DNA MICROARRAY

Zalka Anna – Kromat Kft.
Mol.Biol. Üzletágvezető
Committed to you and your future

• We’re here to help accelerate real progress…
Agilent Technologies

Field of focus in measurement systems... from the beginnings of HP

Test and Measurement  
Semiconductor Products  
Life Sciences and Chemical Analysis

Agilent Laboratories
Agilent in Life Sciences

Product and market focus
Agilent’s Life Science & Chemical Analysis Group

- Microfluidics
- Liquid Chromatography
- Mass Spectrometry
- Microarrays

- RNA
- DNA
- Protein
- Cell

- Quadrupole
- Ion Trap
- AP Maldi
- ESI TOF

- cDNA Microarrays
- Oligo Microarrays
- DNA Microarray Scanner
- Informatics

Networked Data Systems - Informatics
Consumables and Services
AGILENT KEY PRODUCTS

HPLC, GC, /MS

Microarray scanner

Bioanalyzer

Real-time PCR, PCR
Background
Biology, Cells and DNA

• All living organisms consist of cells. Humans have trillions of cells; Yeast - one cell.
• Cells are of many different types (blood, skin, nerve), but all arose from a single cell (the fertilized egg).
• Each cell contains a complete copy of the genome (the program for making the organism), encoded in DNA.
• A gene is a segment of DNA that specifies how to make a protein. Human DNA has about 30-35,000 genes; Rice has about 50-60,000, but shorter genes.
• Genomics: investigation of structure and function of very large numbers of genes simultaneously (studying many, many genes at once vs. ‘gene-by-gene’).
A Dynamic View
Cell activity varies with time!

- metabolites
- protein
- mRNA
- forwards-propagated correlations

event

20/11/2008 A.Zalka – Microarray Technology/ELTE
DNA Microarray System
DNA MICROARRAY

• genomic tool to monitor simultaneously thousands of genes (structure or activity) to entire genome on one slide/chip – **SEQUENCE DATABASE**
  • chip contains immobilizing/detection probes (up to 500-1000K spots/chip) – **CHIP MANUFACTURING**
  • probes are artificial oligonucleotides (25-60 bp ss) applied or directly synthetised onto the chip surface – **PROBE MANUFACTURING**
  • immobilizing of sample DNA/RNA fragments is due to the hybridization ability to the matching complementer sequence containing probes – **ARRAY PROCESSING**
  • detection is done by labeling of sample with fluorescent label prior to hybridization – **INSTRUMENTATION**
Microarray is a technological approach

That is why need for:
- standardized, robust design
- sophisticated instrumentation
- calibration, validation
- technology optimization and fine-tuning

Technology is complex and crossdisciplinary (biotechnology, nanotechnology, chemistry, physics, robotics; plus the application areas: molecular biology, oncology, pathology, medicine…)
- needs critical mass of knowledge, experience
- may fail or be mis-operated

Technology and the ways it is applied are still evolving

Computational statistics and data analysis can contribute
DNA Microarrays

Measuring RNA by Complementary Hybridization

A DNA chip is an ARRAY of many different pieces of DNA spotted onto glass.

Each piece (spot) of DNA on the chip can probe for its complement in the sample.
One “spot” on a microarray contains many DNA strands of the same sequence.
Labeled DNA copies of all mRNA from one sample (i.e. from a tumor) is hybridized to array
Shining a laser light at GeneChip causes tagged DNA fragments that hybridized to glow.

Non-hybridized DNA

Hybridized DNA
Microarrays

• Basic Concept
  – Based on Crick-Watson Hybridization

• Different Microarray technologies exist.
  – Probe type (cDNA vs oligo)
  – Spotting vs in-situ synthesis
  – Single vs. dual channel

• Output is a typically an image
  – Sources of errors
  – Image processing is required
  – Images are converted into gene expression matrices for further analysis
MICROARRAY APPLICATIONS

Key Features: Density, Sensitivity, Flexibility

DNA

- aCGH
  - Copy number
    - chromosomal aberrations
    - gene copy number

- CH₃
  - Methylation
    - methylation patterns
    - downstream transcriptional effects

- ChIP
  - Transcription Factors
    - protein/DNA interactions
    - transcription
    - DNA replication
    - DNA repair

- GX
  - mRNA
    - high sensitivity measurements of transcription
    - correlate results with genomic data

RNA

- Splice Variants
  - mRNA isoforms
    - splice forms of specific genes
    - downstream effects on translation

- RNAi
  - RNA interference
    - presence of microRNAs
    - knockout analysis
    - correlate results with transcription data
MICROARRAY HISTORY

• 1975 Southern Blotting technology (Edward Southern)
• 1991 First high-density nylon filter arrays (Lennon, Lehrach)
• 1995 cDNA-Microarrays (Schena et al.)

