4. Data processing by cDNA microarray

4-1. Feature extraction

The image of the microarray generated by the scanner (see materials and methods) was the raw data of experiment. Computer algorithms, known as a feature extraction software (GenePix 4.0), converted the image into the numerical information that quantified gene expression; this was the first step of data analysis. The image processing involved in feature extraction had a major impact on the quality of your data. The first step in the computational analysis of microarray data is to convert the digital TIFF images of hybridization intensity generated by the scanner into numerical measures of the hybridization intensity of each channel on each feature. This process is known as feature extraction. There are four steps below. The first step is to identify the positions of the features on the microarray. The second of processes, for each feature, is to identify the pixels on the image that are part of the feature. The following step is to identify nearby pixels that will be used for background calculation for each feature. Finally, software can calculate the numerical information for the intensity of the feature and the intensity of the background and quality control information.

In our chip, there are 32 grids in an 8 x 4 pattern, and each grid has 15 x 16 features. Therefore, there are a total of 7,680 features on this array. The cDNA microarray hybridization was used in our study to measure the changes in mRNA levels of a general sampling of human genes (172 ESTs) and another set of genes (7633) which were chosen on the basis of known roles at 1 day following treatments. We did many processes presented in Figure 15 to reduce variability in any step of microarray experiment after the extraction of RNA, hybridization, and image processing. Table 11 showed the selection of genes through various processes for further analysis. This table included balancing factors and the numbers of genes maintained in each microarray experiment.

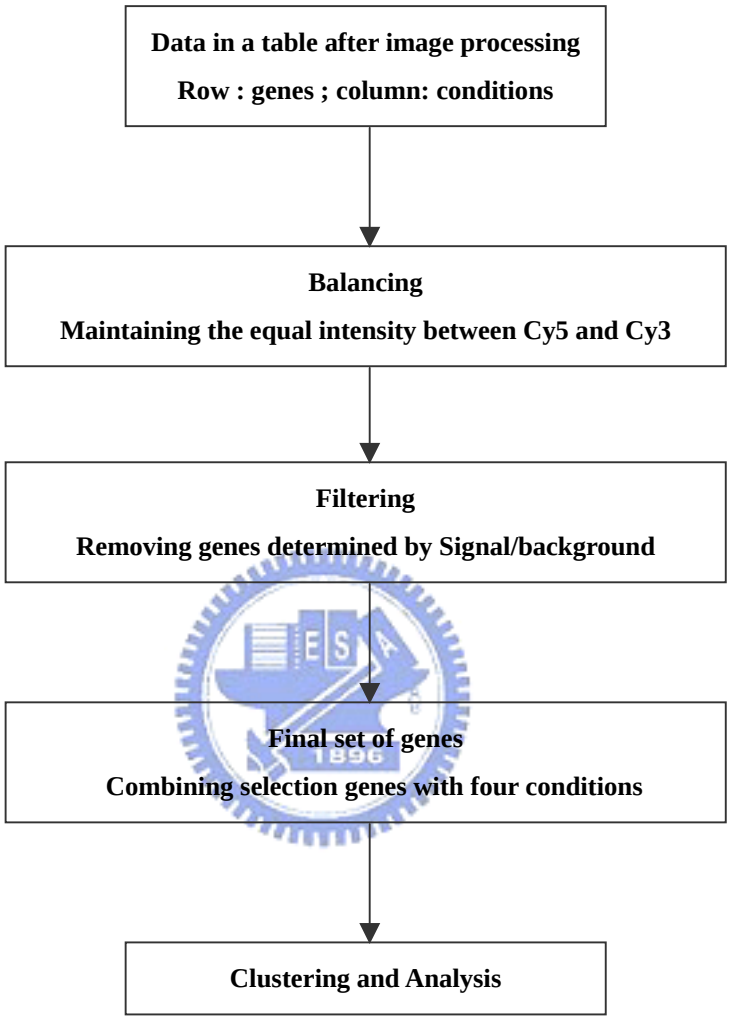


Figure 15. The procedure of data analysis after image processing.

4-2. Balancing

The array was scanned for the two individual fluorescent colors and the signals were corrected for the differences between the Cy3 and Cy5 channels: each probe 2 (Cy5) signal value was multiplied (or divided) by a balance coefficient in order to normalize the gene-expression data relative to probe 1 (Cy3). Data sets were subjected to normalization within each microarray experiment such that the mean of the Cy5 channel was balanced against the ratio of the Cy3 channel. A balance coefficient was calculated as the sum of the mean intensity of the Cy5 channel (F635mean) divided by the sum of the mean intensity of the Cy3 channel (F532mean). Background of Cy3 and Cy5 signal value must be subtracted during the balancing. Figure 16 presented log-intensities for Cy3 and Cy5 with different treatments after balancing processing.

$$\text{Balancing coefficient} = \frac{\sum_{i=1}^{n}(I_{1i} - B_{1i})}{\sum_{i=1}^{n}(I_{2i} - B_{2i})}$$

Where I_{mi} = feature intensity of gene i on array m= Cy5 or Cy3

B_{mi}= background intensity of gene i on array m= Cy5 or Cy3

n = number of genes on the array

4-3. Filtering

Microarray filtering is to exclude genes that may not be interesting in the experiment and to remove genes with very low signal values in both the Cy3 and Cy5 channels. Filtering is dependent on how strict we want to be. In this study, the filtering was calculated as foreground/background and data was to maintain those genes at levels above 1.3 of the relative intensity (foreground/background) in Table 11. The region of

background has a diameter that is 3 times the diameter of the foreground. The foreground is the feature of all pixels analyzed. After the exclusion of some genes 4057 genes were used for further analysis through combinations of genes by different treatments in Figure 17. The scatter plot of Cy5 and Cye3 was in Figure 17 after global normalization.

