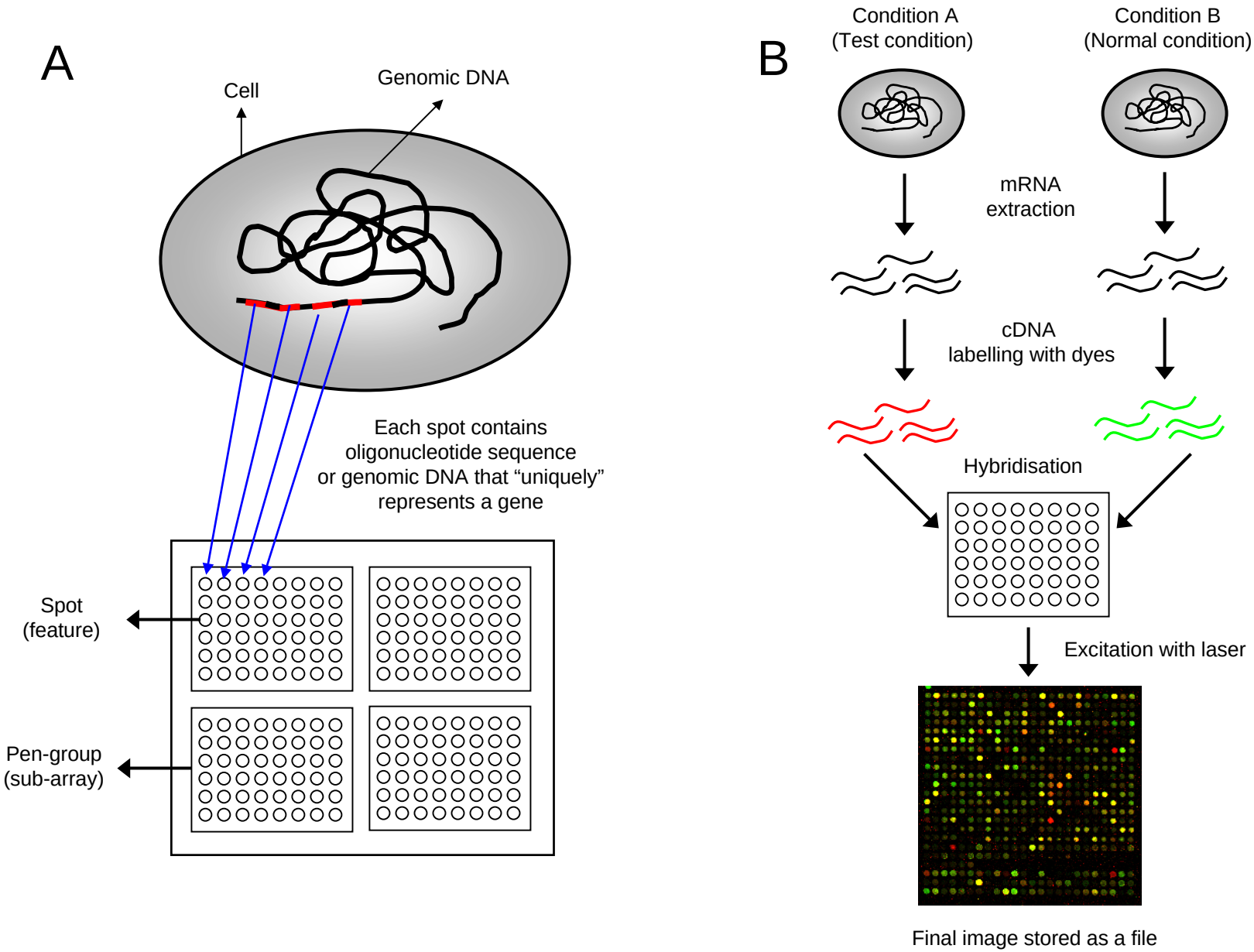

A
Cell
Genomic DNA
Each spot contains
oligonucleotide sequence
or genomic DNA that “uniquely”
represents a gene
Spot
(feature)
Pen-group
(sub-array)
B
Condition A
(Test condition)
Condition B
(Normal condition)
mRNA
extraction
cDNA
labelling with dyes
Hybridisation
Excitation with laser
Final image stored as a file

