Int. J. Cancer: **122**, 325–332 (2008)

Gankyrin oncoprotein overexpression as a critical factor for tumor growth in human esophageal squamous cell carcinoma and its clinical significance

Cristian M. Ortiz[1], Tetsuo Ito[2], Eiji Tanaka[1], Shigeru Tsunoda[1], Satoshi Nagayama[1], Yoshiharu Sakai[1], Hiroaki Higashitsuji[3], Jun Fujita[3] and Yutaka Shimada[4]*

[1]*Department of Surgery, Graduate School of Medicine, Kyoto University, Sakyo-ku, Kyoto, Japan*
[2]*Department of Medicine and Greenebaum Cancer Center, University of Maryland School of Medicine, Baltimore, MD*
[3]*Department of Clinical Molecular Biology, Graduate School of Medicine, Kyoto University, Sakyo-ku, Kyoto, Japan*
[4]*First Department of Surgery, Hyogo College of Medicine, Nishinomiya, Hyogo, Japan*

To elucidate the possible involvement of gankyrin in ESCC progression and the effect of its down-regulation in ESCC, we investigated the expression of gankyrin in ESCC tissues comparing it with the corresponding normal esophageal epithelia and tested a short-hairpin RNA (shRNA) expression vector for gankyrin in ESCC cell lines. Gankyrin protein expression in 11 ESCC cell lines (KYSE series) was examined by RT-PCR and western blot. The expression of gankyrin mRNA in 30 ESCC tissues was compared with the corresponding normal epithelia by Real-time PCR. Expression of gankyrin protein was immunohistochemically analyzed in the ESCC of 103 patients. A gankyrin-shRNA vector was stably transfected into KYSE 170 cells to assess the role of gankyrin in cell motility, invasion and proliferation *in vitro* and tumor formation *in vivo*. Gankyrin expression increased in all 11 ESCC cell lines. Real-time PCR revealed that gankyrin expression was higher in the cancerous tissue for all 30 patients. In immunohistochemistry, gankyrin overexpression was correlated with lower survival rate ($p = 0.0001$), extent of the primary tumor, lymph node metastasis, distant lymph node metastasis and stage ($p = 0.0072$, $p = 0.0004$, $p = 0.0172$ and $p = 0.0002$, respectively). A shRNA vector against gankyrin repressed growth, cell motility, invasiveness *in vitro* and tumor formation *in vivo*. Gankyrin overexpression is associated with poor prognosis. It may play an important role in ESCC tumor progression and could be a potentially important therapeutic gene target in ESCC.
© *2007 Wiley-Liss, Inc.*

Key words: gankyrin; oncogene; esophageal cancer; clinical significance; novel; cell cycle

Despite considerable advances in surgical techniques, perioperative care and neoadjuvant chemoradiotherapy, esophageal squamous cell carcinoma (ESCC) still remains one of the most lethal cancers and the seventh leading cause of cancer death worldwide.[1,2] Gene study is widely performed in ESCC[3–5] and new oncogenes associated with the progression of ESCC have been identified and targeted to improve the survival of patients with this type of refractory cancer.

Gankyrin is a novel oncoprotein with 7 ankyrin repeats.[6–8] Overexpression of gankyrin increases the phosphorylation and degradation of retinoblastoma (RB1) by its RB1-binding motif, suggesting that it promotes tumorigenicity and cancer cell proliferation.[9] Furthermore, binding to CDK4 and counteracting the inhibitory function of the INK4 proteins, including tumor suppressors $p16^{INK4A}$ and $p18^{INK4B}$, gankyrin acts as an accelerator for cell cycle progression.[9–11] Gankyrin also binds to MDM2, facilitating MDM2-p53 interaction and increasing its polyubiquitilation and degradation.[12] These findings suggest the importance of gankyrin in the cell cycle control and tumorigenesis. However, the physiological and physiopathological role of gankyrin in the cell cycle progression of normal cells remains unknown.

Overexpression of gankyrin has been found ubiquitously in hepatocellular carcinoma tissues, liver regeneration and fulminant hepatic failure (FHF),[9] but there are no reports regarding its overexpression in esophageal cancer. To clarify the clinical significance of gankyrin and evaluate the role of this molecule in the progression of ESCC, the following study was carried out by examining the gankyrin mRNA and protein expression from 11 ESCC cell lines (KYSE series). We then examined gankyrin protein expression in 103 ESCC tumors by immunohistochemistry (IHC), analyzing its clinical significance. In addition, to obtain better understanding of the role of gankyrin in ESCC we used a plasmid vector short-hairpin RNA (shRNA) expression system *in vitro* and *in vivo*, that provides a powerful tool used for inducing the loss of function phenotypes by posttranscriptional gene silencing.[13,14]

Material and methods

Surgical specimens and cell culturing

Frozen esophageal tumor tissues and their corresponding normal tissues were obtained for Real-time PCR from 30 patients with primary ESCC who underwent surgery at Kyoto University Hospital from 1991 to 2004. Paraffin-embedded sections were acquired for IHC from 103 patients with primary ESCC who underwent surgery at Kyoto University Hospital from 1997 to 2003. The tumor characteristics are summarized in Table I. The median follow-up time of survival was 30 months. All tumors were confirmed as ESCC by the clinicopathologic department of the hospital. All the cases were diagnosed as formalin-fixed, paraffin-embedded and hematoxilin and eosin (H&E)-stained representative specimens. These were classified according to the fifth edition of the pathological tumor-node-metastasis (TNM) classification.[15–17] Information on the patients' gender, age, stage of disease and histological factors was extracted from the medical records. Written informed consent was obtained from all patients regarding the performance of surgery and the use of resected samples for research (the approval numbers of the Institutional Review Board of Kyoto University are 232 and G48).

All human esophageal squamous carcinoma cell lines (KYSE series) were established in our laboratory and maintained in RPMI 1640 (Life Technologies, Gaithersburg, MD) and Ham's F12 (Nissui Pharmaceutical, Tokyo, Japan) with 2% fetal bovine serum (FBS).[18] The normal esophageal epithelial cell line NEK2 was also established in our laboratory and maintained in a keratinocyte serum-free medium containing 2.5 mg of epidermal growth factor and 25 mg of bovine pituitary extract.[19] HeLa cells were purchased from the American Type Culture Collection (Rockville, MD), cultured in DMEM (Life Technologies) with 10% FCS and used as a positive control.

Purification of total cellular mRNA and reverse transcription-PCR

Total RNA was extracted from KYSE cell lines and the frozen tissues of ESCC patients by the TRIzol reagent (Invitrogen, Carlsbad, CA) according to the manufacturer's protocols.[20,21]

This article contains supplementary material available via the Internet at http://www.interscience.wiley.com/jpages/0020-7136/suppmat.

Grant sponsor: Japanese Ministry of Education, Culture, Sports, Science and Technology; Grant number: 17390363.

***Correspondence to:** First Department of Surgery, Hyogo College of Medicine, 1-1 Mukogawa-cho, Nishinomiya, Hyogo 663-8501, Japan. Fax: +81-79-845-6581. E-mail: sshimada@hyo-med.ac.jp

Received 31 March 2007; Accepted after revision 24 July 2007
DOI 10.1002/ijc.23106
Published online 12 October 2007 in Wiley InterScience (www.interscience.wiley.com).

TABLE I – THE RELATIONSHIP BETWEEN GANKYRIN PROTEIN EXPRESSION AND CLINICOPATHOLOGIC CHARACTERISTICS IN 103 PATIENTS WITH ESCC

Gankyrin expression		− or +	++	+++	*p*-value
Age	<61 yr	11 (10.7)	16 (15.5)	28 (27.2)	0.8073
	≥61 yr	11 (10.7)	15 (14.3)	22 (21.3)	
Gender	Male	17 (16.5)	26 (25.2)	42 (40.8)	0.7753
	Female	5 (4.9)	5 (4.9)	8 (7.8)	
TNM[1]	T1/T2	16 (10.7)	12 (11.7)	17 (16.5)	0.0072
	T3/T4	6 (5.8)	19 (18.4)	33 (32.0)	
	N0	15 (14.3)	9 (8.7)	10 (9.7)	0.0004
	N1	7 (6.8)	22 (21.3)	40 (38.8)	
	M0	22 (21.3)	26 (25.2)	36 (35.0)	0.0172
	M1	0 (0)	5 (4.9)	14 (13.6)	
TNM stage	I/IIA/IIB	19 (18.4)	15 (14.3)	18 (17.5)	0.0002
	III/IV	3 (2.9)	16 (10.7)	32 (31.1)	
Histology	Well/Moderate	3 (2.9)	7 (6.8)	8 (7.8)	0.6097
	Poor	17 (16.5)	21 (20.4)	39 (37.9)	
Location	U	4 (3.9)	9 (8.7)	14 (13.6)	0.6457
	M	9 (8.7)	11 (10.7)	23 (22.3)	
	L	9 (8.7)	11 (10.7)	13 (12.6)	

Values in parentheses indicate percentage.
[1]TNM: T, extent of primary tumor; N, Spread to regional lymph nodes; M, Distant metastasis.

Reverse transcription of total cellular RNA (5 mg) was performed using a First-Strand cDNA Synthesis Kit (Amersham, Buckinghamshire, UK). cDNA was subjected to PCR for 25 cycles of amplification using an Advantage cDNA PCR kit (Becton Dickinson Biosciences, Palo Alto, CA). Each PCR cycle consisted of a denaturation step for 1 min at 94°C and an annealing step for 1 min at 58°C. The final extension step was carried out for 5 min at 72°C. The PCR products were separated on 2.0% agarose gel and visualized by ethidium bromide staining. The PCR primers used for gankyrin were 5′-AGCAGCCAAGGGTAACTTGA-3′ as the forward primer and 5′-TACTTGCTCCTTGGGACACC-3′ as the reverse primer; and for glyceraldehydes-3-phosphate dehydrogenase (G3PDH), 5′-TGGTATCGTGGAAGGACTCATGAC-3′ was used as the forward primer and 5′-ATGCCAGTGAGCTTCCCGTTCAGC-3′ as the corresponding reverse primer. The cDNA from the HeLa cells was used as a positive control for each analysis.

Real-time PCR with SYRB green I

Real-time PCR was carried out using the LightCycler FastStart DNA Master SYBR Green I system (Roche applied science, Penzberg, Germany). The reaction mix contained 2 μl of LightCycler FastStart DNA master mix SYBR Green I (Roche Diagnostics GmbH), 1 μl (0.5 mM) of each primer, 2.4 μl (4 μM) of $MgCl_2$ and 2 μl of DNA template and was amplified with activation of polymerase (95°C for 10 min), followed by 45 cycles of 10 sec at 95°C, 5 sec at 58°C and 10 sec at 72°C. All the samples were measured 3 times independently including 4 gankyrin standards. For the internal control, the targeted gene gankyrin was compared to the housekeeping gene G3PDH. The negative control contained all the elements of the reaction mix without the template DNA. The analysis of fluorescence curves was done by LightCycler software v.3.5 (Maximum second derivate) and expressed in a number of copies.

Immunohistochemical staining

Resected esophageal specimens were fixed in 10% formaldehyde solution, embedded in paraffin, cut in 4-μm thick sections, and mounted on aminopropyltriethoxysilane-coated glass slides. Immunohistochemical staining was carried out using an Envision Kit (Dako Cytomation, Glostrup, Denmark). The tissue sections were deparaffinized and rehydrated, then, the unmasking was performed in an autoclave at 121°C for 5 min. The sections were incubated overnight at 4°C with anti-human gankyrin polyclonal antibody (SantaCruz Biotechnology, Japan, diluted 1:50) in PBS containing 1% bovine serum albumin and then incubated with anti-rabbit labeled polymer horseradish peroxidase for 30 min at room temperature. Then the sections were incubated with 3,3′-diaminobenzidine liquid system (Dako Cymation, Glostrup, Denmark), counterstained with Mayer's haematoxilyn, dehydrated and then mounted. As a negative control, the primary antibody was replaced with normal mouse IgG and another control was carried out without the primary antibody. The immunoreactivity of each slide was confirmed using an anti-human cytokeratin antibody (Dako Cytomation, Glostrup, Denmark).

Evaluation of immunohistochemical staining

Gankyrin immunohistochemical staining was evaluated in 5 areas of each section and the overexpression of gankyrin was divided into 3 groups according to the percentage of gankyrin positive nuclei among the tumor cells: − or positive + when less than 50% of nuclei was stained; positive ++ when 50–70% of nuclei was stained; and positive +++ when more than 70% of nuclei was stained (Fig. 3). All slides were independently evaluated by 2 investigators (C.O. and E.T.) without prior knowledge of each patient's clinical information. When the opinions of the 2 evaluators were different, agreement was reached by careful discussion.

Western blot analysis

Whole-cell extract lysate was made from 1×10^7 cells in a sample buffer (2% sodium dodecyl sulfate [SDS], 10% glycerol, 50 mM Tris-HCl, pH 6.8). It was then electrophoresed on a 12% polyacrylamide gel SDS page and transferred to a polyvinylidene difluoride membrane (Immobilon, Milipore, Bedford, MA) using a semidry transfer blot system (Bio-Rad, Hercules, CA). The membranes were blocked with TBS (20 mM Tris, 150 mM NaCl, pH 7.6) containing 1% Tween 20 and 5% skimmed milk for 1 hr. The membranes were incubated at 4°C overnight with anti-human gankyrin rabbit polyclonal antibody (Santa Cruz Biotechnology, diluted 1:500) or with anti-human β-actin mouse monoclonal antibody (Sigma, St. Louis, MO; diluted 1:2,000) as an internal control. They were washed and then incubated at room temperature for 1 hr with goat anti-rabbit IgG-HRP or anti-mouse IgG (Zymed, San Francisco, CA) respectively, as a secondary antibody, and analyzed using ECL plus reagent (Amersham, Buckinghamshire, UK).

