The *DNMT3B* DNA methyltransferase gene is mutated in the ICF immunodeficiency syndrome

R. Scott Hansen^{*†}, Cisca Wijmenga[‡], Ping Luo^{*}, Ann M. Stanek^{*}, Theresa K. Canfield^{*}, Corry M. R. Weemaes[§], and Stanley M. Gartler^{*†}

Departments of ^{*}Medicine and [†]Genetics, University of Washington, Seattle, WA 98195; [‡]Department of Medical Genetics, Wilhelmina Children's Hospital, 3508AB, Utrecht, The Netherlands; and [§]University Hospital of Nijmegen, 6500HB, Nijmegen, The Netherlands

Contributed by Stanley M. Gartler, October 18, 1999

DNA methylation is an important regulator of genetic information in species ranging from bacteria to humans. DNA methylation appears to be critical for mammalian development because mice nullizygous for a targeted disruption of the *DNMT1* DNA methyltransferase die at an early embryonic stage. No DNA methyltransferase mutations have been reported in humans until now. We describe here the first example of naturally occurring mutations in a mammalian DNA methyltransferase gene. These mutations occur in patients with a rare autosomal recessive disorder, which is termed the ICF syndrome, for immunodeficiency, centromeric instability, and facial anomalies. Centromeric instability of chromosomes 1, 9, and 16 is associated with abnormal hypomethylation of CpG sites in their pericentromeric satellite regions. We are able to complement this hypomethylation defect by somatic cell fusion to Chinese hamster ovary cells, suggesting that the ICF gene is conserved in the hamster and promotes *de novo* methylation. ICF has been localized to a 9-centimorgan region of chromosome 20 by homozygosity mapping. By searching for homologies to known DNA methyltransferases, we identified a genomic sequence in the ICF region that contains the homologue of the mouse *Dnmt3b* methyltransferase gene. Using the human sequence to screen ICF kindreds, we discovered mutations in four patients from three families. Mutations include two missense substitutions and a 3-aa insertion resulting from the creation of a novel 3' splice acceptor. None of the mutations were found in over 200 normal chromosomes. We conclude that mutations in the *DNMT3B* are responsible for the ICF syndrome.

DNA methylation | somatic cell complementation | consanguinity | RNA splicing

ICF syndrome (Mendelian Inheritance in Man 242860) is a rare autosomal recessive disease characterized by a variable immunodeficiency, mild facial anomalies, and centromeric decondensation—chromosomal instability involving chromosomes 1, 9, and 16, (1, 2). Of the few molecular findings reported, the most consistent is hypomethylation of the satellite 2 and satellite 3 regions of chromosomes 1, 9, and 16 (3). Although global methylation patterns appear to be in the normal range (4), other regions of heterochromatin, such as the centromeric α satellite repeats (5) and the inactive X (6, 7), may also be hypomethylated. An ICF locus has been mapped to the proximal long arm of chromosome 20 by examination of consanguineous ICF kindreds (8), but little is known about the etiology of this syndrome.

Genomic methylation patterns play important roles in genome structure and expression, but the processes that determine these complex patterns are, at best, only partially understood (9–12). The ICF syndrome appears to offer an important model for understanding some of these phenomena, prompting us to undertake studies leading to the identification of the defective ICF gene. Because a previous study suggested that the cytogenetic defects in ICF cells could be complemented by fusion to normal cells (13), we explored the possible complementation of the ICF methylation defect by somatic cell hybridization as an approach toward identifying the defective gene. We report here

that introduction of ICF chromosomes into Chinese hamster ovary (CHO) cells by cell fusion results in the *de novo* methylation of their satellite 2 and satellite 3 sequences.

This complementation result led us to examine the ICF locus on chromosome 20 for the presence of a DNA methyltransferase possessing *de novo* activity (i.e., active on duplex DNA with both strands unmethylated). The predominant DNA methyltransferase in mammalian cells is DNMT1 (MCMT1) (14), but it is much more active toward hemimethylated than unmethylated DNA duplexes, and the gene is located on human chromosome 19 (15). In the mouse, *Dnmt3a* and *Dnmt3b* are two DNA methyltransferases that catalyze the methylation of unmethylated and hemimethylated substrates with equal efficiency (16). We identified a human genomic sequence that contains the homologue of the murine *Dnmt3b* gene and maps to the chromosome 20 ICF locus. We then screened ICF families for *DNMT3B* mutations. In four ICF patients from three families, we found three different mutations that cause alterations in, or adjacent to, conserved methyltransferase motifs, and these mutations occur in regions shared by all known splice variants of the enzyme.

