RESEARCH PAPER

Epigenetics 5:6, 1-12; August 16, 2010; © 2010 Landes Bioscience

Tissue specific DNA methylation of CpG islands in normal human adult somatic tissues distinguishes neural from non-neural tissues

Srimoyee Ghosh,[1] Allan J. Yates,[2] Michael C. Frühwald,[3] Jeffrey C. Miecznikowski,[4] Christoph Plass[5] and Dominic J. Smiraglia[1,*]

[1]Department of Cancer Genetics; Roswell Park Cancer Institute; Buffalo, NY USA; [2]Department of Pathology; The Ohio State University; Columbus, OH USA; [3]Department of Pediatric Hematology and Oncology; University Children's Hospital of Münster; Münster, Germany; [4]Department of Biostatistics; State University of New York-Buffalo; Buffalo, NY USA; [5]Department of Epigenomics and Cancer Risk Factors; German Cancer Research Center (DKFZ); Heidelberg, Germany

Key words: Tissue specific methylation, CpG island methylation, neural, brain tissue, grey matter, white matter

Although most CpG islands are generally thought to remain unmethylated in all adult somatic tissues, recent genome-wide approaches have found that some CpG islands have distinct methylation patterns in various tissues, with most differences being seen between germ cells and somatic tissues. Few studies have addressed this among human somatic tissues and fewer still have studied the same sets of tissues from multiple individuals. In the current study, we used Restriction Landmark Genomic Scanning to study tissue specific methylation patterns in a set of twelve human tissues collected from multiple individuals. We identified 34 differentially methylated CpG islands among these tissues, many of which showed consistent patterns in multiple individuals. Of particular interest were striking differences in CpG island methylation, not only among brain regions, but also between white and grey matter of the same region. These findings were confirmed for selected loci by quantitative bisulfite sequencing. Cluster analysis of the RLGS data indicated that several tissues clustered together, but the strongest clustering was in brain. Tissues from different brain regions clustered together, and, as a group, brain tissues were distinct from either mesoderm or endoderm derived tissues which demonstrated limited clustering. These data demonstrate consistent tissue specific methylation for certain CpG islands, with clear differences between white and grey matter of the brain. Furthermore, there was an overall pattern of tissue specifically methylated CpG islands that distinguished neural tissues from non-neural.

Introduction

For over 30 years, the role of DNA methylation in mammals has been under investigation. It has been clearly demonstrated that DNA methylation is required for survival beyond early embryogenesis in mouse.[1,2] What is not yet clear is the exact nature of this absolute requirement for DNA methylation. It is well established that DNA methylation plays a critical role in X chromosome inactivation in females,[3] and in allele specific expression of imprinted genes.[4] In addition, a large body of data has shown correlation between lack of methylation of tissue specific genes in the expressing tissues and methylation of these genes in non-expressing tissues.[5,6] These data suggest that DNA methylation may play a critical role in mammalian development.

Of all CpG dinucleotides in vertebrate genomes, 60%–90% are methylated.[7,8] The majority of the CpG dinucleotides in the mammalian genome are scattered throughout bulk chromatin, generally heavily methylated, and thought to contribute to the repression of transcription from repetitive elements such as Alus and retrotransposons.[9] In human DNA, approximately 15% of the CpG dinucleotides are found in CpG islands,[7,8] defined as stretches of DNA >200 bp, with a G + C content greater than 50%, and an observed to expected ratio of CpG ≥0.6.[10] CpG islands are often found in the promoters of both "house keeping" genes and tissue specific genes. Aside from the inactivated X chromosome and imprinted genes, the thinking has been that CpG islands are normally unmethylated, at least in germ cells,[7] and perhaps throughout the developing and adult organism[5] for most CpG islands.

Tissue specific genes with non-CpG island promoters have been shown to exhibit tissue specific methylation. However, tissue specific genes with CpG island promoters initially received less attention. Earlier studies using candidate gene approaches led to conflicting ideas about tissue specific methylation of CpG islands. One study looking at the methylation status of seven tissue specific gene promoters found little or no methylation in the promoters of four genes with CpG island-like characteristics,[11] supporting the idea that tissue specific genes with CpG island promoters are not methylated, regardless of developmental state, tissue or expression. Other data, however, have shown that CpG islands can indeed become methylated in normal adult tissues, and this phenomenon may be related to age.[12] Recently a number

*Correspondence to: Dominic J. Smiraglia; Email: Dominic.Smiraglia@roswellpark.org
Submitted: 03/29/10; Accepted: 05/04/10
Previously published online: www.landesbioscience.com/journals/epigenetics/article/12228
DOI: 10.4161/epi.5.6.12228

of genomic-based studies looking at very large numbers of CpG islands in normal somatic tissues have been published that provide strong evidence that tissue-specific transcription is controlled, in part, by tissue-specific differentially methylated regions, both CpG island and non-CpG island.

Using bisulfite DNA sequencing, Eckhart et al. (2006)[13] reported high-resolution methylation profiles of human chromosomes 6, 20 and 22 from 12 different tissues. These data indicated that evolutionarily conserved regions are the predominant sites for differential DNA methylation and that methylation specifically within a core region surrounding the transcriptional start site is an informative surrogate for promoter wide methylation. Another study took a more genome wide approach at a lower resolution[14] to generate an epigenomic map of DNA methylation, RNA polymerase II occupancy and chromatin state for 16,000 promoters in human primary fibroblasts and mature sperm. The results showed that promoter sequence and gene function are major predictors of promoter methylation states and suggested that weak CpG islands are predisposed to de novo methylation during differentiation while strong CpG island promoters are mostly unmethylated, even when inactive.

A comparative genomic approach identified evolutionarily conserved tissue specific CpG island methylation between mouse and human. Such CpG islands identified in a genome wide scan of DNA from various mouse tissues[15] were studied in a candidate gene approach in human tissues.[16] Of the 14 mouse CpG islands with tissue specific methylation, six were found to conserve similar patterns of methylation in the human orthologs. Five of these correlated with gene expression patterns, providing further evidence that human gene expression could be regulated by tissue specific CpG island methylation-mediated gene silencing. Multiple additional groups have recently applied various genome wide CpG island methylation approaches[17-20] to confirm and broaden the ideas that CpG island methylation plays an important role in cellular identity and tissue specific gene expression and these regions mostly have a wide range of CpG densities.

Most studies have not addressed the question of inter individual variability of tissue specific CpG island methylation in human tissues. Only a few studies used the same tissues from multiple individuals at the candidate gene level,[16,18,20] but none of these address the issue by using genomic approaches on the same tissues from multiple individuals to determine if tissue specific methylation of the CpG islands was consistent. More recent genomics based approaches have addressed this issue and shown that CpG island methylation patterns in normal human tissues are more similar among the same tissue studied from multiple individuals than tissues from the same individual compared to each other.[21] Another study found that inter-individual differences in DNA methylation patterns could be correlated to age and environmental exposures such as tobacco use.[22]

Focusing on three regions of the brain, a recent study addressed the issue of inter-individual differences in tissue specific methylation.[23] This study determined the quantitative DNA methylation levels of 1,505 CpG sites across 76 brain samples representing cerebral cortex (n = 35), cerebellum (n = 34) and pons (n = 7). A set of 156 loc showed statistically significant methylation differences among brain regions, and unsupervised hierarchical analysis showed strong clustering of the cerebra distinct from the clustering of the cerebella, and pons. Their data suggest that DNA methylation signatures distinguish brain regions, are consistent across many individuals, and may help account for region-specific functional specialization.

