scientific reports

Check for updates

OPEN

Genome-wide methylation profiling reveals extracellular vesicle DNA as an ex vivo surrogate of cancer cell-derived DNA

Kyung-A Kim[1,2,8], Sunmin Kim[3,4,5,8], Inbal Wortzel[6,8], Suho Lee[1], Yoon Dae Han[7], Tae-Min Kim[3,4,5]✉ & Han Sang Kim[1,2]✉

Extracellular vesicle-derived DNA (evDNA) encapsulates the complete genome and mutational status of cells; however, whether cancer cell-derived evDNA mirrors the epigenetic features of parental genomic DNA remains uncertain. This study aimed to assess and compare the DNA methylation patterns of evDNA from cancer cell lines and primary cancer tissues with those of the nuclear genomic DNA. We isolated evDNA secreted by two cancer cell lines (HCT116 and MDA-MB-231) from various subcellular compartments, including the nucleus and cytoplasm. Additionally, we obtained evDNA and nuclear DNA (nDNA) from the primary cancer tissues of colon cancer patients. We conducted a comprehensive genome-wide DNA methylation analysis using the Infinium Methylation EPIC BeadChip, examining > 850,000 CpG sites. Remarkable similarities were observed between evDNA and nDNA methylation patterns in cancer cell lines and patients. This concordance extended to clinical cancer tissue samples, showcasing the potential utility of evDNA methylation patterns in deducing cellular origin within heterogeneous populations through methylation-based deconvolution. The observed concordance underscores the potential of evDNA as a noninvasive surrogate marker for discerning tissue origin, particularly in cancer tissues, offering a promising future for cancer diagnostics. This finding enhances our understanding of cellular origins and would help develop innovative diagnostic and therapeutic strategies for cancer.

Keywords DNA methylation, Extracellular vesicles, Exosome, EV-DNA, Epigenetic modifications, Cancer

Epigenetic alterations are considered among the earliest genomic changes in carcinogenesis[1]. DNA methylation, a crucial epigenetic process, primarily occurs at cytosine bases within cytosine-guanine dinucleotides (CpGs)[2]. It plays a pivotal role in regulating several cellular processes, such as gene expression, genomic imprinting, and maintaining genomic stability[3,4]. Cancer is often characterized by distinct aberrant DNA methylation patterns[5,6], involving widespread hypomethylation and localized hypermethylation of specific genomic regions[7,8]. The stability and detectability of DNA methylation across various sample types, including tumor tissues and body fluids, further underscore its significance[9–11]. Accordingly, leveraging DNA methylation patterns as biomarkers has emerged as a promising strategy for improving cancer diagnosis[12].

Extracellular vesicles (EVs), membrane-bound entities with 30–150 nm size, are released by virtually all cell types and are ubiquitous in bodily fluids[13]. They play a crucial role in intercellular communication by transferring various cellular constituents into recipient cells[14–16]. Accumulating evidence establishes a connection between

[1]Department of Internal Medicine, Graduate School of Medical Science, Brain Korea 21 Project, Yonsei University College of Medicine, Seoul, Republic of Korea. [2]Division of Medical Oncology, Department of Internal Medicine, Yonsei Cancer Center, Yonsei University College of Medicine, 50-1 Yonsei-Ro, Seodaemungu, Seoul 03722, Republic of Korea. [3]Department of Medical Informatics, College of Medicine, Catholic University of Korea, 222, Banpo-Daero, Seocho-Gu, Seoul 06591, Republic of Korea. [4]College of Medicine, Cancer Research Institute, The Catholic University of Korea, Seoul, Republic of Korea. [5]Department of Biomedicine and Health Sciences, Graduate School, The Catholic University of Korea, Seoul, Republic of Korea. [6]Children's Cancer and Blood Foundation Laboratories, Department of Pediatrics, Cell and Developmental Biology, Drukier Institute for Children's Health, Meyer Cancer Center, Weill Cornell Medicine, New York, NY, USA. [7]Division of Colorectal Surgery, Department of Surgery, Yonsei University College of Medicine, Seoul, Republic of Korea. [8]These authors contributed equally: Kyung-A Kim, Sunmin Kim and Inbal Wortzel. ✉email: tmkim@catholic.ac.kr; modeerfhs@yuhs.ac

EVs and several critical aspects of tumor biology, including tumor growth[17], metastasis[18,19], and immune regulation[20,21]. Tumor-specific biomolecules such as DNA, microRNAs, and proteins have been identified in EVs in cell cultures and liquid biopsy specimens from patients with cancer[16,22]. Notably, double-stranded DNA (dsDNA) within EVs mirrors the mutational landscape of the originating tumor[23,24]. Previous studies have demonstrated that extracellular vesicle DNA (evDNA) in human plasma can reveal the mutational status at early stages[25–28]. Our hypothesis posits that the methylation patterns of evDNA recapitulate those of the cells of origin, providing an ex vivo resource for DNA methylation analysis.

We conducted a genome-wide analysis of DNA methylation to examine the methylation status of evDNA, cytoplasmic DNA (cytoDNA), and nuclear DNA (nDNA) derived from human cancer cell lines. Furthermore, we compared the methylation patterns of evDNA with gDNA from cancer patients. Given the growing interest in utilizing extracellular vesicles for liquid biopsies and the feasibility of performing methylation analysis in body fluids, our study offers valuable insights into the cellular origin of cancer diagnostics.

Results

Characterization of EVs and subcellular fractions

We investigated whether cancer cell-derived evDNA exhibits methylation patterns similar to the original cells. Initially, we prepared whole cells, cytoplasmic and nuclear fractions, and EVs from the same cell lines (Fig. 1A) using human colorectal cancer (CRC; HCT116) and human breast cancer (MDA-MB-231) cell lines.

Transmission electron microscopy (TEM) showed that EVs exhibited a cup-shaped morphology and measured approximately 50–150 nm in size (Fig. 1B). The characterization of EVs involved assessing their size distribution through nanoparticle tracking analysis (NTA). The most frequent sizes (mode sizes) of EVs from HCT116 and MDA-MB-231 cells were 119 nm and 121 nm, respectively, within the overall size range of EVs (Fig. 1C).

Furthermore, we evaluated the EV expression of tetraspanins (CD9, CD81, and CD63), commonly expressed on EV surfaces. Fluorescence microscopy confirmed the expression of these tetraspanins in the EVs from both cell lines (Fig. 1D). Western blot analysis was conducted to ensure the purity of the isolated fractions and EVs (Fig. 1E), with Lamin B1 serving as a marker for nuclear fractions, CD9 for EVs, and GAPDH as the loading control. These results affirm the successful separation and purity of each fraction, rendering them suitable for downstream analyses.

Subsequently, we isolated DNA from each fraction and analyzed DNA size and distribution using the TapeStation system (Fig. 1F). The analysis revealed that the main peaks of evDNA were predominantly within the 250 to 15 kilobase pairs (kbp) range. In contrast, the main peaks of total DNA, nDNA, and cytoDNA were predominantly larger than approximately 60 kbp. These findings underscore a distinct size distribution of evDNA compared to total DNA and nDNA.

Genome-wide DNA methylation patterns of evDNA and cellular DNAs

To determine the DNA methylation profiles of evDNA and cellular DNA (total DNA, nDNA, and cytoDNA), we performed a bead array-based CpG methylation assay to interrogate 850,000 CpG sites for methylation. The CpG methylation levels in all types of DNA displayed a bimodal distribution (Fig. 2A), with the majority classified as either 'hypomethylated' (normalized β-values < 0.25) or 'hypermethylated' (normalized β-values > 0.75), constituting 24.8% and 54.7% of probes, respectively. The normalized β-value density distribution of evDNA resembled that of total DNA and nDNA in both cell lines, suggesting that the methylation patterns of evDNA and nDNA are comparable. Hypermethylated regions showed a slight increase in the cytoDNA compared to those in the total DNA, nDNA, and evDNA in both cell lines. Additionally, we observed the enrichment of hypomethylated CpGs in HCT116 cells compared to that in MDA-MB-231 cells (62.2% and 29.4%, respectively), indicating a significant intercellular difference in DNA methylation (Fig. 2A).

Subsequently, we categorized CpGs based on their proximity to neighboring CpG islands: islands (regions within 2 kb from a CpG island), shores (regions up to 2 kb from a CpG island), and shelves (regions from 2 to 4 kb from a CpG island). We then examined their methylation levels (Fig. 2B). CpG islands tended to be hypomethylated across both cell lines compared to regions outside the CpG islands. Furthermore, we scrutinized differential methylation patterns in the promoter-proximal regions, including gene bodies, TSS1500 (1500 bp upstream of the transcription start site [TSS]), TSS200 (200 bp upstream of TSS), the first exon, and the 5′ untranslated region (UTR). Our findings revealed that at TSS1500, TSS200, the first exon, and the 5′ UTR, all DNA types were more hypomethylated than gene bodies (Fig. 2C). These observations suggest that methylation levels are influenced more by the location of CpG islands and specific CpG sites than by the differences between cell lines or cellular DNA components.

