tion profiles of the cardiovascular system. They should also provide a basis for further characterization of the importance of epigenetic alterations in the development of cardiac hypertrophy or heart failure. Application of DCS to human heart specimens has the potential to highlight such information of clinical relevance.

Acknowledgments

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, Culture, Sports, Science, and Technology of Japan and by a grant from Salt Science Research Foundation (#04C7).

References

1. Baylin SB, Herman JG. DNA hypermethylation in tumorigenesis: epigenetics joins genetics. *Trends Genet.* 2000;16:168–174.
2. van Leeuwen F, Gottschling DE. Genome-wide histone modifications: gaining specificity by preventing promiscuity. *Curr Opin Cell Biol.* 2002; 14:756–762.
3. Burgess-Beusse B, Farrell C, Gaszner M, Litt M, Mutskov V, Recillas-Targa F, Simpson M, West A, Felsenfeld G The insulation of genes from external enhancers and silencing chromatin. *Proc Natl Acad Sci U S A.* 2002;99:16433–16437.
4. Carrozza MJ, Utley RT, Workman JL, Cote J. The diverse functions of histone acetyltransferase complexes. *Trends Genet.* 2003;19:321–329.
5. Verdin E, Dequiedt F, Kasler HG. Class II histone deacetylases: versatile regulators. *Trends Genet.* 2003;19:286–293.
6. Gusterson RJ, Jazrawi E, Adcock IM, Latchman DS. The transcriptional co-activators CREB-binding protein (CBP) and p300 play a critical role in cardiac hypertrophy that is dependent on their histone acetyltransferase activity. *J Biol Chem.* 2003;278:6838–6847.
7. Iezzi S, Di Padova M, Serra C, Caretti G, Simone C, Maklan E, Minetti G, Zhao P, Hoffman EP, Puri PL, Sartorelli V. Deacetylase inhibitors increase muscle cell size by promoting myoblast recruitment and fusion through induction of follistatin. *Dev Cell.* 2004;6:673–684.
8. Zhang CL, McKinsey TA, Chang S, Antos CL, Hill JA, Olson EN. Class II histone deacetylases act as signal-responsive repressors of cardiac hypertrophy. *Cell.* 2002;110:479–488.
9. Kook H, Lepore JJ, Gitler AD, Lu MM, Wing-Man Yung W, Mackay J, Zhou R, Ferrari V, Gruber P, Epstein JA. Cardiac hypertrophy and histone deacetylase-dependent transcriptional repression mediated by the atypical homeodomain protein Hop. *J Clin Invest.* 2003;112:863–871.
10. Antos CL, McKinsey TA, Dreitz M, Hollingsworth LM, Zhang CL, Schreiber K, Rindt H, Gorczynski RJ, Olson EN. Dose-dependent blockade to cardiomyocyte hypertrophy by histone deacetylase inhibitors. *J Biol Chem.* 2003;278:28930–28937.
11. Kuwahara K, Saito Y, Ogawa E, Takahashi N, Nakagawa Y, Naruse Y, Harada M, Hamanaka I, Izumi T, Miyamoto Y, Kishimoto I, Kawakami R, Nakanishi M, Mori N, Nakao K. The neuron-restrictive silencer element-neuron-restrictive silencer factor system regulates basal and endothelin 1-inducible atrial natriuretic peptide gene expression in ventricular myocytes. *Mol Cell Biol.* 2001;21:2085–2097.
12. Kaneda R, Toyota M, Yamashita Y, Koinuma K, Choi YL, Ota J, Kisanuki H, Ishikawa M, Takada S, Shimada K, Mano H. High-throughput screening of genome fragments bound to differentially acetylated histones. *Genes Cells.* 2004;9:1167–1174.
13. Yamamoto K, Dang QN, Kelly RA, Lee RT. Mechanical strain suppresses inducible nitric-oxide synthase in cardiac myocytes. *J Biol Chem.* 1998;273:11862–11866.
14. Kent WJ. BLAT–the BLAST-like alignment tool. *Genome Res.* 2002;12: 656–664.
15. Rat Genome Sequencing Project Consortium. Genome sequence of the Brown Norway rat yields insights into mammalian evolution. *Nature.* 2004;428:493–521.
16. Marks AR. Cardiac intracellular calcium release channels: role in heart failure. *Circ Res.* 2000;87:8–11.
17. Kapinya KJ, Harms U, Harms C, Blei K, Katchanov J, Dirnagl U, Hortnagl H. Role of NAD(P)H:quinone oxidoreductase in the progression of neuronal cell death in vitro and following cerebral ischaemia in vivo. *J Neurochem.* 2003;84:1028–1039.
18. Shawber C, Nofziger D, Hsieh JJ, Lindsell C, Bogler O, Hayward D, Weinmaster G. Notch signaling inhibits muscle cell differentiation through a CBF1-independent pathway. *Development.* 1996;122: 3765–3773.
19. Henderson C, Brancolini C. Apoptotic pathways activated by histone deacetylase inhibitors: implications for the drug-resistant phenotype. *Drug Resist Updat.* 2003;6:247–256.
20. Chen QM, Tu VC, Wu Y, Bahl JJ. Hydrogen peroxide dose dependent induction of cell death or hypertrophy in cardiomyocytes. *Arch Biochem Biophys.* 2000;373:242–248.
21. Onan D, Pipolo L, Yang E, Hannan RD, Thomas WG. Urotensin II promotes hypertrophy of cardiac myocytes via mitogen-activated protein kinases. *Mol Endocrinol.* 2004;18:2344–2354.
22. Fingert JH, Heon E, Liebmann JM, Yamamoto T, Craig JE, Rait J, Kawase K, Hoh ST, Buys YM, Dickinson J, Hockey RR, Williams-Lyn D, Trope G, Kitazawa Y, Ritch R, Mackey DA, Alward WL, Sheffield VC, Stone EM. Analysis of myocilin mutations in 1703 glaucoma patients from five different populations. *Hum Mol Genet.* 1999;8: 899–905.
23. Bernstein BE, Kamal M, Lindblad-Toh K, Bekiranov S, Bailey DK, Huebert DJ, McMahon S, Karlsson EK, Kulbokas EJ 3rd, Gingeras TR, Schreiber SL, Lander ES. Genomic maps and comparative analysis of histone modifications in human and mouse. *Cell.* 2005;120:169–181.
24. Weinmann AS, Yan PS, Oberley MJ, Huang TH, Farnham PJ. Isolating human transcription factor targets by coupling chromatin immunoprecipitation and CpG island microarray analysis. *Genes Dev.* 2002;16:235–244.
25. Turner BM, O'Neill LP. Histone acetylation in chromatin and chromosomes. *Semin Cell Biol.* 1995;6:229–236.
26. Strahl BD, Allis CD. The language of covalent histone modifications. *Nature.* 2000;403:41–45.

Carcinogenesis vol.26 no.12 pp.2078–2085, 2005
doi:10.1093/carcin/bgi184
Advance Access publication July 20, 2005

Screening for genomic fragments that are methylated specifically in colorectal carcinoma with a methylated *MLH1* promoter

Koji Koinuma[1,2], Ruri Kaneda[1], Minoru Toyota[3], Yoshihiro Yamashita[1], Shuji Takada[1], Young Lim Choi[1], Tomoaki Wada[1], Masaki Okada[2], Fumio Konishi[4], Hideo Nagai[2] and Hiroyuki Mano[1,5,*]

[1]Division of Functional Genomics and [2]Department of Surgery, Jichi Medical School, Tochigi 329-0498, Japan, [3]Department of Molecular Biology, Cancer Research Institute, Sapporo Medical University, Hokkaido 060-8556, Japan, [4]Department of Surgery, Omiya Medical Center of Jichi Medical School, Saitama 330-8503, Japan and [5]CREST, Japan Science and Technology Agency, Saitama 332-0012, Japan

*To whom correspondence should be addressed. Tel: +81 285 58 7449; Fax: +81 285 44 7322;
E-mail: hmano@jichi.ac.jp

A subset of colorectal carcinomas (CRCs) is associated with microsatellite instability (MSI) of the genome. Although extensive methylation of CpG islands within the promoter regions of DNA mismatch repair genes such as *MLH1* is thought to play a central role in tumorigenesis for MSI-positive sporadic CRCs, it has been obscure whether such aberrant epigenetic regulation occurs more widely and affects other cancer-related genes *in vivo*. Here, by using methylated CpG island amplification coupled with representational difference analysis (MCA–RDA), we screened genomic fragments that are selectively methylated in CRCs positive for *MLH1* methylation, resulting in the identification of hundreds of CpG islands containing genomic fragments. Methylation status of such CpG islands was verified for 28 genomic clones in 8 CRC specimens positive for *MLH1* methylation and the corresponding paired normal colon tissue as well as in 8 CRC specimens negative for methylation. Many of the CpG islands were preferentially methylated in the *MLH1* methylation-positive CRC specimens, although methylation of some of them was more widespread. These data provide insights into the complex regulation of the methylation status of CpG islands in CRCs positive for MSI and *MLH1* methylation.

Introduction

Colorectal carcinoma (CRC) is one of the leading causes of cancer death in humans. Evidence indicates the existence of two major types of genomic instability in CRCs: chromosomal instability and microsatellite instability (MSI) (1). Whereas chromosomal instability is associated with an abnormal DNA content (such as aneuploidy), inactivation of the tumor suppressor gene *TP53*, and activation of oncogenes (2), MSI is associated with defects in the DNA mismatch repair (MMR) system that result in frameshift mutations in microsatellite repeats, and thereby affect the structure of genes containing such repeats (3).

Although germ-line mutations of MMR genes have been detected in the genome of individuals with hereditary non-polyposis colorectal cancer (4–6), many sporadic CRCs positive for MSI are associated with epigenetic silencing of non-mutated MMR genes (7,8). MSI-positive CRCs are characterized by specific clinicopathologic features and gene mutations. Such tumors occur with a higher frequency in women than in men, develop in the right side of the colon, and manifest a mucinous or poorly differentiated histopathology. Many of the CpG dinucleotides within the promoter region of the MMR gene *MLH1* are methylated in MSI-positive CRCs (9,10), and the *BRAF* gene frequently contains activating mutations in these cancers (11–13). Some genomic fragments have been found to be methylated specifically in such CRCs (7), and an entity of CRC with a CpG island methylator phenotype (CIMP) has been proposed (14). However, the profiles of genes and genomic fragments that become methylated in CRC specimens positive for *MLH1* methylation have remained uncharacterized.

With the use of methylated CpG island amplification coupled with representational difference analysis (MCA–RDA) (15), we have now performed a global screening of pooled genomic DNA from CRC specimens positive or negative for *MLH1* methylation in order to identify differentially methylated genomic fragments. With this approach, we identified hundreds of CpG islands whose methylation was specific to CRCs with a methylated *MLH1* promoter.

Abbreviations: CRC, colorectal carcinoma; COBRA, combined bisulfite restriction analysis; CIMP, CpG island methylator phenotype; EGF, epidermal growth factor; GDF, growth-differentiation factor; MSI, microsatellite instability; MMR, mismatch repair; PCR, polymerase chain reaction; TGF-β, transforming growth factor-β.

Materials and methods

Tumor specimens and cell lines

Tumor specimens were obtained from patients with sporadic CRC who underwent surgical treatment in Jichi Medical School Hospital. Informed consent was obtained from each patient, and the study was approved by the ethics committee of Jichi Medical School. Normal portion of colon tissue was excised from a region >5 cm distant from the cancerous region in every case. Genomic DNA was extracted with the use of a QIAmp DNA Mini kit (Qiagen, Valencia, CA). The MSI status of each tumor was determined on the basis of analysis of nine microsatellite repeat loci as previously described (8). The methylation status of the *MLH1* promoter was also examined in each sample (11).

Colon carcinoma cell lines (Caco-2, HCT116, SW480) were obtained from American Type Culture Collection (Manassas, VA). Caco2 cells were cultured in Dulbecco's modified Eagle's medium (DMEM)-F12 (Invitrogen, Carlsbad, CA) supplemented with 10% fetal bovine serum (FBS) (Invitrogen) and 2 mM L-glutamine. HCT116 and SW480 cells were maintained in McCoy's 5A medium (Invitrogen) and Leibovitz's L-15 medium (Invitrogen), both supplemented with 10% FBS, respectively.

MCA–RDA

Genomic DNA from four tumor specimens positive for *MLH1* methylation was mixed and used as a 'tester' sample, whereas that from four specimens negative for *MLH1* methylation was used as a 'driver.' MCA–RDA was performed with the two pooled DNA samples as previously described (15). In brief, both the tester and the driver DNA samples (5 μg of each) were digested first with the SmaI endonuclease (New England Biolabs, Beverly, MA) and then with

XmaI (New England Biolabs), and the resulting fragments were ligated to the RMCA adapter (15). Amplification of methylated CpG islands was achieved by polymerase chain reaction (PCR) with the RMCA24 primer. The amplified fragments of the tester-DNA and driver-DNA were digested with XmaI and SmaI, respectively. The tester amplicons were then ligated to the JMCA adapter and subjected to annealing with an excess amount of the driver amplicons. PCR with the JMCA24 primer then amplified only the tester-specific amplicons. Another round of amplification was performed with the NMCA adapter and the NMCA24 primer (15). The final products were digested with XmaI and ligated into XmaI-digested pBlueScript (Stratagene, La Jolla, CA) for nucleotide sequencing.

Combined bisulfite restriction analysis (COBRA)

The methylation status of isolated clones was tested by the COBRA method (16). The genomic DNA was denatured, incubated for 16 h at 55°C in 3.1 M sodium bisulfite, and then subjected to PCR to amplify CpG islands. The PCR products were digested with a restriction endonuclease, and the resulting fragments were fractionated by polyacrylamide gel electrophoresis (PAGE). The gel was stained with SYBR Green I (Takara Bio, Shiga, Japan) and scanned with an LAS3000 imaging system (Fuji Film, Tokyo, Japan). Genomic fragments were determined to be positive for CpG methylation if ≥10% of the PCR products were cleaved by the restriction endonuclease. The PCR primers and endonucleases used for COBRA are shown in the Supplementary Table online.

