Available online at www.sciencedirect.com
ScienceDirect
Genomics 90 (2007) 314–323
GENOMICS
www.elsevier.com/locate/ygeno

# Global, comparative analysis of tissue-specific promoter CpG methylation ☆

Elmar Schilling, Michael Rehli *

*Department of Hematology and Oncology, University Hospital, 93042 Regensburg, Germany*

Received 10 January 2007; accepted 28 April 2007
Available online 20 June 2007

**Abstract**

Understanding cell-type-specific epigenetic codes on a global level is a major challenge after the sequencing of the human genome has been completed. Here we applied methyl-CpG immunoprecipitation (MCIp) to obtain comparative methylation profiles of coding and noncoding genes in three human tissues, testis, brain, and monocytes. Forty-four mainly testis-specific promoters were independently validated using bisulfite sequencing or single-gene MCIp, confirming the results obtained by the MCIp microarray approach. We demonstrate the previously unknown somatic hypermethylation at many CpG-rich, testis-specific gene promoters, in particular in ampliconic areas of the Y chromosome. We also identify a number of miRNA genes showing tissue-specific methylation patterns. The comparison of the obtained tissue methylation profiles with corresponding gene expression data indicates a significant association between tissue-specific promoter methylation and gene expression, not only in CpG-rich promoters. In summary, our study highlights the exceptional epigenetic status of germ-line cells in testis and provides a global insight into tissue-specific DNA methylation patterns.


*Keywords:* DNA methylation; Epigenetics; Gene regulation; Methylation profiling

The epigenetic code comprises a second layer of information on top of the genetic code. Whereas the latter provides the framework for RNA and protein structure, the epigenetic code controls the packaging of DNA and ultimately regulates gene expression [1,2]. In mammals, methylation of cytosine residues in CpG dinucleotides is a frequent epigenetic mark that is required for normal embryonic development. DNA methylation plays important roles in gene regulation, e.g., in X-chromosome inactivation, imprinting, or silencing of retrotransposons, and is mostly associated with gene repression [1,2]. Global methylation patterns are established at the time of implantation and usually only a small portion of the genome remains unmodified [3]. Regions that are generally free of methylated CpG dinucleotides include the so-called CpG islands that are believed to be protected from methylation through *cis*-acting elements [4]. In contrast, CpG-depleted regions that are involved in gene regulation often show a dynamic, tissue-specific pattern of CpG methylation [1,2].

The extent of tissue-specific hyper- or hypomethylation throughout the genome is largely unknown, and its role in gene regulation and development has been controversially discussed [5,6]. The key to a better understanding of DNA methylation will be a more global picture of its tissue-specific and interindividual variations. However, few methods can be applied to detect DNA methylation genome-wide. Most currently used techniques to map global methylation differences involve the restriction of DNA with methylation-sensitive enzymes [7,8], thereby limiting these approaches to a relatively small number of unique restriction sites. Some improvement of those strategies may be offered by combining traditional methylation-sensitive enzymes with McrBC, an enzyme complex from *Escherichia coli* that degrades methylated DNA between two methylated half-sites of the form (G/A)mC separated by 40–3000 bp, as demonstrated in some recent studies [9,10]. The most accurate methylation profiling techniques are based on bisulfite treatment of DNA. However, those approaches are expensive, laborious, and not readily applicable on a genome-wide level. Nevertheless, the Human

☆ Sequence data from this article have been deposited with the GEO Data Library under Series Entry GSE5548.

* Corresponding author. Fax: +49 941 944 5593.
*E-mail address:* michael.rehli@klinik.uni-regensburg.de (M. Rehli).

0888-7543/$ - see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.ygeno.2007.04.011

Epigenome Project has started to map human CpG methylation using bisulfite technology, and the analysis of about 2500 amplicons across three chromosomes in 12 different tissues has been published recently [11]. Although this approach significantly advances our understanding of human epigenetics, it is clear that cheaper and faster high-throughput profiling platforms are needed to expedite the decoding of the human epigenome.

The separation of relatively small pieces of genomic DNA on the basis of their methylation density might represent an interesting technical alternative to current approaches. Here, we describe an adaptation of the methyl-CpG immunoprecipitation technique [12] for the detection of hypomethylated DNA fragments. We present global comparative methylation patterns for gene promoters in three human tissues and identify many promoters showing tissue-specific hypermethylation in brain, testis, or monocytes. In combination with whole-genome microarrays, our approach is suited to detecting differences in tissue- or cell-type-specific DNA methylation patterns in nonrepetitive regions throughout the genome.

## Results and discussion

### *Genome-wide detection of tissue-specific hypomethylation by methyl-CpG immunoprecipitation*

Methyl-CpG immunoprecipitation (MCIp) is based on a recombinant MBD2–Fc fusion protein that was originally designed by our lab to detect globally disease-related hypermethylation in CpG islands [12,13]. In contrast to another recently developed immunoprecipitation approach using 5-methyl-specific antibodies (called MeDIP or mDIP) that specifically enriches for methylated fragments [14,15], MCIp can divide the bulk of genomic DNA fragments into separate fractions of increasing methylation density [12]. Therefore, it should allow the simultaneous, quantitative analysis of the whole range of DNA-methylation density, including both hyper- and hypomethylated DNA.

It is currently believed that tissue-specific variations in DNA methylation largely occur in CpG-depleted promoters, whereas most CpG island promoters remain unmethylated in all tissues [2]. We hypothesized that most tissue-specific methylation differences may therefore be detected in the nonbinding fraction of a methyl-CpG immunoprecipitation that contains DNA fragments with no or only very few methylated CpG dinucleotides. We modified the previously described procedure to separate unmethylated DNA fragments rapidly as follows. In contrast to the original approach, genomic DNA was fragmented to a mean of 400–500 bp by ultrasonication rather than restriction digestion to avoid bias. To allow the direct labeling of DNA fragments for subsequent microarray analysis without prior amplification, we used 4 μg, instead of 300 ng, of genomic DNA fragments that were processed in large spin columns, and we processed and analyzed the unbound DNA fraction that was obtained immediately after the binding reaction. The salt concentration of the binding buffer was adjusted to allow the flowthrough of fragments methylated at few sites (one or two mCpG's). The final strategy for purification and subsequent detection of the fractionated material is summarized in Fig. 1, which also shows a typical MCIp fractionation of a maternally imprinted gene (*SNRPN*). As demonstrated in this example, the modified MCIp approach efficiently separates the methylated and the unmethylated alleles in somatic tissues, whereas the majority of this locus, as expected, appears unmethylated in testis.