Table 11. Results in the summary of gene selection procedure.

Steps of processing	Description of methods	Treatments			
		PTCA	**Re188**	**PTCA/Re**	**Re/PTCA**
Balancing					
	Balancing coefficient	1.011	0.602	1.047	0.653
Filtering					
	F635/B635[1]& F532/B532[2] > 1.2				
	The selection of genes	7145	7062	5976	5767
	F635/B635& F532/B532 > 1.3				
	The selection of genes	6732	6343	5234	4754
	F635/B635& F532/B532 > 1.4				
	The selection of genes	6133	5415	4332	3893
Global Normalization					
	Normalization coefficient	1.038	1.035	1.152	1.154

[1], Foreground 635 nm/background 635 nm; [2], foreground 532 nm/background 532 nm

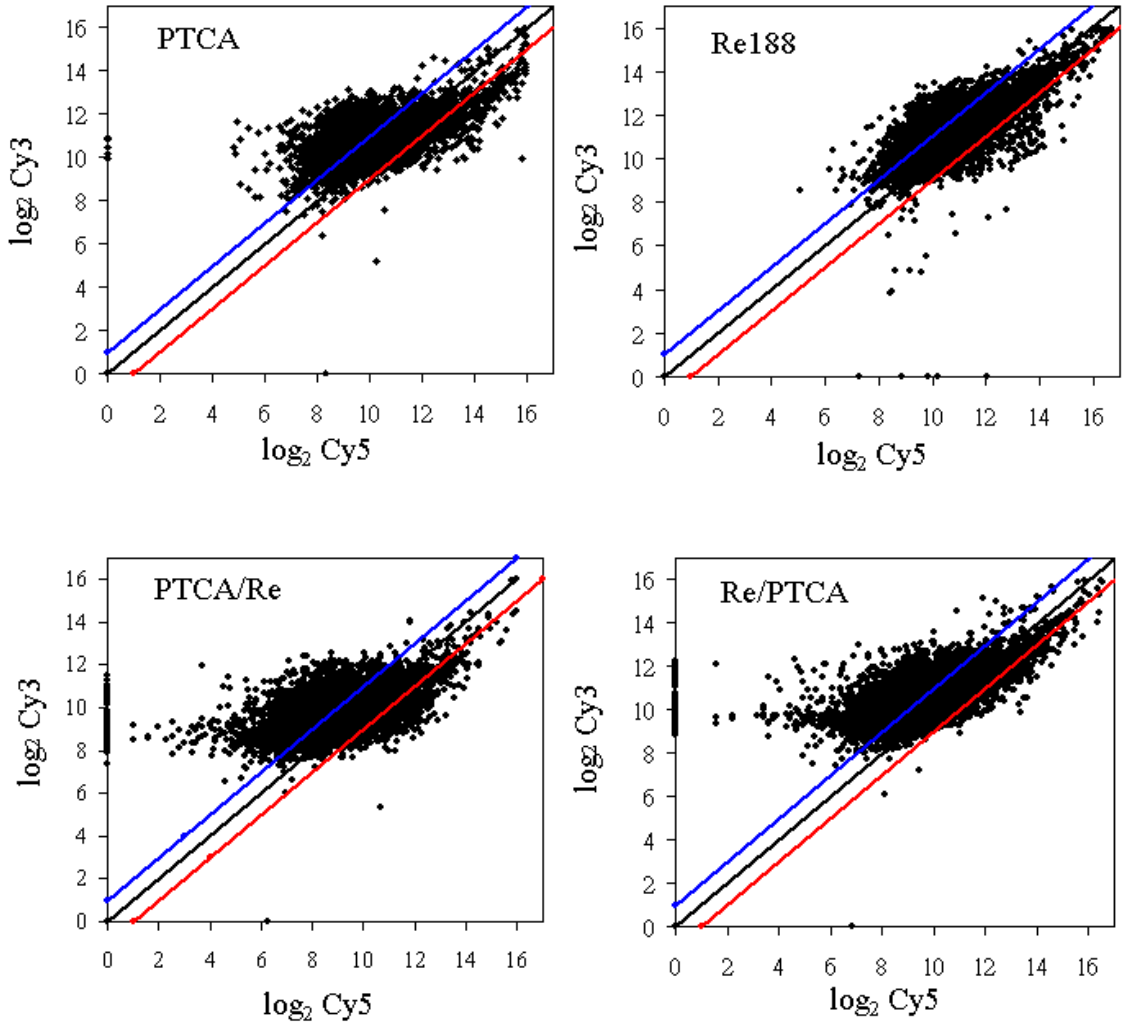


Figure 16. Scatter plot for comparison of expression levels between log intensities for Cy3 and Cy5 without filtering. Each dot represents the hybridization intensity of each gene (7333 elements). Porcine coronary arterial cells were exposed to PTCA, Re188, PTCA prior to Re188 (the PTCA/Re group), and PTCA after Re188 (the Re/PTCA group). In each graph, Cy5 signals were shown in the X-axis, and Cy3 signals in the Y-axis. Intensity of each gene did not be filtered after image processing

