6054–6064 *Nucleic Acids Research, 2010, Vol. 38, No. 18* *Published online 27 May 2010*
doi:10.1093/nar/gkq456

A single amino acid substitution confers enhanced methylation activity of mammalian Dnmt3b on chromatin DNA

Li Shen[1,2], Ge Gao[3,*], Ying Zhang[1,2], He Zhang[3], Zhiqiang Ye[3], Shichao Huang[1,2], Jinyan Huang[4] and Jiuhong Kang[1,2,4,*]

[1]Laboratory of Molecular Cell Biology and Center of Cell Signaling, Institute of Biochemistry and Cell Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, [2]Graduate School of the Chinese Academy of Sciences, Shanghai 200031, [3]Center for Bioinformatics, National Laboratory of Protein Engineering and Plant Genetic Engineering, College of Life Sciences, Peking University, Beijing 100871 and [4]Key Laboratory of Signaling and Disease Research, School of Life Science and Technology, Tongji University, Shanghai 200092, P.R. China

Received January 1, 2010; Revised May 6, 2010; Accepted May 10, 2010

ABSTRACT

Dnmt3a and Dnmt3b are paralogous enzymes responsible for *de novo* DNA methylation but with distinguished biological functions. In mice, disruption of Dnmt3b but not Dnmt3a causes global DNA hypomethylation, especially in repetitive sequences, which comprise the large majority of methylated DNA in the genome. By measuring DNA methylation activity of Dnmt3a and Dnmt3b homologues from five species, we found that mammalian Dnmt3b possessed significantly higher methylation activity on chromatin DNA than Dnmt3a and non-mammalian Dnmt3b. Sequence comparison and mutagenesis experiments identified a single amino acid substitution (I662N) in mammalian Dnmt3b as being crucial for its high chromatin DNA methylation activity. Further mechanistic studies demonstrated this substitution markedly enhanced the binding of Dnmt3b to nucleosomes and hence increased the chromatin DNA methylation activity. Moreover, this substitution was crucial for Dnmt3b to efficiently methylate repetitive sequences, which increased dramatically in mammalian genomes. Consistent with our observation that Dnmt3b evolved more rapidly than Dnmt3a during the emergence of mammals, these results demonstrated that the I662N substitution in mammalian Dnmt3b conferred enhanced chromatin DNA methylation activity and contributed to functional adaptation in the epigenetic system.

INTRODUCTION

DNA cytosine methylation is an important epigenetic modification crucially involved in many cellular processes, such as transcription regulation and chromatin organization (1). Aberrant DNA methylation contributes to carcinogenesis and other diseases (2–3). Two types of DNA methyltransferases (DNMTs) are mainly responsible for this modification (4). The maintenance DNMT preferentially recognizes hemimethylated DNA and methylates the newly synthesized strand after DNA replication, while *de novo* DNMTs create new methylation patterns on non-methylated DNA.

In mammals, two *de novo* DNMTs, namely Dnmt3a and Dnmt3b, are known for the establishment of initial chromatin DNA methylation patterns (4,5). Recombinant Dnmt3a and Dnmt3b have similar DNA methylation activity *in vitro* when poly(dG-dC)-poly(dG-dC) or other naked DNA is used as the substrate (6), and have little or no intrinsic sequence specificity beyond the CpG dinucleotide (4). However, the methylation activity of both Dnmt3a and Dnmt3b substantially decreases when the DNA substrate is packaged into nucleosomes, the basic unit of chromatin in eukaryotic cells (7–9), and Dnmt3b was reported to have higher nucleosomal DNA methylation activity than Dnmt3a (8). Moreover, although Dnmt3a and Dnmt3b have overlapped targets *in vivo*, Dnmt3a is specifically required for the methylation of some single copy genes, such as imprinted genes in germ cells and the *Xist* gene on the X chromosome (10–12), while Dnmt3b is irreplaceable in the methylation of minor satellite repeats and other different repetitive sequences depending on the cell types (5,11,13,14).

*To whom correspondence should be addressed. Tel: +86 21 5492 1373; Fax: +86 21 5492 1372; Email: jhkang@sibs.ac.cn
Correspondence may also be addressed to Ge Gao. Tel/Fax: +86 10 62751861; Email: gaog@mail.cbi.pku.edu.cn

Notably, repetitive sequences have increased dramatically in the mammalian genomes (15,16), and these repetitive sequences represent a large part of the genome and contain the large majority (>90%) of 5-methylcytosines (m^5C) in the genome (17). Consistently, mutations in the *DNMT3B* gene cause ICF (immunodeficiency, centromeric instability, facial anomalies) syndrome, which shows hypomethylation of satellite DNA and chromatin instability (14,18,19). Disruption of Dnmt3b in mice led to global DNA hypomethylation and embryonic lethality, while Dnmt3a deficient mice developed to term without significant changes in global DNA methylation although died at about 4 weeks of age (5).

In contrast with their distinguished *in vivo* functions, Dnmt3a and Dnmt3b share remarkable sequence homology and both have three conserved domains: the N-terminal regulatory region contains a PWWP domain that targets the enzyme to chromatin, a cysteine-rich PHD zinc-finger domain which interacts with transcriptional repressors, histone deacetylases and histone methyltransferases generating inactive chromatin markers; the C-terminal region is the catalytic domain which is highly conserved (4,20–22). The similar domain organization and remarkable sequence homology of Dnmt3a and Dnmt3b suggest that they are paralogs generated from duplication in a common ancestor. Gene duplication is widely recognized as one of the primary sources for evolutionary novelties, and theoretical models predict that after the duplication event, one copy of the duplicated genes would be silenced quickly unless functional novelty or partitioning had occurred by sequence divergence (23–26). The distinguished biological functions of Dnmt3a and Dnmt3b are consistent with the theoretical models. However, the contribution of evolutionary sequence divergence of Dnmt3a and Dnmt3b to their *in vivo* DNA methylation activity and biological functions after gene duplication remains largely untouched.

In this study, we found that mammalian Dnmt3b had higher chromatin DNA methylation activity than Dnmt3a and non-mammalian Dnmt3b by using budding yeast as the *in vivo* test tube, and a single amino acid substitution (I662N) of mammalian Dnmt3b in the highly conserved catalytic domain conferred the high chromatin DNA methylation activity. Moreover, this substitution was crucial for Dnmt3b to efficiently methylate repetitive sequences, which increased dramatically in mammalian genomes and are the major targets of DNA methylation. We also observed that Dnmt3b evolved more rapidly than Dnmt3a in vertebrates, during the emergence of mammals. These results provide the first example on how sequence divergence after gene duplication contributes to functional novelty and adaptation in the epigenetic system.

MATERIALS AND METHODS

In silico analysis

Protein sequences of Dnmt3 homologues were aligned using T-Coffee (27). The phylogenetic tree was inferred using NJ method with JTT model in MEGA (28). Synonymous and non-synonymous substitution rates between human Dnmt3a/Dnmt3b and their orthologous genes in other species were further derived by NG model (29), and molecular clock hypothesis test was performed as proposed by Tajima (30).

Cloning of anole lizard Dnmt3a and Dnmt3b

Total RNA of anole lizard tissue was extracted with Trizol Reagent (Invitrogen), and reverse-transcribed using SMART RACE cDNA Amplification Kit (Takara). Then 5′-RACE and 3′-RACE PCR were performed according to the manufacturer's instructions. For Dnmt3a, the gene-specific primers (GSPs) used were: 5′-TCCCGTC CCCGAACCACATGACCCA-3′ (5′-RACE) and 5′-GCAAGGACCAGCACTTCCCTGTCTTCAT-3′ (3′-RACE). For Dnmt3b, the GSPs used were: 5′-GCACC TGAAACAACTGGCATTGTCACAAAGGA-3′ (5′-RACE) and 5′-TGAACGGCAAAGACGACGACTT GTGGTGC-3′ (3′-RACE). The PCR products were cloned into pMD19-T vector (Takara) and sequenced, which have been submitted to GeneBank (GQ121004 and GQ121005).

