27. Boland CR, Thibodeau SN, Hamilton SR, Sidransky D, Eshleman JR, Burt RW, Meltzer SJ, Rodriguez-Bigas MA, Fodde R, Ranzani GN, Srivastava S. A National Cancer Institute workshop on microsatellite instability for cancer detection and familial predisposition: development of international criteria for the determination of microsatellite instability in colorectal cancer. Cancer Res 1998;58:5248–57.
28. Konishi M, Kikuchi-Yanoshita R, Tanaka K, Muraoka M, Onda A, Okumura Y, Kishi N, Iwama T, Mori T, Koike M, Ushio K, Chiba M, et al. Molecular nature of colon tumors in hereditary nonpolyposis colon cancer, familial polyposis, and sporadic colon cancer. Gastroenterology 1996;111:307–17.
29. Jass JR, Biden KG, Cummings MC, Simms LA, Walsh M, Schoch E, Meltzer SJ, Wright C, Searle J, Young J, Leggett BA. Characterisation of a subtype of colorectal cancer combining features of the suppressor and mild mutator pathways. J Clin Pathol 1999;52:455–60.
30. Suraweera N, Duval A, Reperant M, Vaury C, Furlan D, Leroy K, Seruca R, Iacopetta B, Hamelin R. Evaluation of tumor microsatellite instability using five quasimonomorphic mononucleotide repeats and pentaplex PCR. Gastroenterology 2002;123:1804–11.
31. Kolch W. Meaningful relationships: the regulation of the Ras/Raf/MEK/ERK pathway by protein interactions. Biochem J 2000;351(Pt 2):289–305.
32. Peyssonnaux C, Eychene A. The Raf/MEK/ERK pathway: new concepts of activation. Biol Cell 2001;93:53–62.
33. Barnier JV, Papin C, Eychene A, Lecoq O, Calothy G. The mouse B-raf gene encodes multiple protein isoforms with tissue-specific expression. J Biol Chem 1995;270:23381–9.
34. Yuen ST, Davies H, Chan TL, Ho JW, Bignell GR, Cox C, Stephens P, Edkins S, Tsui WW, Chan AS, Futreal PA, Stratton MR, et al. Similarity of the phenotypic patterns associated with BRAF and KRAS mutations in colorectal neoplasia. Cancer Res 2002;62:6451–5.
35. Herman JG, Umar A, Polyak K, Graff JR, Ahuja N, Issa JP, Markowitz S, Willson JK, Hamilton SR, Kinzler KW, Kane MF, Kolodner RD, et al. Incidence and functional consequences of hMLH1 promoter hypermethylation in colorectal carcinoma. Proc Natl Acad Sci USA 1998;95:6870–5.
36. Esteller M, Toyota M, Sanchez-Cespedes M, Capella G, Peinado MA, Watkins DN, Issa JP, Sidransky D, Baylin SB, Herman JG. Inactivation of the DNA repair gene O6-methylguanine-DNA methyltransferase by promoter hypermethylation is associated with G to A mutations in K-ras in colorectal tumorigenesis. Cancer Res 2000;60:2368–71.
37. Esteller M, Risques RA, Toyota M, Capella G, Moreno V, Peinado MA, Baylin SB, Herman JG. Promoter hypermethylation of the DNA repair gene O(6)-methylguanine-DNA methyltransferase is associated with the presence of G:C to A:T transition mutations in p53 in human colorectal tumorigenesis. Cancer Res 2001;61:4689–92.
38. Lee JT Jr, McCubrey JA. The Raf/MEK/ERK signal transduction cascade as a target for chemotherapeutic intervention in leukemia. Leukemia 2002;16:486–507.
39. Sebolt-Leopold JS. Development of anticancer drugs targeting the MAP kinase pathway. Oncogene 2000;19:6594–9.

High-throughput screening of genome fragments bound to differentially acetylated histones

Ruri Kaneda[1,2], Minoru Toyota[3], Yoshihiro Yamashita[1], Koji Koinuma[1], Young Lim Choi[1], Jun Ota[1,4], Hiroyuki Kisanuki[1], Madoka Ishikawa[1], Shuji Takada[1], Kazuyuki Shimada[2] and Hiroyuki Mano[1,4,*]

Divisions of [1]Functional Genomics, and [2]Cardiology, Jichi Medical School, 3311-1 Yakushiji, Kawachigun, Tochigi 329-0498, Japan
[3]Department of Molecular Biology, Cancer Research Institute, Sapporo Medical University, Sapporo 060-8556, Japan
[4]CREST, Japan Science and Technology Agency, Saitama 332-0012, Japan

Although acetylation-deacetylation of histones contributes to regulation of gene expression, few methods have been available to determine the whole-genome histone acetylation profile in specific cells or tissues. We have now developed a genome-wide screening method, differential chromatin scanning (DCS), to isolate genome fragments embedded in histones subject to differential acetylation. This DCS screening was applied to a human gastric cancer cell line incubated with or without an inhibitor of histone deacetylase (HDAC) activity, resulting in the rapid identification of more than 250 genome fragments. Interestingly, a number of cancer-related genes were revealed to be the targets of HDAC in the cancer cells, including those for tumour protein 73 and cell division cycle 34. Such differential acetylation of histone was also shown to be linked to the regulation of transcriptional activity of the corresponding genes. Among the isolated genome fragments, 94% (32/34) of them were confirmed to be bound to differentially acetylated histones, and the genes corresponding to 78% (7/9) of them exhibited differential transcriptional activity consistent with the level of histone acetylation. With its high fidelity, the DCS method should open a possibility to rapidly compare the genome-wide histone acetylation profiles and to provide novel insights into molecular carcinogenesis.

Introduction

The genomic DNA of eukaryotes is compactly packaged by complexation with histones, rendering it physically difficult for the nuclear machinery responsible for DNA transcription or replication to access the DNA. However, this three-dimensional structure of genomic DNA is subject to dynamic and rapid regulation by various chemical modifications of the DNA itself or of its surrounding histones. These epigenetic modifications include DNA methylation as well as histone acetylation, methylation, and phosphorylation (Baylin & Herman 2000; Burgess-Beusse *et al.* 2002; van Leeuwen & Gottschling 2002).

Acetylation of histones is mediated by histone acetyltransferases (HATs) and takes place on the ε-amino group of conserved lysine residues located in the NH_2-terminal tail of core histones (Carrozza *et al.* 2003). Such histone modification is tightly linked to transcriptional regulation by either chromatin remodeling or the provision of binding sites for other proteins. HAT activity in cells is rapidly counteracted by the activity of histone deacetylases (HDACs) (Verdin *et al.* 2003), with the result that the turnover time of histone acetylation is as short as a few minutes (Waterborg 2002).

The importance of histone acetylation in the regulation of gene expression has been demonstrated for a variety of cellular processes including cell differentiation, cell cycle progression, DNA repair, and carcinogenesis (Kouzarides 1999; Yasui *et al.* 2003). Furthermore, a histone acetylation profile (HAP) of a certain type of acute myeloid leukaemia was shown to be directly related to the sensitivity of the malignancy to chemotherapeutic agents (Grignani *et al.* 1998; Lin *et al.* 1998). Elucidation of the HAP of human malignancies may therefore provide a basis for novel approaches to cancer treatment. This notion is supported by the observation that inhibitors of HDAC activity are effective in the treatment of haematological malignancies such as multiple myeloma and myelodysplastic syndrome (Gore *et al.* 2001; Catley

Communicated by: Fuyuki Ishikawa
**Correspondence*: E-mail: hmano@jichi.ac.jp

DOI: 10.1111/j.1365-2443.2004.00804.x

et al. 2003). The means to establish a genome-wide HAP have been lacking, however (Gabrielli *et al.* 2002).

To characterize the HAP of any given cell or tissue type, we have now developed a new screening method, termed differential chromatin scanning (DCS). Application of the DCS method to a human gastric cancer cell line has readily identified hundreds of target fragments of histone deacetylase (HDAC), including a number of cancer-related genes such as *tumour protein p73* (*TP73*), *cell division cycle 34* (*CDC34*), and *bone morphogenetic protein 7* (*BMP7*).

With its high fidelity, DCS enables us a genome-wide screening of DNA fragments embedded in histones with differential acetylation level between a given pair of samples.

Results

The DCS method

The DCS procedure is schematically shown in Fig. 1. We verified the screening ability of this method by identifying HDAC targets in the genome of the human gastric cancer cell line MKN28. The cells were treated with trichostatin A (TSA), a specific inhibitor of HDAC activity, for use as a 'tester' sample, whereas MKN28 cells not subjected to TSA treatment were used as a 'driver'. Inhibition of HDAC activity by TSA would be expected to increase the acetylation level of target histones compared with that apparent in the driver sample. We thus attempted to rapidly isolate genomic fragments bound to histones that were acetylated only in the tester sample.

After the cross-linking of DNA to histones with the use of formaldehyde, both tester and driver cells were separately lysed and subjected to mild DNA shearing by sonication for a short period. Complexes of DNA and acetylated histones were then specifically immunoprecipitated with antibodies to acetylated histone H3, after which the DNA fragments were released from such complexes into solution.

The nonspecific binding of residual RNA was minimized by treating the DNA solution with RNase A, and the DNA fragments were then rendered blunt-ended. The DNA was digested maximally with *RsaI* to obtain fragments with a relatively uniform size of several hundred base pairs. A TAG adaptor was ligated to both ends of the DNA fragments, and subsequent PCR amplification with a TAG primer yielded amplicons with an *XmaI*/*SmaI* site at each end.

The tester DNA was then digested with *XmaI* (thereby generating cohesive ends), whereas the driver DNA was digested with *SmaI* (generating blunt ends). The tester DNA was ligated with the first subtraction adaptor (Toyota *et al.* 1999) through its cohesive ends and was then annealed to an excess amount of the driver DNA. Under this condition, DNA fragments present only in the tester sample undergo self-annealing and thereby generate a binding site for the first subtraction primer at both ends. Subsequent PCR amplification with this primer thus selectively amplified the tester-specific DNA fragments.

To exclude DNA fragments that possess endogenous (probably nonspecific) binding sites for the first subtraction primer, we digested the first subtraction products with *XmaI* and ligated the resulting molecules with the second subtraction adaptor. A second round of subtraction amplification was then performed with the second subtraction primer, yielding DNA fragments that were associated with acetylated histones specifically in the tester cells.

HDAC targets in a cancer cell line

From the products of the second round of subtraction PCR, we randomly selected 288 DNA fragments and determined their nucleotide sequence. The size of the genome fragments was ≥ 50 bp in 265 clones (mean size, ~270 bp), and these DNA sequences were subjected to subsequent analyses. The sequences were first screened with the BLAT search program (Kent 2002) against the nucleotide sequence database assembled as of July 2003 by the Genome Bioinformatics Group of the University of California at Santa Cruz (http://genome.ucsc.edu/). Among the 265 DNA fragments examined, 200 sequences showed > 95% identity to the human genome sequence and 198 of these were located either within a protein-coding gene (demonstrated or predicted) or in the vicinity (within 5 kbp) of such a gene (Table 1 and Table S1).

To verify the fidelity of the tester-specific amplification by our DCS method, we chose 34 DNA fragments and quantified their abundance in the immunoprecipitates prepared from both tester and driver cells with the antibodies to acetylated histone H3. The amount of each DNA fragment relative to that of the *glyceraldehyde-3-phosphate dehydrogenase* (*GAPDH*) gene was determined by quantitative real-time polymerase chain reaction (PCR). As shown in Table 1, selective amplification by DCS was proved to be highly reliable, with 32 out of 34 clones exhibiting tester-specific precipitation (tester/driver ratio ≥ 1.5).

