Hypermethylation of human DNA sequences in embryonal carcinoma cells and somatic tissues but not in sperm

Xian-Yang Zhang, Paul T.Loflin, Charles W.Gehrke[1], Peter A.Andrews[2] and Melanie Ehrlich

Department of Biochemistry, Tulane Medical School, New Orleans, LA 70112, [1]Department of Biochemistry, Experiment Station Chemical Laboratories, University of Missouri, Columbia, MO 65201 and [2]The Wistar Institute of Anatomy and Biology, Philadelphia, PA 19104, USA

ABSTRACT

Certain human DNA sequences are much less methylated at CpG sites in sperm than in various adult somatic tissues. The DNA of term placenta displays intermediate levels of methylation at these sequences (Sp-0.3 sequences). We report here that pluripotent embryonal carcinoma (EC) cells derived from testicular germ cell tumors are hypermethylated at the three previously cloned Sp-0.3 sequences and seven newly isolated sequences that exhibit sperm-specific hypomethylation. In contrast to their hypermethylation in EC cells, the Sp-0.3 sequences are hypomethylated in a line of yolk sac carcinoma cells, which like placenta, represent an extraembryonic lineage. These DNA sequences, therefore, appear to be subject to coordinate changes in their methylation during differentiation, probably early in embryogenesis, despite their diversity in copy number (1 to 10^4) and primary structure. Two of these Sp-0.3 sequences are highly homologous to DNA sequences in human chromosomal regions that might be recombination hotspots, namely, a cryptic satellite DNA sequence at a fragile site and the downstream region of the β-globin gene cluster.

INTRODUCTION

DNA methylation appears to affect the expression of certain genes in vertebrate cells (1 - 3) and might also influence chromatin structure (4, 5), DNA conformation (6), chromosomal recombination (7, 8), DNA replication (9), and DNA repair (10). Overall levels of 5-methylcytosine (m^5C) in vertebrate genomes (11 - 13) and patterns of methylation of many individual genes (1) vary from tissue to tissue. Genes in sperm are often, but not always, more highly methylated at the examined CpG sites than in one or more somatic DNAs (14 - 20). This observation and the reports of decreases in DNA methylation (demethylation) in early rabbit trophoblast (21) and in murine embryonal carcinoma cells induced to differentiate (22, 23) indicate that demethylation of DNA occurs at very many sites during vertebrate development. In contrast, several murine and human satellite sequences are much less methylated in sperm than in various somatic DNAs (24 - 28). The hypomethylation of these sperm DNA sequences implies that increases in DNA methylation (de novo methylation; 3) are also part of the developmental program for these vertebrates.

Previously, we cloned three human DNA sequences that are much less methylated in sperm DNA than in a wide variety of somatic DNAs (29). Here we show that these sequences as well as seven newly cloned human DNA sequences displaying sperm-specific hypomethylation are highly methylated in undifferentiated embryonal carcinoma cells. In contrast, all of these sequences have low to moderately low methylation levels in placenta and in a line of yolk sac carcinoma cells, both of which represent extrasomatic cell lineages (29, 30). Despite the variety of primary structures and copy numbers of these sequences, they appear to be part of a large group of sequences whose methylation levels change coordinately during differentiation.

MATERIALS AND METHODS

Sources of DNA. High-molecular-weight DNA was isolated from sperm and postnatal human tissues as previously described (11). Unless otherwise specified, these human tissues were from adult male autopsy samples. Term placenta and amnion were also isolated and used for DNA extraction (29). Cultured WISH (31) and HeLa cells were propagated for DNA extraction as above in Minimal Essential Medium (GIBCO) supplemented with 10% fetal calf serum. The cloned cell line, NTERA-2 cl.D1 (NT2/D1), its mother line TERA-2, and three other human embryonal carcinoma (EC) cell lines 833KE, 1156QE, and a cloned derivative of 2102 Ep called 2102Ep cl.4D3 were maintained under conditions favoring retention of their undifferentiated character (32, 33). NT2/D1 and 2102Ep are cloned lines that express many of the properties identified as characteristic of human EC cells, e.g., Surface Antigen phenotype SSEA-3$^+$/SSEA-4$^+$/TRA-1-60$^+$/SSEA-1$^-$ (34, 35). 1156QE and 833KE are uncloned lines that resemble human EC cells in many respects except that 833KE expresses relatively high levels of SSEA-1. All the above EC cell lines were derived from testicular germ cell tumors. The 1411H cell line, although similarly derived, exhibits a yolk sac carcinoma phenotype and does not express the surface antigen phenotype typical of human EC cells (30, unpublished observations). JAR is an uncloned cell line derived from a gestational choriocarcinoma (36), which has a phenotype distinct from that of human EC cells. One batch of NT2/D1 cells was induced to differentiate by a modification of a previously described procedure (37). Confluent cultures were split 1:2 and incubated with medium containing 10^{-5} M retinoic acid. One week later, the culture was split again 1:6 and incubated for another week in retinoic acid-containing medium followed by incubation in drug-free medium for various lengths of time.

The cloned human DNA used as hybridization probes and for sequence analysis were derived from from 0.3 to 1 kb fragments of a HpaII digest of human sperm DNA cloned in M13mp8 DNA (29). Three of these sequences (Sp-0.3-12, -15, and -16) had been previously cloned (29). The other seven Sp-0.3 sequences described in Table 1 were newly cloned by the same methods.

DNA blot hybridization. For characterization of genomic methylation patterns, human DNA inserts (Sp-0.3 sequences) were radiolabeled by a method which gave much stronger hybridization signals than those obtained from these M13 recombinants in a previous study (29). Recombinant viral DNA (1 µg) was used as a template with a pentadecamer primer (Sequencing Primer; New England Biolabs) in a reaction catalyzed by the large (Klenow) fragment of Escherichia coli DNA polymerase I (Bethesda Research Labs). The reaction was performed under standard conditions (38) except that the amount of $[\alpha-{}^{32}P]dATP$ was limiting (100 pmol) and incubation was for 90 min at 24°C. After extraction with organic solvents, precipitation with ethanol, digestion with HindIII and BamHI or EcoRI, and electrophoresis in a 5% polyacrylamide gel, the insert-containing band (usually ∿0.3-0.5 kilobase-pairs; kb) was excised, electro-eluted, purified and used as a probe for blot hybridization (29). As previously, an aliquot of each restriction digest was checked for the absence of inhibitors of digestion using λ $[{}^{32}P]DNA$ as an internal control (29). The cloned probes showing sperm-specific hypomethylation were also used to determine their genomic copy number by dot-blot hybridization to different amounts of human placental DNA. These probes and a standard, single-copy probe from the 5' region (∿515 base-pairs; bp) of the human thymidine kinase gene that was cloned in M13mp8 (39) were prepared identically.

DNA sequence analysis and the determination of the m^5C content. Nine DNA inserts were sequenced by standard methods (40) using $[{}^{35}S]deoxyadenosine$ triphosphate as the source of radiolabel and the above-mentioned pentadecamer primer. For sequencing the two strands, the insert was sequenced in M13mp8 and in the reverse orientation in M13mp19. For inserts longer than 350 bp, synthetic oligonucleotides complementary to part of the sequenced insert were used as primers to allow sequencing of the entire insert. Comparison of Sp-0.3 sequences to sequences in GENBANK was done with the assistance of a computer program (MicroGenie; Beckman).

The m^5C content of DNA from cultured cells was determined by high performance liquid chromatography (HPLC) analysis of DNA digested to deoxy-nucleosides as previously described (11, 41). All determinations were done in duplicate.
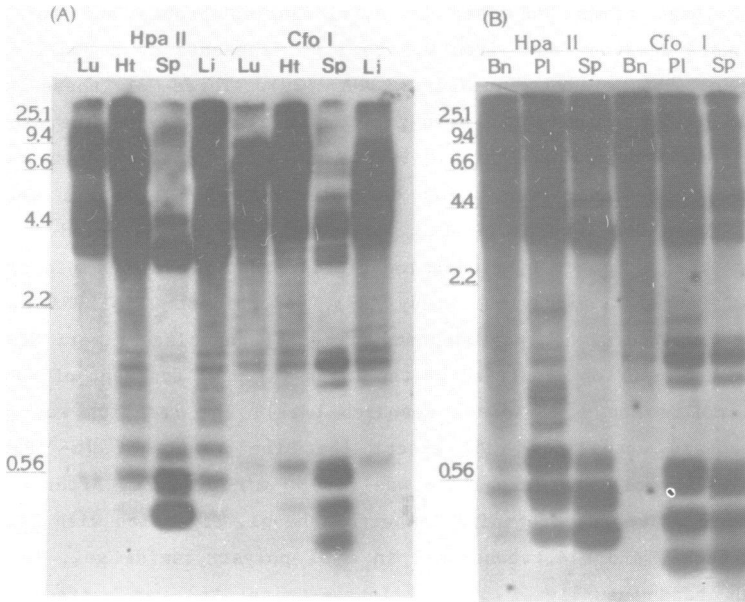
Fig. 1. Analysis of sperm-specific hypomethylation of the cloned sequence
Sp-0.3-23 at HpaII (CCGG) and CfoI (GCGC) sites. Digests of 10 µg of genomic
DNA were prepared, electrophoresed, and subjected to blot hybridization and
autoradiography using the cloned, sperm-derived sequence Sp-0.3-23 as a probe
(see Materials and Methods). The genomic DNAs used for this analysis were
human lung (Lu), heart (Ht), liver (Li), brain (Bn), sperm (Sp), and term
placenta (Pl). Each of the first four tissues was from a different adult
male. Sperm was from semen pooled from four individuals. Placenta refers to
placenta stripped of the chorionic plate, amnion, and smooth chorion, and the
brain DNA was isolated from the cerebellum.

