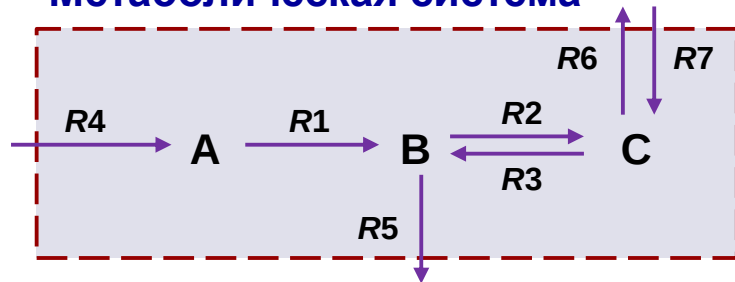


Системная компьютерная биология и биоинформатика: моделирование и компьютерный дизайн экспериментов по созданию штаммов с целевыми свойствами

Н.А. Колчанов, С.Е. Пельтек , А.С. Розанов, С.А. Лашин, Ф.В. Казанцев
ФИЦ Институт цитологии и генетики СО РАН
Новосибирский Государственный Университет
Г. Новосибирск

Равновесные потоковые модели (Flux-balance analysis): расчет величин стационарных потоков метаболитов в метаболических сетях

Метаболическая система



Реакции метаболической системы

Внутренние

R1: -1 A → 1 B

R2: -1 B → 1 C

R3: -1 C → 1 B

Обменные

R4: 1 A

R5: -1 B

R6: -1 C

R7: 1 C

Стехиометрическая матрица

$$S = \begin{array}{c|ccccccc} & R1 & R2 & R3 & R4 & R5 & R6 & R7 \\ \hline A & -1 & 0 & 0 & 1 & 0 & 0 & 0 \\ B & 1 & -1 & 1 & 0 & -1 & 0 & 0 \\ C & 0 & 1 & -1 & 0 & 0 & -1 & 1 \end{array}$$

1: метаболит продуцируется в реакции

-1: метаболит расходуется в реакции

0: метаболит не участвует в реакции

1. Состояние метаболической сети описывается вектором концентраций метаболитов $\mathbf{X} = [x_1, \dots, x_m]$

2. Изменение концентраций $\mathbf{X}$ описывается системой дифференциальных уравнений:

$$\frac{d\mathbf{x}}{dt} = \mathbf{S}\mathbf{v}$$

Здесь:

$\mathbf{S}$ – матрица стехиометрических коэффициентов,

$\mathbf{v} = [v_1, \dots, v_i, \dots, v_n]$ – вектор потоков метаболической сети, соответствующий n реакциям; поток $v_i(t)$ равен скорости, при которой i -я реакция протекает в момент времени t ;

3. Рассматривается равновесное состояние системы в n -мерном пространстве стационарных потоков $\mathbf{R}_n$:

$$\frac{d\mathbf{x}}{dt} = \mathbf{S}\mathbf{v} \longrightarrow \mathbf{S}\mathbf{v} = 0$$

Равновесное
состоянии
системы

Пространство
решений без
ограничений

Нахождение потоков $\{v_i\}$ стационарного состояния системы сводится к решению системы из m линейных алгебраических уравнений. Как правило, число реакций n в системе больше числа метаболитических компонент (m), т.е. число переменных больше числа уравнений ($n > m$), поэтому система не имеет единственного решения.

Ограничения:

- 1) $0 = S_v$
- 2) $a_i < v_i < b_i$

Накладывая на потоки v_i ограничения, обусловленные термодинамикой протекания реакций, генетической регуляцией, границами скоростей ферментативных реакций, величинами входных и выходных потоков метаболитов и др., можно определить выпуклый многогранный конус C в n -мерном пространстве R^n - область допустимых стационарных потоков

Ограниченная
область
стационарных
потоков C

Расчёт потоков,
максимизирующих
целевую функцию Z

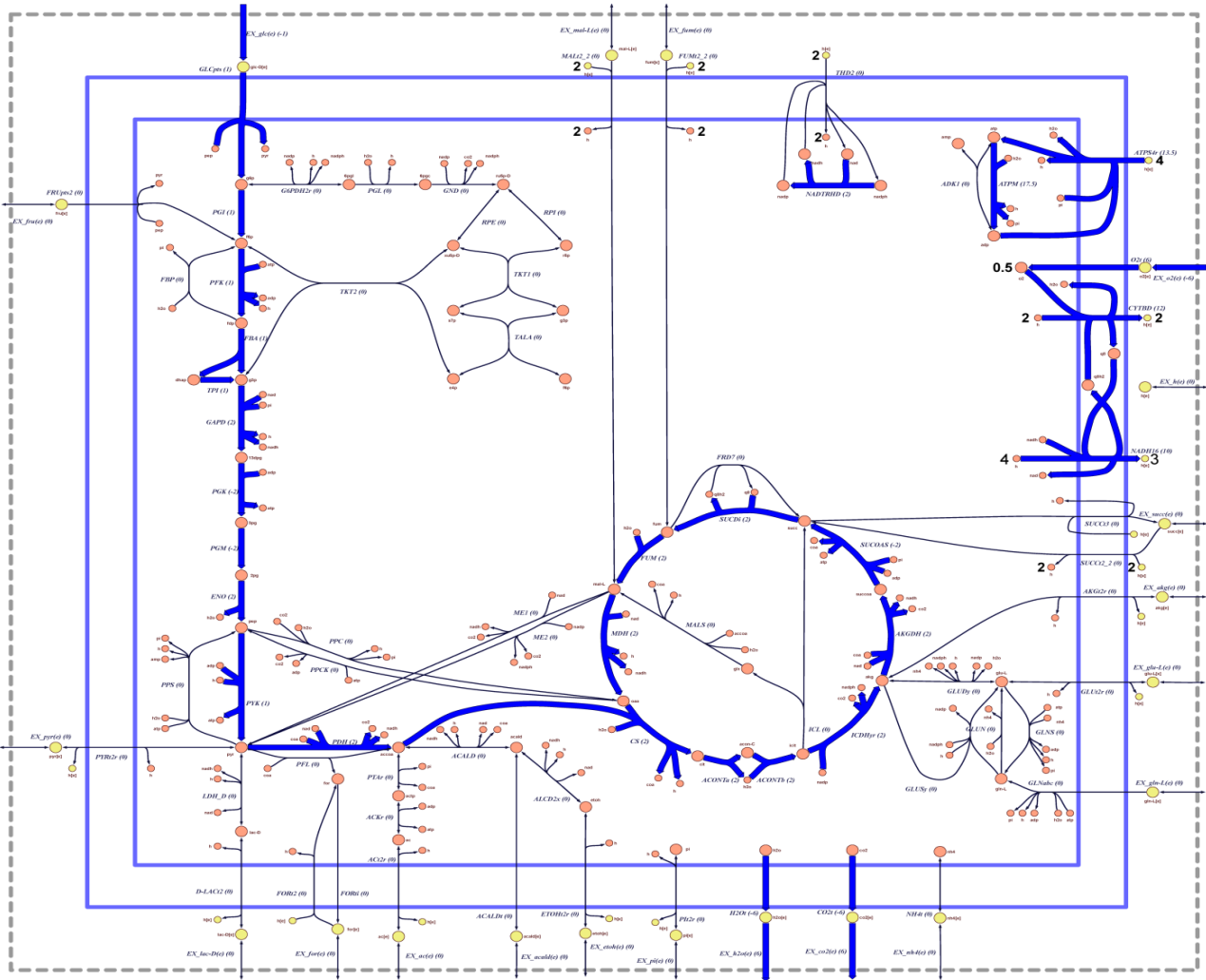
В пределах пространства C с помощью аппарата линейного программирования осуществляется поиск решения (вектора потоков), соответствующего максимуму (минимуму) целевой функции

$$Z = c_1^* v_1 + c_2^* v_2 \dots,$$

которая представляет собой некоторую взвешенную комбинацию потоков v_i . c_i^* - i -я компонента вектора весов, показывающая, насколько велик вклад i -й реакции в целевую функцию. При этом экстремум функции Z соответствует одной из вершин многогранника C . Целевая функция Z может быть предназначена для: (а) максимизации выхода биомассы культуры клеток на конкретном субстрате; (б) максимизации активности конкретного метаболитического пути; (в) минимизации активности группы целевых метаболитических путей и др.

Оптимальное
решение

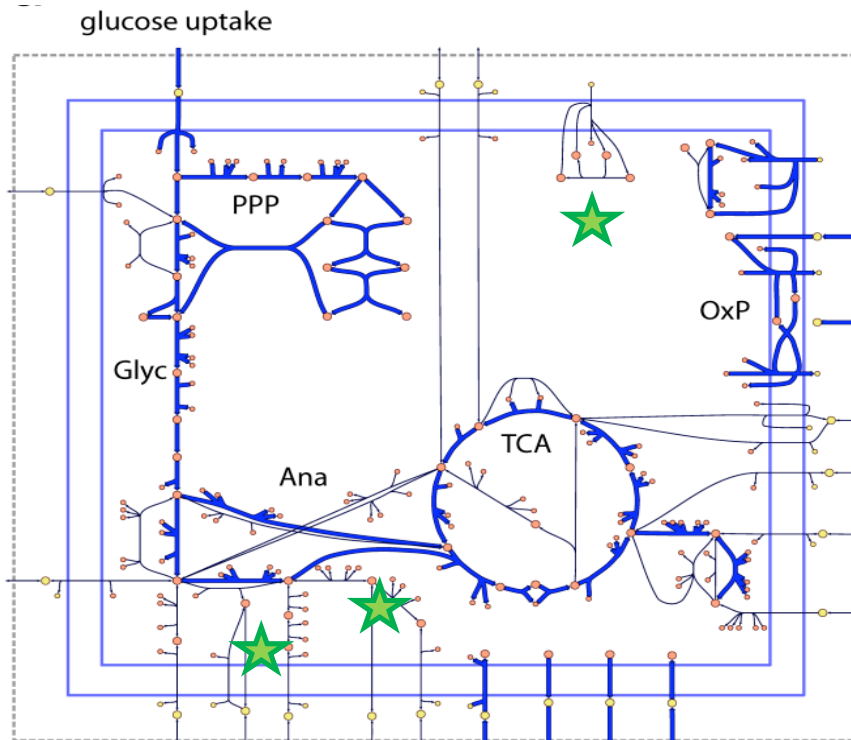
Nature Biotechnology, Supplementary Tutorial.