Plasmids

Full length cDNAs of Dnmt3a/Dnmt3b of human, mouse, chicken and zebrafish were kindly provided by Dr Arthur D. Riggs, Dr Guo-liang Xu, Dr Hiroyuki Sasaki, Dr Nobuyoshi Shimoda, respectively. For yeast transformation, full lengths of DNMTs were cloned into modified pYES3/CT in-frame with a FLAG tag at the N-terminus. Plasmids containing cDNA encoding mouse Dnmt3a and Dnmt3b chimeric proteins were generated similar to previously described (12). Single amino acid mutants of mouse Dnmt3a and Dnmt3b were generated by QuickChange Site Directed Mutagenesis (Stratagene). For Dnmt3L co-expression, pYES3/CL (modified from pYES3/CT) was used. All constructs were confirmed by sequencing. For ES cell transfection, N-terminal FLAG-tagged full lengths of mouse Dnmt3b and its N662I mutant were subcloned into pCAG-IRESblast expression vector (11; kindly provided by Dr Taiping Chen).

Yeast methods

Wild-type and Set1 knockout YPH499 yeast (*Saccharomyces cerevisiae*) strains were kind gifts from Dr Guoliang Xu. Yeast was transformed using a standard lithium acetate heat-shock procedure. Transformants were selected on plates of appropriate selective medium (SC) with 2% glucose. To induce expression of the Dnmt3 proteins, single colonies were inoculated into appropriate SC + 2% glucose liquid medium and grew overnight at 30°C with shaking. Then yeast cells were spun down, washed with sterilized water and resuspended into induction medium (SC + 2% galactose) to obtain an OD_{600} of 1.0, and continued growing for 24 h. After the 24-h induction, yeast cells were harvested for further analysis or protein purification.

HPLC analysis

Genomic DNA was prepared from yeast or ES cells, digested and analyzed essentially as described (31) using a Perkin-Elmer HPLC with an UV/VIS detector. Deoxyribonucleotide 5′-monophosphates in DNA digests were resolved using an octadecylsilane-bonded silica column (Ultimate AQ-C18, 5μm, 4.6 × 250 mm, Welch Materials) eluted at 20°C with 10 mM potassium orthophosphate (pH 3.7) at 1.0 ml/min.

Protein purification

DNMT-expressing yeast cells were lysed with a Mini-Beadbeater (for small scale) or a Microfluidizer (for large scale) in cold Lysis Buffer (50 mM Tris–HCl, pH 7.4, 500 mM NaCl, 2 mM $MgCl_2$, 10% glycerol, 0.5% NP-40, 1 mM EDTA, 0.5 mM dithiothreitol) containing freshly added complete protease inhibitor tablets (Roche) and 0.5 mM phenylmethylsulfonyl fluoride. Lysates were pre-cleared by centrifugation for 30 min at 20 000*g* and cleared a second time for 1 h at 92 000*g*. The soluble extract (5–10 mg/ml protein) was incubated with pre-equilibrated anti-FLAG M2 Affinity Gel (Sigma) at 4°C for 2 h. After washing the beads for three times with lysis buffer, the DNMT proteins were eluted by two sequential 30-min incubations at 4°C with elution buffer (20 mM Tris–HCl, pH 7.4, 200 mM KCl, 0.2 mM dithiothreitol, 200 μg/ml 3× FLAG peptide). Purified DNMT proteins were electrophoresed in 8% SDS–PAGE gel, stained with Coomassie blue and quantitated using BSA as a protein standard.

Preparation of mono-nucleosomes and naked DNA

Mono-nucleosomes were prepared from HeLa cells as previously described (8). The corresponding naked DNA was extracted from the mono-nucleosomes by phenol/chloroform method.

In vitro DNMT assay

DNMT activity was assayed essentially as previously described (8). Briefly, the immuneprecipitates (from lysates of 10 OD Dnmt3-expressing yeast cells) or 100 ng purified Dnmt3 proteins, the indicated substrate [500 ng for yeast genomic DNA, 100 ng for poly(dG-dC)-poly(dG-dC), 75 ng for nucleosomal DNA or naked DNA] and 2 μl [methyl-^{3}H]AdoMet (81.9 Ci/mmol, Perkin-Elmer) were incubated in a total volume of 25 μl reaction buffer (20 mM Tris–HCl, pH 7.4, 20 mM KCl, 60 ng/μl BSA, 2.7 M glycerol, 1 mM EDTA and 0.2 mM DTT) at 37°C for 2 h (immunoprecipitates) or 1 h (purified Dnmt3 proteins). After the incubation, the radioactivity ($[^3H]$-CH_3) incorporated into DNA was determined with a liquid scintillation counter.

Electrophoretic mobility shift assay

Electrophoretic mobility shift assay (EMSA) was performed essentially as previously described (32). Briefly, 30 nM mono-nucleosomes or naked DNA extracted from the mono-nucleosomes were incubated with 120 nM purified proteins in reaction buffer (20 mM Tris–HCl, pH 7.4, 20 mM KCl, 60 ng/μl BSA, 1 mM EDTA and 0.2 mM DTT) containing 0.2 mM sinefungin for 30 min and applied to 4% polyacrylamide gel run in 0.5× TBE buffer. Then the gel was stained with SYTO 60 fluorescent nucleic acid stain (Invitrogen) and scanned with Odyssey infrared imager (LI-COR).

ES cell culture and stable expression

Wild-type J1 and *Dnmt3a*$^{-/-}$*3b*$^{-/-}$ mutant ES cells (kindly provided by Dr En Li) were maintained on 0.1% gelatin (Sigma) in Dulbecco's Modified Eagle's medium (Invitrogen) supplemented with 15% fetal bovine serum (HyClone), 0.1 mM non-essential amino acids (Invitrogen), 0.1 mM β-mercaptoethanol, 1000 U of leukemia inhibitory factor (LIF) per milliliter. For stable expression, pCAG-IRESblast expression vector encoding GFP, mDnmt3b and mDnmt3b (N662I) were electroporated into *Dnmt3a*$^{-/-}$*3b*$^{-/-}$ ES cells of passage 70, and subsequently selected in blasticidin-containing (10 μg/ml, Sigma) medium for 7 days. Individual blasticidin-resistant colonies were examined for protein expression by western blot with polyclonal anti-Dnmt3b antibody (kindly provided by Dr Guo-liang Xu).

Bisulfite genomic sequencing

Genomic DNA (200 ng) was restricted with EcoRI and treated with sodium bisulfite as previously described (33). The treated DNA was subjected to PCR amplification for regions of interest, and PCR products were cloned into pMD19-T vector (Takara) and sequenced. Primer sequences for bisulfite sequencing were designed using Methyl Primer Express software. For yeast genomic region 1 (chr7: 155 658–156 092): 5′-TGAAAAAATTTA GTAGTTTTTGGG-3′ and 5′-ACCACACATCCCATA ACTCTAT-3′; for yeast genomic region 2 (chr15: 1 085 279–1 085 671): 5′-TTGTTGTTGTTGGAGATT TGTA-3′ and 5′-CCTATCCATTACACCAATTCTAA A-3′; for mouse *Pgk-2*: 5′-GTTGATGAGTTTAAGTTT TTGTTG-3′ and 5′-ATCCAATTCCTTCTTCAT AAAA-3′; for mouse *β-globin:* 5′-TTTTGATTTGTTTT TTTTTTT-3′ and 5′-TACATTCAAAAAAAACCA TATT-3′.

Southern blot

Genomic DNA was extracted from various ES cell lines by the classical phenol/chloroform method. A total of 1.5 μg of DNA was digested with the indicated methylation-sensitive or -insensitive enzymes (New England Biolabs) overnight and subsequently run on a 1% agarose gel. Then the gel was transferred to Hybond-N+ membrane (GE Healthcare) and hybridized to different probes. The probes used in this study were as follows: major satellite repeats (5′-TAT GGC GAG GAA AAC TGA AAA AGG TGG AAA ATT TAG AAA TGT CCA CTG TAG GAC GTG GAA TAT GGC AAG-3′: 11), minor satellite repeats (5′-GAC TGA AAA ACA CAT TCG TTG GAA ACG GGA TTT GTA GAA CAG TGT ATA TCA ATG AGT TAC AAT GAG-3′; 34), Line1 repeats (5′-AGG ATC TTT ACC TAT CCT AAA TCA GAT

AGG GGA CTA ATA TCC AAC ATA TAT AAA GAA CTC-3′; 35).

Statistical analysis

Tajima's test was used for synonymous and non-synonymous substitution analysis. Other data are shown as mean ± SEM and were analyzed by Student's *t*-test.

