

# Supplementary Information

## Mutations in DNMT1 cause hereditary sensory neuropathy with dementia and hearing loss

Christopher J. Klein<sup>1</sup>, Maria-Victoria Botuyan<sup>2\*</sup>, Yanhong Wu<sup>3\*</sup>, Christopher J. Ward<sup>4\*</sup>, Garth A. Nicholson<sup>5</sup>, Simon Hammans<sup>6</sup>, Kaori Hojo<sup>7</sup>, Hiromitch Yamanishi<sup>7</sup>, Adam R. Karpf<sup>8</sup>, Douglas C. Wallace<sup>9</sup>, Mariella Simon<sup>9</sup>, Cecilie Lander<sup>10</sup>, Lisa A. Boardman<sup>11</sup>, Julie M. Cunningham<sup>3</sup>, Glenn E. Smith<sup>12</sup>, William J. Litchy<sup>1</sup>, Benjamin Boes<sup>13</sup>, Elizabeth J. Atkinson<sup>14</sup>, Sumit Middha<sup>14</sup>, P. James Dyck<sup>1</sup>, Joseph E. Parisi<sup>15</sup>, Georges Mer<sup>2</sup>, David I. Smith<sup>3</sup> & Peter J. Dyck<sup>1</sup>

### Affiliations:

<sup>1</sup>Mayo Clinic, Department of Neurology, Division of Peripheral Nerve Diseases, Rochester, MN, USA. <sup>2</sup>Mayo Clinic, Biochemistry and Molecular Biology, Rochester MN, USA. <sup>3</sup>Mayo Clinic, Laboratory Medicine and Pathology, Rochester, MN, USA. <sup>4</sup>Mayo Clinic, Nephrology and Hypertension Research, Rochester MN, USA. <sup>5</sup>University of Sydney, Molecular Medicine Laboratory & ANZAC Research Institute, Australia. <sup>6</sup>Southampton University Hospitals NHS Trust, Department of Neurology, Southampton, United Kingdom, <sup>7</sup>Harima Sanatorium, Division of Neuropsychiatry, Hyogo, Japan. <sup>8</sup>Roswell Park Cancer Institute, Department of Pharmacology and Therapeutics, Buffalo, NY, USA. <sup>9</sup>Center for Molecular & Mitochondrial Medicine and Genetics, University of California, Irvine, USA. <sup>10</sup>Queensland Health, Royal Brisbane Hospital, Herston, Australia. <sup>11</sup>Mayo Clinic, Division of Gastroenterology, Rochester, MN, USA. <sup>12</sup>Mayo Clinic, Division of Psychology, Rochester, MN. <sup>13</sup>Roche Applied Science Genomic Sequencing, Indianapolis, IN, USA. <sup>14</sup>Mayo Clinic, Biomedical informatics and Statistics, Rochester, MN, USA. <sup>15</sup>Mayo Clinic, Division of Neuropathology Rochester, MN, USA.

\* Authors contributed equally

### Authors Contributions:

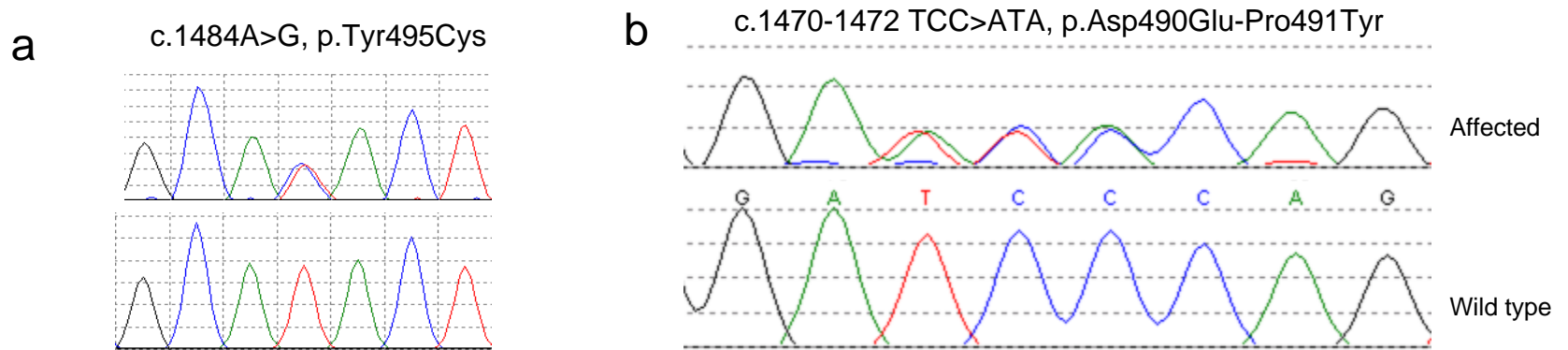
C.J.K., D.I.S. and P.J.D. directed the study. C.J.K. wrote the paper. C.J.K., P.J.D., G.A.N., S.H., K.H., H.Y., D.C.W., M.S., C.L., L.A.B., G.E.S., W.J.L., evaluated or collated patient data. E.J.A. did the linkage and haplotype analysis. S.M. and B.B. did next generation sequencing analysis. C.J.K., C.J.W. and Y.W. did the cell culture and protein expression studies, gene sequencing and southern blot analysis. J.M.C. and A.R.K. did the methylation analysis. M.V.B. and G.M. did the mutagenesis, bacterial protein expression and structural analysis. J.E.P. provided pathologic analysis of autopsy material.

