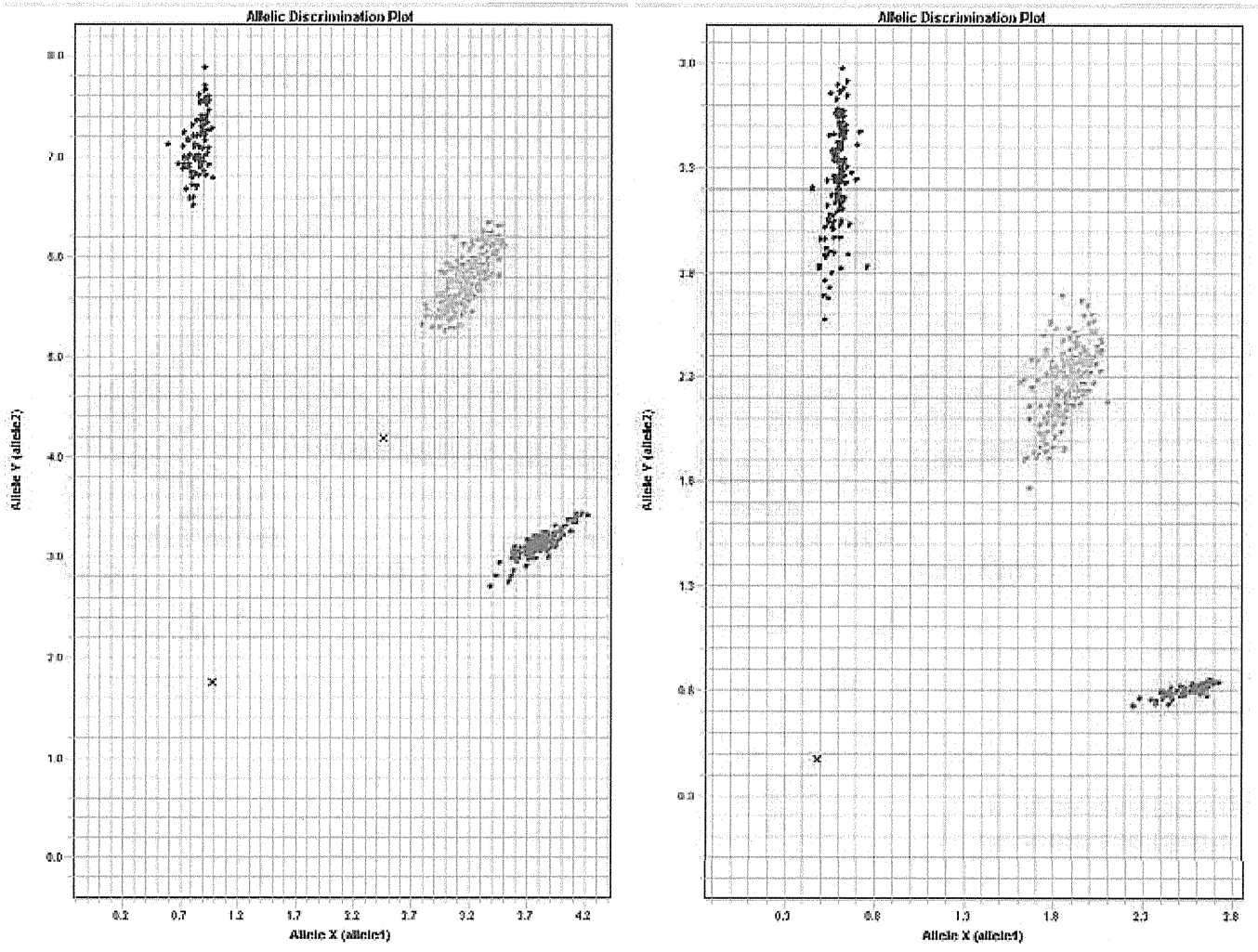
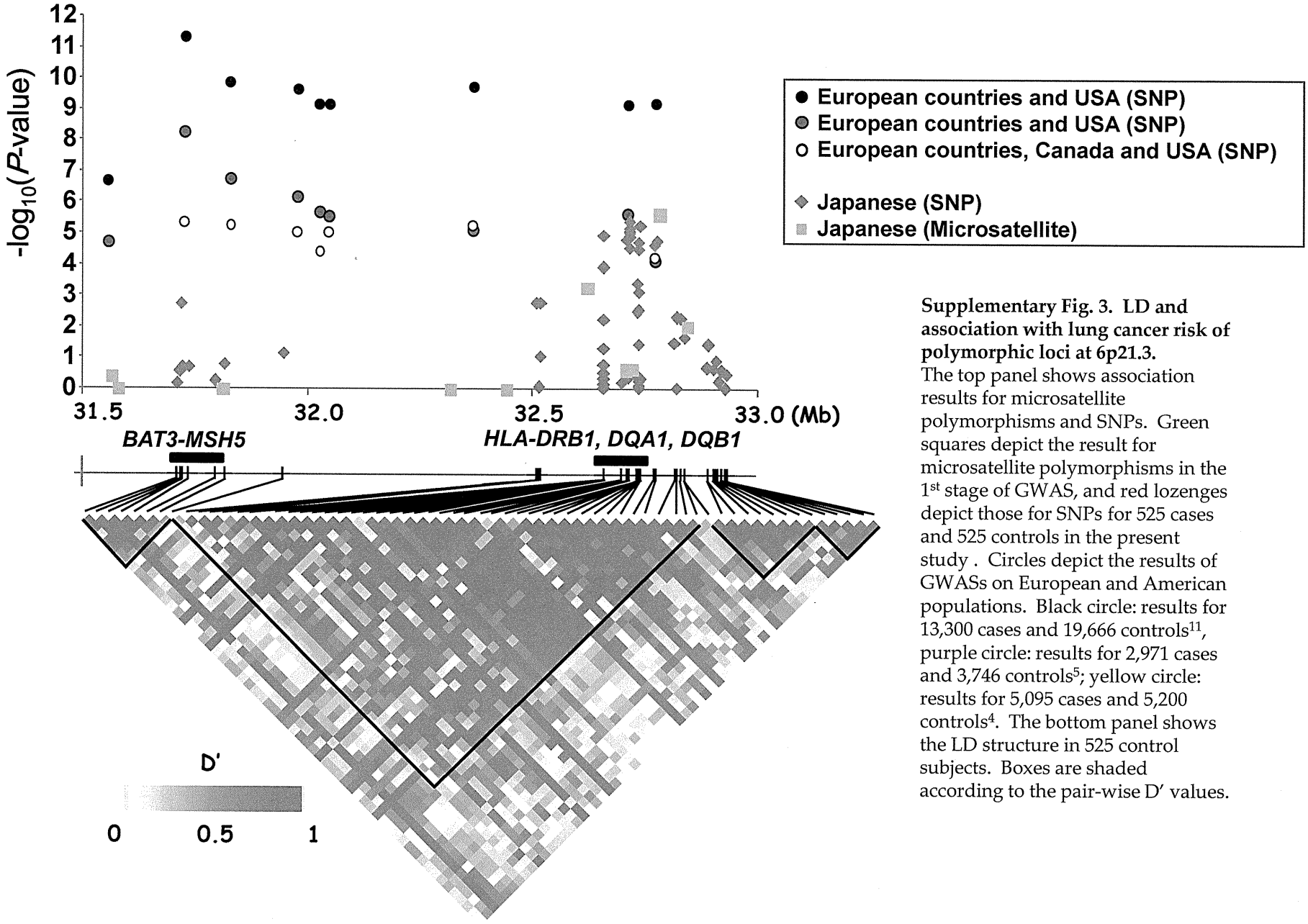
rs34843907 **rs17426593**



Supplementary Fig. 2. The Taqman cluster plots for the rs34843907 and rs17426593 SNPs.
Left: rs34843907. Red: homozygotes for the A allele; Green: heterozygotes; Blue: homozygotes for the C allele.
Right: rs17426593. Red: homozygotes for the C allele; Green: heterozygotes; Blue: homozygotes for the T allele.
Undetermined genotypes are labeled X.



Supplementary Fig. 3. LD and association with lung cancer risk of polymorphic loci at 6p21.3. The top panel shows association results for microsatellite polymorphisms and SNPs. Green squares depict the result for microsatellite polymorphisms in the 1st stage of GWAS, and red lozenges depict those for SNPs for 525 cases and 525 controls in the present study . Circles depict the results of GWASs on European and American populations. Black circle: results for 13,300 cases and 19,666 controls[11], purple circle: results for 2,971 cases and 3,746 controls[5]; yellow circle: results for 5,095 cases and 5,200 controls[4]. The bottom panel shows the LD structure in 525 control subjects. Boxes are shaded according to the pair-wise D′ values.

Oncogene (2009) 28, 2903–2909

www.nature.com/onc

ORIGINAL ARTICLE

Significance of RRM1 and ERCC1 expression in resectable pancreatic adenocarcinoma

H Akita[1], Z Zheng[2], Y Takeda[1], C Kim[1], N Kittaka[1], S Kobayashi[1], S Marubashi[1], I Takemasa[1], H Nagano[1], K Dono[1], S Nakamori[3], M Monden[1], M Mori[1], Y Doki[1] and G Bepler[2]

[1]*Department of Surgery, Osaka University Graduate School of Medicine, Osaka, Japan;* [2]*Department of Thoracic Oncology, H. Lee Moffitt Cancer Center and Research Institute, University of South Florida, Tampa, FL, USA and* [3]*Department of Surgery, National Hospital Organization Osaka National Hospital, Osaka, Japan*

The identification of molecular markers, useful for therapeutic decisions in pancreatic cancer patients, is crucial for advances in disease management. Gemcitabine, although a cornerstone of current therapy, has limited efficacy. *RRM1* is a key molecule for gemcitabine efficacy and is also involved in tumor progression. We determined *in situ* *RRM1* and *excision repair cross complementation group 1* (*ERCC1*) protein levels in 68 pancreatic cancer patients. All had R0 resections without preoperative therapy. Protein levels were determined by automated quantitative analysis (AQUA), a fluorescence-based immunohistochemical method. The relationship between protein expressions and clinical outcomes, including response to gemcitabine at the time of disease recurrence, was determined. Patients with high *RRM1* showed significantly better overall survival than patients with low expression ($P=0.0196$). There was a trend toward better overall survival for patient with high ERCC1 ($P=0.0552$). When both markers were considered together, patients with both high *RRM1* and *ERCC1* faired the best in terms of overall and disease-free survival ($P=0.0066$, $P=0.0127$). In addition, treatment benefit from gemcitabine in patients with disease recurrence was observed only in patients with low *RRM1*. The combination of *RRM1* and *ERCC1* expression is prognostic in pancreatic cancer patients after a complete resection. On disease recurrence, only patients with low *RRM1* derive benefit from gemcitabine.
Oncogene (2009) **28,** 2903–2909; doi:10.1038/onc.2009.158; published online 22 June 2009

Keywords: pancreatic cancer; *RRM1*; *ERCC1*; AQUA; prognosis; gemcitabine

Introduction

Pancreatic cancer is one of the leading causes of tumor-related mortalities. The prognosis of patients after complete resection is poor, and more than 50% of patients develop tumor recurrence at distant or locoregional sites, with an estimated 5-year survival of only 20% (Kayahara *et al.*, 1993; Nitecki *et al.*, 1995; Staley *et al.*, 1996; Sener *et al.*, 1999; Li *et al.*, 2004). The addition of chemotherapy and radiotherapy to surgical resection is important, and gemcitabine, a pyrimidine nucleotide analogue, has become the standard chemotherapeutic agent in such programs (Burris *et al.*, 1997; Oettle *et al.*, 2007) (Rothenberg *et al.*, 1996). However, the clinical response rate to gemcitabine remains modest, mainly because of the profound chemoresistance inherent in pancreatic cancer. The selection of patients who derive a true benefit from gemcitabine could be an important stepping stone toward improvement of outcome of pancreatic cancer.

RRM1, the gene that encodes the regulatory subunit of ribonucleotide reductase, is a key determinant of gemcitabine efficacy. In various cancers, we and others have described that overexpression of the *RRM1* gene is strongly associated with gemcitabine resistance (Cao *et al.*, 2003; Rosell *et al.*, 2004; Bergman *et al.*, 2005; Bepler *et al.*, 2006; Nakahira *et al.*, 2007). However, there is no clinical study that investigated the correlation between *RRM1* protein expression and gemcitabine resistance.

On the other hand, the expression of *RRM1* was also reported to correlate with the tumorigenic and metastatic potential of lung cancer (Gautam *et al.*, 2003), and an oncogenic ras-transformed cell line with high expression of an *RRM1* transgene had reduced metastatic potential (Fan *et al.*, 1997). Furthermore, high expression of *RRM1* in transgenic mice is associated with resistance to carcinogen-induced lung tumorigenesis (Gautam and Bepler, 2006). Recently, overexpression of *RRM1* and the *excision repair cross-complementation group 1* (*ERCC1*) gene product was reported to correlate with favorable prognosis in non-small-cell lung cancer (Zheng *et al.*, 2007).

The present study was designed to evaluate the protein expression of *RRM1* and *ERCC1* in pancreatic cancer by automated quantitative analysis (AQUA). We describe the relationship between *RRM1* and *ERCC1* expression, the association between the expression of these proteins and prognosis, as well as the response to

Correspondence: Dr Y Takeda, Department of Surgery, Osaka University Graduate School of Medicine, 2-2, Yamadaoka, Suita-city, Osaka 5650871, Japan.
E-mail: ytakeda@gesurg.med.osaka-u.ac.jp
Received 24 November 2008; revised 23 March 2009; accepted 14 May 2009; published online 22 June 2009

gemcitabine therapy. To our knowledge, this study is the first to examine both the prognostic and predictive aspects of *RRM1* in the same clinical samples.

Results

RRM1 and ERCC1 expression characteristics

We constructed a tissue microarray using triplicate 0.6-mm cores from formalin-fixed and paraffin-embedded specimens of the primary tumor. Immunostaining showed a granular nuclear pattern for *RRM1*, and a fine granular pattern for *ERCC1* (Figure 1). Next, we used AQUA to analyse the expression levels of *RRM1* and *ERCC1* in specimens obtained from 68 patients. The scores of *RRM1* ranged from 116 to 1644 (median, 539; mean, 546) for all specimens, and the scores of *ERCC1* ranged from 55 to 1469 (median 382, mean 412).

The average score of triplicate tissues from each patient was used for analysis of the association between staining and clinical parameters. The AQUA scores for *RRM1* did not correlate significantly with those of *ERCC1* ($r=0.172$, $P=0.1610$) (Figure 2). The median values of *RRM1* and *ERCC1* expression levels were used to divide the patients into high and low expression groups. There were no significant differences between patients with high and low tumoral *RRM1* expression or high and low tumoral *ERCC1* expression with respect to age, sex, histopathological type (well/mod/poor), tumor size, tumor location (head/body/tail), pathological depth of tumor (pT1/T2/T3), the total number of resected lymph nodes, pathological lymph node metastasis (negative/positive) and the number of metastatic

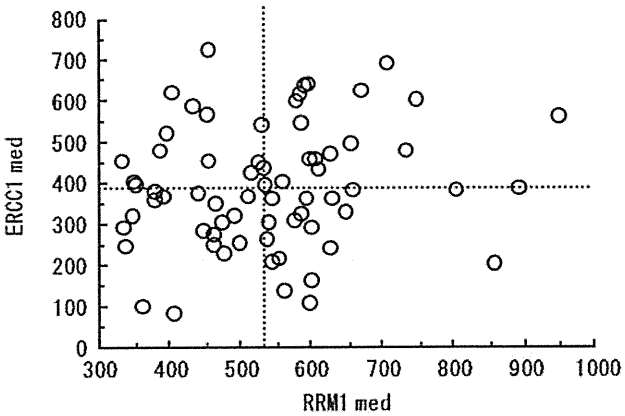


Figure 2 Relationship between automated quantitative analysis (AQUA) scores of *RRM1* and *excision repair cross-complementation group 1* (*ERCC1*) expression. *RRM1* expression did not correlate with that of *ERCC1* ($r=0.172$, $P=0.161$).

