Egg-specific expression of protein with DNA methyltransferase activity in the biocarcinogenic liver fluke *Clonorchis sinensis*

SEON-HEE KIM[1], HYE-JEONG CHO[1], WOON-MOK SOHN[2], CHUN-SEOB AHN[3], YOON KONG[3], HYUN-JONG YANG[4] *and* YOUNG-AN BAE[1]*

[1] *Department of Microbiology, Graduate School of Medicine, Gachon University, 191 Hambakmoe-ro, Yeonsu-gu, Incheon 406-799, Republic of Korea*
[2] *Department of Parasitology and Institute of Health Sciences, Gyeongsang National University College of Medicine, Jinju, Republic of Korea*
[3] *Department of Molecular Parasitology, Sungkyunkwan University School of Medicine and Center for Molecular Medicine, Samsung Biomedical Research Institute, Suwon, Republic of Korea*
[4] *Department of Parasitology, Ewha Womans University School of Medicine, Seoul, Republic of Korea*

(*Received 11 February 2015; revised 21 April 2015; accepted 3 May 2015*)

SUMMARY

Despite recent reports regarding the biology of cytosine methylation in *Schistosoma mansoni*, the impact of the regulatory machinery remains unclear in diverse platyhelminthes. This ambiguity is reinforced by discoveries of DNA methyltransferase 2 (DNMT2)-only organisms and the substrate specificity of DNMT2 preferential to RNA molecules. Here, we characterized a novel DNA methyltransferase, named CsDNMT2, in a liver fluke *Clonorchis sinensis*. The protein exhibited structural properties conserved in other members of the DNMT2 family. The native and recombinant CsDNMT2 exhibited considerable enzymatic activity on DNA. The spatiotemporal expression of CsDNMT2 mirrored that of 5-methylcytosine (5 mC), both of which were elevated in the *C. sinensis* eggs. However, CsDNMT2 and 5 mC were marginally detected in other histological regions of *C. sinensis* adults including ovaries and seminal receptacle. The methylation site seemed not related to genomic loci occupied by progenies of an active long-terminal-repeat retrotransposon. Taken together, our data strongly suggest that *C. sinensis* has preserved the functional DNA methylation machinery and that DNMT2 acts as a genuine alternative to DNMT1/DNMT3 to methylate DNA in the DNMT2-only organism. The epigenetic regulation would target functional genes primarily involved in the formation and/or maturation of eggs, rather than retrotransposons.

Key words: *Clonorchis sinensis*, DNA methylation, DNMT2, 5-methylcytosine, DNA-dependent methyltransferase, egg development.

INTRODUCTION

Clonorchis sinensis is a parasitic trematode that causes clonorchiasis in humans, which is highly prevalent in China, Korea and Vietnam. This food-borne parasite afflicts more than 35 million people worldwide and thus, causes a great socio-economic burden in endemic regions (Keizer and Utzinger, 2009). Humans are infected with the parasite by consuming raw or undercooked freshwater fish containing infective *C. sinensis* metacercariae. The metacercariae excyst in the duodenum and migrate into the bile ducts, where they grow into reproductive adults. The main clinical manifestation of clonorchiasis is obstructive cholangiopathy including irregular dilatation and cystic degeneration of the bile ducts, glandular hyperplasia and metaplasia of mucin secreting biliary epithelial cells and subsequent periductal fibrosis (Hong, 2003). Most importantly, chronic infection with *C. sinensis* is likely to be closely involved with the generation of cholangiocarcinoma, a malignant tumour arising in the bile duct epithelium (Shin *et al.* 2010). Together with *Opisthorchis viverrini*, *C. sinensis* has been categorized as a Group I biocarcinogen by the World Health Organization (Bouvard *et al.* 2009).

The draft genome of *C. sinensis* was recently generated and analysed by several Chinese research groups (Wang *et al.* 2011; Huang *et al.* 2013). The last assembled version demonstrated that the liver fluke has a 547-Mb genome encompassing 13 634 protein-coding genes. The overall genome size and coding capacity of *C. sinensis* were apparently greater than those of other trematodes including *Schistosoma japonicum* (398 Mb and 12 657 genes), which can be partially attributed to increases in exon numbers and intron lengths of the protein-coding genes (The *Schistosoma japonicum* Genome Sequencing and Functional Analysis Consortium, 2009; Huang *et al.* 2013), or to the higher copy numbers as observed in a series of structural genes (Cai *et al.* 2008; Bae *et al.* 2013*a*, *b*). Interestingly, however, the fraction of repeating regions was found to be lower in the *C. sinensis* genome

* Corresponding author: Department of Microbiology, Graduate School of Medicine, Gachon University, 191 Hambakmoe-ro, Yeonsu-gu, Incheon 406-799, Republic of Korea. E-mail: yabae03@gmail.com

Parasitology (2015), **142**, 1228–1238. © Cambridge University Press 2015
doi:10.1017/S0031182015000566

https://doi.org/10.1017/S0031182015000566 Published online by Cambridge University Press

(32·25% *vs* 40·13%); the distributions of short interspersed nuclear elements and long-terminal-repeat (LTR) retrotransposons were greatly reduced, whereas those of other repeats including long interspersed nuclear elements were slightly greater (The *Schistosoma japonicum* Genome Sequencing and Functional Analysis Consortium, 2009; Huang *et al*. 2013). These genomic characters were reduced significantly in cestodes, which are phylogenetically more closely related to trematodes than other parasitic groups (Tsai *et al*. 2013). Taken together, these findings suggest that diverse platyhelminthes have developed different molecular mechanisms involved in chromosome/genome evolution.

DNA methylation is well known for its involvement in regulating gene expression, condensing chromatin and maintaining genome stability (reviewed in Law and Jacobsen, 2010). Methylation of the 5′ carbon of cytosine (5-methylcytosine, 5 mC) in cytosine- and guanine-rich DNA sequences is frequently associated with promoters and transposons. The 5 mC-dependent regulatory machinery is widely conserved in various animal and plant taxa, where DNA methyltransferase (DNMT) 1 and DNMT3 are responsible for the maintenance and establishment, respectively, of the genome-wide DNA methylation map (i.e. methylome). Meanwhile, the discovery of DNMT2-only or DNMT-free organisms has raised questions concerning the universal role of DNA methylation and/or the specific function(s) of DNMT2 (Jeltsch, 2010; Schaefer and Lyko, 2010; Raddatz *et al*. 2013).

Recently, cytosine methylation was investigated in *Schistosoma mansoni* (Geyer *et al*. 2011). The blood fluke was shown to possess functional DNA methylation machinery, where a single DNMT2-like protein (SmDNMT2) was responsible for the relevant enzymatic activity. The DNMT2-orthologous proteins, together with the methyl-CpG binding domain proteins (MBDs), are well conserved throughout the phylum Platyhelminthes (Geyer *et al*. 2013). It was also demonstrated that representative species of four platyhelminth classes contain 5 mC in their genomic DNAs. However, no definitive DNA methylation pattern was detected in a comprehensive methylome analysis of *S. mansoni*, suggesting that DNA methylation might be dispensable in the parasite (Raddatz *et al*. 2013). In an effort to address the controversial issue in platyhelminthes, we examined DNA methylation-related phenomena in *C. sinensis* expressing a single DNMT protein homologous to DNMT2 family members. The spatiotemporal expression pattern of *Clonorchis* DNMT2 (CsDNMT2), which was largely restricted to eggs, coincided well with the histological distribution of 5 mC and exhibited considerable methyltransferase activity toward DNA. These results support the notion that DNA methylation-dependent activity is functional primarily during early developmental stages of trematode parasites.

