APPLICATION OF cDNA MICROARRAY TECHNIQUE TO DETECTION OF GENE EXPRESSION IN HOST CELLS INFECTED WITH VIRUSES

J.H. NAM[1], C.H. YU[1], K.A. HWANG[1], J.S. KIM[2], S.H. AHN[2], J.Y. SHIN[1], W.Y. CHOI[1], Y.R. JOO[1], K.Y. PARK[1*]

[1]Department of Virology, Korean National Institute of Health, 5 Nokbun-dong, Eunpyung-gu, Seoul, 122-701, Korea;
[2]Genomictree, Taejeon, Korea

Received June 14, 2002; accepted August 20, 2002

Summary. – cDNA microarray technique was used to monitor changes in mRNA levels in cells after Hantaan virus (HTNV) infection. The values of the ratio of medians for HTNV and Japanese encephalitis virus (JEV) at the early stage of infection were compared and found similar, suggesting that the same or similar genes are associated with the early events of infection with either virus. The reproducibility of values of the "ratio of medians" for HTNV was examined. We found that applying cluster analysis to the gene expression data groups efficiently together genes with the same function. Therefore, in analyzing the effects of viral infection on host cells by the cDNA microarray technique, clustering data appear to be necessary for gaining biological meaning from a dump of gene expression profiles obtained from virus-infected cells.

Key words: cDNA microarray; cluster analysis; viral infection

Introduction

Recently, a cDNA microarray technique has been developed to survey the expression of many cellular genes simultaneously. This technique has been used to study the effects of viral infection on the expression of cellular genes.

Interactions between viral pathogens and host cells are very complex but well regulated. Viral pathogenesis is dependent upon several factors including virus-host interactions. However, each virus evolves its own particular strategy for manipulating the host, to optimize viral persistence and transmission, and detailed analyses of specific virus-host interactions have not yet clarified the unique mechanisms and molecules involved (Manger and Relman, 2000). The levels of cellular mRNAs may be affected by virus infection. Therefore, cellular mRNA levels may reflect (i) the pathways involved or the cellular factors that are induced or repressed during viral infection and (ii) the molecular mechanisms of pathogenesis. Standard methodologies of molecular biology including Northern blot analysis, S1 nuclease protection, differential display, sequencing cDNA libraries, and serial analysis of gene expression can be used to detect and quantify the cellular gene expression (Kim *et al.*, 2001). However, these methods are laborious and can focus on only small number of cellular gene products at a time. Recent development of the cDNA microarray techniques allows simultaneous measurement of expression of many cellular genes (Brown and Botstein, 1999; Duggan *et al.*, 1999). These techniques have been used to study the effects of various viruses on the expression of cellular genes (Zhu *et al.*, 1998; Geiss *et al.*, 2000; Bigger *et al.*, 2001).

In the present study, we used the cDNA microarray technique to investigate the expression of human lung cell (A549) genes as influenced by infection with HTNV (the test virus) and JEV (the control virus), because the human lungs are one of the main target organs of HTNV infection. A cluster analysis of the genome-wide expression data derived from cDNA microarray analysis uses standard statistical algorithms to arrange genes according to similarities in their patterns of

*Corresponding author. E-mail: Jnam66@yahoo.com; fax: +822-3801499.

Abbreviations: HTNV = Hantaan virus; JEV = Japanese encephalitis virus; p.i. = post infection; SDS = sodium dodecyl sulfate; SSC = saline sodium citrate buffer