RESULTS

Mammalian Dnmt3b shows higher chromatin DNA methylation activity than Dnmt3a and non-mammalian Dnmt3b *in vivo*

Based on genome-wide *in silico* analysis, Dnmt3a and Dnmt3b homologues in 14 metazoan genomes were identified (Supplementary Table S1). Phylogenetic analysis suggested that Dnmt3a and Dnmt3b were generated by gene duplication event around the separation of vertebrates (Supplementary Figure S1). To experimentally examine the functional divergence of Dnmt3a and Dnmt3b after gene duplication, we tested the chromatin DNA methylation activity of Dnmt3a and Dnmt3b of zebrafish, lizard, chicken, mouse and human, using budding yeast as the *in vivo* test tube. Budding yeast is an ideal system to investigate eukaryotic DNMTs, because it provides chromatin-packed genomic DNA that mimics the *in vivo* substrate of Dnmt3 proteins, but has neither endogenous DNMTs nor any background DNA methylation (36–38). Indicated Dnmt3 proteins were firstly introduced into yeast by galactose-controlled expression, and expressed at similar levels (Figure 1A). Then yeast genomic DNA was extracted and subjected to HPLC analysis. As shown in Figure 1B, the *in vivo* chromatin DNA methylation was much higher in yeast expressing mammalian Dnmt3b proteins than other Dnmt3 proteins (Dnmt3a and non-mammalian Dnmt3b). No DNA methylation was detected in yeast expressing chicken Dnmt3b, which has been predicted to be enzymatically inactive due to the natural loss of an exon (39), a phenomenon appears to be conserved in birds (Supplementary Figure S2). The Dnmt3 proteins were also immunoprecipitated from equal amounts of yeast lysates and subjected to *in vitro* DNMT assays using poly(dG-dC)-poly(dG-dC) as the substrate (6). As shown in Figure 1C, all the Dnmt3 proteins (except chicken Dnmt3b) showed comparable DNA methylation activity to mammalian Dnmt3b, similar results were obtained when extracted yeast genomic DNA was used as the substrate (Supplementary Figure S3). These results demonstrated that the relatively low chromatin DNA methylation activity of Dnmt3a and non-mammalian Dnmt3b was neither due to loss of enzymatic activity by the possible incorrect protein folding when expressed in yeast, nor to their possible different preference for the DNA sequence of yeast genome. In addition, immunofluorescence analysis revealed that both mouse Dnmt3a and Dnmt3b (mDnmt3a, mDnmt3b) were correctly localized in the yeast nucleus without significant difference (Supplementary Figure S4). Thus, our results

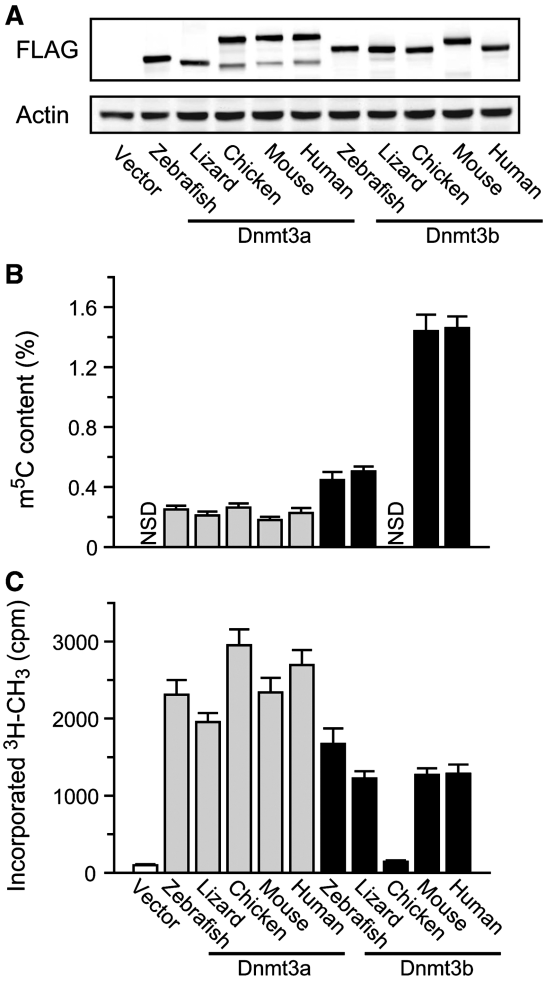

Figure 1. Mammalian Dnmt3b possesses higher chromatin DNA methylation activity than other Dnmt3 proteins. (A) Dnmt3 proteins were ectopically expressed in yeast at similar levels. (B) The m^5C contents (m^5C/total cytosine) in yeast chromatin DNA were measured by HPLC. NSD, no signal detected. (C) The Dnmt3 proteins were immunoprecipitated from equal amounts of yeast lysates and subjected to *in vitro* DNMT assays using poly(dG-dC)-poly(dG-dC) as the substrate. Data are presented as mean ± SEM, $n = 3$.

indicated that mammalian Dnmt3b had higher chromatin DNA methylation activity than Dnmt3a and non-mammalian Dnmt3b *in vivo*, which might be attributed to the sequence divergence during evolution.

An evolutionary amino acid substitution (I662N) in mammalian Dnmt3b is crucial for its high chromatin DNA methylation activity

To determine which domain of mammalian Dnmt3b contributes to its high chromatin DNA methylation activity, we generated mouse 3aN-3bC and 3bN-3aC chimeras, and measured their chromatin DNA methylation activity in yeast (Figure 2A). Across the entire sequence, the 3bN-3aC chimera is more similar to mDnmt3b than the 3aN-3bC chimera (91% versus 53% sequence identity to mDnmt3b). Nevertheless, it is the less similar 3aN-3bC chimera that displayed high chromatin DNA methylation

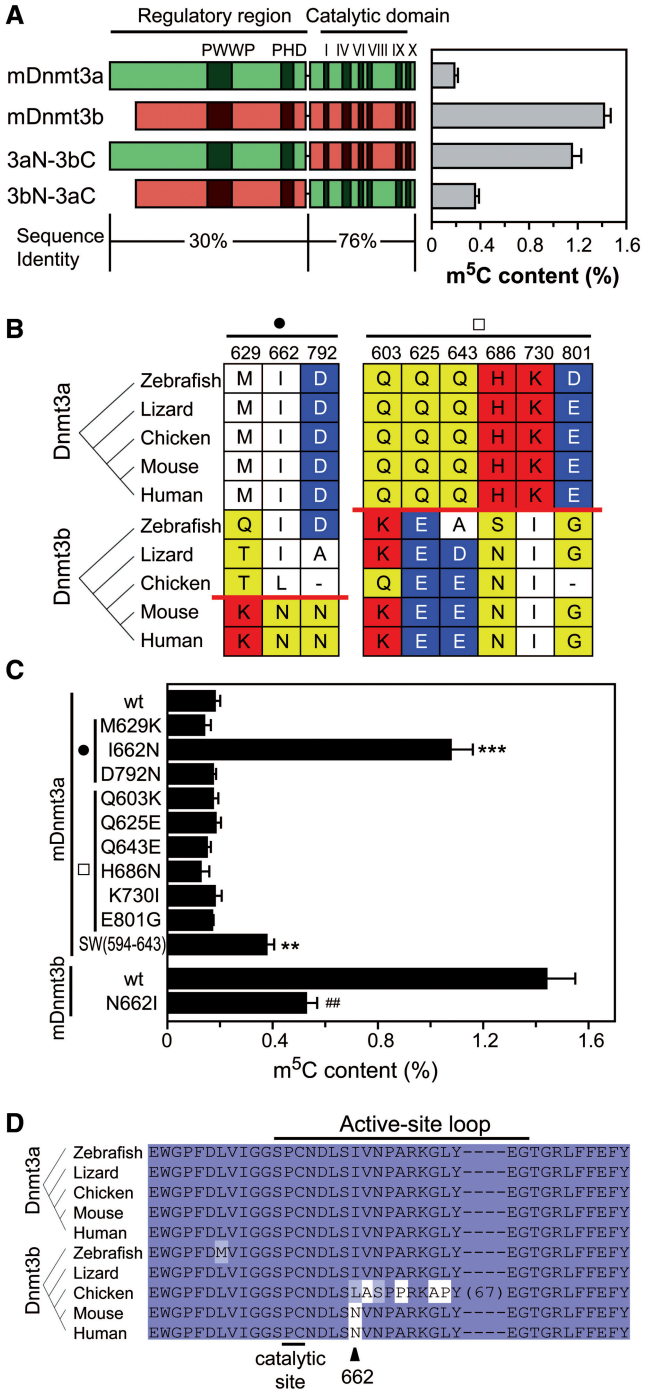

Figure 2. The I662N substitution in mammalian Dnmt3b is crucial for the high chromatin DNA methylation activity. (**A**) The diagram of 3aN-3bC and 3bN-3aC chimeras is shown and their chromatin DNA methylation activity was measured as the m^5C content in yeast chromatin DNA. (**B** and **C**) Radical amino acid substitutions between mammalian Dnmt3b and other Dnmt3 proteins (filled circle) or between Dnmt3a and Dnmt3b (open square) in the catalytic domain were listed (B), and their effects on the chromatin DNA methylation activity was measured as the m^5C content in yeast chromatin DNA (C). Amino acids are classified into four groups: positive charged (red), negative charged (blue), hydrophilic (yellow) and hydrophobic (white). The numbering in all figures, if not specified, refers to mDnmt3b sequence. Wt, wild-type. Data are presented as mean ± SEM, $n = 3$. $^{***}P < 0.001$, $^{**}P < 0.01$, versus mDnmt3a; $^{\#\#}P < 0.01$, versus mDnmt3b. (**D**) The I662N substitution of Dnmt3b is located on the active-site loop of the catalytic domain and restricted in mammals.