Figure 1

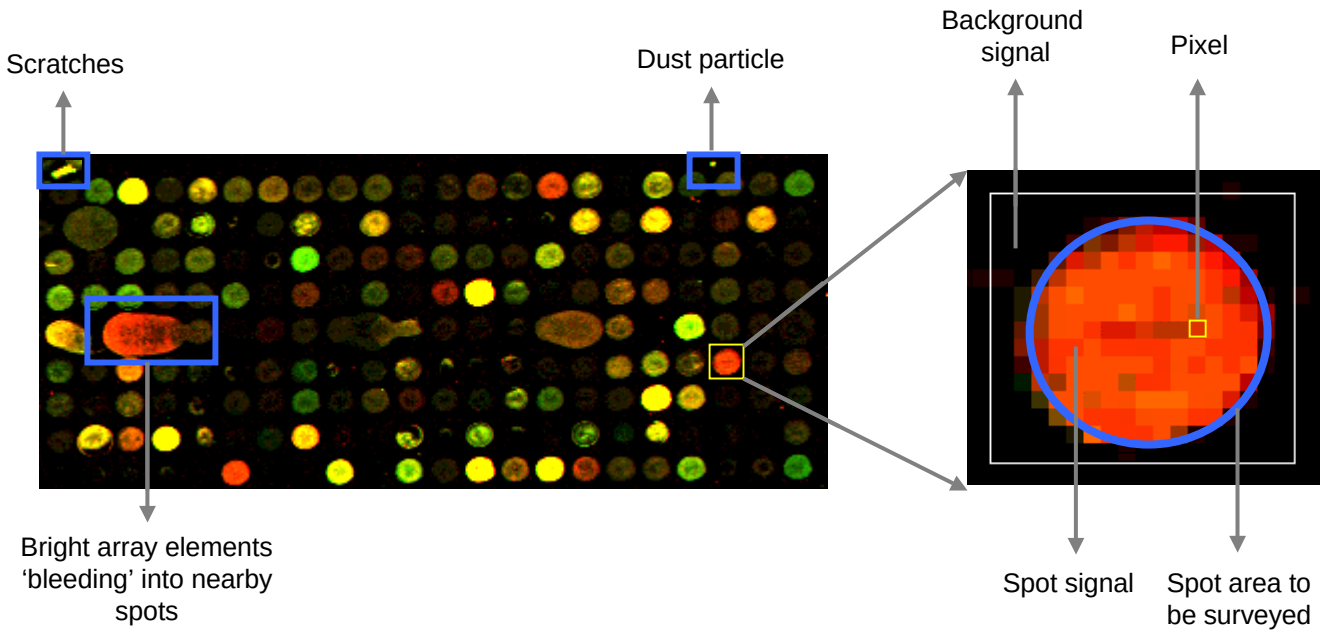
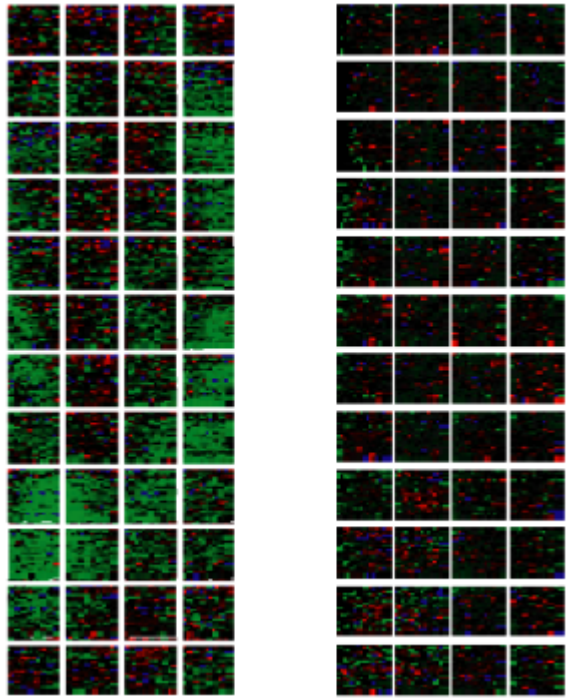

