

COMBINING MORPHOLOGICAL AND MIGRATION PROFILES OF *IN VITRO* TIME-LAPSE DATA

Tim Becker* Juan C. Caicedo* Shantanu Singh* Markus Weckmann† Anne E. Carpenter*

* Imaging Platform, Broad Institute of Harvard and MIT, Cambridge, USA

†Department of Pediatric Allergy and Pulmonology,
Childrens Hospital University of Luebeck, Germany

ABSTRACT

Neutrophil granulocytes belong to the first responders of the innate immune system and are characterized by the capability to migrate towards a site of bacterial infections or inflammation. Several image-based *in vitro* assays exist to record and quantify their movement using migration parameters like speed, directionality or mean squared displacement. In this work, we propose to add morphological parameters to the analysis of time-lapse data. We analyzed a previously published data set of neutrophil granulocytes and combined morphological and migration profiles using the similarity network fusion (SNF) algorithm. To assess the information (gain) stored in the morphological, migration and combined data, we propose to use the signal strength as an objective measure. We conclude that morphological profiling can be combined with classical migration parameters to improve the readout of *in vitro* migration assays.

Index Terms— time-lapse, migration, morphology, chemotaxis, SNF

1. INTRODUCTION

As first-responders to bacterial infection and inflammation, neutrophil granulocytes play an important role in the innate immune system. Their main characteristic is the ability to migrate towards a site of infection, mainly directed and controlled by a process denoted as chemotaxis. To study this specific migration pattern, several *in vitro* assays can mimic chemotaxis using chemical gradients, or that use microfluidic channels to induce a neutrophil movement in a controlled environment. In recent years, several image-based assays have been developed to observe and characterize migration patterns. These assays include a diverse set of technologies, such as the highly specialized lab-on-chip based microfluidic devices [1] that can capture and isolate neutrophils from a single drop of blood [2]. Other very flexible assays have been proposed to easily change several migration conditions like matrices, chemoattractants or drug treatment [3]. Importantly, all these different image and time-lapse based assays offer single cell resolution for automated migration analysis.

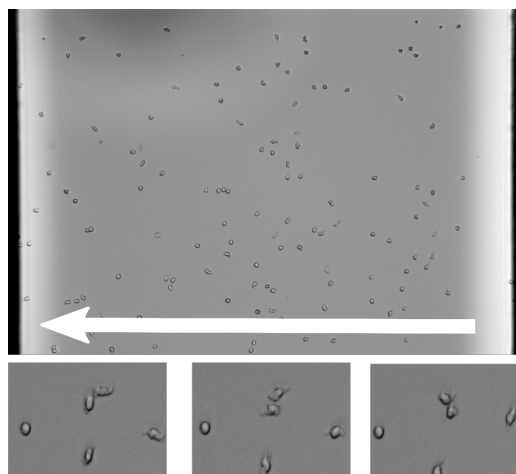


Fig. 1. Neutrophil granulocytes are analyzed in an *in vitro* migration assay where chemotaxis is induced using the chemoattractant Interleukin-8. The neutrophils show a directed migration towards the left channel side and are imaged using brightfield microscopy (121 image per time series, 30 second interval, pixel resolution $1.3\mu m$, $10\times$ magnification, 1.3MP image size). The three cropped images illustrate the varying morphology of the moving cells.

Automated and semi-automated segmentation methods can be used to detect and track cells and neutrophils. For example, researchers can select a representative neutrophil that is used as a reference object in a pattern matching approach to find the position of all other neutrophils in the image [2]. This approach results in good tracking accuracy and allows quantifying the migration using parameters like speed, directionality, mean squared error and forward migration index. While this approach results in a good description of the migration profile, it fails to incorporate information about changes in cellular morphology related to neutrophil response. In this work we aim to investigate the value of incorporating morphological information in the analysis of migration data. We compute morphological features of individual cells in each frame of the time-lapse data, and then create a morphological profile by aggregating these features