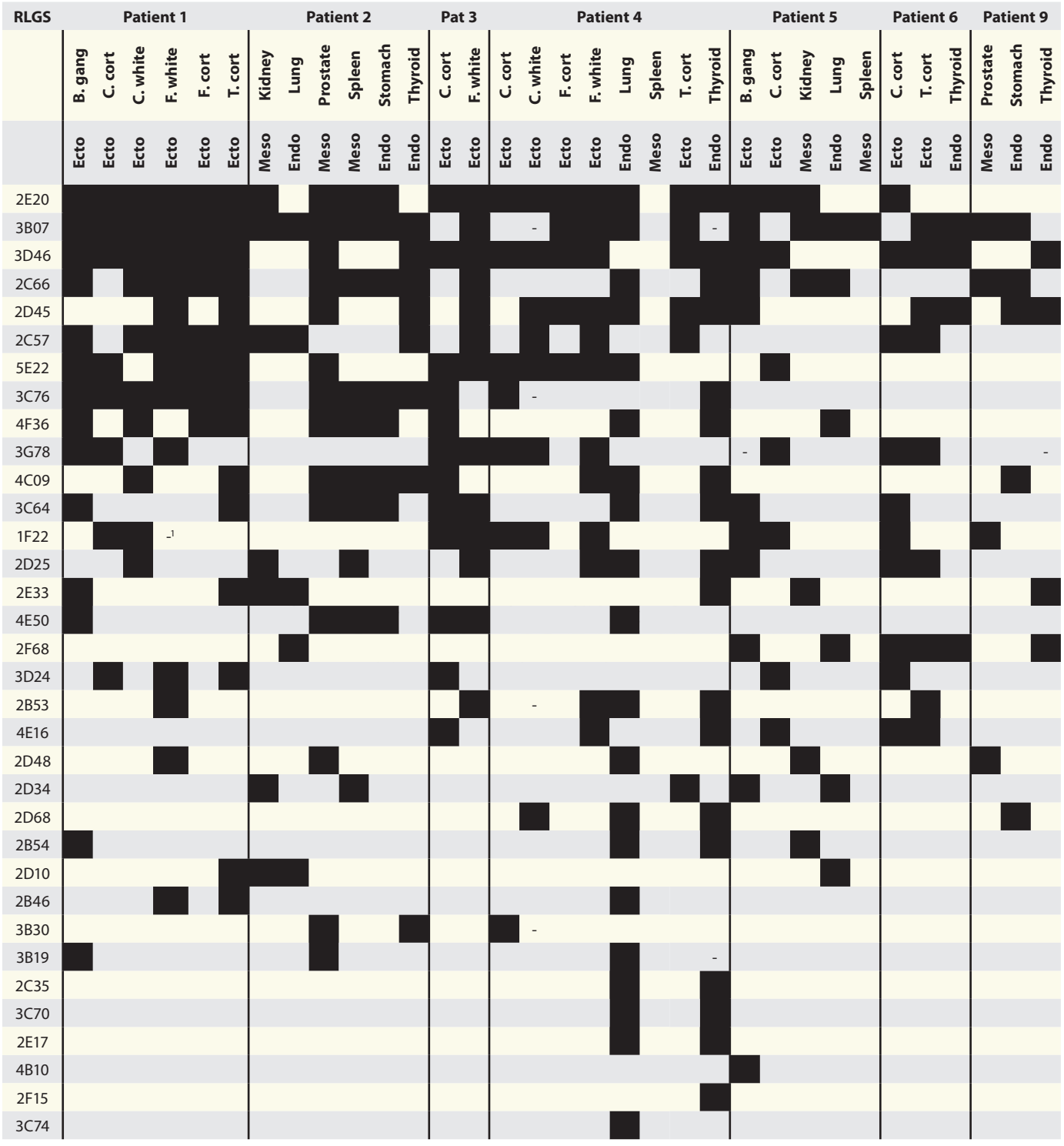
In the present study, we have analyzed an average of approximately 1500 *Not*I sites by Restriction Landmark Genomic Scanning (RLGS), from 12 different tissues obtained from 7 individuals. Each tissue was obtained from a minimum of two individuals and a maximum of five. Our data clearly show the existence of tissue specific CpG island methylation among regions of the brain as well as among other tissue types and that many of these methylation differences are preserved among different individuals. Furthermore, we demonstrate that differences in tissue specific methylation not only exists in different regions of the brain, as previously reported,[23] but also between the grey and white matter of the same brain region. Despite differences among brain regions, neural tissues were clearly distinguished from non-neural.

Results

Tissue specific methylation differences found in multiple individuals. RLGS profiles were prepared for DNA from 12 tissues collected at autopsy from seven patients. Distinct brain regions were collected including the basal ganglia, the cerebellum, and the frontal and temporal lobes of the cerebrum. Both the cerebellum and the frontal lobe of the cerebrum were separated into grey matter and white matter. Non-brain tissues included kidney, lung, prostate, spleen, stomach and thyroid. All profiles from a single patient were compared against each other, and differences in the two dimensional patterns were noted. Not all tissues were collected from or produced high quality DNA from each patient. **Table 1** shows the tissues for which RLGS gels were analyzed from each patient. A total of 33 RLGS profiles were analyzed in this way. Profiles between the seven patients were not directly compared in order to avoid noting pattern differences due to genetic polymorphism. However, differences noted within a single patient's set of profiles were studied in all 33 RLGS profiles from all tissues in each patient.

The RLGS fragments are named according to the numbering system for our master RLGS profile previously described.[29] A total of 141 RLGS spots showed tissue specific methylation differences within the tissues of at least one patient. Of these, 82 RLGS spots were present, or only partially methylated, in at least one tissue from every patient, therefore ruling out genetic polymorphism as an explanation for RLGS spot loss. It is important to note that we made an exception for patient 3. The spots that were present in at least one tissue in all patients but not present in patient 3 (only 19 such spots), were still considered non-polymorphic for this patient because we only have data from two brain tissues of this patient. These 82 spots, listed in **Supplemental Table 1**, comprise our set of non-polymorphic RLGS spots showing differential methylation among normal tissues.

Table 1. Tissue specific methylation of cloned RLGS spots



Black boxes indicate methylation. [1]RLGS spot could not be analyzed. Abbreviations: B. gang, basal ganglia; C. cort, Cerebellar cortex; C. white, Cerebellar white matter; F. cort, Frontal lobe cortex; F. white, Frontal lobe white matter; T. cort, Temporal lobe cortex; Ecto, ectoderm derived tissue; Endo, endoderm derived tissue; Meso, mesoderm derived tissue.

Of these 82 RLGS spots, we have identified the sequence represented by these spots for 34 by various RLGS spot cloning approaches.[24,28] **Table 1** shows the methylation patterns for each of those 34 RLGS spots in each tissue in each patient, with black boxes indicating methylation. The genomic location and gene homology information for each of the 34 RLGS spots is shown in **Table 2**. **Supplemental Table 1** shows the data for all 82 non-polymorphic RLGS spots (including those whose genomic location has not been identified) that exhibited tissue specific methylation differences.

Table 2. Genomic context of RLGS spots

RLGS	[1]NotI +/- 200 bp	[2]Gene homology	Context	[5]CpG island	GC%	O/E CpG
2E20	chr10:26545049-26545449	GAD2	[3]5' end	Y	62	0.83
3B07	chr7:154975702-154976102	CNPY1	[4]non-5' end	Y	66	0.77
3D46	chr17:69861884-69862284	KIF19	non-5' end	Y	70	0.72
2C66	chr14:55654263-55654663	PELI2	5' end	Y	72	0.99
2D45	chr19:36534086-36534486	TSHZ3	5' end	Y	58	1.06
2C57	chr10:103579943-103580343	KCNIP2	non-5' end	Y	63	0.89
5E22	chr12:50603467-50603867	ACVRL1	non-5' end	N	-	-
3C76	chr12:122322149-122322549	CDK2AP1	5' end	Y	52	1.31
4F36	chr1:44655757-44656157	RNF220	non-5' end	Y	68	0.84
3G78	chr1:47463932-47464332	TAL1	non-5' end	Y	58	1.06
4C09	chr3:9570143-9570543	LHFPL4	5' end	Y	66	1.07
3C64	chr7:1253663-1254063	-	-	Y	70	0.81
1F22	chr9:125817263-125817663	LHX2	5' end	Y	58	0.77
2D25	chr10:135086929-135087329	SPRN	5' end	Y	67	0.81
2E33	chr1:110412400-110412800	ALX3	5' end	Y	65	0.83
4E50	chr15:83326046-83326446	PDE8A	5' end	Y	61	1.12
2F68	chr14:23850460-23850860	LTB4R	5' end	Y	68	0.83
3D24	chr17:55534607-55535007	LOC653653	5' end	N	-	-
2B53	chr22:17658760-17659160	CLTCL1	5' end	Y	61	1.04
4E16	chr4:174326770-174327170	GALNT7	5' end	Y	68	0.99
2D48	chr7:27231338-27231738	-	-	Y	70	0.98
2D34	chr8:81947988-81948388	ZNF704	non-5' end	Y	56	1.23
2D68	chr10:104394156-104394556	TRIM8	5' end	Y	65	0.93
2B54	chr10:131659907-131660307	-	-	Y	57	0.89
2D10	chr5:979471-979871	-	-	Y	68	0.83
2B46	chr15:29406195-29406595	KLF13	5' end	Y	66	1.06
3B30	chr15:40816055-40816455	CDAN1	5' end	Y	69	0.84
3B19	chr2:23461579-23461979	KLHL29	5' end	Y	53	1.11
2C35	chr10:23502265-23502665	LOC729385	5' end	Y	68	0.86
3C70	chr12:122321009-122321409	CDK2AP1	5' end	Y	53	1.31
2E17	chr17:52026137-52026537	NOG	5' end	Y	68	0.85
4B10	chr15:32181122-32181522	PGBD4	5' end	Y	63	0.78
2F15	chr17:75398454-75398854	-	-	Y	60	0.8
3C74	chr5:81082153-81082553	SSBP2	5' end	Y	57	1.04

