



# **Исследование качественных и количественных характеристик транскриптома (3)**

*Катохин Алексей Вадимович, к.б.н.*

Кафедра информационной биологии ФЕННГУ



# Методы массового исследования транскриптов. ДНК-биочипы (микроматрицы)

## ПРИМЕНЕНИЕ ДНК-БИОЧИПОВ В БИОЛОГИИ

АНАЛИЗ ПОЛИМОРФИЗМА

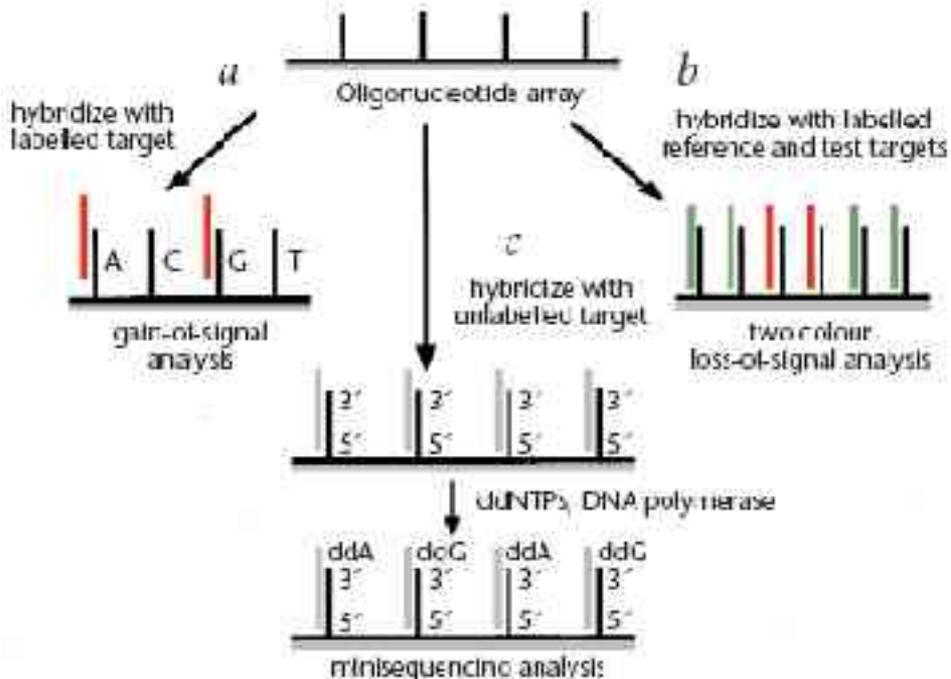
АНАЛИЗ ЭКСПРЕССИИ ГЕНОВ

СРАВНИТЕЛЬНЫЙ АНАЛИЗ ГЕНОМОВ

# Методы массового исследования транскриптов. ДНК-биочипы (микроматрицы)

## АНАЛИЗ ПОЛИМОРФИЗМА В ГЕНОМНОЙ ДНК

Мутационный анализ и ресеквенирование с помощью олигонуклеотидных биочипов.



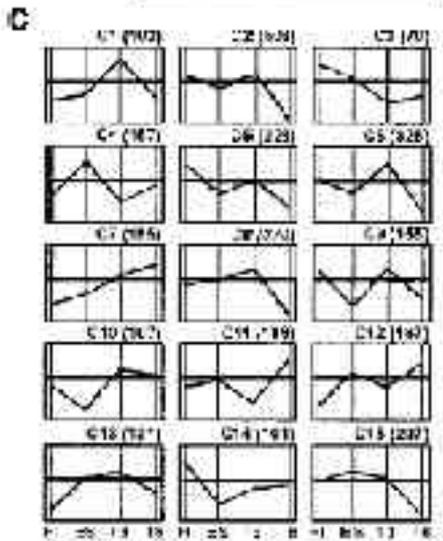
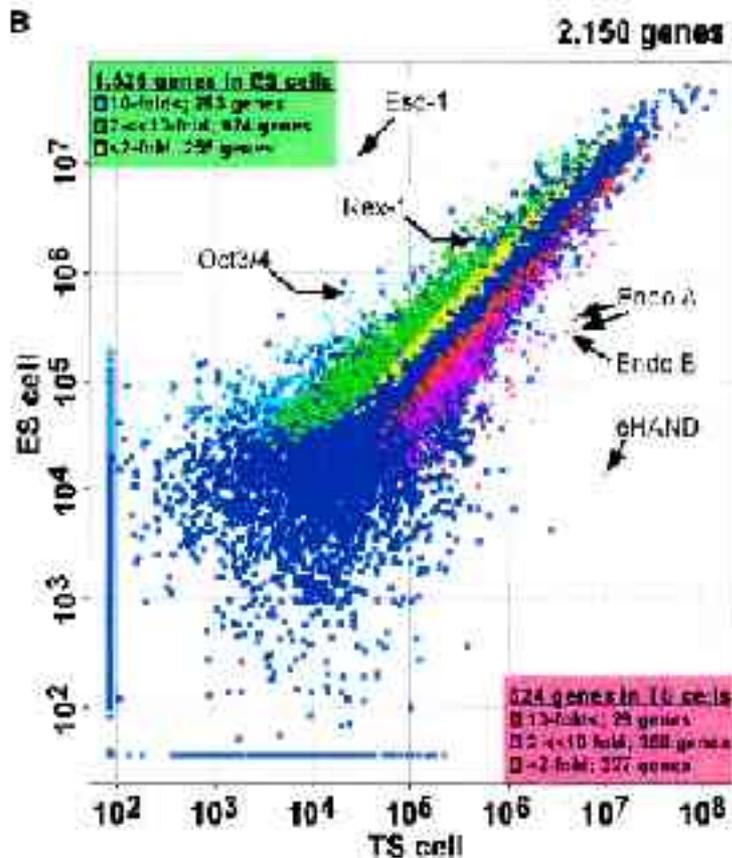
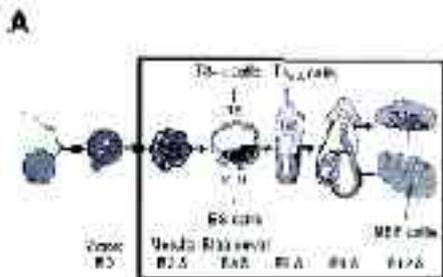
Миллионы нуклеотидных позиций, варьируемых у разных особей, могут быть скринированы с помощью специально разработанных биочипов. SNP-биочипы используются для исследования (1) сцепления между маркерами, (2) неравновесия по сцеплению (linkage disequilibrium), (3) потери гетерозиготности (loss of heterozygosity).

Три экспериментальные стратегии для анализа последовательностей ДНК с помощью биочипов.



# Методы массового исследования транскриптов. ДНК-биочипы (микроматрицы)

## АНАЛИЗ ЭКСПРЕССИИ ГЕНОВ Молекулярный паспорт стволовых клеток.



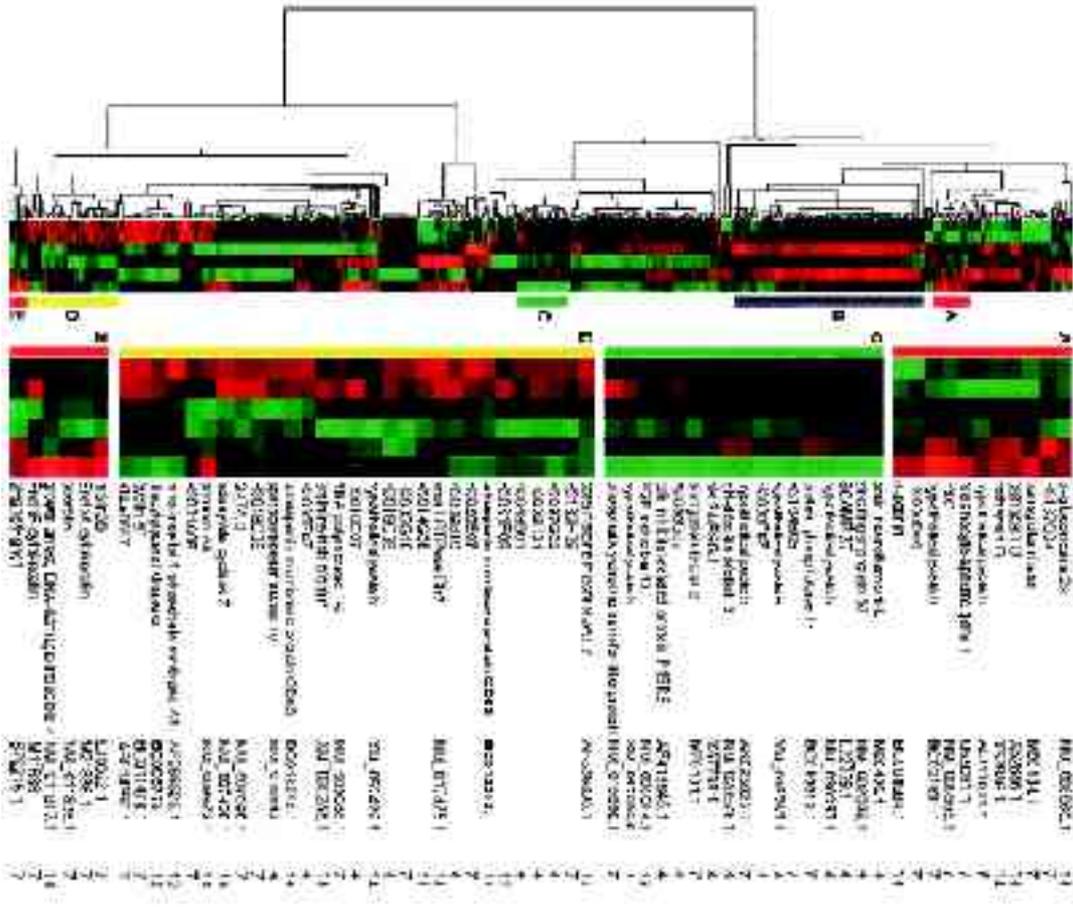
Образцы четырёх клеточных линий:  
ES (эмбрионально-стволовые),  
TS<sub>3,5</sub> и TS<sub>6,5</sub> (трофобласт-стволовые),  
MEF (мышинные эмбриональные фибробласты).  
Мышиный кДНК-биочип NIA 15К.

Гены, экспрессия которых достоверно ( $P < 0.05$ )  
????? ? ??? ? ??????  
?????????, ?????????? ???????.

Кластер-анализ методом «к-средних» дифференциально экспрессирующихся генов: 15 кластеров

# Методы массового исследования транскриптов. ДНК-биочипы (микроматрицы)

## АНАЛИЗ ЭКСПРЕССИИ ГЕНОВ Молекулярный паспорт стволовых клеток.



Иерархическая кластеризация 346 генов, специфично экспрессирующихся в образцах.

???????? A ? E - TS ?????; B - MEF; C – ?????? ? ?????? ?????? ??????????; D – ES ?????.

???????? 4 ????????????????? ES  
 ?????? ?????? ??????, 7 - TS  
 ?????? ??????, 14 - MEF-  
 ?????? ??????, 12 – ??? ?  
 ??? ES ? TS ??????

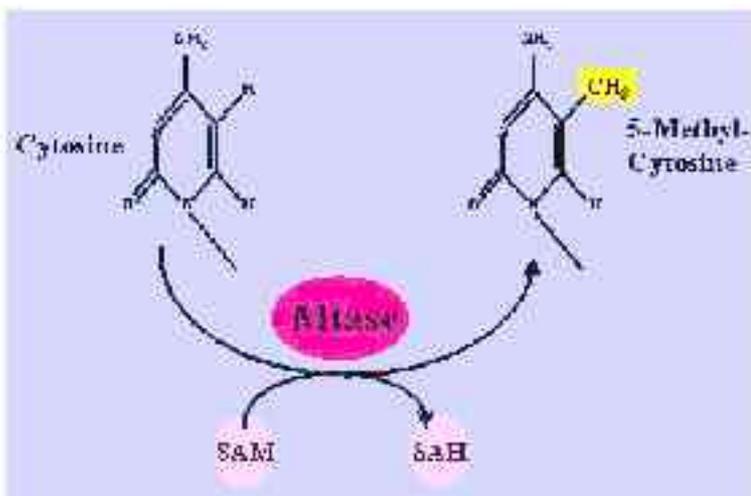




# Методы массового исследования транскриптов. ДНК-биочипы (микроматрицы)

## АНАЛИЗ ЭКСПРЕССИИ ГЕНОВ

Исследование связи между экспрессией генов и эпигенетическим состоянием геномных районов вокруг генов.



Метилирование ДНК – один из эпигенетических механизмов. Основной мишенью для метилирования в геноме млекопитающих является цитозин. Чаще всего метилирование происходит в контексте динуклеотидов CpG, хотя CpNG, CC(a/t)GG, CpA и CpT также могут быть метилированы. «CpG-островки» охватывают промоторы и первые экзоны генов.

Метилированное состояние «CpG-островков» часто ведет к подавлению экспрессии генов.

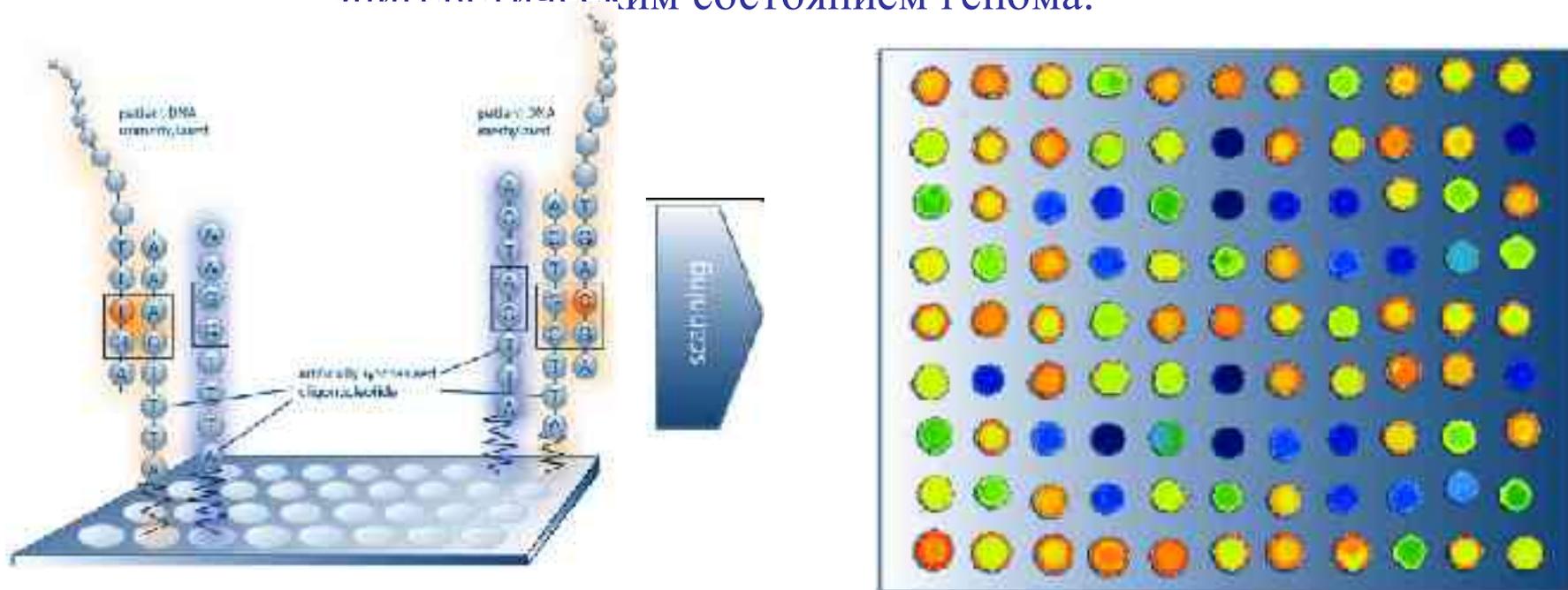
Распределение метилированных сайтов ДНК в нормальных и трансформированных клетках значительно различается – в опухолевых клетках наблюдается гипометилирование одних и гиперметилирование других.



