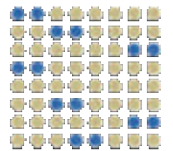


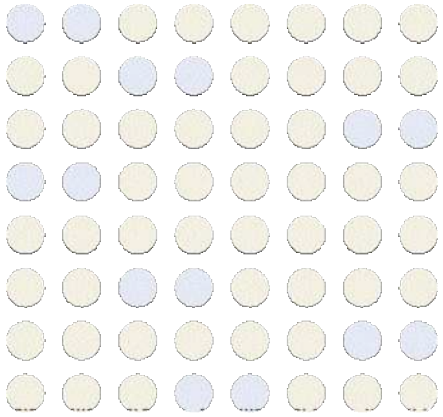
Introduction to Microarray Technology

Neil Winegarden

International Association for Food
Protection Annual Meeting - 2002

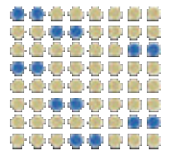


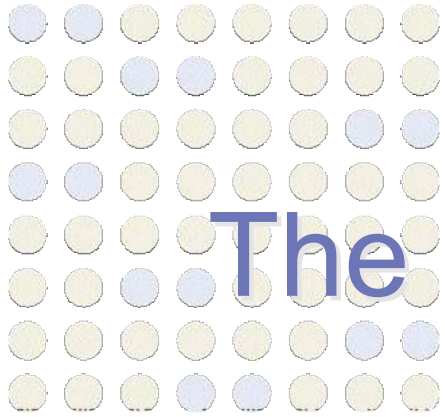
UHN Microarray Centre



Nomenclature 101

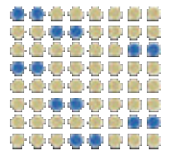
- Microarray: An array of DNA molecules, be it cDNA or oligonucleotides that can be used to study gene-expression and DNA composition.
- GeneChip™: Affymetrix trademarked name for their proprietary type of microarray.
- Target: Used either to describe the material on the array, or the material being hybridized to the array. Leads to much confusion.
- Probe: Same problem as Target.

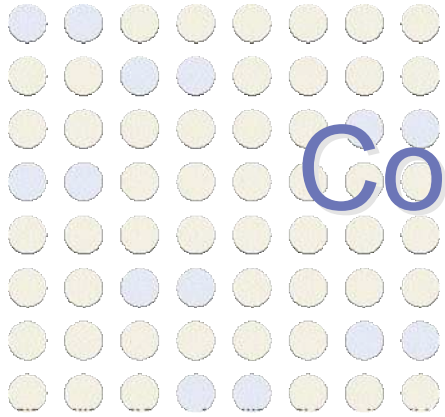




The Importance of Standards

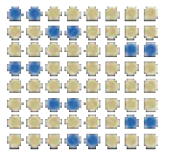
- MGED – Microarray Gene Expression Data Working Group
 - An international group focusing on developing standards for sharing microarray data
- MIAME – Minimal Information About a Microarray Experiment
 - A Standard put forth by MGED to allow for researchers in one group to duplicate work done by another group
 - Standard set of ontologies or terms to describe a microarray experiment

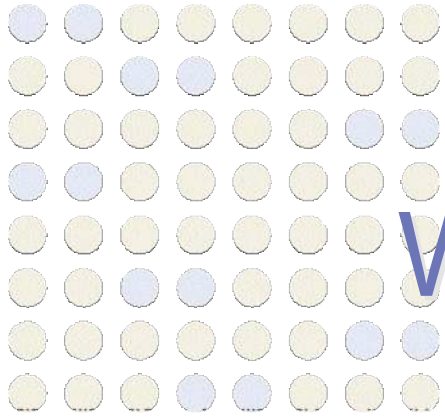




Conventions used in this presentation:

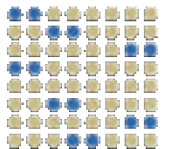
- Target will refer to the material which is hybridized to the array (e.g. the labelled cDNA).
 - Also referred to as **sample**.
- Probe will refer to the material which comprises the array (the arrayed cDNA or oligonucleotides).
 - Also referred to as **features**.

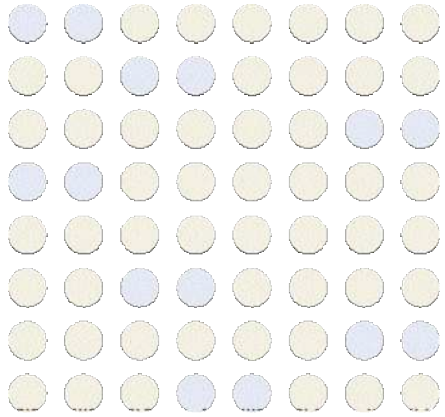




What is a microarray?

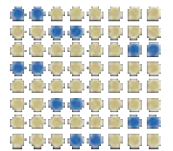
- A microarray is an ordered collection of biological material (typically DNA) printed onto a solid or semi solid substrate.
- Microarrays allow for highly parallel processing of biological analyses.
- Microarrays can allow for parallel analysis of RNA, DNA, proteins, cells, tissues and small molecules.

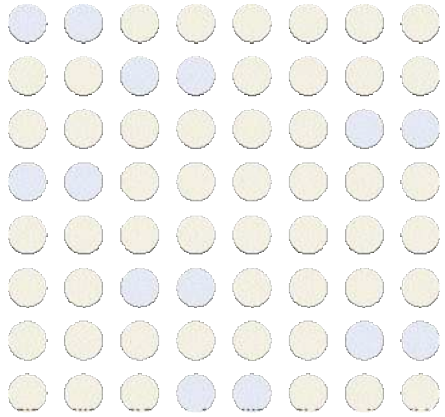




DNA Arrays

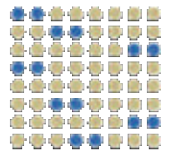
- DNA Arrays are far and away the most common type of array.
- DNA Arrays can be used to analyse either DNA (mutation analysis) or RNA (gene expression).
- DNA arrays may be made from cDNA/ESTs or oligonucleotides.

