Table 2. Cell Cycle Perturbations Induced by Pemetrexed (MTA), Cisplatin (CDDP), and Their Combinations for PA1 and MCF7 Cells at 48 h

	MTA + CDDP (24 h)				MTA (24 h) → CDDP (24 h)				CDDP (24 h)→ MTA (24 h)			
Cell Cycle (%)	Control	MTA	CDDP	MTA + CDDP	Control	MTA	CDDP	MTA + CDDP	Control	MTA	CDDP	MTA + CDDP
PA1 cells												
Sub-G_1	3.6	2.4	42.9	2.1	4.3	3.1	8.9	15.3	2.9	2.2	45.1	41.8
G_1	56.2	64.1	7.3	67.1	58.1	65.3	5.8	4.4	57.3	60.1	6.9	10.6
S	15.6	26.7	17.2	19.1	10.4	25.9	48.4	38.7	11.0	30.4	15.8	20.1
G_2/M	24.6	6.8	19.1	11.7	27.2	5.7	36.9	41.6	28.8	7.3	32.2	27.5
MCF-7 cells												
Sub-G_1	4.2	17.5	3.9	5.8	5.3	11.1	2.9	16.8	5.1	10.3	3.6	2.5
G_1	57.6	53.4	28.8	63.7	55.8	61.3	22.3	60.6	58.8	57.2	27.9	25.8
S	16.8	26.9	4.7	21.4	19.1	22.1	21.2	13.8	16.4	28.6	5.0	20.4
G_2/M	21.4	2.2	62.6	9.1	25.1	5.5	53.6	8.8	19.7	3.9	63.5	51.3

early S phase, in which cells are sensitive to cisplatin (20). This may explain the synergistic effects of sequential exposure to pemetrexed followed by cisplatin. On the contrary, one agent may reduce the cytotoxicity of the other agent by preventing cells from entering the specific phase in which the cells are most cytotoxic to the other agent. It has been shown that cisplatin elicits cytotoxic effects by blocking cells in G_2/M phase (20), while pemetrexed does by blocking cells in S phase (21). Indeed, simultaneous exposure to pemetrexed and cisplatin produced antagonistic effects, which were caused by the cancellation of cisplatin-induced G_2/M arrest by coexisting pemetrexed in PA1 and MCF7 cells. This was also the case with sequential exposure with cisplatin first followed by pemetrexed.

Our findings suggest that the sequential administration of pemetrexed followed by cisplatin may be the optimal schedule for these combinations. For example, administrations of pemetrexed on day 1 and cisplatin on day 2 would be worthy of clinical investigations. The simultaneous administration of pemetrexed and cisplatin and the sequential administration of cisplatin followed by pemetrexed may be inadequate. However, it must be noted that there are a number of difficulties in the translation of results from in vitro models to clinical therapy. The drug metabolism and pharmacokinetics under in vivo and in vitro conditions are different. Clinical outcome includes both the antitumor effects and normal tissue toxicity that results from a variable drug exposure, whereas in vitro models represent only antitumor effects at a constant drug exposure.

Teicher et al. studied the combination of pemetrexed with cisplatin in vivo against EMT-6 murine mammary carcinoma by a tumor cell survival assay (26). They observed that pemetrexed administered four times over 48 h with cisplatin administered with the third dose of pemetrexed produced an additive or more than additive tumor response. Teicher et al. further studied the combination of pemetrexed with cisplatin in human tumor xenografts (27). Administration of pemetrexed (days 7–11, days 14–18) along with cisplatin (day 7) produced greater-than-additive effects for human lung cancer H460 and Calu-6 tumor growth delay. Because experimental systems, schedules of drug administrations, and evaluating methods for synergism are different, it is difficult to compare their findings and ours.

A clinical and pharmacokinetic phase I study of pemetrexed in combination with cisplatin has been reported by Thordtmann et al. (15). They observed that this combination was clinically active and simultaneous administration of both agents on day on 1 (pemetrexed intravenously over 10 min and cisplatin over 2 h) every 21 days was less toxic than a sequential administration of pemetrexed on day 1 and cisplatin on day 2. They recommended the simultaneous administration of pemetrexed at 500 mg/m^2 plus cisplatin at 75 mg/m^2 on day 1 every 21 days for this combination. Phase II and III studies of the same schedules have been started for this combination and encouraging results have been obtained so far (16–18).

Our in vitro findings are not contradictory to clinical findings. In our study, simultaneous exposure to pemetrexed and cisplatin produced additive effects in WiDr cells and antagonistic effects in A549, MCF7, andPA1 cells. Most data points fell in the area of sudadditivity in MCF7 and PA1 cells, suggesting that the combination is superior to each drug alone but "sub-optimal." The simultaneous administration of pemetrexed and cisplatin was less toxic than the sequential administration, probably due to antagonistic interaction in the simultaneous exposure. Our isobologram shows that the doses of both agents in the pemetrexed–cisplatin sequence required

for IC_{80} or IC_{50} levels were much less (40–90%) than of those in simultaneous exposure (Fig. 3). Pemetrexed at 500 mg/m^2 and cisplatin at 75 mg/m^2, the optimal dose for the simultaneous administration, would be overdosed for the sequential administration of pemetrexed followed by cisplatin, which produced synergistic effects.

In conclusion, the present findings show that the interaction of pemetrexed and cisplatin is definitely schedule dependent. Sequential exposure to pemetrexed followed by cisplatin produced synergistic effects, whereas simultaneous exposure to the two agents and sequential exposure to cisplatin followed by pemetrexed produced antagonistic effects. These findings suggest that the optimal schedule of pemetrexed in combination with cisplatin at the cellular level is the sequential administration of pemetrexed followed by cisplatin. Although the simultaneous administration of pemetrexed and cisplatin on day 1 is more convenient and less toxic for patients than the sequential administration of pemetrexed on day 1 and cisplatin on day 2, the former schedule may be suboptimal and may not improve the clinical efficacy to "originally expected" level for this combination. It would be important to conduct dose-finding clinical trials in sequential administration of pemetrexed and cisplatin.

ACKNOWLEDGMENTS: This work was supported in part by a Grant-in-Aid for Cancer Research (11-8) from the Ministry of Health and Welfare and by a Grant-in-Aid for Research on the Second-Term Comprehensive 10-Year Strategy for Cancer Control from the Ministry of Health and Welfare of Japan.

REFERENCES

1. Taylor, E. C.; Kuhnt, D.; Shih, C.; Rinzel, S. M.; Grindey, G. B.; Barredo, J.; Jannatipour, M.; Moran, R. G. A dideazatetrahydrofolate analogue lacking a chiral center at C-6 N-[4-[2-(2-amino-34-dihydro-4-oxo-7H-pyrrolo[23-d]pyrimidin-5- yl)ethyl]benzoyl]-L-glutamic acid is an inhibitor of thymidylate synthase. J. Med. Chem. 35:4450–4454; 1992.
2. Habeck, L. L.; Mendelsohn, L. G.; Shih, C.; Taylor, E. C.; Colman, P. D.; Gossett, L. S.; Leitner, T. A.; Schultz, R. M.; Andis, S. L.; Moran, R. G. Substrate specificity of mammalian folypolyglutamate synthetase for 510-dideazatetrahydrofolate analogs. Mol. Pharmacol. 48: 326–333; 1995.
3. Shih, C.; Habeck, L. L.; Mendelsohn, L. G.; Chen, V. J.; Schultz, R. M. Multiple folate enzyme inhibition: Mechanism of a novel pyrrolopyrimidine-based antifolate LY231514 (MTA). Adv. Enzyme Regul. 38:135–152; 1998.
4. Shih, C.; Thornton, D. E. Preclinical pharmacology studies and the clinical development of a novelmultitargeted antifolate MTA (LY231514) In: Jackman, A. L., ed. Anticancer drug development guide: Antifolate drugs in cancer therapy. Totowa, NJ: Humana Press; 1998:183–201.
5. McDonald, A. C.; Vasey, P. A.; Adams, L.; Walling, J.; Woodworth, J. R.; Abrahams, T.; McCarthy, S.; Bailey, N. P.; Siddiqui, N.; Lind, M. J.; Calvert, A. H.; Twelves, C. J.; Cassidy, J.; Kaye, S. B. A phase I and pharmacokinetic study of LY231514 the multitargeted antifolate. Clin. Cancer Res. 4:605–610; 1998.
6. Rinaldi, D. A. Overview of phase I trials of multitargeted antifolate (MTA LY231514). Semin. Oncol. 26(Suppl. 6): 82–88; 1999.
7. Rusthoven, J. J.; Eisenhauer, E.; Butts, C.; Gregg, R.; Dancey, J.; Fisher, B.; Iglesias, J. Multitargeted antifolate LY231514 as first-line chemotherapy for patients with advanced non-small-cell lung cancer: A phase II study. National Cancer Institute of Canada Clinical Trials Group. J. Clin. Oncol. 17:1194–1199; 1999.
8. John, W.; Picus, J.; Blanke, C. D.; Clark, J. W.; Schulman, L. N.; Rowinsky, E. K.; Thornton, D. E.; Loehrer, P. J. Activity of multitargeted antifolate (pemetrexed disodium LY231514) in patients with advanced colorectal carcinoma: Results from a phase II study. Cancer 88:1807–1813; 2000.
9. Hanauske, A. R.; Chen, V.; Paoletti, P.; Niyikiza, C. Pemetrexed disodium: A novel antifolate clinically active against multiple solid tumors. Oncologist 6:363–373; 2001.
10. Pivot, X.; Raymond, E.; Laguerre, B.; Degardin, M.; Cals, L.; Armand, J. P.; Lefebvre, J. L.; Gedouin, D.; Ripoche, V.; Kayitalire, L.; Niyikiza, C.; Johnson, R.; Latz, J.; Schneider, M. Pemetrexed disodium in recurrent locally advanced or metastatic squamous cell carcinoma of the head and neck. Br. J. Cancer 85:649–655; 2001.
11. Shepherd, F. A. Pemetrexed in the treatment of non-small cell lung cancer. Semin. Oncol. 29(Suppl. 18):43–48; 2002.
12. Calvert, H. Pemetrexed (Alimta): A promising new agent for the treatment of breast cancer. Semin. Oncol. 30 (Suppl. 3):2–5; 2003.
13. Scagliotti, G. V.; Shin, D. M.; Kindler, H. L.; Scagliotti, G. V.; Shin, D. M.; Kindler, H. L.; Vasconcelles, M. J.; Keppler, U.; Manegold, C.; Burris, H.; Gatzemeier, U.; Blatter, J.; Symanowski, J. T.; Rusthoven, J. J. Phase II study of pemetrexed with and without folic acid and vitamin B12 as front-line therapy in malignant therapy in malignant pleural mesothelioma. J. Clin. Oncol. 21:1556–1561; 2003.
14. Hanna, N.; Shepherd, F. A.; Fossella, F. V.; Pereira, J. R.; De Marinis, F.; von Pawel, J.; Gatzemeier, U.; Tsao, T. C.; Pless, M.; Muller, T.; Lim, H. L.; Desch, C.; Szondy, K.; Gervais, R.; Shaharyar; Manegold, C.; Paul, S.; Paoletti, P.; Einhorn, L.; Bunn, Jr., P. A. Randomized phase III trial of pemetrexed versus docetaxel in patients with non-small-cell lung cancer previously treated with chemotherapy. J. Clin. Oncol. 22:1589–1597; 2004.
15. Thodtmann, R.; Depenbrock, H.; Dumez, H.; Blatter, J.; Johnson, R. D.; van Oosterom, A.; Hanauske, A. R. Clinical and pharmacokinetic phase I study of multitargeted antifolate (LY231514) in combination with cisplatin. J. Clin. Oncol. 17:3009–3016; 1999.
16. Manegold, C.; Gatzemeier, U.; von Pawel, J.; Pirker, R.; Malayeri, R.; Blatter, J.; Krejcy, K. Front-line treatment of advanced non-small-cell lung cancer with MTA (LY231514 pemetrexed disodium ALIMTA) and cisplatin: A multicenter phase II trial. Ann. Oncol. 11:435–440; 2000.
17. Shepherd, F. A.; Dancey, J.; Arnold, A.; Neville, A.; Rusthoven, J.; Johnson, R. D.; Fisher, B.; Eisenhauer, E. Phase II study of pemetrexed disodium a multitargeted antifolate and cisplatin as first-line therapy in patients with advanced non small cell lung carcinoma: A study

