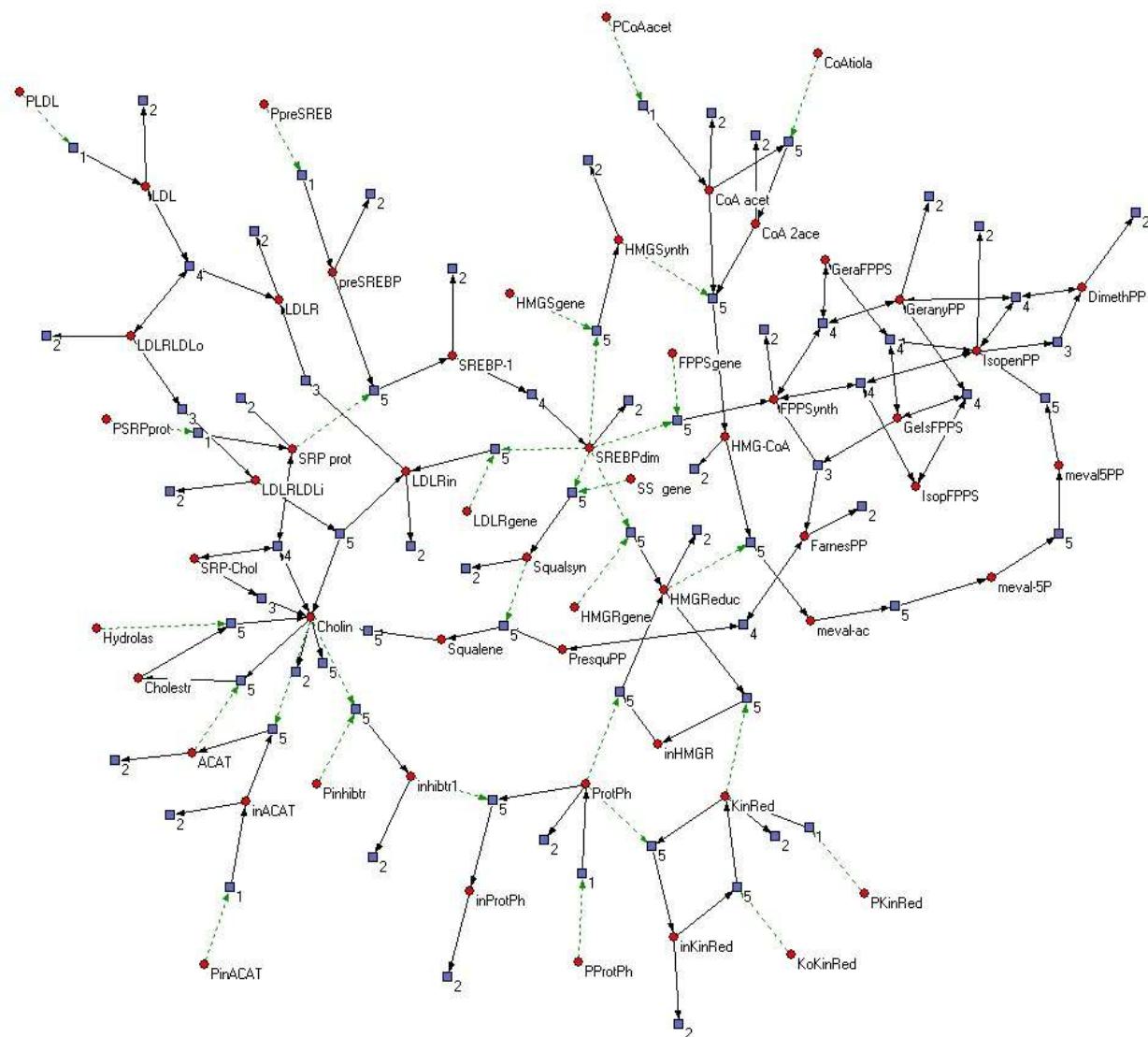


Ферментативная реакция





Представление генных сетей в виде двудольных орграфов



**Двудольный граф
компьютерной модели
гаплоидной генной сети
регуляции биосинтеза
холестерина в клетке**

- 39 динамических переменных
- 74 процесса

- - участие в процессе с ненулевой стехиометрией
- - участие в процессе с нулевой стехиометрией
- 1 - конститutивный синтез
- 2 - утилизация, либо деградация
- 3 - мономолекулярная реакция
- 4 - бимолекулярная реакция
- 5 - обобщенная реакция Михаэлиса-Ментен



