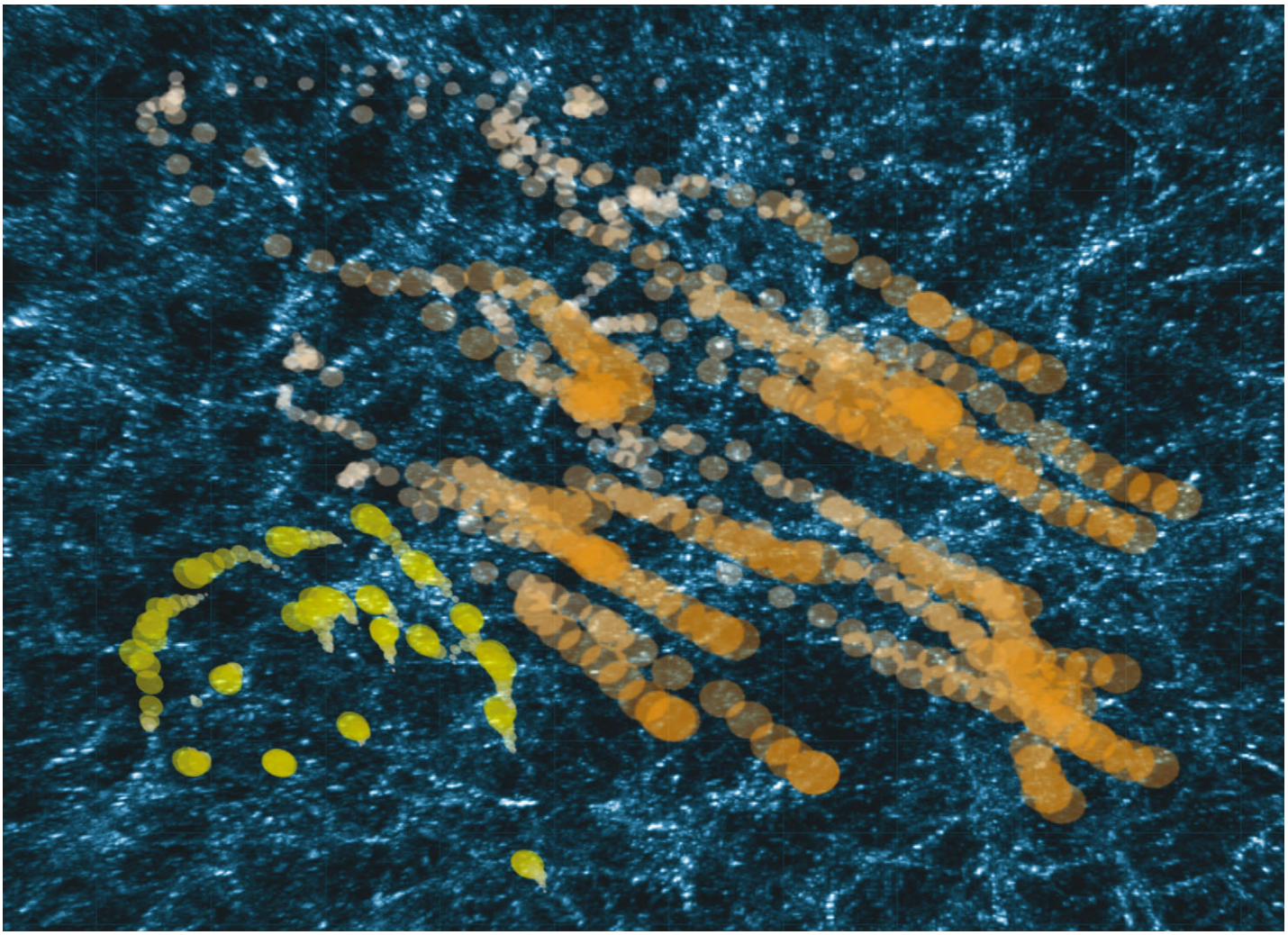

Featuring work from the Zaman Lab, at the Department of Biomedical Engineering at Boston University, whose work probes cell motility in 3-dimensional matrices.

Collective motion of mammalian cell cohorts in 3D

Collective migration of cells is suspected to be predominant in cancer metastasis. This article describes experimental and analytical techniques for collective cell motility. These studies on dynamic properties of cellular clusters reveal the heterogeneity and emergence inherent within collective cellular systems.

As featured in:

See Yasha Sharma, Muhammad H. Zaman *et al.*, *Integr. Biol.*, 2015, **7**, 1526.

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/ibiology

Registered charity number: 207890

Integrative Biology

PAPER

View Article Online
View Journal | View Issue

Cite this: *Integr. Biol.*, 2015, 7, 1526

Collective motion of mammalian cell cohorts in 3D†

Yasha Sharma,[a] Diego A. Vargas,[a] Adrian F. Pegoraro,[b] David Lepzelter,[a] David A. Weitz[b] and Muhammad H. Zaman*[a]

Received 18th August 2015,
Accepted 26th October 2015

DOI: 10.1039/c5ib00208g

www.rsc.org/ibiology

Collective cell migration is ubiquitous in biology, from development to cancer; it occurs in complex systems comprised of heterogeneous cell types, signals and matrices, and requires large scale regulation in space and time. Understanding how cells achieve organized collective motility is crucial to addressing cellular and tissue function and disease progression. While current two-dimensional model systems recapitulate the dynamic properties of collective cell migration, quantitative three-dimensional equivalent model systems have proved elusive. To establish such a model system, we study cell collectives by tracking individuals within cell cohorts embedded in three dimensional collagen scaffolding. We develop a custom algorithm to quantify the temporal and spatial heterogeneity of motion in cell cohorts during motility events. In the absence of external driving agents, we show that these cohorts rotate in short bursts, <2 hours, and translate for up to 6 hours. We observe, track, and analyze three dimensional motion of cell cohorts composed of 3–31 cells, and pave a path toward understanding cell collectives in 3D as a complex emergent system.