TP73 as a target of HDAC

Interestingly, certain of these 32 clones corresponded to sequences either within or close to human cancer-related

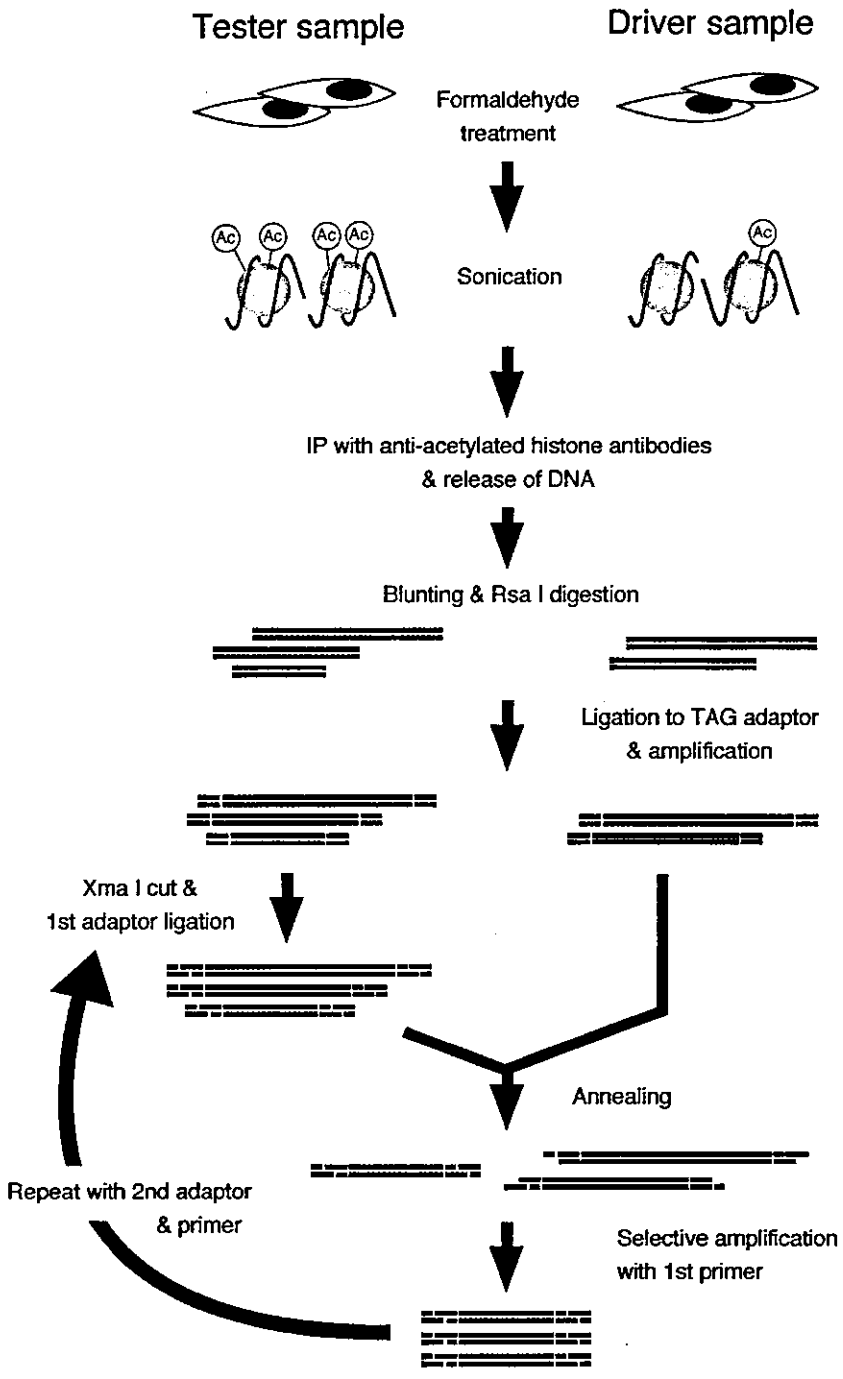

Figure 1 The DCS method. DNA fragments bound to acetylated (Ac) histones are purified by immunoprecipitation (IP) and subjected to TAG adaptor ligation (green bars) and PCR amplification. The tester DNA is then digested with *Xma*I, ligated to the first subtraction adaptor (red bars), and annealed with an excess amount of the driver DNA. Given that only the tester-specific fragments self-anneal, PCR with the first subtraction primer selectively amplifies these fragments. The products are subjected to a second round of subtraction PCR with the second subtraction adaptor and primer to ensure the fidelity of subtraction.

genes, including those for TP73, CDC34, and BMP7. One of the clones (10sBKR_D10) was, for instance, mapped to a region located ~1.2 kbp downstream of the last exon of *TP73* (GenBank accession number, NM_005427) (Fig. 2A). TP73 is a close homologue of a well-known tumour suppressor, TP53. The *TP73* gene is localized at 1p36 locus which is subject to recurrent loss of heterozygosity in various human cancers (Kaghad *et al.* 1997). An array of isoforms for TP73 are generated *in vivo* through alternative splicing mechanism of its mRNA, and those forms exert a variety of (sometimes, opposite) actions on tumour growth (Benard *et al.* 2003).

Thus, it was intriguing to examine whether, in this gastric caner cell line, histone acetylation takes place

Table 1 Genome fragments isolated through the DCS screening

Clone ID	Chromosome position	Tester/driver DNA ratio	GenBank accession no.	Tester/driver mRNA ratio	Annotation
10sBKR_A02	20q13.31	4.79	NM001719		Homo sapiens, bone morphogenetic protein 7 (osteogenic protein 1)
10sBKR_A03	11p15.5	3.66	AK056774		Homo sapiens gene duplee encoding H19
10sBKR_A08	1p36.33	14.62	NM024848		Homo sapiens hypothetical protein FLJ13941 (FLJ13941), mRNA
10sBKR_A10	1 random	11.08	NM017891		Homo sapiens cDNA FLJ20584 fis, clone KAT09532
10sBKR_B10	11p15.5	3.66	NM173584		Homo sapiens hypothetical protein MGC45840 (MGC45840), mRNA
10sBKR_C04	6p21.33	1.14*	NM000434	22.94	Homo sapiens sialidase 1 (NEU1), mRNA
10sBKR_C07	20q13.33	23.92	NA		Homo sapiens expressed gene rerty
10sBKR_C09	22q12.2	16.80	NM030758		Homo sapiens oxysterol binding protein 2 (OSBP2), mRNA
10sBKR_C12	5p15.33	16.45	NA		Homo sapiens expressed gene rujy
10sBKR_D01	12q14.1	13.64	NM000785		Homo sapiens cytochrome P450, family 27, subfamily B, polypeptide 1, mRNA
10sBKR_D02	1p36.11	14.83	NM016124		Homo sapiens Rhesus blood group, D antigen (RHD), transcript variant 1, mRNA
10sBKR_D07	11q12.2	9.19	NM017870	2.58	Homo sapiens likely ortholog of rat GRP78-binding protein (GBP)
10sBKR_D10	1p36.32	15.03	NM005427	101.13	Homo sapiens tumour protein p73 (TP73), mRNA
10sBKR_E01	2q37.3	6.15	NM016510		Homo sapiens selenocysteine lyase (SCLY), mRNA
10sBKR_E04	17q21.2	9.19	NA		Homo sapiens expressed gene risiru
10sBKR_E05	3p22.2	2.17	NM033027		Homo sapiens AXIN1 up-regulated 1 (AXUD1), mRNA
10sBKR_E06	19p13.11	4.47	NM005919		Homo sapiens MADS box transcription enhancer factor 2
10sBKR_E07	20q13.32	4.72	NM001336	0.59*	Homo sapiens cathepsin Z (CTSZ), mRNA
10sBKR_E11†	19p13.3	12.91	NM004359	2.35	Homo sapiens cell division cycle 34 (CDC34), mRNA
			NM005317	1.739×10^6	Homo sapiens granzyme M (lymphocyte met-ase 1) (GZMM), mRNA
10sBKR_F01	7q22.1	9.13	NA		Homo sapiens expressed gene smerkar
10sBKR_F04	7p13	7.16	NM001220		Homo sapiens calcium/calmodulin-dependent protein kinase II beta, mRNA
10sBKR_F03	11q13.1	15.73	NM015080		Homo sapiens neurexin 2 (NRXN2), transcript variant alpha-1, mRNA
10sBKR_F05	6p21.31	7.78	NM003322		Homo sapiens tubby like protein 1 (TULP1), mRNA
10sBKR_F06	1 random	6.15	NA		Homo sapiens expressed gene zoyplaw
10sBKR_F09	17q21.33	13.09	NM002507	1144.10	Homo sapiens nerve growth factor receptor (NGFR), mRNA
10sBKR_F11	14q32.33	1.09*	NM001519		Homo sapiens BRF1 homolog, subunit of RNA polymerase III
10sBKR_F12	11p15.5	11.31	NM003957	0.89*	Homo sapiens serine/threonine kinase 29 (STK29), mRNA
10sBKR_G02	16p13.3	21.71	NM003834		Homo sapiens regulator of G-protein signalling 11 (RGS11), mRNA
10sBKR_G06	3p25.1	461.44	NM001998	4482.23	Homo sapiens fibulin 2 (FBLN2), mRNA
10sBKR_G07	Xp22.13	1.58	NA		Homo sapiens gene CA5B encoding carbonic anhydrase VB, mitochondrial
10sBKR_G09	19q13.43	7.84	NA		Homo sapiens gene muwaru encoding hypothetical protein MGC2752
10sBKR_G11	17q25.3	4.38	NM005993		Homo sapiens tubulin-specific chaperoned (TBCD), mRNA
10sBKR_G12	1q42.13	26.91	NM024554		Homo sapiens piggyBac transposable element derived 5 (PGBD5), mRNA
10sBKR_H05	19q13.42	4.44	NM013333		Homo sapiens epsin 1 (EPN1), mRNA

*Ratio of < 1.5. †The 10sBKR_E11 was mapped to a loci of two distinct genes. NA, not assigned.

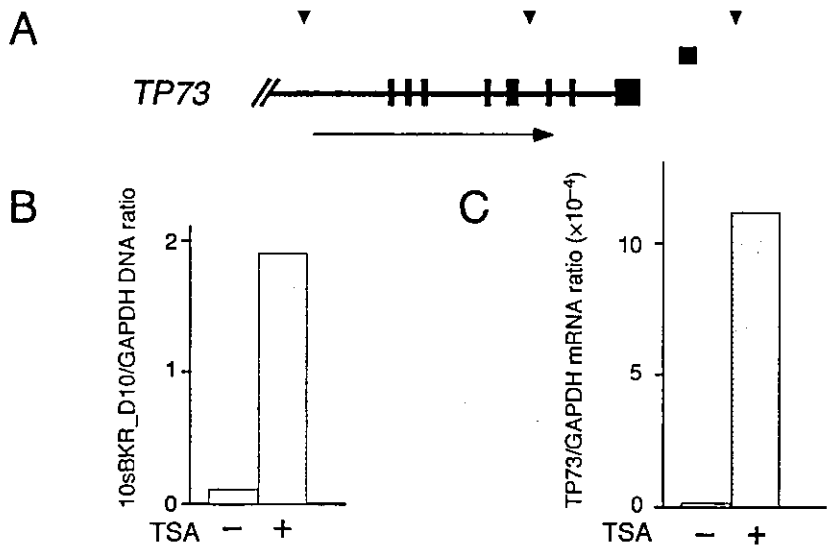

Figure 2 Identification of *TP73* as a target of HDAC activity. (A) one of the clones (10sBKR_D10; red rectangle) identified by DCS screening was mapped to chromosome 1p36.32, located ~1.2 kbp downstream of the last exon of *TP73*. Exons are denoted by black boxes, arrows indicate the direction of transcription, and green triangles depict distance markers of 5 kbp. (B) the amount of DNA corresponding to the 10sBKR_D10 sequence relative to the amount of that derived from *GAPDH* was measured by real-time PCR in the immunoprecipitates prepared with antibodies to acetylated histone H3 from MKN28 cells treated (+) or not (−) with TSA. (C) the amounts of *TP73* mRNAs relative to that of *GAPDH* mRNA in MKN28 cells treated or not with TSA were determined by real-time RT-PCR.

