Математическое моделирование динамики генных сетей

А.В. Ратушный

Кафедра информационной биологии ФЕННГУ

Цели лекции

- Рассмотреть математические модели конкретных генных сетей:
 - генная сеть, контролирующая гомеостаз холестерина в клетке (отрицательные обратные связи)
 - генная сеть регуляции дифференцировки эритроидной клетки (положительные обратные связи)
- Построение математических моделей, возникновение различных моделей одной и той же биологической системы
- Проблемы математического моделирования генных сетей
- Верификация параметров математических моделей генных сетей (решение обратной задачи)
- Моделирование динамики генных сетей при разнообразных воздействиях на них в норме и при различных мутациях
- Исследование «мутационных портретов» генных сетей
- Моделирование и анализ эволюции генных сетей
- Применение математических моделей генных сетей на практике и для исследования фундаментальных проблем

Генная сеть, контролирующая гомеостаз холестерина в клетке

- 1 - gene;
- 2 - active protein;
- 3 - inactive protein;
- 4 - multimeric complex;
- 5 - nonprotein substance;
- 6 - direct reaction;
- 7 - indirect reaction;
- 8 - regulatory event (switch on);
- 9 - regulatory event (increase);
- 10 - regulatory event (decrease)

Cells maintain cholesterol concentration in balance by sensing its level in the endoplasmic reticulum and transducing this information into the expression of multiple homeostatic genes (Steck T.L. and Lange Y., Dev Cell. 2002 Sep;3(3):306-8)

GeneNet Database:

<http://www.mgs.bionet.nsc.ru/mgs/systems/genenet>

Генная сеть регуляции дифференцировки эритроидной клетки

Система регуляции поступления ионов Fe^{3+} в эритроидную клетку и контролирующая функция белка регулируемого железом IRP (iron regulating protein). еALAS – эритроидная 5-аминолевулинат синтетаза

Формальное описание элементарных процессов в генных сетях

Monomolecular reaction:

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

Enzymatic reaction:

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

Bimolecular reaction:

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

Фрагмент системы дифференциальных уравнений:

математическая модель биосинтеза холестерина в клетке

.....

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

$$d(\text{LDLR} \text{____})/dt = + \{-Kclaw1 * \text{LDLR} \text{____} * \text{LDL} \text{____} + Kclaw2 * \text{LDLRLDLi}\} + \{Kfre \text{____} * \text{LDLRin}\} - \{KLDLRdeg * \text{LDLR} \text{____}\}$$

$$d(\text{LDL} \text{____})/dt = + \{-Kclaw1 * \text{LDLR} \text{____} * \text{LDL} \text{____} + Kclaw2 * \text{LDLRLDLi}\} + \{KsynLDL * \text{PLDL} \text{____}\} - \{KutiLDL * \text{LDL} \text{____}\}$$

$$d(\text{LDLRin})/dt = + \{KISBldlr * \text{LDLRgene} * \text{SREBPdim} / (KMSBldlr + \text{LDLRgene} + \text{SREBPdim})\} - \{Kfre \text{____} * \text{LDLRin}\} + \{Kfree * \text{LDLRLDLi}\} - \{KdegLRin * \text{LDLRin}\}$$

.....

Константы мономолекулярных реакций, с ⁻¹	$KISBfree = 1 * 10^{-02}$, $Ksrpcdeg = 2.5 * 10^{-05}$, $Kfre \text{____} = 1 * 10^{-02}$, $KLDLRdeg = 1.93 * 10^{-6}$, $KsynLDL = 3.32 * 10^{-02}$, $KutiLDL = 3.21 * 10^{-06}$, $KdegLRin = 1.93 * 10^{-06}$	
Константы бимолекулярных реакций	л*ммоль ⁻¹ *с ⁻¹	с ⁻¹
	$Kclaw1 = 3.3 * 10^{-01}$	$Kclaw2 = 1 * 10^{-03}$
Константы ферментативных реакций	Константы оборота фермента, с ⁻¹	
	$KISBestr = 1 * 10^{+02}$, $KISBhydr = 1.9 * 10^{+03}$, $KISBldlr = 1 * 10^{-1}$	$KMSBestr = 6.67 * 10^{-03}$, $KMSBhydr = 7.6 * 10^{-01}$, $KMSBldlr = 3.3 * 10^{-05}$

