

# What is Microarrays

Lecture 1

MBI401-High throughput Data analysis

Mamta Sagar

Department of Bioinformatics

UIET-IBSBT, CSJM University, Kanpur

# 1. INTRODUCTION

- Functional genomics involves the analysis of large datasets of information derived from various biological experiments. One such type of large-scale experiment involves monitoring the expression levels of thousands of genes simultaneously under a particular condition, called gene expression analysis.

- Microarray technology makes this possible and the quantity of data generated from each experiment is enormous, dwarfing the amount of data generated by genome sequencing projects

- Nucleic acid arrays or more simply DNA arrays are a group of technologies in which specific DNA sequences are either deposited or synthesized in a 2-D (or sometimes 3-D) array on a surface in such a way that the DNA is covalently or non-covalently attached to the surface.

- In typical use, a DNA array is used to probe a solution of a mixture of labeled nucleic acids and the binding (by hybridization) of these “targets” to the “probes” on the array is used to measure the relative concentrations of the nucleic acid species in solution.

- By generalizing to a very large number of spots of DNA, an array can be used to quantify an arbitrarily large number of different nucleic acid sequences in solution.
- There are other means of quantifying different nucleic acid sequences in a sample, including quantitative PCR, “digital PCR” (Sykes et al., 1992; Vogelstein and Kinzler, 1999), and hybridization to optically tagged “probe sequences” (Geiss et al., 2008).

- It should be noted that the use of microarrays to measure the relative concentrations of nucleic acid sequences in solution is rapidly being supplanted by high throughput sequencing methods

# The early history of DNA arrays

- The early history of DNA arrays was started by the colony hybridization method of Grunstein and Hogness (Grunstein and Hogness, 1975). In this procedure, DNA of interest was randomly cloned into E. coli plasmids that were plated onto agar petri plates covered with nitrocellulose filters. Replica plating was used to produce additional agar plates.



# History

- The colonies on the filters were lysed and their DNA's were denatured and fixed to the filter to produce a random and unordered collection of DNA spots that represented the cloned fragments. Hybridization of a radiolabeled probe of an DNA or RNA of interest was used to rapidly screen 1000's of colonies to identify clones containing DNA that was complimentary to the probe .

# History

- In 1979, this approach was adapted to create ordered arrays by Gergen et. al. (Gergen et al., 1979) who picked colonies into 144 well microplates. They created a mechanical 144 pin device and a jig that allowed them to replicate multiple microtiter plates on agar and produce arrays of 1728 different colonies in a 26 × 38 cm region. An additional transfer of colonies to squares of Whatman filter paper followed by a growth, lysis, denaturation and fixing of the DNA to the filter, allowed the production of DNA arrays on filters that could be re-used multiple times.

- During the next decade, filter based arrays and protocols similar to these were used in a variety of applications including: cloning genes of specific interest, identifying SNP's (Miller and Barnes, 1986), cloning genes that are differentially expressed between two samples (Crampton et al., 1980) and physical mapping(Craig et al., 1990).

# Automation

- In the late 1980's and early 1990's Hans Lehrach's group automated these processes by using robotic systems to rapidly array clones from microtiter plates onto filters (Craig et al., 1990; Lennon and Lehrach, 1991).

# The birth of the modern DNA array

In the late 90's and 2000's, DNA array technology progressed rapidly as both new methods of production and fluorescent detection were adapted to the task.

It should also be noted that during this time, there was a gradual transition from spotting relatively long DNA's on arrays to producing arrays using 25-60bp oligos.

Three basic types of arrays came into play during this time frame, spotted arrays on glass, in-situ synthesized arrays and self assembled arrays

# Spotted arrays

- Spotted arrays—In 1996 Derisi et. al. published a method which allowed very high density DNA arrays to be made on glass substrates(DeRisi et al., 1996). Poly-lysine coated glass microscope slides provided good binding of DNA and a robotic spotter was designed to spot multiple glass slide arrays from DNA stored in microtiter dishes. By using slotted pins (similar to fountain pens in design) a single dip of a pin in DNA solution could spot multiple slides. Spotting onto glass, allowed one to fluorescently label the sample. Fluorescent detection provided several advantages relative to the radioactive or chemiluminescent labels common to filter based arrays.

