博士論文

Analysis of resistant factors for DNA methylation acquisition by Epstein-Barr virus infection

（EBウイルス感染が誘導するエピゲノム異常におけるDNAメチル化抵抗因子の解析）

南波　宏枝

Contents

Contents 2

1. Abstract 4

2. Abbreviations 5

3. Introduction 6

3-1. Epigenetics 6

3-1-1. Histone modifications 6

3-1-2. DNA modifications 6

3-2. Epigenomic alterations in cancer 7

3-3. DNA methylator epigenotype 8

3-4. EBV-associated gastric cancer 10

3-4-1. Epstein-Barr virus (EBV) 11

3-5. Technique for DNA methylation analysis 11

3-5-1. Bisulfite conversion 11

3-5-2. Infinium BeadChip array 12

3-5-3. Pyrosequencing 13

3-5-4. Methylated DNA immunoprecipitation sequencing (MeDIP-seq) 13

3-6. Akata system 14

3-7. Purpose of this study 14

4. Results 16

4-1. Transcriptome analysis 16

4-1-1. Transcriptome analysis by RNA-sequencing 16

4-1-2. Validation by real-time RT-PCR and immunoblotting analyses 16

4-2. Downregulation mechanism of *TET2* 18

4-3. Hydroxymethylation target genes of TET2 21

4-4. Knockdown of *TET2* accelerated *de novo* methylation during EBV infection 23

4-5. No methylation was induced by *TET2* knockdown alone 29

5. Discussion 31

5-1. Downregulation mechanism of *TET2* 31

5-2. TET2 overexpression analysis 32

5-3. Analysis of *TET2* knockdown and EBV infection 33

5-4. *TET2* knockdown itself 33

5-5. DNA methylation in other malignancies 34

6. Conclusions 36

7. Materials and methods 37

7-1. Cell culture and treatment 37

7-2. Plasmid construction 37

7-3. Knockdown by shRNA 38

7-4. Real-time RT-PCR 38

7-5. Hydroxymethylated and methylated DNA immunoprecipitation sequencing (hMeDIP-seq / MeDIP-seq) 39

7-6. RNA-sequencing (RNA-seq) analysis 42

7-7. Infinium assays 42

7-8. Pyrosequencing analysis 44

7-9. Analysis of miRNA 44

7-10. Statistical analyses 45

8. Acknowledgements 52

9. References 54

1. Abstract

Extensive DNA methylation is observed in gastric cancer with Epstein-Barr virus (EBV) infection, and EBV infection is the cause to induce this extensive hypermethylation phenotype in gastric epithelial cells. However, some 5′ regions of genes do not undergo *de novo* methylation, despite the induction of methylation in surrounding regions, suggesting the existence of a resistance factor against DNA methylation acquisition. We conducted an RNA-seq analysis of gastric epithelial cells with and without EBV infection and found that TET family genes, especially *TET2*, were repressed by EBV infection at both mRNA and protein levels. *TET2* was found to be downregulated by EBV transcripts, e.g. BARF0 and *LMP2A*, and also by seven human miRNAs targeting *TET2*, e.g., miR-93 and miR-29a, which were upregulated by EBV infection, and transfection of which into gastric cells repressed *TET2*. Hydroxymethylation target genes by TET2 were detected by hydroxymethylated DNA immunoprecipitation sequencing (hMeDIP-seq) with and without TET2 overexpression, and overlapped significantly with methylation target genes in EBV-infected cells. When *TET2* was knocked down by shRNA, EBV infection induced *de novo* methylation more severely, including even higher methylation in methylation-acquired promoters or *de novo* methylation acquisition in methylation-protected promoters, leading to gene repression. *TET2* knockdown alone without EBV infection did not induce *de novo* DNA methylation. These data suggested that TET2 functions as a resistance factor against DNA methylation in gastric epithelial cells and repression of *TET2* contributes to DNA methylation acquisition during EBV infection.

2. Abbreviations

hMeDIP: hydroxymethylated DNA immunoprecipitation

MeDIP: methylated DNA immunoprecipitation

DNMT: DNA methyltransferase

TSS: transcription start sites

EBV: Epstein-Barr virus

RT-PCR: reverse transcription polymerase chain reaction

TET: Ten-eleven-translocation

FBS: fetal bovine serum

cDNA: complementary DNA

CMV: cytomegalovirus

miRNA: microRNA

WT: wild type

OE: overexpression

IP buffer: immunoprecipitation buffer

Tris: tris (hydroxymethyl) aminomethane

EDTA: ethylenediaminetetraacetic acid

TE: tris-EDTA

mRNA: messenger RNA

SDS: sodium dodecyl sulphate

3. Introduction

3-1. Epigenetics

Epigenetics is heritable information during cell division other than DNA sequences, including various histone modifications and DNA modifications. Epigenetic status regulates gene expressions leading to phenotypic changes [1].

3-1-1. Histone modifications

A histone is a octamer consisting of 2 molecules of H2A, H2B, H3 and H4 respectively. 145-147 bp of DNA wraps around histone proteins to form a nucleosome core, and the linker histone H1 binds the nucleosome at the entry and the exit sites of the DNA [2-10]. The core histones are mainly globular with the exception of their N-terminal tails called histone tails [11]. Amino acid residue of histone tails acquire various modifications such as lysine methylation, lysine acetylation, arginine methylation, serine phosphorylation, threonine phosphorylation, lysine ubiquitylation, lysine sumoylation and so on, influencing transcription, repair and replications [11].

3-1-2. DNA modifications

One of the important DNA modifications is methylation of cytosine at CpG nucleotide, which is mainly regulated by three DNA methyltransferases, DNMT1, DNMT3a and DNMT3b. DNMT3a and DNMT3b work as *de novo* methyltransferases and DNMT1 works as a maintenance methyltransferase during mitosis but is also reported to work as a *de novo* methyltransferase in cancer cells [12, 13]. DNA methylation often occurs on CpG sites at CpG islands. CpG island is a genomic DNA region often located around the promoter regions, in which the frequency of the CG sequence is high compared to other regions [14]. DNA

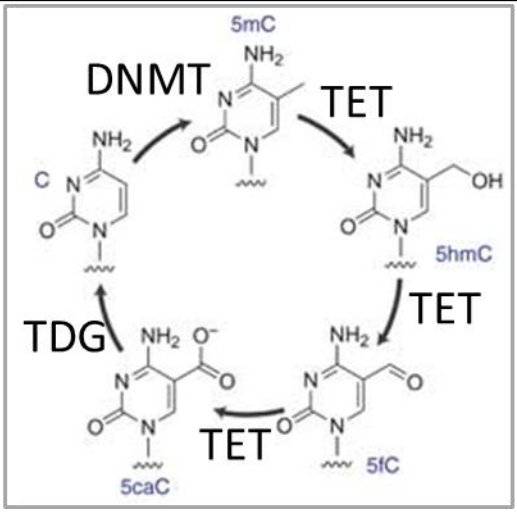

Fig. 1: 5mC conversion to 5hmC, 5fC and 5caC by TET proteins. Modified figure from reference [19].

methylation in CpG islands of the gene promoter regions blocks transcription factors from binding to DNA, and also gathers proteins which hold Methyl-CpG-binding domain, leading to chromatin structure changes and gene repressions [15]. 5-methylcytosine (5mC) at methylated CpG dinucleotide can be oxidized to 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC) by TET family proteins, which are known as demethylating enzymes, and 5fC and 5caC are directly turned unmodified cytosine through base excision repair by thymine DNA glycosylase (TDG) (Fig. 1) [16-20]. TET family proteins consist of TET1, TET2 and TET3, which are reported to require α-ketoglutarate and Fe^{2+} as coenzymes in oxidation process.

3-2. Epigenomic alterations in cancer

Aberrant DNA methylation is one of the major epigenomic alterations in cancer, and DNA hypermethylation of gene promoter regions inactivates tumor suppressor genes and strongly affects cancer development [21-25]. Cancer with high

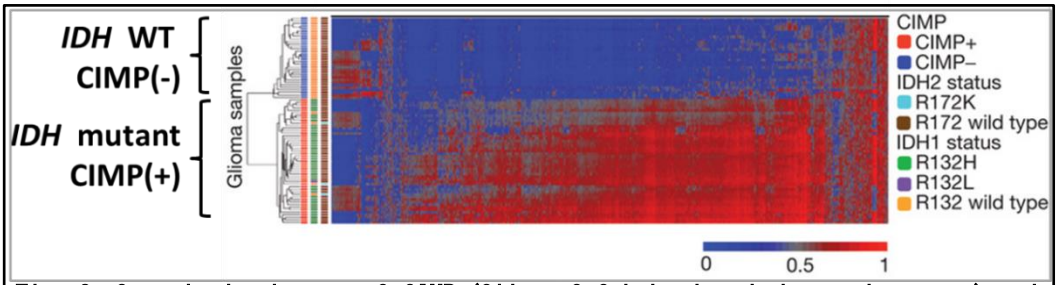

Fig. 2: Correlation between G-CIMP (Glioma CpG island methylator phenotype) and *IDH1* mutations. Modified figure from [28].

methylation level in gene promoter regions, known as CIMP (CpG island methylator phenotype), has been found in colorectal cancer [26], gastric cancer [27], glioma [28-32], adult T cell leukemia/lymphoma (ATLL) [33], and many other cancers [34]. One of TET family genes, *TET1*, was first found as a fused protein to MLL in acute myeloid leukemia [35, 36] and 15% of various myeloid cancer patients have *TET2* mutations which cause DNA hypermethylation and induce leukemogenesis [33, 37-40]. The subgroup of glioma with extensive promoter hypermethylation is known as G-CIMP (Glioma CpG island methylator phenotype) and more than 70% of lower-grade gliomas (grades II and III) possess mutations of *IDH1* and *IDH2* which produce α-ketoglutarate, a cofactor of TET family (Fig. 2) [28-30, 32].

3-3. DNA methylator epigenotype

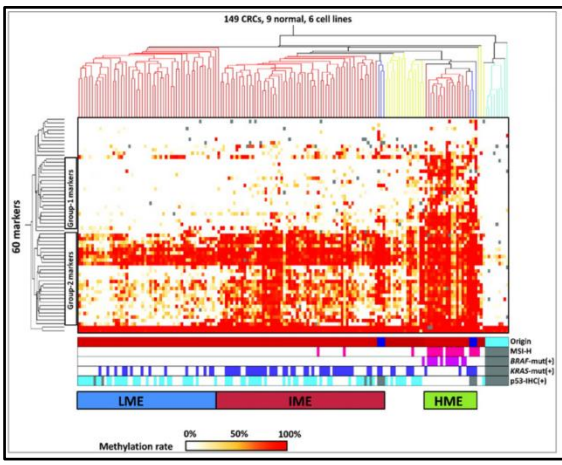

Fig. 3: Three major epigenotypes of colorectal cancer. Colorectal cancer is classified into high-methylation epigenotype (HME), intermediate-methylation epigenotype (IME) and low-methylation epigenotype (LME) [41].

Colorectal cancer can be stratified into 3 major subgourps with unique DNA methylation epigenotypes: high-methylation epigenotype (HME), intermediate-methylation epigenotype (IME) and low-methylation epigenotype (LME) [41]. HME correlated strongly with MSI-high and *BRAF*-mutation(+), whereas IME correlated with *KRAS*-mutation(+) (Fig. 3).

For gastric cancer, it also has 3 methylation epigenotypes, EBV-/low methylation, EBV-/high methylation and EBV+/high methylation subtypes. The 3

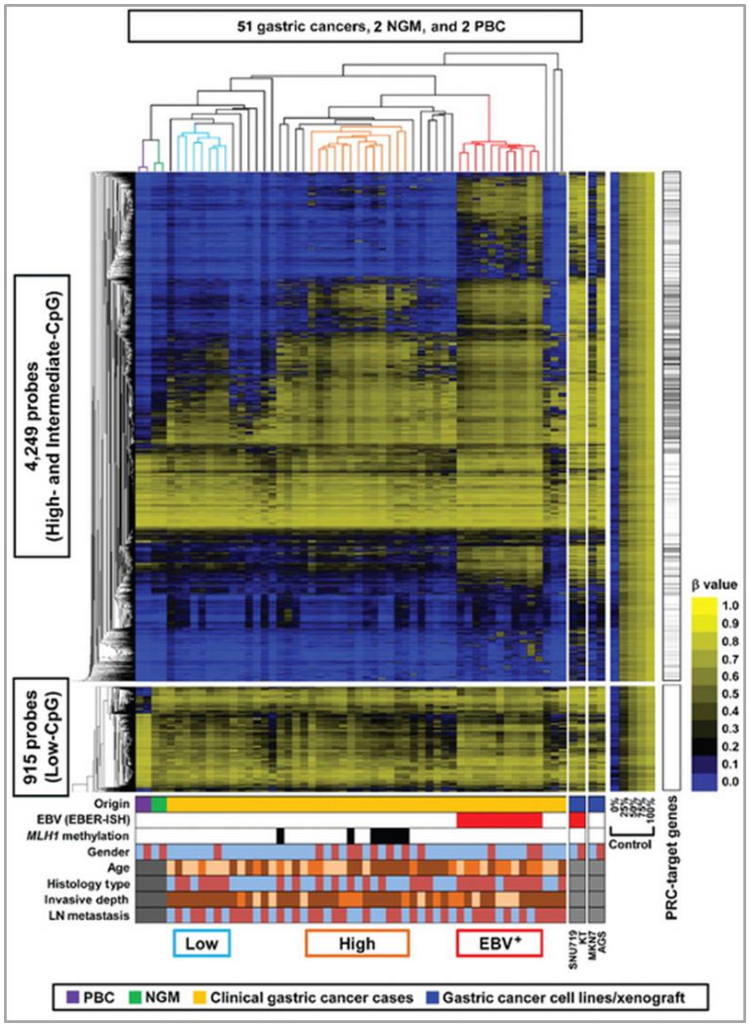

Fig. 4: Three epigenotypes of gastric cancer, EBV-/low, EBV-/high and EBV+/high. PRC-target genes in embryonic stem cells are significantly enriched in EBV-/high and EBV-/low methylated genes, suggesting a different methylation mechanism for EBV+ tumors [42].

epigenotypes are characterized by 3 kinds of genes: genes methylated only EBV+ tumors, genes methylated EBV+ and EBV-/high tumors and genes methylated all 3 kinds of tumors [42]. Polycomb repressive complex (PRC) target genes in embryonic stem cells are significantly enriched in EBV-/high and EBV-/low methylated genes, suggesting different methylation mechanisms for EBV+ tumors (Fig. 4).

3-4. EBV-associated gastric cancer

One of the most hypermethylated malignancies is EBV-associated gastric cancer [42-48]. Among gastric cancer, 7-15% are EBV-associated gastric cancer throughout the world without geographical divergence, and EBV is basically found in every malignant cell [45]. Clinical features of EBV-associated gastric cancer include male predominance, appearance in the upper part of the stomach and abundance of inflammatory cell infiltration [49, 50]. Mutation of *ARID1A*, a member of SWI/SNF family of a nucleosome remodeling complex is also reported [51], and *MLH1* is not methylated in the promoter region so microsatellite instability (MSI) is not detected in EBV associated gastric cancer [47, 52]. Host miRNA expression is changed by EBV infection. EBV infection downregulates the expression of host miR-200 family whose targets are *ZEB1* and *ZEB2*, and the upregulation of ZEB1 and ZEB2 causes *CDH1* repression [53]. Latent infection of EBV upregulates DNMTs of host cells and extensive methylation is observed first in the EBV genome and subsequently in host genome [42, 54]. Although it was reported that EBV expressing protein LMP2A induces DNMT1 by phosphorylating STAT3 [55], molecular mechanisms of methylation induction during EBV infection is largely unknown.