of the National Cancer Institute of Canada Clinical Trials Group. Cancer 92:595–600; 2001.
18. Vogelzang, N. J.; Rusthoven, J. J.; Symanowski, J.; Denham, C.; Kaukel, E.; Ruffie, P.; Gatzemeier, U.; Boyer, M.; Emri, S.; Manegold, C.; Niyikiza, C.; Paoletti, P. Phase III study of pemetrexed in combination with cisplatin versus cisplatin alone in patients with malignant pleural mesothelioma. J. Clin. Oncol. 21:2636–2644; 2003.
19. Scagliotti, G. V.; Kortsik, C.; Dark, G. G.; Price, A.; Manegold, C.; Rosell, R.; O'Brien, M.; Peterson, P. M.; Castellano, D.; Selvaggi, G.; Novello, S.; Blatter, J.; Kayitalire, L.; Crino, L.; Paz-Ares, L.; Go, R. S. Pemetrexed combined with oxaliplatin or carboplatin as first-line treatment in advanced non-small cell lung cancer: A multicenter, randomized, phase II trial. Clin. Cancer Res. 11(2 Pt. 1):690–696; 2005.
20. Adjei, A. A. Review of the comparative pharmacology and clinical activity of cisplatin and carboplatin. J. Clin. Oncol. 17:409–422; 1999.
21. Jackel, M.; Kopf-Maier, P. Influence of cisplatin on cell-cycle progression in xenografted human head and neck carcinomas. Cancer Chemother. Pharmacol. 27:464–471; 1991.
22. Tonkinson, J. L.; Marder, P.; Andis, S. L.; Schultz, R. M.; Gossett, L. S.; Shih, C.; Mendelsohn, L. G. Cell cycle effects of antifolate antimetabolites: Implications for cytotoxicity and cytostasis. Cancer Chemother. Pharmacol. 39: 521–531; 1997.
23. Tonkinson, J. L.; Worzalla, J. F.; Teng, C. H.; Mendelsohn, L. G. Cell cycle modulation by a multitargeted antifolate, LY231514, increases the cytotoxicity and antitumor activity of gemcitabine in HT29 colon carcinoma. Cancer Res. 59:3671–3676; 1999.
24. Schultz, R. M.; Dempsey, J. A. Sequence dependence of Alimta (LY231514, MTA) combined with doxorubicin in ZR-75-1 human breast carcinoma cells. Anticancer Res. 21:3209–3214; 2001.
25. Kano, Y.; Akutsu, M.; Tsunoda, S.; Izumi, T.; Mori, K.; Fujii, H.; Yazawa, Y.; Mano, H.; Furukawa, Y. Schedule-dependent synergism and antagonism between pemetrexed and paclitaxel in human carcinoma cell lines in vitro. Cancer Chemother. Pharmacol. 54:505–513; 2004.
26. Teicher, B. A.; Alvarez, E.; Liu, P.; Lu, K.; Menon, K.; Dempsey, J.; Schultz, R. M. MTA (LY231514) in combination treatment regimens using human tumor xenografts and the EMT-6 murine mammary carcinoma. Semin. Oncol. 28:55–62; 1999.
27. Teicher, B. A.; Chen, V.; Shih, C.; Menon, K.; Forler, P. A.; Phares, V. G.; Amsrud, T. Treatment regimens including themultitargeted antifolate LY231514 in human tumor xenografts. Clin. Cancer Res. 6:1016–1023; 2000.
28. Kano, Y.; Sakamoto, S.; Kasahara, T.; Akutsu, M.; Inoue, Y.; Miura, Y. In vitro effects of amsacrine in combination with other anticancer agents. Leukemia Res. 15:1059–1064; 1991.
29. Steel, G. G.; Peckham, M. J. Exploitable mechanisms in combined radiotherapy-chemotherapy: the concept of additivity. Int. J. Radiat. Oncol. Biol. Phys. 5:85–91; 1979.
30. Kano, Y.; Ohnuma, T.; Okano, T.; Holland, J. F. Effects of vincristine in combination with methotrexate and other antitumor agents in human acute lymphoblastic leukemia cells in culture. Cancer Res. 48:351–356; 1988.
31. Kano, Y.; Akutsu, M.; Tsunoda, S.; Mano, H.; Sato, Y.; Honma, Y.; Furukawa, Y. In vitro cytotoxic effects of a tyrosine kinase inhibitor STI571 in combination with commonly used antileukemic agents. Blood 97:1999–2007; 2001.
32. Kano, Y.; Akutsu, M.; Tsunoda, S.; Suzuki, K.; Adachi, K. In vitro schedule-dependent interaction between paclitaxel and SN-38 (the active metabolite of irinotecan) in human carcinoma cell lines. Cancer Chemother. Pharmacol. 42:91–98; 1998.
33. Furukawa, Y.; Iwase, S.; Kikuchi, J.; Nakamura, M.; Terui, Y.; Yamada, H.; Kano, Y.; Matsuda, M. Phosphorylation of bcl-2 protein by cdc2 Kinase during G2/M phases and its role in cell cycle regulation. J. Biol. Chem. 275:21661–21667; 2000.

Leukemia & Lymphoma, January 2006; 47(1): 9–14

Taylor & Francis
Taylor & Francis Group

REVIEW

DNA micro-array analysis of myelodysplastic syndrome

HIROYUKI MANO

Division of Functional Genomics, Jichi Medical School, Kawachigun, Tochigi, Japan

Abstract
Myelodysplastic syndrome (MDS) is an enigmatic disorder characterized by ineffective hematopoiesis and dysplastic morphology of blood cells. The clinical course of MDS consists of distinct stages, with early stages often progressing to advanced ones or to acute myeloid leukemia (AML). Little is known of the molecular pathogenesis of MDS or of the mechanism of its stage progression. DNA micro-array analysis, which allows simultaneous monitoring of the expression levels of tens of thousands of genes, has the potential to provide insight into the pathophysiology of MDS. Several studies have applied this new technology to compare gene expression profiles either between MDS and the healthy condition, among the different stages of MDS or between MDS-derived AML and *de novo* AML. Selection of an appropriate hematopoietic fraction is important for such studies, which to date have been performed with differentiated granulocytes, $CD34^+$ progenitors and $CD133^+$ immature cells. These studies have revealed that each stage of MDS has its own 'molecular signature', indicating the feasibility of differential diagnosis of MDS based on gene expression profile. They have also demonstrated that the current clinical diagnosis of MDS results in the misclassification of patients with regard to these molecular signatures.

Keywords: *Myelodysplastic syndrome, DNA micro-array, acute myeloid leukemia, stage progression, gene expression profile*

Introduction

Myelodysplastic syndrome (MDS) is an enigmatic disorder that is characterized by 2 clinical manifestations: ineffective hematopoiesis (cytopenia in peripheral blood despite hyper- or normal cellularity in bone marrow) and dysplastic morphology of blood cells [1]. MDS mostly affects the elderly, with an incidence of 15–50 cases per 100 000 people per year [2]. Clonality in multiple lineages of blood cells is found in individuals with MDS, suggesting that MDS is a clonal disorder of multi-potent stem cells in bone marrow [3].

An important aspect of MDS is that it comprises different clinical stages. According to the World Health Organization (WHO) classification of MDS [4], affected individuals whose bone marrow contains $< 5\%$ blasts are diagnosed with refractory anemia (RA), RA with ringed sideroblasts (RARS), refractory cytopenia with multi-lineage dysplasia (RCMD) or refractory cytopenia with multi-lineage dysplasia and ringed sideroblasts (RCMD-RS), whereas those whose bone marrow contains 5–9% or 10–19% blasts are diagnosed with RA with excess blasts (RAEB)-1 or RAEB-2, respectively. About 10–30% of MDS patients at the early stages (RA, RARS, RCMD or RCMD-RS) will eventually undergo stage progression to RAEB and, subsequently, to acute myeloid leukemia (AML).

Despite the relatively high incidence of MDS, its molecular pathogenesis is poorly understood (Figure 1). Gene mutations or other genomic alterations that might give rise to RA or RCMD remain to be identified and the ineffective hematopoiesis apparent in MDS patients remains to be characterized at the molecular biological level. It is also not known what triggers progression of early stages of MDS to advanced ones in some individuals but not others.

Cytopenia in peripheral blood is also found in patients with aplastic anemia (AA). Although the bone marrow of most individuals with AA is characterized by hypocellularity, the difference in marrow cellularity between patients with AA and those with RA or RCMD is not always clear. Antithymocyte globulin, a standard treatment for AA, is also effective in a sub-set of patients at the early

Correspondence: Hiroyuki Mano, MD, PhD, Division of Functional Genomics, Jichi Medical School, 3311-1 Yakushiji, Kawachigun, Tochigi 329-0498, Japan. Tel: +81-285-58-7449. Fax: +81-285-44-7322. E-mail: hmano@jichi.ac.jp

Received for publication 15 July 2005.

ISSN 1042-8194 print/ISSN 1029-2403 online © 2006 Taylor & Francis
DOI: 10.1080/10428190500264231

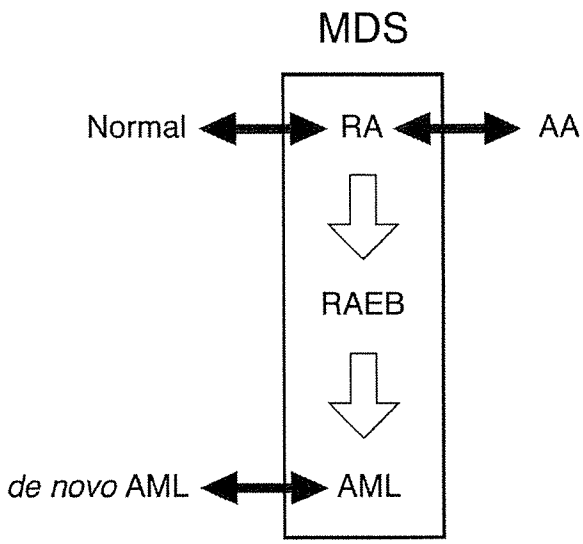

Figure 1. Stage progression of MDS and its relation to other conditions. Little is currently known of the molecular pathogenesis of MDS or of the similarities or differences between the early stages of MDS and aplastic anemia (AA) and between MDS-derived AML and *de novo* AML.

stages of MDS [5], further complicating the distinction between the two disorders.

It is widely believed that AML evolved from MDS has a poorer prognosis than does *de novo* AML, suggesting that the 2 clinical entities might be distinct. However, patients with MDS-associated AML are often older than are those with *de novo* AML and tend to possess karyotypes associated with high risk. It, thus, remains unclear whether the prognosis of *de novo* AML does indeed differ from that of MDS-associated AML if patient age and karyotype are matched.

The Human Genome Project is now close to completion, with 99% of DNA in euchromatin having been sequenced at 99.999% accuracy (http://www.ncbi.nlm.nih.gov/genome/guide/human/). Annotation of the human genome has revealed an unexpectedly small number (20 000 – 25 000) of protein-coding genes [6] compared with the numbers identified in the nematode (~19 000 genes) [7] and fruit fly (~14 000 genes) genomes [8]. The development of DNA micro-array analysis now allows simultaneous measurement of the level of expression of tens of thousands of genes in a given sample [9,10]. With this approach, it is thus possible to obtain a total gene expression profile or 'transcriptome' for each of the various stages of MDS and then to compare these profiles either among MDS stages or with those of the healthy condition, AA or *de novo* AML. Such analysis has the potential both to identify novel molecular markers for the differential diagnosis of MDS vs AA or *de novo* AML as well as to reveal genes that contribute to the pathogenesis of MDS. It might also be possible to determine whether MDS should be treated as a clinical entity distinct from AA or *de novo* AML.

It is important to bear in mind, however, that blood cells of different lineages and differentiation levels possess markedly different transcriptomes, even within the same individual. Any shift in cell composition in the specimens analysed will, thus, greatly influence the gene expression profile determined by micro-array studies [11].

Normal vs MDS

Several studies have attempted to compare transcriptomes between healthy individuals and patients with MDS in order to identify differences in gene expression. Pellagatti et al. [12] isolated RNA from differentiated granulocytes of 7 healthy individuals and 21 patients with MDS (17 with RA, 2 with RARS and 2 with RAEB according to the French-American-British (FAB) classification). The RNA was subjected to hybridization with a cDNA micro-array harboring probes corresponding to ~6000 human genes and the researchers identified 12 genes whose expression was frequently up-regulated (ratio of >2.0 in $\geqslant 9$ patients) or down-regulated (ratio of <0.5 in $\geqslant 10$ patients) in MDS. The relevance of these genes to the molecular diagnosis of MDS is unknown.

In contrast, Hofmann et al. [13] isolated $CD34^+$ progenitor fractions from the bone marrow of 4 healthy subjects and 11 MDS patients (7 low-risk and 4 high-risk according to the International Prognostic Scoring System [14]) for analysis with micro-arrays containing >12 000 human probe sets. They identified 161 genes whose expression was down-regulated (ratio of <0.2) in low-risk MDS patients compared with healthy subjects. They also detected 117 genes whose expression was up-regulated (ratio of >5) in MDS patients and 27 of these genes encoded regulators of hematopoiesis, including acute myeloid leukemia 1 (AML1), activating transcriptional factor 3 (ATF3), homeobox 7 (HOX7) and Delta-like homolog 1 (DLK1).

Chen et al. [15] also chose $CD34^+$ cells for comparison of transcriptomes between healthy controls and MDS patients, specifically those with monosomy 7 or trisomy 8. $CD34^+$ progenitor cells were purified from the bone marrow of 4 control subjects, 4 MDS patients with trisomy 8 and 2 MDS patients with monosomy 7 and RNA isolated from these cells was subjected to hybridization with the same type of arrays (Affymetrix GeneChip HGU95Av2) as those used by Hofmann et al. [13]. Comparison of gene expression profiles among the subjects revealed that genes important in immune function and inflammation were frequently over-expressed in the MDS patients

with trisomy 8, consistent with previous findings implicating autoimmune activity in such patients [16]. In contrast, genes important in cell growth were often down-regulated in the patients with monosomy 7. These findings, thus, suggested that gene expression profiles differ between MDS blasts with trisomy 8 and those with monosomy 7.

Stage progression in MDS

A substantial proportion of MDS patients in early stages of the disease, especially those with an unfavorable karyotype [17], undergoes progression to advanced stages or to AML. Given that currently available chemotherapeutic regimens for advanced MDS are of limited efficacy, it would be clinically advantageous to block stage progression in MDS. Alterations of several oncogenes and tumor suppressor genes have been implicated in the progression of MDS [18]. Activating mutations of *RAS* genes are thought to be the most prevalent (affecting 10 – 30% of cases) of such changes in MDS [19,20]. It remains unclear, however, whether *RAS* mutation occurs at the early or late stages of MDS. Inactivation of the p53 gene is also apparent in 5 – 10% of MDS patients [21]. Again, however, it is not known whether loss of p53 function is an early or late event during MDS progression. In addition, epigenetic silencing of the p15 gene and shortening of telomeres have been detected in bone marrow cells of MDS patients [22]. None of these gene alterations is specific to MDS and it is unclear which changes are the cause of MDS itself and which are associated with stage progression.

To obtain insight into the mechanism of stage progression of MDS, in their comparison of gene expression profiles among healthy controls, low-risk MDS patients and high-risk MDS patients, Hofmann et al. [13] applied the class membership prediction method to identify genes whose expression was linked to separation of the 3 classes. They identified 11 such genes (Table I) and a simple 2-way clustering analysis of the study subjects based on the expression patterns of these 11 genes clearly separated the 3 major classes. Furthermore, a similar clustering analysis of a second set of subjects ($n = 8$) also separated individuals with high-risk MDS from those with low-risk MDS. Although the number of study subjects was small, these data support the notion that each stage of MDS has a characteristic gene expression profile or 'molecular signature'.