Математические модели генной сети, контролирующей гомеостаз холестерина в клетке

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 - конститутивный синтез
 - 2 - утилизация, либо деградация
 - 3 - мономолекулярная реакция
 - 4 - бимолекулярная реакция
 - 5 - обобщенная реакция Михаэлиса-Ментен

Математические модели генной сети, контролирующей гомеостаз холестерина в клетке

Выделение регуляторной части

Математические модели генной сети, контролирующей гомеостаз холестерина в клетке

Неоднозначность протекания метаболической цепочки реакций от ланостерина до холестерина

<http://www.chem.qmul.ac.uk>

Математические модели генной сети, контролирующей гомеостаз холестерина в клетке

*Двудольный граф модели
генной сети с портретным
описанием метаболических
путей синтеза холестерина*

● - 230 динамических
переменных

■ - 431 процесс

→ - 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/>

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	Protein
Name_brief	IL-12
Name_full	Interleukin-12
Organism	Mouse – mus musculus
Cells	
StageCellDifferentiation	Peritoneal macrophages
OrganismStatus	Terminally differentiated
ExpressionDetectionDevice	
Reference	Norm

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

Динамические данные

ID	Time		Product	
	PointOf Measurement	Units	Concentration n%	Standard 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		Concentration		
			(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

Верификация параметров математических моделей динамики генных сетей

Эксперимент 1

Эксперимент N

▲ эксперимент — модель

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

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

Верификация параметров математических моделей технология построения сценариев

Верификация параметров математических моделей технология построения сценариев

Верификация параметров моделей генной сети, контролирующей гомеостаз холестерина в клетке

Связывание ЛНП на поверхности клетки

Experiment conditions: Monolayers of cells were grown in the lipoprotein-deficient serum (LPDS) for 2 days and then were incubated in the medium containing the indicated concentration of ^{125}I -LDL at 37°C . After 5 hr, the amount of ^{125}I -LDL bound to the cell was measured.

Подавление синтеза рецепторов ЛНП

Experiment conditions: Monolayers of cells were grown in LPDS for 3 days. On day 4 of cell growth (zero time). Each dish of nonconfluent cells received 2 ml of growth medium containing $10\ \mu\text{g}$ protein/ml of unlabeled LDL. At the indicated time, the medium was replaced with 2 ml of medium containing $25\ \mu\text{g}$ protein/ml of ^{125}I -LDL. After incubation at 37°C for 2 hr, the specific heparin releasable ^{125}I radioactivity was determined.

Верификация параметров моделей генной сети, контролирующей гомеостаз холестерина в клетке

Действие возрастания концентрации ЛНП в среде на содержание свободного холестерина и его эфиров в нормальных фибробластах человека

Experiment conditions: The cells were incubated in fresh growth medium containing 10% fetal calf serum. After 3 days, the medium was replaced with 2 ml of fresh growth medium containing 5% human LPDS. After 24 hr, the medium was replaced with 2 ml of fresh growth medium containing 5% human LPDS and indicated concentration of LDL. After 24 hr, each cell monolayer was washed and harvested and the sterols content were measured.

Верификация параметров моделей генной сети регуляции дифференцировки эритроидной клетки

Dynamics of the main components on the system regulating the erythroid cell differentiation. By TfR/Tf are denoted transferrin receptors bound to transferrin (*calculations by a model*)

beginning of erythropoietin

effect

Hemoglobin is actively accumulated starting from the stage of basophilic erythroblast. Hemoglobin and other components are usually synthesized during G_1 and the beginning of S phases of DNA synthesis (Kozinets & Goldberg, 1982). The total (*transit*) time from proerythroblast to reticulocyte formation is 120 h (5 days). Two fifth of the overall cell cycle occur during G_1 and the beginning of S phases (Fedorov, 1976). Thus, the overall syntheses of internal components in a differentiating erythroid cell require approximately 50 h.

Верификация параметров моделей генной сети

регуляции дифференцировки эритроидной клетки

The number of transferrin receptors on the surface of an immature erythrocyte amounts to approximately 10^4 to 10^5 per cell; they are not detected in the progenitor cell (Kuhn, 1994).

The rates of hemoglobin synthesis in the cells at the stages of proerythroblasts and basophilic erythroblasts equal approximately 0.5 pg/h, that is, about 4.7×10^6 molecules/h for an average cell with a diameter of 10 μm ; at the stage of reticulocyte, the rate decreases fivefold, that is, to about 10^6 molecules/h (Kozinets & Goldberg, 1982). Each erythrocyte contains about 280 million hemoglobin molecules (Fedorov, 1976).

