Molecular Workshop

Shane Byrne
Sheryl Maher
Overview

- Background Molecular Biology
- Introduction to Southern Blotting and PCR
- Assay Design
- Commercial PCR
- Controls and Validation

BREAK

- Setting Up a Lab
- Next Gen Sequencing and Hybridisation Arrays
- Future Developments
What is DNA

- DNA is the building block of life
- DNA is the genetic material we use to be us
- DNA is found in everything
- DNA is a sugar + a nitrogen base + some phosphate
- DNA is found in a cell nucleus
- DNA is made of nucleotides
- DNA contains four different bases A, C, G, T
DNA Structure

What does 5’ 3’ mean?

(a) adenosine monophosphate (AMP) (a nucleotide)
Bases

- Thymine
- Adenine
- Cytosine
- Guanine
What does DNA do in our cells

- DNA makes up genes
- Genes encode proteins
- Proteins have a function
  - Support, enzyme, transport etc.
Where do we find DNA

- DNA occurs in all living things
  - We have DNA
  - Parasites have DNA
  - Bacteria have DNA
  - Everything that is able to make a copy of its self has DNA (except RNA viruses)
Fun things about DNA

- DNA had two antiparallel strands.
- A will always bind with T
- G will always bind with C
- C will always bind with G
- T will always bind with A

- This complementarity means if we know the order of bases on one strand we can determine the sequence of the second strand.
- DNA has a net negative charge
- DNA strands are antiparallel
How can we use DNA in a clinical setting?

- Make copies of DNA
  - PCR
    - Real time
    - End point
- Hybridise DNA to a known sequence
  - Southern blots
  - Microarrays (bead or platform)
- Sequence it
Southern Blotting

Molecular Workshop

AIMS 1914-2014
Past. Present. Future
4-6 September
Sydney
Southern blotting

- Developed in the 1970’s
- Isolate DNA
  - Separate based on size (gel electrophoresis)
- Transfer to nylon membrane
- Fix the DNA to the surface
- Label a short fragment of DNA (gene or synthetic sequence)
- Hybridise the labelled probe (in solution) to the membrane (DNA)
  - Complementary base pairing allows DNA to bind
Southern procedure

- Run gel
- Transfer to membrane
- Hybridise probe (over night)
- Wash
- Detect
Southern Transfer
Southern Blot Analysis

http://www.mun.ca/biology/scarr/Southern_Blot_analysis.html
Melting Temperature (Tm)

- For a given sequence of DNA it is possible to calculate the temperature at which 50% of the hydrogen bonds will be broken.
- Is calculated to determine optimal hybridization conditions

\[ T_m = 69.3^\circ C + 0.41(\% \text{ G} + \text{ C})^\circ C \]
Importance of Tm

- Determine hybridisation temperature
- Help enhance specificity
- Prevent background signals.
Ways to modify specificity of DNA hybridisation

- DNA strands are held together by weak hydrogen bonds.
- Ways to alter DNA binding
  - Modify temperature
  - Alter pH
  - Change salt, or salt concentration
  - Introduce a detergent
  - Include organic compounds (denaturants)
  - Alter probe length
  - Alter probe G+C content
Temperature

- Altering the temperature of DNA hybridisation will effect the binding of the probe to the membrane.
- Increased temperatures leads to increase specificity
- Decreased temperatures leads to reduced specificity
pH

- Changes to the pH of the hybridisation solution will alter the binding of the DNA on the membrane and the probe.
- Above pH 10 the DNA is denatured and unable to form hydrogen bonds
- EDTA can protect against nuclease degradation
Salts

- Cations (Na⁺ K⁺) interact with the negatively charged phosphates to reduce electrostatic repulsion between the DNA strands
- Increases the speed of hybridisation
- Decreases the Tm of the probe sequence

- Destabilisation of the GC and AT base pairs at very high concentrations.