3-4-1. Epstein-Barr virus (EBV)

Epstein-Barr virus (EBV) is a double stranded DNA virus belonging to Herpesviridae family and 90% of adult people are latently infected all over the world [56]. EBV infection causes infectious mononucleosis in initial infection and is also involved in malignant tumors such as Hodgkin's lymphoma, Burkitt lymphoma, T-cell and NK-cell lymphomas, opportunistic lymphoma in immunocompromised host, nasopharyngeal carcinoma, and Gastric carcinoma [57-61]. When EBV is infected to host cells, genome of EBV is hardly integrated into host genome but bound to host genome by two dimers of Epstein–Barr nuclear antigen 1 (EBNA1) protein forming episome [62, 63]. Although EBV genome encodes approximately 100 genes, only limited genes are expressed depends on types of latency, from type I to type III. The latency types depend on types of host cells and for gastric cancer, EBV shows latency type I as with Burkitt lymphoma. Latency type I is the most limited expression and only a few genes such as *EBNA1* and *LMP2A* are expressed. Other repressed genes are regulated by DNA methylation [64].

3-5. Technique for DNA methylation analysis

In this study, we conducted 3 types of DNA methylation analyses, Infinium BeadChip array, pyrosequencing and methylated DNA immunoprecipitation sequencing (MeDIP-seq). Bisulfite conversion is conducted before Infinium BeadChip array and pyrosequencing.

3-5-1. Bisulfite conversion

Bisulfite conversion is the most common method for DNA methylation analysis. Unmethylated cytosine is sulfonated and sulfonated cytosine is

deaminated by hydrolysis, and it turns to sulfonated uracil. After removal of sulfone group by alkali treatment, it turns into uracil which is read as thymine by PCR. On the other hand, 5mC doesn't change in response to bisulfite treatment and is read as cytosine itself. Difference of thymine and cytosine enables us to distinguish methylation patterns [65]. Since neither 5hmC nor 5mC changes in response to bisulfite treatment, 5hmC and 5mC cannot be distinguished in a bisulfite-based methylation assay [66, 67].

3-5-2. Infinium BeadChip array

The Infinium HumanMethylation450 BeadChip (Illumina, WG-314-1003) contains more than 485,000 individual CpG probes covering 99% of RefSeq genes with an average of 17 CpG sites per gene. In each CpG site, ratio of fluorescent signal called β value was measured by a methylated probe relative to the sum of

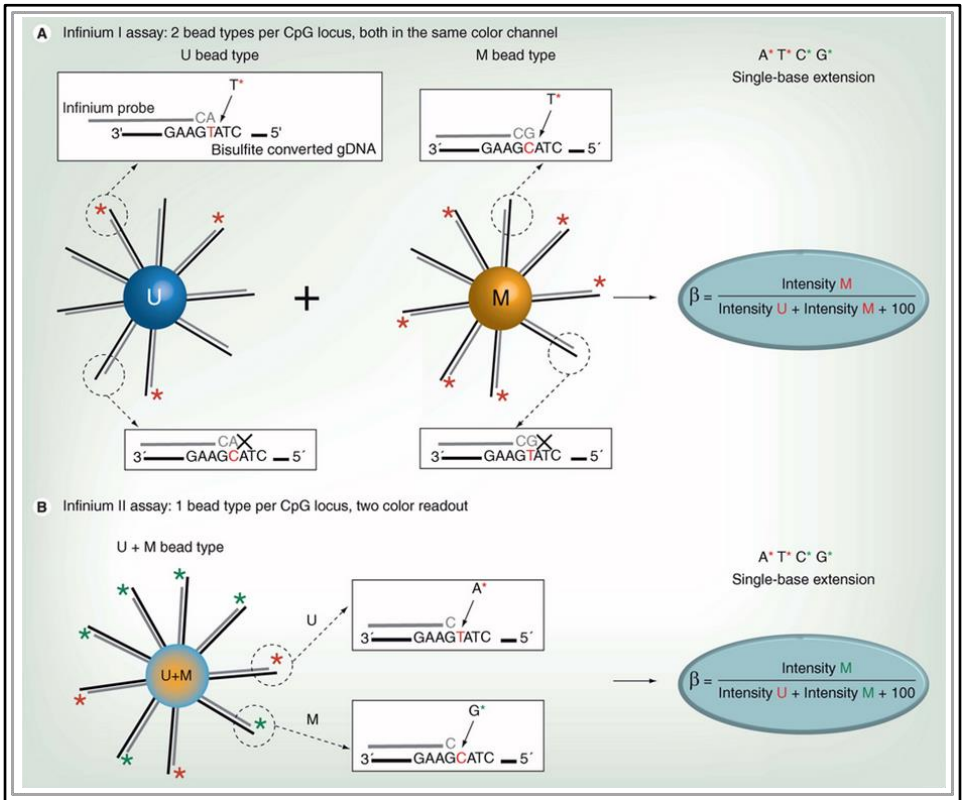

Fig. 5: The principal of Illumina Infinium HumanMethylation450 BeadChip. There are two types of assays, Infinium I assay and Infinium II assay [68].

both methylated and unmethylated probes [68]. The β value, ranging from 0.00 to 1.00, reflects the methylation level of each CpG site from low to high. While Infinium 27K uses one type of assay, Infinium 450K combines two kinds of assays, Infinium I assay and Infinium II assay (Fig. 5). In Infinium I assay, there are two different probes located on two different bead types. One is for methylated strand, and the other is for unmethylated strand. For Infinium II assay, there is one type of probe which can bind to both methylated and unmethylated DNA. Guanine with green signal is incorporated if DNA is methylated, and adenine with red signal is incorporated if DNA is unmethylated. In both assays, the percentage of DNA methylation is reported as β value.

3-5-3. Pyrosequencing

Pyrosequencing enables us to know which nucleotide is incorporated to template strand [69]. When one of the nucleotides is incorporated to template strand, released pyrophosphate is converted to ATP by ATP sulfurylase. Luciferase generates light signals proportional to ATP molecules released. There are 4 assay types in Pyrosequencing; Genotyping (SNP), Allele Quantification (AQ), Sequence Analysis (SQA) and Methylation Analysis (CpG). In the Methylation Analysis, the ratio of methylation level of specific CpG sites in a region is quantitatively detected. When analyzing bisulfite-treated DNA, 5mC is detected as cytosine and unmodified cytosine is detected as thymine. Methylation levels are detected as (cytosine / cytosine + thymine) × 100 (%).

3-5-4. Methylated DNA immunoprecipitation sequencing (MeDIP-seq)

In MeDIP-seq, fragmented DNA was immunoprecipitated by anti-5mC antibody. The enriched 5mC regions are read by next generation sequencing. This

method enables differentiation between 5mC and 5hmC, that bisulfite based methods cannot do, by using different antibodies for 5mC and 5hmC.

3-6. Akata system

Normally the frequency of EBV infection to epithelial cells is very low, Akata system, developed by Takada et al., enables to construct EBV infected epithelial cells efficiently [70]. B cell lymphoblast cell line, Akata, derived from EBV+ Burkitt lymphoma is infected by EBV latently. By stimulating immunoglobulin G, which is expressed on the surface of Akata cells, by anti-IgG antibody, latent infection turns into lytic infection and Akata cells release an abundance of EBV. By co-culturing lytic Akata cells with gastric epithelial cells, EBV infected gastric epithelial cells are constructed. One characteristic of Akata cells is that after EBV is dropping out by rounds of passages, reinfection can be conducted. Akata cells used in this study were reinfected with EBV holding neomycin resistant gene so that the selection of EBV infected cells by G418 is possible. By infecting low-methylation gastric cancer cells with EBV applying Akata system, previous studies have demonstrated that EBV infection itself causes the induction of extensive hypermethylation [42, 70, 71].

3-7. Purpose of this study

Normally in cells, induction and resistance factors for DNA methylation might work in balance to maintain methylation status, and aberrant DNA methylation may occur when the balance is collapsed for example with EBV infection. To investigate the molecular mechanisms by which extensive hypermethylation is induced in EBV-positive gastric cancer, a transcriptome analysis of epithelial

gastric cancer cells with and without EBV infection was conducted, and *TET2* was found to be one of the downregulated genes by EBV infection. In this study, we investigated if *TET2* works as a resistance factor against DNA methylation acquisition during EBV infection.

4. Results

4-1. Transcriptome analysis

4-1-1. Transcriptome analysis by RNA-sequencing

To identify candidate resistance factors for methylation acquisition during EBV infection, we performed an RNA-seq analysis using a low-methylation gastric cancer cell line, MKN7 (MKN7_WT), and three previously established EBV-infected MKN7 clones (MKN7_EB#1, EB#2, and EB#3) [42]. Downregulated genes in response to EBV infection included *TET1* and *TET2*, which encode TET family demethylation enzymes (Fig. 6).

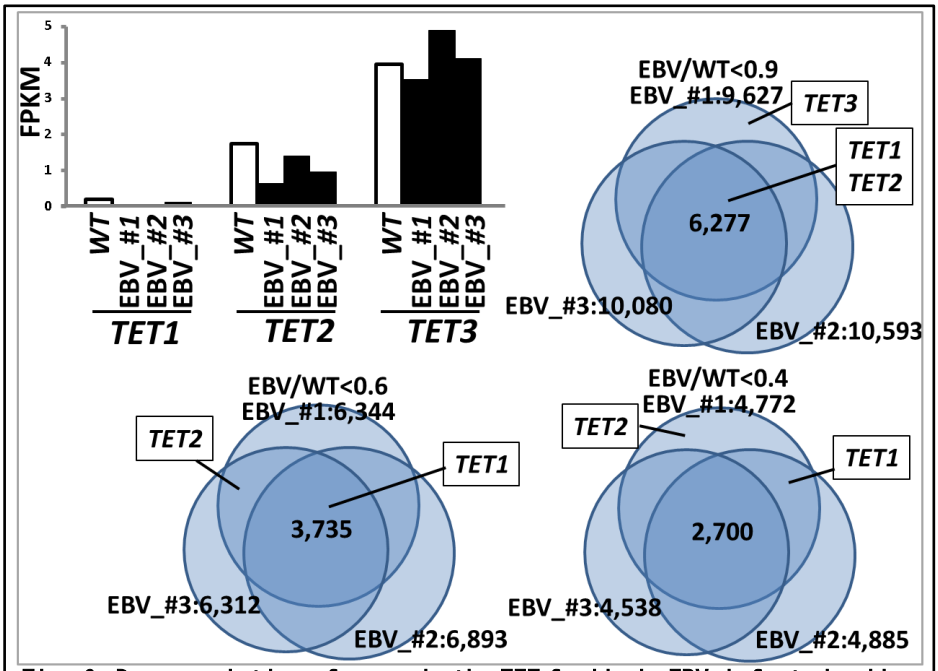

Fig. 6: Downregulation of genes in the TET family in EBV-infected cells. Genome-wide gene expression levels were analyzed by RNA-seq. The diagram indicates the numbers of genes downregulated by <0.9-fold, <0.6-fold and <0.4-fold respectively in three EBV-infected clones compared with MKN7_WT. For example for <0.60-fold change, *TET1* was included in the 3,735 genes that were downregulated in all the three clones, and *TET2* was downregulated in MKN7_EB#1 and EB#3.

4-1-2. Validation by real-time RT-PCR and immunoblotting analyses

Quantitative RT-PCR to analyze TET family genes was performed using MKN7_WT and MKN7_EB#1. *TET2* was markedly downregulated after EBV

infection, and *TET1* was expressed at low levels in both cells (Fig. 7A). Another gastric epithelial cell line, GES1 (GES1_WT), established from normal gastric epithelial cells, was also infected with EBV (GES1_EBV). This cell line acquires extensive hypermethylation in response to *in vitro* EBV infection (Matsusaka et al., unpublished data). Based on a quantitative RT-PCR analysis, the three TET family genes were downregulated in GES1_EBV, especially *TET2* (Fig. 7B). Immunoblotting analyses also showed that TET2 protein expression was significantly repressed by EBV infection in both MKN7 (Fig. 7C) and GES1 cells (Fig. 7D). Since TET2 expression was markedly decreased after EBV infection in both MKN7 and GES1 cells among the three TET family genes, and TET2 is involved in cytosine hydroxymethylation, we hypothesized that *TET2* downregulation contributes to methylation, at least partially.

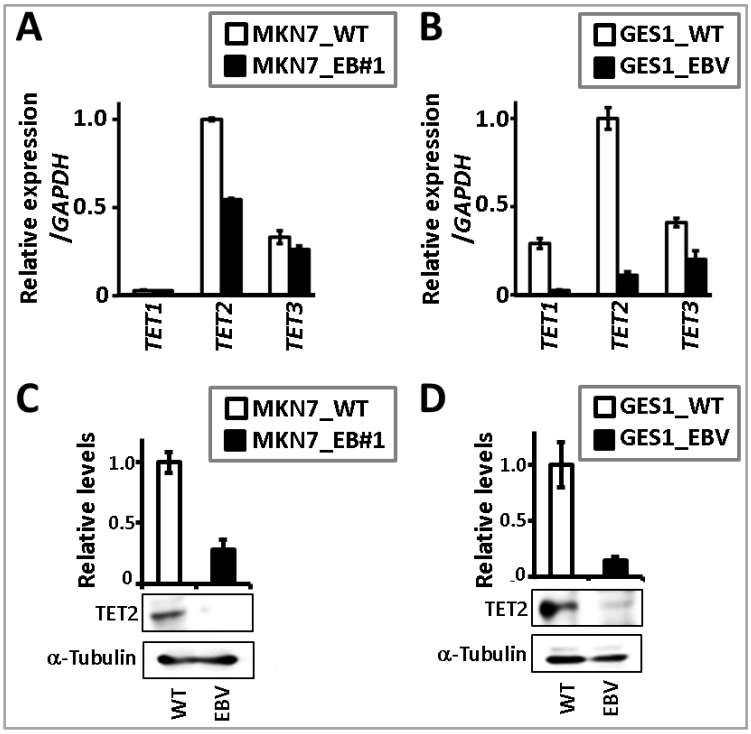

Fig. 7: Expression of TET family. (A, B) Expression of TET family in mRNA level in MKN7 (A) and GES1 cell lines (B). (C, D) Protein expression was also downregulated after EBV infection in MKN7 (C) and GES1 cell lines (D).

4-2. Downregulation mechanism of *TET2*

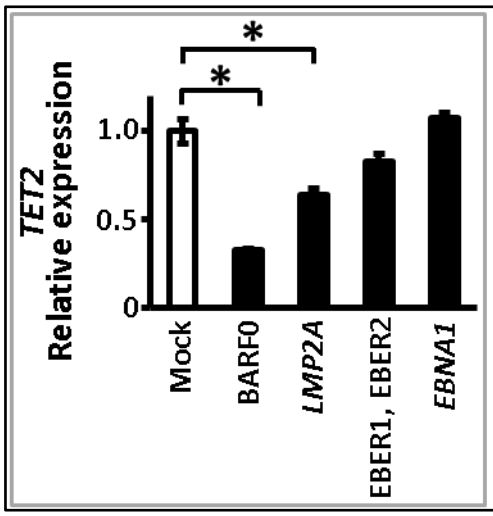

Fig. 8: Viral factors to downregulate *TET2*. Expression level of *TET2* was analyzed in MKN7 overexpressing EBV latent genes, *LMP2A*, *ENBA1*, BARF0 and EBER1/2, and normalized against that of *GAPDH*. BARF0 and *LMP2A* significantly downregulated *TET2*.

To investigate the mechanism which downregulates *TET2* expression during EBV infection, we first examined if EBV encoded transcripts contribute to decrease of *TET2*. It is known that most of EBV genome is dense methylated in latent infection in gastric epithelial cells, and the limited number of protein-coding genes, *LMP2A* and *EBNA1*, and non-coding transcripts, BARF0 and EBER1/2, are allowed to express [49, 54]. These are called latent genes, and were overexpressed in MKN7 cells. It was found that *TET2* was downregulated to 0.35-fold by BARF0, and 0.65-fold by *LMP2A* (Fig. 8).

