

Спецкурс "Информационная Биология"

Лекция 11-ая «Механизмы регуляции транскрипции: описание в компьютерных базах данных»

Лектор - **к.б.н., с.н.с. Е.В.Игнатьева.**

Тема лекции: «Механизмы регуляции транскрипции: описание в компьютерных базах данных»

Слайд 2

Ген - протяженный участок ДНК, включает участки, соответствующие экзонам и интронам, старт транскрипции, сайт терминации транскрипции и регуляторные районы. **Экспрессия генов** - последовательность реакций на пути от гена к белку. Включает:

- транскрипцию
- сплайсинг
- транспорт через ядерные поры
- трансляцию

Транскрипция - синтез РНК. Образуется незрелая РНК, содержащая участки, соответствующие экзонам и интронам **Сплайсинг** процесс вырезания участков, соответствующих интронам, с образованием зрелой матричной РНК Осуществляется в ядре.

Трансляция - синтез белка в цитоплазме клетки при участии рибосом

Слайд 3

Регуляция экспрессии генов может осуществляться на всех этапах. Транскрипция - первый этап, на котором осуществляется выбор, будет ген экспрессироваться или нет.

Рассмотрим три вопроса

- как осуществляется транскрипция
 - чем определяется старт транскрипции
 - какие механизмы в клетке регулируют интенсивность транскрипции

-

Слайд 4

ДНК и РНК - полимеры, состоящие из нуклеотидов. Нуклеотид включает фосфат, основание и сахар. В молекуле ДНК нуклеотиды включают сахар дезоксирибозу, а в молекуле РНК нуклеотид включает сахар рибозу.

Молекула ДНК состоит из двух комплиментарных и антипараллельных друг другу полинуклеотидных нитей. РНК - включает одну полинуклеотидную цепь

Нуклеотиды в ДНК и РНК соединены фосфодиэфирными связями, которые соединяют 3' углеродный атом одного нуклеотида с 5' фосфатной группой следующего нуклеотида. Таким образом каждая цепь в ДНК или РНК имеет полярность, которая обозначается как 5' и 3' конец

Слайд 5

Транскрипция - синтез молекулы РНК, комплиментарной и антипараллельной одной из цепи ДНК (матричной цепи)

Соответственно, эта цепь ДНК по которой строится комплиментарная цепь РНК называется **матричной цепью**. Другая цепь ДНК называется **кодирующей цепью**, поскольку ее последовательность идентична последовательности РНК. При этом необходимо помнить, что вместо основания Т (тимин) в РНК включается основание U (урацил)

Пример:

ДНК

Нематричная (кодирующая) цепь ТАСGGАТА

Матричная цепь АТGCСТАТ

РНК, которая синтезируется на основе этого участка: UACGGАUА

Видно, что РНК комплементарна матричной цепи ДНК и идентична кодирующей цепи ДНК

Слайд 6

Транскрипция осуществляется при участии фермента - **РНК полимеразы**.

РНК-полимераза в ходе транскрипции выполняет следующие функции :

- расплетание и заплетание ДНК

- синтез РНК

- движение вдоль цепи ДНК

На слайде РНК-полимераза изображена схематически в виде цилиндра.

РНК-полимераза содержит специальные каналы внутри соответствующие по толщине нитям ДНК и РНК

Шириной 12-15 ангстрем глубиной 20 А

шириной 25 ангстрем глубиной 5-10 А

Скорость синтеза 40 нуклеотидов в секунду при 37 градусах

Слайд 7 Этапы транскрипции на примере прокариотического организма (клетка не имеет ядра). Представитель - *E.coli* кишечная палочка.