# Методы массового исследования транскриптов. ДНК-биочипы (микроматрицы)

## АНАЛИЗ ЭКСПРЕССИИ ГЕНОВ

Исследование связи между экспрессией генов и эпигенетическим состоянием генома.



Образцы из фрагментов геномной ДНК обрабатываются бисульфитом натрия, превращающего неметилированные цитозины в урацилы, затем метятся в процессе амплификации и гибридизуются с олигонуклеотидными биочипами. Для каждого потенциального сайта метилирования разрабатывается пара олигонуклеотидов.

Novik K.L., *et al.*, Epigenomics: genome-wide study of methylation phenomena.

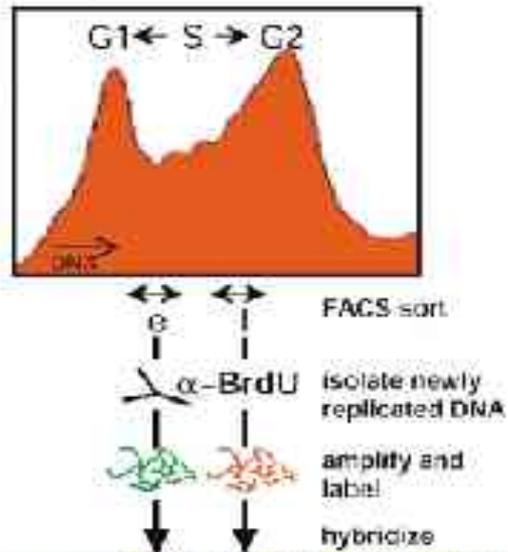
Curr Issues Mol Biol. 2002 4(4):111-128.



# Методы массового исследования транскриптов. ДНК-биочипы (микроматрицы)

## АНАЛИЗ ЭКСПРЕССИИ ГЕНОВ

Исследование связи между экспрессией генов и временем репликации соответствующих районов генома.



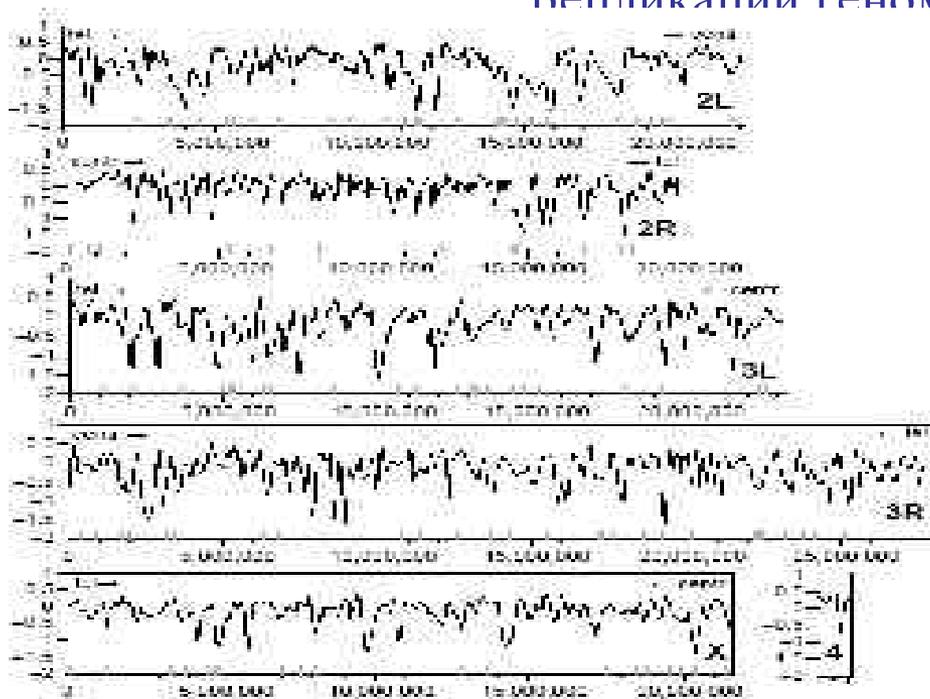
- Профиль клеточного цикла, выявляемый после импульсного введения в клетки BrdU, окрашивания клеток пропидиум иодидом (ДНК-специфическим красителем)
- Сортировка клеток по содержанию ДНК с применением FACS (fluorescence-activated cell sorting).
- Иммунопреципитация новосинтезированной ДНК антителами против BrdU, амплификация и введение флуоресцентной метки.
- Гибридизация с кДНК-биочипом, содержащим 6500 генов. Три биологические реплики.



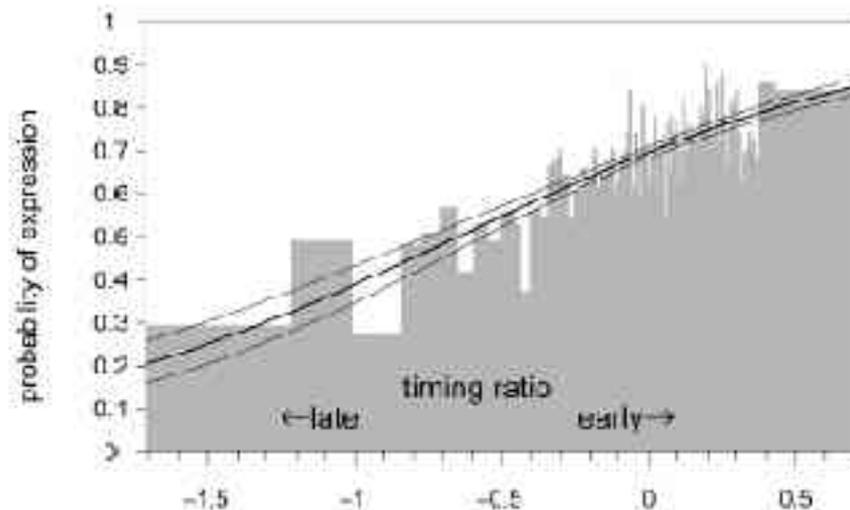
# Методы массового исследования транскриптов. ДНК-биочипы (микроматрицы)

## АНАЛИЗ ЭКСПРЕССИИ ГЕНОВ

Исследование связи между экспрессией генов и временем репликации генома.



Репликационный профиль для 6500 генов секвенированной части генома *D. melanogaster*.



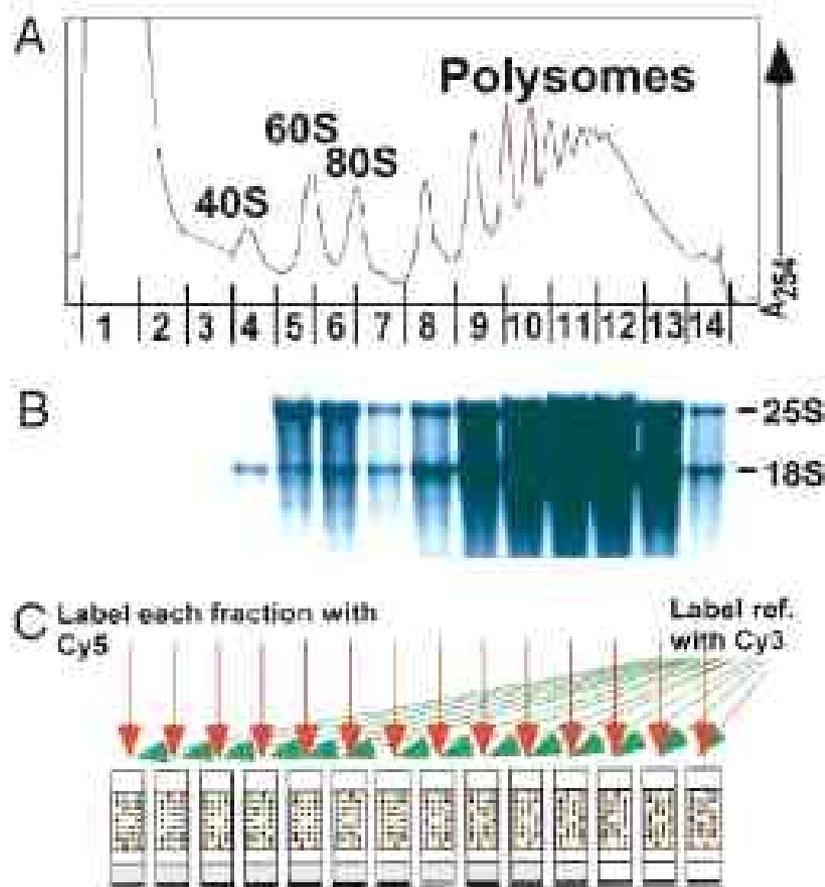
Корреляция между транскрипционной активностью гена и временем его репликации.



# Методы массового исследования транскриптов. ДНК-биочипы (микроматрицы)

## АНАЛИЗ ЭКСПРЕССИИ ГЕНОВ

Исследование связи между транскрипционной и трансляционной активностями генов.



Чем больше уровень трансляции мРНК, тем больше на ней полисом.

– Экспериментальная процедура получения полисомных профилей – седиментация мРНК в сахарозном градиенте.  
– Отбор образцов из разных фракций.

Обратная транскрипция и мечение образцов.

– Гибридизация с ДНК-биочипами, содержащими все открытые рамки считывания

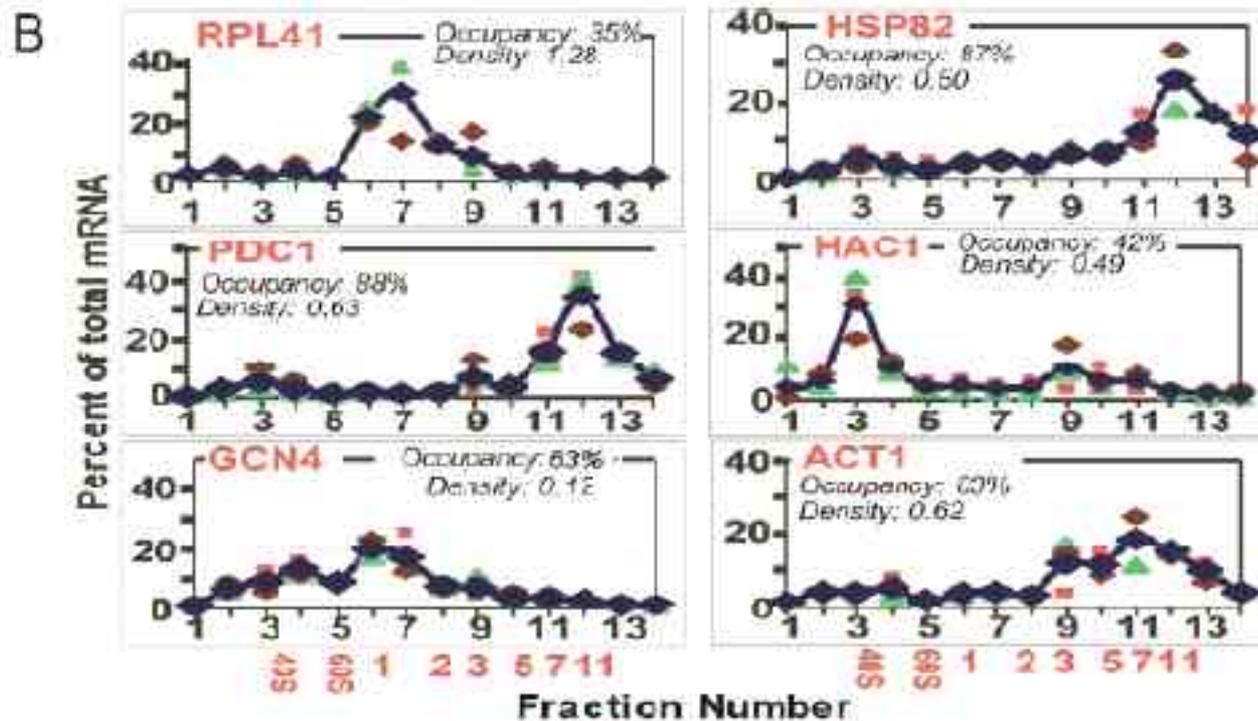
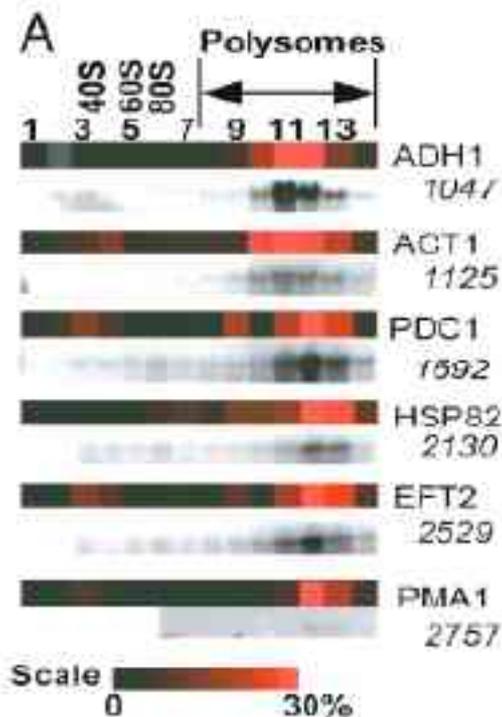
*Saccharomyces cerevisiae*.