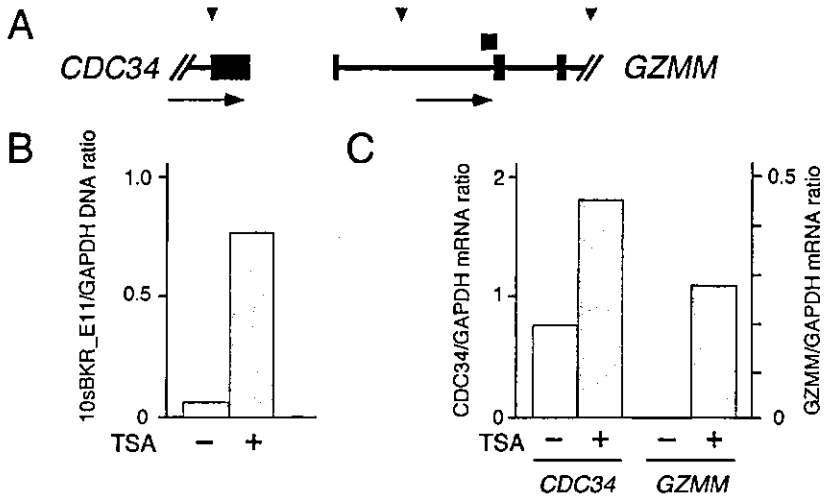

Figure 3 Identification of *CDC34* and *GZMM* as targets of HDAC activity. (A) one of the clones (10sBKR_E11; red rectangle) identified by DCS screening was mapped to chromosome 19p13.3, spanning the first intron and second exon of *GZMM* and located ~5 kbp downstream of the last exon of *CDC34*. The position of the 10sBKR_E11 in the genome is schematically shown as in Figure 2A. (B) The amount of DNA corresponding to the 10sBKR_E11 sequence relative to the amount of that derived from *GAPDH* was measured by real-time PCR. (C) the amounts of *CDC34* and *GZMM* mRNAs relative to that of *GAPDH* mRNA in MKN28 cells treated or not with TSA were determined by real-time RT-PCR.

on the *TP73* gene, and thereby regulates its transcriptional activity. Quantity of the genome fragment corresponding to the 10sBKR_D10 clone was measured by the real-time PCR method, revealing that its tester/driver ratio was 15.03 (Fig. 2B). Therefore, the extent of histone acetylation in this genome region in the tester cells was 15.03 times that in the driver cells. We also quantified the mRNA amount of *TP73* genes in both tester and driver cells by real-time reverse transcription (RT)-PCR method. As demonstrated in Fig. 2C, acetylation of the 10sBKR_D10 region was accompanied with a profound activation in the transcription level of *TP73*. These lines of evidence indicate that the *TP73* gene is epigenetically regulated by HDAC in a gastric cancer cell line.

Other targets of HDAC

One of our clones (10sBKR_E11) was mapped to a region which contains the first intron and the second exon of the *granzyme M* (*GZMM*) gene (GenBank accession number, NM_005317), and is also located ~5 kbp downstream of the last exon of *CDC34* (GenBank accession number, NM_004359) (Fig. 3A). Therefore, it was possible that histone acetylation of this region may affect the transcription of both genes. We first quantified the tester/driver ratio for the relative amount of the 10sBKR_E11 fragment in the acetylated histone H3 immunoprecipitates. As shown in Fig. 3B, histone at this region was more heavily acetylated in the tester cells compared to the driver cells.

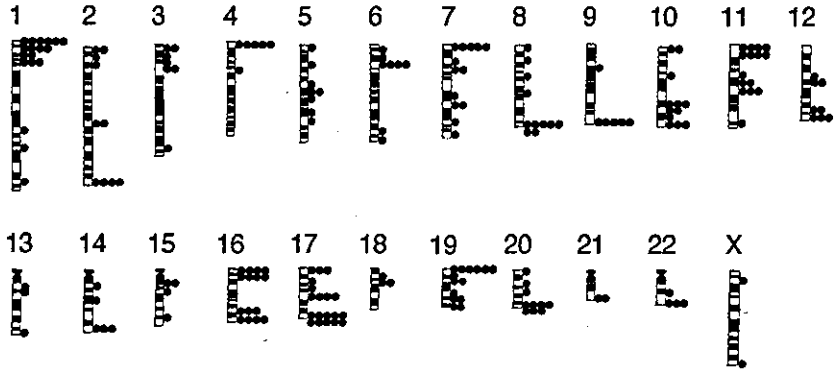

Figure 4 Chromosomal distribution of HDAC targets. The genome fragments (red dots) isolated by the DCS method were mapped to human chromosomes. The MKN28 cell line was established from a female and therefore does not contain the Y chromosome.

To examine how histone acetylation in this area affected the transcriptional activity of *CDC34* and *GZMM*, we measured the amounts of the corresponding mRNAs in both tester and driver cells by real-time RT-PCR. As demonstrated in Fig. 3C, histone acetylation up-regulated the expression of both *CDC34* and *GZMM*, although the extent of gene activation was greater for the latter than for the former. These data suggest that the histone acetylation status of a given genomic locus may regulate the transcription of distinct genes in a differential manner. Our data do not exclude the possibility, however, that histones in other regions of *CDC34* or *GZMM* were also acetylated in a TSA-dependent manner and that these regions possess an enhancer activity greater than that of the 10sBKR_E11 region.

CDC34 encodes a ubiquitin E2 ligase, and its activity is essential for G_1- to S-phase transition in cell cycle (Pagano *et al.* 1995). CDC34 catalyses ubiquitination of the inhibitors for cyclin-dependent kinases (CDKI) which are thereby subject to proteolysis. Therefore, it was interesting to find that *CDC34* gene was a target of HDAC in a cancer cell line.

We selected seven additional genes for quantification of the relative amounts of their mRNAs. Of the total of nine genes examined, seven genes were preferentially expressed in the TSA-treated cells compared with the non-treated cells (tester/driver ratio ≥ 1.5) (Table 1), consistent with the notion that histone acetylation results in the recruitment of transcription factors and consequent transcriptional activation.

Among the 136 clones that mapped within or close to known genes, only eight pairs were assigned to the same genes. Given the high fidelity of our DCS screening, these data suggest that HDAC complexes simultaneously act at hundreds of independent target histones within the same cell. To visualize directly the genome-wide HAP, we mapped to human chromosome figures our genomic clones whose chromosomal positions are known (Fig. 4). The HDAC targets were distributed widely throughout the human genome, although some 'hot spots' for deacetylation were apparent. Seven of our DCS clones were, for instance, mapped to the same 8q24.3 locus. Detailed mapping indicated that all these clones were located within a region spanning 2.5 Mbp. Therefore, it may be possible that regional alterations of chromatin structure lead to a coordinated transcriptional regulation of genes within the affected region.

Discussion

We have developed a rapid and highly reliable method for the selective isolation of genome fragments bound to histones whose level of acetylation differs between samples. Application of the DCS method to MKN28 cell line has readily revealed that many cell proliferation-associated genes are epigenetically regulated in cancer cells, including those for BMP7, TP73, GBP, AXUD1, CDC34, NGFR, RGS11 (see Table 1), JUNB, SIDEA, and MAFG (see Table S1).

Recently, similar genome-wide comparison methods have been developed by coupling chromatin immunoprecipitation (ChIP) with either comparative genomic hybridization (CGH) or analysis with a DNA microarray containing thousands of CpG islands or promoter fragments (Weinmann *et al.* 2002; Ballestar *et al.* 2003; Odom *et al.* 2004). However, the CGH analysis is not able to provide information on the precise genome sequences differentially immunoprecipitated, and the CpG array can only identify genome fragments that contain CpG islands. Interestingly, the majority of our clones identified by DCS do not satisfy the criteria of CpG islands (for instance, none of the clones corresponding to *TP73*, *CDC34-GZMM*, or *BMP7* was rich in the CpG sequence). Additionally, screening of whole promoter fragments is still hampered by the incomplete

human genome annotation and by its inability to investigate the enhancer regions. The emerging number of non-coding RNAs also makes it difficult to precisely map 'promoter' sequences in human genome. Therefore, for the identification of HAT-HDAC targets, our DCS method is likely to provide a non-biased means of genome-wide screening.

Although the study in this manuscript has revealed only the HDAC targets in a cancer cell line, the DCS method should be useful to directly identify HAT targets as well. For instance, a DCS comparison between chemotherapy-sensitive cancer cells and chemotherapy-resistant ones would enable to isolate genes whose associated histones are acetylated in a chemoresistance-dependent manner. In addition, DCS method can be applied to other non-malignant cells/tissues. We could indeed identify HDAC targets in cardiomyocytes (R.K., personal communication).

Furthermore, modification of our method may open up the possibility of a general ChIP-based subtraction strategy. For instance, DCS performed with antibodies to a given histone- or DNA-binding protein, rather than with antibodies to acetylated histone, should readily allow the isolation of target genomic fragments differentially recognized by such proteins.

Experimental procedures

Cell culture

MKN28 cells were obtained from the Japanese Collection of Research Bioresources (http://cellbank.nihs.go.jp/) and were maintained in RPMI 1640 medium (Invitrogen) supplemented with 10% foetal bovine serum. For preparation of the tester sample, cells were incubated for 24 h with 300 nM TSA (Wako, Tokyo, Japan). Satoh *et al.* (2002) have already demonstrated that, in MKN28 cells, histones of several genes become acetylated in a TSA-dependent manner.

The DCS method

Both formaldehyde treatment of cells and immunoprecipitation with antibodies to acetylated histone H3 were performed with the ChIP assay kit (#06-599; Upstate Biotechnology). Sonication of DNA was performed in one 10-s pulse to minimize the extent of DNA shearing (the resulting DNA fragments should migrate as a smear corresponding to a size of 10–20 kbp in gel electrophoresis). Immunoprecipitation was achieved with protein A-Sepharose beads (Sigma) suspended in TE (10 mM Tris-HCl, 7.4, and 0.1 mM EDTA) containing 0.2 mg/mL tRNA (Roche Biochemicals) and 0.5 mg/mL bovine serum albumin (New England Biolabs), instead of with the protein A beads supplied with the kit. DNA fragments were recovered from the immunoprecipitates and treated with RNase A to remove residual RNA. The DNA fragments were digested with *Rsa*I (New England Biolabs), and ligated to the TAG adaptor (5′-CCACCGCCATCCGAGCCTTTCTGCCCGGG-3′/3′-GAAAGACGGGCCC-5′). After PCR amplification with the TAG primer (5′-CCACCGCCATCCGAGCCTTTCTGC-3′), the tester DNA and the driver DNA were digested with *Xma*I and *Sma*I, respectively. Only the tester DNA (0.5 μg) was ligated to the first subtraction adaptor (5′-GTGAGGGTCGGATCT-GGCTGGCTC-3′/3′-CGACCGAGGGCC-5′). The adaptor-ligated tester DNA was then annealed with 40 μg of the driver DNA at 67 °C for 20–24 h before PCR amplification 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 XmaI followed by a second round of subtraction PCR with the second subtraction adaptor (5′-GTTAGCGGACACAGGGCGGGTCAC-3′/3′-GCCCAG-TGGGCC-5′) and second subtraction primer (5′-GTTAGCG-GACACAGGGCGGGTCAC-3′). The final products were digested with *Xma*I and ligated into pBlueScript (Stratagene) for nucleotide sequencing. The detailed protocol for the DCS method is available as a supplemental protocol online.

