Gene expression analysis

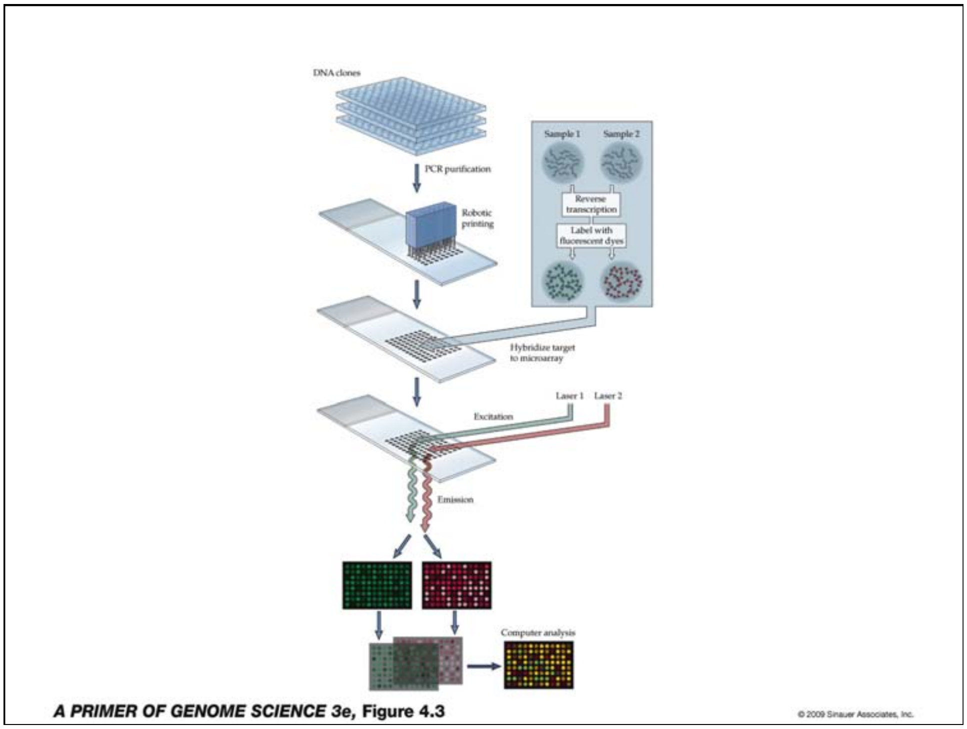
Biosciences 741: Genomics

Fall, 2013

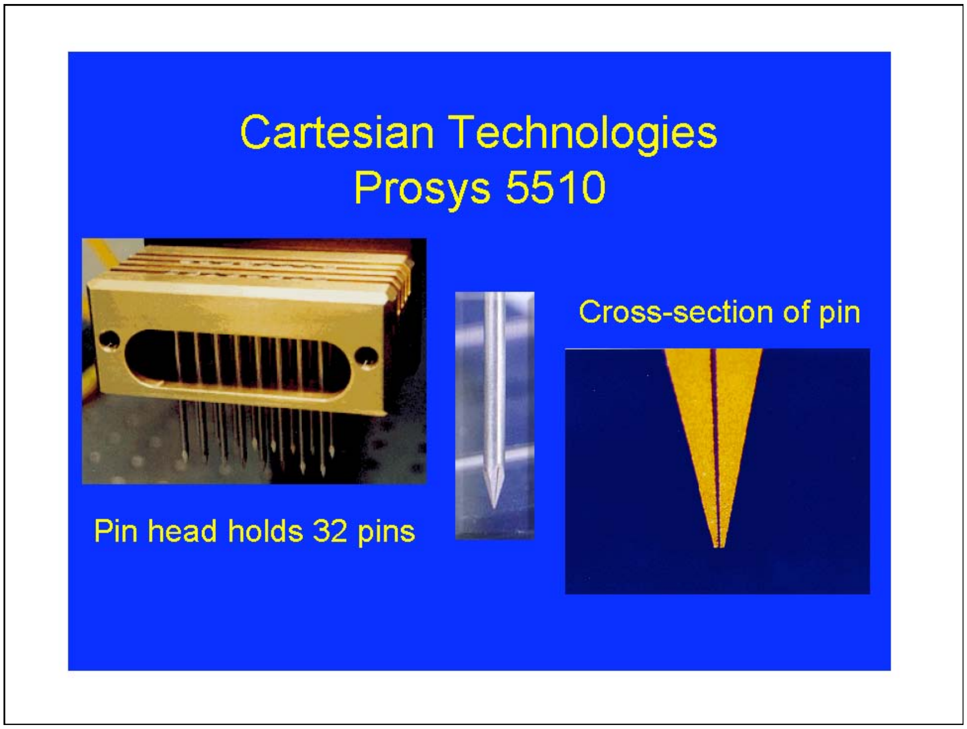
Week 5

Gene expression analysis

- From EST clusters to spotted cDNA microarrays
- Long vs. short oligonucleotide microarrays vs. RT-PCR
- Methods of DNA microarray data analysis
- Serial analysis of gene expression (SAGE) and RNA-seq
- Promoter analysis
- Chromatin immunoprecipitation (ChIP-seq)



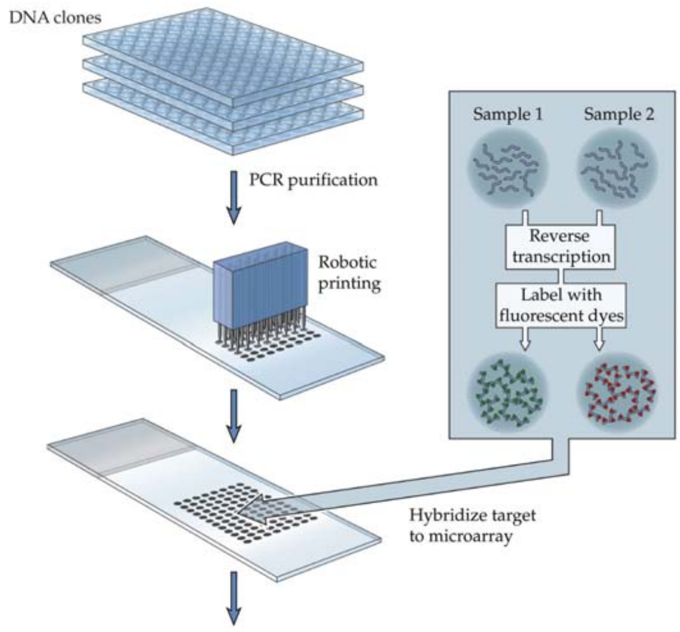
A PRIMER OF GENOME SCIENCE 3e, **Figure 4.3**



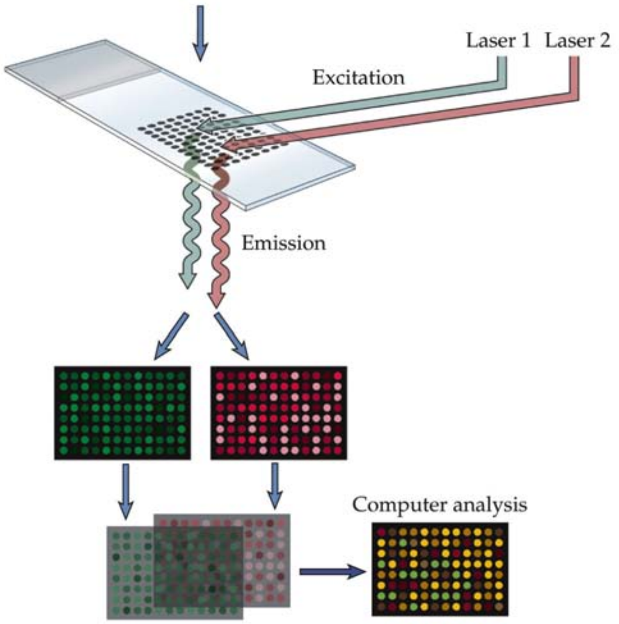
Cartesian Technologies Prosys 5510

Pin head holds 32 pins

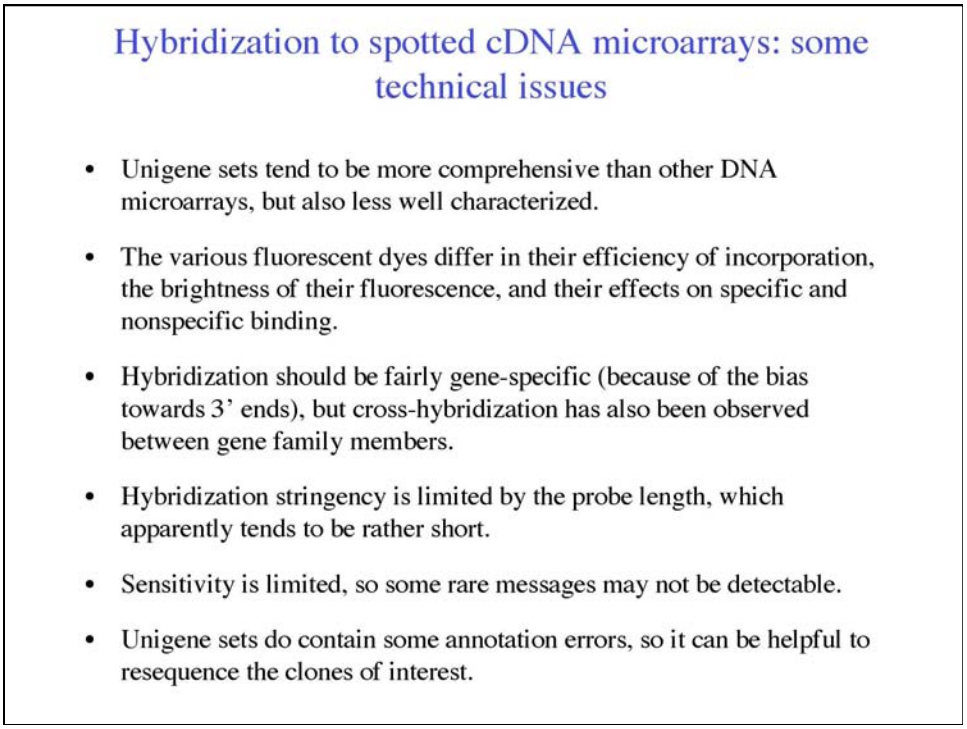
Cross-section of pin



***A PRIMER OF GENOME SCIENCE 3e*, Figure 4.3 (Part 1)**

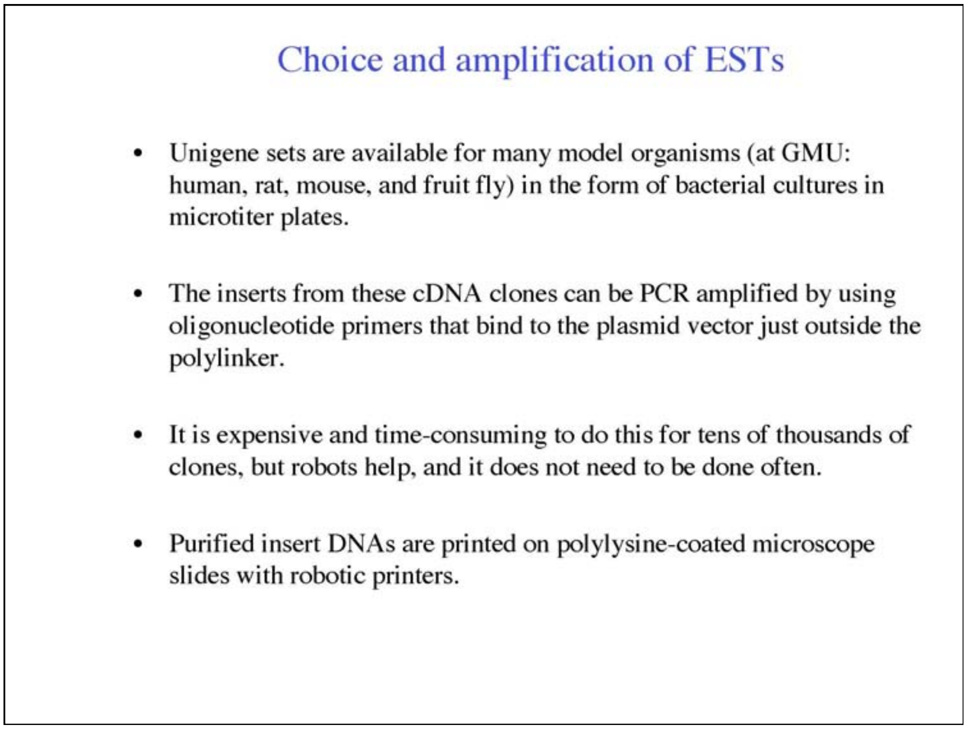


***A PRIMER OF GENOME SCIENCE 3e*, Figure 4.3 (Part 2)**



Hybridization to spotted cDNA microarrays: some technical issues

- Unigene sets tend to be more comprehensive than other DNA microarrays, but also less well characterized.
- The various fluorescent dyes differ in their efficiency of incorporation, the brightness of their fluorescence, and their effects on specific and nonspecific binding.
- Hybridization should be fairly gene-specific (because of the bias towards 3' ends), but cross-hybridization has also been observed between gene family members.
- Hybridization stringency is limited by the probe length, which apparently tends to be rather short.
- Sensitivity is limited, so some rare messages may not be detectable.
- Unigene sets do contain some annotation errors, so it can be helpful to resequence the clones of interest.