For bisulfite sequencing of the *BMP3* promoter, genomic DNA isolated from cancer specimens or cell lines was treated with sodium bisulfite (11) and then subjected to PCR with the primers 5′-AGTTAGAGAGYGAAAGAATTAAG-3′ and 5′-ATACAACRAAATAACRACCAACC-3′. The PCR product was ligated into pGEMT-easy (Promega, Madison, WI).

Real-time reverse transcription-polymerase chain reaction (RT–PCR)

Total RNA was prepared from tester or driver samples with an RNeasy Mini column (Qiagen), treated with RNase-free DNase (Qiagen), and subjected to reverse transcription with PowerScript reverse transcriptase (BD Biosciences Clontech, San Jose, CA) and an oligo(dT) primer. Portions of the resulting cDNAs were subjected to PCR with a QuantiTect SYBR Green PCR kit (Qiagen). The amplification protocol comprised incubations at 94°C for 15 s, 63°C (64°C for *KIT* cDNA) 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 7900HT 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 DNA molecules corresponding to the β-actin gene (*ACTB*) and to genomic fragments of interest were used to calculate the abundance of the latter relative to that of the former. The oligonucleotide primers for PCR were 5′-CCATCATGAAGTGTGACGTGG-3′ and 5′-GTCCGCCTAGAAGCATTTGCG-3′ for *ACTB*, 5′-AAGTCAACTCCTTGGCCATCTGT-3′ and 5′-TGGAAAAGGTAACCTCCTCTTTGG-3′ for the bone morphogenetic protein 3 gene (*BMP3*), and 5′-TGACGTCTGGTCCTATGGGATTT-3′ and 5′-TACATTTCAGCAGGTGCGTGTTC-3′ for *KIT*.

Results

MCA–RDA screening

A total of 249 cases with CRC were examined for their MSI status as well as for methylation in the promoter region of *MLH1*. The majority ($n = 213$) of the tumor specimens from these patients were negative for MSI, while the rest ($n = 36$) were MSI-positive. Also, most of them ($n = 226$) were shown to have no methylation within the *MLH1* promoter (unmethylated group), whereas the remainder ($n = 23$) had a methylated promoter (methylated group) (11). Altogether 16 patients were positive for MSI, but did not have a methylated *MLH1* promoter. On the other hand, three patients were negative for MSI despite the presence of the methylated *MLH1* promoter. The characteristics of the patients examined in the present study are summarized in Table I.

To isolate genes that were specifically methylated in the methylated group, we selected specimens from four men of each group: ID nos 2, 17, 20 and 77 (mean age, 70.0 years) from the methylated group, and ID nos 1, 8, 13 and 31 (mean age, 73.0 years) from the unmethylated group. We have excluded female subjects from the initial screening, since an intense methylation of one X chromosome in female cells may have yielded a large number of pseudopositive clones, methylation status of which may not have linked to the clinical classes, but to lyonization.

Table I. Patient characteristics

ID no.	*MLH1* methylation	MSI status	Age (years)	Sex	Dukes grade
2[a]	Yes	MSS	70	Male	A
17[a]	Yes	MSI	65	Male	B
20[a]	Yes	MSI	76	Male	B
77[a]	Yes	MSI	69	Male	A
225	Yes	MSI	83	Female	C
263	Yes	MSI	86	Female	C
280	Yes	MSI	83	Female	C
305	Yes	MSI	74	Male	B
318	Yes	MSI	76	Female	B
336	Yes	MSI	68	Male	B
413	Yes	MSI	69	Female	A
416	Yes	MSI	76	Female	B
1[a]	No	MSS	65	Male	A
8[a]	No	MSS	89	Male	B
13[a]	No	MSS	74	Male	B
31[a]	No	MSS	64	Male	B
238	No	MSS	74	Male	A
249	No	MSS	62	Male	B
255	No	MSS	69	Female	C
278	No	MSS	73	Male	C
295	No	MSS	71	Female	C
298	No	MSS	70	Male	D
307	No	MSS	80	Female	C
308	No	MSS	62	Male	B
481	No	MSS	59	Male	C

[a]Specimens used for MCA–RDA screening. MSS, microsatellite stable.

Equal amounts of genomic DNA isolated from the four selected tumor specimens of each group were mixed and subjected to MCA–RDA analysis, with the pooled DNA of the methylated group as the tester and that of the unmethylated group as the driver. A total of 384 clones were randomly picked up from the resultant MCA–RDA products, and digestion of the purified plasmid DNA with restriction endonucleases revealed that 294 out of the 384 clones carried the insert fragments. Nucleotide sequencing of such 294 clones indicated that 209 of the clones were found to contain CpG islands. Screening of human genome databases (http://www.ncbi.nlm.nih.gov/BLAST/ and http://genome.ucsc.edu/cgi-bin/hgBlat) with these sequences revealed that 186 of them were localized within or in close proximity to characterized or uncharacterized human genes (112 independent genes).

Candidate genes for differential methylation

The GenBank accession numbers and annotation information for the 112 genes identified by MCA–RDA are shown in Table II. Multiple clones were isolated for a single gene in 33 instances, whereas only one MCA–RDA product was obtained for the remaining genes. The genes listed in Table II were thus candidates for genes whose CpG islands are methylated in a manner dependent on *MLH1* methylation.

We then examined the methylation status of the isolated fragments mapped to the promoter regions. First, we tried to amplify individually by PCR all 70 fragments mapped to the promoter regions in Table II, and could successfully amplify 35 fragments from the pooled DNAs used in MCR–RDA.

Table II. Genes identified by MCA–RDA with CRCs positive or negative for *MLH1* methylation

Gene	GenBank accession no.	No. of MCA–RDA clones	Position of MCA–RDA clones	COBRA data
Homo sapiens, solute carrier family 38, member 3	BC042875	11	Promoter	N.D.
Homo sapiens hypothetical protein MGC29643[a]	AK075487	10	Promoter	compatible
Human progesterone receptor	M15716	6	Exon 1	N.D.
Homo sapiens alpha-1 type XV collagen	L25286	6	Promoter	N.D.
Homo sapiens K-Cl co-transporter KCC4	AF105365	5	Promoter	N.D.
Homo sapiens CGI-150 protein	AF151908	5	3′ region	N.D.
Homo sapiens cDNA FLJ37615	AK094934	4	3′ region	N.D.
Human mRNA and promoter DNA for progesterone receptor	X51730	3	Promoter	N.D.
Human arachidonate 12-lipoxygenase mRNA	M62982	3	Promoter	not compatible
Homo sapiens clone IMAGE:5173621	BC031660	3	Promoter	N.D.
Homo sapiens Ras and Rab interactor 1	BC014417	3	3′ region	N.D.
Homo sapiens papilin (PAPLN)[a]	BC042057	3	Promoter	compatible
Homo sapiens F-box and leucine-rich repeat protein 7 (FBXL7)[a]	AB020647	3	Promoter	compatible
Homo sapiens fibronectin type 3 and ankyrin repeat domains 1	BC024189	3	Promoter	not compatible
BX444427 *Homo sapiens* ADULT BRAIN *Homo sapiens* cDNA clone	BX444427	3	Promoter	N.D.
UI-H-BW0-aju-d-09-0-UI.s1 NCI_CGAP_Sub6 *Homo sapiens* cDNA clone	AW297872	2	N.D.	N.D.
Human dystrobrevin-delta	U26742	2	Near 5′ end	N.D.
Homo sapiens, clone IMAGE:5728979	BC035731	2	Promoter	compatible
Homo sapiens ras interactor (RIN1)	L36463	2	Promoter	N.D.
Homo sapiens partial mRNA for doublesex-mab-3 (DM) domain	AJ301580	2	Promoter	not compatible
Homo sapiens mRNA; cDNA DKFZp586D0619	AL834130	2	Promoter	N.D.
Homo sapiens mRNA for NTAK[a]	AB005060	2	Promoter	compatible
Homo sapiens mRNA for FLJ00396 protein	AK090474	2	3′ region	N.D.
Homo sapiens mRNA for dihydropyrimidinase related protein-3 (DPYSL3)[a]	D78014	2	Promoter	compatible
Homo sapiens hypothetical protein MGC35308	BC034775	2	Exon 1	N.D.
Homo sapiens hypothetical protein MGC33600	BC035022	2	N.D.	N.D.
Homo sapiens growth differentiation factor 7 (GDF7)[a]	AF522369	2	Promoter	compatible
Homo sapiens gene NXN encoding nucleoredoxin	NM_022463	2	N.D.	N.D.
Homo sapiens clone 24649	AF070591	2	3′ region	N.D.
Homo sapiens cDNA: FLJ21511	AK025164	2	Promoter	not compatible
Homo sapiens solute carrier family 30, member 10 (SLC30A10)[a]	AK090806	2	Promoter	compatible
Homo sapiens beta-parvin	AF237769	2	Promoter	N.D.
603031612F1 NIH_MGC_115 *Homo sapiens* cDNA clone IMAGE:5172891	BI489865	2	Promoter	N.D.
Mus musculus adult male spinal cord cDNA, clone:A330088B02	AK079637	1	N.D.	N.D.
Mus musculus 9.5 days embryo parthenogenote cDNA, clone:B130054P17	AK045275	1	N.D.	N.D.
Mus musculus 12 days embryo spinal ganglion cDNA, clone:D130032D18	AK051302	1	N.D.	N.D.
IL5-EN0086-281100-292-f08 EN0086 *Homo sapiens* cDNA	BF851362	1	N.D.	N.D.
Human receptor tyrosine kinase ligand LERK-7 precursor	U26403	1	Promoter	not compatible
Human solute carrier family 30, member 3 (SLC30A3)[a]	U76010	1	Promoter	compatible
Human platelet-derived growth factor receptor alpha	M21574	1	3′ region	N.D.
Human pephBGT-1 betaine-GABA transporter	U27699	1	3′ region	N.D.
Human oxytocin mRNA	M25650	1	Promoter	N.D.
Human mRNA for KIAA0222 gene	D86975	1	3′ region	N.D.
Human c-kit proto-oncogene (KIT)[a]	X06182	1	Promoter	compatible
Human cell 12-lipoxygenase	M35418	1	Promoter	N.D.
Human bone morphogenetic protein-3 (BMP3)[a]	M22491	1	Promoter	compatible
Human (clone hST3O-1) sialyltransferase	L29555	1	Promoter	N.D.
Homo sapiens, Similar to parathyroid hormone receptor 1	BC031578	1	3′ region	N.D.
Homo sapiens, potassium channel, subfamily K, member 13 (KCNK13)[a]	BC012779	1	Promoter	compatible
Homo sapiens, clone MGC:50339	BC043386	1	N.D.	N.D.
Homo sapiens, clone IMAGE:6041910	BC040712	1	Promoter	N.D.
Homo sapiens, clone IMAGE:5590527	BC040874	1	Promoter	N.D.
Homo sapiens Wilms tumor 1	BC032861	1	Promoter	not compatible
Homo sapiens TRALPUSH[a]	AF399708	1	Promoter	compatible
Homo sapiens Sry-related HMG-box protein	AF270652	1	Exon 1	N.D.
Homo sapiens sialyltransferase 4A, transcript variant 2 (SIAT4A)[a]	BC018357	1	Promoter	compatible
Homo sapiens protein tyrosine phosphatase, receptor type, N polypeptide 2	BC034040	1	3′ region	N.D.
Homo sapiens prostaglandin E2 receptor	L25124	1	Promoter	N.D.
Homo sapiens polyamine modulated factor-1	AF141310	1	3′ region	N.D.
Homo sapiens nuclear receptor subfamily 5, group A, member 1	BC032501	1	3′ region	N.D.
Homo sapiens NEL-like 2 (NELL2)[a]	BC020544	1	Promoter	compatible
Homo sapiens mRNA; cDNA DKFZp667I0324	AL832828	1	N.D.	N.D.
Homo sapiens mRNA; cDNA DKFZp564L0472	AL080101	1	3′ region	N.D.
Homo sapiens mRNA; cDNA DKFZp564G1482	AL136698	1	Promoter	N.D.
Homo sapiens mRNA, chromosome 1 specific transcript KIAA0495	AB007964	1	Promoter	N.D.
Homo sapiens chromosome 13 open reading frame 21 (C13orf21)[a]	BC029067	1	Promoter	compatible
Homo sapiens mRNA for SOX7 protein[a]	AJ409320	1	Promoter	compatible
Homo sapiens mRNA for nephrosis 2, idiopathic, steroid-resistant (NPHS2)[a]	AJ279254	1	Promoter	compatible
Homo sapiens mRNA for MDC2 alpha, MDC2 beta	AB009671	1	Promoter	N.D.
Homo sapiens mRNA for KIAA0641 protein	AB014541	1	N.D.	N.D.

