From bloodjournal.hematologylibrary.org at University of Tokyo on April 30, 2014. For personal use only.

analysis for 96 samples in several hours, at most only 250 unidentified protein peaks were detectable. In the present study, we integrated the proteomics server for the huge data set "Expressionist" (Genedata A.G., Basel, Switzerland) with high-end mass spectrometers to maximize the quality and quantity of protein catalogs transferred from mass spectrometers. We first describe the discovery phase providing a panel of novel diagnostic molecules from quantification of 14 064 peptides and identification of 4763 proteins. As the functional validation phase, we further examined the physiological potential of an identified diagnostic marker candidate, calpain-2 (CAN2), particularly concerning the association of its activity with survival or progression of ATL cells.

Materials and methods

PBMCs and cell lines

Peripheral blood mononuclear cells (PBMCs) from 6 normal donors, 5 asymptomatic carriers, and 9 HAM/TSP patients used in the screening analysis were collected in the St. Marianna University School of Medicine. Those from 9 ATL patients were collected in the Imamura Bun-in Hospital. PBMCs from 4 ATL patients used for the validation experiments were provided by the Joint Study on Predisposing Factors of ATL Development. The others from 4 HAM/TSP patients were collected in the St. Marianna University School of Medicine. The use of these human specimens in this study was approved by individual institutional ethical committees: the Ethical Committee of Yokohama Institute, RIKEN (approval code Yokohama H22-3); the Ethical Committee of St. Marianna University School of Medicine; the Institutional Review Board of Imamura Bun-in Hospital; and the Ethical Committee of the University of Tokyo (approval code 10-50). This study was conducted in accordance with the Declaration of Helsinki.

SO-4, KOB, and KK1 cells were kindly provided by Dr Yasuaki Yamada, cultured in RPMI 1640 supplemented with 10% fetal bovine serum (Cell Culture Bioscience, Tokyo, Japan), 100 kU/L interleukin 2 (Cell Science & Technology Institute Inc., Tokyo, Japan), and 1 × antibiotic-antimycotic solution (Sigma-Aldrich, MO). Jurkat, SUP-T1, CCRF-CEM, and MOLT-3 cells were cultured in RPMI 1640 supplemented with 10% fetal bovine serum and 1 × antibiotic-antimycotic solution. All cell lines were grown at 37°C in 5% CO_2. $CD3^+CD4^+CD25^+CCR4^+$ T cells were isolated with anti-CD3-FITC (eBioscience, San Diego, CA), anti-CCR4-PE (Becton Dickinson, CA), anti-CD4-Cy7 (eBioscience), and anti-CD25-APC (eBioscience) on a Cell Sorter JSAN (Bay Bioscience, Hyogo, Japan).

Sample preparation for mass spectrometric analysis

The $CD4^+CD25^+CCR4^+$ T cells were washed with phosphate-buffered saline 3 times and lysed in denaturation buffer (8 M urea in 50 mM ammonium bicarbonate). After sonication, reduction with 5 mM tris(2-carboxyethyl) phosphine (Sigma-Aldrich) at 37°C for 30 minutes, and alkylation with 25 mM iodacetamide (Sigma-Aldrich) at room temperature for 45 minutes, lysates were digested with Trypsin GOLD (Promega, WI) with protein/enzyme ratio of 25:1 at 37°C for 12 hours. The digested peptides were desalted with Oasis HLB μElution plate (Waters, MA). The collected samples were dried up with a Vacuum Spin Drier (TAITEC Co. Ltd., Saitama, Japan) and subjected to mass spectrometric analyses.

Liquid chromatography tandem mass spectrometry (LC/MS/MS)

The digested peptides were separated on a 0.1 × 200 mm homemade C_{18} column using a 2-step linear gradient, 2% to 35% acetonitrile for 95 minutes and 35% to 95% acetonitrile for 15 minutes in 0.1% formic acid with a flow rate of 200 nL/min. The eluting peptides were analyzed with a QSTAR-Elite mass spectrometer (AB Sciex, CA) in the smart information-dependent acquisition mode of Analyst QS software 2.0 (AB Sciex). The other parameters on QSTAR-Elite were shown as follows: DP = 60, FP = 265, DP2 = 15, CAD = 5, IRD = 6, IRW = 5, curtain gas = 20, and ion spray voltage = 2000 V.

Two-dimensional (2D) LC/MS/MS

Tryptic digests of $CD4^+CD25^+CCR4^+$ T cells were dissolved in 10 mM ammonium formate in 25% acetonitrile and fractionated by a 0.2 × 250 mm monolith strong cation exchange column (GL Science, Tokyo, Japan). Peptides were eluted with an ammonium formate gradient from 10 mM to 1 M in curve = 3 mode for 70 minutes using a Prominence high-performance liquid chromatography (HPLC) system (Shimadzu Corporation, Kyoto, Japan). The eluate was fractionated into 20 fractions and analyzed individually by LTQ-Orbitrap-Velos mass spectrometer (Thermo Scientific, Bremen, Germany) accompanied with the Ultimate 3000 nano-HPLC system. The fractionated peptide samples were separated with the same gradient used in the QSTAR-Elite system described previously and analyzed by LTQ-Orbitrap-Velos acquiring a full MS scan on Fourier-transition mode with MS resolution = 60 000 and simultaneously MS/MS scans for the 20 most intense precursor ions in each MS spectrum on ion-trap mode with regular resolution. Other important parameters for LTQ-Orbitrap-Velos were as follows: capillary temp = 250, source voltage = 2 kV, MS scan range = mass-to-charge ratio (m/z) 400 to 1600, acquire data dependent CID MS/MS for top-20 intense precursors, and dynamic exclusion enabled during 30 seconds. For protein identification, all MS/MS spectra were searched against SwissProt database version 2012_06 (20 232 human protein sequences) using SEQUEST algorithm on ProteomeDiscoverer 1.3 software (Thermo Scientific) with the following parameters: MS tolerance = 3 ppm, MS/MS tolerance = 0.8 Da, maximum missed cleavages = 2, enzyme = trypsin, taxonomy = *Homo sapiens*, fixed modification = carbamidomethylation on cysteine, and variable modification = oxidation on methionine. We accepted the protein identification satisfying the false discovery rate <1% by Percolator false discovery rate estimation algorithm on ProteomeDiscoverer.

Label-free quantification analysis

The LC/MS/MS raw data were imported into the Expressionist RefinerMS module and subjected to the following data processing and relative quantification steps. The total work flow on the RefinerMS module is shown in supplemental Figure 1 (see the *Blood* Web site). The LC/MS/MS raw data set from 29 clinical samples was displayed in 2D planes (m/z vs retention time [RT]). The chromatogram grid was applied to all planes: scan counts = 10, polynom order = 3, and RT smoothing = 0. The planes were simplified by subtracting background noises using chromatogram chemical noise subtraction: RT window = 50 scans, quantile subtraction = 50%, and RT smoothing = 3 scans. After the noise subtraction, data points with intensity <10 were clipped to zero. The RT variety among 29 planes was adjusted by chromatogram RT alignment: RT transformation window = 0.2 minutes, RT search interval = 5 minutes, m/z window = 0.1 Da, and gap penalty = 1. Peaks were detected by chromatogram summed peak detection: summation window = 5 scans, overlap = 50, minimum peak size = 4 scans, maximum merge distance = 10 points, peak RT splitting = true, intensity profiling = max, gap/peak ratio = 1%, refinement threshold = 5, consistency threshold = 0.8, and signal/noise threshold = 1. The detected peaks were grouped into isotopic clusters derived from each molecule using 2-step chromatogram isotopic peak clustering. The first parameters were as follows: minimum charge = 1, maximum charge = 10, maximum missing peaks = 0, first allowed gap position = 3, RT window = 0.1 minute, m/z tolerance = 0.05 Da, isotope shape tolerance = 10, and minimum cluster size ration = 1.2. The second parameters were as follows: minimum charge = 1, maximum charge = 10, maximum missing peaks = 0, first allowed gap position = 3, RT window = 0.1 minute, m/z tolerance = 0.05 Da, and minimum cluster size ration = 0.6.

Expression vectors and siRNA

For the Δ_{19}CAN2 construct, the *CAPN2* fragment was amplified with primers 5′-CATGTCGACTCCCACGAGAGGGCCATCAAGT-3′ and 5′-CATTCTAGATCAAAGTACTGAGAAACAGAGCC-3′ from pBlueBacIII *CAPN2* and cloned into pEFBOS-Myc. Prior to the overexpression experiments, we confirmed that the sequence of the inserted *CAPN2* fragment was identical to the Mammalian Gene Collection sequence (accession number

From bloodjournal.hematologylibrary.org at University of Tokyo on April 30, 2014. For personal use only.

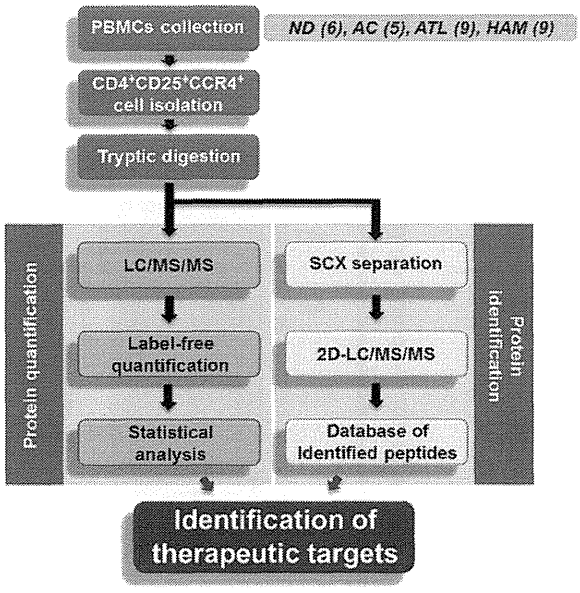

Figure 1. Schematic overview of proteomic profiling for $CD4^+CD25^+CCR4^+$ cells. PBMCs were collected from 6 normal donors, 5 asymptomatic carriers, 9 ATL patients, and 9 HAM/TSP patients, followed by isolation of the $CD4^+CD25^+CCR4^+$ subset using the cell-sorting system. The statistical candidate selection steps, including LC/MS/MS data processing, label-free quantification, and statistical analysis, were performed on the Expressionist proteome server. The protein identification database was separately established based on 2D LC/MS/MS analysis. ND, normal donors; AC, asymptomatic carriers.

BC021303). The 5-μg vector DNA was transfected to 1×10^6 cells. The small interfering RNAs (siRNAs) against *SPTAN1*, *PTMS*, *HSPE1*, and SHMT2 and siRNA universal negative control were purchased from Sigma-Aldrich. The 500-pmol siRNA oligo was transfected into 1×10^6 cells. The vectors and siRNAs were transfected into all cell lines except CCRF-CEM by Amaxa Nucleoportator transfection Kit V (Lonza, Cologne, Germany) and CCRF-CEM by Kit C (Lonza).

Cell cycle analysis and proliferation assay

For the cell cycle analysis, 1×10^5 to 2×10^5 cells were washed and agitated in 0.1% Triton-X (Sigma-Aldrich) with 100 ng/mL of ribonuclease (Sigma-Aldrich). Following addition of 1 μg/mL propidium iodide, the flow cytometric analysis was performed on FACScalibur (Becton Dickinson). The data analysis was performed using FlowJo software (Tree Star Inc., OR). Doublet events were eliminated from analyses by proper gating on FL2-W/FL2-A primary plots before histogram analysis of DNA content. Cell proliferation was estimated by measuring cell metabolic activity using Cell Counting Kit-8 (Dojindo, Kumamoto, Japan) following the manufacturer's recommendation.

Western blotting

Cells were lysed in lysis buffer [1% NP-40, 2 mM EGTA, 2 mM $MgCl_2$, 150 mM NaCl, 20 mM tris(hydroxymethyl)aminomethane–HCl (pH 7.5), 10% glycerol, containing the protease inhibitor cocktail Complete (Roche, IN)] and subjected to sodium dodecyl sulfate–polyacrylamide gel electrophoresis and transferred onto PVDF membranes. Following blocking with 4% Block Ace (Yukijirushi Nyugyo Inc., Tokyo, Japan), membranes were incubated with anti-myc (9E10; Sigma-Aldrich) or anti–α-II spectrin (Abcam, Cambridge, UK) antibodies. Membranes were then incubated with horseradish peroxidase–conjugated anti-mouse IgG (GE Healthcare, NJ) or anti-rabbit IgG (GE Healthcare), respectively, and visualized with Western Lightning kit (Perkin Elmer, MA).

Multiple reaction monitoring (MRM)

$CD4^+$ T cells were isolated from PBMCs using flow cytometry. The tryptic digests of the isolated cells were analyzed by 4000 Q-TRAP mass spectrometer (AB Sciex) accompanied with Ultimate 3000 nano-HPLC system. The LC gradient was as follows: 2% to 30% acetonitrile for 10 minutes and 30% to 95% acetonitrile for 5 minutes in 0.1% formic acid with a flow rate of 300 nL/min. The MRM transitions monitored were m/z 409.7/375.2 for α-II spectrin (SPTA2); m/z 538.3/889.5 for parathymosin (PTMS); m/z 507.3/147.1 for heat shock 10-kDa protein, mitochondrial (CH10); m/z 490.3/147.1 for serine hydroxymethyltransferase, mitochondrial (GLYM); and m/z 581.3/919.5 for β-actin, respectively. Individual peak areas were normalized by the peak area of β-actin. Data acquisition was performed with ion spray voltage = 2300 V, curtain gas = 10 psi, nebulizer gas = 10 psi, and an interface heating temperature = 150°C. The parameters were set as follows: declustering potential = 60, entrance potential = 10, collision cell exit potential = 10, and dwell time for each transition = 10 seconds. Collision energy was optimized to achieve maximum intensity for each MRM transition as follows: 34.03 V for m/z 409.7/175.1, 24.68 eV for m/z 538.3/889.5, 23.32 eV for m/z 507.3/147.1, 37.57 eV for m/z 490.3/147.1, and 31.58 eV for m/z 581.3/919.5.

Results

Quantitative proteome profiling of $CD4^+CD25^+CCR4^+$ T cells

A schematic overview of the screening approach is shown in Figure 1. To identify diagnostic markers expressed in HTLV-1–infected T cells, a $CD4^+CD25^+CCR4^+$ subset of PBMCs from 6 uninfected volunteers, 5 asymptomatic carriers, 9 HAM/TSP patients, and 9 ATL patients was isolated by flow cytometry (Figure 2). The averaged proportion of $CD4^+CD25^+CCR4^+$ cells in $CD4^+$ T cells from 4 clinical groups was 6.48 ± 2.46%, 13.17 ± 13.06%, 20.55 ± 10.73%, and 55.83 ± 22.40%, respectively, indicating that the occupancy of viral reservoir cells varied drastically among both pathological groups and even individuals within a group. Enrichment of the infected cells was confirmed by viral load measurement of the used samples (supplemental Figure 2). As reported previously,[10] the viral load of $CD4^+CD25^+CCR4^+$ cells (37.91 copies/100 cells on average) was ~10 times higher than that of $CD4^+CD25^-CCR4^-$ cells (4.12 copies/100 cells on average), indicating that the former cells were evidently the HTLV-1–enriched fraction. This fact strongly supports the importance of

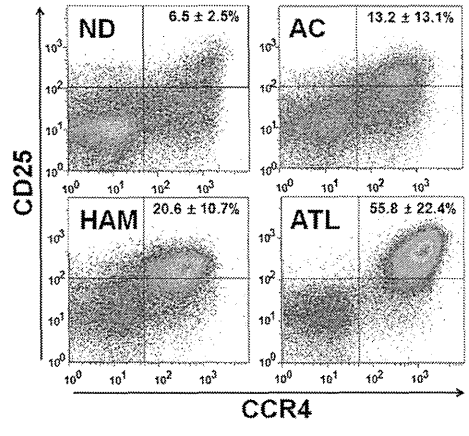

Figure 2. Representative sorting results of $CD4^+CD25^+CCR4^+$ cells. After labeling with anti-CD3-FITC, anti-CD4-Cy7, anti-CD25-APC, and anti-CCR4-PE, the $CD3^+CD4^+CD25^+CCR4^+$ fraction was isolated. The averaged content ± standard deviation (%) of $CD25^+CCR4^+$ cells out of $CD3^+CD4^+$ cells was calculated for each clinical group and is displayed in the upper right section of the panels.