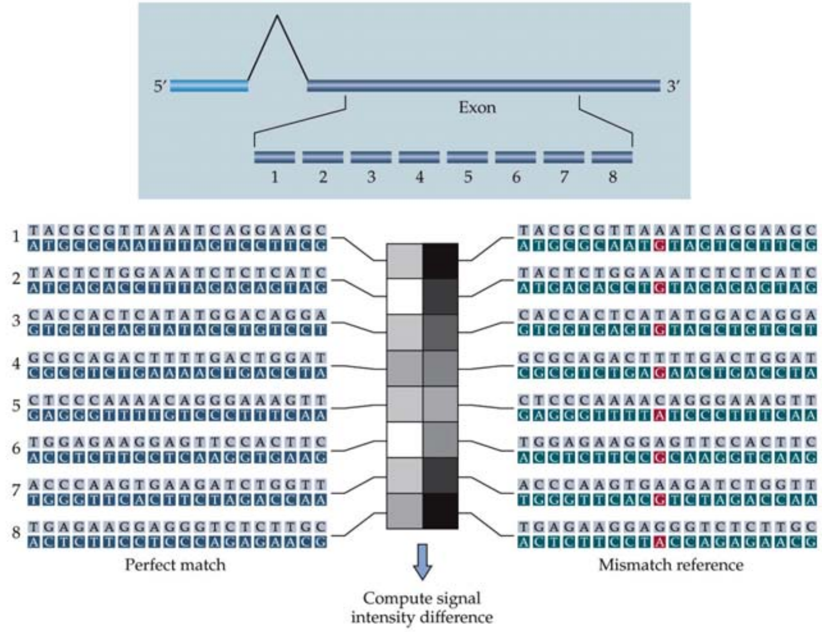


Choice and amplification of ESTs

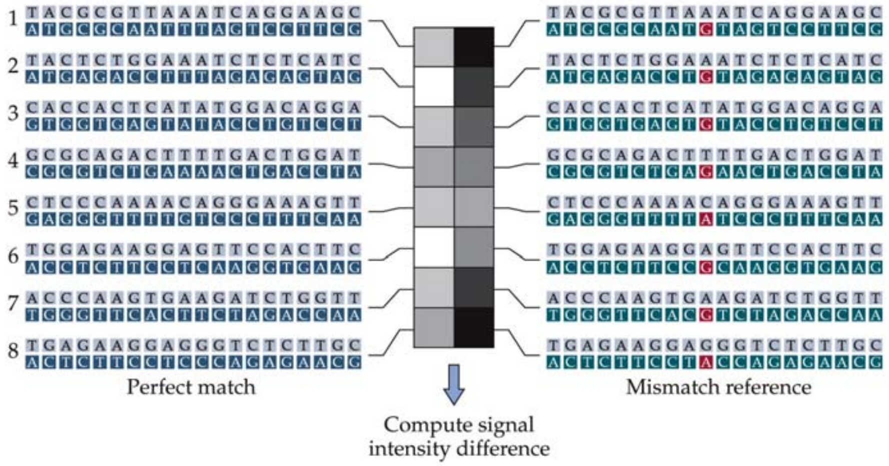
- Unigene sets are available for many model organisms (at GMU: human, rat, mouse, and fruit fly) in the form of bacterial cultures in microtiter plates.
- The inserts from these cDNA clones can be PCR amplified by using oligonucleotide primers that bind to the plasmid vector just outside the polylinker.
- It is expensive and time-consuming to do this for tens of thousands of clones, but robots help, and it does not need to be done often.
- Purified insert DNAs are printed on polylysine-coated microscope slides with robotic printers.

Gene expression analysis

- From EST clusters to spotted cDNA microarrays
- Long vs. short oligonucleotide microarrays vs. RT-PCR
- Methods of DNA microarray data analysis
- Serial analysis of gene expression (SAGE) and RNA-seq
- Promoter analysis
- Chromatin immunoprecipitation (ChIP-seq)



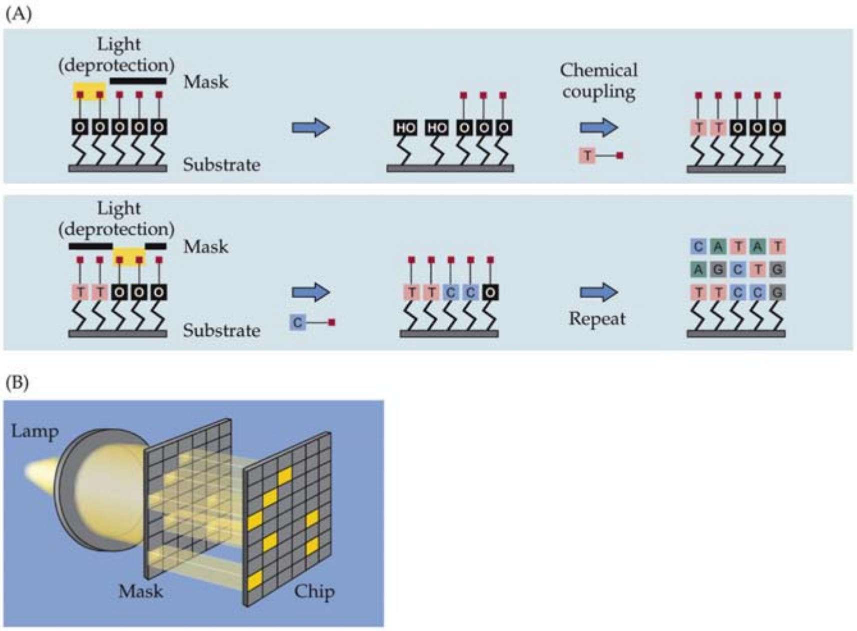
***A PRIMER OF GENOME SCIENCE* 3e, Figure 4.5**



***A PRIMER OF GENOME SCIENCE 3e*, Figure 4.5 (Part 2)**

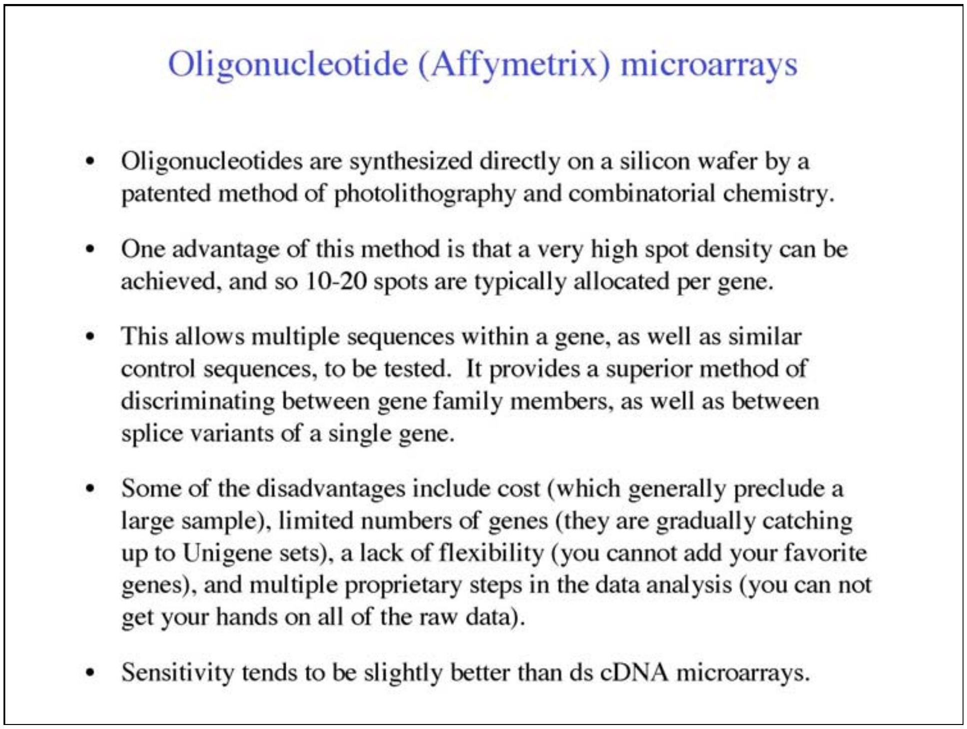
© 2009 Sinauer Associates, Inc.

Figure 4.6 Construction of Affymetrix oligonucleotide arrays



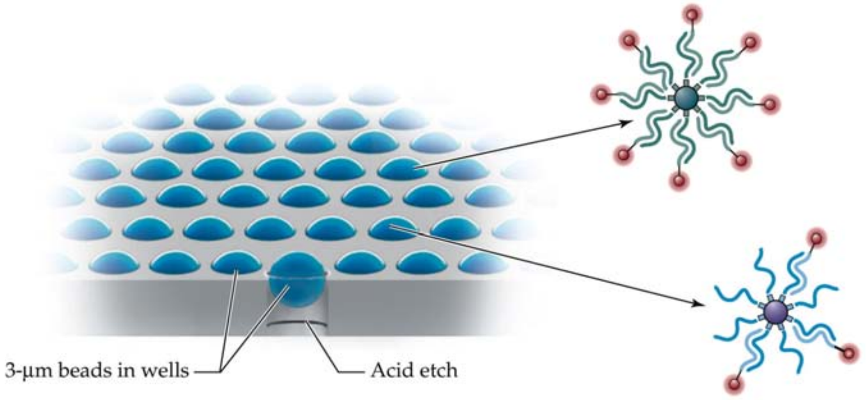
***A PRIMER OF GENOME SCIENCE 3e*, Figure 4.6**

© 2009 Sinauer Associates, Inc.



