



Теория генных сетей

В.А. Лихошвай¹

С.И. Фадеев² С.В. Николаев²

¹ Институт цитологии и генетики СО РАН likho@bionet.nsc.ru

² Институт математики им. С.Л.Соболева СО РАН fadeev@math.nsc.ru



Генная сеть – группа генов, координированно функционирующих при выполнении молекулярных, биохимических, физиологических и других функций организма



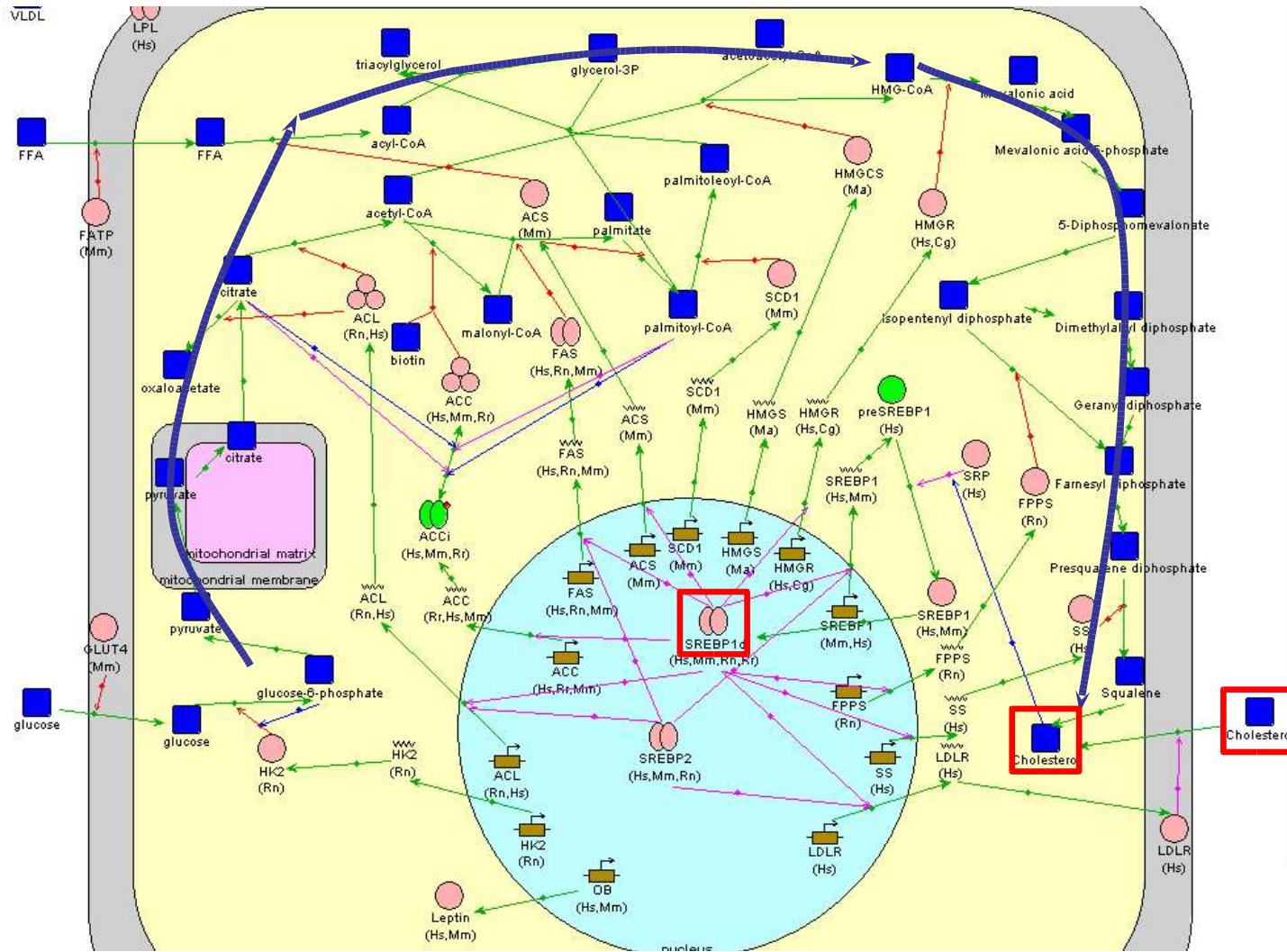
Компоненты генной сети:

2. Группа координированно функционирующих генов
3. Белки и мРНК, кодируемые этими генами
(структурные белки, ферменты, транскрипционные факторы и т.д.)
4. Пути передачи сигналов
5. Регуляторные контуры: отрицательные и положительные обратные связи
6. Внешние сигналы, гормоны, метаболиты и т.д.

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



Пример генной сети описанной в базе GeneNet

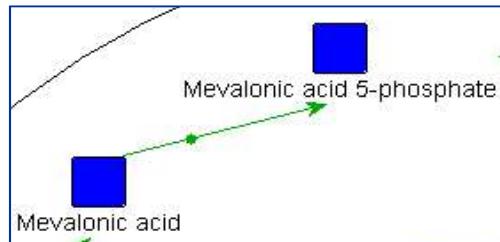


A fragment of a gene network regulating functioning of adipocyte
authors: Proscura A.L., Ignatieva E.V.

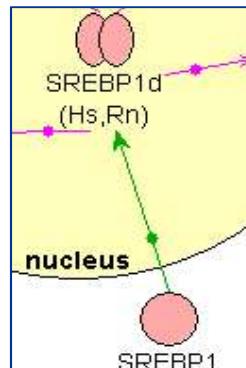


Примеры химико-кинетического описания элементарных событий в генной сети липидного метаболизма

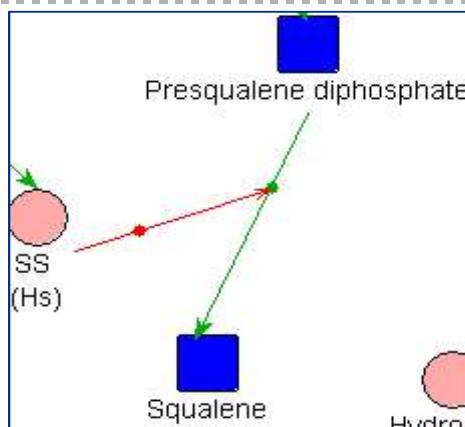
Мономолекулярная реакция



Бимолекулярная реакция



Реакция ферментативного синтеза (уравнение Михаэлиса-Ментена)

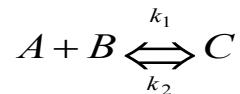




Закон суммирования скоростей элементарных реакций

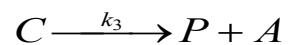


1 reaction



$$\begin{aligned} da/dt &= db/dt = -dc/dt = \\ &= -k_1 ab + k_2 c \end{aligned}$$

2 reaction



$$\begin{aligned} dc/dt &= -dp/dt = -da/dt = \\ &= -k_3 c \end{aligned}$$

Model

$$\begin{cases} da/dt = (-k_1 ab + k_2 c) + (k_3 c) \\ db/dt = (-k_1 ab + k_2 c) \\ dc/dt = -(-k_1 ab + k_2 c) - (k_3 c) \\ dp/dt = (k_3 c) \end{cases}$$



Фрагмент системы дифференциальных уравнений математической модели регуляции биосинтеза холестерина в клетке

$d(\text{Cholin})/dt = + \{KISBin * \text{Squalene}\} - \{KISBestr * \text{Cholin} * \text{ACAT}/(KMSBestr + \text{Cholin} + \text{ACAT})\} + \{KISBhydr * \text{Cholestr} * \text{Hydrolas}/(KMSBhydr + \text{holester} + \text{Hydrolas})\} + \{1785 * KISBfree * \text{LDLRLDLi}\} + \{-in1prot * \text{Cholin} * \text{SRPprot} + Kin2prot * \text{SRP-Chol}\} + \{Ksrpcdeg * \text{SRP-Chol}\} - \{Kutichol * \text{Cholin}\}$

$d(\text{LDLR}_{\text{---}})/dt = + \{-Kclaw1 * \text{LDLR}_{\text{---}} * \text{LDL}_{\text{---}} + Kclaw2 * \text{LDLRLDLo}\} + \{Kfre_{\text{---}} t * \text{LDLRin}\} - \{KLDLRdeg * \text{LDLR}_{\text{---}}\}$

