Negative correlation between average speed and average turning angle naturally arises for sparsely sampled cell movement data

Vitaly V. Ganusov[1,2], Viktor S. Zenkov[3], and Barun Majumder[1]

[1]Department of Microbiology, University of Tennessee, Knoxville, TN 37996, USA
[2]Department of Mathematics, University of Tennessee, Knoxville, TN 37996, USA
[3]Department of Electrical Engineering and Computer Science, University of Tennessee, Knoxville, TN 37996, USA

December 30, 2020

Abstract

Mechanisms regulating cell movement are not fully understood. One major feature that may allow cells to displace far from an initial location is persistence, the ability to perform movements in a direction similar to the previous movement direction. The level of cell persistence is determined by the turning angles (angles between two sequential movement vectors). Several recent studies, including one in eLife [1], found that a cell's average speed and average turning angle are negatively correlated, and suggested a fundamental cell-intrinsic movement program whereby cells with lower turning ability are intrinsically faster. Using simulations, we show that a negative correlation between the measured average cell's speed and average turning angle naturally arises due to sampling when the frequency of sampling is lower than that of the cell's decisions to change movement direction. Interestingly, assuming heterogeneity in the cell's persistence ability (determined by the concentration parameter of the von Mises-Fisher distribution) but not in the cell's speed results in high differences in measured average speed per cell and a strong negative correlation between cell's speed and average turning angle. Our results thus suggest that the observed negative correlation between a cell's speed and turning angle need not arise due to cell's intrinsic program, but can be a simple consequence of experimental measurements. Additional analyses show, however, that while theoretically it is possible to discriminate between the alternatives when recording cell movements at high frequencies, experimentally high imaging frequencies are associated with increased noise, and this represents a major barrier to determine whether a cell-intrinsic correlation between cell's speed and its turning ability truly exists.

Results and Discussion

Lymphocyte movement is important to locate infections at peripheral sites, but how cells coordinate movements to achieve such a goal remains poorly defined [2, 3]. Typically, movement of cells in tissues

in vivo is recorded using intravital microscopy at a particular frame rate (e.g., a small 3D volume of the liver of $500 \times 500 \times 50$ μm can be scanned every 20 sec, [3, 4]) and by segmenting individual time frames with software (e.g., ImageJ from the NIH or Imaris from Bitplane), 3D coordinates of multiple cells over time can be obtained. Several parameters such as mean square displacement and movement length distribution can be then generated from such digitized data, sometimes providing important insights into cell movement strategies [3, 5]. From cell displacements per time interval or from overall displacement of the cell over the course of the whole movie, one can calculate the cell's speed. Furthermore, one can also calculate turning angles – angles that a cell takes when making sequential movements as recorded in the experimental movie [6, Figure S1]. Average turning angle as well as the distribution of turning angles are informative about the overall pattern of cell movement; that is, when the average turning angle is $90^0 = \pi/2$ (and turning angle distribution is describe by a sin function), cell is performing a Brownian random walk, while smaller average turning angles imply that the cell is performing a correlated random walk. It is important to emphasize, however, that both the cell's average speed and the cell's average turning angle are parameters, estimated from the movement data, and may not match directly the cell-intrinsic speed and turning ability.

Authors of several recent studies accurately measured movement of cells in 2D (*in vitro*) or in 3D (*in vivo*) over time [1, 7, 8] and calculated speeds and turning angles for individual cells. Interestingly, there was a strong negative correlation between these parameters observed for different cell types, different conditions, and with using gene-deficient cells suggesting that such correlation is a fundamental property of moving cells: cells that turn less have intrinsically faster speeds [1]. Jerison & Quake [1] also suggested that cell heterogeneity in movement speeds may further contribute to the overall negative correlation between cell's speed and average turning angle.

Because cell's intrinsic programs for speed and ability to turn are not known, a cell's speed and average turning angle are estimated using measurements. Geometrically, it is clear that if the sampling of a cell's positions occurs rarer than the times of decisions cell makes to turn, a negative correlation between estimated speed and estimated turning angle may naturally arise because cells that turned less by chance are likely to be observed to displace further, and thus, have a higher estimated speed (Figure S1). To check this intuition, we ran a series of stochastic simulations.

