

Automated characterization of cell tracks and classification by using their position coordinates and velocity components

Shariful Islam¹, Manuel Nietert¹, Michael Haberl², Alexander Flügel², Tim Friede¹, Tim Beißbarth¹

¹Dept. Medical Statistics; ²Dept. of Neuroimmunology, UM Göttingen; shariful@med.uni-goettingen.de

Introduction

- Cell motility is an important factor in many biological processes
- Cell motility is crucial during immunosurveillance and for mounting immunological response
- Different immune cell populations in different milieus may exhibit distinct migratory patterns that are dependent of their functional status

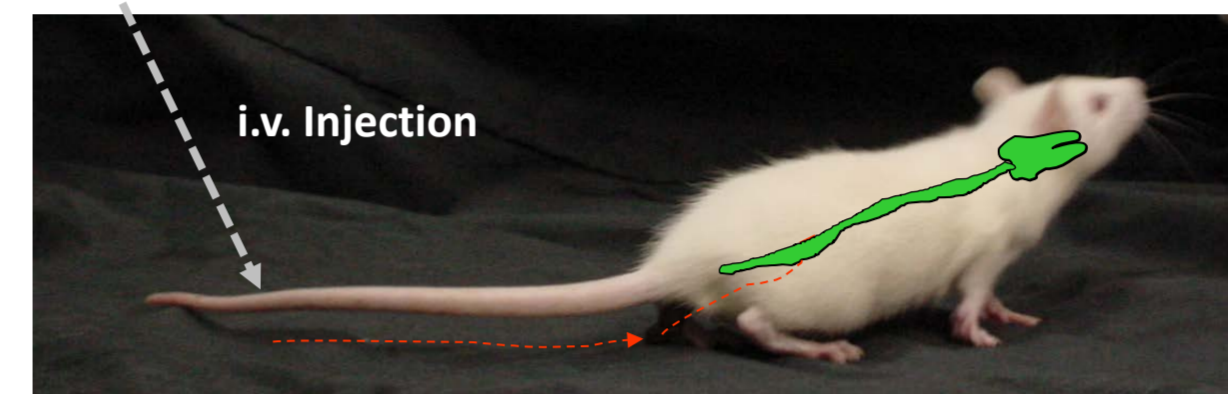
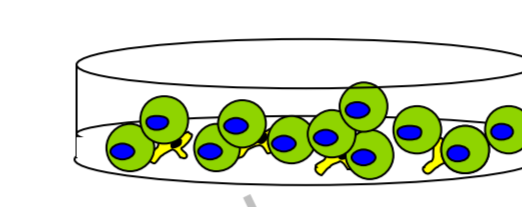
Aims

- Develop an unbiased computational approach to analyze these complex biological processes