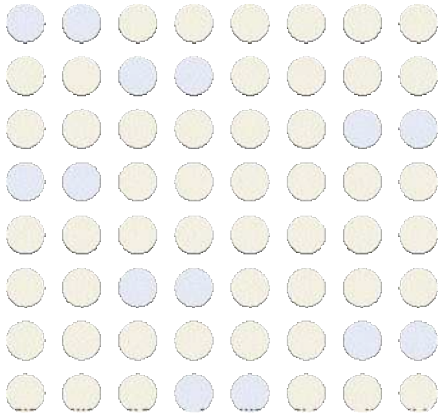




Protein Arrays

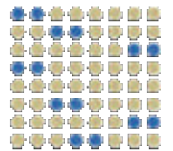
- There are two primary types of protein arrays: antibody arrays or functional protein arrays.
 - Antibody arrays are the protein analogue of DNA arrays. They are used to determine relative abundances of expressed protein in a sample.
 - Functional protein arrays are used to test proteins for a variety of activities. Examples are kinase assays, ligand binding assays etc...
- There are two main challenges:
 - Getting enough material to spot down.
 - Keeping the proteins in a functional state.

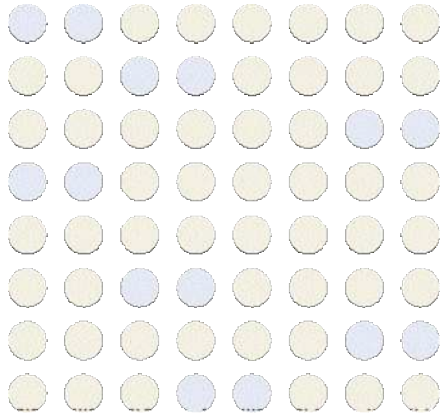




Cell Arrays

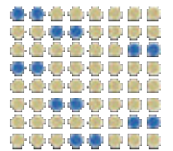
- Cell arrays do not actually involve printing of cells onto the substrate.
 - Rather some material such as DNA is printed onto the slide.
 - Cells are then grown onto the slide, and at each of the discrete locations that the material was spotted, the cells take up that material.
 - The effect on the cells can then be monitored.
 - In the case of DNA, a reporter gene can be added to monitor transfection efficiency.

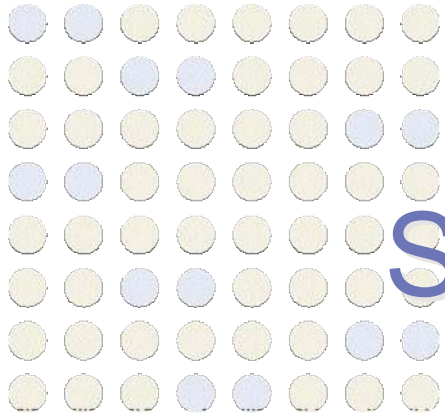




Tissue Arrays

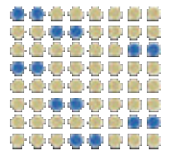
- Tissue Arrays allow for high-throughput pathology
- Dozens to hundreds of tissue samples are arrayed onto slides to allow for high throughput staining
- A pathologist is still required to analyse each sample

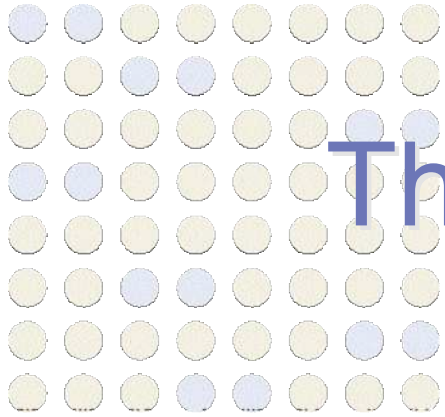




Small Molecule Arrays

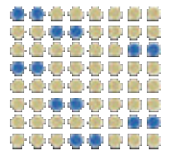
- Small molecules can also be arrayed
- These small molecules can then be probed with a protein of interest to look for potential drug candidates
- This is generally most useful when a particular drug target has already been identified

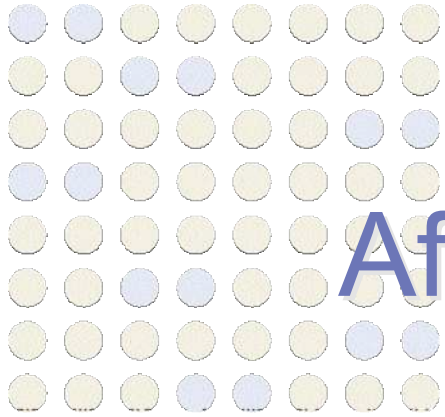




The Many Faces of DNA Microarrays

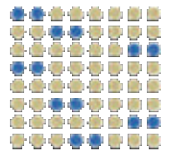
- Several different types of DNA Microarrays exist
 - Affymetrix Arrays (GeneChips™)
 - Spotted Oligonucleotide Arrays
 - Synthesised *in situ*
 - Synthesised *ex situ*
 - Spotted cDNA Arrays
- Spotted arrays may be printed using either contact or non-contact methodologies

