



OPEN **O-GlcNAcylation on serine 40 of histone H2A promotes proliferation and invasion in triple-negative breast cancer**

Yoko Uno & Koji Hayakawa

Triple-negative breast cancer (TNBC) is characterized by resistance to conventional treatment and a poor prognosis. The *O*-linked β -N-acetylglucosamine (*O*-GlcNAc) modification of proteins has been reported to affect cancer progression. However, the key *O*-GlcNAc proteins involved in TNBC phenotypes remain unclear. Our previous study demonstrated that serine 40 of histone H2A was modified by *O*-GlcNAcylation (H2AS40Gc). Since S40 is located inside the globular domain of H2A, H2AS40Gc may be involved in the regulation of gene expression by altering chromatin conformation and could serve as the molecular basis for TNBC. The present study showed that H2AS40Gc levels were significantly higher in TNBC than in the other breast cancer subtypes. Using TNBC cells in which H2AS40Gc levels were depleted, we found that H2AS40Gc is required to promote cell proliferation and migration. The underlying mechanism of this promotion involves the accumulation of H2AS40Gc in the promoter region of *KDM5B*, a demethylase for lysine 4 of histone H3 (H3K4) that represses the expression of *KDM5B*, resulting in increased H3K4 trimethylation and elevated expression of genes related to proliferation and migration. Our findings clearly indicate that H2AS40Gc functions to promote proliferation and migration through *KDM5B* suppression and provide new insights into potential therapeutic approaches for TNBC.

According to the latest data from GLOBOCAN 2022, breast cancer is the most frequently diagnosed primary malignancy in women globally¹. Approximately 15% of all invasive breast cancers are triple-negative breast cancers (TNBC), which are resistant to endocrine therapy and anti-HER2 targeted therapy and have poor prognosis and metastasis rates^{2,3}. This highlights the need to understand the mechanisms underlying the TNBC phenotype in order to develop new therapeutic strategies for better patient outcomes.

O-linked β -N-acetylglucosamine (*O*-GlcNAc) modification occurs on the serine/threonine residues and mediates cellular signal transduction^{4,5}. Recent studies have indicated that *O*-GlcNAc plays a critical role in maintaining cellular homeostasis and is also linked to various diseases including cancer^{6,7}. Increased levels of *O*-GlcNAc transferase (OGT) protein and total *O*-GlcNAc have been observed in breast tumors of high histological grade^{8,9}. Based on these reports, OGT could be a potential therapeutic target for breast cancer¹⁰. Nevertheless, potential side effects are a concern because there are numerous target proteins of OGT. To address this, it is necessary to identify *O*-GlcNAcylated proteins specifically involved in the development of the TNBC phenotype.

Various *O*-GlcNAcylated proteins have been identified, including those located in the signaling pathways associated with cancer progression. For instance, the expression and activity of proteins involved in the MEK/ERK pathway have been reported to be modulated by *O*-GlcNAc in cancer. In breast cancer, MEK2 is directly modified by OGT at Thr13, which enhances MEK2 phosphorylation and ERK1/2 activation, thereby promoting cancer cell proliferation¹¹. β -catenin, which is situated in the Wnt/ β -catenin signaling pathway and functions in cancer survival, is directly modified by OGT at Thr41, which increases protein stability involved in cell proliferation¹². This evidence suggests the existence of a unique *O*-GlcNAcylated protein that plays a central role in TNBC.

Post-translational modifications of histones are involved in many biological processes and are closely related to gene transcription¹³. According to multiple studies, more than ten types of histone *O*-GlcNAcylation sites have been identified^{14,15}. Among these, *O*-GlcNAcylation of serine 40 on H2A (H2AS40Gc) was confirmed based on the detection of endogenous *O*-GlcNAc by mass spectrometry analysis^{16–18}. Using a specific monoclonal antibody for H2AS40Gc, the following characteristics and functions of H2AS40Gc have been clarified: (1) The

Laboratory of Genetics, Faculty of Veterinary Medicine, Okayama University of Science, Imabari-shi, Ehime 7948555, Japan. ✉email: k-hayakawa@ous.ac.jp

S40 residue is located in the globular domain of H2A, (2) H2AS40Gc is restricted to placental animals, and (3) H2AS40Gc maintains genome integrity through the DNA repair mechanism in coordination with γ H2AX and Acetylated H2AZ. These findings suggest that H2AS40Gc regulates chromatin dynamics related to transcription in mammals, and is likely to play a specific role in TNBC.

In this study, we conducted functional analysis of H2AS40Gc in TNBC to determine whether H2AS40Gc could serve as a potential molecular target for treating TNBC. Our findings demonstrated that H2AS40Gc levels were noticeably higher in TNBC than in the other subtypes. Additionally, H2AS40Gc targeted the gene encoding the histone H3 lysine 4 (H3K4) demethylase *KDM5B* and promoted cell proliferation and migration by suppressing *KDM5B* expression in TNBC.

Materials and methods

Reagents

Reagents were purchased from FUJIFILM Wako Pure Chemicals, unless otherwise stated. Primers and oligonucleotides were purchased from FASMAC (Table S1). The antibodies used in this study were listed in Table S2.

Cell culture

The cell lines used in this study were purchased from the American Type Culture Collection (ATCC) and cultured according to the protocols provided by the ATCC. The composition of the medium used for the cell culture is shown in Table S3.

Cell proliferation assay

Cell proliferation was assessed using Cell Counting Kit-8 (CCK-8) reagent (DoJINDo). Cells were seeded at 1×10^4 cells/well in 24-well dishes. An OGT inhibitor (OSMI-1, Abcam) was added to achieve a final concentration of 20 μ M, 24 h post-seeding. 200 μ L fresh medium and 10 μ L CCK-8 were added every 24 h, and the cells were incubated at 37 °C with 5% CO₂ for 2 h. Subsequently, 100 μ L of the medium was collected, and the absorbance was measured to calculate the cell numbers based on the standard values.

Cell invasion assay

The invasion assay was performed using the Cell Invasion Assay Kit (Abcam), according to the manufacturer's guidelines. An 8- μ m pore polycarbonate membrane coated with 1 mg/mL fibronectin was pretreated with serum-free DMEM for 2 h at 37 °C and 5% CO₂. After removing the medium, 1×10^3 cells were seeded in 100 μ L serum-free DMEM. After 24 h, cells on the lower surface of the chamber were fixed with 4% paraformaldehyde. After three washes with wash buffer, the samples were stained, and the fluorescence intensity was measured using a Spark microplate reader (TECAN). Cell numbers were calculated based on standard values.

Transfection

Cells were seeded in 6-well dishes at a density of 4×10^5 cells/well. Cells were transfected with 2 μ g of plasmid DNA and 2 μ L of jetOPTIMUS (Polyplus) per well. Cells were passaged at 24 h after transfection into 10-cm dishes and cultured in the presence of 10 μ g/ μ L blasticidin. The medium was changed at 48 h after transfection, and blasticidin-resistant cells were collected at 72 h for each analysis.