Построение компьютерных динамических моделей метаболических путей

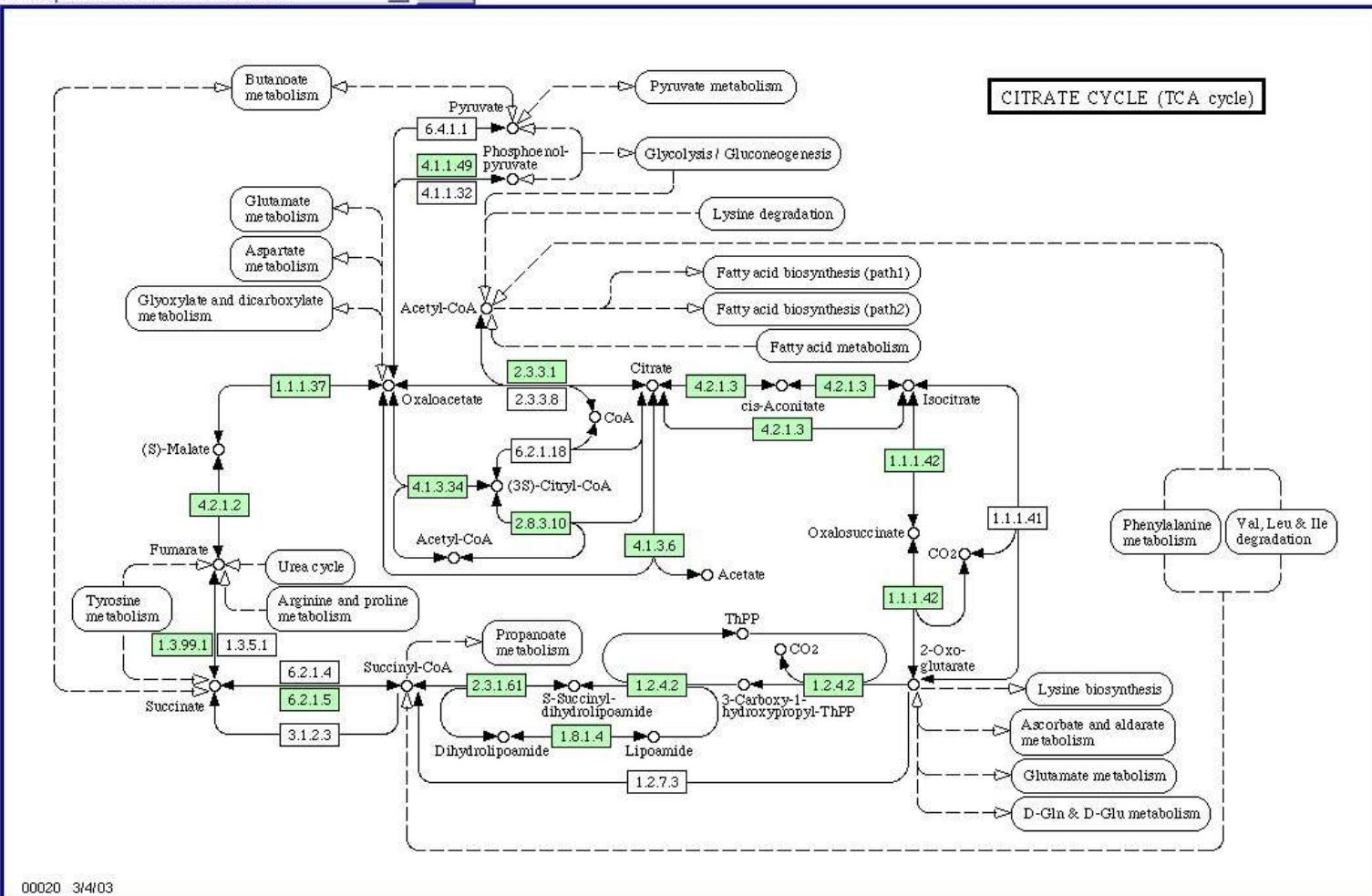
**Автоматическое конструирование компьютерных динамических моделей метаболических путей на основе информации, представленной в базе данных KEGG
(<http://www.genome.ad.jp/>)**

Triple cycle (TCA cycle) - *Escherichia coli* K-12 MG1655

Go to: [[LinkDB search](#) | [Ortholog Table](#)]

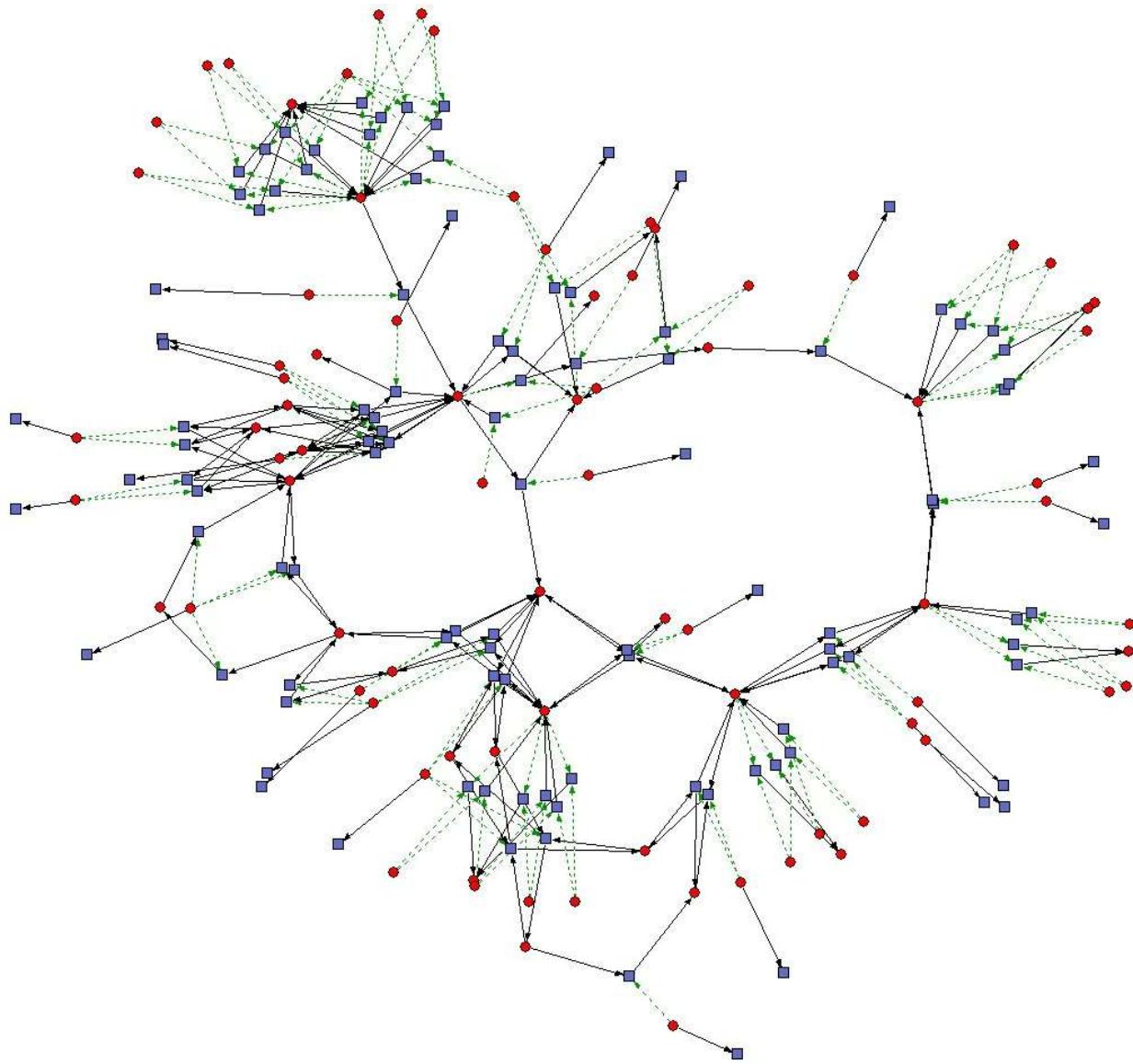
Go to: Escherichia coli K-12 MG1655

Exec





Построение компьютерных динамических моделей метаболических путей



*Двудольный граф
компьютерной модели
цикла трикарбоновых кислот
Escherichia coli K-12*