Толстые голубые стрелки соответствуют активным потокам

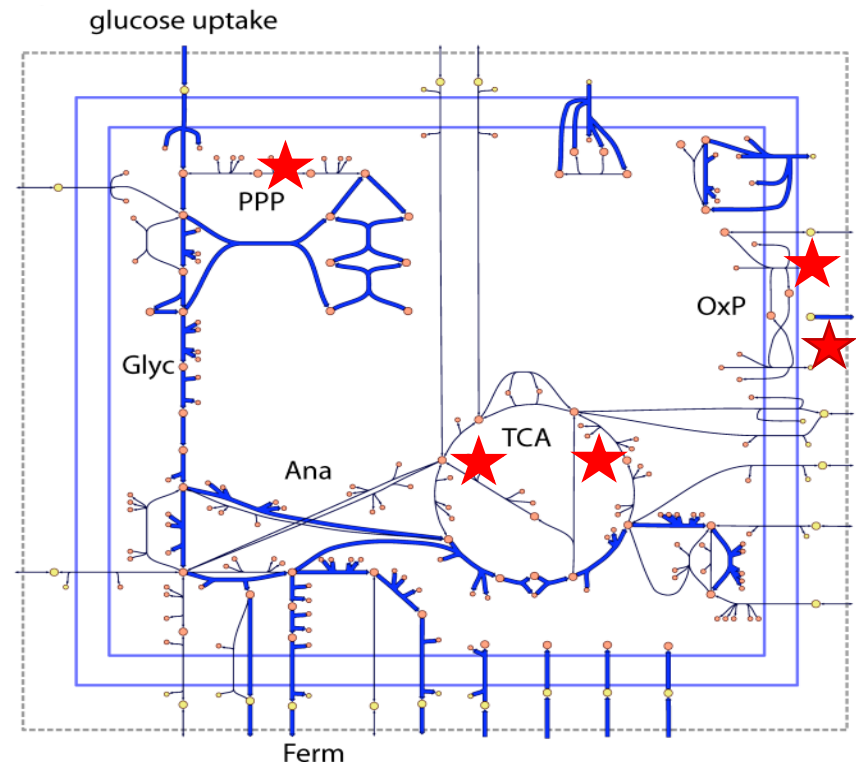
Метод FBA: распределение метаболических потоков, максимизирующее рост биомассы E.Coli в аэробных и анаэробных условиях

J.D. Orth, I. Thiele, Pallson. What is flux balance analysis, Nat. Biotechnol., 2010, 28 (3), 245 – 248;
Nature Biotechnology, Supplementary Tutorial



Аэробные условия:

- $\text{glc uptake} = 18.5 \text{ mmol gDW}^{-1}\text{hr}^{-1}$;
- $\text{O}_2\text{ uptake} = 1000$ (неограничено)
- скорость роста: **1.6531 hr⁻¹**



Анаэробные условия:

- $\text{glc uptake} = 18.5 \text{ mmol gDW}^{-1}\text{hr}^{-1}$;
- $\text{O}_2\text{ uptake} = 0$
- скорость роста: **0.4706 hr⁻¹**

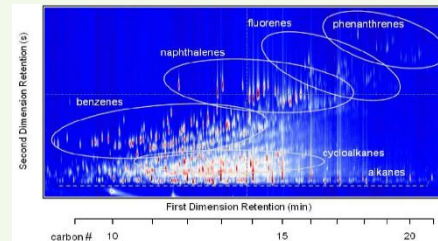
Толстые голубые стрелки соответствуют активным потокам, а тонкие стрелки – неактивным реакциям. Различия в активности метаболических путей для двух этих условий отмечены зелеными и красными звёздочками.

МЕТОДЫ АНАЛИЗА МЕТАБОЛОМНОГО СОСТАВА КЛЕТОК

(включая аминокислоты и их производные, пептиды, сахара, амингликозиды, производные сахаров, липиды, стероидные соединения, флавоноиды, жирные кислоты, алколоиды, пентозофосфаты, ароматические соединения, куйноны, каротиноиды и др.)

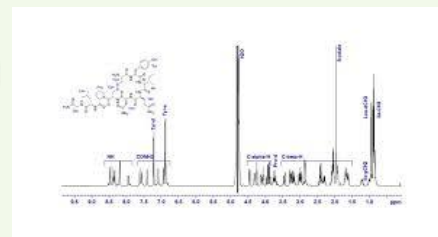
GC-MS, GCxGC-MS

Газовая хроматография – масс-спектрометрия



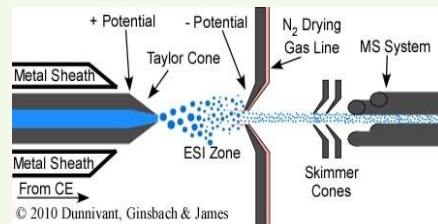
LC-MS, LC-NMR

ВЭЖХ-(масс-спектрометрия, ЯМР)



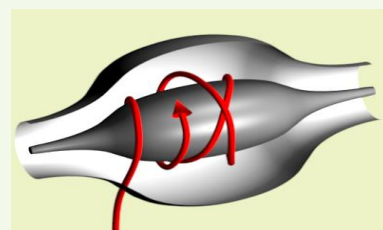
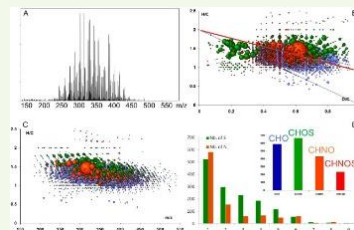
CE-MS

Капиллярной электрофорез-масс-спектрометрия



Ultra High resolution MS

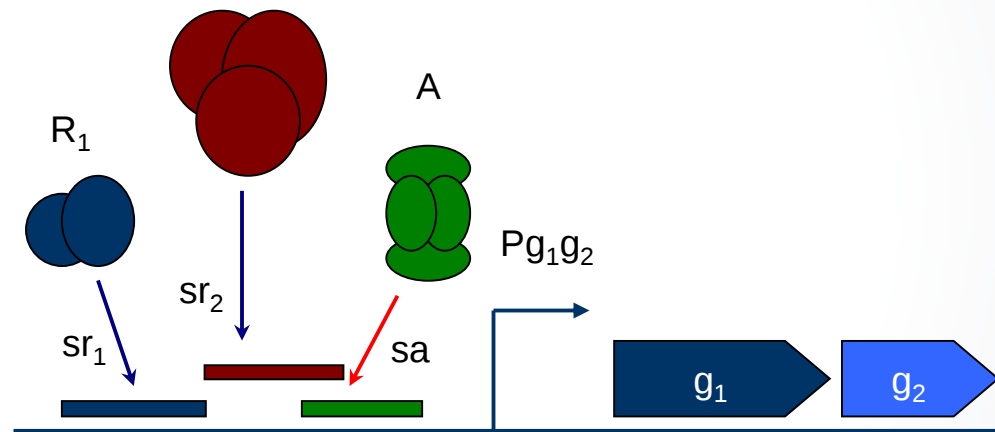
Высокоразрешающая масс-спектрометрия



Химико-кинетический подход к моделированию генетически регулируемых метаболических сетей, основанный на использовании функций Хилла (ИЦиГ СО РАН)

Пример: моделирование регуляции экспрессии оперона

Структура промотора $P_{g_1g_2}$ бицистронного оперона, регулируемого активатором (A) и двумя репрессорами (R_1 , R_2). При этом сайт связывания активатора A и репрессора R_2 пересекаются.



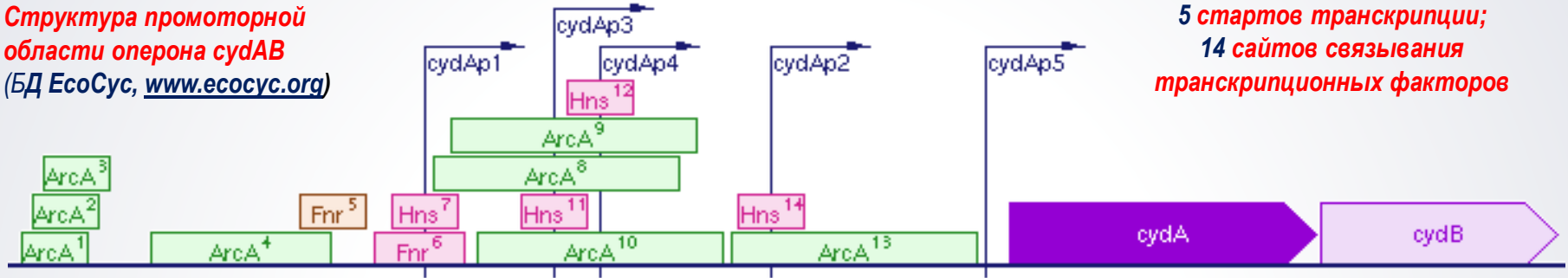
Фреймовая модель кинетики синтеза белков G_1 и G_2 , основанная на функциях Хилла

$$\frac{dG_1(t + T_{G_1})}{dt} = \frac{dG_2(t + T_{G_2})}{dt} = k_{g12} \cdot \frac{\delta_0 + \left(l_a \frac{A}{K_a}\right)^{h_a}}{1 + \left(\frac{A}{K_a}\right)^{h_a} + \left(\frac{R_1}{K_{r1}}\right)^{h_{r1}} + \left(\frac{R_2}{K_{r2}}\right)^{h_{r2}} + \frac{A^{h_a} \cdot R_2^{h_{r2}}}{(K_{ar2})^{h_a + h_{r2}}}} \star$$

Likhoshvai V., Ratushny A. "Generalized hill function method for modeling molecular processes." *Journal of bioinformatics and computational biology* (2007): 521-531.

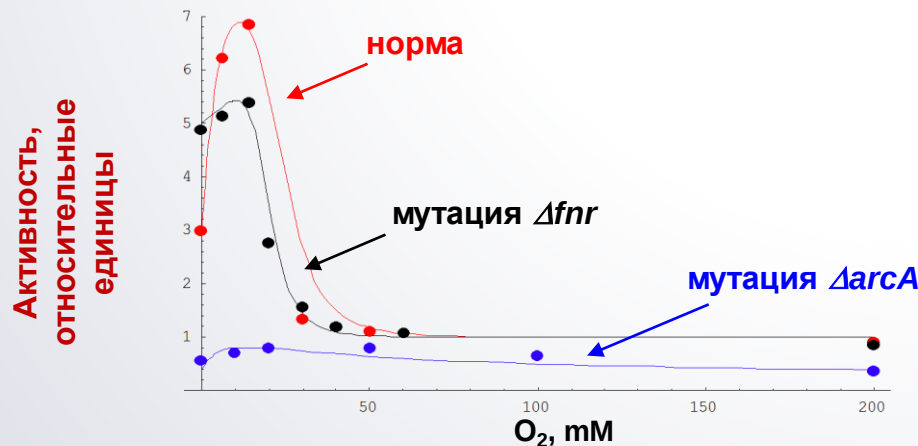
Моделирование регуляции экспрессии оперона *cydAB* в терминах обобщенных функций Хилла

Структура промоторной области оперона *cydAB*
(БД EcoСус, www.ecocysc.org)



Обобщенная функция Хилла, описывающая функционирование множественного промотора

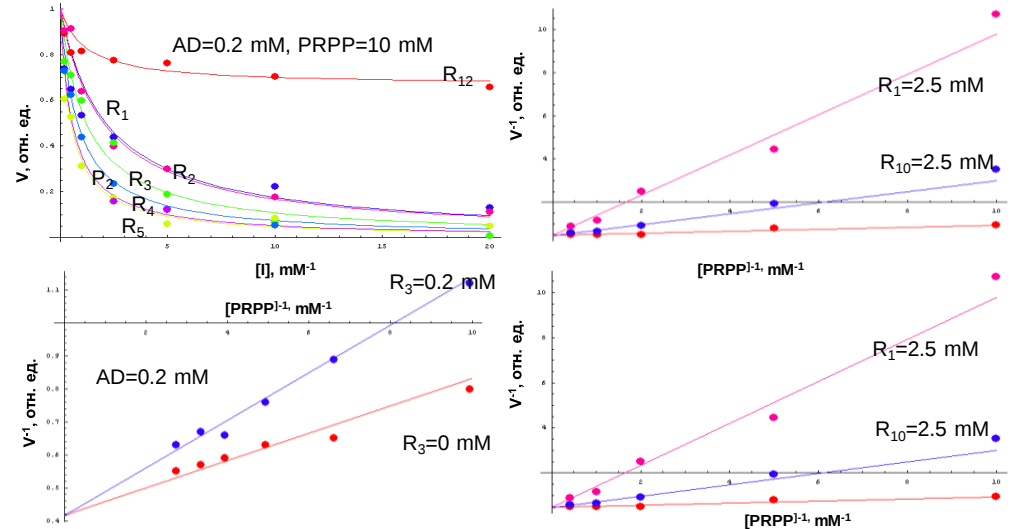
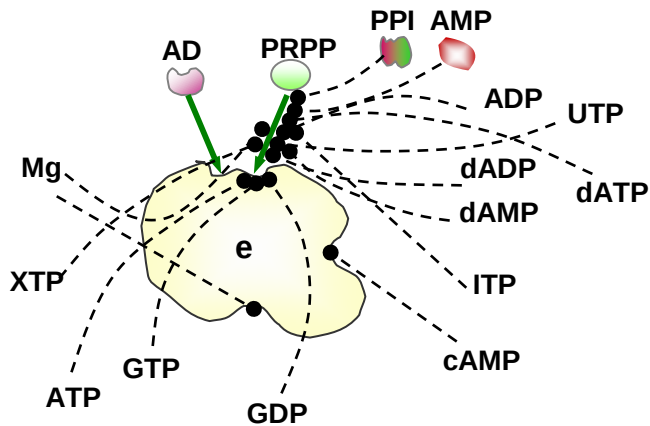
$$f_{cyd} = \frac{k_0 + \left(\frac{Fn}{k_F}\right)^{nF} + \left(\frac{A}{k_A}\right)^{nA}}{1 + \left(\frac{A}{k_A}\right)^{nA} + \left(\frac{H}{k_H}\right)^{nH} + \left(\frac{Fn}{k_{Fn}}\right)^{nFn} + \frac{Fn^{nFn1} * A^{nA1}}{k_{Fn,A}^{nFn1 + nA1}} + \frac{A^{nA2} * H^{nH1}}{k_{A,H}^{nA2 + nH1}} + \frac{H^{nH2} * Fn^{nFn2}}{k_{H,Fn}^{nH2 + nFn2}} + \frac{Fn^{nFn3} * H^{nH3} * A^{nA3}}{k_{Fn,H,A}^{nFn3 + nH3 + nA3}}}$$



- Точки на графике – экспериментальные данные из работы (Tseng et al., J Bacteriol. 1996).
- Сплошные линии – расчет по модели.

