Classification of differentially activated groups of fibroblasts using morphodynamic and motile features

Cite as: APL Bioeng. **9**, 026116 (2025); doi: 10.1063/5.0250502
Submitted: 25 November 2024 · Accepted: 5 May 2025 ·
Published Online: 15 May 2025

View Online Export Citation CrossMark

Minwoo Kang,[1] Chanhong Min,[1] Somayadineshraj Devarasou,[1] and Jennifer H. Shin[1,2,3,a)]

AFFILIATIONS

[1]Department of Mechanical Engineering, KAIST, 291 Daehak-Ro, Yuseong-Gu, Daejeon 34141, Republic of Korea
[2]Graduate School of Stem Cell and Regenerative Biology, KAIST, 291 Daehak-Ro, Yuseong-Gu, Daejeon 34141, Republic of Korea
[3]KI for Health Science and Technology, KAIST, 291 Daehak-Ro, Yuseong-Gu, Daejeon 34141, Republic of Korea

[a)]**Author to whom correspondence should be addressed:** j_shin@kaist.ac.kr. **Tel.:** +82 42 350 3232

ABSTRACT

Fibroblasts play essential roles in cancer progression, exhibiting activation states that can either promote or inhibit tumor growth. Understanding these differential activation states is critical for targeting the tumor microenvironment (TME) in cancer therapy. However, traditional molecular markers used to identify cancer-associated fibroblasts are limited by their co-expression across multiple fibroblast subtypes, making it difficult to distinguish specific activation states. Morphological and motility characteristics of fibroblasts reflect their underlying gene expression patterns and activation states, making these features valuable descriptors of fibroblast behavior. This study proposes an artificial intelligence-based classification framework to identify and characterize differentially activated fibroblasts by analyzing their morphodynamic and motile features. We extract these features from label-free live-cell imaging data of fibroblasts co-cultured with breast cancer cell lines using deep learning and machine learning algorithms. Our findings show that morphodynamic and motile features offer robust insights into fibroblast activation states, complementing molecular markers and overcoming their limitations. This biophysical state-based cellular classification framework provides a novel, comprehensive approach for characterizing fibroblast activation, with significant potential for advancing our understanding of the TME and informing targeted cancer therapies.

INTRODUCTION

Fibroblasts play key roles in the tumor microenvironment (TME) and can be activated by cancer cells. The activated fibroblasts by cancer cells, cancer-associated fibroblasts (CAFs), have risen to prominence as key players in cancer progression.[1] These cells synthesize and remodel the extracellular matrix (ECM), producing various growth factors. These functions modulate cancer metastasis and aid epithelial–mesenchymal transition (EMT). In addition, CAFs are involved in angiogenesis, modification of tumor metabolism, modulation of the immune system, and even drug access and therapy responses.[2,3] Accordingly, CAFs started to be recognized as reliable anti-cancer targets, so-called anti-CAF therapies.[2]

Researchers initially presumed that CAFs were homogeneous and accelerated tumor progression. However, current evidence clearly shows that CAFs form heterogeneous populations, with some promoting tumor growth and others exerting anti-tumorigenic effects. A growing body of research continues to highlight the complex nature of these cells.[4–9] Researchers have long sought to identify CAF subtypes through molecular markers. For instance, α-smooth muscle actin (α-SMA) is commonly used to identify CAFs, and other markers such as fibroblast activation protein (FAP) and fibroblast-specific protein 1 (FSP1) have also been characterized. However, these markers are often co-expressed across different subtypes, complicating efforts to define distinct subpopulations. To effectively target CAFs for cancer treatment, it is essential to classify heterogeneous subtypes so that pro-tumorigenic ones can be suppressed or reprogrammed into anti-tumorigenic states.[1]

Single-cell RNA sequencing (scRNA-Seq) has become a common technique for classifying CAF subpopulations. However, the limitations of molecular markers, which are shared among subtypes, make it challenging to define distinct groups based solely on molecular data.[10,11] To better delineate CAF heterogeneity, Bartoschek *et al.*

introduced the use of biophysical properties such as cell size and granularity in combination with scRNA-Seq data.[11] Similarly, Pelon *et al.* used scRNA-Seq to identify four CAF subtypes, labeled CAF-S1 to CAF-S4.[5] They focused on two subtypes (CAF-S1 and CAF-S4), which shared myofibroblastic features but exhibited notable functional differences related to tumor progression and cellular traction forces. Although this was not explicitly addressed in their study, these two subtypes displayed distinct morphological features: CAF-S1 cells had a smaller spreading area and larger aspect ratio. In comparison, CAF-S4 cells showed a larger spreading area and smaller aspect ratio. This suggests that incorporating detailed morphological features beyond simple geometric parameters could improve the identification of heterogeneous fibroblast subtypes.

The motility of fibroblasts, like their morphology, also reflects their heterogeneous nature. Costea *et al.* conducted a study on oral squamous cell carcinoma and identified two distinct CAF subtypes with different motility characteristics and effects on cancer cells.[12] The CAF-N subtype, which had a transcriptomic profile similar to normal oral fibroblasts, displayed higher motility than the CAF-D subtype, which had a more divergent transcriptomic profile. These findings highlight the value of combining morphological and motile features with biomolecular markers to provide a more comprehensive characterization of fibroblast activation states. The morphodynamic and motile characteristics of fibroblasts result from complex gene expression patterns, making them valuable holistic descriptors.

Unlike fixed-cell biomolecular analyses, which offer only end point data, live-cell imaging enables the observation of dynamic changes in fibroblasts over time. This temporal information can provide insights into phenotypic transitions and activation states.[13–17] Wang *et al.* demonstrated this benefit by showing that live-cell imaging captured dynamic phenotypic transitions in epithelial cells that static snapshot data failed to reveal.[14]

In this study, we aim to explore the potential of morphodynamic and motile features to complement biomolecular markers in providing a comprehensive understanding of fibroblast activation. To achieve this, we propose an artificial intelligence (AI)-based classification framework specifically designed to classify differentially activated fibroblasts based on morphodynamic and motile features. We extract these features from label-free live-cell imaging data using deep learning and machine learning algorithms. To test our hypothesis, we employed fibroblasts co-cultured with two breast cancer cell lines of varying aggressiveness—MCF-7 (luminal A, lower aggressiveness) and MDA-MB-231 (triple-negative, highly metastatic)—because co-culture provides a reproducible model system where cancer cell–secreted factors induce consistent fibroblast activation patterns within a short period. By maintaining uniform growth conditions and a controlled tumor–stromal interface, this method allows reliable comparisons of fibroblast responses to different cancer cell types, making it well-suited for evaluating the accuracy of our proposed classification framework. Our results show that morphodynamic and motile features effectively distinguish the activation states of fibroblasts, complementing traditional biomolecular markers.

RESULTS

Validation of *in vitro* models of activated fibroblasts with CAF markers

To investigate the activation of fibroblasts, we co-cultured normal human dermal fibroblasts (NHDF) with two breast cancer cell lines: MCF-7 (luminal A, lower aggressiveness) and MDA-MB-231 (triple-negative, highly metastatic) [Fig. 1(a)]. Previous studies have shown that fibroblasts co-cultured with cancer cells *in vitro* become activated and exhibit interactions with cancer cells similar to those of CAFs.[18–20] By using *in vitro* models of activated fibroblasts, we consistently obtained cells with similar activation patterns across experiments, enabling reliable analysis and supporting the verification of our new strategy for studying CAF biology. We analyzed these co-cultured fibroblasts—referred to as NHDF activated by MCF-7 and NHDF activated by MDA-MB-231—to assess their activation status. While these fibroblasts exhibit traits similar to CAFs, they do not fully represent *in vivo* CAFs. We verified their activation status by evaluating the expression of common CAF markers using qPCR (quantitative polymerase chain reaction) [Fig. 1(b), supplementary material Fig. 1 and Table 1].

Our data suggest that fibroblast activation follows distinct trajectories, even within the five-day co-culture period. Rather than a binary switch, fibroblast activation is a progressive process modulated by the tumor microenvironment.[21–23] The short (five-day) co-culture in our experiments likely captures an early phase of this transition. Fibroblasts co-cultured with MDA-MB-231 cells exhibit a pronounced CAF-like phenotype, with significantly lower PDGFRβ (platelet-derived growth factor receptor beta) expression compared to both the control and MCF-7 groups. In contrast, MCF-7–activated fibroblasts display PDGFRβ levels similar to those of the control, suggesting a weaker or delayed activation response at the bulk population level. Since qPCR averages gene expression across the entire sample, it may obscure subpopulations of strongly activated fibroblasts that contribute to early stromal remodeling. Nonetheless, our findings indicate that the highly aggressive, triple-negative MDA-MB-231 cells induce a more consistent and pronounced stromal response, likely enriching for FAP+PDPN+ fibroblasts—a phenotype commonly linked to immunosuppressive functions and ECM remodeling in triple-negative breast cancer (TNBC).[21]

Meanwhile, fibroblasts in the control group exhibited broader variability in PDGFRβ, PDPN, and FAP expression, consistent with a more heterogeneous population in the absence of cancer cell influence. This baseline heterogeneity aligns with previous studies,[24,25] showing that fibroblast populations are inherently diverse but become more uniform following tumor-induced activation. Exposure to MDA-MB-231 cells appears to drive a more homogeneous shift toward a CAF-like phenotype, whereas MCF-7–induced activation remains subtler and more variable. Together, these results suggest that fibroblasts undergo distinct activation patterns depending on the cancer cell type they interact with. Given these molecular trends, we next sought to assess how these differential activation states manifest at the single-cell level by analyzing biophysical properties, including morphology and motility features.

Morphodynamic features for fibroblast classification

Live-cell imaging of single fibroblasts provides detailed spatio-temporal information about cell behavior. Although the imaging spanned 24 h, we analyzed only the first 12 h of image stacks to avoid complications from active cell proliferation occurring later. From four independent live-cell imaging replicates (n = 4), each conducted under consistent passage number, seeding densities, and media conditions, we captured approximately 70 000 morphological states of fibroblasts.

