

Лекция «Исследование качественных и количественных характеристик транскриптома».

Вторая часть.

Слайд 2. Биочип-Технологии

Прежде чем займемся рассмотрением материала о ДНК-биочипах, необходимо отметить, что вообще биочиповые или матричные технологии становятся одним из самых продуктивных способов исследования живых объектов. Главный принцип – расположение в определенном порядке в виде матрицы проб (или зондов), состоящих из известных веществ, и зондирование с помощью такого матричного набора проб гетерогенной смеси веществ, в виде которой обычно представлен биологический материал, для определения качественного состава этой смеси и количественной оценки ее компонентов.

К настоящему времени кроме биочипов с ДНК (gene chips) уже разработаны биочипы, содержащие иммобилизованные белки, углеводы, липиды и т.д., и даже живые объекты – клетки и микроорганизмы.

Однако в рамках представляемых лекций мы будем говорить только о ДНК-биочипах, поэтому и название «биочип» будет относиться только к ДНК-содержащим биочипам.

Слайд 3. ДНК-биочипы (ДНК-микроматрицы)

Итак эта лекция посвящена относительно новому, но очень перспективному методу исследования функциональных проявлений генома на уровне транскриптома – методу ДНК-биочипов или, как иногда их называют в русскоязычной литературе, - ДНК микроматриц.

Необходимо подчеркнуть, что идейной основой количественных исследований транскриптома методом ДНК-биочипов является предположение, что измеренные в биочип-эксперименте интенсивности гибридизации для каждого гена отражают их относительный уровень экспрессии.

Слайд 4. Определение ДНК-биочипов (ДНК-микроматриц)

Итак, ДНК-биочипы – это миниатюризированные матрицы или подложки, на которых в определенном порядке распределены фрагменты ДНК, соответствующие отдельным генам или их частям. Такие организованные микроматрицы позволяют проводить эксперименты по одновременному анализу структуры и экспрессии тысяч генов с помощью параллельной гибридизации.

Высокий уровень методов преобразования результатов этих экспериментов в цифровые данные и методов компьютерной обработки последних обеспечивает возможность анализировать и сопоставлять экспрессию таких массивов генов во множестве экспериментальных условий.

Таким образом получается статическая информация об экспрессии генов (в какой ткани или типе клеток, на какой стадии, при каком воздействии и т.д.) и динамическая информация об экспрессии генов при сопоставлении данных отдельных экспериментов.

Слайд 5. ТИПЫ ДНК-БИОЧИПОВ

Рассмотрим, какие бывают типы ДНК-биочипов.

В английской литературе чаще применяют термин «аггау» для выражения мысли об упорядоченном в виде матрицы расположении проб. Однако в русском языке это слово неудобно и в произношении и в написании. И хотя в английской литературе у термина «chip» более узкое применение, о чем я скажу ниже, в русской литературе этот термин получил большее распространение для описания всех разновидностей «эррейев».

Итак, во-первых, различают ДНК-биочипы в зависимости от размеров – существуют микробиочипы, когда размеры элементов, содержащих пробы, находятся в микрометровом диапазоне, а сами биочипы помещаются на обычном микроскопическом стекле размером 25 мм x 76 мм. Также существуют и активно применяются «макро-биочипы» или «макро-эррейи» с размерами до 25см x 25см, однако такие биочипы, как правило, изготавливаются из гибких мембран.

Во-вторых, различают ДНК-биочипы в зависимости от материала – на твердой или гибкой поверхности. В качестве твердой поверхности используют поверхность стекла или пластика или стекла, покрытого полимерной пленкой. А в качестве гибкой поверхности используют, как правило, нейлоновые мембраны или фильтры. Строго говоря, слово «чип» (а если придется писать или говорить на английском – то обязательно) применяется только к «эррейям» на твердой поверхности.

В третьих, существуют также биочипы, в которых используется метод иммобилизации фрагментов ДНК в объеме небольших капель или блоков геля – этот метод был разработан российскими учеными во главе с академиком А.Д. Мирзабековым фактически как самый первый ДНК-биочип. Фильтровые биочипы с этой точки зрения также являются биочипами с иммобилизованными в объеме пробами, так как структура фильтров пористая.

В-четвертых, различают ДНК-биочипы в зависимости от природы используемых фрагментов ДНК: для производства олигонуклеотидных биочипов используют химически синтезированные одноцепочечные олигонуклеотиды длиной 20-75 н.о.; а для кДНК-биочипов – размноженные (в бактериальных клетках или с помощью ПЦР) двухцепочечные клоны из библиотек кДНК длиной 100-2500 н.о.. Наконец, существуют и, так называемые, геномные ДНК-биочипы, в которых иммобилизованы большие фрагменты, до десятков кб, для анализа физической структуры геномов.

Слайд 6. БИОЧИПЫ МИРЗАБЕКОВА - MAGIChip™ ТЕХНОЛОГИЯ

В самых первых ДНК-биочипах была реализована технология иммобилизация олигонуклеотидов в очень маленьких кусочках полиакриламидного геля. Затем было придумано раскапывать микрокапли раствора с компонентами для геля и с олигонуклеотидами и проводить фотоактивируемую кополимеризацию.

Слайд 7. СПОСОБЫ ИЗГОТОВЛЕНИЯ ДНК-БИОЧИПОВ НА ТВЕРДОЙ ПОВЕРХНОСТИ

Продолжим классификацию разновидностей ДНК-биочипов. Типизацию по масштабу, типу подложки и содержанию мы уже рассмотрели. Теперь рассмотрим типизацию по способу их приготовления. Олигонуклеотидные биочипы могут быть приготовлены или способом синтеза олигонуклеотидов *in situ*, или распечатаны способом, применяемым в принтерах, обычно струйных. кДНК-биочипы, как правило, изготавливаются методом печати и обычно контактной, хотя встречается и струйная, бесконтактная.

Рассмотрим на схемах способы приготовления ДНК-биочипов.

Первый метод – фотолитография, или фотоиндуцируемый синтез олигонуклеотидов *in situ*.

Второй - механическое раскапывание или контактная, матричная печать.

Третий – бесконтактная печать по методу ink-jet принтеров.

Видно, что в первом случае пробы распределены в виде квадратных пятен, а в двух последних – в виде круглых.

Такое разнообразие методов приготовления ДНК-биочипов показывает, что метод находится в стадии бурного развития и поиска оптимальных инженерных и методических решений.

Слайд 8. ИЗГОТОВЛЕНИЕ ДНК-БИОЧИПОВ – на стекле

Так выглядит стеклянный ДНК-биочип, в данном случае, судя по круглым пятнам, - приготовленный методом печати. Видно, какую маленькую область на предметном стекле занимают ряды проб. Видно также, что морфология пятен проб слегка различается – это один из источников статистического разброса в результатах. А после гибридизации сканированное изображение показывает ряды из реплицированных проб, т.к. таким образом этот разброс в какой-то мере преодолевается.

Слайд 9. ПРОЦЕСС ИЗГОТОВЛЕНИЯ ДНК-БИОЧИПОВ - РОБОТЫ

На этом слайде проиллюстрированы различные устройства для приготовления ДНК-биочипов – эррейеры, принтеры или споттеры.