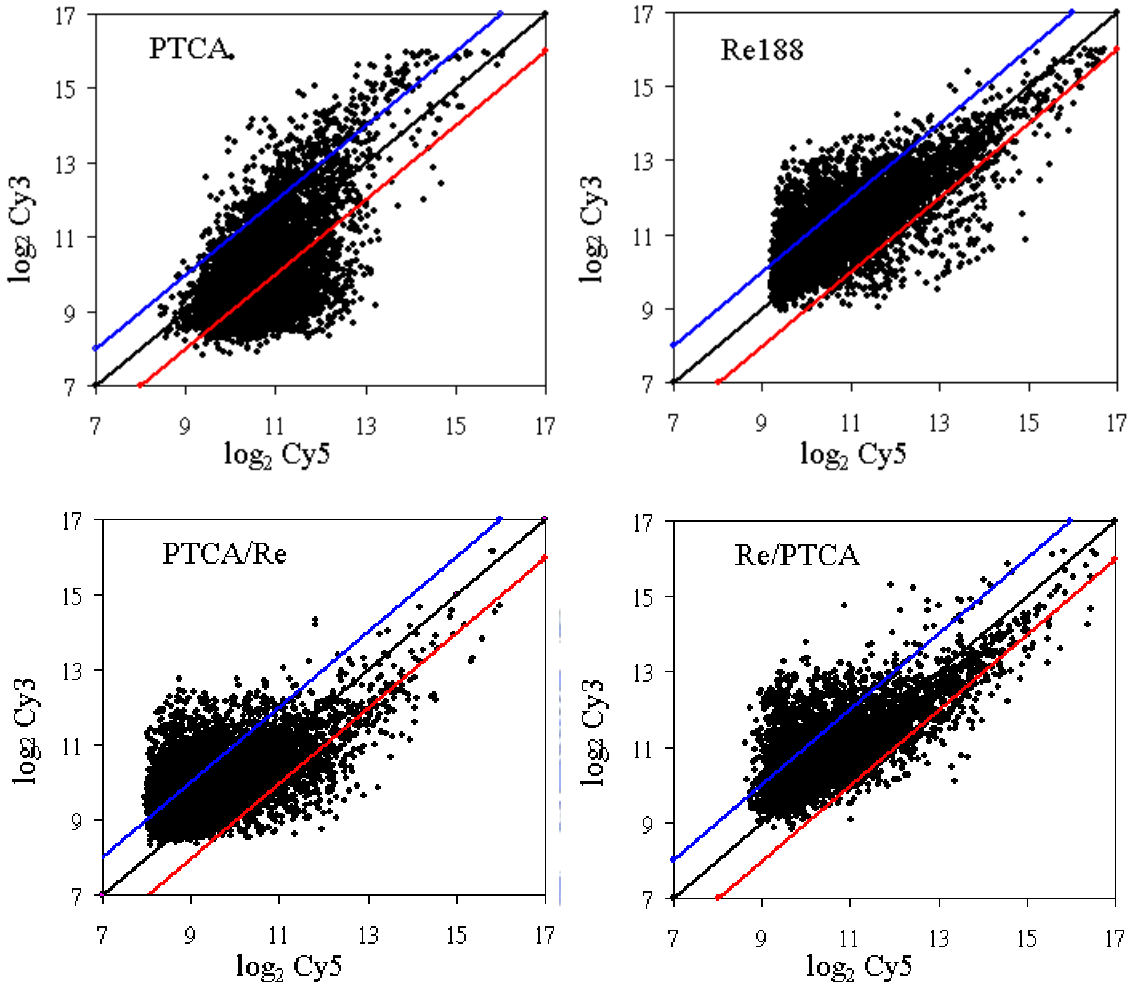


Figure 17. Scatter plot for comparison of expression levels between log intensities for Cy3 and Cy5 with filtering and normalization. Each dot represents the hybridization intensity of each gene (4000 - 6700 elements). Porcine coronary arterial cells were exposed to PTCA, Re188, PTCA prior to Re188 (the PTCA/Re group), and PTCA after Re188 (the Re/PTCA group). In each graph Cy5 signals were shown in the X-axis, and Cy3 signals in the Y-axis. Intensity of each gene had done with filtering and normalization processing.