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

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

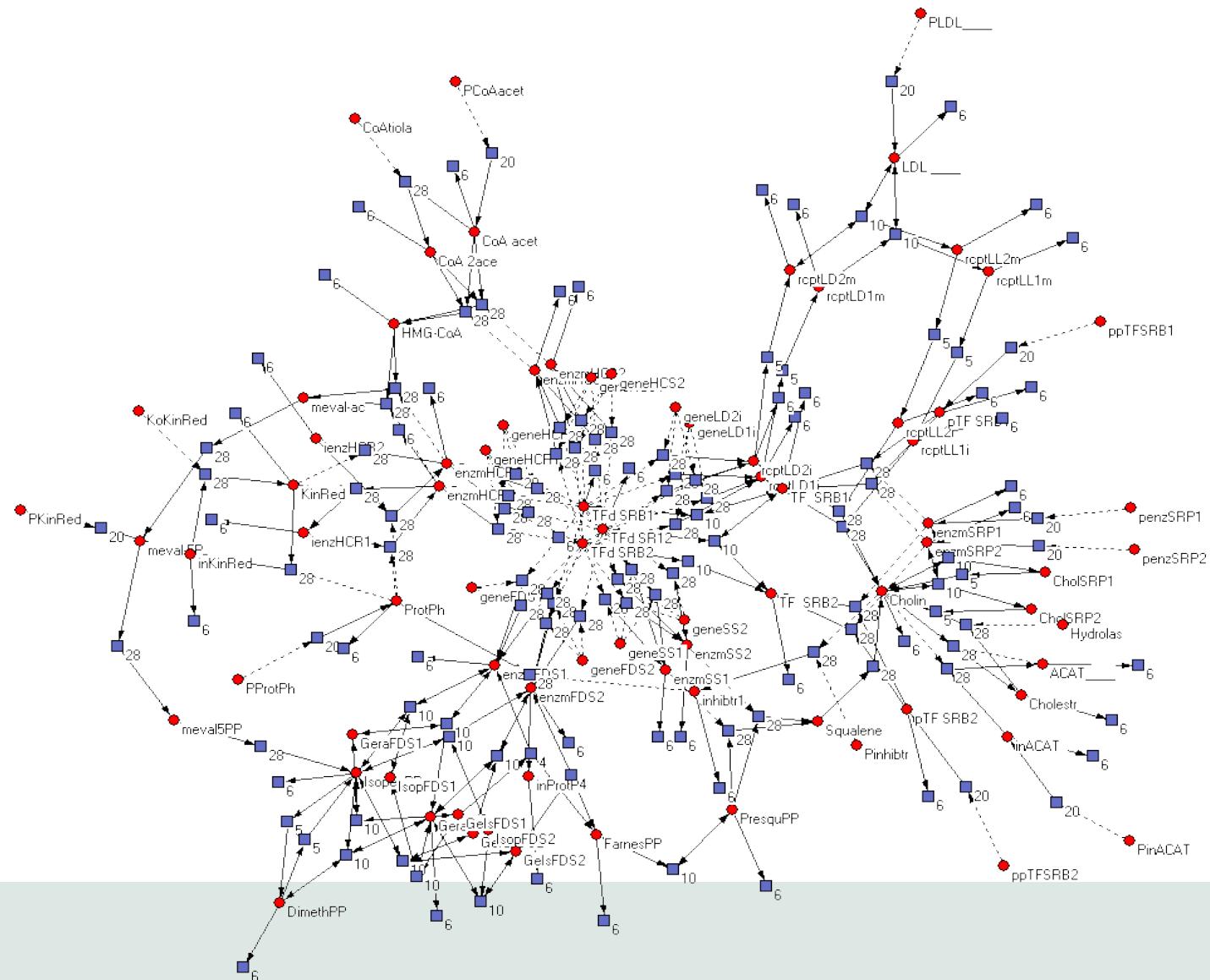
.....

.....

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	$\text{л} * \text{ммоль}^{-1} * \text{с}^{-1}$ $Kclaw1 = 3.3 * 10^{01}$	с^{-1} $Kclaw2 = 1 * 10^{-03}$
Константы ферментативных реакций	Константы оборота фермента, с ⁻¹ $KISBestr = 1 * 10^{+02}, KISBhydr = 1.9 * 10^{+03}, KISBldlr = 1 * 10^1$	Константы Михаэлиса, ммоль/л $KMSBestr = 6.67 * 10^{-03}, KMSBhydr = 7.6 * 10^{-01}, KMSBldlr = 3.3 * 10^{-05}$



Граф модели регуляции синтеза холестерина





Системы обыкновенных дифференциальных уравнений общего вида, описывающие функционирование генных сетей

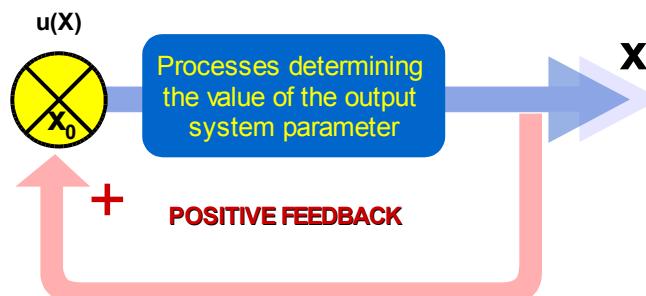
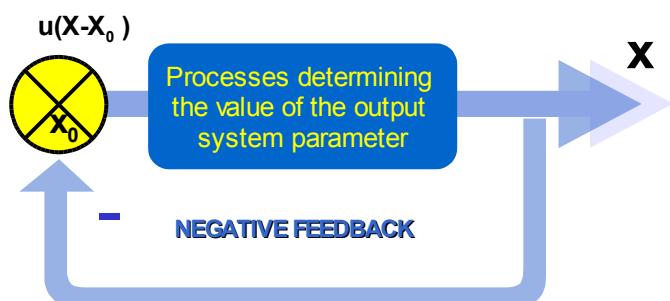
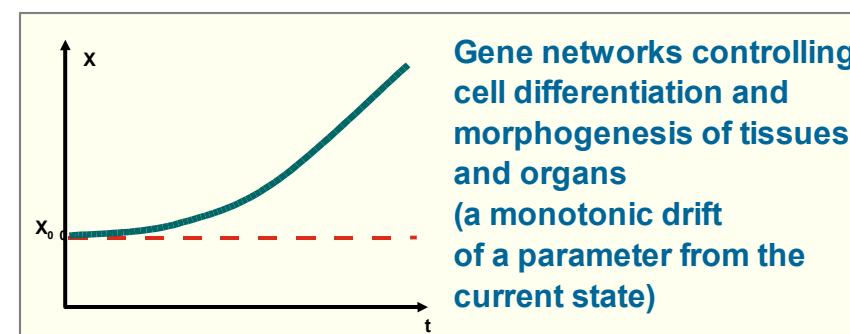
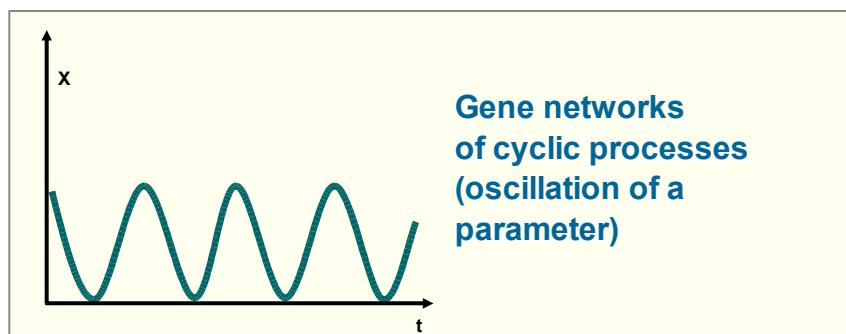
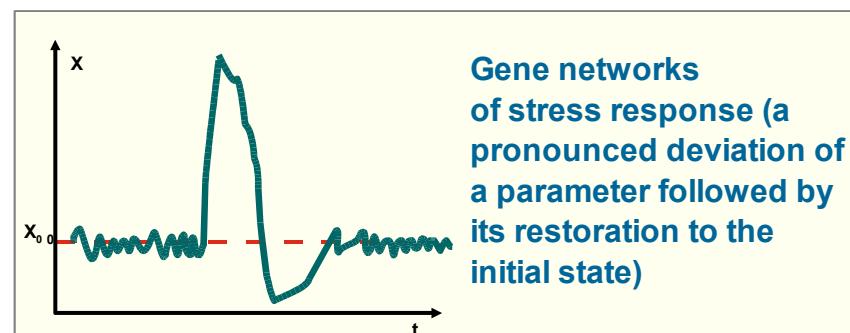
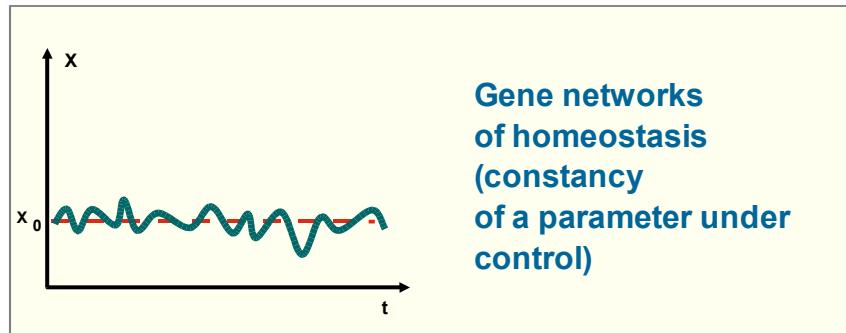
$$\frac{dx_i}{dt} = F_i(X_{i'}, K) - x_i G(X, K), \quad i=1, \dots, n.$$

n – количество динамических переменных модели генной сети, $X=(x_1, \dots, x_n)$ – вектор динамических переменных, $K=(k_1, \dots, k_m)$ – вектор параметров, $X_{i'}=(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)$ – вектор динамических переменных, не содержащий x_i , $F_i(X_{i'}, K)$ и $G_i(X, K)$ – рациональные функции, которые принимают неотрицательные значения при всех неотрицательных значениях переменных и параметров.

Таким образом, изучение динамики функционирования генных сетей можно рассматривать как математическую задачу изучения свойств решений систем данного вида.



Types of dynamics of processes controlled by gene networks





Элементы гипотетических генных сетей

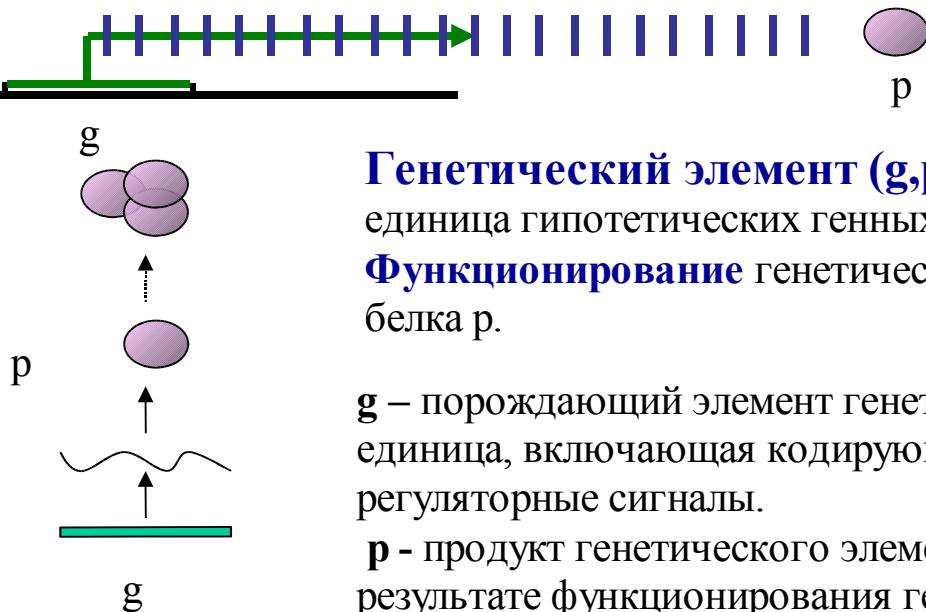


Мультимер
белка

белок

мРНК

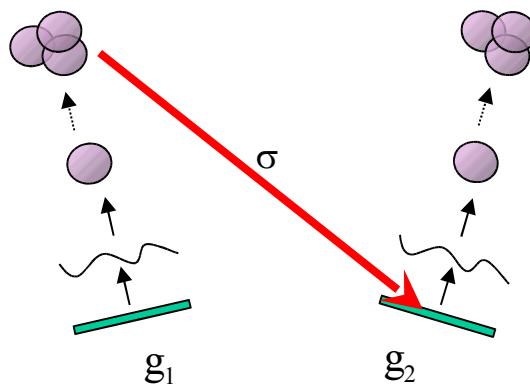
ген



Генетический элемент (g,p)- элементарная структурная единица гипотетических генных сетей,
Функционирование генетического элемента - процесс синтеза белка p.