# Методы массового исследования транскриптов. ДНК-биочипы (микроматрицы)

## АНАЛИЗ ЭКСПРЕССИИ ГЕНОВ

Исследование связи между транскрипционной и трансляционной активностями генов.



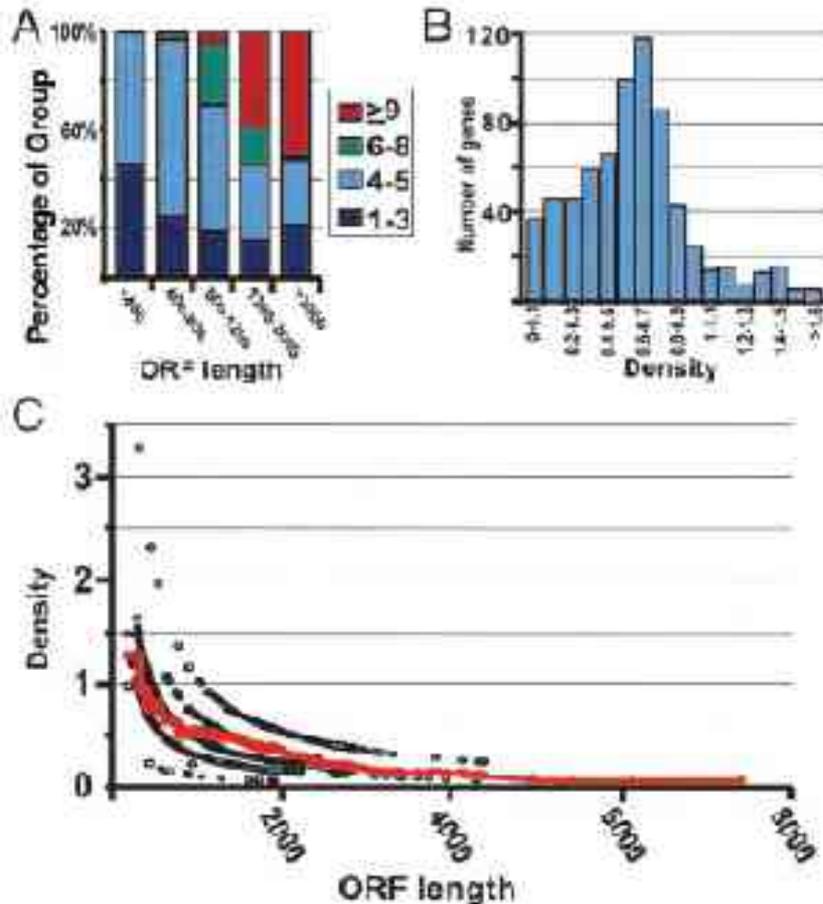
Полисомные профили *Saccharomyces cerevisiae* и их проверка методом нозерн-блот-гибридизации.



# Методы массового исследования транскриптов. ДНК-биочипы (микроматрицы)

## АНАЛИЗ ЭКСПРЕССИИ ГЕНОВ

Исследование связи между транскрипционной и трансляционной активностями ГЕНОВ.



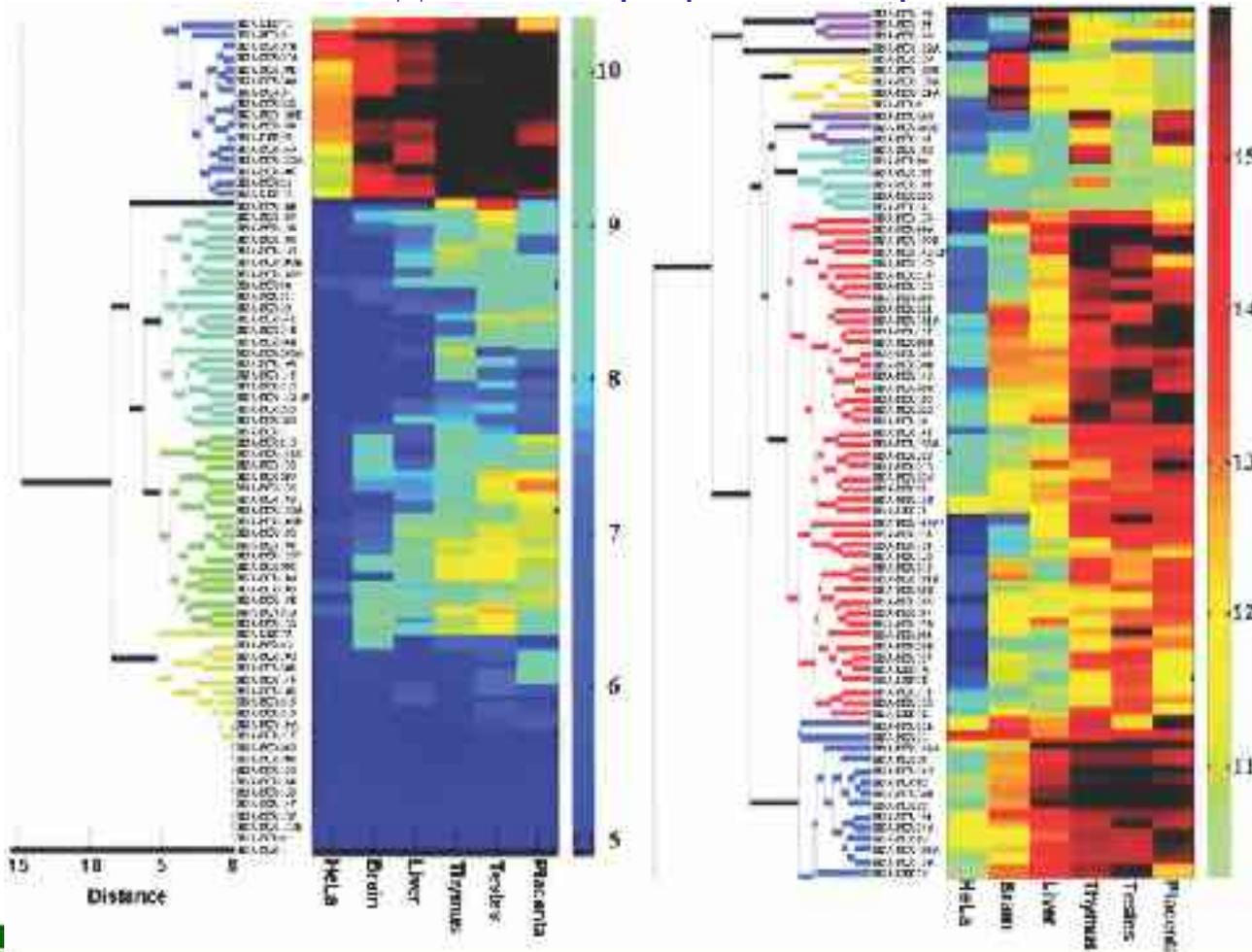
- Число и плотность рибосом, связанных с мРНК:
- A. Гены сгруппированы соответственно длине их ORC
- B. Число генов как функция плотности рибосом
- C. Плотность рибосом как функция от длины ORC



# Методы массового исследования транскриптов. ДНК-биочипы (микроматрицы)

## АНАЛИЗ ЭКСПРЕССИИ ГЕНОВ

Исследование профиля экспрессии миРНК.



Профиль экспрессии 150 миРНК человека в пяти органах и клеточной культуре.

Шкала интенсивности представляет  $\log_2$  значений интенсивности сигнала после вычитания фона и нормализации.



# Методы массового исследования транскриптов. ДНК-биочипы (микроматрицы)

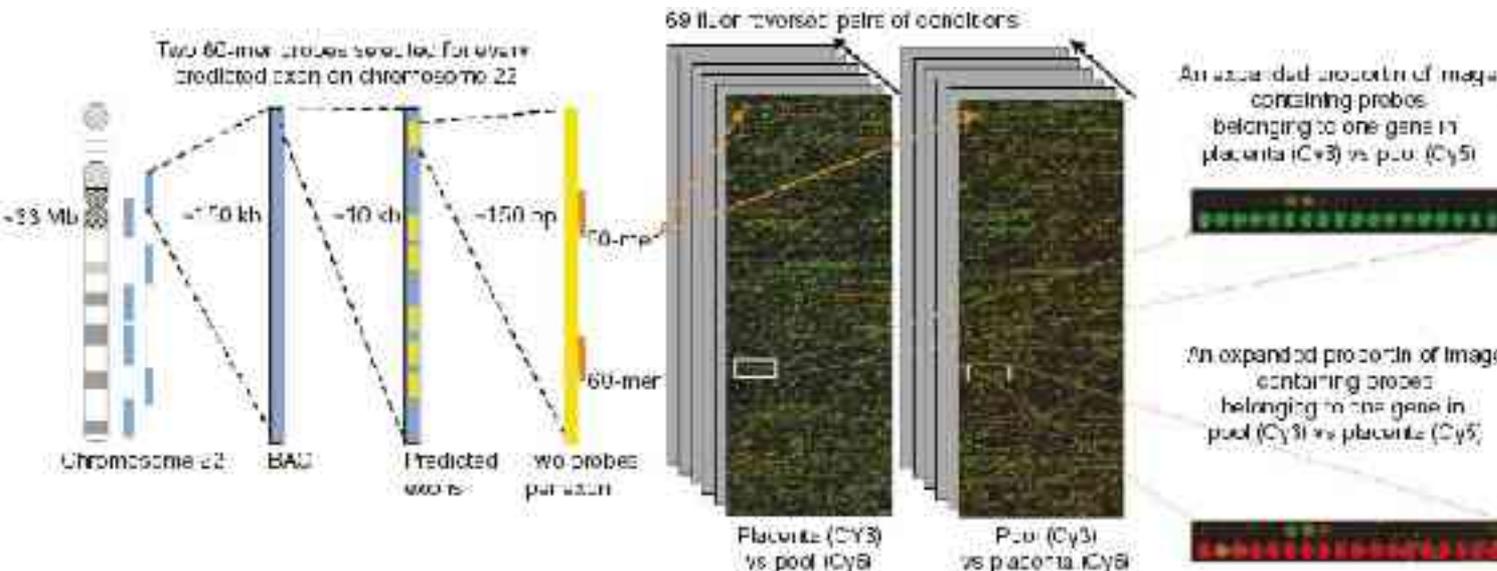
## АНАЛИЗ ЭКСПРЕССИИ ГЕНОВ

Экспериментальная аннотация человеческого генома - оценка компьютерных предсказаний наличия генов и определение полноразмерных транскриптов с помощью выявления одновременной экспрессии их экзонов.

Метод применен для анализа экспрессии генов из 22 хромосомы (8,183 экзонов) в 69 парах экспериментальных условий: 572 группы корегулируемых экзонов или 572

проверенных по экспрессии гена (expression-verified genes - EVGs).

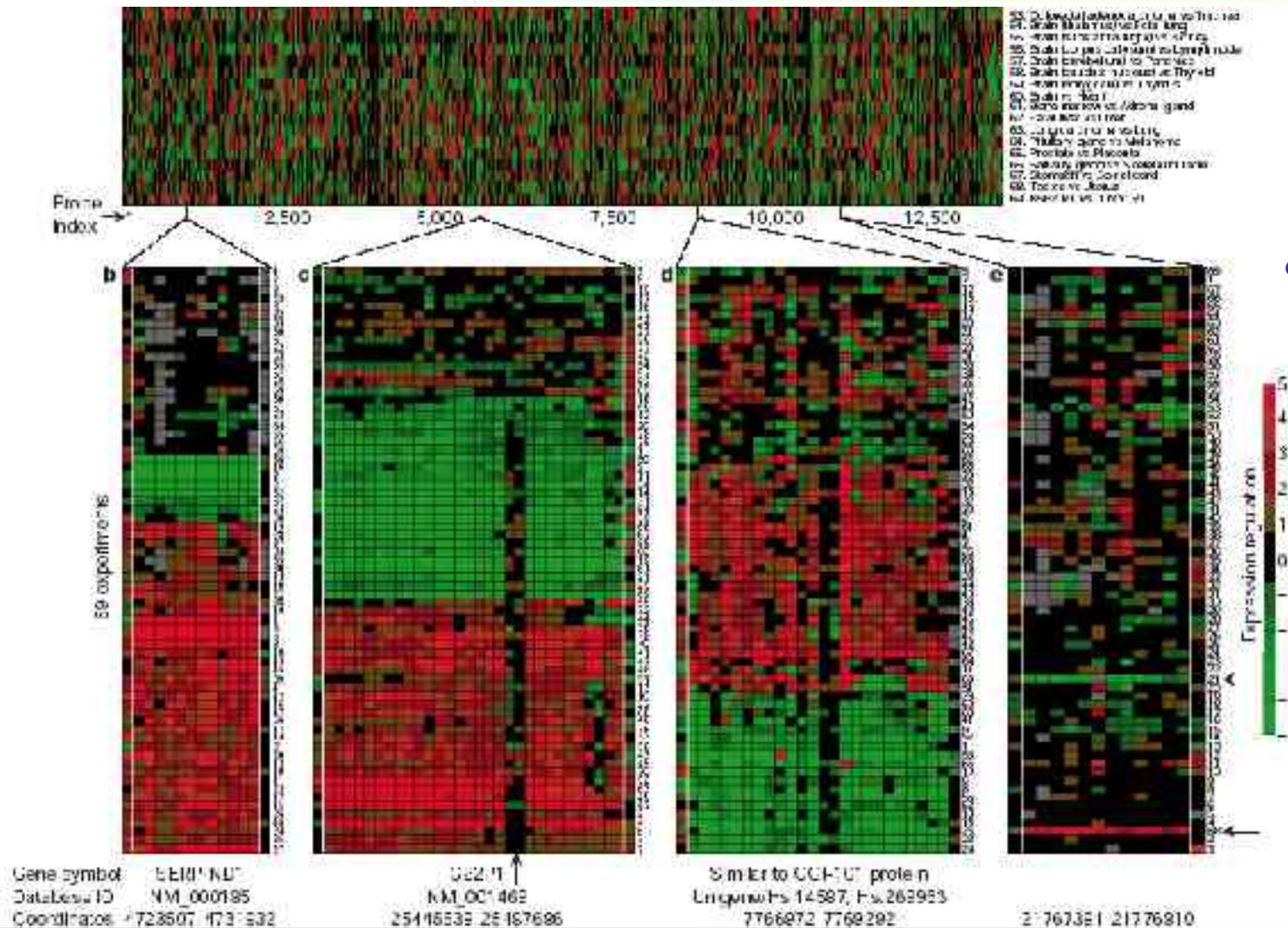
210 (85%) из 247 известных генов и 185 (57%) из 325 предсказанных генов.





