ARID1A loss-of-function induces CpG island methylator phenotype

（ARID1A 機能異常が CpG アイランドメチル化形質を誘発する）

山田　晴美

Cancer Letters 532 (2022) 215587

Contents lists available at ScienceDirect

Cancer Letters

journal homepage: www.elsevier.com/locate/canlet

Original Articles

ARID1A loss-of-function induces CpG island methylator phenotype

Harumi Yamada [a,b], Hideyuki Takeshima [a,**], Ryoji Fujiki [c], Satoshi Yamashita [a], Shigeki Sekine [d], Takayuki Ando [e], Naoko Hattori [a], Atsushi Okabe [c], Takaki Yoshikawa [f], Kazutaka Obama [b], Hitoshi Katai [f], Atsushi Kaneda [c], Toshikazu Ushijima [a,*]

[a] Division of Epigenomics, National Cancer Center Research Institute, 5-1-1 Tsukiji, Chuo-ku, 104-0045, Tokyo, Japan
[b] Department of Gastrointestinal Surgery, Faculty of Medicine, Kyoto University Graduate School of Medicine, 54 Shogoin-kawahara-cho, Sakyo-ku, 606-8507, Kyoto, Japan
[c] Department of Molecular Oncology, Graduate School of Medicine, Chiba University, 1-8-1, Inohana, Chuo-ku, Chiba-shi, 260-8670, Chiba, Japan
[d] Department of Diagnostic Pathology, National Cancer Center Hospital, 5-1-1 Tsukiji, Chuo-ku, 104-0045, Tokyo, Japan
[e] Third Department of Internal Medicine, University of Toyama, 2630 Sugitani, 930-0194, Toyama, Japan
[f] Gastric Surgery Division, National Cancer Center Hospital, 5-1-1 Tsukiji, Chuo-ku, 104-0045, Tokyo, Japan

ARTICLE INFO

Keywords:
Epigenetics
Chromatin remodeling
ARID1A
DNA methylation
CIMP

ABSTRACT

The CpG island methylator phenotype (CIMP) is associated with prognosis and drug sensitivity in multiple cancer types. In gastric cancer, the CIMP is closely associated with Epstein-Barr virus (EBV) infection and AT-rich interactive domain 1A (*ARID1A*) mutations, a component of the SWI/SNF chromatin remodeling complex. However, the involvement of SWI/SNF defects in CIMP induction has been unclear. In this study, we demonstrate a causal role of ARID1A loss-of-function in CIMP induction. Mutations of SWI/SNF components, especially *ARID1A*, was associated with the CIMP, as well as EBV infection, in gastric cancers, and also in uterine endometrial and colorectal cancers, which are not affected by EBV infection. Genome-wide DNA methylation analysis showed that *ARID1A* knockout (KO) in cultured 293FT cells and gastric epithelial cells, GES1, induced aberrant DNA methylation of a substantial number of CpG sites. DNA methylation was induced at genomic regions with high levels of pre-existing histone H3 lysine 27 trimethylation (H3K27me3) and those with acquired H3K27me3 by *ARID1A* KO. These results showed that the *ARID1A* mutation induced aberrant DNA methylation, and this is likely to be one of the potential mechanisms of CIMP induction.

1. Introduction

Epigenetic alterations, especially aberrant DNA methylation of a promoter CpG island, are known to strongly suppress transcription of its downstream gene [1–3], and are deeply involved in the development of various types of cancers [4–6] by silencing of a variety of tumor-suppressor genes. In some cancer cells, a large number of CpG islands are simultaneously methylated, and this phenotype is referred to as the CpG island methylator phenotype (CIMP) [7–12]. The CIMP was originally reported in colorectal cancers [10,12–14], and then in various types of cancers, such as gastric cancers [15–17], glioblastomas [18–20], hepatocellular carcinomas [21–23], lung cancers [24,25], uterine endometrial cancer [26,27], bladder cancer [28,29] and neuroblastomas [30–32]. The CIMP is known to be associated with clinicopathological characteristics, such as patient prognosis and susceptibility to treatment, depending upon cancer types.

Regarding the induction mechanism of the CIMP, in glioblastomas (G-CIMP), a role of *IDH1* mutation, which produces 2-hydroxyglutarate (2-HG) and suppresses TET activity, has been reported [19,20]. In colorectal cancers, age-related DNA methylation of WNT negative regulators and tumor-suppressor genes, such as *CDKN2A*, causes a stem-like state and defects of cellular differentiation, and produces cells capable of $BRAF^{V600E}$-driven transformation [11,12]. In gastric cancers, infection with Epstein-Barr virus (EBV) is associated with very strong CIMP

Abbreviations: CIMP, CpG island methylator phenotype; UCEC, uterine corpus endometrial carcinoma; COAD, colon adenocarcinoma; BLCA, bladder epithelial cancer; HSD, high standard deviation; qRT-PCR, quantitative reverse transcription-polymerase chain reaction; TSS, transcription start site; ChIP, chromatin immunoprecipitation.

* Corresponding author.
** Corresponding author.
E-mail addresses: hitakesh@ncc.go.jp (H. Takeshima), tushijim@ncc.go.jp (T. Ushijima).

https://doi.org/10.1016/j.canlet.2022.215587
Received 30 March 2021; Received in revised form 26 November 2021; Accepted 3 February 2022
Available online 5 February 2022

[33–37]. EBV infection down-regulates *TET2* expression by up-regulating *TET2*-targeting microRNAs [38], and reduces active histone modifications H3K4me3 and H3K27Ac, which lead to induction of aberrant DNA methylation of multiple genes [39]. In addition, EBV was recently shown to induce extensive chromatin topology changes, resulting in extensive expression reprogramming involving multiple enhancers [40]. Gastric cancers with EBV infection frequently have inactivation of *ARID1A*, a component of the SWI/SNF chromatin remodeling complex, by mutations and EBV-encoded microRNAs [37, 41]. However, it is still unclear whether the inactivation of SWI/SNF components, especially *ARID1A*, is involved in CIMP induction.

In this study, we analyze an association between SWI/SNF mutations, represented by *ARID1A* mutations, and the CIMP, and demonstrate a causal role of ARID1A defects in aberrant DNA methylation induction.

2. Materials and methods

2.1. Clinical samples

Fifty gastric cancer samples (n = 50) were obtained from our previous study [42], in which we selected the samples randomly from available surgical and biopsy specimens. EBV infection was evaluated by *in situ* hybridization targeting EBER (EBER-ISH) [43]. The study was approved by the Institutional Review Boards of the National Cancer Center, and all the samples were collected with informed consents.

2.2. Cell lines and culture condition

Human embryonic kidney cell line, 293FT, was purchased from Thermo Fisher Scientific (Waltham, MA), and cultured in Dulbecco's Modified Eagle Medium (DMEM) (Sigma-Aldrich, St. Louis, MO) containing 10% (v/v) fetal bovine serum (FBS) (MP Biomedicals, Santa Ana, CA) and penicillin-streptomycin (GIBCO BRL, Palo Alto, CA). Human normal gastric epithelial cell line, GES1, was obtained from the Beijing Institute for Cancer Research, and cultured in RPMI1640 (Wako, Tokyo, Japan) with 10% (v/v) FBS (GE Healthcare, Chicago, IL) and penicillin-streptomycin (Sigma-Aldrich). Both cell lines were cultured at 37 °C in a 5% CO_2 incubator.

2.3. Establishment of ARID1A-knockout cells

Normal cell lines, 293FT and GES1, were used for establishment of *ARID1A*-knockout cells, since, in cancer cell lines, CpG islands susceptible to methylation induction are already methylated and the number of remaining unmethylated CpG islands is limited [44].

ARID1A KO in 293FT was conducted at the National Cancer Center. Since methylation levels of individual genes are highly diverse among clones within a cell line, 293FT cells were cloned before knockout of the *ARID1A* gene. 293FT clone-2, which showed comparable growth with its parental cell population, was transfected with an Edit-R Cas9 Nuclease mRNA (GE Healthcare Dharmacon Inc, Lafayette, CO), Edit-R tracrRNA (GE Healthcare Dharmacon Inc), and Edit-R crRNA against *ARID1A* (CM-017263-01-0002) (GE Healthcare Dharmacon Inc) using DharmaFECT Duo (GE Healthcare Dharmacon Inc). Then, cells were cloned by limiting dilution, and the presence of knockout *ARID1A* alleles was analyzed by Sanger sequencing of the crRNA target sequence (Exon8). Clones without wild-type (WT) alleles were selected as *ARID1A* KO clones (KO-6, 16, 40, 50 and 78) (Supplementary Table 1), and those with only WT alleles were selected as *ARID1A* WT clones (WT-4, 9, 15, 20, 23, 27, 44, 46, 52 and 63). Depletion of ARID1A protein was confirmed by Western blot analysis.

ARID1A KO in GES1 was conducted at Chiba University by introducing a vector encoding both Cas9 and gRNA. gRNAs against *ARID1A* Exon9 (gRNA-1, 5′-CTACCCCAATATGAATCAAG-3′) and Exon11 (gRNA-2, 5′-CAGACACATAGAGGCGATAG-3′) were cloned into pSpCas9 (BB)-2A-Puro vector (pX459, plasmid #48139) (Addgene, Watertown, MA), which was a gift from Dr. Feng Zhang (Department of Biological Engineering, Massachusetts Institute of Technology). Each of the constructed vectors or empty vector was transfected into GES1 cells using lipofectamine3000 reagent (Thermo Fisher Scientific). The cells were treated with 2.0 μg/ml of puromycin for 48 h, and were cloned by single-cell sorting with FACSAria III (Becton, Dickinson and Company, Franklin Lakes, NJ). Depletion of ARID1A protein was confirmed by Western blot analysis.

2.4. Establishment of ARID1A-rescued cells

pRP expression vector encoding *ARID1A* (pRP-ARID1A) and that encoding enhanced green fluorescent protein (EGFP) (pRP-EGFP) were purchased from Vector Builder (Chicago, IL). pRP-ARID1A or pRP-EGFP was transfected into *ARID1A* KO clones (KO-6, 50 and 78) using lipofectamine3000 reagent. Cells stably overexpressing ARID1A or EGFP were selected with 0.6 μg/ml of puromycin.

2.5. Western blot analysis

Whole cell extract was resolved by SDS-PAGE and transferred to an Immobilon-P nylon membrane (Merck Millipore, Billerica, MA). Each membrane was treated with BlockAce (Dainippon Pharmaceutical Co., Ltd., Suita, Japan), and reacted with rabbit anti-ARID1A antibody (Bethyl Laboratories, Montgomery, TX, A301-041A, 1:2000 for 293FT cells; and Abcam, Cambridge, UK, ab182560, 1:500 for GES1 cells), anti-DNMT1 antibody (ABclonal, Woburn, MA, A16729, 1:1000), anti-DNMT3A antibody (ABclonal, A16834, 1:1000), anti-TET1 antibody (ABclonal, A1506, 1:1000), anti-TET2 antibody (ABclonal, A17306, 1:1000), anti-NUP98 (Cell Signaling Technology, Danvers, MA; 2598, 1:1000) and anti-Actin antibody (Thermo Fisher Scientific, MA5-11869, 1:500). Following four cycles of 5-min washes in PBS or TBS supplemented with 0.5% Tween 20 (PBST or TBST), blots were reacted with a secondary antibody. Chemiluminescence detection was performed using an ECL kit (Biological Industries, Kibbutz Beit Haemek, Israel).

2.6. DNA methylation microarray analysis and data

Genome-wide DNA methylation data of the 50 primary gastric cancers, obtained using an Infinium HumanMethylation450 BeadChip array, were utilized from our previous studies [42,45]. For accurate analysis, 41 gastric cancers with a cancer cell fraction of 40% or more were selected from the 50 gastric cancers using markers for gastric cancer cell fraction [46]. Genome-wide DNA methylation analysis of *ARID1A* WT and KO cells was conducted using an Infinium MethylationEPIC Kit (Illumina, San Diego, CA). Neighboring probes (CpG sites) were assembled into a genomic block, and a genomic block was annotated by their relative locations from a transcription start site (TSS) and against a CpG island. All the 862,927 probes were assembled into 551, 478 genomic blocks, and 538,523 blocks on autosomes were used for the methylation analysis. A DNA methylation level of a genomic block was calculated as a mean β value of all the probes within the genomic block as described [47]. The DNA methylation data were deposited in the Gene Expression Omnibus (GEO) database with the accession # GSE165239.

In addition, DNA methylation data of uterine corpus endometrial carcinomas (UCECs), colon adenocarcinomas (COADs) and bladder epithelial cancers (BLCAs) were obtained from the TCGA (The Cancer Genome Atlas) database. For accurate analysis, cancer samples with a cancer cell fraction of 40% or more were selected. The cancer cell fraction was assessed by a histogram of β values of probes unmethylated in a corresponding normal tissue [48].

2.7. Mutation data

Somatic mutation data of 16 genes encoding SWI/SNF components (*ARID1A*, *ARID1B*, *ARID2*, *PBRM1*, *PHF10*, *SMARCA2*, *SMARCA4*, *SMARCAL1*, *SMARCB1*, *SMARCC1*, *SMARCC2*, *SMARCD1*, *SMARCD3*, *SMARCE1*, *ACTL6A* and *ACTL6B*) were obtained from our previous study using an Ion Personal Genome Machine (PGM) or an Ion Proton Sequencer (Life Technologies, Carlsbad, CA) [49]. Somatic mutation data of UCECs, COADs and BLCAs were obtained from the TCGA database.

2.8. Quantitative reverse transcription-polymerase chain reaction (qRT-PCR)

mRNA expression was analyzed by qRT-PCR as described [50] using primers listed in Supplementary Table 2. The raw copy numbers of *DNMT*s and *TET*s were normalized to that of *GAPDH*.

2.9. Chromatin immunoprecipitation sequencing (ChIP-seq)

Cells were treated with 1% formalin to cross-link DNA and histone, and the reaction was stopped by treatment with 0.16 M glycine. Cross-linked cells were re-suspended with lysis buffer (50 mM Tris-HCl (pH 8.0), 10 mM EDTA, 1% (w/v) SDS) and sonicated to obtain fragmented chromatin. Thirty μg of the chromatin was immunoprecipitated using 2 μg of antibody against trimethylation of histone H3 lysine 27 (H3K27me3) (Cell Signaling Technology; C36B11). Twenty μl of the chromatin was used as input DNA. Using immunoprecipitated and input DNA, sequencing libraries were prepared using a GeneNext NGS Library Prep Kit (TOYOBO, Osaka, Japan), and were sequenced using an Illumina HiSeqX_Ten sequencer in 150 bp pair-end mode. A sequencing depth of 44.1–56.7 million reads per sample was obtained.