Первый этап транскрипции - распознавание матрицы: при участии сигма фактора (у *E.coli*) РНК полимеразы связывается с ДНК и расплетает ДНК в точке инициации транскрипции

Второй этап - инициация:

сигма фактор отсоединился и синтезирована короткая цепь РНК (2-9 пар оснований)

Третий этап - элонгация:

движение РНК полимеразы вдоль ДНК, расплетание ДНК, синтез РНК, заплетание ДНК

Третий этап - терминация:

окончание транскрипции, распад комплекса ДНК-РНК-полимераза. происходит после распознавания терминатора

Слайд 8

Промотор - участок, прилегающий к старту транскрипции. Содержит регуляторные элементы, которые опознаются с белками, обеспечивающими инициацию транскрипции У *E.coli* роль таких белков выполняет сигма фактор.

ГЕНОМ *E.coli* - 4*10⁶ пар оснований

ПРОМОТОР ~60 пар оснований

Как осуществляется распознавание точки инициации транскрипции

-Анализ последовательностей более 100 различных промоторов *E.coli* выявил сходное строение:

1) в районе старта транскрипции - пурины, часто "САТ"

2) в позиции "-10" гексамер ТАТААТ (последовательность "-10")

3) в позиции "-35" гексамер ТТGACA (последовательность "-35")

4) расстояние между последовательностями "-35" и "-10" ~16-18 нуклеотидов

Функцию участков промотора изучали с помощью мутаций

1) нарушение последовательности - 35 замедляет посадку полимеразы на ДНК

2) нарушение последовательности -10 замедляет образование открытого комплекса.

Слайд 9

Слайд 10

полимераза четыре субъединицы: 2 альфа, бета и бета'. Они выполняют разные функции.

Слайд 11

Располагая этими данными, и анализируя промотор с неизвестной функцией, можно предсказать, какой сигма фактор будет с ним взаимодействовать. Следовательно, можно предвидеть ситуацию, в которой будет экспрессироваться ген.

Слайд 12

Процесс транскрипции у эукариот. Его особенности:

1) У эукариот - три типа РНК полимераз и три группы генов:

- pol I транскрибируемые -рибосомные РНК
- pol II транскрибируемые - матричные РНК
- pol III транскрибируемые - транспортные РНК

2) Каждая полимераза имеет много субъединиц, например, pol II включает более 10 субъединиц

3) Для точной посадки полимеразы на участок, прилежащий к старту транскрипции нужны вспомогательные белки - базальные транскрипционные факторы

4) промоторы генов, транскрибируемых каждым типом полимераз имеют характерное строение

В дальнейшем речь пойдет о транскрипции с участием РНК полимеразы II. С ее участием экспрессируется наибольшее количество генов.

Наблюдается огромное разнообразие в строении регуляторных районов генов этой группы. Регуляторные районы включают :

1) **коровый промотор** - короткий участок ДНК (~60 пар оснований), расположенный непосредственно перед стартом транскрипции и, как правило, содержащий ТАТА-бокс, а также Inr-элемент

2) **промоторный район** - 5' регуляторный район, непосредственно примыкающий к старту транскрипции и включающий коровый промотор

2) отдаленные регуляторные районы **энхансеры, сайленсеры.**

Могут быть удалены на большое расстояние , до нескольких тысяч пар оснований и располагаться в 5', 3' участках генов, экзонах и интронах.

Функция корового промотора формирование на нем **преинициационного комплекса** (= базального транскрипционного комплекса), включающего базальные транскрипционные факторы и полимеразу.

Функция отдаленных регуляторных районов - взаимодействие с белками, транскрипционными факторами.

Слайд 13

Существует как минимум две модели формирования преинициационного комплекса

Первая модель - пошаговой сборки. На первом этапе базальный транскрипционный фактор TFIID связывается с коровым промотором.

Вслед за этим осуществляется последовательная посадка других базальных транскрипционных факторов (TFIIB, TFIIF, TFIIE, TFIIH)

Слайд 14

Вторая модель - инициация транскрипции при участии холоэнзима, включающего РНК полимеразу II. Помимо РНК полимеразы II холоэнзим содержит базальные транскрипционные факторы. Другие (не базальные) транскрипционные факторы, которые взаимодействуют с удаленными регуляторными районами гена, стабилизируют преинициационный комплекс. Это оказывается возможным благодаря изгибанию ДНК.