- Typically SSC – Sodium Citrate + Sodium Chloride
Detergents

- Prevent nonspecific binding at lower concentrations (0.1 - 1%)
  - Interaction of DNA with membrane

- Act as blocking agents at higher concentrations (7% SDS)

- SDS, Tween 20, Noniden P-40 commonly used
Denaturing reagents

- Including denaturing reagents (Formamide, Urea, ethylene glycol etc) reduces the temperature required for hybridisation
- Removes the risk of sample or probe degradation (possible when exposed to high temperatures for long times)
- May reduce specificity
Formamide

- Increased accessibility of complementary strands
- Decreased electrostatic repulsion
- Effect independent of nucleotide sequence
- 80% formamide encourages DNA:RNA binding rather than DNA:DNA binding.

**Figure 1.** Effect of formamide on $T_m$ of different DNA samples. (○), *M. smegmatis*; (●), *E. coli*; (Δ), *B. cereus*.

<table>
<thead>
<tr>
<th>Source of DNA</th>
<th>a</th>
<th>b</th>
<th>c</th>
<th>d</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>B. cereus</em></td>
<td>4.75</td>
<td>5.75</td>
<td>6.50</td>
<td>8.50</td>
</tr>
<tr>
<td><em>E. coli</em></td>
<td>4.50</td>
<td>4.80</td>
<td>5.50</td>
<td>8.62</td>
</tr>
<tr>
<td><em>M. smegmatis</em></td>
<td>4.00</td>
<td>5.25</td>
<td>5.50</td>
<td>10.25</td>
</tr>
</tbody>
</table>

(a), 0% formamide; (b), 30% formamide; (c), 50% formamide; (d), 70% formamide.
Application of Tm modification

• Probe is hybridised at fixed conditions
  – 42°C if there is formamide in the buffer
  – 65°C if the probe is in a salt buffer

• Specificity is achieved through washing.

• Changing the probe length or probe GC content can significantly effect the Tm.
  • With long probes (1kb) moving a couple of bases has less effect than with short probes (20bp)
Typical Washing protocol

- Hybridisation (6-24 hours)
  - (0.5M Sodium Phosphate pH 7.3, 1% BSA, 4% SDS)

- First wash (low stringency)
  - 30 mins 2X SSC/0.1% SDS

- Second wash (high stringency)
  - 30 mins 0.5X SSC/0.1% SDS

For a probe with the following...

<table>
<thead>
<tr>
<th>Homology to Target</th>
<th>GC-content</th>
<th>Buffer</th>
<th>Temperature</th>
</tr>
</thead>
<tbody>
<tr>
<td>80–100%</td>
<td>Average (40%)</td>
<td>0.5x SSC + 0.1% SDS</td>
<td>65°C, if probe is &gt; 100 bp</td>
</tr>
<tr>
<td>&lt; 80%</td>
<td>Average (40%)</td>
<td>0.5x SSC + 0.1% SDS</td>
<td>Approx. 60°C*</td>
</tr>
<tr>
<td>80–100%</td>
<td>High (50%)</td>
<td>0.1x SSC + 0.1% SDS</td>
<td>68°C</td>
</tr>
</tbody>
</table>
PCR

- Developed by Kerry Mullis in 1982
  - published in 1983
  - Patent finalised in 1985
- Patent on Taq DNA polymerase expired 2005.
- Allows specific regions of DNA to be replicated (more copies made)
- The region is defined by locations where primers bind, need one primer for each DNA strand.
- A cyclical process of separation of DNA strands (denaturation), binding of the primers (annealing) and synthesis of the new DNA (extension) enables many copies to be made.
PCR
PCR

- Denaturation
- Annealing
- Extension
Key considerations for PCR

**Reaction mixture**
- Template
- Buffer
- Enzyme
- Nucleotides
- Primers

**Reaction conditions**
- Denaturation temperature
- Annealing temperature
- Extension temperature
- Extension time
Template

- The template for PCR is DNA

- Isolated from
  - Cultured organisms
  - Clinical samples

- Synthesised from RNA - cDNA
Buffer

• Everything we considered for Southern optimisation applies
• Need to create the optimal environment for reaction
  • Optimal pH
  • Optimal salt/ion concentration
  • Optimal hybridisation
• Activity of Taq is strongly influenced by Mg$^{2+}$ concentration.
Enzyme

- The standard enzyme for PCR is the DNA polymerase from the *Thermus Aquaticus*
- Adds 1000 base pairs in less than 10 seconds at 72°C
- 1 error per 9000 bases
- Adds an extra A at the end of each strand.