From bloodjournal.hematologylibrary.org at University of Tokyo on April 30, 2014. For personal use only.

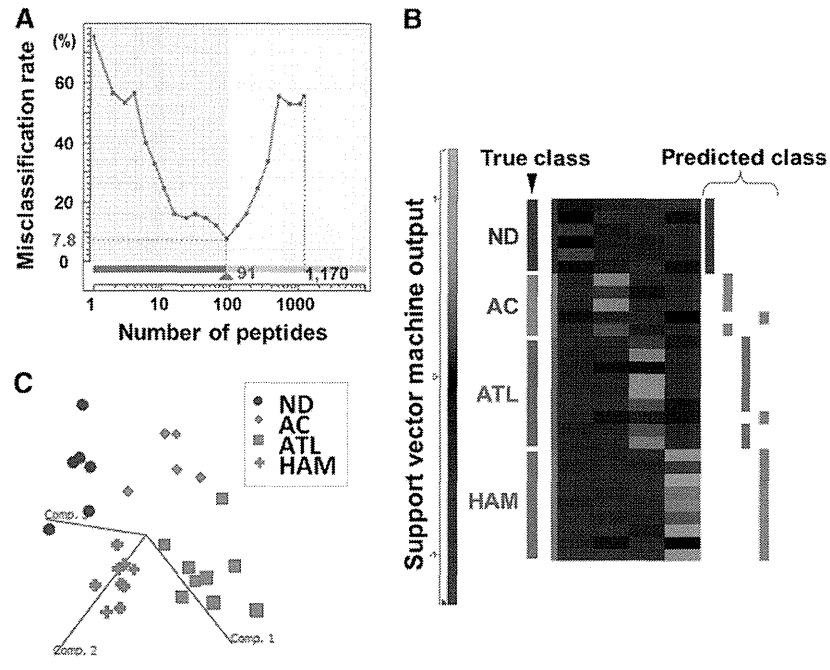

Figure 3. Statistical extraction of candidate therapeutic targets. The 14 064 nonredundant peptides detected were subjected to a 4-group Kruskal-Wallis test (ND, AC, ATL, and HAM), resulting in identification of 1170 first candidates ($P < .01$). ND, normal donors; AC, asymptomatic carriers. (A) Next, the Expressionist ranking method further narrowed down the candidates to 91 peptides based on SVM-REF so that the misclassification rate in the cross-validation test became minimum, 7.8%. (B) The predicted classification result by leave-one-out cross-validation test. The 27 out of 29 cases were successfully classified into the true classes. (C) The three-dimensional plot shows the additional assessment for the classification power of 91 classifiers by principal component analysis. Comp. 1 to 3 indicate principal components 1 to 3.

enriching pathogenic cells for rigorous quantitative biomarker discovery.

An accurately adjusted number of $CD4^+CD25^+CCR4^+$ cells from 29 cases were digested with trypsin and subjected to LC/MS/MS analysis individually. Because recent mass spectrometers often deal with data on the order of hundreds of megabytes per sample, it has been considered almost impossible to calculate a data set larger than a gigabyte from large-scale clinical samples on desktop computers. Hence, we constructed a proteomics server equipped with a 12-core central processing unit, 36 SAS hard disks, and 192-GB physical memories driving the Expressionist, which was designed to combine the database module, the data processing module, and the statistical analysis module into a single integrative platform for genomics, proteomics, and metabolomics. The detailed work flow for data processing and quantification for 29 LC/MS/MS raw data was described in the "Materials and methods" and is illustrated in supplemental Figure 1. Finally, 68 454 nonredundant peaks were detected and grouped into 37 143 isotopic clusters, or molecules. As tryptic peptides should appear as multivalent ions in electrospray ionization mass spectra, 23 079 singly charged ions were removed, resulting in utilization of 14 064 peptide signals for further statistical selection of diagnostic markers.

Statistical identification of candidate diagnostic markers for ATL

A stepwise statistical extraction was employed for the effective identification of proteins, which demonstrated specific up- or downregulation in the ATL group. In the first stage, a 4-group Kruskal-Wallis test was performed to roughly extract the candidates showing a significantly distinct expression level among 4 clinical groups. Here we set the cutoff line at $P < .01$ and obtained 1170 first candidate peptides simply because the isolated peptide set using this criterion showed the best performance in the following prediction model.

Next, we selected the final candidates by the support vector machine–recursive feature elimination algorithm in the Expressionist Analyst module. Support vector machine–recursive feature elimination is a candidate elimination method based on SVM, which enabled us to improve the classification outputs by selecting the best-performing peptide set among initially provided candidates.[11] As a result, a combination of 91 peptides showed the lowest misclassification rate (7.78%) in a leave-one-out cross-validation test (Figure 3A-B). To evaluate the classification efficiency of 91 selected candidates, the principal component analysis was performed. Figure 3C shows the three-dimensional plot of 29 clinical samples based on the 3 best-explainable components, which illustrated statistically clear segregation among the 4 clinical groups. These assessments indicated that the 91 peptides should be a sufficient set of classifiers that closely associated with the pathological characteristics of the 4 clinical groups.

Based on an independently constructed 6279-protein identification database for $CD4^+CD25^+CCR4^+$ cells using 2D LC/MS/MS (see details in "Materials and methods"), 19 peptides among the 91 candidate peptides were successfully assigned to 17 proteins listed in Table 1. The mass spectrometric quantification profiles for the 19 peptides are also shown in Figure 4 (box plots).

Recovering CAN2 activity induced cell death in ATL cells

Our diagnostic marker discovery for ATL identified an enzyme-substrate pair, CAN2 and SPTA2, which demonstrated significantly aberrant expression level in ATL patients (Figure 4). Interestingly, the intensities of the 2 proteins in 27 screening cases (without 2 statistical outliers in Figure 4) showed a clearly inverse correlation ($R^2 = 0.395$, Figure 5A). To examine whether CAN2 downregulation and/or SPTA2 upregulation might be essential for the growth of ATL cells, the enzymatic activity of CAN2 was rescued by overexpressing the constitutively active form of CAN2 (Δ_{19}CAN2) in 3 ATL cell lines, SO-4, KOB, and KK1. After 36 hours of transfection, significant inhibition of cell proliferation (Figure 5B) and induction of sub-G1 transition was observed by activation of CAN2 in 3 ATL cells, but not in 4 non-ATL leukemia cell lines (Figure 5C). Furthermore, overexpression of Δ_{19}CAN2 drastically attenuated the expression level of SPTA2 in the ATL cell

From bloodjournal.hematologylibrary.org at University of Tokyo on April 30, 2014. For personal use only.

Table 1. List of 17 protein classifiers for categorization of normal donors, asymptomatic carriers, HAM/TSP, and ATL

Accession	Protein name	*P* value (Kruskal-Wallis test)	m/z	RT	Charge	Peptide score	Identity or homology threshold	Sequence
LPPL	Eosinophil lysophospholipase	2.3.E-03	409.722	47.4	2	36.3	27	MVQVWR
CH10	Heat shock 10-kDa protein, mitochondrial	2.5.E-03	430.721	40.6	2	26.2	21	GGIMLPEK
PRG2	Bone marrow proteoglycan	2.4.E-03	528.271	64.6	2	31.6	28	RLPFICSY
MOES	Moesin	8.1.E-04	532.253	26.8	2	46.2	29	EKEELMER
MNDA	Myeloid cell nuclear differentiation antigen	9.4.E-03	647.863	69.1	2	67.3	24	SLLAYDLGLTTK
GLYM	Serine hydroxymethyltransferase, mitochondrial	8.7.E-04	408.551	21.6	3	31.1	18	HADIVTTTTHK
PTMS	Parathymosin	9.7.E-04	453.875	17.8	3	41.2	25	AAEEEDEADPKR
TPIS	Triosephosphate isomerase	9.1.E-03	472.266	71.0	3	54.0	28	QSLGELIGTLNAAK
HSP71	Heat shock 70-kDa protein 1A/1B	9.7.E-03	563.307	65.5	3	93.8	21	IINEPTAAAIAYGLDR
CD6	T-cell differentiation antigen CD6	7.7.E-03	592.306	37.8	3	62.7	22	VLCQSLGCGTAVERPK
ANXA1	Annexin A1	4.4.E-04	612.347	61.5	3	57.0	17	RKGTDVNVFNTILTTR
ANXA6	Annexin A6	2.3.E-03	669.017	70.9	3	54.7	16	AMEGAGTDEKALIEILATR
SPTA2	Spectrin α chain, brain	5.4.E-03	409.718	28.8	2	42.7	30	EAGSVSLR
GLYM	Serine hydroxymethyltransferase, mitochondrial	1.1.E-03	428.240	57.0	2	42.8	27	SGLIFYR
DRB1s	HLA class II histocompatibility antigen, DRB1-1, 4, 10, 11, 13, 15, 16 β chain	1.0.E-02	478.216	25.8	2	55.9	25	AAVDTYCR
CAN2	Calpain-2 catalytic subunit	2.4.E-03	483.253	54.0	2	66.6	29	SDTFINLR
STAT1	Signal transducer and activator of transcription 1-α/β	7.3.E-03	486.290	21.7	2	39.1	29	KILENAQR
PRG2	Bone marrow proteoglycan	9.4.E-04	497.742	49.2	2	31.6	27	FQWVDGSR
CXCL7	Platelet basic protein	1.3.E-03	528.761	43.1	2	51.7	28	ICLDPDAPR

line SO-4 (Figure 5D), but not in the non-ATL leukemia cell line Jurkat (Figure 5E). On the other hand, an additional cell proliferation assay using siRNA against *SPTAN1* revealed that reduction of SPTA2 was not sufficient for the induction of cell death for ATL cells (supplemental Figures 3 and 4).

In addition, 3 proteins (PTMS, CH10, and GLYM) were also found to be upregulated in ATL cells. To address the roles of these proteins, a cell proliferation assay was conducted using 3 ATL cell lines treated with siRNAs against *PTMS*, *HSPE1* (gene symbol of CH10), or *SHMT2* (gene symbol of GLYM) (supplemental Figure 4). As a result, suppression of the *SHMT2* gene induced significant growth inhibition for all 3 ATL cell lines. Although si*HSPE1*-treated KOB cells showed a statistically significant decrease in cell growth rate, si*HSPE1* and si*PTMS* had only partial

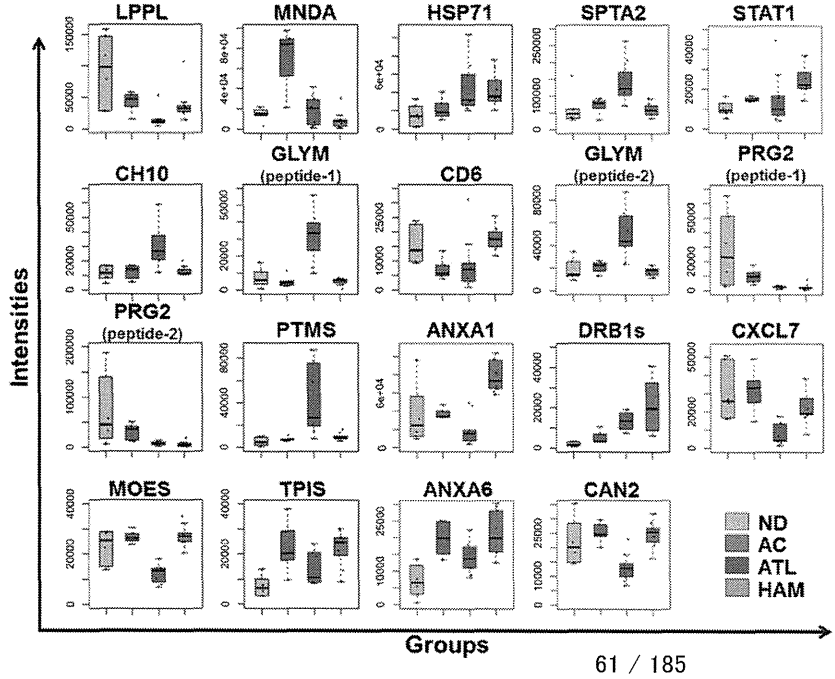

Figure 4. Summary of quantitative features for the 17 protein classifiers identified. The 19 box plots (see Table 1 for protein names) show the results of mass spectrometric quantification and protein identification. We finally identified 19 peptides out of 91 candidates in Figure 3, which were assigned to 17 proteins. Proteins identified from 2 distinct peptides were shown as GLYM (peptides 1 and 2) or PRG2 (peptides 1 and 2). The y-axis indicates normalized relative intensity of peptides in mass spectrometric data. ND, normal donors; AC, asymptomatic carriers.

From bloodjournal.hematologylibrary.org at University of Tokyo on April 30, 2014. For personal use only.

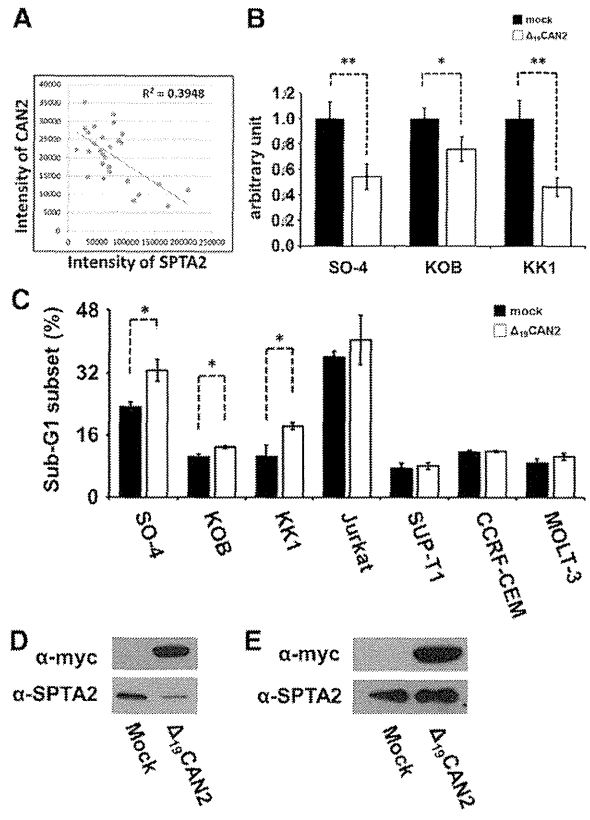

Figure 5. Rescue of CAN2 activity induced cell death in ATL cells. (A) Correlation between CAN2 and SPTA2 expression level in 27 cases. (B) Cell proliferation was measured by MTT assay on SO-4, KOB, and KK1 cells 36 hours after transfection of mock vector or Δ_{19}CAN2. $^{*}P < .05$; $^{**}P < .01$ by Student *t* test. (C) Overexpression of Δ_{19}CAN2 significantly accelerated cell death in 3 ATL (SO-4, KOB, and KK1) and 4 non-ATL (Jurkat, SUP-T1, CCRF-CEM, and MOLT-3) cell lines. $^{**}P < .05$ by Student *t* test. The drastic attenuation of SPTA2 expression was observed after transfection of Δ_{19}CAN2 in SO-4 cells (D), but not in Jurkat cells (E). The immunoblot of anti-myc tag confirmed the expression of exogenous Δ_{19}CAN2.

or no effects on proliferation of ATL cell lines. To further confirm whether the overexpression of SPTA2, PTMS, CH10, or GLYM protein would be an ATL-specific molecular signature, the expression levels of these proteins in 8 clinical samples were evaluated by the mass spectrometric quantification technology MRM (supplemental Figures 5 and 6). Expression of SPTA2, GLYM, and CH10 in cells derived from ATL patients was significantly higher than that in cells derived from HAM/TSP patients. The level of PTMS also showed a clearly increasing tendency in the ATL patient group. Taken together, these results suggested that the deprivation of CAN2 activity and upregulation of GLYM in HTLV-1–infected T cells might have a key role at the onset or progression of ATL.