Construction of expression vectors

Each specific miRNA sequences targeting *H2AC4*, *H2AC7*, *H2AC8* and *H2AW* were cloned into the pcDNA 6.2-GW/EmGFP-miR vector (Thermo Fisher Scientific) for knockdown of the H2AS40 genes. Sequences targeting *KDM5B*- and *LacZ*-encoding mRNA (control) were also cloned. Human full-length *KDM5B* with a 3 $\times$ FLAG-tag was obtained from cDNA of MDA-MB-231 cells by two PCR amplifications and ligated into the pENTR/D-TOPO vector (Thermo Fisher Scientific). BigDye sequencing confirmed appropriate inserts. The 3 $\times$ FLAG-fused genes were subcloned into the pCAG-DEST-PGK-Puromycin-IRES-VENUS-pA vector^{17,19} using Gateway LR Clonase II (Thermo Fisher Scientific). Vectors for overexpression of H2AS40 and H2AS40A mutant were generated as described in a previous study¹⁵. Plasmids were purified using NucleoBond Xtra Midi (Takara), followed by PCI extraction and ethanol precipitation.

Western blotting (WB)

OGT, MGEA5, *KDM5B* and *ACTB* were analyzed using whole-cell fraction proteins extracted using RIPA buffer. Core histones and H2AS40Gc were analyzed using nuclear fraction proteins extracted using the LysoPure Nuclear and Cytoplasmic Extraction Kit. Proteins were separated by 10% or 15% SDS-PAGE, transferred to a PVDF membrane, blocked with 5% skim milk or 5% bovine serum albumin, and incubated with primary antibody. Immunoreactive bands were detected using the Immunostar Zeta and Amersham Imager 680 (Cytiva). *ACTB* and pan-H2A served as internal controls. Band intensities were quantified using ImageJ software (<https://imagej.nih.gov/ij/>).

RNA extraction and cDNA synthesis

RNA extraction was performed using the Monarch Total RNA Miniprep Kit (New England Biolabs). cDNA synthesis was performed using ReverTra Ace qPCR RT Master Mix with gDNA remover (TOYOBO). The cDNA was diluted with TE Buffer (pH 8.0, NIPPON GENE) to 10 ng/ μ L of total RNA equivalent and stored at -20 °C until use.

RNA-seq

RNA-seq library preparation was performed using 1 µg of total RNA and the NEBNext Ultra II RNA Library Prep Kit for Illumina (New England Biolabs) following the manufacturer's guidelines. Sequencing was performed using the Illumina NovaSeq 6000 system. RNA-seq data analysis, such as the identification of differentially expressed genes (DEGs, FDR < 0.05) and Gene Set Enrichment Analysis (GSEA), was performed using RaNA-seq software²⁰. GSEA within the RaNA-seq software is based on the algorithm of fast pre-ranked gene set enrichment analysis using cumulative statistic calculation. Gene ontology was performed using Metascape (<http://metascape.org/>). The analysis parameters by Metascape are Min Overlap, 3; P Value Cutoff, 0.01; Min Enrichment, 1.5.

Chromatin Immunoprecipitation (ChIP)-seq

ChIP was conducted as per previous report¹⁸. ChIP-seq libraries were prepared using 10 ng of ChIP DNA and the NEBNext Ultra II DNA Library Prep Kit for Illumina (New England Biolabs) following the manufacturer's guidelines. Sequencing was performed using the Illumina NovaSeq 6000 system. Sequence reads were aligned to the human genome (hg38) using Bowtie2 and peaks were identified using MACS2 in the Galaxy browser (www.galaxy.psu.edu). Genes exhibiting a significant H2AS40Gc peak within ± 5 kb of the transcriptional start site (TSS) or within the gene body were classified as H2AS40-target genes. Principal component analysis (PCA) and correlation assays were performed using deepTools in the Galaxy browser.

Quantitative PCR (qPCR)

RT- and ChIP-qPCR was performed using the THUNDERBIRD Next SYBR qPCR Mix (TOYOBO) and Quant Studio 5 (Thermo Fisher Scientific). PCR was performed under the following thermocycling conditions: denaturation at 95 °C for 1 min, 40 cycles of denaturation at 95 °C for 10 s, and elongation at 60 °C for 30 s. After the reaction, the Cq values for each gene were calculated. The expression levels were determined using the $\Delta\Delta C_t$ method. Values were normalized to VCP expression. The heatmap was visualized using MeV software²¹. ChIP signals were calculated as ChIP DNA/Input DNA.

Statistical analyses

All experiments were performed at least in biological triplicates, and values were expressed as mean ± standard deviation. Student's t-test was performed for comparison of WB, RT-qPCR, ChIP-qPCR, cell proliferation and invasion assays. ** P-value < 0.01 was considered to be statistically significant.

Results

High H2AS40Gc level in TNBC cells

To elucidate H2AS40Gc levels in TNBC cells, the amount of H2AS40Gc was measured by WB in three breast cancer cell lines categorized as luminal-type BC (MCF-7, BT-474 and MDA-MB-361) and four TNBC cell lines (MDA-MB-231, Hs578T, HCC-1937 and BT-20). The results showed that the levels of H2AS40Gc were significantly higher in all TNBC cells than in luminal-type BC cells by approximately eight-fold (Fig. 1A), suggesting that H2AS40Gc is more utilized in TNBC compared to in luminal-type BC.

To determine why H2AS40Gc levels were higher in TNBCs, the expression levels of OGT and O-GlcNAcase MEPA5 were analyzed. In all cell lines, the expression of the two enzymes did not correlate with H2AS40Gc levels (Fig. 1B). There are 19 genes encoding canonical H2A in the human genome, including four genes in which the 40th amino acid is serine (H2AS40) and 15 genes in which the 40th amino acid is alanine (H2AA40), considering that the expression levels of these isoforms determine H2AS40Gc levels in breast cancer. RT-qPCR results showed that neither H2AS40 nor H2AA40 genes expression levels were correlated with H2AS40Gc levels (Fig. 1C). Thus, high H2AS40Gc levels in TNBC cells were independent of the expression levels of enzymes and H2A genes.

H2AS40Gc promotes the expression of genes involved in TNBC cell proliferation

To elucidate the function of H2AS40Gc, H2AS40Gc was suppressed by knockdown (KD) of the expression of all four H2AS40 genes, *H2AC4*, *H2AC7*, *H2AC8* and *H2AW*, in the TNBC cell line MDA-MB-231. The efficiency of the KD was validated using RT-qPCR. As expected, there was a predominant down-regulation of the four H2AS40 genes, but not H2AA40 genes. Importantly, the decrease in H2AS40Gc without affecting the amount of total H2A was confirmed by WB (Fig. 2A and S1). Three of H2AA40 genes, *H2AC6*, *H2AC18* and *H2AJ*, were significantly up-regulated by the KD, suggesting that the increase in expression occurred to compensate for the decrease in H2A protein levels associated with KD. In the following experiments, we used H2AS40-KD TNBC cells, in which H2AS40Gc was functionally inhibited.

