MCB
Journals.ASM.org

Regionally Specific and Genome-Wide Analyses Conclusively Demonstrate the Absence of CpG Methylation in Human Mitochondrial DNA

Elizabeth E. Hong,[a] Cindy Y. Okitsu,[b] Andrew D. Smith,[a] Chih-Lin Hsieh[b]

Molecular and Computational Biology Section, Department of Biological Sciences,[a] and Department of Urology and Department of Biochemistry and Molecular Biology,[b] University of Southern California, Los Angeles, California, USA

Although CpG methylation clearly distributes genome-wide in vertebrate nuclear DNA, the state of methylation in the vertebrate mitochondrial genome has been unclear. Several recent reports using immunoprecipitation, mass spectrometry, and enzyme-linked immunosorbent assay methods concluded that human mitochondrial DNA (mtDNA) has much more than the 2 to 5% CpG methylation previously estimated. However, these methods do not provide information as to the sites or frequency of methylation at each CpG site. Here, we have used the more definitive bisulfite genomic sequencing method to examine CpG methylation in HCT116 human cells and primary human cells to independently answer these two questions. We found no evidence of CpG methylation at a biologically significant level in these regions of the human mitochondrial genome. Furthermore, unbiased next-generation sequencing of sodium bisulfite treated total DNA from HCT116 cells and analysis of genome-wide sodium bisulfite sequencing data sets from several other DNA sources confirmed this absence of CpG methylation in mtDNA. Based on our findings using regionally specific and genome-wide approaches with multiple human cell sources, we can definitively conclude that CpG methylation is absent in mtDNA. It is highly unlikely that CpG methylation plays any role in direct control of mitochondrial function.

CpG methylation occurs with a genome-wide distribution in vertebrates and has impacts on transcription, repeat element biology, and recombination. Although the presence of CpG methylation in the mitochondrial genome was reported using restriction digestion (1, 2) and radiolabeling methods (3) a few decades ago, two of our unpublished studies showed no CpG methylation in HEK293 cells by the sodium bisulfite sequencing method. The first of our unpublished studies nearly 2 decades ago did not detect any DNA methylation at CpG sites in the 12S and the 16S regions of mitochondrial DNA (mtDNA) from a limited number of molecules sequenced after sodium bisufite treatment of DNA. Our more recent study of three mtDNA regions (nucleotides 560 to 893, 4250 to 4569, and 16381 to 16470) confirmed an absence of CpG methylation in a total of 1,487 CpG sites from 203 molecules examined from HEK293 cells using the sodium bisulfite genomic sequencing method (see Table S1 in the supplemental material). A limited study using bisulfite-PCR/single-stranded DNA conformation polymorphism analysis failed to detect any cytosine methylation in mtDNA (4). Our interest in mtDNA methylation was then renewed when detection of 5-methylcytosine (5mC) and 5-hydroxymethylcytosine (5hmC) at CpG dinucleotides in mtDNA by immunoprecipitation (IP) (5), mass spectrometry (6), and enzyme-linked immunosorbent assay (ELISA) (7) were recently reported. However, these methods do not provide critical information about the sites and the frequency of methylation at each site that is required to fully assess any putative functional impact of DNA methylation on mtDNA.

Shock et al. (5) concluded that the level of 5mC and 5hmC modifications is much higher in human cells than the 3 to 5% previously reported in mouse cells by one group (2). Nass (3) estimated 5mC frequency at CpG sites as between 2.8 and 8.4% in mouse and hamster cells, and Shmookler Reis and Goldstein (1) concluded that 2 to 5% of the mtDNA is methylated at all CCGG sites in human cells. Although methylation frequency in mtDNA was consistently estimated in the 2 to 8% range in all of the reports prior to Shock et al., the pattern of DNA methylation was not clearly known and quite controversial. The restriction digestion method employed by some is limited by an unavoidable small percentage of incompletely digested or indigestible DNA molecules. The IP method using anti-5mC antibody captures the methylated DNA targets and provides a percentage pulldown of the input DNA in a quantitative PCR (qPCR) assay. This method cannot be used to determine the target frequency in a given sample due to the difficulty in determining the efficiency of antibodies in binding specific targets, even though it is widely used to compare target frequencies between different samples. Infantino et al. (6) reported a very high frequency of DNA methylation using mass spectrometry analysis but could not assign methylation to specific sites. A more recent study by Dzitoyeva et al. (7) also detected 5mC and 5hmC in an ELISA; however, this approach does not provide the frequency of methylation because the assay only normalizes against a control standard of unknown sequence in the commercial kit for quantification. Most of these assays used to document cytosine methylation in mtDNA do not provide any information

Received 25 February 2013 Returned for modification 11 April 2013
Accepted 3 May 2013
Published ahead of print 13 May 2013
Address correspondence to Chih-Lin Hsieh, hsieh_c@ccnt.hsc.usc.edu.
Supplemental material for this article may be found at http://dx.doi.org/10.1128/MCB.00220-13.

doi:10.1128/MCB.00220-13

on the sites or frequency of CpG methylation of the targets. Without such information, the impact of DNA methylation on mitochondrial function cannot be discerned.

