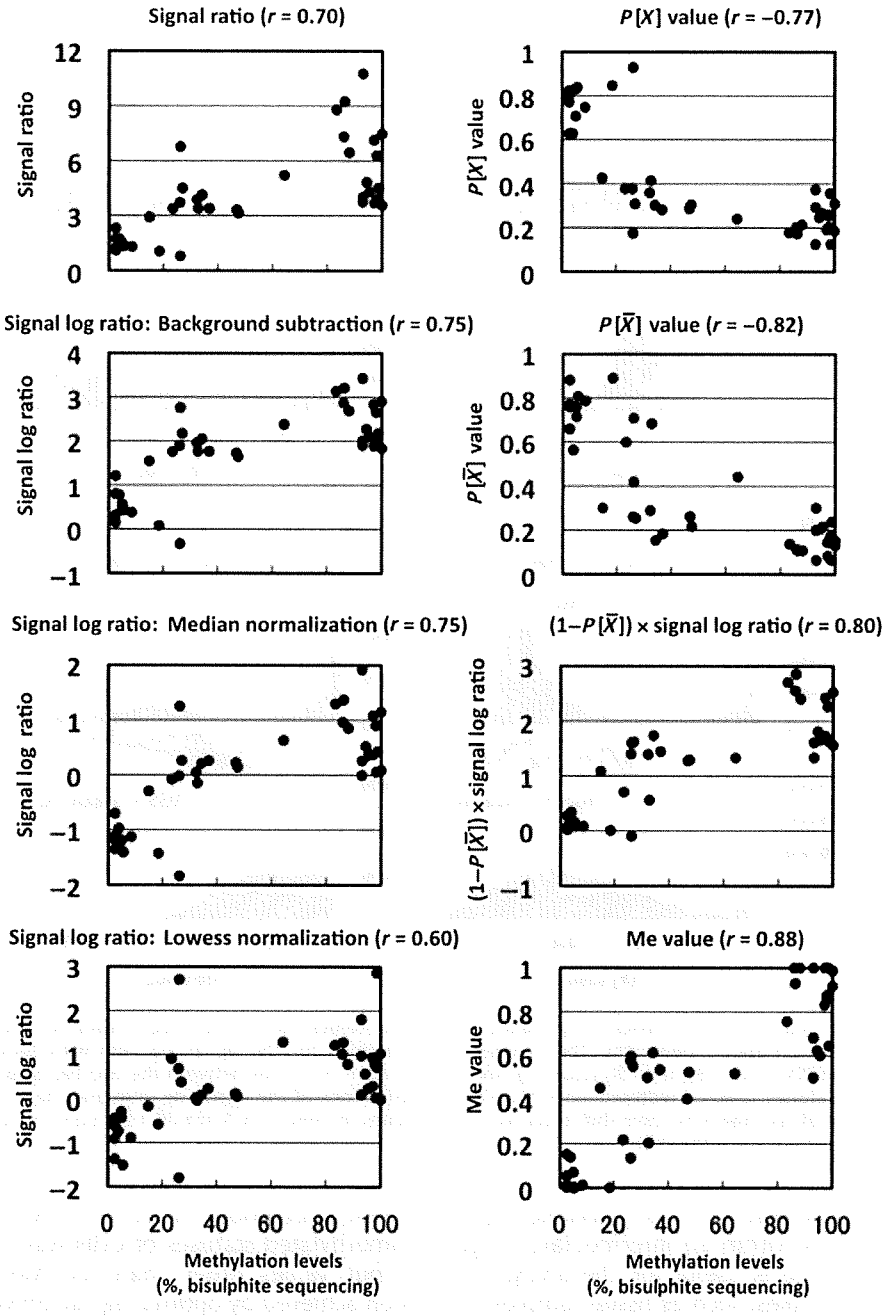

Figure 2. Correlation between an output value and the fraction of methylated CpG sites by bisulphite sequencing. Forty samples with variable methylation levels (11 loci in four cell lines) were analysed by bisulphite sequencing, and the fraction of methylated CpG sites among all the CpG sites within 200 bp of the microarray probes was obtained. The numbers of CpG sites in this region ranged from 7 to 21. The Me value had a linear correlation with the fraction ($r = 0.88$, Pearson's correlation coefficient).

for methylated and unmethylated CGIs with these cut-off values were 71.8 and 83.8%, respectively (Table 1). Since these cut-off CGI Me values worked in the four samples with different global methylation levels, Me values of 0.6 and 0.4 were considered to be suitable to identify highly methylated and unmethylated CGIs, respectively, with reasonable accuracies.

These data showed that, for quantification of methylation levels, the Me value had linearity similar to Batman and MEDME.[14,16] Although samples analysed were different, the coefficient of determination

Table 1. Comparison between methylation statuses determined using the CGI Me value and those by MSP

MSP	MeDIP-CGI microarray	AGS	HSC39	HSC57	KATOIII	Total
Methylated	Highly methylated	64	17	12	36	129
	Moderately methylated	29	13	6	23	71
	Unmethylated	4	3	6	8	21
Unmethylated	Highly methylated	1	6	7	4	18
	Moderately methylated	2	8	11	4	25
	Unmethylated	11	56	61	21	149
Methylated/unmethylated	Highly methylated	0	4	5	8	17
	Moderately methylated	1	4	2	5	12
	Unmethylated	1	2	2	4	9
Not amplified by MSP		0	0	1	0	1
Methylated	Sensitivity	0.660	0.515	0.500	0.537	0.584
	Specificity	0.938	0.875	0.864	0.739	0.848
	Accuracy	0.699	0.770	0.786	0.619	0.718
Unmethylated	Sensitivity	0.786	0.800	0.772	0.724	0.776
	Specificity	0.949	0.884	0.758	0.857	0.884
	Accuracy	0.929	0.832	0.768	0.823	0.838

Using 113 CGIs, we compared methylation statuses determined using the Me value and those by MSP. Methylated/unmethylated, both methylated and unmethylated DNA molecules were detected by MSP.

(R^2, Pearson's) obtained by the Me value was 0.77, being in the same range as Batman (0.82) and MEDME (0.75). The linearity of the Me value was validated also using four samples with different global methylation levels, which has not been done with Batman (one sample) and MEDME (two samples). Further, the CGI Me value enabled us to score completely methylated and unmethylated statuses of CGIs in samples with variable global methylation levels using common cut-off values. It is also of note that the Me value can be conveniently calculated using a spreadsheet and a commercially available software, such as Excel (Microsoft).

3.4. *Application of the Me value to analysis of the association between CGI methylation and gene expression*

Using the Me value, methylation profiles of CGIs and promoters were analysed in four gastric cancer cell lines, AGS cells after 5-aza-dC treatment, and one normal gastric tissue sample (a pool of four samples from four individuals). CGIs (promoters) with CGI (promoter) Me values more than 0.6 were classified as 'methylated'. Of the 30 533 CGIs analysed, 3768–7310 CGIs were methylated in the cancer cell lines, and 3393 CGIs were methylated in the normal sample (Table 2). However, methylation was infrequent in promoters (Table 2, Supplementary Fig. S4A and B). The ratio of methylated promoters in cancer cell lines (6.7–12.5%) was in accordance with that in a previous report (10.9% in lung cancer cell lines).[32] The difference between overall CGIs and promoters was clearer in the normal gastric tissue than in the human gastric cancer cell lines. Interestingly, CGIs in the vicinity of the LINE and SINE repetitive elements had lower Me values than those further away from them in cancer cell lines (Supplementary Fig. S4C and D). A small peak was observed at 300 bp from SINE, and CGIs closer to SINE and LINE were less methylated than those further apart from SINE and LINE. There is a possibility that a boundary function of these repetitive elements prevents methylation beyond the boundary spread across into CGIs. Actually, SINE is reported to have binding sites for the TFIIIC transcription factor, and to protect promoter CGIs from repressive chromatin modifications.[33]

The numbers of methylated CGIs were larger in AGS and KATOIII than in HSC39 and HSC57, which was in accordance with our previous findings.[26] The difference among the four cell lines was observed not only at the overall genome level but also on individual chromosomes (Supplementary Fig. S5). In contrast, the numbers of methylated promoters had the same difference at the overall genome level, but some distortions were present on specific chromosomes. For example, HSC39 had the largest numbers of methylated promoters on chromosomes 15, 18, and 22, and HSC57 had the largest number on chromosome 21. This result suggested that promoters have unique mechanisms to be methylated in cancer cells, such as functional selection of a cell with methylation of specific promoters.