Statistical analysis

The correlation between gankyrin expression and each clinicopathological factor was evaluated by Fisher's exact test. Cumulative survival curves were calculated by the Kaplan–Meier method and analyzed by the log-rank test. Multivariate analysis was done using the Cox's regression model and logistic multivariate regression model. The statistical analysis was performed using the

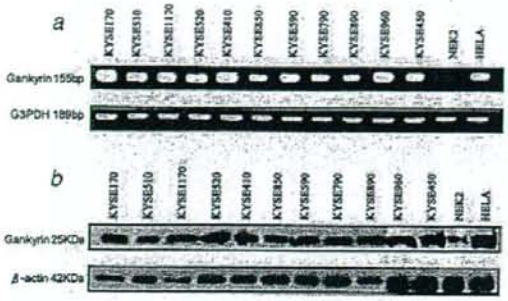


FIGURE 1 – (*a*) RT-PCR analysis in 11 ESCC cell lines (KYSE series). The mRNA of HeLa cells was included as a positive control in the same analysis. (*b*) Western blot analysis in 11 ESCC cell lines (KYSE series). The HeLa cell line was included as a positive control for gankyrin expression. Western blot was probed with anti-gankyrin Ab.

software StatView 4.5 (Abacus Concept, Berkley, CA). A p-value of <0.05 indicated statistical significance.

Construction of gankyrin shRNA expression plasmid vector

To construct a stable vector for gankyrin-shRNA, we used the BLOCK-iT™ U6 RNAi Entry Vector Kit (Invitrogen, Carisbad, CA), and destination vector kit pBLOCK-iT™ 6-DEST Gateway (Invitrogen, Carisbad, CA) that contains 2 recombination sites, *att*R1 and *att*R2. A chemically synthesized oligonucleotide encoding a shRNA for gankyrin and including a loop motif, was inserted into the downstream of the U6 promoter of the plasmid using T4 DNA Ligase (Invitrogen) and was cloned. The sequences of the oligonucleotide targeted to gankyrin are 5′-GCTGTACTCCCT TACATTATG-3′ corresponding to positions 415–436 within the gankyrin mRNA sequence. To confirm the results, we designed a second sequence 5′-GCCTGGGTTTAATACTCAAGA-3′ corresponding to positions 742–763 within the gankyrin mRNA sequence. For the negative control, an empty vector pBLOCK-iT™6-DEST was used. pENTER™-gus and pBLOCK-iT™ 6-GW/U6-laminin^{shRNA} were used as positive controls.

Transfections

The KYSE 170 ESCC cell line and HeLa cell line were stably transfected with the gankyrin-shRNA vector or the empty vector using Lipofectamine 2000 reagent (Invitrogen, Carisbad, CA). Clones were selected using 30 μg/ml of blasticidin (Nacalai Tesque, Kyoto, Japan).

Cell growth assay

Cells were plated in 6-cm dishes (2×10^4 cells/dish) at Day 0 and incubated for 24 hr for sufficient cell growth. Cells were harvested with trypsin/EDTA every 48 hr for 5 days, and counted with cell counter (Coulter Z1, Beckman-coulter, Fullerton, CA). To examine the effect that the suppression of gankyrin expression has on cell growth, we compared it with the control culture.

Flow cytometry for cell cycle analysis

A flow cytrometric analysis of DNA content was done to evaluate the cell cycle phase distribution. The cells were harvested at 70% confluent dish, fixed in 70% ethanol at −20°C and treated with RNase A/PBS. Then, the cells were stained with propidium iodide (50 μg/ml). The evaluation of cell cycle was done using a FACScan flow cytometer (Becton Dickinson, Mountain View, CA) and Cell Quest software (Becton Dickinson). Each experiment was repeated 3 times.

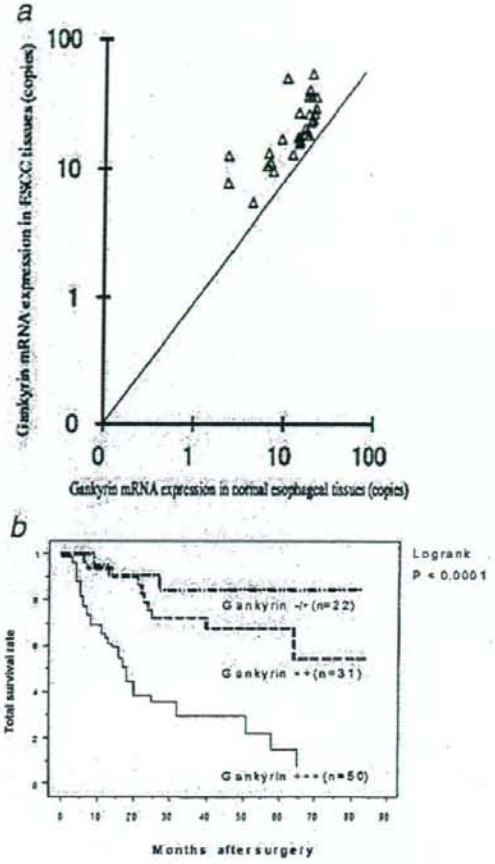


FIGURE 2 – (*a*) Distribution of T/N values. Gankyrin mRNA expression in esophageal tissue of patients with ESCC, measured by Real time PCR method, using the LightCycler FastStart DNA Master SYBR Green I system. Gankyrin mRNA expression in tumoral (*Y*-axis) and normal esophageal tissue (*X*-axis) in 30 patients with ESCC. The figure shows that the overexpression of gankyrin was higher in the tumoral tissues compared with the corresponding normal esophageal tissue in all 30 patients with ESCC. The expression of gankyrin was measured in number of copies, and the average ratio between the expression of gankyrin in the tumoral part and the normal tissue (T/N ratio) was 1.9. The straight-line divides tumoral part from the normal esophageal part. (*b*) The cumulative survival rate of the patients in relation to immunohistochemical gankyrin expression. The survival rate analysis using the Kaplan–Meier method revealed that the prognosis of patients with tumors immunohistochemically positive +++ (Group C) for gankyrin was lower than that of patients with tumors with a −/+ (Group A) or positive ++ (Group B) gankyrin expression ($p < 0.0001$).

Transwell migration and matrigel invasion assay

The migration was determined by a micropore chamber assay and the invasion was determined by a matrigel invasion chamber assay. Cells (2.5×10^4) were seeded onto the top chamber of a 24-well micropore polycarbonate membrane filter with 8 μm pores (Becton Dickinson Labware, Lincoln Park, NJ), and the bottom chamber was filled with Ham's F12/RPMI 1640 containing 10% of FBS as a chemo attractant. After 20 hr of incubation in a 5% CO_2 humidified incubator at 37°C, the membranes were fixed and

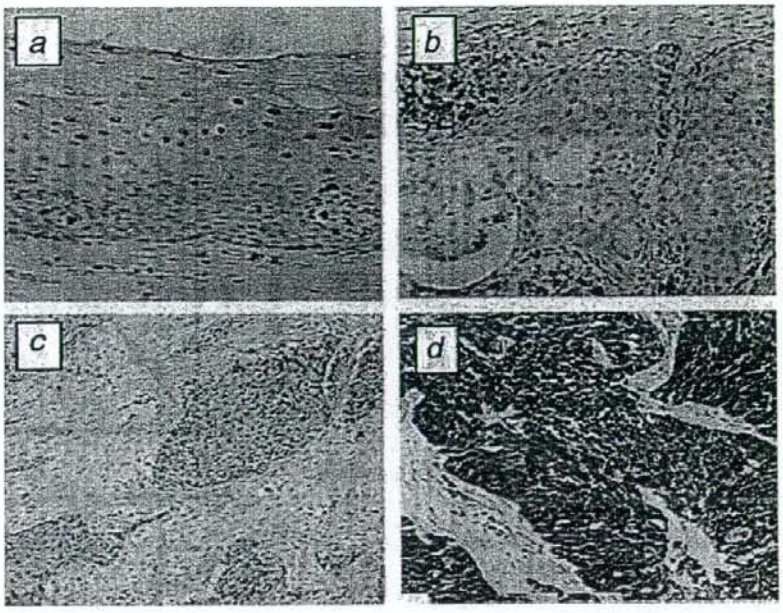


FIGURE 3 – Immunohistochemical staining for gankyrin. (*a*) Normal esophageal epithelium. Gankyrin is expressed in the horny layer, weakly in the nucleus and cytoplasm of the granular layers, and no staining can be found in the spindle and basal layer. ×400. (*b*) Negative or + expression of gankyrin. No staining in the nucleus and cytoplasm. ×200. (*c*) Positive ++ staining for gankyrin, 50% of the nucleus stained and the cytoplasm showed weak expression. ×200. (*d*) Positive +++ staining for gankyrin expression. Strong nuclear and cytoplasm staining ×200.

stained with DiffQuik reagent (International reagents, Kobe, Japan). Migration was quantified by counting the migrated cells in the filter and invasion was quantified by counting under a light microscope all the cells that had invaded through the membrane.

Tumor formation assay in nude mice

Suspensions of 1×10^6 KYSE 170 parental and derivated cells (gankyrin shRNA-transfected cells: 170 G1, 170 G2 and empty vector-transfected cells: 170 mock) in PBS (60 μl) were injected subcutaneously into the right flanks of 5-week-old male BALB/c nu/nu mice (Japan SLC, Shizuoka, Japan) at Day 0. The inoculation was conducted in 8 mice respectively, and the tumor growth was estimated from the average volume tumors by the formula: $1/2 \times L^2 \times W$ (L = length and W = width of the tumor). At 30 days after injection, all mice were sacrificed and the subcutaneous tumors were resected and fixed in 10% formaldehyde/PBS. Then, tumors were paraffin-embedded and stained with H&E and gankyrin. All the animal experiments were performed in accordance with institutional guidelines.

Results

Gankyrin expression in ESCC cell lines

We examined gankyrin expression in 11 ESCC cell lines (KYSE series). RT-PCR showed that gankyrin mRNA in KYSE cell lines was at the same level as the control HeLa cell (Fig. 1*a*). As shown in Figure 1*b*, gankyrin protein expression in all cell lines tested were determined by western blot. There was no relevant difference in the gankyrin expression among KYSE cell lines.

Real-time PCR

The expression of gankyrin was higher in the cancerous tissue for all 30 patients (Fig. 2*a*). The distribution of gankyrin mRNA expression among 30 patients was as follows: 19 tumors had a T/N ratio <1.9 and 11 tumors had a ratio >1.9. The average T/N ratio was 1.9, and the standard deviation was 1.08. The average gankyrin mRNA expression (expressed in number of copies) in the tumoral part was 31 with a median of 19 and a standard deviation of 12.44. In the corresponding normal esophageal epithelium, the average gankyrin mRNA expression was 12.82 with a median of 14 and a standard deviation of 6.02.

Immunohistochemical staining and immunofluorescent staining

After performing Real-time PCR using the surgical specimens, we can assume that the up-regulation of gankyrin may have occurred. We classified 103 ESCC tumors in 3 groups according to the percentage of nuclear staining among the cells (Fig. 3): (*i*) 22 tumors with − or + expression (Group A) (21.4%); (*ii*) 31 tumors with positive expression ++ (Group B) (30%); (*iii*) 50 tumors with positive expression +++ (Group C) (48.5%). The staining pattern of gankyrin protein expression in tumors was examined by the IHC study. Gankyrin protein expression was generally increased in comparison to the normal epithelium. The tumors often showed strong nuclear and cytoplasm staining compared with the normal epithelium. The immunofluorescent staining using KYSE 850 cells and ESCC samples showed that gankyrin was expressed in the cytoplasm and the nucleus (supplementary data 1).

