scientific reports

Check for updates

OPEN

Profiling extracellular matrix-driven heterogeneity of single cell migration and morphology

Euichul Shin[1,4], Jihun Han[2,4], Ara Jung[1], Jahangir Khan[1], Ian Y. Wong[3] & Bomi Gweon[1]✉

Cell migration and morphology are fundamental to various biological processes, including tissue development, wound healing, and cancer progression. In this study, we investigated the migration and morphological characteristics of HeLa cells cultured on different extracellular matrix (ECM) components—laminin (Ln), collagen 1 (Col1), and collagen 3 (Col3)—using machine learning-based image analysis. Cells on Ln exhibited significantly higher motility, increased turning angles, and reduced directional persistence compared to those on Col1 and Col3. Morphologically, these cells showed larger areas and higher solidity, suggesting a distinct attachment pattern compared to those on collagens. In addition, principal component analysis (PCA) effectively classified cells based on motility parameters, with cells on Ln forming a clearly separate cluster in PCA space. Further biological assays demonstrated that cells on Ln exhibited increased cell-cell interactions and smaller but more numerous focal adhesions, indicating that Ln promotes enhanced migratory plasticity and intercellular interactions. These biological factors were identified as key contributors to the distinctions observed in PCA classification. This study highlights the differential effects of ECM components on cellular behavior and demonstrates the value of integrating machine learning-based quantitative analysis with biological interpretation to advance our understanding of cell-microenvironment interactions.

Keywords Cell migration, Extracellular matrix (ECM), Principal component analysis (PCA)

Cellular properties, such as morphology and migration, are vital for understanding cellular physiology[1–7]. These properties reflect fundamental cellular processes, including adhesion, proliferation, and differentiation, as well as are involved in various complex events such as embryonic development, wound healing, and cancer progression[1–11]. In cancer pathophysiology, for example, changes in cell morphology and migration are closely associated to epithelial-to-mesenchymal transition (EMT), a critical process that facilitates invasion and metastasis[12–18]. Similarly, cell migration is indispensable for tissue formation during development and wound repair. Given their significance, various techniques have been developed to investigate cellular behavior through the analysis of morphology and migration.

Many classical image analysis methods are based on image thresholding to detect and segment cells from the background[19–22]. One of the pioneering approaches was introduced by Prewitt et al., who developed early computational methods for segmenting cell nuclei from images[19]. Since then, techniques such as Otsu's method, iterative thresholding, and adaptive thresholding have been widely used to determine optimal intensity thresholds for segmentation[21–25]. These methods numerically detect cell or nuclear edges based on intensity profiles and have been applied to images acquired through various modalities, including fluorescence, bright-field, and phase-contrast microscopy[24,26–29]. They enable quantitative analysis of cellular morphology and facilitate automated assessments of cell migration. For example, these approaches have been used to quantify cell shape and alignment within cellular monolayers[27,28], track melanocyte migration under the influence of cell division[30], and automate cell tracking for drug testing[31,32]. While these thresholding-based techniques have significantly advanced cell segmentation and tracking, they have inherent limitations in capturing the complexity and dynamic nature of cellular behaviors. They often struggle with variable cell shapes, overlapping structures, and inconsistencies in image intensity profiles due to variations in illumination, contrast, and various optical properties.

Recent advancements in artificial intelligence (AI) have greatly enhanced cell image analysis methodologies. Machine learning and deep learning approaches have addressed challenges related to segmentation variability

[1]Department of Mechanical Engineering, Sejong University, Seoul 05006, Korea. [2]Department of Mathematics and Statistics, University at Albany - State University of New York, Albany, NY 12222, USA. [3]School of Engineering, Legoretta Cancer Center, Brown University, Providence, RI 02912, USA. [4]Euichul Shin and Jihun Han contributed equally to this work. ✉email: bgweon@sejong.ac.kr

and improved the accuracy of automated quantification. These AI-based techniques enable accurate, high-throughput analyses of cellular morphology and classification. For instance, convolutional neural networks (CNNs) were employed to automatically identify cell types and classify thousands of cells in a label-free manner[33], as well as to track cell cycle progression in phase-contrast images[34]. In addition, deep learning-based semantic segmentation models incorporating CNN have been developed to identify differentiated cells by predicting smooth muscle actin formation in fibroblasts[35]. AI has also demonstrated great potential in tracking migration and analyzing dynamic cellular behaviors. Kesapragada et al. employed a machine learning method to classify macrophage subtypes by integrating motility patterns with morphological features, demonstrating the feasibility of AI-based methods in distinguishing functional cell states[36]. These approaches have also been applied to predict cancer cell migration, as demonstrated by Garcia-Moreno et al., who employed deep learning to model the dynamics of breast cancer cell migration[37].

As such, with the AI-based image analysis, high-throughput data processing has become feasible, which has facilitated large-scale data collection and improved the ability to analyze heterogeneous populations. These advances have supported the capture of subtle variations in cell behavior, particularly when combined with robust statistical analyses. Notably, integrating AI-assisted cell detection and tracking with statistical approaches such as principal component analysis (PCA), and uniform manifold approximation and projection (UMAP) has demonstrated a synergistic effect in uncovering underlying features and distinguishing subtle differences among heterogeneous cell populations[38–41]. These statistical models have been particularly effective in capturing high-dimensional morphological and motility data, aiding in more precise classification of cell states. For instance, PCA has been applied to classify cells undergoing EMT based on morphodynamic phenotypes, demonstrating its potential in identifying morphological changes linked to metastatic behavior[38]. Alizadeh et al. utilized PCA to quantify morphological metrics in breast cancer and osteosarcoma cell lines and achieved high accuracy in distinguishing metastatic cells from non-cancerous cells[39]. In addition, Kang et al. integrated PCA with deep learning to analyze cancer-associated fibroblasts (CAFs) and demonstrated that morphodynamic and motile features can effectively classify heterogeneous CAF phenotypes[41].

Expanding on these statistical approaches, recent advancements in deep-learning-based representation learning have introduced powerful alternatives for lower-dimensional feature extraction. These include GMVAE, TS2Vec[42,43], and other unsupervised deep clustering or self-supervised frameworks such as VaDE, DEC, and Barlow Twins[44–46]. These methods learn nonlinear transformation of the data that can enhance separability by jointly optimizing latent cluster structures (GMVAE, VaDE, DEC) or by enforcing invariance across augmented views or temporal contexts (Barlow Twins, TS2Vec)[42–46]. However, such approaches generally require substantial training data, careful hyperparameter tuning, and may produce inconsistent or non-reproducible results due to the stochastic nature of deep neural network optimization. Most importantly, the resulting latent representations are typically difficult to interpret biologically, as the learned nonlinear manifolds do not map transparently onto morphological or migratory variables. In contrast, the PCA-based approach used in this study provides transparent, reproducible, and direct interpretable axes of variations that can be directly related to underlying morphological and migration features, which is essential for mechanical insight.

Despite the power of unsupervised classification methods like PCA and UMAP in identifying clusters and classifying heterogeneous cell populations, these methods do not inherently reveal the underlying biological factors responsible for such distinctions. Therefore, the biological rationale behind clustering and classification remains challenging to interpret solely through statistical models. In this study, we investigated the impact of different extracellular matrix (ECM) environments—laminin (Ln), collagen 1 (Col1), and collagen 3 (Col3)—on HeLa cell migration and morphology. As different ECM proteins provide different physical and chemical cues, cells on different ECM proteins behave differently[47]. Our results demonstrated that cells on Ln exhibited significantly higher motility, increased turning angles, and reduced directional persistence, indicating an exploratory migration pattern. In contrast, cells on Col1 and Col3 demonstrated more directed migration with greater net displacement. Morphologically, cells on Ln showed larger areas and higher solidity, while those on Col1 and Col3 displayed more elongated shapes.

To systematically classify these differences, we applied PCA and effectively distinguished cells based on their motility characteristics. Analysis of PCA loadings identified key migration parameters—including turning angle, pause time, and displacement—as primary contributors to the differences between cells on Ln and collagens. Additionally, immunofluorescence imaging and cell-cell interaction analyses revealed distinct behavioral patterns among these groups. Building on these observations, we integrated PCA with biological phenotype analysis, including cytoskeletal and focal adhesion organization, to establish a systematic approach for classifying cell populations while uncovering the biological factors driving these distinctions. This methodology offers a framework for extracting meaningful biological insights from AI-assisted classification, addressing a key challenge in data-driven cellular analysis. Expanding this framework to models with defined pathophysiological characteristics could further enhance our understanding of cellular heterogeneity in disease, and therapeutic responses.

Results

Characteristics of migration on different ECMs

To investigate cellular migration and morphology across different ECMs—Ln, Col1, and Col3—we used HeLa cells. Given that ECM components play a crucial role in regulating cell adhesion, migration, and morphology, we anticipated that modifying the ECM would provide a straightforward yet effective way to influence these characteristics[47]. Cells were seeded on ECM-coated dishes and imaged for 12 h using a microscope equipped with an incubator. Cell detection and tracking in the time-lapse images were performed using YOLOv7 and the SORT algorithm. Detailed methodology is provided in the "Materials and methods" Section.

Figure 1 shows an analysis of cell migration trajectories and motility parameters on Ln, Col1, and Col3. Figure 1A displays representative migration trajectories, where cells on Ln appear to migrate within a more confined area compared to those on Col1 and Col3. To quantitatively evaluate migration dynamics, key parameters such as total migration distance, displacement, directness, turning angle, velocity variability, and minimum enclosing circle diameter (MEC) diameter were extracted from the trajectories (Fig. 1B–L). Schematic illustrations of these migration-associated parameters and turning angle calculations, along with their detailed definitions and numerical values, are provided in Supplementary Fig. S1, Tables S1, and S2.