- Different enzymes are now available
  - Improved fidelity
  - Hot start (improved sensitivity)
  - Faster synthesis
  - Improve proof-reading ability

Brock & Freeze, 1969
Nucleotides

- Need to include ATP, CTP, GTP and TTP in the reaction mix to allow new DNA to be synthesised.
- Typically used at (2 mM per base)
Primers

- Primers are the variable in PCR which determine specificity.
- There are two primers in a PCR, one which binds to each strand of DNA.
- The binding of a primer allows DNA polymerase to add nucleotides.

- If the reaction is kept consistent but the primers are changed it is completely different assay.
PCR
## Reaction set up – Master mixes

<table>
<thead>
<tr>
<th>Component</th>
<th>Volume</th>
<th>Final Concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>10X PCR buffer minus Mg&lt;sup&gt;2+&lt;/sup&gt;</td>
<td>5 µL</td>
<td>1x</td>
</tr>
<tr>
<td>10 mM dNTP mixture</td>
<td>1 µL</td>
<td>0.2 mM each</td>
</tr>
<tr>
<td>50 mM MgCl&lt;sub&gt;2&lt;/sub&gt;</td>
<td>1.5 µL</td>
<td>1.5mM</td>
</tr>
<tr>
<td>Primer mix (10 µM each)</td>
<td>2.5 µL</td>
<td>0.5mM each</td>
</tr>
<tr>
<td>Template DNA</td>
<td>0.5-10 µL</td>
<td></td>
</tr>
<tr>
<td><strong>Taq DNA Polymerase</strong></td>
<td>0.2-0.5 µL</td>
<td>1-2.5 units</td>
</tr>
<tr>
<td>Autoclaved distilled water</td>
<td>To 50 µL</td>
<td></td>
</tr>
</tbody>
</table>

- Pipetting 0.2 µL accurately is very difficult
- Use master mixes
- Use separate rooms
Key considerations for PCR

**Reaction mixture**
- Template
- Buffer
- Enzyme
- Nucleotides
- Primers

**Reaction conditions**
- Denaturation temperature
- Annealing temperature
- Extension temperature
- Extension time
- Number of cycles
Reaction conditions

- Denaturation temperature
  - Separate double stranded DNA
  - 95 °C
- Annealing temperature
  - Determine Tm of primers
    - Deduct 5 °C to determine annealing temperature
  - Varies for each primer pair typically around 55 °C
- Extension temperature
  - Determined by enzyme
    - Typically between 68 °C and 72°C
- Extension time
  - Determined by enzyme and product length
    - 1000 bases per 10 sec at optimal conditions
    - 45-90 secs
- Number of cycles
  - ~25-35 over 40 and non-specific products may be identified
Analysis
In-house PCR assays
Why design your own assays?

**For**
- Cheaper
- More control
- Not a black box procedure
- Can develop assay for any target of interest

**Against**
- Extensive QC required
- Registration
- Time consuming
PCR assay design considerations I

- What are you trying to do?
  - Amplify a single gene (virus, cancer, tissue etc)
  - Amplify all related sequence (virus family, bacteria species)
  - Introduce mutations into sequence
  - Track contamination in the lab
PCR assay design considerations II

- What information do you currently have?
  - Do you have a complete sequence available
  - Do you have multiple sequences from different samples
  - Do you have a partial sequence from a related organism

