INTRODUCTION

DNA methylation is one of the major mechanisms of epigenetic regulation, as it constitutes a heritable cellular memory independently of changes in the genome's nucleotide sequence (Bird, 2002; Suzuki & Bird, 2008). DNA methylation is frequently involved in the regulation of chromosomal functions such as gene expression, DNA repair and recombination (Cedar & Bergman, 2009; Esteller, 2007; Jaenisch & Bird, 2003; Robertson & Wolffe, 2000; Schlissel, 2004). It is introduced by DNA methyltransferases (DNMTs) that catalyze the covalent modification of a methyl group to cytosine to form 5-methylcytosine (5mC) (Goll & Bestor, 2005). In mammals,
most of DNA methylation occurs in a CpG context in many differentiated somatic cells, although it can also occur at cytosines in a non-CpG context in some cell types, including embryonic stem (ES) cells, induced pluripotent stem (iPSC) cells, neurons and glial cells (Jang et al., 2017; Lister et al., 2009; Ziller et al., 2011). DNA methylation in a non-CpG context is much more frequently observed in plant genomes (Feng et al., 2010; Lindroth et al., 2001). CpG methylation in promoter regions often leads to the establishment of repressive histone modifications and the formation of heterochromatin regions, where gene expression is severely repressed (Cedar & Bergman, 2009). On the other hand, the hypermethylation of CpG sites within the gene body is associated with increased expression (Lister et al., 2009; Yang et al., 2014). Although non-CpG methylation has been reported for decades, their functions remain elusive (Jang et al., 2017; Patil et al., 2014).

DNA methylation status is determined by the balance between DNA methylation and demethylation reactions. DNA demethylation can occur either passively through DNA replication or actively through the action of a dioxygenase family known as “Ten-eleven translocation proteins” (TET proteins), including TET1, TET2 and TET3 in mammalian cells (Pastor et al., 2013; Wu & Zhang, 2017). They successively oxidize 5mC to 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC) and finally 5-carboxylcytosine (5caC), which are then removed by thymine DNA glycosylase (TDG). The resulting apurinic/apyrimidinic (AP) site is restored by an unmodified cytosine through base excision repair (BER), thereby completing active demethylation of 5mC (He et al., 2011; Ito et al., 2011; Weber et al., 2016). TET proteins contain a cysteine-rich domain and a double-stranded β-helix (DSBH) domain, which coordinates Fe(II) and interacts with 2-oxoglutarate (2-OG) as cofactors (Figure 1a; DSBH domain, which coordinates Fe(II) and interacts with 2-oxoglutarate (2-OG) as cofactors (Figure 1a). DSBH domains are represented as catalytic domain (CD) 1 and 2). The insertion sites of disruption markers to the genomic loci are indicated. Numbers are the position of amino acid residues. "aa" represents amino acids.

AID-dependent Ig diversification pathways are thought to be triggered by the occurrence of DNA lesions (single strand nicks or double strand breaks) followed by the removal of the AID-deaminated deoxycytidine (i.e., deoxyuracil) either by BER (removal of a uridine by Uracil-DNA glycosylase followed by AP endonuclease which introduces a DNA nick) or by mismatch repair (MMR: repair by mismatch recognition, followed by excision and replacement of the incorrect strand) (Di Noia & Neuberger, 2007). However, the molecular mechanism for the initial event of AID to induce Ig diversification still needs to be more extensively elucidated.

DT40 is a chicken-derived B-cell line where GCV is continuously occurring at the active Ig variable region (Baba et al., 1985; Buerstedde et al., 1990). We previously reported that the frequency of GCV (and SHM) in DT40 is markedly enhanced when DT40 is cultured in a medium containing a histone deacetylase inhibitor trichostatin A (TSA) (Seo et al., 2009; Yang et al., 2014). Although non-CpG methylation has been reported for decades, their functions remain elusive (Jang et al., 2017; Patil et al., 2014).

**Figure 1** Construction of TET1-KO, TET2-KO and TET3-KO DT40 cell lines and the effects on transcription of DNMT1, DNMT3A, IgVλ and AID genes. (a) Schematic diagrams of chicken TET1, TET2 and TET3 proteins with DSBH domains (denoted as CD1 and CD2: catalytic domain 1 and 2). The insertion sites of disruption markers to the genomic loci are indicated. Numbers are the position of amino acid residues. (b) Principle of the sIgM gain assay. The CL18 strain harbors a frameshift in its IgVλ locus, resulting in a sIgM(−) phenotype. The mutation is restored predominantly by GCV (very rarely by SHM) thereby generating sIgM(+) cells. Thus, the reversion rate from sIgM(−) to sIgM(+) is about 20% of all sIgM(−) cells quantified by RT-qPCR. The expression levels were normalized to β-actin and then to the WT levels. The thin vertical bars demonstrate the standard deviation (SD) of four biological replicates. (c) DNMTs mRNA levels in WT, TET1-, TET2- and TET3-KO cells quantified by RT-qPCR. The expression levels were normalized to β-actin and then to the WT levels. The thin vertical bars demonstrate the standard deviation (SD) of four biological replicates. The p-values were calculated with unpaired t test with Welch’s correction (ns; not statistically significant, *p < .05, **p < .01, ***p < .001, ****p < .0001). (d) IgVλ mRNA levels in WT and all single TET mutants quantified by RT-qPCR. The expression levels were normalized as described in (c). The error bars show the SD of at least six biological replicates (from left to right: n = 20, n = 6, n = 9 and n = 6). The p-values were as described above. (e) AID mRNA levels in WT and all single TET mutants quantified by RT-qPCR. The expression levels were shown as described in (c) and (d) (from left to right: n = 4, n = 12, n = 16 and n = 19).
et al., 2005, 2006). This higher frequency of GCV is concomitant with the increase in histone acetylation levels around the Ig variable region and the corresponding induction of IgLC transcription, indicating that TSA-mediated formation of accessible chromatin configuration can promote GCV. In addition, our subsequent studies showed histone deacetylase
HDAC1 and HDAC2 modulate GCV in DT40 cells (Kurosawa et al., 2010; Lin et al., 2008). These results suggest that local epigenetic configuration influences the GCV and SHM frequency. However, little is known about the contribution of local DNA methylation to the regulation of Ig diversification.