Materials and Methods

Clinical and Pathological Phenotypes of the ICF Kindreds. The clinical and pathological phenotypes of the four affected members of the three kindreds studied have been described previously. All affected individuals exhibit the classical features of variable immunodeficiency, chromosomal instability involving chromosomes 1, 9, and 16, and mild facial anomalies. The P1 and P2 ICF cases in the Dutch Caucasian Family 1 and the P3 case in the Turkish Family 2 were described by Wijmenga *et al.* (8); the P4 case (cells designated GM08714 or GM08747) in the North American Caucasian Family 3 (GM08728, mother; GM08729, father) was described by Carpenter *et al.* (17). P1 and P2 came from a consanguineous union, the parents having a common ancestor five generations ago; P3 came from a first cousin marriage; no evidence of consanguinity was found for Family 3 (P4 case) by Carpenter *et al.* (17).

Cells and Cell Culture. The ICF cells used for analysis included lymphoblastoid cells from the P1 male (P1 lymphoblasts), fibroblasts from the P3 female (P3 fibroblast), lymphoblasts (P4 lymphoblast or GM08714) and fibroblasts (P4 fibroblast or GM08747) from the P4 female, lymphoblasts from P4's mother (GM08728 lymphoblasts), and lymphoblasts from P4's father (GM08729 lymphoblasts). GM08747, GM08714, GM08728, and GM08729 were furnished by the Coriell Cell Repositories (Camden, NJ). Control fibroblasts were furnished by G. Martin

Abbreviations: CHO, Chinese hamster ovary; RT, reverse transcription; SNP, single nucleotide polymorphism.

[†]To whom reprint requests should be addressed at: Department of Genetics, Box 357360, University of Washington, Seattle, WA 98195-7360. E-mail: supreme@u.washington.edu.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Table 1. DNMT3B SNPs

Source	GCTG(C/T)TACA, Cys524Cys	AATA(C/T)GAAG, Tyr558Tyr, splice donor site	ACAT(A/G)ATTG, ivs20+139
F3	c/c	c/c	a/a
M3	c/t	c/t	a/g
P4	c/c	c/c	a/a
P1*	t/t	t/t	g/g
10759	c/t	c/t	a/g
188	c/t	c/t	a/g
P3*	c/c	c/c	a/a
dj1085f17 PAC	t	t	g

*Consanguineous family.

(University of Washington) and control lymphoblasts by W. Raskind (University of Washington). Lymphoblasts and fibroblasts were cultured as described (18, 19).

Somatic cell hybrids containing chromosomes with human satellite 2 and satellite 3 sequences from GM08714 ICF lymphoblasts were obtained by polyethylene glycol-mediated fusion of the lymphoblasts with CHO-YH21 hamster cells (20) essentially as described (21). Attached cells were selected for resistance to hypoxanthine, aminopterin, thymidine (HAT) medium, were cloned by limiting dilution into 96-well microtiter plates, and were expanded and analyzed with specific probes for satellite 2 and 3 sequences by Southern blot analysis. Satellite probes were obtained from genomic DNA by linear amplification with the S.2 primer for the satellite 2 probe or the S.3 primer for the satellite 3 probe (22).

Determination of DNMT3B Coding Sequence and Genomic Structure. The protein sequence of the murine *Dnmt3b* gene (GenBank accession no. AF068626) was used to search the Sanger Centre chromosome 20 sequence data for homologous genes by using TBLASTN at their chromosome-specific BLAST server (http://www.sanger.ac.uk/HGP/Chrom_blast_server.shtml). Unfinished sequence data containing 12 unordered segments for the P1 artificial chromosome clone dj1085F17 was found to contain coding sequences with high similarity to the murine gene. The segments were ordered by using homology and gene prediction programs (primarily GENSCAN, <http://CCR-081.mit.edu/GENSCAN.html>) to yield a sequence containing the full length coding region of the human *DNMT3B*. The predicted RNA sequence and genomic structure were used to synthesize oligonucleotides for DNA and cDNA amplification to verify the sequence and to screen ICF samples for mutations (primer sequences are available in Table 2, published as supplemental data on the PNAS web site, www.pnas.org). The ordered genomic sequence was later verified in the finished sequence release (GenBank accession no. AL035071); the deduced mRNA sequence and gene structure for the coding region was verified by reverse transcription (RT)-PCR and is identical to that recently deduced from overlapping cDNA clones (GenBank accession no. AF156488) (23) with the additional finding of two exonic single nucleotide polymorphisms (SNPs) in exon 14 that were frequently observed (see Table 1; the first coding exon was named exon 1 because the upstream exons predicted by GENSCAN and other programs had weak probability scores and could not be verified by RT-PCR analysis).