To whom correspondence should be addressed

E-mail: klein.christopher@mayo.edu

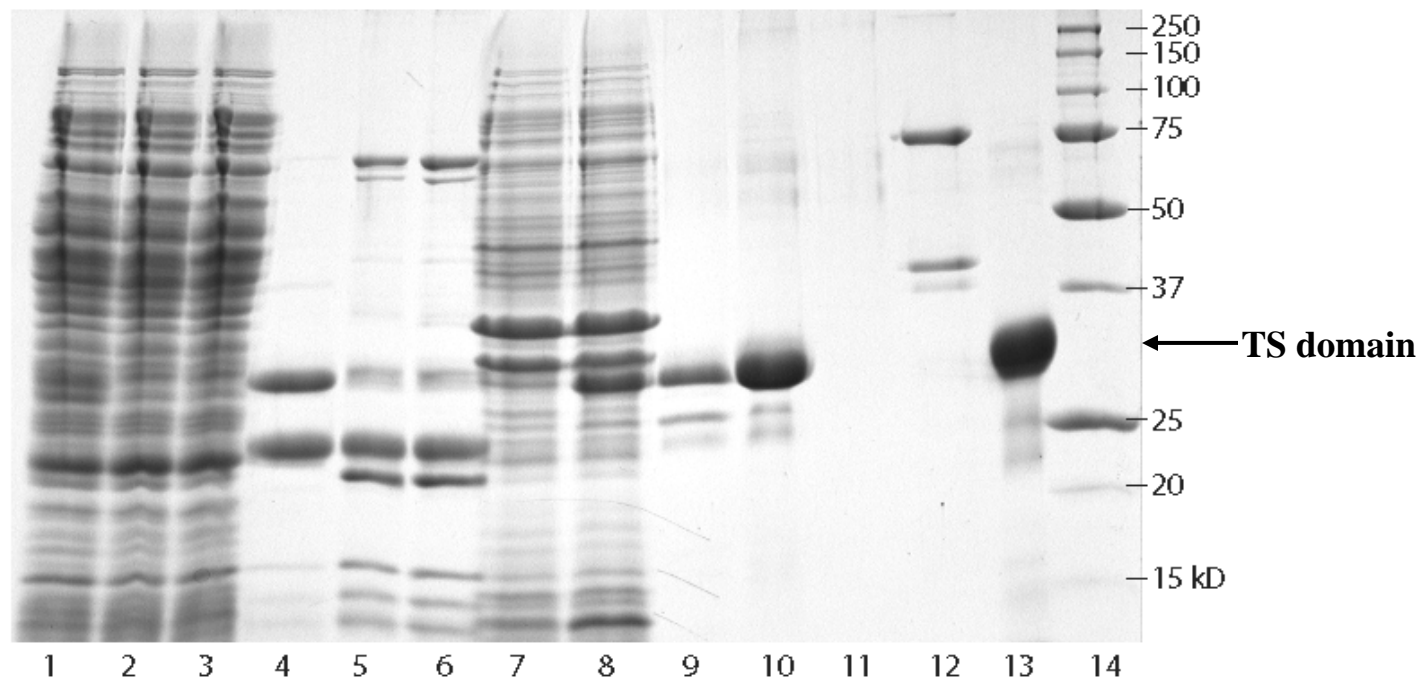
Tel: 507-284-5443

**Supplementary Figure 1. Sanger sequencing traces of c.1484A>G and c.1470-1472 TCC>ATA mutations.**



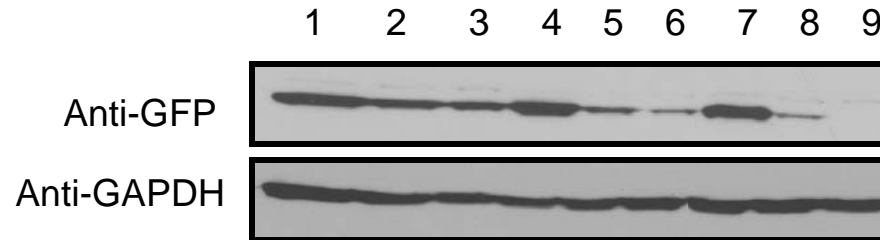
**Supplementary Figure 1.** Sanger sequencing traces of two heterozygous mutations are shown. Mutation **(a)** c.1484A>G, p.Tyr495Cys were found in affected persons from Kindred-1,-3 and -4. Mutation **(b)** c.1470-1472 TCC>ATA, p.Asp490Glu-Pro491Tyr were found in affected persons from Kindred-2.

**Supplementary Figure 2. SDS-PAGE and Coomassie blue staining of TS domain expressed in *E. coli*.**



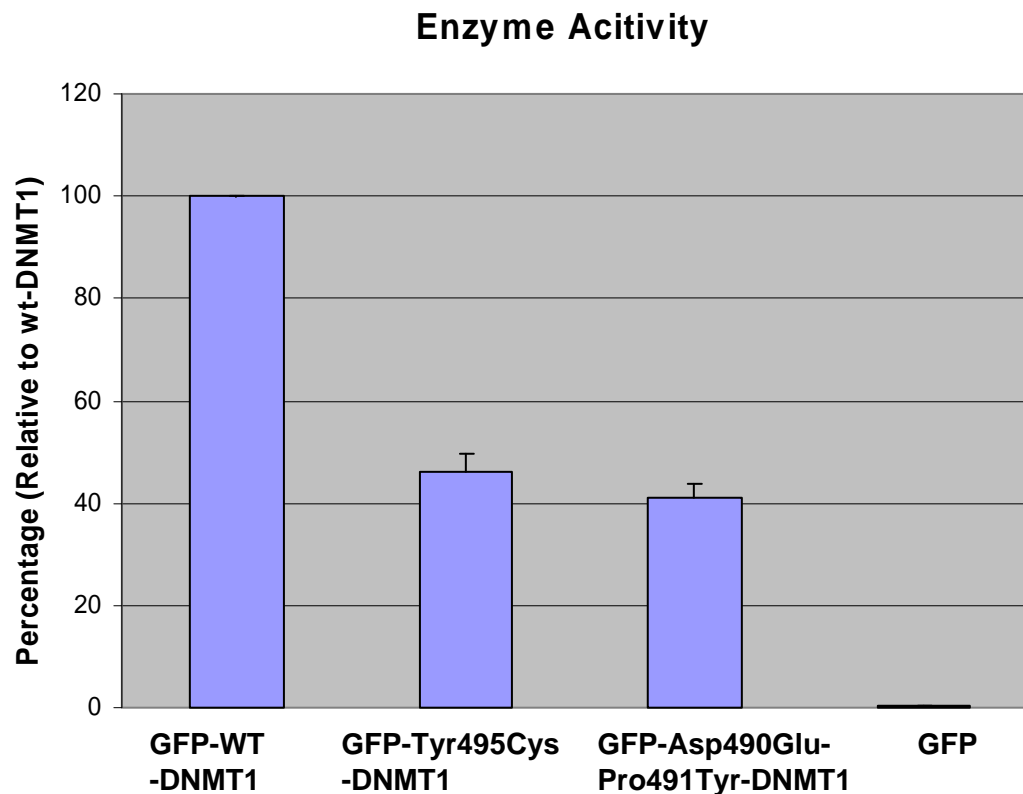
**Supplementary Figure 2.** Soluble fractions of wild type-TS, Y495C- and D490E-P491Y-TS domain before (**lanes 1-3**) and after (**lanes 4-6**) passing through Ni-NTA columns. Insoluble fractions of Y495C and D490E-P491Y TS domain before (**lanes 7, 8**) and after (**lanes 9, 10**) Ni-NTA purification. FPLC fractions (**lanes 11-13**) of Ni-NTA purified wild type TS domain. Standard molecular weight markers (BioRad) are shown in **lane 14**.