Here, we investigated the role of TET proteins in Ig diversification using DT40 and demonstrated that DT40 lacking TET3 exhibited a marked reduction in GCV activity in Ig variable regions, concomitant with a drop in the bulk levels of 5hmC and the local increase in 5mC at non-CpG cytosines in Ig pseudogene segments. Although TET3-KO only had a slight effect on the expression of proteins required for SHM and GCV such as AID, the double KO of TET1 and TET3 resulted in severe AID deficiency. These results suggest that TET3 plays a critical role in avian Ig diversification via the modulation of non-CpG methylation in Ig pseudogenes.

2 | RESULTS

2.1 | Effects of single knockouts for TET1, TET2 and TET3 genes on cell growth and gene expression in chicken B-cell line DT40

To examine the roles of TET proteins in Ig diversification, we conducted homologous gene targeting to establish single knockout (KO) mutants for TET1, TET2 and TET3 genes (TET1-KO, TET2-KO and TET3-KO, respectively) in a DT40-derived line “Clone 18 (CL18)”, which harbors a frameshift in the rearranged Ig light chain variable (IgVλ) segment. The CL18 strain thus exhibits a surface IgM negative (sIgM(−)) phenotype (Buerstedde et al., 1990), while it is converted to a surface IgM positive (sIgM(+)) when the frameshift mutation is restored by gene conversion (GCV) events using intact pseudogene segments as recombination templates (Figure 1b).

Deletions of the TET genes were targeted to their DSBH domain, which has been implicated in Fe(II)/2-OG dependent dioxygenase activity catalyzing an oxidation reaction into its substrate (Figure 1a; DSBH domains are shown as CD1 and 2). The deletions were confirmed by genomic PCR (Figure S1a). The disruption of each TET gene was also confirmed by loss of mRNA expression quantified with RT-qPCR experiments (Figure S1b).

We then examined the effects of TET gene single deletions on cell proliferation (Figure S1c). The cell growth speed and doubling time of each TET KO strain were almost identical to those of the TET-intact wild-type (WT) CL18 cells. Thus, we concluded that cell proliferation of DT40 is not significantly affected by any single KO deletion of TET genes.

Next, we assessed the influence of each single TET KO mutant on the expression of DNA methyltransferases (DNMTs, DNMT1 and DNMT3A) (Figure 1c), since methylation levels are established via the competing enzymatic activities of DNMTs and TET proteins (Ginno et al., 2020; Ravichandran et al., 2018; Verma et al., 2018). We observed a faint but statistically significant increase in the expression of DNMT1 in TET3-KO cells (1.17-fold). All in all, single TET KO mutants seem to have a very limited effect on the expression of DNMT1, and no effects on DNMT3A (Figure 1c).

We also examined the expression levels of active IgVλ (Figure 1d) and AID (an essential player in Ig diversification) (Figure 1e). The effects of single TET KO on the expression of IgVλ were similarly slight, with TET2-KO and TET3-KO exhibiting a 25% reduction in IgVλ expression (Figure 1d). On the other hand, AID expression levels were moderately reduced in TET1-KO (to about 40% of the WT level), but less so in TET2-KO and TET3-KO cells (by about 20%) (Figure 1e). We further examined other genes (CTNNBL1, BACH2, FANCD2, DDX11, PARP1, POLH and RNF8) (Abe et al., 2018) for Ig GCV in TET3-KO (Figure S2), since TET3-KO exhibited marked deficiency in Ig GCV (see below). We did not observe any significant variation of the expression levels of these factors for Ig GCV. In general, it can be concluded that single TET KO mutants had a limited impact on the expression of Ig gene diversification factors.

2.2 | Disruption of TET3 reduces diversification of IgVλ in DT40 cells

To determine whether TET proteins are involved in the Ig diversification, we studied TET1-KO, TET2-KO and TET3-KO strains using the surface IgM (slgM) gain assay. This assay measures the reversion from slgM(−) to slgM(+) phenotype through the repair of the CL18-type frameshift, which is almost exclusively attributable to GCV events around the frameshift mutation (Arakawa et al., 2002; Buerstedde et al., 1990) (Figure 1b). The slgM(−) CL18 subclones from WT and each TET-deficient mutant were first cloned by limiting dilution and the frequency of slgM(+) population of individual subclones was determined by flow cytometry. After about 1 month of clonal expansion, we observed that frequency of conversion to slgM(+) was significantly lower than that of WT in all TET-deficient mutants, with TET3 KO cells being the most severely affected (Figure 2a).