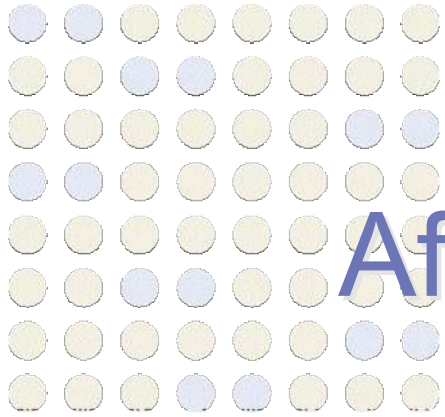




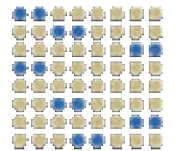
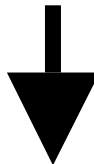
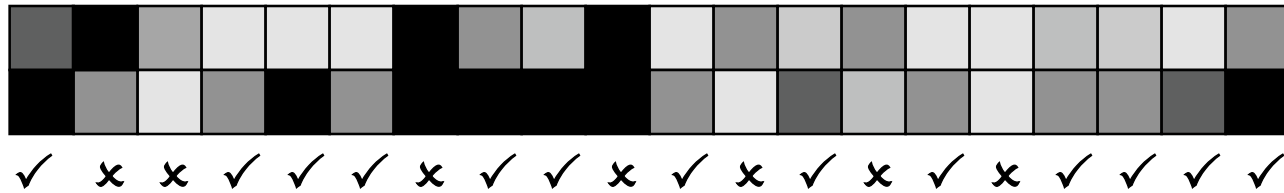
Affymetrix GeneChips™

- Each gene is represented by a set of 20 oligonucleotide pairs (25-mers)
 - The first 20 of these oligos are perfect matches for the gene being interrogated
 - The other 20 oligos have the 13th (middle) base mismatched from the gene being interrogated
- For each pair of oligos, the signal from the perfect match must be higher than the signal from the mismatch to be included

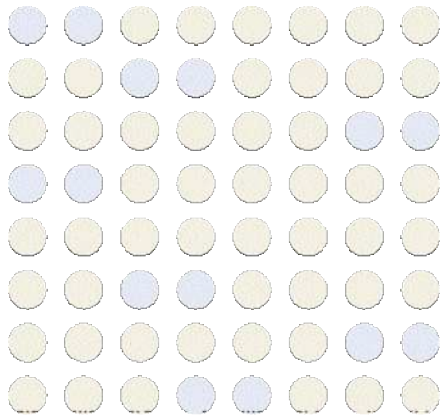




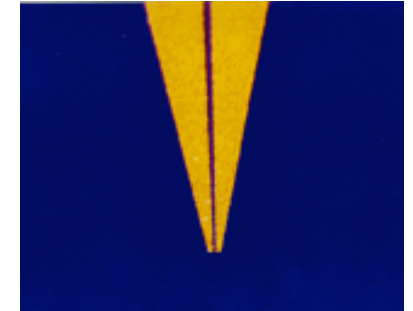
Affymetrix GeneChips™



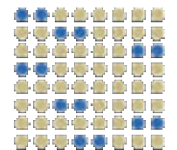
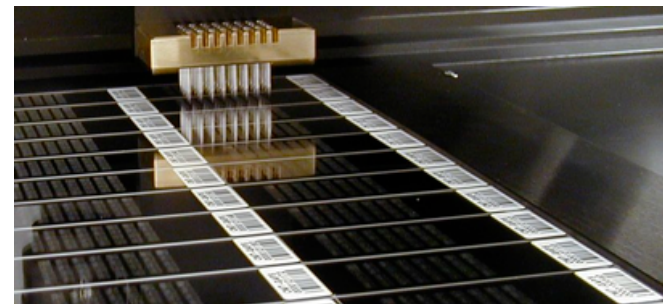
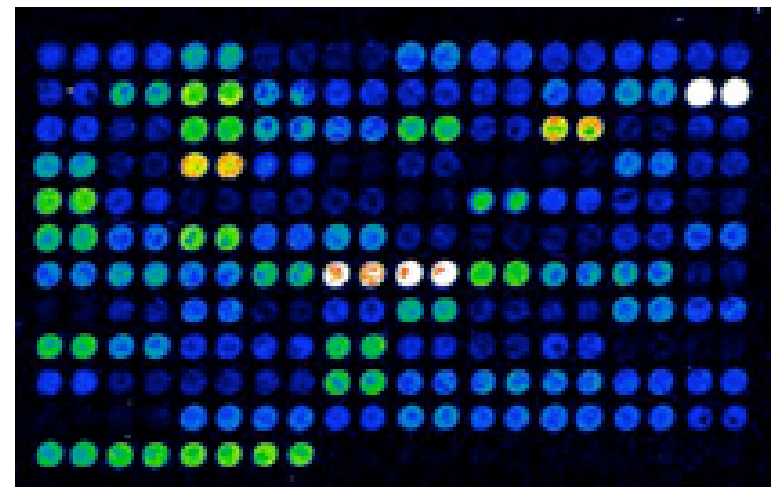
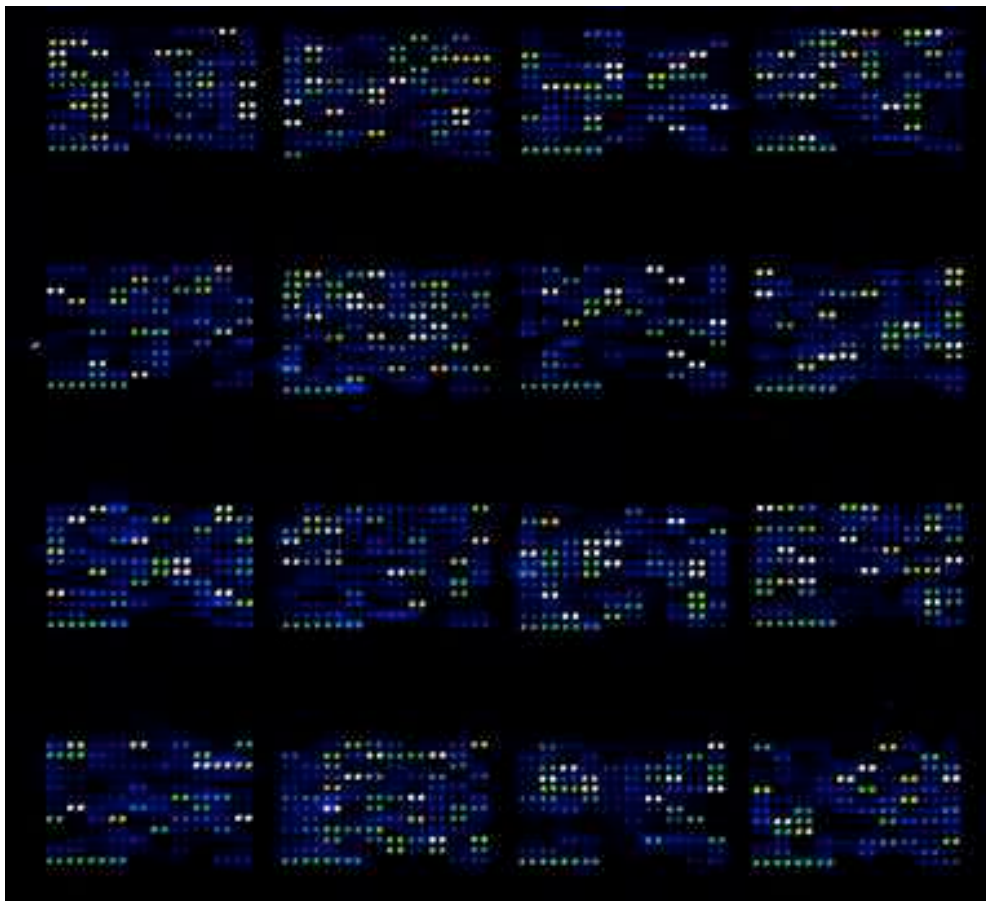
UHN Microarray Centre