[1]BLAT coordinates, March 2006 freeze; [2]Annotated gene within 2 kb of the CpG island or NotI site; [3]Within 2 kb of transcriptional start site and/or including exon 1; [4]With the body of the gene, excluding exon 1, and/or within 2 kb of the 3' end; [5]Is the RLGS spot NotI site within 200 bp of a CpG island (classic Gardiner-Garden and Frommer (1987) definition).

Figure 1A shows an example of the RLGS analysis of all 12 tissues, with brain tissues from patient 1 and non-brain tissues from patient 2. The RLGS fragment 3D24 (see arrow, **Fig. 1A**) is present in the cerebellar white matter (C. White) profile, but is clearly absent from the cerebellar cortex (grey matter) (C. Cort). As indicated in **Table 1** and seen as varying intensity of the 3D24 fragment in **Figure 1A**, certain other nervous system tissues were also scored as showing methylation. We estimate our ability to score methylation by RLGS to be constrained by a lower detection limit of 35% to 45% fragment intensity loss, depending on the quality of the DNA.[29] Subsequent cloning of the 3D24 RLGS fragment (NotI-EcoRV fragment) and use of this fragment as a hybridization probe for Southern blot analysis in **Figure 1B** demonstrates partial methylation in all tissues. Quantitation of the methylation was accomplished by comparing the intensity of the methylated band to the combined intensities of the methylated and unmethylated bands to determine the estimated percent methylation for each tissue (**Fig. 1C**). These data show that the expected sensitivity of RLGS analysis fits well with the subsequent measurement of tissue specific methylation by quantitative Southern blot analysis. Moreover, these data demonstrate high level methylation of an RLGS spot only in cerebellum cortex, but

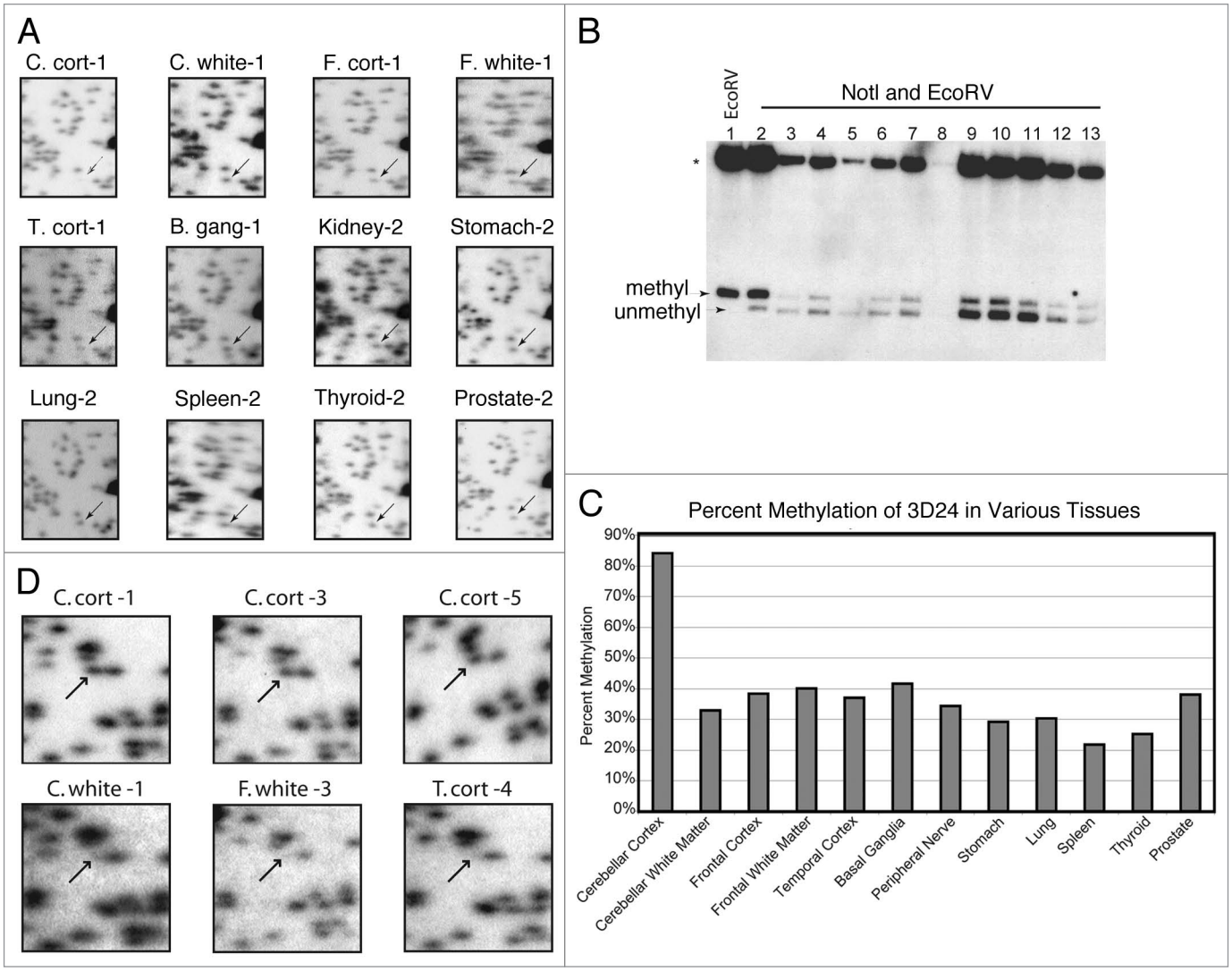


Figure 1. An example of tissue specific methylation identified by RLGS analysis of 12 tissues, and quantitated by Southern blot analysis. (A) Cutouts of the full RLGS analysis on all 12 tissues with arrows indicating fragment 3D24. (B) Quantitative Southern blot analysis using the 3D24 *Not*I-*Eco*RV fragment as hybridization probe. Lane one contains cerebellar cortex DNA digested with *Eco*RV only. Lanes 2–13 are all double digested with *Eco*RV and *Not*I. The DNA's in lanes 2–13 are loaded in the same order as indicated in (A), with peripheral nerve replacing kidney in lane 8. Depending on DNA availability, 0.5 µg (peripheral nerve) to 10 µg (cerebellar cortex) were loaded in each lane. Bands indicating methylation or no methylation are indicated. (C) Quantitation of the percent methylation detected by the Southern blot in (B). For lanes 2–13, percent methylation was determined as described in the *Materials and Methods*. *, 13 kb *Eco*RV sequence without a *Not*I site that is homologous to the probe. BLAST analysis with the probe sequence identified a BAC clone explaining this band.[29] (D) Representative RLGS profilesections showing spot 3B07, and demonstrating consistent spot presence (lack of methylation) in C. cort of multiple patients, but spot absence (methylation) in other brain regions from multiple patients. Abbreviations: B. gang, basal ganglia; C. cort, Cerebellar cortex; C. white, Cerebellar white matter; F. cort, Frontal lobe cortex; F. white, Frontal lobe white matter; T. cort, Temporal lobe cortex.

only low level methylation in other tissues, including the white matter of the cerebellum.