RESULTS

Specific methylation patterns in somatic tissues versus in sperm. We
studied three previously described cloned DNA inserts, Sp-0.3-12, -15 and -16,
(29) and seven newly isolated ones (Sp-0.3-4, -5, -7, -8, -10, -23, -40)
showing sperm-specific hypomethylation (Table 1). When each of these was used
as a probe for blot hybridization, a much lower-molecular-weight distribution
of hybridizing bands was obtained from sperm DNA digested with the CpG me-
thylation-sensitive enzyme HpaII than from analogous digests of adult somatic
DNAs (Figs. 1-3; Table 1). Generally, various different adult somatic DNAs
gave very similar patterns of hypermethylated DNA fragments with a given probe
(29; Table 1). However, some of the cloned sequences had a percentage of their
genomic copies which were equally hypomethylated in all these cells (Fig. 2B
and 3A; Table 1). The seven newly isolated Sp-0.3 DNA clones were also used

as probes for DNA from an extraembryonic organ, term placenta. In each case, hybridizing placental DNA sequences were more HpaII-sensitive than those of adult somatic tissues but less HpaII-sensitive than those of sperm DNA (Figs. 1 & 3; Table 1). In contrast, DNA from term ammion was hypermethylated like adult somatic DNAs (Fig. 4A; Table 1).

The extent of methylation of sperm's CCGG sites within and neighboring the cloned sequences was determined by a comparison of an MspI digest and a HpaII digest of sperm DNA using the above cloned sequences as probes. At most CCGG sites (42), CpG methylation has no effect on MspI digestion, whereas all such methylation inhibits its isoschizomer HpaII. HpaII digests of human sperm DNA probed with Sp-0.3-4, -5, -7, -8, -10, -15, or -40 gave patterns of hybridizing fragments similar to those of MspI digests of human DNA (Fig. 2A & B and data not shown) although, in most cases a small amount of hybridizing high-molecular-weight DNA was seen only in the HpaII digest. In contrast, genomic copies of Sp-0.3-23 sequences in the sperm HpaII digests were approximately equally divided between fragments that were $\leq$0.6 kb and those that were >2 kb (Fig. 1). This heterogeneity was shown to be largely due to residual DNA methylation because only one major band of ∿0.5 kb was seen in an MspI digest of human (placental) DNA.

The seven new Sp-0.3 probes (Table 1) were also used for hybridization to adult somatic DNAs and sperm DNA digested with CfoI, which is inhibited by CpG methylation at GCGC sites. Sperm DNA still gave the lowest-molecular-weight distribution of hybridizing bands (Figs. 1 & 3 and data not shown). As in the HpaII digests, adult somatic DNAs had much of the hybridizing material as high-molecular-weight fragments. Several human DNA samples tested for methylation at other CpG-containing restriction sites gave results similar to those with CfoI and HpaII. For example, many more copies of the genomic Sp-0.3-5 and -23 sequences were hypomethylated at nearby AvaI sites in sperm compared to in heart and liver (Fig. 3C and data not shown). Furthermore, a partial HpaII digest of human sperm DNA which was hybridized to Sp-0.3-23 as a probe, gave some hybridizing bands of only slightly higher molecular-weight than those seen in a complete digest (data not shown). This suggests that many HpaII sites in the vicinity of the Sp-0.3-23 sequences in sperm DNA are unmethylated.

Methylation in embryonal carcinoma cells. A cloned, pluripotent EC cell line NT2/D1 (37) was used as a source of human cells resembling early embryonic cells. The DNA from these cells in the undifferentiated state was digested with HpaII and subjected to blot hybridization with the above Sp-0.3 DNA
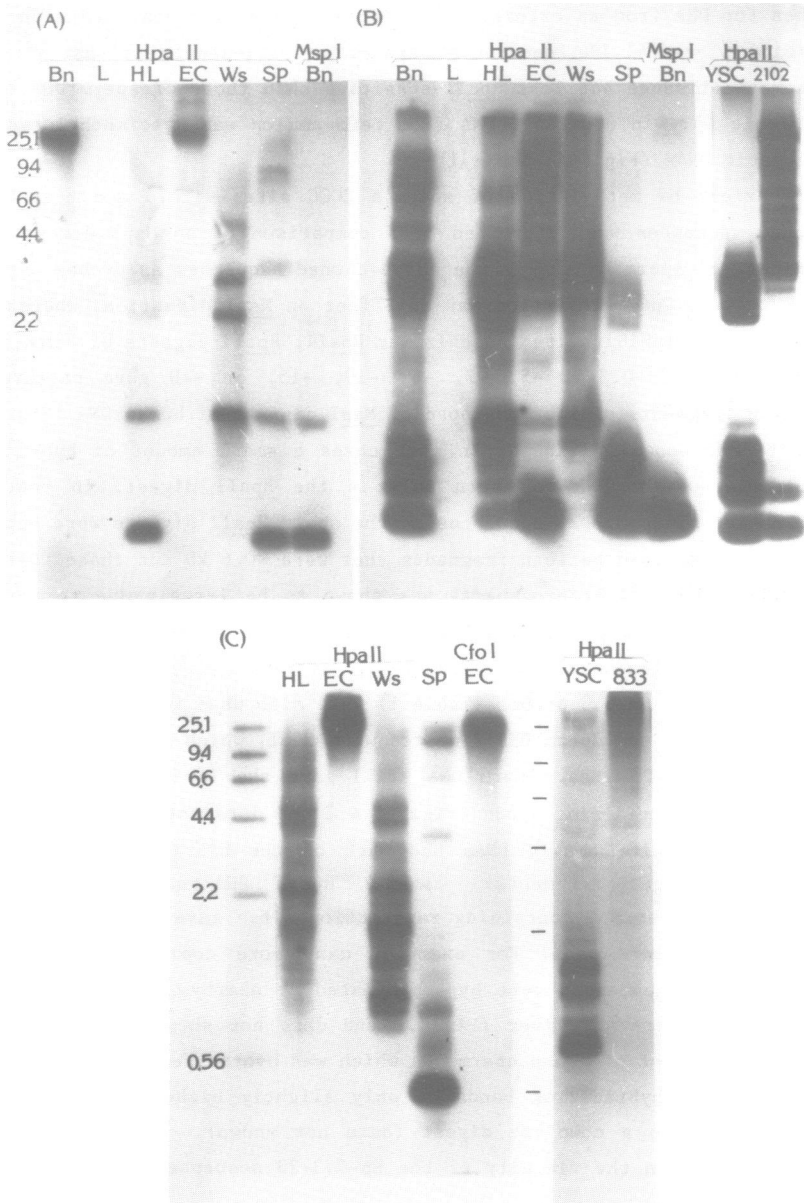
Fig. 2. Hypermethylation of the cloned sequences Sp-0.3-40, -10, and -12 in embryonal carcinoma cells and in several other human cell lines. HpaII, MspI, and CfoI digests of DNA were prepared and subjected to blot hybridization as described in Fig. 1. The DNAs used were from human brain (Bn) and sperm (Sp) as in Fig. 1 and from six types of human cell cultures, undifferentiated NT2/D1 embryonal carcinoma cells (EC), undifferentiated 2102EP and 833KE

embryonal carcinoma cells (2102 and 833), 1411H yolk sac carcinoma cells
(YCS), HeLa cells (HL), and WISH cells (Ws). Also, an established murine cell
line, L cells (L), was used. The positions of HindIII-digested λ [^{32}P]DNA
markers are shown in panels A and C. Panel A, Sp-0.3-40 as the probe; B,
Sp-0.3-10 as the probe; C, Sp-0.3-12 as the probe. The YSC and 2102 samples
in panel B and YSC and 833 samples in panel C were run separately from the
others so that equivalent bands are seen at somewhat different positions.
When Sp-0.3-10 or -40 was used as the probe for HpaII digests of heart, liver,
or lung DNAs, similar results were obtained as from the analogous brain digest
except that, with Sp-0.3-40 as a probe, lung DNA gave hybridizing bands in the
moderate-molecular-weight region and a trace of hybridizing material in the
low-molecular-weight region in addition to the high-molecular-weight
hybridizing sequences (data not shown). The hypermethylation of Sp-0.3-12
sequences in lung, liver, and heart was previously described (29).