This tendency was also observed in the subclones cultured in medium supplemented with TSA, a histone deacetylase inhibitor that promotes chromatin accessibility and enhances the Ig GCV (Seo et al., 2005, 2006) (Figure 2b). In TSA-treated conditions, TET3-KO cells showed a much greater reduction of IgVλ GCV than either TET1-KO or TET2-KO cells, suggesting the importance of TET3 in Ig diversification independently of TSA-induced local chromatin configuration changes.
We then clarified the exact magnitude and spectrum of sequence alterations by sequencing the rearranged IgVλ of WT and TET3-KO cells after a 4-week culture (Figure 2c–e). Consistent with the results of the sIgM gain assay, only 26% of the TET3-KO population had single or double sequence alterations, while 91% of the WT demonstrated single, double, or triple sequence alterations (Figure 2c). Moreover, the mutation rate was reduced in TET3-KO cells to 24% of the WT level (Figure 2d).

Despite the difference in frequency, WT and TET3-KO cells exhibited a very similar spectrum of sequence alterations (Figure 2e). As observed in WT DT40 and natural chicken B cells (McCormack et al., 1991; Reynaud et al., 1987), Ig diversification in TET3-KO DT40 largely relied upon GCV. Previous studies have shown that genes involved in Ig diversification control the different steps of GCV and/or SHM (Abe, Branzei, et al., 2018). The disruption of genes for HR, such as XRCC2/3 (Sale et al., 2001), severely impairs GCV while increasing the frequency of SHM. On the other hand, the genes required for the induction of mutations, such as AID and BACH2 (Arakawa et al., 2002; Budzyńska et al., 2017), and the genes involved in the selection of repair pathway, including PARP1 (Paddock et al., 2010), are needed to sustain both GCV and SHM. Considering these reports and the unaltered spectrum in TET3-KO cells, TET3 is dispensable for the GCV process per se but may be important in the initiation of sequence alterations in IgVλ. The data were obtained and analyzed as described in (a). (c) Frequency of mutations occurred in IgVλ in WT and TET3-KO cells. The total number of the examined sequences is indicated in the center of each chart. (d) Mutation rate per kilobase per generation in WT and TET3-KO cells. The frequencies of mutation events were calculated from sequences analyzed in (c). The number of generations is based on the growth curve of each cell line (Figure S1c). The thin vertical bars show the SD of all sequences analyzed (****p < .0001, unpaired t test with Welch’s correction). (e) Spectrum of Ig sequence diversification in WT and TET3-KO cells. Mutations are classified as described in Experimental procedures. The number of total mutations is indicated in the center of each chart. GCV; gene conversion, Amb; ambiguous, PM; point mutation, Del; deletion, Ins; insertion
or template selection of GCV. This idea is consistent with the results that TET3 deletion did not influence the expression of genes involved in GCV (see Figure S2).

As another possible mechanism for the GCV deficiency in TET3-KO cells, we do not completely rule out the possibility that HR process itself could be impaired by the disruption of TET3 gene. However, this is unlikely, since multiple reports have indicated that HR-deficient mutants often exhibit cell growth retardation (Abe, Ooka, et al., 2018; Oestergaard et al., 2012; Yamamoto et al., 2005), whereas TET3-KO strain does not (Figure S1c). Taken together with the unaltered mutation spectrum in TET3-KO cells (Figure 2e), these results again suggest that TET3 targets earlier stages than the HR reaction process.

2.3 Absence of TET3 leads to altered usage of pseudogene templates for GCV

We further investigated the usage of pseudogene templates in IgVλ GCV. GCV is a type of unidirectional HR that uses upstream IgV pseudogenes as template (donor) sequences. IgVλ has 25 pseudogenes in its upstream region. We analyzed the IgVλ GCV tracts of WT and TET3-KO cells and estimated which pseudogenes were used as GCV templates (Figure 3a, b). In TET3-KO cells, along with the decrease of total GCV events, the diversity in pseudogene template usage was markedly reduced (Figure 3a, b), with only three pseudogenes used as templates (ψ8, ψ11 and ψ24), whereas 14 pseudogenes (ψ3, ψ4, ψ5, ψ7, ψ8, ψ10, ψ12, ψ13, ψ14, ψ15, ψ16, ψ17, ψ18, ψ19, ψ20, ψ21, ψ22, ψ23, ψ24).

![Figure 3](image-url)
We found that the methylation index (mC/C, expressed as a percentage) in TET3-KO cells was significantly higher (about threefold) in pseudogenes ψ8 and ψ7 compared with WT (Figure 4b,c). DNA methylation levels at the IgVλ segment in TET3-KO cells (Figure 4b–e). In this study, we focused on 5 loci including 3 pseudogenes (ψ7, ψ8, and ψ10, which were employed in GCV of WT DT40-CL18, but was less utilized in GCV of WT (Figure 3b). Thus, there seems to be no clear relationship among effects of TET3 deletion, GCV frequency in WT and sequence similarity of pseudogenes to the IgVλ segment, suggesting the possibility that the deletion of TET3 may have differentially affect individual pseudogenes, possibly by altering their DNA methylation status.