activity similar to mDnmt3b. Thus, the sequence divergence of the highly conserved catalytic domain is mainly responsible for the high chromatin DNA methylation activity of mammalian Dnmt3b.

Radical amino acid substitutions (substitutions with amino acids of distinct properties) have been suggested to be the major contributor for diverged protein functions (40–42). In the conserved catalytic domain, we found three radical amino acid substitutions that discriminate between mammalian Dnmt3b and other Dnmt3 proteins (Figure 2B and Supplementary Figure S5), suggesting a possible association with the high chromatin DNA methylation activity. We thus created versions of mDnmt3a in which these positions (the numbering in the study, if not specified, refers to mDnmt3b) were individually changed to the corresponding amino acid of mDnmt3b. These mutations did not significantly change the protein expression levels in yeast (Supplementary Figure S6). Then we tested their chromatin DNA methylation activity in yeast (Figure 2C). Remarkably, the I662N substitution, which is located on the highly conserved active-site loop (32; Figure 2D and Supplementary Figure S9), significantly increased the chromatin DNA methylation activity of mDnmt3a, and the reverse mutation (N662I) at the corresponding site of mDnmt3b significantly decreased its chromatin DNA methylation activity (Figure 2C). In contrast, the reciprocal mutation at position 662 of mDnmt3a and mDnmt3b did not significantly change their *in vitro* DNA methylation activity towards poly(dG-dC)-poly(dG-dC) (Supplementary Figure S7), suggesting a specific role of this residue for the chromatin DNA methylation activity. Since non-mammalian Dnmt3b also showed slightly higher chromatin DNA methylation activity than Dnmt3a (Figure 1B), the amino acids in the catalytic domain that radically discriminate between Dnmt3a and Dnmt3b were also individually substituted (Figure 2B). However, none of these substitutions alone had significant effects on the activity of mDnmt3a, only swapping a short stretch of sequence (position 594–643) containing several of these residues led to a moderate increase in activity (Figure 2C). Thus, our results suggested that these substitutions might function coordinately and have some effects on methylation activity, while the I662N substitution alone played a critical role for the high chromatin DNA methylation activity of mammalian Dnmt3b.

The I662N substitution enhances Dnmt3b's nucleosomal DNA methylation activity

To find out the accurate effects of this substitution on chromatin DNA methylation, we examined DNA methylation status of two stretches of genomic DNA in the yeast expressing mDnmt3a, mDnmt3b and their mutants by bisulfite genomic sequencing. Consistent with the yeast global chromatin DNA methylation observed by HPLC analysis, genomic DNA from mDnmt3b- or mDnmt3a (I662N)-expressing yeast contained substantially more methylated CpG sites than that from mDnmt3a- or mDnmt3b (N662I)-expressing yeast in the two examined sequences (Figure 3A). Notably, in comparison with a

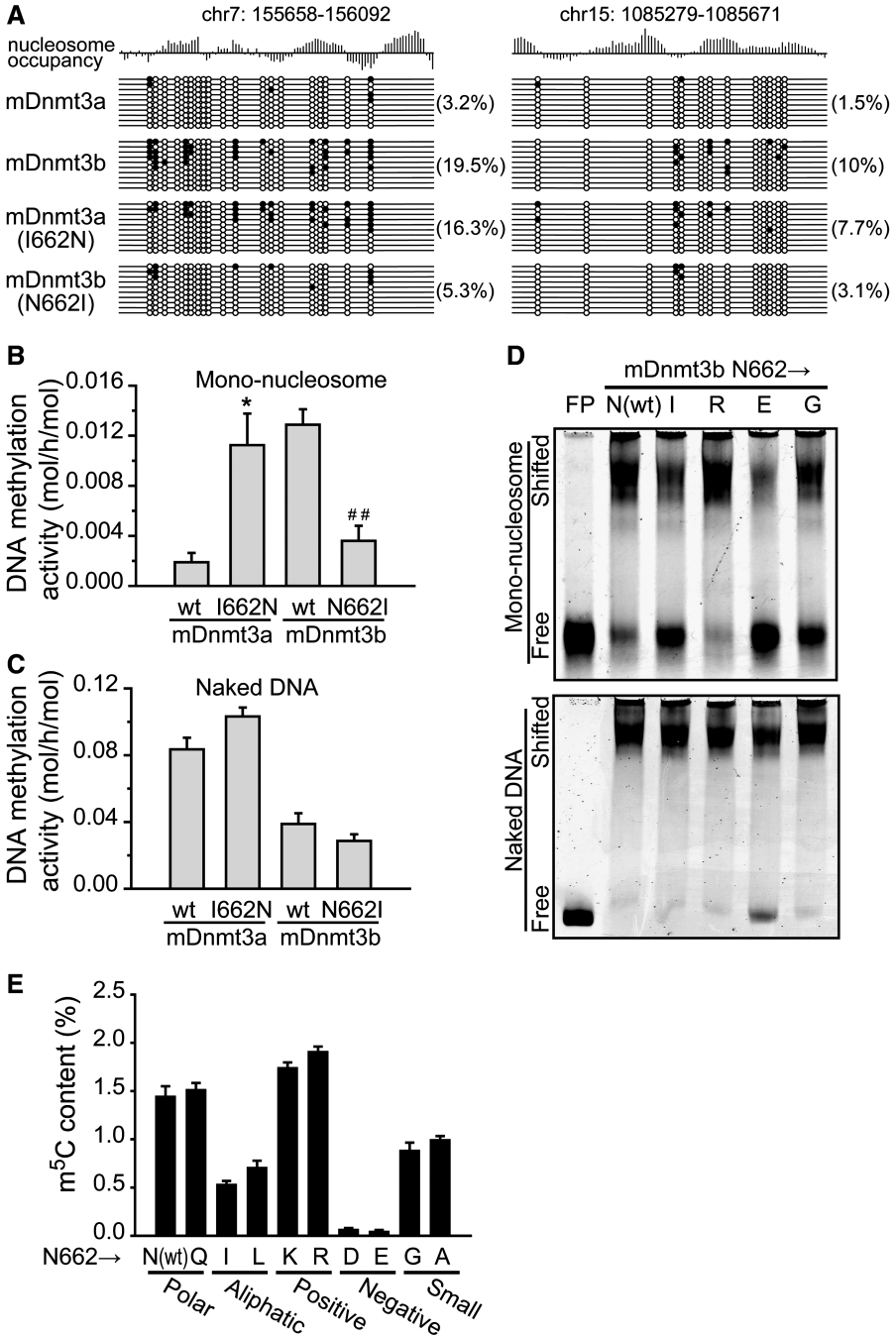

Figure 3. The I662N substitution enhances the binding of Dnmt3b to nucleosomes. (**A**) DNA methylation status of two stretches of genomic DNA was examined by bisulphite genomic sequencing in the yeast expressing different Dnmt3 proteins. Black circles represent methylated sites, and the percentages of methylated CpG sites are indicated in parenthesis. Previously reported nucleosome occupancies of the corresponding regions are indicated at the top, each vertical line represents the probe intensity ratio between nucleosomal and whole-genome DNA at that position, a positive signal represents nucleosome occupancy (43). (**B** and **C**) *In vitro* DNA methylation activity towards mono-nucleosomes (B) or corresponding naked DNA (C) was measured using purified Dnmt3 proteins. Data are presented as mean ± SEM, $n = 3$. $^{*}P < 0.05$, versus mDnmt3a; $^{\#\#}P < 0.01$, versus mDnmt3b. (**D**) EMSA analysis was carried out using mono-nucleosomes or naked DNA as the probe. FP, free probe. (**E**) Chromatin DNA methylation activity of mDnmt3b with different substitutions at position 662 was measured as the m^5C content in yeast chromatin DNA. Data are presented as mean ± SEM, $n = 3$.