Scratches
Dust particle
Background signal
Pixel
Bright array elements 'bleeding' into nearby spots
Spot signal
Spot area to be surveyed

Figure 2

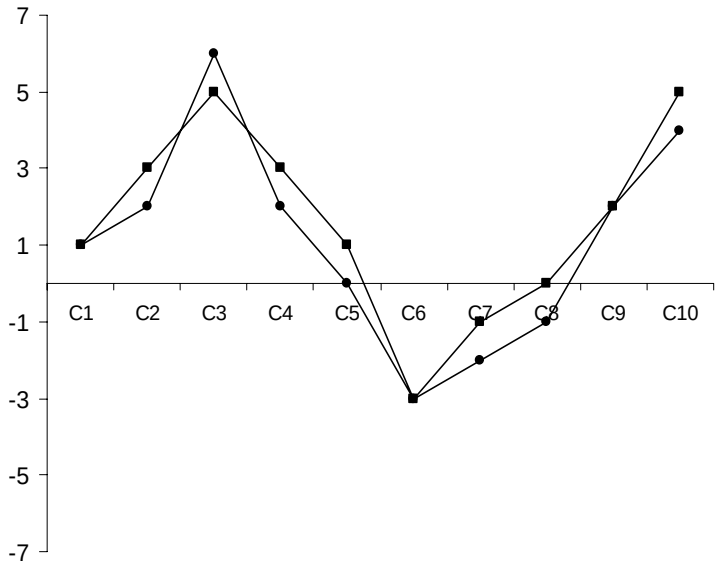


Before normalisation

After normalisation

Figure 3

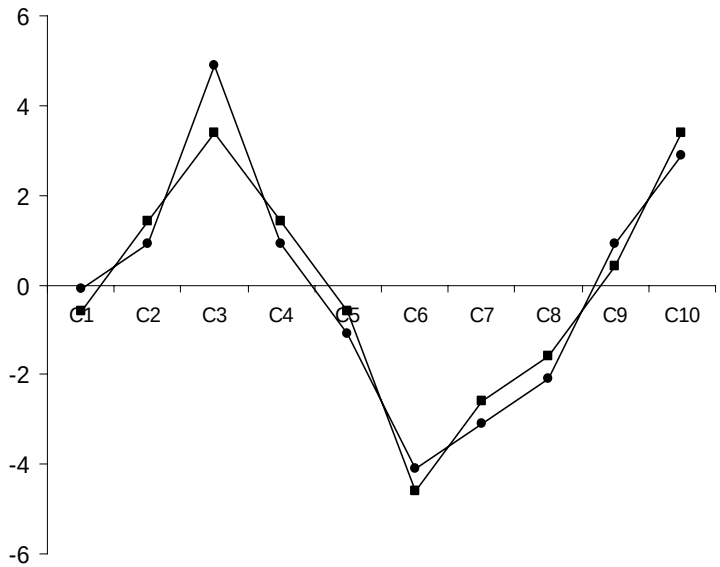
Expression profile before mean centring



Gene A = [1, 2, 6, 2, 0, -3, -2, -1, 2, 4] mean = 1.1

Gene B = [1, 3, 5, 3, 1, -3, -1, 0, 2, 5] mean = 1.6

Expression profile after mean centring



Gene A = [-0.1, 0.9, 4.9, 0.9, -1.1, -4.1, -3.1, -2.1, 0.9, 2.9]

Gene B = [-0.6, 1.4, 3.4, 1.4, -0.6, -4.6, -2.6, -1.6, 0.4, 3.4]

