A short introduction to celltrax analysis tools

What the app does and what doesn’t

The celltrax analysis tools Shiny app is an interactive version of two R packages for analysis of microscopy cell tracing data: the celltrackR (1) and celltrax, offering basic tools with their most common settings to non-programmers. The app:

- computes a set of most relevant track statistic such as displacements, speed and straightness measures
- enables estimation of motion persistence by retrieving displacement vector auto-covariance
- assesses directional movement by cell pair analysis and Hotelling’s test (1,2)
- enables fast, visual analysis of motion mode heterogeneity in the sample with PCA (principle component analysis) (3), UMAP (uniform manifold approximation and projection) (4–6) and MDS (multi-dimensional scaling)
- customization of appearance of analysis result plots and saving the results locally as .pdf images, statistic tables or an .RDa R data file

Due to limitations of external computation power, the app allows for processing of **maximal 1000 tracks or 40000 steps** in total. For larger samples or macroscopic object analysis or any highly customized analysis setup, switching to R programming environment is recommended.

Importing track data

It's highly recommend that you fed pre-processed data into the analysis, e.g. with the celltrax pre-processing tools app, whose output is fully compatible with the required data format in the analysis app. The input data must be in the following format: single, tab-delimited text files with the **track ID, time, x, y and, optionally, z column** is the preferred format (Figure 1). The program does it's best to assign column names to the variables, but make sure if the guesses are correct and adjust them if needed. Important: during import, tracks shorter than 2 steps are silently removed.

id	t	x	y	z
1	48	90.8534	65.3943	-6416.8
1	72	89.5923	64.9042	-6419.93
1	96	88.6958	67.1125	-6421.8
1	120	87.3437	68.2392	-6424.08
1	144	86.274	67.9236	-6425.14
1	168	84.0549	68.2502	-6426.68
1	192	85.9669	68.547	-6426.09

Figure 1: *Required input data format.*

Analysis and results

The data analysis is started by pressing the ‘Launch’ button. In case of any errors, you may simply re-fresh the page by pressing F5. You may adjust settings of analysis and graphical output for each analysis task at the corresponding tabs. **The default settings are supposed to work fine for most input human cell samples scaled in µm – please check out the output units in your sample!**

Track statistics

This tool visualizes displacement vector lengths and speeds (total, median or mean per step), straightness and asphericity metrics (7) for assessment of persistent directional motion and enables to investigate the extent of passive and active cell motility. The estimation of passive and active motility is accomplished by Gaussian modeling as described in more details in the celltrackR package vignette. Briefly, two models of the cell motility are compared with the BIC (Bayesian Information Criterion) measure: a single Gaussian model describing passive, Brownian motion, and a two-Gaussian model more appropriate for active movement. The delta BIC statistic refers simply to the difference in BIC between the one- and two-Gaussian models. Since low BIC values indicate a good model fit, low delta BIC suggests high extent of passive motility. This kind of analysis is facilitated by visual inspection of delta BIC values plotted against the total displacement vector lengths or asphericity provided by our tool: cells moving actively towards a stimulus are expected to have high delta BIC, large displacement vectors and high asphericity.

Auto-covariance

Auto-covariance of displacement vectors provides a handy measure of persistent, both stimulus-specific or -unspecific, motion of the cell. The app offers two ways of computing auto-covariance: with the dot product of subsequent displacement vectors or with the angle between them. Finally, the mean statistic value for particular steps is calculated across the sample. In case of vectors dot products, the interpretation is straightforward: statistic values above 0 are suggestive of persistent motion, yet the values are influenced by the displacement vector lengths. The analysis mode employing the displacement vector angles is insensitive to the vector lengths, statistic values approaching 90 degrees or 1.57 radians indicate lack of auto-covariance.

Pair analysis

Analysis of distances and angles between displacement vectors for cell pairs enables for assessment of directional motion. In brief, low angles between the displacement vectors, independently of the

cell-distance indicates a directional movement of large parts of the sample cells towards a stimulus. The tool implements this kind of analysis in two flavors: by simple analysis of cell pairs (‘cells’) and be a more in-depth investigation of cell pairs at each step (‘steps’). To assess the expected angle values as a function of distance, GAM (generalized additive model) trend with 95% confidence interval is fitted to the data (8). Note: in case no directional motility is present in the sample, angles between the displacement vectors are approximately uniformly distributed between 0 and 180 degree with the expected value of 90 degree.