Mutation Screening. DNA or RNA was extracted from ICF and normal cells, and PCR or RT-PCR analysis was performed as described (24). The *DNMT3B* coding sequence was amplified by PCR from cDNA and/or DNA by using cycling parameters of 94° for 1 min, 60° for 1 min, and 72° for 2 min. Amplimer

sequences were usually determined directly from products purified by using Qiaquick PCR purification kit (Qiagen, Valencia, CA). Cycle sequencing was performed by using the ABI PRISM BigDye Terminator kit (PE Biosystems, Foster City, CA), with either the forward or the reverse primer of the PCR product; the products were analyzed on an ABI PRISM 377XL sequencer (PE Biosystems). In samples containing multiple RT-PCR products because of alternative splicing and/or heterozygous splicing mutations, the products were cloned into the pCR2.1 vector according to the manufacturer's instructions (TA cloning kit, Invitrogen) and were sequenced with a vector primer (M13-R4, 5'-tgtgagcggataacaatttcacacagg-3' or M13-F3, 5'-tccagtcacgacgtgtgtaaagcagc-3'). Alternatively spliced regions yielding low abundance RT-PCR products were also examined by amplification of exonic regions from DNA. In addition to the partial sequences obtained from cDNA products, the first coding exon (exon 1) was also sequenced from a DNA amplimer. Codon numbers and cDNA positions for mutations refer to the largest splice isoform, 3B1 (GenBank accession no. AF156488), with position 1 of the cDNA being the first nucleotide of the initiating ATG codon. A mutation in intron 21 was evaluated for altering RNA splicing by using GENSCAN and NNSPLICE0.9, a neural network splice prediction program (www.fruitfly.org/seq-tools/splice.html).

Rapid screening of control DNAs for the mutations identified was done by restriction analysis of PCR products. Digestion was performed on 10- to 25- μ l PCR product adjusted to 10 mM MgCl₂ by using 7.5–15 units of restriction enzyme (New England Biolabs) for 2 hr at the manufacturer's suggested temperature. Control DNAs were from unrelated, unaffected Caucasians from North America and the Netherlands, totaling 105 individuals for the A603T mutation, 160 individuals for the V726G mutation, and 105 individuals for the intron 21 splice mutation. Haplotype analysis for Family 3 was performed with 17 markers as described (8).

Results

Complementation of the ICF Hypomethylation Defect at Satellite 2 and Satellite 3 Loci. The cytogenetic abnormalities of stretching, breakage, association, and fusion found for satellite 2 and satellite 3 regions of chromosomes 1, 9, and 16 are associated with abnormal hypomethylation of these regions (7). Cytogenetic abnormalities appear to be at least partially mitigated by somatic cell fusion to normal cells (13). To further explore somatic cell complementation in ICF, we fused P4 ICF lymphoblasts (GM08714) with CHO-YH21 fibroblasts and tested hybrid clones for the presence and methylation status of human satellite 2 and satellite 3 sequences. These sequences are hypermethylated in normal fibroblasts and lymphoblasts (Fig. 1). We found seven clones with strong satellite 2 and/or satellite 3 signals by Southern blot analysis, and these sequences had become hypermethylated at *Hpa*II and *Bst*BI sites in all cases (Fig. 1; data not shown). We presumed that this *de novo* methylation resulted from complementation of the defective human ICF gene with a functional hamster homologue and sought to identify and localize genes for DNA methyltransferases with *de novo* activity (active toward substrates with both strands unmethylated).

Sequence and Structure of the Human DNMT3B Gene. MT3 α and MT3 β are two highly homologous mouse DNA methyltransferases with *de novo* activity that were recently identified (16). Because the ICF locus maps to the proximal long arm of chromosome 20, we searched the unfinished chromosome 20 sequence data at the Sanger Centre for similarity to these enzymes and identified a P1 artificial chromosome sequence (dj1085F17) containing a full length coding sequence highly homologous to the murine *Dnmt3b* gene. The Sanger Centre localized the PAC to 20q11.2 by fluorescence *in situ* hybridiza-

Fig. 1. Complementation of the ICF satellite hypomethylation defect after somatic cell hybridization. Satellite 2 sequences were examined for methylation status in the P4 ICF family and in somatic cell hybrid clones derived from fusion of P4 lymphoblasts to CHO cells. DNAs were digested with either the methylation-sensitive enzyme *HpaII* or its methylation-insensitive isoschizimer, *MspI*. In normal lymphoblasts, the satellite sequences are resistant to *HpaII* digestion whereas *MspI* digestion results in a ladder of fragments. This pattern indicates that these satellites are normally hypermethylated at *HpaII* sites. The ICF *HpaII* digestion patterns, in contrast, are very similar to those of *MspI*, indicating that the satellites are hypomethylated. All seven somatic cell hybrid clones derived from P4 lymphoblasts that contained human satellite 2 sequences were more resistant to *HpaII* digestion than in the ICF patient (14-IC3, 14-ID1, 14-ID4, and 14-IA2 are shown), indicating that they had become *de novo* methylated in the CHO cell background.

tion analysis, a region consistent with ICF mapping data. In addition, stSG20534/SHGC-15969 is an STS within the *DNMT3B* gene that has been localized by radiation hybrid mapping to the ICF region near D20S187 (<http://www.sanger.ac.uk/HGP/mapping.shtml>).

