sample size (n = 40), Thuss-Patience et al. [8] reported that second-line CPT-11 monotherapy (250 to 350 mg/m^2, triweekly) significantly prolonged overall survival (OS) compared to best supportive care (BSC); the median survival in the CPT-11 arm was 123 days compared to 72.5 days for BSC; OS, hazard ratio [HR] = 2.85 (95% CI, 1.41–5.79); P = 0.0027. These results indicate that second-line chemotherapy using CPT-11 can now be considered as a treatment option in GC.

There have been two phase II studies evaluating IP therapy. Boku et al. [9] reported an ORR of 26.7% (4/15), and Ajani et al. [10] reported an ORR of 31% (9/29) and an MST of 5 months for AGC refractory to 5-FU therapy. In addition, in a retrospective study, Ueda et al. [11] reported a 28% (8/28) ORR, a progression-free survival (PFS) of 3.4 months, and an MST of 9.4 months.

Our present study had a selection bias, with comparatively few patients (23%) having peritoneal metastases, because such cases tend to be treated with taxanes; however, our results (28.6% ORR, TTP of 4.3 months, MST of 9.4 months, and 34.6% 1-year survival rate) indicate that second-line IP therapy for AGC appears to provide almost the same efficacy as that seen in other second-line trials, even for patients who have experienced S-1 failure.

We also demonstrated greater feasibility for IP therapy by using it in a second-line rather than first-line setting. In the first-line setting, IP therapy did not show statistically significant superiority to 5-FU because of its toxicity; more than 30% of patients receiving IP therapy discontinued for toxicity-related reasons, as opposed to fewer than 10% stopping for toxicity due to 5-FU and S-1 [1]. In the present study, grade 3/4 leukopenia or neutropenia were relatively mild, and only 10% of patients stopped treatment because of toxicity-related reasons. The reasons for these results may be that first, the duration of IP therapy is shorter in the second-line setting than in the first-line setting, and, second, in this study, dose reduction and discontinued treatment were carried out exactly according to protocol.

Another recent well-known IP regimen is biweekly CPT-11+CDDP. Koizumi et al. [12] reported on their phase I/II study using it as first-line therapy, where CPT-11 (60 mg/m^2) and CDDP (30 mg/m^2) were administered on days 1 and 15. In 2008, Nakae et al. [13] reported a phase II study of biweekly IP after S-1 failure. The ORR was 28.6%, and the median OS was 389 days. The most common grade 3/4 toxicities were: neutropenia (22.9%), anemia (11.4%), anorexia (14.3%), fatigue (8.6%), and diarrhea (2.9%). The efficacy and toxicity of this biweekly regimen were almost the same as those seen in our study. The biweekly regimen is available for outpatients; however, there are no phase III data on this regimen.

Currently, either CPT-11 (monotherapy or combined with CDDP) or taxanes [14–16] would be selected as second-line chemotherapy for AGC. However, it is not yet clear which of these regimens is most effective. Therefore, there are some currently ongoing clinical trials on the second-line treatment of AGC (after S-1 or S-1+CDDP failure), including phase III studies comparing CPT-11 and paclitaxel, CPT-11 alone and CPT-11+CDDP, CPT-11 alone and CPT-11+S-1, and so on. The results are anxiously awaited.

In conclusion, the combination of CPT-11 and CDDP as second-line chemotherapy for AGC appears to be effective and feasible, and should therefore be considered as a promising treatment option for patients who have experienced S-1 failure. Because CDDP has been widely used as first-line treatment for AGC patients, this regimen is suitable for patients who failed S-1 monotherapy used as adjuvant chemotherapy.

References

1. Boku N, Yamamoto S, Fukuda H, Shirao K, Doi T, Sawaki A, et al. Fluorouracil versus combination of irinotecan plus cisplatin versus S-1 in metastatic gastric cancer: a randomised phase 3 study. Lancet Oncol 2009;10:1063–9.
2. Koizumi W, Narahara H, Hara T, Takagane A, Akiya T, Takagi M, et al. S-1 plus cisplatin versus S-1 alone for first-line treatment of advanced gastric cancer (SPIRITS trial): a phase III trial. Lancet Oncol 2008;9:215–21.
3. Imamura H, Iiishi H, Tsuburaya A, Hatake K, Imamoto H, Esaki T, et al. Randomized phase III study of irinotecan plus S-1 (IRIS) versus S-1 alone as first-line treatment for advanced gastric cancer (GC0301/TOP002). Gastrointestinal Cancers Symposium 2008: abstract 5.
4. Sakuramoto S, Sasako M, Yamaguchi T, Kinoshita T, Fujii M, Nashimoto A, et al. Adjuvant chemotherapy for gastric cancer with S-1, an oral fluoropyrimidine. N Engl J Med 2007;357: 1810–20.
5. Hamaguchi T, Ohtsu A, Hyodo I, Arai Y, Takiuchi H, Fujii H, et al. A phase II study of biweekly irinotecan and mitomycin C combination therapy in patients with fluoropyrimidine-resistant advanced gastric cancer: The Japan Clinical Oncology Group trial (JCOG0109). J Clin Oncol, 2004 ASCO Annual Meeting Proceedings 2004;22(14S):4071.
6. Giuliani F, Molica S, Maiello E, Battaglia C, Gebbia V, Di Bisceglie M, et al. Irinotecan (CPT-11) and mitomycin-C (MMC) as second-line therapy in advanced gastric cancer: a phase II study of the Gruppo Oncologico dell' Italia Meridionale (prot. 2106). Am J Clin Oncol 2005;28:581–5.
7. Futatsuki K, Wakui A, Nakao I, Sakata Y, Kambe M, Shimada Y, et al. Late phase II study of irinotecan hydrochloride (CPT-11) in advanced gastric cancer. Gan To Kagaku Ryoho 1994;21:1033–8.
8. Thuss-Patience PC, Kretzschmar A, Deist T, Hinke A, Bichev D, Lebedinzew B, et al. Irinotecan versus best supportive care (BSC) as second-line therapy in gastric cancer: a randomized phase III study of the Arbeitsgemeinschaft Internistische Onkologie (AIO). J Clin Oncol 2009;27:(Suppl; abstract 4540).
9. Boku N, Ohtsu A, Shimada Y, Shirao K, Seki S, Saito H, et al. Phase II study of a combination of irinotecan and cisplatin against metastatic gastric cancer. J Clin Oncol 1999;17:319–23.
10. Ajani JA, Baker J, Pisters PW, Ho L, Mansfield PF, Feig BW, et al. Irinotecan/cisplatin in advanced, treated gastric or gastroesopha-

geal junction carcinoma. Oncology (Williston Park) 2002;16: 16–8.
11. Ueda S, Hironaka S, Boku N, Fukutomi A, Yoshino T, Onozawa Y. Combination chemotherapy with irinotecan and cisplatin in pretreated patients with unresectable or recurrent gastric cancer. Gastric Cancer 2006;9:203–7.
12. Koizumi W, Kurihara M, Satoh A, Takiuchi H, Tanabe S, Shimada K, et al. Phase I/II study of bi-weekly irinotecan plus cisplatin in the treatment of advanced gastric cancer. Anticancer Res 2005;25:1257–62.
13. Nakae S, Hirao M, Kishimoto T, Iijima S, Ishida H, Morimoto T, et al. Phase II study of bi-weekly CPT-11+CDDP for patients with gastric cancer refractory to S-1 (OGSG 0504 study). J Clin Oncol, 2008 ASCO Annual Meeting Proceedings 2008;26(15S):4571.
14. Arai T, Hamaguchi T, Shirao K, Shimada Y, Yamada Y, Muro K, et al. Weekly paclitaxel in patients with heavily treated advanced gastric cancer. Proc Am Soc Clin Oncol 2003;22 Abstract 1291.
15. Hironaka S, Zenda S, Boku N, Fukutomi A, Yoshino T, Onozawa Y. Weekly paclitaxel as second-line chemotherapy for advanced or recurrent gastric cancer. Gastric Cancer 2006;9:14–8.
16. Jo JC, Lee JL, Ryu MH, Sym SJ, Lee SS, Chang HM, et al. Docetaxel monotherapy as a second-line treatment after failure of fluoropyrimidine and platinum in advanced gastric cancer: experience of 154 patients with prognostic factor analysis. Jpn J Clin Oncol 2007;37:936–41.

Open

The Pharmacogenomics Journal (2010), 1–9

www.nature.com/tpj

ORIGINAL ARTICLE

Three-gene predictor of clinical outcome for gastric cancer patients treated with chemotherapy

HK Kim[1,2], IJ Choi[2], CG Kim[2], HS Kim[2], A Oshima[1], Y Yamada[3], T Arao[4], K Nishio[4], A Michalowski[1] and JE Green[1]

[1]*Laboratory of Cancer Biology and Genetics, National Cancer Institute, Bethesda, MD, USA;* [2]*National Cancer Center, Goyang, Gyeonggi, Republic of Korea,* [3]*National Cancer Center, Tokyo, Japan and* [4]*Kinki University School of Medicine, Osaka-Sayama, Japan*

Correspondence:
Dr JE Green, Laboratory of Cancer Biology and Genetics, National Cancer Institute, 37 Convent Drive, Bethesda, MD 20892, USA.
E-mail: JEGreen@mail.nih.gov

Received 23 March 2010; revised 10 October 2010; accepted 17 October 2010

To identify transcriptional profiles predictive of the clinical benefit of cisplatin and fluorouracil (CF) chemotherapy to gastric cancer patients, endoscopic biopsy samples from 96 CF-treated metastatic gastric cancer patients were prospectively collected before therapy and analyzed using high-throughput transcriptional profiling and array comparative genomic hybridization. Transcriptional profiling identified 917 genes that are correlated with poor patient survival after CF at $P<0.05$ (poor prognosis signature), in which protein synthesis and DNA replication/recombination/repair functional categories are enriched. A survival risk predictor was then constructed using genes, which are included in the *poor prognosis signature* and are contained within identified genomic amplicons. The combined expression of three genes—*MYC*, *EGFR* and *FGFR2*—was an independent predictor for overall survival of 27 CF-treated patients in the validation set (adjusted $P=0.017$), and also for survival of 40 chemotherapy-treated gastric cancer patients in a published data set (adjusted $P=0.026$). Thus, combined expression of *MYC*, *EGFR* and *FGFR2* is predictive of poor survival in CF-treated metastatic gastric cancer patients.
The Pharmacogenomics Journal advance online publication, 21 December 2010; doi:10.1038/tpj.2010.87

Keywords: gastric; cancer; chemotherapy; gene; expression

Introduction

Although the emerging area of targeted anticancer agents holds great promise, cytotoxic chemotherapy remains the primary treatment option for many cancer patients. Identifying patients who likely will or will not benefit from cytotoxic chemotherapy through the use of biomarkers could greatly improve clinical management by better defining appropriate treatment options for patients. None of the molecules experimentally identified to cause chemotherapy resistance *in vitro* was sufficiently validated in primary tumors and thus clinically applicable,[1] underscoring the importance of well-designed, clinical study to identify clinically relevant mechanisms for chemotherapy resistance. In fact, however, such predictors derived to date from high-throughput transcriptional profiling of primary tumors, especially gastrointestinal tract cancers, have not shown satisfactory performance.[2–5] It may be primarily owing to the high rate of false-positive discovery in high-throughput data, in addition to the high degree of genetic variation of individual tumor compared with limited number of samples available for the study.

To provide insight into clinically relevant mechanisms for chemotherapy resistance in gastric cancer, we prospectively collected and analyzed 123 endoscopic biopsy samples before cisplatin and fluorouracil (CF) chemotherapy from patients with extended follow-up, using high-throughput transcriptional profiling and comparative genomic hybridization (CGH) analyses. We could identify functional categories enriched in genes correlated with patient outcome, and develop a genomic predictor that was validated in two independent data sets.

Materials and methods

Patients

Sample collection, treatment and follow-up were performed according a protocol approved by the Institutional Review Board of the National Cancer Center Hospital in Goyang, Korea (NCCNHS01-003). All patients signed an Institutional Review Board-approved informed consent form. Eligibility for enrollment into the study included the following parameters: (1) age $\geqslant$ 18 years; (2) histologically confirmed gastric adenocarcinoma; (3) clinically documented distant metastasis; (4) no previous or concomitant malignancies other than the gastric cancer; (5) no previous history of chemotherapy, either adjuvant or palliative; and (6) adequate function of all major organs. Patients who were lost to follow-up before completing six cycles of chemotherapy, except for documented progressive disease, were excluded from this study.

Sample size calculation

Overall survival was the primary clinical end point of this study. As a minimum of 91 events were estimated to be required for the number of training set samples[6] at $\alpha = 0.001$, $\beta = 0.05$, τ (standard deviation of log intensity) $= 0.75$ and δ (hazard ratio (HR) associated with one-unit change of log intensity) $= 2$, we used the 96 samples collected until January 2005 as the training set for development of the predictor.

Ninety-six eligible patients who were treated with CF by one medical oncologist (HK) from August 2001 to January 2005 were used for the expression profiling training set. A second group of 27 eligible patients was used as the array validation cohort. Twenty-two patients in the validation cohort were treated with CF, and five patients were treated with cisplatin plus oral capecitabine (a fluorouracil pro-drug considered equivalent to fluorouracil; CX),[7] by another group of medical oncologists in the same institution between February 2005 and April 2006. Tissue procurement and processing were the same for the training and validation samples.

