Electronic Supplementary Material (ESI) for Soft Matter.

Supporting Information

1. Generation and structural properties of the random network model

The random network models used in the simulations are generated based on the maximally random jammed (MRJ) packing configurations, which are obtained using the Lubachevsky-Stillinger algorithm [1]. In particular, starting from a low-density configuration of small non-overlapping spheres randomly distributed in the simulation box, each sphere is given a random velocity. The particles follow Newton's law of motion, colliding with each other as they move and each particle also grows (i.e., isotopically expand) with a prescribed growth rate. The simulated is terminated once a jammed packing of the spheres, characterized by a diverging collision rate, is obtained (see **Fig. S1**).

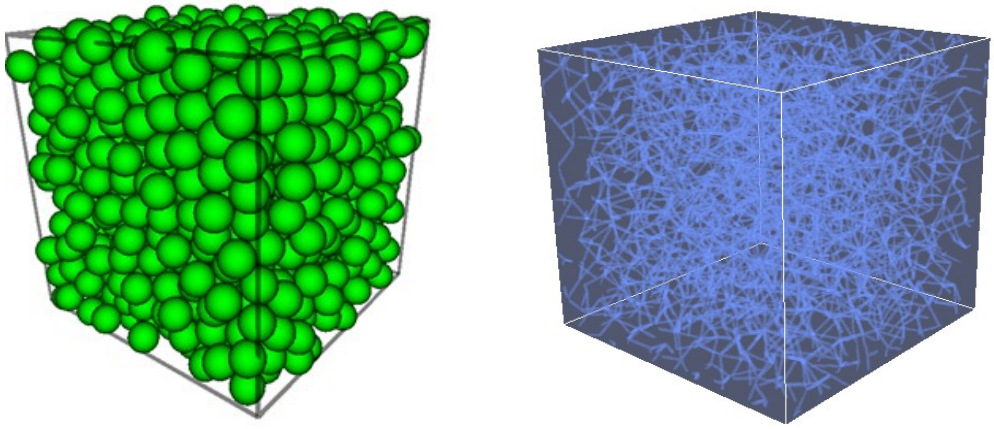


Fig. S1: Left panel: A typical maximally randomly jammed packing configuration of equal-sized spheres generated using the Lubachevsky-Stillinger algorithm. Right panel: A random network model can be derived from this packing by connecting the spheres touching each other.

Once the MRJ packings are obtained, the sphere centers are mapped to the nodes of a network, and each pair of the contacting spheres are connected by a bond. This leads to an isostatic random network with an average coordination number $Z \approx 6$. **Fig. S2** shows the structural statistics of random networks derived from MRJ packings, including the distribution of coordination number (left panel), and the distribution of bond length (right panel). It can be seen that the bond lengths are very uniform and the overall network is statistically isotropic.