Table II. *Continued*

Gene	GenBank accession no.	No. of MCA–RDA clones	Position of MCA–RDA clones	COBRA data
Homo sapiens mRNA for transcription factor 7-like 1 (TCF7L1)[a]	AB031046	1	Promoter	compatible
Homo sapiens mRNA for calmegin (CLGN)[a]	D86322	1	Promoter	compatible
Homo sapiens mRNA for ADAMTS19 protein[a]	AJ311904	1	Promoter	compatible
Homo sapiens microtubule-associated protein 1B, transcript variant 1	NM_005909	1	Promoter	N.D.
Homo sapiens matrix metalloproteinase-21	AF520613	1	Exon 2	N.D.
Homo sapiens KIAA0534 protein	BC047716	1	Promoter	N.D.
Homo sapiens hypothetical protein MGC49007	BC041175	1	N.D.	N.D.
Homo sapiens hypothetical protein LOC284801	BC036201	1	Promoter	N.D.
Homo sapiens hypothetical protein LOC283887[a]	BC023651	1	Promoter	compatible
Homo sapiens hypothetical protein BC009980	BC009980	1	Promoter	N.D.
Homo sapiens GATA binding protein 2	BC051342	1	Promoter	not compatible
Homo sapiens forkhead-related transcription factor FREAC-9	AF042832	1	Exon 1	N.D.
Homo sapiens erythroblast macrophage protein EMP	AF084928	1	Promoter	N.D.
Homo sapiens Enah/Vasp-like (EVL)[a]	BC023997	1	Promoter	compatible
Homo sapiens cytokine receptor-like factor 1	BC044634	1	Promoter	N.D.
Homo sapiens cystathionine beta-synthase (CBS)[a]	L14577	1	Promoter	compatible
Homo sapiens clone DNA68818 PSST739	AY358393	1	Promoter	N.D.
Homo sapiens ceh-10 homeodomain containing protein	AY336059	1	Exon 1	N.D.
Homo sapiens cDNA FLJ42875	AK124865	1	Promoter	N.D.
Homo sapiens cDNA FLJ41549[a]	AK123543	1	Promoter	compatible
Homo sapiens cDNA FLJ38293	AK095612	1	Promoter	N.D.
Homo sapiens cDNA FLJ37464[a]	AK094783	1	Promoter	compatible
Homo sapiens cDNA FLJ33739	AK091058	1	3′ region	N.D.
Homo sapiens cDNA FLJ14238	AK024300	1	Promoter	N.D.
Homo sapiens cDNA clone IMAGE:6025756	BC064906	1	Promoter	N.D.
Homo sapiens bridging integrator protein-1	U68485	1	3′ region	N.D.
Homo sapiens brain tumor associated protein LRRC4[a]	AF196976	1	Promoter	compatible
Homo sapiens apoptosis-associated tyrosine kinase, mRNA	BC047378	1	3′ region	N.D.
ho64c01.x1 Soares_NFL_T_GBC_S1 *Homo sapiens* cDNA clone	AW873619	1	Promoter	N.D.
Helix pomatia sulfatase 1 precursor	AF109924	1	Promoter	N.D.
H.sapiens mitogen inducible gene mig-2 (MIG2)[a]	Z24725	1	Promoter	compatible
GENCOURT_8532095 NIH_MGC_113 *Homo sapiens* cDNA clone	BU899260	1	N.D.	N.D.
EST379664 MAGE resequences, MAGJ *Homo sapiens* cDNA	AW967589	1	N.D.	N.D.
BX394700 *Homo sapiens* NEUROBLASTOMA COT 25-NORMALIZED	BX394700	1	Promoter	N.D.
AL545903 *Homo sapiens* PLACENTA COT 25-NORMALIZED	AL545903	1	Promoter	N.D.
AGENCOURT_8219616 Lupski_sympathetic_trunk *Homo sapiens* cDNA clone	BQ722471	1	N.D.	N.D.
AGENCOURT_7546470 NIH_MGC_70 *Homo sapiens* cDNA clone	BQ218409	1	Promoter	N.D.
7k34e06.x1 NCI_CGAP_Ov18 *Homo sapiens* cDNA clone IMAGE:3477227	BF058764	1	N.D.	N.D.
602695383F1 NIH_MGC_97 *Homo sapiens* cDNA clone IMAGE:4827284	BG722892	1	N.D.	N.D.
Homo sapiens mRNA for KIAA0711 protein	AB018254	1	3′ region	N.D.
Homo sapiens E2F binding protein	AY152547	1	Exon 1	N.D.
Homo sapiens cDNA FLJ38336	AK095655	1	3′ region	N.D.

N.D., not determined. COBRA data indicated that methylation level of the MCA–RDA clones in CRCs positive for *MLH1* methylation was increased (compatible) or not (not compatible) compared with that in CRCs negative for *MLH1* methylation.
[a]Clones analyzed in a test set of samples.

By using COBRA, their CpG methylation status was assessed among the samples used in MCR–RDA. As indicated in Table II, 28 fragments out of 35 were preferentially methylated in the tester DNA, while 7 of them were not.

The methylation status of such 28 fragments was further tested in clinical specimens that had not been used for the initial screening. This test set included eight cancer specimens positive for *MLH1* methylation and the paired normal colon tissue as well as eight cancer specimens negative for *MLH1* methylation (Table I).

The methylation status of each genomic fragment in the clinical specimens is shown color-coded in Figure 1; fragments with a methylation level of $\geq$10% as determined by COBRA are indicated in red, whereas those with a methylation level of $<$10% are shown in blue. Most genomic fragments were extensively methylated in most or all of the cancer specimens positive for *MLH1* methylation, but not in those negative for *MLH1* methylation. The difference in CpG methylation for the MCA–RDA products between the methylated and unmethylated groups of patients was thus confirmed in a distinct test set of CRC specimens.

A more detailed inspection of the data in Figure 1, however, indicates that the MCA–RDA products can be separated into three subgroups on the basis of their methylation profiles. The genomic fragments in the first group (*MIG2* to *NPHS2* in Figure 1) were also methylated in $\geq$25% of the paired normal colon tissue samples. The fragment corresponding to *MIG2*, for instance, was methylated in all of the *MLH1* methylation-positive cancer specimens and the respective normal tissue. Methylation of these genomic regions thus probably occurred in each patient before the development of CRC and might be related to the aging process.

The genomic fragments in the second group (*BMP3* to *DPYSL3*) were not methylated in normal colon tissue but were methylated in $\geq$25% of cancer specimens that were negative for *MLH1* methylation. The methylation of these

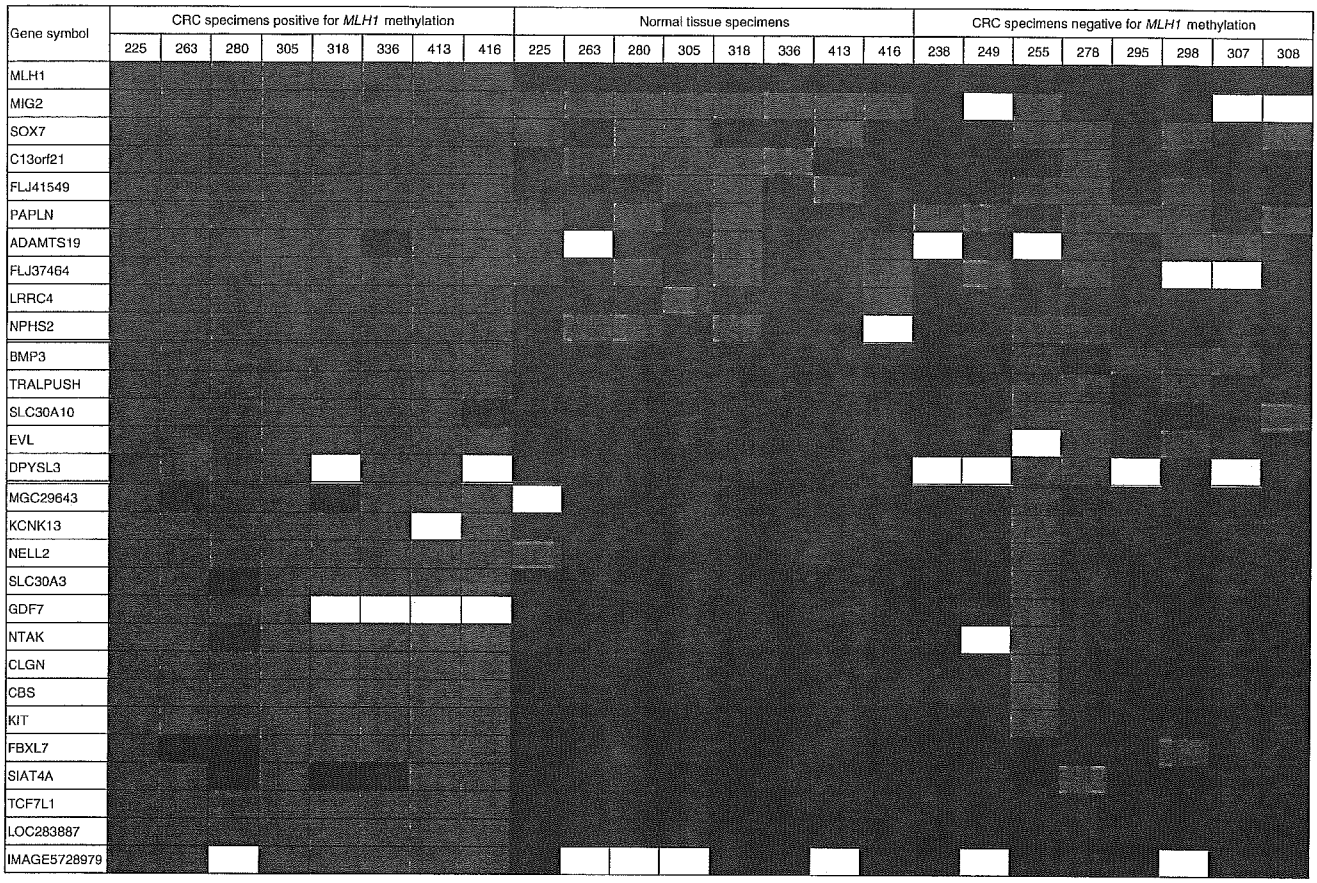

Fig. 1. Gene methylation profiles of CRC specimens. Twenty-eight clones were randomly chosen from the MCA–RDA products of the selected study specimens, and their methylation status (plus that of *MLH1*) was determined by COBRA in CRC specimens positive for the methylation of the *MLH1* promoter ($n = 8$), their paired normal colon tissue samples, and CRC specimens negative for *MLH1* methylation ($n = 8$). Each column represents a clinical specimen (ID numbers are shown), and each row indicates a gene corresponding to an MCA–RDA product. Red box, methylated gene; blue box, unmethylated gene; white box, not examined.

fragments thus appeared to be specific to the cancerous state with a slightly increased prevalence among *MLH1* methylation-positive CRC.

The fragments in the third group (*MGC29643* to *IMAGE5728979*) were methylated in <25% both of normal specimens and of cancer specimens negative for *MLH1* methylation. The methylation of these genes thus appears to be regulated in concert with that of *MLH1*.

Analysis of BMP3

Among the genomic clones analyzed, we first focused on that corresponding to BMP3. BMP3 is a member of the transforming growth factor-β (TGF-β) superfamily of proteins that also includes TGF-β1, TGF-β2, TGF-β3, Mullerian inhibitory substance, BMP2A, BMP2B, BMP6, growth-differentiation factor (GDF) 5, GDF6 and GDF7 (17,18). Members of this protein superfamily exert inhibitory effects on various human cancers through activation of their cognate receptors and SMAD proteins (19,20). BMP2, for instance, induces both the activation of the p38 isoform of mitogen-activated protein kinase and apoptosis in medulloblastoma cells (21). Although little is known of the physiological functions of BMP3, it is possible that this protein also possesses antitumor activity and that its expression is epigenetically regulated in cancer cells. Interestingly, Dai *et al.* (22) have recently reported that *BMP3* promoter is methylated frequently (~50%) in non-small-cell lung carcinoma, which many imply that dysfunction of BMP3 may be commonly involved in the carcinogenesis of a wide range of human tumors.

The COBRA assay revealed that the MCA–RDA clone corresponding to the promoter region of *BMP3* was methylated in CRC specimens that were positive or negative for *MLH1* methylation (Figures 1 and 2A). Further, as shown in Figure 2B, detailed analysis of the methylation status of the *BMP3* promoter by sequencing of DNA fragments after sodium bisulfite treatment revealed extensive hemi- or biallelic methylation of the promoter in CRC specimens positive for *MLH1* methylation (ID nos 225, 318 and 481) but not in one negative for *MLH1* methylation (ID no. 249). CpG methylation throughout the promoter fragment was also evident in CRC cell lines positive (HCT116) or negative (Caco2) for *MLH1* methylation, but not in the *MLH1* methylation-negative line SW480. Together with the COBRA data in Figure 1, these results suggest that the promoter region of *BMP3* is methylated in all clinical specimens and cell lines positive for methylation of the *MLH1* promoter as well as in some specimens and cell lines negative for *MLH1* methylation.

We then examined whether the epigenetic changes in the *BMP3* promoter affected its transcriptional activity. Quantitative real-time RT–PCR analysis revealed that *BMP3*