Discussion

In the past decade, proteomics technologies have developed dramatically for the purpose of obtaining more and more comprehensive and sensitive proteome maps in cells or clinical specimens. The performance of mass spectrometers in particular has exhibited remarkable progress; however, as for sensitivity and throughput, it has still been difficult to identify biomarkers from crude samples including body fluids or total cell lysate. A major reason could be that the range of protein concentration in the analyte is indeed much larger than the dynamic range of recent mass spectrometers.[12] The other essential factor to be improved for clinical proteomics is the capacity of the bioinformatics platform to allow analysis of a sufficient number of clinical samples in order to statistically overcome the significant individual variability.[13]

Concerning the first issue, we previously developed and applied various focused proteomic applications targeting molecular biochemical features including glycan structure biomarkers[14-16] and low-molecular-weight peptide biomarkers.[17] The preenrichment of subproteome fractions effectively reduces the complexity of crude samples and allowed us to identify potential serum cancer biomarkers successfully. Through our previous knowledge, we provide an approach for investigating infectious diseases by employing virus-infected cell-focused proteomics. In addition to HTLV-1, for instance, isolation of HIV-infected cells is highly desired because the frequency of these cells in AIDS patients' PBMCs is ~1 out of 10^4 to 10^5 cells.[18] Actually, we successfully demonstrated the effect of HTLV-1–infected cell isolation on the elimination of individual variability (Figure 2, supplemental Figure 2) and reliable identification of disease state–associated proteins (Figures 4 and 5). We further showed the potential of the next-generation bioinformatics platform Expressionist to remove the constraint on the capacity of data size acquired from high-end mass spectrometers. Expressionist covered whole discovery steps from processing of raw mass spectrometer data to statistical analyses (Figures 1 and 3, and supplemental Figure 1) and, importantly, could perform quantification analysis using a basically unlimited number of clinical samples. Hence, in parallel with the development of mass spectrometers, high-specification and inexpensive OMICS server systems are necessary for future diagnostic marker and therapeutic target discoveries using hundreds or thousands of clinical specimens.

In this study, we focused on the $CD4^+CD25^+CCR4^+$ T-cell subpopulation in which T helper 2, T helper 17, and regulatory T (Treg) cells were mainly involved.[10] The purpose for which we used this subset was to technically enrich the preferential viral reservoir cells and to strengthen reliability of screening results. However, investigating proteome behaviors of these subtypes in HTLV-1–associated diseases is also important physiologically because it has been frequently reported that deregulated Treg plays significant roles in pathogenesis of ATL and HAM/TSP. Indeed, aberrant proliferation of Treg cells is considered the main cause of immunodeficiency in ATL patients because of their innate immunosuppressive functions,[19] whereas abnormal production of interferon γ from infected Treg cells might induce chronic spinal inflammation in HAM/TSP patients.[20] Given the list of our 17 classifier proteins, activation of signal transducer and activator of transcription 1-α/β is the well-known key factor for HAM/TSP,[21] whereas upregulation of heat shock 70-kDa protein 1A/1B, CH10, and PTMS were reported in many other types of tumors.[22-24] The association of these 4 proteins with the etiology of HAM/TSP and ATL would be evident according to the previous work, supporting that our other candidates might similarly have a direct impact on the transformation of Treg cells after infection of HTLV-1. Particularly, the specific upregulation of GLYM in ATL cells represents the first evidence that excessive folate metabolism might be essential for the progression or survival of ATL cells because GLYM is a fundamental enzyme catalyzing the supply of glycine accompanying the conversion of tetrahydrofolate to 5,10-methylenetetrahydrofolate.[25] Indeed, the suppression of GLYM expression, which was confirmed to be upregulated in ATL patients, resulted in significant reduction of cell growth. This observation suggests that diminishing GLYM

From bloodjournal.hematologylibrary.org at University of Tokyo on April 30, 2014. For personal use only.

expression or enzyme activity could be a promising strategy for molecular-targeting treatment of ATL. Together with the down-regulation of CAN2 in the ATL cells shown in Figure 5, the proteins listed in Table 1 could provide the molecular basis for not only interpretation of physiological mechanisms in ATL or HAM/TSP but also development of novel therapeutic agents for HTLV-1–associated diseases.

CAN2 belongs to a Ca^{2+}-regulated cytosolic cysteine protease family, which includes 14 calpain isoforms.[26] The enzymatic activity of calpain is implicated in diverse physiological processes, such as cytoskeletal remodeling, cellular signaling, and apoptosis.[26] As an example of a spectrin-mediated apoptosis pathway, it was reported that CAN2 produced SPTA2 breakdown products following traumatic brain injury.[27] Because SPTA2 interacts with calmodulin and constructs the membrane cytoskeletons, its breakdown is considered a process of membrane structural changes during cell death.[28,29] This fact is concordant with our finding in ATL, suggesting that accumulation of SPTA2 in ATL cells can be attributed to the suppression of CAN2 expression and contribute to circumvent apoptosis. In the analysis of basal levels of CAN2 and SPTA2 in 7 cell lines (supplemental Figure 7), 3 ATL cell lines showed endogenous expression of CAN2 and moderate levels of SPTA2. On the other hand, 4 non-ATL leukemia cells demonstrated very high expression of SPTA2 and undetectable levels of CAN2. Although we found the downregulation of CAN2 and accumulation of SPTA2 in ATL cells, this tendency might be more distinctive in HTLV-1 (–) leukemia cells. Taken together, even though the expression level of CAN2 was indeed suppressed in ATL cells, the CAN2-SPTA2 apoptotic pathway itself might remain normal. In contrast, this pathway was considered to be impaired at multiple stages in HTLV-1 (–) leukemia cells because CAN2 expression was completely diminished (supplemental Figure 7) and overexpression of CAN2 could not reactivate the CAN2-SPTA2 apoptotic pathway (Figure 5B-E). In these cells, not only genetic downregulation of CAN2 but also inhibition of CAN2 enzymatic activity might be involved in the carcinogenesis.

In conclusion, comprehensive proteomic profiling of HTLV-1–infected T cells provided 17 disease-associated signature proteins, which have great potential for future clinical use as diagnostic biomarkers. As we described regarding the relationship between the CAN2-SPTA2 pathway and ATL phenotypes, further individual functional analyses will contribute to understanding the detailed molecular mechanisms involved in the onset or progression of HAM/TSP and ATL.

Acknowledgments

The authors thank Dr Hiroyuki Sorimachi for kindly providing pBlueBacIIICAPN2 vector.

This work was supported by Research on Measures for Intractable Diseases, the Ministry of Health Labour and Welfare Japan.

Authorship

Contribution: M.I. and K.U. designed the study, performed experiments, analyzed results, and wrote the manuscript; A.T. and N.S. performed experiments; N.A., T.S., A.U., and Y.Y. collected the clinical samples and performed flow cytometric experiments; Y.N. and H.N. revised the manuscript; and all authors discussed the results and commented on the manuscript.

Conflict-of-interest disclosure: The authors declare no competing financial interests.

Correspondence: Koji Ueda, Laboratory for Biomarker Development, Center for Genomic Medicine, RIKEN, General Research Building 6F, Institute of Medical Science, 4-6-1 Shirokanedai, Minato-ku, Tokyo, Japan, 1088639; e-mail: k-ueda@riken.jp.

References

1. Yamashita M, Ido E, Miura T, Hayami M. Molecular epidemiology of HTLV-I in the world. *J Acquir Immune Defic Syndr Hum Retrovirol.* 1996;13(suppl 1):S124-S131.
2. Asquith B, Zhang Y, Mosley AJ, et al. In vivo T lymphocyte dynamics in humans and the impact of human T-lymphotropic virus 1 infection. *Proc Natl Acad Sci U S A.* 2007;104(19):8035-8040.
3. Sakashita A, Hattori T, Miller CW, et al. Mutations of the p53 gene in adult T-cell leukemia. *Blood.* 1992;79(2):477-480.
4. Beltran B, Quiñones P, Morales D, Cotrina E, Castillo JJ. Different prognostic factors for survival in acute and lymphomatous adult T-cell leukemia/lymphoma. *Leuk Res.* 2011;35(3):334-339.
5. Ishida T, Joh T, Uike N, et al. Defucosylated anti-CCR4 monoclonal antibody (KW-0761) for relapsed adult T-cell leukemia-lymphoma: a multicenter phase II study. *J Clin Oncol.* 2012; 30(8):837-842.
6. Semmes OJ, Cazares LH, Ward MD, et al. Discrete serum protein signatures discriminate between human retrovirus-associated hematologic and neurologic disease. *Leukemia.* 2005;19(7):1229-1238.
7. Kirk PD, Witkover A, Courtney A, et al. Plasma proteome analysis in HTLV-1-associated myelopathy/tropical spastic paraparesis. *Retrovirology.* 2011;8:81.
8. Rahman S, Quann K, Pandya D, Singh S, Khan ZK, Jain P. HTLV-1 Tax mediated downregulation of miRNAs associated with chromatin remodeling factors in T cells with stably integrated viral promoter. *PLoS ONE.* 2012;7(4):e34490.
9. Polakowski N, Gregory H, Mesnard JM, Lemasson I. Expression of a protein involved in bone resorption, Dkk1, is activated by HTLV-1 bZIP factor through its activation domain. *Retrovirology.* 2010;7:61.
10. Yamano Y, Araya N, Sato T, et al. Abnormally high levels of virus-infected IFN-gamma+ CCR4+ CD4+ CD25+ T cells in a retrovirus-associated neuroinflammatory disorder. *PLoS ONE.* 2009;4(8):e6517.
11. Oh JH, Gao J, Nandi A, Gurnani P, Knowles L, Schorge J. Diagnosis of early relapse in ovarian cancer using serum proteomic profiling. *Genome Inform.* 2005;16(2):195-204.
12. Anderson NL, Anderson NG. The human plasma proteome: history, character, and diagnostic prospects. *Mol Cell Proteomics.* 2002;1(11): 845-867.
13. Nordon IM, Brar R, Hinchliffe RJ, Cockerill G, Thompson MM. Proteomics and pitfalls in the search for potential biomarkers of abdominal aortic aneurysms. *Vascular.* 2010;18(5):264-268.
14. Ueda K, Katagiri T, Shimada T, et al. Comparative profiling of serum glycoproteome by sequential purification of glycoproteins and 2-nitrobenzensulfenyl (NBS) stable isotope labeling: a new approach for the novel biomarker discovery for cancer. *J Proteome Res.* 2007;6(9): 3475-3483.
15. Ueda K, Fukase Y, Katagiri T, et al. Targeted serum glycoproteomics for the discovery of lung cancer-associated glycosylation disorders using lectin-coupled ProteinChip arrays. *Proteomics.* 2009;9(8):2182-2192.
16. Ueda K, Takami S, Saichi N, et al. Development of serum glycoproteomic profiling technique; simultaneous identification of glycosylation sites and site-specific quantification of glycan structure changes. *Mol Cell Proteomics.* 2010;9(9): 1819-1828.
17. Ueda K, Saichi N, Takami S, et al. A comprehensive peptidome profiling technology for the identification of early detection biomarkers for lung adenocarcinoma. *PLoS ONE.* 2011;6(4): e18567.
18. Bahbouhi B, al-Harthi L. Enriching for HIV-infected cells using anti-gp41 antibodies indirectly conjugated to magnetic microbeads. *Biotechniques.* 2004;36(1): 139-147.
19. Matsubar Y, Hori T, Morita R, Sakaguchi S, Uchiyama T. Delineation of immunoregulatory properties of adult T-cell leukemia cells. *Int J Hematol.* 2006;84(1):63-69.
20. Best I, López G, Verdonck K, et al. IFN-gamma production in response to Tax 161-233, and frequency of CD4+ Foxp3+ and Lin HLA-

From bloodjournal.hematologylibrary.org at University of Tokyo on April 30, 2014. For personal use only.

DRhigh CD123+ cells, discriminate HAM/TSP patients from asymptomatic HTLV-1-carriers in a Peruvian population. *Immunology.* 2009; 128(1, pt 2):e777-e786.

21. Nakamura N, Fujii M, Tsukahara T, et al. Human T-cell leukemia virus type 1 Tax protein induces the expression of STAT1 and STAT5 genes in T-cells. *Oncogene.* 1999;18(17): 2667-2675.
22. Alaiya AA, Al-Mohanna M, Aslam M, et al. Proteomics-based signature for human benign prostate hyperplasia and prostate adenocarcinoma. *Int J Oncol.* 2011;38(4): 1047-1057.
23. Cappello F, Rappa F, David S, Anzalone R, Zummo G. Immunohistochemical evaluation of PCNA, p53, HSP60, HSP10 and MUC-2 presence and expression in prostate carcinogenesis. *Anticancer Res.* 2003;23(2B): 1325-1331.
24. Letsas KP, Vartholomatos G, Tsepi C, Tsatsoulis A, Frangou-Lazaridis M. Fine-needle aspiration biopsy-RT-PCR expression analysis of prothymosin alpha and parathymosin in thyroid: novel proliferation markers? *Neoplasma.* 2007; 54(1):57-62.
25. Anderson DD, Quintero CM, Stover PJ. Identification of a de novo thymidylate biosynthesis pathway in mammalian mitochondria. *Proc Natl Acad Sci USA.* 2011;108(37): 15163-15168.
26. Storr SJ, Carragher NO, Frame MC, Parr T, Martin SG. The calpain system and cancer. *Nat Rev Cancer.* 2011;11(5):364-374.
27. Liu MC, Akle V, Zheng W, et al. Comparing calpain- and caspase-3-mediated degradation patterns in traumatic brain injury by differential proteome analysis. *Biochem J.* 2006;394(pt 3): 715-725.
28. Wallis CJ, Wenegieme EF, Babitch JA. Characterization of calcium binding to brain spectrin. *J Biol Chem.* 1992;267(7): 4333-4337.
29. Liu X, Van Vleet T, Schnellmann RG. The role of calpain in oncotic cell death. *Annu Rev Pharmacol Toxicol.* 2004;44: 349-370.