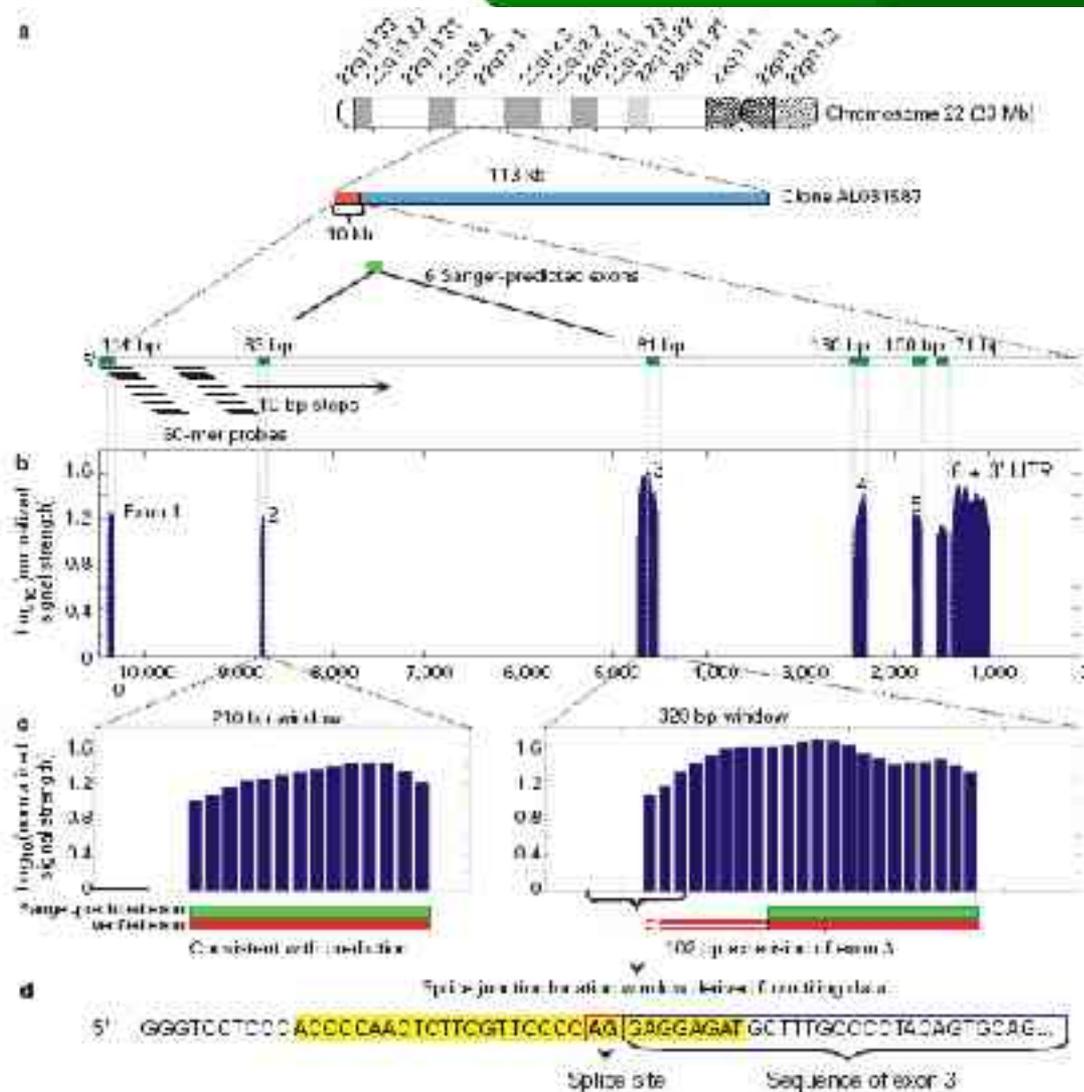
# Методы массового исследования транскриптов. ДНК-биочипы (микроматрицы)

Псевдоцветное изображение взвешенного по ошибкам  $\log_{10}$  отношения (кр/зел) экспрессии каждого из ~8000 экзонов.



Shoemaker DD, *et al.*, Experimental annotation of the human genome using microarray technology. Nature. 2001 409:922-7.

# Методы массового исследования транскриптов. ДНК-биочипы (микроматрицы)

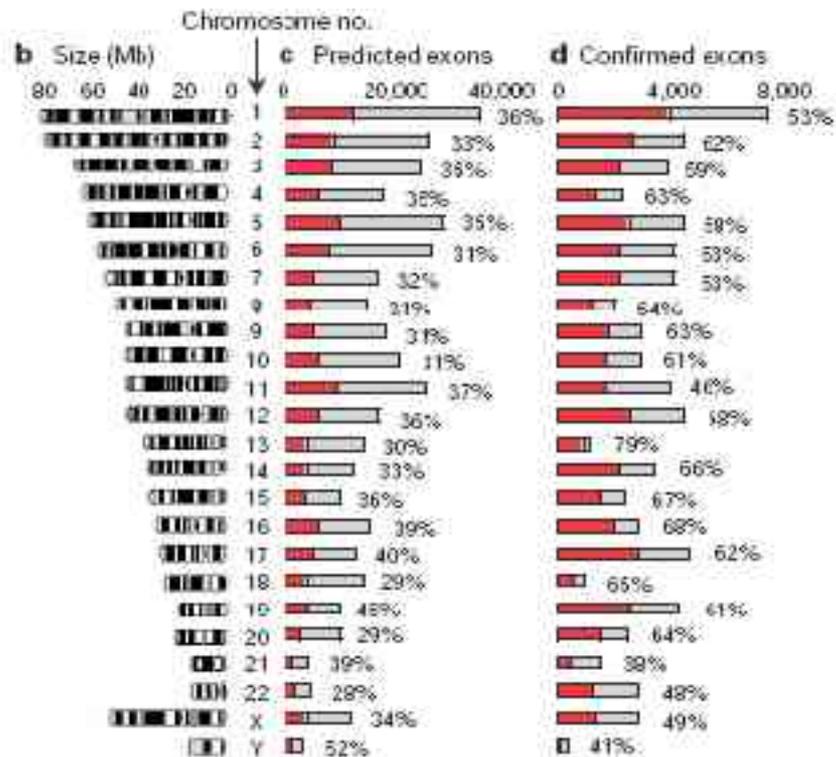
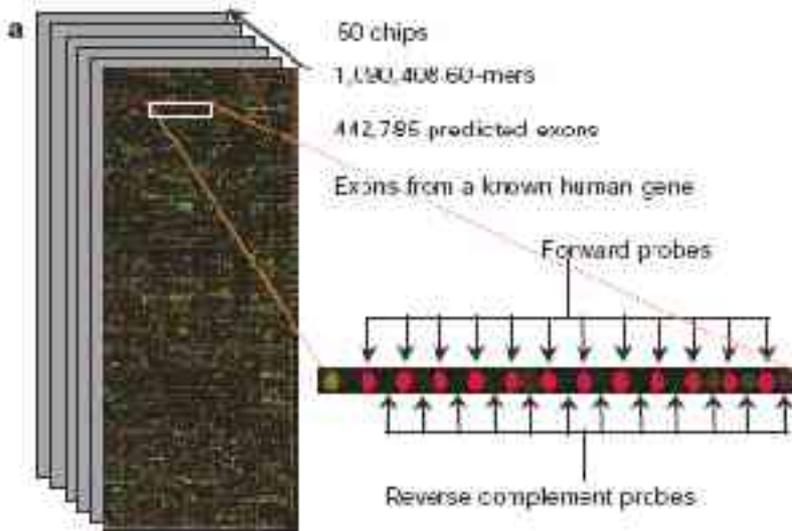


Перекрывающиеся ряды проб, соответствующих обеим цепям отдельных геномных районов хромосомы 22.



# Методы массового исследования транскриптов. ДНК-биочипы (микроматрицы)

Сканирование экзонов в масштабах всего генома: 1 090 408 проб, 110 000 обратнo-комплементарных проб, 50 биочипов, два экспериментальных условия (РНК из двух клеточных линий). Красные полосы – доля экспериментально (для этих условий) проверенных экзонов





# Методы массового исследования транскриптов. ДНК-биочипы (микроматрицы)

## ПРИМЕНЕНИЕ ДНК-БИОЧИПОВ В МЕДИЦИНЕ

ВХОД

ВЫХОД



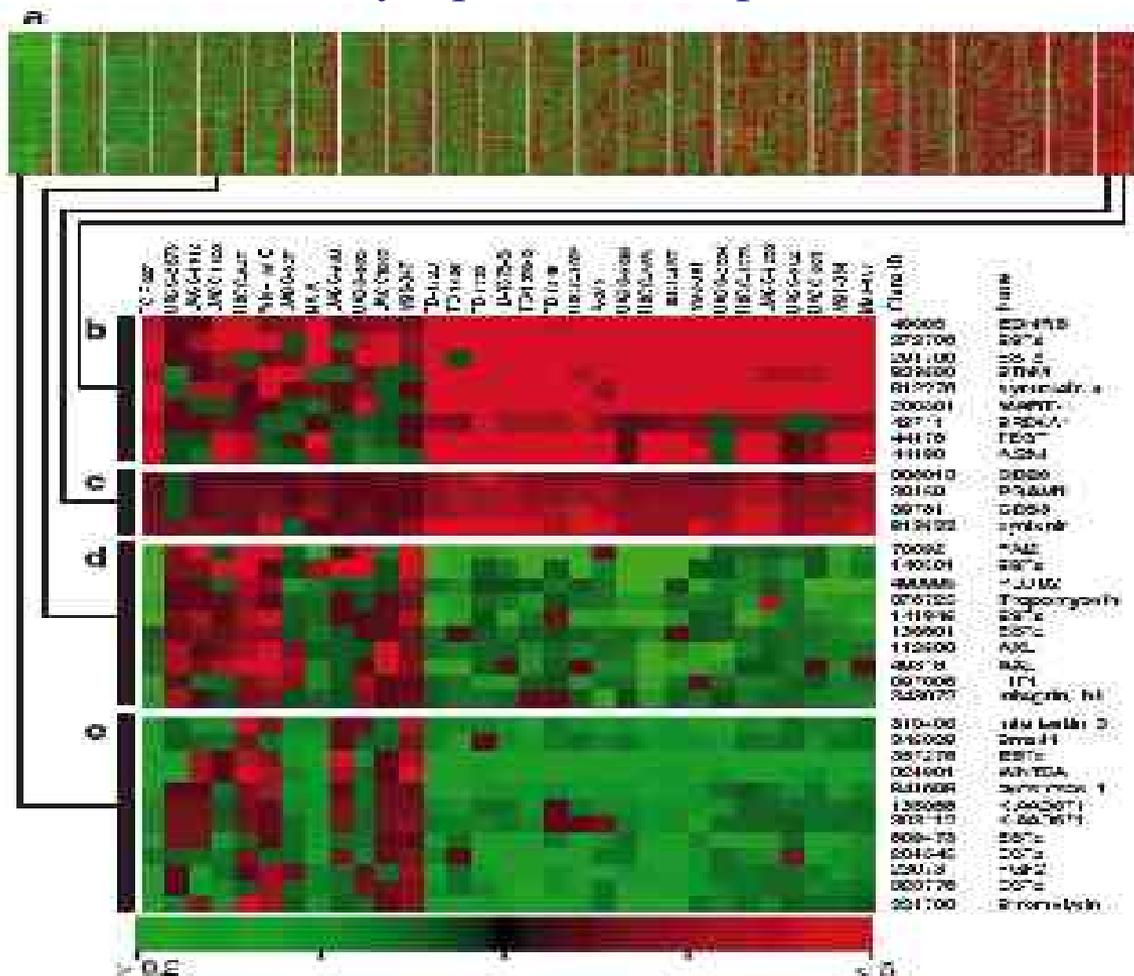


# Методы массового исследования транскриптов. ДНК-биочипы (микроматрицы)



## АНАЛИЗ ЭКСПРЕССИИ ГЕНОВ

Молекулярная классификация злокачественных опухолей.



- Селекция кластеров меланом
- б. Двумерный кластерный анализ образцов опухолей и генов;
- с. Четыре самых контрастных по экспрессии генов в кластерах и характеристические гены в них.

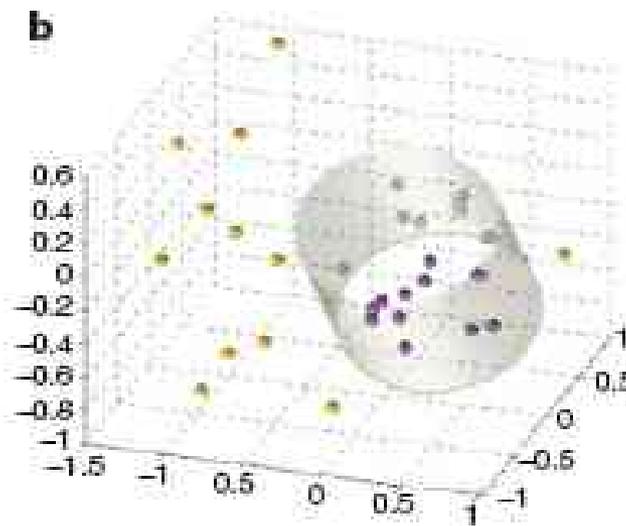
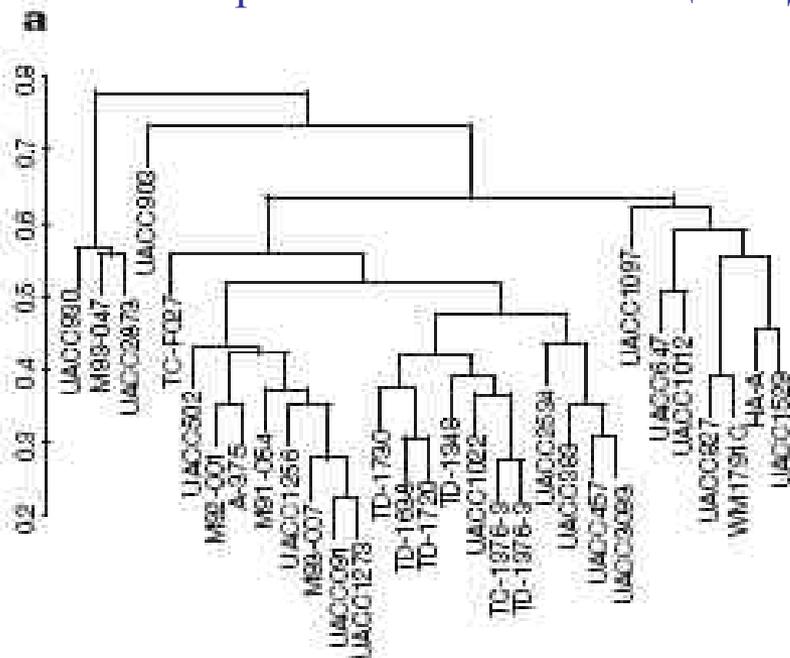


# Методы массового исследования транскриптов. ДНК-биочипы (микроматрицы)

## АНАЛИЗ ЭКСПРЕССИИ ГЕНОВ

Молекулярная классификация злокачественных опухолей.