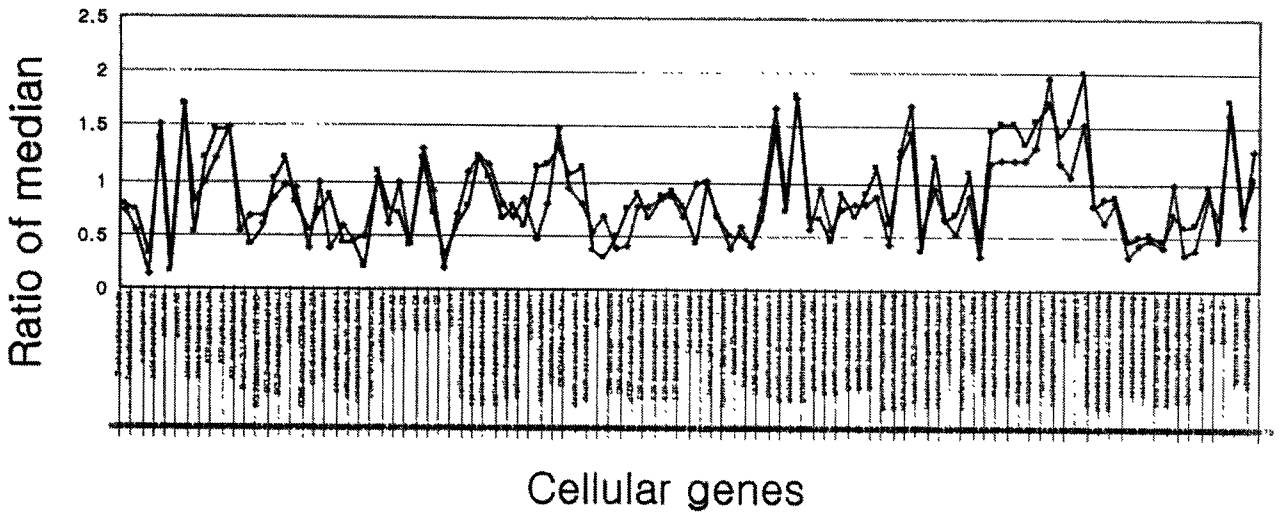

Fig. 1

Reproducibility of the ratio of median measurements for 100 cDNAs in the chip

The ratios for duplicate spots from two separate microarrays corresponding to 8 hrs p.i. were normalized. The normalized values (y axis) for 100 cDNAs (x axis) are superimposed on the line spot. Each data set is indicated by a differently colored line.

expression (Eisen *et al.*, 1998). Therefore, we applied the cluster analysis to cDNA microarray data obtained from HTNV infection. Furthermore, we tested the reproducibility of cDNA microarray data from repeated experiments with HTNV infection and compared the cDNA microarray data from HTNV and JEV infections.

Materials and Methods

Viruses and cells. Hantaan virus (HTNV) strain 76-118 and Japanese encephalitis virus (JEV) K94P05 were propagated in cultured Vero E6 cells (Lee *et al.*, 1978; Chung *et al.*, 1996).

Infectious virus titers (PFU) in the supernatant of cell cultures were determined as described earlier (Tanishita *et al.*, 1984). The A549 cell line, which retains many features of type II alveolar epithelial cells, including the synthesis of surfactant and cytoplasmic multilamellar inclusion bodies, was used. The cells were grown in Dulbecco's Modified Eagle's Medium (DMEM) supplemented with 10% of fetal bovine serum (FBS) in 5% CO_2 at 37°C.

Virus infection. A549 cells were infected with HTNV or JEV at a multiplicity of infection of 1 PFU/cell. Two hours post infection (p.i.) DMEM was added.

Total RNA was extracted at 2, 8, 24, and 72 hrs p.i. and purified (Lieber *et al.*, 1976).

cDNA microarray analysis. cDNA microarray slides were obtained from Genomictree (Daejeon, Korea). These consisted of silanized glass slides with the surface containing 480 different cDNAs corresponding to the genes for apoptosis, cell cycle, tumor suppressor and oncogenes. The hybridization was performed according to the manufacturer's protocol, described in detail at http://www.genomictree.com). Briefly, total cellular RNA (100 μg), prepared using the Trizol reagent (Gibco/BRL), was used as the template for the first-strand cDNA synthesis in a reaction with oligo(dT) primer (5'-T_{20}VN-3') or Cy3-dUTP or Cy5-dUTP (NEN), and Superscript™ II reverse transcriptase (Gibco/BRL). The labeled cDNA probe was then purified through a Microcon-30 column (Amicon), resuspended in 80 μl of hybridization solution (3′ saline sodium citrate buffer (SSC) with 0.3% sodium dodecyl sulfate (SDS)). The probe was then denatured at 100°C for 2 mins and applied to the DNA chip at 65°C for 16 hrs in a humidified chamber. Finally, the slide was washed in 2 × SSC for 2 mins, 0.1 × SSC with 0.1% SDS for 5 mins, and 0.1 × SSC for 5 mins, then spun-dried and scanned in a GenePix 4000B scanner (Axon Instruments Inc.) at room temperature.

