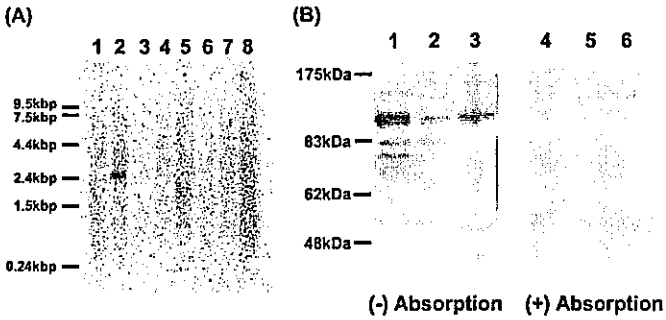

FIG. 1. **Tissue distribution of Oatp14.** *A*, Northern blotting. Commercially available rat multiple tissue Northern blots containing 2 μg of poly(A)$^+$ RNA was hybridized for 3 h using the Oatp14 fragment as a probe. *Lane 1*, heart; *lane 2*, brain; *lane 3*, spleen; *lane 4*, lung; *lane 5*, liver; *lane 6*, skeletal muscle; *lane 7*, kidney; *lane 8*, testis. *B*, Western blotting. Choroid plexus (*lanes 1* and *4*, 50 μg), brain homogenate (*lanes 2* and *5*, 50 μg), and isolated brain capillary (*lanes 3* and *6*, 50 μg) were separated by SDS-PAGE (10% separating gel). Oatp14 was detected by anti-Oatp14 polyclonal antibody.

porter 1 (BSAT1) using gene microarray techniques by comparing the gene expression profile of cDNA from the brain capillary with that from the liver and kidney. BSAT1 cDNA consisted of 2148 bp that encoded a 716-amino acid residues protein with 12 putative membrane-spanning domains. BSAT1 was highly enriched in the brain capillary compared with brain homogenate, liver, and kidney. Comparison of the cDNA sequences of BSAT1 revealed that it is the 14th member of the Oatp/OATP family (Oatp14). Although the localization at the BBB and the substrates of this isoform remain unknown, BBB-specific expression prompted us to hypothesize that Oatp14 accounts for the efflux of organic anions including $E_2$17βG via the BBB, together with Oatp2. The purpose of the present study is to characterize the substrate specificity and spectrum of inhibitors of Oatp14, as well as its tissue distribution and localization. Through this study, we found out that thyroxine (T_4) is a good substrate of Oatp14, and the expression level of Oatp14 in the BBB is affected by plasma thyroid conditions. The results of the present study suggest that Oatp14 plays an important role in regulating the concentration of T_4 in the central nervous system and in brain development.

EXPERIMENTAL PROCEDURES

Chemicals—[^{3}H]Leu-enkephalin was purchased from American Radiolabeled Chemicals (St. Louis, MO). [^{3}H]Pravastatin was kindly donated from Sankyo (Tokyo, Japan), [^{14}C]cerivastatin was from Bayer AG (Wuppertal, Germany), and [^{14}C]E3040 glucuronide and [^{14}C]E3040 sulfate were from Eisai (Tokyo, Japan). [^{3}H]Taurolithocholate sulfate (TLCS), [^{35}S]4-methylumbelliferone sulfate (4-MUS), and [^{35}S]troglitazone sulfate (TRO-S) were synthesized according to a method described previously (20, 21). The radiochemical purity of [^{3}H]TLCS, [^{35}S]4-MUS, and [^{35}S]TRO-S prepared by this method were more than 95%. Other labeled compounds were purchased from PerkinElmer Life Science. Unlabeled pravastatin, troglitazone, and its conjugated metabolites were kindly donated from Sankyo, unlabeled cerivastatin was from Bayer AG, and unlabeled E3040 glucuronide and E3040 sulfate were from Eisai. All other chemicals were commercially available, of reagent grade, and were used without any purification.

Capillary Isolation—Rat brain capillaries were isolated using a modification of the procedure of Boado *et al.* (22). All steps in the isolation procedure were carried out at 4 °C in pregassed (95% O_2-5% CO_2) solutions. Briefly, pieces of gray matter were gently homogenized in three volumes (v/w) of an artificial extracellular fluid buffer and, after addition of dextran (final concentration 15%), the homogenate was centrifuged at low speed. The resulting pellet was resuspended in Buffer B (103 mM NaCl, 25 mM $NaHCO_3$, 10 mM D-glucose, 4.7 mM KCl, 2.5 mM $CaCl_2$, 1.2 mM $MgSO_4$, 1.2 mM K_2HPO_4, and 15 mM HEPES, 1 mM sodium pyruvate, 0.5% (w/v) bovine serum albumin, pH 7.4) and then filtered through a 200-μm nylon mesh. The filtrate was passed over a column of glass beads, and after washing with Buffer B, the capillaries adhering to the beads were collected by gentle agitation.

FIG. 2. **Immunohistochemical staining of Oatp14 for brain slices.** Frozen sections of rat brain were used for immunohistochemical detection with peroxidase to probe for Oatp14 with a polyclonal antibody. The *lined* and *dotted arrows* represent luminal and abluminal sides of brain capillary endothelial cells, respectively. Positive labeling was only found in the border of brain capillary endothelial cells.

Northern Blot Analysis—A commercially available hybridization blot containing poly(A)$^+$ RNA from various rat tissues (rat multi-tissue Northern blot; Clontech) was used for the Northern blot analysis. A fragment (position numbers 1–807) from Oatp14 was used as a probe, and its nucleotide sequence showed less than 60% identity with other members of the Oatp family. The master blot filter was hybridized with the ^{32}P-labeled probe at 68 °C according to manufacturer's instructions. The filter was washed finally under high stringency conditions (0.1× SSC (1× SSC = 0.15 M NaCl and 0.015 M sodium citrate)) and 0.1% SDS at 65 °C and then exposed to Fuji imaging plates (Fuji Photo Film, Kanagawa, Japan) for 3 h at room temperature and examined using an imaging analyzer (BAS 2000; Fuji Photo Film).

Western Blot Analysis—Antiserum against Oatp14 was raised in rabbits against a synthetic peptide consisting of the 17 carboxyl-terminal amino acids of Oatp14. Antiserum was purified by affinity column chromatography using the antigen and used for subsequent analyses. Choroid plexus, brain homogenate, and isolated brain capillary samples were diluted with Loading Buffer (BioLabs, Hertfordshire, United Kingdom). They were then boiled for 3 min and loaded onto an 8.5% SDS-polyacrylamide electrophoresis gel with a 3.75% stacking gel. Proteins were electroblotted onto a polyvinylidene difluoride membrane (Pall Filtran, Karlstein, Germany) using a blotter (Trans-blot; Bio-Rad) at 15 V for 1 h. The membrane was blocked with TBS-T (Tris-buffered

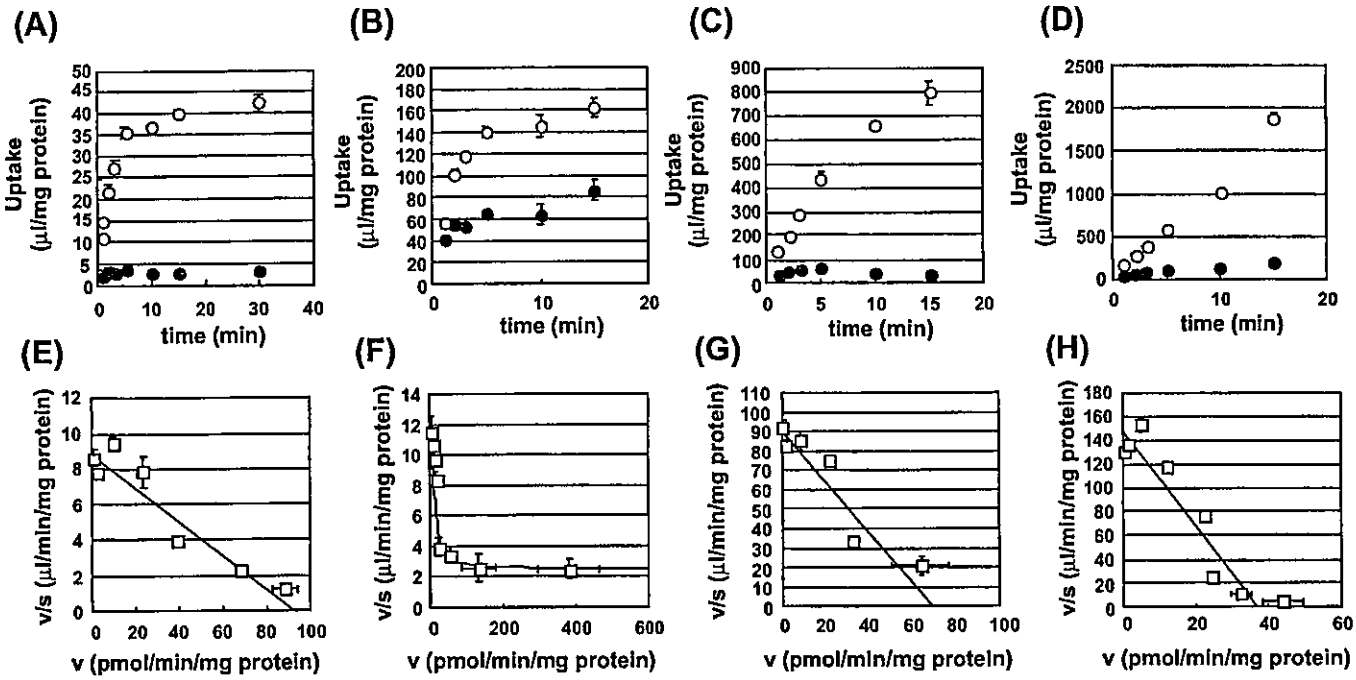

FIG. 3. **Time profiles and concentration dependence of the uptake of [^{3}H]$E_2$17βG, [^{14}C]cerivastatin, [^{35}S]TRO-S, and [^{125}I]T_4 by Oatp14-transfected HEK293 cells.** The uptake of [^{3}H]$E_2$17βG (*A* and *E*), [^{14}C]cerivastatin (*B* and *F*), [^{35}S]TRO-S (*C* and *G*), and [^{125}I]T_4 (*D* and *H*) by Oatp14-transfected HEK293 cells was examined at 37 °C. The *upper graphs* show the time profiles. *Open* and *closed circles* represent the uptake by Oatp14-transfected cells and vector-transfected cells, respectively. The *lower graphs* show the concentration dependence. Specific uptake was obtained by subtracting the uptake by vector-transfected cells from that by Oatp14-transfected cells. Each *point* represents the mean ± S.E. (n = 3).

saline containing 0.05% Tween 20) and 5% skimmed milk for 1 h at room temperature. After washing with TBS-T, the membrane was incubated with the antibodies (dilution 1:1000). The membrane was allowed to bind a horseradish peroxidase-labeled anti-rabbit IgG antibody (Amersham Biosciences) diluted 1:5000 in TBS-T for 1 h at room temperature followed by washing with TBS-T.

Immunohistochemical Study—Frozen sections from male Sprague-Dawley rats were prepared for the immunohistochemical study after fixing in acetone (−20 °C). The brain slices adhered to the glass cover slips were washed with PBS and fixed for 10 min on ice in acetone. After washing with PBS, the capillaries were permearized in 0.2% (v/v) Triton X-100 in PBS and incubated with peroxidase blocking reagent (DAKO, Carpinteria, CA) for 10 min at room temperature to block nonspecific peroxidase. Slices were incubated with anti-Oatp14 antibody (1:100) for 60 min at room temperature, washed three times with PBS, and subsequently incubated for 60 min at room temperature with the horseradish peroxidase-labeled anti-rabbit secondary antibody (Envision+ system; DAKO). The immune reaction was visualized using diaminobenzidine and then nuclei were stained with hematoxylin (DAKO). The specificity of the antibody reaction was verified by negative controls, which were incubated with polyclonal antibody that had been blocked with the antigenic peptide.

Cloning of Rat Oatp14 cDNA—Based on the nucleotide sequence reported by Li *et al.* (19) (GenBank™ accession number NM 053441), the following primers were designed to isolate Oatp14 cDNA encoding a full open reading frame of Oatp14: forward primer, 5′-ggaattccgccaccatggacacttcatccaaaga-3′ and reverse primer, 5′-ggattccttaaagtcggtctccttgc-3′. PCR was performed using cDNA prepared from rat brain as template according to the following protocol: 96 °C for 1 min, 55 °C for 1 min, and 72 °C for 2 min; 50 cycles. PCR products were subcloned to pGEM-T Easy Vector (Promega, Madison, WI) and sequenced. The nucleotide sequence of rat Oatp14 cDNA was identical as being that of BSAT1 except for one base change (A175G) resulting in a change of amino acid (T59A). However, it was confirmed that this change was not because of an error accumulated during PCR by sequencing the RT-PCR products directly.

Stable Expression of Oatp14 cDNA in HEK293 Cells—The Oatp14-cDNA was subcloned into the pcDNA3.1(+) (Invitrogen) and introduced into HEK293 cells by lipofection with FuGENE 6 (Roche Diagnostics) according to the manufacturer's protocol and were selected by culturing them in the presence of G418 sulfate (800 μg/ml; Invitrogen). HEK293 cells were grown in minimum essential medium (Invitrogen) supplemented with 10% fetal bovine serum, penicillin (100 units/ml), streptomycin (100 μg/ml), and G418 sulfate (400 μg/ml) at 37 °C with 5% CO_2 and 95% humidity. Cells were incubated for 24 h before starting the experiments with culture medium supplemented with sodium butyrate (5 mM).

Transport Study—Uptake was initiated by adding the radiolabeled ligands to the incubating buffer in the presence and absence of inhibitors after cells had been washed three times and preincubated with Krebs-Henseleit buffer (142 mM NaCl, 23.8 mM $NaHCO_3$, 4.83 mM KCl, 0.96 mM KH_2PO_4, 1.20 mM $MgSO_4$, 12.5 mM HEPES, 5 mM glucose, and 1.53 mM $CaCl_2$, adjusted to pH 7.4) at 37 °C for 15 min. For the efflux study, cells were preincubated with [^{125}I]T_4 at 37 °C for 15 min and washed three times with ice-cold Krebs-Henseleit buffer, followed by incubation in the absence of [^{125}I]T_4 with Krebs-Henseleit buffer at 37 °C. The uptake and efflux were terminated at designed times by adding ice-cold Krebs-Henseleit buffer. The radioactivity associated with the cells and medium specimens was determined in a liquid scintillation counter. The remaining aliquots of cell lysates were used to determine protein concentrations by the method of Lowry (23) with bovine serum albumin as a standard. Ligand uptake is given as the cell-to-medium concentration ratio determined as the amount of ligand associated with the cells divided by the medium concentration. Specific uptake was obtained by subtracting the uptake by vector-transfected cells from that by Oatp14-expressed cells.

Kinetic Analyses—Kinetic parameters were obtained from the following Michaelis-Menten equation, $v = V_{max}S/(K_m + S)$, where v is the uptake rate of the substrate (pmol/min/mg protein), S is the substrate concentration in the medium (μM), K_m is the Michaelis-Menten constant (μM), and V_{max} is the maximum uptake rate (pmol/min/mg protein). To obtain the kinetic parameters, the equation was fitted to the initial uptake velocity. The experimental data were fitted to the equation by nonlinear regression analysis with weighting as the reciprocal of the observed values, and the Damping Gauss Newton Method algorithm was used for fitting. Inhibition constants (K_i) for Oatp14-mediated transport were calculated assuming competitive inhibition.

Production of Hyperthyroid and Hypothyroid Conditions—Male Sprague-Dawley rats, weighing 200–220 g, were purchased from Japan SLC (Shizuoka, Japan). Rats had free access to food and water at all times during the study. Production of hyperthyroid and hypothyroid conditions involved a modification of the procedure of Burmeister *et al.* (24). Hypothyroidism was induced by the addition of 0.05% methimazole (MMI), an inhibitor for thyroid hormone synthesis in the thyroid gland, to the drinking water or thyroidectomy. Hypothyroidism was assessed clinically by failure to gain weight at the expected rate and

TABLE I
Substrate specificity of Oatp14

Oatp14-expressed and vector-transfected HEK293 cells were grown to confluence. After 24 h of incubation in culture medium including 5 mM sodium butyrate, uptake was examined in Krebs-Henseleit buffer.