To examine effects of cellular transcripts on *TET2* expression, we next

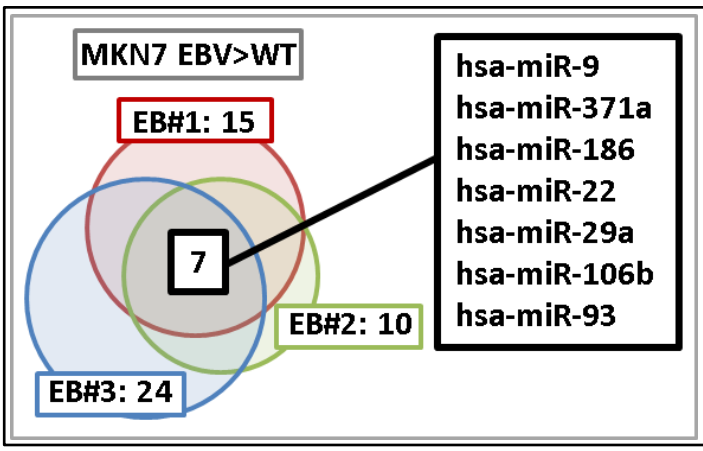

Fig. 9: MiRNA microarray analysis. Expression of 2,549 human miRNAs in MKN7_WT and the three EBV-infected clones were analyzed using a microarray. Among 87 miRNAs that were expected to target *TET2* based on an *in silico* analysis, 15, 10, and 24 miRNAs were upregulated >1.3-fold in MKN7_EB#1, EB#2, and EB#3, respectively, and 7 miRNAs were upregulated in all the three clones.

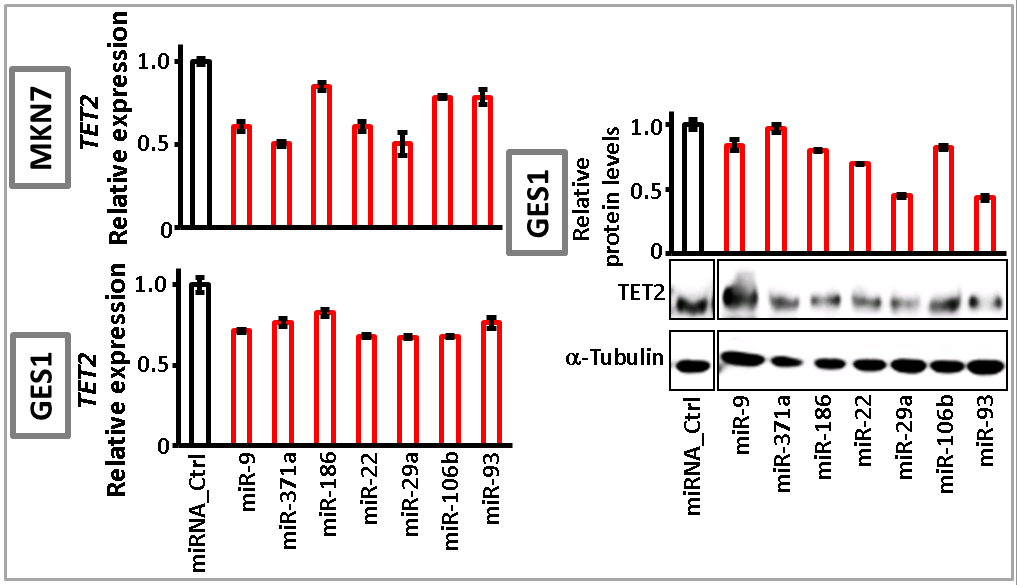

Fig. 10: Validation of miRNA microarray. The 7 miRNAs were transfected into MKN7 and GES1, and real-time RT-PCR showed that *TET2* expression levels decreased by 50–85% after 48 h. Immunoblotting analysis was also conducted for TET2 and α-Tubulin expression in GES1 transfected with the 7 miRNAs.

conducted miRNA microarray analysis for human miRNA expression in MKN7_WT and three EBV-infected MKN7 clones. Of 83 candidate miRNAs that targeted *TET2* according to *in silico* data (http://microrna.org/), 7 miRNAs were commonly upregulated in the three EBV-infected MKN7 clones compared to MKN7_WT (Fig. 9). To validate whether these 7 miRNAs decrease *TET2* expression, we transfected the miRNAs into MKN7 and another cell line GES1 and performed quantitative RT-PCR to analyze *TET2*. All the 7 miRNAs decreased *TET2* expression levels to 50–85% in MKN7 as well as GES1 cells, suggesting that the upregulation of these 7 miRNAs downregulates *TET2*, at least partly (Fig. 10). Downregulation of TET2 protein was also confirmed by immunoblotting analysis (Fig. 10). Transfection of miR-29a and miR-93 induced more marked downregulation in protein level than in mRNA level, which is consistent with reports explaining that miRNA works for not only mRNA cleavage but also translational repression [72]. The predicted regions for the 7 miRNAs to bind 3′ UTR of *TET2* were analyzed using microRNA.org [73]

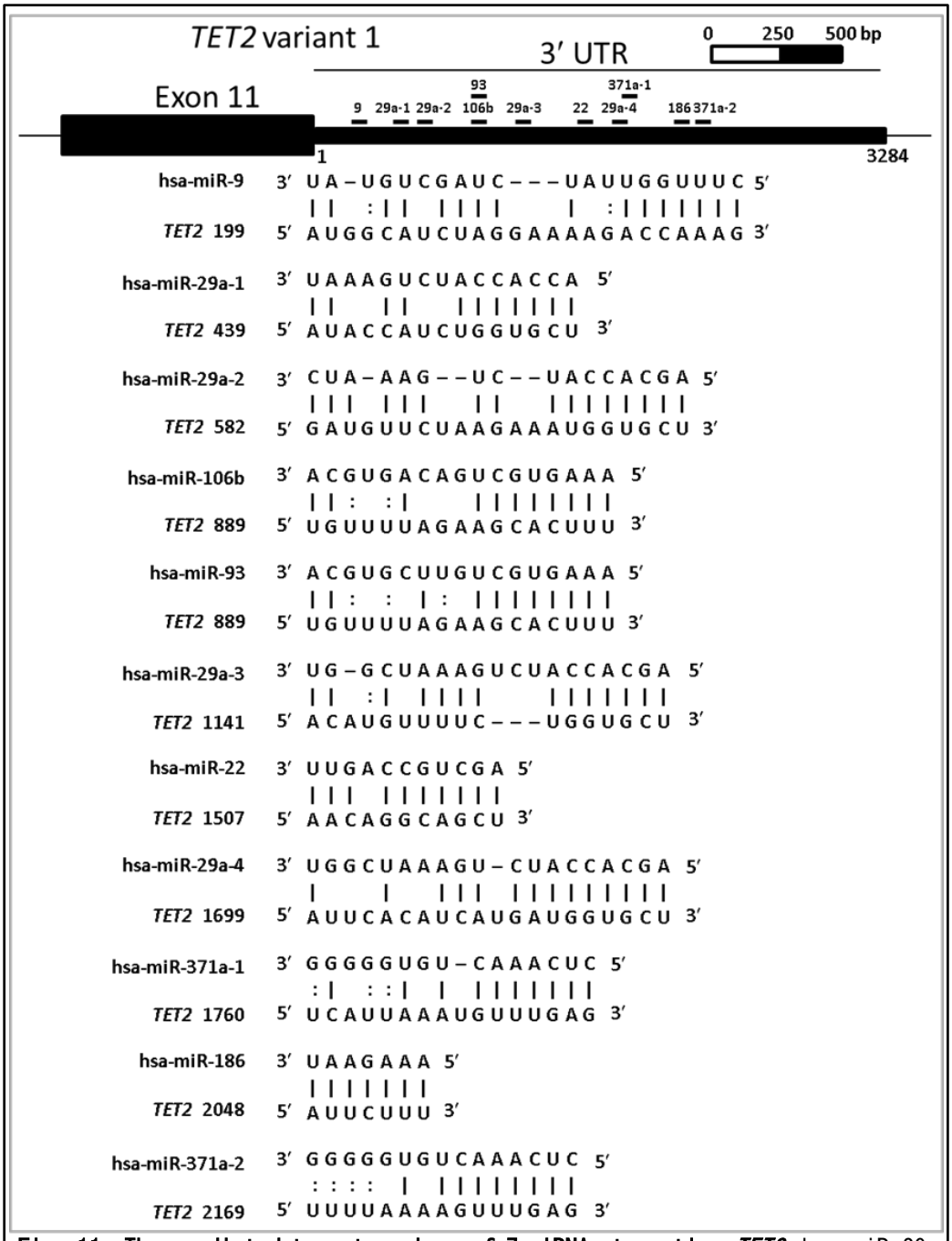

Fig. 11: The predicted target regions of 7 miRNAs targeting *TET2*. hsa-miR-29a has 4 sites and has-miR-371a has 2 sites to bind 3′ UTR of *TET2*.

and shown (Fig. 11).

We next searched for human genes targeted by viral miRNAs encoded in EBV genome, using computational software Vir-Mir database [74, 75]. No EBV miRNA was found to target human *TET2* (Supplementary Table S1).

4-3. Hydroxymethylation target genes of TET2

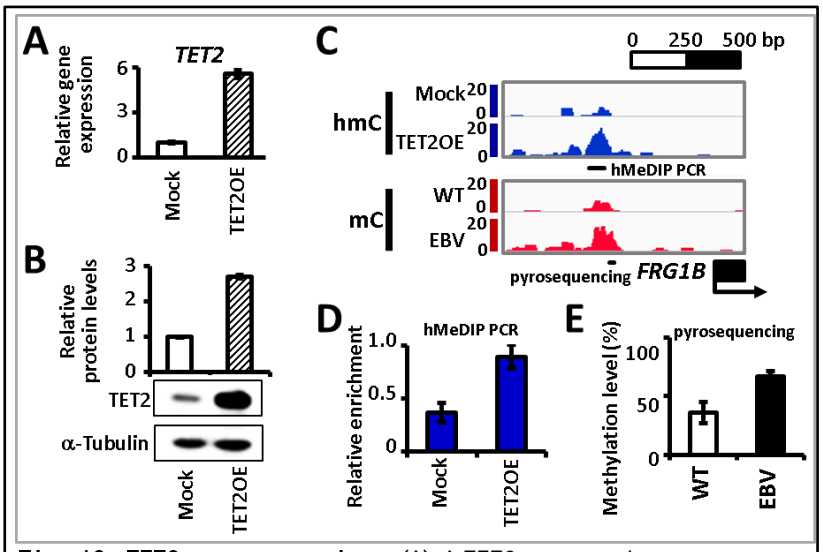

Fig. 12: TET2 overexpression. (A) A TET2-expressing vector was transfected into GES1 and the expression level of *TET2* relative to GAPDH at 30 days after transfection was analyzed by real-time RT-PCR. (B) Immunoblotting analysis was conducted for TET2 and α-Tubulin expressed in Mock and TET2OE cells. (C) Representative results of hMeDIP-seq and MeDIP-seq around *FRG1B* are shown. The hydroxymethylation level of the region was increased in cells with TET2 overexpression, whereas the methylation level was increased in EBV infection. (D) hMeDIP was repeated, and increase of hmC in 5′ region of *FRG1B* was validated by hMeDIP-PCR at the region indicated in Figure 12C, and normalized against a positive control region *NEDD9*. (E) Increase of mC was validated by quantitative pyrosequencing assay at the region indicated in Figure 12C.

To identify hydroxymethylation target genes in response to TET2, we performed hMeDIP-seq and MeDIPseq analyses using GES1 cells transfected with a TET2-overexpression vector (TET2OE) or mock vector (Mock) (Fig. 12). We detected 2,619 hydroxymethylation target genes that showed 5hmC peaks in the hMeDIP-seq analysis in both Mock and TET2OE cells. Among 2,619 genes, a significant number of genes overlapped with the 3,029 methylation target genes in EBV-infected cells (527 genes, $P<1\times10^{-15}$, χ^2 test) (Fig. 13A). These results suggested that many genes remain unmethylated owing to hydroxymethylation by TET2 before EBV infection, leading to methylation via TET2 depression after EBV infection. However, other enzymes, in addition to TET2, may produce 5hmC. To

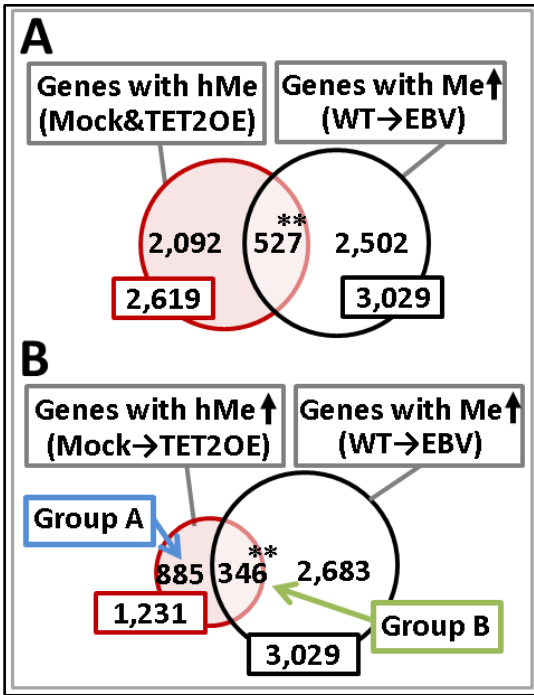

Fig. 13: Overlap of hydroxymethylated genes and methylated genes by EBV infection. (A) Among 2,619 hydroxymethylation target genes showing hydroxymethylation peaks within ±1 kb of the TSS in both Mock and TET2OE cells, 527 genes (20.1%) were methylation target genes during EBV infection ($P<1\times10^{-15}$). (B) Among hydroxymethylation target genes in TET2OE, 1,231 genes showing increased hydroxymethylation from Mock to TET2OE were extracted as hydroxymethylation target genes by TET2. As many as 346 genes (28.1%) were methylation target genes during EBV infection ($P<1\times10^{-15}$).

specifically analyze TET2 target genes, we focused on 1,231 genes that did not possess 5hmC peaks in Mock cells, but acquired 5hmC peaks after TET2 overexpression. Among these 1,231 genes, more significant overlap with methylation target genes was found (346 genes, $P<1\times10^{-15}$, χ^2 test) (Fig. 13B). These

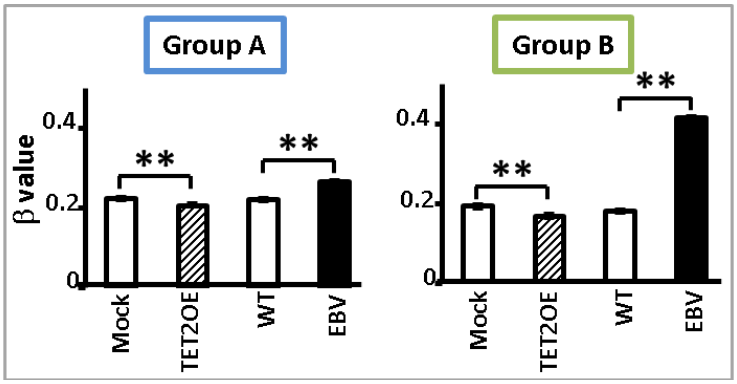

Fig. 14: Increased methylation in EBV infected cells and decreased methylation in TET2 overexpressing cells. Methylation levels of hydroxymethylation target genes by TET2 were analyzed by Infinium, and average β values are shown. The 346 methylation target genes during EBV infection (Group B) showed marked increases of β values in GES1_EBV, while the other 885 genes that were not extracted as methylation target genes during EBV infection (Group A) showed slight, but still significant increases of β values in GES1_EBV. Both genes in Group A and Group B showed decreases of β values in TET2OE.