2.4 | Involvement of TET3 in the maintenance of low methylation at non-CpG cytosines in pseudogenes

We next assessed the effects of TET KOs on global levels of 5-hydroxymethylcytosine (5hmC) by dot blot analysis using anti-5hmC antibodies (Figure 4a). In contrast with TET1-KO and TET2-KO cells, which exhibited a slight increase or decrease in global 5hmC levels compared with the WT level, respectively, we observed a marked decrease in global 5hmC levels in TET3-KO cells. These results imply that TET3 plays a primary role to convert 5mC to 5hmC and regulate the DNA methylation levels of DT40 genomic DNA. The reason why TET1-KO and TET2-KO mutants are less affected is unknown, but it is consistent with the previous observations that each TET protein has distinct tissue-specific or cell type-specific functions (Ito et al., 2010; Rasmussen & Helin, 2016; Wu & Zhang, 2017).

To uncover the targets of TET3-mediated DNA demethylation, we performed bisulfite sequencing at several candidate loci in WT and TET3-KO cells (Figure 4b–e). In this study, we focused on 5 loci including 3 pseudogenes (ψ7, ψ8, and ψ10, which were employed in GCV of WT or TET3-KO cells as explained in the previous section), the IgVλ segment, and the core region of the DIVAC (diversification activator) element, downstream of IgVλ required for the enhancement of SHM and GCV (Blagodatski et al., 2009; Köhler et al., 2012; Kothapalli et al., 2008).

We found that the methylation index (mC/C, expressed as a percentage) in TET3-KO cells was significantly higher (about threefold) in pseudogenes ψ8 and ψ7 compared with WT (Figure 4b,c). DNA methylation levels at the IgVλ and DIVAC loci were slightly increased in TET3-KO cells, but remained at very low levels (2.5% and 2.0% on average, respectively).

Further in-depth analysis of the effects on DNA methylation by distinguishing between CpG sites and non-CpG sites (Figure 4d,e) revealed that DNA methylation in CpG sites was less significantly affected in TET3-KO cells. More importantly, the cytosines in a non-CpG context were persistently hypermethylated, particularly at all three pseudogenes in TET3-KO cells (Figure 4e). The methylation levels at non-CpG cytosines in pseudogenes showed stark differences from those at the IgVλ and DIVAC. It should be noted that most cytosines in the tested loci are in a non-CpG context (84%–90% in each target, Figure S3), meaning that the higher levels of unclassified DNA methylation in TET3-KO cells can be largely explained by the observed increase of DNA methylation in non-CpG cytosines. These results support the notion that TET3-mediated DNA demethylation at non-CpG cytosines in individual pseudogenes affects pseudogene usage during Ig GCV.

2.5 | TET1 and TET3 cooperatively upregulate AID expression and promote GCV

Although TET3-KO cells did not show a remarkable difference in the mRNA levels of GCV-associated genes (Figure 1d,e, and Figure S2), TET1-KO cells demonstrated a moderate but significant reduction in AID expression (Figure 1e). This result can explain the lowered slgM gain in TET1-KO cells (Figure 2a,b) and suggests that TET1 may play a role in promoting AID expression. Interestingly, previous studies have shown that double TET deletions (TET2/3 DKO mice) cause more detrimental effects on B cells than single TET deletions do (Domínguez et al., 2018; Lio et al., 2016; Orlanski et al., 2016).

We thus tested whether disruption of both TET1 and TET3 show more consequent effects on AID expression. We found that TET1 and TET3 double knockouts (TET1/3 DKO cells) showed a dramatic reduction of AID mRNA levels (about 17% of WT levels) (Figure 5a), which was much more severe than that in single TET1-KO or TET3-KO cells, respectively (Figure 1e). To explore the potential of TET1 and TET3-mediated DNA demethylation on AID expression, we performed bisulfite sequencing at the putative promoter region of AID (Figure S4). While TET1/3 DKO cells had slightly higher methylation levels than WT, the difference was not enough to account for the impaired expression of AID in TET1/3 DKO cells. Although the underlying mechanism remains to be clarified, it is likely that TET1 and TET3 redundantly and cooperatively regulate AID expression in DT40 cells. On the other hand, IgVλ mRNA levels were not significantly
FIGURE 4 TET3-dependent maintenance of genome-wide 5hmC levels and protection of pseudogenes from methylation at non-CpG sites in Ig pseudogenes. (a) Total 5hmC levels of WT, TET1-KO, TET2-KO and TET3-KO cells assessed by anti-5hmC dot blot. The left panel shows the dot blot results of the genomic DNA isolated from representative clones, and the quantification of them is shown in the right panel. The levels of 5hmC were normalized to the WT levels, and the thin vertical bars indicate the SD of triplicates (ns; not statistically significant (p > .05), *p < .05, **p < .01, ****p < .0001, unpaired t test with Welch’s correction). (b) Representative methylation patterns of WT and TET3-KO cells in the pseudogene ψ8, estimated by the bisulfite-sequencing analysis. Each row represents the sequence from one clone. Here, we show 20 representative clones (out of 70) which exhibited median levels of methylation. Unmethylated or methylated cytosines are represented by open or filled marks. Black and red marks correspond to cytosines in a non-CpG context or CpG context, respectively. (c) Quantification of the methylation levels of cytosines in pseudogenes (ψ7, ψ8 and ψ10), IgVλ2 and DIVAC element from WT and TET3-KO cells. The ratio of methylated cytosines to all cytosines (mC/C in %) is shown. The methylation levels were calculated from the results of bisulfite sequencing. The boxes extend from the 25th to 75th percentiles of each population with Tukey whiskers (ns; not statistically significant, *p < .05, ****p < .0001, Mann–Whitney test). (d) The ratio of methylated cytosines to cytosines in a CpG context (mC/CpG in %) in the target regions. Data were obtained and analyzed as described in (c). (e) The ratio of methylated cytosines to cytosines in a non-CpG context (mC/non-CpG in %) in the target regions. Data were obtained and analyzed as described in (c).
Importantly, the bisulfite sequencing conducted in this study revealed that TET3-KO had a marked increase in 5mC levels in some pseudogenes. In addition, the induced levels of 5mC were much higher (15%—30%) compared with those in IgVλ and DIVAC (less than 5%). It is thus likely that TET3 is targeted to pseudogene segments to repress local DNA methylation. A reduction in local DNA methylation is supposed to increase local DNA accessibility, so that pseudogenes with less 5mC are supposed to be preferentially utilized as Ig GCV templates.