NEOPLASIA
www.neoplasia.com

Volume 15 Number 11 November 2013 pp. 1251–1261 **1251**

The Oncogenic Polycomb Histone Methyltransferase EZH2 Methylates Lysine 120 on Histone H2B and Competes Ubiquitination[1,2]

Masaharu Kogure[*,†], Masashi Takawa[*], Vassiliki Saloura[‡], Kenbun Sone[‡], Lianhua Piao[‡], Koji Ueda[§], Reem Ibrahim[*], Tatsuhiko Tsunoda[¶], Masanori Sugiyama[†], Yutaka Atomi[†], Yusuke Nakamura[‡] and Ryuji Hamamoto[*,‡]

[*]Laboratory of Molecular Medicine, Human Genome Center, Institute of Medical Science, The University of Tokyo, Minato-ku, Tokyo, Japan; [†]Department of Surgery, Kyorin University School of Medicine, Mitaka, Tokyo, Japan; [‡]Section of Hematology/Oncology, Department of Medicine, The University of Chicago, Chicago, IL; [§]Laboratory for Biomarker Development, RIKEN, Minato-ku, Tokyo, Japan; [¶]Laboratory for Medical Informatics, RIKEN, Yokohama, Kanagawa, Japan

Abstract

The histone methyltransferase enhancer of zeste 2 (EZH2) is known to be a polycomb protein homologous to *Drosophila* enhancer of zeste and catalyzes the addition of methyl groups to histone H3 at lysine 27 (H3K27). We previously reported that EZH2 was overexpressed in various types of cancer and plays a crucial role in the cell cycle regulation of cancer cells. In the present study, we demonstrated that EZH2 has the function to monomethylate lysine 120 on histone H2B (H2BK120). EZH2-dependent H2BK120 methylation in cancer cells was confirmed with an H2BK120 methylation-specific antibody. Overexpression of EZH2 significantly attenuated the ubiquitination of H2BK120, a key posttranslational modification of histones for transcriptional regulation. Concordantly, knockdown of EZH2 increased the ubiquitination level of H2BK120, suggesting that the methylation of H2BK120 by EZH2 may competitively inhibit the ubiquitination of H2BK120. Subsequent chromatin immunoprecipitation–Seq and microarray analyses identified downstream candidate genes regulated by EZH2 through the methylation of H2BK120. This is the first report to describe a novel substrate of EZH2, H2BK120, unveiling a new aspect of EZH2 functions in human carcinogenesis.

Neoplasia (2013) 15, 1251–1261

Introduction

Enhancer of zeste 2 (EZH2) belongs to polycomb group (PcG) protein and is a member of the polycomb repressor complex 2 (PRC2) that methylates histone H3 at lysine 27 (H3K27), a repressive mark that maintains epigenetic silencing of genes. EZH2 is active only when it is associated with other PRC2 core components embryonic ectoderm development (EED), suppressor of zesta 12 homolog (SUZ12), and RBBP4 (retinoblastoma binding protein 4; RbAp48). The PRC2 complex is responsible for repressing of a large number of genes that are essential for development and differentiation. PcG proteins specify positional information such as antero-posterior patterning, through activating or repressing the stable state of *Hox* gene expression. In addition to these well-established functions in embryonic developmen

Address all correspondence to: Ryuji Hamamoto, PhD, Section of Hematology/Oncology, Department of Medicine, The University of Chicago, 5835 S Cottage Grove Ave, MC2115 Chicago, IL 60637. E-mail: rhamamoto@medicine.bsd.uchicago.edu

[1]This work was supported by a grant-in-aid for Young Scientists (22681030) from the Japan Society for the Promotion of Science and Project for Development of Innovation Research on Cancer Therapeutics.

[2]This article refers to supplementary materials, which are designated by Tables W1 to W6 and Figures W1 to W8 and are available online at www.neoplasia.com.

Received 6 August 2013; Revised 16 October 2013; Accepted 21 October 2013

Copyright © 2013 Neoplasia Press, Inc. All rights reserved 1522-8002/13/$25.00
DOI 10.1593/neo.131436

a series of studies have suggested that PcG proteins may influence both Hox-dependent and Hox-independent downstream pathways that control cell proliferation. We previously demonstrated that EZH2 was overexpressed in various types of human cancer, and its overexpression was correlated with a negative outcome in patients with non-small-cell lung carcinoma (NSCLC) after surgical resection [1]. Importantly, selective inhibitors targeting EZH2 have recently developed, and they showed the growth-inhibitory effects of tumor cells [2,3]. These results indicate that EZH2 is a promising target for development of cancer treatment. Although the transcriptional regulation mechanism by EZH2 through methylation of H3K27 has already been well studied, additional functions of EZH2 through methylation of other substrates still remained unclear.

Posttranslational modification of the four core histones is a commonly important process during the regulation of gene activation and repression. Histone modifications are also involved in various cellular processes, including DNA damage response and alternative splicing. Histone H2B is one of the four core histones involved in chromatin formation in eukaryotic cells. Featuring a main globular domain and a long N-terminal tail, H2B is involved with the structure of the nucleosomes of the "beads on a string" structure. As posttranslational modifications of histone H2B, acetylation, phosphorylation, ubiquitination, and sumoylation have already been reported [4–8]. Among these modifications, histone H2B lysine 120 monoubiquitination (H2BK120ub) is known to be a key histone modification that plays critical roles in the transcriptional regulation as well as higher order chromatin organization in many species [9]. H2BK120ub is associated with a high level of gene expression in human cells [10]. This histone modification is also induced after DNA damage and has been indicated to have a critical role in the maintenance of replication-dependent histone mRNA 3′-end processing [11]. The human ring finger protein 20 (RNF20)/RNF40 complex is the major H2B E3 ligase [12]. At the structure level, monoubiquitination of H2BK120 interferes with compaction of chromatin, resulting in open chromatin fibers that display greater accessibility to transcription factors and their coregulators [13]. Several studies implicate H2BK120ub in developmental processes including that optical embryonic stem cell differentiation requires dynamic changes in H2B ubiquitination patterns in a timely and well-coordinated manner [7,14].

In this study, we identified that the histone methyltransferase EZH2 methylates H2BK120 and competitively inhibits ubiquitination of this lysine residue in cancer cells. This is the first report describing identification of histone H2B as a novel substrate of EZH2 and demonstrating the biologic significance of the histone H2B methylation in human carcinogenesis.

Materials and Methods

Cell Culture

CCD-18Co, HFL1, HCT116, SW480, RT4, MCF7, HeLa, and 293T cells were from American Type Culture Collection (ATCC, Manassas, VA) in 2001 and 2003 and tested and authenticated by DNA profiling for polymorphic short tandem repeat (STR) markers. SBC5, RERF-LC-AI, and Alexander cells were from Japanese Collection of Research Bioresources (JCRB) in 2001 and tested and authenticated by DNA profiling for polymorphic STR markers.

All cell lines were grown in monolayers in the following appropriate media: Dulbecco's modified Eagle's medium for RERF-LC-AI/ and 293T cells; Eagle's minimum essential medium for CCD-18Co, SBC5, MCF7, and HeLa cells; Leibovitz's L-15 for SW480 cells; McCoy's 5A medium for RT4 and HCT116 cells; RPMI 1640 medium for Alexander cells; and Ham's F-12K Medium for HFL1 cells supplemented with 10% FBS and 1% antibiotic/antimycotic solution (Life Technologies, Carlsbad, CA). All cells were maintained at 37°C in humid air with 5% CO_2 condition (CCD-18Co, HFL1, SBC5, RERF-LC-AI, HCT116, SW480, Alexander, RT4, MCF7, HeLa, and 293T) or without CO_2 (SW480). Cells were transfected with FuGENE 6 (Roche Applied Science, Penzberg, Germany) according to the manufacturer's protocols [15,16].

Antibodies

The following primary antibodies were deployed: anti-FLAG [M2; Sigma-Aldrich (St Louis, MO); dilution used in Western blot (WB): 1:5000, immunocytochemistry (ICC): 1:100], anti-EZH2 (NCL-L-EZH2; Leica, Solms, Germany; dilution used in WB: 1:500), anti–histone H2B (ab1790; Abcam, Cambridge, United Kingdom; dilution used in WB: 1:1000), anti–Ubiquityl-Histone H2BK120 (No. 5546; Cell Signaling Technology, Danvers, MA; dilution used in WB: 1:1000, ICC: 1:50), and anti–actin, beta (ACTB; No. 4967; Cell Signaling Technology; dilution used in WB: 1:1000). The anti–K120-methylated H2B antibody [Sigma-Aldrich; dilution used in WB: 1:500, ICC: 1:100, immunohistochemistry (IHC): 1:5000] was produced in rabbit immunized with a synthetic peptide.

In Vitro *Methyltransferase Assay*

For the *in vitro* methyltransferase assay, recombinant Histone H2B (No. 14-410; Millipore, Billerica, MA) was incubated with EZH2 enzyme complex (No. 51004; BPS Bioscience, San Diego, CA) using 2 μCi of S-adenosyl-L-[methyl-^{3}H] methionine (Perkin-Elmer, Waltham, MA) as the methyl donor in a mixture of 10 μl of methylase activity buffer (50 mM Tris-HCl at pH8.8, 10 mM DTT, and 10 mM $MgCl_2$), for 1 hour at 30°C. Proteins were resolved on a 5% to 20% sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE) gel (Ready Gel; Bio-Rad Laboratories, Hercules, CA) and visualized by MemCode Reversible Stain (Thermo Fisher Scientific, Waltham, MA) and fluorography.

Mass Spectrometry

Histone H2B samples reacted with EZH2 were separated on SDS-PAGE and stained with Coomassie brilliant blue. The excised Histone H2B bands were reduced in 10 mM tris(2-carboxyethyl) phosphine (Sigma-Aldrich) with 50 mM ammonium bicarbonate (Sigma-Aldrich) for 30 minutes at 37°C and alkylated in 50 mM iodoacetamide (Sigma-Aldrich) with 50 mM ammonium bicarbonate for 45 minutes in the dark at 25°C. Trypsin GOLD (Promega, Fitchburg, WI) solution was added with the enzyme-to-protein ratio at 1/50 (wt/wt) and incubated at 37°C for 16 hours. The resulting peptides were extracted from gel fragments and separated on a 0.1 × 200 mm homemade C_{18} column using 45-minute linear gradient from 2% to 35% acetonitrile in 0.1% formic acid with flow rate at 200 nl/min. The eluting peptides were analyzed with QSTAR Elite QqTOF system (AB Sciex, Framingham, MA) in the smart information-dependent acquisition mode of the Analyst QS software 2.0 (AB Sciex). The acquired mass spectrometry (MS) and tandem mass spectrometry (MS/MS) peak lists were analyzed with in-house Mascot server version

2.4.01 (Matrix Science, Boston, MA) to identify peptide sequences. We finally accepted the assigned peptides with expectation value less than 0.05 as the positive identification in Mascot Database search.

Western Blot Analysis

Samples were prepared from the cells lysed with CelLytic M Cell Lysis Reagent (Sigma-Aldrich) [17], and whole-cell lysates or immunoprecipitation (IP) products were transferred to nitrocellulose membrane. Protein bands were detected by incubating with HRP-conjugated antibodies (GE Healthcare, Little Chalfont, United Kingdom) and visualizing with Enhanced Chemiluminescence (GE Healthcare). We declare that our blots were evenly exposed in each membrane and that the blots are not clopped to the bands.

Immunocytochemistry

Cultured cells were fixed in 4% paraformaldehyde in 0.1 M phosphate buffer (pH 7.4) at room temperature for 30 minutes, permeabilized in 0.1% Triton X-100 (Sigma-Aldrich) for 3 minutes on ice, and blocked with BLOCK ACE (Yukijirushi, Sapporo, Japan) for 1 hour at room temperature. Fixed cells were incubated with primary antibodies overnight at 4°C. Then incubated with Alexa Fluor–conjugated second antibody (Life Technologies) and observed using a Leica confocal microscopy [18,19].

Immunohistochemical Analysis

The K120 monomethylation status of histone H2B in clinical tissues were examined by immunohistochemical analysis [1,17]. EnVision+ kit/HRP kit (Dako, Carpinteria, CA) was applied, and slides of paraffin-embedded lung tumor specimens were processed under high pressure (125°C, 30 seconds) in antigen-retrieval solution, high pH 9 (S2367; Dako), treated with peroxidase-blocking regent, and then treated with protein-blocking regent (K130, X0909; Dako). Tissue sections were incubated with a rabbit anti-H2BK120me1 polyclonal antibody, followed by HRP-conjugated secondary antibody (Dako). Antigen was visualized with substrate chromogen (Dako Liquid DAB Chromogen; Dako). Finally, tissue specimens were stained with Mayer haematoxylin (Hematoxylin QS; Vector Laboratories, Burlingame, CA) for 20 seconds to discriminate the nucleus from the cytoplasm. The intensity of H2BK120 methylation was evaluated using the following criteria: strong positive (scored as 2+), brown staining in >50% of tumor cells completely obscuring cytoplasm; weak positive (1+), any lesser degree of brown staining appreciable in tumor cell cytoplasm; and absent (scored as 0), no appreciable staining in tumor cells. Cases were accepted as strongly positive if two or more investigators independently defined them as such. Detailed clinical information of lung and colony tissues is described in Tables W1 and W2.

siRNA Transfection

Small interfering RNA (siRNA) oligonucleotide duplexes were purchased from Sigma-Aldrich for targeting the human *EZH2* transcript. siNegative control (siNC) was used as a control siRNA. The siRNA sequences are described in Table W3. siRNA duplexes (100 nM final concentration) were transfected into bladder and lung cancer cell lines with Lipofectamine 2000 (Life Technologies) [20].

Immunoprecipitation

Transfected 293T or SBC5 cells were lysed with CelLytic M Cell Lysis Reagent (Sigma-Aldrich) containing a complete protease inhibitor cocktail (Roche Applied Science). In a typical immunoprecipitation reaction, 300 μg of whole-cell extract was incubated with an optimum concentration of primary antibody. After the beads had been washed three times in 1 ml of TBS buffer (pH 7.6), proteins that bound to the beads were eluted by boiling in Lane Marker Reducing Sample Buffer (Thermo Fisher Scientific).

Chromatin Immunoprecipitation–Seq

Chromatin immunoprecipitation (ChIP) assays were performed using ChIP Assay kit (Millipore) according to the manufacturer's protocol. Briefly, the fragment of EZH2 and chromatin complexes was immunoprecipitated with anti-H2BK120me1 (Sigma-Aldrich) and anti-FLAG (M2; Sigma-Aldrich) antibodies 36 or 48 hours after transfection with siRNAs and pCAGGS-n3FC (Mock) or pCAGGS-n3FC-EZH2 (3× FLAG-EZH2) vectors, respectively. ChIP-DNA samples were end repaired, A tailed, and finally ligated with DNA adaptors. Subsequently, samples were size fractionated (175-225 bp) on a 2% tris-borate-EDTA (TBE) agarose gel. Fractionated ChIP-DNA samples were amplified using KAPA Library Preparation Kit (Kapa Biosystems, Woburn, MA). After a final polymerase chain reaction amplification step (18 cycles), the resulting DNA libraries were analyzed by an Agilent Technologies 2100 Bioanalyzer (Santa Clara, CA), and DNA libraries were sequenced on Illumina GAIIx (San Diego, CA).