5. Data analysis by cDNA microarray

5-1. Microarray analysis of injured vessels

To determine which transcripts increased and decreased in injured vessel after treatments, we performed microarray analysis. The ratio of means in 440 transcripts was increasing 2 folds between Cy3 and Cy5 signals at 1 day after the treatment of PTCA. The ratio of means in 246 transcripts was induced to increase 2 folds between Cy3 and Cy5 signals at 1 day upon ^{188}Re exposure. The ratio of means in 429 transcripts was increasing 2 folds between Cy3 and Cy5 signals at 1 day in the PTCA/Re group. The ratio of means in 234 transcripts was induced to increase 2 folds at 1 day in the Re/PTCA group. The ratio of means in 918 transcripts was decreasing 2 folds between Cy3 and Cy5 signals at 1 day after the treatment of PTCA. The ratio of means in 612 transcripts was decreasing 2 folds between Cy3 and Cy5 signals at 1 day upon ^{188}Re exposure. The ratio of means in 707 transcripts was decreasing 2 folds between Cy3 and Cy5 signals at 1 day in PTCA/Re group. The ratio of means in 653 transcripts was decreasing 2 folds between Cy3 and Cy5 signals at 1 day in the Re/PTCA group. The list of the number of up and downregulated genes in PTCA, Re188, PTCA/Re, and Re/PTCA was shown in Table 12. This table included various categories, such as transcription factors, cell cycle, inflammation, extracellular matrix, proliferation, kinase, ESTs.

Table 13 showed upregulated genes by the treatment of PTCA but downregulated genes upon ^{188}Re exposure. Table 14 showed downregulated genes by the treatment of PTCA but upregulated genes upon ^{188}Re exposure.

Table 12. List of number of upregulated and downregulated genes in various kinds of function.

Categories	**PTCA**	**Re188**	**P/Re**	**Re/P**
Upregulated genes				
Transcription factor	16	10	18	10
Cell cycle	8	7	12	2
Inflammation and adhesion molecules	7	7	11	4
Extracellular matrix	74	29	76	44
Proliferation	6	0	4	6
Kinase	15	11	12	10
Growth factor and receptor	15	11	18	7
EST	22	11	21	8
Other	277	160	257	143
Total number of upregulated genes	440	246	429	234
Downregulated genes				
Transcription factor	80	51	59	63
Cell cycle	25	14	14	16
Inflammation and adhesion molecules	29	11	12	14
Extracellular matrix	135	89	106	105
Proliferation	8	11	11	6
Kinase	37	26	38	26
Growth factor and receptor	57	36	47	48
EST	38	23	32	23
Other	509	351	388	352
Total number of downregulated genes	918	612	707	653

5-2. Upregulated genes by the treatment of PTCA and downregulated genes upon Re-188 exposure

The pathogenesis of proliferative arteriopathies, including restenosis after PTCA and stenting, originates with injury to the endothelium. Injured endothelial cells secrete growth and chemotactic factors, which recruit mononuclear cells that secrete inflammatory cytokines. It has been suggested that the cytokine-induced activation of VSMCs is the most critical cellular event in the development of both cardiac allograft vasculopathy and interventional restenosis. Therefore, the identification and functional characterization of genes that regulate these processes is a promising approach for the identification of targets to combat vascular proliferative disorders.

Allograft inflammatory factor-1 (AIF-1) is a 143 amino acid, cytoplasmic, calcium-binding protein [Autieri, 1996]. Studies from several diverse systems demonstrated that AIF-1 expression was shown in a variety of inflammatory situations. For example, its expression was observed in infiltrating macrophages in rat cardiac allografts [Utans et al, 1995], and lesions of experimental autoimmune encephalomyelitis. Studies had previously reported the acute and transient expression of AIF-1 in medial and intimal VSMCs in several models of arterial injury in rat and swine [Autieri et al, 2000]. AIF-1 is not present in normal arteries but is expressed in neointimal VSMCs in injured arteries. AIF-1 is not expressed in unstimulated cultured human VSMCs but is strongly induced in response to inflammatory cytokines and T lymphocyte-conditioned media [Autieri et al, 2000]. A functional interaction with actin enhanced migration of VSMCs that overexpress AIF-1 and suggested that this enhanced migratory activity is attributable to the AIF-1-induced activation of the Rac1 GTPase [Autieri et al, 2003]. The current hypothesis is that the expression of AIF-1 in injured arteries may participate in the progression of cardiac allograft vasculopathy by its ability to promote growth of VSMCs, implicating AIF-1 as a target for treating

proliferative arteriopathies. In our study, the expression of AIF-1 by the treatment of PTCA was 7 folds higher than that upon ^{188}Re exposure and 2.2 folds greater than that in the PTCA/Re group. Expression of AIF-1 in coronary arteries may be promoting their proliferation of SMCs and lead to the dysregulation of cell cycle protein expression and subsequent shortening of the cell cycle [Autieri, 2002].

Ring-box 1 (Rbx1), an evolutionarily conserved protein, contains a RING-H2 finger like motif and interacts with cullins. The yeast homolog of Rbx1 is a subunit and potent activator of the Cdc53-containing SCF^{Cdc4} ubiquitin ligase. SCF complexes, consisting of Skp1, cullins/Cdc53, and F-box proteins, have been implicated in multiple phosphorylation-dependent proteolysis pathways, including the destruction of cyclin-dependent kinase inhibitor, such as p27 in mammals and Sci1 in yeasts by SCF^{Cdc4}, a process required for the G_1 to S transition associated proteins. Rbx1 stimulates Sic1 ubiquitination in vitro. Besides, the SAG/ROC/Rbx is required for yeast cell growth [Kamura et al, 1999] and can promote cell growth and S phase entry on serum starvation if overexpressed [Duan et al, 2001]. Besides, sporulation and tetrad dissection showed 2:0 segregation for viability, indicating that rbx is an essential gene. In our study, the expression of rbx1 had 7 folds higher by the treatment of PTCA than that upon ^{188}Re exposure and 3.5 folds greater than that in the PTCA/Re group in Table 13. Therefore, expression of rbx1 may be inhibited by the radiation, attenuating the progress of cell cycle.