First, we simulated 500 cells moving randomly with movement lengths following a thin-tailed distribution (Pareto distribution with $\alpha = 5.5$ and $\bar{r} = 2$, [3, see Materials and Methods]). When the cells' positions were sampled at the frequency the cells were changing movement direction, we observed no correlation between average speed and average turning angle as assumed (Figure 1A). However, if only every 10th movement was sampled, we observed a small but statistically significant negative correlation between measured average speed and average turning angle (Figure 1B). The average speed of cells in such rarely sampled data was also much smaller than the assumed speed, because cells were turning and thus not displacing far from the initial location.

Second, to simulate a correlated random walk we used our newly proposed methodology of sampling random vectors with a bias towards a particular direction using von Mises-Fisher (VMF) distribution [9, see Materials and Methods]. Interestingly, for a moderately biased correlated random walk (with concentration parameter $\kappa_t = 1$ in VMF distribution) we similarly observed that rarer but not regular sampling of cell's positions resulted again in a statistically significant negative correlation between measured cell's speed and average turning angle per individual cell (Figure 1C-D). Both of the sets of simulations assumed that all cells have the same intrinsic speed (determined by $\bar{r}$) and turning ability (determined by κ_t) and resulted in relatively small variability in estimated average

speeds for individual cells.

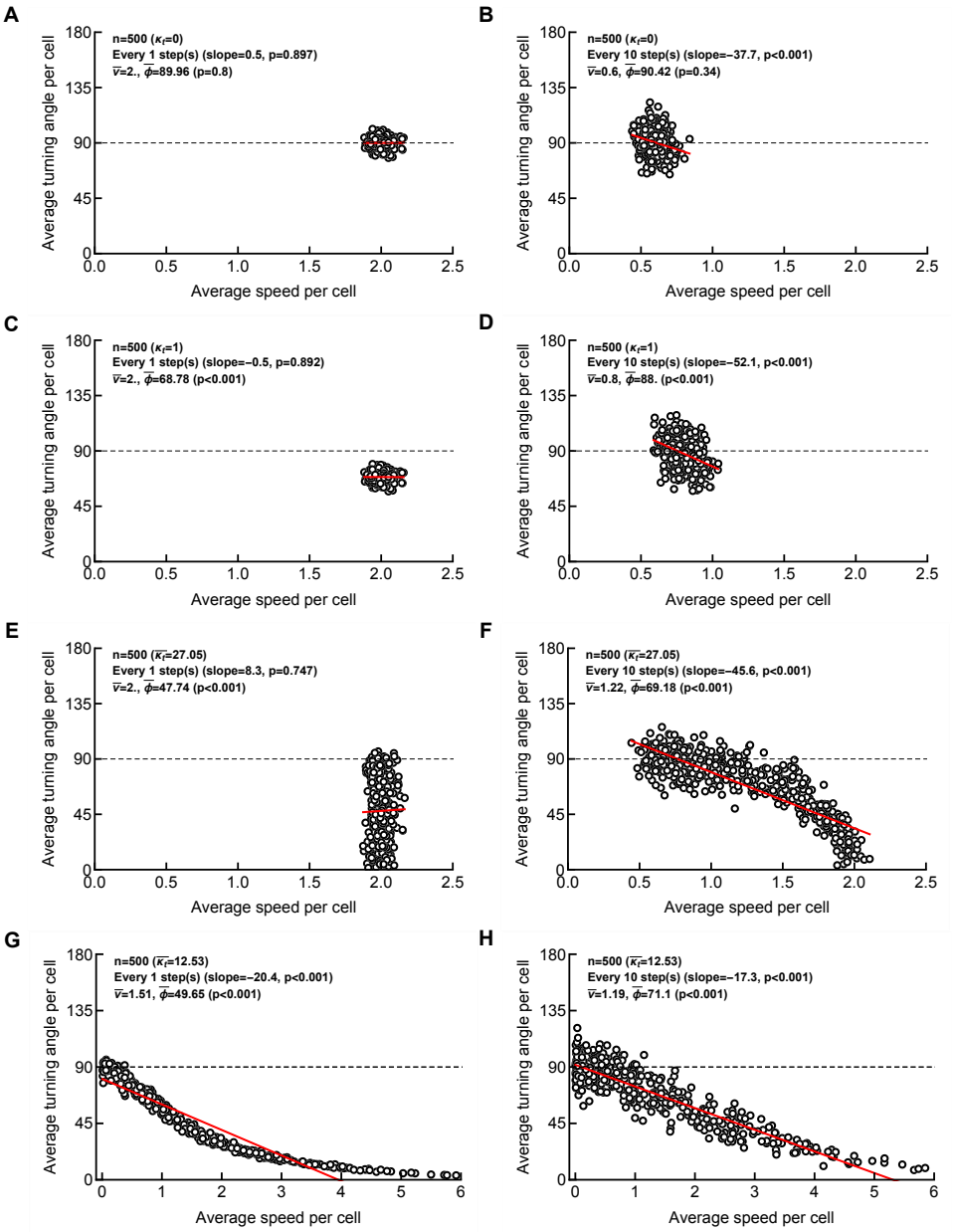

Figure 1: Correlation between average speed and average turning angle arises in the absence of cell-intrinsic link between cell's speed and cell's turning angles. We simulated movement of 500 cells assuming either no preference for movement (A-B), with some persistence for forward movement ($\kappa_t = 1$, C-D), when there is a heterogeneity in cell's persistence of movement (κ_t was sampled from a lognormal distribution, E-F) and sampled movement trajectories either every step (A,C,E) or every tens step (B,D,F). Other details of simulations are given in Materials and Methods. Each panel contains information on the average speed for all cells ($\bar{v}$), average turning angle for all cells ($\bar{\phi}$), the result of linear regression of the average speed per cell and average speed per cell (denoted as "slope" and shown by red line) with p value from the t-test. We also provide a test if the average turning angle of cells in the population is different from 90^0 (Mann-Whitney test). We also simulated cell movements assuming a direct relationship between cell's κ_t and cell's movement speed (G-H). Note different scales in A-F and G-H (because there are few fast moving cells in simulations in G-H).