Таксономия меланом на основе данных профилирования экспрессии генов с помощью кДНК-биочипа (6,971 генов).



а. Дендрограмма иерархической кластеризации 19 меланом по данным экспрессии генов;

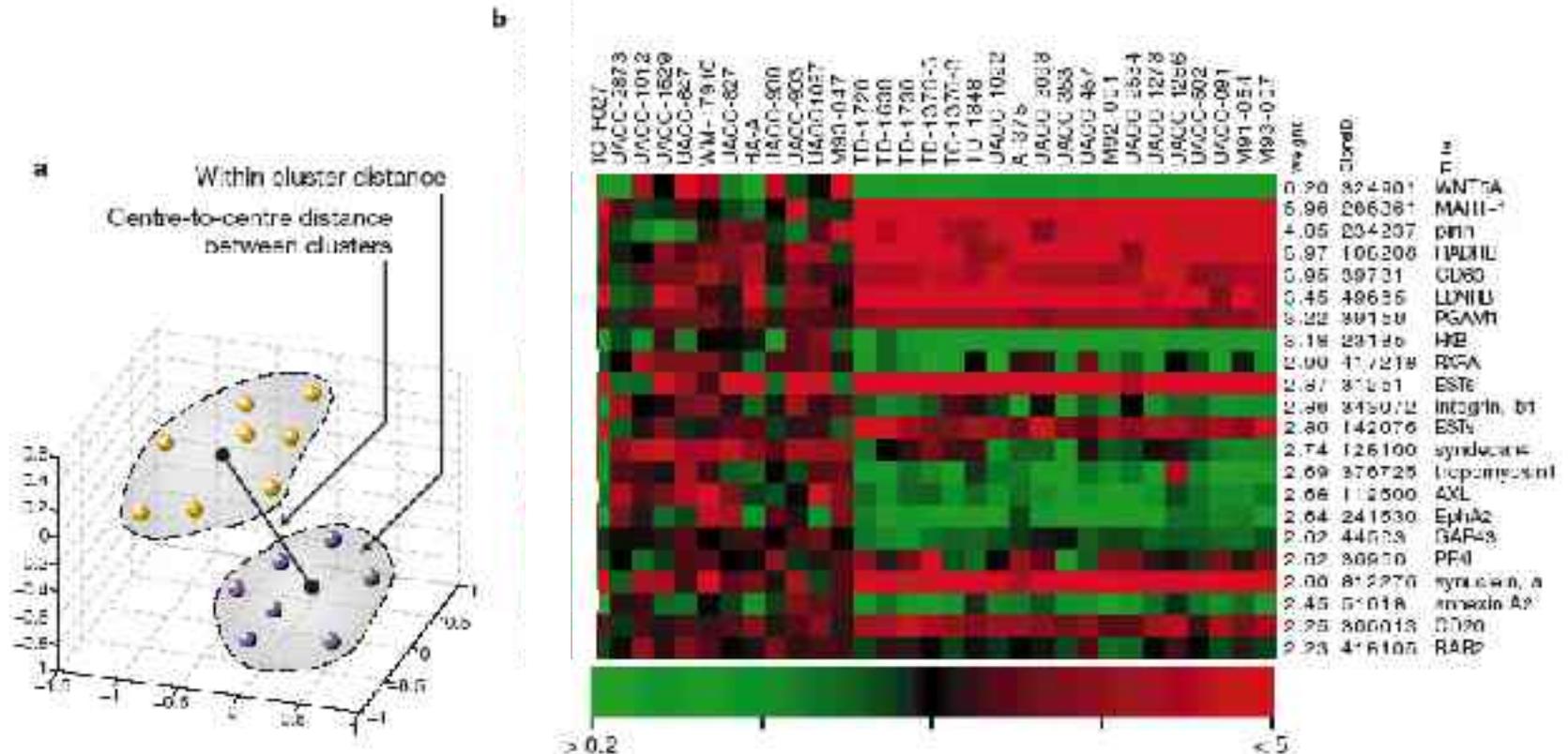
б. Трехмерный график значений MDS для 31 меланомы. Показан кластер из 19 образцов.



# Методы массового исследования транскриптов. ДНК-биочипы (микроматрицы)

## АНАЛИЗ ЭКСПРЕССИИ ГЕНОВ

Молекулярная классификация злокачественных опухолей.



Выявление генов, экспрессия которых позволяет различать кластеры меланом  
 а. MDS-анализ распределения генов по их вкладу в минимизацию объема кластера и максимизацию межкластерной дистанции; б. Экспрессия 22 самых заметных генов в кластерах.

Bittner M. *et al.*, Molecular classification of cutaneous malignant melanoma by gene expression profiling.

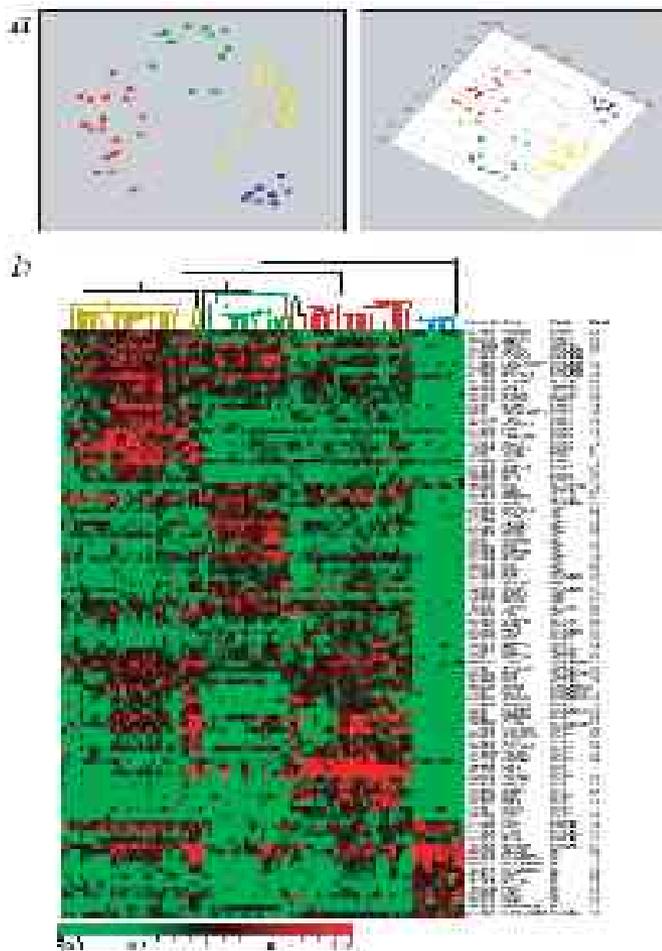
Nature. 2000 406(6795):536-540.



# Методы массового исследования транскриптов. ДНК-биочипы (микроматрицы)

## АНАЛИЗ ЭКСПРЕССИИ ГЕНОВ

Молекулярная классификация злокачественных опухолей.



Таксономия нейробластом (NB), рабдомиосарком (RMS), лимфом Ходжкина (NHL) и опухолей семейства Эуинга (EWS) на основе данных профилирования экспрессии генов с помощью кДНК-биочипа (3789 кДНК и 2778 EST).

Анализ экспрессии генов в образцах:  
b. две проекции графика значений MDS,  
c. иерархическая кластеризация данных по экспрессии генов, выявление характеристических генов.



## Исследование качественных и количественных характеристик транскриптома



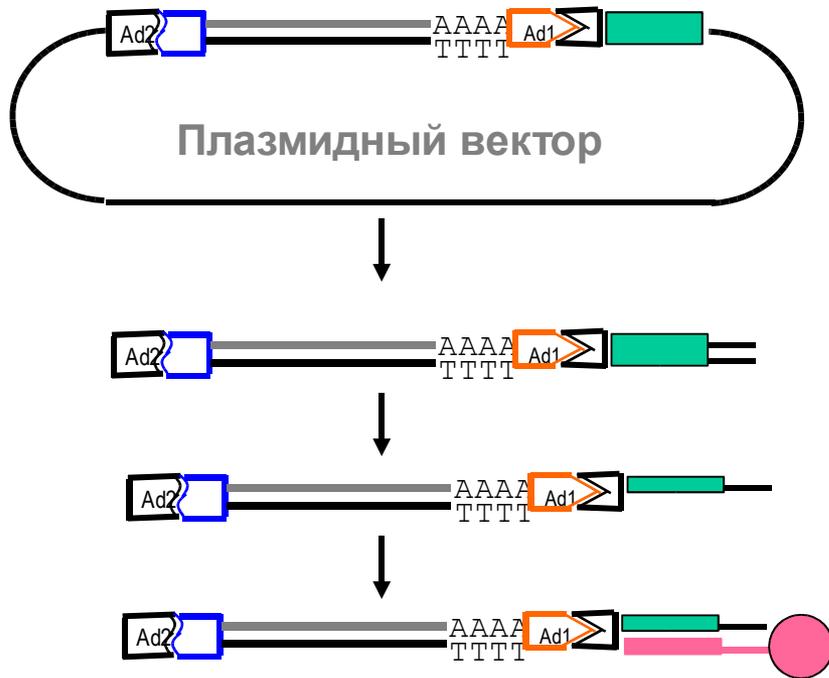
### Массовое одновременное секвенирование характерных фрагментов Massively Parallel Signature Sequencing (MPSS)

MPSS-технология позволяет идентифицировать почти все ДНК-фрагменты в образце.

- Продуцирует количественный профиль экспрессии генов на основе подсчета всех мРНК в образце, и эти цифровые данные удобны для создания реляционных баз данных
- Высокая чувствительность, способность детектировать редкие транскрипты. В каждый MPSS-анализ вовлечены более миллиона транскриптов, что обеспечивает исключительный динамический диапазон: от менее чем 10 транскриптов на миллион до 50,000 транскриптов на миллион.
- Не нужны предварительные молекулярно-генетические знания об организме

# Исследование качественных и количественных характеристик транскриптома

## Массовое одновременное секвенирование характерных фрагментов Massively Parallel Signature Sequencing (MPSS)



1. Приготовление кДНК из мРНК и клонирование в вектор, содержащий набор из  $1.67 \times 10^7$  разных 32-мерных олигонуклеотидных меток. Пул кДНК матриц, представляющих  $3-4 \times 10^4$  разных транскриптов, после клонирования формирует пул  $5-7 \times 10^{11}$  конъюгатов. Далее берется образец, включающий только 1% ( $\sim 1.6 \times 10^5$ ) от общего числа меток.

Каждый транскрипт, таким образом, конъюгирован с уникальной меткой. Даже единичный транскрипт представлен с вероятностью более 99%.

Образец из конъюгатов амплифицируется в ПЦР, метки делают одноцепочечными. Конъюгаты гибридизуются с популяцией микрошариков, к которым присоединены все комплименты меток. 1% микрошариков будет связан с конъюгатами. Только эти будут отсортированы с помощью FACS (Fluorescence-Activated Cell Sorter).



# Исследование качественных и количественных характеристик транскриптома

## Массовое одновременное секвенирование характерных фрагментов Massively Parallel Signature Sequencing (MPSS)

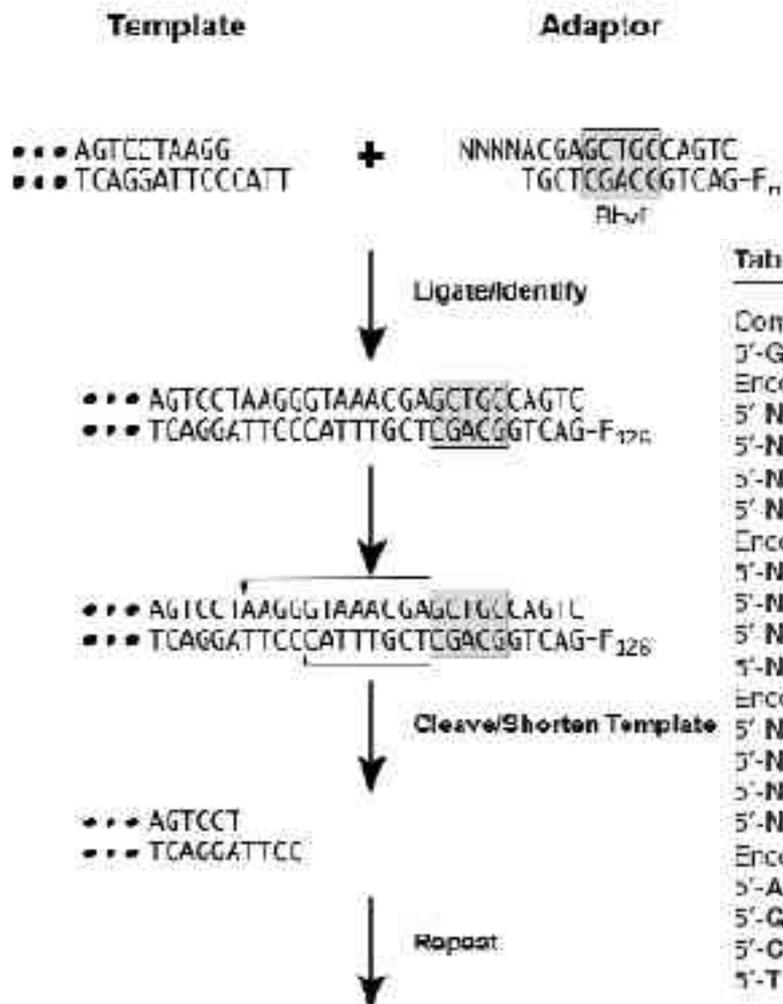


Table 1. Sequences of encoded adaptors<sup>a</sup>

Common strand:	
3'	GACTGGCAGCTG
Encoded adaptors for detecting base 1	
5'	NNNAACGAGCTGCCAGTCGaatfaggog
5'	NNNGAAGAGCTGCCAGT Gctgartaocq
5'	NNNCAGGAGCTGCCAGT Caccaataocq
5'	NNNTACGAGCTGCCAGT Cgottttag
Encoded adaptors for detecting base 2	
5'	NNANACGAGCTGCCAGT Cgggacttggg
5'	NNGNACGAGCTGCCAGT Ctggtgigtat
5'	NNCNACGAGCTGCCAGT Caccggcattc
5'	NNTNACGAGCTGCCAGT Cgallclclanc
Encoded adaptors for detecting base 3	
5'	NANNACGAGCTGCCAGT Cggsagagagto
5'	NGNACGAGCTGCCAGT Ctggtctctct
5'	NCNACGAGCTGCCAGT Ctagggactt
5'	NTNACGAGCTGCCAGT Cggogataact
Encoded adaptors for detecting base 4	
5'	ANNACGAGCTGCCAGT Cgcatccact
5'	GNNACGAGCTGCCAGT Ccaactcgtca
5'	CNNACGAGCTGCCAGT Ccaccapcaaca
5'	TNACGAGCTGCCAGT Cgtrajtjtle

- Смесь адаптеров, включающих все возможные четырехбуквенные выступы, отождествляется к целевой последовательности так, чтобы лигировался только тот, который имеет совершенный комплементарный выступ.