clones. The genomic copies of these sequences, which are hypermethylated at
HpaII sites in adult somatic tissues, were also hypermethylated at these sites
in this EC DNA (Figs. 2 - 5; Table 1). In many cases, these DNA sequences
appeared even more highly methylated in EC cells than in adult tissues (Figs.
3B & 4A; Table 1). CfoI digests of this genomic DNA gave similar results
(Figs. 2C & 3C). Also, DNA from the parent EC cell line TERA-2, and from
three other EC cell lines 2102EP, 1156QE, and 833KE, was highly methylated in
all of these Sp-0.3 sequences (Figs. 2B, 2C, & 5; Table 1). The Sp-0.3-7 and
-10 DNA sequences and, to a lesser extent, Sp-0.3-5, which appear in two
fractions in somatic cells, a hypermethylated fraction and a fraction with the
low methylation level of sperm DNA, similarly appeared in both a highly
methylated fraction and an undermethylated fraction in EC DNA (Figs. 2B, 3B,
3C & 4A). In contrast to EC cells, two other established human cell lines,
HeLa and WISH, did not show this extreme hypermethylation of DNA sequences
that are undermethylated in sperm (Fig. 2). No significant homology to mouse
L cell DNA was observed using any of these sequences as hybridization probes.

Because of the high degree of methylation (HpaII-resistance) in EC cells
at each examined DNA sequence exhibiting sperm-specific hypomethylation, we
determined the total m^{5}C content of EC cell DNA by HPLC (11, 41). The EC
cells had rather high genomic m^{5}C contents as follows: NT2/D1, 1.06 mol%
(Table 2); TERA-2, 1.01 mol%; 2102EP, 1.12 mol%; 1156QE, 1.01 mol%; and 833KE,
1.03 mol%. The m^{5}C level in the DNA of EC cells should be compared to those
of various adult human tissues, which range from 0.87 to 1.00 mol% depending
upon the type of tissue (11). These EC cell lines have a similar or slightly
higher genomic m^{5}C content than brain (1.00 mol%; 11) and much higher than
that of sperm (0.84 mol%; 11) although all the teratocarcinomas from which
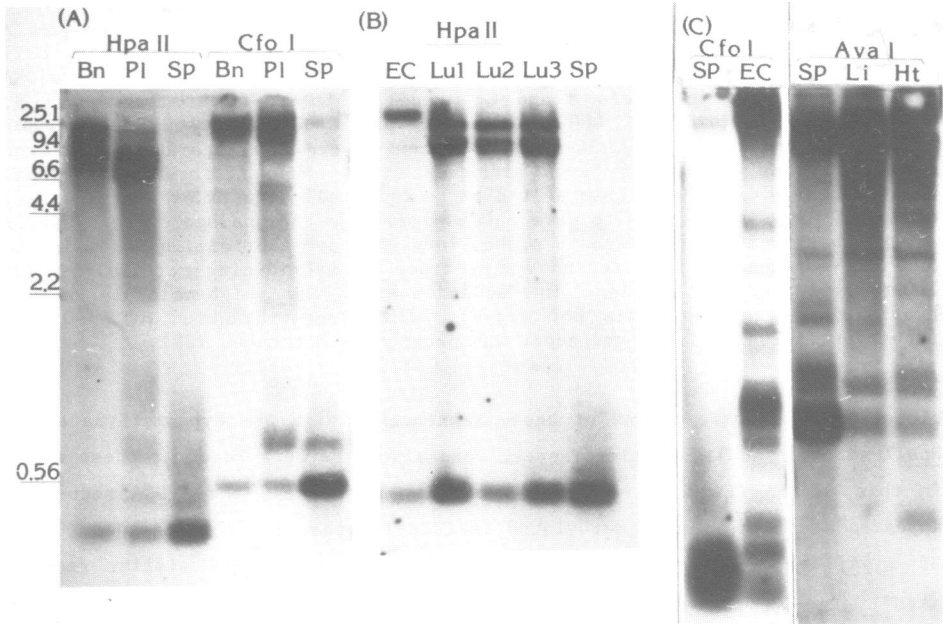they were derived were testicular in origin (33). Also, the genomic m^{5}C

Fig. 3. Analysis of sperm-specific hypomethylation of the cloned sequences Sp-0.3-7 and -5 and their hypermethylation in embryonal carcinoma cells. HpaII, CfoI, and AvaI digests of human brain (Bn), placenta (Pl), sperm (Sp), lung (Lu), liver (Li), heart (Ht), or undifferentiated NT2/D1 embryonal carcinoma cell (EC) DNA were prepared and analyzed by blot hybridization as described in the previous figures. The probe for panels A and B was Sp-0.3-7 and for panel C was Sp-0.3-5. In panel B, Lu 1, Lu 2, and Lu 3 represent DNA from a 4-year-old female, a 33-year-old male, and a 41-year-old male, respectively. With Sp-0.3-5 as a probe, HpaII digests of various adult somatic tissue DNAs gave indistinguishable patterns of hybridizing bands ranging in size from ~0.8-15 kb; only a very small percentage of the hybridization signal was in a 0.6-kb fragment. An analogous HpaII digest of NT2/D1 DNA had even higher-molecular-weight bands in addition to the persisting lower-molecular-weight bands. In contrast, almost all the hybridizing DNA from a sperm HpaII digest was in the 0.6-kb band (data not shown).

contents of other analyzed established human cell lines was much lower than those of the EC lines, namely, HeLa, WISH, JAR, and 1411H had 0.79, 0.73, 0.82, and 0.74 mol% m^5C, respectively. Although the yolk sac carcinoma cell line 1411H had a similar amount of genomic m^5C as in HeLa (a cervical carcinoma derivative) and WISH (amnion derivative) cells, it was the only cell line found to be hypomethylated in all tested Sp-0.3 sequences as compared with those from adult somatic tissues (Figs. 2B, 2C, & 5; Table 1).

The EC cell line NT2/D1, which of all the tested EC lines is capable of the greatest extent of in vitro differentiation, was induced to differentiate

Table 1. Ten human·M13mp8 recombinant DNA clones (0.3-0.55 kb inserts) that show sperm-specific hypomethylation in genomic DNA[a]

Name of sequence	Copy no.	Relative extent of methylation[b] at CCCG sites in						Composition		
		Sperm	Placenta	Adult somatic tissues	Amnion	Three lines of EC cells	Yolk sac carcin. cells	% G + C	% CpG	% CpG obs.[d] / % CpG exp.
Sp-0.3-4	1	-	ND	+++	++	+++	ND	55	2.4	0.31
Sp-0.3-5	20	-/+	-/+	++[c]	++	+++[c]	-/+	64	5.4	0.52
Sp-0.3-7	1	-	++	+++[c]	ND	+++[c]	ND	ND	ND	ND
Sp-0.3-8	100	-	+	++	ND	+++	-	64	4.7	0.48
Sp-0.3-10	400	-/+	+	++[c]	ND	++[c]	-/+	47	5.1	0.93
Sp-0.3-12	50	-	++	+++	+++	+++	+	59	3.0	0.35
Sp-0.3-15	10^3	-	++	+++	+++	+++	+	69	10.6	0.90
Sp-0.3-16	10^4	-/+	+	+++	+++	++++	ND	61	4.6	0.49
Sp-0.3-23	10^3	-/+	+	++[c]	++	+++	+	54	5.3	0.76
Sp-0.3-40	~4	-/+	++	+++	+++	+++	-/+	41	2.6	0.61

[a]From 22 moderately repeated or single-copy clones from the 0.3 to 1 kb fraction of a HpaII digest of human sperm DNA, only these ten showed much less methylation at HpaII sites in sperm than in somatic tissues. Six other analogously cloned DNAs showed only slight hypomethylation in sperm; four others displayed identical methylation patterns at HpaII sites in brain, placental, and sperm DNA; and two were hypermethylated in placenta as described in the text. The methylation patterns of Sp-0.3-12, -15, and -16 in sperm and somatic tissues had been previously described (29) and are included for comparison to those of cultured embryonal carcinoma cells (EC) and a line of yolk sac carcinoma cells.