Слайд 10. ПРОЦЕСС ИЗГОТОВЛЕНИЯ ДНК-БИОЧИПОВ - видео

В этом видео-ролике показана работа контактного принтера ДНК-биочипов.

Слайд 11. КОНТАКТНАЯ ПЕЧАТЬ (Pin Printing Technology)

Теперь рассмотрим подробнее способы приготовления ДНК-биочипов с помощью печати. Печать ДНК-биочипов заключается в доставке малых количеств ДНК-проб на твердую поверхность.

Обычно доставляемый объем составляет нанолитровый и пиколитровый диапазон, поэтому применяются особые методы, сходные с матричной или струйной высокоразрешающей печатью.

Как уже было сказано, есть два основных способа печати:

- бесконтактная
- контактная

Рассмотрим способ приготовления ДНК-биочипов с помощью контактной печати (pin printing technology). Суть его заключается в погружении тонких металлических стержней или пинов в образцы, размещенные в планшетах и перенесение малого количества образца на стекло на кончике пина. Размер пятна определяется диаметром и формой пина, а также поверхностными энергиями взаимодействий между веществами пина, жидкости и предобработанного стекла.

Ограничения, присущие для этого способа – это малый объем переносимого образца, низкая скорость и низкая воспроизводимость. Однако возможность производить адаптируемые для различных нужд ДНК-биочипы разной емкости в условиях отдельной лаборатории является настолько привлекательной, что этот способ изготовления ДНК-биочипов очень широко применим.

Слайд 12. РАЗНОВИДНОСТИ ПИНОВ

К тому же были разработаны инженерные решения, направленные на совершенствование контактной печати – это использование более сложно устроенных пинов. Слева – исходный ординарный пин, в центре – пин в виде пинцета, позволяющий заправляться большим объемом пробы, справа – расщепленный пин, также с большей емкостью.

Слайд 13. РАЗНОВИДНОСТИ ПИНОВ: TeleChem's

Разработаны также пины фирмы TeleChem's (Micro Spotting pins), которые работают наподобие чернильных штампов и могут переносить до 0.2-1.0 мкл пробы, отпечатывать пятна размером 75-360 μm в диаметре с расстоянием между ними 140 μm . Снизу показано, что получаемые ряды пятен довольно униформны.

Слайд 14. РАЗНОВИДНОСТИ ПИНОВ: Pin-and and Ring

Следующая разработка в направлении совершенствования контактной печати – так называемый Pin-and-Ring, пин с кольцом, применяемый в аппаратах фирмы Genemachine. Еще больше увеличивается емкость пина и, соответственно, количество контактных отпечатков с одной заправки, что приводит к увеличению производительности споттера, т.е. количеству стекол, изготовленных в более однообразных условиях.

Слайд 15. БЕСКОНТАКТНАЯ ПЕЧАТЬ (INK-JET)

Приготовление ДНК-биочипов с помощью бесконтактной печати (ink-jet). На слайде показано, как выглядят устройства для такого способа:

- пьезоэлектрическое устройство, в котором стеклянный капилляр с раствором пробы подвергается обжимающему воздействию пьезо-кристалла, объем которого регулируется электрическим сигналом, в результате из капилляра выдавливается капля пробы определенного объема;

- шприц-соленоидное (syringe-solenoid) устройство, в котором раствор пробы находится под давлением и процесс выброса капли пробы из капилляра регулируется высокоскоростным соленоидным клапаном.

На этом слайде показаны ряды ³²P-меченых проб, нанесенных одноканальным пьезоэлектрическим диспенсером. Расстояние между пятнами 750 μm , размер пятен 145 μm . Объем капель – сотни пиколитров (т.е. 10-4 мкл). Высокая скорость – несколько тысяч капель в секунду. Ряды выглядят довольно равномерными.

Также показаны ряды проб, нанесенных шприц-соленоидным (syringe-solenoid) устройством. Расстояние между пятнами 500 μm , размер пятен 325 μm . Объем капель – 4-8 нанолитров (т.е. 10-3 мкл). Высокая скорость – несколько тысяч капель в секунду. Ряды выглядят довольно равномерными, хотя и заметны различия некоторых пятен по диаметру.

Слайд 16. ЗАВИСИМОСТЬ КАЧЕСТВА ПЕЧАТИ ОТ СПОСОБА ОБРАБОТКИ СТЕКЛА И РАСТВОРИТЕЛЯ ДЛЯ ПРОБ

Мы рассмотрели только что зависимость качества и воспроизводимости печати от формы пина или типа устройства.

Теперь рассмотрим зависимость от поверхностных энергий взаимодействий между веществами пина, жидкости и предобработанного стекла, в частности от состава, которым обрабатывают стекла, и состава пробы. Видно, что при использовании полилизина или аminosилана есть небольшие различия, но еще больше различия получаются, если использовать, например, буферы для ПЦР рекций или воду или фосфатный буфер и т.д..

Слайд 17. Процесс фотолитографии по методу Affymetrix

Теперь рассмотрим подробнее процесс фотоиндуцируемого синтеза олигонуклеотидов *in situ* или фотолитографии по методу фирмы Affymetrix, которая является ведущим производителем олигонуклеотидных биочипов высокой плотности.

Сначала проектируется и изготавливается набор масок, трафаретов, которые через которые затем облучают поверхность стеклянной пластины. В экспонированных точках поверхности стекла происходит фотоиндуцируемая реакция присоединения одного нуклеотида, присутствующего в реакционной смеси, омывающей поверхность. После смены маски и нуклеотида в смеси подается новый импульс света и присоединение еще одного нуклеотида.

Весь набор проб синтезируется одновременно, что повышает воспроизводимость, т.е. равномерность содержания проб. Пластина затем делится на отдельные биочипы и упаковывается в картриджи, которые служат для защиты проб и как камеры для последующей гибридизации.

Теперь мы можем отметить ограничения этого способ приготовления ДНК-биочипов – во-первых, дороговизна, обусловленная именно первым этапом – дизайн и изготовления масок на каждый акт присоединения одного нуклеотида. Т.е. если мы хотим иметь пробы длиной 20 н.о., то мы должны создать набор из 20 масок, которые можно применить для производства олигонуклеотидных биочипов только одного уникального содержания; во-вторых – невозможность оперативно менять дизайн олигонуклеотидных биочипов, т.е. содержание и расположение проб в зависимости от нужд исследуемой биологической модели, от появления новых результатов или данных и т.д..

Однако эти ограничения преодолеваются самой фирмой Affymetrix в основном путем массового производства широкого ассортимента олигонуклеотидных биочипов для множества организмов.

Слайд 18. Affymetrix: Структура элементов биочипа

Слайд 19. Affymetrix: Структура элементов биочипа

Рассмотрим структуру элементов олигонуклеотидных биочипов высокой плотности фирмы Affymetrix. В таких биочипах одному гену соответствуют набор из 11-20 проб. В каждом элементе биочипа (probe cell) миллионы копий олигонуклеотидных проб, и каждый олигонуклеотид

представляет собой 25-мер. Каждая проба представлена парой олигонуклеотидов, а именно олигонуклеотидом, являющимся «пробой совершенного соответствия» (perfect match (PM) probe), и олигонуклеотидом, являющимся «пробой нарушенного соответствия» (mismatch (MM) probe), которая идентична PM-пробе за исключением 13-ой позиции посередине, куда введена однонуклеотидная замена (А на Т или С на G).