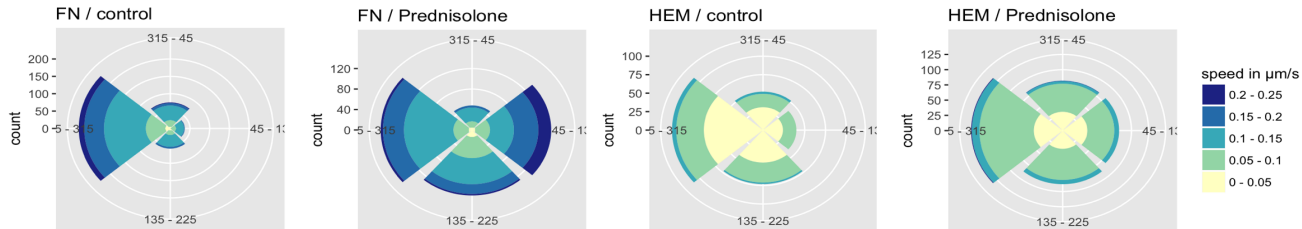


Fig. 2. The migration profiles of all experimental groups are shown as a sector analysis (360 degree histogram of the trajectories distribution). Additionally, the speed distribution in each sector is color coded similar to a windrose plot. A directed migration is represented by a dominant migration towards the chemoattractant Interleukin-8 on the left. The strongest effect of the Prednisolone treatment can be observed in Fibronectin (FN) as opposed to a collagen-rich matrix (HEM).

through time for each trajectory. The ability of morphological features to detect changes in phenotype of migrating cells is compared with that of migration features alone. We also present a method that combines cellular morphologies with migration information, and show that they can complement each other for improved analysis of migration experiments.

2. MATERIAL AND METHODS

2.1. Time-lapse data of neutrophil granulocytes

We analyzed 20 time-lapse movies that were acquired using the simplified migration assay (SiMA) [3]. In this assay, an Interleukin-8 gradient is used to induce a controlled migration (chemotaxis). Of these 20 videos, 10 show a migration of the neutrophils in Fibronectin (FN) and the other 10 show a migration in a collagen-rich matrix (HEM). In each set, 5 time-lapse experiments were recorded with and without treatment (Prednisolone)(summarized in Table 1, for details see [3]). The chemotaxis gradient was allowed to stabilize in the first 10 minutes; the corresponding images were removed. Images were cropped to remove the left and right border visible in the field of view (see Fig. 1).

2.2. Segmentation and tracking

The analysed time-lapse data consists of brightfield images with varying image contrast (Fig. 1). To segment, detect, and track the cells we used CellProfiler 3.0 [4]. The implemented pipeline consists of illumination correction, image segmentation based on a morphological gradient, and cell tracking. For tracking, the CellProfiler implementation of the LAP (linear assignment problem)[5] algorithm is used and optimized following the scheme described in [6]. Segmentation errors result in broken trajectories with a short lifetime. To remove these tracking artefacts, all trajectories with a lifetime shorter than 10 minutes / 20 frames were removed from the analysis. As an objective measure of quality, we used the VOT (valid observation time) which is defined as the total length of all analyzed trajectories divided by the total length of all tracks (including removed tracks). Or in other words: the VOT rep-

resents the fraction of . Table 1 presents these measurements for all experimental groups analyzed in this dataset.

2.3. Morphological Profiling

Cellular phenotypes are quantified in each frame of each movie using morphological profiling [7]. For each cell, 49 features corresponding to shape and appearance are measured, and trajectory profiles are created by averaging each feature across time. Finally, condition profiles are created by summarizing trajectory profiles using mean and standard deviation, resulting in 98 measurements in total. To normalize all data, the features are z-scored with respect to the control condition, prior to computing trajectory profiles. To reduce redundancy, we filter the features so that no two features have a correlation greater than a specified threshold (here, .95 Pearson). Finally, we remove features that have near-zero variance. After all these steps, morphological profiles result in vectors of 52 features for each trajectory. The profiling pipeline and downstream analysis was implemented in R version 3.4.1 using the cytominer package version 0.1¹.

2.4. Migration Profiling

Migration features are calculated based on the centroid position of the tracked cells. For each trajectory, a eight dimen-

¹Methods for Image-Based Cell Profiling [R package cytominer version 0.1.0]. [cited 2017 Oct 2]; Available from: <https://CRAN.R-project.org/package=cytominer>

group	matrix	treatment	videos	tracks	VOT
1	FN	control	5	406	0.88
2	FN	Prednisolone	5	536	0.85
3	HEM	control	5	274	0.96
4	HEM	Prednisolone	5	385	0.96

Table 1. The analyzed image series consists of four experimental groups showing neutrophil migration under different conditions. Each group consists of five measurements using neutrophils of different patients. The tracking metric VOT [3] indicates that the migration data in HEM is higher-quality.

