Acquired resistance to inhibitors of anaplastic lymphoma kinase (ALK) is a major clinical challenge for ALK fusion-positive non-small-cell lung cancer (NSCLC). In the absence of secondary ALK mutations, epigenetic reprogramming is one of the main mechanisms of drug resistance, as it leads to phenotype switching that occurs during the epithelial-to-mesenchymal transition (EMT). Although drug-induced epigenetic reprogramming is believed to alter the sensitivity of cancer cells to anticancer treatments, there is still much to learn about overcoming drug resistance. In this study, we used an in vitro model of ceritinib-resistant NSCLC and employed genome-wide DNA methylation analysis in combination with single-cell (sc) RNA-seq to identify cytidine deaminase (CDA), a pyrimidine salvage pathway enzyme, as a candidate drug target. CDA was hypomethylated and upregulated in ceritinib-resistant cells. CDA-overexpressing cells were rarely but definitively detected in the naive cell population by scRNA-seq, and their abundance was increased in the acquired-resistance population. Knockdown of CDA had antiproliferative effects on resistant cells and reversed the EMT phenotype. Treatment with epigenome-related nucleosides such as 5-formyl-2′-deoxycytidine selectively ablated CDA-overexpressing resistant cells via accumulation of DNA damage. Collectively, our data suggest that targeting CDA metabolism using epigenome-related nucleosides represents a potential new therapeutic strategy for overcoming ALK inhibitor resistance in NSCLC.

Experimental & Molecular Medicine (2022) 54:1236–1249; https://doi.org/10.1038/s12276-022-00836-7

INTRODUCTION
Anaplastic lymphoma kinase (ALK) is a receptor tyrosine kinase that is expressed in the nervous system, testes, and small intestine in adult humans. As documented thus far, chromosomal rearrangements of ALK result in fusions with more than 20 different genes, and ALK fusion proteins drive tumorigenesis in many different cancers. A fusion between EML4 (echinoderm microtubule-associated protein-like 4) and ALK was identified in non-small-cell lung cancer (NSCLC) in 2007. Since then, ALK fusions have been detected in ~3–7% of patients with NSCLC and have been associated with nonsmoking and younger age among patients.

ALK inhibitors are highly effective at treating patients with ALK fusion-positive NSCLC, but the inevitable emergence of chemotherapeutic drug resistance limits their utility. Genetic mechanisms such as secondary mutations in the ALK kinase domain explain the drug resistance observed in 20 and 50% of patients with first-generation (crizotinib) and second-generation (ceritinib and alectinib) treatment, respectively. This suggests that nongenetic mechanisms contribute to drug resistance by modulating transcriptome plasticity. It was recently suggested that drug-induced epigenetic reprogramming can alter the sensitivity of cancer cells to anticancer treatments. Indeed, epigenome characteristics such as DNA methylation, histone modifications, and noncoding RNAs contribute to gene expression changes for adaptation in response to anticancer therapies, but there is still much to learn about how epigenetic mechanisms contribute to drug resistance.

Acquisition of nucleotides is critical for DNA replication in proliferating cells. Nucleotides are produced either by de novo synthesis or by salvage pathways that recycle nucleobases from intracellular degradation processes or are acquired via extra-cellular uptake. Cytidine deaminase (CDA) is an enzyme of the pyrimidine salvage pathway that catalyzes the deamination of cytidine and deoxycytidine to produce uridine and deoxyuridine, respectively. Epigenome-related nucleosides such as...
5-hydroxymethyl-2′-deoxycytidine (5hm5C) and 5-formyl-2′-deoxycytidine (5f5C) are produced by ten-eleven translocation (TET) methylcytosine dioxygenases. Most cells can scavenge epigenome-related nucleosides and maintain genomic integrity, but cancer cells overexpressing CDA convert 5hm5C and 5f5C into variants of uridine, which accumulate in DNA and result in...
DNA damage and even cell death. Therefore, epigenome-related nucleosides may present a new strategy for targeting tumors overexpressing CDA.

We previously reported that enhancer remodeling and microRNA alterations can drive the epithelial-to-mesenchymal transition (EMT) and promote chemoresistance to ALK inhibitors. In the present study, we examined genome-wide changes in DNA methylation in acquired resistance to the anticancer drug ceritinib and identified CDA among the hypomethylated and upregulated genes in drug-resistant cells. Single-cell RNA sequencing of CD3122 and LR cells revealed CDA overexpressed in cells with acquired resistance. Finally, we found that treatment with epigenome-related nucleosides such as 5fdC are a promising therapeutic strategy for overcoming ALK inhibitor resistance in CDA-overexpressing NSCLC.

**MATERIALS AND METHODS**

**Chemicals**

Ceritinib (a.k.a., LDK378; S7083, Selleckchem, TX), ShmC (PY7589, Berry & Associates, MI), and 5fdC (PY7589, Berry & Associates) were dissolved in dimethyl sulfoxide (DMSO; D2650, Sigma–Aldrich, MO). Trihydroxuridine (S84222, Merck Millipore, NJ) was dissolved in water.

**Cell lines**

The EML4-ALK-positive NSCLC cell line H3122 was obtained as previously described in ref. 14 and cultured in RPMI-1640 medium (LM 011-01, Welgene, Gyeongsan, Korea) supplemented with 1% antibiotics (15240062, Gibco, NY) and 10% fetal bovine serum (5001-07, Welgene). Ceritinib (LDK378)-resistant (LR) cells were established previously14. Briefly, H3122 cells were cultured with increasing concentrations of ceritinib starting at the IC50 value. Resistant cells were obtained after ~6 months of culture with continuous 1 µM ceritinib treatment. All cells were maintained at 37 °C in a humidified atmosphere containing 5% CO2.