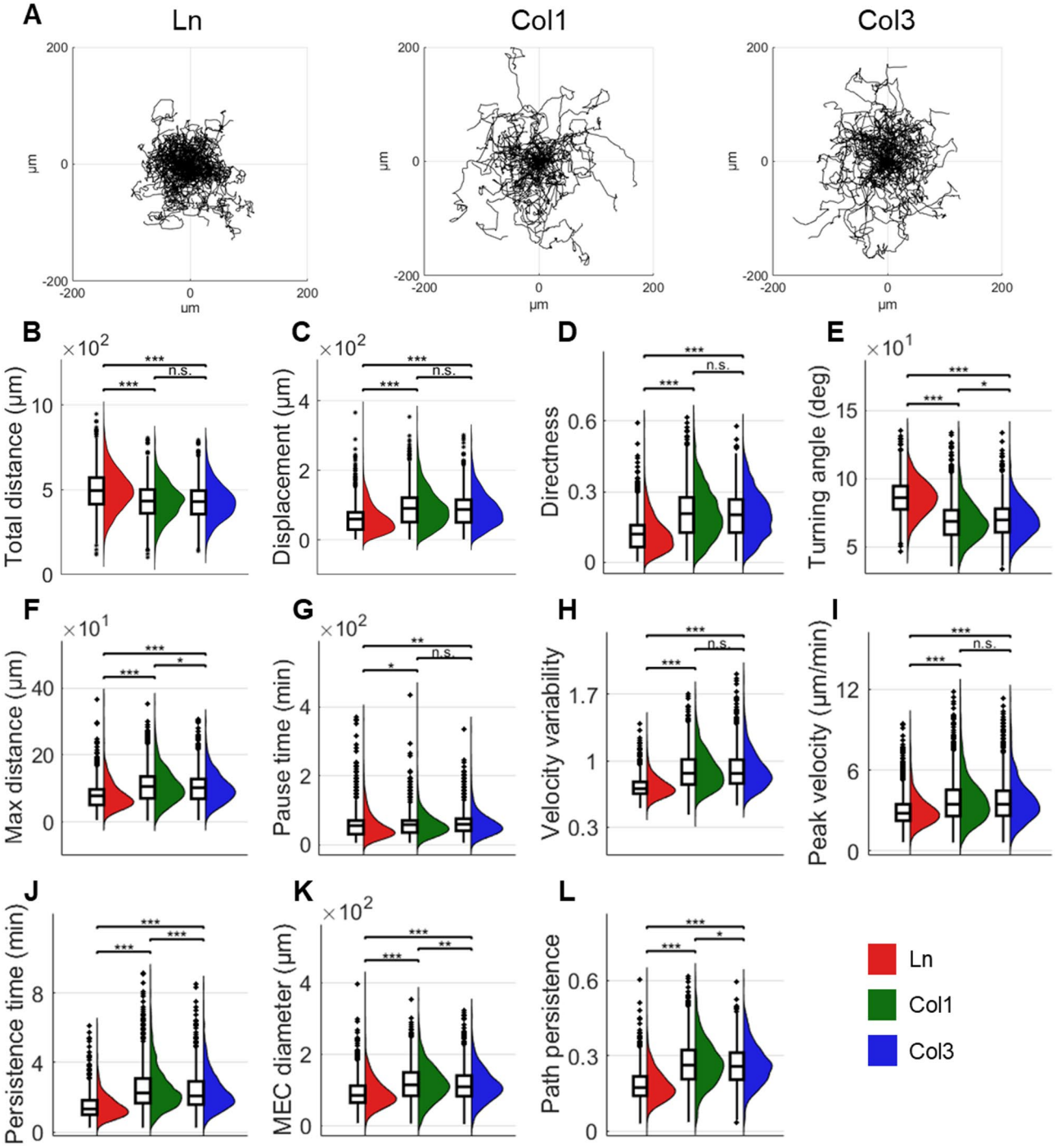

Fig. 1. Quantitative analysis of HeLa cell migration dynamics on Laminin (Ln) and Collagen (Col1, Col3) substrates. (**A**) Representative trajectories of HeLa cells migrating on Ln, Col1, and Col3-coated surfaces ($n = 54$ cells randomly selected per group). (**B**–**L**) Violin plots illustrating the distributions of 11 migration-associated parameters. Data were pooled from three independent experiments (Total cells analyzed: Ln, $n = 1{,}960$; Col1, $n = 1{,}602$; Col3, $n = 1{,}680$). Significant differences are marked by $^{*}P < 0.05$, $^{**}P < 0.01$, and $^{***}P < 0.001$, whereas n.s. represents non-significant results.

Contrary to the initial visual impression from the trajectories, Fig. 1B shows that cells on Ln exhibited significantly greater motility, with a total migration distance of 493.84 ± 121.3 μm, which is approximately 14% and 15% greater than the distances on Col1 (430.38 ± 107.64 μm) and Col3 (428.58 ± 106.35 μm), respectively. However, as shown in Fig. 1C, the net displacement of cells on Ln was markedly lower than on Col1 and Col3. This discrepancy suggests that cells on Ln exhibited less directional migration, and this trend was corroborated by the significantly reduced directness observed in Fig. 1D. Turning angle and mean squared displacement (MSD) analysis further supports these findings. Cells on Ln exhibited an increased mean turning angle of 86.1 ± 12.6°, compared to 69.0 ± 14.7° on Col1 and 70 ± 13.7° on Col3 (Fig. 1E). The higher turning angle reflects more frequent directional changes, as visualized in rose plots (Supplementary Fig. S2A). These trends were also evident in the MSD analysis, where the computed slope was 0.91 for cells on Ln, indicating sub-diffusive migration, compared to 1.17 and 1.22 for Col1 and Col3, respectively (Supplementary Fig. S2B).

Additionally, the maximum distance traveled from the initial position (i.e., maximum displacement, Fig. 1F) was shortest for cells on Ln, reflecting their constrained migratory behavior. Cells on Ln exhibited the shortest pause time and the smallest speed variability. Additionally, their peak velocity, defined as the highest speed reached by each cell during the observation period, was the lowest among the three conditions, consistent with a relatively moderate but stable migration pattern (Fig. 1G and H, and 1I). Consistent with these observations, cells on Ln exhibited the lowest persistence time (Fig. 1J), MEC diameter (Fig. 1K), and path persistence (Fig. 1L) compared to those on Col1 and Col3. These results reinforce that cells on Ln undergo frequent directional changes and occupy a smaller spatial region during migration. Interestingly, while significant migratory differences were observed between cells on Ln and those on Col1 and Col3, the differences between Col1 and Col3 were minimal overall. Although the turning angle and maximum displacement differences between Col 1 and Col 3 were statistically significant, these differences were minor and did not substantially alter the overall similarity in their migratory behaviors. Thus, the migratory characteristics of cells on Col 1 and Col 3 remained largely comparable, yet distinct from those on Ln.

Characteristics of morphology on different ECMs

From the same time-lapse images used in the migration analysis, cellular morphology was assessed at the 12-hour time point, corresponding to the final frame of the time-lapse sequence. Cell segmentation was performed using the Cellpose 2.0 algorithm (cellpose.org), which was customized for enhanced accuracy by training with a dataset of 362 manually annotated phase-contrast images of HeLa cells obtained from independent experiments. A detailed description of the Cellpose 2.0-based segmentation methodology is provided in the "Materials and methods" Section. The segmented cell boundaries are shown in Fig. 2C, corresponding to the phase-contrast images in Fig. 2B. As shown in Fig. 2A, cells on Ln exhibited a more spread-out morphology with broader attachment to the surface, resulting in larger, less elongated shapes compared to cells on Col1 and Col3. To quantitatively evaluate these morphological differences, parameters including cell area, perimeter, solidity, and aspect ratio were extracted from binary cell segmentation masks (Fig. 2D–G).

As can be seen in Fig. 2D, cells on Ln had the largest area (2194.9 ± 989.2 μm^2), significantly greater by approximately 22% and 28% than those on Col1 (1721.5 ± 954.9 μm^2) and Col3 (1754.0 ± 979.3 μm^2), respectively. Similarly, the perimeter of cells on Ln (247.1 ± 77.3 μm) was greater by 8% and 6% compared to Col1 (228.7 ± 97.3 μm) and Col3 (232.4 ± 101.0 μm), respectively (Fig. 2E). In addition, solidity was highest in cells on Ln. The solidity of cells on Ln (0.72) was approximately 3% higher than that of cells on Col1 (0.70) and Col3 (0.70) (Fig. 2F). Figure 2G demonstrates that cells on Ln had the lowest aspect ratio, indicating a less elongated morphology. While the differences in both solidity and aspect ratio were relatively small, the large sample size—approximately 1600 cells per experimental group—provided robust statistical power to confirm these differences as significant. Overall, these results suggest a meaningful relationship between substrate type and cell morphology. Furthermore, these quantitative findings are consistent with the morphological observations in Fig. 2A, where cells on Ln appear larger and less elongated compared to those on Col1 and Col3.

PCA of migration and morphology parameters

The results demonstrate significant differences in migration and morphology parameters among the experimental groups, particularly between Ln and collagens (Col1 and Col3). Cells on Ln exhibit fast yet sub-diffusive migration and larger spreading areas with a more rounded morphology, which distinguish them from cells on Col1and Col3. These observations motivated us to employ an integrative method to classify the experimental groups by utilizing 18 distinct parameters representing migration and morphology.

To analyze these 18-dimensional features of migration and morphology across three groups, we applied Principal Component Analysis (PCA), a dimensionality reduction method that projects high-dimensional data into a lower-dimensional space defined by principal components (PCs). PCA identifies the directions of maximum variance within the data. The first two components, PC1 and PC2, accounted for 27.95% and 22.68% of the total variance, respectively (Fig. 3A). Feature weightings (Fig. 3B) revealed that PC1 is primarily influenced by morphological traits such as circularity, solidity, perimeter, and complexity, while PC2 is dominated by migration features, including tortuosity, turning angle, and maximum distance. These contributions are further visualized in the correlation biplot (Fig. S3).

Figure 3C shows the distribution of data in pairs of PCs, with covariance ellipses for each group drawn at a standard deviation of 2. Covariance ellipses represent the region where approximately 95% of the data points are expected to lie, based on the spread and orientation of the data in the PC space. In the PC1-PC2 space, Ln exhibited a distribution that overlapped substantially with collagens along the PC1 axis but was notably distinct along the PC2 axis (Fig. 3C). This positioning suggests that Ln shares characteristics with the collagens in terms of PC1 (morphology) while being distinct in PC2, with motion patterns characterized by higher tortuosity and turning behavior.