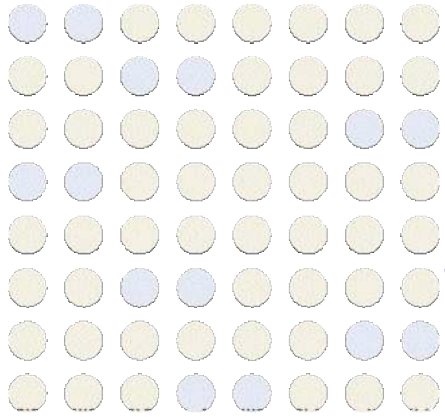
Spotted Arrays



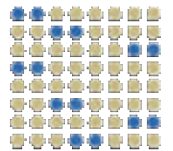
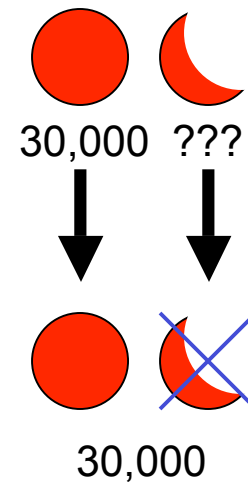
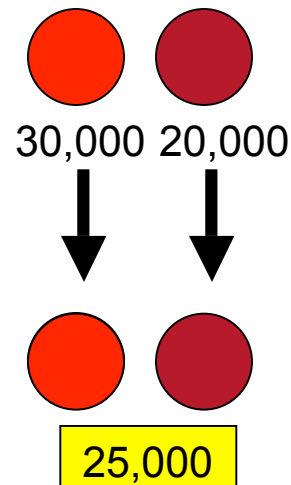
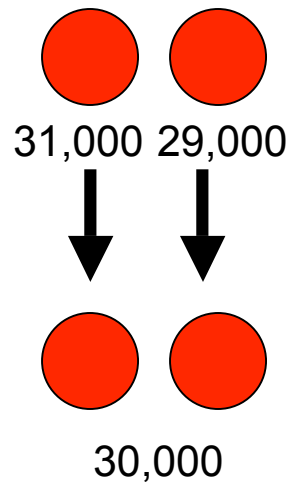
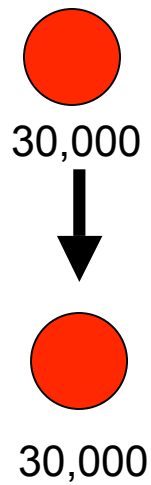
ArrayIt Chipmaker™

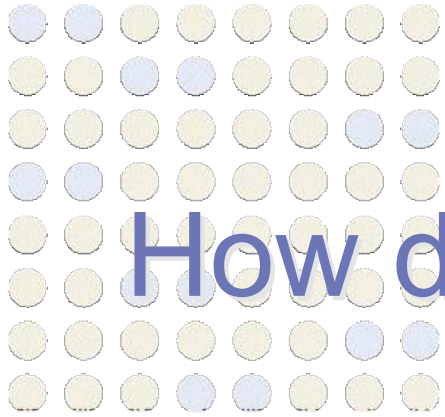


UHN Microarray Centre



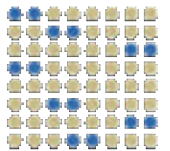
Spotted DNA Arrays





How do DNA microarrays work?