The most dramatic and consistent methylation differences across multiple patients occurred between brain regions and even between grey and white matter of the same region. Although many loci showed methylation specifically in a limited number of tissues, such as 3D24 described above, which was methylated in four out of five patients' cerebellar cortex (**Table 1**), other loci showed a lack of methylation consistently in one tissue. For example, RLGS spot 3B07, which is found at the 3' end of the CNPY1 gene, is specifically unmethylated in each case of cerebellar cortex (n = 5), but methylated in most other brain regions including the white matter of the cerebellum. **Figure 1D** shows representative RLGS gel regions demonstrating the presence of the spot in the cerebellar cortex gels of three patients, and absence of the spot in other brain regions.

We performed hierarchical cluster analysis on the total RLGS data set, shown in **Figure 2**, to look for patterns of methylation. We found that while some tissues tended to cluster together based on methylation patterns, such as C. cort, more striking was that neural tissues strongly clustered distinctly from non-neural tissues. The cluster analysis used a "manhattan" distance (absolute

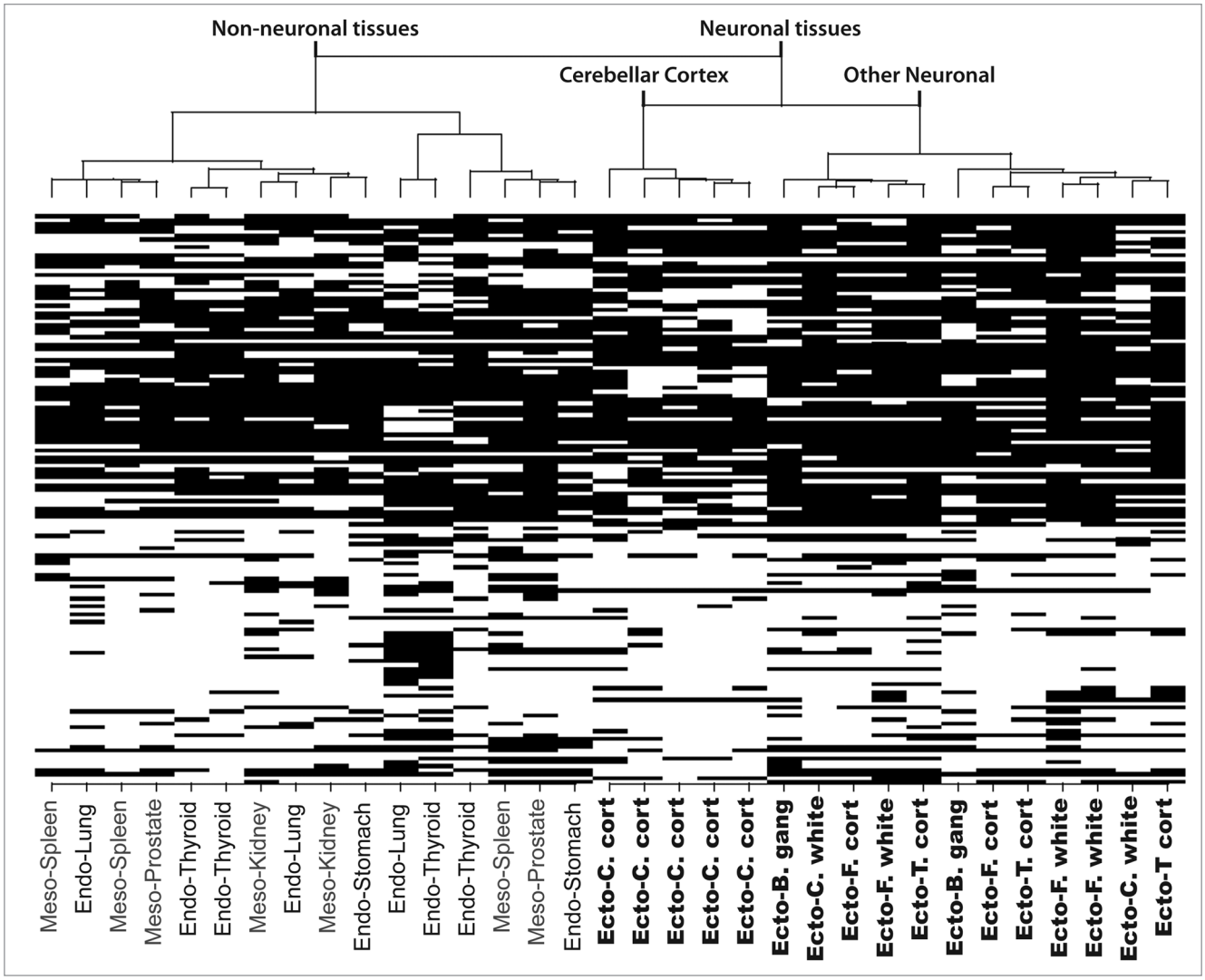


Figure 2. Hierarchical clustering of RLGS data. The heat map and dendogram (top) show the "manhattan" distance (absolute distance) between samples using the "Ward" method for clustering. The RLGS spots are arranged on the Y-axis. Black boxes indicate RLGS spots that were positive for methylation in the indicated tissue. Tissues are arranged on the X-axis. Abbreviations: B. gang, basal ganglia; C. cort, Cerebellar cortex; C. white, Cerebellar white matter; F. cort, Frontal lobe cortex; F. white, Frontal lobe white matter; T. cort, Temporal lobe cortex; Ecto, ectoderm derived tissue; Meso, mesoderm derived tissue; Endo, endoderm derived tissue.

distance) between samples and the "Ward" method for clustering. In short, the "manhattan" distance between two samples is equal to the number of loci that show methylation differences while the "Ward" method for clustering attempts to minimize the sum of squares of any two (hypothetical) clusters than can be formed at each step in the clustering algorithm. From examining the dendogram, it appears there are two major groups, one group consists of the mesoderm and endoderm derived tissues while the other group consists of only the brain tissues, which are ectoderm derived.

Figure 3A shows representative RLGS profiles pointing out spot 3D46 for six ectoderm derived tissues from three patients. The spot is absent or barely detectable in all six profiles. However, this spot is present in mesoderm and endoderm derived tissues including stomach of patients 2 and 9 and spleen of patient 5, lung from patients 2 and 5 and the thyroid of patient 8 (**Fig. 3B**). Given the limited number of mesoderm or endoderm derived tissues in the study and that the only ectoderm derived tissues used are neural, we cannot distinguish whether or not the differences in methylation pattern are related to the germ layer from which the tissues are derived, or more specifically are related to differences between neural and non-neural cell lineages.

MAQMA confirms tissue and neural specific methylation. The RLGS data presented above provide information about the methylation status of the NotI restriction sites and relies upon a subjective determination of spot loss, which can vary in degree due to partial methylation (**Fig. 1**). In order to verify the RLGS results and to obtain more extensive methylation information around the area of the NotI site with quantitative data we used bisulfite sequencing by Mass Array Quantitative Methylation Analysis (MAQMA) on the Sequenome platform.[31,32] MAQMA gives a quantitative value of percent methylation at each

informative CpG dinucleotide. A diagrammatic representation of the quantitative data is shown in **Figure 4A** for RLGS spot 1F22, which generally confirms the RLGS results shown in **Table 1**, with the most methylation seen in the cerebellar cortex and the frontal lobe white matter. In order to get a single value for comparative purposes between tissue types or the neural or non-neural nature of the tissues, the average percent methylation of all the CpG dinucleotides within the sequenced CpG island was calculated and the results were plotted (**Fig. 4B–E**).