ChIP-seq data were aligned to the human reference genome (hg19 version) using bowtie2 (v2.4.1) [51] with the parameters -D15 –R 2 –N 0 – L 22 –I s, 1, 1.15 -x hg19. Peaks were called using MACS2 (v2.1.0) with the broadpeak mode and with the parameters -g hs -p 1e-6 -f BAMPE -B [52]. Fragment pileup of individual samples was normalized to reads per million mapped reads (rpm), and an average read was obtained for individual genomic blocks. ChIP-seq data of three WT (WT-4, 20 and 23) and KO (KO-6, 50 and 78) clones were averaged, respectively, and was displayed using the Integrative Genomics Viewer (IGV) [53].

2.10. Cluster analysis

Unsupervised cluster analysis was conducted using R 4.0.3. with Heatplus package in Bioconductor.

2.11. Statistical analysis

Significant difference of the mutation frequency between CIMP-positive and -negative groups was evaluated by Fisher's test. An FDR *q*-value was obtained using R 4.0.3 with q value package. Significance of difference in the average β value between CIMP-positive and -negative groups was analyzed using the unpaired Student's t-test. $P < 0.05$ by two-sided test was considered significant. Significance of difference in Δβ value between the group with low H3K27me3 and that with high H3K27me3 was analyzed using the Mann-Whitney *U* test. $P < 0.05$ was considered significant.

3. Results

3.1. Gastric cancers were classified into two groups based on DNA methylation profiles

To evaluate any association between the CIMP status and the SWI/SNF mutations, DNA methylation and mutation profiles were analyzed in 41 primary gastric cancers with a cancer cell fraction of 40% or more. Unsupervised cluster analysis was conducted using 5000 genomic blocks in CpG islands with high standard deviation (HSD). The 41 gastric cancers were classified into two major branches, corresponding to CIMP-positive and -negative cancers (Fig. 1). This classification was in accordance with that in previous reports using classical CIMP markers, such as *MINTs*, *MLH1*, *CDK2A, FLNc*, *LOX*, *PAX6* and *THBD* [15,54]. The number of hypermethylated ($\Delta\beta > 0.1$) genomic regions was greater in the CIMP-positive cancers regarding CpG islands and non-CGIs (Shore, Shelf and Open sea) (Supplementary Fig. 1).

Mutations of SWI/SNF component genes were associated with the CIMP status [CIMP-positive (10/17) vs CIMP-negative (4/24); $p = 0.0079$]. *ARID1A* was most frequently mutated among the SWI/SNF components, and its mutation as a single gene was also associated with CIMP status [CIMP-positive (5/17) vs -negative (0/24); $p = 0.0083$]. Four of the five gastric cancers with *ARID1A* mutations were associated with EBV infection. As in many previous reports, *ARID1A* mutation and EBV infection highly overlapped, and it was difficult to conclude that *ARID1A* mutation itself was associated with the CIMP in gastric cancer.

3.2. ARID1A mutation was associated with CIMP in some cancers

To analyze the association between the mutations of SWI/SNF components, especially *ARID1A*, and the CIMP in cancers not affected by EBV infection, the association was analyzed in cancer types with *ARID1A* mutation of 15% or more; namely, uterine corpus endometrial carcinoma (UCEC), colon adenocarcinoma (COAD) and bladder epithelial cancer (BLCA). One hundred cases were randomly obtained from the TCGA database for each cancer type, and using samples with a cancer cell fraction of 40% or more, unsupervised cluster analysis using 5000 genomic blocks in CpG islands with HSD was performed. In UCEC, *ARID1A* mutation was strongly associated with the CIMP status [CIMP-positive (36/49) vs -negative (5/25); $p = 1.65 \times 10^{-5}$] (Fig. 2A). In COAD, the association was also observed, although weak [CIMP-positive (10/26) vs -negative (8/51); $p = 0.044$] (Fig. 2B). In BLCA, the association was not observed [CIMP-positive (6/12) vs -negative (16/71); $p = 0.073$] (Supplementary Fig. 2). Among the CIMP-positive cancers, 20% of gastric cancers, 58% of UCEC and 50% of COAD had variant allele frequencies of 50% (40–60%) in cancer cells estimated by cancer cell fraction, suggesting that *ARID1A* mutations had a haploinsufficient function in CIMP induction, as previously reported [55,56] (Supplementary Fig. 3). These results suggested the possibility that *ARID1A* mutation itself could cause CIMP depending upon cancer types.

3.3. ARID1A knockout induced aberrant DNA methylation

To analyze the causal role of ARID1A loss-of-function in the induction of aberrant DNA methylation, multiple clones with complete *ARID1A* KO, along with control (*ARID1A* WT) clones, were isolated first from 293FT-2 cells with high transfection efficiency (Fig. 3A and B). A volcano plot was drawn for three KO clones (KO-6, 50 and 78) and three WT clones (WT-27, 46 and 52) using 538,523 genomic blocks on autosomes. Hypermethylation was induced as early as '0′ weeks, approximately 4 weeks after introduction of gRNA and Cas9 nuclease ('0′ week), and increased at 4 and 20 weeks of culture (Fig. 3C). To visualize hypermethylated genomic blocks, we first selected 269,533 genomic blocks with small inter-clonal variation [$\Delta\beta$ value ≤ 0.05 among three WT clones (WT-4, 20, and 23)]. Of the 269,533 genomic blocks, 597 genomic blocks were hypermethylated in WT clones at '0′ week, showing the degree of inter-clonal variation. In contrast, 2,116, 3481 and 5816 genomic blocks were hypermethylated in KO clones at 0, 4 and 20 weeks of culture, respectively (Fig. 3D). Furthermore, we introduced exogenous *ARID1A* into three KO clones. Resultantly, the number of hypermethylated genomic blocks slightly decreased by *ARID1A* rescue (Supplementary Fig. 5). These results showed that aberrant DNA

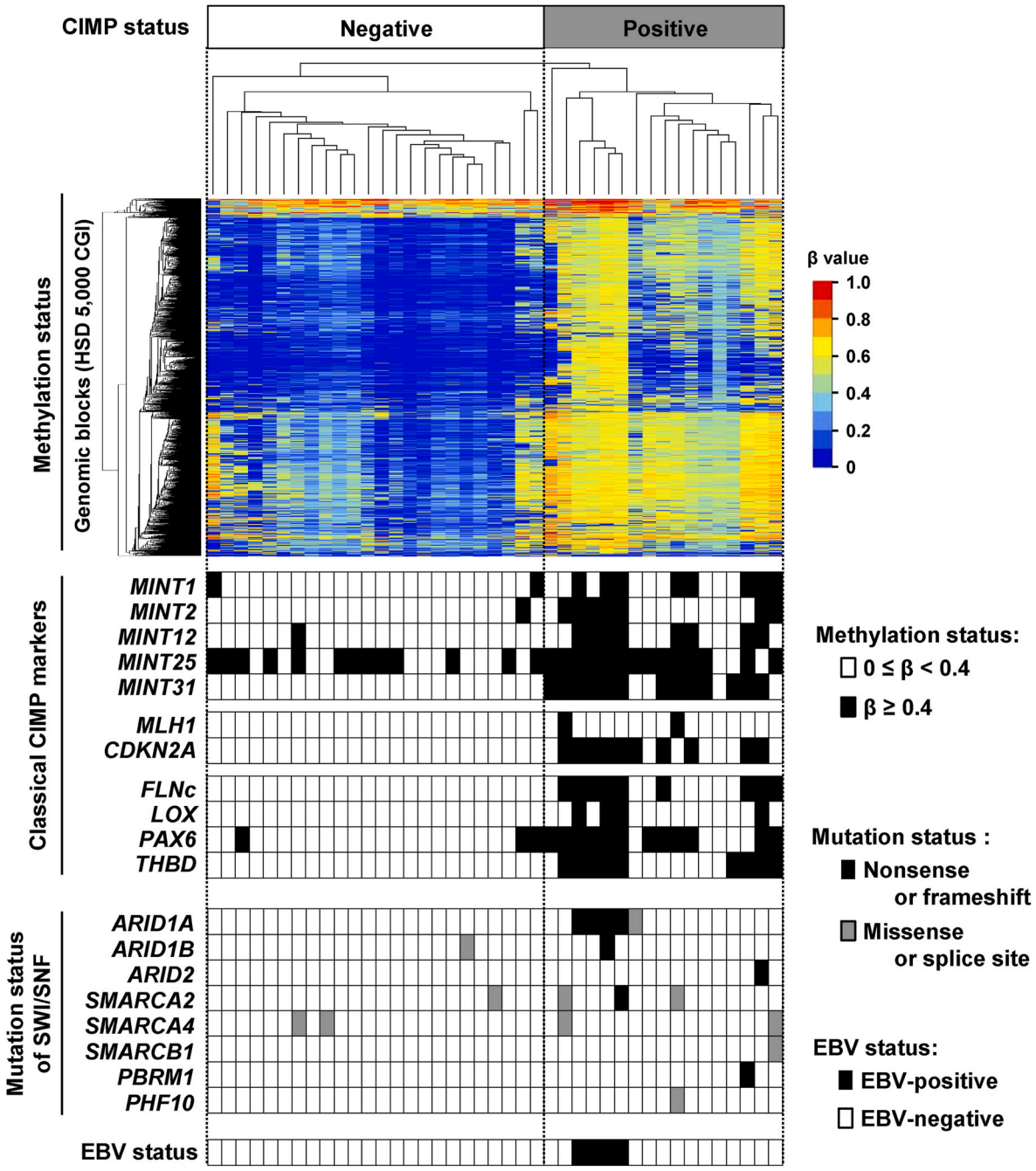


Fig. 1. Association between SWI/SNF component mutations and CIMP status in gastric cancer. Unsupervised cluster analysis was performed using 5000 genomic blocks in CpG islands with HSD. 41 gastric cancers with a cancer cell fraction of 40% or more were classified into two major branches, CIMP-positive and -negative cancers. This classification was in accordance with that in previous reports using classical CIMP markers, such as *MINTs*, *MLH1*, *CDK2A, FLNc*, *LOX*, *PAX6* and *THBD*. Mutations of SWI/SNF component genes were associated with the CIMP status ($p = 0.0079$), and *ARID1A* mutation as a single gene was also associated with the CIMP status ($p = 0.0083$). Four of the five gastric cancers with *ARID1A* mutations were associated with EBV infection.

methylation was induced by *ARID1A* loss-of-function.

To confirm aberrant DNA methylation induction by ARID1A loss-of-function in gastric epithelial cells, normal gastric epithelial cells, GES1, were additionally analyzed (Supplementary Figs. 6A and 6B). Among 349,106 genomic blocks with a Δβ value of 0.05 or less among two WT clones, 3936 genomic blocks were hypermethylated in WT clones at '0' week, showing the degree of inter-clonal variation. In contrast, 7148 and 7167 genomic blocks were hypermethylated in KO clones at 0 and 4 weeks of culture (Supplementary Fig. 6C). The number was 1.8-fold larger than the inter-clonal variation, and showed that hypermethylation had already been induced by *ARID1A* KO at '0' week. The lack of increase between 0 and 4 weeks of culture was considered due to saturation of methylation induction.

Characteristics of genomic regions hypermethylated by ARID1A loss-of-function were then analyzed. Compared with the distribution of all the genomic blocks analyzed, those frequently hypermethylated by *ARID1A* KO were located in gene body regions without CpG islands (Fig. 3E, Supplementary Fig. 4, Supplementary Fig. 6D). Although CpG islands were resistant to methylation induction, the number of hypermethylated genomic blocks in a CpG island increased depending upon the culture period (Fig. 3E). By limiting to CpG islands in gene promoter regions, which are known to silence their downstream genes

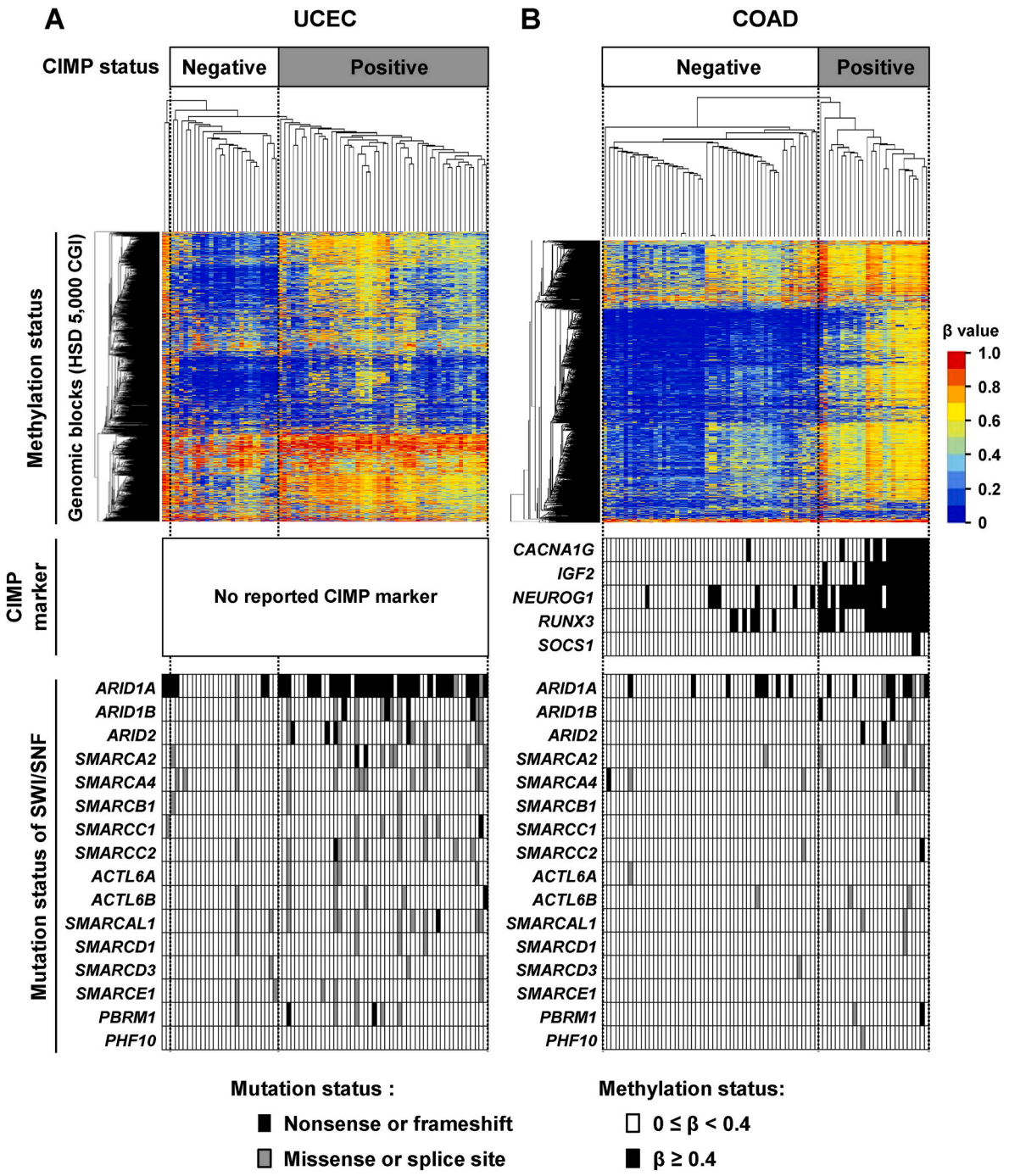


Fig. 2. Association between SWI/SNF component mutations and CIMP status in some cancers. The data of 100 randomly selected UCECs and COADs were obtained from the TCGA database, and cancers with a cancer cell fraction of 40% or more were selected. Unsupervised cluster analysis was performed using 5000 genomic blocks in CpG islands with HSD. (A) In UCEC, *ARID1A* mutation was strongly associated with the CIMP status ($p = 1.65 \times 10^{-5}$). Two samples from the left-side edge, which showed a large number of methylated genomic blocks, were clustered in a subgroup different from the CIMP-negative subgroup. (B) In COAD, *ARID1A* mutation was weakly but significantly associated with the CIMP status ($p = 0.044$). The classification based on genome-wide methylation analysis was in accordance with that in previous reports using classical CIMP markers, such as *CACNA1G*, *IGF2*, *NEUROG1*, *RUNX3* and *SOCS1*.

consistently [1,3], 32 genes located in TSS200 CpG island were hypermethylated by *ARID1A* KO. These genes contained a number of transcription factors (Supplementary Table 3) [57].