- - 73 динамических переменных
- - 139 процессов

- - участие в процессе с ненулевой стехиометрией
- - - → - участие в процессе с нулевой стехиометрией



Построение компьютерных динамических моделей метаболических путей

В базе данных KEGG представлены 101 метаболических путей, со спецификацией на на различные организмы

Carbohydrate Metabolism:

- Glycolysis / Gluconeogenesis
- Citrate cycle (TCA cycle)
- Pentose phosphate pathway
- Pentose and glucuronate interconversions
- Fructose and mannose metabolism
- Galactose metabolism
- Ascorbate and aldarate metabolism
- Pyruvate metabolism
- Glyoxylate and dicarboxylate metabolism
- Propanoate metabolism
- Butanoate metabolism
- C5-Branched dibasic acid metabolism

Amino Acid Metabolism

- Glutamate metabolism
- Alanine and aspartate metabolism
- Glycine, serine and threonine metabolism
- Methionine metabolism
- Cysteine metabolism
- Valine, leucine and isoleucine degradation
- Valine, leucine and isoleucine biosynthesis
- Lysine biosynthesis
- Lysine degradation
- Arginine and proline metabolism
- Histidine metabolism
- Tyrosine metabolism
- Phenylalanine metabolism
- Tryptophan metabolism
- Phenylalanine, tyrosine and tryptophan biosynthesis
- Urea cycle and metabolism of amino groups

Energy Metabolism:

- Carbon fixation
- Reductive carboxylate cycle (CO₂ fixation)
- Methane metabolism
- Nitrogen metabolism
- Sulfur metabolism

Lipid Metabolism:

- Fatty acid biosynthesis (path 1)
- Fatty acid biosynthesis (path 2)
- Fatty acid metabolism
- Sterol biosynthesis
- Bile acid biosynthesis
- Androgen and estrogen metabolism

Metabolism of Cofactors and Vitamins:

- Thiamine metabolism
- Riboflavin metabolism
- Vitamin B6 metabolism
- Nicotinate and nicotinamide metabolism
- Pantothenate and CoA biosynthesis
- Biotin metabolism
- Folate biosynthesis
- One carbon pool by folate
- Porphyrin and chlorophyll metabolism
- Ubiquinone biosynthesis

Nucleotide Metabolism:

- Purine metabolism
- Pyrimidine metabolism
- Nucleotide sugars metabolism

Metabolism of Other Amino Acids:

- beta-Alanine metabolism
- Taurine and hypotaurine metabolism
- Aminophosphonate metabolism
- Selenoamino acid metabolism
- Cyanoamino acid metabolism
- D-Glutamine and D-glutamate metabolism
- D-Arginine and D-ornithine metabolism
- D-Alanine metabolism
- Glutathione metabolism

Biodegradation of Xenobiotics

- Caprolactam degradation
- gamma-Hexachlorocyclohexane degradation
- 1,1,1-Trichloro-2,2-bis(4-chlorophenyl)ethane (DDT) degradation
- 2,4-Dichlorobenzoate degradation
- Tetrachloroethene degradation
- Styrene degradation
- 1,4-Dichlorobenzene degradation
- Nitrobenzene degradation
- Ethylbenzene degradation
- Fluorene degradation
- Carbazole degradation
- Benzoate degradation via CoA ligation
- Benzoate degradation via hydroxylation

Metabolism of Complex Carbohydrates:

- Starch and sucrose metabolism
- N-Glycans biosynthesis
- Aminosugars metabolism
- Lipopolysaccharide biosynthesis
- Peptidoglycan biosynthesis
- Glycosaminoglycan degradation

Metabolism of Complex Lipids:

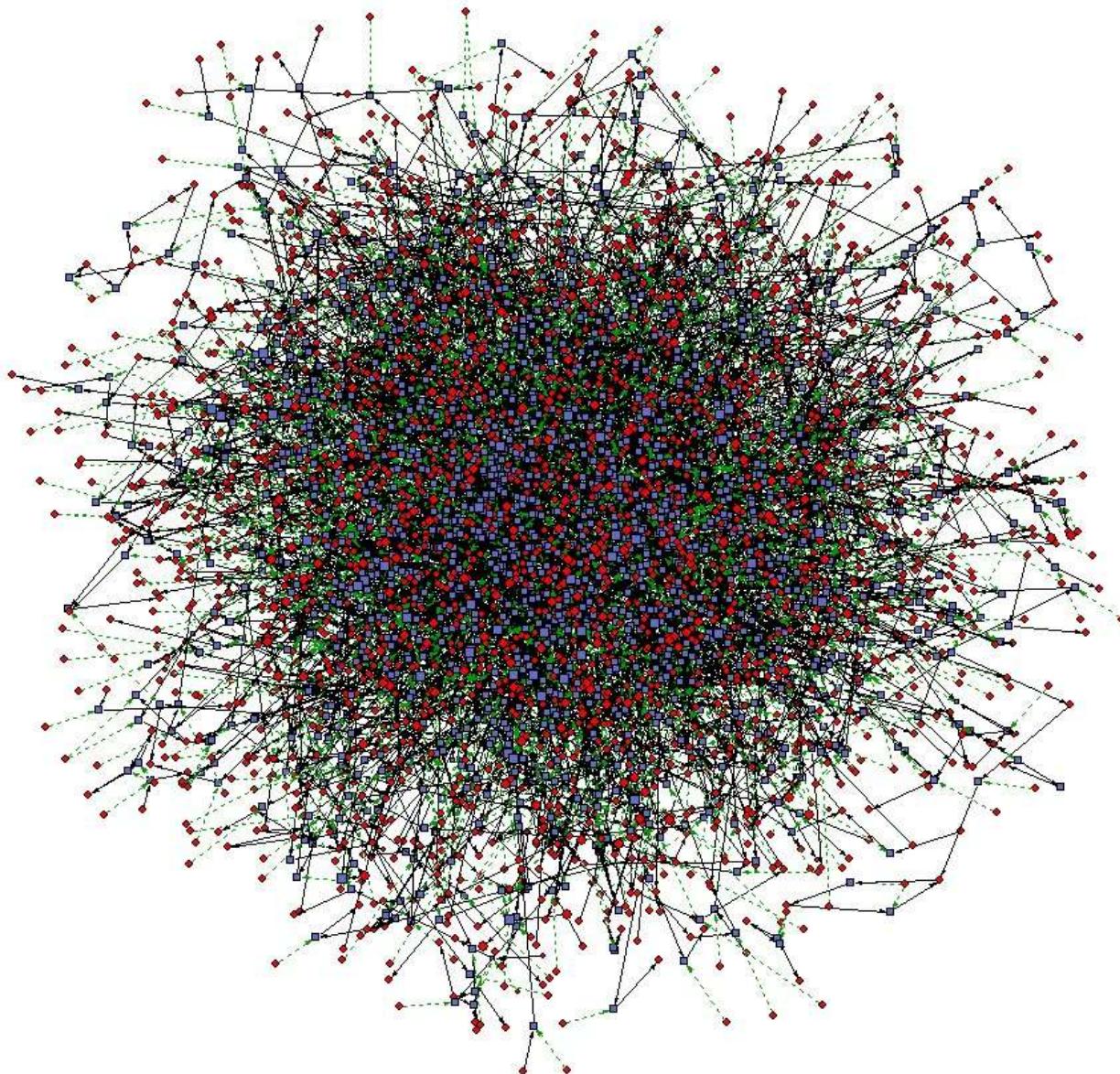
- Glycerolipid metabolism
- Inositol phosphate metabolism
- Phospholipid degradation
- Sphingoglycolipid metabolism
- Blood group glycolipid biosynthesis - neolact series
- Globoside metabolism
- Ganglioside biosynthesis