Слайд 20. Affymetrix: Схема технологии

Общая схема GeneChip-технологии такова: есть изготовленные на фирме биочипы, пользователю надо самому приготовить меченые образцы и провести гибридизацию с помощью специальных фирменных установок, гарантирующих выполнение всех необходимых оптимальных условий.

Для гибридизации с микробиочипом выделенная мРНК сначала конвертируется в одноцепочечную кДНК, затем в двухцепочечную кДНК, которая амплифицируется до некоторой степени и затем используется как матрица в *in vitro* транскрипции с промотора для РНК полимеразы фага Т7 для производства меченых биотином кРНК (комплементарных РНК). Именно этот пул меченых кРНК используется для гибридизации.

Выявление результатов гибридизации производится с помощью стрептавидин-фикоэритринового конъюгата и сканирования сигнала от фикоэритрина

Слайд 21. Affymetrix: Вид сигналов гибридизации

В результате каждого сканирования сигнала гибридизации получается псевдоаналоговое изображение микробиочипа.

Значения интенсивности сигнала для каждого пикселя этого изображения представлены в виде .DAT-файла.

На изображении сигнала от одного набора проб видно, что не все пробы для гена одинаково информативны – иногда ММ-проба светится также ярко, как и РМ-проба, что означает, что сигнал от этой пары мы будем вынуждены отбросить. Иногда одинаково не светятся ни РМ-, ни ММ-пробы, что может означать, что в образце нет именно этой изоформы РНК, последовательности которой соответствует рассматриваемая пара проб. Или процесс приготовления кДНК прошел неудачно, и участки мРНК, которым соответствует рассматриваемая пара проб, не оказались представлены в пулах кДНК и кРНК.

Эти примеры показывают, насколько важно знать и уметь разбираться в структуре проб Affymetrix –биочипов, чтобы извлекать из результатов экспериментов надежные и осмысленные заключения о поведении изучаемого биологического объекта.

Слайд 22. Affymetrix: Вычисление сигнала

При вычисления сигнала или корректировки фона (background adjustment) измеряется суммарная интенсивность сигнала по пикселям каждой ячейки как одно значение для каждого элемента (cell), и полученные данные представляются в .CEL-файле. Для этого проводятся преобразования данных по следующему алгоритму:

- Каждый элемент состоит из 10x10 пикселей;
- Для каждого элемента (и РМ, и ММ) из расчетов удаляются внешние 36 пикселей для снижения влияния фона, остаются 8x8 пикселей;
- За значение сигнала какого-либо элемента берется 75-й перцентиль от суммарного значения для 8x8 пикселей;
- За значение фона в одном из 16-ти секторов биочипа принимается среднее от 2% элементов в этом секторе с самыми низкими значениями;
- За значение сигнала для каждой пары проб берется разница сигналов РМ и ММ.

Было разработано несколько вариантов алгоритмов вычисления сигналов в соответствии с разработанными несколькими моделями генерации общего и локального фона или шума. При этом вычисляются коэффициенты поправок на шум (вариации между пикселями в ячейках) и фон с

разных участках биочипа (автофлуоресценция и сигнал неспецифической гибридизации в промежутках между ячейками).

Слайд 22. Affymetrix: Типы файлов с результатами

Суммируем, какие типы файлов и с каким содержанием генерируются в результате проведения экспериментов с Affymetrix-биочипами:

- 1) .DAT файл - изображение, Содержит значения сигнала для ~107 пикселей (~50 Мб).
- 2) .CEL файл - Cell intensity file. Содержит значения РМ и ММ для элементов чипа (~7 Мб).
- 3) .CDF файл - Chip Description File. Содержит данные о положении проб, пар проб и наборов проб на чипе (~7 Мб).
- 4) .EXP файл - Experiment Description File. Содержит описание образцов, установки для проточной системы реактивов и для сканера (~1 кб).
- 5) .RPT файл - Quality report file. Содержит описание параметров качества данных (~2 кб).

Данные экспериментов с Affymetrix-биочипами являются самыми сложными и трудными для использования неподготовленными биологами, стремящимися извлечь из этих экспериментов новые сведения о своей исследуемой биологической системе. Поэтому и важно в этом разобраться, чтобы незнание структуры и типа этих данных не становилось препятствием для их использования в своей работе.

Слайд 24. Affymetrix: Причины применения нормализации

Видя, насколько процесс извлечения информации с помощью Affymetrix-биочипов (да и других тоже) многостадийный и сложный, невольно задаешься вопросом: а можно ли доверять тем различиям в экспрессии генов, которые выявляются в этих экспериментах? Действительно, можно ожидать, что мы наблюдаем два типа различий:

- Интересующие нас – это различия в экспрессии генов
- и мешающие нам – это различия в приготовлении образцов, различия в обработке биочипов, различия в сканировании биочипов и т.д.

Даже в в одном эксперименте с многими отдельными гибридизациями между биочипами могут быть различия, например, по общей яркости биочипов или общим и локальным значениям фона.

Систематические отклонения могут быть также и из-за различий в количествах РНК в образцах, эффективности гибридизации и настройках или нестабильности сканера. Как можно сравнивать такие результаты?

На графиках показаны статистические параметры распределений для данных, полученных в шести гибридизаций одного эксперимента. Видно, насколько они разные и по положению медианы, и по положению стандартных отклонений, и по положению наименьших и наибольших значений, задающих шкалу массиву данных.

Поэтому к таким данным обязательно необходимо применять давно разработанные методы статистической обработки экспериментальных данных, в частности, процедуры нормировки, стандартизации или нормализации.

Слайд 25. Графическое представление сигналов

Для того, чтобы оценить качество получаемых в результате биочип-экспериментов данных и обнаружить какие-либо отклонения, принято представлять данные биочип-экспериментов в виде графиков, когда по одной оси откладываются значения интенсивности сигнала для генов в одной гибридизации, а по другой оси – значения для этих же генов в другой гибридизации, часто являющейся просто повтором, технической репликой. Тогда по виду «облака» значений можно много заключить о качестве данных.

Другим способом представить данные – это по абсциссе отложить сумму значений сигналов для каждого гена, а по оси ординат – разницу между значениями. В таком представлении еще нагляднее отклонения, т.к. в идеале «облако» значений должно быть симметрично расположено вокруг прямой со нулевым значением.

Слайд 26. Наблюдение систематического отклонения

Если же средняя интенсивность сигнала (красная линия) отклоняется от нулевого значения, это означает присутствие систематического отклонения значений сигналов.

Слайд 27. Affymetrix. Графики распределения данных

На этом слайде показаны результаты для пяти Affymetrix-биочипов (HG-U95A) после гибридизации с мРНК из культур клеток печени (Bolstad et al. 2003) в виде так называемых графиков «М vs. А», где М and А - РМ-значения для каждой пары проб (10 попарных графиков для РМ-значений). Видно, насколько отличаются распределения значений в каждой гибридизации и насколько необходима предварительная статистическая обработка, прежде чем будут выводиться биологически значимые заключения.

