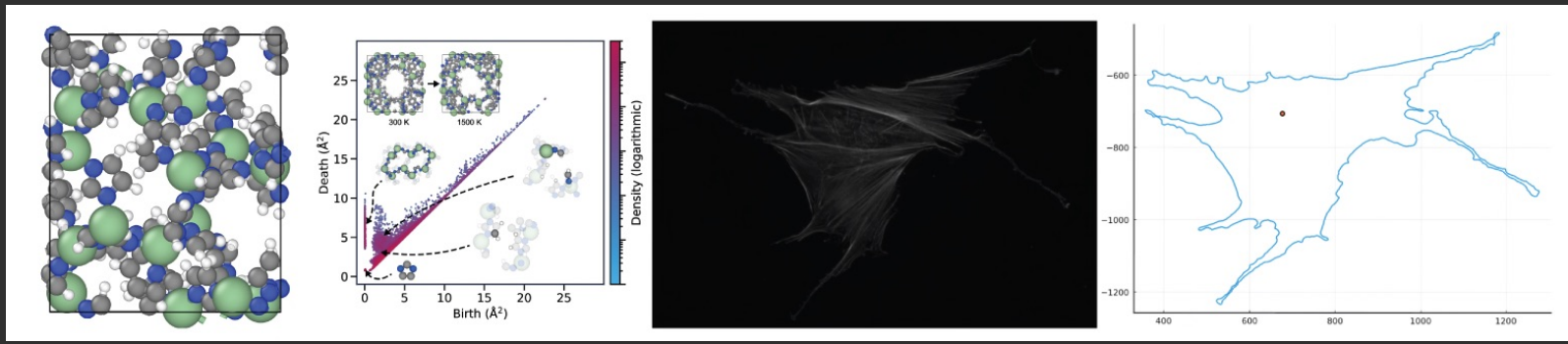
# Microstructure Analysis using Geometry and Topology




Yossi Bokor Bleile
Department of Mathematical Sciences
Aalborg University

Joint with:

1) Rasmus Christensen, Søren S. Sørensen, Anne-Marie Svane, Matteo Pegoraro, Christophe A. N. Biscio, Lisbeth Fajstrup, and Morten M. Smedskjaer
2) Florian Rehfeldt, Patrice Koehl, Stephan Tillmann

# Biology Crash Course

Human mesenchymal stem cells (hMSC) are multi potent stem cells, which can differentiate into various *lineages* of cells, including:
* osteocytes
* adipocytes
* chondrocytes
* neurocytes
* hepatocytes

Differentiation into lineages can be triggered by environmental stimuli such as the rigidity of the extracellular matrix they are cultured on.

**Question: can we distinguish between these lineages based on morphological features of a cell?**

# Oh no, problem!

The preferred method for collecting hMSCs is from bone marrow, and unfortunately, it is not possible to avoid the presence of other cells, such as bone barrow fibroblasts, in the sample.

We anticipate about 5% of non-hMSCs to be present in a given sample.

Hence, we need to *clean* our population of cells, to ensure we are only learning information about the morphology of hMSCs.

**Due to issues in the culturing and imagining process, we expect up to 7% of cells to display *abnormal* growth patterns.**

# New Aim

**Due to the the presence of the ~5-12% of cells which are not hMSCs/display abnormal morphology, we can not immediate identify morphological features to characterise lineages. So, we have a new problem.**

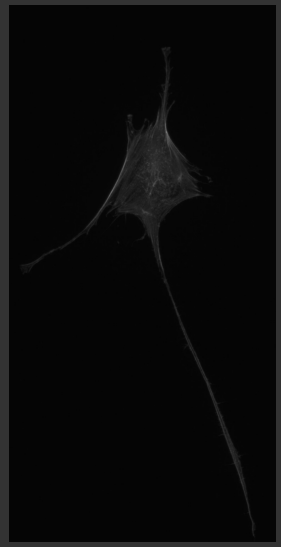
Question: can we identify the 2 sub-populations present in a sample?