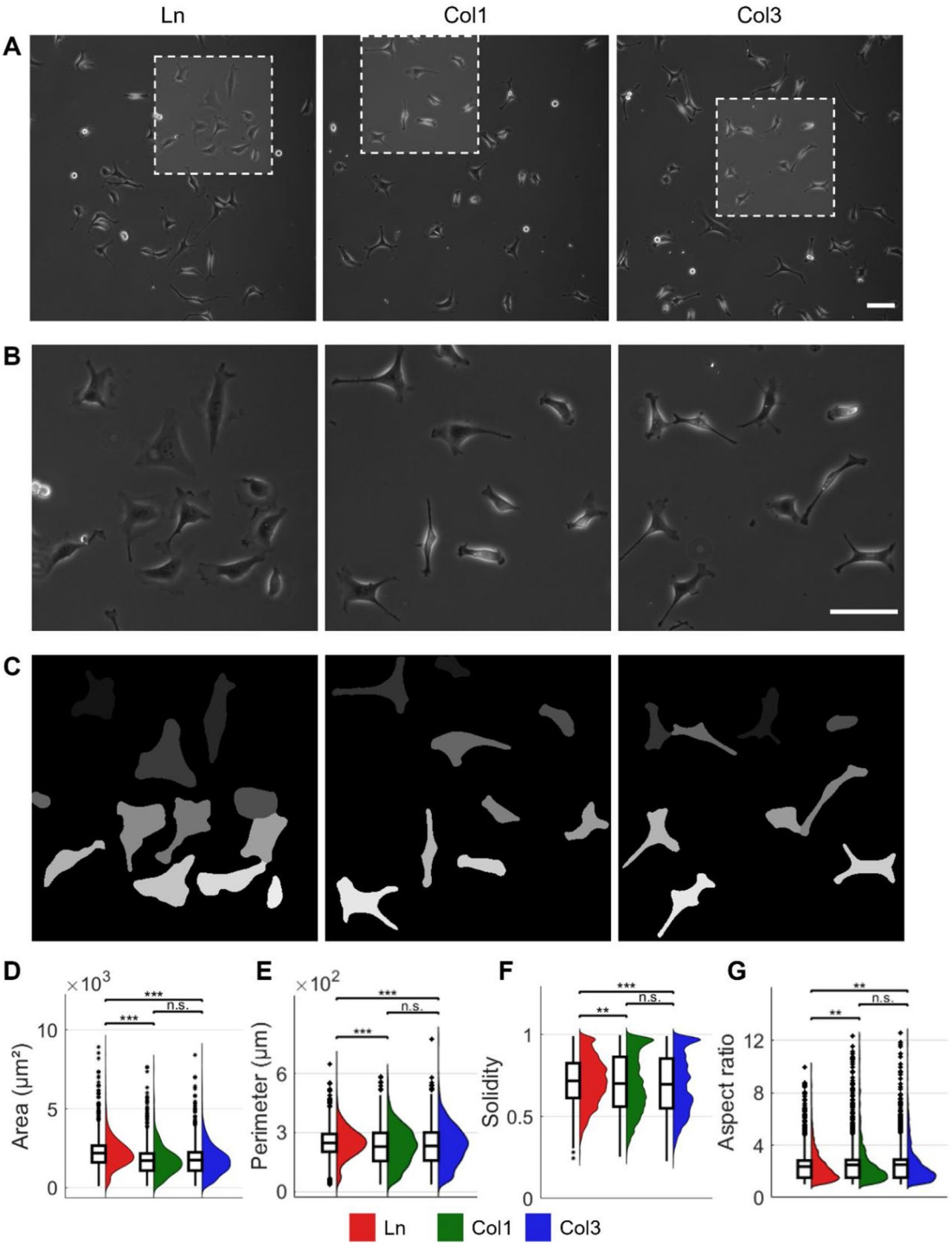

Fig. 2. Morphological changes in HeLa cells. (**A**) Phase-contrast images of HeLa cells on different ECM proteins with dotted boxes indicating magnified regions. Scale bar, 50 μm. (**B**) Magnified views of cells showing detailed morphology. Scale bar, 50 μm. (**C**) Segmentation of HeLa cells detected using Cellpose, trained on 362 manually annotated HeLa cell images. (**D**–**G**) Violin plots illustrate key morphological characteristics of cells. Total cell numbers analyzed were 1,957 cells for Ln, 1,601 cells for Col1, and 1,680 cells for Col3, obtained from three independent experiments.

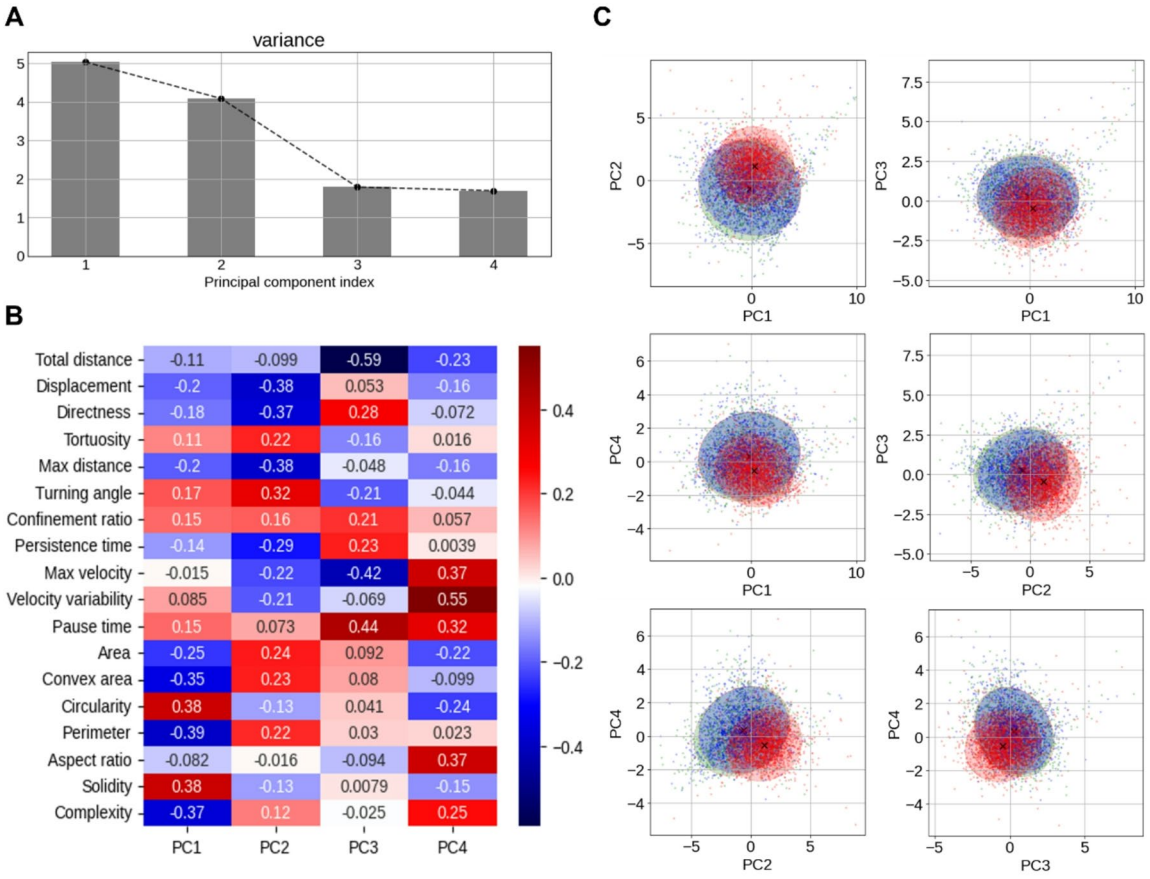

Fig. 3. Embedding of 18 combined migration and morphology features into a 4-dimensional principal component (PC) space. (**A**) Variance explained by each of the first four PCs, illustrating their contributions to the total variability. (**B**) Feature weightings for each PC, representing the contributions of the 18 migration and morphology features. (**C**) Two-dimensional projections of the data onto each pair of PCs, visualized with covariance ellipses to indicate group distributions (red: Ln, green: Col1, blue: Col3). Total cell numbers analyzed were 1,957 cells for Ln, 1,601 cells for Col1, and 1,680 cells for Col3, obtained from three independent experiments.

Separate PCA of each of migration and morphology parameters

Given that Ln was more distinct from collagens along the PC2 axis in the previous analysis, we hypothesized that migration features, which dominate PC2, contribute more significantly to differentiating Ln from Col1 and Col3. To explore this hypothesis, we analyzed migration, and morphology features separately in the principal component space, as illustrated in Fig. S4. In both cases, Col1 and Col3 exhibit similar distributions that are broader, and more variable compared to Ln. For morphology features (Fig. S4A), Ln's distribution is almost entirely contained within that of the collagens, indicating that Ln shares a subset of morphological characteristics with Col1 and Col3. In contrast, for migration features, Ln's distribution is more distinct from the collagens. Ln clearly separates from Col1 and Col3 in the PC1-PC2 space (top left panel of Fig. S4B or Fig. 4C), with similarly sized covariance ellipses and collagens and Ln's centroids are positioned in the second and fourth quadrants, respectively, suggesting distinct migration patterns.

Figure 4 provides details of the PCA results for migration features. Figure 4A illustrates the variance explained by the first four principal components, with PC1 and PC2 accounting for 38.16% and 16.49% of the total variance, respectively, capturing a substantial portion of the migration data variability. Figure 4B shows the weightings of individual features, indicating their contributions to each principal component. Figure 4C visualizes the data projection in the PC1-PC2 space.

PC1 distinguishes between features such as displacement, maximum distance, and directness, which correspond to lower PC1 values, and turning angle and tortuosity, which are associated with higher PC1 values. Thus, cells with higher PC1 values can be considered to exhibit more curved, tortuous movement patterns, while those with lower PC1 values can be considered to demonstrate linear, longer-distance motion. PC2 captures variations where lower PC2 values are linked to greater total distance traveled, and higher PC2 values correspond to increased pause time, directness, and variability in velocity. Therefore, movements with higher PC2 values can be considered to involve frequent pausing and variable speeds, while lower values reflect more continuous and uniform motion.