• 1996 Affymetrix Genechip Technology

• 2002 Agilent open array technology

• 2003 Illumina Bead Array technology
MACROARRAYS: cDNAs or colonies spotted onto nylon slot blots

**Advantage:**
Robust, straightforward protocol
Simple 32P incorporation and hybridization
Filters can be reused
Relatively inexpensive
Four orders of magnitude detection
Radioactivity more linear at low intensities than fluorescence

**Disadvantage:**
Only one sample can be hybridized at a time
Difficult to get accurate ratios-more qualitative
Larger format requires more material to label
Less sensitive than microarrays; 1 in 20,000

Macroarray vs Microarray
Miniaturization, Increased throughput, Two-color comparisons, More reliable ratios, Higher density spotting, Higher sensitivity
Different Array Formats & Technologies

- cDNA spotted microarray
- Nylon membrane
- Illumina Bead Array
- Agilent: Long oligo Ink Jet
- GeneChip Affymetrix
SPOTTED cDNA ARRAYS: contact or non-contact printing

Arrayed Library
(96 or 384-well plates of bacterial glycerol stocks)

PCR amplification
Directly from colonies with SP6-T7 primers in 96-well plates

Consolidate into 384-well plates

Spot as microarray on glass slides

(Ngai Lab, UC Berkeley)
SPOTTING cDNAs

cDNA microarrays
• DNA copies of mRNA (cDNA) arrayed onto glass microscope slides
• ~10,000 cDNAs per slide
• Annotated with sequence
384 well plate
Contains cDNA probes

Glass Slide
Array of bound cDNA probes
4x4 blocks = 16 print-tip groups

Print-tip group 7

Pins collect cDNA from wells

Print-tip group 1

cDNA clones
Spotted in duplicate

Print-tip group 7

Pins collect cDNA from wells
Expression profiling with DNA microarrays

- **cDNA “A”**: Cy5 labeled
- **cDNA “B”**: Cy3 labeled

**Hybridization**

**Scanning**

**Analysis**

**Image Capture**
titration control ->

< induced

repRESSED ->

< unchanged
Spotted chips

Advantages:
- Spot any size product-70 bp-100 kb
- Hybridize two samples at a time: all data is produced as relative ratio.
- 1:100,000 sensitivity (except Agilent)
- Flexible, easy to generate re-arrays and incorporate new products/create custom chips
- Relatively inexpensive to obtain chips

Disadvantages:
- Must generate products for spotting; time-consuming and expensive
- More variability in production; academic and corporate
- Longer front-end learning curve
- No standardization
AFFYMETRIX IN SITU SYNTHETISED ARRAYS

Affymetrix
• Short DNA strands about 25 nucleotides arrayed onto glass surfaces on computer chips
• Allow specific hybridization
• One spot = population of 25 nt strands all the same sequence
• Up to 65,000-400,000 genes per chip.
Affymetrix’s GeneChip
PHOTOLITOGRAPHIC *in-situ* SYNTHESIS

Affymetrix

1. Light is shined through a mask onto a wafer that has initial starting strands where the DNA will be built from.

2. The mask has specific tiny openings that allow the light to come in contact with the wafer at specific sections (in this diagram there are 5 probes only and each could represent a different feature).

3. Any place where light hits, removes a “protective” group from the strands.

4. Free nucleotides (the red T) are washed over the wafer and the nucleotides will combine with any strand that had lost its’ protective group in the previous step.

5. This is then repeated (shine light through a mask, deprotect the strands, add free nucleotides) numerous times until a each strand built is 25 base pairs long.