Insight, innovation, integration

Collective cell migration is crucial in multiple biological processes and disease; biochemical, mechanical and physical factors combined can lead to diverse physical behaviors. 2D model systems for collective cell migration yield promising insights into the link between cell-dynamics and biological function, but are unable to capture the *in vivo* complexity. In order to address the inherent challenges of 2D systems, we present a 3-dimensional experimental model for the motion of cellular cohorts in native like environments. We show that cell cohorts have spatiotemporal heterogeneity that can be mapped and quantified. We have devised quantitative algorithms that treat each cohort as unique, and isolate 'motility events' that occur over the period of observation. These techniques allow for systematic processing of large amounts of data, and probe heterogeneity between and within cellular cohorts.

Introduction

From early development to morphogenesis, wound healing, and even in cancer pathologies, biological function hinges on collective cell migration. Currently, studies of the inter-cellular dynamics of collective cell motion are primarily conducted *via* two dimensional (2D) monolayer experiments. Here we present a 3-dimensional study of cell collectives.

Cells must coordinate adherence and motility to maintain organized coherent motion;[1] extensive work continues to probe how cells establish such communication and organization.[2,3] Cells migrate collectively to build and vascularize tissues, heal wounds, and occasionally in tumor metastases. A comprehensive review of collective cell motility establishes the many modes of migration available to cell collectives as well as the forces that drive motility.[4] A broad classification for 3D collective migration modes can comprise of two categories—in the first, cells never dissociate from their original tissue, as in the case of branching morphogenesis, angiogenesis, and multicellular strand invasion of cancers; and in the second, a detached cluster moves through ECM and other non-motile cells as observed in cancer metastases and drosophila border cells.[5,6]

From a mechanical and physical standpoint, time-lapse imaging and immunohistochemistry reveal relevant characteristics of collective cell migration (CCM).[1,7,8] For example, in *Drosophila*, E-cadherin is essential for collective direction sensing,[9] and tissue rotation is essential for building an extracellular-matrix (ECM) to control egg shape.[10] In a 3D example,

[a] *Department of Biomedical Engineering, Boston University, Boston, Massachusetts 02215, USA. E-mail: zaman@bu.edu*
[b] *School of Engineering and Applied Sciences, Harvard University, Cambridge, Massachusetts 02138, USA*

† Electronic supplementary information (ESI) available. See DOI: 10.1039/c5ib00208g

human mammary cells embedded in 3D gels reveal that rotation is essential to the formation of breast acini – it does not occur for cancerous breast cells, and when disrupted within normal cells, acini do not form.[11] Tracking assays on monolayers reveal density-dependent phase transitions,[12] substrate dependence,[13] and the forces driving CCM.[14,15] A comprehensive study of the mechanical properties of epithelial monolayers identified E-cadherin and P-cadherin as key proteins contributing to intercellular forces.[16] Heterogeneity emerges within groups of cells exhibiting collective behavior – functionally distinct populations of cells are termed leader and follower cells.[4] Leader cells are located at the front of a moving collective; they are responsible for receiving cues and directing the collective. Well-defined leader cells are found in cases of sprouting morphogenesis and angiogenesis.[17] In sheet migration, key molecules are upregulated to form leader cells at the leading edge; removal of these leader cells disrupts migration.[18]

Cancer pathologies are not amenable to direct observations of coherent translation due to diagnostic limitations; however, indirect evidence from *in vivo* measurements demonstrates that cancer metastases can migrate through tissue layers as collective masses.[5,6,19] Clinically, patients with epithelial-originating cancers or carcinomas present with circulating tumor microemboli, or clusters of circulating tumor cells up to 8 cells large.[20,21] Typical 3D studies of cell collectives involve immunohistochemistry assays and invasion assays of immortalized cancerous and non-cancerous cell lines. Immunohistochemistry has elucidated biochemical markers crucial to the emergence of leader-follower heterogeneity[22] in cancer cell lines. Invasion assays involve seeding a large spheroid (>200 μm in diameter) of cancerous or non-cancerous cells into a 3D matrix; the subsequent invasion of the spheroid into the matrix can take the form of single cell invasion or multicellular strand invasion. Time-lapse microscopy conducted on invasion assays highlights cell dynamics, leader-cell formation,[23] and cell jamming;[24] together these data suggest that cancer cells have inherent plasticity of migration modes and the ability to transition between these modes.[25]

The dynamics of collective cell motility are essential to understanding collective processes and function. In 2D environments, epithelial cells and fish keratocytes[26] have been used as model systems to study the dynamic aspects of collective cell migration. Here, we present a model system for quantifying 3D collective migration using mammalian cell cohorts comprised of three to thirty-one cells. This can serve as a tool for understanding the motility of detached cellular clusters that have been observed in cancer metastases *in vitro* and found as circulating tumor microemboli *in vivo*. We track individual cells within cohorts embedded in a 3D scaffold and identify events of emergent collective behavior in the absence of external driving agents. Our system serves as a 3D experimental model for collective motility of cells that is able to analyze each cell cohort as an individual entity. It is a first step toward a physical understanding of collective cell motility in 3D, including cancer cell invasion and the critical conditions that lead toward collective metastasis.

Materials and methods

Cell culture

MDCK Type II epithelial cells are propagated in monolayers at 37 °C, 5% CO_2, and ~70% humidity; monolayers are cultured in DMEM media supplemented with 10% fetal bovine serum and 1% penicillin/streptomycin. Cells are stably transfected to express Nuclear Localization Signal (NLS) bound to Green Fluorescent Protein (GFP). Transfection is accomplished by a GFP-NLS plasmid (Clontech, Takara Bio, Japan) of Lipofectamine 2000 (Invitrogen, Life Technologies, Grand Island, NY). To maintain fluorescence 0.5 mg ml^{-1} G418 is added to the media; Fluorescence Activated Cell Sorting (FACS) selects for the brightest 1% of cells. Identical procedures are followed on MDCK GFP-Ecad cell lines.[27]

3D Cluster protocol

Single cell suspensions are formed by immersing cells in media after trypsinization; cells are passed through a 40 μm cell strainer (BD Biosciences, San Jose, CA); cells are then seeded onto a 10 cm diameter Ultra Low Attachment Dish (Corning, Corning NY) with 10 ml media. After 48 hours, clusters are extracted by passing the solution through a 100 μm cell strainer followed by a 40 μm cell strainer, retaining clusters of 10–20 cells. These are resuspended and centrifuged at 800 rcf, and then immersed in a collagen solution for 3D culture and imaging.