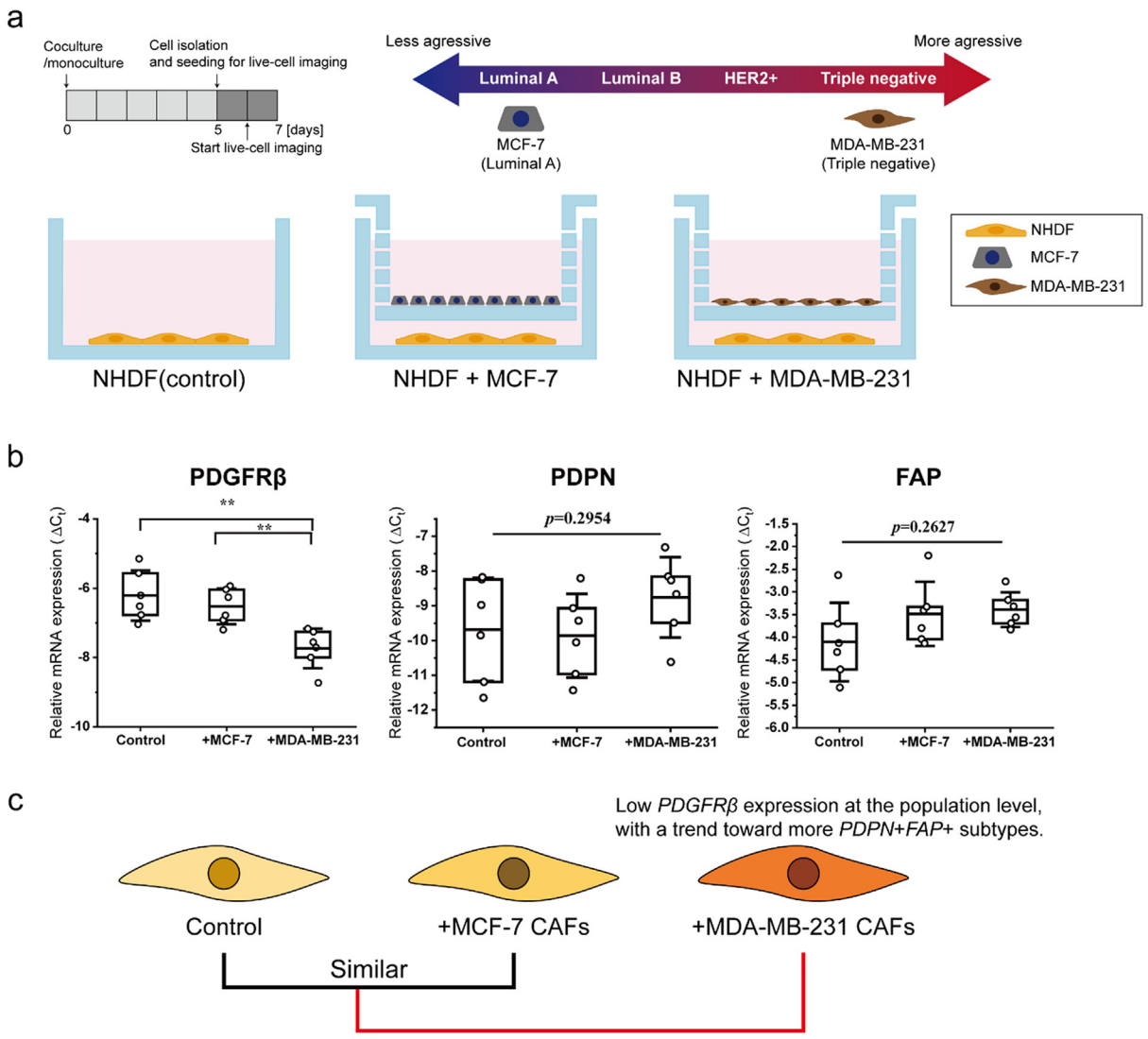

FIG. 1. Establishing *in vitro* models of activated fibroblasts and validation with CAF markers. (a) We induced *in vitro* models of activated fibroblasts by co-culturing them with breast cancer cell lines of varying aggressiveness. (b) Relative mRNA expressions of CAF markers of normal fibroblasts and *in vitro* models of activated fibroblasts. (c) Assumed similarity between control and *in vitro* models of activated fibroblasts based on the biomolecular assay. Control, +MCF-7, and +MDA-MB231 stand for normal fibroblasts, fibroblasts co-cultured with MCF-7, and fibroblasts co-cultured with MDA-MB-231, respectively. n = 6, and data represent mean ± SD.

To assess potential batch effects, we performed principal component analysis (PCA) (supplementary material Fig. 2), which confirmed that replicates within each condition exhibited overlapping distributions in PCA space, indicating minimal batch-to-batch variability. Based on these results, we pooled the data to enhance statistical power and provide a more comprehensive classification of fibroblast morphodynamics. Figure 2(a) shows representative images of fibroblasts, including normal fibroblasts and those differentially activated by two distinct cancer cell types. As demonstrated, distinguishing these cells based solely on visual inspection was not feasible. Therefore, we applied specific morphological descriptors to analyze the morphodynamic states of the cells.

In addition to observing basic cell morphology, the phase-contrast images provided texture information based on variations in light intensity. Also, Zernike moments have been used in biomedical imaging applications as effective features because of their orthogonal characteristics, noise robustness, and their rotation and scale invariance.[26–31] To capture the complexity of these morphological characteristics, we extracted 52 features from three morphological feature groups, including basic morphology, texture features, and Zernike moments (supplementary material Fig. 3; see the Methods section for details). Although each individual morphological feature (supplementary material Fig. 4) can differentiate certain aspects of control vs activated fibroblasts, focusing on a single dimension still

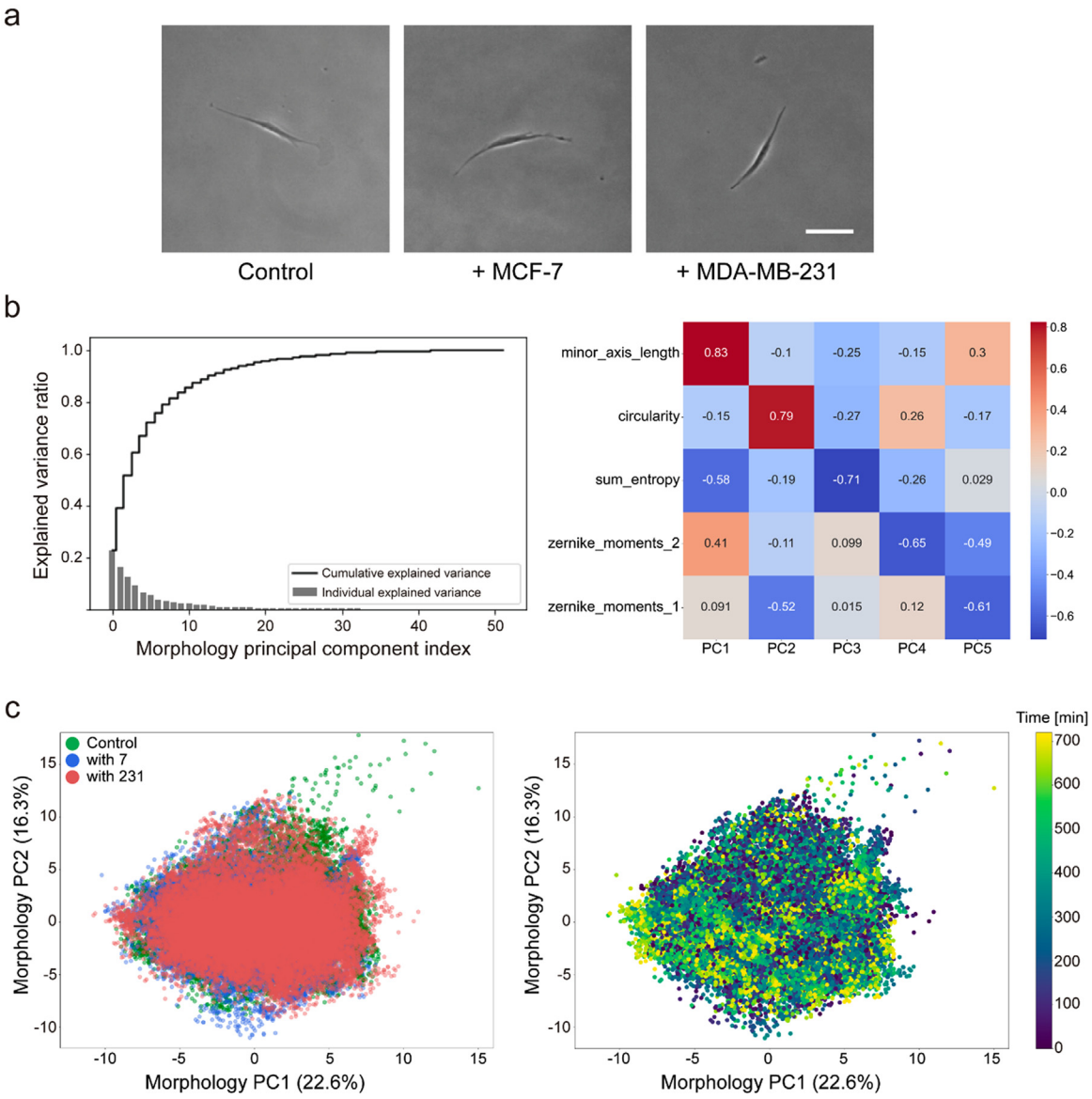

FIG. 2. Embedding of morphology state of cells onto 2D morphology state space. (a) Random snapshots of normal fibroblasts and *in vitro* models of activated fibroblasts from live-cell imaging. Control, +MCF-7, and +MDA-MB-231 stand for normal fibroblasts, fibroblasts co-cultured with MCF-7, and fibroblasts co-cultured with MDA-MB-231, respectively. Scalebar: 100 μm. (b) Dimensionality reduction of morphological features with PCA. (Left panel) The individual explained the variance of PCs and the cumulative explained variance. (Right panel) Top 5 PCs and corresponding variables. (c) Embedding of morphology state of cells onto 2D morphology state space. (Left panel) Morphological states of normal fibroblast (control), +MCF-7 CAF (with7), and +MDA-MB-231 CAF (with231). (Right panel) Temporal fluctuation of morphological states.