Слайд 28. NimbleGen's Maskless Array

Affymetrix-технология, хотя и самая распространенная, не единственная для типа олигонуклеотидных биочипов. Недавно появилась еще одна разновидность фотолитографического способа приготовления олигонуклеотидных биочипов - NimbleGen's Maskless Array Synthesizer (MAS) Photolithography. В этом способе используются микрозеркала, поэтому формирование «временных масок» происходит в соответствии с их предварительным дизайном под контролем компьютера. Понятно, что такой подход способен преодолеть вышеописанные для технологии Affymetrix ограничения.

Слайд 29. Олигонуклеотидные биочипы Agilent

Здесь показаны олигонуклеотидные биочипы еще одной фирмы – Agilent. Особенностью этой технологии является использование *in situ* синтеза длинных 60-мерных олигонуклеотидов, которые обладают очень высокой специфичностью и поэтому для которых уже нет необходимости размещать на биочипе ММ-пробы. В этой технологии также используется конкурентная двухцветная гибридизация.

Слайд 30. Олигонуклеотидные биочипы Amersham CodeLink Gene Expression Bioarray System

Еще одна технологии олигонуклеотидных биочипов представлена биочипами CodeLink Gene Expression Bioarray System фирмы GE Amersham. Ее особенность – иммобилизация 30-мерных олигонуклеотидов в объеме, в микрокаплях геля (3-D gel matrix).

Слайд 31. Amersham membrane (filter) array

Фирма GE Amersham производит также и мембранные макробиочипы. Так выглядят результаты гибридизации этих макробиочипы с радиоактивно мечеными образцами.

Слайд 32. SuperArray membrane (filter) array

Еще одна фирма, производящая мембранные макробиочипы – SuperArray. Эти макробиочипы содержат около 100 проб, как правило представляющих гены, относящиеся к какой-либо генной сети, сигнальному или метаболическому пути. Благодаря своей дешевизне и нетребовательности к дорогим сканирующим устройствам, эта технология успешно существует на рынке биочипов, предоставляя средство уже не общегеномного скрининга, а целенаправленного слежения за поведением конкретных генов в различных экспериментальных условиях.

Слайд 33. СХЕМА ИСПОЛЬЗОВАНИЯ кДНК-БИОЧИПОВ

Рассмотрим теперь общую схему применения кДНК-биочипов.

Первый этап - приготовление биочипов. В автоматическом роботизированном режиме клоны их библиотеки кДНК нарабатываются до необходимых количеств и распределяются на поверхности стекла специальными аппаратами, называемыми arrayer или spotter.

Второй этап - приготовление образца (sample) для гибридизации. Из двух источников клеток – тестовых (или испытываемых) и референсных (или эталонов для сравнения), это могут быть клетки двух разных тканей, или клетки одной ткани в двух разных состояниях и т.д., выделяются пулы мРНК, которые обратно транскрибируются в пулы кДНК. В каждый пул кДНК вводится своя флуоресцентная метка, например, в тестовый пул красного цвета, а в референсный – зеленого. Затем оба пула объединяют и получают образец для гибридизации.

Условия гибридизации почти не отличаются от тех, что используются в методах Саузерн- или нозерн- гибридизации, только проводят ее в очень малых объемах – не больше полумиллилитра – и поэтому в специальных камерах или устройствах. Меченные молекулы образца гибридизуются с пробами или мишенями (targets), распределенными на биочипе. Затем проводят отмывку неспецифически гибридизовавшихся молекул.

Затем проводят регистрацию сигналов гибридизации с помощью сканирования. В специальном устройстве (scanner или reader) лазеры генерируют излучение света определенной длины волны, в ответ на которое меченные молекулы образца, удержанные молекулами проб, испускают флуоресцентное свечение также определенной длины волны, которое регистрируется в цифровом виде. В результате получают изображение, снимок поверхности стекла. Необходимо подчеркнуть, что регистрация проводится отдельно по каждой длине волны, т.е. каналу, поэтому изображение получается в черно-белой шкале, отражающей интенсивность флуоресцентного ответа. Затем происходит обработка изображений с помощью компьютера, например, псевдораскрашивание изображений в красный и зеленый цвет, суперпозиция изображений, позволяющая наглядно проявить дифференциальное содержание тех или иных молекул в тестовом или референсном пулах кДНК, выделенных из образцов клеток.

Слайд 34. ПРИГОТОВЛЕНИЕ МЕЧЕННОГО ОБРАЗЦА И ГИБРИДИЗАЦИЯ ДЛЯ кДНК-БИОЧИПОВ

Поподробнее рассмотрим стадию приготовления меченного образца или мишени для кДНКовых биочипов.

Каждый пул кДНК, приготовленный из двух разных образцов, делится на две равные части молекул, затем в молекулы вводятся меченые флуоресцентными группировками нуклеотиды. Это прямое включение. Есть методы и непрямого включения, когда вводятся нуклеотиды, модифицированные активными группировками, к которым затем химически присоединяют флуорохромы. Так поступают для того, чтобы избежать ингибирующего влияния на реакцию полимеризации слишком громоздких флуорохромных группировок или дифференциального влияния разных флуорохромных группировок.

Наиболее распространено использование зеленого цианина3 (Cyanine3 (Cy3)) и красного цианина5 (Cyanine5 (Cy5)).

Как правило, получения данных о дифференциальной экспрессии генов используют двухцветную совместную конкурентную гибридизацию, когда оба меченых образца объединяются.

Слайд 35. ТИПЫ ФАЙЛОВ С РЕЗУЛЬТАТАМИ СКАНИРОВАНИЯ кДНК-БИОЧИПОВ

Как правило, результаты двухцветной гибридизации представлены в виде GPR-файла.

Это файл содержит результатов квантификации по обоим каналам (на дли волн 635 нм и 543 нм) и вычисления статистических параметров, осуществленной программным обеспечением сканеров.

GPR-файл содержит также общую информацию о методе получения изображения и его анализа, о настройках сканера, а также данные, полученные для каждого индивидуального элемента.

Слайд 36. ХАРАКТЕРНЫЕ ТЕРМИНЫ И ПАРАМЕТРЫ

По поводу терминов, используемых в области ДНК-биочиповых экспериментов для обозначения того, что с чем гибридизуется, до сих пор есть разнобой.

Некоторые авторы заявляют, что они склонны придерживаться принципов, установившихся для классических методов Саузерн- и нозерн-блот-гибридизаций: иммобилизованный партнер гибридизации – это мишень или образец, а растворенный партнер гибридизации – это проба или зонд, набор конкретных молекул, с помощью которых характеризуют мишень или образец. Однако, поскольку в методе ДНК-биочипов ситуация прямо противоположная классической: ведь иммобилизованными становятся именно наборы конкретных молекул, то большинство авторов считают, что термины проба или зонд должны относиться к этим молекулам, т.е. факт их предварительной охарактеризованности важнее, чем характер их физического состояния с эксперименте. Соответственно, термины мишень или образец продолжают относиться к совокупности неидентифицированных молекул, которые надо охарактеризовать, хотя они и в виде раствора.