Basically, the whole process occurs through the controlling of the light to hit the specific features on the chip!
PHOTOLITOGRAPHIC SYNTHESIS

- A little more on the primary process used in the building of the specific probes on the chip – photolithography

- The idea of photolithography is to manipulate light to direct the chemical synthesis of DNA strands which are the probes

- Light is shined through a mask, which has tiny openings that allow light through at specific spaces

- These specific points are features (represented by the squares in the diagram)

- Remember, tiny little probes need to be built vertically up from each feature

- Before being hit with light, the chip is “protected” by a light-sensitive protection group at each probe in the feature, so nothing can be added to these probes

- When light hits the probes in the feature, they are “deprotected” by the removal of the light-sensitive protection group

- This allows the chemical synthesis used in the next step to build the nucleotide chain (probe) on those unprotected DNA chains
A TYPICAL MASK

• Here is an example of a mask
• From a distance, it looks like tiny opening (each black open space)
• Actually, when you look at it up close, each black opening represents where a chip will be made
• Within each opening is actually a pattern of smaller openings that will be used to manipulate the light when building the chip (almost like masks within a mask)
SUBSTRATE CLEANING AND COATING

These pictures show some of the early preparations to the wafer for chip manufacturing:

- The substrate (basically a glass wafer) is cleaned, then coated with polymide.
- The polymide helps with the scanning of the final chip as it reduces refraction and background noise that makes the chips hard to read during the scanning.
- This slide mainly illustrates the cleaning process.
SUBSTRATE CLEANING AND COATING

- More coating of the chip
- If you look closely, you can see the yellow polyimide being applied to the glass wafer in the 4 top pictures
- This is applied to one side of the wafer (and the probes will be built on the other side)
- This is then spun evenly among the wafer, which then takes on the yellow color
- The technician is dipping the wafers in a separate chemical to make the coating permanent
CHEMICAL SYNTHESIS STATION

• The next step is to begin the chemical synthesis of the probes
• Each time this is done, the wafer is placed vertically into a rectangular station which is flooded with the chemical needed at that step
• You can see in the 4 pictures at the bottom that the liquid is coming in and slowly filling up the compartment
• The top left shows the full synthesis station and top right two pictures show the wafer being placed into the synthesis compartment
Once the chips have been completed, it is time to turn to the next step in the chip manufacturing process – chip packaging.

First, the chips must be cut out of the glass wafer.

This is done with this machine which uses diamond blades to make precise cuts.
CHIP PACKAGING

• Next, the separated chips are moved to a machine that will carefully place them into a small cartridge

• The movement of the chip uses a small robotic arm to take the chip out of the tray and place it into the small cartridge.
CARTRIDGE ASSEMBLY

- Here the chip is placed into one half of the cartridge
- The top portion of the cartridge is then placed on top and sealed, to hold the chip in place
- The cartridge is then labeled (it has a bar code on it to help with organization of the chips when doing later experiments with them)
- The cartridges are then placed into individual packages to keep them protected from the outside elements and increase their “shelf-life”
- They are ready to be shipped!
Synthesis by photolithography
COMPLETED SYNTHESIS PRODUCT
Illumina Bead Chip

Infinium II - Two-color assay

12 Sections
>890,000 features per section
Average 30 fold redundancy
STRUCTURE OF ILLUMINA BEADARRAY

Each silica bead is 3 microns in diameter.

700,000 copies of same probe sequence are covalently attached to each bead for hybridisation & decoding.
ILLUMINA ARRAY FORMAT

Illumina multi-sample array formats
ILLUMINA BEADCHIP- EXPRESSION ARRAYS

• RefSeq BeadChip
  • 8 or 12 samples/slide
  • >22,000 probes
  • Human, Mouse, Rat

• Whole Genome BeadChip
  • 6 samples/slide
  • 48,000 probes
  • Human, Mouse

• 12 samples/slide
• >48K probes
AGILENT OPEN GENOMICS ARRAY
AGILENT in situ SYNTHESIS
Flexibility of Ink Jet Printing

60-mers Offer vs 25-mers: 5-8 fold sensitivity improvement, Lower Limit of Detection of 0.001 pM (5x), Broad hybridization area & Probe specificity in situ synthesis: unparallel flexibility with array design

~240K probes/slide
AGILENT Ink Jet Printing- SuRePrint Technology

- No contact with glass
- Precision printing/volume control
- Uniform feature size/shape
- Increased pixel statistics
- QC and monitoring process
In Situ Synthesized Oligonucleotide Arrays