ChIP-Seq Data Analysis

Sequencing data were acquired from the Illumina GAIIx sequencer [21]. ChIP-Seq reads were aligned using Burrows-Wheeler Aligner software (version 0.62-r126) to the human genome (University of California, Santa Cruz, CA; hg19) with default settings [22]. Burrows-Wheeler Aligner's default mismatch criteria for creating the mismatches was set up that mismatch allowance was 3%, which means that allowance of InDels (insert and deletion) was at least more than 3 bp in 100-bp reads. When same reads mapped to the same genomic position, a maximum of two reads mapping to the same position were used. For enriched-region (peak) identification (peak calling), we used the Model-based Analysis of ChIP-Seq (MACS; Harvard University, Boston, MA; version 1.4.1) algorithm [23,24]. We set up a band with 300 bp, *P* value of 1.00e-06, model fold between 10 and 30 (default) to call peaks representing enriched H2BK120me1 marks and EZH2-located regions. To prospect the peak detection of enhancer region, tags detected were aligned in each gene set within 10 kbp upstream from transcription start sites or 10 kbp downstream from transcription terminal sites. We used the statistical software DESeq (European Molecular Biology Laboratory, Hamburg, Germany), an R package, to normalize the number of tags detected in each gene region and calculated the ratio of samples and controls [25]. Normalized intensity value of genes was ordered as a heatmap using hierarchical clustering in Cluster 3.0 software (Stanford University, Stanford, CA).

Microarray Hybridization and Statistical Analysis

Microarray analysis to identify downstream genes was described previously [20,26,27]. Purified total RNA was labeled and hybridized onto Affymetrix GeneChip U133 Plus 2.0 oligonucleotide arrays (Affymetrix, Santa Clara, CA) according to the manufacturer's instructions. Probe signal intensities were normalized by robust multichip average (RMA) and Quantile (using R and Bioconductor).

ChIP Assay

ChIP assays were performed using ChIP Assay kit (17-295; Millipore) according to the manufacturer's protocol. Briefly, the fragment of K120-ubiquitinated H2B and chromatin complexes in HEK293 cells was immunoprecipitated with an anti–Ubiquityl-Histone H2BK120 antibody. After DNA fragments bound to K120-ubiquitinated H2B were eluted out, an aliquot was subjected to quantitative real-time polymerase chain reaction reactions. Protein A Agarose/Salmon Sperm DNA (16-157; Millipore) was used as a negative control. Primers were designed near the transcriptional start site of each gene, and their sequence information is shown in Table W4.

Results

EZH2 Methylates Lysine 120 on Histone H2B

To identify a novel substrate of EZH2, we conducted *in vitro* methyltransferase assays using recombinant histone proteins as well as nonhistone proteins that play important roles in human carcinogenesis. Consequently, we identified that EZH2 appeared to methylate histone H2B in addition to histone H3 (Figure W1). To validate this result, we performed *in vitro* methyltransferase assays using different amount of EZH2 enzymes and confirmed histone H2B methylation by EZH2 in a dose-dependent manner (Figure 1*A*). Subsequently, we conducted liquid chromatography (LC)-MS/MS

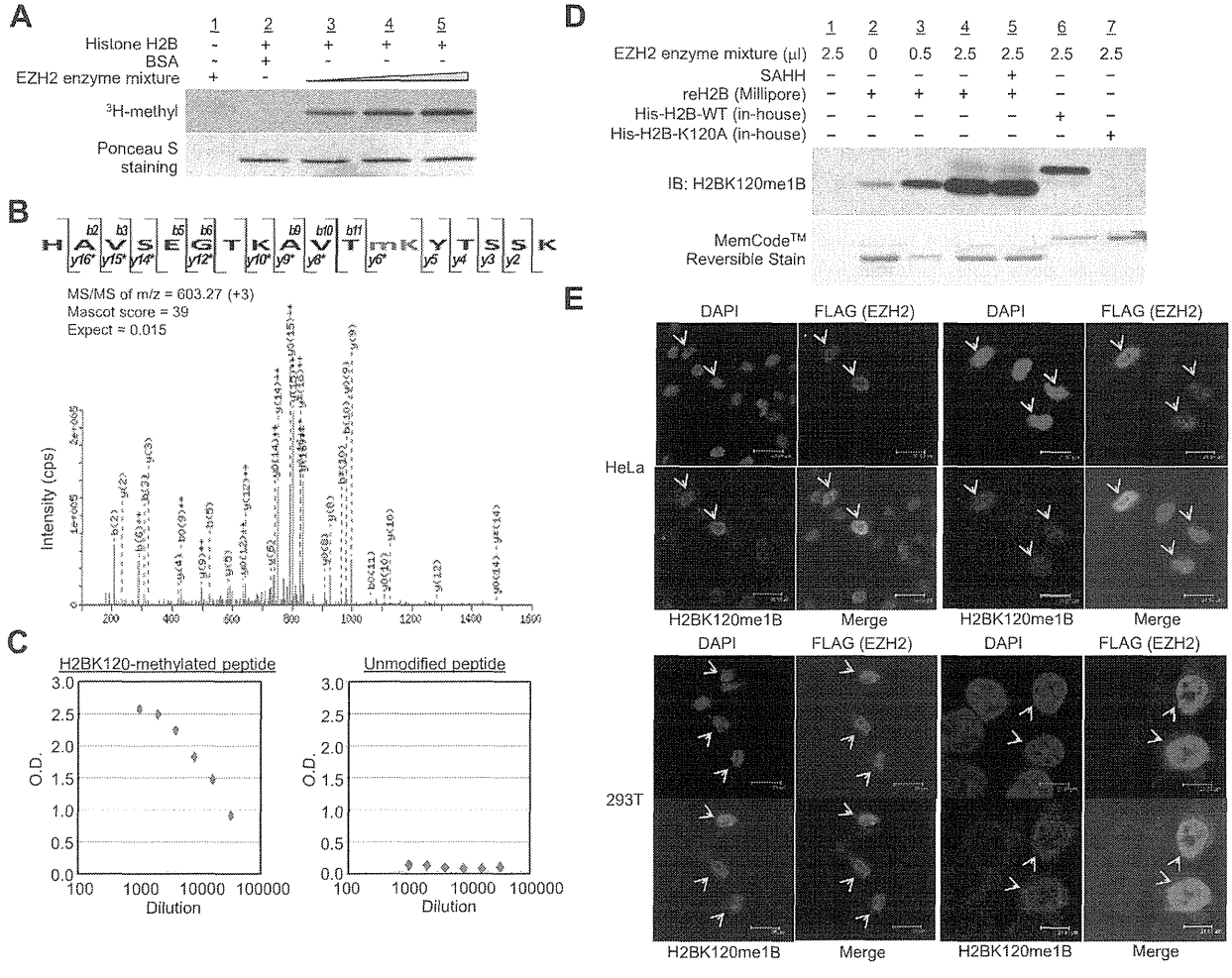

Figure 1. EZH2 methylates lysine 120 of histone H2B both *in vitro* and *in vivo*. (A) *In vitro* methyltransferase assay of H2B. Recombinant H2B and ^{3}H-SAM were incubated in the presence or absence of EZH2 enzyme complex (No. 51004; BPS Bioscience), and the reaction products were analyzed by SDS-PAGE, followed by fluorography (upper panel). The membrane was stained with Ponceau S (Sigma-Aldrich; lower panel). BSA was used as a control. (B) The MS/MS spectrum corresponding to the monomethylated H2B 109-125 peptide. The 14-Da increase of the lysine 120 residue was observed, demonstrating the monomethylated lysine 120. Score and Expect show Mascot Ion Score and Expectation value in Mascot Database search results, respectively. (C) Determination of the titer and specificity of the anti-monomethylated K120 H2B antibody analyzed by ELISA. (D) Validation of an anti-monomethylated K120 H2B antibody. Recombinant H2B-WT or H2B-K120A proteins and ^{3}H-SAM were incubated in the presence or absence of EZH2 enzyme complex, and the reaction products were analyzed by SDS-PAGE, followed by Western blot analysis with an anti-monomethylated K120 histone H2B antibody (upper panel) and stained with MemCode Reversible Stain (lower panel). SAHH, S-adenosyl-L-homocysteine hydrolase. (E) Immunocytochemical analysis of HeLa and 293T cells. Cells were stained with an anti-FLAG antibody [M2; Sigma-Aldrich; Alexa Fluor 594 (red)], an anti-H2BK120me1 antibody [Sigma-Aldrich; Alexa Fluor 488 (green)] and 4′,6′-diamidine-2′-phenylindole dihydrochloride [DAPI (blue)].

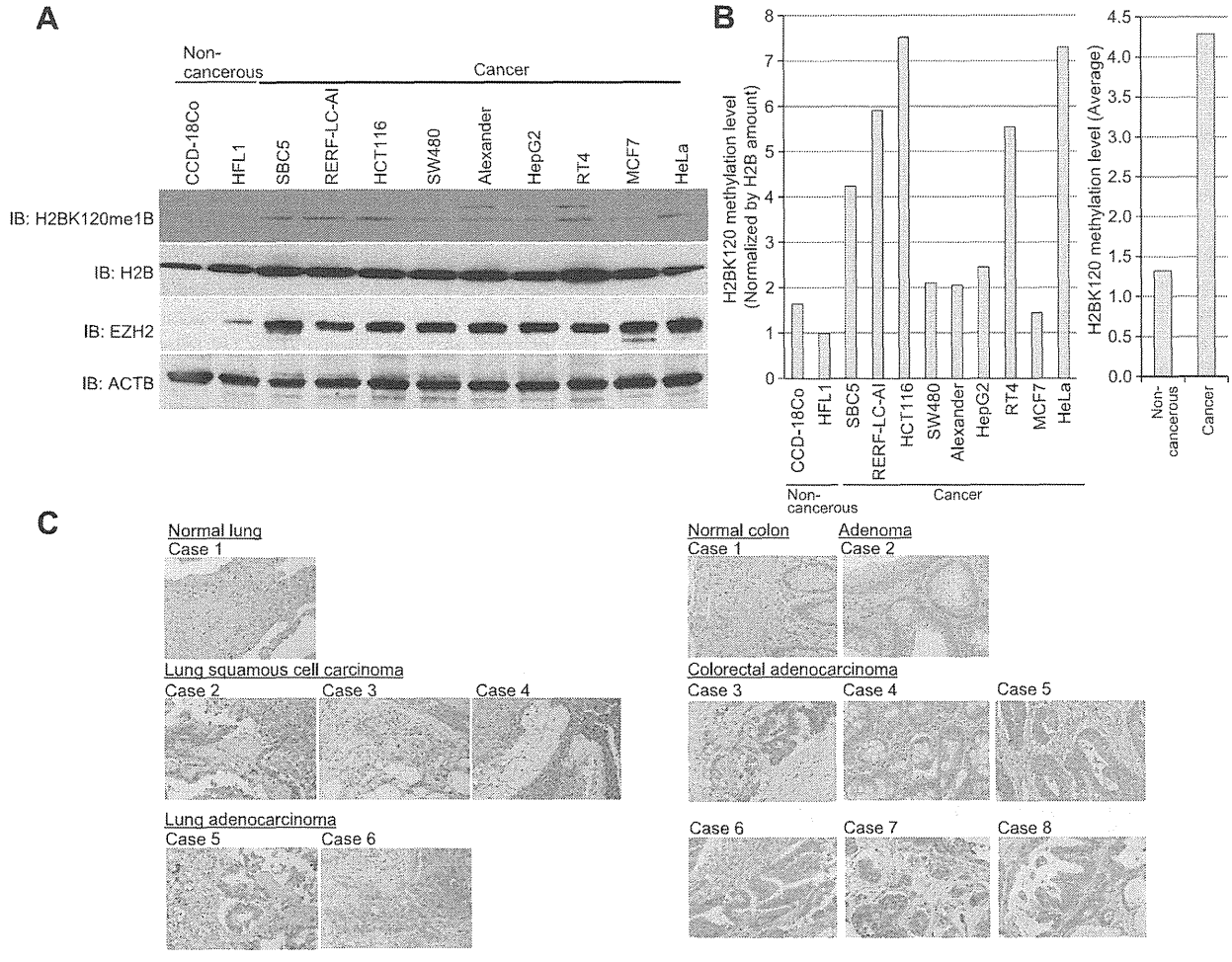

Figure 2. H2BK120 methylation is increased in cancer cells. (A) Validation of H2BK120 methylation status in various cell lines. Lysates from noncancerous cell lines (CCD-18Co and HFL1) and cancer cell lines (SBC5, RERF-LC-AI, HCT116, SW480, Alexander, HepG2, RT4, MCF7, and HeLa) were immunoblotted with anti-monomethylated K120 H2B, anti–histone H2B (ab1790; Abcam), anti-EZH2 (NCL-L-EZH2; Leica), and anti-ACTB (No. 4967; Cell Signaling Technology) antibodies. Information of certificated cell lines is described in Table W4. (B) Quantitative analysis of H2BK120 methylation levels. X-ray films were scanned with GS-800 calibrated densitometer (Bio-Rad Laboratories). The intensity of each H2BK120 monomethylation signal was normalized by the corresponding H2B signal and averaged. (C) Tissue microarray images of lung and colon tissues stained by standard immunohistochemistry for the methylation status of H2BK120. Clinical information for each section is represented in Tables W1 and W2. All tissue samples were purchased from BioChain (Newark, CA). Original magnification, ×200.

analysis to define the methylation site on H2B and found that H2BK120 was monomethylated by EZH2 (Figure 1*B*). We then generated anti–H2BK120-methylated antibodies using an H2BK120-methylated peptide (Figure W2). Among four different lots of the antibodies we obtained, we confirmed by slot blot analysis that H2BK120me1B showed the best quality of the result (Figure W3). In addition, ELISA analysis using this H2BK120-methylation–specific antibody clearly indicated that this antibody recognized the K120-methylated peptide but did not react to the unmethylated peptide, supporting the high specificity of this antibody against H2BK120 methylation (Figure 1*C*). To further verify the quality of the H2BK120-methylation antibody, we prepared both wild-type H2B (H2B-WT) and K120-substituted H2B (H2B-K120A) recombinant proteins and performed *in vitro* methyltransferase assays using EZH2. The H2BK120-methylation signal was increased in a dose-dependent manner, and the signal was completely diminished when H2B-K120A protein was used as a substrate (Figure 1*D*). Using the methylation-specific antibody, we performed immunocytochemical analysis after transfection of EZH2 vectors into HeLa cells to confirm EZH2-dependent H2BK120 methylation *in vivo* and detected the strong staining of methylated H2BK120 in EZH2-overexpressing cells (Figure 1*E*). Taken together, EZH2 can methylate H2BK120, and this methylation is also observed *in vivo*.

H2BK120 Monomethylation Is Increased in Cancer Cells

Because we previously found that the histone methyltransferase EZH2 was overexpressed in various types of cancer [1], we examined whether the methylation level of H2BK120 was enhanced in human cancer cells. Firstly, we prepared cell extracts of two noncancerous cell lines (CCD-18Co and HFL1) and nine cancer cell lines (SBC5, RERF-LC-AI, HCT116, SW480, Alexander, HepG2, RT4, MCF7,

and HeLa; Table W5), and conducted Western blot analysis using the specific antibody (Figure 2*A*). As we expected, the methylation levels of H2BK120 in cancer cell lines expressing high levels of EZH2 were higher than in noncancerous cell lines in which EZH2 expression was very low (Figures 2*B* and W4). We subsequently conducted immunohistochemical analysis of lung and colon tissues (one normal lung, three lung squamous cell carcinomas, and two lung adenocarcinomas as well as one normal colon, one colon adenoma, and six colorectal adenocarcinomas) using the methylation-specific antibody and found higher methylation levels of H2BK120 in cancer cells than normal cells (Figure 2*C*, *left* for lung and *right* for colon).