Substrates	Oatp14	pcDNA	Ratio, Oatp14/pcDNA
	μl/mg protein/15 min		
CA	5.66 ± 0.30	5.24 ± 0.43	1.1 ± 0.1
GCA	18.7 ± 2.3	15.7 ± 1.6	1.2 ± 0.2
TCA	8.09 ± 0.52	6.06 ± 0.10	1.3 ± 0.1[a]
LCA	576 ± 6	535 ± 18	1.1 ± 0.0
CDCA	65.7 ± 6.7	84.1 ± 11.5	0.8 ± 0.1
UDCA	10.5 ± 0.1	9.71 ± 0.35	1.1 ± 0.0
TLCS	61.5 ± 1.9	43.7 ± 1.7	1.4 ± 0.1
Estradiol	204 ± 4	186 ± 13	1.1 ± 0.1
Testosterone	52.3 ± 1.5	39.2 ± 1.4	1.3 ± 0.1[b]
Dihydrotestosterone	147 ± 15	106 ± 5	1.4 ± 0.2[b]
Corticosterone	24.7 ± 1.5	20.9 ± 0.5	1.2 ± 0.1
Estrone	303 ± 15	248 ± 6	1.2 ± 0.1[a]
DHEAS	9.68 ± 0.24	7.72 ± 0.48	1.3 ± 0.1[a]
Estrone-sulfate	11.1 ± 0.8	6.3 ± 0.1	1.7 ± 0.1[a]
$E_2$17βG	50.1 ± 4.7	2.4 ± 0.2	21.2 ± 2.9[b]
LTC4	14.5 ± 0.6	13.4 ± 0.1	1.1 ± 0.0[b]
LTD4	19.6 ± 0.8	18.0 ± 2.0	1.1 ± 0.1
LTE4	30.5 ± 1.3	26.0 ± 0.9	1.2 ± 0.1[a]
PGD2	3.50 ± 0.15	3.61 ± 0.12	1.0 ± 0.1
PGE2	6.99 ± 0.46	5.80 ± 0.24	1.2 ± 0.1
Leu-Enkephalin	54.2 ± 2.9	43.3 ± 2.5	1.3 ± 0.1[a]
CCK-8	2.58 ± 0.21	1.81 ± 0.13	1.4 ± 0.2[a]
T3	951 ± 16	733 ± 4	1.3 ± 0.0[b]
Reverse T3	1397 ± 79	71 ± 5	19.7 ± 1.7[b]
T4	1456 ± 10	124 ± 3	11.8 ± 0.3[b]
Ketoprofen	9.53 ± 0.42	6.91 ± 0.26	1.4 ± 0.1[b]
Ibuprofen	3.18 ± 0.11	3.99 ± 1.18	0.8 ± 0.2
Indomethacin	31.3 ± 0.9	34.3 ± 2.0	0.9 ± 0.1
Benzylpenicillin	5.76 ± 0.47	5.45 ± 0.12	1.1 ± 0.1
OchratoxinA	8.58 ± 1.29	5.81 ± 0.14	1.5 ± 0.2[b]
Qunidine	1390 ± 34	1274 ± 150	1.1 ± 0.1
Cerivastatin	105 ± 2	33 ± 1	3.1 ± 0.1[b]
Pravastatin	5.62 ± 0.38	3.50 ± 0.75	1.6 ± 0.4[a]
Digoxin	8.70 ± 0.13	9.96 ± 0.22	0.9 ± 0.0
E3040	106 ± 3	93 ± 5	1.1 ± 0.1
E3040G	10.7 ± 1.0	2.12 ± 0.22	5.1 ± 0.7[a]
E3040S	4.78 ± 0.35	1.69 ± 0.26	2.8 ± 0.5[b]
4MUS	1.82 ± 0.35	0.90 ± 0.03	2.0 ± 0.1[b]
Troglitazone-sulfate	64.1 ± 14.3	8.4 ± 0.4	7.6 ± 1.7[b]

[a] Statistically significant uptake is indicated. $p < 0.05$.
[b] Statistically significant uptake is indicated. $p < 0.01$.

TABLE II
K_m, V_{max} and V_{max}/K_m values for Oatp14

The K_m and V_{max} values were determined by nonlinear regression analysis using data shown in Fig. 3.

Substrate	K_m	V_{max}	V_{max}/K_m
	μM	*pmol/min/mg protein*	*μl/min/mg protein*
$E_2$17βG	10.7 ± 1.6	93.4 ± 10.4	8.73 ± 1.63
Cerivastatin	1.34 ± 0.25	14.5 ± 2.2	10.8 ± 2.6
TRO-S	0.76 ± 0.09	69.0 ± 6.7	91.3 ± 14.2
T4	0.18 ± 0.03	32.1 ± 2.5	147 ± 14

could be observed within 2 weeks of the beginning MMI treatment and within 1 week after thyroidectomy. Hyperthyrodism was produced by giving L-T3 (50 μg/100 g body weight, subcutaneously, daily) 4 days before capillary isolation.

RESULTS

Tissue Distribution of Oatp14—The expression of Oatp14 mRNA in rat tissues was investigated by Northern blot analysis (Fig. 1*A*). A band was detected at 2.6 kbp, predominantly in the brain. No hybridization signals were detected in mRNA isolated from other tissues, including the heart, spleen, lung, liver, skeletal muscle, kidney, and testis.

Immunoblot and Immunohistochemical Staining of Oatp14—The expression of Oatp14 in the choroid plexus, brain homogenate, and brain capillary was examined by Western blot analysis (Fig. 1*B*). Immunoreactive protein was detected at ~90 kDa in the choroid plexus, brain homogenate, and brain capillary. These bands were abolished when preabsorbed polyclonal antibody for Oatp14 was used, suggesting that the positive bands were specific for the antigen peptide.

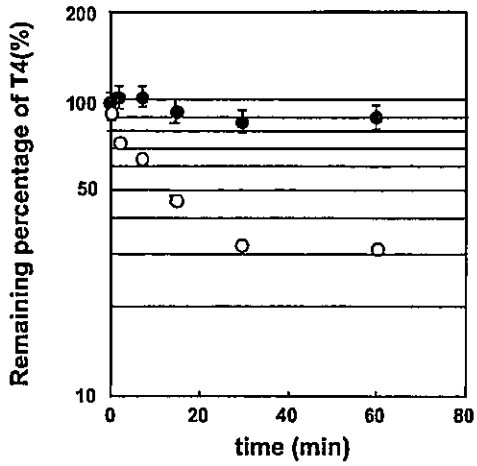

FIG. 4. **Time profiles of the efflux of [^{125}I]T_4 from Oatp14-transfected HEK293 cells.** The efflux of preloaded [^{125}I]T_4 from Oatp14-transfected HEK293 cells was examined at 37 °C. *Open* and *closed circles* represent the efflux from Oatp14- and vector-transfected cells, respectively. Each *point* represents the mean ± S.E. ($n = 3$).

To investigate the localization of Oatp14 in brain capillary endothelial cells, immunohistochemical staining was carried out using anti-Oatp14 polyclonal antibody (Fig. 2). Positive signals for anti-Oatp14 polyclonal antibody were detected in brain capillary endothelial cells. The signals were detected along the plasma membrane of brain capillary endothelial cells. The signal was abolished by preincubating the polyclonal antibody of Oatp14 with antigen (data not shown).

Transport Properties of Oatp14—Fig. 3 shows the time profiles of the uptake of [^{3}H]$E_2$17βG (*A*), [^{14}C]cerivastatin (*B*), [^{35}S]TRO-S (*C*), and [^{125}I]T_4 (*D*) by Oatp14-expressed HEK293 cells and vector-transfected HEK293 cells. Their uptake by Oatp14-expressed cells is markedly greater than that by vector-transfected cells. This Oatp14-mediated uptake showed saturation kinetics and followed the Michaelis-Menten equation (Fig. 3, *E-H*). The kinetic parameters for the uptake by Oatp14 were determined by nonlinear regression analysis and summarized in Table I. The uptake of various organic anions by Oatp14 was investigated, and the results are summarized in Table II. The uptake of [^{14}C]E3040 glucuronide, [^{14}C]E3040 sulfate, [^{14}C]4-MUS, and [^{125}I]reverse T_3 by Oatp14-expressed cells was significantly greater compared with that by vector-transfected (Table II). Although the triiodothyronine (T_3) uptake by Oatp14-expressed cells was significantly greater than that by vector-transfected cells, the Oatp14-mediated uptake for T_3 was ~6-fold smaller than that of T_4 and reverse T_3 by Oatp14 (Table II). The difference in the uptake of [^{3}H]taurocholate, [^{3}H]TLCS, [^{3}H]testosterone, [^{3}H]dihydrotestosterone, [^{3}H]estrone, [^{3}H]estrone sulfate (ES), [^{3}H]dehydroepiandrosterone sulfate, [^{3}H]leukotriene E_4 (LTE_4), [^{3}H]Leu-enkephalin, [^{3}H]cholecystokinin-octapeptide (CCK-8), [^{125}I]T_3, [^{3}H]pravastatin, [^{3}H]ketoprofen, and [^{3}H]ochratoxin A was statistically significant between Oatp14-expressed and vector-transfected cells, although the rates of uptake were very low (Table II).

To investigate whether Oatp14 can mediate bidirectional transport, cells were preloaded with [^{125}I]T_4 for 15 min followed by incubation in the absence of [^{125}I]T_4. The radioactivity associated with cell specimens was rapidly reduced in Oatp14-expressed HEK293 cells compared with that in vector-trans-

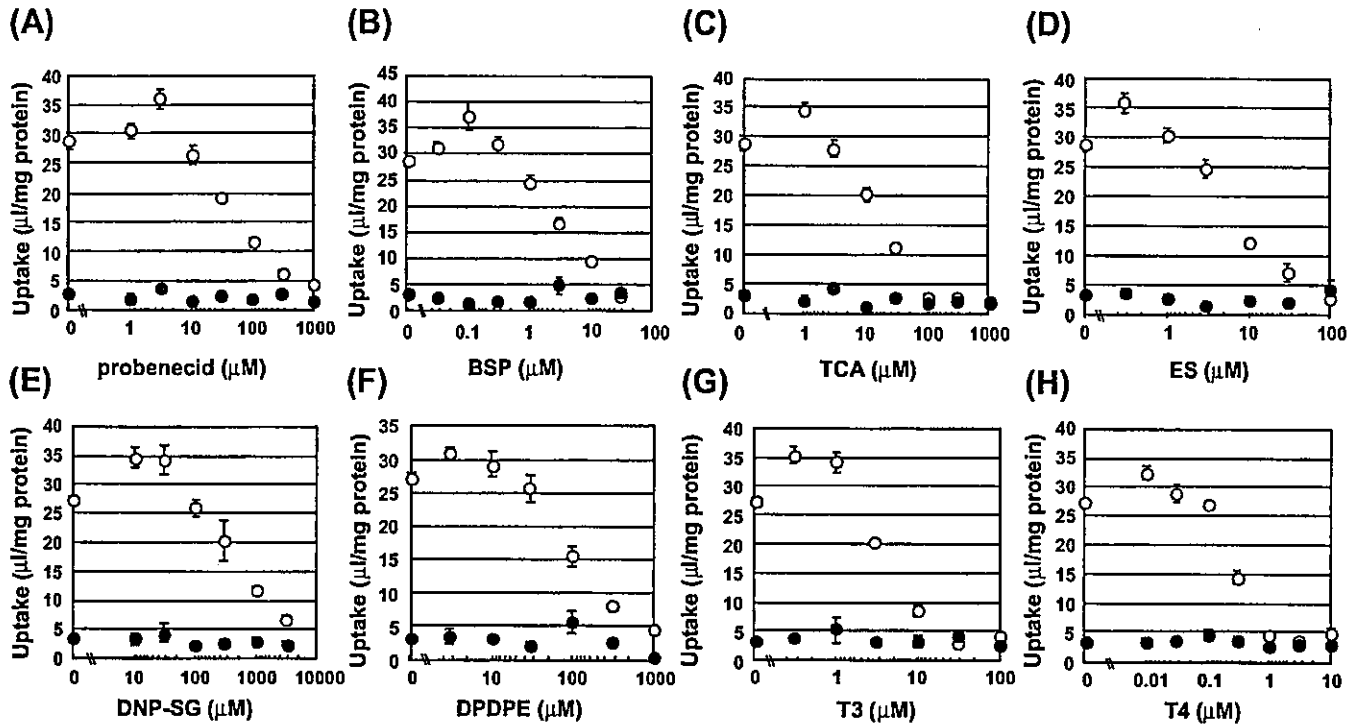

FIG. 5. **Effects of unlabeled probenecid, BSP, trichloroacetic acid, ES, DNP-SG, DPDPE, T_3, and T_4 on the uptake of [^{3}H]$E_2$17βG by Oatp14-transfected HEK293 cells.** The effects of unlabeled probenecid (*A*), BSP (*B*), taurocholate (*TCA*; *C*), ES (*D*), DNP-SG (*E*), DPDPE (*F*), T_3 (*G*), and T_4 (*H*) on the uptake of [^{3}H]$E_2$17βG by Oatp14-transfected HEK293 cells were examined at 37 °C. The specific uptake was obtained by subtracting the uptake by vector-transfected cells from that by gene-transfected cells. *Open* and *closed circles* represent the uptake by Oatp14- and vector-transfected cells, respectively. Each *point* represents the mean ± S.E. (n = 3).

fected cells, and the elimination rate constants were 0.032 ± 0.002 and 0.006 ± 0.001 min^{-1}, respectively (Fig. 4).

cis-inhibitory effects on the Oatp14-mediated uptake of [^{3}H]$E_2$17βG were investigated (Fig. 5). Sulfobromophthalein (BSP), pravastatin, ES, and trichloroacetic acid were potent inhibitors of Oatp14, whereas probenecid was a moderate inhibitor (Fig. 5). *p*-Aminohippurate and cimetidine, typical substrates of organic anion and cation transporters, had no effect on the Oatp14-mediated uptake, whereas benzylpenicillin was a weak inhibitor (Fig. 6). Leukotriene C_4 (*LTC_4*) and glutathione (*GSH*) had no effect, but dinitrophenyl-s-glutathione (*DNP-SG*) was a weak inhibitor (see Figs. 5 and 6). No inhibitory effect by folates (methotrexate, folate, and 5-methyltetrahydrofolate) or tetraethylammonium was observed (Fig. 6). The K_i values of probenecid, BSP, trichloroacetic acid, ES, DNP-SG, DPDPE, T_3, and T_4 for the uptake of [^{3}H]$E_2$17βG by Oatp14-expressed HEK cells are summarized in Table III.

Effects of Hyperthyroid and Hypothyroid Conditions on the Expression of Oatp14 in the Brain Capillary—The effects of hyper- and hypothyroid conditions on the expression of Oatp14 in the brain capillary were investigated by RT-PCR and Western blotting (Fig. 7, *A* and *B*). RT-PCR and Western blotting analyses revealed that the expression levels of Oatp14 mRNA and protein were up- and down-regulated under hypothyroid and hyperthyroid conditions, respectively.

DISCUSSION

In the present study, we reported the substrate specificity of Oatp14, as well as its tissue distribution and localization in the brain. Oatp14 is expressed in the brain capillary and choroid plexus. It mediated the uptake of T_3, T_4, and reverse T_3, as well as organic anions such as $E_2$17βG, cerivastatin, and TRO-S, suggesting its involvement in the membrane transport of these ligands in the brain capillary.

T_3 and its prohormone, T_4, are produced in the thyroid gland and released into the blood. T_3 plays an essential role in brain development via binding to specific nuclear receptors (thyroid hormone receptor) (25). Deficiency of thyroid hormones particularly during fetal and neonatal period in the brain causes mental retardation and cretinism (26, 27). T_3 is supplied to the brain and peripheral tissues as T_4 from which T_3 is enzymatically produced by type 2 iodothyronine deiodinase (D2) (25). Therefore, the brain uptake process of T_4 from the circulating blood is the first step in all subsequent reactions of thyroid hormone in the brain. Whether there is a specific transport mechanism(s) for T_4 in brain capillary endothelial cells remains controversial. The brain uptake of T_4 was saturable in dogs (28) but not in mice (29). Analysis of the transport and molecular properties of Oatp14 should help us resolve this.