overlooks the broader morphological context. For example, increased circularity or aspect ratio may reflect partial activation in some fibroblasts, yet other cells within the same group may present different—sometimes even opposing—trends in other shape descriptors. Therefore, to fully characterize the interplay of all shape metrics and capture coordinated morphological changes across the entire population, we integrated all 52 features for each cell into a high-dimensional descriptor set. This multivariate approach enables us to explore subtle but consistent patterns of fibroblast activation—patterns that might be missed or appear less robust when examining each feature independently. Consequently, we performed principal component analysis (PCA) to condense these features into a lower-dimensional space that can highlight the combined morphological shifts underlying CAF-like phenotypes. The first two components (PC1 and PC2) explained 22.6% and 16.3% of the variance, and we plotted the top 5 PCs and the corresponding highest correlated variables [Fig. 2(b)]. The right panel of Fig. 2(b) shows that the first two principal components (PC1 and PC2) are most strongly associated with minor axis length and

circularity, both of which are part of the basic morphology group. In contrast, features from the other two groups—Haralick texture (e.g., sum entropy) and Zernike moments (e.g., zernike moments1 and 2)—contribute more modestly to PCs 3–5. This suggests that basic morphological descriptors such as cell elongation and roundness capture the dominant variance in fibroblast shape, whereas higher-order features contribute to subtler aspects of morphological variation.

Next, we embedded the morphological states of all fibroblasts into morphology PC1–PC2 state space [Fig. 2(c)]. Even after dimensionality reduction via PCA, we observed substantial overlap among normal and differentially activated fibroblasts in the morphology PC1–PC2 space [Fig. 2(c), left panel]. This finding highlights the partial and heterogeneous nature of fibroblast activation within our co-culture system, where cells exhibit varying degrees of morphological shifts rather than transitioning into a single distinct phenotype. Additionally, we observed temporal fluctuations in morphology, but the graph did not reveal a clear or consistent pattern of change over time [Fig. 2(c) right panel].

To quantify the heterogeneity of the morphological states, we applied k-means clustering to define seven distinct morphological states [Fig. 3(a)]. We then visualized the distribution of these states across the samples using a heatmap [Fig. 3(b)]. Interestingly, the control group and the NHDF activated by MCF-7 group exhibited similar patterns, as confirmed by unsupervised hierarchical clustering based on Euclidean distance.

We further quantified the morphological heterogeneity of each group using Shannon entropy, with lower values indicating a more homogeneous phenotype [Fig. 3(c)]. The NHDF activated by the MDA-MB-231 group exhibited the lowest Shannon entropy, indicating that it was the most homogeneous group. This analysis demonstrated that the defined morphology states, along with the extracted morphological features, successfully distinguished the phenotypic characteristics of the differentially activated fibroblasts.

As mentioned, we captured the temporal fluctuations of the morphological features [Fig. 2(c) right panel]. To further investigate how fibroblasts transition among different morphological states over time, we quantified the temporal distribution of states for each group [Fig. 4(a)]. While some fibroblasts exhibited transient fluctuations, the limited 12-h observation period made it difficult to determine whether these changes represented true state transitions or minor stochastic variations in cell shape. To incorporate the temporal component more rigorously, we applied time-series k-means clustering with dynamic time warping (DTW), enabling the identification of three distinct morphodynamic states [Fig. 4(b)].

Similar to our earlier visualization of morphological states, we plotted the distribution of these morphodynamic states in a heatmap [Fig. 4(c)]. As expected, the control group and NHDFs activated by MCF-7 displayed similar distribution patterns, reinforcing our earlier findings from morphological analysis. Unsupervised hierarchical clustering based on Euclidean distance further confirmed that NHDFs activated by MDA-MB-231 were more distinct, supporting the idea that morphodynamic states capture biologically relevant differences in fibroblast activation. Nonetheless, the heterogeneity of the NHDFs activated by MDA-MB-231 and the control groups appeared similar, limiting our ability to clearly distinguish the most homogeneous MDA-MB-231-activated fibroblasts [Fig. 4(d)]. Within the scope of our dataset, these findings suggest that morphological descriptors at individual time points provide sufficient discriminatory power for classifying activated fibroblasts in this short timeframe, whereas transient fluctuations do not necessarily consolidate into additional discrete clusters over the 12-h imaging period. However, longer-term or more frequent sampling could better clarify whether these ephemeral fluctuations reflect true intermediate activation states or merely short-term cytoskeletal rearrangements. Future studies extending the observation period beyond 12 h and incorporating real-time molecular markers may help determine whether such transient morphodynamic changes correspond to functional shifts in fibroblast activation.

Motility features for fibroblast classification

As observed in the previous section, each cell exhibited distinct morphological characteristics, and the same was valid for their motility. In line with employing multiple features to quantify cellular morphology, we sought to comprehensively capture the diverse motile characteristics of fibroblasts. To achieve this, we chose 50 motile features outlined in the Methods section, encompassing general, mean square displacement (MSD), and random walk features. We subsequently extracted these features from each individual fibroblast (supplementary material Fig. 5). We excluded cells that migrated out of the

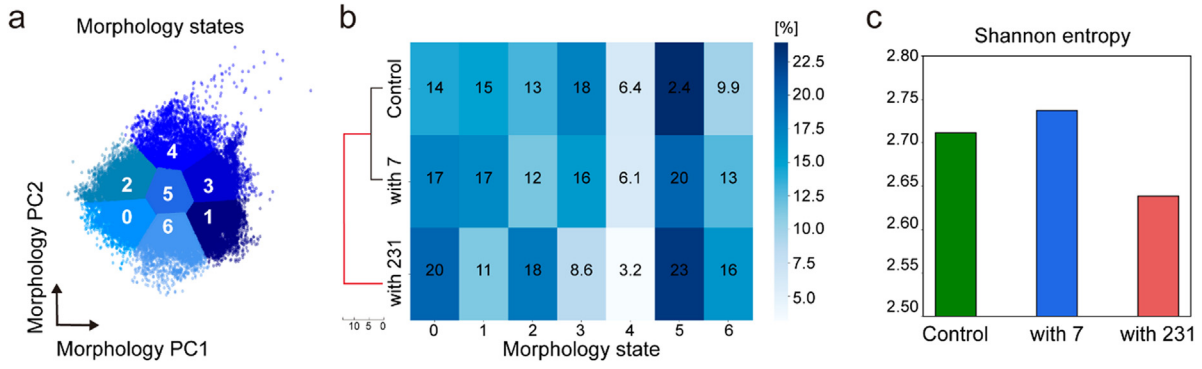

FIG. 3. Defined morphology states in a 2D morphology state space and the distribution of morphology states of each group. (a) 7 defined morphology states identified by unsupervised clustering. (b) Defined morphology state distribution for each group and unsupervised hierarchical clustering of fibroblasts. (c) Morphological heterogeneity within the groups represented by Shannon entropy.

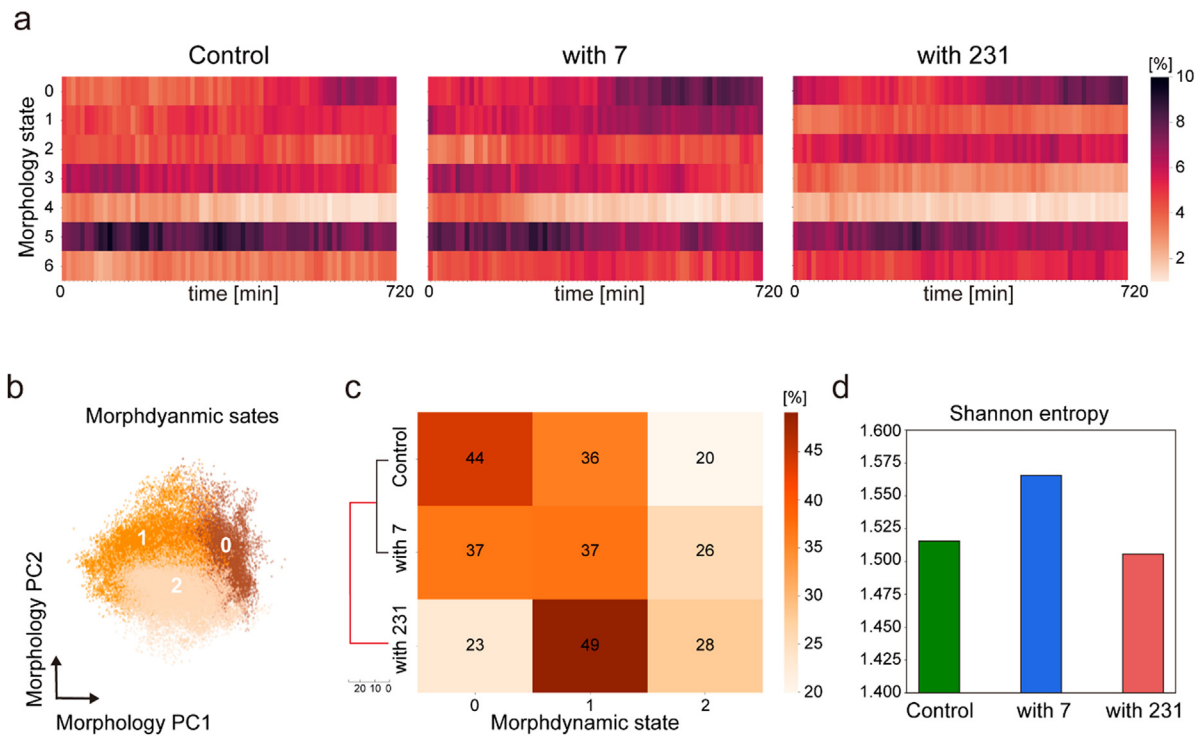

FIG. 4. Defined morphodynamic states in a 2D morphology state space and the distribution of morphodynamic states of each group. (a) Temporal fluctuation of morphology states in each group. (b) Three defined morphodynamic states by time-series k-means clustering. (c) Defined morphodynamic state distribution of each group and unsupervised hierarchical clustering of fibroblasts. (d) Morphodynamical heterogeneity within the groups represented by Shannon entropy.

frame or proliferated during imaging, resulting in a final dataset of approximately 940 fibroblasts.