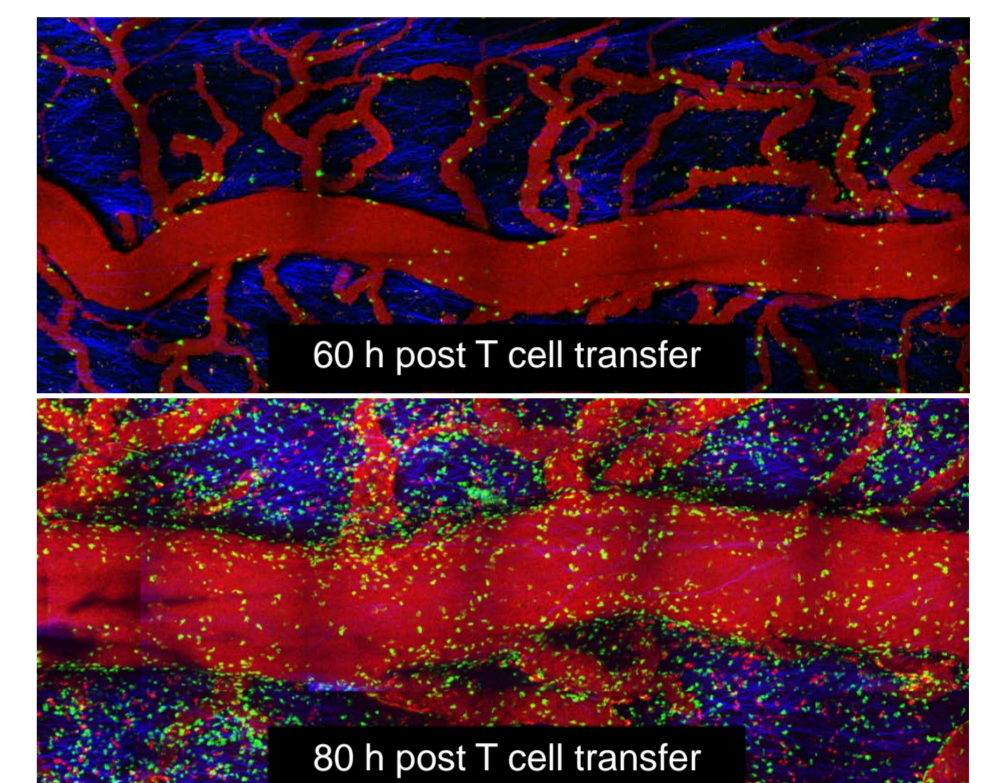
Biological Model

- Lewis rat model of Multiple Sclerosis: **E**xperimental **A**utoimmune **E**ncephalomyelitis (EAE) with adoptive transfer of encephalitogenic T cells (T_{MBP})
- Intravital 2-photon microscopy of leptomeningeal vessels of the spinal cord during different phases of EAE