Transfection of Oatp14 cDNA into HEK293 cells resulted in a marked increase in the uptake of T_4, as well as reverse T_3, an inactive metabolite of T_4 produced by type 3 iodothyronine deiodinase. Although the uptake of T_3 by Oatp14-expressed cells was significantly greater than that by vector-transfected cells (Table II), T_3 was extensively taken up by vector-transfected cells (Table II). Whether the uptake in vector-transfected cells is ascribed to specific transport system(s) for T_3 or passive diffusion remains unknown. The transport activity for T_3 exhibited by Oatp14, obtained by subtraction of the uptake by vector-transfected cells from that by Oatp14-expressed cells, was ~6-fold lower than that for T_4 and reverse T_3 by Oatp14 (Table II), although the chemical structures of T_3 and reverse T_3 are quite similar. The K_i value of T_3 for Oatp14 was 25-fold greater than that of T_4 (Table III), and this may result in apparent low transport activity for T_3 by Oatp14. Oatp14 can also mediate bidirectional transport, because the efflux of T_4 is facilitated in Oatp14-expressed cells (Fig. 4), and it is possible that Oatp14 is involved in both the uptake and efflux of its ligands through the brain capillary (*i.e.* BBB).

Involvement of Oatp14 in maintaining homeostasis of T_4 in the brain was also supported by the change in expression levels

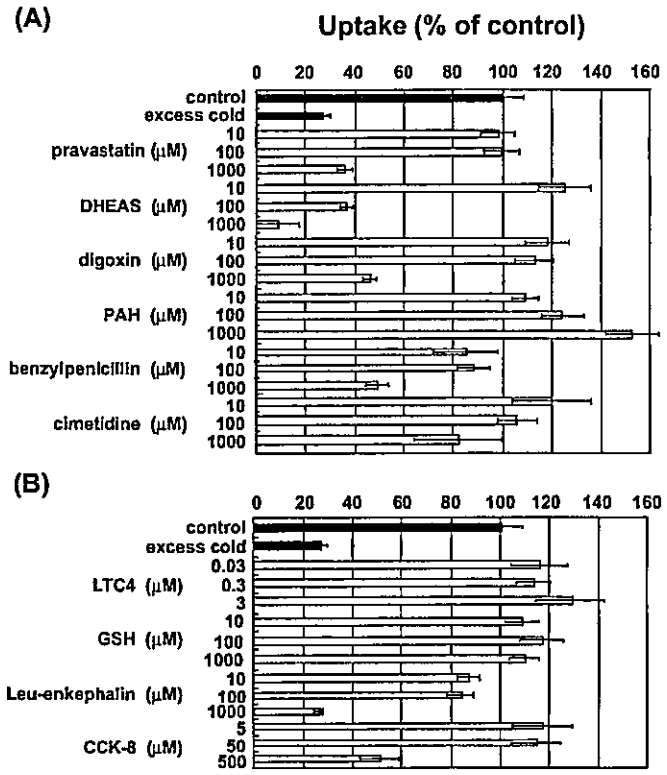

FIG. 6. **Effects of several unlabeled compounds on the uptake of [^{3}H]$E_2$17βG by Oatp14-transfected HEK293 cells.** The effects of several unlabeled compounds on the uptake of [^{3}H]$E_2$17βG by Oatp14-transfected HEK293 cells were examined at 37 °C. Results are given as a ratio with respect to the control values determined in the absence of unlabeled compounds. Each *point* represents the mean ± S.E. (n = 3). *DHEAS*, dehydroepiandrosterone sulfate; *PAH*, *p*-aminohippurate; *GSH*, glutathione; *CCK-8*, cholecystokinin-octapeptide.

TABLE III
K_i values for Oatp14

The K_i values were determined by nonlinear regression analysis using data shown in Fig. 5.

Inhibitor	K_i
	μM
Probenecid	39.5 ± 8.3
BSP	4.18 ± 1.02
Taurocholate	7.24 ± 3.33
ES	6.63 ± 1.62
DNP-SG	467 ± 67
DPDPE	86.3 ± 11.2
T3	2.46 ± 0.96
T4	0.11 ± 0.04

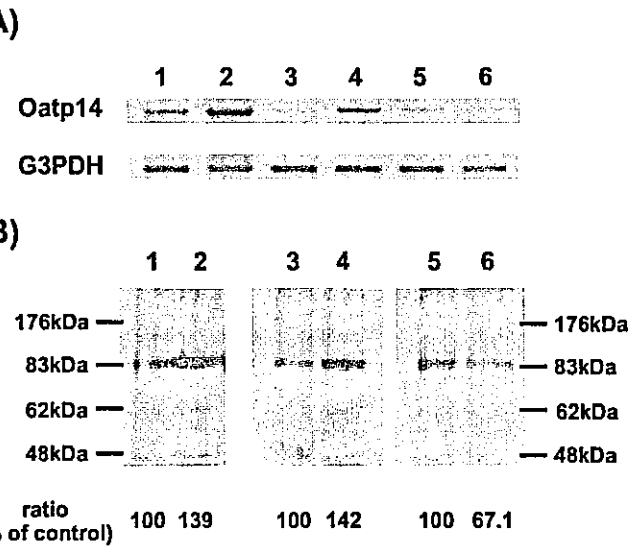

FIG. 7. **Effect of hyperthyroid and hypothyroid conditions on the expression of Oatp14 on the BBB.** *A*, RT-PCR analysis. mRNA samples were prepared from isolated brain capillary from control rats (*lanes 1*, *3*, and *5*), hypothyroid rats (MMI-treated or thyroidectomized rats; *lanes 2* and *4*, respectively) and hyperthyroid rats (T_3-treated rats; *lane 6*). PCR products stained with ethidium bromide were visualized under UV light. *G3PDH*, glucose-3-phosphate dehydrogenase. *B*, Western blotting. Brain capillaries (50 μg/*lane*) isolated from control rats (*lanes 1*, *3*, and *5*), hypothyroid rats (MMI-treated or thyroidectomized rats; *lanes 2* and *4*, respectively), and hyperthyroid rats (T_3-treated rats; *lane 6*) were separated by SDS-PAGE (10% separating gel). Oatp14 was detected by anti-Oatp14 polyclonal antibody.

of Oatp14 in the brain capillary under hypo- and hyperthyroid conditions (Fig. 7). The expression of Oatp14 in the brain capillary changed as if Oatp14 was responsible for maintaining the concentration of T_4 in the central nervous system: up- and down-regulated under hypothyroid and hyperthyroid conditions, respectively (Fig. 7). This pattern is similar to that observed in D2 expression (25). Increased D2 expression increases the conversion of T_4 to T_3 to compensate for the decrease in the local brain concentration of T_4 and vice versa. Therefore, we hypothesize that Oatp14 is involved in the uptake of T_4 through the brain capillaries.

In addition to Oatp14, Oatp2, the other isoform of rat Oatp family, is also the candidate transporter for T_3 and T_4 uptake by the brain from the circulating blood in rodents. The uptake of both T_3 and T_4 was significantly increased in Oatp2-cRNA injected oocytes with similar K_m values (~5–7 μM) (12). Oatp2 has been identified both in the luminal and abluminal membrane of brain capillary endothelial cells (15). It is possible that Oatp14 and Oatp2 serve high and low affinity sites for T_4 in the brain capillary, because the K_m values of Oatp2 were ~30-fold greater than that of Oatp14 (12). Following uptake from the circulating blood into endothelial cells, T_4 has to cross the abluminal membrane to reach the brain interstitial space and brain parenchymal cells. Whether this process is carrier-mediated remains unknown. Bidirectional nature of Oatp2-mediated transport has been reported in Oatp2-cRNA-injected oocytes (30). Oatp14 and Oatp2 are candidate transporters involved in the abluminal secretion of thyroid hormones. Further studies are necessary to identify the exact localization of Oatp14 in the brain capillary and to evaluate its contribution to the total brain uptake of T_4 into the brain. Pardridge *et al.* (31) demonstrated that the brain uptake of T_3 was saturable and inhibited by T_4 using carotid arterial bolus injection technique of Oldendorf, and Oatp2 may account for the brain uptake of T_3 in the brain capillary.

Oatp14 was detected in the choroid plexus by Western blot analysis (Fig. 1). The choroid plexus is located in the lateral, third and fourth ventricles, and the interface between the cerebrospinal fluid and the circulating blood acting as a barrier to protect the central nervous system, in conjunction with the BBB (32, 33). The brain distribution of T_3 and T_4 after intracerebroventricular administration is limited to ependymal cells and circumventricular organs and, thus, transport via the choroid plexus could account for the brain distribution near the ventricles (34). As speculated in the case of brain capillary endothelial cells, it is possible that Oatp14 acts as an uptake system to supply T_4 to ependymal cells and circumventricular organs in the choroid plexus.

In addition to thyroids, Oatp14 accepts certain types of amphipathic organic anions, such as $E_2$17βG, cerivastatin, and TRO-S, as substrates although their transport activity was markedly lower than that of T_4, except TRO-S (see Fig. 3 and Table II). Because Oatp14 can mediate the bidirectional trans-

port, it is possible that Oatp14 is involved in the efflux of organic anions such as $E_2 17\beta G$ from the brain when it is microinjected into the cerebral cortex and possibly in the efflux of excess T_4 and reverse T_3 from the brain. The spectrum of inhibitors of Oatp14 was consistent with the transporter hypothesized based on *in vivo* studies (18), but further investigations will be required to confirm this speculation.

Whether the results obtained using cDNA from rodents can be applied to the human situation is an important issue. Human OATP-F, an isoform in which Oatp14 exhibits high homology (84% in amino acid level), has a similar substrate specificity to Oatp14 (35). Northern blot analysis demonstrated abundant expression of OATP-F in the brain and testis and, to a lesser extent, heart, but the localization in the brain remains unidentified. In terms of substrate specificity and homology, OATP-F is supposed to be the human ortholog of Oatp14, and it may be suggested that OATP-F is also involved in the uptake of T_4 from the circulating blood into the central nervous system though the brain capillary and choroid plexus. In view of the importance of supplying T_4 to the brain during development, it is possible that functional loss of the OATP-F gene may be associated with a thyroid hormone-related neuronal disorder characterized by resistance to thyroid hormone treatment.

In conclusion, we have characterized Oatp14 in terms of its substrate specificity and localization in the brain and demonstrated that Oatp14 accepts T_4, as well as organic anions, including certain glucuronide and sulfate conjugates. Oatp14 is localized on the plasma membrane of brain capillary endothelial cells and involved in the uptake of T_4 from the blood to the central nervous system. Oatp14 is one of the mechanisms for maintaining homeostasis of T_4 and, ultimately, T_3 in the brain.

REFERENCES

1. Rapoport, S. I. (1976) *Exp. Neurol.* **52**, 467–479
2. Pardridge, W. M. (1991) *Semin. Cell Biol.* **2**, 419–426
3. Minn, A., Ghersi-Egea, J. F., Perrin, R., Leininger, B., and Siest, G. (1991) *Brain Res. Brain Res. Rev.* **16**, 65–82
4. Strazielle, N., and Ghersi-Egea, J. F. (1999) *J. Neurosci.* **15**, 6275–6289
5. Suzuki, H., Terasaki, T., and Sugiyama, Y. (1997) *Adv. Drug. Deliv. Rev.* **25**, 257–285
6. Kusuhara, H., and Sugiyama, Y. (2001) *Drug Discov. Today* **6**, 150–156
7. Lee, G., Dallas, M., Hong, M., and Bendayan, R. (2001) *Pharmacol. Rev.* **52**, 569–596
8. Kullak-Ublick, G. A., Stieger, B., Hagenbuch, B., and Meier, P. J. (2000) *Semin. Liver Dis.* **20**, 273–292
9. Kullak-Ublick, G. A., Ismair, M. G., Stieger, B., Landmann, L., Huber, R., Pizzagalli, F., Fattinger, K., Meier, P. J., and Hagenbuch. B. (2001) *Gastroenterology* **120**, 525–533
10. Jacquemin, E., Hagenbuch, B., Stieger, B., Wolkoff, A. W., and Meier, P. J. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 133–137
11. Noe, B., Hagenbuch, B., Stieger, B., and Meier, P. J. (1997) *Proc. Natl. Acad. Sci. U. S. A.* **94**, 10346–11035
12. Abe, T., Kakyo, M., Sakagami, H., Tokui, T., Nishio, T., Tanemoto, M., Nomura, H., Hebert, S. C., Matsuno, S., Kondo, H., and Yawo, H. (1998) *J. Biol. Chem.* **273**, 22395–22401
13. Nishio, T., Adachi, H., Nakagomi, R., Tokui, T., Sato, E., Tanamoto, M., Fujiwara, K., Okabe, M., Onogawa, T., Suzuki, T., Nakai, D., Shiiba, K., Suzuki, M., Ohtani, H., Kondo, Y., Unno, M., Ito, S., Iinuma, K., Nunoki, K., Matsuno, S., and Abe, T. (2000) *Biochem. Biophys. Res. Commun.* **275**, 831–838
14. Kullak-Ublick, G. A., Hagenbuch, B., Stieger, B., Schteingart, C. D., Hofmann, A. F., Wolkoff, A. W., and Meier, P. J. (1995) *Gastroenterology* **109**, 1274–1282
15. Gao, B., Stieger, B., Noe, B., Fritschy, J. M., and Meier, P. J. (1999) *J. Histochem. Cytochem.* **47**, 1255–1264
16. Gao, B., Hagenbuch, B., Kullak-Ublick, G. A., Benke, D., Aguzzi, A., and Meier, P. J. (2000) *J. Pharmacol. Exp. Ther.* **294**, 73–79
17. Dagenais, C., Ducharme, J., and Pollack, G. M. (2001) *Neurosci. Lett.* **301**, 155–158
18. Sugiyama, D., Kusuhara, H., Shitara, Y., Abe, T., Meier, P. J., Sekine, T., Endou, H., Suzuki, H., and Sugiyama, Y. (2001) *J. Pharmacol. Exp. Ther.* **298**, 316–322
19. Li, J. Y., Boado, R., and Pardridge, W. M. (2001) *J. Cereb. Blood Flow Metab.* **21**, 61–68
20. Izumi, T., Hosiyama, K., Enomoto, S., Sasahara, K., and Sugiyama, Y. (1997) *J. Pharmacol. Exp. Ther.* **280**, 1392–40021
21. Akita, H., Suzuki, H., Ito, K., Kinoshita, S., Sato, N., Takikawa, H., and Sugiyama, Y. (2001) *Biochim. Biophys. Acta* **1511**, 7–16
22. Boado, R. J., and Pardridge, W. M. (1991) *J. Neurochem.* **57**, 2136–2139
23. Lowry, O. (1951) *J. Biol. Chem.* **193**, 265–273
24. Burmeister, L. A., Pachucki, J., and Germain, D. L. S. (1997) *Endocrinology* **138**, 5231–5237
25. Forrest, D., Reh, T. A., and Rusch, A. (2002) *Curr. Opin. Neurol.* **12**, 49–56
26. Oppenheimer, J. H., and Schwartz, H. L. (1997) *Endocr. Rev.* **18**, 462–475
27. Shi, Y. B., Ritchie, J. W. A., and Taylor, P. M. (2002) *Pharmacol. Ther.* **94**, 235–251
28. Hagen, G. A., and Solberg, L. A., Jr. (1974) *Endocrinology* **95**, 1398–1410
29. Banks, W. A., Kastin, A. J., and Micahls, E. A. (1985) *Life Sci.* **37**, 2407–2414
30. Li, L., Meier, P. J., and Ballatori, N. (2000) *Mol. Pharmacol.* **58**, 335–340
31. Pardridge, W. M. (1979) *Endocrinology* **105**, 605–612
32. Groothuis, D. R., and Levy, R. M. (1997) *J. Neurovirol.* **3**, 387–400
33. Segal, M. B. (2000) *Cell. Mol. Neurobiol.* **20**, 183–196
34. Dratman, M. B., Chrutchfield, F. L., and Schoenhoff, M. B. (1991) *Brain Res.* **554**, 229–236
35. Pizzagalli, F., Hagenbuch, B., Stieger, B., Klenk, U., Folkers, G., and Meier, P. J. (2002) *Mol. Endocrinol.* **16**, 2283–2296

J Nutr Sci Vitaminol, 49, 125–132, 2003

Analysis of the Comprehensive Effects of Polyunsaturated Fatty Acid on mRNA Expression Using a Gene Chip

Yoko FUJIWARA[1,2], Masayo YOKOYAMA[1], Rumi SAWADA[1], Yousuke SEYAMA[1], Masami ISHII[3], Shunichi TSUTSUMI[3], Hiroyuki ABURATANI[3], Satoko HANAKA[2,4], Hiroshige ITAKURA[2,5] and Akiyo MATSUMOTO[2,6,*]

[1] *Department of Food Science and Nutrition, Ochanomizu University, Tokyo 112–8610, Japan*
[2] *Division of Clinical Nutrition, National Institute of Nutrition and Health, Tokyo 162–8636, Japan*
[3] *Genome Sciences, Research Center for Advanced Science and Technology, The University of Tokyo, Tokyo, Japan*
[4] *Medical School of Teikyo University, Tokyo 153–8904, Japan*
[5] *Department of Life Science, Ibaraki Christian University, Ibaraki, Japan*
[6] *Department of Clinical Dietetics and Human Nutrition, Josai University, Saitama 350–0295, Japan*

(Received October 3, 2002)

Summary To investigate the comprehensive effects of polyunsaturated fatty acids (PUFA) on gene expression, we analyzed changes of mRNA expression in PUFA-treated HepG2 cells using a DNA micro array. We incubated HepG2 cells for 24 h with or without 0.25 mM oleic acid (OA), arachidonic acid (AA), eicosapentaenoic acid (EPA) or docosahexaenoic acid (DHA), and then compared the expression profiles of thousands of genes using a GeneChip. PUFA influenced the expression of various genes related to cell proliferation, growth and adhesion, as well as for many transcription factors including sterol regulatory element binding proteins (SREBP). Treatments with AA, EPA, and DHA repressed the expression of genes related to cholesterol and lipid metabolism. Moreover, data from gene chip analysis proved that PUFA reduced the expression of prostasin, which is a serine protease. By measuring the mRNA levels of SREBPs, mevalonate pyrophosphatase and prostasin using quantitative RT-PCR, we confirmed the effect of PUFA revealed by gene chip analysis. These data might provide useful clues with which to explore novel functions of PUFA.