To handle the high-dimensional motility data, we applied principal component analysis (PCA) for dimensionality reduction. The first two components (PC1 and PC2) accounted for 19.6% and 16.6% of the variance, respectively. We plotted the top five principal components (PCs) along with their most strongly correlated variables [Fig. 5(a)]. From the right panel of Fig. 5(a), we observed a notable correlation where PC1 exhibited a strong association with the turning motion (average theta 2) and PC2 is correlated with speed. This observation implies that the general features among three motile feature groups (general features, MSD feature, and random walk features) contribute to the motile representation of normal fibroblasts and differentially activated fibroblasts.

Next, we embedded the motility states of all fibroblasts into the motility PC1–PC2 state space [Fig. 5(b), left panel]. We applied k-means clustering and identified three distinct motility states that captured the heterogeneity within the fibroblast populations [Fig. 5(b), right panel]. We visualized the distribution of these motility states using a heatmap [Fig. 5(c)]. Consistent with the morphological analysis, the control group and NHDF activated by MCF-7 exhibited similar distribution patterns, as confirmed by unsupervised hierarchical clustering based on Euclidean distance. Additionally, the motility features effectively identified the NHDF activated by MDA-MB-231 as the most homogeneous group, which was further supported by the lowest Shannon entropy values [Fig. 5(d)].

Although both morphological and motility features individually provided valuable insights into the biophysical states of fibroblasts, integrating these two sets of features enhances classification robustness and ensures a more comprehensive characterization of their activation states. Fibroblast activation is a complex and dynamic process, involving both cytoskeletal remodeling and migratory behavior, which may change independently. Some cells may undergo significant morphological alterations without notable shifts in motility, while others may exhibit distinct migration patterns despite maintaining similar overall morphology. By combining morphological and motility features, we can capture a broader spectrum of activation states, ensuring that subtle or transient transitions are not overlooked. Additionally, projecting onto only two principal components inevitably results in information loss. While morphology-only and motility-only classifications produced largely consistent clustering outcomes, the integration of these features provides a more holistic representation of fibroblast activation dynamics. This approach may be particularly relevant in tumor microenvironments, where CAFs exhibit diverse and evolving activation states due to prolonged interactions with cancer cells, stromal factors, and ECM remodeling. Capturing both morphodynamic and motility characteristics allows for a more accurate reflection of fibroblast functional diversity, particularly in heterogeneous tumor–stromal interactions.

Morphodynamics–motility composite descriptors for fibroblast classification

We developed morphodynamics–motility composite descriptors by integrating morphodynamic features with motility features

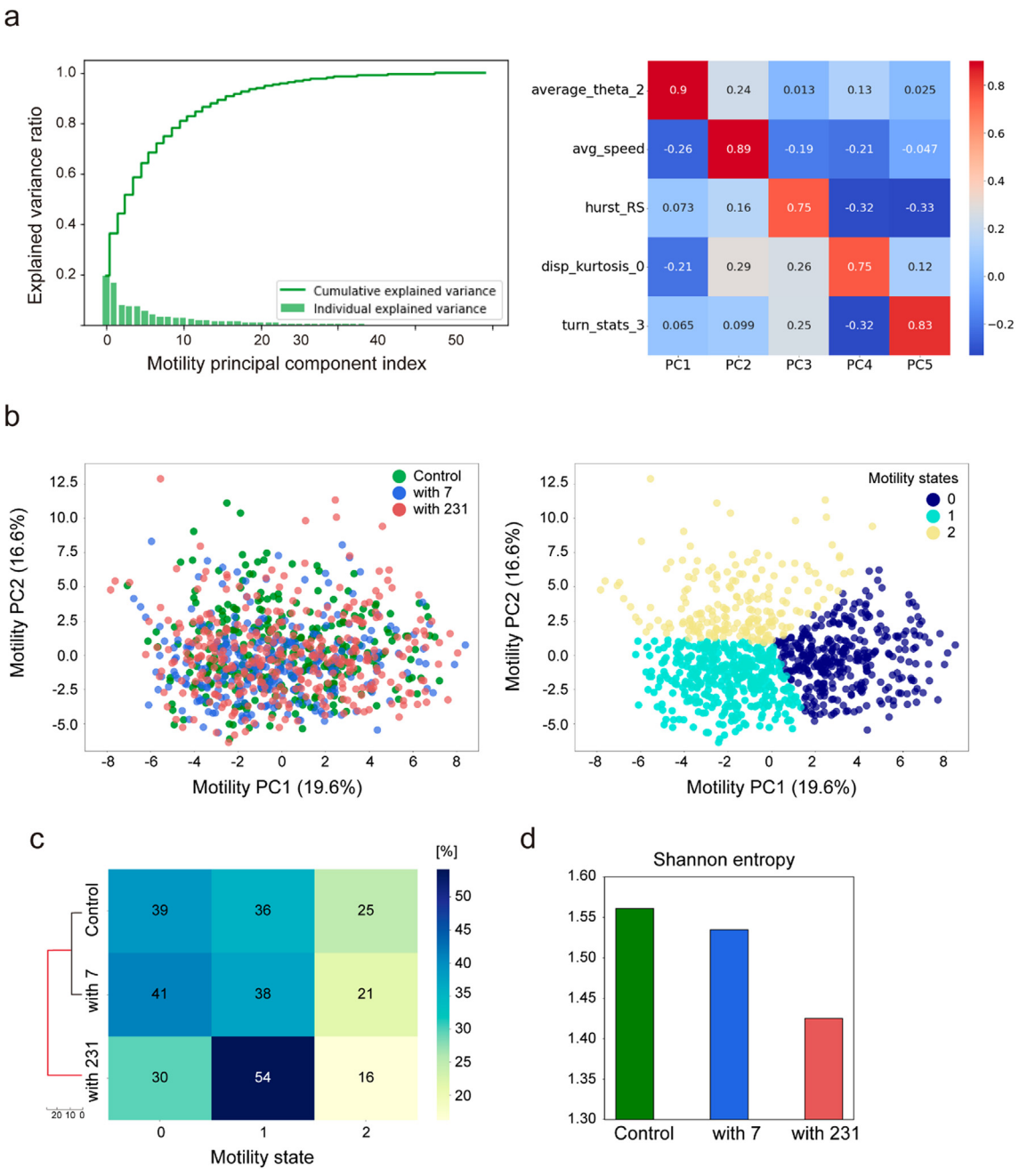

FIG. 5. Embedding of motility state of cells onto 2D motility state space, defining motility states, and the distribution of motility states of each group. (a) Dimensionality reduction of motility features with PCA. (Left panel) The individual variance explained by PCs, and the cumulative explained variance. (Right panel) The top 5 PCs and corresponding variables. (b) Embedding of motility state of cells onto 2D motility state space. (Left panel) Motility states of normal fibroblast (control), +MCF-7 CAF (with 7), and +MDA-MB-231 CAF (with 231). (Right panel) 3 defined motility states by unsupervised clustering. (c) Defined motility state distribution of each group and unsupervised hierarchical clustering of fibroblasts. (d) Motility heterogeneity within the groups represented by Shannon entropy.

[Fig. 6(a)]. To achieve this, we extracted the top principal components (PCs) that explained 95% of the variance in both morphodynamic and motility features, resulting in 20 morphology PCs and 23 motility PCs. To align their dimensionality, we time-averaged the morphology PCs. We then combined the mean values of the morphology and motility PCs into a single feature vector.

To visualize and embed the morphodynamics–motility composite features in a two-dimensional space, we used uniform manifold