The three human TACC genes encode a family of proteins that are suspected to play a role in carcinogenesis in previous studies. A novel gene, TACC1 (transforming acidic coiled coil-containing gene 1), is located close to FGFR1 within a region amplified in breast cancer on human chromosome 8p11. TACC1 belongs to a family of three paralogous genes [Still et al, 1999; Popovici et al, 2001] that also includes TACC2/AZU1/ECTACC and TACC3.

Their function is not precisely known; a Xenopus TACC protein called Maskin was

involved in translational control, while D-TACC in Drosophila is a centrosomal protein required for normal spindle function in the early embryo [Gergely et al, 2000]. TACC1 was original identified as a potential oncogene. In previous studies TACC1 mRNA gene expression was downregulated in various types of tumors. According to our data, the expression of TACC1 had 5.4 folds higher in the PTCA group and 5 folds higher in the PTCA/Re group than that in the Re188 group in Table 13.

Table 13. List of upregulated genes by the treatment of PTCA and downregulated genes upon Re-188 exposure.

Genbank Accession No.	Abbreviation	Functional Category	Treatments			
			PTCA	Re188	PTCA/Re	Re/PTCA
H63077	ANXA1	Membrane protein	2.020	0.427	2.614	2.690
H09065	BAP1	Tumor suppression	2.076	0.308	0.844	1.863
AA699317	RBX1	Proliferation	2.106	0.302	0.637	1.510
AA598796	TACC1	Extracellular matrix	2.383	0.414	2.079	1.094
AA193116	GPD1	M-dehydrogenase	2.539	0.396	0.554	1.808
AI969670	LDHB	Metabolism	2.583	0.472	1.961	1.926
AA425205	DKFZp761F0118	Binding protein/transcription	2.679	0.436	0.844	1.032
R52654	CYCS	Oxidase/transporter	2.799	0.481	1.927	1.104
T52693	PCTAIRE2BP	Extracellular matrix	2.814	0.420	0.776	1.401
W69953	AIF1	Binding protein/calcium	3.238	0.480	1.457	2.655
AA126911	HNRPA1	Transport	4.033	0.481	3.605	2.549

Table 14. List of downregulated genes by the treatment of PTCA and upregulated genes upon Re-188 exposure.

Genbank Accession No.	Abbreviation	Functional Category	Treatments			
			PTCA	Re188	P/Re188	Re188/P
AA504455	COG2	Receptor	0.352	2.012	3.095	1.269
AA477091	MGC27165	Extracellular matrix	0.443	2.038	1.309	0.651
AA843592	MIPEP	m-peptidase	0.237	2.153	2.196	0.788
AA421687	TSTA3	m-reductase/transporter	0.231	2.214	2.864	0.773
AA894557	CKB	Kinase	0.347	2.300	1.394	1.248
H09230	DTNA	Binding protein/muscle	0.468	2.342	0.559	0.215
AA451817	CCNH	Cell cycle	0.392	2.424	2.148	1.091
AA481209	PRKCE	Channel/transporter	0.353	2.437	1.092	1.322
AA410896	C6orf9	Membrane protein	0.352	2.710	3.282	1.242
AA099394	SSR1	Receptor	0.356	4.701	1.879	0.609
AA219045	MAP1B	Extracellular matrix	0.296	5.766	4.163	0.734

5-3. Data mining by cluster analysis

The purpose of cluster analysis is to divide the given data into groups using same characteristics. It is usually attempted to find “natural groups” in the data. However it is difficult to define what the natural group is. Therefore, these groups are separated and dependent on the combination of data in every experiment. Except for grouping genes, genes in the same group are co-expressed, which gives hints that the unknown genes may have functions of the respective groups they cluster. Furthermore, we can reveal the unknown mechanisms of genes in the signal transduction step-by-step throught microarray and cluster analysis.

Unsupervised clustering using agglomerative and divisive cluster analyses is performed to reveal relationships between genes. In our study, one of the agglomerative approaches was performed using hierarchical clustering by Eisen et al.

Upregulated genes, 781 transcripts, were produced through image processing, balancing, filtering and normalization. The number of genes was divided into 4 groups with the program of Standfort’s software in Figure 18. Downregulated genes, 1437 transcripts, were produced through the same processes. The downregulated genes were grouped into four in Figure 19. Therefore, there were eight groups contributed by upregulated and downregulated genes. However, all of the 4057 genes divided by the clustering program were not presented because we found that they provided us for less understanding of their characteristics. These weak characteristics resulted from the effect of many experiment biases. Therefore, we only identified the clear genes, such as up and downregulated genes to complete the clustering.