[b]DNA from sperm pooled from several donors, term placenta or amnion, adult brain, heart, lung, and liver, and from three relatively undifferentiated, human EC cell lines, NT2/D1, 2102EP and 833KE, was used for digestion with HpaII and blot hybridization with the probe indicated in the first column. The yolk sac carcinoma cell line that was tested was 1411H. Symbols: -, almost no detectable methylation at HpaII sites; -/+, most, but not all, copies are undermethylated at HpaII sites; for sequences that are largely methylated, the relative extent of methylation is indicated as + to +++; ND, not determined.

[c]Although most of the indicated genomic copies were highly methylated, some were hypomethylated.

[d]Observed percentage of dinucleotides as CpG divided by the analogous percentage expected for a DNA of the indicated base composition containing a random base sequence. In Figs. 6 & 7 are given the exact length and sequence of each of these cloned inserts.
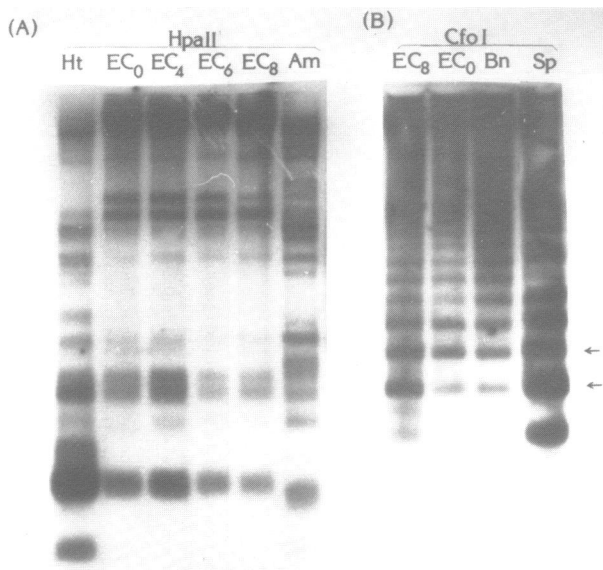
Fig. 4. Effects of in vitro differentiation of NT2/D1 embryonal carcinoma cells on methylation of Sp-0.3-5 and alphoid satellite DNA. Cells were harvested before (EC$_0$) or 4, 6, or 8 weeks after induction of differentiation with retinoic acid (EC$_4$, EC$_6$, EC$_8$). Heart, brain, amnion, and sperm DNA (Ht, Bn, Am, Sp) were included for comparison. The probes used were Sp-0.3-5 in panel A and the EcoRI-dimer, a tandem alphoid satellite DNA repeat (58) in panel B and the genomic DNA was digested with HpaII or CfoI, as indicated, for this blot hybridization analysis. The arrows in panel B show the two bands whose relative intensity was most obviously changed in the differentiated EC cell population compared to in the analogous undifferentiated cell population. As discussed previously, the CfoI or HhaI digest gives an underestimate of the hypomethylation of this satellite DNA (25).

by treatment with retinoic acid. By four weeks after the addition of retinoic acid, most of the cells have undergone differentiation into a variety of cell types including neuronal-type cells and giant cells; very few EC cells persist in such cultures (33, 37). When DNA from cells obtained four to eight weeks after retinoic acid addition was digested with HpaII and hybridized to probes Sp-0.3-4, -5, -7, -12, 13, -23, or -40, similar patterns of hybridizing fragments were seen as with the analogous DNA digests from undifferentiated EC cells (Fig. 4A; data not shown). However, by HPLC analysis, a decrease in the total m^{5}C content of the NT2/D1 cell DNA was detectable at eight weeks after the addition of retinoic acid (Table 2). The average genomic m^{5}C content of that differentiated culture was approximately 10% lower (about 10^7 fewer m^{5}C residues per diploid genome) than that of the undifferentiated cells.
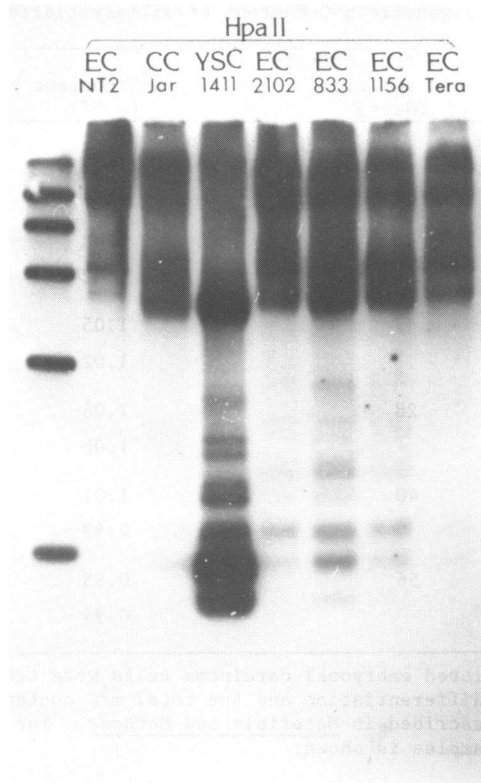
Fig. 5. Methylation patterns of the Sp-0.3-23 repeat in embryonal carcinoma cells, choriocarcinoma cells and yolk sac carcinoma cells. DNA from five embryonal carcinoma cell lines (NT2/D1, 2102Ep, 833KE, 1156QE, and TERA-2), a choriocarcinoma (CC) cell line (JAR), and a yolk sac carcinoma (YSC) cell line (1411H) was digested with HpaII and subject to blot hybridization with the Sp-0.3-23 probe. This blot can be compared to those of Fig. 1, which also involved Sp-0.3-23 as a probe. No differences were observed between these different DNA samples in the analogous MspI digests (data not shown).

Furthermore, by blot hybridization analysis, the human alphoid satellite DNA (EcoRI-dimer), which was more methylated in adult somatic tissues (25) and undifferentiated EC cells (Fig. 4B) than in sperm DNA, became slightly hypomethylated in NT2/D1 cells eight weeks after the initiation of retinoic acid treatment (Fig. 4B). The small degree of demethylation of this EcoRI-dimer upon differentiation of the EC cells might be related to small differences in the extent of methylation of several of the Sp-0.3 sequences in different tissues (Fig. 1 & legend to Fig. 2).

Table 2. Total genomic m^{5}C content of differentiating EC cells

Time after retinoic acid addition[a] (days)	m^{5}C content (mol%)
0	1.07
	1.04
2	1.03
	1.02
7	1.05
	1.02
28	1.06
	1.06
40	1.01
	0.99
56	0.95
	0.94

[a]The undifferentiated embryonal carcinoma cells were treated with retinoic acid to induce differentiation and the total m^{5}C content of their DNA determined as described in Materials and Methods. The m^{5}C content of duplicate DNA samples is shown.

Sequence analysis of Sp-0.3 inserts. Nine of the Sp-0.3 inserts were sequenced (Figs. 6 & 7). Only Sp-0.3-15, -16, and -23 had significant homology with previously determined DNA sequences. The entire Sp-0.3-15 sequence, which has a copy number of ~10^3, is ~95% homologous to part of a newly discovered, cryptic complex satellite DNA (Sst I repeat; 43) with a repeat length of ~2.5 kb (Fig. 7). Consistent with a satellite DNA organization, genomic sequences hybridizing with Sp-0.3-15 gave distinct bands in PstI, KpnI, CfoI, or TaqI digests, which were sometimes multimers of a 2.5-kb repeat length.