Precision Jetting Enables High Levels of Spatial Multiplexing and Flexible Designs

~ 244,000 Features on 1”x3” Slide
High Density and Multi-pack Arrays

- **High Density**
  - 244K features/Array

- **Medium Density**
  - 2 x 105K
  - 4 X 44K

- **Low Density**
  - 8 x 15K
Agilent Arrays, Gene Content

Broad genome coverage

Current, accurate, comprehensive genome representation from leading public sources to map out novel genes and pathways

Focused multi-array format

A focused subset of genes, printed multiple times on the same slide for increased throughput in research on well defined systems to minimize costs
eArray
Delivering Custom Content and Formats

Name Your Design
Select Array Format
MICROARRAY EXPERIMENT

- Biological question
- Experimental design
- Microarray experiment
- Image analysis
- Normalization

Analysis
- Estimation
- Testing
- Clustering
- Discrimination

Biological verification and interpretation
Two Color Approach

One Color Approach

One Color vs. Two Color

Basic Approaches
MICROARRAY WORKFLOW (GX)

- **EXPERIMENT DESIGN**
  - (Agilent array, Agilent custom array, standard slide)

- **NUCLEIC ACID ISOLATION**
  - (Agilent, Qiagen, etc kits)

- **RNA INTEGRITY CONTROL**
  - (Agilent 2100 Bioanalyzer)

- **DIRECT LABELING (CY3, CY5)**

- **LINEAR AMPLIFICATION, LABELING (CY3, CY5)**
  - (Agilent Linear Amplification kit)

- **HYBRIDIZATION, SSPE WASH**
  - (Agilent Hybridization Oven)

- **MICROARRAY SCANNING**
  - (Agilent Microarray Scanner)

- **DATA COLLECTION, NORMALIZATION, QUALITY CONTROL OF MICROARRAY DATA — transportable raw data**
  - (Agilent Feature Extraction Software - internal)
Sample preparation
RNA 6000 Nano LabChip kit
Analysis of Total RNA Integrity

High quality total RNA

Typical first QC step during cDNA or cRNA sample prep for microarrays

Partially degraded total RNA

2100 bioanalyzer: single lane gel-like image

2100 bioanalyzer: electropherogram
MICROARRAY HYBRIDIZATION

1. Target loading

2. Assemble hyb.chamber

3. incubation

4. wash
Agilent’s Genomics Platform
Agilent Technologies Microarray Scanner
Dual-mode, open system with dynamic autofocus

- Dyes Supported: Cy3 and Cy5
- Excitation wavelengths:
  - SHG-YAG laser, 532nm
  - HeNe laser 633nm
- Laser Power - 20mW each
- Supported Glass Format:
  25x75mm (1”x3”)slides
- Scan time:
  8 mins per slide
- Scan Window Max: 71mm x 21.6mm
- Dynamic Range: 10^4 (16-bit)
- Pixel size: 5µm; 10 µm
- Approximate dimensions (h x w x d):
  610 x 914 x 610mm (24” x 36” x 24”)

Automation: 48 DNA Microarray carousel
Fastest scans: 7 minutes for 2 dyes in 20mm x 60mm
Scanner Workflow
Automated Feature Extraction

“Dead Reckoning”

“Cookie Cutter” Method
Bright spots near the injection port on the segmented chambers

• Caused by wash fluid remaining stuck after the drying step. Visible on ~30% of arrays tested.

Thoroughly clean the hybridization chamber including the injection port
Check the nitrogen pressure
Bright Spots affect on data

- Feat Num 65268
- BG Pop OL, BG Nonunif OL
Uniformity Test Slide

- **Agilent**
- **Brand Y**
- **Brand Z**

- **Slide Defect**

-10 %

+10 %
Agilent’s Genomics Platform
MICROARRAY DATA ANALYSIS

- Scan the arrays
- Quantities each spot
- Subtract background
- Filter out bad data
- Normalization
- Data presentation and data mining
Feature Extraction Data Workflow

Agilent and non-Agilent microarrays scanned on Agilent Scanner

16-bit Tiff Image (uncompressed)

Agilent Array Application Specific Analytics Software

Import

QC Report
Shape
Text
JPEG
MAGE-ML
GEML

Feature Extraction Result Files

Feature Extraction Software

Grid and Measure Spots
Reject Outlier Pixels
Subtract Background
Correct Dye Biases