- What information can you reasonably get?
PCR assay design considerations III

- How often will you run this assay?
- How important is this assay?
  - What are the consequences if you get it wrong?
- What type of specimen will you be applying this test to?
- Will you always be receiving the same type of sample?
  - Will it be a pure culture isolation, or direct from a clinical sample?
- Will it be a conventional PCR assay or a real-time PCR assay
- How fast do you need the results?
Assay design (primer design)

- From your target sequence, identify a short nucleotide sequence (~20 bp),
How to design good primers

- Length
- Region to be amplified
- Tm matching
- Self similarity
- Target similarity
Primer length

- Ideally primers are ~20 nt
- Adjust length of primers to modify the Tm of the primer
- Doesn’t need to be exactly 20, can be 18 can be longer.
Region to be amplified

- Length of product will be largely influenced by type of assay
  - Real-time PCR
    - Short products are better 100-200 bp
  - End point PCR
    - <2000 bp
- Also determined by good primer sites
Tm matching

- Want primers to bind at the same temperature
  - Annealing temperature determined by primer Tm.
  - Match within 1-2 degrees
  - Add extra bases to improve similarity
  - Only include region which will be binding to template (overhanging regions do not affect annealing temperature)

- Annealing temperatures Tm – 7-9 °C
  - Use 7 °C for better specificity
  - Use 9 °C for better chance of reaction working
Minimise self similarity

• Risks
  • Primer dimer formation
    • Primer binding to itself or the other primer – rather than the template
  • Secondary structure
    • Stem loop formation
Target similarity

- Want to have primers match the target sequence exactly
- Want the primers to have no similarity with other regions of the genome.
- Want to minimise potential for the primer to cross react with potential contaminants in the sample
- Single base pairs mismatches can prevent primers from effectively binding
GC content

- At the 3’ end G or C bases improves the strength of binding
- Strings of G’s and C’s disrupt stable base pairing
Other considerations

- Minimise strings of consecutive bases
- Minimise secondary structure
- Tm around 56 °C to 72 °C
Design tools

- NCBI
- Primer 3
- Who ever you order your primers from
NCBI tools

- **How to: Design PCR primers and check them for specificity**
- **Starting with ...**
- **ONE OR MORE PRIMER SEQUENCES**
  - Go to the [Primer BLAST](http://www.ncbi.nlm.nih.gov/guide/howto/design-pcr-primers/) submission form.
  - Enter one or both primer sequences in the Primer Parameters section of the form. If only one primer is available, a template sequence is also required. See "A Target Template Sequence..." below.
  - In the Primer Pair Specificity Checking Parameters section, select the appropriate source Organism and the smallest Database that is likely to contain the target sequence. These settings give the most precise results. For broadest coverage, choose the nr database and do not specify an organism.
  - Click the "Get Primers" button to submit the search and retrieve template and specificity information.
- **A TARGET TEMPLATE SEQUENCE OR ACCESSION NUMBER**
  - Go to the [Primer BLAST](http://www.ncbi.nlm.nih.gov/guide/howto/design-pcr-primers/) submission form.
  - Enter the target sequence in FASTA format or an accession number of an NCBI nucleotide sequence in the PCR Template section of the form. If the NCBI mRNA reference sequence accession number is used, the tool will automatically design primers that are specific to that splice variant.
  - If one or both primer sequences are to be used in the search, enter these in the Primer Parameters section of the form. Primer BLAST performs only a specificity check when a target template and both primers are provided.
  - In the Primer Pair Specificity Checking Parameters section, select the appropriate source Organism and the smallest Database that is likely to contain the target sequence. These settings give the most precise results. For broadest coverage, choose the nr database and do not specify an organism.
  - Click the "Get Primers" button to submit the search and retrieve specific primer pairs.
NCBI


- Ebola virus genome [NC_002549.1](https://www.ncbi.nlm.nih.gov/nuccore/NC_002549.1) (find the reference number using genome or gene search)
### Primer-BlAST: Finding primers specific to your PCR template (using Primer3 and BLAST)