Key Words DNA micro array, polyunsaturated fatty acid, HepG2 cells

Fatty acids (FA) are important not only as an energy source but also as components of the cell membrane because fatty acid composition influences membrane fluidity and the function of receptors or channels (*1*). Polyunsaturated fatty acids (PUFA) decrease plasma triacylglycerols and cholesterol levels and influence lipid metabolism (*2*). Many studies within the past decade have shown that PUFA function as mediators of gene transcription. One of the mechanisms is explained by the down-regulation of the sterol regulatory element binding protein (SREBP), which regulates the intracellular cholesterol metabolism (*3*, *4*). On the other hand, peroxisome proliferator activated receptor (PPAR), which binds FA and its metabolites, regulates the gene expression concerned with the β-oxidation of fatty acids (*5*). Several other factors such as liver X receptor (LXR) (*6*), hepatocyte nuclear factor 4 (HNF4) (*7*), *c-fos* and *nur-77* (*8*) are also thought to regulate the expression of genes that respond to FA.

The mechanism of how PUFA controls gene expression was summarized as follows (*9*). Fatty acids or their derivatives function as ligands for a transcription factor (TF), which then binds DNA at the FA response element and activates or represses transcription. Fatty acids or their derivatives thereby modify transcriptional potency and initiate a signal transduction cascade to induce covalent modification of a TF. Fatty acids act indirectly via

* To whom correspondence should be addressed.
E-mail: amatsu@josai.ac.jp

Abbreviations: AA, arachidonic acid; ACAT, acyl CoA: cholesterol acyltransferase; BSA, bovine serum albumin; CETP, cholesteryl ester transfer protein; DHA, docosahexaenoic acid; DMEM, Dulbecco's modified Eagle's medium; EPA, eicosapentaenoic acid; FA, fatty acid; HNF4, hepatocyte nuclear factor 4; HTGL, hepatic triglyceride lipase; LCAT, lecithin-cholesterol acyltransferase; LPDS, lipoprotein-deficient serum; LXR, liver X receptor; MPD, mevalonate pyrophosphate decarboxylase; OA, oleic acid; PPAR, peroxisome proliferator activated receptor; PUFA, polyunsaturated fatty acids; ROS, reactive oxygen species; RT-PCR, reverse transcriptase-polymerase chain reaction; SA, stearic acid; SDS, sodium dodecyl sulfate; SRE, sterol regulatory element; SREBP, sterol regulatory element-binding protein; SSPE, buffer consisting of saline, sodium dihydrogen phosphate-ethylenediamine tetraacetate.

alterations in either TF mRNA stability or gene transcription, resulting in variations of de novo TF synthesis with impact on the transcription rate of genes encoding proteins related to FA transport and metabolism (9).

Despite much recent progress, the mechanism(s) by which FA modulate gene transcription remains largely unknown. DNA micro array is known to be a powerful tool to investigate mRNA expression profiles, and has been widely applied in the field of drug innovation to analyze functions and side effects of new compounds. In this study, to investigate the comprehensive effects of PUFA on gene regulation, we analyzed mRNA expression profiles in PUFA-treated HepG2 cells using a DNA micro array.

MATERIALS AND METHODS

Cell culture with fatty acids. The human hepatoma cell line, HepG2, purchased from the Riken Gene Bank (Tsukuba, Japan) was cultured in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal calf serum (Intergen Co., Purchase, NY, USA) at 37°C under a 5% CO_2 atmosphere. Cells seeded in 60-mm diameter collagen-coated culture dishes (Sumitomo Bakelite Co., Ltd., Tokyo, Japan) at a density of 1×10^6/dish were cultured to approximately 90% confluence and then incubated with 0.25 mM oleic acid (18 : 1, OA), arachidonic acid (20 : 4, AA), eicosapentaenoic acid (20 : 5, EPA) or docosahexaenoic acid (22 : 6, DHA) in DMEM supplemented with 10% fetal calf lipoprotein-deficient serum (LPDS), which eliminates FA in the serum (all from Sigma-Aldrich Chemical Co., St Louis, MO, USA). The FAs were dissolved in essential FA-free bovine serum albumin (BSA). Control HepG2 cells were incubated with the same 10% LPDS/DMEM without FA. After a 24-h incubation at 37°C, the cells were harvested. Total RNA was isolated from the cells using a RNeasy Mini Kit (Qiagen Inc., Valencia, CA, USA) according to the manufacturer's instructions. Cells were pooled from triplicate dishes and then poly(A)$^+$ RNA was purified using a QuickPrep micro mRNA purification kit (Amersham Pharmacia Biotech Inc., Piscataway, NJ, USA).

Micro array analysis of expression profiles. Studies using the GeneChip proceeded according to the technical manual supplied with the Affymetrix GeneChip Expression Analysis System (*10, 11*). First-strand cDNA was generated with 1 μg of poly(A)$^+$ RNA and 0.1 μmol of T7-linked oligo(dT)$_{24}$ primer (Amersham Pharmacia Biotech Inc.) using SuperScript Choice System (Life Technologies, Rockville, MD, USA). After second-strand synthesis, in vitro transcription was performed using biotinylated UTP and CTP (Enzo Diagnostics, Inc., Farmingdale, NY, USA). The amplified cRNA was purified by passage through a column containing affinity resin (RNeasy Mini Kit, Qiagen Inc.), and quantified by absorbance at 260 nm. Biotinylated cRNA (25 μg) was fragmented into 50- to 150-nt units before overnight hybridization to GeneChips (HuGene FL Array, Affymetrix Inc., Santa Clara, CA, USA) that contain oligonucleotide probe sets for approximately 6,000 human genes. Fragmented cRNA and sonicated herring sperm DNA, up to 0.1 mg/mL, were added to hybridization buffer containing 1 M NaCl, 10 mM Tris-HCl (pH 7.6), and Triton X-100 (ST-T). The mixture was denatured at 99°C for 5 min, incubated at 45°C for 5 min, and then a hybridization cocktail was injected into the probe array cartridge. Hybridization proceeded at 45°C for 16 h with rotary shaking at 60 rpm. Thereafter, the hybridization solution was removed from the array, which was filled with non-stringent washing buffer (5×SSPE, 0.01% Tween-20, 0.05% antifoam). The hybridized array was stained with 5.0 μg/mL streptavidin/phycoerythrin (Molecular Probes, Eugene, OR, USA) and 2.0 mg/mL acetylated BSA (Sigma, St. Louis, MO, USA) in 1×ST-T at 40°C for 15 min. The probe array was scanned twice at 3 μm of resolution using an HP GeneArray Scanner (Affymetrix Inc.).

The intensity of each feature of the array was captured by Affymetrix GeneChip Software (Affymetrix Inc.) according to standard Affymetrix procedures (*10, 11*) with a class AB mask file. This file is designed to exclude inappropriate probe pairs that represent introns or reverse sequences. A single expression level for each gene was derived from 20 probe pairs representing each gene, that is, ~20 perfectly matched (PM) and mismatched (MM) control probes. The MM probes act as specific controls that allow the direct subtraction of background and cross-hybridization signals. The average difference (Avg. diff.) representing PM−MM for each gene-specific probe set shows the quantitative mRNA levels. The fold change of the transcripts between the control and the experimental sample was calculated as follows:

$$\text{Fold change} = \{\text{Avg. diff. change}/\max[\min(\text{Avg. diff.}_{\text{exp.}}, \text{Avg. diff.}_{\text{control}}), Q_M * Q_C]\} + (+1 \text{ if Avg. diff.}_{\text{exp.}} \geq \text{Avg. diff.}_{\text{control}} \; -1 \text{ if Avg. diff.}_{\text{exp.}} \leq \text{Avg. diff.}_{\text{control}})$$

Where

$$\text{Avg. diff. change} = \text{Diff.}_{\text{exp.}} - \text{Avg. diff.}_{\text{control}}$$

$$Q_C = \max(Q_{\text{exp.}}, Q_{\text{control}})$$

And the Q multiplier

Q_M=2.1 in 50 μm feature arrays and 2.8 in 24 μm feature arrays.

This equation permits the expression of fold change as a positive number when the transcript increased over its control state, and as a negative number when the transcript level decreased. If the noise (Q Q_M) of either array was greater than the Avg. diff. of the transcript (in either the control or experimental data), the fold change was calculated using the noise.

Quantitative RT-PCR. Single-strand cDNA was synthesized from 1 μg of total RNA using random hexamer and TaqMan Reverse Transcription Reagents (Applied Biosystems, Foster City, CA, USA). Primers for PCR were designed using Primer Express (Applied Biosystems) software. Primer sequences were 5′-GCAAGGCCATCG-

ACTACATTC-3′ (forward) and 5′-TTGCTTTTGTGGAC-AGCAGTG-3′ (backward) for SREBP-1; 5′-AGGCGGA-CAACCCATAATATCA-3′ (forward) and 5′-GACTTGTG-CATCTTGGCGTCT-3′ (backward) for SREBP-2; 5′-CGA-GTCACACTGGCCTGAACT-3′ (forward) and 5′-CACGG-TACTGCCTGTCAGCTT-3′ (backward) for MPD; 5′-CTT-GATCTTTGAGCCCATTCTTC-3′ (forward) and 5′-TCT-GGCCATTGCTACAGGC-3′ (backward) for prostasin; 5′-TTCAGAAAACACAGATGACCTACTACTTC-3′ (forward) and 5′-CTGATCTTTCGCTTTGATGTTTTAGAC-3′ (backward) for hepatic triglyceride lipase (HTGL); 5′-AAATT-CCATGGCACCGTCA-3′ (forward) and 5′-AGCATCGCC-CCACTTGATT-3′ (backward) for GAPDH. Real-time quantitative RT-PCR proceeded in a reaction mixture containing 10 ng of first-strand cDNA, 300 nM of each primer set in a final volume of 25 μl and SYBR Green PCR core reagent (Applied Biosystems). The results were analyzed using a GeneAmp 5700 sequence detection system (Applied Biosystems). All values were expressed as mean±SD. Significance of the difference between PUFA treatments ($p<0.05$) was determined by analysis of variance (ANOVA) using a Stat View software (Abucas Concepts, Inc., Berkeley, CA, USA).

RESULTS

Gene chip analysis

Table 1 shows the effect of FA on genes except for those related to cholesterol and lipoprotein metabolism. PUFA suppressed the mRNA levels of SREBPs, as well as of those expressing transcription factors such as NF-κB p65, nuclear factor I-X (NFI-X), PPARs and Rad2. PUFA repressed the expression of lipogenic genes such as fatty acid synthase and stearoyl-CoA desaturase. Although gene expression related to FA oxidation was not changed, PUFA up-regulated 2-oxoglutarate dehydrogenase, isocitrate dehydrogenase and succinyl-CoA synthetase, all of which are involved in the TCA cycle. However, the expression of LXR, HNF, *c-fos*, and *nur*-77, which were thought to responde to FA, did not change. All FA largely increased mRNA levels of metallothionein-IG, ventricular/slow twitch myosin alkali light chain (MLC-1V/Sb isoform), and deleted split hand/split foot 1 (DDS1). In addition, PUFA affected the expression of genes involved in cell differentiation and proliferation.

Table 2 shows changes in genes related to cholesterol and lipoprotein metabolism by FA treatments. PUFA suppressed the mRNA levels of LDL receptor, HMG CoA synthase and HMG CoA reductase, all of which are SREBP targets (*12–14*). Furthermore, the expression of mevalonate pyrophosphate decarboxylase (MPD) and squalene epoxidase, which function in the cholesterol synthetic pathway were down regulated. Lysosomal acid lipase was one of the PUFA responsive genes that had not been reported. Results of gene chip showed no change of the expression of LXR-alpha and undetectable levels of the mRNA expression of acyl CoA: cholesterol acyltransferase (ACAT), and cholesteryl ester transfer protein (CETP), which are thought to be regulated by PUFA (*15, 16*).

Quantitative RT-PCR

To confirm the effects of PUFA observed by DNA micro array analysis, we measured the mRNA levels of SREBPs and MDP in PUFA-treated HepG2 cells using real-time RT-PCR. AA, EPA, and DHA reduced the mRNA levels of SREBP-1 and -2, whereas OA did not affect the mRNA levels (Fig. 1). Generally, *n*-6 and *n*-3 PUFA similarly affect the expression level of SREBP. Figure 2 shows that PUFA also decreased the mRNA levels of MPD by 50%. We also confirmed the effect of PUFA on the mRNA levels of prostasin (Fig. 3). PUFA repressed the expression of prostasin, but changes in the mRNA levels were smaller than those shown by gene chip analysis. Figure 4 shows up-regulation by PUFA on the mRNA expression of HTGL.

DISCUSSION

The present study investigated the extensive effects of PUFA on gene expression in HepG2 cells and explored thenovel functions of PUFA. Few genes were induced or reduced by PUFA more than 8-fold among the approximately 6,000 gene probes on the oligonucleotide chip. The effects of PUFA on gene expression were moderate compared with those elicited by other drugs or chemicals (*17, 18*). This may be because PUFA are nutrients that are naturally catabolized to produce energy. As the MPD mRNA level was decreased by −9.5 (11%) with DHA by gene chip analysis, we measured the mRNA levels of MPD using a real time RT-PCR of HepG2 cells incubated with PUFA. The gene chip analysis showed reduced mRNA levels by −1.7, −1.2, −1.8 in SREBP-1, and −1.5, −1.9, −1.7 in SREBP-2 with AA, EPA, and DHA, respectively. We also analyzed mRNA levels of SREBPs in FA-treated HepG2 cells using quantitative RT-PCR (Fig. 1). We confirmed that PUFA suppressed MPD, and SREBPs but the magnitude of the effect was not equivalent between the data obtained from the gene chip analysis and that from RT-PCR. However, the gene chip analysis reflects the tendencies of PUFA, and is a useful tool for investigating the comprehensive effect. In the tables, therefore, we listed the data obtained using the gene chip with the absolute value of the fold change >2 and the similar tendencies among PUFA treatments.

The results of gene chip analysis showed the changes of SREBP and PPAR resulting from PUFA, but the analysis did not detect changes in HNF, *c-fos* and *nur*-77, which have been previously reported (*8*). More detailed measurements of each gene are required to accurately quantify changes in the mRNA levels. It may be related to the expression levels of some genes in HepG2 cells. Although HepG2 cells are familiar for investigating the lipid metabolism, the expression of PPAR was low whereas the expressions of apolipoproteins AI, AII, and E were high (data not shown).