In contrast, the Sp-0.3-16 probe gave only smears of hybridizing material in BamHI, MboI, TaqI, BglI or PstI genomic digests with very light bands superimposed on the smears in some of these digests (data not shown and ref. 29). This suggests that Sp-0.3-16, which has a copy number of ~10^4 is part of a dispersed repeat. Furthermore, Sp-0.3-16 is 87% homologous over a 65-bp region to an internally repetitious interspersed repeat previously found in clone p16 from human DNA by Yang et al. (44). In clone p16 are eight tandem copies of a 66-bp sequence, the first 25 bp of which are almost the same as

```
Sp-0.3-4 .          .          .          .          .          .          .          .          .          .
CGGGGAAGGAAGGAGTTGAGAATCAAGAGTTAAAGGGCTAGAAACCTTTGCATGTCCCTAGGCTGCTGACTGCCCCCTAATGACAGGCTACAGAAGGCAG  100

GGGTAAGTCTCTCTCTTCCTCAACCACTGAAGGGGTCCATTTCATCTGGTGAAAGGTCAGCCCCCGACCAAGGCTTGCGGGAAGGCCAGGACGCCATCAC  200

AGACCAGGCAGACATCAGGGCGAGATGGAGTGCTCACCTCAGGGATTCGGGTAAGCCTTCCCTGCCTCACTTAAGGCCACCAGAGAGGAGGTGTGGAGTT  300

GAGTAAAGGCCTCTGTGGGTCTGTCTGCTCCATCCTCTCTTCTCACCATTCAACGTGCATTGCAGATTACCGAAGAACCG  380


Sp-0.3-5 .          .          .          .          .          .          .          .          .          .
CGGGTTGCTCTGTGCTGGTCGGTGCCCTTGTACAGGTGGCAGCTGCGAGCTGACGCTCTCAGTAGAGGTCGGCAGGGTTTGTCCCAGAGAGCCTGAGGAT  100

CGCGGTGTGCACCACCCTCCCAGCCTAGGGTGCACCTCTTCCTTGGCACGCGCCCAGAGCTCGGGGTTTCGGGCGCTGGGCCCTGTGCACGGCTGTCCAGA  200

ATAGGCTGTGCGGCTGGTTCCCGCCCTGGCAAGGCATCCAGCCATGGAATCTGCACTGCTGTTGGGGGCAGGCAAGGTCGGGGGATGGGGGTGTGGTTTC  300

ACACTTGCTAACGGGCGCCACCTGGCATCGGTAGCTGCAGCTGAGAGCATGGCAGAGGCTGGCAGGGCTGGTCCAGACACCTGAGGTCGCTGAGTGCACC  400

GCCCTACCACCCTAGAGTCTCCTGTTCCTTAGACTGCTCCCAGGACGTGGTCTGCGAGCGCTAGACACTGAGCAGCCTCCAGGATGGGGCTGAGCGGCCG  500

ATTCCCGCCTTGCCACAGCTACAGTGTGAATTGAGGCGCCACCGCATTAGTCTGGCCCTGGGGTTCGTGCTACTGGTGGCATGGACAGAGATGGGGGTCG  600

CCACAGCTGCTATGGGCTGAGCAGCCGATTCCCGCCCTCTTGCGCTATGGGACCG  655


Sp-0.3-8 .          .          .          .          .          .          .          .          .
CGGCCAGAATTTCACGGAAGGACATGGGCAGAGAGAGGTCGGAGGGCTCCCGTGCACCTCAGCCGGCCTGTGCACTGAGGAATGTGCAGCCAGAAGGGCT  100

TCCCGGACAGCCAGCCAGCGGTTCTTATAAAGGCCCACAGGCAGGCAGGCTCCACCACTTCATGAATGGCGGTGAGCCCTGGGACAGCTCGCCCCACCAC  200

AGAAGGGTACCAGGGCGTCGAGTCCTGCGGCGCGGGGTGTGGTGGTGTGTGTGGGGGGAGGGCGTGGTGATGGTGGTGGTGGGGACCGGAGAGATGAAGAGGAA  300

GGGTGCTAGGGGGAAGGGGTGAGGGTGGCCCTGTTTCGGGGGCTGGCTCTCCGCAACTCCCTAGGAATCCCGTGGGAACTGGAAGAGGCTCTCAGGGCTCC  400

CAAGCGTCTTCAGCAGGGAAAAACCG  426


Sp-0.3-10.          .          .          .          .          .          .          .          .
CGGCTCTCACTGCACGTTGAGAGAGTCACAAAAGCGCTAGTTCACAACAGGAAAAAACGGCAGCTAACGTGTCTTGGGGAAAATAGACGTCTTCCTGAAA  100

ACTGGGGATTTCTACTTCACCTGAAAAGAAAGACATACGAGAAAGGAAAAACACGAACAAAACAAAACAGAACGAAACAAAACAAGCCAACAAACACGGG  200

CCAAGGCGCCGTCCCTGGAAATCTTAAGTGAGCAAAGTTATTAGTTTTCAGAAAGCGTTTTCTATTTGGGCCAAGTACTGAGAAGGCCCAGACTAGAGCC  300

GTGGCGCCCTCGCCCAYTGTGAAACTCTGCTGGCCG  335
```

Fig. 6. Sequences of four human DNA clones showing sperm-specific hypomethylation. The 5' to 3'
sequence of only one strand is given. Some directly repeated oligonucleotide sequences that are present
in tandem or in tight clusters are overlined. A 27-bp sequence which is present in three copies in
Sp-0.3-5 showing 78-89% homology to one another (allowing a 1-bp insert in one copy) is doubly underlined
and another 40-bp sequence repeated once within Sp-0.3-5 with 82% homology between the copies (including
a 1-bp insert in one of them) is underlined. Also, three sequences resembling the GC box (45) in Sp-0.3-5
are overlined. In Sp-0.3-4, three unusually long polypurine or polypyrimidine tracts are underlined.