Agilent Technologies © Agilent Technologies All rights reserved

Agilent G4175AA
GeneSpring ...... Software
www.agilent.com/onmrs/anaalystios

Feature Extraction Data Workflow

Grid Template
Grid File
FE Protocol

16-bit Tiff Image (uncompressed)

Agilent and non-Agilent microarrays scanned on Agilent Scanner

Feature Extraction Software

Grid and Measure Spots
Reject Outlier Pixels
Subtract Background
Correct Dye Biases

Agilent Technologies © Agilent Technologies All rights reserved

Agilent G2567AA
Feature Extraction Software
www.agilent.com
DNA Microarray Applications

Target: DNA

- aCGH: Copy number
  - chromosomal aberrations
  - gene copy number

- CH$_3$: Methylation
  - methylation patterns
  - downstream transcriptional effects

- ChIP: Transcription Factors
  - protein/DNA interactions
  - transcription
  - DNA replication
  - DNA repair

RNA

- GX: mRNA
  - high sensitivity measurements of transcription
  - correlate results with genomic data

- Splice Variants: splice forms of specific genes downstream effects on translation

- RNAi: mRNA isoforms
  - presence of microRNAs
  - knockout analysis
  - correlate results with transcription data
GENE EXPRESSION ANALYSIS: PREDICTING CELL ACTIVITY

Idea: measure the amount of mRNA to see which genes are being expressed in (used by) the cell.

Measuring protein might be better, but is currently harder.
One Color vs. Two Color

Basic Approaches

Two Color Approach

One Color Approach

Sample 1

Sample 2

Sample 1

Sample 2

EXAMPLES OF REACTIONS

PAIR OF COMPLEMENTARY BASES

cDNA FROM TREATED CELLS

cDNA FROM UNTREATED CELLS

EXPERIMENTS
From Microarray images to Gene Expression Matrices

- Images
- Spots
- Spot/Image quantitations
- Intermediate data
- Samples
- Final data
- Gene Expression Matrix
- Genes
- Gene expression levels
Other views of the expression data
Find differentially expressed genes

- Choose Condition(s): Organ/Organism part Artery, Treatment type IL1beta

- Fold Difference: 2

- Interactive Update

- 1,837 out of 15,774 genes pass filter
The Demo Human genome window
Querying for Expression Data
Applications on Agilent Scanner

Key Features: Density, Sensitivity, Flexibility

DNA

- aCGH
- CH₃
- ChIP
- GX

RNA

- Splice Variants
- RNAi

DNA

- Copy number
- Methylation
- Transcription Factors

RNA

- mRNA
- mRNA isoforms
- RNA interference

- chromosomal aberrations
- gene copy number
- methylation patterns
- downstream transcriptional effects
- protein/DNA interactions
- transcription
- DNA replication
- DNA repair
- high sensitivity measurements of transcription
- correlate results with genomic data
- splice forms of specific genes downstream effects on translation
- presence of microRNAs
- knockout analysis
- correlate results with transcription data

Methylation

Transcription Factors

20/11/2008  A.Zalka – Microarray Technology/ELTE
Array-Based Comparative Genomic Hybridization

A method to identify and quantify DNA copy number changes across the genome in a single experiment.

- High-resolution
- High-throughput
- Quantitative
- Tremendous flexibility

from T. Ried, NEJM, 2004

High level amplification of v-myc in colon carcinoma cells
# aCGH vs Other Technologies

<table>
<thead>
<tr>
<th></th>
<th>Resolution</th>
<th>Coverage</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>a Cytogenetics</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Karyotyping</td>
<td>&gt; 10 Mb</td>
<td>Complete</td>
</tr>
<tr>
<td>SKY</td>
<td>&gt; 2 Mb</td>
<td>Complete</td>
</tr>
<tr>
<td>Traditional CGH</td>
<td>&gt; 2 Mb (cytoband)</td>
<td>Complete</td>
</tr>
<tr>
<td>FISH (interphase)</td>
<td>≥ 20 Kb</td>
<td>Probe Specific</td>
</tr>
<tr>
<td>FISH (metaphase)</td>
<td>≥ 100 Kb</td>
<td>Probe Specific</td>
</tr>
<tr>
<td><strong>b aCGH</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BAC</td>
<td>100 Kb (Spectral Genomics - 2 Mb)</td>
<td>Complete</td>
</tr>
<tr>
<td>cDNA</td>
<td>2 Kb</td>
<td>Genes Only</td>
</tr>
<tr>
<td>Oligo (60-mer)</td>
<td>0.06 Kb</td>
<td>Complete</td>
</tr>
</tbody>
</table>