To determine whether CpG methylation plays a role in mitochondrial biology, the location and frequency of the methylated sites must be clearly determined. Also, whether there is any non-CpG methylation in mtDNA needs to be investigated, as is potentially indicated by the results of Infantino et al. (6). Such information can be obtained using the more definitive method of sodium bisulfite genomic sequencing, which allows analysis of each molecule for methylation at each cytosine (8). The sodium bisulfite method generally has a 0.5% nonconversion rate as the assay background and cannot convert either 5mC or 5hmC (9). The claimed levels of DNA methylation in mitochondria are much higher than the assay background; therefore, the sodium bisulfite method should easily identify the sites of methylation or hydroxymethylation at CpG and non-CpG sites and provide the needed information for mtDNA methylation. Four mtDNA regions—12S ribosome (12S), 16S ribosome (16S), cytochrome *c* oxidase subunit II (COII), and ATP synthase F0 subunit 6 (ATP6)—in human HCT116 cells reported in Shock et al. (5) are the only studies of specific regions of methylation. These four regions were examined in the present study using the sodium bisulfite sequencing method to assess cytosine methylation sites and to estimate the frequency of CpG methylation in mtDNA. To examine global DNA methylation status, as well as sites of methylation in mtDNA, 10 published data sets from genome-wide sodium bisulfite sequencing of DNA from many other cell sources, including HCT116 cells, were analyzed for mtDNA methylation (10–13). We further confirmed the findings from the published data set generated from reduced representation bisulfite sequencing of HCT116 cells with sodium bisulfite sequencing of total DNA from HCT116 using next-generation sequencing to obtain unbiased information on mtDNA methylation. The study here is the most comprehensive and detailed analysis for cytosine modification of human mtDNA, and it concludes the lack of DNA methylation in CpG and non-CpG sites in human mtDNA unequivocally. Based on this finding, the conclusion that DNA methylation does not directly involve the regulation of mitochondrial biology can be clearly drawn.

MATERIALS AND METHODS

DNA sources. HTC116 cells (a gift from Bert Vogelstein, Johns Hopkins University, Baltimore, MD) were maintained in McCoy 5A medium with 10% fetal bovine serum. DNA was extracted from HEK293 cells and HTC116 cells by using proteinase K digestion, phenol-chloroform extraction, and isopropanol precipitation. Total DNA from primary cells without cell culture was isolated from whole blood obtained from a normal male subject by guanidinium thiocyanate extraction method after lysing the red blood cells.

Bisulfite genomic sequencing. Bisulfite genomic sequencing was carried out as described previously (14). DNA was digested with restriction enzyme HindIII and purified by phenol-chloroform extraction and ethanol precipitation. One microgram of HindIII-digested DNA was then treated with sodium bisulfite overnight at 55°C in 30 μl of 1% agarose after denaturing with NaOH. The agarose containing the treated DNA was washed with water, alkaline treated, washed with water again, and equilibrated with Tris-EDTA (TE). The agarose was melted after all treatments at 70°C, and an equal volume of TE was added. PCR was carried out with 0.8 μl of the treated DNA and the appropriate primers (Table 1). The PCR products were ligated into the TOPO TA cloning vector (Invitrogen). Multiple clones from each ligation were sequenced. Both strands of each clone were sequenced using an Excel II sequencing kit (Epicentre) and analyzed on a Li-Cor sequencer (model 4200).

TABLE 1 PCR primers and regions of amplification for 12S, 16S, COII, and ATP6 regions of mtDNA

Primer	Sequence (5′–3′)	Nucleotide[a]	
		Start	End
MT12StopFor	GATTAAAAGGAATAAGTATTAAGTA	740	764
MT12StopRev	TCTAATCCCAATTTAAATCTTAAC	1083	1060
MT16StopFor	ATATGTTAAGATTTTATTAGTTAAAG	2851	2876
MT16StopRev	AAACCTTATTTCTCTTATCCTTTC	3150	3127
MTCOX2topFor	ATTGGTTATTAATGGTATTGAATTTA	7883	7908
MTCOX2topRev	CTCCACAAATTTCAAAACATTAAC	8189	8166
MTATP6topFor	TTTATTTATTTAATTAATAGTTTTGGT	8970	8996
MTATP6topRev	CCCCTATTAAAAATCATAAACTAA	9162	9138

[a] Based on the nucleotide positions of the human mtDNA map at www.mitomap.org.

Next-generation sequencing and analysis of published data sets. Sodium bisulfite-treated total DNA from HCT116 cells was sequenced using the Illumina next-generation sequencing platform by BGI Tech (Hong Kong). In brief, adaptors were ligated to the sonicated total DNA before sodium bisulfite treatment (Zymo EZ DNA methylation-gold kit). The treated DNA was amplified, size selected, and pair end sequenced. DNA from enterobacteria phage lambda was spiked into HCT116 DNA as a control for the effectiveness of the sodium bisulfite treatment. A total of 13,333,334 sequencing reads (paired end, 90 bases per end, 2.4e09 bases total) were mapped as previously described (15–17) to the hg19 genome assembly downloaded from the FTP site associated with the UCSC Genome Browser. The unconverted Cs at non-CpG sites in the mtDNA reads and the nuclear genome reads were counted separately to obtain the nonconversion rate of the sodium bisulfite treatment in each of the two genomes. The reads mapped to the phage lambda, which does not have DNA methylation, were analyzed separately to determine the background nonconversion rate. The unconverted Cs at the CpG sites were counted to derive the frequency of DNA methylation both in the reads of mtDNA and the nuclear sequences. Ten published data sets were downloaded from the NCBI Sequence Read Archive (18) and mapped to either hg18 or hg19 genome assembly. The nonconversion rate of the nuclear and mitochondrial genome and the methylation frequency at CpG sites of mtDNA were determined as described above.

RESULTS

Lack of appreciable CpG methylation in mtDNA from HCT116 cells in specific regions. It is possible that 5mC is present in mtDNA from HTC116 but not in HEK293 cells or that 5mC is present in the regions of mtDNA analyzed in the study by Shock et al. (5) but not in the regions that we analyzed, as previously described. To minimize any uncertainty, we examined cytosine methylation in the same mtDNA regions of the same cell line as Shock et al. (5). Unlike an unmodified cytosine that gives rise to a T, 5mC and 5hmC do not react with sodium bisulfite and remain as Cs in the PCR following the chemical treatment of DNA. Based on the estimation of 2 to 8% methylation at CpG sites in mtDNA by the previous reports, sodium bisulfite sequencing of mtDNA should reveal both the pattern and the frequency of cytosine modification in mtDNA.