Таким образом, при чтении статей с применением метода ДНК-биочипов, следует сразу разобраться с тем, как авторы используют эти термины.

Теперь продолжим характеристику типов ДНК-биочипов с использованием введенных терминов:

- кДНК-биочипы, как правило, состоят из 40000 кДНК проб длиной 600-2400 н.п..
- Олигонуклетидные биочипы высокой плотности (high-density oligonucleotide arrays) могут содержать до 500,000 пар проб на одном стекле, причем одному гену соответствуют 11-20 проб.

Слайд 37. СКАНИРОВАНИЕ И СРАВНЕНИЕ КАЧЕСТВА РЕЗУЛЬТАТОВ ГИБРИДИЗАЦИИ

Теперь перейдем к стадии сканирования результатов гибридизации и получение изображений.

На слайде показаны результаты сканирования нескольких ДНК-биочипов с одинаковым содержанием, но изготовленных по разным технологиям и на разных аппаратах. Видно, насколько разными получаются изображения в результате сканирования результатов гибридизации: на одних картинках большой фон, на других - неоднородный фон, на третьих – неровные пятна и т.д.. Собственно на этом этапе видно суммирование малозаметных случайных технических неравномерностей процессов на всех предшествующих этапах, приводящее к значительному статистическому разбросу данных при проведении ДНК-биочиповых экспериментов.

Слайд 38. СКАНИРОВАНИЕ РЕЗУЛЬТАТОВ ГИБРИДИЗАЦИИ И ПОЛУЧЕНИЕ ИЗОБРАЖЕНИЙ

На этой стадии оператор уже может сделать заключение о том, достаточно ли качественно проведен биочип-эксперимент и стоит ли проводить обработку данных.

На верхней картинке показано изображение нормальной биочип-гибридизации. На средней показано, как выглядят дефекты отмывки. На нижней - как выглядит гибридизация, если образец РНК деградирован.

Слайд 39. СРАВНЕНИЕ РАЗНЫХ БИОЧИПОВ

А на этой картинке показаны изображения, полученные для разных биочипов:

- A. напечатанный кДНК-биочип
- B. Agilent-биочип 60-меры
- C. Amersham 3-D 30-меры
- D. Affymetrix-биочип высокой плотности.

Можно заметить, насколько различаются по техническим параметрам эти биочипы. Поэтому задача приведения результатов экспериментов с использованием разных типов биочипов к некоему единому формату, позволяющему сравнивать результаты и использовать их, независимо от технологий, очень сложна, но и потребность в решении этой задачи трудно переоценить!

Слайд 40. ОБРАБОТКА ИЗОБРАЖЕНИЙ

Итак, рассмотрим этап обработки изображений, с которого начинается анализ цифровой информации.

Исходные данные, получаемые после сканирования по каждому каналу, - это, как вы помните, псевдоаналоговое изображения определенного района поверхности стекла в черно-белой шкале. Обычно это 16-битные TIFF (Tagged Information File Format) изображения.

Они преобразовываются в цифровые данные об интенсивности сигнала гибридизации после:

- определения центра пробы (регистрации)
- выделение пикселей картины, относящихся к пробе и не-пробе (сегментация)
- определение значений интенсивности сигнала от пробы (как суммированной величины значений для пикселей каждого сегмента пробы) и определение значений фона (как суммированной величины значений для пикселей каждого сегмента не-пробы) (квантификация).

Слайд 41. ПРЕОБРАЗОВАНИЕ ДАННЫХ, ПОЛУЧЕННЫХ МЕТОДОМ ДНК-БИОЧИПОВ

Рассмотрим теперь в общем виде, как же получают данные биочип-экспериментов. После сканирования изображений результатов гибридизаций получают сначала матрицы измерений интенсивностей сигналов гибридизаций для набора экспериментов. Это первичные данные, генерируемые сканерами. Затем после процедур математической и статистической обработок получают матрицы данных по экспрессии генов, в которой сведены данные какого-либо набора экспериментов.

Слайд 42. ИСТОЧНИКИ РАЗБРОСА ХАРАКТЕРИСТИК

Рассмотрим в общем виде источники разброса характеристик сканированных изображений.

Систематические ошибки происходят от различий в:

- количествах образцов
- эффективности выделения РНК
- эффективности обратной транскрипции
- введения метки
- эффективности детекции сигнала.

Т.е. эти факторы имеют сходный эффект на многие измерения и поддаются коррекции на основе анализа данных для калибровки отдельных этапов технологического процесса ДНК-биочипового эксперимента.

Стохастические ошибки связаны с различиями в:

- успешности ПЦР и качестве ДНК в пробах
- эффективности раскапывания/печати, отражающейся на размере пятен и их морфологии, а также в присутствии в какой-то мере кросс-гибридизации и неспецифической гибридизации.

Эти факторы случайны и не учитываемы, представляют собой естественный «шум» и фон, и поддаются коррекции с помощью моделирования ошибки.

Слайд 43. ПРИМЕНЕНИЕ ДНК-БИОЧИПОВ В БИОЛОГИИ

Теперь отвлечемся от технологических проблем и рассмотрим, как используются ДНК-биочипы в биологии.

Самыми заметными является их использование при анализе полиморфизма, при сравнительном анализе геномов и анализе экспрессии генов. Последнее приложение я не буду рассматривать в рамках этого цикла лекций, поскольку эта обширная тема нуждается в более подробном рассмотрении в рамках цикла, посвященного современным методам анализа экспрессии генов.

Слайд 44. АНАЛИЗ ПОЛИМОРФИЗМА В ГЕНОМНОЙ ДНК

Методы общегеномного анализа полиморфизма в геномных последовательностях генов, особенно их регуляторных областей, с помощью ДНК-биочипов позволяют делать выводы об их влиянии на изменения в транскриптоме, т.е. собственно о функциональных проявлениях геномного полиморфизма.

Самый распространенный тип полиморфизма в геноме – это однонуклеотидные замены или ОНП. Английский термин – SNP (single nucleotide polymorphism).

К настоящему времени предложено большое количество методов выявления SNP, которые можно разбить на категории в зависимости от принципа биохимической реакции, метода детекции или формата анализа (Syvanen, 2001). Биохимическими механизмами, положенными в основу различения аллелей, могут быть:

- гибридизация с аллель-специфическими олигонуклеотидами,
- эндонуклеазное расщепление,
- удлинение праймера,
- присоединение к праймеру только одного нуклеотида (минисеквенирование),
- лигирование двух олигонуклеотидов, инвазивное расщепление и др.

Слайд 45. АНАЛИЗ ПОЛИМОРФИЗМА В ГЕНОМНОЙ ДНК - ТРИ СТРАТЕГИИ

Миллионы нуклеотидных позиций, вариабельных у разных особей (single nucleotide polymorphism – SNP), могут быть скринированы с помощью специально разработанных ДНК-биочипов. SNP-биочипы используются для исследования (1) сцепления между маркерами, (2) неравновесия по сцеплению (linkage disequilibria), (3) потери гетерозиготности (loss of heterozygosity).