Слайд 15

В состав базального транскрипционного фактора TFIID входит TBP (ТАТА связывающий белок). При взаимодействии с белком TBP ДНК изгибается

Слайд 16

Известно множество различных транскрипционных факторов. У человека их около 2000. Транскрипционные факторы можно расклассифицировать по функциям.

I Присутствуют в клетке постоянно

II Экспрессируются при определенных условиях

1) Экспрессируются в развитии в определенных клетках (GATA, HNF, Pit1, MyoD)

2) Экспрессия зависит от сигнала

а) Стероидные рецепторы GR, ER, PR, TR, RARs, RXR, PPAR

б) Активируются внутренними сигналами (SREBP, p53 ...)

в) Активируются сигналами, передаваемыми от мембраны клетки:

(SRF, CREB, FOS-JUN, STAT, NF-kB, NFAT)

Слайд 17

Регуляторные районы генов могут иметь сложное строение и включать сайты связывания различных транскрипционных факторов.

В качестве примера приведены регуляторные районы, контролирующие транскрипцию гена аполипопротеина В человека.

Слайд 18

Еще один пример сложного строения регуляторных районов - ген ароматазы человека.

Слайд 19

Сложное строение регуляторных районов генов и многообразие сочетаний сайтов различных типов обеспечивает большое разнообразие вариантов регуляции экспрессии генов. Еще одним дополнительным механизмом регуляции экспрессии генов является регуляция активности транскрипционных факторов. На слайде представлены семь возможных механизмов регуляции активности транскрипционных факторов.

Слайд 20

ПРИМЕР:

холестерин регулирует активность транскрипционного фактора SREBP.

Этот фактор присутствует в клетке в форме неактивного предшественника preSREBP.

PreSREBP связан с мембранами эндоплазматического ретикулума. Стерол-регулируемая протеаза (SRP) расщепляет PreSREBP, в участке "1". Затем другая протеаза расщепляет белок в участке "2" и активный SREBP высвобождается и проникает в ядро. Происходит димеризация фактора, а затем активация транскрипции целого ряда генов, содержащих в регуляторных районах сайты связывания фактора SREBP.

Высокий уровень холестерина ингибирует активность стерол-регулируемой протеазы. Таким образом реализуется механизм отрицательной обратной связи, обеспечивающий постоянный уровень холестерина в клетке.

Слайд 21

У эукариот ДНК хранится в ядре в комплексе с белками, образуя компактную структуру - хроматин. Хроматин имеет несколько уровней организации.

1) Нуклеосомы - комплекс ДНК с **гистоновыми** белками

2) 30-нм фибриллы

3) петли длиной 300 нм

4) фибриллы толщиной 250 нм

5) хроматиды в хромосоме

Слайд 22

Нуклеосомная укладка влияет на интенсивность экспрессии. Как ??? Гистоновые белки, которые обеспечивают компактную укладку ДНК, препятствуют (затрудняют) взаимодействие транскрипционных факторов и белков базального транскрипционного комплекса с ДНК

Слайд 23

Транскрипционные факторы делятся на две группы по способности связываться с ДНК-сайтами в составе нуклеосомы:

- обладающие такой способностью
- не способных к этому

Слайд 24

ПРИМЕР:

промотор гена MMTV (Mouse mammary tumor virus)

Экспрессия этого гена регулируется транскрипционными факторами, способными взаимодействовать ДНК-сайтами в составе нуклеосомы. И более того, наибольшая активация транскрипции возможна только в случае, если ДНК изогнута на поверхности нуклеосомы. Если нуклеосомная укладка отсутствует, два транскрипционных фактора GR и NF1 конкурируют друг с другом за связывание с близко расположенными сайтами их связывания на ДНК.