Treatment

Patients continued therapy indefinitely until they experienced unacceptable toxicities or progressive disease was documented. CF-treated patients received cisplatin 60 mg m^{-2} intravenously on day 1 and fluorouracil 1000 mg m^{-2} intravenously on days 1–5 of a 3-week schedule. The treatment schedule for fluorouracil could be shortened at the discretion of the oncologist to 3 instead of 5 days for elderly patients ($\geqslant$70 years) or patients with poor performance status (Eastern Cooperative Oncology Group performance status $\geqslant$2). Chemotherapy doses were reduced according to toxicities and the patient's performance status. Specific dose modification schemes for the subsequent cycle were left to the discretion of treating oncologist. Five patients (18.5%) in the validation group received oral capecitabine (Xeloda; Roche, Basel, Switzerland; 1250 mg m^{-2} twice a day for 2 weeks) instead of intravenous infusion of fluorouracil. Time to progression was measured from the initiation of chemotherapy to the progressive disease. In patients without any measurable lesions, time to progression was measured to the time when a change in therapy was required because unmeasurable lesions (such as ascites) unequivocally progressed.

Gene expression and CGH microarray analyses

Tissue samples were collected and processed for RNA and DNA extraction as described previously,[8] only if samples contained at least 50% tumor cells. Affymetrix (Santa Clara, CA, USA) HG-U133A gene expression microarray data were analyzed with survival analysis algorithms of BRB-ArrayTools (version 3.6, National Cancer Institute, http://linus.nci.nih.gov/BRB-ArrayTools.html).[9] The survival risk groups were constructed using a predictive index based on the supervised principal component method of Bair and Tibshirani.[10] A three-gene predictive index percentile was generated based on the weighted average of the log intensities of the three genes (*FGFR2* (211401_s_at), *EGFR* (210984_x_at) and *c-MYC* (202431_s_at)), using a proportional hazards regression on the first two principal components of the log intensities of those three genes, in which a high value of the predictive index corresponds to a high risk of death. If the predictive index of a sample in the validation set corresponded to the median predictive index of the training set, the sample was assigned a 50% predictive index. We specified the number of risk groups as 2 (high and low) and the predictive index percentile for defining the two risk groups as 67%, using a 67.1% rate of clinical benefit (partial response and stable disease) and 32.9% rate of progressive disease in the training set. We also performed Cox regression analyses using this three-gene predictive index percentile as a continuous variable, in which HRs for survival were calculated according to each percentile increase in three-gene predictive index percentile (from 0 to 100%). Array CGH data were generated using Agilent (Santa Clara, CA, USA) 4 $\times$ 44k HD-CGH Microarrays and analyzed using CGH Analytics software (version 3.5.14). Aberrations with average tumor/normal $\log_2$ ratio >2.0 were defined as amplifications. Experimental details are provided in Supplementary Materials and Methods.

Analyses of published DNA microarray data

The entire set of published Affymetrix U133 Plus 2.0 DNA microarray data[4] ($n = 40$) was combined with our training set data ($n = 96$), using common probe set IDs. MAS5 data of

the combined data set were $\log_2$ transformed, normalized using the median over the entire arrays and analyzed for survival risk prediction using BRB-ArrayTools 3.6, as described above.

Publicly accessible microarray data for surgically treated gastric cancer patients generated by the Stanford Functional Genomics Facility were obtained from the NCBI GEO database (GSE4007) and included about 30 300 genes common to these data sets. The microarray data were generated and normalized as described in Leung *et al.*[11] Batch effects in gene expression were removed with probe-wise mean centering and missing data were imputed with the nearest-neighbor averaging method.[12] The array cDNA clones were annotated using SOURCE (Stanford Microarray Database) and the Entrez GeneID was used as the mapping identifier for the Affymetrix HG-U133A array. A combined data set of our training set data ($n=96$) and GSE4007 data ($n=88$) was analyzed for survival risk prediction using BRB-ArrayTools 3.6 as described above.

Results

Genes correlated with poor survival after CF therapy

As primary gastric cancer lesions cannot be reliably measured by diagnostic imaging, patient survival, not radiographic response, was used as the primary clinical covariate to which gene expression was correlated to identify a predictor of response to CF therapy. To define a gene expression signature that correlates with overall survival, we used expression array data of 96 pretreatment biopsy samples as the training set to develop a predictor (Supplementary Table 1). Ninety-five out of 96 patients (99%) in the training set cohort died with follow-up for one survivor at 39.4 months. None of the clinicopathological or treatment factors listed in Table 1, including second-line chemotherapy, were significantly correlated with survival time of the patients in the training set.

To identify a transcriptional profile related to clinical benefit from CF therapy, the survival times of patients in the array training set were correlated with the mRNA expression levels measured by microarray. One thousand five hundred and sixty-five genes were significantly correlated with the overall survival of the 96 patients (P-value <0.05). Among them, 917 genes had an HR higher than 1 (poor prognosis signature) and 648 genes had an HR lower than 1 (good prognosis signature). We performed gene ontology analyses on this 'poor prognosis signature' using Ingenuity Pathway Analysis (www.ingenuity.com). The role of *BRCA1* in DNA damage response (*BRCA2, E2F5, FANCE, MSH2, NBN, PLK1, RFC, SMARCA4, SLC19A1*), nucleotide excision repair (*ERCC2, POLR2C, POLOR2J, RAD23A, RAD23B*) and estrogen receptor signaling were highly represented canonical pathways. Many of these *poor prognosis signature* genes belonging to these three pathways are previously linked to *in vitro* cisplatin resistance.[13–15] Overexpression of *ERCC2* ($P=0.007$ in our data) is associated with cisplatin resistance in lung cancer cell lines.[13] Silencing of hHR23A ($P=0.022$ in our data) decreases the nuclear DRP1 level and cisplatin resistance in lung adenocarcinoma cells.[14] Disruption of the Fanconi anemia–BRCA pathway is reported in cisplatin-sensitive ovarian tumors.[15] Thus, this gene ontology analysis supports the clinical relevance of these DNA repair

Table 1 Clinicopathological characteristics of patients

	Training set (n = 96)	*Validation set (n = 27)*
Baseline clinicopathological characteristic		
Age, no. (%)		
<70 years	90 (93.8%)	25 (92.6%)
⩾70 years	6 (6.2%)	2 (7.4%)
Sex, no. (%)		
Male	73 (76.0%)	23 (85.2%)
Female	23 (24.0%)	4 (14.8%)
PS, no. (%)		
ECOG PS 0 or 1	91 (94.8%)	25 (92.6%)
ECOG PS 2 or 3	5 (5.2%)	2 (7.4%)
Histological type, no. (%)		
Lauren's intestinal	40 (41.7%)	9 (33.3%)
Lauren's diffuse	56 (58.3%)	18 (66.6%)
Location of primary lesion, no. (%)		
Upper 1/3	14 (14.6%)	2 (7.4%)
Middle 1/3	28 (29.2%)	10 (37.0%)
Lower 1/3	49 (51.0%)	15 (55.6%)
Entire stomach	5 (5.2%)	0
Distant metastasis, no. (%)	96 (100%)	27 (100%)
Tumor cell percentage in sample (%)		
Median	60	70
Interquartile range	50–70	55–80
Treatment and outcome		
Chemotherapy regimen, no. (%)		
Cisplatin/fluorouracil	96 (100%)	22 (81.5%)
Cisplatin/capecitabine	0 (0%)	5 (18.5%)
Relative dose intensity (%)		
Median	79	81
Interquartile range	73–88	72–87
Number of chemotherapy cycles		
Median	4	7
Interquartile range	3–9	5–13
Response (WHO criteria), no. (%)		
PR	38 (44.7%)	12 (48.0%)
SD	19 (22.4%)	9 (36.0%)
PD	28 (32.9%)	4 (16.0%)
Non-measurable disease	11	2
Second-line chemotherapy, no. (%)	69 (71.9%)	19 (70.4%)
Median follow-up for survivors (months)	39.4	30.4
Overall survival (months)		
Median	8.1	12.6
Interquartile range	5.6–15.9	7.4–30.4
Time to progression (months)		
Median	3.9	6.3
Interquartile range	2.2–8.3	3.9–14.6

Abbreviations: ECOG, Eastern Cooperative Oncology Group; PD, progressive disease; PR, partial response; PS, performance status; SD, stable disease; WHO, World Health Organization.

canonical pathways, which were shown to be associated with *in vitro* cisplatin resistance.

Ingenuity Pathway Analysis functional categories enriched in poor prognosis signature were: protein synthesis, DNA replication/recombination/repair and cancer (Supplementary Table 2). The protein synthesis category includes ribosomal subunit mRNAs (*RPL13, RPL18, RPL24, RPL30, RPL38, RPLS, RPL7, RPL7A, RPL8, RPS2, RPS5*) and eukaryotic translation initiation factors (*EIF1, EIF2B2, EIF2B4, EIF2S1, EIF3B, EIF3C, EIF3D, EIF3E, EIF3F, EIF3H, EIF3I, EIF4A1, EIF4A3, EIF4B, EIF4EBP1, EIF5, EIF5B*). This result suggests that the most prominent feature of poor prognosis signature is increased protein synthesis, presumably resulting from activation of oncogenes, such as *EGFR, FGFR2* and *MYC* (Supplementary Table 2). MYC-induced transcriptional activation of protein synthesis-related genes is previously shown by a microarray report that the majority of genes responsive to MYC overexpression are involved in macromolecular synthesis, protein turnover and metabolism, including 30 ribosomal protein genes.[16]

Infinitesimal perturbation analysis canonical pathways enriched in 648 genes in good prognosis signature were antigen presentation pathway, B-cell development and interleukin-15 production. Enriched functional categories were gastrointestinal disease, inflammatory disease and genetic disorder.

Development of the three-gene predictor

Although such a gene ontology analysis of the whole signature provides some insight into clinically relevant mechanisms for chemotherapy resistance, this large number of genes is not readily amenable to clinical application. Therefore, we wished to narrow down 917 genes in the whole poor prognosis signature to the smaller number of genes, which may have driven the expression of majority of genes in the signature. Focusing on such 'driver gene' candidates would also minimize the chance of including false-positive discovery in a genomic predictor. For this purpose, a second tier of genomic analysis was performed to identify genes that could be functionally important in gastric cancer cells.

Genomic DNA from samples available from the training set patients was analyzed by array CGH to identify gene amplifications. Age, sex and overall survival were similar between the 30 patients (31.3%) whose samples were analyzed by array CGH and the other patients in the training set. Using very conservative criteria (average tumor/normal $\log_2$ ratio >2.0 for $\geqslant 5$ consecutive CGH probes), nine amplicons were identified in 11 patients (Table 2). We identified genes found in both the 1565 gene expression signature whose transcriptional levels correlated with poor survival of 96 training set patients (P-value <0.05) and that are also located within the nine amplicons identified by the array CGH. Three genes—*MYC* (8q24.13–24.21), *EGFR* (7p11.2) and *FGFR2* (10q26)—were identified in the amplicons (Table 2) whose expression array signal values significantly correlated with the survival time of the 96 patients in the training set (Figure 1). Patients with *EGFR* and *FGFR2* amplifications had higher expression levels of each gene (8.4 and 10.2 ± 0.8 (mean $\pm$ s.d.), for *EGFR* and *FGFR2*, respectively) than tested patients without the amplification of these genes (5.9 ± 1.0 and 5.2 ± 1.1, for *EGFR* and *FGFR2*, respectively). One of the two patients with *MYC* amplification had higher expression than patients without amplification (10.9 vs 9.5 ± 0.9).

The mRNA expression array signal values of these three genes were correlated with the short survival time with P-values of 0.0154, 0.0096 and 0.0057, for *MYC, EGFR* and *FGFR2*, respectively. The expression patterns of these three

Table 2 Amplicons identified using array CGH[a]

Cytoband	*Start*	*End*	*Target gene*	*No. of patients*
3q27.1	185 763 900	185 763 959	*EPHB3*	1
5q33.1	149 481 646	149 514 673	*PDGFRB*	1
7p11.2	54 746 103	55 363 004	*EGFR*	1
8q24.13–24.21	126 357 675	128 822 455	*MYC*	2
9p13.3	33 745 689	33 961 753	*PRSS3, UBE2R2, UBAP2*	1
10q26	123 264 724	13 123 458 467	*FGFR2*	2
17q12	35 046 052	35 282 145	*ERBB2*	2
17q21.2	36 110 139	36 230 022	*KRT24, KRT25A, KRT25C, KRT25D, KRT10*	2
17q21.2	36 569 493	36 888 515	*KRTAP4-4, KRTAP4-10, KRTAP9-9, KRTAP9-4, KRTAP17-1, KRTHA3A, KRTHA3B, KRTHA4, KRTHA1, KRTHA7, KRTHA8, KRTHA2, KRTHA5*	1

Abbreviation: CGH, comparative genomic hybridization.
[a]Defined by aberrations with average tumor/normal $\log_2$ ratio >2.0 for $\geqslant 5$ consecutive probes.