#### PCR Template
- **Enter accession, gi, or FASTA sequence** (Refseq record is preferred)

Or, upload FASTA file

#### Primer Parameters
- **Use my own forward primer** (5'-3' on plus strand)
- **Use my own reverse primer** (5'-3' on minus strand)
- **PCR product size**
- **# of primers to return**
- **Primer melting temperatures (°C)**

#### Exon/intron selection
- **Exon junction span**
- **Exon junction match**
- **Intron Inclusion**
- **Intron length range**

#### Primer Pair Specificity Checking Parameters
- **Specificity check**
- **Search mode**
- **Database**
- **Organism**
- **Exclusion (optional)**
- **Entrez query (optional)**
- **Primer pair similarity calculation**
Cross-species reactivity
Trouble shooting PCRs

- Touchdown
- Hot start
- Modify annealing temperature
- Change magnesium concentration
- Try a more sensitive detection method
- Add cycles
- Denature the template for longer or at higher temperature
- Try adding enhancing agents
Trouble shooting

- Smear or multiple products
  - Increase annealing temperature
  - Decrease total number of cycles
  - Vary component concentration
Commercial PCR assays
Sequencing technologies
Common technologies

- Traditional
  - Sanger (dideoxy) sequencing
- Sample prep
  - Fluidgm
- Next generation sequencing
  - Pyrosequencing
  - Illumina
  - SOLiD
  - Ion Torrent
Cost per human genome

Cost per Genome

Moore's Law

NIH National Human Genome Research Institute
genome.gov/sequencingcosts
<table>
<thead>
<tr>
<th>Sequencing Chemistry</th>
<th>Ion Torrent</th>
<th>454 Sequencing Qiagen</th>
<th>Illumina</th>
<th>SOLiD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ion semiconductor sequencing</td>
<td>Ion semiconductor sequencing</td>
<td>Pyrosequencing</td>
<td>Polymerase-based sequence-by-synthesis</td>
<td>Ligation-based sequencing</td>
</tr>
<tr>
<td>Amplification approach</td>
<td>Emulsion PCR</td>
<td>Emulsion PCR</td>
<td>Bridge amplification</td>
<td>Emulsion PCR</td>
</tr>
<tr>
<td>Mb per run</td>
<td>100</td>
<td>100</td>
<td>600,000</td>
<td>170,000</td>
</tr>
<tr>
<td>Time per run</td>
<td>1.5 hours</td>
<td>7 hours</td>
<td>9 days</td>
<td>9 days</td>
</tr>
<tr>
<td>Read length</td>
<td>200 bp</td>
<td>400 bp</td>
<td>2x150 bp</td>
<td>35x75 bp</td>
</tr>
<tr>
<td>Cost per run</td>
<td>$ 350 USD</td>
<td>$ 8,438 USD</td>
<td>$ 20,000 USD</td>
<td>$ 4,000 USD</td>
</tr>
<tr>
<td>Cost per Mb</td>
<td>$ 5.00 USD</td>
<td>$ 84.39 USD</td>
<td>$ 0.03 USD</td>
<td>$ 0.04 USD</td>
</tr>
<tr>
<td>Cost per instrument</td>
<td>$ 50,000 USD</td>
<td>$ 500,000 USD</td>
<td>$ 600,000 USD</td>
<td>$ 595,000 USD</td>
</tr>
</tbody>
</table>

From wikipedia

The HiSeq X Ten is the world’s first sequencing platform to break the US$1000 barrier for 30x coverage of a human genome. When used at scale, the HiSeq X Ten delivers a US$1000 genome, inclusive of instrument depreciation, sequencing consumables, DNA extraction, library preparation, and estimated labour for a typical high-throughput genomics laboratory.

The HiSeq X Ten platform comprises 10 HiSeq X sequencing instruments. Each instrument is capable of generating up to 1.8 terabases of data every 3 days, which is sufficient to sequence at least 15 whole human genomes at 30x coverage. Combined the HiSeq X Ten platform has capacity to sequence 320 genomes per week or 18,000 genomes per year.