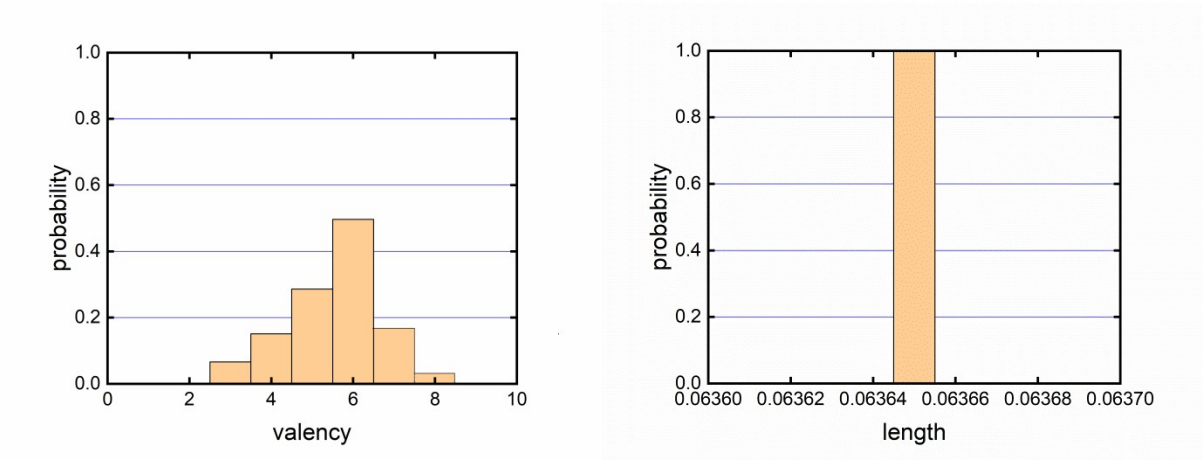


Fig. S2: Structural statistics of random networks derived from MRJ packings. Left panel: distribution of coordination number. Right panel: distribution of bond length. It can be seen that in this case, the bond lengths fall into a very narrow bin, indicating very uniform distribution of bond length. The bond length is measured with respect to the edge length of the cubic simulation box.

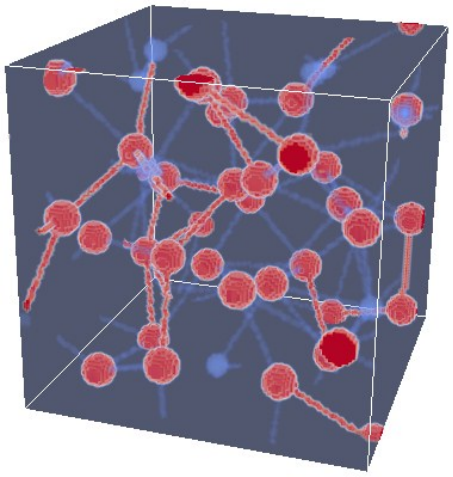


Fig. S3: Force network (carried by the high-stress fibers highlighted using red color) generated by contractile particles (shown as red spheres) on the network. For better visualization, only a small sub-network is shown here.

Figure S3 shows an example of the force network (i.e., carried by the high stress fibers highlighted using red color) generated by contractile particles (shown as red spheres) on the network. Based on the APN dynamics described in the text in detail, the particles will self-organize regulated by the force network. On the other hand, the re-organization of the particles leads to re-configuration of the force network. This feedback loop determines the overall collective dynamics of the system.

2. Evolution of clusters and radical statistics

In the main text, we only show schematically the particle cluster configurations, and radial statistics for three selected representative particle number densities, i.e., $\rho = 0.05$, $\rho = 0.10$, $\rho = 0.15$, which are respectively below, close to and above the critical density $\rho_c = 0.114$. Here, we show the results for other cell number densities.

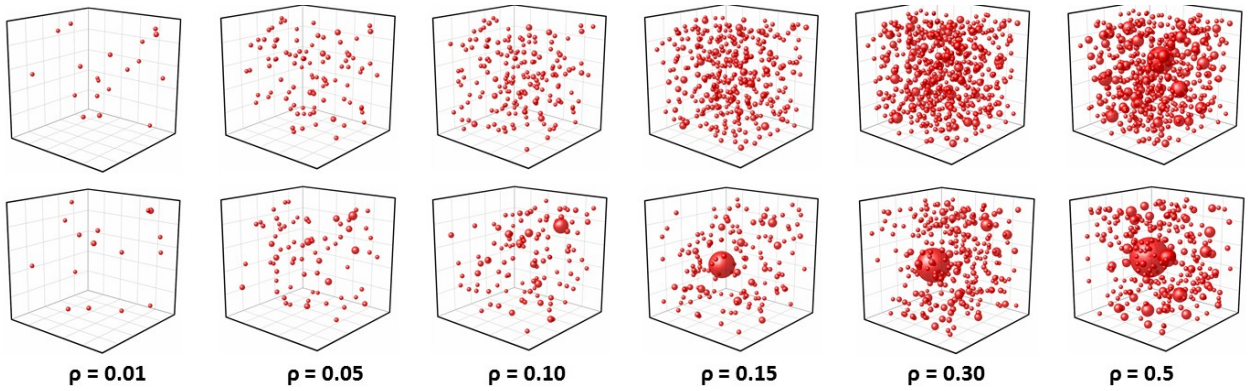


Fig. S4: Evolution of cluster configurations in the APN system as the particle number density increases. In these plots, a cluster is represented by a sphere for better visualization with the center coinciding with the center of the cluster and the radius representing the cluster size.