After the cluster of upregulated genes, we got 4 groups called cluster 1, 2, 3, and 4, respectively in Table 15. There were 313 genes in the cluster 1. Upregulation of 58 genes resulted from the treatment of angioplasty alone, 21 genes and 63 genes were

upregulated in the PTCA/Re group, and in the Re/PTCA group, respectively. It indicated that genes in cluster 1 were highly similar among PTCA, PTCA/Re, and Re/PTCA groups. There were 130 upregulated genes in the cluster 2. All of the upregulated genes were only contributed by the angioplasty treatment alone. There were 216 upregulated genes in the cluster 3. These genes were primary in the PTCA/Re group. Seventy-eight upregulated genes were in the PTCA/Re group. In addition, three-fourth of 78 genes were upregulated in the Re188 group whereas one-twenty fifth of 78 genes were upregulated in PTCA group, and genes were not upregulated in Re/PTCA group. This indicated that the expression profile of downstream genes might be similar between the PTCA/Re and Re188 groups instead of among the PTCA/Re, Re/PTCA and PTCA groups in the cluster 3. Therefore, the upregulated genes with radiation treatment alone were similar to that treated with angioplasty prior to radiation. There were 52 upregulated genes exposured by radiation whereas there were 5, and 1 upregulated genes in the PTCA only and PTCA/Re groups, respectively. According to the genes in the cluster 4, this indicated that genes couldn't be upregulated in other groups but the upregulated genes were dominant in the ^{188}Re radiation. Therefore, the results showed that genes damaged by radiation leaded to the unique expression profile rather than angioplasty and combination of angioplasty and radiation did in the cluster 4 (Figure 18).

After the cluster of downregulated genes, we got 4 groups called cluster 5, 6, 7, and 8, respectively in the Table 15. There were 326 genes in the cluster 5. PTCA prior to radiation attributed to most of the upregulated genes in cluster 5. Five, 96, and 21 upregulated genes were in the Re188 only, PTCA/Re, and Re/PTCA groups. Zero gene was upregulated in the PTCA group. This cluster indicated that the downregulated genes mainly contributed by PTCA/Re were more related to in Re/PTCA group than that in Re188 group. There were 316 genes in the cluster 6, including 78 downregulated genes in PTCA

group and 34 genes in PTCA/Re group. This cluster represented that downregulated genes in the PTCA group and PTCA/Re group were closer. There were 559 genes in the cluster 7 and approximate one-second of 559 genes were in the PTCA group. Besides, eighty-four downregulated genes were in the Re/PTCA group. This cluster indicated that the relationship of downregulated genes in PTCA and Re/PTCA group was closer than that in other groups. There were 236 downregulated genes in the cluster 8. One hundred thirty-six genes and 1 gene were downregulated in the Re188 and PTCA/Re group, respectively. This result indicated that radiation leaded to dominant downregulated-genes in the cluster 8 (Figure 19). We found whether genes were upregulated or downregulated, the number of upregulated and downregulated genes was dominant in the cluster 4 and 8. Therefore, these results showed that the brachytherapy caused the clear expression profile of genes in our study.

In Table 16, upregulated and downregulated genes in each cluster were divided into several groups which is dependent on the function, including transcription factor, cell cycle, inflammation and adhesion molecules, extracellular matrix, receptor and growth factors, and kinases, and EST. The sum of upregulated genes was 781 and the sum of downregulated genes was 1437.

Table 15. Upregulated and downregulated genes divided through the software program cluster

Cluster	Groups															Descriptions of Clustering
	1	2	3	4	1 &2	1 &3	1 &4	2 &3	2 &4	3 &4	1,2 & 3	1,2, & 4	2,3, & 4	1,3, &4	1,2,3, &4	
Upregulated genes																
1	58	0	21	63	0	23	36	0	0	28	0	0	0	81	3	Uniformly distributed in group 1, 3, and 4
2	87	0	0	0	15	15	2	0	0	0	4	0	0	6	1	Primarily distributed in group 1
3	3	16	78	0	0	22	0	71	0	2	16	0	2	3	3	Distributed in group 2 and 3. Group 3 dominates.
4	5	52	1	0	15	1	0	5	0	0	42	0	1	0	0	Mostly distributed in group 2
Downregulated genes																
5	0	5	96	21	0	3	0	9	10	43	2	1	25	32	79	Distributed in group 3 more than those in group 4
6	78	1	34	0	10	40	3	5	0	0	50	0	0	39	56	Distributed in group 3 more than those in group 1
7	215	0	0	84	25	1	86	0	5	7	4	24	0	43	65	Distributed in group 4 less than those in group 1
8	0	136	1	0	20	0	0	29	6	0	20	0	2	0	22	Primarily distributed in group 2

Group 1 indicates PTCA treatment only, group 2 indicates radiation treatment only, group 3 indicates that PTCA is prior to radiation, and group 4 indicates PTCA after radiation

Table 16. The number of upregulated and downregulated genes divided by functional categories

