trials may conclusively answer these questions. However, they may need hundreds or thousands of patients and many years to obtain statistical differences.[16] In addition, our PTA procedures seem to be evolving, as shown in this study. If we enter into a head-to-head trial between surgical resection and PEIT and there is a 5%–8% survival difference, as suggested in this study, the conclusion obtained from comparing RFA and surgical resection or PEIT and resection might differ.

To obtain scientific evidence, a randomized controlled study design is absolutely essential. However, we face patients who seek the best and safest treatment currently in use in clinical practice. We are obliged to offer all available data to patients so they can select the treatment modality that best suits them. In that context, we definitely need a staging and scoring system that provides good discriminating prognosis. The classification of Okuda et al[17] has been widely used. This classification is based on 2 features: tumor and background liver damage. With the classification of Okuda et al[17] as a milestone, several other staging systems have been proposed.[18–20] As part of a working group on the classification, cause, and surveillance of HCC for the World Congress of Gastroenterology held in Bangkok, Thailand, in 2002 (principal investigator, M.O.), we analyzed various staging and scoring systems.[21] The staging of Okuda et al[17] was used worldwide and certainly provided discriminating prognosis in areas in which the prognosis of patients with HCC is poor and median survival is only 1–2 years. However, in countries where much longer survivals are expected, other systems, for example, the CLIP score system, are superior in prognostic discriminating ability because the grading for tumor character is more precise and compatible with current imaging diagnoses compared with that of Okuda et al[17], which was based on autopsy gross findings. In our study, CLIP score and prognosis by PEIT were closely related, confirming the clinical utility of this scoring system (Figure 7). Thus, one can provide patients with information on the possibility of 5-year survival based on the score. For example, if a patient has a CLIP I score and 5-year survival by PEIT of 59% and by surgical resection of 62%, the patient might choose PTA.

If cohorts can be well matched using these prognostic factors, survival could be compared according to each treatment method within each group, each hospital, or even each country. To accomplish this, parameters to be incorporated into this scoring system should be simple and easily applicable to different institutions or geographic areas. We now propose a Tokyo score that needs only bilirubin and serum albumin levels and imaging data on tumor number and size.[22] This staging and scoring system may provide better clues and information for patients to choose treatment once data are obtained in a large number of patients treated by different modalities and validated in different institutions.

We used PTA to treat our patients, the majority of whom have tumors small enough to be treatable. Our "local" recurrence (the tumor recurs close to the targeted primary tumor) is now as low as 1.5%/y. Therefore, we believe we can control most tumor nodules by means of PTA. However, recurrence in the remote area of the liver (de novo recurrence) reaches almost 70%–80% in 5 years (data not shown). Therefore, we need a means to prevent the recurrence that hampers long-term survival of patients treated by means of PTA. This is particularly important among those infected with HCV.

Recently, we were able to show that the eradication of HCV by interferon in patients who were treated by PTA ("secondary" prevention) resulted in a 5-year survival rate of 78%, whereas the rate in those who failed to respond to treatment was 48%, similar to the figure (45%) for ablation alone.[23] These data clearly indicate there is definite need for expertise to effectively treat tumor nodules, but that alone is insufficient, and therefore better means of treatment of HCV is needed that may eventually resolve background fibrosis and decrease de novo carcinogenesis. The combination of antiviral treatment and PTA may prove to be effective and accomplish the same outcome as liver transplantation, both of which might cure the 2 diseases of chronic hepatitis C and HCC.[24]

References

1. Omata M, Yoshida H. Prevention and treatment of hepatocellular carcinoma. Liver Transpl 2004;10(Suppl):S111–S114.
2. Shiratori Y, Shiina S, Imamura M, Kato N, Kanai F, Okudaira T, Teratani T, Tohgo G, Toda N, Ohashi M, Ogura K, Niwa Y, Kawabe T, Omata M. Characteristic difference of hepatocellular carcinoma between hepatitis-B and -C viral infection in Japan. Hepatology 1995;22:1027–1033.
3. Takano S, Yokosuka O, Imazeki F, Tagawa M, Omata M. Incidence of hepatocellular carcinoma in chronic hepatitis B and C: a prospective study of 251 patients. Hepatology 1995;21:650–655.
4. Koike Y, Shiratori Y, Sato S, Obi S, Teratani T, Imamura M, Hamamura K, Imai Y, Yoshida H, Shiina S, Omata M. Risk factors for recurring hepatocellular carcinoma differ according to infected hepatitis virus—an analysis of 236 consecutive patients with a single lesion. Hepatology 2000;32:1216–1223.
5. Omata M, Yoshida H. Resolution of liver cirrhosis and prevention of hepatocellular carcinoma by interferon therapy against chronic hepatitis C. Scand J Gastroenterol 2003;237(Suppl):S47–S51.
6. Shiina S, Tagawa K, Unuma T, Takanashi R, Yoshiura K, Komatsu Y, Hata Y, Niwa Y, Shiratori Y, Terano A. Percutaneous ethanol injection therapy for hepatocellular carcinoma. A histopathologic study. Cancer 1991;68:1524–1530.

7. Shiina S, Imamura M, Omata M. Percutaneous ethanol injection therapy (PEIT) for malignant liver neoplasms. Semin Interv Radiol 1997;14:295–303.
8. Shiina S, Hata Y, Niwa Y, Komatsu Y, Tanaka T, Yoshiura K, Hamada E, Ohshima M, Mutoh H, Kurita M. Multiple-needle insertion method in percutaneous ethanol injection therapy for liver neoplasms. Gastroenterol Jpn 1991;26:47–50.
9. Shiina S, Teratani T, Obi S, Hamamura K, Koike Y, Omata M. Nonsurgical treatment of hepatocellular carcinoma: from percutaneous ethanol injection therapy and percutaneous microwave coagulation therapy to radiofrequency ablation. Oncology 2002; 62(Suppl):S64–S68.
10. Rossi S, Di Stasi M, Buscarini E, Cavanna L, Quaretti P, Squassante E, Garbagnati F, Buscarini L. Percutaneous radiofrequency interstitial thermal ablation in the treatment of small hepatocellular carcinoma. Cancer J Sci Am 1995;1:73–81.
11. Mazzaferro V, Regalia E, Doci R, Andreola S, Pulvirenti A, Bozzetti F, Montalto F, Ammatuna M, Morabito A, Gennari L. Liver transplantation for the treatment of small hepatocellular carcinomas in patients with cirrhosis. N Engl J Med 1996;334:693–699.
12. Ebara M, Ohto M, Sugiura N, Kita K, Yoshikawa M, Okuda K, Kondo F, Kondo Y. Percutaneous ethanol injection for the treatment of small hepatocellular carcinoma. Study of 95 patients. J Gastroenterol Hepatol 1990;5:616–626.
13. Shiina S, Tagawa K, Niwa Y, Unuma T, Komatsu Y, Yoshiura K, Hamada E, Takahashi M, Shiratori Y, Terano A. Percutaneous ethanol injection therapy for hepatocellular carcinoma: results in 146 patients. AJR Am J Roentgenol 1993;160;1023–1028.
14. Lencioni R, Pinto F, Armillotta N, Bassi AM, Moretti M, Di Giulio M, Marchi S, Uliana M, Della Capanna S, Lencioni M, Bartolozzi C. Long-term results of percutaneous ethanol injection therapy for hepatocellular carcinoma in cirrhosis: a European experience. Eur Radiol 1997;7:514–519.
15. Livraghi T, Giorgio A, Marin G, Salmi A, de Sio I, Bolondi L, Pompili M, Brunello F, Lazzaroni S, Torzilli G. Hepatocellular carcinoma and cirrhosis in 746 patients: long-term results of percutaneous ethanol injection. Radiology 1995;197:101–108.
16. Hoshida Y, Shiratori Y, Omata M. Difficulties in conducting controlled trials in radical therapies for non-advanced hepatocellular carcinoma [letter]. Hepatology 2000;32:877–880.
17. Okuda K, Ohtsuki T, Obata H, Tomimatsu M, Okazaki N, Hasegawa H, Nakajima Y, Ohnishi K. Natural history of hepatocellular carcinoma and prognosis in relation to treatment. Study of 850 patients. Cancer 1985;56:918–928.
18. The Cancer of the Liver Italian Program (CLIP) Investigators. A new prognostic system for hepatocellular carcinoma: a retrospective study of 435 patients: the Cancer of the Liver Italian Program (CLIP) investigators. Hepatology 1998;28:751–755.
19. Llovet JM, Bru C, Bruix J. Prognosis of hepatocellular carcinoma: the BCLC staging classification. Semin Liver Dis 1999;19:329–338.
20. Kudo M, Chung H, Osaki Y. Prognostic staging system for hepatocellular carcinoma (CLIP score): its value and limitations, and a proposal for a new staging system, the Japan Integrated Staging score (JIS score). J Gastroenterol 2003;38:207–215.
21. Omata M, Dan Y, Daniele B, Plentz R, Rudolph KL, Manns M, Piratvisuth T, Chen DS, Tateishi R, Chutaputti A. Clinical features, etiology, and survival of hepatocellular carcinoma among different countries. J Gastroenterol Hepatol 2002;17(Suppl):S40–S49.
22. Tateishi R, Yoshida H, Shiina S, Imamura H, Hasegawa K, Teratani T, Obi S, Sato S, Koike Y, Hamamura K, Akamatsu M, Fujishima T, Kanda M, Kimbara T, Ishikawa T, Kawabe T, Makuuchi M, Omata M. Proposal of a new prognostic model for hepatocellular carcinoma—Tokyo score [abstract]. Gastroenterology 2003;124:708A–708A.
23. Shiratori Y, Shiina S, Teratani T, Imamura M, Obi S, Sato S, Koike Y, Yoshida H, Omata M. Interferon therapy after tumor ablation improves prognosis in patients with hepatocellular carcinoma associated with hepatitis C virus. Ann Intern Med 2003;138: 299–306.
24. Omata M, Okita K. Therapy for viral hepatitis and prevention of hepatocellular carcinoma. Tokyo: Springer, 2004:1–295.

Address requests for reprints to: Masao Omata, MD, Department of Gastroenterology, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo, Japan. e-mail: omata-2im@h.u-tokyo.ac.jp; fax (81) 3-5800-8808.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Biochemical and Biophysical Research Communications 318 (2004) 461–469
BBRC
www.elsevier.com/locate/ybbrc

Interaction of the hepatitis B virus X protein (HBx) with heat shock protein 60 enhances HBx-mediated apoptosis

Yasuo Tanaka,[a] Fumihiko Kanai,[a,b,*] Takayuki Kawakami,[a] Keisuke Tateishi,[a] Hideaki Ijichi,[a] Takao Kawabe,[a] Yoshihiro Arakawa,[b] Takao Kawakami,[c,1] Toshihide Nishimura,[c,1] Yumiko Shirakata,[d] Katsuro Koike,[d] and Masao Omata[a,b]

[a] *Department of Gastroenterology, Graduate School of Medicine, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8655, Japan*
[b] *Clinical Research Center, University of Tokyo Hospital, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8655, Japan*
[c] *Structural and Analytical Sciences Unit, Chemistry Department, Research Division, GlaxoSmithKline K.K. Tsukuba Research Laboratories, 43 Wadai, Tsukuba, Ibaraki 300-4247, Japan*
[d] *Department of Gene Research, The Cancer Institute of the Japanese Foundation for Cancer Research, 1-37-1 Kami-Ikebukuro, Toshima-ku, Tokyo 170-8455, Japan*

Received 17 February 2004
Available online 20 April 2004

Abstract

Understanding the function of the hepatitis B virus X protein (HBx) is fundamental to elucidating the underlying mechanisms of hepatitis and hepatocarcinogenesis caused by hepatitis B virus (HBV) infection. We identified heat shock protein 60 (Hsp60) as a novel cellular target of HBx by the combination of affinity purification and mass spectrometry. Physical interaction between HBx and Hsp60 was confirmed by standard immunoprecipitation and immunoblot methods. Analysis of HBx deletion constructs showed that amino acids 88–117 of HBx were responsible for the binding to Hsp60. Confocal laser microscopy demonstrated that HBx and Hsp60 colocalized in mitochondria. Furthermore, terminal deoxynucleotidyl transferase-mediated dUTP end labeling (TUNEL) revealed that the introduction of Hsp60 into cells facilitated HBx-induced apoptosis. These findings suggest the importance of the molecular chaperon protein Hsp60 to the function of HBV viral proteins.

Keywords: Hepatitis B virus; Hepatitis B virus X protein; Heat shock protein 60; Affinity purification; Mass spectrometry; Apoptosis

Hepatitis B virus (HBV), a member of the hepadnavirus family, is closely associated with the development of acute and chronic hepatitis and hepatocarcinogenesis [1]. Chronic infection with HBV is a leading cause of cirrhosis and hepatocellular carcinoma. The number of patients with persistent HBV infection exceeds 350 million worldwide and is increasing [2]. However, an effective treatment for chronic HBV infection is not yet established [3]. HBV has four open reading frames (ORFs) that encode viral proteins. One ORF, named X, encodes the hepatitis B virus X protein (HBx), which consists of 154 amino acids. This protein is required for viral infection and replication [4,5], and is considered a major factor in hepatocarcinogenesis [6].

HBx has been proposed to have many cellular functions and has been reported to interact with a number of host factors, including transcriptional factors and molecules involved in intracellular signaling and apoptosis [7]. HBx binds basal transcription factors, such as RNA polymerase subunit RPB5 [8], and acts as a transcriptional coactivator. HBx also has been reported to enhance transcription by directly interacting with factors such as Smad4 in TGFβ signaling [9]. Furthermore, HBx interacts with a tumor suppressor gene, p53, and inhibits p53 entry into the nucleus [10]. HBx also binds human voltage-dependent anion channel 3 (HVDAC3),

* Corresponding author. Fax: +81-3-3814-0021.
E-mail address: kanaif-int@h.u-tokyo.ac.jp (F. Kanai).
[1] Present address: Clinical Proteome Center, Tokyo Medical University, Shinjuku Sumitomo Bldg. 3F, 2-6-1 Nishi-shinjuku, Shinjuku-ku, Tokyo 163-0203, Japan.