that of the next 25 bp and the last 16 bp of which are distinct (44). The
similar region in Sp-0.3-16 contains more than 80% homology to one of the
25-bp units (positions 16 to 40; Fig. 7) followed by the 16-bp unit (positions
41 to 56; Fig. 7) described in clone p16. In Sp-0.3-16, these 41 bp are

```
Sp-0.3-12.          .           .           .           .           .           .           .
CGGCGAGACTGGCCTTGTCTTCTAGGACAAGTGCATTTCCCCTGCCGCTTCCTGTCTCACTCTTGAGGGACATCCTCTCCTCCGCTCACAGGTGGACAGAC  100

                .           .           .           .           .           .           .           .
TCCCTGGATCTTTTGGCTGTAACGAATGTCAGGAAACAAAGGGACTGCGCTGGGACTGGGGCTGGGGCTGCGTGCAGGGGAGGTTGCGTCAGGGCTAGCT  200

    .           .           .           .           .           .           .           .
GGGTGGTGGAGGGTTGGGGGTGGGGTGACTGTTGCAGAAACCTCTGTGGTCCTCTGGCCAGGCATTTCAACATGTGGCTTGGCTGAGGCACAAGCCCCGTC  300

CAGGTTCCCAGGTCTTCTTTGTGTTCCCTTGGCACTCAGGGAAAGGCCATTGATCCCCTTTCGACCG  367


Sp-0.3-15                                                       .           .           .
CGGGCAACGGTGGCTTTCAGTGTGACCCAACGGCAGTCGGGCCTCGCGCATGCTGGCGAGCGACTCCCTCGCTCCACCCCCCCTTACTCCGCAGAGTCAG  100

        .           .           .           .           .           .           .           .
GCTGCGACCCTTTAAAAAATGGCGGCGACGCGGCGGGCTGCGGGGCCTGGGCGGCGGGTGCTGGAGGTTGCGGCGGCGGCGGGCGAAGGCGAGCCGAGGCGG  200

        .           .           .           .           .           .           .           .
CGGGTGGGAAGAGGACTGCCCAGAGGGGCCTGCGGGAGACCCAGGGTCGGACCCCATAGAGTCCTGTGTGAGACCTCCTTGATCGCTCTCTGCTCGGTTCCA  300

CTAAGAGAGCTTCGGGGTGCCG  322


Sp-0.3-16.      1   .           .           .       2   .           .           .
Sp-0.3-16   CGGGAGCAGCTAGATACAGAGTGTCCATTGGTGCATTCACAAACCTTGAGCTAAACACAGGGTGCTGATTGGTGTATTTACAATCCCTGAGCTAGACATA  100
β-3'                    --------CTG-----------T-------------G------A----------A---C-----------T-T--------C-
P16                     --------CTG--------G--T---------T-----G-------A--------------G---

Sp-0.3-16   AAGACTCTCCACGTCCCCACCAGACTCAGGAGCCCAGCTGGCTTCACCCAGTCGATCCCACAGCGGGGTCACAGGTGGAGCTGCCTGCCAGTGCTGTGCC  200
β-3'        --AGT------A--------T---

Sp-0.3-16   ATGCGCTAGCACTCCTCAGCCCTTGGGCGTCGATGGGACTGGGCCCGTGGAGCAAGGGGCGGCATTCGTTGGGGAGGCTCGGGCTGCACAGAAACCACGG  300

Sp-0.3-16   AGGCGGAGGAAGGCTAGGCATGGCGGGCTGCAGTCCCGAGGCCTGCCCGCTGGAAGGCAGCTACAACCCG   370


Sp-0.3-23.          .           .           .           .           .           .           .
Sp-0.3-23   CGGGCTGGTCTCGAACTCCTGACCTCGGGATGACAGACGTGAGCCACTGCGTTCAGTGTACAGTACCATTTCTTAGAAATCACTCCTCACGGGAACACAC  100
Alu         -A--C------------.---------A--

Sp-0.3-23   ACTTATGGGTGACGTGTAGAGATTTTAGTTAGTTAGTTAGTTAGTTAGTTATTATGTGCGGGGGAGGTGGGGGGACGGAGTTTGGCTCTTGCTGCCCAG  200
Alu                                                                   --T-----C-C----X--T-------

Sp-0.3-23   GCTACAGTGCAATGGCCTAGGGGACTCAAGGAGTCAACCTATGGCAGAGAGGACACGTCATTCTGAGCGTAAGGGCCGCAGCGAAAGGTGGCAGGGCCCG  300
Alu         ---GG-----------

Sp-0.3-23   CGCTTTTAAAGGCTGAAATCCCG   323


Sp-0.3-40.          .           .           .           .           .           .           .
CGGTTAATACGGCCTCATGGTCCCACTTTATGGGTGTAACTCACTTAAAGACAATAGATTCTCTAAACAGAAGACATGGGCCTGGAGGAGGTTGGCTCAT  100

        .           .           .           .           .           .           .           .
ATTCGCGTGTAAAACACTAAACTCCATTCCTGGTTCATGGCCCTTGTGGAATATCCTAGGAGAGGTGAACTGATCAAGACGTATCCCAGAATTTGTGTTT  200

        .           .           .           .           .           .           .           .
CTGAAACAAATTCCATTCTTTCCAAGCATTTTGGTCAGAAAAATACTGAGTGGAAGTGCAAAAAGAAAAACAAACAAATAAACTAAGACTGAGCCTGAAA  300

AATCTATTCCCGTTGAAGGTTGCAAACTGTATCGCTACAAACTTCTGGTTTCATTTCCCTTAAAATTGAACCACGGTCTGCATGCCG  387
```
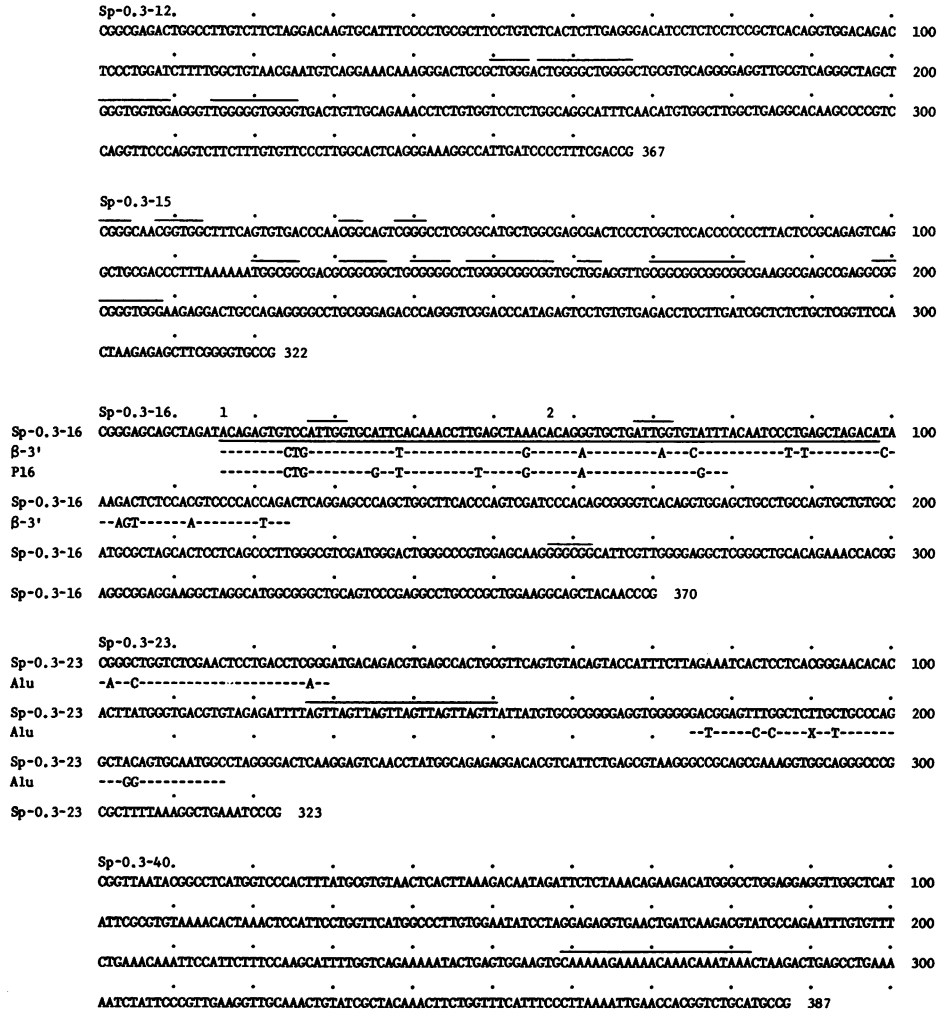
Fig. 7. Sequences of five human DNA clones showing sperm-specific hypomethylation. Tandem or clustered oligonucleotide sequences, including an A-rich region with imperfect repeats in Sp-0.3-40 and (C/T)GG in Sp-0.3-15, are overlined as in Fig. 6. All of Sp-0.3-15 is 95% homologous to the tandem, cryptic satellite Sst I repeat from positions 864 to 1185 (43). Part of the DNA sequences for the 3' region of the human β-globin cluster (β-3', positions 478-370 on the reverse complement of the reported strand; ref. 46) and the human DNA repeat clone p16 (positions 612-676; ref. 43) are aligned beneath Sp-0.3-16 for maximum homology; the nonidentical bases are indicated. The numbers on top of sequence Sp-0.3-16 indicate the start of its two tandem copies of a 41-bp repeat. The complement to the sequence CCAAT and the sequence GGGCGG (the minimal consensus sequence of the GC box) are overlined for Sp-0.3-16. Beneath the Sp-0.3-23 sequence are shown two parts of the consensus sequence for the Alu repeat (positions 87-58, which includes the overlapping B box consensus sequence and transcription directing element for RNA polymerase III, and positions 279-239; ref. 48). A 1-bp gap (X) for maximizing homology is indicated.

repeated once, in tandem, with 78% homology between the copies (Fig. 7). Each copy of the 41-bp repeated element contains the sequence CCAAT and in another part of this cloned insert there is the sequence GGGCGG (Fig. 7); these oligo-
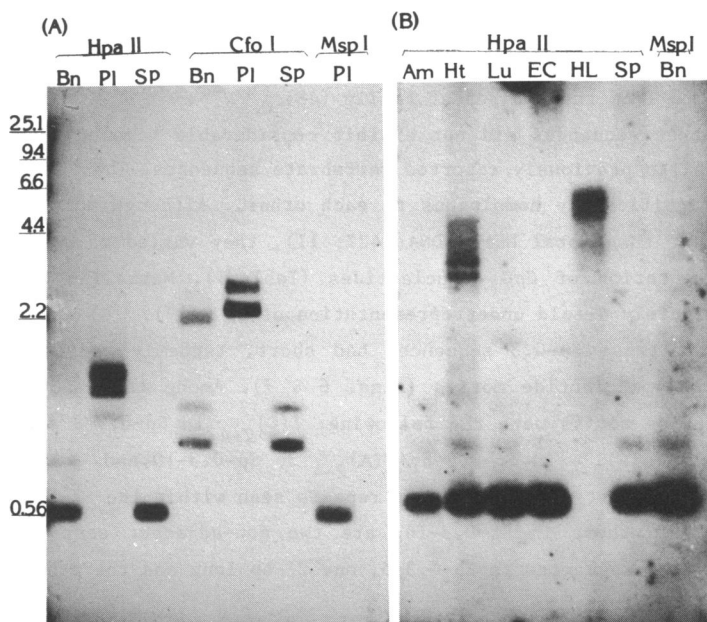
Fig. 8. Analysis of hypermethylation of the cloned sequence Sp-0.3-13 in human placenta. Blot hybridization analysis of DNA from sperm and various human tissues or cultured cell lines was conducted as described in Figs. 1 and 2 except that DNA from term amnion (Am) was included.

nucleotide sequences are associated with many vertebrate promoters (45). One other human sequence homologous to this 66-bp repeat of p16 has been reported and is found 30 kb downstream or the human β-globin gene cluster (46). This sequence has seven copies of a tandem repeat comprised mostly of the 41-bp unit seen in Sp-0.3-16 rather than the longer 66-bp unit of p16 (44, 46). Furthermore, a single copy of a 24-bp sequence is present within this 3' β-globin region repeat, which shows no homology to p16's 66-bp repeat (46) but which is 79% homologous to positions 101 to 124 of Sp-0.3-16 (Fig. 7). This 24-bp sequence of Sp-0.3-16 follows almost immediately after its second 41-bp repeat and is part of a 109-bp region shared with 85% homology between Sp-0.3-16 and the 3' β-globin gene cluster region (Fig. 7).

In Sp-0.3-23, two sequences (30 bp and 41 bp) are 87% or 81% homologous (Fig. 7) to two separate parts of the consensus sequence of the ∿300-bp Alu repeat (47), an interspersed repeat which constitutes ∿7% of the human genome. Within the 30-bp sequence, Sp-0.3-23 has a 17-bp subsequence (positions 7 - 23; Fig. 7) that is almost completely homologous to the overlapping B box

consensus sequence for RNA polymerase III and the Alu repeat-associated transcription-directing element, which appears to be essential and sufficient for transcription from the Alu repeat family (48).