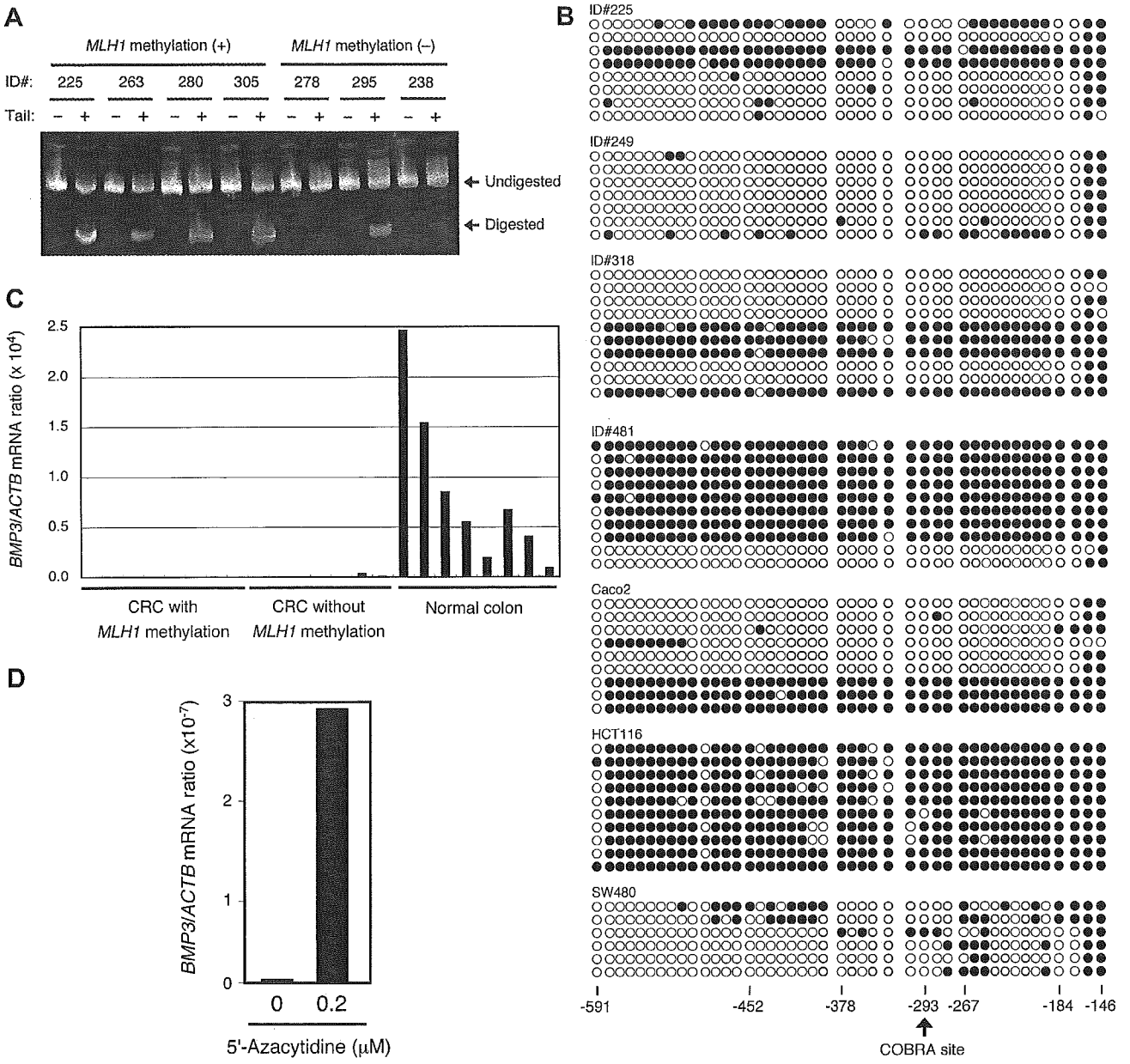

Fig. 2. Methylation status of the *BMP3* promoter and *BMP3* expression in CRC specimens. (**A**) The methylation status of the promoter region of *BMP3* in the indicated clinical specimens was examined by COBRA. Sensitivity of PCR products to digestion with TaiI is indicative of methylation of the CpG island examined. (**B**) Genomic DNA of the indicated clinical specimens and CRC cell lines (Caco2, HCT116 and SW480) was treated with sodium bisulfite, after which the *BMP3* promoter region was amplified by PCR and sequenced. Closed and open circles indicate methylated and unmethylated CpG islands, respectively. The nucleotide positions of the CpG islands (numbered relative to the transcriptional start site) are indicated at the bottom, and the TaiI digestion site for COBRA in (A) is shown by the arrow. (**C**) The level of expression of *BMP3* relative to that of *ACTB* in clinical specimens was determined by quantitative RT–PCR. (**D**) HCT116 cells were incubated for 72 h with or without 0.2 μM of 5′-azacytidine, and were then subjected to RT–PCR analysis for determination of the amount of *BMP3* mRNA relative to that of *ACTB* mRNA.

was transcriptionally silent in CRC specimens positive for *MLH1* methylation (Figure 2C), in which the *BMP3* promoter was also extensively methylated. In contrast, *BMP3* mRNA was abundant in the paired normal colon tissue samples. Although *BMP3* expression was detected in some CRC specimens negative for *MLH1* methylation, the level of expression was greatly reduced compared with that in normal colon tissue. These data thus indicate that *BMP3* transcription is suppressed in most CRCs.

In order to directly examine the relationship between promoter methylation and gene silencing of *BMP3*, HCT116 cells were incubated for 3 days with 5′-azacytidine, an inhibitor of *de novo* methylation of genomic DNA. Interestingly, treatment with 5′-azacytidine markedly induced the amount of *BMP3* mRNA in the cells (Figure 2D) and demethylation of its promoter region as well (data not shown). Therefore, extensive methylation of the *BMP3* promoter region should be directly linked to the suppression of its transcription.

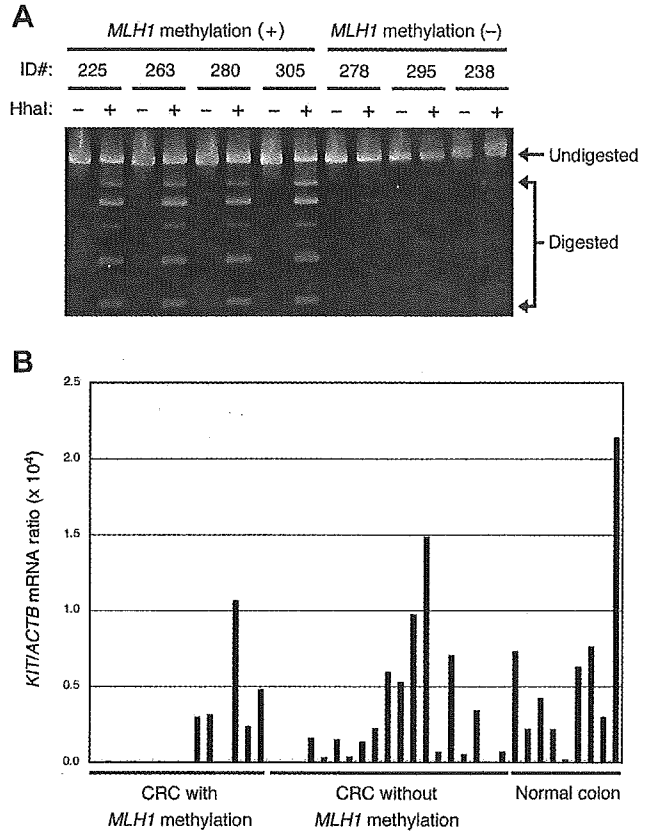

Fig. 3. Methylation status of the *KIT* promoter and *KIT* expression in CRC specimens. (**A**) The methylation status of the promoter region of *KIT* in the indicated clinical specimens was examined by COBRA. Sensitivity of PCR products to digestion with HhaI is indicative of methylation of the CpG island examined. (**B**) The level of expression of *KIT* relative to that of *ACTB* in clinical specimens was determined by quantitative RT–PCR.

Analysis of KIT

KIT encodes a receptor tyrosine kinase for stem cell factor. Point mutations in *KIT* that increase the kinase activity of the encoded protein have been identified in human gastrointestinal stromal tumors (23), suggestive of a causative role for KIT in these tumors. The expression and activation status of KIT in CRCs have been unclear, however (24,25). We therefore analyzed the methylation status of the *KIT* promoter region in our samples.

Methylation of the *KIT* promoter was highly restricted to CRC specimens positive for *MLH1* methylation (Figures 1 and 3A). However, the abundance of *KIT* mRNA did not necessarily match the methylation status of the *KIT* promoter. Despite extensive methylation of the promoter in one CRC sample (ID no.77), for instance, the amount of *KIT* mRNA was relatively high (*KIT/ACTB* mRNA ratio = 4.78×10^{-5}), indicating that promoter methylation might not be a major determinant of transcriptional activity. It is possible, however, that our COBRA analysis revealed the methylation status of a CpG site that is not important for *KIT* transcription.

The mean expression level of *KIT* in the *MLH1* methylation-positive CRC specimens [*KIT/ACTB* mRNA ratio, $1.72 \times 10^{-5} \pm 3.02 \times 10^{-5}$ (mean ± SD)] was significantly lower than that in normal colon tissue ($6.03 \times 10^{-5} \pm 6.29 \times 10^{-5}$; $P = 0.038$, Student's *t*-test). The level of *KIT* expression in CRCs negative for *MLH1* methylation ($2.92 \times 10^{-5} \pm 1.61 \times 10^{-5}$) was also lower than that in normal colon tissue, but this difference was not significant ($P = 0.123$). It is therefore likely that *KIT* is not overexpressed in CRCs.

Discussion

We have screened for genomic fragments whose CpG islands are selectively methylated in CRC specimens positive for methylation of the *MLH1* promoter region. We could readily identify hundreds of genomic fragments with CpG islands that were expected to be differentially methylated between CRCs with or without *MLH1* methylation. Twenty-eight such clones (Table II) were indeed proved to be preferentially methylated in the four CRC specimens positive for *MLH1* methylation compared with the four samples negative for *MLH1* methylation, both of which were used in the original MCA–RDA screening (data not shown).

To verify the selective methylation of these clones, we performed COBRA with a different set of specimens including eight CRCs with *MLH1* methylation and their paired normal tissue samples as well as eight CRCs without *MLH1* methylation. Although all the 28 clones examined were preferentially methylated in the CRC specimens positive for *MLH1* methylation, their methylation profiles among the specimens were not identical, indicating that all CpG methylation observed in MSI-positive CRCs was not specific to this subtype of tumor.

The methylation of certain genomic fragments (14 out of 28 clones examined), however, was highly specific to CRCs that manifested *MLH1* methylation. Almost 50% of the genes were thus methylated in a parallel manner to the CpG methylation of the *MLH1* gene, indicating that a subset of genes is specifically methylated in a subset of CRCs. Our data thus support the existence of CIMP-positive CRCs (14), while it would be mandatory for the better characterization of CIMP-positive tumors to further collect co-methylated genes and to define precisely the hallmark genes for the identification of CIMP (26). It would be interesting to examine whether such clearly defined CIMP is associated with certain clinical manifestations.

Genes corresponding to the co-methylated genomic fragments in our assay included those whose function relates to cell proliferation or differentiation. The predicted structure of NELL2, for example, contains epidermal growth factor (EGF)-like repeats (27), which are present in diverse proteins involved in regulation of the cell cycle, cell proliferation, and developmental processes. NTAK is a member of the EGF family of proteins and is a ligand and activator of ErbB protein tyrosine kinases (28). In addition, GDF7 is a member of the TGF-β superfamily (18), and *TCF7L1* is highly homologous to TCF1 which is a target gene of the Wnt-β-catenin signaling pathway (29), and which plays an important role in CRC carcinogenesis. Aberrant epigenetic regulation of these genes may thus contribute to the pathogenesis or clinical features of CRCs positive for *MLH1* methylation.

The MCA–RDA method thus proved to be highly effective for the identification of differentially methylated genes among fresh clinical specimens. Given the high fidelity of this approach, it is likely that a large number of genes (or genomic fragments) are methylated in CRCs in concert with methylation of the *MLH1* promoter. Our study provides a basis for further characterization of the molecular pathogenesis of CRCs classified as MSI. Together with the results of other

studies (7,30), it also suggests the possibility of development of a stratification scheme for CRCs based on genome methylation profile.

Supplementary material

Supplementary material can be found at: http://www.carcin.oxfordjournals.org/

Acknowledgements

This study was supported by a Grant-in-Aid for Third-Term Comprehensive Control Research for Cancer from the Ministry of Health, Labor, and Welfare of Japan, and by a grant for 'High-Tech Research Center' Project for Private Universities: Matching Fund Subsidy (2002–2006) from the Ministry of Education, Culture, Sports, Science, and Technology of Japan.

Conflict of Interest Statement: None declared.

References

1. Lengauer,C., Kinzler,K.W. and Vogelstein,B. (1998) Genetic instabilities in human cancers. *Nature*, **396**, 643–649.
2. Kinzler,K.W. and Vogelstein,B. (1996) Lessons from hereditary colorectal cancer. *Cell*, **87**, 159–170.
3. Ionov,Y., Peinado,M.A., Malkhosyan,S., Shibata,D. and Perucho,M. (1993) Ubiquitous somatic mutations in simple repeated sequences reveal a new mechanism for colonic carcinogenesis. *Nature*, **363**, 558–561.
4. Fishel,R., Lescoe,M.K., Rao,M.R., Copeland,N.G., Jenkins,N.A., Garber,J., Kane,M. and Kolodner,R. (1993) The human mutator gene homolog *MSH2* and its association with hereditary non-polyposis colon cancer. *Cell*, **75**, 1027–1038.
5. Bronner,C.E., Baker,S.M., Morrison,P.T. *et al.* (1994) Mutation in the DNA mismatch repair gene homologue *hMLH1* is associated with hereditary non-polyposis colon cancer. *Nature*, **368**, 258–261.
6. Papadopoulos,N., Nicolaides,N.C., Wei,Y.F. *et al.* (1994) Mutation of a mutL homolog in hereditary colon cancer. *Science*, **263**, 1625–1629.
7. Toyota,M., Ahuja,N., Ohe-Toyota,M., Herman,J.G., Baylin,S.B. and Issa,J.P. (1999) CpG island methylator phenotype in colorectal cancer. *Proc. Natl Acad. Sci. USA*, **96**, 8681–8686.
8. Miyakura,Y., Sugano,K., Konishi,F., Ichikawa,A., Maekawa,M., Shitoh,K., Igarashi,S., Kotake,K., Koyama,Y. and Nagai,H. (2001) Extensive methylation of h*MLH1* promoter region predominates in proximal colon cancer with microsatellite instability. *Gastroenterology*, **121**, 1300–1309.
9. Cunningham,J.M., Christensen,E.R., Tester,D.J., Kim,C.Y., Roche,P.C., Burgart,L.J. and Thibodeau,S.N. (1998) Hypermethylation of the h*MLH1* promoter in colon cancer with microsatellite instability. *Cancer Res.*, **58**, 3455–3460.
10. Veigl,M.L., Kasturi,L., Olechnowicz,J. *et al.* (1998) Biallelic inactivation of h*MLH1* by epigenetic gene silencing, a novel mechanism causing human MSI cancers. *Proc. Natl Acad. Sci. USA*, **95**, 8698–8702.
11. Koinuma,K., Shitoh,K., Miyakura,Y. *et al.* (2004) Mutations of BRAF are associated with extensive h*MLH1* promoter methylation in sporadic colorectal carcinomas. *Int. J. Cancer*, **108**, 237–242.
12. Wang,L., Cunningham,J.M., Winters,J.L., Guenther,J.C., French,A.J., Boardman,L.A., Burgart,L.J., McDonnell,S.K., Schaid,D.J. and Thibodeau,S.N. (2003) BRAF mutations in colon cancer are not likely attributable to defective DNA mismatch repair. *Cancer Res.*, **63**, 5209–5212.
13. Oliveira,C., Pinto,M., Duval,A. *et al.* (2003) BRAF mutations characterize colon but not gastric cancer with mismatch repair deficiency. *Oncogene*, **22**, 9192–9196.
14. Issa,J.P. (2004) Opinion: CpG island methylator phenotype in cancer. *Nat. Rev. Cancer*, **4**, 988–993.
15. Toyota,M., Ho,C., Ahuja,N., Jair,K.W., Li,Q., Ohe-Toyota,M., Baylin,S.B. and Issa,J.P. (1999) Identification of differentially methylated sequences in colorectal cancer by methylated CpG island amplification. *Cancer Res.*, **59**, 2307–2312.
16. Xiong,Z. and Laird,P.W. (1997) COBRA: a sensitive and quantitative DNA methylation assay. *Nucleic Acids Res.*, **25**, 2532–2534.
17. Hogan,B.L. (1996) Bone morphogenetic proteins: multifunctional regulators of vertebrate development. *Genes Dev.*, **10**, 1580–1594.
18. Davidson,A.J., Postlethwait,J.H., Yan,Y.L., Beier,D.R., van Doren,C., Foernzler,D., Celeste,A.J., Crosier,K.E. and Crosier,P.S. (1999) Isolation of zebrafish gdf7 and comparative genetic mapping of genes belonging to the growth/differentiation factor 5, 6, 7 subgroup of the TGF-beta superfamily. *Genome Res.*, **9**, 121–129.
19. Miyazono,K., Kusanagi,K. and Inoue,H. (2001) Divergence and convergence of TGF-beta/BMP signaling. *J. Cell. Physiol.*, **187**, 265–276.
20. Bottinger,E.P., Jakubczak,J.L., Haines,D.C., Bagnall,K. and Wakefield,L.M. (1997) Transgenic mice overexpressing a dominant-negative mutant type II transforming growth factor beta receptor show enhanced tumorigenesis in the mammary gland and lung in response to the carcinogen 7,12-dimethylbenz-[a]-anthracene. *Cancer Res.*, **57**, 5564–5570.
21. Hallahan,A.R., Pritchard,J.I., Chandraratna,R.A., Ellenbogen,R.G., Geyer,J.R., Overland,R.P., Strand,A.D., Tapscott,S.J. and Olson,J.M. (2003) BMP-2 mediates retinoid-induced apoptosis in medulloblastoma cells through a paracrine effect. *Nat. Med.*, **9**, 1033–1038.
22. Dai,Z., Popkie,A.P., Zhu,W.G. *et al.* (2004) Bone morphogenetic protein 3B silencing in non-small-cell lung cancer. *Oncogene*, **23**, 3521–3529.
23. Hirota,S., Isozaki,K., Moriyama,Y. *et al.* (1998) Gain-of-function mutations of c-kit in human gastrointestinal stromal tumors. *Science*, **279**, 577–580.
24. Bellone,G., Silvestri,S., Artusio,E., Tibaudi,D., Turletti,A., Geuna,M., Giachino,C., Valente,G., Emanuelli,G. and Rodeck,U. (1997) Growth stimulation of colorectal carcinoma cells via the c-kit receptor is inhibited by TGF-beta 1. *J. Cell. Physiol.*, **172**, 1–11.
25. Sammarco,I., Capurso,G., Coppola,L. *et al.* (2004) Expression of the proto-oncogene c-KIT in normal and tumor tissues from colorectal carcinoma patients. *Int. J. Colorectal Dis.*, **19**, 545–553.
26. Yamashita,K., Dai,T., Dai,Y., Yamamoto,F. and Perucho,M. (2003) Genetics supersedes epigenetics in colon cancer phenotype. *Cancer Cell*, **4**, 121–131.
27. Watanabe,T.K., Katagiri,T., Suzuki,M., Shimizu,F., Fujiwara,T., Kanemoto,N., Nakamura,Y., Hirai,Y., Maekawa,H. and Takahashi,E. (1996) Cloning and characterization of two novel human cDNAs (NELL1 and NELL2) encoding proteins with six EGF-like repeats. *Genomics*, **38**, 273–276.
28. Higashiyama,S., Horikawa,M., Yamada,K. *et al.* (1997) A novel brain-derived member of the epidermal growth factor family that interacts with ErbB3 and ErbB4. *J. Biochem.*, **122**, 675–680.
29. Castrop,J., van Norren,K. and Clevers,H. (1992) A gene family of HMG-box transcription factors with homology to TCF-1. *Nucleic Acids Res.*, **20**, 611.
30. Toyota,M., Ohe-Toyota,M., Ahuja,N. and Issa,J.P. (2000) Distinct genetic profiles in colorectal tumors with or without the CpG island methylator phenotype. *Proc. Natl Acad. Sci. USA*, **97**, 710–715.