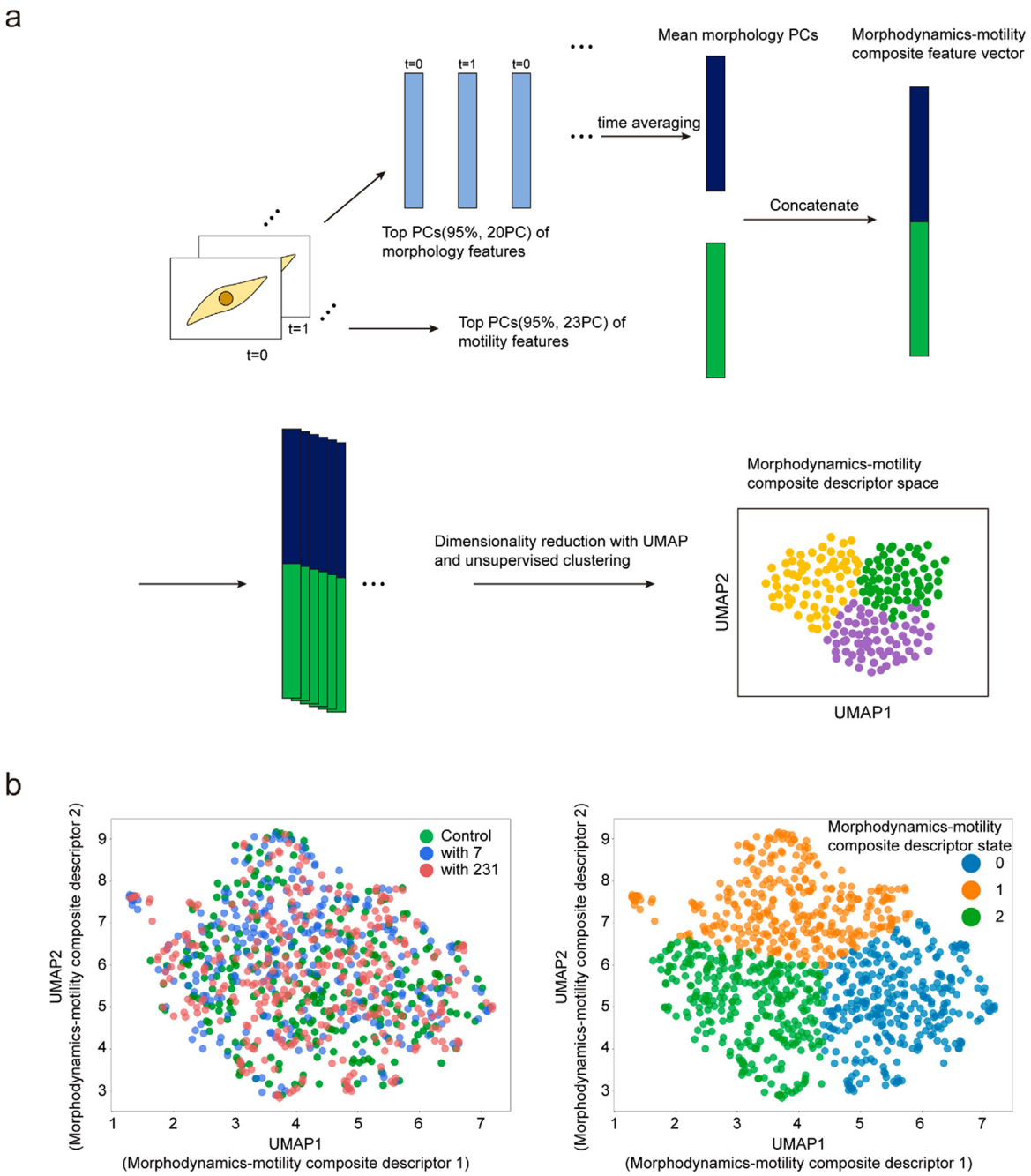

FIG. 6. Establishing morphodynamics–motility composite descriptors and embedding the composite descriptors of cells onto the 2D composite descriptor space. (a) Using machine learning algorithms to establish morphodynamics–motility composite descriptors by incorporating the morphodynamic features with motile features. (b) Embedding of morphodynamics–motility composite descriptor state of cells onto a 2D morphodynamics–motility composite descriptor space. (Left panel) Morphodynamics–motility composite descriptors of normal fibroblast (control), +MCF-7 CAF (with 7), and +MDA-MB-231 CAF (with 231). (Right panel) 3 defined morphodynamics–motility composite descriptor states.

approximation and projection (UMAP) instead of PCA. This decision was based on the fact that the PCs from the composite feature vector exhibited only slight differences in the variance explained (supplementary material Fig. 6). As a result, we determined that using only the first two PCA components would not adequately preserve the information in the data. Instead, we visualized the composite descriptors of individual cells for each group in a two-dimensional UMAP space, which we defined as the morphodynamics–motility composite

descriptor space [Fig. 6(b), left panel]. Next, we applied k-means clustering and identified three distinct morphodynamics–motility composite descriptor states [Fig. 6(b) right panel]. To verify that UMAP effectively preserved both morphodynamic and motility features, we plotted the cells labeled with these composite descriptor states onto the two-dimensional space of various features (mean morphology PC1, PC2, and motility PC1, PC2). The figure shows that the data points clustered consistently based on the composite descriptors across most feature spaces [Fig. 7(a)].

Following this, we visualized the distribution of fibroblasts using a heatmap [Fig. 7(b)]. The composite descriptors successfully captured the similarity between the control group and NHDF activated by MCF-7. The overall heterogeneity of each group, quantified using Shannon entropy, supported these findings. The Shannon entropy values for the control group and NHDF activated by MCF-7 were similar, while the NHDF activated by MDA-MB-231 group exhibited the lowest entropy, indicating that the composite descriptors effectively classified the differentially activated fibroblasts and identified the most homogeneous group [Fig. 7(c)]. This classification aligns with the molecular distinctions observed in our PCR analysis (Fig. 1), reinforcing the validity of these composite descriptors in distinguishing fibroblast activation states (Fig. 8).

DISCUSSION

The heterogeneity of CAFs has long been recognized, with scRNA-seq revealing multiple subpopulations defined by distinct gene expression profiles. However, these molecular classifications provide an incomplete picture of CAF complexity, as they do not directly capture dynamic phenotypic traits, such as cell shape, motility, and ECM interactions.[32] Our study aimed to bridge this gap by classifying differentially activated fibroblasts *in vitro* using biophysical descriptors derived from live-cell imaging.

In this study, we employed a controlled *in vitro* model using normal human dermal fibroblasts (NHDFs) co-cultured with two breast cancer cell lines, MCF-7 and MDA-MB-231, to examine early fibroblast activation events. Given the relatively short co-culture period (5 days), we likely captured an initial phase of fibroblast activation. While the precise functional consequences of these early activation states require further validation, our results align with previous findings that highly aggressive triple-negative breast cancer (TNBC) cells, such as MDA-MB-231, drive strong stromal remodeling and ECM reorganization, potentially accelerating cancer cell invasion.[21] Future studies incorporating longer-term co-cultures, 3D spheroid invasion assays, or *in vivo* xenograft models could help determine whether these early activated fibroblasts sustain pro-tumorigenic or therapy-resistant behaviors over extended timescales.

Our analysis revealed that both morphology and motile features effectively distinguished control fibroblasts from activated fibroblasts, with the most pronounced differences observed in MDA-MB-231–co-cultured cells. Although morphology and motility alone provided robust classifications, integrating both features can enhance classification robustness by capturing a broader range of activation states. Fibroblast activation is inherently dynamic, with cells undergoing cytoskeletal remodeling and changes in migration behavior at different stages. A fibroblast may exhibit minimal morphological alterations while displaying significant motility shifts, or vice versa. By integrating both aspects, we can mitigate the risk of overlooking transitional activation states and capture a broader spectrum of fibroblast behaviors.

Given the dynamic nature of fibroblast activation process, it is important to consider the temporal limitations of our experimental setup. The short co-culture timeframe (5 days) or 12 h imaging window may not fully capture transient morphodynamic shifts or long-term phenotypic changes. For instance, Fig. 4 illustrates brief fluctuations in fibroblast morphology and motility that did not consolidate into stable subpopulations, suggesting that some observed states may reflect transient activation phases or partial reversion states rather than stable phenotypes. To confirm whether these short-lived morphodynamic transitions correspond to meaningful biological processes—such as early activation, partial reversion, or functional priming—rather than minor stochastic variations in shape, future experiments should extend the imaging period to 24–72 h and integrate real-time molecular reporters (e.g., fluorescently tagged α-SMA or PDPN). These modifications would allow tracking of morphodynamic fluctuations alongside gene expression changes, clarifying whether dynamic remodeling of fibroblast shape and motility plays a functional role in tumor progression or therapy resistance.

While our study focused on a controlled *in vitro* model, it underscores the potential of biophysical state-based classification as a powerful complement to transcriptomic profiling. Morphology and motility descriptors—such as aspect ratio, migration speed, and directional persistence—capture cytoskeletal organization and ECM interactions, both of which are hallmarks of fibroblast activation and drivers of tumor microenvironment remodeling. In a clinical context, this classification framework could be extended to patient-derived CAFs, enabling real-time phenotypic screening in high-throughput settings. Unlike fixed-cell assays, which capture only static molecular snapshots, live-cell imaging could reveal dynamic shifts in fibroblast states over time. Stratifying patients based on predominant fibroblast phenotypes in their tumors—such as elongated, highly motile CAFs vs more static, ECM-producing fibroblasts—could help identify stromal-targeting therapeutic strategies aimed at either disrupting pro-tumor CAFs or amplifying anti-tumor ones. That said, our current study relies on a short co-culture model and bulk qPCR validation, necessitating further studies with single-cell transcriptomic comparisons and longer-term analyses to fully link biophysical descriptors with gene expression.

To deepen these insights, integrating our biophysical classification with single-cell transcriptomic methods will be crucial. By examining how individual fibroblasts transition between morphodynamic states while simultaneously profiling their gene expression, we can establish more direct links between observed phenotypes and the molecular pathways that govern them. For example, aspect ratio and migration speed might correlate with ECM-remodeling or immune-modulatory gene programs, while turning angles and directional persistence could reflect pathways regulating cell–cell and cell–matrix interactions. Furthermore, performing single-cell analyses on *ex vivo*–expanded CAFs or 3D organoid systems derived from patient tumors would clarify whether the states detected in our NHDF co-cultures emerge under clinically relevant conditions. This multi-omics strategy—combining biophysical descriptors and transcriptional data—has the potential to refine stromal classification systems, revealing which fibroblast subsets most strongly promote tumor progression, immune evasion, or therapy resistance.