Our data showed that PUFA repressed the expression of hepatic lipogenic genes and almost all of the genes associated with the cholesterol synthetic pathway. In this study using HepG2 cells, one of the major mechanisms of mRNA expression down-regulation caused by PUFA

Table 1. Fold changes of mRNA levels by fatty acid treatment in HepG2 cells.

Gene	Accession	Average difference[a]	Fold change				Function
			OA	AA	EPA	DHA	
Interferon-gamma receptor alpha chain	U19247	226	−1.1	−2.3	−2.3	−2.2	antiviral activity
Mitochondrial NADH dehydrogenase	U65579	407	1.8	2.0	3.1	2.7	aspiratory chain
Heparan sulfate proteoglycan (HSPG2)	M85289	146	−1.5	1.3	5.3	1.3	cell adhesion
cdc25Hs	M34065	−26	2.7*	1.9*	1.4*	2.1*	cell differentiation
Interleukin 1 alpha (IL 1)	M28983	−51	2.6*	2.0*	2.7*	1.8*	cell differentiation
MAC30	L19183	1,769	1.0	−2.8	−2.1	−1.8	cell differentiation
Protein tyrosine phosphatase (PTP-PEST)	M93425	73	2.0*	1.9*	−1.0*	3.1*	cell differentiation
Small proline-rich protein 2 (SPRR2B)	L05188	−110	2.9*	2.8*	3.0*	1.8*	cell differentiation
SWI/SNF complex 155 kDa subunit (BAF155)	U66615	197	1.4	1.5*	2.3	2.2*	cell differentiation
Drosophila female sterile homeotic (FSH)	X62083	4	1.5	1.2	13.3*	2.9*	cell proliferation
Glial growth factor 2		394	−5.5*	−3.0	−5.3*	−5.6*	cell proliferation
Membrane-associated protein (HEM-1)	M58285	193	1.8	2.5*	2.9	2.4*	cell proliferation
Sec23A isoform	X97064	51	3.3*	1.1*	2.1*	2.6*	cell proliferation
Sec23B isoform	X97065	230	2.3	1.8	2.4	2.1	cell proliferation
S-lac lectin L-14-II (LGALS2)	M87860	−3	1.5*	2.0*	3.2*	4.5*	cell proliferation
Microsomal glutathione *S*-transferase (GST-II)	U77604	2,836	1.0	−1.1	1.2	−2.0	detoxification
FDXR gene (adrenodoxin reductase)	M58509	287	1.2	1.5	1.6	2.2	electron transport system
Uncoupling protein homolog (UCPH)	U94592	169	2.8	−2.7	2.2	−2.0	energy consumption
Fatty acid synthase	S80437	4,358	−1.0	−2.1	−2.1	−2.3	fatty acid synthesis
Stearoyl-CoA desaturase		1,416	1.1	−2.9	−2.9	−3.1	fatty acid synthesis
Liver fatty acid binding protein (FABP)	M10050	6,859	1.1	−2.0	−1.6	−1.5	fatty acid transport
Ceruloplasmin (ferroxidase)	M13699	309	1.2	−1.9	−2.9	−3.1	Fe oxidation
Galactokinase (GALK1)	L76927	120	2.5	2.8	−1.9*	3.1	glycogenesis/glycolysis
RASF-A PLA2	M22430	122	2.0	2.2	1.5	2.7	inflamation
S-lac lectin L-14-II (LGALS2)	M87860	−3	1.5*	2.0*	3.2*	4.5*	lectin
Deleted in split hand/split foot 1 (DSS1)	U41515	102	1.4	3.7	4.4	4.8	limb development
Urokinase-type plasminogen activator receptor	U09937	−19	1.6*	1.*	4.9*	2.4*	platelet coagulation
Metallothionein-IG (MT1G)	J03910	195	1.8	3.6	2.3	5.6	protection against heavy metal toxicity
Inter-alpha-trypsin inhibitor subunit 3	X16260	238	−1.1	−4.8*	−4.2*	−2.8	proteinase inhibitor
Vacuolar proton pump, 116-kDa subunit	U45285	41	2.0*	4.4*	4.3*	7.1*	proton pump
Prostasin	L41351	749	−1.2	−2.6	−8.3	−3.8	serine proteinase
Extracellular-superoxide dismutase (SOD3)	J02947	124	1.7	1.3	3.6*	2.2	superoxiside scavenger
Manganese superoxide dismutase (SOD2)	X65965	611	−1.2	−1.2	−2.0	−1.1	superoxiside scavenger
2-Oxoglutarate dehydrogenase	D10523	143	1.4	−1.0	2.0	1.4	TCA cycle
Isocitrate dehydrogenase	Z68129	202	1.5	1.9	2.5	2.0	TCA cycle
Succinate dehydrogenase (SDH)	L21936	496	1.9	1.4	2.1	2.5	TCA cycle
Succinyl-CoA synthetase	Z68204	6	2.1*	1.2*	2.1*	2.1*	TCA cycle
LXR-alpha	U22662	67	1.4*	−1.3*	2.1*	1.3*	transcription factor
NF-kappa-B p65 subunit	L19067	201	2.3	1.7	1.4	1.8	transcription factor
Nuclear factor I-X	L31881	94	1.4	−1.2	4.4	1.4	transcription factor
PPAR alpha	L02932	4	−1.4*	1.2*	1.5*	1.1*	transcription factor
PPAR gamma	L40904	99	2.0	−1.6	−1.1	1.0	transcription factor
Rad2		40	2.6	2.1*	3.5*	2.6*	transcription factor
SREBP-1	U00968	1,105	1.0	−1.7	−1.2	−1.8	transcription factor
SREBP-2	U02031	559	1.1	−1.5	−1.9	−1.7	transcription factor
KIAA0030	D21063	45	2.5*	2.2*	8.1*	1.9*	unknown
KIAA0092	D42054	325	−1.1	−1.3	−1.7	−3.8	unknown
KIAA0219	D86973	96	4.2	1.9	2.7	3.2	unknown
Inducible protein	L47738	−55	3.9*	3.5*	4.5*	3.5*	unknown

HepG2 cells were treated with 0.25 mM of oleic acid (OA), arachidonic acid (AA), eicosapentaenoic acid (EPA), or docosahexaenoic acid (DHA) for 24 h.

[a] Average differences were expressed the intensities of the mRNA levels in control HepG2 cells.

* The value of fold change was calculated using the noise, since the noise of either array was greater than the average difference of the transcript in both the control and the FA-treated groups.

Table 2. Changes of mRNA levels in genes related to cholesterol and lipoprotein metabolism by FA-treatment.

Gene	Accession	Average difference[a]	Fold change			
			OA	AA	EPA	DHA
Repressed						
HMG-CoA reductase	M11058	614	−1.5	−2.9	−2.2	−3.1
HMG-CoA synthase	L25798	226	−1.5	−2.9	−2.4*	−2.0
Mevalonate kinase	M88468	276	−1.2	−1.2	−2.7	−1.1
Mevalonate pyrophosphate decarboxylase	U49260	1,638	−1.3	−3.4	−1.9	−9.5
Squalene epoxidase	D78129	1,782	−1.0	−2.0	−1.2	−2.2
2,3-Oxidosqualene-lanosterol cyclase	U22526	200	−1.0	−2.5	−2.8*	−4.8*
LDL receptor	L00352	1,358	−1.1	−2.6	−2.1	−2.3
Lysosomal acid lipase	U04285	957	−1.1	−1.6	−2.2	−1.6
Induced						
Hepatic triglyceride lipase	M29194	−1	1.7*	1.2*	1.8*	2.6*
Apolipoprotein(a)	X06290	89	2.2	1.3	2.4	−1.3 *
ICAM-2	M32334	5	2.0*	1.3*	3.1*	−1.3*
No change						
Apolipoprotein AI regulatory protein (ARP-1)	M64497	40	1.2*	1.1*	−1.7*	1.1*
Ear-3		75	1.0	1.1*	−1.1*	−1.5*
Lectin-like oxidized LDL receptor	D89050	−21	1.1*	−2.0*	−1.1*	1.1*
Lipoprotein lipase	M15856	52	−1.4*	−1.0*	−1.3*	−1.3*
Scavenger receptor type I	D13264	−13	−1.2*	−1.3*	1.0*	1.2*
CLA-1 (SR-BI)	Z22555	0	0.0*	0.0*	0.0*	0.0*
CD36	Z32765	731	1.1	−1.3	1.3	1.5
HDL binding protein	M64098	942	1.2	1.2	1.2	1.4
CD6 ligand (ALCAM/HB2)	L38608	87	1.0	−1.8*	−1.5*	1.2
Cdc42 GTPase-activating protein	U02570	310	1.3	1.2	1.2	1.4
LCAT	M12625	741	1.0	1.1	−1.2	1.2
ACAT	L21934	−12	1.4*	1.1*	1.2*	1.2*
CETP	M30185	−140	−2.9 *	−1.2 *	1.3*	−1.9 *
Phospholipid transfer protein		245	1.4	−1.2	1.4	−1.0
MTP	X91148	0	0.0*	0.0*	0.0*	0.0*

HepG2 cells were treated with 0.25 mM of oleic acid (OA), arachidonic acid (AA), eicosapentaenoic acid (EPA), or docosahexaenoic acid (DHA) for 24 h.

[a] Average differences were expressed the intensities of the mRNA levels in control HepG2 cells.

* The value of fold change was calculated using the noise, since the noise of either array was greater than the average difference of the transcript in both the control or the FA-treated groups.

may be mediated through SREBPs. PUFA reduce the mRNA expression of SREBPs (*4*, *19–22*), which regulate lipogenic gene transcription (SREBP-1) and control cholesterol metabolism (SREBP-2) (*12–14*). Sakakura et al. reported that SREBP regulates the gene expression of all of the enzymes involved in cholesterol synthesis including MPD (*23*). These results indicate that PUFA down-regulates the entire cholesterol synthetic pathway.

Yoshikawa et al. reported that the PUFA suppression of SREBP-1c expression is mediated through competition with liver X factor receptor (LXR) ligand during activation of the ligand-binding domain of LXR (*24*). On the other hand, Tobin et al. (*6*) reported that fatty acids induced the LXR alpha expressions that regulate the fatty acid and cholesterol metabolism. LXR alpha was not changed in our gene chip data. Although we need to analyze LXR further, Cyp7A1 was up-regulated by PUFA using RT-PCR (data not shown).

The PUFA response region is located in the promoter of the stearoyl-CoA desaturase 1 (SCD1) gene (*25*). SREBP may play an important role to regulate the SCD1 because its rate of down-regulation was similar to those of the genes related to cholesterol metabolism. However, Kim et al. (*26*) recently demonstrated that cholesterol overrides the PUFA-mediated repression of the SCD1 gene and regulates SCD1 gene expression through a mechanism independent of SREBP-1 maturation in vivo. The detailed mechanism of the down-regulation of SCD1 caused by PUFA has not been resolved. Furthermore, Matsuzaka et al. (*27*) reported that Δ6-desaturase and Δ5-desaturase expression is dually regulated by SREBP-1c and PPAR-α. At least, PUFA are thought to also autoregulate their biosynthesis through SREBP. In addition, CETP might interact with SREBP-1 (*28*) and is down-regulated by PUFA (*16*). As the expression level of CETP in HepG2 cells is very low, we could not evaluate the effect of PUFA on its expression using the oligonucleotide chip system (Table 1).

On the other hand, the gene expression of enzymes that catabolize FA (*29–32*), namely carnitine: palmitoyl-CoA acyltransferase 1 (CPT1), acyl-CoA oxidase

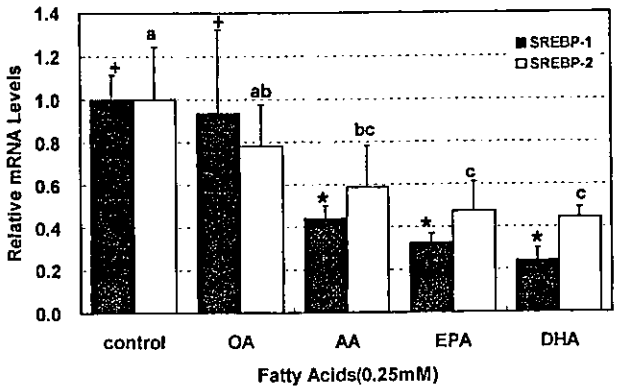

Fig. 1. Effect of PUFA on sterol regulatory element-binding protein (SREBP) mRNA expression in HepG2 cells. HepG2 cells were incubated with PUFA (0.25 mM) for 24 h. Total RNA was extracted, then mRNA expression levels of SREBPs were measured using real time RT-PCR as described in Materials and Methods. Relative mRNA levels were normalized to those of GAPDH. Values are means±SD (n=3). Mean values with different superscript letters in SREBP-1 expressions are significantly different (p<0.05). Different symbols ($^+$ and *) show significant differences in the SREBP-2 expressions (p<0.05).

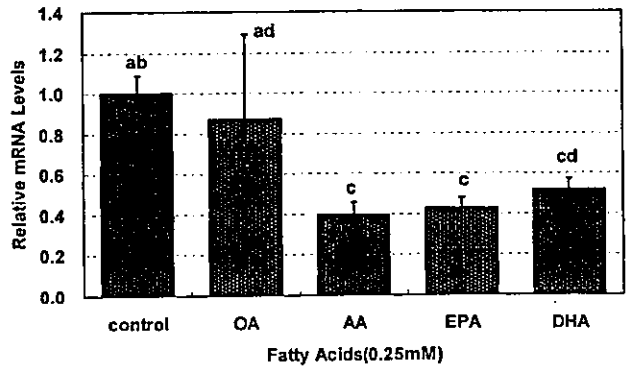

Fig. 2. Effect of PUFA on mevalonate pyrophosphate decarboxylase (MPD) expression in HepG2 cells. HepG2 cells were incubated with FA (0.25 mM) for 24 h. Total RNA was extracted, then mRNA expression of MPD were measured using real time RT-PCR as described in Materials and Methods. Relative mRNA levels were normalized to those of GAPDH. Values are means±SD (n=3). Mean values with different letters show significant differences in PUFA treatments (p<0.05).

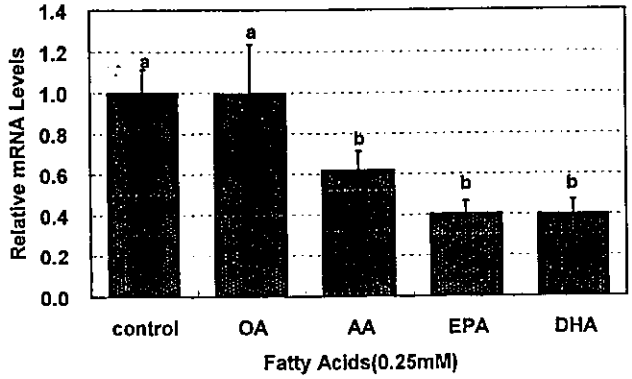

Fig. 3. Effect of PUFA on prostasin expression in HepG2 cells. HepG2 cells were incubated with FA (0.25 mM) for 24 h. Total RNA was extracted, then mRNA expression of prostasin were measured using real time RT-PCR as described in Materials and Methods. Relative mRNA levels were normalized to those of GAPDH. Values are means±SD (n=3). Mean values with different letters show significant differences in PUFA treatments (p<0.05).

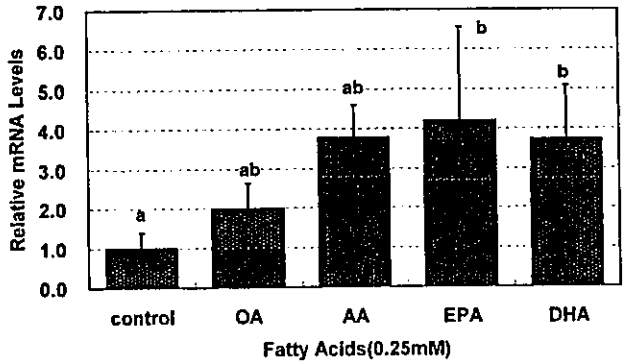

Fig. 4. Effect of PUFA on HTGL expression in HepG2 cells. HepG2 cells were incubated with FA (0.25 mM) for 24 h. Total RNA was extracted, then mRNA expression of HTGL were measured using real time RT-PCR as described in Materials and Methods. Relative mRNA levels were normalized to those of GAPDH. Values are means±SD (n=3 for control, OA, EPA and DHA, n=2 for AA). Mean values with different letters show significant differences in PUFA treatments (p<0.05).