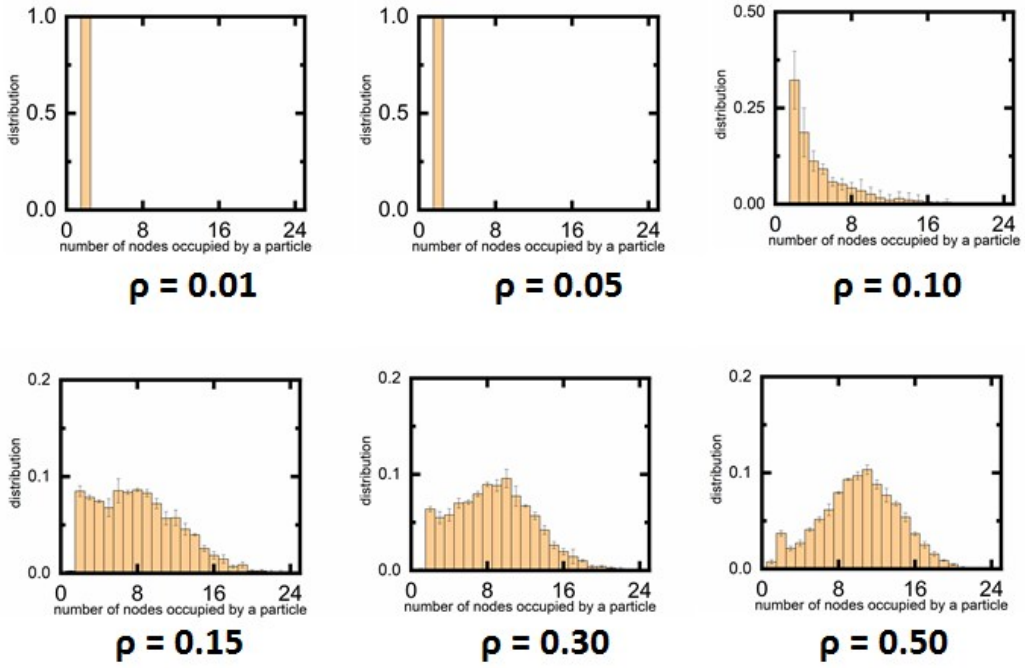


Fig. S5: Statistics of the number of distinct nodes visited by a particle for s = 24 successive steps for different number densities ρ. As ρ approaches ρ_c from below, a subset of highly dynamic particles emerge which are able to visit many distinct nodes for a given number of steps and are referred to as "radicals".

Figures S4 and **S5** respectively show the evolution of cluster configurations and radical statistics as number density increases. In the case of cluster configurations, we can clearly see that below ρc, the particles only aggregate into isolated small stationary clusters. As ρ increases beyond ρc, the majority of particles aggregate into a large dynamic cluster. Similarly, for the radical statistics, a subset of highly dynamic particles emerge which are able to visit many distinct nodes for a given number of steps. Such radicals dominates the re-organization of the dynamic clusters.

3. Percolation probability analysis

The percolation threshold associated with the influence sphere is determined using a percolation probability analysis. In particular, for a given particle number density $\rho = N_p/M_n$ (where N_p is the number of particles and M_n is the total number of available nodes in the network), a total of X_C = 100 independent configurations are generated. Each configuration is generated by randomly selecting Np nodes from the given network and placing an influence sphere with radius R_{in} centered at the selected node.