3.4. Expression levels of DNA methylation-related genes were not changed by ARID1A KO

To examine the molecular mechanisms of DNA methylation induction by *ARID1A* KO, expression changes of DNA methylation writers, *DNMT*s (*DNMT1*, *DNMT3A*, and *DNMT3B*), and erasers, *TET*s (*TET1*, *TET2*, and *TET3*), were analyzed at mRNA and protein levels. However, expression changes of the DNA methylation writers or erasers were within 1.2-fold by *ARID1A* KO (Supplementary Fig. 8), suggesting that their expression changes were not involved in DNA methylation induction by *ARID1A* KO.

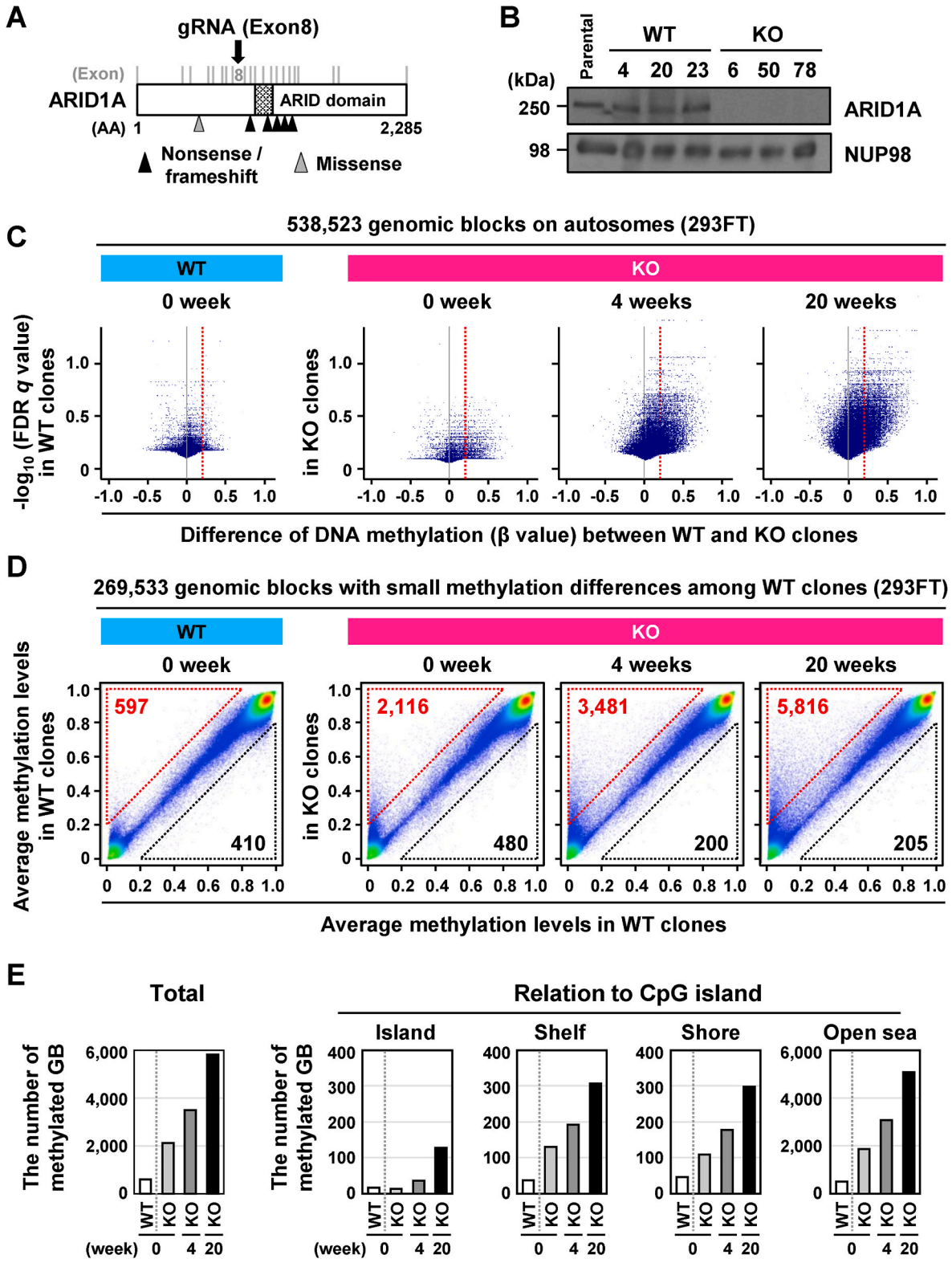


Fig. 3. ARID1A knockout induced aberrant hypermethylation in 293FT cells. (A) The location of *ARID1A* KO gRNA. The target sequence was located in exon 8. The locations of mutations detected in the 41 gastric cancers are shown by arrowheads (one cancer had two mutations). AA; amino acid. (B) ARID1A protein levels in KO clones. Complete depletion of ARID1A protein was confirmed in *ARID1A* KO clones (KO-6, 50 and 78). (C) Volcano plot analysis of methylation differences. The number of hypermethylated ($\Delta\beta > 0.2$) genomic regions with a high -log10 (FDR q value) increased dependent upon the culture period. (D) DNA methylation induction at genomic blocks not affected by variation among clones. Among the 269,533 genomic blocks with a $\Delta\beta$ value of 0.05 or less among three WT clones, 597 genomic blocks were hypermethylated in WT clones at the time point of KO clone isolation ('0' week; approximately 4 weeks after introduction of gRNA and Cas9 nuclease). In contrast, 2,116, 3481 and 5861 genomic blocks were hypermethylated in KO clones after 0, 4 and 20 weeks of culture, respectively (shown by the triangle with the dotted red line). (E) The number of genomic blocks hypermethylated by *ARID1A* KO against a CpG island. The number of hypermethylated genomic blocks increased in and outside of a CpG island, depending upon culture period.

3.5. Aberrant DNA methylation was induced at regions with H3K27me3 by ARID1A KO

To examine the characteristics of target genomic regions, the distribution of H3K27me3, a premark for DNA methylation induction [58–61] was analyzed by ChIP-seq (Fig. 4A). Based upon the distribution of H3K27me3 levels in a parental clone (pre-existing H3K27me3) (Supplementary Figs. 9A), 538,523 genomic blocks on autosomes were categorized into three groups; genomic blocks without pre-existing H3K27me3 (level = 0 rpm), those with low pre-existing H3K27me3 (≤0.1 rpm), and those with high (>0.1 rpm). Genomic blocks with high levels of pre-existing H3K27me3 were more likely to be hypermethylated in *ARID1A* KO clones after 20 weeks of culture [$p < 0.0001$] (Fig. 4B; a representative region in Fig. 4C).

Also, to analyze the function of acquired H3K27me3 by *ARID1A* KO, 12,527 genomic blocks without H3K27me3 (level = 0 rpm) in the parental clone and WT clones were selected. According to the increased H3K27me3 levels in the *ARID1A* KO clones (acquired H3K27me3) (Supplementary Fig. 9B), these regions were categorized into three groups; genomic blocks with no increase (increased level = 0 rpm), those with small increase (≤0.02 rpm), and those with large increase (>0.02 rpm). Genomic blocks with a large increase were more likely to be hypermethylated in *ARID1A* KO clones after 20 weeks of culture [$p < 0.0001$] (Fig. 4D; a representative region in Fig. 4E). These results showed that DNA methylation was induced at genomic blocks with high levels of pre-existing H3K27me3 and with high levels of H3K27me3 acquired by *ARID1A* KO.

4. Discussion

ARID1A mutation was associated with the CIMP, as well as EBV infection, in gastric cancers, and also in UCEC and COAD, which are not affected by EBV infection. However, the association was weak in COAD and was not observed in BLCA. Functionally, depletion of ARID1A in 293FT embryonic kidney cells and GES1 gastric epithelial cells induced aberrant DNA methylation. The causal role of ARID1A inactivation in methylation induction is supported by a previous report that showed the presence of *ARID1A* mutation in non-cancerous gastric mucosae with

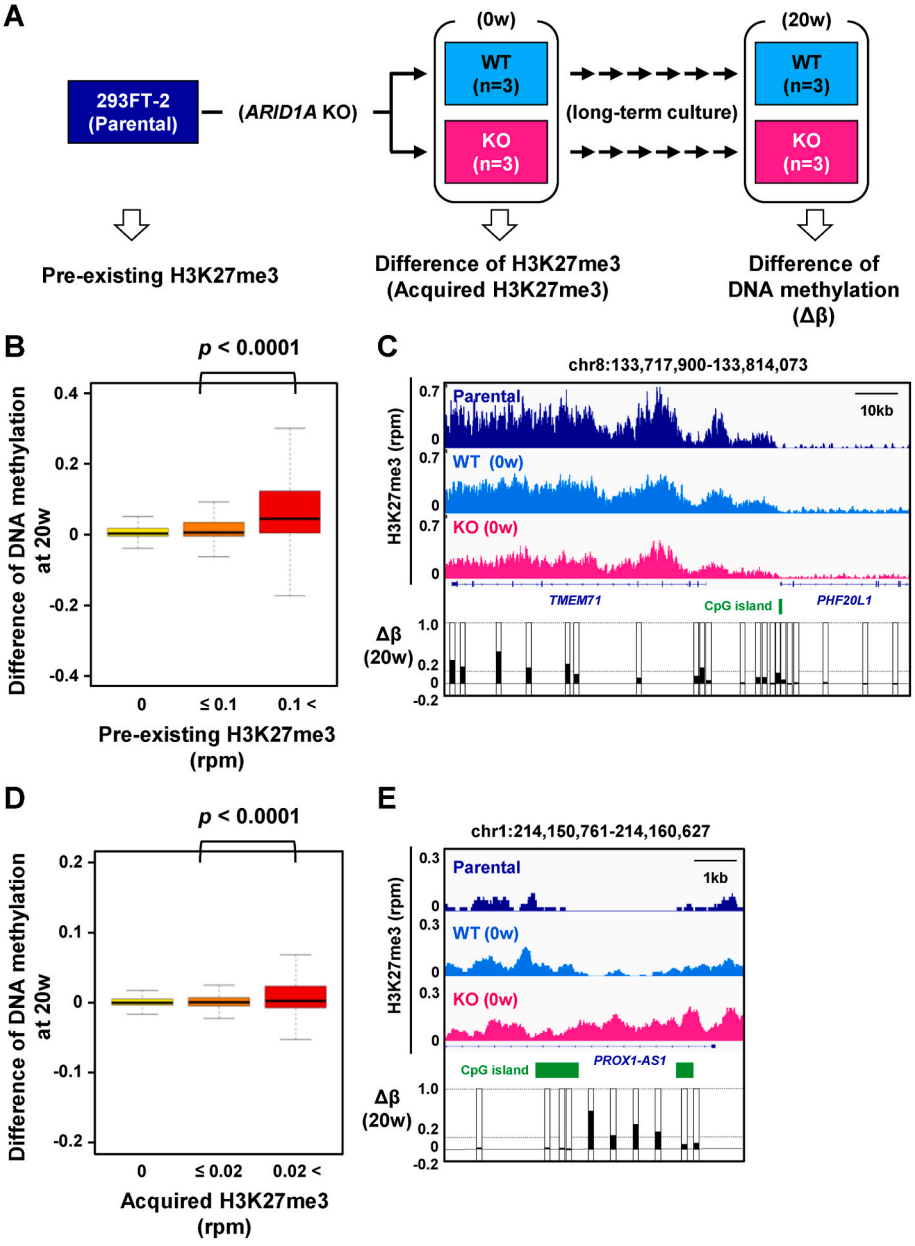


Fig. 4. The susceptibility of genomic regions with pre-existing and acquired H3K27me3 to aberrant DNA methylation induction by ARID1A KO. (A) Analytical strategy. Pre-existing H3K27me3 was obtained in parental cells (293FT-2) by ChIP-seq. Acquired H3K27me3 was assessed as the difference of average H3K27me3 levels in KO clones (KO-6, 50 and 78) compared to that in WT clones (WT-4, 20 and 23) when the clones were isolated ('0' week of culture). DNA methylation induction was further analyzed in the WT (WT-4, 20 and 23) and KO clones (KO-6, 50 and 78) at 20 weeks of culture. (B) The effect of pre-existing H3K27me3 on DNA methylation induction at 20 weeks. 538,523 genomic blocks on autosomes were categorized into three groups; genomic blocks without pre-existing H3K27me3 (level = 0 rpm, n = 189,010), those with low pre-existing H3K27me3 (≤0.1 rpm, n = 267,943), and those with high (>0.1 rpm, n = 48,880). Genomic blocks with high levels of pre-existing H3K27me3 were found to be more strongly hypermethylated in *ARID1A* KO clones ($p < 0.0001$). Box-and-whisker plots showed the median and the first, second, third and fourth quartile. (C) A representative genomic region with pre-existing H3K27me3 and DNA methylation induction. The gene body of *TMEM71* had high levels of H3K27me3 in parental cells, and DNA methylation was induced in the regions at 20 weeks, as shown by the increases in β values (positive Δβ values). The gene body of *PHF20L1* had low levels of H3K27me3 in both parental and *ARID1A* KO clones, and DNA methylation was not induced in the regions. (D) The effect of acquired H3K27me3 at '0' week on DNA methylation induction at 20 weeks. 12,527 genomic blocks without H3K27me3 in both parental and all WT clones were categorized into three groups; genomic blocks without acquired H3K27me3 (level = 0 rpm, n = 3565), those with a small increase (≤0.02 rpm, n = 6701), and those with a large increase (>0.02 rpm, n = 2261). Genomic blocks with high levels of acquired H3K27me3 were more strongly hypermethylated in *ARID1A* KO clones ($p < 0.0001$). Box-and-whisker plots show the median and the first, second, third and fourth quartile. (E) A representative genomic region with acquired H3K27me3 and DNA methylation induction. Intron 1 of *PROX1-AS1* acquired H3K27me3 by *ARID1A* KO, and was hypermethylated afterward.