Если нуклеосомная укладка имеет место, GR и NF1 одновременно связываются с ДНК, что способствует наибольшей активации транскрипции

Слайд 25

Следующий механизм регуляции транскрипции - ацетилирование гистонов. Гистоновые белки имеют положительно заряженные концевые фрагменты, которые могут взаимодействовать с отрицательно заряженными фосфатами сахарофосфатного остова ДНК. Это взаимодействие стабилизирует комплекс ДНК с белками. Ацетилирование (присоединение ацетильной группы CO₂CO) концевых фрагментов гистоновых белков нейтрализует их положительный заряд. Как следствие, электростатическое взаимодействие белков с ДНК ослабляется. Появляется возможность взаимодействия ДНК с некоторыми транскрипционным факторами.

Слайд 26

Транскрипционные факторы участвуют в регуляции уровня ацетилирования гистонов. На слайде представлен такой механизм на примере гетеродимерного транскрипционного фактора RAR/RXR. Механизмом активации этого транскрипционного фактора является связывание с лигандом (см. слайд 19)

В отсутствие лиганда гетеродимерный фактор RAR/RXR связывается с ДНК а также белками корепрессорами (mSin3A, SMRT, HDAC-1). Корепрессоры осуществляют деацетилирование гистоновых белков, входящих в состав нуклеосом. Прочное связывание ДНК с нуклеосомными белками препятствует формированию базального транскрипционного комплекса в районе старта транскрипции. Совсем другая картина наблюдается в случае, если рецепторы ретиноевой кислоты связаны с лигандом. При этом транскрипционные-факторы связываются с другими белками - коактиваторами (CBP/p300, p/CAF, SRC/ACTR).

Коактиваторы осуществляют ацетилирование гистонов. При этом связь ДНК с нуклеосомными белками ослабляется, что облегчает формирование базального транскрипционного комплекса в районе корового промотора. В результате наблюдается повышение уровня транскрипции гена.

Слайд 27

Следующий механизм регуляции транскрипции - метилирование ДНК. В эукариотических клетках уровень метилирования ДНК часто коррелирует с уровнем экспрессии гена. Метилированные участки ДНК транскрибируются менее активно, чем неметилированные участки. Наиболее часто метилированию подвергаются цитозиновые нуклеотиды, стоящие после гуаниновых

Слайд 28

Пример:

Ген проэнкефалина человека содержит сайт связывания транскрипционного фактора AP-2 в районе -80/-70 от старта транскрипции. В случае, если сайт находится в деметилированном состоянии, AP-2 эффективно взаимодействует с этим сайтом, что приводит к активации транскрипции гена. Метилирование нарушает связывание AP-2 с ДНК и транскрипция снижается.

Слайд 29

Следующий механизм регуляции транскрипции - наличие локус-контролирующих районов. **Локус-контролирующий район (LCR)** - удаленный регуляторный район, обеспечивающий включение и выключение генов за счет попеременного взаимодействия с регуляторными районами генов, находящихся в кластере

Слайд 30

Пример: локус-контролирующий район кластера гемоглобиновых генов человека.

Размер бета-глобинового кластера человека составляет 70 килобаз. Кластер содержит гены гемоглобинов ϵ , γ , δ , β и псевдоген $\psi\beta^1$. Экспрессия гемоглобинов осуществляется на ранних этапах онтогенеза в пренатальный период в желточном мешке, а затем в эмбриональной печени, в постнатальный период в костном мозге. При этом LCR последовательно активирует гены, входящие в состав кластера. В желточном мешке экспрессируется ген ϵ , в эмбриональной печени экспрессируются гемоглобины γ и δ , а в постнатальный период в костном мозге идет экспрессия гемоглобинов δ и β . Последовательное включение новых генов происходит при одновременной инактивации ϵ - и γ генов на соответствующих стадиях развития.

Слайд 31

Следующий механизм регуляции транскрипции - наличие инсуляторов. **Инсулятор** - участок ДНК, который, будучи помещенным между двумя регуляторными элементами может препятствовать активирующему либо подавляющему действию одного элемента на другой.

Слайд 32

Роль инсулятора может выполнять участок прикрепления к ядерному матриксу (MAR = matrix attachment regions). При включении такого инсулятора между энхансером и промотором (см. слайд) два регуляторных района оказываются в различных доменах, и не способны взаимодействовать.