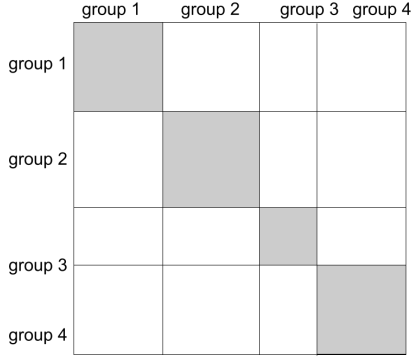


Fig. 3. Structure of the affinity matrix. The submatrix for each of the four experimental groups listed in Table 1 is shown in gray and has size $n_i \times n_i$, where n_i is the number of trajectories in a group. The lower, white six submatrices were used to calculate the null distribution. The fraction strong for each group is defined as the percentage of trajectories with significantly increased values compared to this null distribution.

sional feature vector was calculated including directionality, distance traveled, forward migration index (in direction x and y), speed (speed, speed in x and in y direction) and the chemotactic index as $\cos(\alpha)$ where α is the angle of the trajectory. Similarly to the morphological profiles, all migration features were z-scored with respect to the control condition.

2.5. Combining morphological and migration features

To combine morphological and migration data, we use the similarity network fusion (SNF) algorithm [8]. SNF was developed to combine data sets with different characteristics, such as scales or levels of noise. Intuitively, SNF creates a graph by combining the evidence of connected data points in each modality separately to produce more consistent neighborhoods. First, a similarity network for each data type (in our case migration or morphological profiles) is constructed in form of an affinity matrix W . The matrix W is of size $n \times n$ and each entry $W(i, j)$ is the affinity between two data points x_i, x_j . Using the euclidean distance, $W(i, j)$ is calculated as

$$W_{i,j} \simeq \exp\left(\frac{\|x_i - x_j\|^2}{\sigma \epsilon_{i,j}}\right) \quad (1)$$

where σ is a hyperparameter that can be chosen and $\epsilon_{i,j}$ is a normalization factor. To finally fuse both data sets, the two affinity matrices are iteratively fused resulting in one similarity network, that is represented as a new affinity matrix.

2.6. Signal strength of a profile

The goal of our experiment is to identify conditions where migration and morphology patterns change. If one condition has

a large number of similar cell trajectories that are not commonly seen in other experimental groups, we can conclude that the effect of that condition is significant. We propose to quantify this effect using a statistical test that compares the similarity of trajectories of the same condition against the similarity of trajectories drawn from different random conditions.

To quantify the strength of the phenotype induced in a given condition, we measure how often two trajectory profiles from the condition are more similar than expected at random. We thus build a null distribution comprising the similarity between pairs of trajectories between different conditions (white submatrices in Fig. 3). A threshold θ is calculated as the 95th quantile of this distribution. The phenotype strength is defined for each condition as the fraction of trajectory pairs (i, j) of the same group that satisfy $W(i, j) > \theta$ (each experimental group is shown as a gray submatrix in Fig. 3).

3. RESULTS AND DISCUSSION

We calculated two affinity matrices W_{mo} and W_{mi} for the morphological features and the migration features respectively. The values for the neighbourhood K and σ were found experimentally using grid search to identify the pair with better signal strength (for the control / FN condition). Based on this, we selected $K = 26$ and $\sigma = 0.4$. These values were used to create affinity matrices for morphology, migration data and the combination of both. Two methods were used to combine the morphology and migration data. First, we used the mean affinity value $W_{mi+mo} = \frac{1}{2}(W_{mi} + W_{mo})$ as a baseline of the potential improvement of both sources of information. Then, the two matrices were fused using the SNF algorithm and a combined matrix W_{SNF} was generated. The signal strength of all representations ($W_{mi}, W_{mo}, W_{mi+mo}, W_{SNF}$) was calculated for each of the four experimental groups.

3.1. Control group vs. Prednisolone treatment

We find that both profiles, the morphological and migration features, show a stronger phenotype signal in the control group compared to the Prednisolone treated neutrophils. This result is in line with earlier observation that Prednisolone weakens and partly reverses chemotaxis (compare Fig. 2).

3.2. Migration medium Fibronectin vs. HEM

Next, we compared the migration in FN against HEM. The previous analysis revealed a lowered mean speed of $0.04 \mu m/s$ in HEM compared to $0.13 \mu m/s$ for neutrophils migrating in FN. However, we found a slightly improved signal of the migration features measured in HEM. Surprisingly, the migration medium does not have a big influence on the phenotype strength. This shows that the strength of migration

profiles is not correlated to the mean speed of the tracked objects.

3.3. Combining profiles

We used two methods for combining profiles to compare the gain of information and test the hypothesis that morphology complements migration data. We find that the signal strength of the signal of W_{mi+mo} is better for most but not all conditions compared to the best corresponding W_{mi} or W_{mo} , showing that combining modalities is not a trivial problem (Table 2). In contrast, the signal of W_{SNF} results in an improved value in all conditions (up to $2.6\times$), suggesting that SNF is a good strategy to combine morphology and migration data.