DNA quantification

Genome fragments immunoprecipitated by antibodies to acetylated histone H3 were subjected to PCR with a QuantiTect SYBR Green PCR Kit (Qiagen). The amplification protocol comprised incubations at 94 °C for 15 s, 62 °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 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 corresponding to the *GAPDH* gene and genome fragments of interest were used to calculate the abundance of the latter fragments relative to that of the former. The oligonucleotide primers for PCR were 5′-TCGGTGCGTGCCCAGTTGAACC-3′ and 5′-ATGCGGCTGACTGTCGAACAGGAG-3′ for *GAPDH*, 5′-ACCAGTCGCTGCTTTTAAATAAGG-3′ and 5′-GCTAGGAGCTTCCCCTACTAACAT-3′ for 10sBKR_D10, and 5′-CTGATGTGCGTTTTGAAGGACT-3′ and 5′-GAAACTTTTAGGCTGATACGTGTG-3′ for 10sBKR_E11.

mRNA quantification

Total RNA was prepared from the tester or driver cells with an RNeasy Mini column (Qiagen), treated with RNase-free DNase (Qiagen), and subjected to reverse transcription with PowerScript reverse transcriptase (BD Biosciences Clontech) with an oligo(dT) primer. Portions of the resulting cDNAs were subjected to PCR with a QuantiTect SYBR Green PCR Kit. The amplification protocol comprised incubations at 94 °C for 15 s, 65 °C for 30 s, and 72 °C for 60 s. The annealing temperature of PCR was changed to 60 °C for the cDNA of *CDC34* and 66 °C for that of *GZMM*. The oligonucleotide primers for PCR were 5′-AGAACATCATCCCTGCCTCTACT-3′ and 5′-

ATATTTGGCAGGTTTTTCTAGACG-3′ for *GAPDH*, 5′-GACGAGGACACGTACTACCTTCA-3′ and 5′-GTAGGTGACTCGGCCTCTGTAG-3′ for *TP73*, 5′-AAGATGTGGCACCCTAACATCTAC-3′ and 5′-AGGGAGATCACACTCAGGAGAAT-3′ for *CDC34*, and 5′-ATGTGTAACAACAGCCGCTTCT-3′ and 5′-CTTGAAGATGTCAGTGCAGACC-3′ for *GZMM*.

Acknowledgements

We thank Dr Kohei Miyazono for his helpful suggestions. This work was supported in part by a Grant-in-Aid for Research on the Second-Term Comprehensive 10-Year Strategy for Cancer Control from the Ministry of Health, Labor, and Welfare of Japan; a grant from Mitsubishi Pharma Research Foundation; a grant from Takeda Science Foundation; and a grant from Sankyo Foundation of Life Science.

Supplementary material

The following supplementary material is available from: http://www.blackwellpublishing.com/products/journals/suppmat/GTC/GTC804/GTC804sm.htm. **Appendix S1** The detailed protocol of the DCS method. **Table S1** All genome fragments isolated through the DCS screening.

References

Ballestar, E., Paz, M.F., Valle, L., *et al.* (2003) Methyl-CpG binding proteins identify novel sites of epigenetic inactivation in human cancer. *EMBO J.* **22**, 6335–6345.

Baylin, S.B. & Herman, J.G. (2000) DNA hypermethylation in tumorigenesis: epigenetics joins genetics. *Trends Genet.* **16**, 168–174.

Benard, J., Douc-Rasy, S. & Ahomadegbe, J.C. (2003) TP53 family members and human cancers. *Hum. Mutat.* **21**, 182–191.

Burgess-Beusse, B., Farrell, C., Gaszner, M., *et al.* (2002) The insulation of genes from external enhancers and silencing chromatin. *Proc. Natl. Acad. Sci. USA* **99** (Suppl. 4), 16433–16437.

Carrozza, M.J., Utley, R.T., Workman, J.L. & Cote, J. (2003) The diverse functions of histone acetyltransferase complexes. *Trends Genet.* **19**, 321–329.

Catley, L., Weisberg, E., Tai, Y.T., *et al.* (2003) NVP-LAQ824 is a potent novel histone deacetylase inhibitor with significant activity against multiple myeloma. *Blood* **102**, 2615–2622.

Gabrielli, B.G., Johnstone, R.W. & Saunders, N.A. (2002) Identifying molecular targets mediating the anticancer activity of histone deacetylase inhibitors: a work in progress. *Curr. Cancer Drug Targets* **2**, 337–353.

Gore, S.D., Weng, L.J., Zhai, S., *et al.* (2001) Impact of the putative differentiating agent sodium phenylbutyrate on myelodysplastic syndromes and acute myeloid leukemia. *Clin. Cancer Res.* **7**, 2330–2339.

Grignani, F., De Matteis, S., Nervi, C., *et al.* (1998) Fusion proteins of the retinoic acid receptor-alpha recruit histone deacetylase in promyelocytic leukaemia. *Nature* **391**, 815–818.

Kaghad, M., Bonnet, H., Yang, A., *et al.* (1997) Monoallelically expressed gene related to p53 at 1p36, a region frequently deleted in neuroblastoma and other human cancers. *Cell* **90**, 809–819.

Kent, W.J. (2002) BLAT—the BLAST-like alignment tool. *Genome Res.* **12**, 656–664.

Kouzarides, T. (1999) Histone acetylases and deacetylases in cell proliferation. *Curr. Opin. Genet. Dev.* **9**, 40–48.

van Leeuwen, F. & Gottschling, D.E. (2002) Genome-wide histone modifications: gaining specificity by preventing promiscuity. *Curr. Opin. Cell Biol.* **14**, 756–762.

Lin, R.J., Nagy, L., Inoue, S., Shao, W., Miller, W.H. Jr & Evans, R.M. (1998) Role of the histone deacetylase complex in acute promyelocytic leukaemia. *Nature* **391**, 811–814.

Odom, D.T., Zizlsperger, N., Gordon, D.B., *et al.* (2004) Control of pancreas and liver gene expression by HNF transcription factors. *Science* **303**, 1378–1381.

Pagano, M., Tam, S.W., Theodoras, A.M., *et al.* (1995) Role of the ubiquitin-proteasome pathway in regulating abundance of the cyclin-dependent kinase inhibitor p27. *Science* **269**, 682–685.

Satoh, A., Toyota, M., Itoh, F., *et al.* (2002) DNA methylation and histone deacetylation associated with silencing DAP kinase gene expression in colorectal and gastric cancers. *Br. J. Cancer* **86**, 1817–1823.

Toyota, M., Ho, C., Ahuja, N., *et al.* (1999) Identification of differentially methylated sequences in colorectal cancer by methylated CpG island amplification. *Cancer Res.* **59**, 2307–2312.

Verdin, E., Dequiedt, F. & Kasler, H.G. (2003) Class II histone deacetylases: versatile regulators. *Trends Genet.* **19**, 286–293.

Waterborg, J.H. (2002) Dynamics of histone acetylation in vivo. A function for acetylation turnover? *Biochem. Cell Biol.* **80**, 363–378.

Weinmann, A.S., Yan, P.S., Oberley, M.J., Huang, T.H. & Farnham, P.J. (2002) Isolating human transcription factor targets by coupling chromatin immunoprecipitation and CpG island microarray analysis. *Genes Dev.* **16**, 235–244.

Yasui, W., Oue, N., Ono, S., Mitani, Y., Ito, R. & Nakayama, H. (2003) Histone acetylation and gastrointestinal carcinogenesis. *Ann. NY. Acad. Sci.* **983**, 220–231.

Received: 21 August 2004
Accepted: 16 September 2004

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Biochemical and Biophysical Research Communications 320 (2004) 1328–1336
BBRC
www.elsevier.com/locate/ybbrc

Cardiac function-related gene expression profiles in human atrial myocytes ☆

Ruri Ohki-Kaneda [a,b], Jun Ohashi [c], Keiji Yamamoto [b], Shuichi Ueno [a,b], Jun Ota [a,d], Young Lim Choi [a], Koji Koinuma [a], Yoshihiro Yamashita [a], Yoshio Misawa [e], Katsuo Fuse [e], Uichi Ikeda [b], Kazuyuki Shimada [b], Hiroyuki Mano [a,d,*]

[a] *Division of Functional Genomics, Jichi Medical School, 3311-1 Yakushiji, Kawachigun, Tochigi 329-0498, Japan*
[b] *Division of Cardiology, Jichi Medical School, 3311-1 Yakushiji, Kawachigun, Tochigi 329-0498, Japan*
[c] *Department of Human Genetics, Graduate School of Medicine, University of Tokyo, Tokyo 113-0033, Japan*
[d] *CREST, JST, Saitama 332-0012, Japan*
[e] *Division of Cardiovascular Surgery, Jichi Medical School, 3311-1 Yakushiji, Kawachigun, Tochigi 329-0498, Japan*

Received 6 May 2004
Available online 2 July 2004

Abstract

To obtain insights into the molecular pathogenesis of heart failure in humans, we have analyzed the expression profiles of >12,000 genes in a total of 17 human specimens of right atrial myocytes. From this large data set, we here tried to identify gene clusters, expression level of which is correlated precisely with clinical parameter values of cardiac function. We could reveal that cardiac myocytes with normal sinus rhythm were clearly differentiated, in the point of view of gene expression, from those with atrial fibrillation. Further, an expression profile-based prediction of arrhythmia by a newly developed "weighted-distance method" could efficiently diagnose our samples. We could even construct calculation formulae for the values of left ventricular ejection fraction based on the expression level of selected genes. To our best knowledge, this is the first report to indicate that pumping ability of heart can be predicted by any measures of atrium.

Keywords: DNA microarray; Heart failure; Ejection fraction; Correspondence analysis

The pumping failure of heart, or heart failure, is induced by a variety of pathological conditions, such as ischemic heart diseases, arrhythmias, and sustained pressure- or volume-overload onto heart [1]. It is still highly difficult to restore the contractile ability of cardiac myocytes in once failed heart, and, therefore, heart failure is still one of the main causes for human death [2]. Unfortunately, little is still understood for the molecular mechanisms underlying why and how heart failure is developed and becomes irreversible.

Although cardiac myocytes in culture may be a useful means to analyze the intracellular signaling pathways responsible for hypertrophic changes in heart, such in vitro systems cannot efficiently simulate the long-term changes of myocytes found in failed heart. To decipher the molecular pathology of heart failure, therefore, it would be indispensable to directly investigate the human specimens of failed heart or to analyze myocytes in model animals for heart failure.

☆ *Abbreviations:* Af, atrial fibrillation; EF, ejection fraction; TR, tricuspid valve regurgitation; MR, mitral valve regurgitation; LAD, left atrial diameter; RT-PCR, real-time reverse transcription-polymerase chain reaction; GAPDH, glyceraldehyde-3-phosphate dehydrogenase.

* Corresponding author. Fax: +81-285-44-7322.
E-mail address: hmano@jichi.ac.jp (H. Mano).

0006-291X/$ - see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.bbrc.2004.06.090

Toward this goal, we have recently collected human specimens of right atrial myocytes from patients with various heart disorders, and obtained the expression profiles of >12,000 genes in the samples by using high-density oligonucleotide microarrays (R.O. et al., submitted). Since the patients in the analysis could be classified into two major groups, i.e., individuals with either normal sinus rhythm or an arrhythmia, chronic atrial fibrillation (Af), we first tried to identify gene sets, expression of which were associated with the presence/absence of chronic Af in vivo. A total of 33 genes were shown to be induced (>1.5-fold) in Af samples compared to those with normal sinus rhythm, while 63 genes were to be suppressed (<0.5-fold) in the Af-ones.

However, such analyses did not fully address a fundamental question as to whether atrial myocytes with Af have a distinct gene expression profile, or "molecular signature" [3], to those with normal sinus rhythm. In other words, with regard to gene expression profiles, are the myocytes with Af distinct from those with normal rhythm?

In this paper, by applying sophisticated statistical methods, we have tried to address these issues. Furthermore, we here tested whether it was possible to identify gene sets, expression level of which was precisely correlated with detailed clinical parameter values of cardiac function. The results were highly successful, and we could even construct "prediction formulae for the values of left ventricular ejection fraction (EF)" based on the expression level of selected genes. To our best knowledge, this is the first report to prove that pumping ability of heart can be predicted by any measures of atrium.

Materials and methods

Microarray analysis. Detailed information on the patients registered for this study and on the microarray experiments will be reported elsewhere (R.O. et al., submitted). Left ventricular EF values in all patients was calculated by dividing the stroke volume (end-diastolic volume minus end-systolic volume) by the end-diastolic volume (both obtained angiographically). From the patients who underwent cardiac surgery in Jichi Medical School, right atrial appendages were obtained with written informed consent. Total RNA was extracted from the specimens with the use of RNAzol B (Tel-Test, Friendswood, TX), and a portion (20 μg) of the RNA was converted to double-stranded cDNA by the SuperScript Choice System (Life Technologies, Gaithersburg, MD). Biotin-labeled cRNA was prepared from the resulting cDNA with the use of ENZO BioArray RNA labeling kit (ENZO Diagnostics, Farmingdale, NY), and hybridized with the GeneChip Human U95Av2 array (Affymetrix, Santa Clara, CA) harboring the oligonucleotides corresponding to >12,000 human genes. Detection of the hybridized cRNAs was performed with the GeneChip instrument system according to the manufacturer's protocol. The fluorescence intensity of each gene was normalized relative to the median fluorescence value for all genes in each array hybridization.