Oligonucleotide (Affymetrix) microarrays

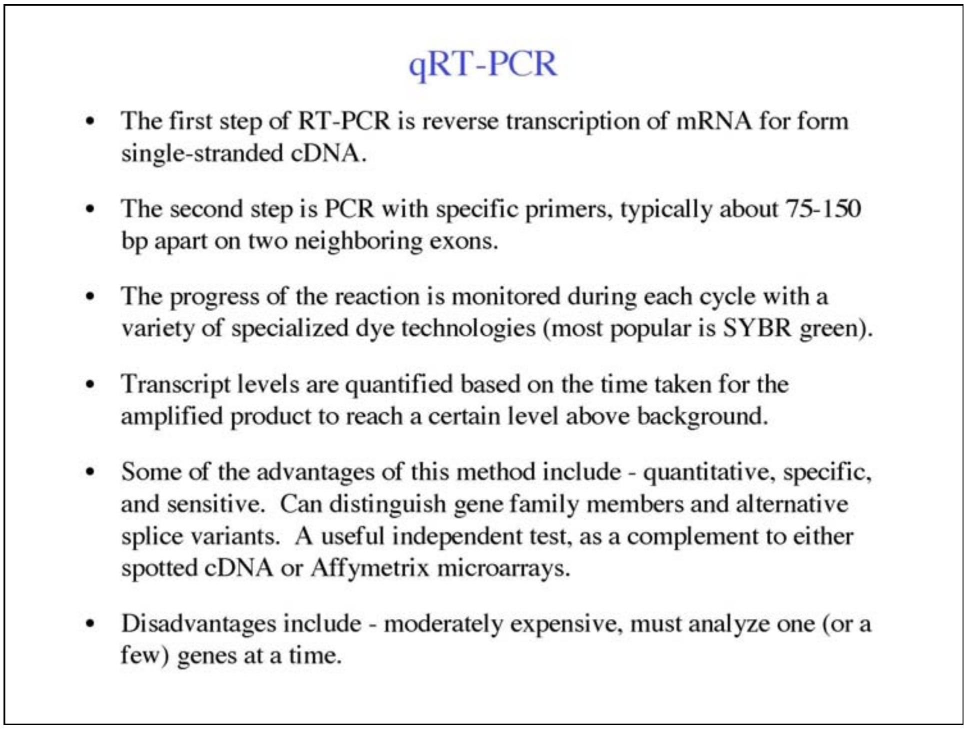
- Oligonucleotides are synthesized directly on a silicon wafer by a patented method of photolithography and combinatorial chemistry.
- One advantage of this method is that a very high spot density can be achieved, and so 10-20 spots are typically allocated per gene.
- This allows multiple sequences within a gene, as well as similar control sequences, to be tested. It provides a superior method of discriminating between gene family members, as well as between splice variants of a single gene.
- Some of the disadvantages include cost (which generally preclude a large sample), limited numbers of genes (they are gradually catching up to Unigene sets), a lack of flexibility (you cannot add your favorite genes), and multiple proprietary steps in the data analysis (you can not get your hands on all of the raw data).
- Sensitivity tends to be slightly better than ds cDNA microarrays.



A PRIMER OF GENOME SCIENCE 3e, Figure 4.4

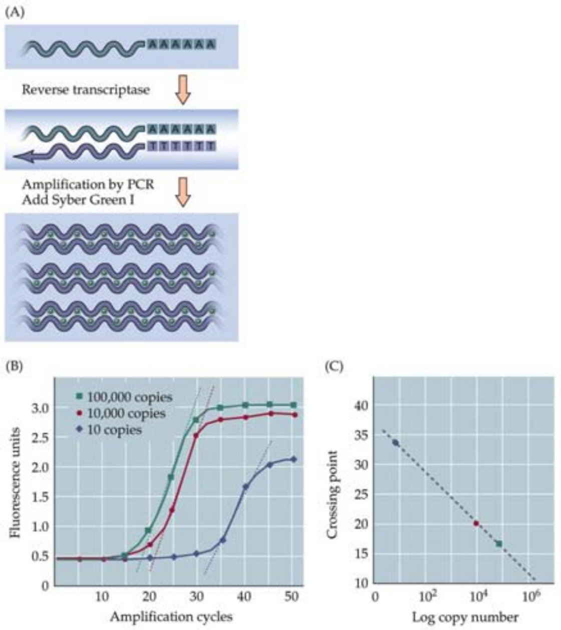
Long oligonucleotide microarrays

- Long oligonucleotide microarrays are increasingly popular. These are based on conventional methods of oligonucleotide synthesis and microarray printing (on polylysine or etc) of 60-80 base oligos, typically one per gene.
- These in many respects represent a compromise between the advantages and disadvantages of spotted cDNAs vs. proprietary short oligos.
- Long oligonucleotide microarrays have moderate cost, good sensitivity, and good reproducibility.
- Some long oligonucleotide microarrays are commercially available. It is also possible to make your own, and hence control the gene content of your assay.
- In general, the correlations between microarray platforms are low to moderate, due to a variety of factors (3' bias in probes, probe length, hybridization stringency, diffusion limits, etc).



qRT-PCR

- The first step of RT-PCR is reverse transcription of mRNA for form single-stranded cDNA.
- The second step is PCR with specific primers, typically about 75-150 bp apart on two neighboring exons.
- The progress of the reaction is monitored during each cycle with a variety of specialized dye technologies (most popular is SYBR green).
- Transcript levels are quantified based on the time taken for the amplified product to reach a certain level above background.
- Some of the advantages of this method include - quantitative, specific, and sensitive. Can distinguish gene family members and alternative splice variants. A useful independent test, as a complement to either spotted cDNA or Affymetrix microarrays.
- Disadvantages include - moderately expensive, must analyze one (or a few) genes at a time.



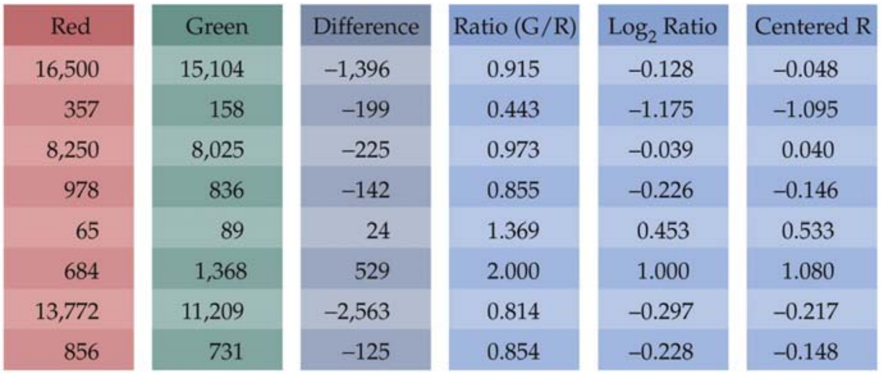
***A PRIMER OF GENOME SCIENCE 3e*, Figure 4.17**

Gene expression analysis

- From EST clusters to spotted cDNA microarrays
- Long vs. short oligonucleotide microarrays vs. RT-PCR
- Methods of DNA microarray data analysis
- Serial analysis of gene expression (SAGE) and RNA-seq
- Promoter analysis
- Chromatin immunoprecipitation (ChIP-seq)

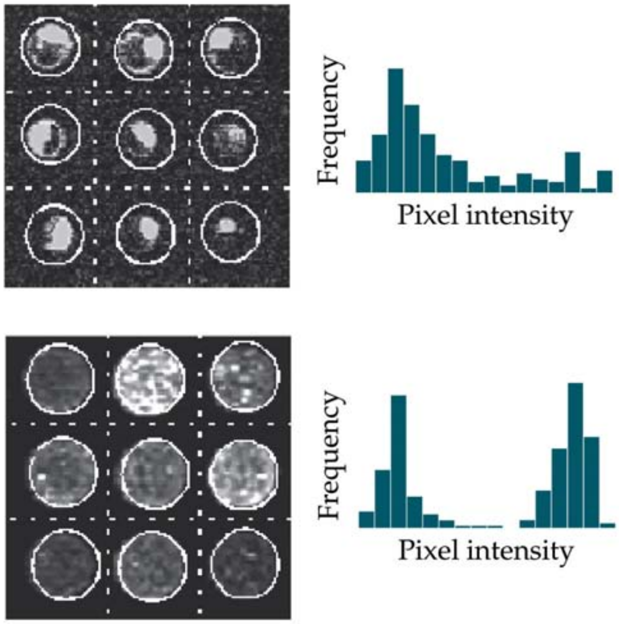
Microarrays: initial data analysis

- Affymetrix data is based on single-color hybridizations, computed as (perfect match - mismatch), which are then log-normalized. Thus, a separate slide must be purchased if you wish to use a reference sample.
- A similar method of initial data analysis can be used for spotted cDNA microarrays. This approach eliminates problems with dyes and reference samples. It should also facilitate data mining.
- Alternatively, some labs prefer competitive two-color hybridizations (red/green). This can help to control for variation between printed spots, as well as variation in hybridization conditions.
- Some of the disadvantages of two-color hybridizations are that experiments can only be compared if they used the same reference sample, genes not expressed in the reference sample can not be analyzed, and ratios cause serious statistical problems.
- LOESS normalization, sometimes used with either one- or two-color spotted cDNA arrays, can correct for variations between printer pins and/or regional variations in background.



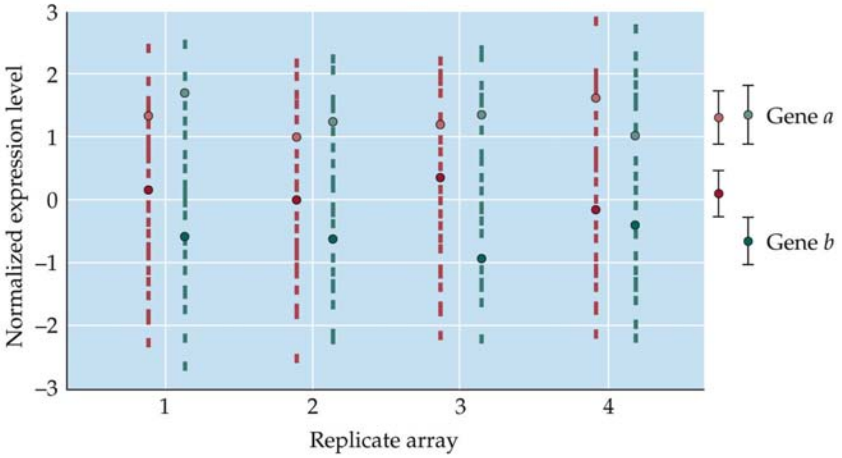
Red	Green	Difference	Ratio (G/R)	Log_2 Ratio	Centered R
16,500	15,104	–1,396	0.915	–0.128	–0.048
357	158	–199	0.443	–1.175	–1.095
8,250	8,025	–225	0.973	–0.039	0.040
978	836	–142	0.855	–0.226	–0.146
65	89	24	1.369	0.453	0.533
684	1,368	529	2.000	1.000	1.080
13,772	11,209	–2,563	0.814	–0.297	–0.217
856	731	–125	0.854	–0.228	–0.148

***A PRIMER OF GENOME SCIENCE 3e*, Figure 4.8**



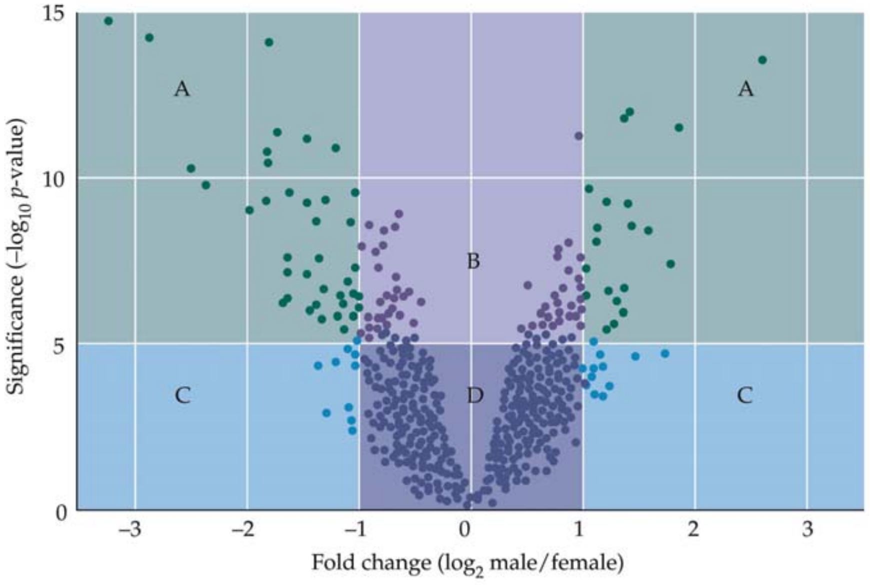
A PRIMER OF GENOME SCIENCE 3e, Box 4.1, Figure A