346 methylation target genes during EBV infection (Group B in Fig. 13B) showed significant decreases in β values when TET2 was overexpressed ($P<1\times10^{-15}$), and significant increases after EBV infection ($P<1\times10^{-15}$) (Fig. 14). The other 885 genes that were not extracted as methylation target genes during EBV infection (Group A in Fig. 13B) showed slight, but significant increases in β values in GES1_EBV ($P<1\times10^{-15}$) (Fig. 14). These results suggested that hydroxymethylation by TET2 protects the unmethylated status of genes before EBV infection, and decreased TET2 via EBV infection could promote the methylation of these target genes.

4-4. Knockdown of *TET2* accelerated *de novo* methylation during EBV infection

If TET2 is a resistance factor for methylation acquisition, the knockdown of

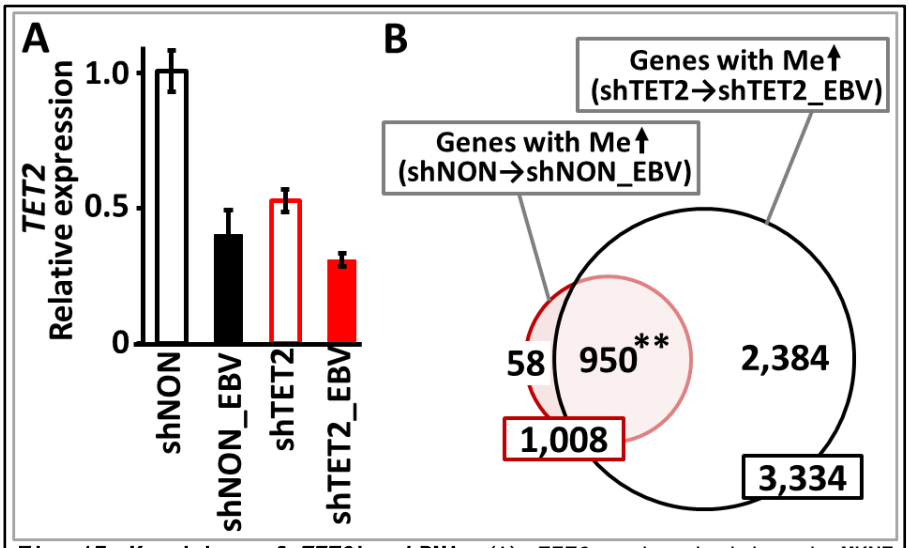

Fig. 15: Knockdown of *TET2* by shRNA. (A) *TET2* was knocked down in MKN7 by infection with a shRNA lentivirus targeting *TET2* (shTET2). shNON, MKN7 cells infected with non-targeting shRNA. Real-time RT-PCR analysis of *TET2*, normalized against *PPIA*, showed that the expression level of *TET2* was lower in shTET2 than shNON cells. After EBV infection of shNON and shTET2 cells (shNON_EBV and shTET2_EBV, respectively), *TET2* levels were lower in shNON_EBV than shNON, and decreased further in shTET2_EBV compared with shTET2. (B) While 1,008 genes acquired promoter methylation by EBV infection in shNON cells, 3,334 genes showed *de novo* methylation by EBV infection in shTET2 cells, and 950 genes (94.2%) overlapped in the two cell types. "*De novo* methylation" was defined as >2 probes with β <0.2 that increased to β >0.4 within ±1 kb of TSS after EBV infection.

TET2 might accelerate methylation acquisition during EBV infection. We therefore knocked down *TET2* in MKN7 cells by shRNA (shTET2) and infected shTET2 cells with EBV (shTET2_EBV) (Fig. 15A). Non-targeting shRNA lentivirus was also transfected to obtain control cells (shNON). Seventy two days after EBV infection, DNA methylation levels were quantitatively investigated using the Infinium 450k BeadChip array. When shNON cells were infected with EBV (shNON_EBV), 1,008 genes acquired *de novo* promoter methylation, whereas 3,334 genes acquired *de novo* promoter methylation when shTET2 cells were infected with EBV (shTET2_EBV), including as many as 950 of the 1,008 genes in shNON-EBV. While most (94%) of the methylation target genes in shNON cells were also methylated in shTET2 cells, 2,384 genes were newly methylated in *TET2*-depleted cells (Fig. 15B).

Next, among 10,829 genes that were defined as unmethylated in shNON_EBV cells, 1,419 genes acquired *de novo* methylation in shTET2_EBV cells. In contrast, among 7,953 genes defined as unmethylated in shTET2_EBV cells, only 28 showed *de novo* methylation in shNON_EBV ($P<1\times10^{-15}$) (Fig. 16). There are 2 kinds of *de*

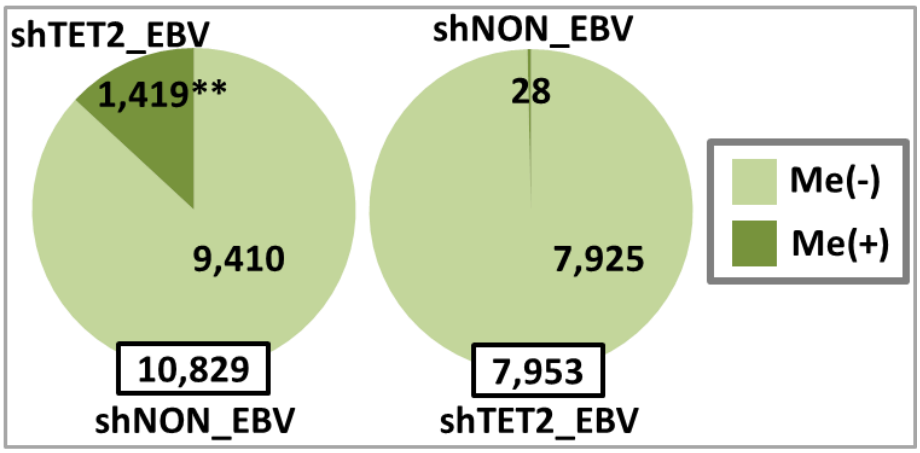

Fig. 16: Increase in methylation target genes by EBV infection in *TET2*-depleted cells. Among 10,829 genes defined as unmethylated in shNON_EBV cells, 1,419 genes acquired methylation in shTET2_EBV. In contrast, only 28 genes showed differential methylation in shNON_EBV among 7,953 genes that were unmethylated in shTET2_EBV ($P<1\times10^{-15}$).

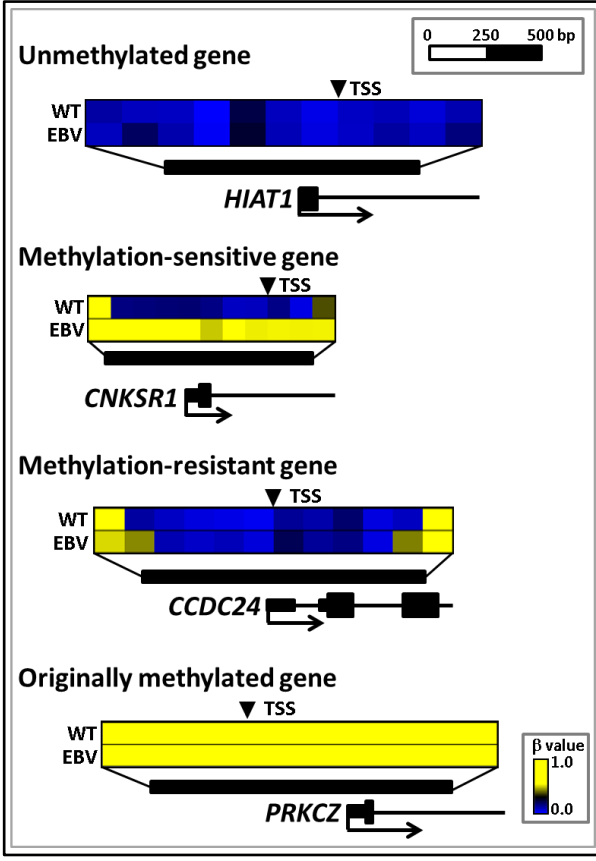

Fig. 17: Representation of an unmethylated gene (*HIAT1*), a methylation-sensitive gene (*CNKSR1*), a methylation-resistant gene (*CCDC24*), and an originally methylated gene (*PRKCZ*) during EBV infection. When *de novo* methylation is induced by EBV infection, there are two types of methylation acquisition around promoter regions. Methylation-sensitive genes undergo complete methylation around the promoter regions; methylation-resistant genes undergo methylation in regions surrounding their promoters, but maintain an unmethylated status in narrow regions around the TSS (Funata et al., unpublished data).

novo methylation induced by EBV infection. One is methylation-sensitive which acquire complete methylation in their promoter regions, and the other is methylation-resistance which underwent *de novo* methylation in the region surrounding the TSS, but maintained an unmethylated status in narrow regions around the TSS (Funata et al., unpublished data) (Fig. 17). Among the 1,419 genes that showed *de novo* methylation in shTET2_EBV cells, 498 were methylation-sensitive genes (Fig. 18A), and 550 were methylation-resistant genes (Fig. 18B). The 498 methylation-sensitive genes, e.g., *EGFR*, showed decreased expression in shTET2_EBV cells, but the 550 methylation-resistant genes, e.g., *C1orf109*, did not show decreased expression in shTET2_EBV cells (Fig. 18).

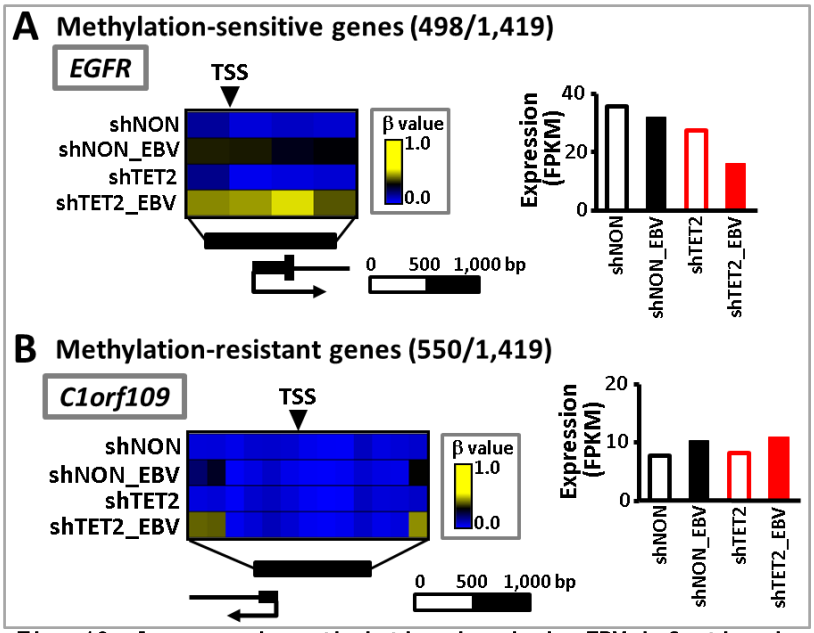

Fig. 18: Increase in methylation levels by EBV infection in *TET2*-depleted cells. (A) Among the 1,419 genes showing differential methylation in shTET2_EBV (Fig. 16), 498 genes became methylation-sensitive genes, acquiring complete methylation in the promoter regions. Representative methylation levels in the *EGFR* promoter region are shown. (B) Among the 1,419 genes showing differential methylation in shTET2_EBV (Fig. 16), 550 were methylation-resistant genes, acquiring methylation in the regions surrounding promoters, but maintaining an unmethylated status in narrow regions around the TSS. The methylation levels of the *C1orf109* promoter region are representatively shown.

When methylation-sensitive genes (i.e., genes that were completely methylated in their promoter regions) in shNON_EBV and shTET2_EBV cells were compared, 298 overlapped (Fig. 19A and B) and 1,210 were methylated only in shTET2_EBV. The 1,210 genes that acquired complete methylation in shTET2_EBV cells only (Group C in Fig. 19A), exhibited significantly higher methylation levels ($P<1\times10^{-15}$, Wilcoxon signed-rank test) and lower expression levels ($P=2\times10^{-12}$, Wilcoxon signed-rank test) in shTET2_EBV cells than shNON_EBV cells (Fig. 20). Interestingly, the 298 genes that acquired complete methylation in both shNON_EBV and shTET2_EBV cells (Group D in Fig. 19) also showed significantly higher methylation levels ($P<1\times10^{-15}$, Wilcoxon signed-rank test) and lower

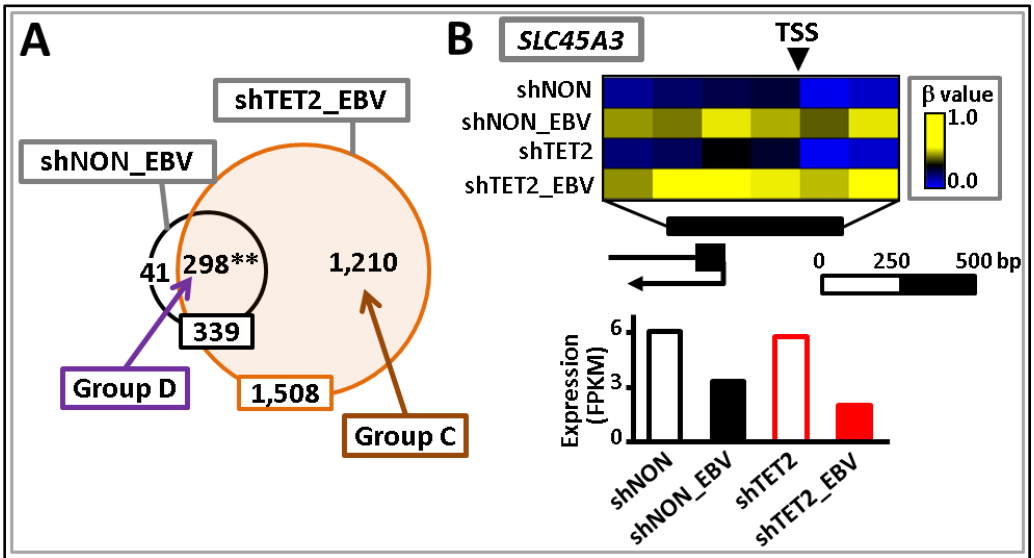

Fig. 19: Increase in methylation levels, even in methylation-sensitive genes, by *TET2* knockdown. (A) The number of methylation-sensitive genes was 339 in shNON_EBV, and this increased to 1,508 in shTET2_EBV, with 298 overlapping genes (87.9%). (B) *SLC45A3* was one of the 298 genes showing complete promoter methylation in both shNON_EBV and shTET2_EBV. The methylation level was even higher and the expression level was even lower in shTET2_EBV compared with shNON_EBV.

expression levels (P=7×10^{-5}, Wilcoxon signed-rank test) in shTET2_EBV cells than shNON_EBV cells (Fig. 20).

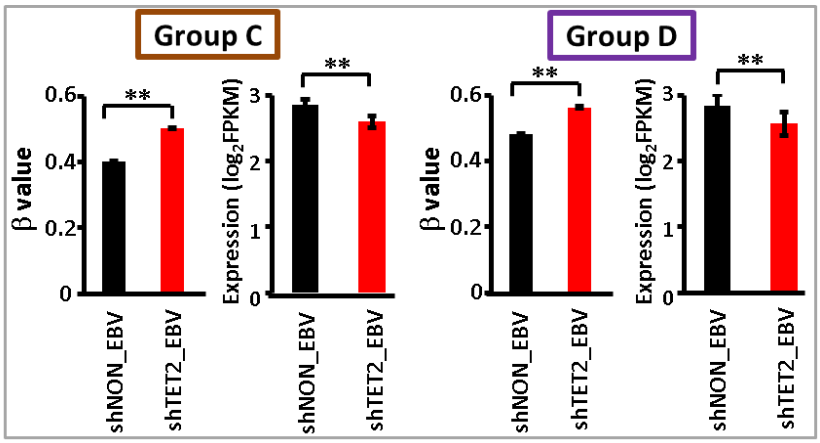

Fig. 20: Increase in methylation levels and decrease in expression levels, even in methylation-sensitive genes, by *TET2* knockdown. The 1,210 methylation-sensitive genes in shTET2_EBV only (Group C in Fig. 19A) showed significant increases in methylation levels (P<1×10^{-15}) and significant decreases in expression levels (P=2×10^{-12}) in comparison with the levels observed in shNON_EBV. The 298 methylation-sensitive genes in both shNON_EBV and shTET2_EBV (Group D in Fig. 19A) also showed significant increases in methylation levels (P<1×10^{-15}) and significant decreases in expression levels (P=7×10^{-5}) in comparison with the levels observed in shNON_EBV.