The scanned image was analyzed using a GenePix Pro 3.0 (Axon Instruments Inc.). Signal intensity values were determined by subtracting the median background value from the median intensity value for each spot. Expression values were normalized by a single multiplicative normalization factor and applied to all Cy5/Cy3 ratios so that the median normalized Cy5/Cy3 ratio became 1.0. All values were calculated for two individual samples under the same conditions. Some values were excluded due to a relative error rate of more than 0.5-fold change. Microarray data were clustered using the J-express 2.1 from http://ihome.cuhk.edu.hk/~b400559/arraysoft.html. Clustering was performed by the complete linkage method. The aim of this algorithm was to compute a dendrogram that would assemble all elements into a single tree. The clustering data were employed for construction of gene graphs.

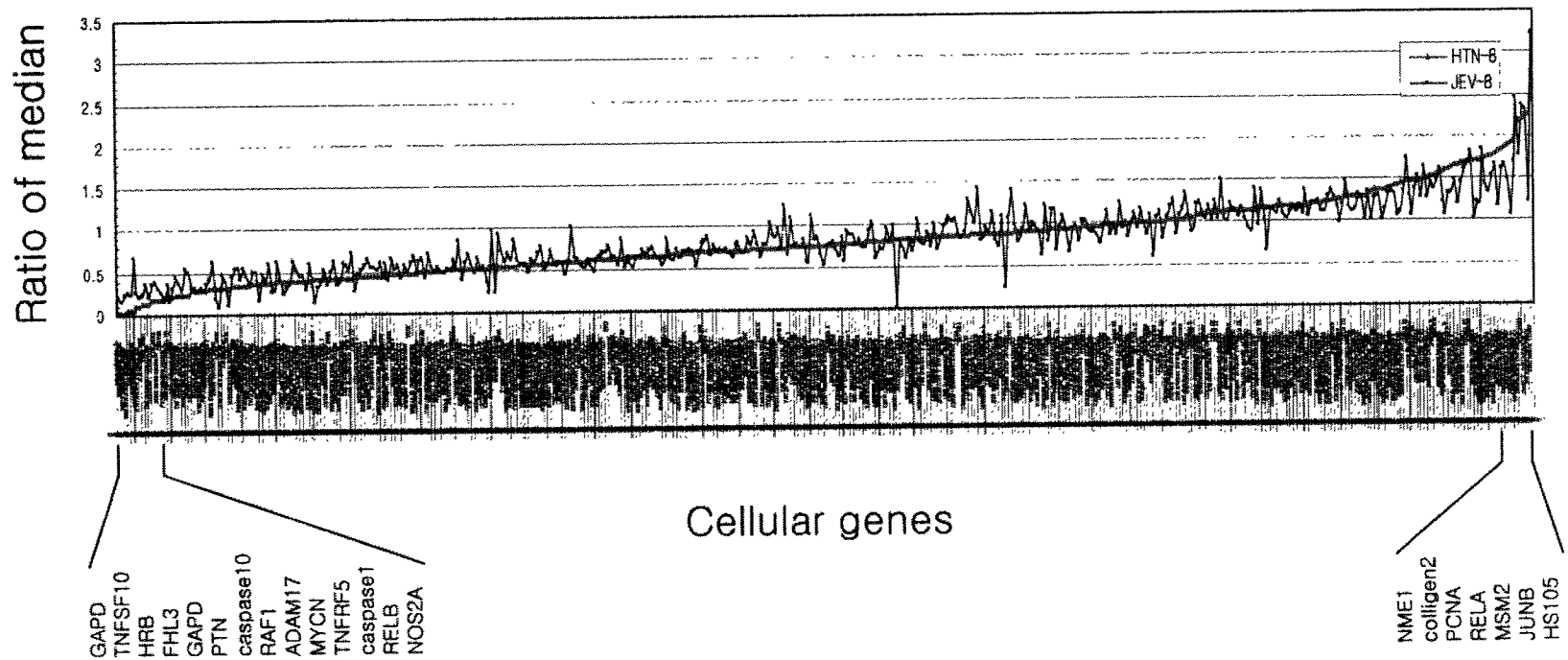

Fig. 2

Comparison of ratio of medians values for HTNV and JEV at 8 hrs p.i.