A 2 mg ml^{-1} collagen dilution is obtained by mixing equal volumes of collagen Type 1 stock (BD Biosciences, San Jose, CA) solution and neutralizing buffer (100 mM Hepes in 2× PBS, pH 7.3) with PBS. Cell clusters are added to the 2 mg ml^{-1} collagen solution; this cluster-collagen suspension is seeded onto several wells of a 24-well plate or 96 well-plate (MatTek, Ashland MA). The plates are incubated at 37 °C, 5% CO_2 and ~70% humidity for 2 hours until the collagen has polymerized, after which ~1–2 ml of growth media is added to each well. FluoSpheres® Carboxylate-Modified Microspheres in 1.0 μm (Invitrogen) with red fluorescence (580/605) are diluted to $\sim 10^8$ beads per ml collagen when used.

Imaging and tracking

Images are acquired with a DMI600B Microscope (Leica, Solms, Germany) and ImagEM EM-CCD Camera (Hamamatsu Photonics, Hamamatsu, Japan) using a Spinning Disk Confocal setup (Yokogawa, Tokyo, Japan). Micro-Manager 1.4 Software (http://www.micro-manager.org) employs a 10× 0.3 NA objective lens to image $\sim 560 \times 560 \times 100$ μm^3 fields of view. 3D stacks are acquired in the *XY* plane with a Z-step of 4 μm, every 10 minutes, for ~48 hours over a 100–200 μm depth. For the experiment with GFP-Ecad a 20× 0.4 NA objective lens was used with a Z-step of 2 μm over a 42 μm depth. Cells can sense the substrate beneath the 3D matrix from inside the collagen gel,[28] and cells that are closer to the glass bottom show very high proliferation along with sheet-like dynamics. Since the scope of this work pertains to 3D cell morphology, acquisition and analysis is restricted to cells clusters located >100 μm from the glass bottom. 3D morphology is further verified before tracking and analysis.

ImageJ (NIH) is used to estimate the average nuclear diameter (~8 μm). This diameter is input into a Matlab (MathWorks, Natick, MA) spot-tracking algorithm designed by the Kilfoil group;[29] the algorithm is modified for 3D nucleus detection and tracking. Parameters such as nucleus diameter, mask, minimum track time, and maximum displacement between consecutive time points, are optimized until varying each parameter independently has a minimal effect on the output. This metric is optimized until on average, 93% of all nuclei identified are assigned to a track. New cell identifiers are assigned when tracking is lost after 3 consecutive time points or after cell division. For a single track, if a cell is missing for up to two time points, its position is interpolated by assuming a straight path. A de-drifting algorithm designed by the Kilfoil group[30] eliminates net motion that is common to all cells in the 3-dimensional field of view to account for stage drift, which is ~25–40 μm every 24 hours.

Clustering algorithm

A custom hierarchical clustering algorithm is written in Matlab to sort cells into cohorts, or groups of cells that are physically attached to each other. The algorithm is agglomerative – each cell is initially assigned a unique cluster identifier. For the first cell considered, all cells positioned within 35 μm are assigned to its cluster; all cells positioned within 35 μm of those cells are then assigned the same cluster. The process repeats until there are no cells that could be grouped into the same cluster; then the next cell with a unique cluster ID is considered. For these cells, a cutoff distance of 35 μm is empirically determined to be ~1.5 3D cell lengths; thus minimizing the likelihood of skipping an adherent neighbor. Cutoff distances ranging between 25–45 μm do not affect the output data. A custom function autocorrelates cluster IDs between consecutive time-points to ensure that each cluster has a unique ID for the duration of the experiment. A cluster is reassigned the same ID if it retains a majority of cell IDs from the previous time point, such that if two cohorts merge, the new cluster is labeled as the larger of the two.

Displacement squared quartiles and order

Displacements of each cell in a cohort between time $t + 0.5 \times T_{int}$ and $t - (0.5 \times T_{int} + \Delta T)$ are calculated across the entire timespan of the experiment, where T_{int} = 1 h and ΔT = 10 min (gap between consecutive time points). This results in a distribution with as many values as number of cells in the cohort at each time point. Displacements are squared, and the median, upper-quartiles, and lower-quartiles of this distribution are evaluated for all time points of the experiment. To calculate order parameter,[31] a smoothing function is run on *XYZ* position data between consecutive time points according to eqn (1) where x represents position and t represents time; the interval between consecutive data points is 10 minutes.

$$\vec{x}_t = \frac{\vec{x}_{t-\Delta T} + \vec{x}_t + \vec{x}_{t+\Delta T}}{3} \quad (1)$$

The T_{int} order parameter is calculated for the cohort between time $t + 0.5 \times T_{int}$ and $t - (0.5 \times T_{int} + \Delta T)$ as shown in eqn (2) where v is velocity and N is the number of cells in the cohort.

$$\varphi(t) = \frac{\left|\sum_{i=1}^{N} \vec{v}_{iT_{int}}\right|}{\sum_{i=1}^{N} \left|\vec{v}_{iT_{int}}\right|} \quad (2)$$

T_{int} is selected by studying Mean Squared Displacement (MSD) *vs.* time interval plots (data not shown) for all cells in the experiment. MSD plots suggest that the cells in these experiments have high heterogeneity of behavior over intervals as low as 30 minutes. In order to account for bias induced by tracking, de-drifting, and noise, we doubled this number to set T_{int} = 1 h.

