

Лекция 3

Слайд 1

Название лекции

Математическое моделирование динамики молекулярно-генетических систем

Слайд 2

Цели лекции

- Обобщенный химико-кинетический метод моделирования (формальное представление моделей)
- Спецификации языка моделирования
- Процесс сборки математических моделей
- Моделирование многокомпартментных структур
- Моделирование ДНК- и РНК-специфичных процессов
- Моделирование полиплоидных генных сетей
- Различные средства в рамках ОХКММ (заключение)

Слайд 3

Динамика молекулярно-генетических систем в ОХКММ формально описывается моделями вида:

$$\frac{dx_i}{dt} = F_i(X, K) - x_i G_i(X, K)$$

где n – количество динамических переменных модели,

$X = (x_1, \dots, x_n)$ – вектор динамических переменных, как правило имеющих смысл концентраций,

$K = (k_1, \dots, k_m)$ – вектор параметров,

$F_i(X, K)$ и $G_i(X, K)$ – рациональные функции, принимающие неотрицательные значения при всех неотрицательных значениях переменных и параметров.

Слайд 4 и 5

Формализация в рамках ОХКММ осуществляется на основе блочного подхода, который подразумевает проведение иерархической декомпозиции исходной системы до элементарных структур и процессов, описание элементарных процессов в терминах элементарных блоков с последующим конструированием из элементарных процессов моделей элементарных структур, из которых в дальнейшем проводится сборка моделей генетической системы и ее вариантов.

Слайды 6, 7, 8, 9

Каждое элементарное событие в рамках ОХКММ описывается посредством формализованных блоков изолировано от других процессов, т.е. выделяется в самостоятельный блок.

Спецификация структуры элементарного блока:

[*/]

[/*]

[{bb}] [] <имя блока> [].

[]

[{pp}] [] <имя продукта 1> [] <имя продукта 2> [] ... <имя продукта n> [].

[]

[{ee}] [] <имя расширения 1> [] <имя расширения 2> [] ... <имя расширения k> [].

[]

[{ss}] [] <имя параметра 1> [] <имя параметра 2> [] ... <имя параметра m> [].

Z1 Z2...Zm

[]

[{ee}] []<имя расширения 1>[]<имя расширения 2>[]...<имя расширения k>].

[{cc}] []<имя параметра 1>[]<имя параметра 2>[]...<имя параметра m>].

Z1 Z2...Zm

[{zz}]<имя окончания записи блока>.

Блоки описываются последовательно, один за другим. Каждый блок начинается со строки, содержащей имя блока (имя является списком атрибутов, которые в листе располагаются между символами "<" и ">". Атрибут имеет следующий общий вид записи: наименование атрибута(значение атрибута) (см. слайд 10)). Данное имя несет информацию о выбранном блоке, а также дополнительную информацию, которая необходима для корректного использования данного процесса в моделях. Признаком распознавания, что данная последовательность является именем блока служит наличие перед "(" символа "B".

В последующих строках указываются: список имен продуктов, список имен расширений, список имен параметров (констант реакций) и их значений. Концом блока является строка, содержащая стандартное имя <E(++++)>.

Признаком окончания строки является символ "." Все строки, которые содержат во второй и третьей позиции символы /* или */ опознаются как комментарии и игнорируются при построении моделей.

Для описания каждого типа реакции существуют определенные виды стандартных блоков. Ниже представлены имена некоторых из них:

<B(<=>)> - бимолекулярная обратимая реакция (пример приведен на слайде 7)

<B(1->1)> - мономолекулярная необратимая реакция (из 1-го реагента 1 продукт)

<B(1->N)> - мономолекулярная необратимая реакция (из 1-го реагента N продуктов)

<B(1=>1)> – конститутивный синтез

<B(1->0)> – утилизация

<B(XM0V)> – одна из обобщенных схем Михаэлиса для описания ферментативных реакций

В общем случае формальные блоки могут реализовывать любой закон переработки входящей информации и банк блоков может пополняться по мере возникновения необходимости.

В настоящее время банк содержит 51 блок, которые в совокупности позволяют проводить очень широкий спектр действий. Например, имеется блок формирования точек запоминания динамических характеристик системы, имеется блок запуска новой задачи Коши, имеются блоки арифметических действий и т.д.