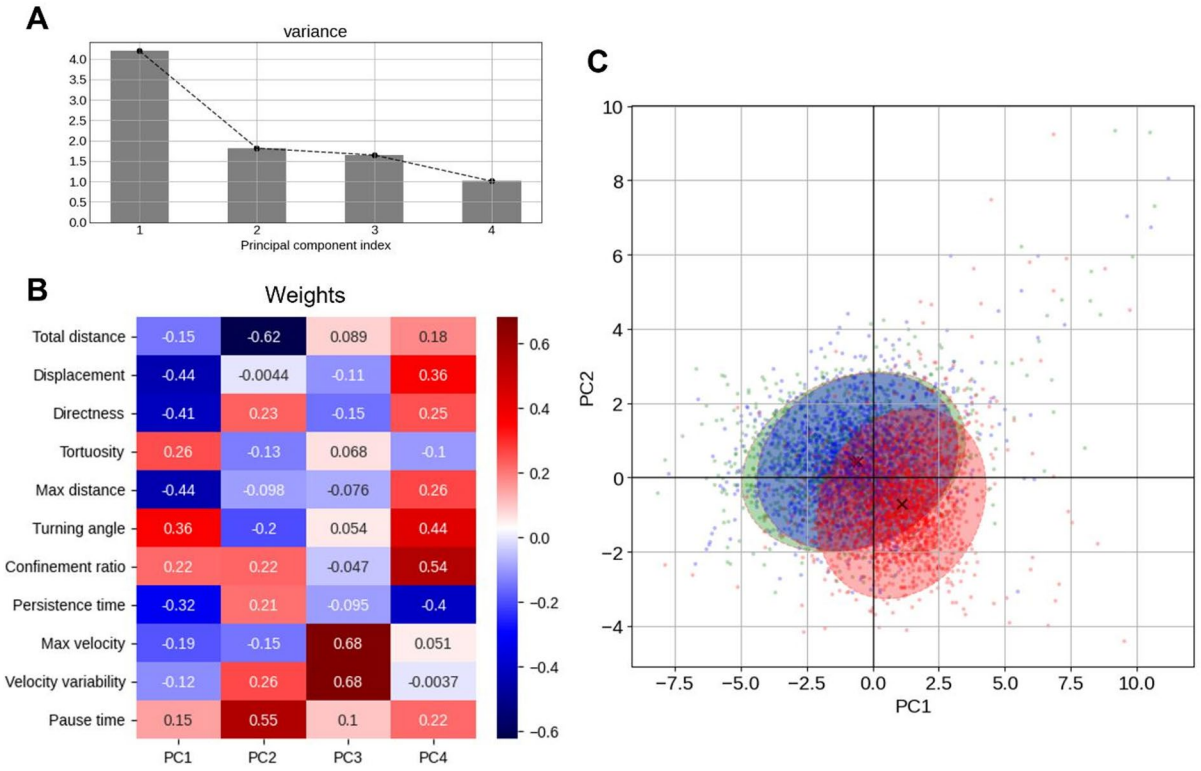

Fig. 4. PCA for migration features. (**A**) Variance explained by each of the first four PCs, representing their contributions to the total variability. (**B**) Feature weightings for each PC, highlighting the contributions of the 11 migration features. (**C**) Projections on PC1-PC2 space.

In the PC1-PC2 space, Col1 and Col3 are predominantly located in the second quadrant (low PC1, high PC2), indicating shorter distances, frequent pausing, and variable speeds, characteristic of confined or intermittent motion. Ln, in contrast, is primarily found in the fourth quadrant (high PC1 and low PC2), reflecting movements with greater tortuosity and frequent turning but reduced overall distances traveled. These distinctions reveal the behavioral differences captured by the PCA and are consistent with our earlier findings: cells on Ln exhibit exploratory migration with increased turning angles and reduced directional persistence, while cells on Col1 and Col3 display more directed movement. To further accentuate these divergent characteristics, we focused on representative subsets within the PCA space, which yielded a more distinct separation of migration profiles between the groups (Figs. S5 and S6).

Validation of PCA as a classifying method between cells on different ECMs

We applied kernel density estimation (KDE), a non-parametric technique, to estimate the probability density of each group. This approach provides a more detailed understanding of the data distribution beyond what is captured by covariance ellipses. Figure 5 presents the KDE results, showing the separation between Ln and collagens. In addition, the peak densities distinctly located in the second and fourth quadrants, consistent with the distributions observed through covariance ellipses in the PCA plot. This consistency reinforces the validity of our approach using covariance ellipses to capture group separation in the PCA space.

To further validate the KDE results, we applied the same method to an independent dataset, depicted as dashed lines in Fig. 5. This independent dataset was obtained under slightly different experimental conditions, where twice as many cells (2×10^4 cells/dish) were seeded compared to the experiments earlier (1×10^4 cells/dish). This change in cell density may have influenced the distance between cells. Despite the altered experimental conditions, the KDE results for the independent dataset intriguingly show a similar separation between the groups and comparable marginal distributions along PC1 and PC2, confirming the robustness and reliability of our classification methods via PCA. This KDE results confirm the distinct characteristics of Ln and collagens in migration and morphology, effectively revealing the underlying structure of the data.

In addition, to confirm whether the classifications identified through PCA are consistent with other dimensionality reduction methods and to validate our findings, we applied t-distributed Stochastic Neighbor Embedding (t-SNE). t-SNE is a nonlinear dimensionality reduction technique that focuses on preserving the local structure of data. The data was embedded into a two-dimensional t-SNE space, as shown in Fig. S7. Unlike PCA, which emphasizes capturing global variance, t-SNE focuses on preserving the local structure of the data, making it particularly effective for revealing clusters and relationships at finer scales. Figure S7 shows the t-SNE embedding for Ln and collagens, with data from PCA Regions 1 (Ln-exclusive) and 3 (collagens-exclusive)

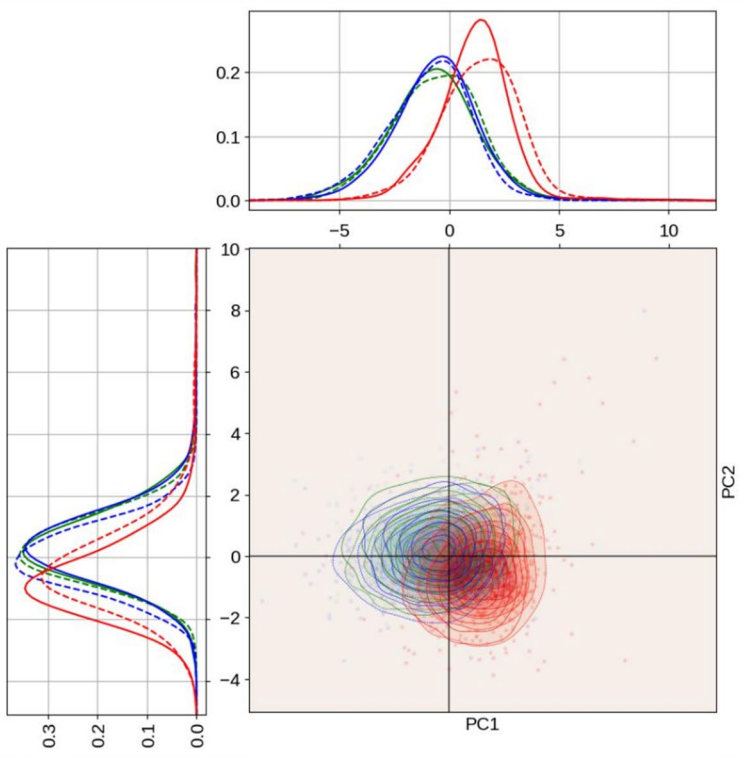

Fig. 5. Kernel density estimation (KDE) of migration features for each group in the PC1-PC2 space. Bold lines depict the contour lines of the estimated densities, while dashed lines represent the contour lines from an independent dataset. The side plots on the left and bottom show the marginal distributions along the PC2 and PC1 axes, confirming the robustness of the estimated densities.

clearly separated. This distinct separation in t-SNE reinforces the observations from PCA, demonstrating that the differentiation between Ln and collagens is robust across different dimensionality reduction methods.

To evaluate the versatility of our analytical framework, we extended the analysis to the non-cancerous mammary epithelial cell line, MCF10A. While the overall response to the extracellular environment shared similarities with HeLa cells, a notable cell-type-specific difference was observed in total distance; MCF10A cells on Ln exhibited significantly shorter migration distances compared to those on collagens (Fig. S8). This behavior likely reflects the inherent nature of normal epithelial cells, which tend to establish a more stable, stationary phenotype upon contact with laminin, a key component of the basement membrane[48,49]. Nevertheless, PCA effectively captured the environmental influence, with MCF10A cells displaying distinct clusters on the PC1-PC2 plane for Ln versus collagens (Fig. S9). These results demonstrate that our PCA-based analysis is an effective tool for characterizing variations in cell migration patterns induced by different extracellular environments, confirming its robustness across different cell models.

Biological phenotype underlying the distinct classification in PCA method

Based on the distinct separation observed in the PCA and t-SNE analyses, we aimed to investigate how the differences between cells on Ln and collagens correlate with their underlying biological phenotypes. As shown in Fig. 4, larger turning angles, longer pause times, and decreased displacement in cells on Ln were key factors contributing to their separation from the cells on collagens in the PC1-PC2 space. Based on these, we hypothesized that cells on Ln might exhibit a higher frequency of interactions with neighboring cells compared to those on collagens. To test this, we analyzed cell-cell interaction behaviors by evaluating the frequency and duration of contact during migration. Using time-lapse imaging of HeLa cells over a 12-hour period, we selected a cell in the initial frame and manually counted the number of contacts and their durations (see supplementary videos S1, S2, and S3). Time-lapse images were captured at 5-minute intervals, and cell-cell interactions were categorized into two types: transient contacts and sustained contacts. Transient contacts were defined as direct cell-cell contacts that were visible for less than one frame interval (i.e., less than 5 min). Sustained contacts were defined as interactions lasting for more than two consecutive frames (i.e., lasting longer than 5 min).