g – порождающий элемент генетического элемента: генетическая единица, включающая кодирующую часть, промотор и его регуляторные сигналы.

p - продукт генетического элемента: белок синтезируемый в результате функционирования генетического элемента



Регуляторный механизм (σ) - элементарный процесс, посредством которого некоторая совокупность генетических элементов, например (g₁,p₁), действуют на конкретный генетический элемент (g₂,p₂), уменьшая или увеличивая скорость синтеза его продукта p₂.



ГИПОТЕТИЧЕСКОЙ ГЕННОЙ СЕТЬЮ

будем называть конечную совокупность

ГЕНЕТИЧЕСКИХ ЭЛЕМЕНТОВ,

которые связаны в единую функциональную систему посредством заданных

РЕГУЛЯТОРНЫХ МЕХАНИЗМОВ.



Основная проблема теории гипотетических генных:



Исследование взаимосвязи структуры регуляторных связей гипотетической генной сети с предельными режимами ее функционирования

Под предельными режимами мы понимаем такие состояния гипотетических генных сетей, которые устанавливаются на бесконечном интервале времени.

Априори ясно, что предельными режимами могут быть стационарные точки, циклы, а также аттракторы более сложной природы.

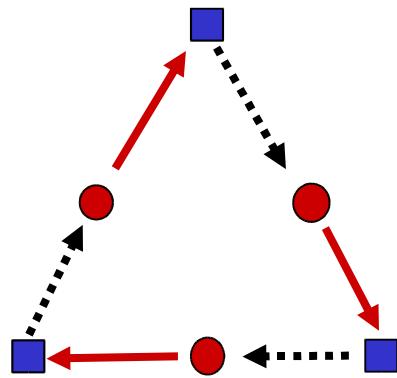
Однако заранее совершенно не ясно как они могут быть связаны со структурой регуляторных связей. В частности, возникает вопрос, что мы можем сказать о возможных режимах поведения гипотетической генной сети, если нам задан структурный граф.



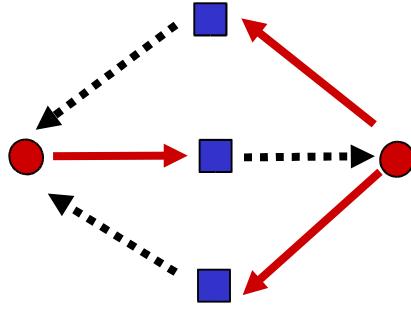
Орграфы гипотетических генных сетей



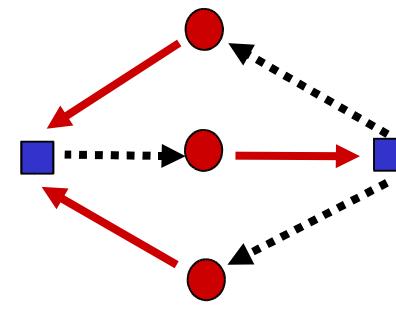
а)



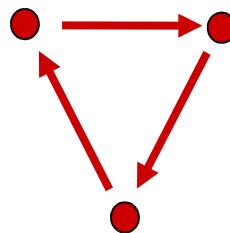
б)



в)



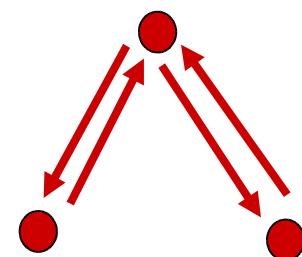
г)



д)

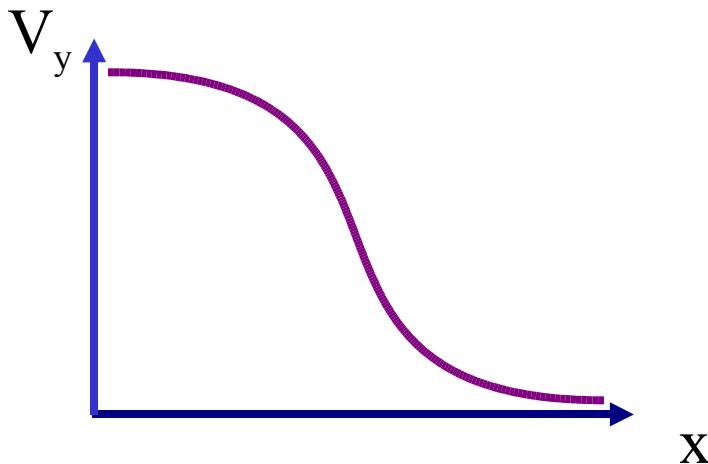


е)





Отрицательная регуляция



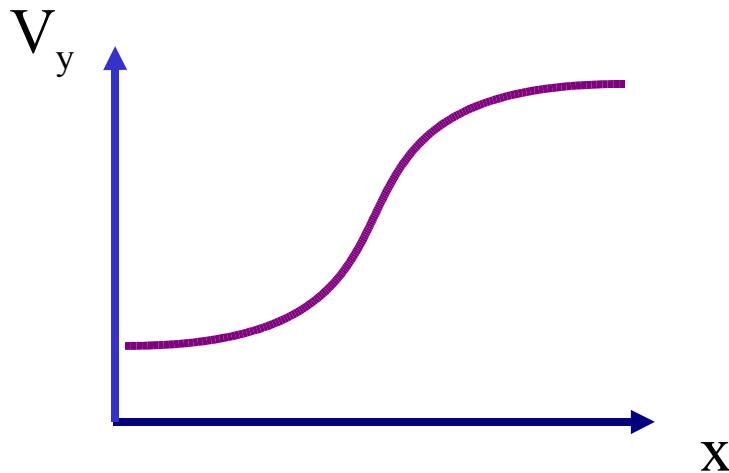
Скорость синтеза белка убывает с ростом концентрации x

Основное уравнение, описывающее отрицательную
регуляторную связь

$$V_y = \alpha / (1 + K_1^\gamma x^\gamma)$$



Положительная регуляция



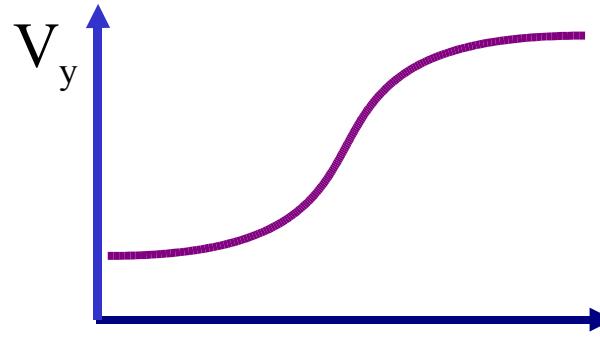
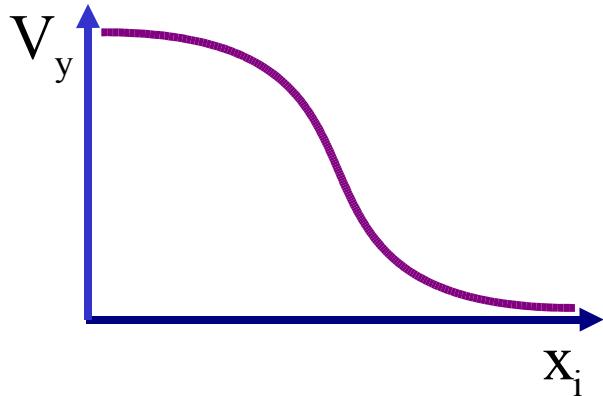
Скорость синтеза белка у возрастает с ростом концентрации x

Простейшее уравнение, описывающее положительную
регуляторную связь

$$V_y = [\alpha + K_1^h x^h] / [1 + K_2^\gamma x^\gamma]$$



Множественная регуляция

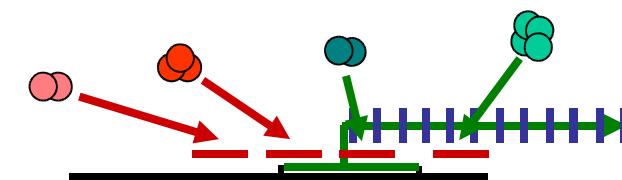
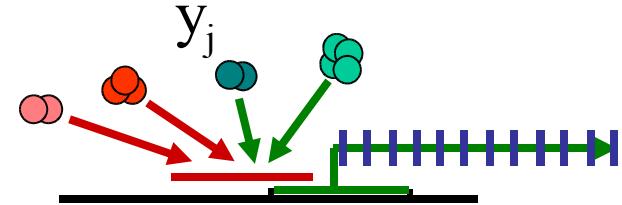
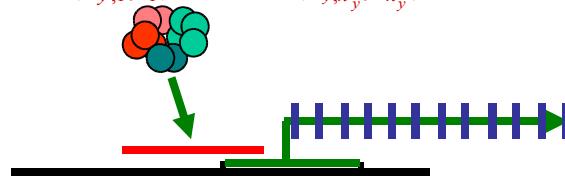


$$1. \quad V_y = \frac{\alpha + (k_{y,1}y_1)^{h_1} + \dots + (k_{y,n_y}y_{n_y})^{h_{n_y}}}{1 + (s_{x,1}x_1)^{\gamma_1} + \dots + (s_{x,n_x}x_{n_x})^{\gamma_{n_x}} + (s_{y,1}y_1)^{h_1} + \dots + (s_{y,n_y}y_{n_y})^{h_{n_y}}}$$

$$2. \quad V_y = \frac{\alpha + ky_1^{h_1} \dots y_n^{h_n}}{1 + sy_1^{h_1} \dots y_n^{h_n}}$$