**Bulk RNA-seq analysis**

Bulk RNA-seq was performed as described in ref. 14. The mRNA-seq library was prepared using the TruSeq RNA Sample Prep kit (Illumina), and sequencing was performed using the Illumina HiSeq2000 platform to generate 100-bp paired-end reads. The sequence reads were then mapped to the human genome (hg19) using STAR (v.2.5.1) and gene expression was quantified using the count module in STAR. The edgeR package (v.3.12.1) was employed to select differentially expressed genes (DEGs) from RNA-seq count data. The “trimmed mean of M-values” normalized value for each gene (in counts per million, cpm) was set at 1 and Log2 transformed for further analysis.

Quantitative reverse transcription-PCR (qRT–PCR) RNA was isolated using the RNeasy Plus Mini kit (74136, Qiagen, CA) and assessed using a NanoDrop ND-1000 spectrophotometer (Agilent, CA). CDNA was synthesized from 1 µg RNA using the iScript cDNA Synthesis kit (1708890, Bio-Rad, CA). Real-time PCR was performed in triplicate for each sample according to the manufacturer’s instructions (170-8880AP, Bio-Rad). The value for each gene was normalized to the β-actin signal. Supplementary Table 1 lists the sequences of the primers used.

DNA methylation microarray analysis Genome-wide DNA methylation was analyzed in duplicate using the Infinium methylation 450K bead chip array (Illumina), and the resulting DNA methylated/unmethylated signal intensity data were imported into R (v.3.4.2) for analysis. The data were normalized using the subset-quantile within array normalization method with background correction with the minfi (v.1.30.0) R package. CpG methylation values were calculated as average β values. Measurements with β values < 0.2 were considered significant above background. To identify differentially methylated CpGs between H3122 and LR cells, statistical analysis was performed using the DMRcate R package (v.1.20.0). Differentially methylated CpGs with P < 0.05 and an average difference >15% were selected.

**Preparation and sequencing of the scRNA-seq library**

Samples were prepared as outlined in the 10x Genomics Single Cell 3’ v2 Reagent kit user guide. The single-cell RNA-seq (scRNA-seq) library was prepared using the Chromium Single Cell 3’ Library and Gel Bead Kit V2 (PN-120237), Chromium Single Cell 3’ Chip kit V2 (PN-120236), and Chromium I7 Multiplex kit (PN-120262) with the 10x Genomics Chromium instrument. Samples were sequenced using a HiSeq 2500 with the following run parameters: read 1, 26 cycles; read 2, 98 cycles; index 1, 8 cycles. A median sequencing depth of 60,000 reads/cell was targeted for each sample. Supplementary Table 2 lists the 10x Genomics web summaries for each sample profiled.

**Bisulfite sequencing**

Genomic DNA was isolated using the DNeasy Blood and Tissue kit (69506, Qiagen). Unmethylated cytosines were converted to uracil with sodium bisulfite using the EZ DNA Methylation-Gold kit (DS005, Zymo Research, CA). Modified DNA was amplified with primers targeting the CpG sites (cg04087271, cg20619374, and cg06984156) of CD4 (Supplementary Table 3). Gel-purified bands were extracted using a Gel Extraction kit (28706, Qiagen) and cloned into the pGEM-T Easy vector (A1360, Promega, WI). Multiple plasmid DNA was isolated using the HTS Plasmid kit (PHTS-30, Core Bio System, Korea), and Sanger DNA sequencing was performed by GenoTech (Daejeon, Korea).

Western blot analysis Western blotting was performed as described in ref. 14. Antibodies were diluted in 5% skim milk or 5% bovine serum albumin in Tris buffered saline containing 0.3% Tween-20. The antibodies and dilutions used were as follows: anti-CDA (SAB1300717, Sigma–Aldrich, 1:500), anti-E-cadherin (1:1000, 3195, CST, MA), anti-N-cadherin (1:1000, 4061, CST), anti-vimentin (1:1000, 5741, CST), anti-GAPDH (1:2000, sc-47724, Santa Cruz Biotechnology, CA), anti-PARP (1:1000, 5174, CST), goat anti-mouse IgG (1:5000, 31430, Invitrogen, CA), and goat anti-rabbit IgG (1:5000, 31460, Invitrogen). Immunopositive bands were detected using an ECL kit (K-12045-D50, K-12042-D10, Advansta, CA) and visualized using a Fujifilm LAS-4000 (Tokyo, Japan).
Preparation and sequencing of the scATAC-seq library

Samples were prepared as outlined in the 10x Genomics Single Cell ATAC Reagent kit v1.1 user guide. A single-cell assay of the transposase-accessible chromatin (scATAC) library was prepared using the Chromium Next GEM Single Cell ATAC Library and Gel Bead kit v1.1 (PN-1000175), Chromium Next GEM Chip H Single Cell kit v1.1 (PN-1000161), and Single Index kit (PN-1000212) with the 10x Genomics Chromium instrument. Samples were sequenced using a HiSeq 2500 with the following run parameters.
parameters: read 1, 50 cycles; read 2, 50 cycles; index 1, 8 cycles; index 2, 16 cycles. A median sequencing depth of 25,000 reads/nucleus was targeted. The 10x Genomics web summary can be found in Supplementary Table 5.

**In vivo tumorigenicity assay**

All animal experiments were conducted as approved by the Committee on Animal Experimentation of the Korea Research Institute of Bioscience & Biotechnology (KRIIB, Daejeon, Korea). Four-week-old female BALB/c nude mice were purchased from Orient Bio. LR cells (5 × 10^6) in matrigel suspension (354248, Corning) were subcutaneously injected into the flanks of 5-week-old mice. Tumor size was measured using Vernier calipers every 3 days. Tumor volume was calculated as volume (mm^3) = 0.522 × length × width^2. When the tumor size reached 100–150 mm^3, the mice were randomly divided into control and 5fdC groups. 5fdC was prepared in 20% polyethylene glycol 400 (807485, Sigma–Aldrich) and 3% Tween 80 (822187, Sigma–Aldrich) and administered at 100 mg/kg every 24 h for 13 days. Control mice were treated on the same schedule as vehicle.

