List of Figures

1.1 The structure of DNA ........................................ 2
1.2 Central-dogma of biology ..................................... 3
1.3 The steps in microarray technology .......................... 5
1.4 Microarray spotting device at the Institute for Genomic Research (top) and examples of commonly used print heads (bottom) ........ 6
1.5 Scan of a cDNA microarray containing the whole yeast genome. Each spot represents a different gene and the colours represent the relative abundance of the mRNA from the red channel versus the mRNA from the green channel. ........................................ 7
1.6 Steps in analysis of microarray data .......................... 9

2.1 An example of boxplot of five arrays .......................... 21
2.2 An example of volcano plot of microarray data .............. 22
2.3 An example of heatmap along with the dendrogram of genes and arrays ...... 24
2.4 An illustration of MAplot ...................................... 24

3.1 Graphs showing (a) lognormal density function in comparison to normal density function (b) lognormal density functions for different parameter values ........................................ 54
3.2 The number of genes selected from the SpikeIn experiment by various tests under different FDR q levels .......................... 64
3.3 The p value for the F test for testing variances of samples .......... 66
3.4 The number of genes selected based on t-test method and generalized p value method at different q levels ........................................... 69

4.1 Visualization of Sporulation data in (a) first two and (b) in first three components ......................................................... 84

4.2 Visualization of SVD of sporulation data. Plots of relative variance (a), the first (b), second (c), and the third (d) eigen genes are shown. . . 85

4.3 The rand index and adjusted rand index by principal components instead of original variables. The k-means algorithm of clustering was applied. ................................................................. 87

4.4 The rand index and adjusted rand index by principal components instead of original variables. The hierarchical average link algorithm of clustering was applied. ................................................................. 88

4.5 Gene transcriptional response of five genes selected at random from the simulated dataset: a) noisy sine wave genes b) noisy exponential genes ................................................................. 89

4.6 Visualization of the simulated data a) The relative contribution of singular values b) first c) second and d) third eigen genes. .................. 90

4.7 Visualization of the simulated data by a) first two and b) first three PC's. .................................................................................. 91

4.8 Visualization of the Leukemia data by first two partial least squares loadings. Group 1 and 2 are ALL and AML samples respectively. . . 93

5.1 The membership values of the 38 samples in (a) cluster 1, AML and (b) in cluster 2, ALL obtained by FCM clustering. ......................... 103

6.1 Probability density plots of $AL(\theta, \alpha, \beta)$ .................................................................................................................. 114

6.2 Histogram of random numbers of $AL(\theta = 0.91, \alpha = 10.25, \beta = 2.35)$ . 116

6.3 The Histogram of gene expression across genes on a single array from Potato ........................................................................... 118

6.4 The Histogram of gene expression across genes from selected arrays from the time course experiment of a compatible interaction between potato and Phytophthora infestans ................................. 119

6.5 The Simulated sample path for the ALAR(1) process ............................. 125

6.6 The histogram of marginal distribution of the ALAR(1) process ........... 126