MATERIALS AND METHODS

Isolation of CsDNMT2

A partial 320-amino acid protein, annotated as CsDNMT2 (GAA54195), was isolated from the *C. sinensis* database of GenBank (http://www.ncbi.nlm.nih.gov/genbank/) through a series of BLAST searches using the amino acid sequences of human DNMTs (HsDNMT1, NP_001124295; HsDNMT2, NP_004403; HsDNMT3a, NP_783328; HsDNMT3b, NP_008823). A coding DNA sequence (CDS) corresponding to the CsDNMT2 fragment (963 bp) was extracted from a *C. sinensis* genomic scaffold (DF143779). Two primers were prepared from the upstream (reverse direction, 5′-CGGAGGGCTCA TGCTCCAC-3′) and downstream (forward direction, 5′-GACTTTTTGGACGATAAC-3′) regions of the CDS. The primers were used in combination with T3 and T7 promoter primers to amplify the 5′- and 3′-regions of *CsDNMT2* transcript by polymerase chain reaction (PCR) from a *C. sinensis* egg cDNA library (provided by Division of Malaria and Parasitic Diseases, Korea Centers for Disease Control and Prevention). The nucleotide sequences of both amplicons were determined and overlapped to obtain a cDNA contig. Integrity of the overlapped sequence was confirmed by amplifying the full open reading frame (ORF) of *CsDNMT2* in the cDNA library with an ORF-specific primer pair (CsDNMT-ORF-F, 5′-ATGCGTGTGCTGGA GTTGTATTC-3′ and CsDNMT-ORF-R, 5′-T CAGTTAGTCTTTTGTGCCGAC-3′).

Sequence analysis

The structure of chromosomal *CsDNMT2* gene was determined by aligning genomic and full-length cDNA sequences. The ORF was identified using the ORF Finder program at the NCBI web site (http://www.ncbi.nlm.nih.gov/gorf/gorf.html). The translated amino acid sequence was subject to a functional domain search against the InterPro database of protein families using InterProScan 5 (http://www.ebi.ac.uk/interpro/). The presence of an N-terminal hydrophobic sequence was examined by the SignalP 4·1 program (http://www.cbs.dtu.dk/services/SignalP/). The primary structure of the CsDNMT2 protein was compared with those of other platyhelminth orthologs, as well as those of previously characterized DNMT2s (Dong *et al*. 2001; Schultz *et al*. 2012; Li *et al*. 2013), using the Clustal X program (ver. 2·1).

Prediction of secondary and tertiary structures

The secondary and tertiary structures of CsDNMT2 were predicted by the I-TASSER program (ver. 3·0; http://zhanglab.ccmb.med.umich.edu/I-TASSER/), which combines threading, *ab initio* modelling and

https://doi.org/10.1017/S0031182015000566 Published online by Cambridge University Press

structural refinement (Zhang, 2008). The template modelling-score (TM-score) and root mean square deviation (RMSD) were calculated between CsDNMT2 and the reference models, and were then used to evaluate the quality of the predicted tertiary structure(s). The stereochemical quality of the models was validated by subjecting PDB files obtained by I-TASSER to the PROCHECK server (http://www.ebi.ac.uk/thornton-srv/databases/pdbsum/Generate.html). The resulting model was visualized with the PyMol program (DeLano, 2002).

Generation of recombinant CsDNMT2$_{Ab}$ and specific antiserum

A cDNA segment (612 bp) corresponding to the middle-region of CsDNMT2 was amplified from the egg cDNA library using a specific primer pair (5′-CAGAATTCGAGGTGACAGCATTCAATG C-3′ with *Eco*R I site and 5′-GACTCGAGACAA GGACGAACAATGTCCAGG-3′ with *Xho* I site). The PCR product was digested with *Eco*R I and *Xho* I, ligated into the *Eco*R I/*Xho* I site of pET-28a plasmid (Novagen, Madison, WI, USA), then introduced into *Escherichia coli* DH5α cells. The expression fidelity of transformed clones was validated by automated DNA sequencing. The plasmid was expressed in *E. coli* BL21 (DE3) cells in the presence of 0·5 mM isopropyl-β-D-thiogalactopyranoside for 4 h at 37 °C. The recombinant protein (rCsDNMT2$_{Ab}$) was purified under denaturing conditions by nickel-nitrilotriacetic acid (Ni-NTA) agarose chromatography (Qiagen, Valencia, CA, USA). The eluents were examined by 12% SDS-PAGE under reducing conditions.

Purified rCsDNMT2$_{Ab}$ (30 μg) was mixed with the complete (for the first injection) or incomplete (for the second and third injections) Freund's adjuvants (Sigma-Aldrich, St. Louis, MO, USA) and subcutaneously injected into BALB/*c* mice three times at 2-week intervals. The mice were finally boosted through the tail vein with 10 μg protein and sacrificed 7 days later. The blood was collected by heart puncture, centrifuged for 10 min at 3000 ***g*** at 4 °C and stored at −80 °C until use. The protocol was approved by the Institutional Review Board and conducted in the Laboratory Animal Research Center of Sungkyunkwan University, Korea (protocol 13–17).

Preparation of parasitic materials

Adult *C. sinensis* worms were isolated from the hepatobiliary tract of experimentally infected rats, as described in a previous report (Bae *et al*. 2013*b*). The worms were homogenized in Triton Extraction Buffer (TEB: PBS containing 0·5% [v/v] Triton X-100 and a protease inhibitor cocktail [complete; Roche Diagnostics GmbH, Mannheim, Germany]) using a Dounce tissue grinder (Wheaton, Millville, NJ, USA). The homogenate was centrifuged at 10 000 ***g*** for 20 min at 4 °C and the supernatant was stored as a cytosolic protein extract (A-CyE). After washing with TEB, the nuclei pellet was suspended in 0·2 N HCl and incubated overnight at 4 °C. The solution was centrifuged at 20 000 ***g*** for 5 min at 4 °C. The resulting supernatant was designated as a nuclear protein extract (A-NuE).

Another set of cytosolic and nuclear protein extracts were prepared without use of any probable denaturant such as detergent and HCl following the experimental protocol described by Sigma-Aldrich (http://www.sigmaaldrich.com/). Briefly, frozen *C. sinensis* adults were ground to powder on liquid nitrogen using a mortar and pestle. Worm powder was dissolved into a lysis buffer (10 mM HEPES, pH 7·9, with 1·5 mM $MgCl_2$ and 10 mM KCl) containing 1 mM DTT and a protease inhibitor cocktail (Complete; Roche Diagnostics GmbH, Mannheim, Germany). After centrifugation at 10 000 ***g*** for 20 min at 4 °C, the supernatant was taken as a cytosolic extract (W-CyE). The nuclei pellet was suspended in an extraction buffer (20 mM HEPES, pH 7·9, with 1·5 mM $MgCl_2$, 0·42 M NaCl, 0·2 mM EDTA and 25% [v/v] glycerol) with 1 mM DTT and the protease inhibitors. The solution was shaken gently at 4 °C for 30 min and centrifuged at 20 000 ***g*** at 4 °C for 5 min. The resulting supernatant (nuclear extract, W-NuE) and the cytosolic extract were dialyzed against PBS (pH 7·0).

Live *C. sinensis* adults were incubated in phenol red-free RPMI-1640 (pH 7·2) at 37 °C overnight. Eggs were collected from the medium and examined under a dissection microscope to remove any fragment from dead worm bodies. After washing with PBS (pH 7·0) more than five times, the eggs were ground under liquid nitrogen in a mortar and pestle. Total RNA, genomic DNA and proteins were extracted from the egg powder using the TriZol reagent according to the manufacturer's instructions (Invitrogen, Carlsbad, CA, USA). A similar approach was used to isolate RNA and DNA from adult *C. sinensis* worms. Genomic DNA was also extracted from normal rat liver as positive control using the Wizard Genomic DNA purification Kit (Promega, Madison, WI, USA).

Examination of native CsDNMT2 expression

Total RNAs (1 μg) extracted from *C. sinensis* eggs and adults were reverse-transcribed into single-stranded cDNAs using a RNA PCR kit (AMV, ver. 3·0; Takara, Shiga, Japan) in a 10 μL reaction volume. The cDNAs (1 μL) were used to amplify *CsDNMT2* using the gene-specific primer pairs (5′-GAGCATGAGCCCTCCGTGTCAAC-3′ and 5′-CAACGGGACTTACGATCACAAGG-3′ for *CsDNMT2$_{Ab}$*; 5′-ATGCGTGTGCTGGAGTT

https://doi.org/10.1017/S0031182015000566 Published online by Cambridge University Press

GTATTC-3′ and 5′-TCAGTTAGTCTTTTGT GCCGAC-3′ for $CsDNMT2_{ORF}$) under the following conditions: 2 min at 94 °C (preheat); 50 s at 94 °C, 50 s at 60 °C, 1·5 min at 72 °C (26 cycles); 5 min at 72 °C (final extension). Primers specific to a tropomyosin gene (*CsTrop*, L43918; 5′-TGAGT CTCGTCTAGAAGCTGCTG-3′ and 5′-GGT GAAATACGTAGGTTTGAACAC-3′) were used as an internal control.