Figure 4

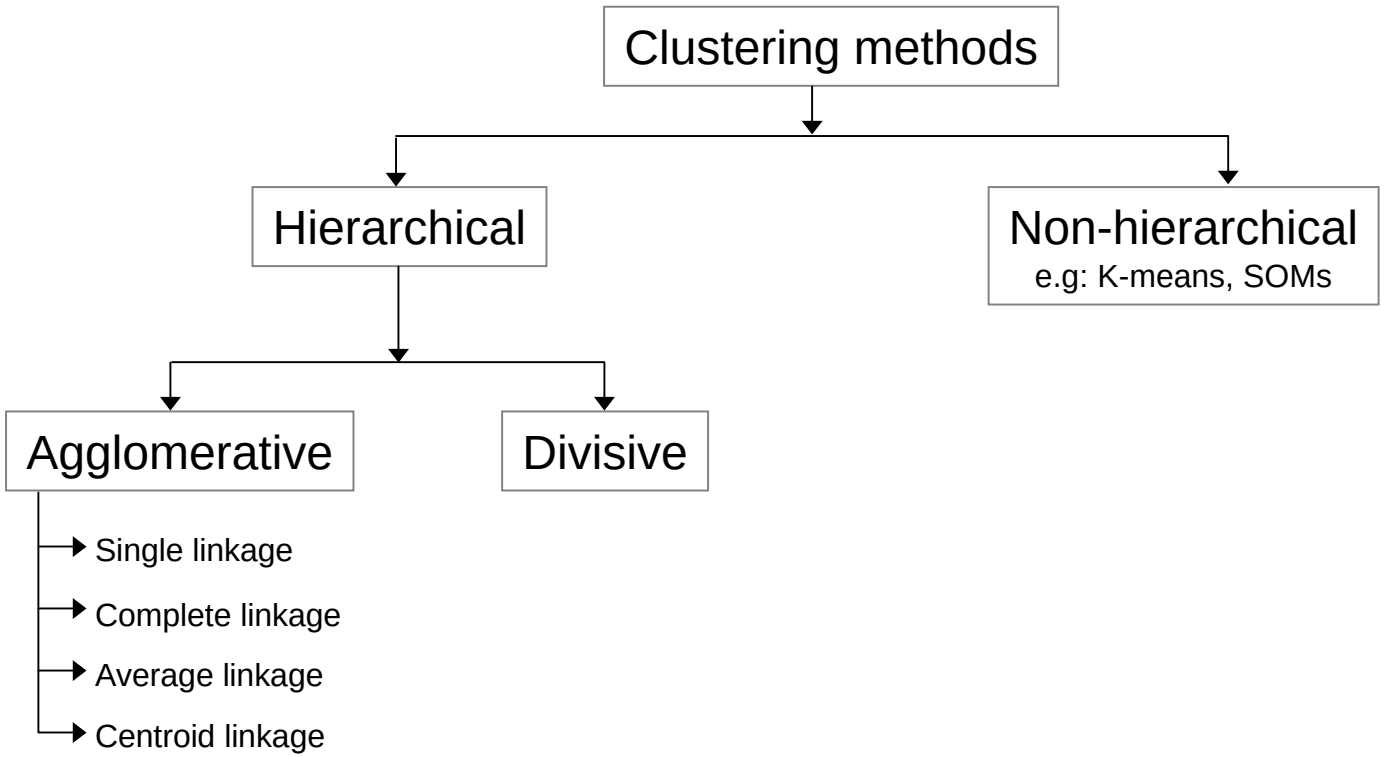

Clustering methods
Hierarchical
Non-hierarchical
e.g: K-means, SOMs
Agglomerative
Divisive
Single linkage
Complete linkage
Average linkage
Centroid linkage