Table I. Genes used for class separation of healthy individuals and low- or high-risk patients with MDS [13].

Gene symbol	Accession number	Chromosomal position
TACSTD2	J04152	1p32
UQCRC1	L16842	3p21.3
TNNC1	M37984	3p21.3
KDELR	M88458	7p
CLC	L01664	19q13.1
H-PKL	M5422	7
RGS19	Z91809	
ATF3	L19871	1
FARP1	AI701049	
GNG7	AW051450	
TPD52L2	U44429	6q22-q23

CD133 (also known as AC133) is a cell surface protein that is expressed exclusively on $CD34^+$ $CD38^-$ hematopoietic stem cells (HSCs) [23,24]. Many AML blasts also express CD133 [25], indicating that the differentiation of these cells is blocked at a highly immature stage. The existence of 'cancer stem cells' for AML and solid tumors has been recently demonstrated and such cells express CD133 in brain tumors [26], as do $CD34^+CD38^-$ cells in AML [27]. Analysis of $CD133^+$ HSC-like fractions among MDS patients may, thus, reveal the character of 'MDS stem cells'. Analysis of such fractions also has the advantage of eliminating from micro-array data the influence of variation in cell composition of bone marrow, which is especially important given that different stages of MDS are characterized by different numbers of immature blasts within marrow.

Ueda et al. [28] performed micro-array analysis with $CD133^+$ cells isolated from the bone marrow of 2 healthy individuals, 11 patients with RA, 5 patients with RAEB and 14 patients with MDS-associated AML. Comparison of the gene expression profiles among the different stages of MDS led to the identification of 11 late stage (RAEB, MDS-associated AML)-specific genes and 6 early stage (healthy controls, RA)-specific genes. The latter set of genes included that for PIASy, which catalyses sumoylation of substrate proteins [29]. Loss of expression of the PIASy gene in advanced MDS suggested that the encoded protein might possess anti-tumor activity. Consistent with this notion, forced expression of PIASy in a mouse myeloid cell line resulted in rapid induction of apoptosis when the cells were cultured in the presence of granulocyte colony-stimulating factor (G-CSF) (Figure 2) [28]. These results suggest that PIASy functions to restrain cell growth and that loss of its expression may facilitate stage progression in MDS. Loss of PIASy expression has also been implicated in stage progression of chronic myeloid leukemia [30].

MDS-derived AML vs *de novo* AML

Although dysplastic morphology of blood cells is a hallmark of MDS and MDS-derived AML, such

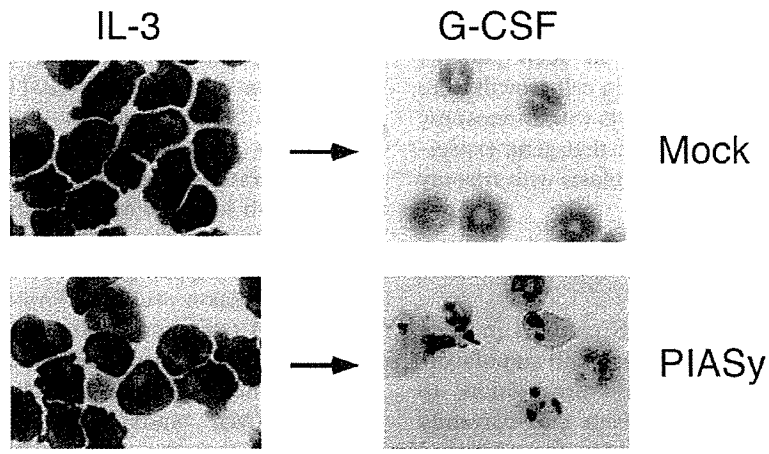

Figure 2. Induction of apoptosis by PIASy. 32Dcl3 cells were infected with a control retrovirus (Mock) or with a virus that encodes human PIASy and were then either maintained in the presence of interleukin-3 (IL-3) or exposed to G-CSF. Whereas control cells incubated with G-CSF underwent gradual differentiation into granulocyte-like cells, those expressing PIASy died rapidly by apoptosis on exposure to G-CSF. Reproduced with permission from Ueda et al. [28].

dysplasia is also apparent in the blood cells of some healthy older individuals. It is, thus, sometimes difficult to differentiate *de novo* AML from MDS-derived AML in the elderly, especially in the absence of a clinical history of the patient. Further complicating the issue, some younger patients with *de novo* AML also manifest blood cell dysplasia [31,32].

To provide insight into the differences or similarities between these clinical classes, Oshima et al. purified $CD133^+$ cells from the bone marrow both of 10 patients with *de novo* AML of the M2 sub-type, according to the FAB classification [33] and of 10 patients with MDS-derived AML corresponding to the M2 sub-type [34]. They then subjected RNA from these cells to micro-array analysis with HGU95Av2 micro-arrays. Comparison of samples matched for FAB sub-type was performed to minimize the influence of the differentiation ability of the blasts on gene expression profile; any differences in gene expression identified with this approach would, thus, be expected to be related with a high probability to the difference in the nature of MDS-derived AML from that of *de novo* AML.

Statistical analysis of the resulting expression data (Welch's ANOVA, $p < 0.01$; effect size of $\geqslant 5.0$ units) identified a total of 57 probe sets corresponding to genes whose expression was associated with diagnosis. However, a simple 2-way clustering analysis of the subjects based on the expression pattern for these probe sets failed to separate them into diagnosis-related sub-groups. To visualize the similarity or difference between the 2 classes, the researchers applied correspondence analysis, a method for the decomposition of multi-dimensional data [35]. This approach allows not only a low-dimensional projection of the expression profiles of numerous genes but also measurement both of the contribution of each gene to a given extracted dimension and of the contribution of each extracted dimension to the total complexity.

Correspondence analysis of the expression data for the 57 probe sets reduced the complexity from 57 to 3 dimensions. The specimens were then projected into a virtual space on the basis of their calculated 3-dimensional (3D) co-ordinates (Figure 3(*a*)). Most subjects with *de novo* AML were localized in a region of the space distinct from that occupied by those with MDS-derived AML. However, 2 individuals with *de novo* AML localized with those with MDS-derived AML. These observations indicated that the transcriptome of MDS-derived AML is distinct from that of *de novo* AML, but that current clinical diagnosis does not completely correlate with the difference in transcriptomes.

A similar analysis was performed on a larger scale by Tsutsumi et al. [36]. These researchers isolated $CD133^+$ cells from the bone marrow of patients with MDS-derived AML ($n = 11$), with *de novo* AML without dysplasia ($n = 15$), with *de novo* AML with multi-lineage dysplasia ($n = 11$) [32,37] or with therapy-related AML ($n = 2$). The study subjects were not limited to a specific FAB sub-type, however. The transcriptomes of these clinical classes were compared with the use of HGU95Av2 arrays. Comparison of *de novo* AML without dysplasia and MDS-derived AML led to the identification of 30 probe sets corresponding to genes whose expression was related to diagnosis. Correspondence analysis

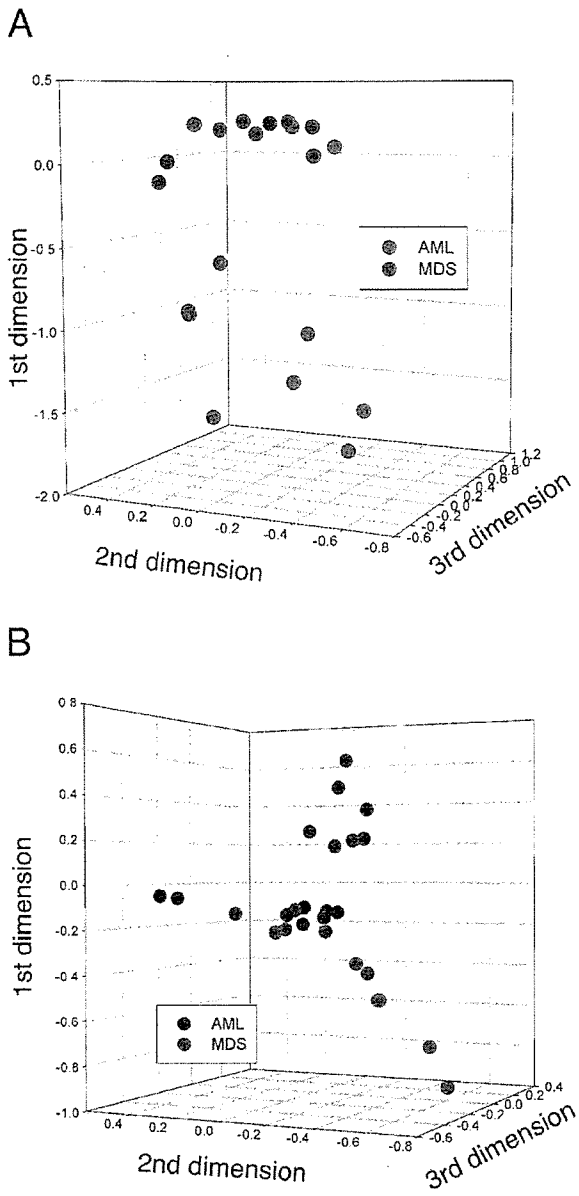

Figure 3. Analysis of the difference between the transcriptomes of MDS-related AML and *de novo* AML by 3D projection of study subjects. (*a*) Patients with *de novo* AML of the M2 sub-type (green) and those with MDS-derived AML of the M2 sub-type (red) were projected into a virtual space on the basis of co-ordinates calculated by correspondence analysis from the expression profiles of 57 probe sets shown in Oshima et al. [34]. (*b*) Patients with *de novo* AML without dysplasia (blue) and those with MDS-derived AML (red) were projected as in (*a*) on the basis of co-ordinates calculated from the expression profiles of 30 probe sets. Modified with permission from Tsutsumi et al. [36].

and 3D projection confirmed the distinct, but partially overlapping, gene expression profiles for the individuals diagnosed clinically with MDS-derived AML and those diagnosed with *de novo* AML without dysplasia (Figure 3(*b*)).

Future directions

DNA micro-array analysis has provided new insight into MDS. Such analysis allows comparison of transcriptomes between the healthy condition and MDS, among the distinct stages of MDS and between MDS-derived AML and other types of AML. The results of such studies are likely to lead to the development of accurate means of differential diagnosis as well as to the identification of genes that are important determinants of the various characteristics of MDS. DNA micro-array analysis and other genomics approaches also may clarify whether MDS is a single clinical entity as well as its relations to other clonal disorder of HSCs.

As of now, however, our knowledge of MDS provided by DNA micro-array analysis is relatively limited and based on preliminary data. Studies with much larger numbers of subjects than those performed to date are required. It remains to be determined both how many sub-groups of MDS can be defined on the basis of gene expression profiles and how these sub-groups are related to clinical characteristics, especially to long-term prognosis of patients and to the chemosensitivity of blasts. Given that the bone marrow of individuals at distinct stages of MDS contains different proportions of immature blasts, fractionation of bone marrow (for isolation of $CD34^+$ or $CD133^+$ cells, for example) before micro-array analysis is desirable. The combination of DNA micro-array analysis and other genomics tools should then shed new light on this enigmatic and intractable disorder.

References

1. Heaney ML, Golde DW. Myelodysplasia. New England Journal of Medicine 1999;340:1649–1660.
2. Aul C, Giagounidis A, Germing U. Epidemiological features of myelodysplastic syndromes: results from regional cancer surveys and hospital-based statistics. International Journal of Hematology 2001;73:405–410.
3. Boultwood J, Wainscoat JS. Clonality in the myelodysplastic syndromes. International Journal of Hematology 2001;73: 411–415.
4. Harris NL, Jaffe ES, Diebold J, Flandrin G, Muller-Hermelink HK, Vardiman J, et al. World Health Organization classification of neoplastic diseases of the hematopoietic and lymphoid tissues: report of the Clinical Advisory Committee meeting—Airlie House, Virginia, November 1997. Journal of Clinical Oncology 1999;17: 3835–3849.
5. Molldrem JJ, Leifer E, Bahceci E, Saunthararajah Y, Rivera M, Dunbar C, et al. Antithymocyte globulin for treatment of the bone marrow failure associated with myelodysplastic syndromes. Annals of Internal Medicine 2002;137: 156–163.
6. International Human Genome Sequencing Consortium. Finishing the euchromatic sequence of the human genome. Nature 2004;431:931–945.