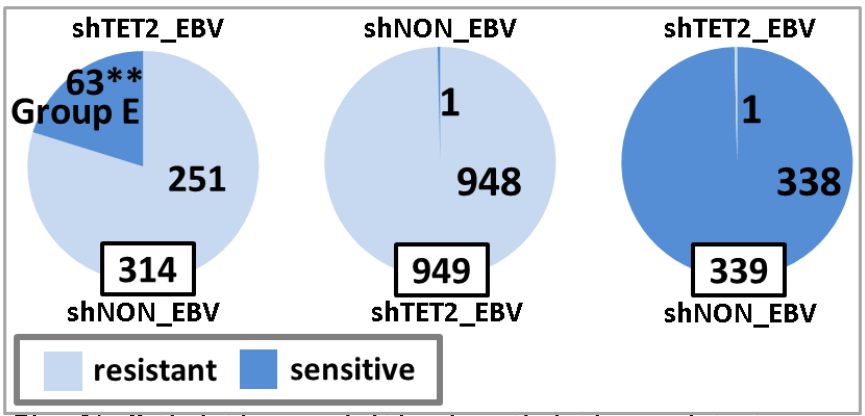

Fig. 21: Methylation acquisition in methylation-resistant genes in shNON_EBV. Among 314 methylation-resistant genes in shNON_EBV, which showed protection of the unmethylated status around the TSS in shNON_EBV, 63 genes acquired complete promoter methylation in shTET2_EBV, i.e., they became methylation-sensitive by *TET2* depletion (*left*). Among 949 methylation-resistant genes in shTET2_EBV, in contrast, only 1 gene was methylation-sensitive in shNON_EBV (*middle*). Among 339 methylation-sensitive genes in shNON_EBV, only 1 gene became methylation-resistant in shTET2_EBV (*right*).

Among the 314 methylation-resistant genes in shNON_EBV cells that underwent *de novo* methylation in the region surrounding promoters, but maintained an unmethylated status in narrow regions around the TSS, 63 acquired complete methylation in the promoter regions (i.e., were classified as methylation-sensitive genes) in shTET2_EBV (Group E in Fig. 21). These genes did not show decreased expression in shNON_EBV cells, presumably because their TSS was protected from methylation, but they were significantly downregulated in shTET2_EBV cells, presumably because their TSS acquired complete methylation (P=0.006) (Fig. 22A). Complete acquisition of methylation in *HTRA1* and its repression in shTET2_EBV cells were representatively shown (Fig. 22B), and the methylation changes were validated by pyrosequencing (Fig. 22C). Among 949

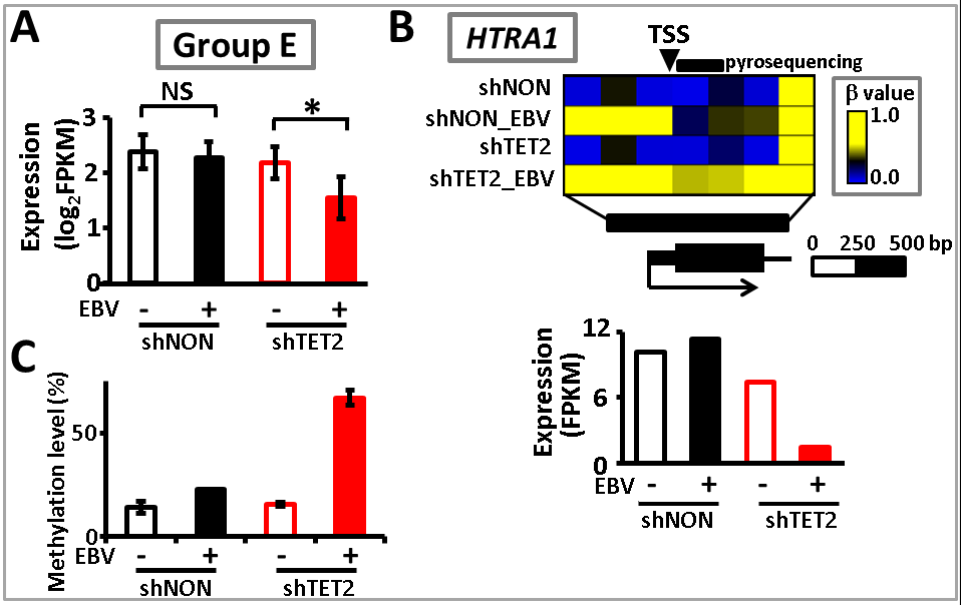

Fig. 22: Expression and a representative gene of methylation acquired gene. (A) The 63 genes that became methylation-sensitive in shTET2_EBV (Group E in Fig. 21) showed significant decreases in expression by EBV infection, only when transfected into *TET2*-depleted cells (P=0.006). (B) Representative methylation status of *HTRA1* is shown. *HTRA1* promoter remained unmethylated in shNON_EBV, but acquired complete promoter methylation in shTET2_EBV. Expression was retained in shNON_EBV, but silenced in shTET2_EBV cells. (C) Quantitative pyrosequencing assay was conducted to validate unmethylated status in shNON_EBV and acquisition of promoter methylation in shTET2_EBV at the 5′ region of *HTRA1* indicated in Figure 22B.

methylation-resistant genes in shTET2_EBV cells, only 1 gene acquired complete methylation in the promoter region in shNON_EBV cells ($P<1\times10^{-15}$, χ^2 test) (Fig. 21). Additionally, only 1 gene became methylation-resistant in shTET2_EBV cells among 339 methylation-sensitive genes in shNON_EBV cells ($P<1\times10^{-15}$, χ^2 test) (Fig. 21).

4-5. No methylation was induced by *TET2* knockdown alone

To analyze the effect of TET2 knockdown, shTET2 and shNON cells were cultured without EBV infection, and methylation alterations were analyzed using the Infinium BeadChip array (Fig. 23). Knockdown of *TET2* expression was confirmed in mRNA and protein levels (Fig. 23A). Of 13,150 genes that were defined

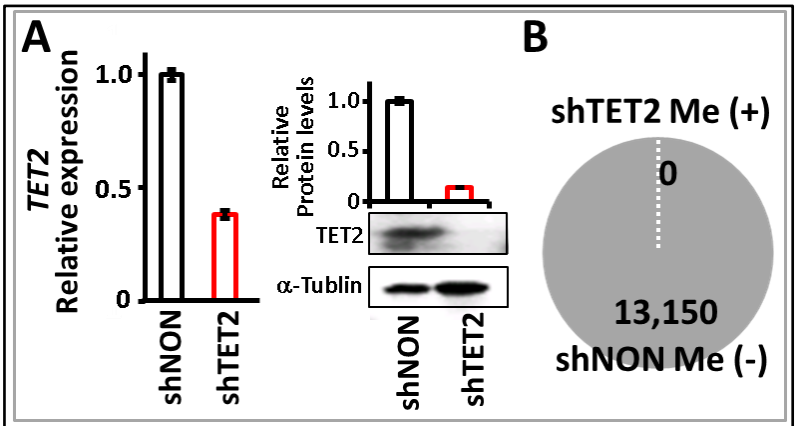

Fig. 23: No methylation induction by *TET2* knockdown itself, without EBV infection. (A) Real-time RT-PCR showed a decrease in TET2 expression in shTET2 compared with shNON cells. Immunoblotting analysis was conducted for TET2 and α-Tubulin expression in shNON and shTET2 cells. (B) Among 13,150 genes that were unmethylated in shNON, none showed methylation in shTET2, i.e., *de novo* methylation did not occur by *TET2* knockdown alone.

as unmethylated in shNON cells, none acquired *de novo* methylation in shTET2 cells (Fig. 23B). No induction of methylation was detected in unmethylated genes, methylation-sensitive genes, or methylation-resistant genes (Fig. 24). These results suggested that *TET2* depletion is not sufficient to induce *de novo* methylation, and that EBV infection might also trigger other mechanisms in addition to repression of a resistant factor.

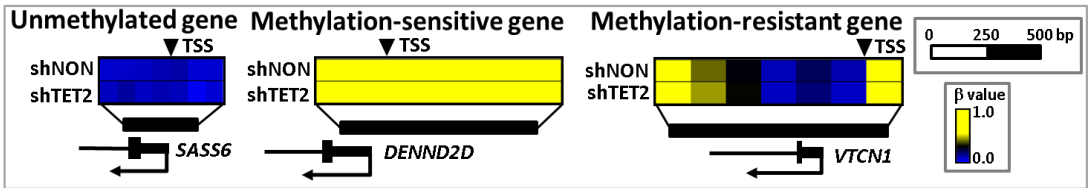

Fig. 24: Three representative genes showing no methylation alterations in response to the knockdown of *TET2*.

5. Discussion

The previous study showed that EBV infection causes extensive DNA methylation in gastric epithelial cells [42, 54]. In this study, we performed a transcriptome analysis to identify candidate critical factor(s) that contribute to the epigenetic alterations and found that *TET2* was downregulated during EBV infection.

5-1. Downregulation mechanism of *TET2*

For myeloid cancer patients, 15% have *TET2* mutation and it leads to DNA hypermethylation [33, 37-40]. For EBV infected gastric cancer cells in this study, there are 2 kinds of possible mechanisms to downregulate *TET2* expression; EBV derivative factors such as EBV encoded transcripts and miRNA, and human derivative factors such as human miRNA, histone modification changes and DNA methylation changes. Firstly for EBV encoded transcripts, protein-coding genes, *LMP2A* and *EBNA1*, and non-coding transcripts, BARF0 and EBER1/2, are allowed to overexpress in MKN7 cell line, and it was found that *TET2* was downregulated by BARF0 and *LMP2A* (Fig. 8). Secondly for EBV encoded miRNA, none was found to target human *TET2* in this study (Supplementary Table S1). Thirdly for human encoded miRNA, 7 miRNAs were upregulated by EBV infection in all 3 clones investigated, and all the 7 miRNAs decreased *TET2* expression in MKN7 as well as GES1 cells, suggesting that the upregulation of these 7 miRNAs downregulates *TET2*, at least partly (Fig. 10). Transfection of miR-29a and miR-93 induced more marked downregulation in protein level than in mRNA level, which is consistent with reports explaining that miRNA works for not only mRNA cleavage but also

translational repression [72]. Fourthly for histone changes, active marks H3K4me3 and H3K27ac decreased after EBV infection in the promoter region of *TET2,* and it may contribute to *TET2* downregulation (Funata et al., unpublished data). Lastly for DNA methylation changes, methylation level was slightly increased in *TET2* promoter region (unpublished data). It is required to infect EBV to TET2 overexpressing cells. If DNA methylation doesn't increase after EBV infection in TET2 overexpressing cells, downregulation of *TET2* may contribute to *TET2* promoter methylation itself.

5-2. TET2 overexpression analysis

Since a ChIP-seq-grade anti-TET2 antibody is not available, the detection of TET2 binding regions by ChIP-seq was not possible; accordingly, hydroxymethylated DNA regions by TET2 overexpression were detected by hMeDIP-seq. We performed a hMeDIP-seq analysis using Mock and TETOE cells, and identified hydroxymethylated promoter regions that showed 5hmC peaks in both cells. These hydroxymethylated genes overlapped significantly with methylation target genes in EBV-infected cells. However, hydroxymethylation of these genes is not necessarily caused by TET2, but might be related to other TET family proteins [16, 17, 20]. To detect hydroxymethylated genes converted by TET2, we focused on genes that showed no 5hmC peaks in Mock cells, but were hydroxymethylated in TET2OE cells. These hydroxymethylation target genes by TET2 showed markedly significant overlap with methylation target genes in EBV-infected cells. These results suggested that hydroxymethylation by TET2 is involved in protecting DNA from methylation, and that the repression of this

resistance factor against DNA methylation may contribute to *de novo* methylation acquisition during EBV infection.

5-3. Analysis of *TET2* knockdown and EBV infection

When *TET2* was knocked down, more genes acquired *de novo* methylation by EBV infection, including genes that were resistant to *de novo* methylation around TSS in shNON_EBV cells (Fig. 15-22). This also indicated that TET2 functions as a resistance factor against DNA methylation during EBV infection, and *TET2* depletion leads to a loss of protective mechanisms and the acquisition of *de novo* methylation. But these genes were not necessarily hydroxymethylation target genes in shNON_EBV cells (data not shown). It is possible that these regions could be further oxidized to 5fC and 5caC, which would be excised by thymine DNA glycosylase to be converted to unmodified cytosine [16-18, 20]. Furthermore, TET1, TET2, and TET3 reportedly bind to DNA without the appearance of 5mC or 5hmC, indicating that they may protect DNA from methylation by physically binding to DNA, regardless of their catalytic activity [76, 77]. Further investigations are necessary to clarify how TET2 works as a resistant factor against methylation acquisition.

5-4. *TET2* knockdown itself

Interestingly, *TET2* knockdown itself did not induce methylation. It has been suggested that additional factor(s) are required to induce *de novo* methylation owing, for example, to the increase in methylation pressure via upregulated DNMTs [42], and that TET2 might function in resistance against those factors. EBV

infection might also trigger this pressure, and methylation might not be induced without EBV infection, even if the resistant factor is depleted. In transcriptome analysis, histone deacetylase *HDAC8* was downregulated in EBV-infected clones, and genes encoding Polycomb group proteins such as *SUZ12* and *BMI1* were upregulated in EBV-infected clones. Expression changes of these epigenetic factors might also be important for epigenomic alteration after EBV infection. To fully clarify the molecular mechanism underlying the unique epigenomic phenotype with extensive hypermethylation in EBV-positive gastric cancer, further investigations are necessary to identify factors that induce methylation pressure and to determine how the factors are activated and recruited to methylation target genes.

5-5. DNA methylation in other malignancies

Other than EBV infected gastric cancer, *Helicobacter pylori* (*H. pylori*) is also associated with gastric carcinogenesis [78, 79] and its infection induces abberant hypermethylation in promoter regions of tumor-suppressor genes such as *p16*, *LOX* and *CDH1* [80]. A specific single nucleotide polymorphism of *IL-1*β increases gastric cancer risk and higher incidence of *CDH1* promoter methylation [81, 82]. TET2 is a member of TET family proteins converting DNA 5mC to 5hmC, 5fC, and 5caC [16-18]. *TET2* mutations in various myeloid cancer [37, 38], or mutations of *IDH1* and *IDH2* which can inhibit DNA hydroxymethylation by TET enzymes [28-32], in low-grade gliomas, have been reported and cause DNA hypermethylation phenotype. Here, we showed that *TET2* could be downregulated by EBV infection in gastric epithelial cells via EBV transcripts and upregulation of human miRNAs targeting *TET2*. Other TET family genes are also of great interest since several

comprehensive genomic analyses in gastric cancer revealed mutation of *TET1* in microsatellite-stable gastric cancers [47], and *TET1* was also downregulated by EBV infection (Fig. 7).

6. Conclusions

In summary, we found that TET2 is downregulated during EBV infection via expression of EBV transcripts and upregulation of human miRNAs targeting *TET2*, and that TET2 may function as a resistance factor against DNA methylation in gastric epithelial cells. While *TET2* depletion itself does not increase the methylation level, the downregulation of *TET2* contributes to methylation induction by EBV infection.