Biosynthesis of Secondary Metabolites:

- Terpenoid biosynthesis
- Flavonoids, stilbene and lignin biosynthesis
- Alkaloid biosynthesis I
- Alkaloid biosynthesis II
- Streptomycin biosynthesis
- Erythromycin biosynthesis
- Tetracycline biosynthesis



Компьютерная динамическая модель метаболизма клетки (**KEGG**)



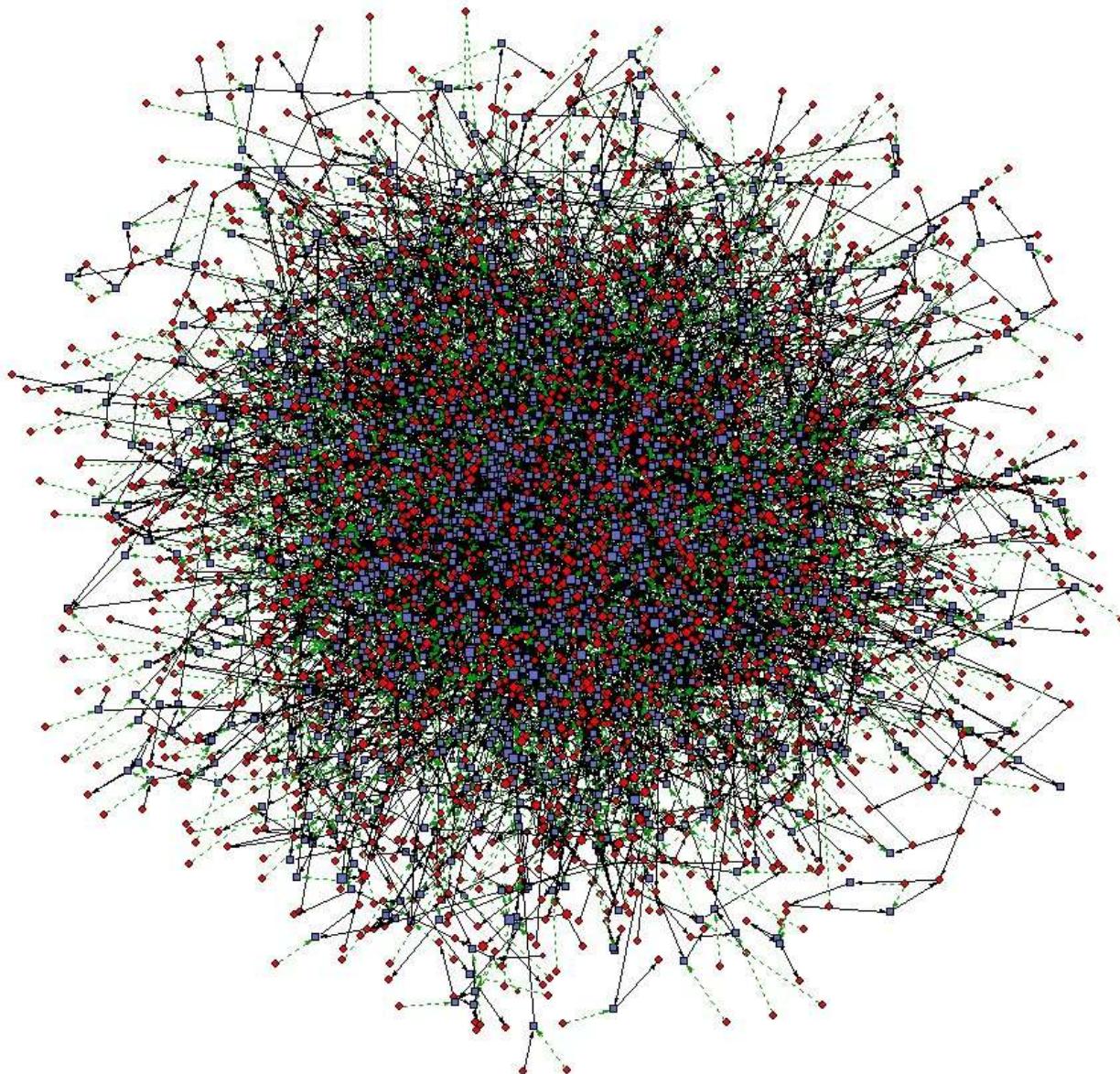
*Двудольный граф
компьютерной динамической
модели метаболизма клетки,
со спецификацией на организм
Escherichia coli K-12*