Correlation between gankyrin protein overexpression and clinicopathologic findings

Gankyrin positive expression was correlated with the extent of the primary tumor, lymph node metastasis, distant lymph node metastasis and stage ($p = 0.007$, $p = 0.0004$, $p = 0.017$ and $p = 0.0002$, respectively). However, there was no correlation with age, gender, histological grade or location (Table I). The logistic regression model analysis showed that lymph node metastasis,

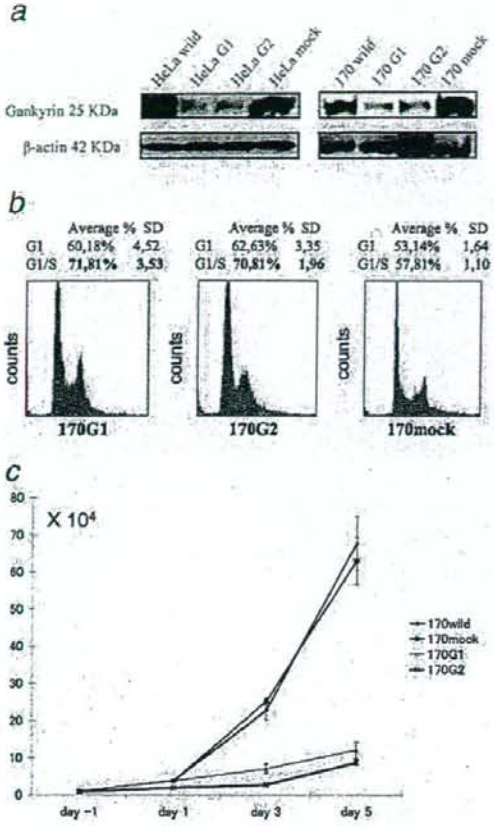


FIGURE 4 – Vector-based shRNA for gankyrin down-regulation in ESCC cell line. (*a*) Western blot analysis for gankyrin expression in parental KYSE170 cells and stable transfectant cells with an empty vector (170 mock) or gankyrin-shRNA vector (170G1 and 170G2); the HeLa cells were used as a control. (*b*) Cell cycle assay performed by flow cytometry for 170 mock, 170G1 and G2. The ratios of cell populations in G1/S are indicated as the average of 3 results, with the corresponding standard deviation (SD). (*c*) *In vitro* growth assay in KYSE170 cells and stable transfectant cells with an empty vector (170 mock) or gankyrin-shRNA vector (170G1 and 170G2). The cells were counted every 2 days and the experiment was repeated 3 times.

distant lymph node metastasis and the extent of the primary tumor correlated with the gankyrin protein overexpression ($p = 0.007$, $p = 0.02$ and $p = 0.05$, respectively, and Odds ratio 3.31, 3.71 and 2.17, respectively).

Association between gankyrin protein expression and patient prognosis

Overall survival analysis using the Kaplan–Meyer method showed that the prognosis of the patients in Groups B and C was significantly lower than that of the patients in Group A ($p = 0.0001$; Fig. 2*b*). The multivariate analysis revealed that the extent of the primary tumor, distant lymph node metastasis and gankyrin (+++) status ($p = 0.02$, $p = 0.01$ and $p = 0.0001$, respectively) were independent poor prognostic factors (Risk ratio 2.09, 2.26 and 3.97, respectively).

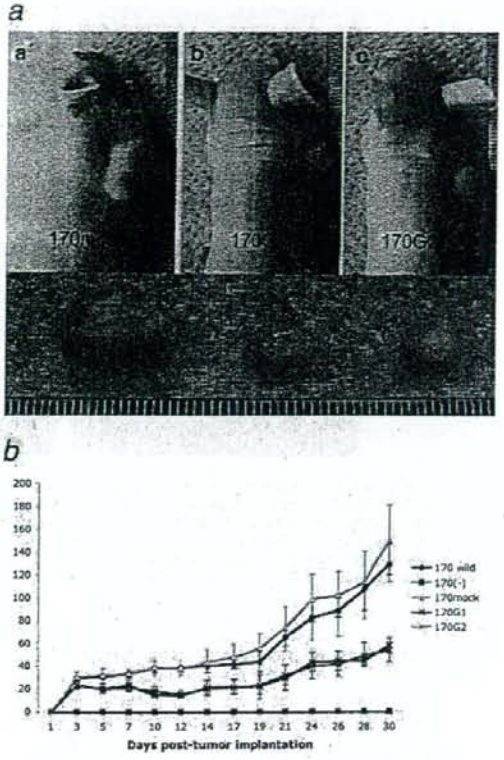


FIGURE 5 – The effect of gankyrin down-regulation on tumor formation model. (*a*) Tumor subcutaneous implantation after 30 day of inoculation. (*a'*) The empty vector-transfected cells (170 mock) were transfected and injected in the right flank. (*b'*,*c'*) Gankyrin shRNA transfected cells (170G1 and 170G2) were injected subcutaneously in the right flank. (*d'*) Comparative tumor size of each group (170 mock, 170G1 and 170G2). (*b*) The tumor growth curves of the mice injected with parental cells (170 W), empty vector-transfected cells (170 mock) and gankyrin shRNA-transfected cells (170G1 and 170G2). The assay was performed using 8 mice/group.

Effect of gankyrin down-regulation on proliferation, motility and invasion in vitro

To evaluate the role of gankyrin overexpression in ESCC cells, we established stable clones by transfection of vector-based shRNA for gankyrin in KYSE 170 cells and HeLa cells. Gankyrin expression level was efficiently reduced by 80–85%, and the transfection of the empty vector shRNA (170 mock) did not reduce gankyrin expression level (Fig. 4*a*). To elucidate in detail the growth suppression caused by gankyrin down-regulation, we examined the cell cycle changes in 170G1, 170G2 and 170 mock by using flow cytometry. Cell populations in G1/S phase were significantly larger in 170G1 (71.81%) and 170G2 (70.81%) than in the 170 mock (57.1%) (Fig. 4*b*). We investigated the effect of gankyrin down-regulation on proliferation in 1 ESCC cell line (KYSE 170). With the suppression of gankyrin, the cell growth was inhibited by 80% in 170G1 and by 86% in 170G2, compared with the empty vector shRNA transfected cell line (170 mock) and wild cell line (170 W). ($p < 0.01$ and $p < 0.01$, respectively) (Fig. 4*c*). Next, we examined the effect of gankyrin down-regulation on cell motility and invasion *in vitro*. Cell migration was reduced by 26% in 170G1 and 27% in 170G2 ($p < 0.0001$ and $p < 0.001$,

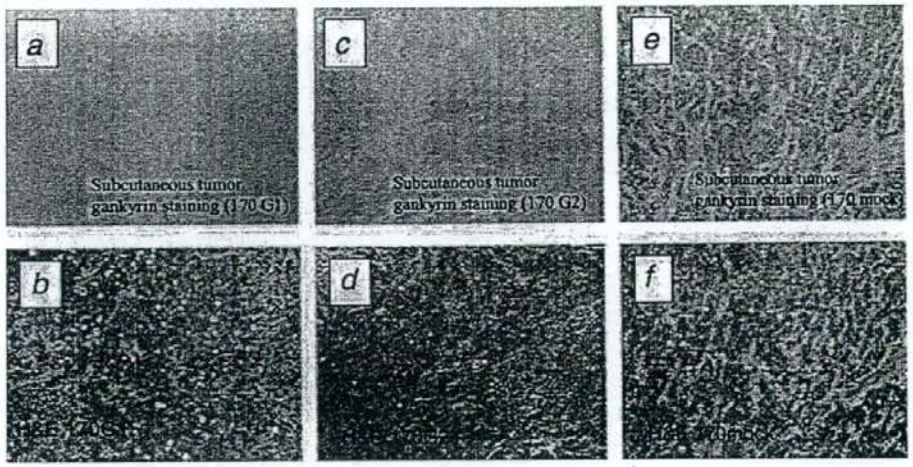


FIGURE 6 – (*a*) Gankyrin immunohistochemical staining of subcutaneous tumors after 30 days postinoculation of gankyrin shRNA transfected cells (170G1). (*b*) H&E staining of subcutaneous tumors after 30 days postinoculation of gankyrin shRNA transfected cells (170G1). (*c*) Gankyrin immunohistochemical staining of subcutaneous tumors after 30 days postinoculation of gankyrin shRNA transfected cells (170G2). (*d*) H&E staining of subcutaneous tumors after 30 days postinoculation of gankyrin shRNA transfected cells (170G2). (*e*) Gankyrin immunohistochemical staining of subcutaneous tumors after 30 days postinoculation of empty-vector transfected cells (170 mock). (*f*) H&E staining of subcutaneous tumors after 30 days postinoculation of empty-vector transfected cells (170 mock).

respectively). Cell invasion was suppressed by 27% in 170G1 and 27% in 170G2 ($p < 0.003$ and $p < 0.002$, respectively).

Effect of gankyrin down-regulation on tumor formation in vivo

To investigate the possible effect of gankyrin-shRNA on tumor formation *in vivo*, we performed a subcutaneous tumor formation assay in nude mice. The tumor formations caused by gankyrin-shRNA transfected cells (170G1 and 170G2) were evidently smaller than those formed by 170 mock (Fig. 5*a*). We evaluated tumor volume and weight after 30 days of inoculation (Fig. 5*b*). The average tumor volume was significantly suppressed in 170G1 and 170G2 (61.4% and 64.2% respectively), compared with 170 mock ($p < 0.0001$ and $p < 0.0001$ respectively). The average tumor weight was inhibited, as well, in 170G1 and 170G2 by 60 and 63%, respectively ($p < 0.0003$ and $p < 0.0002$ respectively). Finally, all tumors were stained with H&E and gankyrin, and gankyrin protein expression was reduced to a minimum in tumors of inoculated mice with gankyrin shRNA transfected cells, but not in tumors of empty vector transfected cells (Fig. 6).

Discussion

In the present study, the results for IHC showed that protein overexpression of gankyrin in ESCC tumors was significantly associated with poor prognosis. The strong overexpression of gankyrin was an independent prognostic factor, and suggests that gankyrin overexpression may be a useful marker for patient prognosis in ESCC. Gankyrin staining intensity was always increased in the tumoral part compared with the corresponding normal epithelium part in ESCC specimens.

In all the specimens studied by Real-time PCR, the gankyrin mRNA expression level in the tumoral part was higher than that in the corresponding normal epithelium. However, we found that its up-regulation in the tumoral part was proportional to the basal amount of mRNA in the corresponding normal tissue. This data supported our idea that gankyrin up-regulation is an abnormality in ESCC, and that it might be overexpressed in the tumor samples.

All ESCC cell lines (KYSE series) demonstrated gankyrin overexpression compared with the control HeLa cells. Obviously, 1 normal esophageal cell line is by no means sufficient to compare the gankyrin expression between tumoral cells and normal cells, but our preliminary comparison between 11 KYSE cell lines with the normal esophageal epithelial cell line (NEK2) showed that the expression of gankyrin mRNA in KYSE cell lines was higher than that in the normal esophageal epithelial cell line. However, further experiments should be carried out with other normal cell lines to verify the results.

Our clinicopathological analysis by IHC showed that gankyrin overexpression was significantly associated with poor patient survival. Moreover, gankyrin was the strongest prognostic factor of the patients among 6 major parameters. These results suggest that gankyrin could be a useful prognostic marker for patients with ESCC. Gankyrin overexpression in the tumors was shown to be correlated with the extent of the primary tumor, regional lymph node metastasis, distant lymph node metastasis and stage. These findings suggest that gankyrin might be important during the development of malignancy potential in ESCC, and may play a role in its progression.

Several studies have demonstrated that the p53 gene is one of the most mutated genes in a variety of human cancers,[22–26] and its involving with gankyrin has recently been reported.[27,28] Previously, we have published articles classifying the p53 gene mutations in ESCC[29,30] and we presumed that gankyrin might be overexpressed in ESCC according to the types of p53 gene mutations discovered.

Overexpression of gankyrin has been reported in hepatocellular carcinoma (HCC).[12,31–34] However, so far there have been no reports on the role of gankyrin in either esophageal cancer or in other gastrointestinal cancers. This study is the first report on gankyrin expression in ESCC.

Using shRNA to target gankyrin, it has been demonstrated that a significant down-regulation of gankyrin inhibits the expression of cell cycle molecules and induces apoptosis.[35] We were able to suppress gankyrin expression using a vector-based shRNA for 2 different target sequences in 1 ESCC cell line, obtaining similar results in both examinations *in vitro* and *in vivo*.

It was previously reported that gankyrin has a physiological and physiopathological role in the cell-cycle progression by increasing the phosphorylation and degradation of Rb protein, and binding with CDK4.[1,36] We found the coexpression of gankyrin and CDK4 by immunofluorescent staining in ESCC samples embedded in paraffin and ESCC cell lines (KYSE series) (supplementary data 1). However, the role of gankyrin in cell cycle progression in ESCC is still unclear. Moreover, based on the studies on apoptotic growth factors that have suggested that G1/S stages of the cell cycle may be apoptotic insensitive stages[37] and that S6 ATPase-gankyrin-Rb-E2F1 interactions may control E2F1 stability in relation to the G1/S transition,[38] we demonstrated that the down-regulation of gankyrin expression inhibited proliferation in an ESCC cell line, inducing an accumulation of cells in the G1/S phase *in vitro*. Furthermore, we clearly demonstrated that the down-regulation of gankyrin inhibited tumor growth *in vivo*. Our findings suggested that gankyrin has a crucial role in the proliferation of cancer cells and it is an excellent target for halting proliferation in ESCC.

It has been reported that MMP expressions are high in esophageal cancer,[39,40] and appear to play an important role in the ability of cells to migrate and invade the extracellular matrix that may be associated with tumorigenicity.[41,42] Gankyrin down-regulation partially decreased the motility and invasiveness in an ESCC cell line. Kan *et al.* demonstrated that gene expression of MMPs was remarkably high in esophageal cancer tissues and very low in cancer cell lines.[43] We found that in mock cells the expression of MMP-2 was weak, whereas in gankyrin down-regulated cells it was virtually inexistent (supplementary Fig. 2). At the same time, the MMP-2 expression of both mock and gankyrin down-regulated cells in mice tumor samples was high (supplementary Fig. 3). These results might suggest that gankyrin may mediate tumor cell migration and invasion *in vitro* by regulating the expression of MMP-2. On the other hand, the expression of MMP-2 *in vivo* might be enhanced by the interaction with stromal cells. The mechanisms by which gankyrin increases motility and invasive potential are not fully understood.