T_{MBP} cell blasts

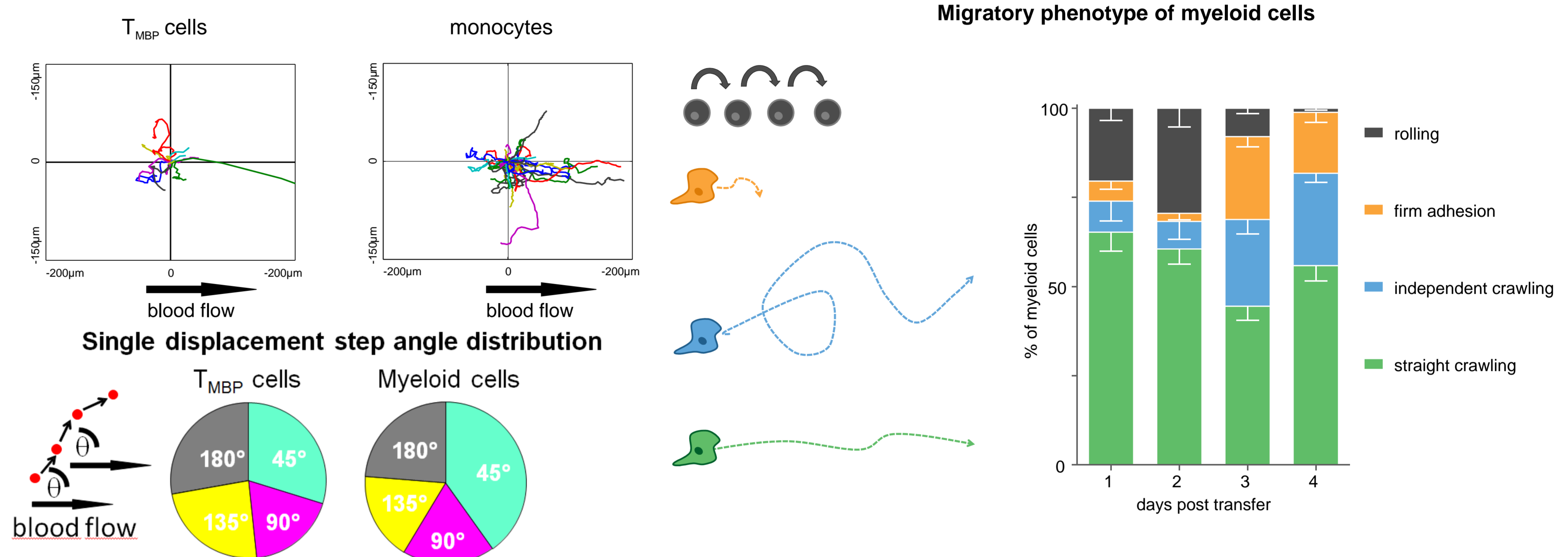
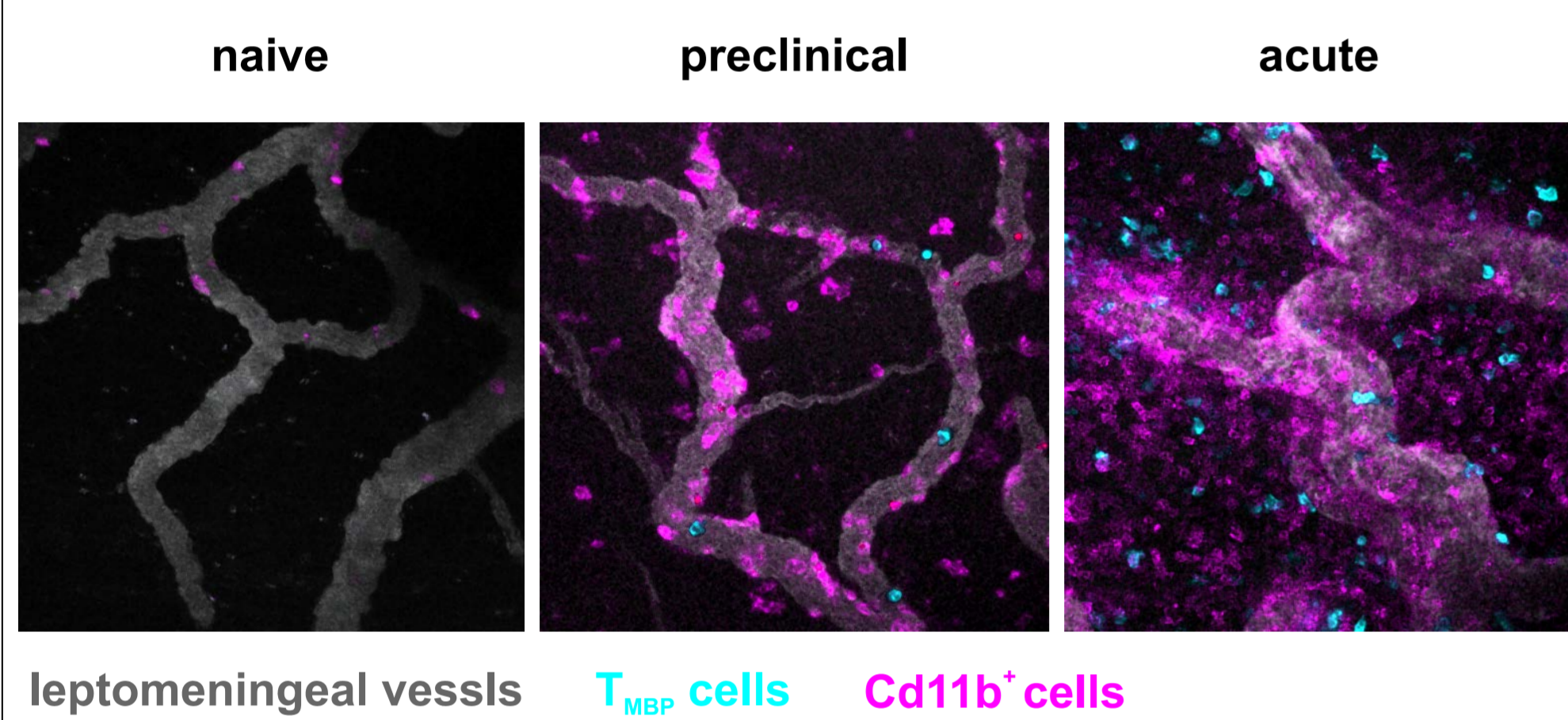


Transfer EAE



Myeloid Cells During EAE

- Myeloid cells accumulate over the course of EAE in leptomeningeal vessels
- They display distinct migratory phenotypes inside the blood vessels
- Change in migratory phenotypes over the course of EAE



Feature Selection

We computed several features as described in Mokhtari¹ et al., 2013:

$$\frac{|p_1 p_l|}{|p_1 p_2| + |p_2 p_3| + \dots + |p_{l-1} p_l|}, \frac{|p_l p_l|}{\max\{|p_1 p_2|, |p_2 p_3|, \dots, |p_{l-1} p_l|\}}, \frac{\max\{|p_1 p_2|, |p_2 p_3|, \dots, |p_{l-1} p_l|\}}{|p_1 p_l| + |p_2 p_3| + \dots + |p_{l-1} p_l|}$$