Automated event selection

To analyze individual cohorts, a custom algorithm is written in Matlab to detect motility events from median displacement squared data. Initially, Matlab's built-in peak finding algorithm is used to find all peaks in the data. Peaks are merged if the valley between them $> 0.5 \times P_{min}$ and the time gap between them $< 1.5 \times T_{int}$. Then peaks with width $< T_{int}$ or height $> P_{min}$ are eliminated. P_{min}, or the minimum peak height for a motility event, is conservatively set at 60 μm^2, in order to track motion of ~1 3D nucleus diameter and minimize the loss of relevant information.

Pairwise correlations

Once an event is identified, smoothed positions of cells for that event inform correlation functions between all possible cell pairs within a cohort. This correlation function is represented by eqn (3) where i and j are the cell pair, τ is time difference, t is time, and v is the velocity.[32]

$$C_{i,j}(\tau) = \left\langle \left(\frac{\vec{v}_i}{|\vec{v}_i|}\right)_t \cdot \left(\frac{\vec{v}_j}{|\vec{v}_j|}\right)_{t+\tau} \right\rangle \quad (3)$$

This correlation function reaches a maximum peak value at a lag time τ_c; when a peak has a height > 0.5, a correlation is considered significant and τ_c is retained. For positive τ_c, cell i lags cell j with duration τ_c; conversely, for negative τ_c, cell j lags cell i with duration τ_c.

Results

To investigate long term behavior and heterogeneity of motion in time, 3D cell tracking is performed on representative cell cohorts comprising 3–31 cells every 10 minutes over a duration of 48 hours. Positions, cell IDs, and cluster IDs for twelve cell cohorts are obtained from two different 2 mg ml^{-1} collagen gels and five independent fields of view. Cohorts are dynamic and exhibit spatial and temporal heterogeneity; behavior may include seemingly random movement, collective rotation, or collective translation. Planar projections and renderings of nuclear tracking at 0 h, 24 h and 48 h are represented in red, green and blue, respectively, in Fig. 1. Two of the cohorts merge between 24 and 48 h, as visible in Fig. 1C and D. For this

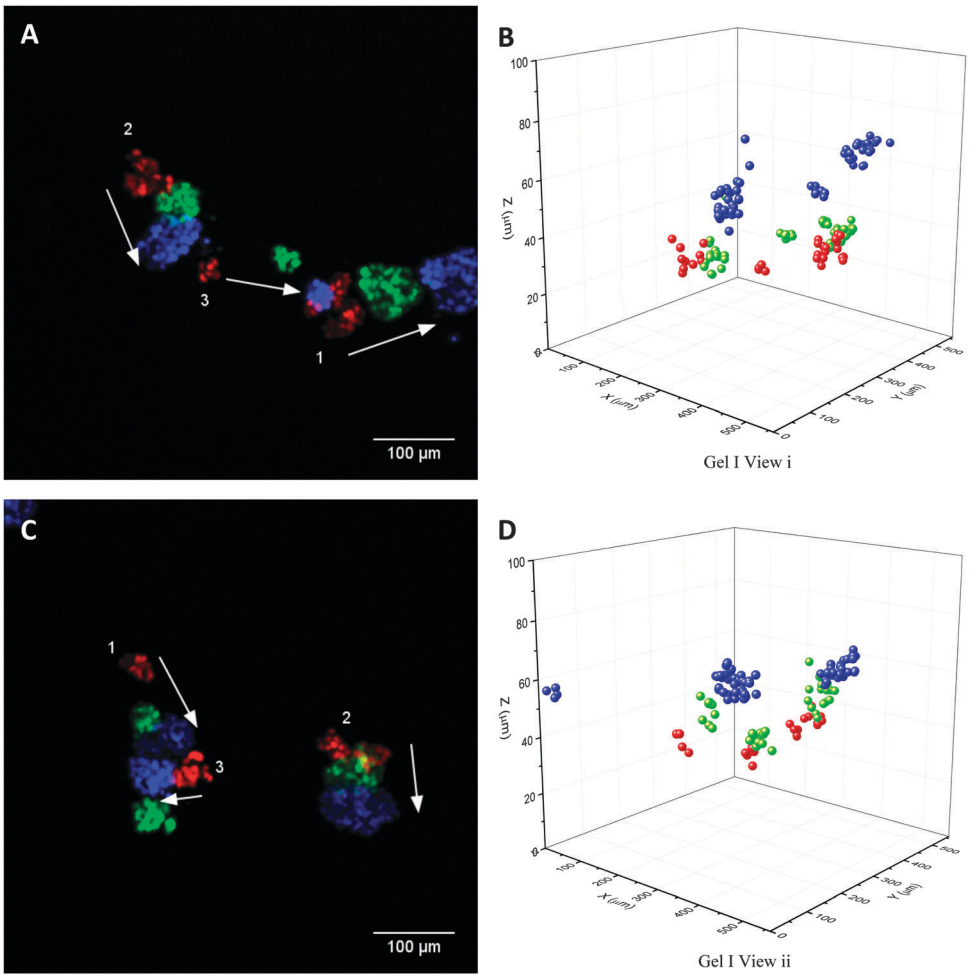

Fig. 1 Panels A and C show Z-projections of 2 fields of view from the same 3D collagen gel with MDCK GFP-NLS cell cohorts at 0 h (red), 24 h (green) and 48 h (blue). The numbers in white indicate cohort number as determined by a clustering algorithm. Panels B and D are 3D renderings of nuclear tracking corresponding to Panel A and C respectively.

merging event, one cohort changes direction of motion in the second half of the experiment. Data for the other 3 fields of view is represented in ESI,† Fig. S1; planar projection time-lapse videos and renderings of nuclear tracking are presented in ESI† Videos 1–5. The videos qualitatively demonstrate rotation and translation of individual cohorts at various time-points; individual cohorts within a field of view may or may not be correlated to each other. To verify that motion is not caused by external ECM deformation, an experiment is performed using 1 μm fluorescent beads embedded alongside the cells in the collagen matrix (ESI,† Video 6). This video suggested that the matrix is relatively stable except for perturbations caused in the vicinity of cohorts, likely due to pulling of the gel. GFP-Ecad planar projections are in ESI,† Video 7, illustrating that E-cadherin is membrane bound in the cohorts.