In conclusion, gankyrin overexpression in ESCC tumor is associated with poor survival of patients, distant lymph node metastasis and tumor staging. Down-regulation of gankyrin in an ESCC cell line using shRNA vector decreased cell proliferation, motility and invasiveness *in vitro*, and tumor formation *in vivo*. These findings suggest that gankyrin may play an important role in ESCC tumor progression and could be a potentially important therapeutic gene target in ESCC.

Acknowledgements

The authors cordially thank Ms. Sakiko Shimada for her kind assistance in immunohistochemical staining and providing and culturing the ESCC cell lines; and Mr. Fabricio Loaiza, Ms. Kumi Kodama, Ms. Takako Murai, Ms. Mayumi Kawashima and Ms. Akane Iwase for their technical assistance.

References

1. Pisani P, Parkin DM, Bray F, Ferlay J. Estimates of the worldwide mortality from 25 cancers in 1990. Int J Cancer 1999;83:870–3.
2. Visser BC, Venook AP, Patti MG. Adjuvant and neoadjuvant therapy for esophageal cancer: a critical reappraisal. Surg Oncol 2003;12:1–7.
3. Kan T, Shimada Y, Sato F, Ito T, Kondo K, Watanabe G, Maeda M, Yamasaki S, Meltzer SJ, Imamura M. Prediction of lymph node metastasis with use of artificial neural network based on gene expression profiles in esophageal squamous cell carcinoma. Ann Surg Oncol 2004;11:1070–8.
4. Kanda Y, Nishiyama Y, Shimada Y, Imamura M, Nomura H, Hiai H, Fukumoto M. Analysis of gene amplification and overexpression in human esophageal-carcinoma cell lines. Int J Cancer 1994;58:291–7.
5. Tanaka H, Shimada Y, Imamura M, Shibagaki I, Ishizaki K. Multiple types of aberrations in the p16 (INK4a) and the p15(INK4b) genes in 30 esophageal squamous-cell-carcinoma cell lines. Int J Cancer 1997; 70:437–42.
6. Yuan C, Li J, Mahajan A, Poi MJ, Byeon IJ, Tsai MD. Solution structure of the human oncogenic protein gankyrin containing seven ankyrin repeats and analysis of its structure-function relationship. Biochemistry 2004;43:12152–61.
7. Manjasetty BA, Quedenau C, Sievert V, Bussow K, Niesen F, Delbruck H, Heinemann U. X-ray structure of human gankyrin, the product of a gene linked to hepatocellular carcinoma. Proteins 2004;55: 214–17.
8. Krzywda S, Brzozowski AM, Al-Safty R, Welchman R, Mee M, Dawson S, Fujita J, Higashitsuji H, Mayer RJ, Wilkinson AJ. Crystallization of gankyrin, an oncoprotein that interacts with CDK4 and the S6b (rpt3) ATPase of the 19S regulator of the 26S proteasome. Acta Crystallogr D Biol Crystallogr 2003;59:1294–5.
9. Iwai A, Marusawa H, Kiuchi T, Higashitsuhi H, Fujita J, Tanaka K, Chiba T. Role of a novel oncogenic protein, gankyrin, in hepatocyte proliferation. J Gastroenterol 2003;38:751–8.
10. Hori T, Kato S, Saeki M, DeMartino GN, Slaughter CA, Takeuchi J, Tohe A, Tanaka K. cDNA cloning and functional analysis of p28 (Nas6p) and p40.5 (Nas7p), two novel regulatory subunits of the 26S proteasome. Gene 1998;216:113–22.
11. Li J, Tsai MD. Novel insights into the INK4-CDK4/6-Rb pathway: counter action of gankyrin against INK4 proteins regulates the CDK4-mediated phosphorylation of Rb. Biochemistry 2002;41: 3977–83.
12. Higashitsuji H, Higashitsuji H, Itoh K, Sakurai T, Nagao T, Sumitomo Y, Masuda T, Dawson S, Shimada Y, Mayer RJ, Fujita J. The oncoprotein gankyrin binds to MDM2/HDM2, enhancing ubiquitylation and degradation of p53. Cancer Cell 2005;8:75–87.
13. Bernards R, Brummelkamp T, Beijersbergen R. shRNA libraries and their use in cancer genetics. Nat Methods 2006;3:701–6.
14. McAnuff MA, Rettig GR, Rice KG. Potency of siRNA mediated knockdown in vivo. J Pharm Sci 2007;22:1–9.
15. Sobin LH, Fleming ID. TNM classification of malignant tumors, fifth edition (1997). Cancer 1997;80:1803–4.
16. Sobin LH. TNM, sixth edition: new developments in general concepts and rules. Semin Surg Oncol 2003;21:19–22.
17. Fleming ID, Cooper JS, Henson DE, eds. American Joint Committee on Cancer. AJCC cancer staging manual, 5th edn. Philadelphia: J.B. Lippincott, 1997.
18. Shimada Y, Imamura M, Wagata T, Yamaguchi N, Tobe T. Characterization of 21 newly established esophageal cancer cell lines. Cancer 1992;69:277–84.
19. Okumura T, Shimada Y, Imamura M, Yasumoto S. Neurotrophin receptor p75(NTR) characterizes human esophageal keratinocyte stem cells in vitro. Oncogene 2003;22:4017–26.
20. Chomczynski P, Sacchi N. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Anal Biochem 1987;162:156–9.
21. Chomczynski P. A reagent for the single-step simultaneous isolation of RNA, DNA and proteins from cell and tissue samples. Biotechniques 1993;15:532–4, 536–7.
22. Pipas J, Levine A. Role of T antigen interactions with p53 in tumorigenesis. Semin Cancer Biol 2001;1:23–30.
23. Vogelstein B, Lane D, Levine AJ. Surfing the p53 network. Nature 2000;408:307–10.
24. Lozano G, Zambetti G. What have animal models taught us about the p53 pathway? J Pathol 2005;205:206–20.
25. Mendrysa SM, McElwee MK, Michalowski J, O'Leary KA, Young KM, Perry ME. mdm2 Is critical for inhibition of p53 during lymphopoiesis and the response to ionizing irradiation. Mol Cell Biol 2003;23:462–72.
26. Bond GL, Hu W, Bond EE, Robins H, Lutzker SG, Arva NC, Bargonetti J, Bartel F, Taubert H, Wuerl P, Onel K, Yip L, et al. A single nucleotide polymorphism in the MDM2 promoter attenuates the p53 tumor suppressor pathway and accelerates tumor formation in humans. Cell 2004;119:591–602.
27. Higashitsuji H, Liu Y, Mayer RJ, Fujita J. The oncoprotein gankyrin negatively regulates both p53 and RB by enhancing proteasomal degradation. Cell Cycle 2005;4:1335–7.
28. Lozano G, Zambetti GP. Gankyrin: an intriguing name for a novel regulator of p53 and RB. Cancer Cell 2005;8:3–4.
29. Tanaka H, Shibagaki I, Shimada Y. Characterization of p53 gene mutations in esophageal squamous cell carcinoma cell lines: increased frequency and different spectrum of mutations from primary tumors. Int J Cancer 1996;65:372–6.
30. Wagata T, Shibagaki I, Imamura M, Shimada Y, Toguchida J, Yandell DW, Ikenaga M, Tobe T, Ishizaki K. Loss of 17p mutation of the p53 gene, and overexpression of p53 protein in esophageal squamous cell carcinomas. Cancer Res 1993;53:846–50.
31. Nagao T, Higashitsuji H, Nonoguchi K, Sakurai T, Dawson S, Mayer RJ, Itoh K, Fujita J. MAGE-A4 interacts with the liver oncoprotein

gankyrin and suppresses its tumorigenic activity. J Biol Chem 2003;278:10668–74.
32. Sakurai T, Itoh K, Higashitsuji H, Nagao T, Nonoguchi K, Chiba T, Fujita J. A cleaved form of MAGE-A4 binds to Miz-1 and induces sapoptosis in human cells. J Biol Chem 2004;279:15505–14.
33. Krzywda S, Brzozowski AM, Higashitsuji H, Fujita J, Welchman R, Dawson S, Mayer RJ, Wilkinson AJ. The crystal structure of gankyrin, an oncoprotein found in complexes with cyclin-dependent kinase 4 a 19 S proteasomal ATPase regulator, and the tumor suppressors Rb and p53. J Biol Chem 2004;279:1541–5.
34. Padmanabhan B, Adachi N, Kataoka K, Horikoshi M. Crystal structure of the homolog of the oncoprotein gankyrin, an interactor of Rb and CDK4/6. J Biol Chem 2004;279:1546–52.
35. Li H, Fu X, Chen Y, Hong Y, Tan Y, Cao H, Wu M, Wang H. Use of adenovirus-delivered siRNA to target oncoprotein $p28^{gank}$ in hepatocellular carcinoma. Gastroenterology 2005;128:2029–41.
36. Higashitsuji H, Itoh K, Nagao T, Dawson S, Nonoguchi K, Kido T, Mayer RJ, Arii S, Fujita J. Reduced stability of retinoblastoma protein by gankyrin, an oncogenic ankyrin-repeat protein overexpressed in hepatomas. Nat Med 2000;6:96–9.
37. Collins N, Reginato M, Paulus JK, Sgroi DC, Labaer J, Brugge JS. G1/S cell cycle arrest provides anoikis resistance through erk-mediated bim suppression. Mol Cell Biol 2005;25:5282–91.
38. Dawson S, Apcher S, Mee M, Higashitsuji H, Baker R, Uhle S, Dubiel W, Fujita J, Mayer RJ. Gankyrin is an ankyrin-repeat oncoprotein that interacts with CDK4 kinase and the S6 ATPase of the 26 S proteasome. J Biol Chem 2002;277:10893–902.
39. Koyama H, Iwata H, Kuwabara Y, Iwase H, Kobayashi S, Fujii Y. Gelatinolytic activity of matrix metalloproteinase-2 and -9 in esophageal carcinoma: a study using in situ zymography. Eur J Cancer 2000;36:2164–70.
40. Duncan ME, Richardson JP, Murray GI, Melvin WT, Fothergill JE. Human matrix metalloproteinase-9: activation by limited trypsin treatment and generation of monoclonal antibodies specific for the activated form. Eur J Biochem 1998;258:37–43.
41. Ismail RS, Baldwin RL, Fang J, Browning D, Karlan BY, Gasson JC, Chang DD. Differential gene expression between normal and tumor-derived ovarian epithelial cells. Cancer Res 2000;60:6744–49.
42. Wellmann A, Thieblemont C, Pittaluga S, Sakai A, Jaffe ES, Siebert P, Raffeld M. Detection of differentially expressed genes in lymphomas using cDNA arrays: identification of clusterin as a new diagnostic marker for anaplastic large-cell lymphomas. Blood 2000;96:398–404.
43. Kan T, Shimada Y, Sato F, Maeda M, Kawabe A, Kaganoi J, Itami A, Yamasaki S, Imamura M. Gene expression profiling in human esophageal cancers using cDNA microarray. Biochem Biophys Res Commun 2001;286:792–801.

J Hum Genet (2008) 53:515–523
DOI 10.1007/s10038-008-0279-x

ORIGINAL ARTICLE

Global correlation analysis for micro-RNA and mRNA expression profiles in human cell lines

Yoshinao Ruike · Atsuhiko Ichimura · Soken Tsuchiya · Kazuharu Shimizu · Ryo Kunimoto · Yasushi Okuno · Gozoh Tsujimoto

Received: 19 February 2008 / Accepted: 29 February 2008 / Published online: 10 May 2008

Abstract Microribonucleic acids (miRNAs) are small noncoding RNAs that negatively regulate gene expression at the posttranscriptional level. Although considerable progress has been made in studying the function of miRNAs, they still remain largely unclear, mainly because of the difficulty in identifying target genes for miRNA. We performed a global analysis of both miRNAs and mRNAs expression across 16 human cell lines and extracted negatively correlated pairs of miRNA and mRNA which indicate miRNA-target relationship. The many of known-target of *miR-124a* showed negative correlation, suggesting our analysis were valid. We further extracted physically relevant miRNA-target gene pairs, applying computational target prediction algorithm with inverse correlations of miRNA and messenger RNA (mRNA) expression. Furthermore, gene-ontology-based annotation and functional enrichment analysis of the extracted miRNA-target gene pairs made it possible to indicate putative functions of miRNAs. The data collected here will be of value for further studies into the function of miRNA.

Keywords Micro-RNA · Microarray · Transcriptome · Correlation analysis · GO analysis

Electronic supplementary material The online version of this article (doi:10.1007/s10038-008-0279-x) contains supplementary material, which is available to authorized users.