Another finding of interest is that TET3-mediated demethylation is often found at cytosines in a non-CpG context, which was first described in the plant genome (Feng et al., 2010; Lindroth et al., 2001). In mammals, the majority of DNA methylation occurs at cytosines in a CpG context. Non-CpG comprises a very small portion of total methylated sites in differentiated somatic human cells, except for embryonic stem cells and neuronal cells (Jang et al., 2017; Ziller et al., 2011). It was reported that in chicken breast muscle tissue, genome-wide methylated CpG content among methylated cytosines is about 95% (Zhang et al., 2017). The frequency of methylated CHG or methylated CHH (where H is A, C or T) is about 1% and 4%, respectively (Zhang et al., 2017). Thus, mammalian and avian TET proteins are believed, for the most part, to catalyze demethylation of non-CpG cytosines and is involved in the establishment of non-CpG methylation (Guo, Su, et al., 2014; Ravichandran et al., 2018). Our results illustrate that the increased levels of 5mC in TET3-KO is largely attributable to methylation at non-CpG sites rather than that at CpG sites.

The molecular mechanism of TET3-targeting to non-CpG sites in the Ig pseudogenes is unknown. Previously, however, Guo et al. reported that strand-specific non-CpG methylation occurs in short interspersed nuclear elements (SINEs) and long interspersed nuclear elements (LINEs) in human-induced pluripotent stem (iPS) cells (W. Guo et al., 2014). It was also reported that TET proteins bind to multiple transposable elements and regulate their expression (de la Rica et al., 2016).

To note, while the interaction of TET proteins with cytosines is largely confined to CpG sites, DNMT3A can methylate non-CpG cytosines and is involved in the establishment of non-CpG methylation (Guo, Su, et al., 2014; Ramsahoye et al., 2000). In addition, recent studies have demonstrated that genomic methylation patterns are dynamically regulated by the competitive activities of TET proteins and DNMT3 (Charlton et al., 2020; Ginno et al., 2020). These reports imply the potential role of TET proteins in regulating non-CpG methylation in repetitive sequences, most likely achieved by counteracting de novo DNMTs. Future research directions will investigate the mechanisms underlying dynamics of non-CpG methylation in repetitive sequences such as transposons, retroelements and pseudogene clusters.

### 3.2 | Function of TET3 and TET1 in gene regulation or cis-acting elements for Ig diversification

The second possibility is that TET3 is involved in the activation of genes for Ig GCV and SHM. Indeed, TET proteins have previously been implicated in the regulation of early embryogenesis, cell differentiation and various neuronal functions (Pastor et al., 2013; Wu & Zhang, 2017). However, the results of this study clearly indicate that this hypothesis is unlikely, since many of the central players required for GCV and SHM were mostly unchanged in TET3-KO cells (Figure S2).

On the other hand, it should be noted that TET1-KO did lead to a moderate reduction in AID expression (Figure 1e), a gene that plays critical roles in the diversification of Ig variable segments in activated B cells. Furthermore, double KO of TET1 and TET3 (TET1/3 DKO) exhibited more severe reduction in AID expression (Figure 5a). The mouse Aicda locus encoding AID contains four transcriptional controlling regions which contain cis-acting sequences for NF-κB, C/EBP, Smad3/4, STAT6, Sp1, HoxC4 and Pax5 proteins (Zan & Casali, 2013). Lio et al. reported that the double knockout mouse of Tet2 and Tet3 results in reduction in AID expression due to increased DNA methylation in the superenhancer for AID (Lio et al., 2019). Although similar regulatory elements are not identified yet in the chicken AID locus, these results suggest that TET1 and TET3 redundantly and cooperatively function to upregulate AID in DT40 cells. However, TET3 per se is dispensable for AID activation; hence, the loss of GCV in TET3-KO is not simply due to the reduction in AID expression.