Concordance of DNA methylation between evDNA and nDNA

Next, we performed correlative analyses to investigate the concordance in CpG methylation profiles. Hierarchical clustering analysis of DNA methylation across cellular compartments revealed that evDNA methylation closely resembled that of total DNA and nDNA rather than cytoDNA. Pearson's correlation coefficient (R) measured 0.99 in HCT116 cells and 1.00 in MDA-MB-231 cells between evDNA and nDNA (Fig. 3A). The methylation pattern demonstrated a high correlation between evDNA and nDNA in the HCT116 (R = 0.991) and MDA-MB-231 (R = 0.997) cell lines (Fig. 3B). Furthermore, we observed notable similarities between evDNA and nDNA methylation across various types of CpG islands, including islands, shores, and shelves (Supplementary Fig. 1A,B). High correlations were also evident in promoter-proximal regions, such as gene bodies, TSS1500, TSS200, the first exon, exon boundary, and the 3′ and 5′ UTRs (Supplementary Fig. 1C,D). These findings indicate a substantial concordance in the DNA methylation patterns between evDNA and nDNA.

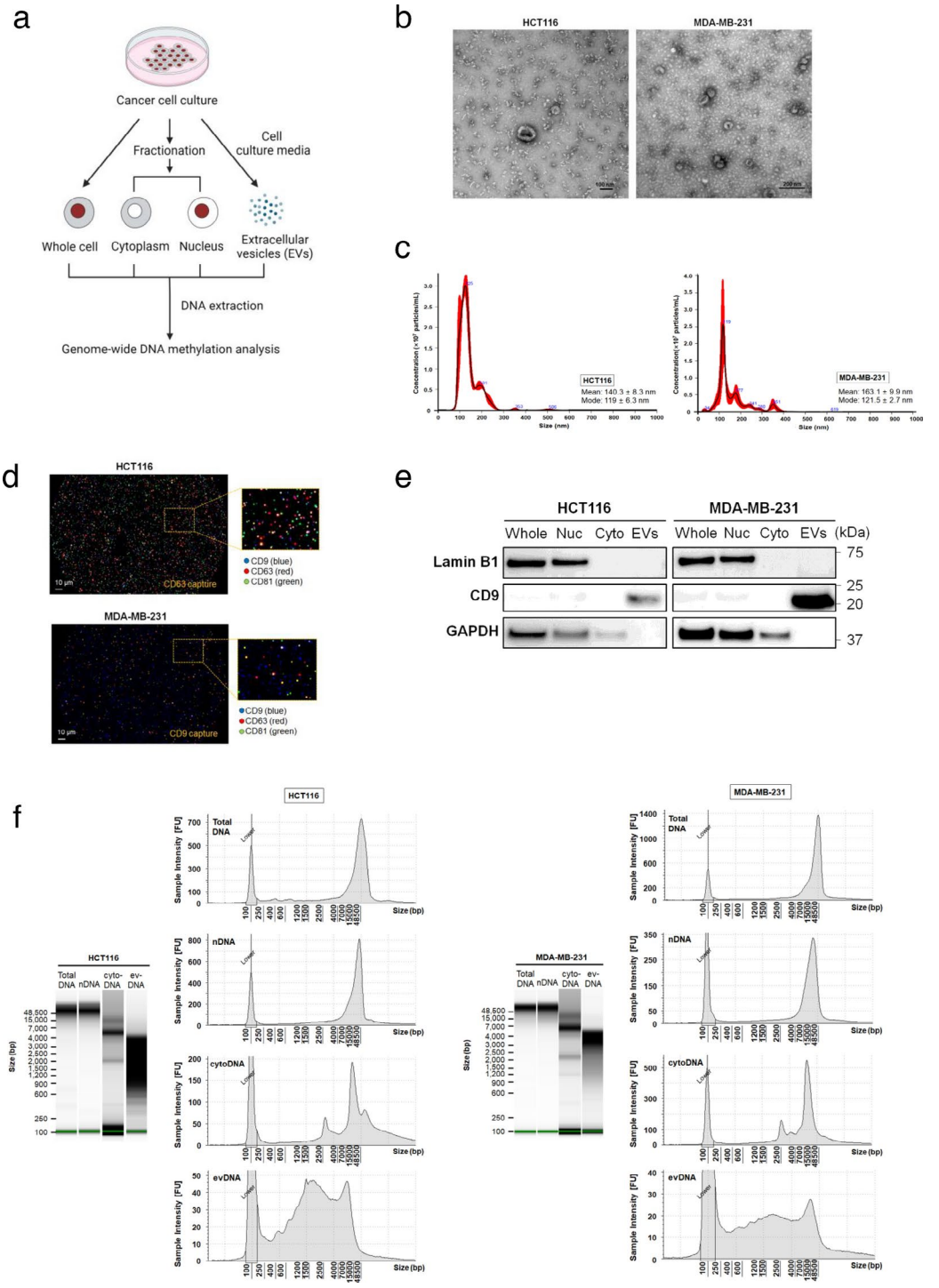


Fig. 1. EV and cellular fraction characterization. (**a**) Methodological workflow for this study. (**b**) Transmission Electron Microscopy (TEM) analysis of extracellular vesicles (EVs)-derived from HCT116 and MDA-MD-231 cell lines. (**c**) Particle size distribution of EV samples derived from HCT116 (top) and MDA-MB-231 (bottom) cell line. Mean size, mode size (the most frequent size), and standard deviation (SD) values of EV samples were determined by NanoSight Tracking Analysis (NTA). (**d**) Representative fluorescence images of EV samples derived from HCT116 (left side) and MDA-MB-231 (right side) cell line. EV samples were captured by specific antibody-coated chips against anti-CD9 (blue), anti-CD63 (red), and anti-CD81 (green) antibodies. (**e**) Western blot analysis of whole cell lysates (whole), nucleic fraction (nuc), cytoplasmic fraction (cyto), and EVs from HCT116 and MDA-MB-231 cell lines. The membranes were probed for Lamin B1 (nuclear marker), CD9 (EV marker), and GAPDH (loading control). (**f**) Electropherograms and virtual gel images showing the profiles of total genomic DNA, nuclear DNA (nDNA), cytoplasmic DNA (cytoDNA), and extracellular vesicle-derived DNA (evDNA) from HCT116 (left side) and MDA-MB-231 (right side) cell lines.