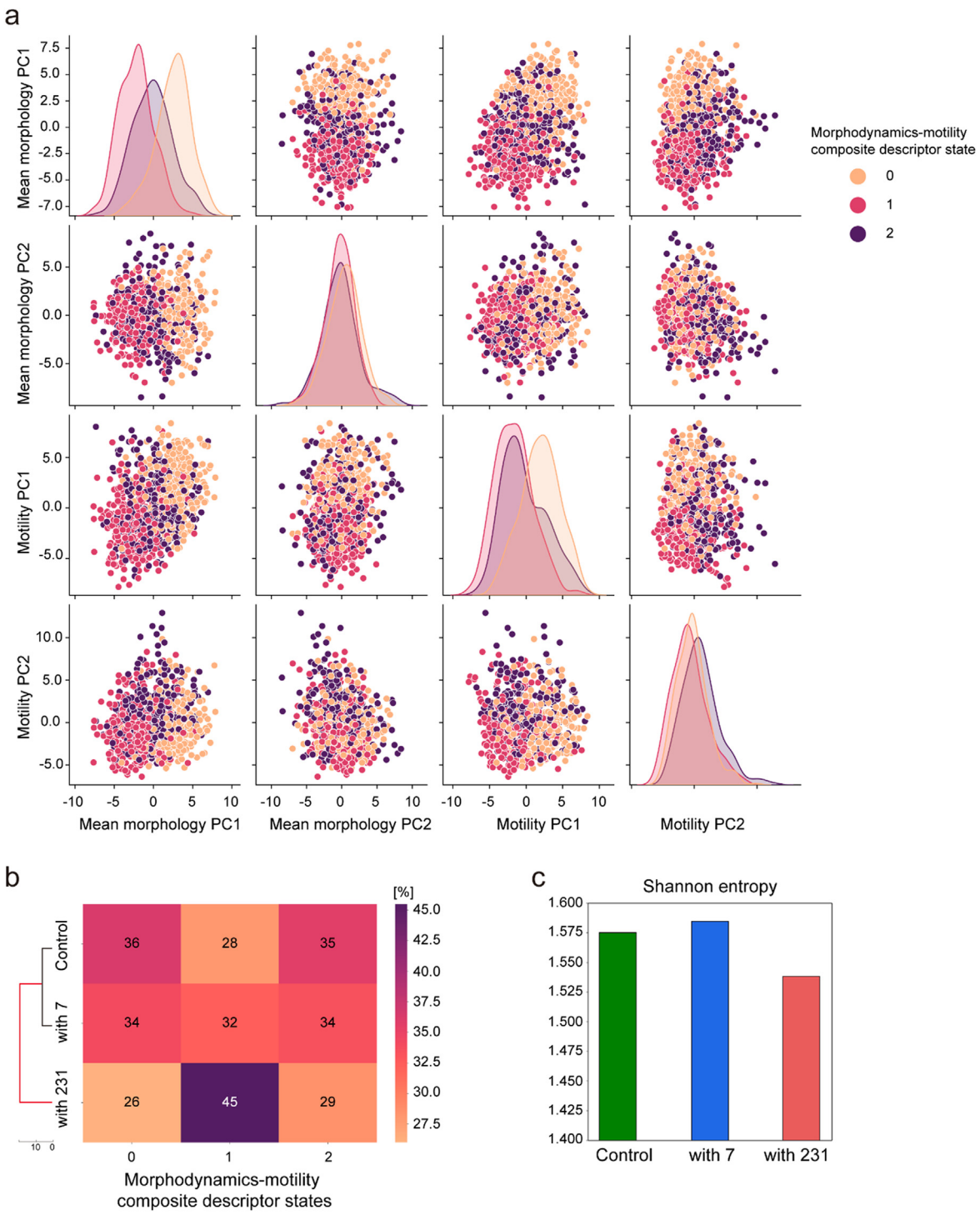

FIG. 7. Verification of morphodynamics–motility composite descriptors and distribution of morphodynamics–motility composite descriptor states of each group. (a) Distributions of cells labeled with morphodynamics–motility composite descriptor state onto various feature spaces (mean morphology PC1, 2 and motility PC1, 2). The diagonal plots are kernel density estimation (KDE) of the univariate distribution of the corresponding features. (b) Defined composite descriptor states distribution of each group and the unsupervised hierarchical clustering of fibroblasts. (c) Morphodynamics–motility heterogeneity within the groups represented by Shannon entropy.

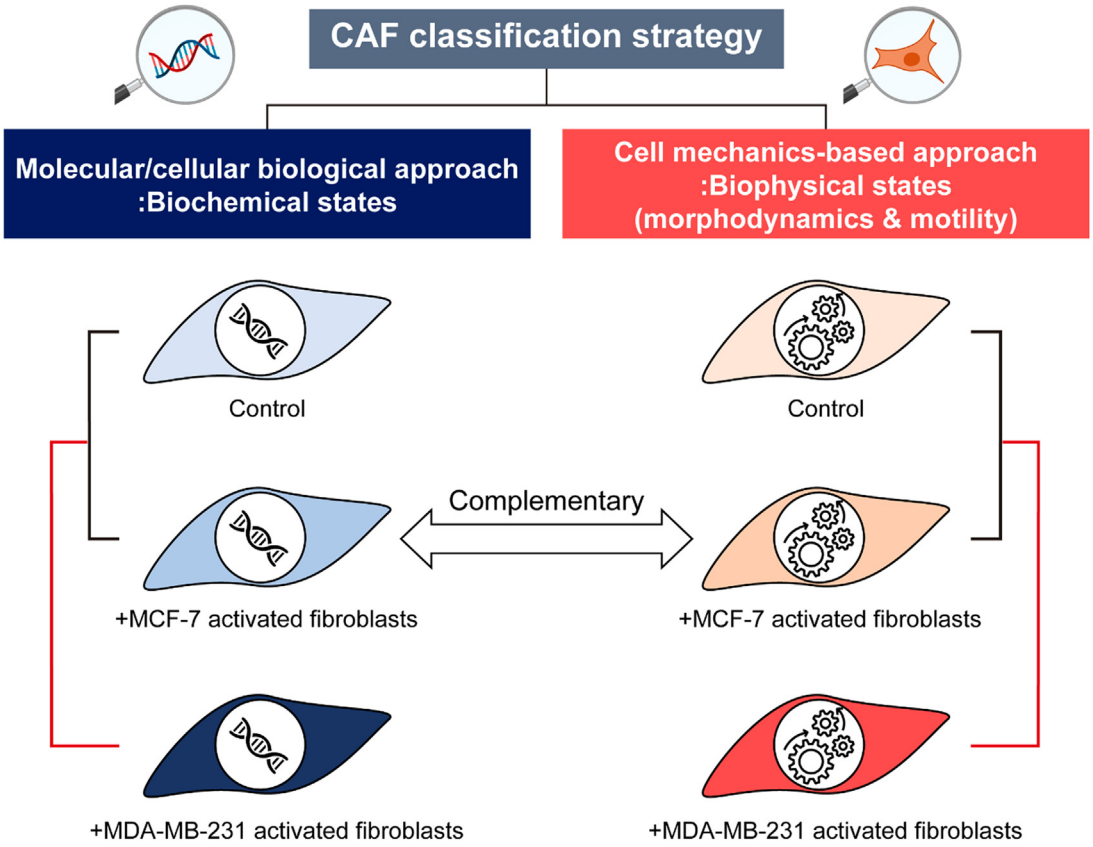

FIG. 8. Comparison between different CAF classification strategies. Biophysical state-based CAF classification predicted the same result as biochemical state-based classification.

By identifying the biophysical states that correspond to specific molecular networks, future research can inform targeted interventions against the most pro-tumor fibroblast phenotypes while preserving anti-tumor or homeostatic fibroblast populations. For instance, inhibitors or biologics targeting cytoskeletal regulators or ECM-remodeling enzymes might be selectively applied to CAF subpopulations that exhibit invasive morphodynamics. Ultimately, bridging the gap between traditional CAF markers (e.g., FAP, PDPN) and real-time morphological or motile phenotypes could catalyze the development of more precise stromal-targeting therapies, enhancing patient outcomes.

In summary, our study demonstrates that biophysical features—morphodynamics and motility—offer a complementary strategy for classifying fibroblast activation states, analogous to molecular profiling via transcriptomic markers. We show that highly aggressive cancer cells like MDA-MB-231 induce distinct mechanical changes in fibroblasts, underscoring the importance of biophysical classification as an independent and complementary tool to molecular profiling. By integrating morphodynamic and motile features, our approach provides a functional readout of fibroblast activation, refining classification systems while offering mechanistic insights into how cytoskeletal reorganization and motility reflect underlying activation programs. Future studies combining molecular and biophysical classifications may further improve fibroblast-based patient stratification, facilitating more precise stromal-targeting therapies to disrupt tumor-supportive fibroblast functions while preserving beneficial stromal populations.

METHODS

Cell culture and establishment of differentially activated fibroblast models

We cultured normal human adult dermal fibroblasts (NHDFs) (Lonza), MCF-7, and MDA-MB-231 cells separately in DMEM (Dulbecco's Modified Eagle Medium; WelGene) supplemented with 10% fetal bovine serum (FBS) and 1% penicillin–streptomycin (WelGene). To generate differentially activated fibroblasts, we co-cultured NHDFs with either MCF-7 or MDA-MB-231 cells in co-culture dishes (SPL, 209260) for five days. The seeding density for NHDFs was 4650 cells/cm^2, and for cancer cells, 2325 cells/cm^2. The co-culture experiments involved low-passage (<P6) NHDFs, all derived from the same initial stock, ensuring that cells did not undergo extensive splitting cycles. We conducted the four co-culture experiments on different days but followed identical protocols for seeding density, medium composition, incubation times, and imaging conditions, keeping all experiments within a tightly controlled three-to-four-week timeframe. To assess whether batch effects influenced fibroblast activation patterns, we examined PCA distributions across replicates

(supplementary material Fig. 2) and found no distinct clustering by experiment, confirming that we could analyze the data collectively. To verify that the five-day co-culture period induced activation, we performed real-time qPCR to measure the expression of multiple fibroblast markers.

Live-cell imaging of single fibroblasts

Live-cell imaging was conducted using an Axio Observer.Z1/7 (Carl Zeiss) microscope equipped with a climate-controlled chamber (37 °C, 5% CO_2). We captured phase-contrast images of single fibroblasts every 10 min for 24 h using a 5× objective lens and a 1× Optovar magnification changer. To ensure we could track individual cells and minimize overlap during migration, we seeded fibroblasts from each condition (monoculture or co-culture) sparsely at 2000 cells per well in six-well plates.