To determine which genes are controlled by H2AS40Gc in TNBC, GSEA using RNA-seq data of H2AS40-KD TNBC cells and MDA-MB-231 cells expressing *LacZ*-targeted miRNA (control) was performed. GSEA revealed the most significant term to be “pathway in cancer,” which belongs to the KEGG pathway database (Fig. 2B). Indeed, most genes belonging to this pathway showed variable expression in H2AS40-KD TNBC cells (Fig. S2A,B). To identify genes whose expression was significantly altered in this pathway, the number of DEGs was calculated. A total of 486 genes were down-regulated and 367 genes were up-regulated in response to H2AS40-KD compared to the control (Fig. S2C).

Seventeen genes belonging to KEGG term “pathways in cancer” were significantly down-regulated by H2AS40-KD (Fig. 2C). This down-regulation was confirmed by RT-qPCR in H2AS40-KD TNBC cells (Fig. S2D). The expression of these 17 genes was characterized as higher in TNBC cells than in luminal-type BC cells (Fig. 2D). Among the 17 genes, *FN1*, *STAT1*, *BMP4*, *SLC2A1*, *LAMC1*, and *CXCR4* are known to be involved in

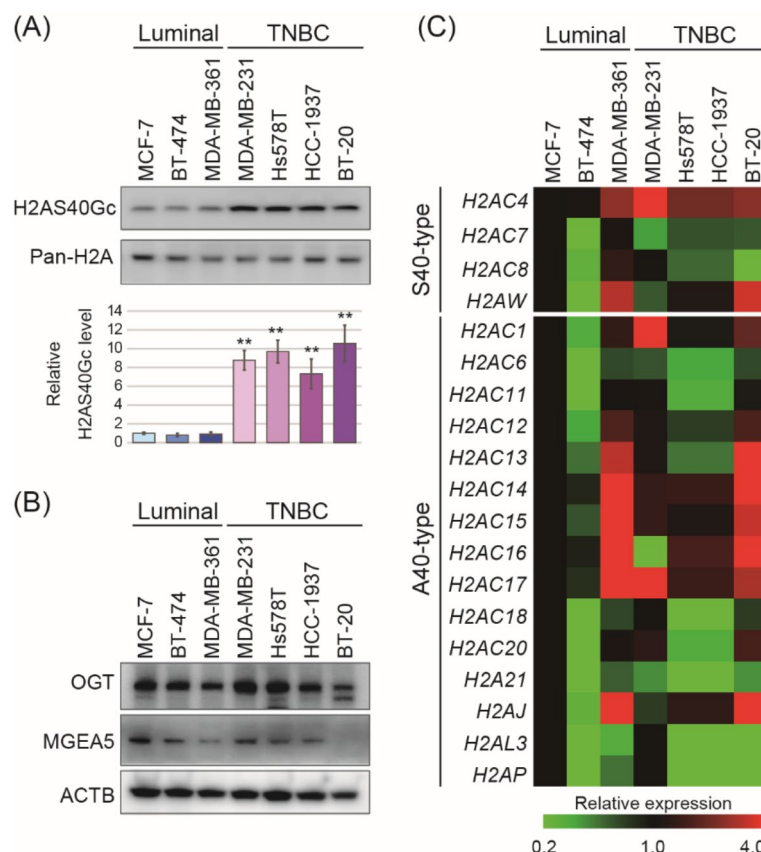


Fig. 1. Serin 40 of H2A is highly O-GlcNAcylated in TNBC cells. **(A)** H2AS40Gc levels in luminal-type BC cell lines (MCF-7, BT-474 and MDA-MB-361) and TNBC cell lines (MDA-MB-231, Hs578T, HCC-1937 and BT-20). Top, western blotting (WB) was performed using nucleoproteins. Bottom, H2AS40Gc levels were normalized to H2A from WB band intensities. Relative values were based on the expression in MCF-7 cells equaling 1. The original blots are presented in Supplementary Fig. S8. **(B)** Protein levels of O-GlcNAc transferase (OGT) and O-GlcNAcase MGEA5 in TNBC and luminal-type BC cell lines. **(C)** Results from RT-qPCR determining mRNA levels of 19 canonical H2A genes. The values are visualized as a heatmap. Color scale bars indicate individual gene expression relative to expression in MCF-7 cells.

breast cancer growth^{22–28}, suggesting that H2AS40Gc promotes cell proliferation by regulating the expression of these genes. Hereafter, this gene set is referred to as the 17 downregulated genes.

H2AS40Gc accelerates proliferation and invasion of TNBC cells

The findings from RNA-seq analysis implied that H2AS40Gc may play a role in the proliferation of TNBC cells. OGT is reportedly important for cell proliferation⁶, and its importance has been confirmed in both MCF-7 and MDA-MB-231 cells (Figs. S3A–C). To test the importance of H2AS40Gc, a cell proliferation assay was conducted using H2AS40-KD TNBC cells. The results showed a decrease in cell proliferation relative to the control (Fig. 3A). A more pronounced decrease in cell proliferative capacity due to H2AS40-KD occurred in MDA-MB-231 than in luminal-type BC cell MCF-7 (Fig. 3A, S3D,E). The rescue experiment confirmed that overexpression of H2AS40, but not H2AS40A mutant, recovered the proliferation activity in H2AS40-KD TNBC cells (Fig. 3B). In parallel, overexpression of H2AS40, but not H2AS40A, restored the expression of 17 downregulated genes in H2AS40-KD TNBC cells (Fig. S3F). These findings provided strong evidence that H2AS40Gc contributes to cell proliferation.

Gene ontology (GO) analysis of genes that were downregulated by H2AS40-KD showed a significant enrichment of the term “positive regulation of cell migration” (Fig. 3C and S4A). As TNBC is known to exhibit a higher migration potential than luminal-type BC²⁶, H2AS40Gc may also function in invasive activity. The results of the invasion assay showed a significant reduction in the number of invading cells in H2AS40-KD TNBC cells, and this reduction was rescued by the overexpression of exogenous H2AS40 but not H2AS40A-mutant (Fig. 3D). Six genes were common between genes in the KEGG term “pathways in cancer” and GO term “positive regulation of cell migration” and were downregulated by H2AS40-KD (Fig. S4B). These analyses clearly demonstrated that H2AS40Gc promotes proliferation and invasive potential by controlling the expression of genes.