The ratio of median values for HTNV-infected cells (red) were arranged from low to high, then those of JEV-infected cells (blue) were matched with those for HTNV for the same genes on the graph. Some cellular genes at both extremes are shown on the graph. GAPD = glyceraldehyde-3-phosphate dehydrogenase; TNFSF10 = tumor necrosis factor (ligand) superfamily member 10; HRB = HIV-1 Rev-binding protein; FHL3 = four and a half LIM domains 3; GAPD = glyceraldehyde 3 phosphate dehydrogenase; PTN = pleiotrophin (heparin-binding growth factor 8, neurite growth-promoting factor 1); RAF1 = v-raf-1, murine leukemia virus oncogene homolog 1; ADAM17 = a disintegrin and metalloproteinase domain 17 (tumor necrosis factor alpha-converting enzyme); MYCN = v-myc, myelocytomatosis virus-related oncogene; TNFSF5 = tumor necrosis factor (ligand) superfamily member 5; RELB = v-rel, reticuloendotheliosis virus oncogene homolog B; NOS2A = nitric oxide synthase 2A (inducible, hepatocytes); NME1 = non-metastatic cells 1 protein (NM23A); PCNA = proliferating cell nuclear antigen; RELA = v-rel, reticuloendotheliosis virus oncogene homolog A; MSH2 = mutS homolog 2 (colon cancer); JUNB = jun B proto-oncogene; HS105 = heat-shock protein 105.

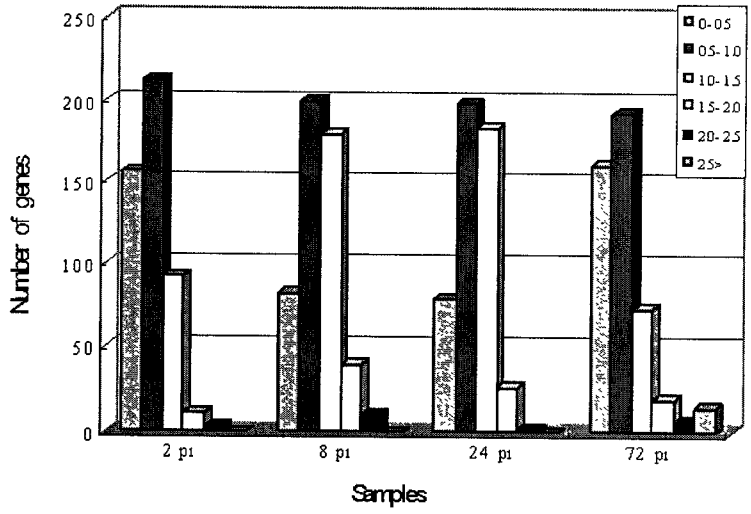

Fig. 3

Distribution of ratio of medians values for HTNV at 2, 8, 24, and 72 hrs p.i.

