How many of you have checked out the web site on protein-DNA interactions?
Example of an approximately 40,000 probe spotted oligo microarray with enlarged inset to show detail.
Find and be ready to discuss how DNA microarrays were used to study a particular problem. Some keywords that might help are genome-wide expression profiling, cDNA microarrays
Uses of DNA microarrays

<table>
<thead>
<tr>
<th>Technology or Application</th>
<th>Synopsis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gene expression profiling</td>
<td>In an mRNA or gene expression profiling experiment the expression levels of thousands of genes are simultaneously monitored to study the effects of certain treatments, diseases, and developmental stages on gene expression. For example, microarray-based gene expression profiling can be used to identify genes whose expression is changed in response to pathogens or other organisms by comparing gene expression in infected to that in uninfected cells or tissues.</td>
</tr>
</tbody>
</table>
DNA microarray technology
### Cluster analysis or clustering

Cluster analysis or clustering is the assignment of a set of observations into subsets (called clusters) so that observations in the same cluster are similar in some sense. Clustering is a method of unsupervised learning, and a common technique for statistical data analysis used in many fields, including machine learning, data mining, pattern recognition, image analysis and bioinformatics. Besides the term clustering, there are a number of terms with similar meanings, including automatic classification, numerical taxonomy, botryology and typological analysis.
Example of both a heat map and of clustering
### Uses of DNA microarrays

<table>
<thead>
<tr>
<th><strong>Comparative genomic hybridization</strong></th>
<th>Assessing genome content in different cells or closely related organisms. [5][6]</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>GeneID</strong></td>
<td>Small microarrays to check IDs of organisms in food and feed (like GMO [1]), mycoplasms in cell culture, or pathogens for disease detection, mostly combining PCR and microarray technology.</td>
</tr>
</tbody>
</table>
Comparative Genomics

1. Labeling of genomic tumor DNA and normal genomic control DNA by Nick translation
   - Biotin-labeled tumor DNA
   - Digoxigenin-labeled control DNA

2. Simultaneous hybridization of differentially labeled tumor and control DNAs to normal human metaphase spreads

3. Fluorescence detection of the hybridized DNAs
Comparative Genomics

4. Result

- balanced DNA content
- overrepresentation of the whole chromosome within the tumor DNA
- underrepresentation of the long arm within the tumor DNA
- high level amplification
## Uses of DNA microarrays

<p>| Chromatin immunoprecipitation on Chip | DNA sequences bound to a particular protein can be isolated by immunoprecipitating that protein (ChIP), these fragments can be then hybridized to a microarray (such as a tiling array) allowing the determination of protein binding site occupancy throughout the genome. Example protein to immunoprecipitate are histone modifications (H3K27me3, H3K4me2, H3K9me3, etc), Polycomb-group protein (PRC2:Suz12, PRC1:YY1) and trithorax-group protein (Ash1) to study the epigenetic landscape or RNA Polymerase II to study the transcription landscape. |</p>
<table>
<thead>
<tr>
<th>Uses of DNA microarrays</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Tiling array</strong></td>
</tr>
<tr>
<td><strong>SNP detection</strong></td>
</tr>
</tbody>
</table>
Fabrication of Arrays

• Spotting

probes are synthesized prior to deposition on the array surface and are then "spotted" onto glass. A common approach utilizes an array of fine pins or needles controlled by a robotic arm that is dipped into wells containing DNA probes and then depositing each probe at designated locations on the array surface.

Video clip
Fabrication of Arrays

- photolithography

Oligonucleotide arrays are produced by printing short oligonucleotide sequences designed to represent a single gene or family of gene splice-variants by synthesizing this sequence directly onto the array surface instead of depositing intact sequences. On a silica substrate where light and light-sensitive masking agents are used to "build" a sequence one nucleotide at a time across the entire array. Each applicable probe is selectively "unmasked" prior to bathing the array in a solution of a single nucleotide, then a masking reaction takes place and the next set of probes are unmasked in preparation for a different nucleotide exposure.
Methods for Working with DNA and RNA

4. DNA sequencing
   A. Sanger dideoxy method
      Use chain terminators and DNA polymerase
      Enzymatic approach
   B. Maxam-Gilbert method
      Chemical approach: nucleotide-specific modification
Methods for Working with DNA and RNA

Sanger dideoxy method

i. use of dideoxynucleotides as chain terminators will have four separate reactions for each chain terminator

ii. Klenow fragment of DNA polymerase

iii. single-stranded DNA template annealed to an oligonucleotide primer

iv. alpha P-32 dATP typically used to internally labeled synthesized DNA figure 7-15
Automated DNA sequencing

Fluorescent tagged DNA
Dye-labeled segments of DNA, copied from template with unknown sequence

DNA migration

Dye-labeled segments are applied to a capillary gel and subjected to electrophoresis.

Detector

Laser beam

Laser

Computer-generated result after bands migrate past detector
Methods for Working with DNA and RNA

5. DNase I footprinting and Interference Assay
   A. radiolabel one unique 5' end of the DNA using T4 polynucleotide kinase and alpha-P32 ATP.
   B. after binding protein to DNA, add in DNasel (an endonuclease) to cut the regions of DNA not bound by protein
   C. It is key to cut the DNA only once per DNA molecule
Solution of identical DNA fragments radioactively labeled ◊ at one end of one strand.