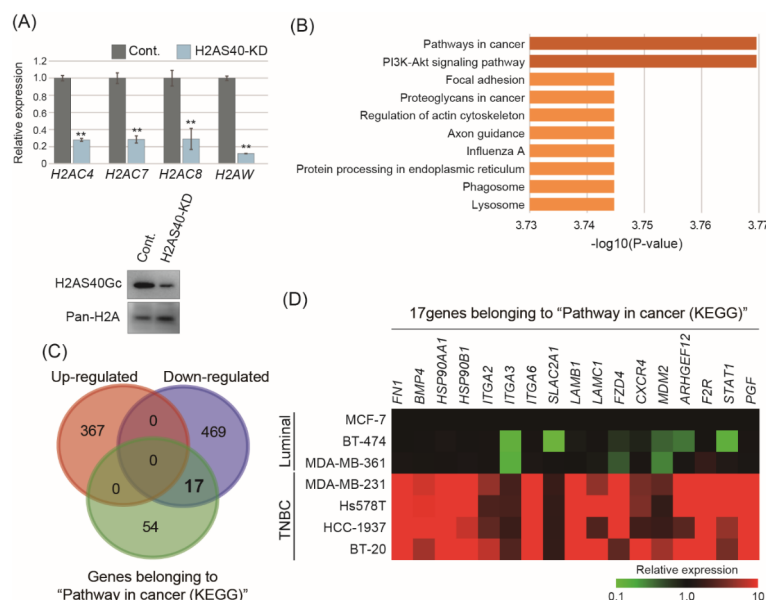


Fig. 2. H2AS40Gc controls expression of genes related to cell proliferation. **(A)** Depletion of H2AS40Gc by the knockdown (KD) of H2AS40 genes in MDA-MB-231 cells. The control cells were transfected with a vector expressing miRNA against *lacZ* (Cont.). Top, Results from RT-qPCR determining mRNA levels of four H2AS40 genes. The relative values are based on the expression of Cont. equaling 1. Bottom, H2AS40Gc levels in H2AS40Gc-KD TNBC cells by WB. The original blots are presented in Supplementary Fig. S8. **(B)** Gene Set Enrichment Analysis (GSEA) using RNA-seq data of H2AS40-KD TNBC cells. The vertical axis represents the top 10 ranked KEGG pathways with the smallest *P*-values that were significantly different between Cont. and H2AS40-KD. The related results are shown in Fig. S2A,B. **(C)** The number of genes belonging to KEGG term "pathways in cancer" and upregulated and downregulated DEGs (Cont. vs H2AS40-KD). **(D)** Results from RT-qPCR determining mRNA of 17 downregulated genes in TNBC and lumina-type BC cells. The values are visualized as a heatmap. Color scale bars indicate individual gene expression relative to expression in MCF-7 cells.

H2AS40Gc targets *KDM5B*, which encodes a histone demethylase

To investigate whether the 17 downregulated genes are directly regulated by H2AS40Gc, ChIP-seq data using an antibody specific to H2AS40Gc in TNBC cell line MDA-MB-231 was obtained. Contrary to expectations, when modification levels were observed in the genome browser, no significant peaks were observed in any of the 17 downregulated genes, either near the promoter region (± 5 kb from the TSS) or within the gene-body region (Fig. S5). Thus, these genes were indirectly regulated by H2AS40Gc.

To search for H2AS40Gc target genes that regulate the expression of the 17 downregulated genes, we first analyzed whether H2AS40Gc is a histone modification that promotes or represses gene expression in TNBC. ChIP-seq data for trimethylation on H3K4 (H3K4me3) and trimethylation on lysine 36 of H3 (H3K36me3), which are known as positive histone marks, as well as trimethylation on lysine 27 of H3 (H3K27me3), which is known as a repressive histone mark, were obtained, and compared to ChIP-seq data of H2AS40Gc in MDA-MB-231 cells. PCA using genomic localization information showed that H2AS40Gc was closest to H3K27me3 (Fig. 4A). The correlation coefficient showed that H2AS40Gc was mostly positively correlated with H3K27me3 (Fig. 4B) and H3K27me3 was enriched on the genomic region around H2AS40Gc-peaks (Fig. S6), suggesting that H2AS40Gc has a repressive effect on gene expression.

Based on the results of ChIP-seq analysis, the direct target genes of H2AS40Gc may be up-regulated by H2AS40-KD. Next, we explored for genes shared by candidate factors that bind to the promoter regions of 17 downregulated genes (236 genes from the ChIP-Atlas database²⁹), genes that were targets of H2AS40Gc (3,912 genes from ChIP-seq data) and genes up-regulated by H2AS40-KD (365 genes from RNA-seq data). As result, one gene, *KDM5B*, was identified, which encodes a demethylase of H3K4 (Fig. 4C). There were significant peaks around the TSS of *KDM5B* in MDA-MB-231 cells as well as in other TNBC line BT-20 cells, H2AS40Gc were enriched at the promoter of *KDM5B* (Fig. 4D), indicating that accumulation of H2AS40Gc in the *KDM5B* region is common in TNBC cells. *KDM5B* mRNA and protein expression was lower in TNBC cells than in luminal-type BC cells (Fig. 4E and S7A), and a decrease in *KDM5B* localization at the genic region of all 17 downregulated genes was observed in TNBC cells (Fig. 4F), suggesting that H2AS40Gc functions to repress the expression of *KDM5B*.

H2AS40Gc represses *KDM5B* expression in TNBC cells

When H2AS40Gc levels at the promoter region of the *KDM5B* locus decreased (Fig. 5A), the total *KDM5B* expression and *KDM5B* binding status at the genic region of all 17 downregulated genes were elevated (Fig. 5A,B

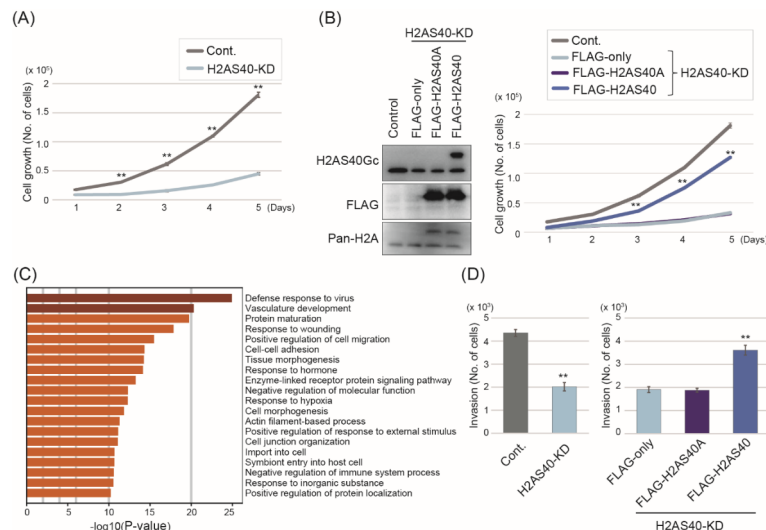


Fig. 3. H2AS40Gc is involved in the proliferation and invasion of TNBC cells. **(A)** Cell proliferation was measured for five day in H2AS40-KD TNBC cells. **(B)** Rescue experiment of cell proliferation assay in H2AS40-KD TNBC cells overexpressing H2AS40. Cells overexpressing miRNA against lacZ (Cont.), FLAG-only or H2AS40A mutant were used as the controls. Left, exogenous H2AS40A mutant and H2AS40 were detected by WB. Right, cell proliferation was measured for five days. The original blots are presented in Supplementary Fig. S8. **(C)** Gene ontology (GO) analysis of down-regulated DEGs in H2AS40-KD TNBC cells. GO analysis was categorized based on the biological process using Metascape software. **(D)** Invasion assay in H2AS40-KD TNBC cells. Left, cell invasion activity was measured using an invasion chamber. Right, the rescue experiment for the invasion assay was performed using H2AS40-KD TNBC cells overexpressing H2AS40.

and S7B). Rescue experiments by overexpression of H2AS40 in H2AS40-KD TNBC cells strongly showed suppressive effect of H2AS40Gc against *KDM5B* expression (Fig. S7C). Thus, H2AS40Gc at the promoter region repressed the expression of *KDM5B*.