0006-291X/$ - see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.bbrc.2004.04.046

a mitochondrial protein, and alters the mitochondrial transmembrane potential [11]. In the apoptotic pathway, the function of HBx is still under debate. Recently, it was reported that HBx interacts with c-FLIP and abrogates its apoptosis-inhibition function [12]. Another group revealed that HBx binds the survivin–HBXIP (hepatitis B X-interacting protein) complex and suppresses caspase activation [13].

Despite these reports, a precise role of HBx in HBV infection remains unknown. In the present study, we performed affinity purification and mass spectrometry to identify Hsp60 as a novel HBx-binding protein. Hsp60 is a molecular chaperon that assists with the correct folding of proteins, facilitates the proteolytic degradation of misfolded or denatured proteins, and stabilizes unfolded proteins [14]. Our data suggest that the interaction between Hsp60 and HBx provides a mechanism for the control of HBx's pro-apoptotic function.

Materials and methods

Cell lines and transfection. A human primary hepatocyte cell culture was obtained from Applied Cell Biology Research Institute (Kirkland, WA) and was maintained in CS-C Complete Medium (Cell Systems, Kirkland, WA). The human embryonic kidney cell line 293 and the human hepatoma cell line Huh7 were purchased from the Riken Cell Bank (Tsukuba Science City, Japan) and were maintained in Dulbecco's modified Eagle's medium (Sigma, St. Louis, MO) containing 10% heat-inactivated fetal bovine serum. Transfection was performed with Effectene transfection reagent (Qiagen, Hilden, Germany) according to the manufacturer's instructions.

Antibodies. Mouse monoclonal anti-HA antibody (12CA5) and mouse monoclonal anti-GFP antibody were obtained from Roche Diagnostics (Indianapolis, IN) and Santa Cruz Biotechnology (Santa Cruz, CA), respectively. Mouse monoclonal anti-Hsp60 antibody was purchased from Stressgen (Victoria, Canada). Rabbit polyclonal anti-HBx antibody (U23) was described previously [15]. Alexa Fluor 488 goat anti-mouse IgG secondary antibody and Texas red conjugated goat anti-mouse IgG secondary antibody were obtained from Molecular Probes (Eugene, OR) and Southern Biotechnology Associates (Birmingham, AL), respectively. Anti-mouse antibody and anti-rabbit antibody conjugated with horseradish peroxidase (HRP) were purchased from Amersham Biosciences (Uppsala, Sweden).

Plasmids. The EGFP-fused HBx and its deletion constructs, pEGFP-X, pEGFP-X(1–117), pEGFP-X(1–87), and pEGFP-X(68–117), were described previously [16]. The HBx deletion fragment XΔ(5–87) was derived from pHBVX-dAval/*Eco*II [17], blunted by Klenow fragment, and cloned into pEGFP-C1 (Clontech, Palo Alto, CA), resulting in pEGFP-XΔ(5–87). The pCMV-HA-X, an N-terminal HA-tagged HBx expression plasmid, was described previously [18]. Using pCMV-HA-X as a template, HBx was amplified by PCR. The amplified fragment with the HA-tag at the N-terminus of HBx(1–87) was cloned into the pCMVβ vector (Clontech), producing pCMV-HA-X(1–87). The Hsp60 expression plasmid, pCMV-Hsp60, was provided by Dr. Naohiko Seki (Chiba University). Using pCMV-Hsp60 as a template, the HA-tag was inserted at the C-terminus of Hsp60 by PCR and cloned into the pcDNA3 (Invitrogen, Carlsbad, CA) to create pcDNA3-Hsp60-HA. HBx was amplified by PCR and cloned into the pGEX-4T-1 (Amersham Biosciences), resulting in pGEX-HBx. All constructs were verified by DNA sequencing.

Affinity purification. The preparation of bead-immobilized GST fusion proteins was performed as described [19]. In brief, *Escherichia coli* BL21(DE3) (Stratagene, La Jolla, CA) that had been transformed by pGEX-HBx or pGEX-4T-1 were induced with 1 mM IPTG at 37 °C for 3 h, pelleted, and lysed in phosphate-buffered saline (PBS) containing 5 mM EDTA, 1 mM DTT, 1 mM PMSF, 100 nM pepstatin, and 1% Triton X-100. The lysates were sonicated and clarified by centrifugation (8000*g* for 10 min at 4 °C). Fusion proteins were purified from the cleared bacterial lysates using glutathione–Sepharose 4B (Amersham Biosciences).

Subconfluent human primary hepatocytes (approximately 3×10^7) were lysed in lysis buffer I [20 mM phosphate buffer, pH 7.4, 1% Nonidet P-40, 5% glycerol, 10 mM β-mercaptoethanol, 150 mM NaCl, and the protease inhibitor cocktail Complete Mini (Roche, Mannheim, Germany)] and centrifuged at 15,000*g* for 20 min at 4 °C. Lysates (60 mg) were incubated with bead-immobilized GST-HBx or GST control beads in lysis buffer I for 4 h at 4 °C. The beads were washed with lysis buffer and bound proteins were eluted by boiling in SDS sample buffer. Eluted proteins were separated by 7.5% SDS–PAGE and visualized with silver staining according to the procedure of Blum et al. [20].

Protein identification by mass spectrometry. Protein bands were excised from the gel plate and subjected to in-gel tryptic digestion according to the procedure of Shevchenko et al. [21]. Peptide mixtures from the digested proteins were extracted from the gel pieces, and the recovered peptides were analyzed by nanoflow liquid chromatography (nanoLC)–nanoelectrospray ionization (nanoESI)–iontrap MS/MS as previously described [22]. All MS/MS spectra were searched using the MOWSE scoring algorithm, MASCOT, against the MSDB database (ftp://ftp.ncbi.nih.gov/repository/MSDB/msdb.nam), Release 20032804, with the following criteria: tryptic digestion (hydrolysis of the peptide bonds following lysine and arginine residues) and variable modifications of methionine (oxidation, +16.0) and cysteine (*S*-carboxyamidomethylation, +57.0).

Immunoprecipitation and immunoblot. Immunoprecipitation and immunoblot methods were performed as described [19]. First, 48 h after transfection, 2×10^6 transfected 293 cells were lysed in lysis buffer I, and cellular proteins were immunoprecipitated with antibody and protein G–Sepharose (Zymed, San Francisco, CA). After the beads were washed, the bound proteins were eluted, separated by 10% SDS–PAGE, and electrophoretically transferred to the PVDF Hybond-P (Amersham Biosciences). The membrane was probed with primary antibody (anti-GFP at 1:500, anti-HA at 1:1000, and anti-Hsp60 at 1:5000) and with secondary antibody (anti-mouse antibody and anti-rabbit antibody conjugated with HRP at 1:1000), and the HRP signal was detected with enhanced chemiluminescence ECL-plus (Amersham Biosciences).

Immunofluorescence microscopy. Huh7 cells were seeded in a slide flask (Nalge Nunc, Naperville, IL) at a density of 3×10^4 cells/cm^2, and 24 h later cells were transfected with the expression vectors. For the staining of mitochondria, cells were incubated in medium containing 500 nM MitoTracker Red CMXRos (Molecular Probes) for 45 min. Forty eight hours after transfection, cells were fixed with 4% paraformaldehyde for 30 min at room temperature, rinsed with PBS, and permeabilized with a 0.1% Triton X-100/0.1% sodium citrate solution for 2 min on ice. After two rinses with PBS, the cells were incubated with mouse anti-Hsp60 monoclonal antibody (1:200 dilution) for 1 h at room temperature, followed by incubation with Alexa Fluor 488 goat anti-mouse IgG secondary antibody (1:200 dilution) or Texas Red-conjugated goat anti-mouse IgG secondary antibody (1:100 dilution) for 45 min at room temperature. Fluorescence images were obtained using a Leica TCS SL confocal laser scanning microscope (Leica Microsystems, Wetzler, Germany). Pictures were taken with Leica Confocal Software (Leica Microsystems).

In vitro detection of apoptosis. The TUNEL assay was performed using the In Situ Cell Death Detection Kit (Roche Diagnostics) as described previously [18]. Briefly, Huh7 cells seeded in a slide flask were transfected with the expression vector. After 24 h, the cells were fixed, permeabilized, and treated with the TUNEL reaction mixture in a humidified chamber for 60 min at 37 °C in the dark. To visualize the HBx or Hsp60 protein, the cells were incubated with mouse anti-HA

monoclonal antibody (1:100 dilution) and Texas red-conjugated goat anti-mouse antibody (1:100 dilution). Cells were analyzed using fluorescence microscopy (Leica Microsystems). The TUNEL-positive cells among the HBx- or Hsp60-positive cells in five independent fields were counted. Independent experiments were performed at least three times.

Statistics. The results of the TUNEL assay were analyzed using analysis of variance (ANOVA) with a post hoc Scheffe test (StatView, Abacus Concepts, Berkeley, CA). The data represent means ± SD calculated from three independent experiments. P values <0.05 were considered significant.

Results

Identification of host proteins interacting with HBx by mass spectrometry

To search for novel host proteins that interact with HBx, we performed affinity purification. Lysates of human hepatocytes were applied to Sepharose beads that had been conjugated with bacterially expressed recombinant GST-HBx fusion protein or control GST protein. After the beads had been washed, proteins were eluted, separated by SDS–PAGE, and silver stained. As shown in Fig. 1, 56-kDa (Figs. 1A and B) and 46-kDa (Figs. 1A and C) proteins appeared to specifically bind to recombinant HBx protein when hepatocyte lysates were incubated with GST-HBx beads (Figs. 1A–C, lane 2). These proteins were not seen on GST-HBx beads without hepatocyte lysates (Figs. 1A–C, lane 4) or control lanes of GST beads (Figs. 1A–C, lanes 1 and 3). These 56- and 46-kDa bands in lane 2 and, as a control, the corresponding bands in lane 4 were excised from the gel and subjected to in-gel tryptic digestion. The resultant peptides were extracted from the gel slice and analyzed by nanoLC–nanoESI–MS/MS to obtain the m/z values of the respective peptides and their collision-induced dissociation (CID) fragments. Amino acid sequence databases of human proteins were searched for the m/z values using the MASCOT search program, and the authentic protein was identified from the search results. Both 56- and 46-kDa proteins were identified and Table 1 summarizes the identified peptide sequences of a 56-kDa band. The most significant hit for the 56-kDa protein was obtained with the human chaperonin GroEL precursor (Hsp60, PIR Accession No. A32800). Five of the peptides obtained were derived from human Hsp60, not bacterial GroEL, and all of these peptides had unique human Hsp60 sequences. No protein was identified from the control band in lane 4 (data not shown). The interaction between 46-kDa protein (Fig. 1C) and HBx is under investigation.

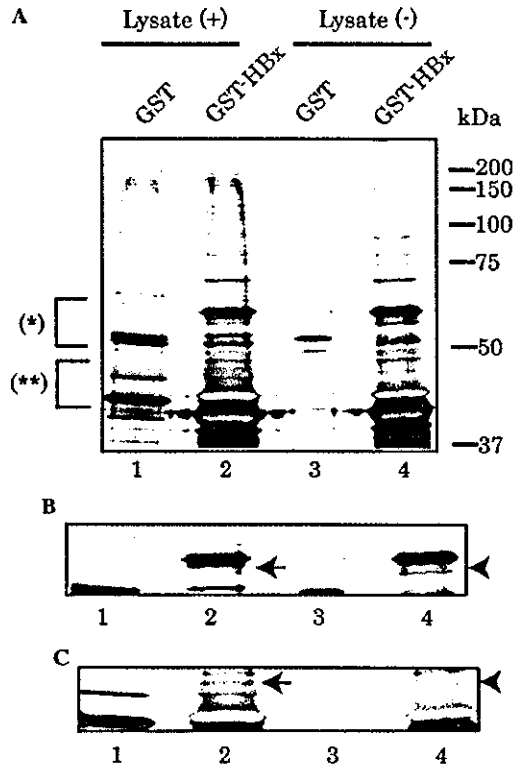


Fig. 1. Affinity-purified proteins that bind to recombinant HBx. (A) Lysates of human hepatocytes were applied to the Sepharose beads conjugated with bacterially expressed recombinant GST-HBx fusion proteins or control GST proteins. After the beads were washed, proteins were eluted from the beads, separated by SDS-PAGE, and silver stained. Lane 1, GST incubated with hepatocyte lysate; lane 2, GST-HBx incubated with hepatocyte lysate; lane 3, GST without hepatocyte lysate; and lane 4, GST-HBx without hepatocyte lysate. The molecular weights of the protein standards are indicated. (B) Higher magnification of gel plate (*). (C) Higher magnification of gel plate (**). The 56- and 46-kDa bands indicated by the arrow in lane 2 and, as a control, the corresponding bands in lane 4 (indicated by the arrow head) were excised from the gel and subjected to in-gel tryptic digestion.