Matrix	Treatment	W_{mo}	W_{mi}	W_{mi+mo}	W_{SNF}
FN	control	7.8%	9.5%	9.4%	10.6%
FN	Predn.	7.4%	6.9%	7.9%	17.9%
HEM	control	9.7%	10.6%	12.2%	15.8%
HEM	Predn.	7.4%	7.3%	9.3%	16.1%

Table 2. For each of the four experimental groups the signal strength is calculated using four different profiles (morphology W_{mo} , migration W_{mi} , average combination W_{mi+mo} and SNF combination W_{SNF}). Numbers are the fraction of cell trajectory pairs that have a similarity significantly larger than the null distribution. Larger is better.

3.4. Conclusion

The framework we developed extends the traditional approach of analyzing time-lapse tracking data by adding morphological features to migration features. To measure the information in the data modalities, we propose the use of a statistical test to estimate the signal strength of experimental groups, which indicates how reliably different each condition is from control. Using this estimation, we found that combining morphological and migration features using SNF improves the overall signal, suggesting that they can complement each other to improve the downstream analysis of tracking data. Signal strength can also be used to find the optimal parameters of the SNF algorithm. Using this framework, we compared the migration of neutrophils in two different matrices: Fibronectin and HEM, a complex, collagen rich matrix. Compared to Fibronectin, the migration in HEM shows a better signal for all three analyzed profiles (morphological, migration and combined). We therefore conclude that complex matrices such as HEM are better suited for studying neutrophil granulocytes behaviour.

4. ACKNOWLEDGEMENTS

Tim Becker was funded by DFG Research Fellowship 5728.

5. REFERENCES

- [1] Leo Boneschansker, Jun Yan, Elisabeth Wong, David M Briscoe, and Daniel Irimia, “Microfluidic platform for the quantitative analysis of leukocyte migration signatures,” *Nat. Commun.*, vol. 5, pp. 4787, Sept. 2014.
- [2] Eric Karl-Heinz Sackmann, Erwin Berthier, Elizabeth A Schwantes, Paul S Fichtinger, Michael D Evans, Laura L Dziadzio, Anna Huttenlocher, Sameer K Mathur, and David J Beebe, “Characterizing asthma from a drop of blood using neutrophil chemotaxis,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 111, no. 16, pp. 5813–5818, Apr. 2014.
- [3] Markus Weckmann, Tim Becker, Gyde Nissen, Martin Pech, and Matthias V Kopp, “SiMA: A simplified migration assay for analyzing neutrophil migration,” *Cytometry A*, May 2017.
- [4] Lee Kamensky, Thouis R Jones, Adam Fraser, Mark-Anthony Bray, David J Logan, Katherine L Madden, Vebjorn Ljosa, Curtis Rueden, Kevin W Eliceiri, and Anne E Carpenter, “Improved structure, function and compatibility for CellProfiler: modular high-throughput image analysis software,” *Bioinformatics*, vol. 27, no. 8, pp. 1179–1180, Apr. 2011.
- [5] Khuloud Jaqaman, Dinah Loerke, Marcel Mettlen, Hiro-taka Kuwata, Sergio Grinstein, Sandra L Schmid, and Gaudenz Danuser, “Robust single-particle tracking in live-cell time-lapse sequences,” *Nat. Methods*, vol. 5, no. 8, pp. 695–702, Aug. 2008.
- [6] Mark-Anthony Bray and Anne E Carpenter, “CellProfiler tracer: exploring and validating high-throughput, time-lapse microscopy image data,” *BMC Bioinformatics*, vol. 16, pp. 368, Jan. 2015.
- [7] Juan C Caicedo, Sam Cooper, Florian Heigwer, Scott Warchal, Peng Qiu, Csaba Molnar, Aliaksei S Vasilevich, Joseph D Barry, Harmanjit Singh Bansal, Oren Kraus, Mathias Wawer, Lassi Paavolainen, Markus D Herrmann, Mohammad Rohban, Jane Hung, Holger Hennig, John Concannon, Ian Smith, Paul A Clemons, Shantanu Singh, Paul Rees, Peter Horvath, Roger G Lington, and Anne E Carpenter, “Data-analysis strategies for image-based cell profiling,” *Nat. Methods*, vol. 14, no. 9, pp. 849–863, Aug. 2017.
- [8] Bo Wang, Aziz M Mezlini, Feyyaz Demir, Marc Fiume, Zhuowen Tu, Michael Brudno, Benjamin Haibe-Kains, and Anna Goldenberg, “Similarity network fusion for aggregating data types on a genomic scale,” *Nat. Methods*, vol. 11, no. 3, pp. 333–337, Mar. 2014.