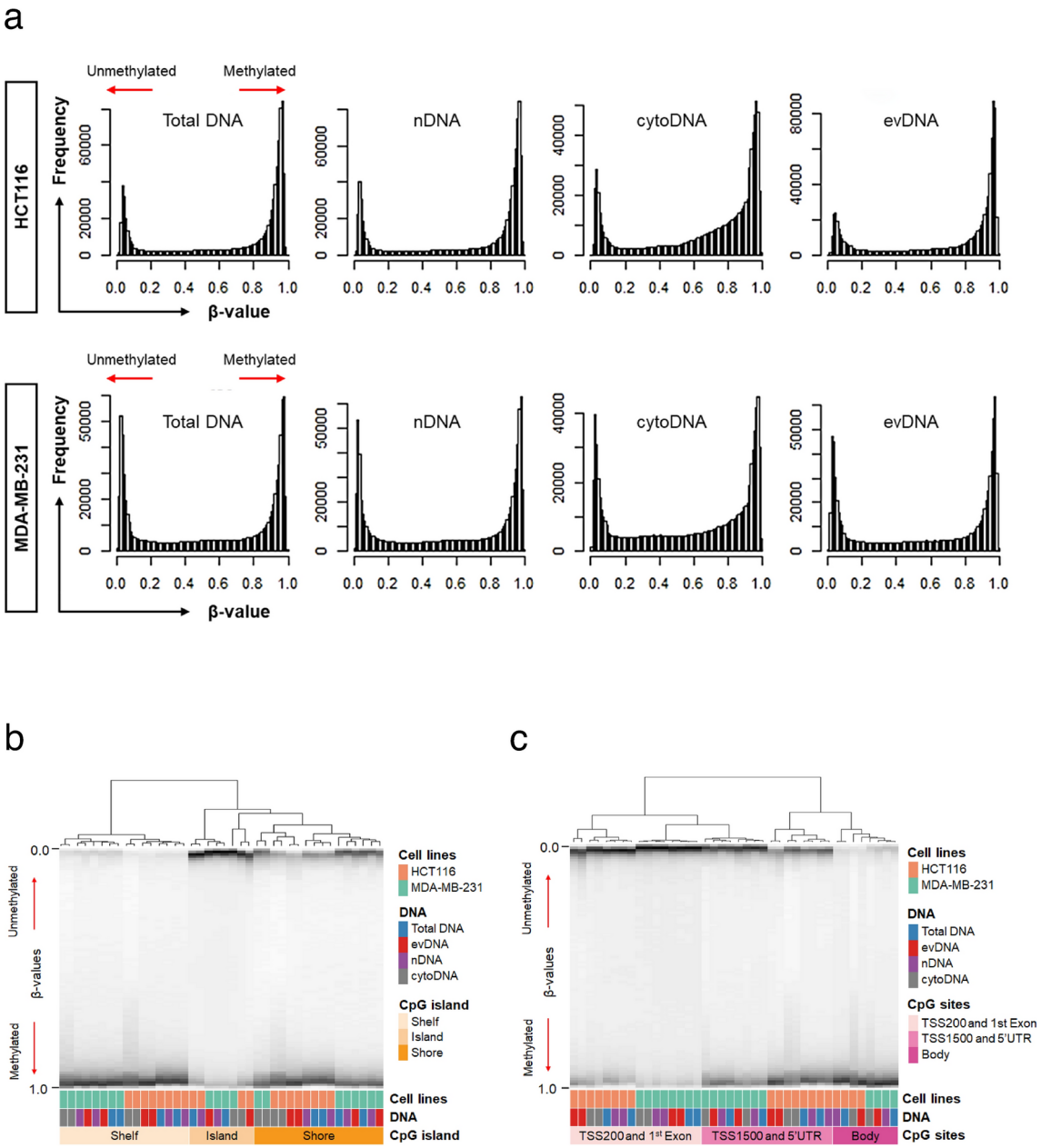


Fig. 2. Comparative DNA methylation levels in EV and cellular fractions. (**a**) Histograms illustrating the distribution of DNA methylation levels (β-value) for total DNA, nuclear DNA (nDNA), cytoplasmic DNA (cytoDNA), and extracellular vesicle DNA (evDNA) in HCT116 (top) and MDA-MB-231 (bottom) cell lines. Methylation levels are shown as normalized distributions of average β-values, with the two peaks indicating distinct methylation states: β-values below 0.2 represent 'unmethylated' regions, while β-values above 0.7 correspond to 'fully methylated' regions. (**b**) Unsupervised hierarchical clustering of CpG Islands based on β-values, revealing DNA methylation patterns across different DNA fractions in both cell lines. The heatmap illustrates differential methylation in total DNA (blue), evDNA (red), nDNA (purple), and cytoDNA (gray) for HCT116 (orange) and MDA-MB-231 (green). Annotations for CpG regions, including shelf (light orange), island (yellow-orange), and shore (orange), are provided at the bottom of the heatmap. (**c**) Unsupervised hierarchical clustering of individual CpG sites, similarly grouped by β-values. Annotations for CpG sites include transcription start site regions (TSS200 and 1st Exon in light pink, TSS1500 and 5'UTR in pink), and the gene body (dark pink), displayed at the bottom of the heatmap.

To elucidate the genomic and epigenetic configurations associated with additional DNA methylation changes acquired by evDNA, we analyzed the differential methylation levels between evDNA and nDNA according to DNA functional regions, annotated by the Encyclopedia of DNA Elements (ENCODE). (Fig. 3C). Our study revealed that both evDNA and nDNA of cell lines exhibited similar methylation patterns in enhancer regions (e.g., 'weak/poised enhancer' and 'strong enhancer' states). This pattern was consistently observed across DNA

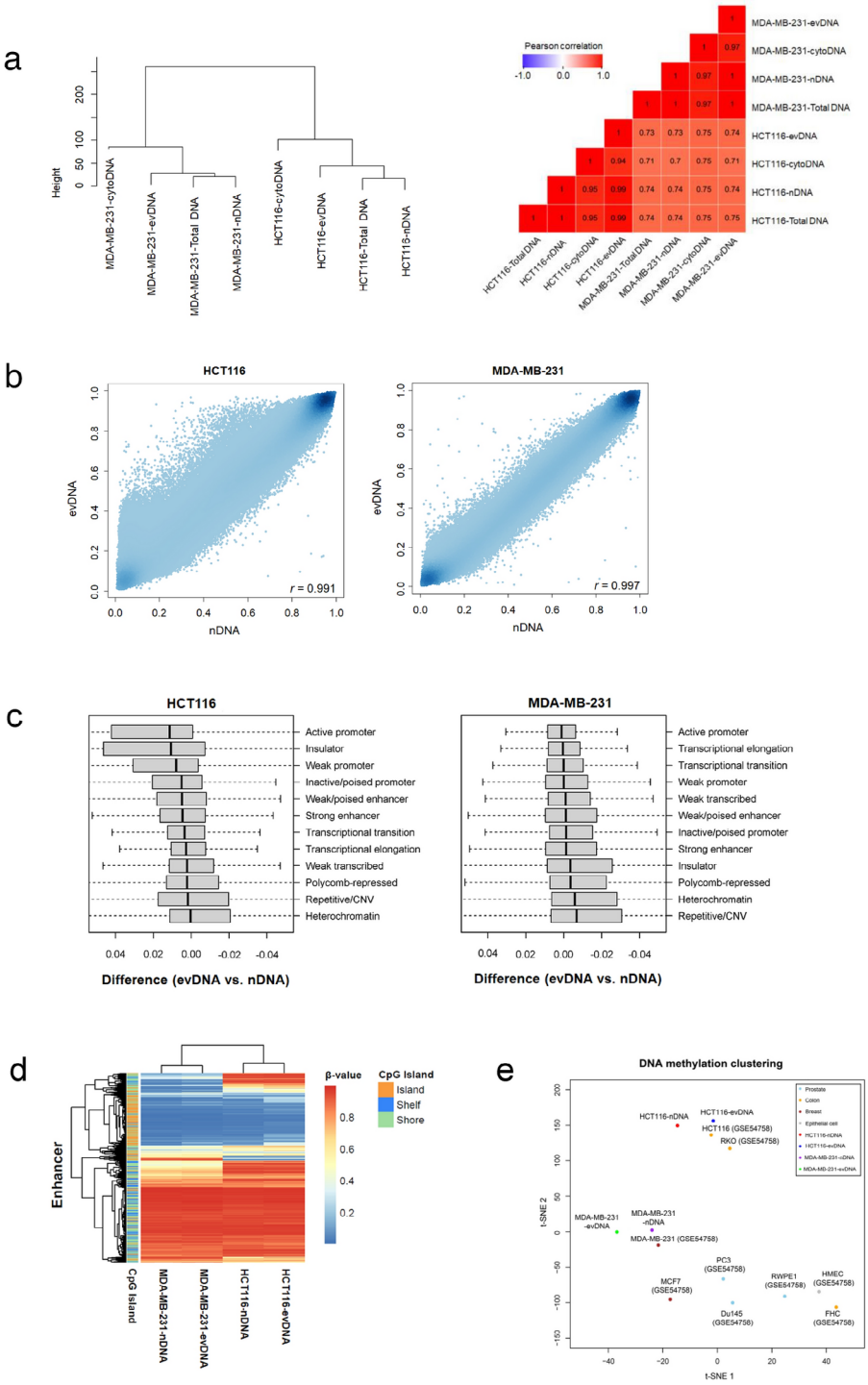


Fig. 3. Concordance in DNA methylation patterns in cancer cell lines. (**a**) Dendrogram of hierarchical clustering (left side) in total DNA, nDNA, cytoDNA, and evDNA from HCT116 and MDA-MB-231 cell lines. Heatmap of pairwise correlation of methylation profiles (right side). The heatmap colors indicate the correlation coefficients based on the pairwise Pearson correlation of methylation levels in total DNA, nDNA, cytoDNA, and evDNA from HCT116 and MDA-MB-231 cell lines. (**b**) Scatter plots of methylation levels between nDNA and evDNA from HCT116 (left side) and MDA-MB-231 (right side) cell lines. (**c**) The difference in the 12-DNA functional state using the ChromHMM model is based on the DNA methylation level of evDNA and nDNA from HCT116 and MDA-MB-231 cell lines. (**d**) Hierarchical clustering based on the β-values and enhancers at CpG island, including island, shelf, and shore in total nDNA and evDNA from HCT116 and MDA-MB-231 cell lines. (**e**) t-distributed stochastic neighbor embedding (tSNE) clustering of DNA methylation patterns of evDNA and nDNA from HCT116 and MDA-MB-231 cell lines with reference methylation data. Reference data (GSE54758): colon cancer, HCT116, and RKO; breast cancer, MDA-MB-231, and MCF7; prostate cancer, PC3, and Du145; normal epithelial cells, RWPE1, HMEC, and FHC.

functional regions derived from cell lines of diverse lineages, including human embryonic stem cell line H1 (H1-hESC), hepatoblastoma cell line (HepG2), human mammary epithelial cells (HMEC), mononucleated muscle cells (HSMM), human umbilical vein endothelial cells (HUVEC), human myelogenous leukemia cell line (K562), normal human epidermal keratinocytes (NHEK), and normal human lung fibroblasts (NHLF) (Supplementary Fig. 2).