HBx interacts with Hsp60 in vivo

The protein identified by mass spectrometry, Hsp60, has been reported to associate with viral proteins, including HBV polymerase [23,24] and human immunodeficiency virus (HIV) integrase [25]. These interactions were shown to be important to the functions of the viral proteins. Therefore, we further analyzed the interaction between HBx and Hsp60. To investigate whether HBx binds to Hsp60 in a cellular context, expression vectors for GFP-tagged HBx (GFP-HBx) and for HA-tagged Hsp60 (Hsp60-HA) were introduced into 293 cells. As shown in Figs. 2A and B, immunoprecipitation followed by immunoblotting demonstrated that GFP-HBx, and not GFP alone, interacted with Hsp60-HA in 293 cells. The interaction between HBx and Hsp60 was confirmed using flag-tagged HBx (HBx-flag) and Hsp60-HA constructs (data not shown). To analyze the interaction between HBx and endogenous Hsp60, pEGFP-X or pEGFP-C1 was transfected into 293 cells, and immunoprecipitation was performed. This experiment

Table 1
Mass spectrometric identification of HBx-associated protein

Protein name	PIR Accession No.	Molecular mass (kDa)		Peptides matched	Peptide assignment				
		Observed	Calculated		Observed mass (charge)	Expected mass (Da)	Calculated mass (Da)	Sequence	Peptide position
Human chaperonin GroEL precursor (Hsp60)	A32800	56	57.963	5	695.32 (+2)	1388.62	1388.70	(R)GYISPYFINTSK	222–233
					752.75 (+2)	1503.49	1503.75	(K)TLNDELEIIEGMK	206–218
					801.47 (+2)	1600.92	1600.74	(K)C*EFQDAYVLLSEK	236–249
					843.10 (+2)	1684.18	1683.90	(R)AAVEEGIVLGGGC*ALLR	430–446
					886.63 (+2)	1771.24	1770.85	(R)C*IPALDSLTPANEDQK	447–462

PIR accession number is derived from the Protein Information Resources (PIR) database (http://pir.georgetown.edu/).
'C*' in the peptide sequences stands for carboxyamidomethyl cysteine.

revealed that GFP-HBx bound to endogenous Hsp60 in vivo (Fig. 2C).

Mapping the interaction site for Hsp60 on HBx

To further characterize the HBx binding site involved in the interaction with Hsp60, we performed immunoprecipitation using the following deletion mutants of HBx: pEGFP-XΔ(5–87), pEGFP-X(1–117), pEGFP-X(1–87), and pEGFP-X(68–117) (Fig. 3A). All of the constructs except X(1–87) bound endogenous Hsp60 (Fig. 3B), indicating that amino acids 88–117 of HBx were necessary for binding to Hsp60. This region (amino acids 88–117) contains the component for the transactivation function of HBx [15] and the putative leucine-rich nuclear export signal (NES) motif (amino acids 89–100) [26].

HBx and Hsp60 colocalized in mitochondria

To assess the subcellular localization of HBx and Hsp60, immunocytochemistry using confocal microscopy was performed. The mitochondrion is the major cellular location of Hsp60. The localization of Hsp60 (green) in mitochondria (red) was consistent with previous reports (Fig. 4A) [14]. To investigate the in vivo localization of HBx, GFP-HBx was transfected into Huh7 cells, and mitochondria or endogenous Hsp60 was stained. HBx (green) colocalized in mitochondria (red), as we had previously reported (Fig. 4B). Furthermore, GFP-HBx (green) colocalized with Hsp60 (red), which localized to mitochondria (Fig. 4C). Although HBx 1–87 did not bind to Hsp60 (Fig. 3B), the majority of HBx 1–87 still localized to mitochondria [16], suggesting that association of HBx and Hsp60 is not required for the mitochondrial localization of HBx.

Hsp60 facilitates HBx-induced apoptosis

We and other groups previously reported that HBx induces apoptosis or sensitizes cells to apoptotic stimuli such as tumor necrosis factor α (TNFα) [16,27,28]. Therefore, we investigated the effect of Hsp60 on HBx-mediated apoptosis. Huh7 cells were transfected with the HBx or Hsp60 expression vector, and the TUNEL assay was performed. Fig. 5A shows a representative image of the nucleus of an apoptotic cell (middle, green) induced by HBx transduction (left, red). The level of HBx expressed using the cytomegalovirus promoter-enhancer is considered to be within the range in HBV infected cells [28,29]. Approximately 40% of the cells underwent apoptosis after transduction with 0.26 μg HBx expression plasmid. The HBx 1–87 mutant, which lacks the Hsp60 binding region, caused a similar degree (33%) of cell death. In contrast, Hsp60 did not produce

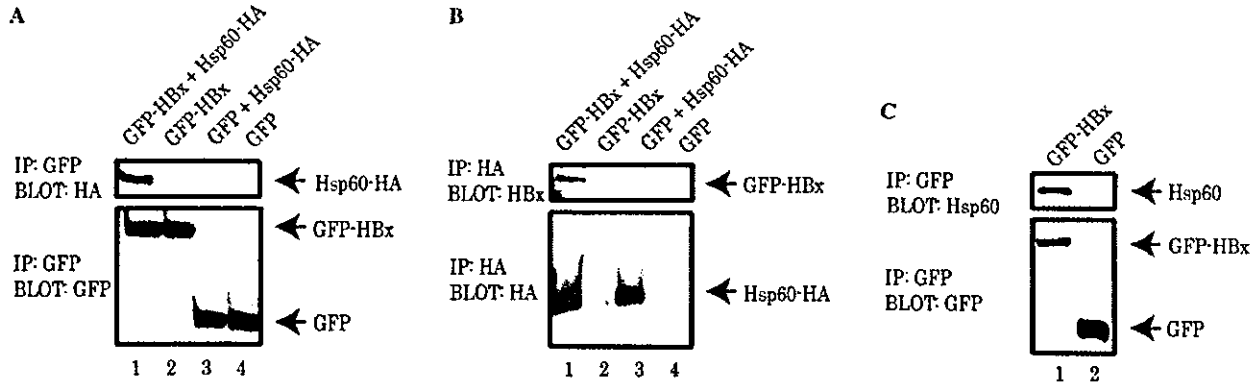


Fig. 2. Interaction of HBx and Hsp60 in vivo. 293 cells were transfected with the following combinations of plasmids: lane 1, pEGFP-X 1 μg and pcDNA3-Hsp60-HA 1 μg; lane 2, pEGFP-X 1 μg and pcDNA3 1 μg; lane 3, pEGFP-C1 1 μg and pcDNA3-Hsp60-HA 1 μg; and lane 4, pEGFP-C1 1 μg and pcDNA3 1 μg. After 48 h, cells were harvested, and the lysates were immunoprecipitated with anti-GFP (A) or anti-HA (B) antibody followed by immunoblot with corresponding antibodies. (C) The following plasmids were transfected into 293 cells: lane 1, pEGFP-X 2 μg; and lane 2, pEGFP-C1 2 μg. Cell lysates were immunoprecipitated with anti-GFP antibody followed by immunoblot with corresponding antibodies.

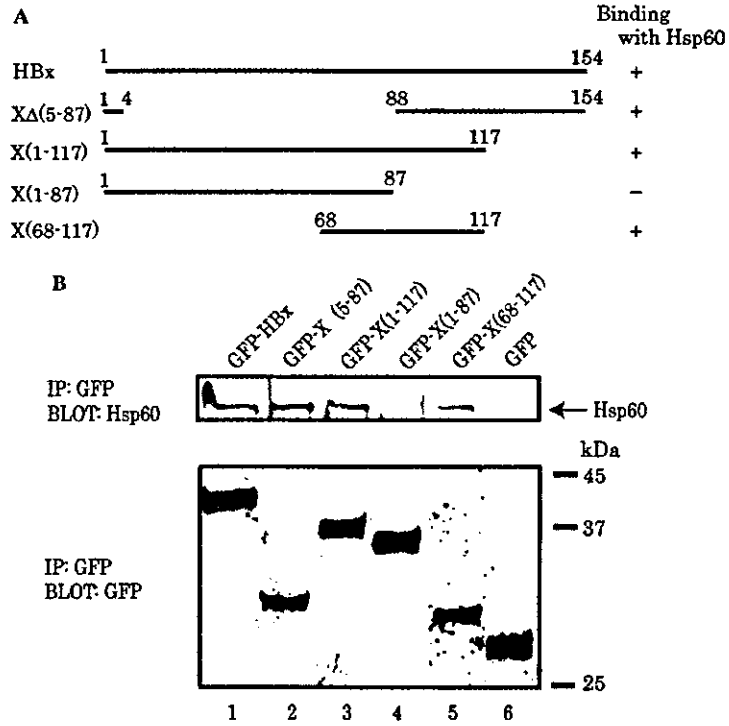


Fig. 3. Mapping of the Hsp60-interacting region of HBx. (A) Schematic presentation of the HBx and HBx deletion mutants that were analyzed. The binding of HBx with Hsp60 is shown. (B) 293 cells were transfected with the following plasmids: lane 1, pEGFP-X 2 μg; lane 2, pEGFP-XΔ(5–87) 2 μg; lane 3, pEGFP-X(1–117) 2 μg; lane 4, pEGFP-X(1–87) 2 μg; lane 5, pEGFP-X(68–117) 2 μg; and lane 6, pEGFP-C1 2 μg. Cell lysates were immunoprecipitated with anti-GFP antibody followed by immunoblot with corresponding antibodies. The molecular weights of the protein standards are indicated.

significant cell death (11%) (Fig. 5B). As a control, the β-galactosidase expression plasmid, pCMVβ, was transfected into cells, and TUNEL positive cells were counted [16]. The percentage of apoptotic cells transfected with pCMV β was similar to that of Hsp60 (data not shown).

To further investigate the effect of Hsp60 on HBx-mediated apoptosis, Huh7 cells were co-transfected with 0.06 μg of the HBx expression plasmid with and without 0.20 μg of the Hsp60 expression plasmid (about three-fold the amount of HBx). Under the conditions of this assay, 18% of cells expressing HBx underwent apoptosis,

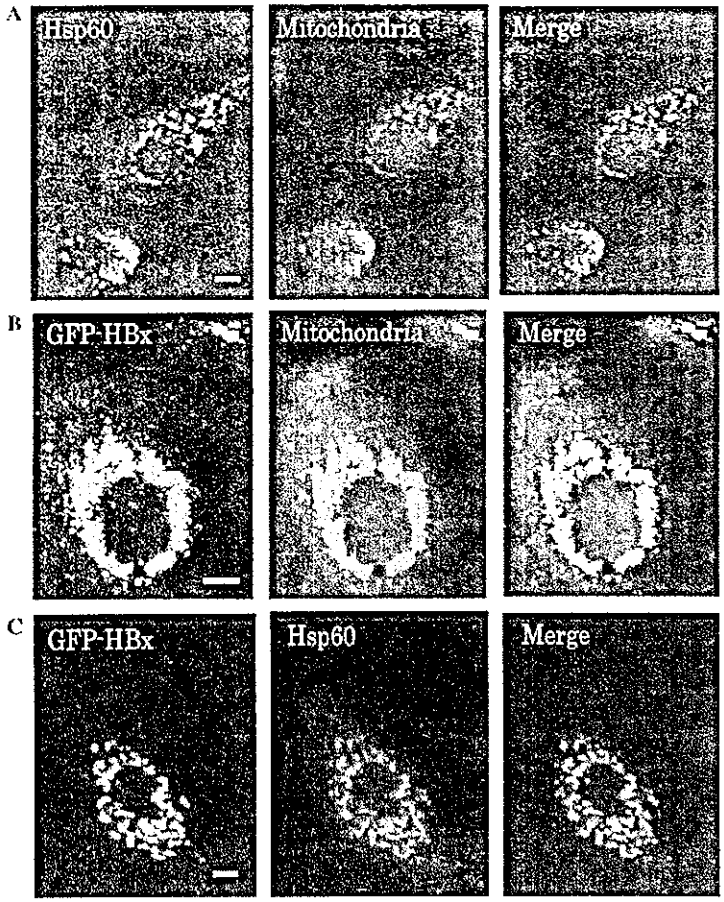


Fig. 4. Subcellular localization of HBx and Hsp60 in Huh7 cells. (A) Cells were immunostained with anti-Hsp60 antibody and secondary antibody Alexa Fluor 488 (green), and mitochondria were stained with CMXRos (red). (B) Localization of transfected GFP-HBx (green) and mitochondria (red). pEGFP-X was transfected, and after 48 h, mitochondria were stained with CMXRos. (C) Localization of GFP-HBx (green) and endogenous Hsp60 (red). Cells were transfected with pEGFP-X, and after 48 h, endogenous Hsp60 was immunostained with anti-Hsp60 antibody and Texas red-conjugated goat anti-mouse IgG secondary antibody. Merged images are shown in the right panel. Bar = 10 μm.

and the co-expression of Hsp60 with HBx enhanced the effect of HBx-mediated apoptosis more than twofold (42%). This effect of Hsp60 was not seen in cells expressing the HBx mutant that lacks the Hsp60 binding domain (HBx 1–87) (Fig. 5C). Introduction of Hsp60 alone did not induce significant levels of apoptosis. In spite of the presence of endogenous Hsp60 in Huh7 cells, there is no evident difference in percentages of apoptotic cells induced by HBx or HBx 1–87. One possibility is that overexpressed HBx needs much amount of Hsp60 to be refolded correctly. These data suggest that Hsp60 binds to HBx and enhances HBx-mediated apoptosis in mitochondria.

Discussion

In the present study, we identified Hsp60 as one of the proteins that interact with HBx by searching a protein database for the tandem mass spectra obtained from the nanoLC–nanoESI–MS/MS analysis of tryptic peptides from a 56-kDa protein band. We also demonstrated that amino acids 88–117 of HBx play an important role in the association with Hsp60. Our results indicate that HBx and Hsp60 interact in mitochondria, and that this interaction enhances the promotion of cell death by HBx.