Helicobacter pylori infection [62]. The study here is the first report to show the causal involvement of ARID1A loss-of-function in aberrant DNA methylation induction, at least in some types of cells.

At the same time, the association between *ARID1A* loss-of-function and the CIMP appeared to be dependent upon tissue types. In COAD, the association was weak, and the importance of pre-existing DNA methylation in normal colonic tissue to accommodate cells with a *BRAF* mutation has been reported [11], and this may have weakened the effect of *ARID1A* mutations. There is also a possibility that the mutator phenotype due to *MLH1* methylation in the CIMP-positive COAD could lead to an *ARID1A* mutation. In BLCA, frequent mutations of other epigenetic modifiers, KDM6A and KMT2D, have been reported [63], and they may have changed a phenotype due to *ARID1A* mutations. The redundant function of ARID1A and ARID1B dependent upon tissues [64] may have influenced the different degrees of the associations between the *ARID1A* mutation and the CIMP in different cancer types.

DNA methylation levels increased at genomic regions with high levels of H3K27me3, which is known as a premark of aberrant DNA methylation [58–61]. In addition, H3K27me3 levels themselves were increased at some of the genomic regions by ARID1A depletion, and aberrant DNA methylation was also induced in such genomic regions with acquired H3K27me3. The SWI/SNF complex is known to antagonize EZH2, which is an enzyme responsible for H3K27me3 [65,66]. The SWI/SNF loss-of-function may have led to increased EZH2 activity in the genomic regions with acquired H3K27me3, and H3K27me3, once induced, was considered to change the genomic susceptibility to aberrant DNA methylation induction as it does in aging and chronic inflammation [67–69].

Although CpG islands are the target genomic regions of the CIMP, gene body regions without CpG islands were more frequently hypermethylated by ARID1A depletion in cultured cells than CpG islands. Nevertheless, at least 128 CpG islands were hypermethylated by the ARID1A depletion. The majority of past studies on CIMP focused upon CpG islands, and information on methylation induction in genomic regions outside CpG islands was mostly missing. There remains a possibility that, although methylation of CpG islands is functionally important, methylation of other genomic regions is also induced in CIMP-positive cancers. Indeed, CIMP-positive gastric cancers had a higher methylation level than CIMP-negative gastric cancers in genomic blocks outside CpG islands (Supplementary Fig. 1). Since CpG islands are resistant to aberrant methylation induction [3,70], longer culture periods or additional molecular defects, such as TET inactivation [44], might induce more methylation of CpG islands.

Four-week culture of *ARID1A* KO clones of GES1 did not lead to a further increase of hypermethylated genomic blocks compared with methylation status at approximately 4 weeks after introduction of gRNA and Cas9 nuclease, and this was considered due to saturation of methylation induction. This was supported by the methylation profile of GES1 parental cells, which clustered with gastric cancer cell lines, not with normal gastric epithelial tissues (Supplementary Fig. 7). This showed that, in GES1, despite their recognition as 'normal' cells, most genomic regions susceptible to methylation induction were already methylated. *ARID1A* rescue in the *ARID1A* KO clones resulted in only a small suppression of the number of hypermethylated genomic blocks. This was considered to be because aberrant DNA methylation has already been induced in most susceptible regions during establishment of the KO clone, and rescue of *ARID1A* expression after the establishment resulted in suppression of DNA methylation induction only in the remaining small number of regions.

In conclusion, ARID1A was causally involved in aberrant DNA methylation induction in some types of cancers.

Author contributions

H.Y., H.T., K.O. and T.U. designed the research. H.Y., H.T., N.H., R. F., A.O. and A.K. carried out the experiments. S.Y. conducted the bioinformatics analysis. S.S., T.A., T.Y. and H.K. collected clinical samples. H.Y., H.T. and T.U. wrote the manuscript. All authors read and approved this manuscript.

Funding

This research was supported by AMED [grant Number 20ck0106552h0001 and 20gm1310006h0001] for T.U. and by JSPS KAKENHI [grant Number 19K08456] for H.Y.

Ethical approval

The study was approved by the Institutional Review Boards of the National Cancer Center.

Consent

All the samples were collected with informed consents.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors thank the National Cancer Center Biobank, Tokyo, Japan, for formalin-fixed paraffin-embedded (FFPE) sample collection; and to Ms. Mika Wakabayashi who offered excellent technical assistance. The results here are partly based upon data generated by the TCGA Research Network (https://www.cancer.gov/tcga).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.canlet.2022.215587.

References

[1] J.C. Lin, S. Jeong, G. Liang, D. Takai, M. Fatemi, Y.C. Tsai, G. Egger, E.N. Gal-Yam, P.A. Jones, Role of nucleosomal occupancy in the epigenetic silencing of the MLH1 CpG island, Cancer Cell 12 (2007) 432–444.

[2] D.L. Taylor, A.U. Jackson, N. Narisu, G. Hemani, M.R. Erdos, P.S. Chines, A. Swift, J. Idol, J.P. Didion, R.P. Welch, L. Kinnunen, J. Saramies, T.A. Lakka, M. Laakso, J. Tuomilehto, S.C.J. Parker, H.A. Koistinen, G. Davey Smith, M. Boehnke, L. J. Scott, E. Birney, F.S. Collins, Integrative analysis of gene expression, DNA methylation, physiological traits, and genetic variation in human skeletal muscle, Proc. Natl. Acad. Sci. U. S. A. 116 (2019) 10883–10888.

[3] M.V.C. Greenberg, D. Bourc'his, The diverse roles of DNA methylation in mammalian development and disease, Nat. Rev. Mol. Cell Biol. 20 (2019) 590–607.

[4] S.B. Baylin, P.A. Jones, A decade of exploring the cancer epigenome - biological and translational implications, Nat. Rev. Cancer 11 (2011) 726–734.

[5] S.E. Bates, Epigenetic therapies for cancer, N. Engl. J. Med. 383 (2020) 650–663.

[6] A.P. Feinberg, M.A. Koldobskiy, A. Gondor, Epigenetic modulators, modifiers and mediators in cancer aetiology and progression, Nat. Rev. Genet. 17 (2016) 284–299.

[7] J.P. Issa, CpG island methylator phenotype in cancer, Nat. Rev. Cancer 4 (2004) 988–993.

[8] H. Suzuki, E. Yamamoto, R. Maruyama, T. Niinuma, M. Kai, Biological significance of the CpG island methylator phenotype, Biochem. Biophys. Res. Commun. 455 (2014) 35–42.

[9] A.D. Kelly, H. Kroeger, J. Yamazaki, R. Taby, F. Neumann, S. Yu, J.T. Lee, B. Patel, Y. Li, R. He, S. Liang, Y. Lu, M. Cesaroni, S.A. Pierce, S.M. Kornblau, C.E. Bueso-Ramos, F. Ravandi, H.M. Kantarjian, J. Jelinek, J.P. Issa, A CpG island methylator phenotype in acute myeloid leukemia independent of IDH mutations and associated with a favorable outcome, Leukemia 31 (2017) 2011–2019.

[10] D.J. Weisenberger, G. Liang, H.J. Lenz, DNA methylation aberrancies delineate clinically distinct subsets of colorectal cancer and provide novel targets for epigenetic therapies, Oncogene 37 (2018) 566–577.

[11] Y. Tao, B. Kang, D.A. Petkovich, Y.R. Bhandari, J. In, G. Stein-O'Brien, X. Kong, W. Xie, N. Zachos, S. Maegawa, H. Vaidya, S. Brown, R.W. Chiu Yen, X. Shao, J. Thakor, Z. Lu, Y. Cai, Y. Zhang, I. Mallona, M.A. Peinado, C.A. Zahnow, N. Ahuja, E. Fertig, J.P. Issa, S.B. Baylin, H. Easwaran, Aging-like spontaneous