7. The C. elegans Sequencing Consortium. Genome sequence of the nematode *C. elegans*: a platform for investigating biology. Science 1998;282:2012–2018.
8. Adams MD, Celniker SE, Holt RA, Evans CA, Gocayne JD, Amanatides PG, et al. The genome sequence of *Drosophila melanogaster*. Science 2000;287:2185–2195.
9. Duggan DJ, Bittner M, Chen Y, Meltzer P, Trent JM. Expression profiling using cDNA microarrays. Natural Genetics 1999;21:10–14.
10. Cheung VG, Morley M, Aguilar F, Massimi A, Kucherlapati R, Childs G. Making and reading microarrays. Natural Genetics 1999;21:15–19.
11. Miyazato A, Ueno S, Ohmine K, Ueda M, Yoshida K, Yamashita Y, et al. Identification of myelodysplastic syndrome-specific genes by DNA microarray analysis with purified hematopoietic stem cell fraction. Blood 2001;98: 422–427.
12. Pellagatti A, Esoof N, Watkins F, Langford CF, Vetrie D, Campbell LJ, et al. Gene expression profiling in the myelodysplastic syndromes using cDNA microarray technology. British Journal of Haematology 2004;125:576–583.
13. Hofmann WK, de Vos S, Komor M, Hoelzer D, Wachsman W, Koeffler HP. Characterization of gene expression of $CD34^+$ cells from normal and myelodysplastic bone marrow. Blood 2002;100:3553–3560.
14. Greenberg P, Cox C, LeBeau MM, Fenaux P, Morel P, Sanz G, et al. International scoring system for evaluating prognosis in myelodysplastic syndromes. Blood 1997;89: 2079–2088.
15. Chen G, Zeng W, Miyazato A, Billings E, Maciejewski JP, Kajigaya S, et al. Distinctive gene expression profiles of CD34 cells from patients with myelodysplastic syndrome characterized by specific chromosomal abnormalities. Blood 2004;104: 4210–4218.
16. Sloand EM, Kim S, Fuhrer M, Risitano AM, Nakamura R, Maciejewski JP, et al. Fas-mediated apoptosis is important in regulating cell replication and death in trisomy 8 hematopoietic cells but not in cells with other cytogenetic abnormalities. Blood 2002;100:4427–4432.
17. Nösslinger T, Reisner R, Koller E, Gruner H, Tuchler H, Nowotny H, et al. Myelodysplastic syndromes, from French-American-British to World Health Organization: comparison of classifications on 431 unselected patients from a single institution. Blood 2001;98:2935–2941.
18. Fenaux P. Chromosome and molecular abnormalities in myelodysplastic syndromes. International Journal of Hematology 2001;73:429–437.
19. Mano H, Ishikawa F, Hirai H, Takaku F. Mutations of N-ras oncogene in myelodysplastic syndromes and leukemias detected by polymerase chain reaction. Japanese Journal of Cancer Research 1989;80:102–106.
20. Paquette RL, Landaw EM, Pierre RV, Kahan J, Lubbert M, Lazcano O, et al. N-ras mutations are associated with poor prognosis and increased risk of leukemia in myelodysplastic syndrome. Blood 1993;82:590–599.
21. Jonveaux P, Fenaux P, Quiquandon I, Pignon JM, Lai JL, Loucheux-Lefebvre MH, et al. Mutations in the p53 gene in myelodysplastic syndromes. Oncogene 1991;6:2243–2247.
22. Tien HF, Tang JH, Tsay W, Liu MC, Lee FY, Wang CH, et al. Methylation of the p15(INK4B) gene in myelodysplastic syndrome: it can be detected early at diagnosis or during disease progression and is highly associated with leukaemic transformation. British Journal of Haematology 2001;112: 148–154.
23. Hin AH, Miraglia S, Zanjani ED, Almeida-Porada G, Ogawa M, Leary AG, et al. AC133, a novel marker for human hematopoietic stem and progenitor cells. Blood 1997; 90:5002–5012.
24. Kratz-Albers K, Zuhlsdorp M, Leo R, Berdel WL, Buchner T, Serve H. Expression of AC133, a novel stem cell marker, on human leukemic blasts lacking CD34-antigen and on a human $CD34^+$ leukemic line: MUTZ-2. Blood 1998;92: 4485–4487.
25. Bühring H-J, Seiffert M, Marxer A, Weiß B, Faul C, Kanz L, Brugger W. AC133 antigen expression is not restricted to acute myeloid leukemia blasts but is also found on acute lymphoid leukemia blasts and on a subset of $CD34^+$ B-cell progenitor. Blood 1999;94:832–833.
26. Singh SK, Hawkins C, Clarke ID, Squire JA, Bayani J, Hide T, et al. Identification of human brain tumour initiating cells. Nature 2004;432:396–401.
27. Bonnet D, Dick JE. Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell. Natural Medicine 1997;3:730–737.
28. Ueda M, Ota J, Yamashita Y, Choi YL, Ohki R, Wada T, et al. DNA microarray analysis of stage progression mechanism in myelodysplastic syndrome. British Journal of Haematology 2003;123:288–296.
29. Sachdev S, Bruhn L, Sieber H, Pichler A, Melchior F, Grosschedl R. PIASy, a nuclear matrix-associated SUMO E3 ligase, represses LEF1 activity by sequestration into nuclear bodies. Genes Development 2001;15:3088–3103.
30. Ohmine K, Ota J, Ueda M, Ueno S-I, Yoshida K, Yamashita Y, et al. Characterization of stage progression in chronic myeloid leukemia by DNA microarray with purified hematopoietic stem cells. Oncogene 2001;20:8249–8257.
31. Brito-Babapulle F, Catovsky D, Galton DAG. Clinical and laboratory features of *de novo* acute myeloid leukemia with trilineage myelodysplasia. British Journal of Haematology 1987;66:445–450.
32. Taguchi J, Miyazaki Y, Yoshida S, Fukushima T, Moriuchi Y, Jinnai I, et al. Allogeneic bone marrow transplantation improves the outcome of de novo AML with trilineage dysplasia (AML-TLD). Leukemia 2000;14:1861–1866.
33. Bennett JM, Catovsky D, Daniel MT, Flandrin G, Galton DA, Gralnick HR, Sultan C. Proposed revised criteria for the classification of acute myeloid leukemia. A report of the French-American-British Cooperative Group. Annals of Internal Medicine 1985;103:620–625.
34. Oshima Y, Ueda M, Yamashita Y, Choi YL, Ota J, Ueno S, et al. DNA microarray analysis of hematopoietic stem cell-like fractions from individuals with the M2 subtype of acute myeloid leukemia. Leukemia 2003;17:1990–1997.
35. Fellenberg K, Hauser NC, Brors B, Neutzner A, Hoheisel JD, Vingron M. Correspondence analysis applied to microarray data. Proceedings of the National Academy of Sciences (USA) 2001;98:10781–1076.
36. Tsutsumi C, Ueda M, Miyazaki Y, Yamashita Y, Choi YL, Ota J, et al. DNA microarray analysis of dysplastic morphology associated with acute myeloid leukemia. Experimental Hematology 2004;32:828–835.
37. Kuriyama K, Tomonaga M, Matsuo T, Kobayashi T, Miwa H, Shirakawa S, et al. Poor response to intensive chemotherapy in *de novo* acute myeloid leukaemia with trilineage myelodysplasia. Japan Adult Leukaemia Study Group (JALSG). British Journal of Haematology 1994;86:767–773.

Experimental trial for diagnosis of pancreatic ductal carcinoma based on gene expression profiles of pancreatic ductal cells

Madoka Ishikawa,[1] Koji Yoshida,[2] Yoshihiro Yamashita,[1] Jun Ota,[1,3] Shuji Takada,[1] Hiroyuki Kisanuki,[1] Koji Koinuma,[1] Young Lim Choi,[1] Ruri Kaneda,[1] Toshiyasu Iwao,[4] Kiichi Tamada,[5] Kentaro Sugano[5] and Hiroyuki Mano[1,3,6]

[1]Division of Functional Genomics, Jichi Medical School, 3311-1 Yakushiji, Kawachi-gun, Tochigi 329-0498; [2]Department of Medicine, Kawasaki Medical School, Okayama 701-0192; [3]CREST, Japan Science and Technology Agency, 4-1-8, Honcho, Kawaguchi-shi, Saitama 332-0012; [4]Gastroenterological Center, Aizu Central Hospital, 1-1 Tsuruga-machi, Aizu-Wakamatsu, Fukushima 956-8611; and [5]Division of Gastroenterology, Jichi Medical School, 3311-1 Yakushiji, Kawachi-machi, Kawachi, Tochigi 329-0498, Japan

(Received November 18, 2004/Revised April 25, 2005/Accepted April 27, 2005/Online Publication July 22, 2005)

Pancreatic ductal carcinoma (PDC) remains one of the most intractable human malignancies, mainly because of the lack of sensitive detection methods. Although gene expression profiling by DNA microarray analysis is a promising tool for the development of such detection systems, a simple comparison of pancreatic tissues may yield misleading data that reflect only differences in cellular composition. To directly compare PDC cells with normal pancreatic ductal cells, we purified MUC1-positive epithelial cells from the pancreatic juices of 25 individuals with a normal pancreas and 24 patients with PDC. The gene expression profiles of these 49 specimens were determined with DNA microarrays containing >44 000 probe sets. Application of both Welch's analysis of variance and effect size-based selection to the expression data resulted in the identification of 21 probe sets corresponding to 20 genes whose expression was highly associated with clinical diagnosis. Furthermore, correspondence analysis and 3-D projection with these probe sets resulted in separation of the transcriptomes of pancreatic ductal cells into distinct but overlapping spaces corresponding to the two clinical classes. To establish an accurate transcriptome-based diagnosis system for PDC, we applied supervised class prediction algorithms to our large data set. With the expression profiles of only five predictor genes, the weighted vote method diagnosed the class of samples with an accuracy of 81.6%. Microarray analysis with purified pancreatic ductal cells has thus provided a basis for the development of a sensitive method for the detection of PDC. (*Cancer Sci* 2005; 96: 387–393)

Pancreatic ductal carcinoma (PDC), arising from the pancreatic ductal cells, accounts for more than 85% of all pancreatic malignancies, and is one of the most intractable malignancies in humans.[1,2] Effective therapy for PDC is hampered by the lack of specific clinical symptoms, with a 5-year survival rate of only 20 to 30%. An increase in the serum concentration of the protein CA19-9 is a reliable marker for PDC, but such an increase is only apparent in the advanced stages of disease.[3] Furthermore, although activating mutations of the *KRAS* oncogene have been detected in PDC cells, such mutations are also associated with other conditions, including chronic pancreatitis.[4,5]

DNA microarray analysis allows the simultaneous monitoring of the expression level of thousands of genes[6,7] and is therefore a potentially suitable approach for the identification of novel molecular markers for detection of the early stages of PDC. However, caution is warranted in simple comparisons between normal and cancerous pancreatic tissues. Because normal pancreatic tissue is composed mostly of exocrine and endocrine cells, and cancerous pancreatic tissue consists mostly of tumor cells that arise from ductal epithelial cells, a simple comparison between these two tissues tends to identify cell lineage-dependent gene expression differences.[8]

To minimize such misleading data that are attributable to population-shift effects, we have set up a depository for pancreatic ductal cells purified from pancreatic juice collected from patients during endoscopic retrograde cholangiopancreatography (ERCP). Comparison of such pancreatic ductal cell preparations between control individuals and PDC patients by DNA microarray analysis has the potential to identify specific gene markers for the latter. Indeed, an initial screening of a limited number of samples (from three individuals with a normal pancreas and six with PDC) with a DNA microarray of 3456 genes yielded candidates for new PDC marker genes.[8]

We have now expanded this project by using a larger number of specimens: 25 from individuals with a normal pancreas and 24 from PDC patients. Each purified preparation of pancreatic ductal cells was subjected to microarray experiments with Affymetrix HGU133 A&B GeneChips, which contain >44 000 probe sets corresponding to ~33 000 human genes. The application of sophisticated bioinformatics techniques to this large data set (a total of 2 156 000 data points) resulted in the establishment of an algorithm to differentiate transformed ductal cells from normal ones.

Materials and Methods

Preparation of pancreatic ductal cells. The study subjects comprised individuals who underwent ERCP and collection of pancreatic juice for cytological examination. The subjects gave informed consent and the study was approved by the institutional review board of Jichi Medical School. Diagnosis of patients was confirmed on the basis of the combination of results obtained by ERCP, cytological examination of pancreatic juice, abdominal computed tomography, and measurement of the serum concentration of CA19-9, as well as of follow-up observations. Approximately one-third of each specimen of pancreatic juice was used to purify $MUC1^+$ ductal cells.[9]

Cells were collected from the pancreatic juice by centrifugation and were resuspended in 1 mL of MACS binding buffer (150 mM NaCl, 20 mM sodium phosphate [pH 7.4], 3% fetal bovine serum, 2 mM ethylenediamine tetraacetic acid). They were then incubated for 30 min at 4°C with 0.5 μg of a mouse

[6]To whom correspondence should be addressed. E-mail: hmano@jichi.ac.jp
Abbreviations: ACTB, β-actin; EPPK1, epiplakin 1; ERCP, endoscopic retrograde cholangiopancreatography; H2BFB, H2B histone family, member B; KNN, *k* nearest neighbor; NRCAM, neuronal cell adhesion molecule; PCR, polymerase chain reaction; PDC, pancreatic ductal carcinoma; PLOD2, procollagen-lysine, 2-oxoglutarate 5-dioxygenase 2; RASAL2, RAS protein activator-like 2; SCGB3A1, secretoglobin, family 3A, member 1; SST, somatostatin; WV, weighted vote.

 doi:10.1111/j.1349-7006.2005.00064.x

monoclonal antibody to MUC1 (Novocastra Laboratories, Newcastle upon Tyne, UK), washed with MACS binding buffer, and mixed with MACS MicroBeads conjugated with antibodies to mouse immunoglobulin G (Miltenyi Biotec, Auburn, CA, USA). The resulting mixture was subjected to chromatography on a miniMACS magnetic cell separation column (Miltenyi Biotec), and the eluted MUC1$^+$ cells were divided into portions and stored at −80°C. Portions of the unfractionated cells as well as the isolated MUC1$^+$ cells of each individual were stained with Wright–Giemsa solution to examine the purity of the ductal cell-enriched fractions.

Microarray experiments. Total RNA was extracted from the MUC1$^+$ cell preparations with the use of an RNeasy Mini column and RNase-free DNase (Qiagen, Valencia, CA, USA) and was subjected to two rounds of mRNA amplification with T7 RNA polymerase.[10] The high fidelity of the amplification step has been demonstrated previously.[11] One microgram of the amplified cRNA was then converted to double-stranded cDNA by PowerScript reverse transcriptase (BD Biosciences Clontech, Palo Alto, CA, USA), and the cDNA was used to prepare biotin-labeled cRNA with an ENZO BioArray Transcript Labeling Kit (Affymetrix, Santa Clara, CA, USA). Hybridization of the labeled cRNA with GeneChip HGU133 A&B microarrays, which contain >44 000 probe sets, was performed with the GeneChip system (Affymetrix). The mean expression intensity of the internal positive control probe sets[12] was set to 500 arbitrary units (U) in each hybridization, and the fluorescence intensity of each test gene was normalized accordingly. All normalized array data are available at the Gene Expression Omnibus web site (http://www.ncbi.nlm.nih.gov/geo) under the accession number GSE1542.