**Immunohistochemistry**

Authorization for the use of paraffin-embedded human lung cancer samples for research purposes and ethical approval were obtained from the Institutional Review Board of Seoul National University Hospital (Seoul, Korea). Tissue sections for immunohistochemistry were cut at a 4-μm thickness from paraffin blocks of patient samples. The sections were stained with anti-CD8 (ab222515, Abcam) using a Benchmark XT autostainer (Ventana Medical Systems, AZ). CDA expression was determined using the tumor proportion score, which is the percentage of viable tumor cells showing cytoplasmic staining. All samples were independently reviewed by a pathologist (J.K.) in a blinded manner.

**Statistical analysis**

The number of biological replicates (n) is described in the figure legends. Sample size determination based on statistics was not applied in this study. Statistical analysis was performed using R software (v. 3.6.3). Data are presented as the mean ± SEM or SD, as indicated in figure legends. The researchers involved in this study were not completely blinded to the animal experiments but were blinded to the human data analyses. Statistical significance between the two groups was evaluated using Student’s t-test or the Mann–Whitney U-test, as appropriate. A P value of <0.05 was considered significant.

**RESULTS**

**Changes in the DNA methylome are associated with acquired resistance to ceritinib**

We previously established an in vitro model of ceritinib-resistant NSCLC and confirmed that the model had no genetic aspects that could potentially confound analysis, such as secondary mutations in the ALK tyrosine kinase domain. To explore changes in the DNA methylome and transcriptome during the development of acquired resistance to an ALK inhibitor in NSCLC, we performed 450 K BeadChip DNA methylation analysis and scRNA-seq of the EML4-ALK fusion-positive lung cancer cell line H3122 and the ceritinib-resistant cell line LR established (Fig. 1a, b). Previous bulk RNA-seq revealed dramatic differences in global gene expression between H3122 and LR cells (Fig. 1c) among the genes detected by RNA-seq (n = 10,519), 13% (n = 1,403) were upregulated and 14% (n = 1,535) downregulated in LR cells (fold change (FC) > 2, false-discovery rate (FDR) < 0.05; Fig. 1c). CYP4F11 (cytochrome P450 4F11), EDIL3 (EGF-like repeats and discoidin domains 3), and AXL (AXL receptor tyrosine kinase) were the most strongly upregulated genes, whereas DUSP6 (dual-specificity phosphatase 6) and GLB1L2 (galactosidase beta 1-like 2) were the most strongly downregulated genes. DNA methyltransferases and TETs were also downregulated in LR cells (Fig. 1d), suggesting that acquired resistance may involve changes in not only the transcriptome but also the DNA methylome.

To determine whether the transcriptional changes observed are associated with DNA methylation changes, we analyzed genome-wide methylation changes using the Infinium Human Methyltion 450 K array. A total of 23,426 CpGs were hypermethylated and 17,797 CpGs hypomethylated in LR cells (β value change > 0.15, P < 0.05; Fig. 1e). Integrated analyses of DNA methylation and gene expression revealed 1,002 genes, including XYLTI (xylosyltransferase 1) and DUSP6, to be hypermethylated and downregulated, and 809 genes, including ANKR2D (ankyrin repeat domain 2) and AXL, to be hypomethylated and upregulated (Fig. 1f). The hypermethylated and downregulated genes are associated with apoptosis, cell proliferation, and MAPK (mitogen-activated protein kinase) signaling; the hypomethylated and upregulated genes are associated with cell adhesion, cell migration, and Hippo signaling (Fig. 1g). These data suggest that transcriptional plasticity may drive stable epigenetic changes such as DNA methylation during the development of ceritinib resistance.

To examine differentially methylated and expressed genes more closely, we first focused on DUSP6. Reactivation of MAPK signaling is a hallmark of acquired resistance to ALK inhibitors in NSCLC, and decreased expression of DUSP6, a MAPK phosphatase, promotes resistance to ALK inhibitors. Interestingly, CpG sites in exon 3 of DUSP6, but not the promoter region, were heavily methylated in LR cells (Fig. 1h), suggesting that these sites may be critical regulatory regions for DUSP6 transcription. Other negative regulators of MAPK signaling, including SPRY1, DAB2IP, ARBR1, DMD, CNKSR3, PTPRR, NLRP12, WNK2, SLC9A3R1, ERF1L, and SPRY4, were also hypermethylated and downregulated in LR cells (Supplementary Fig. 1). AXL is a receptor tyrosine kinase associated with tumor cell proliferation, metastasis, EMT, and drug resistance. We previously showed that AXL activation during EMT is a primary feature of acquired resistance to ALK.
inhibitors. The AXL promoter region was demethylated in cells with acquired resistance (Fig. 1). These data suggest that drug-induced changes in DNA methylation have a key role in converting a transient transcriptional state to a stable resistant state.

scRNA-seq reveals CDA-overexpressing cells in both resistant and nonresistant cells