Received April 29, 2005; revised July 10, 2005; accepted July 12, 2005

ELSEVIER

European Journal of Cancer 41 (2005) 2170–2175

European Journal of Cancer

www.ejconline.com

Retroviral expression screening of oncogenes in pancreatic ductal carcinoma

Hiroyuki Kisanuki [a], Young Lim Choi [a], Tomoaki Wada [a], Ryozo Moriuchi [b], Shin-ichiro Fujiwara [a], Ruri Kaneda [a], Koji Koinuma [a], Madoka Ishikawa [a], Shuji Takada [a], Yoshihiro Yamashita [a], Hiroyuki Mano [a,c,*]

[a] *Division of Functional Genomics, Jichi Medical School, 3311-1 Yakushiji, Kawachigun, Tochigi 329-0498, Japan*
[b] *Department of Molecular Microbiology and Immunology, Nagasaki University Graduate School of Medicine, Nagasaki, Japan*
[c] *CREST, Japan Science and Technology Agency, Saitama, Japan*

Received 4 March 2005; received in revised form 3 May 2005; accepted 10 May 2005
Available online 25 August 2005

Abstract

Pancreatic ductal carcinoma (PDC) remains one of the most intractable malignancies in humans. In order to clarify the molecular events underlying the carcinogenesis in PDC, we constructed a retroviral cDNA expression library from a PDC cell line, and used it to screen transforming genes in PDC by a focus formation assay with mouse 3T3 fibroblasts. We could obtain a total of 30 transformed cell foci in the screening, and one of the cDNA inserts harvested from such cell clones turned out to encode a wild-type human ARAF1. Unexpectedly, a long terminal repeat-driven overexpression of *ARAF1* mRNA was confirmed to induce transformed foci in fibroblasts. The oncogenic potential of ARAF1 was examined by injecting the transformed fibroblasts into athymic nude mice. Importantly, *ARAF1* mRNA was highly expressed in pancreatic ductal cell specimens purified from patients with PDC. These results have unveiled the transforming potential of ARAF1 protein, and also suggest that quantity of intracellular ARAF1 may be important in carcinogenesis of various human cancers.

Keywords: Pancreatic ductal cell carcinoma; Retrovirus; ARAF1; cDNA library; Oncogene

1. Introduction

Pancreatic ductal carcinoma (PDC) originates from pancreatic ductal cells, 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. At the time of diagnosis, most patients are no longer candidates for surgical resection, and, even in individuals who do undergo such surgery, the 5-year survival rate is only 20–30% [1].

Vast efforts have been made to elucidate molecular events responsible for the carcinogenesis of PDC. Mutations of *TP53* gene can be, for instance, found in PDC specimens [3], and in the intraductal *in situ* regions as well [4]. Similarly, inactivation has been found for other tumour-suppressor genes, such as DPC, RB1 and p16 [5].

As for oncogenes, activating mutations in the *KRAS2* gene has been reported to be frequently associated with PDC [6]. The same *KRAS2* mutations could be, however, found in the samples for chronic pancreatitis [7], making their pathogenetic role uncertain. Additionally, an increased telomerase activity was shown to be present only in PDC, but not in nonmalignant pancreatic disorders [8]. Again, however, others could detect an elevated telomerase activity in chronic pancreatitis and normal

* Corresponding author. Tel.: +81 285 58 7449; fax: +81 285 44 7322.
E-mail address: hmano@jichi.ac.jp (H. Mano).

0959-8049/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.ejca.2005.05.014

pancreas [9]. Therefore, it is yet to be revealed which transforming genes truly promote a clonal growth of pancreatic ductal cells.

For an efficient isolation of tumour-promoting genes in PDC, it would be desirable to conduct functional screening based on transforming ability. Focus formation assays with mouse 3T3 fibroblasts have been highly successful for the identification of oncogenes in human cancer [10]. In such screening, genomic DNA is isolated from cancer specimens, and used to transfect 3T3 cells to obtain transformed cell foci. It should be noted, however, that, since expression of any genes in these experiments are driven by their own promoters/enhancers, oncogenes in PDC can exert their effects in 3T3 cells only when the promoter/enhancer regions of such genes are active in fibroblasts, which is not always guaranteed.

To ensure the sufficient expression of cDNAs in 3T3 cells, their transcription should be regulated by an exogenous promoter fragment. We have therefore constructed a retroviral cDNA expression library from a PDC cell line, MiaPaCa-2, which was used to infect 3T3 cells. In the preparation of cDNA library, we further took advantage of the SMART polymerase chain reaction (PCR) system (Clontech, Palo Alto, CA), which preferentially amplifies full-length cDNAs. A focus formation assay with the library resulted in an identification of a transforming *ARAF1* gene.

2. Materials and methods

2.1. Cells and culture

MiaPaCa-2, Capan-2, PANC-1, 3T3, and BOSC23 [11] cell lines were obtained from American Type Culture Collection, and maintained in Dulbecco's modified Eagle medium/F12 (DMEM/F12; Invitrogen, Carlsbad, CA) containing 10% fetal bovine serum (Invitrogen) and 2 mM L-glutamine.

The fresh clinical specimens were obtained from patients who gave informed consent, and the study was approved by the institutional review board of Jichi Medical School. Cells were collected from the pancreatic juice by centrifugation, labeled with anti-MUC1 antibody [12] (Novocastra Laboratories, Newcastle upon Tyne, UK), and subjected to chromatography on a miniMACS magnetic cell separation column (Miltenyi Biotec, Auburn, CA) [13]. The purity of the resultant $MUC1^+$ cell fraction was confirmed by staining with Wright Giemmsa solutions and microscopy examination for each case (data not shown).

2.2. Retrovirus library construction

Total RNA was extracted from MiaPaCa-2 cells by an RNeasy Mini column with RNase-free DNase (Qiagen, Valencia, CA), and the first strand cDNA was synthesized by PowerScript reverse transcriptase (Clontech) with SMART IIA oligonucleotide and CDS primer IIA (both from Clontech). The cDNAs were then amplified for 12 cycles with 5′ PCR primer IIA according to the instruction of the SMART PCR cDNA synthesis kit (Clontech) except a substitution of LA Taq polymerase (Takara Bio, Shiga, Japan) for Advantage 2 DNA polymerase provided with the kit. Resultant cDNAs were treated with proteinase K, blunt-ended by T4 DNA polymerase, and ligated to the BstXI-adaptor (Invitrogen). Unbound adaptors were removed through the cDNA size fractionation column (Invitrogen), and the cDNAs were finally ligated to the pMXS retroviral plasmid (a kind gift of Dr. T. Kitamura at Institute of Medical Science, University of Tokyo) [14] digested with BstXI. The pMXS-cDNA plasmids were introduced into ElecroMax DH10B cells (Invitrogen) with electroporation.

2.3. Focus formation assay

Generation of recombinant retroviral library and focus formation assay was conducted as described previously [15]. Briefly, BOSC23 cells were transfected with Lipfectamin reagent (Invitrogen) and 2 μg of retroviral plasmid together with 0.5 μg of pGP plasmid 0.5 μg of pE-eco plasmid (both from Takara Bio). Two days after the transfection, polybrene (Sigma, St. Louis, MI) was added to the culture supernatant at a concentration of 4 μg/ml, and the supernatant was subsequently used to infect 3T3 cells for 48 h. For the focus formation assay, the culture medium of 3T3 cells was then changed to DMEM-high glucose (Invitrogen) supplemented with 5% calf serum and 2 mM L-glutamine. Transformed foci were picked up after 3 weeks of culture. Genomic DNA was extracted from each transformed focus, and was subjected to PCR with 5′ PCR primer IIA (Clontech) and LA Taq polymerase for 50 cycles of 98 °C for 20 s and 68 °C for 6 min. Amplified genome fragments were purified for nucleotide sequencing. For tumorgenicity assay in nude mice, transformed 3T3 cells were injected into each shoulder of nu/nu Balb-c mice (6 weeks old). Tumour formation was assessed after 4 weeks.

2.4. "Real-time" RT-PCR

Portions of oligo(dT)-primed cDNA were subjected to PCR with a QuantiTect SYBR Green PCR Kit (Qiagen). The amplification protocol was comprised of 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 7900HT sequence detection system (PE Applied Biosystems, Foster City, CA), thereby allowing determination of the threshold cycle (C_T) at which exponential amplification of products

begins. The C_T values for cDNAs corresponding to the β-actin gene (*ACTB*) and to the *ARAF1* gene were used to calculate the abundance of the latter mRNA relative to that of the former. The oligonucleotide primers for PCR were as follows: 5′-CCATCATGAAGTGTGACGTGG-3′ and 5′-GTCCGCCTAGAAGCATTTGCG-3′ for *ACTB*, and 5′-ACTACCTCCATGCCAAGAACATCA-3′ and 5′-GACGTCTGACTGGAAGCTGTAGGG-3′ for *ARAF1*.

3. Results

3.1. Screening with focus formation assay

From mRNA of MiaPaCa-2 cells, full-length cDNAs were selectively amplified and ligated to a retroviral vector pMXS. We could obtain a total of 2.1×10^6 colony forming units (cfu) of independent plasmid clones. Thirty clones were randomly selected from the library, and examined for the incorporated cDNAs. Twenty-seven (90%) out of the 30 clones contained inserts with an average length of 2.05 kbp.

By introducing the plasmid DNA into a packaging cell line, we generated a recombinant ecotropic retrovirus library that was subsequently used to infect mouse 3T3 fibroblasts. Infection experiments were repeated for a total of 6 times. After 3 weeks of culture, 30 transformed foci were observed (Fig. 1(b)). No foci could be found among the cells infected with an empty virus (Fig. 1(a)), while numerous foci were easily identified in the cells infected with a virus expressing v-Ras oncoprotein (Fig. 1(c)).

Each focus was isolated, expanded independently, and was subject to the extraction of genomic DNA. We then tried to recover retroviral inserts from the genomic DNA by PCR amplification with the primer used originally to amplify the cDNAs in the construction of the library. In most cases, two to three DNA fragments were recovered from each genome (Fig. 2(a)), implying a multiple retroviral infection on the recipient 3T3 cells.

We obtained a total of 56 cDNA fragments by PCR, all of which were subjected to nucleotide sequencing from both ends. Screening of the cDNA sequences against human genome sequence database assembled as of July 2003 by the Genome Bioinformatics Group of the University of California at Santa Cruz (http://genome.ucsc.edu) revealed that the 56 fragments correspond to 13 independent genes, eleven of which could be matched at >95% identity to the human genome sequence (Table 1). Among the 11 genes, 7 of them were known genes while the rest 4 were unknown. Each cDNA clone was ligated to pMXS in both directions, and a recombinant retrovirus was generated from each resultant plasmid. Transforming ability of the cDNAs was thus confirmed by a focus formation assay with the recombinant virus.