We verified that the *DNMT3B* sequence predicted by using murine homology and GENSCAN information was accurate by sequencing RT-PCR and PCR products amplified with derived primers (see *Materials and Methods*). Xie *et al.* (23) recently described the same coding sequence obtained from their overlapping cDNA clones and also localized the gene to 20q11.2. The gene was also mapped to 20q11.2 by Robertson, *et al.* (25). The gene is composed of 22 coding exons and two major regions of alternative splicing (Fig. 2). The different splice forms we

Fig. 2. Structure of the human *DNMT3B* gene. Genomic sequence, protein homology to the murine *Dnmt3b* gene, and cDNA sequence from RT-PCR analysis were used to determine the genomic structure of the human *DNMT3B* gene. Alternative splice forms were detected by RT-PCR and are depicted according to data obtained from mouse (16) and human (23, 25) cDNA clones. The exon-intron boundaries that we determined agree with those recently described by Xie *et al.* (23). Splice forms were named according to Robertson *et al.* (25).

Fig. 3. Homozygous *DNMT3B* mutation in ICF Family 1. Family 1 is a consanguineous Dutch kindred with two affected brothers (P1 and P2) that was described previously (8, 40). (A) A homozygous T to G transversion in codon 726 (2177T → G) was detected in the cDNA of P1 (mt3b-14f:15r product) that changes valine to glycine. (B) An assay was developed to rapidly screen human populations for this mutation. Digestion of the 3bx19 gt-f:mt3b-20r PCR product with *EcoNI* only occurs with the mutant allele, as seen for Family 1 members in which the father (F1) and mother (M1) are heterozygous and the two affected children (P1 and P2) are homozygous for the mutation (DNAs were derived from blood cells, except P1-LB, which is from a lymphoblastoid cell line used in screening cDNA for mutations; nrml is a normal control).

observed agree with those reported previously in mouse (16) and human (23, 25) cells (Fig. 2). The predominant splice variant we detected in lymphoblastoid cells and normal fibroblasts lacks exons 9, 20, and 21 (the 3B3 isoform).

ICF Patients Have Mutations in *DNMT3B*. We examined the *DNMT3B* gene for mutations in ICF patients primarily by sequencing RT-PCR products from derived lymphoblasts and fibroblasts. The amount of *DNMT3B* expression was normal in ICF patients for all products examined. Exonic sequence was also examined by DNA amplification in regions of alternative splicing. Several silent exonic and intronic SNPs were discovered in normal and ICF cells that appear to be highly polymorphic (Table 1). The *DNMT3B* gene is not imprinted in lymphoblasts or fibroblasts because we found expression of exonic SNPs to be biallelic in all informative cases.

In addition to these SNPs, three ICF-specific mutations were discovered: a 2177T → G mutation in exon 19 corresponding to a Val726Gly missense substitution that is homozygous in the two affected brothers of Family 1 (Fig. 3); an IVS21-11G → A transition at a CpG dinucleotide in intron 21 that is homozygous in the P3 patient of Family 2 (Fig. 4) and heterozygous in the P4 patient of Family 3 (Fig. 5); and a 1807G → A transition at a CpG dinucleotide in exon 16 that corresponds to an Ala603Thr missense mutation that is heterozygous in P4 (Fig. 5).

The GENSCAN and NNSPLICE computer programs predict that the intron 21 mutation creates a new splice acceptor site for exon 22 that is nine nucleotides 5' of the normal splice acceptor site, thus resulting in a SerThrPro insertion before codon 807. The P4 compound heterozygote received the Ala603Thr mutation from the mother, but the splice site mutation was not detected in the father (Fig. 5), suggesting that it occurred in his germline. Paternity for P4 was verified by haplotype analysis with 15 informative markers on chromosome 20, and the splice site mutation in P4 was verified to be on the paternal chromosome by *DNMT3B* sequence analysis in P4-derived somatic cell hybrids (data not shown). We developed methods for restriction enzyme analysis of these mutations to rapidly screen DNAs for their