Statistical analysis. In the comparison of normal sinus rhythm- vs Af-specimens, t statistic and the effect size (difference in the mean of expression levels between the two classes) [4] were calculated for each gene. When the gene showed $|t| > 4.073$ (corresponding to a significance level of 0.001 in t test with 15 degrees of freedom) and |effect size| > 0.5 arbitrary units (U), the difference in expression level between the two classes was regarded as statistically significant. Correspondence analysis [5] was carried out for all genes showing the significant difference by using the ViSta computer software (http://www.visualstats.org/). Each sample was plotted in three dimensions based on the coordinates of the three major dimensions from correspondence analysis.

To examine whether these informative genes could predict the class of the present samples, we have here developed a novel class prediction means, "weighted-distance method," with the points of dimensions obtained from correspondence analysis. There were a total of 17 samples: 10 specimens with sinus rhythm and 7 with Af arrhythmia. A correspondence analysis was conducted once for all 17 samples, and the class prediction was performed for each sample (say sample X) under the condition that the classes of only remaining 16 samples were known. The weighted distance from sample X to sample Y ($Y \neq X$) is defined as $D = \sqrt{\sum_{i=1}^{k} [v_i (d_i^X - d_i^Y)^2]}$, where k is the number of dimensions to be considered, v_i indicates the contribution of the ith dimension, and d_i^X and d_i^Y represent points of the ith dimension for sample X and sample Y, respectively. Because v_1 exceeded 70% in our analysis (i.e., the cumulative contribution to the 1st dimension was very large), k was set to be 1 in this study (i.e., $D = \sqrt{v_1 (d_1^X - d_1^Y)^2}$). In general, k should be determined by the cumulative contribution to the kth dimension. Let D_S and D_A be the averages of weighted distance (D) from sample X to sinus rhythm samples and from sample X to the Af samples, respectively. When $D_S/(D_S + D_A)$ was below a threshold value, T, the sample X was assigned a "sinus rhythm" class, and when $D_A/(D_S + D_A)$ was below T, the sample X was assigned an "Af" one. In this study, T was set to be 0.4.

Correlation to the clinical grade of tricuspid valve regurgitation (TR) or mitral valve regurgitation (MR) was calculated for the expression level of each gene by a paired Spearman's rank test, and was considered as statistically significant when $P < 0.001$. Similarly, Pearson's correlation coefficient (r) for EF (in percents) or left atrial diameter (LAD) (in millimeters) was examined for the expression level of each gene.

For each of the EF-linked genes identified above, we conducted the simple regression analysis to generate an EF-calculation formula "$y = Ax + B$," where y is EF (%) and x is the expression level (U) of each gene.

All raw microarray data as well as the details of the genes shown in the figures are available as supplementary information at the web site of *Biochem. Biophys. Res. Commun.*

Real-time reverse transcription-polymerase chain reaction (RT-PCR). Portions of nonamplified cDNA were subjected to PCR with a QuantiTect SYBR Green PCR Kit (Qiagen, Valencia, CA). The amplification protocol comprised incubations at 94 °C for 15 s, 60 °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 7700 sequence detection system (PE Applied Biosystems, Foster City, CA), thereby allowing determination of the threshold cycle (C_T) at which exponential amplification of PCR products begins. The C_T values for cDNAs corresponding to the glyceraldehyde-3-phosphate dehydrogenase (GAPDH) and *Calpactin I* genes were used to calculate the abundance of *Calpactin I* mRNA relative to that of *GAPDH* mRNA. The oligonucleotide primers for PCR were 5′-GTCAGTGGTG GACCTGACCT-3′ and 5′-TGAGCTTGACAAAGTGG TCG-3′ for *GAPDH* cDNA and 5′-TCTGGCTGTGGACAAAATAATG-3′ and 5′-TTTCCCTTCTGCTTCATGTGTA-3′ for *Calpactin I* cDNA.

Results

Isolation of genes linked to arrhythmia

Among the 17 samples analyzed, 7 were obtained from the patients with chronic, persistent Af and the rest

Table 1
Characteristics of the study subjects

Patient ID	Age	Sex	Disease	MR grade	TR grade	LAD (mm)	Rhythm	EF (%)
1	44	M	AR	0	1	39.7	Sinus	49
2	70	F	MR	3	1	50.1	Sinus	48
3	75	M	IHD	1	1	39	Sinus	61
4	65	M	AS	1	0	45.1	Sinus	71
5	60	M	AS	2	0	39.7	Sinus	33
6	51	F	MS	0	3	71	Af	78
7	55	M	MR	4	2	69.4	Af	55
9	64	M	MR	4	2	77	Af	68
10	74	F	MSR, ASR	3	2	61.1	Af	66
11	75	M	ASR, MR	3	2	59	Af	37
12	64	F	MR	3	3.5	48	Sinus	65
13	64	F	MR, AR	4	3.5	42	Sinus	80
14	14	M	ASD	0	2	32.1	Sinus	67
15	59	F	MS, AR	2	4	76.7	Af	70
16	69	F	MSR, AR	3	4	54	Af	57.5
17	63	M	AR, MR	2	3	52	Sinus	50
18	37	F	ASD	2	0	35	Sinus	60

M, male; F, female; AR, aortic regurgitation; IHD, ischemic heart disease; AS, aortic stenosis; MS, mitral stenosis; MSR, mitral stenosis/regurgitation; ASR, aortic stenosis/regurgitation; ASD, atrial septal defect.

10 were from those with normal sinus rhythm (Table 1). Identification of any genes constantly activated or inactivated in Af samples would shed new light on the pathophysiology of this most prevalent arrhythmia [6]. First, identified were the genes whose expression levels were statistically different (Student's t test, $P<0.001$) between the two classes, i.e., normal sinus rhythm and Af. However, expression levels of many such genes were very low throughout the samples, making the reliability of "arrhythmia-dependence" obscure. Therefore, we further applied another "selection window" that the effect size should be >0.5 U between the two classes, leading to the identification of 11 arrhythmia-dependent genes, expression profile of which should most effectively contrast Af status from normal sinus rhythm.

A dendrogram or "gene tree" is demonstrated in Fig. 1A for these genes including those encoding for sarcoglycan ε (SGCE; GenBank Accession No. NM_003919), zinc finger protein 103 (ZFP103; NM_005667), and hepatic leukemia factor (HLF; NM_002126). Here in the tree, genes with similar expression pattern across the samples were placed near each other. HLF is a helix–loop–helix type transcriptional factor that functions to protect cells from apoptosis [7]. Decreased expression of *HLF* in the myocytes with Af arrhythmia may reflect the apoptotic change found in atrial remodeling.

Expression profile-based prediction of Af

In the point of view of gene expression profile, are the myocytes with Af arrhythmia different from those with normal sinus rhythm, and, if so, how different? We tried to address this issue by the correspondence analysis to extract three major dimensions from the expression patterns of the 11 arrhythmia-dependent genes. By the calculated coordinates for each sample in these dimensions, the patient samples were projected into the virtual space with the three dimensions. Surprisingly, as shown in Fig. 1B, all samples with Af were clustered in an area clearly far from that for the samples with sinus rhythm. It is, therefore, likely that all Af samples have a common "gene expression signature" which is distinct from that for sinus rhythm samples.

The clear separation of Af and sinus rhythm specimens in Fig. 1B also indicates a possibility of an expression profile-based prediction for Af. As demonstrated in Fig. 1B, class-separation was most prominent in the 1st dimension. Indeed, the contribution ratio of each dimension calculated by the correspondence analysis was 72.74%, 9.5%, and 6.29% for the 1st, 2nd, and 3rd dimensions, respectively. Therefore, we tried to conduct the class prediction (i.e., Af or sinus rhythm) for each sample by using the coordinates for the 1st dimension. The relative distance of a given sample to the Af or sinus rhythm group (excluding the sample for prediction) was calculated, and the sample was assigned a class when the relative distance to the class was $\leqslant 0.4$. As shown in Table 2, the class prediction based on the correspondence analysis with the 11 genes in Fig. 1A was successful for all cases.

To accurately measure the prediction power of our weighted distance method, we conducted a cross-validation trial (i.e., "drop-one-out" format) for the diagnosis of normal sinus rhythm or chronic Af class. To predict the class of sample X, "arrhythmia-associated genes" were extracted from the comparison of remaining 16 samples according to the criteria used in Fig. 1A ($P<0.001$ in Student's t test, and |effect size|>0.5 U). Correspondence analysis was carried out for the expression profiles of the arrhythmia-associated genes, and

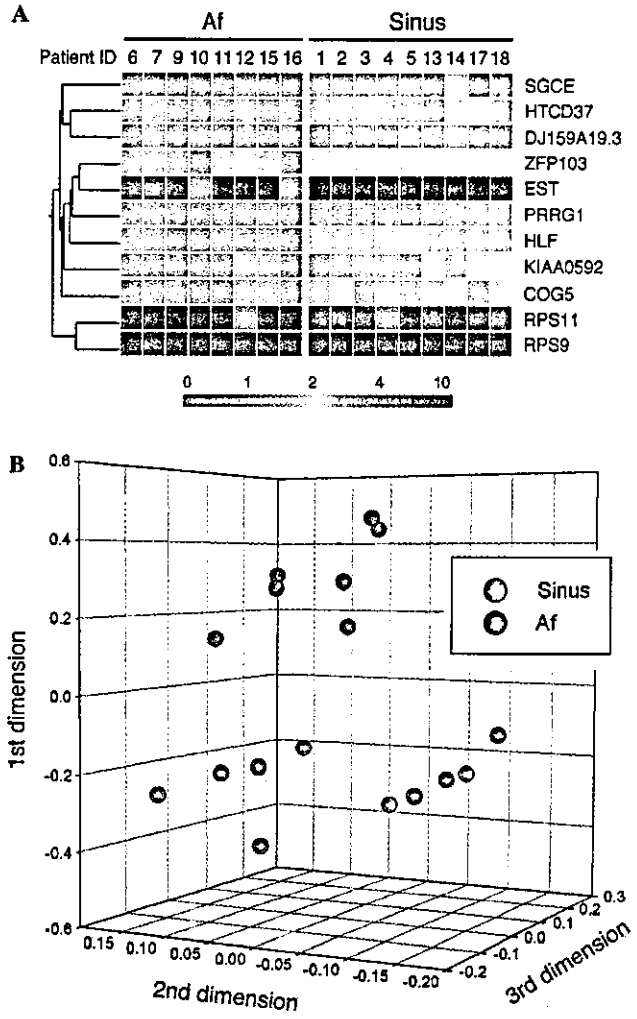

Fig. 1. Correspondence analysis of the data set for the differentiation between Af vs normal sinus rhythm. (A) The expression profiles of 11 arrhythmia-dependent genes are demonstrated as a dendrogram, shown color-coded as indicated by the scale at the bottom. Each column represents a separate patient sample as indicated by the ID number at the top, and each row corresponds to a single gene. The gene symbols are shown at the right. (B) The correspondence analysis of the arrhythmia-dependent genes has identified three major dimensions in their expression profiles. Projecting the samples into a virtual space with these three dimensions demonstrates that specimens of heart with normal sinus rhythm (Sinus) are placed apart from those with Af (Af).

was used to calculate the relative weighted distance of the "dropped" sample X to either normal sinus rhythm or Af class. A clinically matched prediction was obtained for 13 (76.5%) out of 17 cases (not shown). Therefore, even in a cross-validation assay, the weighted distance method could achieve a high accuracy.

Genes linked to the clinical grades of MR or TR

Given the efficient isolation of genes tightly linked to the occurrence of Af, our expression data set could be also used for the characterization of other clinical parameters of heart. For instance, echocardiographic studies can assess various aspects of cardiac function, such as the severity of TR, MR, and atrium dilation (as LAD), and the contraction ability of heart (as left ventricular EF).

