Laboratory Techniques

Vincenzo Cirigliano PhD

Molecular Genetics
Labco Diagnostics, Barcelona
vincenzo.cirigliano@labco.eu
Traditional technologies

- Lab techniques in Prenatal Diagnosis
  - Rapid Aneuploidy Detection

- Molecular cytogenetics
  - Array platforms

- Introduction to NGS

Pros, cons, things to know…
Cytogenetic Analysis

- Needs culture of fetal cells
- Staining of fixed metaphase chromosomes
- Light Microscopy
- Analysis of number and structure of chromosomes

Karyotype Accuracy

99.4 – 99.8% AF
97.5 – 99.6% CVS
Cytogenetics

• Main Drawbacks

  • Long and highly specialised training
  • Long culture time leads to parental anxiety
  • Low resolution
Fluorescence In Situ Hybridisation

- Uses fluorescence to identify chromosome regions of interest
- Fluorescently labelled locus specific DNA probe
- Higher resolution than Karyotype/Microdeletions
Interphase FISH

- DNA Probes can hybridize to nuclei so...
  - Cell culture is not necessary
- Takes 1-2 days
Interphase FISH

Main Drawbacks

- Skilled Operators
- Labour Intensive
- Sample Amount / Gestational Age
- Needs specific indications
QF-PCR detection of T21
Main Advantages of QF-PCR

- Multiple Loci Analysed/partial imbalances
- Maternal Contamination
- Mosaicism
- Inclusion of Single Gene Defects (CF, HbS)
- Rapid, Objective & Easy to Automate
- Reduced Costs / High Throughput

Standard RAD in Europe since 2001
Cytogenetics

**Chromosome Analysis**
- Trisomy 21
- Deletion 11q

322 genes
46 Mb

Molecular Cytogenetics

**FISH**
- Prader-Willi syndrome
- Deletion 8p

30 genes
4 Mb

**CGH**
- Deletion 12q

50 genes
8 Mb

**Array-CGH**
- Deletion 12q

>15 genes
~ 1.5 Mb
Molecular Cytogenetics: CMA

Advantages:

- Similar to simultaneous FISH experiment with thousands or millions of probes
- Phenotypic information is not necessary as it is a whole-genome analysis
- Identification of both microscopic and submicroscopic unbalances
- Objective, gains or losses easily identified in known genomic locations
- Different platforms available for resolution in density and genomic coverage
Molecular Cytogenetics: CMA

Drawbacks:

- Does not identify balanced rearrangements
- May uncover unwanted information
  - Adult onset condition in a prenatal setting
  - AOH which may indicate consanguinity
- May identify variants of unclear clinical significance
  - Counseling dilemmas and parental anxiety
Array CGH

Patient

Reference
Array CGH

Patient

Reference
Array CGH

Patient

Reference

Chromosome 21
Array CGH

Patient

Reference

Chromosome X
Array CGH

Array Scan 3\(\mu\)m (0.003 mm)
Unbalanced translocation t(4;21) (p16.3; q22.13) resulting in 3Mb deletion of 4p16.3 and 9Mb gain of 21q22.13.
Uncultured Amniotic Fluid

Characterization of mosaic ring marker chromosome (i12p)
Karyotype
Molecular Karyotype
SNP array: Isodisomy for Chromosome 7

Normal 2 copy dosage by aCGH

Complete Homozygosity

A or B

Courtesy Lisa G. Shaffer
## aCGH vs SNP Arrays

<table>
<thead>
<tr>
<th></th>
<th>Array CGH</th>
<th>SNP Array</th>
</tr>
</thead>
<tbody>
<tr>
<td>Copy gains and losses</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Triploidy</td>
<td>Not Always</td>
<td>YES</td>
</tr>
<tr>
<td>Mosaicism</td>
<td>YES</td>
<td>Easier to visualize</td>
</tr>
<tr>
<td>Uniparental Disomy</td>
<td>NO</td>
<td>Isodisomy</td>
</tr>
<tr>
<td>AOH</td>
<td>NO</td>
<td>YES</td>
</tr>
</tbody>
</table>
Arrays: Platforms