Movement directionality

The simplest way of directionality assessment is a visual analysis of cell tracks and total displacement vectors as implemented in the app. The normalized displacement vector plot, i. e. with the vector starting points placed at (0, 0), the mean velocity vector is displayed as well. To get a quantitative measure of such motility, Hotelling’s test was proposed (2) and implemented in the app, currently in the two-dimensional (X and Y) version.

Movement heterogeneity

Subsets of cells in the sample may display differing modes of motility in respect to the track straightness, extent of passive motility, direction, speed and track length. The app enables for rapid visual exploration of such heterogeneity by visualizing of results of dimensionality reduction analysis such as PCA (principal component analysis) (3), UMAP (uniform manifold approximation and projection) (4–6) and MDS (multi-dimensional scaling). Variables used in such analysis include step number, tracking duration, delta BIC (passive versus active motility), total displacement, mean speed, overall angle between total displacement vectors, dot product of total displacement vectors, track asphericity, X and Y unit vector displacements. The statistic data set consisting of complete records only is pre-processed by normalization (Z score) with mean or median centering. Note, that the distance metric entry is valid only for UMAP and MDS!

Of importance, the dimensionality reduction tool implemented in the app is a very simple one. The user is recommended to resort to R and to explore possibilities provided by celltrackR, celltrax and clustTools packages for larger data or another set of input track statistics.

Analysis results

The app provides several possibilities to get the analysis results:

- plots may be downloaded separately with the ‘download plot’ button for each drawing

- Track statistics, auto-covariance measures and Hotelling's test results can be downloaded as an Excel table (note: there are multiple tabs!)
- The results of dimensionality reduction analysis can be downloaded as an Excel table along with the analysis algorithm parameters
- The most advanced type of data is the R data file (.RDa) containing the input data (trax/tracks object), track statistics, dimensionality reduction analysis results and the entire set of plots. After loading of the data file into R, additional steps of plot customization and analyses may be easily performed. In this case, you would need to install the celltrax and clustTools development package.

New analysis

To start a new analysis, refresh the page (F5) or click the 'Reset form' button in the side panel.

Credits

The app developers would like to acknowledge authors, programmers and contributors of the packages celltrackR (1), tidyverse (9), rlang, stringi, umap (4–6), shiny, shinyWidgets, shinyjs, and waiter whose tools were used in the app development.

References

1. Wortel IMN, Liu AY, Dannenberg K, Berry JC, Miller MJ, Textor J. CelltrackR: An R package for fast and flexible analysis of immune cell migration data. *ImmunoInformatics* (2021) **1**–**2**:100003. doi:10.1016/J.IMMUNO.2021.100003

2. Textor J, Peixoto A, Henrickson SE, Sinn M, Von Andrian UH, Westermann J. Defining the quantitative limits of intravital two-photon lymphocyte tracking. *Proc Natl Acad Sci U S A* (2011) **108**:12401–12406. doi:10.1073/PNAS.1102288108/-/DCSUPPLEMENTAL/PNAS.201102288SI.PDF

3. Croux C, Filzmoser P, Oliveira MR. Algorithms for Projection-Pursuit robust principal component analysis. *Chemom Intell Lab Syst* (2007) **87**:218–225. doi:10.1016/j.chemolab.2007.01.004

4. Becht E, McInnes L, Healy J, Dutertre CA, Kwok IWH, Ng LG, Ginhoux F, Newell EW. Dimensionality reduction for visualizing single-cell data using UMAP. *Nat Biotechnol 2018 371* (2018) **37**:38–44. doi:10.1038/nbt.4314

5. Konopka T. umap: Uniform Manifold Approximation and Projection. (2022) Available at: https://cran.r-project.org/web/packages/umap/index.html [Accessed June 1, 2022]

6. McInnes L, Healy J, Melville J. UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. (2018) Available at: https://arxiv.org/abs/1802.03426v3 [Accessed February 21, 2022]

7. Mokhtari Z, Mech F, Zitzmann C, Hasenberg M, Gunzer M, Figge MT. Automated characterization and parameter-free classification of cell tracks based on local migration behavior. *PLoS One* (2013) **8**: doi:10.1371/JOURNAL.PONE.0080808

8. Wood SN. Generalized additive models: An introduction with R, second edition. *Gen Addit Model An Introd with R, Second Ed* (2017)1–476. doi:10.1201/9781315370279/GENERALIZED-ADDITIVE-MODELS-SIMON-WOOD

9. Wickham H, Averick M, Bryan J, Chang W, McGowan L, François R, Grolemund G, Hayes A, Henry L, Hester J, et al. Welcome to the Tidyverse. *J Open Source Softw* (2019) **4**:1686. doi:10.21105/joss.01686