To demonstrate the feasibility of this method to detect tissue-specific differences in methylation levels on a global scale, we separated unmethylated DNA fragments from three different tissues (testis, brain, monocytes) by MCIp. The purified DNA fragments were labeled directly with fluorescent dyes. Testis DNA fragments were cohybridized with either monocyte or brain DNA fragments onto human promoter arrays containing 44,000 oligonucleotide probes (45- to 60-mer) corresponding to nonrepetitive sequences and representing approximately 17,000 annotated genes and approximately 20,000 promoters. To limit interindividual differences and to increase the significance of our analysis, we performed two hybridizations per comparison using two independent testis samples (biological replicates). Fig. 2 shows representative scatter plots of MCIp microarray analyses. On a global scale, testis DNA fragments reproducibly appeared to be less methylated than brain or monocyte DNA, and brain was usually less different from testis than monocytes.

We assumed that DNA methylation in proximal promoter regions (−600 to +200 relative to the transcription start site) would be most influential in terms of gene regulation and focused our analysis on oligonucleotide probes located in proximal promoter regions that were enriched significantly ($p<0.01$ for processed signals) and more than 2.4-fold different between samples to ensure a major difference in the degree of methylation. At a mean fragment size of 400–500 bp, a single oligonucleotide probe should detect methylation events within a radius of at least 350 bp. Since we allowed the flowthrough of fragments methylated at one or two mCpG's during the MCIp binding reaction, a given DNA fragment should contain at least four or five CpG residues within the above radius, to allow the reproducible separation of unmethylated and methylated fragments. Therefore, we calculated the number of CpG residues in a region 350 bp upstream and downstream of each probe (CpG index) and restricted our analysis to probes with a CpG index above 4, thereby excluding less than 10% of all probes in proximal promoter regions and 4% of all proximal promoters. A proximal gene promoter was considered comparatively less methylated (hypomethylated) if at least one promoter probe was consistently different in both hybridization experiments using the above criteria. A promoter was grouped as hypomethylated in testis if it was less methylated compared to both somatic tissues. Promoters were grouped as hypomethylated in brain (monocytes) if they appeared less methylated in the brain (monocyte)/testis comparisons but not significantly different in the monocyte (brain)/testis hybridizations. In total, we identified 283 gene promoters that were specifically hypomethylated in testis. In brain and monocytes, 123 and 190 gene promoters, respectively, were specifically

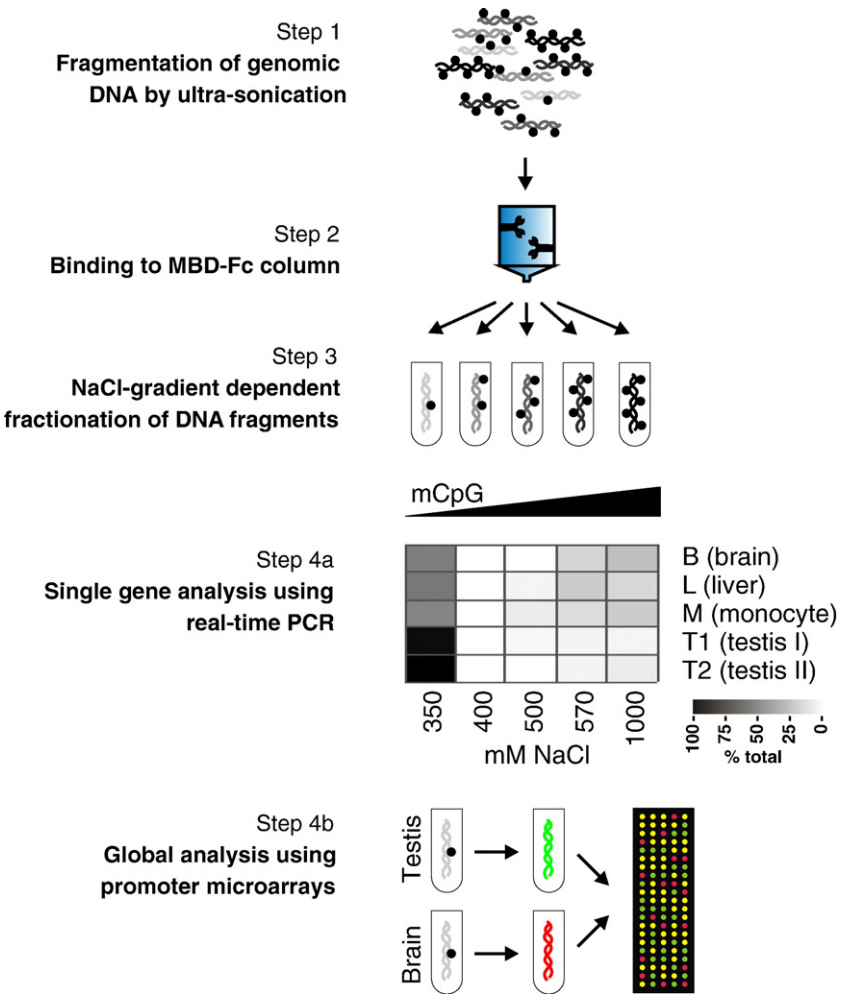



Fig. 1. MCIp-on-Chip approach to detect promoter regions with tissue-specific hypomethylation. Genomic DNA of tissue or purified cell populations is fragmented to a mean size of 400–500 bp using ultrasonication (Step 1). The material is bound to MBD–Fc Sepharose beads at 350 mM NaCl for 3 h (Step 2) and fractionated using wash/elution buffers containing increasing concentrations of NaCl (400–1000 mM, Step 3). For single-gene analysis, the amount of a single gene fragment may be quantified by real-time PCR in each fraction (Step 4a). The data of an exemplary analysis of the *SNRPN* promoter are graphically shown for the five tissue samples brain, liver, monocytes, testis sample 1, and testis sample 2 as a heat map. The percentage of amplified material in a given fraction is represented by gray coloring. Since the *SNRPN* promoter is subject to maternal imprinting, approximately half of the amplified material in somatic tissues is present in the 350 mM fraction (unmethylated). The second half elutes mainly above 570 mM NaCl, representing the methylated alleles. As expected, germ-line cells in testis lack the maternal imprint and approximately 75% of the *SNRPN* fragments are found in the unmethylated fraction. As an alternative to single-gene analysis, individual fractions (here the unmethylated 350 mM NaCl fraction) of two tissues may be used for dual-color microarray analysis (Step 4b).