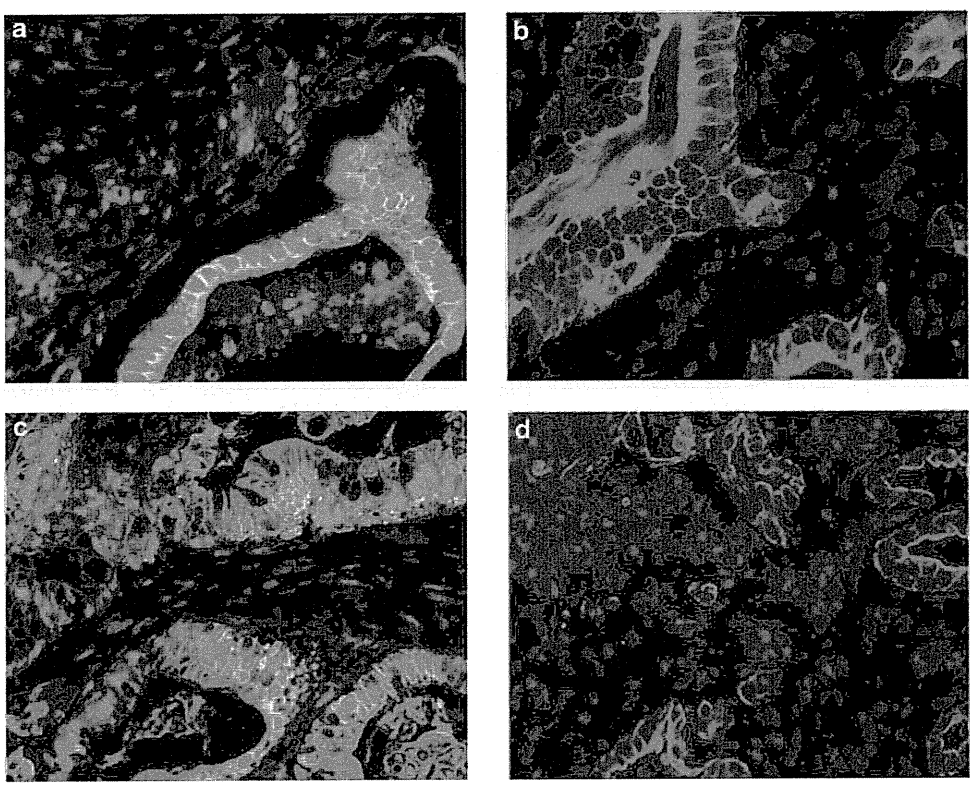


Figure 1 Staining for *RRM1* and *excision repair cross-complementation group 1* (*ERCC1*) proteins. (**a**) *RRM1*-positive sample. Note the granular nuclear pattern. Nucleus, blue; cytoplasm, red; *RRM1*, green; and merged, light blue to light green. (**b**) *RRM1*-negative sample. Nucleus, blue; and cytoplasm, red. (**c**) *ERCC1*-positive sample. Note the fine granular pattern in the nucleus. Nucleus, blue; cytoplasm, red; *ERCC1*, green; and merged, light blue to light green. (**d**) *ERCC1*-negative sample. Nucleus, blue; and cytoplasm, red.

Table 1 Relationship between protein expression levels and clinicopathological factors

	RRM1 expression level			*ERCC1 expression level*		
	High	*Low*	P-*value*	*High*	*Low*	P-*value*
Age (years) (mean ± s.d.)	66.8 ± 7.6	64.4 ± 7.9	0.220	64.6 ± 7.7	66.6 ± 7.8	0.283
Sex (male/female)	15/19	18/16	0.628	15/19	18/16	0.628
Histopathology (well/mod/poor)	17/14/3	9/18/7	0.102	12/19/3	14/13/7	0.237
Tumor size (cm) (mean ± s.d.)	27.4 ± 9.3	26.7 ± 8.2	0.752	25.2 ± 8.2	28.9 ± 8.9	0.077
Tumor location (head/body/tail)	27/6/1	27/4/3	0.497	27/4/3	27/6/1	0.497
pT (T1/T2/T3)	1/1/32	1/0/33	0.602	1/1/32	1/0/33	0.602
Total number of resected lymph node	34.4 ± 12.9	30.3 ± 13.6	0.243	30.8 ± 10.6	34.3 ± 15.7	0.330
PN (positive/negative)	12/22	17/17	0.327	18/16	11/23	0.141
Total number of metastatic lymph node	1.6 ± 1.9	1.0 ± 1.7	0.202	1.1 ± 1.7	1.5 ± 1.9	0.315
Gem therapy (+/−)	14/20	14/20	0.999	13/21	15/19	0.806

Abbreviation: *ERCC1, excision repair cross-complementation group 1.*

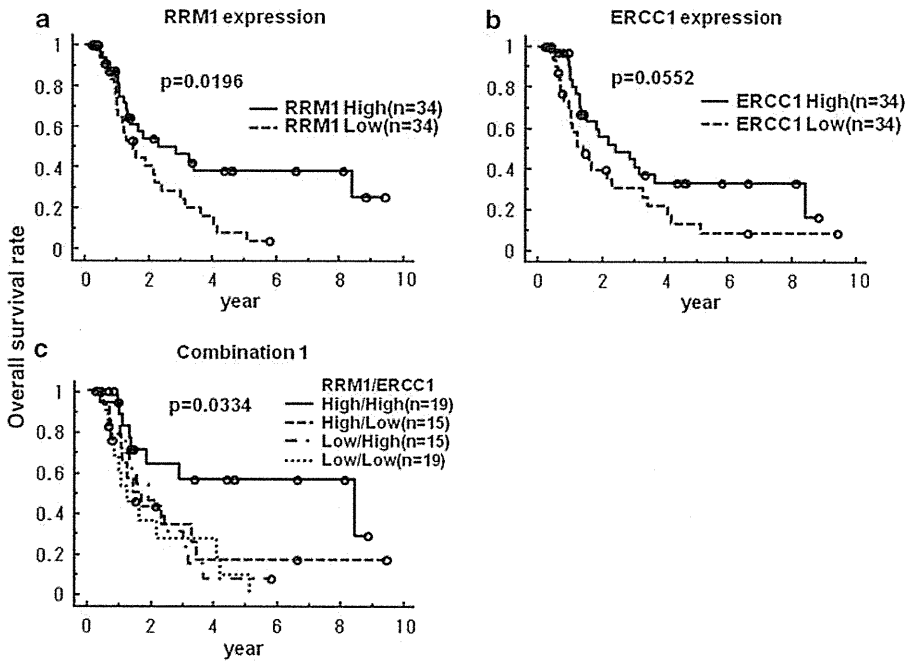


Figure 3 Relationship between *RRM1* and *excision repair cross-complementation group 1* (*ERCC1*) expression levels and overall survival rate. (**a**) Relationship between *RRM1* and overall survival is significant (3-year survival; 46.3 versus 28.6%, $P=0.0196$). (**b**) Relationship between *ERCC1* and overall survival is marginal ($P=0.0552$). (**c**) Relationship between the combination of *RRM1* and *ERCC1* expression levels in the same tumor and overall survival rate. Only high expression levels of *RRM1* and *ERCC1* in the same tumor related with the improvement of overall survival rate ($P=0.0334$).

lymph nodes, and whether or not gemcitabine was used as chemotherapy (Table 1).

Relationship between RRM1/ERCC1 expression and prognosis

The median overall survival of all patients was 16.3 months (4.3–113) and the median disease-free survival was 10.3 months (2–106). The Kaplan–Meier overall survival estimates were significantly better for patients with high *RRM1* expression compared with those having low *RRM1* expression levels (3-year survival; 46.3 versus 28.6%, $P=0.0196$) (Figure 3a). Likewise, patients with high *ERCC1* expression had a better overall survival than those with low levels of expression; although this difference was only marginally significant ($P=0.0552$) (Figure 3b). When we divided the 68 patients into four groups; that is, high tumoral expression of both proteins (High/High, $n=19$), high expression of only *RRM1* (High/Low, $n=15$), high expression of only *ERCC1* (Low/High, $n=15$) and low expression of both proteins (Low/Low, $n=19$); only patients of the High/High group had a significantly better prognosis than the others (3-year survival; 56.7 versus 30.5%, $P=0.0066$) (Figure 3c, Supplementary Figure 1).

With regard to disease-free survival, high *ERCC1* expression levels were significantly associated with better outcome (3-year survival; 30.2% for high versus

23.1% for low, $P=0.0454$). There was no significant difference in disease-free survival between the high and low *RRM1* expression groups (Supplementary Figures 2A and B). With respect to the combination of *RRM1* and *ERCC1*, only the High/High group showed a significantly better disease-free survival compared with the other groups (3-year survival, 43.2 versus 19.2%, $P=0.0127$) (Supplementary Figures 2C and D).

Univariate and multivariate analysis of factors associated with prognosis

We investigated the prognostic significance of various clinicopathological factors in pancreatic cancer patients who underwent radical resection. Univariate analysis showed that only the pathological type and absence or presence of lymph node metastases, were prognostically significant for disease-free survival ($P=0.034$, 0.025, respectively), and both parameters had marginal significance for overall survival ($P=0.078$, 0.084, respectively) (Table 2). Multivariate analysis identified the *RRM1* expression level as the only independent determinant of overall survival (hazard ratio (HR) 1.89, $P=0.046$), and none of the parameters tested was selected by the analysis as a significant prognostic factor in disease-free survival.

RRM1 expression and response to gemcitabine

Of all the 68 patients, 28 received therapy with single-agent gemcitabine. In 23 patients, this treatment was initiated at the time of tumor recurrence. To elucidate the relationship between *RRM1* expression level and gemcitabine therapy, we used survival after recurrence, which represented the period from starting gemcitabine therapy or other therapies in 50 patients with relapse, until death. First, we examined the survival benefit of gemcitabine. The 23 patients who were treated with gemcitabine had a significantly better survival than those who did not ($P=0.0074$) (Supplementary Figure 3). After dividing patients that were treated with gemcitabine into high and low *RRM1* expression groups, only patients with low *RRM1* expression benefited from gemcitabine therapy ($P=0.0010$) (Figure 4b). The survival of patients with high *RRM1* expression treated with gemcitabine was not significantly better than of those not treated with gemcitabine ($P=0.3309$) (Figure 4a). The interaction term between *RRM1* expression and gemcitabine treatment was significant for survival after recurrence ($P=0.0109$).

Discussion

Ribonucleotide reductase, composed of the regulatory subunit *RRM1* and the catalytic subunit *RRM2*, is a key enzyme involved in DNA synthesis, catalyzing the biosynthesis of deoxyribonucleotides from the corresponding ribonucleotides (Wright *et al.*, 1990; Hurta and Wright, 1992). *ERCC1*, a structure-specific DNA repair endonuclease responsible for the 5′ incision, has a key role in the removal of adducts from genomic DNA

Table 2 Prognostic factors for postoperative survival by Cox's proportional hazard model

	Univariate analysis				*Multivariate analysis*			
	DFS		*OS*		*DFS*		OS	
	HR	P-*value*	*HR*	P-*value*	*HR*	P-*value*	*HR*	P-*value*
Histology (poor, mod/well)	1.91	0.034	1.75	0.078	1.77	0.066	1.56	0.172
PN (positive/negative)	2.00	0.025	1.76	0.084	1.73	0.107	1.50	0.256
RRM1 expression (low/high)	1.55	0.129	2.04	0.022	1.39	0.265	1.89	0.046
ERCC1 expression (low/high)	1.75	0.048	1.78	0.059	1.42	0.265	1.54	0.194

Abbreviations: DFS, disease-free survival; *ERCC1*, excision repair cross-complementation group 1; HR, hazard ratio and OS, overall survival.

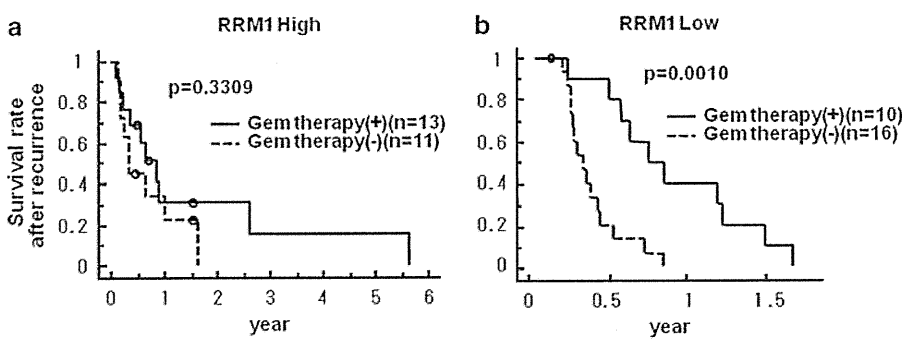


Figure 4 Relationship between survival after recurrence and patients treated with or without gemcitabine (**a**) in high *RRM1* expression group, and (**b**) in low expression group. Only patients with low *RRM1* expression benefited from gemcitabine therapy ($P=0.0010$).

through the nucleotide excision repair pathway (Reardon *et al.*, 1999; Niedernhofer *et al.*, 2004; Ceppi *et al.*, 2006). *RRM1* is reported to influence cell survival, probably through interaction with the *phosphatase and tensin homolog* (*PTEN*), which is an inhibitor of cell proliferation, and suppresses cell migration and invasion by reducing the phosphorylation of focal adhesion kinase (Gautam *et al.*, 2003; Bepler *et al.*, 2004). In lung cancer, the expression levels of *RRM1* and *ERCC1* are significantly correlated (Bepler *et al.*, 2006; Ceppi *et al.*, 2006).

Gemcitabine is the first line cytotoxic agent for treatment of patients with advanced pancreatic cancer, and it is the only agent with proven benefit in a large adjuvant clinical trial (Oettle *et al.*, 2007). However, it is estimated that only 25% of patients benefit from gemcitabine (Burris *et al.*, 1997). *RRM1* expression appears to be the key determinant of gemcitabine resistance (Dumontet *et al.*, 1999; Goan *et al.*, 1999; Jung *et al.*, 2001). This is partially due to expansion of the dNTP pool, which competitively inhibits the incorporation of gemcitabine triphosphate into DNA (Plunkett *et al.*, 1996). Another mechanism is the direct interaction between *RRM1* and gemcitabine with *RRM1* acting as a 'molecular sink' for gemcitabine (Davidson *et al.*, 2004; Bergman *et al.*, 2005). *ERCC1* is reported to be associated with the repair of cisplatin-induced DNA adducts in ovarian cancer (Li *et al.*, 2000), gastric cancer (Metzger *et al.*, 1998), colorectal cancer (Shirota *et al.*, 2001), lung cancer (Olaussen *et al.*, 2006) and esophageal cancer (Joshi *et al.*, 2005; Kim *et al.*, 2008).