Figure 4.9 Analysis of variance (ANOVA) for gene expression data



A PRIMER OF GENOME SCIENCE 3e, Figure 4.9

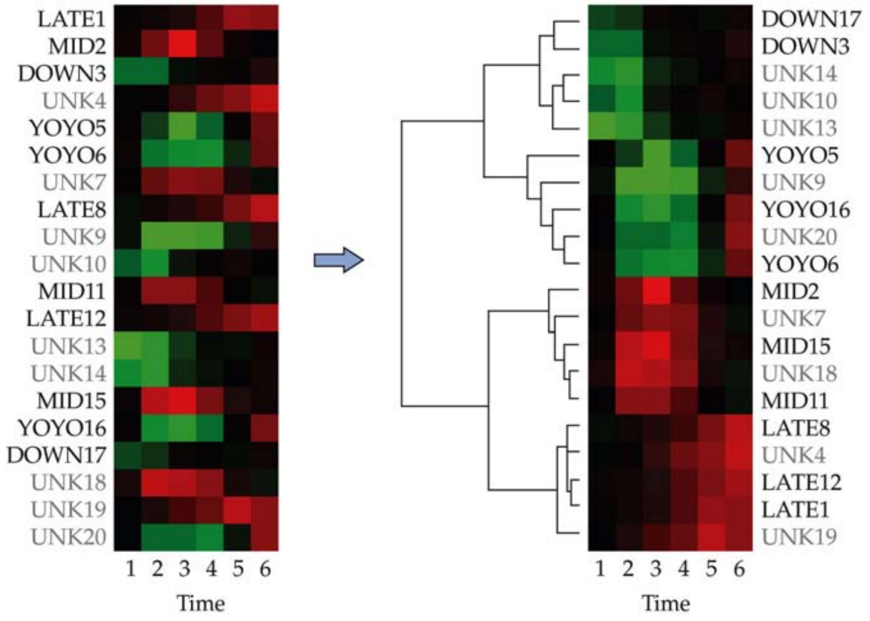
Figure 4.10 - the “volcano” plot of statistical significance versus $\log_2$ of fold-change in gene expression values.



***A PRIMER OF GENOME SCIENCE 3e*, Figure 4.10**

© 2009 Sinauer Associates, Inc.

Figure 4.11 - Hierarchical clustering of gene expression patterns.



***A PRIMER OF GENOME SCIENCE 3e*, Figure 4.11**

© 2009 Sinauer Associates, Inc.

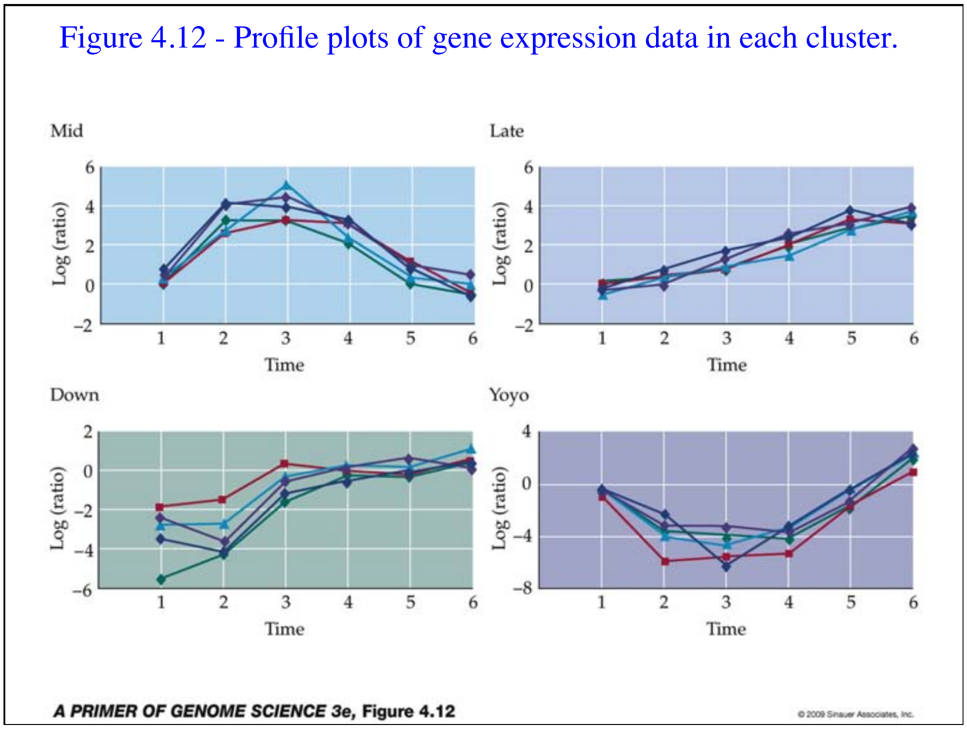
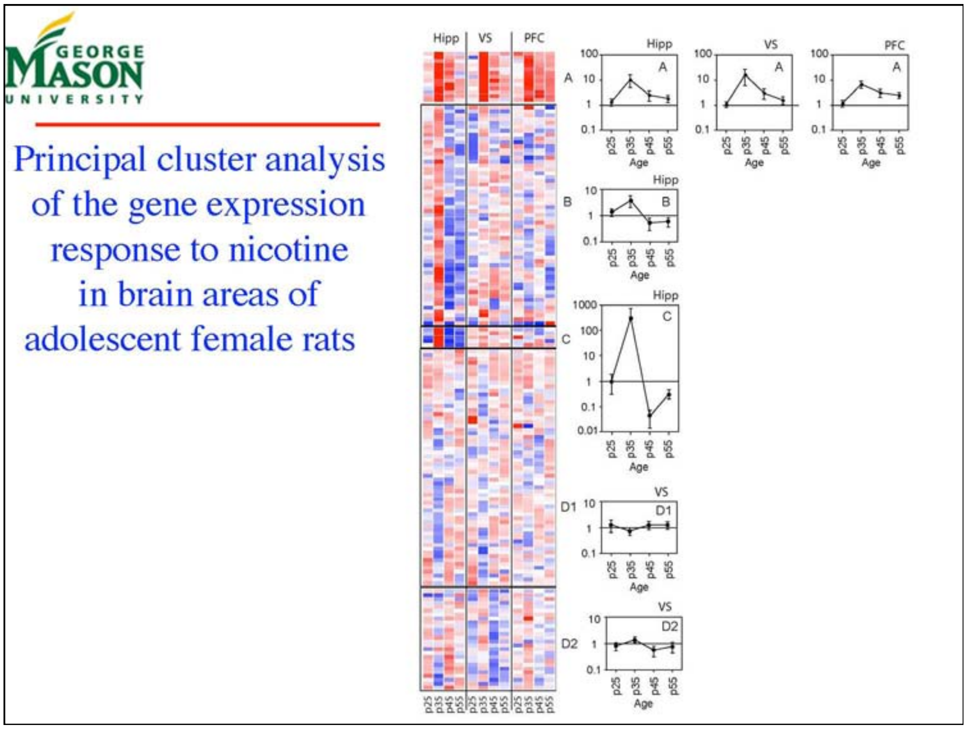


Figure 4.12 - Profile plots of gene expression data in each cluster.

A PRIMER OF GENOME SCIENCE 3e, **Figure 4.12**

© 2009 Sinauer Associates, Inc.



Principal cluster analysis of the gene expression response to nicotine in brain areas of adolescent female rats

Cluster A is specifically induced by nicotine.

Clusters B & C may be due, in part, to the delay of normal developmental profiles.

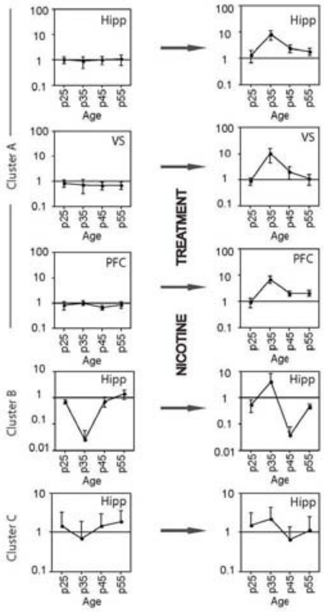
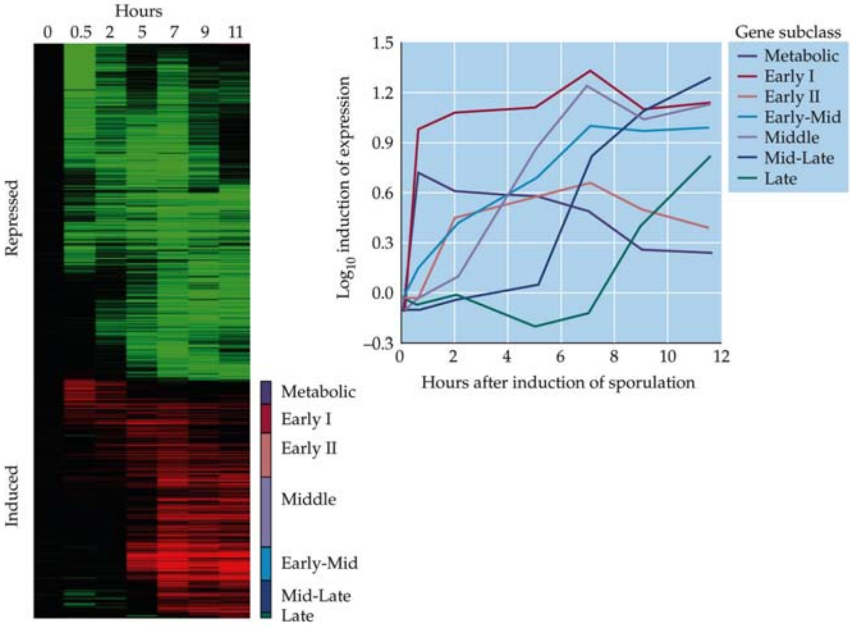


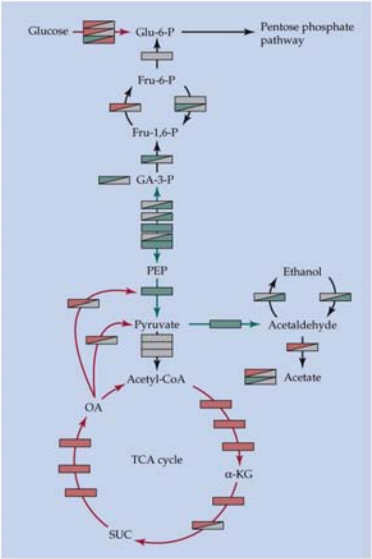
Figure 4.18 - microarray analysis of sporulation in budding yeast.



A PRIMER OF GENOME SCIENCE 3e, **Figure 4.18**

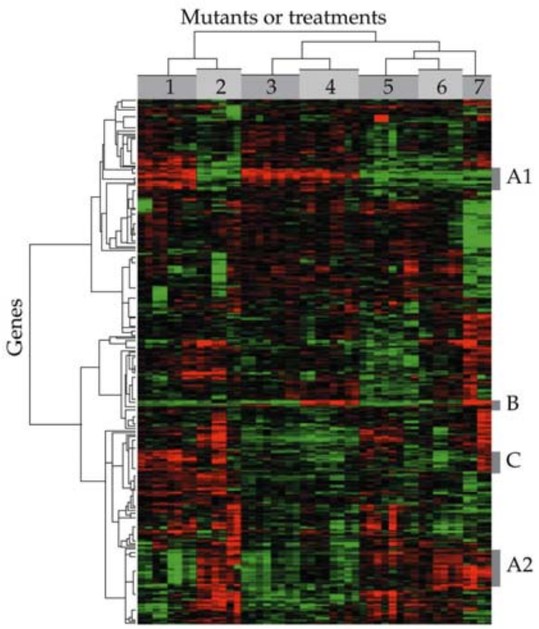
© 2009 Sinauer Associates, Inc.