Gene expression profiles were also obtained by expression microarray analysis of the same four cancer cell lines and the normal gastric tissue sample. When genes were classified by their expression levels, genes with high expression had lower methylation levels in their promoters than

Table 2. DNA methylation profiles of CGIs according to their positions against TSSs obtained by MeDIP-CGI microarray

	Total number analysed	Number of methylated CpG islands and promoters					
		AGS	HSC39	HSC57	KATOIII	AGS + 5-aza-dC (1 μM)	Normal gastric tissue
CpG islands							
All CGI	30 533	7310 (23.9%)	3768 (12.3%)	4663 (15.3%)	6460 (21.2%)	4 (0.0%)	3393 (11.1%)
Upstream	10 709	1911 (17.8%)	937 (8.7%)	794 (7.4%)	1556 (14.5%)	1 (0.0%)	260 (2.4%)
Gene body (+1 to +1 k)	10 654	1607 (15.1%)	867 (8.1%)	665 (6.2%)	1328 (12.5%)	0 (0.0%)	194 (1.8%)
Gene body (+1 to +5 k)	2050	694 (47.0%)	508 (24.8%)	598 (29.2%)	809 (39.5%)	1 (0.0%)	353 (17.2%)
Gene body (more than +5 k)	4431	2438 (55.0%)	1099 (24.8%)	2171 (49.0%)	2155 (48.6%)	2 (0.0%)	2360 (53.3%)
Downstream	1186	561 (47.3%)	310 (26.1%)	376 (31.7%)	509 (42.9%)	0 (0.0%)	216 (18.2%)
Divergent	1503	99 (6.6%)	47 (3.1%)	59 (3.9%)	103 (6.9%)	0 (0.0%)	10 (0.7%)
Promoters	9624	1205 (12.5%)	792 (8.2%)	641 (6.7%)	1142 (11.9%)	3 (0.0%)	113 (1.2%)
Probes	237 202	70 027 (29.5%)	43 825 (18.5%)	48 292 (20.4%)	65 192 (27.5%)	723 (0.3%)	27 017 (11.4%)

Definitions of the CGI positions are described in Section 2.

those with low expression in all cell lines (Fig. 3A). When genes were classified by methylation levels (unmethylated, moderately methylated, and highly methylated), genes with high methylation in their promoters had lower expression than those with low methylation (left panel in Fig. 3B). In contrast, genes with high methylation in their gene bodies (5 kb or more downstream of TSSs) had slightly, but significantly, higher expression levels (right panel). These results clearly showed that methylation of gene body CGIs was associated with increased gene expression, as in previous reports.[34–37] Finally, genes with low expression in normal gastric tissue had higher methylation levels of promoters in cancer cell lines (Supplementary Fig. S6). This result was in line with the fact that genes with low expression are susceptible to DNA methylation.[20,38,39]

3.5. *Application of the Me value to analyse the effect of a demethylating agent*

Treatment of cells with 5-aza-dC induces demethylation of various CGIs, but the degree depending upon their positions against TSSs has not been clarified. The relationship between the degree of demethylation of promoters and that of expression induction has not been clarified, either. To address these two issues, we analysed methylation and gene expression levels in AGS before and after 5-aza-dC treatment.

Demethylation of the genes with methylated CGIs or promoters (Me value > 0.6) before 5-aza-dC treatment was analysed. The average degree of demethylation was not influenced by the positions of CGIs against their TSS (Fig. 4A), whereas the degree of demethylation of individual CGIs was highly variable (representative genes in Fig. 4B). The average degree was not influenced by the distance between a CGI and repetitive elements (LINE, SINE), either (data not shown). There was no correlation ($r = 0.12$, Fig. 4C) between the degree of demethylation of promoters (decrease of Me value in MeDIP-CGI microarray) and that of induction of gene expression (fold increase of signal log ratio in expression microarray analysis). The majority of genes with methylation of their promoters showed little or no increase of expression after 5-aza-dC treatment. The number of methylated promoters identified by MeDIP-microarray analysis was much larger than that of genes identified as silenced by expression microarray analysis after 5-aza-dC treatment (Table 3).

These results showed that expression cannot be induced for the majority of genes with methylation of their promoters even with a demethylating agent, possibly due to the lack of transcriptional factors or the presence of inactive histone modifications. Genes with low transcription tend to become methylated,[20,38,39] and such genes are unlikely to be expressed even if the methylation is removed. Caution is necessary when the relationship between methylation of the promoter and the expression of a gene is interpreted.

3.6. *Application of the Me value in future studies*

The differential role of DNA methylation in promoters and gene bodies and the lack of association between the degree of demethylation of promoters and that of induction of gene expression were clearly shown due to the accuracy of the Me value. We here focused on CGIs using a CGI microarray. Roles of methylation of CpG-poor genomic regions

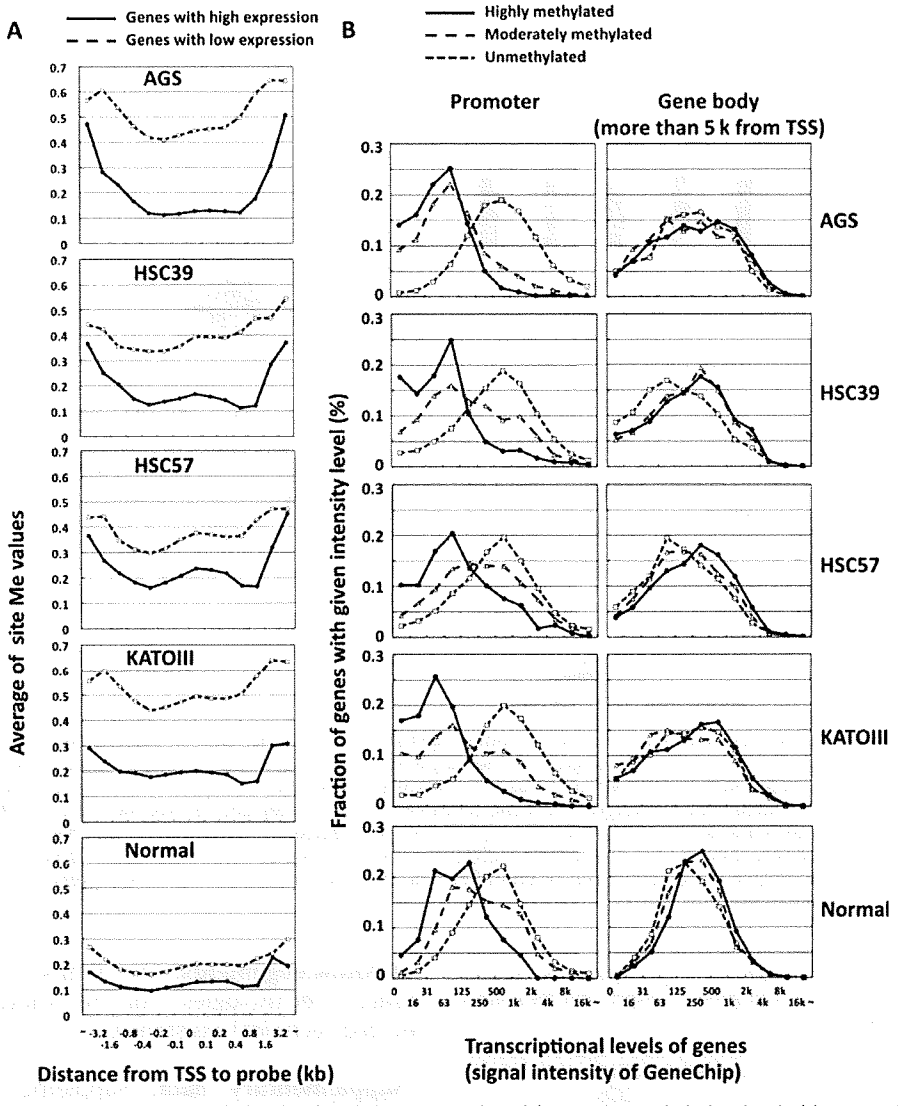

Figure 3. Association between DNA methylation level and gene expression. (A) Average methylation levels (site Me values) of genes with high and low expression. Genes with low expression (dashed line) showed significantly higher methylation levels than those with high expression (solid line) in positions close to TSSs. (B) Gene expression levels according to methylation statuses of the promoters (left panel) and gene bodies (right panel). Genes were classified into unmethylated (CGI Me value <0.4, dotted line), moderately methylated ($0.4 \leq$ CGI Me value ≤ 0.6, dashed line), and highly methylated (CGI Me value >0.6, solid line). Methylation of promoters was strongly associated with low expression ($P = 8 \times 10^{-203}$ in AGS, *t*-test of expression of methylated genes and unmethylated genes), and methylation of CGIs in gene bodies was associated with higher expression ($P = 6 \times 10^{-4}$ in AGS).