Protein extracts (50 μg) from eggs and adults (A-CyE and A-NuE) were separated on 12% SDS-PAGE gels under reducing conditions. The protein bands were transferred onto nitrocellulose membranes (Schleicher & Schuell Bioscience, Dassel, Germany) and the membranes were incubated with the anti-$rCsDNMT2_{Ab}$ antibody (1:1000 dilution). Positive reactions were visualized with horseradish peroxidase (HRP)-conjugated rabbit anti-mouse IgG antibody (Bethyl Laboratories, Montgomery, TX, USA) and an enhanced chemiluminescence (ECL) detection system (GE Healthcare, Pittsburgh, PA, USA). The proteins (10 μg) were similarly examined by the Western blot analysis with an anti-histone H3 antibody (CT, Pan, clone A3S; Millipore, Billerica, Massachusetts, USA).

Whole mounts of *C. sinensis* adults were prepared as previously described (Bae *et al*. 2013*b*), then incubated with the mouse anti-$rCsDNMT2_{Ab}$ antibody overnight at 4 °C followed by a fluorescein isothiocyanate (FITC)-conjugated goat anti-mouse IgG antibody (Sigma-Aldrich). The slides were counterstained with 4′,6-diamidino-2-phenylindole (DAPI, 10 mg mL^{-1}; Invitrogen) for 5 min at 4 °C in the dark and observed under a fluorescence microscope (IX-70; Olympus, Tokyo, Japan). The worm sections were also incubated with a pooled normal mouse serum (negative control) or a 5 mC-specific antibody (33D3; Epigentek, Farmingdale, NY, USA).

Detection of 5 mC in the genomic DNA of C. sinensis

Genomic DNA (5 μg) prepared from adult worms was digested with *Msp* I or *Hpa* II, fractionated on 1% agarose gels and then either visualized by ethidium bromide staining or processed for Southern blotting. The blots were hybridized with *CsRn1* LTR probe enzymatically labelled with an ECL Direct Labelling Kit (Amersham Pharmacia Biotech) and signals were visualized with the ECL Detection Kit and X-ray film.

The presence of 5 mC was quantitatively examined in the genomic DNA of *C. sinensis* eggs and adults (100 and 400 ng/well) using a methylated DNA quantification kit (MethylFlash™; Epigentek) according to the manufacturer's instructions. Rat genomic DNA and pET-28a plasmid, that had been linearized by *Eco*R I digestion, were included in the analysis as positive and negative controls, respectively. The assay was performed in triplicate and the relative amounts of the 5 mC were presented as the mean ± standard deviation (S.D.) of absorbance at 450 nm.

DNA methyltransferase activity of CsDNMT2

DNA fragment corresponding to the ORF region of *CsDNMT2* was amplified with the ORF-specific primers to generate the full-length recombinant CsDNMT2 protein ($rCsDNMT2_{ORF}$). The protein expressed in inclusion bodies was refolded via a series of dialysis against decreasing concentrations of guanidine hydrochloride solution, as previously described (Thomson *et al*. 2012). The refolded protein was purified by Ni-NTA agarose chromatography and finally dialyzed against PBS (pH 7·0).

Presence of CsDNMT2 and histone H3 in the whole worm extracts (W-CyE and W-NuE; each 20 μg) was examined by Western blot analysis. The protein samples (20 μg/reaction) and $rCsDNMT2_{ORF}$ (0·1, 0·2 and 0·4 μg/reaction) was used in the measurement of the DNMT activity using a DNA methyltransferase activity assay kit (EpiQuik™; Epigentek). The mean ± S.D. of absorbance at 450 nm, which was determined from triplicate reactions, was expressed as a relative enzyme activity.

RESULTS

Isolation of C. sinensis gene homologous to DNMT2

A BLASTp search with HsDNMT2 enabled retrieval of a partial 320-aa sequence (GAA54195) from the *Clonorchis* proteome database, which had been annotated as DNA (cytosine-5-)-methyltransferase (coverage 83%, *E*-value 8e-58, and identity 38%). The HsDNMT1 and HsDNMT3 sequences did not have any significant match in the database. Whole genome screening using HsDNMTs and the *Clonorchis* sequence by tBLASTn also isolated only a single nucleotide sequence (BADR02002681) containing the putative *Clonorchis* DNMT2 gene. Based on the nucleotide information, a full-length cDNA sequence was isolated from an egg cDNA library of *C. sinensis* and named *CsDNMT2*. The *CsDNMT2* ORF was 1143 bp in length encoding a 380-aa protein. Consistent with the BLAST results, functional analysis of the protein by InterProScan identified the CsDNMT2 as a C-5 cytosine methyltransferase (InterPro; IPR001525), a DNA_methylase (PFAM; PF00145) and a C5METTRFRASE (PRINTS; PR00105). The *CsDNMT2* sequence was deposited in the GenBank with the accession number KM491714.

Structural characterization of CsDNMT2

The amino acid sequence of CsDNMT2 was aligned with those of DNMT2 proteins in protozoa, platyhelminthes and human (Fig. 1). All sequences examined shared well-conserved motifs found in the

https://doi.org/10.1017/S0031182015000566 Published online by Cambridge University Press