**Supplementary Figure 3. Degradation assay of GFP-wt-DNMT1, GFP-Tyr495Cys-DNMT1 and GFP-Asp490Glu-Pro491Tyr-DNMT1.**



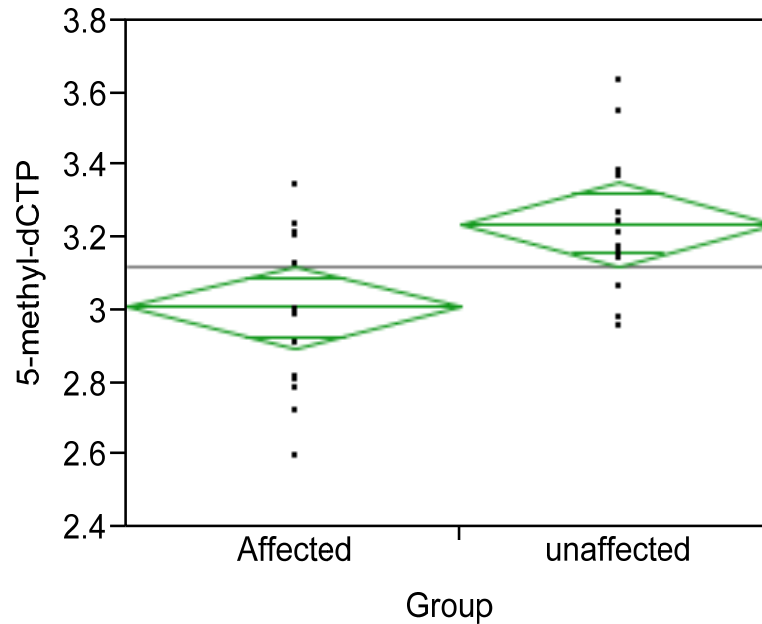
**Supplementary Figure 3.** HeLa cells were transfected with plasmids containing GFP-tagged full length wild type *DNMT1* (**1, 4, 7**), p.Tyr495Cys-*DNMT1* (**2, 5, 8**) or p.Asp490Glu-Pro491Tyr-*DNMT1* (**3, 6, 9**). After 24 hrs of transfection, cells were treated with 0.05mg/ml of cycloheximide and collected at 0hr (**1-3**), 2hr (**4-6**) and 4hr (**7-9**) time points. Fifty ug of cell lysates were analyzed on Western blot using monoclonal anti-GFP antibody. The blot was stripped and probed with GAPDH to ensure equal loading. The size of GFP-DNMT1 is ~220kD. Shown is one representative blot of three independent experiments.