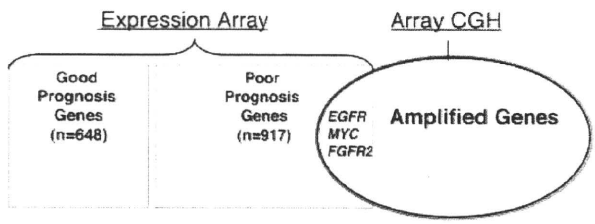

Figure 1 Three genes—*EGFR, FGFR2* and *MYC*—overlap between genes whose array expression levels correlated with survival times (96 training set patients, $P<0.05$) and gene copy number changes determined by array comparative genomic hybridization (CGH) (tumor/normal $\log_2$ ratio >2 for $\geqslant 5$ consecutive probes).

genes along with the cumulative survival data for all patients are depicted in the heatmap in Figure 2. None of the three genes had significantly different expression levels between those patients who received second-line chemotherapy and those who did not. Quantitative real-time RT-PCR and immunohistochemical staining for the three genes validated the array expression data (Supplementary Figures 1 and 2).

A three-gene predictive index percentile was then calculated for each of the 27 patients in the validation cohort, based on the weighted average of the log intensities of these three genes for each sample (designated as the three-gene predictor). Patterns of *MYC*, *EGFR* and *FGFR2* expression in these 27 patients, together with the predictive index, are graphically displayed in Figure 2. As a continuous variable, the three-gene predictive index percentile is an independent predictor for poor survival in the validation set by Cox regression analyses, after considering age, performance status, histological type and second-line chemotherapy (adjusted $P=0.017$) (Table 3). Patients predicted to have poor survival after CF using a predictive index percentile $\geqslant 67\%$ had a significantly shorter median survival than patients with a predictive index percentile $<67\%$ (7.4 months for the high-risk group vs 16.8 months for the

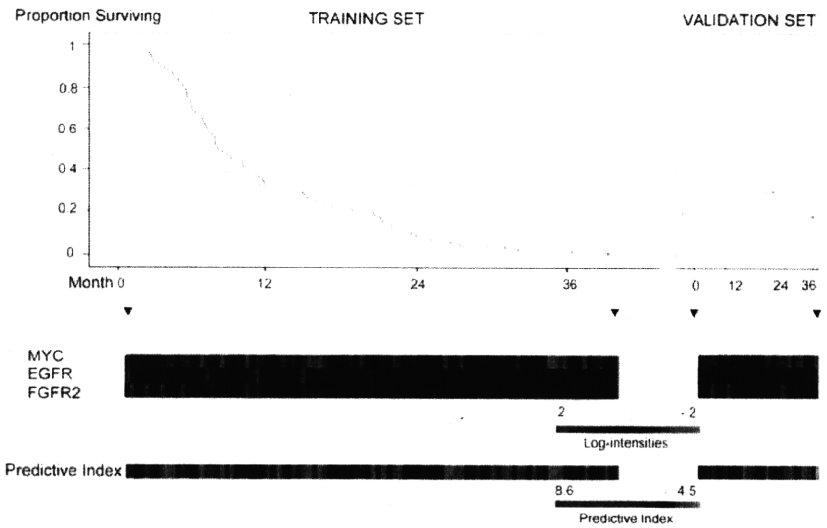

Figure 2 Affymetrix array expression levels of *MYC*, *EGFR* and *FGFR2* in 96 training set samples (left) and 27 validation set samples (right), shown with Kaplan–Meier plots for overall survival. Samples are ordered by the increasing survival period of patient from left to right, for the training and validation sets, respectively. A three-gene predictive index for each patient based on the three-gene predictor is indicated below.

Table 3 Cox regression analyses of the three-gene predictive index percentile, as a continuous variable, for 27 patients in the validation set

	Overall survival		*Time to progression*	
	P-value	*HR (95% CI)*	*P-value*	*HR (95% CI)*
Univariate				
Three-gene predictive index percentile[a]	0.050	1.015[b] (1.000–1.030)	0.026	1.017 (1.002–1.031)
Multivariate				
Three-gene predictive index percentile	0.017	1.023 (1.004–1.042)	0.014	1.023 (1.005–1.043)
Age $\geqslant 70$ years[c]	0.027	7.614 (1.257–46.130)	0.144	3.605 (0.646–20.112)
Poor performance status (ECOG PS 2 or 3)	0.346	2.130 (0.442–10.258)	0.074	4.829 (0.861–27.086)
Second-line chemotherapy	0.041	4.231 (1.064–16.831)	0.011	5.992 (1.502–23.902)
Diffuse histological type	0.773	1.164 (0.415–3.263)	0.280	1.774 (0.626–5.025)

Abbreviations: CI, confidence interval; ECOG PS, Eastern Cooperative Oncology Group performance status; HR, hazard ratio.
[a]Computed based on weighted average of log intensities of the three genes (*EGFR*, *FGFR2* and *MYC*) obtained using a proportional hazards regression on the first two principal components of the log signal intensities of those three genes.
[b]HR for each percentile increase in three-gene predictive index percentile. For example, a predictive index percentile of 100 (the highest predictive index) is associated with an HR of 4.4 ($=1.015^{100}$), compared with a predictive index percentile of 0 (the lowest predictive index). The median predictive index (50%) is associated with HRs of 2.1 ($=1.015^{50}$), compared with the lowest predictive index.
[c]For patients aged $\geqslant 70$ years, the treatment schedule for fluorouracil could be shortened at the discretion of the oncologist to 3 instead of 5 days.

low-risk cohort; $P=0.047$) (Figure 3a). As a class, the high-risk group predicted by the three-gene predictor (patient group with a predictive index percentile $\geqslant 67\%$) was associated with an adjusted HR of 3.1 (95% CI, 1.2–8.4; $P=0.022$). In addition, the three-gene predictive index percentile is also an independent predictor for the time to progression, which is a more specific indicator of the clinical responsiveness to systemic therapy than overall survival[17] (adjusted $P=0.014$) (Table 3). We therefore show that, independent of old age ($\geqslant 70$ years), poor performance status (Eastern Cooperative Oncology Group performance status $\geqslant 2$) and second-line chemotherapy, the three-gene predictive index is predictive of the benefit from CF to metastatic gastric cancer patients. An adjusted HR for time to progression according to each percentile increase in three-gene predictive index percentile was 1.023 (95% CI, 1.005–1.043) (that is, 100, 75 and 50% predictive indices are associated with an HR of 9.7 ($=1.023^{100}$), 5.5 ($=1.023^{75}$) and 3.1 ($=1.023^{50}$), respectively, compared with a 0% predictive index).

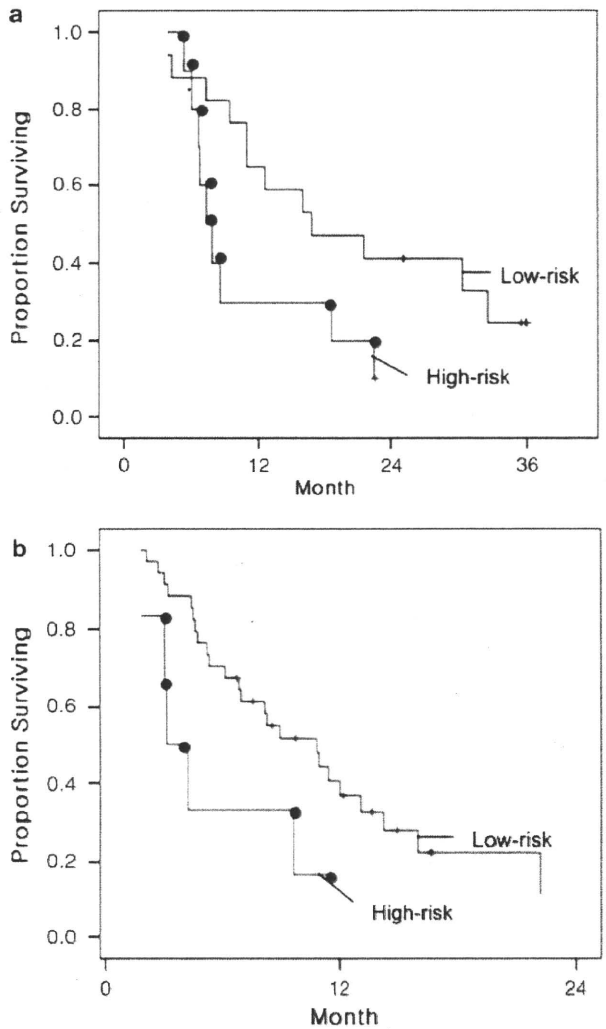

Figure 3 (**a**) Kaplan–Meier survival curves for the two risk groups of the validation cohort predicted by three-gene predictor. Patients at a high risk (predictive index percentile $\geqslant 67\%$; $n=10$) had significantly shorter median survival than patients at a low risk ($n=17$) (7.4 vs 16.8 months; log rank $P=0.047$). Green and blue lines represent overall survival curves for the predicted high- and low-risk groups, respectively. (**b**) Kaplan–Meier survival curves for the two risk groups of the published microarray data set from 40 metastatic gastric cancer patients treated with either fluorouracil-based regimens or cisplatin/irinotecan combination chemotherapy regimen. Patients at a high risk (predictive index percentile $\geqslant 67\%$; $n=6$) had shorter median survival than patients at a low risk ($n=34$), at a borderline significance (3.1 vs 10.8 months; log rank $P=0.056$). Green and blue lines represent overall survival curves for the predicted high- and low-risk groups, respectively. The color reproduction of the figure is available on the html full text version of the manuscript.

Three-gene predictor predicts survival of patients in the second validation set

To extend these results, we wished to test the predictive power of the three-gene predictor in other independent data sets. After the three-gene predictor was validated in 27 patient samples in our validation set, another microarray study with a comparable study design to our study was published in the literature.[4] These data were only one published microarray data set that could be used to determine whether the three-gene predictor could predict the outcome of metastatic gastric cancer patients treated with either cisplatin or fluorouracil. This data set contains pretreatment expression array data for 40 patients who subsequently received either fluorouracil-based chemotherapy ($n=24$) or cisplatin/irinotecan combination chemotherapy ($n=16$) and patient survival data. We applied the same three-gene predictor to this published microarray data set, just as we did to our 27 patient data in the first validation set. The three-gene predictive index percentile, as a continuous variable, was found to be significantly associated with poor survival of these 40 patients ($P=0.047$; HR according to each percentile increase in three-gene predictive index percentile = 1.014 (95% confidence interval, 1.000–1.027)). Cox multivariate analysis showed that the three-gene predictive index percentile is an independent predictor for poor survival, after considering performance status, age, sex and the chemotherapy regimen (adjusted $P=0.026$; adjusted HR = 1.017 (1.002–1.032)) (Table 4, Figure 3b). Thus, the predictive power of the three-gene predictor is consistent across two validation sets, that is, one from our study patients and the other from published data.

Interestingly, the three-gene predictor was found to be an independent predictor for poor survival, when the same Cox regression analysis was performed only on a subset of these patients ($n=16$) treated with cisplatin in combination with irinotecan, a topoisomerase I inhibitor (adjusted $P=0.011$; adjusted HR = 1.038 (1.008–1.068)). Patients treated with irinotecan were not included in the original training set patients. Hence, the predictive power of three-gene predictor may not be specifically associated with only CF therapy, although further large-scale studies need to be

performed to address the predictive value of the three-gene predictor for other therapeutic regimens.

Three-gene predictive index and radiographic response

Although the radiographic tumor response was not the main end point of this study, we also evaluated the association between the three-gene predictive index and radiographic response of patients with measurable disease. When published data[4] were also included, 104 patients had either partial response or stable disease (clinical benefit) as the best response, whereas 46 patients had progressive disease. The three-gene predictive index was significantly associated with radiographic response at a univariate *P*-value of 0.039, which is higher than the Cox regression *P*-value for the overall survival of all study patients (Table 5). This statistical association was at borderline significance in a multivariate regression analysis.

Table 4 Cox regression analyses of the three-gene predictive index percentile, as a continuous variable, for published DNA microarray data from 40 metastatic gastric cancer patients treated with either FU-based chemotherapy or cisplatin/irinotecan combination chemotherapy

	Overall survival	
	P-value	*HR (95% CI)*
Univariate		
Three-gene predictive index percentile	0.047	1.014 (1.000–1.027)
Multivariate		
Three-gene predictive index percentile	0.026	1.017[a] (1.002–1.032)
Performance status ⩾1	0.028	3.008 (1.129–8.016)
Age[b]	0.766	0.995 (0.961–1.030)
Male	0.538	1.359 (0.512–3.605)
FU-based chemotherapy regimen[c]	0.744	0.854 (0.332–2.199)

Abbreviations: CI, confidence interval; FU, fluorouracil; HR, hazard ratio.
[a]Adjusted HR for each percentile increase in three-gene predictive index percentile. For example, a predictive index percentile of 100 (the highest predictive index) is associated with an HR of 5.4 ($=1.017^{100}$), compared with a predictive index percentile of 0 (the lowest predictive index).
[b]As a continuous variable.
[c]As compared with the irinotecan/cisplatin combination chemotherapy regimen.

Three-gene predictor is not prognostic but predictive

Although we showed that the three-gene predictor predicted time to progression and overall survival for CF-treated patients, we wished to further address whether it represents a prognostic signature, using the published data set from 88 gastric cancer patients who were treated by surgery alone and not with chemotherapy.[11] The three-gene predictive index percentile was not a prognostic factor in this data set as a continuous variable ($P=0.506$). There was no difference in survival in the surgically treated patients between the high- and low-risk groups predicted by the three-gene predictor ($P=0.972$). These results strongly suggest that the three-gene predictor is not a predictor of prognosis for gastric cancer patients, but is predictive of the patient response to chemotherapy.