- See www.microarrays.ca/support/tut.html



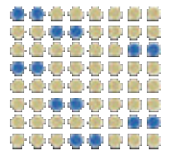
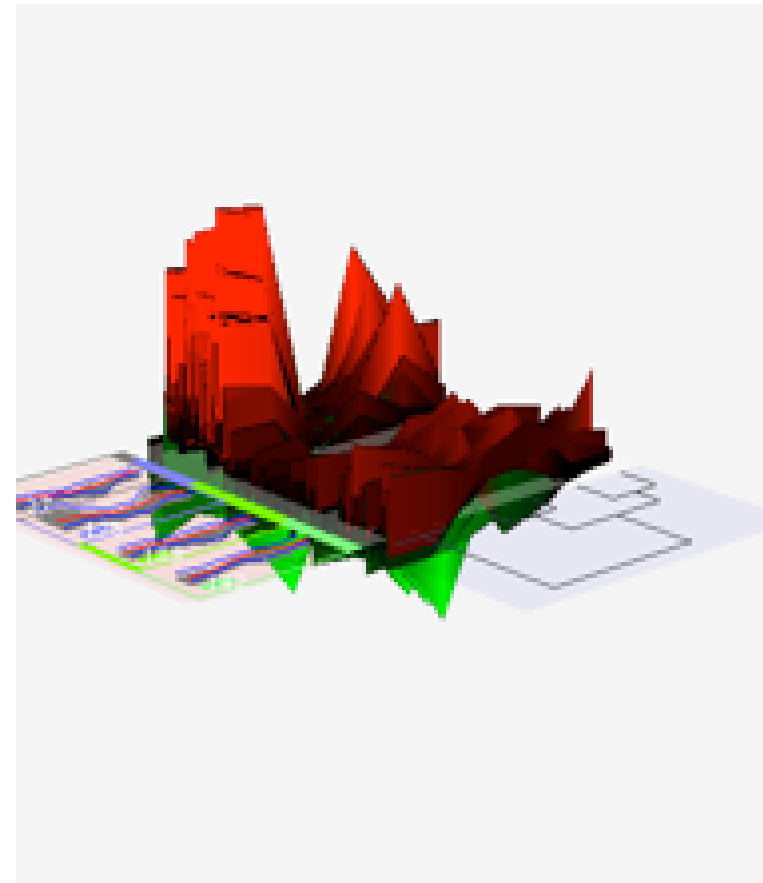
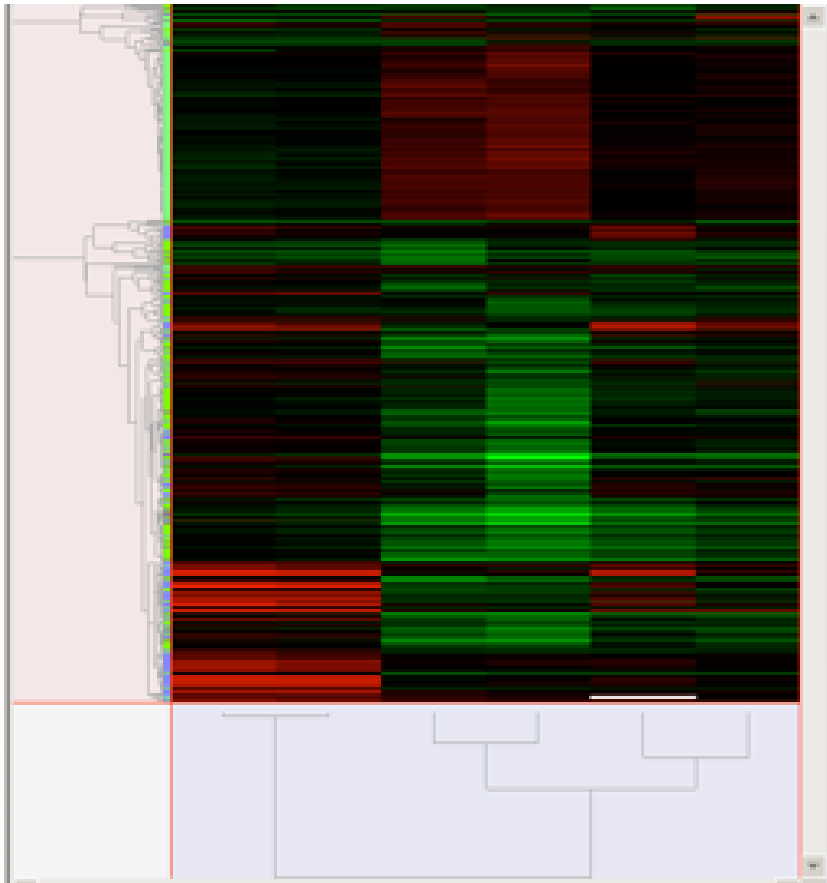
UHN Microarray Centre