Likhoshvai V.A., Ratushny A.V. In silico cell I. Hierarchical approach and generalized hill functions in modeling enzymatic reactions and gene expression regulation. Proc. BGRS'2006 , 2006, v2. 3-18.

Моделирование в терминах функций Хилла регуляции ферментативной реакции, осуществляемой ферментом аденин-фосфорибозилтрансферазой



Точки – эксп. данные из работы (Hochstadt-Ozer & Stadtman, J. Biol. Chem., 1971).
Сплошные линии – результаты численных расчетов по мат. модели.

СУБСТРАТЫ: AD – adenine,
PRPP – phosphoribosylpyro-phosphate
ПРОДУКТЫ: PPI – pyrophosphate,
AMP – adenosine-5'-phosphate

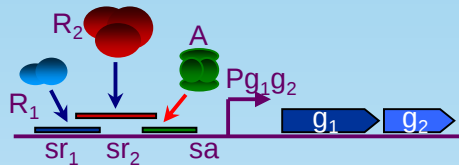
12 регуляторов:
ADP; dADP; ATP; dATP;
dAMP; GTP; ITP; XTP;
UTP; GDP; Mg; cAMP

Стационарная скорость ферментативной реакции:

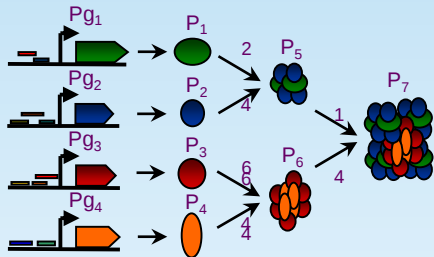
$$V = \frac{k_f \cdot e_0 \cdot \frac{S_1}{K_{m,S_1}} \cdot \frac{S_2}{K_{m,S_2}}}{\left(1 + \frac{S_1}{K_{m,S_1}}\right) \cdot \left(1 + \frac{S_2}{K_{m,S_2}} + \frac{P_1}{k_{i,P_1}} + \frac{P_2}{k_{i,P_2}} + \sum_{j=1}^5 \frac{R_j}{k_{i,R_j,S_2}} + \sum_{j=6}^{10} \left(\frac{R_j}{k_{i,R_j,S_2}}\right)^2 + \frac{R_{11}}{k_{i,R_{11},S_2}}\right)} \cdot \frac{1}{1 + kl_{R_{12}} \cdot \frac{R_{12}}{k_{R_{12}} + R_{12}}}$$

Программное обеспечение для компьютерного анализа и моделирования генетически регулируемых метаболических процессов в рамках химико-кинетического подхода (ИЦиГ СО РАН)

Модели регуляции экспрессии генов



Модели регуляции формирования активных ферментативных комплексов

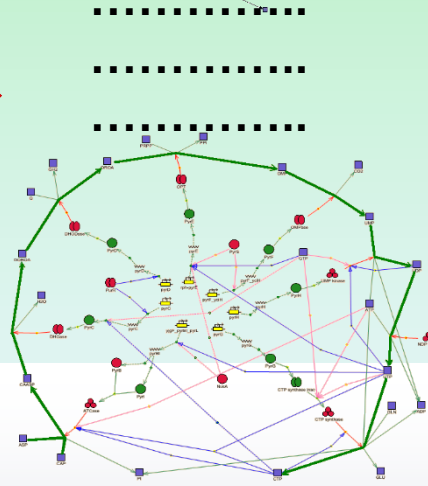
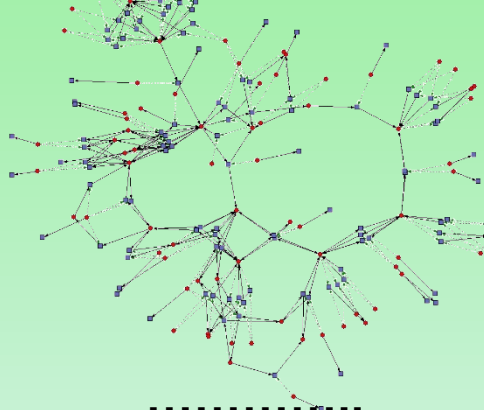


Модели регуляции ферментативных реакций



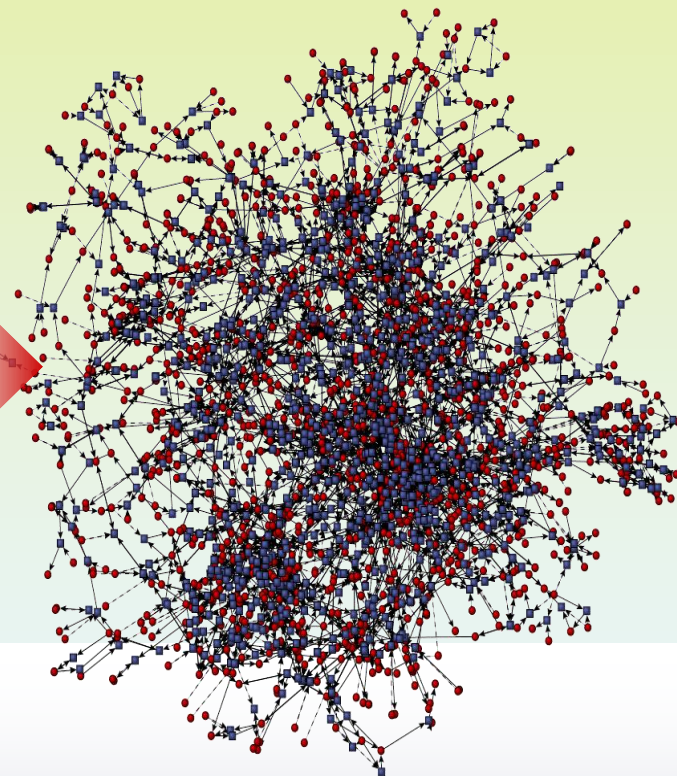
Верификация моделей элементарных подсистем

Математические модели генетически регулируемых локальных метаболических сетей



Верификация моделей локальных метаболических сетей

Математические модели крупномасштабных блоков генетически регулируемых метаболических путей

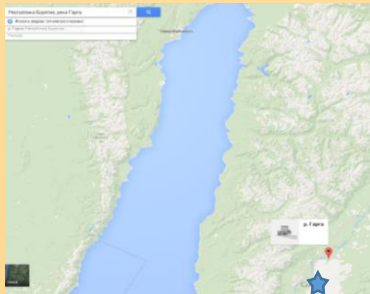


Верификация крупномасштабных моделей генных сетей и метаболических путей

Отработка ключевых блоков Платформы на примере получения штамма – продуцента молочной кислоты

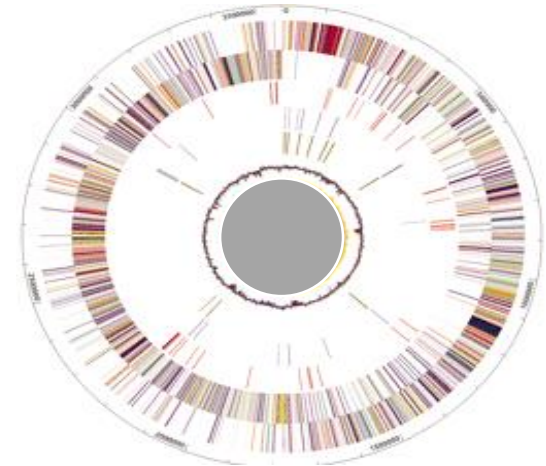
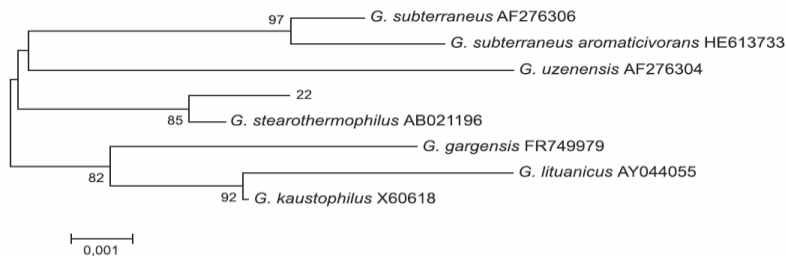
Новый термофильный штамм *Geobacillus* 22 из коллекции ИЦиГ СО РАН, перспективный для биотехнологических приложений

- широко представлен в гидротермах; представители рода способны жить в температурном диапазоне от 40 до 75°C;
- высокая скорость роста при повышенных температурах как в аэробных, так и в анаэробных условиях;
- большинство видов рода *Geobacillus* перерабатывают такие сахара, как D-глюкоза, D-ксилоза и L-арабиноза до смеси продуктов, содержащей лактат, формиат, ацетат и этанол.



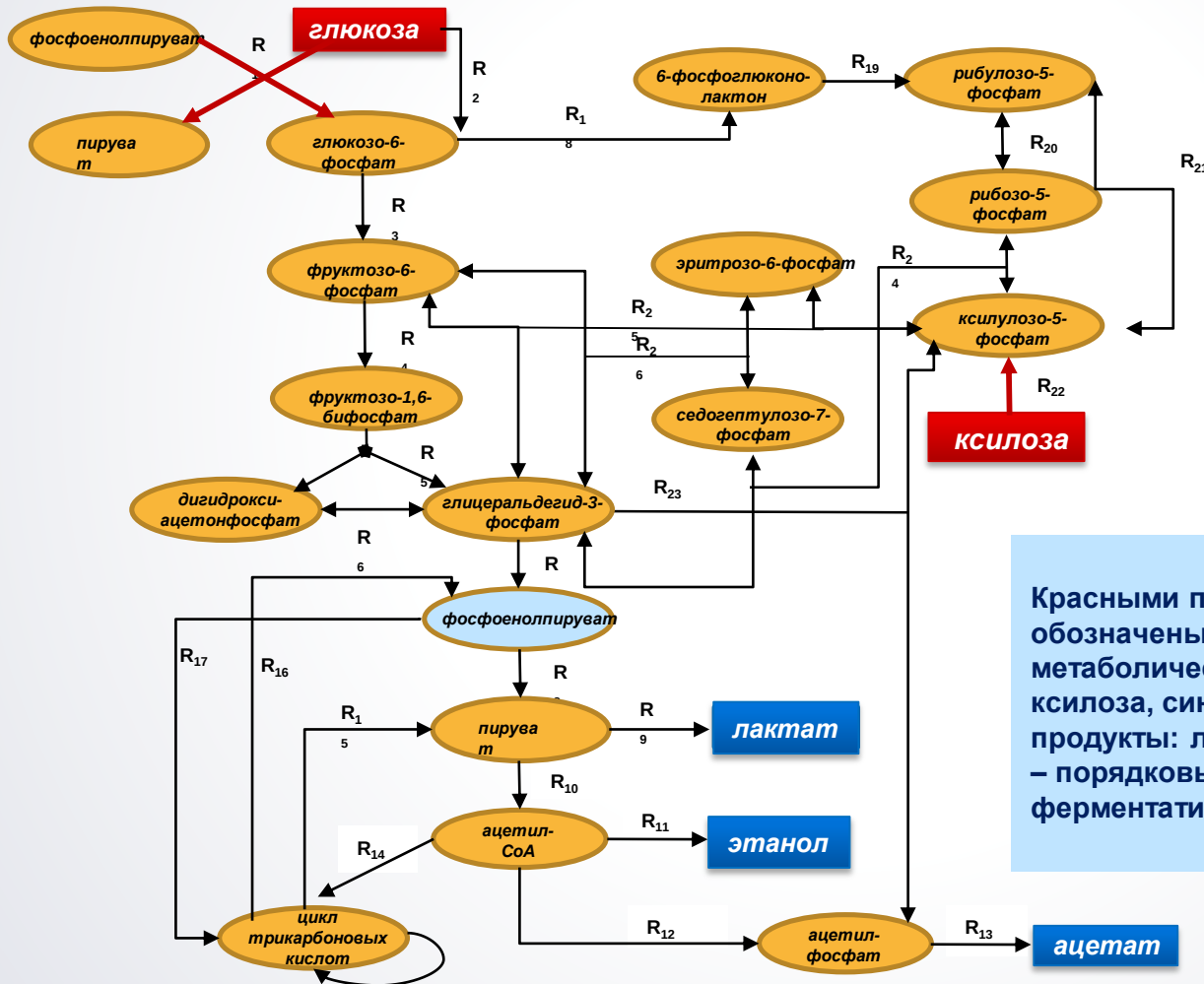
Сформированы микробиологические коллекции нового перспективного штамма *Geobacillus stearothermophilus* 22 из горячих источников Северного Прибайкалья