In Fig. 6A, the number of contacts was significantly higher on Ln (8.1 ± 4.3 contacts/cell) compared to Col1 (6.2 ± 3.5 contacts/cell) and Col3 (6.1 ± 2.8 contacts/cell). Transient contacts were most frequent on Ln (1.08 ± 1.13 contacts/cell) compared to Col1 (0.88 ± 1.24 contacts/cell) and Col3 (0.52 ± 0.76 contacts/cell) (Fig. 6B). Additionally, the duration of cell contacts was longest on Ln (281.8 ± 176.5 min) compared to Col1 (183.3 ± 131.1 min) and Col3 (251.4 ± 154.3 min) (Fig. 6C). While some differences between Ln and collagens were not statistically significant, the overall trends confirmed our hypothesis: cells on Ln engaged in more cell-

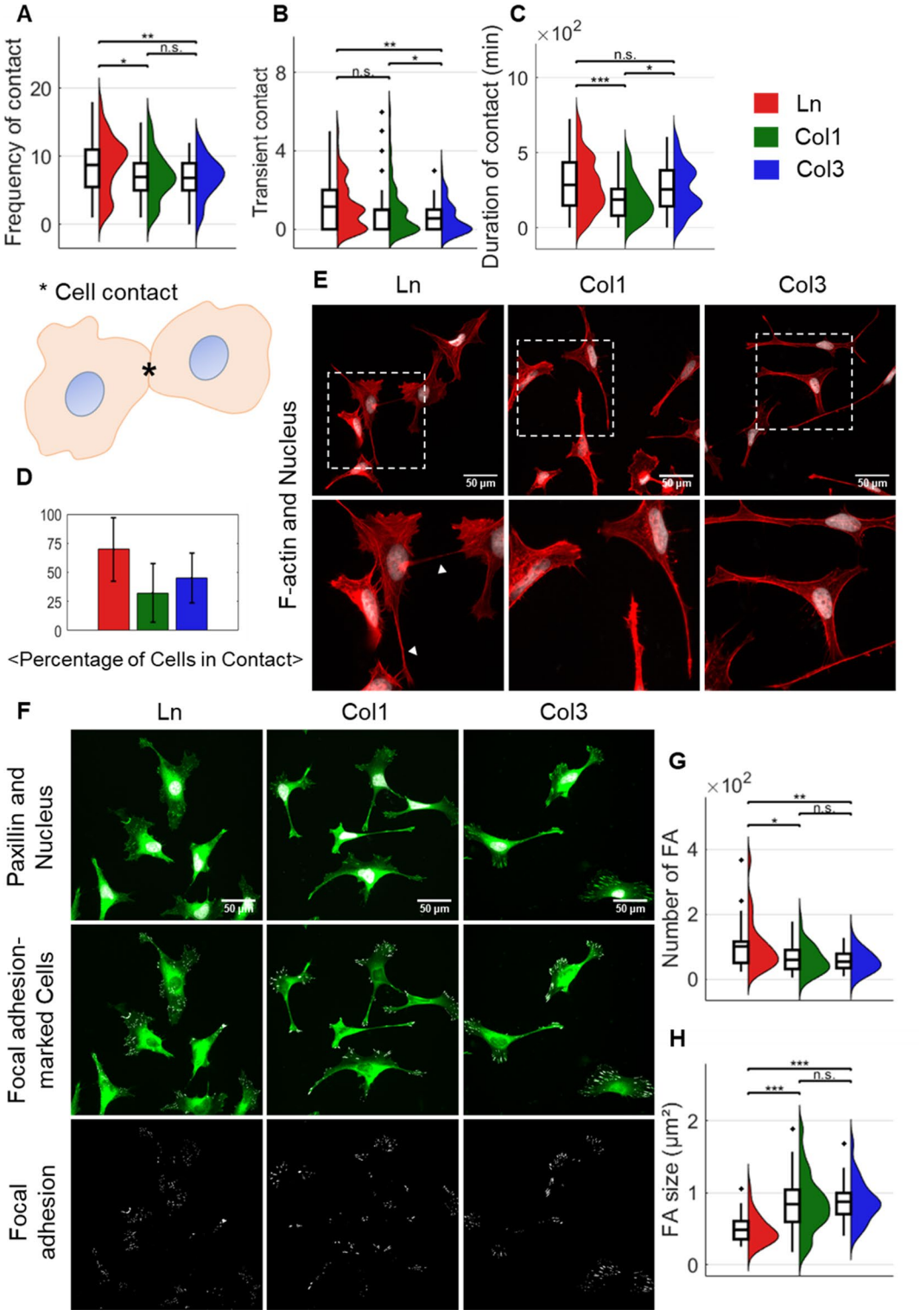

Fig. 6. Changes in the cytoskeleton, cell-cell interactions, and focal adhesions of HeLa cells under different ECM conditions. (**A**) The Number of cell-cell contacts during 12 h of migration. (**B**) The number of transient cell-cell contacts, defined as interactions lasting less than 5 min. (**C**) Duration of cell-cell contacts during cell migration. (**D**) Proportion of cells in contact with neighboring cells, as determined from actin-stained images. (**E**) Actin-stained images of HeLa cells. (**F**) Quantification of focal adhesions using paxillin immunofluorescence staining. (**G**) Comparison of the number of focal adhesions per cell across different ECM conditions. (**H**) Comparison of focal adhesion size across different ECM conditions. Data for (**A**–**D**) were obtained from approximately 60 cells per group, and data for (**G**, **H**) were obtained from approximately 30 cells per group.

cell interactions by contacting neighboring cells more frequently and for longer durations compared to those on collagens.

To examine the cellular structures involved in cell-cell interactions more closely, we then stained and imaged actin stress fibers using phalloidin and focal adhesions using paxillin antibody. In Fig. 6E, cells on Ln exhibited more protrusive filopodia at their periphery (indicated by white arrowheads), which extended toward neighboring cells. In contrast, cells on collagens displayed fewer and shorter filopodia, with smoother boundaries compared to those on Ln. To quantitatively correlate these filopodia structures and cell-cell interactions, we analyzed the proportion of cells in contact with their neighbors through filopodia, based on the actin images. As can be seen in Fig. 6D, cells on Ln exhibited a significantly higher proportion of filopodia mediated interactions compared to the other groups. These results align well with previous results related to the cell-cell contact frequency and duration.

Since focal adhesions (FAs) connect the actin stress fibers to the ECMs, they play a crucial role in force transmission and are essential for cell migration[50]. The characteristics of FAs are therefore closely linked to cellular morphology and migration. To further investigate this, we analyzed the FA properties of cells across the three different ECMs. Paxillin, a key component of FAs, was immunostained and imaged to visualize the structures. Figure 6F presents the immunofluorescence images of paxillin and outlines the quantification process for FAs. To quantify the number and size of FAs, we manually annotated paxillin-positive dots from the fluorescence images using ImageJ. These annotated regions were converted into binary masks for further analysis. As shown in Fig. 6G and H, cells on Ln (100.7 ± 76.8) exhibited a significantly higher number of FAs than those on collagens (59.4 ± 39.7 in Col1 and 54 ± 29.2 in Col3). However, the size of these FAs on Ln (0.5 ± 0.2 μm^2) was significantly smaller, nearly half the size of those observed on Col1 (0.8 ± 0.4 μm^2) and Col3 (0.9 ± 0.3 μm^2).

Discussion

This study demonstrates the impact of different ECMs—Ln, Col1, and Col3—on the migration, and morphology of HeLa cells. Using machine learning-based analyses of migration and morphology parameters, we identified key features of cellular behavior across these ECMs. PCA effectively classified cells, particularly based on migration characteristics, and revealed a clear distinction between cells on Ln and those on collagens. This classification further helped uncover latent biological patterns, such as increased cell-cell interactions and the prominence of filopodia of the cells on Ln.

While migration parameters demonstrated a clear separation between cells on Ln and collagens, the distinction based on morphological features was more subtle. This subtle distinction may be attributed to the limitations of phase-contrast imaging, which provides an overall view of cell morphology through refractive index contrast with the surrounding medium but lacks the resolution to capture fine structural details such as filopodia and lamellipodia (Fig. 6D). Incorporating additional factors into morphological analyses, such as the temporal changes in cell morphology over time, could provide additional insights into cellular behavior. For instance, tracking the dynamic evolution of cell morphology during migration could reveal critical features, such as the formation and retraction of protrusions, that are closely linked to cell-cell interaction events. Previously Kang et al. demonstrated that cancer-associated fibroblast (CAF) subsets could be effectively classified using a multifaceted framework that combined motility, and morphodynamic metrics[41]. In addition, Elbez et al. also introduced a machine learning framework utilizing morphodynamic metrics to classify cancer cells based on time-dependent morphology[38]. These studies show the potential of integrating time-resolved morphological data to reveal the plasticity of cell states and the heterogeneity within cancer populations. Building on such strategies, incorporating morphodynamic analyses into our PCA or other classification methods could improve group separation and provide a more comprehensive understanding of how cells interact with and adapt to different ECM environments.

Looking ahead, the integration of advanced computer vision architectures, such as the Vision Transformer (ViT), could serve as a crucial framework for future studies seeking to capture these complex morphodynamics[51]. Recent shifts toward generalist models, including Segment Anything (SAM), microscopy-adapted variants (e.g., micro-sam), and hybrid pipelines like Cellpose-SAM, represent a promising shift toward more robust and promptable segmentation in bioimage analysis[52–55]. Furthermore, as transformer-based object detectors (e.g., DETR, DINO, and RT-DETR) and video-capable frameworks (e.g., SAM 2) continue to mature, their future application could enable the high-accuracy automated tracking of dynamic morphological transitions that remain challenging to capture with the analytical framework employed in this study[56–59].

In a broader biological context, the increased interactions observed between cells on Ln in this study are consistent with findings from previous research. For instance, Noel et al. demonstrated that Ln or laminin-enriched basement membranes promote cell-cell interactions by facilitating the formation of increased cell-cell contacts[60]. Furthermore, previous studies have shown that Ln enhances filopodia formation and supports dynamic motility through interactions with the α6β4 integrin and actin stabilization[61]. This adhesive interaction between Ln and α6β4 integrin has been identified as a critical mechanism driving tumor invasion and metastasis. Consistent with this, multiple studies have reported that Ln is upregulated in the tumor microenvironment[62–64]. A study by Fullár indicated that this upregulation could result from increased secretion by CAFs, as observed in cervical cancers[62]. Elevated Ln not only augment cancer cell adhesion and motility but also stimulates matrix metalloproteinases (MMPs) secretion[62,63,65]. These MMPs degrade the ECM, thereby facilitating cancer cell invasion and significantly increasing the risk of metastasis. These findings can be linked to the studies revealing that Ln-rich microenvironments have been shown to promote epithelial-to-mesenchymal transition (EMT), a key process in cancer progression and dissemination[63–66].

In contrast to laminin-binding integrins, collagens are primarily recognized by integrins such as α1β1, α2β1, and α11β1, which are crucial in mediating stable cell-ECM adhesion and tension-bearing functions[67–69]. In

particular, one of the key collagen-binding integrins, α2β1 integrin, is known to be involved in focal adhesion formation and its role in generating cellular contractility[69,70]. These prior reports align well with our observations of increased FA size and more directed migration in cells on collagens. We speculate that the distinct migratory phenotypes observed on Col1 and Col3 may be driven by the activation of collagen-specific integrins, resulting in a more stable adhesion–migration balance compared to the dynamic, exploratory motility supported by laminin-α6β4 interactions.