Discussion

Cytotoxic chemotherapy prolongs the median survival of metastatic gastric cancer patients from 3–5 to 9–11 months compared with best supportive care, with a response rate of 40–50%.[18–21] Combination CF constitutes the backbone for chemotherapy regimens commonly used for gastric cancers.[19,22] We also reported that CF in combination with low-dose docetaxel is active for metastatic gastric cancer with tolerable toxicity profile.[18] The ability to predict the primary resistance of common solid tumors to cytotoxic

Table 5 Logistic regression analysis on the three-gene predictive index for radiographic response of 150 patients with measurable disease, including patients represented by the published data set

	Radiographic response[a]		*Time to progression*		*Overall survival*	
	P-value	*OR (95% CI)*	*P-value*[b]	*HR (95% CI)*	*P-value*[c]	*HR (95% CI)*
Univariate						
Three-gene predictive index[d]	0.039	2.001 (1.036–3.864)	0.020	1.304 (1.042–1.631)	0.030	1.288 (1.026–1.618)
Multivariate						
Three-gene predictive index	0.059	1.902 (0.976–3.704)	0.019	1.309 (1.045–1.641)	0.018	1.316 (1.048–1.654)
Age ⩾70 years	0.914	1.069 (0.318–3.598)	0.791	1.119 (0.486–2.577)	0.113	1.600 (0.895–2.862)
Poor performance status (ECOG PS 2 or 3)	0.336	0.513 (0.132–1.999)	0.026	2.192 (1.097–4.381)	0.048	1.921 (1.004–3.677)

Abbreviations: CI, confidence interval; ECOG PS, Eastern Cooperative Oncology Group performance status; HR, hazard ratio; OR, odds ratio; WHO, World Health Organization
[a]No clinical benefit (progressive disease according to the WHO criteria; $n=46$) vs clinical benefit (partial response and stable disease; $n=104$).
[b]Result of Cox regression analysis on the three-gene predictive index for the time to progression of 123 patients in the training and the first validation sets.
[c]Result of Cox regression analysis on the three-gene predictive index for the overall survival of all of 163 study patients including published data set.
[d]Computed based on weighted average of log intensities of the three genes (*EGFR*, *FGFR2* and *MYC*) obtained using a proportional hazards *survival* regression on the first two principal components of the log signal intensities of those three genes.

chemotherapy is currently lacking, but would significantly improve patient care by identifying those who would best be treated by alternative strategies. This study has identified a three-gene predictor that distinguishes gastric cancer patients likely to receive a therapeutic benefit from CF from those who will not.

Most previous studies attempting to identify predictors of chemoresistance in gastric cancer have examined only individual genes such as *TS* or *ERCC1*.[23,24] High-throughput DNA microarray analyses to identify gene expression signatures predictive of chemotherapy or chemoradiotherapy resistance in gastrointestinal cancer patients have been limited by the small number of samples,[2,3] heterogeneous treatment[4] or were not prospectively designed.[5] In contrast to these previous studies, our study uses high-throughput genomic approaches, is prospective with a large, pre-defined number of training set patients, separate validation cohorts and survival data during an extended follow-up period. Although previously reported *TS* and *ERCC1* tend to be associated with poor prognosis of our patients, the association was not significant enough for them to be considered for our predictive model ($P=0.073$ and 0.076, for *TS* and *ERCC1*, respectively). Notably, the outcome discrimination predicted by the classifier was statistically significant on two validation groups, including the only available published microarray data set from chemotherapy-treated gastric cancer patients.[4] Although the sample size of our validation set is relatively small, it is nonetheless large enough to show that our three-gene predictor provides a statistically significant discrimination of patient outcome in multivariate survival analyses. The study design we employed is consistent with an allocation of two-thirds to one-third training-to-test set sample allocation as recommended by statisticians.[25]

We combined analyses of gene expression changes identified by expression profiling with the identification of DNA copy number changes using array CGH to develop a predictor composed of a much smaller number of critical genes that potentially could be of clinical utility. We identified *MYC*, *EGFR* and *FGFR2* in regions of amplification, as well as in the gene expression signature related to clinical outcome after CF therapy, suggesting that these genes might be functionally involved in determining resistance. Amplification of *MYC*, *EGFR* and *FGFR2* have previously been observed in gastric cancer at frequencies 4.8–15.5%,[26] 2.3–13.3%[27] and 3–10%,[26,28] respectively, suggesting that, in some cases, tumors amplify these regions for selective advantage. Combined expression of these three genes could predict overall survival and time to progression of CF-treated gastric cancer patients. Thus, combining array CGH analysis with relevant transcriptional changes is a feasible approach for building a predictive model using functionally important genes and reducing the likelihood of false biomarker discovery. Transcriptional levels of genes other than *MYC*, *EGFR* and *FGFR2* identified in the amplified genomic loci were not associated with the survival of the 96 training set patients (for example, $P=0.313$ for *ERBB2*).

Primary gastric tumors are not easily measurable by current radiographic techniques, and often there are no metastatic lesions that are readily quantifiable in metastatic gastric cancer patients. To develop a predictor from the general population of gastric cancer patients in an unbiased way, this study was designed to correlate gene expression profiling of the tumors with overall survival and time to progression, not radiographic response. Overall survival is the ultimate measure of the treatment benefit afforded to a patient and is a particularly appropriate gauge for patients with metastatic gastric cancer, as radiographic assessment is problematic in such patients. The fact that both the time to progression as well as overall survival are predicted by our three-gene predictor in CF-treated patients, but not surgically treated patients, suggests that the three-gene predictor is a predictive indicator for the clinical benefit from CF.

Although *EGFR* and *FGFR2* expression have been reported to have prognostic value for gastric cancer patients treated surgically,[29,30] we did not find the three-gene predictive index to be prognostic for surgically treated patients with gastric cancer. Our findings are consistent with previously reported experimental data on chemoresistance. Inhibitors of *EGFR* act synergistically with cisplatin[31] and fluorouracil,[32] whereas an *FGFR2* inhibitor is synergistic with fluorouracil.[33] MYC has been linked to cisplatin resistance in several *in vitro* models.[34–37]

Taken together, combined expression of *MYC*, *EGFR* and *FGFR2* is predictive of poor survival in CF-treated metastatic gastric cancer patients. More focused prospective trials that are designed to test the clinical utility of this three-gene predictor are warranted.

Conflict of interest

The authors declare no conflict of interest.

Acknowledgments

The work was supported in part by National Institute of Health Intramural Program, Center for Cancer Research, National Cancer Institute, Korean National Cancer Center Grant 0910570 and Converging Research Center Program through the Ministry of Education, Science and Technology of Korea (2010K001121). We thank Dr Richard Simon for valuable discussions and critical reading of the manuscript, Dr Lyuba Varticovski for critical review of the manuscript and Dr Chang-Hee Kim, Ms Susie Korolevich, Ms Eunbyul Lee, Ms Eugene Kim and Ms Ilji Jeon for technical help.

Transcript Profiling

All expression microarray data is available at Gene Expression Omnibus (accession number GSE14210; http://www.ncbi.nlm.nih.gov/geo).

References

1 Uchida K, Danenberg PV, Danenberg KD, Grem JL. Thymidylate synthase, dihydropyrimidine dehydrogenase, ERCC1, and thymidine phosphorylase gene expression in primary and metastatic gastro-

intestinal adenocarcinoma tissue in patients treated on a phase I trial of oxaliplatin and capecitabine. *BMC Cancer* 2008; **8**: 386.
2 Ghadimi BM, Grade M, Difilippantonio MJ, Varma S, Simon R, Montagna C *et al.* Effectiveness of gene expression profiling for response prediction of rectal adenocarcinomas to preoperative chemoradiotherapy. *J Clin Oncol* 2005; **23**: 1826–1838.
3 Rimkus C, Friederichs J, Boulesteix AL, Theisen J, Mages J, Becker K *et al.* Microarray-based prediction of tumor response to neoadjuvant radiochemotherapy of patients with locally advanced rectal cancer. *Clin Gastroenterol Hepatol* 2008; **6**: 53–61.
4 Yamada Y, Arao T, Gotoda T, Taniguchi H, Oda I, Shirao K *et al.* Identification of prognostic biomarkers in gastric cancer using endoscopic biopsy samples. *Cancer Sci* 2008; **99**: 2193–2199.
5 Allen WL, Coyle VM, Jithesh PV, Proutski I, Stevenson L, Fenning C *et al.* Clinical determinants of response to irinotecan-based therapy derived from cell line models. *Clin Cancer Res* 2008; **14**: 6647–6655.
6 Simon R, Radmacher MD, Dobbin K. Design of studies using DNA microarrays. *Genet Epidemiol* 2002; **23**: 21–36.
7 Kang Y, Kang WK, Shin DB, Chen J, Xiong J, Wang J *et al.* Randomized phase III trial of capecitabine/cisplatin vs continuous infusion of 5-FU/cisplatin as first-line therapy in patients with advanced gastric cancer: efficacy and safety results. *J Clin Oncol* 2006; **24** (Suppl): 18s (abstract 4018).
8 Kim HK, Choi IJ, Kim HS, Kim JH, Kim E, Park IS *et al.* DNA microarray analysis of the correlation between gene expression patterns and acquired resistance to 5-FU/cisplatin in gastric cancer. *Biochem Biophys Res Commun* 2004; **316**: 781–789.
9 Simon R, Lam A, Li MC, Ngan M, Menenzes S, Zhao Y. Analysis of gene expression data using BRB-Array Tools. *Cancer Informatics* 2007; **2**: 11–17.
10 Bair E, Tibshirani R. Semi-supervised methods to predict patient survival from gene expression data. *PLoS Biol* 2004; **2**: E108.
11 Leung SY, Chen X, Chu KM, Yuen ST, Mathy J, Ji J *et al.* Phospholipase A2 group IIA expression in gastric adenocarcinoma is associated with prolonged survival and less frequent metastasis. *Proc Natl Acad Sci USA* 2002; **99**: 16203–16208.
12 Troyanskaya O, Cantor M, Sherlock G, Brown P, Hastie T, Tibshirani R *et al.* Missing value estimation methods for DNA microarrays. *Bioinformatics* 2001; **17**: 520–525.
13 Weaver DA, Crawford EL, Warner KA, Elkhairi F, Khuder SA, Willey JC. ABCC5, ERCC2, XPA and XRCC1 transcript abundance levels correlate with cisplatin chemoresistance in non-small cell lung cancer cell lines. *Mol Cancer* 2005; **4**: 18–25.
14 Chiang YY, Chen SL, Hsiao YT, Huang CH, Lin TY, Chiang IP *et al.* Nuclear expression of dynamin-related protein 1 in lung adenocarcinomas. *Mod Pathol* 2009; **22**: 1139–1150.
15 Taniguchi T, Tischkowitz M, Ameziane N, Hodgson SV, Mathew CG, Joenje H *et al.* Disruption of the Fanconi anemia–BRCA pathway in cisplatin-sensitive ovarian tumors. *Nat Med* 2003; **9**: 568–574.
16 Guo QM, Malek RL, Kim S, Chiao C, He M, Ruffy M *et al.* Identification of c-myc responsive genes using rat cDNA microarray. *Cancer Res* 2000; **60**: 5922–5928.
17 Balko JM, Black EP. A gene expression predictor of response to EGFR-targeted therapy stratifies progression-free survival to cetuximab in KRAS wild-type metastatic colorectal cancer. *BMC Cancer* 2009; **9**: 145–154.
18 Park SR, Chun JH, Yu MS, Lee JH, Ryu KW, Choi IJ *et al.* Phase II study of docetaxel and irinotecan combination chemotherapy in metastatic gastric carcinoma. *Br J Cancer* 2006; **94**: 1402–1406.
19 Park SR, Chun JH, Yu MS, Lee JH, Ryu KW, Choi IJ *et al.* Phase II study of low-dose docetaxel/fluorouracil/cisplatin in metastatic gastric carcinoma. *Am J Clin Oncol* 2005; **28**: 433–438.
20 Cunningham D, Rao S, Starling N, Iveson T, Nicolson M, Coxon F *et al.* Randomised multicentre phase III study comparing capecitabine with fluorouracil and oxaliplatin with cisplatin in patients with advanced oesophagogastric (OG) cancer: The REAL 2 trial. *J Clin Oncol* 2006; **24**: 18s (abstr. 4017).
21 Pyrhönen S, Kuitunen T, Nyandoto P, Kouri M. Randomised comparison of fluorouracil, epidoxorubicin and methotrexate (FEMTX) plus supportive care with supportive care alone in patients with non-resectable gastric cancer. *Br J Cancer* 1995; **71**: 587–591.
22 Tetzlaff ED, Cen P, Ajani JA. Emerging drugs in the treatment of advanced gastric cancer. *Expert Opin Emerg Drugs* 2008; **13**: 135–144.
23 Lenz HJ, Leichman CG, Danenberg KD, Danenberg PV, Groshen S, Cohen H *et al.* Thymidylate synthase mRNA level in adenocarcinoma of the stomach: a predictor for primary tumor response and overall survival. *J Clin Oncol* 1996; **14**: 176–182.
24 Metzger R, Leichman CG, Danenberg KD, Danenberg PV, Lenz HJ, Hayashi K *et al.* ERCC1 mRNA levels complement thymidylate synthase mRNA levels in predicting response and survival for gastric cancer patients receiving combination cisplatin and fluorouracil chemotherapy. *J Clin Oncol* 1998; **16**: 309–316.
25 Dobbin KK, Simon RM. Optimally splitting cases for training and testing high dimensional classifiers (submitted for publication) http://linus.nci.nih.gov/~brb/reprints.htm.
26 Hara T, Ooi A, Kobayashi M, Mai M, Yanagihara K, Nakanishi I. Amplification of c-myc, K-sam, and c-met in gastric cancers: detection by fluorescence *in situ* hybridization. *Lab Invest* 1998; **78**: 1143–1153.
27 Moutinho C, Mateus AR, Milanezi F, Carneiro F, Seruca R, Suriano G. Epidermal growth factor receptor structural alterations in gastric cancer. *BMC Cancer* 2008; **8**: 10–15.
28 Yoshida T, Sakamoto H, Terada M. Amplified genes in cancer in upper digestive tract. *Semin Cancer Biol* 1993; **4**: 33–40.
29 Lieto E, Ferraraccio F, Orditura M, Castellano P, Mura AL, Pinto M *et al.* Expression of vascular endothelial growth factor (VEGF) and epidermal growth factor receptor (EGFR) is an independent prognostic indicator of worse outcome in gastric cancer patients. *Ann Surg Oncol* 2008; **15**: 69–79.
30 Yasui W, Oue N, Aung PP, Matsumura S, Shutoh M, Nakayama H. Molecular–pathological prognostic factors of gastric cancer: a review. *Gastric Cancer* 2005; **8**: 86–94.
31 Fan Z, Baselga J, Masui H, Mendelsohn J. Antitumor effect of anti-epidermal growth factor receptor monoclonal antibodies plus *cis*-diamminedichloroplatinum on well established A431 cell xenografts. *Cancer Res* 1993; **53**: 4637–4642.
32 Hiraishi Y, Wada T, Nakatani K, Uchida J, Iwasa T, Yoshida T *et al.* Synergistic antitumor effect of S-1 and the epidermal growth factor receptor inhibitor gefitinib in non-small cell lung cancer cell lines: role of gefitinib-induced down-regulation of thymidylate synthase. *Mol Cancer Ther* 2008; **7**: 599–606.
33 Yashiro M, Shinto O, Nakamura K, Tendo M, Matsuoka T, Matsuzaki T *et al.* Synergistic anti-tumor effects of FGFR2 inhibitor with 5-fluorouracil on scirrhous gastric carcinoma. *Int J Cancer* 2010; **126**: 1004–1016.
34 Leonetti C, Biroccio A, Candiloro A, Citro G, Fornari C, Mottolese M *et al.* Increase of cisplatin sensitivity by c-myc antisense oligodeoxynucleotides in a human metastatic melanoma inherently resistant to cisplatin. *Clin Cancer Res* 1999; **5**: 2588–2595.
35 Xie XK, Yang DS, Ye ZM, Tao HM. Recombinant antisense C-myc adenovirus increase *in vitro* sensitivity of osteosarcoma MG-63 cells to cisplatin. *Cancer Invest* 2006; **24**: 1–8.
36 Knapp DC, Mata JE, Reddy MT, Devi GR, Iversen PL. Resistance to chemotherapeutic drugs overcome by c-Myc inhibition in a Lewis lung carcinoma murine model. *Anticancer Drugs* 2003; **14**: 39–47.
37 Biroccio A, Benassi B, Fiorentino F, Zupi G. Glutathione depletion induced by c-Myc downregulation triggers apoptosis on treatment with alkylating agents. *Neoplasia* 2004; **6**: 195–206.