7. Materials and methods

7-1. Cell culture and treatment

The gastric cancer cell line MKN7 was obtained from Riken BioResource Center Cell Bank and was authenticated by the cell bank using short tandem repeat PCR. The normal fetal gastric mucosal cell line GES1, which was immortalized by SV40, was obtained from the Beijing Institute for Cancer Research [83]. These cells were cultured in RPMI 1640 (Wako, Tokyo, Japan). with 10% fetal bovine serum (HyClone SH30910.03; GE Healthcare, Chicago, IL, USA) and penicillinstreptomycin (Sigma-Aldrich, St. Louis, MO, USA) at 37 °C in a 5% CO2 incubator. For *in vitro* EBV infection of MKN7 and GES1 cells, the Akata system was used as previously reported [70]. DNA was extracted using the QIAamp DNA Mini Kit (Qiagen, Hilden, Germany). RNA was extracted using the RNeasy Mini Kit (Qiagen) following the manufacturer's protocol, and treated with DNaseI (Qiagen).

7-2. Plasmid construction

cDNA of TET2 transcript variant 1 was inserted into the EcoRV site of the pcDNA5/TO vector. The CMV promoter region of pcDNA5/TO was changed to a CAG promoter and hygromycin resistance gene was changed to puromycin resistance gene. Beginning at 16 h before transfection, cells were cultured in RPMI 1640 medium with 10% fetal bovine serum without penicillin-streptomycin in 6-well dishes, and 2 µg of the TET2-expressing vector or mock vector was transfected using Lipofectamine 2000 (Invitrogen, Carlsbad, CA, USA). The transfectants were selected using 2 µg/mL puromycin (Sigma-Aldrich).

7-3. Knockdown by shRNA

To knock down *TET2*, double-stranded oligonucleotide DNA encoding small hairpin RNA (shRNA) against *TET2* was cloned into the pLKO.1 vector between EcoRI and AgeI sites. Oligonucleotide sequences for shRNA against *TET2* (shTET2) and control non-target shRNA (shNON) are listed in Table 1. Viral packaging for shRNA retrovirus vectors was performed using 293T cells and FuGENE 6 (Promega, Madison, WI, USA), and medium containing the virus was collected 48 h after transfection.

Table 1: shRNA target sequences in lentivirus

target gene	strand	sequence
TET2	top	CCGGCAGATGCACAGGCCAATTAAGCTCGAGCTTAATTGGCCTGTGCATCTGTTTTTG
	bottom	AATTCAAAAACAGATGCACAGGCCAATTAAGCTCGAGCTTAATTGGCCTGTGCATCTG
no target	top	CCGGCAACAAGATGAAGAGCACCAACTCGAGTTGGTGCTCTTCATCTTGTTGTTTTTG
	bottom	AATTCAAAAACAACAAGATGAAGAGCACCAACTCGAGTTGGTGCTCTTCATCTTGTTG

Oligonucleotide sequences of shRNA against *TET2* (shTET2) and control non-target shRNA (shNON), inserted into pLKO.1 vector between EcoRI and AgeI sites, are shown, and 21-mer target sequences are underlined.

7-4. Real-time RT-PCR

cDNA was prepared from 1 μg of total RNA using SuperScript III Reverse Transcriptase (ThermoFisher, Waltham, MA, USA). Real-time RT-PCR was performed using SYBR Green and CFX96 Touch Real-Time PCR (Bio-Rad Laboratories, Hercules, CA, USA). The quantity of mRNA for each gene in a sample was estimated by comparisons with standard samples that contained 101 to 106 gene copies. These levels were normalized to those of GAPDH and PPIA, as previously described [84]. The PCR primers and conditions are presented in Table 2.

Table 2: Primer sequences for real-time RT-PCR

Gene	Primer types	Primer Sequence	Anneal (°C)	Product (bp)
TET1	Forward	CTTACAGAGTTTGGCTACACGAT	60	93
	Reverse	CGGGCAACATTTTCATATTCCAC		
TET2	Forward	ATTGAAACAAGACCAAAAGGCTA	60	109
	Reverse	CTTATCACTCAAATCGGAGACA		
TET3	Forward	CCACGGCTTCCAGACAGACCAC	60	91
	Reverse	CAGGGGCAGCTTTCTCCGTTCCCA		
GAPDH	Forward	CCAGGTGGTCTCCTCTGACTTC	60	108
	Reverse	TCATACCAGGAAATGAGCTTGACA		
PPIA	Forward	ACAGTGCTTGCTGGCAGTTAGA	60	130
	Reverse	CAAATCCGCCACCTCTAGGATAG		

7-5. Hydroxymethylated and methylated DNA immunoprecipitation sequencing (hMeDIP-seq / MeDIP-seq)

DNA regions with 5hmC and 5mC were analyzed by hMeDIP and MeDIP, respectively. Fragmentation of genomic DNA was performed using a Picoruptor (Diagenode, Seraing, Belgium) for 10 sets of 30 s on and 30 s off; 20 µg were prepared for hMeDIP and 4 µg were prepared for MeDIP. Fragmented DNA was separated into two tubes, and the tubes were incubated at 95 °C for 10 min followed by 10 min on ice. Immunoprecipitation (IP) Buffer was added to reach 500 µL; subsequently, 2 µL of anti-5hmC antibody (#39769; Active Motif, Carlsbad, CA, USA) was added for hMeDIP or 4 µL of anti-5mC antibody (#33D3; Diagenode) was added for MeDIP. The components of the buffer are described in Table 3. The tubes were rotated at 4 °C overnight. Either 50 µL of 50% Protein A or G Sepharose (GE Healthcare) was added to each tube, followed by rotation at 4 °C for 2 h. The DNA-bead mixture was moved to columns (Corning, New York, NY, USA) and

Table 3: Buffers for DIP experiment

IP buffer	20 mM tris-HCl pH8.0 2 mM EDTA 150 mM NaCl 1.0% Triton-X100
Wash buffer	20 mM Tris-HCl pH8.0 2 mM EDTA 300 mM NaCl 0.1% SDS 1.0% Triton-X100
Elution buffer	25 mM Tris-HCl pH8.0 10 mM EDTA 0.5% SDS 100 mM NaCL

centrifuged at 1,000 × *g* for 1 min at 4 °C. After the flow-through was discarded, the beads were subjected to washing steps using 500 µL of IP Buffer twice, 500 µL of Wash Buffer 5 times, and 500 µL of TE Buffer twice. Beads were transferred to tubes with 400 µL of Elution Buffer and treated with 5 µL of Proteinase K (New England Biolabs, Ipswich, MA, USA) at 55 °C for 1 h. After phenol and chloroform DNA purification, DNA was eluted with 20 µL of distilled water. Enrichment of genomic regions in samples after hMeDIP and MeDIP was validated by real-time PCR using primers listed in Table 4. The hMeDIP and MeDIP DNA was used to prepare library samples using the NEBNext ChIP-Seq Library Preparation Set for Illumina (New England Biolabs) following the manufacturer's protocol. Deep sequencing was performed on the Illumina HiSeq 1500 or NextSeq 500 system using the TruSeq Rapid SBS Kit (Illumina, San Diego, CA, USA) in 50-base single-end mode according to the manufacturer's protocol. The FASTQ reads were mapped to the hg19 reference sequence (UCSC) using BWA with default settings.

The numbers of uniquely mapped reads for hMeDIP samples were 18,046,940 (Mock) and 16,171,239 (TET2 overexpression). For MeDIP samples, there were 16,322,113 (WT) and 17,482,843 (EBV) uniquely mapped reads. These hMeDIP-seq and MeDIP-seq data were submitted to the NCBI BioSample database (http://www.ncbi.nlm.nih.gov/biosample), and the accession numbers are GSM2253669 - GSM2253672. To count the number of reads that were mapped to within ±1 kb of transcription start sites (TSS), Count Reads version 0.2 was used with a window size of 300 bp, and the read count for each window was divided by the total read count and expressed as reads per million mapped sequence reads (RPM), as previously described [85]. Only high-CpG promoter genes (CpG score ≥0.48) [86] were analyzed.

Table 4: Primers for DIP-PCR

Gene	Primer types	Primer Sequence	Anneal (°C)	Product (bp)	hMe/Me
Positive control regions					
NEDD9	Forward	GCTTCATGCCTGCTACATATACTCC	60	106	hMe, Me
	Reverse	TGTGGTCTGTTTGTGTGTAAGAGG			
HMGA2	Forward	CATGCAGAGCAGCGATTTAAA	60	99	hMe
	Reverse	GCCTTTCTTCTTGCTTTAGGACAA			
ACTB	Forward	GAGGGAAATGAGGGCAGGACTT	60	127	hMe
	Reverse	GCTGCCCTGAGGCACTCTTC			
AJAP1	Forward	CGAGCCAGGTCTGAGGC	60	77	Me
	Reverse	GGTCGCTCACCTGAGTCCTA			
Negative control region					
HBB	Forward	GGGCTGAGGGTTTGAAGTCC	60	89	hMe, Me
	Reverse	CCACAGGGTGAGGTCTAAGTG			
Validation region					
FRG1B	Forward	CTTCCTCCATCTCTATCGGGCAA	60	73	hMe
	Reverse	CTAATCCTGGAGGCGGCATCTGA			

7-6. RNA-sequencing (RNA-seq) analysis

Libraries for RNA-seq were prepared using the TruSeq Stranded mRNA Sample Prep Kit (Illumina), following the manufacturer's protocol. Deep sequencing was performed on the Illumina HiSeq 1500 or NextSeq 500 platform using the TruSeq Rapid SBS Kit (Illumina) in 50-base single-end mode according to the manufacturer's protocol. The RNA-seq data were submitted to the NCBI BioSample database (http://www.ncbi.nlm.nih.gov/ biosample), and the accession numbers are GSM2253673 - GSM2253676. TopHat was used to map FASTQ reads and Cufflinks was used for transcript assembly. Gene expression levels were expressed as fragments per kilobase of exon per million mapped sequence reads (FPKM). When expression alterations were analyzed, expression levels are presented as log2 FPKM values, excluding genes with log2 FPKM ≤ 0.

7-7. Infinium assays

Bisulfite conversion was performed using the Zymo EZ DNA Methylation Kit (Zymo Research, Irvine, CA, USA) with 500 ng of genomic DNA for each sample. Whole genome amplification, labeling, hybridization, and scanning were performed according to the manufacturer's protocols. Genes were classified into the following four types based on methylation alterations within ±1 kb of a TSS. (i) "Unmethylated genes" contained >2 probes with $\beta < 0.2$ in wild-type cells and no probes with $\beta > 0.4$ after EBV infection. (ii) "Methylated genes" contained >2 probes with $\beta < 0.2$ in wild-type cells and >2 probes with β from <0.2 to >0.4 after EBV infection. Among "Methylated genes," (iii) "Methylation-sensitive genes" were those in which all probes showed $\beta > 0.2$ after EBV infection, and (iv)

“Methylation-resistant genes” were those in which >2 probes in a row showed β <0.2 even after EBV infection (Fig. 17).

In a bisulfite-based methylation assay, 5hmC and 5mC cannot be distinguished [66, 67]. Unmethylated cytosine changes to uracil, which is read as thymine by PCR, but neither 5hmC nor 5mC changes in response to bisulfite treatment, and both and are read as cytosine. The quantity of 5hmC, however, is much smaller than that of 5mC. In this study, the average β value of hydroxymethylated genes with only a 5hmC peak and no 5mC peak was only 0.07, whereas that of methylated genes with only a 5mC peak and no 5hmC peak was 0.54. The results for representative genes are shown in Fig. 25. Because the β value for 5hmC was sufficiently small, an Infinium assay was performed for 5mC detection.

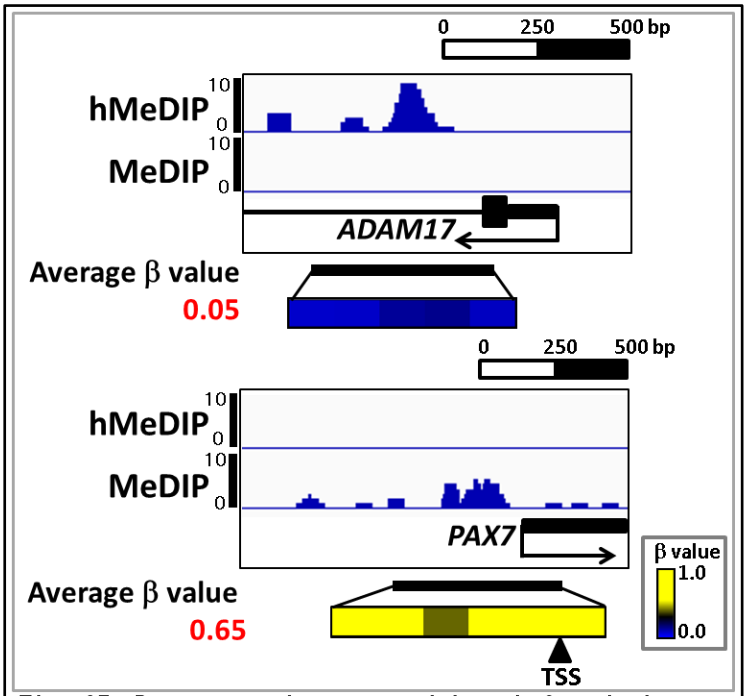

Fig. 25: Representative genes with an hmC peak, but no mC peak, e.g., *ADAM17*, and those with mC peak and no hmC peak, e.g., *PAX7*. The β values of the former genes were very low, e.g., 0.05 for *ADAM17*. The β values of the latter genes were markedly higher, e.g., 0.65 for *PAX7*.

7-8. Pyrosequencing analysis

Validation for methylated locus was carried out by pyrosequencing as described previously [42]. Primers for pyrosequencing were designed by Pyrosequencing Assay Design Software ver.2.0 (QIAGEN) to amplify bisulfite-treated DNA region containing several CpG sites. Primer sequences are listed in Table 5.

Table 5: Primers for pyrosequencing

Gene	strand	Primer types	Primer Sequence	Anneal (°C)	Product (bp)
FRG1B	bottom	Forward	GAGTGTTGTGGGAATGTGG	56	101
		Reverse*	TCCTCAAATACCCCCTCCAAAATTAAACT		
		Sequenceing	GGTGGAGAGTTATGGT		
HTRA1	bottom	Forward*	GGAGAGTGTAGGAGGGTTT	58	134
		Reverse	CCATCCCACCAACCCCCATC		
		Sequenceing	CACCCCTACCAAACCAATAAACT		

* Primers with 5′-biotin tag.

7-9. Analysis of miRNA

For the microRNA (miRNA) expression analysis, Direct-zol RNA MiniPrep (Zymo Research) was used to extract total RNA, including miRNAs. The miRNA Microarray System with miRNA Complete Labeling and Hyb Kit Version 2.4 (Agilent Technologies, Santa Clara, CA, USA) was used with 100 ng of total RNA as an input, following the manufacturer's instructions. After washing, scanning was conducted using a DNA Microarray Scanner (Agilent Technologies) and the resulting image data were converted to numerical form using Feature Extraction ver. 10.7.1.1 (Agilent Technologies). The numerical data were normalized using

GeneSpring GX 12.0 (Agilent Technologies). For validation, candidate miRNA in mature form was obtained (Bioneer, Daejeon, Republic of Korea) and transfected into MKN7 and GES1 with a final concentration of 20 nM using Lipofectamine 2000 (Invitrogen).

7-10. Statistical analyses

Statistical analyses of gene expression and methylation levels based on β values were performed using Wilcoxon signed-rank test. Gene counts were compared using the χ^2 test. R program (www.r-project. org/) was implemented in those testing. $P<0.05$ was considered to be statistically significant.