Next, we identify the largest cluster formed by the overlapping influence spheres and check whether this cluster spans the entire system. If a system spanning cluster is identified, the system is considered percolated. The total number of percolated systems for a given particle number density is then obtained, i.e., $X_p(\rho)$, and the percolation probability for the system at number density ρ is computed as: $P(\rho) = X_p(\rho)/X_C$. Following this procedure, the percolation probability P(ρ) for all ρ of interest can be computed (as shown in Fig. 4 of the main paper).

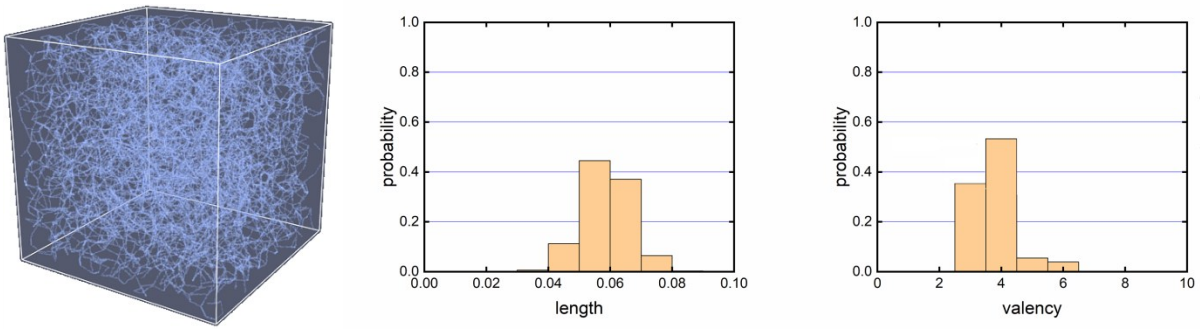
To estimate the percolation threshold ρ* for the system, we approximate the percolation probability P(ρ) with the following function [2]:

$$P(\rho) = \frac{1}{2}\left[1 + \tanh\left(\frac{\rho - \rho^*}{\Delta}\right)\right] \tag{A1}$$

where Δ is the effective "transition width", which decreases with system size. By fitting the numerical data for P(ρ) using Eq. (A1), we can obtain ρ*accordingly, i.e., ρ*=0.0672.

3. Collective dynamics of active particles on other networks

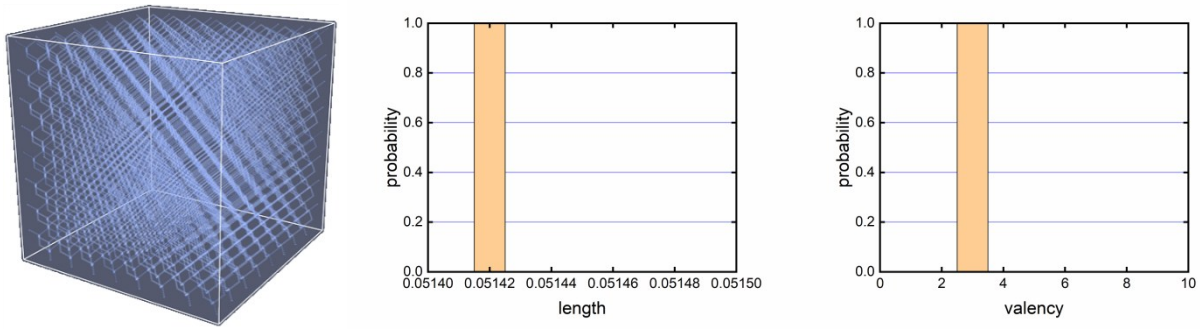
In the main text, we focused on the collective dynamics of active particle on model networks derived from MRJ packings of equal-sized spheres. In this section, we report the self-organizational behaviors of active particles on two different classes of networks, including networks derived from confocal images and derived from diamond lattice.



(a) **(b)** **(c)**

Fig. S6: (a) A realization of the reconstructed network, mimicking the 3D collagen network within a cubic volume of 100 μm by 100 μm by 100 μm. (b) Distribution of bond lengths, which are measured with respect to the edge length of the cubic simulation box. (c) Distribution of the coordination number.