DNA extracted from HCT116 cells was treated with sodium bisulfite to convert unmodified cytosines and to identify sites of 5mC and 5hmC. Four regions—12S, 16S, COII, and ATP6—of the mtDNA, where 5mC and 5hmC were reported to be present, were amplified and sequenced after sodium bisulfite treatment. The primers used to amplify these DNA regions do not include any CpG sites to ensure that the PCRs do not bias against or for

TABLE 2 Summary of unconverted Cs observed in four regions of mtDNA from HCT116 cells after sodium bisulfite treatment[a]

Locus (size [bp])	No. of Cs	No. of CpG sites	N	Uncoverted C at CpG sites			Mutation			Uncoverted C at non-CpG sites		
				No. of sites	No. of mol	%[b]	No. of sites	No. of mol	F×10^{-4}[c]	No. of sites	No. of mol	%[d]
12S (292)	79	11	16	0	0		3	3		0	0	
			14	0	0		3	3		2	2	
			9	0	0		0	0		1	1	
			15	0	0		4	4		0	0	
Total	4,266	594	54	0	0	<0.17	10	10	6.3	3	3	0.08
16S (250)	61	13	14	1	1		6	5		1	1	
			15	0	0		7	6		0	0	
			13	0	0		2	1		1	1	
Total	2,562	546	42	1	1	0.18	15	12	14.3	2	2	0.10
COII (258)	81	17	13	0	0		1	1		1	1	
			5	0	0		0	0		0	0	
			16	0	0		4	4		2	2	
Total	2,754	578	34	0	0	<0.17	5	5	5.7	3	3	0.14
ATP6 (247)	82	8	15	2	2		3	3		1	1	
			15	0	0		4	4		3	3	
			15	0	0		0	0		2	2	
			15	1	1		7	2		2	2	
Total	4,920	480	60	3	3	0.63	14	9	9.4	8	8	0.18

[a] Each line represents an independent PCR and cloning experiment. Cs, cytosines at non-CpG sites; N, total number of molecules examined; mol, molecules.
[b] That is, the percent frequency of unconverted cytosine at CpG sites (i.e., the number of unconverted cytosines divided by the total number of CpG sites examined).
[c] F×10^{-4}, frequency of mutation × 10^{-4}.
[d] That is, the percent frequency of unconverted cytosines at non-CpG sites (i.e., the number of unconverted cytosines at non-CpG sites divided by the total number of non-CpG sites examined).

DNA with unconverted Cs at CpG sites. Two independent PCRs were carried out for each of the four regions, and multiple clones were sequenced from each of the two PCRs after TA cloning. A total of 54 molecules were sequenced from the 12S region, and no methylation at any of the 594 CpG sites was detected (Table 2). This finding suggests that the CpG methylation level in the 12S region is <0.17%. In the 16S region, a single unconverted cytosine was detected from a total of 42 molecules with 546 CpG sites, giving the region a 0.18% methylation frequency (Table 2). No unconverted cytosine is detected in any of the 34 molecules (578 CpG sites) sequenced from the COII region of the mtDNA; thus, the methylation frequency of this region is <0.17% (Table 2). A frequency of 0.63% methylation was calculated in the ATP6 region of mtDNA, based on a single unconverted cytosine detected on each of the three molecules among a total of 60 molecules (480 CpG sites) sequenced (Table 2). None of the unconverted C was at the same CpG site, indicating the randomness of the occurrence. Other single-base mutations were observed on two of these three molecules harboring a single unconverted cytosine at CpG methylation sites. There are some molecules from all four regions that have unconverted cytosines at non-CpG sites, most likely due to incomplete conversion (Table 2). The nonconversion rate is the highest in the ATP6 region at 0.18% and the lowest in the 12S region at 0.08%. These nonconversion rates are much lower than the generally acceptable sodium bisulfite nonconversion background of 0.5%. It is noteworthy that the mutation frequency, which includes mutations that can arise in the HCT116 cells in culture, during PCR, or in *Escherichia coli* after TA cloning, was almost 2-fold higher in the 16S (14.3×10^{-4}) and the ATP6 (9.5×10^{-4}) regions than in the 12S and the COII regions (Table 2). Overall, we observe a much lower level of unconverted cytosines at CpG sites than expected based on the prediction by Shock et al. (5) at the 16S and the ATP6 regions. No unconverted cytosine was found in the 12S and the COII regions of the mtDNA in HCT116 cells. We conclude that HCT116 cells have minimal, if any, cytosine methylation in mtDNA. Also, the modified cytosines observed are sporadic in occurrence and random in nature, instead of at specific sites across the entire molecule, as suggested by Shmookler and Goldstein (1).

Lack of appreciable CpG methylation in mtDNA from primary human cells in specific regions. To further examine whether CpG methylation is present in mtDNA from primary human tissue, the same four regions of the mtDNA from whole blood were analyzed by the sodium bisulfite sequencing method. In the COII region, no unconverted cytosine was detected at any CpG sites on the 16 molecules sequenced, indicating a methylation frequency of <0.37% (Table 3). Methylation frequencies ranging from 0.27 to 0.66% were calculated based on the detection of a single unconverted cytosine at all CpG sites examined from each of the three other regions (Table 3). Rates of nonconversion at non-CpG sites range from 0.02 to 0.06%. The mutation rates range from 7.8×10^{-4} in the 12S region to 40.4×10^{-4} in the ATP6 region (Table 3). Unconverted cytosines at CpG sites were detected in three of the four mtDNA regions examined from primary cells and were at a level of <0.7%. Again, DNA methylation only occurs infrequently, if it exists at all, and the sites of methylation are random in primary human cells, as was found in HCT116 cells.