Treat with DNase under conditions in which each strand is cut once (on average). No nicks are made in the area where RNA polymerase has bound.

Isolate labeled DNA fragments and denature. Only labeled strands are detected in next step.

Separate fragments by polyacrylamide gel electrophoresis and visualize radiolabeled bands on x-ray film.

Uncut DNA fragment

Missing bands indicate where RNA polymerase was bound to DNA.
Methods for Working with DNA and RNA

6. cDNA

- an oligo dT primer is annealed to the poly A portion of the mRNA template
- the enzyme reverse transcriptase catalyzes the synthesis of DNA from an RNA template
- reverse transcriptase backs around to place a "hook" at the end - DNA-DNA-DNA primer
- RNA is removed by alkaline hydrolysis
- DNA synthesis is completed with DNA polymerase
- This process is used in molecular cloning see next section
mRNA template is annealed to synthetic oligonucleotide (oligo d(T)) primer.

Reverse transcriptase and dNTPs yield a complementary DNA strand.

mRNA is degraded with alkali.

DNA polymerase I and dNTPs yield double-stranded DNA.
7. PCR or polymerase chain reaction  
   A. An effective method for in vitro amplification of a particular DNA sequencing  
      as much as a million times amplification with 25 cycles  
   B. relies on using a thermostable version of DNA polymerase  
      example is Taq DNA polymerase  
   C. can be used for  
      molecular cloning (also RT-PCR, combine PCR with synthesis of cDNA)  
      site-directed mutagenesis  
      rapid identification of clone  
      rapid DNA sequencing to name just a few  
   D. Extreme care needs to be taken to prevent contamination  
      even dust can be significant
Region of target DNA
to be amplified

1. Heat to separate strands.
2. Cool; add synthetic oligonucleotide primers.

3. Add thermostable DNA polymerase to catalyze $5' \rightarrow 3'$ DNA synthesis.

Repeat steps 1 and 2.

DNA synthesis (step 3) is catalyzed by the thermostable DNA polymerase (still present).

(a)
DNA synthesis (step ③) is catalyzed by the thermostable DNA polymerase (still present).

Repeat steps ① through ③.

After 25 cycles, the target sequence has been amplified about $10^6$-fold.

(a)
Questions

• How does changing the ratio of dideoxynucleotide to deoxynucleotide (i.e. ddATP vs dATP) effect Dideoxy (Sanger) Method of DNA sequencing? Could this effect the range that can be sequenced and why?

• How can PCR be used to monitor which version of a particular gene or large alteration in that gene is present in your colony of cells? How could PCR be used to isolate genes that are related to each other be a conserved domain?

• How is DNA supercoiling measured? Which amino acids could be used to recognize the major and minor groove of DNA and to contact the phosphate backbone? How does the lac repressor protein bind DNA?
examples

- [http://books.google.com/books?id=baLqFvWJz30C&pg=PA41&lpg=PA41&dq=protein-DNA+interactions+recognition&source=web&ots=4R5Atfm1qd&sig=wDJv4Dfz6oudJKPIGYBRrDKVZm0&hl=en&sa=X&oi=book_result&resnum=10&ct=result#PPA42,M1](http://books.google.com/books?id=baLqFvWJz30C&pg=PA41&lpg=PA41&dq=protein-DNA+interactions+recognition&source=web&ots=4R5Atfm1qd&sig=wDJv4Dfz6oudJKPIGYBRrDKVZm0&hl=en&sa=X&oi=book_result&resnum=10&ct=result#PPA42,M1)


What is non-specific versus specific binding of DNA?

- Also example of protein bending DNA: TBP protein

What DNA sequence might be preferred for this kind of protein?
(a) Heat to separate strands.
(2) Anneal primers containing noncomplementary regions with cleavage site for restriction endonuclease.

(b) Clone by insertion at an EcoRI site in a cloning vector.
Chemical Synthesis of DNA

Solid-phase chemistry based
Synthesis of oligonucleotide
Phosphoramidite Method

DMT = dimethoxytrityl
Removed by acid treatment

Nucleoside protected at 5' hydroxyl

1. Nucleoside attached to silica support

2. Protecting group removed
DMT – O

Cyanoethyl protecting group
NC–(CH₂)₂–O–P

Diisopropylamino activating group
(CH₃)₂CH–N⁺–CH(CH₃)₂

Base₁

Next nucleotide added
(CH₃)₂CH–N–CH(CH₃)₂

Diisopropylamine byproduct

Base₂

Nucleotide activated at 3' position
Repeat steps 2 to 4 until all residues are added.

- Remove protecting groups from bases
- Remove cyanoethyl groups from phosphates
- Cleave chain from silica support

Oligonucleotide chain
Questions

• Why are cDNA libraries still so important in the era of genomics?
• What information can be derived from cDNA that are not obtained generally by DNA microarrays?
High throughput DNA sequencing

here is an example of what has been called 454 sequencing

http://www.roche-applied-science.com/publications/multimedia/genome_seq uencer/amplicon_07/wbt.htm