Theoretical numerical calculations according to this model demonstrate the following:

Oscillating synthesis dynamics are typical of heme, α and β globins, and the components of this molecular genetic system associated with the IRP control function and involved into Fe ions influx into the erythroid cell (Figure, b,c,d,e). This is due to the network of negative and positive feedbacks.

No excess of free heme is accumulated in the differentiating erythroid cell (Figure, f). First, it is due to control of heme biosynthesis by IRP, which binds eALAS mRNA, and eALAS is, in turn, the enzyme involved in heme synthesis at the initial stage. Second, the excess of heme has a positive effect on globin biosyntheses (Ponka, 1997). The number of transferrin receptors on the surface of an immature erythrocyte reaches approximately 10^4 per cell, whereas these receptors are initially absent in the progenitor cell (Figure, c).

The rate of hemoglobin synthesis several hours after the beginning of erythropoietin effect equals approximately 5×10^6 molecules/h; maximum (8×10^6) is observed at 31 hours; in ten hours after the genes are switching off, approximately 10^6 molecules/h. Note that after 25 hours, the rate of synthesis demonstrates oscillating dynamics. An erythrocyte contains about 280 million molecules of hemoglobin on cessation of its synthesis at 74 hours (Figure, d).

Анализ влияния мутаций на генную сеть регуляции дифференцировки эритроидной клетки

Сравнение динамики синтеза гемоглобина при нормальном (1) и мутантных (2,3) типах системы дифференцировки эритроидной клетки (расчет по модели).

2 – мутация гена **EKL** (прекращения синтеза транскрипционного фактора EKL);

3 – мутация гена **GATA-1** (снижение активности транскрипционного фактора GATA-1 на 30%).

Hb – гемоглобин.

По оси **X** – время в часах, по оси **Y** – число молекул на клетку ($1 = 10^8$ молекул).

Анализ влияния мутаций на генную сеть, контролирующей гомеостаз холестерина в клетке

Моделирование отклика генной сети на увеличенное в 2 раза поступления ЛНП в плазму крови в течении 8-ми часов при мутации, уменьшающей скорость экспрессии гена ЛНП рецепторов в 2 раза

Анализ влияния мутаций на генную сеть, контролирующей гомеостаз холестерина в клетке

Моделирование отклика генной сети на увеличенное в 2 раза поступления ЛНП в плазму крови в течении 8-ми часов при мутации, понижающей в 5 раз способность рецепторов связывать ЛНП

НОРМА

МУТАЦИЯ

Анализ влияния мутаций на генную сеть, контролирующей гомеостаз холестерина в клетке

Моделирование отклика генной сети на увеличенное в 2 раза поступления ЛНП в плазму крови в течении 8-ми часов при мутации, понижающей способность рецепторов освобождать ЛНП в эндосомах, приводящей к повышенной деградации рецептора в 10 раз

Анализ мутационного портрета генной сети, контролирующей гомеостаз холестерина в клетке

Для более полного исследования влияния мутаций на систему, с целью выявления ключевых процессов генной сети и анализа поведения биологической системы в разных патологических ситуациях, необходимо проводить исследование мутаций по всем звеньям генной сети с разной степенью выраженности.

“Мутационный портрет” генной сети – набор ее стационарных состояний и динамических характеристик, полученный при “мутационном” варьировании в заданных границах скоростей протекания каждого элементарного процесса генной сети.

Исследовалось влияние одиночных мутаций разной интенсивности по всем параметрам модели. Проведено ~2000 расчетов

Анализ мутационного портрета геной сети, контролирующей гомеостаз холестерина в клетке

Изменение содержания свободного холестерина в клетке в зависимости от мутационного изменения констант элементарных процессов в геной сети биосинтеза холестерина

(1) Константа оборота фермента SRP;

(2) Константа обратной реакции димеризации SREBP1;

(3) Константа Михаэлиса-Ментен фермента ацетоацетил КоА тиолазы;

(4) Константа оборота фермента АХАТ (ацил-КоА: холестерин ацилтрансфераза)

Анализ мутационного портрета геной сети, контролирующей гомеостаз холестерина в клетке

Геной сеть с указанием чувствительности стационарного содержания свободного холестерина в клетке к мутационным изменениям параметров

- - изменение скоростей данных процессов в значительной степени сказывается на стационарной концентрации холестерина, которая может меняться от 0 до более чем 200% относительно нормы;
- - стационарная концентрация холестерина меняется не более чем на 35% от нормы;
- - стационарная концентрация холестерина меняется не более чем на 25% от нормы.