in genome regulation are not as clear as those of CGIs,[10,40] and focusing on CpG-rich regions is advantageous to isolate genes of interest. Also, the collection efficiency of methylated DNA by MeDIP or MBD is low in CpG-poor genomic regions,[7,9,41] and analysis of both CpG-poor and -rich regions leads to a complicated algorithm for normalization. The selective analysis of CGIs enabled us to develop a simple, but reliable output value.

The use of MeDIP allowed us to analyse methylation levels of any genomic regions regardless of the presence of restriction sites, and we were able to focus on promoters in some part of this study. We did not amplify the IP DNA during preparation of hybridization probes to avoid any amplification bias, and this was enabled by the use of a microarray platform with a high sensitivity and low signal/noise ratio. Under these optimal conditions, we achieved a high reproducibility ($r = 0.98$) and simplicity for the quantitative methylation microarray analysis.

In conclusion, we developed a new output value for microarray analysis suitable for quantitative

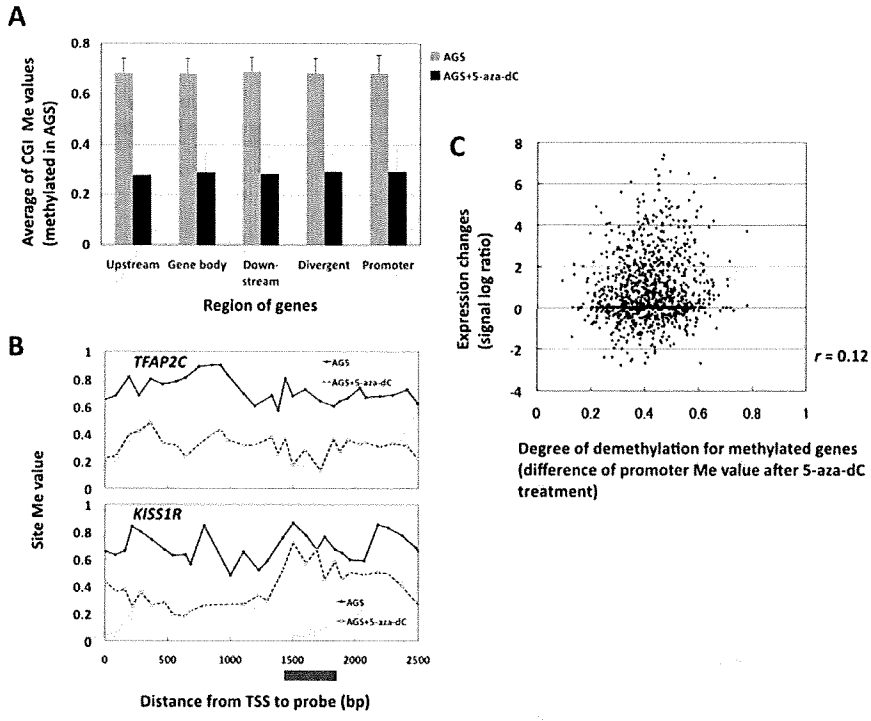

Figure 4. Degree of demethylation according to the positions against TSSs, and the lack of association between the degree of demethylation and that of re-expression. (A) Degree of demethylation according to the positions against TSSs. Methylation levels were analysed for CGIs methylated in AGS without 5-aza-dC treatment. 5-Aza-dC treatment induced demethylation of upstream CGIs, gene body CGIs, downstream CGIs, divergent CGIs and promoters to similar degrees. (B) Site Me value of *TFAP2C* and *KISS1R* in AGS and AGS treated with 5-aza-dC. In 1400–1800 bp against TSS, the degree of demethylation was high in *TFAP2C* and low in *KISS1R*, which representatively showed variation among individual genes. (C) Degree of demethylation of promoters and degree of re-expression in AGS. The degree of demethylation of promoters was not associated with that of re-expression of methylated genes after 5-aza-dC treatment.

Table 3. The number of methylated promoters by MeDIP-microarray analysis and silenced genes by expression microarray after 5-aza-dC treatment

	Number of methylated promoter CGIs			
	AGS	HSC39	HSC57	KATOIII
MeDIP-CGI microarray[a]	1205	792	641	1142
Expression microarray after 5-aza-dC treatment[b] (Moriguchi et al.[26])	25	4	1	41

[a]Promoters identified as methylated by MeDIP-microarray analysis, among 9624 genes arrayed on CpG island microarray (Agilent).

[b]The genes in which a 16-fold increase of expression after treatment of 5-aza-dC (optimized concentrations in each cell line) was observed by expression microarray, which confirmed its methylation in promoter by MSP, among 19 421 genes arrayed on GeneChip Human Genome 133 Plus 2.0 (Affymetrix).

assessment of DNA methylation levels. The Me value will be useful in genome-wide screening using heterogeneous samples.

Acknowledgements: We thank Dr Toshiki Taya (Agilent Technologies) for his knowledgeable and devoted technical assistance.

Supplementary data: Supplementary data are available at www.dnaresearch.oxfordjournals.org.

Funding

This work was supported by a Grant-in-Aid for the Third-term Cancer Control Strategy Program from the Ministry of Health, Labour and Welfare, Japan. K.G. and H.T. are recipients of a Research Resident Fellowship from the Foundation for Promotion of Cancer Research.

References

1. Jones, P.A. and Baylin, S.B. 2002, The fundamental role of epigenetic events in cancer, *Nat. Rev. Genet.*, **3**, 415–28.
2. Ushijima, T., Watanabe, N., Okochi, E., Kaneda, A., Sugimura, T. and Miyamoto, K. 2003, Fidelity of the