Next Generation Sequence preparation

COMPLETE EXPERIMENTS IN FIVE EASY STEPS

1. Samples and primers are loaded onto the 48.48 Access Array IFC, which is mounted on an SBS-compatible carrier, allowing sample loading with an 8-channel pipette or a liquid dispensing robot.

2. In the pre-PCR IFC Controller AX, samples and primers are automatically combined into 2,304 unique PCR reactions.

3. The IFC is placed in the FC1™ Cycler for target amplification.

4. After PCR, products from each sample are pooled on-chip in the post-PCR IFC Controller AX and pumped out for collection.

5. The 48 amplified and tagged products can be collected using a multi-channel pipette.

Five hands-on steps to generate 48 sequencer-ready libraries
Sanger Sequencing
Good sequence
Illumina

1. PREPARE GENOMIC DNA SAMPLE
   Randomly fragment genomic DNA and ligate adapters to both ends of the fragments.

2. ATTACH DNA TO SURFACE
   Bind single-stranded fragments randomly to the inside surface of the flow cell channels.

3. BRIDGE AMPLIFICATION
   Add unlabeled nucleotides and enzyme to initiate solid-phase bridge amplification.

4. FRAGMENTS BECOME DOUBLE STRANDED
   The enzyme incorporates nucleotides to build double-stranded bridges on the solid-phase substrate.

5. DENATURE THE DOUBLE-STRANDED MOLECULES
   Denaturation leaves single-stranded templates anchored to the substrate.

6. COMPLETE AMPLIFICATION
   Several million dense clusters of double-stranded DNA are generated in each channel of the flow cell.
Illumina

1. **DETERMINE FIRST BASE**
   - First chemistry cycle: to initiate the first sequencing cycle, add all four labeled reversible terminators, primers, and DNA polymerase enzymes to the flow cell.
   - After laser excitation, capture the image of emitted fluorescence from each cluster on the flow cell. Record the identity of the first base for each cluster.

2. **IMAGE FIRST BASE**
   - Second chemistry cycle: to initiate the next sequencing cycle, add all four labeled reversible terminators and enzymes to the flow cell.

3. **DETERMINE SECOND BASE**
   - After laser excitation, collect the image data as before. Record the identity of the second base for each cluster.

4. **IMAGE SECOND CHEMISTRY CYCLE**
   - Repeat cycles of sequencing to determine the sequence of bases in a given fragment, a single base at a time.

5. **SEQUENCE READS OVER MULTIPLE CHEMISTRY CYCLES**
   - Align data, compare to a reference, and identify sequence differences.

6. **ALIGN DATA**
Pyrosequencing

Step 1
A sequencing primer is hybridized to a single-stranded PCR amplicon that serves as a template, and incubated with the enzymes, DNA polymerase, ATP sulfurylase, luciferase, and apyrase as well as the substrates, adenosine 5’ phosphosulfate (APS), and lucifern.

Step 2
The first deoxyribonucleotide triphosphate (dNTP) is added to the reaction. DNA polymerase catalyzes the incorporation of the deoxyribonucleotide triphosphate into the DNA strand, if it is complementary to the base in the template strand. Each incorporation event is accompanied by release of pyrophosphate (PPi) in a quantity equimolar to the amount of incorporated nucleotide.

Step 3
ATP sulfurylase converts PPi to ATP in the presence of adenosine 5’ phosphosulfate (APS). This ATP drives the luciferase-mediated conversion of lucifern to oxylucifern that generates visible light in amounts that are proportional to the amount of ATP. The light produced in the luciferase-catalyzed reaction is detected by a charge coupled device (CCD) chip and seen as a peak in the raw data output (Pyrogram). The height of each peak (light signal) is proportional to the number of nucleotides incorporated.

Step 4
Apyrase, a nucleotide-degrading enzyme, continuously degrades unincorporated nucleotides and ATP. When degradation is complete, another nucleotide is added.

