

СЕМИНАР 10

Стехиометрические модели. Определение фермента по аминокислотной последовательности. Использование баз данных для реконструкции метаболических путей. Постановка задачи линейного программирования для потоковых моделей.

Метаболизм клетки включает в себя сотни и тысячи реакций и является сложной саморегулирующейся системой. В настоящее время накопленные экспериментальные данные о метаболических процессах в различных клетках позволяют строить математические модели метаболических сетей. Как правило, один и тот же продукт биосинтеза может быть получен несколькими метаболическими путями. Математическое моделирование позволяет найти и проанализировать возможные варианты получения целевых продуктов биосинтеза и найти среди них тот, который является оптимальным для поставленной задачи. На основе результатов оптимизации и анализа модели методами генной инженерии можно вносить соответствующие изменения в генотип клетки и получать такое перераспределение метаболических потоков, которое приводит к увеличению выхода целевого метаболита.

В последние годы построены генетические карты многих простейших организмов. Расшифровка геномов разных организмов дала толчок к возникновению новой дисциплины — функциональной геномики, направленной на изучение роли определенных последовательностей генов для функционирования организма. В традиционных науках, химии белка и биохимии, появились новые направления, такие как протеомика и метаболомика.

Информация о последовательности генов собрана в различные базы данных [<http://www.brenda.uni-koeln.de>, <http://biocyc.org>], что

позволяет воспроизвести массив соответствующих метаболических путей для отдельного организма.

Метаболическое моделирование является одним из самых многообещающих подходов системной биологии для предсказания *in silico* функционирования клетки на основе взаимосвязи и взаимодействия всех клеточных компонентов.

СТЕХИОМЕТРИЧЕСКИЕ (ПОТОКОВЫЕ) МОДЕЛИ

Одним из способов описания метаболических реакций клетки является построение потоковых моделей, основанных на *допущении о стационарном состоянии концентраций метаболитов*. Стационарное состояние метаболической системы реализуется в клетках, растущих с постоянной скоростью, и соответствует экспоненциальной фазе роста. При таком подходе модель метаболизма представляет собой систему линейных алгебраических уравнений, переменными которой являются сами скорости метаболических реакций. Для построения математической модели стационарного метаболизма необходимо знать только список метаболитов, участвующих в рассматриваемой метаболической системе, и реакции, связывающие их.

Проиллюстрируем построение стехиометрической модели на простом примере. Рассмотрим схему метаболических реакций, v_1, \dots, v_6 , в результате которых исходный метаболит S_1 превращается в метаболит S_4 , являющихся конечным продуктом (рис. 10.1).

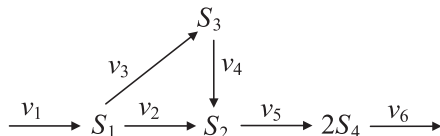


Рис. 10.1. Схема модельного метаболического пути. S_i — метаболиты, v_i — скорости реакций

Будем считать, что скорости, направленные на образование метаболита, имеют положительный знак, а скорости реакций, в которых метаболит расходуется, имеют отрицательный знак.

Построим матрицу стехиометрических коэффициентов, соответствующую приведенной схеме. Столбцы матрицы соответствуют скоростям реакций, строки — метаболитам.

$$N = \begin{pmatrix} v_1 & v_2 & v_3 & v_4 & v_5 & v_6 \\ 1 & -1 & -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & -1 \end{pmatrix} \begin{matrix} S_1 \\ S_2 \\ S_3 \\ S_4 \end{matrix}$$

Проанализируем строки матрицы.

Коэффициенты первой строки означают, что метаболит S_1 образуется в реакции v_1 и расходуется в реакциях v_2 и v_3 .

Коэффициенты второй строки означают, что метаболит S_2 образуется в реакции v_2 и v_4 и расходуется в реакции v_5 .

Коэффициенты третьей строки означают, что метаболит S_3 образуется в реакции v_3 и расходуется в реакции v_4 .

Коэффициенты четвертой строки означают, что две молекулы метаболита S_4 образуются в реакции v_5 , при этом в реакции v_6 расходуется только одна молекула метаболита S_4 .