Results and Discussion

Reproducibility of cDNA microarray data

Although the cDNA microarray technique has many advantages, one problem that arises in data analysis is that the value of the ratio of medians can fluctuate, depending on several factors. Therefore, all microarray data must be normalized to "housekeeping" genes and be adjusted with a "normalization factor" for the ratio of medians (Geiss *et al.*, 2000). We compared the cellular mRNA levels in two independent samples under the same conditions. Eighteen per cent of 480 genes had a relative error of more than 0.5-fold change of the ratio of medians. This result is similar to that from a cDNA microarray analysis of an HIV-1 infection, which showed a 15% relative error for the most highly expressed genes (Geiss *et al.*, 2000). Fig. 1 shows the reproducibility of the ratio of medians for cellular genes. According to these data, cDNA microarray experiments should be repeated more than twice to ensure that the data are reliable.

Comparison of cDNA microarray data from cells infected with HTNV and JEV

Total RNA from the cells infected with HTNV or JEV were isolated at 8 hrs p.i. and hybridized to cDNA microarray slides. The values of the ratio of medians were normalized and these values for the HTNV- and JEV-infected cells were compared graphically (Fig. 2). The values for HTNV were arranged from low to high, and the values for JEV were then matched with those for HTNV for the same genes. The patterns of cellular gene expression for HTNV and JEV were similar, suggesting that, during the early stages of viral infection, both HTNV and JEV altered the expression of the same cellular genes engaged in host defense, transcription, and cell proliferation.

Distribution of the ratio of medians values

A value of 1.0 for the ratio of medians indicates that the values for Cy5 and Cy3 are the same, so that the expression of that gene is unchanged in the cells after virus infection. A ratio of medians value lower than 0.5 means that the expression of that gene is down-regulated more than twofold, and a ratio of medians value higher than 2.0 means that the gene is up-regulated more than twofold. Most genes examined had ratio of median values within the range of 0.5–2.0 (Fig. 3). Within 24 hrs p.i. only 15 genes had values higher than 2.0. In contrast, 21 genes had values higher than 2.0 at 72 hrs p.i. Fifteen of these had values higher than 2.5. However, 156, 82, 79, and 160 of the 480 genes examined exhibited twofold down-regulation at 2, 8, 24, and 72 hrs p.i., respectively. This suggests that many cellular genes are down-regulated in response to viruses at an early stage of infection, and that the expression of some cellular genes is then up-regulated to defend the host cell or control viral replication and assembly. However, because viral pathogenesis and its interpretation are very complex, this hypothesis requires further investigation.