(AOX) and acyl-CoA synthetase (ACS), which are induced by PUFA, did not change in our study. These enzymes are related to fatty acid oxidation and are generally believed to be regulated by PPAR (*5*, *33*, *34*). PPAR-α and -γ are located in the liver and adipocytes, respectively (*5*). The expression level of PPAR-α in human liver (*35*, *36*) is much lower than that in mouse liver, and over-expression of PPAR-α in HepG2 cells shows the induction of mitochondrial HMG-CoA synthase, CPT, and ACS mRNA (*37*). The present study detected only weak expression of PPAR-α and PPAR-γ in HepG2 cells (Table 2). Therefore, the effects might be obvious without regulation mediated by PPAR-α. We also showed that enzymes involved in the TCA cycle were up-regulated. Therefore, PUFA suppressed the synthesis of cholesterol and lipogenesis, but induced ATP generation by activation of the TCA cycle in HepG2 cells.

Takahashi et al. (*38*) have recently examined the effect of dietary fish oil on the gene expression profile in mouse liver using high-density oligonucleotide arrays. Although our findings were similar to theirs, they showed that immune reaction-related genes, antioxidant genes (several glutathione transferase, uncoupling protein 2 and Mn-superoxide dismutase) and genes involved in lipid catabolism were significantly up-regulated, indicating that dietary fish oil down-regulated the endogenous PPAR-α-activation system and increased the antioxidant gene expression that protects against excess ROS. Our data also suggested that PUFA induce

antioxidant genes, such as metallothionein-IG and extracellular-superoxide dismutase (SOD3). However, the overall response to oxidation was much less and the expression of microsomal glutathione *S*-transferase and manganese superoxide dismutase (SOD2) were not significantly changed (Table 2). We believe that little oxidative stress was induced by adding PUFA to HepG2 cells even though the PUFA were extremely pure (99%). The induction of immunological and antioxidant genes in their study might have been caused by adaptation to excess ROS production, since they fed the diet containing a very high concentration of fish oil (60% of total energy intake) for 6 mo.

Prostasin is a new serine protease that was purified from seminal fluid, and its cDNA has been sequenced (*39*). Prostasin is expressed in the human prostate, kidney, and lung, as well as in body fluids, including seminal fluid and urine (*40*). The relationship between prostasin and prostate cancer has been investigated (*41–44*). Prostasin might act as an extracellular regulator of epithelial sodium channels (*44*). However its physiological role in humans is not known. Prostasin was significantly suppressed by PUFA in this study and an SRE was located in its upstream region of the gene (*45*), suggesting that prostasin plays an important role in processing some proteins in response to cellular cholesterol concentrations.

PUFA also affected the genes involved in cell proliferation and differentiation. Further analysis using the data obtained by this study is needed in order to clarify the mechanism. PUFA are thought to control gene transcription through several steps. Together with their metabolites, PUFA play important roles in signal transduction cascades, and as ligands for transcription factors. Gene chip analysis might provide useful clues to investigate not only continuous regulation, but also the interaction between many transcription factors, such as SREBP and LXR.

Acknowledgments

This work was supported in part by grants from the Ministry of Health, Labor and Welfare for Aging and Health, the Japan Health Sciences Foundation for Health Sciences focusing on Drug Innovation, the Danone Institute Japan and Grants-in Aid for Scientific Research (No. 14580138) from the Ministry of Education, Science, Culture, Sports and Technology of Japan.

REFERENCES

1) Spector AA, York MA. 1985. Membrane lipid composition and cellular function. *J Lipid Res* **26**: 1015–1035.
2) Nestel PJ. 1990. Effect of *n*-3 fatty acids on lipid metabolism. *Annu Rev Nutr* **10**: 149–167.
3) Brown MS, Goldstein JL. 1997. The SREBP pathway: Regulation of cholesterol metabolism by proteolysis of a membrane-bound transcription factor. *Cell* **89**: 331–340.
4) Worgall TS, Sturley SL, Seo T, Osborne TF, Deckelbaum RJ. 1998. Polyunsaturated fatty acids decrease expression of promoters with sterol regulatory elements by decreasing levels of mature sterol regulatory element-binding protein. *J Biol Chem* **273**: 25537–25540.
5) Desvegne B, Wahli W. 1999. Peroxisome proliferator-activated receptors: Nuclear control of metabolism. *Endocr Rev* **20**: 649–688.
6) Tobin KA, Steinger HH, Alberti S, Spydevold O, Auwerx J, Gustafsson JA, Nebb HI. 2000. Cross-talk between fatty acid and cholesterol metabolism mediated by liver X receptor-alpha. *Mol Endocrinol* **14**: 741–752.
7) Hertz R, Magenheim J, Berman I, Bar-Tana J. 1998. Fatty acyl-CoA thioesters are ligands of hepatic nuclear factor-4 alpha. *Nature* **392**: 512–516.
8) Roche E, Buteau J, Anieto I, Reig JA, Soria B, Prentki M. 1999. Palmitate and oleate induce the immediate-early response genes *c-fos* and *nur-77* in the pancreatic β-cell line INS-1. *Diabetes* **47**: 2007–2014.
9) Duplus E, Glorian M, Forest C. 2000. Fatty acid regulation of gene transcription. *J Biol Chem* **275**: 30749–30752.
10) Lockhart DJ, Dong H, Byrne MC, Follettie MT, Gallo MV, Chee MS, Mittmann M, Wang C, Kobayashi M, Horton H, Brown EL. 1996. Expression monitoring by hybridization to high-density oligonucleotide arrays. *Nat Biotechnol* 14: 1675–1680.
11) Lee CK, Klopp RG, Weindruch R, Prolla TA. 1999. Gene expression profile of aging and its retardation by caloric restriction. *Science* **285**: 1390–1393.
12) Horton JD, Shimano H, Hamilton RL, Brown MS, Goldstein JL. 1999. Disruption of LDL-receptor gene in transgenic SREBP-1a mice unmasks hyperlipidemia resulting from production of lipid-rich VLDL. *J Clin Invst* **103**: 1067–1076.
13) Shimomura I, Bashmakov Y, Horton JD. 1999. Increased levels of nuclear SREBP-1c associated with fatty livers in two mouse models of diabetes mellitus. *J Biol Chem* **274**: 30028–30032.
14) Horton JD, Shimomura I, Brown MS, Hammer RE, Goldstein JL, Shimano H. 1998. Activation of cholesterol synthesis in preference to fatty acid synthesis in liver and adipose tissue of transgenic mice overproduction sterol regulatory element-binding protein-2. *J Clin Invest* **101**: 2331–2339.
15) Seo T, Oelkers P, Giattina MR, Worgall TS, Sturley SL, Deckelbaum RJ. 2001. Differential modulation of ACAT1 and ACAT2 transcription and activity by long chain free fatty acids in cultured cell. *Biochemistry* **40**: 4756–4762.
16) Hirano R, Igarashi O, Kondo K, Itakura H, Matsumoto A. 2001. Regulation by long-chain fatty acids of the expression of cholesteryl ester transfer protein in HepG2 cells. *Lipids* **36**: 401–406.
17) Takabe W, Mataki C, Wada Y, Ishii M, Izumi A, Aburatani H, Kamakubo T, Niki E, Kodama T, Noguchi N. 2000. Gene expression induced by BO-654, probucol and BHQ in human endothelial cells. *J Atheroscler Thromb* **7**: 223–230.
18) Akiyoshi S, Ishii M, Nemoto N, Kawabata M, Aburatani H, Miyazono K. 2001. Targets of transcriptional regulation by transforming growth factor-beta: expression profile analysis using oligonucleotide arrays. *Jpn J Cancer Res* **92**: 257–268.
19) Xu J, Nakamura MT, Cho HP, Clarks SD. 1999. Sterol regulatory element binding protein-1 expression is suppressed by dietary polyunsaturated fatty acids. *J Biol Chem* **274**: 23577–23583.
20) Kim HJ, Takahashi M, Ezaki O. 1999. Fish oil feeding

decreases mature sterol regulatory element-binding protein-1 (SREBP-1) by down-regulation of SREBP-1c mRNA in mouse liver. *J Biol Chem* **274**: 25892–25898.

21) Yahagi H, Shimano H, Hatsu A, Amemiya-Kudo M, Okazaki H, Tamura Y, Izuka Y, Shinoiri F, Ohashi K, Osuga J, Harada K, Gotoda T, Nagai R, Ishibashi S, Yamada N. 1999. A crucial role of sterol regulatory element-binding protein-1 in the regulation of lipogenic gene expression by polyunsaturated fatty acids. *J Biol Chem* **274**: 35840–35844.

22) Xu J, Teran-Garcia M, Park JH, Nakamura T, Clarke SD. 2001. Polyunsaturated fatty acids suppress hepatic sterol regulatory element-binding protein-1 expression by accelerating transcript decay. *J Biol Chem* **276**: 9800–9807.

23) Sakakura Y, Shimano H, Sone H, Takahashi A, Inoue N, Toyoshima H, Suzuki S, Yamada N, Inoue K. 2001. Sterol regulatory element-binding proteins induce an entire pathway of cholesterol synthesis. *Biochem Biophys Res Commun* **286**: 176–183.

24) Yoshikawa T, Shimano H, Yahagi N, Ide T, Amemiya-Kudo M, Matsuzaka T, Nakakuki M, Tomita S, Okazaki H, Tamura Y, Iizuka Y, Ohashi K, Takahashi A, Sone H, Osuga J, Gotoda T, Ishibashi S, Yamada N. 2002. Polyunsaturated fatty acids suppress sterol regulatory element binding protein 1c promoter activity by inhibition of liver X receptor (LXR) binding to LXR response elements. *J Biol Chem* **277**: 1705–1711.

25) Zhang L, Ge L, Tran T, Stenn K, Prouty SM. 2001. Isolation and characterization of the human stearoyl-CoA desaturase gene promoter: requirement of a conserved CCAAT *cis*-element. *Biochem J* **357**: 183–193.

26) Kim H-J, Miyazaki M, Ntambi JM. 2002. Dietary cholesterol opposes PUFA-mediated repression of the steroyl-CoA desaturase-1 gene by SREBP-1 independent mechanism. *J Lipid Res* **43**:1750–1757.

27) Matsuzaka T, Shimano H, Yahagi N, Amemiya-Kudo M, Yoshikawa T, Hasty AH, Tamura Y, Osuga J, Okazaki H, Iizuka Y, Takahashi A, Sone H, Gotoda T, Ishibashi S, Yamada N. 2002. Dual regulation of mouse Δ5- and Δ6 desaturase gene expression by SREBP-1 and PPARα. *J Lipid Res* **43**: 107–114.

28) Gauthier B, Robb M, Gaudet F, Ginsburg GS, MacPherson R. 1999. Characterization of a cholesterol response element (CRE) in the promoter of the cholesteryl ester transfer protein gene: functional role of the transcription factors SREBP-1a, and-2, and YY1. *J Lipid Res* **40**: 1284–1293.

29) Chatelain F, Kohl C, Esser V, McGarry JD, Girard J, Pegorier J-P. 1996. Cyclic AMP and fatty acids increase carnitine palmitoyltransferase I gene transcription in cultured fetal rat hepatocytes. *Eur J Biochem* **235**: 789–798.

30) Berthou L, Saladin R, Yaqoob P, Branellec D, Calder P, Frunchart J-C, Denefle P, Auwerx J, Staels B. 1995. Regulation of rat liver apolipoprotein A-I, apolipoprotein A-II and acyl-coenzyme A oxidase gene expression by fibrates and dietary fatty acids. *Eur J Biochem* **232**: 179–187.

31) Flatmark T, Niilsson A, Krannes J, Eikhom TS, Fukami NH. 1998. On the mechanism of Induction of the enzyme systems for peroxisomal B-oxidation of fatty acids in rat liver by diets rich in partially hydrogenated fish oil. *Biochim Biophys Acta* **962**: 122–130.

32) Ide T. 2001. Effect of dietary alpha-linolenic acid on the activity and gene expression of hepatic fatty acid oxidation enzymes. *Biofactors* **13**: 9–14.

33) Clarke SD, Jump DB. 1996. Polyunsaturated fatty acid regulation of hepatic gene transcription. *Lipids* **31**: S-7–S-11.

34) Price PT, Nelson CM, Clarke SD. 2000. Omega-3 polyunsaturated fatty acid regulation of gene expression. *Curr Opin Lipidol* **11**: 3–7.

35) Palmer CN, Hsu MH, Griffin KJ, Raucy JL, Johnson EF. 1998. Peroxisome proliferator activated receptor-alpha expression in human liver. *Mol Pharmacol* **53**: 14–22.

36) Gevois P, Torra IP, Chinetti G, Grotzinger T, Dubois G, Fruchart JC, Fruchart-Najib J, Leitersdorf E, Staels B. 1999. A truncated human peroxisome proliferator-activated receptor alpha splice variant with dominant negative activity. *Mol Endocrinol* **13**: 1535–1549.

37) Hsu MH, Savas U, Griffin KJ, Johnson EF. 2001. Identification of peroxisome proliferators-responsive human genes by elevated expression of the peroxisome proliferators-activated receptor α in HepG2 cells. *J Biol Chem* **276**: 27950–27958.

38) Takahashi M, Tsuboyama-Kasaoka N, Nakatani T, Ishii M, Tsutsumi S, Aburatani H, Ezaki O. 2002. Fish oil feeding alters liver gene expressions to defend against PPARα and ROS production. *Am J Physiol Gastrointest Liver Physiol* **282**: G338–G348.

39) Yu JX, Chao L, Chao J. 1995. Molecular cloning, tissue-specific expression, and cellular localization of human prostasin mRNA. *J Biol Chem* **270**: 13483–13489.

40) Laribi A, Berteau P, Gala J, Eschwege P, Benoit G, Tombal B, Schmitt F, Loric S. 2001. Blood-borne RT-PCR assay for prostasin-specific transcripts to identify circulating prostate cells in cancer patients. *Eur Urol* **39**: 65–71.

41) Chen LM, Hodge GB, Guarda LA, Welch JL, Greenberg NM, Chai KX. 2001. Down-regulation of prostasin serine protease: a potential invasion suppressor in prostate cancer. *Prostate* **48**: 93–103.

42) Mok SC, Chao J, Skates S, Wong K, Yiu GK, Muto MG, Berkowitz RS, Cramer DW. 2001. Prostasin, a potential serum marker for ovarian cancer: identification through microarray technology. *J Natl Cancer Inst* **93**: 1458–1464.

43) Donaldson SH, Hirsh A, Li DC, Holloway G, Chao J, Boucher RC, Gabriel SE. 2002. Regulation of the epithelial sodium channel by serine proteases in human airways. *J Biol Chem* **277**: 8338–8345.

44) Narikiyo T, Kitamura K, Adachi M, Miyoshi T, Iwashita K, Shiraishi N, Nonoguchi H, Chen LM, Chai KX, Chao J, Tomita K. 2002. Regulation of prostasin by aldosterone in the kidney. *J Clin Invest* **109**: 401–408.

45) Yu JX, Chao L, Ward DC, Chao J. 1996. Structure and chromosomal localization of the human prostasin (PRSS8) gene. *Genomics* **32**: 334–340.

Physiol Genomics 13: 31–46, 2003.
First published January 7, 2003; 10.1152/physiolgenomics.00116.2002.