When attempting to identify tumor cell heterogeneity or driver cell populations, the traditional bulk RNA-seq method has limitations because it analyzes the gene expression profile of a
mixture of cells. Recently, scRNA-seq technologies have allowed the investigation of RNA expression differences on a cell-by-cell basis. To explore the heterogeneity of naive and drug-resistant cells, we performed scRNA-seq using 10x Genomics Single Cell 3’ Solution and obtained gene expression profiles for 9401 cells. Clustering analysis of scRNA-seq data divided H3122 cells (n = 4835) and LR cells (n = 4566) into 12 clusters based on uniform manifold approximation and projection (Fig. 2a). Clusters 10 and 5 consisted mostly of H3122 cells, whereas Clusters 2, 7, 8, and 11 included mostly LR cells. Most LR cells were in the G2/M or S phase of the cell cycle (Fig. 2b) and Supplementary Fig. 2a), whereas most H3122 cells were in G1, suggesting that LR cells are more proliferative than H3122 cells. Differential expression analysis with bulk RNA-seq data allows for the comparison of a limited number of biological replicates, but scRNA-seq can identify key players in subpopulations of cells. We identified 761 upregulated genes and 401 downregulated genes in LR cells (Log2FC > 0.25, FDR < 0.01; Fig. 2c). EMT, cell cycle, and drug metabolism pathways were enriched in LR cells (Supplementary Fig. 2b); response to endoplasmic reticulum (ER) stress and cell adhesion molecules were enriched in H3122 cells (Supplementary Fig. 2c). CDA was one of the most significantly upregulated genes in LR cells (Fig. 2c). Intriguingly, although CDA-overexpressing cells were detected mainly among LR cells, they were also found rarely (yet definitively) in the naive H3122 cell population (Fig. 2d, e). CDA is a nucleoside metabolism enzyme involved in the homeostasis of the cellular pyrimidine pool. CDA-overexpressing cells in the naive cell population showed enriched expression of cancer stem cell-related genes such as S100A10 (S100 calcium-binding protein A10)10, LGALS1 (galectin-1)11, and SH3BGR1L3 (SH3 domain-binding glutamate-rich protein-like)3 (Supplementary Table 6). These data show that CDA-overexpressing cells may possess a growth advantage during ceritinib treatment. CDA mRNA expression was increased 12-fold in LR cells, and methylation was reduced in the promoter and putative enhancer regions of CDA (Fig. 2f).

To investigate whether CDA expression is increased in acquired resistance to other tyrosine kinase inhibitors (TKIs), we analyzed bulk RNA-seq data for H3122 cells treated with crizotinib. Consistent with our observations of ceritinib resistance, CDA and AXL were upregulated and DUSP6 downregulated in crizotinib-resistant cells (Supplementary Fig. 2d). Furthermore, we analyzed scRNA-seq data for PC9 cells treated with erlotinib, an EGFR TKI, for 0, 1, 2, 4, 9, and 11 days. Astonishingly, CDA was upregulated even on Day 1, whereas AXL was upregulated only after 9 days (Fig. 2g). On Day 3 of erlotinib treatment, CDA expression was increased in the overall clusters, whereas AXL expression was decreased compared with Day 0 (Fig. 2h). These data demonstrate that CDA-overexpressing cells may have a growth advantage during TKI treatment, causing them to be selected during treatment and thereby contributing to acquired resistance.

We identified 25 common genes that showed a high correlation (R > 0.3) with CDA expression in our scRNA-seq data and increased expression in both ceritinib-resistant (Fig. 1c) and crizotinib-resistant cells (Supplementary Fig. 2d, i). These included genes involved in EMT (CAV1, CAV2, BAP1, COL1A1, COL1A2, COL5A1, ACTA2, ACTG1, TNFRSF1A, and TNFRSF1B), cell adhesion (CCN1 and CCN3), drug metabolism (CYP1A1, CYP1B1, CYP2C9, CYP2C18, CYP2J2, CYP2U1, and CYP3A5), and cell cycle (CDKN2A, CDKN2B, CDKN2C, CDKN2E, and CDKN2F). These genes may drive LR cell transformation. We hypothesized that CDA expression was increased in LR cells because CDA is a nucleoside metabolism enzyme involved in the homeostasis of the cellular pyrimidine pool. CDA-overexpressing cells were detected mainly among LR cells, they were also found rarely (yet definitively) in the naive H3122 cell population (Fig. 2d, e). Intriguingly, although CDA-overexpressing cells were detected mainly among LR cells, they were also found rarely (yet definitively) in the naive H3122 cell population (Fig. 2d, e). CDA is a nucleoside metabolism enzyme involved in the homeostasis of the cellular pyrimidine pool. CDA-overexpressing cells in the naive cell population showed enriched expression of cancer stem cell-related genes such as S100A10 (S100 calcium-binding protein A10)10, LGALS1 (galectin-1)11, and SH3BGR1L3 (SH3 domain-binding glutamate-rich protein-like)3 (Supplementary Table 6). These data show that CDA-overexpressing cells may possess a growth advantage during ceritinib treatment. CDA mRNA expression was increased 12-fold in LR cells, and methylation was reduced in the promoter and putative enhancer regions of CDA (Fig. 2f).

To investigate whether CDA expression is increased in acquired resistance to other tyrosine kinase inhibitors (TKIs), we analyzed bulk RNA-seq data for H3122 cells treated with crizotinib. Consistent with our observations of ceritinib resistance, CDA and AXL were upregulated and DUSP6 downregulated in crizotinib-resistant cells (Supplementary Fig. 2d). Furthermore, we analyzed scRNA-seq data for PC9 cells treated with erlotinib, an EGFR TKI, for 0, 1, 2, 4, 9, and 11 days. Astonishingly, CDA was upregulated even on Day 1, whereas AXL was upregulated only after 9 days (Fig. 2g). On Day 3 of erlotinib treatment, CDA expression was increased in the overall clusters, whereas AXL expression was decreased compared with Day 0 (Fig. 2h). These data demonstrate that CDA-overexpressing cells may have a growth advantage during TKI treatment, causing them to be selected during treatment and thereby contributing to acquired resistance.