TABLE 3 Summary of unconverted Cs observed in four regions of mtDNA from human primary cells after sodium bisulfite treatment[a]

Locus (size [bp])	No. of Cs	No. of CpG sites	N	Uncoverted C at CpG sites			Mutation			Uncoverted C at non-CpG sites		
				No. of sites	No. of mol	%[b]	No. of sites	No. of mol	F×10^{-4}[c]	No. of sites	No. of mol	%[d]
12S (292)	79	11	11	0	0		2	2		0	0	
			11	1	1		3	3		1	1	
Total	1,738	242	22	1	1	0.41	5	5	7.8	1	1	0.07
16S (250)	61	13	13	0	0		3	2		0	0	
			15	1	1		8	5		1	1	
Total	1,708	364	28	1	1	0.27	11	7	15.7	1	1	0.07
COII (258)	81	17	11	0	0		7	4		1	1	
			5	0	0		0	0		1	1	
Total	1,296	272	16	0	0	<0.37	7	4	17	2	2	0.20
ATP6 (247)	82	8	14	1	1		11	7		1	1	
			5	0	0		8	4		0	0	
Total	1,558	152	19	1	1	0.66	19	11	40.4	1	1	0.07

[a] Each line represents an independent PCR and cloning experiment. Cs, cytosines at non-CpG sites; N, total number of molecules examined; mol, molecules.
[b] That is, the percent frequency of unconverted cytosine at CpG sites (i.e., the number of unconverted cytosines divided by the total number of CpG sites examined).
[c] F×10^{-4}, frequency of mutation × 10^{-4}.
[d] That is, the percent frequency of unconverted cytosines at non-CpG sites (i.e., the number of unconverted cytosines at non-CpG sites divided by the total number of non-CpG sites examined).

The lack of appreciable methylation in the four regions of mtDNA is not due to cross-contamination of numt sequences in the nuclear genome. Mitochondrial DNA is known to insert into the nuclear genome as "numts " (nuclear DNA sequences of mitochondrial origin). It is known that there are between 1,000 to 10,000 copies of mtDNA in each mammalian cell (19). It is important to consider that the high copy number of mtDNAs per cell can lead to incorrect conclusions when studying these sequences in the nuclear genome, but it is much less likely that these sequences in the nuclear genome would contribute significantly, if at all, to the study of mtDNA. Despite this being unlikely to complicate our study, we carried out a sequence comparison to identify the mismatches between the numts and the mtDNA regions that we studied.

There are 391 numts in the human genome considered as orthologous for which identical loci were found in chimpanzee (20). There are an additional 61 nonorthologous numts that arose after the divergence of human and chimpanzee. Of the 452 numts in the human genome, 35 (34 orthologous and 1 nonorthologous) overlap with the four regions of mtDNA studied here (nucleotides 740 to 1080, 2851 to 3150, 7883 to 8166, and 8970 to 9255). Mismatches between the four specific regions of mtDNA and these 35 numts were examined, and substantial differences were found in all 35 numts. These mismatches would not likely allow amplification from numts (lack of overlap with one of the primers or a high number of mismatches in priming sites) or would allow the origin of the DNA fragments to be clearly identified as 25 is the minimum number of mismatches found in the internal sequences of the amplicons (Table 4). None of the molecules sequenced in the current study was of numt origin, based on the sequences at the sites of mismatch. To further ensure that the DNA fragments examined were not from any region of the nuclear genome that is not previously known to match the mtDNA sequence, the sequences of the four regions studied here were entered into the NCBI BLAST search to identify potential nuclear genome matches. Again, there are sufficient mismatches to allow the identification of origin if the DNA fragments were from the nuclear genome (Table 5). All of the DNA molecules sequenced here were perfect matches to the mtDNA, except for a few sporadic point mutations, likely due to infidelity of the *Taq* polymerase. This analysis clearly shows that none of the molecules sequenced in the present study were from the nuclear genome. This finding is completely consistent with the fact that the probability of identical sequence in the nuclear genome being amplified when mtDNA is the target sequence is extremely low because of the vast excess of mtDNA copy number relative to the nuclear genome in each cell. It is clear that DNA fragments examined in the present study were unequivocally of mtDNA origin and were without CpG methylation and non-CpG methylation.

Analysis of published genome-wide bisulfite sequencing data sets indicates the lack of cytosine methylation in mtDNA. It is remotely possible that cytosine methylation exists in regions of mtDNA, even though we found no evidence of it in the regional specific sodium bisulfite sequencing analysis in our experiments. To be absolutely certain, 10 publicly available genome-wide sodium bisulfite sequencing data sets from four published papers were analyzed for mtDNA methylation (10–13). The nonconversion rate of cytosines at non-CpG sites in the nuclear DNA in these studies ranged from 0.2 to 0.7%, indicating that the background nonconversion rate of the assay is within the general acceptable rate of 0.5% (Table 6). The CpG methylation frequency in mtDNA based on unconverted Cs at CpG sites range from 0.2 to 0.8% in the 10 data sets analyzed, indicating the lack of appreciable methylation above the background level of the assay (Table 6). The frequencies of unconverted Cs at all cytosine sites of mtDNA in these 10 data sets ranged from 0.08 to 1.01%, which is not much

TABLE 4 Summary of the number of mismatches between numts and the four mtDNA regions sequenced