The image-based random networks (IBRN) are obtained via stochastic reconstruction based on input statistics extracted from 3D confocal imaging data of 2mg/ml collagen gels, see Ref. [3] for details. The input statistics include collagen concentration (number density of nodes), fiber segment length (bond length) distribution, coordination number distribution and bond-angle distribution. **Figure S6a** shows a realization of the reconstructed network, mimicking the 3D collagen network within a cubic volume of 100 μm by 100 μm by 100 μm, containing ~ 5,000 nodes and ~10,000 bonds, with an average coordination number $Z \approx 4$. The same linear elastic model used in the main paper is used for the image-based random networks. The bond length distribution and coordination number distribution are respectively shown in **Fig. S6b and S6c**.



(a) **(b)** **(c)**

Fig. S7: (a) A realization of the diamond-lattice network, constructed as described in the text. (b) Distribution of bond lengths, which are measured with respect to the edge length of the cubic simulation box. (c) Distribution of the coordination number.

The diamond-lattice networks (DLN) are obtained by connecting the nearest neighbors of a diamond lattice and randomly perturbing the position of the nodes by a tiny amount. This allows us to avoid degenerate force distribution on the bond (i.e., more than more bonds connecting to

the node with a contracting particle possess the same tensile force) while maintain the overall symmetry and order of the lattice. **Figure S7a** shows a realization of the diamond-lattice network, which contains ~5,000 nodes and ~10,000 bonds and each node has the coordination number Z = 4. **Figure S7b and S7c** respectively show the bond length distribution and coordination number distribution. It can be seen that the bond lengths associated with the perturbed nodes still fall into a very narrow bin. The same linear elastic model used in the main paper is used for the image-based random networks.

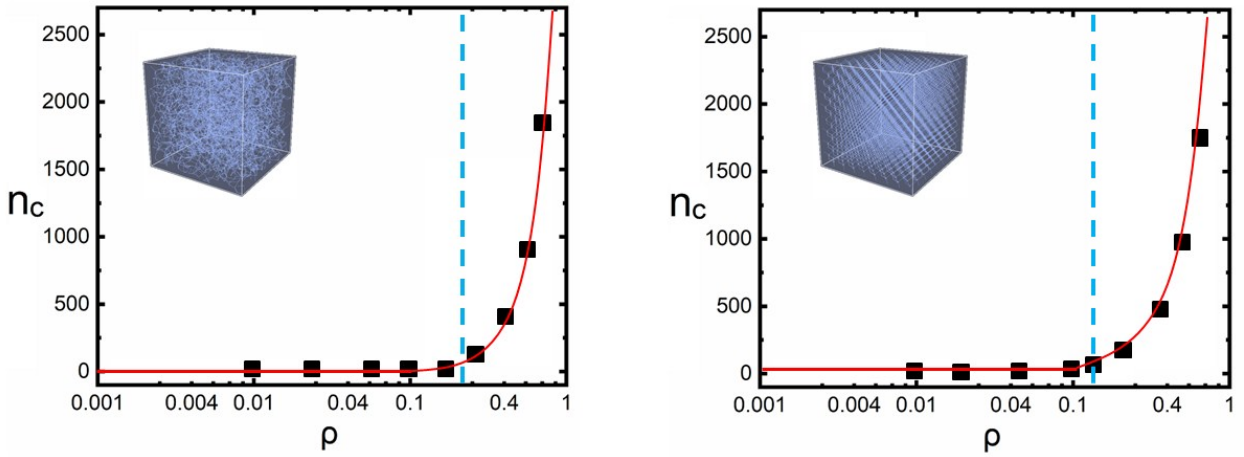


Fig. S8: The maximal cluster size n_c as a function of the particle number density ρ for the image-based random networks (IBRN) (left) and diamond-lattice networks (DLN) (right).