**Supplementary Figure 4. Enzymatic activity of GFP-wt-DNMT1, GFP-Tyr495Cys-DNMT1 or GFP-Asp490Glu-Pro491Tyr-DNMT1**



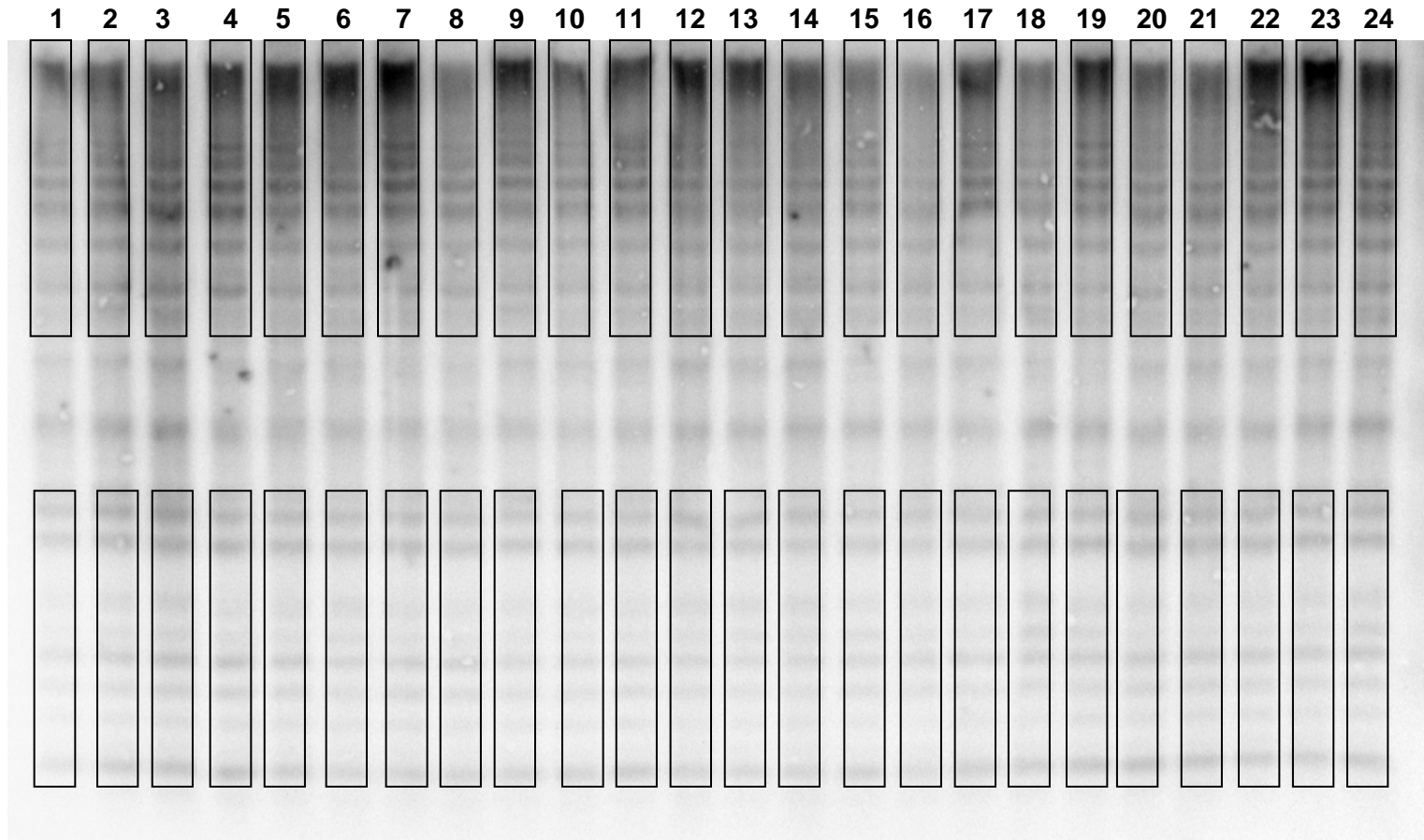
**Supplementary Figure 4.** Enzymatic activity assay measuring binding ability of hemimethylated trapping DNA substrate. Expressed GFP-full length wt-DNMT1, Tyr495Cys-DNMT1 or Asp490Glu-Pro491Tyr-DNMT1 were pull down by GBP-beads. Equal amount of proteins were incubated with 0.1uM of florescencely labeled trapping DNA substrate containing mechanism based inhibitor 5-aza-dCTP. After washing steps, the signals were read on fluorescence spectrometer. The relative binding ability is represented as a percentage comparing to the binding ability of wild type DNMT1. The plasmid contining GFP only was used as a negative control. Three independent experiments were performed. Error bars, means  $\pm$  s.d.

**Supplementary Figure 5. Comparison of genomic 5-mdC content in the affected and unaffected group.**



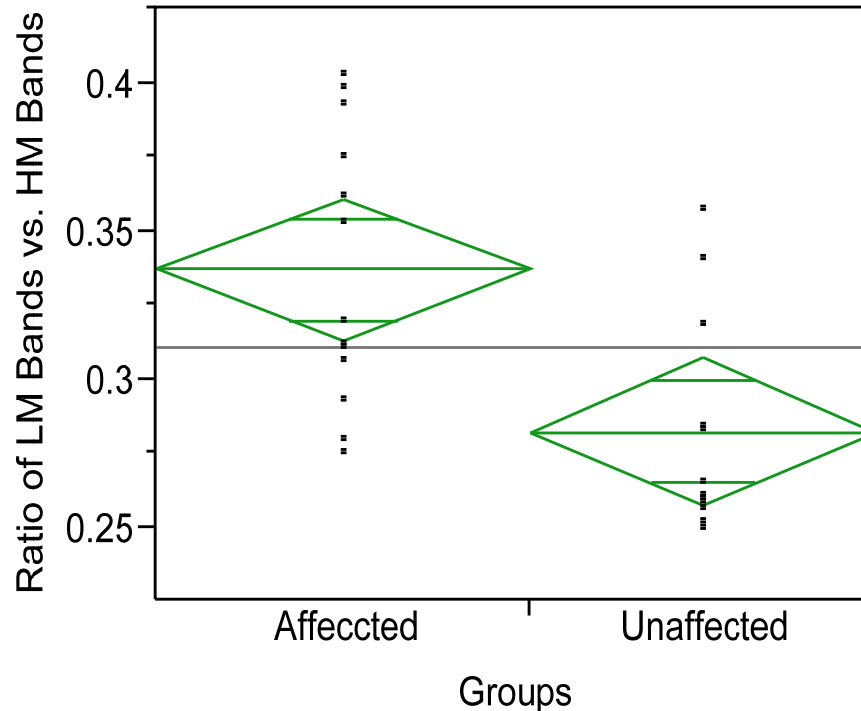
**Supplementary Figure 5.** Eight percent reduction of genomic 5-mdC content (represented as the percentage of total nucleotides) was observed in affected (n=12) group comparing to the unaffected group (n=12) ( $p < 0.001$ ). The 5-mdC content of genomic DNA was measured using liquid chromatography-electrospray ionization mass spectrometry (LC-ESI-MS/MS). All samples were analyzed in duplicate and proper controls were tested before and after sample measurement to ensure the accuracy. JMP 8.0 (SAS) was used to perform two tailed Wilcoxon rank-sum test. The image from JMP was shown here. The line across each green colored diamond represents the group mean. The vertical span of each diamond represents the 95% confidence interval for each group.