Instance segmentation, tracking, and extraction of morphodynamic features

The code to analyze cell image data was written in Python version 3.8. For instance, in the segmentation of single fibroblasts, we performed semantic segmentation using the U-net model with an adapted backbone.[33] Specifically, to create the training data, we manually annotated 1016 fibroblasts from 26 randomly selected images using ImageJ, generating binary mask images for semantic segmentation. We cropped these phase-contrast cell images and their corresponding binary masks (1792 × 1024 pixels) into 256 × 256 pixel patches for training. A total of 728 patches of phase-contrast images and their corresponding binary masks underwent data augmentation via random flipping, rotation, distortion, and brightness/contrast adjustments to generate 50 000 training samples. Resnet34 served as the encoder for the adapted U-net (supplementary material Fig. 7).[34] The model was trained using binary cross-entropy and Jaccard loss with Adam optimizer. The training parameters were a batch size of 32 and 100 epochs. The model's performance was evaluated using the IoU (Intersection over Union) score, with a validation score exceeding 0.7 [supplementary material Fig. 8(a)]. Although the training IoU was higher than the validation IoU, indicating potential overfitting, visual inspection of the segmented images confirmed that the model performed well despite occasional mispredictions [supplementary material Fig. 8(b)]. To improve the accuracy of the extracted features, we manually reviewed each segmented mask against its corresponding phase-contrast image and corrected any segmentation errors in ImageJ. This step was particularly crucial for cells exhibiting extreme elongation, overlapping boundaries, or irregular perimeters, as automatic predictions occasionally failed to capture their true shape. By refining each final mask to more accurately represent the cell contours, we minimized downstream errors in the extracted morphological and motility features.

During instance segmentation, we automatically removed binary masks smaller than 200 pixels or those touching the image frame boundary, as they did not represent valid cells. We then applied bounding boxes to label each cell and extracted 52 morphological features—including basic morphology, texture, and Zernike moments—for each labeled cell (supplementary material Fig. 3). We extracted ten basic morphological descriptors using the scikit-image library (v1.1),[35] categorizing them into three functional groups: (i) size-related metrics (area, perimeter, equivalent diameter, minor axis length); (ii) polarity/elongation descriptors (major axis length, aspect ratio, circularity); and (iii) compactness/occupancy indicators (extent, solidity, compactness). These features capture key aspects of fibroblast morphology, including spreading, elongation, and membrane protrusions—hallmarks of cytoskeletal remodeling and ECM interactions during fibroblast activation.[2] To analyze light-intensity variations in phase-contrast images, we computed Haralick texture features[36] from gray-level co-occurrence matrices (GLCM) using the mahotas library (v1.4.13).[37] Each feature was averaged across four orientations (0°, 45°, 90°, and 135°) to maintain rotational invariance. Zernike moments are a set of orthogonal polynomials defined on a unit circle, offering a rotation-invariant description of an object's shape. They are recognized as valuable quantitative shape descriptors for individual cells.[38,39] In this study, we computed Zernike polynomials up to degree 9, which yielded 30 descriptive values derived from their complex coefficients. By focusing on the magnitudes of these coefficients, we obtained robust shape features that are insensitive to rotation. Calculating Zernike moments via the mahotas library complements our basic morphological features by capturing finer contour details (e.g., membrane protrusions and indentations), which may reflect cytoskeletal remodeling in activated fibroblasts.

We tracked single cells over time by matching labeled cells in adjacent frames using the linear assignment problem (LAP) methodology. We constructed a pairwise cost matrix based on centroid displacement and cell size differences, applying threshold values of 100 pixels for displacement and a twofold difference for size, as previously described (expressions 1 and 2).[15] We then relabeled cells in subsequent frames according to their matches with previous frames, label terminations, or the appearance of new objects (supplementary material Video 1),

$$cost_{ij} = \left(r_{ij}^{size} + \frac{1}{r_{ij}^{size}} \right) \times {d_{ij}}^2, \tag{1}$$

$$r_{ij}^{size} = \frac{s_i}{s_j}, \tag{2}$$

where d_{ij} is the distance between the centroid of cell i and cell j from two adjacent frames, s_i is the size of cell i.

Extraction of motility features

We extracted cell centroids from each frame and calculated 50 motility features using the Heteromotility tool.[40] We grouped these features into three categories: general features, mean square displacement (MSD) features, and random walk features (supplementary material Fig. 5). The 33 general features included metrics such as net distance, linearity, progressivity, and speed-related statistics. We used MSD-derived α values to categorize cell motion as random, directed, or confined. Additionally, we extracted 16 random walk features, including displacement kurtosis, Hurst exponent, autocorrelation, and partial autocorrelation.

Dimensionality reduction of morphology and motility features with principal component analysis (PCA) and UMAP for morphodynamics–motility composite descriptor

We extracted 52 morphological features from each cell and applied PCA to reduce dimensionality. We then embedded a total of

70 018 morphology states into a two-dimensional morphology space (morphology PC1–PC2). Using the same procedure, we processed motility features from 937 cells and embedded them into a two-dimensional motility state space (motility PC1–PC2). To define the morphodynamics–motility composite descriptor based on the morphology and motility PCs, we applied UMAP (version 0.5.2) with default parameters.

Unsupervised clustering to define biophysical states

To identify the heterogenous subtypes of fibroblasts in the two-dimensional biophysical state space (morphology PC1–PC2 or motility PC1–PC2 spaces), an unsupervised clustering algorithm, k-means clustering, was implemented. The optimal number of clusters was determined by a combination of the elbow method, silhouette coefficient, and Davies–Bouldin Index (DBI), balancing quantitative metrics with biological interpretability (supplementary material Fig. 9). For morphology, which contained approximately 70 000 observations, the elbow plot exhibited plausible inflection points at $k = 3$ and $k = 7$. However, $k = 3$ was deemed too coarse to capture the extensive shape diversity, while silhouette analysis supported $k = 6$–7. Although $k = 6$ had a slightly lower DBI than $k = 7$ (0.8828 vs 0.8895), the marginal numerical difference, combined with the large dataset size, favored $k = 7$. For motility, based on ∼1000 tracked cells, the elbow and silhouette plots pointed to $k = 3$ or $k = 6$ as potential options. $k = 3$ emerged as the best choice, given its peak in silhouette score and a lower DBI (0.8913) compared to $k = 4$ (0.9461). Splitting into six or seven clusters risked over-fragmenting a more modest dataset. For the morphodynamics–motility composite descriptor space, which combined time-averaged morphology and motility principal components, the dataset size was the same as motility-only dataset (∼1000 observations). The elbow and silhouette graphs again suggested $k = 3$, and $k = 3$ had a favorable DBI (0.8854), indicating meaningful partitioning without generating overly small clusters. In addition to these quantitative metrics, qualitative validation was performed. We examined representative cells near the centroid of each morphology cluster in the PC space and observed consistent shape traits, such as elongated vs rounded morphologies (supplementary material Fig. 10). For motility clusters, we analyzed time-lapse videos of single cells from each cluster to confirm differences in speed and turning behavior (supplementary material Videos 2–4). In the composite feature space, clusters reflected subtle shape–motion correlations, and selection of k relied primarily on quantitative metrics rather than additional cell-level visualizations. For the clustering of time-series data, time-series k-means clustering with dynamic time warping (DTW) was implemented.

After defining the clusters (the defined biophysical states), the fibroblasts' distribution of clusters was calculated and visualized with a heatmap. Then, unsupervised hierarchical clustering with average-linkage and Euclidean distance was carried out. To quantify the general heterogeneity of fibroblasts, Shannon entropy was used,

$$\text{Shannon entropy} = -\sum p(x)\log p(x), \tag{3}$$

where $p(x)$ is the % distribution of the clusters.

Reverse transcription quantitative PCR

After co-culturing for five days, the total RNA of fibroblasts was isolated with RNAiso reagent (Takara Bio, Japan) according to the manufacturer's instructions. We reverse-transcribed the extracted RNAs into cDNA using iScript cDNA Synthesis Kits (Bio-Rad, USA) and a Biometra T-personal Thermal Cycler. We then performed real-time qPCR in duplicates using iQ SYBR Green Supermix (Bio-Rad, USA) on a Bio-Rad CFX96 real-time detection system. We used glyceraldehyde 3-phosphate dehydrogenase (GAPDH) as the reference gene. For hypothesis testing, we used ΔCt values to express relative mRNA expression normalized to GAPDH. $\Delta Ct = Ct(\text{reference gene}) - Ct(\text{gene of interest})$, where Ct corresponds to the threshold cycle.

We used the following primers:

GAPDH (for: CTGGGCTACACTGAGCACC, Rev: AAGTGGTC GTTGAGGGCAATG),
ACTA2 (for: CTATGAGGGCTATGCCTTGCC, Rev: GCTCAG CAGTAGTAACGAAGGA),
FAP (for: TGAACGAGTATGTTTGCAGTGG, Rev: GGTCTTT GGACAATCCCATGT),
FSP1 (for: GATGAGCAACTTGGACAGCAA, Rev: CTGGGCTG CTTATCTGGGAAG),
PDPN (for: AACCAGCGAAGACCGCTATAA, Rev: CGAATG CCTGTTACACTGTTGA),
PDGFRα (for: TTGAAGGCAGGCACATTTACA, Rev: GCGACA AGGTATAATGGCAGAAT),
PDGFRβ (for: TGATGCCGAGGAACTATTCATCT, Rev: TTTCT TCTCGTGCAGTGTCAC).

STATISTICAL ANALYSIS

Prior to statistical hypothesis testing, normality was assessed using the Shapiro–Wilk test, and homogeneity of variances was evaluated with Levene's test. We performed statistical hypothesis testing using Welch's one-way ANOVA followed by Games–Howell post hoc tests. We considered p-values < 0.05 statistically significant. Statistical analyses were performed using jamovi version 1.6.23.0 [The jamovi project (2021), https://www.jamovi.org], and custom Python scripts utilizing the pingouin and statannotations packages, including Welch's ANOVA and Games-Howell post hoc analysis.