Tandem MS instruments, such as the ion-trap used in this study, are routinely applied in an LC–MS/MS equipped with a nanoelectrospray ionization (ESI) source in order to generate peptide fragment ion spectra suitable for protein identification from sequence databases. Instrumental control for further fragmentation of automatically selected peptide ions, called data-dependent collision-induced dissociation (CID), is a notable trend with recent MS/MS instruments. The MASCOT search program is based on the MOWSE algorithm, but has a greater selectivity and sensitivity than the earlier algorithms, which simply counted the number of peptide matches. MOWSE takes into account the relative abundances of the peptides of a given mass in the

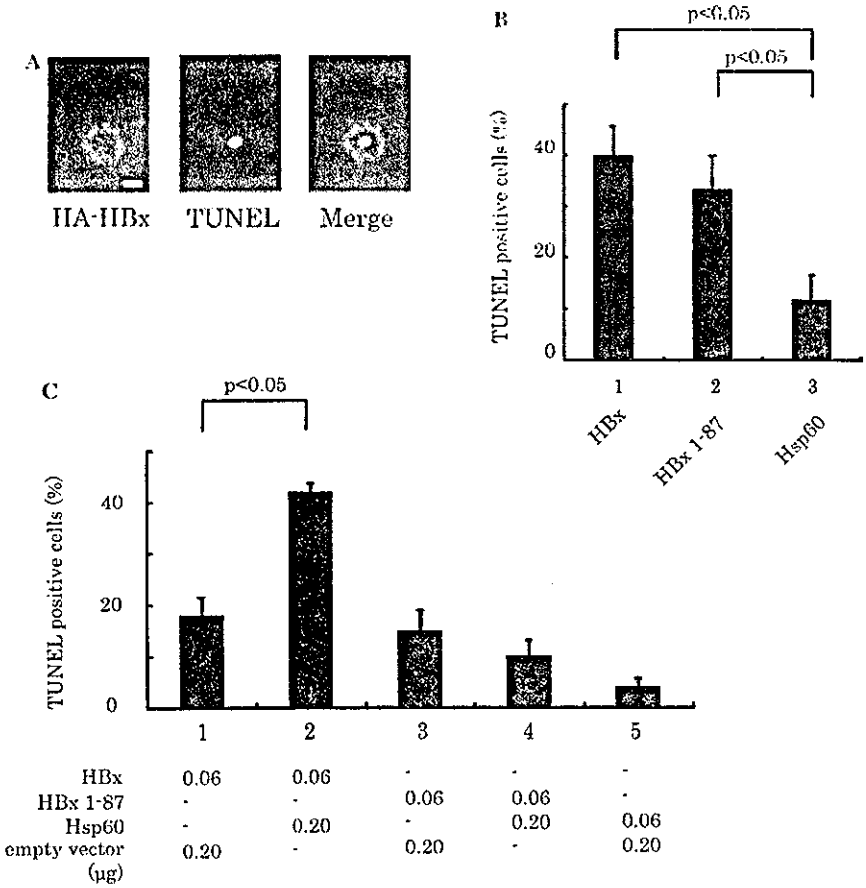


Fig. 5. Hsp60 facilitates HBx-induced apoptosis. (A) A representative image of a TUNEL positive cell after transduction with HBx. HBx and TUNEL positive nuclei were visualized by red (left) and green (middle), respectively. Merged image is shown in the right panel. Bar = 10 μm. (B) The effects of HBx and Hsp60 on apoptosis. Huh7 cells were transfected with: lane 1, pCMV-HA-X 0.26 μg; lane 2, pCMV-HA-X(1–87) 0.26 μg; and lane 3, pcDNA3-Hsp60-HA 0.26 μg. After 24 h, the TUNEL assay was performed and apoptotic cells were counted. The results are means ± SD from three separate experiments. (C) The following plasmids were introduced into Huh7 cells: lane 1, pCMV-HA-X 0.06 μg and pCMV 0.20 μg; lane 2, pCMV-HA-X 0.06 μg and pCMV-Hsp60 0.20 μg; lane 3, pCMV-HA-X(1–87) 0.06 μg and pCMV 0.20 μg; lane 4, pCMV-HA-X(1–87) 0.06 μg and pCMV-Hsp60 0.20 μg; and lane 5, pcDNA3-Hsp60-HA 0.06 μg and pCMV 0.20 μg. The number of apoptotic cells was analyzed as described above. Results are means ± SD from three separate experiments.

database. MASCOT sets the benchmark for protein identification using CID spectra because it is highly reliable, automatically analyzes data from the entire LC–MS/MS experiment, and requires no user interpretation. As shown in Table 1, Hsp60 was the only protein identified by MASCOT as having a significant hit score, which was the total of hit scores from five Hsp60 peptides, in a search using CID spectra generated by the analysis of peptide fragments. Although many *E. coli* proteins remained on the surface of the bacterially expressed recombinant GST-HBx, we were able to identify two human proteins from a small amount of material using this system. This method is widely applicable to finding novel protein–protein interactions.

Hsp60 is homologous to the eubacteria chaperonin GroEL. GroEL consists of 14 identical subunits of 58 kDa, each arranged in two stacked heptameric rings, each of which contains a central cavity [30]. Each subunit consists of three domains: the equatorial, the intermediate, and the apical. The equatorial domain has the ATP binding site, and the hydrophobic surfaces of the apical domain appear to interact with substrates. The intermediate domain connects the equatorial and apical domains. Owing to difficulty in making Hsp60 deletion constructs that retain the ability to correctly form a complex, we were unable to map the domain of Hsp60 that interacts with HBx. However, based upon previous reports, HBx probably binds to the apical domain of Hsp60 in a manner similar to that of other Hsp60 substrates. Although the HBx expression level by using the cytomegalovirus promoter–enhancer is reported to be within the limits of HBV infected cells [28,29], the interaction with Hsp60 in physiological level of HBx also remains to be investigated.

GroEL substrates generally consist of two or more domains with αβ-folds, which contain α-helices and buried β-sheets with extensive hydrophobic surfaces [31]. HBx has two putative alpha helical regions (amino acids 75–88 and 109–131) that seem to aid in the mitochondrial targeting of the protein [32]. The Hsp60 binding region of HBx (amino acids 88–117) contains part of one alpha helical region. The lack of this alpha helical region in the HBx construct that contains amino acids 1–87 might contribute to its inability to bind Hsp60.

Several investigators have reported the subcellular distribution of HBx as predominantly cytosolic with a small amount localized to the nuclear fraction [33]. We and Siddiqui's group reported that the mitochondrion was the target of HBx [16,32]. We have recently demonstrated that amino acids 68–117 of HBx are important in localizing the protein to mitochondria. As previously reported, Hsp60 is synthesized as a precursor protein with an N-terminal mitochondrial sequence, and cleavage of the targeting sequence requires the import of the precursor protein into the matrix compartment [14]. Thus, the association of HBx and Hsp60 in mitochondria appears to be an intrinsic property of this viral protein rather than an artifact of overexpression.

The mitochondrion is a key organelle in the control of apoptosis. We have reported that the ectopic expression of HBx causes mitochondrial aggregation, a decrease in mitochondrial membrane potential, and membrane blebbing in HBx transfected cells [18]. Furthermore, we have recently demonstrated that HBx constructs that localize to mitochondria can induce apoptosis and that HBx 1–87, which localizes to mitochondria, induces apoptosis at a level similar to that produced by HBx [16]. Rahmani et al. reported that HBx directly interacts with an outer mitochondrial human voltage-dependent anion channel, HVDAC3 [11]. This association leads to a decrease in the mitochondrial membrane potential and constitutively activated transcription factors STAT-3 and NF-κB via oxidative stress [34]. Recently, two novel mechanisms for the pro-apoptotic function of HBx have been reported. The first is that HBx interacts with c-FLIP and inhibits its anti-apoptotic function [12]. The second is that HBx alters intracellular calcium signaling, which is an important event in HBx-induced apoptosis [35]. Inappropriate apoptosis has been considered to be associated with human diseases, including inflammation and cancer. Moreover, the expression of an oncogenic protein can induce or increase cell sensitivity to apoptosis. During chronic HBV infection, HBx expression may induce apoptosis in infected hepatocytes and cause inflammation. In this study, the expression of HBx causes apoptosis, and HBx-induced apoptosis was enhanced by the binding of Hsp60 (Fig. 5C). Our results suggest that both the localization of HBx in mitochondria and the interaction of HBx with Hsp60 are important for HBx-mediated apoptosis.

On the other hand, anti-apoptotic functions of HBx have also been reported. HBx appears to inhibit p53-mediated apoptosis by sequestering p53 in the cytosol [10,36]. HBx has also been reported to inhibit caspase 3 activity and to thereby block apoptosis [37]. These data suggest that HBx alters the balance between programmed cell death and proliferation, and participates in the apoptotic and anti-apoptotic pathways.

Hsp60 also interacts with HBV polymerase and HIV integrase in vivo. It has been reported that Hsp60 interacts with HBV polymerase via two minimal sites and converts it into the active state [24]. Hsp60 has been demonstrated to bind to the catalytic domain of HIV integrase and stimulate its in vitro processing and joining activities [25]. Although HBx does not appear to have enzymatic activity like that of HBV polymerase or HIV integrase, Hsp60 might promote the correct folding of HBx and thereby enhance HBx-mediated apoptosis. Based on these previous reports, the interaction between HBV and Hsp60 may be vital to the virological cycle of HBV.

In conclusion, we have shown that chaperonin Hsp60 interacts with HBx and enhances HBx-mediated apoptosis. The precise mechanism of apoptosis induced by HBx and Hsp60 and the biological significance of this interaction during HBV infection remain to be investigated.

Acknowledgments

We thank Mitsuko Tsubouchi for technical assistance, Dr. Hideki Taguchi (Chemical Resources Laboratory, Tokyo Institute of Technology, Yokohama, Japan) for helpful discussions, and Dr. Naohiko Seki (Chiba University) for the gift of pCMV-Hsp60. This work was supported in part by the Health Science Research Grants for Medical Frontier Strategy Research from the Ministry of Health, Labor, and Welfare of Japan, and Grants-in-Aid for scientific research from the Ministry of Education, Culture, Sports, Science and Technology, Japan.

References

[1] D. Ganem, H.E. Varmus, The molecular biology of the hepatitis B viruses, Annu. Rev. Biochem. 56 (1987) 651–693.
[2] W.M. Lee, Hepatitis B virus infection, N. Engl. J. Med. 337 (1997) 1733–1745.
[3] M. Omata, Treatment of chronic hepatitis B infection, N. Engl. J. Med. 339 (1998) 114–115.
[4] F. Zoulim, J. Saputelli, C. Seeger, Woodchuck hepatitis virus X protein is required for viral infection in vivo, J. Virol. 68 (1994) 2026–2030.
[5] M.J. Bouchard, L.H. Wang, R.J. Schneider, Calcium signaling by HBx protein in hepatitis B virus DNA replication, Science 294 (2001) 2376–2378.