Fig. 4. Homozygous *DNMT3B* mutation in ICF Family 2. Family 2 is a consanguineous Turkish kindred with an affected female proband (P3) described by Wijmenga *et al.* (8). In P3 of Family 2, a homozygous CpG to CpA transition (IVS21-11G → A) was detected 11 nucleotides upstream from the normal 5' splice acceptor site for exon 22 in the mt3b-23f:23r PCR product. The mutation can be detected by *HpaI* analysis of the product; normal alleles are digested by *HpaI*, but the mutant one is not. The father (F2) and mother (M2) of P3 are both heterozygous for the mutation and the two unaffected daughters homozygous for normal alleles. The mutation creates an AG dinucleotide that is predicted to create a new splice acceptor site that retains nine intronic nucleotides at the 5' end of exon 22.

presence (Figs. 3B, 4B, and 5C), and none were found in over 200 chromosomes from unrelated individuals (data not shown).

Fig. 5. Heterozygous *DNMT3B* mutations in ICF Family 3. Family 3 is a North American kindred with a female proband (P4) and no evidence of consanguinity. (A) The IVS21-11G → A mutation described in Fig. 4 for Family 2 was also found on one P4 allele in the mt3b-23f:23r PCR product. (B) A CpG to CpA transition in codon 603 (1807G → A) was detected on one P4 allele (mt3b-11f:11r RT-PCR product) that changes alanine to threonine. This mutation is present in the mother (M3) but not the father (F3). (C) The IVS21-11G → A mutation was analyzed by *HpaI* digestion as described in Fig. 4. The 1807G → A mutation can be detected by the ability of *Tsp45I* to digest the mt3b-19f:19r PCR product; normal products are not digested.

The three mutations are predicted to affect all major splice forms of *DNMT3B*, and this was verified for most of the variants by RT-PCR analysis. As in normal cells, the 3B3 isoform is the major splice variant detected in ICF cells. The intron 21 mutation was found to correlate with the presence of an abnormal 3B3 RT-PCR product that migrates slower than normal in agarose gels, consistent with the prediction of a new splice site (Fig. 6A). The mutant splice site is apparently very efficient because the normal 3B3 product is absent in the P3 homozygote and the mutant product is about as abundant as the normal one in P4. Sequence analyses of the 3B1, 3B3, and 3B4 RT-PCR products from P3 and P4 reveal that the aberrant band contains nine extra nucleotides (TACCCACAG) at the 5' end of exon 22, in agreement with GENSCAN and NNSPLICE predictions (Fig. 6B; data not shown). NNSPLICE predicts a good score of 0.75 for the normal splice site in either the normal or mutant contexts, but the new splice site in the mutant has the much better score of 0.97 using mt3b-23f:23r input sequences. Because the normal splice site was not used in cells from the P3 homozygote, the mutation appears to create an efficient splice acceptor that out competes the normal one downstream.

All of the ICF mutations occur in gene regions that are invariant among the DNMT3-like methyltransferases that have been identified in zebrafish, mice, and humans and are within or near regions of similarity to certain bacteriophage methyltransferases (Fig. 7; data not shown). In addition, these mutations are in the catalytic domain of the enzyme and are within or near motifs present in most ^{m5}C-methyltransferases (23, 26). The Ala603Thr mutation occurs just adjacent to methyltransferase motif I, a region thought to be involved in binding the AdoMet cofactor. The V726G mutation occurs in a region with weak similarity to motif VIII, and the in-frame SerThrPro insertion occurs in the middle of motif IX. The motif VIII to IX region is known to be important for interaction with DNA in many methyltransferases and is the major determinant of target site specificity for monospecific and multispecific ^{m5}C-methyltransferases (27).

Discussion

We have detected mutations in the human *DNMT3B* gene in four ICF patients. None of the mutations were found in over 200 chromosomes of unaffected, unrelated individuals. The syndrome and the gene map to the same 9-centimorgan region on chromosome 20 and the basic molecular abnormality found in ICF is hypomethylation. All of the detected mutations are expected to affect DNA methyltransferase activity as they occur in, or adjacent to, conserved cytosine methyltransferase motifs, and the two missense mutations are at invariant amino acid positions in all known DNMT3-like enzymes present in zebrafish, mice, and humans. We conclude, therefore, that the *DNMT3B* mutations observed in patients are the likely underlying cause of their ICF phenotypes.

The only other mammalian DNA methyltransferase mutations known are those made in *Dnmt1* knockout mice, and the null mutation is embryonic lethal (28). Because the *DNMT3B* methyltransferase should have major *de novo* activity (16), one might also expect lethality from a null mutation in this gene. The ICF phenotype, however, is relatively mild, implying that the ICF mutations may be hypomorphic. This idea is consistent with our observations of normal *DNMT3B* mRNA levels in ICF patients and mutations that are missense or in-frame insertions rather than truncations or reading frame alterations.