epigenetic silencing facilitates Wnt activation, stemness, and Braf(V600E)-induced tumorigenesis, Cancer Cell 35 (2019) 315–328, e316.
[12] T. Ushijima, H. Suzuki, The origin of CIMP, at last, Cancer Cell 35 (2019) 165–167.
[13] M. Toyota, N. Ahuja, M. Ohe-Toyota, J.G. Herman, S.B. Baylin, J.P. Issa, CpG island methylator phenotype in colorectal cancer, Proc. Natl. Acad. Sci. U. S. A. 96 (1999) 8681–8686.
[14] R. Jover, T.P. Nguyen, L. Pérez-Carbonell, P. Zapater, A. Payá, C. Alenda, E. Rojas, J. Cubiella, F. Balaguer, J.D. Morillas, J. Clofent, L. Bujanda, J.M. Reñé, X. Bessa, R. M. Xicola, D. Nicolás-Pérez, A. Castells, M. Andreu, X. Llor, C.R. Boland, A. Goel, 5-Fluorouracil adjuvant chemotherapy does not increase survival in patients with CpG island methylator phenotype colorectal cancer, Gastroenterology 140 (2011) 1174–1181.
[15] M. Toyota, N. Ahuja, H. Suzuki, F. Itoh, M. Ohe-Toyota, K. Imai, S.B. Baylin, J. P. Issa, Aberrant methylation in gastric cancer associated with the CpG island methylator phenotype, Cancer Res. 59 (1999) 5438–5442.
[16] S.Y. Park, M.C. Kook, Y.W. Kim, N.Y. Cho, N. Jung, H.J. Kwon, T.Y. Kim, G. H. Kang, CpG island hypermethylator phenotype in gastric carcinoma and its clinicopathological features, Virchows Arch. 457 (2010) 415–422.
[17] H. Zouridis, N. Deng, T. Ivanova, Y. Zhu, B. Wong, D. Huang, Y.H. Wu, Y. Wu, I. B. Tan, N. Liem, V. Gopalakrishnan, Q. Luo, J. Wu, M. Lee, W.P. Yong, L.K. Goh, B. T. Teh, S. Rozen, P. Tan, Methylation subtypes and large-scale epigenetic alterations in gastric cancer, Sci. Transl. Med. 4 (2012) 156ra140.
[18] H. Noushmehr, D.J. Weisenberger, K. Diefes, H.S. Phillips, K. Pujara, B.P. Berman, F. Pan, C.E. Pelloski, E.P. Sulman, K.P. Bhat, R.G. Verhaak, K.A. Hoadley, D. N. Hayes, C.M. Perou, H.K. Schmidt, L. Ding, R.K. Wilson, D. Van Den Berg, H. Shen, H. Bengtsson, P. Neuvial, L.M. Cope, J. Buckley, J.G. Herman, S.B. Baylin, P.W. Laird, K. Aldape, Identification of a CpG island methylator phenotype that defines a distinct subgroup of glioma, Cancer Cell 17 (2010) 510–522.
[19] S. Turcan, D. Rohle, A. Goenka, L.A. Walsh, F. Fang, E. Yilmaz, C. Campos, A. W. Fabius, C. Lu, P.S. Ward, C.B. Thompson, A. Kaufman, O. Guryanova, R. Levine, A. Heguy, A. Viale, L.G. Morris, J.T. Huse, I.K. Mellinghoff, T.A. Chan, IDH1 mutation is sufficient to establish the glioma hypermethylator phenotype, Nature 483 (2012) 479–483.
[20] M.J. Dabrowski, B. Wojtas, Global DNA methylation patterns in human gliomas and their interplay with other epigenetic modifications, Int. J. Mol. Sci. 20 (2019).
[21] L. Shen, N. Ahuja, Y. Shen, N.A. Habib, M. Toyota, A. Rashid, J.P. Issa, DNA methylation and environmental exposures in human hepatocellular carcinoma, J. Natl. Cancer Inst. 94 (2002) 755–761.
[22] J. Cheng, D. Wei, Y. Ji, L. Chen, L. Yang, G. Li, L. Wu, T. Hou, L. Xie, G. Ding, H. Li, Y. Li, Integrative analysis of DNA methylation and gene expression reveals hepatocellular carcinoma-specific diagnostic biomarkers, Genome Med. 10 (2018) 42.
[23] G. Li, W. Xu, L. Zhang, T. Liu, G. Jin, J. Song, J. Wu, Y. Wang, W. Chen, C. Zhang, X. Chen, Z. Ding, P. Zhu, B. Zhang, Development and validation of a CIMP-associated prognostic model for hepatocellular carcinoma, EBioMedicine 47 (2019) 128–141.
[24] K. Shinjo, Y. Okamoto, B. An, T. Yokoyama, I. Takeuchi, M. Fujii, H. Osada, N. Usami, Y. Hasegawa, H. Ito, T. Hida, N. Fujimoto, T. Kishimoto, Y. Sekido, Y. Kondo, Integrated analysis of genetic and epigenetic alterations reveals CpG island methylator phenotype associated with distinct clinical characters of lung adenocarcinoma, Carcinogenesis 33 (2012) 1277–1285.
[25] Y. Saito, G. Nagae, N. Motoi, E. Miyauchi, H. Ninomiya, H. Uehara, M. Mun, S. Okumura, F. Ohyanagi, M. Nishio, Y. Satoh, H. Aburatani, Y. Ishikawa, Prognostic significance of CpG island methylator phenotype in surgically resected small cell lung carcinoma, Cancer Sci. 107 (2016) 320–325.
[26] B.P. Whitcomb, D.G. Mutch, T.J. Herzog, J.S. Rader, R.K. Gibb, P.J. Goodfellow, Frequent HOXA11 and THBS2 promoter methylation, and a methylator phenotype in endometrial adenocarcinoma, Clin. Cancer Res. 9 (2003) 2277–2287.
[27] M. Yanokura, K. Banno, M. Adachi, D. Aoki, K. Abe, Genome-wide DNA methylation sequencing reveals miR-663a is a novel epimutation candidate in CIMP-high endometrial cancer, Int. J. Oncol. 50 (2017) 1934–1946.
[28] R. Maruyama, S. Toyooka, K.O. Toyooka, K. Harada, A.K. Virmani, S. Zöchbauer-Müller, A.J. Farinas, F. Vakar-Lopez, J.D. Minna, A. Sagalowsky, B. Czerniak, A. F. Gazdar, Aberrant promoter methylation profile of bladder cancer and its relationship to clinicopathological features, Cancer Res. 61 (2001) 8659–8663.
[29] I. Ibragimova, E. Dulaimi, M.J. Slifker, D.Y. Chen, R.G. Uzzo, P. Cairns, A global profile of gene promoter methylation in treatment-naïve urothelial cancer, Epigenetics 9 (2014) 760–773.
[30] M. Abe, M. Ohira, A. Kaneda, Y. Yagi, S. Yamamoto, Y. Kitano, T. Takato, A. Nakagawara, T. Ushijima, CpG island methylator phenotype is a strong determinant of poor prognosis in neuroblastomas, Cancer Res. 65 (2005) 828–834.
[31] K. Asada, M. Abe, T. Ushijima, Clinical application of the CpG island methylator phenotype to prognostic diagnosis in neuroblastomas, J. Hum. Genet. 58 (2013) 428–433.
[32] D. Capper, N.W. Engel, D. Stichel, M. Lechner, S. Glöss, S. Schmid, C. Koelsche, D. Schrimpf, J. Niesen, A.K. Wefers, D.T.W. Jones, M. Sill, O. Weigert, K.L. Ligon, A. Olar, A. Koch, M. Forster, S. Moran, O.M. Tirado, M. Sáinz-Jaspeado, J. Mora, M. Esteller, J. Alonso, X.G. Del Muro, W. Paulus, J. Felsberg, G. Reifenberger, M. Glatzel, S. Frank, C.M. Monoranu, V.J. Lund, A. von Deimling, S. Pfister, R. Buslei, J. Ribbat-Idel, S. Perner, V. Gudziol, M. Meinhardt, U. Schüller, DNA methylation-based reclassification of olfactory neuroblastoma, Acta Neuropathol. 136 (2018) 255–271.
[33] M. Tokunaga, C.E. Land, Y. Uemura, T. Tokudome, S. Tanaka, E. Sato, Epstein-Barr virus in gastric carcinoma, Am. J. Pathol. 143 (1993) 1250–1254.
[34] A.P. Burke, T.S. Yen, K.M. Shekitka, L.H. Sobin, Lymphoepithelial carcinoma of the stomach with Epstein-Barr virus demonstrated by polymerase chain reaction, Mod. Pathol. 3 (1990) 377–380.
[35] M. Fukayama, R. Hino, H. Uozaki, Epstein-Barr virus and gastric carcinoma: virus-host interactions leading to carcinoma, Cancer Sci. 99 (2008) 1726–1733.
[36] K. Matsusaka, A. Kaneda, G. Nagae, T. Ushiku, Y. Kikuchi, R. Hino, H. Uozaki, Y. Seto, K. Takada, H. Aburatani, M. Fukayama, Classification of Epstein-Barr virus-positive gastric cancers by definition of DNA methylation epigenotypes, Cancer Res. 71 (2011) 7187–7197.
[37] T.V. Bass Aj, I. Shmulevich, S.M. Reynolds, M. Miller, B. Bernard, T. Hinoue, P. W. Laird, C. Curtis, H. Shen, D.J. Weisenberger, N. Schultz, R. Shen, N. Weinhold, D.P. Kelsen, R. Bowlby, A. Chu, K. Kasaian, A.J. Mungall, A.G. Robertson, P. Sipahimalani, A.D. Cherniack, G. Getz, Y. Liu, M.S. Noble, C. Pedamallu, C. Sougnez, A. Taylor-Weiner, R. Akbani, J.S. Lee, W. Liu, G.B. Mills, D. Yang, W. Zhang, A. Pantazi, M. Parfenov, M. Gulley, M.B. Piazuelo, B.G. Schneider, J. Kim, A. Boussioutas, M. Sheth, J.A. Demchok, C.S. Rabkin, J.E. Willis, S. Ng, K. Garman, D.G. Beer, A. Pennathur, B.J. Raphael, H.T. Wu, R. Odze, H.K. Kim, J. Bowen, K.M. Leraas, T.M. Lichtenberg, S. Weaver, M. McLellan, M. Wiznerowicz, R. Sakai, G. Getz, C. Sougnez, M.S. Lawrence, K. Cibulskis, L. Lichtenstein, S. Fisher, S.B. Gabriel, E.S. Lander, L. Ding, B. Niu, A. Ally, M. Balasundaram, I. Birol, R. Bowlby, D. Brooks, Y.S. Butterfield, R. Carlsen, A. Chu, J. Chu, E. Chuah, H.J. Chun, A. Clarke, N. Dhalla, R. Guin, R.A. Holt, S.J. Jones, K. Kasaian, D. Lee, H.A. Li, E. Lim, Y. Ma, M.A. Marra, M. Mayo, R.A. Moore, A.J. Mungall, K. L. Mungall, K. Ming Nip, A.G. Robertson, J.E. Schein, P. Sipahimalani, A. Tam, N. Thiessen, R. Beroukhim, S.L. Carter, A.D. Cherniack, J. Cho, K. Cibulskis, D. DiCara, S. Frazer, S. Fisher, S.B. Gabriel, N. Gehlenborg, D.I. Heiman, J. Jung, J. Kim, E.S. Lander, M.S. Lawrence, L. Lichtenstein, P. Lin, M. Meyerson, A. I. Ojesina, C. Sekhar Pedamallu, G. Saksena, S.E. Schumacher, C. Sougnez, P. Stojanov, B. Tabak, A. Taylor-Weiner, D. Voet, M. Rosenberg, T.I. Zack, H. Zhang, L. Zou, A. Protopopov, N. Santoso, M. Parfenov, S. Lee, J. Zhang, H. S. Mahadeshwar, J. Tang, X. Ren, S. Seth, L. Yang, A.W. Xu, X. Song, A. Pantazi, R. Xi, C.A. Bristow, A. Hadjipanayis, J. Seidman, L. Chin, P.J. Park, R. Kucherlapati, R. Akbani, S. Ling, W. Liu, A. Rao, J.N. Weinstein, S.B. Kim, J. S. Lee, Y. Lu, G. Mills, P.W. Laird, T. Hinoue, D.J. Weisenberger, M.S. Bootwalla, P. H. Lai, H. Shen, T. Triche Jr., D.J. Van Den Berg, S.B. Baylin, J.G. Herman, G. Getz, L. Chin, Y. Liu, B.A. Murray, M.S. Noble, B.A. Askoy, G. Ciriello, G. Dresdner, J. Gao, B. Gross, A. Jacobsen, W. Lee, R. Ramirez, C. Sander, N. Schultz, Y. Senbabaoglu, R. Sinha, S.O. Sumer, Y. Sun, N. Weinhold, V. Thorsson, B. Bernard, L. Iype, R.W. Kramer, R. Kreisberg, M. Miller, S.M. Reynolds, H. Rovira, N. Tasman, I. Shmulevich, S. Ng, D. Haussler, J.M. Stuart, R. Akbani, S. Ling, W. Liu, A. Rao, J.N. Weinstein, R.G. Verhaak, G.B. Mills, M.D. Leiserson, B. J. Raphael, H.T. Wu, B.S. Taylor, A.D. Black, J. Bowen, J.A. Carney, J.M. Gastier-Foster, C. Helsel, K.M. Leraas, T.M. Lichtenberg, C. McAllister, N.C. Ramirez, T. R. Tabler, L. Wise, E. Zmuda, R. Penny, D. Crain, J. Gardner, K. Lau, E. Curely, D. Mallery, S. Morris, J. Paulauskis, T. Shelton, C. Shelton, M. Sherman, C. Benz, J. H. Lee, K. Fedosenko, G. Manikhas, O. Potapova, O. Voronina, D. Belyaev, O. Dolzhansky, W.K. Rathmell, J. Brzezinski, M. Ibbs, K. Korski, W. Kycler, R. Laźniak, E. Leporowska, A. Mackiewicz, D. Murawa, P. Murawa, A. Spychała, W. M. Suchorska, H. Tatka, M. Teresiak, M. Wiznerowicz, R. Abdel-Misih, J. Bennett, J. Brown, M. Iacocca, B. Rabeno, S.Y. Kwon, R. Penny, J. Gardner, A. Kemkes, D. Mallery, S. Morris, T. Shelton, C. Shelton, E. Curley, I. Alexopoulou, J. Engel, J. Bartlett, M. Albert, D.Y. Park, R. Dhir, J. Luketich, R. Landreneau, Y.Y. Janjigian, D.P. Kelsen, E. Cho, M. Ladanyi, L. Tang, S.J. McCall, Y.S. Park, J.H. Cheong, J. Ajani, M.C. Camargo, S. Alonso, B. Ayala, M.A. Jensen, T. Pihl, R. Raman, J. Walton, Y. Wan, J.A. Demchok, G. Eley, K.R. Mills Shaw, M. Sheth, R. Tarnuzzer, Z. Wang, L. Yang, J.C. Zenklusen, T. Davidsen, C.M. Hutter, H.J. Sofia, R. Burton, S. Chudamani, J. Liu, Comprehensive molecular characterization of gastric adenocarcinoma, Nature 513 (2014) 202–209.
[38] H. Namba-Fukuyo, S. Funata, K. Matsusaka, M. Fukuyo, B. Rahmutulla, Y. Mano, M. Fukayama, H. Aburatani, A. Kaneda, TET2 functions as a resistance factor against DNA methylation acquisition during Epstein-Barr virus infection, Oncotarget 7 (2016) 81512–81526.
[39] S. Funata, K. Matsusaka, R. Yamanaka, S. Yamamoto, A. Okabe, M. Fukuyo, H. Aburatani, M. Fukayama, A. Kaneda, Histone modification alteration coordinated with acquisition of promoter DNA methylation during Epstein-Barr virus infection, Oncotarget 8 (2017) 55265–55279.
[40] A. Okabe, K.K. Huang, K. Matsusaka, M. Fukuyo, M. Xing, X. Ong, T. Hoshii, G. Usui, M. Seki, Y. Mano, B. Rahmutulla, T. Kanda, T. Suzuki, S.Y. Rha, T. Ushiku, M. Fukayama, P. Tan, A. Kaneda, Cross-species chromatin interactions drive transcriptional rewiring in Epstein-Barr virus-positive gastric adenocarcinoma, Nat. Genet. 52 (2020) 919–930.
[41] K. Kase, M. Saito, S. Nakajima, D. Takayanagi, K. Saito, L. Yamada, M. Ashizawa, H. Nakano, H. Hanayama, H. Onozawa, H. Okayama, H. Endo, S. Fujita, W. Sakamoto, Z. Saze, T. Momma, K. Mimura, S. Ohki, K. Shiraishi, T. Kohno, K. Kono, ARID1A deficiency in EBV-positive gastric cancer is partially regulated by EBV-encoded miRNAs, but not by DNA promotor hypermethylation, Carcinogenesis 42 (2021) 21–30.
[42] Y. Yoda, H. Takeshima, T. Niwa, J.G. Kim, T. Ando, R. Kushima, T. Sugiyama, H. Katai, H. Noshiro, T. Ushijima, Integrated analysis of cancer-related pathways affected by genetic and epigenetic alterations in gastric cancer, Gastric Cancer 18 (2015) 65–76.
[43] R.F. Ambinder, R.B. Mann, Detection and characterization of Epstein-Barr virus in clinical specimens, Am. J. Pathol. 145 (1994) 239–252.
[44] H. Takeshima, T. Niwa, S. Yamashita, T. Takamura-Enya, N. Iida, M. Wakabayashi, S. Nanjo, M. Abe, T. Sugiyama, Y.J. Kim, T. Ushijima, TET repression and increased