H2BK120 Monomethylation Competitively Inhibits Monoubiquitination

As H2BK120 was shown to be monoubiquitinated and play a crucial role in the transcriptional regulation, we examined the relationship between H2BK120 methylation and ubiquitination [9]. Immunocytochemial analysis was conducted after FLAG-tagged EZH2 expression vectors were transfected into HeLa cells. EZH2-transfected cells showed significantly lower signals of H2BK120 ubiquitination than control cells (Figure 3*A*). This result indicated that EZH2-dependent H2BK120 methylation appears to competitively inhibit the ubiquitination. In addition, we knocked down the expression of EZH2 in SBC5 cells and conducted the Western blot analysis using purified histone samples to validate the relationship between H2BK120 methylation and ubiquitination. Consistently, H2BK120 methylation status of purified histone H2B proteins was increased after knockdown of EZH2 (Figure 3*B*). Furthermore, we examined the status of H2BK120 ubiquitination in normal and cancer cells and confirmed that ubiquitination levels in cancer cells were lower than those in noncancerous cells (Figure W5). These results reveal that the histone methyltransferase EZH2 is likely to inhibit the

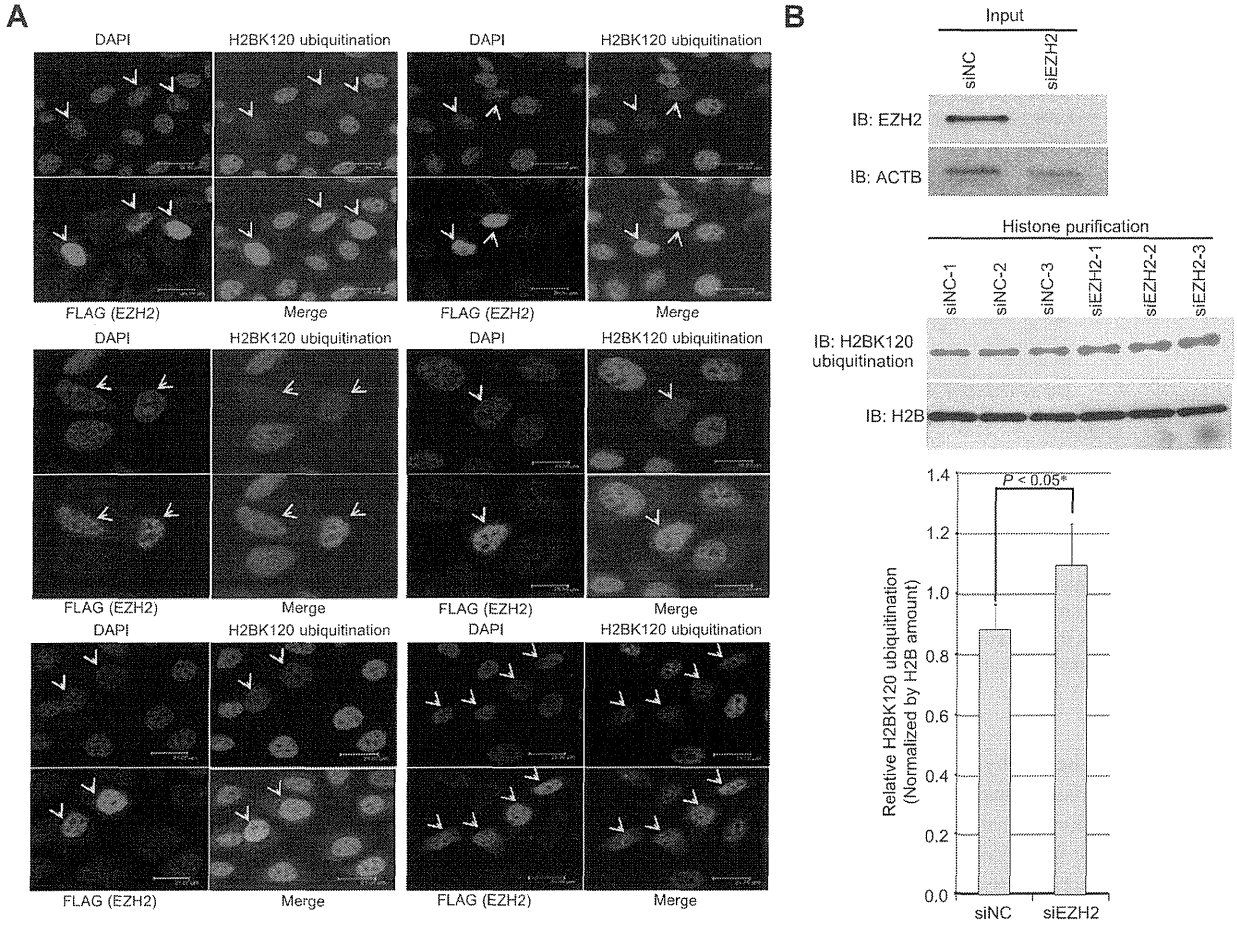

Figure 3. H2BK120 methylation competitively inhibited the ubiquitination. (A) Immunocytochemical analysis of HeLa cells. Cells were stained with an anti–Ubiquityl-Histone H2BK120 antibody (No. 5546; Cell Signaling Technology), an anti-FLAG antibody (M2; Sigma-Aldrich), and 4′,6′-diamidine-2′-phenylindole dihydrochloride [DAPI (blue)]. Alexa Fluor 488 (green) and Alexa Fluor 594 (red) antibodies were used as secondary antibodies. HeLa cells were fixed in 4% paraformaldehyde 24 hours after transfection of FLAG-tagged EZH2 expression vectors. (B) Effects of EZH2 knockdown on the ubiquitination of H2BK120 in SBC5 cells. SBC5 cells were lysed 48 hours after treatment with siNC and siEZH2. Samples were immunoblotted with anti-EZH2 (NCL-L-EZH2; Leica, upper) and anti–Ubiquityl-Histone H2BK120 (middle) antibodies. Expressions of ACTB (No. 4967; Cell Signaling Technology) and H2B (ab1790; Abcam) were the internal controls. The ubiquitination level of H2BK120 was calculated by GS-800 (Bio-Rad Laboratories, lower). Results are the means ± SD of three independent experiments. The *P* value was calculated using Student's *t* test.

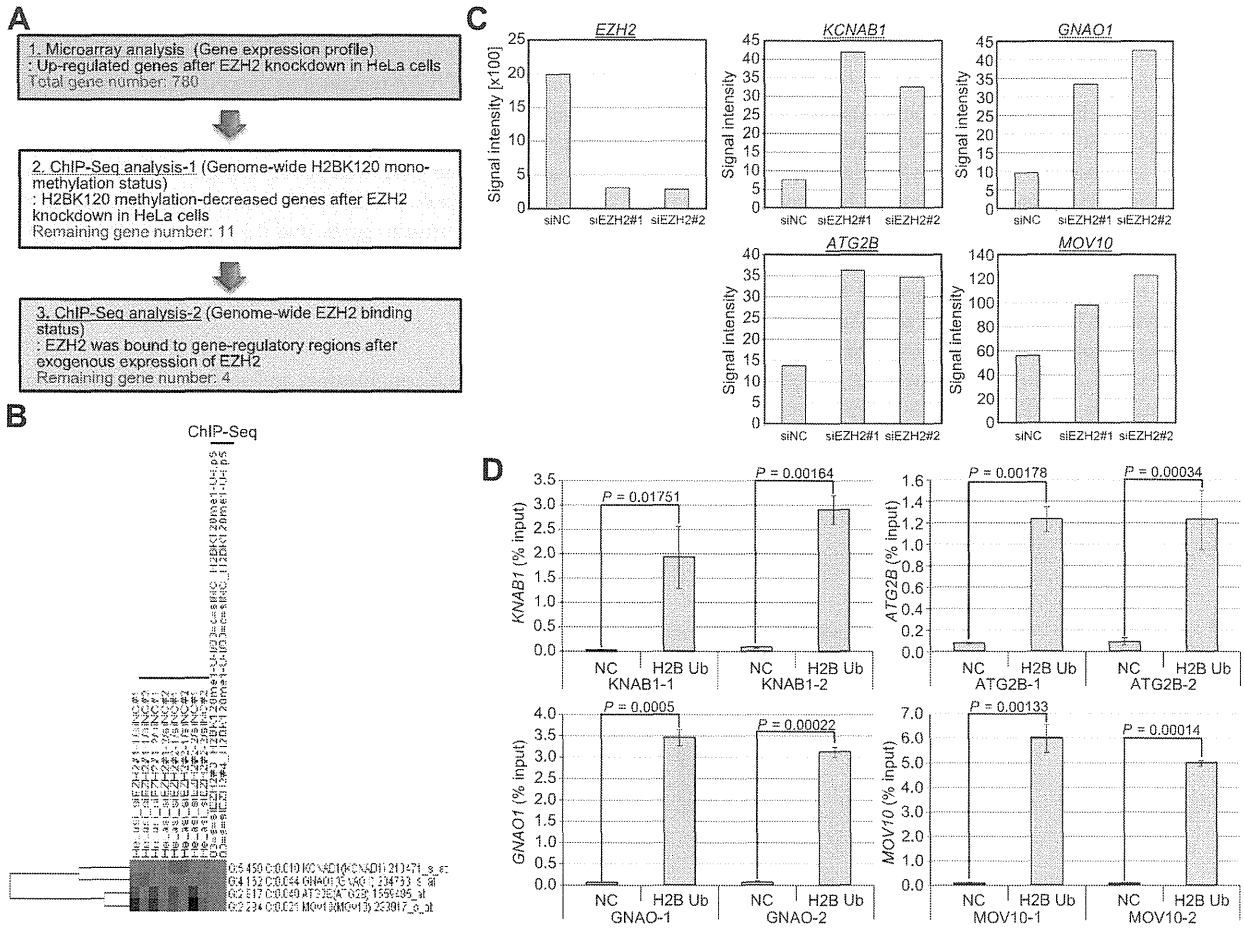

Figure 4. Direct downstream genes of EZH2 through the methylation of H2BK120. (A) Schematic diagram of the strategy to identify direct downstream genes of EZH2. Detailed procedures of ChIP-Seq and microarray analyses were described in Materials and Methods section. 1) Gene expression profile analysis was conducted using the Affymetrix GeneChip system after knockdown of EZH2 in HeLa cells. Seven hundred eighty genes were identified to be downregulated by EZH2 knockdown (>2.0). 2) Chip-Seq analysis to examine the methylation status of histone H2BK120 monomethylation in gene-regulatory regions. Among 780 genes identified by microarray analysis, the methylation status was decreased in 11 genes. 3) ChIP-Seq analysis to identify binding regions of EZH2 proteins after exogenous expression of EZH2. Finally, four genes remained as direct downstream genes of EZH2 through the methylation of H2BK120. (B) The heatmap of four genes directly regulated by EZH2 through the inhibition of histone H2BK120 methylation. Affymetrix GeneChip data (microarray analysis) and ChIP-Seq data were combined. G, the expression ratio of each target gene between siEZH2-treated and control cells (siEZH2/Control). C, the ratio of methylation status in the regulatory region of each gene (siEZH2/Control). Probe Set ID (HG-U133_Plus_2.na33.annotation) is shown with gene symbol. (C) Expression levels of *EZH2*, *KCNAB1*, *GNAO1*, *ATG2B*, and *MOV10* in HeLa cells were analyzed by Affymetrix GeneChip system. (D) ChIP analysis using primer pairs of downstream candidate genes as described under Materials and Methods section. Cross-linked and sheared chromatin was immunoprecipitated with an anti–Ubiquityl-Histone H2BK120 antibody (No. 5546; Cell Signaling Technology). Protein A Agarose/Salmon Sperm DNA (Millipore) was used as a negative control. Results are the means ± SD of three independent experiments and shown as a percentage of the input chromatin. The *P* value was calculated using Student's *t* test.

ubiquitination of H2BK120 in a competitive manner through the methylation of the lysine residue.

Mechanisms in Transcriptional Regulation through EZH2-Dependent H2BK120 Methylation

As it is known that H2BK120 ubiquitination plays a crucial role in the transcriptional regulation, we next examined the transcriptional regulation mechanism of EZH2-dependent H2BK120 methylation in cancer cells. To identify direct downstream genes of EZH2 through the methylation of H2BK120, we conducted cDNA microarray and ChIP-Seq analyses and combined the data. At first, cDNA microarray analysis using the Affymetrix GeneChip system was performed, and expression levels of 780 genes were increased after treatment with siEZH2 in HeLa cells (Figure 4*A*, more than two times compared with control). Then, we planned to do ChIP-Seq analysis to examine the genome-wide H2BK120 monomethylation status. Firstly, we validated the quality of the antibody whether it is available for immunoprecipitation or not. FLAG-conjugated wild-type H2B (H2B-WT) and K120-substituted H2B (H2B-K120A) expression vectors were

transfected into 293T cells, and cell lysates were immunoprecipitated with anti-H2BK120 methylation antibodies, followed by Western blot analysis using an anti-FLAG antibody (Figure W6*A*). Among three lots of anti-H2BK120 methylation antibodies, H2BK120me1B showed the best quality as well as slot blot analysis. To evaluate the availability of this antibody for the immunoprecipitation of endogenous H2BK1200-methylated proteins, Western blot analysis using the H2BK120me1B antibody was performed after immunoprecipitation of SBC5 cell lysates using the same antibody (Figure W6*B*). Subsequently, we confirmed the specific signal of K120-methylated H2B proteins, implying that this antibody is also available for the immunoprecipitation of endogenous K120-methylated histone H2B. On the basis of this result, ChIP-Seq analysis was conducted after treatment with siNC (negative control) and two independent siEZH2s, and among 780 genes selected by microarray analysis, the methylation status of gene-regulatory regions of 11 genes was significantly decreased (Figure 4*A*). In addition, we also examined the genome-wide EZH2-binding status to identify the gene-regulatory regions that EZH2 was bound by ChIP-Seq analysis. Finally, four genes (*KCNAB1*, *GNAO1*, *ATG2B*, and *MOV10*) remained as direct downstream genes of EZH2 through the methylation of H2BK120 after combining the data of microarray and ChIP-Seq analyses (Figure 4*B*). Detailed expression profile data of the four genes were shown in Figure 4*C*. These results reveal that EZH2-dependent H2BK120 methylation can regulate the transcription of downstream genes, and it is a novel mechanism of human carcinogenesis mediated by EZH2 overexpression.