methylation pattern and its variation in the genome, *Genome Res.*, **13**, 868–74.
3. Jones, P.A. and Baylin, S.B. 2007, The epigenomics of cancer, *Cell*, **128**, 683–92.
4. Robertson, K.D. 2005, DNA methylation and human disease, *Nat. Rev. Genet.*, **6**, 597–610.
5. Suzuki, M.M. and Bird, A. 2008, DNA methylation landscapes: provocative insights from epigenomics, *Nat. Rev. Genet.*, **9**, 465–76.
6. Ushijima, T. 2005, Detection and interpretation of altered methylation patterns in cancer cells, *Nat. Rev. Cancer*, **5**, 223–31.
7. Weber, M., Davies, J.J., Wittig, D., et al. 2005, Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells, *Nat. Genet.*, **37**, 853–62.
8. Zhang, X., Yazaki, J., Sundaresan, A., et al. 2006, Genome-wide high-resolution mapping and functional analysis of DNA methylation in arabidopsis, *Cell*, **126**, 1189–201.
9. Keshet, I., Schlesinger, Y., Farkash, S., et al. 2006, Evidence for an instructive mechanism of de novo methylation in cancer cells, *Nat. Genet.*, **38**, 149–53.
10. Weber, M., Hellmann, I., Stadler, M.B., et al. 2007, Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome, *Nat. Genet.*, **39**, 457–66.
11. Zilberman, D., Gehring, M., Tran, R.K., Ballinger, T. and Henikoff, S. 2007, Genome-wide analysis of *Arabidopsis thaliana* DNA methylation uncovers an interdependence between methylation and transcription, *Nat. Genet.*, **39**, 61–9.
12. Hayashi, H., Nagae, G., Tsutsumi, S., et al. 2007, High-resolution mapping of DNA methylation in human genome using oligonucleotide tiling array, *Hum. Genet.*, **120**, 701–11.
13. Jacinto, F.V., Ballestar, E., Ropero, S. and Esteller, M. 2007, Discovery of epigenetically silenced genes by methylated DNA immunoprecipitation in colon cancer cells, *Cancer Res.*, **67**, 11481–6.
14. Down, T.A., Rakyan, V.K., Turner, D.J., et al. 2008, A Bayesian deconvolution strategy for immunoprecipitation-based DNA methylome analysis, *Nat. Biotechnol.*, **26**, 779–85.
15. Cheng, A.S., Culhane, A.C., Chan, M.W., et al. 2008, Epithelial progeny of estrogen-exposed breast progenitor cells display a cancer-like methylome, *Cancer Res.*, **68**, 1786–96.
16. Pelizzola, M., Koga, Y., Urban, A.E., et al. 2008, MEDME: an experimental and analytical methodology for the estimation of DNA methylation levels based on microarray derived MeDIP-enrichment, *Genome Res.*, **18**, 1652–9.
17. Straussman, R., Nejman, D., Roberts, D., et al. 2009, Developmental programming of CpG island methylation profiles in the human genome, *Nat. Struct. Mol. Biol.*, **16**, 564–71.
18. Buck, M.J. and Lieb, J.D. 2004, ChIP-chip: considerations for the design, analysis, and application of genome-wide chromatin immunoprecipitation experiments, *Genomics*, **83**, 349–60.
19. Yamashita, S., Tsujino, Y., Moriguchi, K., Tatematsu, M. and Ushijima, T. 2006, Chemical genomic screening for methylation-silenced genes in gastric cancer cell lines using 5-aza-2′-deoxycytidine treatment and oligonucleotide microarray, *Cancer Sci.*, **97**, 64–71.
20. Nakajima, T., Yamashita, S., Maekita, T., Niwa, T., Nakazawa, K. and Ushijima, T. 2009, The presence of a methylation fingerprint of Helicobacter pylori infection in human gastric mucosae, *Int. J. Cancer*, **124**, 905–10.
21. Lee, T.I., Rinaldi, N.J., Robert, F., et al. 2002, Transcriptional regulatory networks in *Saccharomyces cerevisiae*, *Science*, **298**, 799–804.
22. Pokholok, D.K., Harbison, C.T., Levine, S., et al. 2005, Genome-wide map of nucleosome acetylation and methylation in yeast, *Cell*, **122**, 517–27.
23. Hahn, M.A., Hahn, T., Lee, D.H., et al. 2008, Methylation of polycomb target genes in intestinal cancer is mediated by inflammation, *Cancer Res.*, **68**, 10280–9.
24. Dong, H., Yauk, C.L., Rowan-Carroll, A., et al. 2009, Identification of thyroid hormone receptor binding sites and target genes using ChIP-on-chip in developing mouse cerebellum, *PLoS One*, **4**, e4610.
25. Ke, X.S., Qu, Y., Rostad, K., et al. 2009, Genome-wide profiling of histone h3 lysine 4 and lysine 27 trimethylation reveals an epigenetic signature in prostate carcinogenesis, *PLoS One*, **4**, e4687.
26. Moriguchi, K., Yamashita, S., Tsujino, Y., Tatematsu, M. and Ushijima, T. 2007, Larger numbers of silenced genes in cancer cell lines with increased de novo methylation of scattered CpG sites, *Cancer Lett.*, **249**, 178–87.
27. Kaneda, A., Kaminishi, M., Yanagihara, K., Sugimura, T. and Ushijima, T. 2002, Identification of silencing of nine genes in human gastric cancers, *Cancer Res.*, **62**, 6645–50.
28. Miyamoto, K., Asada, K., Fukutomi, T., et al. 2003, Methylation-associated silencing of heparan sulfate D-glucosaminyl 3-*O*-sulfotransferase-2 (3-OST-2) in human breast, colon, lung and pancreatic cancers, *Oncogene.*, **22**, 274–80.
29. Abe, M., Watanabe, N., McDonell, N., et al. 2008, Identification of genes targeted by CpG island methylator phenotype in neuroblastomas, and their possible integrative involvement in poor prognosis, *Oncology*, **74**, 50–60.
30. Yamashita, S., Takahashi, S., McDonell, N., et al. 2008, Methylation silencing of transforming growth factor-beta receptor type II in rat prostate cancers, *Cancer Res.*, **68**, 2112–21.
31. Eckhardt, F., Lewin, J., Cortese, R., et al. 2006, DNA methylation profiling of human chromosomes 6, 20 and 22, *Nat. Genet.*, **38**, 1378–85.
32. Fukasawa, M., Kimura, M., Morita, S., et al. 2006, Microarray analysis of promoter methylation in lung cancers, *J. Hum. Genet.*, **51**, 368–74.
33. Tomilin, N.V. 2008, Regulation of mammalian gene expression by retroelements and non-coding tandem repeats, *Bioessays*, **30**, 338–48.

34. Hellman, A. and Chess, A. 2007, Gene body-specific methylation on the active X chromosome, *Science*, **315**, 1141–3.
35. Smith, J.F., Mahmood, S., Song, F., et al. 2007, Identification of DNA methylation in 3′ genomic regions that are associated with upregulation of gene expression in colorectal cancer, *Epigenetics*, **2**, 161–72.
36. Ball, M.P., Li, J.B., Gao, Y., et al. 2009, Targeted and genome-scale strategies reveal gene-body methylation signatures in human cells, *Nat. Biotechnol.*, **27**, 361–8.
37. Rauch, T.A., Wu, X., Zhong, X., Riggs, A.D. and Pfeifer, G.P. 2009, A human B cell methylome at 100-base pair resolution, *Proc. Natl Acad. Sci. USA*, **106**, 671–8.
38. Song, J.Z., Stirzaker, C., Harrison, J., Melki, J.R. and Clark, S.J. 2002, Hypermethylation trigger of the glutathione-S-transferase gene (GSTP1) in prostate cancer cells, *Oncogene*, **21**, 1048–61.
39. de Smet, C., Loriot, A. and Boon, T. 2004, Promoter-dependent mechanism leading to selective hypomethylation within the 5′ region of gene MAGE-A1 in tumor cells, *Mol. Cell Biol.*, **24**, 4781–90.
40. Irizarry, R.A., Ladd-Acosta, C., Wen, B., et al. 2009, The human colon cancer methylome shows similar hypo- and hypermethylation at conserved tissue-specific CpG island shores, *Nat. Genet.*, **41**, 178–86.
41. Rauch, T., Li, H., Wu, X. and Pfeifer, G.P. 2006, MIRA-assisted microarray analysis, a new technology for the determination of DNA methylation patterns, identifies frequent methylation of homeodomain-containing genes in lung cancer cells, *Cancer Res.*, **66**, 7939–47.

Chapter 10

Methylation-Sensitive Representational Difference Analysis (MS-RDA)

Toshikazu Ushijima and Satoshi Yamashita

Abstract

Methylation-sensitive representational difference analysis (MS-RDA) is a genome subtraction method that isolates DNA fragments differentially methylated between two genomes. It can be performed in any organism, even in those for which no microarray products are available. An important characteristic of MS-RDA is that it enriches unmethylated CpG-rich regions of the genome (amplicon), most of which are unique sequences. DNA fragments differentially methylated between two DNA samples will be present in one amplicon, but not in the other. The difference can be identified by RDA. Most technical difficulties reside in the RDA procedure, and many fine techniques are necessary for a successful application of this powerful technology.

Key words: MS-RDA, genome-wide screening, epigenome, DNA methylation, epigenetics.

1. Introduction

Methylation-sensitive representational difference analysis (MS-RDA) is a genome subtraction method that isolates DNA fragments differentially methylated between two genomes *(1–3)*. In an era when microarray analyses are extensively performed, the advantage of MS-RDA is that it can be performed to analyze any organism, even those for which no microarrays are available. Even if microarrays are available for the species of interest, MS-RDA does not require any fancy instruments, such as expensive scanners. If an investigator is familiar with the RDA procedure itself, MS-RDA can be easily performed.