Figure 5

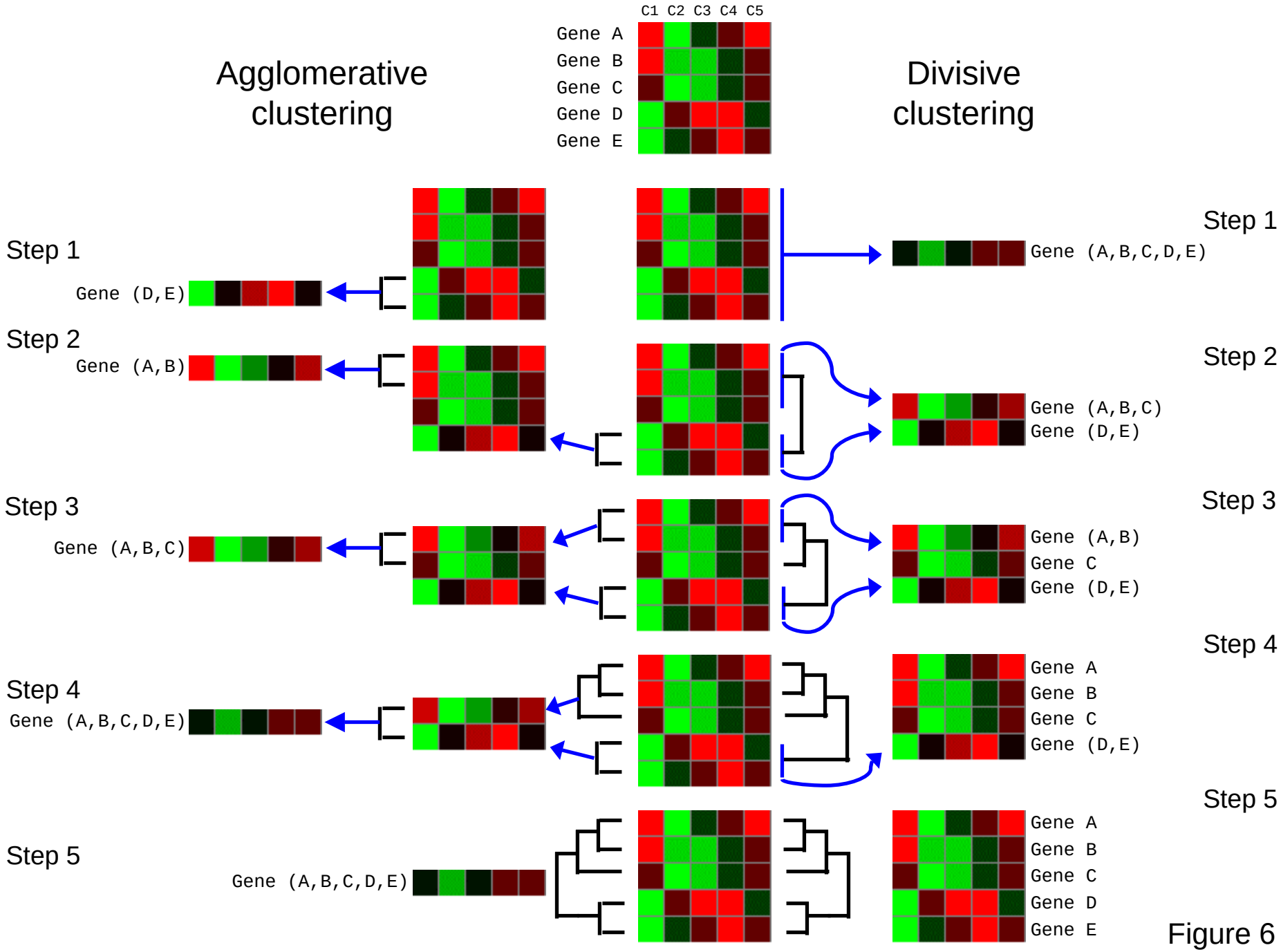

Agglomerative clustering
Divisive clustering
C1 C2 C3 C4 C5
Gene A
Gene B
Gene C
Gene D
Gene E
Step 1
Gene (D,E)
Step 2
Gene (A,B)
Step 3
Gene (A,B,C)
Step 4
Gene (A,B,C,D,E)
Step 5
Gene (A,B,C,D,E)
Step 1
Gene (A,B,C,D,E)
Step 2
Gene (A,B,C)
Gene (D,E)
Step 3
Gene (A,B)
Gene C
Gene (D,E)
Step 4
Gene A
Gene B
Gene C
Gene (D,E)
Step 5
Gene A
Gene B
Gene C
Gene D
Gene E