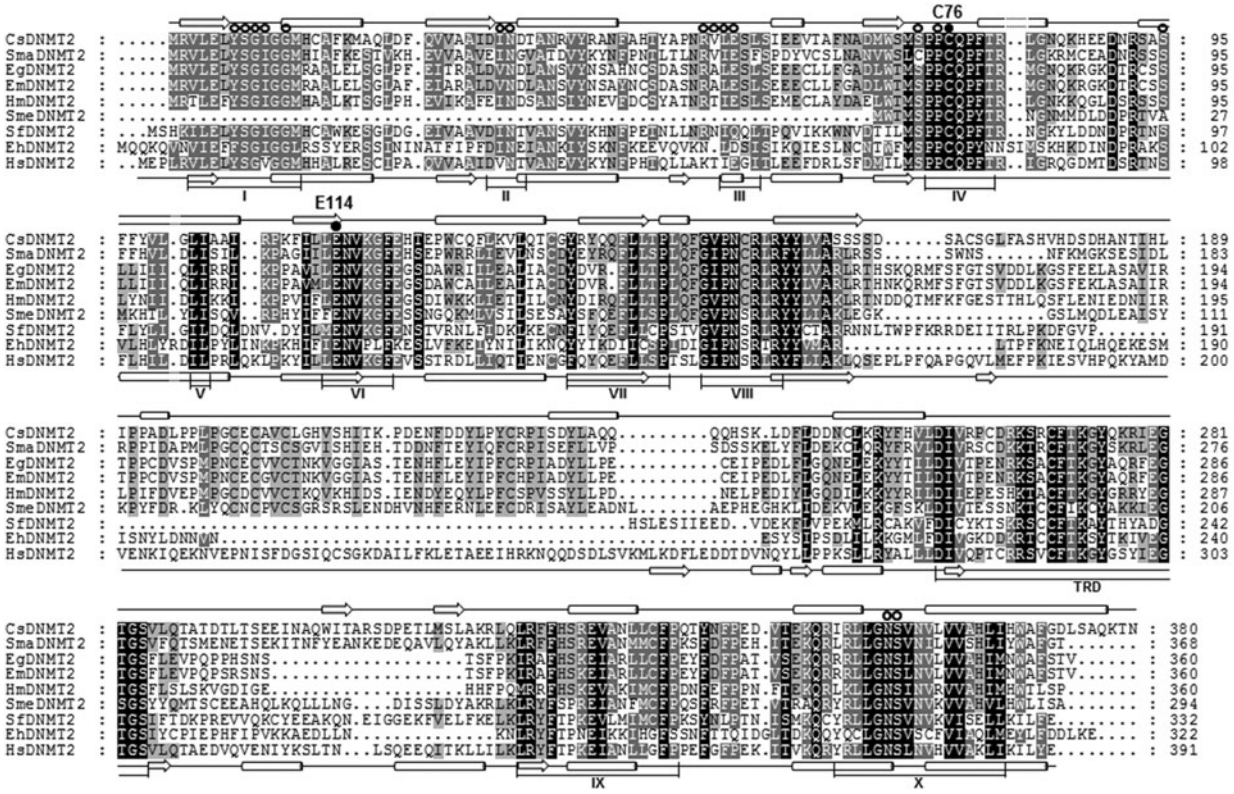

Fig. 1. Sequence alignment of *Clonorchis sinensis* DNMT2 with its orthologs. Dots represent gaps introduced during the alignment. Different shades of grey indicate the degrees of similarity among sites. Secondary structural elements (rods for α-helix and arrows for β-strand) of CsDNMT2 and HsDNMT2 are presented at the top and bottom of the alignment, respectively. Conserved sequence motifs of the DNMT2 family are marked with Roman numerals (I–X). Amino acids involved in substrate binding are distinguished with open circles. Closed circles mark cysteine and glutamic acid acting as a catalytic nucleophile in DNA (C76) and RNA (E114) methyltransferases, respectively. TRD indicates the target recognition domain. The sequences used in the alignment were retrieved from the GenBank database: EgDNMT2, *Echinococcus granulosus* (CDJ23387); EhDNMT2, *Entamoeba histolytica* (XP_655267); EmDNMT2, *Echinococcus multilocularis* (CDI70800); HmDNMT2, *Hymenolepis microstoma* (CDJ13147); HsDNMT2, *Homo sapiens* (NP_004403); SfDNMT2, *Spodoptera frugiperda* (AFS64716); SmaDNMT2, *Schistosoma mansoni* (XP_002575636); SmeDNMT2, *Schmidtea mediterranea* (AHB82105).

DNMT2 protein family (motif I–X), even though hypervariable regions were observed between motif VIII and the target recognition domain. The aa residues involved in substrate binding could be recognized at their respective positions (open circles in Fig. 1; Dong *et al*. 2001). Cysteine and glutamic acid, which are known to act as catalytic nucleophiles for the initial attack of DNA methyltransferase and RNA methyltransferase, respectively (Jeltsch *et al*. 2006; Jurkowski *et al*. 2008), were conserved in motifs IV and VI (C7 and E114 marked by closed circles). The SignalP program could not recognize an N-terminal hydrophobic sequence. The exon-intron architectures of the DNMT2 genes were also highly comparable among the platyhelminth and human homologs; five introns of the *HsDNMT2* gene were orthologous to those of platyhelminth genes (62·5%; Fig. 2A).

A total of three models were suggested for the tertiary structure of CsDNMT2 protein by I-TASSER using multiple threading templates including *Spodoptera frugiperda* (PDB no. 4h0nA), *Entamoeba histolytica* (PDB no. 3qv2A) and human (PDB no. 1g55A) DNMT2s (Supplementary Table S1). The confidence scores (C-score) of the simulated models were found to be −0·53 (model 1), −0·58 (model 2) and −0·59 (model 3). The TM-score and RMSD of model 1 showing the highest confidence were determined to be 0·65 ± 0·13 and 7·9 ± 4·4 Å, respectively. A Ramachandran plot was generated by the PDBsum server to evaluate the stereochemical quality of model 1. As shown in Fig. 2B, amino acids residing in most favoured regions, additional allowed regions, generously allowed regions and disallowed regions were estimated to be 85·4% (292 residues), 10·5% (36 residues), 2% (7 residues) and 2% (7 residues), respectively (see legend to Fig. 2B). The statistical quality of the simulated structure was substantially enhanced when HsDNMT2 was used as the guide template, which occupied the top of the PDB enzyme homolog list obtained by functional prediction (C-score 0·23, TM-score 0·74 ± 0·11 and RMSD 6·2 ± 3·8 Å). The tertiary structure of CsDNMT2 was

https://doi.org/10.1017/S0031182015000566 Published online by Cambridge University Press

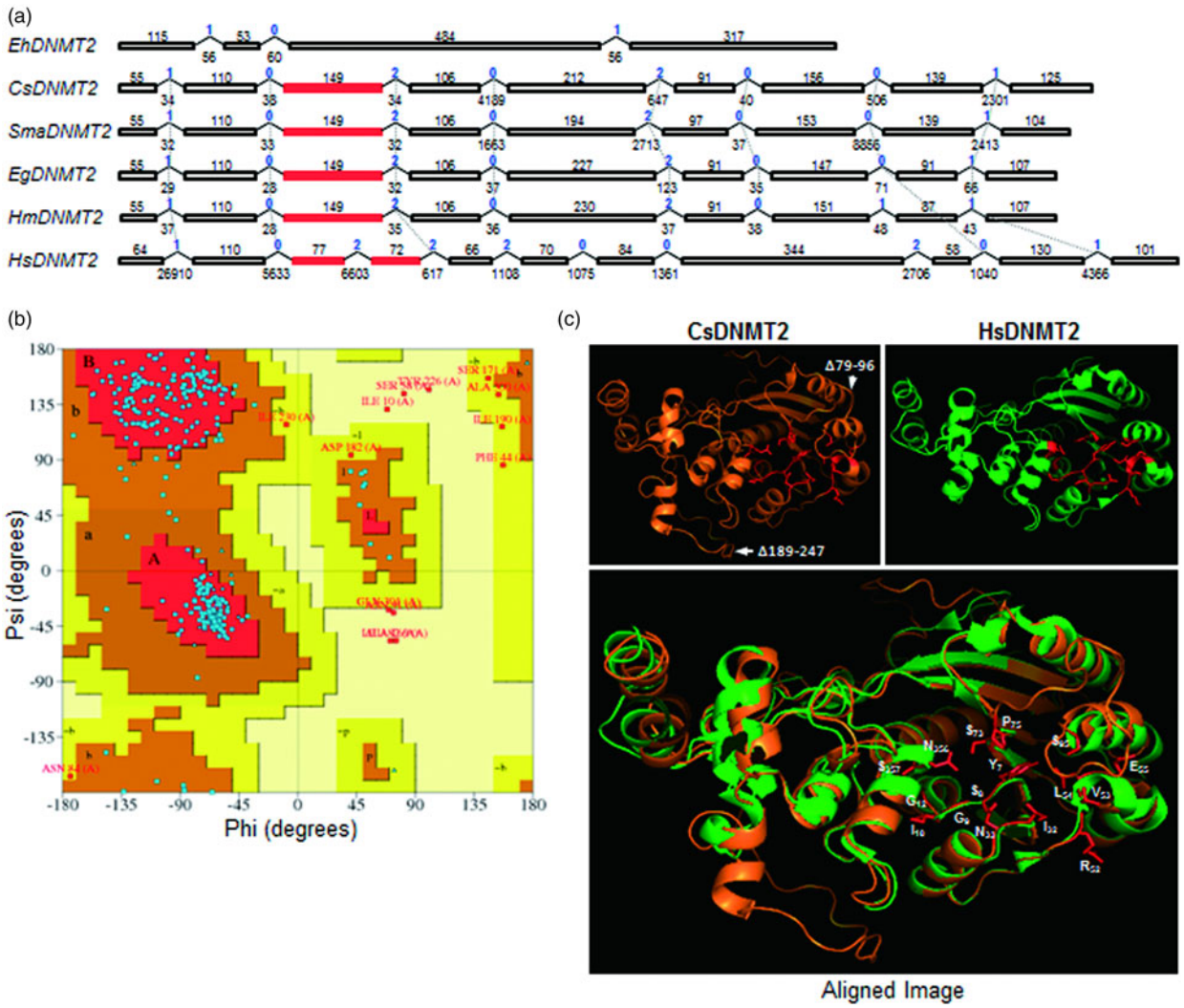

Fig. 2. Structural conservation of *Clonorchis sinensis* DNMT2 gene (*CsDNMT2*) and its protein product. (A) Exon-intron architecture of platyhelminth DNMT2 genes (*SmaDNMT2* of *Schistosoma mansoni; EgDNMT2* of *Echinococcus granulosus; HmDNMT2* of *Hymenolepis microstoma*) were compared with those of protozoan (*EhDNMT2* of *Entamoeba histolytica*) and human (*HsDNMT2*) genes. The coding sequences (CDS) are presented with open squares in proportion to their relative sizes and the intervening introns are marked by indentations with a fixed length. The phase of each intron (blue Arabic numerals), as well as the length of the exons and introns in base pairs, is also marked at the corresponding position. The dotted vertical lines connect introns occupied in an orthologous position. Red squares mark the exons of platyhelminth genes, which have been further split in the vertebrate orthologs. (B) A Ramachandran plot was generated to evaluate the stereochemical quality of the predicted CsDNMT2 tertiary structure. Most favoured regions (A, B and L) are coloured in red, while additional allowed (a, b, l and p), generously allowed (-a, -b, -l and –p) and disallowed regions are shown as brown, yellow and light yellow, respectively. Blue squares and triangles represent non-glycin and glycin residues in most favoured and additional allowed regions. Red squares mark non-glycin residues in generously allowed and disallowed regions. (C) The tertiary structures of CsDNMT2 (ocher) and HsDNMT2 (yellowish green) were predicted with the I-TASSER program and aligned with each other using the PyMol program. The amino acids composing the substrate binding site are marked in red. White arrows in the CsDNMT2 model indicate amino acid stretches corresponding to those of human protein that were deleted during crystallization (Dong *et al*. 2001).