### Supplementary Figure 6a. Scanned Image of Southern blot



**Supplementary Figure 6a.** Genomic DNA (5ug) was digested with the CpG methylation-sensitive enzyme *BstB1* (NEB) overnight and separated on the 1% agarose gel. The digested DNA was then transferred to the blot. A satellite-2 consensus oligonucleotide (5'TCGAGTCCATTCGATGAT3') was end-labeled with  $^{32}\text{p}$  using T4-kinase (NEB) and hybridized to the blot. The blot was scanned by Storm phosphorimager and the signal was measured using ImageQuant software (Molecular Dynamics). Although no major shift was observed, the extent of hypomethylation was confirmed by calculating ratios of intensities of low molecular bands (lower boxed area) vs. high molecular weight bands (higher boxed area). All 24 samples are from kindred 1, from left to right, they are V-7,V-9, V-12, V-14, V-13,V-15,V-17,V-16,VI-3,VI-5,VI-11,VI-10,VI-12,VI-14,VI-30,VI-23,VI-24,VI-33,VI-26,VII-13,VII-15,VII-14,VII-18,VII-20.

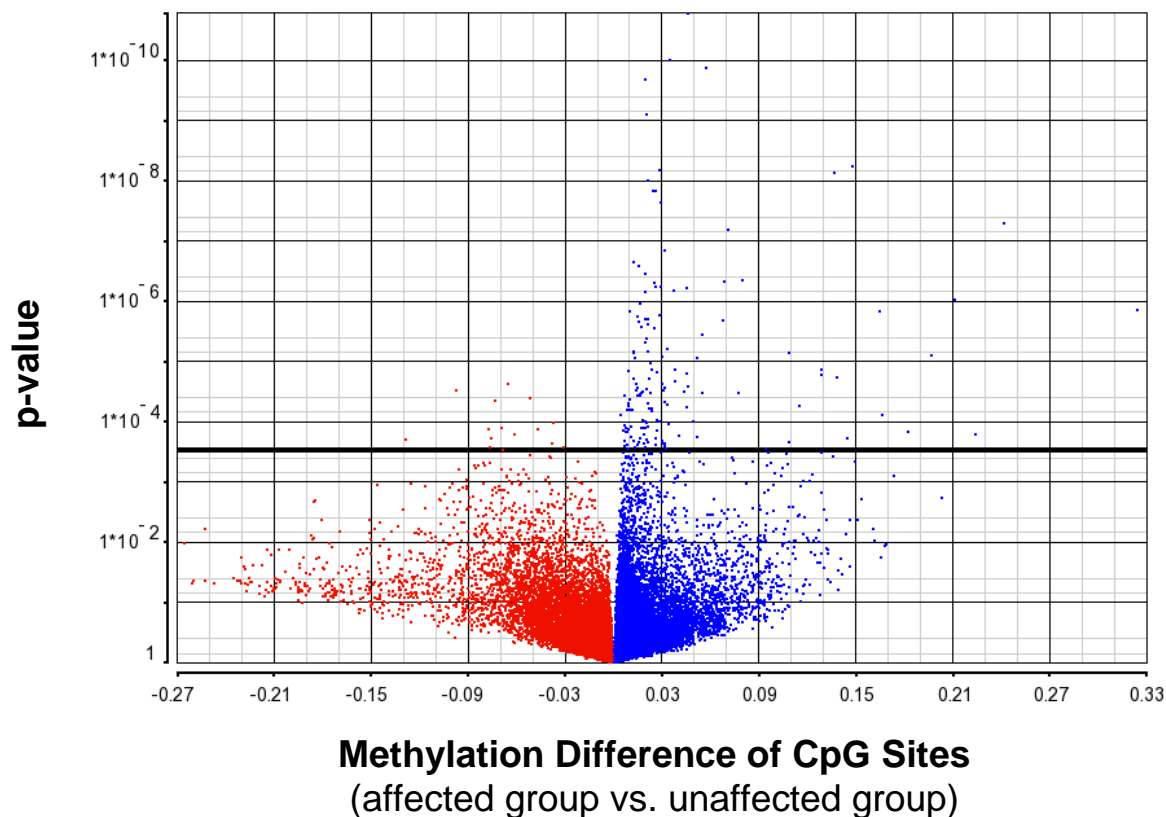
**Supplementary Figure 6b. Comparison of signal ratio of low molecular weight bands vs. high molecular weight bands in the affected vs. unaffected group.**



**Supplementary Figure 6b.** Comparison of signal ratio of low molecular weight vs. high molecular weight bands between the affected (n=12) and unaffected group (n=12) (p=0.004). JMP 8.0(SAS) was used to perform two tailed Wilcoxon rank-sum test. The image from JMP analysis was shown here. The line across each green colored diamond represents the group mean. The vertical span of each diamond represents the 95% confidence interval for each group.



**Supplementary Figure 7. Volcano plot of methylation difference between the affected group and unaffected group.**



**Supplementary Figure 7.** Volcano plot of methylation difference between affected group and unaffected group from Kindred-1. Y-axis represents the p-value for two-way ANOVA test. X-axis is the methylation difference between two groups for each CpG site. Black line indicates the significant p-value using FDR adjustment ( $p=0.000294$ ). Red colored dots represent methylation level is lower in affected group. Blue dots represent methylation level is higher in affected group. The overall pattern is slightly tilted toward left (red colored dots), but methylation differences with significant p-value (above the black line) are mostly increased methylation (blue colored dots), suggesting a moderate global hypomethylation profile with local hypermethylation in affected group.

### Supplementary Table 1. Summary of Exome Sequencing

<b>Sequencing Analysis Result within Linked loci at 19p13.2 from Illumina GAI and Roche GS FLX Titanium</b>	
Total Variants Identified (including known SNPs)	106
Synonymous	32
Nonsynonymous	26
Non-Genic or UTR	48
Novel Variants	4
<b>Novel Non-synonymous Heterozygous</b>	<b>1 (DNMT1 p.Tyr495&gt;Cys)</b>
Number of Uncovered CCDS exons	27 (directly sequenced and confirmed negative)

<b>Illumina GAI Sequencing Statistics</b> (Mapped by MAQ, Annotated by SIFT & SeattleSEQ)				
<b>Coverage of Targeted Region (37.6Mb)</b>	5x	10x	20x	30x
	coverage	coverage	coverage	coverage
	92.46%	86.45%	74.79%	63.80%
Number of mapped bases:	4,608,921,150			
Number of mapped reads:	61,452,282			
Read Length	75bp			
Total Variants Read	23046			
Coding Variants (known and novel)	Total			
Synonymous	7868			
Nonsynonymous	6784			
Novel Variants (coding& non-coding)	1453			