Figure 6

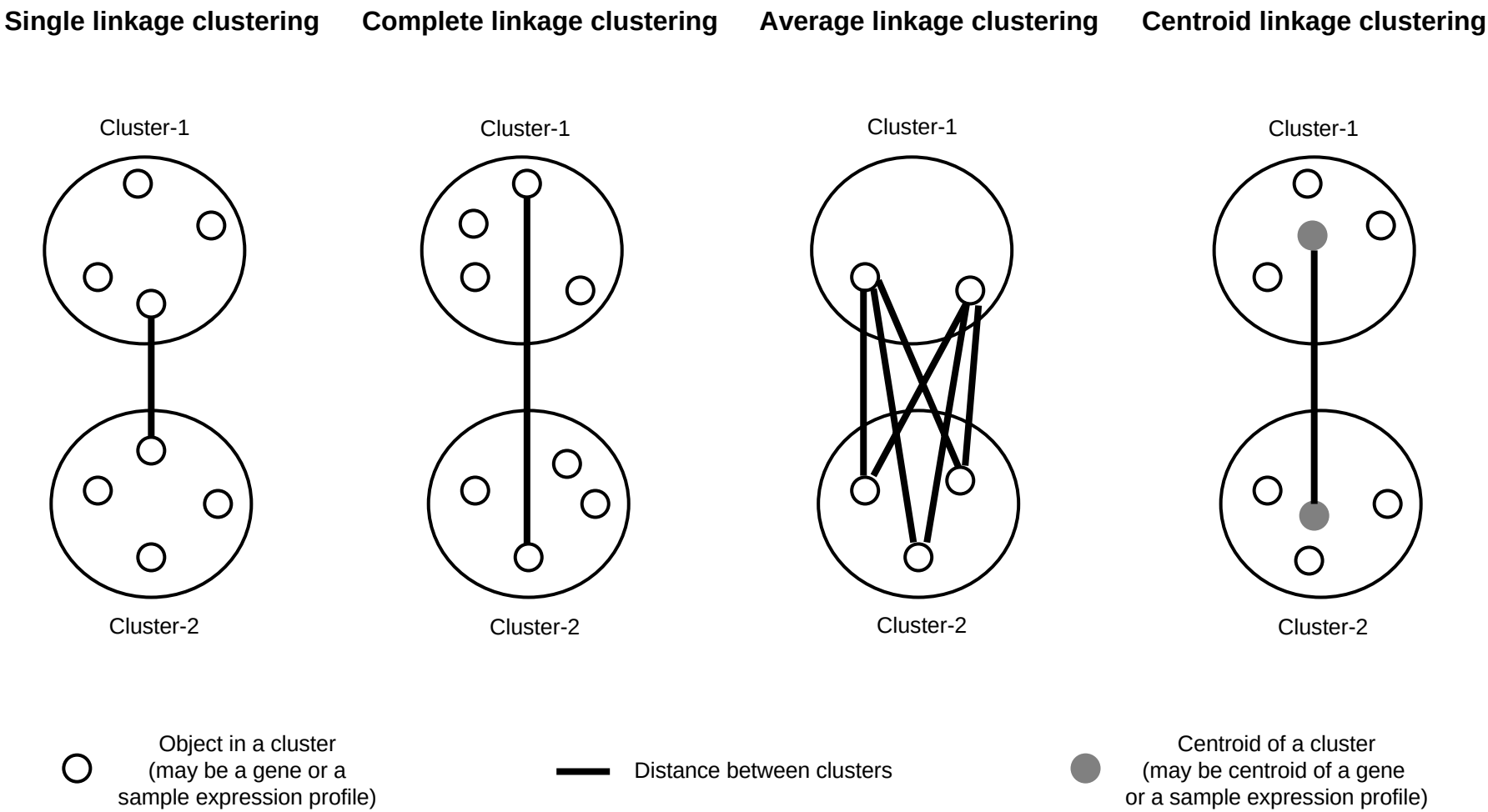
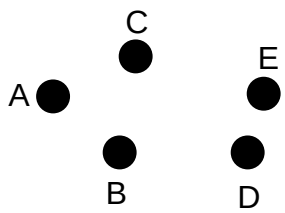
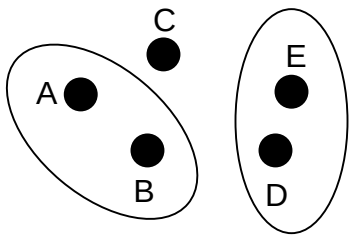

Single linkage clustering
Complete linkage clustering
Average linkage clustering
Centroid linkage clustering
Cluster-1
Cluster-2
Object in a cluster
(may be a gene or a
sample expression profile)
Distance between clusters
Centroid of a cluster
(may be centroid of a gene
or a sample expression profile)