	No. of mismatches[b]											
	12S			16S			COII			ATP6		
Index[a]	F	R	I	F	R	I	F	R	I	F	R	I
1							1	1	5			
9							3	8	80	4	5	69
26	2	1	60	9	4	38						
50							4	2	76	6	3	57
53	1	15	58	2	11	44						
54	3	1	58									
64							6	6	84	6	3	62
66	NA	0										
83							13	7	84	6	4	68
85										4	5	68
116	NA	2										
123				6	4	34						
127	5	3	60	6	7	34				5	3	66
144							3	2	27	3	1	25
150	3	0	46									
172							6	4	69	5	8	63
182							4	4	79			
193	4	0	57	7	NA							
201	9	0	61	9	6	42						
206	3	0	72	8	4	41						
216	NA	2										
223				8	2	36	5	NA				
226	4	2	51	8	NA							
251				5	3	51						
254							6	6	77			
273	4	2	53									
274	4	2	53									
278	3	2	76	10	1	39						
310	5	1	58									
311				6	2	38						
353	NA	3		7	2	34						
375	3	0	54	7	2	32						
391	NA	2										
376	NA	2										
4N*							5	NA				

[a] That is, the index number for numts in Table S1 in the supplemental material of Hazkani-Covo and Graur (20). *, new numts in Table 2 of Hazkani-Covo and Graur (20).
[b] F, site of forward primer; R, site of reverse primer; I, internal DNA between the two primers. NA, one of the priming sites is not available.

TABLE 5 Summary of nucleotide mismatches between the mtDNA fragments sequenced and the two closest matches in the nuclear genome

Fragment[a]	No. of mismatches
12S	
chr5	16
chr11	24
16S	
chr17	33
chrX	37
COII	
chr1	7
chr17	41
ATP6	
chr1	4
chr5	29

[a] chr, chromosome.

higher than the background level of the assay for nonconversion (Table 6). This result indicates the lack of methylation at non-CpG sites in mtDNA. This analysis includes a variety of human cell lines, including HCT116, and primary human cells. It is clear that there is no appreciable methylation at CpG and non-CpG sites in mtDNA.

Next-generation sequencing confirms the lack of appreciable global CpG methylation in mtDNA from HCT116 cells. The published sodium bisulfite sequencing data set on HCT116 cells was obtained using the reduced representation bisulfite sequencing method (10). Although unlikely, the enrichment method may potentially bias the regions of mtDNA being sequenced. To further confirm the methylation status of mtDNA in this published data set, we carried out independent next-generation sequencing experiments using HCT116 cells. Total DNA from HCT116 cells was subjected to sodium bisulfite treatment and sequenced by next-generation sequencing using the Illumina platform. Among the 13,333,334 pair-end reads obtained, 29,471 reads were uniquely mapped to the mitochondrial genome in the hg19 ge-

TABLE 6 Analysis of published datasets of genome-wide sodium bisulfite sequencing for cytosine methylation in mtDNA

Reference	Cell type	% CpG methylation in mtDNA	% sodium bisulfite nonconversion	
			At all Cs in mtDNA	At non-CpGs in nuclear DNA
Akalin et al. (10)[a]	Acute myeloid leukemia	0.259	0.264	0.319
	Healthy $CD34^+$	0.244	0.143	0.266
	Colon cancer cell line HCT116	0.207	0.082	0.333
	Mixed-lineage leukemia	0.305	0.131	0.203
Heyn et al. (11)	$CD4^+$ T cells from centenarian	0.502	0.470	0.456
	$CD4^+$ T cells from newborn	0.483	0.406	0.475
	PBMCs from middle-aged man	0.609	0.708	0.782
Heyn et al. (12)	Healthy B cell	0.225	0.225	0.255
Hon et al. (13)	Breast cancer cell line HCC1954	0.779	1.009	0.677
	Healthy mammary epithelial cells	0.566	0.832	0.556

[a] Data were generated using reduced representation bisulfite sequencing.

nome assembly, with potential amplification duplicates removed based on the pair-end sequences. Based on reads mapped to the lambda phage control, 0.46% of the Cs were unconverted, reflecting the background nonconversion rate of the specific sodium bisulfite treatment. A similar fraction of unconverted Cs was observed at non-CpG sites in nuclear chromosomes (0.55%) and in the mtDNA (0.47%). The fraction of unconverted Cs at CpG sites (either strand) in the nuclear chromosomes is 0.738, indicating a 73.8% average frequency of CpG methylation in the nuclear genome. In contrast, the methylation level determined similarly for mtDNA is 0.48%, which is almost identical to the rates of nonconversion detected in phage DNA control, nuclear DNA, and mtDNA in the sequencing library constructed from HCT116 cells. These 29,471 reads in the final analysis cover the mitochondrial genome to the depth of 94 times without duplications. Therefore, the lack of DNA methylation observed in the regionally specific analysis is not due to the "jackpot" effect of PCR amplification, and it is clearly a mitochondrial genome-wide phenomenon. It is also noteworthy that there are ~4,000 copies mtDNA per HCT116 cell if the number of sequencing reads of mtDNA origin in this next-generation sequencing experiment is an unbiased and quantitative representation of the DNA content in the cells. This estimation of mtDNA copy number per cell is consistent with previous estimations (19). The CpG methylation frequency of 73.8% found in the nuclear DNA of HCT116 is in line with the methylation frequency of >50% reported in human cells (21, 22, 23). These findings clearly show that DNA methylation, if it exists at all in mtDNA, is not any higher than the background nonconversion rate of the assay. Consistent with our findings in the four specific regions of mtDNA described above and the analysis of published data sets, the unbiased next-generation sequencing experiment carried out here conclusively indicates that there is no cytosine methylation in the mtDNA in HCT116 or other human cells at either CpG sites or non-CpG sites.