However, it is possible that cells differ in their ability to turn, e.g., either because of cell-intrinsic program or because of the environmental constrains by physical barriers or chemical cues in specific

locations of the tissue [3]. Therefore, in the third set of simulations we allowed every cell to have an intrinsic turning ability (defined by individual κ_t) drawn from a lognormal distribution (to allow for a broader range of κ_t). Importantly, while there was no correlation between speed and turning angle for frequently measured cell movements (Figure 1E), we observed a strong negative correlation between speed and turning angle for a larger span of the speeds (Figure 1F and see Figures S2-S4). Similarly as with "simpler" simulations, cells that had smaller κ_t, had a higher propensity to turn, resulting in smaller overall displacement, and thus, in smaller measured speeds. In contrast, cells that turn little (high κ_t) have estimated speeds close to intrinsic ($\bar{r} = 2$). Interestingly, allowing for speeds to be a cell's property (i.e., when $\bar{r}$ for individual cells was sampled from lognormal distribution) with random turning angles ($\kappa_t \approx 0$) did not result in the negative correlation between speed and turning angle suggesting that the ability to turn being cell-specific is the key to the observed correlation (Figure S5).

Fourth and finally, we tested how the frequency of sampling of cell movements influences the observed correlation between cell speed and average turning angle when there is an intrinsic link between instantaneous speed of the cell and its turning ability. We therefore simulated cell movement by sampling κ_t from a lognormal distribution and then linked cell's turning ability (defined by κ_t) to the cell's intrinsic speed (assuming $\bar{r} = \ln(1 + \kappa_t)$). Interestingly, the frequency of sampling had a moderate effect on the negative correlation between average speed and average turning angle (Figure 1G-H and Figure S6). Taken together, these results strongly suggest that because the intrinsic cell speed, intrinsic turning ability, or frequency at which any cell makes decisions of turning are not known, a negative correlation between measured speed of cells and average turning angle need not arise from an intrinsic program, but may be a simple consequence of sampling [1].