These insights provide important context for interpreting the unique FA characteristics observed in cells on different ECM substrates. Specifically, smaller FA sizes in cells on Ln are likely associated with their enhanced migration speed and frequent directional changes. The FAs in cells on Ln appear to remain small due to faster turnover rates, which likely facilitate the rapid adjustments required for dynamic motility. Typically, nascent FAs either undergo turnover or maturation, and previous studies have shown that rapid FA disassembly or turnover is essential for efficient cell migration[71–74]. In line with this, several studies also have demonstrated that larger FAs are commonly accompanied by slower cell migration[73,75,76]. In our study, cells on Ln, characterized by smaller FAs, exhibited faster migration, whereas cells on collagens, characterized by larger FAs, are shown to have slower and more directed migration. Similarly, Meehan et al. reported that cells with larger FAs exhibits slower and more directed migration[75]. These reports support our finding in FA size and turnover in driving adaptive migratory behaviors, especially within Ln-enriched microenvironments.

This study demonstrates the utility of a machine learning-integrated platform for analyzing cancer cell dynamics, leveraging PCA to combine morphological and migratory parameters into a unified framework. By effectively distinguishing cell behaviors across different extracellular matrix conditions, this method offers a scalable and robust approach to studying complex cellular processes. The integration of high-dimensional analysis with automated imaging and tracking techniques establishes a powerful tool for investigating cancer cell behavior in response to microenvironmental cues. The ability of this platform to identify distinct phenotypic patterns, such as those observed on Ln, highlights its potential for characterizing metastatic potential in cancer cells. Beyond its current application, we believe that this methodology can be extended to drug discovery and testing, particularly for evaluating the impact of therapeutic candidates on cell migration and adhesion dynamics.

Conclusions

In this study, we presented a versatile and quantitative platform for cell mechanics analysis, providing a robust framework for investigating cellular heterogeneity and behavior across different microenvironmental conditions. By integrating machine learning-based statistical analysis with biological interpretation, we provided a detailed quantitative characterization of how ECM composition shapes cancer cell dynamics, demonstrating the platform's capacity to capture subtle behavioral shifts. While this methodology was applied here to simplified ECM variations, its feasibility suggests broader applicability to more complex, physiologically relevant systems. Moving forward, this platform could be expanded to study ECM remodeling in disease progression, tumor metastasis, and therapeutic response. Ultimately, this work offers an effective analytical framework for advancing our understanding of cell-microenvironment interactions, with potential for both fundamental research and translational applications.

Materials and methods

Cell culture

HeLa cells used in this experiment were purchased from Korea Cell Line Bank. The cells were cultured in Modified Eagle's Medium (MEM, Welgene; #LM007-01) supplemented with 10% fetal bovine serum (FBS, Sigma; #F2442-500 ml) and 1% antibiotic-antimycotic (genDEPOT; #CA002-010) at 37 °C in a 5% CO_2 environment. Cells were maintained in 100 mm dishes and were washed with PBS and trypsinized using Trypsin-EDTA (Welgene; #LS010-10) every 2–3 days. Cell confluency was maintained below 80% throughout the culture period to ensure optimal growth conditions.

Surface coating by ECM

Type l collagen (Advanced Biomatrix; #5005-B), Type lll collagen (Advanced Biomatrix; #5021), and Laminin (Gibco; #23017015) were coated onto confocal dishes (SPL; #101350). Each ECM solution was diluted to a concentration of 50 μg/mL in PBS, and 465 μL was applied to the glass surface of each dish. The ECM-coated dishes were incubated overnight at 4 °C and washed with PBS prior to cell seeding.

Live cell imaging

HeLa cells were seeded at a density of 1×10^4 cells/dish and incubated for 2 h. The dishes were then transferred to a live-cell imaging chamber maintained at 37 °C with 5% CO_2 and appropriate humidity control. Imaging was performed using a Leica DMi8 microscope (Leica, Germany) equipped with a 10× objective lens. During each imaging session, six confocal dishes were analyzed, with approximately 15 positions per dish captured at 5-minute intervals over 12 h. This imaging procedure was repeated across three independent experiments.

Paxillin-immunofluorescence staining

HeLa cells were seeded at a density of 2×10^4 cells per ECM-coated dish. After 14 h of incubation, the cells were fixed with 3.7% paraformaldehyde (AMRESCO; #0493-500 ml) at room temperature (RT) for 15 min and permeabilized with 0.2% Triton X-100 (Daejung; #9002-93-1) for another 15 min at RT. Blocking was performed with 3% BSA in PBS for 1 h at RT. The primary antibody, Paxillin (GeneTex; #GTX125891), was diluted 1:1000 in 1% BSA/PBS, and 300 μL of the solution was applied to each dish for overnight incubation at 4°C. The secondary antibody, Goat Anti-Rabbit IgG (H&L) Alexa Fluor 488 (abcam; #ab150077), was diluted 1:500 in 1% BSA/PBS

and applied for 1 h at RT. Rhodamine phalloidin (Invitrogen; #R415) was diluted at 1:400 in PBS and applied for 30 min at RT, followed by Hoechst 33342 (Invitrogen) staining at a dilution of 1:2000 in PBS for 10 min. After adding the mounting solution (Invitrogen; #P36980), samples were covered with glass coverslips. Between each step, dishes were washed three times with PBS. Imaging was performed with a Leica DMi8 microscope using a 40× water immersion objective lens (Leica).

Quantitative analysis of focal adhesions

Focal adhesions were quantified from immunofluorescence images stained for paxillin[77]. The analysis was performed using ImageJ by manually masking focal adhesions and calculating their numbers and areas. All focal adhesions, regardless of size or circularity, were included in the analysis. The resolution of the images used in our analysis was 1296 × 966 pixels (346.88 μm × 258.48 μm), and the focal adhesion size was calculated using a conversion factor of 0.0716 μm^2 per pixel.

Analysis of cell-cell contact frequency via live cell imaging

Phalloidin-stained images containing at least four cells per frame were selected for cell-cell contact analysis. The percentage of cell-to-cell contact was calculated from the selected images, and the overall contact percentage was averaged across all representative images. To measure contact during cell migration, additional live-cell imaging was performed. HeLa cells were seeded at a density of 2×10^4 cells per ECM-coated dish and imaged 2 h post-incubation. Four randomly selected cells per position were analyzed in the acquired TIF images, and the number and duration of contacts were manually counted using ImageJ.

Dataset preparation and model training

To achieve high-fidelity cell detection and tracking, we implemented two object detection models trained separately for HeLa and MCF10A cells based on the YOLOv7 (You Only Look Once, version 7) architecture, reflecting their distinct morphological characteristics. The models were trained using in-house datasets consisting of phase-contrast microscopy images acquired in our laboratory. For HeLa cell detection, we utilized a dataset of approximately 5,000 images, of which ~ 40% contained HeLa cells. For MCF10A detection, the dataset comprised approximately 1,700 images, also with ~ 40% composition of MCF10A cells. All images were manually annotated with bounding boxes using the LabelImg tool (Tzutalin, 2015). For both datasets, the images were randomly partitioned into training (80%) and validation (20%) sets for model development and hyperparameter optimization.

For cell segmentation aimed at morphological feature extraction, we curated a separate dataset of 362 phase-contrast images of HeLa cells acquired in our laboratory at the same magnification (10× objective lens) as the experimental time-lapse sequences. Pixel-level segmentation masks were manually annotated to serve as ground truth labels for training the Cellpose model.

Cell detection and tracking

Cell detection aimed at quantifying cell motility was automated using the YOLOv7 object detection model[78]. The YOLOv7 models were trained from scratch without pretrained weights. Training was conducted for 300 epochs with a batch size of 16 and input image size of 640 × 640 pixels. Stochastic gradient descent (SGD) with a momentum of 0.937 was used as the optimizer. The learning rate (LR) was initialized at 0.01 and decayed to 0.001 using a cosine annealing schedule, with a warm-up period of 3 epochs. Weight decay of 0.0005 was applied for regularization. Data augmentation was applied during training, including HSV-based photometric adjustments (hue ± 1.5%, saturation ± 70%, value ± 40%), random geometric transformations (translation ± 10%, scaling ± 50%, horizontal flipping, probability 0.5), and mosaic augmentation combining four images (probability 1.0).

Detected cells were tracked across time-lapse frames using the SORT (Simple Online and Real-time Tracking) algorithm. SORT applies a Kalman filter for motion prediction and uses Intersection over Union (IoU) matching of bounding boxes to maintain cell identities across frames. Cell motility parameters were calculated from the centroid trajectories of tracked cells using custom MATLAB code. Detailed mathematical definitions and formulations for all motility parameters are provided in supplementary table (Table S1).

Cell morphological segmentation

Cellpose 2.0 was employed for cell boundary detection and segmentation[79,80]. We fine-tuned the pretrained "cyto" model from the Cellpose 2.0 library, which was originally developed as a generalist algorithm for diverse imaging modalities. Fine-tuning was performed for 100 epochs with Adam optimizer with a LR of 0.1 and weight decay of 0.0001. Training utilized Cellpose's built-in data augmentation, which includes random rotations and intensity variations. The trained Cellpose model was applied to the final frames (12-hour time point) of the time-lapse sequences to generate binary segmentation masks. Cell morphological features were then subsequently computed from these masks using custom MATLAB code.

Performance evaluation

To assess YOLOv7 model performance on unseen data, we prepared an independent evaluation set comprising more than 500 manually annotated cells from HeLa or MCF10A cell images that were not included in the training or validation datasets. Model performance was quantified using the Jaccard index (J), a cell-level, count-form overlap metric defined as:

$$J = \frac{TP}{TP + FP + FN}$$

where TP, FP, and FN denote true positives, false positives, and false negatives at the level of individual cells, respectively (Fig. S10). When evaluated at an IoU matching threshold of 0.5, this metric is denoted $J_{0.5}$.

For YOLOv7 bounding box detection, raw model outputs were first filtered using a confidence threshold of 0.25. A predicted bounding box was counted as a TP if its bounding-box IoU with an unmatched ground-truth box was ≥ 0.5, using greedy one-to-one assignment (each ground-truth box matched at most once); unmatched predictions were counted as FP and unmatched ground-truth boxes as FN. Using this protocol, the trained YOLOv7 models achieved $J_{0.5}$ values of 0.88 for HeLa cell detection and 0.85 for MCF10A cell detection.