The heterogeneity of cohorts in time is determined by studying the individual cell displacements for each cohort. For the 48 hour experiment, 1 hour displacements for all cells in a cohort are calculated; these displacements are squared ($|d^2|$), and the median of the resulting distribution is plotted corresponding to the left y-axis of Fig. 2. The upper and lower quartiles for the same distributions are plotted in the gray regions around the lines representing the median. The motility events isolated are depicted in shaded vertical strips in Fig. 2. Upon isolating motility events for each individual cohort, we find evidence of coherent rotation and translation within intervals ranging from 1 to 6 hours. For these 12 cohorts, a total of 61 motility events are obtained; five cohorts have 1–3 events, whereas those depicted in Fig. 2 are motile for almost the entire duration of the experiment. Displacements and events for ten other cohorts are displayed in ESI,† Fig. S2.

Order parameters provide a quantitative metric to measure the collectivity of systems. Establishing an order parameter for this system identifies the presence of translation, and also distinguishes between rotation and translation. For cells translating collectively, the order parameter is $\sim$1, and for cells rotating collectively this parameter is low, between 0–0.5. For a cohort rotating about an axis in the center of the cohort, the order parameter is 0. For cohorts rotating about an off-center or external axis the order parameter is higher. This is because when the axis of rotation is in the middle of the cohort there is an average velocity of 0 within the cohort.

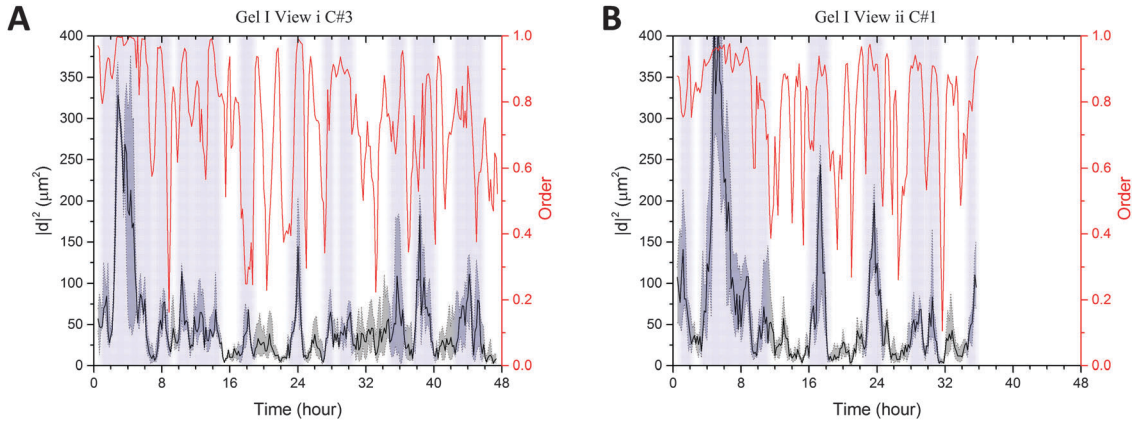

Fig. 2 Panel A has temporal analysis for Gel I View i Cohort #3, and Panel B has the same for Gel I View ii Cohort #1. The left *y*-axis corresponds to the 1 h squared displacement distribution: black line is the median, grey shaded regions are upper and lower quartiles. Vertical shaded regions represent motility events. The right *y*-axis and the red line correspond to 1 h order parameter of the cohort.

For the 48 h experiment, 1 hour order parameters are plotted corresponding to the right *y*-axis in Fig. 2. All motility events are identified *via* peaks in displacement; translation events are accompanied by peaks in the order parameter; rotation events are accompanied by fluctuations or valleys in the order parameter (Fig. 2 and Table 1). While translation occurs over durations of 1 to 6 hours, rotation only occurs in bursts of 1 to 2 hours (Fig. 2, 3 and Table 1). For selected motility events, metrics such as total displacement of the cohort, average order, and average number of cells are depicted in Table 1; metrics for all 61 events are in ESI,† tables.

The highest observed translation event has a duration of 6 hours and measures ~90 μm, corresponding to ~8 3D cell diameters. To visualize a few motility events, cell trajectories derived from raw data with events from four different cohorts are mapped in Fig. 3. Panels A and B show rotation, Panels C and D show translation. Spots are colored to indicate the initial and final time point of cell tracks within the event.

Directional correlation functions, defined in eqn (3), probe the inter-cellular dynamics within each cohort during a motility event. For the four representative events depicted in Fig. 3, leading and lagging times are calculated using pairwise correlation functions; normalized occurrence frequencies of these times are displayed in Fig. 4. For all cohorts in motility events, most cell pairs have 0 time lag between trajectories. For the two translating events in panels C and D, some cell-pairs have lags up to 200 minutes. For the shorter rotating events, the range of lag times is relatively smaller. There is no clear correlation between the magnitude of this time lag and the distance between cells.

Table 1 Parameters of four motility events. *C* is Cohort ID, T_i is the initial time, T_f is the final time, $|d|$ is the displacement of the cohort, $\langle\varphi\rangle$ is the average order, and N_c is the average number of cells in the cohort

Gel	View	*C*	T_i (h)	T_f (h)	$\lvert d\rvert$ (μm)	$\langle\varphi\rangle$	N_c
I	i	3	17:10	18:50	3.35	0.39	7
I	ii	1	3:20	9:30	89.06	0.92	10
II	i	1	1:30	5:50	53.09	0.84	19
I	ii	3	32:30	33:50	6.33	0.72	23