На схеме представлены три экспериментальные стратегии для анализа последовательностей ДНК с помощью биочипов. Одноцветный анализ “gain-of-signal” позволяет после гибридизации с чипом, в котором для каждого SNP предусмотрены четыре различающихся по центральной вариабельной позиции аллель-специфичных олигонуклеотида, обнаруживать присутствие нового аллеля благодаря появлению нового сигнала. Двухцветный анализ “loss-of-signal” позволяет после конкурентной совместной гибридизации испытуемого и референсного образцов с SNP-биочипом с большей специфичностью обнаруживать присутствие нового аллеля. Наконец, метод минисеквенирования состоит в проведении реакции аллель-специфичного удлинения олигонуклеотида на один из четырех меченых дидезоксинуклеотидов прямо на чипе.

Слайд 46. АНАЛИЗ ПОЛИМОРФИЗМА В ГЕНОМНОЙ ДНК - ДВУХЦВЕТНЫЙ МЕТОД – ПОТЕРЯ СИГНАЛА

На слайде показана схема мутационного анализа с помощью олигонуклеотидных биочипов в варианте двухцветной гибридизации и детекции по принципу выявления потери сигнала (по материалам статьи Ji et al., Microarray-based method for genotyping of functional single nucleotide polymorphisms using dual-color fluorescence hybridization. Mutat Res. 2004).

После того как выделенные из двух источников образцы геномной ДНК метятся каждый своим флуорохромом, равные количества образцов объединяются для совместной гибридизации. После сканирования по одному каналу видно, как какой-либо ряд из трех реплик проб, соответствующих последовательности одного аллельного варианта, как бы исчезает. А при сканировании по другому каналу именно этот ряд виден, зато другие отсутствуют. Это интерпретируется как гомозигота по конкретному аллельному варианту. Если же ряд не исчезает при сканировании по обоим каналам (на объединенной картинке выглядит желтым), то это гетерозигота. На сопутствующих секвенограммах видны эти ситуации. Однако для того, чтобы получить информацию о состоянии всех полиморфных позиций в протяженных участках генома пришлось бы проделать гигантскую работу по обычному секвенированию, а с помощью ДНК-биочипов, опираясь на полученные в предыдущих исследованиях сведения о расположении и характере ОНП, можно в одном эксперименте получить полный профиль полиморфизма.

Слайд 47. СЕКВЕНИРОВАНИЕ С ПОМОЩЬЮ ОЛИГОНУКЛЕОТИДНЫХ БИОЧИПОВ

Если же у нас нет предварительных сведений о расположении и характере ОНП в интересующем нас участке генома, то необходимо сначала провести выявление таких позиций. Это можно сделать с помощью олигонуклеотидного биочипа, содержащего в качестве проб перекрывающиеся короткие (пентамеры или гексамеры) олигонуклеотиды. Результаты гибридизации можно расшифровать с помощью специально разработанного алгоритма и программного обеспечения, которые будут генерировать варианты переходов от одного олигонуклеотида к другому с соблюдением требований связанности, непрерывности текста.

Слайд 48. СЕКВЕНИРОВАНИЕ С ПОМОЩЬЮ ОЛИГОНУКЛЕОТИДНЫХ БИОЧИПОВ

Однако у метода секвенирования с помощью биочипов с короткими олигонуклеотидами есть существенный недостаток – в участках с повторенными мотивами алгоритм не в состоянии сгенерировать правильную последовательность.

Поэтому используют различные модификации для повышения информативности биочип-секвенирования. Например, применяют наборы вспомогательных меченых разными флуорохромами олигонуклеотидов, взаимодействующих с матрицей встык с иммобилизованной пробой с последующим лигированием. Этот прием позволяет расширить окно расшифровки последовательности. Также этим способом можно детектировать не только нуклеотидные замены, но также небольшие вставки или делеции.

Слайд 49. СЕКВЕНИРОВАНИЕ С ПОМОЩЬЮ MAGIChip™ ТЕХНОЛОГИИ

Эта методика активно применяется в MAGIChip™ технологии, восходящей к разработкам академика Д.А. Мирзабекова. Именно объемный формат олигонуклеотидных биочипов позволяет использовать микрокапли геля как реакционные микрофлюидные камеры для проведения повторяющихся процедур гибридизации, лигирования, отмывки и детекции.

Слайд 50. ПРИМЕНЕНИЕ ДНК-БИОЧИПОВ В БИОЛОГИИ

Теперь перейдем к рассмотрению методов сравнительного анализа генома с помощью ДНК-биочипов.

Слайд 51. СРАВНИТЕЛЬНЫЙ АНАЛИЗ ГЕНОМА С ПОМОЩЬЮ ДНК-БИОЧИПОВ

ДНК- биочипы в этой области функциональной геномики используются для:

- Сравнительно-эволюционного анализа генов;
- Исследования распределения связанных транскрипционных факторов;
- Картирования химических модификаций в хроматине и ДНК;
- Исследования процессов репликации и репарации;
- Выявления различий в числе копий ДНК-фрагментов.

Кроме непосредственного измерения содержания транскриптов множества генов ДНК-биочип-технология позволяет исследовать зависимость транскрипции от таких фундаментальных процессов, как состояние хроматина множества генов, временем репликации этих генов, уровнем трансляции их транскриптов – и все это в масштабах всего генома.

Слайд 52. СРАВНИТЕЛЬНО-ЭВОЛЮЦИОННЫЙ АНАЛИЗ ГЕНОВ

Рассмотрим сравнительно-эволюционный анализ генов на примере исследования механизмов эволюции генов, вовлеченных в развитие болезней человека (Adjaye J, et al., Cross-species hybridisation of human and bovine orthologous genes on high density cDNA microarrays. BMC Genomics. 2004;5(1):83).

Слайд 53. СРАВНИТЕЛЬНО-ЭВОЛЮЦИОННЫЙ АНАЛИЗ ГЕНОВ

На слайде показаны результаты гибридизации специально разработанного кДНК-биочипа с 349 хорошо охарактеризованными генами человека. Каждая проба нанесена 20 раз для обеспечения высокой достоверности квантификации сигналов. Образцами для совместной гибридизации были экстракты РНК из мозга человеческих и коровьих эмбрионов. Были получены выводы об очень высокой корреляции в экспрессии ортологичных генов ($CC = 0.94$)

Слайд 54. СПРАВНИТЕЛЬНО-ЭВОЛЮЦИОННЫЙ АНАЛИЗ ГЕНОВ

Эти результаты, очень важные для выяснения молекулярных механизмов наследственных человеческих нейродегенеративных заболеваний, были проверены независимыми и более точными нацеленными методами, такими как нозерн-блот-гибридизация и реально-временная ПЦР.

Слайд 55. КАРТИРОВАНИЕ СВЯЗАННЫХ ТРАНСКРИПЦИОННЫХ ФАКТОРОВ

На этом слайде показана схема выполнения экспериментов по картированию связанных транскрипционных факторов с помощью ДНК-биочипов, так называемого метода ChIP-on-chip, т.е. хроматиновая иммунопреципитация в сочетании с биочипами.