We first tried to identify the genes whose expression level was linked to the stages of MR or TR. In the clinical settings, the severity of MR and TR is represented as widely accepted clinical grades, from normal "grade 0" to the highly damaged "grade 4" [8,9]. The MR or TR grade indicates the severity of volume-overload onto left or right atrium, respectively.

Spearman's rank test has identified a total of nine genes, expression of which was significantly linked to the clinical grade of MR (Fig. 2A). All genes thus identified had a reverse correlation toward the grade of MR; in other words, expression of these genes became suppressed along with the stage progression of MR. Intriguingly, some protein products in the gene list had been shown to be directly involved in the regulation of apoptosis, such as the nucleoside diphosphate kinase type 6 (NM23-H6; GenBank Accession No. NM_005793), CUG triplet repeat RNA binding protein 2 (CUGBP2; NM_006561), and caspase-8- and FADD-like apoptosis regulator (CFLAR; NM_003879).

Similar statistical analysis has also identified a total of 37 genes, expression of which was significantly linked to the clinical grades of TR (Fig. 2B). Among them, the highest correlation coefficient (forward correlation) was found with the genes for general transcription factor IIIC, polypeptide 1, α (GTF3C1; GenBank Accession No. NM_001520), and thioredoxin 2 (TXN2; NM_012473), and the lowest coefficient (reverse correlation) was with those for general transcription factor IIA2 (GTF2A2; NM_004492) and NF-E2-like basic leucine zipper transcriptional activator (NFE2L2; NM_006164).

Genes linked to the LAD size

Next, among the clinical parameters for cardiac function, we searched for genes whose expression was significantly correlated to continuous variates, such as LAD or EF. Pearson's correlation coefficient to the size of LAD was examined for the expression level of each gene, and we could identify a total of five genes with statistically significant correlation ($P<0.001$) to LAD. In Fig. 3 demonstrated are the expression levels of such five genes among the samples as a dendrogram format. Expression of all these genes became suppressed as the size of LAD increased. Interestingly, one of such genes encodes for SGCE that links cytoskeleton proteins to extracellular matrix. SGCE is a member of dystrophin–glycoprotein complex which is essential in the stability of muscle fiber membrane [10]. Mutations in the *SGCE* gene are responsible for the familial myoclonus–dystonia syndrome [11]. Loss of expression in *SGCE* may thus be involved in the

Table 2
Prediction of arrhythmia by the weighted-distance method

Patient ID	Clinical diagnosis	Distance to normal sinus rhythm	Distance to Af	Prediction
1	Sinus	0.144	0.856	Sinus
2	Sinus	0.128	0.872	Sinus
3	Sinus	0.125	0.875	Sinus
4	Sinus	0.139	0.861	Sinus
5	Sinus	0.205	0.795	Sinus
6	Af	0.731	0.269	Af
7	Af	0.834	0.166	Af
9	Af	0.840	0.160	Af
10	Af	0.786	0.214	Af
11	Af	0.839	0.161	Af
12	Sinus	0.132	0.868	Sinus
13	Sinus	0.153	0.847	Sinus
14	Sinus	0.289	0.711	Sinus
15	Af	0.668	0.332	Af
16	Af	0.802	0.198	Af
17	Sinus	0.127	0.873	Sinus
18	Sinus	0.250	0.750	Sinus

The relative weighted-distance to the normal sinus rhythm group or to the Af group was calculated as $D_S/(D_S+D_A)$ or $D_A/(D_S+D_A)$, respectively (see Materials and methods).

injured muscle integrity in dilated left atrium. It should be noted that expression of *SGCE* was also linked to the presence/absence of Af (Fig. 1A). Therefore, activation of *SGCE* gene may reflect the pathological changes associated with atrial remodeling.

Calculation of EF values by gene expression profiles

One of our goals in this study was to test the possibility of profile-based prediction of cardiac pump function or patients' prognosis. There have been no reports for such investigation, and it is yet unclear to what extent atrial myocytes reflect the overall function of heart.

To address these issues, here we have tried to calculate the exact digits of left ventricular EF (in percents) from the gene expression profiles in right atrium myocytes. First, Pearson's correlation coefficient to EF was calculated for each gene, and we could isolate 15 genes whose expression was correlated to EF ($P<0.001$) (Fig. 4A). Among them, only the expression of fragile histidine triad gene (FHIT; GenBank Accession No. NM_002012) was induced along with the increase of EF, while the expression of the rest genes had a reverse correlation to the EF level.

A simple regression analysis was applied to these 15 EF-linked genes, constructing for each gene an EF-calculation formula, "$y=Ax+B$," where "y" is EF (%) and "x" is the expression level (U) of each gene. In Table 3 shown are the resultant calculation formulae that gave the lowest P values for the coefficient "A." The EF-calculation ability of the expression levels in these four genes was highly potent, indicating that the "gene expression signature" of right atrial muscle cells reflects the character of left ventricular myocytes, at least, with regard to the pumping function. To demonstrate the accuracy of the calculated EF by the formulae, clinically observed EF and the calculated one by each formula were compared in a graph (Fig. 4B). It was to us surprising that such first-degree formulae with a single variate based on the expression intensities of the genes could predict precisely the exact digits of EF. Therefore, these four genes may not only be the molecular markers for cardiac pump function, but also be directly involved in the molecular pathology of heart failure.

We have also tried the multiple regression analysis to calculate a multivariate formula with a better prediction power than that by the simple regression analysis. However, the P values for any multivariate formulae were larger than the P values listed in Table 3. Therefore, transcription of the four genes in Table 3 may be under the regulation of common mechanisms.

To confirm the reliability of such EF-calculation approach, we set aside, from the calculation, three "test" samples, ID#5 (with the lowest EF value), ID#16 (with the 7th lowest EF value), and ID#9 (with the 13th lowest EF value), and obtained a set of genes whose expression level correlated with the EF values for the remaining 14 "training" samples (data not shown). Regression analysis was applied to such genes including one for S100A10 (GenBank Accession No. AI201310), generating an EF-calculation formula ($y=-10.017x+80.327$, $P=0.0127$) for the S100A10 mRNA level. With this formula, the EF values were then predicted for the test samples. The EF value (33%) for the ID#5 was predicted to be 39.7%, 57.5% for ID#16 was predicted to be 60.6%, and 68% for ID#9 was to be 66.0%. Similarly, regression analysis was conducted on another EF-related gene in the training samples, that for OSF-4 (GenBank

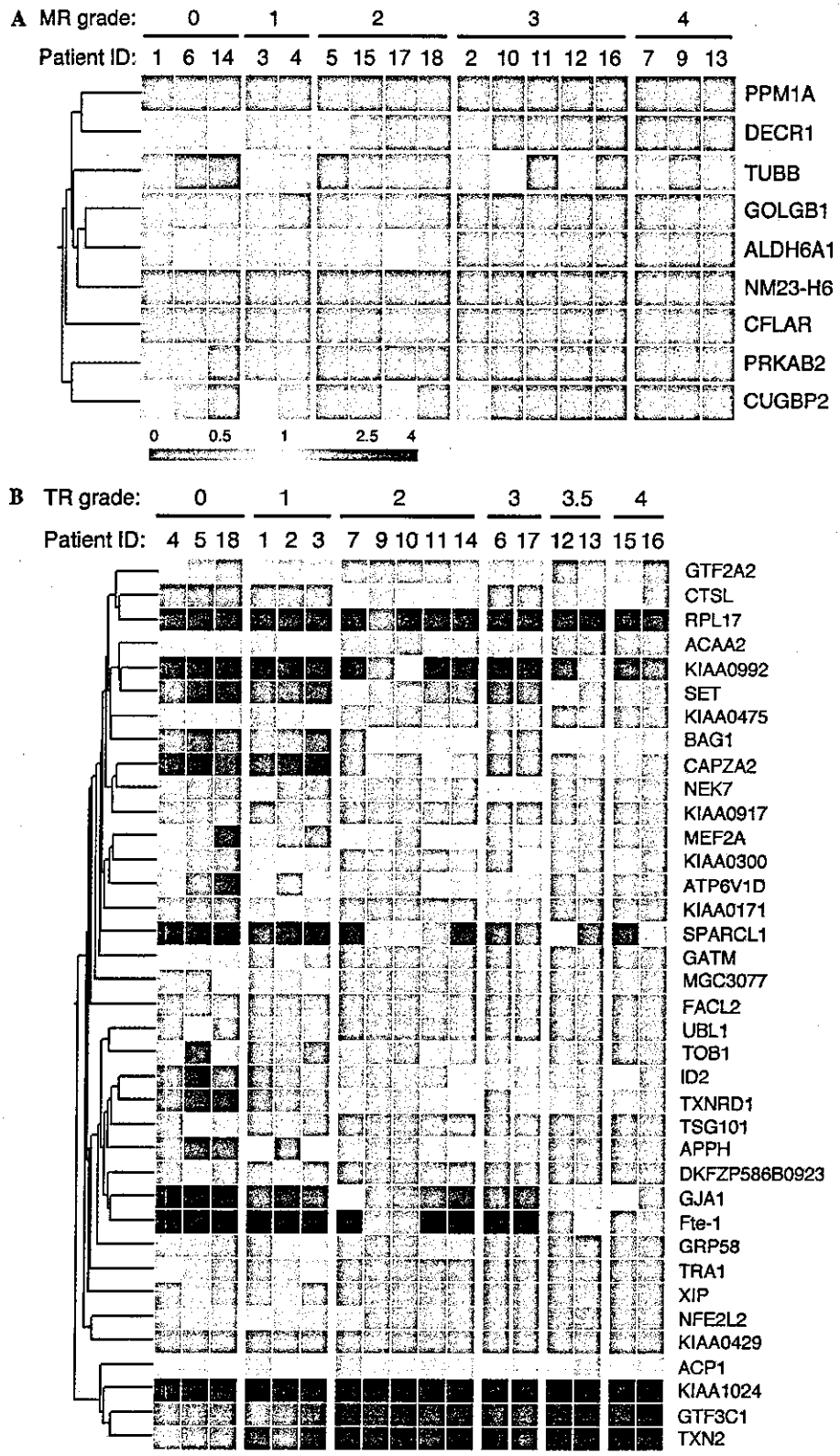

Fig. 2. Genes related to the clinical grades of MR and TR. (A) A dendrogram indicates the expression profiles of nine genes, expression of which was significantly linked to the clinical grade of MR. Each column represents a separate patient sample as indicated by the ID number at the top, and each row corresponds to a single gene. Expression level of each gene in given sample is shown color-coded as indicated by the scale at the bottom. (B) Expression profiles of 37 TR-dependent genes are demonstrated as in (A).

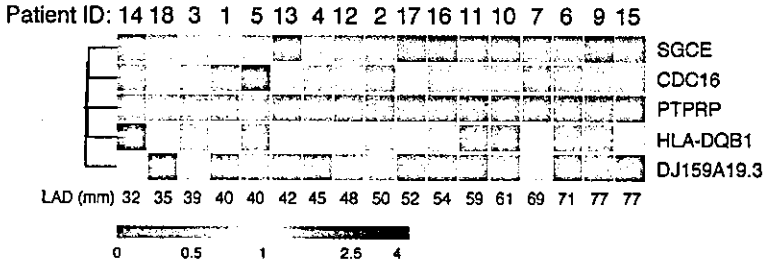

Fig. 3. Genes linked to the LAD size. Expression levels of five genes with statistically significant correlation ($P<0.001$) to the LAD size are indicated as a dendrogram as in Fig. 2A. The patient ID and its corresponding LAD size (in millimeters) are indicated at the top and the bottom, respectively.