Quantitative analysis of gene expression in pancreatic cancer is challenging because it contains more stromal tissue than other cancers (Sato *et al.*, 2004; Bachem *et al.*, 2005; Infante *et al.*, 2007), which makes laser microdissection a necessity to obtain gene expression of tumor tissue (Giovannetti *et al.*, 2006). Quantitative analysis of the *RRM1* protein had been difficult because of technical limitations. However, an automated, quantitative *in situ* assessment of protein expression was developed recently (Camp *et al.*, 2002), and applied for objective and practical evaluation of *RRM1* and *ERCC1* protein expression levels in tumor specimens (Zheng *et al.*, 2007). In this study, we used the above mentioned technology for gene expression analysis in pancreatic cancer specimens.

We found that the expression levels of *RRM1* and *ERCC1* affected the clinical outcome similar to that described in non-small-cell lung cancer (Zheng *et al.*, 2007). Patients with high levels of expression of both proteins had the best prognosis, including both disease-free survival and overall survival. However, once treatment with gemcitabine was initiated at the time of recurrence, it was only the group of patients with low levels of *RRM1* that benefited significantly from this intervention. In other words, patients with high tumoral *RRM1* levels may as well be treated with other agents, such as S-1 or oxaliplatin plus 5-fluorouracil plus leukovorin (CONKO-003), instead of gemcitabine (Ueno *et al.*, 2005; Okusaka *et al.*, 2008; Saif, 2008). In contrast, patients with low tumoral *RRM1* levels showed improved survival following treatment with gemcitabine (Moore *et al.*, 2007; Boeck and Heinemann, 2008). Many clinical trials of anticancer drugs, including molecular targeting agents, did not result in the improvement of outcome when conducted in unselected groups of patients (Heinemann *et al.*, 2006; Herrmann *et al.*, 2007; Cascinu *et al.*, 2008). However, if patients can be divided into groups with high or low likelihood of benefit from gemcitabine, a more rational design of future trials becomes available (Simon *et al.*, 2007). We believe that future treatment strategies for pancreatic cancer should be more precise and tailored to individual patients, and *RRM1* may be one of the candidate molecules for the stratification. We found that *RRM1* and *ERCC1* were not significantly coexpressed in pancreatic cancer, which is different from several previous reports in non-small-cell lung cancer (Ceppi *et al.*, 2006; Zheng *et al.*, 2007). This discrepancy may be explained by differences in tissue of origin and mechanisms of carcinogenesis between pancreatic cancer and lung cancer.

It is important to carry out prospective tailored therapeutic trials in pancreatic cancer with the goal of improving the clinical outcome, and it is our opinion that *RRM1* and *ERCC1* could play an important role in the design of such trials.

Materials and methods

Patients

Between January 1992 and March 2008, 166 patients underwent surgery for pancreatic cancer at Osaka University Hospital. We excluded 84 patients for the following reasons: (1) tumors were not resectable in 26 patients because of liver metastases or peritoneal carcinomatosis, (2) surgery resulted in R1 (residual microscopic cancer) or R2 (residual macroscopic cancer) resections in 21 patients, (3) chemotherapy or chemoradiotherapy was provided preoperatively to 37 patients and (4) lack of neutral-buffered formalin-fixed and paraffin-embedded tumor blocks or/and clinical follow-up information for study purposes in 14 cases. As the natural history of variant pancreatic neoplasms differs from the usual pancreatic ductal adenocarcinoma, patients with intraductal papillary mucinous neoplasms, mucinous cystic adenocarcinomas and medullary adenocarcinomas were excluded from this study. Supplementary Table 1 summarizes the characteristics of the 68 patients who were enrolled in this study. They included 33 men and 35 women with a mean age of 60.7 ± 7.8 years ($\pm$ s.d.). All patients had R0 (no residual cancer) resections by pancreaticoduodenectomy in 54 patients, distal pancreatectomy in 12 patients and other resections in 2 patients. The histopathological grading showed poorly, moderately, and well-differentiated adenocarcinoma in 10, 32 and 26 patients, respectively. The UICC-TNM classification was 2, 1 and 65 patients with pT1, pT2 and pT3; 29, 33 and 6 patients with pN0, pN1 and pM1lym; and 1, 1, 27, 33 and 6 patients with stage IA, IB, IIA, IIB and IV, respectively. None of the patients had received neoadjuvant therapy preoperatively. All 68 patients were followed until disease recurrence and/or death. The median follow-up period was 16.3 months (range, 4.3–113), the 5-year survival rate was 23.4%, and the recurrence of disease was observed in 50 patients. Treatment with gemcitabine was carried out in 28 patients; 5 patients

received it as adjuvant chemotherapy and 23 patients received it after disease recurrence. Radiation therapy was not carried out during all the follow-up period.

Immunofluorescence and automated quantitative analysis
We carried out immunostaining after constructing a tissue microarray. Immunofluorescence combined with AQUA was used to assess *in situ* expression of the target molecules as described previously (Zheng *et al.*, 2007). Antigens were retrieved by incubating the tissue in a microwave oven. Optimal concentrations of antisera and antibodies were used to detect *RRM1*, *ERCC1* and cytokeratin. The antiserum to *RRM1* was generated from rabbits and affinity-purified (R1AS-6) as described previously (Zheng *et al.*, 2007). Commercially available antibodies were used for the analysis of *ERCC1* (Ab-2 clone 8F1, MS-671-R7, Laboratory Vision Corporation, Fremont, CA, USA) and cytokeratin (antihuman pancytokeratin AE1/AE3, M3515 and Z0622, Dako Cytomation, Glostrup, Denmark) (Zheng *et al.*, 2007). They were visualized with the use of fluorochrome-labeled secondary antibodies. The final slides were scanned with SpotGrabber (HistoRx, New Haven, CT, USA), and images were analysed with AQUA (version 1.6, PM-2000, HistoRx). The AQUA scores ranged from 0 (no expression) to 3000 (maximal observed expression).

Statistical analysis and ethical issues
Data are expressed as mean ± s.d. Differences in continuous values were evaluated by the Student's *t*-test (Table 1). The Fisher's exact probability test was used to compare discrete variables (Table 1). We evaluated correlations between AQUA scores of *RRM1* and *ERCC1* by Pearson's correlation coefficient (Figure 2). Disease-free and overall survival rates were estimated by the Kaplan–Meier method and compared using the log-rank test (Table 1, Figures 3 and 4). Cox's proportional hazard regression model with stepwise comparisons was used to analyse independent prognostic factors (Table 2). The predictive value of *RRM1* was studied by testing the interaction between *RRM1* expression and gemcitabine treatment in the same Cox model. A P-value <0.05 was used to indicate statistical significance.

This study was analysed by the statistical expert in our laboratory and the study protocol was approved by the Human Ethics Review Committee of Osaka University, and a signed consent form was obtained from each subject.

Conflict of interest

The authors declare no conflict of interest.

Acknowledgements

This work was partially supported by National Institutes of Health (NIH) grant R01-CA129343 to GB and by a grant-in-aid for cancer research from the Ministry of Culture and Science of Japan.

References

Bachem MG, Schunemann M, Ramadani M, Siech M, Beger H, Buck A *et al.* (2005). Pancreatic carcinoma cells induce fibrosis by stimulating proliferation and matrix synthesis of stellate cells. *Gastroenterology* **128**: 907–921.

Bepler G, Kusmartseva I, Sharma S, Gautam A, Cantor A, Sharma A *et al.* (2006). RRM1 modulated *in vitro* and *in vivo* efficacy of gemcitabine and platinum in non-small-cell lung cancer. *J Clin Oncol* **24**: 4731–4737.

Bepler G, Sharma S, Cantor A, Gautam A, Haura E, Simon G *et al.* (2004). RRM1 and PTEN as prognostic parameters for overall and disease-free survival in patients with non-small-cell lung cancer. *J Clin Oncol* **22**: 1878–1885.

Bergman AM, Eijk PP, Ruiz van Haperen VW, Smid K, Veerman G, Hubeek I *et al.* (2005). *in vivo* induction of resistance to gemcitabine results in increased expression of ribonucleotide reductase subunit M1 as the major determinant. *Cancer Res* **65**: 9510–9516.

Boeck S, Heinemann V. (2008). Second-line therapy in gemcitabine-pretreated patients with advanced pancreatic cancer. *J Clin Oncol* **26**: 1178–1179; author reply 1179.

Burris III HA, Moore MJ, Andersen J, Green MR, Rothenberg ML, Modiano MR *et al.* (1997). Improvements in survival and clinical benefit with gemcitabine as first-line therapy for patients with advanced pancreas cancer: a randomized trial. *J Clin Oncol* **15**: 2403–2413.

Camp RL, Chung GG, Rimm DL. (2002). Automated subcellular localization and quantification of protein expression in tissue microarrays. *Nat Med* **8**: 1323–1327.

Cao MY, Lee Y, Feng NP, Xiong K, Jin H, Wang M *et al.* (2003). Adenovirus-mediated ribonucleotide reductase R1 gene therapy of human colon adenocarcinoma. *Clin Cancer Res* **9**: 4553–4561.

Cascinu S, Berardi R, Labianca R, Siena S, Falcone A, Aitini E *et al.* (2008). Cetuximab plus gemcitabine and cisplatin compared with gemcitabine and cisplatin alone in patients with advanced pancreatic cancer: a randomised, multicentre, phase II trial. *Lancet Oncol* **9**: 39–44.

Ceppi P, Volante M, Novello S, Rapa I, Danenberg KD, Danenberg PV *et al.* (2006). ERCC1 and RRM1 gene expressions but not EGFR are predictive of shorter survival in advanced non-small-cell lung cancer treated with cisplatin and gemcitabine. *Ann Oncol* **17**: 1818–1825.

Davidson JD, Ma L, Flagella M, Geeganage S, Gelbert LM, Slapak CA. (2004). An increase in the expression of ribonucleotide reductase large subunit 1 is associated with gemcitabine resistance in non-small cell lung cancer cell lines. *Cancer Res* **64**: 3761–3766.

Dumontet C, Fabianowska-Majewska K, Mantincic D, Callet Bauchu E, Tigaud I, Gandhi V *et al.* (1999). Common resistance mechanisms to deoxynucleoside analogues in variants of the human erythroleukaemic line K562. *Br J Haematol* **106**: 78–85.

Fan H, Huang A, Villegas C, Wright JA. (1997). The R1 component of mammalian ribonucleotide reductase has malignancy-suppressing activity as demonstrated by gene transfer experiments. *Proc Natl Acad Sci USA* **94**: 13181–13186.

Gautam A, Bepler G. (2006). Suppression of lung tumor formation by the regulatory subunit of ribonucleotide reductase. *Cancer Res* **66**: 6497–6502.

Gautam A, Li ZR, Bepler G. (2003). RRM1-induced metastasis suppression through PTEN-regulated pathways. *Oncogene* **22**: 2135–2142.

Giovannetti E, Del Tacca M, Mey V, Funel N, Nannizzi S, Ricci S *et al.* (2006). Transcription analysis of human equilibrative nucleoside transporter-1 predicts survival in pancreas cancer patients treated with gemcitabine. *Cancer Res* **66**: 3928–3935.

Goan YG, Zhou B, Hu E, Mi S, Yen Y. (1999). Overexpression of ribonucleotide reductase as a mechanism of resistance to 2,2-difluorodeoxycytidine in the human KB cancer cell line. *Cancer Res* **59**: 4204–4207.

Heinemann V, Quietzsch D, Gieseler F, Gonnermann M, Schonekas H, Rost A *et al.* (2006). Randomized phase III trial of gemcitabine

plus cisplatin compared with gemcitabine alone in advanced pancreatic cancer. *J Clin Oncol* **24**: 3946–3952.

Herrmann R, Bodoky G, Ruhstaller T, Glimelius B, Bajetta E, Schuller J *et al.* (2007). Gemcitabine plus capecitabine compared with gemcitabine alone in advanced pancreatic cancer: a randomized, multicenter, phase III trial of the Swiss Group for Clinical Cancer Research and the Central European Cooperative Oncology Group. *J Clin Oncol* **25**: 2212–2217.

Hurta RA, Wright JA. (1992). Alterations in the activity and regulation of mammalian ribonucleotide reductase by chlorambucil, a DNA damaging agent. *J Biol Chem* **267**: 7066–7071.

Infante JR, Matsubayashi H, Sato N, Tonascia J, Klein AP, Riall TA *et al.* (2007). Peritumoral fibroblast SPARC expression and patient outcome with resectable pancreatic adenocarcinoma. *J Clin Oncol* **25**: 319–325.

Joshi MB, Shirota Y, Danenberg KD, Conlon DH, Salonga DS, Herndon II JE *et al.* (2005). High gene expression of TS1, GSTP1, and ERCC1 are risk factors for survival in patients treated with trimodality therapy for esophageal cancer. *Clin Cancer Res* **11**: 2215–2221.

Jung CP, Motwani MV, Schwartz GK. (2001). Flavopiridol increases sensitization to gemcitabine in human gastrointestinal cancer cell lines and correlates with down-regulation of ribonucleotide reductase M2 subunit. *Clin Cancer Res* **7**: 2527–2536.

Kayahara M, Nagakawa T, Ueno K, Ohta T, Takeda T, Miyazaki I. (1993). An evaluation of radical resection for pancreatic cancer based on the mode of recurrence as determined by autopsy and diagnostic imaging. *Cancer* **72**: 2118–2123.

Kim MK, Cho KJ, Kwon GY, Park SI, Kim YH, Kim JH *et al.* (2008). Patients with ERCC1-negative locally advanced esophageal cancers may benefit from preoperative chemoradiotherapy. *Clin Cancer Res* **14**: 4225–4231.

Li D, Xie K, Wolff R, Abbruzzese JL. (2004). Pancreatic cancer. *Lancet* **363**: 1049–1057.

Li Q, Yu JJ, Mu C, Yunmbam MK, Slavsky D, Cross CL *et al.* (2000). Association between the level of ERCC-1 expression and the repair of cisplatin-induced DNA damage in human ovarian cancer cells. *Anticancer Res* **20**: 645–652.

Metzger R, Leichman CG, Danenberg KD, Danenberg PV, Lenz HJ, Hayashi K *et al.* (1998). 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* **16**: 309–316.

Moore MJ, Goldstein D, Hamm J, Figer A, Hecht JR, Gallinger S *et al.* (2007). Erlotinib plus gemcitabine compared with gemcitabine alone in patients with advanced pancreatic cancer: a phase III trial of the National Cancer Institute of Canada Clinical Trials Group. *J Clin Oncol* **25**: 1960–1966.

Nakahira S, Nakamori S, Tsujie M, Takahashi Y, Okami J, Yoshioka S *et al.* (2007). Involvement of ribonucleotide reductase M1 subunit overexpression in gemcitabine resistance of human pancreatic cancer. *Int J Cancer* **120**: 1355–1363.

Niedernhofer LJ, Odijk H, Budzowska M, van Drunen E, Maas A, Theil AF *et al.* (2004). The structure-specific endonuclease Ercc1-Xpf is required to resolve DNA interstrand cross-link-induced double-strand breaks. *Mol Cell Biol* **24**: 5776–5787.

Nitecki SS, Sarr MG, Colby TV, van Heerden JA. (1995). Long-term survival after resection for ductal adenocarcinoma of the pancreas. Is it really improving? *Ann Surg* **221**: 59–66.

Oettle H, Post S, Neuhaus P, Gellert K, Langrehr J, Ridwelski K *et al.* (2007). Adjuvant chemotherapy with gemcitabine vs observation in patients undergoing curative-intent resection of pancreatic cancer: a randomized controlled trial. *Jama* **297**: 267–277.