Previous studies demonstrated that cis-acting DIVAC elements are needed for AID targeting to Ig variable regions (Blagodatski et al., 2009; Kohler et al., 2012). The DIVAC elements reside in the region spanning from the transcription start site to 9.8kb downstream of DT40 IgVλ. Thus, it is plausible that TET3 is targeted to DIVAC to modulate local DNA methylation levels. Here, TET3-KO resulted in a slight increase of 5mC levels in DIVAC (Figure 4c). However, the induced levels of 5mC remained very low (at most 1%—2%), ruling out critical involvement of TET3 in the regulation of Ig diversification through the regulation of methylation in the DIVAC elements.

### 3.3 | Does TET3 induce DNA breaks in IgVλ segments?

The third possibility is that DNA lesions are induced in the Ig variable region as a by-product of the TET3-mediated demethylation process. TET family proteins catalyze the successive oxidation of 5mC to convert it to 5hmC, then
to 5-formylcytosine (5fC), and 5-carboxylcytosine (5caC), which are eventually removed and restored to cytosine through the action of base excision repair (BER) (Wu & Zhang, 2017). The BER pathway includes the removal of a damaged base by DNA glycosylases, leaving an abasic site which is further processed by an apurinic/apyrimidinic (AP) endonuclease. AP endonuclease finally creates a nick in the phosphodiester backbone of DNA (Krokan & Bjoras, 2013). Those DNA lesions are repaired by short- or long-patch DNA synthesis, and likely trigger other mutagenic events such as gene conversion-type recombination or error-prone repair synthesis.

It should be noted that a similar mechanism has been proposed to explain the contribution of AID cytidine deaminase to induction of SHM and GCV in the Ig variable region (Di Noia & Neuberger, 2007). Thus, it is speculated that TET3 may catalyze DNA lesions within the Ig variable regions in B cells to induce Ig GCV and SHM. If this is true, the absence of TET3 should lead to an increase in 5mC within the IgVλ segment. However, in TET3-KO cells, DNA methylation levels at the IgVλ only slightly increased, remaining at very low levels (at most 2%–3%) (Figure 4c). This result rules out the role of TET3 in forming recombination-activating DNA breaks in the IgVλ segment.

4 | EXPERIMENTAL PROCEDURES

4.1 | Cell culture

DT40 cells were cultured at 39.5°C and 5% CO₂ in IMDM supplemented with GlutaMax-1 (Thermo Fisher Scientific), 10% fetal calf serum (BioWest), 1% chicken serum (Thermo Fisher Scientific), 55 μM β-mercaptoethanol and 1% penicillin/streptomycin (Thermo Fisher Scientific) as described before (Hashimoto et al., 2016; Hashimoto et al., 2019; Seo et al., 2020). For the sIgM gain assay with TSA treatment, cells were maintained in medium containing 10 nM of trichostatin A (Fujifilm).

4.2 | Generation of TET-deficient DT40 cells

All knockout vectors for TET1, TET2 and TET3 were derived from p3LoxNeo provided by Dr. Hiroshi Arakawa and constructed as follows. The upstream and downstream regions of the DSBH domain of each TET gene were amplified by PCR from DT40 genomic DNA using KOD-plus DNA polymerase (Toyobo). The PCR primers for each TET gene were designed based on the NCBI database (TET1: NC_006093.3, TET2: NC_006091.3, TET3: NC_006109.3). The sequences of all PCR primers used (T1KO-1, 2, 3 and 4, and T2KO-1, 2, 3 and 4, and T3KO-1, 2, 3 and 4) are provided in Table S1. The upstream or downstream PCR products for the DSBH domain of each TET gene were then inserted into the upstream or downstream of neomycin marker cassette in p3LoxNeo, respectively, using DNA Ligation Kit Mighty Mix (Toyobo) or 5x In-Fusion HD Enzyme Premix (Takara Bio). The TET1 knockout vector with gpt marker cassette, the TET2 knockout vector with hisD or neomycin marker cassette, and the TET3 knockout vector with hisD or zeocin marker cassette were also constructed as described above, using the plasmids in which the neomycin cassette of p3LoxNeo is replaced by the relevant marker cassette. Transfection of the knockout vectors into CL18 DT40 cells was performed by electroporation using Gene Pulser (550 V, 25 μF) (Bio-Rad), as previously described (Hashimoto et al., 2016; Lin et al., 2008). The transfectants were transferred to selection medium containing 1,500 μg/mL G418, 5 μl/mL mycopheronolic acid and either 450 μg/mL HisD or 1,600 μg/mL zeocin, in accordance with their selection marker, dispensed into 96-well plates and cultured for about 1 week. Targeted integration of the vectors was confirmed by genomic PCR using the primers T1-1 and G-1, N-1 and T1-2, T2-1 and N-2, H-1 and T2-2, N-1 and T3-1, and H-1 and T3-1 (Table S1).