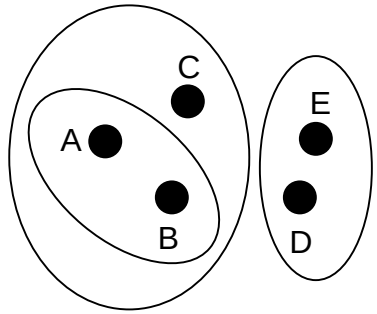
Figure 7



	A	B	C	D	E
A	0	1	2	3	4
B		0	3	4	5
C			0	4	5
D				0	1
E					0



	A,B	C	D,E
A,B	0	2	3
C		0	4
D,E			0



	A,B,C	D,E
A,B,C	0	3
D,E		0

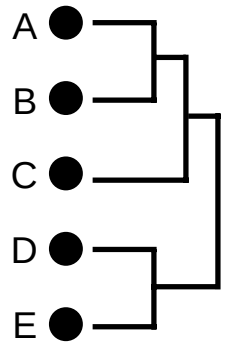


Figure 8

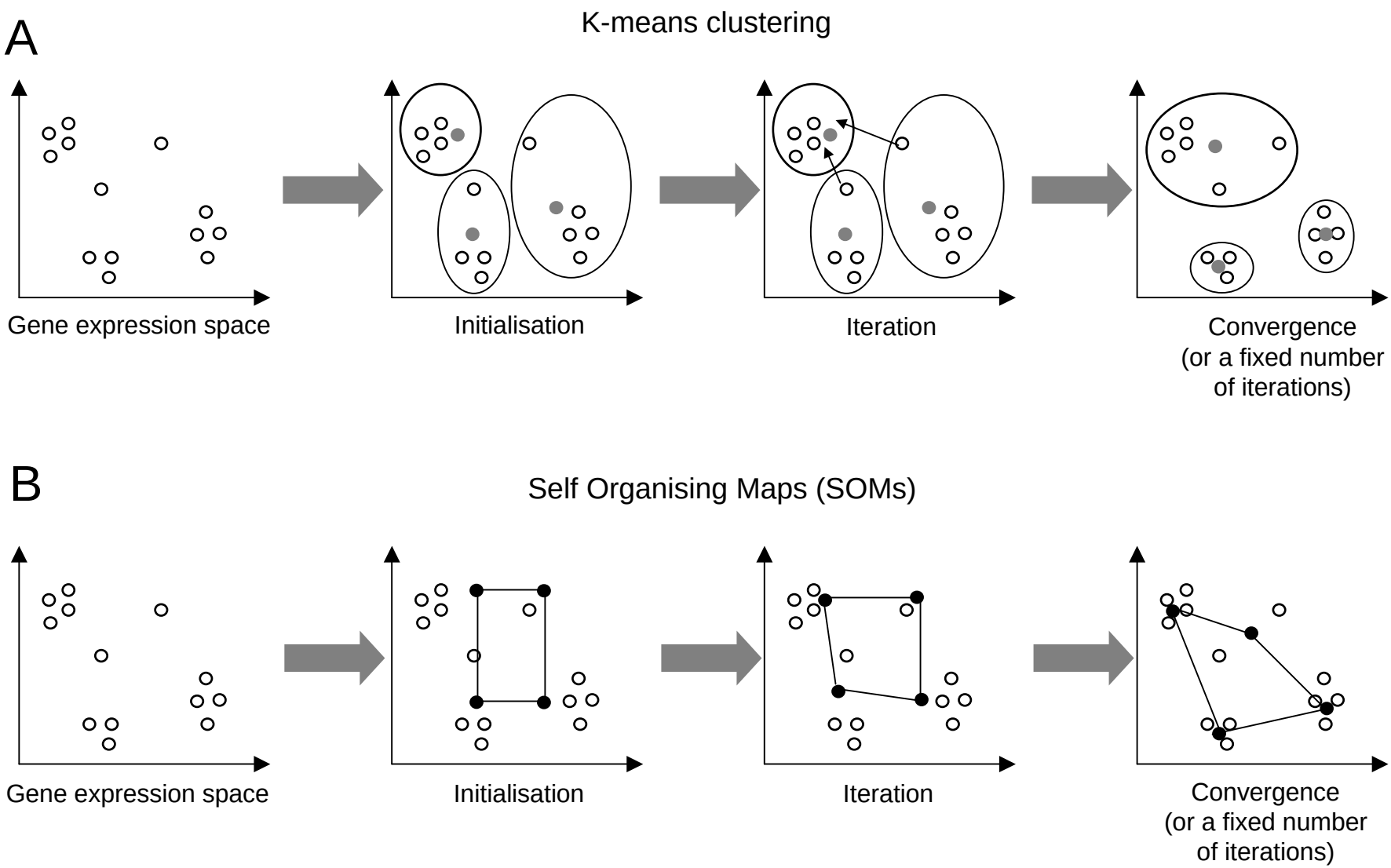

A
K-means clustering
Gene expression space
Initialisation
Iteration
Convergence
(or a fixed number
of iterations)
B
Self Organising Maps (SOMs)
Gene expression space
Initialisation
Iteration
Convergence
(or a fixed number
of iterations)

Figure 9

Table 1.A: Absolute measurement

	C1	C2	C3	C4
Gene A	10	80	40	20
Gene B	100	200	400	200
Gene C	30	240	60	60
Gene D	20	160	80	80

Table 1.B: Relative measurement

	C1/C4	C2/C4	C3/C4
Gene A	0.50	4.00	2.00
Gene B	0.50	1.00	2.00
Gene C	0.50	4.00	1.00
Gene D	0.25	2.00	1.00

Table 1.C: $\log_2$(relative measurement)

	$\log_2$ (C1/C4)	$\log_2$ (C2/C4)	$\log_2$ (C3/C4)
Gene A	-1	2	1
Gene B	-1	0	1
Gene C	-1	2	0
Gene D	-2	1	0

Table 1.D: Discrete values

	D [$\log_2$ (C1/C4)]	D [$\log_2$ (C2/C4)]	D [$\log_2$ (C3/C4)]
Gene A	0	1	0
Gene B	0	0	0
Gene C	0	1	0
Gene D	-1	0	0

Table 2: List of software available for academic use

	Software	URL
1	**Express Yourself** - An automated, online microarray data processing platform, where you can upload image files and carry out data processing and data analysis.	http://array.mbb.yale.edu/analysis/
2	**Expression Profiler** - A set of tools for clustering, analysis and visualization of gene expression and other genomic data. Tools in the Expression Profiler allow to perform cluster analysis, pattern discovery, pattern visualization, study and search Gene Ontology categories, generate sequence logos, extract regulatory sequences, study protein interactions, as well as to link analysis results to external databases.	http://ep.ebi.ac.uk/EP/