All that work for a spreadsheet

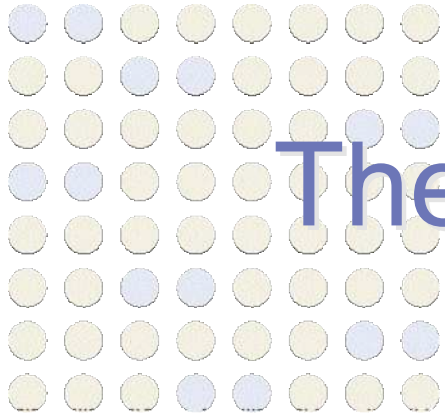
| | A | B | C | D | E | F | G | H | I | J | K | L | M | N | O | P | Q |
|------|-----------|-----------|-----------|-----|--------|--------------|-------|-------|------------|-----------|------------|-----------|------------|-----------|------------|-----------|--------|
| 3870 | Open Data | | | | | | | | | | | | | | | | |
| 3879 | Number | Array Row | Array Col | Row | Column | Name | X Loc | Y Loc | ch1 Intens | ch1 Backg | ch1 Intens | ch1 Backg | ch2 Intens | ch2 Backg | ch2 Intens | ch2 Backg | ignore |
| 3880 | 1 | 1 | 1 | 1 | 1 | 1 H15967 IA | 1180 | 700 | 1920.44 | 443.25 | 2189.7 | 514.542 | 573.935 | 47.0333 | 408.505 | 88.7948 | |
| 3881 | 2 | 1 | 1 | 1 | 2 | 2 H15967 IA | 1350 | 790 | 1958.03 | 455.25 | 1837.85 | 727.768 | 877.032 | 172.25 | 474.16 | 257.833 | |
| 3882 | 3 | 1 | 1 | 1 | 3 | 3 H15318 IA | 1560 | 780 | 2119.65 | 357.417 | 2277.59 | 342.25 | 753.817 | 182.5 | 381.555 | 214.946 | |
| 3883 | 4 | 1 | 1 | 1 | 4 | 4 H15318 IA | 1750 | 800 | 2442.32 | 437.583 | 2256.24 | 368.417 | 777.868 | 138 | 388.098 | 165.205 | |
| 3884 | 5 | 1 | 1 | 1 | 5 | 5 R52868 IA | 1980 | 800 | 5550.03 | 384 | 1563.61 | 382.436 | 3293.84 | 128.887 | 1858.27 | 508.334 | |
| 3885 | 6 | 1 | 1 | 1 | 6 | 6 R52868 IA | 2180 | 800 | 3111.88 | 243.917 | 1493.34 | 210.812 | 2954.24 | 58.25 | 1948.2 | 71.0368 | |
| 3886 | 7 | 1 | 1 | 1 | 7 | 7 R60388 IA | 2350 | 810 | 3322.41 | 293.917 | 1425.7 | 287.984 | 2682.26 | 157.5 | 1338.84 | 203.093 | |
| 3887 | 8 | 1 | 1 | 1 | 8 | 8 R60388 IA | 2550 | 800 | 2657.62 | 160.583 | 1191.26 | 111.252 | 2193.2 | 111.333 | 1209.13 | 149.204 | |
| 3888 | 9 | 1 | 1 | 1 | 9 | 9 H08288 IA | 2760 | 790 | 2056.29 | 295 | 1267.42 | 283.708 | 991.088 | 161.75 | 503.473 | 231.195 | |
| 3889 | 10 | 1 | 1 | 1 | 10 | 10 H08288 IA | 2950 | 800 | 1708.98 | 278.667 | 1035.52 | 241.878 | 896.497 | 53.5833 | 610.031 | 53.3453 | |
| 3890 | 11 | 1 | 1 | 1 | 11 | 11 H16544 IA | 3150 | 800 | 1959.25 | 319.5 | 1200.73 | 238.599 | 1305.89 | 138.917 | 708.268 | 191.739 | |
| 3891 | 12 | 1 | 1 | 1 | 12 | 12 H16544 IA | 3350 | 790 | 2031.07 | 363.917 | 1189.86 | 381.423 | 1338.82 | 161.333 | 774.593 | 239.472 | |
| 3892 | 13 | 1 | 1 | 1 | 13 | 13 T84853 IM | 3560 | 800 | 1578.08 | 281.417 | 597.424 | 147.182 | 1075.58 | 223.083 | 435.034 | 294.82 | |
| 3893 | 14 | 1 | 1 | 1 | 14 | 14 T84853 IM | 3750 | 790 | 1808.54 | 285.667 | 677.131 | 231.88 | 1074.31 | 135.333 | 598.071 | 162.038 | |
| 3894 | 15 | 1 | 1 | 1 | 15 | 15 R14194 IA | 3950 | 790 | 3887.08 | 624.833 | 1495.55 | 794.95 | 2825.76 | 158.5 | 1317.34 | 248.381 | |
| 3895 | 16 | 1 | 1 | 1 | 16 | 16 R14194 IA | 4150 | 790 | 3776.02 | 247.167 | 1555.84 | 238.948 | 2540.88 | 89.5833 | 1339.58 | 82.2985 | |
| 3896 | 17 | 1 | 1 | 1 | 17 | 17 R17499 IA | 4350 | 800 | 4807.7 | 335.333 | 2035.4 | 298.798 | 52445.4 | 149.883 | 18990 | 171.123 | |
| 3897 | 18 | 1 | 1 | 1 | 18 | 18 R17499 IA | 4550 | 800 | 4521.8 | 435.667 | 1431.4 | 353.298 | 48936.7 | 94.8833 | 18825.4 | 167.874 | |
| 3898 | 19 | 1 | 1 | 1 | 19 | 19 R20374 IA | 4760 | 790 | 5532.67 | 483.667 | 1948.01 | 394.788 | 4174.27 | 62.9167 | 2019.27 | 118.78 | |
| 3899 | 20 | 1 | 1 | 1 | 20 | 20 R20374 IA | 4980 | 790 | 4887.88 | 351.167 | 2198.84 | 206.674 | 3940.32 | 88 | 2294.08 | 88.0157 | |
| 3900 | 21 | 1 | 1 | 2 | 1 | 1 R07120 IA | 1170 | 1010 | 2590.11 | 227.917 | 1114.53 | 144.542 | 1634.76 | 118.25 | 893.51 | 193.028 | |
| 3901 | 22 | 1 | 1 | 2 | 2 | 2 R07120 IA | 1380 | 1000 | 2352.27 | 299.417 | 1112.98 | 391.868 | 1740.88 | 84.25 | 1094.34 | 125.3 | |
| 3902 | 23 | 1 | 1 | 2 | 3 | 3 R25968 IA | 1570 | 1000 | 6224.88 | 583.75 | 2782.56 | 399.056 | 3339.11 | 96.9167 | 1770.9 | 132.058 | |
| 3903 | 24 | 1 | 1 | 2 | 4 | 4 R25968 IA | 1760 | 990 | 4954.78 | 382.083 | 2550.07 | 311.805 | 2832.34 | 34.5 | 1585.98 | 71.1535 | |
| 3904 | 25 | 1 | 1 | 2 | 5 | 5 R32895 IA | 1960 | 1000 | 5549.04 | 281.5 | 2295.97 | 147.001 | 5008.87 | 82.9167 | 2644.07 | 86.4423 | |
| 3905 | 26 | 1 | 1 | 2 | 6 | 6 R32895 IA | 2180 | 1000 | 5691.9 | 275.333 | 2374.4 | 222.438 | 5138.71 | 119.917 | 3092.82 | 214.746 | |
| 3906 | 27 | 1 | 1 | 2 | 7 | 7 R78857 IA | 2350 | 990 | 1420.51 | 241.333 | 1539.94 | 222.021 | 803.295 | 214.333 | 378.893 | 201.801 | |
| 3907 | 28 | 1 | 1 | 2 | 8 | 8 R78857 IA | 2550 | 990 | 2362.12 | 337.417 | 2091.02 | 240.861 | 904.02 | 234 | 546.448 | 256.907 | |
| 3908 | 29 | 1 | 1 | 2 | 9 | 9 H23380 IA | 2760 | 1000 | 1090.52 | 269.187 | 635.403 | 210.538 | 447.556 | 34.9333 | 296.333 | 50.4324 | |
| 3909 | 30 | 1 | 1 | 2 | 10 | 10 H23380 IA | 2980 | 1000 | 880.182 | 226.917 | 517.722 | 132.581 | 588.732 | 143.417 | 276.231 | 198.078 | |
| 3910 | 31 | 1 | 1 | 2 | 11 | 11 T87178 IM | 3180 | 1000 | 1112.88 | 248.917 | 482.555 | 128.425 | 658.079 | 158.933 | 391.387 | 216.785 | |
| 3911 | 32 | 1 | 1 | 2 | 12 | 12 T87178 IM | 3380 | 1000 | 1195.43 | 308 | 438.131 | 187.91 | 723.688 | 115.583 | 348.058 | 189.729 | |
| 3912 | 33 | 1 | 1 | 2 | 13 | 13 T84561 IM | 3570 | 1000 | 1793.33 | 332.833 | 650.882 | 293.889 | 1174.51 | 118.5 | 673.688 | 218.881 | |
| 3913 | 34 | 1 | 1 | 2 | 14 | 14 T84561 IM | 3770 | 1000 | 1827.09 | 267.417 | 1038.4 | 218.47 | 1301.6 | 89.75 | 737.278 | 86.8877 | |
| 3914 | 35 | 1 | 1 | 2 | 15 | 15 T95249 IM | 3980 | 990 | 1420.42 | 287.25 | 1089.15 | 187.717 | 824 | 140 | 375.165 | 188.604 | |
| 3915 | 36 | 1 | 1 | 2 | 16 | 16 T95249 IM | 4160 | 1000 | 2163.85 | 345.5 | 930.152 | 233.307 | 1328.99 | 70.75 | 788.26 | 176.845 | |