Focus formation assays were conducted for the 26 independent viruses (all 13 independent genes for both directions), and one of them, expressing ARAF1 protein (GenBank accession number, NM_001654), gave transformed foci in repeated experiments. ARAF1 belongs to

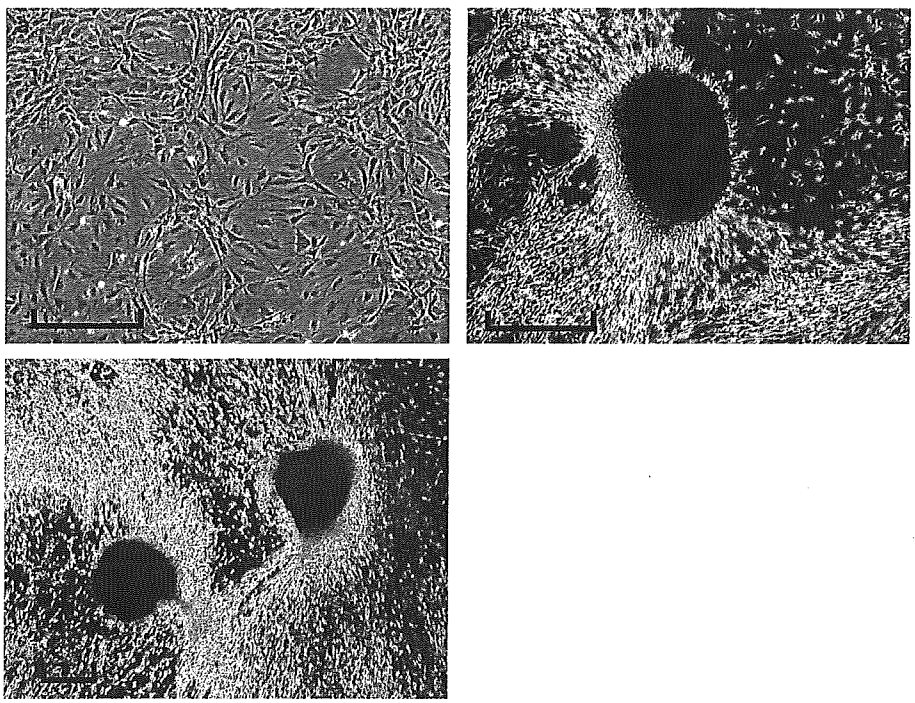

Fig. 1. Focus formation assay with retroviral library. Mouse 3T3 cells were infected with (a) an empty virus, (b) retroviruses from the MiaPaCa-2 library, or (c) a retrovirus expressing v-Ras as a positive control. Pictures were taken after 3 weeks of culture. Scale bar, 100 μm.

Table 1
cDNAs isolated from 3T3 transformants

Clone #	Gene symbol	GenBank number	Covering full ORF
1	Unknown	AK026325	ND
2	No homologues sequences	ND	ND
3	Unknown	AF318370	Yes
4	PITPNM1	NM_004910	No
5	Unknown	BC022099	Yes
6	DECR2	NM_020664	Yes
7	ARAF1	NM_001654	Yes
8	BCLG	NM_138724	No
9	No homologues sequences	ND	ND
10	Unknown	AL607122	ND
11	JAG1	NM_000214	Yes
12	MRPL43	NM_176792	Yes
13	PLOD3	NM_001084	No

ORF, open reading frame; ND, not determined.

the RAF family of serine/threonine kinases, and phosphorylates MEK1 [16]. It had not been known whether an overexpression of wild ARAF1 protein has a transforming activity.

3.2. *ARAF1* as an oncogene

We thus determined the whole nucleotide sequence of our *ARAF1* cDNA (cDNA clone ID #7). The sequence is 2441 bp, and contains an open reading frame (spanning nucleotide position 126–1943) encoding a protein of 606 amino acids, which is identical to ARAF1 (Fig. 2(b)). Within the protein-coding region, there is only one nucleotide mutation compared to the published *ARAF1* cDNA sequence; a "T" at nucleotide position 1550 in the reported sequence (NM_001654) is replaced with a "C" in our sequence. The codon sequence "TTG" at the amino acid position 450 of ARAF1 is thus changed to "CTG" in our cDNA. However, both codons encode the same leucine residue, and thus the mutation does not affect the protein sequence.

To confirm that mere overexpression of wild ARAF1 protein has a transforming activity, we repeated the focus formation assay with the retrovirus generated from our *ARAF1* cDNA. As shown in Fig. 2(c), the recombinant virus reproducibly induced transformed foci (30–50 foci per microgram of the input plasmid) in the recipient 3T3 cells. The transforming ability of ARAF1 was also tested by the tumorgenicity assay with athymic nude mice. The 3T3 cells infected with the empty virus or retrovirus expressing ARAF1 or v-Ras were inoculated subcutaneously into nude mice. As shown in Fig. 2(d), tumour formation was observed in all mice for the latter 2 cases, arguing that ARAF1 has oncogenic potential.

3.3. Expression of *ARAF1* in PDC

We finally measured the expression level of *ARAF1* mRNA in PDC by the quantitative real-time reverse

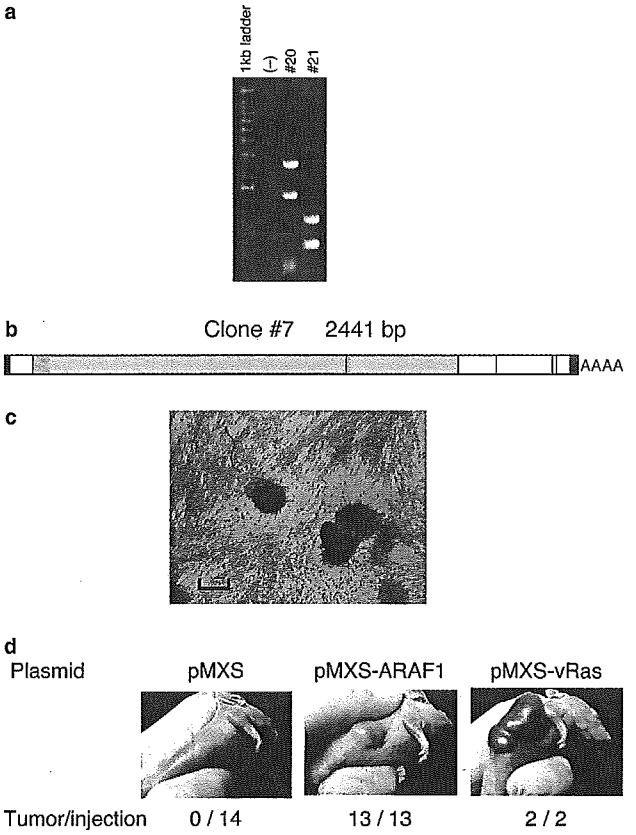

Fig. 2. Identification of the transforming *ARAF1* gene. (a) Genomic DNAs isolated from the transformed foci (cell clone ID #20 and #21) were subjected to PCR. A PCR without DNA template was also conducted as a negative control (−). DNA size markers (1 kb DNA ladder; Invitrogen) are electrophoresed at the left. (b) A 3T3 clone yielded a PCR product of 2441 bp long (cDNA clone ID #7). The cDNA has a protein-coding region (gray) for 606 amino acids that was identical to human ARAF1 protein. Nucleotides that did not match the published *ARAF1* cDNA are indicated by red lines. (c) Our ARAF1 cDNA was ligated to pMXS, and used to generate recombinant virus. Infection with the virus induced multiple transformed foci in 3T3 cells. Scale bar, 100 μm. (d) 3T3 cells (5×10^5) were cultured for two days with retrovirus made from pMXS, pMXS-ARAF1 or pMXS-vRas plasmid, and were injected subcutaneously into nu/nu mice. Tumour formation was examined after 4 weeks.

transcription (RT)-PCR method. As shown in Fig. 3, all 3 PDC cell lines express similar amounts of *ARAF1* mRNA. In addition, we quantified *ARAF1* mRNA in human clinical specimens. Pancreatic juice from patients with PDC contains cancer cells (transformed ductal cells) in addition to normal ductal cells and blood cells. The former two fractions were purified, by an affinity column for MUC1 surface protein [12], from pancreatic juice of PDC patients ($n = 14$). Such purified fractions should be highly enriched for PDC cells [13]. Similar MUC1-positive fractions were also purified from the pancreatic juice of healthy individuals ($n = 7$). Quantification of *ARAF1* mRNA revealed that its mRNA level was highly elevated in six out of the 14 patient, but not in the specimens from healthy individuals. These data

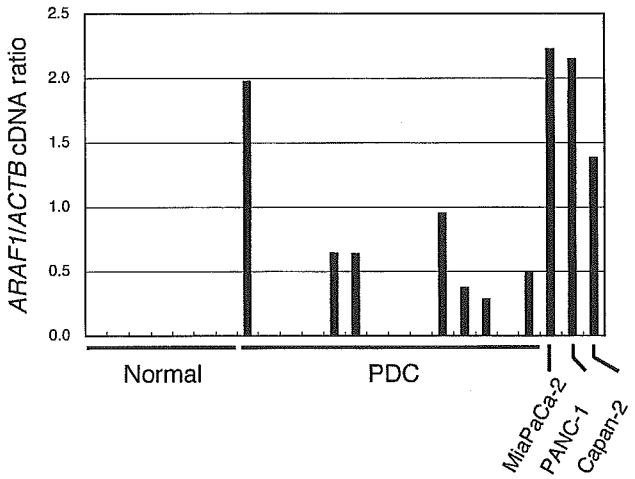

Fig. 3. Quantification of *ARAF1* mRNA. Complementary DNA was prepared from PDC cell lines (MiaPaCa-2, PANC-1 and Capan-2) or $MUC1^+$ ductal cell preparations purified from healthy individuals (Normal) or patients with PDC, and were then subjected to real-time RT-PCR analysis with primers specific for the *ARAF1* or *ACTB* genes. The ratio of the abundance of *ARAF1* mRNA to that of *ACTB* mRNA was calculated as 2^n, where n is the C_T value for *ACTB* cDNA minus the C_T value for *ARAF1* cDNA.

indicate that the transcription of *ARAF1* is selectively activated in PDC cells.

4. Discussion

In this manuscript, we have constructed a retroviral cDNA expression library of PDC. Since 90% (27/30) of the viral plasmids carried cDNA inserts and since the overall clone number was >2 millions, our library should cover nearly all transcriptome in MiaPaCa-2 cells.

RAF family is composed of RAF1, ARAF1 and BRAF in humans. All these serine/threonine kinases are believed to act downstream of RAS-family proteins, and to phosphorylate and regulate downstream MAP kinase kinases (MAPKKs). Many studies have revealed transforming potentials of RAF family proteins. RAF was originally identified as a cellular homologue of viral oncoprotein, v-Raf [17]. Deletion of amino-terminal regions unmasks the transforming ability of RAF1 [18] and ARAF1 [19]. On the other hand, somatic point mutations have been found in the *BRAF* gene among clinical specimens of colorectal carcinoma [20]. Such mutations were shown to induce transforming activity in BRAF protein. In contrast to *BRAF*, somatic point mutations are rarely found in *ARAF1* gene [21].

Overexpression of wild forms of RAF1 or BRAF failed to exert a transforming activity [18,20]. Although deletion/truncation of amino terminal regions of ARAF1 induced transformed foci in 3T3 fibroblasts [22] and abrogated cytokine-dependency in hematopoietic cells [19], it has not been tested whether wild ARAF1 protein has transforming potential. In this manuscript, however, it was unexpectedly revealed that a long terminal repeat-driven expression of ARAF1 induces transformed foci in 3T3 cells, which subsequently generated tumours in immunocompromised mice. Therefore, it has been unveiled here that an overexpression of wild ARAF1 is directly linked with cellular transformation process.

These data also indicate the importance of measuring protein/mRNA amounts of ARAF1 in various human cancers. In this context, it was interesting to find a high expression of *ARAF1* mRNA in fresh clinical specimens of PDC. Our findings shed new light on the understanding of RAF family kinases, and open up the possibility that ARAF is involved in carcinogenesis in human cancers through a previously unexpected mechanism.

Conflict of interest statement

None declared.

Acknowledgments

This work was supported in part by a grant for Third-Term Comprehensive Control Research for Cancer from the Ministry of Health, Labor, and Welfare of Japan; and a grant for "High-Tech Research Center" Project for Private Universities: Matching Fund Subsidy (2002–2006) from the Ministry of Education, Culture, Sports, Science, and Technology of Japan.