Сравнительный анализ по 16S рРНК показал, что штамм *Geobacillus* 22 наиболее близок к типовому штамму *G. Stearothermophilus*.



Осуществлено секвенирование генома *Geobacillus stearothermophilus*

Реконструкция гликолитического пути у штамма *G. stearothermophilus* 22:

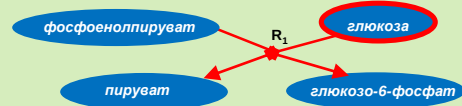


Структурная модель биосинтеза этанола, лактата и ацетата в клетке *Geobacillus spp* на основе интеграции данных KEGG и биоинформатического анализа

Красными прямоугольниками обозначены начальные субстраты в метаболической системе: глюкоза и ксилоза, синими - конечные продукты: лактат, этанол и ацетат. R_i – порядковый номер ферментативной реакции R

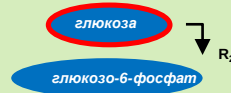
Кинетическая модель биосинтеза этанола, лактата и ацетата в клетке *G. stearothermophilus*

PTS-транспорт глюкозы в клетку



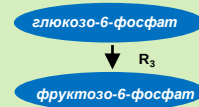
$$R_1 = \frac{V_{pts} * C_{glc} * C_{pep}}{C_{pyr} * \left(1 + \frac{C_{g6p}}{K_{pts4}}\right) * \left(\frac{C_{glc} * C_{pep}}{C_{pyr}} + C_{glc} * K_{pts3} + \frac{C_{pep} * K_{pts2}}{C_{pyr}} + K_{pts1}\right)}$$

Фосфорилирование глюкозы



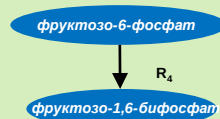
$$R_2 = \frac{V_{gik} * C_{atp} * C_{glc}}{(C_{atp} + K_{gik1}) * (C_{glc} + K_{gikm})}$$

Изомеризация глюкозо-6-фосфата во фруктозо-6-фосфат



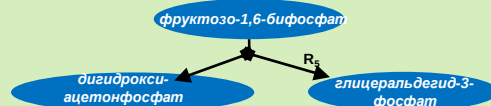
$$R_3 = \frac{V_{pgi} * \left(C_{g6p} - \frac{C_{f6p}}{K_{pgieq}}\right)}{K_{pgi1} * \left(1 + \frac{C_{g6p}}{K_{pgi3}} + \frac{C_{f6p}}{K_{pgi2} * \left(1 + \frac{C_{f6p}}{K_{pgi4}}\right)}\right) + C_{g6p}}$$

Фосфорилирование фруктозо-6-фосфат



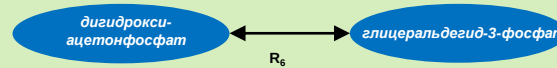
$$R_4 = \frac{V_{pfk} * C_{f6p} * K_{atp1}}{K_{atp2} * \left(C_{f6p} + \frac{K_{pfk1} * \left(K_{adp2} + \frac{C_{pep}}{K_{pfk2}}\right)}{K_{adp1}}\right) * \left(1 + \frac{L_{pfk}}{\left(1 + \frac{C_{f6p} * K_{adp1}}{K_{pfk1} * \left(K_{adp2} + \frac{C_{pep}}{K_{pfk2}}\right)}\right)^4}\right)}$$

Альдольное расщепление фруктозо-1,6-бисфосфата



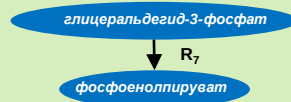
$$R_5 = \frac{V_{aldo} * \left(C_{fdp} - \frac{C_{dhap} * C_{gap}}{K_{aldoeq}}\right)}{K_{aldo1} + C_{fdp} + \frac{C_{dhap} * C_{gap}}{K_{aldoeq} * V_{blf}} + \frac{C_{dhap} * K_{aldo2}}{K_{aldoeq} * V_{blf}} + \frac{C_{fdp} * C_{gap}}{K_{aldo4}} + \frac{C_{gap} * K_{aldo3}}{K_{aldoeq} * V_{blf}}}$$

Превращение дигидроксиацетонфосфата в глицеральдегид-3-фосфат



$$R_6 = \frac{V_{tis} * \left(C_{dhap} - \frac{C_{gap}}{K_{tiseq}}\right)}{K_{tis1} * \left(1 + \frac{C_{gap}}{K_{tis2}}\right) + C_{dhap}}$$

Образование фосфоенолпирувата



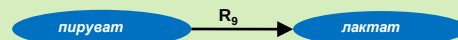
$$R_7 = \frac{V_{gapdh} * \left(C_{gap} - \frac{C_{nad} * C_{pep}}{C_{nad} * K_{gapdheq}}\right)}{\left(C_{gap} + K_{gapdh1} * \left(1 + \frac{C_{pep}}{K_{gapdh2}}\right)\right) * \left(1 + \frac{K_{gapdh3} * \left(1 + \frac{C_{nad}}{K_{gapdh4}}\right)}{C_{nad}}\right)}$$

Субстратное фосфорилирование АДФ с образованием пирувата



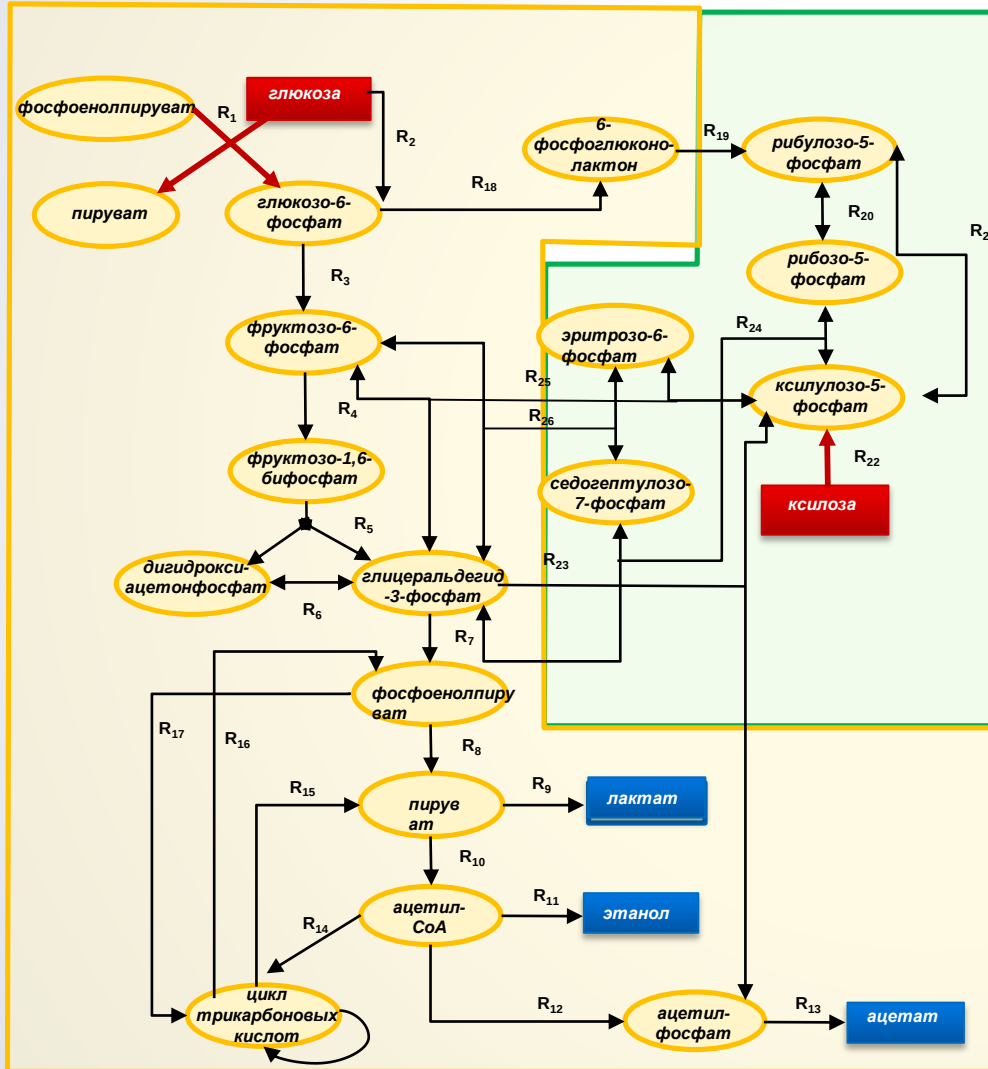
$$R_8 = \frac{V_{pyk} * C_{adp} * C_{pep} * \left(1 + \frac{C_{pep}}{K_{pyk1}}\right)^3}{K_{pyk1} * (C_{adp} + K_{pyk3}) * \left(L_{pyk} * \left(\frac{1 + \frac{C_{atp}}{K_{pyk2}}}{1 + \frac{C_{amp}}{K_{pyk4}} + \frac{C_{fdp}}{K_{pyk5}}}\right)^4 + \left(1 + \frac{C_{pep}}{K_{pyk1}}\right)^4\right)}$$

СИНТЕЗ МОЛОЧНОЙ КИСЛОТЫ



$$R_9 = \frac{V_{ldh2} * C_{pyr}}{C_{pyr} * \left(1 + \frac{C_{iac1}}{K_{ldh2i}}\right) + K_{ldh2m} * \left(1 + \frac{K_{ldh2a}}{C_{xli}}\right)}$$

Решение обратной задачи для определения параметров путей биосинтеза этанола, лактата и ацетата в клетках *Geobacillus spp* 22



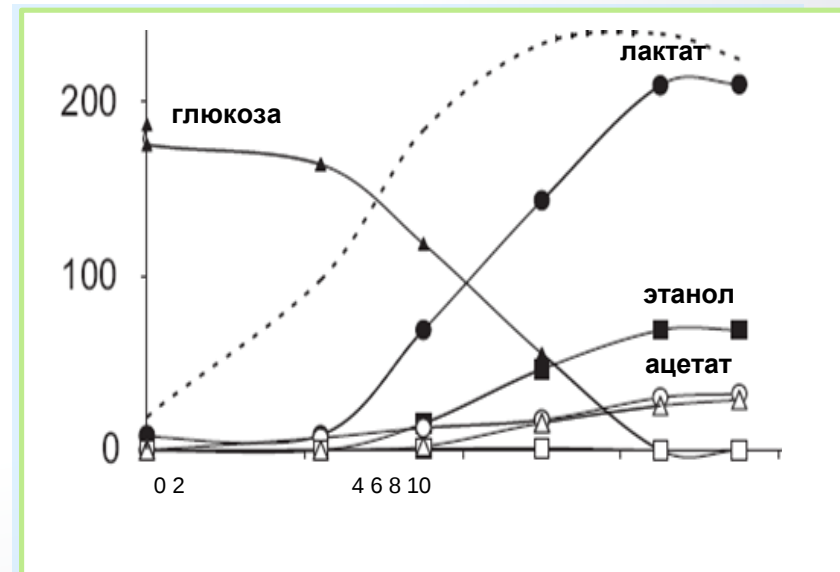
Значения 103 параметра взяты из метаболической сети *Escherichia coli*

Значения 24 параметров взяты из метаболической сети *Lactococcus lactis*



Значения 7-и параметров были неизвестны

Экспериментальная кинетика а потребления глюкозы клетками дикого типа *G. stearotherophilus* и синтеза конечных продуктов (Cripps et al., 2009)



Параметрическая обратная задача решалась с помощью программного комплекса DEEP (Санкт-Петербургский Политехнический Университет, Kozlov and Samsonov, 2011) на высокопроизводительном кластере ЦКП «Биоинформатика» ИЦиГ СО РАН

Основой комплекса является модифицированный метод дифференциальной эволюции – стохастический итеративный алгоритм многомерной математической оптимизации с применением идей генетических алгоритмов.