SUPPLEMENTARY MATERIAL

See the supplementary material for the following: expression analysis of additional CAF markers by qPCR (supplementary material Fig. 1); principal component analysis and replicate validation via kernel density plots (supplementary material Fig. 2); summary of extracted morphological, texture, and Zernike features (supplementary material Fig. 3); comparative analysis of fibroblast morphology across co-culture conditions (supplementary material Fig. 4); summary of extracted motility features (supplementary material Fig. 5); dimensionality reduction of morphodynamics–motility composite descriptors with PCA (supplementary material Fig. 6); deep learning architecture for single-cell segmentation (supplementary material Fig. 7); validation of segmentation model performance (supplementary material Fig. 8); clustering analyses for morphology, motility, and composite features (supplementary material Fig. 9); representative cell images from each morphology cluster (supplementary material Fig. 10); statistical validation of CAF marker expression analysis (supplementary material Table 1); and single-cell segmentation demonstration (supplementary material Video 1) and representative cell movies of motility clusters 0–2 (supplementary material Videos 2–4).

ACKNOWLEDGMENTS

This research was supported by the National Research Foundation of Korea Grant funded by the Korean Government (NRF-2021R1A2C3008408) and the Korean Fund for Regenerative Medicine Grant (KFRM 22B0401L1), funded by the Korean Government (the Ministry of Science and ICT, the Ministry of Health and Welfare).

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Ethics Approval

Ethics approval is not required

Author Contributions

Minwoo Kang: Conceptualization (lead); Investigation (lead); Methodology (equal); Software (equal); Visualization (lead); Writing – original draft (lead); Writing – review & editing (equal). **Chanhong Min:** Methodology (equal); Software (equal); Visualization (supporting). **Somayadineshraj Devarasou:** Investigation (supporting); Writing – review & editing (supporting). **Jennifer H. Shin:** Funding acquisition (lead); Investigation (supporting); Supervision (lead); Writing – review & editing (equal).

DATA AVAILABILITY

The data that support the findings of this study are openly available in GitHub, Ref. 41.

REFERENCES

[1]K. A. Gieniec, L. M. Butler, D. L. Worthley, and S. L. Woods, "Cancer-associated fibroblasts-heroes or villains?," Br. J. Cancer **121**, 293–302 (2019).
[2]E. Sahai *et al.*, "A framework for advancing our understanding of cancer-associated fibroblasts," Nat. Rev. Cancer **20**, 174–186 (2020).
[3]D. Ganguly *et al.*, "Cancer-associated fibroblasts: Versatile players in the tumor microenvironment," Cancers **12**, 2652 (2020).
[4]H. Li *et al.*, "Reference component analysis of single-cell transcriptomes elucidates cellular heterogeneity in human colorectal tumors," Nat. Genet. **49**, 708–718 (2017).
[5]F. Pelon *et al.*, "Cancer-associated fibroblast heterogeneity in axillary lymph nodes drives metastases in breast cancer through complementary mechanisms," Nat. Commun. **11**, 404 (2020).
[6]A. N. Hosein *et al.*, "Cellular heterogeneity during mouse pancreatic ductal adenocarcinoma progression at single-cell resolution," JCI Insight **4**, e129212 (2019).
[7]C. X. Dominguez *et al.*, "Single-cell RNA sequencing reveals stromal evolution into LRRC15(+) myofibroblasts as a determinant of patient response to cancer immunotherapy," Cancer Discovery **10**, 232–253 (2020).
[8]E. Elyada *et al.*, "Cross-species single-cell analysis of pancreatic ductal adenocarcinoma reveals antigen-presenting cancer-associated fibroblasts," Cancer Discovery **9**, 1102–1123 (2019).
[9]D. Ohlund *et al.*, "Distinct populations of inflammatory fibroblasts and myofibroblasts in pancreatic cancer," J. Exp. Med. **214**, 579–596 (2017).
[10]G. Friedman *et al.*, "Cancer-associated fibroblast compositions change with breast cancer progression linking the ratio of S100A4(+) and PDPN$^+$ CAFs to clinical outcome," Nat. Cancer **1**, 692 (2020).
[11]M. Bartoschek *et al.*, "Spatially and functionally distinct subclasses of breast cancer-associated fibroblasts revealed by single cell RNA sequencing," Nat. Commun. **9**, 5150 (2018).
[12]D. E. Costea *et al.*, "Identification of two distinct carcinoma-associated fibroblast subtypes with differential tumor-promoting abilities in oral squamous cell carcinoma," Cancer Res. **73**, 3888–3901 (2013).
[13]W. Wang, D. Poe, Y. Yang, T. Hyatt, and J. Xing, "Epithelial-to-mesenchymal transition proceeds through directional destabilization of multidimensional attractor," Elife **11**, e74866 (2022).
[14]W. Wang *et al.*, "Live-cell imaging and analysis reveal cell phenotypic transition dynamics inherently missing in snapshot data," Sci. Adv. **6**, eaba9319 (2020).
[15]Z. Wu *et al.*, "DynaMorph: Self-supervised learning of morphodynamic states of live cells," Mol. Biol. Cell **33**, ar59 (2022).
[16]F. J. Hartmann *et al.*, "Single-cell metabolic profiling of human cytotoxic T cells," Nat. Biotechnol. **39**, 186–197 (2021).
[17]D. A. Van Valen *et al.*, "Deep learning automates the quantitative analysis of individual cells in live-cell imaging experiments," PLoS Comput. Biol. **12**, e1005177 (2016).
[18]M. Majety, L. P. Pradel, M. Gies, and C. H. Ries, "Fibroblasts influence survival and therapeutic response in a 3D co-culture model," PLoS One **10**, e0127948 (2015).
[19]E. Prieto-Garcia *et al.*, "Tumor-stromal interactions in a co-culture model of human pancreatic adenocarcinoma cells and fibroblasts and their connection with tumor spread," Biomedicines **9**, 364 (2021).
[20]S. Y. Jeong, J. H. Lee, Y. Shin, S. Chung, and H. J. Kuh, "Co-culture of tumor spheroids and fibroblasts in a collagen matrix-incorporated microfluidic chip mimics reciprocal activation in solid tumor microenvironment," PLoS One **11**, e0159013 (2016).
[21]E. Elwakeel and A. Weigert, "Breast cancer CAFs: Spectrum of phenotypes and promising targeting avenues," Int. J. Mol. Sci. **22**, 11636 (2021).
[22]V. Bernard *et al.*, "Single-cell transcriptomics of pancreatic cancer precursors demonstrates epithelial and microenvironmental heterogeneity as an early event in neoplastic progression," Clin. Cancer Res. **25**, 2194–2205 (2019).
[23]E. Elwakeel *et al.*, "Phenotypic plasticity of fibroblasts during mammary carcinoma development," Int. J. Mol. Sci. **20**, 4438 (2019).
[24]M. L. Zou *et al.*, "Fibroblasts: Heterogeneous cells with potential in regenerative therapy for scarless wound healing," Front. Cell Dev. Biol. **9**, 713605 (2021).
[25]S. Mahmoudi *et al.*, "Heterogeneity in old fibroblasts is linked to variability in reprogramming and wound healing," Nature **574**, 553–558 (2019).
[26]Z. Chen and S. K. Sun, "A Zernike moment phase-based descriptor for local image representation and matching," IEEE Trans. Image Process. **19**, 205–219 (2010).
[27]C. Kan and M. D. Srinath, "Invariant character recognition with Zernike and orthogonal Fourier-Mellin moments," Pattern Recognit. **35**, 143–154 (2002).
[28]A. Khotanzad and J. H. Lu, "Classification of invariant image representations using a neural network," IEEE Trans. Acoust. Speech Signal Process. **38**, 1028–1038 (1990).
[29]Z. Iscan, Z. Dokur, and T. Ölmez, "Tumor detection by using Zernike moments on segmented magnetic resonance brain images," Expert Syst. Appl. **37**, 2540–2549 (2010).
[30]A. Tahmasbi, F. Saki, and S. B. Shokouhi, "Classification of benign and malignant masses based on Zernike moments," Comput. Biol. Med. **41**, 726–735 (2011).
[31]Y. Kumar, A. Aggarwal, S. Tiwari, and K. Singh, "An efficient and robust approach for biomedical image retrieval using Zernike moments," Biomed. Signal Process. Control **39**, 459–473 (2018).
[32]S. Devarasou, M. Kang, and J. H. Shin, "Biophysical perspectives to understanding cancer-associated fibroblasts," APL Bioeng. **8**, 021507 (2024).
[33]O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," Lect. Notes Comput. Sci. **9351**, 234–241 (2015).
[34]K. M. He, X. Y. Zhang, S. Q. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (IEEE, 2016), pp. 770–778.
[35]S. van der Walt *et al.*, "scikit-image: Image processing in Python," PeerJ **2**, e453 (2014).
[36]R. M. Haralick, "Statistical and structural approaches to texture," Proc. IEEE **67**, 786–804 (1979).
[37]L. P. Coelho, "Mahotas: Open source software for scriptable computer vision," J. Open Res. Software **1**, e3 (2013).

[38]E. Alizadeh, S. M. Lyons, J. M. Castle, and A. Prasad, "Measuring systematic changes in invasive cancer cell shape using Zernike moments," Integr. Biol. **8**, 1183–1193 (2016).
[39]Z. Pincus and J. A. Theriot, "Comparison of quantitative methods for cell-shape analysis," J. Microsc. **227**, 140–156 (2007).
[40]J. C. Kimmel, A. Y. Chang, A. S. Brack, and W. F. Marshall, "Inferring cell state by quantitative motility analysis reveals a dynamic state system and broken detailed balance," PLoS Comput. Biol. **14**, e1005927 (2018).
[41]See https://github.com/shinlab-kaist for "GitHub."