[6] C.M. Kim, K. Koike, I. Saito, T. Miyamura, G. Jay, HBx gene of hepatitis B virus induces liver cancer in transgenic mice, Nature 351 (1991) 317–320.
[7] S. Murakami, Hepatitis B virus X protein: a multifunctional viral regulator, J. Gastroenterol. 36 (2001) 651–660.
[8] J.H. Cheong, M. Yi, Y. Lin, S. Murakami, Human RPB5, a subunit shared by eukaryotic nuclear RNA polymerases, binds human hepatitis B virus X protein and may play a role in X transactivation, EMBO J. 14 (1995) 143–150.
[9] D.K. Lee, S.H. Park, Y. Yi, S.G. Choi, C. Lee, W.T. Parks, H. Cho, M.P. de Caestecker, Y. Shaul, A.B. Roberts, S.J. Kim, The hepatitis B virus encoded oncoprotein pX amplifies TGF-beta family signaling through direct interaction with Smad4: potential mechanism of hepatitis B virus-induced liver fibrosis, Genes Dev. 15 (2001) 455–466.
[10] H. Ueda, S.J. Ullrich, J.D. Gangemi, C.A. Kappel, L. Ngo, M.A. Feitelson, G. Jay, Functional inactivation but not structural mutation of p53 causes liver cancer, Nat. Genet. 9 (1995) 41–47.
[11] Z. Rahmani, K.W. Huh, R. Lasher, A. Siddiqui, Hepatitis B virus X protein colocalizes to mitochondria with a human voltage-dependent anion channel, HVDAC3, and alters its transmembrane potential, J. Virol. 74 (2000) 2840–2846.
[12] K.H. Kim, B.L. Seong, Pro-apoptotic function of HBV X protein is mediated by interaction with c-FLIP and enhancement of death-inducing signal, EMBO J. 22 (2003) 2104–2116.
[13] H. Marusawa, S. Matsuzawa, K. Welsh, H. Zou, R. Armstrong, I. Tamm, J.C. Reed, HBXIP functions as a cofactor of survivin in apoptosis suppression, EMBO J. 22 (2003) 2729–2740.
[14] B. Bukau, A.L. Horwich, The Hsp70 and Hsp60 chaperone machines, Cell 92 (1998) 351–366.
[15] S. Takada, K. Koike, Three sites of the hepatitis B virus X protein cooperatively interact with cellular proteins, Virology 205 (1994) 503–510.
[16] Y. Shirakata, K. Koike, Hepatitis B virus X protein induces cell death by causing loss of mitochondrial membrane potential, J. Biol. Chem. 278 (2003) 22071–22078.
[17] M. Arii, S. Takada, K. Koike, Identification of three essential regions of hepatitis B virus X protein for trans-activation function, Oncogene 7 (1992) 397–403.
[18] S. Takada, Y. Shirakata, N. Kaneniwa, K. Koike, Association of hepatitis B virus X protein with mitochondria causes mitochondrial aggregation at the nuclear periphery, leading to cell death, Oncogene 18 (1999) 6965–6973.
[19] F. Kanai, P.A. Marignani, D. Sarbassova, R. Yagi, R.A. Hall, M. Donowitz, A. Hisaminato, T. Fujiwara, Y. Ito, L.C. Cantley, M.B. Yaffe, TAZ: a novel transcriptional co-activator regulated by interactions with 14-3-3 and PDZ domain proteins, EMBO J. 19 (2000) 6778–6791.
[20] H. Blum, H. Beier, H.J. Gross, Improved silver staining of plant proteins, RNA and DNA in polyacrylamide gels, Electrophoresis 8 (1987) 93–99.
[21] A. Shevchenko, M. Wilm, O. Vorm, M. Mann, Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels, Anal. Chem. 68 (1996) 850–858.
[22] T. Kawakami, T. Nagata, A. Muraguchi, T. Nishimura, Alteration of protein composition in mouse thymocytes by signals through T-cell receptor, Electrophoresis 21 (2000) 1846–1852.
[23] S.G. Park, G. Jung, Human hepatitis B virus polymerase interacts with the molecular chaperonin Hsp60, J. Virol. 75 (2001) 6962–6968.
[24] S.G. Park, S.O. Lim, G. Jung, Binding site analysis of human HBV pol for molecular chaperonin, hsp60, Virology 298 (2002) 116–123.
[25] V. Parissi, C. Calmels, V.R. De Soultrait, A. Caumont, M. Fournier, S. Chaignepain, S. Litvak, Functional interactions of human immunodeficiency virus type 1 integrase with human and yeast HSP60, J. Virol. 75 (2001) 11344–11353.
[26] M. Forgues, A.J. Marrogi, E.A. Spillare, C.G. Wu, Q. Yang, M. Yoshida, X.W. Wang, Interaction of the hepatitis B virus X protein with the Crm1-dependent nuclear export pathway, J. Biol. Chem. 276 (2001) 22797–22803.
[27] O. Terradillos, A. de La Coste, T. Pollicino, C. Neuveut, D. Sitterlin, H. Lecoeur, M.L. Gougeon, A. Kahn, M.A. Buendia, The hepatitis B virus X protein abrogates Bcl-2-mediated protection against Fas apoptosis in the liver, Oncogene 21 (2002) 377–386.
[28] F. Su, R.J. Schneider, Hepatitis B virus HBx protein sensitizes cells to apoptotic killing by tumor necrosis factor alpha, Proc. Natl. Acad. Sci. USA 94 (1997) 8744–8749.
[29] M. Dandri, P. Schirmacher, C.E. Rogler, Woodchuck hepatitis virus X protein is present in chronically infected woodchuck liver and woodchuck hepatocellular carcinomas which are permissive for viral replication, J. Virol. 70 (1996) 5246–5254.
[30] K. Braig, Z. Otwinowski, R. Hegde, D.C. Boisvert, A. Joachimiak, A.L. Horwich, P.B. Sigler, The crystal structure of the bacterial chaperonin GroEL at 2.8 Å, Nature 371 (1994) 578–586.
[31] W.A. Houry, D. Frishman, C. Eckerskorn, F. Lottspeich, F.U. Hartl, Identification of in vivo substrates of the chaperonin GroEL, Nature 402 (1999) 147–154.
[32] K. Huh, A. Siddiqui, Characterization of the mitochondrial association of hepatitis B virus X protein, HBx, Mitochondrion 1 (2002) 349–359.
[33] F. Henkler, J. Hoare, N. Waseem, R.D. Goldin, M.J. McGarvey, R. Koshy, I.A. King, Intracellular localization of the hepatitis B virus HBx protein, J. Gen. Virol. 82 (2001) 871–882.
[34] G. Waris, K.W. Huh, A. Siddiqui, Mitochondrially associated hepatitis B virus X protein constitutively activates transcription factors STAT-3 and NF-kappa B via oxidative stress, Mol. Cell. Biol. 21 (2001) 7721–7730.
[35] M. Chami, D. Ferrari, P. Nicotera, P. Paterlini-Brechot, R. Rizzuto, Caspase-dependent alterations of Ca^{2+} signaling in the induction of apoptosis by hepatitis B virus X protein, J. Biol. Chem. 278 (2003) 31745–31755.
[36] X.W. Wang, M.K. Gibson, W. Vermeulen, H. Yeh, K. Forrester, H.W. Sturzbecher, J.H. Hoeijmakers, C.C. Harris, Abrogation of p53-induced apoptosis by the hepatitis B virus X gene, Cancer Res. 55 (1995) 6012–6016.
[37] K. Gottlob, M. Fulco, M. Levrero, A. Graessmann, The hepatitis B virus HBx protein inhibits caspase 3 activity, J. Biol. Chem. 273 (1998) 33347–33353.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Biochemical and Biophysical Research Communications 315 (2004) 1088–1096
BBRC
www.elsevier.com/locate/ybbrc

A low-density cDNA microarray with a unique reference RNA: pattern recognition analysis for IFN efficacy prediction to HCV as a model☆

Akito Daiba,[a] Niro Inaba,[a] Satoshi Ando,[a] Naoki Kajiyama,[a] Hiroshi Yatsuhashi,[b] Hiroshi Terasaki,[a] Atsushi Ito,[a] Masanori Ogasawara,[a] Aki Abe,[a] Junichi Yoshioka,[a] Kazuhiro Hayashida,[c,1] Shuichi Kaneko,[d] Michinori Kohara,[e] and Satoru Ito[a,*,2]

[a] *JGS Japan Genome Solutions, Inc. 51 Komiya-cho, Hachioji, Tokyo 192-0031, Japan*
[b] *National Nagasaki Medical Center, Kubara 2-1001-1 Omura City, Nagasaki 856-8562, Japan*
[c] *Medicine and Biosystemic Science, Kyushu University Graduate School of Medical Sciences, 3-1-1 Maidashi, Higashi-ku, Fukuoka 812-8582, Japan*
[d] *Kanazawa University Graduate School of Medical Sciences, 13-1 Muro-machi, Kanazawa 920-8641, Japan*
[e] *Tokyo Metropolitan Institute of Medical Science, 3-18-22 Hon-komagome, Bunkyo-ku, Tokyo 113-8613, Japan*

Received 9 January 2004

Abstract

We have designed and established a low-density (295 genes) cDNA microarray for the prediction of IFN efficacy in hepatitis C patients. To obtain a precise and consistent microarray data, we collected a data set from three spots for each gene (mRNA) and using three different scanning conditions. We also established an artificial reference RNA representing pseudo-inflammatory conditions from established hepatocyte cell lines supplemented with synthetic RNAs to 48 inflammatory genes. We also developed a novel algorithm that replaces the standard hierarchical-clustering method and allows handling of the large data set with ease. This algorithm utilizes a standard space database (SSDB) as a key scale to calculate the Mahalanobis distance (MD) from the center of gravity in the SSDB. We further utilized sMD (divided by parameter k: MD/k) to reduce MD number as a predictive value. The efficacy prediction of conventional IFN mono-therapy was 100% for non-responder (NR) vs. transient responder (TR)/sustained responder (SR) ($P < 0.0005$). Finally, we show that this method is acceptable for clinical application.

Keywords: Low-density microarray; Artificial reference RNA; Efficacy prediction; Mahalanobis distance

Large data sets can be collected from cDNA microarrays for the study of expression profiles in biological systems. The large amount of data generated can be especially useful to help cluster genes into interest groups. Such genome-wide information can be used for clinical applications (see reviews, [1–5]), for example, for the identification of the cDNA expression patterns associated with different stages of tumor development. However, translating complex microarray data into practical clinical applications has been difficult. New algorithms are being developed to solve this problem, for example for the prognosis of cancer treatment [6,7]. Also, low-density microarrays with selected genes of interest can simplify the analysis of microarray data.

Another critical issue in understanding microarray data is the level of precision in the data set. Solid phase DNA hybridization is not as quantitative as hybridization in solution, and scanners have limited dynamic

☆ *Abbreviations:* SAGE, serial analysis of gene expression; SSDB, standard space database; MD, Mahalanobis distance; sMD, scaled MD; FL, firefly luciferase gene; RL, *Renilla* luciferase gene; GP, baculovirus glycoprotein gene; LD, lambda DNA; MEP, microcapillary electrophoretic; aRNA, amplified RNA; NR, non-responder; SR, sustained responder; TR, transient responder.

* Corresponding author. Fax: +81-426-45-0461.
E-mail address: sr-itou@jgs-inc.co.jp (S. Ito).
[1] Present address: Sasebo Kyosai Hospital, 10-17 Shimaji, Sasebo, Nagasaki 857-8575, Japan.
[2] Visiting position: Medical Research Institute, Tokyo Medical and Dental University.

0006-291X/$ - see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.bbrc.2004.01.160

ranges. In addition, sample variability can result from variations in sampling conditions, RNA amplification, RNA degradation, and cDNA labeling conditions. These factors are not well understood, and, currently, precision is primarily controlled at the level of data collection [8–10].

One means of enhancing the precision of data collection is to use reference controls for each individual. This can be accomplished by laser-captured microdissection of tissues into diseased and counter part areas [11,12]. In case of hepatitis C, this is difficult because inflammatory damage occurs throughout the liver. Another possible approach is to utilize artificial reference RNAs as a reference [13–15] in conjunction with RNA from established cell lines, such as hepatocellular carcinoma cell lines. However, the stages and characteristics of the disease in vivo and in vitro can differ, and expression of some RNAs of interest, especially some inflammatory genes, may be too low in the cultured cells to produce a satisfactory signal-to-noise ratio.

In the current studies, we identified inflammatory genes that can be used for a low-density microarray to predict the efficacy of INF treatment in hepatitis C patients. We found sufficient levels of expression for these genes in patient samples, but very low levels of expression in established cell lines. We overcame this problem by using a low-density cDNA microarray system in conjunction with a unique artificial reference RNA. In addition, we describe an algorithm for analysis of the microarray data.

Methods

cDNA Microarray. We selected 295 genes for the cDNA microarray based on publicly available data, including SAGE analysis and other DNA microarray data from HCV patients and a normal subject [16,17]. From 2000 candidate genes, we omitted low frequency-tag genes based on the SAGE data. Genes previously identified as predictive host factors for IFN efficacy [18–20] were given a high priority. For most of the genes, each cDNA was designed approximately 500–600 bp and within approximately 1 kb region from the 3′-poly(A) tail and all cDNAs for microarray probe were cloned into the pGEM vector (Promega, Madison, WI). We also selected and cloned external control genes (approximately 0.5–1 kb) into the pGEM vector to establish the dynamic range of the microarray. These genes were firefly luciferase (FL; negative control), *Renilla* luciferase (RL) [21], baculovirus envelope gp64 (GP), and lambda phage DNA (LD). All clones for capture probe on microarray were sequenced and validated by comparison with the GenBank sequence. The aminosilane surfaces of SuperAmine glass slides (TeleChem International, Sunnyvale, CA) were stamped with triplicate spots of cDNA probe corresponding to each of the remaining 295 genes. The average spot size was 150 μm and separated each other with a distance (500 μm) as shown in Fig. 2B.

Reference RNA preparation. Extracted total RNAs from four hepatocellular carcinoma cell lines (HepG2, Hep3B, Huh7, and IMY-4 [22]) purified through RNeasy column (Qiagen, Hilden, Germany) were mixed as a reference source. In order to find a mixing ratio of four cell derived RNAs and provide reliable reference source, we measured the copy number of certain genes in each cell line by real-time PCR method. Using real-time PCR with the PRISM 7900HT system (Applied Biosystems, Foster City, CA), we measured the copy number of several genes, including the IFN-α/β receptor, double-strand RNA-activated protein kinase (PKR), 2′,5′-oligoadenylate synthetase (2,5-AS), interferon regulatory factor-1 (IRF-1), interferon regulatory factor-3 (IRF-3), and glyceraldehyde-3-phosphate dehydrogenase (GAPDH). The primer sets were as follows: IFN-α/β receptor (forward: GTGACCTCACAGATGAGTGG; reverse: CCCTCTGA CTGTTCTTCAATGA; and probe: CACCGTCCTAGAAGGA TTCAGCGG), PKR (forward: CCTGTCCTCTGGTTCTTTTG; reverse: TGTCAGGAAGGTCAAATCTG; and probe: CTACGTGT GAGTCCCAAAGCAAC), 2,5-AS (forward: CTCAGAAATAC CCCAGCCAAATC; reverse: GTGGTGAGAGGACTGAGGAA; and probe: CCAGGTCAGCGTCAGATCGGCCTC), IRF-1 (forward: GCAAGGCCAAGAGGAAGTCA; reverse: TCATCAGG CAGAGTGGAGCT; and probe: TTCCAGCCCTGATACCTTC TCTGATGG), IRF-3 (forward: AAGGAAGGAGGCGTGTTTGA; reverse: ATTTCTACCAAGGCCCTGAGG; and probe: CGTCCGC TTCCTTCCGTGAAGGTAAT), GAPDH (forward: GAAGGTGA AGGTCGGAGT; reverse: GAAGATGGTGATGGGATTTC; and probe: CAAGCTTCCCGTTCTCAGCC). The RNA mixture was amplified using MessageAmp aRNA kit (Ambion, Austin, TX). The resulting aRNA was used as the reference aRNA. Moreover, we cloned genes (RL, GP, and LD; ~1000 bp) into pCRII TOPO vector (Invitrogen, Carlsbad, CA) as scanning range markers as well as 48 genes (in the same TOPO vector) of inflammatory genes to spike into reference aRNA. Each cloning region was designed to be larger than the size of capture probes on the microarray. Then three external control RNAs and additional spike RNAs of 48 genes were synthesized by Megascript T7 kit (Ambion, Austin, TX). Three external control RNAs were mixed as spike control mixture in both target sample and reference aRNAs. Other 48 spike RNAs were added to the reference aRNA.