Геометрическая интерпретация метода в случае двух переменных



Функционал минимизации:

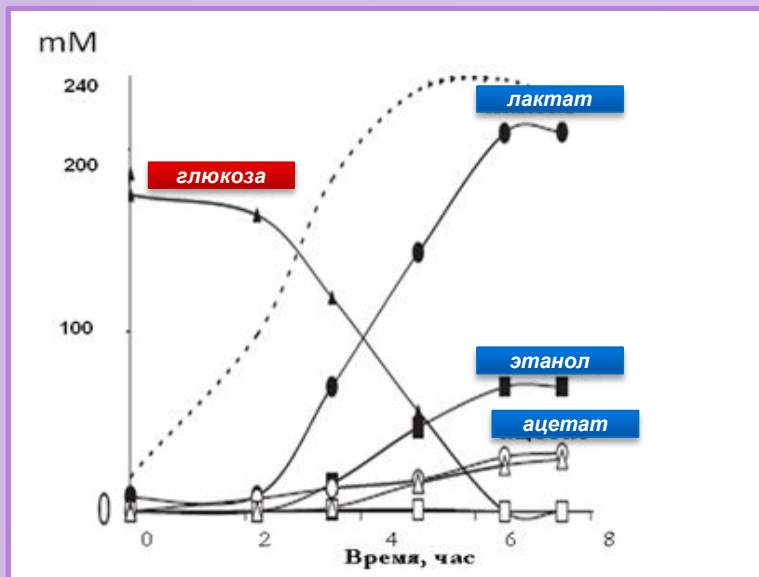
$$F = \frac{1}{2} \times \sum_{i=1}^n (y_{exp} - y_{num})^2$$

y_{exp} - экспериментальное значение переменной в точке i
 y_{num} - значение переменной в точке i , рассчитанное в модели

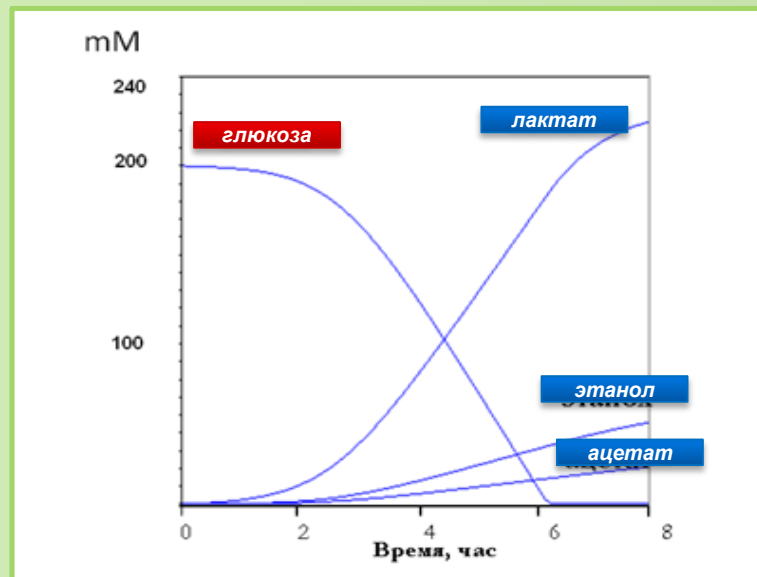
Множество векторов параметров, полученных случайным образом q_i , $i=1, \dots, NP$. Размер NP фиксирован

Экспериментальная кинетика потребления глюкозы клетками дикого типа *G. stearothermophilus* и синтеза конечных продуктов (Cripps et al., 2009) в сравнении с результатами моделирования для оптимальных значения параметров

Экспериментальные данные



Данные *in silico*

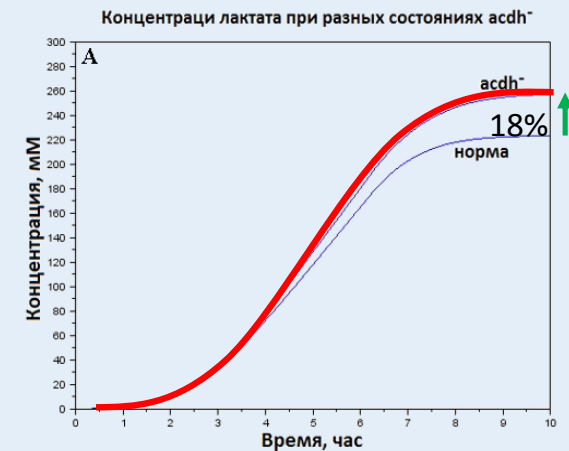
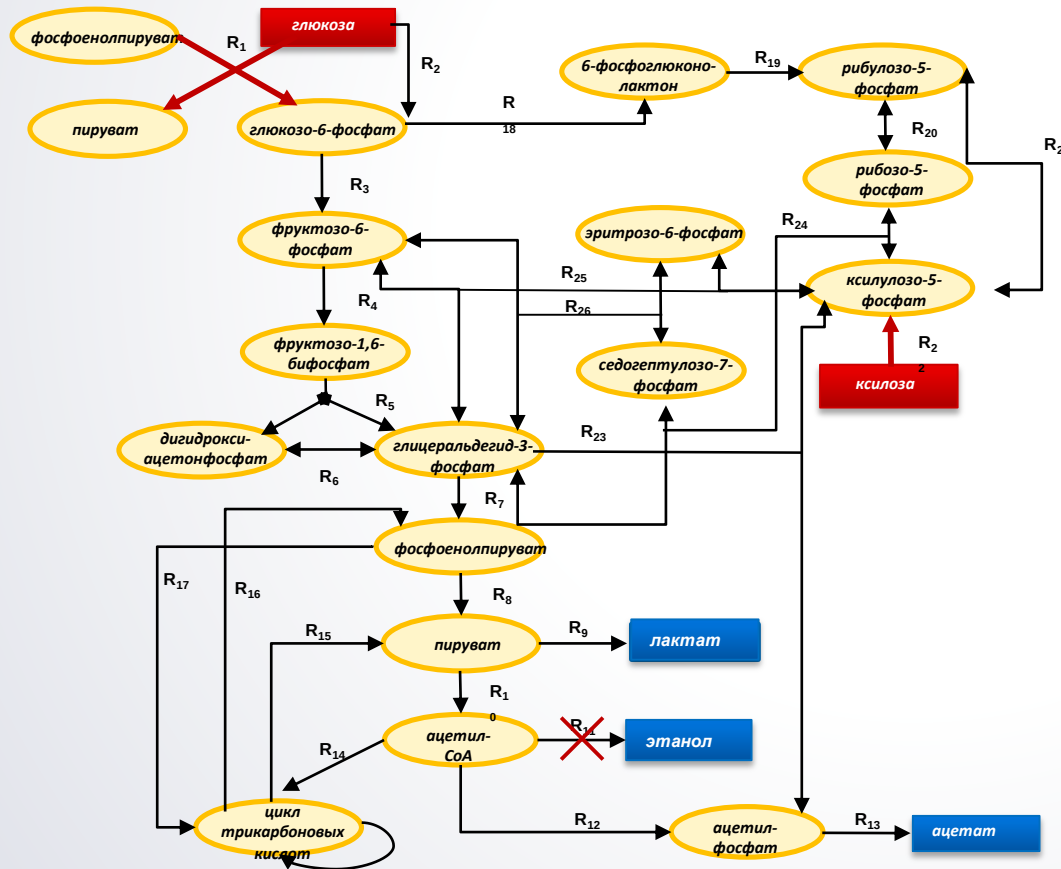


Были найдены значения констант скоростей семи указанных ниже реакций, обеспечивающие наилучшее соответствие экспериментальных и рассчитанных данных:

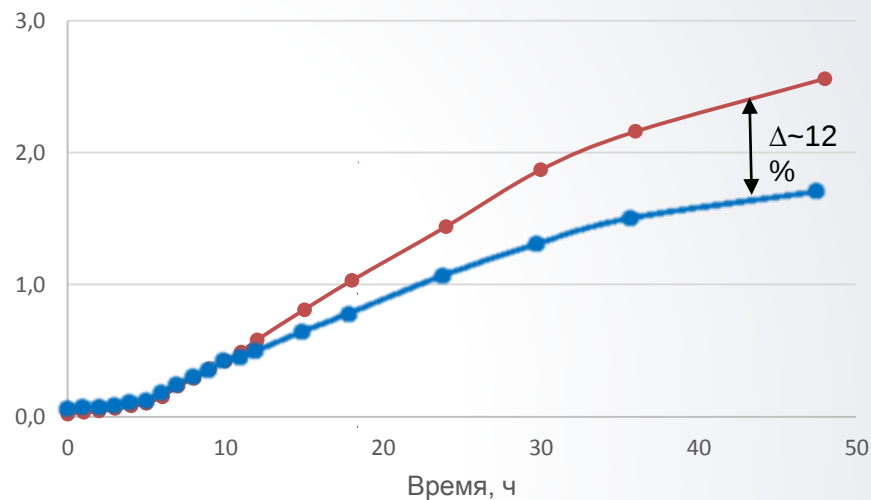
1. PTS-транспорт глюкозы (R_1)
2. фосфорилирование глюкозы (R_2)
3. Изомеризация глюкозо-6-фосфата (R_3)
4. Фосфорилирование фруктозо-6-фосфат (R_4)
5. Превращение дигидроксиацетонфосфата в глицеральдегид-3-фосфат (R_6)
6. Образование фосфоенолпирувата (R_7)
7. Субстратное фосфорилирование АДФ с образованием пирувата (R_8)

Метаболическая инженерия in silico: поиск одиночных нокаутов в геноме штамма *G. Stearothermophilus*, максимизирующих наработку молочной кислоты (лактата)

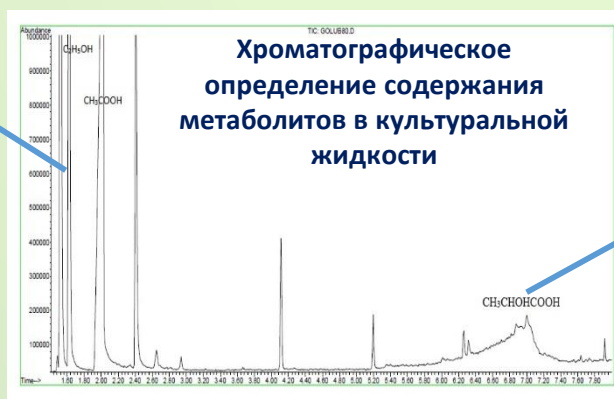
IN silico анализ показал, что повышение выхода молочной кислоты может быть достигнуто путем нокаута гена алкоголь дегидрогеназы: сравнение рассчитанной кинетики наработки лактата для штамма дикого типа и мутантного штамма с нокаутом гена *acdH*-