Сначала выделяются образцы из двух разных источников. Например один источник является клетками, подвергнутыми сначала «нокауту» по какому-то гену транскрипционного фактора (удалению), а затем трансфекции конструкцией с вариантом этого гена с добавлением участка для кодирования определенного эпитопа в составе транслируемого белка. Другой же источник является стандартными клетками. После проведения *in vivo* кросс-линкирования выделяют хроматин из образцов, проводят раздельно иммунопреципитацию с помощью антител к введенному эпитопу и другим эпитопам. Затем фрагменты ДНК в составе отселектированного хроматина выделяются, метятся и используются для гибридизации с биочипами. В результате выявляются различия в профилях гибридизации, на основании которых делаются выводы о распределении изучаемого транскрипционного фактора.

Слайд 56. КАРТИРОВАНИЕ ХИМИЧЕСКИХ МОДИФИКАЦИЙ В ХРОМАТИНЕ И ДНК

На этом слайде показана схема выполнения экспериментов по картированию химических модификаций в хроматине и ДНК, например, метилированию цитозинов или метилированию, ацетилированию гистонов, и т.д.

Принцип здесь такой же как в предыдущем методе – главное придумать, как сделать контрастными изучаемые различия и выявить эти различия с помощью гибридизации с биочипами.

Например, здесь показано, что если мы будем использовать чувствительные к метилированным основаниям эндонуклеазы рестрикции, то паттерн фрагментов рестрикции в разных образцах геномной ДНК, про которые мы подозреваем, что они различаются по степени метилированности промоторов ряда генов, будет различаться. И это можно проявить с помощью специально разработанного ДНК-биочипа. На основании это сделать заключение о роли метилирования ДНК как одного из эпигенетических механизмов в регуляции транскрипционной активности генов.

Слайд 57. КАРТИРОВАНИЕ МЕТИЛИРОВАННЫХ CpG-ОСТРОВКОВ

Рассмотрим, как еще можно с помощью ДНК-биочипов исследовать связь между экспрессией генов и эпигенетическим состоянием геномных районов вокруг них, например, исследовать степень метилированности ДНК.

Метилирование ДНК – один из эпигенетических механизмов у человека и млекопитающих. Основной мишенью для метилирования в геноме человека и млекопитающих является цитозин. Чаще всего метилирование происходит в контексте динуклеотидов CpG, хотя CpNG, CC(a/t)GG, CpA и CpT также могут быть метилированы. Так называемые «CpG-островки», как правило, охватывают промоторы и первые экзоны генов. Метилированное состояние «CpG-островков» часто ведет к подавлению экспрессии генов. Показано во многих работах, что распределение метилированных сайтов ДНК в нормальных и трансформированных клетках значительно различается – в опухолевых клетках наблюдается гипометилирование одних сайтов ДНК и гиперметилирование других.

Слайд 58. КАРТИРОВАНИЕ МЕТИЛИРОВАННЫХ CpG-ОСТРОВКОВ

На этом слайде показан принцип исследования связи между экспрессией генов и степенью метилированности их ДНК по материалам статьи Novik K.L., et al., 2002, Epigenomics: genome-wide study of methylation phenomena (Curr Issues Mol Biol. 4(4):111-128). Образцы из фрагментов геномной ДНК обрабатываются бисульфитом натрия, превращающего неметилированные цитозины в урацилы, затем эти образцы метятся в процессе амплификации и гибридизуются с олигонуклеотидными биочипами. Для каждого потенциального сайта метилирования разрабатывается пара олигонуклеотидов. В случае если сайт метилирован, то в анализируемой позиции образца останется цитозин, и сигнал после гибридизации будет от олигонуклеотида, содержащего в соответствующей позиции гуанозин (правая часть левого рисунка). А если сайт не метилирован, то сигнал будет от олигонуклеотида, содержащего в соответствующей позиции аденозин (левая часть левого рисунка). Если использовать двухцветную совместную гибридизацию испытываемого и референсного образцов, то анализ расположения зеленых, красных или желтых (т.е. образец представляет собой или гетерозиготу, или гетерогенную смесь клеток) сигналов на ДНК-биочипе позволит определить состояние метилированности каждого потенциального сайта метилирования ДНК каждого исследуемого гена. А сопоставление этих данных с результатами параллельного биочип-эксперимента по определению количества мРНК этих исследуемых генов в этих же образцах позволило выявить искомые корреляции между процессами метилирования ДНК генов и их уровнем транскрипции.

Слайд 59. ИССЛЕДОВАНИЕ ДИНАМИКИ РЕПЛИКАЦИИ

Рассмотрим, как можно исследовать связь между экспрессией генов и временем репликации соответствующих районов генома на примере статьи Schübeler D., et al., 2002, Genome-wide DNA replication profile for *Drosophila melanogaster*: a link between transcription and replication timing. Nature Genet. 32:438-442.

На рисунке показан профиль клеточного цикла, выявляемый после импульсного введения в клетки бром-дезокс-уридина (BrdU) и окрашивания клеток ДНК-специфическим красителем (пропидиум иодидом). Используя метод FACS (fluorescence-activated cell sorting), клетки затем сортируются по содержанию ДНК в зависимости от флуоресценции пропидиум иодида. Затем проводят иммунопреципитацию новосинтезированной ДНК антителами против BrdU, амплификацию осажденных фрагментов ДНК и введение в ампликоны флюорометки, в данном случае образцы из клеток, находящихся на стадии ранней репликации, помечены зеленым Cy3, а образцы из клеток, находящихся на стадии поздней репликации, помечены красным Cy5.

Совместная двухцветная гибридизация меченых образцов с кДНК-биочипом, содержащим 6500 генов дает возможность выявить момент репликации каждого конкретного гена. В этом эксперименте были использованы три биологические реплики, т.е. для приготовления образцов все процедуры и манипуляции с клетками и ДНК проводились независимо три раза.

Слайд 60. ИССЛЕДОВАНИЕ ДИНАМИКИ РЕПЛИКАЦИИ

После обработки данных гибридизации меченых образцов с кДНК-биочипом был получен репликационный профиль для 6500 генов из секвенированной части генома *D. melanogaster*. А сопоставление этих данных с результатами параллельного биочип-эксперимента по определению количества мРНК этих генов в этих же образцах позволило выявить корреляции между транскрипционной активностью гена и временем его репликации: чем выше уровень транскрипции гена, тем раньше он реплицируется, и наоборот.

Слайд 61. ВЫЯВЛЕНИЕ РАЗЛИЧИЙ В ЧИСЛЕ КОПИЙ ДНК-ФРАГМЕНТОВ

На этом слайде показаны схемы выявления различий в числе копий ДНК-фрагментов с помощью ДНК-биочипов. Мы видим стандартную схему двухцветной гибридизации. В этом приложении используются возможности биочип-технологий давать количественную информацию о присутствии

в образце отдельных фрагментов ДНК. Справа показан профиль интенсивности сигнала гибридизации, который интерпретируется как профиль копийности отдельных участков генома, что часто сопровождает процесс малигнизации клеток или опухолеобразования.

Слайд 62. СРАВНИТЕЛЬНАЯ ГЕНОМНАЯ ГИБРИДИЗАЦИЯ на биочипах (array CGH)

Этот метод очень широко применяется для исследования структурных и функциональных (поскольку изменения в числе копий генов – одна из причин изменений экспрессии генов!) преобразований геномов и имеет обозначение array CGH или aCGH. В качестве пособия для более детального ознакомления приведен обзор (Pinkel D., Albertson D.G. Comparative Genomic Hybridization. Annu. Rev. Genomics Hum. Genet. 2005).