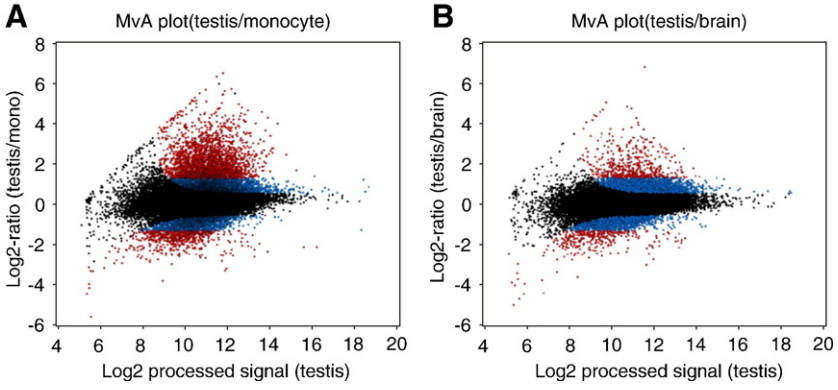



Fig. 2. Global profiling of tissue DNA methylation using human promoter microarrays. MvA plots (log2 signal ratio versus log2 processed signal testis) are shown for representative hybridization experiments of human testis versus (A) monocytes or (B) brain. Black spots represent probes that were not significantly changed ($p > 0.01$), blue spots were significantly enriched (ratio <2.4), and red spots were significantly enriched ($p < 0.01$) and >2.4-fold different.

hypomethylated. A detailed list of all 782 probes representing 596 identified promoters is given in Supplementary Table 1.

*Structure of tissue-specifically hypomethylated promoters*

Assuming that a single oligonucleotide probe should detect methylation events within a radius of at least 350 bp (as detailed above) we expected that methylated CpG islands or *Alu* repeats may affect a probe signal within a 250-to 350-bp distance. We therefore determined the presence of classical CpG islands [16] or *Alu* repeats [17] in the vicinity of individual probes at radii of 250, 350, and 850 bp around each probe. Notably, strikingly more hypomethylated promoters were associated with CpG islands in testis than in either somatic tissue (see Fig. 3A) at all distances ($p<0.0001$). This observation is in line with the previously postulated general absence of tissue-specific variation in somatic CpG island methylation. The marked association of testis-specific hypomethylation with CpG island promoters extends earlier studies reporting the somatic methylation of individual, so-called cancer-testis antigens, which are generally controlled by CpG-rich promoters [18,19].

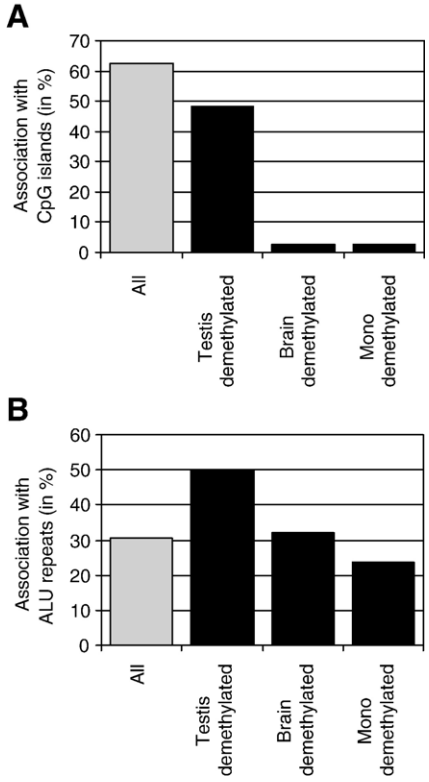



Fig. 3. Association of hypomethylated promoters with CpG islands and *Alu* repeats. Graphical presentations of a probe (50 bp), detected fragments (mean of 450 bp), and distances used to determine overlaps with (A) CpG islands and (B) *Alu* repeats. The presence of CpG islands or *Alu* repeats in the vicinity of proximal promoter probes (±250, 350, or 850 bp) was calculated for all individual promoter regions with a CpG index above 4 (All; 19,145 promoters in total) or within each hypomethylated group (testis, brain, and monocytes with 283, 123, and 190 promoters, respectively). The percentage of association of hypomethylated probes within a given radius with either (A) CpG islands or (B) *Alu* repeats is shown relative to all promoters in a group.

Previous studies indicated that *Alu* repeats in testis may be less methylated in general than those in somatic cells; however, little information on the positions of hypomethylated *Alu's* is available [20]. Oligonucleotide probes on the microarray do not correspond directly to repetitive sequences; however, they may be located next to repeats and detect their methylation status as detailed above. When comparing the percentage of hypomethylated promoters that are associated with *Alu* repeats, we noticed a significantly higher incidence of *Alu* repeats in the vicinity of promoter sequences that were hypomethylated in testis compared to the other two tissues. The difference was obvious at all distances but was most significant at 250 bp around individual hypomethylated probes ($p<0.0001$). This represents further indication of a role for *Alu* repeats in testis-specific gene regulation.

Genes with tissue-specifically hypomethylated promoters were generally found on all chromosomes (data not shown). An accumulation of testis-specific hypomethylation that has not been observed previously was detected on the Y chromosome, on which several of the annotated genes appeared hypomethylated compared to both somatic tissues (see Fig. 4A and below) and, as described earlier, on the X chromosome [21].

*Validation of MCIp microarray results*

To validate the MCIp microarray data, we initially focused on the Y chromosome and analyzed all CpG-rich promoters of testis-specific genes in ampliconic areas [22] and a number of additional, randomly chosen genes of the Y chromosome by MCIp and subsequent real-time PCR. As shown in Fig. 4B, single-gene MCIp data corresponded well with the microarray data confirming the somatic hypermethylation of most testis-specific genes, especially in ampliconic areas of the Y chromosome.

A similar degree of correlation was observed with tissue-specifically hypomethylated gene promoters on other chromosomes. The extended analysis of 44 gene promoters using MCIp and/or bisulfite sequencing is summarized in Table 1 (heat maps and additional bisulfite sequencing data are available as Supplementary Figs. 1 and 2, respectively). A number of previous studies already noted the testis-specific hypomethylation of selected genes. Our analysis confirmed the testis-specific methylation status published for *DDX4* [7] and *ACTL7B* [23], as well as *TAF7L* and *LUZP4* [21]. In our microarray analysis, *MAGEA1* and *TKTL1*, which had been analyzed previously by bisulfite sequencing [18,21], were detected as hypomethylated in only one of the two testis samples tested. The latter was also true for two of three genes (*TYROBP* and *MIRN363*) that tested as false negatives on the microarray. The causes of discrepant results in some cases are unknown. Sources of potential errors include the possibly inefficient or nonspecific detection of individual probes on the microarray, normalization artifacts during data processing, or experimental variation, e.g., through interexperimental differences in ultrasonication efficiency. In total, 35 of 40 validated gene promoters demonstrated the same methylation profile in global and single-gene analysis. Taken