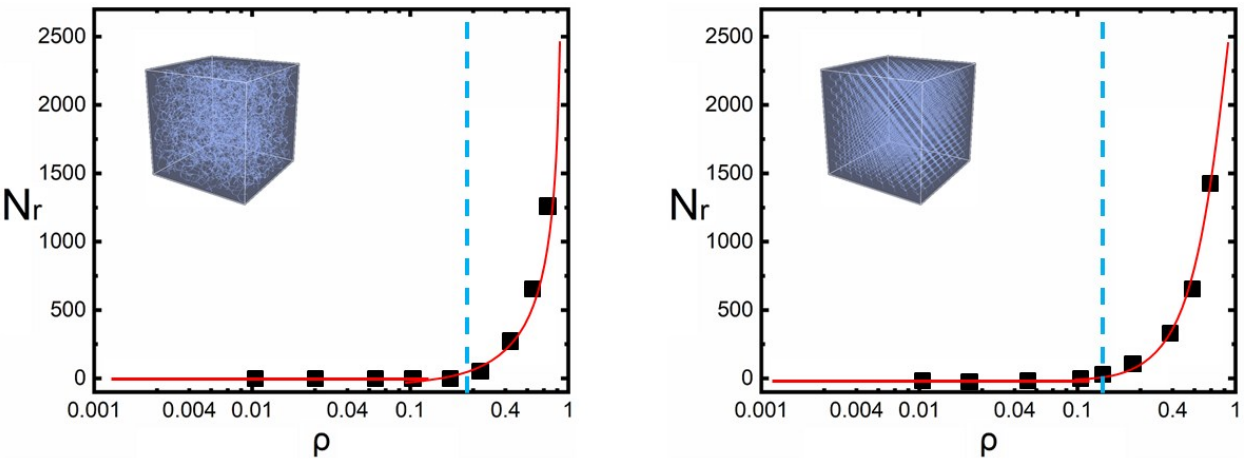


Fig. S9: The radical number N_r as a function of the particle number density ρ for the image-based random networks (IBRN) (left) and diamond-lattice networks (DLN) (right).

Once the network models are obtained, we use the same APN procedure described in the main text to investigate the collective dynamics and self-organization behaviors of active particles on these networks. **Fig. S8** shows the evolution of the maximal cluster size n_c as the particle number

density ρ increases for the image-based random networks (IBRN) (left panel) and diamond-lattice networks (DLN) (right panel). It can be seen that both systems exhibit clear transition behaviors as observed in the packing-based random networks described in the main text. In particular, the critical number density for IBRN is $\rho_c^{IBRN} \approx 0.28$ and $\rho_c^{DLN} \approx 0.12$. These values are also consistent with the estimates based on the statistics of number of radicals N_r as a function of ρ, as shown in **Fig. S9**. We also employ the mean-field theory and perform the percolation analysis for these two systems as described above and in the main text. The estimated critical densities (i.e., as the percolation density of the influence sphere) are respectively $\rho_*^{IBRN} \approx 0.306$ and $\rho_*^{DLN} \approx 0.14$, which are in relatively reasonable agreement with the numerical results.

These results, together with the results for the packing-based random network presented in the main text, indicate that the dynamic phase transition occurs on networks with different degrees of topological and geometrical order, and connectivity. However, the transition density ρ_c depends on the network microstructure. Interestingly, the IBRN system possesses a significantly higher ρ_c than the other two systems. One possible reason is that the IBRN possesses a wide distribution of bond lengths. In our APN model, for the same particle contraction, the shorter bonds carry larger tensile forces and thus, effective attract particles to the nodes they connect. This would stabilize the isolated clusters even at relatively higher density.

In general, the collective dynamics of the active particles are determined by the force network carried by the physical network due to particle contraction. The force network in turns depends on the distribution of particles in the system as well as the network microstructure. For networks with higher degree of geometrical and topological order (such as the packing-based and diamond-lattice networks), the particle distribution would play a more important role in determining the force network. On the other hand, for very disordered networks with large heterogeneities as such the image-based network, the local heterogeneities would lead to very non-uniform distribution of tensile forces on the bonds. In such cases, the network microstructure would also play an important role in determining the collective dynamics of the systems. We will investigate these problems in detail in our future work.