DISCUSSION

Our previous examination of three regions of mtDNA that account for ca. 4% of the mtDNA genome from HEK293 cells using the sodium bisulfite genomic sequencing method did not detect any cytosine modification at a total of 1,487 CpG sites from 203 molecules (see Table S1 in the supplemental material). This indicates a <0.1% overall frequency of DNA methylation in mtDNA. To rule out the possibility of a cell type or DNA regionally specific methylation, the same regions of mtDNA reported to have CpG methylation were examined here in the sodium bisulfite-treated DNA from HCT116 cells. The lack of an appreciable amount of unconverted cytosine (no higher than 0.66%) at CpG sites in the four regions of mtDNA examined from HCT116 cells indicates no significant cytosine methylation. Further examination of these same regions of mtDNA from primary human cells also failed to detect appreciable amounts of CpG methylation. Analysis of 10 published genome-wide sodium bisulfite sequencing data sets derived from several cell lines and primary cell sources (10–13) provided further support for this finding. Unbiased sodium bisulfite sequencing of total DNA from HCT116 cells using next-generation sequencing further confirms the lack of cytosine methylation at both CpG sites and non-CpG sites. Taking findings from our sequencing experiments and from analyses of published data sets, a general lack of cytosine methylation in human mtDNA is clear. Furthermore, the sites of CpG methylation, if the very few unconverted Cs at CpG sites are considered to be actual DNA methylation rather than assay false positives, are random, not the nonrandom and whole molecule methylation suggested by Shmookler Reis and Goldstein (1). Our findings here unequivocally demonstrate that there is no evidence of appreciable CpG methylation in mtDNA. Although the <1% unconverted Cs cannot be ruled out as true DNA methylation, random DNA methylation at such a low level is unlikely to have biological significance as proposed in previous reports (5–7).

Sodium bisulfite sequencing has been used to examine cytosine modification for 20 years and remains the "gold standard" for DNA methylation analysis (8). In single-stranded DNA, sodium bisulfite treatment does not convert 5mC and 5hmC but converts unmodified cytosine to uracil. Therefore, the 5mC- and 5hmC-containing mtDNA would give rise to unconverted cytosines at CpG sites after sodium bisulfite treatment. This method can provide precise sites and frequency estimations of methylation at CpG sites, even though the frequency estimation could suffer the potential "jackpot" effect in the PCR and cloning preference, which we have minimized by

doing multiple PCR and cloning experiments. Even the extremely low level of unconverted cytosine detected here is likely to be an overestimate (rather than an underestimate) of the true CpG methylation frequency due to mutations and nonconversion events that can give rise to a cytosine in the final sequence. The highest level of methylation detected in the present study is 0.66% in the ATP6 region of mtDNA from primary human cells. This methylation frequency is based on a single unconverted cytosine on a single molecule. Therefore, the overall methylation frequency in mtDNA is likely to be much less than 0.66%. Our next-generation sequencing experiment showed a nonconversion rate of <0.5% for phage control, mtDNA, and nuclear DNA in the sequencing library constructed from the same sodium bisulfite treatment. The sites of these unconverted Cs are entirely random, indicating the potential nature of incomplete conversion or mutation instead of true methylation sites. Unlike many other assays for DNA methylation, the background of the sodium bisulfite sequencing assay can be determined and the sites of such background can also be clearly identified; therefore, a definitive result and interpretation can be derived.

The findings in previous studies reporting DNA methylation and the present study lead to two critical issues: (i) the reliability and sensitivity of the detection methods and (ii) the biological significance of methylation in mtDNA at the level detected. The percent pulldown by anti-5mC and anti-5hmC antibodies reported by Shock et al. (5) cannot be determined from the descriptions in that study. Our experiments with two different concentrations of the anti-5mC antibody for IP of *in vitro* methylated DNA showed a linear relation between the number of methylated CpG sites on the DNA fragment and the percentage pull-down (see Fig. S1 in the supplemental material). However, when fewer than 10 pairs of CpG sites (one CpG on each strand as a pair) are methylated on a DNA fragment, the recovery of DNA targets by IP using anti-5mC antibody is neither reliable nor efficient. With increasing amounts of antibody in the assay, the efficiency of the pulldown increases for both methylated targets and the background. These findings are consistent with the results of Weber et al. (24; unpublished data). There are 427 CpG sites on each strand of the mtDNA; therefore, the DNA fragments of 400 bp in the IP experiment of Shock et al. (5) would have an average of 10 pairs CpG sites on each fragment. Unless more than half of the CpG pairs on the DNA fragment are methylated, the DNA fragment is not likely to be captured by the anti-5mC antibody. Therefore, considering the efficiency of IP, a >50% frequency of methylation at CpG sites should be observed on much more than 5% of the molecules in the sodium bisulfite sequencing of mtDNA, based on the observation of Shock et al. (5). Infantino et al. (6) reported a 0.25 ratio of 5mC to the sum of 5mC and deoxycytidine in the control cells and a 0.13 ratio in patient cells. There are a total of 7,231 cytosines (Cs) and a total of 854 CpG sites on the two strands of mtDNA. If DNA methylation only occurs at CpG sites and all 854 CpG sites are methylated, it would indicate that 2562 unmodified Cs were detected in the mass spectrometry analysis. Any less methylation at CpG sites would indicate even fewer Cs detected in the assay. Therefore, up to 3,815 Cs may have other modifications. The other possible interpretation of the high 5mC ratio detected is that methylation of Cs at non-CpG sites is much more frequent than CpG methylation in mtDNA. The report by Dzitoyeva et al. (7) describes about 0.1 to 0.25 U of 5mC and ~0.05 U of 5hmC in mtDNA isolated from mouse brain using a commercial ELISA kit. Unfortunately, the frequency of methylation cannot be clearly discerned from that publication. If the units in their study are taken as percentage (suggested by Epigentek, manufacturer of the kit that they used) and considering that only 1% of the mouse genome consists of CpG sites (25), then the 0.1% in the study by Dzitoyeva et al. would translate to 10% methylation of CpG sites. The interpretations of results from all of these studies would clearly predict detection of a high number of unconverted Cs in sodium bisulfite assays, well above the 0.5% nonconversion background, and this is not what we observed.