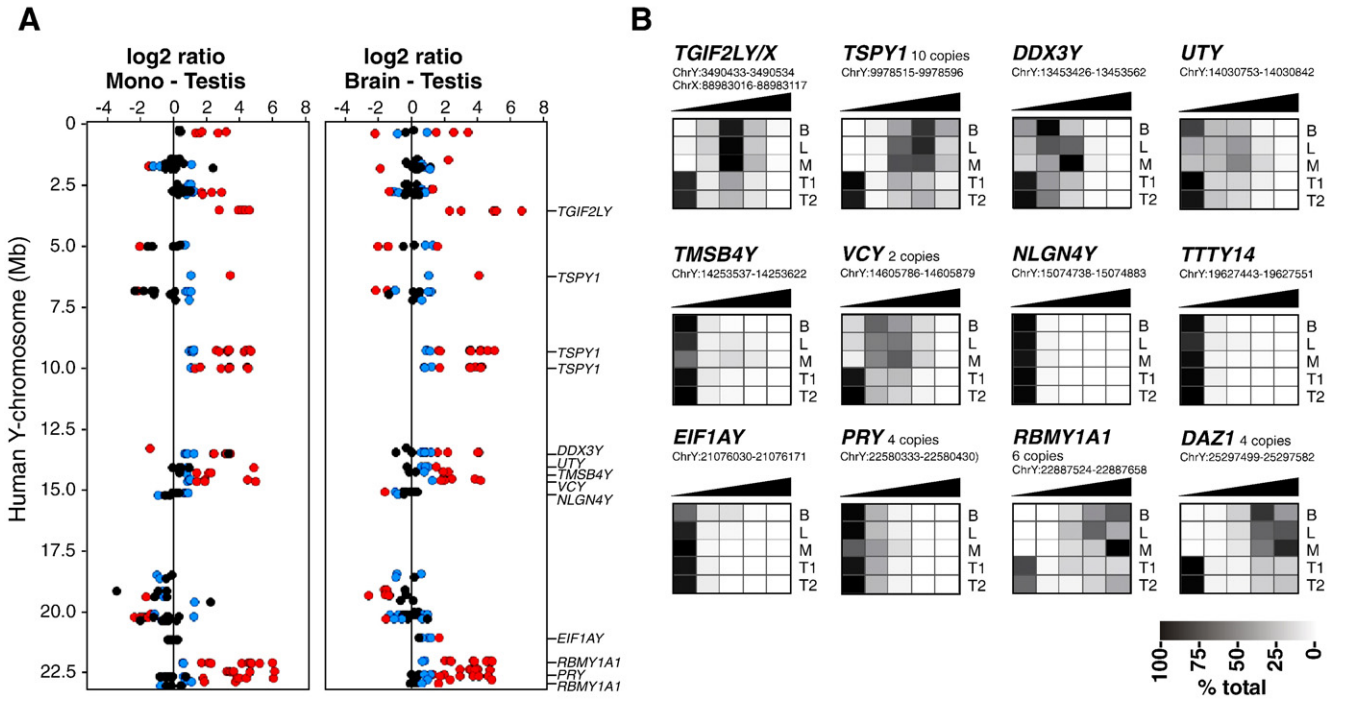



Fig. 4. Tissue-specific methylation profiles of the human Y chromosome. (A) Schematic representation of microarray results from one of two hybridizations for the human Y chromosome. Log2 signal ratios for each probe are represented relative to their absolute chromosomal position (based on NCBI build 35). Black spots represent probes that were not significantly changed ($p > 0.01$), blue spots were significantly enriched but less than 2.4-fold different, and red spots were significantly enriched ($p < 0.01$) and >2.4-fold different. The positions of genes shown in (B) are indicated. (B) Graphical presentation of single-gene MCIp real-time PCR data for several tested Y-chromosome genes. Gene symbols, copy numbers, and representative chromosomal localizations of the amplified product (based on NCBI build 35 of the human genome sequence) are given above the heat maps, which represent the percentages of amplified material in a given NaCl fraction by gray coloring for DNA samples of brain (B), liver (L), monocytes (M), testis sample 1 (T1), and testis sample 2 (T2) as described for Fig. 1.

together, the MCIp microarray approach demonstrated a high reproducibility.

It is noteworthy that tissue-specific variations in promoter methylation were also detected in several noncoding microRNAs as exemplified by *MIRN127, 142, 338*, and *363* (see Table 1 and Supplementary Fig. 1). A recent publication described the testis-specific demethylation of *MIRN127* and its activation in somatic tissues by chromatin-modifying drugs [24], confirming our MCIp result.

*Evidence for promoter-specific mechanisms establishing the hypermethylation of some CpG islands in somatic tissues*

In addition to the large number of testis-specific gene promoters that demonstrated hypermethylation in somatic tissues, we also noted a large group of testis-specific genes that did not exhibit a differential methylation status in any of the studied tissues, suggesting that only a subgroup of testis-specific genes is affected by somatic hypermethylation. For example, *AKAP3*, *DMRT1*, *TEKT3*, or *TTTY14* gene promoters were generally unmethylated in all tissues examined (see Fig. 4B and Supplementary Fig. 1). We also noted that copies of Y-linked genes like DAZL (chromosome 3) or VCX (chromosome X) that share almost identical promoter regions with their gene homologues on chromosome Y (DAZ and VCY, respectively) also show somatic hypermethylation, whereas homologous genes that are regulated by different promoter sequences (data not shown), like *RBMX* (chromosome X) or *TSPYL5* (chromosome 8), show a different methylation profile compared to their Y-chromosome homologues *RBMY1A1* and *TSPY1* (see Fig. 4B and Supplementary Fig. 1). The *SYBL1* genes on both X and Y chromosomes appear to represent an exception in having almost identical promoters but being solely methylated on the Y chromosome [25]. Based on the methylation characteristics of *SYBL1*, the existence of Y-chromosome inactivation resembling the well-known X-chromosome inactivation in females was previously proposed [25]. However, according to the methylation pattern of all other Y-chromosome genes that were analyzed in our study, this kind of phenomenon appears to be restricted to the short pseudo-autosomal region on Yq, which is flanked by a large heterochromatic region.