Additionally, we analyzed genomic regions, focusing on the enhancers of evDNA and nDNA within the HCT116 and MDA-MB-231 cell lines. Our findings indicated minor variations in DNA methylation levels at the enhancer regions between evDNA and nDNA (Fig. 3D). Furthermore, evDNA and nDNA exhibited some differences in the DNA methylation levels of CpG islands.

To further assess the extent of epigenetic similarities between nDNA and evDNA derived from the same cell line, we compared the methylation profiles of evDNA and nDNA with those of human cell lines of different lineages. We integrated our methylation data with publicly available data (Gene Expression Omnibus accession, GSE54758) and performed t-SNE analysis on the combined dataset. Nine human cell lines, including two breast cancer cell lines (MDA-MB-231 and MCF-7), two CRC cell lines (HCT116 and RKO), two prostate cancer cell lines (Du1-45 and PC-3), and three normal epithelial cell lines (RWPE-1, FHC, and HMEC), were used to assess epigenetic similarity. As shown in Fig. 3E, the DNA methylation profiles distinctly segregated the cell lines based on lineage, forming four clusters—colon, breast, prostate, and normal epithelial cells. Interestingly, the evDNA originating from HCT116 cells clustered together with CRC cell lines (HCT116 and RKO), whereas that derived from MDA-MB-231 cells clustered with breast cancer cell lines (MDA-MB-231 and MCF-7). These results support the hypothesis that evDNA, exhibiting epigenetic characteristics similar to nDNA, can potentially serve as a lineage-specific marker indicative of the cell of origin.

evDNA is located both externally and internally on the EV surface[29]. Next, we evaluated whether internalized evDNA has a similar methylation pattern to genomic DNA. To minimize the presence of external DNA, we treated EVs with DNase before DNA extraction, allowing for the extraction of internal evDNA. We applied a reference-based non-negative least-squares (NNLS) algorithm to estimate the cellular composition of DNA mixtures. The proportions of colon cancer were 41.4%, 41.2, and 42.0% in 116-gDNA, 116-eV-DNA, and 116-inter-evDNA (HCT116-internalized evDNA), respectively. We detected breast cancer proportions of 28.3%, 28.1%, and 28.6% in the 231-gDNA, 231-eV-DNA, and 231-inter-evDNA (MDA-MB-231-internalized evDNA) samples, respectively (Fig. 4). We also analyzed the DNA methylation distribution across gDNA, evDNA, and internalized evDNA (inter-evDNA) in HCT116 and MDA-MB-231 cells. Our results revealed similar patterns in the DNA methylation distributions of gDNA, evDNA, and internalized evDNA in HCT116 and MDA-MB-231 cells (Supplementary Fig. 3A,B).

Deconvolution of tumor composition using evDNA methylation patterns

To explore the potential utility of evDNA as a cancer diagnostic biomarker, we analyzed the methylation patterns of evDNA and nDNA sourced from the tissues of patients with CRC (Supplementary Fig. 4). Colon cancer tissue-derived EVs from CRC patients exhibited a characteristic round shape with diameters ranging from 50 to 200 nm by TEM and NTA. (Fig. 5A,B). Patient-derived EVs expressed EV marker CD9 and β-actin (Fig. 5C). A high concordance was observed between the methylation profiles of evDNA and nDNA extracted from CRC tumors

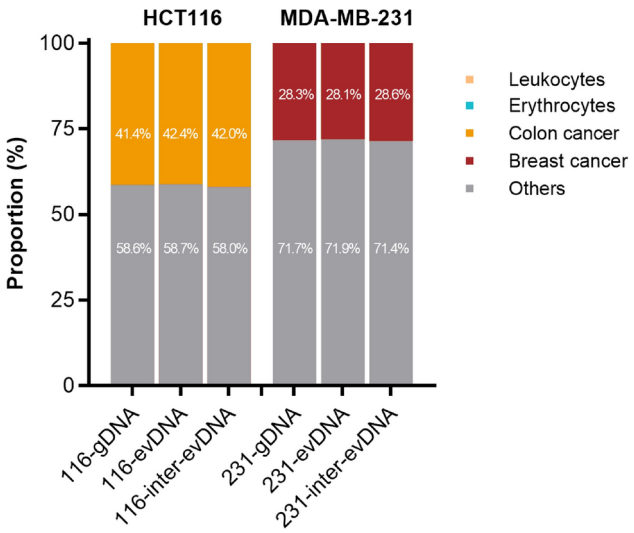


Fig. 4. Comparison of DNA methylation patterns between internalized evDNA and gDNA. (**a**) Bar plot of cell proportion across evDNA, internalized evDNA, and gDNA. DNA samples are colored by deconvoluted cell types: leukocytes (yellow), erythrocytes (mint), colon cancer (orange), breast cancer (burgundy), and others (grey). To obtain internalized evDNA, EVs isolated from HCT116 and MDA-MB-231 cells were subjected to DNase treatment, and the internalized DNA was subsequently extracted (116-inter-evDNA and 231- inter-evDNA).

(Fig. 5D). In the differential methylation analysis of CpG islands, we noted greater hypomethylation at CpG islands than at the regions outside these islands in primary tumors. Additionally, consistent hypomethylation was observed at TSS1500, TSS200, the first exon, and the 5′ UTR compared to that in gene bodies in both primary tumors (Fig. 5E), aligning with the in vitro findings depicted in Fig. 2B,C. To explore the possibility of differentiating between patient-derived nDNA proportions and evDNA based on methylation patterns in simulated heterogeneous tissue mixtures, we generated mimicking mixtures using a public dataset (GSE122126; setting of cell-free DNA). These mixtures, simulating heterogeneous tissues, contained immune cells, hepatocytes, and cancerous epithelial cells. The public dataset comprised 100 mixtures with defined proportions of immune cells (leukocytes and $CD45^+$) and hepatocytes, i.e., 96% and 4%, respectively. The composition of cancerous epithelial cells varied based on evDNA content (ranging from 0 to 50%; Fig. 5F); additionally, simulation data for different nDNA contents were also generated (ranging from 0 to 50%; Supplementary Fig. 5A,B). The simulation data demonstrated a high concordance between mixed evDNA and observed methylation (Fig. 5G). To further validate and evaluate the potential clinical applicability of evDNA, we collected blood samples from patients with colon cancer as well as from individuals with healthy control (Supplementary Fig. 4). The methylation patterns of genomic DNA extracted from tumor tissue, tumor-derived EVs, blood-derived cell-free DNA (cfDNA), and blood-derived EVs were compared using the NNLS method to analyze the different cell type proportions in the blood-derived samples. This analysis revealed that although evDNA from the blood of CRC patients exhibited only a low percentage of colon cancer cells (0.7–1.3%), it was substantially higher compared to healthy control (0.0–0.2%) (Fig. 5H, and Supplementary Fig. 6). These results suggest the feasibility of predicting the proportion of tumor cell-derived evDNA in the plasma or cfDNA methylation analysis.

The potential of EVs for diagnostic applications

To evaluate the diagnostic potential of tumor-derived EVs, we quantified the levels of carcinoembryonic antigen (CEA), a tumor marker, on EVs from tumor tissue of CRC patients using an enzyme-linked immunosorbent assay (ELISA). EV samples from tumor tissue were mixed with EV samples from control blood, yielding tumor-derived EV concentrations ranging from 0 to 30 ng/μL of protein. ELISA detection of CEA levels in the mixture samples showed a limit of detection (LOD) ranging from 4.22 to 6.11 ng/μL of protein, indicating that EVs from tumor tissue could potentially serve as a diagnostic tool for cancer detection (Supplementary Fig. 7).

Discussion

Our investigation of genome-wide DNA methylation patterns in evDNA revealed significant similarities between its DNA methylation profiles and those of nDNA in cancer cells. Initially, we characterized EVs originating from two cancer cell lines, HCT116 (colon cancer) and MDA-MB-231 (breast cancer), using NTA, western blotting, and ExoView imaging. Consistent with previous findings[23], our results indicate that tumor-derived EVs encapsulate dsDNA, predominately ranging between 250 bp and 15.0 kbp. This size distribution suggests the presence of dsDNA within these EVs. Considering the stability of the global methylation patterns in dsDNA[30–32], our findings support the use of evDNA in methylation-based clinical diagnostics.