<b>Roche GS FLX Titanium Sequencing Statistics</b> (Analyzed by Roche GS Mapper )		
<b>Coverage of Targeted Region (30Mb)</b>	<b>69.63%</b>	Weighted Mean Coverage Depth
		10.0x
Number of mapped bases:	930,106,045	
Number of mapped reads:	2,435,028	
Mode Read Length	490bp	
Average Read length	382bp	
Total Variants Read	11677	
Known Coding Variants	Total	
Synonymous	5836	
Nonsynonymous	4740	
Novel Coding Variants	Total	
Synonymous	224	
Nonsynonymous	366	

**Supplementary Table 2. Comparison of 5-mdC content between the same sex siblings and cousins from Kindred-1**

<b>AFFECTED</b>	<b>UNAFFECTED</b>	<b>RELATIONSHIP</b>	<b>DIFFERENCE (IN PERCENTAGE) OF GLOBAL 5-mdC CONTENT</b>
V-7	V-9	sister	-14.6% ± 1.9%
VI-11	VI-10	sister	-14.2% ± 1.8%
VI-23	VI-26	brother	-14.1% ± 2.8%
VI-30	VI-26	brother	-7.7% ± 0.9%
VI-33	VI-26	same sex cousin	-14.1% ± 1.9%
VI-36	VI-26	same sex cousin	-12.4% ± 1.5%

**Supplementary Table 2.** Comparison of the genomic 5-mdC content between three pairs of same sex siblings and between two pairs of same sex cousins with similar age (<5 years difference) from Kindred-1. The results showed consistent 7.7-14.6% reduction of 5-methylcytosine content in all affected sibling and cousins.

**Supplementary Table 3. Pyrosequencing Methylation Assay of Repetitive Elements**

a. Sequence of the tested CpG sites:

<p><b>Satellite-2-CpG sites:</b></p> <p>AATGGAATTATTATCGAATGGAATCGAATGGTATTAAACGGAAAAAACGGAAT  TAT <b>CG(1)</b>AATGGAAT <b>CG(2)</b>AAGAGAATTTT <b>CG(3)</b>AACGGATT <b>CG(4)</b>AATGGAATTA  TTAATGGAATGGAATGGAATAATTTATTGGATTTCGAATGTAATTATT</p>
<p><b>Line-1-CpG sites:</b></p> <p>CAGGGAGTTCCTTTCCGAGTCAAAGAAAGGGGTGA <b>CG(1)</b>GA <b>CG(2)</b>CACCTGGAA  AAT <b>CG(3)</b>GGTCTCTCCACC <b>CG(4)</b>AATATTG <b>CG(5)</b>CTTT <b>CG(6)</b>GAC <b>CG(7)</b>GCTTAAA  AAA <b>CG(8)</b>G <b>CG(9)</b>CAC <b>CG(10)</b><b>CG(11)</b>AGATTATATCTTGCACCTGGCTAGGAGGGTC  CTACGCCACGGAGTCTCGCT</p>
<p><b>Satellite-α-CpG sites:</b></p> <p>ACTTCTTTGTGATGTTGACATTCAACTGACAGAGGTGAACCTTCCCTTGTGAGTTC  AGGTTGAAA <b>CG(1)</b>CTCCTTT <b>CG(2)</b>TAGCATCTGCAAGTGGAGATTTGGAA <b>CG(3)</b>CT  ATGAGGCCTA <b>CG(4)</b>GTAGTAAAGGAAACAGCTTCATGTAAAACTGGACAGA</p>
<p><b>Alu-CpG sites:</b></p> <p>AGCAGCAGATCCCATTTAGGGAGCACCAACTTCATGCCAGGCCCCACATGAAGAT  CCCATAATCTCCCCAGTGACCTGT <b>CG(1)</b>AGGGGA <b>CG(2)</b>TACTCTCACATACTCCA  TTATAAAGAGGAGAAAACCTGAGGTCCCAAAGGGATGGAA</p>

b. Pyrosequencing Primers

Repetitive Elements	GenBank Accession Number	Primer Sequence
Line-1	Line-1-PCR forward	GGGAAGAGTAAGGGGTTAGGGA
	Line-1-PCR reverse	CCCTCCCCAACCTTACTAC
	Line-1-sequencing	AGTTAAAGAAAGGGGTG
Sat-α	Sat-a-PCR forward	M38468 GTGGATATTGGGATTTTTTTGAGAATT
	Sat-a-PCR reverse	CCACCCAAAAAATATTCAACTCTATAA
	Sat-a-sequencing	TTTTTTTTGTGAGTTTAGGT
Alu	Alu-PCR forward	AGTAGTAGATTTTATTTAGGGAGTATTAA
	Alu-PCR reverse	TTCCATCCCTTTAAACCTCAATTTTC
	Alu-sequencing	GAAGATTTTATTAATTTTTTTAGTG
Satellite 2	Sat2-PCR forward	X72623 TTATTGAATGGAAATGAAAGGGGTTAT
	Sat2-PCR forward	CCAATAAATTATTCCATTCCATTCCATTAA
	Sat2-sequencing	TTTTAATGGAAAGGAATGG