The above observations strongly suggest the presence of sequence-specific regulatory mechanisms controlling either the specific demethylation of some promoters in testis or germ-line cells or their global methylation in somatic tissues. Common *cis*-acting sequences responsible for this phenomenon are unknown so far and we failed to associate specific DNA motifs with either group of promoters. The methylation of CpG islands in lymphocytes was recently shown to correlate with DNA sequence, repeats, and predicted DNA structure using bioinformatics scoring and prediction methods, suggesting that multiple factors may predispose CpG islands for DNA methylation [26]. Recent studies have shown that

Table 1
Profiles of promoter hypomethylation and mRNA expression for selected genes in brain, testis, and monocytes

| Gene symbol | CGI promoter[a] | Chromosome | MCIp array | MCIp PCR[b] | Bisulfite[c] | mRNA expression[d] |
|---|---|---|---|---|---|---|
| *BACH* (ACOT7) | No | 1 | Brain[e] | Brain[e] | Brain[e] | Brain |
| *CORT* | No | 1 | Brain | Brain | — | Ubiquitous |
| *DAZL* | Yes | 3 | Not diff.[f] | Testis | — | Testis |
| *MORC1* | Yes | 3 | Testis | Testis | — | Testis |
| *TSPYL5* | Yes | 8 | Testis, brain | Testis, brain | — | Testis, brain, blood |
| *DMRT1* | Yes | 9 | Not diff. | Not diff. | — | Testis |
| *TLR4* | No | 9 | Mono | Mono | Mono | Blood |
| *LMO2* | No | 11 | Mono | Mono | — | Blood, endothelium |
| *CLPB* | Yes | 11 | Testis, brain | Testis, brain | — | Testis |
| *SPTBN2* | No | 11 | Brain | Brain | — | Testis, brain, other |
| *AKAP3* | No | 12 | Not diff. | Not diff. | — | Testis |
| *FGF6* | Yes | 12 | Testis | Testis | — | Testis, muscle |
| *H1FNT* | No | 12 | Testis | Testis | — | Testis |
| *MGP* | No | 12 | Mono | Not diff. | — | Ubiquitous (not blood) |
| *MIRN127* | Yes | 14 | Testis | Testis | — | Testis |
| *APBA2* | No | 15 | Brain | Brain | — | Brain, blood |
| *SNRPN* | Yes | 15 | Testis | Testis | — | Ubiquitous (imprinted) |
| *PRM1* | No | 16 | Not diff. | Not diff. | — | Testis |
| *MIRN142* | No | 17 | Mono | Mono | — | n.a. |
| *MIRN338* | No | 17 | Brain | Brain | — | n.a. |
| *TEKT3* | Yes | 17 | Not diff. | Not diff. | — | Testis |
| *AURKC* (up)[g] | Yes | 19 | Testis | Testis | — | Testis |
| *TYROBP* | No | 19 | Not diff. (mono)[h] | Mono | — | Blood |
| *ZSCAN5* | Yes | 19 | Testis | — | Testis | Testis |
| *PRAME* | Yes | 22 | Testis | Testis | — | Testis |
| *DCX* | No | X | Mono | Not diff. | — | Fetal brain |
| *MAGEB10* | No | X | Testis | Testis | Testis | n.a. |
| *MIRN363* | No | X | Not diff. (brain) | Brain | — | n.a. |
| *RBMX* | Yes | X | Not diff. | Not diff. | — | Ubiquitous |
| *TAF7L* | No | X | Testis | Testis | — | Testis |
| *VCX* | No | X | n.a. | Testis | — | Testis |
| *DAZ1* | Yes | Y | n.a. | Testis | — | Testis |
| *DDX3Y* (up)[g] | No | Y | Testis | Testis | — | Testis, blood |
| *EIF1AY* | No | Y | Not diff. | Not diff. | — | Blood |
| *NLGN4Y* | Yes | Y | Not diff. | Not diff. | — | Ubiquitous |
| *PRY* | No | Y | Not diff. | Not diff. | — | Testis |
| *RBMY1A1* | Yes | Y | Testis | Testis | — | Testis |
| *SYBL1* | Yes | Y | n.a. | Testis | — | Testis |
| *TGIF2LY* | No | Y | Testis | Testis | Testis | Testis |
| *TMSB4Y* | Yes | Y | Not diff. | Not diff. | — | Ubiquitous |
| *TSPY1* | Yes | Y | Testis | Testis | — | Testis |
| *TTTY14* | Yes | Y | n.a. | Not diff. | — | Testis |
| *UTY* | No | Y | Testis, brain | Testis, brain | — | Ubiquitous |
| *VCY* | No | Y | Testis | Testis | — | Testis |

[a] The presence or absence of a CpG island in the 5′-proximal promoter region (CGI promoter) was determined using the RepeatMasker function in the UCSC Genome browser (http://genome.ucsc.edu/).

[b] Single-gene MCIp data are shown in Fig. 4 and Supplementary Fig. 1. Data obtained with the liver DNA sample were not included.

[c] Bisulfite data are shown in Supplementary Fig. 2.

[d] Expression data indicate the primary source of expression according to the reference databases SymAtlas and RefEXA (see Materials and methods).

[e] The indicated tissue(s), e.g., brain, was found to be less methylated (hypomethylated) compared to the two other tissues used in the microarray study, e.g., testis and monocytes. Data obtained with the liver DNA sample in single gene MCIp were not included.

[f] Abbreviations used: not diff., not significantly different; n.a., not available.

[g] Single-gene MCIp PCR primers for *AURKC* and *DDX3Y* were located >600 bp upstream (up) of the putative transcription start site in a region that was differentially methylated as indicated. The proximal region was not significantly different in both cases.