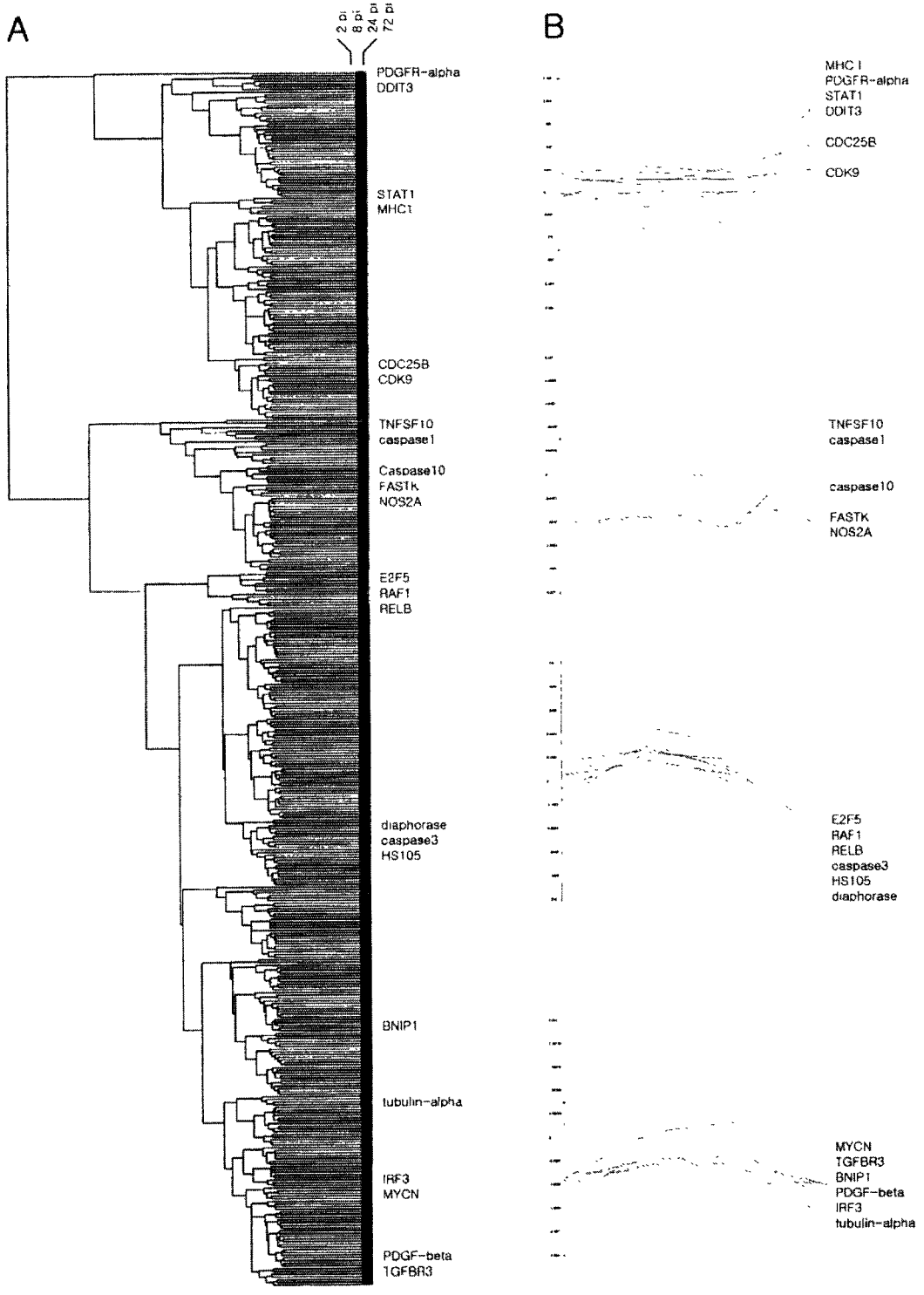

Fig. 4

Clustered display (A) and gene graph (B) of the data for HTNV-infected cells

The data were collected using a cDNA microarray with elements representing approximately 480 distinct human genes. Each gene is represented by a single row of colored boxes; each time point is represented by a single column. MHC I = major histocompatibility protein class I; PDGFR-alpha = platelet-derived growth factor receptor alpha; STAT1 = signal transducer and activator of transcription of 1.91 K; DDIT3 = DNA-damage-inducible transcript 3; CDC25B = cell division cycle 25B; CDK9 = cyclin-dependent kinase 9 (CDC2-related kinase); TNFSF10 = tumor necrosis factor (ligand) superfamily member 10; FASTK = Fas-activated serine/threonine kinase; NOS2A = nitric oxide synthase 2A (inducible, hepatocytes); E2F5 = E2F transcription factor 5, RELB = v-rel, reticuloendotheliosis virus oncogene homolog B; HS105 = heat-shock protein 105; MYCN = v-myc, myelocytomatosis virus-related oncogene; TGFBR3 = transforming growth factor beta receptor III (betaglycan of 300 K); BNIP1 = BCL2, adenovirus E1B 19 K-interacting protein 1; PDGF-beta = platelet-derived growth factor beta; IRF3 = interferon regulatory factor 3.

Clustering of genes

To evaluate which cellular genes are affected by factors such as viral infection, stress or chemicals, the obvious first step is to identify genes with significantly different patterns of expression in two differently treated samples or at different times after a specific treatment, preferably utilizing the full potential of genome-scale experiments (Eisen *et al.*, 1998). Several methods for such an analysis have been developed, one of which is the clustering method. In this study, we constructed a cluster tree from the expression patterns of cellular genes during the time-course of a viral infection (Fig. 4A). Examples of these gene expression profiles were also shown in Fig. 4B. The relationships between genes are represented by a tree, the branch lengths of which reflect the degree of similarity between the expression profiles of genes, as assessed by a pairwise similarity function. This clustering of genes that have unrelated sequences but similar patterns of expression indicates that the genes with similar functions clustered tightly together.