Supplementary Information accompanies the paper on the The Pharmacogenomics Journal website (http://www.nature.com/tpj)

J Cancer Res Clin Oncol (2010) 136:911–921
DOI 10.1007/s00432-009-0733-y

ORIGINAL PAPER

Nucleolin as cell surface receptor for tumor necrosis factor-α inducing protein: a carcinogenic factor of *Helicobacter pylori*

Tatsuro Watanabe · Hideaki Tsuge · Takahito Imagawa · Daisuke Kise · Kazuya Hirano · Masatoshi Beppu · Atsushi Takahashi · Kensei Yamaguchi · Hirota Fujiki · Masami Suganuma

Received: 22 October 2009 / Accepted: 13 November 2009 / Published online: 5 January 2010

Abstract
Purpose Tumor necrosis factor-α inducing protein (Tipα) is a unique carcinogenic factor released from *Helicobacter pylori* (*H. pylori*). Tipα specifically binds to cells and is incorporated into cytosol and nucleus, where it strongly induces expression of *TNF-α* and *chemokine* genes mediated through NF-κB activation, resulting in tumor development. To elucidate mechanism of action of Tipα, we studied a binding protein of Tipα in gastric epithelial cells.
Methods Tipα binding protein was found in cell lysates of mouse gastric cancer cell line MGT-40 by FLAG-pull down assay and identified to be cell surface nucleolin by flow cytometry using anti-nucleolin antibody. Incorporation of Tipα into the cells was determined by Western blotting and expression of *TNF-α* gene was quantified by RT-PCR.
Results Nucleolin was co-precipitated with Tipα-FLAG, but not with del-Tipα-FLAG (an inactive mutant). After treatment with Tipα-FLAG, incorporated Tipα was co-immunoprecipitated with endogenous nucleolin using anti-nucleolin antibody. The direct binding of Tipα to recombinant His-tagged nucleolin fragment (284–710) was also confirmed. Although nucleolin is an abundant non-ribosomal protein of the nucleolus, we found that nucleolin is present on the cell surface of MGT-40 cells. Pretreatment with anti-nucleolin antibody enhanced Tipα-incorporation into the cells through nucleolin internalization. In addition, pretreatment with tunicamycin, an inhibitor of N-glycosylation, decreased the amounts of cell surface nucleolin and inhibited both internalization of Tipα and expression of *TNF-α* gene.
Conclusions All the results indicate that nucleolin acts as a receptor for Tipα and shuttles Tipα from cell surface to cytosol and nuclei. These findings provide a new mechanistic insight into gastric cancer development with Tipα.

Keywords Gastric cancer · TNF-α · NF-κB · *Helicobacter pylori* · Tumor promotion

Abbreviations

H. pylori	*Helicobacter pylori*
Tipα	TNF-α inducing protein
CagA	Cytotoxin associated antigen
LC–MS	Liquid chromatography–mass spectrometry
RBD	RNA binding domain
NF-κB	Nuclear factor-kappa B
NEMO	NF-κB essential modulator

T. Watanabe · A. Takahashi · M. Suganuma (✉)
Saitama Cancer Center, Research Institute for Clinical Oncology, Kitaadachi-gun, Saitama 362-0806, Japan
e-mail: masami@cancer-c.pref.saitama.jp

T. Watanabe · A. Takahashi
Graduate School of Science and Engineering, Saitama University, Sakura-ku, Saitama 338-8570, Japan

H. Tsuge · T. Imagawa
Institute for Health Sciences, Tokushima Bunri University, Tokushima 770-8514, Japan

D. Kise · H. Fujiki
Faculty of Pharmaceutical Sciences, Tokushima Bunri University, Tokushima 770-8514, Japan

K. Hirano · M. Beppu
Tokyo University of Pharmacy and Life Science, Hachioji, Tokyo 192-0392, Japan

K. Yamaguchi
Hospital, Saitama Cancer Center, Saitama 362-0806, Japan

Introduction

Helicobacter pylori (*H. pylori*) is a gram-negative bacterium that colonizes in the mucosa of human stomach, resulting in induction of chronic gastritis, peptic ulcer, and stomach cancer (IARC Working Group on the Evaluation of Carcinogenic Risks to Humans 1994, Peek and Blaser 2002). Key criteria of these clinical outcomes are the severity and persistence of inflammation caused by *H. pylori*-infection, associated with strong induction of inflammatory cytokines, such as tumor necrosis factor-α (TNF-α), interleukine-1 (IL-1) and chemokines (El-Omar et al. 2000; Peek 2008; Snaith and El-Omar 2008). It is well accepted that inflammatory cytokines contribute to maintain cancer microenvironment (Balkwill 2009; El-Omar et al. 2003), and among the inflammatory cytokines, TNF-α plays a master role as an endogenous tumor promoter in carcinogenesis (Balkwill 2009; Moore et al. 1999; Suganuma et al. 1999). Moreover, TNF-α released from the cells acts as an instigator of a cytokine network sequence, from TNF-α to IL-1 and IL-6 and back to TNF-α, maintaining inflammation in the process of tumor promotion (Suganuma et al. 2002).

To extend the concept, a new gene, *TNF-α inducing protein* (*Tipα*) gene, was cloned from the genome of *H. pylori* strain 26695. Tipα directly induces *TNF-α* gene expression in gastric epithelial cells (Suganuma et al. 2005, 2006, 2008). The unique features of Tipα protein are as follows: (1) *H. pylori* lacking *Tipα* gene reduced the colonization levels of *H. pylori* in the stomach of mice (Godlewska et al. 2008); (2) Vaccination with Tipα significantly reduced colonization of *H. pylori* in mice associated with high levels of Tipα-specific antibody (Inoue et al. 2009); (3) Tipα protein is secreted from *H. pylori* but not mediated through Type IV secretion system (Suganuma et al. 2005); and (4) clinical isolates of *H. pylori* obtained from gastric cancer patients secreted Tipα protein in larger amounts than did *H. pylori* from patients with simple gastritis (Suganuma et al. 2008), strongly suggesting that Tipα plays an important role in *H. pylori*-induced inflammation and cancer development in human stomach (Balkwill 2009). All these features are different from those of other virulence factors, such as the cag pathogenicity island (cagPAI), CagA (cytotoxin associated antigen) and VacA (vacuolating cytotoxin A).

Members of the *Tipα* gene family include *Tipα* itself, *H. pylori-membrane protein 1* (*HP-MP1*), and *jph0543*, and these do not have any obvious homologues in other species (Suganuma et al. 2005; Yoshida et al. 1999). Tipα protein consists of 172 amino acids with a molecular weight of 19 kDa, and it forms a homodimer via two disulfide bonds with two cysteine residues in the N-terminal region. We previously reported that homodimer formation of Tipα is essential for induction of *TNF-α* gene expression in gastric epithelial cells (Suganuma et al. 2008) and also for transformation of Bhas 42 (v-H-*ras* transfected BALB/3T3) cells (Suganuma et al. 2005). To extend our experiments, we made two inactive Tipα mutants: a deletion mutant of Tipα (del-Tipα) that deleted six amino acids including two cysteine residues from native Tipα, and C5A/C7A double mutant (C5A/C7A-Tipα), two cysteine residues of Tipα are replaced by two alanines. The two mutated Tipα proteins induced *TNF-α* gene expression less strongly than native Tipα did (Suganuma et al. 2008). The crystal structures of del-Tipα and truncated forms of Tipα were recently reported by three independent groups, which revealed that they take dimerized forms, although they do not have the full length of protein (Jang et al. 2009; Tosi et al. 2009; Tsuge et al. 2009). If so, it is understandable that del-Tipα has weak activity.

We also found that fluorescence-labeled Tipα specifically binds to the surface of MGT-40 cells and enters into the cytosol and nuclei, whereas del-Tipα and C5A/C7A-Tipα bind weakly to the cells (Suganuma et al. 2008). In the light of this evidence, we think that homodimers of Tipα can easily bind to a specific receptor molecule on the cell surface of gastric epithelial cells. We identified nucleolin as a specific receptor of Tipα on the cell surface using pull-down assay with anti-FLAG antibody against FLAG-tagged Tipα protein. Nucleolin is a well-known major non-ribosomal protein consisting of 710 amino acids in nucleolus, and it has three different structural domains: an N-terminal domain containing highly acidic residues, a central domain containing four RNA recognition motifs, and a C-terminal domain containing Arg-Gly-Gly (RGG) repeats (Ginisty et al. 1999). Nucleolin is known to have multi-functions, including chromatin remodeling, DNA recombination, DNA replication, RNA transcription by RNA polymerase I and II, rRNA processing, mRNA stabilization, cytokinesis and apoptosis (Ginisty et al. 1999; Storck et al. 2007). Furthermore, recent evidence indicates that nucleolin is present on the surface of a wide range of cancer cells and some other types of the cells, and acts as receptors for several molecules (Hirano et al. 2005; Hoja-Lukowicz et al. 2009; Hovanessian et al. 2000; Legrand et al. 2004; Reyes-Reyes and Akiyama 2008). To investigate the specific interaction of nucleolin with Tipα, we conducted experiments to characterize localization of nucleolin, and studied the internalization of Tipα and subsequent induction of *TNF-α* gene expression. This paper reports for the first time that nucleolin is clearly involved in the carcinogenic process of *H. pylori* as a major cellular receptor of Tipα protein.

Materials and methods

Cell culture and reagents

Mouse gastric cancer cell line MGT-40 was maintained in DMEM with 10% fetal bovine serum (JRH Bioscience) and MITO+ serum extender (Becton–Dickinson Labware), as described previously (Ichinose et al. 1998). Human gastric cancer cell line MKN-1 and human monocytic leukemia cell line THP-1 were grown in RPMI 1640 medium with 10% fetal bovine serum. Anti-Tipα antibody was raised in rabbits by immunizing a synthetic peptide of 19 amino acids (from 11 to 29) of Tipα, and anti-nucleolin antibody (anti-NUC295) was raised in rabbits by immunizing a syntheic peptide of eight amino acids (from 295 to 302) of nucleolin, as described previously (Hirano et al. 2005; Suganuma et al. 2005). Other anti-nucleolin antibodies were purchased from Santa Cruz Biotethnology and Bethyl Laboratories, Inc. Anti-HSP90, anti-epidermal growth factor (EGF) receptor, anti-TNF receptor 2 and anti-lamin B antibodies were purchased from Santa Cruz Biotethnology. Anti-FLAG antibody was obtained from Sigma.