Y. Ruike · A. Ichimura · R. Kunimoto · G. Tsujimoto (✉)
Department of Genomic Drug Discovery Science,
Graduate School of Pharmaceutical Sciences,
Kyoto University, 46-29 Yoshida Shimoadachi-cho Sakyo-ku,
Kyoto 606-8501, Japan
e-mail: gtsuji@pharm.kyoto-u.ac.jp

S. Tsuchiya · K. Shimizu
Department of Nanobio Drug Discovery,
Graduate School of Pharmaceutical Sciences,
Kyoto University, Kyoto, Japan

Y. Okuno
Department of Pharmacoinformatics,
Graduate School of Pharmaceutical Sciences,
Kyoto University, Kyoto, Japan

Introduction

Microribonucleic acids (miRNAs) are a class of small (approximately 22 nucleotides) noncoding RNAs that negatively regulate gene expression at the posttranscriptional level. They play profound and pervasive roles in manipulating gene expression involved in cell development, proliferation and apoptosis in various eukaryotes (Sevignani et al. 2006). Also, recent evidence indicated that aberrant miRNA expression in the pathogenesis of several human diseases (van Rooij et al. 2006; Takamizawa et al. 2004; Iorio et al. 2005), revealing that miRNA genes could be a potential target for drug discovery (Sevignani et al. 2006; Liu et al. 2007; Tsuchiya et al. 2006).

In the past few years, several hundred miRNAs were identified in animals and plants, although it is estimated that miRNAs account for ~1% of predicted genes in higher eukaryotic genomes (Griffiths-Jones 2004). Only a handful of miRNA, however, have been functionally characterized. These findings, together with the complicated expression patterns and large number of predicted targets, imply that miRNAs may regulate a broad range of physiological and developmental process.

Idnetifying targets of each miRNA is crucial for understanding the biological functions of miRNAs, as miRNA-directed regulation occurs at a posttranscriptional level via interaction with their target messenger RNA (mRNA)

to elicit degradation or translational repression of complementary mRNA targets (Behm-Ansmant et al. 2006; Chendrimada et al. 2007). Also, miRNAs are known to be only partially complementary to their targets and can induce significant degradation of many mRNA targets in animals (Bagga et al. 2005; Jing et al. 2005; Giraldez et al. 2006; Rehwinkel et al. 2006). This ambiguity makes it difficult to predict the targets of miRNAs [up to 30% of genes have been predicted to be regulated by miRNAs (Lewis et al. 2003)]. However, most of the predicted miRNA-target pairs remain to be biologically verified, because a simple, high throughput method to biologically validate miRNA-targets does not exist. Several recent studies revealed spatial or temporal avoidance of miRNA coexpression with target genes (Farh et al. 2005; Sood et al. 2006). Thus, genes preferentially expressed at the same time and place as an miRNA have evolved to selectively avoid sites matching the miRNA (Farh et al. 2005). These findings may support the negative correlation between miRNA and their target mRNA expression level as a whole. Based on the idea that negative correlation between miRNA and mRNA expression levels may reflect miRNA-target relationship, we performed a global analysis of both miRNA and mRNA expression across 16 human cell lines. Global correlation analysis of the expression profiles collected revealed a useful approach to identify the miRNA-target interactions.

Materials and methods

Cell lines and RNA purification

We used 16 well-studied, various organ-derived cell lines in this study: human lung carcinoma A549 (Giard et al. 1973), fibrosarcoma HT1080 (Rasheed et al. 1974), cervix carcinoma Henrietta Lacks (HeLa) (Scherer et al. 1953), cervix carcinoma HeLaS3 (subclone of HeLa) (Puck et al. 1956), hepatocellular carcinoma Huh7 (Nakabayashi et al. 1985), breast adenocarcinoma MCF7 (Soule et al. 1973), breast adenocarcinoma MDAMB231 (Cailleau et al. 1974), embryonal kidney HEK293T (subclone of HEK293) (DuBridge et al. 1987), colon adenocarcinoma HT29 (Fogh et al. 1977), hepatocellular carcinoma HepG2 (Aden et al. 1979), neuroblastoma SKNMC (Biedler et al. 1973), colon adenocarcinoma $Caco_2$ (Fogh et al. 1977), embryonal kidney HEK293 (Graham et al. 1977), and colon carcinoma HCT116 (Brattain et al. 1981). Cells were cultured in Dulbecco's modified Eagle's medium (DMEM) containing 10% fetal bovine serum (FBS). Human T-cell lukemia Jurkat (Schneider et al. 1977) and chronic myeloid leukemia K562 (Lozzio and Lozzio 1973) cell lines were cultured in Roswell Park Memorial Institute (RPMI) medium containing 10% FBS. Exponentially growing cells were harvested, and total RNA was collected by standard procedures using ISOGEN (Nippon Gene) and chloroform and precipitated with isopropanol. Integrity and purity of RNA was verified with an Agilent Bioanalyzer.

Quantification of miRNAs using stem-loop real-time PCR

Expression of 148 miRNAs was measured by reverse transcription followed by the real-time polymerase chain reaction (RT-PCR) assay as previously described (Gaur et al. 2007). This method uses stem-loop primers for reverse transcription followed by RT-PCR (TaqMan MicroRNA Assays; Applied Biosystems). ABI PRISM™ 7700 Sequence Detector was used to detect amplification. This stem-loop RT-PCR method detects specifically mature, but not precursor, miRNA. RNA input was normalized by measuring U6 expression using the following TaqMan probes.

5′FAM-CCCCTGCGCAAGGA-MGB3′, forward primer: 5′-TGGAACGATACAGAGAAGATTAGCA-3′, reverse primer: 5′-AACGCTTCACGAATTTGCGT-3′. Expression data of a few miRNA was confirmed to be well correlated to those obtained by Northern analysis (data not shown).

Microarray experiments

Genome-wide mRNA expression profiles of 16 human cell lines were obtained by microarray analysis with the Affymetrix GeneChip Human Genome U133 Plus 2.0 Array, according to the manufacturer's instructions. Briefly, double-stranded complementary DNA (cDNA) was synthesized from total RNA. An in vitro transcription reaction was then carried out to produce biotin-labeled complementary RNA (cRNA) from the cDNA. The cRNA was then fragmented and used for hybridization. The hybridized probe array was subsequently stained and scanned by a Genechip Scanner 3000. We used the robust multiarray analysis (RMA) expression measure that represents the log transform of (background corrected and normalized) intensities of the GeneChips (Gautier et al. 2004). RMA measures were computed using the R package program (http://www.bioconductor.org). The probe sets with the lowest maximal expression across the samples in the data set (30%) were removed.

Data analysis

Experimentally normalized ΔCt values for the miRNA profiles, or normalized microarray data sets for mRNA profiles, were used to classify cells by agglomerative hierarchical clustering based on Euclidean distances between data sets. The ΔCt values for the miRNA profiles

were used to evaluate the 16 cell lines by agglomerative hierarchical clustering using average linkage and correlation similarity and verified for significance by multiscale bootstrap resampling analysis (Suzuki and Shimodaira 2006). Classical Pearson's correlation tests were performed to verify the relationships between expression profiles of miRNA and mRNA. The significance of each correlation was assessed by assuming that the distribution of correlations under the null hypothesis of no correlation follows a t distribution with $n-2$ degrees of freedom, where n is the number of measurements in the expression profile.

Target prediction and GO term enrichment analysis

Potential targets for miRNAs were predicted among 3′ untranslated region (UTR) sequences of inversely correlated target transcripts using the miRanda algorithm, which is associated with the Sanger MIRBASE (Enright et al. 2003). algorithm parameters were set as follows: score threshold at 50, energy threshold at −20 kcal/mol, scaling parameter at 4, gap-open penalty at −2, and gap-extend penalty at −8. The significance of enrichment of a list of target genes with genes belonging to a gene ontology (GO) group were scored using weight algorithm (Ashburner et al. 2000; Alexa et al. 2006). The algorithm for GO group scoring was implemented in the R programming language. The results were obtained using R version 2.5.0 and the libraries provided by the Bioconductor project, version 1.14 (Gentleman et al. 2004).

Results

Micro-RNA expression profiles in 16 human cell lines

We performed quantitative measurement of 155 kinds of mature human miRNAs in 16 human cell lines. The probe sets for miRNA with minimal Ct values >35 in all cell lines were excluded from the following analysis. Seven miRNAs were thus excluded, and the remaining 148 different mature miRNAs were used for further analysis (expression data in a text format is provided in Table S1).

To characterize the expression patterns of miRNAs, hierarchical clustering was performed using normalized ΔCt values (Fig. 1). The most prominent feature of the clustered data was that many miRNAs displayed similar expression pattern among the majority of samples, although some miRNAs displayed very specific expression patterns. Hierarchical clustering of 16 cell lines resulted in a dendrogram, which contained small clusters reflecting common tissue origin (e.g., HepG2 and Huh7) or subclones (e.g., HeLa and HeLaS3). We also found some cell lines showed characteristic expression patterns independent of their tissue origins. In particular, $Caco_2$ had a very different expression pattern from HT-29 or HCT-116, although all of them originated from colon carcinoma. Thus, some miRNAs were specifically expressed in $Caco_2$ (e.g., miR-372 and -373), but not in either HT-29 or HCT-116. Determining the differential expression of miRNAs from similar biological backgrounds would be of value to investigate the functions of miRNA.

To classify expression patterns of miRNAs, we employed hierarchical clustering based on correlation similarity between each miRNA (Fig. 2). This analysis revealed the relationships between miRNAs that had similar expression patterns among cell lines. To assess the robustness of these relationships, we conducted a multiscale bootstrap resampling anaysis (Suzuki and Shimodaira 2006) and verified the significance (Fig. 2). The resulting dendrogram had several statistically significant clusters of miRNAs, suggesting that their expression might be commonly regulated. Noticeably, many miRNAs that coexpressed formed genomic clusters. Frequent coexpression between neighboring miRNAs that formed genomic clusters were previously reported, analyzing miRNA expression patterns across 24 normal human organs (Baskerville and Bartel 2005). To test coexpression between neighboring miRNAs, we examined the relationship between genomic distance and correlation coefficients between neighboring miRNAs. Among the miRNAs measured, miRNAs with several genomic origins, which can be indicated by the presence of multiple premature miRNA, were excluded in the following analysis. Micro-RNAs from the same chromosome and oriented in the same direction were defined as "paired". For the resulting 224 pairs (miRNA pairs are summarized in Table S2), we examined the relationship between the genomic distance of each miRNA pair and the correlation coefficient of the miRNA expression pattern (Fig. 3). Most pairs of miRNAs separated by <100 kb showed highly positive correlation.

Genome-wide gene expression analysis

Next, we obtained a global gene expression profile by employing microarray analysis using the same RNA samples used for quantifying miRNA (microarray data was deposited in Gene Expression Omnibus accession number GSE10021). The miRNAs located in the intronic region of genes are usually coordinately expressed with their host gene mRNA (Baskerville and Bartel 2005). Among 148 miRNAs measured in this study, 40 were derived from the intronic region of genes and examined for their correlation with the corresponding host genes