high-resolution atlas of nucleosome occupancy in budding yeast (43) (Figure 3A), we found that methylated sites in mDnmt3a- and mDnmt3b (N662I)-expressing yeast were mostly restricted to some specific CpG sites in the linker regions between nucleosomes, while mDnmt3b and mDnmt3a (I662N) methylated CpGs in both linker and nucleosome occupied regions. Combined with the previous report that mDnmt3b possessed higher nucleosomal DNA methylation activity than mDnmt3a (8), we speculated that the N662 residue might contribute to mDnmt3b's high chromatin DNA methylation activity through improving its methylation capability towards

nucleosomal DNA. To test this hypothesis, we prepared mono-nucleosomes from mammalian cells (Supplementary Figure S8) and measured the DNA methylation activity of purified mDnmt3a, mDnmt3b, mDnmt3a (I662N), mDnmt3b (N662I) *in vitro*, using mono-nucleosomes as well as naked DNA extracted from the nucleosomes as substrates. As expected, mDnmt3b and mDnmt3a (I662N) showed significantly higher nucleosomal DNA methylation activity than mDnmt3b (N662I) or mDnmt3a (Figure 3B), while the reciprocal mutations at position 662 of mDnmt3b and Laboratory of Molecular Cell Biology and Center of Cell Signaling, mDnmt3a did not significantly change their DNA methylation activity towards naked DNA (Figure 3C). These results demonstrated that I662N substitution could confer an intrinsic change of enzymatic activity towards nucleosomal DNA, and largely excluded the possibility that lack of proper histone modifications in yeast was the cause of the difference in methylation activity between mDnmt3a and mDnmt3b.

A polar or positive charged amino acid at position 662 enhances the binding of Dnmt3b to nucleosomes

To explore the underlying mechanism for the function of residue N662, we created and purified mDnmt3b mutants with their N662 residue substituted to amino acids with distinct properties, and carried out EMSA. In this assay, sinefungin, an analogue of the methyl group donor AdoMet, was used to ensure the formation of the enzyme–DNA complex without methyl group transfer (32). As shown in Figure 3D, mutants with N662 substituted to I (aliphatic), E (negative charged) or G (small) displayed weaker nucleosome binding than wild-type mDnmt3b, while the mutant with positive charged amino acid R displayed even stronger binding to nucleosomes; meanwhile, the binding to naked DNA was not significantly affected except when N662 was substituted to E, which might strongly repel the nearby phosphate group of DNA backbone. Taken together, we proposed that a polar or positive charged amino acid at position 662 could facilitate the formation of active enzyme-DNA complex for nucleosomal DNA so that further methyl group transfer could proceed effectively. Consistent with this proposal, substitution of residue 662 with polar or positive charged amino acids conferred high chromatin DNA methylation activity in yeast, while those with negative charged amino acids almost abolished this activity (Figure 3E).

The I662N substitution is crucial for Dnmt3b to efficiently methylate repetitive sequences in mammalian cells

To evaluate the biological importance of I662N substitution for Dnmt3b's functions in mammals, wild-type mDnmt3b and mDnmt3b (N662I) were introduced into late passage *Dnmt3a*$^{-/-}$*3b*$^{-/-}$ embryonic stem (ES) cells, which are severely hypomethylated throughout the genome (11). Clones expressing GFP (GFP3, GFP7), mDnmt3b (b1 and b8) and mDnmt3b N662I (bM2 and bM9) were derived, where the exogenous expression levels of mDnmt3b and mDnmt3b (N662I) were similar to the endogenous Dnmt3b expression in wild-type ES cells (J1) (Figure 4A). As revealed by HPLC analysis, wild-type mDnmt3b restored DNA methylation of *Dnmt3a*$^{-/-}$*3b*$^{-/-}$ ES cells to an average of 78% of the normal level, while the N662I mutant could only restore DNA methylation to 30% of the normal level, further confirming the critical role of residue N662 (Figure 4B).

Since measurement of global DNA methylation levels reveals mostly the methylation status of repetitive sequences, the major targets of DNA methylation (17), we next examined the methylation status of several repetitive sequences, including minor satellite (centromeric tandem repeats), major satellite (pericentromeric tandem repeats) and Line1 (long interspersed repeats). By using methylation-sensitive restriction enzymes and Southern blot with specific probes, we found that all these repetitive sequences were highly methylated in wild-type ES cells but severely hypomethylated in GFP-expressing *Dnmt3a*$^{-/-}$*3b*$^{-/-}$ ES cells. And as expected, the N662I mutant methylated these repetitive sequences much less efficiently than wild-type mDnmt3b (Figure 4C). Then we also checked the methylation status of specific genomic loci of single copy genes. The phosphoglycerate kinase 2 (*Pgk-2*) and *β-globin* genes were reported to be highly methylated in wild-type ES cells but hypomethylated in late passage *Dnmt3a*$^{-/-}$*3b*$^{-/-}$ ES cells (11), so we examined their methylation status by bisulfite genomic sequencing. Both wild-type mDnmt3b and the N662I mutant partially restored the DNA methylation on these foci, however, no significant difference was observed between ES cells expressing wild-type mDnmt3b and the N662I mutant (Figure 4D). These results indicated a specific role of the I662N substitution in the methylation of repetitive sequences, which might result from potentially unequal chromosomal accessibility between repetitive sequences and single copy genes.

DISCUSSION

In this study, we firstly showed that the chromatin DNA methylation activity of mammalian Dnmt3b was greatly enhanced due to an evolutionary amino acid substitution in the catalytic domain. Moreover, this substitution is crucial for Dnmt3b to efficiently methylate the repetitive sequences in mammalian genomes, which is important for genome stability and gene regulation (4). It is noteworthy that the I662N substitution appears to be restricted to mammals (Figure 2D), of which the genomes are generally characterized by containing the largest percentage of repetitive sequences among all metazoans (15,16). Meanwhile, Dnmt3b proteins in birds are enzymatically inactive (Figure 1B and C; Supplementary Figures S2 and S3), and birds' genomes are characterized by the low density of repetitive sequences (16,44). These results suggest an association of the chromatin DNA methylation activity of Dnmt3b with the density of repetitive sequences in the genome. In addition, Tajima's test rejected the equal substitution rate hypothesis (i.e. molecular clock hypothesis) on the non-synonymous sites (but not on synonymous sites) for Dnmt3a and Dnmt3b (Figure 5A), and