Preparation of three different *Tipα* genes tagged with FLAG

Three genes encoding Tipα-FLAG, del-Tipα-FLAG and C5A/C7A-Tipα-FLAG were obtained by PCR of pET28(a)$^+$-Tipα (Suganuma et al. 2005) containing oligonucleotide primers: Tipα-FLAG_F (5′-AGAGCATATGCT GCAGGCTTGCACTTGCCC) and Tipα-FLAG_R (5′-GG ATCCTACTTATCGTCGTCATCCTTGGTAGTCCATG GCTATAGG), del-Tipα-FLAG_F (5′-AGAGCATATGC CAAACACTTCACAAAGGAA), del-Tipα-FLAG_R (5′-G GATCCTACTTATCGTCGTCATCCTTGGTAGTCCAT GGCTATAGG), and C5A/C7A-Tipα-FLAG_F (5′-CA GCCATATGCTGCAGGCTGCCACTGCCCCAAACA C), C5A/C7A-Tipα-FLAG_R (5′-GGATCCTACTTATCG TCGTCATCCTTGGTAGTCCATGGCTATAGG). Three amplified fragments were separately cloned into a pET28(a)$^+$ expression vector (Invitrogen).

Preparation of three FLAG-tagged Tipα proteins

Each FLAG-tagged protein was expressed in *E. coli* (DE3) transfected pET28(a)$^+$ expression vector containing each of the corresponding genes above mentioned. They were induced with isopropyl-β-D-thiogalactopyranoside, and then purified by Ni^{2+} chelating resin (Ni Sepharose 6 Fast Flow, GE Healthcare), as reported previously (Suganuma et al. 2005). Tipα-FLAG, del-Tipα-FLAG and C5A/C7A-Tipα-FLAG all carry a tag of six histidines at the N-terminal region and also a FLAG-tag at the C-terminal region (Fig. 1a). All three recombinant Tipα proteins were more than 98% pure on SDS-PAGE. To conduct Ni^{2+} affinity pull-down assay, His-tag-removed Tipα-FLAG and His-tag-removed C5A/C7A-Tipα-FLAG proteins were prepared as follows: His-tagged Tipα-FLAG and C5A/C7A-Tipα-FLAG proteins were cleaved at a thrombin cleavage cite

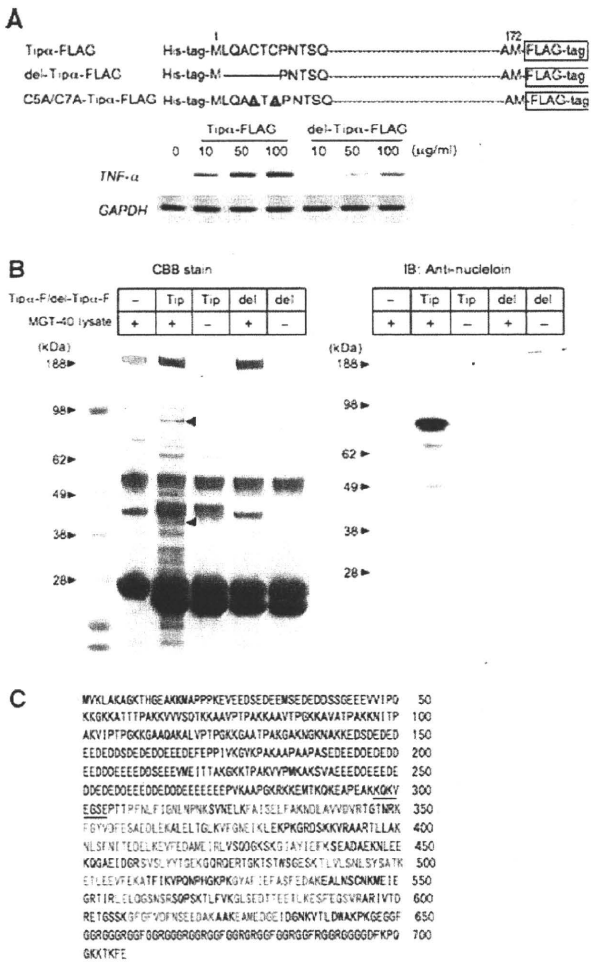

Fig. 1 Identification of nucleolin as Tipα binding protein. **a** Schematic representation of Tipα-FLAG, del-Tipα-FLAG and C5A/C7A-Tipα-FLAG proteins (*top*). Induction of *TNF-α* gene expression with Tipα-FLAG and with del-Tipα-FLAG in MGT-40 cells (*bottom*). Total RNAs were isolated from MGT-40 cells 1 h after treatment with Tipα-FLAG and with del-Tipα-FLAG, and the levels of TNF-α and GAPDH mRNAs were determined by semi-quantitative RT-PCR, as described in Materials and methods. **b** Representative results of FLAG pull-down assay. After incubation of MGT-40 cell lysates with Tipα-FLAG (Tip) and with del-Tipα-FLAG (del), Tipα-FLAG and del-Tipα-FLAG were immunoprecipitated with anti-FLAG antibody. The polypeptides that co-immunoprecipitated with Tipα-FLAG and with del-Tipα-FLAG were resolved in 4–12% NuPAGE and then stained with Quick CBB (*left panel*) and immunoblotted with anti-nucleolin antibody (IB: *right panel*). **c** Amino acid sequence of mouse nucleolin. Amino acids with *red characters* are assigned to the sequences determined by LC–MS analysis. *Underlined sequences* are recognition sites of anti-NUC295

using Thrombin cleavage capture kit (Novagen). Then the cleaved His-tag-peptide and uncleaved protein were separated using Ni^{2+} chelating resin (Tsuge et al. 2009).

Preparation of His-tagged nucleolin protein fragment

His-tagged nucleolin gene fragment (*NUC284*), containing both residues from 284 to 710 of human nucleolin and C-terminal His-tag, was expressed in *E. coli* transfected pBAD/Thio-E/NUC284 expression vector, and purified by Ni^{2+} chelating resin, as described previously (Hirano et al. 2005).

Expression of *TNF-α* gene

MGT-40 and THP-1 cells were treated with recombinant protein for 1 h, and total RNAs obtained from the cells were isolated with ISOGEN reagent (Nippon Gene). Expressions of *TNF-α* gene and *glyceraldehyde-3-phosphate dehydrogenase* (*GAPDH*) gene as a control were determined by both semi-quantitative RT-PCR and real-time RT-PCR, as described previously (Suganuma et al. 2005). The values are expressed as the average of three separate experiments.

FLAG pull-down assay

MGT-40 cell lysates were prepared with NP-40 lysis buffer containing 20 mM Tris–HCl (pH 7.4), 100 mM NaCl, 0.5% NP-40, 10% glycerol, 1 mM PMSF, 1 μg/ml aprotinin, and 1 μg/ml leupeptin, and the lysates (600 μg/ml) were incubated with Tipα-FLAG (200 μg/ml) and del-Tipα-FLAG (200 μg/ml) in buffer A containing 50 mM Tris–HCl (pH 7.4), 150 mM NaCl, 1 mM PMSF, 1 μg/ml aprotinin, and 1 μg/ml leupeptin at 4°C for 2 h. After addition of 20 μl anti-FLAG M2 Gel (Sigma), the mixture was further incubated at 4°C for 2 h, and then the resin was washed with buffer A containing 1% Triton X-100. The polypeptides associated with the resin were resolved in 4–12% NuPAGE (Invitrogen), and were determined using staining with Quick CBB (Wako). The control experiments were similarly conducted without using Tipα-FLAG and del-Tipα-FLAG.

LC–MS analysis

Gel sections containing polypeptides co-precipitated with Tipα-FLAG were subjected to proteolysis with 2 μg/ml trypsin (Wako) at 25°C overnight, and the digestion was stopped by adding an elution solution (50% acetonitrile, 5% formic acid). Each sample was analyzed using Nano-ESI-Ion trap MS (HCT plus, Bruker Daltonics), according to manufacturer's instruction (Bruker application note). The data were analyzed by a protein database search on MASCOT (Matrix Science).

Ni^{2+} affinity pull-down assay

His-tagged nucleolin fragment (NUC284) was incubated with His-tag-removed Tipα-FLAG or His-tag-removed C5A/C7A-Tipα-FLAG in NP-40 lysis buffer containing 10 mM imidazole at 4°C for 2 h. Twenty microlitre of Ni^{2+} chelating resin was then added to the mixture, which was further incubated at 4°C for 2 h. After washing the resin with NP-40 lysis buffer containing 40 mM imidazole, the complex of nucleolin fragment (NUC284) with Tipα-FLAG or with C5A/C7A-Tipα-FLAG were determined by Western blotting using anti-nucleolin (H-250, Santa Cruz) and anti-Tipα antibodies, respectively.

Incorporation of Tipα into cells

MGT-40 cells were treated with Tipα, and then lysed in lysis buffer containing 20 mM Tris–HCl (pH 8.0), 150 mM NaCl, 1% Triton X-100, 0.1% SDS, 1% sodium deoxycholate, 1 mM PMSF, 1 μg/ml aprotinin, and 1 μg/ml leupeptin. Cell lysates were resolved in 12% SDS-PAGE. Incorporation of Tipα into the cells was determined by Western blotting using anti-Tipα antibody (Suganuma et al. 2008).

Analysis of subcellular fractionation

Homogenates of MGT-40 and THP-1 cells were fractionated into membrane, cytosol, and nuclei using Qproteome cell compartment kit (Qiagen), according to the manufacturer's instruction. Each fraction was subjected to Western blotting, using anti-nucleolin, anti-HSP90 (a marker for cytosol), anti-EGFR or anti-TNF receptor 2 (for membrane) and anti-lamin B antibodies (for nuclei).

Immunoprecipitation

MKN-1 and THP-1 cells were treated with Tipα-FLAG and del-Tipα-FLAG at a concentration of 100 μg/ml at 37°C for 1 h, and then lysed as described above. Cell lysates (about 400 μg) were incubated with anti-nucleolin antibody (A300-711A, Bethyl Lab, Inc.) at 4°C for 2 h. The immunocomplex was captured with protein A sepharose (GE Healthcare) at 4°C overnight, and then washed with NP-40 lysis buffer. The immunocomplex was applied to 12% SDS-PAGE. Tipα-FLAG, del-Tipα-FLAG and nucleolin were determined by Western blotting using anti-FLAG and anti-nucleolin antibodies (MS-3, Santa Cruz).

Flow cytometry

MGT-40 and THP-1 cells (1×10^6 cells/ml) in PBS were incubated with 2 μg/ml anti-NUC295 antibody and 10 μg/ml Alexa Fluor 488-conjugated goat anti-rabbit IgG (Invitrogen) on ice for 30 min. Then cells were subjected to flow cytometry (Epics XL, Beckman Coulter).

Statistical analysis

The data were compared using Student's *t* test.

Results

Identification of nucleolin as a Tipα-binding protein

To characterize the nature of the specific binding protein for Tipα, Tipα tagged with FLAG at C-terminus (Tipα-FLAG) and del-Tipα tagged with FLAG at C-terminus (del-Tipα-FLAG)—the latter with six amino acids deleted including two cysteine residues from N-terminal region of Tipα—were used for the experiments (Fig. 1a). Tipα-FLAG protein induced *TNF-α* gene expression in mouse gastric cancer cells (MGT-40), while del-Tipα-FLAG was over ten times weaker than Tipα-FLAG. Thus, Tipα-FLAG and del-Tipα-FLAG showed the same biological activity as did recombinant Tipα and del-Tipα (Fig. 1a).

The mixtures of MGT-40 cell lysates with Tipα-FLAG and with del-Tipα-FLAG were separately subjected to pull-down assay using resin conjugated with anti-FLAG antibody. Thirteen polypeptide bands on SDS-PAGE were found to be co-precipitated with Tipα-FLAG, but not with del-Tipα-FLAG (Fig. 1b). Each polypeptide band was subjected to LC–MS analysis after tryptic digestion, and it turned out that the amino acid sequences of two polypeptides, with 88 and 40 kDa, were similar to that of mouse nucleolin, as shown in Fig. 1c. The results showed that the polypeptide with 88 kDa is nucleolin and the other polypeptide with 40 kDa is a fragment of nucleolin. Three polypeptides with less than 40 kDa were derived from Tipα, and another polypeptide with less than 50 kDa was identical to ribosomal protein L4 fragment; the others could not be confirmed by LC–MS.

The polypeptide with 88 kDa was further confirmed to be nucleolin using immunoblot analysis with anti-nucleolin antibody, but the polypeptide with 40 kDa did not react with anti-nucleolin antibody (Fig. 1b), probably because the latter peptide did not contain recognition sites of the antibody. Although several polypeptides with 50–70 kDa reacted with anti-nucleolin antibody, we think that there were degradation fragments of nucleolin co-precipitated with Tipα-FLAG. The results strongly suggest that nucleolin acts as a specific binding protein of Tipα.