4. Experimental verification

To verify the model predictions, we also design *in vitro* experiments to investigate the collective dynamics of invasive MDA-MB-231 breast cancer cells in collagen gel. The experimental details and results are reported elsewhere [4] and we will only briefly describe the observations. Specifically, similar collective aggregation behaviors have also been observed in an in vitro system of invasive breast cancer cells (MDA-MB-231), initially co-cultured with non-invasive MCF-7 cells on a 2D substrate, which is then covered by a 100% Matrigel (see **Fig. S10A-C**). It was observed that the invasive tumor cells (MDA-MB-231) can migrate into the Matrigel, as they developed various self-organization patterns for different cell number densities. As shown in **Fig. S10D**, at small cell densities, the invasive tumor cells in the Matrigel form isolated clusters. As the cell number density increases (see **Fig. S10F**), eventually a single large cluster

emerge, which is consistent with the APN model predictions. Subsequent control experiments suggested that the significant effective pulling among the invasive cell generated due to cell contraction as they migrate into the Matrigel layer played an important role in giving rise to the observed collective patterns.

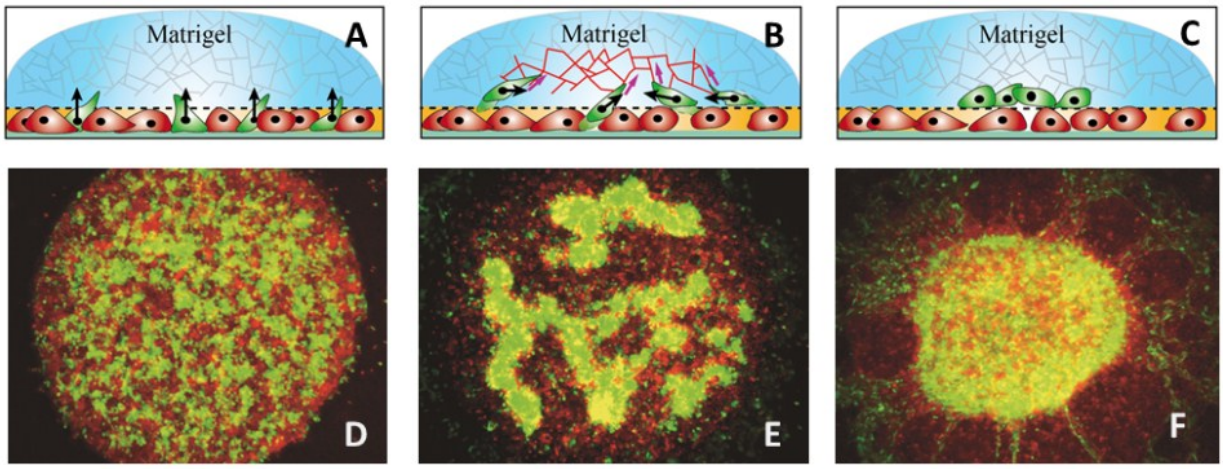


Fig. S10: Upper panels: Schematic illustration of the experimental set up and collective aggregation dynamics in the *in vitro* co-culture system. Low panels: Confocal images of the self-organization patterns in the system at different cell number densities. The cell densities increase from D to F. The size of the image is 1.2 mm by 1.5 mm.

References:

[1] B. D. Lubachevsky and F. H. Stillinger, J. Statistical Physics **60**, 561 (1990).

[2] M. D. Rintoul and S. Torquato, J. Phys. A: Math. Gen. **30**, 585 (1997).

[3] H. Nan, L. Liang, G. Chen, L. Liu, R. Liu, and Y. Jiao, Phys. Rev. E 97, 033311 (2018).

[4] X. Wang, S. Chen, F. Ye, Y. Jiao, and L. Liu, Abnormal and rapid aggregation of invasive breast cancer cells induced by collective polarization and ECM mediated forces. 2018. Submitted.