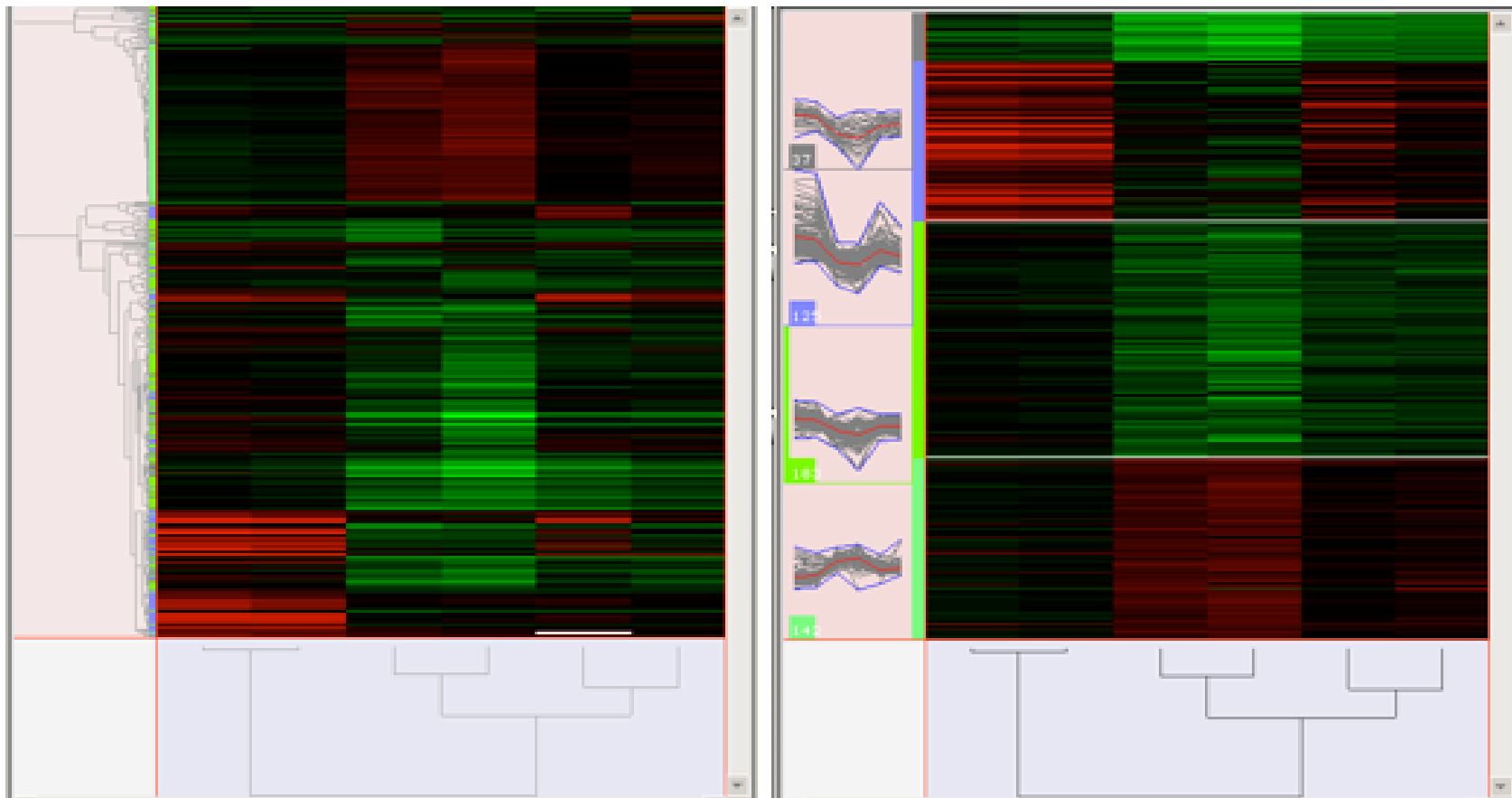
Data Mining and Visualization: Clustering