well aligned with that of HsDNMT2, where aa residues constituting the substrate-binding site of each molecule were similarly spaced (Fig. 2C).

Examination of CsDNMT2 expression and chromosomal 5 mC

The *CsDNMT2* transcripts were readily amplified by RT-PCR using *C. sinensis* egg RNA, whereas the mRNA was substantially reduced in a whole adult worm sample (Fig. 3A). The differential expression of *CsDNMT2* between egg and adult stages was further evidenced by Western blot analysis of *C. sinensis* proteins with the rCsDNMT2$_{Ab}$-specific antibody (Fig. 3B). A strong positive signal was detected between 40 and 50 kDa on the lane loaded with the egg proteins, of which molecular weight was comparable with that

https://doi.org/10.1017/S0031182015000566 Published online by Cambridge University Press

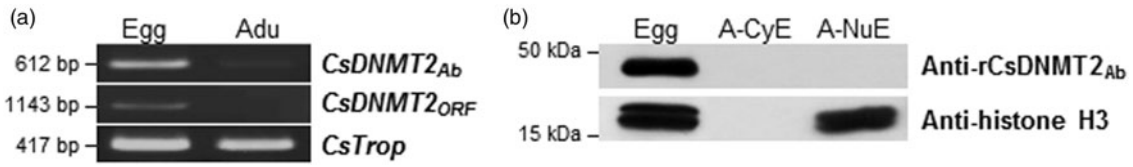

Fig. 3. Expression pattern of *CsDNMT2*. (A) The relative expression level of *CsDNMT2* was examined by reverse transcription PCR, using total RNAs extracted from *Clonorchis sinensis* eggs and adults (Adu). $CsDNMT2_{Ab}$ corresponded to a middle region of the *CsDNMT2* coding sequence (positions 181–792), while $CsDNMT2_{ORF}$ covered the full coding sequence. Amplification of *CsTrop* was included as an internal control. (B) Presence of native CsDNMT2 in the egg and adult protein extracts was determined by Western blot analysis with the anti-rCsDNMT2$_{Ab}$ antibody. The cytosolic (A-CyE) and nuclear (A-NuE) proteins of adult *C. sinensis*, which were extracted using the Dounce homogenizer to minimize contamination by egg materials, were used for the examination. Antibody specific to histone H3 was applied in the examination as a control.

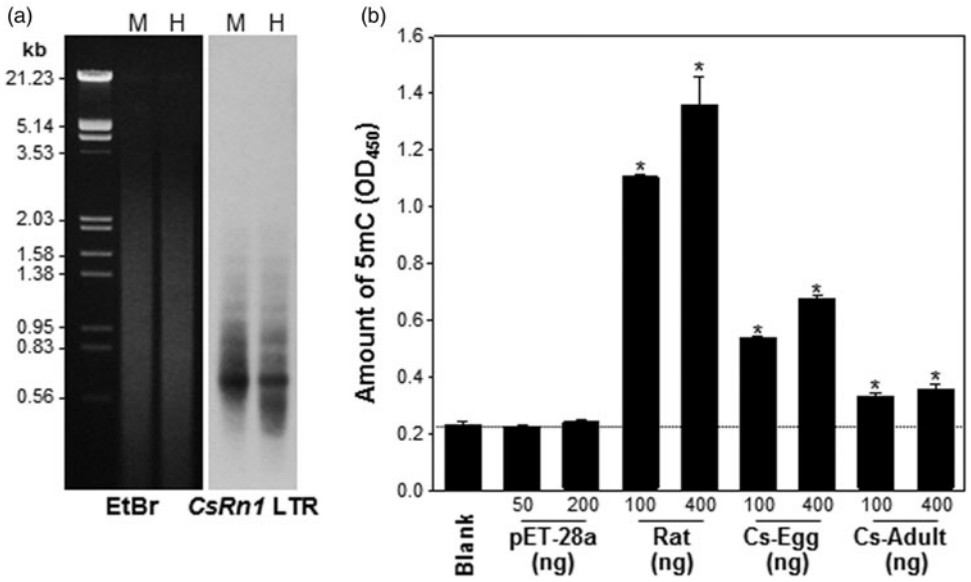

Fig. 4. Detection of 5-methylcytosine (5 mC) in the chromosomal DNA of *Clonorchis sinensis*. (A) Adult genomic DNA was digested with the 5 mC-sensitive *Hpa* II (H) or 5 mC-insensitive *Msp* I (M) isoschizomers. After electrophoresis on 1% agarose gels, the restriction fragments were visualized by ethidium bromide staining (EtBr) or hybridized with the long-terminal-region sequence of *CsRn1* (*CsRn1* LTR) in a Southern blot. The signals were developed using an ECL detection system. DNA size standards are shown on the left. (B) Relative abundance of 5 mC in the genomic DNAs of *C. sinensis* eggs and adults was assayed using ELISA. Relative amounts of 5 mC were estimated as the average absorbance at 450 nm ± standard deviation from triplicate reactions. Rat and pET-28a plasmid DNAs were included as positive and negative controls. Concentrations of DNAs added in the measurements are presented in ng at the bottom of the histograms. Asterisks indicate reactions with one-tailed p values less than 0·0001 based on Student's t test compared with the reactions without DNA.

of CsDNMT2 (43·3 kDa), however, lanes with the adult protein extracts (A-CyE and A-NuE) did not show any considerable reactivity with the antibody. An anti-histone H3 antibody was found to react similarly with the egg proteins and adult nuclear proteins (lower panel of Fig. 3B).

Restriction patterns of adult genomic DNA with *Msp* I and *Hpa* II were similar when examined by ethidium bromide staining or Southern blot analysis with the *CsRn1* LTR probe (Fig. 4A). The presence of 5 mC in the genomic DNAs of *C. sinensis* adults and eggs was further examined using the methylated DNA quantification assay kit. Rat genomic DNA and pET-28a plasmid DNA were used as positive and negative controls, respectively. The parasite DNAs exhibited considerable reactivity against the 5 mC-specific antibody in the ELISA-based analysis (P-values of one-tailed t-test <0·0001; Fig. 4). The signal was dose-dependent relative to DNA concentration and substantially higher in egg-stage DNA. Rat DNA was also found to be highly reactive against the antibody, whereas that of the bacterial plasmid was negligible. These findings coincided well with the histological localities of CsDNMT2

https://doi.org/10.1017/S0031182015000566 Published online by Cambridge University Press