c. Assay Results Summary

Repetitive elements	Case Group		Control Group		Difference between affected and unaffected group	p-Value
	Mean methylation level (%)	S.E.	Mean Methylation level	S.E.		
Sat-2-CpG-1	63.2	1.2	69.0	1.2	-5.8	0.009*
Sat-2-CpG-2	23.6	1.1	25.4	1.1	-1.8	0.50
Sat-2-CpG-3	63.6	0.7	68.9	0.8	-5.3	0.02*
Sat-2-CpG-4	81.0	0.9	85.6	0.9	-4.6	0.02*
Line-1-CpG-1	38.7	0.7	39.8	0.8	-1.1	0.33
Line-1-CpG-2	57.5	0.5	57.7	0.5	-0.2	0.75
Line-1-CpG-3	72.9	0.9	75.4	0.9	-2.5	0.06
Line-1-CpG-4	10.2	0.5	11.0	0.5	-0.8	0.32
Line-1-CpG-5	48.9	1.0	50.4	1.1	-1.5	0.26
Line-1-CpG-6	32.5	2.8	32.8	3.2	-0.3	0.57
Line-1-CpG-7	10.1	0.5	12.1	0.6	-2.0	0.05
Line-1-CpG-8	14.7	1.2	17.2	1.4	-2.5	0.18
Line-1-CpG-9	23.1	0.9	22.4	1.0	+0.7	0.45
Line-1-CpG-10	48.6	1.0	51.0	1.2	-2.4	0.42
Line-1-CpG-11	17.4	0.9	18.9	1.0	-1.5	0.08
Sat- $\alpha$ -CpG-1	67.8	1.7	68.6	1.8	-0.8	0.62
Sat- $\alpha$ -CpG-2	97.1	1.7	98.9	1.9	-1.8	0.98
Sat- $\alpha$ -CpG-3	83.5	2.5	85.4	2.6	-1.9	0.35
Sat- $\alpha$ -CpG-4	67.7	2.9	69.6	2.8	-1.9	0.25
Alu-CpG-1	92.0	1.8	90.6	1.9	+1.4	0.35
Alu-CpG-2	87.9	2.1	87.7	2.3	+0.2	0.32

**Supplementary Table 4. Sequencing Primers for DNMT1 Exons**

dNMT1Exon1f	ATCCCCATCACACCTGAAAG
dNMT1Exon1r	GCCCGTCTGTCAGCAGC
dNMT1Exon2f	AAACTTGTTTGTGTCCAAAACCTTC
dNMT1Exon2r	TGCAAAATCCATTTAAAGAAAAC
dNMT1Exon3f	TTGACTGAGACAGCATTGCC
dNMT1Exon3r	AGGACAGCTGGGGATCTTG
dNMT1Exon4-5f	TTGACTGAGACAGCATTGCC
dNMT1Exon4-5r	GCTTCTAGGTTGAAAATGAGCC
dNMT1Exon6f	AACCTGGCTTCCTGCAATAG
dNMT1Exon6r	ACGGAAGACAGAATTGCCAC
dNMT1Exon7f	GATCGGTGTTTAGACCCGTG
dNMT1Exon7r	TTCTCAAATAAAGTCTCTCCCC
dNMT1Exon8f	CTGCCTACTTCTCGGACCTG
dNMT1Exon8r	TTGCCATGGAAACACACG
dNMT1Exon9f	AATGAAGATTGCTTTTGGGG
dNMT1Exon9r	ACCCATCTTGTTCTTCCACG
dNMT1Exon10f	TCTTTACTCCCACCACTGGAC
dNMT1Exon10r	TAGTGCCCACTGTTCCACAC
dNMT1Exon11f	CTCTGAACCTGGGGAGGAG
dNMT1Exon11r	GATGGGGCTGGACTTGAAC
dNMT1Exon12f	GGAAAGTGAAGCCTCGTGTC
dNMT1Exon12r	CAGGGTCCCCACACATCTAC
dNMT1Exon13f	CATGGGGAGGCATTAGTTTG
dNMT1Exon13r	CATGTGATTCACCCGCTTC
dNMT1Exon14-15f	CAGAGCTCACCTGCTGGC
dNMT1Exon14-15r	GTCTGTGGGAGCAGGAACAC
dNMT1Exon16-17f	AAGGTAAACATCTGCCGGG
dNMT1Exon16-17r	TGGCTCTTATCCACGAAGTG
dNMT1Exon18f	GCTCCAGCCTGAGGGATAG
dNMT1Exon18r	TTTTAATAGAGGCAGGCTCTTG
dNMT1Exon19f	GACTGAGGAGCACCCGAC
dNMT1Exon19r	ACATGGCCTTCTGCAAGC
dNMT1Exon20-21f	ACATTTGGGTACGGGATGAC
dNMT1Exon20-21r	CTACGGGAGAGGTTCCAGC
dNMT1Exon22f	TGGGACAGAGGTAAGGATGC
dNMT1Exon22r	TGAGCAGCCAGAGTCTCAAG
dNMT1Exon23-24f	TGGAGTTTTACTCTTGTCGTCC
dNMT1Exon23-24r	CTCTTCTCAGGGGCAAACAG
dNMT1Exon25f	ATTGACTTCTTAATGAAATCGAGTC
dNMT1Exon25r	CCTTTTCAGTTTTTCATCTAGGGC
dNMT1Exon26f	CTGAAAAGGACGAGTGCTCC
dNMT1Exon26r	TGCCTCCCTTGGGAGATAAG
dNMT1Exon27f	ACTGCTGACGTGCGTTCTG
dNMT1Exon27r	GCCTTTGACGAGCAAGAGAC
dNMT1Exon28f	GCACAAGAAAGCCACCTCTTC
dNMT1Exon28r	TTGGCAATGTCTGTAAGGAGG