The other sequences did not exhibit considerable homology (>80% over at least 50 bp) to previously reported vertebrate sequences. These sequences were also not significantly homologous to each other. Although most had a higher C + G content than total human DNA (42%; 11), they varied in their extent of underrepresentation of CpG dinucleotides (Table 1). Mammalian DNA generally has approximately 4-fold underrepresentation of CpG (49).

Most of these Sp-0.3 sequences had short, tandemly repeated or highly clustered oligonucleotide motifs (Figs. 6 & 7). Among the tandemly repeated oligonucleotide motifs were the following: $T(G)_{2-4}$ in Sp-0.3-8 and -12; $Y(G)_2$ in Sp-0.3-15; $A(G)_{2-4}$ in Sp-0.3-8; $C(A)_{2-4}$ in Sp-0.3-10 and -40; and AGTT in Sp-0.3-23. The most striking longer repeats seen within the Sp-0.3 sequences, in addition to those in Sp-0.3-16, are two non-adjacent sets of triply or doubly repeated sequences in Sp-0.3-5, one 27 bp long and the other 40 bp long (Fig. 6).

Hypermethylation of sequences in placental DNA. Two of the sequences cloned from the 0.3 to 1 kb fraction of HpaII-digested sperm DNA, Sp-0.3-13 and -31, were more methylated at HpaII and CfoI sites in placental DNA than were most of the copies of these sequences in a variety of adult somatic DNAs or sperm DNA (Fig. 8). Although heart DNA contained some hypermethylated copies of sequence Sp-0.3-13, most of its copies were completely unmethylated at its two adjacent HpaII sites in contrast to placental DNA, which contained no detectable unmethylated copies of this sequence (Fig. 8). This is the first description of DNA sequences hypermethylated in placenta and contrasts with the large extent of hypomethylation that generally characterizes human placental DNAs (11). Undifferentiated NT2/D1 EC cells, but not HeLa cells, were also hypomethylated at these sites (Fig. 8B).

DISCUSSION

Evidence for DNA methylation being involved in germ line-associated, differential genetic imprinting of paternal and maternal genomes has recently been reported (50-52). Another germ line-associated DNA methylation pattern is that murine (26, 27), bovine (28), and human (24, 25) DNAs exhibit under-methylation of certain satellite sequences in sperm but not in the respective somatic tissue DNAs. Here, we demonstrate that several single-copy human DNA sequences as well as a variety of repetitive DNA sequences display sperm-

specific hypomethylation and that, in contrast, three human embryonal carcin-
oma cell lines are hypermethylated in all these sequences (Table 1). That
seven cloned sequences in this study and three in a previous study (29) show
generally similar hypermethylation in adult tissues representing derivatives
of all three primary germ layers suggests that de novo methylation of these
sequences (Sp-0.3 sequences) occurs before formation of the primary germ
layers. This hypothesis is consistent with the hypermethylation of these
sequences seen in amnion, which, like the three primary germ layers, derives
from the embryoblast (see below). Support for this hypothesis is also observed
in the de novo methylation of retroviral DNA in pre-implantation mouse embryos
but not in post-implantation embryos (53). The mammalian female germ line is
subject to X chromosome inactivation, which has been linked to DNA methyla-
tion, followed by reactivation, which may involve DNA demethylation (54 - 56).
Because of the above and because the mammalian germ line arises from the
embryoblast, we propose that the Sp-0.3 sequences in the male germ line, and
possibly also in the female germ line, undergo early de novo methylation
together with the rest of the embryoblast followed by demethylation specifi-
cally in the germ line.

The EC cell lines that were used in the present study originated from
testicular germ cell tumors and are thought to resemble early embryonic cells.
One of them, NT2/D1, can differentiate into a variety of somatic, but appar-
ently not trophoblastic cell types (37). The Sp-0.3 DNA sequences (Figs. 2-5;
Table 1) and the alphoid satellite DNA (Fig. 4) in these EC cells may be
hypermethylated as a result of de novo methylation early in their conversion
to teratocarcinoma cells from germ cells. Such methylation might resemble de
novo methylation very early in embryogenesis if the EC cells truly mimic their
embryonic counterparts. Alternatively, these cells may have never undergone
the demethylation associated with sperm formation and so may have retained the
hypermethylation characteristic of these sequences in somatic cell DNA.

The earliest event in mammalian differentiation is the formation of the
trophoblast lineage, which yields most of the tissues of term placenta. This
is distinct from the embryoblast (inner cell mass) lineage, which is the
precursor of the embryo proper plus certain extraembryonic tissues, such as
the amnion. The above Sp-0.3 sequences displayed a methylation level in
placenta intermediate to those in sperm and adult somatic tissues (Table 1;
29). Placental DNA has an m^5C content markedly lower than those of sperm and
adult tissues (11) and shows an overall genomic hypomethylation even early in
embryogenesis (21). Also, the cell line 1411H, a model for yolk sac

components, another extrasomatic cell lineage (30), was hypomethylated at Sp-0.3 sequences (Figs. 2 & 5; Table 1). This may be the result of very early de novo methylation followed by much demethylation of these sequences as these extrasomatic lineages form. Alternatively, de novo methylation may occur to a lesser extent in the trophoblast and yolk sac lineages. In contrast to the hypomethylation in placenta of the Sp-0.3 sequences described in Table 1, two sequences that we cloned were hypermethylated in placenta relative to all the other studied cell populations (Fig. 8). This illustrates how intricate are the patterns of changes in DNA methylation that accompany vertebrate differentiation.

For the sequences displaying sperm-specific hypomethylation, a comparison of various restriction digests suggests that the region of hypermethylation in the adult somatic DNA is moderately long (>2 kb) and generally includes a high degree of methylation at GCGC sites as well as at CCGG sites and at other CpG sites (Figs. 1-3 and Results). Therefore, it is likely that the de novo methylation during embryogenesis, which is responsible for this somatic hypermethylation is a regional, rather than a localized, phenomenon. Local primary structure is probably not a sufficient signal for embryonic de novo methylation of adult somatic tissues because several sequences were found in both methylated and unmethylated DNA fractions (Figs. 2B & 3; Table 1) and because the sequences of the Sp-0.3 clones were quite diverse (Figs. 6 & 7). The determinants of which chromosomal regions are subject to de novo methylation are unclear. Interspersed repeats (Sp-0.3-16) as well as tandem repeats (Sp-0.3-15; Fig. 7) and single-copy sequences (Sp-0.3-4 and -7; Fig. 6) are included in this class of sequences showing sperm-specific hypomethylation. Oligonucleotide sequences resembling RNA polymerase II promoter-associated elements (CCAAT and GC boxes; 45) were seen in Sp-0.3-16 and Sp-0.3-5 and RNA polymerase III promoter-associated sequences (the B box and the transcription directing element; 48) in a 30-bp region of Sp-0.3-23 homologous to the Alu repeat family (Figs. 6 & 7). Short tandem oligonucleotide repeats, especially those with adjacent G or A residues on one strand were common (Figs. 6 & 7) but are of unknown significance. Whatever the determinants for de novo methylation of the Sp-0.3 sequences (Table 1) and alphoid satellite DNA (Fig. 4; 25) during embryogenesis, other regions of human DNA, such as ribosomal RNA genes and CpG-rich regions of certain genes transcribed by RNA polymerase II (49, 57) are undermethylated in somatic as well as germ cells and, therefore, escape this methylation.

Although the physiological importance of these Sp-0.3 sequences is unclear, the evidence for regional de novo methylation in a large number and variety of sequences during early mammalian embryogenesis and, possibly, for the corresponding demethylation in the germ line suggests an important functional significance. The differentiation-associated de novo methylation of these sequences and their demethylation or maintenance in a hypomethylated state in the germ line should have been easily lost during evolution if it served no selective advantage. That rodents, cows, and humans all contain DNA sequences exhibiting this phenomenon (24-29; this paper) argues for essential role(s) for this differential methylation. These might include influencing chromatin structure because a number of satellite DNAs (24 - 28) and other types of DNA repeats (29; this paper) exhibit sperm-specific hypomethylation. Analysis of methylation of a variety of human DNA repeats in sperm and somatic tissues is necessary to determine how frequently sperm-specific hypo-methylation is associated with repeated DNA sequences. This sperm-specific hypomethylation might also affect recombination because the tandem Sst I repeat (2.5 kb), of which Sp-0.3-15 (0.32 kb) is a part (Fig. 7) is present at only two chromosomal locations, chromosome 19 (q13.1-q13.3) and 4 (4q31), one which is a constitutive fragile site (43). Furthermore, Sp-0.3-16 has 85% homology to a 109-bp region downstream of the human β-globin gene cluster and localized to a part of the genome that appears to have a high frequency of nonhomologous, intramolecular recombination resulting in deletions (46). Chromosomal localization of other Sp-0.3 sequences exhibiting sperm-specific hypomethylation and isolation and characterization of larger, homologous inserts from DNA libraries may provide further evidence for association of these sequences with functionally important chromosomal structures. Such analyses may also elucidate determinants for de novo methylation, a phenomenon which can have a large influence on the expression of cloned genes (1 - 3; 49).