To our surprise, the sodium bisulfite sequencing of mtDNA did not detect any unconverted cytosine in the 203 molecules sequenced from HEK293 cells, the three molecules each with a single unconverted cytosine among a total of 190 molecules sequenced from HCT116 cells, or the three molecules each with a single unconverted cytosine among a total of 85 molecules sequenced from primary human cells. Taking into account the potential contribution from mutation and the very low level of nonconversion by sodium bisulfite, the frequency of CpG methylation may be even lower, if it exists at all in mtDNA. The possibility of numts in the nuclear DNA affecting the results of region-specific sodium bisulfite sequencing was clearly ruled out by primer sequence differences, as well as by examining all mismatches in the specific regions to the nuclear genome in the molecules sequenced. Also, sodium bisulfite sequencing of the same regions of mtDNA harvested by the Hirt method (27), which preferentially extracts small circular DNA from mammalian cells, showed similar results to those described in the present study, supporting the mitochondrial origin of the clones sequenced (data not shown). This comparison also suggests that cautions should be taken when studying numt sequences in the nuclear genome due to the high copy number of mtDNA in each cell that can bias the observation as described previously (26). Analyses of 10 published data sets of genome-wide sodium bisulfite sequencing of DNA from several other human cell sources further support such a conclusion. Next-generation sequencing of sodium bisulfite treated HCT116 DNA to 94× coverage of mtDNA without duplication also clearly confirms the lack of DNA methylation in mitochondria. Although the possibility of 0.5% DNA methylation in mtDNA cannot be ruled out due to the background nonconversion of the sodium bisulfite assay, it is doubtful that a biologically significant function should be considered at this extremely low frequency of CpG methylation, especially given that the sites of these potential methylation are random. The conclusion of no appreciable CpG methylation in human mtDNA in the present study is essential for future studies of mitochondrial biology, since direct effects of cytosine methylation on mitochondrial function can now be unequivocally ruled out. The findings of the present study and of the previous studies serve as an important reminder that conclusions may be made due to pitfalls of the assays used, even with diligent effort. Also, many DNA methylation analysis assays, including genome-wide sodium bisulfite sequencing, frequently read out CpG methylation levels and patterns in terms of the overall frequency at each CpG site, each region, or the entire genome. However, understanding of the biological impact of DNA methylation requires the critical information of the DNA methylation pattern, including the methylation status of contiguous sites on each molecule. Therefore, it is crucial to thoroughly examine experimental results using direct methods that give rise to minimal ambiguity, and the

interpretation of biological importance of experimental observations should only be made with clear knowledge of sites of methylation on each molecule analyzed.

ACKNOWLEDGMENTS

We thank M. R. Lieber and R. Mosteller for critical reading of the manuscript.

This work was partially supported by U.S. National Institutes of Health National Human Genome Research Institute grant R01 HG005238 (A.D.S.) and partially supported by a generous gift from J. Aresty and K. Aresty (C.-L.H).