For MS-RDA, genomic DNA is digested with a methylation-sensitive restriction enzyme, and DNA fragments are restricted

Jörg Tost (ed.), *DNA Methylation: Methods and Protocols, Second Edition, vol. 507*

DOI 10.1007/978-1-59745-522-0_10 Springerprotocols.com

into sizes suitable for PCR that are subsequently amplified into an "amplicon" by PCR using a universal adaptor and primer (**Fig. 10.1**). The amplicon is enriched with DNA fragments derived from unmethylated genomic regions with frequent occurrence of the recognition sites of the restriction enzyme. Preparation of an amplicon reduces the complexity of the genome, and genomic subtraction can be efficiently performed using two amplicons from two different genomes. DNA fragments present only in the "tester" sample, but not in the "driver" sample, can be isolated by the following genomic subtraction RDA (**Fig. 10.2**).

Any conventional method for genome-wide screening of differences in DNA methylation needs to prepare a representation

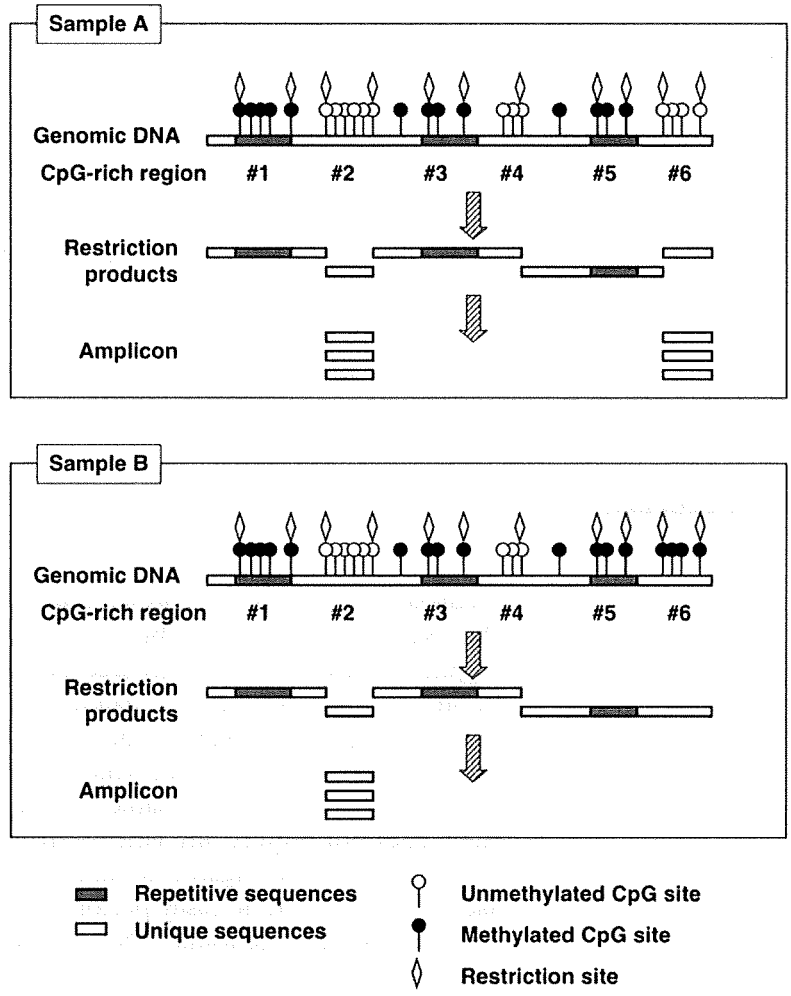

Fig. 10.1. Preparation of amplicon in MS-RDA. Tester and driver amplicons are prepared from samples A and B, respectively. The samples are digested by a methylation-sensitive restriction enzyme, such as *Hpa*II, *Sac*II, or *Nar*I. DNA fragments suitable for PCR are derived from unmethylated CpG-rich regions and amplified into an amplicon by PCR using a universal adaptor.

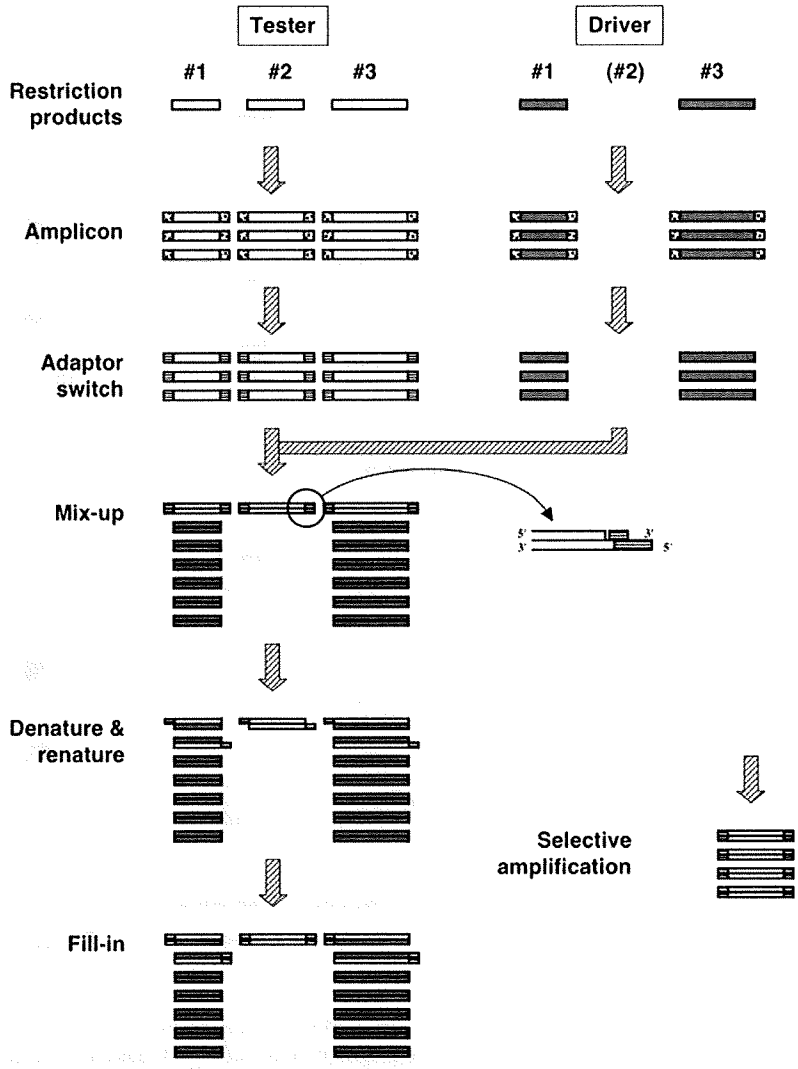

Fig. 10.2. Procedure of RDA. The overall procedure of RDA. Amplicons are prepared from restriction products of the tester and driver DNA (also shown in **Fig. 10.1**). Only for the tester, a new adaptor is ligated. A small amount of the tester is mixed with the driver amplicon, and the mixture is denatured by heat and undergoes re-annealing under stringent conditions. If a DNA fragment in the tester had its counterpart in the driver (fragments #1 and #3), it will re-anneal with a fragment in the driver because the driver is present in large excess. If a DNA fragment does not have its counterpart in the driver (fragment #2), it will re-anneal with a tester fragment. Only such DNA fragments have an adaptor on both ends and can be amplified exponentially in the following selective amplification.

of the genome as the entire genome spans 3×10^9 bp and is too large to be comprehensively analyzed by any conventional method and even by most microarray techniques. The representation that is used for the screening can be prepared by various means such as subtraction (MS-RDA, methylated CpG island amplification-RDA, etc.) or two-dimensional gel-electrophoresis (restriction landmark genomic scanning; RLGS,