Discussion

Histone H2B is a core histone, and so far, acetylation, phosphorylation, ubiquitination, and sumoylation have already been reported as posttranslational modifications of this protein. In the present study, we identified that lysine 120 of histone H2B is methylated by EZH2 and that the methylation competitively inhibits the

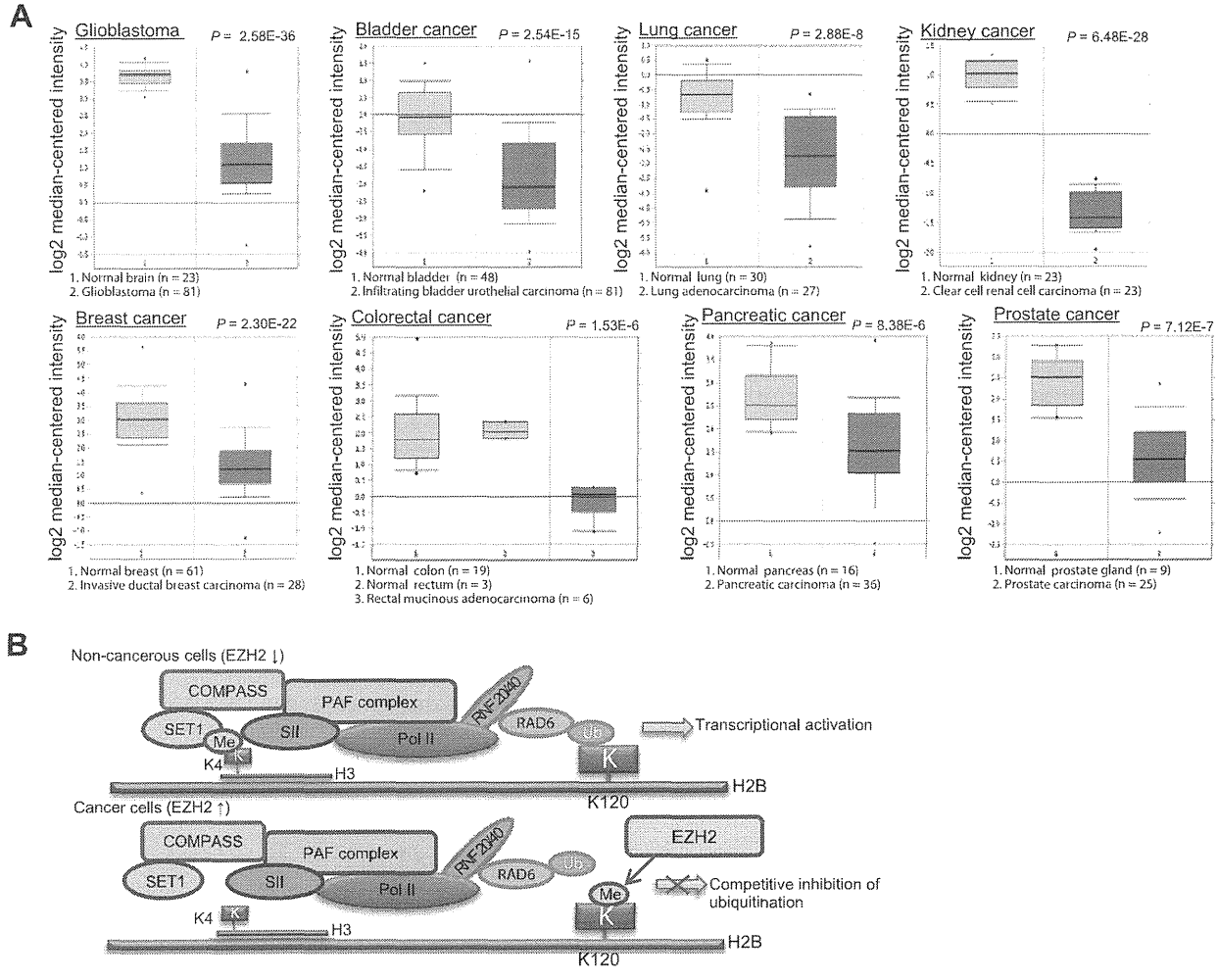

Figure 5. EZH2-dependent transcriptional regulation mechanism through H2BK120 methylation. (A) Expression of *KCNAB1* in cancer tissues. Expression levels of *KCNAB1* are downregulated in various types of cancer. Gene expression data in Oncomine (University of Michigan, Ann Arbor, MI) were analyzed. The thick bars in the boxes are average expression levels, and the boxes represent 95% of the samples. The error bars are above or below the boxes, and the range of expression levels is enclosed by two dots. (B) Proposed model for regulation of H2BK120 posttranslational modification during human carcinogenesis.

ubiquitination of H2BK120. This is the first report to describe the function of histone H2B methylation.

In mammals, H2BK120 monoubiquitination was reported to be preferentially associated with highly transcribed genes [10]. H2BK120 monoubiquitination can cooperate with facilitates chromatin transcription (FACT) and the polymerase associated factor (PAF) complex to regulate transcription elongation by RNA polymerase II [28] and can also facilitate DNA repair [29] and mRNA 3′ end processing [11] in human cells. In addition, Vitaliano-Prunier et al. recently reported that H2B ubiquitination in yeast cells plays a role in mRNA export from the nucleus into the cytoplasm [30]. Importantly, there are several reports describing the relationship between H2BK120 ubiquitination and cancer. The H2B deubiquitinase ubiquitin specific peptidase 22 (USP22) is part of a gene signature predictive of a cancer stem cell tumor phenotype of aggressive growth, metastasis, and therapy resistance [31]. RNF20, the E3 ubiquitin ligase for monoubiquitination of H2BK120, is a putative tumor suppressor [32]; its down-regulation in human cells promotes migration, anchorage independence, and carcinogenesis [32,33]. Furthermore, it was recently reported that H2B ubiquitination levels were decreased in advanced and metastatic breast cancer, parathyroid tumors, and seminoma [34–36]. These data imply that H2BK120 ubiquitination appears to possess tumor-suppressive functions, and its dysfunction causes malignant alteration of human cells.

In this study, we demonstrated that the histone methyltransferase EZH2 methylated H2BK120 and competitively inhibited the ubiquitination and that in cancer cells, the methylation level was increased, and the ubiquitination was decreased at lysine 120 of histone H2B. From these results, H2BK120 ubiquitination shows tumor-suppressive functions, and the oncogenic polycomb histone methyltransferase EZH2 competitively inhibits the ubiquitination through the methylation of H2BK120, which may result in the reduction of tumor-suppressive effects. In addition, we identified direct downstream genes of EZH2 through the methylation of H2BK120 on the basis of ChIP-Seq and microarray analyses (Figure 4*A*). Importantly, ChIP analysis indicated that H2BK120 ubiquitination levels of downstream candidate genes were significantly high in nontumor cells, implying that these genes may be transcriptionally activated in nontumor cells through the H2BK120 ubiquitination (Figure 4*D*). Among direct downstream candidates we identified, expression levels of *KCNAB1* and *GNAO1* are significantly decreased in various types of tumor tissues (Figures 5*A* and W7). *KCNAB1* is a voltage-gated potassium channel β subunit and modulates the activity of the pore-forming α subunit [37,38]. It has already been reported that some components of potassium channel serve as tumor suppressors. *KCTD11* (potassium channel tetramerization domain containing 11) is a novel tumor suppressor gene that inhibits cell growth and is mapping on human chromosome 17p13.2 [39]. According to the research using a panel of 177 human tumor samples and their normal matching samples representing 18 different types of cancer, down-regulation of KCTD11 protein level is a diffusely common event in tumorigenesis [40]. In addition, the putative tumor suppressor *KCNRG* (potassium channel–regulating gene) encodes a protein with a high homology to the tetramerization domain of voltage-gated K+ (Kv) channels [41]. This protein appears to interfere with the normal assembly of the K+ channel proteins by binding to their tetramerization domain, thereby causing the suppression of Kv currents [42]. Because Kv channels are involved in the proliferation of tumor cells [43] and normal lymphocytes [44], while being upregulated in neoplastic hematopoietic cells [45], *KCNRG* seems to exert a tumor suppressor effect [41]. On the contrary, *GNAO1*, a member of the signal-transducing guanine nucleotide-binding (G) protein family, has been implicated in ion channel regulation [46], and mutation of *GNAO1* was reported to cause the malignant alteration of cells [47,48]. *GPRC5a*, a family gene of *GNAO1*, is reported as a tumor suppressor gene, and knockout of GPCR5 leads to NF-κB activation in airway epithelium and promotes lung tumorigenesis [49]. Furthermore, according to expression profile analysis, expression levels of *ATG2B*, one of the autophagy-related genes, in several types of tumor tissues were much lower than those in corresponding normal tissues (Figure W8). Frameshift mutations of ATG2B were found in gastric and colorectal cancers [50]. The truncation mutants of ATG2B may inactivate its autophagy function and/or autophagic cell death, which resembles a typical loss-of-function mutation. These data imply that *KCNAB1*, *GNAO1*, and *ATG2B* are likely to be deregulated in cancer and work as tumor suppressors. Taken together, the oncogenic histone methyltransferase EZH2 is likely to contribute to human carcinogenesis based on the transcriptional regulation of downstream genes through not only H3K27 methylation but also H2BK120 methylation. In line with this, we also conducted the ChIP-Seq analysis and found that *GNAO1*, *ATG2B*, and *MOV10* were regulated by H2BK120 methylation, but the *KCNAB1* gene appears to be regulated by both H2BK120 methylation and H3K27 methylation (Table W6). This implies that some gene is likely to be synergistically regulated by two different histone marks.

All the while, histone H3-K27 has been only the known target of EZH2-dependent methylation, and the importance of EZH2 in human carcinogenesis was described only focusing on H3K27 methylation. In this study, we identified H2BK120 to be a novel target of EZH2-dependent methylation and downstream candidates through the novel methylation site (Figure 5*B*). Because EZH2 is overexpressed in various types of cancer and it plays a critical role in the growth regulation of cancer cells, EZH2 is recognized as an important target of anticancer treatment. In fact, chemical compounds, which inhibit methyltransferase activity of EZH2, have recently been developed as anticancer drugs. According to our findings, we need to take care of H2BK120 methylation besides H3K27 methylation to develop efficient anticancer treatment targeting EZH2, and the identification of this novel methylation site must contribute to unveil the multifunctions of EZH2 in human carcinogenesis. Furthermore, additional functional analysis of H2BK120 methylation may elucidate the significance of biologic function of this methylation, including disease like cancer, and the importance for diagnosis, prognosis, and treatment for cancer.

Acknowledgments

We thank Hyun-Soo Cho, Kzuyuki Hayashi, Kazuhiro Maejima, Yuka Yamane, Yukiko Iwai, and Haruka Sawada for technical assistance.

References

[1] Takawa M, Masuda K, Kunizaki M, Daigo Y, Takagi K, Iwai Y, Cho HS, Toyokawa G, Yamane Y, Maejima K, et al. (2011). Validation of the histone methyltransferase EZH2 as a therapeutic target for various types of human cancer and as a prognostic marker. *Cancer Sci* **102**, 1298–1305.

[2] Knutson SK, Wigle TJ, Warholic NM, Sneeringer CJ, Allain CJ, Klaus CR, Sacks JD, Raimondi A, Majer CR, Song J, et al. (2012). A selective inhibitor of EZH2 blocks H3K27 methylation and kills mutant lymphoma cells. *Nat Chem Biol* **8**, 890–896.

[3] McCabe MT, Ott HM, Ganji G, Korenchuk S, Thompson C, Van Aller GS, Liu Y, Graves AP, Della Pietra A III, Diaz E, et al. (2012). EZH2 inhibition as a therapeutic strategy for lymphoma with EZH2-activating mutations. *Nature* **492**, 108–112.

[4] Schiltz RL, Mizzen CA, Vassilev A, Cook RG, Allis CD, and Nakatani Y (1999). Overlapping but distinct patterns of histone acetylation by the human coactivators p300 and PCAF within nucleosomal substrates. *J Biol Chem* **274**, 1189–1192.

[5] Suka N, Suka Y, Carmen AA, Wu J, and Grunstein M (2001). Highly specific antibodies determine histone acetylation site usage in yeast heterochromatin and euchromatin. *Mol Cell* **8**, 473–479.

[6] Cheung WL, Ajiro K, Samejima K, Kloc M, Cheung P, Mizzen CA, Beeser A, Etkin LD, Chernoff J, Earnshaw WC, et al. (2003). Apoptotic phosphorylation of histone H2B is mediated by mammalian sterile twenty kinase. *Cell* **113**, 507–517.

[7] Zhu B, Zheng Y, Pham AD, Mandal SS, Erdjument-Bromage H, Tempst P, and Reinberg D (2005). Monoubiquitination of human histone H2B: the factors involved and their roles in *HOX* gene regulation. *Mol Cell* **20**, 601–611.

[8] Nathan D, Ingvarsdottir K, Sterner DE, Bylebyl GR, Dokmanovic M, Dorsey JA, Whelan KA, Krsmanovic M, Lane WS, Meluh PB, et al. (2006). Histone sumoylation is a negative regulator in *Saccharomyces cerevisiae* and shows dynamic interplay with positive-acting histone modifications. *Genes Dev* **20**, 966–976.

[9] Sun ZW and Allis CD (2002). Ubiquitination of histone H2B regulates H3 methylation and gene silencing in yeast. *Nature* **418**, 104–108.

[10] Minsky N, Shema E, Field Y, Schuster M, Segal E, and Oren M (2008). Monoubiquitinated H2B is associated with the transcribed region of highly expressed genes in human cells. *Nat Cell Biol* **10**, 483–488.

[11] Pirngruber J, Shchebet A, Schreiber L, Shema E, Minsky N, Chapman RD, Eick D, Aylon Y, Oren M, and Johnsen SA (2009). CDK9 directs H2B monoubiquitination and controls replication-dependent histone mRNA 3′-end processing. *EMBO Rep* **10**, 894–900.

[12] Kim J, Hake SB, and Roeder RG (2005). The human homolog of yeast BRE1 functions as a transcriptional coactivator through direct activator interactions. *Mol Cell* **20**, 759–770.

[13] Fierz B, Chatterjee C, McGinty RK, Bar-Dagan M, Raleigh DP, and Muir TW (2011). Histone H2B ubiquitylation disrupts local and higher-order chromatin compaction. *Nat Chem Biol* **7**, 113–119.

[14] Fuchs G, Shema E, Vesterman R, Kotler E, Wolchinsky Z, Wilder S, Golomb L, Pribluda A, Zhang F, Haj-Yahya M, et al. (2012). RNF20 and USP44 regulate stem cell differentiation by modulating H2B monoubiquitylation. *Mol Cell* **46**, 662–673.

[15] Hamamoto R, Furukawa Y, Morita M, Iimura Y, Silva FP, Li M, Yagyu R, and Nakamura Y (2004). *SMYD3* encodes a histone methyltransferase involved in the proliferation of cancer cells. *Nat Cell Biol* **6**, 731–740.

[16] Hamamoto R, Silva FP, Tsuge M, Nishidate T, Katagiri T, Nakamura Y, and Furukawa Y (2006). Enhanced SMYD3 expression is essential for the growth of breast cancer cells. *Cancer Sci* **97**, 113–118.

[17] Cho HS, Shimazu T, Toyokawa G, Daigo Y, Maehara Y, Hayami S, Ito A, Masuda K, Ikawa N, Field HI, et al. (2012). Enhanced HSP70 lysine methylation promotes proliferation of cancer cells through activation of Aurora kinase B. *Nat Commun* **3**, 1072.

[18] Cho HS, Suzuki T, Dohmae N, Hayami S, Unoki M, Yoshimatsu M, Toyokawa G, Takawa M, Chen T, Kurash JK, et al. (2011). Demethylation of RB regulator MYPT1 by histone demethylase LSD1 promotes cell cycle progression in cancer cells. *Cancer Res* **71**, 655–660.

[19] Takawa M, Cho HS, Hayami S, Toyokawa G, Kogure M, Yamane Y, Iwai Y, Maejima K, Ueda K, Masuda A, et al. (2012). Histone lysine methyltransferase SETD8 promotes carcinogenesis by deregulating PCNA expression. *Cancer Res* **72**, 3217–3227.

[20] Hayami S, Kelly JD, Cho HS, Yoshimatsu M, Unoki M, Tsunoda T, Field HI, Neal DE, Yamaue H, Ponder BA, et al. (2011). Overexpression of LSD1 contributes to human carcinogenesis through chromatin regulation in various cancers. *Int J Cancer* **128**, 574–586.