Concluding remarks

The cDNA microarray technique is a high-throughput approach, and has, therefore, become very useful in the field of biology. In particular, interactions between viruses and host cells, molecular pathogenesis of the virus or the resistance mechanisms of the infected host cells are easily examined by this technique. In the present study, we demonstrated how the cDNA microarray technique can be applied to viral infections. First, to exclude the data noise, cDNA microarray data must originate from at least two independent biological samples processed under the same conditions. Second, although the noise in the microarray data for a single gene is considerable, clustering analysis can be particularly useful insofar as it can group genes with similar functions, even when the genes in a particular group are not obviously biologically related. In this way, similarities in gene expression profiles may present an easy way to provisionally attribute functions on a genomic scale. This conclusion is in agreement with an earlier report by Eisen *et al.* (1998). Therefore, the time-course approach and the cluster analysis are valuable tools of investigation of interactions between viruses and host cells using the cDNA microarray technique.

Acknowledgement. J.H. Nam, C.H. Yu, and K.A. Hwang contributed equally to this work. This study was supported by the Korean National Institute of Health.

References

Bigger CB, Brasky KM, Lanford RE (2001): DNA microarray analysis of chimpanzee liver during acute resolving hepatitis C virus infection. *J. Virol.* **75**, 7059–7066.

Brown PO, Botstein D (1999): Exploring the new world of the genome with DNA microarrays. *Nat. Genet.* **21** (Suppl. 1), 33–37.

Chung YJ, Nam JH, Ban SJ, Cho HW (1996): Antigenic and genetic analysis of Japanese encephalitis virus isolated from Korea. *Am. J. Trop. Hyg.* **55**, 91–97.

Duggan DJ, Bittner M, Chen Y, Meltzer P, Trent JM (1999): Expression profiling using cDNA microarrays. *Nat. Genet.* **21** (Suppl. 1), 10–14.

Eisen MB, Spellman PT, Brown PO, Botstein D (1998): Cluster analysis and display of genome-wide expression patterns. *Proc. Natl. Acad. Sci. USA* **95**, 14863–14868.

Geiss GK, Bumgarner RE, An MC, Agy MB, Wout AB, Hammersmark E, Carter VS, Upchurch D, Mullins JI, Katze MG (2000): Large-scale monitoring of host cell gene expression during HIV-1 infection using cDNA microarrays. *Virology* **266**, 8–16.

Kim JH, Kim HY, Lee YS (2001): A novel method using edge detection for signal extraction from cDNA microarray image analysis. *Exp. Mol. Med.* **33**, 83–88.

Lee HW, Lee PW, Johnson KM (1978): Isolation of the etiologic agent of Korean hemorrhagic fever. *J. Infect. Dis.* **137**, 298–302.

Lieber M, Smith B, Szakal A, Nelson-Rees W, Todaro G (1976): A continuous tumor cell line from a human lung carcinoma with properties of type II alveolar epithelial cells. *Int. J. Cancer* **17**, 62–70.

Manger ID, Relman DA (2000): How the host 'sees' pathogens: global gene expression responses to infection. *Curr. Opin. Immunol.* **12**, 215–218.

Tanishita O, Takahashi Y, Okuno Y, Yamanishi K, Takahashi M (1984): Evaluation of focus reduction neutralization test with peroxidase–antiperoxidase staining technique for hemorrhagic fever with renal syndrome virus. *J. Clin. Microbiol.* **20**, 1213–1215.

Zhu H, Cong J-P, Mamtora G, Gingeras T, Shenk T (1998): Cellular gene expression altered by human cytomegalovirus: global monitoring with oligonucleotide arrays. *Proc. Natl. Acad. Sci. USA* **95**, 14470–14475.