Запишем вектор скоростей реакций:

$$v = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \\ v_6 \end{pmatrix}.$$

Тогда систему уравнений, описывающую превращение метаболитов на рис. 10.1, можно представить в матричном виде:

$$\begin{pmatrix} \frac{dS_1}{dt} \\ \frac{dS_2}{dt} \\ \frac{dS_3}{dt} \\ \frac{dS_4}{dt} \end{pmatrix} = \begin{pmatrix} 1 & -1 & -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & -1 \end{pmatrix} \cdot \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \\ v_6 \end{pmatrix}. \quad (10.1)$$

Уравнения (10.1) называются **уравнениями баланса**, поскольку они отражают все процессы, в результате которых метаболит образуется или расходуется. В стационарном состоянии все потоки, направленные на образование конкретного метаболита, уравновешены потоками, в результате которых концентрация этого метаболита уменьшается. Сумма всех потоков, участвующих в изменении концентрации данного метаболита, таким образом, равна нулю. Для системы дифференциальных уравнений это означает, что сумма всех скоростей реакций, в которых участвует данный метаболит, равна нулю. Тогда все правые части уравнений приравняются нулю, и система дифференциальных уравнений вырождается в систему алгебраических уравнений:

$$\begin{pmatrix} 1 & -1 & -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & -1 \end{pmatrix} \cdot \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \\ v_6 \end{pmatrix} = 0. \quad (10.2)$$

Умножая матрицу стехиометрических коэффициентов на вектор скоростей, получим систему алгебраических уравнений, являющейся системой **уравнений баланса** для стационарного состояния:

$$\begin{cases} v_1 - v_2 - v_3 = 0, \\ v_2 + v_4 - v_5 = 0, \\ v_3 - v_4 = 0, \\ 2v_5 - v_6 = 0. \end{cases} \quad (10.3)$$

ЗАДАЧА ЛИНЕЙНОГО ПРОГРАММИРОВАНИЯ

Полученная система включает 4 уравнения и 6 переменных, поэтому является *недоопределенной*. Такая система имеет бесконечное число решений и не может быть решена однозначно. Однако в исходном пространстве переменных можно выделить такое подпространство, в котором мы сможем найти максимум или минимум функционала, определяемого из нашей системы, и поставить **задачу линейного программирования (задачу оптимизации)**: задать *допустимое множество, целевую функцию* и критерий поиска (ищем максимум или минимум).

Такое пространство выделяется естественным образом, поскольку величины потоков всегда ограничены положительными значениями. Для оптимизации выбирается целевая функция, обычно это один или несколько потоков, суммарная величина которых по биологическому смыслу задачи должна быть максимальной (минимальна). Пусть в рассматриваемой схеме нас интересует максимальный выход конечного продукта S_4 , определяемый скоростью v_6 .

Тогда задача линейного программирования для рассматриваемой схемы будет выглядеть следующим образом.

Найти максимум *целевой функции* z ,

$$z = v_6 \rightarrow \max ,$$

при следующих условиях:

$$\begin{cases} v_1 - v_2 - v_3 = 0, \\ v_2 + v_4 - v_5 = 0, \\ v_3 - v_4 = 0, \\ 2v_5 - v_6 = 0, \end{cases} \quad (10.4)$$

$$0 \leq v_1, \dots, v_6 < \infty.$$

Рассмотрим, каким образом, используя базы данных, можно построить потоковую модель и поставить задачу оптимизации для поиска решений модели.

Задание 10.1

Работа с базой данных BLAST по сравнению последовательностей.

Скопируйте в буфер обмена одну из последовательностей, предложенных преподавателем.

Откройте базу данных BLAST (<http://blast.ncbi.nlm.nih.gov>).

В разделе Web BLAST выберите Protein Blast.

В открывшемся окне “Enter accession number(s), gi(s), or FASTA sequence(s)” введите скопированную аминокислотную последовательность.

Пример аминокислотной последовательности:

```
MAALTRDPQFQKLQQWYREHRSELNLRRLFDANKDRFNHF
SLTLNTNHHGILVDYSKNLVTEDVMRMLVDLAKSRGVEAARER
MFNGEKINYTEGRAVLHVALRNRNTPILVDGKDVMPVKNVL
DKMKSFCQVRVSGDWKGYTGKTITDVINIGIGGSDLGPLMVTEA
LKPYSGGPRVWYVSNIDGTHIAKTLAQLNPESLFIASKTFTTQ
ETITNAETAKEWFLQAAKDPSAVAKHFVALSTNTTKVKEFGIDPQ
NMFEFWDWVGGRYSLWSAIGLSIALHVGFDFNEQLLSGAHWMD
QHFRTTPEKNAPVLLALLGIWYINCFGCETHAMLPYDQYLHRF
AAAYFQQGDMESNGKYITKSGTRVDHQTGPIVWGEPTNGQHAF
YQLIHQGTKMIPCDLIPVQTQHPIRKGLHKKILLANFLAQTEALM
RGKSTEEARKELQAAGKSPEDLERLLPHKVFEGNRPTNSIVFTKL
TPFMLGALVAMYENKIFVQGIWDINSFDQWGVELGKQLAKKIEP
ELDGSAQVTSHDASTNGLINFIKQQREARVQ
```

Дождитесь результата обработки последовательности, который появится в таблице “Sequences producing significant alignments”, и выпишите результат с наилучшим выравниванием (в верхней строке).