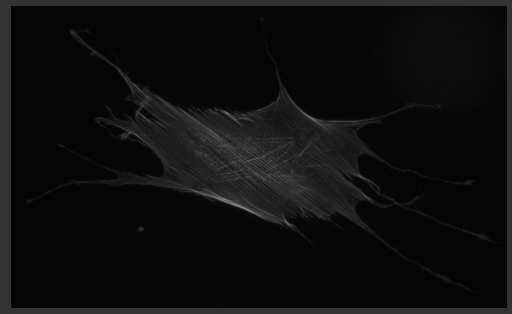
## Spoiler

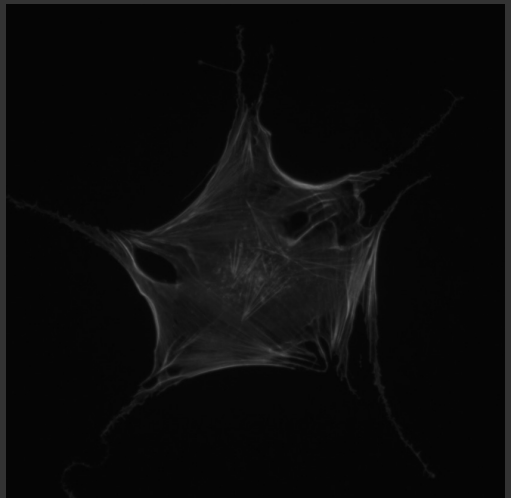
YES

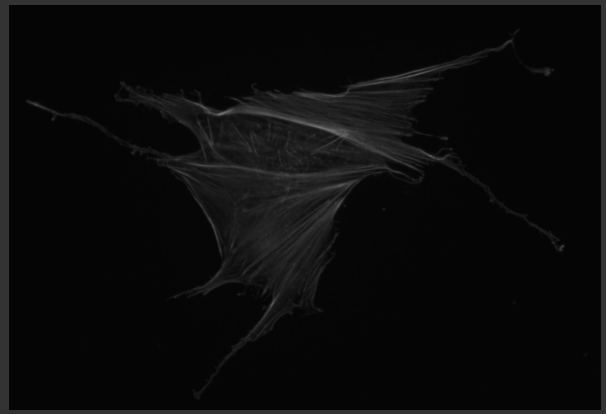
## Identifying sub-populations

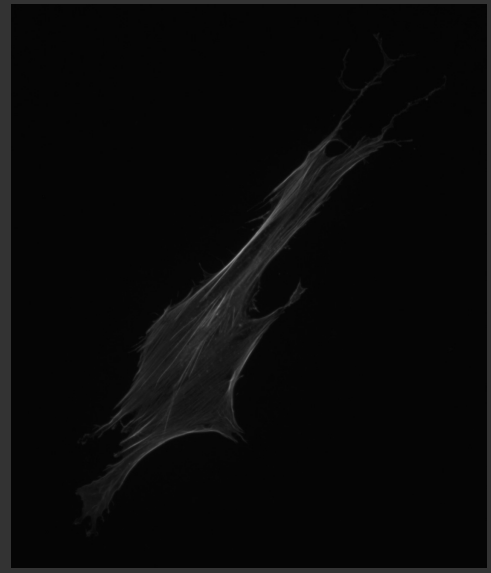
Problem set up:
Given a population of cells cultured in the same environmental conditions (extracellular matrix, treatment, time cultured, staining), can we identify the 5-10% of cells we expect to exhibit abnormal growth patterns?