These data confirmed tissue specific methylation differences for the spots 1F22, 3B07, 3D24 and 3D46. For all four loci, the cerebellar cortex showed the differential methylation with 3B07 being the only case of less methylation. For all four loci, the consistency of methylation level in all five cerebellar cortex samples is striking. For RLGS spot 1F22 (**Fig. 4B**), the cerebellar cortex has an average of 78% methylation (n = 5), while the white matter from the same region exhibited only 35% methylation (n = 3). RLGS spot 3B07 (**Fig. 4C**), showed hypomethylation in grey matter with an average of 31%, while white matter averaged 52%. In this case, even the cerebellar white matter had slightly less methylation than any of the other tissues, which averaged 62% methylation (n = 24). By quantitative Southern blot analysis (**Fig. 1B and C**) spot 3D24 exhibited a high level of methylation in cerebellar cortex but only a moderate methylation level in all other tissues and this is confirmed again in **Figure 4D**.

Additionally, these data demonstrate neural vs. non-neural methylation specificity for RLGS spots 1F22, 3D24 and 3D46, regardless of whether or not samples of cerebellar cortex were included in the analysis. For example, **Figure 4E** shows the methylation status of 3D46. All the brain tissues have a higher level of methylation (51–78%) than the non-neural tissues (both the endoderm derived tissues (18–48%) and the mesoderm derived tissues (17–38%)). Statistical analysis demonstrated that there is significantly ($p << 0.001$) more methylation in the brain tissues than non-neural tissues (n = 17, and 15, respectively). Furthermore, even when the cerebellar cortex samples were removed from the analysis, the difference was still highly significant ($p << 0.001$). Similar neural vs. non-neural methylation specificity is seen for 1F22, and to a lesser extent 3D24. This was not the case for 3B07, which shows hypomethylation of the CpG island only in cerebellar cortex.

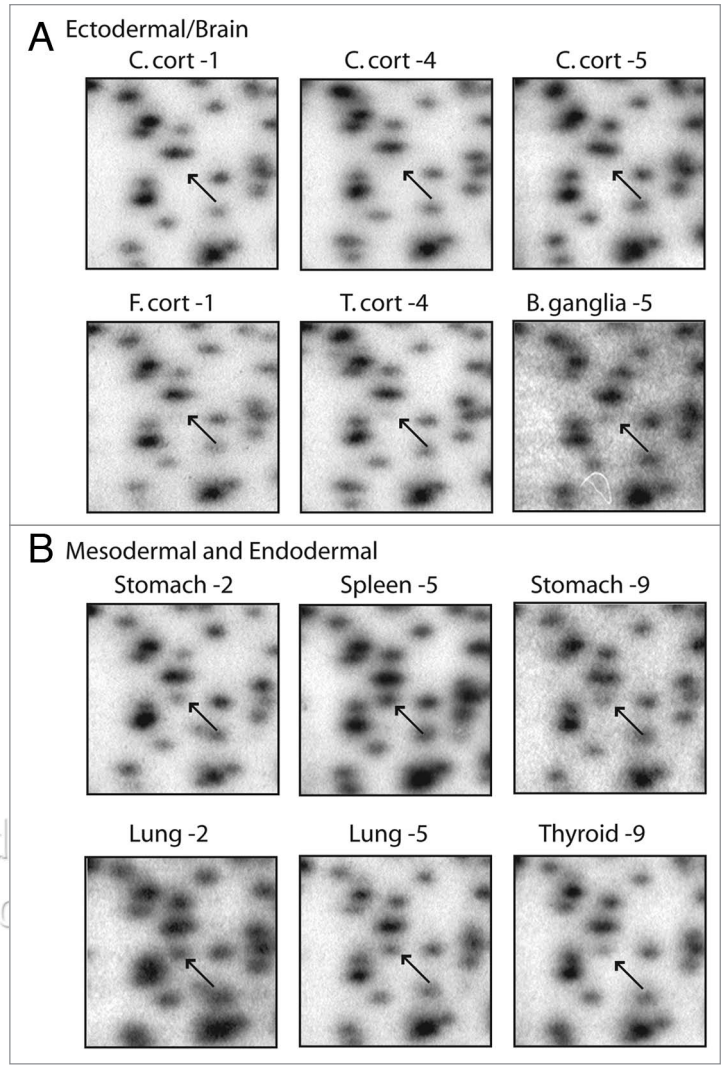


Figure 3. RLGS profiles of spot 3D46 showing neural vs. non-neural tissue specificity. (A) The spot is absent (top) or faintly present (bottom) in ectodermal tissues (C. cortex, T. cortex and B. ganglia). (B) In mesodermal and endodermal tissues (stomach, spleen, lung and thyroid) the spot is clearly present, hence unmethylated. Abbreviations: B. gang, basal ganglia; C. cort, Cerebellar cortex; C. white, Cerebellar white matter; F. white, Frontal lobe white matter; T. cort, Temporal lobe cortex.

Discussion

The data presented in this article provide a global assessment of CpG island methylation differences in normal adult human somatic tissues. Our data demonstrate tissue specific CpG island hypermethylation, confirmed by multiple approaches, for three RLGS loci: 3D24 (pseudogene of AP1S2), 3D46 (KIF19) and 1F22 (LHX2), as well as tissue specific CpG island hypomethylation of 3B07 (CNPY1). We further demonstrate that not only do different regions of the human brain have differential methylation patterns, perhaps related to their specific functions as suggested by the study of Laad-Acosta et al. (2007),[23] but also there are clear differences between grey matter and white matter within the same brain region. This was most clearly demonstrated comparing cerebellar cortex and white matter for RLGS spot 1F22. We propose that methylation differences between the cortex and the white matter reflect differences in the predominant cell types from which the DNA is extracted—the nuclei of neurons in the cortex and of oligodendricytes and astrocytes in the white matter. We also provide evidence suggestive of germ layer specificity in the methylation patterns, with the confirmed examples of the RLGS loci 3D46 (KIF19) and for 1F22 (LHX2). Ectoderm derived (brain) tissues had the highest levels of methylation for these two spots, while the mesodermal tissues and endodermal tissues showed low levels of methylation. A caveat to this