Statistical analysis. Hierarchical clustering of the data set, Welch's analysis of variance (ANOVA), and k nearest neighbor (KNN) method-based class prediction were performed with GeneSpring 6.2 software (Silicon Genetics, Redwood, CA). The weighted vote (WV) method[13] was performed with GeneCluster 2.1.7.[14] Correspondence analysis[15] for all genes showing a significant difference in expression was performed by using ViSta software.[16] Each sample was plotted in three dimensions based on the coordinates obtained from the correspondence analysis. With the exception of the effect-size selection, in which linear values were used for calculation, all normalized expression values were transformed to logarithms prior to analyses.

Real-time PCR analysis. Portions of nonamplified cDNA were subjected to PCR with a QuantiTect SYBR Green PCR Kit (Qiagen). The amplification protocol comprised incubations at 94°C for 15 s, 57°C for 30 s, and 72°C for 60 s. Incorporation of the SYBR Green dye into PCR products was monitored in real time with an ABI PRISM 7900 HT sequence detection system (PE Applied Biosystems, Foster City, CA, USA), thereby allowing determination of the threshold cycle (C_T) at which exponential amplification of products begins. The C_T values for cDNA corresponding to the β-actin gene (*ACTB*) and to the target genes were used to calculate the abundance of target gene mRNA relative to that of *ACTB* mRNA. The oligonucleotide primers for PCR were as follows: 5′-CCATCATGAAGTGTGACGTGG-3′ and 5′-GTCCGCCTAGAAGCATTTGCG-3′ for *ACTB*, 5′-CCCGTGAACCACCTCATAG-3′ and 5′-AGCGTCTTGTCCTCAGGTGTA-3′ for the secretoglobin, family 3A, member 1 gene (*SCGB3A1*), and 5′-GATGAAATGAGGCTTGAGCTG-3′ and 5′-GTTTCTAATGCAAGGGTCTCG-3′ for the somatostatin gene (*SST*).

Results

Transcriptome of pancreatic ductal cells. As demonstrated previously, affinity purification with antibodies to MUC1 yielded an apparently homogeneous preparation of pancreatic ductal cells.[8] With this approach, we purified pancreatic ductal cell specimens from 25 individuals with a normal pancreas and 24 patients with PDC. Clinical characteristics for the latter individuals are summarized in Table 1. All 49 specimens were each subjected to DNA microarray analysis with Affymetrix HGU133 A&B GeneChips, which contain >44 000 probe sets.

Table 1. Clinical characteristics of patients with PDC

Patient	Age (years)	Sex	Cytological examination	Atypical cell proportion*	Clinical stage†
ID073	74	Male	V	H	IVa
ID086	72	Female	IV	M	IVa
ID088	65	Male	V	L	IVb
ID089	70	Female	III	L	III
ID090	72	Male	III	L	IVa
ID095	85	Female	III	L	0
ID096	76	Female	IV	L	IVa
ID098	61	Female	IV	L	IVa
ID103	65	Male	V	H	IVb
ID117	76	Female	IV	L	IVa
ID119	73	Female	V	L	IVa
ID120	70	Female	III	M	0
ID125	75	Male	II	L	I
ID131	67	Female	II	L	IVa
ID142	69	Male	III	H	I
ID147	51	Male	V	L	IVb
ID202	56	Female	III	M	IVa
ID203	73	Male	III	M	I
ID218	51	Male	III	L	0
ID224	71	Male	V	L	IVa
ID225	50	Female	III	L	IVa
ID227	65	Male	I	L	III
ID229	60	Female	IV	M	IVa
ID234	71	Male	III	L	IVa

*Isolated ductal cells contained <20% (L), 20–40% (M) or ≧ 40% (H) of atypical cells. †Clinical stage was determined according to the proposal of Isaji *et al.*[25]

For analysis of the gene expression data, we first set the condition that the expression level of a given probe set should receive the 'Present call' (from Microarray Suite 5.0 software) in at least 30% ($n = 15$) of the samples in order to exclude transcriptionally silent genes from the analysis. A total of 7778 probe sets fulfilled this selection criterion. Unsupervised two-way hierarchical clustering analysis[17] was then applied to the 49 specimens based on the expression profiles of these 7778 probe sets, generating a dendrogram in which the samples are clustered according to the similarity in expression pattern of the probe sets (Fig. 1). Although this dendrogram contained a large branch consisting mostly of PDC patients, normal and cancer specimens did not form separate, diagnosis-dependent branches. The transcriptome of virtually all expressed genes thus did not differ sufficiently between normal and cancerous ductal cells to allow diagnosis.

PDC-specific molecular signature. To capture a PDC-specific molecular signature, we next identified genes whose expression level differed significantly between the normal and cancerous ductal cells. Application of Welch's ANOVA ($P < 0.001$) for this purpose yielded 26 out of the 7778 probe sets examined. However, some of the probe sets thus identified had low absolute expression levels throughout the samples, even though the ratio of the expression levels between the two classes was relatively large. To eliminate such 'nearly silent' genes and to enrich genes whose expression level was markedly increased in at least one of the classes, we further selected those whose effect size (absolute difference in mean expression intensities)[18] between the two classes was ≧ 50 U.

 doi:10.1111/j.1349-7006.2005.00064.x

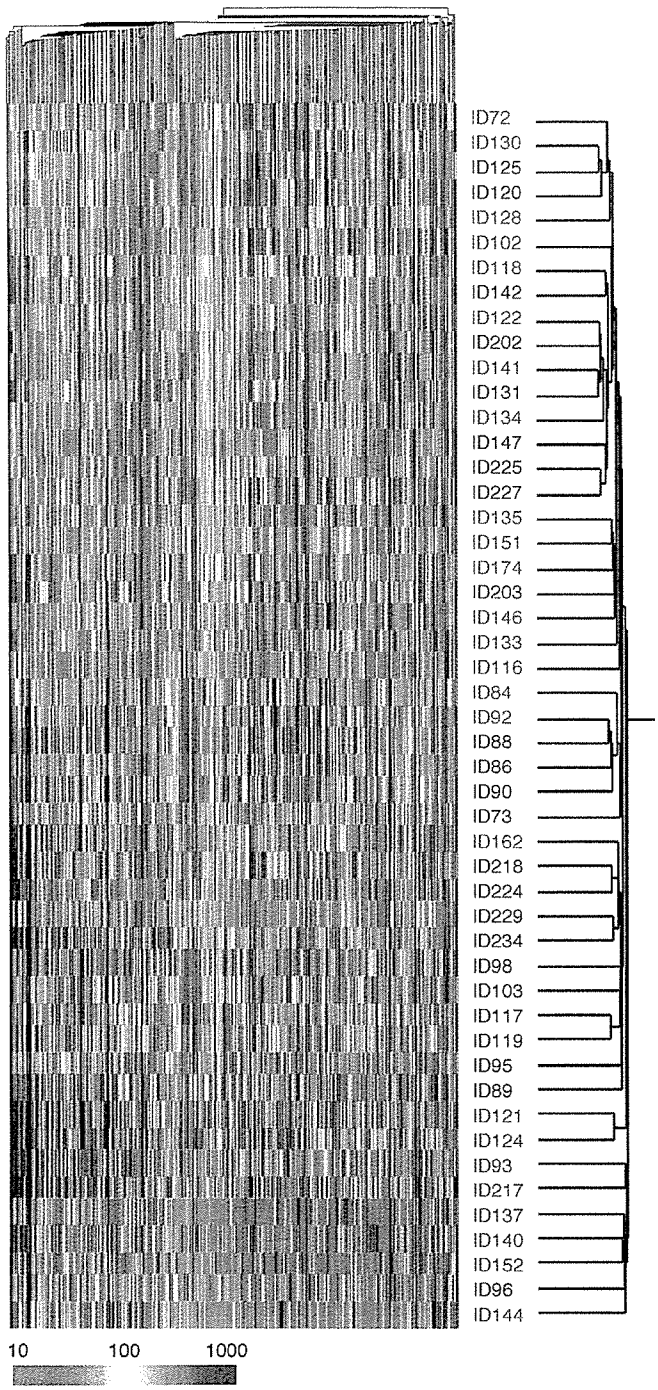

Fig. 1. Gene expression profiles of the purified pancreatic ductal cells. Hierarchical two-way clustering of the study subjects (normal ductal cell specimens [green] and PDC specimens [red]) was performed on the basis of the expression profiles of 7778 probe sets. Each column corresponds to a single probe set, and each row corresponds to a separate subject. The expression level of probe sets is color-coded according to the indicated scale.

With this approach, we identified 21 probe sets (corresponding to 20 independent genes) whose expression levels differed significantly between the two clinical conditions. Construction of a dendrogram for the expression profiles of these 21 probe sets revealed that the subjects were grouped into two major branches (Fig. 2a). Although each branch corresponded approximately to the two clinical classes, a few subjects were still misclassified in both branches. It was not clear, however, whether this failure to clearly separate the two clinical classes was due to an inadequacy of the separation power of the clustering method or to the heterogeneity of the samples within each clinical class. Furthermore, these results did not address whether normal and cancerous ductal cells are truly distinct from each other from the point of view of gene expression profiles.

To address these issues, we attempted to visualize the similarity or difference between the two classes. Correspondence analysis is a relatively new approach to the decomposition of multidimensional data.[15] It allows not only a low-dimensional projection of expression profiles for numerous genes, but also measurement both of the contribution of each gene to a given extracted dimension and of the contribution of each extracted dimension to the total complexity. Correspondence analysis of the expression data of the 21 probe sets shown in Fig. 2a reduced the number of dimensions from 21 to three. On the basis of the calculated 3-D coordinates for each sample, the specimens were then projected into a virtual space (Fig. 2b). Although most of the normal samples were positioned in a region of the space distant from that occupied by the PDC specimens, the two groups were not separated completely. Decomposition of the data set was thus not sufficiently effective to achieve a high accuracy in differential diagnosis.

Supervised class prediction. We next attempted class prediction by using two supervised algorithms. The WV method was recently developed to assign binary classes based on gene expression profiles.[13] A defined number of 'class predictor' genes whose expression contrasts the two classes most effectively are first selected in a training data set. A weighting factor, which reflects how well a gene is correlated with the class distinction, is also calculated for each gene. The expression levels of the class predictors are then quantitated in the test data set, and the 'prediction strength' is determined on the basis of the expression intensities and weighting factors of the predictors. The WV method has been successfully used to differentiate acute myeloid leukemia from acute lymphoid leukemia,[13] as well as diffuse large B cell lymphoma with poor prognosis from that with good prognosis.[19]

The KNN method, like the WV method, first involves the selection of a defined number of predictor genes. It then finds nearest neighbors to the classes based on a distance function for pairs of observations. The KNN method predicts the class of a given test sample based on the majority of votes among the nearest neighbors.[20]

To measure precisely the class prediction ability of these two methods, we performed a cross-validation trial for each with our data set: One sample was therefore set aside and the program was trained with the remaining 48 samples; the class of the withheld test sample was then predicted by the program, and the trial was repeated for each of the 49 samples to calculate the overall accuracy of the program.

For both WV and KNN methods, the cross-validation was performed with the 49 specimens and with different numbers of class predictor genes (n = 1 to 20, 30, 40, 50, 60, 70, 80, 90, or 100). Both methods had similar error rates, with the WV method having a slightly lower error rate than the KNN method (Fig. 3a). The best prediction accuracy (81.6%) was obtained by the WV method with five class predictor genes. In this cross-validation, different sets of five predictors were selected for each leave-one-out trial, with a total of 11 probe sets (corresponding to 10 genes) used as predictors. Two-way clustering of the expression profiles of these 11 probe sets yielded the dendrogram shown in Fig. 3b. It should be noted that two probe sets (DKFZp564I1922 and EPPK1) were selected as the predictors in all 49 leave-one-out trials.