REFERENCES
1. Cooper, D.N. (1983) Hum. Genet. 64, 315-333.
2. Lassar, A.B., Paterson, B.M. and Weintraub, H. (1986) Cell 47, 649-656.
3. Riggs, A.D., and Jones, P.A. (1983) Adv. Cancer Res. 40, 1-30.
4. Keshet, I., Lieman-Hurwitz, J. and Cedar, H. (1986) Cell. 44, 535-543.
5. Schnedl, W., Erlanger, B.F. and Miller, O.J. (1976) Hum. Genet. 31, 21-26.
6. Kylsik, J., Stirdivant, S.M., Singleton, C.K., Zacharas, W. and Wells, R.D. (1983) J. Mol. Biol. 168, 51-71.
7. Hori, T.-A. (1983) Mutat. Res. 121, 47-52.
8. Sutherland, G.R., Parslow, M.I. and Baker, E. (1985) Hum. Genet. 69, 233-237.
9. Jablonka, E., Goitein, R., Marcus, M. and Cedar, H. (1985) Chromosoma 93, 152-156.
10. Hare, J. T. and Taylor, J. H. (1985) Proc. Natl. Acad. Sci. USA 82, 7350-7354.
11. Ehrlich, M., Gama-Sosa, M. A., Huang, L.-H., Midgett, R. M., Kuo, K. C., McCune, R. A. and Gehrke, C. (1982) Nucleic Acids Res. 10, 2709-2721.
12. Gama-Sosa, M.A., Midgett, R.M., Slagel, V.A., Githens, S., Kuo, K.C., Gehrke, C.W. and Ehrlich, M. (1983) Biochim. Biophys. Acta. 740, 212-219.
13. Vanyushin, B.F., Mazin, A.L. Vasilyev, V.K. and Belozersky, A.N. (1973) Biochim. Biophys. Acta 299, 397-403.
14. Cooper, D.N., Errington, L.H. and Clayton, R.M. (1983) DNA 2, 131-139.
15. Libert, F., Vassart, G. and Christophe, D. (1986) Biochem. Biophys. Res. Commun. 134, 1109-1113.
16. Mandel, J.L., and Chambon, P. (1979) Nucleic Acids Res. 7, 2081-2103.
17. Philipsen, J.N.J., Gruber M. and Geert, A.B. (1985) Biochim. Biophys. Acta 826, 186-194.
18. Stein, R., Sciaky-Gallili, N., Razin, A. and Cedar, H. (1983) Proc. Natl. Acad. Sci. USA 80, 2422-2426.
19. van der Ploeg, L.H.T., and Flavell, R.A. (1980) Cell 19, 947-958.
20. Waalwijk, C., and Flavell, R.A. (1978) Nucleic Acids Res. 5, 4631-4639.
21. Manes, C., and Menzel, P. (1981) Nature 293, 589-590.
22. Aujame, L., Craig, J. and McBurney, M.W. (1984) Can. J. Biochem. Cell Biol. 62, 584-591.
23. Bestor, T.H., Hellewell, S.B. and Ingram, V.M. (1984) Molec. Cell. Biol. 4, 1800-1806.
24. Cooke, H.J., Schmidtke, J. and Gosden, J. R. (1982) Chromosoma 87, 491-502.
25. Gama-Sosa, M.A., Wang, R. Y.-H., Kuo, K. C., Gehrke, C.W. and Ehrlich, M. (1983) Nucleic Acids Res. 11, 3087-3095.
26. Ponzetto-Zimmerman, C., and Wolgemuth, D.J. (1984) Nucleic Acids Res. 12, 2807-2822.
27. Sanford, J., Forrester, L. and Chapman, V. (1984) Nucleic Acids Res. 12, 2823-2836.
28. Sturm, K.S., and Taylor, J.H. (1981) Nucleic Acids Res. 9, 4537-4546.
29. Zhang, X.-Y., Wang, R. Y.-H. and Ehrlich, M. (1985) Nucleic Acids Res. 13, 4837-4851.
30. Vogelzang, N.J., Bronson, D., Savino, D., Vessella, R.L., and Fraley, E.F. Cancer 55, 2584-2593.
31. Hayflick, L. (1961) Exp. Cell. Res. 23, 14-20.
32. Andrews, P.W., Bronson, D.L., Benham, F., Strickland, S., and Knowles, B.B. (1980) Int. J. Cancer 26, 269-280.
33. Andrews, P.W., Damjanou, I., Simon, D., Banting, G.S., Carlin, C., Dracopoli, N.C. and J. Fogh. (1984) Lab. Invest. 50, 147-162.
34. Andrews, P.W., Fenderson, B., and Hakomori, S.-I. (1987) Int. J. Androl. 10, 95-104.

35. Fenderson, B.A., Andrews, P.W., Nudelman, E., Clausen, H., and Hakamori, S.-I. (1987) Devel. Biol. 122, 21-34.
36. Pattillo, R.A., Ruckert, A., Hussa, R.O., Bernstein, R., and Delfs, E. (1971) In Vitro 6, 398-399.
37. Andrews, P.W. (1984) Devel. Biol. 103, 285-295.
38. Shenoy, S., Daigle, K., Ehrlich, K.C., Gehrke, C.W. and Ehrlich, M. (1986) Nucleic Acids Res. 14, 4407-4420.
39. Flemington, E., Bradshaw, H., Traina-Dorge, V., Slagel, V. and Deininger, P. (1987) Gene 52, 267-277.
40. Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA. 74, 5463-5467.
41. Gehrke, C.W., McCune, R.A., Gama-Sosa, M.A., Ehrlich, M. and Kuo, K.C. (1984) J. Chromatog. 301, 199-219.
42. Busslinger, M., deBoer, E., Wright, S. Grosveld, F.G. and Flavell, R.A. (1983) Nucleic Acids Res. 11, 3559-3569.
43. Epstein, N.D., Karlsson, S., O'Brien, S., Modi, W., Moulton, A. and Nienhuis, A.W. (1987) Nucleic Acids Res. 15, 2327-2341.
44. Yang, R., Fristensky, B., Deutch, A.H., Huang, R.-C. C., Tan, Y.H., Narang, S.A. and Wu, R. (1983) Gene. 25, 59-66.
45. Kadonaga, J.T., Jones, K.A. and Tjian, R. (1986) Trends Biochem. Sci. 11, 20-23.
46. Henthorn, P.S., Mager, D.L., Huisman, T.H.J. and Smithies, O. (1986) Proc. Natl. Acad. Sci. USA 83, 5194-5198.
47. Deininger, P.L., Jolly, D.J., Rubin, C.M., Friedmann, T. and Schmid, C.W. (1981) J. Mol. Biol. 151, 17-33.
48. Perez-Stable, C., Ayres, T.M. and Shen, C.-K.J. (1984) Proc. Natl. Acad. Sci. USA 81, 5291-5295.
49. Ehrlich, M., and Wang, R.Y.-H. (1981) Science 212, 1350-1357.
50. McGrath, J. and Solter, D. (1984) Cell 37, 179-183.
51. Sapienza, C., Peterson, A.C., Rossant, J., and Balling, R. (1987) Nature 328, 251-254.
52. Reik, W., Collick, A., Norris, M.L., Barton, S.C., and Surani, M.A. (1987) Nature 328, 248-251.
53. Jähner, D., Stuhlmann, H., Stewart, C.L., Harbers, K., Lohler, J., Simon, I. and Jaenisch, R. (1982) Nature 298, 623-628.
54. Kratzer, P.G. and Chapman, (1981) Proc. Nat. Acad. Sci. USA 78, 3093-3097.
55. Lock, L.F., Melton, D.W., Caskey, C.T., and Martin, G.R. (1986) Molec. Cell. Biol. 6, 914-924.
56. Keith, D.H., Singer-Sam, J., and Riggs, A.D. (1986) Molec. Cell. Biol. 6, 4122-4125.
57. Gardiner-Garden, M. and Frommer, M. (1987) J. Mol. Biol. 196, 261-282.
58. Wu, J.C. and Manuelidis, L. (1980) J. Mol. Biol. 142, 363-386.