Volume asphericity $A_d(n, m)$ is given by

$$A_d(n, m) = \frac{1}{d(d-1) \langle R_l^2 \rangle_d^2} \sum_{k=1}^d (R_k^2 - \langle R_l^2 \rangle_d)^2$$

where the average value $\langle R_l^2 \rangle_d = \frac{1}{d} \sum_{l=1}^d R_l^2$

Our data also contain velocity components v_x , v_y , and v_z of each cell track at each time point. So we can express the velocity V of the cell by $\vec{V} = v_x \vec{i} + v_y \vec{j} + v_z \vec{k}$ (see Fig. 2)

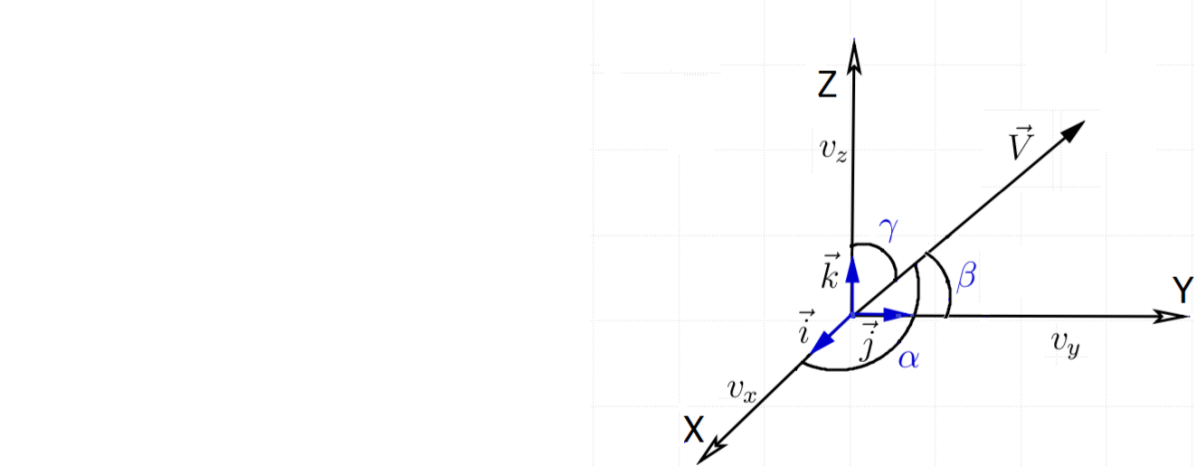


Figure 2. Velocity components of V and their direction cosines which yields

$$\cos \alpha = \frac{\vec{V} \cdot \vec{i}}{\|\vec{V}\|} = \frac{v_x}{\sqrt{v_x^2 + v_y^2 + v_z^2}}$$

$$\cos \beta = \frac{\vec{V} \cdot \vec{j}}{\|\vec{V}\|} = \frac{v_y}{\sqrt{v_x^2 + v_y^2 + v_z^2}}$$

$$\cos \gamma = \frac{\vec{V} \cdot \vec{k}}{\|\vec{V}\|} = \frac{v_z}{\sqrt{v_x^2 + v_y^2 + v_z^2}}$$

Intuitively, we expect the standard deviations of these angles to be very different for 'straight' and 'crawling' cell tracks (see Fig. 3).