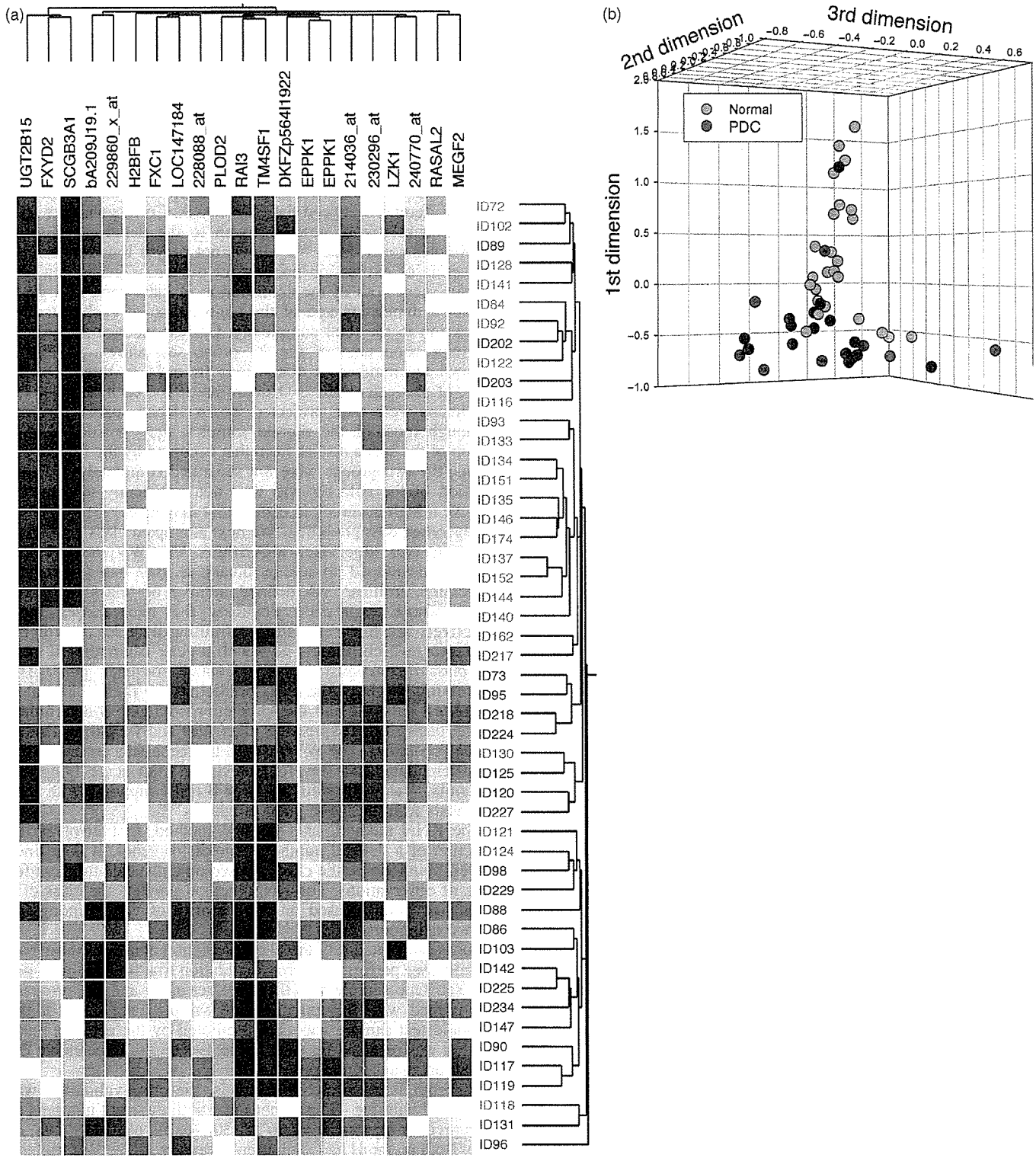

Fig. 2. Isolation of a PDC-specific molecular signature. (a) Dendrogram of the 21 probe sets whose expression level differed significantly (Welch's ANOVA, $P < 0.001$) with an effect size of ≧ 50 U between normal and cancerous specimens. Each row corresponds to a separate subject, and each column to a probe set whose expression is color-coded according to the scale in 1. Gene symbols are shown at the top; 229860_x_at, 228088_at, 214036_at, 230296_at, and 240770_at are expressed sequence tag IDs designated by Affymetrix. Detailed information on the genes and their expression levels is provided in Supplementary Information at the *Cancer Science* web site. (b) Correspondence analysis of the 21 probe sets identified three major dimensions in their expression profiles. Projection of the specimens into a virtual space with these three dimensions revealed that those from individuals with a normal pancreas and those from patients with PDC were partially separated.

 doi:10.1111/j.1349-7006.2005.00064.x

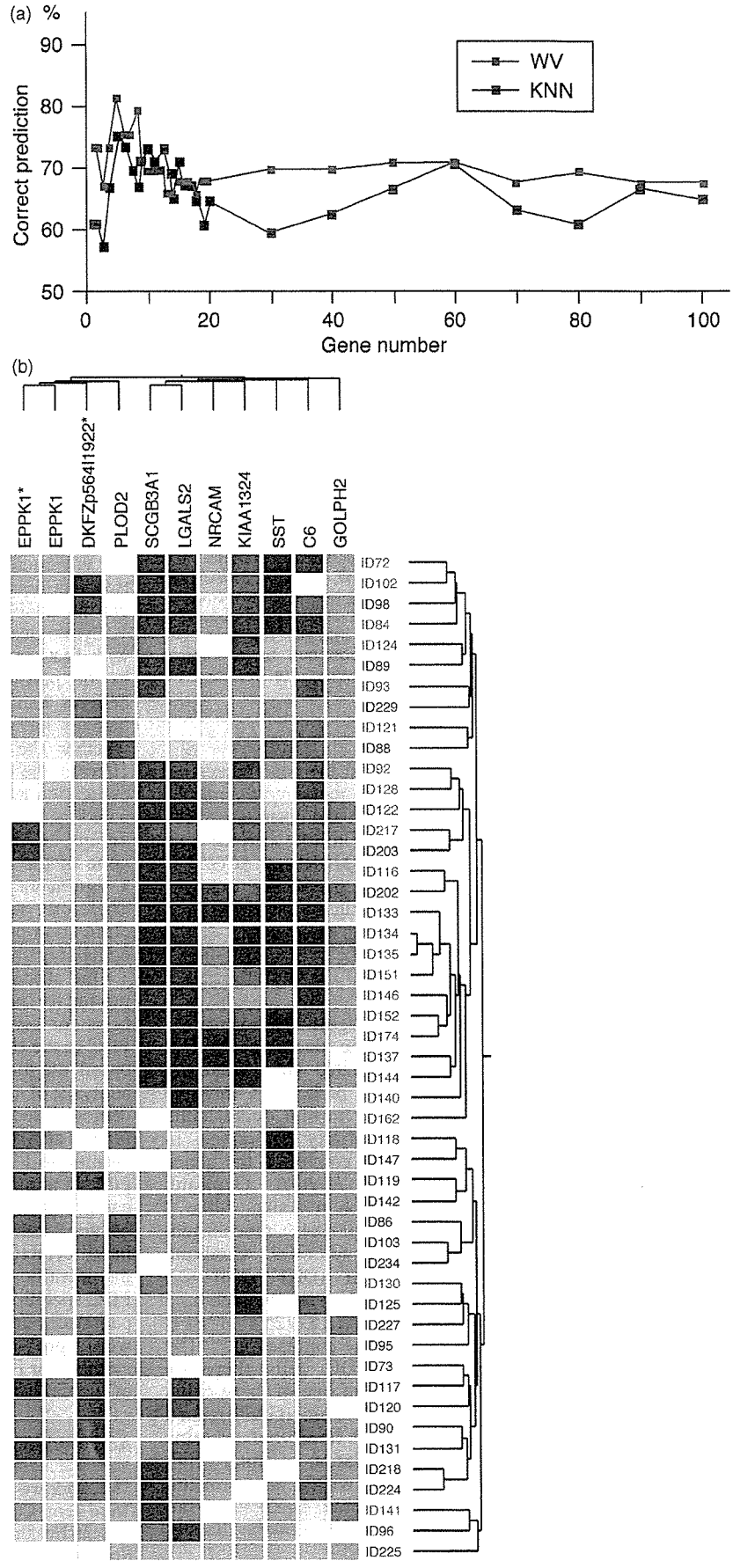

Fig. 3. Supervised class prediction. (a) Cross-validation trials for class prediction of normal or PDC specimens based on various numbers of predictor genes were performed with the WV or KNN methods. Correct prediction rate (%) is plotted for each trial. (b) Expression profiles of 11 probe sets identified by the WV method with five predictors. Samples are clustered according to the similarity in the expression pattern of the 11 probe sets. Asterisks indicate the two probe sets selected in all trials. Detailed information on the genes and their expression levels is provided in Supplementary Information at the *Cancer Science* web site.

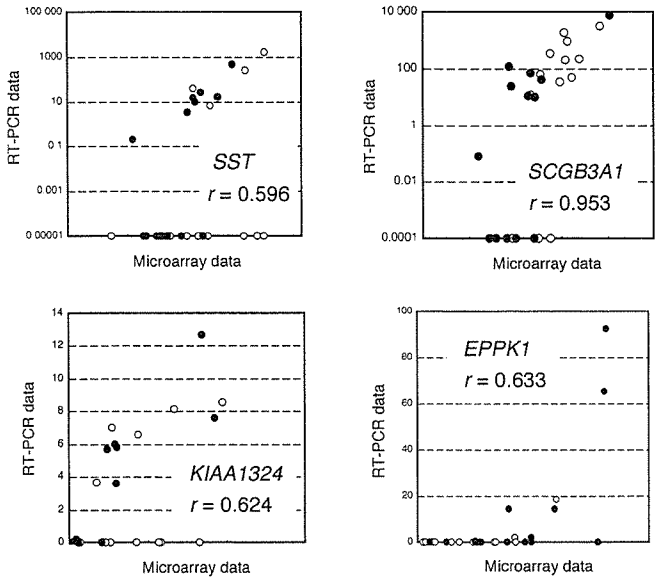

Fig. 4. Validation by reverse transcription and real-time PCR analysis of gene expression profiles obtained by microarray analysis. The relative amounts of mRNA corresponding to *SST, SCGB3A1, KIAA1324* or *EPPK1* in the MUC1$^+$ cells derived from (○) healthy individuals or (●) patients with PDC were determined by reverse transcription and real-time PCR with *ACTB* transcripts as the internal standard. The resulting values are plotted against those obtained by microarray analysis. Pearson's correlation coefficient (*r*) values are provided for each comparison.

Confirmation of expression data. To confirm the gene expression profiles obtained by microarray analysis, we measured the mRNA levels of some genes by reverse transcription and quantitative real-time PCR analysis. The relative amounts of mRNA derived from the *SST* (GenBank accession number NM_001048) or *SCGB3A1* (GenBank accession number AA742697) genes, for example, determined by this latter approach were highly correlated with those quantitated by microarray analysis (Fig. 4).

Discussion

In the present study, we constructed the largest gene expression database available to date for pancreatic ductal cells. Our statistical approach to identify genes associated with a diagnosis of PDC resulted in the extraction of 21 probe sets, three of which were preferentially expressed in normal ductal cells and the remaining 18 were preferentially expressed in cancerous ductal cells. The latter group contained the genes for H2B histone family member B (*H2BFB*; GenBank accession number BC002842), RAS protein activator-like 2 (*RASAL2*; GenBank accession number NM_004841), procollagen-lysine, 2-oxoglutarate 5-dioxygenase 2 (*PLOD2*; GenBank accession number NM_000935), adlican (DKFZp564I1922; GenBank accession number AF245505), and epiplakin 1 (*EPPK1*; GenBank accession number AL137725). H2BFB functions as a linker histone in nucleosome compaction.[21] The increased expression of *H2BFB* in PDC cells therefore probably reflects the increased proliferation rate of these cells. RASAL2 shares a GTPase-activating protein (GAP)-related domain with members of the RAS-GAP family of proteins and is thought to contribute to the regulation of small GTP-binding proteins. *RASAL2* is localized within the prostate cancer susceptibility locus at chromosome 1q25[22], so an altered activity of the encoded protein might thus be directly linked to carcinogenesis.

The expression profile of these disease-associated genes was not, however, sufficient to separate the specimens into the normal or cancer class with a high accuracy. We therefore applied sophisticated algorithms in the supervised mode in an attempt to achieve this goal. In our trials of the WV and KNN methods with various numbers of predictor genes, the WV method trained with five genes gave the best result. The accuracy of correct diagnosis achieved (81.6%) is higher than that obtained by cytological examination of pancreatic juice.[23]

In the 'leave-one-out' trials for all 49 samples, a total of 11 probe sets were chosen by the WV algorithm as the class predictors. These probe sets corresponded to 10 genes, including those for EPPK1, DKFZp564I1922, PLOD2, SCGB3A1, SST, and neuronal cell adhesion molecule (*NRCAM*; GenBank accession number NM_005010). NRCAM belongs to the immunoglobulin (Ig) superfamily of proteins, contains multiple repeats of the Ig domain in its extracellular region, and is expressed at the surface of neuronal cells. The DKFZp564I1922 protein also contains 12 repeats of the Ig domain.[24] Increased expression of these Ig domain-containing proteins may thus be a specific property and a novel molecular marker of PDC.

Among the 10 genes used in the WV analysis, only two (those for EPPK1 and DKFZp564I1922) were chosen as predictors in all 49 trials. In addition, the Welch's ANOVA strategy and the WV method selected five probe sets in common, including two sets for EPPK1, one for SCGB3A1, one for PLOD2, and one for DKFZp564I1922.

Cytological examination revealed that, among the individuals with PDC in our study, 16 patients had <20% of atypical cells in the purified ductal cell specimens ('L' in Table 1), three patients had ≧ 40% of such cells ('H'), and the other five patients had 20–40% of such cells ('M'). We thus examined whether the proportion of atypical cells in the specimens affected the expression intensities of the selected genes. The expression levels of the genes in Fig. 2a was, for instance, compared by Student's *t*-test between the individuals in the L and M groups, and between those in the M and H groups. Surprisingly, none of the genes in Fig. 2a were differentially expressed in a significant manner between these groups (data not shown). Therefore, our microarray-based prediction scheme should be of clinical importance even for patients with pancreatic juice containing small amounts of cancer cells.

Our strategy to identify a PDC-specific gene expression profile for purified pancreatic ductal cells should provide the basis for several possible scenarios for the early detection of PDC in the clinical setting. One scenario would be a microarray-based diagnosis of PDC with a sophisticated algorithm for analysis of the expression of a limited number of genes (as demonstrated in the present study). A second scenario would require an extension of our project to isolate single gene markers specific to PDC; the expression of such genes should be negligible in non-cancerous cells but would be markedly increased in cancerous cells. Such PDC-specific single gene markers would be good candidates for the construction of a sensitive PCR-based detection system for PDC. A third scenario may involve the identification of soluble proteins among the products of PDC-specific genes that could be detected in the serum of patients. Further expansion of our gene expression database would probably facilitate the development of such detection systems for PDC, which would improve the long-term prognosis of individuals with this intractable disease.