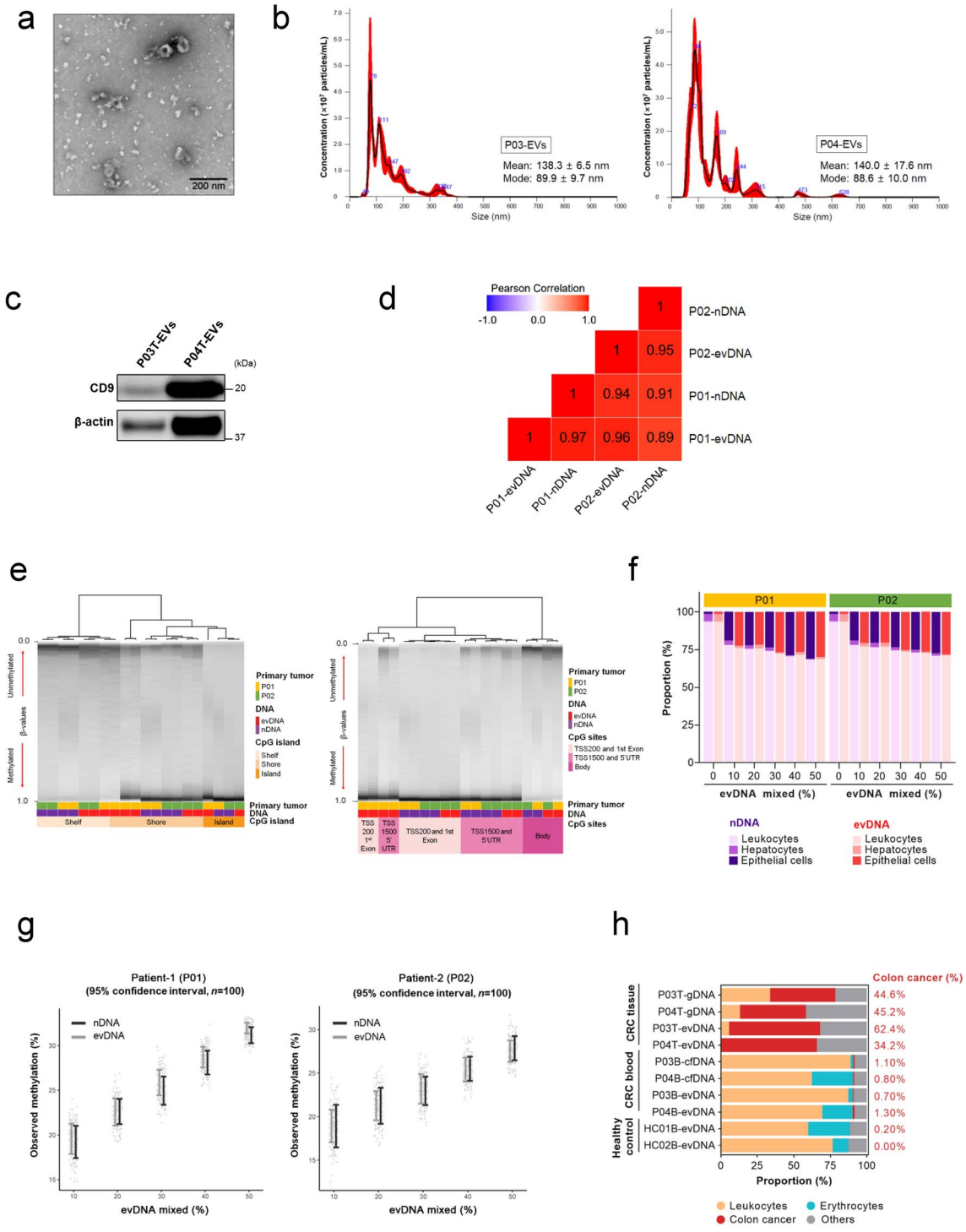
We examined over 850,000 CpG sites across total DNA, nDNA, cytoDNA, and evDNA. Our findings suggest that cytoDNA may exhibit higher hypermethylation levels than nDNA, potentially leading to an incomplete representation of epigenetic changes. In contrast, the methylation pattern of evDNA closely mirrors that of nDNA in both cell lines, indicating that EVs may carry the entire genome and reflect the epigenetic modifications present in their cells of origin. CpG islands, particularly in promoters, are crucial regions for DNA methylation and gene transcription regulation[3,7], and cancer-specific hypermethylation at CpG island (CGI) promoters is a well-documented epigenetic alteration[33]. Our data revealed hypermethylation of CGI promoters in both evDNA and nDNA, independent of the cell line, with differentially methylated sites predominantly located in the gene body. This observation aligns with the findings of Huang et al.[34], who reported that early-stage intestinal epithelial cells exhibit differential hypermethylation, transitioning to global hypomethylation in the gene body as they progress to a cancerous state. Our discovery that evDNA is as hypomethylated as nDNA in the gene body suggests that tumor-related methylated DNA is encapsulated within EVs.

As crucial elements of epigenetic regulation, enhancers may exert extensive and significant control over cellular identity and the advancement of cancer[35]. Previous research using whole-genome bisulfite sequencing has demonstrated aberrant DNA methylation in the super-enhancer regions of human cancers[36]. The results from our differential methylation analysis showed no differences in the methylation levels of enhancers between evDNA and nDNAs. There is a growing need to investigate enhancer methylation and its regulatory functions in human cancer, aiming for a more detailed understanding of the enhancer landscape. Our limitation can be overcome by employing bisulfite sequencing or enhancing the coverage of methylation arrays.

A recent study on glioblastoma EVs underscored the utility of DNA methylation patterns in EVs in identifying the methylation class of parental cells and the original tumors[37]. Our cluster analysis demonstrated that samples with similar lineages clustered closely, indicating that DNA methylation patterns in EVs can effectively distinguish different cell lines. These results align with previous findings[37], supporting that evDNA methylation analysis can serve as a valuable tool for classifying and identifying cancers based on their methylation patterns.

Approximately 80% of evDNA is located on the surface of extracellular vesicles[37,38]. Additionally, whole genome sequencing has shown that over 99% of total evDNA (both surface-bound and internal) is identical to genomic DNA (gDNA) and is not enriched in specific genomic regions[29], indicating that evDNA can be used to trace tumor-derived DNA. Consistent with these findings, our analysis of DNA methylation patterns demonstrated that internalized evDNA is well-preserved compared to total evDNA and gDNA from the cell of origin. This supports the potential of EV-derived DNA as valuable biomarkers for cancer detection.

To validate the findings from these cell lines, we extended our study to include tumor tissues from primary cancer surgery samples. The levels of DNA methylation in tumor tissue-derived evDNA and gDNA from CRC



patients were similar, suggesting that the evDNA from tumor tissues displayed properties comparable to those of gDNA from primary tissues. Deconvolution tests, simulating DNA mixtures containing cell-free or plasma evDNA, facilitated the estimation of the relative proportion of tumor cells based on evDNA methylation patterns. Previous research has indicated a lower proportion of tumor-derived evDNA in the plasma than in cfDNA in patients with lung cancer[39]. Our simulation tests, intriguingly, demonstrated the measurability of cancer abundance as low as < 5% based on evDNA methylation. When combined with high-throughput sequencing and filtering, evDNA analysis holds the potential to discern low-frequency tumor-derived DNA from plasma-based sources.

The key advantage of evDNA lies in its cell-of-origin specificity and protection from DNase and other nucleases. Although the quantity of evDNA is lower than cfDNA, evDNA shows a higher mutational fractional abundance compared to cfDNA. Additionally, evDNA has a positive correlation with the tumor marker CEA in colon cancer[40], indicating it is more tumor-specific and offers potential advantages over current cfDNA approaches. The increased secretion of extracellular vesicles by tumor cells further highlights evDNA as a

◀**Fig. 5**. Methylation pattern concordance in clinical cancer samples. Characterization of EVs derived from patients with CRC. (**a**) Transmission Electron Microscopy (TEM) analysis of tumor tissue-derived EVs identified round-shaped structures ranging from 50 to 200 nm in size as EVs. (**b**) Particle size distribution of tumor tissue-derived EV samples from CRC patients. The mean size, mode size, and standard deviation (SD) values of EVs were determined by NTA. (**c**) Western blot analysis of EVs from CRC patients (P03T-EVs and P04T-EVs). The membranes were probed for CD9 (EVs marker) and β-actin (loading control). (**d**) Heatmap of pairwise correlation of methylation profiles in colon cancer patient samples. The heatmap color codes the pairwise Pearson correlation of methylation levels in total DNA, nDNA, cytoDNA, and evDNA from tumor samples of colon cancer patients (P01 and P02). (**e**) Unsupervised hierarchical clustering based on the β-values at CpG Islands (left side) and CpG sites (right side). Heatmap displaying DNA methylation differences in nDNA (purple) and evDNA (red) from tumor samples of colon cancer patients (P01, yellow; P02, green). Annotations for the shelf (light orange), island (yellow-orange), shore (orange), TSS200 and 1st Exon (light pink), TSS1500 and 5'UTR (pink), and body (dark pink) are shown at the bottom of the heatmap. (**f**) Bar plot of cell proportion across evDNA fractions in simulated DNA mixtures ranging from 0 to 50% at 10% intervals. (**g**) Scatterplot with error bars of evDNA mixed and observed methylation patterns in evDNA and nDNA from tumor samples of colon cancer patients (P01 and P02). (**h**) Bar plot of cell proportion across gDNA, tumor tissue-derived evDNA, and blood-derived evDNA. Cell type proportions for each sample are colored by deconvoluted cell types: leukocytes (yellow), erythrocytes (mint), colon cancer (burgundy), and others (grey).

valuable biomarker for liquid biopsy. Improving methods to isolate tumor-specific EVs will boost the value of evDNA, allowing more precise characterization of tumor-specific information.