Слайд 63. СРАВНИТЕЛЬНАЯ ГЕНОМНАЯ ГИБРИДИЗАЦИЯ на биочипах (array CGH)

Разумеется, такие эксперименты в биочиповом формате очень сложно проводить из-за чрезвычайной сложности генома высших позвоночных, из-за их переполненности повторенными последовательностями, большая часть которых расположена в прицентромерных районах и, до сих пор не расшифрована. Поэтому на процесс гибридизации влияет множество плохо учитываемых факторов, а ответственность за выводы очень высока, особенно в генетико-медицинских исследованиях.

На интенсивность сигнала в элементе биочипа может влиять множество факторов:

- нуклеотидный состав
- присутствие повторенных последовательностей
- количество ДНК в пробе и т.д.

Было обнаружено, что значения интенсивностей могут различаться в 30 раз, даже если различий в копиях гена нет. Следовательно необходимо повысить надежность метода за счет тщательной разработки дизайна экспериментов, контроля за систематическими отклонениями и моделирования случайных отклонений.

Слайд 64. СРАВНИТЕЛЬНАЯ ГЕНОМНАЯ ГИБРИДИЗАЦИЯ на биочипах (array CGH)

На этом слайде рассмотрены некоторые факторы, влияющие на эффективность метода aCGH и эмпирически выведенные зависимости между измеренным отношением сигналов и изменениями в числе копий.

Слайд 65. СРАВНИТЕЛЬНАЯ ГЕНОМНАЯ ГИБРИДИЗАЦИЯ на биочипах (array CGH)

В качестве примера этого метода приведена статья по исследованию профилей копийности геномных последовательностей и профилей экспрессии расположенных в них генов в различных образцах рака легких человека (Tonon G., Wong K-K., Maulik G. et al. High resolution profiles of human lung cancer. PNAS, 2005).

Слайд 66. СРАВНИТЕЛЬНАЯ ГЕНОМНАЯ ГИБРИДИЗАЦИЯ на биочипах (array CGH)

Так выглядит визуализация геномных профилей в образцах первичных легочных аденокарцином (AC), карцином легких (squamous) и клеточных линий из рака легких.

Слайд 67. СРАВНИТЕЛЬНАЯ ГЕНОМНАЯ ГИБРИДИЗАЦИЯ на биочипах (array CGH)

Отдельные геномные районы, кандидаты на роль причины канцерогенеза, проанализированы с большим разрешением и в сочетании с профилированием экспрессии генов. Например, район 3q (от 180 до 199 Мб) демонстрирует значительное различие между линиями AC и SCC как методом aCGH, так и профилированием экспрессии генов.

Слайд 68. ПРИМЕНЕНИЕ ДНК-БИОЧИПОВ В БИОЛОГИИ

Рассмотрим теперь, как анализ экспрессии генов применяется для исследования структуры и функционирования генома.

Слайд 69. АНАЛИЗ ЭКСПРЕССИИ ГЕНОВ: Выявление генов

На следующем слайде показаны пример работы по экспериментальной аннотации человеческого генома (Shoemaker DD, et al., Experimental annotation of the human genome using microarray technology. Nature. 2001 409:922-927).

В этой работе была предпринята попытка оценки компьютерных предсказаний генов и определение полноразмерных транскриптов с помощью выявления одновременной экспрессии их экзонов. Таким способом были проанализирована экспрессия генов из 22 хромосомы (8,183 экзонов) в 69 парах экспериментальных условий. Было выявлено 572 группы корегулируемых экзонов или 572 проверенных по экспрессии гена (expression-verified genes - EVGs). Это составило 210 (85%) из 247 известных генов 22 хромосомы и 185 (57%) из 325 предсказанных генов.

Слайд 70. АНАЛИЗ ЭКСПРЕССИИ ГЕНОВ: Выявление генов

На этом слайде показаны псевдоцветные изображения взвешенного по ошибкам $\log_{10}$ -отношения (красный/зеленый) экспрессии каждого из ~8000 экзонов. Можно видеть, как работает алгоритм выявления коэкспрессии экзонов и «сборка» их в ген.

Слайд 71. АНАЛИЗ ЭКСПРЕССИИ ГЕНОВ: Выявление генов

Затем в этой же работе была проведена проверка полученных результатов с помощью ДНК-биочипов, содержащих 60-мерные фрагменты, перекрывающие геномные последовательности генов. С помощью этих перекрывающихся рядов проб, соответствующих обеим цепям отдельных геномных районов хромосомы 22, были уточнены границы экзонов и самих транскриптов.

Слайд 72. АНАЛИЗ ЭКСПРЕССИИ ГЕНОВ: Выявление генов

В конце концов, в этой работе было проведено сканирование экзонов в масштабах всего генома: было взято 1 090 408 проб, 110 000 обратно-комплементарных проб, 50 биочипов, два экспериментальных условия (РНК из двух клеточных линий). На гистограммах показана в виде красных полос доля экспериментально (для этих условий) проверенных экзонов.

Слайд 73. СООБЩЕСТВО MGED

Учитывая большую сложность, масштабность и зависимость от статистических обработок ДНК-биочиповых данных, можно понять, почему ведущие специалисты в этой области озаботились выработкой стандартов для этих данных. Затратив большие усилия на выработку общих согласованных решений, группа ученых организовала специальное Общество «Microarray Gene Expression Data» (MGED) (<http://www.mged.org>) для установления общих стандартов описания данных по биочип-экспериментам, систем обработки, передачи и хранения данных в общедоступных базах данных.

Среди прочих средств унификации и стандартизации был выработан язык «MicroArray Gene Expression Markup Language» (MAGE-ML) для создания общего формата и достижения сравнимости результатов; протокол «Minimum Information About a Microarray Experiment» (MIAME) для определения типа информации и степени подробности, с которой исследователь обязан ее представить; рабочую группу MGED «Society Ontology Working Group» (<http://www.mged.org/ontology>) для формирования набора контролируемых словарей и онтологий, необходимых для описания биологических образцов и манипуляций с ними, т.е. экспериментальных процедур.

Слайд 74. MINIMUM INFORMATION ABOUT A MICROARRAY EXPERIMENT (MIAME)

Протокол «Minimum Information About a Microarray Experiment» (MIAME) содержит требования к минимальной информации об опубликованном эксперименте, основанном на ДНК-биочип-методе, и включает шесть типов описаний:

1. План эксперимента - набор отдельных гибридизационных экспериментов
2. План биочипа – содержание пятен/ячеек, компоновка по рядам и т.д.
3. Образцы – источник, приготовление экстрактов, способ мечения
4. Гибридизация – процедура и параметры
5. Измерение – характеристики изображений и сканнеров
6. Нормировка – способ, коэффициенты и т.д.

Слайд 75. ПРИМЕНЕНИЕ ДНК-БИОЧИПОВ В МЕДИЦИНЕ

На этом слайде показаны основные области применения ДНК-биочип-технологии в биомедицинских исследованиях. Но об этом разговор пойдет в рамках лекций, посвященных современным методам анализа экспрессии генов.