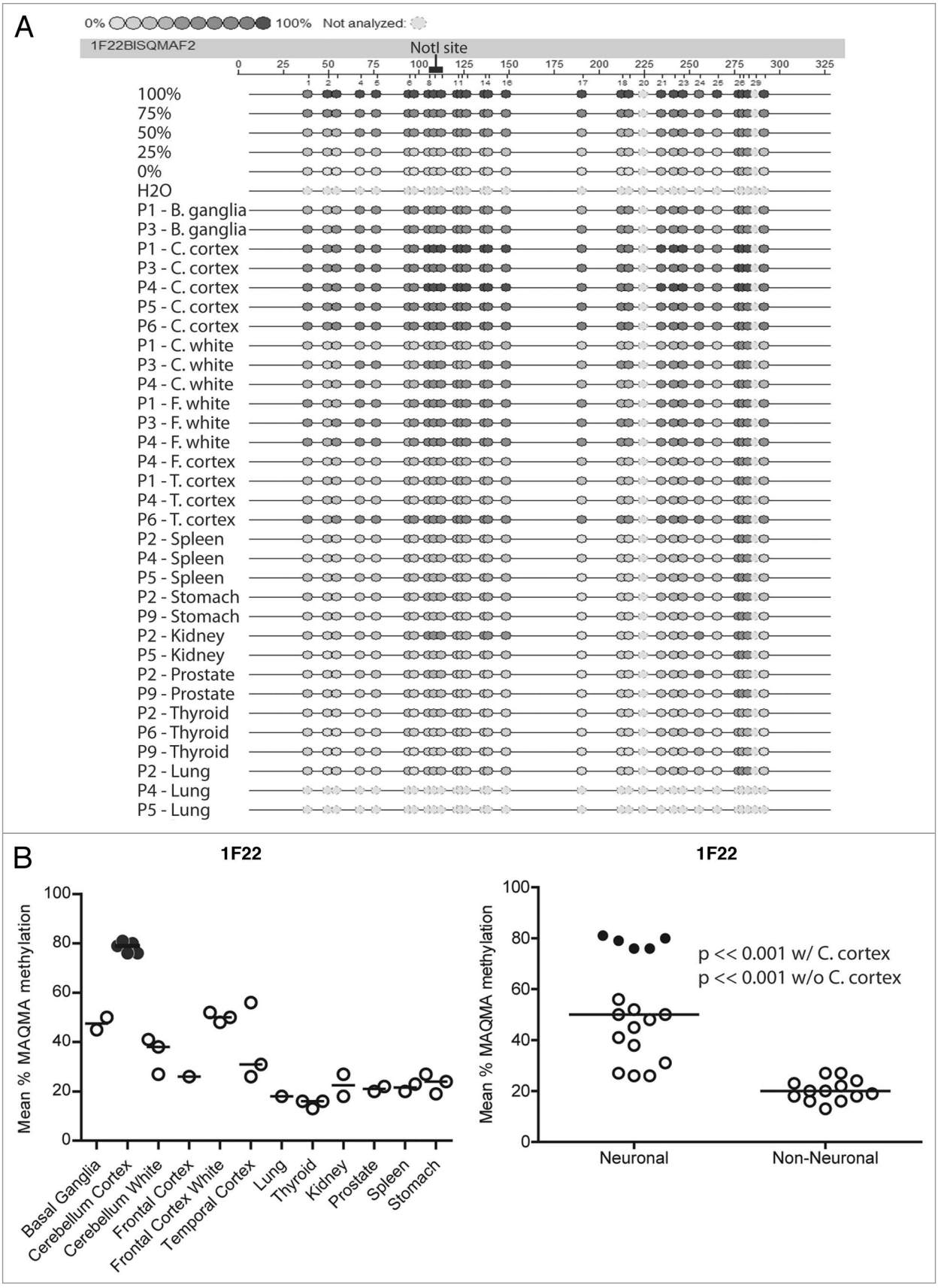


Figure 4. For figure legend, see page 9.

Figure 4 (See opposite page). Confirmation of RLGS data using MAQMA. (A) MAQMA data for the 1F22 CpG island on all normal human tissues. Each line represents a single sample as labeled on the left indicating the patient number and tissue. In the beads-on-a string diagram, each bead represents a CpG dinucleotide. The percent of methylation is indicated by the grey scale shading as shown by the key at the top of the figure with dark grey representing 100% methylation. The two lung samples at the bottom failed to PCR amplify. To test for bisulfite PCR bias we used peripheral blood lymphocyte (PBL) DNA as a 0% methylation control, and an in vitro methylated (IVM) aliquot of the same DNA as a 100% methylated control. These DNAs were mixed at appropriate ratios to generate 75%, 50% and 25% methylated controls and all the controls were bisulfite treated and used as template for bisulfite PCR and MAQMA analysis and shown at the top of the figure. (B–E) The average level of methylation detected by MAQMA across the sequenced fragment of the CpG islands for each samples is shown on the Y axes. This value comes from taking the average of MAQMA values for each CpG dinucleotide sequenced for each sample for each of the four RLGS spots indicated at the top of each figure. The samples are divided up categorically along the X axes by tissue type on the left, and by neural vs. non-neural tissue on the right. Each symbol represents a single sample. The black, filled in circles indicate the Cerebellar cortex samples. p-values were determined using the Wilcoxon Rank Sum test[39] to determine if there were differences in mean methylation levels between neural and non-neural. Values are given for the Wilcoxon Rank Sum test performed both with and without the Cerebellar cortex samples.

interpretation, however, is that we have only neural tissues from ectoderm to compare. Therefore, it is possible that the differences in methylation we demonstrate have more to do with neural vs. non-neural differentiation than the germ layer from which the tissues are derived. Similar conclusions about the relationship between NotI site methylation patterns and developmental similarity of the cell types have been postulated by Sakamoto et al. (2007)[34] in mouse tissues.

A finding of particular interest is the high percentage of methylation specifically in the cerebellar cortex for the gene LHX2 (LIM/homeobox protein LH2) compared to other regions of the brain. The gene LHX2 acts as a transcriptional regulator and has been recently reported to play the role of a classic selector gene in determining cortical identity during brain development in mouse.[35] We found that the LHX2 gene (RLGS spot 1F22) remains highly methylated in the cerebellar cortex (range 76–81%; n = 5) compared to other regions in adult human brain (range 26–56%; n = 12). As shown in **Supplemental Figure 1**, the Allen Brain Atlas Project found very low expression of Lhx2 in mouse cerebellum by in situ hybridization, but high expression in some forebrain structures that include the frontal and temporal lobes (used in our study). This fits well with our observations that LHX2 is highly methylated only in the cerebellum. In addition, LHX2 is believed to play a role in mesoderm (mesenchyme) development in mouse.[36] In our study, the Lhx2 gene shows minimal methylation in mesodermal tissues like kidney, prostate and spleen.

Due to lack of availability of RNA from the autopsy samples we could not directly study gene expression in our study. However, from the GNF expression database we could correlate methylation and expression patterns of some of the genes. In agreement with our methylation findings and the in situ findings from the Allen Brain Atlas Project, the GNF expression database indicates that the Lhx2 gene is highly expressed in most brain regions in both mouse and human except the cerebellum where its expression is comparatively low (**Suppl. Fig. 2**). Interestingly, this gene is not expressed in most non-brain tissues including the ones used in our study, yet we did not detect DNA methylation in these tissues by either RLGS analysis or bisulfite sequencing. This suggests that DNA methylation is not required for silencing of the gene in some tissues. These observations suggest that DNA methylation may only be required to silence this gene in neural cells, while non-neural cells do not have the ability to express LHX2 regardless of the methylation status.

The CNPY1 gene shows exactly the opposite pattern to LHX2. While LHX2 is important in forebrain development, CNPY1 is important in hindbrain development. The Canopy 1 gene was found to be expressed in the midbrain—hindbrain boundary in zebra fish and plays a major role in the development of the midbrain tectum and the cerebellum.[37] The Allen Brain Atlas Project found that mouse Cnpy1 is highly expressed in the cerebellum, but not in most other brain regions (**Suppl. Fig. 1**). The GNF expression database shows a similar result with a high level of expression only seen in the cerebellum (**Suppl. Fig. 2**). Our data show that this gene (RLGS spot 3B07) is highly methylated in all tissues except cerebellum (**Fig. 4**), and is thus consistent with its methylation playing a role in silencing of the gene in non-cerebellar tissues.

Taken together, these data indicate that a limited number of CpG islands display tissue specific methylation and demonstrate a remarkable amount of specificity not only comparing one brain region to another, but also among different regions in the same neuroanatomical structure. A more comprehensive microarray or high-throughput sequencing based approach would certainly identify additional such CpG islands, but the proportion would be expected to be the same. We suggest that these tissue specific methylation patterns are a critical aspect of the regulatory mechanisms of tissue-specific gene expression during different phases of development. Furthermore, the fact that different regions of the brain have differential patterns of methylation supports the notion that DNA methylation is a major determinant of the functional specializations of specific brain regions, and perhaps even its cellular composition. Additionally, our finding of neural specific methylation patterns suggests that methylation of some CpG islands may be determined early in development. This may occur in association with some of the epigenetic differences that have been demonstrated at loci that bind polycomb repressive complex 2 (PRC2) as stem cells differentiate along different lineages.[38]

Materials and Methods

Normal tissue samples. Normal tissues were obtained from seven patients at autopsy performed within four to eight hours of death. Tissues were immediately frozen in liquid nitrogen and stored at -70°C. Tissues were obtained from the following patient donors: Patient 1 was a 54-year-old Caucasian female; Patient 2 was an 80-year-old Caucasian male; Patient 3 was a 38-year-old Caucasian male; Patient 4 was a 69-year-old Caucasian female;