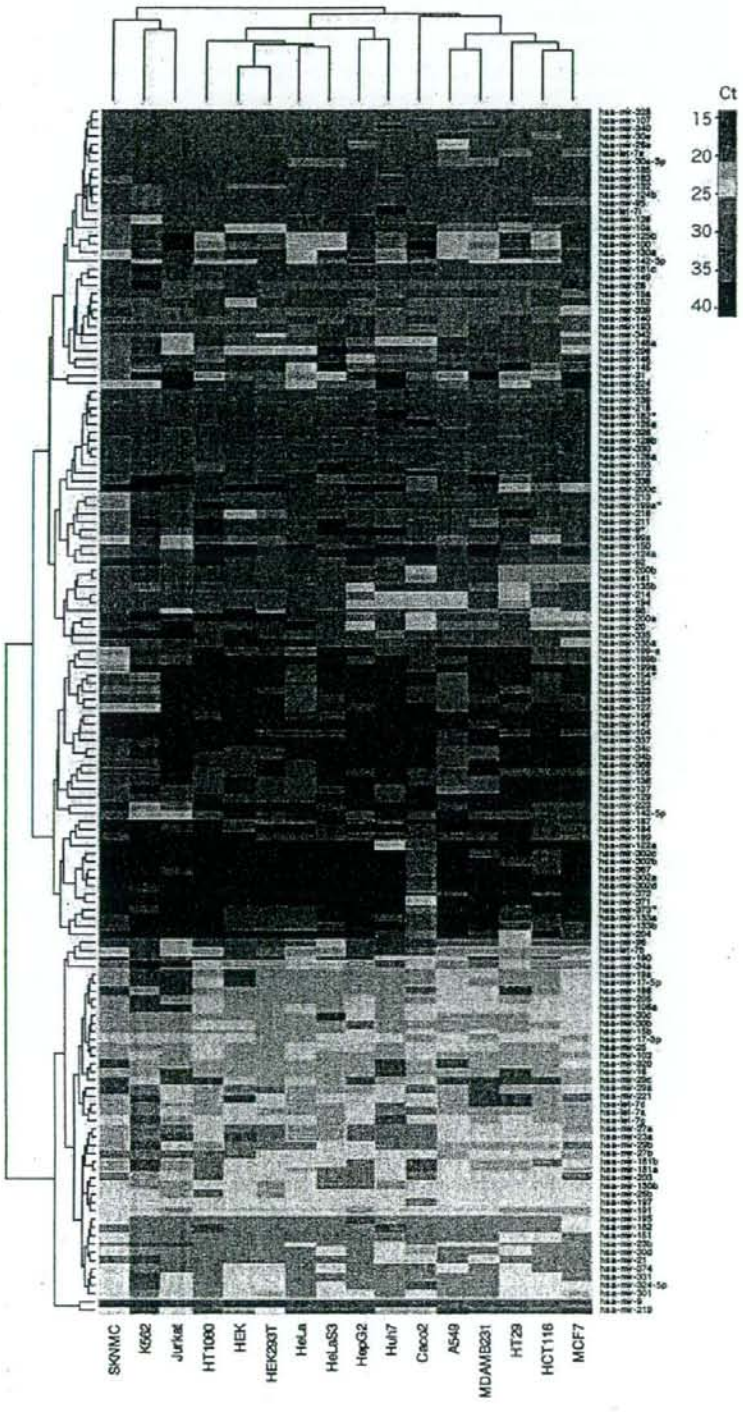
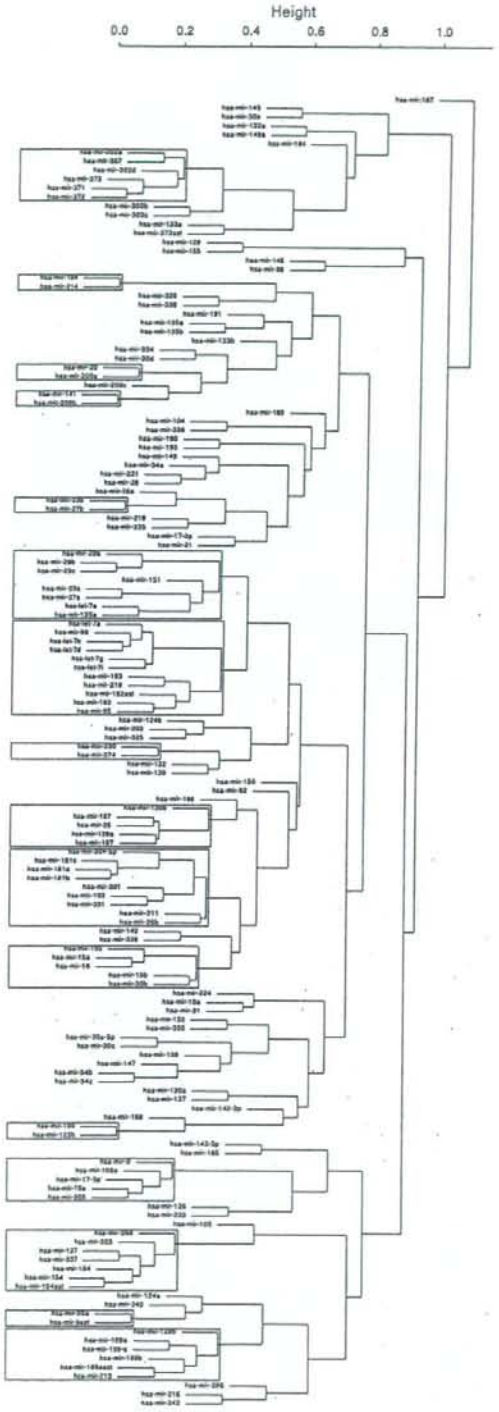


Fig. 1 Hierarchical clustering of 148 microribonucleic acid (miRNA) expression profiles in 16 human cell lines. Expression profiles (ΔCt values) of 148 miRNAs measured in total RNA from cell lines were clustered



◀ **Fig. 2** Hierarchical clustering with bootstrap analysis based on correlational expression profiles among 148 microribonucleic acids (miRNAs). Expression profiles of 148 miRNAs were clustered and verified for significance by multiscale bootstrap resampling analysis (1,000 iterations, sampling with replacement). Clusters were scored as statistically significant in cases in which miRNAs clustered with a 99% or better confidence interval (corresponding to $p < 0.01$), as determined by the bootstrap analysis

using standard Pearson's correlation test (Table 1). The strong correlation between intronic miRNA and host genes was observed, suggesting they are derived from the same precursor genes, which might be under the control of the same promoter.

To determine whether the negative correlation observed between miRNA and gene expression reflected miRNA-target gene relationships, we first examined the relationship using a previously identified miRNA-target gene pairs. We examined *miR-124a*, which has the largest number of known target genes (Lim et al. 2005), as a control. The *miR-124a* expression level measured by RT-PCR analysis was compared with either those of the known target genes or those of all genes, both obtained by microarray analysis in 16 cell lines. Then, each data set was subjected to Pearson's correlation test (Fig. 4). The distribution of correlation coefficient for *miR-124a*-target gene pairs was shifted toward negative side, compared to that of the *miR-124a*-total gene pairs. The mean of correlation coefficients between the two sets was significantly different ($p < 0.01$). Furthermore, Fisher's exact test showed a significant accumulation of target genes

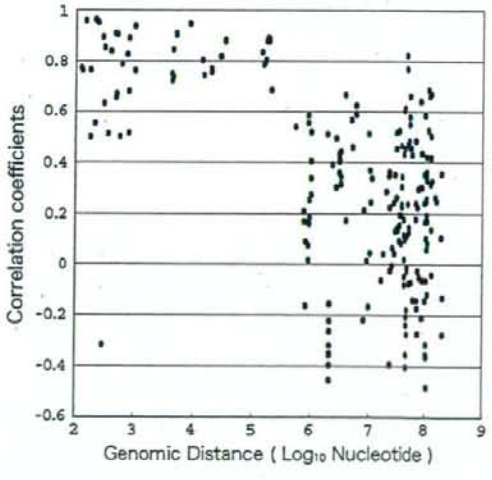


Fig. 3 Relationship between the distance separating mico-RNA (miRNA) loci and their coordinate expression in human tissues. Each miRNA was paired with each of the others lying in the same orientation on the same chromosome. For each pair, the correlation coefficient for their expression was plotted according to the distance between the two loci (*white circle*)

Table 1 Correlation of intronic microribonucleic acid (miRNA) expression with host gene expression

Host gene	RefSeq ID	miRNA	Correlation
MEST	NM_002402	hsa-mir-335	0.92
C9orf3	NM_032823	hsa-mir-23b	0.86
EVL	NM_016337	hsa-mir-342	0.84
C9orf3	NM_032823	hsa-mir-27b	0.83
PTK2	NM_005607	hsa-mir-151	0.82
240284_x_at	–	hsa-mir-141	0.80
GABRE	NM_004961	hsa-mir-224	0.78
RNF130	NM_018434	hsa-mir-340	0.77
EGFL7	NM_016215	hsa-mir-126	0.76
LPP	NM_005578	hsa-mir-28	0.72
229934_at	–	hsa-mir-223	0.70
1554097_a_at	–	hsa-mir-31	0.67
ZNF265	NM_005455	hsa-mir-186	0.66
FAM33A	NM_182620	hsa-mir-301	0.66
EML2	NM_012155	hsa-mir-330	0.60
1558515_at	–	hsa-mir-374	0.60
GPC1	NM_002081	hsa-mir-149	0.59
MGC11257	NM_032350	hsa-mir-339	0.59
C9orf3	NM_032823	hsa-mir-189	0.59
ARPP-21	NM_001025068	hsa-mir-128b	0.58
BIC	NR_001458	hsa-mir-155	0.58
240284_x_at	–	hsa-mir-200c	0.56
AATK	XM_927215	hsa-mir-338	0.54
FSTL1	NM_007085	hsa-mir-198	0.53
SMC4L1	NM_001002799	hsa-mir-15b	0.52
CTDSP1	NM_021198	hsa-mir-26b	0.50
MCM7	NM_005916	hsa-mir-25	0.50
MEG3	NR_002766	hsa-mir-127	0.45
ARRB1	NM_004041	hsa-mir-326	0.39
PDE2A	NM_002599	hsa-mir-139	0.26
TMEM113	NM_025222	hsa-let-7g	0.26
228528_at	–	hsa-mir-29c	0.21
WWP2	NM_007014	hsa-mir-140	0.19
TLN2	NM_015059	hsa-mir-190	0.16
R3HDM1	NM_015361	hsa-mir-128a	0.15
TRPM3	NM_001007470	hsa-mir-204	0.00
COPZ2	NM_016429	hsa-mir-152	−0.10
PANK1	NM_138316	hsa-mir-107	−0.14
DLEU2	NR_002612	hsa-mir-15a	−0.17
C21orf34	NM_001005732	hsa-mir-99a	−0.38

Gene symbol of the host gene, RefSeq ID, intronic miRNA, and correlation coefficient are listed for each intronic miRNA host gene pairs

within the inversely correlated group of genes [odds ratio (OR) 2.5, $p < 0.01$], indicating that the negative correlation observed between miRNA and gene expression in this reflects miRNA-target gene relationship.

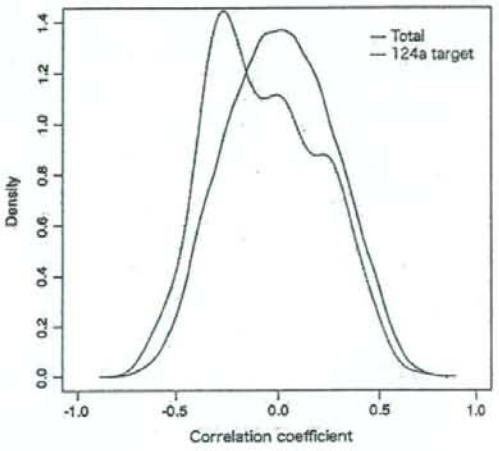


Fig. 4 Distribution of correlation coefficients between *miR-124a* and known targets of *miR-124a*. The estimated densities of distributions for target genes (*red*) or total genes (*black*) were plotted

To extract a global relationship of miRNA-target genes, all miRNA-mRNA pairs were subjected to Pearson's correlation test. Of the total 5.7 million different pairs examined, 184,901 miRNA-mRNA pairs (3.2%) showed a negative correlation.

Applying miRNA-target prediction algorithm

Because negative correlation might also reflect indirect regulation of gene expression by miRNAs, we further examined whether the extracted genes contained the corresponding target sites for miRNA in the 3′ UTRs. To seek miRNA-directed target sites, we adopted the miRanda algorithm (Enright et al. 2003) to a set of 30,664 different transcripts enrolled in RefSeq and monitored them by microarray. This analysis yielded 44,572 miRNA-target gene pairs that fulfilled the conditions that the expression level of genes negatively correlated with that of the miRNA, and that the genes contain the target sequence against the miRNA. Hence, approximately 300 genes per miRNA were extracted on average. Based on the extracted miRNA-target gene pairs, we next tried to predict the functions of miRNA. Thus, we calculated the frequency of gene ontology (GO) term (biological process area) of target genes for each miRNA. The GO terms were chosen if the frequency of the GO term was statistically significantly more than that expected by chance. Table 2 lists the top 30 miRNAs that had the highest GO term frequency. Most miRNA-target gene pairs were found to be novel, and the frequently noted GO terms for each miRNA was a diverse and wide variety in the biological process.

Table 2 Gene ontology (GO) term enrichment analysis for predicted microribonucleic acid (miRNA)-target genes

GO ID	Number of categorized genes	Number of extracted genes	Number of expected genes	*P* value (weight algorithm)	miRNA	GO definition
GO:0006694	139	10	1.1	1.70E-07	hsa-mir-320	Steroid biosynthesis
GO:0006694	139	8	0.79	1.30E-06	hsa-mir-34c	Steroid biosynthesis
GO:0016125	158	9	1.19	3.20E-06	hsa-mir-34b	Sterol metabolism
GO:0030001	663	27	10.95	1.80E-05	hsa-mir-331	Metal ion transport
GO:0006955	991	42	19.2	2.00E-05	hsa-mir-296	Immune response
GO:0045086	19	5	0.36	2.20E-05	hsa-mir-34a	Positive regulation of interleukin-2 biosynthesis
GO:0046321	9	3	0.06	2.40E-05	hsa-mir-99a	Positive regulation of fatty acid oxidation
GO:0015808	4	2	0.01	3.60E-05	hsa-mir-195	L-alanine transport
GO:0001568	369	12	2.95	4.50E-05	hsa-mir-150	Blood-vessel development
GO:0001678	4	2	0.01	7.10E-05	hsa-mir-16	Cell glucose homeostasis
GO:0042104	12	4	0.25	8.20E-05	hsa-mir-324-5p	Positive regulation of activated T-cell proliferation
GO:0016125	158	5	0.44	8.40E-05	hsa-mir-152	Sterol metabolism
GO:0006694	139	7	1.04	8.90E-05	hsa-mir-34b	Steroid biosynthesis
GO:0051085	19	4	0.25	9.70E-05	hsa-mir-133a	Chaperone cofactor-dependent protein folding
GO:0007588	78	9	1.92	0.00013	hsa-mir-328	Excretion
GO:0046006	16	4	0.3	0.00018	hsa-mir-149	Regulation of activated T-cell proliferation
GO:0045471	4	2	0.02	0.00019	hsa-mir-148a	Response to ethanol
GO:0006355	4315	42	24.64	0.00019	hsa-mir-31	Regulation of transcription, DNA dependent
GO:0006955	991	29	13.02	0.00022	hsa-mir-133a	Immune response
GO:0006656	14	2	0.02	0.00026	hsa-mir-368	Phosphatidylcholine biosynthesis
GO:0042552	44	2	0.03	0.00028	hsa-mir-29c	Myelination
GO:0030001	663	33	16.34	0.0003	hsa-mir-328	Metal ion transport
GO:0042104	12	3	0.14	0.00032	hsa-mir-181b	Positive regulation of activated T-cell proliferation
GO:0006825	16	3	0.14	0.00033	hsa-let-7i	Copper ion transport
GO:0006857	14	3	0.14	0.00033	hsa-mir-122a	Oligopeptide transport
GO:0006611	54	3	0.13	0.00033	hsa-mir-21	Protein export from nucleus
GO:0006825	16	3	0.14	0.00034	hsa-mir-132	Copper ion transport
GO:0030041	71	4	0.33	0.00035	hsa-mir-154	Actin filament polymerization
GO:0006235	3	2	0.03	0.00035	hsa-mir-27b	dTTP biosynthesis
GO:0006011	4	2	0.03	0.00035	hsa-mir-325	UDP-glucose metabolism

Top 30 highly enriched GO terms. *P* values were calculated using weight algorithm

Discussion

To investigate biological functions of miRNA, it is critical to identify miRNA-directed target genes. However, currently available computational methods (e.g., miRanda, PicTar, and TargetScan) predict numerous target genes that contain many false positives for miRNA (Mazière and Enright 2007). Also, experimental verification of miRNA-target relationship is complicated by the potential outcome of such an interaction being either translational repression or degradation. Furthermore, miRNAs can target multiple genes, and thereby the biological function of a single miRNA can be diverse. Hence, not only to achieve a higher degree of specificity of the prediction (few false positives) but also to comprehensively understand the function of miRNA, large-scale prediction of targets across a whole genome would be required.