A potential effect of these mutations could be to alter the specificity of the enzyme such that it no longer recognizes certain classes of substrates, such as the satellite sequences cited here. There are several splice variants of *DNMT3B* in mice and humans, and the murine 3B1 and 3B2 isoforms were shown to lack specificity for sequences other than CpG (16). The 3B3

Fig. 6. *DNMT3B* alternative splice variants and sequence analysis of the exon 22 splice mutation. Amplification of cDNA with several different primer sets (mt3b-8f:8r, m3b-13f:13r, mt3b-14f:14r, mt3b-15f:15r, mt3b-16f:16r) suggested that several alternative splice forms of *DNMT3B* are present in normal and ICF cells. The predominant splice variant detected in both lymphoblasts (L) and fibroblasts (F) lacks exons 9, 20, and 21 (the 3B3 isoform). (A) The IVS21-11G→A mutation causes an exon 22 splice alteration. The mt3b-14f:14r RT-PCR products in P3 and P4 cDNA contain a larger 3B3 product than normal, correlating with the presence of the intron 21 mutation. (B) A 9-nt addition to exon 22 in P3 and P4 cDNA. Sequence analyses of RT-PCR products corresponding to 3B3 (mt3b-14f:14r) and 3B4 (mt3b-13f:13r; not shown) indicate that the intron 21 mutation creates a new splice acceptor site that adds nine nucleotides to the 3' end of exon 22, effectively creating an insertion of SerThrPro just before Arg 807.

isoform has not been tested and could be important for satellite 2 and 3 methylation. The Val726Gly substitution and SerThrPro insertion mutations are both adjacent to the region deleted in the 3B3 isoform and could specifically alter the activity of this form of the enzyme. The 3B4 and 3B5 variants are probably not functional because they introduce truncating stop codons and/or reading frame alterations (25).

Fig. 7. *DNMT3B* mutations occur at positions that are invariant in *DNMT3*-like methyltransferases and are within or near motifs that are conserved in most cytosine methyltransferases. *DNMT3*-like methyltransferases identified in zebrafish (GenBank accession no. AF135438), mice (GenBank accession nos. AF068625 for *Dnmt3a* and AF068626 for *Dnmt3b*), and humans (GenBank accession nos. AF067972 for *DNMT3A* and AF156488 for *DNMT3B*) are compared in the regions in which ICF mutations were found. Sequences were aligned with CLUSTALW 1.7 (<http://dot.imgen.bcm.tmc.edu:9331/multi-align/multi-align.html>) and were analyzed by BOXSHADE 3.21 (<http://www.isrec.isb-sib.ch:8080/software/BOX-form.html>). The mutations detected in the 3 ICF families are depicted above the aligned sequences in boxes. Regions that are similar to conserved DNA methyltransferase motifs are also depicted above the alignment (dotted line indicates that motif VIII similarity is weak). Key for amino acid comparisons: red, identical; blue, conservative differences; black, nonconservative differences.

Our complementation studies suggest that *de novo* methylation of satellites by *DNMT3B* in somatic cells serves an important cellular function. The methylation of satellites and other heterochromatic regions may not be as well maintained by *DNMT1* as are other genomic regions, and these regions may require more *de novo* methylation than other regions to maintain a hypermethylated state. Thus, the role of *DNMT3B* in somatic cells could be to restore methylation status at sites that are not faithfully remethylated by the maintenance methylation system. *DNMT3B* could also be targeted to heterochromatic or repetitive regions by specific protein-protein interactions. Mutations in the *mascl* DNA methyltransferase in *Ascobolus*, for example, are associated with the specific loss of the targeted *de novo* methylation of repeated sequences (29).

Though the ICF phenotype is not severely abnormal, affected individuals exhibit a range of alterations. The pericentromeric decondensation is a consistent feature that is readily understood in terms of hypomethylation. In fact, the similarity of the cytogenetic features of ICF to 5-azacytidine-induced cytological abnormalities in normal cells (30, 31) led investigators to look for hypomethylation in centromeric regions of ICF. Abnormal pairing of the affected centromeric regions also occurs in ICF somatic cells. This phenomenon is reminiscent of meiosis in which heavily methylated regions are demethylated to allow for synapsis (32). In the case of somatic cells, however, hypomethylation of limited regions, such as satellites 2 and 3, creates an abnormal cytological conformation. This effect of hypomethylation illustrates the importance of methylation for chromatin organization.