For Cellpose instance segmentation, unlike object detection models that output per-box confidence scores, Cellpose directly produces instance masks through pixel-level flow field clustering and does not generate per-instance confidence scores; therefore, no confidence threshold was applied prior to evaluation. A predicted segmentation mask was counted as a TP if its pixel-wise IoU with an unmatched ground-truth mask was ≥ 0.5, using the same greedy one-to-one assignment. Cellpose performance was assessed on a separate test set of 882 manually annotated HeLa cells, achieving a $J_{0.5}$ of 0.90.

Statistical analysis

Statistical significance between groups was assessed using an unpaired t-test in MATLAB. P-values lower than 0.05 were considered statistically significant, with the symbols (*), (**), and (***) denoting p-values < 0.05, < 0.01, and < 0.001, respectively. Non-significant results ($p > 0.05$) were denoted as n.s.

Data availability

All data generated or analysed during this study are included in this manuscript.

Received: 7 August 2025; Accepted: 26 February 2026
Published online: 08 March 2026

References

1. Friedl, P. & Gilmour, D. Collective cell migration in morphogenesis, regeneration and cancer. *Nat. Rev. Mol. Cell Biol.* **10**, 445–457. https://doi.org/10.1038/nrm2720 (2009).
2. Giampieri, S. et al. Localized and reversible TGFβ signalling switches breast cancer cells from cohesive to single cell motility. *Nat. Cell Biol.* **11**, 1287–1296. https://doi.org/10.1038/ncb1973 (2009).
3. Montell, D. J. Morphogenetic cell movements: Diversity from modular mechanical properties. *Science* **322**, 1502–1505. https://doi.org/10.1126/science.1164073 (2008).
4. Lauffenburger, D. A. & Horwitz, A. F. Cell migration: A physically integrated molecular process. *Cell* **84**, 359–369. https://doi.org/10.1016/S0092-8674(00)81280-5 (1996).
5. Shaw, T. J. & Martin, P. Wound repair at a glance. *J. Cell. Sci.* **122**, 3209–3213. https://doi.org/10.1242/jcs.031187 (2009).
6. Trepat, X., Chen, Z. & Jacobson, K. Cell migration. *Compr. Physiol.* **2**, 2369–2392. https://doi.org/10.1002/cphy.c110012 (2012).
7. Wolfenson, H., Yang, B. & Sheetz, M. P. Steps in mechanotransduction pathways that control cell morphology. *Annu. Rev. Physiol.* **81**, 585–605. https://doi.org/10.1146/annurev-physiol-021317-121245 (2019).
8. Yeung, T. et al. Effects of substrate stiffness on cell morphology, cytoskeletal structure, and adhesion. *Cell. Motil. Cytoskel.* **60**, 24–34. https://doi.org/10.1002/cm.20041 (2005).
9. Stroka, K. M. & Aranda-Espinoza, H. Effects of morphology vs. cell-cell interactions on endothelial cell stiffness. *Cell. Mol. Bioeng.* **4**, 9–27. https://doi.org/10.1007/s12195-010-0142-y (2011).
10. Engler, A. J., Sen, S., Sweeney, H. L. & Discher, D. E. Matrix elasticity directs stem cell lineage specification. *Cell* **126**, 677–689. https://doi.org/10.1016/j.cell.2006.06.044 (2006).
11. Thakar, R. G. et al. Cell-shape regulation of smooth muscle cell proliferation. *Biophys. J.* **96**, 3423–3432. https://doi.org/10.1016/j.bpj.2008.11.074 (2009).
12. Mittal, V. Epithelial Mesenchymal Transition in Tumor Metastasis. *Annu. Rev. Pathol.* **13**, 395–412. https://doi.org/10.1146/annurev-pathol-020117-043854 (2018).
13. Lee, M. Y. & Shen, M. R. Epithelial-mesenchymal transition in cervical carcinoma. *Am. J. Transl. Res.* **4**, 1–13 (2012).
14. Wu, P. H. et al. Single-cell morphology encodes metastatic potential. *Sci. Adv.* **6**, eaaw6938. https://doi.org/10.1126/sciadv.aaw6938 (2020).
15. Nurmagambetova, A., Mustyatsa, V., Saidova, A. & Vorobjev, I. Morphological and cytoskeleton changes in cells after EMT. *Sci. Rep.* **13**, 22164. https://doi.org/10.1038/s41598-023-48279-y (2023).
16. Thiery, J. P. Epithelial-mesenchymal transitions in tumour progression. *Nat. Rev. Cancer* **2**, 442–454. https://doi.org/10.1038/nrc822 (2002).
17. Leggett, S. E. et al. Morphological single cell profiling of the epithelial-mesenchymal transition. *Integr. Biol.* **8**, 1133–1144. https://doi.org/10.1039/c6ib00139d (2016).
18. Leggett, S. E., Hruska, A. M., Guo, M. & Wong, I. Y. The epithelial-mesenchymal transition and the cytoskeleton in bioengineered systems. *Cell. Commun. Signal.* **19**, 32. https://doi.org/10.1186/s12964-021-00713-2 (2021).
19. Prewitt, J. M. & Mendelsohn, M. L. The analysis of cell images. *Ann. N. Y. Acad. Sci.* **128**, 1035–1053. https://doi.org/10.1111/j.1749-6632.1965.tb11715.x (1966).
20. Borst, H., Abmayr, W. & Gais, P. A thresholding method for automatic cell image segmentation. *J. Histochem. Cytochem.* **27**, 180–187. https://doi.org/10.1177/27.1.374573 (1979).
21. Otsu, N. A. Threshold selection method from gray-level histograms. *IEEE Trans. Syst. Man. Cybern.* **9**, 62–66. https://doi.org/10.1109/TSMC.1979.4310076 (1979).
22. Ridler, T. W. & Calvard, S. Picture thresholding using an iterative selection method. *IEEE Trans. Syst. Man. Cybern.* **8**, 630–632. https://doi.org/10.1109/TSMC.1978.4310039 (1978).
23. Sezan, M. I. A peak detection algorithm and its application to histogram-based image data reduction. *Comput. Vis. Graph. Image Process.* **49**, 36–51. https://doi.org/10.1016/0734-189X(90)90161-N (1990).
24. Vicar, T. et al. Cell segmentation methods for label-free contrast microscopy: Review and comprehensive comparison. *BMC Bioinform.* **20**, 360. https://doi.org/10.1186/s12859-019-2880-8 (2019).
25. Ryu, J. et al. Multimodal microscopy for the simultaneous visualization of five different imaging modalities using a single light source. *Biomed. Opt. Express* **12**, 5452–5469. https://doi.org/10.1364/BOE.430677 (2021).
26. Yin, Z., Kanade, T. & Chen, M. Understanding the phase contrast optics to restore artifact-free microscopy images for segmentation. *Med. Image Anal.* **16**, 1047–1062. https://doi.org/10.1016/j.media.2011.12.006 (2012).