Supplementary Table S1: EBV miRNA and the targeting human genes

EBV miRNA		Sequence	Targeting human genes	
MIMAT0003720	ebv-miR-BART20-3p	CAUGAAGGCACAGCCUGUUACC	NM_024693	ECHDC3
MIMAT0003719	ebv-miR-BART20-5p	GUAGCAGGCAUGUCUUCAUUCC	NM_203474	PORCN
			NM_016592	GNAS
			NM_000850	GSTM4
			NM_058217	RAD51C
			NM_018059	FLJ10324
			NM_020956	PRX
			NM_147149	GSTM4
			XM_371176	LOC388550
			XM_371009	LOC388327
			NM_203318	MYO18A
			NM_078471	MYO18A
			NM_173557	RNF152
			NM_000561	GSTM1
			NM_203475	PORCN
			NM_203476	PORCN
			NM_022825	PORCN
			NM_203473	PORCN
			NM_001013617	LOC541469
			NM_014521	SH3BP4
			XM_498465	LOC439924
			NM_000848	GSTM2
			NM_019011	TRIAD3
			NM_207111	TRIAD3
			NM_207116	TRIAD3
			NM_146421	GSTM1
			NM_002442	MSI1
			NM_014476	PDLIM3
			NM_004959	NR5A1
			XM_499595	LOC442764
MIMAT0003716	ebv-miR-BART17-3p	UUGUAUGCCUGGUGUCCCCUUA	NM_018425	PI4KII
MIMAT0003715	ebv-miR-BART17-5p	AAGAGGACGCAGGCAUACAAGG	NM_014964	EPN2
			NM_148921	EPN2

			NM_017921	NPL4
MIMAT0003714	ebv-miR-BART16	AUAGAGUGGGUGUGUGCUCUUG	NM_001858	COL19A1
MIMAT0003713	ebv-miR-BART15	AGUGGUUUUGUUUCCUUGAUAG	NM_173455	PDE8A
			NM_198830	ACLY
			NM_001024211	S100A13
			XM_499148	LOC441443
			NM_000411	HLCS
			XM_499149	LOC441444
			NM_016577	RAB6B
			NM_198321	GALNT10
			NM_002605	PDE8A
			NM_018308	ACOXL
			NM_000860	HPGD
			NM_006610	MASP2
			XM_045421	C20orf194
			NM_005935	AFF1
			NM_020917	KIAA1559
			NM_017540	GALNT10
			NM_006650	CPLX2
			NM_001008220	CPLX2
			NM_005979	S100A13
			NM_004089	TSC22D3
			NM_001012960	C1orf84
			NM_013254	TBK1
			NM_001024210	S100A13
			NM_005736	ACTR1A
			NM_139021	MAPK15
			NM_052919	KIAA1920
			NM_173456	PDE8A
			NM_005639	SYT1
			NM_173457	PDE8A
			NM_031886	KCNA7
			NM_000136	FANCC
			NM_002193	INHBB
			NM_001005388	NFASC
			NM_015090	NFASC

			NM_001096	ACLY
			NM_001024213	S100A13
			NM_001015881	TSC22D3
			NM_001024212	S100A13
			NM_003799	RNMT
			NM_198057	TSC22D3
			NM_014775	SFI1
			NM_001007467	SFI1
			NM_138294	PATE
			NM_080473	GATA5
			NM_173454	PDE8A
			NM_001012993	C9orf152
			NM_001029955	LOC285429
			NM_002507	NGFR
MIMAT0003423	ebv-miR-BART12	UCCUGUGGUGUUUGGUGUGGUUU	NM_178225	FBXW5
			NM_178226	FBXW5
			NM_016734	PAX5
			NM_144599	NIPA1
			NM_031444	C22orf13
			NM_003423	ZNF43
			NM_030792	GDPD5
			NM_052909	KIAA1909
MIMAT0003422	ebv-miR-BART11-3p	ACGCACACCAGGCUGACUGCC		
MIMAT0003421	ebv-miR-BART11-5p	GACAGUUUGGUGCGCUAGUUGU	XM_370758	LOC387978
MIMAT0003415	ebv-miR-BART6-3p	CGGGGAUCGGACUAGCCUUAGA		
MIMAT0003414	ebv-miR-BART6-5p	GGUUGGUCCAAUCCAUAGGCUU	NM_018717	MAML3
			NM_021116	ADCY1
			NM_017617	NOTCH1
			NM_001025778	VRK3
			NM_016440	VRK3
			NM_001037553	AGPAT3
			NM_020132	AGPAT3
			NM_032831	C7orf19
			NM_012407	PRKCABP
			NM_001017995	KIAA1295
			NM_019061	MTMR12

MIMAT0003390	ebv-miR-BART1-3p	UAGCACCGCUAUCCACUAUGUCU	NM_001002836	LOC126208
MIMAT0000998	ebv-miR-BHRF1-3	UAACGGGAAGUGUGUAAGCACAC		
MIMAT0000995	ebv-miR-BHRF1-1	UAACCUGAUCAGCCCCGGAGUU		
MIMAT0003718	ebv-miR-BART19	UGUUUUGUUUGCUUGGGAAUGC	NM_022733	SMAP1L
			NM_006628	ARPP-19
			NM_001001930	PPARA
			NM_004014	DMD
			NM_004012	DMD
			NM_004007	DMD
			NM_032871	TNFRSF19L
			NM_014788	TRIM14
			NM_004010	DMD
			NM_006730	DNASE1L1
			NM_001009932	DNASE1L1
			NM_001009933	DNASE1L1
			NM_001037172	ACPL2
			NM_004011	DMD
			NM_004235	KLF4
			NM_004018	DMD
			NM_152282	ACPL2
			NM_024607	PPP1R3B
			NM_001009934	DNASE1L1
			NM_005036	PPARA
			NM_001001929	PPARA
			NM_020686	ABAT
			NM_000663	ABAT
			XM_379044	LOC284898
			NM_002885	RAP1GA1
			NM_001031717	CRELD1
			NM_001007560	C14orf70
			NM_001001928	PPARA
			NM_005114	HS3ST1
			NM_173551	SAMD6
			NM_004020	DMD
			NM_005587	MEF2A
			NM_004021	DMD

			NM_002396	ME2
			NM_006943	SOX12
			NM_000109	DMD
			NM_004009	DMD
			NM_033221	TRIM14
			NM_004006	DMD
			NM_004015	DMD
			NM_004023	DMD
			NM_015513	CRELD1
			NM_004013	DMD
			NM_152222	TNFRSF19L
			NM_004017	DMD
			NM_004022	DMD
			NM_004016	DMD
MIMAT0003717	ebv-miR-BART18	CAAGUUCGCACUUCCUAUACAG	NM_024989	PGAP1
MIMAT0003426	ebv-miR-BART14-3p	UAAAUGCUGCAGUAGUAGGGAU	NM_139279	MCFD2
			NM_199141	CARM1
MIMAT0003425	ebv-miR-BART14-5p	UACCCUACGCUGCCGAUUUACA	NM_002028	FNTB
			XM_376062	LOC400961
MIMAT0003424	ebv-miR-BART13	UGUAACUUGCCAGGGACGGCUGA		
MIMAT0003420	ebv-miR-BART10	ACAUAACCAUGGAGUUGGCUGU	NM_133625	SYN2
MIMAT0003419	ebv-miR-BART9	UAACACUUCAUGGGUCCCGUAG		
MIMAT0003418	ebv-miR-BART8-3p	GUCACAAUCUAUGGGGUCGUAG	NM_018655	LENEP
MIMAT0003417	ebv-miR-BART8-5p	UACGGUUUCCUAGAUUGUACAG	NM_001514	GTF2B
			NM_001001665	FLJ16008
MIMAT0003416	ebv-miR-BART7	CAUCAUAGUCCAGUGUCCAGGG	NM_000664	ACACA
			NM_198835	ACACA
			NM_013246	CLCF1
			NM_006033	LIPG
			NM_002111	HD
			NM_014732	KIAA0513
			NM_145015	MRGPRF
MIMAT0003413	ebv-miR-BART5	CAAGGUGAAUAUAGCUGCCCAUCG		
MIMAT0003412	ebv-miR-BART4	GACCUGAUGCUGCUGGUGUGCU	NM_014571	HEYL
			NM_001037553	AGPAT3
			NM_020132	AGPAT3

			NM_052847	GNG7
MIMAT0003411	ebv-miR-BART3-3p	CGCACCACUAGUCACCAGGUGU		
MIMAT0003410	ebv-miR-BART3-5p	AACCUAGUGUUAGUGUUGUGCU	NM_021943 NM_016605	ZFAND3 FAM53C
MIMAT0001000	ebv-miR-BART2	UAUUUUCUGCAUUCGCCCUUGC	NM_024671 NM_021973 NM_032385 NM_020452 NM_016348 NM_017586 NM_000809	FLJ23436 HAND2 C5orf4 ATP8B2 C5orf4 C9orf7 GABRA4
MIMAT0000999	ebv-miR-BART1-5p	UCUUAGUGGAAGUGACGUGCUGU		
MIMAT0000997	ebv-miR-BHRF1-2-3p	UAUCUUUUGCGGCAGAAAUUGAA		
MIMAT0000996	ebv-miR-BHRF1-2-5p	AAAUUCUGUUGCAGCAGAUAGC	NM_021998 NM_006137	ZNF6 CD7

8. Acknowledgements

I am deeply grateful to Professor Toshiki Watanabe at the Laboratory of Tumor Cell Biology, Department of Computational Biology and Medical Sciences, Graduate School of Frontier Sciences, the University of Tokyo, for providing me the opportunity to study and belong to the Department of Computational Biology and Medical Sciences.

I also greatly appreciate Professor Koichi Matsuda at the Laboratory of Clinical Genome Sequencing, Department of Computational Biology and Medical Sciences, Graduate School of Frontier Sciences, the University of Tokyo, for becoming my supervisor after Professor Toshiki Watanabe's retirement.

This study was carried out under the supervision by Professor Atsushi Kaneda at Department of Molecular Oncology, Graduate School of Medicine, Chiba University. I would like to express my gratitude to Professor Atsushi Kaneda who has guided me patiently through meaningful discussions from detailed experimental technique to the attitude as a researcher throughout this study.

I also would like to extend my gratitude to the researchers at the Department of Molecular Oncology, Graduate School of Medicine, Chiba University, especially to Dr. Keisuke Matsusaka and Dr. Sayaka Funata for critical advice and instructions on this study, to Dr. Masaki Fukuyo for data analysis and practical advice, to Dr. Kazuko Kita, Dr. Shigeru Sugaya, Dr. Atsushi Okabe, Dr. Ken-ichi Shinohara, Dr. Yasunobu Mano, Dr. Kiyoko Takane and Dr. Bahityar Rahmutulla for helpful advice and encouragement, and to the secretary Ms. Akiko Nagao for the administrative support.

I also greatly appreciate Professor Hiroyuki Aburatani at Genome Science Division, Research Center for Advanced Science and Technology, the University of Tokyo, who kindly provided me *TET2* construct used in this study.

I would like to express my gratitude to Dr. Genta Nagae, Dr. Kiyokazu Shirai, Dr. Asuka Kamio and Ms. Kaori Shina at Genome Science Division, Research Center for Advanced Science and Technology, the University of Tokyo, for instructions and helpful advice on hMeDIP-seq and MeDIP-seq.

Lastly, I am sincerely grateful to my family members who have always supported me and provided me helpful advice and encouragement.

9. References

1. Feinberg AP. Phenotypic plasticity and the epigenetics of human disease. Nature. 2007; 447: 433-440.
2. Bednar J, Horowitz RA, Grigoryev SA, Carruthers LM, Hansen JC, Koster AJ, Woodcock CL. Nucleosomes, linker DNA, and linker histone form a unique structural motif that directs the higher-order folding and compaction of chromatin. Proc Natl Acad Sci U S A. 1998; 95: 14173-14178.
3. Furrer P, Bednar J, Dubochet J, Hamiche A, Prunell A. DNA at the entry-exit of the nucleosome observed by cryoelectron microscopy. J Struct Biol. 1995; 114: 177-183.
4. Hamiche A, Schultz P, Ramakrishnan V, Oudet P, Prunell A. Linker histone-dependent DNA structure in linear mononucleosomes. J Mol Biol. 1996; 257: 30-42.
5. Leuba SH, Bustamante C, van Holde K, Zlatanova J. Linker histone tails and N-tails of histone H3 are redundant: scanning force microscopy studies of reconstituted fibers. Biophys J. 1998; 74: 2830-2839.
6. Leuba SH, Bustamante C, Zlatanova J, van Holde K. Contributions of linker histones and histone H3 to chromatin structure: scanning force microscopy studies on trypsinized fibers. Biophys J. 1998; 74: 2823-2829.
7. Ramakrishnan V. Histone H1 and chromatin higher-order structure. Crit Rev Eukaryot Gene Expr. 1997; 7: 215-230.
8. Toth K, Brun N, Langowski J. Trajectory of nucleosomal linker DNA studied by fluorescence resonance energy transfer. Biochemistry. 2001; 40: 6921-6928.

9. Travers A. The location of the linker histone on the nucleosome. Trends Biochem Sci. 1999; 24: 4-7.

10. Zlatanova J, Leuba SH, van Holde K. Chromatin fiber structure: morphology, molecular determinants, structural transitions. Biophys J. 1998; 74: 2554-2566.

11. Kouzarides T. Chromatin modifications and their function. Cell. 2007; 128: 693-705.

12. Jair KW, Bachman KE, Suzuki H, Ting AH, Rhee I, Yen RW, Baylin SB, Schuebel KE. *De novo* CpG island methylation in human cancer cells. Cancer Res. 2006; 66: 682-692.

13. Denis H, Ndlovu MN, Fuks F. Regulation of mammalian DNA methyltransferases: a route to new mechanisms. EMBO Rep. 2011; 12: 647-656.

14. Takai D, Jones PA. Comprehensive analysis of CpG islands in human chromosomes 21 and 22. Proc Natl Acad Sci U S A. 2002; 99: 3740-3745.

15. Bird A. DNA methylation patterns and epigenetic memory. Genes Dev. 2002; 16: 6-21.

16. Ito S, Shen L, Dai Q, Wu SC, Collins LB, Swenberg JA, He C, Zhang Y. Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. Science. 2011; 333: 1300-1303.

17. He YF, Li BZ, Li Z, Liu P, Wang Y, Tang Q, Ding J, Jia Y, Chen Z, Li L, Sun Y, Li X, Dai Q, et al. Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. Science. 2011; 333: 1303-1307.

18. Maiti A, Drohat AC. Thymine DNA glycosylase can rapidly excise

5-formylcytosine and 5-carboxylcytosine: potential implications for active demethylation of CpG sites. J Biol Chem. 2011; 286: 35334-35338.

19. Kellinger MW, Song CX, Chong J, Lu XY, He C, Wang D. 5-formylcytosine and 5-carboxylcytosine reduce the rate and substrate specificity of RNA polymerase II transcription. Nat Struct Mol Biol. 2012; 19: 831-833.

20. Tahiliani M, Koh KP, Shen Y, Pastor WA, Bandukwala H, Brudno Y, Agarwal S, Iyer LM, Liu DR, Aravind L, Rao A. Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. Science. 2009; 324: 930-935.

21. Esteller M. Epigenetics in cancer. N Engl J Med. 2008; 358: 1148-1159.

22. Jones PA, Baylin SB. The epigenomics of cancer. Cell. 2007; 128: 683-692.

23. Feinberg AP, Tycko B. The history of cancer epigenetics. Nat Rev Cancer. 2004; 4: 143-153.

24. Herman JG, Baylin SB. Gene silencing in cancer in association with promoter hypermethylation. N Engl J Med. 2003; 349: 2042-2054.

25. Ting AH, McGarvey KM, Baylin SB. The cancer epigenome--components and functional correlates. Genes Dev. 2006; 20: 3215-3231.

26. Toyota M, Ahuja N, Ohe-Toyota M, Herman JG, Baylin SB, Issa JP. CpG island methylator phenotype in colorectal cancer. Proc Natl Acad Sci U S A. 1999; 96: 8681-8686.