Expression imbalance map: a new visualization method for detection of mRNA expression imbalance regions

Makoto Kano,[1] Kunihiro Nishimura,[2] Shumpei Ishikawa,[3] Shuichi Tsutsumi,[3] Koichi Hirota,[4] Michitaka Hirose,[4] and Hiroyuki Aburatani[3]

[1]*School of Engineering and* [2]*School of Information Science and Technology, University of Tokyo, Tokyo 113-8655; and* [3]*Genome Science Division, and* [4]*Intelligent Cooperative System, Department of Information Systems, Research Center for Advanced Science and Technology, University of Tokyo, 153-8904, Japan*

Submitted 4 September 2002; accepted in final form 20 December 2002

Kano, Makoto, Kunihiro Nishimura, Shumpei Ishikawa, Shuichi Tsutsumi, Koichi Hirota, Michitaka Hirose, and Hiroyuki Aburatani. Expression imbalance map: a new visualization method for detection of mRNA expression imbalance regions. *Physiol Genomics* 13: 31–46, 2003. First published January 7, 2003; 10.1152/physiolgenomics.00116.2002.—We describe the development of a new visualization method, called the expression imbalance map (EIM), for detecting mRNA expression imbalance regions, reflecting genomic losses and gains at a much higher resolution than conventional technologies such as comparative genomic hybridization (CGH). Simple spatial mapping of the microarray expression profiles on chromosomal location provides little information about genomic structure, because mRNA expression levels do not completely reflect genomic copy number and some microarray probes would be of low quality. The EIM, which does not employ arbitrary selection of thresholds in conjunction with hypergeometric distribution-based algorithm, has a high tolerance of these complex factors. The EIM could detect regionally underexpressed or overexpressed genes (called, here, an expression imbalance region) in lung cancer specimens from their gene expression data of oligonucleotide microarray. Many known as well as potential loci with frequent genomic losses or gains were detected as expression imbalance regions by the EIM. Therefore, the EIM should provide the user with further insight into genomic structure through mRNA expression.

gene expression profiling; allelic imbalance; chromosome mapping; hypergeometric distribution; computing methodologies

THE RECENT DEVELOPMENT of microarray technology has enabled simultaneous measurement of genome-wide expression profiles. Many research studies have revealed strong correlations between the expression profiles and cancer classifications. The next era of gene expression analysis would involve systematic integration of expression profiles and other types of gene information, such as locus, gene function, and sequence information. In particular, integration between expression profiles and locus information should be effective in detecting gene structural abnormalities such as genomic gains and losses.

In general, cancer progression is not a single but a multistep process and includes many genomic structural abnormalities. Among them, genomic gains and losses, particularly deletion of tumor suppressor genes and amplification of oncogenes, are associated with cancer progression and its malignant phenotype, although the affected lesion varies among different types of cancers. Comparative genomic hybridization (CGH) for detecting genome-wide abnormalities such as copy number changes, has been applied to various types of cancers (5), but its low resolution (~20 Mb, corresponding to about 200 genes) makes it difficult to identify the causal genes, the structural alternation of which is critical for cancer biological behavior.

Integration of gene expression profiles and gene locus information might allow detection of copy number changes at a much higher resolution. Several studies using oligonucleotide probe arrays suggested a strong relationship between genomic structural abnormalities and expression imbalances (underexpression or overexpression). Mukasa et al. (7) reported that the expression levels of a significant number of genes in the 1p region were reduced to about 50%, in oligodendrogliomas with 1pLOH. Furthermore, Virtaneva et al. (12) reported that acute myeloid leukemia with trisomy 8 was associated with overexpression of genes on chromosome 8. Recently, a genome-wide transcriptome map of non-small cell lung carcinomas based on gene expression profiles generated by serial analysis of gene expression (SAGE) was conducted (3). However, the simple spatial mapping of the expression profiles on chromosomal location sometimes hardly provides information about genomic structure for the following reasons: *1*) since some microarray probes are of low quality, the microarray signal intensities do not always reflect their target mRNA expression levels; and *2*) mRNA expression level does not completely reflect genomic copy number. The aim of the present study was to develop a new method with high tolerance of such complex factors, designed to detect regionally underexpressed or overexpressed genes in cancer specimens compared with the corresponding normal tissues. The expression imbalance region, constituted by

Article published online before print. See web site for date of publication (http://physiolgenomics.physiology.org).

Address for reprint requests and other correspondence: M. Kano, Tokyo Research Laboratory, IBM Japan, 1623-14 Shimotsuruma, Yamato-shi, Kanawaga 242-8502, Japan (E-mail: mkano@jp.ibm.com).

these genes, likely reflects genomic structural changes such as chromosomal gain and loss.

When developing the methodology that integrates the expression profiles and locus information, two significant problems have to be dealt with. First, a definition of what constitutes an expression imbalance region is not yet clarified. How many base pairs on chromosome should be considered as a genomic region (referred to below as chromosomal proximity)? To consider that a certain gene is differentially expressed in cancer and normal tissue, how much difference in the gene expression level is needed between the two (referred to below as cancer specificity)? It is generally very difficult to determine adequate thresholds for chromosomal proximity and cancer specificity. Arbitrary selection of thresholds would involve a risk of overlooking significant genes (that is, "threshold problem"). In addition, to detect expression imbalance regions, it is necessary to search for genes with both cancer specificity and chromosomal proximity. Because determining these two thresholds synergistically increases the risk of overlooking significant genes, the "threshold problem" is more critical in this case.

When selecting thresholds, several statistical theories such as hypothesis testing are helpful. However, commonly used statistical criteria are also arbitrarily determined. If thresholds are automatically determined based on statistical theory, the user cannot search more genes with potential significance, because the information of genes overlooked is almost unknown. Therefore, to detect as many significant genes as possible, a comprehensive presentation of the distribution of the "false balance" (that is, the balance of false negative and false positive) is quite significant rather than an attempt to seek potentially optimal statistical criterion.

Second, there are many candidate expression imbalance regions. Some of them may be a family of genes that are tandemly repeated and are under similar transcriptional regulations. To confirm that a candidate locus is biologically significant, human curation is necessary, using a variety of biological information. Therefore, it is important to present large genome-wide data in a comprehensive manner, indicating which genes are to be further examined. That is, a broadband interface between humans and computers is essential.

We focused on visualization technology as the key technology to solve these two problems. Visualization is effective in providing, genome-wide, the false-balance distribution and indication of the genes that are worth examining. The visualization used in our report would make it possible to present the images of all genes that have both cancer specificity and chromosomal proximity.

In this study, we developed a novel visualization method for detecting expression imbalance regions at much higher resolution than conventional technologies such as CGH, called the expression imbalance map (EIM). The EIM was applied to gene expression data of lung squamous cell carcinoma measured by oligonucleotide microarray and detected many known as well as potential loci with frequent genomic losses or gains as regional signal images on chromosomes (expression imbalance regions). In addition, the EIM could detect not only the expression imbalance common to all cancer specimens, but also individual differences among cancer specimens.

MATERIAL AND METHODS

Data Sets

In this article, the EIM is illustrated using the gene expression data of lung cancer from the study of Bhattacharjee et al. (1). In this experiment, total mRNA was extracted from histologically defined specimens of squamous cell lung carcinomas (abbreviated here as "SQ"; $n = 21$) and normal lung tissues (abbreviated here as "NL"; $n = 17$). The expression profiles were obtained using human U95A oligonucleotide probe arrays (GeneChip; Affymetrix, Santa Clara, CA). The SQ-NL gene expression data set (SQ, $n = 21$; NL, $n = 17$) was then analyzed using the EIM.

Feature Selection and Logarithmic Transformation

To compensate for distortion in the expression level, changes in the expression level were limited from 1 to 8,000. In addition, 4,083 probes with a mean expression above 50 and CV (CV = mean/standard deviation) above 0.2 were selected to eliminate potential low-quality probes. The common logarithm of the gene expression data was used for the following analysis.

Translation from Probe to UniGene

To associate gene locus information with gene expression profiles, each "probeID" on the U95A array was translated to UniGene, using information on the UniGene web site of the National Center for Biotechnology Information (NCBI), by referring to the corresponding original GenBank accession number of each probe set. Then, 11,334 of 12,533 probes on the U95A array were translated into 8,851 UniGenes.

Gene Locus Information

Gene locus information was obtained from the web sites for Genes On Sequence Map (*Homo sapiens* build 27) of NCBI and is defined as "LocusID." Among the LocusIDs on chromosome 1 to 22 of Genes On Sequence Map, the 12,063 LocusIDs, which had the corresponding UniGenes, were utilized to identify the chromosome locations of genes. Since the gene expression data utilized in this study were obtained from both sexes, the X and Y chromosomes were excluded. However, by using the data obtained from only males or females, the EIM can be applied to the analysis of chromosome X and Y. Since the 12,063 LocusIDs had one-to-one correspondence with UniGenes, they were translated into 12,063 UniGenes. However, only 6,652 of the 12,063 UniGenes were in common with the 8,851 UniGenes translated from the probes on the U95A array (Fig. 1). In this article, these 6,652 UniGenes are called "Key-UniGenes." The distributions of the UniGenes and Key-UniGenes on each arm of the chromosome are shown in Table 1. The number of total Key-UniGenes was defined as U (=6,652).

Quantization of Each Chromosome Arm Region

For easier handling of the gene locus information, each chromosome arm region was quantized by unit region called

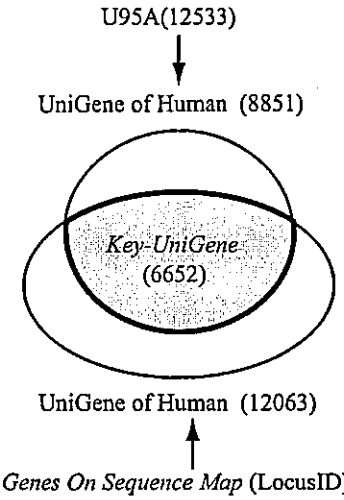

Fig. 1. Correspondence between probeIDs and LocusIDs. To associate gene locus information with gene expression profiles, probeIDs on the Affymetrix U95A oligonucleotide arrays and the LocusIDs on Genes On Sequence Map (*Homo sapiens* build 27) of NCBI were translated into UniGenes. We utilized the 12,063 LocusIDs, which had the corresponding UniGenes, on chromosome 1 to 22 of Genes On Sequence Map. The X and Y chromosomes were excluded, because the gene expression data utilized in this study were obtained from both sexes. Since these 12,063 LocusIDs had one-to-one correspondence with UniGenes, these were translated into 12,063 UniGenes. Out of 12,533 probes on the U95A array, 11,334 were translated into unduplicated 8,851 UniGenes, by referring to the corresponding original GenBank accession number of each probe set. Although the 12,063 UniGenes were obtained from Genes On Sequence Map, only 6,652 of the 12,063 UniGenes were in common with the 8,851 UniGenes translated from the probes on the U95A array. In this article, these 6,652 UniGenes are called "Key-UniGenes."

"bucket" whose length was 100,000 base pairs (100 kbp), and the Key-UniGenes were assigned the corresponding buckets according to their reading position (Fig. 2, *A* and *B*). A reading position indicates the start position for gene transcription and was obtained from Genes On Sequence Map. The number of buckets on chromosome arm *arm* was defined as L_{arm}.

Formation of Locus Cluster

To evaluate the proximity of genes on chromosome arm *arm*, the Key-UniGenes on the *length* neighbor buckets from (*begin*)-th were defined as a cluster $C_{arm\_length\_begin}$ (Fig. 2*A*). Repeating the sufficiently minute changes of *length* and *begin* formed the exhaustive uncertainty cluster sets of Key-UniGenes with chromosomal proximity (Fig. 2*C*). The EIM allows even clusters that overlap each other or include others. Therefore, all neighbor buckets in any area of each chromosome arm were defined as clusters. The number of Key-UniGenes in the cluster $C_{arm\_length\_begin}$ was defined as $n_{arm\_length\_begin}$. $C_{arm\_length\_begin}$ was defined for all

$$arm = 1p, 1q, 2p, 2q, \ldots, 22p, 22q$$

$$length = 2, 3, 4, \ldots [buckets]$$

$$begin = 1, 2, \ldots, (L_{arm} - length + 1)$$

In addition, to avoid considering a region that contains large gaps between genes as "one region," the gaps between the Key-UniGenes that lie next to each other in $C_{arm\_length\_begin}$ were calculated and the maximal gap was defined as $gap_{arm\_length\_begin}$ (Fig. 2*B*). The EIM allows the user to filter out the cluster(s) whose $gap_{arm\_length\_begin}$ is more than gap_{max}, which can be changed interactively. In other words, the user can exclude regions containing large gaps by controlling gap_{max}. When gap_{max} values were 500 kbp, 1 Mbp, 2 Mbp, and 3 Mbp, the percentages of the gaps that were less than gap_{max} were 77.6, 89.4, 96.0, and 98.2%, among all gaps between the Key-UniGenes that lie next to each other.

EIM for Detection of Expression Imbalance Specific To Squamous Cell Carcinomas

Clusters consisting of genes with expression profiles specific to SQs. Probes with expression profiles specific to SQs were extracted as a cluster from 4,083 probes of SQ-NL data sets. Although the EIM does not depend on the type of statistical method used for evaluating the difference between two groups, nonparametric tests such as the Mann-Whitney test have the advantage that no assumption is needed about the distribution of data, compared with parametric tests such as the *t*-test. Thus we explain the case of the Mann-Whitney test as an example.

More specifically, the difference in the level of expression of each gene between two groups (SQs and NLs) was defined using the statistical probability, *P*, of rank sum. Assume that there are two groups (G_a, $n = N_a$; G_b, $n = N_b$) and the rank sums in G_a and G_b are Sum_a and Sum_b, respectively, when all elements ($N_a + N_b$) are sorted in order. For simplicity, assume that Sum_a/N_a is greater than or equal to Sum_b/N_b. *P* is the probability of observing the rank sum of the N_a elements, which are randomly selected from all elements, to be more than Sum_a.

Table 1. *Number of the UniGenes and Key-UniGenes on Genes On Sequence Map*

Chr. Arm	UniGene Number	Key-UniGene Number (L_{arm})	Chr. Arm	UniGene Number	Key-UniGene Number (L_{arm})
1p	715	394	12p	211	107
1q	614	361	12q	488	289
2p	313	179	13p	0	0
2q	485	274	13q	218	127
3p	315	191	14p	0	0
3q	335	171	14q	411	228
4p	111	60	15p	0	0
4q	356	201	15q	379	197
5p	116	61	16p	254	130
5q	472	248	16q	244	123
6p	434	251	17p	218	130
6q	291	158	17q	513	290
7p	180	105	18p	52	34
7q	373	205	18q	135	76
8p	157	95	19p	391	199
8q	262	138	19q	481	249
9p	146	85	20p	122	53
9q	353	193	20q	245	124
10p	104	53	21p	0	0
10q	362	205	21q	137	83
11p	234	129	22p	0	0
11q	502	280	22q	334	176

Distributions of the UniGenes, which were obtained from Genes On Sequence Map (*Homo sapiens* build 27) of NCBI, and Key-UniGenes on each arm of the chromosome. Since the gene expression data utilized in this study were obtained from both sexes, the X and Y chromosomes were excluded. Key-UniGenes are the UniGenes that can be translated into from both the probes on the U95A oligonucleotide arrays and the LocusIDs on chromosome 1 to 22 of the Genes On Sequence Map. The total numbers of the UniGenes and Key-UniGenes are 12,063 and 6,652, respectively. Chr., chromosome; L_{arm}, number of "buckets" on chromosome arm *arm*.