References

1. Bornman PC, Beckingham IJ. Pancreatic tumours. *Br Med J* 2001, **322**, 721–723.
2. Rosewicz S, Wiedenmann B. Pancreatic carcinoma. *Lancet* 1997, **349**, 485–489.
3. Goggins M, Kern SE, Offerhaus JA, *et al.* Progress in cancer genetics: lessons from pancreatic cancer. *Ann Oncol* 1999, **10**(Suppl. 4), 4–8.
4. Barton CM, Staddon SL, Hughes CM, *et al.* Abnormalities of the p53 tumour suppressor gene in human pancreatic cancer. *Br J Cancer* 1991, **64**, 1076–1082.
5. Bramhall SR. The use of molecular technology in the differentiation of pancreatic cancer and chronic pancreatitis. *Int J Pancreatol* 1998, **23**, 83–100.
6. Tada M, Omata M, Ohto M. Clinical application of ras gene mutation for diagnosis of pancreatic adenocarcinoma. *Gastroenterology* 1991, **100**, 233–238.
7. Yanagisawa A, Ohtake K, Ohashi K, *et al.* Frequent c-Ki-ras oncogene activation in mucous cell hyperplasias of pancreas suffering from chronic inflammation. *Cancer Res* 1993, **53**, 953–956.
8. Hiyama E, Kodama T, Shinbara K, *et al.* Telomerase activity is detected in pancreatic cancer but not in benign tumours. *Cancer Res* 1997, **57**, 326–331.
9. Uehara H, Nakaizumi A, Tatsuta M, *et al.* Diagnosis of pancreatic cancer by detecting telomerase activity in pancreatic

juice: comparison with K-ras mutations. *Am J Gastroenterol* 1999, **94**, 2513–2518.
10. Aaronson SA. Growth factors and cancer. *Science* 1991, **254**, 1146–1153.
11. Pear WS, Nolan GP, Scott ML, *et al.* Production of high-titer helper-free retroviruses by transient transfection. *Proc Natl Acad Sci USA* 1993, **90**, 8392–8396.
12. Terada T, Ohta T, Sasaki M, *et al.* Expression of MUC apomucins in normal pancreas and pancreatic tumours. *J Pathol* 1996, **180**, 160–165.
13. 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**, 263–270.
14. Onishi M, Kinoshita S, Morikawa Y, *et al.* Applications of retrovirus-mediated expression cloning. *Exp Hematol* 1996, **24**, 324–329.
15. Yoshizuka N, Moriuchi R, Mori T, *et al.* An alternative transcript derived from the trio locus encodes a guanosine nucleotide exchange factor with mouse cell-transforming potential. *J Biol Chem* 2004, **279**, 43998–44004.
16. Wu X, Noh SJ, Zhou G, *et al.* Selective activation of MEK1 but not MEK2 by A-Raf from epidermal growth factor-stimulated Hela cells. *J Biol Chem* 1996, **271**, 3265–3271.
17. Bonner TI, Kerby SB, Sutrave P, *et al.* Structure and biological activity of human homologs of the raf/mil oncogene. *Mol Cell Biol* 1985, **5**, 1400–1407.
18. Ishikawa F, Sakai R, Ochiai M, *et al.* Identification of a transforming activity suppressing sequence in the c-raf oncogene. *Oncogene* 1988, **3**, 653–658.
19. Hoyle PE, Moye PW, Steelman LS, *et al.* Differential abilities of the Raf family of protein kinases to abrogate cytokine dependency and prevent apoptosis in murine hematopoietic cells by a MEK1-dependent mechanism. *Leukemia* 2000, **14**, 642–656.
20. Davies H, Bignell GR, Cox C, *et al.* Mutations of the BRAF gene in human cancer. *Nature* 2002, **417**, 949–954.
21. Lee JW, Soung YH, Kim SY, *et al.* Mutational analysis of the ARAF gene in human cancers. *APMIS* 2005, **113**, 54–57.
22. Beck TW, Huleihel M, Gunnell M, *et al.* The complete coding sequence of the human A-raf-1 oncogene and transforming activity of a human A-raf carrying retrovirus. *Nucleic Acids Res* 1987, **15**, 595–609.

CMLS, Cell. Mol. Life Sci. 62 (2005) 2140–2146
1420-682X/05/182140-7
DOI 10.1007/s00018-005-5270-5

CMLS Cellular and Molecular Life Sciences

Research Article

Regulation of *Amh* during sex determination in chickens: *Sox* gene expression in male and female gonads

S. Takada[a, b], H. Mano[b] and P. Koopman[a, *]

[a] Institute for Molecular Bioscience, The University of Queensland, Brisbane, Queensland 4072 (Australia), Fax: +61 7 3346 2101, e-mail: p.koopman@imb.uq.edu.au
[b] Division of Functional Genomics, Jichi Medical School, 3311-1 Yakushiji, Minamikawachimachi, Kawachigun, Tochigi 329-0498 (Japan)

Received 18 June 2005; received after revision 22 June 2005; accepted 27 June 2005
Online First 26 August 2005

Abstract. During mammalian sexual development, the SOX9 transcription factor up-regulates expression of the gene encoding anti-Müllerian hormone (AMH), but in chickens, *Sox9* gene expression reportedly occurs after the onset of Amh expression. Here, we examined expression of the related gene *Sox8* in chicken embryonic gonads during the sex-determining period. We found that *cSox8* is expressed at similar levels in both sexes at embryonic day 6 and 7, and only at the anterior tip of the gonad, suggesting that SOX8 is not responsible for the sex-specific increase in *cAmh* gene expression at these stages. We also found that several other chicken *Sox* genes (*cSox3*, *cSox4* and *cSox11*) are expressed in embryonic gonads, but at similar levels in both sexes. Our data suggest that the molecular mechanisms involved in the regulation of *Amh* genes of mouse and chicken are not conserved, despite similar patterns of *Amh* expression in both species.

Key words. Sex determination; Amh; chicken; testis.

Even though molecular mechanisms of patterning and morphogenesis are surprisingly well conserved during metazoan evolution, mechanisms governing sex determination and gonadal development are diverse, even among vertebrates. In mammals, the heterogametic pairing of sex chromosomes (XY) results in male development. The mouse and human sex-determining gene, *Sry/SRY* (sex-determining region on Y chromosome), has been identified [1, 2] and is known to cause the bipotential gonad to differentiate into a testis [3]. However, *Sry* is not a conserved sex-determining gene as it exists only in mammals [4–7]. In contrast to mammals, in birds males are homogametic (ZZ) and females are heterogametic for the sex chromosomes (ZW). Whether avian sex is determined by a master female-determining gene on the W chromosome or by Z chromosome gene dosage is still unclear [8].

There are reports for the mouse that several genes are expressed predominantly in the developing testis but not in the ovary and, therefore, are likely to be important for male sex determination and differentiation. Some of these genes, such as *Amh* (anti-Müllerian hormone) and *Sox9* [9–12] are expressed similarly in mouse and chicken gonads, suggesting that there could be some degree of similarity between the molecular pathways and sexual development in chicken and mouse. However, some differences are evident. For example, *Sf1* (steroidogenic factor 1) and *Gata4* are predominantly expressed in the male gonad in mice [13, 14], while in chicken expression levels of both genes are similar between male and female gonads [12, 15].

* Corresponding author.

Sox genes represent a family related by the *Sry*-type high-mobility group (HMG) box, and function as transcription factors in various developmental processes through binding to a conserved core DNA sequence [16]. Twenty *Sox* genes have been identified in mouse and human, and are classified by their HMG box sequences into subgroups A–H [17, 18]. The expression pattern of each gene tends to be conserved in mouse and chicken. Among them, *Sox9* (group E) is known to act as a sex-determining gene. Mutations of human *SOX9* cause campomelic dysplasia, a severe skeletal malformation syndrome associated in most cases with XY sex reversal [19, 20], and ectopic expression of *Sox9* in XX mouse fetal gonads induces testis formation [21].

AMH plays an important role in inducing the regression of the Müllerian ducts in males, which normally give rise to the uterus, oviducts, upper vagina and fallopian tubes in females [22]. Analysis of mouse and human *Amh* gene regulation has uncovered several factors important for modulating *Amh* expression. For example, SF1 up-regulates *Amh* expression by cooperative interaction with WT1 [23], GATA-4 [24], SOX9 [25] and SOX8 [26]. Mice mutant for the SOX-binding site or the SF1-binding site in the *Amh* promoter revealed that SOX proteins are essential for *Amh* expression, while SF1 enhances the final expression level [27]. Oréal et al. [28] were the first to describe the chicken *Amh* promoter and showed that it has little overall homology with the mouse *Amh* promoter, but contains two putative SOX-binding sites and one SF1-binding site, suggesting that the mouse and chicken *Amh* promoters are similarly regulated. However, chicken *Sox9* is expressed too late to be a *cAmh* regulator [28, 29], but another SOX protein might substitute for SOX9 and, together with SF1, regulate *cAmh* expression.

Previously, we hypothesized that SOX8 might be a candidate transcription factor for regulating the chicken *Amh* gene [26, 30]. Mouse *Sox8* is expressed male specifically during gonad development. Its expression starts prior to the onset of *Amh* gene expression, and encodes a protein product that can up-regulate mouse *Amh* together with SF1 in vitro [26]. Group E Sox genes, consisting of Sox8, Sox9 and Sox10, show moderate levels of amino acid similarity and have similar genomic organization, suggesting that group E *Sox* genes may originate from one ancestral gene [31]. Although expression patterns of *Sox9* and *Sox10* overlap to a limited extent [32, 33], expression of *Sox8* overlaps substantially with expression of *Sox9* [31, 32] and to a lesser extent, *Sox10* [33, 31]. This fact suggests that there is some functional redundancy between SOX8 and SOX9, similar to that published for SOX1, SOX2 and SOX3 in lens formation [34], L-SOX5 and SOX6 in cartilage formation [35] and SOX7, SOX17 and SOX18 in vasculogenesis [36–38].

In this study, we analyzed the expression patterns of chicken *Sox8* in developing gonads during the sex-determining window. If *cSox8* contributes to *cAmh* gene expression, one would expect to find *cSox8* predominantly expressed in the embryonic testis and prior to the onset of *cAmh*. We found this not to be the case, suggesting that SOX8 is not responsible for sex-specific expression of *cAmh* in chicken. We also tested the expression patterns of several other *cSox* genes which are expressed in embryonic testis, and similarly found that they were not expressed male specifically.

Materials and methods

Chicken embryos. Fertilized chicken eggs were obtained from local suppliers (Ingham, Brisbane, Australia and Saitama Experimental Animal Supply, Saitama, Japan) and maintained at 18 °C until use. Eggs were transferred to an incubator at 37.5 °C and allowed to develop for 5, 6 or 7 days. Staging was confirmed at dissection according to Hamburger and Hamilton [39]. The entire urogenital ridge of each embryo was explanted for whole-mount *in situ* hybridization. Sexing was performed by PCR as described elsewhere [40] using genomic DNA purified from a hind limb of each embryo.

Amplification of HMG box sequences. Total RNA was isolated from left and right gonads of day 6 male embryos by standard methods [41]. RNA (0.5 μg) was then treated with DNaseI and first-strand cDNA was synthesized using M-MLV reverse transcriptase (Invitrogen) according to the manufacturer's protocol. For amplification of the HMG box, the PCR reaction was carried out in a solution containing 1 x NH_4 buffer (Bioline), 100 M $MgCl_2$, 100 μM dNTPs, 0.4 μM primers and 0.5 unit Biotaq DNA polymerase (Bioline) with 4.5 min denaturation at 95 °C followed by 40 cycles of amplification at 95 °C for 30 s and 57 °C for 30 s. The PCR product was cloned into pGEM-T Easy (Promega). Sequencing was performed using the BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems) and M13 reverse primer, and electrophoresis was carried out by the Australian Genome Research Facility, Brisbane, Australia. Primer sequences used were as follows; G7A1: 5'-AGC G(A/G)C CCA TGA ACG C(A/C/G/T)T T-3' and G7B1: 5'-CGC (C/T)GG TA(C/T) TT(A/G) TA(A/G) TC(A/C) GGG T-3'. The PCR reaction was also carried out using genomic DNA as a template.

RT-PCR of *Sox* genes, *Amh* and *Actin*. The left and right gonads of staged, sexed embryos were pooled (ten and five embryos of each sex for day 6 and day 7, respectively) to isolate total RNA using the RNeasy Mini kit (Qiagen) with the optional on-column DNase digestion with the RNase-free DNase set. The first-strand cDNA was synthesized from 1 μg of total RNA using Power-

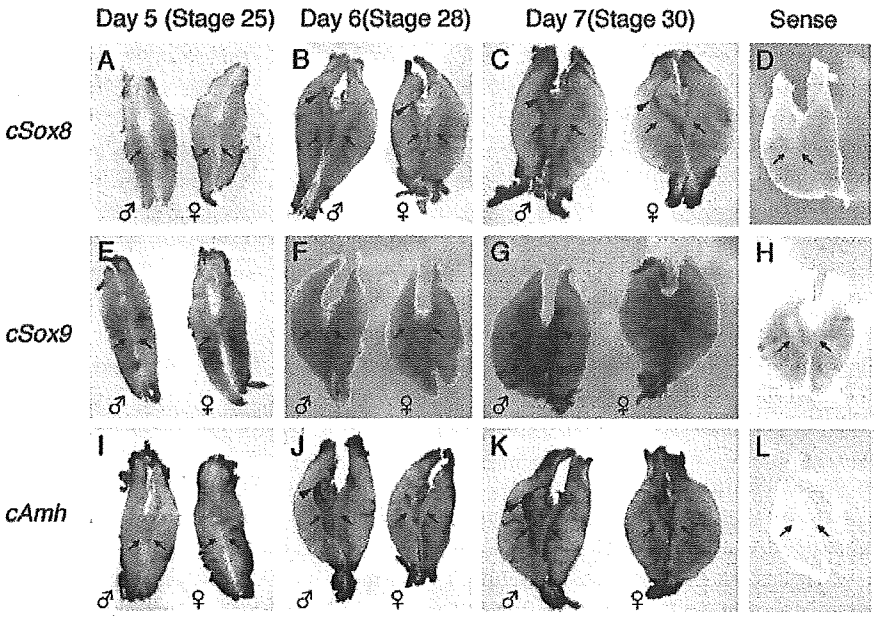

Figure 1. Whole-mount *in situ* hybridization of *cSox8*, *cSox9* and *cAmh* in the chicken embryonic gonad/mesonephros. In each panel, male (♂) and female (♀) gonad/mesonephros are left and right, respectively. Probes used are shown at the left of each panel. The samples in A–C, E–G and I–K were hybridized with antisense probe. In D and L a day 6 sample hybridized with sense probe. In H a day 7 sample was hybridized with sense probe. Developmental stages (days and Hamburger-Hamilton stage) of gonad/mesonephros are labeled above. Arrow shows the position of gonad. Arrowheads point to regions of positive staining.