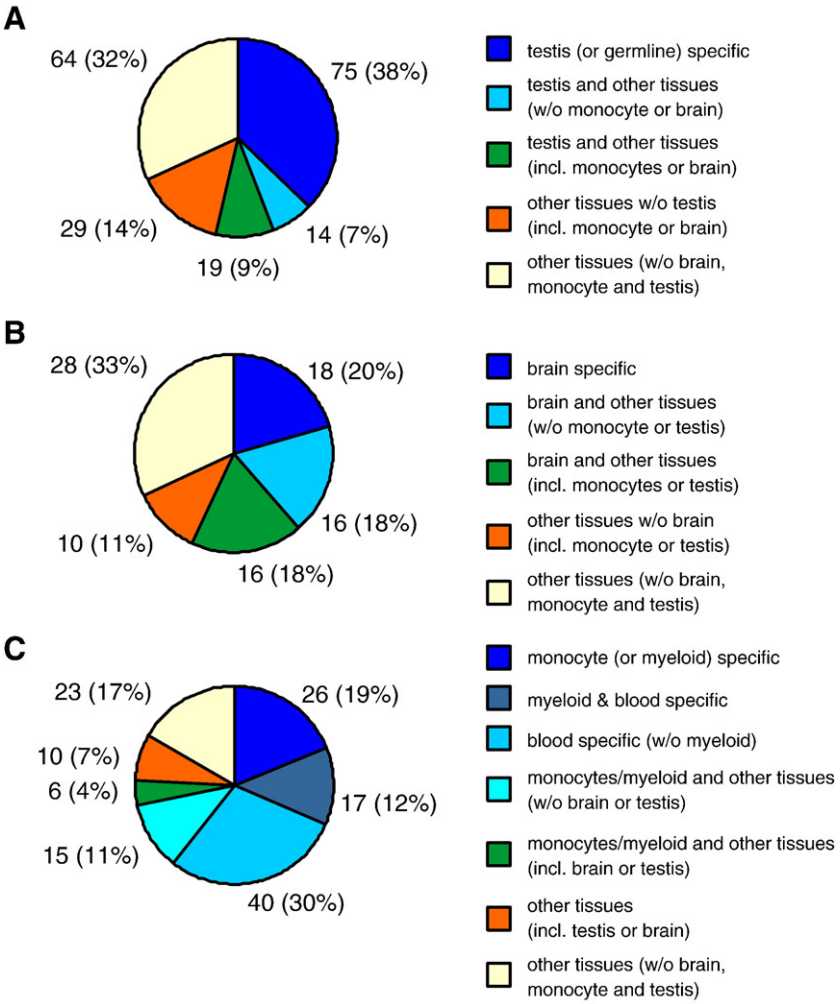
[h] The promoters of *TYROBP* and *MIRN363* were significantly different in only one of two hybridizations; however, they showed the same trend in each case (the respective tissue is given in parentheses).

the testis-specific nuclear protein and the CTCF paralog BORIS/CTCFL [27] can initiate the demethylation of the cancer-testis antigens *MAGEA1* [28] and *NY-ESO-1* [29] in normal somatic cells, suggesting that this transcription factor may be involved in the testis-specific demethylation of at least some of the somatically methylated promoters. Another recent report indicates that E2F6 may be essential for the long-term somatic silencing of certain male-germ-cell-specific genes [30]. Taken together, the above observations favor a model in which specific *cis*-acting sequences control promoter

methylation at the level of somatic hypermethylation rather than testis-specific demethylation.

*Relationship between DNA methylation and gene expression*

In contrast to CpG island promoters, the effect of DNA methylation on CpG-depleted promoters has been controversial. A number of recent studies did not support a major role for DNA methylation in tissue-specific gene regulation [5,6]. However, so far, systematic comparisons on a genome-wide level have not been performed. To correlate our CpG methylation data and gene expression patterns, we extracted RNA expression data from publicly available reference databases (Symatlas, RefEXA) for all corresponding differentially methylated fragments. The distribution of expression patterns for tissue-specifically hypomethylated fragments is summarized in Figs. 5A–5C. The largest group of testis-hypomethylated genes (38%) was specifically expressed in testis. Expression of genes that were hypomethylated in monocytes or brain was also often specific for the respective gene (19 or 20%, respectively). Interestingly, most genes in the monocyte-hypomethylated group were transcribed specifically in the




D

| expressed only in \ demethylated in | testis | brain | mono |
|---|---|---|---|
| testis | 75 | 1 | 5 |
| brain | 2 | 18 | 2 |
| mono | 2 | 1 | 26 |

Fig. 5. Comparison of methylation and expression profiles. (A–C) Distribution of expression patterns for genes with hypomethylated promoters in (A) testis, (B) brain, or (C) monocytes. Only genes with traceable expression profiles were included. (D) Distribution of testis-, brain-, or monocyte-specific genes in each demethylated group.

hemopoietic system, which may suggest an association of DNA methylation patterns with the origin of a cell type (hemopoietic) rather than the specialized cell type itself (monocyte). The coexistence of DNA methylation and transcription in a certain tissue was also detected. Transcription may indeed proceed in the presence of DNA methylation at those sites. However, there are a number of possibilities that may also explain these observations, including alternative promoter usage or simply false-positive microarray signals. A larger number of microarray experiments may be necessary for an in-depth analysis and further analyses of individual genes will be required to clarify these issues.

Tissues are generally composed of different, specialized cell types, which likely results in mixed gene expression and methylation patterns, as exemplified by bisulfite sequences presented in Supplementary Fig. 2. It is therefore difficult, if not impossible, to correlate tissue methylation or expression patterns for all genes, because both signatures may be determined by independent cell populations within the tissue. In the case of highly tissue-specific genes a significant correlation between expression and methylation pattern may exist, although some studies did not support this kind of relationship [5,6]. When we determined the number of genes that showed a restricted expression pattern in testis, monocytes, or brain within each group of hypomethylated genes, the number of genes expressed in a certain tissue was significantly enriched ($p < 0.0001$) in the hypomethylated group of the same tissue. This observation suggests that tissue-specific DNA hypomethylation correlates significantly with tissue-specific transcription (see Fig. 5D).

It is clear that the presented pilot study provides only an initial and limited view of the human methylation landscape. Interindividual methylation variability, which may be key to understanding the role of methylation in gene regulation, has not been addressed in our study and, due to the fact that tissues are composed of different cell types, only differences in dominant cell types are likely being detected in our experiments. Further studies using purified cell populations and including matched methylation and expression profiling will be necessary to understand completely the relationship between DNA methylation and gene expression.

### *Toward decoding the human epigenome*

We present a novel technical approach that allows the detection of hypomethylated DNA fragments on a global level. Our approach is specifically designed to enable the quantitative detection of cell type-specific differences in DNA hypomethylation with a reasonable resolution and should allow routine methylation profiling comparable to microarray-based expression studies. Although the presented study focused on promoter regions, the same strategy can also be used to screen methylation patterns across all nonrepetitive regions of the genome, depending on the availability of microarrays. In combination with existing technologies, the described MCIp microarray approach will likely make a significant contribution toward decoding the human epigenome.

## Materials and methods

### *Reagents*

All chemical reagents were purchased from Sigma–Aldrich (Taufkirchen, Germany) unless otherwise noted. Standard oligonucleotides were synthesized by Operon Biotechnologies GmbH (Cologne, Germany). HPLC-purified oligonucleotides were obtained from Metabion (Planegg–Martinsried, Germany). DNA sequencing was done by Geneart (Regensburg, Germany).