Functional Categories	Clusters							
	The number of upregulated clustering				The number of downregulated clustering			
	1	2	3	4	5	6	7	8
Transcription factor	14	4	14	3	31	25	53	16
Cell cycle	9	4	7	1	6	7	14	7
Inflammation and adhesion molecules	4	1	7	3	6	8	19	8
Extracellular matrix	63	24	29	18	51	47	97	27
Proliferation	8	0	3	0	4	3	3	3
Receptor and Growth factor	9	3	10	6	19	19	39	6
Kinase	9	6	6	8	7	14	23	9
ESTs	10	7	10	6	14	16	16	4
Other genes	187	81	130	77	188	177	295	156
Total genes in each cluster	313	130	216	122	326	316	559	236
The total no. of genes in each cluster	781				1437			

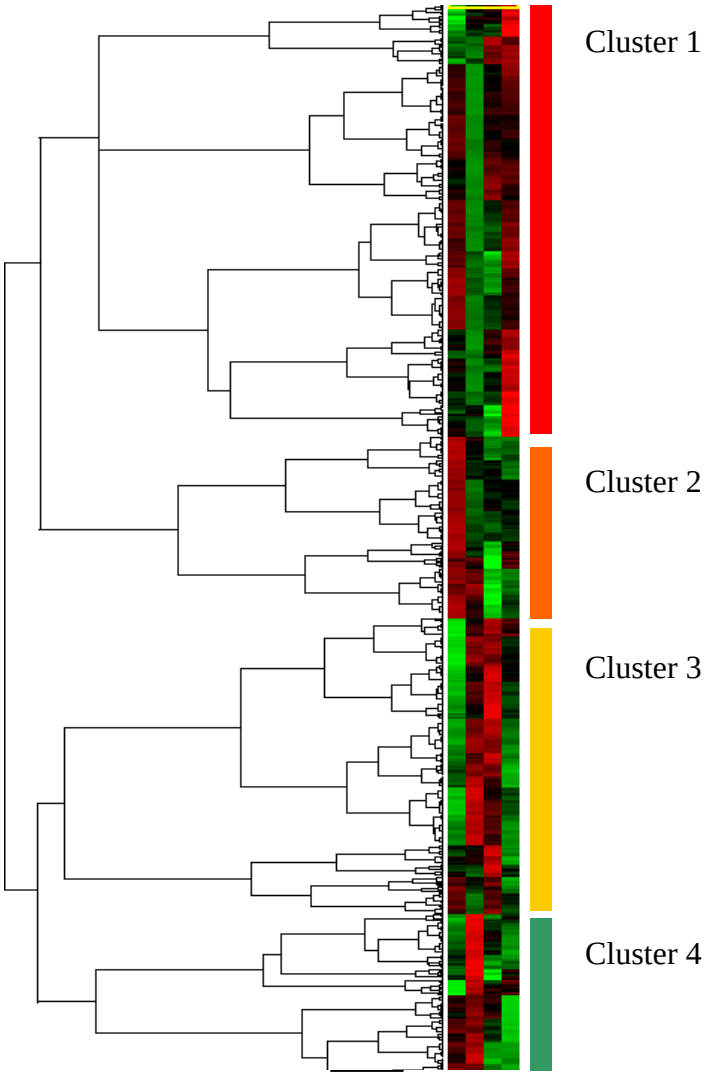


Figure 18. Upregulated genes in coronary arterial cells in response to angioplasty and/or brachytherapy. Genes with high expression levels were selected from the data set. The data set contains combined conditional experiments, including PTCA only group, Re188 only group, PTCA/Re188, and Re188/PTCA group (column from left to right). The hierarchical clustering tree was then generated Eisen et al. The leaves of the resulting tree (genes) have been rearranged by the software reported by Eisen et al. The gene expression profiles were then visualized as lines of color boxes. Each box corresponded to one data point, where red indicates an overexpressed gene (ratio of mean > 2), and green indicates an underexpressed gene (ratio of mean < 0).