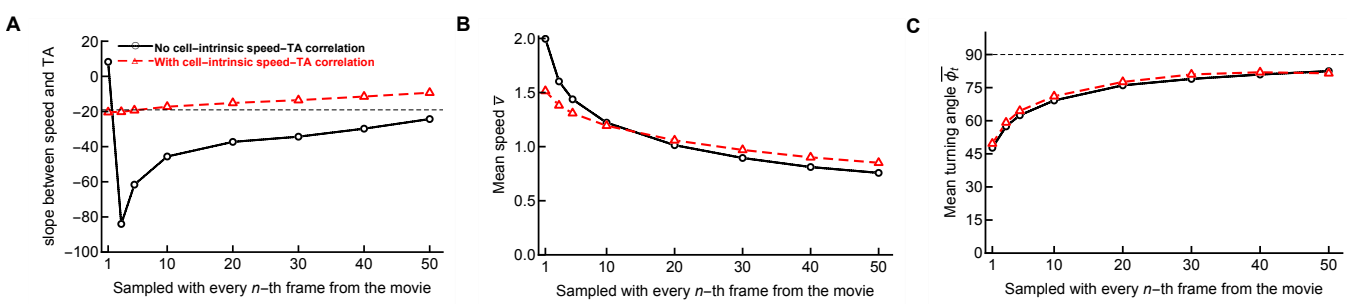

Figure 2: Frequency of imaging does not allow to discriminate between hypotheses explaining negative correlation between measured average speed and average turning angle for a cell. We simulated movement of cells assuming that i) each cell in the population has a preference of moving forward (defined by the concentration parameter κ_t) but all cells have the same intrinsic speed (Figure 1E-F) or ii) for each cell is characterized by a preference to move forward defined by κ_t and intrinsic speed defined as $v^{(i)} = \ln(1 + \kappa_t^{(i)})$ where (i) indicates an ith cell (Figure 1G-H). For different frequencies of recording cell movement we calculated the slope between average speed and average turning angle (TA or $\bar{\phi}_t$) per cell (panel A), the average speed per cell ($\bar{v}$, panel B), or average turning angle per cell ($\bar{\phi}_t$, panel C) for the two hypotheses (without and with cell-intrinsic link between speed and turning angle, shown by different markers and lines). Thin dashed line in panel A denotes the expected slope with lowest possible frequency of imaging ($-90 \arctan(\pi/(2\bar{r}))/(\pi/2) \approx -0.19$ for $\bar{r} = 2$). Values on the x-axes indicate which frames were included in the calculations. For example, $n = 10$ indicates that data at $t = 1, 11, 21, 31 \ldots$ were used in calculations. In simulations concentration parameter κ_t was lognormally distributed between cells with $\mu = 1$ and $\sigma = 2$ (see Figure 1 and Materials and Methods for more detail).

In our simulations it was clear that the frequency of sampling has a major impact on the regression slope between average cell's speed and average turning angle (e.g., Figures S2-S4). Therefore, we

tested if a change in frequency of sampling of cell movement can be used to discriminate between cell-intrinsic vs. observed correlation between speed and turning angle. Therefore, we calculated how the slope between speed and turning angle, average speed, and average turning angle change with reducing the frequency of imaging (Figure 2). Unfortunately, the two models were similar in how these parameters changed with imaging frequency except of the narrow range when imaging frequency would coincide with the frequency at which cells make turn decisions (Figure 2A at $n = 1$). However, very frequent imaging contains many artifacts due to changes in cell shape or fluorescence signal from the cell [10], suggesting that changes in frequency of imaging may not be the method to discriminate between these alternatives.