Based on the above results, we hypothesized that the decrease in cell proliferative and invasive activities caused by H2AS40-KD could be due to the increased expression of *KDM5B* and the concomitant decrease in the expression of 17 downregulated genes. To verify this hypothesis, we carried out H2AS40-KD and *KDM5B*-KD concurrently in the MDA-MB-231 cells (Fig. 5C). In the double-KD cells, restoration of the expression of all 17 downregulated genes was observed (Fig. 5D), along with the simultaneous restoration of cell proliferation and invasion levels (Fig. 5E,F). This suggested that the elevation of *KDM5B* expression was the cause of the H2AS40-KD TNBC phenotype. Overexpression experiments of *KDM5B* in MDA-MB-231 cells showed a decrease in cell proliferation and invasive capacity (Fig. 5G,H), supporting this hypothesis.

As *KDM5B* is a demethylase of H3K4, it was expected that the increase in *KDM5B* expression by H2AS40-KD would result in a decrease in H3K4me3 levels. Finally, analysis of H3K4me3 levels at 17 downregulated gene loci in H2AS40-KD TNBC cells showed that the levels significantly decreased at all loci (Fig. 6A), which was consistent with the *KDM5B* binding state shown in Fig. 5B.

Discussion

The O-GlcNAc regulatory system has been reported to play a crucial role in cellular processes in various types of cancers^{6,7}. However, the majority of these studies have focused on the overall O-GlcNAc level and the expression of the modifying enzyme OGT, without delving into which O-GlcNAcylated proteins, particularly, have a primary role in cancer function. This study aimed to examine the function of H2AS40Gc, specifically focusing on its level and impact on TNBC. The results showed that H2AS40Gc was highly modified in TNBC cells compared to luminal-type BC cells and enhanced the cell proliferative and invasive potential of TNBC cells via the following mechanism: enrichment of H2AS40Gc at the promoter region repressed the expression of *KDM5B*, and thereby H3K4me3 levels were elevated by decreasing *KDM5B* binding at the loci of genes related to cell proliferation and migration (Fig. 6B).

Evaluation of the seven breast cancer cell lines revealed that TNBC cell lines exhibited higher levels of H2AS40Gc compared to luminal-type BC lines, implying its significance as a molecular mechanism in TNBC. Despite the absence of patient samples for analysis in this study, the consistently elevated levels of H2AS40Gc across all TNBC cell lines strongly suggest a similar pattern in vivo. To determine the cause of the high H2AS40Gc levels in TNBC, the expression levels of OGT and MGEA5 were examined; however, there was no correlation with the enzymes. Additionally, gene expression analysis of the S40- and A40-types of canonical H2A failed to explain the H2AS40Gc levels. Mechanisms other than the regulation of enzymes or H2A gene expression may be involved in determining H2AS40Gc levels. OGT contains a domain known as tetratricopeptide repeats, which are crucial for protein interactions^{30,31}. MGEA5 also contains a stalk domain that facilitates protein binding^{32,33}.

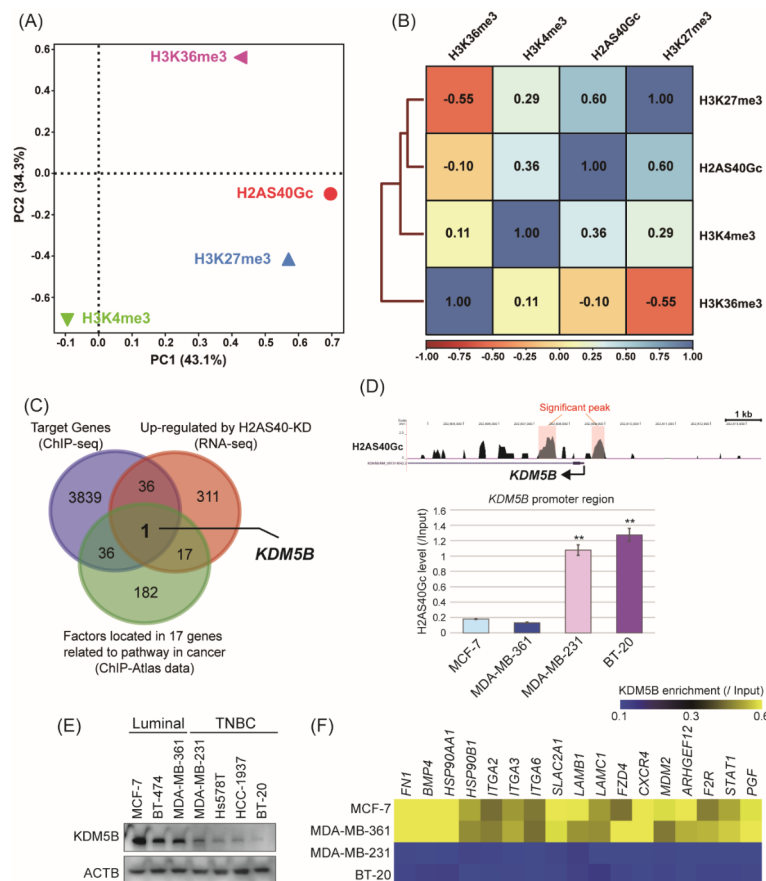


Fig. 4. H2AS40Gc is enriched on the genic region of *KDM5B*. **(A)** Proximity of the genomic localization of H2AS40Gc to H3K4me3, H3K27me3, and H3K36me3. Principal component analysis (PCA) was performed using mapping data of ChIP-seq. **(B)** Correlation coefficients for genomic localization of H2AS40Gc with H3K4me3, H3K27me3 and H3K36me3. **(C)** Number of H2AS40Gc-target genes (ChIP-seq), upregulated DEGs in H2AS40-KD (RNA-seq), and candidate factors that bound around the TSS of 17 downregulated genes (ChIP-Atlas data). **(D)** H2AS40Gc levels around the *KDM5B* locus. Top, peaks of H2AS40Gc around *KDM5B* in TNBC. Bottom, results from ChIP-qPCR determining H2AS40Gc enrichment on the *KDM5B* promoter region in TNBC and luminal-type BC cells. **(E)** KDM5B expression in TNBC and luminal-type BC cells by WB. ACTB was used as an internal control. The original blots are presented in Supplementary Fig. S8. **(F)** Enrichment of KDM5B on the promoter regions of 17 downregulated genes in TNBC and luminal-type BC cells. Heatmaps show ChIP-qPCR data. The values were normalized using the input data. Color scale bars indicate the H2AS40Gc level of each gene.