Discussion and conclusion

Conventional cell tracking research assumes that the behavior of cells in 2D and 3D is homogenous in time; information from all time points is averaged to deduce the timescales for various cellular behavior.[33] This is typically achieved by fitting cell trajectories to a stochastic random walk model[34] to evaluate speed and persistence of cells in 2D or 3D; however the basic assumptions of a persistent random walk fail in a system of cell collectives. Simplified versions of this analysis have been used on cell collectives in 3D to evaluate diffusion coefficients and angular velocity of human mammary cells,[11] however these techniques are inapplicable to data presented here, since the first assumption to evaluating Mean Squared Displacement *vs.* time-lag is temporal homogeneity. As evident in Fig. 2 and ESI,† Fig. S2, for cell cohorts in this study, temporal heterogeneity is observable in patterns of motion. These systems are not correlated within a single field of view over the duration of observation—there are intervals in which clusters move away (ESI,† Video 1, 4–7 s) and intervals in which the same clusters move toward each other and merge (ESI,† Video 1, 33–40 s). Not all cohorts are alike; some exhibit higher translation, rotation, and fluctuation than others. For example, Gel I View i Cohort #3 has fourteen high motility events of both translation and rotation, while Gel II View ii Cohort #1 has only one motility event of translation (Fig. 2, and ESI,† tables). Cells in 3D are smaller than their 2D counterparts – our average cohort diameter is on the order of 2D single cell lengths,[35] but in 3D, it spans ~6–15 cells (Fig. 1, and ESI,† Fig. S1). Thus, displacements on the order of tens of microns, which would not be relevant for 2D studies, mark coherent collective motility in these 3D studies (Fig. 1 and 2).

Cell collective studies of epithelial monolayers typically calculate the velocity correlation length[13,36,37] of monolayers. This correlation length is the length at which a radius-dependent

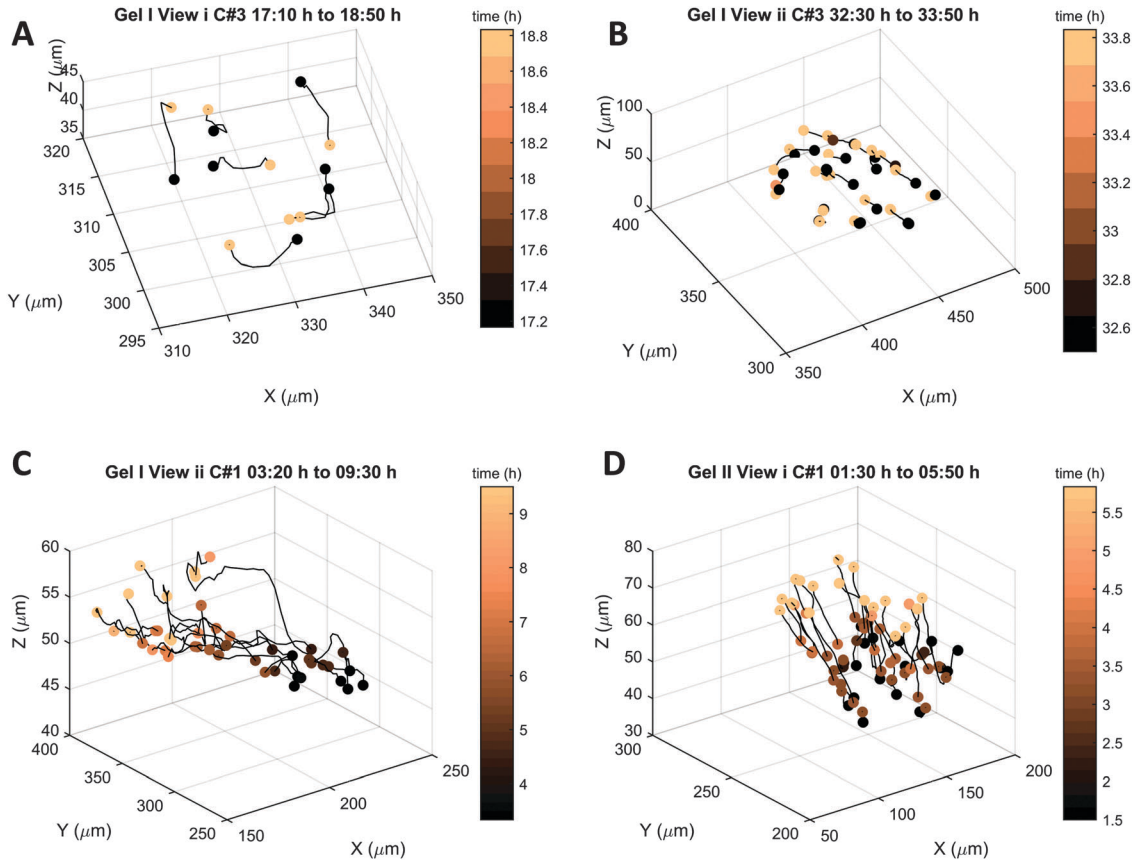

Fig. 3 Cell trajectories extracted from raw data for four different events and cohorts. Each cell track is represented by a single black line. Colors of spots mark initial and final time for each track. Panels A and B depict rotation, panels C and D depict translation.