Step 5
Addition of dNTPs is performed sequentially. It should be noted that deoxyadenosine alfa-thio-triphasate (dATP-S) is used as a substitute for the natural deoxyadenosine triphosphate (dATP) since it is efficiently used by the DNA polymerase, but not recognized by the luciferase. As the process continues, the complementary DNA strand is built up and the nucleotide sequence is determined from the signal peaks in the Pyrogram trace.
Schematic representation of the progress of the enzyme reaction in solid-phase pyrosequencing.

ACCTTGAGTACCATCTAGGAGATCCT

Polymerase

dATP

PPi

ATP-sulfurylase

ATP

Luciferase

Light

Intensity

Nucleotide & Polymerase

Wash

Time

Ronaghi M Genome Res. 2001;11:3-11
Schematic representation of the progress of the enzyme reaction in liquid-phase pyrosequencing.

Ronaghi M Genome Res. 2001;11:3-11
SOLiD sequencing
Illumina in Discussions with 100 Potential Clinical Customers for MiSeq, Plans Cancer Diagnostics

April 25, 2012

By Monica Heger

Over half of the orders taken for Illumina’s MiSeq instrument are for clinical customers and the company is in active discussions with more than a hundred potential customers looking to use the instrument for diagnostics, CEO Jay Flatley said this week during a conference call to discuss the company’s first-quarter earnings.

“The clinical market is emerging very rapidly and will be a key story for 2012,” he said. “We believe clinical sequencing is nearing an inflection point.”

While Flatley did not discuss the number of MiSeqs sold in the quarter, he said that the instrument is well-suited for clinical applications, such targeted cancer panels.
Considerations for equipment choice

- Cost
- Function
- Space
- Usage
- Consider outsourcing
Hybridization Assays

Molecular Workshop

AIMS 1914-2014
Past. Present. Future
4-6 September
Sydney

NATIONAL SCIENTIFIC MEETING
Technology timeline

- 1977 Southern blotting
- 1994 cDNA collection
- 1995 gene expression analysis
- 1996 Commercial arrays
- 1997 first whole genome assay (yeast)
- 2004 who genome on one chip
- 2004 first approval for clinical use (USA)
- 2007 $999 direct to customer genotype (23 and me etc)
- 2010 $99 genotype
- 2014 $1000 human genome
Southern blot
Dot Blots

- Used after HCV PCR for genotyping
Arrays
Sample Preparation
Bead arrays

- Bead read using a laser scanning confocal microscope

Figure 2: Direct Hyb Gene Expression Profiling Bead Design

Illumina array-based technologies support a broad array of RNA expression profiling products for a variety of applications.
Bead arrays

**FIGURE 1: INFINIUM II ASSAY PROTOCOL**

1. Make amplified DNA
2. Incubate amplified DNA
3. Fragment amplified DNA
4. Precipitate & resuspend
5. Prepare BeadChip
6. Hybridize samples on BeadChip
7. Extend/Stain samples on BeadChip
8. Image BeadChip
9. Auto-call genotypes and generate reports

**GENOMIC DNA (750 ng)**

- Indicates stain in red channel
- Indicates stain in green channel
Biofilmchip arrays
Types of Arrays

- Comparative genomic hybridisation (CGH)
  - Copy number variation
  - Use BAC/cDNA clones to identify pathogenic regions
  - 1 data point per 10-100MB
- Gene expression profiling
  - Look at mRNA expression levels and how that changes with disease state
  - 1-5 data points per gene
- Exon arrays
  - Look at expression patterns of the expressed genome.
  - Determine alternative splicing
  - 4 data points per exon
- Resequencing arrays
  - 4 data points per base
- SNP detection arrays
  - 23andme – genome scans
First Diagnostic array

- Determines patient response to drugs
- Identify cytochrome p450 genes transcribed
- Approved by the FDA 2004
- China 2012
- Determine response to Tamoxifen

AmpliChip CYP450 Test

Analysis of CYP2D6 and CYP2C19
Comprehensive detection of gene deletions and duplications – for genes, which play a major role in the metabolism of many medicines. Helps clinicians determine the correct treatment dose for therapeutic drugs and avoid toxic side effects associated with CYP2D6 or CYP2C19 gene products.