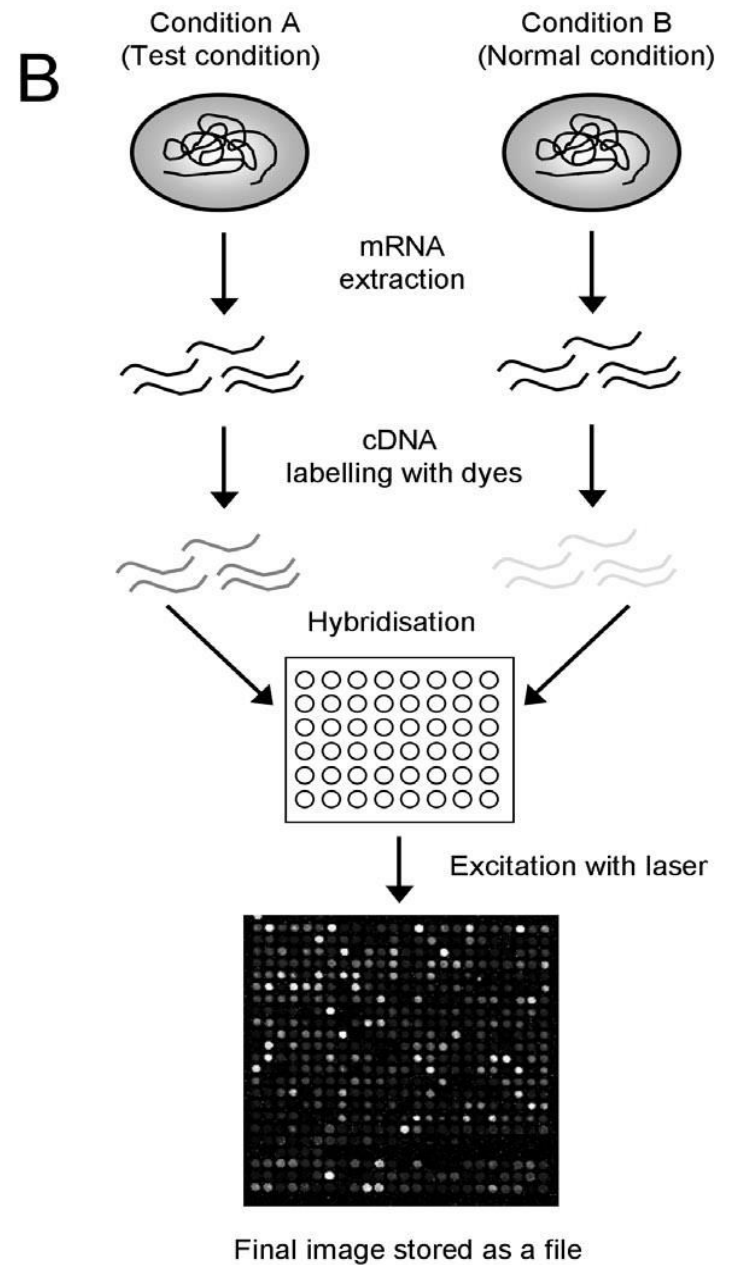
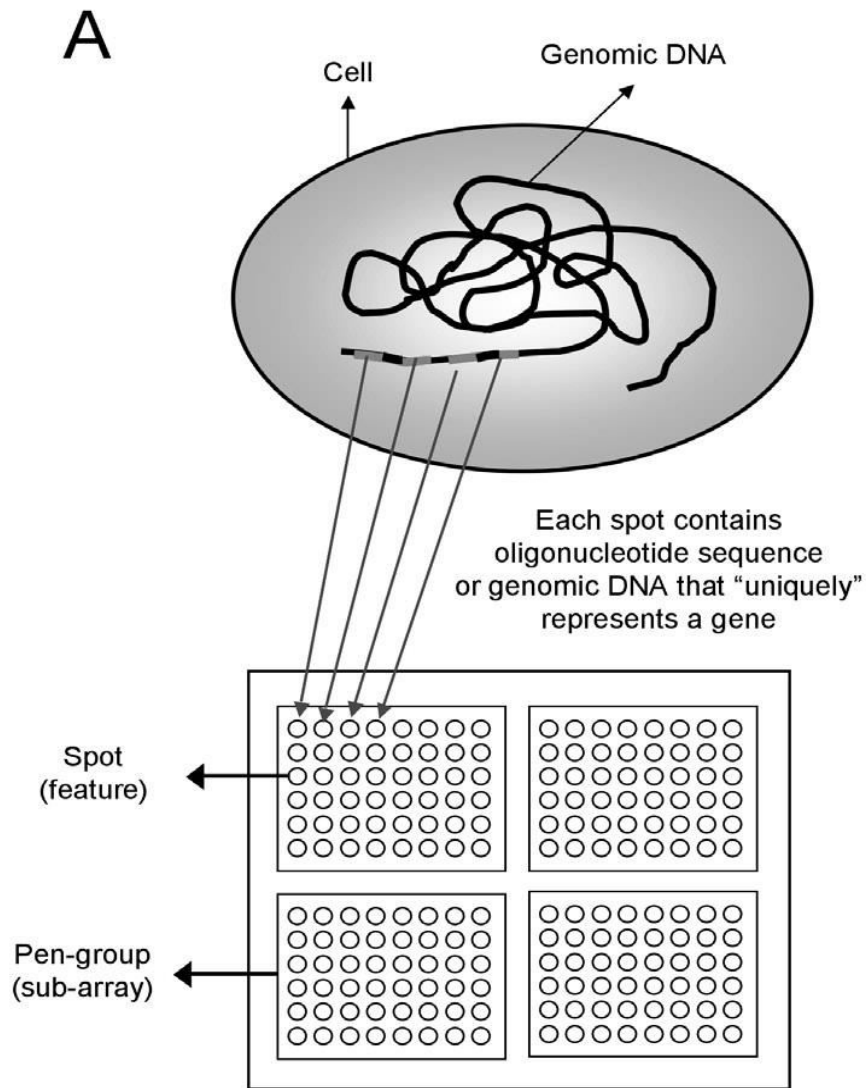
# In-situ, Synthesized arrays