Table 1. Coverage and resolution of different cytogenetic and microarray-based CGH techniques.
<table>
<thead>
<tr>
<th>Chromosome</th>
<th>Median Probe Spacing</th>
<th>Total</th>
<th>Exonic</th>
<th>Intron ic</th>
<th>Intergenic</th>
<th>Telomeric-end*</th>
</tr>
</thead>
<tbody>
<tr>
<td>chr1</td>
<td>196 bp</td>
<td>650,560</td>
<td>8,852</td>
<td>368,785</td>
<td>272,923</td>
<td>4,570</td>
</tr>
<tr>
<td>chr2</td>
<td>189 bp</td>
<td>744,040</td>
<td>6,972</td>
<td>372,640</td>
<td>364,428</td>
<td>9,431</td>
</tr>
<tr>
<td>chr3</td>
<td>186 bp</td>
<td>615,371</td>
<td>5,548</td>
<td>341,766</td>
<td>268,057</td>
<td>10,630</td>
</tr>
<tr>
<td>chr4</td>
<td>178 bp</td>
<td>594,597</td>
<td>4,063</td>
<td>264,531</td>
<td>326,003</td>
<td>6,312</td>
</tr>
<tr>
<td>chr5</td>
<td>186 bp</td>
<td>549,536</td>
<td>4,407</td>
<td>246,794</td>
<td>298,335</td>
<td>7,972</td>
</tr>
<tr>
<td>chr6</td>
<td>186 bp</td>
<td>531,411</td>
<td>4,743</td>
<td>266,394</td>
<td>260,274</td>
<td>9,460</td>
</tr>
<tr>
<td>chr7</td>
<td>190 bp</td>
<td>451,565</td>
<td>4,031</td>
<td>253,996</td>
<td>193,538</td>
<td>6,612</td>
</tr>
<tr>
<td>chr8</td>
<td>189 bp</td>
<td>442,376</td>
<td>3,175</td>
<td>212,277</td>
<td>226,924</td>
<td>9,165</td>
</tr>
<tr>
<td>chr9</td>
<td>196 bp</td>
<td>331,744</td>
<td>3,527</td>
<td>170,535</td>
<td>157,682</td>
<td>9,366</td>
</tr>
<tr>
<td>chr10</td>
<td>198 bp</td>
<td>392,123</td>
<td>3,620</td>
<td>219,447</td>
<td>169,056</td>
<td>10,243</td>
</tr>
<tr>
<td>chr11</td>
<td>198 bp</td>
<td>382,414</td>
<td>4,726</td>
<td>204,716</td>
<td>172,972</td>
<td>8,054</td>
</tr>
<tr>
<td>chr12</td>
<td>190 bp</td>
<td>395,120</td>
<td>5,294</td>
<td>210,931</td>
<td>178,895</td>
<td>10,019</td>
</tr>
<tr>
<td>chr13</td>
<td>182 bp</td>
<td>315,908</td>
<td>1,912</td>
<td>131,143</td>
<td>182,853</td>
<td>1,526</td>
</tr>
<tr>
<td>chr14</td>
<td>191 bp</td>
<td>270,052</td>
<td>2,798</td>
<td>134,064</td>
<td>133,200</td>
<td>1,011</td>
</tr>
<tr>
<td>chr15</td>
<td>200 bp</td>
<td>230,082</td>
<td>3,026</td>
<td>135,709</td>
<td>91,347</td>
<td>2,445</td>
</tr>
<tr>
<td>chr16</td>
<td>212 bp</td>
<td>210,380</td>
<td>3,006</td>
<td>113,762</td>
<td>93,612</td>
<td>10,768</td>
</tr>
<tr>
<td>chr17</td>
<td>212 bp</td>
<td>213,144</td>
<td>4,499</td>
<td>132,663</td>
<td>75,982</td>
<td>10,692</td>
</tr>
<tr>
<td>chr18</td>
<td>188 bp</td>
<td>244,742</td>
<td>1,620</td>
<td>113,437</td>
<td>129,685</td>
<td>13,043</td>
</tr>
<tr>
<td>chr19</td>
<td>237 bp</td>
<td>117,160</td>
<td>3,590</td>
<td>77,546</td>
<td>36,024</td>
<td>10,556</td>
</tr>
<tr>
<td>chr20</td>
<td>212 bp</td>
<td>172,267</td>
<td>2,189</td>
<td>90,835</td>
<td>79,243</td>
<td>12,188</td>
</tr>
<tr>
<td>chr21</td>
<td>190 bp</td>
<td>106,161</td>
<td>989</td>
<td>54,588</td>
<td>50,584</td>
<td>5,533</td>
</tr>
<tr>
<td>chr22</td>
<td>234 bp</td>
<td>86,748</td>
<td>1,511</td>
<td>56,331</td>
<td>28,906</td>
<td>5,057</td>
</tr>
<tr>
<td>chrX</td>
<td>190 bp</td>
<td>346,546</td>
<td>2,935</td>
<td>150,457</td>
<td>193,154</td>
<td>593</td>
</tr>
<tr>
<td>chrY</td>
<td>238 bp</td>
<td>18,148</td>
<td>107</td>
<td>4,293</td>
<td>13,748</td>
<td>599</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>8,412,195</td>
<td>87,140</td>
<td>4,327,630</td>
<td>3,997,425</td>
<td>175,845</td>
</tr>
<tr>
<td>Percentage</td>
<td></td>
<td>100%</td>
<td>1%</td>
<td>51%</td>
<td>48%</td>
<td>2%</td>
</tr>
</tbody>
</table>