Another interesting feature of ICF patients is a generalized immunodeficiency leading to respiratory infections. The rearrangements that are critical in B cell development could be disturbed if nonspecific hypomethylation occurred within the rearrangement regions (33, 34). The other general phenotype in ICF is referred to as facial anomalies, the type of feature that leads pediatricians to call for a cytogenetic analysis. These are the types of abnormalities found in aneuploidies that have been largely explained by dosage phenomena. There is no evidence of aneuploidy in ICF individuals. How, then, can this phenotype result from an abnormal methyltransferase? One possibility is that the regulation of several critical genes is disrupted either directly or indirectly by hypomethylation. For example, the chromatin organization of the satellite regions may have position effects on the expression or repression of closely linked genes important for development, and hypomethylation of these satellites could disturb the expression patterns of such genes. A similar phenomenon may be occurring in other diseases, such as Roberts syndrome (Mendelian Inheritance in Man 268300) (35, 36), that involve abnormalities in centromeric heterochromatin.

DNA methylation is widespread; very few organisms lack it, and, among those, remnants of methyltransferases are present that may now carry out other functions (37, 38). In addition to the ICF syndrome, the Rett syndrome is another human disease that has recently been found to involve an apparent deficiency in the methylation system in that mutations were found in *MECP2*, an epigenetic regulator that specifically binds to methylated DNA (39). One of the most interesting and unexplained features of methylation is the complexity of the methylation patterns (10). Analysis of mutant methyltransferases, such as the *DNMT3B* forms found in the ICF syndrome, may well shed light on these complexities.

We thank Jan A. F. M. Luyten for organizing patient materials, Eric Strengman for technical assistance, Bert van de Heuvel for helpful discussions, and Henk Vietor for cell culture work. We also thank Eric Lynch for useful suggestions and help with preliminary experiments. Wendy Raskind and Karen Stephens kindly supplied control DNAs from unrelated individuals. Lester Goldstein and Albert La Spada contributed useful editorial suggestions for the manuscript. This work was supported by grants from the National Institutes of Health (Grants HD16659 and GM52463) and the Dutch Heart Foundation.