Процедура построения математической модели основана на локальной независимости элементарных процессов. Это, с одной стороны позволяет описывать элементарные процессы независимо друг от друга и создавать их постоянно пополняемый банк, а с другой - объединение элементарных процессов в единую модель выражается в суммировании скоростей изменения концентраций по каждому веществу по всем элементарным процессам, в которых участвует данный компонент модели.

Слайд 10 и 11

Имена динамических переменных и параметров математических моделей состоят из списка атрибутов и их значений:

<Attribute_Name_1(Attribute_Value_1)...Attribute_Name_N(Attribute_Value_N)>

Существуют служебные зарезервированные атрибуты, например:

<B(...)J(...)> – идентификатор блока и режим использования блока;

<...Z(...)> – расширение;

<...O(...)> – ориентация блока или гена, обрабатывается специальной программой;
<...A(...)> – аллельность, обрабатывается специальной программой;
<...G(...)> – принадлежность к определенному участку генетической карты, обрабатывается специальной программой;
<...C(...)> – принадлежность к определенному компартменту;
<...S(...)> – принадлежность к определенному сценарию;

Слайды 12, 13, 14

Для унификации имен считается, что:

ss(...) – атрибут сущности

Примеры:

ss(protein) – белок;

ss(complex) – комплекс;

ss(enzyme) – фермент;

ss(compound) – низкомолекулярное вещество;

I(...) – атрибут наименования

Пример:

<ss(enzyme)I(Citrate synthase)> – фермент цитрат синтаза;

Имена регуляторов транскрипции представляются следующим образом:

<ss(complex)I(имя регулятора)>

Пример:

<ss(complex)I(IHF)> - регулятор Integration Host Factor

Имена белков, синтезируемых с транскриптов представляются следующим образом:

<ss(protein)I(имя гена с большой буквы)>

Пример:

<ss(protein)I(AcnB)> – белок, синтезированный с гена acnB

Имена ферментов, синтезируемых с транскриптов и ферментативных комплексов представляются следующим образом:

<ss(enzyme)I(название фермента из исходного файла)>

Пример:

<ss(enzyme)I(Citrate synthase)> – фермент цитрат синтаза

Имена метаболитов представляются следующим образом:

<ss(compound)I(название вещества из исходного файла)>

Пример:

<ss(compound)I(PYR)> – пируват

Слайд 15

Существует возможность задавать синонимы имен и значений атрибутов, например:

<...A(...)...> = <...Allele(...)...>;

<...C(...)...> = <...Compartment(...)...>

<...ss(compound)...> = <...ss(substrate)...>;

<...I(HMG-CoA reductase)...> = <...I(3-hydroxy-3-methylglutaryl-CoA reductase)...>

Слайд 16

В модели продукты считаются различными, если между их именами нет детального посимвольного совпадения.

В модели имена продуктов имеют глобальный смысл.

В модели параметры также считаются различными, если между их именами нет детального посимвольного совпадения, но в отличие от продуктов, имена и значения параметров каждого блока хранятся отдельно от имен и значений параметров других блоков. Т.е.

параметры имеют локальный смысл. В двух различных блоках модели параметры могут иметь одинаковые имена и различные числовые значения.

Слайд 17

Сборка математической модели осуществляется на основе пакета сценариев, для каждого из которых описывается его компартментная организация и дается служебный список карт, которые содержат перечень имен расширений и которым сопоставлен список моделей. Формальный блок, описывающий конкретный процесс, включается в общую модель, если в соответствующей карте встречается имя расширения, содержащееся в данном блоке в определенном месте.

Слайд 18

Моделирование многокомпарментных структур

Проблема 1

Многие процессы, протекающие в компартментах являются идентичными процессами, а при рассмотрении сложных организмов, идентичными могут являться также и целые компартменты. Это приводит к проблеме описания идентичных процессов протекающих в разных компартментах.

Проблема 2

Между компартментами организма происходит обмен веществами – т.е. существует активный и пассивный транспорт. При этом, типичны ситуации транспорта одного и того же вещества из одного компартмента, сразу в несколько других, не обязательно идентичных, компартментов. Это приводит к проблеме описания процессов транспорта.