3	**Cluster & Treeview** - **Cluster** performs a variety of types of cluster analysis and other types of processing on large microarray datasets. Currently includes hierarchical clustering, self-organizing maps (SOMs), K-means clustering, principal component analysis.**Treeview** can be used to graphically browse results of clustering and other analyses from Cluster.	http://rana.lbl.gov/EisenSoftware.htm
4	**Xcluster** - cross platform software for analysing microarray data.	http://genetics.stanford.edu/~sherlock/cluster.html
5	**J-Express** - A Java implementation of hierarchical clustering, self organized maps, and principal component analysis, with several different viewing options and output formats.	http://www.microarrays.org/software.html
6	**TM4** - A package of Open Source software programs for microarray analysis	http://www.tigr.org/software/
7	**GeneXPress** - A visualization and analysis tool for gene expression data, integrating clustering, gene annotation, and sequence information.	http://genexpress.stanford.edu/
8	**GEPAS** - Gene Expression Pattern Analysis Suite.	http://gepas.bioinfo.cnio.es/tools.html
9	**GenMAPP** - A computer application designed to visualize gene expression data on maps representing biological pathways, and other biologically meaningful groups of genes.	http://www.genmapp.org/
10	**OligoArray** - An application which computes gene specific oligonucleotides for genome-scale oligonucleotide microarray construction.	http://berry.engin.umich.edu/oligoarray/