Okusaka T, Funakoshi A, Furuse J, Boku N, Yamao K, Ohkawa S *et al.* (2008). A late phase II study of S-1 for metastatic pancreatic cancer. *Cancer Chemother Pharmacol* **61**: 615–621.

Olaussen KA, Dunant A, Fouret P, Brambilla E, Andre F, Haddad V *et al.* (2006). DNA repair by ERCC1 in non-small-cell lung cancer and cisplatin-based adjuvant chemotherapy. *N Engl J Med* **355**: 983–991.

Plunkett W, Huang P, Searcy CE, Gandhi V. (1996). Gemcitabine: preclinical pharmacology and mechanisms of action. *Semin Oncol* **23**: 3–15.

Reardon JT, Vaisman A, Chaney SG, Sancar A. (1999). Efficient nucleotide excision repair of cisplatin, oxaliplatin, and Bis-aceto-ammine-dichloro-cyclohexylamine-platinum(IV) (JM216) platinum intrastrand DNA diadducts. *Cancer Res* **59**: 3968–3971.

Rosell R, Danenberg KD, Alberola V, Bepler G, Sanchez JJ, Camps C *et al.* (2004). Ribonucleotide reductase messenger RNA expression and survival in gemcitabine/cisplatin-treated advanced non-small cell lung cancer patients. *Clin Cancer Res* **10**: 1318–1325.

Rothenberg ML, Moore MJ, Cripps MC, Andersen JS, Portenoy RK, Burris III HA *et al.* (1996). A phase II trial of gemcitabine in patients with 5-FU-refractory pancreas cancer. *Ann Oncol* **7**: 347–353.

Saif MW. (2008). New developments in the treatment of pancreatic cancer. Highlights from the '44th ASCO Annual Meeting'. Chicago, IL, USA. May 30–June 3, 2008. *JOP* **9**: 391–397.

Sato N, Maehara N, Goggins M. (2004). Gene expression profiling of tumor-stromal interactions between pancreatic cancer cells and stromal fibroblasts. *Cancer Res* **64**: 6950–6956.

Sener SF, Fremgen A, Menck HR, Winchester DP. (1999). Pancreatic cancer: a report of treatment and survival trends for 100,313 patients diagnosed from 1985–1995, using the National Cancer Database. *J Am Coll Surg* **189**: 1–7.

Shirota Y, Stoehlmacher J, Brabender J, Xiong YP, Uetake H, Danenberg KD *et al.* (2001). ERCC1 and thymidylate synthase mRNA levels predict survival for colorectal cancer patients receiving combination oxaliplatin and fluorouracil chemotherapy. *J Clin Oncol* **19**: 4298–4304.

Simon G, Sharma A, Li X, Hazelton T, Walsh F, Williams C *et al.* (2007). Feasibility and efficacy of molecular analysis-directed individualized therapy in advanced non-small-cell lung cancer. *J Clin Oncol* **25**: 2741–2746.

Staley CA, Lee JE, Cleary KR, Abbruzzese JL, Fenoglio CJ, Rich TA *et al.* (1996). Preoperative chemoradiation, pancreaticoduodenectomy, and intraoperative radiation therapy for adenocarcinoma of the pancreatic head. *Am J Surg* **171**: 118–124; discussion 124–5.

Ueno H, Okusaka T, Ikeda M, Takezako Y, Morizane C. (2005). An early phase II study of S-1 in patients with metastatic pancreatic cancer. *Oncology* **68**: 171–178.

Wright JA, Chan AK, Choy BK, Hurta RA, McClarty GA, Tagger AY. (1990). Regulation and drug resistance mechanisms of mammalian ribonucleotide reductase, and the significance to DNA synthesis. *Biochem Cell Biol* **68**: 1364–1371.

Zheng Z, Chen T, Li X, Haura E, Sharma A, Bepler G. (2007). DNA synthesis and repair genes RRM1 and ERCC1 in lung cancer. *N Engl J Med* **356**: 800–808.

Supplementary Information accompanies the paper on the Oncogene website (http://www.nature.com/onc)

Core fucosylation of E-cadherin enhances cell–cell adhesion in human colon carcinoma WiDr cells

Daisuke Osumi,[1,2,10] Motoko Takahashi,[3,10] Eiji Miyoshi,[4] Shunichi Yokoe,[1] Seung Ho Lee,[1] Katsuhisa Noda,[1] Shoji Nakamori,[5] Jianguo Gu,[6] Yoshitaka Ikeda,[7] Yoshio Kuroki,[3] Kazuo Sengoku,[2] Mutsuo Ishikawa[2] and Naoyuki Taniguchi[1,8,9,11]

[1]Department of Biochemistry, Osaka University Graduate School of Medicine, Osaka; [2]Department of Obstetrics and Gynecology, Asahikawa Medical College, Asahikawa; [3]Department of Biochemistry, Sapporo Medical University School of Medicine, Sapporo; [4]Department of Molecular Biochemistry and Clinical Investigation, Osaka University Graduate School of Medicine, Suita; [5]Department of Surgery, National Osaka Hospital, Osaka; [6]Division of Regulatory Glycobiology, Institute of Molecular Biomembrane and Glycobiology, Tohoku Pharmaceutical University, Sendai; [7]Division of Molecular Cell Biology, Department of Biomolecular Sciences, Saga University Faculty of Medicine, Saga; [8]Department of Disease Glycomics, Research Institute for Microbial Diseases, Center for Advanced Science and Innovation, Osaka University, Suita; [9]Systems Glycobiology Group, Disease Glycomics Team, RIKEN Advanced Science Institute, Wako, Japan

(Received October 13, 2008/Revised January 21, 2009/Accepted January 22, 2009/Online publication March 11, 2009)

α1,6-Fucosyltransferase (Fut8), an enzyme that catalyzes the introduction of α1,6 core fucose to the innermost *N*-acetylglucosamine residue of the *N*-glycan, has been implicated in the development, immune system, and tumorigenesis. We found that α1,6-fucosyltransferase and E-cadherin expression levels are significantly elevated in primary colorectal cancer samples. Interestingly, low molecular weight population of E-cadherin appeared as well as normal sized E-cadherin in cancer samples. To investigate the correlation between α1,6-fucosyltransferase and E-cadherin expression, we introduced α1,6-fucosyltransferase in WiDr human colon carcinoma cells. It was revealed that the low molecular weight population of E-cadherin was significantly increased in α1,6-fucosyltransferase-transfected WiDr cells in dense culture, which resulted in an enhancement in cell–cell adhesion. The transfection of mutated α1,6-fucosyltransferase with no enzymatic activity had no effect on E-cadherin expression, indicating that core fucosylation is involved in the phenomena. In α1,6-fucosyltransferase knock down mouse pancreatic acinar cell carcinoma TGP49 cells, the expression of E-cadherin and E-cadherin dependent cell–cell adhesion was decreased. The introduction of α1,6-fucosyltransferase into kidney epithelial cells from α1,6-fucosyltransferase $^{-/-}$ mice restored the expression of E-cadherin and E-cadherin-dependent cell–cell adhesion. Based on the results of lectin blotting, peptide *N*-glycosidase F treatment, and pulse-chase studies, it was demonstrated that the low molecular weight population of E-cadherin contains peptide *N*-glycosidase F insensitive sugar chains, and the turnover rate of E-cadherin was reduced in α1,6-Fucosyltransferase transfectants. Thus, it was suggested that core fucosylation regulates the processing of oligosaccharides and turnover of E-cadherin. These results suggest a possible role of core fucosylation in the regulation of cell–cell adhesion in cancer. (*Cancer Sci* 2009; 100: 888–895)

It is generally accepted that glycosylation affects many properties of glycoproteins, including their conformation, flexibility, and hydrophilicity. As a result, it regulates protein sorting, stability, and protein–protein interactions.[1–5] *N*-Glycans have a common core structure, and their branching patterns are determined by glycosyltransferases.[6,7] Fut8 is an enzyme that catalyzes the introduction of α1,6 core fucose on the asparagine-branched *N*-acetylglucosamine residue of the chitobiose unit of complex-type *N*-glycans.[8,9] Fut8 has been investigated especially in terms of oncogenesis, since the α1,6-fucosylation of α-fetoprotein is a well-known marker of hepatocellular carcinoma.[10] In previous studies, our group reported that Fut8 expression is markedly enhanced in several types of cancer cell lines[11] rat hepatoma tissues[12] and in ovarian serous adenocarcinoma cells.[13]

E-cadherin is a 120 kDa type I membrane protein, which belongs to the class of calcium-dependent cell adhesion molecules.[14] It mediates cell–cell adhesion through the assembly of multiprotein complexes linked to the actin cytoskeleton.[15] Several models have been proposed to date for the cadherin homophilic interactions. Examples include the "linear zipper model", which involves Trp-mediated *cis* dimers and *trans* interactions between the outermost domains,[16] a Trp-dimer model, which involves the formation of a Trp-mediated *trans*-homophilic bond,[17] a model which involves *cis*-dimerization at the Ca^{2+}-binding site,[18] and a model that invokes extensive overlap between ectodomains in the adhesive binding interface.[19,20] The extracellular domain of human E-cadherin consists of five repeats of about 110 amino acid residues, referred to as EC1 through 5, and contains four potential *N*-glycosylation sites, two each in EC4 and EC5. It is synthesized in the form of a precursor polypeptide that is glycosylated and the precursor is then processed to the mature polypeptide.[21–23] It has previously been reported that cells expressing unprocessed E-cadherin by mutating recognition site(s) for processing protease showed no E-cadherin-dependent mediated adhesion.[24]

Our previous studies demonstrated that the introduction of GnT-III and the addition of bisecting GlcNAc residues, products of GnT-III, to E-cadherin down-regulated tyrosine phosphorylation of β-catenin, enhanced cell–cell adhesion mediated by E-cadherin, and suppressed lung metastasis in mouse melanoma cells.[25,26] Consistent with these results, Guo *et al.* also reported that the overexpression of GnT-V, which competes with GnT-III for biantennary substrates, decreased cadherin-mediated cell–cell adhesion.[27] On the other hand, the overexpression of Fut8 in hepatoma cells suppressed intrahepatic metastasis after splenic injection into athymic mice.[28] Liwosz *et al.* reported that the status of *N*-glycosylation of E-cadherin is altered in a cell density-dependent manner, and the loss of complex type of *N*-glycan reduces the molecular weight of E-cadherin and enhanced its preferential association with the actin cytoskeleton, leading to the stabilization of E-cadherin scaffolds.[29] These

[10]These authors contributed equally to this work.
[11]To whom correspondence should be addressed. E-mail: tani52@wd5.so-net.ne.jp
Abbreviations: AAL, *Aleuria aurantia* lectin; BSA, bovine serum albumin; CHO, Chinese hamster ovary; ConA, *Concanavalia ensiformis*; DMEM, Dulbecco's modified Eagle's medium; DSA, *Datura stramonium* lectin; EDTA, ethylenediamine tetraacetic acid; EGFR, epidermal growth factor receptor; FBS, fetal bovine serum; FcγRIIIA, Fcγ receptor IIIA; Fut8, α1,6-fucosyltransferase; GAPDH, glyceraldehydes-3-phosphate dehydrogenase; GDP-Fucose, guanosine diphosphate-fucose; GlcNAc, *N*-acetyl glucosamine; GnT-III, *N*-acetylglucosaminyltransferase III; LRP-1, lipoprotein receptor-related protein-1; MDCK, Madin-Darby canine kidney; MES-NaOH, 2-morpholinoethane sulfonic acid, memohydrate; PBS, phosphate-buffered saline; PCR, polymerase chain reaction; PNGase F, peptide *N*-glycosidase F; SDS-PAGE, sodium dodecyl sulfate–polyacrylamide gel electrophoresis; TGF-β1, transforming growth factor-β1.

doi: 10.1111/j.1349-7006.2009.01125.x

Table 1. Clinical features of patients with colorectal cancer

Sample number	Age	Sex	Location†	Dukes' stages
1	71	Female	A	C
2	41	Male	R	B
3	77	Male	R	B
4	87	Male	T	B
5	69	Male	S	C
6	71	Male	A	D

†A, ascending colon; R, rectum; S, sigmoid colon; T, transverse colon.

studies suggest that *N*-glycans play a role in modulating E-cadherin status. In the present study, we found that Fut8 and E-cadherin protein levels are significantly increased in colorectal cancer samples. E-cadherin in Fut8 transfected WiDr cells, Fut8 knocked down cells, and Fut8 deficient cells from $Fut8^{-/-}$ mice were examined and our results demonstrate that the activity of Fut8 is involved in the appearance of a low molecular weight population of E-cadherin and regulates the total amount of E-cadherin. We propose the possible involvement of core fucosylation in changing the *N*-glycosylation patterns of E-cadherin, the subsequent stabilization of cell–cell contacts, and the regulation of metastatic potential.

Materials and Methods

Human tissues samples. All experiments were approved by ethical committees both in Osaka University and Osaka National Hospital. Tissues from six cases of primary colorectal cancer were surgically resected (Table 1). Written informed consent was obtained from each patient before surgery. The excised samples were obtained within 1 h after the operation from tumor tissues and corresponding non-tumor tissues 5–10 cm remote from the tumor. All of the excised tissues were placed immediately in liquid nitrogen and stored at –80°C until additional analysis.

Cell lines, culture, and transfection. Human colon carcinoma WiDr cells were obtained from the American Type Culture Collection (Rockville, MD, USA) and were maintained in DMEM supplemented with 10% FBS. A Fut8 expression vector was constructed by inserting the open reading frame of human *Fut8* cDNA into a mammalian expression vector pCXN2 which was regulated by the β-actin promoter. Mutant Fut8, which had no enzymatic activity, was produced by mutating arginine 365 to alanine.[30] WiDr cells were transfected with pCXN2/*Fut8* or pCXN2/R365A *Fut8* or pCXN2 by using Lipofectamine 2000 (Invitrogen, Carlsbad, CA, USA) according to the manufacturer's instructions. Selection was performed by 2-week incubation in medium containing G418, and G418-resistant colonies were isolated and recloned by serial dilution to ensure clonality. Fut8 knocked down mouse pancreatic aciuar cell catcinoma TGP49 cells were prepared as described previously.[31] Fut8 deficient kidney epithelial cells were prepared from $Fut8^{-/-}$ mice as described previously.[32] The cells were serum starved for 8 h before harvest to achieve the cell cycle synchronization.

Activity assay of Fut8. The enzymatic activity of Fut8 was measured by high-performance liquid chromatography using a fluorescence-labeled sugar chain as the substrate, as previously described.[33] A standard mixture included 80 mM MES-NaOH (pH 7.0), 0.5% Triton X-100, 2 μM 4-(2-pyridylamino)butylamine-labeled sugar chain, and 50 μM GDP-Fucose. After incubation at 37°C for 2 h, the reaction was terminated by incubating at 100°C for 1 min. The samples were then centrifuged at 15 000 *g* for 10 min and applied to high-performance liquid chromatography on a TSK-gel, ODS-80TM column (4.6 × 150 mm) (Tosho, Tokyo, Japan). Elution was performed at 55°C with 20 mM sodium acetate buffer, pH 4.0, 0.1% butanol, in an isocratic manner. Fluorescence of the column elutes was detected with a fluorescence spectrometer (model RF 535; Shimadzu Corp., Kyoto, Japan), the excitation and emission wavelengths being 320 and 400 nm, respectively.