- - 4036 динамических переменных
- - 3973 процесса

- - участие в процессе с ненулевой стехиометрией
- - → - участие в процессе с нулевой стехиометрией



Компьютерная динамическая модель метаболизма клетки (**KEGG**)



*Двудольный граф
компьютерной динамической
модели метаболизма клетки,
со спецификацией на организм
Escherichia coli K-12*

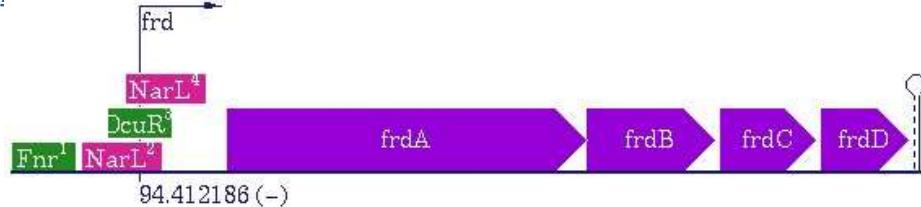
- - 4036 динамических переменных
- - 3973 процесса

- - участие в процессе с ненулевой стехиометрией
- - → - участие в процессе с нулевой стехиометрией



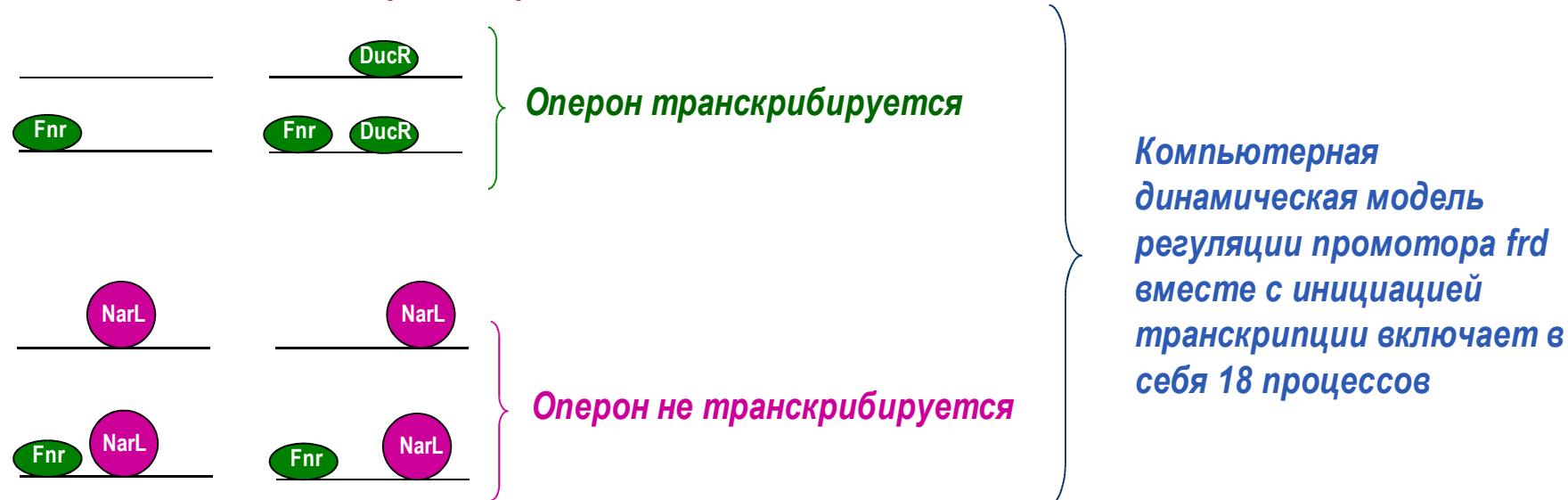
Построение компьютерных динамических моделей регуляции промоторов

Автоматическое конструирование компьютерных динамических моделей регуляции промоторов на основе информации, представленной в базе данных EcoCyc
(<http://www.ecocyc.org>)



Взаимное расположение регуляторных сайтов на промоторе *frd*

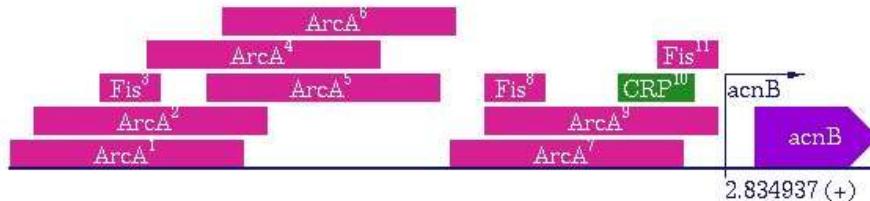
8 возможных состояния промотора:





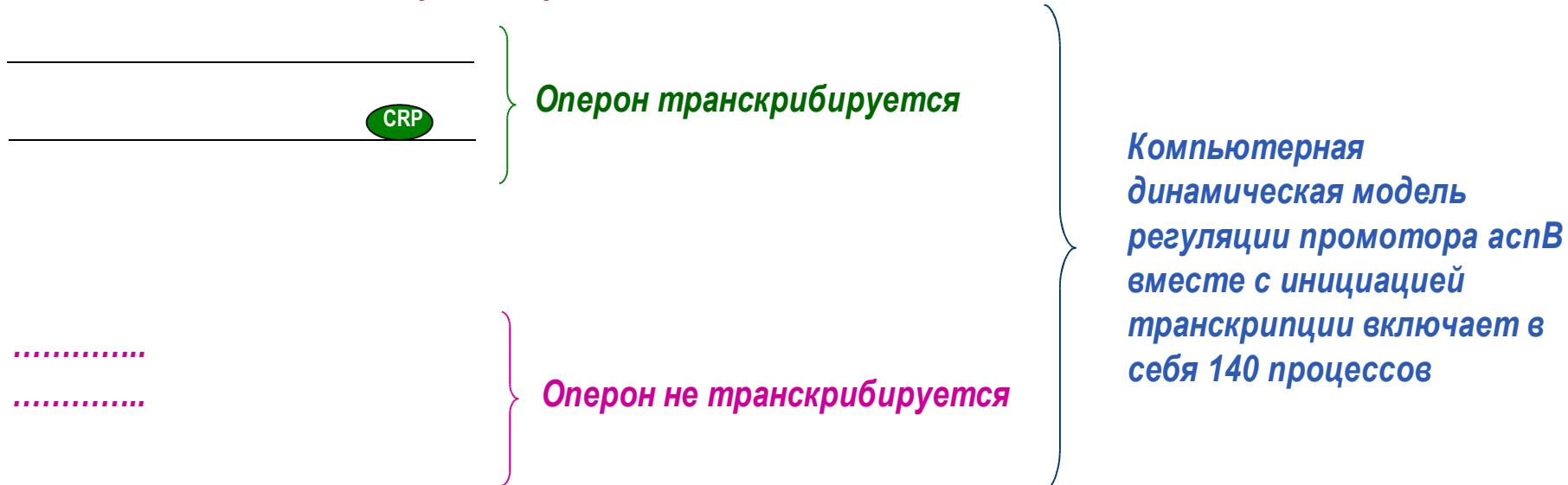
Построение компьютерных динамических моделей регуляции промоторов