Слайд 33

ЧАСТЬ 2: БАЗЫ ДАННЫХ ПО РЕГУЛЯЦИИ ТРАНСКРИПЦИИ

Слайд 34

Первая база, о которой будет рассказано:

[TRANSCRIPTION REGULATORY REGIONS DATABASE \(TRRD\)](http://www.mgs.bionet.nsc.ru/mgs/gnw/trrd/)

Эта база разрабатывается в Институте цитологии и генетики Сибирского отделения Российской академии наук, (Новосибирск-90) с 1993 года

<http://www.mgs.bionet.nsc.ru/mgs/gnw/trrd/>

Слайд 35

Публикация по TRRD в Nucleic Acids Research за 2002 год.

Слайд 36

[Главная страница для входа в TRRD через Интернет](http://www.bionet.nsc.ru/trrd/)

<http://www.bionet.nsc.ru/trrd/>

Слайд 37

TRRD объединяет семейство баз данных. Входы в них представлены на слайде.

Слайд 38

TRRD включает данные о:

1) структурно-функциональной организации транскрипционных регуляторных районов. структурные и функциональные характеристики:

а) сайтов связывания транскрипционных факторов
б) регуляторных районов (промоторов, энхансеров, сайленсеров)

в) локус-контролирующих районов

2) данные о паттернах экспрессии генов

3) данные о транскрипционных факторах

4) общую информацию о генах вместе с перечислением регуляторных элементов всех уровней

Слайд 39

Пример записи из базы TRRDGENES - общее описание гена amyloid beta protein precursor gene человека

Слайд 40

Пример записи из базы TRRDGENES - общее описание гена apolipoprotein A1 человека (начало)

Слайд 41

Общее описание гена apolipoprotein A1 человека в TRRDGENES (продолжение) Приведена информация о двух регуляторных единицах и входящих в них сайтах связывания транскрипционных факторов.

Слайд 42

Общее описание гена apolipoprotein A1 человека в TRRDGENES (продолжение) Приведена информация о композиционном элементе, сайтах связывания транскрипционных факторов и регуляторной единице (промоторе)

Слайд 43

Общее описание гена apolipoprotein A1 человека в TRRDGENES (продолжение)

Слайд 44

Описание транскрипционных регуляторных единиц (промоторов, энхансеров, сайленсеров) в базе TRRDUNITS. Приведены последовательность и позиции промотора гена D2 крысы.

Слайд 45

Еще пример входа в базе TRRDUNITS: данные об энхансере гена apolipoprotein A1 человека

Слайд 46

Генерация блока полей с нуклеотидными последовательностями регуляторной единицы в базе TRRDUNITS осуществляется автоматически с помощью программы TRRD-Pars. Генерация проводится на основе сопоставления данных из TRRDGENES, TRRDSITES и EMBL/GenBank и включает в себя несколько этапов. На первом этапе рассматривается точка отсчета, использованная для описания регуляторной единицы в TRRDGENES (точкой отсчета может быть старт транскрипции, трансляции, начало последовательности). На основе информации о позициях сайтов в последовательности EMBL/GenBank, а также о расстоянии сайтов от точки отсчета, определяется позиция точки отсчета в последовательности EMBL/GenBank. На втором этапе, на основании найденной позиции точки отсчета, а также аннотированных данных о положении границ регуляторной единицы относительно точки отсчета, определяется локализация регуляторной единицы и проводится экстракция соответствующего участка нуклеотидной последовательности из EMBL/GenBank.

Слайд 47

Описание сайта связывания транскрипционного фактора в TRRDSITES. Приведено описание сайта связывания HNF-4 из регуляторных районов гена ApoB человека. Выделены цифровые коды экспериментальных методов исследования сайтов связывания транскрипционных факторов. Стрелки указывают на типичные иллюстрации результатов в экспериментальных статьях.

Слайд 48

Еще один пример описания сайта связывания транскрипционного фактора в TRRDSITES (представление в системе SRS). Это сайт связывания T3Rbeta/RXRalpha в регуляторном районе гена ApoA1 человека.