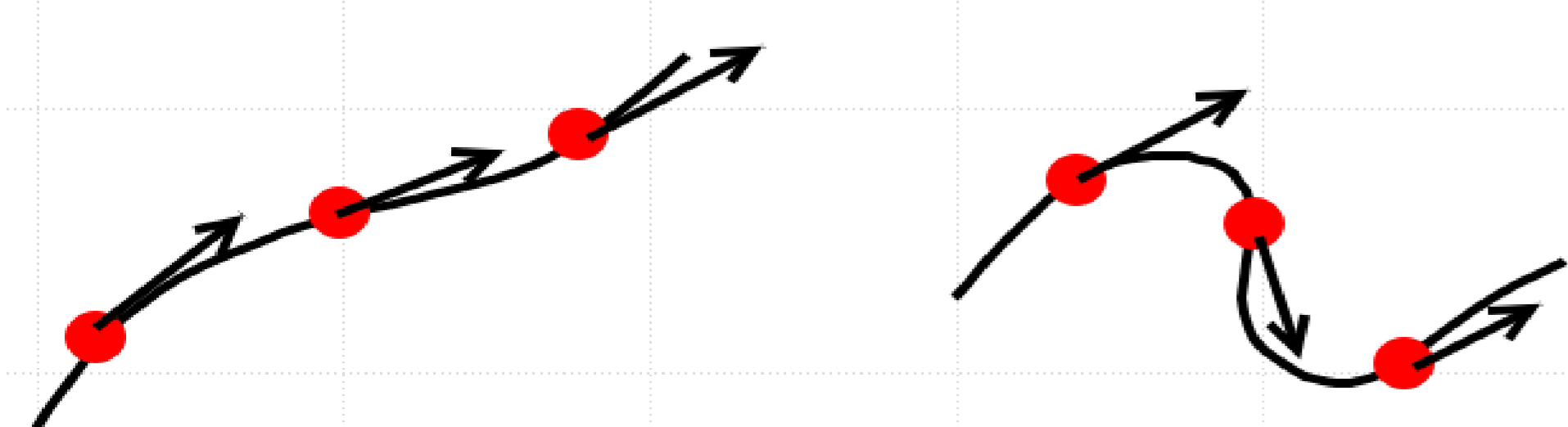


Figure 3. Variation in angles for different types of cell tracks

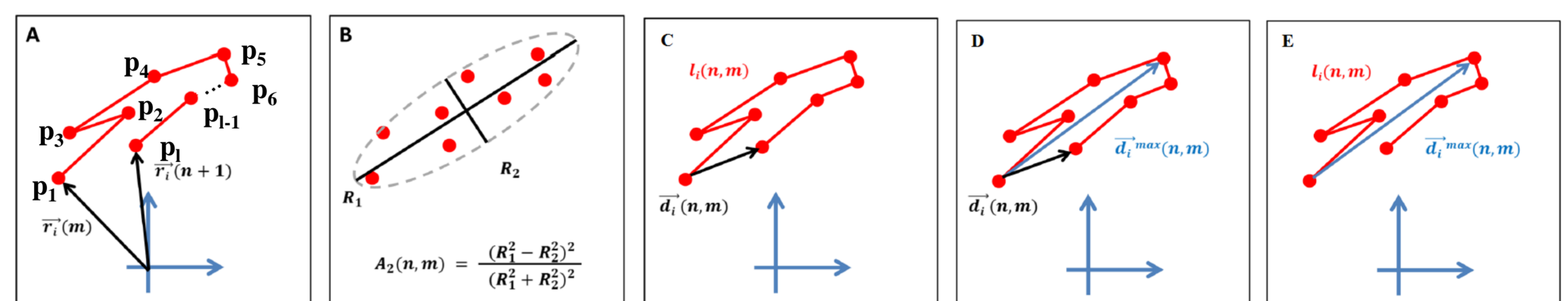
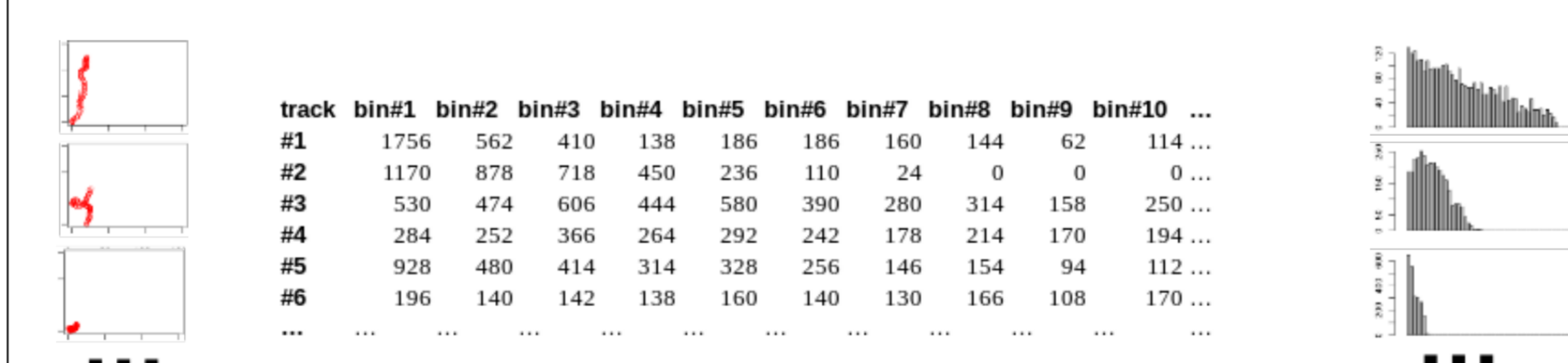