Автоматическое конструирование компьютерных динамических моделей регуляции промоторов на основе информации, представленной в базе данных EcoSус
(<http://www.ecosusc.org>)



Взаимное расположение регуляторных сайтов на промоторе *acnB*

70 возможных состояния промотора:





Построение компьютерных динамических моделей формирования ферментативных комплексов

Конструирование компьютерных динамических моделей формирования ферментативных комплексов на основе информации, представленной в базе данных EcoSус

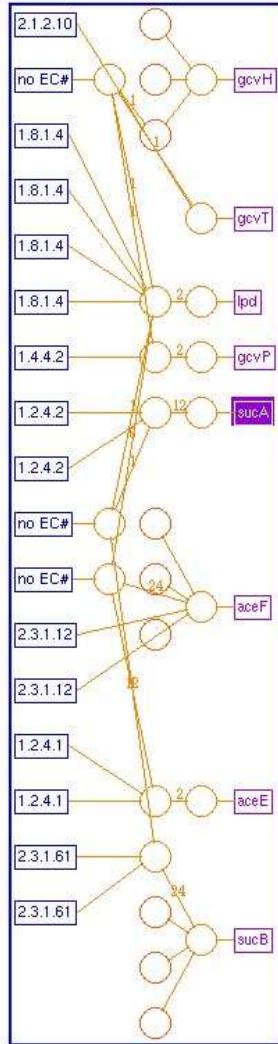
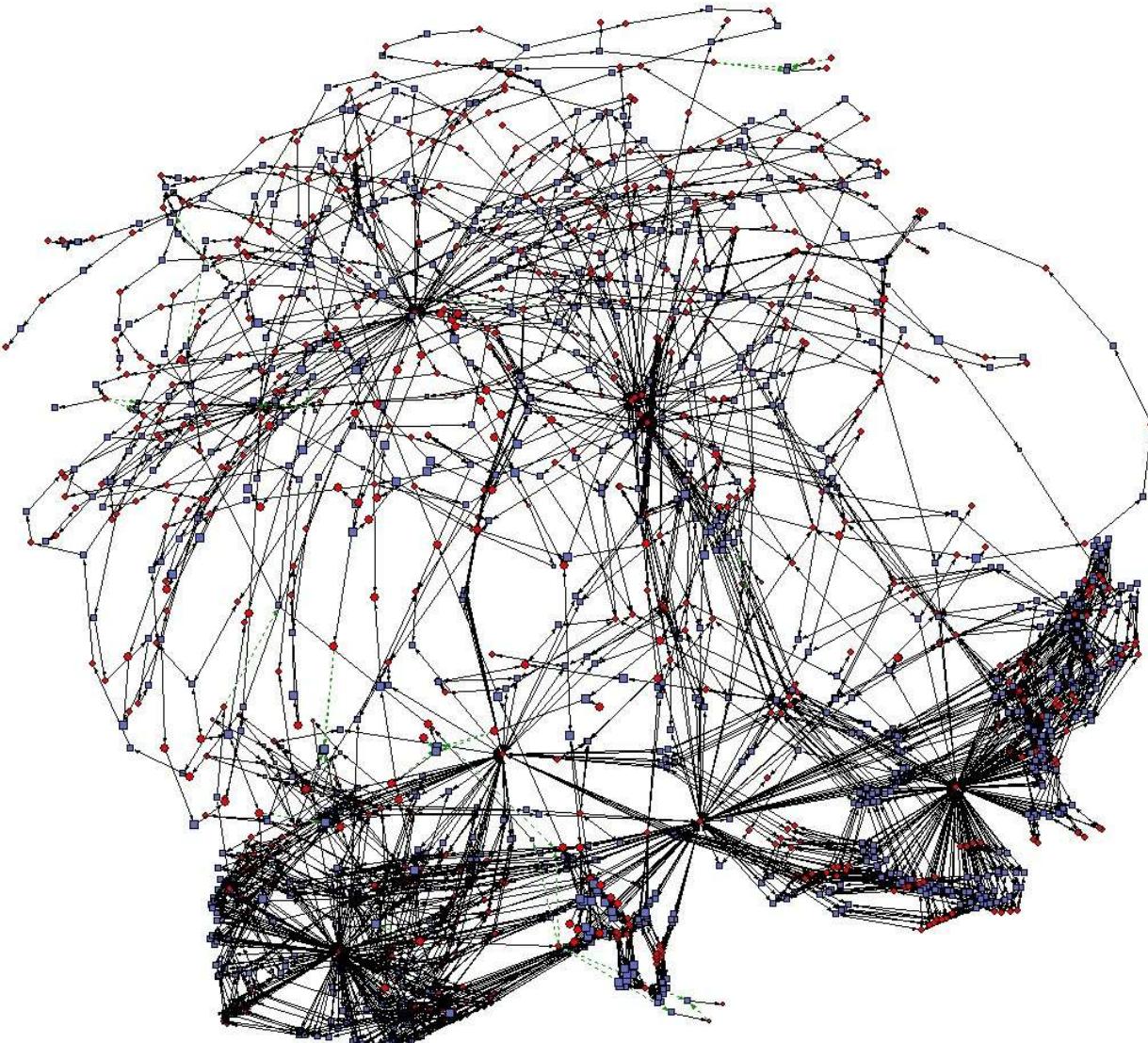