DNMT activity synergistically induce aberrant DNA methylation, J. Clin. Invest. 130 (2020) 5370–5379.
[45] J.G. Kim, H. Takeshima, T. Niwa, E. Rehnberg, Y. Shigematsu, Y. Yoda, S. Yamashita, R. Kushima, T. Maekita, M. Ichinose, H. Katai, W.S. Park, Y.S. Hong, C.H. Park, T. Ushijima, Comprehensive DNA methylation and extensive mutation analyses reveal an association between the CpG island methylator phenotype and oncogenic mutations in gastric cancers, Cancer Lett. 330 (2013) 33–40.
[46] L. Zong, N. Hattori, Y. Yoda, S. Yamashita, H. Takeshima, T. Takahashi, M. Maeda, H. Katai, S. Nanjo, T. Ando, Y. Seto, T. Ushijima, Establishment of a DNA methylation marker to evaluate cancer cell fraction in gastric cancer, Gastric Cancer 19 (2016) 361–369.
[47] N. Iida, Y. Okuda, O. Ogasawara, S. Yamashita, H. Takeshima, T. Ushijima, MACON: a web tool for computing DNA methylation data obtained by the Illumina Infinium Human DNA methylation BeadArray, Epigenomics 10 (2018) 249–258.
[48] T. Nakamura, S. Yamashita, K. Fukumura, J. Nakabayashi, K. Tanaka, K. Tamura, K. Tateishi, M. Kinoshita, S. Fukushima, H. Takami, K. Fukuoka, K. Yamazaki, Y. Matsushita, M. Ohno, Y. Miyakita, S. Shibui, A. Kubo, T. Shuto, S. Kocialkowski, S. Yamanaka, A. Mukasa, T. Sasayama, K. Mishima, T. Maehara, N. Kawahara, M. Nagane, Y. Narita, H. Mano, T. Ushijima, K. Ichimura, Genome-wide DNA methylation profiling identifies primary central nervous system lymphoma as a distinct entity different from systemic diffuse large B-cell lymphoma, Acta Neuropathol. 133 (2017) 321–324.
[49] H. Takeshima, T. Niwa, T. Takahashi, M. Wakabayashi, S. Yamashita, T. Ando, Y. Inagawa, H. Taniguchi, H. Katai, T. Sugiyama, T. Kiyono, T. Ushijima, Frequent involvement of chromatin remodeler alterations in gastric field cancerization, Cancer Lett. 357 (2015) 328–338.
[50] H. Takeshima, S. Yamashita, T. Shimazu, T. Niwa, T. Ushijima, The presence of RNA polymerase II, active or stalled, predicts epigenetic fate of promoter CpG islands, Genome Res. 19 (2009) 1974–1982.
[51] B. Langmead, S.L. Salzberg, Fast gapped-read alignment with Bowtie 2, Nat. Methods 9 (2012) 357–359.
[52] Y. Zhang, T. Liu, C.A. Meyer, J. Eeckhoute, D.S. Johnson, B.E. Bernstein, C. Nusbaum, R.M. Myers, M. Brown, W. Li, X.S. Liu, Model-based analysis of ChIP-seq (MACS), Genome Biol. 9 (2008) R137.
[53] J.T. Robinson, H. Thorvaldsdóttir, W. Winckler, M. Guttman, E.S. Lander, G. Getz, J.P. Mesirov, Integrative genomics viewer, Nat. Biotechnol. 29 (2011) 24–26.
[54] S. Enomoto, T. Maekita, T. Tsukamoto, T. Nakajima, K. Nakazawa, M. Tatematsu, M. Ichinose, T. Ushijima, Lack of association between CpG island methylator phenotype in human gastric cancers and methylation in their background non-cancerous gastric mucosae, Cancer Sci. 98 (2007) 1853–1861.
[55] M.R. Wilson, J.J. Reske, J. Holladay, G.E. Wilber, M. Rhodes, J. Koeman, M. Adams, B. Johnson, R.W. Su, N.R. Joshi, A.L. Patterson, H. Shen, R.E. Leach, J. M. Teixeira, A.T. Fazleabas, R.L. Chandler, ARID1A and PI3-kinase pathway mutations in the endometrium drive epithelial transdifferentiation and collective invasion, Nat. Commun. 10 (2019) 3554.
[56] J.N. Wu, C.W. Roberts, ARID1A mutations in cancer: another epigenetic tumor suppressor? Cancer Discov. 3 (2013) 35–43.
[57] M. Safran, I. Dalah, J. Alexander, N. Rosen, T. Iny Stein, M. Shmoish, N. Nativ, I. Bahir, T. Doniger, H. Krug, A. Sirota-Madi, T. Olender, Y. Golan, G. Stelzer, A. Harel, D. Lancet, GeneCards Version 3: the Human Gene Integrator, Database, Oxford), 2010, p. baq020, 2010.
[58] J.E. Ohm, K.M. McGarvey, X. Yu, L. Cheng, K.E. Schuebel, L. Cope, H. P. Mohammad, W. Chen, V.C. Daniel, W. Yu, D.M. Berman, T. Jenuwein, K. Pruitt, S.J. Sharkis, D.N. Watkins, J.G. Herman, S.B. Baylin, A stem cell-like chromatin pattern may predispose tumor suppressor genes to DNA hypermethylation and heritable silencing, Nat. Genet. 39 (2007) 237–242.
[59] Y. Schlesinger, R. Straussman, I. Keshet, S. Farkash, M. Hecht, J. Zimmerman, E. Eden, Z. Yakhini, E. Ben-Shushan, B.E. Reubinoff, Y. Bergman, I. Simon, H. Cedar, Polycomb-mediated methylation on Lys27 of histone H3 pre-marks genes for de novo methylation in cancer, Nat. Genet. 39 (2007) 232–236.
[60] M. Widschwendter, H. Fiegl, D. Egle, E. Mueller-Holzner, G. Spizzo, C. Marth, D. J. Weisenberger, M. Campan, J. Young, I. Jacobs, P.W. Laird, Epigenetic stem cell signature in cancer, Nat. Genet. 39 (2007) 157–158.
[61] H. Takeshima, D. Ikegami, M. Wakabayashi, T. Niwa, Y.J. Kim, T. Ushijima, Induction of aberrant trimethylation of histone H3 lysine 27 by inflammation in mouse colonic epithelial cells, Carcinogenesis 33 (2012) 2384–2390.
[62] T. Shimizu, H. Marusawa, Y. Matsumoto, T. Inuzuka, A. Ikeda, Y. Fujii, S. Minamiguchi, S. Miyamoto, T. Kou, Y. Sakai, J.E. Crabtree, T. Chiba, Accumulation of somatic mutations in TP53 in gastric epithelium with Helicobacter pylori infection, Gastroenterology 147 (2014) 407–417, e403.
[63] L. Tran, J.F. Xiao, N. Agarwal, J.E. Duex, D. Theodorescu, Advances in bladder cancer biology and therapy, Nat. Rev. Cancer 21 (2021) 104–121.
[64] K.C. Helming, X. Wang, B.G. Wilson, F. Vazquez, J.R. Haswell, H.E. Manchester, Y. Kim, G.V. Kryukov, M. Ghandi, A.J. Aguirre, Z. Jagani, Z. Wang, L.A. Garraway, W.C. Hahn, C.W. Roberts, ARID1B is a specific vulnerability in ARID1A-mutant cancers, Nat. Med. 20 (2014) 251–254.
[65] B.G. Wilson, X. Wang, X. Shen, E.S. McKenna, M.E. Lemieux, Y.J. Cho, E. C. Koellhoffer, S.L. Pomeroy, S.H. Orkin, C.W. Roberts, Epigenetic antagonism between polycomb and SWI/SNF complexes during oncogenic transformation, Cancer Cell 18 (2010) 316–328.
[66] B.G. Bitler, K.M. Aird, A. Garipov, H. Li, M. Amatangelo, A.V. Kossenkov, D. C. Schultz, Q. Liu, M. Shih Ie, J.R. Conejo-Garcia, D.W. Speicher, R. Zhang, Synthetic lethality by targeting EZH2 methyltransferase activity in ARID1A-mutated cancers, Nat. Med. 21 (2015) 231–238.
[67] S. Yamashita, S. Nanjo, E. Rehnberg, N. Iida, H. Takeshima, T. Ando, T. Maekita, T. Sugiyama, T. Ushijima, Distinct DNA methylation targets by aging and chronic inflammation: a pilot study using gastric mucosa infected with Helicobacter pylori, Clin. Epigenet. 11 (2019) 191.
[68] M. Yu, W.D. Hazelton, G.E. Luebeck, W.M. Grady, Epigenetic aging: more than just a clock when it comes to cancer, Cancer Res. 80 (2020) 367–374.
[69] A.R. Maiuri, M. Peng, R. Podicheti, S. Sriramkumar, C.M. Kamplain, D.B. Rusch, C. E. DeStefano Shields, C.L. Sears, H.M. O'Hagan, Mismatch repair proteins initiate epigenetic alterations during inflammation-driven tumorigenesis, Cancer Res. 77 (2017) 3467–3478.
[70] M. Weber, I. Hellmann, M.B. Stadler, L. Ramos, S. Pääbo, M. Rebhan, D. Schübeler, Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome, Nat. Genet. 39 (2007) 457–466.

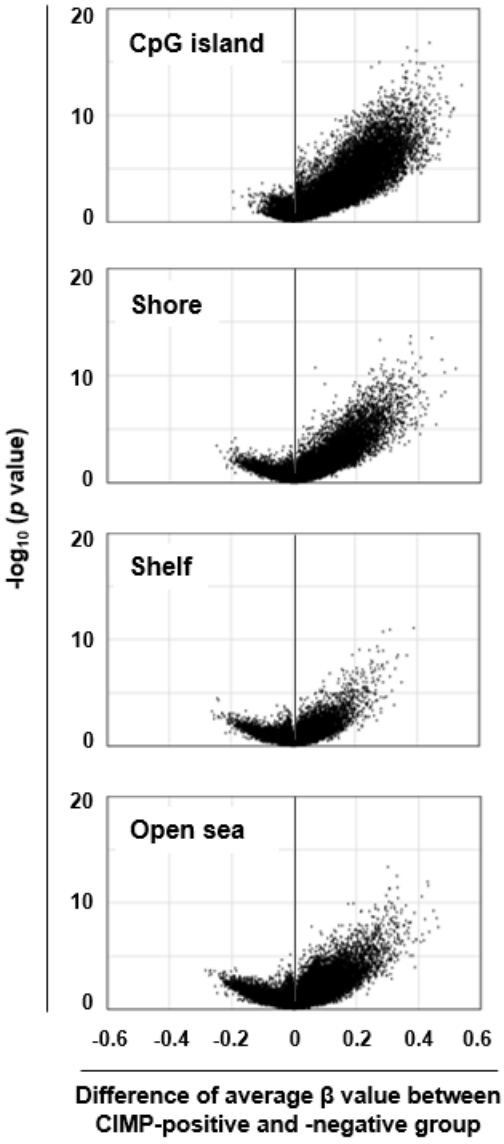
Supplementary Figures for

ARID1A loss-of-function induces the CpG island methylator phenotype

Harumi Yamada[1, 2], Hideyuki Takeshima[1, *], Ryoji Fujiki[3], Satoshi Yamashita[1], Shigeki Sekine[4], Takayuki Ando[5], Naoko Hattori[1], Atsushi Okabe[3], Takaki Yoshikawa[6], Kazutaka Obama[2], Hitoshi Katai[6], Atsushi Kaneda[3] and Toshikazu Ushijima[1, **]

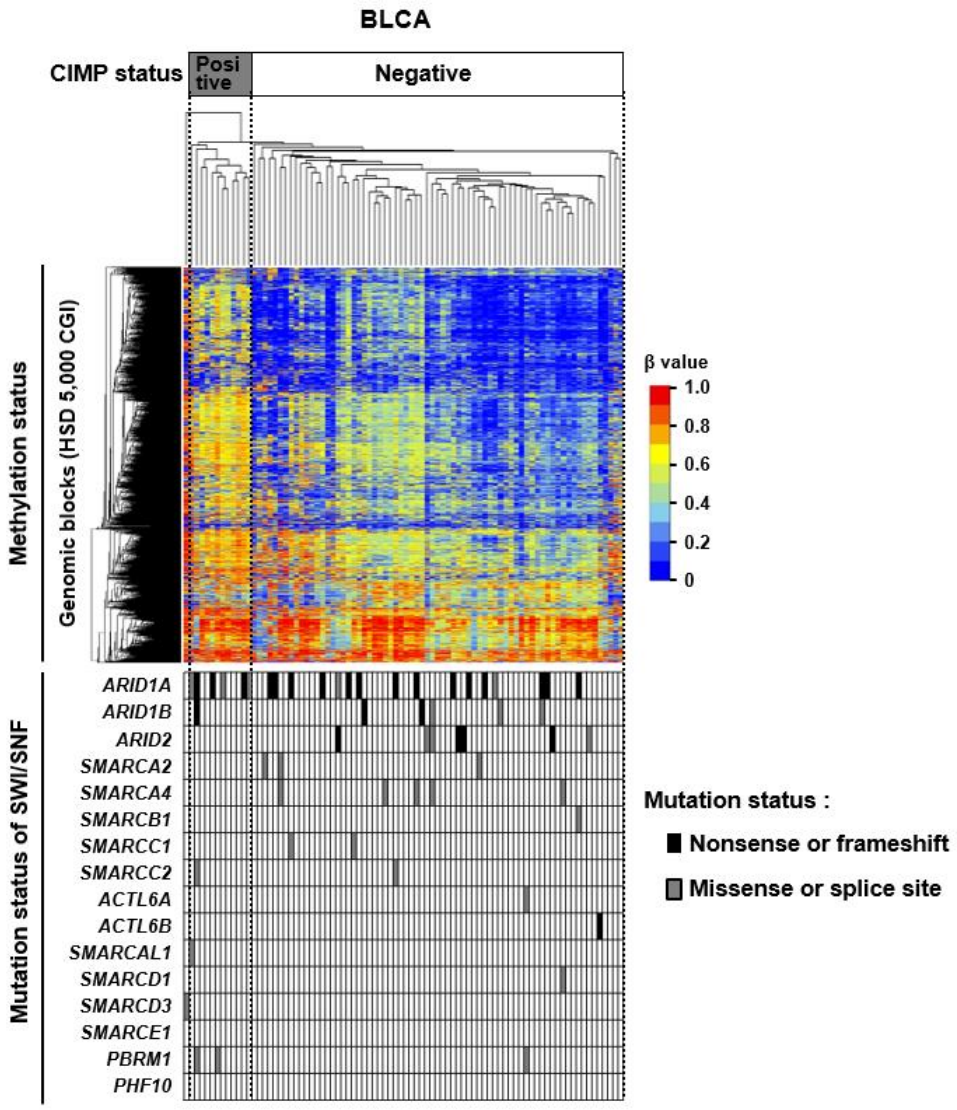
**Corresponding author (main). e-mail: tushijim@ncc.go.jp

*Corresponding author (additional). e-mail: hitakesh@ncc.go.jp



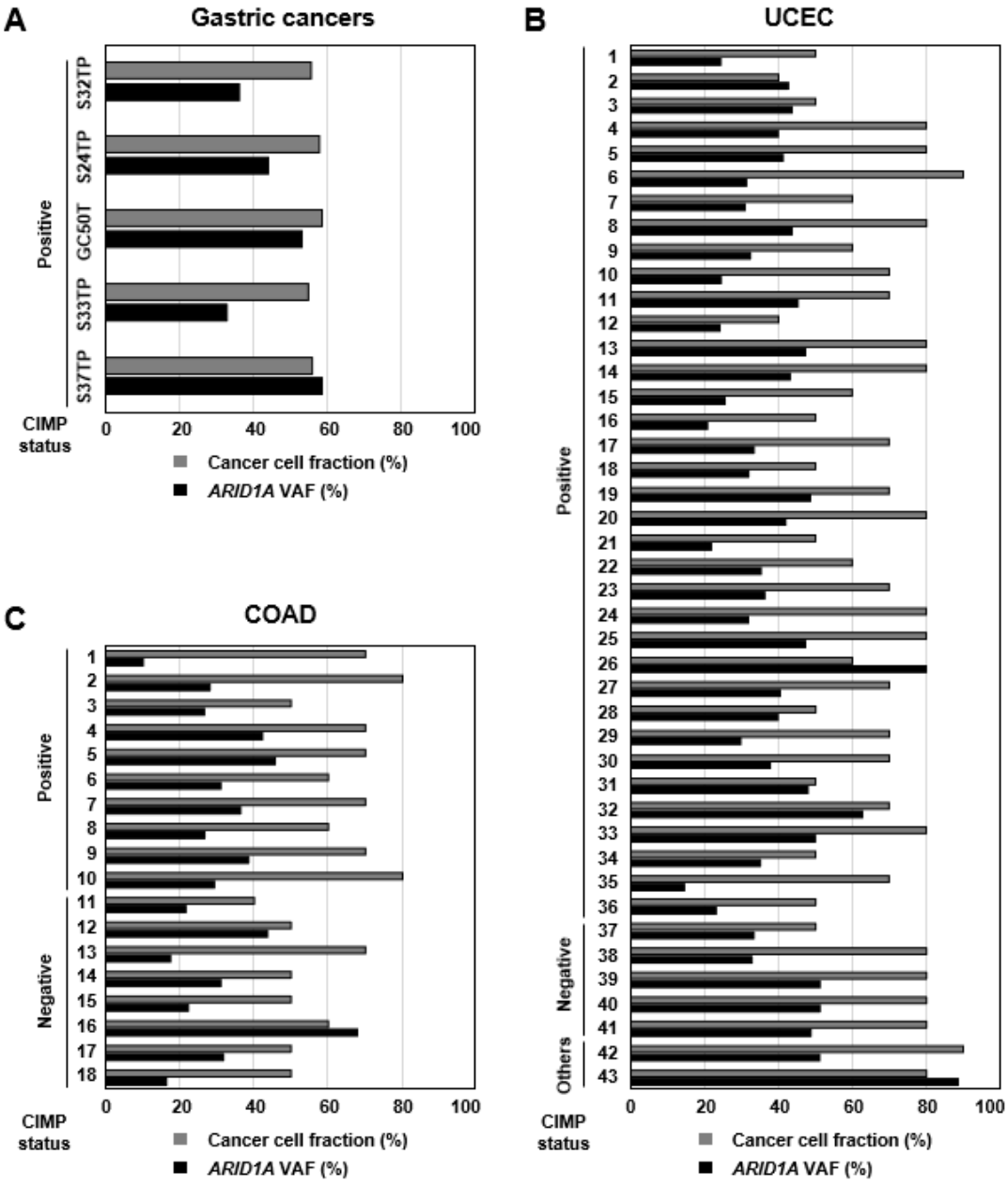
Supplementary Figure 1. Volcano plot analysis of methylation differences between CIMP-positive and -negative primary gastric cancers.

p values for the difference between the average β values of the CIMP-positive and -negative groups was calculated by the unpaired Student's t-test, and plotted in the y-axis against the difference in the x-axis. The number of hypermethylated ($\Delta\beta > 0.1$) genomic regions with higher $-\log_{10}$ (*p* value) and large differences was greater in the CIMP-positive cancers regarding CpG islands and non-CGIs (Shore, Shelf and Open sea).