Interaction of incorporated Tipα with endogenous nucleolin in the cells

The binding of Tipα to nucleolin at cellular levels was examined by immunoprecipitation using anti-nucleolin antibody. Since the affinity of the anti-nucleolin antibody for human nucleolin is higher than that for mouse nucleolin, we used cell lysates of both human gastric cancer cell lines MKN-1 and human monocytic leukemia cell line THP-1 for the experiments. Significant amounts of Tipα-FLAG interacted with the lysates of MKN-1 and THP-1 cells, but the amounts of del-Tipα-FLAG interacted less with their cell lysates, which shows that both Tipα-FLAG and del-Tipα-FLAG were incorporated into the cells (Fig. 2a). Using anti-nucleolin antibody, these cell lysates were further subjected to immunoprecipitation: nucleolin was immunoprecipitated with anti-nucleolin antibody associated with Tipα-FLAG but not del-Tipα-FALG in both MKN-1 and THP-1 cells (Fig. 2a). These results suggest that Tipα directly binds to native and endogenous nucleolin in the cells, and the differences of binding ability between Tipα-FLAG and del-Tipα-FLAG to nucleolin are comparable to their inducing potencies of *TNF-α* gene expression.

Direct interaction of nucleolin with Tipα

Next, we studied whether Tipα directly binds to recombinant human nucleolin fragment NUC284, which consists of amino acids from 284 to 710 containing four RNA binding domains. His-tag-removed Tipα-FLAG was incubated in vitro with NUC284 fragment and Ni^{2+} chelating resin, and we found that Tipα-FLAG significantly co-precipitated with NUC284 fragment, although small amounts of Tipα-FLAG precipitated with Ni^{2+} chelating resin only (Fig. 2b). However, His-tag-removed C5A/C7A-Tipα-FLAG did not co-precipitate with NUC284 fragment, suggesting that the homodimer form of Tipα is necessary for direct binding to nucleolin: We think that the homodimer of Tipα directly binds to two-thirds of C-terminal nucleolin, without any scaffold proteins.

Cell surface localization of nucleolin on MGT-40 cells

We previously reported that FITC-labeled Tipα specifically binds to the cell surface of MGT-40 cells (Suganuma et al. 2008). Since nucleolin is present on both nucleolus and surface of the cells (Barel et al. 2008; Hirano et al. 2005; Hoja-Lukowicz et al. 2009: Hovanessian et al. 2000; Legrand et al. 2004; Reyes-Reyes and Akiyama 2008),

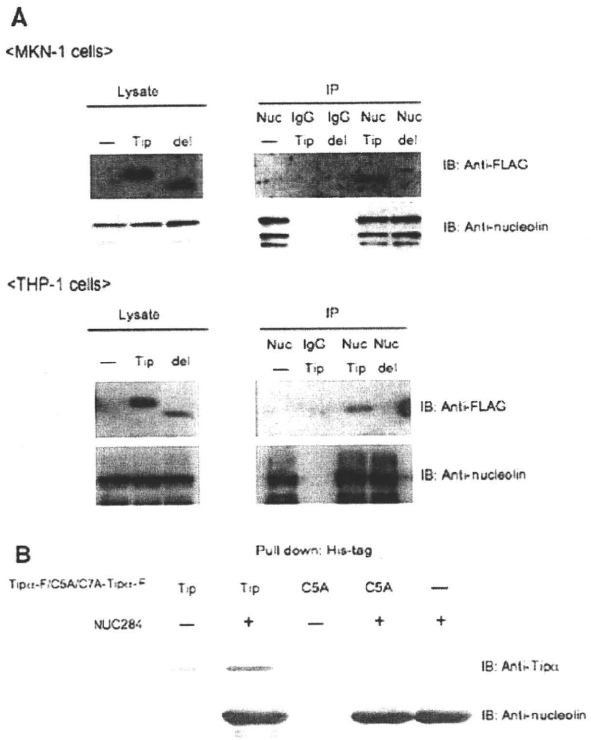

Fig. 2 Direct interaction of nucleolin with Tipα. **a** Tipα was immunoprecipitated with endogenous human nucleolin in MKN-1 and THP-1 cells. MKN-1 and THP-1 cells were treated with 100 μg/ml Tipα-FLAG (Tip) and with del-Tipα-FLAG (del) at 37°C for 1 h. Tipα-FLAG and del-Tipα-FLAG significantly incorporated into the cells (*left panels*). Each cell lysate was immunoprecipitated with anti-nucleolin antibody (NUC) and with rabbit IgG (as a control, IgG). Immunoprecipitates were resolved in 12% SDS-PAGE and immunobotted (IB) with anti-FLAG antibody and anti-nucleolin antibody (*right panels*). **b** Direct interaction of recombinant human nucleolin fragment with Tipα in vitro. His-tag removed Tipα-FLAG (Tip) and His-tag removed C5A/C7A-FLAG (C5A), which were prepared as described in Experimental procedures, were incubated with a 6-His-tag fused recombinant human nucleolin fragment containing 284–710 amino acid residues (NUC284) and then subjected to pull-down assay using Ni^{2+} chelating resins. The precipitates were resolved in 12% SDS-PAGE and analyzed by Western blotting with anti-Tipα antibody and with anti-nucleolin antibody

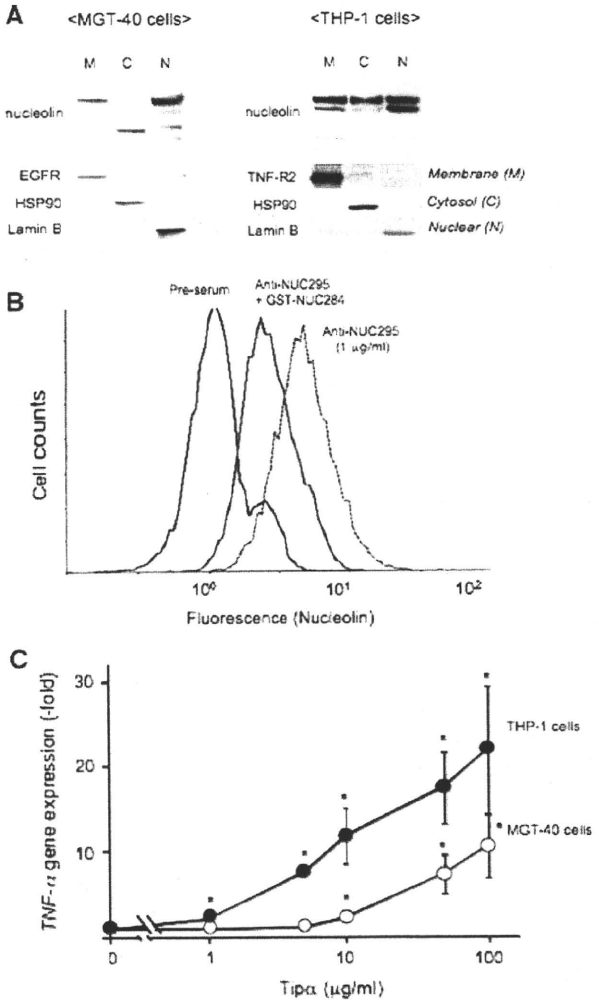

Fig. 3 Localization of nucleolin on cell surface of MGT-40 and THP-1 cells. **a** Subcellular localization of nucleolin analyzed by cell fractionation. MGT-40 and THP-1 cells were fractionated into membrane (M), cytosolic (C) and nuclear (N) fractions, and each fraction was immunoblotted with anti-nucleolin antibody. Each fraction was confirmed by Western blotting with antibodies for fractionation-marker proteins: EGFR for membrane of MGT-40 cells, TNF-R2 for membrane of THP-1 cells, HSP90 for cytosol and lamin B for nuclei. **b** Detection of nucleolin on cell surface shown by flow cytometry. MGT-40 cells were incubated with 1 μg/ml anti-NUC295 (Anti-NUC295) and with pre-immune serum (Pre-serum) as a control in the presence of 10 μg/ml Alexa Fluor 488-conjugated goat rabbit IgG on ice for 30 min. Preincubation of Anti-NUC295 with recombinant nucleolin fragment (Anti-NUC+GST-NUC284) significantly reduced fluorescence. **c** Strong induction of *TNF-α* gene expression with Tipα in THP-1 cells (*filled circle*) and MGT-40 cells (*open circle*). One hour after treatment with Tipα at various concentrations, expression of *TNF-α* and *GAPDH* genes was determined by semi-quantitative RT-PCR. Relative expression of *TNF-α* gene is shown as fold change compared with control after normalization of GAPDH mRNA levels. The results are the averages of three independent experiments. *Bars* indicate standard deviation. Statistical levels between non-treated and Tipα-treated cells were shown to be significant $*P < 0.01$

we first confirmed the sub-cellular localization of nucleolin in MGT-40 cells and THP-1 cells. Although most of the nucleolin was present in the nuclear fraction of MGT-40 cells, significant small amounts of nucleolin were found in membrane and cytosol fractions (Fig. 3a), while THP-1 cells showed large amounts of nucleolin in the membrane fraction (Fig. 3a). Moreover, nucleolin localized on cell surface was further determined by flow cytometry using anti-nucleolin antibody (anti-NUC295). A study of the MGT-40 cells using flow cytometry revealed that the fluorescent peak dramatically shifted to a high fluorescent peak by treatment with anti-NUC295 antibody, and that

this fluorescent peak was significantly reduced by preincubation of anti-NUC295 with recombinant GST-nucleolin fragment (GST-NUC284) containing amino acids from 284 to 710 (Fig. 3b). This indicated that nucleolin on the cell surface had interacted with anti-NUC295. Further, we found that Tipα induced dose-dependently *TNF-α* gene expression in MGT-40 and THP-1 cells, based on the results that nucleolin localized on cell surface of both cells (Fig. 3c). The relationship between the amounts of cell surface nucleolin and the potency of Tipα on *TNF-α* gene expression will be reported elsewhere.

Effects of anti-NUC295 on *TNF-α* gene expression induced by Tipα

We studied how anti-NUC295 antibody affects the induction of *TNF-α* gene expression in MGT-40 cells treated with Tipα. First, the treatment with rabbit IgG and anti-nucleolin H-250 antibodies—the latter of which does not recognize nucleolin on cell surface—did not affect the levels of *TNF-α* gene expression induced by Tipα. However, treatment with anti-NUC295 antibody dose-dependently enhanced the *TNF-α* gene expression induced by Tipα up to twofold (Fig. 4a), and treatment with anti-NUC295 antibody dose-dependently enhanced incorporation of Tipα into the cytosol of MGT-40 cells (Fig. 4b). From our results showing that anti-NUC295 antibody internalized into MGT-40 cells as determined by flow cytometry (data not shown), we think that the complex of nucleolin, Tipα and anti-NUC295 internalized into the cells and then induced *TNF-α* gene expression.

Inhibitory effects of down-regulated cell surface nucleolin on biological activity of Tipα

Nucleolin on the cell surface is a glycoprotein containing *N*- and *O*-glycans (Carpentier et al. 2005), and the N-glycosylation of nucleolin is essential for localization on cell surface (Losfeld et al. 2009). We found that treatment of MGT-40 cells with 5 μg/ml tunicamycin, an inhibitor of the N-linked glycosylation of protein, significantly reduced the amounts of nucleolin on the cell surface as determined by flow cytometry (Fig. 5a): The levels of cell surface nucleolin were reduced by approximately 50%. Moreover, pretreatment with tunicamycin inhibited about 50% *TNF-α* gene expression induced by Tipα because tunicamycin reduced the incorporated amounts of Tipα into MGT-40 cells (Fig. 5b, c). The reduced amounts of nucleolin correlated well with reduction of *TNF-α* gene expression. Cell surface nucleolin is thus a functional receptor of Tipα associated with incorporation of Tipα into the cells and subsequent *TNF-α* gene expression.