In this study, we intended to investigate global miRNA-target relationships, assuming a negative correlation between miRNA and gene expression levels as an indicator for the miRNA-target relationship. The collected genome-wide expression data sets of both miRNA and mRNA enabled global correlation analysis between an miRNA and its target gene. Based on the observation that the expression of genes known as *miR-124a* targets tended to be negatively correlated with *miR-124a* expression, we reasoned that correlation analysis would be useful in investigating miRNA-target relationship. Then, we further extracted miRNA-target gene pairs by collecting genes that contain miRNA-directed target sites within 3′ UTR. The resulting list of miRNA gene pairs may provide useful information in investigating miRNA-target relationship.

Our miRNA-expression study provides important information on endogenous miRNA expression in 16 cell lines. One of the most serious problems in exploring miRNA function in a cell-base assay is the influence of endogenous miRNA expression. For example, in vitro reporter gene assay, a general approach used for target gene validation, has different results in different cell lines examined. This is thought to be due to the differential expression of endogenous miRNAs in the cell line used, as it strongly affects background expression of the reporter gene containing the test sequence. As we examined genetically tractable and well-studied cell lines in this study, the data obtained should be informative for studying functions of miRNA using those cell lines.

In collecting global miRNA and gene-expression profiles, we observed that many of the neighboring miRNA pairs, which were located within ~100-kb region of the chromosome, showed a significantly positive correlation and that the expression of intronic miRNAs was generally correlated with that of host genes. These results were consistent with the idea that clustered miRNAs are processed from the same primary transcript (Suh et al. 2004) and that intronic miRNAs are from the same primary transcript as their host gene (Lau et al. 2001; Sempere et al. 2004). Unexpectedly, however, we observed that miRNAs expressed from the same miRNA precursor (ex, *miR-142-3*p and *-5p*) were not coexpressed. This discrepancy could be explained by the fact that expression of these miRNAs depends on asymmetric selection of mature miRNA strand followed by the cleavage of pre-miRNA (Hutvagner 2005). Although it likely involves differential binding to and then differential retention of the two individual RNA strands by Dicer and its associated proteins (Zeng 2006), the detailed molecular mechanism for this asymmetry is not clear. The differential expression among cell lines we studied may provide a clue to the mechanism of asymmetric selectivity.

Correlated expression of miRNA and corresponding genes may imply their association in function as well. GO-based annotation and functional enrichment analysis of the extracted miRNA-target gene pairs made it possible to indicate putative functions of miRNAs. The approach developed in this study should be of value for future studies into the functions of miRNAs.

Acknowledgments This work was supported in part by research grants from the Scientific Fund of the Ministry of Education, Science, and Culture of Japan (to GT); National Institute of Biomedical Innovation (to GT); the Program for Promotion of Fundamental Studies in Health Sciences of the National Institute of Biomedical Innovation (NIBIO) (to GT); the Japan Health Science Foundation and the Ministry of Human Health and Welfare (to GT); and the 21st Century Center of Excellence Program "Knowledge Information Infrastructure for Genome Science" (to GT and YR). YR is supported as a Research Assistant by 21st Century COE program "Knowledge Information Infrastructure for Genome Science".

References

Aden DP, Fogel A, Plotkin S, Damjanov I, Knowles BB (1979) Controlled synthesis of HBsAg in a differentiated human liver carcinoma-derived cell line. Nature 282:615–616

Alexa A, Rahnenführer J, Lengauer T (2006) Improved scoring of functional groups from gene expression data by decorrelating GO graph structure. Bioinformatics 22:1600–1607

Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G (2000) Gene ontology: tool for the unification of biology. The gene ontology consortium. Nat Genet 25:25–29

Bagga S, Bracht J, Hunter S, Massirer K, Holtz J, Eachus R, Pasquinelli AE (2005) Regulation by let-7 and lin-4 miRNAs results in target mRNA degradation. Cell 122: 553–563

Baskerville S, Bartel DP (2005) Microarray profiling of microRNAs reveals frequent coexpression with neighboring miRNAs and host genes. RNA 11:241–247

Behm-Ansmant I, Rehwinkel J, Doerks T, Stark A, Bork P, Izaurralde E (2006) mRNA degradation by miRNAs and GW182 requires both CCR4:NOT deadenylase and DCP1:DCP2 decapping complexes. Genes Dev 20:1885–1898

Biedler JL, Helson L, Spengler BA (1973) Morphology and growth, tumorigenicity, and cytogenetics of human neuroblastoma cells in continuous culture. Cancer Res 33:2643–2652

Brattain MG, Fine WD, Khaled FM, Thompson J, Brattain DE (1981) Heterogeneity of malignant cells from a human colonic carcinoma. Cancer Res 41:1751–1756

Cailleau R, Young R, Olivè M, Reeves WJ (1974) Breast tumor cell lines from pleural effusions. J Natl Cancer Inst 53:661–674

Chendrimada TP, Finn KJ, Ji X, Baillat D, Gregory RI, Liebhaber SA, Pasquinelli AE, Shiekhattar R (2007) MicroRNA silencing through RISC recruitment of eIF6. Nature 447:823–828

DuBridge RB, Tang P, Hsia HC, Leong PM, Miller JH, Calos MP (1987) Analysis of mutation in human cells by using an Epstein–Barr virus shuttle system. Mol Cell Biol 7:379–387

Enright AJ, John B, Gaul U, Tuschl T, Sander C, Marks DS (2003) MicroRNA targets in Drosophila. Genome Biol 5:R1

Farh KK, Grimson A, Jan C, Lewis BP, Johnston WK, Lim LP, Burge CB, Bartel DP (2005) The widespread impact of mammalian

MicroRNAs on mRNA repression and evolution. Science 310:1817–1821

Fogh J, Wright WC, Loveless JD (1977) Absence of HeLa cell contamination in 169 cell lines derived from human tumors. J Natl Cancer Inst 58:209–214

Gaur A, Jewell DA, Liang Y, Ridzon D, Moore JH, Chen C, Ambros VR, Israel MA (2007) Characterization of microRNA expression levels and their biological correlates in human cancer cell lines. Cancer Res 67:2456–2468

Gautier L, Cope L, Bolstad BM, Irizarry RA (2004) Affy–analysis of Affymetrix GeneChip data at the probe level. Bioinformatics 20:307–315

Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S, Ellis B, Gautier L, Ge Y, Gentry J, Hornil K, Hothorn T, Huber W, Iacus S, Irizarry R, Leisch F, Li C, Maechler M, Rossini AJ, SawitzkiG, Smith C, Smyth G, Tierney L, Yang JY, Zhang J (2004) Bioconductor: open software development for computational biology and bioinformatics. Genome Biol 5:R80

Giard DJ, Aaronson SA, Todaro GJ, Arnstein P, Kersey JH, Dosik H, Parks WP (1973) In vitro cultivation of human tumors: establishment of cell lines derived from a series of solid tumors. J Natl Cancer Inst 51:1417–1423

Giraldez AJ, Mishima Y, Rihel J, Grocock RJ, VanDongen S, Inoue K, Enright AJ, Schier AF (2006) Zebrafish MiR-430 promotes deadenylation and clearance of maternal mRNAs. Science 312:75–79

Graham FL, Smiley J, Russell WC, Nairn R (1977) Characteristics of a human cell line transformed by DNA from human adenovirus type 5. J Gen Virol 36:59–74

Griffiths-Jones S (2004) The microRNA registry. Nucleic Acids Res 32:D109–D111

Hutvagner G (2005) Small RNA asymmetry in RNAi: function in RISC assembly and gene regulation. FEBS Lett 579:5850–5857

Iorio MV, Ferracin M, Liu CG, Veronese A, Spizzo R, Sabbioni S, Magri E, Pedriali M, Fabbri M, Campiglio M, Mènard S, Palazzo JP, Rosenberg A, Musiani P, Volinia S, Nenci I, Calin GA, Querzoli P, Negrini M, Croce CM (2005) MicroRNA gene expression deregulation in human breast cancer. Cancer Res 65:7065–7070

Jing Q, Huang S, Guth S, Zarubin T, Motoyama A, Chen J, DiPadova F, Lin SC, Gram H, Han J (2005) Involvement of microRNA in AU-rich element-mediated mRNA instability. Cell 120:623–634

Lau NC, Lim LP, Weinstein EG, Bartel DP (2001) An abundant class of tiny RNAs with probable regulatory roles in Caenorhabditis elegans. Science 294:858–862

Lewis BP, Shih IH, Jones-Rhoades MW, Bartel DP, Burge CB (2003) Prediction of mammalian microRNA targets. Cell 115:787–798

Lim LP, Lau NC, Garrett-Engele P, Grimson A, Schelter JM, Castle J, Bartel DP, Linsley PS, Johnson JM (2005) Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs. Nature 433:769–773

Liu W, Mao SY, Zhu WY (2007) Impact of tiny miRNAs on cancers. World J Gastroenterol 13:497–502

Lozzio CB, Lozzio BB (1973) Cytotoxicity of a factor isolated from human spleen. J Natl Cancer Inst 50:535–538

Mazière P, Enright AJ (2007) Prediction of microRNA targets. Drug Discov Today 12:452–458

Nakabayashi H, Taketa K, Yamane T, Oda M, Sato J (1985) Hormonal control of alpha-fetoprotein secretion in human hepatoma cell lines proliferating in chemically defined medium. Cancer Res 45:6379–6383

Puck TT, Marcus PI, Cieciura SJ (1956) Clonal growth of mammalian cells in vitro; growth characteristics of colonies from single HeLa cells with and without a feeder layer. J Exp Med 103:273–283

Rasheed S, Nelson-Rees WA, Toth EM, Arnstein P, Gardner MB (1974) Characterization of a newly derived human sarcoma cell line (HT-1080). Cancer 33:1027–1033

Rehwinkel J, Natalin P, Stark A, Brennecke J, Cohen SM, Izaurralde E (2006) Genome-wide analysis of mRNAs regulated by Drosha and Argonaute proteins in Drosophila melanogaster. Mol Cell Biol 26:2965–2975

van Rooij E, Sutherland LB, Liu N, Williams AH, McAnally J, Gerard RD, Richardson JA, Olson EN (2006) A signature pattern of stress-responsive microRNAs that can evoke cardiac hypertrophy and heart failure. Proc Natl Acad Sci USA 103:18255–18260

Scherer WF, Syverton JT, Gey GO (1953) Studies on the propagation in vitro of poliomyelitis viruses. IV. Viral multiplication in a stable strain of human malignant epithelial cells (strain HeLa) derived from an epidermoid carcinoma of the cervix. J Exp Med 97:695–710

Schneider U, Schwenk HU, Bornkamm G (1977) Characterization of EBV-genome negative "null" and "T" cell lines derived from children with acute lymphoblastic leukemia and leukemic transformed non-Hodgkin lymphoma. Int J Cancer 19:621–626

Sempere LF, Freemantle S, Pitha-Rowe I, Moss E, Dmitrovsky E, Ambros V (2004) Expression profiling of mammalian microRNAs uncovers a subset of brain-expressed microRNAs with possible roles in murine and human neuronal differentiation. Genome Biol 5:R13

Sevignani C, Calin GA, Siracusa LD, Croce CM (2006) Mammalian microRNAs: a small world for fine-tuning gene expression. Mamm Genome 17:189–202

Sood P, Krek A, Zavolan M, Macino G, Rajewsky N (2006) Cell-type-specific signatures of microRNAs on target mRNA expression. Proc Natl Acad Sci USA 103:2746–2751

Soule HD, Vazguez J, Long A, Albert S, Brennan M (1973) A human cell line from a pleural effusion derived from a breast carcinoma. J Natl Cancer Inst 51:1409–1416

Suh MR, Lee Y, Kim JY, Kim SK, Moon SH, Lee JY, Cha KY, Chung HM, Yoon HS, Moon SY, Kim VN, Kim KS (2004) Human embryonic stem cells express a unique set of microRNAs. Dev Biol 270:488–498