Cell migration is a complicated process. Cells move randomly, often by correlated random walks as determined by the turning angle distribution. Recent studies under various conditions and constraints have shown that faster cells tend to move straighter because of persistence and slower cells tend to change direction more often [1, 7–10]. It was recently suggested that this negative correlation between average speed of cell and average turning angle might be due to cell's intrinsic program [1]. Here we argue that this negative correlation is a natural consequence of fact that both cell's speed and average turning angle are estimated from the data; however, the frequency of sampling does not allow to discriminate between the alternatives. Propensity of moving cells to keep forward movement when environment and conditions are constant can be postulated as a type of "biological inertia" (or "biological conservation of momentum") specific to the cells [3]. Per such postulate changes in the environment (e.g., shape of the surface on which cells move, cues, or change in nutrients) are responsible for cells to change movement, e.g., to stop and/or turn. It should be noted, however, that such "biological conservation of momentum" or "inertia" should not be confused with physical inertia because viscous forces are much stronger than the inertial forces for biological (small) cells (Reynolds number is of the order of 10^{-4}, which classifies the cell movements under physical processes of low Reynolds number [11]). A more integrated experimental approach is needed to be capable of examining continuous cell movements as we have observed in some experiments [4]. Future studies should attempt to define ways to falsify the simpler explanation of observed negative correlation between cell speed and turning angle as the consequence of sampling.

Materials and Methods

To simulate cell movement, we assumed that cells undergo a correlated (persistent) random walk with the degree of persistence determined by the concentration parameter κ_t in the von Mises-Fisher (VMF) distribution [9, 12], which is a probability distribution on an n-dimensional sphere (in our case, $n = 3$) that chooses a direction with measurable bias toward a given direction. The distribution is

$$P(\chi|\mu, \kappa_t) = \frac{\kappa_t e^{\kappa_t \mu^T \chi}}{2\pi(e^{\kappa_t} - e^{-\kappa_t})}, \tag{1}$$

where μ is the direction vector toward which there is bias (e.g., the previous movement vector), χ is the newly chosen direction vector, κ_t is the concentration (with 0 meaning no bias, positive meaning persistence, and negative meaning aversion), and $|\mu| = |\chi| = 1$. Random (biased) vectors given direction μ and concentration parameter κ_t were generated in Mathematica (v. 12.1) by choosing

a vector with bias toward direction {0,0,1}, which simplifies the process to choosing 1) x and y randomly from a normal distribution $N(0,1)$ (using command `RandomVariate`), and 2) z based on the von Mises-Fisher distribution, chosen by

$$z = 1 + (\ln(r) + \ln(1 + (1 - r)\frac{e^{-2\kappa_t}}{r}))/\kappa_t, \tag{2}$$

where r is chosen uniformly between 0 and 1 [13, 14]. Then x and y are weighted to place the chosen vector on the unit sphere, and then we use a rotation transform (command `RotationTransform`) to adjust the generated vector with respect to the desired bias direction. The native Mathematica code to generate a random vector using VMF distribution is

```
vonMisesFisherRandom[\[Mu]_?VectorQ, \[Kappa]_?NumericQ] :=
 Module[{\[Xi] = RandomReal[], w},
  w = 1 + (Log[\[Xi]] + Log[1 + (1 - \[Xi]) Exp[-2 \[Kappa]]/\[Xi]])/\[Kappa];
  RotationTransform[{{0, 0, 1}, Normalize[\[Mu]]}][Append[Sqrt[1 - w^2]
    Normalize[RandomVariate[NormalDistribution[], 2]], w]]]
```

The length of the movement was drawn randomly from the Pareto (powerlaw) distribution

$$f(r|r_{\min}, \alpha) = \frac{\alpha r_{\min}^{\alpha}}{r^{\alpha+1}}, \tag{3}$$

where $r_{\min}$ and α are the scale and shape parameter, respectively, and $\bar{r} = \alpha r_{\min}/(\alpha - 1)$. In the Pareto distribution, $r \geq r_{\min}$. In simulations, we assumed $\alpha = 5.5$, corresponding to a Brownian-like distribution of movement lengths [3], $\bar{r} = 2$, and $r_{\min} = \bar{r}(\alpha - 1)/\alpha$. Thus, κ_t determines the degree of walk persistence (i.e., the average turning angle) and $r_{\min}$ and α determine the speed of the cell movement. As these quantities are independent, in most of our simulations speed and turning angles truly have no correlation.