Therefore, it is expected that there are factors that interact with these enzymes to regulate *O*-GlcNAc levels in H2AS40-specific manner. Although more ten types of histone *O*-GlcNAcylation have been discovered, the factors involved in establishing *O*-GlcNAcylation of histones remain unclear^{14,15}. To advance H2AS40Gc drug discovery, it is necessary to uncover the establishment system of the *O*-GlcNAc modification of H2AS40.

Among the genes with variable expression by H2AS40-KD, those related to proliferation and migration were predominantly enriched, particularly in GSEA and GO analyses. All 17 downregulated genes were also detectable in luminal-type BC cells although their expression was lower than that in TNBCs, suggesting that this difference in expression, which is dependent on H2AS40Gc levels, is responsible for the difference in the proliferative and invasive potential between luminal-type BC and TNBC. Moreover, these genes have been shown to have critical functions in BC as well as in various cancer types such as lung and pancreatic cancer^{34–38}. According to the findings of this study using TNBC, H2AS40Gc may play a role in regulating the proliferative and invasive potential of cancer cells in various types of cancers with high H2AS40Gc levels. Therefore, H2AS40Gc is a crucial factor for determining the mechanisms that contribute to the phenotypes of specific cancer types including TNBC.

KDM5B was identified as a direct target gene of H2AS40Gc in TNBC cells. KDM5 has been reported to be abnormally expressed in various cancers, and its higher expression is associated with poor prognosis^{39–41}. Notably, KDM5B expression was lower in TNBCs than in luminal-type BC, and elevated KDM5B levels reduced the proliferative and invasive potential of TNBC cells⁴², which was supported by our results. Our findings revealed that KDM5B directly controls the expression of genes associated with proliferation and invasion by modulating the level of H3K4me3, indicating that it is essential to maintain low levels of KDM5B to preserve

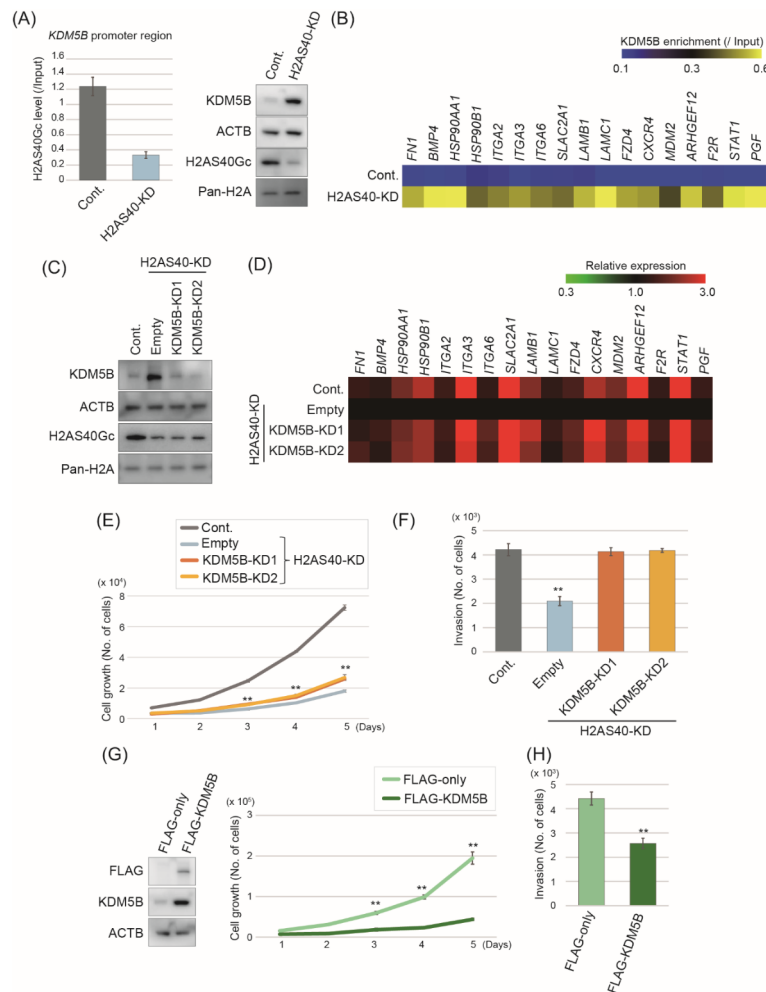


Fig. 5. H2AS40Gc represses *KDM5B* expression. **(A)** *KDM5B* expression in H2AS40-KD TNBC cells. Left, enrichment of H2AS40Gc on *KDM5B* promoter region in H2AS40-KD TNBC cells by ChIP-qPCR. Right, *KDM5B* protein levels in H2AS40-KD TNBC cells by WB. The original blots are presented in Supplementary Fig. S8. **(B)** Enrichment of *KDM5B* on the promoter regions of 17 downregulated genes in H2AS40-KD TNBC cells. Heatmaps show ChIP-qPCR data. The values were normalized using the input data. Color scale bars indicate the H2AS40Gc level of each gene. **(C)** Levels of H2AS40Gc and *KDM5B* in MDA-MB-231 cells knockdowned in both H2AS40 and *KDM5B* (H2AS40-KDM5B-KD TNBC cells). **(D)** Expression levels of 17 downregulated genes in H2AS40-KDM5B-KD TNBC cells by RT-qPCR. The values are visualized as a heatmap. Color scale bars indicate individual gene expression relative to expression in empty. **(E,F)** Cell proliferation **(E)** and invasion activity **(F)** in H2AS40-KDM5B-KD TNBC cells. **(G,H)** *KDM5B* suppresses cell proliferation in TNBC cells. Exogenous *KDM5B* levels by WB and proliferation **(G)** and invasion activity **(H)** in MDA-MB-231 cells overexpressing *KDM5B*.

the proliferative and invasive potential of TNBC. Since H2AS40Gc acts as a suppressor of *KDM5B* transcription by localizing to the *KDM5B* promoter region, the H2AS40Gc-*KDM5B* axis is a crucial regulatory system for the TNBC-specific phenotypes.

KDM5B has been reported to promote metastasis and progression in other types of cancer cells, such as melanoma, which is contrary to its function in TNBCs. This functional disparity may be attributed to differences in the *KDM5B*'s partner proteins. *KDM5B* is associated with nucleosome remodeling and the deacetylase (NuRD) complex, suggesting a collaborative repressive function⁴². Furthermore, *KDM5B* exhibits demethylase-independent transcriptional activity by binding to nuclear receptors, including estrogen receptor, and enhances its transcriptional activity⁴³. In this study, we observed that *KDM5B* was transcriptionally regulated by H2AS40Gc. In addition to elucidating the transcriptional regulatory mechanism of *KDM5B*, understanding which partner proteins cooperate with *KDM5B* in different types of cancers will be essential for the development of anticancer drugs targeting the H2AS40Gc-*KDM5B*-axis in the future.