Биотехнологические испытания созданного штамма: наработка биоэтанола и молочной кислоты рекомбинантным штаммом *Geobacillus stearothermophilus* 22 в вихревом биореакторе «Vortex-5»



Биоэтанол

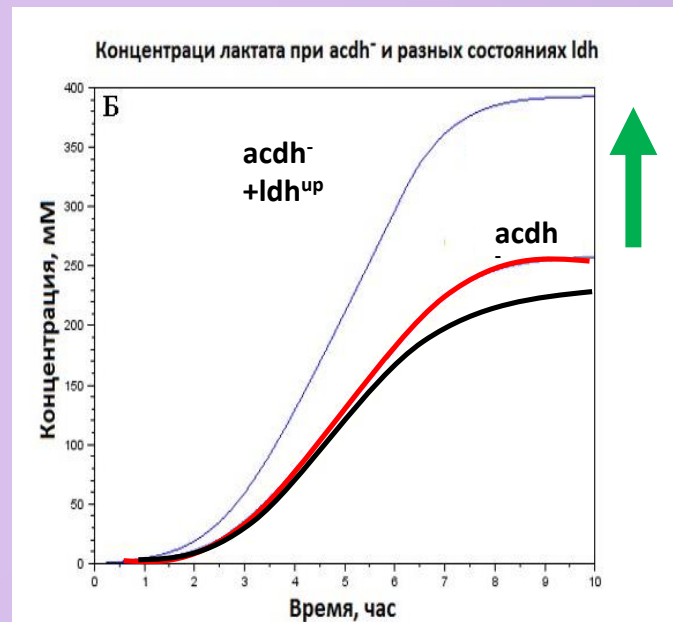
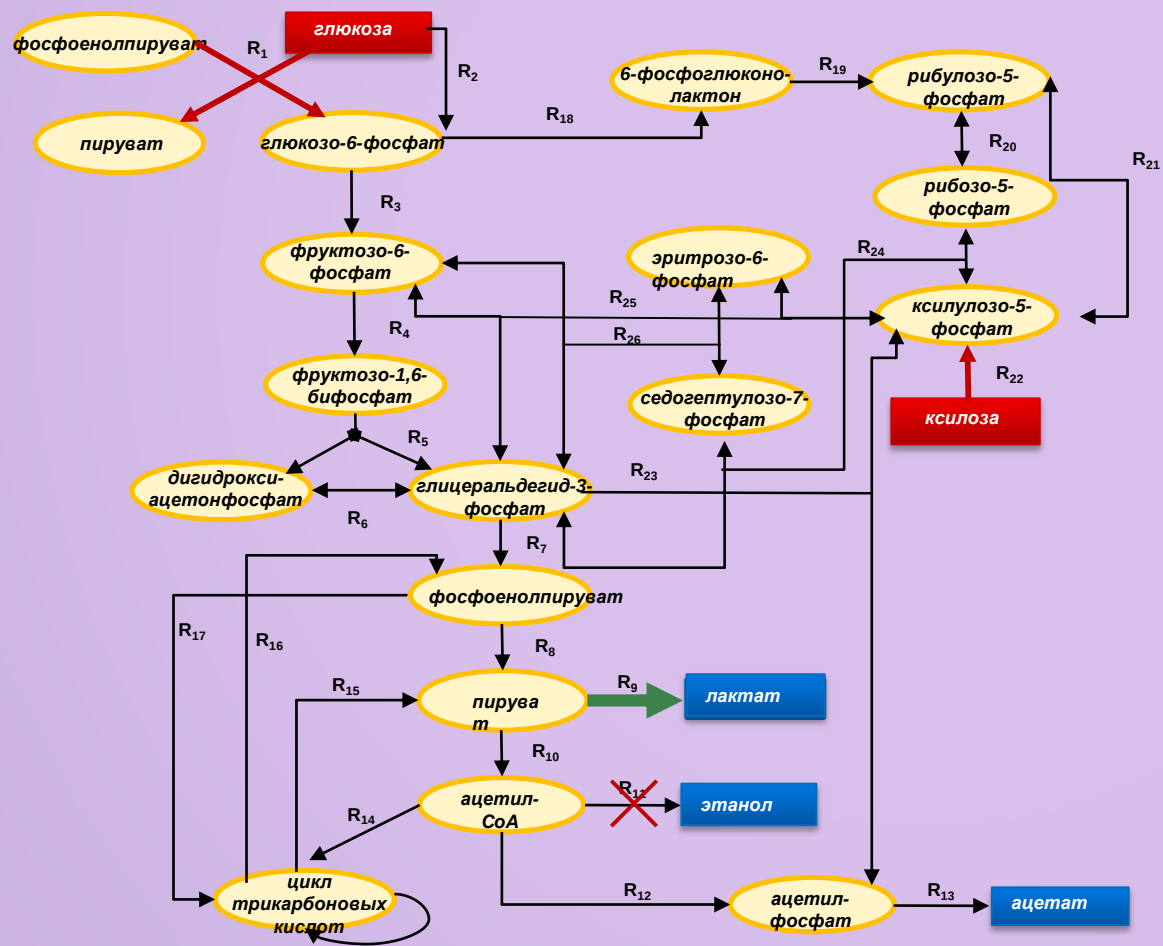


Молочная кислота

- Концентрация молочной кислоты в среде (%) при культивировании клеток дикого типа *Geobacillus stearothermophilus*;
- Концентрация молочной кислоты в среде (%) при культивировании мутантных клеток *acdH*;

***In silico* поиск комплексных изменений в геноме штамма *G. Stearothermophilus*, максимизирующих наработку молочной КИСЛОТЫ:**

помимо нокаута ген *acdH*, дополнительно активирована транскрипция гена *ldh* (за счет встраивания более активного промотора)



База математических моделей элементарных метаболических и регуляторных процессов и подсистем (ФИЦ ИЦиГ СО РАН))



MAMMOTH

Mathematical Models of biomolecular sysTems

<http://mammoth.biomodelsgroup.ru>

Structural models search

succinat Search

All objects help

scheme_name

ADS -> fumarate + AMP

succinate + a quinone <=> fumarate + a quinol

CAIR + ATP + asp -> SAICAR + ADP + Pi + H+

SAICAR -> AICAR + fumarate

IMP + asp + GTP -> ADS + GDP + Pi + H+

Structural model description: [GN0000167]

succinate + a quinone <=> fumarate + a quinol

Model consists of:

	Alias	Short name	Function	SID
	s1	succinate	substrate	SS0000191
	s2	a quinone	substrate	SS0000146
	p1	fumarate	product	SS0000180
	p2	a quinol	product	SS0000147
	r1	SQR	enzyme	SS0000327

Information (press to expand)

Mathematical models

MM0000153

$$r1 \left(-\frac{k_{catr} p1 p2}{k_{mFUM} k_{mQH2}} + \frac{k_{catf} s1 s2}{k_{mQ} k_{mSUC}} \right)$$

$$\left(1 + \frac{p1}{k_{mFUM}} + \frac{r2}{k_{iMAL}} + \frac{s1}{k_{mSUC}} \right) \left(1 + \frac{p2}{k_{mQH2}} + \frac{s2}{k_{mQ}} \right)$$

Data:

Name	Value	Units	Description
kcatf	86	1/s	Catalytic constant for forward reaction
kmSUC	2	mM	Michaelis constant for succinate
kmQ	4.17	mM	Michaelis constant for quinone
kcatr			
kmFUM			
kmQH2			

Parameters set

PS0000145

Описание
структурной
модели

Описание математической
модели

Значения параметров
модели

"MAMMOTH: a new database for curated MATHematical Models of bioMOlecular sysTems" Kazantsev F.V. *et al.* J. Bioinform. Comp. Biol. 2018. DOI: 10.1142/S0219720017400108.

В базе представлены модели ферментативных реакций и процессов регуляции эффективности экспрессии генов.

База содержит 110 элементарных моделей, каждая из которых представляет собой функцию, описывающую скорость соответствующего процесса.

Параметры моделей получены из литературных источников или подобраны при построении моделей.

Разработанный инструмент и интерфейс базы позволяют осуществлять поиск, отбор элементарных моделей и автоматическую генерацию из них более сложных моделей в формате SBML.

- Более 100 базовых моделей
- Более 5000 комплексных моделей

Моделирование функционирования и эволюции микробных сообществ: программный комплекс «Гаплоидный эволюционный конструктор» (ИЦиГ СО РАН)

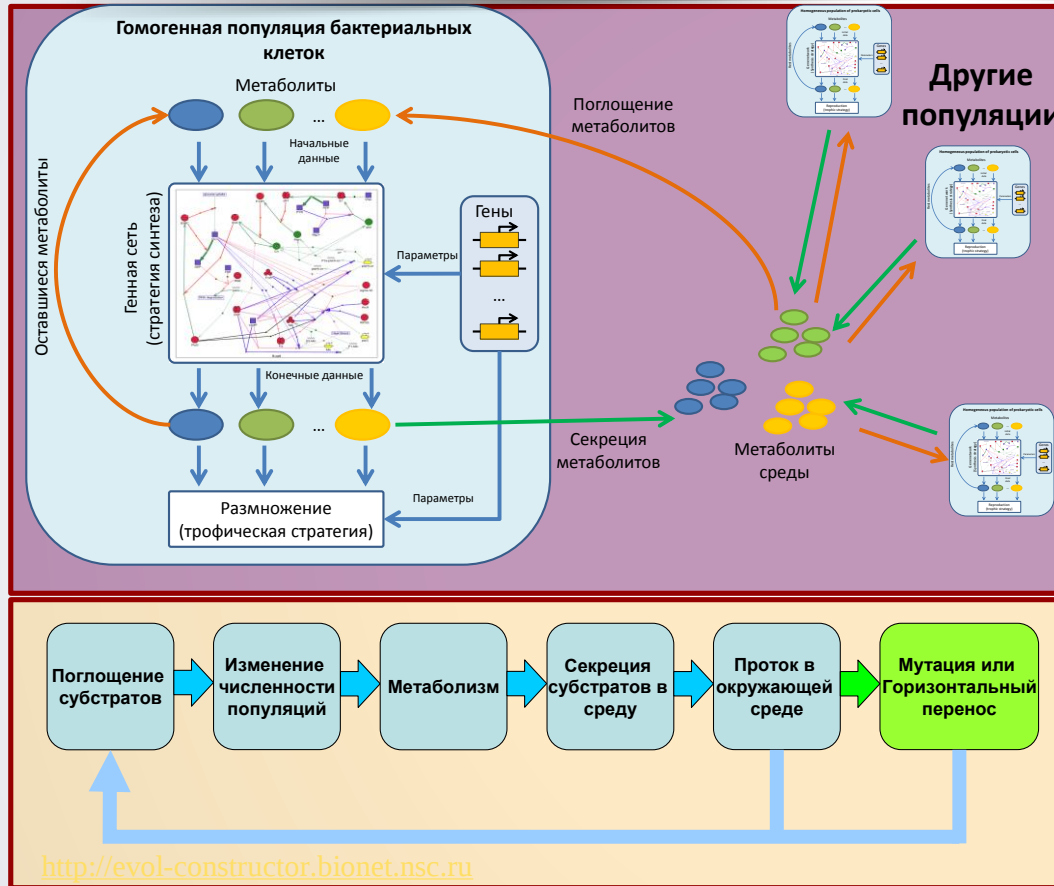
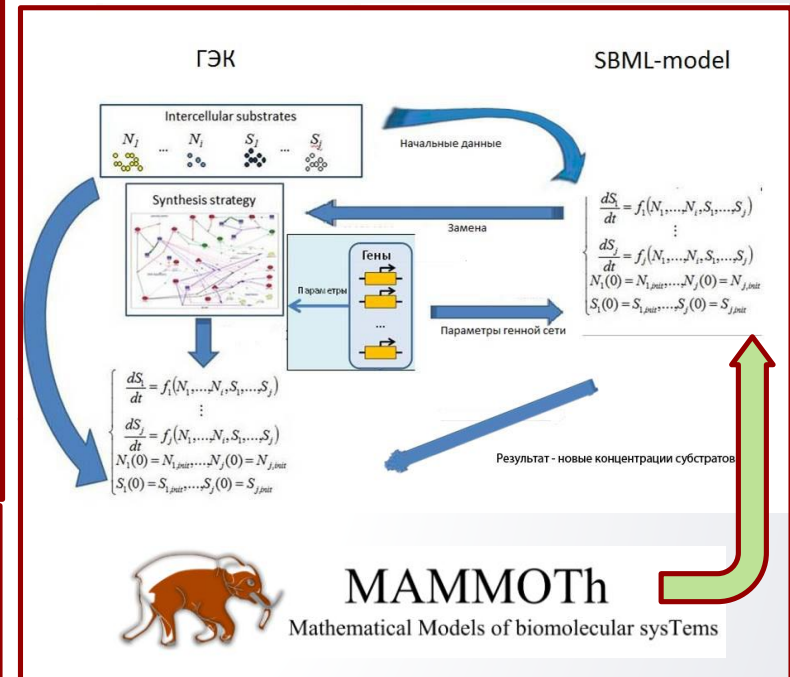