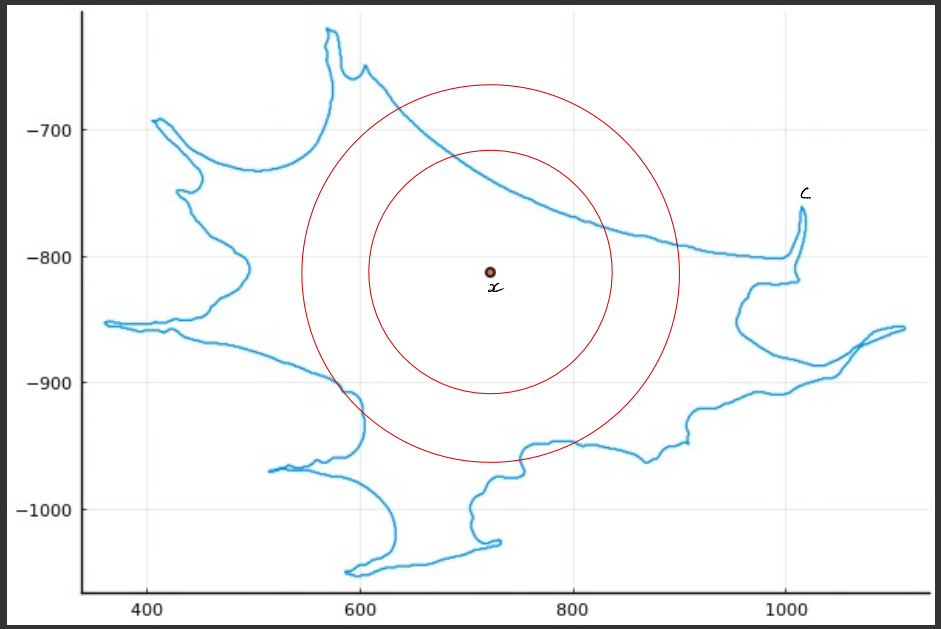








# Identifying subpopulations: feature extraction?



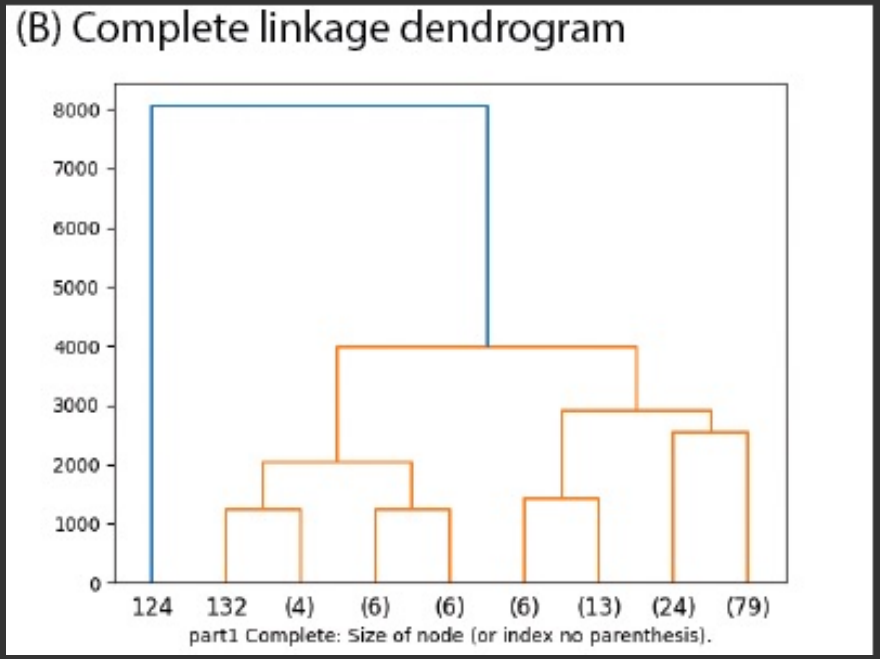

Let $x$ be the center of the nucleus and $f$ the distance to $x$ function.

We then take the persistent homology of $C$ using $f$, and pair the essential 0- and essential 1-cycles to obtain the 0th persistence diagram.