REFERENCES

1. **Shmookler Reis RJ, Goldstein S.** 1983. Mitochondrial DNA in mortal and immortal human cells. J. Biol. Chem. **258:**9078–9085.
2. **Pollack Y, Kasir J, Shemer R, Metzger S, Szyf M.** 1984. Methylation pattern of mouse mitochondrial DNA. Nucleic Acids Res. **12:**4811–4824.
3. **Nass MMK.** 1973. Differential methylation of mitochondrial and nuclear DNA in cultured mouse, hamster and virus-transformed hamster cells in vivo and in vitro methylation. J. Mol. Biol. **80:**155–175.
4. **Maekawa M, Taniguchi T, Higashi H, Sugimura H, Sugano K, Kanno T.** 2004. Methylation of mitochondrial DNA is not a useful marker for cancer detection. Clin. Chem. **50:**1480–1481.
5. **Shock LS, Thakkar PV, Peterson EJ, Moran RG, Taylor SM.** 2011. DNA methyltransferase 1, cytosine methylation, and cytosine hydroxymethylation in mammalian mitochondria. Proc. Natl. Acad. Sci. U. S. A. **108:**3630–3635.
6. **Infantino V, Castegna A, Iacobazzi F, Spera I, Scala I, Andria G, Iacobazzi V.** 2011. Impairment of methyl cycle affects mitochondrial methyl availability and glutathione level in Down's syndrome. Mol. Genet. Metab. **102:**378–382.
7. **Dzitoyeva S, Chen H, Manev H.** 2012. Effect of aging on 5-hydroxymethylcytosine in brain mitochondria. Neurobiol. Aging **33:**2881–2891.
8. **Frommer M, McDonald LE, Millar DS, Collis CM, Watt F, Grigg GW, Molloy PL, Paul CLA.** 1992. Genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. Proc. Natl. Acad. Sci. U. S. A. **89:**1827–1831.
9. **Huang Y, Pastor WA, Shen Y, Tahiliani M, Liu DR, Rao A.** 2010. The behaviour of 5-hydroxymethylcytosine in bisulfite sequencing. PLoS One **5:**e8888. doi:10.1371/journal.pone.0008888.
10. **Akalin A, Garrett-Bakelman FE, Kormaksson M, Busuttil J, Zhang L, Khrebtukova I, Milne TA, Huang Y, Biswas D, Hess JL, Allis CD, Roeder R, Valk GPJ, Löwenberg B, Delwel R, Fernandez HF, Paietta E, Tallman MS, Schroth GP, Mason CE, Melnick A, Figueroa ME.** 2012. Base-pair resolution DNA methylation sequencing reveals profoundly divergent epigenetic landscapes in acute myeloid leukemia. PLoS Genet. **8:**e1002781. doi:10.1371/journal.pgen.1002781.
11. **Heyn H, Li N, Ferreira HJ, Moran S, Pisano DG, Gomez A, Diez J, Sanchez-Mut JV, Setien F, Carmona FJ, Puca AA, Sayols S, Pujana MA, Serra-Musach J, Iglesias-Platas I, Formiga F, Fernandez AF, Fraga MF, Heath SC, Valencia A, Gut IG, Wang J, Esteller M.** 2012. Distinct DNA methylomes of newborns and centenarians. Proc. Natl. Acad. Sci. U. S. A. **109:**10522–10527.
12. **Heyn H, Vidal E, Sayols S, Sanchez-Mut JV, Moran S, Medina I, Sandoval J, Simó-Riudalbas L, Szczesna K, Huertas D, Gatto S, Matarazzo MR, Dopazo J, Esteller M.** 2012. Whole-genome bisulfite DNA sequencing of a DNMT3B mutant patient. Epigenetics **7:**542–550.
13. **Hon GC, Hawkins RD, Caballero OL, Lo C, Lister R, Pelizzola M, Valsesia A, Ye Z, Kuan S, Edsall LE, Camargo AA, Stevenson BJ, Ecker JR, Bafna V, Strausberg RL, Simpson AJ, Ren B.** 2012. Global DNA hypomethylation coupled to repressive chromatin domain formation and gene silencing in breast cancer. Genome Res. **22:**246–258.
14. **Tsai AG, Chen DM, Lin M, Hsieh JC, Okitsu CY, Taghva A, Shibata D, Hsieh C-L.** 2012. Heterogeneity and randomness of DNA methylation patterns in human embryonic stem cells. DNA Cell Biol. **31:**893–907.
15. **Smith AD, Chung WY, Hodges E, Kendall J, Hannon G, Hicks J, Xuan Z, Zhang MQ.** 2009. Updates to the RMAP short-read mapping software. Bioinformatics **25:**2841–2842.
16. **Hodges E, Molaro A, Dos Santos CO, Thekkat P, Song Q, Uren PJ, Park J, Butler J, Rafii S, McCombie WR, Smith AD, Hannon GJ.** 2011. Directional DNA methylation changes and complex intermediate states accompany lineage specificity in the adult hematopoietic compartment. Mol. Cell **44:**1–12.
17. **Molaro A, Hodges E, Fang F, Song Q, McCombie WR, Hannon GJ, Smith AD.** 2011. Sperm methylation profiles reveal features of epigenetic inheritance and evolution in primates. Cell **146:**1029–1041.
18. **Kodama Y, Shumway M, Leinonen R, International Nucleotide Sequence Database Collaboration.** 2012. The Sequence Read Archive: explosive growth of sequencing data. Nucleic Acids Res. **40:**D54–D56.
19. **Lightowlers RN, Chinnery PF, Turnbull DM, Howell N.** 1997. Mammalian mitochondrial genetics: heredity, heteroplasmy, and disease. Trends Genet. **13:**450–455.
20. **Hazkani-Covo E, Graur D.** 2007. A comparative analysis of numt evolution in human and chimpanzee. Mol. Biol. Evol. **24:**13–18.
21. **Ehrlich M, Gama-Sosa MA, Huang LH, Midgett R, Kuo MKC, McCune RA, Gehrke C.** 1982. Amount and distribution of 5-methylcytosine in human DNA from different types of tissues of cells. Nucleic Acids Res. **10:**2709–2721.
22. **Laurent L, Wong E, Li G, Huynh T, Tsirigos A, Ong CT, Low HM, Kin Sung KW, Rigoutsos I, Loring J, Wei CL.** 2010. Dynamic changes in the human methylome during differentiation. Genome Res. **20:**320–331.
23. **Li Y, Zhu J, Tian G, Li N, Li Q, Ye M, Zheng H, Yu J, Wu H, Sun J, Zhang H, Chen Q, Luo R, Chen M, He Y, Jin X, Zhang Q, Yu C, Zhou G, Sun J, Huang Y, Zheng H, Cao H, Zhou X, Guo S, Hu X, Li X, Kristiansen K, Bolund L, Xu J, Wang W, Yang H, Wang J, Li R, Beck S, Wang J, Zhang X.** 2010. The DNA methylome of human peripheral blood mononuclear cells. PLoS Biol. **8:**e1000533. doi:10.1371/journal.pbio.1000533.
24. **Weber M, Davies JJ, Wittig D, Oakeley EJ, Haase M, Lam WL, Schübeler D.** 2005. Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells. Nat. Genet. **37:**853–862.
25. **Jabbar K, Bernardi G.** 2004. Cytosine methylation and CpG, TpG (CpA), and TpA frequencies. Gene **333:**143–149.
26. **Jeon JP, Shim SM, Nam HY, Baik SY, Kim JW, Han BG.** 2007. Copy number increase of 1p36.33 and mitochondrial genome amplification in Epstein-Barr virus-transformed lymphoblastoid cell lines. Cancer Genet. Cytogenet. **173:**122–130.
27. **Hirt B.** 1967. Selective extraction of polyoma DNA from infected mouse cultures. J. Mol. Biol. **26:**365–369.