27. Kim, J. H. et al. Propulsion and navigation within the advancing monolayer sheet. *Nat. Mater.* **12**, 856–863. https://doi.org/10.1038/nmat3689 (2013).
28. Shin, Y. et al. Emulating endothelial dysfunction by implementing an early atherosclerotic microenvironment within a microfluidic chip. *Lab. Chip* **19**, 3664–3677. https://doi.org/10.1039/c9lc00352e (2019).
29. Sailem, H., Bousgouni, V., Cooper, S. & Bakal, C. Cross-talk between Rho and Rac GTPases drives deterministic exploration of cellular shape space and morphological heterogeneity. *Open Biol.* **4**, 130132. https://doi.org/10.1098/rsob.130132 (2014).
30. Letort, V. et al. Quantitative analysis of melanocyte migration in vitro based on automated cell tracking under phase contrast microscopy considering the combined influence of cell division and cell-matrix interactions. *Math. Model. Nat. Phenom.* **5**, 4–33. https://doi.org/10.1051/mmnp/20105101 (2010).
31. Adanja, I. et al. Automated tracking of unmarked cells migrating in three-dimensional matrices applied to anti-cancer drug screening. *Exp. Cell. Res.* **316**, 181–193. https://doi.org/10.1016/j.yexcr.2009.10.004 (2010). https://doi.org/https://doi.
32. Zimmer, C., Labruyere, E., Meas-Yedid, V., Guillen, N. & Olivo-Marin, J. C. Segmentation and tracking of migrating cells in videomicroscopy with parametric active contours: a tool for cell-based drug testing. *IEEE Trans. Med. Imaging* **21**, 1212–1221. https://doi.org/10.1109/TMI.2002.806292 (2002).
33. Meng, N., Lam, E. Y., Tsia, K. K. & So, H. K. H. Large-scale multi-class image-based cell classification with deep learning. *IEEE J. Biomed. Health Inf.* **23**, 2091–2098. https://doi.org/10.1109/JBHI.2018.2878878 (2019).
34. He, Y. R. et al. Cell cycle stage classification using phase imaging with computational specificity. *ACS Photonics* **9**, 1264–1273. https://doi.org/10.1021/acsphotonics.1c01779 (2022).
35. Hillsley, A., Santos, J. E. & Rosales, A. M. A deep learning approach to identify and segment alpha-smooth muscle actin stress fiber positive cells. *Sci. Rep.* **11**, 21855. https://doi.org/10.1038/s41598-021-01304-4 (2021).
36. Kesapragada, M. et al. A data-driven approach to establishing cell motility patterns as predictors of macrophage subtypes and their relation to cell morphology. *PloS One* **19**, e0315023. https://doi.org/10.1371/journal.pone.0315023 (2025).
37. Garcia-Moreno, F. M., Ruiz-Espigares, J., Gutiérrez-Naranjo, M. A. & Marchal, J. A. Using deep learning for predicting the dynamic evolution of breast cancer migration. *Comput. Biol. Med.* **180**, 108890. https://doi.org/10.1016/j.compbiomed.2024.108890 (2024).
38. Elbez, R. et al. Cell-morphodynamic phenotype classification with application to cancer metastasis using cell magnetorotation and machine-learning. *PloS One* **16**, e0259462. https://doi.org/10.1371/journal.pone.0259462 (2021).
39. Alizadeh, E. et al. Cellular morphological features are predictive markers of cancer cell state. *Comput. Biol. Med.* **126**, 104044. https://doi.org/10.1016/j.compbiomed.2020.104044 (2020).
40. Huang, K. et al. Image-based profiling and deep learning reveal morphological heterogeneity of colorectal cancer organoids. *Comput. Biol. Med.* **173**, 108322. https://doi.org/10.1016/j.compbiomed.2024.108322 (2024).
41. Kang, M., Min, C., Devarasou, S. & Shin, J. H. Classification of differentially activated groups of fibroblasts using morphodynamic and motile features. *APL Bioeng.* **9**, 026116. https://doi.org/10.1063/5.0250502 (2025).
42. Dilokthanakul, N. et al. Deep unsupervised clustering with Gaussian mixture variational autoencoders. *arXiv*. https://arxiv.org/abs/1611.02648 (2016).
43. Yue, Z. et al. TS2Vec: Towards universal representation of time series. in *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 36. 8980–8987. https://doi.org/10.1609/aaai.v36i8.20881 (2022).
44. Jiang, Z., Zheng, Y., Tan, H., Tang, B. & Zhou, H. Variational deep embedding: An unsupervised and generative approach to clustering. in *Proceedings of the Twenty Sixth International Joint Conference on Artificial Intelligence*, 1965–1972. https://doi.org/10.24963/ijcai.2017/273 (2017).
45. Xie, J., Girshick, R. & Farhadi, A. Unsupervised deep embedding for clustering analysis. in *Proceedings of The 33rd International Conference on Machine Learning*, Vol. **48**. 478–487 (2016).
46. Zbontar, J., Jing, L., Misra, I., LeCun, Y. & Deny, S. Barlow Twins: Self-supervised learning via redundancy reduction. in *Proceedings of the 38th International Conference on Machine Learning*, Vol. 139. 12310–12320 (2021).
47. Yue, B. Biology of the extracellular matrix: An overview. *J. Glaucoma* **23**, 20–S23. https://doi.org/10.1097/ijg.0000000000000108 (2014).
48. Kiosses, W. B., Hahn, K. M., Giannelli, G. & Quaranta, V. Characterization of morphological and cytoskeletal changes in MCF10A breast epithelial cells plated on Laminin-5: Comparison with breast cancer cell line MCF7. *Cell. Communi. Adhes.* **8**, 29–44. https://doi.org/10.3109/15419060109080705 (2001).
49. Debnath, J., Muthuswamy, S. K. & Brugge, J. S. Morphogenesis and oncogenesis of MCF-10A mammary epithelial acini grown in three-dimensional basement membrane cultures. *Methods* **30**, 256–268. https://doi.org/10.1016/S1046-2023(03)00032-X (2003).
50. Seetharaman, S. & Etienne-Manneville, S. Cytoskeletal crosstalk in cell migration. *Trends Cell Biol.* **30**, 720–735. https://doi.org/10.1016/j.tcb.2020.06.004 (2020).
51. Dosovitskiy, A. et al. An image is worth 16 × 16 words: Transformers for image recognition at scale. *arXiv* https://doi.org/10.48550/arXiv.2010.11929 (2020).
52. Kirillov, A. et al. Segment anything. *arXiv* https://doi.org/10.48550/arXiv.2304.02643 (2023).
53. Pape, C., Wolny, A., Krull, A., Mueller, D. & Kainmueller, D. Segment anything for microscopy. *Nat. Methods* **22**, 579–591. https://doi.org/10.1038/s41592-024-02580-4 (2025).
54. Pachitariu, M., Rariden, M. & Stringer, C. Cellpose-SAM: Superhuman generalization for cellular segmentation. *bioRxiv* https://doi.org/10.1101/2025.04.28.651001 (2025).
55. Marks, M. et al. CellSAM: A foundation model for cell segmentation. *Nat. Methods* **22**, 2585–2593. https://doi.org/10.1038/s41592-025-02879-w (2025).
56. Carion, N. et al. End-to-end object detection with transformers. *arXiv* https://doi.org/10.48550/arXiv.2005.12872 (2020).
57. Ravi, N. et al. SAM 2: Segment anything in images and videos. *arXiv* https://doi.org/10.48550/arXiv.2408.00714 (2024).
58. Zhang, H. et al. DINO: DETR with improved DeNoising anchor boxes for end-to-end object detection. *arXiv* https://doi.org/10.48550/arXiv.2203.03605 (2022).
59. Lv, W. et al. DETRs beat YOLOs on real-time object detection. *arXiv*. https://doi.org/10.48550/arXiv.2304.08069 (2023).
60. Noel, A. et al. Antagonistic effects of laminin and fibronectin in cell-to-cell and cell-to-matrix interactions in MCF-7 cultures. *Vitro Cell. Dev. Biol* **24**, 373–380. https://doi.org/10.1007/BF02628488 (1988).
61. Rabinovitz, I. & Mercurio, A. M. The integrin α6β4 functions in carcinoma cell migration on Laminin-1 by mediating the formation and stabilization of actin-containing motility structures. *J. Cell. Biol.* **139**, 1873–1884. https://doi.org/10.1083/jcb.139.7.1873 (1997).
62. Fullár, A. et al. Remodeling of extracellular matrix by normal and tumor-associated fibroblasts promotes cervical cancer progression. *BMC Cancer* **15**, 256. https://doi.org/10.1186/s12885-015-1272-3 (2015).
63. Pouliot, N. & Kusuma, N. Laminin-511. *Cell. Adhes. Migr.* **7**, 142–149. https://doi.org/10.4161/cam.22125 (2013).
64. Qiu, X., Tan, H., Fu, D., Zhu, Y. & Zhang, J. Laminin is over expressed in breast cancer and facilitate cancer cell metastasis. *J. Cancer Res. Ther.* **14**, S1170–S1172. https://doi.org/10.4103/0973-1482.191035 (2018).
65. Terranova, V. P., Llotta, L. A., Russo, R. G. & Martin, G. R. Role of laminin in the attachment and metastasis of murine tumor cells. *Cancer Res.* **42**, 2265–2269 (1982).
66. Huang, C. & Chen, J. Laminin–332 mediates proliferation, apoptosis, invasion, migration and epithelial–to–mesenchymal transition in pancreatic ductal adenocarcinoma. *Mol. Med. Rep.* https://doi.org/10.3892/mmr.2020.11649 (2021).
67. Leitinger, B. & Hohenester, E. Mammalian collagen receptors. *Matrix Biol.* **26**, 146–155. https://doi.org/10.1016/j.matbio.2006.10.007 (2007).

68. Xu, Y. et al. Multiple binding sites in collagen type I for the integrins alpha1beta1 and alpha2beta1. *J. Biol. Chem.* **275**, 38981–38989. https://doi.org/10.1074/jbc.M007668200 (2000).
69. Eckes, B. et al. Mechanical tension and integrin alpha 2 beta 1 regulate fibroblast functions. *J. Investig. Dermatol. Symp. Proc.* **11**, 66–72. https://doi.org/10.1038/sj.jidsymp.5650003 (2006).
70. Chan, B. M. et al. Distinct cellular functions mediated by different VLA integrin alpha subunit cytoplasmic domains. *Cell* **68**, 1051–1060. https://doi.org/10.1016/0092-8674(92)90077-p (1992).
71. Webb, D. J., Parsons, J. T. & Horwitz, A. F. Adhesion assembly, disassembly and turnover in migrating cells – over and over and over again. *Nat. Cell Biol.* **4**, E97–E100. https://doi.org/10.1038/ncb0402-e97 (2002).
72. Broussard, J. A., Webb, D. J. & Kaverina, I. Asymmetric focal adhesion disassembly in motile cells. *Curr. Opin. Cell. Biol.* **20**, 85–90. https://doi.org/10.1016/j.ceb.2007.10.009 (2008).
73. Kim, D. H. & Wirtz, D. Focal adhesion size uniquely predicts cell migration. *Faseb J.* **27**, 1351–1361. https://doi.org/10.1096/fj.12-220160 (2013).
74. Nagano, M., Hoshino, D., Koshikawa, N., Akizawa, T. & Seiki, M. Turnover of focal adhesions and cancer cell migration. *Int. J. Cell. Biol.*. https://doi.org/10.1155/2012/310616 (2012).
75. Meehan, S. & Nain, A. S. Role of suspended fiber structural stiffness and curvature on single-cell migration, nucleus shape, and focal-adhesion-cluster length. *Biophys. J.* **107**, 2604–2611. https://doi.org/10.1016/j.bpj.2014.09.045 (2014).
76. Fokkelman, M. et al. Cellular adhesome screen identifies critical modulators of focal adhesion dynamics, cellular traction forces and cell migration behaviour. *Sci. Rep.* **6**, 31707. https://doi.org/10.1038/srep31707 (2016).
77. Horzum, U., Ozdil, B. & Pesen-Okvur, D. Step-by-step quantitative analysis of focal adhesions. *MethodsX* **1**, 56–59. https://doi.org/10.1016/j.mex.2014.06.004 (2014).
78. Wang, C. Y., Bochkovskiy, A. & Liao, H. Y. M. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* 7464–7475. https://doi.org/10.1109/CVPR52729.2023.00721 (2023).
79. Stringer, C., Wang, T., Michaelos, M. & Pachitariu, M. Cellpose: A generalist algorithm for cellular segmentation. *Nat. Methods* **18**, 100–106. https://doi.org/10.1038/s41592-020-01018-x (2021).
80. Pachitariu, M. & Stringer, C. Cellpose 2.0: how to train your own model. *Nat. Methods* **19**, 1634–1641. https://doi.org/10.1038/s41592-022-01663-4 (2022).

Acknowledgements

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (RS-2022-NR070365), by a grant of the Korea-US Collaborative Research Fund (KUCRF), funded by the Ministry of Science and ICT and Ministry of Health & Welfare, Republic of Korea (RS-2024-00468873), and by Sejong University.

Author contributions

Conceptualization: J.H and B.G; Data Curation: E.S, J.H, and B.G; Investigation: E.S, J.H, A.J, and J.K; Methodology: E.S, J.H, and A.J; Validation: J.H, I.Y.W, and B.G; Writing – original draft: E.S., J.H, and B.G; Writing – review & editing: J.H, I.Y.W, and B.G; Funding acquisition: I.Y.W, and B.G; Supervision; B.G; All authors critically reviewed the report and approved the final version.

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-026-42530-y.

Correspondence and requests for materials should be addressed to B.G.

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

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