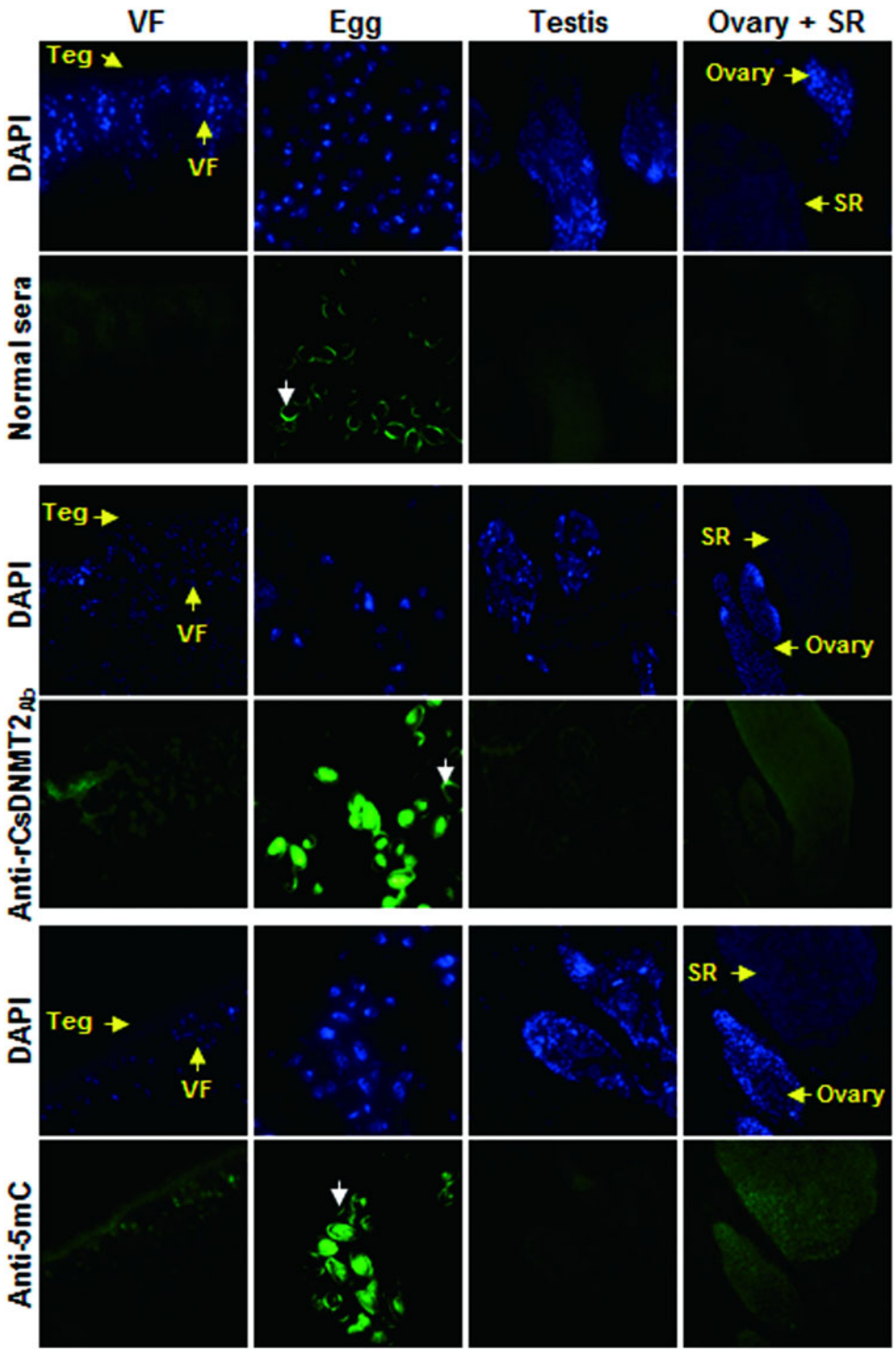

Fig. 5. Histological distributions of CsDNMT2 and 5-methylcytosine (5 mC) in adult *Clonorchis sinensis* sections. The adult worm sections were incubated with antibody specific to CsDNMT2 or 5 mC. Positive binding was visualized by staining the sections with a fluorescein isothiocyanate (FITC)-conjugated secondary antibody. A pooled serum obtained from normal mice was used as a negative control and nuclei were counter-stained with 49,6-diamidino-2-phenolindole (DAPI). SR, seminal receptacle; Teg, tegument; VF, vitelline follicle. White arrows in FITC-stained images indicate epifluorescent eggshells.

and 5 mC in adult worms. Both molecules appeared to be largely restricted within intrauterine eggs, while some marginal signals were detected in vitelline follicles, ovary and the seminal receptacle (Fig. 5).

Methyltransferase activity of rCsDNMT2$_{ORF}$

Since native CsDNMT2 was not detected in the adult proteins extracted using the Dounce homogenizer (Fig. 3B), another set of protein extracts (W-CyE and W-NuE) was prepared from adult worms

https://doi.org/10.1017/S0031182015000566 Published online by Cambridge University Press

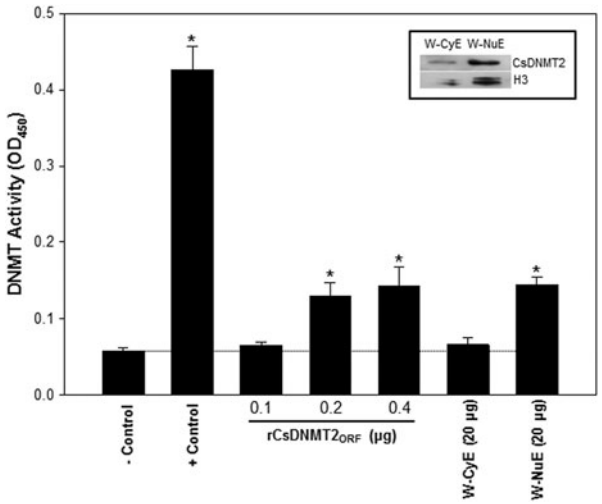

Fig. 6. DNA methyltransferase activity of $rCsDNMT2_{ORF}$. The enzymatic activity of $rCsDNMT2_{ORF}$ was measured against a DNA substrate. Reactions with the cytosolic (W-CyE) and nuclear (W-NuE) proteins of *Clonorchis sinensis* adults, which were prepared by a liquid nitrogen-based method, were included in the assay, as well as those with no protein (−control) or those with DNMT1 (+control). Inset panel shows the Western blot results of W-CyE and W-NuE with the CsDNMT2- and H3-specific antibodies. The relative activity is represented as an average absorbance at 450 nm ± standard deviation from triplicate reactions. Asterisks indicate reactions with one-tailed p values less than 0·0001 based on Student's t test compared with the negative control.

homogenized with a mortar and pestle on liquid nitrogen. The acid-extraction step was excluded to prevent protein denaturation. This nuclear extract contained a substantial amount of native CsDNMT2, although histone H3 was greatly reduced likely resulting from using a more neutral extraction procedure (inset panel in Fig. 6). The $rCsDNMT2_{ORF}$ protein expressed in the inclusion bodies of *E. coli* cells was refolded under *in vitro* conditions to obtain a bioactive form. The protein was used in the determination of DNA methyltransferase activity after purification, together with the W-CyE and W-NuE. As shown in Fig. 6, the 5 mC-specific antibody was highly reactive with DNAs that were pre-incubated with the refolded $rCsDNMT2_{ORF}$ or with the nuclear extract (W-NuE) of *C. sinensis*, as well as control DNMT protein included in the assay kit (+control). The signal intensity increased relative to increasing $rCsDNMT2_{ORF}$ added during the pre-incubation step. Similar to the results from Western blot analysis (inset panel), substrate DNA mixed with the cytosolic proteins (W-CyE) exhibited minor reactivity with the 5 mC-specific primary antibody.

DISCUSSION

Some structural features of partial CsDNMT2 (GAA54195) have recently been described by Geyer *et al.* (2013). In the present study, we determined the full-length *CsDNMT2* gene and proposed a three-dimensional model for the protein. The DNA methyltransferase activity and spatio-temporal expression/distribution pattern of the *Clonorchis* protein were also examined in association with that of 5 mC. The protein shared structural properties with DNMT2 family proteins isolated from various organisms. The expression pattern of CsDNMT2 coincided well with the histological locality of 5 mC, both of which were largely restricted within uterine eggs and only slightly presented in other tissues including seminal receptacle. Comprehensive screening of the parasite's whole genome using the sequences of CsDNMT2 and HsDNMTs demonstrated that only the CsDNMT2 protein might provide enzymatic activity responsible for the conversion of cytosine into 5 mC in the chromosomal DNA of *C. sinensis*.