Script reverse transcriptase (Clontech) and oligodT. PCR amplification was carried out using the QuantiTect SYBR Green PCR kit (Qiagen) with uracil-N-glycosylase and the 7900HT Sequence Detection System (Applied Biosystems). Samples were incubated at 50 °C for 2 min, then 95 °C for 15 min, followed by 40 cycles of amplification at 94 °C for 30 s, 54.2 °C for 1 min and 72 °C for 1 min. For the amplification of *cSox8* and *cSox9*, 85.4 °C for 15sec was added after each 72 °C step of each amplification cycle. Primer sequences used were as follows; cSox3-1: 5'-GCACCAGCACTACCAGAG-3' and cSox3-2: 5'-CGA ATG CGG ACA CGA ACC) for *cSox3* [29], cSox4F: 5'-TCG GGG GAT TGG CTG GAG TC-3' and cSox4R: 5'- CTC AGC CGA TCC TCG TTT CC-3' for *cSox4*, cSox8RTF : 5'-CTA CAA GGC TGA CAG CGG GC-3' and cSox8RTR: 5'-AGG CCG GGC TCT TGT GAG TC-3' for *cSox8*, cSox9F: 5'-CCC CAA CGC CAT CTT CAA-3' and cSox9R: 5'-CTG CTG ATG CCG TAG GTA-3' for *cSox9*, cSox11F2: 5'-AAG CAG GAG GCG GAC GAC GA-3' and cSox11R2: 5'-CGC CCC GCA CCT CCT CGT AG-3' for *cSox11*, cAmh-1: 5'-GTG GAT GTG GCT CCC TAC CC-3' and cAmh-2: 5'-GCA GCA CCG AGG GCT CCT CC-3' for *cAmh* [29] and Actin-1: 5'-TGG ATG ATG ATA TTG CTG C-3' and Actin-2: 5'-ATC TTC TCC ATA TCA TCC C-3' [29] for *Actin*. To calculate the relative amount of gene expression levels, the $\Delta\Delta$ C_T method was used. Three independent pooled samples were analyzed. Maximum average values were set as 100%.

For the RT-PCR amplification of *cSox12* and *cSox14*, the same cDNA for the real-time RT-PCR was used. The PCR reaction was carried out in the same solution as HMG box amplification with 4.5 min denaturation at 95 °C followed by 40 cycles of amplification at 95 °C for 30 s and 50 °C for 30 s. Primer sequences used were as follows; cSox12F: 5'-AGA TCT CCA AGC GCC TGG GTC G-3' and cSox12R: 5'-GGT AGT CGG CCA TGT GCT TG-3' for *cSox12*, cSox14F: 5'-GAG GTT CCT CAC ACC TTG GC-3' and cSox14R: 5'-ACA CGG AGG AAT CCC AGT CC-3' for *cSox14*.

Probes. The previously reported *cAmh* probe [9] was obtained by RT-PCR amplification of an 821-bp fragment, using primers cAMHRTF (5'-ACG GTG CGC GCC CAC TGG CAG G-3') and cAMHRTR (5'-ACG TCG TGA CCT GCA AGC CCT C-3') and RNA prepared from 5.5-day-old whole embryo. The *cSox8* probe was cloned by PCR using primers chSox8C2F (5'-CTG CAG AGC TCC AAC TAC TAC A-3') and chSox8C2R (5'-GAG CTC TGT CCT TTT GGA GAG T-3') and chicken genomic DNA as the template. The fragment corresponds to nucleotides 1228–1589 of the *cSox8* mRNA sequence (accession No. AF228664). PCR products of *cAmh* and *cSox8* were cloned into the pGEM-T Easy vector. The *cSox9* probe was reported previously by Kent et al. [11]. The *cSox11* fragment was cloned by SacI digestion of *cSox11* cDNA and subsequent insertion into pBluescriptII KS vector. The fragment corresponds to nucleotides 667–967 of the *cSox11* mRNA sequence (accession No. AB012237). The *cSox4* probe was obtained by KspI digestion of *cSox4* cDNA and subsequent insertion into pBluescriptII KS vector. The fragment corresponds to nucleotides 576–965 of the chicken *Sox4* mRNA sequence (accession No. AY493693).

***In situ* hybridization.** Sense and antisense digoxygenin-labeled RNA probes were generated by *in vitro* transcrip-

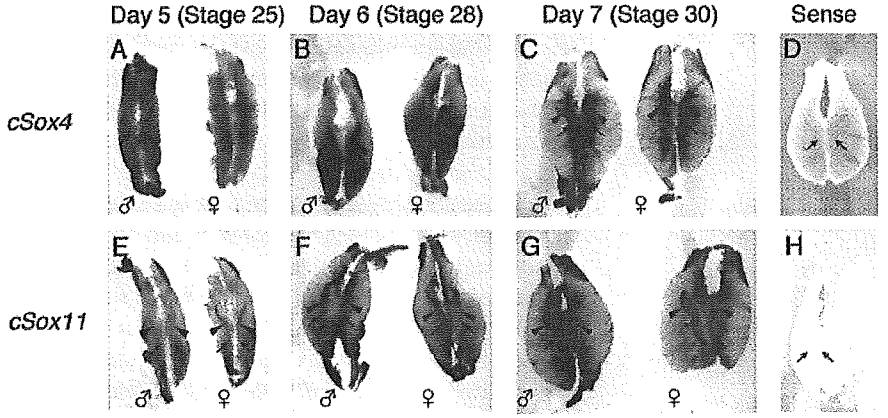

Figure 2. Whole-mount *in situ* hybridization of *cSox4* and *cSox11* in the chicken embryonic gonad/mesonephros. In each panel, male (♂) and female (♀) gonad/mesonephros are left and right, respectively. Probes used are shown at the left of each panel. A–C and E–G were hybridized with antisense probe. D and H show day 6 samples hybridized with sense probe. Developmental stages (days and Hamburger-Hamilton stage) of gonad/mesonephros are labeled above. Arrow shows the position of gonad. Arrowheads point to regions of staining.

tion. Whole mount *in situ* hybridization was performed as described using the maleic acid buffer (MABT) method [42]. Experiments were carried out at least twice for each probe, with similar results.

Results and discussion

To compare the temporospatial expression of *cSox8* and *cAmh* in embryonic gonads, we employed whole-mount *in situ* hybridization analysis using female and male gonad/mesonephros complexes isolated from day 5, 6 and 7 chicken embryos (Hamburger and Hamilton stages 25, 28 and 30, respectively [39]). These stages cover the temporal window at which sexual dimorphism in the gonad first becomes apparent [43]. In addition to providing spatial information relating to gene expression, whole-mount *in situ* hybridization is commonly used as a semiquantitative guide to gene expression levels between different tissues hybridized with the same probe and incubated under the same conditions, as in the experiments described below.

As expected, *cAmh* was expressed at higher levels in male than in female gonads at day 6 and 7 (fig. 1J, K). Expression levels in the right male gonads were higher than in the left, possibly reflecting the asymmetric development of avian genital ridges [43]. We did not observe expression of *cAmh* at day 5 (fig. 1I), even though Oréal et al. [28] reported that *cAmh* is expressed weakly and at similar levels in male and female gonads at day 5 by section *in situ* hybridization. This may reflect lower sensitivity of our whole mount *in situ* hybridization method.

Previous workers have reported that male-specific high-level expression of *cSox9* is preceded by expression of *cAmh* in the chick [28, 29]. We analyzed the temporal expression of *cSox9* in chicken embryos. No signal was observed in male or female gonads at day 5 or day 6 (fig. 1E, F). High levels of *cSox9* expression were observed in day 7 male gonads, while no signal was observed in the day 7 female (fig. 1G). This compares to the results of Oréal et al. [28] who described very faint expression of *cSox9* in day 6 gonads by *in situ* hybridization using sections. Our data demonstrate high levels of *cAmh* expression in day 6 male gonad, preceding the high levels of *cSox9* expression first detected in day 7 male gonad. They suggest that SOX9 is not responsible for the male-specific up-regulation of *cAmh*, but may play a role in the maintenance and/or amplification of *cAmh* expression in the male gonads once transcription is initiated. Our results support the previous observation that the male-specific high levels of *cAmh* expression precede testicular *cSox9* expression [28, 29].

We next examined expression of *Sox8*. At day 5, no *cSox8* expression was observed (fig. 1A). At day 6 and 7, *cSox8* expression was observed at the anterior tip of both male and female right gonads at similar levels (fig. 1B, C). This expression profile is different from that of *cAmh*, both in spatial distribution of transcripts and degree of sex specificity, suggesting that SOX8 could not be responsible for sex-specific up-regulation of *cAmh* in chicken.

The expression patterns of mouse and chicken *Sox8* imply that the functions of SOX8 are conserved in most but not all tissues between the two species. For example, *Sox8* is expressed in brain, skeletal muscle, eye, somite, dermomyotome, limb, digits, gut, spinal cord and dorsal root ganglia in both species [30, 31, 44]. However, some differences are evident in embryonic heart and gonad: in chicken, *cSox8* is expressed in the embryonic heart, testis and ovary, whereas in mouse, *Sox8* expression occurs predominantly in the embryonic testis, but not in the ovary or heart [31, 44]. Given these observations, SOX8 may contribute to the male-specific activation of *Amh* expression during gonadgenesis in mice but not chicken.

In mouse, only two SOX proteins, SOX8 and SOX9, have been identified as regulators of *Amh* expression in the embryonic gonad [26]. Based on our data, and previous studies [28, 29], neither SOX8 nor SOX9 can be responsible for the onset of high levels of *cAmh* expression

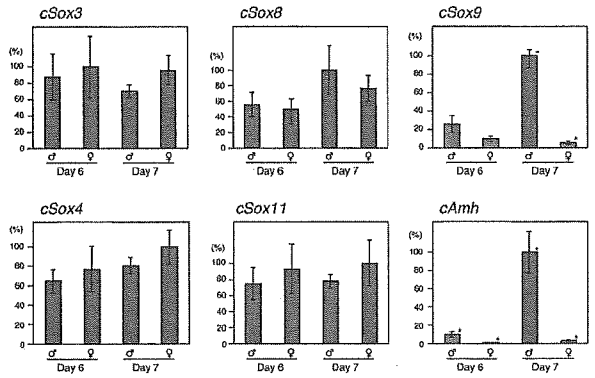

Figure 3. Quantitative, real-time RT-PCR analysis of *Sox* and *Amh* gene expression. Averages of three independent trials are shown as bars, with SEs shown as lines. Values marked with * were significantly different between males (♂) and females (♀) ($p = 0.0017$, 0.0047 and 0.012 for cSox9 at day 7, cAmh at day 6 and 7, respectively, using a two-sample equal variance t-test). Others were not significantly different between males and females ($p = 0.81$, 0.24, 0.74, 0.52, 0.17, 0.67, 0.38, 0.64, 0.49 for *cSox3* at day 6 and 7, *cSox8* at day 6 and 7, *cSox9* at day 6, *cSox4* at day 6 and 7, and *cSox11* at day 6 and 7, respectively).

in chicken, even though two SOX-binding sites are predicted in the *cAmh* promoter region. These observations prompted us to search for other *Sox* genes expressed in chicken male gonads that could be considered as candidate regulators of *Amh* expression. We conducted the analysis at day 6, since at this stage *cAmh* is expressed at high levels in male gonads while *cSox9* is either not expressed or expressed at a very low level.

We utilized degenerate RT-PCR on purified day 6 male gonad RNA using generic *Sry*-type HMG box primers to generate fragments for cloning into a plasmid vector. Twenty independent clones were sequenced, revealing that 12 clones were *cSox4* [45], 7 were *cSox11* [46] and 1 was *cSox9* [47].

One possible explanation for these results is that the degenerate primers used show a bias for amplification of *cSox4* and *cSox11* templates. To examine this possibility, we used the same degenerate primers to amplify *Sox* fragments from genomic DNA, in which all intronless *Sox* genes (Sox1, -2, -3, -4, -11, -12, -14, -21) capable of amplification by the primers are represented in equal proportions. Among ten clones amplified, none was *cSox4* or *cSox11*. Thus, primer bias does not explain our data relating to *cSox4* and *cSox11* expression in developing chicken gonads.

cSox4 and *cSox11* are expressed in male gonads at day 6, prompting us to examine the expression profiles of each in male and female gonads through the sex determination window. If both or either is expressed preferentially in male gonads, they could be considered a candidate for regulation of the *cAmh* gene. To evaluate the expression patterns of *cSox4* and *cSox11* in embryonic gonads, we employed whole-mount *in situ* hybridization analysis at the same stages previously used to profile *cSox8*, *cSox9* and *cAmh* expression. *cSox4* and *cSox11* signals were detected at similar levels in male and female gonads at all stages examined (fig. 2) suggesting that neither of them plays a role in sex-specific regulation of *Amh*.

The identification of *Sox* genes that are expressed in chicken embryonic gonads at day 5, 6 and 7 was previously attempted by McBride et al. [6]. Using RT-PCR amplification of the conserved *Sry*-type HMG box domain from RNA samples prepared from testes with mesonephroi attached, they found expression of *cSox3*, *cSox4*, *cSox9*, *cSox11*, *cSox12* and *cSox14*. Our data confirm that *cSox4*, *cSox9* and *cSox11* are indeed expressed, as is cSox3 (see below); however; we amplified the HMG box from day 6 gonad only, and this difference along with the differences in PCR primers, may explain the discrepancies in the data for *cSox12* and *cSox14*. Moreover, day 6 male gonad expresses *cSox4* and *cSox11* transcripts so abundantly that RT-PCR cloning is difficult for *Sox* genes expressed at low levels.

To examine the levels of gene expression quantitatively, we utilized RT-PCR and real-time RT-PCR analyses using RNAs isolated from pooled, sexed embryonic gonads at days 6 and 7 (fig. 3). As expected, *cAmh* and *cSox9* were expressed at different levels between males and females at day 7. At day 6, the expression levels of *cAmh* were statistically different between males and females ($p < 0.005$) while the expression levels of *cSox9* were not ($p > 0.1$). However, *cSox3*, *cSox4*, *cSox8* and *cSox11* were expressed at similar levels between males and females at days 6 and 7, suggesting that none of these *Sox* genes is responsible for the male-specific up-regulation of *cAmh* expression.

We were unable to amplify *cSox12* and *cSox14* sequences by RT-PCR from chicken embryonic gonads. As a positive control, chicken genomic DNA was included as template. Signals were observed at expected size of 108 bp for *cSox12* and 331 bp for *cSox14* only from genome template, but not from gonad RNA samples, showing that neither gene is expressed in embryonic gonads at day 6 and day 7 (data not shown).

Previous studies have eliminated *cSox3* as a candidate for male-specific up-regulation of *Amh* expression because *cSox3* is expressed at similar levels in the male and female gonads at the sex-determining window [28, 29]. Our present data support this conclusion. We rule out *cSox8* because it is expressed in a different spatial pattern to *Amh*, and *cSox12* and *-14* because they are not expressed in gonads at sex-determining stages at all. We exclude *cSox4* and *cSox11* also, on the basis of equivalent expression levels between male and female. It is formally possible that *cSox4* and *cSox11* might be expressed in Sertoli cells in the male (the site of *Amh* expression) and in another cell lineage in females, in which *Amh* is