Probes Design

BACs
Oligos

Several Different Densities

Number of Probes

1x244K 2x105K 4x44K 8x15K
~244,000 ~105,000 ~44,000 ~15,000

1x1M 2x400K 4x180K 8x60K
~1,000,000 ~400,000 ~180,000 ~60,000

Probes Distribution

Targeted Uniform (Backbone)
Targeted + Backbone
Arrays: Platforms

Probes Design

BACs

Oligos

Number of Probes

Several Different Densities

1x244K 2x105K 4x44K 8x15K

~244,000 2x105,000 ~44,000 ~15,000

1x1M 2x400K 4x180K 8x60K

~1,000,000 ~400,000 ~180,000 ~60,000

Probes Distribution

Targeted Uniform (Backbone)

Targeted + Backbone
<table>
<thead>
<tr>
<th>Study</th>
<th>Clinically relevant CNVs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fiorentino (2011)</td>
<td>6.1%</td>
</tr>
<tr>
<td>Ronenfeld/Shaffer (2012)</td>
<td>6.6%</td>
</tr>
<tr>
<td>Schwartz (2011)</td>
<td>5.7%</td>
</tr>
<tr>
<td>NIHD Study (2012)</td>
<td>6%</td>
</tr>
</tbody>
</table>

Add clinically relevant information if karyotype is normal
• Abnormal Result

• CNV identified as related to a particular phenotype
• Alteration contains dosage-sensitive, disease-causing genes
• Alteration occurs within a region of the genome known to be involved in chromosomal syndromes
• Alterations are statistically enriched in patient populations as compared to controls
Arrays Results Interpretation

• Benign Result

• CNV is not thought to cause an abnormal phenotype
• Alteration is found in control populations in equal frequencies
• May be ethnic-specific or found widely in most populations
Arrays Results Interpretation

- **Variants Of Uncertain Significance**

  - CNV not related with abnormal phenotype but large enough to be of concern
  - Alteration has not been reported in the medical literature, or not found in available databases
  - Alteration does not contain known disease-causing genes
  - Additional testing, or further evaluation of the family history may be advised
  - Such findings can be revisited over time as new studies are published
Increasing resolution to the nucleotide level

Whole Exome Sequencing (WES)
Whole Genome Sequencing (WGS)
Diminishing Cost of Sequencing

**Moore’s Law:** the number of transistors in a dense integrated circuit has doubled approximately every two years.
## NGS Platforms

<table>
<thead>
<tr>
<th>Platform</th>
<th>Chemistry</th>
<th>Read Length</th>
<th>Run Time</th>
<th>Gb/Run</th>
<th>Strenght</th>
</tr>
</thead>
<tbody>
<tr>
<td>454 GS Junior (Roche)</td>
<td>Pyro-sequencing</td>
<td>500</td>
<td>8 hrs.</td>
<td>0.04</td>
<td>Long Read Length</td>
</tr>
<tr>
<td>MiSeq (Illumina)</td>
<td>Reversible Terminator</td>
<td>2x300</td>
<td>27 hrs.</td>
<td>15</td>
<td>Long Read Length</td>
</tr>
<tr>
<td>HiSeq (Illumina)</td>
<td>Reversible Terminator</td>
<td>2x250</td>
<td>2 days (rapid mode)</td>
<td>300 (rapid mode)</td>
<td>High-throughput / cost</td>
</tr>
<tr>
<td>Ion Proton (Life)</td>
<td>Proton Detection</td>
<td>200</td>
<td>4 hrs.</td>
<td>10</td>
<td>Short Run times</td>
</tr>
<tr>
<td>PacBio RS</td>
<td>Real-time Sequencing</td>
<td>3000 (up to 15,000)</td>
<td>20 min</td>
<td>3</td>
<td>No PCR, Longest Read Length</td>
</tr>
</tbody>
</table>

Note: Some figures might be out of date
Common steps in NGS

Massively Parallel Sequencing

• Sanger-sequencing: DNA synthesis and detection are two separate steps (slow and low throughput).