Acknowledgments

This work was supported in part by a Grant-in-Aid for research on the Third Term Comprehensive Control Research for Cancer from the Ministry of Health, Labour, and Welfare of Japan, and by a grant from the Research Foundation for Community Medicine.

 doi:10.1111/j.1349-7006.2005.00064.x

References

1 Bomman PC, Beckingham IJ. Pancreatic tumours. *Br Med J* 2001; **322**: 721–3.
2 Rosewicz S, Wiedenmann B. Pancreatic carcinoma. *Lancet* 1997; **349**: 485–9.
3 Sawabu N, Watanabe H, Yamaguchi Y, Ohtsubo K, Motoo Y. Serum tumor markers and molecular biological diagnosis in pancreatic cancer. *Pancreas* 2004; **28**: 263–7.
4 Kondo H, Sugano K, Fukayama N *et al.* Detection of point mutations in the K-ras oncogene at codon 12 in pure pancreatic juice for diagnosis of pancreatic carcinoma. *Cancer* 1994; **73**: 1589–94.
5 Furuya N, Kawa S, Akamatsu T, Furihata K. Long-term follow-up of patients with chronic pancreatitis and K-ras gene mutation detected in pancreatic juice. *Gastroenterology* 1997; **113**: 593–8.
6 Schena M, Shalon D, Davis RW, Brown PO. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* 1995; **270**: 467–70.
7 Duggan DJ, Bittner M, Chen Y, Meltzer P, Trent JM. Expression profiling using cDNA microarrays. *Nat Genet* 1999; **21**: 10–4.
8 Yoshida K, Ueno S, Iwao T *et al.* Screening of genes specifically activated in the pancreatic juice ductal cells from the patients with pancreatic ductal carcinoma. *Cancer Sci* 2003; **94**: 263–70.
9 Terada T, Ohta T, Sasaki M, Nakamura Y, Kim YS. Expression of MUC apomucins in normal pancreas and pancreatic tumours. *J Pathol* 1996; **180**: 160–5.
10 Van Gelder RN, von Zastrow ME, Yool A, Dement WC, Barchas JD, Eberwine JH. Amplified RNA synthesized from limited quantities of heterogeneous cDNA. *Proc Natl Acad Sci USA* 1990; **87**: 1663–7.
11 Oshima Y, Ueda M, Yamashita Y *et al.* DNA microarray analysis of hematopoietic stem cell-like fractions from individuals with the M2 subtype of acute myeloid leukemia. *Leukemia* 2003; **17**: 1990–7.
12 Affymetrix [web site on the Internet]. Support: Mask files. Available from URL: http://www.affymetrix.com/support/technical/mask_files.affx
13 Golub TR, Slonim DK, Tamayo P *et al.* Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science* 1999; **286**: 531–7.
14 Broad Institute [web site on the Internet]. GeneCluster 2.0. Available from URL: http://www.broad.mit.edu/cancer/software/genecluster2/gc2.html
15 Fellenberg K, Hauser NC, Brors B, Neutzner A, Hoheisel JD, Vingron M. Correspondence analysis applied to microarray data. *Proc Natl Acad Sci USA* 2001; **98**: 10781–6.
16 ViSta. [web site on the Internet]. ViSta: the visual statistics system. Available from URL: http://www.visualstats.org
17 Alon U, Barkai N, Notterman DA *et al.* Broad patterns of gene expression revealed by clustering analysis of tumor and normal colon tissues probed by oligonucleotide arrays. *Proc Natl Acad Sci USA* 1999; **96**: 6745–50.
18 Dhanasekaran SM, Barrette TR, Ghosh D *et al.* Delineation of prognostic biomarkers in prostate cancer. *Nature* 2001; **412**: 822–6.
19 Shipp MA, Ross KN, Tamayo P *et al.* Diffuse large B-cell lymphoma outcome prediction by gene-expression profiling and supervised machine learning. *Nat Med* 2002; **8**: 68–74.
20 Dudoit S, Fridlyand J, Speed TP. [pre-print PDF on the Internet]. Comparison of discrimination methods for the classification of tumors using gene expression data. Available from URL: http://www.stat.berkeley.edu/tech-reports/index.html
21 Albig W, Kardalinou E, Drabent B, Zimmer A, Doenecke D. Isolation and characterization of two human H1 histone genes within clusters of core histone genes. *Genomics* 1991; **10**: 940–8.
22 Noto S, Maeda T, Hattori S, Inazawa J, Imamura M, Asaka M, Hatakeyama M. A novel human RasGAP-like gene that maps within the prostate cancer susceptibility locus at chromosome 1q25. *FEBS Lett* 1998; **441**: 127–31.
23 Nakaizumi A, Tatsuta M, Uehara H *et al.* Cytologic examination of pure pancreatic juice in the diagnosis of pancreatic carcinoma. The endoscopic retrograde intraductal catheter aspiration cytologic technique. *Cancer* 1992; **70**: 2610–4.
24 Strausberg RL, Feingold EA, Grouse LH *et al.* Generation and initial analysis of more than 15 000 full-length human and mouse cDNA sequences. *Proc Natl Acad Sci USA* 2002; **99**: 16 899–903.
25 Isaji S, Kawarada Y, Umemoto S. Classification of pancreatic cancer. Comparison of Japanese and UICC classifications. *Pancreas* 2004; **28**: 231–4.

Supplementary Material

The following supplementary material is available for this article online:

Table S1. Annotation information and expression intensity data for the genes shown in Figure 2A.

Table S2. Annotation information and expression intensity data for the genes shown in Figure 3B.

Genome-Wide Screening for Target Regions of Histone Deacetylases in Cardiomyocytes

Ruri Kaneda, Shuichi Ueno, Yoshihiro Yamashita, Young Lim Choi, Koji Koinuma, Shuji Takada, Tomoaki Wada, Kazuyuki Shimada, Hiroyuki Mano

Abstract—The acetylation status of core histones in cardiomyocytes has been linked to the development of cardiac hypertrophy and heart failure. Little is known, however, of the genes affected by abnormal histone acetylation in such pathological conditions. We recently developed a genome-wide screening method, differential chromatin scanning (DCS), to isolate genomic fragments associated with histones subject to differential acetylation. We have now applied DCS to H9C2 rat embryonic cardiomyocytes incubated with or without trichostatin A (TSA), a specific inhibitor of histone deacetylase (HDAC) activity. About 200 genomic fragments were readily isolated by DCS on the basis of the preferential acetylation of associated histones in TSA-treated cells. Quantitation of the amount of DNA in chromatin immunoprecipitates prepared with antibodies to acetylated histone H3 revealed that 37 of 38 randomly chosen DCS clones were preferentially precipitated from the TSA-treated cells, thus verifying the high fidelity of DCS. Epigenetic regulation of DCS clones was further confirmed in cells treated with sodium butyrate, another HDAC inhibitor, as well as in cardiac myocytes isolated from neonatal rats. The mRNA level of 9 (39%) of 23 genes corresponding to DCS clones changed in parallel with the level of histone acetylation in H9C2 cells. Furthermore, a physiological hypertrophic stimulus, cardiotrophin-1, affected the acetylation level of histones associated with genomic regions corresponding to certain DCS clones. Our data thus establish a genome-wide profile of HDAC targets in cardiomyocytes, which should provide a basis for further investigations into the role of epigenetic modification in cardiac disorders. (***Circ Res.* 2005;97:210-218.**)

Key Words: epigenetics ■ histone acetylation ■ trichostatin A ■ cardiomyocyte

Epigenetic modification of chromatin includes methylation of genomic DNA as well as acetylation, methylation, and phosphorylation of histone proteins. Such epigenetic changes play important roles in the regulation of gene transcriptional activity associated with cell growth and differentiation as well as with organ development.[1–3] Acetylation of core histones is mediated by histone acetyltransferases (HATs) and, in many instances, results in relaxation of chromatin structure and transcriptional activation of associated genes.[4] Histone deacetylases (HDACs) counteract HAT activity by catalyzing the removal of acetyl moieties from lysine residues in histone tails, thereby inducing chromatin condensation and transcriptional repression.[5]

Regulation of histone acetylation has been linked to cardiac hypertrophy. The HAT activity of CREB-binding protein (CBP) and p300 is thus required for the induction of hypertrophic changes in cardiac muscle cells by phenylephrine.[6] Consistent with this observation, inhibition of HDAC activity results in an increase in the size of muscle cells.[7] Furthermore, HDACs of class II (HDAC-4, -5, -7, and -9) suppress cardiac hypertrophy in part by binding to and inhibiting the activity of myocyte enhancer factor 2 (MEF2).[8] In contrast, however, HDAC2 together with Hop was found to promote cardiac hypertrophy in vivo in a manner sensitive to systemic administration of the HDAC inhibitor trichostatin A (TSA).[9] Moreover, HDAC inhibitors prevent hypertrophy and sarcomere organization in cultured cardiac myocytes,[10] suggestive of a positive role for HDACs in cardiac hypertrophy.

These seemingly discrepant findings may be attributable either to differential actions of different classes of HDACs (and, possibly, of HATs) with regard to myocyte hypertrophy or to a dissociation between the deacetylase activity of HDACs and a prohypertrophic function.[8] Clarification of the role of HATs and HDACs in hypertrophy would be facilitated by identification of the genes targeted by these enzymes during the induction of hypertrophic changes. Little is known, however, of the genes regulated by HATs or HDACs in myocytes. Induction of the atrial natriuretic peptide (ANP) gene is associated with acetylation of histones (H3 and H4) located in the 3′ untranslated region of the gene.[11] Histones bound to the β-myosin heavy chain gene have also been shown to be targeted by HATs in myocytes.[8]

Original received February 21, 2005; revision received June 1, 2005; accepted June 24, 2005.

From the Divisions of Functional Genomics (R.K., S.U., Y.Y., Y.L.C., K.K., S.T., T.W., H.M.) and Cardiovascular Medicine (R.K., S.U., K.S.), Jichi Medical School, Tochigi, Japan; and CREST (H.M.), Japan Science and Technology Agency, Saitama.

Correspondence to Dr Hiroyuki Mano, Division of Functional Genomics, Jichi Medical School, 3311-1 Yakushiji, Kawachigun, Tochigi 329-0498, Japan. E-mail hmano@jichi.ac.jp

Circulation Research is available at http://circres.ahajournals.org DOI: 10.1161/01.RES.0000176028.18423.07

We have recently established a new technique, differential chromatin scanning (DCS),[12] for genome-wide screening of DNA regions associated with histones that are differentially acetylated between a given pair of cell or tissue samples. To isolate target genes of HDACs in cardiac myocytes, we have now applied DCS to a rat embryonic heart–derived myogenic cell line, H9C2, treated or not with TSA. More than 200 genomic fragments were readily isolated by DCS, and genomic regions corresponding to 37 clones of 38 examined were confirmed to be associated with differentially acetylated histones. Furthermore, the expression of genes located in or close to such regions paralleled the associated level of histone acetylation.

Materials and Methods

Cell Culture

H9C2 cells were obtained from American Type Culture Collection (Rockville, Md) and were maintained in Dulbecco's modified Eagle's medium (DMEM; Invitrogen) supplemented with 10% fetal bovine serum (Invitrogen) and 2 mmol/L L-glutamine. For preparation of the tester sample, cells were incubated for 24 hours with 300 nM TSA (Wako). For other treatments, cells were incubated with 2 or 4 mmol/L sodium butyrate (Sigma) for 24 hours or with 1 nM cardiotrophin-1 (Calbiochemy) for the indicated times.

Neonatal cardiac myocytes were prepared as described previously.[13] In brief, ventricular tissue was dissected from newborn rats and subjected to digestion overnight at 4°C with trypsin (1 mg/mL; Invitrogen) in Hanks' balanced salt solution (Invitrogen). Myocytes were harvested by subsequent digestion of the tissue with collagenase (Worthington,) and were centrifuged twice at 50*g* to remove less dense cells such as fibroblasts. Myocytes were then cultured in DMEM-F12 (Invitrogen) supplemented with 10% fetal bovine serum and 2 mmol/L L-glutamine.

Differential Chromatin Scanning

HDAC targets were screened in H9C2 cells by DCS as described previously.[12] In brief, both tester and driver cells were fixed and subjected to immunoprecipitation with antibodies to acetylated histone H3 with the use of a chromatin immunoprecipitation (ChIP) assay kit (Upstate Biotechnology). DNA fragments recovered from the immunoprecipitates were digested with *Rsa*I (New England Biolabs), and the digestion products were ligated to the TAG adapter (5′-CCACCGCCATCCGAGCCTTTCTGCCCGGG-3′/3′-GAAAGACGGGCCC-5′). After polymerase chain reaction (PCR)–mediated amplification with the TAG primer (5′-CCACCGCCATCCGAGCCTTTCTGC-3′), the tester and driver DNA samples were digested with *Xma*I and *Sma*I, respectively. The tester DNA (0.5 μg) was ligated to the first subtraction adapter (5′-GTGAGGGTCGGATCTGGCTGGCTC-3′/3′-CGACCGAGGGCC-5′), annealed with 40 μg of the driver DNA at 67°C for 20 to 24 hours, and then subjected to PCR with the first subtraction primer (5′-GTGAGGGTCGGATCTGGCTGGCTC-3′). After digestion of single-stranded DNA with mung-bean nuclease (New England Biolabs), the amplified products were subjected to digestion with *Xma*I followed by a second round of subtraction PCR with the second subtraction adapter (5′-GTTAGCGGACACAGGGCGGGTCAC-3′/3′-GCCCAGTGGGCC-5′) and second subtraction primer (5′-GTTAGCGGACACAGGGCGGGTCAC-3′). The final products were digested with *Xma*I and ligated into pBlueScript (Stratagene). *Escherichia coli* DH5α cell clones transformed with the resulting recombinant plasmids were grown in 96-well plates and subjected to direct plasmid purification in the plates with the use of a Montage Plasmid Miniprep$_{96}$ Kit (Millipore). The nucleotide sequences of the purified plasmids were then determined by Dragon Genomics Center (Mie) and were used to screen, with the BLAT search program,[14] the nucleotide sequence database (http://genome.ucsc.edu/) assembled in June 2003 by the Genome Bioinformatics Group of the University of California at Santa Cruz (UCSC).

Quantitation of DNA

Genomic fragments immunoprecipitated by antibodies to acetylated histone H3 (Upstate Biotechnology) were subjected to PCR with a QuantiTect SYBR Green PCR Kit (Qiagen). The amplification protocol comprised incubations at 94°C for 15 s, 60°C for 30 s, and 72°C for 1 minute. Incorporation of the SYBR green dye into PCR products was monitored in real time with an ABI PRISM 7700 sequence detection system (PE Applied Biosystems), thereby allowing determination of the threshold cycle (C_T) at which exponential amplification of PCR products begins. The C_T values for DNA molecules in the immunoprecipitates and for those in the original sample before immunoprecipitation were used to calculate the abundance of the former relative to that of the latter. The oligonucleotide primers for PCR were 5′-CCGGAAGAGGTGGTTATGTAAA-3′ and 5′-GCTAAGAAGGGACAGGGCTAAC-3′ for the H9C2T-2_D09 clone, 5′-GTTTGTCTGGAGCCTGTACTCTC-3′ and 5′-AAGTTCTCCGTTTCAGGATTCAC-3′ for the H9C2T-2_C06 clone, 5′-CACATCCTTGGTGCTTCTGA-3′ and 5′-GAGGAGGGTGAGGAGCTGAG-3′ for the H9C2T-1_E03-1 clone, and 5′-CCCGGTGTTCTGTACGTAGG-3′ and 5′-ACTGATGGAGCATCCACATTCT-3′ for the H9C2T-S-1-8 clone.