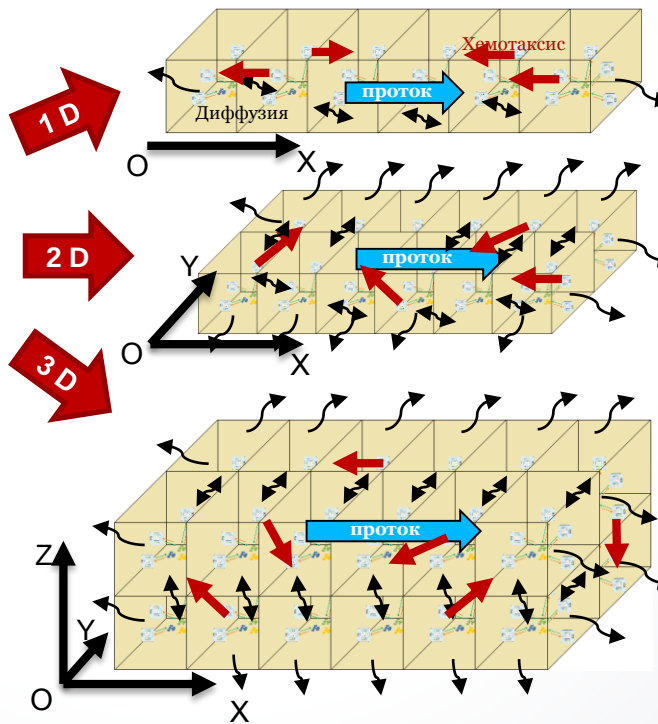
Схема интеграции ГЭК с существующими базами данных



Lashin et al. In Silico Biol. (2007)
 Lashin et al. Math. Bioinf. (2014)
 Klimenko et al. BMC Evol. Biol. (2015)

- Поглощение субстратов
- Метаболизм
- Секретия продуктов
- Самовоспроизведение популяций микроорганизмов
- Мутации
- Потеря генов
- Горизонтальный перенос генетического материала
- Генетическая регуляция
-
-

Моделирование функционирования пространственно распределенных микробных сообществ и эволюции их метагеномов (ИЦиГ СО РАН)



ПРОЦЕССЫ:

- Проток
- Диффузия
- Хемотаксис
- Поглощение субстратов
- Метаболизм
- Секреция продуктов
- Самовоспроизведение популяций микроорганизмов
- Мутации
- Потеря генов
- Горизонтальный перенос генетического материала

Klimenko et al., BMC Evol. biol., 2015

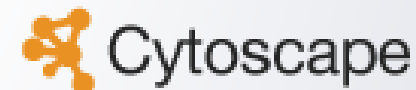
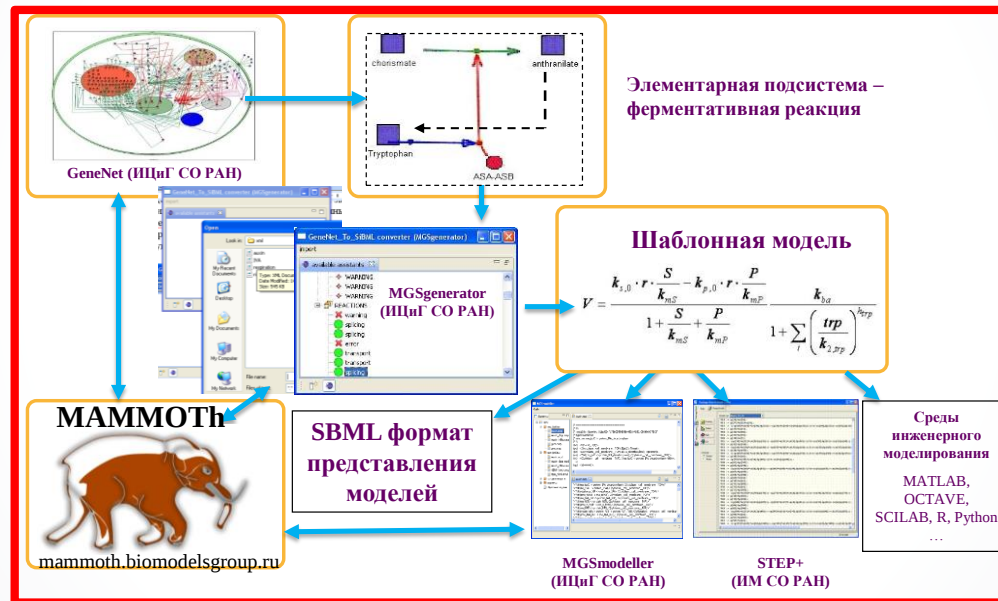
SLON == Simulations Laboratory ONline

Технология автоматического создания моделей генных сетей и метаболических путей (ИЦиГ СО РАН):

The STRING database



Michael Kuhn
EMBL Heidelberg

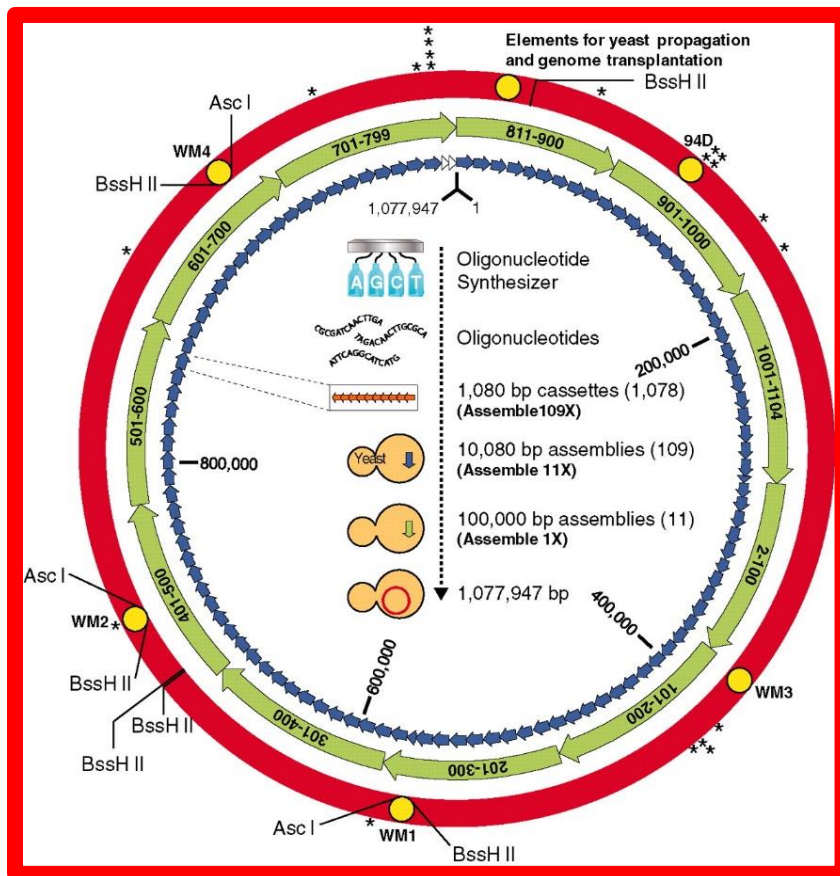


Портал
БИОРЕСУРСНЫХ
КОЛЛЕКЦИЙ



Схема сборки первого синтетического генома бактерии

Gibson, D. G., et al., Creation of a bacterial cell controlled by a chemically synthesized genome.
Science, 2010, 329(5987), 52-56.



Scienceexpress

Research Article

Creation of a Bacterial Cell Controlled by a Chemically Synthesized Genome

Daniel G. Gibson,¹ John I. Glass,¹ Carole Lartigue,¹ Vladimir N. Noskov,¹ Ray-Yuan Chuang,¹ Mikkel A. Algire,¹ Gwynedd A. Benders,² Michael G. Montague,¹ Li Ma,¹ Monzia M. Moodie,¹ Chuck Merryman,¹ Sanjay Vashee,¹ Radha Krishnakumar,¹ Nacyra Assad-Garcia,¹ Cynthia Andrews-Pfannkoch,¹ Evgeniya A. Denisova,¹ Lei Young,¹ Zhi-Qing Qi,¹ Thomas H. Segall-Shapiro,¹ Christopher H. Calvey,¹ Prashanth P. Parmar,¹ Clyde A. Hutchison III,² Hamilton O. Smith,² J. Craig Venter^{1,2*}

¹The J. Craig Venter Institute, 9704 Medical Center Drive, Rockville, MD 20850, USA. ²The J. Craig Venter Institute, 10355 Science Center Drive, San Diego, CA 92121, USA.

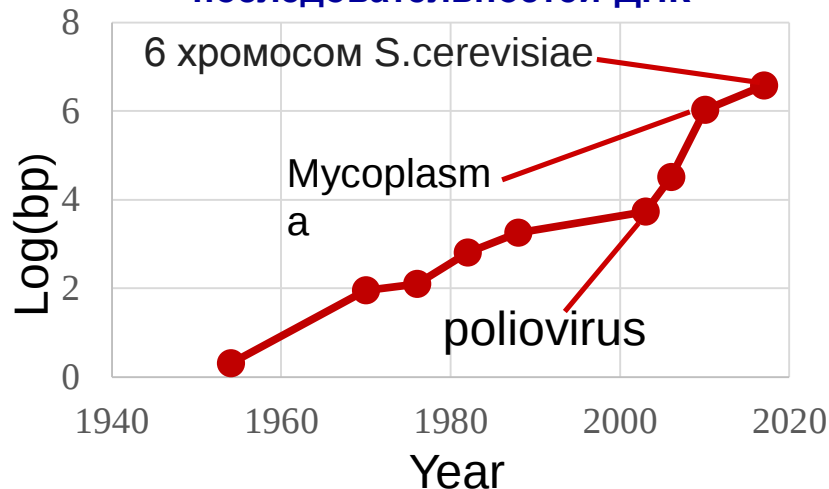
*To whom correspondence should be addressed. E-mail: jcventer@jcv.org

Геном собирали из нуклеотидов, которые были синтезированных колоночным методом.

Все последующие крупные проекты выполняются с использованием олигонуклеотидов, полученных с помощью array технологий.

Технологии синтеза генетических последовательностей

Максимальная длина синтезированных последовательностей ДНК



Технологии синтеза генетических последовательностей непрерывно развиваются. С 1950х годов они прошли большой путь от синтеза коротких олигонуклеотидов до синтеза полных геномов микроорганизмов.