Слайд 49

Обратите внимание на три поля в таблице TRRDSITES. Они внесены в формат недавно. В поле ImportantPos представлены важные для взаимодействия с транскрипционным фактором нуклеотиды в пределах последовательности сайта. Эти нуклеотиды выявлены на основании различных экспериментальных процедур, английские названия которых приведены на слайде. Два других поля SeqContradiction и PosContradiction содержат данные о последовательности сайта или его позиции из экспериментальной статьи, если эти данные не совпадают с таковыми из баз EMBL либо GenBank. В этом случае аннотатор вносит данные из EMBL либо GenBank в основные поля - Sequence и SequencePosition.

Слайд 50

Описание гетеродимерного транскрипционного фактора PPARgamma/RXRalpha в TRRDFACTORS

Слайд 51

Еще один пример описания гетеродимерного транскрипционного фактора T3Ralpha/RXRalpha в TRRDFACTORS (представление в системе SRS).

Слайд 52

Описание локус-контролирующего района в TRRDLCR.

Слайд 53

Формат базы позволяет проводить иерархическое описание локус контролирующего района, включая а) описание кластера генов, б) общую информацию о локус контролирующем районе, в) информацию о функциональных элементах, г) данные о сайтах гиперчувствительности, которые являются маркерами района.

Слайд 54

Примеры строения кластеров генов, контролируемых локус-контролирующими районами. Кластеры генов могут включать как несколько генов (глобиновые кластеры), так и один ген (CD2 человека)

Слайд 55

Структура локус-контролирующего района кластера глобиновых генов человека. В нижней части представлены шкала с позициями и сайты связывания транскрипционных факторов.

Слайд 56

Особенности экспрессии генов описываются в таблице TRRDEXP. Каждый ген может иметь несколько блоков, описывающих экспрессию. В представленной записи содержится информация о том, что ген AAP человека экспрессируется в лимфобластоидных клетках и астроцитах. В лимфобластоидных клетках экспрессия усиливается под действием теплового шока (heat shock). Справа представлены типичные фотографии, представляющие результаты экспериментов.

Слайд 57

Данные об экспериментальной статье в таблице TRRDBIB

Слайд 58

Визуализация информации в виде карт регуляторных районов генов осуществляется с помощью программы TRRD-Viewer. Доступ к вьюеру обеспечивается через запись о гене в таблице TRRDGENES. Представлен пример карты регуляторных районов гена FAS крысы. Изображения регуляторных районов и сайтов связывания транскрипционных факторов являются одновременно гиперссылками на их текстовое описание в соответствующих частях TRRD.

Слайд 59

TRRD - один из информационных модулей системы GENEEXPRESS-2

Слайд 60

Гиперссылки на базы данных и программные модули системы GeneExpress из базы TRRDSITES

Слайд 61

Гиперссылки на другие части TRRD и внешние базы данных из TRRDGENES.

Слайд 62

TRRD регулярно пополняется новыми данными.

Слайд 63

Ввод данных в TRRD сопровождается их синтаксическим и семантическим анализом с использованием оригинальных программ. Первая программа TRRD-INPUT реализована на Visual FoxPro 5.0 using OLE technology and ActiveX elements (Ananko E.A. et al., 1998). Одна из ее функций - проверка вводимых в TRRD терминов на их соответствие с контролируемыми словарями.

Слайд 64

Полная схема процессинга данных. Помимо синтаксического анализа, который осуществляет программа TRRD-INPUT, проводится семантический анализ с помощью программы TRRD-Parc. Это подразумевает динамическую верификацию информации о нуклеотидных последовательностях и позициях сайтов связывания транскрипционных факторов из

TRRDSITES используя последовательности, из EMBL/GenBank, ссылки на которые записаны в TRRD. Программа TRRD-Parc проводит также автоматическую генерацию блока полей с нуклеотидными последовательностями регуляторной единицы в базе TRRDUNITS (см. слайд 46).