Схема демонстрирует образование различных ферментативных комплексов из белковых субъединиц, синтезируемых с генов *gcvH*, *gcvT*, *gcvP*, *lpd*, *sucA*, *sucB*, *aceF*, *aceE*

Компьютерная динамическая модель, описывающая данную схему регуляции формирования ферментативных комплексов включает в себя 104 процесса

Компьютерная динамическая модель генной регуляции цикла трикарбоновых кислот



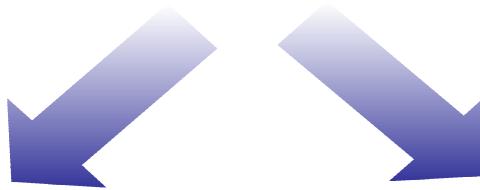
Двудольный граф
компьютерной динамической
модели генной регуляции
цикла трикарбоновых кислот
в организме *Escherichia coli*

● - 757 динамических
переменных
■ - 1882 процесса

→ - участие в процессе с
ненулевой стехиометрией
- - → - участие в процессе с
нулевой стехиометрией



Проблемы моделирования



Математические

(возникают на стадии построения, расчета и адаптации моделей)

- решение систем дифференциальных уравнений
- модели в основном состоят частично из системы ДУ и дискретной части
- решение обратной задачи
- и т.д.

Методологические

(возникают на стадии работы с моделями)

- исследование устойчивости функционирования систем
- решение задач управления
- исследование эволюции моделируемых биологических систем
- и т.д.



Верификация параметров математических моделей динамики генных сетей

Основные проблемы:

- Пробелы в знании структурно-функциональной организации некоторых частей генных сетей (посредники, механизмы протекания процессов и т.д.);
- Нехватка количественных данных, пригодных для адаптации конкретной модели;
- В базах данных в основном накапливаются качественная описательная информация (информация о структурно-функциональной организации генной сети, механизмы различных процессов, организация регуляторных участков генов и т.д.), а также количественная информация статического характера (константы ферментативных реакций, массы белков, длины нуклеотидных последовательностей и т.д.);
- Рассредоточение кинетических данных в огромном количестве научных публикаций (слабоструктурированность данных);
- Гетерогенность данных (наборы экспериментальных динамических данных получаются в своих конкретных условиях, при разных экспериментальных воздействиях в различные моменты времени и т.д.).



Верификация параметров математических моделей динамики генных сетей



- константы многих ферментативных реакций доступны в базах данных

WIT - <http://www-unix.mcs.anl.gov/compbio/>

BRENDA - [http://brenda\(bc.uni-koeln.de/](http://brenda(bc.uni-koeln.de/)

и т.д.

- константы макропроцессов (репликация, транскрипция, трансляция и т.д.)

$V_{\text{трансляции}} \sim 0.1 \text{ сек}^{-1}$

- константы, характеризующие взаимодействие генной сети с ее клеточным и организменным окружением

- равновесные концентрации;

- времена жизни или полураспада компонентов системы;

- интегральные характеристики;

- т.д.



Верификация параметров математических моделей динамики генных сетей

- Математическая модель генной сети характеризуется набором констант k_1, \dots, k_m в системе дифференциальных уравнений
- Как правило, экспериментально измеренные значения известны только для ограниченного числа констант
- Значения остальных констант определяются на основе численного эксперимента



Верификация параметров математических моделей динамики генных сетей

Общая информация

Type
Name_brief
Name_full
Organism
Cells
StageCellDifferentiation
OrganismStatus
ExpressionDetectionDevice

Reference

Peritoneal macrophages were preincubated with LPS for the indicated periods, then washed with PBS twice, and then stimulated with LPS.

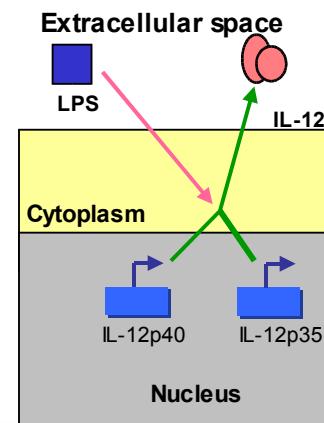
Protein
IL-12
Interleukin-12
Mouse – mus musculus
Peritoneal macrophages
Terminally differentiated
Norm
Relative protein level

Динамические данные

ID	PointOf Measurement	Time		Product	
		Units	Concentratio n%	Standart Error	
0	0	hours	0.0	—	
1	6	hours	30.6	—	
2	7	hours	4.6	—	
3	9	hours	4.0	—	
4	12	hours	2.8	—	
5	18	hours	2.7	—	
6	30	hours	0.2	—	
7	30	hours	0.3	—	
8	30	hours	1.9	—	

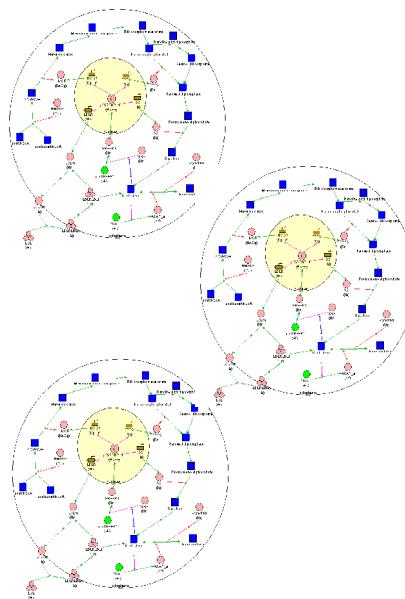
Экспериментальные условия

ID	External Factor	Specification	Time		Concentratio		Unit
			(initial point)	(exposure time)	Units	Value	
1	LPS	Escherichia coli O55:B5	0	6	hours	10	ng/ml
2	LPS	Escherichia coli O55:B5	0	1	hours	100	ng/ml
2	LPS	Escherichia coli O55:B5	1	7	hours	10	ng/ml
3