Features and Benefits

Features
- The first chip-based in vitro diagnosis product, Mark certification and FDA approval
- The most comprehensive CYP2C19 analysis in a single step

Benefits
- Detects up to 33 CYP2D6 and CYP2C19 alleles
- Detects CYP2D6 gene duplication
- AmpliTaq GOLD polymerase protects against cross-contamination
- Proprietary software enables customized reports containing predicted phenotype information

Intended Use
(CE-IVD, 1/2009, Rev. 6.0)
The AmpliChip CYP450 Test performs a rapid, comprehensive analysis of Cytochrome P450 genes and provides a personalized, accurate prediction of drug response.

Registration Status
- CE-IVD
- Japan-IVD
- US-IVD
Diagnostic approvals

China Approves First Microarray Platform For In Vitro Diagnostic Use

By Asian Scientist Newsroom | Tech & Pharma
February 16, 2012

China’s State Food and Drug Administration has given clearance to Affymetrix’s GeneChip® System 3000Dx v.2 for in vitro clinical diagnostic use.

AsianScientist (Feb. 16, 2012) – Affymetrix, Inc. today announced that its GeneChip® System 3000Dx v.2 (GCS 3000Dx v.2) has been approved by China’s State Food and Drug Administration (SFDA) for in vitro diagnostic use.

It represents the first microarray instrument system to be granted SFDA clearance for in vitro diagnostics, and more than 2,000 clinical centers in China will now have access to the platform.
Tissue of origin microarray

- Based on expression profiles, identify original source for cancer
Luminex

- 100 beads of different colour
- Can bind anything to the surface – proteins, antibodies, DNA probes
- Conduct mixed reactions 10 different assays in each well.
- Label sample
Luminex detection
Multiplexed assays

- Assays panels for
  - Allergy testing
  - Autoimmune disease
  - Cystic fibrosis
  - Gastrointestinal pathogen testing
  - HLA testing
  - Infectious disease
  - Pharmacogenomics
  - Respiratory viral testing
Exome sequencing gives cheaper, faster diagnosis in heterogeneous disease, study shows

Science Daily (June 23, 2012) — The first report of the diagnostic use of the technique of exome sequencing, where short sequences of DNA are analysed, shows that it can give good results at low cost, a researcher from The Netherlands will tell the annual conference of the European Society of Human Genetics June 25. The scientists were able to perform a genetic diagnosis in around 20% of 100 cases of patients with intellectual disability (ID) and 50% of the 25 cases of blindness studied. Not only is the exome test cheaper, but results are available more quickly than with Sanger sequencing[1], they say.

Dr. Marcel Nelen, head of the Core Genome Analysis Laboratory of the Genetics Department at Radboud University Medical Centre, Nijmegen, will describe to delegates how he and colleagues at the University examined exome sequences from 262 patients with six heterogeneous diseases: ID, blindness, deafness, movement disorders, cancer, and epilepsy.

[1] Exome sequencing is a type of genetic testing that looks for changes in DNA. It is faster than Sanger sequencing and can provide results more quickly. It is also cheaper, making it a more accessible option for genetic testing.
Exome coverage

From Figure 1: Major exome platforms (Clark et al, Nat. Biotech., 2011)

### Differences in Target Space

First off, a comparison of the declared exome targets for each platform.

A large number of bases (29.45 Mbp), presumably the “meat” of the exome are targeted by all three platforms. Individually, the platforms have 4-28 Mbp of unique target space. **Agilent does better for Ensembl transcripts; Nimblegen has better coverage of miRNAs.** These two platforms share more target space with each other than either did with the Illumina platform. This is primarily because **Illumina goes after untranslated regions (UTRs).**
CGH

- Determine genome aberrations
  - Insertion
  - Deletion
  - Copy number variation

- Cancer
- Epilepsy
- Learning disorders
Analyzing CGH Data

Clones/Chromosomes ⇒
Re-Sequencing arrays

- Infectious diseases
SNP arrays