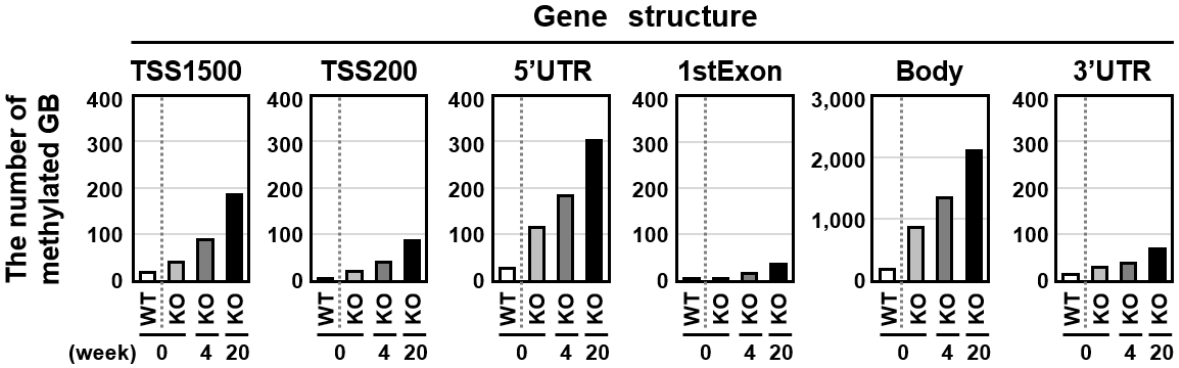
Supplementary Figure 2. Association between SWI/SNF component mutations and CIMP status in bladder epithelial cancer (BLCA), which is not affected by EBV infection.

The data of random 100 BLCAs were obtained from the TCGA database, and cases with a cancer cell fraction of 40% or more were selected. Unsupervised cluster analysis was performed using 5,000 genomic blocks in CpG islands with HSD. In BLCA, *ARID1A* mutation was not associated with the CIMP status (p = 0.073).



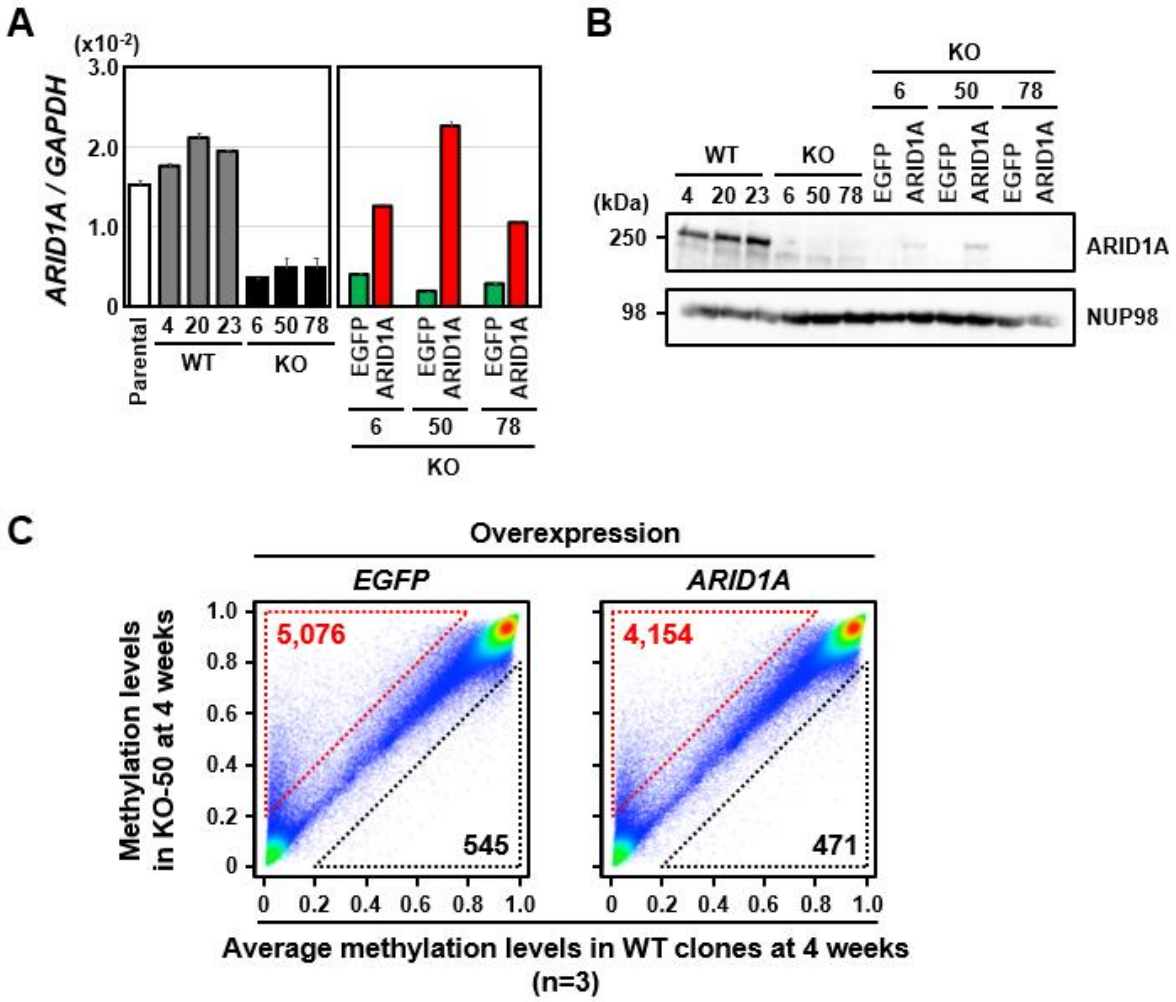
Supplementary Figure 3. The cancer cell fraction and variant allele frequency of *ARID1A* in individual CIMP-positive samples.

Among CIMP-positive cancers, 20% of gastric cancers **(A)**, 58% of UCEC **(B)**, and 50% of COAD **(C)** had variant allele frequencies of 50% (40-60%) in cancer cells estimated by cancer cell fraction.



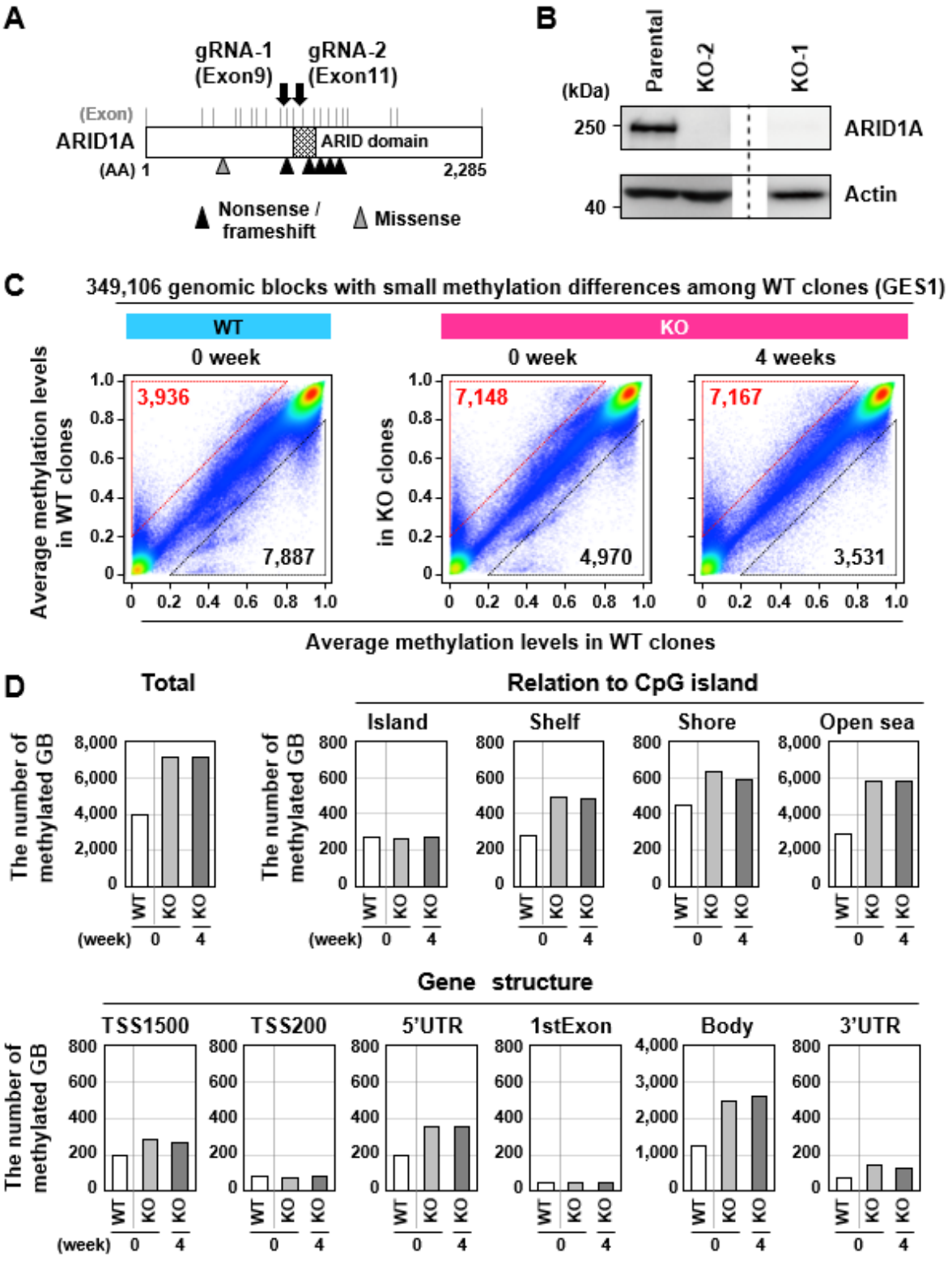
Supplementary Figure 4. Aberrant DNA methylation induction depending upon gene structures in 293FT cells.

The number of genomic blocks hypermethylated by *ARID1A* KO against gene structure. Among 269,533 genomic blocks on autosomes with a Δβ value of 0.05 or less among three WT clones, the number of hypermethylated genomic blocks increased in all the locations against a gene structure, depending upon culture period.



Supplementary Fig. 5 *ARID1A* rescue slightly inhibited induction of aberrant DNA methylation by *ARID1A* knockout.

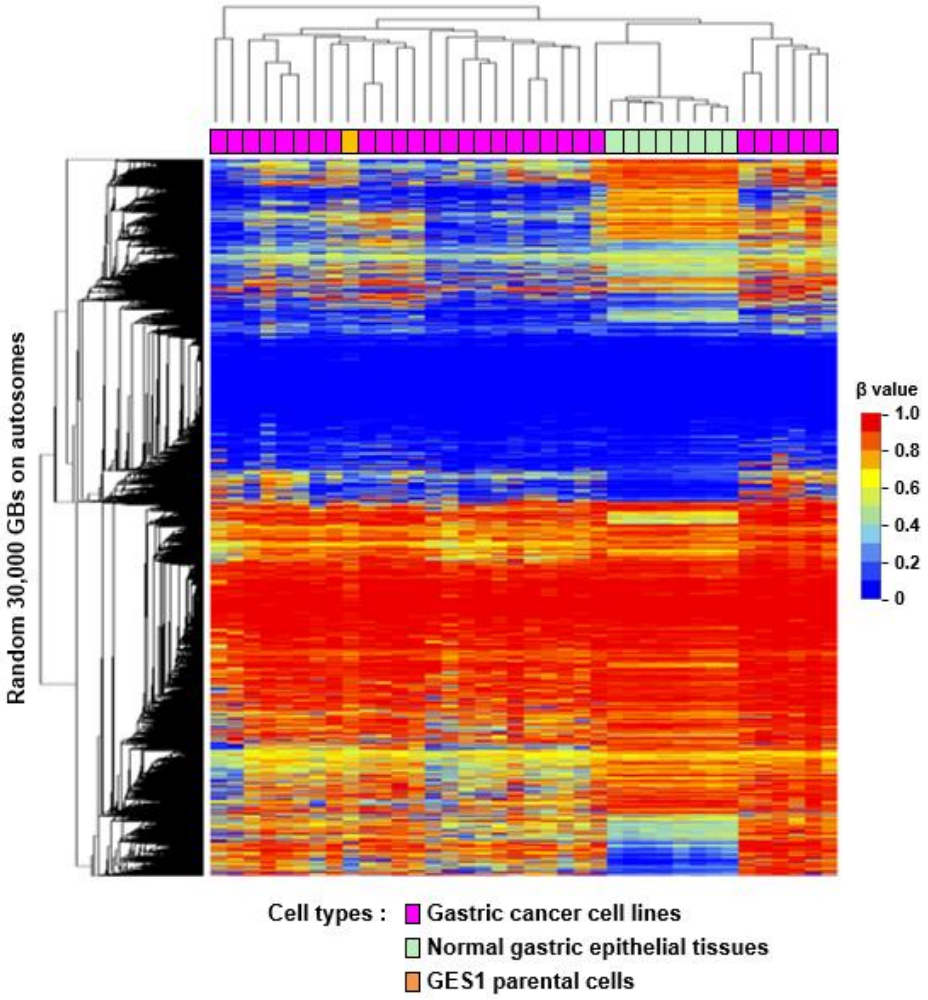
(A) mRNA expression levels of *ARID1A* in *ARID1A*-rescued cells. mRNA expression level of *ARID1A* was sufficient only in KO-50. **(B)** ARID1A protein levels in *ARID1A*-rescued cells. ARID1A protein was sufficiently expressed only in KO-50. **(C)** Comparison of DNA methylation levels between *EGFP*-overexpressed cells and *ARID1A*-rescued cells (KO-50). The number of hypermethylated ($\Delta\beta > 0.2$) genomic blocks (shown by the triangle with the dotted red line) slightly decreased in *ARID1A*-rescued cells (4,154 genomic blocks) compared to *EGFP*-overexpressed cells (5,076 genomic blocks).