TNBC is a malignancy with significant unmet medical needs, and new drugs are urgently required^{44,45}. Our study identified H2AS40Gc as a potential molecular target for TNBC treatment. Since H2AS40Gc is a modified protein, conventional drug discovery strategies for developing inhibitors and agonists cannot be directly applied^{46,47}. Therefore, a strategy that targets regulatory factors related to the establishment of residue-specific

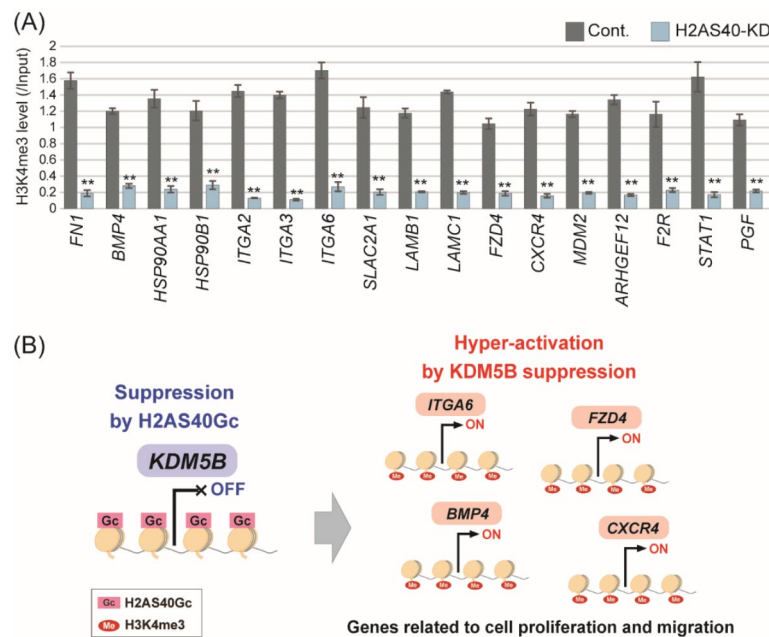


Fig. 6. Regulation of proliferation and invasiveness by H2AS40Gc through KDM5B repression in TNBC cells. **(A)** Results from ChIP-qPCR determining H3K4me3 enrichment on the promoter regions of 17 downregulated genes in H2AS40-KD TNBC cells. The values were normalized using the input data. **(B)** Perspective model for the regulation of cell proliferation and invasion by H2AS40Gc in TNBC cells. H2AS40Gc suppresses KDM5B expression by localizing to the promoter region, which induces H3K4me3 elevation and high expression of genes related to cancer cell proliferation and invasion.

modifications or utilizes molecules that recognize modifier proteins is preferable. Novel intracellular proteolytic methods utilizing antibody molecules have been actively developed^{48,49}. By combining these technologies with antibodies specific for H2AS40Gc, it is possible to regulate H2AS40Gc levels in specific cell types in vivo. In the future, elucidation of the regulatory mechanism of H2AS40Gc and the development of H2AS40Gc-related technologies will likely lead to the eradication of TNBC.

Data availability

All data generated during this study are included in the published article and Supplementary Information files. ChIP- and RNA-seq data are available in GEO with accession number GSE274331.

Received: 9 October 2024; Accepted: 20 March 2025

Published online: 24 March 2025

References

- Bray, F. et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* **74**, 229–263 (2024).
- Griffiths, C. L. & Olin, J. L. Triple negative breast cancer: A brief review of its characteristics and treatment options. *J. Pharm. Pract.* **25**, 319–323 (2012).
- Yin, L., Duan, J. J., Bian, X. W. & Yu, S. C. Triple-negative breast cancer molecular subtyping and treatment progress. *Breast Cancer Res.* **22**, 61 (2020).
- Hardivillé, S. & Hart, G. W. Nutrient regulation of signaling, transcription, and cell physiology by O-GlcNAcylation. *Cell. Metab.* **20**, 208–213 (2014).
- Nagel, A. K. & Ball, L. E. O-GlcNAc transferase and O-GlcNAcase: Achieving target substrate specificity. *Amino Acids* **46**, 2305–2316 (2014).
- Sun, L., Lv, S. & Song, T. O-GlcNAcylation links oncogenic signals and cancer epigenetics. *Discov. Oncol.* **12**, 54 (2021).
- He, X. F., Hu, X., Wen, G. J., Wang, Z. & Lin, W. J. O-GlcNAcylation in cancer development and immunotherapy. *Cancer Lett.* **566**, 216258 (2023).
- Krzeslak, A., Forma, E., Bernaciak, M., Romanowicz, H. & Bryś, M. Gene expression of O-GlcNAc cycling enzymes in human breast cancers. *Clin. Exp. Med.* **12**, 61–65 (2012).
- Barkovskaya, A. et al. O-GlcNAc transferase Inhibition differentially affects breast cancer subtypes. *Sci. Rep.* **9**, 5670 (2019).
- Kozal, K. & Krzeslak, A. Role of O-GlcNAcylation in breast cancer biology. *Cell. Physiol. Biochem.* **57**, 183–197 (2023).
- O-GlcNAcylation of MEK 2 promotes the proliferation and migration of breast cancer cells | Glycobiology | Oxford Academic. <https://academic.oup.com/glycob/article/31/5/571/5976287>
- Harosh-Davidovich, S. B. & Khalaila, I. O-GlcNAcylation affects β -catenin and E-cadherin expression, cell motility and tumorigenicity of colorectal cancer. *Exp. Cell Res.* **364**, 42–49 (2018).
- Allis, D., Jenuwein, T., Reinberg, D., Caparros, M. L. & Epigenetics Second Edition. (2015). <https://cshlpress.com/default.tpl?cart=17225707112301504192&action=full&-eqskudatarq=987>