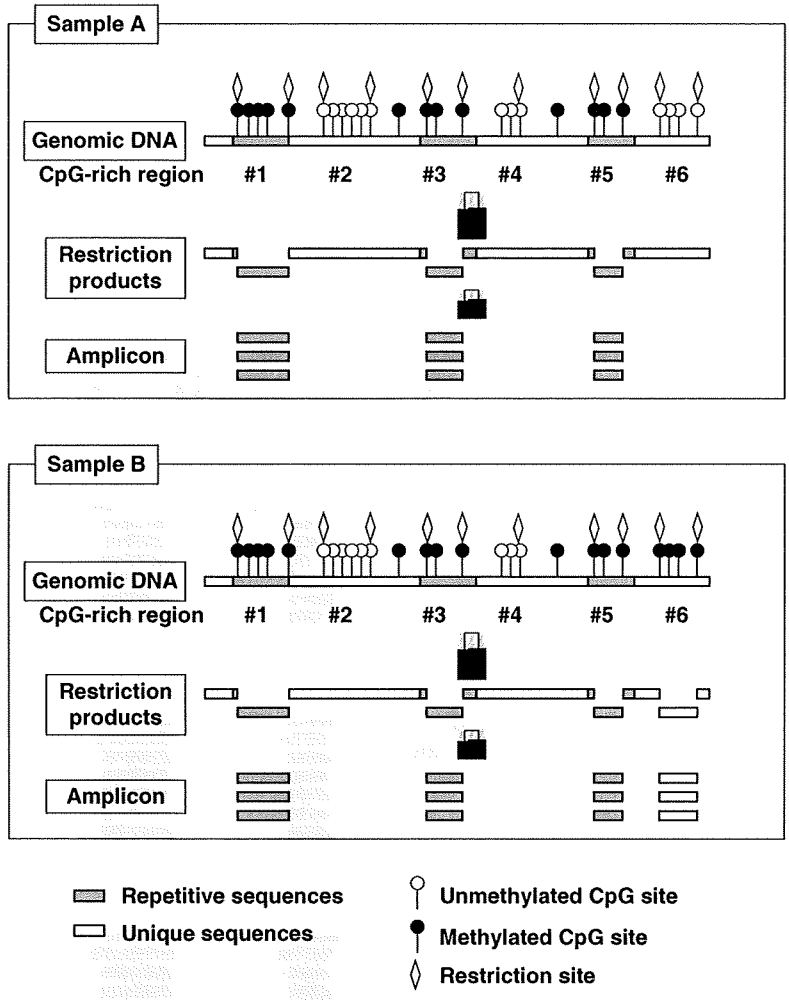

Fig. 10.3. Preparation of amplicons from methylated regions of the genome. Amplicons can also be prepared from methylated regions of the genome. However, if such regions are amplified, the vast majority is derived from repetitive sequences and not suitable for genome subtraction.

Chapter 11) *(3)*. To prepare a representation of the genome, either the unmethylated or methylated regions of the genome can be extracted. Since more than 40% of the human genome is made of repetitive sequences that are CpG rich and highly methylated, a representation constructed from methylated regions of the genome will contain a significant amount of repetitive sequences (**Fig. 10.3**). Repetitive sequences are resistant to removal by genome subtraction and prevent efficient isolation of differentially methylated unique sequences. MS-RDA prepares its representation from unmethylated regions of the genome, and avoids thereby the interference from repetitive sequences.

MS-RDA has been applied to isolate DNA fragments methylated in human lung cancers, gastric cancers, breast cancers, pancreatic cancers, neuroblastomas, and melanomas *(4–10)*. A

tumor-suppressor gene was identified in gastric cancers *(11)*, and a promising prognostic marker was identified in neuroblastomas *(9, 12)*.

2. Materials

2.1. Regular Reagents Used in Molecular Biology

1. 5 M NaCl,
2. Neutralized, water-saturated phenol and chloroform,
3. 3 M NaOAc,
4. 10 M NH_4OAc,
5. Agarose and NuSieve GTG agarose,
6. 0.5 × TBE: 5 mM Tris, 0.45 mM boric acid, and 1 mM EDTA, and
7. Competent cells.

2.2. Oligonucleotides

1. Oligonucleotides to prepare adaptors for the *Hpa*II series:
   RHpa24: 5′-AGC ACT CTC CAG CCT CTC ACC GAC-3′,
   RHpa11: 5′-CGG TCG GTG AG-3′,
   JHpa24: 5′-ACC GAC GTC GAC TAT CCA TGA AAC-3′,
   JHpa11: 5′-CGG TTT CAT GG-3′,
   NHpa24: 5′-AGG CAA CTG TGC TAT CCG AGG GAC-3′,
   NHpa11: 5′-CGG TCC CTC GG-3′.
2. Oligonucleotides to prepare adaptors for the *Sac*II series:
   RSac26: 5′-AGC ACT CTC CAG CCT CTC ACG ACC GC-3′,
   RSac9: 5′-GGT CGT GAG-3′,
   JSac26: 5′-ACC GAC GTC GAC TAT CCA TGA ACC GC-3′,
   JSac9: 5′-GGT TCA TGG-3′,
   NSac26: 5′-AGG CAA CTG TGC TAT CCG AGG ACC GC-3′, and
   NSac9: 5′-GGT CCT CGG-3′.
3. Oligonucleotides to prepare adaptors for the *Nar*I series:
   RNar24: 5′-AGC ACT CTC CAG GCA CTC ACC AGG-3′,
   RNar11: 5′-CGC CTG GTG AG-3′,
   JNar24: 5′-ACC GAC GTC GAC TAT CCA TGA AGG-3′,
   JNar11: 5′-CGC CTT CAT GG-3′,
   NNar24: 5′-AGG CAA CTG TGC TAT CCG AGG AGG-3′, and
   NNar11: 5′-CGC CTC CTC GG-3.
4. Oligonucleotides used as PCR primers.
   For the *Hpa*II series, use RHpa24, JHpa24, and NHpa24.
   For the *Sac*II series, use RSac24: 5′-AGC ACT CTC CAG CCT CTC ACG ACC-3′,
   JSac24: 5′-ACC GAC GTC GAC TAT CCA TGA ACC-3′ and

NSac24: 5′-AGG CAA CTG TGC TAT CCG AGG ACC-3′. For the *Nar*I series, use RNar24, JNar24, and NNar24.

2.3. Amplicon Preparation

1. Methylation-sensitive restriction enzymes *Hpa*II, *Sac*II, and *Nar*I (New England Biolabs, Beverly, MA), and the methylation-insensitive restriction enzyme, *Msp*I (New England Biolabs).
2. T4 ligase (New England Biolabs), and T4 ligase buffer (10 × accompanying T4 ligase). The T4 ligase buffer contains ATP and should be thawed on ice.
3. Taq polymerase. Hot Start Taq polymerases activated by heating at 95°C are "not" suitable.
4. dNTP each 2 mM or 2.5 mM.
5. 10 × PCR buffer III: 67 mM Tris-HCl (pH 8.8), 4 mM $MgCl_2$, 16 mM $(NH_4)_2SO_4$, 10 mM β-mercaptoethanol, and 100 μg/mL bovine serum albumin (BSA). After mixing these reagents, filter with an Acrodisk (0.2 μm), aliquot into several 1-mL stock tubes, and keep at −20°C.
6. 5 M betaine. After dissolution, filter with an Acrodisk (0.2 μm), aliquot into several 1 mL stock tubes, and keep at −20°C.
7. Chroma Spin + TE-200 column (Clontech K1325).

2.4. Competitive Hybridization and Selective Amplification

1. 3 × EE solution: 30 mM EPPS (4-(2-Hydroxyethyl)-1-piperazinepropanesulfonic acid, pH 8.0), and 3 mM EDTA. EPPS is known to have a consistent pH at high temperatures *(13)*, and is stable for years once the pH is adjusted.
2. Mung bean nuclease (New England Biolabs).

2.5. Cloning of the MS-RDA Product

1. TA cloning kit using for example pGEM-T Easy (Promega).

3. Methods

The difficulty of MS-RDA mainly resides in the technical complexity of RDA. Since there are many steps in RDA, it is evident that even if 80% efficiency is achieved in each step, the final yield will go down to 11% after 10 steps. If 90% efficiency is achieved, it will be 35%; and at 95% efficiency, it will be 60%. Therefore, it is essential to maximize the efficiency at each step by maintaining DNA fragments that should be maintained and by eliminating those that should be eliminated.

The following procedures describe the *Hpa*II series of MS-RDA. Except for the optimal ratio of the mixture of the tester and driver amplicons, the *Sac*II series and *Nar*I series can be

performed in the same manner using the appropriate adaptors (**Section 2.2**).