A
gRNA-1 (Exon9)
gRNA-2 (Exon11)
(Exon)
ARID1A
ARID domain
(AA) 1
2,285
Nonsense / frameshift
Missense
B
Parental
KO-2
KO-1
(kDa)
250
ARID1A
40
Actin
C
349,106 genomic blocks with small methylation differences among WT clones (GES1)
WT
KO
0 week
0 week
4 weeks
3,936
7,887
7,148
4,970
7,167
3,531
Average methylation levels in WT clones
in KO clones
Average methylation levels in WT clones
D
Total
Relation to CpG island
Island
Shelf
Shore
Open sea
The number of methylated GB
(week) 0 4
Gene structure
TSS1500
TSS200
5'UTR
1stExon
Body
3'UTR

1

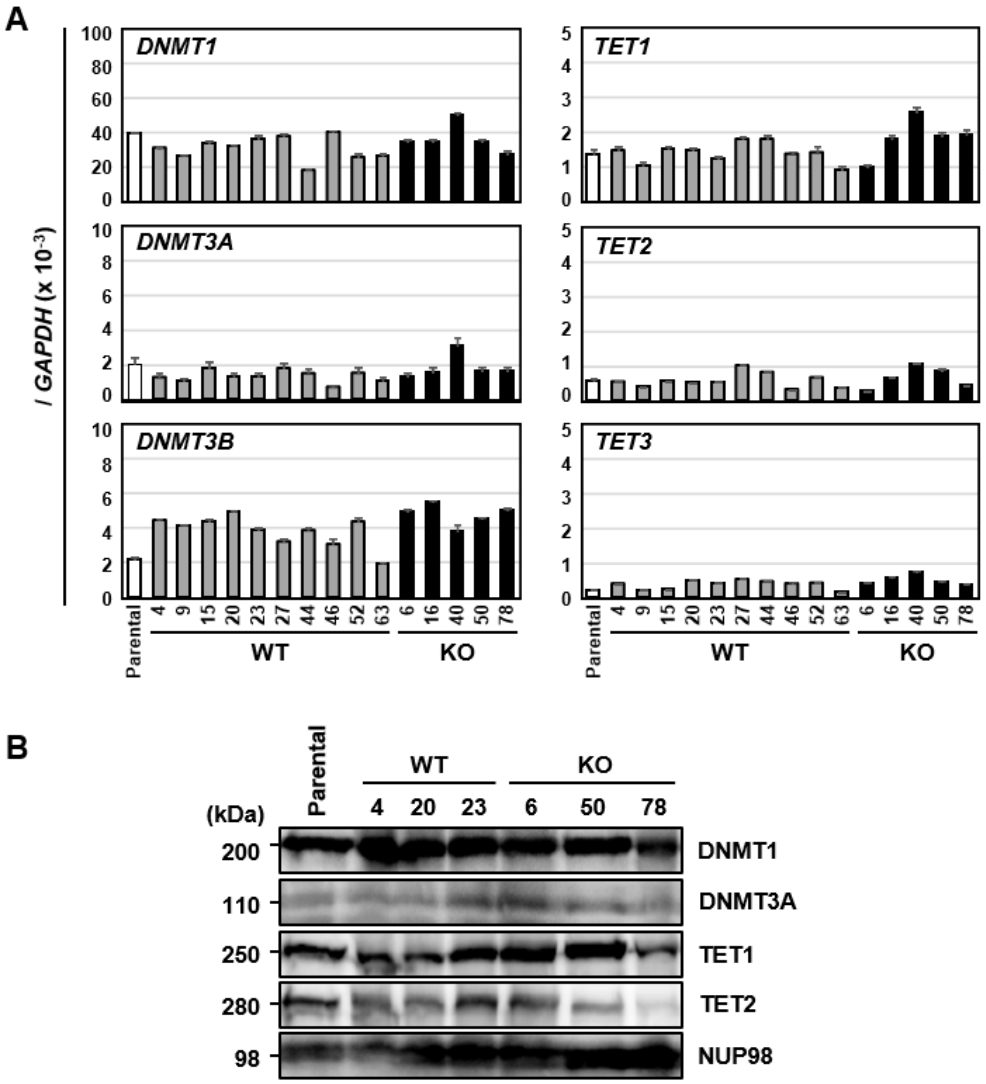
Supplementary Figure 6. *ARID1A* knockout induced aberrant DNA methylation in GES1 cells.

(A) The location of *ARID1A* KO gRNA. The target sequences were located in exon 9 (gRNA-1) and Exon 11 (gRNA-2). **(B)** ARID1A protein levels in KO clones. Complete depletion of ARID1A protein was confirmed in two *ARID1A* KO clones (KO-1 and KO-2). **(C)** Comparison of average DNA methylation levels between WT and KO clones. Among the 349,106 genomic blocks with a $\Delta\beta$ value of 0.05 or less among two WT clones, 3,936 genomic blocks were hypermethylated in WT clones at the time point of KO clone isolation ('0' week; approximately 4 weeks after introduction of gRNA and Cas9 nuclease). 7,148 and 7,167 genomic blocks were hypermethylated ($\Delta\beta > 0.2$) at 0 and 4 weeks of culture, respectively (shown by the triangle with the dotted red line). **(D)** The number of hypermethylated genomic blocks by *ARID1A* KO in GES1 cells against a CpG island and gene structure. Hypermethylation was induced by *ARID1A* KO in most locations against a CpG island and a gene structure at '0' week.



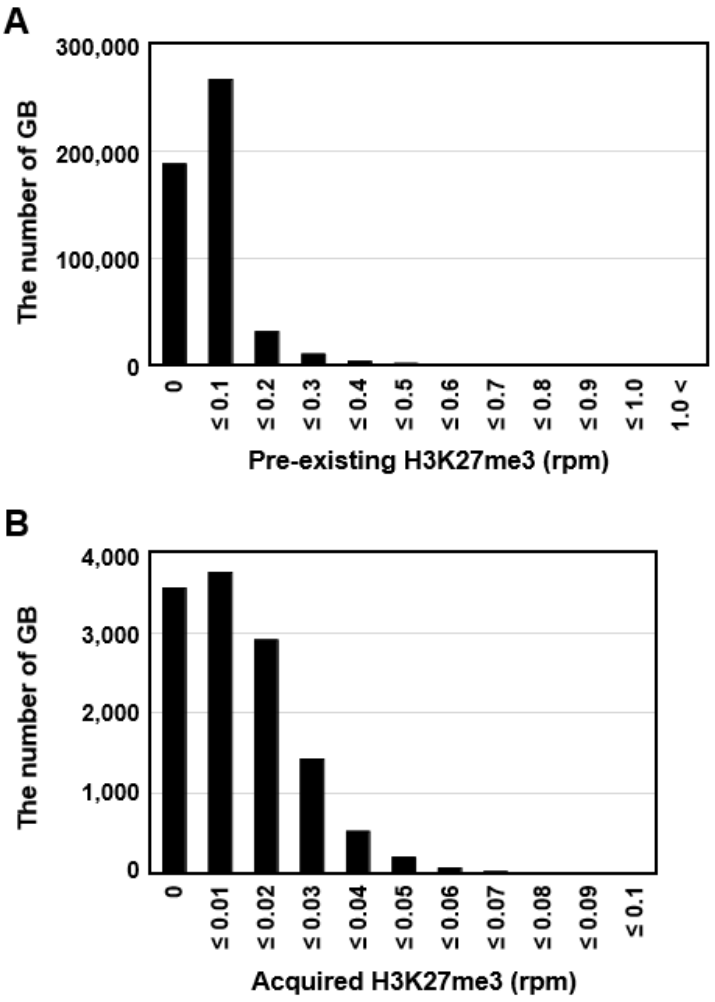
Supplementary Fig. 7. Methylation profile of GES1 parental cells.

Unsupervised cluster analysis was performed using 30,000 randomly selected genomic blocks on autosomes among 29 gastric cancer cell lines, 8 normal gastric epithelial tissues, and GES1 parental cells. The GES1 parental cells clustered with gastric cancer cell lines, not with normal gastric epithelial tissues, showing that aberrant DNA methylation has been already induced in GES1 cells, although the degree was small.



Supplementary Fig. 7 Expression levels of DNA methylation writers and erasers.

(A) mRNA expression levels of DNA methylation-related genes. mRNA expression levels of DNA methylation writers (*DNMT1*, *DNMT3A* and *DNMT3B*) and erasers (*TET1*, *TET2* and *TET3*) were analyzed by qRT-PCR. No expression levels of any genes were changed by *ARID1A* KO. **(B)** Protein expression levels of DNA methylation-related genes. Protein expression levels of DNA methylation writers (DNMT1 and DNMT3A) and erasers (TET1 and TET2) were analyzed by western blotting. No expression levels of any genes were changed by *ARID1A* KO.



Supplementary Figure 8. The distribution of pre-existing and acquired H3K27me3 level.

(A) The distribution of pre-existing H3K27me3 level in parental cells. The number of genomic blocks without H3K27me3 (levels = 0 rpm) was 189,010, those with H3K27me3 levels less than 0.1 rpm was 267,943, and those with H3K27me3 levels more than 0.1 rpm was 48,880. **(B)** The distribution of acquired H3K27me3 level by *ARID1A* KO for 12,527 genomic blocks without H3K27me3 in both parental and all WT clones. 3,565 genomic blocks were without acquired H3K27me3, 6,701 were with acquired H3K27me3 less than 0.02 rpm, and 2,261 were with acquired H3K27me3 more than 0.02 rpm.

Supplementary Table 1 ; Sequence of *ARID1A* KO clones.

Sample		Sequence	
Target sequence (exon8)		CTTATGGCCC**TAACATGGCCAATATGCCAC**CTCAGGTTGGGTCAGGGATG	
Sequence in KO clones			
KO-6	Sequence1	CTTATGGCCCTAAAACATGGCCAATATGCCACCTCAGG	+2bp
	Sequence2	CTTATGGCCCTAA(CATGGCCAATATGC)CACCTCAGG	-14bp
KO-16	Sequence1	CTTATGGCCCTAAAACATGGCCAATATGCCACCTCAGG	+2bp
	Sequence2	CTTATG(GCCCTAACATGGCCAATATG)CCACCTCAGG	-20bp
KO-40	Sequence1	CTTATGGCCCTAAACATGGCCAATATGCCACCTCAGG	+1bp
	Sequence2	C(TTATGGCCCTAACATGGCC)AATATGCCACCTCAGG	-20bp
KO-50	Sequence1	CTTATGGCCCTAAATCATGGCCAATATGCCACCTCAGG	+2bp
	Sequence2	CTTATGGCC(CTAACATGGCC)AATATGCCACCTCAGG	-11bp
KO-78	Sequence1	CTTATGGCCCTAAACATGGCCAATATGCCACCTCAGG	+1bp

Supplementary Table 2 ; Primer sequences and conditions for qRT-PCR.

Gene	Forward primer (5' -> 3')	Reverse primer (5' -> 3')	Amplicon length (bp)	Annealing temparature (˚C)
DNMT1	CTTCAATTCGCGCACCTACTC	AGCGCTTGAAGGAGACAAAG	82	62
DNMT3A	CATGGGCGTTAGTGACAAGAG	TGCAGCTGACACTTCTTTGG	82	62
DNMT3B	GAAGCACGAGGGGAATATC	CTTTCCTGGCTGGATTCAC	76	58
TET1	GCCTTCGTCACTGCCAACCTTA	TTCACTGGGTGAGGAGCGGATG	144	61
TET2	AGCAGCAGCCAATAGGACAT	CCCTCAACATGGTTGGTTCT	132	58
TET3	CACCAAGAGTCTGCTGGACA	GGCCAGATCCCAAGTGAGTA	122	62
ARID1A	ATGGCAATCAGTTCTCCACC	GGCCGCTTGTAATTCTGCTG	100	62
ARID1B	CGTGGGCTTTGGACACTATT	CTGCGTTGTGATCAAGTGCT	192	62
SMARCA2	AGACGGCTCTCAACTCCAAA	ATTCCTGGTGTTTCTGACGG	132	60
SMARCA4	TTCAACGTCTTGCTGACGAC	AGCTTGCAGTGGTGGTTCTT	122	62
EZH2	CGGACAGCCAGGTAGCACGG	TTCGTGCCCTTGTGTGATAGC	82	62
GAPDH	AGGTGAAGGTCGGAGTCAACG	AGGGGTCATTGATGGCAACA	82	54

Supplementary Table 3 ; List of methylated TSS200 CpG islands by *ARID1A* KO.

Gene	Gene ontology annotation*
FEV	DNA-binding transcription factor activity and transcription corepressor activity
SSTR5	G protein-coupled receptor activity and somatostatin receptor activity
LOC146336	-
KCNH2	Protein homodimerization activity and obsolete signal transducer activity
WFDC2	Serine-type endopeptidase inhibitor activity and cysteine-type endopeptidase inhibitor activity
GNG13	Obsolete signal transducer activity and G-protein beta-subunit binding
LOXL3	Copper ion binding and oxidoreductase activity, acting on the CH-NH2 group of donors, oxygen as acceptor
EMX2OS	-
ICAM4	Integrin binding
NTN3	Signaling receptor binding
CLSTN1	Calcium ion binding and amyloid-beta binding
CYBA	Protein heterodimerization activity and SH3 domain binding
TFAP2A	DNA-binding transcription factor activity and sequence-specific DNA binding
LIMS2	Single organismal cell-cell adhesion
TP73	DNA-binding transcription factor activity and identical protein binding
GP1BB	Transmembrane signaling receptor activity ;
LYL1	Protein dimerization activity
CRYBA2	Protein homodimerization activity and structural constituent of eye lens
C10orf82	-
RASL10A	GTP binding and GTPase activity
LOC146336	-
DLL3	Alcium ion binding and Notch binding
FAAP20	-
CHRNB1	Extracellular ligand-gated ion channel activity and channel activity
DERL3	-
DOK7	Protein kinase binding and insulin receptor binding
ISG15	Protein tag
ENTPD2	Hydrolase activity and nucleoside-diphosphatase activity
C10orf53	-
ZNF233	Nucleic acid binding
LOC100287216	-
LHX6	DNA-binding transcription factor activity and sequence-specific DNA binding

* Gene ontology annotation was searched by GeneCards®: The Human Gene Database (https://www.genecards.org/)** .

**Safran *et al.* , Database (Oxford), 5; baq020, 2010.