Protein extraction, immunoprecipitation, and western blotting. Frozen tissue samples were homogenized in 5 vol. of lysis buffer (20 mM Tris-HCl, pH 7.4, 150 mM NaCl, 5 mM ethylenediaminetetraacetic acid, 1% [w/v] Nonidet P-40, 10% [w/v] glycerol, 5 mM sodium pyrophosphate, 10 mM NaF, 1 mM sodium orthovanadate, 10 mM β-glycerophosphate, 1 mM phenylmethylsulfonyl fluoride, 2 μg/mL aprotinin, 5 μg/mL leupeptin, and 1 mM dithiothreitol) using Polytron homogenizer (Kinematica, Littau-Luzern, Switzerland). After centrifugation (15 000 *g*) for 20 min at 4°C the supernatant were collected. Cell cultures were rinsed twice with ice-cold PBS and harvested in lysis buffer. Protein concentrations were determined using a Protein Assay CBB kit (Nacalai Tesque, Kyoto, Japan). For the immunoprecipitation of E-cadherin, whole cell lysates (500 μg) were incubated with 4 μg of mouse anti-E-cadherin antibody (610182; BD Bioscience, San Jose, CA, USA) for 2 h at 4°C, and then with 20 μL of Protein G Sepharose 4 Fast Flow (GE Healthcare Biosciences, Buckinghamshire, UK) for 4 h at 4°C. For western blot analysis, protein samples or immunoprecipitates were subjected to SDS-PAGE, and transferred to nitrocellulose membranes (Schleicher & Schuell, Dassel, Germany). Mouse anti-E-cadherin antibody (610182; BD Bioscience) or mouse anti-Fut8 antibody[34] was used as primary antibodies. Immunoreactive bands were visualized using an enhanced chemiluminescence kit (GE Healthcare Biosciences).

Cell surface biotinylation and immunoprecipitation of E-cadherin. Cell surface biotinylation was performed as described previously.[35] Briefly, cells were incubated with sulfosuccinimidobiotin (s-NHS-biotin; Pierce, Rockford, IL, USA) (1 mg/mL) for 10 min on ice, and the reaction was quenched with 50 mM NH_4Cl. The cell lysate was immunoprecipitated with anti-E-cadherin antibody as described above. The biotinylated proteins were visualized using a Vectastain ABC kit (Vector Laboratories, Burlingame, CA, USA) and an enhanced chemiluminescence kit.

Lectin blot analysis. Lectin blot analysis was performed as described previously.[36] The immunoprecipitated E-cadherin was electrophoresed on 8% SDS-PAGE and transferred to nitrocellulose membranes. The membrane was blocked with 5% BSA (w/v) and then incubated with 2 μg/mL of biotinylated AAL, DSA, or ConA lectin (Seikagaku Corp., Tokyo, Japan) for 30 min at room temperature. After washing, lectin reactive proteins were detected using a Vectastain ABC kit and an enhanced chemiluminescence kit.

Reverse transcription–polymerase chain reaction (RT-PCR) and quantitative real-time PCR. Total RNA was prepared from WiDr cells. cDNAs were synthesized using an SYBY RT-PCR kit (Perfect Real Time; Takara-Bio Inc., Otsu, Japan) and Reverse Transcription Reagent (Takara-Bio Inc.) according to the manufacturer's instructions. A random hexamer was used for cDNA synthesis. Real-time PCR was performed using the SYBR RT-PCR kit and was analyzed on Smart Cycler II system (Cepheid, Sunnyvale, CA, USA). Human E-cadherin was amplified using the primers, sense (5′-GGATTGCAAATTCCTGCCATTC-3′) and antisense (5′-AACGTTGTCCCGGGTGTCA-3′). GAPDH was amplified as a control using the primers, sense (5′-ATTG-CCCTCAACGACCACTT-3′) and antisense (5′-AGGTCCACCA-CCCTGTTGCT-3′). The levels of gene expression were determined using a Delta-Delta Ct method.[37]

Immunofluorescence microscopy. Cells were plated on poly L-lysine-coated glass bottom dish, fixed by incubation with PBS containing 4% paraformaldehyde for 10 min at room temperature and permeabilized in PBS containing 0.1% Triton X-100 for 1 min. After washing with PBS three times for 15 min each, the cells were incubated in ECCD-2 (monoclonal antibody to mouse

E-cadherin M108; Takara-Bio Inc.) (1:100 dilution) for 1 h at room temperature. Primary antibody binding was detected with a fluorescein isothiocyanate-labeled goat antibody to mouse IgG. Glass bottom dishes were mounted under Permafluor aqueous mounting medium and the stained cells were viewed with a laser scanning confocal microscope (Carl Zeiss, Jena, Germany).

Cell aggregation assay. Cells were washed twice with PBS and dissociated by incubation with PBS containing 2 mM EDTA for 30 min at 37°C. Single cell suspensions were prepared, washed, and resuspended in DMEM containing 1% (w/v) BSA. 1×10^6 cells were incubated on a rotation apparatus for 3 h at 37°C. In some experiments, 2 mM EDTA or 100 μg/mL HECD-1 (M106 monoclonal antibody to human E-cadherin; Takara-Bio Inc.) was added. At the end of the incubation, cells were diluted into single wells of a 24-well plate to prevent further aggregation. After allowing cells to settle for 40 min at 37°C, an equal volume of 7.4% formaldehyde in PBS was added to each well and the plate was incubated for 10 min at room temperature. Photos were taken at random under a phase contrast microscope to count single cells or cell aggregates (four or more cells).

Peptide N-glycosidase F (PNGase F) treatment. Whole cell lysates were boiled in 0.1 M 2-mercaptoethanol and 0.1% SDS for 10 min. After boiling, the 50 μg of proteins were incubated for 16 h at 37°C with 100 mM Tris-HCl (pH 8.6), 1% NP-40, and 40 mU/mL PNGase F (Takara Bio Inc.). Then the samples were subjected to 6% SDS-PAGE as described above.

Metabolic labeling and pulse chase study. Eighty-percent confluent monolayers of Fut8 and mock transfectants in 6-well dishes were preincubated for 2 h at 37°C with methionine, cysteine-free DMEM (Sigma) containing dialyzed 10% FBS. For pulse chase studies, L-[^{35}S]methionine and L-[^{35}S]cysteine (Promix; GE Healthcare Biosciences) were added at a concentration of 200 μCi/mL each to the culture media and incubated for 20 min at 37°C for protein labeling. After rinsing three times with PBS, the cells were incubated at 0, 8, 24, and 32 h in DMEM with 10% FBS. Another experiment was performed under conditions of a long-pulse and a long-chase. For the long-pulse, L-[^{35}S]methionine and L-[^{35}S]cysteine were added at a concentration of 50 μCi/mL each to the culture media, followed by incubation for 24 h at 37°C for protein labeling, and after rinsing, the cells were incubated at 0, 24, and 48 h in DMEM with 10% FBS. E-cadherin was immunoprecipitated and subjected to 8% SDS-PAGE. After electrophoresis, the gel was autoradiographed using imaging plates and a BAS-2500 system (FujiFilm, Tokyo, Japan).

Results

E-cadherin and Fut8 expression were increased in primary colorectal cancers. Abnormal cell–cell adhesion that is observed in many types of cancer could result from the changes in E-cadherin expression. We examined Fut8 and E-cadherin expression levels in primary colorectal cancer, and found that they are significantly increased in five out of six examined samples (Fig. 1). Low molecular weight population of E-cadherin appeared as well as normal sized E-cadherin only in the cancer samples. The relative expression levels of Fut8 and E-cadherin in five samples are 3.5–29 and 3.3–11, respectively.

Establishment of WiDr clones stably expressing Fut8. To investigate the correlation between Fut8 and E-cadherin expression, we introduced Fut8 in WiDr human colon carcinoma cells in which Fut8 expression levels are low. WiDr cells were transfected with pCXN2/*FUT8* or pCXN2, and G418-resistant clones were selected as described under "Materials and Methods". The selected Fut8 transfected clones showed elevated enzymatic activities, as 560, 470 and 190 nmol/h/mg protein, respectively. The following experiments were performed with three clones and similar results were observed for all data.

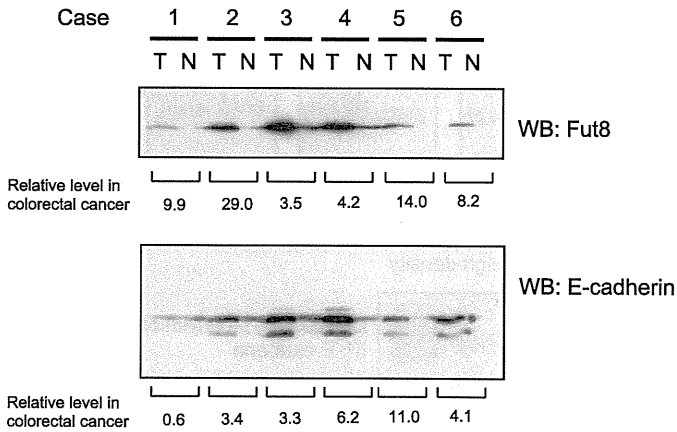


Fig. 1. α1,6-Fucosyltransferase (Fut8) and E-cadherin expression levels were increased in primary colorectal cancer. Total protein lysate was prepared from matched samples of tumor (T) and adjacent non-tumor tissue (N). 100 μg of protein from each pair were subjected to 10% sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE), transferred to nitrocellulose membranes, and Fut8 and E-cadherin were detected using an anti-Fut8 antibody and an anti E-cadherin antibody. WB, Western blotting.

Analysis of E-cadherin in the Fut8 transfected WiDr cells. Western blotting showed that the expression level of E-cadherin in *Fut8* transfectants was increased, especially in high-density cultures (~11×10^4 cells/cm^2) when compared to low density cultures (~3×10^4 cells/cm^2) (Fig. 2a). In high-density cultures, a low molecular weight population of E-cadherin appeared in the *Fut8* transfectants, in addition to the band with the same molecular weight as that in mock transfectants. We also established mutated *Fut8* (R365A *Fut8*) which had no enzymatic activity, and examined its transfectants. Western blotting showed that the lower band was not expressed, even in high-density cultures, suggesting that the Fut8 activity was involved in the appearance of the low molecular weight form of E-cadherin (Fig. 2a). It was confirmed that both of the bands in *Fut8* transfectants are expressed at the cell surface of *Fut8* transfectants (Fig. 2b). A lectin blot analysis indicated that both bands in the *Fut8* transfectants reacted with AAL, which binds preferentially to fucose linked α1,6 to GlcNAc although it binds to fucose linked α1,3 to *N*-acetyllactosamine as well (Fig. 2c). The results suggested that the both forms were core-fucosylated. We examined the reactivity toward other lectins such as DSA or ConA, which react with terminal galactose linked β1,6 to GlcNAc residues or mannose, respectively; however, there were no differences between mock and *Fut8* transfectants (data not shown). The mRNA levels of E-cadherin were evaluated by RT-PCR and quantitative real-time PCR. The results indicated that there was no significant difference between the *Fut8* and mock transfectants (data not shown), suggesting that post-translational modification is involved in the increase in E-cadherin expression in the *Fut8* transfectants.

Accumulation of E-cadherin at the cell–cell border in the Fut8 transfected WiDr cells. Morphologically, *Fut8* transfected WiDr cells appeared as clusters with tight cell–cell contacts, whereas mock transfectants had relatively loose contacts. E-cadherin expression was examined immunohistochemically. The *Fut8* transfectants showed more intense fluorescence with condensation at the cell–cell contacts compared to mock transfectants, indicating that the expression of E-cadherin was elevated in *Fut8* transfectants (Fig. 2d).

Increased cell aggregation in the Fut8 transfected WiDr cells. To examine whether increased E-cadherin in Fut8 transfectants functions, we performed a cell aggregation assay. As shown in

doi: 10.1111/j.1349-7006.2009.01125.x

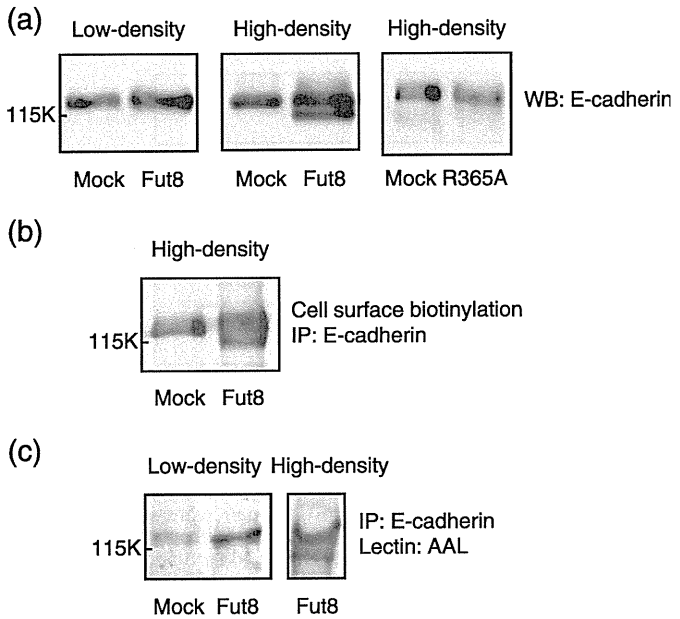


Fig. 2. Effect of α1,6-fucosyltransferase (Fut8) transfection on the characteristics and the expression of E-cadherin in colon carcinoma WiDr cells. (a) Western blotting analysis of E-cadherin in mock, Fut8, and R365A mutated Fut8 transfectants. A whole cell lysate (20 μg) was prepared from low (~3×10^4 cells/cm^2) or high-density (~11×10^4 cells/cm^2) cultures and subjected to 8% sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE), transferred to a nitrocellulose membrane, and E-cadherin was detected using an anti-E-cadherin antibody. WB, Western blotting. (b) Cell surface expression of E-cadherin in the mock and Fut8 transfectants. E-cadherin was immunoprecipitated from whole cell lysate of surface biotinylated mock and Fut8 transfectants, and subjected to 8% SDS-PAGE, transferred to nitrocellulose membranes and the biotinylated E-cadherin was visualized using a Vectastain ABC kit and an enhanced chemiluminescence kit. (c) Lectin blot analysis of E-cadherin in the mock and Fut8 transfectants. E-cadherin was immunoprecipitated from 400 μg of whole cell lysate, subjected to 8% SDS-PAGE, and transferred to nitrocellulose membranes, which were probed by *Aleuria aurantia* lectin (AAL). IP, immunoprecipitation. (d) Distribution of E-cadherin in mock and Fut8 transfectants. E-cadherin was detected by a laser scanning confocal microscopy using an anti-E-cadherin antibody, ECCD-2 (scale bar, 10 μm).

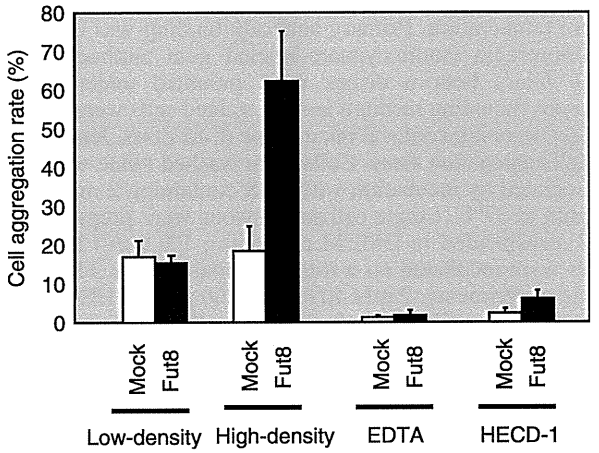


Fig. 3. Effect of α1,6-fucosyltransferase (Fut8) transfection on E-cadherin-dependent cell–cell adhesion in WiDr cells. Cell–cell aggregation was assayed with or without EDTA or the E-cadherin inhibiting antibody, HECD-1, as described in 'Materials and Methods'. Data represent the mean (±SD) of six experiments.