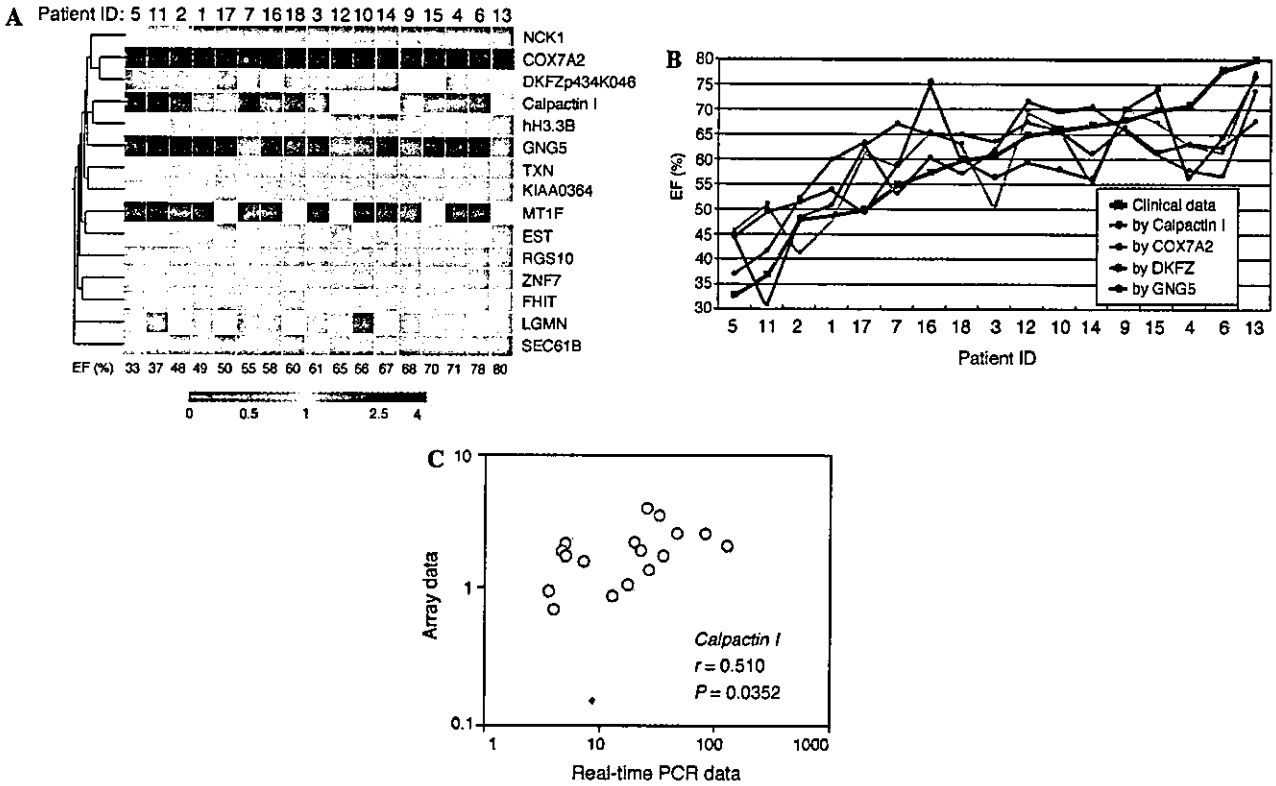

Fig. 4. Prediction of EF values by gene expression intensities. (A) Expression levels of 15 genes with statistically significant correlation ($P<0.001$) to the EF values are indicated as a dendrogram as in Fig. 2A. The patient ID and its corresponding EF value (in percents) are indicated at the top and the bottom, respectively. (B) The clinically observed EF values are compared with the predicted ones by the expression levels of four genes including that for DKFZp434K046 (DKFZ). Samples are sorted according to the clinically observed EF values. (C) Expression levels of Calpactin I mRNA are compared between the normalized microarray data and the normalized real-time RT-PCR data. Pearson's correlation coefficient (r) between the two data sets was 0.510, and this correlation was statistically significant ($P=0.0352$).

Table 3
Prediction of EF on the basis of gene expression levels

Gene	GenBank Accession No.	Prediction formula	P value for prediction
Calpactin I	AI201310	$y=-11.18x+82.03$	0.0005
COX7A2	AA978033	$y=-1.75x+91.12$	0.0006
DKFZP434K046	AC004382	$y=-27.58x+99.94$	0.001
GNG5	AI541042	$y=-4.03x+74.99$	0.0004

y represents EF (%) and x represents gene expression level (U).

Accession No. D21254). The prediction of the EF values by the OSF-4 expression level was again significantly accurate; 29.8%, 66.0%, and 71.9% for ID#5, ID#16, and ID#9, respectively. These data imply that left ventricular function can be predicted at a substantial accuracy by measures of right atrium.

Reliability of our gene expression data was confirmed by another mRNA quantitation method, real-time RT-PCR. The relative expression levels of *Calpactin I* to that of *GAPDH* were, for instance, calculated by real-time RT-PCR analysis, and were compared with the normalized expression intensities of *Calpactin I* in the microarray data (Fig. 4C). The correlation coefficient between the two data sets was 0.510, and this correlation was statistically significant ($P=0.0352$). Similar quantitation was also conducted on other genes including those for ribosomal protein L17 (*RPL17* in Fig. 2B; GenBank Accession No. X53777) ($r=0.68$, $P=0.003$), 90 kDa heat-shock protein (*HSPCA*; X15183) ($r=0.613$, $P=0.009$), and high mobility group protein 14 (*HMGN1*; J02621) ($r=0.444$, $P=0.074$).

Discussion

In this paper, with human atrial myocytes as a test case, we have tried to address whether large scale expression profiling can identify genes whose expression level is closely associated with various clinical parameters of disorders. As demonstrated in the paper, the results with an extensive statistical analysis were surprisingly informative.

We could isolate, for instance, a total of 11 genes that were regulated in a dependent manner to the presence of chronic Af arrhythmia (Fig. 1A), and that were highly potent to differentiate heart specimens with normal sinus rhythm from those with Af (Fig. 1B and Table 2). The clear separation of these two classes in a virtual space argues that protein products of these Af-dependent genes may function in the pathogenesis and/or maintenance of Af. Many genes in this list are yet only known to encode hypothetical proteins, such as HTCD37 (GenBank Accession No. AI829890), DJ159A19.3 (NM_015699), KIAA0592 (AB011164), and an EST (W28612).

A number of genes have been reported to be involved in the maintenance of Af and heart remodeling [12], including those encoding L-type Ca^{2+} channels, potassium channels, Connexins, and proteolytic enzymes (such as Calpain I). In our data set, however, we could not detect dependent changes in the expression levels of the genes for a voltage-gated potassium channel Kv11 or calcium channel proteins. The genes for other potassium channels and Calpain I were not included on the Affymetrix HGU95Av2 array.

Severe MR often induces volume-overload onto left atrium. It is still unknown to what extent such stress on left atrium affects the gene expression in right atrial myocytes. It was therefore interesting to find a number of apoptosis-related genes in the "MR-dependent" gene list (Fig. 2A). CFLAR, for instance, binds to and inhibits the activity of, caspase-8, or FLICE, and thereby protects cells from apoptotic changes [13]. CUGBP2 or NAPOR, a binding protein to CUG repeats, is known to regulate splicing mechanism of RNAs, and the *CUGBP2* message is induced during apoptosis [14]. Intriguingly, *CUGBP2* is a candidate gene whose product is responsible for heart developmental defects observed in the patients with monosomy 10p [15]. NM23-H6 functions as a nucleoside diphosphate kinase and is involved in a wide range of cell signaling. It is localized in mitochondria, and overexpression of this kinase results in growth suppression and cell cycle arrest [16]. Loss of expression of these genes may, therefore, play a role in the apoptotic changes found in the enlarged atrium associated with MR.

Apoptosis-related genes were again found in the TR-related gene list (Fig. 2B). Bcl2-associated athanogene 1 (BAG1) binds to and enhances the anti-apoptotic activity of BCL-2 [17]. Thioredoxin reductase 1 (TXNRD1; GenBank Accession No. NM_003330) functions in detoxification within cells by its reducing activities and also helps to prevent cells from apoptosis.

Finally, we have tried to construct "EF-calculation formulae" based on the expression level of EF-linked genes. In the analysis of the whole samples, a total of four such genes were used, including those for Calpactin I or S100A10, cytochrome *c* oxidase subunit VIIa polypeptide 2 (COX7A2; AA978033), guanine nucleotide binding protein γ5 (GNG5; AI541042), and a hypothetical protein (DKFZp434K046; AC004382). Given the very high prediction ability of these genes (Table 3), expression levels of them (and their products) may be important determinants of cardiac pump function.

Calpactin I is an intracellular protein that binds to and regulates the activity of Annexin II [18], a well-known substrate of Src protein–tyrosine kinase. Although the in vivo function of Calpactin I is still obscure, a recent report indicates the Calpactin I regulation of a sodium ion channel, NaV1.8 [19]. GNG5 is a γ subunit of heterotrimeric G proteins and should function to transduce upstream signals from cell surface receptor proteins [20]. Heterotrimeric G proteins play pivotal roles in the signal transduction mechanism of soluble reagents that regulate cardiac functions, such as catecholamines, endothelins, and angiotensins. Given the fact that overexpression of certain types of heterotrimeric G proteins induces cardiac hypertrophy in vivo [21], GNG5-driven intracellular signals may also directly participate in the regulation of myocyte contraction. It would be an intriguing question as to which cell surface receptors work in concert with GNG5.

Parts of atrial appendages are routinely removed in a wide array of heart surgery. Therefore, it would be intriguing if these removed samples can "tell" pathophysiological status of not only atrium but of whole heart. Our data analysis has proved that the transcriptome of right atrium provides a wealth of information in this aspect, probably more so than people had expected.

Further validation of the in vivo functions for the genes thus identified would shed a new light on the molecular pathology of heart failure. In addition, our statistical approach should be applicable not only to the investigation of heart functions, but to that of a wide variety of human disorders. Gene expression profiling is likely a very potent means to assess the detailed clinical parameters.

Acknowledgment

This work was supported in part by a Grant-in-Aid for Scientific Research on Priority Areas (C) "Medical Genome Science" from the Ministry of Education, Science, Sports and Culture of Japan.