27. Toyota M, Ahuja N, Suzuki H, Itoh F, Ohe-Toyota M, Imai K, Baylin SB, Issa JP. Aberrant methylation in gastric cancer associated with the CpG island methylator phenotype. Cancer Res. 1999; 59: 5438-5442.

28. Turcan S, Rohle D, Goenka A, Walsh LA, Fang F, Yilmaz E, Campos C,

Fabius AW, Lu C, Ward PS, Thompson CB, Kaufman A, Guryanova O, et al. IDH1 mutation is sufficient to establish the glioma hypermethylator phenotype. Nature. 2012; 483: 479-483.

29. Yan H, Parsons DW, Jin G, McLendon R, Rasheed BA, Yuan W, Kos I, Batinic-Haberle I, Jones S, Riggins GJ, Friedman H, Friedman A, Reardon D, et al. IDH1 and IDH2 mutations in gliomas. N Engl J Med. 2009; 360: 765-773.

30. Noushmehr H, Weisenberger DJ, Diefes K, Phillips HS, Pujara K, Berman BP, Pan F, Pelloski CE, Sulman EP, Bhat KP, Verhaak RG, Hoadley KA, Hayes DN, et al. Identification of a CpG island methylator phenotype that defines a distinct subgroup of glioma. Cancer Cell. 2010; 17: 510-522.

31. Parsons DW, Jones S, Zhang X, Lin JC, Leary RJ, Angenendt P, Mankoo P, Carter H, Siu IM, Gallia GL, Olivi A, McLendon R, Rasheed BA, et al. An integrated genomic analysis of human glioblastoma multiforme. Science. 2008; 321: 1807-1812.

32. Cohen AL, Holmen SL, Colman H. IDH1 and IDH2 mutations in gliomas. Curr Neurol Neurosci Rep. 2013; 13: 345.

33. Kataoka K, Nagata Y, Kitanaka A, Shiraishi Y, Shimamura T, Yasunaga J, Totoki Y, Chiba K, Sato-Otsubo A, Nagae G, Ishii R, Muto S, Kotani S, et al. Integrated molecular analysis of adult T cell leukemia/lymphoma. Nat Genet. 2015; 47: 1304-1315.

34. Hughes LA, Melotte V, de Schrijver J, de Maat M, Smit VT, Bovee JV, French PJ, van den Brandt PA, Schouten LJ, de Meyer T, van Criekinge W, Ahuja N, Herman JG, et al. The CpG island methylator phenotype: what's

in a name? Cancer Res. 2013; 73: 5858-5868.

35. Ono R, Taki T, Taketani T, Taniwaki M, Kobayashi H, Hayashi Y. LCX, leukemia-associated protein with a CXXC domain, is fused to MLL in acute myeloid leukemia with trilineage dysplasia having t(10;11)(q22;q23). Cancer Res. 2002; 62: 4075-4080.

36. Lorsbach RB, Moore J, Mathew S, Raimondi SC, Mukatira ST, Downing JR. TET1, a member of a novel protein family, is fused to MLL in acute myeloid leukemia containing the t(10;11)(q22;q23). Leukemia. 2003; 17: 637-641.

37. Delhommeau F, Dupont S, Della Valle V, James C, Trannoy S, Masse A, Kosmider O, Le Couedic JP, Robert F, Alberdi A, Lecluse Y, Plo I, Dreyfus FJ, et al. Mutation in TET2 in myeloid cancers. N Engl J Med. 2009; 360: 2289-2301.

38. Rasmussen KD, Jia G, Johansen JV, Pedersen MT, Rapin N, Bagger FO, Porse BT, Bernard OA, Christensen J, Helin K. Loss of TET2 in hematopoietic cells leads to DNA hypermethylation of active enhancers and induction of leukemogenesis. Genes Dev. 2015; 29: 910-922.

39. Abdel-Wahab O, Mullally A, Hedvat C, Garcia-Manero G, Patel J, Wadleigh M, Malinge S, Yao J, Kilpivaara O, Bhat R, Huberman K, Thomas S, Dolgalev I, et al. Genetic characterization of TET1, TET2, and TET3 alterations in myeloid malignancies. Blood. 2009; 114: 144-147.

40. Langemeijer SM, Kuiper RP, Berends M, Knops R, Aslanyan MG, Massop M, Stevens-Linders E, van Hoogen P, van Kessel AG, Raymakers RA, Kamping EJ, Verhoef GE, Verburgh E, et al. Acquired mutations in TET2 are common in myelodysplastic syndromes. Nat Genet. 2009; 41: 838-842.

41. Yagi K, Akagi K, Hayashi H, Nagae G, Tsuji S, Isagawa T, Midorikawa Y, Nishimura Y, Sakamoto H, Seto Y, Aburatani H, Kaneda A. Three DNA methylation epigenotypes in human colorectal cancer. Clin Cancer Res. 2010; 16: 21-33.

42. Matsusaka K, Kaneda A, Nagae G, Ushiku T, Kikuchi Y, Hino R, Uozaki H, Seto Y, Takada K, Aburatani H, Fukayama M. Classification of Epstein-Barr virus-positive gastric cancers by definition of DNA methylation epigenotypes. Cancer Res. 2011; 71: 7187-7197.

43. Tokunaga M, Land CE, Uemura Y, Tokudome T, Tanaka S, Sato E. Epstein-Barr virus in gastric carcinoma. Am J Pathol. 1993; 143: 1250-1254.

44. Fukayama M, Hayashi Y, Iwasaki Y, Chong J, Ooba T, Takizawa T, Koike M, Mizutani S, Miyaki M, Hirai K. Epstein-Barr virus-associated gastric carcinoma and Epstein-Barr virus infection of the stomach. Lab Invest. 1994; 71: 73-81.

45. Shibata D, Weiss LM. Epstein-Barr virus-associated gastric adenocarcinoma. Am J Pathol. 1992; 140: 769-774.

46. Kang GH, Lee S, Cho NY, Gandamihardja T, Long TI, Weisenberger DJ, Campan M, Laird PW. DNA methylation profiles of gastric carcinoma characterized by quantitative DNA methylation analysis. Lab Invest. 2008; 88: 161-170.

47. Wang K, Yuen ST, Xu J, Lee SP, Yan HH, Shi ST, Siu HC, Deng S, Chu KM, Law S, Chan KH, Chan AS, Tsui WY, et al. Whole-genome sequencing and comprehensive molecular profiling identify new driver mutations in gastric

cancer. Nat Genet. 2014; 46: 573-582.

48. Comprehensive molecular characterization of gastric adenocarcinoma. Nature. 2014; 513: 202-209.

49. Fukayama M, Ushiku T. Epstein-Barr virus-associated gastric carcinoma. Pathol Res Pract. 2011; 207: 529-537.

50. Maeda E, Akahane M, Uozaki H, Kato N, Hayashi N, Fukayama M, Ohtomo K. CT appearance of Epstein-Barr virus-associated gastric carcinoma. Abdom Imaging. 2009; 34: 618-625.

51. Abe H, Maeda D, Hino R, Otake Y, Isogai M, Ushiku AS, Matsusaka K, Kunita A, Ushiku T, Uozaki H, Tateishi Y, Hishima T, Iwasaki Y, et al. ARID1A expression loss in gastric cancer: pathway-dependent roles with and without Epstein-Barr virus infection and microsatellite instability. Virchows Arch. 2012; 461: 367-377.

52. Gulley ML. Genomic assays for Epstein-Barr virus-positive gastric adenocarcinoma. Exp Mol Med. 2015; 47: e134.

53. Shinozaki A, Sakatani T, Ushiku T, Hino R, Isogai M, Ishikawa S, Uozaki H, Takada K, Fukayama M. Downregulation of microRNA-200 in EBV-associated gastric carcinoma. Cancer Res. 2010; 70: 4719-4727.

54. Kaneda A, Matsusaka K, Aburatani H, Fukayama M. Epstein-Barr virus infection as an epigenetic driver of tumorigenesis. Cancer Res. 2012; 72: 3445-3450.

55. Hino R, Uozaki H, Murakami N, Ushiku T, Shinozaki A, Ishikawa S, Morikawa T, Nakaya T, Sakatani T, Takada K, Fukayama M. Activation of DNA methyltransferase 1 by EBV latent membrane protein 2A leads to

promoter hypermethylation of PTEN gene in gastric carcinoma. Cancer Res. 2009; 69: 2766-2774.

56. Niedobitek G, Agathanggelou A, Steven N, Young LS. Epstein-Barr virus (EBV) in infectious mononucleosis: detection of the virus in tonsillar B lymphocytes but not in desquamated oropharyngeal epithelial cells. Mol Pathol. 2000; 53: 37-42.

57. Epstein MA, Barr YM. CULTIVATION IN VITRO OF HUMAN LYMPHOBLASTS FROM BURKITT'S MALIGNANT LYMPHOMA. Lancet. 1964; 1: 252-253.

58. Young LS, Rickinson AB. Epstein-Barr virus: 40 years on. Nat Rev Cancer. 2004; 4: 757-768.

59. Nalesnik MA, Jaffe R, Starzl TE, Demetris AJ, Porter K, Burnham JA, Makowka L, Ho M, Locker J. The pathology of posttransplant lymphoproliferative disorders occurring in the setting of cyclosporine A-prednisone immunosuppression. Am J Pathol. 1988; 133: 173-192.

60. Duval A, Raphael M, Brennetot C, Poirel H, Buhard O, Aubry A, Martin A, Krimi A, Leblond V, Gabarre J, Davi F, Charlotte F, Berger F, et al. The mutator pathway is a feature of immunodeficiency-related lymphomas. Proc Natl Acad Sci U S A. 2004; 101: 5002-5007.

61. zur Hausen H, Schulte-Holthausen H, Klein G, Henle W, Henle G, Clifford P, Santesson L. EBV DNA in biopsies of Burkitt tumours and anaplastic carcinomas of the nasopharynx. Nature. 1970; 228: 1056-1058.

62. Reisinger J, Rumpler S, Lion T, Ambros PF. Visualization of episomal and integrated Epstein-Barr virus DNA by fiber fluorescence in situ

hybridization. Int J Cancer. 2006; 118: 1603-1608.

63. Lu F, Wikramasinghe P, Norseen J, Tsai K, Wang P, Showe L, Davuluri RV, Lieberman PM. Genome-wide analysis of host-chromosome binding sites for Epstein-Barr Virus Nuclear Antigen 1 (EBNA1). Virol J. 2010; 7: 262.

64. Allday MJ, Kundu D, Finerty S, Griffin BE. CpG methylation of viral DNA in EBV-associated tumours. Int J Cancer. 1990; 45: 1125-1130.

65. Hayatsu H. Discovery of bisulfite-mediated cytosine conversion to uracil, the key reaction for DNA methylation analysis--a personal account. Proc Jpn Acad Ser B Phys Biol Sci. 2008; 84: 321-330.

66. Huang Y, Pastor WA, Shen Y, Tahiliani M, Liu DR, Rao A. The behaviour of 5-hydroxymethylcytosine in bisulfite sequencing. PLoS One. 2010; 5: e8888.

67. Jin SG, Kadam S, Pfeifer GP. Examination of the specificity of DNA methylation profiling techniques towards 5-methylcytosine and 5-hydroxymethylcytosine. Nucleic Acids Res. 2010; 38: e125.

68. Dedeurwaerder S, Defrance M, Calonne E, Denis H, Sotiriou C, Fuks F. Evaluation of the Infinium Methylation 450K technology. Epigenomics. 2011; 3: 771-784.

69. Colella S, Shen L, Baggerly KA, Issa JP, Krahe R. Sensitive and quantitative universal Pyrosequencing methylation analysis of CpG sites. Biotechniques. 2003; 35: 146-150.

70. Imai S, Nishikawa J, Takada K. Cell-to-cell contact as an efficient mode of Epstein-Barr virus infection of diverse human epithelial cells. J Virol. 1998; 72: 4371-4378.

71. Takada K. Cross-linking of cell surface immunoglobulins induces

Epstein-Barr virus in Burkitt lymphoma lines. Int J Cancer. 1984; 33: 27-32.

72. Esquela-Kerscher A, Slack FJ. Oncomirs - microRNAs with a role in cancer. Nat Rev Cancer. 2006; 6: 259-269.

73. Betel D, Wilson M, Gabow A, Marks DS, Sander C. The microRNA.org resource: targets and expression. Nucleic Acids Res. 2008; 36: D149-153.

74. Li SC, Pan CY, Lin WC. Bioinformatic discovery of microRNA precursors from human ESTs and introns. BMC Genomics. 2006; 7: 164.

75. Li SC, Shiau CK, Lin WC. Vir-Mir db: prediction of viral microRNA candidate hairpins. Nucleic Acids Res. 2008; 36: D184-189.

76. Williams K, Christensen J, Pedersen MT, Johansen JV, Cloos PA, Rappsilber J, Helin K. TET1 and hydroxymethylcytosine in transcription and DNA methylation fidelity. Nature. 2011; 473: 343-348.

77. Deplus R, Delatte B, Schwinn MK, Defrance M, Mendez J, Murphy N, Dawson MA, Volkmar M, Putmans P, Calonne E, Shih AH, Levine RL, Bernard O, et al. TET2 and TET3 regulate GlcNAcylation and H3K4 methylation through OGT and SET1/COMPASS. Embo j. 2013; 32: 645-655.

78. Uemura N, Okamoto S, Yamamoto S, Matsumura N, Yamaguchi S, Yamakido M, Taniyama K, Sasaki N, Schlemper RJ. Helicobacter pylori infection and the development of gastric cancer. N Engl J Med. 2001; 345: 784-789.

79. Ohnishi N, Yuasa H, Tanaka S, Sawa H, Miura M, Matsui A, Higashi H, Musashi M, Iwabuchi K, Suzuki M, Yamada G, Azuma T, Hatakeyama M. Transgenic expression of Helicobacter pylori CagA induces gastrointestinal and hematopoietic neoplasms in mouse. Proc Natl Acad Sci U S A. 2008;

105: 1003-1008.

80. Matsusaka K, Funata S, Fukayama M, Kaneda A. DNA methylation in gastric cancer, related to Helicobacter pylori and Epstein-Barr virus. World J Gastroenterol. 2014; 20: 3916-3926.

81. El-Omar EM, Carrington M, Chow WH, McColl KE, Bream JH, Young HA, Herrera J, Lissowska J, Yuan CC, Rothman N, Lanyon G, Martin M, Fraumeni JF, Jr., et al. Interleukin-1 polymorphisms associated with increased risk of gastric cancer. Nature. 2000; 404: 398-402.

82. Chan AO, Chu KM, Huang C, Lam KF, Leung SY, Sun YW, Ko S, Xia HH, Cho CH, Hui WM, Lam SK, Rashid A. Association between Helicobacter pylori infection and interleukin 1beta polymorphism predispose to CpG island methylation in gastric cancer. Gut. 2007; 56: 595-597.

83. Ke Y, Ning T, Wang B. [Establishment and characterization of a SV40 transformed human fetal gastric epithelial cell line-GES-1]. Zhonghua Zhong Liu Za Zhi. 1994; 16: 7-10.

84. Kaneda A, Fujita T, Anai M, Yamamoto S, Nagae G, Morikawa M, Tsuji S, Oshima M, Miyazono K, Aburatani H. Activation of Bmp2-Smad1 signal and its regulation by coordinated alteration of H3K27 trimethylation in Ras-induced senescence. PLoS Genet. 2011; 7: e1002359.

85. Kaneda A, Nonaka A, Fujita T, Yamanaka R, Fujimoto M, Miyazono K, Aburatani H. Epigenomic Regulation of Smad1 Signaling During Cellular Senescence Induced by Ras Activation. Methods Mol Biol. 2016; 1344: 341-353.

86. Weber M, Hellmann I, Stadler MB, Ramos L, Paabo S, Rebhan M, Schubeler

D. Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. Nat Genet. 2007; 39: 457-466.