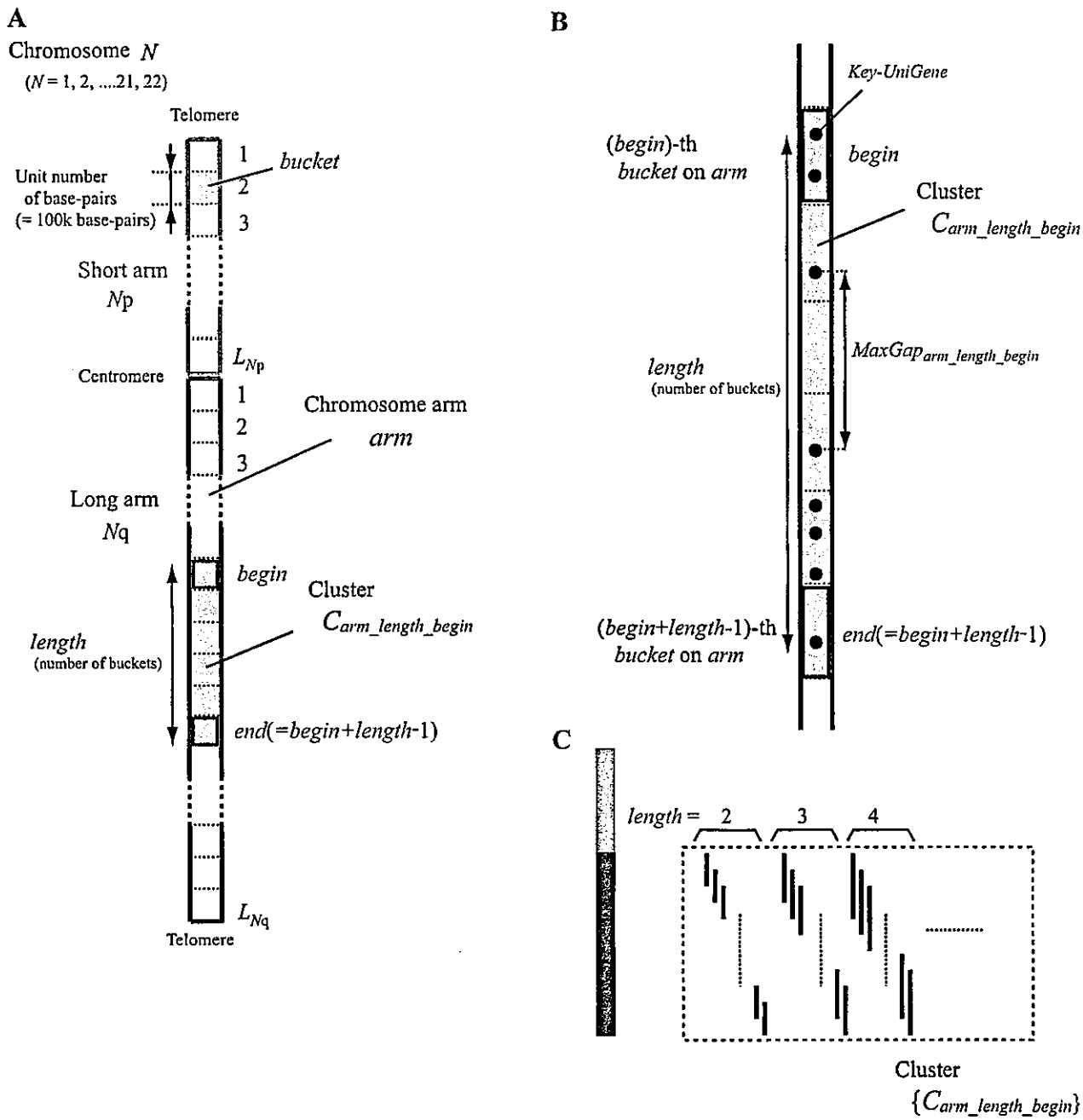

Fig. 2. Formation of clusters of genes with chromosomal proximity. *A*: for easier handling of the gene locus information, each chromosome arm region was quantized by unit region called "bucket" whose length was 100 kbp, and the Key-UniGenes were assigned the corresponding buckets according to their reading positions, which were obtained from Genes On Sequence Map (*Homo sapiens* build 27) of NCBI. The number of buckets on chromosome arm *arm* was defined as L_{arm}. To evaluate the proximity of genes on chromosome arm *arm*, the Key-UniGenes on the *length* neighbor buckets from (*begin*)-th were defined as a cluster $C_{arm\_length\_begin}$. *B*: to avoid considering a region containing large gaps between genes as "one region," the gaps between Key-UniGenes which lie next to each other in $C_{arm\_length\_begin}$ were calculated and the maximal gap was defined as $gap_{arm\_length\_begin}$. The expression imbalance map (EIM) allows the user to filter out the clusters whose $gap_{arm\_length\_begin}$ is more than gap_{max}, which can be changed interactively. In other words, the user can exclude regions containing large gaps by controlling gap_{max}. *C*: repeating the sufficiently minute changes of *length* and *begin* formed the exhaustive uncertainty cluster set of locus information. The EIM allows even the clusters that overlap each other or include others. Therefore, all neighbor buckets in any area of each chromosome arm were defined as clusters.

Based on this P value, the differential level $D_1(g)$ in which g is the probe name was defined as follows

$$D_1(g) = -\log_{10}P \quad (1)$$

Probes whose differential level D_1 was equal to or more than *diff* were defined as a cluster of probes with expression profiles specific to SQs, C_{sign_diff} (Fig. 3). The suffix *sign* indicates a differential direction (+, overexpression; −, underexpression in SQs). Repeating the sufficiently minute changes of *diff* formed the exhaustive uncertainty set of the clusters specific to SQs. C_{sign_diff} was defined for all

$$sign = -, +$$

$$diff = 2, 3, 4, \ldots$$

For example, C_{+3} was a cluster of probes whose differential level $D_1(g)$ of overexpression was 3 or more. The EIM was constructed by all the clusters C_{sign_diff} with *diff* greater than or equal to the minimum acceptable differential level d_{min} (Fig. 3). Since the default value of d_{min} is 2, all the clusters, C_{sign_diff}, would be utilized. The EIM allows the user to control d_{min} interactively for narrowing down the probes, if needed.

The numbers of probes, UniGenes, and Key-UniGenes of each cluster are shown in Table 2; n_{sign_diff} is the number of Key-UniGenes translated from probes of C_{sign_diff}. When multiple probes in a cluster could be mapped to a single UniGene, only the probe with the highest D_1 value was adopted. In addition, Fig. 3 shows probe permutations whose differential levels are 2 or more, arranged in the order of the differential level. Probes with under- and overexpression are arranged on the left and the right of Fig. 3, respectively.

Construction of the EIM. To detect the expression imbalance regions, it is necessary to search for genes with both cancer specificity and chromosomal proximity. The fundamental algorithm of the EIM is to statistically evaluate the overlaps between clusters of genes with cancer specificity and clusters of genes with chromosomal proximity. The clusters specific to the group of SQs, C_{sign_diff}, are arranged on the abscissa, and the locus clusters, $C_{arm\_length\_begin}$, are on the ordinate, as shown in Fig. 4. The variable k is the number of common Key-UniGenes between C_{sign_diff} and $C_{arm\_length\_begin}$.

The variable k could be evaluated using the hypergeometric probability, H, for observing at least k common elements between randomly selected n_1 and n_2 elements among all U elements as follows, where n_1 is n_{sign_diff} and n_2 is $n_{arm\_length\_begin}$.

Table 2. *Clusters of probes with expression profiles specific to the group of squamous cell lung carcinomas*

Differential Direction	Cluster Name (C_{sign_diff})	Probe Number	Key-UniGene Number (n_{sign_diff})
Underexpression (SQ < NL)	C_{-2}	1,007	668
	C_{-3}	844	567
	C_{-4}	642	429
	C_{-5}	448	301
	C_{-6}	283	188
	C_{-7}	83	61
Overexpression (SQ > NL)	C_{+2}	958	613
	C_{+3}	759	480
	C_{+4}	543	329
	C_{+5}	334	205
	C_{+6}	143	95
	C_{+7}	13	8

The probes (on the Affymetrix U95A arrays) whose expression profiles show significant difference between squamous cell lung carcinomas (SQs) and normal lung (NLs) were extracted as clusters, C_{sign_diff}. The suffix *sign* indicates the differential direction ("+" = overexpression; "−" = underexpression in SQs), and *diff* indicates a differential level D_1 in gene expression profiles between SQs and NLs. For example, C_{+3} is a cluster of probes whose differential level of overexpression is 3 or more. Repeating the sufficiently minute changes of *diff* formed the exhaustive set of the clusters consisting of genes with expression profiles specific to SQs. The numbers of probes and Key-UniGenes for each cluster are shown.

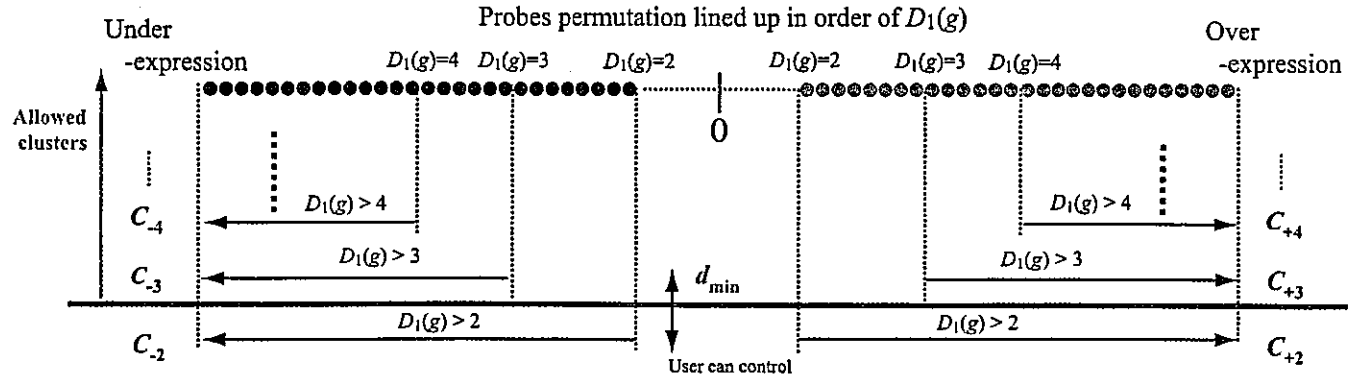

Fig. 3. Probe permutation arranged in order of the difference in gene expression level between squamous cell lung carcinomas (SQs) and normal lungs (NLs). Probes on the U95A arrays are lined up in order of the $D_1(g)$ level, which represents the difference in the gene expression level between SQs and NLs. Only probes with differential levels of 2 or more were arranged. Probes with underexpression and overexpression in SQs are arranged on the *left* and *right* side, respectively. Probes whose differential level $D_1(g)$ is equal to or more than *diff*, are defined as a cluster of probes with expression profiles specific to SQs, C_{sign_diff}. The suffix *sign* indicates the differential direction (+, overexpression; −, underexpression in SQs). Repeating the sufficiently minute changes of *diff* formed the exhaustive uncertainty set of the clusters specific to SQs. The EIM was constructed by all clusters C_{sign_diff} with *diff* that were greater than or equal to the minimum acceptable differential level d_{min}. Since the default value of d_{min} is 2, all the clusters, C_{sign_diff}, would be utilized. The EIM allows the user to control d_{min} interactively for narrowing down the probes, if needed.

$$H(U,n_1,n_2,k) = 1 - \sum_{i=0}^{k-1} \frac{\binom{n_2}{i} \cdot \binom{U-n_2}{n_1-i}}{\binom{U}{n_1}} \quad (2)$$

When the H value is small, the overlap between C_{sign_diff} and $C_{arm\_length\_begin}$ is considered statistically significant. That is, if the H value is small, then the overlap did not occur accidentally. Thus the evaluation value, E, is defined as follows

$$E(U,n_1,n_2,k) = -\log_{10}H(U,n_1,n_2,k) \quad (3)$$

For any combination of C_{sign_diff} and $C_{arm\_length\_begin}$, if both (*begin*)-th and (*begin* + *length* − 1)-th buckets of $C_{arm\_length\_begin}$ have the Key-UniGenes that are included in C_{sign_diff}, then their E values were calculated. This calculation was preprocessing for the EIM. Then, in real-time processing, if both C_{sign_diff} and $C_{arm\_length\_begin}$ met d_{min} and gap_{max}, respectively, then the E value was represented in the intersection area $R_{sign\_diff\_arm\_length\_begin}$ as a gray scale. The user can control d_{min} and gap_{max} interactively. The area where the multiple $R_{sign\_diff\_arm\_length\_begin}$ values overlapped is overwritten at the maximum E value (Fig. 4*B*). A flowchart that details these steps is shown in Fig. 5. The EIM for detecting expression imbalance specific to SQs is shown in Fig. 6. In addition, Fig. 7 shows chromosome 3 of the EIM and the influence of gap_{max} and d_{min} on the detection of the expression imbalance regions specific to SQs.

EIM for Detection of Individual Differences in Expression Imbalance Among SQs

It is effective to extract probes with expression profiles specific to the group of cancers using statistical analyses, such as the Mann-Whitney analysis. However, because this type of analysis treats all specimens with the same pathological diagnosis as one group, the variation in a group is unobservable. This is sometimes a significant problem because cancer specimens generally have a great number of variations. Thus we also developed the EIM for detecting individual differences in expression imbalance among SQs.

Clusters of probes with expression imbalance in each SQ. The first step in the development of the EIM for detecting individual differences in expression imbalance among SQ specimens was to extract probes with under- or overexpression compared with NL specimens, in each SQ specimen independently. Assuming that the expression levels of a certain probe, g, in NL specimens have a lognormal distribution, if the expression level of a SQ specimen, S_i, is included in $100p\%$ of sections on both sides of NL's distributions, its differential level D_2 was defined as follows

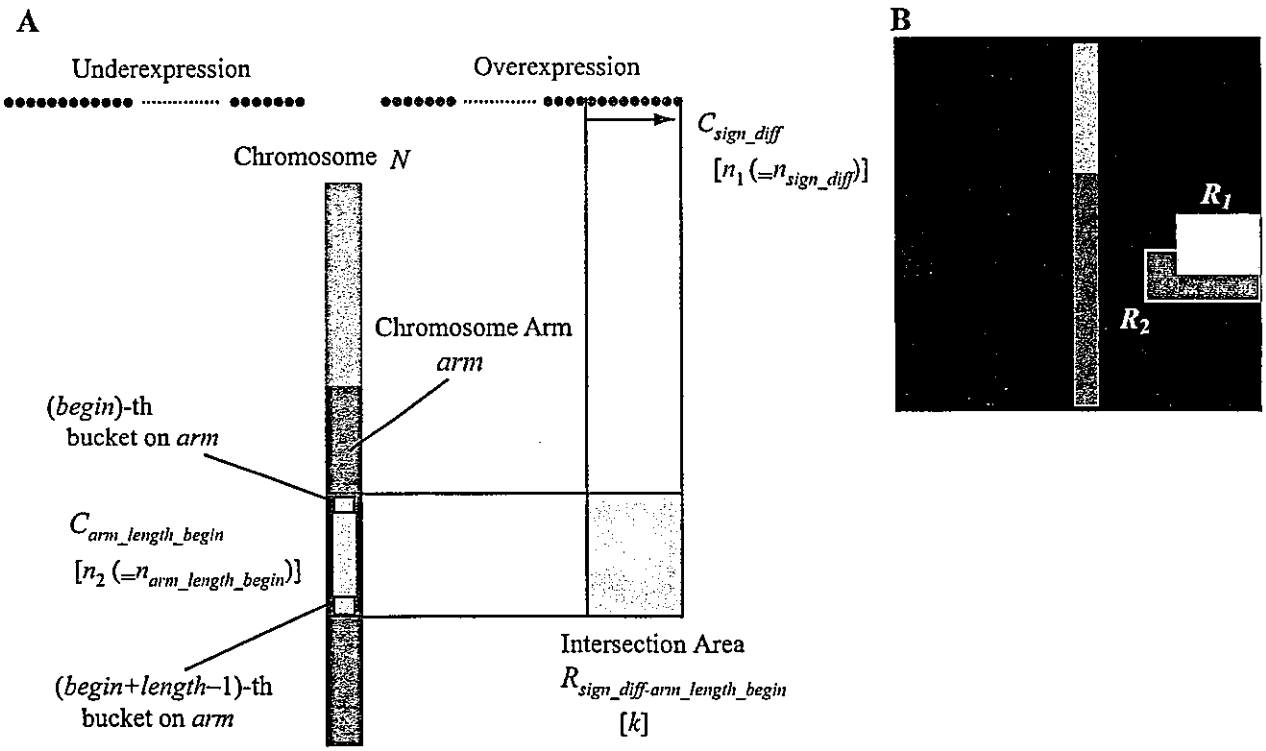

Fig. 4. Clusters of genes specific to the group of SQs vs. clusters of genes with proximity on chromosomes. *A*: to detect expression imbalance regions, it is necessary to search for genes with both cancer specificity and chromosomal proximity. The fundamental algorithm of the EIM is to evaluate statistically the overlaps between clusters of genes with cancer specificity and clusters of genes with chromosomal proximity. The clusters of probes with expression specific to the group of SQ, C_{sign_diff}, are arranged on the abscissa, and those of Key-UniGenes with proximity on chromosomes, $C_{arm\_length\_begin}$, on the ordinate. Among C_{sign_diff} values, the clusters of probes with underexpression and overexpression in SQs are arranged on the *left* and *right* side, respectively. The n_{sign_diff} and $n_{arm\_length\_begin}$ are the numbers of Key-UniGenes in C_{sign_diff} and $C_{arm\_length\_begin}$, respectively; k is the number of common Key-UniGenes both in C_{sign_diff} and $C_{arm\_length\_begin}$. The statistical significance of the overlap between C_{sign_diff} and $C_{arm\_length\_begin}$ was visualized in the intersection area $R_{sign\_diff\_arm\_length\_begin}$ as a gray scale. *B*: the area where the multiple $R_{sign\_diff\_arm\_length\_begin}$ overlapped was overwritten at the maximum E value. Therefore, when the E value of R_1 is higher than that of R_2, the area where R_1 and R_2 overlapped is overwritten at that of R_1.