Sample RNA preparation, labeling, hybridization, and scanning. Total RNA of liver biopsy samples was isolated by Isogen (Nippon Gene, Tokyo, Japan) extraction according to manufacturer's instruction. The total RNA quality was confirmed with a Bioanalyser 2100 microcapillary-electrophoretic (MEP) analyzer (Agilent Technologies, Palo Alto, CA). The 28S/18S ratio of the total RNA was >1.0. Then total RNA (1–2 μg) was transcribed and amplified to produce amplified sample RNA (aRNA) using the MessageAmp aRNA kit (Ambion, Austin, TX) according to manufacturer's instructions. Next, an external control RNA mixture (LD, GP, and RL) was added to both the sample and reference aRNAs. These mixed sample and reference aRNAs were labeled using SuperScript II kit with random hexamer (TaKaRa, Kyoto, Japan) with Cy3-dUTP and Cy5-dUTP (Perkin–Elmer, Boston, MA), respectively. Competitive hybridization of Cy3-labeled sample and Cy5-labeled reference cDNAs on the microarray was carried out according to Brown and coworkers [23]. Slides were scanned three times with ScanArray 5000 (Perkin–Elmer, Boston, MA). Each scan was carried out using the external spike level around 30,000. The data were converted from tif image data to signal using ImaGene software (BioDiscovery, El Segundo, CA) for further statistical analysis. Three file data of each three spot data of each gene were merged to establish the single representative data for each gene (Patent pending: PCT/JP03/06677). The Cy3 (patient sample)/Cy5 (reference sample) ratio of each mRNA signal was calculated for further analysis.

Patients. Liver biopsy samples from five patients receiving IFN-α monotherapy and 10 patients receiving a combination therapy (a mixture of IFN-α, IFN-β, and IFN-α/β) were obtained during 1992–2000 from Kyushu University Hospital and Nagasaki National Medical Center, respectively. Biopsy samples were stored at −80 °C. Informed consent was obtained from all patients in accordance with the Helsinki protocol. The samples were classified into three groups: sustained responders (SR) had an absence of serum HCV RNA both during the therapy and 6 months after the completion of therapy, non-responders (NR) were positive for serum HCV RNA during the

therapy, and transient responders (TR) had an absence of serum HCV RNA during the therapy or at the end of IFN treatment, but has serum HCV RNA after cessation of the therapy. Because RNA degradation may have occurred during storage, and because this can be a major source of variation in microarray data [24], we verified the quality of the extracted RNA by assessing the ribosomal RNA 28S/18S ratio.

Statistical data analysis. The merged data were subjected for hierarchical clustering to noise reduction and normalization (patent pending, PCT/JP03/06677) using the reference control and then analyzed with Genomic Profiler software (Mitsui Knowledge Industry, Tokyo, Japan). In addition, we developed a novel algorithm to calculate the Mahalanobis distance (MD) for the data from 15 patients using a standard space database (SSDB) (Eqs. (1)–(5) and Fig. 1). To establish the SSDB, we searched a gene set representing the differences between the SR/TR and NR groups. The necessary genes for the SSDB and MD calculations were selected using MATLAB (MathWorks, Natick, MA). We have calculated a graded scale utilizing variance–covariance to valuate dispersion and correlation of the standard group as a training set to establish the center of the gravity of SSDB. Once the SSDB was established, new test sample data were applied to the equations to calculate the MD. We utilize sMD as a predictive value. The sMD was presented from the center of gravity of SSDB (0:zero) along its scale to theoretically ∞. This method is one of the pattern recognition analysis dealing with correlation of multi-parameters [25].

$$d_{xi} = \frac{D_{xi} - \overline{D_x}}{\sigma_x} \quad \text{(auto scale)}, \quad (1)$$

$$S_{xx'} = \frac{\sum_{i=1}^{n} (d_{xi} - \overline{d_x})(d_{x'i} - \overline{d_{x'}})}{n-1} \quad \text{(variance–covariance)}, \quad (2)$$

$$S = \begin{bmatrix} S_{11} & S_{12} & \cdots & S_{1(k-1)} & S_{1k} \\ S_{21} & S_{22} & \cdots & S_{2(k-1)} & S_{2k} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ S_{(k-1)1} & S_{(k-1)2} & \cdots & S_{(k-1)(k-1)} & S_{(k-1)k} \\ S_{k1} & S_{k2} & \cdots & S_{k(k-1)} & S_{kk} \end{bmatrix}$$

(variance–covariance matrix), (3)

$$\mathrm{MD}^2 = [d_1 \ \cdots \ d_k] S^{-1} \begin{bmatrix} d_1 \\ \vdots \\ d_k \end{bmatrix} \quad \text{(Mahalanobis distance)}, \quad (4)$$

$$\mathrm{sMD} = \frac{\mathrm{MD}^2}{k} \quad \text{(scaled Mahalanobis distance)}. \quad (5)$$

Results and discussion

The low-density cDNA microarray

High-density microarray data were so hard to handle its huge data for analysis and difficult to understand their meaning. One approach is to minimize gene set for collection of mRNA profiling data to each category of research field. Chang et al. [26] have described the selection of data from high-density microarrays for prediction of "docetaxel" therapy efficacy of breast cancer. Specifically, they omitted low level signals of genes from the high-density microarray data at first. We followed a similar approach to select genes on our microarray, also omitting low frequency tag genes from the SAGE data. This ensured a steady state signal amongst the target samples. Based on this selection, we chose 295 genes for a low-density microarray system. The DNA sequence of each cloned gene fragment (500–600 bp) was validated by comparison with the published sequences in the GenBank database. To provide replicate data, the cDNAs were spotted in triplicate on the aminosilane-coated slide glass [8]. Scanning electron microscopy confirmed that the spots were round, smooth, and homogeneous without any doughnut features (Fig. 2). To obtain stable signals, we used three independent internal RNA references, including RL for the high expression range, GP for the middle range, and LD for the low range. The signals from the microarray were adjusted so that the ScanArray would give reliable signal of 30,000, which should be within the linear and stable range of the scanner (maximum signal = 65,535). We also carried out noise reduction and normalization of data using the artificial external spike genes as well as some housekeeping genes. Validation of the low-density cDNA microarray system was carried out using the RNA from HepG2 (data not shown).

Adjustment of reference RNA

Selection of an appropriate control reference is another important factor for accuracy in microarray analysis. One method has been to use laser microdissection to select normal tissue from the same patient as a reference. Although this is useful for single patients, it cannot be applied to comparison of multiple patients' samples because the baseline expression of specific genes can vary from patient to patient. Therefore, conditions including the duration of disease, the medications used, sampling conditions, storage conditions, and life style differences can cause variability in the microarray data. To eliminate this problem, we have used an artificial reference RNA isolated from cell lines as a reference. When we screened the signal levels of both reference and patient samples, we found 48 genes out of 295 genes in the microarray that were expressed in the patients but not or background level in the RNA mixture from the four cell lines. Typical data from the four cell reference mixture are shown in Fig. 3A. The IFN-receptor and some other well-known IFN-inducible factors are indicated. The graph shows that the levels of these mRNAs are in the low signal range. This includes the IFN-α/β receptor, even though it has been proposed as a possible marker for the prediction of IFN efficacy [18,19]. The problem in this case appears to be high variability in IFN-α/β receptor gene expression between different reference RNA preparations. To avoid this problem, we produced a large single preparation of reference RNA for future analyses. In addition, we have produced 48 synthetic RNAs, which we added to the reference RNA mixture. These synthetic RNAs were designed to be larger than the size of the capture probes on the microarray. The design and purity of some these synthetic

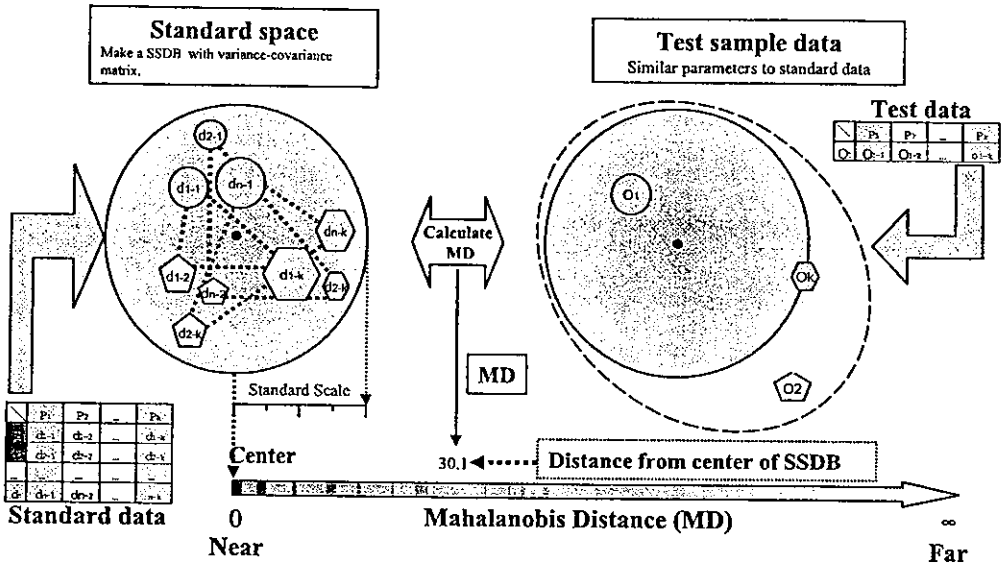


Fig. 1. Pattern recognition for establishment of SSDB and MD calculations. The standard space database (SSDB) was established based on a training data set. The parameter (1 to k) represents each data point (signal level for each gene) in the patient sample. A second parameter (1 to n) represents each patient. Both parameters were utilized along with the distinguishing genes for each group that were used in a variance–covariance matrix calculation to create the SSDB. Next, the test sample data were calculated to obtain the Mahalanobis distance (MD), where the MD represents the distance of new test sample data from the center of gravity in the SSDB. In theory, MD can be from 0 to ∞. A high MD value means that its distance is far away from the SSDB center of gravity.

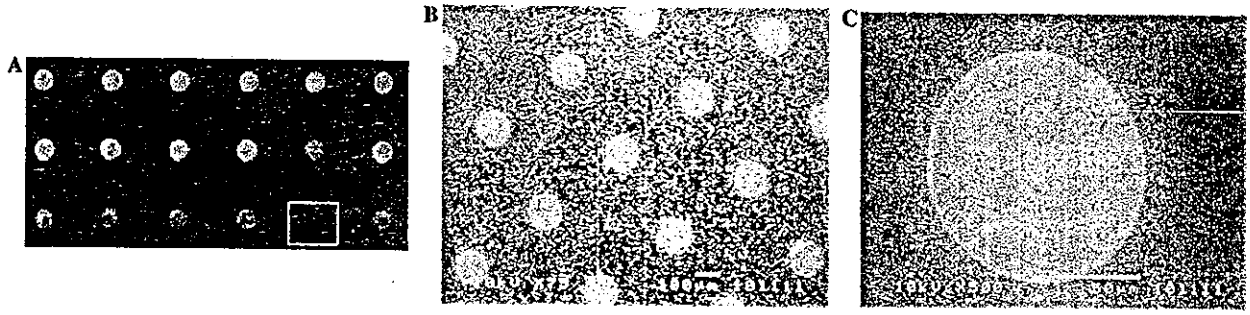


Fig. 2. Spot geometry on the cDNA microarray. (A) A typical fluorescent image was presented. The negative control (firefly luciferase) gene was spotted inside of the white square. Scanning electron microscopic images at (B) low power (75×) and (C) higher power (500×) are shown.

RNAs is shown in Fig. 3B. The level of the synthetic reference RNAs is shown as a scatter plot in Fig. 3C. These results show that the level of the synthetic RNA corresponds to the range of pseudo-inflammatory conditions. Thus, we spiked these 48 synthetic RNAs to the average signal level of patients which was surveyed at first. Without this synthetic reference RNA, it would be difficult to analyze and categorize the microarray data and use it to predict the efficacy of IFN treatment. Thus, the signal below the negative cut-off level will be treated as zero for further ratio calculation leads to nonsense.

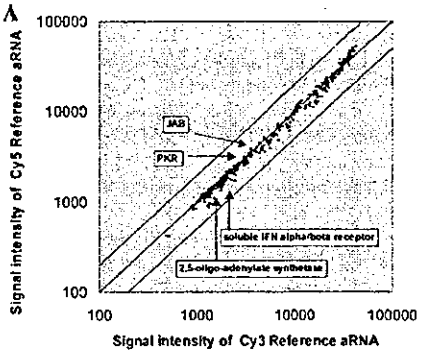
Statistical analysis

A variety of normalization techniques have been used for the analysis of DNA microarray data [9,27–30]. Many of these techniques rely on the expression of housekeeping genes. However, it is difficult to find suitable candidates, and it would be difficult to integrate a large set of housekeeping genes onto the low-density cDNA microarray [30–32]. For these reasons, we have added synthetic non-human genes as external controls. We have also utilized some type of housekeeping genes for normalization of the microarray data (patent pending, PCT/JP03/06677). Furthermore, to minimize variability in the calculations due to variability in the fluorescence measurements, we used six data files (three scans of each Cy3/Cy5 wavelength) to merge into a single representative data for each gene expression analysis.