Fig. 3, the cell aggregation rate of the *Fut8* transfectants was significantly higher than that of the mock transfectants under high-density condition (~11×10^4 cells/cm^2), whereas there were no significant differences in aggregation status under low-density condition (~3×10^4 cells/cm^2). This increase in aggregation rate was inhibited in the presence of a calcium chelator, EDTA, or by an anti-E-cadherin monoclonal antibody, indicating that it was dependent on E-cadherin.

Reduction of E-cadherin and cell adhesion in the Fut8 knocked down TGP49 cells. We previously established *Fut8* knocked down mouse pancreatic cancer cells TGP49.[31] The clones showed a low expression of Fut8 and enzymatic activity was not detectable. Lectin blotting confirmed that the two bands of E-cadherin in wild-type cells reacted with AAL lectin, whereas the band in the *Fut8* knocked down cells did not (Fig. 4a). Morphologically, *Fut8* knocked down cells show more loose cell–cell contacts, and the expression levels of E-cadherin at the cell–cell contacts were decreased compared to wild-type cells (Fig. 4b). Consistently, E-cadherin dependent cell aggregation decreased significantly in *Fut8* knocked down cells (Fig. 4c).

Increase in E-cadherin and cell adhesion by restoring Fut8 in Fut8$^{-/-}$ cells. Fut8 deficient kidney epithelial cells were prepared from *Fut8*$^{-/-}$ mice as described previously.[32] The cell surface expression of E-cadherin (Fig. 5a) increased significantly in *Fut8* restored cells compared to *Fut8*$^{-/-}$ cells. Lectin blotting confirmed that the both of the two E-cadherin bands in the *Fut8* restored cells reacted with AAL lectin (Fig. 5b), although the lower band is relatively difficult to be detected. It was considered that the composition of glycosylation of kidney epithelial cells might be different from WiDr colon carcinoma cells. Morphologically, *Fut8*$^{-/-}$ cells show loose cell–cell contacts and *Fut8* restoring rescued them. An immunohistochemical study showed that the expression level of E-cadherin at the cell–cell contacts in *Fut8*$^{-/-}$ cells was also rescued by *Fut8* transfection (Fig. 5c). E-cadherin dependent cell aggregation (Fig. 5d) increased significantly in *Fut8* restored cells compared to *Fut8*$^{-/-}$ cells.

Peptide *N*-glycosidase F (PNGase F) treatment of *N*-glycan of E-cadherin in the Fut8 transfected WiDr cells. To examine the *N*-glycosylation status, *Fut8* and mock transfectants were treated with PNGase F, which cleaves the *N*-glycan between the innermost *N*-acetylglucosamine and asparagines residues. As shown in Fig. 6(a), the upper band of E-cadherin in the *Fut8* transfectant appeared to be *N*-glycosylated to a similar extent as the case of the mock transfectants. On the other hand, no decrease in molecular weight was observed in the PNGase F-treated lower band in the Fut8 transfectants, although the band was reactive with AAL lectin (Fig. 2c), and notably, the molecular weight of the lower band appeared to be even smaller than that of the PNGase-F digest of the upper band in the *Fut8* transfectants. By molecular mass calculation, it was found that the molecular mass of the upper band is around 125 kDa, PNGase F digested upper band is around 110 kDa, and the

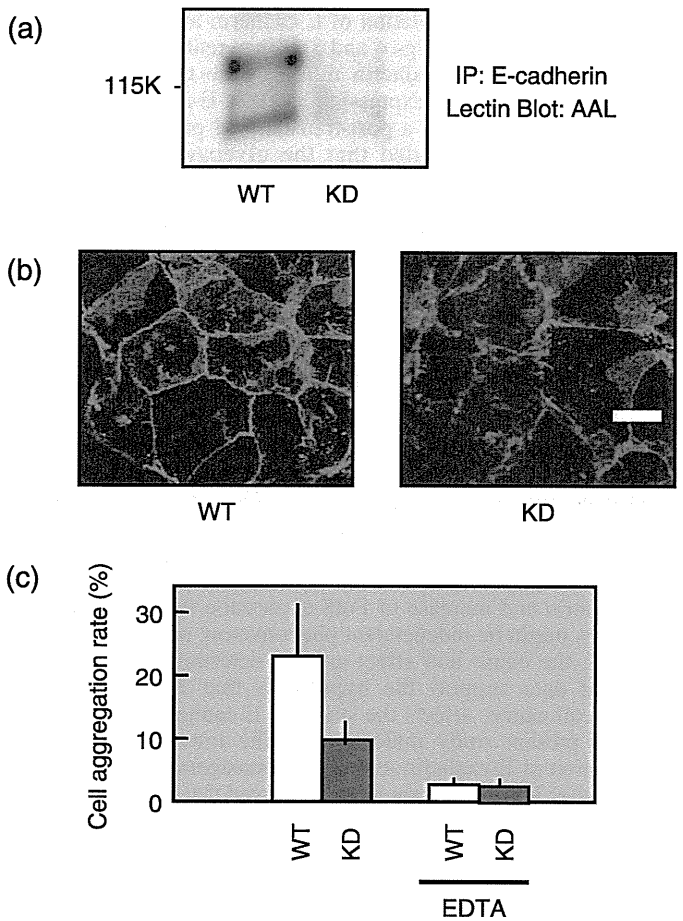


Fig. 4. Changes of E-cadherin expression and E-cadherin-dependent cell–cell adhesion in α1,6-fucosyltransferase (Fut8) knock down cells. (a) Lectin blot analysis of E-cadherin in Fut8 knocked down TGP49 cells and wild-type TGP49 cells. E-cadherin was immunoprecipitated from 400 μg of whole cell lysate, subjected to 8% sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE), and transferred to nitrocellulose membranes, which were probed by *Aleuria aurantia* lectin (AAL). IP, immunoprecipitation. (b) Distribution of E-cadherin in Fut8 knocked down TGP49 cells and wild-type TGP49 cells. E-cadherin was detected by a laser scanning confocal microscopy using an anti-E-cadherin antibody, ECCD-2. WT, wild-type TGP49 cells; KD, Fut8 knocked down TGP49 cells (scale bar, 10 μm). (c) Cell–cell aggregation was assayed with or without EDTA. Data represent the mean (±SD) of six experiments.

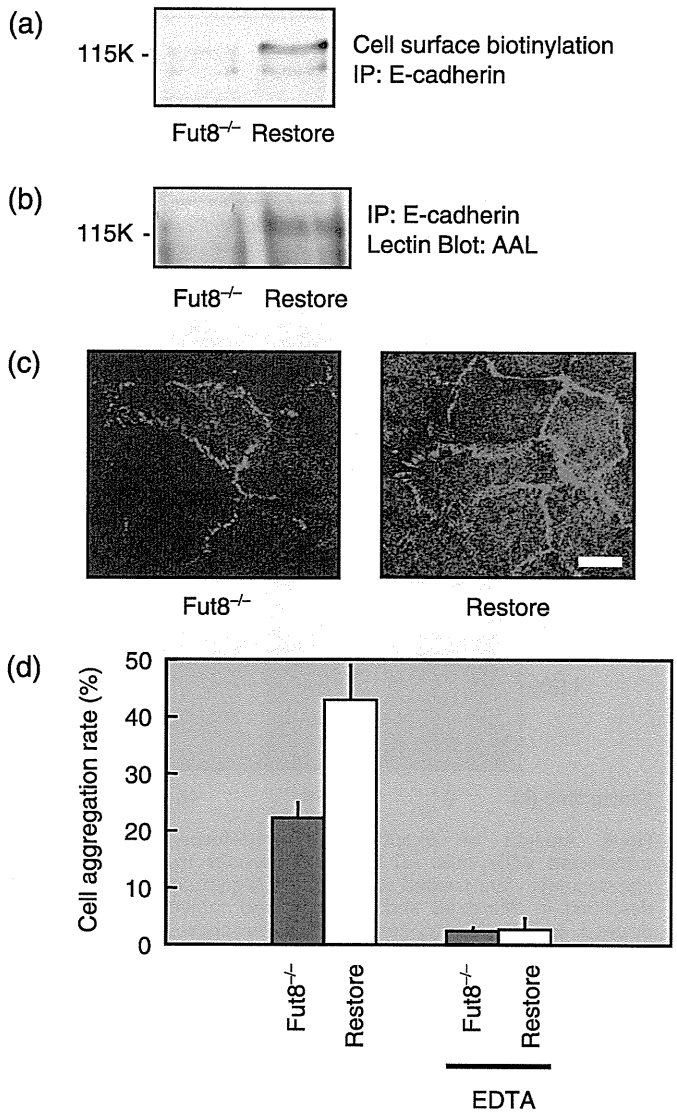


Fig. 5. Changes of E-cadherin expression and E-cadherin-dependent cell–cell adhesion in α1,6-fucosyltransferase (Fut8)$^{-/-}$ cells. (a) Cell surface expression of E-cadherin in the Fut8$^{-/-}$ mouse kidney epithelial cells and Fut8 restored cells. E-cadherin was immunoprecipitated from whole cell lysate of surface biotinylated Fut8$^{-/-}$ cells and Fut8 restored cells, and subjected to 8% sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE), transferred to nitrocellulose membranes and the biotinylated E-cadherin were visualized using a Vectastain ABC kit and an enhanced chemiluminescence kit. (b) Lectin blot analysis of E-cadherin in the Fut8$^{-/-}$ cells and Fut8 restored cells. E-cadherin was immunoprecipitated from 400 μg of whole cell lysates, subjected to 8% SDS-PAGE and transferred to nitrocellulose membranes, which were probed by *Aleuria aurantia* lectin (AAL). (c) Distribution of E-cadherin on Fut8$^{-/-}$ cells and Fut8 restored cells. E-cadherin was detected by a laser scanning confocal microscopy using an anti-E-cadherin antibody, ECCD-2. Fut8$^{-/-}$, Fut8$^{-/-}$ mouse kidney epithelial cells; Restore, Fut8 restored Fut8$^{-/-}$ cells (scale bar, 10 μm). (d) Cell–cell aggregation was assayed with or without EDTA. Data represent the mean (±SD) of six experiments. IP, immunoprecipitation.

lower band in Fut8 transfectant is around 105 kDa. Together with the lectin blotting results, these results indicate that the E-cadherin corresponding to the lower band contains a form of *N*-glycan that is resistant to PNGase F digestion, such as α1-3 fucosylated *N*-glycan.

Turnover of E-cadherin in the Fut8 transfected WiDr cells. To elucidate the mechanisms by which the low molecular weight population of E-cadherin is produced and the total expression levels of E-cadherin are increased in *Fut8* transfectants, the turnover of E-cadherin was examined in pulse-chase studies. As shown in Fig. 6(b), the turnover rate of E-cadherin in *Fut8* transfectants seemed slightly reduced and the band remained longer time. In this experiment, the upper band and the lower band in the *Fut8* transfectants could not be clearly distinguished. When a long-pulse and long-chase experiment was performed, the low molecular weight form of E-cadherin was detected after 24 h chase but not at 0 h in the *Fut8* transfectants (Fig. 6c), whereas the normal form of E-cadherin appeared 20 min after the pulse (data not shown).

Discussion

Changes in glycosylation status have been implicated in pathological status, especially cancer.[38] We have been studying the functional regulation of signaling molecules by *N*-glycosylation.[36,39–43] E-cadherin plays a central role in cancer

doi: 10.1111/j.1349-7006.2009.01125.x

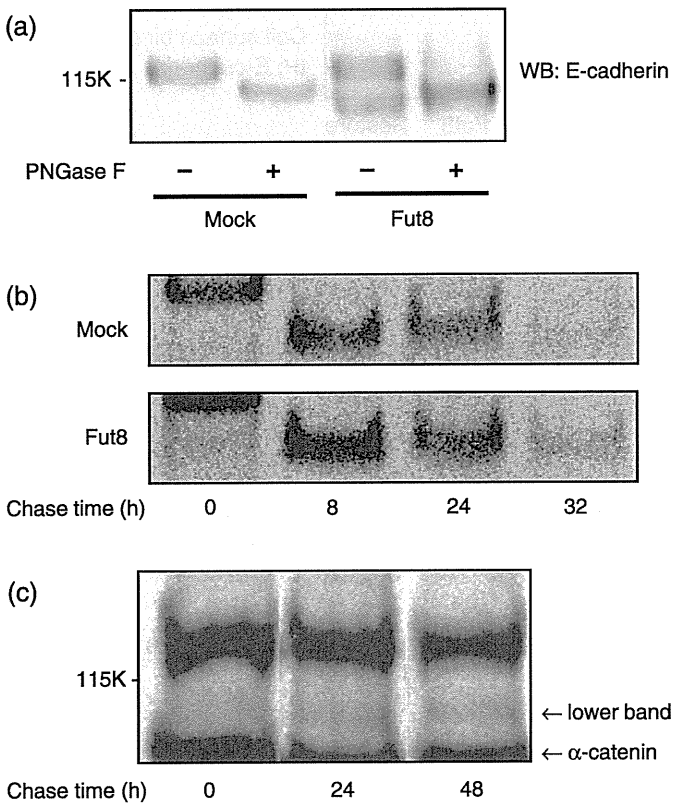


Fig. 6. Analysis of E-cadherin in α1,6-fucosyltransferase (Fut8) transfected WiDr cells. (a) Whole cell lysates from mock and Fut8 transfectants were treated with peptide *N*-glycosidase F (PNGase F) as described in 'Materials and Methods' and subjected to 6% sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE) and transferred to nitrocellulose membrane, probed by an anti-E-cadherin antibody. WB, Western blotting. (b) Mock and Fut8 transfectants were radiolabeled with L-[^{35}S]methionine and L-[^{35}S]cysteine for 20 min. At 0, 8, 24, and 32 h after pulse labeling, E-cadherin was immunoprecipitated from cell lysates using anti-E-cadherin antibody. After separation of the immunoprecipitated E-cadherin by 6% SDS-PAGE, the gel was dried and autoradiographed for 2 days using imaging plates and a BAS-2500 system. The results were reproducible in three independent experiments. (c) Long-pulse and long-chase study. Fut8 transfectants were radiolabeled with L-[^{35}S]methionine and L-[^{35}S]cysteine for 24 h. At 0, 24, and 48 h after pulse labeling, E-cadherin was immunoprecipitated from cell lysates using anti-E-cadherin antibody. After separation of the immunoprecipitated E-cadherin by 6% SDS-PAGE, the gel was dried and autoradiographed for 2 days using imaging plates and a BAS-2500 system. The results were reproducible in three independent experiments.

metastasis, and it has been reported that its function is also affected by the modification of *N*-glycans.[25,26,44] In this study, we found that Fut8 and E-cadherin expression levels are significantly increased in primary colorectal cancer samples. We established the Fut8 transfectants and examined the E-cadherin. Since Fut8 is widely expressed in human tissues and human cancer cell lines, WiDr cells in which Fut8 expression levels are significantly low, and has malignant potential, were selected as the host cells. We have found that Fut8 activity is involved in the appearance of a low molecular weight population of E-cadherin in high-density culture, and regulates the total amount of cell surface E-cadherin (Fig. 2a,b). E-cadherin expression at cell–cell borders and E-cadherin-dependent cell aggregation were significantly enhanced in the *Fut8* transfectants (Figs 2d and 3). Studies with *Fut8* knocked down cells and *Fut8*$^{-/-}$ cells supported the conclusion that Fut8 activity was closely related to the appearance of low molecular weight population of E-cadherin and total expression levels of E-cadherin (Figs 4 and 5). Since real-time PCR showed that the levels of E-cadherin mRNA were not changed in the *Fut8* transfectants, the expression levels of E-cadherin were most likely up-regulated via a post-translational process. The studies with PNGase F revealed that the glycosylation status was different in the lower band in high-density culture of Fut8 transfectants (Fig. 6a). A study of Liwosz *et al.* indicated that the *N*-glycosylation of E-cadherin with complex *N*-glycans is reduced in dense cultures, and that this change facilitates its association with the actin cytoskeleton, leading to the stabilization of E-cadherin scaffolds.[29] They observed that unstable adherens junctions in sparse cells contained E-cadherin primarily modified with complex *N*-glycans, whereas increased amounts of Triton-insoluble E-cadherin in dense cultures correlated with its modification with high mannose/hybrid oligosaccharides, which are small in size. Since CHO cells and MDCK cells, used as host cells in the report, have quite high Fut8 activity, it can be assumed that E-cadherin was core fucosylated. We propose that Fut8 activity is involved in the glycosylation changes of E-cadherin in dense culture and subsequent alterations in cell–cell adhesion. Appearance of low molecular weight population of E-cadherin and increase of Fut8 expression in colorectal cancer samples might be independent phenomenon, and even if they are related, the cause and effect are not determined. However, the present data support the hypothesis that Fut8 increased in colorectal cancer affects the status of E-cadherin.