14. Hirosawa, M., Hayakawa, K., Shiota, K. & Tanaka, S. Histone O-GlcNAcylation and potential biological functions. *OBM Genet.* **2**, 1–18 (2018).
15. Dupas, T., Lauzier, B. & McGraw, S. O-GlcNAcylation: The sweet side of epigenetics. *Epigenet. Chromatin* **16**, 49 (2023).
16. Hirosawa, M. et al. Novel O-GlcNAcylation on Ser(40) of canonical H2A isoforms specific to viviparity. *Sci. Rep.* **6**, 31785 (2016).
17. Hayakawa, K. et al. H2A O-GlcNAcylation at serine 40 functions genomic protection in association with acetylated H2AZ or γ H2AX. *Epigenet. Chromatin* **10**, 51 (2017).
18. Hayakawa, K., Nishitani, K., Tanaka, S. & Kynurenine 3-OH-kynurenine, and anthranilate are nutrient metabolites that alter H3K4 trimethylation and H2AS40 O-GlcNAcylation at hypothalamus-related loci. *Sci. Rep.* **9**, 19768 (2019).
19. Hayakawa, K. et al. Epigenetic switching by the metabolism-sensing factors in the generation of orexin neurons from mouse embryonic stem cells. *J. Biol. Chem.* **288**, 17099–17110 (2013).
20. Prieto, C. & Barrios, D. RaNA-Seq: interactive RNA-Seq analysis from FASTQ files to functional analysis. *Bioinformatics*. <https://doi.org/10.1093/bioinformatics/btz854> (2019).
21. Chu, V. T., Gottardo, R., Raftery, A. E., Bumgarner, R. E. & Yeung, K. Y. MeV + R: using MeV as a graphical user interface for bioconductor applications in microarray analysis. *Genome Biol.* **9**, R118 (2008).
22. Chen, C., Ye, L., Yi, J., Liu, T. & Li, Z. FN1 mediated activation of aspartate metabolism promotes the progression of triple-negative and luminal a breast cancer. *Breast Cancer Res. Treat.* **201**, 515–533 (2023).
23. Koromilas, A. E. & Sexl, V. The tumor suppressor function of STAT1 in breast cancer. *JAKSTAT* **2**, e23353 (2013).
24. Choi, S. et al. BMP-4 enhances epithelial mesenchymal transition and cancer stem cell properties of breast cancer cells via Notch signaling. *Sci. Rep.* **9**, 11724 (2019).
25. Oh, S., Kim, H., Nam, K. & Shin, I. Glut1 promotes cell proliferation, migration and invasion by regulating epidermal growth factor receptor and integrin signaling in triple-negative breast cancer cells. *BMB Rep.* **50**, 132–137 (2017).
26. Wang, Y. Q., Huang, G., Chen, J., Cao, H. & Xu, W. T. LncRNA SNHG6 promotes breast cancer progression and epithelial-mesenchymal transition via miR-543/LAMC1 axis. *Breast Cancer Res. Treat.* **188**, 1–14 (2021).
27. Xu, C., Zhao, H., Chen, H. & Yao, Q. CXCR4 in breast cancer: Oncogenic role and therapeutic targeting. *Drug Des. Devel. Ther.* **9**, 4953–4964 (2015).
28. Neve, R. M. et al. A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes. *Cancer Cell.* **10**, 515–527 (2006).
29. Zou, Z., Ohta, T., Oki, S. & ChIP-Atlas 3.0: A data-mining suite to explore chromosome architecture together with large-scale regulome data. *Nucleic Acids Res.* **52**, W45–W53 (2024).
30. Haltiwanger, R. S., Holt, G. D. & Hart, G. W. Enzymatic addition of O-GlcNAc to nuclear and cytoplasmic proteins. Identification of a uridine diphospho-N-acetylglucosamine:peptide beta-N-acetylglucosaminyltransferase. *J. Biol. Chem.* **265**, 2563–2568 (1990).
31. Jinek, M. et al. The superhelical TPR-repeat domain of O-linked GlcNAc transferase exhibits structural similarities to importin alpha. *Nat. Struct. Mol. Biol.* **11**, 1001–1007 (2004).
32. Ito, R. et al. TET3-OGT interaction increases the stability and the presence of OGT in chromatin. *Genes Cells* **19**, 52–65 (2014).
33. Stephen, H. M., Adams, T. M. & Wells, L. Regulating the regulators: Mechanisms of substrate selection of the O-GlcNAc cycling enzymes OGT and OGA. *Glycobiology* **31**, 724–733 (2021).
34. Huang, Y. et al. Systematic analysis of the prognostic significance and roles of the integrin alpha family in non-small cell lung cancers. *Adv. Ther.* **40**, 2186–2204 (2023).
35. McCarthy, J. B., Basara, M. L., Palm, S. L., Sas, D. F. & Furcht, L. T. The role of cell adhesion proteins? Laminin and fibronectin? In the movement of malignant and metastatic cells. *Cancer Metast. Rev.* **4**, 125–152 (1985).
36. Barginear, M. F. et al. The heat shock protein 90 chaperone complex: An evolving therapeutic target. *Curr. Cancer Drug Targ.* **8**, 522–532 (2008).
37. Queiroz, K. C. S. et al. Protease-activated receptor-1 drives pancreatic cancer progression and chemoresistance. *Int. J. Cancer* **135**, 2294–2304 (2014).
38. Tagliaferri, P. et al. Pharmacological modulation of peptide growth factor receptor expression on tumor cells as a basis for cancer therapy. *Anticancer Drugs* **5**, 379–393 (1994).
39. Lu, P. J. et al. A novel gene (PLU-1) containing highly conserved putative DNA/chromatin binding motifs is specifically up-regulated in breast cancer. *J. Biol. Chem.* **274**, 15633–15645 (1999).
40. Yamane, K. et al. PLU-1 is an H3K4 demethylase involved in transcriptional repression and breast cancer cell proliferation. *Mol. Cell.* **25**, 801–812 (2007).
41. Yamamoto, S. et al. JARID1B is a luminal lineage-driving oncogene in breast cancer. *Cancer Cell.* **25**, 762–777 (2014).
42. Klein, B. J. et al. The histone-H3K4-specific demethylase KDM5B binds to its substrate and product through distinct PHD fingers. *Cell. Rep.* **6**, 325–335 (2014).
43. Catchpole, S. et al. PLU-1/JARID1B/KDM5B is required for embryonic survival and contributes to cell proliferation in the mammary gland and in ER + breast cancer cells. *Int. J. Oncol.* **38**, 1267–1277 (2011).
44. Popovic, L. S. et al. Targeting triple-negative breast cancer: A clinical perspective. *Oncol. Res.* **31**, 221–238 (2023).
45. Greco, S. et al. Update on classic and novel approaches in metastatic triple-negative breast cancer treatment: A comprehensive review. *Biomedicines* **11**, 1772 (2023).
46. Kumar, A., Emdad, L., Fisher, P. B. & Das, S. K. Targeting epigenetic regulation for cancer therapy using small molecule inhibitors. *Adv. Cancer Res.* **158**, 73–161 (2023).
47. Conery, A. R., Rocnik, J. L. & Trojer, P. Small molecule targeting of chromatin writers in cancer. *Nat. Chem. Biol.* **18**, 124–133 (2022).
48. Clift, D. et al. A method for the acute and rapid degradation of endogenous proteins. *Cell* **171**, 1692–1706e18 (2017).
49. Lim, S. et al. BioPROTACs as versatile modulators of intracellular therapeutic targets including proliferating cell nuclear antigen (PCNA). *Proc. Natl. Acad. Sci. U S A* **117**, 5791–5800 (2020).
50. Kanehisa, M., Furumichi, M., Sato, Y., Matsuura, Y. & Ishiguro-Watanabe, M. KEGG: Biological systems database as a model of the real world. *Nucleic Acids Res.* **53**, D672–D677 (2025).

Acknowledgements

We thank to Dr. Nobuyuki Fujiwara (Okayama University of Science) for technical assistance of WB and cell proliferation assay.

Author contributions

K.H. designed the study. Y.U. and K.H. performed all experiments. Y.U. and K.H. prepared the manuscript.

Funding

This work was supported by the Japan Society for the Promotion of Science (KAKENHI grant numbers 22H02543 and 24K21925), the Lotte Shigemitsu Prize, the Lotte Research Promotion Grant, Promotion of Okayama University of Science Research Project (OUS-RP-22-2), the Medical Research Grant of Takeda Science Foundation.

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-025-95394-z>.

Correspondence and requests for materials should be addressed to K.H.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025