Suzuki R, Shimodaira H (2006) Pvclust: an R package for assessing the uncertainty in hierarchical clustering. Bioinformatics 22:1540–1542

Takamizawa J, Konishi H, Yanagisawa K, Tomida S, Osada H, Endoh H, Harano T, Yatabe Y, Nagino M, Nimura Y, Mitsudomi T, Takahashi T (2004) Reduced expression of the let-7 microRNAs in human lung cancers in association with shortened postoperative survival. Cancer Res 64:3753–3756

Tsuchiya S, Okuno Y, Tsujimoto G (2006) MicroRNA: biogenetic and functional mechanisms and involvements in cell differentiation and cancer. J Pharmacol Sci 101:267–270

Zeng Y (2006) Principles of micro-RNA production and maturation. Oncogene 25:6156–6162

J Cancer Res Clin Oncol (2008) 134:1363–1369
DOI 10.1007/s00432-008-0412-4

ORIGINAL PAPER

Microarray-based identification of CUB-domain containing protein 1 as a potential prognostic marker in conventional renal cell carcinoma

Yasuo Awakura · Eijiro Nakamura · Takeshi Takahashi · Hirokazu Kotani · Yoshiki Mikami · Tadashi Kadowaki · Akira Myoumoto · Hideo Akiyama · Noriyuki Ito · Toshiyuki Kamoto · Toshiaki Manabe · Hitoshi Nobumasa · Gozoh Tsujimoto · Osamu Ogawa

Received: 11 January 2008 / Accepted: 28 April 2008 / Published online: 16 May 2008

Abstract
Purpose Renal cell carcinoma (RCC) is characterized by a variable and unpredictable clinical course. Thus, accurate prediction of the prognosis is important in clinical settings. We conducted microarray-based study to identify a novel prognostic marker in conventional RCC.
Patients and methods The present study included the patients surgically treated at Kyoto University Hospital. Gene expression profiling of 39 samples was carried out to select candidate prognostic markers. Quantitative real-time PCR of 65 samples confirmed the microarray experiment results. Finally, we evaluated the significance of potential markers at their protein expression level by immunohistochemically analyzing 230 conventional RCC patients.
Results Using expression profiling analysis, we identified 14 candidate genes whose expression levels predicted unfavorable disease-specific survival. Next, we examined the expression levels of nine candidate genes by quantitative real-time PCR and selected CUB-domain containing protein 1 (CDCP1) for further immunohistochemical analysis. Positive staining for CDCP1 inversely correlated with disease-specific and recurrence-free survivals. In multivariate analysis including clinical/pathological factors, CDCP1 staining was a significant predictor of disease-specific and recurrence-free survivals.
Conclusions We identified CDCP1 as a potential prognostic marker for conventional RCC. Further studies might be required to confirm the prognostic value of CDCP1 and to understand its function in RCC progression.

Keywords CDCP1 · Microarray · Prognosis · Renal cell carcinoma

Electronic supplementary material The online version of this article (doi:10.1007/s00432-008-0412-4) contains supplementary material, which is available to authorized users.

Y. Awakura · E. Nakamura · T. Takahashi · N. Ito · T. Kamoto · O. Ogawa (✉)
Department of Urology,
Kyoto University Graduate School of Medicine,
54 Kawahara-cho, Shogoin, Sakyo-ku,
Kyoto 606-8507, Japan
e-mail: ogawao@kuhp.kyoto-u.ac.jp

H. Kotani · Y. Mikami · T. Manabe
Laboratory of Anatomic Pathology,
Kyoto University Hospital, Kyoto, Japan

T. Kadowaki · G. Tsujimoto
Department of Genomic Drug Discovery Science,
Kyoto University Graduate School of Pharmaceutical Sciences,
Kyoto, Japan

A. Myoumoto · H. Akiyama · H. Nobumasa
New Frontiers Research Laboratories,
Toray Indutries Inc., Kamakura, Japan

Introduction

Renal cell carcinoma (RCC) has a wide variety of clinical presentations. Approximately 30% patients present with metastatic disease, and up to 40% patients with localized RCC have recurrences after undergoing nephrectomy (Lam et al. 2005). Although most patients with metastatic disease have an unfavorable prognosis, frequently resulting in death within a year of diagnosis, some patients survive for years while the disease slowly progresses. At present, the histological stage, grade, and performance status are considered to be the most reliable prognostic factors (Shimazui et al. 1996; Mejean et al. 2003). Recently, to enhance the prediction of RCC prognosis, several groups have developed comprehensive integrated

staging systems based on clinical and histological factors (Lam et al. 2005).

Treatment of metastatic RCC remains challenging because it is highly resistant to chemotherapy and radiation. Although immunotherapy with interleukin-2 and/or interferons has been the mainstay for almost 20 years, the response rates are generally poor, ranging from 4 to 25% (Gleave et al. 1998; Motzer and Russo 2000). Recently, new therapeutic approaches based on molecular targeting have been reported, and several agents have shown encouraging results in clinical trials (Stadler 2005). In the near future, as various options for the treatment of metastatic RCC become available, better understanding of the prognostic factors would enable better selection of patients who might benefit from either conventional immunotherapy or new therapeutics. Therefore, it is important to identify reliable molecular markers for RCC prognosis because only clinical/histological factors are inadequate for the prognostication of RCC patients (Mejean et al. 2003; Maruyama et al. 2006).

High-throughput transcriptional profiling has emerged as a powerful approach for simultaneously screening a large number of genes. Recently, several studies have reported a set of genes that is associated with RCC prognosis (Takahashi et al. 2001; Vasselli et al. 2003; Jones et al. 2005; Kosari et al. 2005; Sultmann et al. 2005). However, the routine use of microarrays for predicting the RCC outcome is impractical because microarray facilities are expensive and are not commonly available. As an alternative method, quantitative real-time PCR assay or immunohistochemistry are particularly feasible in clinical settings. In addition, the scope of microarrays is limited in terms of gene expression quantification, and their results need to be validated using more sensitive experimental procedures including the above-mentioned methods (Chuaqui et al. 2002). In the present study, we analyzed the expression profile of 39 RCC samples to select candidate prognostic markers. Next, we used quantitative real-time PCR assay to confirm whether the expression levels of nine candidate genes are associated with the prognosis of 65 RCC patients. Thus, we identified CUB-domain containing protein-1 (CDCP1) as a potential prognostic marker. Finally, we examined the prognostic significance of CDCP1 protein expression by immunohistochemically studying 230 patients.

Materials and methods

RCC patients and samples

From January 1991 to December 2002, 240 patients underwent surgery and were histologically diagnosed with conventional RCC at Kyoto University Hospital. Of these 240 patients, 230 were enrolled in the present study. Ten patients were excluded because three patients had bilateral RCC, and formalin-fixed paraffin blocks were not available for the remaining seven patients. Of the 230 patients, 39 and 65 patients were enrolled in the microarray and quantitative real-time PCR experiments, respectively. There was an overlap of 25 patients, who were included in both experiments. Immunohistochemical analysis was conducted in all 230 patients. Mean age was 61.0 years (range: 29–90 years), and the male-to-female ratio was 2.8:1. The overall median follow-up was 41.4 months. A total of 31 patients died and median follow-up in surviving patients was 45.0 months. Tumors were staged and graded according to the International TNM classification system and the Fuhrman grading system (Goldstein 1999). Each tissue specimen was snap-frozen at −80°C immediately after surgical resection. Table 1 shows the characteristics of the patients with conventional RCC included in the microarray study ($n = 39$), quantitative real-time PCR ($n = 65$), and immnohistochemical analysis ($n = 230$). Informed consent was obtained from each patient. The institutional review board of the Kyoto University Graduate School of Medicine approved the present study.

RNA extraction and microarray experiment

Total RNA was isolated using RNeasy Kit and RNase-Free DNase (Qiagen, Valencia, CA, USA) according to the manufacturer's protocol. RNA integrity was verified by

Table 1 Characteristics of RCC patients included in microarray ($n = 39$), quantitative real-time PCR ($n = 65$), and immunohistochemical ($n = 230$) studies

	No. of patients (%)		
	Microarray study	Real-time PCR	Immunohistochemical study
Age (years)			
<60	11 (28.2)	19 (29.2)	88 (38.3)
60 or older	28 (71.8)	46 (70.8)	142 (68.7)
Gender			
Male	27 (69.2)	45 (69.2)	169 (73.5)
Female	12 (31.8)	20 (30.8)	61 (26.5)
Stage			
T1a	11 (28.2)	25 (38.5)	101 (43.9)
T1b&T2	16 (40.0)	23 (35.4)	80 (34.8)
>T2	12 (31.8)	17 (26.2)	49 (21.3)
N0M0	30 (76.9)	51 (78.5)	195 (84.8)
>N0 or M1	9 (23.1)	14 (21.5)	35 (15.2)
Grade			
G1&G2	31 (79.5)	45 (69.2)	179 (77.8)
G3&G4	8 (20.5)	20 (30.8)	51 (22.2)
Dead of RCC	9 (23.1)	13(20.0)	31 (13.5)

using the Agilent Bioanalyzer (Agilent Technologies, Palo Alto, CA, USA). Gene expression experiments were carried out using oligonucleotide-DNA microarrays. Detailed experimental procedures are described in Supplementary materials and methods.

Quantitative real-time PCR

cDNA was synthesized using First-Strand cDNA Synthesis kit (Amersham Biosciences, Piscataway, NJ, USA) as described in the manufacturer's protocol. The experimental procedures for real-time PCR are described in Supplementary materials and methods. The primer sequences are shown in Supplementary Table 1S.

Immunohistochemistry

Formalin-fixed paraffin sections (4 μm thick) were deparaffinized in xylene and rehydrated with graded ethanol. Endogenous peroxidase was inactivated by incubating the sections in 3% hydrogen peroxide for 30 min at room temperature. Antigen retrieval involved heat treatment for CDCP1 staining. Subsequently, the sections were incubated overnight at 4°C with polyclonal goat antibody against CDCP1 (Abcam, Cambridge, UK; 100:1 dilution), which was used in the previous study (Ikeda et al. 2006). After the primary antibody was washed, the sections were stained with Dako LSAB$^+$ system (Dako, Glostrup, Denmark) and counterstained with Meyer's hematoxylin. Biotin Blocking System (Dako) was used to examine CDCP1 expression in the normal kidney. The primary antibody was not used in the staining procedure for the control experiments and no specific staining was confirmed. CDCP1 staining was scored as follows: negative, <10% carcinoma cells stained and positive, ≥10% carcinoma cells stained. The percentage of stained cells was estimated in several fields (200×). Immunohistochemical scoring was performed in a single laboratory and was interpreted without any knowledge of the clinical data.

Statistics

Disease-specific survival was defined as the interval between surgery and death from RCC. Recurrence-free survival was defined as the interval between surgery and the subsequent appearance of local recurrence or metastatic disease.

To identify candidate genes whose expression levels correlated with disease-specific survival, the Cox proportional hazard model was fitted for each gene in the microarray experiment, as described previously (Neben et al. 2004). The cut-off level was defined as a hazard ratio (HR) of >2.5 and $P < 0.01$ in order to identify a set of genes associated with unfavorable survival.

The relationship between CDCP1 staining and clinical/pathological parameters was evaluated using Fisher's exact test in an immunohistochemical study. Survival curves were generated by the Kaplan–Meier method. The log-rank test was used to compare the survival curves and for univariate analysis. Multivariate analysis was conducted using the Cox proportional hazards model with a forward stepwise selection procedure. Statistical significance was set at $P < 0.05$. Statistical analyses were performed using the software package R ' (Ihaka and Gentlaman 1996) or Dr. SPSS II (SPSS, Chicago, IL, USA).

Results

Identification of candidate genes related to RCC survival by using microarray and quantitative real-time PCR

By analyzing the expression profile of 39 RCC samples, we identified 14 genes whose expression levels predicted unfavorable disease-specific survival (HR > 2.5, $P < 0.01$; Supplementary Table 2S). Of the 14 genes, we selected 9 genes, which are known to be involved in cell proliferation, apoptosis, or cancer development, and evaluated their expression in 65 samples by quantitative real-time PCR. According to the expression levels of the target gene, 65 patients were divided into two groups, and the disease-specific survival rate was compared (Supplementary Fig. 1S). Of the 9 genes analyzed, only CDCP1 was significantly associated with prognosis ($P = 0.039$). Thus, we performed immunohistochemical analysis on CDCP1 in the 230 patients, which was described as follows.

Immunohistochemical analysis of CDCP1

The patterns of CDCP1 expression in the carcinoma cells were observed as membranous and/or cytoplasmic staining (Fig. 1). The heterogeneous distribution of CDCP1 in both plasma membrane and cytoplasm has also been observed in colon cancers (Scherl-Mostageer et al. 2001). In the present study, CDCP1 staining was not detected in the glomerular and tubular cells of normal kidney. CDCP1 staining was positive in 77 cases (33.5%) and was significantly associated with the T stage, the presence of metastasis and the histological grade ($P = 0.028$, $P = 0.002$, and $P < 0.001$, respectively; Table 2).

Disease-specific survival

Univariate analysis of potential prognostic impact of clinical, pathological, and immunohistochemical parameters identified the presence of symptoms, T stage >2, the presence of nodal or distant metastasis, histological grade ≥3,