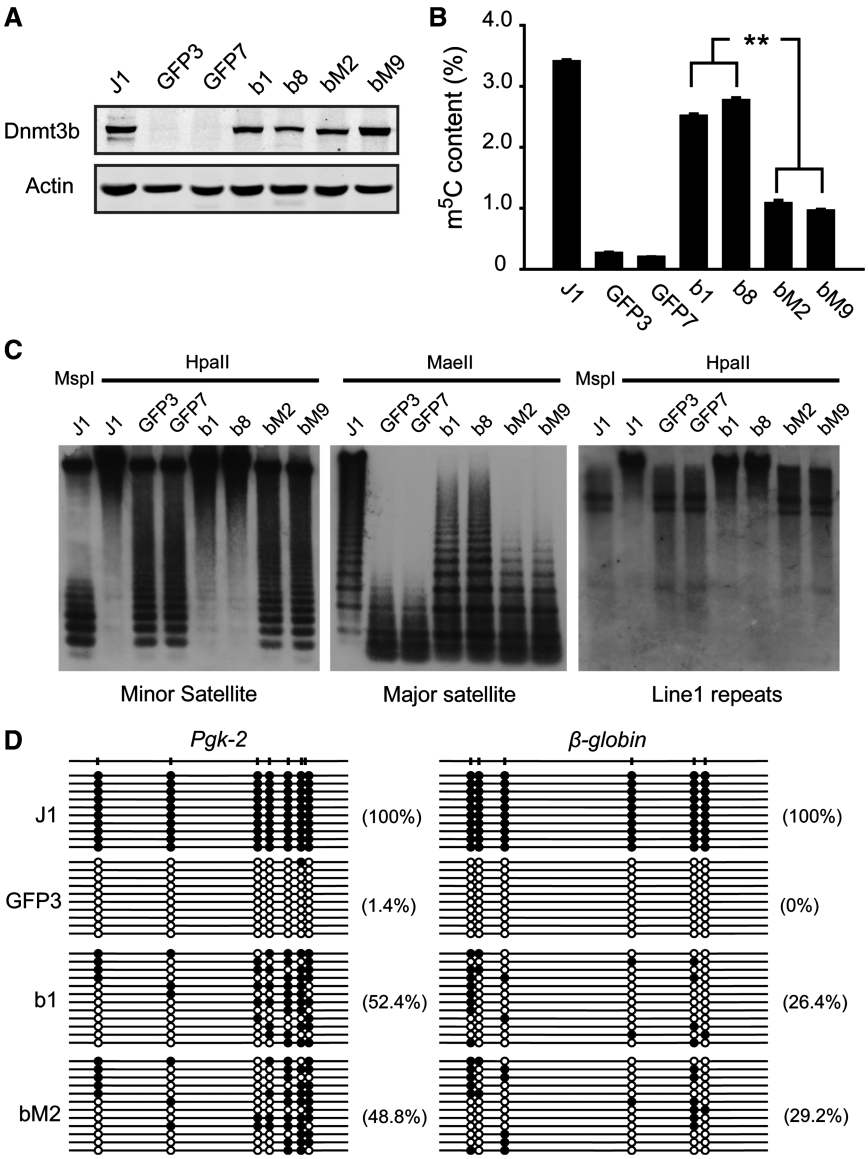

Figure 4. The reverse substitution (N662I) severely compromises mammalian Dnmt3b's efficiency in methylating repetitive sequences. (**A**) GFP (GFP3, GFP7), mDnmt3b (b1, b8) and mDnmt3b (N662I) (bM2, bM9) were stably expressed in *Dnmt3a*$^{-/-}$*3b*$^{-/-}$ ES cells. (**B**) Global DNA methylation levels were assessed by HPLC. Data are presented as mean ± SEM, $n = 3$. $^{**}P < 0.01$. (**C**) DNA methylation status of minor satellite, major satellite and Line1 repetitive sequences was assessed by southern blot. MspI was the complete digestion control for HpaII. (**D**) DNA methylation status of *β–globin* and *Pgk-2* genes was analyzed by bisulphite genomic sequencing.

Dnmt3b showed a significant faster non-synonymous substitution rate than Dnmt3a in vertebrates, during the emergence of mammals (Figure 5B). All of these observations provide consistent evidence suggesting that Dnmt3b evolved more rapidly than Dnmt3a during the emergence of mammals and gained enhanced chromatin DNA methylation activity as a functional adaptation to the dramatically increased repetitive sequences in mammalian genomes.

Based on our findings, a scenario for the evolution of Dnmt3a and Dnmt3b after their generation by gene duplication can be speculated. After the duplication of Dnmt3 gene around the separation of vertebrates, Dnmt3b quickly gained moderately higher chromatin DNA methylation activity than Dnmt3a (Figure 1B), while the latter probably retained a more ancestral state. Invertebrate genomes are fractionally methylated (mosaic methylation), while vertebrate genomes are globally methylated (45,46). The great increase of DNA methylation level in the genome should correspondingly require higher chromatin DNA methylation activity, so that the reservation of both copies of the duplicated Dnmt3 could be advantageous. Another challenge was the dramatically increased repetitive sequences in mammalian genomes. The enhanced chromatin DNA methylation activity of mammalian Dnmt3b conferred by I662N substitution might have provided the selective advantage and further led to its fixation in mammals. Interestingly, Dnmt3L, a

A

	Synonymous			Non-Synonymous		
	S_{3a}	S_{3b}	P-value	N_{3a}	N_{3b}	P-value
Human	221	228	0.377	294	343	0.025*
Chimpanzee	222	227	0.413	294	343	0.025*
Dog	227	236	0.340	298	342	0.041*
Cow	230	222	0.366	299	347	0.029*
Mouse	220	228	0.342	300	345	0.040*
Rat	227	229	0.460	298	340	0.049*
Opossum	247	257	0.319	298	357	0.011*
Chicken	226	208	0.196	307	350	0.047*
Lizard	231	229	0.463	305	353	0.031*

B

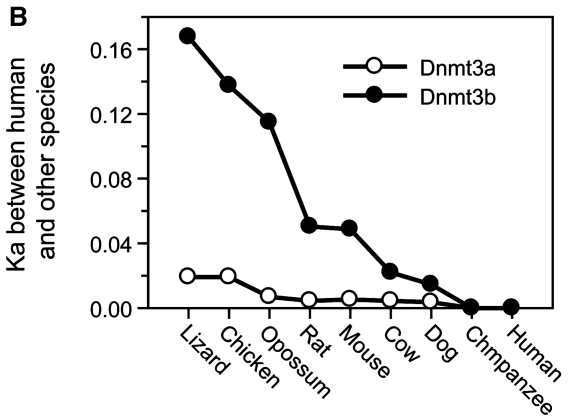

Figure 5. Dnmt3b evolved more rapidly than Dnmt3a during the emergence of mammals. (**A**) Tajima's molecular clock test for synonymous substitutions (S_{3a}, S_{3b}) and non-synonymous substitutions (N_{3a}, N_{3b}) in Dnmt3a and Dnmt3b of representative species. Dnmt3 of amphioxus was used as the outgroup. $^*P < 0.05$. (**B**) Non-synonymous substitution rates (*Ka*) of Dnmt3a and Dnmt3b from representative species against their corresponding human orthologs.

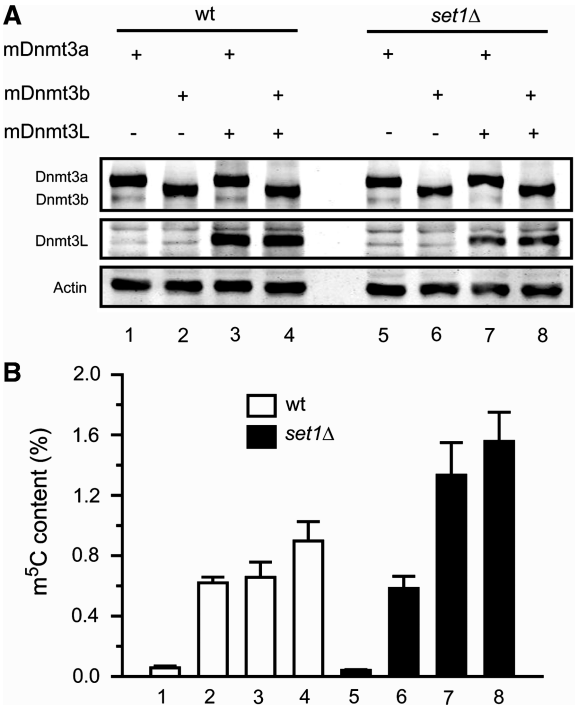

Figure 6. Co-expression of Dnmt3L with Dnmt3a or Dnmt3b enhances chromatin DNA methylation activity. (**A**) Western blot of the DNMTs expressed in different yeast strains. Each strain was transformed with one or two expression constructs (+) or empty vector (−) as indicated on the top. Set1*Δ*, Set1 knockout. (**B**) The m^5C contents in indicated yeast chromatin DNA were measured by HPLC. Data are presented as mean ± SEM, $n = 3$.

regulatory factor for Dnmt3a and Dnmt3b, also seems to appear only in mammals (39). And we found that Dnmt3L could enhance the chromatin DNA methylation activity of both Dnmt3a and Dnmt3b (Figure 6), probably through (i) stabilizing the conformation of the active-site loop to facilitate the formation of active enzyme–DNA complex, especially for Dnmt3a (32); and (ii) targeting them to chromatin regions where methylation of histone H3 lysine 4 is absent (47,48). Thus, both I662N substitution in Dnmt3b and appearance of Dnmt3L greatly enhanced chromatin DNA methylation activity in mammals.