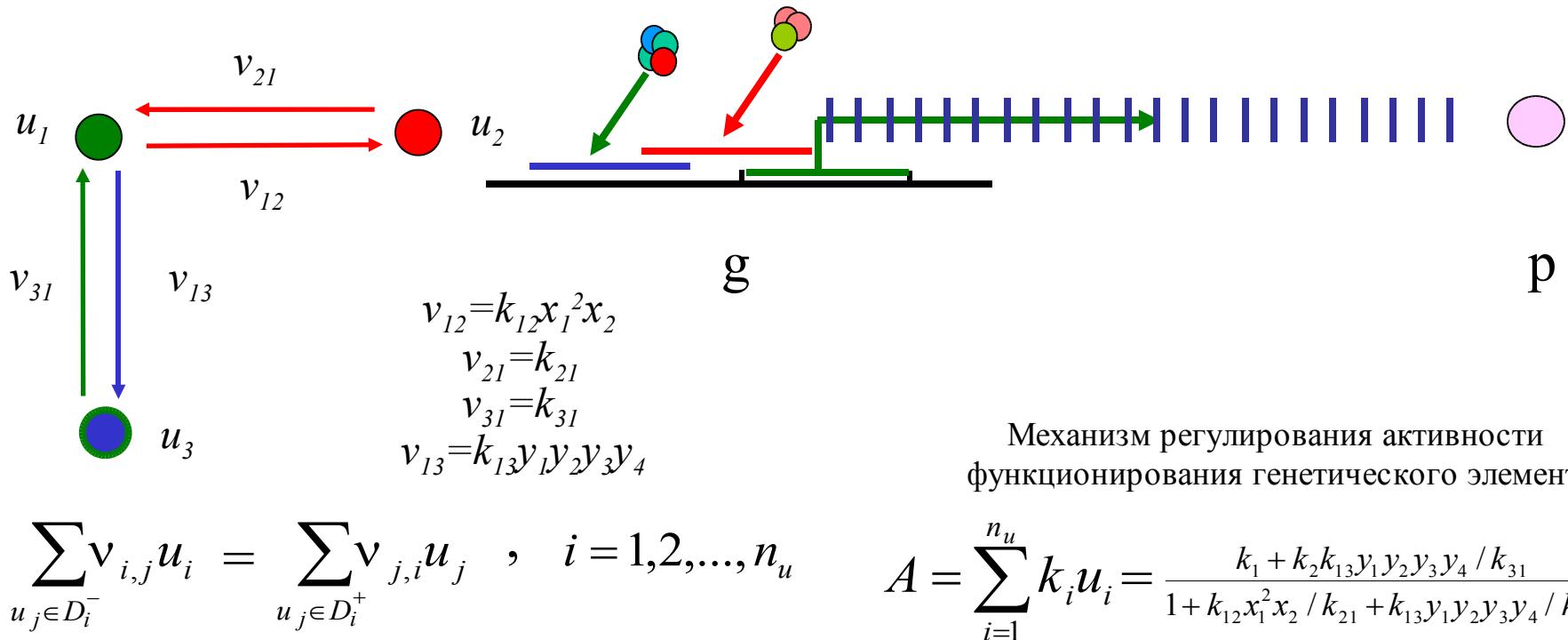
$$3. \quad V_y = \alpha \frac{\alpha_1 + (k_{y,1}y_1)^{h_1}}{1 + (s_{y,1}y_1)^{h_1}} \cdot \dots \cdot \frac{\alpha_{n_y} + (k_{y,n_y}y_{n_y})^{h_{n_y}}}{1 + (s_{y,n_y}y_{n_y})^{h_{n_y}}} \cdot \frac{1}{1 + (s_{x,1}x_1)^{\gamma_1}} \cdot \dots \cdot \frac{1}{1 + (s_{x,n_x}x_{n_x})^{\gamma_{n_x}}}$$

$$4. \quad V_y = \frac{\alpha_{y,1} + (k_{y,1}y_1)^{h_1}}{1 + (s_{y,1}y_1)^{h_1}} + \dots + \frac{\alpha_{y,n_y} + (k_{y,n_y}y_{n_y})^{h_{n_y}}}{1 + (s_{y,n_y}y_{n_y})^{h_{n_y}}} + \frac{\alpha_{x,1}}{1 + (s_{x,1}x_1)^{\gamma_1}} + \dots + \frac{\alpha_{x,n_x}}{1 + (s_{x,n_x}x_{n_x})^{\gamma_{n_x}}}$$





2. Описание регуляции генетического элемента



Механизм регулирования активности функционирования генетического элемента

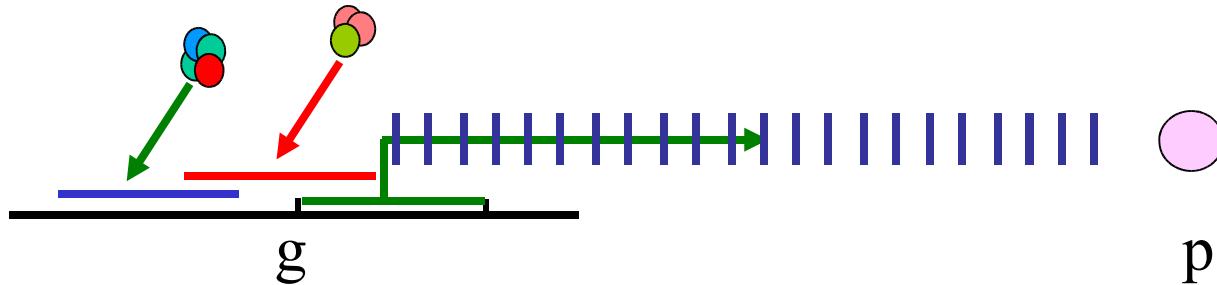
Порождающий элемент g может пребывать в нескольких альтернативных состояниях u_i , $i=1,\dots,n_u$. Считаем, что известна схема перехода состояний друг в друга. Эту схему будем задавать ориентированным графом, в котором вершины представляют альтернативные состояния, а дуги – переходы между ними. Каждый переход осуществляется с некоторой скоростью $v(u_i, u_j)$, которая зависит от концентраций ряда белков-регуляторов, что выражается в конкретном виде зависимостей $v(u_i, u_j)$ от концентраций белков-регуляторов.



Системы, описывающие функционирование ГГС



1. Уравнения, описывающие функционирование генетического элемента



$$\frac{dx_1}{dt} = A - k_1 x_1$$

$$\frac{dx_i}{dt} = k_{i-1} x_{i-1} - k_i x_i$$

$$\frac{dx}{dt} = k_{n_g} x_{n_g} - \beta x$$

n_g – количество промежуточных стадий синтеза белка, x_i – концентрации промежуточных продуктов синтеза белка, x - концентрация белка, α - константа скорости инициации конститутивного синтеза, k_i – константы скорости перехода из i -го промежуточного состояния в $(i+1)$ -ое промежуточное состояние, β - константа деградации белка p .



Модели функционирования ГГС



Описание функционирования ГГС автономными системами

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

Здесь n – количество различных белков,

m – количество различных порождающих элементов,

$n_{j,i}$ – количество промежуточных стадий в синтезе i -го белка j -м порождающим элементом,

$k_{j,i,s}^{(n_{j,i})}$ – константа скорости перехода s -й промежуточной стадии в $(s+1)$ -ю промежуточную стадию при синтезе i -го белка j -м порождающим элементом,

β_i – константа распада i -го белка,

P_j – множество номеров белков, которые кодируются j -м порождающим элементом,

R_j – множество номеров белков, которые регулируют активность функционирования j -го порождающего элемента, G_i - множество номеров порождающих элементов, которые кодируют i -й белок. Класс этих систем будем обозначать через МГГС.



2. Описание функционирования ГГС системами дифференциальных уравнений с запаздывающими аргументами



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

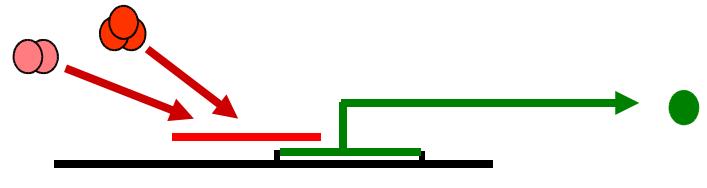
$$\tau_{j,i} = \sum_{s=1}^{n_{j,i}} 1/k_{j,i,s}^{(n_{j,i})}$$



Классы гипотетических генных сетей

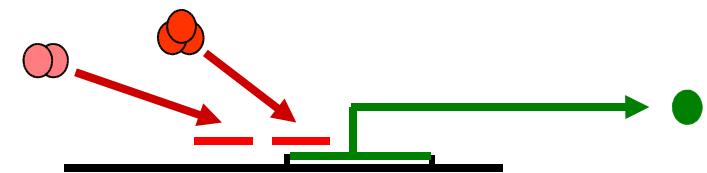
Класс 1.

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



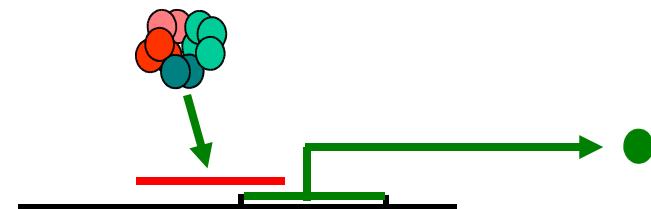
Класс 2.

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



Класс 3.

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



Класс 4.