3.1. Amplicon Preparation

1. Genomic DNA (10 μg) is digested twice with an excess amount (100 U) of a methylation-sensitive restriction enzyme such as *Hpa*II, *Sac*II, or *Nar*I (*see* **Note 1**). The digestion product is extracted twice with phenol, twice with chloroform, and precipitated with ethanol using NH_4OAc. Dissolve the pellet in 20 μL of TE and quantify the solution.
2. Prepare adaptors for PCR amplification in parallel to step 1. To prepare 100 μM RHpa adaptor solution, dilute both RHpa24 and RHpa11 oligonucleotides with TE to final a concentration of 100 μM. Heat the mixture at 70°C for 5 min and cool down to 10°C. Prepare JHpa and NHpa adaptors in a similar manner (*see* **Note2**).
3. Ligate the RHpa adaptor to the purified *Hpa*II-restricted DNA. The ligation mixture (30 μL) contains 500 pmol of RHpa adaptor, 1 μg of the restricted DNA, 1 × T4 ligase buffer, and 800 U of T4 ligase. Keep the mixture at 16°C overnight (*see* **Note 3**).
4. Prepare a PCR mix (400 μL per sample) containing 3 μL of the ligation mixture, 1 × PCR buffer III, 300 μM dNTPs, 1 M betaine (*see* **Note 4**), and 1 μM RHpa24 primer. Prepare one 500 μL Eppendorf tube for the tester sample, and 10 tubes for the driver sample (*see* **Note 5**).
5. Start the PCR reaction by heating tubes at 72°C and addition of 3 μL of Taq polymerase (5 U/μL) to each tube (*see* **Note 6**). Then, perform 20 cycles of PCR (95°C for 1 min and 72°C for 3 min), and check the degree of amplification by running 10 μL of the PCR solution in a 0.9% agarose gel. If the PCR product (smear) has an intensity comparable to the DNA marker (250 ng/lane), the amplification is sufficient (**Fig. 10.4**). If not, add two to ten more cycles.
6. When sufficient amplification is achieved (*see* **Note 7**), purify the PCR product by two phenol extractions, two chloroform extractions, and ethanol precipitation with NH_4OAc. Dissolve the precipitate in 30 μL (tester) and 300–400 μL (driver) of TE, and quantify the samples (*see* **Note 8**).
7. Digest 20 μg of tester amplicon and 200 μg of driver amplicon with 100 U and 1000 U of *Msp*I, respectively. By running 1 μg of the digested DNA in a 3% NuSieveGTG agarose gel, confirm that the adaptor is completely restricted (**Fig. 10.5**).
8. Inactivate the restriction enzyme by extracting the digestion solution with phenol. Remove the digested adaptor by applying the extracted solution to a Chroma Spin + TE-200 column (one column for the tester, and four columns for the driver). For the tester, only the first elute (E1) should be used for the following steps. For the driver, the amplicon remaining in the

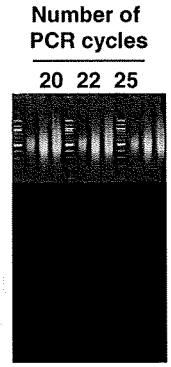

Fig. 10.4. Typical preparation of an amplicon. With 20 cycles of PCR the smear was not intense enough, and two cycles were added (22 cycles). If a further three cycles were added (25 cycles), the smear extended into a high molecular weight and was considered to have a significant amplification bias.

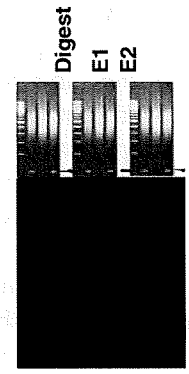

Fig. 10.5. Elimination of the digested adaptor. The digested amplicon showed a strong band for the 24-bp adaptor (shown by an arrow), and the band has completely disappeared in E1. However, the band still appeared in E2, and this E2 was again column purified and used as a driver.

column can be washed out (E2) by adding another TE for elution (*see* **Note 9**). By running 5 μL of E1 (and E2) in 3% NuSieveGTG agarose gel, confirm that the digested adaptor is at least for E1 completely removed (**Fig. 10.5**).

3.2. Competitive Hybridization

1. Ligate a new adaptor (JHpa adaptor) only to the tester amplicon E1. The ligation solution (30 μL) contains 500 pmol of JHpa adaptor, 200 ng of the tester amplicon (E1), 1 × ligation buffer, and 800 U of T4 ligase. Keep the solution at 16°C overnight.
2. Add 70 μL of TE to the ligation solution, and extract the solution once with phenol. Mix the extracted solution (100 μL) with 40 μg of the driver amplicon (E1 and E2). Extract the mixture once with phenol, twice with chloroform, and precipitate it with NaOAc (*see* **Note 10**).

3. Centrifuge the tube and rinse the pellet with 70% ethanol. Dry the pellet to a degree that it is dry but still retains water (*see* **Note 11**). Dissolve the pellet in 4 μL of 3 × EE solution (*see* **Note 12**).
4. Denaturation and re-annealing are successively performed. Place one drop of mineral oil onto the 4 μL 3 × EE solution in tube C. Denature the DNA mixture by heating at 96°C for 10 min in a thermal cycler. Add 1 μL of 5 M NaCl, paying attention so that the tube is "not" cooled during the handling (*see* **Note 13**), and keep the tube at 67°C for 16–24 h for re-annealing.

3.3. Selective Amplification

1. The denatured and re-annealed mixture is diluted by adding 45 μL of 1 M NaCl preheated at 67°C. To prevent evaporation, 1 M NaCl should be prepared in a large volume (400–500 μL). Great care should be taken that the mixture does not cool down.
2. Prepare a PCR solution (387 μL for now, but final 400 μL) that contains 1 × PCR III buffer, 300 μM dNTP, 1 M betaine, and 15 U of Taq polymerase; keep the solution at 72°C.
3. To the PCR solution, add 5 μL of the diluted mixture (step 1) at 72°C avoiding cool down of the mixture (*see* **Note 14**). This step is necessary to fill in 3′ ends of re-annealed products.
4. Add 8 μL of JHpa24 primer (50 μM) to the PCR solution at the first 95°C step, and perform 10 cycles of PCR (95°C for 1 min and 70°C for 3 min). After 10 cycles of PCR, keep the PCR solution at 72°C. Do not cool it down.
5. Take the tubes out of the thermal cycler and add immediately 40 μL of 10 × mung bean buffer to the PCR tube (*see* **Note 15**). Then add 10 μL of mung bean nuclease (10 U/μL). Keep the tube at 30°C for 30 min.
6. Extract the solution twice with phenol, twice with chloroform, and precipitate it with ethanol using NH_4OAc (*see* **Note 16**). Dissolve the pellet in 30 μL of TE.
7. Prepare a PCR solution that contains 1 × PCR III buffer, 300 μM dNTPs, 15 U Taq polymerase, 1 M betaine, and 3 μL of the above solution (step 6) (*see* **Note 17**). Heat the tube to 95°C, and then add 8 μL of JHpa24 primer (*see* **Note 18**). Perform 20 cycles of PCR (95°C for 1 min and 70°C for 3 min).
8. Check the degree of PCR amplification by running 10 μL of the PCR solution in a 2% NuSieve agarose gel (**Fig. 10.6**). If the amplification is not sufficient, add 2–10 cycles of PCR.
9. Extract the PCR solution twice with phenol, twice with chloroform, and precipitate it with ethanol using NH_4OAc. Dissolve the pellet in 50 μL of TE, and quantify the solution. This is the product of the first cycle of competitive hybridization (C1).

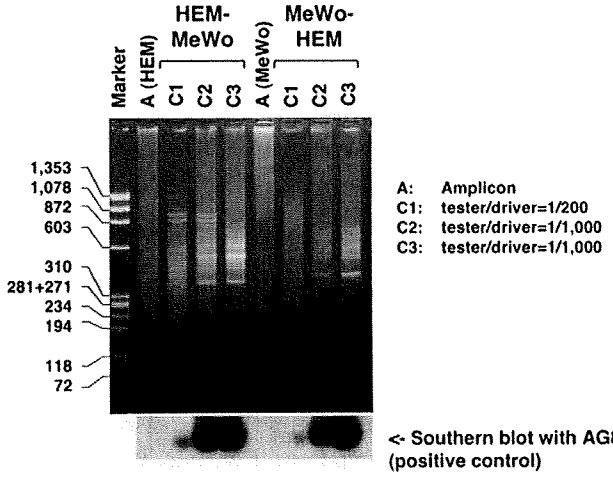

Fig. 10.6. A typical course of MS-RDA. Amplicons were prepared from a normal human embryonic melanocyte (HEM) and a melanoma cell line (MeWo). Two series of MS-RDA were performed using HEM as the tester (HEM – MeWo) and MeWo as the tester (MeWo – HEM). In C1, some differentially methylated fragments seem to be enriched. In C2 and C3, differentially methylated fragments became clearly visible by ethidium bromide staining of the entire DNA. Appropriate enrichment of differentially methylated fragments was confirmed by adding an unmethylated and methylated DNA fragment (AG8) to the tester and driver, respectively, at a concentration of one copy per haploid genome, and by observing its amplification by Southern blot analysis.