5mC and related molecules involved in gene silencing and suppression of transposable elements seem to have evolved in the last eukaryotic common ancestor (Jurkowski and Jeltsch, 2011). DNMT3 and DNMT1 are believed to be responsible for the establishment of organism-specific methylomes, although different sets of these enzymes have been identified in several organisms (Zemach and Zilberman, 2010). However, DNMT2-only or DNMT-free organisms contradicted the widely accepted role of DNA methylation in eukaryotes (Jeltsch, 2010; Raddatz *et al.* 2013). The genomes of parasitic platyhelminthes examined here were also found to encode a single DNMT2 protein (Fig. 1; Geyer *et al.* 2013). Geyer *et al.* (2011) described detection of chromosomal 5 mC and its biological implication, especially those associated with the oviposition processes, in *S. mansoni*. However, no clearly defined pattern such as CpG specificity was recognized in the DNMT2-dependent methylome of the blood fluke, which might propose that 5 mC does not function in the regulation of gene expression (Raddatz *et al.* 2013). The CsDNMT2 protein and chromosomal 5 mC were largely localized within eggs in adult *C. sinensis* (Fig. 5). Adult proteins extracted using the Dounce homogenizer, which rarely disrupts eggshells, did not contain any detectable amount of CsDNMT2 (Fig. 3B). Therefore, the detection of 5 mC (Fig. 4) and DNMT activity (Fig. 6) in adult-derived DNA and protein, respectively, are likely due to the presence of egg DNAs and proteins that were released during the isolation steps. Considering that the schistosome DNA used in the previous methylome analysis was prepared from adult males (Raddatz *et al.* 2013), DNMT2 orthologs appear to express their methylation activity during egg stages in parasitic trematodes.

DNA methylation is also responsible for the prevention of intra-genomic multiplication of Class I

https://doi.org/10.1017/S0031182015000566 Published online by Cambridge University Press

transposable elements (i.e. retrotransposons). The first retrotransposition step, transcription, is blocked by a 5 mC-mediated molecular mechanism within or near the promoter of autonomous elements (Yoder *et al*. 1997). In mammals, retrotransposons expand with increased transcriptional activity, after genome-wide DNA methylation is abolished during an early stage of embryogenesis (Kano *et al*. 2009). The chromosomal DNA of *C*. *sinensis* adults exhibited a similar susceptibility to the methylation-insensitive *Msp* I and methylation-sensitive *Hpa* II isoschizomers. Southern blot analysis of the DNA fragments with the LTR probe of *CsRn1*, an active *Gypsy*-like LTR retrotransposon residing in the *C*. *sinensis* genome (Bae *et al*. 2001; Bae and Kong, 2003), also demonstrated the negligible level of DNA methylation in the *CsRn1* promoter region (Fig 4A). Egg DNA, which showed a strong 5 mC-specific signal (Figs 4B and 5), could not be examined by the restriction pattern analysis due to insufficient egg DNA. Methylome analysis of egg DNA and identification of target genes with defined DNA methylation patterns would help address this issue in DNMT2-only platyhelminthes.

Despite structural conservation among DNMT families, the role of DNMT2 has been highly enigmatic, mainly due to gene mutants displaying only minor phenotypes. The role of DNMT2 as a DNA methyltransferase in DNMT1/DNMT3-deficient eukaryotes has also been challenged by its tRNA-biased substrate specificity (Goll *et al*. 2006; Schaefer and Lyko, 2010). DNMT proteins are subject to posttranslational modification including methylation and phosphorylation to increase their stability and/or activity (Kinney and Pradhan, 2011). However, the biochemical properties of DNMT2 homologs presently available were obtained largely by using recombinant proteins expressed in *E*. *coli* cells (Fisher *et al*. 2004; Li *et al*. 2013; Müller *et al*. 2013). It is possible that posttranslational modification events induce conformational changes to affect positions of catalytic nucleophiles (i.e. C76 for DNA and E114 for RNA; Fig. 1), which alters substrate preference (Jeltsch *et al*. 2006; Jurkowski *et al*. 2008). For example, examination of the CsDNMT2 sequence with NetPhos (http://www.cbs.dtu.dk/services/NetPhos/) suggested multiple serine, threonine and tyrosine residues as targets for phosphorylation (data not shown). Alternatively, the general plasticity of the target recognition could be dependent on the surrounding conditions including the absence/presence of specific cofactor protein(s). DNMT1 associated protein (GAA32832; Rountree *et al*. 2000) and methyl binding domain protein (MBD) 1 (GAA50068)/MBD2 (GAA57483) encoded in the DNMT2-only *Clonorchis* genome could be candidates for the future examination. Investigations on the substrate preference and effects of transacting proteins with rCsDNMT2 prepared in a eukaryotic system could provide evidences to solve these controversial issues.

In conclusion, although we have only limited data on the substrate preference of CsDNMT2 and other platyhelminth orthologs, we conclude that CsDNMT2 functions as a DNA methyltransferase based on the following: (1) CsDNMT2 exhibits sequence homology to members of DNMT family; (2) the spatiotemporal expression pattern of the protein corresponds well with the presence of 5 mC; (3) the protein is significantly accumulated in the nuclear portion of the parasite's cells and; (4) rCsDNMT2 shows methyltransferase activity upon exposure to DNA substrate. The mature gamete- and egg-specific expression of CsDNMT2 further suggests its role during embryogenesis. Future investigations of trematode DNMT2 activity may provide a target for controlling parasitic diseases such as clonorchiasis that pose great public health burdens in endemic regions.

SUPPLEMENTARY MATERIAL

To view supplementary material for this article, please visit http://dx.doi.org/S0031182015000566.

FINANCIAL SUPPORT

This work was supported by the Gachon University research fund of 2012 (GCU-2012-M081 to A-YB).

REFERENCES

Bae, Y. A., Ahn, D. W., Lee, E. G., Kim, S. H., Cai, G. B., Kang, I., Sohn, W. M. and Kong, Y. (2013*a*). Differential activation of diverse glutathione transferases of *Clonorchis sinensis* in response to the host bile and oxidative stressors. *PLoS Neglected Tropical Diseases* **7**, e2211.

Bae, Y. A., Cai, G. B., Kim, S. H., Sohn, W. M. and Kong, Y. (2013*b*). Expression pattern and substrate specificity of *Clonorchis sinensis* tyrosinases. *International Journal for Parasitology* **43**, 891–900.

Bae, Y. A. and Kong, Y. (2003). Evolutionary course of *CsRn*1 long-terminal-repeat retrotransposon and its heterogeneous integrations into the genome of the liver fluke, *Clonorchis sinensis*. *Korean Journal of Parasitology* **41**, 209–219.

Bae, Y. A., Moon, S. Y., Kong, Y., Cho, S. Y. and Rhyu, M. G. (2001). *CsRn*1, a novel retrotransposon in a parasitic trematode, *Clonorchis sinensis*, discloses a new phylogenetic clade of *Ty3/gypsy*-like LTR retrotransposons. *Molecular Biology and Evolution* **18**, 1474–1483.

Bouvard, V., Baan, R., Straif, K., Grosse, Y., Secretan, B., El Ghissassi, F., Benbrahim-Tallaa, L., Guha, N., Freeman, C., Galichet, L. and Cogliano, V., WHO International Agency for Research on Cancer Monograph Working Group (2009). A review of human carcinogens - Part B: biological agents. *Lancet Oncology* **10**, 321–322.

Cai, G. B., Bae, Y. A., Kim, S. H., Sohn, W. M., Lee, Y. S., Jiang, M. S., Kim, T. S. and Kong, Y. (2008). Vitellocyte-specific expression of phospholipid hydroperoxide glutathione peroxidases in *Clonorchis sinensis*. *International Journal for Parasitology* **38**, 1613–1623.

DeLano, W. L. (2002). *The PyMOL Molecular Graphics System*. DeLano Scientific, San Carlos, CA, USA.

Dong, A., Yoder, J. A., Zhang, X., Zhou, L., Bestor, T. H. and Cheng, X. (2001). Structure of human DNMT2, an enigmatic DNA methyltransferase homolog that displays denaturant-resistant binding to DNA. *Nucleic Acids Research* **29**, 439–448.