Figure 4.19 Gene expression changes in yeast - boxes represent genes that are repressed (green) or induced (red) at least 2-fold, either after glucose limitation (upper left quadrant), or 250 generations of adaptive evolution (lower right quadrant).



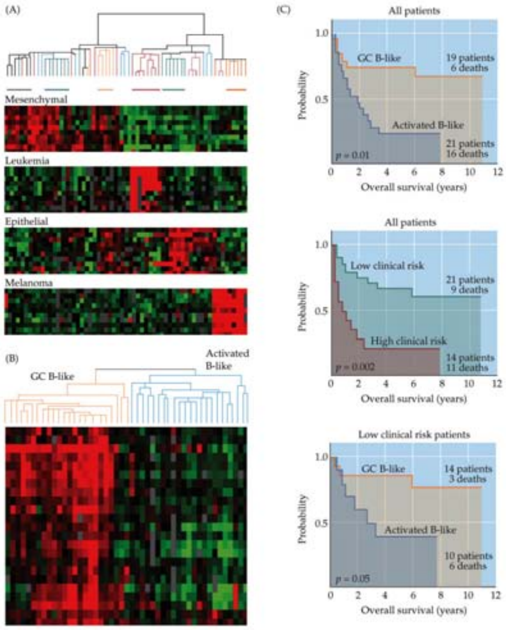
***A PRIMER OF GENOME SCIENCE 3e*, Figure 4.19**

Fig. 4.20 - The compendium approach - can cluster both genes and treatments.



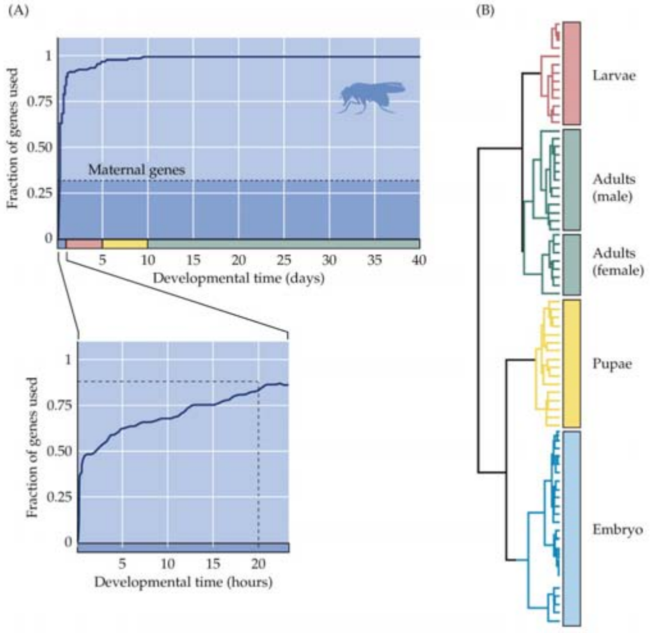
***A PRIMER OF GENOME SCIENCE 3e*, Figure 4.20**

Figure 4.21 - the molecular pharmacology of cancer.



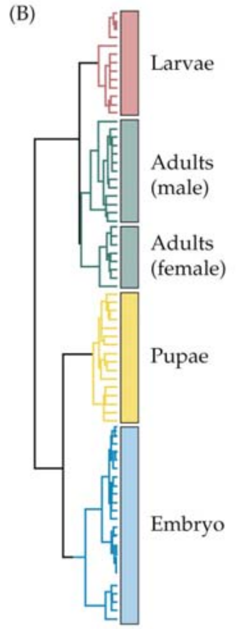
A PRIMER OF GENOME SCIENCE 3e, Figure 4.21

Figure 4.22 Developmental changes in gene expression in *Drosophila*



A PRIMER OF GENOME SCIENCE 3e, Figure 4.22

Figure 4.22 Developmental changes in gene expression in *Drosophila*



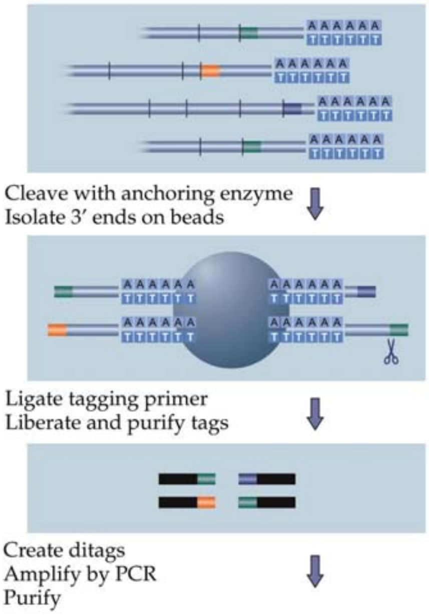
***A PRIMER OF GENOME SCIENCE 3e*, Figure 4.22 (Part 2)**

Gene expression analysis

- From EST clusters to spotted cDNA microarrays
- Long vs. short oligonucleotide microarrays vs. RT-PCR
- Methods of DNA microarray data analysis
- Serial analysis of gene expression (SAGE) and RNA-seq
- Promoter analysis
- Chromatin immunoprecipitation (ChIP-seq)

Serial Analysis of Gene Expression (SAGE)

- SAGE essentially amounts to an accelerated version of 3’ EST analysis.
- Briefly, double-stranded cDNA is cleaved with a restriction enzyme that has a 4-bp recognition sequence, 3’UTRs purified on streptavidin beads, ligated in pairs by using synthetic oligonucleotide adaptors, then cloned as 1 kb concatemers and sequenced.
- Some of the advantages of this method are that it is extremely sensitive, specific, and quantitative. All annotated genes are included. Comparison of different samples (data mining) is straightforward.
- Some of the disadvantages of this method are similar to those of EST analysis - non-specific transcripts will be included, alternative polyadenylation will confuse gene identities, cloning artifacts and sequencing errors further complicate the analysis.
- SAGE is slow and expensive. Hence it is usually not used in studies that require analysis of multiple samples.

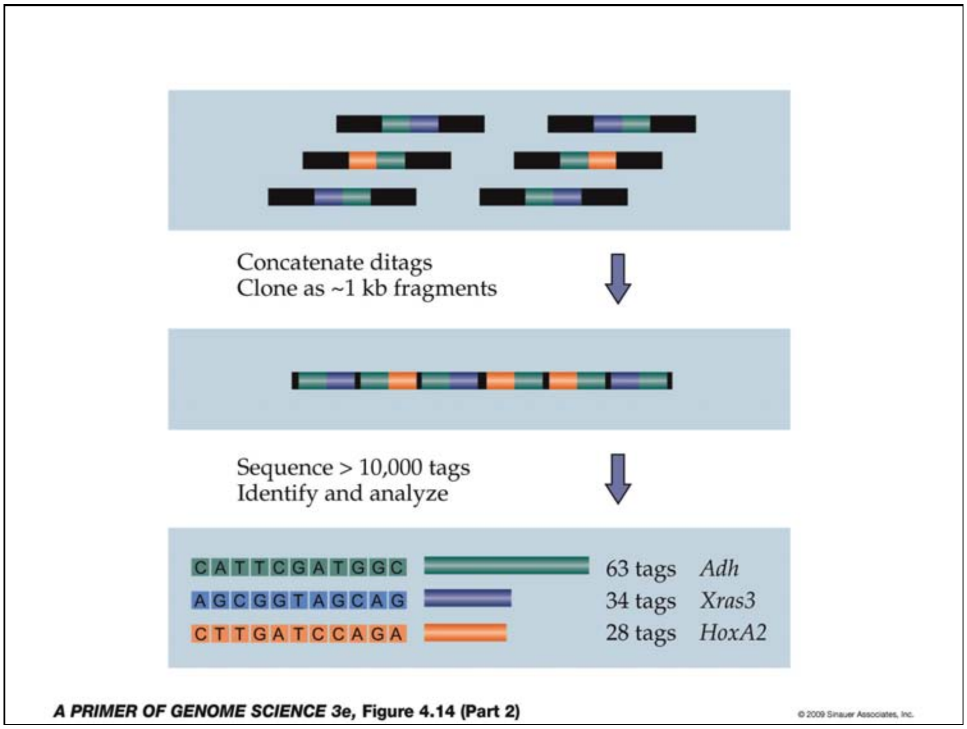


Cleave with anchoring enzyme
Isolate 3′ ends on beads

Ligate tagging primer
Liberate and purify tags

Create ditags
Amplify by PCR
Purify

***A PRIMER OF GENOME SCIENCE 3e*, Figure 4.14 (Part 1)**



Concatenate ditags
Clone as ~1 kb fragments

Sequence > 10,000 tags
Identify and analyze

CATTCGATGGC 63 tags *Adh*
AGCGGTAGCAG 34 tags *Xras3*
CTTGATCCAGA 28 tags *HoxA2*

***A PRIMER OF GENOME SCIENCE 3e*, Figure 4.14 (Part 2)**

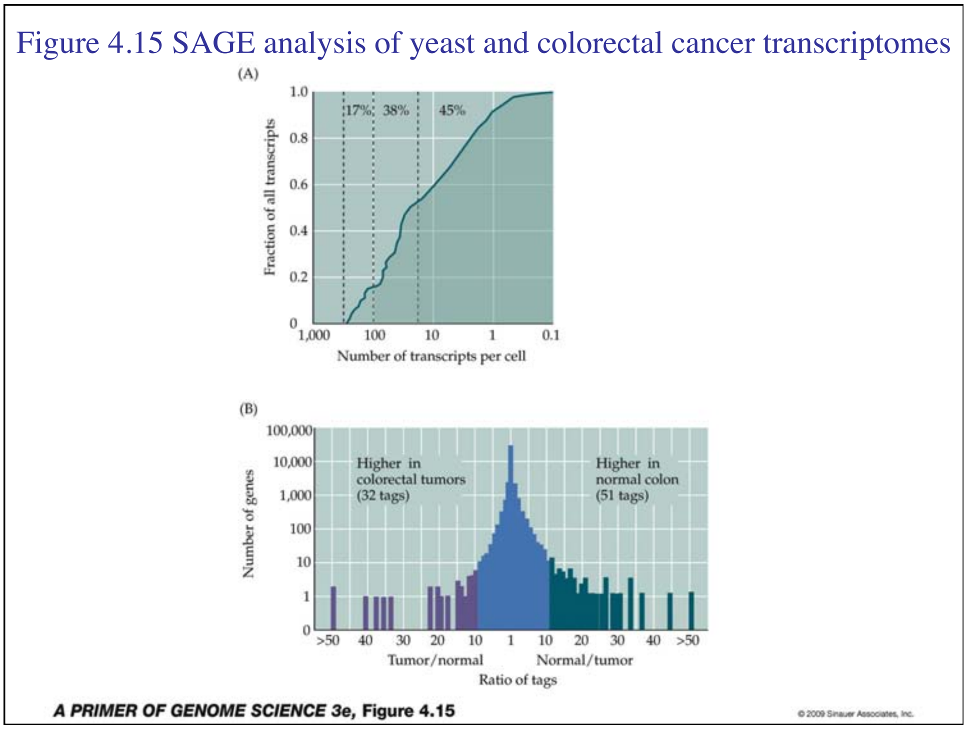
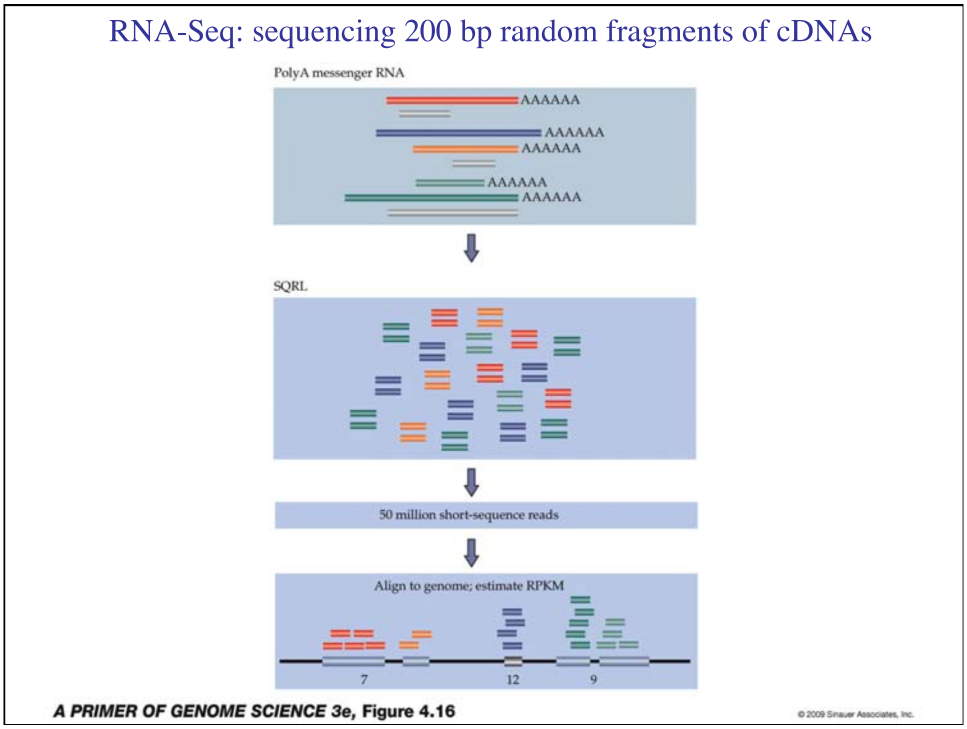