Figure 1. A: Example of a cell track segment. B: The **volume asphericity** is determined by the cell positions of the track segment that are viewed as uncorrelated data points. C: The **confinement ratio** is determined by the displacement over the length of the cell track segment. D: The **displacement ratio** is determined by the displacement over the maximal displacement of the cell track segment. E: The **outreach ratio** is determined by the maximal displacement over the length of the cell track segment.

Statistical Methods and Results

Translate Coordinates to Descriptors



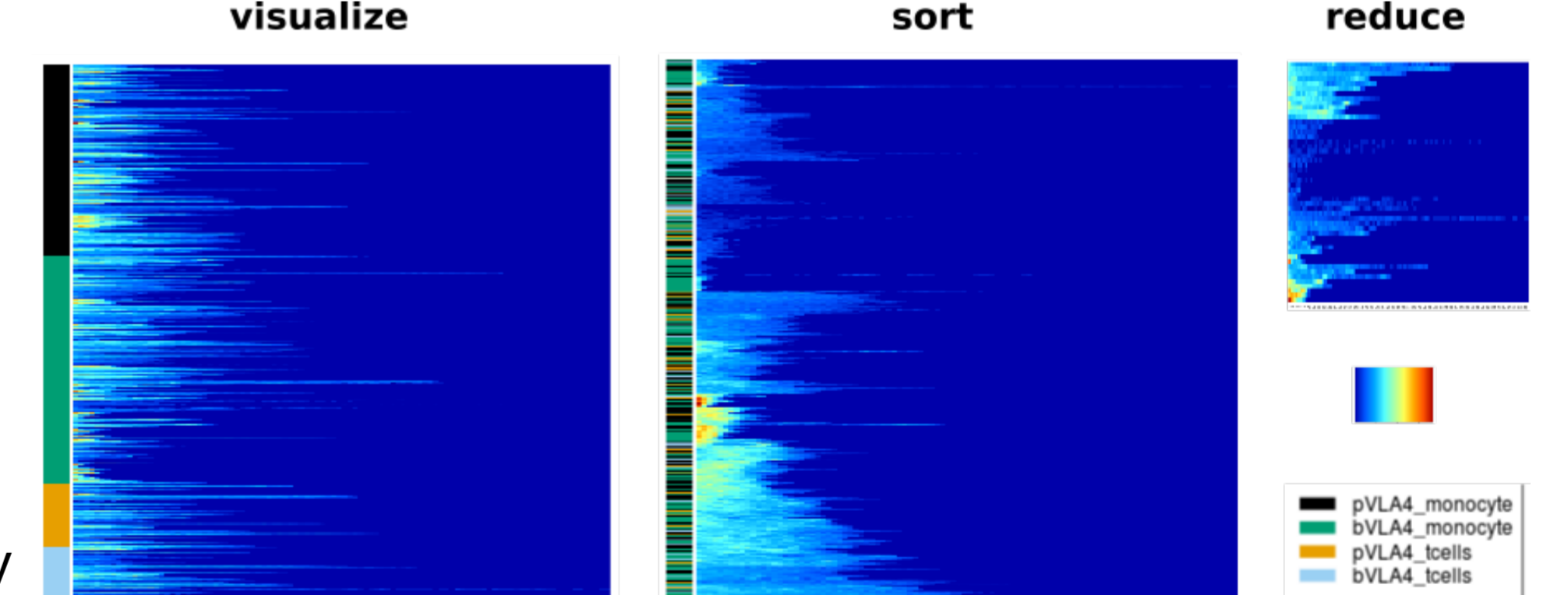
Modularize descriptor derivation for flexibility/exchangeability

Supervised Learning/Modelling

- quality checking** if automatically assigned classes concord to manually defined training sets
- iteratively improving the manually derived terminology** using class definitions from unsupervised clustering to distinguish populations and further quantify substructure compositions