The present study indicates that the low molecular weight population of E-cadherin in the Fut8 transfectants was PNGase-F insensitive; however, lectin blotting showed that it was glycosylated and core-fucosylated (Figs 6a and 2c). The results of long-pulse and long-chase studies suggested that the low molecular weight E-cadherin is produced from the normal form (Fig. 6c). Cell–cell adhesions mediated by E-cadherin could activate signaling pathways such as receptor tyrosine kinase signaling or Wnt signaling, and it might be possible that cell–cell adhesion-derived signaling affects the processing of *N*-glycan with core fucose. Changes in glycosylation patterns according to cell density were also observed in our previous work[45] and we consider that cell-cell adhesion-derived signaling is involved in processing of glycoproteins. Thus, it is suggested that the *N*-glycans processing and turnover rate of E-cadherin were altered, which caused the accumulation of low molecular population of E-cadherin and total E-cadherin, in Fut8 transfectants in dense culture. It is also possible that increase in E-cadherin-dependent cell aggregation in Fut8 transfectants is due to enhancement of E-cadherin cell adhesive activity, and we must consider the both possibilities at present.

As seen in Fig. 6a, the PNGase-F treated upper band of E-cadherin seems still larger than the low molecular weight population. It is possible that this occurred as the result of deamidation in the upper band, since SDS-PAGE often reflects amino acid modifications.[46]

We previously reported that E-cadherin turnover is significantly delayed in melanoma cells transfected with *GnT-III*, which is involved in the regulation of branch formation in *N*-glycans.[25] We also showed that the EGFR or Src-mediated tyrosine phosphorylation of β-catenin was decreased in GnT-III transfectants.[26] Guo *et al.* reported that the EGF-induced tyrosine phosphorylation of β-catenin and P120^{ctn} increased in GnT-V transfectants, which led to a reduction in cell–cell adhesion.[27] The fact that GnT-III activity was not changed by *Fut8* transfection suggests that Fut8 activity regulates the expression of E-cadherin independent of GnT-III, and the accumulation of both normal form and the low molecular weight E-cadherin enhances cell–cell adhesion.

It has been revealed that core fucosylation catalyzed by Fut8 is involved in various biological phenomena. *Fut8* transgenic

mice caused steatosis in the liver and kidney due to a decreased lysosomal acid lipase activity.[47] Core fucose deficient IgG1 showed an improved binding to FcγRIIIA, and consequently, antibody-dependent cellular cytotoxicity activity was up-regulated.[48,49] We developed *Fut8*$^{-/-}$ mice and reported that the mice showed semi lethality, growth retardation, and emphysema-like changes, and the experiments indicated that dysfunction of TGF-β1 receptor[32], impairment in the low-density LRP-1[41] and functional changes of α3β1 integrin[50] are involved in the phenomena. The regulation of the cell surface expression levels of E-cadherin, reported herein, could be involved in the pathology observed in *Fut8* transgenic mice and *Fut8* knock out mice. Metastasis analysis using those animals would indicate the role of Fut8 in cell–cell adhesion *in vivo*.

Acknowledgments

We thank Dr Tadashi Suzuki for helpful comments and discussions. This work was supported by grants from the 21st Century COE Program from Japan Society for the Promotion of Science; Core Research for Evolutional Science and Technology from Japan Science and Technology Agency; and Special Coordination Funds for Promoting Science and Technology and Scientific Research (A) from the Ministry of Education, Culture, Sports, Science and Technology, Japan.

References

1 Hakomori S. Tumor malignancy defined by aberrant glycosylation and sphingo (glyco) lipid metabolism. *Cancer Res* 1996; **56**: 5309–18.
2 Rudd PM, Elliott T, Cresswell P, Wilson IA, Dwek RA. Glycosylation and the immune system. *Science* 2001; **291**: 2370–6.
3 Taniguchi N, Ekuni A, Ko JH *et al.* A glycomic approach to the identification and characterization of glycoprotein function in cells transfected with glycosyltransferase genes. *Proteomics* 2001; **1**: 239–47.
4 Lau KS, Partridge EA, Grigorian A *et al.* Complex N-glycan number and degree of branching cooperate to regulate cell proliferation and differentiation. *Cell* 2007; **129**: 123–34.
5 Taniguchi N, Miyoshi E, Gu J, Honke K, Matsumoto A. Decoding sugar functions by identifying target glycoproteins. *Curr Opin Struct Biol* 2006; **16**: 561–6.
6 Taniguchi N, Miyoshi E, Ko JH, Ikeda Y, Ihara Y. Implication of N-acetylglucosaminyltransferases III and V in cancer: gene regulation and signaling mechanism. *Biochim Biophys Acta* 1999; **1455**: 287–300.
7 Miyoshi E, Noda K, Yamaguchi Y *et al.* The alpha1–6-fucosyltransferase gene and its biological significance. *Biochim Biophys Acta* 1999; **1473**: 9–20.
8 Uozumi N, Yanagidani S, Miyoshi E *et al.* Purification and cDNA cloning of porcine brain GDP-L-Fuc: N-acetyl-beta-D-glucosaminide alpha1–>6fucosyltransferase. *J Biol Chem* 1996; **271**: 27810–7.
9 Yanagidani S, Uozumi N, Ihara Y *et al.* Purification and cDNA cloning of GDP-L-Fuc: N-acetyl-beta-D-glucosaminide: alpha1-6 fucosyltransferase (alpha1–6 FucT) from human gastric cancer MKN45 cells. *J Biochem (Tokyo)* 1997; **121**: 626–32.
10 Shimizu K, Katoh H, Yamashita F *et al.* Comparison of carbohydrate structures of serum alpha-fetoprotein by sequential glycosidase digestion and lectin affinity electrophoresis. *Clin Chim Acta* 1996; **254**: 23–40.
11 Miyoshi E, Uozumi N, Noda K *et al.* Expression of alpha1-6 fucosyltransferase in rat tissues and human cancer cell lines. *Int J Cancer* 1997; **72**: 1117–21.
12 Noda K, Miyoshi E, Uozumi N *et al.* High expression of alpha-1-6 fucosyltransferase during rat hepatocarcinogenesis. *Int J Cancer* 1998; **75**: 444–50.
13 Takahashi T, Ikeda Y, Miyoshi E *et al.* alpha1,6fucosyltransferase is highly and specifically expressed in human ovarian serous adenocarcinomas. *Int J Cancer* 2000; **88**: 914–19.
14 Takeichi M. Cadherin cell adhesion receptors as a morphogenetic regulator. *Science* 1991; **251**: 1451–5.
15 Gumbiner BM. Regulation of cadherin-mediated adhesion in morphogenesis. *Nat Rev Mol Cell Biol* 2005; **6**: 622–34.
16 Shapiro L, Fannon AM, Kwong PD *et al.* Structural basis of cell-cell adhesion by cadherins. *Nature* 1995; **374**: 327–37.
17 Boggon TJ, Murray J, Chappuis-Flament S *et al.* C-cadherin ectodomain structure and implications for cell adhesion mechanisms. *Science* 2002; **296**: 1308–13.
18 Pertz O, Bozic D, Koch AW *et al.* A new crystal structure, Ca^{2+} dependence and mutational analysis reveal molecular details of E–cadherin homoassociation. *Embo J* 1999; **18**: 1738–47.
19 Chappuis-Flament S, Wong E, Hicks LD, Kay CM, Gumbiner BM. Multiple cadherin extracellular repeats mediate homophilic binding and adhesion. *J Cell Biol* 2001; **154**: 231–43.
20 Zhu B, Chappuis-Flament S, Wong E *et al.* Functional analysis of the structural basis of homophilic cadherin adhesion. *Biophys J* 2003; **84**: 4033–42.
21 Peyrieras N, Hyafil F, Louvard D, Ploegh HL, Jacob F. Uvomorulin: a nonintegral membrane protein of early mouse embryo. *Proc Natl Acad Sci USA* 1983; **80**: 6274–7.
22 Vestweber D, Kemler R. Some structural and functional aspects of the cell adhesion molecule uvomorulin. *Cell Differ* 1984; **15**: 269–73.
23 Shore EM, Nelson WJ. Biosynthesis of the cell adhesion molecule uvomorulin (E-cadherin) in Madin-Darby canine kidney epithelial cells. *J Biol Chem* 1991; **266**: 19672–80.
24 Ozawa M, Kemler R. Correct proteolytic cleavage is required for the cell adhesive function of uvomorulin. *J Cell Biol* 1990; **111**: 1645–50.
25 Yoshimura M, Ihara Y, Matsuzawa Y, Taniguchi N. Aberrant glycosylation of E-cadherin enhances cell–cell binding to suppress metastasis. *J Biol Chem* 1996; **271**: 13811–5.
26 Kitada T, Miyoshi E, Noda K *et al.* The addition of bisecting N-acetylglucosamine residues to E-cadherin down-regulates the tyrosine phosphorylation of beta-catenin. *J Biol Chem* 2001; **276**: 475–80.
27 Guo HB, Lee I, Kamar M, Pierce M. N-acetylglucosaminyltransferase V expression levels regulate cadherin-associated homotypic cell-cell adhesion and intracellular signaling pathways. *J Biol Chem* 2003; **278**: 52412–24.
28 Miyoshi E, Noda K, Ko JH *et al.* Overexpression of alpha1-6 fucosyltransferase in hepatoma cells suppresses intrahepatic metastasis after splenic injection in athymic mice. *Cancer Res* 1999; **59**: 2237–43.
29 Liwosz A, Lei T, Kukuruzinska MA. N-glycosylation affects the molecular organization and stability of E-cadherin junctions. *J Biol Chem* 2006; **281**: 23138–49.
30 Takahashi T, Ikeda Y, Tateishi A *et al.* A sequence motif involved in the donor substrate binding by alpha1,6-fucosyltransferase: the role of the conserved arginine residues. *Glycobiology* 2000; **10**: 503–10.
31 Li W, Nakagawa T, Koyama N *et al.* Down-regulation of trypsinogen expression is associated with growth retardation in alpha1,6-fucosyltransferase-deficient mice: attenuation of proteinase-activated receptor 2 activity. *Glycobiology* 2006; **16**: 1007–19.
32 Wang X, Inoue S, Gu J *et al.* Dysregulation of TGF-beta1 receptor activation leads to abnormal lung development and emphysema-like phenotype in core fucose-deficient mice. *Proc Natl Acad Sci U S A* 2005; **102**: 15791–6.
33 Uozumi N, Teshima T, Yamamoto T *et al.* A fluorescent assay method for GDP-L-Fuc: N-acetyl-beta-D-glucosaminide alpha1-6fucosyltransferase activity, involving high performance liquid chromatography. *J Biochem (Tokyo)* 1996; **120**: 385–92.
34 Ito Y, Miyauchi A, Yoshida H *et al.* Expression of alpha1,6-fucosyltransferase (FUT8) in papillary carcinoma of the thyroid: its linkage to biological aggressiveness and anaplastic transformation. *Cancer Lett* 2003; **200**: 167–72.
35 Hobert ME, Kil SJ, Medof ME, Carlin CR. The cytoplasmic juxtamembrane domain of the epidermal growth factor receptor contains a novel autonomous basolateral sorting determinant. *J Biol Chem* 1997; **272**: 32901–9.
36 Sato Y, Takahashi M, Shibukawa Y *et al.* Overexpression of N-acetylglucosaminyltransferase III enhances the epidermal growth factor-induced phosphorylation of ERK in HeLaS3 cells by up-regulation of the internalization rate of the receptors. *J Biol Chem* 2001; **276**: 11956–62.
37 Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2 (-Delta Delta C(T) method. *Methods* 2001; **25**: 402–8.
38 Hakomori S. Glycosylation defining cancer malignancy: new wine in an old bottle. *Proc Natl Acad Sci U S A* 2002; **99**: 10231–3.
39 Shibukawa Y, Takahashi M, Laffont I, Honke K, Taniguchi N. Down-regulation of hydrogen peroxide-induced PKC delta activation in N-acetylglucosaminyltransferase III-transfected HeLaS3 cells. *J Biol Chem* 2003; **278**: 3197–203.
40 Takahashi M, Tsuda T, Ikeda Y, Honke K, Taniguchi N. Role of N-glycans in growth factor signaling. *Glycoconj J* 2004; **20**: 207–12.
41 Lee SH, Takahashi M, Honke K *et al.* Loss of core fucosylation of low-density lipoprotein receptor-related protein-1 impairs its function, leading to the upregulation of serum levels of insulin-like growth factor-binding protein 3 in Fut8$^{-/-}$ mice. *J Biochem (Tokyo)* 2006; **139**: 391–8.
42 Yokoe S, Takahashi M, Asahi M *et al.* The Asn418-linked N-glycan of ErbB3 plays a crucial role in preventing spontaneous heterodimerization and tumor promotion. *Cancer Res* 2007; **67**: 1935–42.

doi: 10.1111/j.1349-7006.2009.01125.x

43 Li W, Takahashi M, Shibukawa Y *et al.* Introduction of bisecting GlcNAc in N-glycans of adenylyl cyclase III enhances its activity. *Glycobiology* 2007; **17**: 655–62.
44 Przybylo M, Hoja-Lukowicz D, Litynska A, Laidler P. Different glycosylation of cadherins from human bladder non-malignant and cancer cell lines. *Cancer Cell Int* 2002; **2**: 6.
45 Iijima J, Zhao Y, Isaji T *et al.* Cell–cell interaction-dependent regulation of N-acetylglucosaminyltransferase III and the bisected N-glycans in GE11 epithelial cells. Involvement of E-cadherin-mediated cell adhesion. *J Biol Chem* 2006; **281**: 13038–46.
46 Fujiwara N, Miyamoto Y, Ogasahara K *et al.* Different immunoreactivity against monoclonal antibodies between wild-type and mutant copper/zinc superoxide dismutase linked to amyotrophic lateral sclerosis. *J Biol Chem* 2005; **280**: 5061–70.
47 Wang W, Li W, Ikeda Y *et al.* Ectopic expression of alpha1,6 fucosyltransferase in mice causes steatosis in the liver and kidney accompanied by a modification of lysosomal acid lipase. *Glycobiology* 2001; **11**: 165–74.
48 Shields RL, Lai J, Keck R *et al.* Lack of fucose on human IgG1 N-linked oligosaccharide improves binding to human Fcgamma RIII and antibody-dependent cellular toxicity. *J Biol Chem* 2002; **277**: 26733–40.
49 Shinkawa T, Nakamura K, Yamane N *et al.* The absence of fucose but not the presence of galactose or bisecting N-acetylglucosamine of human IgG1 complex-type oligosaccharides shows the critical role of enhancing antibody-dependent cellular cytotoxicity. *J Biol Chem* 2003; **278**: 3466–73.
50 Zhao Y, Itoh S, Wang X *et al.* Deletion of core fucosylation on alpha3beta1 integrin down-regulates its functions. *J Biol Chem* 2006; **281**: 38343–50.