Our study has several limitations. First, while our findings highlight the potential of evDNA for use in liquid biopsy, additional research is required to identify tumor-specific methylation markers, refine assay development, and assess sensitivity and specificity in larger patient cohorts. Second, since our study focuses on colon and breast cancer models, the observed similarities in methylation patterns between evDNA and nDNA may not necessarily extend to other cancer types, limiting the generalizability of our results. Third, the small sample size emphasizes the need for further validation in larger, more diverse cohorts across multiple cancer types to fully evaluate the clinical utility of evDNA.

In conclusion, our study establishes that cancer cell-derived evDNA manifests a methylation pattern akin to that observed in the nDNA of the corresponding cancer cells. These findings bolster the potential use of DNA methylation patterns in EVs as cancer biomarkers. Our research provides a foundation for future studies dedicated to developing DNA methylation-based biomarkers for early cancer diagnosis and prognosis monitoring.

Methods

Cell lines and culture

HCT116 and MDA-MB-231 cell lines were obtained from the American Type Culture Collection (ATCC, USA). The cells were cultivated in Dulbecco's modified Eagle's medium (DMEM; Hyclone, USA) supplemented with 10% (v/v) heat-inactivated fetal bovine serum (FBS; Hyclone) and 1 × penicillin/streptomycin (P/S; Hyclone). Cultures were maintained at 37 °C in a humidified atmosphere containing 5% CO_2 and passaged biweekly. For EV collection, cells were cultured for 72 h in DMEM supplemented with 10% EV-depleted FBS and 1 × P/S. Subsequently, the culture media underwent processing for EV isolation.

Human specimens and processing

The study protocol was reviewed and approved by the Institutional Review Board (IRB) of the Severance Hospital Ethics Committee (IRB number 4-2019-0811). A written informed consent was obtained from all patients. All information regarding the human sample was managed using anonymous numerical codes, and the experiments were conducted by the Declaration of Helsinki.

Fresh human tumor tissues were obtained from patients who underwent colorectal resection at Severance Hospital, Republic of Korea. The surgical tissues were sectioned into small pieces and cultured for 16 h in serum-free RPMI, supplemented with 1 × P/S. The resultant culture media were processed for EV isolation.

Human blood samples were collected from patients or healthy controls in EDTA tubes. Whole blood was centrifugated at 1900 × g for 15 min at 4 °C, and the supernatant was transferred to clean tubes. Then, tubes were centrifugated at 1900 × g for 15 min at 4 °C and stored at -80 °C until cfDNA extraction.

Isolation of EVs via ultracentrifugation

For EV isolation, culture media and plasma samples were collected in conical tubes and centrifuged at 500×*g* for 10 min, followed by centrifugation at 3000×*g* for 20 min. The supernatant was then transferred to a clear polycarbonate tube and centrifuged at 12,000×*g* for 20 min using a Beckman Type 70 Ti rotor (Beckman Coulter, UK). The supernatant was transferred to a new tube and centrifuged at 100,000×*g* for 70 min to collect the EV pellet. The pellet was resuspended in 2 mL ice-cold PBS and centrifuged at 100,000×*g* for 70 min. After washing, the purified EV pellets were eluted with 200 µL ice-cold PBS. All steps were performed at 4 °C.

Characterization of EVs

The size distribution of EVs was determined through NTA using NanosSight NS300 (Marvern Panalytical, UK). Purified EVs were diluted in deionized water, and NTA was performed according to the manufacturer's instructions. NTA measurement conditions included imaging each sample thrice for 60 s, with the detection

threshold set at 3. Recorded data were analyzed using the NTA software (version 3.4) to calculate size distribution and concentration.

To assess the surface protein tetraspanins (CD81, CD63, and CD9) of EVs, we employed the ExoView Tetraspanin kit and ExoView R100 system (NanoView Biosciences, USA). Tetraspanins in EVs were assayed according to the manufacturer's instructions. EVs were diluted to 10^7–10^8 particles/mL in Solution A of the ExoView Tetraspanin kit. Tetraspanin chips were placed in a 24-well plate, and EVs were meticulously deposited onto the chips. The plates were incubated for 16 h at room temperature (RT) in the dark. Post-incubation, chips were washed and exposed to fluorescent-conjugated antibodies against CF488-anti-CD9 (blue), CF555-anti-CD63 (red), and CF647-anti-CD81 (green) for 1 h at RT. Subsequently, the chips were subjected to fluorescence imaging using the ExoView R100 system.

To verify the size and shape of patient-derived EVs, samples at a protein concentration of 0.1 mg/mL were prepared for transmission electron microscopy (TEM) with negative staining, according to methods described in a previous study[41].

The EVs suspended in PBS were mounted on a grid with a carbon-coated formvar film for 1 min. Excess liquid was gently removed using filter paper, and the sample was then negatively stained with four drops of 1.5% uranyl acetate (w/v). Following washing and air-drying, the grids were analyzed with a JEOL JSM 1400 transmission electron microscope (JEOL, USA) at 100 kV. Images were acquired using a Veleta 2 K × 2 K CCD camera (Olympus, Germany).

Preparation of nucleic and cytoplasmic fractions from culture cells and tissues

Nuclear and cytoplasmic fractions were obtained through subcellular fractionation using the Pierce Mitochondrial Isolation Kit (Thermo Fisher Scientific, USA). Fractionation was performed according to the manufacturer's instructions. Briefly, cells or tissues were washed thrice with PBS and centrifuged for 2 min at 850×*g*. The cells or tissues were resuspended in Reagent A and incubated on ice for 2 min. Subsequently, they were resuspended in Reagent C and homogenized. The mixture underwent a 10 min centrifugation at 700×*g*, forming a pellet (nuclear fraction) that underwent three washes with PBS. The supernatant was re-centrifuged for 15 min at 12,000×*g*, yielding a final supernatant (cytoplasmic fraction). All procedures were conducted on ice, using pre-cooled buffers.

Western blotting

Total cells, nuclear and cytoplasmic fractions, and purified EVs were lysed on ice in RIPA buffer (Thermo Fisher Scientific) containing 1 × protease inhibitors and 1 mM phenylmethylsulphonyl fluoride (PMSF). The supernatant was transferred to a new tube following centrifugation at 14,000×*g* and 4 °C. All cell lysates underwent preparation for western blot analysis. Protein concentration was determined using a BCA Protein Assay kit (Thermo Fisher Scientific); the absolute concentration was calculated by adding bovine serum albumin to each measurement set.

Protein samples (30 µg) were electrophoresed on a 12% SDS–polyacrylamide gel and transferred onto a polyvinylidene fluoride (PVDF) membrane using an iBlot system (Invitrogen, USA). Subsequently, the membranes were blocked with 5% skim milk at RT and then incubated with the primary antibody against Lamin B1 (nuclear marker; Cat. 2118S, Cell Signaling Technology, USA) and anti-GAPDH (loading control; Cat. 12586, Cell Signaling Technology) and anti-CD9 antibodies (EVs marker; NBP2-22187, Novus Biologicals, USA). Then, the membranes were incubated with horseradish peroxidase-conjugated anti-rabbit or anti-mouse secondary antibodies. Immunoblots were developed using chemiluminescence (ECL, Thermo Fisher Scientific) kit and detected using the ImageQuant LAS 4000 luminescent image analyzer (Fuji Film, Japan).

DNA preparation

DNA was extracted from total cells, tissues, nuclear fractions, and cytoplasmic fractions using the QIAamp DNA Mini Kit (No. 51304; Qiagen, USA) according to the manufacturer's instructions.

The evDNA was extracted using magnetic AMPure beads (Beckman Coulter, USA) per the manufacturer's protocol. Briefly, EVs were lysed in lysis buffer (AL) containing proteinase K (Qiagen). The lysates were mixed with AMPure beads, PEG-NaCl solution, and isopropanol and incubated at RT. Subsequently, the samples were placed on a magnetic rack for clarification. After washing them twice with fresh 80% ethanol and carefully removing the supernatant, the beads were air-dried and eluted with nuclease-free water. cfDNA was extracted with QIAamp Circulating Nucleic Acid Kit (Qiagen) according to the manufacturer's instructions.

To obtain internalized evDNA, EVs were treated with DNase before evDNA extraction. A 100 µL aliquot of the EV was briefly treated with 1 × Baseline-ZERO DNase (EN0771; Thermo Fisher Scientific) for 15 min at 37 °C. Following incubation, the DNase was inactivated by adding a stop solution and incubating for 10 min at 65 °C. We conducted evDNA extraction from DNase-treated EVs.

The integrity, quantity, and purity of all DNA samples were assessed on a 4200 TapeStation System using a DNA High-sensitivity chip (Agilent Technologies, USA).