### *DNA preparation and fragmentation*

Human monocytes from male donors were isolated as previously described [31]. Genomic DNA (gDNA) from monocytes was prepared using the Blood and Cell Culture DNA Midi Kit from Qiagen (Hilden, Germany). Male gDNA from brain, liver, and two testis samples was purchased from BioChain Institute (Hayward, CA, USA). For MCIp, gDNA was fragmented as follows: To reduce viscosity, gDNA was initially sheared using a 0.4 × 19-mm needle attached to a 2-ml syringe (BD, Drogheda, Ireland) before quantification using either PicoGreen dsDNA quantitation reagent (Molecular Probes, Eugene, OR, USA) or the NanoDrop ND 1000 spectrophotometer (Peqlab, Erlangen, Germany). Ultrasonication to a mean fragment size of 400–500 bp was carried out with the Branson Sonifier 250 (Danbury, CT, USA) using the following settings: duty cycle 30%, output 3, sonication time 1 min, and 5 μg DNA in 500 μl TE buffer. Fragment range was controlled using agarose gel electrophoresis.

### *Methyl-CpG immunoprecipitation*

MBD–Fc protein was produced as previously described [12]. For single-gene analysis, typically 22.5 μg purified MBD–Fc protein per 40 μl Protein A–Sepharose 4 Fast Flow beads (Amersham Biosciences, Piscataway, NJ, USA) for a single assay was rotated in 2 ml TBS overnight at 4°C. On the next day, the MBD–Fc beads (40 μl/assay) were transferred and dispersed into 0.5-ml Ultrafree-MC centrifugal filter devices (Millipore, Billerica, MA, USA) and spin-washed twice with buffer A (20 mM Tris–HCl, pH 8.0, 2 mM $MgCl_2$, 0.5 mM EDTA, 350 mM NaCl, 0.1% NP-40). Sonicated DNA (300 ng) was added to the washed MBD–Fc beads in 350 μl buffer A and rotated for 3 h at 4°C. Beads were centrifuged to recover unbound DNA fragments (350 mM fraction) and subsequently washed twice with 200 μl/150 μl of buffers containing increasing NaCl concentrations (400, 500, 570, and 1000 mM). The flowthrough of each washing step was collected in separate tubes and desalted using a QIAquick PCR Purification Kit (Qiagen). In parallel, 300 ng sonicated input DNA was resuspended in 350 μl buffer E (20 mM Tris–HCl, pH 8.0, 2 mM $MgCl_2$, 0.5 mM EDTA, 1000 mM NaCl, 0.1% NP-40) and desalted using a QIAquick PCR Purification Kit (Qiagen) as a control. This MCIp protocol was scaled up to generate DNA fragments for direct microarray hybridization. Here, for each sample, 84 μg purified MBD–Fc protein was added to 200 μl Protein A–Sepharose beads (Amersham Biosciences) in 15 ml TBS and rotated overnight at 4°C. For the precipitation, 2-ml Ultrafree-MC centrifugal filter devices (Millipore) were used and 4 μg of sonicated DNA was added to the washed MBD–Fc beads in 2000 μl buffer A. Unbound DNA fragments (350 mM fraction) were collected and desalted using a QIAquick PCR Purification Kit (Qiagen) before subsequent labeling.

### *Real-time genomic PCR*

Primers were designed using PerlPrimer Software (http://perlprimer.sourceforge.net) and controlled using PCR and BLAT functions of the UCSC Genome Browser (http://genome.ucsc.edu/). Enrichment of a specific fragment in the MCIp eluate was detected and quantified relative to the genomic input by real-time PCR on a Mastercycler EP Realplex (Eppendorf, Hamburg, Germany) using the QuantiTect SYBR Green PCR Kit (Qiagen) according to the manufacturer's instructions. Primer sequences are given in Supplementary Table 2. Cycling parameters were denaturation at 95°C for 15 min and amplification at 95°C for 12 s, 57°C for 20 s, 72°C for 25 s, for 50 cycles. The specificity of the PCR was controlled by melting-curve analysis. The relative amounts of

amplified fragment were calculated from a standard curve plotting four different concentrations of log dilutions of the genomic input DNA against the PCR cycle number at which the measured fluorescence intensity reached a fixed value. For each gene analyzed, at least two independent MCIp preparations were measured in duplicate. For the monocyte samples, two independent samples revealed the same small degree of deviation as technical replicates from liver and brain and were therefore analyzed in combination. Heat maps were created using Spotfire Decision Site Software 7.0 (Spotfire, Göteborg, Sweden). The mean percentage of totally amplified fragment (sum of all five MCIp fractions) in each individual MCIp fraction is graphically displayed. Black represents 100%, white represents 0%, and intermediate values are displayed in gray.

### *Microarray handling and analysis*

To generate fluorescently labeled DNA for microarray hybridization the 350 mM fraction of a MCIp immunoprecipitation from somatic and testis cells, genomic DNA was labeled directly with Alexa Fluor 555–aha–dCTP and Alexa Fluor 647–aha–dCTP, respectively, using the BioPrime Plus Array CGH Genomic Labeling System (Invitrogen, Carlsbad, CA, USA). The labeling reaction was carried out according to the manufacturer's manual. The differently labeled genomic DNA fragments of two tissues were combined to a final volume of 150 μl, supplemented with 50 μg Cot-1 DNA (Invitrogen), 50 μl of Agilent blocking agent (10-fold) (Agilent Technologies, Böblingen, Germany), and 250 μl Agilent hybridization buffer (2-fold) as supplied in the Agilent oligo aCGH Hybridization Kit. The sample was heated to 95°C for 3 min, mixed, and subsequently incubated at 37°C for 30 min and spun down afterward for 1 min. Hybridization on Human Proximal Promoter microarrays (Agilent) was then carried out at 65°C for 40 h using an Agilent SureHyb chamber and an Agilent hybridization oven. Slides were washed in Wash I (6× SSPE, 0.005% *N*-lauroylsarcosine) at room temperature for 5 min and in Wash II (0.06× SSPE) for an additional 5 min. Afterward slides were dried and incubated using acetonitrile and Agilent ozone protection solution, respectively, for 30 s. Images were scanned immediately and analyzed using a DNA microarray scanner (Agilent). Microarray images were processed using Feature Extraction Software 8.5 (Agilent) using the standard CGH protocol and a rank consistent list of normalization genes. The latter contained all probes with a CpG index above 78 (more than 78 CpG residues in a region of 350 bp upstream and downstream of a probe). Processed data were imported into Microsoft Office Excel 2003 for further analysis. Graphical presentations of datasets were obtained using Spotfire Decision Site Software 7.0 (Spotfire).

### *Statistical calculations*

The statistical significance of log ratio values obtained for individual microarray probes was calculated using the "most conservative error model" function of Feature Extraction Software 8.5 (Agilent), by which both a propagated and a Universal error model are evaluated and the higher $p$ value (more conservative estimate of error) is reported. Error models are described in detail in the Feature Extraction Software manual (Agilent). For the analysis of concordance or differences of proportions of categorical values, numbers were tabulated in contingency tables and analyzed by the $\chi^2$ test using the GraphPad Prism software (GraphPad Software, San Diego, CA, USA).