- Каждый из этих 256 адаптеров имеет уникальный мотив, F<sub>n</sub>, который может быть обнаружен после лигирования.

- Следующий цикл начинается с расщепления BbvI, чтобы высвободить следующие четыре основания матрицы

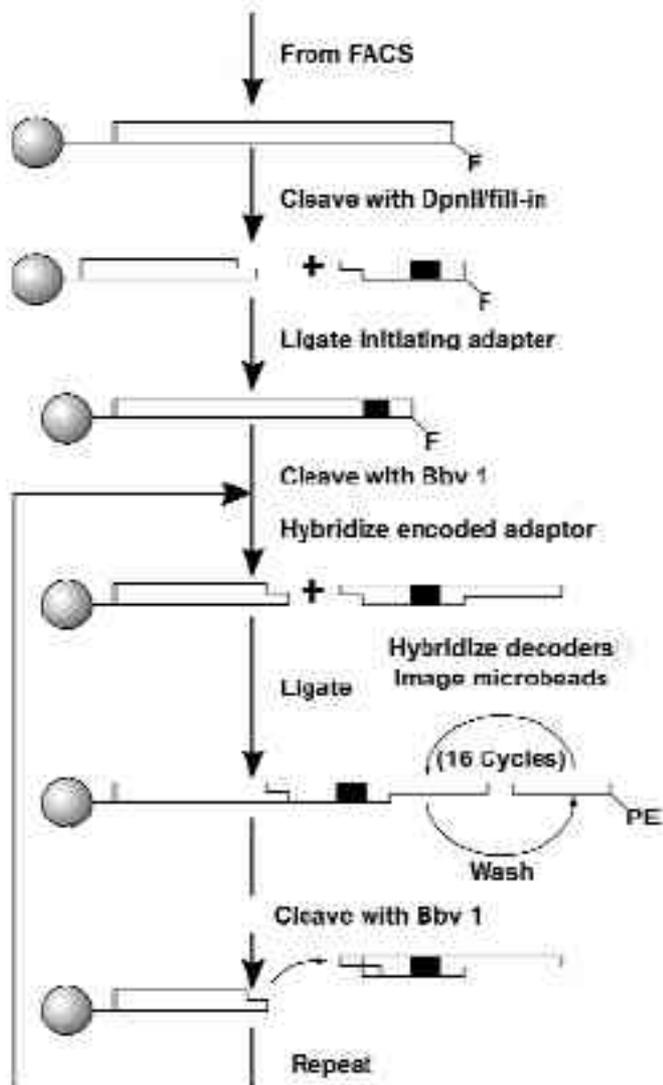
Brenner S, Johnson M, Bridgham J, et al., Gene expression analysis by massively parallel signature sequencing (MPSS) on microbead arrays. Nat Biotechnol. 2000; 18(6):630-634.



# Исследование качественных и количественных характеристик транскриптома

## Массовое одновременное секвенирование характерных фрагментов

### Massively Parallel Signature Sequencing (MPSS)



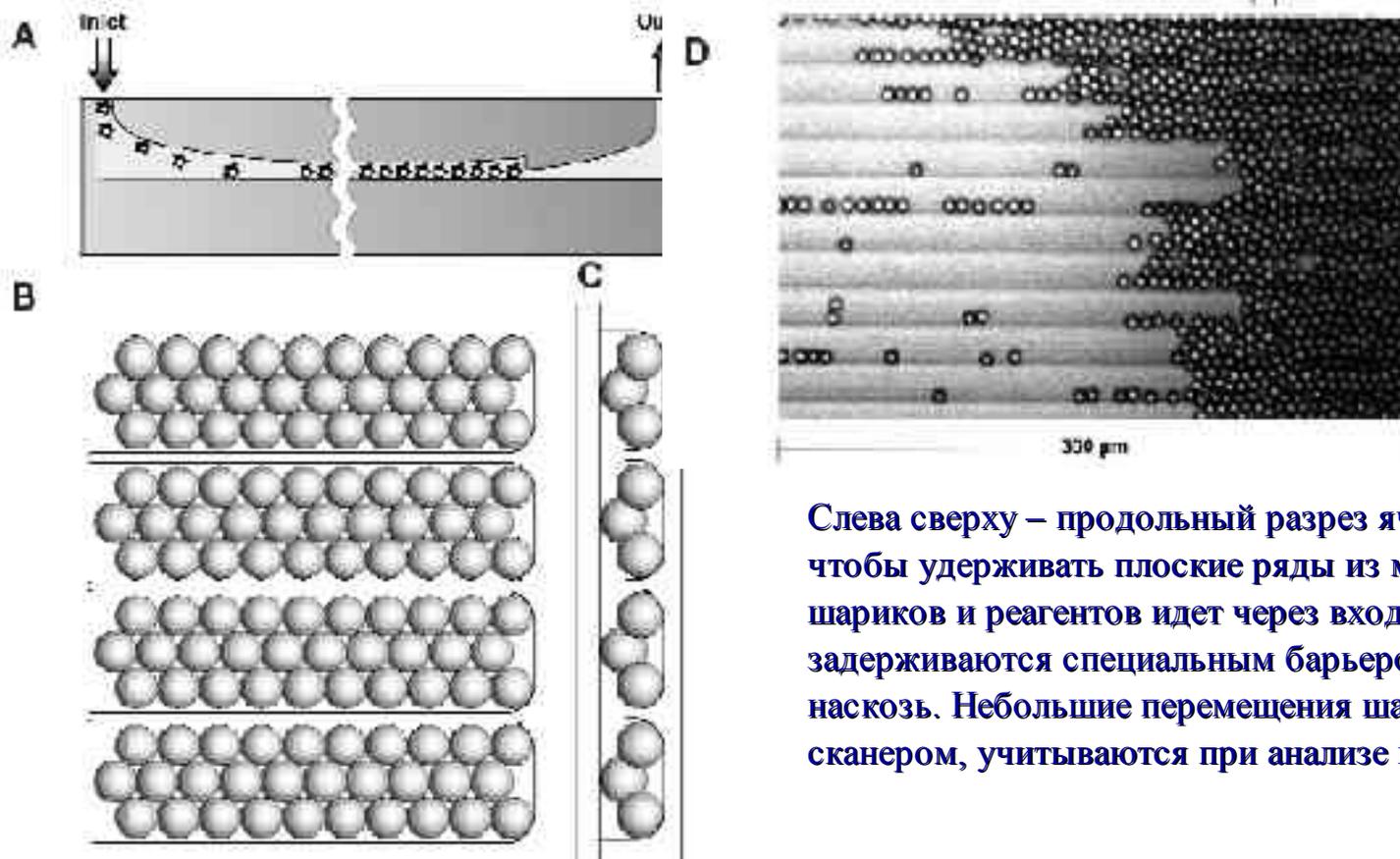
- После того как микрошарики, загруженные флуоресцентно помеченными (F) кДНК, изолированы с помощью FACS, кДНК разрезаются рестриктазой DpnII, чтобы высвободить четыре основания в виде выступа. Этот выступ достраивается до выступа с тремя основаниями.
- Флуоресцентно помеченный (F) иницирующие адаптеры, содержащие BbvI-сайты, лигируются к кДНК в отдельной реакции, после чего микрошарики загружаются в специальные ячейки.
- Затем кДНК разрезаются BbvI, кодирующие адаптеры, прогибризовавшиеся с ними, лигируются.
- Шестнадцать помеченных фикоэритрином (PE) зонда по отдельности гибридизуются с декодирующими мотивами кодирующих адаптеров. После каждой гибридизации ячейка сканируется, изображение рядов микрошариков анализируется для идентификации оснований выступов.
- Кодирующие адаптеры снова обрабатываются BbvI, который расщепляет кДНК и высвобождает четыре новых основания для следующего цикла лигирования и расщепления.

Brenner S, Johnson M, Bridgham J, et al., Gene expression analysis by massively parallel signature sequencing (MPSS) on microbead arrays. Nat Biotechnol. 2000; 18(6):630-634.



# Исследование качественных и количественных характеристик транскриптома

## Массовое одновременное секвенирование характерных фрагментов Massively Parallel Signature Sequencing (MPSS)



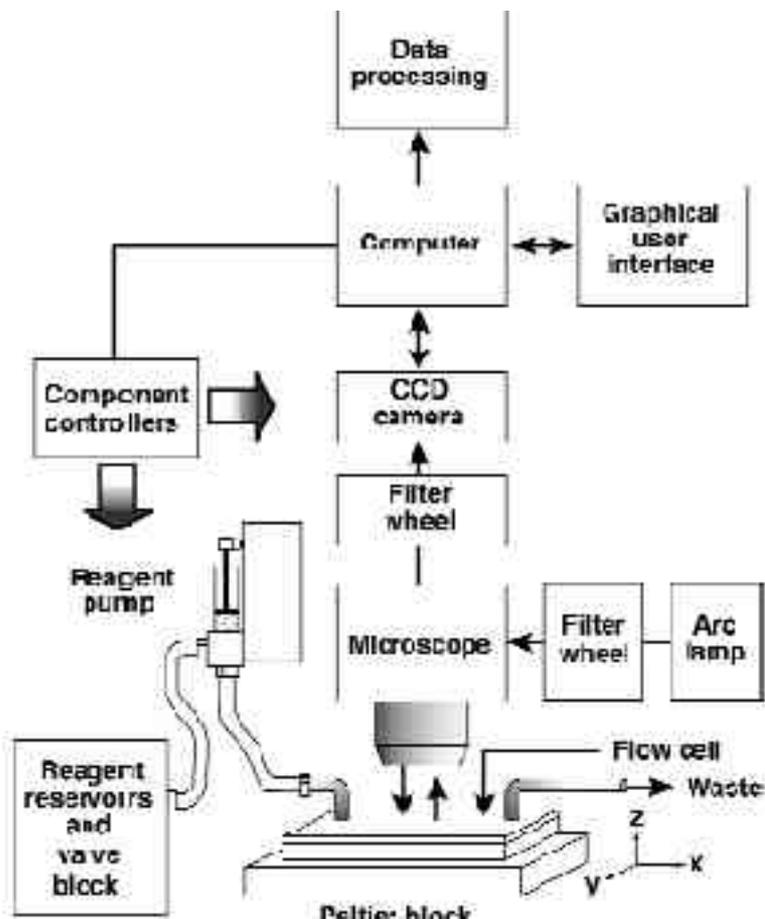
Слева сверху – продольный разрез ячейки. Она сделана так, чтобы удерживать плоские ряды из микрошариков. Загрузка шариков и реагентов идет через входное отверстие. Шарики задерживаются специальным барьером, а растворы проходят насквозь. Небольшие перемещения шариков, зафиксированные сканером, учитываются при анализе изображений.

Brenner S, Johnson M, Bridgham J, et al., Gene expression analysis by massively parallel signature sequencing (MPSS) on microbead arrays. Nat Biotechnol. 2000; 18(6):630-634.



# Исследование качественных и количественных характеристик транскриптома

## Массовое одновременное секвенирование характерных фрагментов Massively Parallel Signature Sequencing (MPSS)



Проточная ячейка монтируется на терморегулируемом блоке Пельтье в конфокальном микроскопе, снабженном ксеноновой лампой, фильтрами и CCD-камерой.

Для сканирования изображения ряды микрошариков разделены на 18 секций с 62,000 шариков в каждой. Специальное программное обеспечение обрабатывает изображения, выделяя центр каждого микрошарика и суммируя флуоресцентный сигнал с определенных пикселей изображения.

Как правило, удается получить надежные данные о последовательности 16-20 нуклеотидов кДНК-матрицы. Далее начинается сказываться «шум» из-за ошибок рестрикции и лигирования.

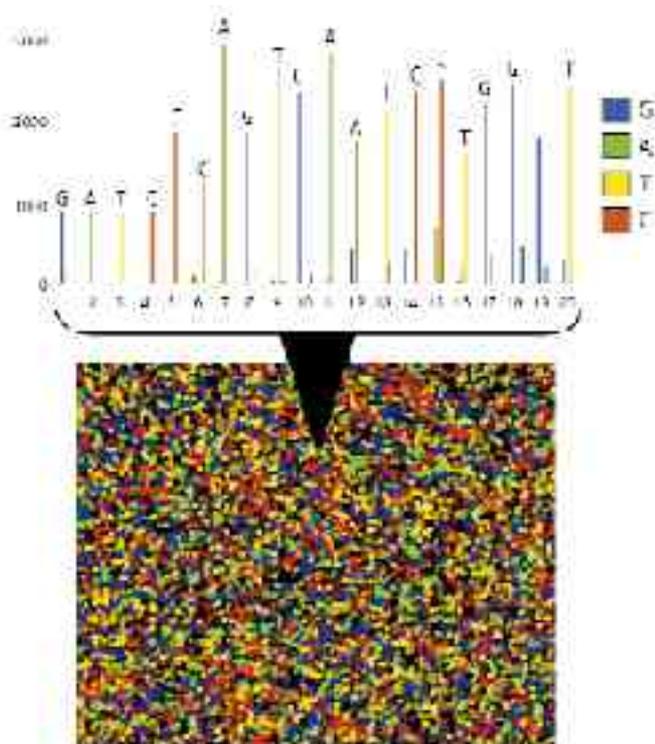
Полученные 16-20-нуклеотидные мотивы потом преобразуются в сигнатуры с помощью специальных программ.