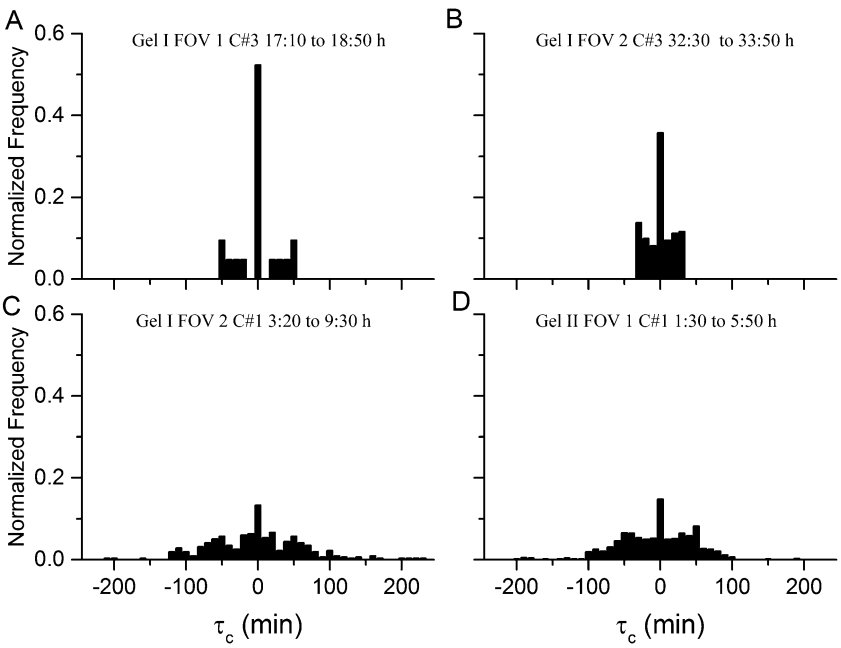

Fig. 4 For four events, these histograms show the normalized occurrence frequency of leading and lagging times τ_c. τ_c is obtained from pair-wise correlation functions, and it represents the time lag at which the function has a peak above 0.5, indicating the delay between cell trajectories.

velocity-correlation function equals zero on average for the monolayer. These analyses are effective at characterizing properties and differences between 2D cellular systems; however the underlying assumption is constant cellular density. Constant density

is neither feasible nor interesting for 3D-collectives, since it does not pertain to any known 3D collective motility modes and cannot account for the ECM. Therefore analysis techniques from the field of collective motion[31] are adapted here, specifically order parameters[1] and directional correlation functions, which have successfully been used to characterize heterogeneity in pigeon flocks.[32] Order parameters are easily adapted for this study—instead of calculating the order for the entire field of view, order parameters can be calculated on a cohort-by-cohort basis. The larger trajectory lags between cell-pairs that occur within a translating cohort as observed in our system may represent information transfer and polarity along the cohort. The order parameter of cohorts is seldom lower than 0.2, and the quartiles follow the same trends as the median (Fig. 2); thus there is a tendency for cells to stay within cohorts instead of splitting apart and invading the matrix. These results imply that cell–cell junctions contribute largely to cohort integrity and function. Indeed, GFP E-cadherin MDCK cell cohort experiments displayed E-cadherin localized at cell boundaries in 3D (ESI,† Video 6).

Many modes of collective cell motility have been observed in the presence of an external driving agent.[4] The results presented here suggest that the system of cell–matrix interactions is complicated and diverse enough to drive collective motion. Emergent motility events arise in the absence of external or forced driving agents, and are stochastic, as in the case of the two clusters that merged (Fig. 1C and D). There are examples of other similarly sized clusters that do not merge and in fact move in opposite directions (ESI,† Fig. S1E and F). The transient nature of these events suggests that this system displays stochasticity and plasticity, both suspected to occur in cancer pathologies.[38] Our setup provides a model system that allows for characterization of inter-cohort and intra-cohort dynamics as well as identification and analysis of emergent motility events. The techniques presented here could be applied to cancer explants, which are known to show coordinated collective motion *in vitro*.[5,6] As opposed to visually searching for motility, our work presents quantitative algorithms to isolate, observe, and characterize it. Since the methods operate on a cohort-by-cohort basis, they can be applied to a large amount of data and automated to extract motility events and compare different cellular cohorts.

Considering the balance of adherence and motility required for collective motion, translation over a few cell lengths in the absence of an external agent provides a promising model for the study of emergent phenomena and collective dynamics. Our results suggest that in the absence of external driving agents, interactions between cohorts and a collagen matrix are sufficient to drive collective cell motility. We show, for the first time, that cells spontaneously rotate in short bursts and translate for several hours; our analyses lay the foundation for quantitatively identifying supracellular polarity. The short bursts of rotation and comparatively larger spans of translation suggest that an internal stimulus arises within the dynamic cell–matrix system that attempts to drive collective translation. This work presents a quantitative approach to 3D cell collectives that have dynamic spatiotemporal heterogeneity – each cellular cohort is unique, and the algorithm finds motility events on a cohort-by-cohort basis. We built our custom algorithm using empirical data; however, it can be used for other cell types and experimental set-ups in order to probe questions of 3D collective mechanics, function, and efficiency. Our approaches can be expanded to study a range of phenomena in 3D, including collective cancer migration, density-dependent phase transitions, cell jamming, and emergent systems.

Acknowledgements

The authors would like to thank Dr Jeffrey Fredberg for comments on a previous draft of this paper. MDCK GFP-Ecad cell lines were a gift from Dr James Nelson at Stanford. This work was supported by the National Science Foundation (DMR-1206335 and DMR-1310266), the National Institute of Health (BRP 1R01 HL107561-01, P01HL120839), the Harvard Materials Research Science and Engineering Center (DMR-1420570), and Boston University's XTNC fellowship of the NIH under Award Number NIH R25 CA153955. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

References

1 E. Mehes and T. Vicsek, Collective motion of cells: from experiments to models, 2014, 21, *arXiv Prepr. arXiv1403.1127*. DOI: 10.1039/C4IB00115J.
2 P. Rørth, Collective cell migration, *Annu. Rev. Cell Dev. Biol.*, 2009, **25**, 407–429.
3 O. Ilina and P. Friedl, Mechanisms of collective cell migration at a glance, *J. Cell Sci.*, 2009, **122**, 3203–3208.
4 A. Haeger, K. Wolf, M. M. Zegers and P. Friedl, Collective cell migration: guidance principles and hierarchies, *Trends Cell Biol.*, 2015, **25**, 556–566.
5 P. Friedl and D. Gilmour, Collective cell migration in morphogenesis, regeneration and cancer, *Nat. Rev. Mol. Cell Biol.*, 2009, **10**, 445–457.
6 P. Friedl, *et al.*, Migration of coordinated cell clusters in mesenchymal and epithelial cancer explants *in vitro*, *Cancer Res.*, 1995, **55**, 4557–4560.
7 C. J. Weijer, Collective cell migration in development, *J. Cell Sci.*, 2009, **122**, 3215–3223.
8 P. Rørth, Collective guidance of collective cell migration, *Trends Cell Biol.*, 2007, **17**, 575–579.
9 D. Cai, *et al.*, Mechanical feedback through E-cadherin promotes direction sensing during collective cell migration, *Cell*, 2014, **157**, 1146–1159.
10 S. L. Haigo and D. Bilder, Global tissue revolutions in a morphogenetic movement controlling elongation, *Science*, 2011, **331**, 1071–1074.
11 K. Tanner, H. Mori, R. Mroue, A. Bruni-Cardoso and M. J. Bissell, Coherent angular motion in the establishment of multicellular architecture of glandular tissues, *Proc. Natl. Acad. Sci. U. S. A.*, 2012, **109**, 1973–1978.