### *Databases for sequence analysis*

The relative locations of microarray probe, CpG islands, and *Alu* repeats were determined using the UCSC Genome Browser (http://genome.ucsc.edu/). Tissue expression profiles were obtained from the two reference databases for gene expression, GNF SymAtlas (http://symatlas.gnf.org/SymAtlas/) and RefEXA (http://www.lsbm.org/site_e/database/).

## Acknowledgments

This work was supported by a grant from the Deutsche Forschungsgemeinschaft (Re1310/7). The authors acknowledge the excellent technical assistance of Lucia Schwarzfischer and Sabine Pape. They also thank Stefan Krause (Department of Hematology, Regensburg) for advice on statistical calculations, Peter Oefner (Institute for Functional Genomics, Regensburg) for reading the manuscript, Andreas Polten (Agilent Technologies, Böblingen) for his support in processing and analyzing microarrays, and lab members for helpful discussions.

## Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ygeno.2007.04.011.

## References

[1] A. Bird, DNA methylation patterns and epigenetic memory, Genes Dev. 16 (2002) 6–21.
[2] R. Jaenisch, A. Bird, Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals, Nat. Genet. 33 (2003) 245–254 (Suppl).
[3] W. Reik, W. Dean, J. Walter, Epigenetic reprogramming in mammalian development, Science 293 (2001) 1089–1093.
[4] S. Gidekel, Y. Bergman, A unique developmental pattern of Oct-3/4 DNA methylation is controlled by a cis-demodification element, J. Biol. Chem. 277 (2002) 34521–34530.
[5] C.P. Walsh, T.H. Bestor, Cytosine methylation and mammalian development, Genes Dev. 13 (1999) 26–34.
[6] P.M. Warnecke, S.J. Clark, DNA methylation profile of the mouse skeletal alpha-actin promoter during development and differentiation, Mol. Cell Biol. 19 (1999) 164–172.
[7] F. Song, et al., Association of tissue-specific differentially methylated regions (TDMs) with differential gene expression, Proc. Natl. Acad. Sci. USA 102 (2005) 3336–3341.
[8] A. Schumacher, et al., Microarray-based DNA methylation profiling: technology and applications, Nucleic Acids Res. 34 (2006) 528–542.
[9] R.A. Rollins, et al., Large-scale structure of genomic methylation patterns, Genome Res. 16 (2006) 157–163.
[10] Y. Yamada, et al., A comprehensive analysis of allelic methylation status of CpG islands on human chromosome 21q, Genome Res. 14 (2004) 247–266.
[11] F. Eckhardt, et al., DNA methylation profiling of human chromosomes 6, 20 and 22, Nat. Genet. 38 (2006) 1378–1385.
[12] C. Gebhard, et al., Genome-wide profiling of CpG methylation identifies novel targets of aberrant hypermethylation in myeloid leukemia, Cancer Res. 66 (2006) 6118–6128.
[13] C. Gebhard, et al., Rapid and sensitive detection of CpG-methylation using methyl-binding (MB)-PCR, Nucleic Acids Res. 34 (2006) e82.
[14] I. Keshet, et al., Evidence for an instructive mechanism of de novo methylation in cancer cells, Nat. Genet. 38 (2006) 149–153.
[15] M. Weber, et al., Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells, Nat. Genet. 37 (2005) 853–862.
[16] M. Gardiner-Garden, M. Frommer, CpG islands in vertebrate genomes, J. Mol. Biol. 196 (1987) 261–282.
[17] J. Jurka, Repbase update: a database and an electronic journal of repetitive elements, Trends Genet. 16 (2000) 418–420.
[18] C. De Smet, C. Lurquin, B. Lethe, V. Martelange, T. Boon, DNA methylation is the primary silencing mechanism for a set of germ line- and tumor-specific genes with a CpG-rich promoter, Mol. Cell Biol. 19 (1999) 7327–7335.
[19] M. Koslowski, et al., Frequent nonrandom activation of germ-line genes in human cancer, Cancer Res. 64 (2004) 5988–5993.
[20] S. Kochanek, D. Renz, W. Doerfler, DNA methylation in the Alu sequences of diploid and haploid primary human cells, EMBO J. 12 (1993) 1141–1151.
[21] M. Koslowski, U. Sahin, C. Huber, O. Tureci, The human X chromosome

is enriched for germline genes expressed in premeiotic germ cells of both sexes, Hum. Mol. Genet. 15 (2006) 2392–2399.
[22] H. Skaletsky, et al., The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes, Nature 423 (2003) 825–837.
[23] M. Hisano, H. Ohta, Y. Nishimune, M. Nozaki, Methylation of CpG dinucleotides in the open reading frame of a testicular germ cell-specific intronless gene, Tact1/Actl7b, represses its expression in somatic cells, Nucleic Acids Res. 31 (2003) 4797–4804.
[24] Y. Saito, et al., Specific activation of microRNA-127 with downregulation of the proto-oncogene BCL6 by chromatin-modifying drugs in human cancer cells, Cancer Cell 9 (2006) 435–443.
[25] M.R. Matarazzo, et al., Allelic inactivation of the pseudoautosomal gene SYBL1 is controlled by epigenetic mechanisms common to the X and Y chromosomes, Hum. Mol. Genet. 11 (2002) 3191–3198.
[26] C. Bock, et al., CpG island methylation in human lymphocytes is highly correlated with DNA sequence, repeats, and predicted DNA structure, PLoS Genet. 2 (2006) 243–252.
[27] D.I. Loukinov, et al., BORIS, a novel male germ-line-specific protein associated with epigenetic reprogramming events, shares the same 11-zinc-finger domain with CTCF, the insulator protein involved in reading imprinting marks in the soma, Proc. Natl. Acad. Sci. USA 99 (2002) 6806–6811.
[28] S. Vatolin, et al., Conditional expression of the CTCF-paralogous transcriptional factor BORIS in normal cells results in demethylation and derepression of MAGE-A1 and reactivation of other cancer-testis genes, Cancer Res. 65 (2005) 7751–7762.
[29] J.A. Hong, et al., Reciprocal binding of CTCF and BORIS to the NY-ESO-1 promoter coincides with derepression of this cancer-testis gene in lung cancer cells, Cancer Res. 65 (2005) 7763–7774.
[30] M. Pohlers, et al., A role for E2F6 in the restriction of male-germ-cell-specific gene expression, Curr. Biol. 15 (2005) 1051–1057.
[31] S.W. Krause, et al., Differential screening identifies genetic markers of monocyte to macrophage maturation, J. Leukocyte Biol. 60 (1996) 540–545.