In simulations, each cell moves in a random direction (determined by κ_t in eqn. (1) and by the previous movement vector) and by a random distance (determined by $r_{\min}$ and α in eqn. (3)). However, if cell movements are measured at a lower frequency that cell is moving, then the measured cell movement speed and average turning angles are calculated from the "assumed" trajectory that takes into account only some of the cell's positions. For example, in simulating cell movement for 100 steps we can only count every 10 steps as a movement, thus calculating the cell's speed by taking positions $1, 11, \ldots 91$ and then calculating the average speed $\bar{v}$ and average turning angle $\bar{\phi}_t$ for these actual positions.

In some simulations, we assumed that every cell has an inherent concentration parameter κ_t which determines cell's persistence ability. We sampled values of κ_t from a lognormal distribution defined as

$$p(\kappa_t|\mu, \sigma) = \frac{1}{\sqrt{2\pi}\kappa_t\sigma} \mathrm{e}^{-\frac{(\ln(\kappa_t)-\mu)^2}{2\sigma^2}}, \tag{4}$$

with $\mu = 1$ and $\sigma = 2$.

We also simulated cell movement with an intrinsic cell's movement speed (determined by $\bar{r}$) and its persistence in the walk (determined by κ_t) being correlated. We sampled κ_t for each cell from a lognormal distribution (eqn. (4)) with parameters $\mu = 1$ and $\sigma = 2$ and let the average movement length for each cell be $\bar{r} = \ln(1 + \kappa_t)$. Then, setting $\alpha = 5.5$, we let for every cell $r_{\min} = \bar{r}(\alpha - 1)/\alpha$ in eqn. (3).

When simulating a distribution of cell-intrinsic speeds we assumed that for each cell $\bar{r}$ in Pareto distribution follows a lognormal distribution as for κ_t (eqn. (4)) with $\mu = 1$ and $\sigma = 2$ (and $\alpha = 5.5$).

Data sources

No new data

Code sources

All analyses have been primarily performed in Mathematica (ver 12) and codes used to generate most of figures in the paper are provided on GitHub:
`https://github.com/vganusov/correlated_random_walk_simulations`

Ethics statement

No animal or human experiments performed.

Author contributions

The question for the study arose during discussions of the cell movement data between all authors. VVG ran the simulations and discussed results with other authors. BM wrote the first draft of the paper, and all authors contributed the final draft of the paper.

Acknowledgments

This work was supported by the NIH grant (R01 GM118553) to VVG.

References

1. Jerison, E. R. & Quake, S. R. 2020 Heterogeneous T cell motility behaviors emerge from a coupling between speed and turning in vivo. *eLife*, **9**. doi:10.7554/eLife.53933.

2. Krummel, M. F., Bartumeus, F. & Gérard, A. 2016 T cell migration, search strategies and mechanisms. *Nature reviews. Immunology*, **16**, 193–201.

3. Rajakaruna, H., O'Connor, J., Cockburn, I. A. & Ganusov, V. V. 2020 Environment-imposed constraints make Brownian walkers efficient searchers. *bioRxiv*.

4. McNamara, H. A., Cai, Y., Wagle, M. V., Sontani, Y., Roots, C. M., Miosge, L. A., O'Connor, J. H., Sutton, H. J., Ganusov, V. V. *et al.* 2017 Up-regulation of LFA-1 allows liver-resident memory T cells to patrol and remain in the hepatic sinusoids. *Science Immunology*, pp. 1–10.

5. Harris, T. H., Banigan, E. J., Christian, D. A., Konradt, C., Wojno, E. D. T., Norose, K., Wilson, E. H., John, B., Weninger, W. *et al.* 2012 Generalized Lévy walks and the role of chemokines in migration of effector CD8+ T cells. *Nature*, **486**(7404), 545–548.

6. Beltman, J. B., Marée, A. F. M. & de Boer, R. J. 2009 Analysing immune cell migration. *Nat Rev Immunol*, **9**(11), 789–798.