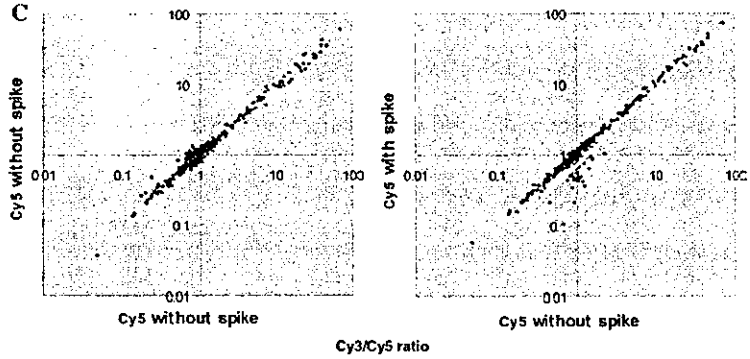
Hierarchical clustering by the classical method

Hierarchical clustering of the merged data was carried out using Euclid distance and Ward method with

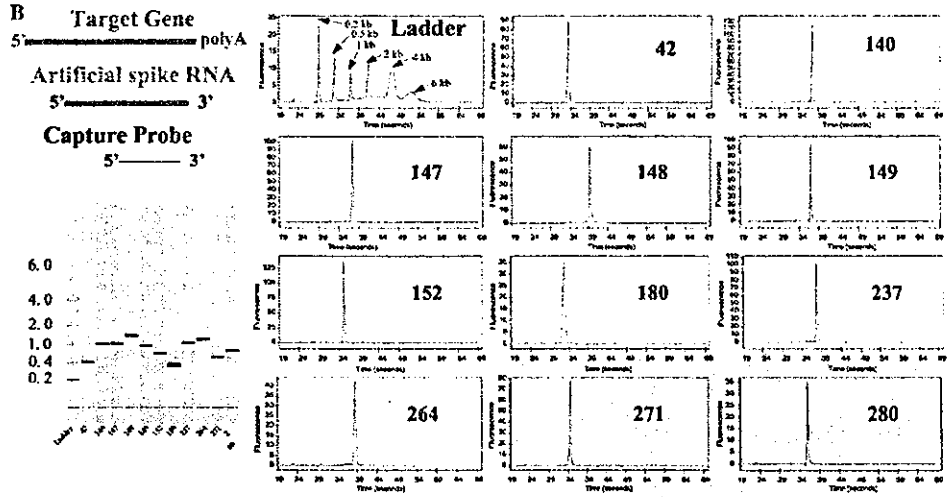
③ A


Signal intensity of Cy5 Reference aRNA
100000
10000
1000
100
100
1000
10000
100000
Signal intensity of Cy3 Reference aRNA
JAB
PKR
soluble IFN alpha/beta receptor
2,5-oligo-adenylate synthetase

C


Cy5 without spike
Cy5 without spike
Cy5 with spike
Cy5 without spike
Cy3/Cy5 ratio

B


Target Gene
5'
polyA
Artificial spike RNA
5'
3'
Capture Probe
5'
3'
6.0
4.0
2.0
1.0
0.4
0.2
Ladder
42
140
147
148
149
152
180
237
264
271
280

④

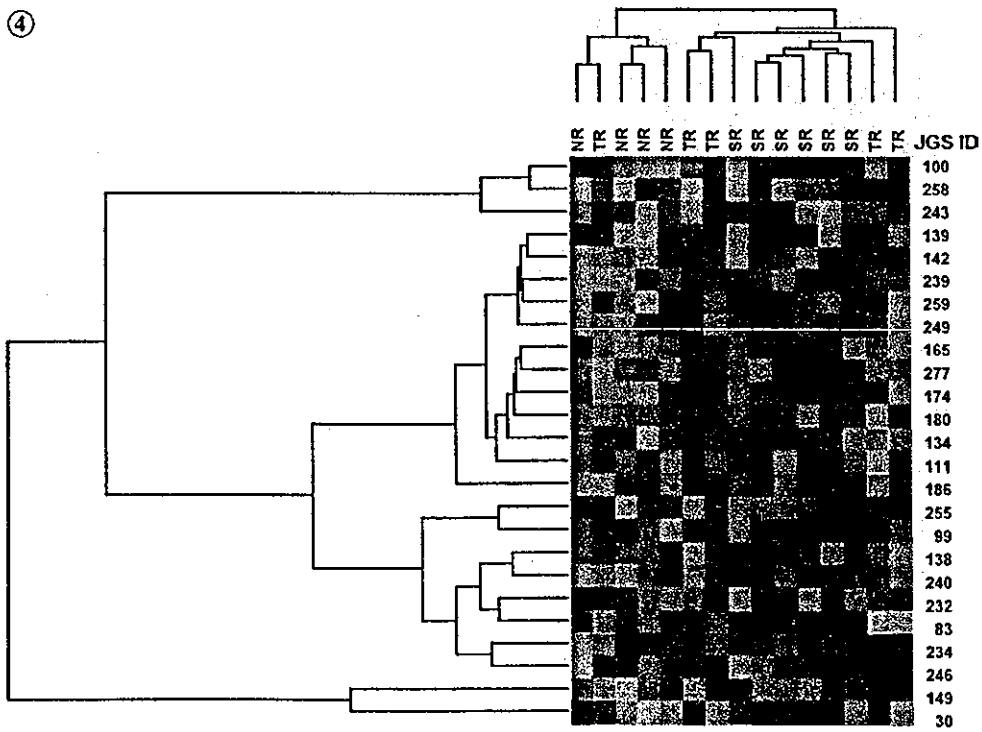

NR
TR
NR
NR
NR
TR
TR
SR
SR
SR
SR
SR
TR
TR
JGS ID
100
258
243
139
142
239
259
249
165
277
174
180
134
111
186
255
89
138
240
232
83
234
246
149
30

Fig. 3. Establishment of the artificial reference RNA. (A) Scatter plot of the mRNA signal level in the four cell RNA mixture. The RNA mixture from four cell lines was covalently modified with Cy3 and Cy5 dyes. The signal levels for typical inflammatory genes, including 2,5-oligoadenylate synthetase (2,5-AS), IFN-α/β receptor, IFN-α receptor, and PKR, are indicated with arrows. (B) Design and purity of the synthetic RNA. The general design of the synthetic RNA (positional relation), including size and position of the probe and the artificial RNA, is shown on the left. Thus, each reference RNA was designed to be longer than the captured probe on the microarray, but not to exceed the size of the target RNA. The purity of the 11 of the synthetic RNA samples is shown. The corresponding gene numbers on our cDNA microarray are shown in each panel and are as follows (GenBank accession number in parentheses): 42, gamma-G globin (X55656); 140, T cell activation antigen (CD27) (M63928); 147, (2′,5′)-oligoadenylate synthetase (D00068); 148, p68 kinase (M35663); 152, CIS3 (AB006967); 180, calcium-binding protein in macrophages (MRP-14), also known as macrophage migration inhibitory factor-related protein (X06233); 237, interleukin 2 receptor β chain (p70-75) (M26062); 264, interferon-induced protein 44 (IFI44) (NM_006417); 271, interleukin 4 (M13982); and 280, hepatocyte growth factor (X16323). (C) Synthetic spiked RNA signal level. The panel on the left represents the scatter plot without any synthetic RNA added to the reference RNA, while the panel on the right shows the reference with added synthetic RNAs. The plot shows the ratio of sample Cy3/Cy5 rather than real signal level. The spots in the white rectangle represent the level of the added synthetic RNAs.

Fig. 4. Hierarchical clustering. cDNA microarray data of 15 patients' samples were analyzed with Genomic Profiler software (MKI, Japan). For clustering, normalization, filtering, and *T* test were essential. Because of interest in predicting clinical outcomes of IFN treatment, we tried to classify the data into two groups, including non-responders (NR) and transient responders (TR)/sustained responders (SR). The accuracy of this prediction was >93%. The corresponding microarray number and according GenBank accession number of the genes responsible for clustering are shown on the right and include: 100, cytoplasmic dynein light chain 1 (U32944); 258, thymosin β-10 (M92383); 243, stathmin (X53305); 139, homeobox 1.4 protein (M74297); 142, cAMP-dependent protein kinase regulatory subunit RI-bet (M65066); 239, alternatively spliced interferon receptor (IFNAR2) (L42243); 259, eukaryotic translation initiation factor 2, subunit 1 α, 35 kDa (BC002513); 249, brain-derived neurotrophic factor precursor (BDNF) (M61176); 165, interleukin 2 (X01586); 277, natural killer cell stimulatory factor (NKSF) (M65290); 174, IFN-responsive transcription factor subunit (M87503); 180, calcium-binding protein in macrophages (MRP-14) also known as macrophage migration inhibitory factor-related protein (X06233); 134, lunatic fringe U94354); 111, protein tyrosine kinase (Syk) (L28824); 186, leukocyte-associated molecule-1 α subunit (LFA-1 α subunit) (Y00796); 255, FLICE-like inhibitory protein short form (U97075); 99, CDK4-inhibitor (p16-INK4) (L27211); 138, α 7B integrin (X74295); 240, interferon-stimulated T-cell α chemoattractant precursor(AF030514); 232, Charcot–Leyden crystal protein (L01664); 83, NADH:ubiquinone oxidoreductase MLRQ subunit (U94586); 234, apoptotic cysteine protease Mch4 (Mch4) (U60519); 246, metallothionein-III (M93311); 149, interferon regulatory factor 1 (X14454); and 30, heat shock 70 kDa protein 1A (BC002453).

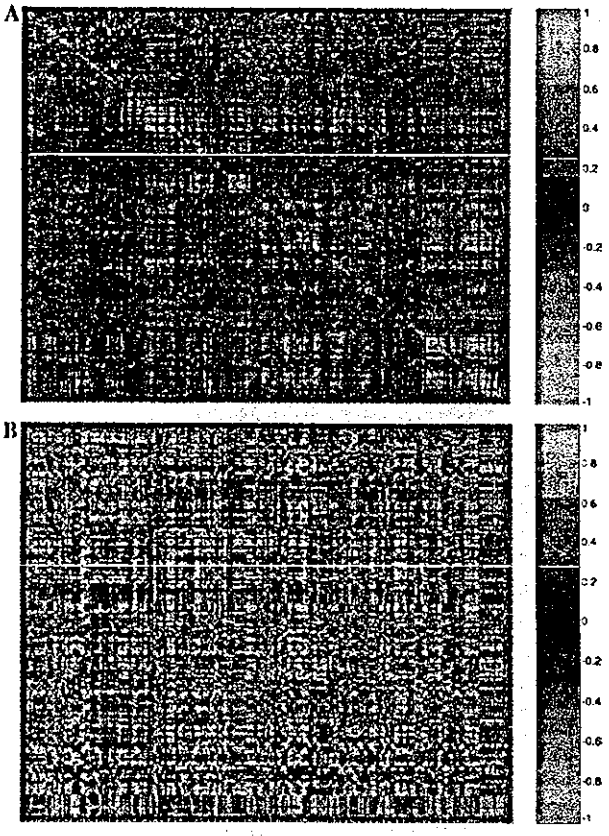


Fig. 5. Typical expression pattern: correlation matrix of 295 gene signals. (A) This expression pattern represented the established standard data as SSDB. Each axis represents genes in consideration. (B) The expression pattern shows an example of NR data as new test sample data. Color and brightness are adjusted according to Eqs. (1)–(3).

Patient No	Prediction	Clinical outcome	sMD (MD/k)*1	
10	NR	NR	6.31	so far from
9	NR	NR	5.97	SSDB center of gravity
8	NR	NR	4.38	
15	NR	NR	3.53	
4	SR or TR	SR	1.96	threshold =2.0
12	SR or TR	TR	1.90	
13	SR or TR	SR	1.44	
1	SR or TR	SR	1.36	
14	SR or TR	TR	1.09	
6	SR or TR	TR	0.98	
3	SR or TR	SR	0.73	
7	SR or TR	SR	0.67	
11	SR or TR	TR	0.38	
5	SR or TR	TR	0.24	
2	SR or TR	SR	0.24	SSDB center of gravity

		Clinical Outcome			
		SR or TR	NR	Total	%
Prediction	SR or TR	11	0	11	100%
	NR	0	4	4	100%
	Total	11	4	15	
	%	100%	100%		

$P < 0.0005$

Fig. 6. Mahalanobis distance (MD) and classification for efficacy prediction. The resulting MD data were arranged along the scale based on numbers of sMD value. The first column shows the patient number, the second column shows the prediction from our microarray analysis, and the third column shows the actual clinical outcome. Blue represents the NR group, purple represents TR, and red represents the SR group. We have set a threshold at the sMD level of 2.0 to set two groups from the sMD calculation. The hit ratio of prediction to clinical outcome was also shown.

Genomic Profiler software. Because of our interest in predicting IFN efficacy in hepatitis C, we compared the NR group against SR/TR group [33–35]. We chose genes with a 5% discriminated expression pattern (T test) between two groups ([SR vs. NR]∩[(SR + TR) vs. NR]). The hierarchical tree for these genes is shown in Fig. 4. Appropriate groups were assigned to all but one case, a TR case that was classified into the NR group.

Development of a new algorithm for hierarchical clustering

Although the classical method dealing with multi-parameters allowed satisfactory classification of the patients into two groups, this method is not useful for practical purposes. Thus, the entire data set is necessary for interpretation of the results from the classical analysis even for the analysis of a single test sample. In general, clinical diagnosis requires the common scale to compare the analytical data from samples. But a classical classification presents only a relative scale among samples for comparison.

For these reasons, we have developed a new algorithm. This method is based on the calculation of MD. The concept of the MD calculation is outlined in Fig. 1. This is one of the pattern recognition analysis and determines how close or how far from the standard group of interest. Thus, it deals with multi-parameters leading to single parameter, Mahalanobis distance (MD) as a scale, from the center of gravity of SSDB established by a training set shown in Fig. 1. Prior to the MD calculation, it was necessary to establish a SSDB with a training data set randomly selected but with clear and known clinical outcome. The standard expression pattern obtained from the SR/TR group, which was the established data source for SSDB, is shown in Fig. 5A and the new test sample pattern of the NR group is shown in Fig. 5B. The red color represents the higher expression profile and green depicts the lower expression profile with an interrelating style at the same time. Then, the pattern recognition algorithm Eqs. (1)–(3) was applied to compare the two groups. Among these expression patterns, we selected the stably and differentially expressed gene set data. Then, every selected gene expression pattern was correlated to each other like connecting a network. Thus, based on Eqs. (1)–(3), the elements that distinguish the groups shown were selected to create the SSDB. The SSDB was created based on variance and covariance. Once we established the SSDB and the center of gravity of the SSDB, we calculated a MD value for each new test sample (Eq. (4)). We utilized sMD value (divided the MD value by the number of parameters) to reduce MD value (Eq. (5)) and simplify understanding of the results[25]. The classification of IFN efficacy prediction to hepatitis C patients is shown in Fig. 6 and clearly shows that this analysis generates the NR and SR/TR groups and they are predicted with 100% accuracy ($p < 0.005$). The sample size (15 cases) was too small for statistical validation. Further detailed analyses and validation are ongoing in our laboratory and will be reported elsewhere.