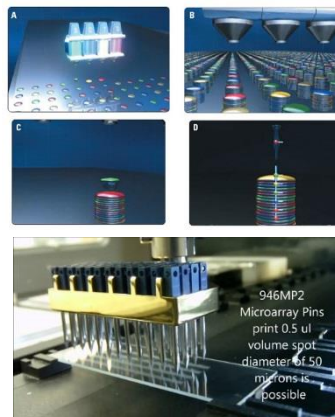
Ручной синтез



Автоматический колоночный синтез



Матричный синтез олигонуклеотидов

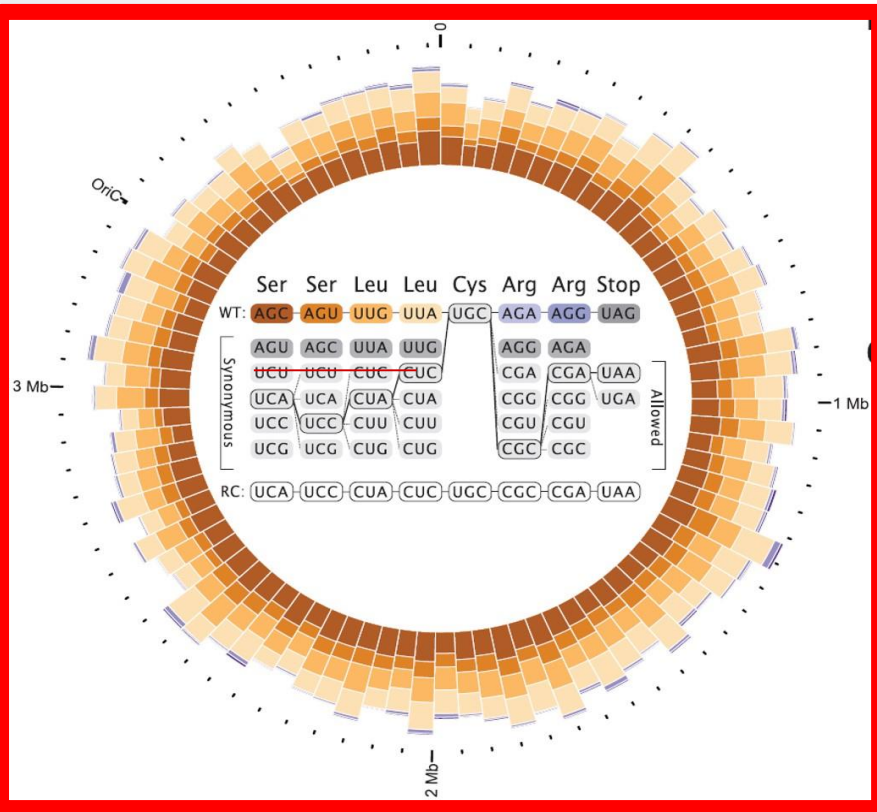


Автоматизация сборки протяженных последовательностей ДНК



Создание *E.coli* с синтетическим геномом, содержащем 57 кодонов из 64

Ostrov N., et al. Design, synthesis, and testing toward a 57-codon genome. Science, 2016, 353(6301), 819-822.



Столбиками обозначены фрагменты из которых собирали геном

В результате полного синтеза генома *E.coli* получен штамм, содержащий на 7 кодонов меньше по сравнению с диким типом.

Аналоги синтезированного генома не встречаются в природе.

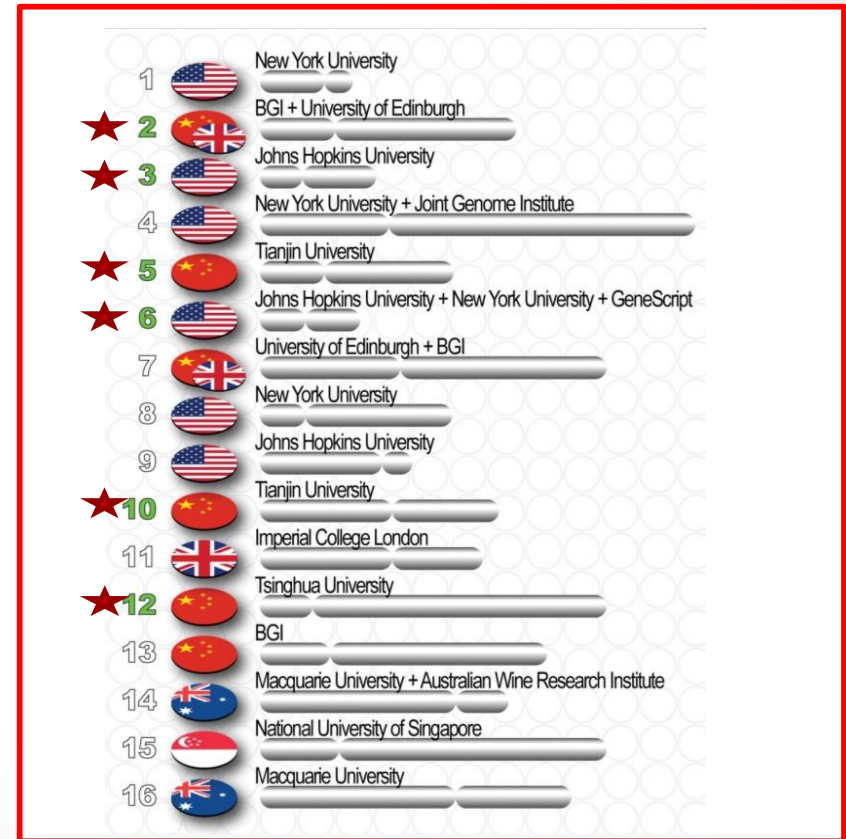
Полученный штамм не способен обмениваться генетическим материалом с окружающей средой из за меньшего числа используемых кодонов.

Это первый пример получению безопасных штаммов которые не смогут передавать полученные в лаборатории свойства в окружающую среду и не смогут получать из нее информацию для собственной эволюции.

Проект Sc2.0: конструирование синтетического генома дрожжей *S. Cerevisiae*, содержащего 6 синтетических хромосом из 16, контролирующего биологические процессы также, как природный геном *S. Cerevisiae*

Krishna Kannan and Daniel G. Gibson, *Science* 355 (6329), 1024-1025

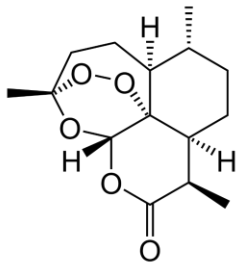
Важнейшее значение для выполнения проекта Project Sc2.0 имело создание компьютерной системы BioStudio, обеспечившей тщательный контроль за сохранением в искусственном геноме дрожжей *S. cerevisiae* генетической информации, контролирующей молекулярно-генетические процессы, реализующиеся в природном варианте *S. cerevisiae*



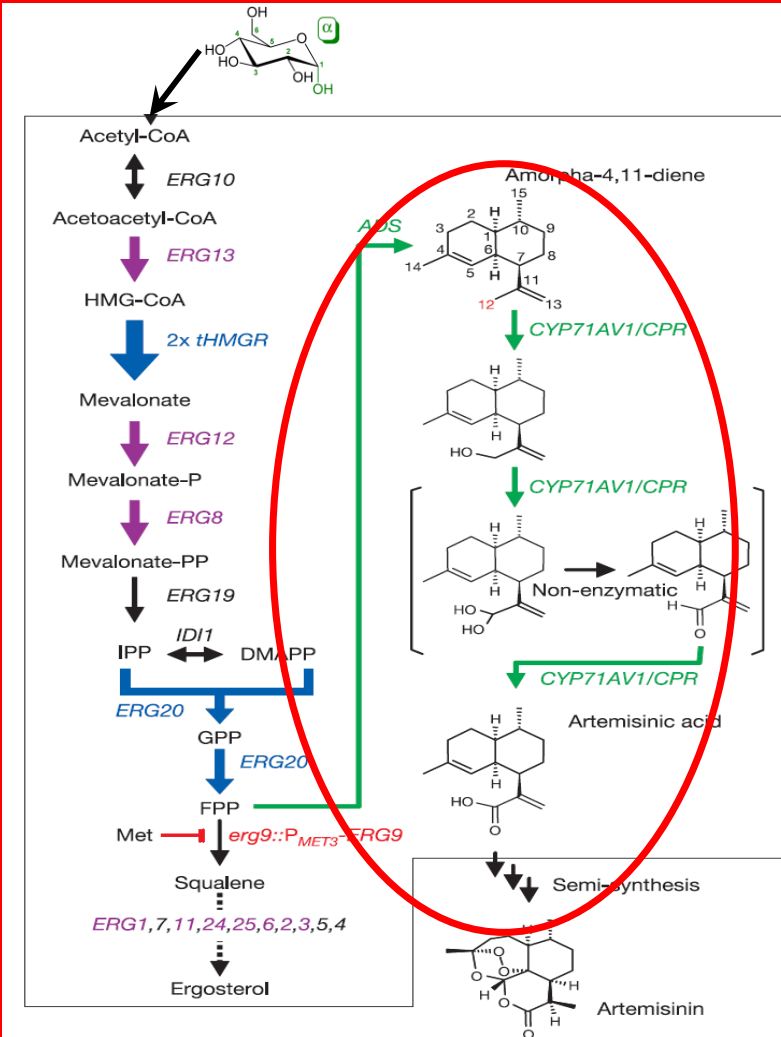
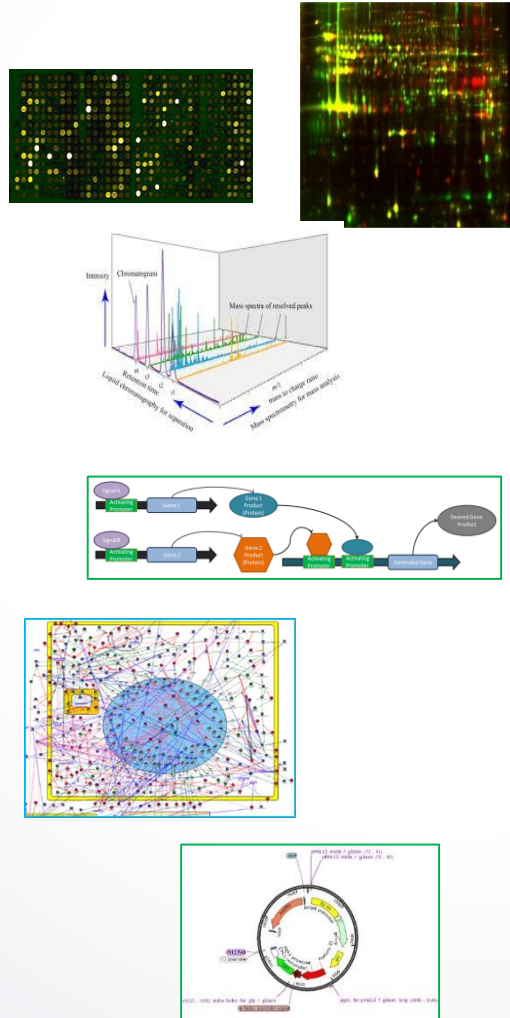
[S. M. Richardson *et al.*, *Science* 355, 1045 (2017), L. A. Mitchell *et al.*, *Science* 355, eaaf4831 (2017), G. Mercy *et al.*, *Science* 355, eaaf4597 (2017), Y. Wu *et al.*, *Science* 355, eaaf4706 (2017), W. Zhang *et al.*, *Science* 355, eaaf3981 (2017)]

Конструирование дрожжей *S. cerevisiae*, в геном которых встроены гены растения *Artemisia annua*, кодирующие ферменты метаболического пути, обеспечивающего синтез артемизина - биологически активного вещества противомаларийного действия

Ro, D. K., Paradise, et. al., Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature*, 2006, 440(7086), 940-943.



**Artemisia annua –
производитель артемизина –
биологически активного
вещества, обладающего
противомаларийным
действием**



**Благодарности сотрудникам
Отдела системной биологии
ФИЦ ИЦиГ СО РАН**



Ведущий научный сотрудник, д.б.н.
Виталий Александрович
Лихошвай



Заместитель директора
по научной работе, к.б.н.
Пельтек Сергей
Евгеньевич



Зав. сектором, к.б.н.
Лашин Сергей
Александрович



Младший научный
сотрудник Казанцев
Фёдор Владимирович



Зав. сектором, к.б.н.
Розанов Алексей
Сергеевич