*Probes located within 3 MB from either chromosome end.
Agilent aCGH Probe Map
eArray - Delivering Custom Content and Formats

- Customized content with desirable resolution

(eArray: web interface)

~ 8 million computationally validated CGH probes (average resolution ~200bp)
Schematic aCGH Protocol

Genomic DNA Isolation from Samples

* Whole Genome Amplification, Cleanup

QC amp on BioAnalyzer **

* Genomic Labeling, Cleanup

QC labeling on BioAnalyzer **

Add blocking reagents (e.g. Cot-1, tRNA, etc.)

Hybridize to aCGH oligo microarray *

Wash, Dry, Scan

Data Analysis* Training & Services *

* New products ** New applications
Experimental array CGH (aCGH)

Reference DNA (green CY3) → Test DNA (red CY5)

Amplification

Deletion

Fluorescent ratios are calculated and normalized:

median log₂ ratio = 0

median log₂ ratio = 0
Two-color aCGH Workflow

1. Prepare cDNA Probe
   - "Normal"
   - Tumor
   - RT / PCR
   - Label with Fluorescent Dyes
   - Combine Equal Amounts

2. Prepare Microarray
   - Hybridize probe to microarray

3. Scan

Microarray Technology

Reference:
- Total Genomic DNA
- DNA-Cy5
- DNA-Cy3

- Ratios:
  - X: log2
  - Y: Chromosome location

Software:
- CGH Analytics software
- Agilent CGH Microarrays

Experimental:
- 50-500 ng
Initial Explorations

**Agilent Human 1A gene expression arrays**
- Designed for Expression Profiling
- 17,086 human genes
- 11,136 probes unique in genome

**HT29 colon cancer cell line**

**BAC array**
Snidjers et al., Nat Genet 2001

Log2 (fluorescence ratio) vs Position along chromosome 8 (MB)
Basic Layout

User Interface Controls

Genome View

Transcript View

Chromosome View

Microarray Feature Data
Human Genome CGH Microarray 44A

Performance validation data from *TGen*
Phi29 amplified DNA from 10 ng total genomic HT29 DNA

Known aberrations on Chr 8 in HT29

CGH Analytics - data analysis and visualization tool
Graphical and Text Aberration Report

HT29 colon cancer cell line
DNA Microarray System
Bioinformatic Tools: GeneSpring sw

Products

www.agilent.com/chem/chem/informatics

Informatics

Genotyping  Location Analysis  CGH  Gene Expression

GS-GT  GS-LOC  GS-CGH  GS GX

20/11/2008  A.Zalka – Microarray Technology/ELTE
Goal: Gain new insights by linking data sets from different applications.