Although the MD method is popular in even biological system publications [36,37], an application of MD method to microarray data is not so popular yet. The current studies are not the first published report where MD for analysis of gene expression data [38]. However, that report focused on the differential expression of a causative gene in conjunction with a standard hierarchical clustering algorithm. In our studies, we have attempted to scale the test sample position as a simple understandable parameter with a new pattern recognition algorithm. Once the SSDB scale has been established for a project, the MD can be easily used to classify new data according to the NR and TR/SR groups on an absolute scale (Fig. 6). This system will be acceptable for clinicians as a simple system for understanding the microarray data.

Conclusions

Besides the technical issues, there are many factors that control the variability within a microarray system, including individual differences between patients, the duration of the disease of each patient, different therapeutic protocols, and complications with other diseases. Some of these factors interact with each other, while others are independent. Therefore, an algorithm that allows some variability in the measurements is needed for prediction of therapeutic outcomes. The algorithm presented here appears to satisfy this requirement and it simplifies handling of large data sets. Finally, this algorithm should be generally applicable to the prediction of therapeutic outcome of diseases.

Acknowledgments

We thank Professor Kazunari Akiyoshi and Dr. Akihiko Watanabe (Tokyo Medical and Dental University) for scanning electron microscopy and Professor Kouji Matsushima (University of Tokyo) for helpful discussions.

References

[1] M. Sanchez-Carbayo, Use of high-throughput DNA microarrays to identify biomarkers for bladder cancer, Clin. Chem. 49 (2003) 23–31.
[2] P.F. Macgregor, J.A. Squire, Application of microarrays to the analysis of gene expression in cancer, Clin. Chem. 48 (2002) 1170–1177.
[3] C.H. Chung, P.S. Bernard, C.M. Perou, Molecular portraits and the family tree of cancer, Nat. Genet. 32 (Suppl.) (2002) 533–540.

[4] N.L. Harris, H. Stein, S.E. Coupland, M. Hummel, R.D. Favera, L. Pasqualucci, W.C. Chan, New approaches to lymphoma diagnosis, Hematology (Am Soc Hematol Educ Program) (2001) 194–220.
[5] L.T. Lam, O.K. Pickeral, A.C. Peng, A. Rosenwald, E.M. Hurt, J.M. Giltnane, L.M. Averett, H. Zhao, R.E. Davis, M. Sathyamoorthy, L.M. Wahl, E.D. Harris, J.A. Mikovits, A.P. Monks, M.G. Hollingshead, E.A. Sausville, L.M. Staudt, Genomic-scale measurement of mRNA turnover and the mechanisms of action of the anti-cancer drug flavopiridol, Genome Biol 2 (2001) research0041.
[6] D.G. Beer, S.L. Kardia, C.C. Huang, T.J. Giordano, A.M. Levin, D.E. Misek, L. Lin, G. Chen, T.G. Gharib, D.G. Thomas, M.L. Lizyness, R. Kuick, S. Hayasaka, J.M. Taylor, M.D. Iannettoni, M.B. Orringer, S. Hanash, Gene-expression profiles predict survival of patients with lung adenocarcinoma, Nat. Med. 8 (2002) 816–824.
[7] E. Huang, S.H. Cheng, H. Dressman, J. Pittman, M.H. Tsou, C.F. Horng, A. Bild, E.S. Iversen, M. Liao, C.M. Chen, M. West, J.R. Nevins, A.T. Huang, Gene expression predictors of breast cancer outcomes, Lancet 361 (2003) 1590–1596.
[8] M.L. Lee, F.C. Kuo, G.A. Whitmore, J. Sklar, Importance of replication in microarray gene expression studies: statistical methods and evidence from repetitive cDNA hybridizations, Proc. Natl. Acad. Sci. USA 97 (2000) 9834–9839.
[9] G.C. Tseng, M.K. Oh, L. Rohlin, J.C. Liao, W.H. Wong, Issues in cDNA microarray analysis: quality filtering, channel normalization, models of variations and assessment of gene effects, Nucleic Acids Res. 29 (2001) 2549–2557.
[10] F. Diehl, S. Grahlmann, M. Beier, J.D. Hoheisel, Manufacturing DNA microarrays of high spot homogeneity and reduced background signal, Nucleic Acids Res. 29 (2001) E38.
[11] L. Luo, R.C. Salunga, H. Guo, A. Bittner, K.C. Joy, J.E. Galindo, H. Xiao, K.E. Rogers, J.S. Wan, M.R. Jackson, M.G. Erlander, Gene expression profiles of laser-captured adjacent neuronal subtypes, Nat. Med. 5 (1999) 117–122.
[12] O. Kitahara, Y. Furukawa, T. Tanaka, C. Kihara, K. Ono, R. Yanagawa, M.E. Nita, T. Takagi, Y. Nakamura, T. Tsunoda, Alterations of gene expression during colorectal carcinogenesis revealed by cDNA microarrays after laser-capture microdissection of tumor tissues and normal epithelia, Cancer Res. 61 (2001) 3544–3549.
[13] L. Assersohn, L. Gangi, Y. Zhao, M. Dowsett, R. Simon, T.J. Powles, E.T. Liu, The feasibility of using fine needle aspiration from primary breast cancers for cDNA microarray analyses, Clin. Cancer Res. 8 (2002) 794–801.
[14] D.A. Wigle, I. Jurisica, N. Radulovich, M. Pintilie, J. Rossant, N. Liu, C. Lu, J. Woodgett, I. Seiden, M. Johnston, S. Keshavjee, G. Darling, T. Winton, B.J. Breitkreutz, P. Jorgenson, M. Tyers, F.A. Shepherd, M.S. Tsao, Molecular profiling of non-small cell lung cancer and correlation with disease-free survival, Cancer Res. 62 (2002) 3005–3008.
[15] T.C. Van Der Pouw Kraan, F.A. Van Gaalen, T.W. Huizinga, E. Pieterman, F.C. Breedveld, C.L. Verweij, Discovery of distinctive gene expression profiles in rheumatoid synovium using cDNA microarray technology: evidence for the existence of multiple pathways of tissue destruction and repair, Genes Immun. 4 (2003) 187–196.
[16] T. Yamashita, S. Hashimoto, S. Kaneko, S. Nagai, N. Toyoda, T. Suzuki, K. Kobayashi, K. Matsushima, Comprehensive gene expression profile of a normal human liver, Biochem. Biophys. Res. Commun. 269 (2000) 110–116.
[17] T. Yamashita, S. Kaneko, S. Hashimoto, T. Sato, S. Nagai, N. Toyoda, T. Suzuki, K. Kobayashi, K. Matsushima, Serial analysis of gene expression in chronic hepatitis C and hepatocellular carcinoma, Biochem. Biophys. Res. Commun. 282 (2001) 647–654.
[18] H. Yatsuhashi, K. Yamasaki, T. Aritomi, P. Maria, D. Carmen, O. Inoue, M. Koga, M. Yano, Quantitative analysis of interferon alpha/beta receptor mRNA in the liver of patients with chronic hepatitis C: correlation with serum hepatitis C virus-RNA levels and response to treatment with interferon, J. Gastroenterol. Hepatol. 12 (1997) 460–467.
[19] E. Mizukoshi, S. Kaneko, M. Yanagi, H. Ohno, K. Kaji, S. Terasaki, A. Shimoda, E. Matsushita, K. Kobayashi, Expression of interferon alpha/beta receptor in the liver of chronic hepatitis C patients, J. Med. Virol. 56 (1998) 217–223.
[20] F.L. Dumoulin, U. Wennrich, H.D. Nischalke, L. Leifeld, H.P. Fischer, T. Sauerbruch, U. Spengler, Intrahepatic mRNA levels of interferon gamma and tumor necrosis factor alpha and response to antiviral treatment of chronic hepatitis C, J. Hum. Virol. 4 (2001) 195–199.
[21] H.F. Kawai, S. Kaneko, M. Honda, Y. Shirota, K. Kobayashi, Alpha-fetoprotein-producing hepatoma cell lines share common expression profiles of genes in various categories demonstrated by cDNA microarray analysis, Hepatology 33 (2001) 676–691.
[22] T. Ito, K. Yasui, J. Mukaigawa, A. Katsume, M. Kohara, K. Mitamura, Acquisition of susceptibility to hepatitis C virus replication in HepG2 cells by fusion with primary human hepatocytes: establishment of a quantitative assay for hepatitis C virus infectivity in a cell culture system, Hepatology 34 (2001) 566–572.
[23] A.B. Khodursky, B.J. Peter, N.R. Cozzarelli, D. Botstein, P.O. Brown, C. Yanofsky, DNA microarray analysis of gene expression in response to physiological and genetic changes that affect tryptophan metabolism in *Escherichia coli*, Proc. Natl. Acad. Sci. USA 97 (2000) 12170–12175.
[24] M. Ellis, N. Davis, A. Coop, M. Liu, L. Schumaker, R.Y. Lee, R. Srikanchana, C.G. Russell, B. Singh, W.R. Miller, V. Stearns, M. Pennanen, T. Tsangaris, A. Gallagher, A. Liu, A. Zwart, D.F. Hayes, M.E. Lippman, Y. Wang, R. Clarke, Development and validation of a method for using breast core needle biopsies for gene expression microarray analyses, Clin. Cancer Res. 8 (2002) 1155–1166.
[25] G. Taguchi, J. Rajesh, The Mahalanobis-Taguchi strategy, ed., Wiley, New York, 2002.
[26] J.C. Chang, E.C. Wooten, A. Tsimelzon, S.G. Hilsenbeck, M.C. Gutierrez, R. Elledge, S. Mohsin, C.K. Osborne, G.C. Chamness, D.C. Allred, P. O'Connell, Gene expression profiling for the prediction of therapeutic response to docetaxel in patients with breast cancer, Lancet 362 (2003) 362–369.
[27] C. Workman, L.J. Jensen, H. Jarmer, R. Berka, L. Gautier, H.B. Nielser, H.H. Saxild, C. Nielsen, S. Brunak, S. Knudsen, A new non-linear normalization method for reducing variability in DNA microarray experiments, Genome Biol 3 (2002) research0048.
[28] J.H. Kim, D.M. Shin, Y.S. Lee, Effect of local background intensities in the normalization of cDNA microarray data with a skewed expression profiles, Exp. Mol. Med. 34 (2002) 224–232.
[29] Y.H. Yang, S. Dudoit, P. Luu, D.M. Lin, V. Peng, J. Ngai, T.P. Speed, Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation, Nucleic Acids Res. 30 (2002) e15.
[30] X. Wang, S. Ghosh, S.W. Guo, Quantitative quality control in microarray image processing and data acquisition, Nucleic Acids Res. 29 (2001) E75.
[31] E.M. Glare, M. Divjak, M.J. Bailey, E.H. Walters, β-Actin and GAPDH housekeeping gene expression in asthmatic airways is variable and not suitable for normalising mRNA levels, Thorax 57 (2002) 765–770.
[32] P.D. Lee, R. Sladek, C.M. Greenwood, T.J. Hudson, Control genes and variability: absence of ubiquitous reference transcripts in diverse mammalian expression studies, Genome Res. 12 (2002) 292–297.
[33] S. Nishiguchi, T. Kuroki, S. Nakatani, H. Morimoto, T. Takeda, S. Nakajima, S. Shiomi, S. Seki, K. Kobayashi, S. Otani, Randomised trial of effects of interferon-alpha on incidence of

hepatocellular carcinoma in chronic active hepatitis C with cirrhosis, Lancet 346 (1995) 1051–1055.
[34] K. Kuwana, T. Ichida, T. Kamimura, S. Ohkoshi, N. Ogata, T. Harada, K. Endoh, H. Asakura, Risk factors and the effect of interferon therapy in the development of hepatocellular carcinoma: a multivariate analysis in 343 patients, J. Gastroenterol. Hepatol. 12 (1997) 149–155.
[35] M.R. Brunetto, F. Oliveri, M. Koehler, F. Zahm, F. Bonino, Effect of interferon-alpha on progression of cirrhosis to hepatocellular carcinoma: a retrospective cohort study. International Interferon-alpha Hepatocellular Carcinoma Study group, Lancet 351 (1998) 1535–1539.
[36] K.R. Coombes, H.A. Fritsche Jr., C. Clarke, J.N. Chen, K.A. Baggerly, J.S. Morris, L.C. Xiao, M.C. Hung, H.M. Kuerer, Quality control and peak finding for proteomics data collected from nipple aspirate fluid by surface-enhanced laser desorption and ionization, Clin. Chem. 49 (2003) 1615–1623.
[37] O. Samek, H.H. Telle, D.C. Beddows, Laser-induced breakdown spectroscopy: a tool for real-time, in vitro and in vivo identification of carious teeth, BMC Oral Health 1 (2001) 1.
[38] A. Chilingaryan, N. Gevorgyan, A. Vardanyan, D. Jones, A. Szabo, Multivariate approach for selecting sets of differentially expressed genes, Math. Biosci. 176 (2002) 59–69.