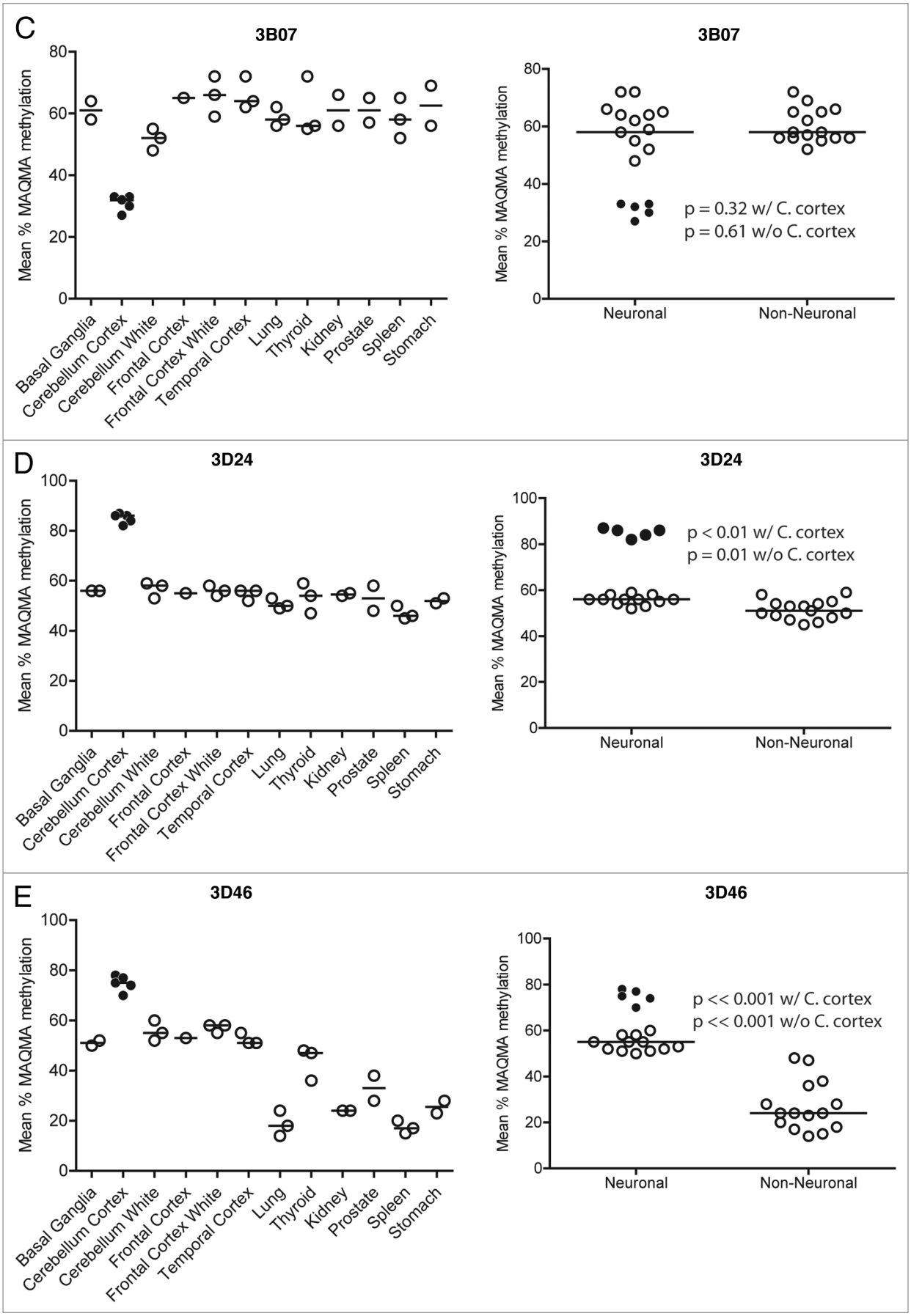

C
3B07
Mean % MAQMA methylation
80
60
40
20
0
Basal Ganglia
Cerebellum Cortex
Cerebellum White
Frontal Cortex
Frontal Cortex White
Temporal Cortex
Lung
Thyroid
Kidney
Prostate
Spleen
Stomach
3B07
Mean % MAQMA methylation
p = 0.32 w/ C. cortex
p = 0.61 w/o C. cortex
Neuronal
Non-Neuronal
D
3D24
Mean % MAQMA methylation
100
80
60
40
20
0
Basal Ganglia
Cerebellum Cortex
Cerebellum White
Frontal Cortex
Frontal Cortex White
Temporal Cortex
Lung
Thyroid
Kidney
Prostate
Spleen
Stomach
3D24
Mean % MAQMA methylation
p < 0.01 w/ C. cortex
p = 0.01 w/o C. cortex
Neuronal
Non-Neuronal
E
3D46
Mean % MAQMA methylation
100
80
60
40
20
0
Basal Ganglia
Cerebellum Cortex
Cerebellum White
Frontal Cortex
Frontal Cortex White
Temporal Cortex
Lung
Thyroid
Kidney
Prostate
Spleen
Stomach
3D46
Mean % MAQMA methylation
p << 0.001 w/ C. cortex
p << 0.001 w/o C. cortex
Neuronal
Non-Neuronal

Figure 4. Confirmation of RLGS data using MAQMA. (B–E) The average level of methylation detected by MAQMA across the sequenced fragment of the CpG islands for each samples is shown on the Y axes. This value comes from taking the average of MAQMA values for each CpG dinucleotide sequenced for each sample for each of the four RLGS spots indicated at the top of each figure. The samples are divided up categorically along the X axes by tissue type on the left, and by neural vs. non-neural tissue on the right. Each symbol represents a single sample. The black, filled in circles indicate the Cerebellar cortex samples. p-values were determined using the Wilcoxon Rank Sum test[39] to determine if there were differences in mean methylation levels between neural and non-neural. Values are given for the Wilcoxon Rank Sum test performed both with and without the Cerebellar cortex samples.

Patient 5 was a 60-year-old male; Patient 6 was a 58 year old male; Patient 9 was a 40 year old male. For all patients there was no evidence of neoplastic or autoimmune disease at the time of death. Tissues from patients 1–4 were collected with approval of the institutional review board of The Ohio State University, which determined that the specimens used in this research were from non-living persons and therefore, according to the United States Code of Federal Regulations § 46.102(f), does not constitute human subjects research. Tissues from patients 5, 6 and 9 were collected with approval by the ethics committee of the board of physicians at the Westfalian Wilhelms University in Münster Germany (IRB # 2007-254-f-S).

Restriction landmark genomic scanning and spot cloning. The protocol for extraction of genomic DNA from tissues was described previously.[24] The published protocol of Dai et al. (2003)[25] was followed for RLGS gels. RLGS spots of interest were cloned as previously described.[24,26-28] RLGS gel analysis was performed as previously described, with visual inspection of profiles.[29,30] We have previously demonstrated, and confirm in this study, that the this approach accurately identifies DNA methylation as loss of spot intensity relative to surrounding spots on the same gel when DNA methylation is approximately 40% or greater at the NotI site.[29,30]

Southern blot analysis. Genomic DNAs were digested with the appropriate restriction endonucleases for 4 hours and subsequently Southern blot analysis was carried out as previously described.[24] Southern blots were quantitated using a PhosphorImager and the Imagequant software (Molecular Dynamics, Inc.,) to quantitate individual bands. In each lane, the total pixel intensity of an equal area was quantitated in the position of the both the methylated band and the unmethylated band. To represent percentage of methylation, the value for the methylated band was divided by the sum of the methylated and unmethylated band, and multiplied by 100. Since all calculation and measurements are done within each individual lane, unequal loading of the lanes is controlled for.

Sodium bisulfite treatment. Sodium bisulfite treatment to convert unmethylated cytosine to thymidine was completed following the manufacturer's protocol EZ-96 DNA Methylation Kit (Zymo Research, Orange, CA) using 750 ng in 50 µl of distilled water and M-Dilution Buffer. The treated samples were resuspended in 75 µl of M-Elution Buffer and stored at -20°C.

MassARRAY quantitative methylation analysis (MAQMA). MassArray Quantitative Methylation Analysis (MAQMA) was performed using the MassARRAY Compact system developed by the Sequenome Company, as previously described.[31] Primer sequences are available upon request. This system utilizes mass spectrometry (MS) for the detection and quantitative analysis of DNA methylation. This approach has been shown to a be highly accurate and reproducible way to quantitate methylation.[32]

Statistics. For each CpG site locus, a Wilcoxon Rank Sum test was used to compare the mean percentage of methylation as measured by MAQMA across tissue types. Significance was assessed by univariate p-values less than 0.05 (**Fig. 4**). The heat map and dendogram shown in **Figure 2** were produced using the "image" function in R with a "manhattan" distance (absolute distance) between samples and the "Ward" method for clustering.[33] In short, the "manhattan" distance between two samples is equal to the number of loci that show methylation differences while the "Ward" method for clustering attempts to minimize the sum of squares of any two (hypothetical) clusters than can be formed at each step in the clustering algorithm. The R computing language was used to create **Figure 2** and perform all statistical testing.[33]

Acknowledgements

This work was supported in part by the National Cancer Institute through the Roswell Park Cancer Center Support Grant, [CA 16056] (D.J.S., S.G., J.C.M.).