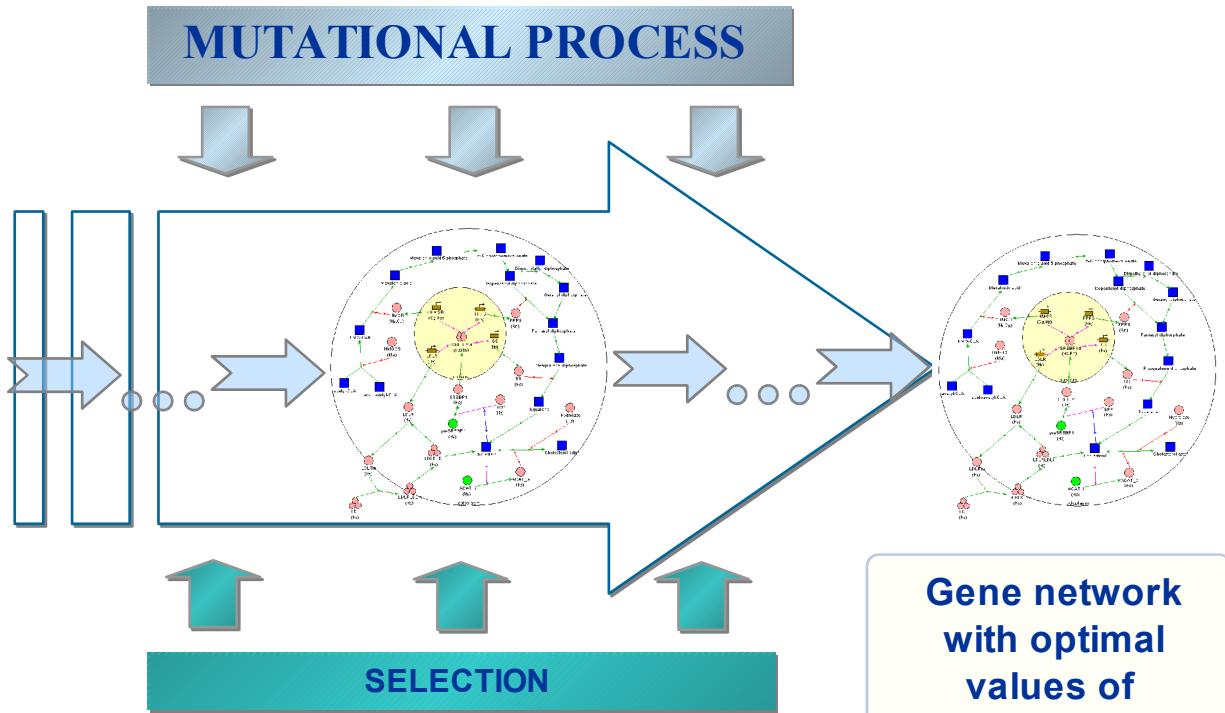




Верификация параметров математических моделей динамики генных сетей (эволюционный алгоритм)



Initial population of gene networks:
not all the constants k_1, \dots, k_m are known



Optimized functional: $W=1/F$, where

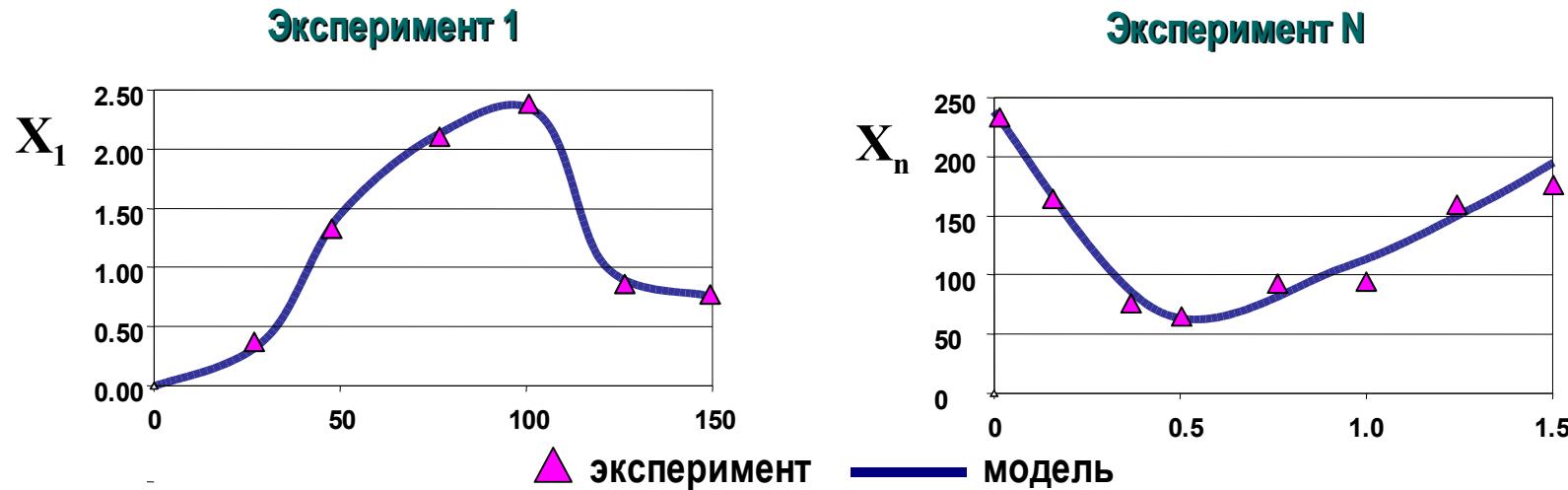
W – is an adaptability of an organism in particular environmental conditions;

F – is the measure of deviation of calculated characteristics from the ordered ones

Gene network with optimal values of coefficients
 k_1, \dots, k_m



Верификация параметров математических моделей динамики генных сетей



Проводится поиск значений констант k_1, \dots, k_m , которые обеспечивают максимальное соответствие между расчетным и наблюдаемым динамическим поведением генной сети для большого количества экспериментов

$$F(k_1, \dots, k_m) = \sum_{i,j(i)} \sqrt{(X_{\text{эксп.}}^{ij} - X_{\text{теор.}(k_1 \dots k_m)}^{ij})^2}$$

$$F(k_1, \dots, k_m) = \sum_{i,j(i)} \left[\frac{X_{\text{эксп.}}^{ij}}{X_{\text{теор.}(k_1 \dots k_m)}^{ij}} + \frac{X_{\text{теор.}(k_1 \dots k_m)}^{ij}}{X_{\text{эксп.}}^{ij}} - 2 \right]$$



Верификация параметров математических моделей технология построения сценариев