Для приведенного примера последовательности это glucose-6-phosphate isomerase isoform 3.

ЗАДАНИЕ 10.2

Работа с базой данных по метаболическим путям — BioСус.

Откройте базу данных BioСус (<http://biocус.org/>).

Введите в строку поиска название фермента (например, glucose-6-phosphate isomerase) и проанализируйте результат.

Выпишите номер исследуемого фермента в системе ЕС (Enzyme Commission system) и реакцию, которую он катализирует.

Для рассматриваемого примера, glucose-6-phosphate isomerase, его номер в ЕС: 5.3.1.9.

Реакция: D-glucopyranose 6-phosphate \rightleftharpoons β -D-fructofuranose 6-phosphate.

В таблице “Biological Process” выберите один из метаболических путей, в котором участвует фермент, например гликолиз (glycolysis).

Найдите путь гликолиза. Для этого используйте опции
Search

Search Pathways

В строку поиска введите название пути — glycolysis.

ЗАДАНИЕ 10.3

Построение стехиометрической модели.

Зарисуйте часть найденного пути, используя сокращения для названий метаболитов.

В полученной схеме реакций обозначьте скорости реакций как $v_1, v_2 \dots$.

Постройте стехиометрическую матрицу, соответствующую схеме реакций.

Запишите уравнения баланса.

Сколько переменных в системе?

Сколько уравнений в системе?

Сколько решений может иметь задача?

ЗАДАНИЕ 10.4

Постановка задачи оптимизации.

Пусть задачей исследования является увеличение выхода ментола (рис. 10.2).

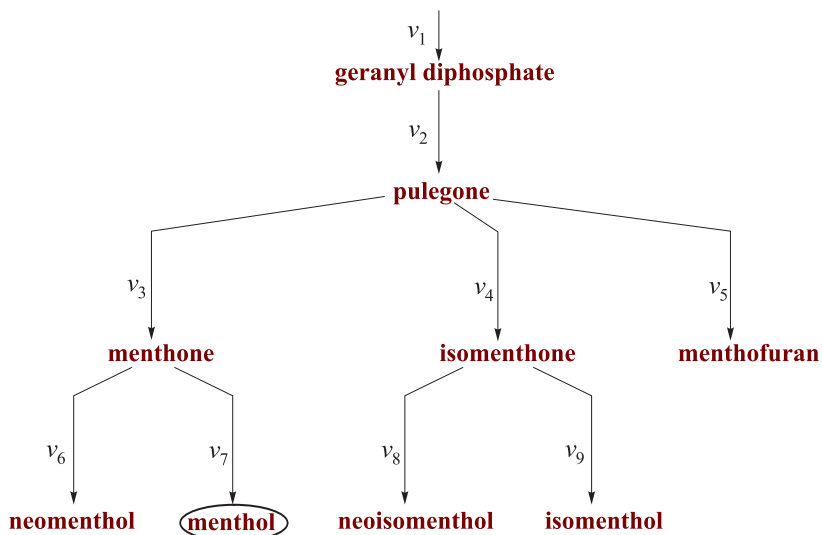


Рис. 10.2. Схемы метаболического пути синтеза ментола

Для заданной схемы синтеза ментола сформулируйте задачу линейного программирования (задачу оптимизации):

- запишите уравнения баланса,
- запишите ограничения на потоки,
- запишите целевую функцию.

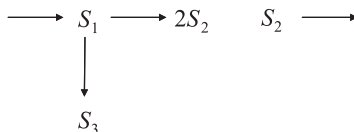
КОНТРОЛЬНЫЕ ВОПРОСЫ

1. Какие базы данных можно использовать для построения метаболических моделей?
2. Как строится матрица стехиометрических коэффициентов по метаболическому пути?
3. Какие приближения используются для решения уравнений баланса?
4. В каких случаях ставится задача линейного программирования?
5. Какие условия необходимо сформулировать для постановки задачи линейного программирования?

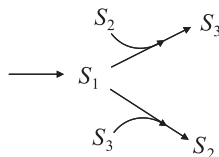
ЗАДАЧИ К СЕМИНАРУ 10

10.1. Запишите уравнения баланса для следующих метаболических путей:

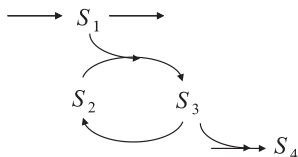
а)



б)



в)



10.2. Поставьте задачу линейного программирования для метаболических путей из задания 10.1, если целевым продуктом является метаболит с наибольшим индексом.