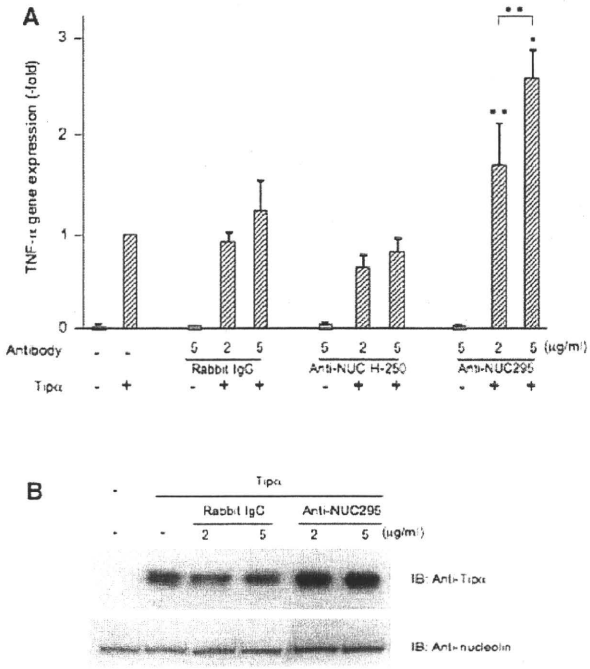

Fig. 4 Significant enhancement of Tipα-induced *TNF-α* gene expression and Tipα incorporation in cells induced by anti-NUC295. **a** MGT-40 cells were previously incubated with rabbit IgG, with anti-NUC H-250 and anti-NUC295 antibodies at 4°C for 1 h, and further treated with 50 μg/ml Tipα at 37°C for 1 h. Relative *TNF-α* gene expression is shown as fold change compared with that of cells treated with 50 μg/ml Tipα after normalization of *GAPDH* gene expression levels. The results are the averages of three independent experiments. *Bars* indicate standard deviation. Statistical significance of effects of anti-NUC295 in TNF-α induction by Tipα compared with non-treated were shown as $*P < 0.01$ and $**P < 0.05$, and the difference between 2 and 5 μg/ml of anti-NUC295 was significant at the level of $**P < 0.05$. **b** Incorporation of Tipα was determined by Western blotting with anti-Tipα antibody. Nucleolin levels were also determined by anti-nucleolin antibody

Discussion

Considering our 1993 discovery that TNF-α is an endogenous tumor promoter in carcinogenesis (Komori et al. 1993), we first cloned a new gene of TNF-α inducing protein from *H. pylori* genome (Suganuma et al. 2005). We also reported that the active form of Tipα is a homo-dimer that induces *TNF-α* gene expression in the cells, resulting in a cancer microenvironment (Suganuma et al. 2006, 2008). Furthermore, Tipα is now widely accepted as a carcinogenic factor of *H. pylori* (Balkwill 2009). This paper reports that Tipα directly binds to nucleolin on the cell surface, and that the complex of Tipα with nucleolin then internalizes into the cells. The results suggest that cell surface nucleolin acts as a receptor of Tipα: nucleolin is mainly localized in the nucleolus, but significant amounts are present on the cell surface, including various cancer

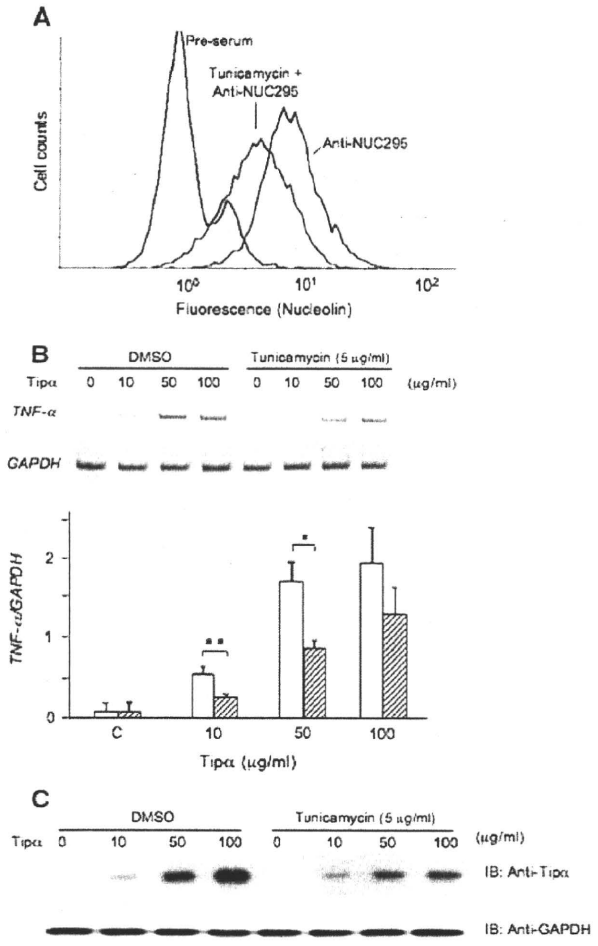

Fig. 5 Inhibition of Tipα-induced *TNF-α* gene expression and Tipα incorporation in cells induced by down-regulation of cell surface nucleolin. **a** MGT-40 cells were treated with or without 5 μg/ml tunicamycin in DMSO at 37°C for 5 h. The cell surface nucleolin on MGT-40 cells was visualized using flow cytometry with anti-NUC295, as described in Materials and methods. **b** Inhibition of *TNF-α* gene expression induced by Tipα after pretreatment with tunicamycin. After pretreatment of MGT-40 cells with or without 5 μg/ml tunicamycin in DMSO at 37°C for 5 h, the cells were treated with various concentrations of Tipα at 37°C for 1 h. The levels of *TNF-α* and *GAPDH* gene expression in MGT-40 cells were determined by semi-quantitative RT-PCR. The results are the averages of three independent experiments. *Bars* indicate standard deviation. Statistical levels were significant $*P < 0.01$ and $**P < 0.05$. **c** Inhibition of Tipα-incorporation into MGT-40 cells by pretreatment with tunicamycin. After treatment with tunicamycin, both MGT-40 cells treated with various concentrations of Tipα and cell lysates were resolved in 12% SDS-PAGE solution and analyzed by Western blotting with anti-Tipα antibody and anti-nucleolin antibody (IB)

cells and proliferating cells (Hirano et al. 2005; Hoja-Lukowicz et al. 2009; Hovanessian et al. 2000; Legrand et al. 2004; Reyes-Reyes and Akiyama 2008). It is well-known that cell surface nucleolin has an important role as a receptor for various extracellular ligands, including human immunodeficiency virus (HIV) particles (Nisole et al. 2002), midkine (Hovanessian 2006; Said et al. 2002), and elongation factor-TU of *Francisella tularensis* (Barel et al. 2008), lactoferrin (Legrand et al. 2004), and endostatin (Shi et al. 2007): Nucleolin acts as a shuttling molecule between cell surface, cytoplasm, and nucleus (Borer et al. 1989). Moreover, it is of interest to note that a specific DNA aptamer of nucleolin, AS1411, is the most well-investigated anti-cancer aptamer, which initially binds to cell surface nucleolin and then internalizes into the cells (Ireson and Kelland 2006). Based on evidence, we think that nucleolin shuttles Tipα, which is supported by the results that some Tipα are present in the nuclei of MGT-40 cells after treatment with Tipα protein. Although the precise function of Tipα in nucleus is not well understood, we found that Tipα directly binds to DNA oligomers in Biacore assay (Kuzuhara et al. 2007). Thus, our understanding on the Tipα function is extended by the several findings, such as nucleolin as the receptor, the translocation of Tipα into the nuclei, and the induction of *TNF-α* gene expression in the cells.

Although it is not well-known how nucleolin translocates across the membrane and how it attaches to the cell surface, glycosylation is assumed to be an essential biochemical modification for nucleolin to localize on the cell surface (Losfeld et al. 2009). In our experiments, tunicamycin significantly reduced the level of cell surface nucleolin in MGT-40 cells—although the nucleolin in nucleoli was not much reduced—and then inhibited both internalization of Tipα and *TNF-α* gene expression. But it is still not clear whether N-glycosylation of nucleolin is involved in Tipα binding because Tipα directly binds to recombinant nucleolin fragment (NUC284) without any glycosylation. Since NUC284 fragment contains four RNA binding domains (RBDs) and a RGG domain—which are well conserved in human, mouse, and rat (Ginisty et al. 1999)—we think that Tipα binds to one of these domains. To understand more precisely the nature of Tipα and nucleolin binding, we used anti-nucleolin antibody (Anti-NUC295) (Hirano et al. 2005): pretreatment with anti-NUC295 unexpectedly enhanced internalization of Tipα and induction of *TNF-α* gene expression by Tipα, indicating that the Tipα binding site of nucleolin is different from epitope of nucleolin and that anti-NUC295 enhances internalizing of nucleolin with Tipα. These findings support the previously reported results with another anti-nucleolin antibody (mAb D3), which recognizes cell surface nucleolin and induces clustering and internalization of nucleolin together with mAb D3 antibody (Hovanessian et al. 2000). Since, we obtained the results that pretreatment with methyl-β-cyclodextrin, which inhibits endocytosis by depletion of cholesterol from membrane, inhibited induction of *TNF-α* gene expression with Tipα (unpublished results), we think that anti-NUC295

enhances endocytosis of Tipα. To prove evidence that nucleolin acts as a specific receptor of Tipα, we conducted a knockdown experiment with shRNA using lentiviral vector, and the growth of THP-1 cells was inhibited, by the complete down-regulation of nucleolin (data not shown).

Several polypeptides co-precipitated with Tipα-FLAG, but did not with del-Tipα-FLAG, in pull-down assay. Although the specificity of the co-precipitation was relatively high, we found that ribosomal protein L4 is an additional binding protein. That is well-known to be a protein interacting with nucleolin. As for interaction of Tipα with nucleolin, the N-terminal portion of Tipα is thought to be an important domain: (1) disulfide bond formation in the N-terminal of Tipα is essential for the interaction, and monomer of del-Tipα does not bind to nucleolin, and (2) we successfully identified nucleolin as the Tipα-binding protein because we used FLAG-tagged at the C-terminal position of Tipα as bait, but did not use the His-tagged at the N-terminal position.

The TNF-α inducing activity of Tipα should be briefly mentioned in connection with *H. pylori*-infection in human stomach cancer development. Tipα-deficient *H. pylori* reduces colonization in mouse gastric mucosa (Godlewska et al. 2008), and vaccinations with Tipα and del-Tipα also effectively prevented colonization of *H. pylori* in the stomach of mice (Inoue et al. 2009). Therefore, we think that targeting molecules, which inhibit the interaction of Tipα and cell surface nucleolin, will be useful tools for the prevention of inflammation induced by *H. pylori* infection and of *H. pylori*-infection itself. For example, lactofferin, which binds to nucleolin (Legrand et al. 2004), is effective in suppression of *H. pylori* colonization (Okuda et al. 2005), which suggests that lactofferin inhibits the binding of Tipα and nucleolin.

How nucleolin is involved in the induction of *TNF-α* gene expression induced by Tipα is an important subject, since TNF-α is a major mediator of cancer-related inflammation in the cancer microenvironment (Balkwill 2009; Komori et al. 1993; Suganuma et al. 1999). A specific DNA aptamer of nucleolin, AS1411 (Ireson and Kelland 2006; Soundararajan et al. 2008) blocks both TNF-α induced- and constitutive-NF-κB activation in human cancer cell lines by forming a complex of nucleolin with an NF-κB essential modulator (NEMO) (Girvan et al. 2006). This indicates that nucleolin regulates NF-κB activation through interaction with NEMO, so it is possible that Tipα incorporated with nucleolin interferes in the interaction of nucleolin with NEMO and affects NF-κB signaling.

Tipα family genes and protein products show carcinogenic activity in combination with v-H-*ras* oncogene: Transfection of *HP-MP1* gene into Bhas 42 cells (v-H-ras transfected BALB/3T3) induces highly malignant transformed cells (Bhas/mp-1): these cells have strong tumorigenicity associated with a high grade of angiogenesis in nude mice (Suganuma et al. 2001). Interestingly, it has been reported that overexpression of nucleolin cooperates with oncogenic mutant Ras in a rat embryonic fibroblast transformation assay (Takagi et al. 2005). And midkine and pleiotrophin, which are ligands of nucleolin, transformed cells (Muramatsu 2002). Nucleolin is specifically expressed on the cell surface in proliferating endothelial cells (Shi et al. 2007), and it is also well-known that nucleolin protein is expressed at high levels on the cell surface of rapid proliferation cells, including cancer cells such as MCF-7 (breast cancer) (Soundararajan et al. 2008), HeLa (cervical cancer) (Li et al. 2009), colo-320 (colon adenocarcinoma) (Reyes-Reyes and Akiyama 2008) and THP-1 cells (Barel et al. 2008; Hirano et al. 2005). We also found that nucleolin is expressed on the surface of mouse and human gastric cancer cell lines (MGT-40 and MKN-1). The study on the expression levels and localization of cell surface nucleolin during development of gastric cancer by *H. pylori*-infection will surely provide a new insight into the identification of high-risk *H. pylori* carriers in more detail. Since 50% of the world population is infected with *H. pylori* (Snaith and El-Omar 2008), our results with Tipα indicate that nucleolin on the cell surface will prove useful as a high-risk biomarker for gastric cancer. This paper is the first report that nucleolin serves as a receptor of Tipα, the carcinogenic factor of *H. pylori*: further results with a complex of Tipα with nucleolin will intensify the understanding of this new carcinogenic mechanism on gastric cancer development in humans.

Acknowledgments We thank Ms. Kaori Suzuki and Ikuko Shiotani, Research Institute for Clinical Oncology, Saitama Cancer Center for their technical assistance. We also thank Dr. Masae Tatematsu for supplying of MGT-40 cells. This work was supported by Japan Society for the Promotion of Science, and Smoking Research Fund.

References

Balkwill F (2009) Tumour necrosis factor and cancer. Nat Rev Cancer 9:361–371

Barel M, Hovanessian AG, Meibom K, Briand JP, Dupuis M, Charbit A (2008) A novel receptor—ligand pathway for entry of *Francisella tularensis* in monocyte-like THP-1 cells: interaction between surface nucleolin and bacterial elongation factor Tu. BMC Microbiol 8:145

Borer RA, Lehner CF, Eppenberger HM, Nigg EA (1989) Major nucleolar proteins shuttle between nucleus and cytoplasm. Cell 56:379–390

Carpentier M, Morelle W, Coddeville B, Pons A, Masson M, Mazurier J, Legrand D (2005) Nucleolin undergoes partial N- and O-glycosylations in the extranuclear cell compartment. Biochemistry 44:5804–5815

El-Omar EM, Carrington M, Chow WH, McColl KE, Bream JH, Young HA, Herrera J, Lissowska J, Yuan CC, Rothman N, Lanyon G, Martin M, Fraumeni JF Jr, Rabkin CS (2000)