# Исследование качественных и количественных характеристик транскриптома



## Массовое одновременное секвенирование характерных фрагментов Massively Parallel Signature Sequencing (MPSS)

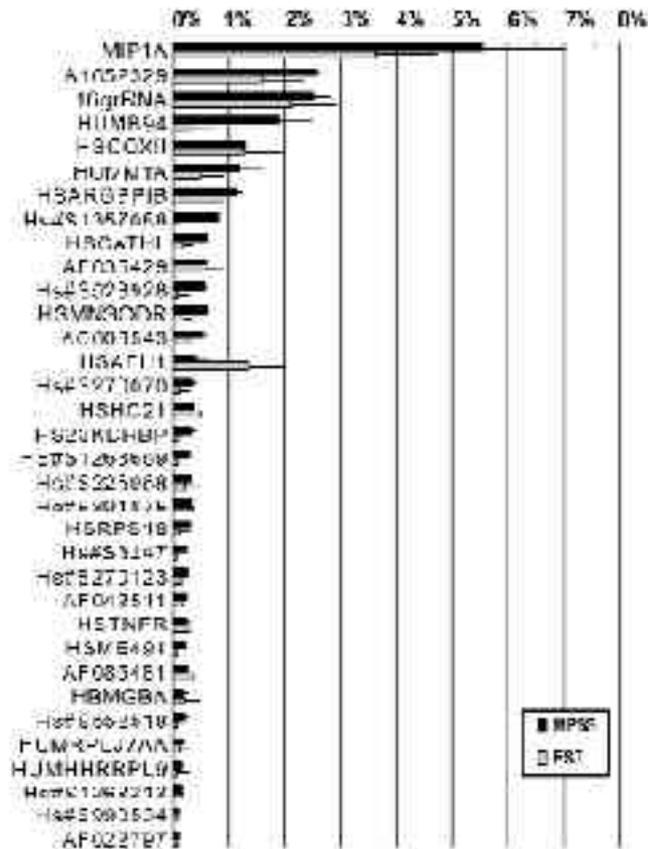


Так выглядит псевдоцветное компьютерное изображение некоторого количества миктошариков и изображения гистограмм значений флуоресцентного сигнала и расшифрованной последовательности некоего микрошарика.



# Исследование качественных и количественных характеристик транскриптома

## Массовое одновременное секвенирование характерных фрагментов Massively Parallel Signature Sequencing (MPSS)



Comparison of MPSS analysis with expressed sequence tag (EST) sequencing. A total of 1,839 randomly picked clones from a cDNA library derived from induced THP-1 cells were sequenced (PE Biosystems, Model 377 DNA Sequencer). The sequences were clustered and searched against GenBank with BLAST. The MPSS signatures corresponding to the EST clusters were then tabulated from the MPSS database of signatures from the same cDNA library. Percentage total for MPSS data is the average abundance measured, with actual measured error shown. Percentage total for the EST data is the number of sequences clustered out of the 1,839 selected clones, with the 99% confidence interval used for error. A 1% relative abundance corresponds to about 2,500 microbead signatures for MPSS data and to about 18 sequences for the EST data..

Brenner S, Johnson M, Bridgham J, et al., Gene expression analysis by massively parallel signature sequencing (MPSS) on microbead arrays. Nat Biotechnol. 2000; 18(6):630-634.



# Исследование качественных и количественных характеристик транскриптома



## АНАЛИЗ ЭКСПРЕССИИ ГЕНОВ: ОТ-ПЦР с детекцией в реальном времени

- Позволяет следить за приращением количества специфического продукта в каждом цикле реакции.
- Не требует выделения амплифицированных фрагментов НК.
- Позволяет определять относительное и абсолютное количество последовательности-мишени.
- Обладает высокой чувствительностью и расширенным динамическим диапазоном.



# Исследование качественных и количественных характеристик транскриптома

АНАЛИЗ ЭКСПРЕССИИ ГЕНОВ: ОТ-ПЦР с детекцией в реальном времени

real-time PCR hardware

**iCycler**  
BioRad

**LightCycler**  
Roche

**5700**  
Applied Biosystems

**7700**  
Applied Biosystems

**FluorTracker**  
Stratagene

**FluorImager**  
Molecular Dynamics



# Исследование качественных и количественных характеристик транскриптома



АНАЛИЗ ЭКСПРЕССИИ ГЕНОВ: ОТ-ПЦР с детекцией в реальном времени

Типы real-time PCR: источники флуоресценции.

- интеркалирующие флуоресцентные красители – чаще всего SYBR Green 1;
- флуоресцентные пробы/зонды – TaqMan и Molecular Beacons;
- флуоресцентные затравки ( primers) с элементом зонда – Scorpions.



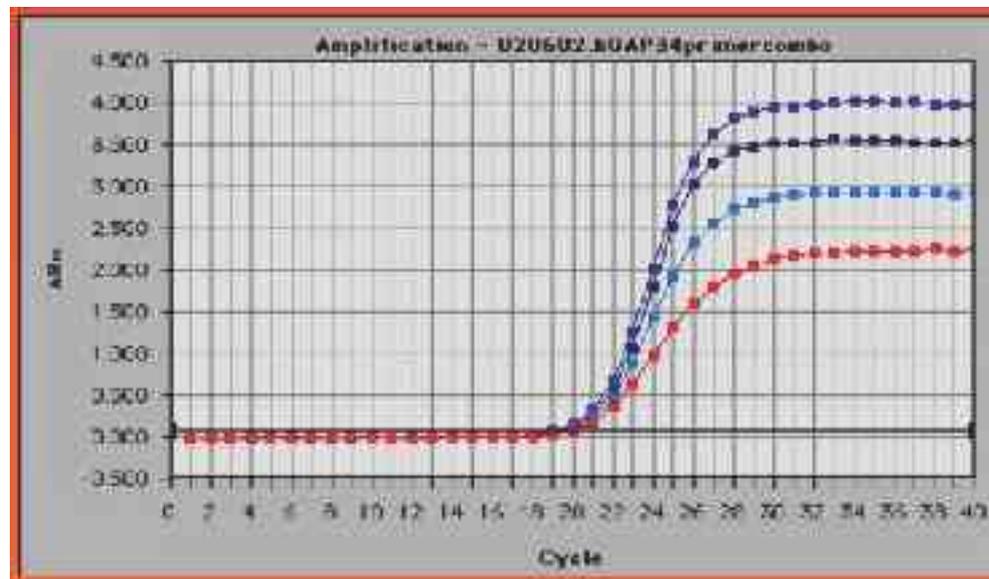
# Исследование качественных и количественных характеристик транскрипта