• NGS: synthesis and detection performed at the same time (sequencing by synthesis)

• Multiple simultaneous reactions (Massively Parallel Sequencing).

• 3rd Generation. Single DNA molecule sequencing, no DNA synthesis increasing sequencing reads
Library Preparation

• DNA randomly fragmented, platform-specific adaptors added at the ends to produce a “library”.

• Library is amplified through PCR (Platform-specific amplification e.g. beads or glass).

• Desynchronization of reads in sequencing / detection cycle main cause of shorter reads and errors.

• Amplification Introduces Bias:
  • Amplification bias against GC rich regions.
  • Alteration of representational abundances (duplicates).

Experimental design
DNA Cleaved into Small Fragments
Adapters Linked to DNA Fragments
Multiplexing by Indexing

Patient 1

Patient 2
Illumina Flow Cell Design
Cluster Generation: Bridge Amplification

Single DNA libraries are hybridized to primer lawn

Bound libraries are then extended by polymerases

Surface of flow cell coated with a lawn of oligo pairs

3’ extension

Adapter sequence
Single-stranded molecule flips over and forms a bridge by hybridizing to adjacent, complementary primer

Hybridized primer is extended by polymerases
Double-stranded bridge is formed
Double-stranded bridge is denatured

Result:
Two copies of covalently bound single-stranded templates
Cluster Generation: Bridge Amplification

Process cyclically repeated
Reverse strands are cleaved and washed away, leaving a cluster with forward strands only ready for SBS
Illumina Cluster Generation

thousands of copies of the same DNA strand
Illumina Sequencing by Synthesis
Illumina Sequencing by Synthesis
Illumina Sequencing by Synthesis
Simultaneous Capture of Millions Clusters
Simultaneous Reading of Clusters
Paired-end Sequencing
Improved alignment

Fragment length

CGCTAGAAG ________________ GAAGTCGCG
ATTCCCGCGATCTTCCCGTTCCGACTGCAGACCTTCAGCGCAGCATATATCGCTAGCATACCGTTATAC

Count:

← Human Genome →

Alignment
Better accuracy
2 x 36 bp

Fragment Size
determination
Paired-end Sequencing
Less NES more useful counts

Improved Allignment, Increased Depth and Resolution
Maternal and fetal cfDNA size distribution
Distribution of cfDNA fragment sizes for NIPT
NGS-Based Diagnostic Tests

- **MPSS based NIPT**
  - Random sequencing of maternal plasma DNA for fetal aneuploidy screening
- **Targeted gene panels**
  - Targeting a specific set of genes related to a defined clinical phenotype
- **Clinical exome**
  - Comprehensive but targeted sequencing of regions within the exome, *known* to be associated with clinical phenotypes
- **Whole exome sequencing (WES)**
  - Sequence the protein-coding regions in a genome (entire exome), both *known* and *unknown* associations to phenotypes
- **Whole genome sequencing (WGS)**
  - Sequence the entire genome including intronic regions and exome
Alignment to Reference Genome
The Exome

• The average human exon is 145 bp long.
• The average gene contains 8.8 exons.

• Exome = 1-2% of the Genome ~50,000,000 bp.
• Contains ~ 85% of disease-causing mutations.
• ~7000 Mendelian disorders known
• While most are individually rare, collectively they account for ~20% of infant mortality
WES Interpretation Workflow

- **Exome**: 60,000 variants
  - Automated Filter by allele frequency
  - Rare Coding Variants: 1,000 variants
    - 1 hour Filter by inheritance & phenotype
  - The Shortlist: 50 variants
    - 14 hours Manual curation & interpretation of each variant
  - The Answer: 1-2 variants

Requires hours of expert time per case
Cost of Interpretation > Cost of Sequencing

RUNNING THE ASSAY AND SEQUENCING THE SAMPLE IS JUST THE START
Gràcies!

Vincenzo Cirigliano PhD

Molecular Genetics
Labco Diagnostics, Barcelona
vincenzo.cirigliano@labco.eu