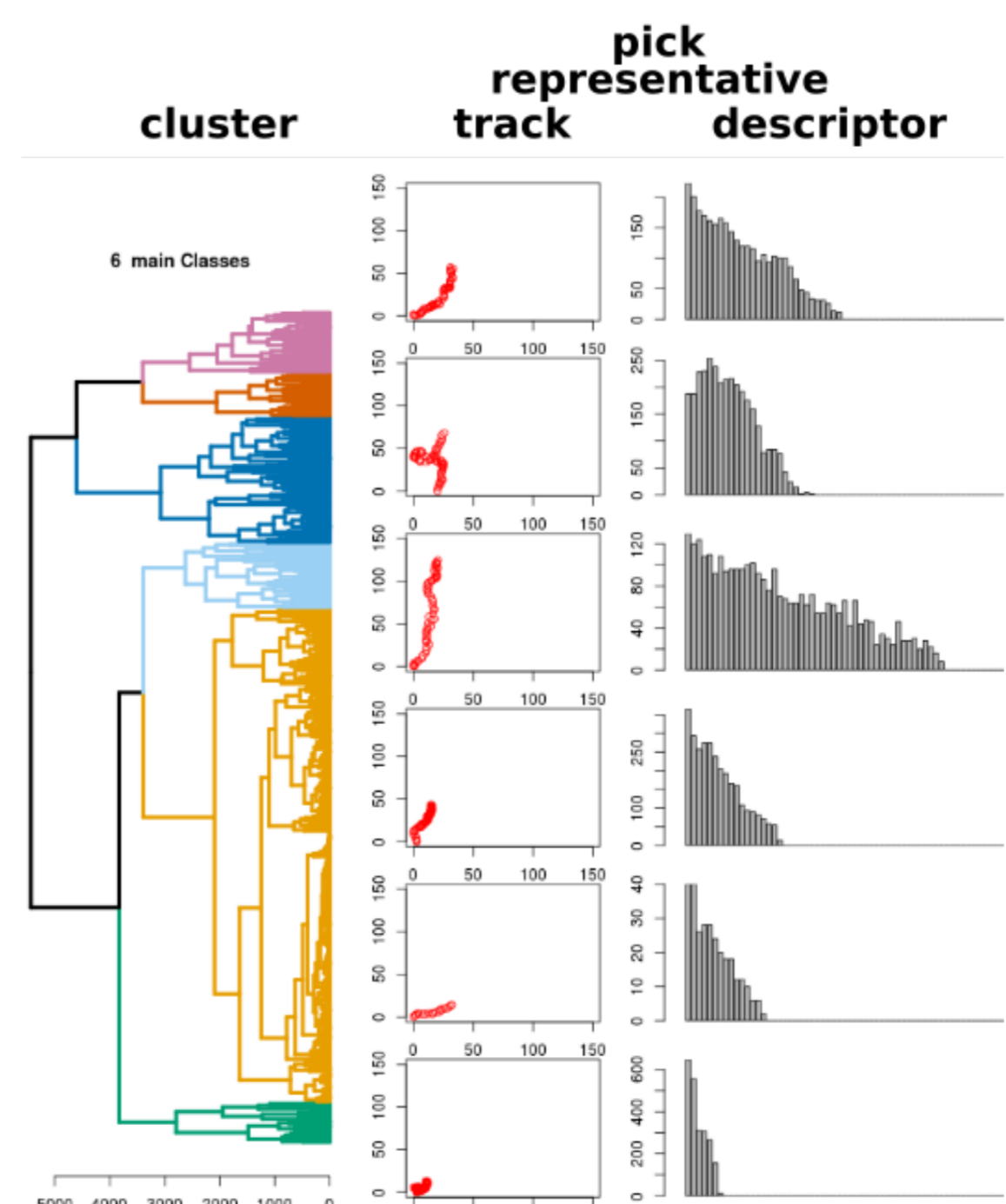
Decision Tree Results The accuracy of a supervised model to predict the manually assigned track class memberships turned out to be 94%. We used a decision tree as a predictive model, which includes *volume asphericity*, the three ratios and the standard deviations of the angles computed from the direction cosines as features.

Explore the Descriptor Space(s)



- test mapping** of trajectory space onto descriptor space, **checking** for retained **information content**;
- picking reductions/representative** to aid manual classification input

Unsupervised Clustering Cluster Tracks into Classes



Differential Analysis yields distinct **subpopulation fingerprints** per condition

Compare Cluster Prevalence