АНАЛИЗ ЭКСПРЕССИИ ГЕНОВ: ОТ-ПЦР с детекцией в реальном времени



ТaqMan система

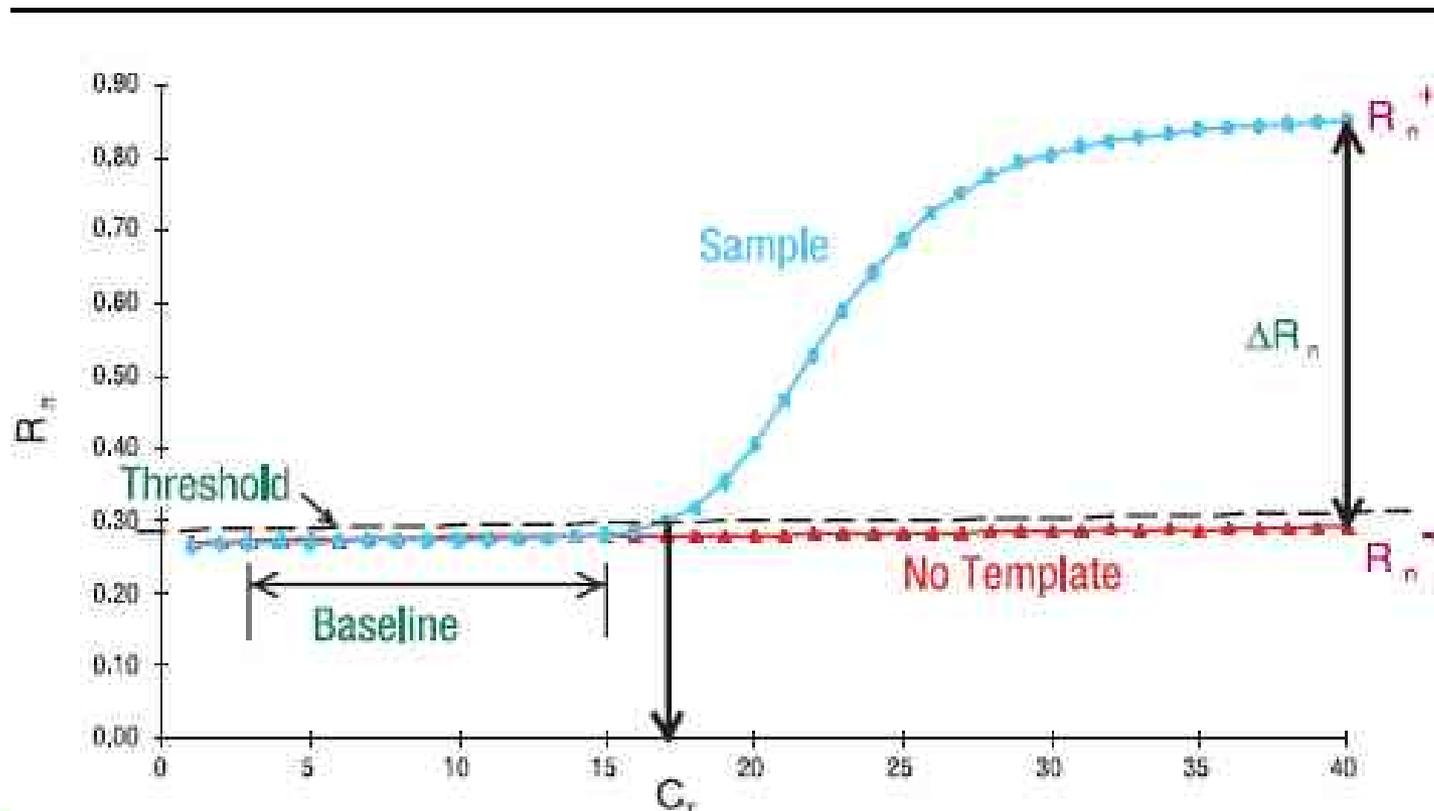




# Исследование качественных и количественных характеристик транскриптома

АНАЛИЗ ЭКСПРЕССИИ ГЕНОВ: ОТ-ПЦР с детекцией в реальном времени

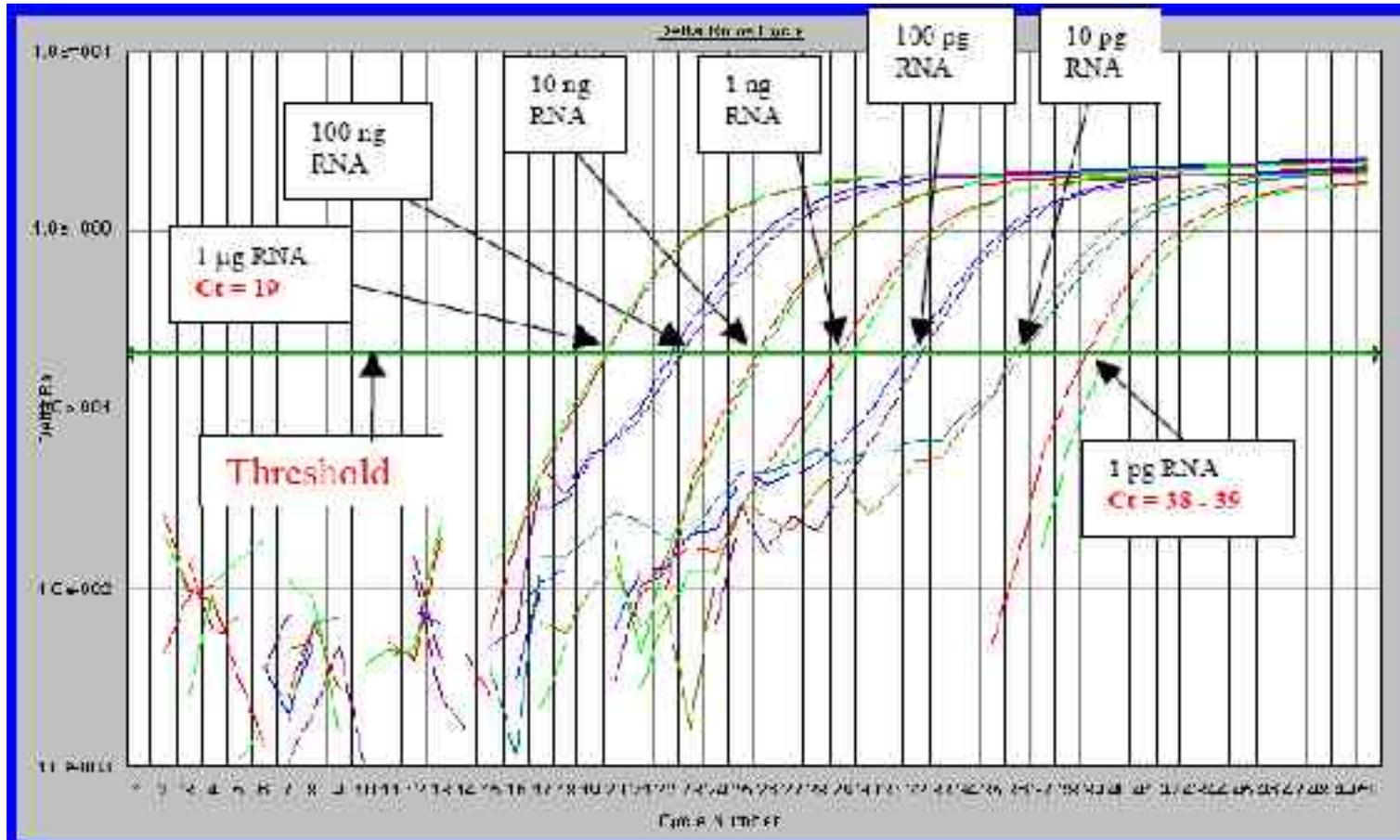
## TaqMan система





# Исследование качественных и количественных характеристик транскрипта

## АНАЛИЗ ЭКСПРЕССИИ ГЕНОВ: ОТ-ПЦР с детекцией в реальном времени TaqMan система



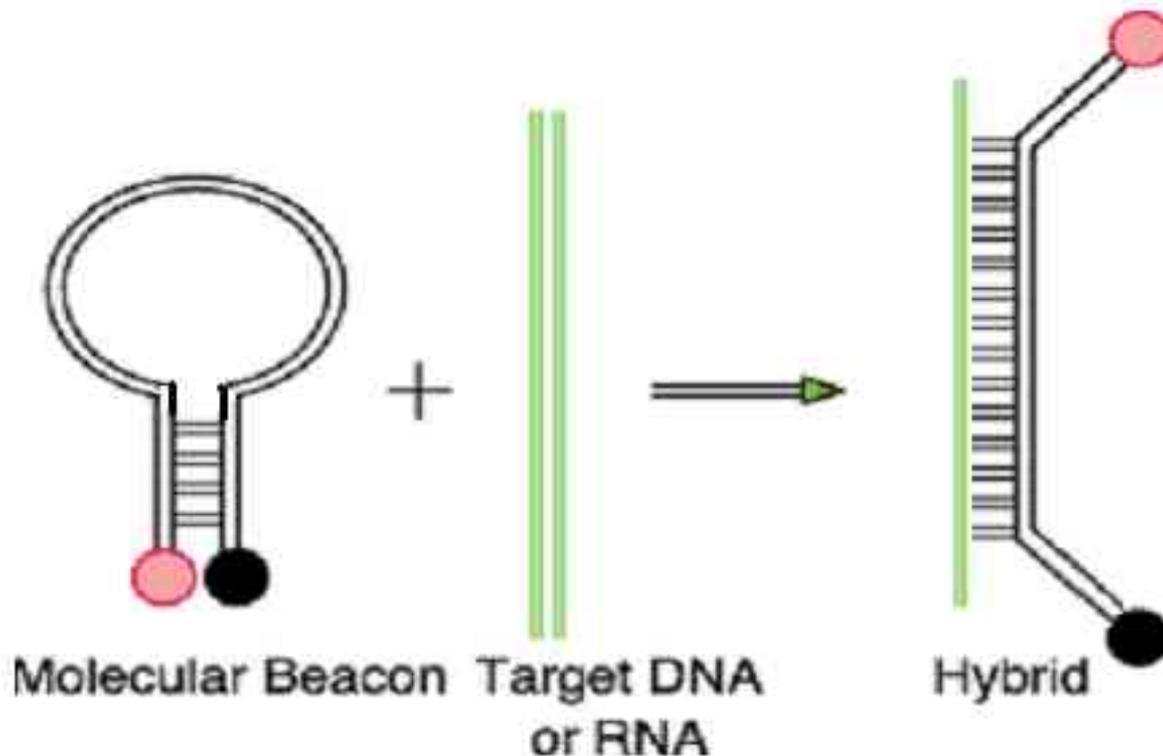


# Исследование качественных и количественных характеристик транскрипта



АНАЛИЗ ЭКСПРЕССИИ ГЕНОВ: ОТ-ПЦР с детекцией в реальном времени

## Molecular Beacons



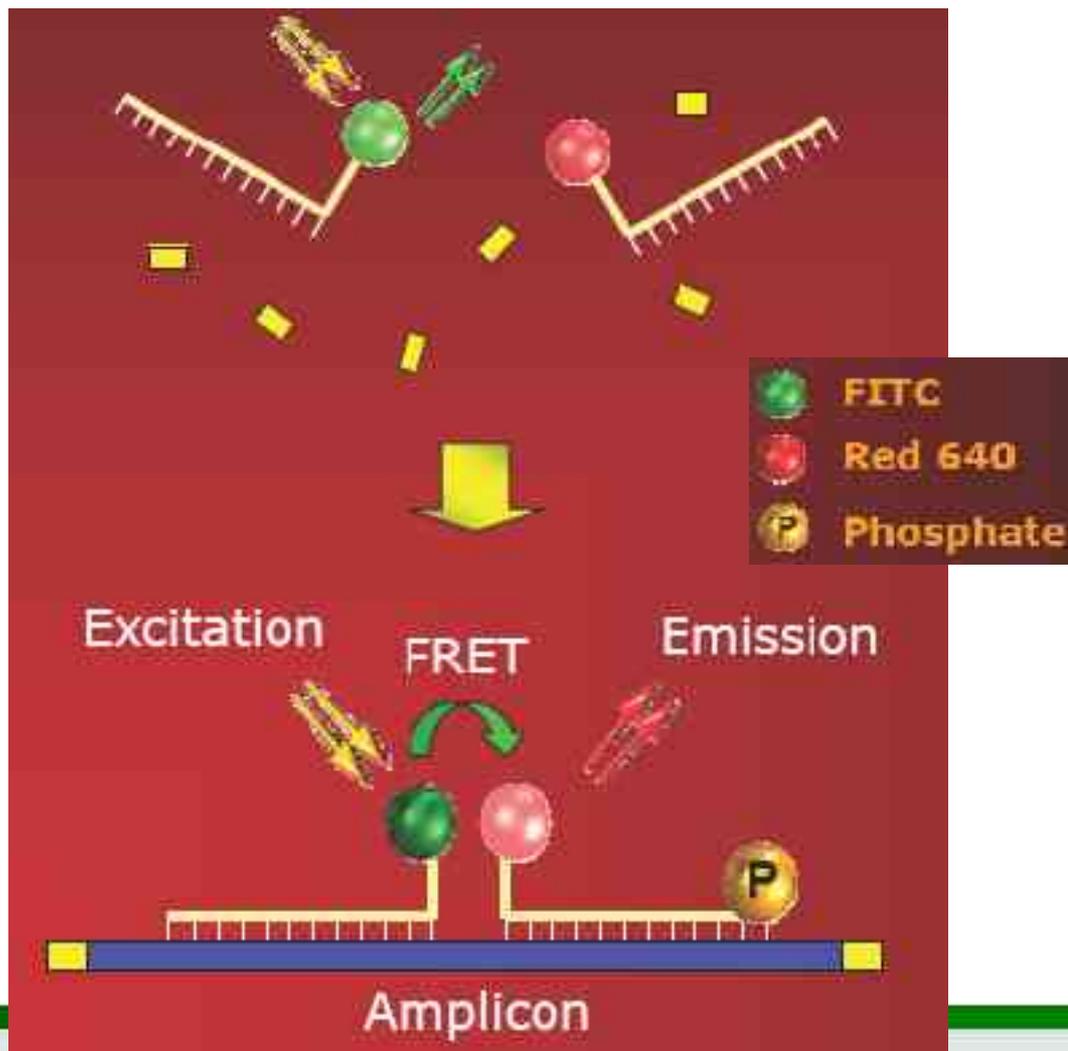


# Исследование качественных и количественных характеристик транскрипта



АНАЛИЗ ЭКСПРЕССИИ ГЕНОВ: ОТ-ПЦР с детекцией в реальном времени

LightCycler – FRET  
dual probes



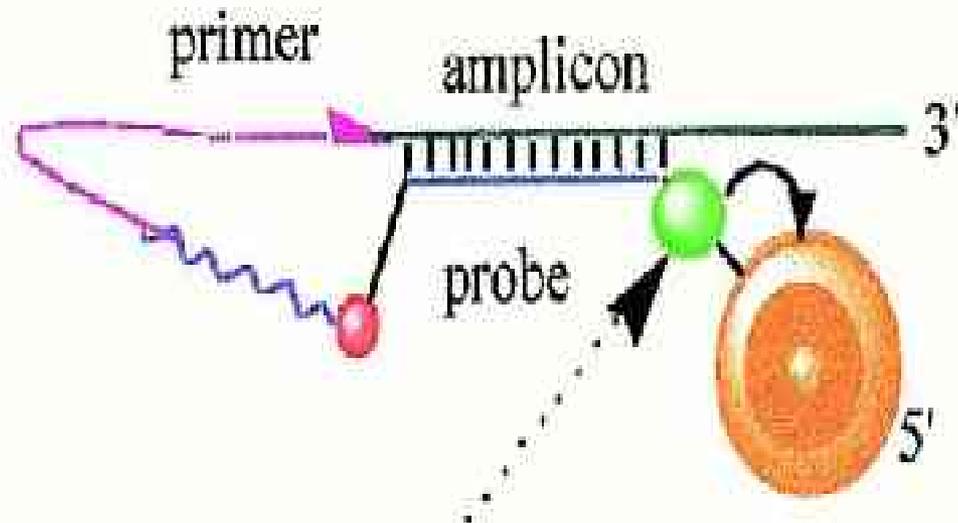
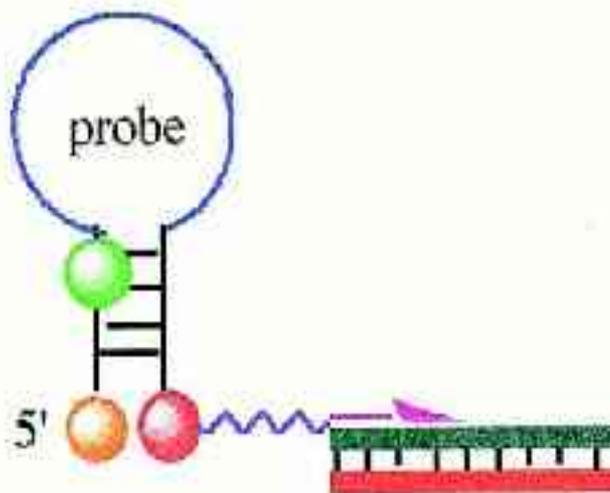


# Исследование качественных и количественных характеристик транскрипта



АНАЛИЗ ЭКСПРЕССИИ ГЕНОВ: ОТ-ПЦР с детекцией в реальном времени

## Scorpions

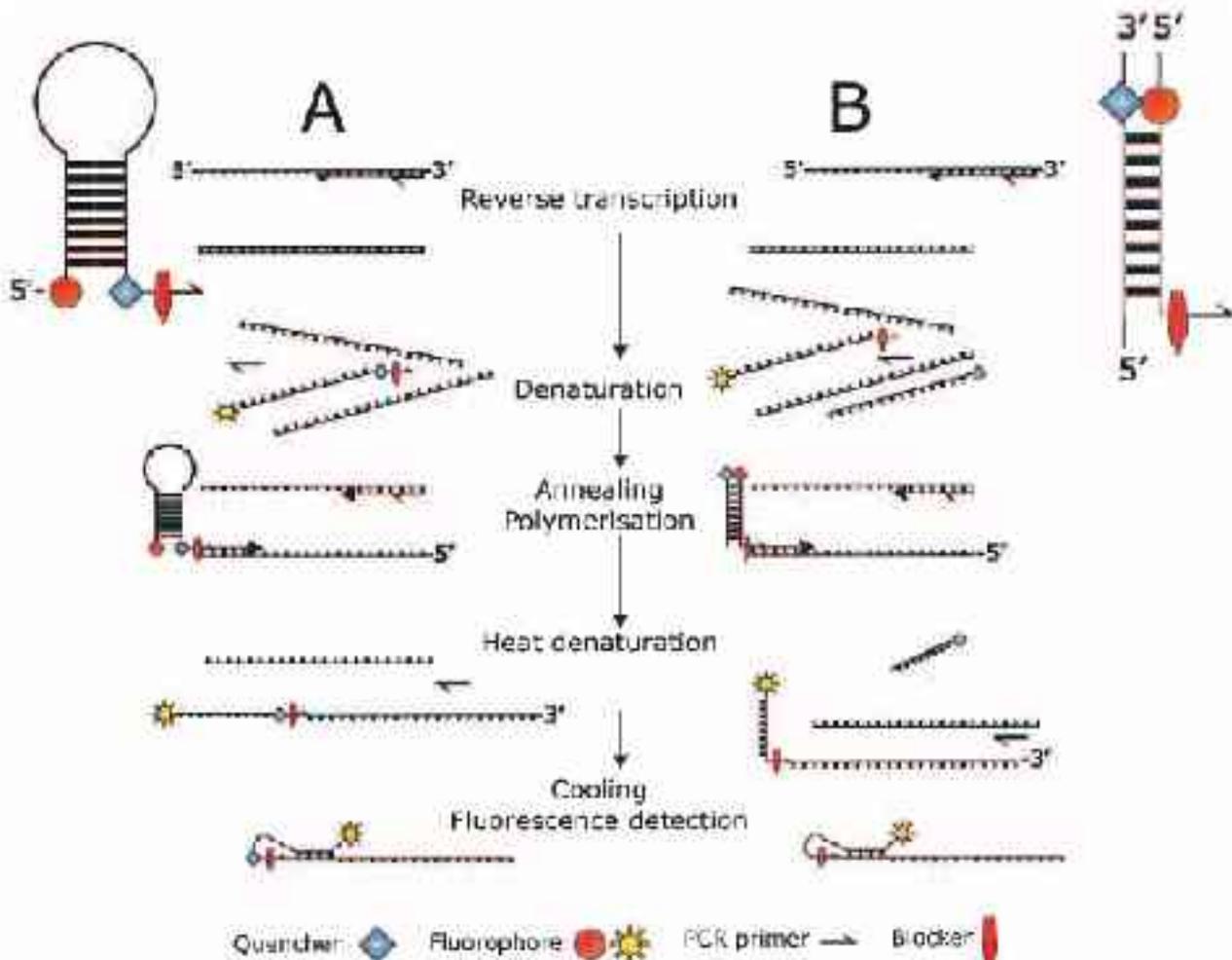




# Исследование качественных и количественных характеристик транскрипта



## АНАЛИЗ ЭКСПРЕССИИ ГЕНОВ: ОТ-ПЦР с детекцией в реальном времени



Scorpions



# Исследование качественных и количественных характеристик транскриптома

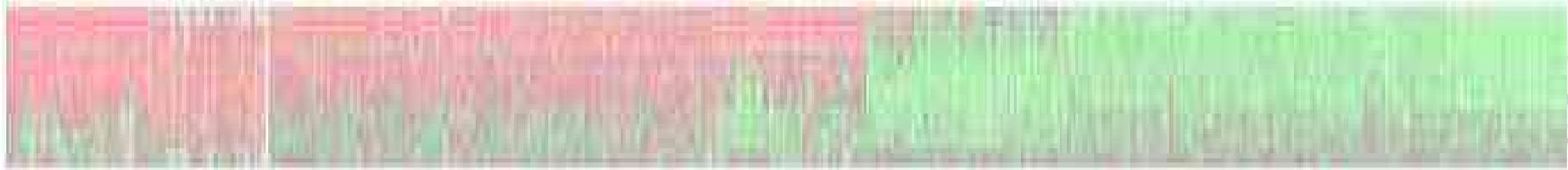


## Характеристика методов детекции и оценки количества

транскриптом метод	Доверие относит. и абсолют., сравнимость	количество генов	чувствительность и динамический диапазон	производительность и эффективность
Прямая детекция мРНК гибридизацией транскриптов с одноцеп. ДНК- или РНК-пробами				
Нозерн-блот	+	(1-5) x (5- 20)	--	--
гибридизация Защита от рибонуклеазы	++	(10-15) x (5- 20)	+	--
Детекция кДНК, полученной ОТ-ПЦР из мРНК				
Количественная ОТ-ПЦР, ОТ-ПЦР с дет. в реал. времени	++	десятки	++ (!)	++
Дифференц. дисплей	+	десятки	++ (!)	++
Олиго/кДНК микрочипы	++ (!)	тысячи	+++	+++
Компьютерный анализ сиквенсов кДНК , « <i>in silico</i> гибридизация» транскриптов				
SAGE	+++	тысячи	+++	+++
MPSS	++++	тысячи	++++	++++
EST	+	тысячи	+++ (!)	+++



# Применение ДНК-биочипов (микроматриц) для исследования экспрессии генов



Спасибо за внимание.