Western blotting confirmed an increased level of cellular CDA in resistant cells (Fig. 3a). To elucidate the function of CDA in acquired resistance, we depleted CDA in LR cells using three different CDA siRNAs. All three siRNAs reduced both CDA mRNA and protein levels (Fig. 3b, c). Notably, CDA depletion reduced LR cell proliferation (Fig. 3d). In addition, treatment with tetrahydrouridine (THU) (0 μM) of LR cells (Fig. 3a).
scATAC-seq reveals gene regulatory networks in cells with acquired resistance

The accessible chromatin landscape of resistant LR cells was established using scATAC-seq, which yielded profiles for 7753 nuclei, with a median of 6821 fragments mapped per nucleus (Supplementary Table 5). To identify cis-regulatory elements controlling gene expression in LR cells, we performed an integrated analysis of scRNA-seq with scATAC-seq. Based on scRNA-seq data, scATAC-seq data for each cluster were predicted with a gene activity score using SnapATAC software (Fig. 4a). To
evaluate the relationship between open chromatin regions (OCRs) and gene expression, interaction scores were calculated by comparing the expression level of each gene with the chromatin accessibility of each OCR located within ±250 kb of the transcription start site (TSS). The OCRs were then categorized into four groups based on the interaction score: high, mid, low, and unrelated (Fig. 4b). The high group consisted of 13.5% of all OCRs and included more promoter regions than any other group. Gene Ontology analysis revealed that OCRs of LR cells are proximal to genes involved in cellular responses to stress, regulation of cell death, and cell migration (Fig. 4c). To dissect causal TFs responsible for the transcriptome profile of LR cells, we performed
TF motif analysis of OCRs using HOMER\textsuperscript{30}, revealing enrichment of TF-binding signatures such as the ATF3, BATF, and NFE2L1 sequence motifs (Fig. 4d).

With regard to CDA specifically, we identified 10 OCRs (R1-R10) within ±250 kb of the TSS significantly associated with CDA expression (Supplementary Fig. 3a, b). OCRs R3–R7 are located in the CDA gene body, and R6 displayed the most significant correlation with CDA expression (Fig. 4e and Supplementary Fig. 3c), suggesting that it may be a strong enhancer of CDA. Supplementary Fig. 3d, e lists the top TFs predicted to bind the 10 OCRs. Among the CDA-associated TFs (Supplementary Table 7), we focused on FOXM1, TEAD1, and SMAD3. FOXM1, a proliferation-specific TF, mediates EMT-associated EGFR TKI resistance\textsuperscript{31}, it promotes rapid cancer cell proliferation in small-cell lung cancer and is associated with poor prognosis\textsuperscript{32}. SMAD3 mediates transcriptional activation of EMT target genes in the TGFβ signaling pathway\textsuperscript{33}. TEAD1 is a key TF in various oncogenic signaling pathways, including the Hippo, Wnt, TGFβ, and EGFR pathways, and plays critical roles in EMT, metastasis, drug resistance, and cancer stem cells\textsuperscript{34}. To determine the effect of these TFs on the expression of CDA, we depleted the TFs using siRNAs and found that depletion of any one TF reduced CDA expression (Fig. 4e and Supplementary Fig. 3a, b). OCRs R3 and R5 displayed the least significant correlation with CDA expression (Supplementary Fig. 3a, b). OCRs R1–R10 showed a decreasing trend in CDA expression with increasing distance from the CDA TSS (Supplementary Fig. 3a, b). OCRs R9 and R10 showed significant correlation with CDA expression (Supplementary Fig. 3a, b). OCRs R1–R10 displayed the most significant correlation with CDA expression (Supplementary Fig. 3a, b).

Taken together, these data indicate that an open chromatin structure may be formed in the promoter and enhancer regions of CDA at least in part via DNA demethylation in cells with acquired resistance. Furthermore, TFs such as TEAD1, SMAD3, and FOXM1 may be recruited to the regulatory region and induce over-expression of CDA, which promotes acquired resistance to ALK inhibitors (Fig. 4g).

5-Formyl-2′-deoxycytidine (5fdC) selectively ablates CDA-overexpressing cells

CDA converts ShmDC and 5fdC to ShmDU and 5fdU, respectively, both of which induce cytotoxicity when incorporated into DNA. Therefore, cytidine variants such as ShmDC and 5fdC have been suggested as drug treatments for CDA-overexpressing cancers\textsuperscript{12}. We examined whether ShmDC or 5fdC inhibits LR cell proliferation. As expected, treatment of LR cells with ShmDC or 5fdC decreased proliferation in a dose-dependent manner. Notably, high doses of 5fdC (10 µM) significantly decreased the proliferation of LR cells but not H3122 cells, suggesting that it selectively inhibits CDA-overexpressing resistant cell proliferation (Fig. 5a). We then examined the effect of 5fdC on the survival of LR cells and found that 5fdC attenuated their colony-forming ability (Fig. 5b) and increased apoptosis (Fig. 5c). To determine whether 5fdC causes DNA damage specifically in LR cells, we carried out immunofluorescence staining for γ-H2AX, a marker of double-stranded DNA breaks. LR cells treated with 10 µM 5fdC showed a 2.7-fold increase in the number of cells with DNA damage (Fig. 5d). To investigate the effect of 5fdC on tumor growth, a xenograft assay was carried out using LR cells (Fig. 5e). Beginning on Day 13 after injection of LR cells, 5fdC (100 mg/kg) was administered each day by intraperitoneal injection. Compared with control mice, 5fdC-treated mice did not show significant changes in body weight but did exhibit reductions in tumor weight (Fig. 5e). Furthermore, tumors in 5fdC-injected mice showed a decrease in the number of proliferating cells and an increase in the number of cells with DNA damage (Fig. 5f). These results demonstrate that CDA-overexpressing resistant cells are vulnerable to 5fdC owing to the accumulation of DNA damage and that 5fdC use is a potential strategy for overcoming ceritinib resistance.