Bisulfite conversion

Bisulfite conversion was conducted with the EZ DNA Methylation-Gold Kit (Zymo Research, USA) following the manufacturer's instructions. Briefly, the DNA was subjected to denaturation with a CT conversion reagent. The resultant converted samples were loaded onto the columns and centrifuged. The bisulfite-converted DNA was eluted and stored at − 20 °C until further use.

Genome-wide DNA methylation profiling

Genome-scale DNA methylation profiles were generated using the Infinium MethylationEPIC BeadChips (Illumina Inc., USA). This assay enables the evaluation of DNA methylation status at over 853,307 CpG sites, representing ~3% of the total CpG dinucleotides in the human genome[42]. The methylation assay was performed according to the manufacturer's guidelines. Processing of genome-wide methylation raw data (IDAT files) was carried out using the GenomeStudio software (v2011.1) and R packages (methylumi and lumi). Background-corrected signal intensities were determined using a set of negative controls. Each methylation data point is represented by fluorescent signals corresponding to the methylated (M) and unmethylated (U) alleles. The signal intensities were used to calculate beta (β)-values, derived from the mean M and U signals for each CpG site using the formula ($\beta = (M)/(U + M + 100)$). Subsequently, β-values were normalized using beta-mixture quantile normalization (BMIQ)[43].

Deconvolution of DNA methylation profiles

Hierarchical clustering utilized Euclidean distances between BMIQ-normalized beta values and complete linkage. Concordance levels were assessed through Pearson correlation coefficients. We used t-distributed stochastic neighbor embedding (t-SNE) analysis[44,45] to investigate the association between the methylation profiles of evDNA and nDNA in nine cell lines encompassing two breast cancer (MDA-MB-231 and MCF7), two CRC (HCT116 and RKO), two prostate cancer (Du145 and PC3), and three normal epithelial cell lines (RWPE1, HMEC, and FHC). DNA methylation data for these cell lines were obtained from the GSE database (GSE54758)[46]. Integrating DNA methylation data for nDNA and evDNA from the HCT116 and MDA-MB-231 cell lines with publicly available data preceded subsequent t-SNE analyses.

To estimate the mean DNA methylation levels of specific CpGs from different cell types, we utilized a non-negative least squares (NNLS) deconvolution algorithm. Reference CpGs were chosen through an adaptation of the algorithm developed by Luo et al.[47]. First, CpGs for each of the cell types were selected based on Benjamini–Hochberg adjusted significant *p*-values from Welch two-sample t-tests (with a 5% significance level), comparing methylation values between the target group and all other groups. The highest p-value among all pairwise comparisons was chosen to ensure high cell specificity. From these selected CpGs, the 100 with the greatest mean methylation differences between the target group and the other groups were selected further.

Mixture profiles were simulated using methylation profiles from a dataset (GSE122126) containing a mixture of immune cells ($CD45^+$) and hepatocytes in defined ratios of 96% and 4%, respectively[10]. Integration of DNA methylation profiles of evDNA and nDNA from primary colorectal cancers with reference data occurred with defined ratios ranging from 0 to 50%. We generated 100 simulation mixtures per ratio by introducing noise values, following a previously described method[48]. For NNLS regression, we used the methylation atlas[10] as signatures; the median of the regression coefficients in the simulated mixture determined the abundance of corresponding signatures, presented with 95% confidence intervals. All methylation data analyses were performed using the R software.

Epigenetic configurations

The ENCODE chromatin state segmentation by ChromHMM for nine cell lines was obtained from the UCSC Genome Browser (https://genome.ucsc.edu/cgi-bin/hgFileUi?db=hg19&g=wgEncodeBroadHmm). To investigate the association between DNA methylation level and chromatin states, we calculated the methylation level differences between nDNA and evDNA for all CpG sites in both HCT116 and MDA-MB-231 cell lines. A comparative analysis examined the differences in methylation levels and annotated chromatin states across all chromosomal regions of the lymphoblastoid cell line GM12878. Furthermore, to enhance the robustness of our analysis, we also incorporated data from eight other cell lines: H1-hESC, HepG2, HMEC, HSMM, HUVEC, K562, NHEK, and NHLF.

ELISA

Extracellular vesicle (EVs) samples from tumor tissue (P01T-EVs, P02T-EVs, and P05T-EVs) were mixed with EV samples from benign blood. These mixtures yielded tumor-derived EV concentrations ranging from 0 to 30 μg/mL. Microtitre plates (Nunc Cell Culture, Thermo Fisher Scientific) were coated with varying concentrations from 0 to 30 μg/mL in 50 μL of 1 × coating buffer and incubated overnight at 4 °C. The plates were then blocked with 100 μL of 1% (w/v) bovine serum albumin (BSA; Sigma-Aldrich) in PBS for 2 h at RT. Following three washes with PBS-T, CEA diluted 1:500 was added into the wells and incubated overnight at 4 °C. After that, the plates were washed three times with PBS-T and incubated with HRP-conjugated anti-mouse IgG secondary antibody at a 1:2000 dilution (100 μL/well) for 2 h at RT. After five washes with PBS-T, 100 μL of 3,3′,5,5′-tetramethylbenzidine (TMB, Sigma-Aldrich) substrate was added to each well. After a 30 min incubation at RT, 100 μL of stop solution (2 M H_2SO_4, Thermo Fisher Scientific) was added, and the optical absorbance at 450 nm was measured using a plate reader.

Data availability

The datasets used and analyzed during the current study are available from the corresponding author upon reasonable request.

Received: 11 December 2023; Accepted: 3 October 2024
Published online: 15 October 2024