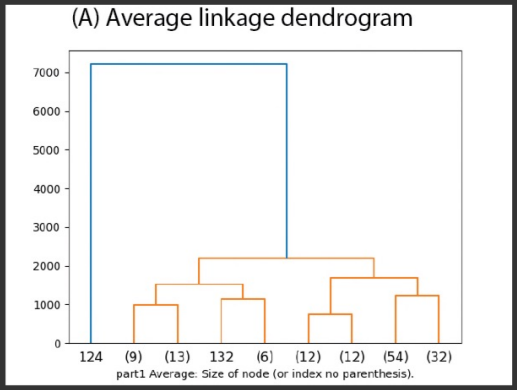
# Identifying subpopulations: clustering?

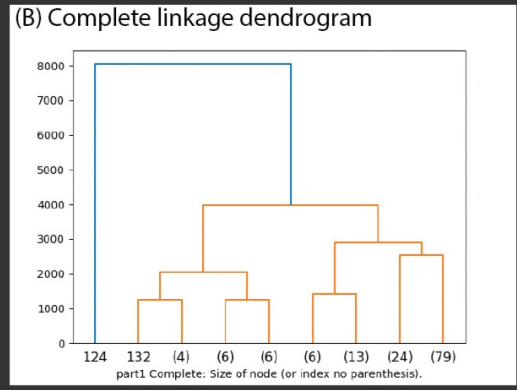
Comparing the contours of the cells in the population, we construct a dissimilarity matrix, and then perform hierarchical clustering with average, complete, single and ward linkage.

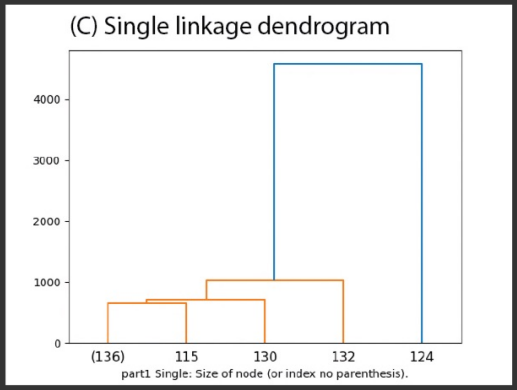
We use the 2-Wasserstein distance on persistence diagrams as our dissimilarity score between two cells.