Слайды 19, 20, 21

На данных слайдах проиллюстрирована методология моделирования в рамках ОХКММ многокомпарментных структур. Математические модели строятся автоматически на основе компартментной организации каждого из сценариев. Более подробное описание и работа с такими моделями будет проводиться на практических занятиях по данному курсу.

Слайд 22

1. Все рассмотренные модели построены на основе единого банка элементарных процессов, которые изначально не привязаны ни к какому конкретному компартменту.
2. Для построения конкретной модели необходимо задать компартментную структуру и указать принадлежность процессов и веществ к компартментам.
3. Построение модели состоит в многократном экстрагировании элементарных процессов из банка с учетом информации из п.2.
4. Привязка каждого вещества к компартменту состоит во включении в его имя наименования компартмента. В результате каждое имя конкретного вещества становится составным и однозначно характеризует его в контексте конкретной модели

Слайд 23

Моделирование ДНК- и РНК-специфичных процессов

Основная проблема:

Описание процессов, протекающих на отдельных участках матрицы.

Для ДНК-специфичных процессов:

- описание процессов связывания веществ с выделенными участками ДНК;
- описание ориентированных матричных процессов (репликация, транскрипция);
- описание взаимодействия матричных процессов с различными состояниями участков ДНК;
- описание взаимодействия ориентированных матричных процессов между собой;
- описание матричных процессов с учетом замкнутости/линейности геномов;
- описание рекомбинации, кроссинговера, делеций и вставок.

Для РНК-специфичных процессов:

- описание трансляции как матричного процесса;
- описание сплайсинга.

Слайды 24, 25, 26, 27

На данных слайдах проиллюстрирована методология моделирования в рамках ОХКММ ДНК- и РНК-специфичных процессов. Математические модели строятся автоматически на основе служебных списков имен – «генетических карт». Более подробное описание и работа с такими моделями будет проводиться на практических занятиях по данному курсу.

Слайд 28

1. Все рассмотренные модели генетических процессов построены на основе единого банка элементарных матричных процессов.
2. Для построения конкретной модели генетической регуляции и матричного синтеза необходимо задать генетическую структуру (генетическую карту) и указать принадлежность процессов и веществ к генам.
3. Построение модели состоит в многократном экстрагировании процессов из банка с учетом информации из п.2.
4. Привязка процессов к генам заключается в добавлении к наименованию продукта идентификатора местоположения гена в составе генома. В результате имя каждого продукта становится составным и однозначно характеризует его в контексте конкретной модели.

Слайд 29

Моделирование полиплоидных генных сетей

Основная проблема

Описание реакций и процессов, которые осуществляются комплексами, состоящими из белков, кодируемых разными полиаллельными вариантами одного и того же гена, порождает проблему комбинаторного перебора всех полиаллельных вариантов веществ. Комбинаторных вариантов тем больше, чем больше глубина реакций.

Слайды 30, 31

На данных слайдах проиллюстрирована методология моделирования в рамках ОХКММ полиплоидных генных сетей. Математические модели строятся автоматически на основе специальной технологией, реализованной в рамках данного подхода. Более подробное описание и работа с такими моделями будет проводиться на практических занятиях по данному курсу.

Слайд 32

1. Все рассмотренные особенности строения живых систем модели построены на основе единого банка элементарных процессов, в которых полиплоидность в явном виде не описывается.
2. Для построения модели необходимо задать какие белки, кодируются полиплоидными вариантами одного и того же гена.
3. Построение модели состоит в экстрагировании элементарных процессов из банка и комбинаторном построении полиплоидной модели на основе информации из п.2.
4. Полиплоидный состав вещества фиксируется путем добавления к его наименованию идентификаторов аллелей. В результате каждое имя, характеризующее вещество, становится составным и однозначно характеризует его в контексте конкретной полиплоидной модели.