Анализ мутационного портрета геной сети, контролирующей гомеостаз холестерина в клетке

Характеристика мутационного спектра геной сети регуляции биосинтеза холестерина в клетке

Тип мутации	Доля процессов, затрагиваемая данным типом мутации
повреждающие	~15%
слабоповреждающие	~45%
нейтральные	~40%

- (i) Наличие нелимитирующих стадий в биохимических путях рассматриваемой сети.
- (iii) Наличие двух процессов, ответственных за появление дополнительных количеств холестерина в клетке: (а) собственный биосинтез холестерина в клетке и (б) его транспорт из плазмы крови в клетку посредством ЛНП рецепторов.
- (v) Возможное шунтирование некоторых биохимических реакций в пути биосинтеза холестерина.
- (vii) Регуляция внутриклеточной концентрации холестерина по механизму отрицательной обратной связи.

Моделирование эволюции геной сети, контролирующей гомеостаз холестерина в клетке

Исходная популяция геной сетей

Оптимизируемый функционал: $W=1/(1+F)$, где W – приспособленность организма к определенным условиям внешней среды;

F – мера отклонения рассчитываемых характеристик от заданных

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

Адаптированная геной сеть к определенным условиям

Моделирование эволюции геной сети, контролирующей гомеостаз холестерина в клетке

Моделирование эволюции геной сети при переходе из нормальной среды в среду с низким уровнем поступления ДНП в организм

$$W_N=1$$

Изменение условий среды

$$V_{L1} = V_N/2$$

$$W_{L1} \sim 0.004$$

адаптация

Оптимизировался функционал, построенный на основе экспериментальных данных по нормальным фибробластам человека из работ [Goldshtein *et al.*, 1975, Goldshtein *et al.*, 1977, Brown *et al.*, 1979]

$$W_{L1} \sim 0.75$$

Изменение условий среды

$$V_{L2} = V_{L1}/2$$

$$W_{L2} \sim 0.002$$

адаптация

$$W_{L2} \sim 0.54$$

Применение динамических компьютерных моделей генных сетей

Application of computer dynamic models of gene networks

Studying and understanding the pathological processes at the cell and organism level (various diseases, syndromes, inborn and acquired errors, mutations, etc);

The model of gene network regulating cholesterol biosynthesis in the cell is applicable to:

- *Atherosclerosis;*
- *Coronary heart disease (CHD);*
- *Smith-Lemli-Opitz syndrome;*
- *Mevalonic aciduria;*
- *Desmosterolosis;*
- *CHILD syndrome;*
- *Different kinds of dysplasia;*

Identification of the genetic and biochemical defects and analysis of their effects on functions of gene networks;

Development of optimal methods for influencing systems for normalizing their functions (e.g. for medical treatment using therapeutics and so on); and

Investigating basic biological problems, such as evolution, etc.

The model of gene network regulating erythroid cell differentiation is applicable to different kinds of anemia and hypoxia:

- *Dyserythropoietic anemia;*
- *Sierophile anemia;*
- *Erythroblastic anemia;*
- *Asiderotic (iron-deficiency) anemia;*
- *and many others;*

Применение динамических компьютерных моделей генных сетей

- Математические и компьютерные модели генных сетей являются незаменимым инструментом для анализа функционирования сложных биологических систем.
- С их помощью возможно предсказывать:
 - динамику тех или иных процессов в рассматриваемых системах;
 - исследовать механизмы их протекания;
 - выявлять ключевые звенья в генных сетях;
 - анализировать влияние различных мутаций на живые системы;
 - и т.д.
- Особую ценность результаты исследования мутационных спектров представляют для выявления мишеней фармакологической регуляции в генных сетях. Действительно, величины констант реакций любой генной сети могут быть в рамках современной фармакологии изменены путем введения высокоспецифичных активаторов или ингибиторов, действующих на определенные элементарные события в генных сетях.
- На этой основе анализа воздействия различных мутаций и выявления ключевых процессов в генных сетях как объекта регуляции могут создаваться стратегии коррекции патологических состояний организма с учетом его генотип-специфических особенностей.