- In-situ, Synthesized arrays—In 1991 Fodor et.al. published a method for light directed, spatially addressable chemical synthesis which combined photolabile protecting groups with photolithography to perform chemical synthesis on a solid substrate (Fodor et al., 1991). In this initial work, the authors demonstrated the production of arrays of 10-amino acid peptides and, separately, arrays of di-nucleotides. In 1994, Fodor et.al. at the recently formed company of Affymetrix demonstrated the ability to use this technology to generate DNA arrays consisting of 256 different octa-nucleotides (Pease et al., 1994). By 1995-1996,

# Self assembled arrays

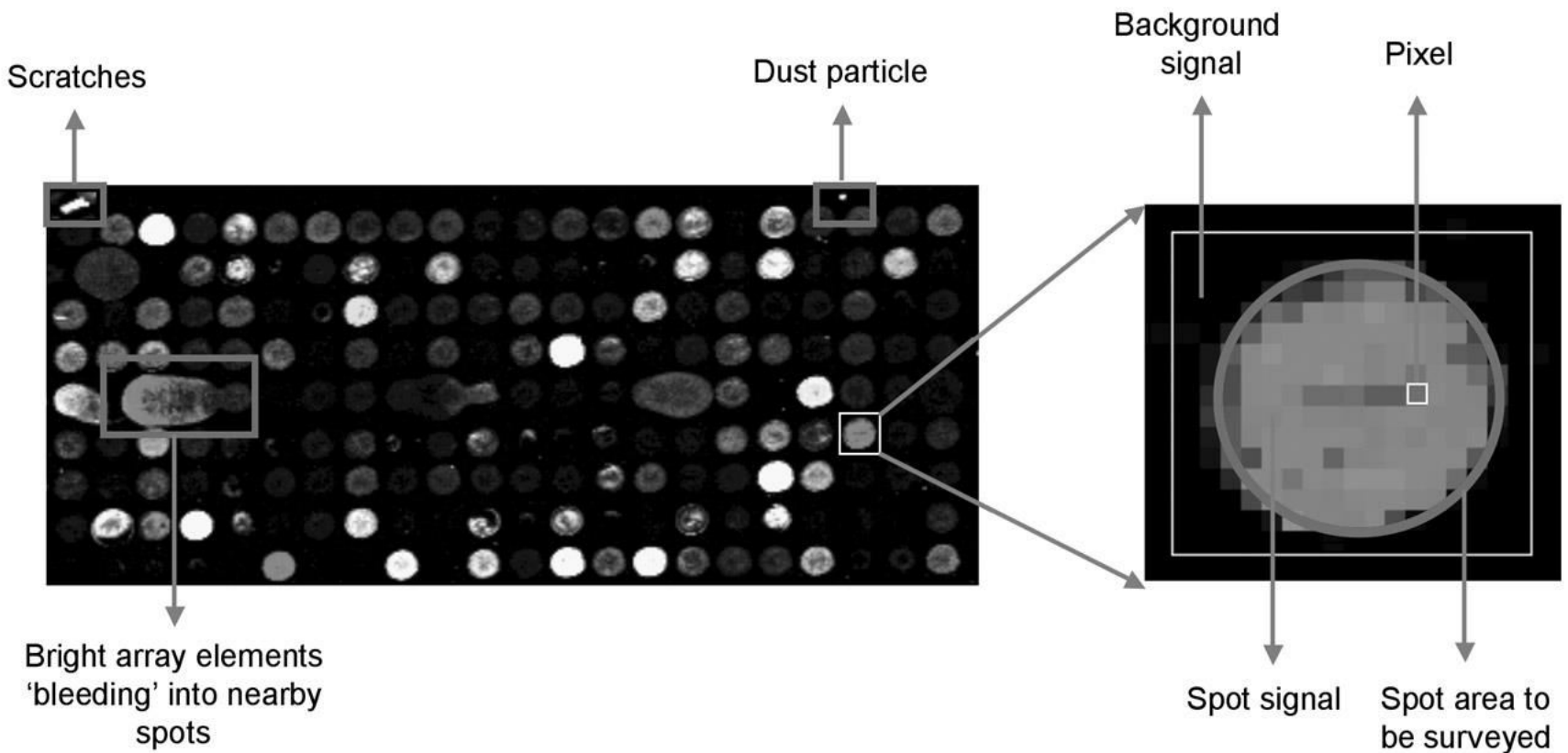
- Self assembled arrays—An alternative approach to the construction of arrays was created by the group of David Walt at Tufts University (Ferguson et al., 2000; Michael et al., 1998; Steemers et al., 2000; Walt, 2000) and ultimately licensed to Illumina. Their method involved synthesizing DNA on small polystyrene beads and depositing those beads on the end of a fiber optic array in which the ends of the fibers were etched to provide a well that is slightly larger than one bead. Different types of DNA would be synthesized on different beads and applying a mixture of beads to the fiber optic cable would result in a randomly assembled array.