Слайд 33, 34, 35, 36, 37

Рассмотренные примеры показывают, что специфика строения биологических систем действительно порождает ряд проблем, которые накладывают определенные требования к технологии моделирования компартментных структур организма, генетических процессов и

полиаллельности. Одновременно примеры подсказывают и направление развития адекватных средств моделирования. Во-первых, во всех рассмотренных примерах модели строились методом блочного конструирования имен веществ. Иными словами, чтобы зафиксировать местоположение вещества в конкретном компартменте, на конкретном участке ДНК, или, его аллельный состав к собственному имени вещества добавляются соответствующие идентификаторы. Немаловажным при этом оказывается возможность разделить процедуры описания процессов и процедуры привязки этих процессов. Элементарные процессы в примерах приводились вне связи с конкретной структурой живых систем. Сама привязка процессов и веществ к конкретным структурам моделируемой системы выносилась на более поздней стадии, на которой задавалась информация о составе компартментов моделируемой системы, о составе процессов в каждом компартменте, о генетическом составе каждого компартмента, если таковой имеется, а также давалась информация об аллельном составе генов.

Следовательно построение модели живой системы может быть организовано в виде алгоритма последовательного разворачивания элементарных процессов, в ходе выполнения которого происходит конструирование имен веществ и параметров. Все этапы конструирования модели: разбор аллельности, генетическая привязка, привязка к компартментам, могут быть выполнены последовательно друг за другом в определенном порядке. В результате выполнения такого алгоритма полученная модель автоматически учтет все перечисленные особенности строения биологической системы.

К настоящему времени разработано программное обеспечение, позволяющее строить модели многокомпарментных биологических систем, вплоть до организменного уровня.

Развитое к настоящему времени программное обеспечение позволяет:

1. Моделировать генетические процессы с учетом взаимного расположения генов и генетических элементов относительно друг друга и их ориентации в составе генома;
2. Рассматривать линейные, кольцевые и фрагментированные геномы;
3. Строить модели с учетом полиаллельности, отдельных генов, которая может быть любого порядка от 1 и выше;
4. Имеет удобные средства привязки процессов к отдельным компартментам и описания процессов транспортировки веществ из одного компартмента в другие;
5. Процессы могут описываться как в терминах биохимических реакций, так и в более общей форме (например, законы взаимопревращения веществ могут описываться уравнениями практически любой природы: полиномами, рациональными функциями, формулами, получаемыми в результате упрощения сложных химических процессов, могут применяться алгебраические уравнения, булевы и пороговые функции, вероятностной и детерминированной природы). В принципе арсенал законов, которые могут быть применены для описания элементарных процессов и действий не ограничен, так как блочный принцип, положенный в основу метода моделирования, позволяет легко наращивать арсенал средств практически без каких либо ограничений.
6. Моделирование может проводиться в терминах функциональных активностей, что значительно расширяет доступный арсенал моделирования.
7. Модели могут снабжаться произвольным количеством средств динамического мониторинга любых интегральных характеристик функционирования системы, что важно при определении фенотипических и морфологических характеристик моделируемого объекта, так как они, как правило, носят комплексный характер.
8. Программное обеспечение позволяет строить математические модели, воспроизводящие условия экспериментов, что важно при сопоставлении данных полученных в натурных экспериментах с данными численных расчетов.
9. Моделирование является распределенным. Т.е. модели процессов могут храниться в разных файлах в банке моделей элементарных процессов. Банк может постоянно пополняться и любой процесс, размещенный в банке, доступен для использования при

построении моделей. Также могут накапливаться банки генетических карт и банки карт сопряжения компартментов.

10. Все модели, разрабатываемые средствами языка SiBML, являются совместимыми между собой. Поэтому процесс построения математических моделей очень сложных систем может проводиться блочно, с последующим объединением блоков в единую математическую модель.
11. Разработанный стандарт языка предусматривает наличие синонимов для наименований веществ, аллелей, генов, компартментов. Это также расширяет возможности эффективного моделирования, в том числе позволяет разрабатывать эффективные средства автоматизированного построения математических моделей на основании информации, содержащихся в базах, в которых накапливается информация о структурно-функциональной организации генетических программ, контролирующих различные аспекты функционирования организмов, баз данных, в которых накапливаются данные о генетической регуляции экспрессии генов, баз данных, в которых накапливаются данные о метаболических реакциях.