4.3 | sIgM gain assay

Subclones of each mutant were isolated by limiting dilution in 96-well plates. Colonies derived from a single cell were transferred into 24-well plates and further expanded for 5–7 days. For flow cytometric analysis, 1 x 10⁶ DT40 cells were washed with FCM buffer (PBS containing 0.5% bovine serum albumin (BSA) and 2 mM EDTA) and stained with FITC-conjugated anti-chicken IgM antibody (Bethyl laboratories, A30-102F) diluted 250-fold in FCM buffer at 4°C for 15 min. After washing twice with FCM buffer, cells were resuspended in FCM buffer and analyzed by NovoCyte flow cytometer (Agilent Technologies).

4.4 | Analysis of IgVλ diversification

After 28 days of culturing, genomic DNA was extracted with NucleoSpin Tissue (MACHEREY-NAGEL) and the IgVλ region was amplified by PCR using KOD-plus DNA polymerase (Toyobo) with the primers VL-1 and 2. The purified PCR products were cloned into pBluescript II SK(−) using 5x In-Fusion HD Enzyme Premix (Takara Bio), and then Sanger sequenced with VL-seq. All primers used are listed in Table S1. Sequence alignment was performed and mutations were identified by Geneious 7.1.9 (Biomatters). Each mutation was classified as previously described (Romanello et al., 2016; Sale et al., 2001). Briefly, ≥8 bp sequences
containing more than one mutation and matching any pseudogenes was categorized as gene conversion (GCV). Any \( \geq 8 \) bp sequence that matched a pseudogene but harboring a single mutation that cannot be identified as either gene conversion or point mutation was classified as ambiguous (Amb). Single base substitutions around which the 8 bp sequence did not match any pseudogenes were categorized as point mutations (PM).

4.5 | RT-qPCR analysis

Total RNA was extracted with NucleoSpin RNA (MACHEREY-NAGEL) according to the manufacturer’s instructions, and reverse transcription was performed with the PrimeScript RT reagent Kit (Perfect Real Time) (Takara Bio). The obtained cDNA was subjected to qPCR with the KAPA SYBR FAST qPCR Master Mix (2X) ABI Prism (Kapa Biosystems) using the StepOnePlus System (Thermo Fisher Scientific). Primers used are listed in Table S1.

4.6 | Dot blot assay

Genomic DNA was extracted with NucleoSpin Tissue (MACHEREY-NAGEL), and the concentration was measured with the Qubit dsDNA HS Assay Kit (Thermo Fisher). DNA was denatured by boiling at 95°C for 5 min followed by snap-freezing on ice. Denatured genomic DNA (25 ng) was spotted onto a nitrocellulose membrane (Roche) and ultraviolet (UV) cross-linked by DNA-FIX (ATTO). The membrane was subsequently incubated in blocking buffer (5% skim milk in TBST (Tris buffer saline containing 0.1% Tween 20)) for 2 hr at room temperature, washed three times with TBST and then incubated overnight at 4°C with a primary antibody against 5hmC diluted 10,000-fold in blocking buffer (Active Motif, Cat # 39769). After four washes with TBST, the membrane was incubated with horseradish peroxidase (HRP)-conjugated anti-rabbit IgG diluted 10,000-fold (GE healthcare) at room temperature for 2 hr. After washing four times with TBST, the blots were developed with Amersham ECL Western Blotting Analysis System (GE Healthcare) and the enhanced chemiluminescence (ECL)-based signals were detected by ImageQuant LAS4000 mini (GE Healthcare). The signals were quantified using ImageJ software (Schneider et al., 2012).

4.7 | Bisulfite sequencing

Genomic DNA was extracted with NucleoSpin Tissue, and then, bisulfite conversion was performed using the MethylEasy Xceed Rapid DNA Bisulphite Modification Kit (Human Genetic Signatures). The target region was amplified by PCR using EpiMark Hot Start Taq DNA Polymerase (New England Biolabs) with primers p7-1 and 2, p8-1 and 2, p10-1 and 2, VLbs-1 and 2, or DIVbs-1 and 2. The PCR products were then subjected to nested PCR using KOD-plus DNA polymerase (Toyobo) with the primers p7N-1 and 2, p8N-1 and 2, p10N-1 and 2, VLbsN-1 and 2, or DIVbsN-1 and 2, respectively. The purified nested PCR products were cloned into pBluescript II SK(−) and sequenced with the M13R primer as described above (Analysis of IgV\( \lambda \) diversification). Sequence alignment and identification of 5mC were performed with Geneious 7.1.9.

4.8 | Statistical analysis

Statistical analyses were performed with GraphPad Prism 8 for macOS (GraphPad Software) and are described in the relevant figure legends.

ACKNOWLEDGMENTS

We would like to thank Dr. K. Kurosawa and S. Kondo (The University of Tokyo) for construction of the TET1-KO strain and TET2-KO strain, experimental setup of bisulfite sequencing and their technical support; Dr. K. Hashimoto (The University of Tokyo) for technical support and helpful discussion on construction of the TET3-KO strain, slgM gain assay, IgV\( \lambda \) sequence analysis; A. Murayama (The University of Tokyo) for technical support on plasmid preparation and Dr. H. Arakawa (Helmholtz Zentrum München—German Research Center for Environmental Health) for the generous gift of p3loxNeo plasmid. This work was supported by a Grant-in-Aid for Basic Science (20K06598). K.O. is also supported by JST CREST Grant Number JPMJCR18S3, Japan.