Fisher, O., Siman-Tov, R. and Ankri, S. (2004). Characterization of cytosine methylated regions and 5-cytosine DNA methyltransferase (Ehmeth) in the protozoan parasite *Entamoeba histolytica*. *Nucleic Acids Research* **32**, 287–297.

https://doi.org/10.1017/S0031182015000566 Published online by Cambridge University Press

Geyer, K. K., Chalmers, I. W., Mackintosh, N., Hirst, J. E., Geoghegan, R., Badets, M., Brophy, P. M., Brehm, K. and Hoffmann, K. F. (2013). Cytosine methylation is a conserved epigenetic feature found throughout the phylum Platyhelminthes. *BMC Genomics* **14**, 462.

Geyer, K. K., Rodríguez López, C. M., Chalmers, I. W., Munshi, S. E., Truscott, M., Heald, J., Wilkinson, M. J. and Hoffmann, K. F. (2011). Cytosine methylation regulates oviposition in the pathogenic blood fluke *Schistosoma mansoni*. *Nature Communications* **2**, 424.

Goll, M. G., Kirpekar, F., Maggert, K. A., Yoder, J. A., Hsieh, C. L., Zhang, X., Golic, K. G., Jacobsen, S. E. and Bestor, T. H. (2006). Methylation of $tRNA^{Asp}$ by the DNA methyltransferase homolog Dnmt2. *Science* **311**, 395–398.

Hong, S. T. (2003). *Clonorchis sinensis*. In *International Handbook of Foodborne Pathogens* (Eds. Miliotis, M. D. and Bier, J. W.), pp. 581–592, Marcel Dekker Inc., New York.

Huang, Y., Chen, W., Wang, X., Liu, H., Chen, Y., Guo, L., Luo, F., Sun, J., Mao, Q., Liang, P., Xie, Z., Zhou, C., Tian, Y., Lv, X., Huang, L., Zhou, J., Hu, Y., Li, R., Zhang, F., Lei, H., Li, W., Hu, X., Liang, C., Xu, J., Li, X. and Yu, X. (2013). The carcinogenic liver fluke, *Clonorchis sinensis*: new assembly, reannotation and analysis of the genome and characterization of tissue transcriptomes. *PLoS ONE* **8**, e54732.

Jeltsch, A. (2010). Phylogeny of methylomes. *Science* **328**, 837–838.

Jeltsch, A., Nellen, W. and Lyko, F. (2006). Two substrate are better than one: dual specificities for DNMT2 methyltansferases. *Trends in Biochemical Sciences* **31**, 306–308.

Jurkowski, T. P. and Jeltsch, A. (2011). On the evolutionary origin of eukaryotic DNA methyltransferases and Dnmt2. *PLoS ONE* **6**, e28104.

Jurkowski, T. P., Meusburger, M., Phalke, S., Helm, M., Nellen, W., Reuter, G. and Jeltsch, A. (2008). Human DNMT2 methylates $tRNA^{Asp}$ molecules using a DNA methyltransferase-like catalytic mechanism. *RNA* **14**, 1663–1670.

Kano, H., Godoy, I., Courtney, C., Vetter, M. R., Gerton, G. L., Ostertag, E. M. and Kazazian, H. H., Jr. (2009). *L*1 retrotransposition occurs mainly in embryogenesis and creates somatic mosaicism. *Genes & Development* **23**, 1303–1312.

Keizer, J. and Utzinger, J. (2009). Food-borne trematodiases. *Clinical Microbiology Reviews* **22**, 466–483.

Kinney, S. R. M. and Pradhan, S. (2011). Regulation of expression and activity of DNA (cytosine-5) methyltransferases in mammalian cells. *Progress in Molecular Biology and Translational Science* **101**, 311–333.

Law, J. A. and Jacobsen, S. E. (2010). Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nature Reviews Genetics* **11**, 204–220.

Li, S., Du, J., Yang, H., Yin, J., Ding, J. and Zhong, J. (2013). Functional and structural characterization of DNMT2 from *Spodoptera frugiperda*. *Journal of Molecular Cell Biology* **5**, 64–66.

Müller, S., Windhof, I. M., Maximov, V., Jurkowski, T., Jeltsch, A., Förstner, K. U., Sharma, C. M., Gräf, R. and Nellen, W. (2013). Target recognition, RNA methylation activity and transcriptional regulation of the *Dictyostelium discoideum* Dnmt2-homoloue (DnmA). *Nucleic Acids Research* **41**, 8615–8627.

Raddatz, G., Guzzardo, P. M., Olova, N., Fantappié, M. R., Rampp, M., Schaefer, M., Reik, W., Hannon, G. J. and Lyko, F. (2013). Dnmt2-dependent methylomes lack defined DNA methylation patterns. *Proceedings of the National Academy of Science of the United States of America* **110**, 8627–8631.

Rountree, M. R., Bachman, K. E. and Baylin, S. B. (2000). DNMT1 binds HDAC2 and a new co-repressor, DMAP1, to form a complex at replication foci. *Nature Genetics* **25**, 269–277.

Schaefer, M. and Lyko, F. (2010). Solving the Dnmt2 enigma. *Chromosoma* **119**, 35–40.

Schultz, E. C., Roth, H. M., Ankri, S. and Finer, R. (2012). Structure analysis of *Entamoeba histolytica* DNMT2 (EhMeth). *PLoS ONE* **7**, e38728.

Shin, H. R., Oh, J. K., Masuyer, E., Curado, M. P., Bouvard, V., Fang, Y. Y., Wiangnon, S., Sripa, B. and Hong, S. T. (2010). Epidemiology of cholangiocarcinoma: an update focusing on risk factors. *Cancer Science* **101**, 579–585.

Song, J., Rechkoblit, O., Bestor, T. H. and Patel, D. J. (2011). Structure of DNMT1-DNA complex reveals a role for autoinhibition in maintenance DNA methylation. *Science* **331**, 1036–1040.

Takeshita, K., Suetake, I., Yamashita, E., Suga, M., Narita, H., Nakagawa, A. and Tajima, S. (2011). Structural insight into maintenance methylation by mouse DNA methyltransferase 1 (Dnmt1). *Proceedings of the National Academy of Science of the United States of America* **108**, 9055–9059.

The *Schistosoma japonicum* Genome Sequencing and Functional Analysis Consortium (2009). The *Schistosoma japonicum* genome reveals features of host-parasite interplay. *Nature* **460**, 345–352.

Thomson, C. A., Olson, M., Jackson, L. M. and Schrader, J. W. (2012). A simplified method for the efficient refolding and purification of recombinant human GM-CSF. *PLoS ONE* **7**, e49891.

Tsai, I. J., Zarowiecki, M., Holroyd, N., Garciarrubio, A., Sanchez-Flores, A., Brooks, K. L., Tracey, A., Bobes, R. J., Fragoso, G., Sciutto, E., Aslett, M., Beasley, H., Bennett, H. M., Cai, J., Camicia, F., Clark, R., Cucher, M., De Silva, N., Day, T. A., Deplazes, P., Estrada, K., Fernández, C., Holland, P. W. H., Hou, J., Hu, S., Huckvale, T., Hung, S. S., Kamenetzky, L., Keane, J. A., Kiss, F., The *Taenia solium* Genome Consortium *et al*. (2013). The genomes of four tapeworm species reveal adaptations to parasitism. *Nature* **496**, 57–63.

Wang, X., Chen, W., Huang, Y., Sun, J., Men, J., Liu, H., Luo, F., Guo, L., Lv, X., Deng, C., Zhou, C., Fan, Y., Li, X., Huang, L., Hu, Y., Liang, C., Hu, X., Xu, J. and Yu, X. (2011). The draft genome of the carcinogenic human liver fluke *Clonorchis sinensis*. *Genome Biology* **12**, R107.

Wojciechowski, M., Czapinska, H. and Bochtler, M. (2013). CpG underrepresentation and the bacterial CpG-specific DNA methyltransferase M.MpeI. *Proceedings of the National Academy of Science of the United States of America* **110**, 105–110.

Yoder, J. A., Walsh, C. P. and Bestor, T. H. (1997). Cytosine methylation and the ecology of intragenomic parasites. *Trends in Genetics* **13**, 335–40.

Zemach, A. and Zilberman, D. (2010). Evolution of eukaryotic DNA methylation and the pursuit of safer sex. *Current Biology* **20**, R780–R785.

Zhang, Y. (2008). I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics* **9**, 40.

https://doi.org/10.1017/S0031182015000566 Published online by Cambridge University Press