[21] Kogure M, Takawa M, Cho HS, Toyokawa G, Hayashi K, Tsunoda T, Kobayashi T, Daigo Y, Sugiyama M, Atomi Y, et al. (2013). Deregulation of the histone demethylase JMJD2A is involved in human carcinogenesis through regulation of the G_1/S transition. *Cancer Lett* **336**, 76–84.

[22] Li H and Durbin R (2009). Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**, 1754–1760.

[23] Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W, et al. (2008). Model-based analysis of ChIP-Seq (MACS). *Genome Biol* **9**, R137.

[24] Feng J, Liu T, and Zhang Y (2011). Using MACS to identify peaks from ChIP-Seq data. *Curr Protoc Bioinformatics* **34**, 2.14.1–2.14.14. Chapter 2, Unit 2 14.

[25] Kadota K, Nishiyama T, and Shimizu K (2012). A normalization strategy for comparing tag count data. *Algorithms Mol Biol* **7**, 5.

[26] Hayami S, Yoshimatsu M, Veerakumarasivam A, Unoki M, Iwai Y, Tsunoda T, Field HI, Kelly JD, Neal DE, Yamaue H, et al. (2010). Overexpression of the JmjC histone demethylase KDM5B in human carcinogenesis: involvement in the proliferation of cancer cells through the E2F/RB pathway. *Mol Cancer* **9**, 59.

[27] Yoshimatsu M, Toyokawa G, Hayami S, Unoki M, Tsunoda T, Field HI, Kelly JD, Neal DE, Maehara Y, Ponder BA, et al. (2011). Dysregulation of PRMT1 and PRMT6, type I arginine methyltransferases, is involved in various types of human cancers. *Int J Cancer* **128**, 562–573.

[28] Pavri R, Zhu B, Li G, Trojer P, Mandal S, Shilatifard A, and Reinberg D (2006). Histone H2B monoubiquitination functions cooperatively with FACT to regulate elongation by RNA polymerase II. *Cell* **125**, 703–717.

[29] Moyal L, Lerenthal Y, Gana-Weisz M, Mass G, So S, Wang SY, Eppink B, Chung YM, Shalev G, Shema E, et al. (2011). Requirement of ATM-dependent monoubiquitylation of histone H2B for timely repair of DNA double-strand breaks. *Mol Cell* **41**, 529–542.

[30] Vitaliano-Prunier A, Babour A, Hérissant L, Apponi L, Margaritis T, Holstege FC, Corbett AH, Gwizdek C, and Dargemont C (2012). H2B ubiquitylation controls the formation of export-competent mRNP. *Mol Cell* **45**, 132–139.

[31] Zhang XY, Varthi M, Sykes SM, Phillips C, Warzecha C, Zhu W, Wyce A, Thorne AW, Berger SL, and McMahon SB (2008). The putative cancer stem cell marker USP22 is a subunit of the human SAGA complex required for activated transcription and cell-cycle progression. *Mol Cell* **29**, 102–111.

[32] Shema E, Tirosh I, Aylon Y, Huang J, Ye C, Moskovits N, Raver-Shapira N, Minsky N, Pirngruber J, Tarcic G, et al. (2008). The histone H2B-specific ubiquitin ligase RNF20/hBRE1 acts as a putative tumor suppressor through selective regulation of gene expression. *Genes Dev* **22**, 2664–2676.

[33] Shema E, Kim J, Roeder RG, and Oren M (2011). RNF20 inhibits TFIIS-facilitated transcriptional elongation to suppress pro-oncogenic gene expression. *Mol Cell* **42**, 477–488.

[34] Chernikova SB, Razorenova OV, Higgins JP, Sishc BJ, Nicolau M, Dorth JA, Chernikova DA, Kwok S, Brooks JD, Bailey SM, et al. (2012). Deficiency in mammalian histone H2B ubiquitin ligase Bre1 (Rnf20/Rnf40) leads to replication stress and chromosomal instability. *Cancer Res* **72**, 2111–2119.

[35] Hahn MA, Dickson KA, Jackson S, Clarkson A, Gill AJ, and Marsh DJ (2012). The tumor suppressor CDC73 interacts with the ring finger proteins RNF20 and RNF40 and is required for the maintenance of histone 2B monoubiquitination. *Hum Mol Genet* **21**, 559–568.

[36] Prenzel T, Begus-Nahrmann Y, Kramer F, Hennion M, Hsu C, Gorsler T, Hintermair C, Eick D, Kremmer E, Simons M, et al. (2011). Estrogen-dependent gene transcription in human breast cancer cells relies upon proteasome-dependent monoubiquitination of histone H2B. *Cancer Res* **71**, 5739–5753.

[37] Leicher T, Roeper J, Weber K, Wang X, and Pongs O (1996). Structural and functional characterization of human potassium channel subunit β1 (KCNA1B). *Neuropharmacology* **35**, 787–795.

[38] Schultz D, Litt M, Smith L, Thayer M, and McCormack K (1996). Localization of two potassium channel β subunit genes, KCNA1B and KCNA2B. *Genomics* **31**, 389–391.

[39] Di Marcotullio L, Ferretti E, De Smaele E, Argenti B, Mincione C, Zazzeroni F, Gallo R, Masuelli L, Napolitano M, Maroder M, et al. (2004). REN^{KCTD11} is a suppressor of Hedgehog signaling and is deleted in human medulloblastoma. *Proc Natl Acad Sci USA* **101**, 10833–10838.

[40] Mancarelli MM, Zazzeroni F, Ciccocioppo L, Capece D, Po A, Murgo S, Di Camillo R, Rinaldi C, Ferretti E, Gulino A, et al. (2010). The tumor suppressor gene $KCTD11^{REN}$ is regulated by Sp1 and methylation and its expression is reduced in tumors. *Mol Cancer* **9**, 172.

[41] Birerdinc A, Nohelty E, Marakhonov A, Manyam G, Panov I, Coon S, Nikitin E, Skoblov M, Chandhoke V, and Baranova A (2010). Pro-apoptotic and anti-proliferative activity of human *KCNRG*, a putative tumor suppressor in 13q14 region. *Tumour Biol* **31**, 33–45.

[42] Ivanov DV, Tyazhelova TV, Lemonnier L, Kononenko N, Pestova AA, Nikitin EA, Prevarskaya N, Skryma R, Panchin YV, Yankovsky NK, et al. (2003).

A new human gene *KCNRG* encoding potassium channel regulating protein is a cancer suppressor gene candidate located in 13q14.3. *FEBS Lett* **539**, 156–160.

[43] Rybalchenko V, Prevarskaya N, Van Coppenolle F, Legrand G, Lemonnier L, Le Bourhis X, and Skryma R (2001). Verapamil inhibits proliferation of LNCaP human prostate cancer cells influencing K+ channel gating. *Mol Pharmacol* **59**, 1376–1387.

[44] Lewis RS and Cahalan MD (1995). Potassium and calcium channels in lymphocytes. *Annu Rev Immunol* **13**, 623–653.

[45] Smith BC and Denu JM (2009). Chemical mechanisms of histone lysine and arginine modifications. *Biochim Biophys Acta* **1789**, 45–57.

[46] Murtagh JJ Jr, Eddy R, Shows TB, Moss J, and Vaughan M (1991). Different forms of Go alpha mRNA arise by alternative splicing of transcripts from a single gene on human chromosome 16. *Mol Cell Biol* **11**, 1146–1155.

[47] Garcia-Marcos M, Ghosh P, and Farquhar MG (2011). Molecular basis of a novel oncogenic mutation in GNAO1. *Oncogene* **30**, 2691–2696.

[48] Kan Z, Jaiswal BS, Stinson J, Janakiraman V, Bhatt D, Stern HM, Yue P, Haverty PM, Bourgon R, Zheng J, et al. (2010). Diverse somatic mutation patterns and pathway alterations in human cancers. *Nature* **466**, 869–873.

[49] Deng J, Fujimoto J, Ye XF, Men TY, Van Pelt CS, Chen YL, Lin XF, Kadara H, Tao Q, Lotan D, et al. (2010). Knockout of the tumor suppressor gene *Gprc5a* in mice leads to NF-κB activation in airway epithelium and promotes lung inflammation and tumorigenesis. *Cancer Prev Res (Phila)* **3**, 424–437.

[50] Kang MR, Kim MS, Oh JE, Kim YR, Song SY, Kim SS, Ahn CH, Yoo NJ, and Lee SH (2009). Frameshift mutations of autophagy-related genes *ATG2B*, *ATG5*, *ATG9B* and *ATG12* in gastric and colorectal cancers with microsatellite instability. *J Pathol* **217**, 702–706.

Table W1. Characteristics of Lung Tissues.

Case No.	Age	Gender	Pathology	Grade	Stage (TNM)	Nature
Case 1	34	Male	Normal	Normal		
Case 2	58	Male	Squamous cell carcinoma	II	T2N0M0	Malignant
Case 3	42	Male	Squamous cell carcinoma	II	T2N0M0	Malignant
Case 4	57	Male	Squamous cell carcinoma	III	T2N0M0	Malignant
Case 5	62	Male	Adenocarcinoma	II	T2N0M0	Malignant
Case 6	73	Male	Adenosquamous carcinoma	III	T2N1M0	Malignant

Table W2. Characteristics of Colon Tissues.

Case No.	Age	Gender	Histology	Grade	Stage (TNM)	Anatomic Site
Case 1	49	Female	Normal			Colon
Case 2	65	Male	Adenoma			Colon
Case 3	51	Male	Adenocarcinoma	II	T3N1M0	Colon
Case 4	39	Female	Adenocarcinoma	I_II	T3N1M0	Colon
Case 5	49	Female	Adenocarcinoma	I_II	T3N0M0	Colon
Case 6	62	Female	Adenocarcinoma	I_II	T2N1M0	Colon
Case 7	60	Male	Adenocarcinoma	II	T2N1M0	Colon
Case 8	34	Male	Adenocarcinoma	I	T3N0M0	Colon

Table W3. siRNA Sequences.

siRNA Name	Sequence
siNC (cocktail)	
Target No. 1	Sense: 5′ AUCCGCGCGAUAGUACGUA3′ Antisense: 5′ UACGUACUAUCGCGCGGAU 3′
Target No. 2	Sense: 5′ UUACGCGUAGCGUAAUACG 3′ Antisense: 5′ CGUAUUACGCUACGCGUAA 3′
Target No. 3	Sense: 5′ UAUUCGCGCGUAUAGCGGU 3′ Antisense: 5′ ACCGCUAUACGCGCGAAUA 3′
siEZH2 No. 1	Sense: 5′ CUAACCAUGUUUACAACUA 3′ Antisense: 5′ UAGUUGUAAACAUGGUUAG 3′
siEZH2 No. 2	Sense: 5′ GACAGAAGAGGGAAAGUGU 3′ Antisense: 5′ ACACUUUCCCUCUUCUGUC 3′

Table W4. Primer Sequences for ChIP Analysis.

Name	Sequence	Length	Tm (°C)	MW
KCNAB1 Set 1 (KCNAB1-1): Amplification Size = 123				
KCNAB1-ChIP-f1	TTC CGT GTT CGA AGA TAC CAC	21	59.5	6381.2
KCNAB1-ChIP-r1	AAC CTT ATC CTG CCA CAA AGC	21	59.5	6319.2
KCNAB1 Set 2 (KCNAB1-2): Amplification Size = 119				
KCNAB1-ChIP-f2	ATT TGT CAG AAG TGC TGG GAG G	22	62.1	6870.5
KCNAB1-ChIP-r2	TCT TTG AAT GTC AGT GAA CCA C	22	58.4	6709.4
GNAO1 Set 1 (GNAO1-1): Amplification Size = 139				
GNAO1-ChIP-f1	AGC CTC GGG TGT CAC ATA TTA G	22	62.1	6750.4
GNAO1-ChIP-r1	CTC AGG AAA CGC GAT GTG GTA G	22	64.2	6824.5
GNAO1 Set 2 (GNAO1-2): Amplification Size = 118				
GNAO1-ChIP-f2	ATT CCG ACC CAC TAC CAC ATC	21	61.2	6255.1
GNAO1-ChIP-r2	GGG CCG GCT CTC CAT CTT GTC	21	67.3	6365.2
ATG2B Set 1 (ATG2B-1): Amplification Size = 122				
ATG2B-ChIP-f1	TCG GAG CCG GAA CTG CTC CAG	21	67.3	6432.2
ATG2B-ChIP-r1	CTC CTG GCG CGT TCA CGA GAC	21	67.3	6383.2
ATG2B Set 2 (ATG2B-2): Amplification Size = 98				
ATG2B-ChIP-f2	CCA GGA TTA AGC GAG CGT ATG	21	61.2	6495.3
ATG2B-ChIP-r2	CCC CGC CTC ATT CAG GTA TTG	21	63.2	6333.2
MOV10 Set 1 (MOV10-1): Amplification Size = 149				
MOV10-ChIP-f1	TTC CCA CTG ACA TTG CAT TTC	21	57.5	6307.2
MOV10-ChIP-r1	AGG CCA CAC ACT CAA TCT ACG	21	61.2	6344.2
MOV10 Set 2 (MOV10-2): Amplification Size = 104				
MOV10-ChIP-f2	TTA CTG TGT ATC CTG GCA GAG C	22	62.1	6741.4
MOV10-ChIP-r2	CAT AAG GGT CAA AGA AGT TTG G	22	58.4	6847.6

Tm indicates melting temperature; MW, molecular weight.

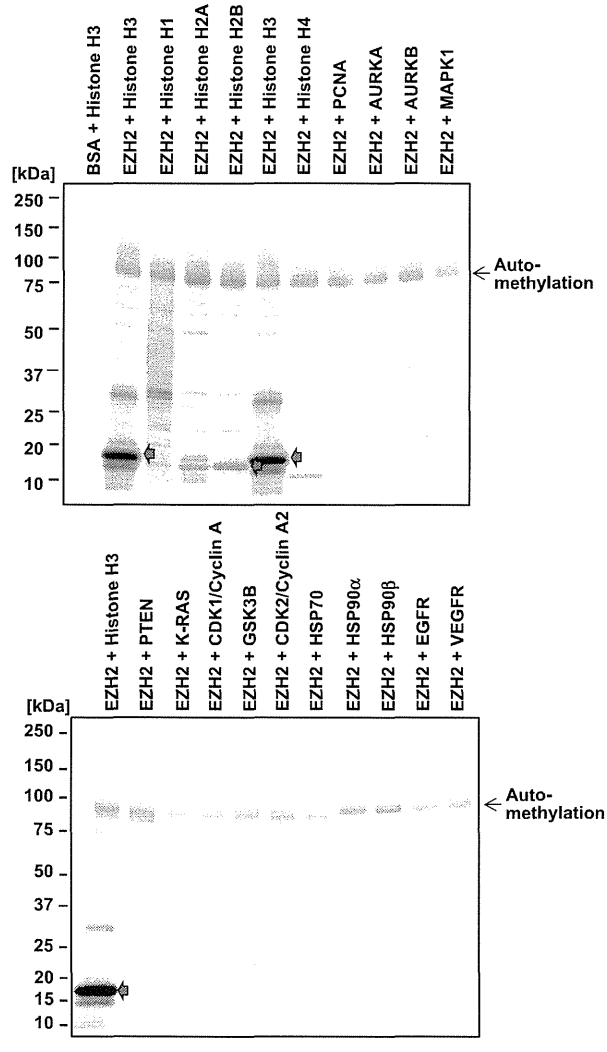

Figure W1. EZH2 methylates histone H2B. *In vitro* methyltransferase of EZH2 was performed using various types of recombinant proteins as substrates. Methylated proteins were visualized with fluorography.