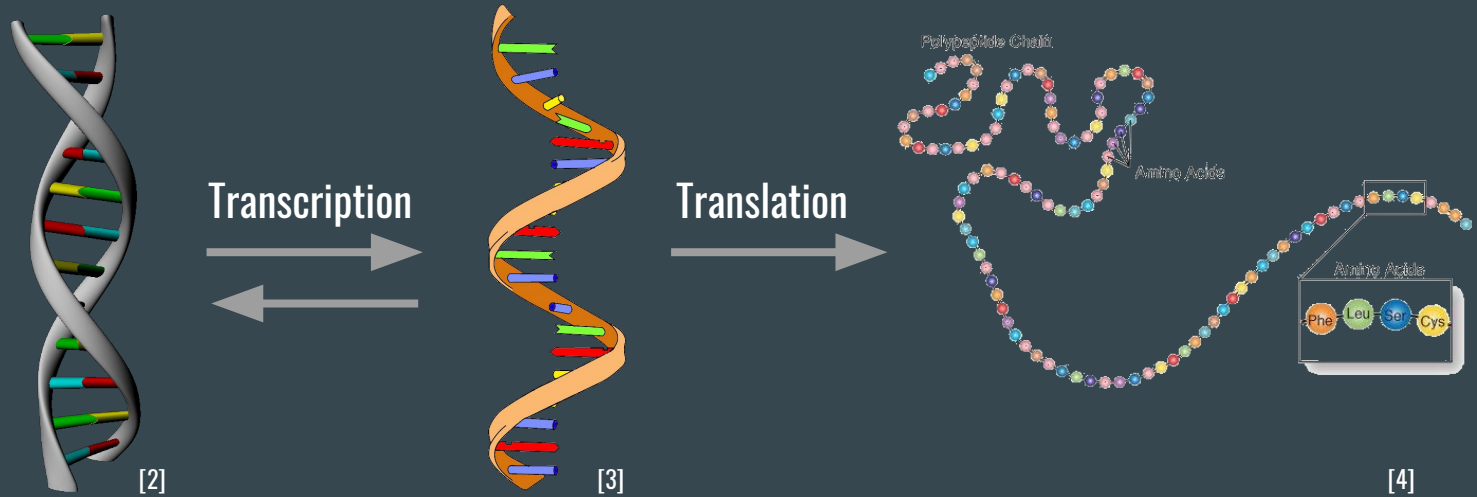
Microarrays

By Jake Sauter

- Background
- Introduction
- Fabrication
  - Deposition Techniques
  - In Situ Synthesis
- Applications
- Challenges
  - Noise
  - Variability
  - . . .

Central Dogma of Molecular Biology

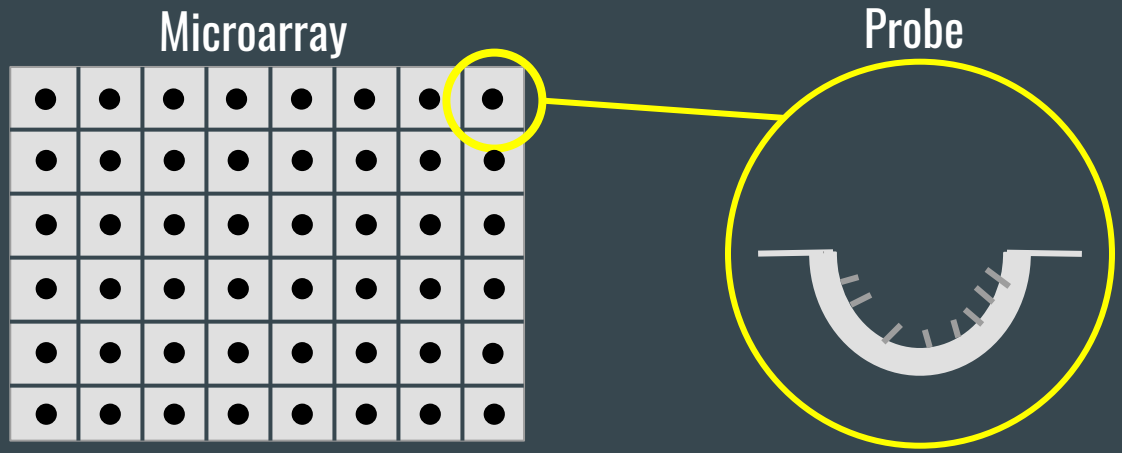


[2] [3] [4]

DNA RNA Protein

Microarray Concept

- DNA Microarray - Usually a substrate (nylon membrane, glass or plastic) on which one deposits single stranded DNAs (ssDNAs) with various sequences.
- Different configurations of deposits (in the probe) can be implemented depending on the purpose of the study



Measuring Gene Expression with Microarrays

- Probes contain DNA material complementary to the the target DNA
- Target DNA material is fluorescently labelled and will hybridize (pair to) the complementary DNA (cDNA) in the wells
- The level of hybridization (fluorescence level) of each probe can easily be measured in a scanner and can indicate the level of expression of a gene corresponding to the cDNA in the well.
- Expression levels from DNA samples of different tissues can be compared using multiple microarrays or on the same microarray (competitive hybridization)