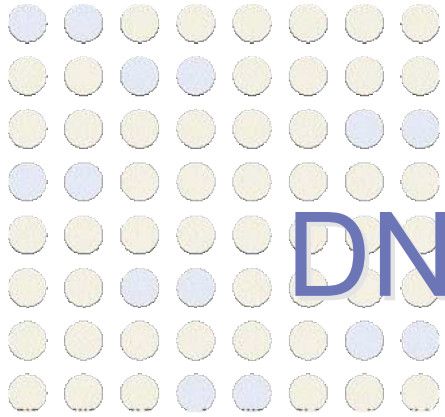
UHN Microarray Centre



The Importance of Multiple Methods

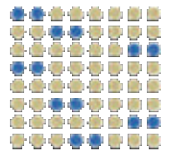


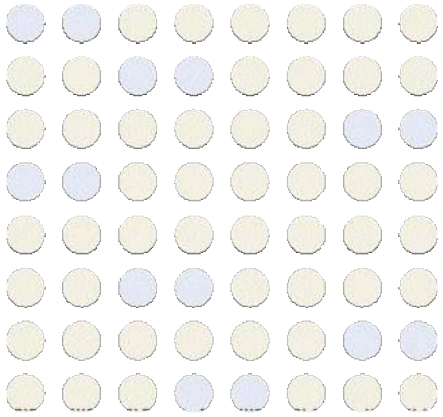
UHN Microarray Centre



DNA Arrays to Study DNA

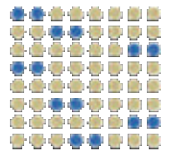
- DNA Microarrays can be used to study the DNA of an organism
 - Point mutations (SNPs) can be studied using oligonucleotide probes
 - Gross mutations such as deletions, additions, and multiplications can be studied using comparative genomic hybridization (CGH)





Arrays for CGH

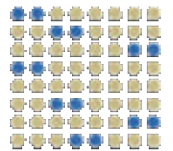
- CGH – Comparative Genomic Hybridization
 - The *sample* is derived from DNA
 - The *features* on the array may be oligos, cDNAs or larger pieces of DNA from BACs (Bacterial Artificial Chromosomes) YACs (Yeast Artificial Chromosomes) etc...
 - The technology does not even require that the genome of the organism is fully described
 - **Genomotyping**, an extension of CGH has been demonstrated using bacteria for which the entire genome can be spotted (*C. jejuni* for example).



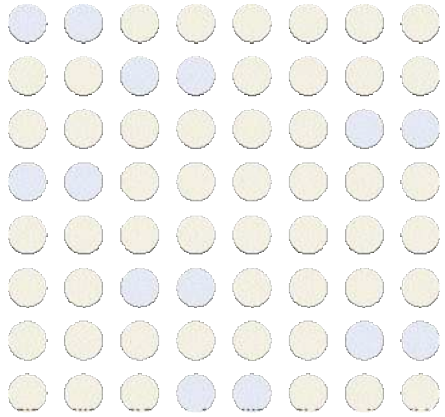


Potential Applications of Arrays to the Study of Food Protection

- Comparison of virulent to non-virulent strains
 - Identifying key regions of the genome (GGH)
 - Identifying key genes or proteins (gene expression, antibody arrays)
- Diagnostics
 - Testing samples for presence of virulent organisms (CGH, SNPs, Gene Expression, Protein Arrays)
- Drug treatments
 - Identification of potential drugs for the treatment of food borne illness (Small Molecule arrays, Cell Arrays)

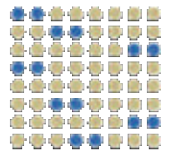


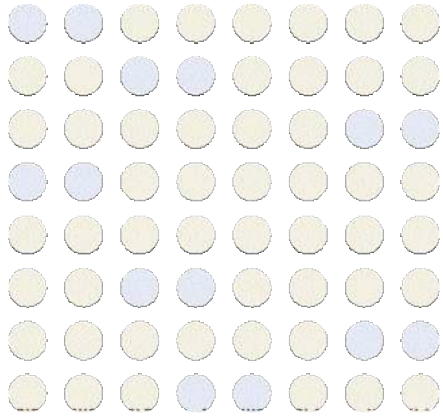
UHN Microarray Centre



Summary

- Microarrays allow for high-throughput analysis of DNA, RNA, Protein etc...
- There are two main types of DNA arrays
 - Affymetrix GeneChips™ - a patented technology, turn-key solution
 - Spotted Arrays – somewhat more flexible, but less turn-key methodology
- DNA arrays can be used to analyse gene expression with an oligo/cDNA array, or DNA mutations using genomic DNA

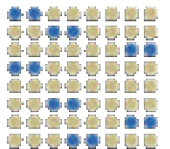




Thank you for your attention

Visit our website at:

www.microarrays.ca



UHN Microarray Centre