3.4. The Second Cycle

1. Digest 5–20 μg of C1 with 50–200 U of *Msp*I to switch the JHpa adaptor to the NHpa adaptor.
2. In a similar manner to the adaptor switch from RHpa to JHpa (step 8 of **Section 3.1** amplicon preparation), remove the digested JHpa adaptor, and confirm that the removal was complete.
3. To the C1 whose adaptor was removed (C1/MspI/E1), ligate the NHpa adaptor in a 30 μL solution that contains 500 pmol of the NHpa adaptor, 200 ng of the C1/MspI/E1, 1 × T4 ligase buffer, and 800 U of T4 ligase. Keep the solution at 16°C overnight.
4. Add 70 μL of TE to the ligation solution, and extract the solution once with phenol. Mix 40 ng (20 μL of the diluted solution) of the diluted ligation solution with 40 μg of driver amplicon (E1 and E2). Extract the mixture once with phenol, twice with chloroform, and precipitate it with ethanol using NaOAc.
5. Perform competitive hybridization and selective amplification as described for C1 in steps 3 and 4 of **Section 3.2** (competitive hybridization) and steps 1–9 of **Section 3.3** (selective amplification), except that the NHpa24 primer is used (*see* **Note 19**). This will produce C2. For C2, 20 cycles are usually enough for the second PCR as described in step 7

of **Section 3.3** (selective amplification). Confirm that selective amplification took place by running C2 in a 2% NuSieve agarose (**Fig. 10.6**).

3.5. The Third Cycle (OPTIONAL)

1. If selective amplification is not satisfactory in C2, a third cycle could be performed. In this case, switch the adaptor from NHpa back to JHpa, and perform competitive hybridization and selective amplification as described above (*see* **Note 20**).

3.6. Analysis of the MS-RDA Product

1. Clone the product of the last selective amplification (C2 or C3) into pGEM-T Easy vectors (*see* **Note 21**). After transformation into competent cells, pick up 96 clones and sequence them (*see* **Note 22**, *see also* **Chapter 14**). The cloned fragments have the final adaptor on their ends, and their sequence could be searched in genome databases.

4. Notes

1. This step is the only step that detects the DNA methylation status of the various genomic regions in the tester and driver. Therefore, the digestion by a methylation-sensitive restriction enzyme must be complete. Complete digestion can be confirmed by electrophoresing an aliquot of the digestion solution in a 0.9% agarose gel and observing the smear pattern. Also, some DNA samples are contaminated with RNA, and, if so, RNA must be eliminated during the digestion by RNase treatment. This will allow an appropriate amount of genomic DNA to be used for the following steps to prepare an amplicon.
2. Once an adaptor is prepared, it can be stored at −20°C for years.
3. The ligation time can be shortened down to 4 h. Since ATP is necessary for the ligation reaction, 10 × T4 ligase buffer should be handled on ice.
4. Betaine is known to facilitate amplification of G+C-rich sequences *(14)*.
5. These can be prepared at room temperature, while some of the later steps need careful temperature control. The ligation solution can be used without any purification.
6. To prevent nonspecific amplification, Taq polymerase should be added at a high temperature. A hot-start using an engineered-Taq polymerase, such as AmpliTaq Gold® (Applied Biosystems), is not suitable because the 3′ ends of the ligation products must be filled in at the initial incubation step at 72°C (**Fig. 10.2**).

7. For the following steps, a sufficient quantity of amplicon should be prepared. However, if too many cycles of PCR are performed, the amplification bias among DNA fragments will become significant. Therefore, the number of PCR cycles should be kept to a minimum within the range that yields a sufficient quantity of amplicon.
8. Typically, the driver amplicon yields 200–400 μg of amplified DNA.
9. Since 40 μg of driver amplicon without the adaptor are necessary for a competitive hybridization, 80–120 μg of the driver amplicon are necessary for a complete MS-RDA procedure. Use of E2 is sometimes inevitable to secure sufficient amounts of driver amplicon without the adaptor. However, for the tester, only E1 should be used because any residual adaptor can be ligated again and impairs the efficiency of competitive hybridization and selective amplification.
10. Since the efficiency of competitive hybridization and re-annealing is critical for the success of the experiment, the mixture is purified completely.
11. As with usual ethanol precipitation, the pellet should be dried to an appropriate degree. If it is too dry, its dissolution will become very difficult. If it contains much ethanol, the volume becomes larger and the efficiency of re-annealing will be impaired.
12. Ethanol precipitation is typically performed in a 1.5 mL Eppendorf tube, and the next step will be performed in a 0.5 mL Eppendorf tube. To dissolve the pellet of 40 μg of DNA (in tube A) completely:
    (a) take two 0.5 mL tubes (B, C),
    (b) put 5 μL of 3 × EE in tube B,
    (c) move 2 μL of 3 × EE from tube B to the pellet in tube A,
    (d) vortex tube A for more than 1 min paying attention that the solution is on the pellet, and spin it down (the solution is very sticky),
    (e) move the 2(+α) μL of the solution in tube A to tube C,
    (f) add fresh 2 μL of 3 × EE from tube B to tube A without changing the tip,
    (g) vortex tube A for more than 1 min, and spin it down,
    (h) move the 2 μL in tube A to tube C without changing the tip, and
    (i) vortex tube C, and spin it down.
13. The 5 M NaCl can be preheated, but care should be taken to avoid evaporation. Wear gloves to avoid burns while handling a hot tube. Cooling down of the tube accelerates non-specific annealing, and will impair the efficiency of RDA.

14. The remaining solution (45 μL) from step 1 is usually kept at 67°C until the success of step 8 is confirmed. If the pellet is lost at step 6, it can be started over from this point.
15. While adding the mung bean buffer to the PCR solution, its temperature goes down to 30–40°C. Do not cool it down too much.
16. The pellet is very tiny, and the highest care must be taken that it is not lost. This step is one of the most difficult steps in the entire RDA procedure.
17. At this step, for the first time, all the DNA molecules in the tube are completely double stranded because all the single-stranded DNA molecules were digested by mung bean nuclease. Therefore, taking care of the temperature is not necessary during preparation of the PCR solution.
18. Only for this PCR, hot-start PCR can be used, if addition of a primer at 95°C is troublesome. The entire solution can be prepared at room temperature using an engineered-Taq polymerase, such as AmpliTaq Gold®.
19. The product of the first 10 cycles of PCR of the second competitive hybridization is usually much more abundant than the one in the first competitive hybridization, and the risk of loss is lower.
20. An addition of a fourth cycle does not improve the experiment. If selective amplification is not obtained in the third cycle, the quality of the initial samples, selection of the initial samples, and technical errors should be considered.
21. Any TA cloning vector is fine.
22. The number of clones sequenced is dependent upon the diversity of the final product. If redundant clones are observed by sequencing, sequencing can be suspended at that time point. If only nonredundant clones are observed, more clones should be sequenced.

Acknowledgements

The protocol is improved through the support of a Grant-in-Aid for the Third-Term Comprehensive Cancer Control Strategy from the Ministry of Health, Labour and Welfare, Japan.

References

1. Ushijima, T., Morimura, K., Hosoya, Y., et al. (1997) Establishment of methylation-sensitive representational difference analysis and isolation of hypo- and hypermethylated genomic fragments in mouse liver tumors. *Proc Natl Acad Sci USA* 94, 2284–2289.
2. Kaneda, A., Takai, D., Kaminishi, M., et al. (2003) Methylation-sensitive representational