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





Двух параметров достаточно для обеспечения сложного поведения гипотетических генных сетей

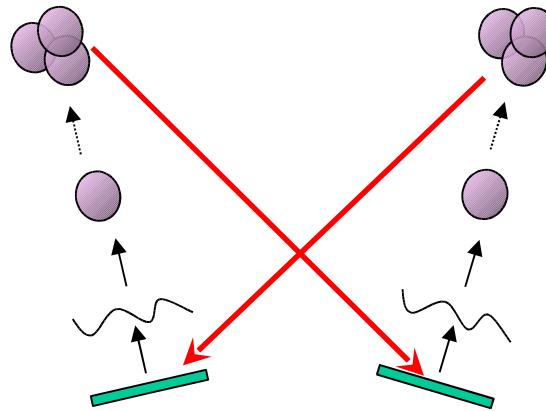


$$\alpha_i = \alpha, \beta_i = 1, \beta_{i,j} = 1, \gamma_i = \gamma.$$

$$\frac{dx_i}{dt} = \frac{\alpha}{1 + \sum_{j \in D_i} x_j^\gamma} - x_i, i = 1, n$$



Генная сеть S(2,2), состоящая из двух генетических элементов и двух регуляторных связей



Графическое представление генной сети: граф G(2,2).

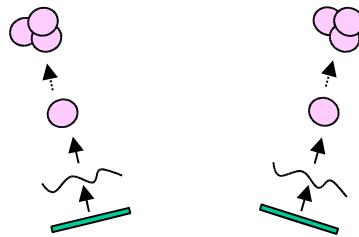


Математическая модель M(2,2), описывающая функционирование генной сети S(2,2)

$$\frac{dx_1}{dt} = \alpha / (1 + x_2^\gamma) - x_1$$

$$\frac{dx_2}{dt} = \alpha / (1 + p x_1^\gamma) - x_2$$

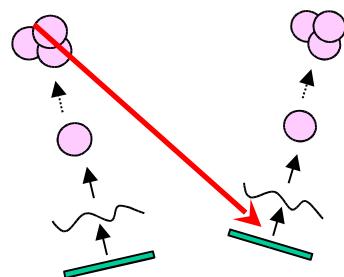
Возможные режимы:
две устойчивые точки



$$\frac{dx_1}{dt} = \alpha - x_1$$

$$\frac{dx_2}{dt} = \alpha - x_2,$$

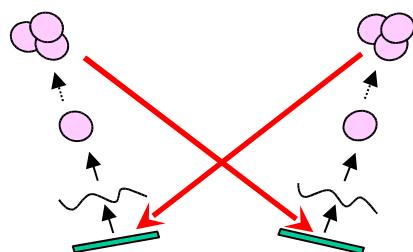
Одна точка покоя: $x_1 = x_2 = \alpha$



$$\frac{dx_1}{dt} = \alpha - x_1$$

$$\frac{dx_2}{dt} = \alpha/(1+x_1^\gamma) - x_2$$

Одна точка покоя: $x_1 = \alpha$, $x_2 = \alpha/(1+\alpha^m)$



$$\frac{dx_1}{dt} = \alpha/(1+x_2^\gamma) - x_1$$

$$\frac{dx_2}{dt} = \alpha/(1+x_1^\gamma) - x_2$$

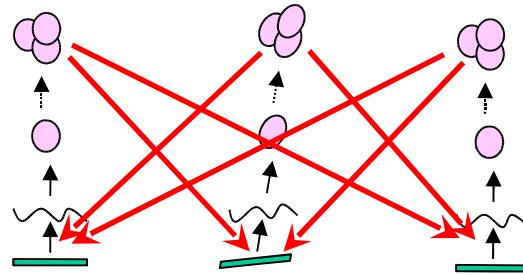


Общее число связных ГГС, содержащих три генетических элемента - 13, из них:

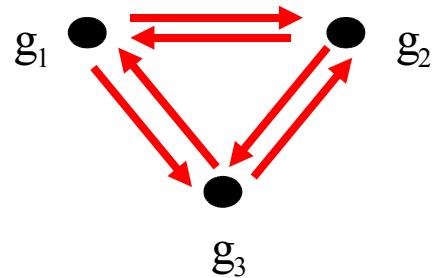
- 7 имеют одну точку покоя;**
- 4 имеют две точки покоя;**
- 1 имеет три точки покоя;**
- 1 имеет предельный цикл;**



Генная сеть S(3,3)



Графическое представление генной сети: Граф G(3,3)



Математическая модель M(3,3)

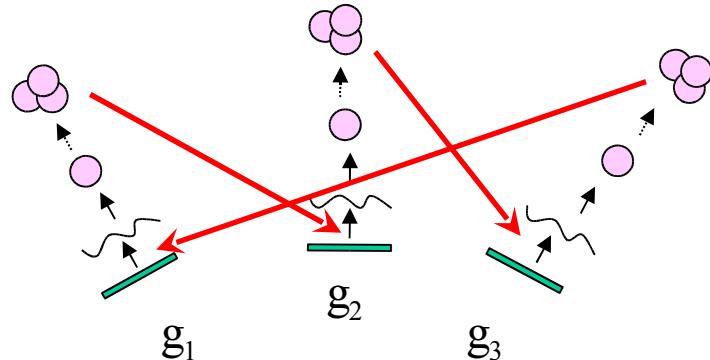
$$\frac{dx_1}{dt} = \alpha / (1 + x_2^\gamma + x_3^\gamma) - x_1$$

$$\frac{dx_2}{dt} = \alpha / (1 + x_1^\gamma + x_3^\gamma) - x_2$$

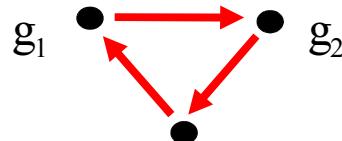
Возможные режимы: три устойчивые точки покоя
 $\frac{dx_3}{dt} = \alpha / (1 + x_1^\gamma + x_2^\gamma) - x_3$



Генная сеть S(3,2), состоящая из трех генетических элементов и трех регуляторных связей



Графическое представление генной сети. Граф G(3,2)



Математическая модель $M(3,2)$, описывающая функционирование генной сети $S(3,2)$.

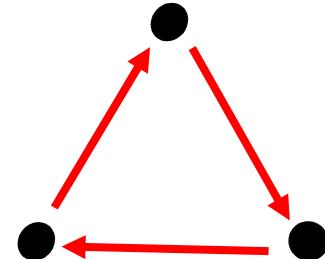
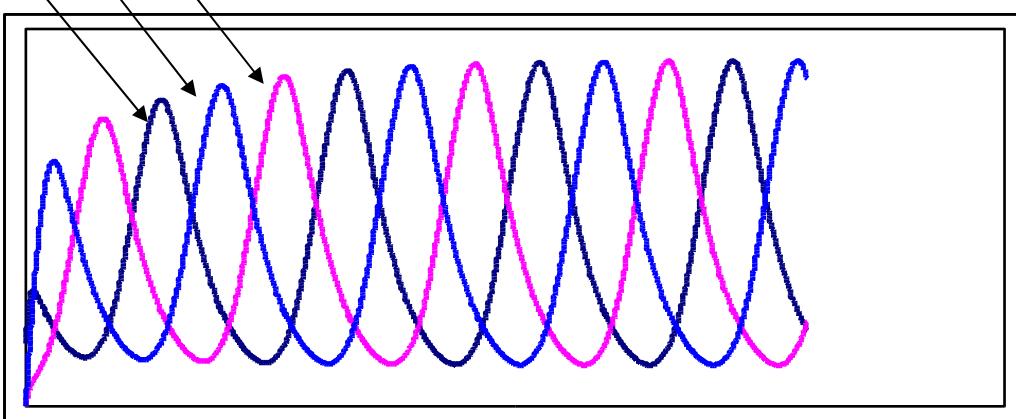
$$\frac{dx_1}{dt} = \alpha / (1 + x_3^\gamma) - x_1$$

$$\frac{dx_2}{dt} = \alpha / (1 + x_1^\gamma) - x_2$$

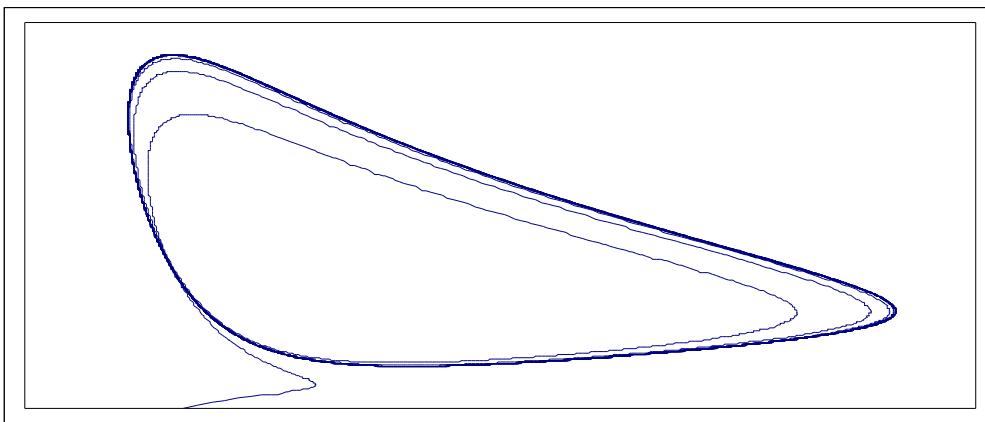
$\frac{dx_3}{dt} = \alpha / (1 + x_2^\gamma) - x_3$
Возможные режимы: один устойчивый предельный цикл



Временные кривые



Фазовая кривая в плоскости (x_1, x_2)





Общее число связных ГГС класса 1, содержащих четыре генетических элемента - 201

из них:

106 имеют одну точку покоя;