Although the single amino acid substitution I662N in the catalytic domain is mainly responsible for mammalian Dnmt3b's high nucleosomal DNA methylation activity observed in this study, we also noticed several radical amino acid substitutions located in the N-terminal regulatory region, which was reported to interact with regulatory factors (e.g. Dnmt3L, chromatin remodeling factors, histone deacetylases and methyltransferases; 4,20). It has been reported that deficiency in certain ATP-dependent chromatin remodeling factors (e.g. Lsh in mouse, ATRX in human) results in the disruption of chromatin DNA methylation patterns (35,49,50). The post-translational modified histones and other chromatin-associated proteins also play crucial roles in the targeting of DNMTs (51–55). Thus, instead of altering the catalytic property, those substitutions in the N-terminal regulatory region might be involved in the interactions between Dnmt3 proteins and their various regulatory factors and contribute to the regulation of chromatin DNA methylation. Further exploring their roles is necessary to fully understand the evolution of Dnmt3 gene family.

By superimposing the structure of Dnmt3a catalytic domain (32) and nucleosome (56) onto the structure of M.HhaI–DNA complex (57), we found that the residue corresponding to N662 of mDnmt3b seemed to be located at the edge of the minor groove of DNA helix and close to the phosphate group of DNA backbone (Supplementary Figure S9). It is possible that residue N662 facilitates the formation of active enzyme–DNA complex through its interaction with the phosphate group of DNA backbone and stabilizing the interaction between the enzyme's active site loop and DNA, which might be especially important when substrate DNA is wrapped into nucleosomes with low accessibility to binding proteins. These observations support our conclusion that the different methylation activity towards nucleosomal DNA between Dnmt3a and Dnmt3b is their intrinsic enzymatic property. However, we do not exclude that regulatory factors together with various histone modifications play important roles in regulating the methylation activity of Dnmt3a and Dnmt3b *in vivo*,

especially in specific physiological conditions, so that both of them could methylate the chromatin efficiently in a well-controlled manner and establish proper DNA methylation patterns.

It is very common to infer the possibility of adaptive evolution by comparing the non-synonymous and synonymous substitution rate. However, these analyses not only predict a significant number of false positives but also fail to predict many positively selected sites. Thus, the importance of potential adaptive amino acid substitution must be verified in proper functional assays, and it is essential to establish such functional assays for the genes to be studied (58,59). In this study, we used budding yeast as the *in vivo* test tube to measure the chromatin DNA methylation activity of Dnmt3 proteins and proved the adaptive significance of I662N substitution for mammalian Dnmt3b. Due to the ease of genetic manipulations and possessing most of the histone modifications of higher eukaryotes, the budding yeast system we used in this study will also be a powerful tool to study the cross-talk between DNA methylation and various histone modifications during evolution, including the evolution of related histone modification enzymes.

Taken together, our study not only provides novel insight into the evolution process of Dnmt3a and Dnmt3b after gene duplication and the molecular mechanism of the diverged functions of Dnmt3a and Dnmt3b, but also gives the first example that minor changes in sequence could provide the functional novelty conferring a selective advantage and contribute to functional adaptation in the epigenetic system.

SUPPLEMENTARY DATA

Supplementary data are available at NAR Online.

ACKNOWLEDGEMENTS

We are grateful to Dr Arthur D. Riggs, Dr Guo-liang Xu, Dr Hiroyuki Sasaki, Dr Nobuyoshi Shimoda, Dr Jin-qiu Zhou, Dr En Li and Dr Taiping Chen for sharing materials. We would like to thank Dr Gang Pei, Dr Liping Wei for helpful discussions, Dr Jian Zhao, Dr Kun He, Dr Xu Xu, Shun-mei Xin and Shan Chen for technical assistance, Dr Bin Wei and Dr Letian Tao for proofreading of this article.

FUNDING

Ministry of Science and Technology (2005CB522406, 2007CB947904, 2006AA02Z334, 2007AA02Z165, 2007CB946904, 2010CB944900); National Natural Science Foundation of China (30623003, 30621091, 30625014, 30871285, 90713047, 90919028); Chinese Academy of Sciences (SIBS2008001). Funding for open access charge: Ministry of Science and Technology.

Conflict of interest statement. None declared.

REFERENCES

1. Bird,A. (2002) DNA methylation patterns and epigenetic memory. *Genes Dev.*, **16**, 6–21.
2. Jones,P.A. and Baylin,S.B. (2002) The fundamental role of epigenetic events in cancer. *Nat. Rev. Genet.*, **3**, 415–428.
3. Robertson,K.D. (2005) DNA methylation and human disease. *Nat. Rev. Genet.*, **6**, 597–610.
4. Goll,M.G. and Bestor,T.H. (2005) Eukaryotic cytosine methyltransferases. *Annu. Rev. Biochem.*, **74**, 481–514.
5. Okano,M., Bell,D.W., Haber,D.A. and Li,E. (1999) DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell*, **99**, 247–257.
6. Aoki,A., Suetake,I., Miyagawa,J., Fujio,T., Chijiwa,T., Sasaki,H. and Tajima,S. (2001) Enzymatic properties of de novo-type mouse DNA (cytosine-5) methyltransferases. *Nucleic Acids Res.*, **29**, 3506–3512.
7. Robertson,A.K., Geiman,T.M., Sankpal,U.T., Hager,G.L. and Robertson,K.D. (2004) Effects of chromatin structure on the enzymatic and DNA binding functions of DNA methyltransferases DNMT1 and Dnmt3a in vitro. *Biochem. Biophys. Res. Commun.*, **322**, 110–118.
8. Takeshima,H., Suetake,I., Shimahara,H., Ura,K., Tate,S. and Tajima,S. (2006) Distinct DNA methylation activity of Dnmt3a and Dnmt3b towards naked and nucleosomal DNA. *J. Biochem.*, **139**, 503–515.
9. Takeshima,H., Suetake,I. and Tajima,S. (2008) Mouse Dnmt3a preferentially methylates linker DNA and is inhibited by histone H1. *J. Mol. Biol.*, **383**, 810–821.
10. Kaneda,M., Okano,M., Hata,K., Sado,T., Tsujimoto,N., Li,E. and Sasaki,H. (2004) Essential role for de novo DNA methyltransferase Dnmt3a in paternal and maternal imprinting. *Nature*, **429**, 900–903.
11. Chen,T., Ueda,Y., Dodge,J.E., Wang,Z. and Li,E. (2003) Establishment and maintenance of genomic methylation patterns in mouse embryonic stem cells by Dnmt3a and Dnmt3b. *Mol. Cell Biol.*, **23**, 5594–5605.
12. Oka,M., Rodic,N., Graddy,J., Chang,L.J. and Terada,N. (2006) CpG sites preferentially methylated by Dnmt3a in vivo. *J. Biol. Chem.*, **281**, 9901–9908.
13. Dodge,J.E., Okano,M., Dick,F., Tsujimoto,N., Chen,T., Wang,S., Ueda,Y., Dyson,N. and Li,E. (2005) Inactivation of Dnmt3b in mouse embryonic fibroblasts results in DNA hypomethylation, chromosomal instability, and spontaneous immortalization. *J. Biol. Chem.*, **280**, 17986–17991.
14. Ueda,Y., Okano,M., Williams,C., Chen,T., Georgopoulos,K. and Li,E. (2006) Roles for Dnmt3b in mammalian development: a mouse model for the ICF syndrome. *Development*, **133**, 1183–1192.
15. International-Human-Genome-Sequencing-Consortium. (2001) Initial sequencing and analysis of the human genome. *Nature*, **409**, 860–921.
16. Organ,C.L., Shedlock,A.M., Meade,A., Pagel,M. and Edwards,S.V. (2007) Origin of avian genome size and structure in non-avian dinosaurs. *Nature*, **446**, 180–184.
17. Yoder,J.A., Walsh,C.P. and Bestor,T.H. (1997) Cytosine methylation and the ecology of intragenomic parasites. *Trends Genet.*, **13**, 335–340.
18. Jeanpierre,M., Turleau,C., Aurias,A., Prieur,M., Ledeist,F., Fischer,A. and Viegas-Pequignot,E. (1993) An embryonic-like methylation pattern of classical satellite DNA is observed in ICF syndrome. *Hum. Mol. Genet.*, **2**, 731–735.
19. 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) Localization of the ICF syndrome to chromosome 20 by homozygosity mapping. *Am. J. Hum. Genet.*, **63**, 803–809.
20. Chen,T. and Li,E. (2006) Establishment and maintenance of DNA methylation patterns in mammals. *Curr. Top. Microbiol. Immunol.*, **301**, 179–201.
21. Okano,M., Xie,S. and Li,E. (1998) Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases. *Nat. Genet.*, **19**, 219–220.