1. Tiepolo, L., Maraschio, P., Gimelli, G., Cuoco, C., Gargani, G. F. & Romano, C. (1979) *Hum. Genet.* **51**, 127–137.
2. Hulten, M. (1978) *Clin. Genet.* **14**, 294.
3. Jeanpierre, M., Turleau, C., Aurias, A., Prieur, M., Ledeist, F., Fischer, A. & Viegas-Pequignot, E. (1993) *Hum. Mol. Genet.* **2**, 731–735.
4. Ji, W., Hernandez, R., Zhang, X. Y., Qu, G. Z., Frady, A., Varela, M. & Ehrlich, M. (1997) *Mutat. Res.* **379**, 33–41.
5. Miniou, P., Jeanpierre, M., Bourc'his, D., Coutinho Barbosa, A. C., Blanquet, V. & Viegas-Pequignot, E. (1997) *Hum. Genet.* **99**, 738–745.
6. Bourc'his, D., Miniou, P., Jeanpierre, M., Molina Gomes, D., Dupont, J., De Saint-Basile, G., Maraschio, P., Tiepolo, L. & Viegas-Pequignot, E. (1999) *Cytogenet. Cell Genet.* **84**, 245–252.
7. Miniou, P., Jeanpierre, M., Blanquet, V., Sibella, V., Bonneau, D., Herbelin, C., Fischer, A., Niveleau, A. & Viegas-Pequignot, E. (1994) *Hum. Mol. Genet.* **3**, 2093–2102.
8. Wijmenga, C., van den Heuvel, L. P., Strengman, E., Luyten, J. A., van der Burgt, I. J., de Groot, R., Smeets, D. F., Draaisma, J. M., van Dongen, J. J., De Abreu, R. A., et al. (1998) *Am. J. Hum. Genet.* **63**, 803–809.
9. Turker, M. S. & Bestor, T. H. (1997) *Mutat. Res.* **386**, 119–130.
10. Colot, V. & Rossignol, J. L. (1999) *BioEssays* **21**, 402–411.
11. Li, E., Beard, C., Forster, A. C., Bestor, T. H. & Jaenisch, R. (1993) *Cold Spring Harbor Symp. Quant. Biol.* **58**, 297–305.
12. Ramsahoye, B. H., Davies, C. S. & Mills, K. I. (1996) *Blood Rev.* **10**, 249–261.
13. Schuffenhauer, S., Bartsch, O., Stumm, M., Buchholz, T., Petropoulou, T., Kraft, S., Belohradsky, B., Hinkel, G. K., Meitinger, T. & Wegner, R. D. (1995) *Hum. Genet.* **96**, 562–571.
14. Bestor, T. H. & Verdine, G. L. (1994) *Curr. Opin. Cell Biol.* **6**, 380–389.
15. Yen, R. W., Vertino, P. M., Nelkin, B. D., Yu, J. J., el-Deiry, W., Kumaraswamy, A., Lennon, G. G., Trask, B. J., Celano, P. & Baylin, S. B. (1992) *Nucleic Acids Res.* **20**, 2287–2291.
16. Okano, M., Xie, S. & Li, E. (1998) *Nat. Genet.* **19**, 219–220.
17. Carpenter, N. J., Filipovich, A., Blaese, R. M., Carey, T. L. & Berkel, A. I. (1988) *J. Pediatr.* **112**, 757–760.
18. Hansen, R. S., Canfield, T. K., Stanek, A. M., Keitges, E. A. & Gartler, S. M. (1998) *Proc. Natl. Acad. Sci. USA* **95**, 5133–5138.
19. Hansen, R. S., Ellis, N. A. & Gartler, S. M. (1988) *Mol. Cell. Biol.* **8**, 4692–4699.
20. Rosenstraus, M. & Chasin, L. A. (1975) *Proc. Natl. Acad. Sci. USA* **72**, 493–497.
21. Davidson, R. L. & Gerald, P. S. (1977) *Methods Cell Biol.* **15**, 325–338.
22. Tagarro, I., Fernandez-Peralta, A. M. & Gonzalez-Aguilera, J. J. (1994) *Hum. Genet.* **93**, 383–388.
23. Xie, S., Wang, Z., Okano, M., Nogami, M., Li, Y., He, W. W., Okumura, K. & Li, E. (1999) *Gene* **236**, 87–95.
24. Sasaki, T., Hansen, R. S. & Gartler, S. M. (1992) *Mol. Cell. Biol.* **12**, 3819–3826.
25. Robertson, K. D., Uzvolgyi, E., Liang, G., Talmadge, C., Sumegi, J., Gonzales, F. A. & Jones, P. A. (1999) *Nucleic Acids Res.* **27**, 2291–2298.
26. Kumar, S., Cheng, X., Klimasauskas, S., Mi, S., Posfai, J., Roberts, R. J. & Wilson, G. G. (1994) *Nucleic Acids Res.* **22**, 1–10.
27. Trautner, T. A., Pawlek, B., Behrens, B. & Willert, J. (1996) *EMBO J.* **15**, 1434–1442.
28. Li, E., Bestor, T. H. & Jaenisch, R. (1992) *Cell* **69**, 915–926.
29. Malagnac, F., Wendel, B., Goyon, C., Faugeron, G., Zickler, D., Rossignol, J. L., Noyer-Weidner, M., Vollmayr, P., Trautner, T. A. & Walter, J. (1997) *Cell* **91**, 281–290.
30. Viegas-Pequignot, E. & Dutrillaux, B. (1976) *Hum. Genet.* **34**, 247–254.
31. Schmid, M., Haaf, T. & Grunert, D. (1984) *Hum. Genet.* **67**, 257–263.
32. Sanford, J., Forrester, L., Chapman, V., Chandley, A. & Hastie, N. (1984) *Nucleic Acids Res.* **12**, 2823–2836.
33. Mostoslavsky, R., Singh, N., Kirillov, A., Pelanda, R., Cedar, H., Chess, A. & Bergman, Y. (1998) *Genes Dev.* **12**, 1801–1811.
34. Engler, P., Weng, A. & Storb, U. (1993) *Mol. Cell. Biol.* **13**, 571–577.
35. Van Den Berg, D. J. & Francke, U. (1993) *Am. J. Med. Genet.* **47**, 1104–1123.
36. Sinha, A. K., Verma, R. S. & Mani, V. J. (1994) *Hum. Hered.* **44**, 121–126.
37. Pinarbasi, E., Elliott, J. & Hornby, D. P. (1996) *J. Mol. Biol.* **257**, 804–813.
38. Hung, M.-S., Karthikeyan, N., Huang, B., Koo, H.-C., Kiger, J. & Shen, C.-K. J. (1999) *Proc. Natl. Acad. Sci. USA* **96**, 11940–11945.
39. Amir, R. E., Van Den Veyver, I. B., Wan, M., Tran, C. Q., Francke, U. & Zoghbi, H. Y. (1999) *Nat. Genet.* **23**, 185–188.
40. Smeets, D. F., Moog, U., Weemaes, C. M., Vaes-Peters, G., Merckx, G. F., Niehof, J. P. & Hamers, G. (1994) *Hum. Genet.* **94**, 240–246.