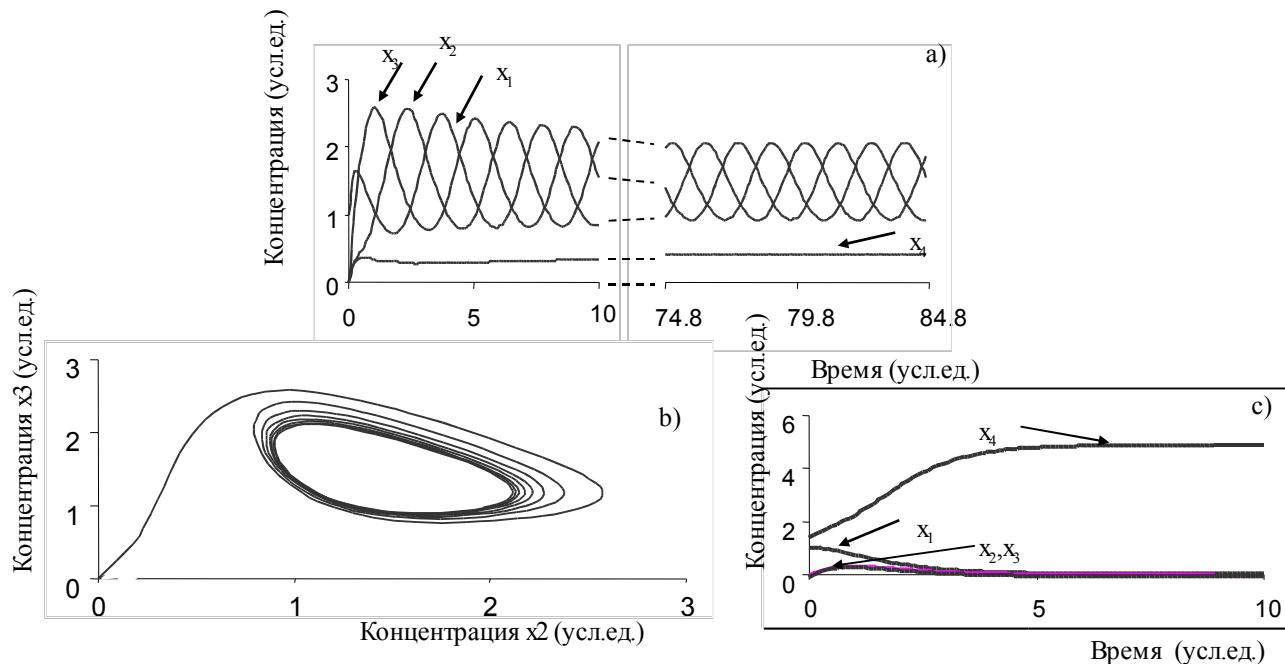
69 имеют две точки покоя;

13 имеет 3 точки покоя;

1 имеет 4 точки покоя;

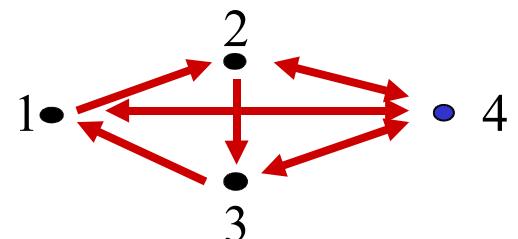
11 имеют один предельный цикл;

1 имеет предельный цикл и точку покоя одновременно.



Расчет динамики функционирования генной сети при различных начальных условиях:

- a, b) $x_1=1$, $x_2=x_3=x_4=0$;
c) $x_1=1$, $x_2=x_3=0$, $x_4=1.5$

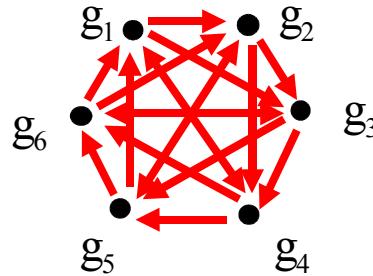




Число генетических элементов	общее количество связных ГГС класса 1
5	9 390
6	1 531 336
7	880 492 496
8	1 792 477 159 408



Графическое представление генной

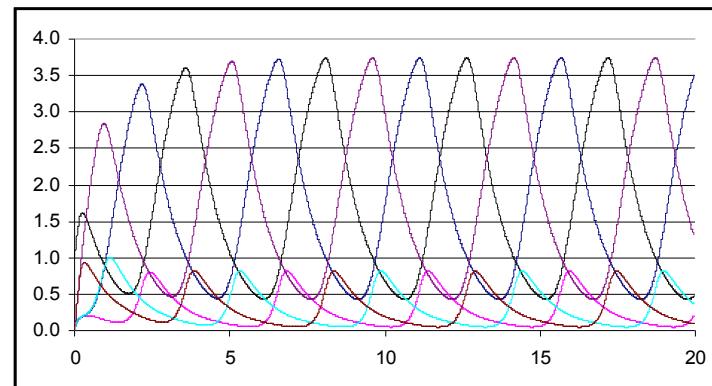


Математическая модель

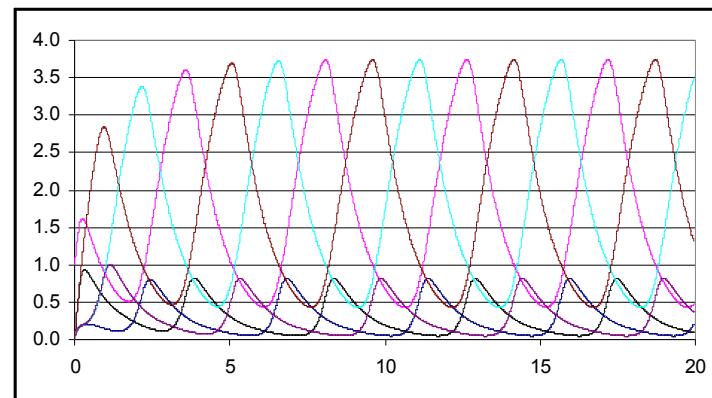
$$\begin{aligned} \frac{dx_1}{dt} &= \alpha / (1 + x_4^\gamma + x_5^\gamma + x_6^\gamma) - x_1 \\ \frac{dx_2}{dt} &= \alpha / (1 + x_1^\gamma + x_5^\gamma + x_6^\gamma) - x_2 \\ \frac{dx_3}{dt} &= \alpha / (1 + x_2^\gamma + x_1^\gamma + x_6^\gamma) - x_3 \\ \frac{dx_4}{dt} &= \alpha / (1 + x_3^\gamma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