Figure 4.15 SAGE analysis of yeast and colorectal cancer transcriptomes

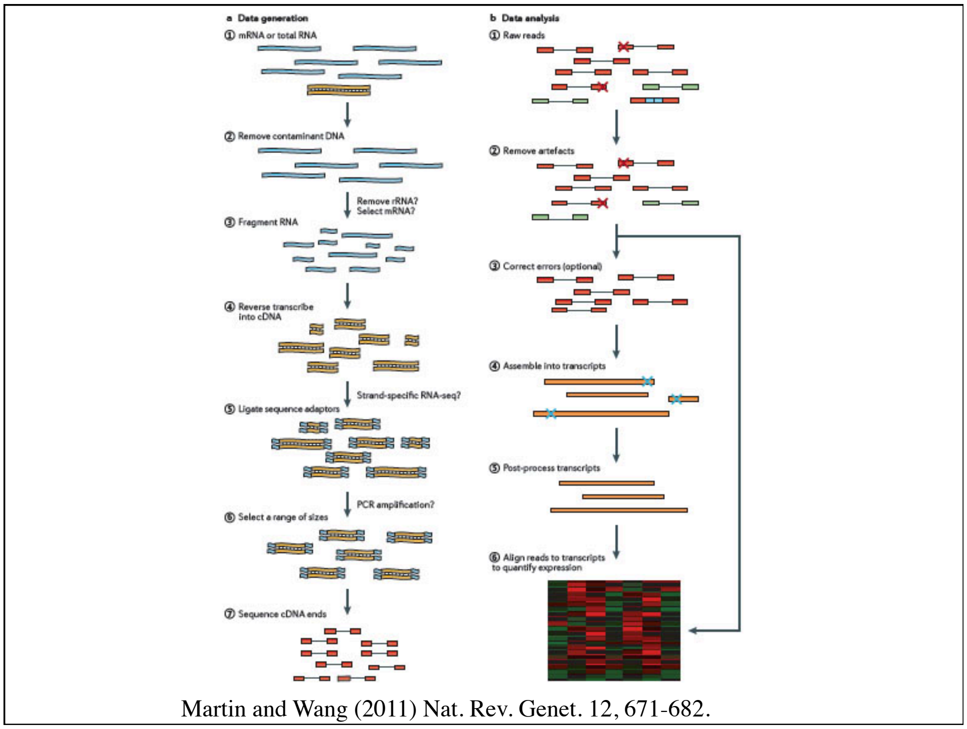
***A PRIMER OF GENOME SCIENCE 3e*, Figure 4.15**



RNA-Seq: sequencing 200 bp random fragments of cDNAs

PolyA messenger RNA
AAAAAA
AAAAAA
AAAAAA
AAAAAA
AAAAAA
SQRL
50 million short-sequence reads
Align to genome; estimate RPKM
7
12
9

***A PRIMER OF GENOME SCIENCE 3e*, Figure 4.16**



a Data generation
① mRNA or total RNA
② Remove contaminant DNA
Remove rRNA?
Select mRNA?
③ Fragment RNA
④ Reverse transcribe into cDNA
Strand-specific RNA-seq?
⑤ Ligate sequence adaptors
PCR amplification?
⑥ Select a range of sizes
⑦ Sequence cDNA ends
b Data analysis
① Raw reads
② Remove artefacts
③ Correct errors (optional)
④ Assemble into transcripts
⑤ Post-process transcripts
⑥ Align reads to transcripts to quantify expression

Martin and Wang (2011) Nat. Rev. Genet. 12, 671-682.

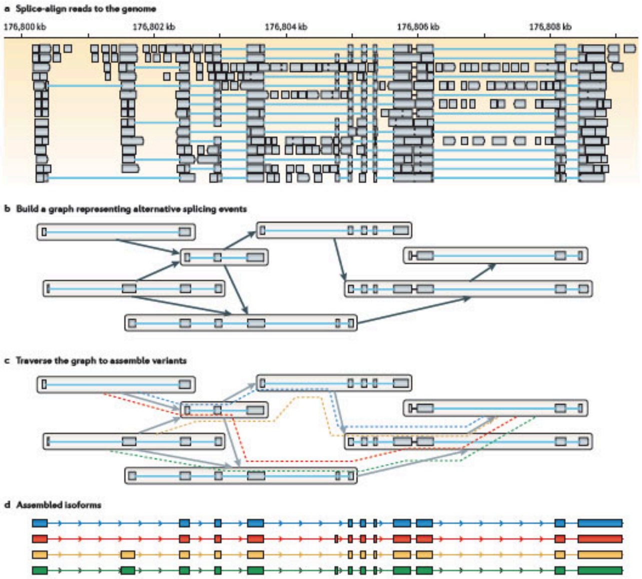
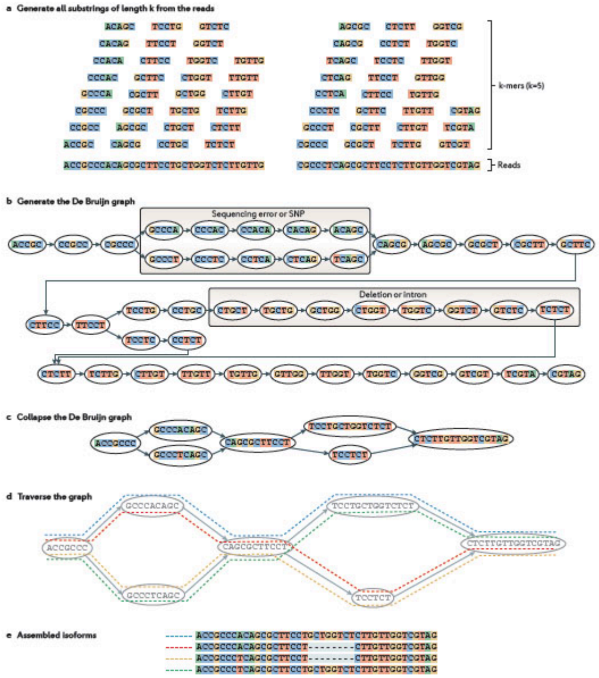

a Splice-align reads to the genome
176,800 kb
176,802 kb
176,804 kb
176,806 kb
176,808 kb
b Build a graph representing alternative splicing events
c Traverse the graph to assemble variants
d Assembled isoforms

Figure 2 | **Overview of the reference-based transcriptome assembly strategy.** The steps of the reference-based transcriptome strategy are shown using an example of a maize gene (GRMZM2O060216). **a** | Reads (grey) are first splice-aligned to a reference genome. **b** | A connectivity or splice graph is then constructed to represent all possible isoforms at a locus. **c,d** | Finally, alternative paths through the graph (blue, red, yellow and green) are followed to join compatible reads together into isoforms.

Martin and Wang (2011) Nat. Rev. Genet. 12, 671-682.


a Generate all substrings of length k from the reads
k-mers (k=5)
ACCGCCCACAGCGCTTCCTGCTGGTCTCTTGTTG
CGCCCTCAGCGCTTCCTCTTGTTGGTCGTAG
Reads
b Generate the De Bruijn graph
Sequencing error or SNP
Deletion or intron
c Collapse the De Bruijn graph
ACCGCCC
GCCCACAGC
GCCCTCAGC
CAGCGCTTCCT
TCCTGCTGGTCTCT
TCCTCT
CTCTTGTTGGTCGTAG
d Traverse the graph
e Assembled isoforms
ACCGCCCACAGCGCTTCCTGCTGGTCTCTTGTTGGTCGTAG
ACCGCCCACAGCGCTTCCT--------CTTGTTGGTCGTAG
ACCGCCCTCAGCGCTTCCT--------CTTGTTGGTCGTAG
ACCGCCCTCAGCGCTTCCTGCTGGTCTCTTGTTGGTCGTAG

Martin and Wang (2011) Nat. Rev. Genet. 12, 671-682.

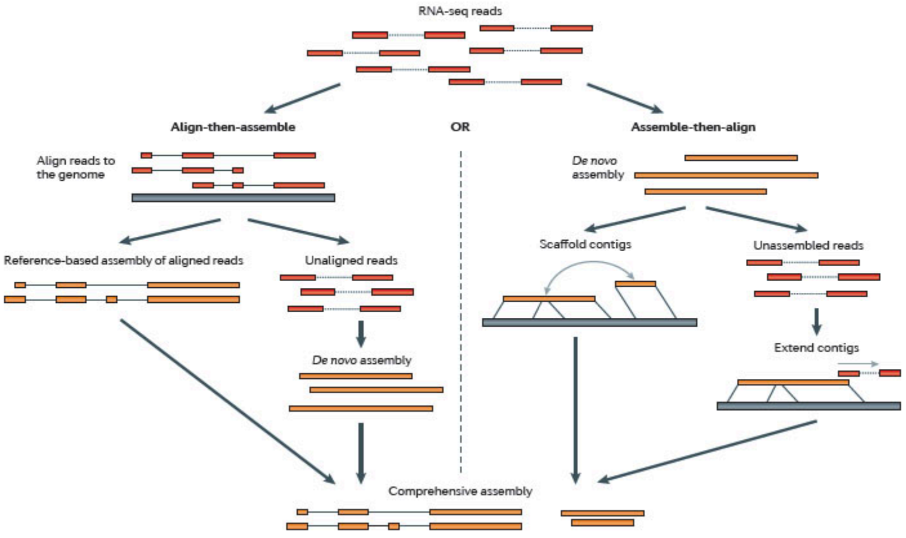
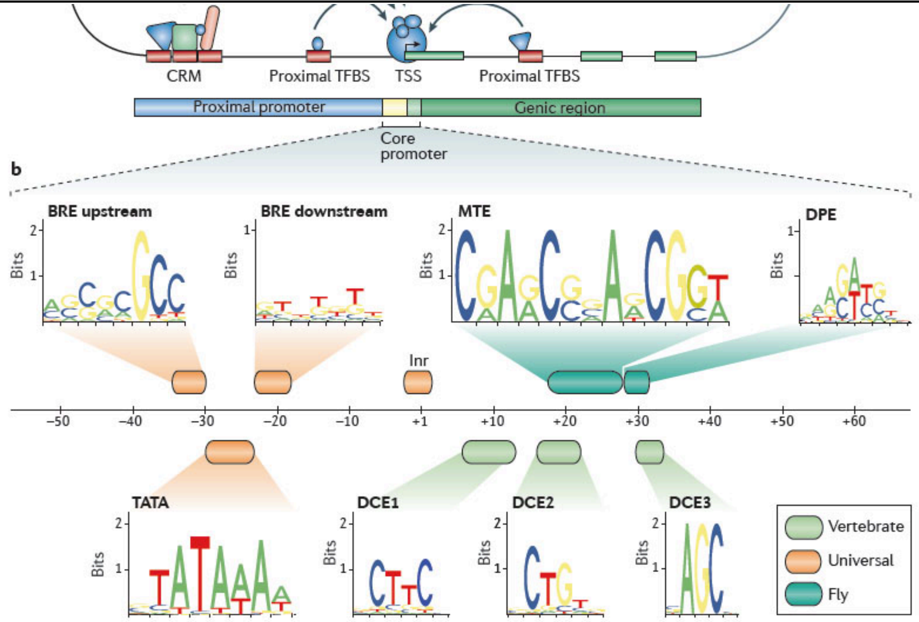