Clinical relevance of CDA in lung cancer patients

To explore the clinical relevance of CDA in lung cancer, we analyzed CDA expression in lung adenocarcinomas using GEPIA\textsuperscript{35}, a tool for analyzing the Cancer Genome Atlas (TCGA) and Genotype-Tissue Expression (GTEX) databases. Although median CDA expression was similar between tumor samples (n = 483, obtained from TCGA) and normal tissue samples (n = 347, obtained from GTEX), ~25% of the tumor samples overexpressed CDA compared with normal samples (Fig. 6a). In addition, the overall survival of patients with lung adenocarcinoma with high CDA expression was significantly lower than that of patients with low CDA expression (Fig. 6b).

To determine whether CDA expression is associated with the EMT signature in lung adenocarcinoma, we investigated EMT scores in low and high CDA expression groups by surveying a 16-gene signature of canonical EMT from TCGA datasets\textsuperscript{36}. Notably, the high CDA expression group had significantly higher EMT scores than the low group (Fig. 6c). Thus, CDA appears to have a key role in EMT during lung cancer progression.

We further examined CDA expression in primary cancer cells from NSCLC patients with ALK rearrangement who exhibited acquired resistance to crizotinib\textsuperscript{37}. Notably, CDA expression was higher in patient-derived resistant cells (SNU-2535, -2550, -2563) than in naïve ALK-positive NSCLC patient-derived cells (SNU-3166) at both mRNA and protein levels (Fig. 6d and Supplementary Table 8). We obtained NSCLC tumor biopsies from 11 patients with EML4-ALK rearrangements before and/or after ALK inhibitor therapy. Contrary to our expectation, immunohistochemistry of CDA showed no significant difference between pre- and post-ALK inhibitor therapy (Fig. 6e, left), which might be attributable to inter- or intratumoral heterogeneity of CDA expression even in the same patient\textsuperscript{38}. Representative CDA expression in tumors from pre- or post-TKI therapy is shown in Fig. 6e.

DISCUSSION

Proliferating cancer cells undergo metabolic adaptations to survive in the harsh tumor microenvironment\textsuperscript{39,40}. Therefore, targeting cancer-specific metabolism may be an effective therapeutic strategy\textsuperscript{6,41}. Our current results demonstrate epigenetic activation of CDA during the development of ceritinib-resistant cells. In this study, we identified CDA as a potential target for overcoming ALK inhibitor resistance in lung cancer. Our results suggest that CDA-overexpressing resistant cells are vulnerable to 5fdC owing to the accumulation of DNA damage and that 5fdC use is a potential strategy for overcoming ceritinib resistance.
resistance in NSCLC with EML4-ALK fusion. scRNA-seq analysis identified CDA upregulation as one of the primary characteristics of ALK inhibitor resistance. CDA-overexpressing cells have a growth advantage during ceritinib treatment and are thus become selected and propagated, contributing to acquired resistance. Decreased methylation of the CDA promoter and enhancer, along with the recruitment of EMT-related TFs, can at least partially explain the increased expression of CDA in resistant cells. We propose that targeting CDA-directed metabolism with epigenome-related nucleosides such as 5fdC represents a new strategy for ablating ALK inhibitor-resistant cells via accumulation of DNA damage leading to cell death (Fig. 6f).
As de novo pyrimidine biosynthesis is an energetically expensive pathway for cell growth and development, CDA may provide an energetically efficient bypass for rapidly proliferating LR cells via a salvage pathway using either intracellular nucleic acid degradation products or extracellular nucleosides. Consistent with this expectation, LR cells showed 1.2-fold more rapid cell growth than H3122 cells (Fig. 5a) and were mainly in the G2 M or S phase of the cell cycle (Fig. 2b). We also found that inhibition of CDA using siRNA or THU reduced LR cell proliferation (Fig. 3). Many studies have shown that reprogramming of pyrimidine metabolism is closely related to cancer progression. Beyond DNA and RNA biosynthesis, recycling of cytidine and uridine is involved in phospholipid synthesis for cell membrane biogenesis and protein glycosylation for cell–cell and cell–matrix adhesion (Fig. 6f). Furthermore, the pyrimidine salvage pathway is involved in the ER stress response and may increase the adaptive capacity of cells to drug treatment.

CDA also has critical roles in EMT. CDA knockdown in LR cells reversed EMT and reduced cell migration and invasion (Fig. 3c). Clinically, CDA expression is associated with the EMT signature in lung cancer patients (Fig. 6c). EMT is a dynamic process in which tumor cells can occupy intermediate EMT states (partial EMT) and can revert to a more epithelial phenotype through the reverse process, i.e., the mesenchymal-to-epithelial transition. Epigenetic changes such as DNA methylation and histone modifications direct this dynamic process. We found the promoter and enhancer regions of CDA to be demethylated in cells with acquired resistance, forming an open chromatin structure to bind TFs such as SMAD3, TEAD1, and FOXM1 (Fig. 4). CDA expression was also linked to EMT-related genes such as CAV2, TXNRD1, and HISTH2B2 (Fig. 2). EMT has been associated with both metastasis and drug resistance. CDA is frequently overexpressed in many cancers, including pancreatic, stomach, testicular, and vaginal cancers. Moreover, CDA overexpression can mediate resistance to chemotherapy based on cytidine analogs such as gemcitabine. Based on our findings using cetinib-resistant cells, we searched public data to determine whether upregulation of CDA is associated with resistance to other anticancer drugs. Intriguingly, CDA upregulation has been found in cancer cells resistant to palbociclib (CDK 4/6 inhibitor), trametinib (MEK inhibitor), and gefitinib (EGFR inhibitor) (Supplementary Table 9). Therefore, it appears that CDA has the potential to promote cancer cell survival and proliferation during the presence of anticancer drugs. Although much biological and clinical validation is needed, targeting CDA with THU or epigenome-related nucleosides may enhance the effectiveness of current strategies for overcoming resistance to these targeted therapies.