При $\alpha > 3$ и $\gamma > 4$ система имеет два устойчивых предельных цикла

Первый устойчивый цикл

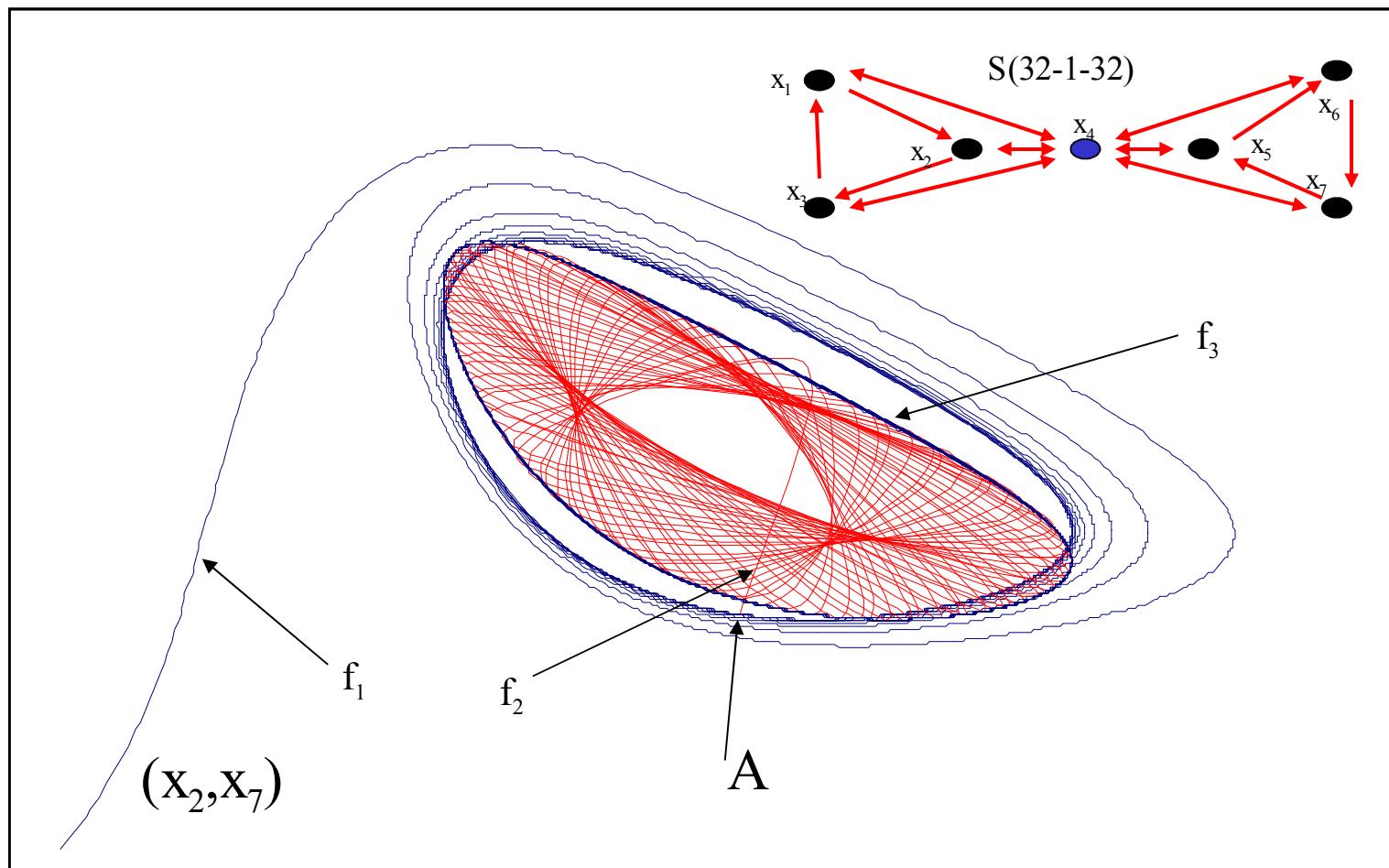


Второй устойчивый цикл

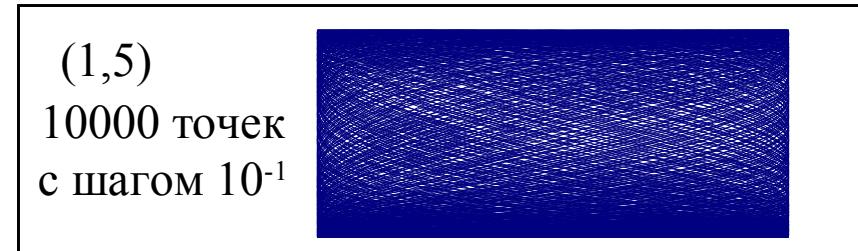
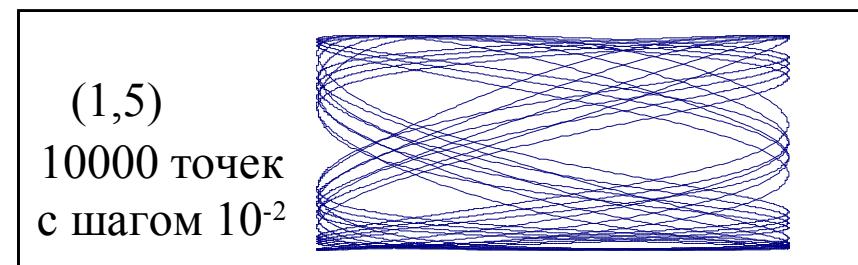
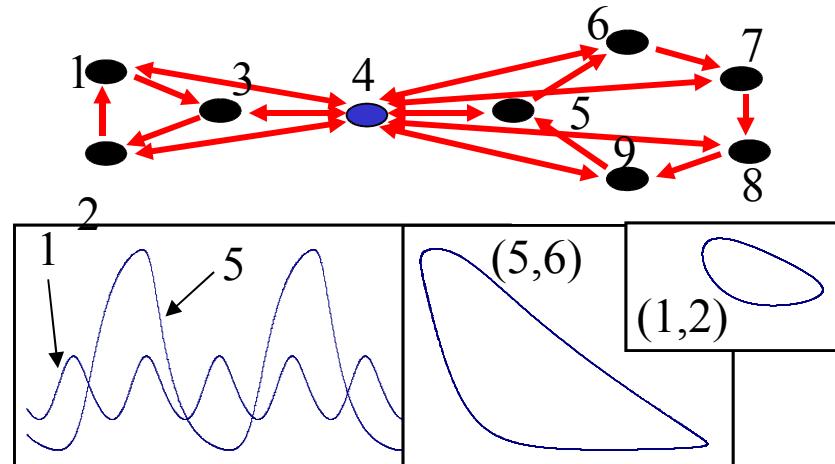




Одна устойчивая точка покоя, три устойчивых цикла, пять неустойчивых циклов

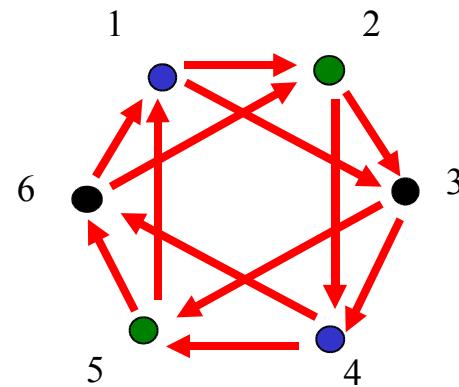
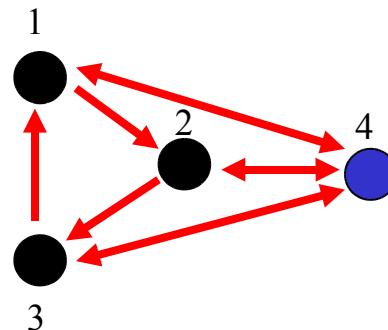
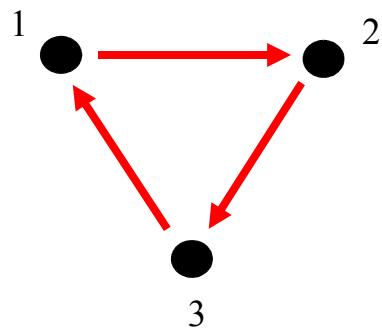
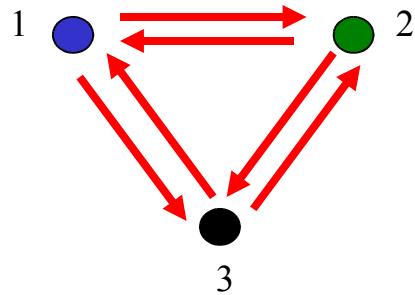


Кривая f_1 стартует и остается в гиперпространстве $x_1=x_5$, $x_2=x_6$, $x_3=x_7$, притягиваясь к неустойчивому циклу





Необходимые сведения из теории орграфов



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

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



Гипотеза о числе устойчивых состояний канонических ГГС классов 1-4



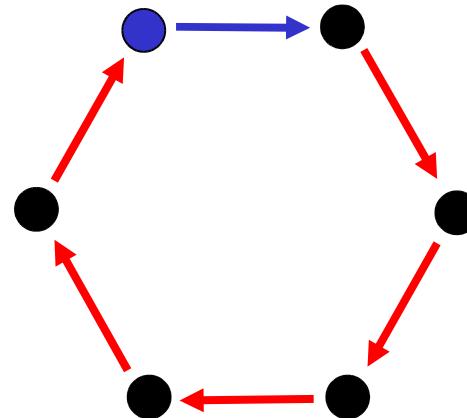
Пусть дан структурный орграф G . Пусть $M_i(G)$ обозначает модель канонической ГГС i -го класса, построенную на орграфе G . Тогда существуют такие γ_0 и α_0 , что при любых $\gamma > \gamma_0$ и $\alpha > \alpha_0$

1. для канонических ГГС первого и второго классов количество имеющихся у них устойчивых состояний равняется числу 1-баз орграфа G ;
2. для канонических ГГС третьего и четвертого классов количество имеющихся у них устойчивых состояний равняется числу 1m - баз орграфа G .

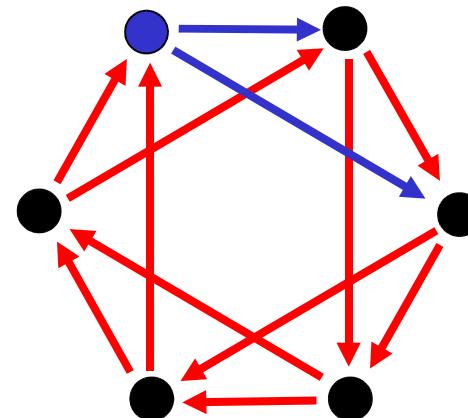
3. если каждая вершина G принадлежит хотя бы одной 1-базе, то $M_1(G)$ не имеет других устойчивых состояний;
4. если каноническая гипотетическая генная сеть 1-го класса, построенная на ограничении структурного орграфа G на совокупность вершин, которые не содержатся ни в одной его 1-базе, не имеет устойчивых колебательных режимов функционирования ни при каких значениях параметров γ и α , то и исходная каноническая ГГС $S_1(G)$ не имеет устойчивых колебательных режимов функционирования.
5. если каноническая гипотетическая генная сеть 1-го класса, при достаточно больших значениях параметров γ и α , содержит устойчивые колебательные режимы функционирования, то каноническая гипотетическая генная сеть, построенная на ограничении структурного орграфа G на совокупность вершин, которые не содержатся ни в одной его 1-базе, также содержит при достаточно больших значениях параметров γ и α устойчивые колебательные режимы функционирования.



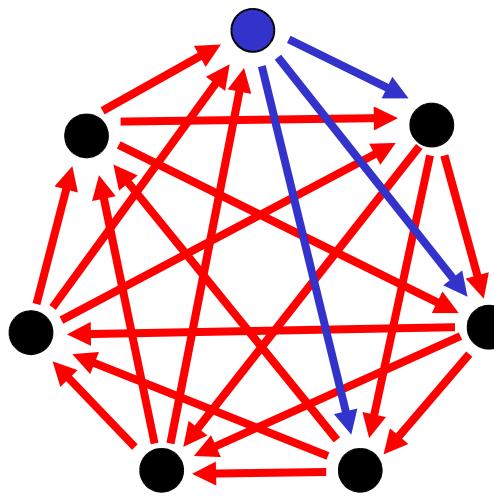
Симметричные орграфы $G(n,k)$



$G(6,2)$



$G(6,3)$



$G(7,4)$



(n,k)-Критерий

об устойчивых предельных режимах функционирования моделей $M_i(n,k)$.



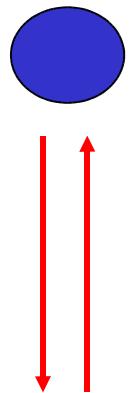
Пусть d - наибольший общий делитель n и k .

Существуют такие γ_0 и α_0 , что при любых $\gamma > \gamma_0$ и $\alpha > \alpha_0$

1. Если $d=k$, то каждая из моделей $M_i(n,k)$ имеет k устойчивых стационарных решений, а все другие предельные режимы функционирования $M_i(n,k)$ в указанной параметрической области неустойчивы.
2. Если $d \neq k$, то модель $M_i(n,k)$ имеет d устойчивых предельных циклов, При этом в указанной параметрической области все другие предельные режимы функционирования гипотетической генной сети 1 класса неустойчивы.