Figure 4 | **Alternative approaches for combined transcriptome assembly.** The left choice depicts the align-then-assemble strategy, in which reference-based assembly is followed by *de novo* assembly of reads that failed to align to the genome. The right choice depicts the assemble-then-align strategy, in which the reads are first *de novo* assembled and then scaffolded and extended using a reference genome. RNA sequencing (RNA-seq) reads are shown in red, and assembled transcripts are shown in orange.

Martin and Wang (2011) Nat. Rev. Genet. 12, 671-682.

Gene expression analysis

- From EST clusters to spotted cDNA microarrays
- Long vs. short oligonucleotide microarrays vs. RT-PCR
- Methods of DNA microarray data analysis
- Serial analysis of gene expression (SAGE) and RNA-seq
- Promoter analysis
- Chromatin immunoprecipitation (ChIP-seq)



Lenhard et al. (2012) Nat. Rev. Genet. 13, 233-245.

Types of promoters in Metazoa

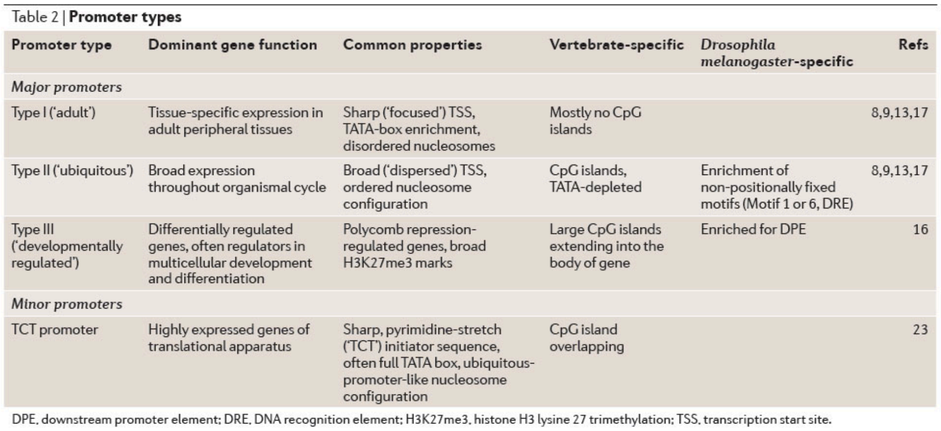


Table 2 | **Promoter types**

Promoter type	Dominant gene function	Common properties	Vertebrate-specific	*Drosophila melanogaster*-specific	Refs
Major promoters					
Type I ('adult')	Tissue-specific expression in adult peripheral tissues	Sharp ('focused') TSS, TATA-box enrichment, disordered nucleosomes	Mostly no CpG islands		8,9,13,17
Type II ('ubiquitous')	Broad expression throughout organismal cycle	Broad ('dispersed') TSS, ordered nucleosome configuration	CpG islands, TATA-depleted	Enrichment of non-positionally fixed motifs (Motif 1 or 6, DRE)	8,9,13,17
Type III ('developmentally regulated')	Differentially regulated genes, often regulators in multicellular development and differentiation	Polycomb repression-regulated genes, broad H3K27me3 marks	Large CpG islands extending into the body of gene	Enriched for DPE	16
Minor promoters					
TCT promoter	Highly expressed genes of translational apparatus	Sharp, pyrimidine-stretch ('TCT') initiator sequence, often full TATA box, ubiquitous-promoter-like nucleosome configuration	CpG island overlapping		23

DPE, downstream promoter element; DRE, DNA recognition element; H3K27me3, histone H3 lysine 27 trimethylation; TSS, transcription start site.

Lenhard et al. (2012) Nat. Rev. Genet. 13, 233-245.

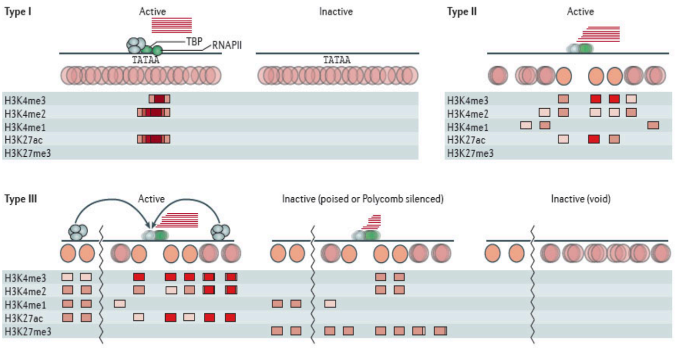


Figure 2 | **Features of the main functional classes of metazoan promoters.** Based on the configuration of promoter signals, transcription start site (TSS) positions, nucleosome positions and their epigenetic marks, most metazoan promoters can be categorized into three general types. This classification represents a broad generalization and should still be considered to be a work in progress. Type I ('adult') promoters are most often used for genes that are specifically expressed in terminally differentiated peripheral tissues of an adult. Type II ('ubiquitous') promoters are active in all cell types. Type III promoters are characteristic of genes with expression that is developmentally regulated and coordinated across multiple cells. TABLE 2 gives more detail about each of the three main types. The DNA is the horizontal black line; the red lines correspond to 5′ ends of individual transcripts and indicate different TSS precision in different promoter types. Sequence-specific transcription factor complexes are in grey, and general transcription factors are in green. The 'fuzziness' represents postulated RNA polymerase II (RNAPII) positioning for TATA-independent initiation on 'broad' promoters. The complexes by no means represent a comprehensive inventory of components at core promoters. Beneath each promoter, nucleosomes are represented by red circles; the 'fuzziness' represents the degree of nucleosome positioning. Histone modifications are shown below the nucleosomes, and the depth of colour represents the prevalence of the modification (with red being the most prevalent). For type III promoters, scenarios with poised RNAPII and/or Polycomb silencing and without RNAPII are shown. H3K27ac, histone H3 lysine 27 acetylation; H3K4me1, histone H3 lysine 4 methylation; TBP, TATA-box-binding protein.

Lenhard et al. (2012) Nat. Rev. Genet. 13, 233-245.

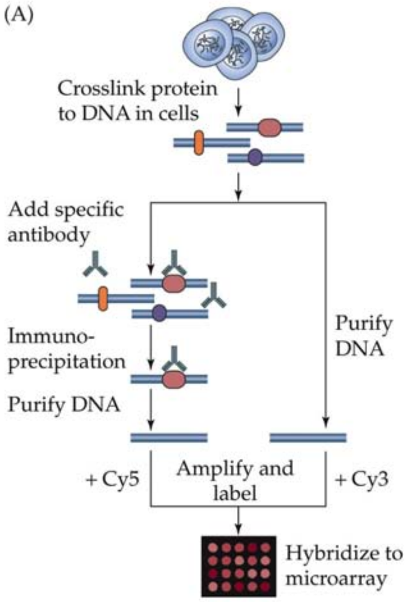
Gene expression analysis

- From EST clusters to spotted cDNA microarrays
- Long vs. short oligonucleotide microarrays vs. RT-PCR
- Methods of DNA microarray data analysis
- Serial analysis of gene expression (SAGE) and RNA-seq
- Promoter analysis
- Chromatin immunoprecipitation (ChIP-seq)

Chromatin immunoprecipitation (ChIP on chip)

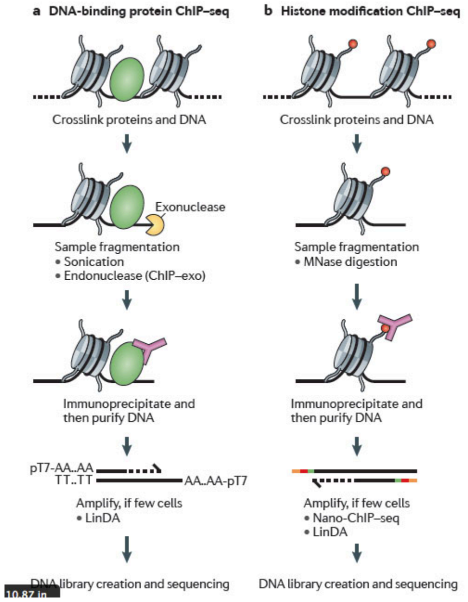
- One goal for future microarray analysis methods is to use gene regulatory pathways to constrain (or inform) cluster analysis.
- A popular method that attempts to discover gene regulatory pathways at the whole genome level involves chromatin immunoprecipitation (first crosslink proteins to DNA in living cells, then shear the DNA and add specific antibody to one transcription factor, and immunoprecipitate the complexes).
- DNA fragments that were purified in this way can be fluorescently labeled, and hybridized to genomic microarrays (or promoter microarrays).
- The result is a high-resolution, physical map of the binding sites of a particular transcription factor, to all gene targets in the genome, under physiological conditions, in the living cell.
- Additional experiments are required to establish whether this binding has a positive, negative, or no effect on transcription.