References

[1] M.A. James, A.M. Saadeh, J.V. Jones, Wall stress and hypertension, J. Cardiovasc. Risk 7 (2000) 187–190.
[2] E. Braunwald, Shattuck lecture—cardiovascular medicine at the turn of the millennium: triumphs, concerns, and opportunities, New. Engl. J. Med. 337 (1997) 1360–1369.
[3] S. Ramaswamy, K.N. Ross, E.S. Lander, T.R. Golub, A molecular signature of metastasis in primary solid tumors, Nat. Genet. 33 (2003) 49–54.
[4] S.M. Dhanasekaran, T.R. Barrette, D. Ghosh, R. Shah, S. Varambally, K. Kurachi, K.J. Pienta, M.A. Rubin, A.M. Chinnaiyan, Delineation of prognostic biomarkers in prostate cancer, Nature 412 (2001) 822–826.
[5] K. Fellenberg, N.C. Hauser, B. Brors, A. Neutzner, J.D. Hoheisel, M. Vingron, Correspondence analysis applied to microarray data, Proc. Natl. Acad. Sci. USA 98 (2001) 10781–10786.
[6] F.D. Murgatroyd, A.J. Camm, Atrial arrhythmias, Lancet 341 (1993) 1317–1322.
[7] T. Inaba, W.M. Roberts, L.H. Shapiro, K.W. Jolly, S.C. Raimondi, S.D. Smith, A.T. Look, Fusion of the leucine zipper gene HLF to the E2A gene in human acute B-lineage leukemia, Science 257 (1992) 531–534.
[8] R.D. Sellers, M.J. Levy, K. Amplatz, C.W. Lillehei, Left retrograde cardioangiography in acquired cardiac disease: technic, indications and interpretation in 700 cases, Am. J. Cardiol. 14 (1964) 437–448.
[9] K. Miyatake, M. Okamoto, N. Kinoshita, M. Ohta, T. Kozuka, H. Sakakibara, Y. Nimura, Evaluation of tricuspid regurgitation by pulsed Doppler and two-dimensional echocardiography, Circulation 66 (1982) 777–789.
[10] A.J. Ettinger, G. Feng, J.R. Sanes, epsilon-Sarcoglycan, a broadly expressed homologue of the gene mutated in limb-girdle muscular dystrophy 2D, J. Biol. Chem. 272 (1997) 32534–32538.
[11] A. Zimprich, M. Grabowski, F. Asmus, M. Naumann, D. Berg, M. Bertram, K. Scheidtmann, P. Kern, J. Winkelmann, B. Muller-Myhsok, L. Riedel, M. Bauer, T. Muller, M. Castro, T. Meitinger, T.M. Strom, T. Gasser, Mutations in the gene encoding epsilon-sarcoglycan cause myoclonus–dystonia syndrome, Nat. Genet. 29 (2001) 66–69.
[12] B.J. Brundel, R.H. Henning, H.H. Kampinga, I.C. Van Gelder, H.J. Crijns, Molecular mechanisms of remodeling in human atrial fibrillation, Cardiovasc. Res. 54 (2002) 315–324.
[13] M. Irmler, M. Thome, M. Hahne, P. Schneider, K. Hofmann, V. Steiner, J.-L. Bodmer, M. Schroter, K. Burns, C. Mattmann, D. Rimoldi, L.E. French, J. Tschopp, Inhibition of death receptor signals by cellular FLIP, Nature 388 (1997) 190–195.
[14] D.K. Choi, T. Ito, F. Tsukahara, M. Hirai, Y. Sakaki, Developmentally-regulated expression of mNapor encoding an apoptosis-induced ELAV-type RNA binding protein, Gene 237 (1999) 135–142.
[15] P. Lichtner, T. Attie-Bitach, S. Schuffenhauer, J. Henwood, P. Bouvagnet, P.J. Scambler, T. Meitinger, M. Vekemans, Expression and mutation analysis of BRUNOL3, a candidate gene for heart and thymus developmental defects associated with partial monosomy 10p, J. Mol. Med. 80 (2002) 431–442.
[16] H. Tsuiki, M. Nitta, A. Furuya, N. Hanai, T. Fujiwara, M. Inagaki, M. Kochi, Y. Ushio, H. Saya, H. Nakamura, A novel human nucleoside diphosphate (NDP) kinase, Nm23-H6, localizes in mitochondria and affects cytokinesis, J. Cell Biochem. 76 (1999) 254–269.
[17] S. Takayama, T. Sato, S. Kuajewski, K. Kochel, S. Irie, J.A. Millan, J.C. Reed, Cloning and functional analysis of BAG-1: a novel Bcl-2-binding protein with anti-cell death activity, Cell 80 (1998) 279–284.
[18] E. Kube, K. Weber, V. Gerke, Primary structure of human, chicken, and *Xenopus laevis* p11, a cellular ligand of the Src-kinase substrate, annexin II, Gene 102 (1991) 255–259.
[19] K. Okuse, M. Malik-Hall, M.D. Baker, W.Y. Poon, H. Kong, M.V. Chao, J.N. Wood, Annexin II light chain regulates sensory neuron-specific sodium channel expression, Nature 417 (2002) 653–656.
[20] K.J. Fisher, N.N. Aronson Jr., Characterization of the cDNA and genomic sequence of a G protein gamma subunit (gamma 5), Mol. Cell. Biol. 12 (1992) 1585–1591.
[21] B.J. Aronow, T. Toyokawa, A. Canning, K. Haghighi, U. Delling, E. Kranias, J.D. Molkentin, G.W. Dorn II, Divergent transcriptional responses to independent genetic causes of cardiac hypertrophy, Physiol. Genomics 6 (2001) 19–28.

ELSEVIER

International Journal of Cardiology 96 (2004) 381–387

International Journal of Cardiology

www.elsevier.com/locate/ijcard

Transcriptional profile of genes induced in human atrial myocardium with pressure overload

Ruri Ohki[a], Keiji Yamamoto[a,*], Shuichi Ueno[a], Hiroyuki Mano[b], Yoshio Misawa[c], Katsuo Fuse[c], Uichi Ikeda[a], Kazuyuki Shimada[a]

[a] *Division of Cardiovascular Medicine, Jichi Medical School, Minamikawachi-Machi, Tochigi 329-0498, Japan*
[b] *Division of Functional Genomics, Jichi Medical School, Minamikawachi-Machi, Tochigi 329-0498, Japan*
[c] *Division of Cardiovascular Surgery, Jichi Medical School, Minamikawachi-Machi, Tochigi 329-0498, Japan*

Received 19 March 2003; received in revised form 11 July 2003; accepted 25 July 2003

Abstract

Background: The molecular response of human myocardium to mechanical stimuli, particularly the difference between pressure or volume overload cardiac hypertrophy, remains incompletely defined. *Methods:* We investigated the transcriptional profile of genes induced in human pressure- or volume-overloaded myocardium with DNA microarray technology. We used right atrial tissue from patients who underwent cardiac surgery. On the basis of pressure data and echocardiographic findings, the patients were divided into three groups: control group ($n=3$), pressure overload group (mean right atrial pressure of >7 mm Hg, $n=3$), and volume overload group (moderate or severe tricuspid regurgitation, $n=3$). Expression profiles of 2139 human genes were investigated with mRNA obtained from the samples. *Results:* In the pressure overload group, expression of genes of cyclin-dependent kinase inhibitor 1A (CDKI1A, 11.7 ± 3.1-fold vs. control), and mitogen-activated protein kinase phosphatase-1 (MKP-1, 26.2 ± 2.1-fold) was significantly increased compared with those in control or volume overload group ($P<0.05$). The specificity of these gene expressions was confirmed by a quantitative "real-time" polymerase chain reaction (PCR) analysis. In addition, mechanical strain induced CDKI1A and MKP-1 protein expressions in neonatal rat cardiac myocytes in an amplitude-dependent manner. In contrast, transcripts of growth factors did not significantly increase. *Conclusions:* This study demonstrated that gene expressions of CDKI1A and MKP-1, but not growth factors, are induced in chronic pressure-overloaded myocardium. These findings suggest that suppressors of the cell cycle or cell proliferation may play a critical role in the pathophysiology of pressure overload.

Keywords: Gene expression; Myocardium; Mechanical stress; Myocytes

1. Introduction

Cardiac hypertrophy is an independent risk factor of cardiac morbidity and mortality [1]. There are two distinct forms of cardiac hypertrophy. Pressure overload conditions such as aortic stenosis and hypertension result in concentric hypertrophy, which is characterized by an increase in ventricular wall thickness, little or no chamber dilation, and the parallel addition of sarcomeres [2,3]. Conversely, volume overload conditions such as mitral regurgitation promote an eccentric (dilated) form of hypertrophy, which is characterized by relatively little increase in wall thickness, a disproportionately large increase in chamber volume, and the serial addition of sarcomeres [4,5]. Recently, we have demonstrated that there are differences in molecular signal transduction between pressure and volume overload in vitro [6]. However, the molecular response in human pressure- or volume-overloaded myocardium remains incompletely defined.

In contrast to differential display, DNA microarray technology allows expression monitoring of hundreds or thousands of genes simultaneously and provides a format for parallel gene expression studies [7,8]. In addition to identifying large clusters of genes that respond to a given stimulus, DNA microarray technology may be used to identify a few genes that comprise highly specific molecular responses [9,10].

In the present study, we investigated the transcriptional profile of genes induced in human pressure- or volume-overloaded atrial myocardium with DNA microarray

* Corresponding author. Tel.: +81-285-58-7344; fax: +81-285-44-5317.
E-mail address: kyamamoto@jichi.ac.jp (K. Yamamoto).

0167-5273/$ - see front matter © 2003 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.ijcard.2003.07.025

technology. We found that gene expressions of cyclin-dependent kinase inhibitor 1A (CDKI1A) and mitogen-activated protein kinase phosphatase-1 (MKP-1), but not growth factors, are induced in chronic pressure-overloaded atrial myocardium. These results suggest that suppressors of the cell cycle or cell proliferation may play a critical role in the pathophysiology of pressure overload.

2. Methods

2.1. Materials

Fibronectin, fetal calf serum, and Hanks' balanced salt solution were purchased from Life Technologies (Rockville, MD). All other chemicals were purchased from Sigma (St. Louis, MO).

2.2. Subjects

This study group consisted of five men and four women (mean age 59 ± 19 years) who underwent cardiac surgery. Hemodynamic studies were performed the morning after an overnight fast. Vasodilators were withheld for at least 24 h before evaluation. Chronic, stable doses of digoxin and diuretics were continued but were administered on an evening schedule. Right and left heart studies, including measurement of pressure, biplane left ventriculography and coronary angiography, were performed using a percutaneous cathether. The severity of tricuspid regurgitation was estimated by color Doppler echocardiography with a 2.5-MHz transducer (SONOS 2500 system, Hewlett Packard, Palo Alto, CA). The severity of tricuspid regurgitation was graded on a four-point scale, based on the distance reached by the abnormal signals from the tricuspid orifice toward the posterior wall in the parasternal four-chamber view [11]. Tricuspid regurgitation was classified as trivial, mild, moderate, or severe. This study was approved by our institutional human investigations committee, and written informed consent was obtained from all patients before participation.

2.3. Atrial myocardium samples

Right atrial appendages were obtained from the patients during cardiac surgery. Pieces of right atrial appendage weighing 200–1400 mg were frozen immediately in liquid nitrogen and stored at −80 °C.

2.4. Transcriptional profiling

Total RNA was isolated by the guanidinium thiocyanate and phenol chloroform method [12]. Preparation of biotin-labeled cRNA probes were performed with the following steps; conversion of RNA to single strand cDNA by reverse transcriptase reactions and synthesis of double strand cDNA template using SuperScript II kit (Life Technologies), purification of double-strand cDNA template using QIAquick PCR purification kit (QIAGEN, Valencia, CA), and transcription of biotin-labeled cRNA (AmpliScribe kit, Epicentre Technologies, Madison, WI) according to the manufacturer's instructions. The DNA microarray hybridization of biotin-labeled cRNA was performed using ExpressChip HO1 and HO2 arrays (Mergen, San Leandro, CA) according to the manufacturer's instructions. The ExpressChip HO1 and HO2 arrays have 2139 well-characterized genes with putative functions. A complete listing of genes contained within ExpressChip HO1 and HO2 can be found at http://www.mergen-ltd.com/HO1/HO1abc_genelist.htm and http://www.mergen-ltd.com/HO2/HO2abc_genelist.htm. The chips were subjected to the laser scanning and signal detection by the GMS418 Array Scanner (Takara Biomedicals, Shiga, Japan). The intensity of emission signals in each oligonucleotide hybridization was normalized to that of the glyceraldehyde-3 phosphate dehydrogenase (GAPDH) signal and analyzed using the GeneSpring software package (Silicon Genetics, Redwood City, CA).

2.5. Real-time reverse transcription–polymerase chain reaction (PCR) analysis

For reverse transcription, RNA was reverse transcribed using T7-dT primer (5′ -TCT AGT CGA CGG CCA GTG AAT TGT AAT ACG ACT CAC TAT AGG GCG TTT TTT TTT TTT TTT TTT TTT-3′) and Superscript II reverse transcriptase (Life Technologies). Real-time quantitative PCR was performed in optical tubes in a 96-well microtiter plate (Perkin-Elmer/Applied Biosystems, Foster City, CA) with an ABI PRISM 7700 Sequence Detector Systems (Perkin-Elmer/Applied Biosystems) according to the manufacturer's instructions. By using the SYBR Green PCR Core Reagents Kit (Perkin-Elmer/Applied Biosystems, P/N 4304886), fluorescence signals were generated during each PCR cycle via the 5′ - to 3′ - endonuclease activity of Taq Gold [13] to provide real-time quantitative PCR information. The oligonucleotide primers used for real-time PCR analysis were shown in Table 1. No template controls as well as the samples were added in a total volume of 50 μl/reaction. Potential PCR product contamination was digested by uracil-*N*-glycosylase, because dTTP is substituted by dUTP [13]. All PCR experiments were performed with the hot start method. In the reaction system, uracil-*N*-glycosylase and Taq Gold (Perkin-Elmer/Applied Biosystems) were applied according to the manufacturer's instructions [13,14]. Denaturing and annealing reactions were performed 40 times at 95 °C for 15 s, and at 60 °C for CDKI1A, 54 °C for MKP-1, 56 °C for glutathione S-transferase theta 1 and 62 °C