Note

Supplementary materials can be found at:
www.landesbioscience.com/supplement/GhoshEPI5-6-Sup.pdf

References

1. Jaenisch R. DNA methylation and imprinting: why bother? Trends Genet 1997; 13:323-9.
2. Li E, Bestor TH, Jaenisch R. Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. Cell 1992; 69:915-26.
3. Panning B, Jaenisch R. RNA and the epigenetic regulation of X chromosome inactivation. Cell 1998; 93:305-8.
4. Li E, Beard C, Jaenisch R. Role for DNA methylation in genomic imprinting. Nature 1993; 366:362-5.
5. Brandeis M, Ariel M, Cedar H. Dynamics of DNA methylation during development. Bioessays 1993; 15:709-13.
6. Eden S, Cedar H. Role of DNA methylation in the regulation of transcription. Curr Opin Genet Dev 1994; 4:255-9.
7. Bird AP. CpG-rich islands and the function of DNA methylation. Nature 1986; 321:209-13.
8. Ng HH, Bird A. DNA methylation and chromatin modification. Curr Opin Genet Dev 1999; 9:158-63.
9. Turker MS, Bestor TH. Formation of methylation patterns in the mammalian genome. Mutat Res 1997; 386:119-30.
10. Gardiner-Garden M, Frommer M. CpG islands in vertebrate genomes. J Mol Biol 1987; 196:261-82.
11. Walsh CP, Bestor TH. Cytosine methylation and mammalian development. Genes Dev 1999; 13:26-34.
12. Ahuja N, Li Q, Mohan AL, Baylin SB, Issa JP. Aging and DNA methylation in colorectal mucosa and cancer. Cancer Res 1998; 58:5489-94.
13. Eckhardt F, Lewin J, Cortese R, Rakyan VK, Attwood J, Burger M, et al. DNA methylation profiling of human chromosomes 6, 20 and 22. Nat Genet 2006; 38:1378-85.

14. Weber M, Hellmann I, Stadler MB, Ramos L, Paabo S, Rebhan M, et al. Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. Nat Genet 2007; 39:457-66.
15. Song F, Smith JF, Kimura MT, Morrow AD, Matsuyama T, Nagase H, et al. Association of tissue-specific differentially methylated regions (TDMs) with differential gene expression. Proc Natl Acad Sci USA 2005; 102:3336-41.
16. Kitamura E, Igarashi J, Morohashi A, Hida N, Oinuma T, Nemoto N, et al. Analysis of tissue-specific differentially methylated regions (TDMs) in humans. Genomics 2007; 89:326-37.
17. Illingworth R, Kerr A, Desousa D, Jorgensen H, Ellis P, Stalker J, et al. A novel CpG island set identifies tissue-specific methylation at developmental gene loci. PLoS Biol 2008; 6:22.
18. Rakyan VK, Down TA, Thorne NP, Flicek P, Kulesha E, Graf S, et al. An integrated resource for genome-wide identification and analysis of human tissue-specific differentially methylated regions (tDMRs). Genome Res 2008; 18:1518-29.
19. Schilling E, Rehli M. Global, comparative analysis of tissue-specific promoter CpG methylation. Genomics 2007; 90:314-23.
20. Shen L, Kondo Y, Guo Y, Zhang J, Zhang L, Ahmed S, et al. Genome-wide profiling of DNA methylation reveals a class of normally methylated CpG island promoters. PLoS Genet 2007; 3:2023-36.
21. Byun HM, Siegmund KD, Pan F, Weisenberger DJ, Kanel G, Laird PW, et al. Epigenetic profiling of somatic tissues from human autopsy specimens identifies tissue- and individual-specific DNA methylation patterns. Hum Mol Genet 2009; 18:4808-17.
22. Christensen BC, Houseman EA, Marsit CJ, Zheng S, Wrensch MR, Wiemels JL, et al. Aging and environmental exposures alter tissue-specific DNA methylation dependent upon CpG island context. PLoS Genet 2009; 5:1000602.
23. Ladd-Acosta C, Pevsner J, Sabunciyan S, Yolken RH, Webster MJ, Dinkins T, et al. DNA methylation signatures within the human brain. Am J Hum Genet 2007; 81:1304-15.
24. Smiraglia DJ, Fruhwald MC, Costello JF, McCormick SP, Dai Z, Peltomaki P, et al. A new tool for the rapid cloning of amplified and hypermethylated human DNA sequences from restriction landmark genome scanning gels. Genomics 1999; 58:254-62.
25. Dai Z, Zhu WG, Morrison CD, Brena RM, Smiraglia DJ, Raval A, et al. A comprehensive search for DNA amplification in lung cancer identifies inhibitors of apoptosis cIAP1 and cIAP2 as candidate oncogenes. Hum Mol Genet 2003; 12:791-801.
26. Yu L, Liu C, Bennett K, Wu YZ, Dai Z, Vandeusen J, et al. A NotI-EcoRV promoter library for studies of genetic and epigenetic alterations in mouse models of human malignancies. Genomics 2004; 84:647-60.
27. Zardo G, Tiirikainen MI, Hong C, Misra A, Feuerstein BG, Volik S, et al. Integrated genomic and epigenomic analyses pinpoint biallelic gene inactivation in tumors. Nat Genet 2002; 32:453-8.
28. Smiraglia DJ, Kazhiyur-Mannar R, Oakes CC, Wu YZ, Liang P, Ansari T, et al. Restriction Landmark Genomic Scanning (RLGS) spot identification by second generation virtual RLGS in multiple genomes with multiple enzyme combinations. BMC Genomics 2007; 8:446.
29. Costello JF, Fruhwald MC, Smiraglia DJ, Rush LJ, Robertson GP, Gao X, et al. Aberrant CpG-island methylation has non-random and tumour-type-specific patterns. Nat Genet 2000; 24:132-8.
30. Smiraglia DJ, Rush LJ, Fruhwald MC, Dai Z, Held WA, Costello JF, et al. Excessive CpG island hypermethylation in cancer cell lines versus primary human malignancies. Hum Mol Genet 2001; 10:1413-9.
31. Ehrich M, Nelson MR, Stanssens P, Zabeau M, Liloglou T, Xinarianos G, et al. Quantitative high-throughput analysis of DNA methylation patterns by base-specific cleavage and mass spectrometry. Proc Natl Acad Sci USA 2005; 102:15785-90.
32. Coolen MW, Statham AL, Gardiner-Garden M, Clark SJ. Genomic profiling of CpG methylation and allelic specificity using quantitative high-throughput mass spectrometry: critical evaluation and improvements. Nucleic Acids Res 2007; 35:119.
33. Computing TRFfS. The R Statistical Computing Language. www.r-project.org 2005.
34. Sakamoto H, Suzuki M, Abe T, Hosoyama T, Himeno E, Tanaka S, et al. Cell type-specific methylation profiles occurring disproportionately in CpG-less regions that delineate developmental similarity. Genes Cells 2007; 12:1123-32.
35. Mangale VS, Hirokawa KE, Satyaki PR, Gokulchandran N, Chikbire S, Subramanian L, et al. Lhx2 selector activity specifies cortical identity and suppresses hippocampal organizer fate. Science 2008; 319:304-9.
36. Kolterud A, Wandzioch E, Carlsson L. Lhx2 is expressed in the septum transversum mesenchyme that becomes an integral part of the liver and the formation of these cells is independent of functional Lhx2. Gene Expr Patterns 2004; 4:521-8.
37. Hirate Y, Okamoto H. Canopy1, a Novel Regulator of FGF Signaling around the Midbrain-Hindbrain Boundary in Zebrafish. Current Biology 2006; 16:421-7.
38. Lee TI, Jenner RG, Boyer LA, Guenther MG, Levine SS, Kumar RM, et al. Control of developmental regulators by Polycomb in human embryonic stem cells. Cell 2006; 125:301-13.
39. Welch BL. On the comparison of several mean values: an alternative approach. Biometrika 1951; 38:330-6.