Figure 4.13 Chromatin immunoprecipitation and regulatory pathways



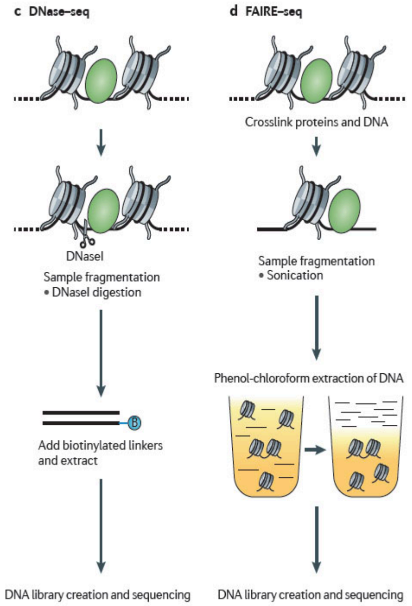
A PRIMER OF GENOME SCIENCE 3e, Figure 4.13 (Part 1)

ChIP for transcription factors or histone modifications


a DNA-binding protein ChIP–seq
b Histone modification ChIP–seq
Crosslink proteins and DNA
Crosslink proteins and DNA
Exonuclease
Sample fragmentation
• Sonication
• Endonuclease (ChIP–exo)
Sample fragmentation
• MNase digestion
Immunoprecipitate and then purify DNA
Immunoprecipitate and then purify DNA
pT7-AA..AA
TT..TT
AA..AA-pT7
Amplify, if few cells
• LinDA
Amplify, if few cells
• Nano-ChIP–seq
• LinDA
10.87 in
DNA library creation and sequencing
DNA library creation and sequencing

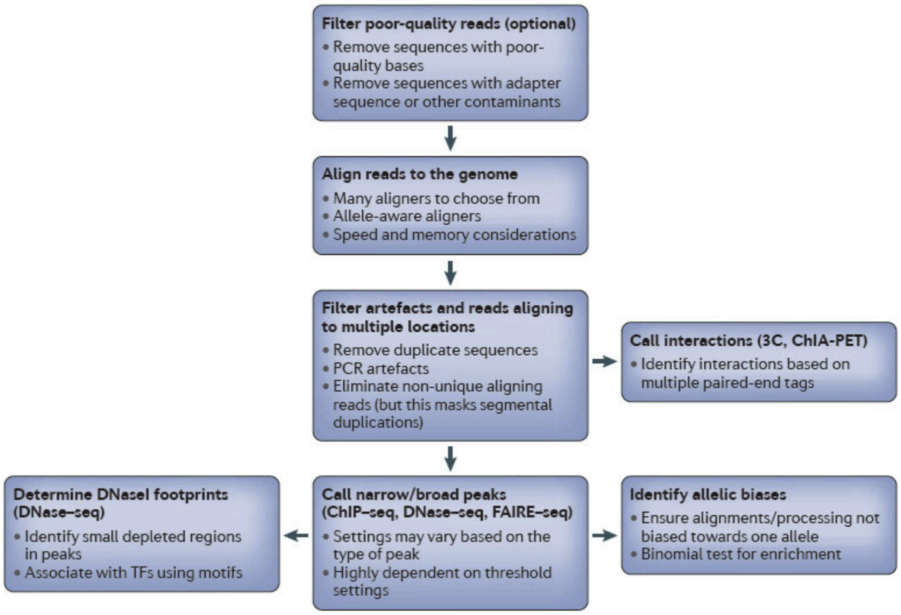
Furey (2012) Nat. Rev. Genet. 13, 840-852.

Dnase-hypersensitive domains, formaldehyde cross-linking


c DNase–seq
d FAIRE–seq
Crosslink proteins and DNA
DNaseI
Sample fragmentation
• DNaseI digestion
Sample fragmentation
• Sonication
Phenol-chloroform extraction of DNA
B
Add biotinylated linkers and extract
DNA library creation and sequencing
DNA library creation and sequencing

Furey (2012) Nat. Rev. Genet. 13, 840-852.

ChIP data analysis



Filter poor-quality reads (optional)
- Remove sequences with poor-quality bases
- Remove sequences with adapter sequence or other contaminants

Align reads to the genome
- Many aligners to choose from
- Allele-aware aligners
- Speed and memory considerations

Filter artefacts and reads aligning to multiple locations
- Remove duplicate sequences
- PCR artefacts
- Eliminate non-unique aligning reads (but this masks segmental duplications)

Call interactions (3C, ChIA-PET)
- Identify interactions based on multiple paired-end tags

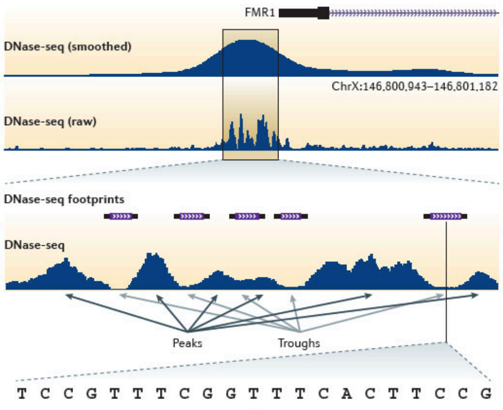
Determine DNaseI footprints (DNase–seq)
- Identify small depleted regions in peaks
- Associate with TFs using motifs

Call narrow/broad peaks (ChIP–seq, DNase–seq, FAIRE–seq)
- Settings may vary based on the type of peak
- Highly dependent on threshold settings

Identify allelic biases
- Ensure alignments/processing not biased towards one allele
- Binomial test for enrichment

Furey (2012) Nat. Rev. Genet. 13, 840-852.

Dnase-seq

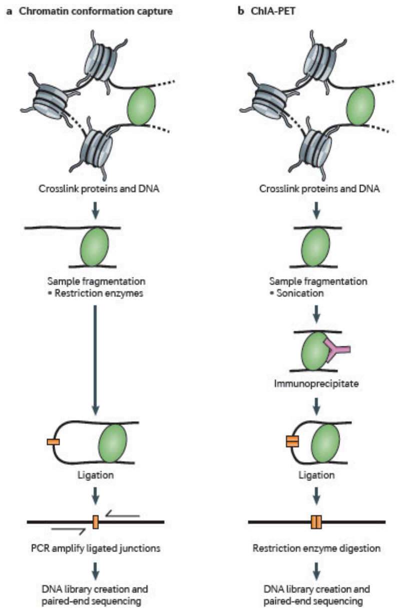


Motifs from JASPAR database

Model name	Score	Relative score	Start	End	Strand	Predicted site sequence
IRF1	12.986	0.904279917181229	3	14	–1	GAAACCGAAACG
IRF2	17.216	0.907706906384892	4	21	–1	CGGAAGTGAAACCGAAAC
SPIB	4.820	0.806987596140569	5	11	–1	ACCGAAA
BRCA1	4.228	0.802287513481405	8	14	–1	GAAACCG

Furey (2012) Nat. Rev. Genet. 13, 840-852.

Chromatin-chromatin interactions


a Chromatin conformation capture
b ChIA-PET
Crosslink proteins and DNA
Crosslink proteins and DNA
Sample fragmentation
• Restriction enzymes
Sample fragmentation
• Sonication
Immunoprecipitate
Ligation
Ligation
PCR amplify ligated junctions
Restriction enzyme digestion
DNA library creation and paired-end sequencing
DNA library creation and paired-end sequencing

Furey (2012) Nat. Rev. Genet. 13, 840-852.

Detection of allele-specific bias in ChIP-seq

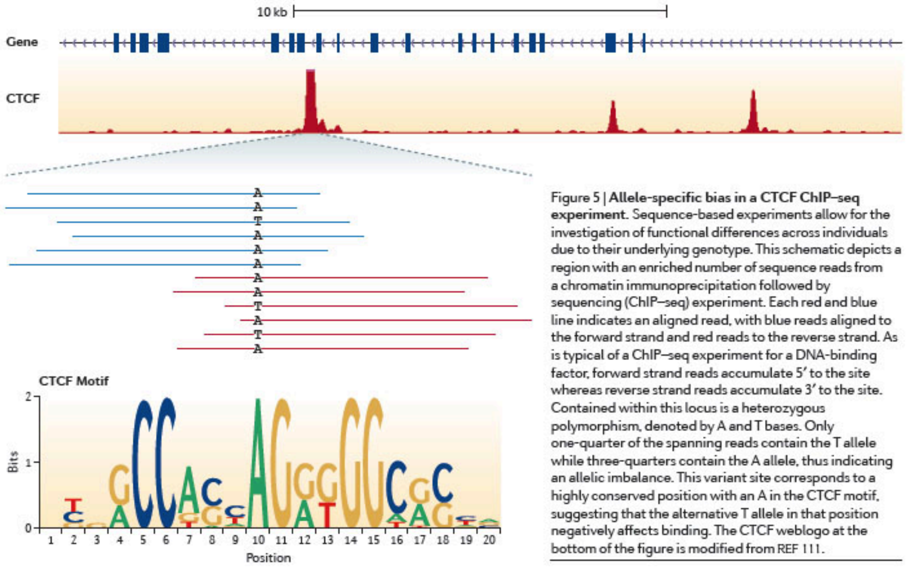
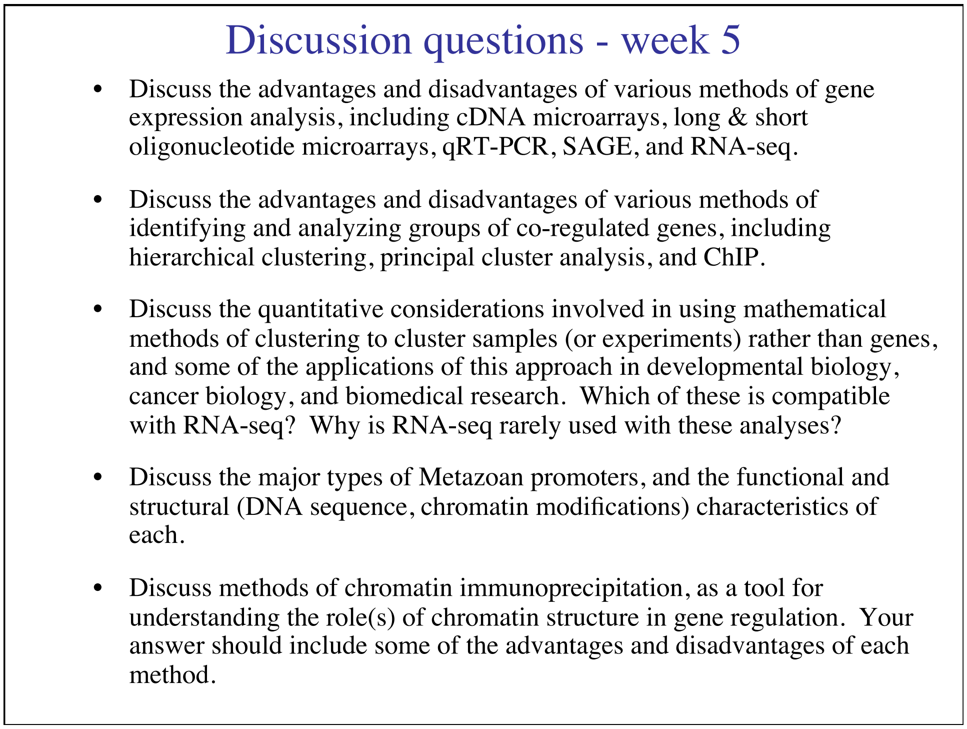


Figure 5 | Allele-specific bias in a CTCF ChIP–seq experiment. Sequence-based experiments allow for the investigation of functional differences across individuals due to their underlying genotype. This schematic depicts a region with an enriched number of sequence reads from a chromatin immunoprecipitation followed by sequencing (ChIP–seq) experiment. Each red and blue line indicates an aligned read, with blue reads aligned to the forward strand and red reads to the reverse strand. As is typical of a ChIP–seq experiment for a DNA-binding factor, forward strand reads accumulate 5′ to the site whereas reverse strand reads accumulate 3′ to the site. Contained within this locus is a heterozygous polymorphism, denoted by A and T bases. Only one-quarter of the spanning reads contain the T allele while three-quarters contain the A allele, thus indicating an allelic imbalance. This variant site corresponds to a highly conserved position with an A in the CTCF motif, suggesting that the alternative T allele in that position negatively affects binding. The CTCF weblogo at the bottom of the figure is modified from REF 111.

Furey (2012) Nat. Rev. Genet. 13, 840-852.



Discussion questions - week 5

- Discuss the advantages and disadvantages of various methods of gene expression analysis, including cDNA microarrays, long & short oligonucleotide microarrays, qRT-PCR, SAGE, and RNA-seq.
- Discuss the advantages and disadvantages of various methods of identifying and analyzing groups of co-regulated genes, including hierarchical clustering, principal cluster analysis, and ChIP.
- Discuss the quantitative considerations involved in using mathematical methods of clustering to cluster samples (or experiments) rather than genes, and some of the applications of this approach in developmental biology, cancer biology, and biomedical research. Which of these is compatible with RNA-seq? Why is RNA-seq rarely used with these analyses?
- Discuss the major types of Metazoan promoters, and the functional and structural (DNA sequence, chromatin modifications) characteristics of each.
- Discuss methods of chromatin immunoprecipitation, as a tool for understanding the role(s) of chromatin structure in gene regulation. Your answer should include some of the advantages and disadvantages of each method.