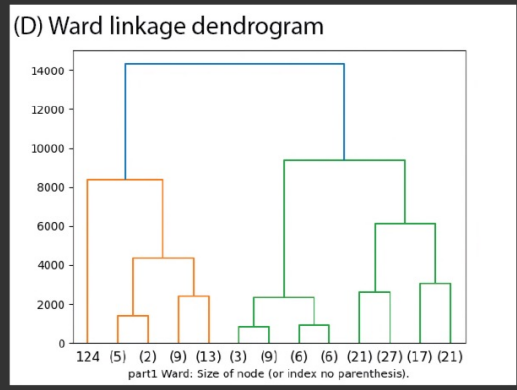


# Results - set 1

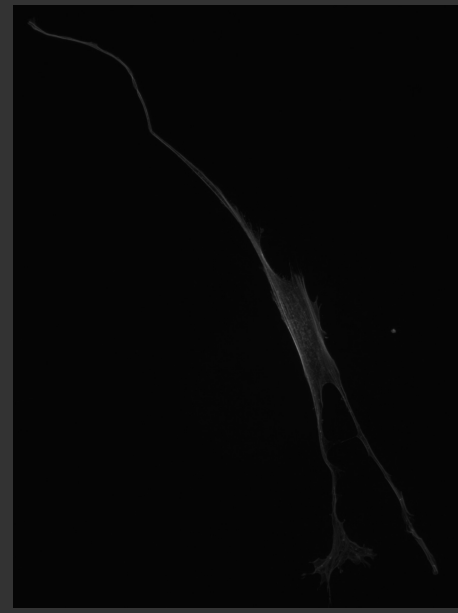


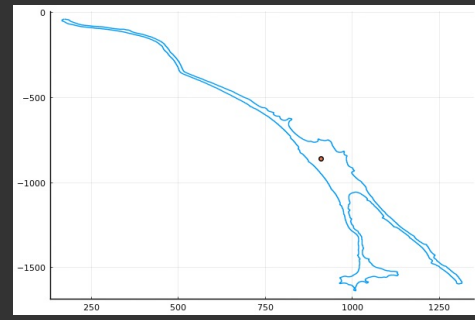



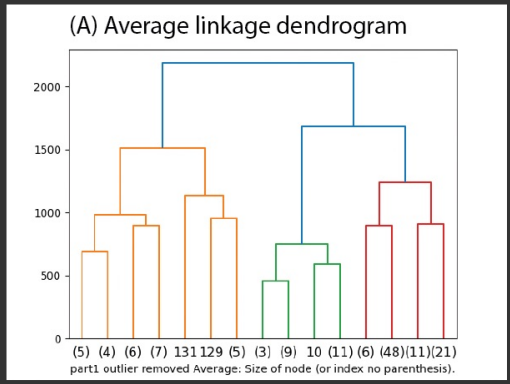



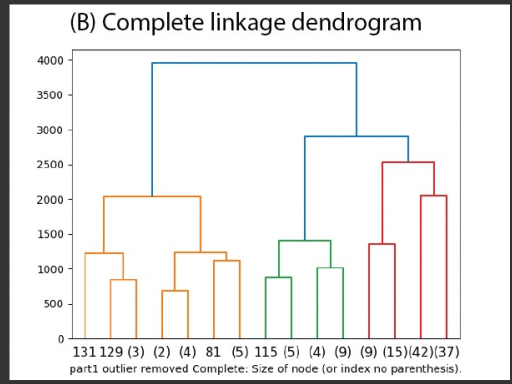

Outlier

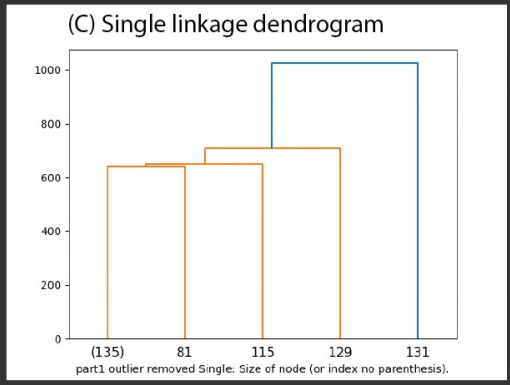


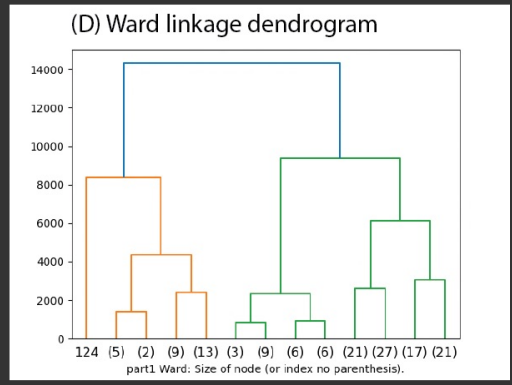


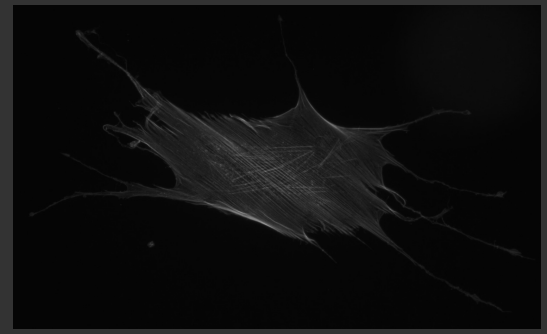
# Results - set 1 no outlier

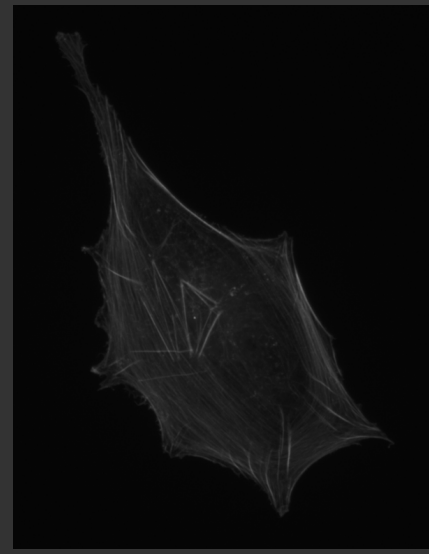


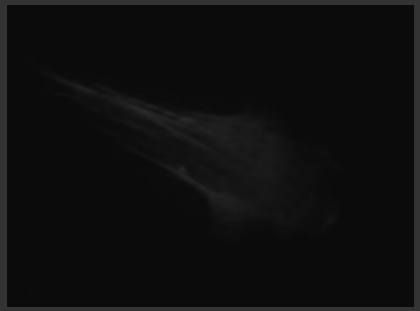
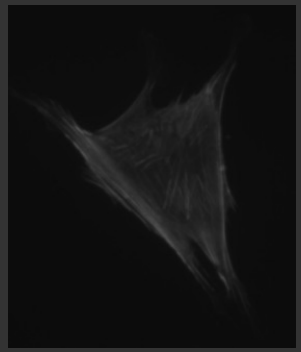
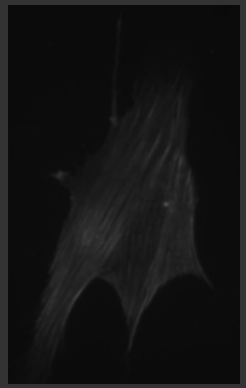
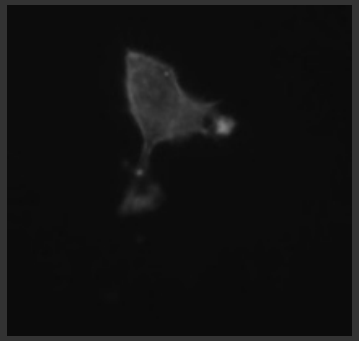
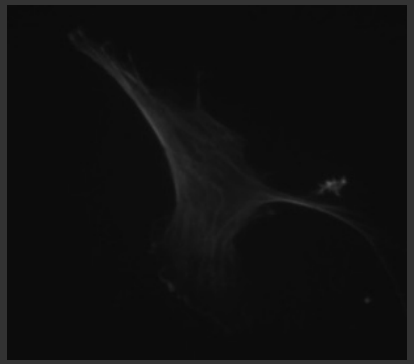
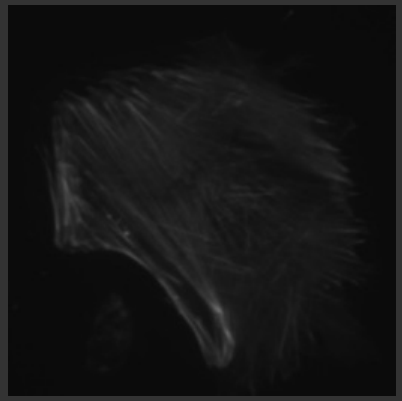





# Back to initial problem

Given populations of human mesenchymal cells cultured in the  different environmental conditions (extracellular matrix, treatment, time cultured, staining), can we classify their lineages based on topological and geometric information?

# Back to initial problem

The jury is still out, about the success of persistent homology to seperate these lineages.

Next steps are to use summaries including

* inscribed ellipse
* inscribing ellipse
* least square ellipse
* Willmore energy
* persistent homology (including rank functions and APF)
* conformal map classes

and the distances between

* distances between ? ellipse
* Willmore distance
* Frechet distance
* Wasserstein distance between persistence diagrams
* PCA on rank functions
* PCA on APF

to understand the space of lineages.

Persistence of radial function not as discerning as we would have hoped.

Potentially due to ‘global’ outliers.