W166–W169 *Nucleic Acids Research, 2009, Vol. 37, Web Server issue* *Published online 4 June 2009*
doi:10.1093/nar/gkp483

Gendoo: Functional profiling of gene and disease features using MeSH vocabulary

Takeru Nakazato[1,2,*], Hidemasa Bono[1], Hideo Matsuda[2] and Toshihisa Takagi[1]

[1]Database Center for Life Science (DBCLS), Research Organization of Information and Systems (ROIS), Faculty of Engineering Building 12, The University of Tokyo, 2-11-16 Yayoi, Bunkyo-ku, Tokyo 113-0032 and [2]Department of Bioinformatic Engineering, Graduate School of Information Science and Technology, Osaka University, 1-3 Machikaneyama, Toyonaka, Osaka 560-8531, Japan

Received February 15, 2009; Revised April 27, 2009; Accepted May 19, 2009

ABSTRACT

Genome-wide data enables us to clarify the underlying molecular mechanisms of complex phenotypes. The Online Mendelian Inheritance in Man (OMIM) is a widely employed knowledge base of human genes and genetic disorders for biological researchers. However, OMIM has not been fully exploited for omics analysis because its bibliographic data structure is not suitable for computer automation. Here, we characterized diseases and genes by generating feature profiles of associated drugs, biological phenomena and anatomy with the MeSH (Medical Subject Headings) vocabulary. We obtained 1 760 054 pairs of OMIM entries and MeSH terms by utilizing the full set of MEDLINE articles. We developed a web-based application called Gendoo (gene, disease features ontology-based overview system) to visualize these profiles. By comparing feature profiles of types 1 and 2 diabetes, we clearly illustrated their differences: type 1 diabetes is an autoimmune disease (*P*-value = 4.55×10^{-5}) and type 2 diabetes is related to obesity (*P*-value = 1.18×10^{-15}). Gendoo and the developed feature profiles should be useful for omics analysis from molecular and clinical viewpoints. Gendoo is available at http://gendoo.dbcls.jp/.

INTRODUCTION

The major aims of omics analysis are to identify disease-relevant genes and to understand their mechanisms. Genome sequences and transcriptomics provide large amounts of data, and researchers have attempted to interpret these genetic data in conjunction with clinical phenotypes (1–3). To analyze these data, we can easily obtain gene information such as gene names and genomic location, and their features in the form of Gene Ontology (GO) terms (4) from Entrez Gene (5,6) and Ensembl (7). Additionally, as a disease database, we generally refer to the Online Mendelian Inheritance in Man (OMIM: http://www.ncbi.nlm.nih.gov/omim/) (8,9).

OMIM contains nearly 18 000 detailed entries for human genes and genetic disorders. OMIM is a useful resource for obtaining information about diseases. However, it is difficult to utilize OMIM's data for omics analysis because almost all of its sections are written in natural language, namely English sentences (10). To enable computers to handle OMIM data, certain studies (11–15) have organized OMIM by selecting terms referred to in the Clinical Synopsis (CS) section as keywords. The CS section describes clinical features of disorders and their mode of inheritance such as 'autosomal dominant'. Some of the terms in the CS section for Prader–Willi syndrome (OMIM ID: #176270) are shown in Table 1 as an example. Previous studies (12,14) characterized diseases according to corresponding tissue and etiology with CS terms. By using these terms, researchers do not have to use text mining techniques to automatically extract disease information from OMIM for omics analysis. However, even though OMIM includes detailed biological and genetic descriptions, CS terms are mainly clinical and diagnostic terms so that it is difficult to decipher disease information in conjunction with biological process data such as gene expression data. In addition, CS terms, such as 'Cardiac' and 'Cardiovascular', are ambiguous because the assigned terms are often defined by the author's original description of the cited articles (8).

Here, to organize the disease features referred to in OMIM, we attempted to use the MeSH (Medical Subject Headings) controlled vocabulary (16). MeSH contains >20 000 keywords and hierarchically categorized into 15 concepts including 'disease', 'chemicals and drugs' and 'anatomy'. It is originally curated for indexing MEDLINE articles by National Library of Medicine (NLM). In our previous study (17), to annotate genes from biological viewpoint excluded by GO such as disease and drug fields, we assigned MeSH to each gene by using

*To whom correspondence should be addressed. Tel: +81 3 5841 6754; Fax: +81 3 5841 8090; Email: nakazato@dbcls.rois.ac.jp

Table 1. Symptoms referred to in OMIM Clinical Synopsis section for Prader–Willi syndrome (partial)

Inheritance:
Isolated cases
Growth:
Height
Mean adult male height, 155 cm
Mean adult female height, 147 cm
Steady childhood growth
Weight
Onset of obesity from 6 months to 6 years
Central obesity
Respiratory:
Hypoventilation
Hypoxia
Skeletal:
Osteoporosis
Osteopenia
Endocrine features:
Hyperinsulinemia
Growth hormone deficiency
Hypogonadotropic hypogonadism
Miscellaneous:
Food related behavioral problems include excessive appetite and obsession with eating
Temperature instability
High pain threshold
Molecular basis:
Microdeletion of 15q11 in 70% of patients confirmed by fluorescent in situ hybridization
Remainder of cases secondary to maternal disomy
Rare cases secondary to chromosome translocation

Clinical features of a disorder are listed in the Clinical Synopsis (CS) section of the OMIM database. The CS section mainly describes morphologies and events in clinical and diagnostic fields. Each feature is itemized, but a controlled vocabulary is not used.

Entrez Gene as gene data. In this article, we therefore generated feature profiles of diseases by applying MeSH to OMIM data with the method previously described (17). By comparing these feature profiles of genes developed (17) and diseases derived from this work, we aim to assist to interpret omics data from the molecular and clinical aspects.

METHODS

Data collection

We retrieved OMIM data available in February 2008 by downloading from the National Center for Biotechnology Information (NCBI) FTP site (ftp://ftp.ncbi.nih.gov/repository/OMIM/) and by using the web service with Entrez Programming Utilities (http://eutils.ncbi.nlm.nih.gov/entrez/query/static/eutils_help.html). We obtained MeSH terms (2008 release) from the NLM web site (http://www.nlm.nih.gov/mesh/meshhome.html).

Articles extraction related to each OMIM entry

To generate OMIM–MeSH associations, we need to retrieve articles referred to in each OMIM entry because MeSH terms are not assigned to OMIM entries directly, but to MEDLINE. A schematic view of the pipeline for generating OMIM–MeSH associations is shown in Supplementary Figure S1. We retrieved PubMed IDs (PMIDs) cited in the reference section of OMIM (Supplementary Figure S1a) and extracted OMIM IDs described in the abstracts in MEDLINE (Supplementary Figure S1b). We also retrieved PMIDs by searching PubMed by inputting disease names (Supplementary Figure S1c). One of the problems is that one disease often has many names (18), e.g. 'type 2 diabetes', 'non-insulin dependent diabetes' and 'NIDDM'. Another problem is that the same abbreviation may refer to several diseases, genes and drugs (19); for example, 'EVA' refers to 'enlarged vestibular aqueduct' (disease), 'epithelial V-like antigen' (gene) and 'ethylene vinyl acetate' (chemical). We therefore created abbreviation/long-form pairs for disease names such as 'PWS' and 'Prader–Willi syndrome' and searched MEDLINE for articles co-occurring with both names. Accordingly, we retrieved 426 141 unique OMIM ID and PMID pairs and generated 1 760 054 OMIM–MeSH pairs.

Scoring of associations between OMIM entries and MeSH terms

OMIM contains gene entries as molecular mechanisms and disease entries as their phenotypes (8). These types are indicated by symbols prefixed to the OMIM ID. We divided the OMIM entries into three groups according to these types: sequence known (*, +), locus known (%) and phenotype (#, none). We then calculated *P*-values as a score of OMIM–MeSH pairs in each group. The *P*-value is the probability of the actual or a more extreme outcome under the null-hypothesis. The lower *P*-value means the larger significance of association. We also calculated information gain to rank the associations of the OMIM–MeSH pairs as described in (17). Briefly, information gain refers to the frequency of co-occurrence of a disease name and a MeSH term and also refers to the specificity of the MeSH term.

Data visualization

We updated the web-based software application called Gendoo (gene, disease features ontology-based overview system) to visualize associations between OMIM entries and relevant MeSH terms. It was originally developed to visualize gene–MeSH associations (17). Gendoo accepts OMIM IDs, OMIM titles, Entrez Gene IDs, gene names and MeSH terms as input queries. For disease names, Gendoo currently uses descriptions of 'title' and 'alternative titles; symbols' sections of OMIM, so that not all synonyms are included in the disease name dictionary. We will increase the synonyms by involving the canonical name and synonyms (entry terms) of corresponding MeSH terms, and extracting disease names from MEDLINE and OMIM resources with text mining approach. Gendoo generates high-scoring lists that display relevant MeSH terms for diseases, drugs, biological phenomena and anatomy together with their scores (Supplementary Figure. S2a). These MeSH terms are sorted according to their information gain, and the background color of each association indicates its *P*-value. Gendoo also gives a hierarchical-tree view of MeSH terms associated with diseases of interest by using

JavaScript and cascading style sheet (CSS) resources from the Yahoo! User Interface (YUI) library (http://developer.yahoo.com/yui/) (Supplementary Figure S2b).

RESULTS

Table 2 lists top-three keywords related to Prader–Willi syndrome for the features of the 'Disease', 'Chemicals and Drugs', 'Biological Phenomena' and 'Anatomy' fields. Prader–Willi syndrome results from deletion of paternal copies of the imprinted SNRPN (small nuclear ribonucleoprotein polypeptide N) and necdin genes within chromosome 15 (20). Gendoo shows the keyword phrases clearly reflecting the features of Prader–Willi syndrome, including 'Chromosomes, Human, Pair 15', 'Genomic Imprinting' and 'Ribonucleoproteins, Small Nuclear'. Gendoo illustrates the disease features from not only a clinical perspective, but also a biological one, unlike the symptoms referred to in the CS section shown in Table 1. To retrieve more clinical and diagnostic features with MeSH, we can increase the number of novel associations by using terms from the 'Analytical, Diagnostic and Therapeutic Techniques and Equipment' category of MeSH.

We applied this analysis to types 1 and 2 diabetes (OMIM IDs are %222100 and #125853, respectively). Figure 1 summarizes the feature profiles; type 1 diabetes is closely related to 'Autoimmune Diseases' and 'Spleen' (their *P*-values are 4.55×10^{-5} and 5.53×10^{-7}, respectively), whereas type 2 diabetes is associated with 'Obesity' (*P*-value = 1.18×10^{-15}) and 'Adipocytes' (*P*-value = 5.17×10^{-5}). Type 1 diabetes is involved in immune systems, and type 2 diabetes is a metabolic disorder (21). This result suggests that the MeSH profiles produced by Gendoo can clarify the differences and similarities in features between OMIM entries.

We provide more practical results shown in Supplementary Table S1.

Table 2. Lists of top-three keywords related to Prader–Willi syndrome

MeSH terms	*P*-value
Diseases	
Prader–Willi syndrome	0
Angelman syndrome	4.05×10^{-140}
Obesity	6.94×10^{-128}
Chemicals and Drugs	
Human growth hormone	5.86×10^{-68}
Ribonucleoproteins, small nuclear	4.29×10^{-62}
Ghrelin	1.58×10^{-50}
Biological Phenomena	
Chromosomes, human, pair 15	0
Genomic imprinting	2.47×10^{-131}
Obesity	1.69×10^{-121}
Anatomy	
Chromosomes, human, pair 15	0
Chromosomes, human, 13–15	1.25×10^{-30}
Adipose tissue	3.93×10^{-13}

We generated feature profiles by using the MeSH vocabulary. Unlike the symptoms referred to in the CS section of OMIM (Table 1), these profiles give not only clinical, but also biological information about the disease.

The Mendelian Inheritance in Man (MIM) is an excellent knowledge bank that has been annotated by Dr McKusick and his colleagues for >40 years, and its online version, OMIM, is accessible through the internet from NCBI (22). However, its bibliographic data structure has prevented OMIM from being fully exploited for omics analysis. To alleviate this problem, we comprehensively characterized human genes and genetic disorders referred to in OMIM with the MeSH vocabulary, and this will enable researchers to decipher their genome-wide data in conjunction with clinical phenotypes by using Gendoo. For example, the developed feature profiles can be applied to analyses of disease-relevant genes by comparing the similarities among profiles of OMIM entries and groups of genes such as those found in the clustering results of gene expression data. Researchers can also make overviews of features of unfamiliar diseases with Gendoo (Supplementary Table S1c and d).

AVAILABILITY

Gendoo can be openly accessed at http://gendoo.dbcls.jp/. Every association file including Entrez Gene/OMIM IDs, MeSH and their scores is available from the web site. Dictionary files including gene/disease names, synonyms and IDs are also downloadable. These web service and files are freely available under a Creative Commons Attribution 2.1 Japan license (http://creativecommons.org/licenses/by/2.1/jp/deed.en).

CONCLUSIONS

We characterized diseases and genes by generating feature profiles of associated drugs, biological phenomena and anatomy with the MeSH vocabulary and developed a web-based application called Gendoo to visualize these

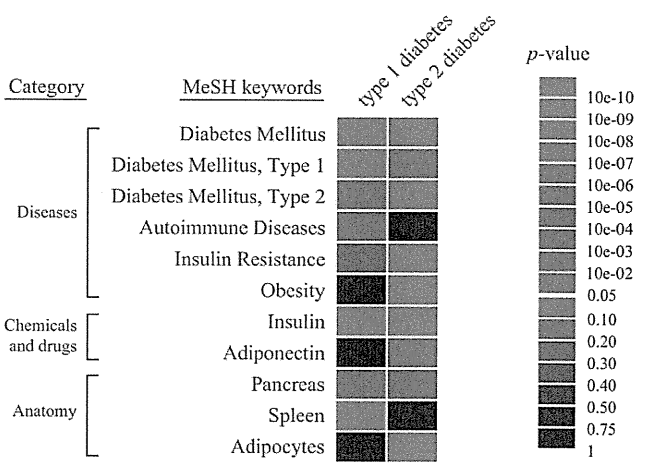


Figure 1. Differences and similarities between feature profiles of types 1 and 2 diabetes. Typical features and scores of types 1 and 2 diabetes are shown. The background colors of each association reflect the *P*-value. Type 1 diabetes is an autoimmune disorder, whereas type 2 diabetes is a metabolic disorder. These profiles clarify the differences between the features of these diseases.