На основе контролируемых словарей генерируются тезаурусы органов и тканей млекопитающих. База данных установлена в Интернет и запросы к ней возможны на основе тезаурусов и системы SRS (см. ниже)

Слайд 65

Представлена иерархическая организация контролируемых словарей морфологических терминов в TRRD на примере органа "мозг" и ткани "нервная ткань"

Слайд 66.

Поиск генов, экспрессируемых в определенном органе, например, в почке возможен на основе тезаурусов органов и тканей млекопитающих. При этом благодаря иерархической структуре тезаурусов в найденных входах будут обнаружены как записи "почка", так и записи, обозначающие ее части (кора, канальцы, гломерулы и т.д.)

Слайд 67

Основным средством поиска данных по ключевым словам в TRRD является система SRS (Sequence Retrieval System). Общее количество индексированных полей, по которым возможен поиск через систему SRS составляет 131

Слайд 68

Главное окно для поиска в TRRDGENES через поисковую систему SRS. На слайде представлен результат поиска по комбинации apolipoprotein (часть названия гена) и human (видовая принадлежность).

Слайд 69

В TRRD предусмотрена возможность поиска генов через браузеры по названиям генов и видов. Быстрый доступ к генам из функционально значимых групп, возможен через тематические разделы TRRD.

Слайд 70 (анимированный слайд - для продолжения нажмите на левую кнопку)

Представлены названия тематических секций TRRD.

Слайд 71

Поиск в базе последовательностей регуляторных районов, гомологичных анализируемой последовательности ДНК, возможен с помощью программы BLAST (http://www.mgs.bionet.nsc.ru/mgs/systems/fastprot/units_blast.html)

Слайд 72

На основе TRRD возможно формировать выборки эукариотических промоторов либо сайтов связывания транскрипционных факторов.

Слайд 73

Рабочая группа TRRD включает более двадцати сотрудников.

Слайд 74

Далее речь пойдет о других специализированных молекулярно-биологических базах данных.

SWISS-PROT - содержит данные о белках. Поскольку транскрипционные факторы являются

белками, эта база данных полезна при анализе механизмов регуляции транскрипции. Описание белка в SWISS-PROT включают ссылку на ген, аминокислотную последовательность, разбивку полипептидной цепи на структурные или функциональные домены.

Слайд 75

База TRANSFAC - является специализированной базой по транскрипционным факторам, помимо характеристики белковых молекул транскрипционных факторов здесь приведены данные о последовательностях их сайтов связывания с указанием генов, в которых они локализованы. Т.о. данные этой базы частично перекрываются с данными TRRD.

Слайд 76

В базе TRANSFAC можно найти также классификацию транскрипционных факторов на основе сходства строения ДНК связывающего домена.

Слайд 77

TRANSFAC содержит также S/MARt DB - данные о районах прикрепления к ядерному матриксу (MAR), а также белках, которые с ними взаимодействуют.

Слайд 78

Показана выдача из этой части базы по команде "browse"

Слайд 79

Один из входов S/MARt DB.

Слайд 80

Далее будут коротко охарактеризованы базы, содержащие данные о паттернах экспрессии. Напомню, что первая база данных, о которой шла речь в этой лекции, содержит часть TRRDEXP, в которой аккумулированы сведения об особенностях экспрессии конкретного гена Эти данные занесены в TRRD экспертами биологами на основании анализа экспериментальных статей. Базы, о которых пойдет речь далее, составляются на основе автоматической регистрации данных экспериментов с использованием технологии "microarray"

Слайд 81

Главная страничка базы ArrayExpress

Слайд 82

Один из входов в базе ArrayExpress

Слайд 83

Главная страничка организации "Microarray Expression Data Society"

Слайд 84

Главная страничка базы HugeIndex Database

Слайд 85

В HugeIndex Database имеется поисковая система, позволяющая проводить поиск генов, по разному экспрессирующихся в двух различных органах.

Слайд 86

Результат запроса , показанного на предыдущем слайде - 3726 генов