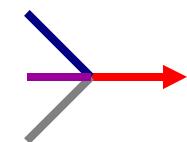
The element of hgn summarizing 2 lowest bits of the binary numbers



Molecular trigger- the elementary memory cell



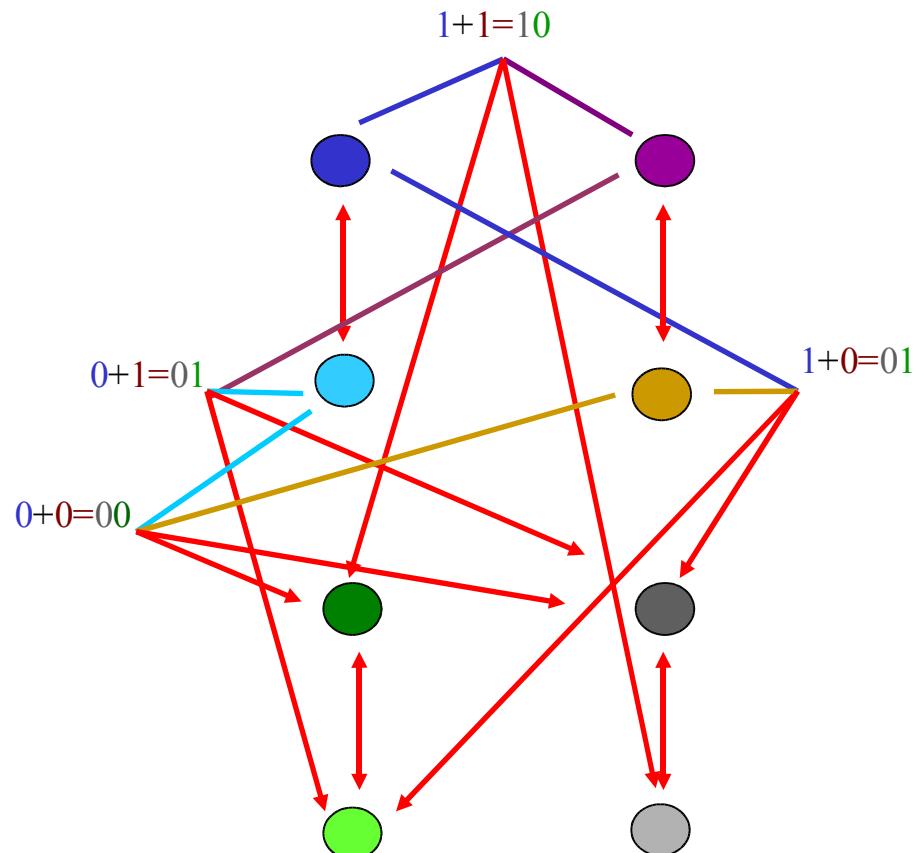
Genetic element



3-d class repression



Repression





The element of hgn summarizing 2 upper bits of the binary numbers



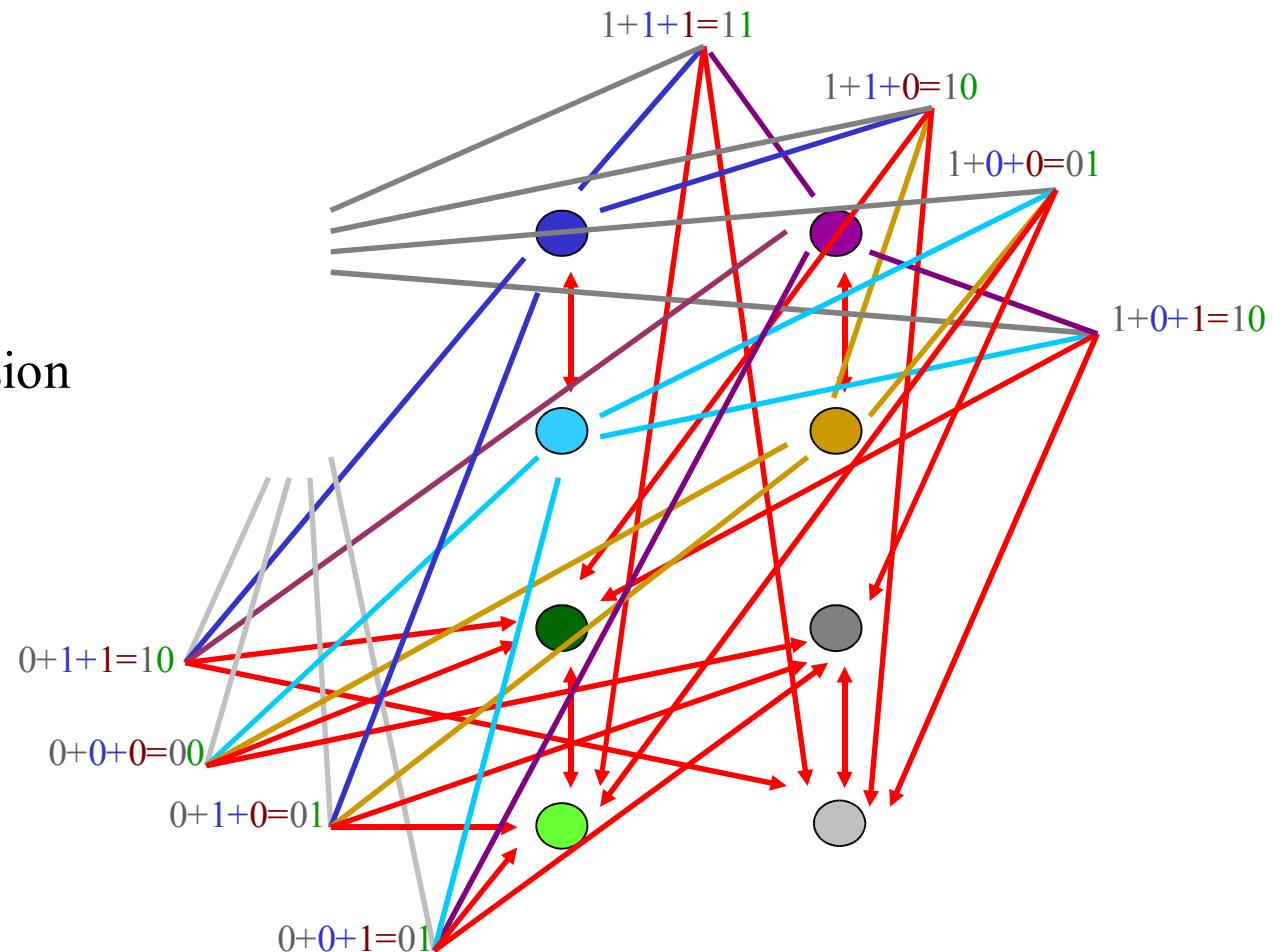
Genetic element



3-d class repression

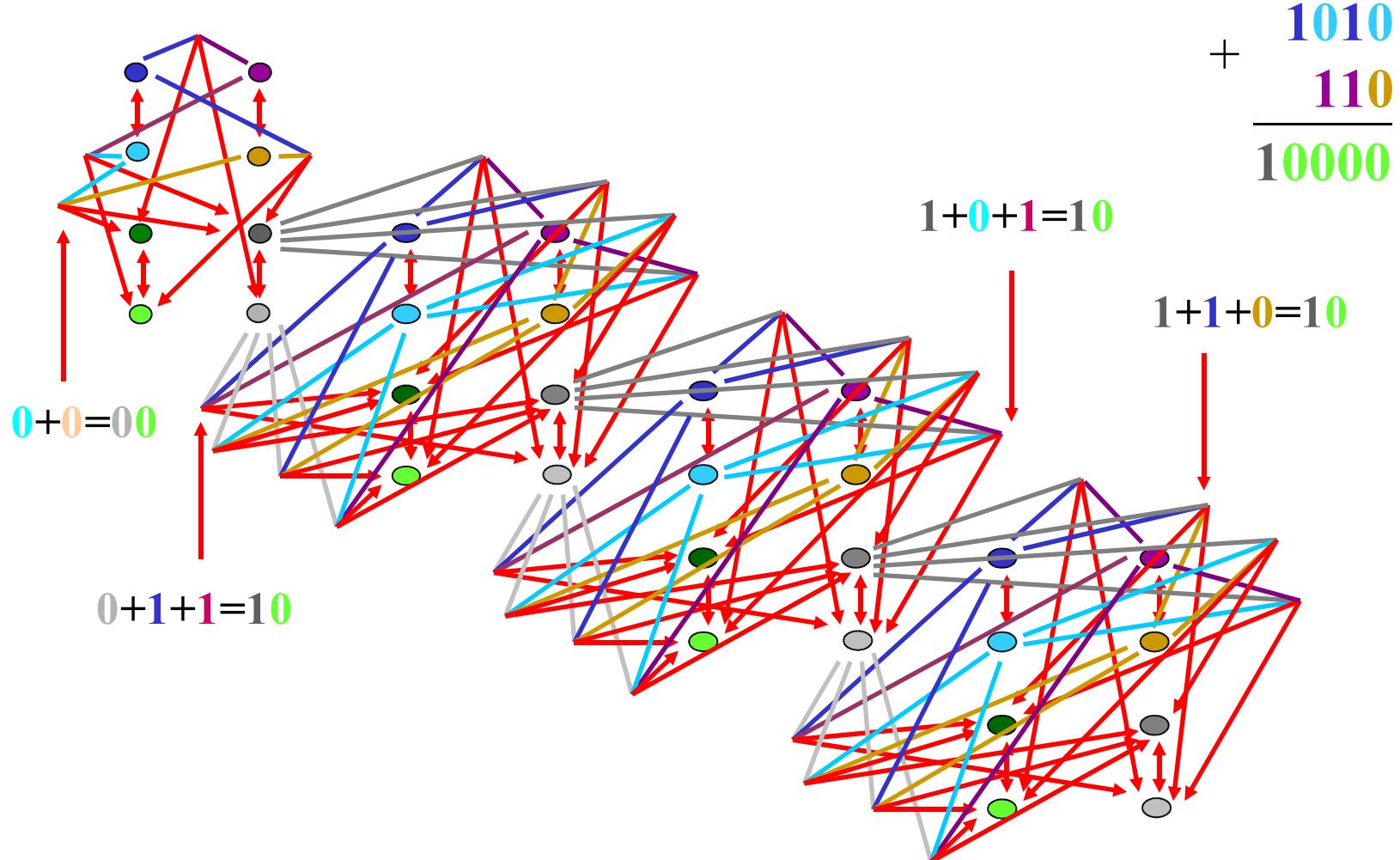


Repression





The hgn summarizing 2 binary 4 bits numbers





The hgn summarizing two binary 8 bits numbers