12 P. Szabó, M. Nagy and T. Vicsek, Transitions in a self-propelled-particles model with coupling of accelerations, *Phys. Rev. E: Stat., Nonlinear, Soft Matter Phys.*, 2009, **79**, 021908.
13 K. Doxzen, *et al.*, Guidance of collective cell migration by substrate geometry, *Integr. Biol.*, 2013, **5**, 1026–1035.
14 A. Brugués, *et al.*, Forces driving epithelial wound healing, *Nat. Phys.*, 2014, **10**, 683–690.
15 T. J. Shaw and P. Martin, Wound repair at a glance, *J. Cell Sci.*, 2009, **122**, 3209–3213.
16 E. Bazellières, *et al.*, Control of cell–cell forces and collective cell dynamics by the intercellular adhesome, *Nat. Cell Biol.*, 2015, **17**, 409–420.
17 R. Riahi, *et al.*, Notch1-Dll4 signalling and mechanical force regulate leader cell formation during collective cell migration, *Nat. Commun.*, 2015, **6**, 6556.
18 N. Yamaguchi, T. Mizutani, K. Kawabata and H. Haga, Leader cells regulate collective cell migration *via* Rac activation in the downstream signaling of integrin β1 and PI3K, *Sci. Rep.*, 2015, **5**, 7656.
19 T. S. Deisboeck and I. D. Couzin, Collective behavior in cancer cell populations, *BioEssays*, 2009, **31**, 190–197.
20 J.-M. Hou, *et al.*, Circulating tumor cells as a window on metastasis biology in lung cancer, *Am. J. Pathol.*, 2011, **178**, 989–996.
21 A. Carlsson, *et al.*, Circulating Tumor Microemboli Diagnostics for Patients with Non–Small-Cell Lung Cancer, *J. Thorac. Oncol.*, 2014, **9**, 1111–1119.
22 A. A. Khalil and P. Friedl, Determinants of leader cells in collective cell migration, *Integr. Biol.*, 2010, **2**, 568–574.
23 C. M. Kraning-Rush, S. P. Carey, J. P. Califano, B. N. Smith and C. a. Reinhart-King, The role of the cytoskeleton in cellular force generation in 2D and 3D environments, *Phys. Biol.*, 2011, **8**, 015009.
24 A. Haeger, M. Krause, K. Wolf and P. Friedl, Cell jamming: collective invasion of mesenchymal tumor cells imposed by tissue confinement, *Biochim. Biophys. Acta*, 2014, **1840**, 2386–2395.
25 P. Friedl and S. Alexander, Cancer invasion and the microenvironment: plasticity and reciprocity, *Cell*, 2011, **147**, 992–1009.
26 J. L. Rapanan, K. E. Cooper, K. J. Leyva and E. E. Hull, Collective cell migration of primary zebrafish keratocytes, *Exp. Cell Res.*, 2014, **326**, 155–165.
27 C. L. Adams, Mechanisms of Epithelial Cell-Cell Adhesion and Cell Compaction Revealed by High-resolution Tracking of E-Cadherin-Green Fluorescent Protein, *J. Cell Biol.*, 1998, **142**, 1105–1119.
28 W. S. Leong, *et al.*, Thickness sensing of hMSCs on collagen gel directs stem cell fate, *Biochem. Biophys. Res. Commun.*, 2010, **401**, 287–292.
29 Y. Gao and M. L. Kilfoil, Accurate detection and complete tracking of large populations of features in three dimensions, *Opt. Express*, 2009, **17**, 4685–4704.
30 V. Pelletier, N. Gal, P. Fournier and M. L. Kilfoil, Microrheology of Microtubule Solutions and Actin-Microtubule Composite Networks, *Phys. Rev. Lett.*, 2009, **102**, 188303.
31 T. Vicsek and A. Zafeiris, Collective motion, *Phys. Rep.*, 2012, **517**, 71–140.
32 M. Nagy, Z. Akos, D. Biro and T. Vicsek, Hierarchical group dynamics in pigeon flocks, *Nature*, 2010, **464**, 890–893.
33 M. H. Zaman, *et al.*, Migration of tumor cells in 3D matrices is governed by matrix stiffness along with cell-matrix adhesion and proteolysis, *Proc. Natl. Acad. Sci. U. S. A.*, 2006, **103**, 10889–10894.
34 R. B. Dickinson and R. T. Tranquillo, Optimal estimation of cell movement indices from the statistical analysis of cell tracking data, *AIChE J.*, 1993, **39**, 1995–2010.
35 N. Sepúlveda, *et al.*, Collective Cell Motion in an Epithelial Sheet Can Be Quantitatively Described by a Stochastic Interacting Particle Model, *PLoS Comput. Biol.*, 2013, **9**, e1002944.
36 D. T. Tambe, *et al.*, Collective cell guidance by cooperative intercellular forces, *Nat. Mater.*, 2011, **10**, 469–475.
37 T. Das, *et al.*, A molecular mechanotransduction pathway regulates collective migration of epithelial cells, *Nat. Cell Biol.*, 2015, **17**, 276–287.
38 P. Friedl, Y. Hegerfeldt and M. Tusch, Collective cell migration in morphogenesis and cancer, *Int. J. Dev. Biol.*, 2004, **48**, 441–449.