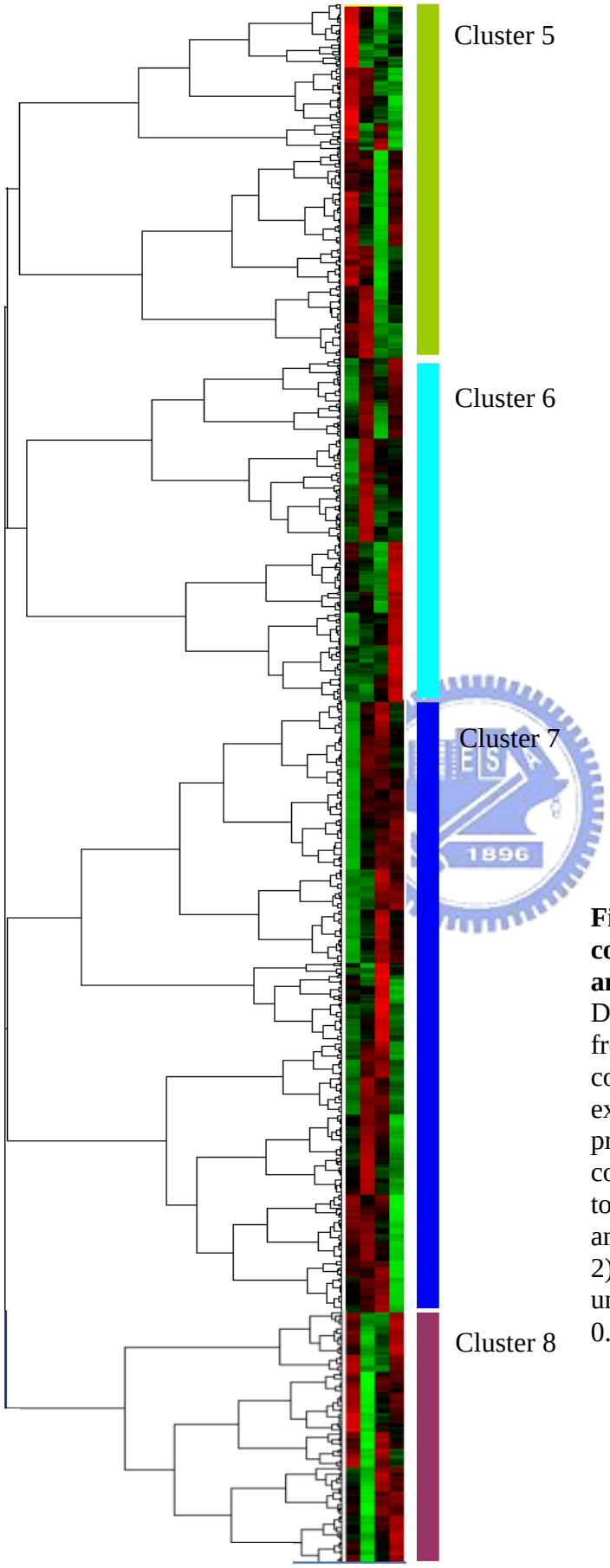
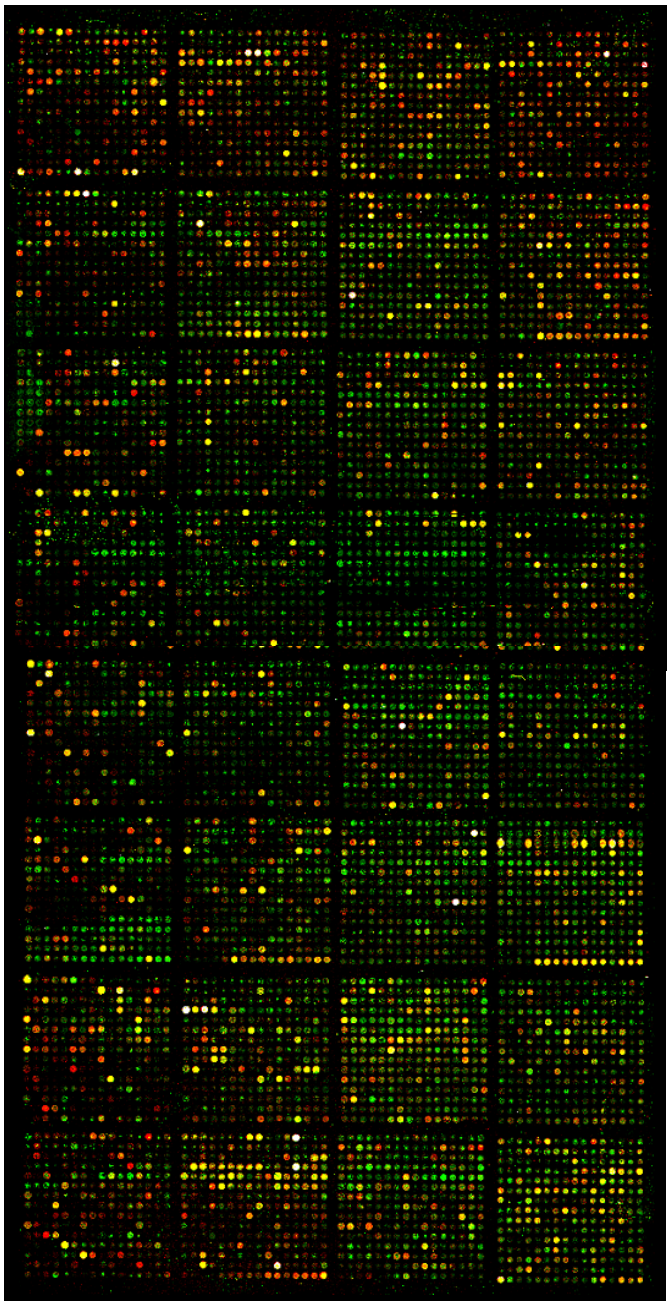


Figure 19. Downregulated genes in coronary arterial cells in response to angioplasty and/or brachytherapy. Downregulated genes were selected from the data set. The data set contains combined conditional experiments. The gene expression profles were then visualized as lines of colour boxes. Each box corresponded to one data point, where red indicates an overexpressed gene (ratio of mean > 2), and green indicates an underexpressed gene (ratio of mean < 0.5).

Figure 20. Image in cDNA microarrays after scanning in the PTCA group



The membrane shown was spotted with 7680 genes, hybridized to fluorescence-labeled cDNA, and exposed to GenePix 4000 for image capture. Red spots (ratio of mean > 2) representes up-regulated genes, green spots (ratio of mean < 0.5) representes downregulated genes in PTCA compared with normal vessel. Yellow spots were related to genes of which gene expression was similar (ratio of mean from 2 to 0.5).