References

1. Sharma, S., Kelly, T. K. & Jones, P. A. Epigenetics in cancer. *Carcinogenesis* **31**, 27–36. https://doi.org/10.1093/carcin/bgp220 (2010).
2. Michalak, E. M., Burr, M. L., Bannister, A. J. & Dawson, M. A. The roles of DNA, RNA and histone methylation in ageing and cancer. *Nat. Rev. Mol. Cell Biol.* **20**, 573–589. https://doi.org/10.1038/s41580-019-0143-1 (2019).
3. Smith, Z. D. & Meissner, A. DNA methylation: Roles in mammalian development. *Nat. Rev. Genet.* **14**, 204–220. https://doi.org/10.1038/nrg3354 (2013).
4. Bird, A. DNA methylation patterns and epigenetic memory. *Genes Dev.* **16**, 6–21. https://doi.org/10.1101/gad.947102 (2002).
5. Koestler, D. C. et al. Distinct patterns of DNA methylation in conventional adenomas involving the right and left colon. *Mod. Pathol.* **27**, 145–155. https://doi.org/10.1038/modpathol.2013.104 (2014).
6. Luo, Y. et al. Differences in DNA methylation signatures reveal multiple pathways of progression from adenoma to colorectal cancer. *Gastroenterology* **147**, 418–429. https://doi.org/10.1053/j.gastro.2014.04.039 (2014).
7. Schubeler, D. Function and information content of DNA methylation. *Nature* **517**, 321–326. https://doi.org/10.1038/nature14192 (2015).
8. Jones, P. A. Functions of DNA methylation: Islands, start sites, gene bodies and beyond. *Nat. Rev. Genet.* **13**, 484–492. https://doi.org/10.1038/nrg3230 (2012).
9. Esteller, M. et al. Inactivation of the DNA-repair gene MGMT and the clinical response of gliomas to alkylating agents. *N. Engl. J. Med.* **343**, 1350–1354. https://doi.org/10.1056/NEJM200011093431901 (2000).
10. Moss, J. et al. Comprehensive human cell-type methylation atlas reveals origins of circulating cell-free DNA in health and disease. *Nat. Commun.* **9**, 5068. https://doi.org/10.1038/s41467-018-07466-6 (2018).
11. Chung, W. et al. Detection of bladder cancer using novel DNA methylation biomarkers in urine sediments. *Cancer Epidemiol. Biomark.* **20**, 1483–1491. https://doi.org/10.1158/1055-9965.Epi-11-0067 (2011).
12. Heyn, H. & Esteller, M. DNA methylation profiling in the clinic: Applications and challenges. *Nat. Rev. Genet.* **13**, 679–692. https://doi.org/10.1038/nrg3270 (2012).
13. Raposo, G. & Stoorvogel, W. Extracellular vesicles: Exosomes, microvesicles, and friends. *J. Cell Biol.* **200**, 373–383. https://doi.org/10.1083/jcb.201211138 (2013).
14. Maas, S. L. N., Breakefield, X. O. & Weaver, A. M. Extracellular vesicles: Unique intercellular delivery vehicles. *Trends Cell Biol.* **27**, 172–188. https://doi.org/10.1016/j.tcb.2016.11.003 (2017).
15. Skog, J. et al. Glioblastoma microvesicles transport RNA and proteins that promote tumour growth and provide diagnostic biomarkers. *Nat. Cell Biol.* **10**, 1470-U1209. https://doi.org/10.1038/ncb1800 (2008).
16. Hoshino, A. et al. Extracellular vesicle and particle biomarkers define multiple human cancers. *Cell* **182**, 1044. https://doi.org/10.1016/j.cell.2020.07.009 (2020).
17. Costa-Silva, B. et al. Pancreatic cancer exosomes initiate pre-metastatic niche formation in the liver. *Nat. Cell Biol.* **17**, 816. https://doi.org/10.1038/ncb3169 (2015).
18. Hoshino, A. et al. Tumour exosome integrins determine organotropic metastasis. *Nature* **527**, 329. https://doi.org/10.1038/nature15756 (2015).
19. Wortzel, I., Dror, S., Kenific, C. M. & Lyden, D. Exosome-mediated metastasis: Communication from a distance. *Dev. Cell* **49**, 347–360. https://doi.org/10.1016/j.devcel.2019.04.011 (2019).
20. Bobrie, A. & Thery, C. Exosomes and communication between tumours and the immune system: Are all exosomes equal?. *Biochem. Soc. Trans.* **41**, 263–267. https://doi.org/10.1042/Bst20120245 (2013).
21. Pelissier Vatter, F. A. et al. Extracellular vesicle- and particle-mediated communication shapes innate and adaptive immune responses. *J. Exp. Med.* https://doi.org/10.1084/jem.20202579 (2021).
22. Chevillet, J. R. et al. Quantitative and stoichiometric analysis of the microRNA content of exosomes. *Proc. Natl. Acad. Sci. USA* **111**, 14888–14893. https://doi.org/10.1073/pnas.1408301111 (2014).
23. Williams, C. et al. Double-stranded DNA in exosomes: A novel biomarker in cancer detection. *Cell Res.* **24**, 766–769. https://doi.org/10.1038/cr.2014.44 (2014).
24. Kahlert, C. et al. Identification of double-stranded genomic DNA spanning all chromosomes with mutated KRAS and p53 DNA in the serum exosomes of patients with pancreatic cancer. *J. Biol. Chem.* **289**, 3869–3875. https://doi.org/10.1074/jbc.C113.532267 (2014).
25. Allenson, K. et al. High prevalence of mutant KRAS in circulating exosome-derived DNA from early-stage pancreatic cancer patients. *Ann. Oncol.* **28**, 741–747. https://doi.org/10.1093/annonc/mdx004 (2017).
26. Wan, Y. et al. Nanoscale extracellular vesicle-derived DNA is superior to circulating cell-free DNA for mutation detection in early-stage non-small-cell lung cancer. *Ann. Oncol.* **29**, 2379–2383. https://doi.org/10.1093/annonc/mdy458 (2018).
27. Bernard, V. et al. Circulating nucleic acids are associated with outcomes of patients with pancreatic cancer. *Gastroenterology* **156**, 108–118. https://doi.org/10.1053/j.gastro.2018.09.022 (2019).
28. Yamamoto, H. et al. BARHL2 methylation using gastric wash DNA or gastric juice exosomal DNA is a useful marker for early detection of gastric cancer in an h-pylori-independent manner. *Clin. Transl. Gastroenterol.* https://doi.org/10.1038/ctg.2016.40 (2016).
29. Thakur, B. K. et al. Double-stranded DNA in exosomes: A novel biomarker in cancer detection. *Cell Res* **24**, 766–769. https://doi.org/10.1038/cr.2014.44 (2014).
30. Egger, G., Wielscher, M., Pulverer, W., Kriegner, A. & Weinhausel, A. DNA methylation testing and marker validation using PCR: Diagnostic applications. *Expert Rev. Mol. Diagn.* **12**, 75–92. https://doi.org/10.1586/Erm.11.90 (2012).
31. Noehammer, C. et al. Strategies for validation and testing of DNA methylation biomarkers. *Epigenomics-Uk* **6**, 603–622. https://doi.org/10.2217/Epi.14.43 (2014).
32. Laird, P. W. The power and the promise of DNA methylation markers. *Nat. Rev. Cancer* **3**, 253–266. https://doi.org/10.1038/nrc1045 (2003).
33. Zheng, Y. Y. et al. A pan-cancer analysis of CpG Island gene regulation reveals extensive plasticity within Polycomb target genes. *Nat. Commun.* **12**, 2485. https://doi.org/10.1038/s41467-021-22720-0 (2021).
34. Huang, K. K. et al. Genomic and epigenomic profiling of high-risk intestinal metaplasia reveals molecular determinants of progression to gastric cancer. *Cancer Cell* **33**, 137–150. https://doi.org/10.1016/j.ccell.2017.11.018 (2018).
35. Sur, I. & Taipale, J. The role of enhancers in cancer. *Nat. Rev. Cancer* **16**, 483–493. https://doi.org/10.1038/nrc.2016.62 (2016).
36. Heyn, H. et al. Epigenomic analysis detects aberrant super-enhancer DNA methylation in human cancer. *Genome Biol.* **17**, 11. https://doi.org/10.1186/s13059-016-0879-2 (2016).
37. Maire, C. L. et al. Genome-wide methylation profiling of glioblastoma cell-derived extracellular vesicle DNA allows tumor classification. *Neuro. Oncol.* **23**, 1087–1099. https://doi.org/10.1093/neuonc/noab012 (2021).
38. Lazaro-Ibanez, E. et al. DNA analysis of low- and high-density fractions defines heterogeneous subpopulations of small extracellular vesicles based on their DNA cargo and topology. *J. Extracell. Vesicles* **8**, 1656993. https://doi.org/10.1080/20013078.2019.1656993 (2019).
39. Abbosh, C. et al. Phylogenetic ctDNA analysis depicts early-stage lung cancer evolution. *Nature* **545**, 446–451. https://doi.org/10.1038/nature22364 (2017).
40. Choi, J. et al. Detection of circulating KRAS mutant DNA in extracellular vesicles using droplet digital PCR in patients with colon cancer. *Front. Oncol.* **12**, 1067210. https://doi.org/10.3389/fonc.2022.1067210 (2022).

41. Bojmar, L. et al. Protocol for cross-platform characterization of human and murine extracellular vesicles and particles. *Star. Protoc.* **5**, 102754. https://doi.org/10.1016/j.xpro.2023.102754 (2024).
42. Moran, S., Arribas, C. & Esteller, M. Validation of a DNA methylation microarray for 850,000 CpG sites of the human genome enriched in enhancer sequences. *Epigenomics* **8**, 389–399. https://doi.org/10.2217/epi.15.114 (2016).
43. Teschendorff, A. E. et al. A beta-mixture quantile normalization method for correcting probe design bias in Illumina Infinium 450 k DNA methylation data. *Bioinformatics* **29**, 189–196. https://doi.org/10.1093/bioinformatics/bts680 (2012).
44. Hinton, G. E. & Roweis, S. Stochastic neighbor embedding. *Adv. Neural Inf. Process. Syst.* **15**, 857–864 (2002).
45. Van der Maaten, L. & Hinton, G. Visualizing data using t-SNE. *J. Mach. Learn. Res.* **9**, 2759–2605 (2008).
46. Takeshima, H., Wakabayashi, M., Hattori, N., Yamashita, S. & Ushijima, T. Identification of coexistence of DNA methylation and H3K27me3 specifically in cancer cells as a promising target for epigenetic therapy. *Carcinogenesis* **36**, 192–201. https://doi.org/10.1093/carcin/bgu238 (2014).
47. Luo, Q. et al. A meta-analysis of immune-cell fractions at high resolution reveals novel associations with common phenotypes and health outcomes. *Genome Med.* **15**, 59. https://doi.org/10.1186/s13073-023-01211-5 (2023).
48. Newman, A. M. et al. Robust enumeration of cell subsets from tissue expression profiles. *Nat. Methods* **12**, 453–457. https://doi.org/10.1038/nmeth.3337 (2015).

Acknowledgements

The schematic in Fig. 1A was created using BioRender.com

Author contributions

HSK, IW, and TK contributed to the conceptualization and design of this study, data curation, and the development of the methodology. KK and SK wrote the original drafts of the manuscript. KK and SL performed experiments and created figures. YDH provided surgical tissue for experiments. TK, IW, and HSK reviewed and edited the manuscript. TK and HSK supervised this study.

Funding

This study was supported in part by a grant from the National Research Foundation of Korea (NRF) funded by the Korean government (MSIT) (NRF-2019R1C1C1006709, NRF-2022R1A2C4001879 to H.S.K. and NRF-2019M3E5D3073104 to T.K.), the Bio & Medical Technology Development Program of the National Research Foundation (NRF) & funded by the Korean government (MSIT) (NRF-2022M3A9F3016364 to H.S.K.) and a grant from the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health & Welfare, Republic of Korea (No. HI22C0353 and RS-2023-00304686 to H.S.K.). Additionally, this work was partly supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF), funded by the Ministry of Education (NRF-2022R1I1A1A01072740 to K.K.).

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-024-75287-3.

Correspondence and requests for materials should be addressed to T.-M.K. or H.S.K.