- Figure 1. (A) A microarray may contain thousands of 'spots'. Each spot contains many copies of the same DNA sequence that uniquely represents a gene from an organism. Spots are arranged in an orderly fashion into P engroups.
- (B) Schematic of the experimental protocol to study differential expression of genes. The organism is grown in two different conditions (a reference condition and a test condition). RNA is extracted from the two cells, and is labelled with different dyes (red and green) during the synthesis of cDNA by reverse transcriptase.
- Following this step, cDNA is hybridized onto the microarray slide, where each cDNA molecule representing a gene will bind to the spot containing its complementary DNA sequence.
- The microarray slide is then excited with a laser at suitable wavelengths to detect the red and green dyes. The final image is stored as a fi le for further analysis. Colour fi gure at: <http://www.mrc-lmb.cam.ac.uk/genomes/madanm/microarray/>.

Figure 2. Zooming onto a spot on the microarray slide. The spot area and the background area are depicted by a blue circle and a white box, respectively. A pixel in the spot area is also shown. Any pixel within the blue circle will be treated as a signal from the spot. Pixels outside the blue circle but within the white box will be treated as a signal from the background. One can see that the images are not perfect, as it is often the case, which leads to many problems with spurious signals from dust particles, scratches, bright arrays, etc. This image was retrieved from Stanford Microarray Database. Colour figure at: <http://www.mrc-lmb.cam.ac.uk/genomes/madanm/microarray/>.



- If the gene was expressed to the same extent in both conditions, one would find the spot to be yellow, and if the gene was not expressed in both conditions, the spot would be black.
- Thus, what is seen at the end of the experimental stage is an image of the microarray, in which each spot that corresponds to a gene has an associated fluorescence value representing the relative expression level of that gene.

# References

Lecture was prepared using following study material

- Roger Bumgarner, DNA microarrays: Types, Applications and their future Curr Protoc Mol Biol. 2013 January ; 0 22: Unit–22.1.. doi:10.1002/0471142727.mb2201s101.
- Madan Babu, M., 2015. *An Introduction to Microarray Data Analysis*. [online] Mrc-lmb.cam.ac.uk. Available at: <<http://www.mrc-lmb.cam.ac.uk/genomes/madanm/microarray/>> [Accessed 3 December 2015].