Fabrication Techniques

- Two fabrication types are widely used
  - Deposition
  - In Situ Synthesis
- Each has its pros/cons that we will discuss later
- Deposition tends to result in longer chains of DNA
- In Situ synthesis allows for man made oligonucleotides (strands of nucleotides) but tend to be shorter (~20-50 bp) to prevent errors

Deposition Techniques

- DNA is prepared away from the chip
- Robots dip thin pins into the solution containing DNA material and then touch the pins onto the surface of the array
- Spotted arrays use small sequences to whole genes and even PCR products (clones)
- Gene expression in most eukaryotes (nucleus containing cells) is studied by utilizing complementary DNA (cDNA) clones, which allow for amplification of sufficient of sufficient quantities of DNA for deposition
    - Introns are also removed in this process as prokaryotic DNA does not contain introns

Deposition Techniques - Cloning

- Mature RNA (mRNA) is reverse transcribed into short cDNAs and introduced into bacterial hosts, which are grown isolated, then selected out if they carry foreign DNA
- The bacteria are prokaryotic and do not contain introns in their DNA
  - introns are intragenic material that does not make it through the transcription process to mRNA
- Two methods of cloning are used
  - ESTs - (expressed sequence tags) are cheap single pass sequences of entire clone libraries and result in partial sequences of clone inserts that are long enough to uniquely identify gene fragments
  - PCR - (polymerase chain reactions) are used to amplify clones containing desired fragments, after which they are purified and result in better clones than ESTs

In Situ Synthesis

- In Situ is latin for "on site" or "locally"
- Short synthesized oligonucleotides are attached to the solid support of the microarray
- Probes can be designed to detect multiple variant regions of a transcript (splice variants) and can be short enough to detect specific exons (intragenic material than is transcription to mRNA)
- This measurement of the abundance of splice variants is not possible with spotted DNA arrays due to probes of variable length which more than one different splice variant may hybridize to

In Situ Synthesis - Strengths

- Probe selection is based on sequence information alone. This means that every probe synthesized on the array is know in contrast to cDNA arrays, which deal with ESTs, and in many cases the function of the sequence to a spot is unknown

- This technology can distinguish and quantitatively monitor closely related genes for the simple fact that it can avoid identical sequences among gene family members due to the precisely synthesized oligonucleotides

In Situ Synthesis - Manufacturing Techniques

- Photolithographic (Affymetrix)
  - Uses photolithographic masks for each base. The mask allows for the correct base to be deposited in all locations that require it
  - Allows for high density arrays but the length of DNA sequence constructed is limited due to a non-zero probability of error at each step
  - **The microarrays used in the observed study were produced from Affymetrix and were high-density oligonucleotide microarrays**
- Ink Jet
  - Employs the technology of ink-jet color printers
  - Four cartridges contain the standard nucleotides (A, C, G, T) and the print head moves across the substrate, depositing nucleotides as needed
- Electrosynthesis
  - Electrodes in each substrate are turned on when the nucleotide solution is currently on the chip is needed to bind to that specific well

In Situ Synthesis - Novel Assay Technique

- A company Illumina makes a DASL - c**D**NA mediated **A**nnealing, **S**election, **E**xtension and Ligation assay that is revolutionary
  - Due to the processes involved in the assay, short (~50 bp) sequences can be used, allowing for the use of degraded samples, but this is not novel for in situ synthesis
  - There is estimated to be ~400 million FFPE (usually degraded) samples archived in North America for cancer alone, with associated clinical outcomes.
  - the sensitivity of the platform is enhanced by PCR amplification using common primers and by having 30 replicate beads per probe

Comparison of Techniques

Deposition	In Situ Synthesis
Long sequences	Short sequences due to limitation of synthesis technology
Spot unknown sequences	Spot known sequences
More variability in the system	More reliable data
Easier to analyze with appropriate experimental design	More difficult to analyze

* Though the field is dynamic and shifts can occur rapidly

Applications

- Sequencing
- SNP detection
- Genotyping
- Disease Association
- Genetic Linkage
- Genomic Loss and Amplification
- Detection of chromosomal rearrangement

Challenges

- Noise
  - Introduced at each step in the complex process
- Normalization
  - Not always performed in the same manner
- Experimental Design
  - Not always thoughtfully designed
- Large number of genes
  - Sometimes finding the one influential gene in excess of 7000 is equivalent to finding a needle in a haystack
- Significance
  - Classical techniques (e.g. χ squared test) cannot be applied because the number of variables is much greater than the number of experiments

Challenges

- Biological factors
  - The expression level is the amount of protein produced, not the amount of mRNA (what is detected by the microarray)
  - Other tools can simply not be replaced by microarrays
  - Gene regulation to biological impact is a complicated non-linear mapping
- Array quality assessment
  - Sources of variability can include mRNA preparation, transcription and labelling processes

References

[1] Book

[2] Vong, Andrew. “3D DNA Strand Rendering in MAYA.” av_designs, 27 Mar. 2015, www.andrewvong.com/3d-dna-strand-rendering-in-maya/.

[3]“Ribonucleic Acid (RNA).” RNA Extraction, Quinnipiac University, qu.edu.iq/bt/wp-content/uploads/2015/12/RNA-extraction.pdf.

[4] charity-stanton. “What Organic Molecule Is DNA?” SlideServe, 12 Nov. 2014, www.slideserve.com/charity-stanton/what-organic-molecule-is-dna.