ORCID

Natsuki Takamura https://orcid.org/0000-0002-2671-4971

REFERENCES

Abe, T., Branzei, D., & Hirota, K. (2018). DNA damage tolerance mechanisms revealed from the analysis of immunoglobulin v gene diversification in avian DT40 cells. Genes, 9(12), 614. https://doi.org/10.3390/genes9120614

Abe, T., Ooka, M., Kawasumi, R., Miyata, K., Takata, M., Hirota, K., & Branzei, D. (2018). Warsaw breakage syndrome DDX11 helicase acts jointly with RAD17 in the repair of bulky lesions and replication through abasic sites. Proceedings of the National Academy of Sciences of the United States of America, 115(33), 8412-8417. https://doi.org/10.1073/pnas.1803110115

Arakawa, H., Hauschild, J., & Buerstedde, J. M. (2002). Requirement of the activation-induced deaminase (AID) gene for immunoglobulin gene conversion. Science, 295(5558), 1301–1306. https://doi.org/10.1126/science.1067308
Baba, T. W., Giroir, B. P., & Humphries, E. H. (1985). Cell lines derived from avian lymphomas exhibit two distinct phenotypes. *Virology, 144*(1), 139–151. https://doi.org/10.1016/0042-6822(85)90312-5

Bird, A. (2002). DNA methylation patterns and epigenetic memory. *Genes and Development, 16*(1), 6–21. https://doi.org/10.1101/gad.947102

Błagodatski, A., Batrak, V., Schmidl, S., Schoetz, U., Caldwell, R. B., Arakawa, H., & Buerstedde, J. M. (2009). A cis-acting diversification activator both necessary and sufficient for AID-mediated hypermethylation. *PLoS Genetics, 5*(1), 1–11. https://doi.org/10.1371/journal.pgen.1000332

Budzyńska, P. M., Kyläniemi, M. K., Kallonen, T., Soikkeli, A. I., Nera, K. P., Lassila, O., & Alinekula, J. (2017). Bach2 regulates AID-mediated immunoglobulin gene conversion and somatic hypermutation in DT40 B cells. *European Journal of Immunology, 47*(6), 993–1001. https://doi.org/10.1002/eji.201646895

Buerstedde, J. M., Reynaud, C. A., Humphries, E. H., Olson, W., Ewert, D. L., & Weill, J. C. (1990). Light chain gene conversion continues at high rate in an ALV-induced cell line. *EMBO Journal, 9*(3), 921–927. https://doi.org/10.1002/j.1460-2075.1990.tb08190.x

Cedar, H., & Bergman, Y. (2009). Linking DNA methylation and histone modification: Patterns and paradigms. *Nature Reviews Genetics, 10*(5), 295–304. https://doi.org/10.1038/nrg2540

Charlton, J., Jung, E. J., Mattei, A. L., Bailly, N., Liao, J., Martin, E. J., Giesselmann, P., Brändl, B., Stamenova, E. K., Müller, F.-J., Kiskinis, E., G rifle, A., Smith, Z. D., & Meissner, A. (2020). TETs compete with DNMT3 activity in pluripotent cells at thousands of methylated somatic enhancers. *Nature Genetics, 52*(8), 819–827. https://doi.org/10.1038/s41588-020-0639-9

Cummings, W. J., Bednarski, D. W., & Maizels, N. (2007). Chromatin structure regulates gene conversion activator both necessary and sufficient for AID-mediated hypermethylation. *PLoS Genetics, 5*(1), 1–11. https://doi.org/10.1371/journal.pgen.1000332

Guo, J. U., Su, Y., Shin, J. H., Shin, J., Li, X., Xie, B., Zhong, C., Hu, S., Le, T., Fan, G., Zhu, H., Chang, Q., Gao, Y., Ming, G.-L., & Song, H. (2014). Distribution, recognition and regulation of non-CpG methylation in the adult mammalian brain. *Nature Neuroscience, 17*(2), 215–222. https://doi.org/10.1038/nn.3607

Hashimoto, K., Kurosawa, K., Murayama, A., Seo, H., & Ohta, K. (2016). B cell-based seamless engineering of antibody Fc domains. *PLoS One, 11*(12), 1–22. https://doi.org/10.1371/journal.pone.0167232

Jaenisch, R., & Bird, A. (2003). Epigenetic regulation of gene expression: How the genome integrates intrinsic and environmental signals. *Nature Genetics, 33*(3S), 245–254. https://doi.org/10.1038/ng1089

Jang, H. S., Shin, W. J., Lee, J. E., & Do, J. T. (2017). CpG and non-CpG methylation in epigenetic gene regulation and brain function. *Genes, 8*(6), 2–20. https://doi.org/10.3390/genes8060148

Kita, H., Kimura, M., Yamamoto, K., Seo, H., Namikoshi, K., Agata, Y., Ohta, K., & Takata, M. (2008). Regulation of histone H4 acetylation by transcription factor E2A in Ig gene conversion. *International Immunology, 20*(2), 277–284. https://doi.org/10.1093/intimm/dxm140

Sadler, K. C., Pradhan, S., Pellegrini, M., & Jacobsen, S. E. (2010). Conservation and divergence of methylation patterning in plants and animals. *Proceedings of the National Academy of Sciences of the United States of America, 107*(19), 8689–8694. https://doi.org/10.1073/pnas.1002720107