22. Hermann,A., Gowher,H. and Jeltsch,A. (2004) Biochemistry and biology of mammalian DNA methyltransferases. *Cell Mol. Life Sci.*, **61**, 2571–2587.
23. Ohno,S. (1970) *Evolution by Gene Duplication*. Springer, New York, USA.
24. Force,A., Lynch,M., Pickett,F.B., Amores,A., Yan,Y.L. and Postlethwait,J. (1999) Preservation of duplicate genes by complementary, degenerative mutations. *Genetics*, **151**, 1531–1545.
25. Prince,V.E. and Pickett,F.B. (2002) Splitting pairs: the diverging fates of duplicated genes. *Nat. Rev. Genet.*, **3**, 827–837.
26. Taylor,J.S. and Raes,J. (2004) Duplication and divergence: the evolution of new genes and old ideas. *Annu. Rev. Genet.*, **38**, 615–643.
27. Notredame,C., Higgins,D.G. and Heringa,J. (2000) T-Coffee: a novel method for fast and accurate multiple sequence alignment. *J. Mol. Biol.*, **302**, 205–217.
28. Kumar,S., Nei,M., Dudley,J. and Tamura,K. (2008) MEGA: a biologist-centric software for evolutionary analysis of DNA and protein sequences. *Brief Bioinform.*, **9**, 299–306.
29. Nei,M. and Gojobori,T. (1986) Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Mol. Biol. Evol.*, **3**, 418–426.
30. Tajima,F. (1993) Simple methods for testing the molecular evolutionary clock hypothesis. *Genetics*, **135**, 599–607.
31. Ramsahoye,B.H. (2002) Measurement of genome wide DNA methylation by reversed-phase high-performance liquid chromatography. *Methods*, **27**, 156–161.
32. Jia,D., Jurkowska,R.Z., Zhang,X., Jeltsch,A. and Cheng,X. (2007) Structure of Dnmt3a bound to Dnmt3L suggests a model for de novo DNA methylation. *Nature*, **449**, 248–251.
33. Olek,A., Oswald,J. and Walter,J. (1996) A modified and improved method for bisulphite based cytosine methylation analysis. *Nucleic Acids Res.*, **24**, 5064–5066.
34. Tada,M., Tada,T., Lefebvre,L., Barton,S.C. and Surani,M.A. (1997) Embryonic germ cells induce epigenetic reprogramming of somatic nucleus in hybrid cells. *EMBO J.*, **16**, 6510–6520.
35. Dennis,K., Fan,T., Geiman,T., Yan,Q. and Muegge,K. (2001) Lsh, a member of the SNF2 family, is required for genome-wide methylation. *Genes Dev.*, **15**, 2940–2944.
36. Bulkowska,U., Ishikawa,T., Kurlandzka,A., Trzcinska-Danielewicz,J., Derlacz,R. and Fronk,J. (2007) Expression of murine DNA methyltransferases Dnmt1 and Dnmt3a in the yeast Saccharomyces cerevisiae. *Yeast*, **24**, 871–882.
37. Proffitt,J.H., Davie,J.R., Swinton,D. and Hattman,S. (1984) 5-Methylcytosine is not detectable in Saccharomyces cerevisiae DNA. *Mol. Cell Biol.*, **4**, 985–988.
38. Colot,V. and Rossignol,J.L. (1999) Eukaryotic DNA methylation as an evolutionary device. *Bioessays*, **21**, 402–411.
39. Yokomine,T., Hata,K., Tsudzuki,M. and Sasaki,H. (2006) Evolution of the vertebrate DNMT3 gene family: a possible link between existence of DNMT3L and genomic imprinting. *Cytogenet. Genome Res.*, **113**, 75–80.
40. Zhang,J., Zhang,Y.P. and Rosenberg,H.F. (2002) Adaptive evolution of a duplicated pancreatic ribonuclease gene in a leaf-eating monkey. *Nat. Genet.*, **30**, 411–415.
41. Maizel,A., Busch,M.A., Tanahashi,T., Perkovic,J., Kato,M., Hasebe,M. and Weigel,D. (2005) The floral regulator LEAFY evolves by substitutions in the DNA binding domain. *Science*, **308**, 260–263.
42. Gu,X. (2006) A simple statistical method for estimating type-II (cluster-specific) functional divergence of protein sequences. *Mol. Biol. Evol.*, **23**, 1937–1945.
43. Lee,W., Tillo,D., Bray,N., Morse,R.H., Davis,R.W., Hughes,T.R. and Nislow,C. (2007) A high-resolution atlas of nucleosome occupancy in yeast. *Nat. Genet.*, **39**, 1235–1244.
44. International-Chicken-Genome-Sequencing-Consortium. (2004) Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. *Nature*, **432**, 695–716.
45. Tweedie,S., Charlton,J., Clark,V. and Bird,A. (1997) Methylation of genomes and genes at the invertebrate-vertebrate boundary. *Mol. Cell Biol.*, **17**, 1469–1475.
46. Suzuki,M.M. and Bird,A. (2008) DNA methylation landscapes: provocative insights from epigenomics. *Nat. Rev. Genet.*, **9**, 465–476.
47. Ooi,S.K., Qiu,C., Bernstein,E., Li,K., Jia,D., Yang,Z., Erdjument-Bromage,H., Tempst,P., Lin,S.P., Allis,C.D. *et al.* (2007) DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA. *Nature*, **448**, 714–717.
48. Hu,J.L., Zhou,B.O., Zhang,R.R., Zhang,K.L., Zhou,J.Q. and Xu,G.L. (2009) The N-terminus of histone H3 is required for de novo DNA methylation in chromatin. *Proc. Natl Acad. Sci. USA*, **106**, 22187–22192.
49. Gibbons,R.J., McDowell,T.L., Raman,S., O'Rourke,D.M., Garrick,D., Ayyub,H. and Higgs,D.R. (2000) Mutations in ATRX, encoding a SWI/SNF-like protein, cause diverse changes in the pattern of DNA methylation. *Nat. Genet.*, **24**, 368–371.
50. Zhu,H., Geiman,T.M., Xi,S., Jiang,Q., Schmidtmann,A., Chen,T., Li,E. and Muegge,K. (2006) Lsh is involved in de novo methylation of DNA. *EMBO J.*, **25**, 335–345.
51. Vire,E., Brenner,C., Deplus,R., Blanchon,L., Fraga,M., Didelot,C., Morey,L., Van Eynde,A., Bernard,D., Vanderwinden,J.M. *et al.* (2006) The Polycomb group protein EZH2 directly controls DNA methylation. *Nature*, **439**, 871–874.
52. Brenner,C., Deplus,R., Didelot,C., Loriot,A., Vire,E., De Smet,C., Gutierrez,A., Danovi,D., Bernard,D., Boon,T. *et al.* (2005) Myc represses transcription through recruitment of DNA methyltransferase corepressor. *EMBO J.*, **24**, 336–346.
53. Lehnertz,B., Ueda,Y., Derijck,A.A., Braunschweig,U., Perez-Burgos,L., Kubicek,S., Chen,T., Li,E., Jenuwein,T. and Peters,A.H. (2003) Suv39h-mediated histone H3 lysine 9 methylation directs DNA methylation to major satellite repeats at pericentric heterochromatin. *Curr. Biol.*, **13**, 1192–1200.
54. Fuks,F., Hurd,P.J., Deplus,R. and Kouzarides,T. (2003) The DNA methyltransferases associate with HP1 and the SUV39H1 histone methyltransferase. *Nucleic Acids Res.*, **31**, 2305–2312.
55. Zilberman,D., Coleman-Derr,D., Ballinger,T. and Henikoff,S. (2008) Histone H2A.Z and DNA methylation are mutually antagonistic chromatin marks. *Nature*, **456**, 125–129.
56. Richmond,T.J. and Davey,C.A. (2003) The structure of DNA in the nucleosome core. *Nature*, **423**, 145–150.
57. Klimasauskas,S., Kumar,S., Roberts,R.J. and Cheng,X. (1994) HhaI methyltransferase flips its target base out of the DNA helix. *Cell*, **76**, 357–369.
58. Yokoyama,S., Tada,T., Zhang,H. and Britt,L. (2008) Elucidation of phenotypic adaptations: Molecular analyses of dim-light vision proteins in vertebrates. *Proc. Natl Acad. Sci. USA*, **105**, 13480–13485.
59. Nozawa,M., Suzuki,Y. and Nei,M. (2009) Reliabilities of identifying positive selection by the branch-site and the site-prediction methods. *Proc. Natl Acad. Sci. USA*, **106**, 6700–6705.