dNMT1Exon29f	CCTTAATTCATCAGGTGCTTGAC
dNMT1Exon29r	AACAACGTGGGTGCTATGC
dNMT1Exon30f	TTACCCTGCAGTTCCTGAG
dNMT1Exon30r	GCCAACCACCCACTTCTTAC
dNMT1Exon31-32f	CCCCACTGAGGGAGAAATTAAG
dNMT1Exon31-32r	GCCTCGGAAGGAGATTCTTG
dNMT1-E33f	GAACCTGGGAGGCAGAGC
dNMT1-E33r	AATGACCACTGCTGACATGC
dNMT1Exon34f	CACAGCTCAGCTCTCACCAG
dNMT1Exon34r	AGGCCTATGCCATTGAACC
dNMT1Exon35f	CGACTCAGGCTGCTGACC
dNMT1Exon35r	AGAGTGCCATGTGGCAGAG
dNMT1-E36fa	TGCTGTGATCTCGGGAGAAG
dNMT1-E36ra	CAGATTCCATGTCTCCCCTG
dNMT1Exon37f	GACTGTCTCTTTCCCGATGG
dNMT1Exon37r	AGGCTTGGTGTGTCTGTGC
dNMT1Exon38f	AGAACGAGGATTGTTGGCTG
dNMT1Exon38r	GTCCTGGGGTGCTGTCC
dNMT1Exon39f	GGACAAGCTCATAGCCAAG
dNMT1Exon39r	CTG AGA GTG ATG GGG CTA CC
dNMT1Exon40f	CCAGGTTGCCTCCATCT G
dNMT1Exon40r	ACTCAATCCTCACAGCAGCC
dNMT1Exon41f	TCCCTTAGCACTCTGCCAC
dNMT1Exon41r	TGCTAGCTTCAAACCTCCACG

**Supplementary Table 5. Primers for site-directed mutagenesis**

<b>Primers for site-directed mutagenesis</b>	
DNMT1_Y495C_f	CCCAGTCCCGAGTGTGCGCCCATATTTGG
DNMT1_Y495C_r	CCAAATATGGGCGCACACTCGGGACTGGG
DNMT1_D490E-P491Y_f	GCCGAATACATTCTGATGGAATACAGTCCCGAGTATGCGCC
DNMT1_D490E-P491Y_r	GGCGCATACTCGGGACTGTATTCCATCAGAATGTATTCGGC

**Supplementary Table 6. Primers for preparing trapping DNA substrate**

<b>Primers for preparation of double stranded trapping DNA substrate (M-5-methycytocine)</b>	
MG-Upper primer	5'-CTCAACAATACTACCATCMGGACCAGAAGAGTCATCATGG-3'
Fill-In-primer	5'-MAXN-CCATGATGACTCTTCTGGTC-3'



## Supplementary Notes

### a. Pull-down of GFP-tagged Proteins

HeLa Cells were transiently transfected with expression plasmids as described above. Twenty-four hours after transfection, about  $1 \times 10^7$  cells were lysed with 200  $\mu$ l of NP-40 lysis buffer (20 mM Tris-HCl pH 7.5, 150 mM or 1M NaCl, 0.5 mM EDTA, 2 mM PMSF, 0.5% NP40, 1 $\times$  protease inhibitor cocktail) for 30 min on ice. The cell debris was cleared by centrifugation for 10 min at 20,000g in 4°C. Supernatants were collected and mixed with 500  $\mu$ l dilution buffer (lysis buffer without NP40), then incubated with 1  $\mu$ g GFP-binding protein conjugated beads (Allele Biotechnology) for 2 h at 4°C with end-to-end mixing. Beads were harvested and washed twice with 1mL of dilution buffer containing 500 mM NaCl and resuspended in the assay buffer. Equal amount of pull-down proteins measured by DNMT1 ELISA assay (Epigentak) were used in the binding assay after resuspending beads in 100  $\mu$ l assay buffer (100 mM KCl, 10 mM Tris-HCl pH 7.6, 1 mM EDTA, 1 mM DTT).

### b. Enzyme Activity Assay (DNMT1 binding assay to trapping DNA substrate)

The methyltransferase activity was measured using a high sensitivity and specificity assay by measuring the binding ability to its trapping DNA substrate that contains the mechanism-based inhibitor 5-aza-dC<sup>35</sup>. The DNMT1 trapping DNA substrate was prepared as described previously<sup>35</sup>. Briefly, DNA oligonucleotides were synthesized (IDT) and the sequences are listed in **Supplementary Table 6**. Double-stranded DNA substrates were synthesized by primer extension using the Klenow fragment (Roche), dTTP, dGTP, dATP (Applied Biosystems) at final concentration of 1mM and 50 $\mu$ M 5-aza-dCTP (Jena Bioscience, Germany). The Fill-In primer was labeled with MAX-NHS-ester and extended to produce hemimethylated DNA trapping substrates containing 5-aza-dC at the CpG site. The pull-down beads with GFP tagged proteins were equilibrated with 200 $\mu$ l of assay buffer (100 mM KCl, 10 mM Tris-HCl pH 7.6, 1mM EDTA, 1 mM DTT) supplemented with 160ng/ $\mu$ l BSA, 100 $\mu$ M S-adenosyl-L-methionine (AdoMet), and 0.1  $\mu$ M trapping DNA substrate. For qualitative determination of DNA methyltransferase activity, trapping were performed at 37°C for 90 min with constant mixing. After washing twice with 1ml assay buffer to remove unbound substrate, beads were resuspended in 100  $\mu$ l assay buffer and transferred into a 96-well microplate. The

binding of trapping DNA substrate was determined by fluorescence spectrometer and results were converted to percentage relative to wild type DNMT1 binding ability. The enzymatic activity was presented as percentage relative to the activity of wild type DNMT1.

#### c. Pyrosequencing n Analysis

One microgram of genomic DNA was converted with sodium bisulfate which converts unmethylated cytosine into uracil and PCR amplified as thymidine (Zymo). After bisulfite treatment and PCR, the degree of each methylation at each CpG position in a sequence is determined using the ratio of T and C. Primers were designed using Pyrosequencing Assay Design Software (Qiagen) for Sat-2, Line-1, Alu and Sat-a. Amplification was carried out on 20 ng of bisulfate treated DNA using TaqGold DNA polymerase (Applied Biosystems) under the following conditions: 10 min at 95°C, followed by 45 cycles of 35 sec at 95°C, 35 sec at 60°C, and 35 sec at 72°C, then 5 minutes at 72°C. The PCR products were checked by gel-electrophoresis to confirm the specificity and size of the band. Pyrosequencing reactions were performed on Biotage PyroMark MD System (Qiagen) according to manufacturer's protocols. Raw data were analyzed using the Pyro Q-CpG 1.0.9 analysis software (Qiagen).