Homework #1 - Statistics for Bioinformatics
Stefan Zdraljevic

Microarray Technologies

The advent of microarray technology has revolutionized biomedical sciences. As with most biomedical technologies, microarrays adapt biological processes toward achieving a functional output, which is primarily measuring gene expression levels or genotyping organisms.

Overview of microarrays:
mRNA isolation —> cDNA/cRNA production —> Add FL. labels —> Hybridize to chip and measure fluorescence —> Analysis

DNA microarrays are made possible because of the chemical principles that dictate hybridization of two DNA strands. DNA exists as a double helix of two complementary polymers, Guanine binds with Cytosine while Adenine binds with Thymine through hydrogen bonding. mRNA sequences are produced inside organisms through a process called transcription, which converts DNA into mRNA. mRNA is essentially a copy of DNA (uracil replaces thymine and differences in backbone) that can be easily isolated from the organism of interest. A given mRNA transcript corresponds to a gene (ignoring splice variants and isoforms) and its abundance is often highly correlated with the activity of the gene. Therefore, by isolating cellular mRNA one could measure a genes expression levels or polymorphisms that are present in the genome.

Isolation of cellular mRNA is the first step when conducting a microarray based experiment. Because mRNA is not very stable it needs to be converted to cDNA through a processes called reverse transcription, which converts RNA to DNA (an additional

step can be taken to generate cRNA). cDNA/cRNA is then labelled with fluorescent dyes, which are used for visualizing hybridization to DNA probes on the microarray chip. A microarray “chip” contains thousands of unique nucleotide polymers in blocks that are complementary to cDNA/cRNA sequences generated from mRNA isolated from the experimental organism. Fluorescently labelled cDNA/cRNA (target sequence) generated from mRNA is washed onto the microarray chip, if the complementary sequence is present on the chip the cDNA will hybridize with it. Once hybridization of probe and target is completed, multiple wash steps are performed to eliminate any residual fluorescently labelled target sequences (eliminates background fluorescence). Now the chip is ready for fluorescence measurement, which quantifies the number of mRNA transcripts present in the original sample by quantifying fluorescence associated with generated cDNA/cRNA.

Traditionally, microarray experiments were done using two fluorescent dyes, one to label cDNA/cRNA from control conditions and another to label cDNA/cRNA from experimental conditions. Both sets of cDNA/cRNA are then washed onto the microarray chip and allowed to hybridize with probe DNA. This allows for the quantification of relative levels of mRNA in experimental vs. control conditions. Newer versions of microarrays (Affymetrix) only use one fluorescent label to visualize transcript abundance and have a few other differences as well.

Affymetrix microarray technologies differ from traditional microarrays in a few ways. First, cRNA generated from cDNA is labelled with biotin and fragmented into smaller

polymers, referred to as oligonucleotides. This leads to their being multiple probes per gene and also leads to smaller probes (26 nucleotides). Second, experimental and control conditions are not mixed. Finally, there is streptavidin-PE wash step included in the Affymetrix protocol for labeling of the biotinylated cRNA, which allows for quantification of cRNA levels.

The main two applications for microarray technologies are for measuring mRNA abundance and for identifying polymorphisms in genomic sequence. Gene activity is typically highly correlated with mRNA abundance. Therefore, by measuring activity of different genes in different conditions, one can start to get an understanding for a genes function. Results from microarray experiments typically drive further investigation into differentially expressed genes. For example, a microarray experiment can identify genes differentially expressed in a tumor, which can then help design experiments to determine if those genes are causal to developing cancer. The second main application of microarrays is genotyping. Single nucleotide polymorphisms can be identified using allele specific oligonucleotides. This technique relies on the differential binding affinities between completely complementary probe-target pairs and pairs with SNPs. SNP detection has applications in identifying disease risk, potential negative reactions to drugs, and forensics.