Quantitation of mRNA

Total RNA was prepared from the tester and driver cells with an RNeasy Mini column (Qiagen) and was subjected to reverse transcription (RT) with PowerScript reverse transcriptase (BD Biosciences Clontech) and an oligo(dT) primer. Portions of the resulting cDNA were subjected to PCR with a QuantiTect SYBR Green PCR Kit. The amplification protocol comprised incubations at 94°C for 15 s, 60°C for 30 s, and 72°C for 1 minute. The C_T values for cDNAs corresponding to glyceraldehyde-3-phosphate dehydrogenase (GAPDH) mRNA and the mRNAs of interest were used to calculate the abundance of the latter relative to that of the former. The C_T values for GAPDH mRNA determined with 10 μg of total RNA from TSA-treated or nontreated cells were 18.5±0.8 and 18.2±0.2 (mean±SD), respectively, validating the use of GAPDH mRNA as an internal control.

The oligonucleotide primers for PCR were 5′-AATGTATCCGTTGTGGATCTGAC-3′ and 5′-ATTGTCATACCAGGAAATGAGCTT-3′ for GAPDH, 5′-GCCTGTGATACTCTGCTTATGTGT-3′ and 5′-CTTGAGGATTTCCTCTTTCTTCTG-3′ for the inositol 1,4,5-trisphosphate receptor type 3 gene (Itpr3), 5′-CAGTACCCTGTTGAGTCATCTCTG-3′ and 5′-GAAAGCAAGGTCTTCTTATTCTGG-3′ for the NAD(P)H dehydrogenase, quinone 1 gene (Nqo1), 5′-GCCTTCTACCTGCATACTACCAAG-3′ and 5′-AGTCTCAAGATACCGGAGCACA-3′ for the metastasis-associated 1 gene (Mta1), and 5′-CTGTTGGTACCTGTGCTGTGTAG-3′ and 5′-ACTGGTAGAGTACGTCCTTGTGG-3′ for the Jagged2 precursor gene (Jag2).

Statistical Analysis

Quantitation of DNA or mRNA was performed in triplicate in at least 2 independent experiments, and data are presented as mean±SD. The statistical significance of differences was analyzed by Student *t* test, with a probability value of <0.05 being considered significant.

Results

DCS in H9C2 Cells

Given that the ANP gene is a known target of HDAC in myocytes,[11] we first examined the effect of TSA on the acetylation level of histones bound to this gene in H9C2 cells. Real-time PCR analysis revealed that the amount of the 3′ untranslated region of the ANP gene that was precipitated by antibodies to acetylated histone H3 from TSA-treated cells was 7.85 times that precipitated from nontreated cells (data

not shown), indicating that the ANP gene is indeed a target of HDAC activity in H9C2 cells.

With the use of TSA-treated cells as the tester and nontreated cells as the driver, we then performed DCS, which in effect couples ChIP with subtraction PCR. After the second round of subtraction PCR, we sequenced the isolated genomic clones in a 96-well plate format. Analysis of 3 such plates thus yielded the nucleotide sequences of 288 DCS products. Among these randomly selected products, 222 clones contained >50 bp and were used to screen the nucleotide sequence database of UCSC with the use of the BLAT program.

A total of 195 clones showed >95% sequence identity to the rat genome sequence, and 178 of these clones were located either within protein-coding genes (demonstrated or predicted) or in the vicinity (within 10 kbp) of such genes (119 independent genes) (Table 1; Table SI available in the online supplement at http://circres.ahajournals.org). Forty-two (23.6%) of the 178 clones were assigned to a region spanning the promoter (0 to −2000 bp relative to the transcriptional start site), the first exon, and the first intron of the corresponding genes. Given that protein-coding genes account for only a few percent of the rat genome,[15] our data suggest that histone acetylation occurs preferentially at regions of the genome involved directly in transcriptional regulation.

Eleven DCS clones were assigned to overlapping sequences upstream of the Oct11 gene, and 7 clones were mapped to overlapping sequences at chromosomal position 8q24, a region with no annotation information (data not shown). The isolation of such multiple clones for individual genomic regions suggests that the DCS products isolated may represent most HDAC targets in H9C2 cells.

HDAC Targets in a Cardiomyocyte Cell Line

To verify the fidelity of DCS, we randomly selected 38 DCS clones and quantified the corresponding genomic fragments in immunoprecipitates prepared from both tester and driver cells with antibodies to acetylated histone H3. The amount of each DNA fragment in the immunoprecipitate relative to that in the original sample before ChIP was determined by real-time PCR. Selective amplification by DCS proved to be highly reliable (Table 1), with 37 of the 38 clones exhibiting tester-selective precipitation (tester/driver ratio of ≥1.5). It is therefore likely that DCS indeed identified targets of HDAC in myocytes.

To visualize directly the genome-wide distribution of HDAC targets, we mapped to rat chromosome figure our genomic clones whose chromosomal positions were known (Figure 1). The HDAC targets were distributed widely throughout the rat genome, although some "hot spots" for deacetylation were apparent. For example, 7 of the DCS clones mapped to chromosomal position 5q36, and detailed mapping revealed that all of these clones were located within a region spanning 27 Mbp. It is thus possible that regional alterations of chromatin structure result in coordinated transcriptional regulation of genes within the affected region.

Some of the clones listed in Table 1 correspond to loci within or close to rat genes whose products function in intracellular calcium mobilization or antioxidant processes. One such clone (H9C2T-2_D09), for instance, mapped to a region encompassing intron 21 and exon 22 of Itpr3 (Figure 2A), which encodes a receptor for inositol 1,4,5-trisphosphate that plays an important role in Ca^{2+}-mediated signal transduction. The cytosolic concentration of Ca^{2+} directly regulates muscle contraction and cardiac rhythm and is a determinant of myocyte hypertrophy and heart failure.[16] The amount of the genomic fragment corresponding to the H9C2T-2_D09 clone was 6.6-fold greater in the ChIP product of TSA-treated cells than in that of untreated cells (Figure 2B), indicating that the extent of histone acetylation in this region of the genome of the tester cells was 6.6 times that in the driver cells. Furthermore, inhibition of HDAC activity was accompanied by an increase in the amount of Itpr3 mRNA (Figure 2C). These data suggest that HDAC actively deacetylates a chromosomal region corresponding to Itpr3 and thereby suppresses the transcriptional activity of the gene.

Another clone (H9C2T-2_C06) was mapped to the first intron of Nqo1 (Figure 2D), which encodes a reductase that contributes to detoxification of quinones and to regulation of apoptosis.[17] We examined whether the acetylation of associated histones and the expression of Nqo1 are regulated by HDAC activity in cardiomyocytes. As with Itpr3, the acetylation level of histones bound to Nqo1 was increased by TSA treatment in H9C2 cells (Figure 2E), and this epigenetic change was accompanied by an increase in the amount of Nqo1 mRNA (Figure 2F).

Nineteen (10.7%) of the 178 clones whose chromosomal location was known were assigned to loci corresponding to at least 2 genes in the rat genome. One such clone, H9C2T-1_E03-1, was mapped to a region corresponding to the first intron of Mta1 and to the last exon of Jag2 (Figure 3A). Histone acetylation in this region might thus affect the transcription of both genes simultaneously. The level of histone H3 acetylation in this region was confirmed to be greater in the tester cells than in the driver cells (Figure 3B). However, although inhibition of HDAC activity by TSA resulted in upregulation of the amount of Jag2 mRNA (Figure 3C), it had no effect on the abundance of Mta1 mRNA (Figure 3D). Histone acetylation in the genomic region that encompasses both Jag2 and Mta1 thus appears to regulate the transcriptional activity of the former gene but not that of the latter. Jag2 is a ligand for the receptor Notch1 and is abundant in the heart.[18] Coculture of fibroblasts expressing human JAG2 with murine C2C12 myoblasts resulted in inhibition of myogenic differentiation of the latter cells, implicating JAG2 in regulation of this process.

We selected an additional 19 DCS clones for quantitation of the corresponding mRNAs. Among the genes examined, 6 were preferentially expressed (tester/driver ratio of ≥1.5) in the TSA-treated cells compared with the nontreated cells (Table 1).

Regulation of Histone Acetylation in Neonatal Rat Cardiac Myocytes

Our DCS analysis identified HDAC targets in a cardiomyocyte cell line. To investigate whether the level of histone

Fragments of the Rat Genome Isolated by DCS in H9C2 Cells

Clone ID	Tester/Driver DNA Ratio	Annotation	GenBank Accession No. (or Ensemble Gene ID)	Locus	Tester/Driver mRNA Ratio	Position Relative to Corresponding Genes
H9C2T-S-2–8	13.7	EST	BQ204614	19q12	5.34	Intron 2
H9C2T-S-1-5b	10.92	Skn-1a	L23862	8q22		4 kbp upstream of exon 1
H9C2T-1_D09	10.7	Putative G protein–coupled receptor (SENR) gene (Senr)	NM_020537	10q32.3	1.09×10^{8}	Exon 1
H9C2T-S-1-2b	9.36	NA	NA	20p12		
H9C2T-1_B05	8.43	No description	(ENSRNOG00000013959)	8q24		Last exon+last intron
H9C2T-S-1-8	7.96	Myocilin (Myoc)	NM_030865	13q22	2.03	2.5 kbp upstream of exon 1
H9C2T-1_E03-1*	7.85	Mta1 (mta1)	NM_022588	6q32	0.91	Intron 1
		Jagged2 precursor	U70050		8.43	Last exon
H9C2T-2_E05	7.7	Protein kinase C and casein kinase substrate in neurons 1 (Pacsin1)	NM_017294	20p12	26.45	Intron 1
H9C2T-1_C04-2	7.69	Carbohydrate (chondroitin 6/keratan) sulfotransferase 3 (Chst3)	NM_053408	20q11	0.25	Intron 1
H9C2T-S-1-2a	6.99	Phosphodiesterase	L27061	16p14	1.12×10^{8}	1.3 kbp upstream of exon 1
H9C2T-S-1-4	6.76	EST	CB741111	14p22		Intron 3
H9C2T-S-1-5a	6.64	Brain and reproductive organ–expressed protein	NM_199270	6q13		Intron 10
H9C2T-1_C04-1	6.61	EST	C0562897	3q43		Exon 1
H9C2T-2_D09	6.6	Inositol 1,4,5-trisphosphate receptor 3 (Itpr3)	NM_013138	20p12	1.66	Intron 21+exon 22
H9C2T-2_E11	6.51	EST	CV077786	3q43		Last exon
H9C2T-1_H05-1	6.29	Microtubule-associated protein tau (Mapt)	NM_017212	10q32.1	602.17	Intron 7
H9C2T-1_G12	5.83	No description	(ENSRNOG00000004217)	10q12		Intron 2
H9C2T-1_H01	5.36	A disintegrin and metalloproteinase domain 1 (fertilin alpha) (Adam1)	NM_020078	12q16	0.46	Last intron
H9C2T-2_F10-1	5.13	Translocase of outer mitochondrial membrane 20 homolog (yeast) (Tomm20)	NM_152935	19q12		Last intron
H9C2T-S-1-7	5.06	EST	BF282351	4q34	0.68	2 kbp upstream of exon 1
H9C2T-1_H10	4.92	Collagen, type V, alpha 1 (Col5a1)	NM_134452	3p12	0.86	Intron 1
H9C2T-2_H07	4.4	Runt-related transcription factor 1 (Runx1)	NM_017325	11q11	1.21	Intron 4
H9C2T-S-1-3	4.36	No description	(ENSRNOG00000021887)	8q32		Last exon+last intron
H9C2T-S-2-3	4.35	EST	C0557128	1q43		0.5 kbp upstream of exon 1
H9C2T-2_A01′	3.9	Fos-like antigen 1 (Fosl1)	NM_012953	1q43	0.39	1 kbp downstream of last exon
H9C2T-S-2-7	3.86	Arylsulfatase B (ARSB)	D49434	2q12	0.48	Intron 4
H9C2T-1_F10	3.85	Phosphatidylinositol 4-kinase (Pik4cb)	NM_031083	2q34	0.88	Exon 4
H9C2T-2_C06	3.78	NAD(P)H dehydrogenase, quinone 1 (Nqo1)	NM_017000	19q12	2.82	Intron 1
H9C2T-1_D03	3.79	Cytokine-inducible SH2-containing protein	AF065161	8q32	1.47	Last exon
H9C2T-1_F09-1*	3.74	EST	CK597511	15p16		Immediately upstream of exon 1
		Normalized rat muscle, Bento Soares Rattus sp. cDNA clone RMUBG18 3′ end	AI171102			Intron 1
		EST	CK598708			Exon 1
H9C2T-2_B07	3.62	Cyclin D1 (Ccnd1)	NM_171992	1q42	0.59	Intron 3+exon 4
H9C2T-1_C03-1	3.58	G protein beta subunit–like (Gbl)	NM_022404	10q12	0.77	Last exon
H9C2T-2_A11	3.4	EST	C0557128	1q43		0.5 kbp upstream of exon 1
H9C2T-S-2-1	3.37	Period1 (rper1), partial cds	AB092976	10q24	0.52	Introns 15, 16+exon 16
H9C2T-2_E03-2	2.94	Interleukin-11 (Il11)	NM_133519	1q12	1	Intron 5
H9C2T-S-1-1	2.69	No description	(ENSRNOG00000001730)	11q22		Intron 3
H9C2T-1_C09	1.83	No description	(ENSRNOG00000025448)	10q32.1		0.5 kbp upstream of exon 1

EST indicates expressed sequence tag; NA, not assigned. *Fragments that mapped to loci of multiple genes.