Modified nucleosides that are common in the epigenome can disrupt the regulation of gene expression if they are recycled and incorporated into DNA, and thus incorporation must be prevented in most healthy cells. In the case of 5-methyl-2′-deoxycytidine (5MdC), it can be recycled in a different form, i.e., deoxoythymidine triphosphate (dTTP), through deamination. In contrast, the oxidized epigenetic nucleosides 5HmdC and 5Fdc cannot be converted to canonical nucleotides in normal cells because cytidine monophosphate kinase 1 (CMPK1) phosphorylates only unmodified dCMP. In CDA-overexpressing cancer cells, however, 5HmdC and 5Fdc can be deaminated to yield 5HmdU and 5Fdu, respectively, which can be incorporated into DNA, leading to cell cycle arrest and/or accumulation of cytotoxic double-stranded DNA breaks that cause cell death. We demonstrated that administration of 5Fdc inhibits the proliferation of ALK inhibitor-resistant NSCLC cells in vitro and in vivo (Fig. 5). Although our in vivo study was carried out with a very small number of mice, 5Fdc selectively abolished CDA-overexpressing cells with no adverse effects on the animals. Zauri et al. have demonstrated the safety of 5HmdC and 5Fdc administration based on a lack of a change in behavior or body weight in mice. They also showed a lack of histological side effects in CDA-expressing organs such as the kidney and intestine and suggested that the cytotoxic thresholds of 5HmdC and 5Fdc would be reached only in highly proliferating and CDA-overexpressing cells. Nonetheless, more extensive studies are needed to fully establish the efficacy and safety of 5HmdC and 5Fdc as drugs for the treatment of cancer.

Taken together, this study provides proof of concept that single-cell transcriptome analysis can identify key players in acquired resistance to cancer therapy and that metabolic mechanisms may provide a vulnerability in such cancer cells that can be utilized to overcome resistance to targeted therapies. In this study, we focused on CDA because CDA-overexpressing cells were present in the naïve cell population and were propagated during ceritinib treatment. Future studies will be required to dissect the details of the role played by CDA in the EMT and the survival of therapy-resistant cancer cells during targeted therapy.

**DATA AVAILABILITY**

RNA-seq data were deposited in the NCBI Gene Expression Omnibus (GEO) under accession number GSE81484. Raw sequence tags were deposited in the NCBI Short Read Archive under accession number SRP075253. scRNA-seq and DNA methylation data were deposited in GEO under accession number GSE139388. scATAC-seq data were deposited in GEO under accession number GSE81484. The raw data for this study were deposited in the Korean Nucleotide Archive (KoNA) under accession number PRJKA210059.

**REFERENCES**

1. Morris, S. W. et al. Fusion of a kinase gene, ALK, to a nucleolar protein gene, NPM, in non-Hodgkin’s lymphoma. Science 263, 1281–1284 (1994).
2. Lin, J. J., Reilly, G. J. & Shaw, A. T. Targeting ALK: precision medicine takes on drug resistance. Cancer Discov. 7, 137–155 (2017).
3. Soda, M. et al. Identification of the transforming EML4-ALK fusion gene in non-small-cell lung cancer. Nature 448, 561–566 (2007).
4. Shave, A. T. et al. Clinical features and outcome of patients with non-small-cell lung cancer who harbor EML4-ALK. *J. Clin. Oncol.* 27, 4247–4253 (2009).

5. Dardaei, L. et al. SHP2 inhibition restores sensitivity in ALK-rearranged non-small-cell lung cancer resistant to ALK inhibitors. *Nat. Med.* 24, 512–517 (2018).

6. Shaffer, S. M. et al. Rare cell variability and drug-induced reprogramming as a mode of cancer drug resistance. *Nature* 546, 431–435 (2017).

7. Flicek, P., Li, W., Gaskell, E. & Bernstein, B. E. Epigenetic plasticity and the hallmark of cancer. *Science* 357, eaal2380 (2017).

8. Khalig, M. & Fallahi-Sichani, M. Epigenetic mechanisms of escape from Braf oncogene dependency. *Cancer* 111, 1480 (2019).

9. Lane, A. N. & Fan, T. W. Regulation of mammalian nucleotide metabolism and biosynthesis. *Nucleic Acids Res.* 43, 2466–2485 (2015).

10. Cacciamani, T. et al. Purification of human cytidine deaminase: molecular and enzymatic characterization and inhibition by synthetic pyrimidine analogs. *Arch. Biochem. Biophys.* 290, 285–292 (1991).

11. Branco, M. R., Ficz, G. & Reik, W. Uncovering the role of 5-hydroxymethylcytosine in the epigenome. *Nat. Rev. Genet.* 13, 7–13 (2011).

12. Zauri, M. et al. CDA directs metabolism of epigenetic nucleosides revealing a therapeutic window in cancer. *Nature* 524, 114–118 (2015).

13. Luengo, A., Gui, D. Y. & Vander Heiden, M. G. Targeting metabolism for cancer therapy. *Cell Chem. Biol.* 24, 1161–1180 (2017).

14. Yung, M. M. et al. Enhancer remodeling and microRNA alterations are associated with acquired resistance to ALK inhibitors. *Cancer Res.* 78, 3350–3362 (2018).

15. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29, 15–21 (2013).

16. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26, 139–140 (2010).

17. Wu, W., Haderk, F. & Bivona, T. G. Non-canonical thinking for targeting ALK-positive lung cancer. *Cancers* 9, 164 (2017).

18. Zhu, C., Wei, Y. & Wei, X. AXL receptor tyrosine kinase as a promising anti-cancer drug target. *Cancers* 10, 27255 (2018).

19. Shaw, A. T. et al. Clinical features and outcome of patients with non-small-cell lung cancer to transcriptional inhibition. *Cancer Res.* 78, 2465–2485 (2018).

20. Liang, S. K. et al. FOXM1 is required for small cell lung cancer tumorigenesis and sensitive to 5′-deoxy-5′-fluorocytidine (‘5’DFCR). *J. Exp. Clin. Cancer Res.* 40, 138 (2021).

21. Park, J. S. et al. Mechanical regulation of glycolysis via cytoskeleton architecture. *Nature* 578, 621–626 (2020).

22. Sivanand, S. & Wellen, K. E. Molecular biology: salvaging the genome. *Nature* 524, 40–41 (2015).

23. Lu, H. et al. Chemotherapy-induced S100A10 recruits KDM6A to facilitate OCT4-mediated breast cancer stemness. *Oncotarget* 9, 4330–4466 (2019).

24. Tang, Z. et al. GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. *Nucleic Acids Res.* 45, W98–W102 (2017).

25. Gibbons, D. L. & Creighton, C. J. Pan-cancer survey of epithelial-mesenchymal transition markers across the cancer genome atlas. *Dev. Dyn.* 247, 555–564 (2018).

26. Kim, S. et al. Heterogeneity of genetic changes associated with acquired crizotinib resistance in ALK-rearranged lung cancer. *J. Thorac. Oncol.* 8, 415–422 (2013).

27. Gao, Y. et al. Chemotherapy-induced CDA expression renders resistant non-small cell lung cancer cells sensitive to 5′-deoxy-5′-fluorocytidine (‘5’DFCR). *Front. Oncol.* 11, 684961 (2021).

28. Simmonds, H. A. Enzymes of nucleotide biosynthesis: differences between intact and lysed cells as well as between species and tissues can be important. *Biochem. Soc. Trans.* 23, 877–879 (1995).

29. Pedroza-Garcia, J. A. et al. Role of pyrimidine salvage pathway in the maintenance of organellar and nuclear genome integrity. *Plant J.* 97, 430–446 (2019).

30. Wang, W., Cui, J., Ma, H., Lu, W. & Huang, J. Targeting pyrimidine metabolism in the era of precision cancer medicine. *Front. Oncol.* 11, 684961 (2021).

31. Nilsson, M. B. et al. A YAP/FOXM1 axis mediates EMT-associated EGFR inhibitor resistance in ALK-positive lung cancer. *Nat. Rev. Drug Discov.* 19, 671–684 (2020).

32. Shin, G. et al. GENT: gene expression database of normal and tumor tissues. *Cancer Info.* 10, 149–157 (2011).

33. Neff, T. & Blau, C. A. Forced expression of cytidine deaminase confers resistance to cytosine arabinoside and gemcitabine. *Exp. Hematol.* 24, 1340–1346 (1996).

34. Lypova, N. et al. Targeting palbociclib-resistant estrogen receptor-positive breast cancer cells via oncolytic virotherapy. *Cancers* 11, 684 (2019).

35. Wagner, S. et al. Suppression of interferon gene expression overcomes resistance to MEK inhibition in KRAS-mutant colorectal cancer. *Oncogene* 38, 1717–1733 (2019).

36. Boumahdi, S. & de Sauvage, F. J. The great escape: tumour cell plasticity in resistance to targeted therapy. *Nat. Rev. Drug Discov.* 19, 39–56 (2020).

37. Hay, S. et al. Epigenetic plasticity and the hallmarks of cancer. *Science* 328, 1157–1161 (2010).

38. H.Heo et al. approved the manuscript for publication. H.H., Jong-Hwan K., and M.K. conceived the in vitro model. H.H., Jong-Hwan K., and M.K. wrote the manuscript. All authors read and approved the final manuscript.

ACKNOWLEDGEMENTS

This work was supported by National Research Foundation of Korea (NRF) grants funded by the Korean government [NRF-2017M3C9A5028693 and 2019R1A2C1087104 to M.K.; Systemic Industrial Infrastructure Projects through the Ministry of Trade, Industry and Energy (MOTIE) [P0009796, 2019 to S.-Y.K.]; the University of Science and Technology [2020YS13 to H.H.]; and the Korea Research Institute of Bioscience & Biotechnology (KORIB) Research Initiative Program [KGM5192113 to M.K.].

AUTHOR CONTRIBUTIONS

H.H., Jong-Hwan K., S.-Y.K., and M.K. generated the hypothesis and designed the study. H.H., H.J.L., Jeong-Hwan K., and Y.-J.S. performed the experiments. M.K. and J.K. performed immunohistochemistry and contributed to its data analysis. J.-Y.I., B.-K.K., and M.W. contributed to the cell proliferation and cell death analyses. H.H., Jong-Hwan K., H.J.L., H.K.H., S.-Y.K., and M.K. contributed to the data analyses. M.R.Y. and B.C.C. conceived the in vitro model. H.H., Jong-Hwan K., and M.K. wrote the manuscript with contributions from all authors. M.K. and S.-Y.K. supervised the research. M.K. and S.-Y.K. are the guarantors of this work. All authors have read and approved the final manuscript.

COMPETING INTERESTS

The authors declare no competing interests.