7. Maiuri, P., Rupprecht, J.-F., Wieser, S., Ruprecht, V., Bénichou, O., Carpi, N., Coppey, M., De Beco, S., Gov, N. *et al.* 2015 Actin flows mediate a universal coupling between cell speed and cell persistence. *Cell*, **161**, 374–386. doi:10.1016/j.cell.2015.01.056.

8. Gorelik, R. & Gautreau, A. 2015 The arp2/3 inhibitory protein arpin induces cell turning by pausing cell migration. *Cytoskeleton (Hoboken, N.J.)*, **72**, 362–371. doi:10.1002/cm.21233.

9. Zenkov, V. S., O'Connor, J., Cockburn, I. A. & Ganusov, V. V. 2020 A minority of liver-localized CD8 T cells demonstrate difficult-to-detect attraction to Plasmodium-infected hepatocytes. *bioRxiv*.

10. Gorelik, R. & Gautreau, A. 2014 Quantitative and unbiased analysis of directional persistence in cell migration. *Nature protocols*, **9**, 1931–1943. doi:10.1038/nprot.2014.131.

11. Purcell, E. M. 1977 Life at low reynolds number. *American Journal of Physics*, **45**(1), 3–11. doi:10.1119/1.10903.

12. Hillen, T., Painter, K. J., Swan, A. C. & Murtha, A. D. 2017 Moments of von Mises and Fisher distributions and applications. *Mathematical biosciences and engineering*, **14**, 673–694. doi: 10.3934/mbe.2017038.

13. Ulrich, G. 1984 Computer generation of distributions on the m-sphere. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, **33**(2), 158–163.

14. discontentment, J. M. 2020 Vectors with a certain magnitude in Mathematica. doi: https://mathematica.stackexchange.com/questions/13038/vectors-with-a-certain-magnitude-in-mathematica/13042.

Negative correlation between average speed and average turning angle naturally arises for sparsely sampled cell movement data

Vitaly V. Ganusov, Viktor S. Zenkov, and Barun Majumder

Supplemental Information

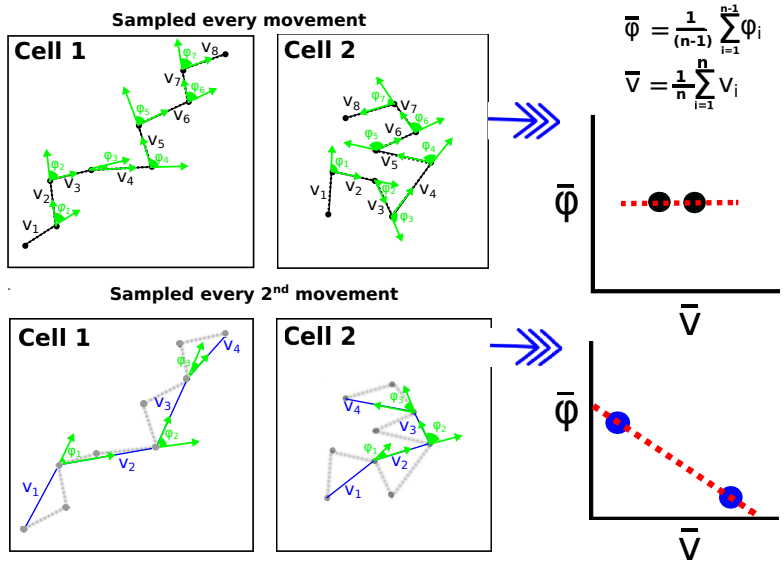

Figure S1: Schematic representation of how frequency of sampling of cell movement influences the observed relationship between average speed and average turning angle per cell. We plot trajectories for 2 cells and show the speed of every movement (denoted as v_i) and turning angle (ϕ_i). When sampling is done at the same frequency as cells make decision to turn (top panels), per model assumption average speed ($\bar{v}$) and average turning angle ($\bar{\phi}$) do not correlate. However, when sampling is done less frequently (e.g., every 2nd movement, bottom panels) the cell 1 that turned less have a lower average turning angle and higher observed speed as compared to the cell 2 that turned a lot.

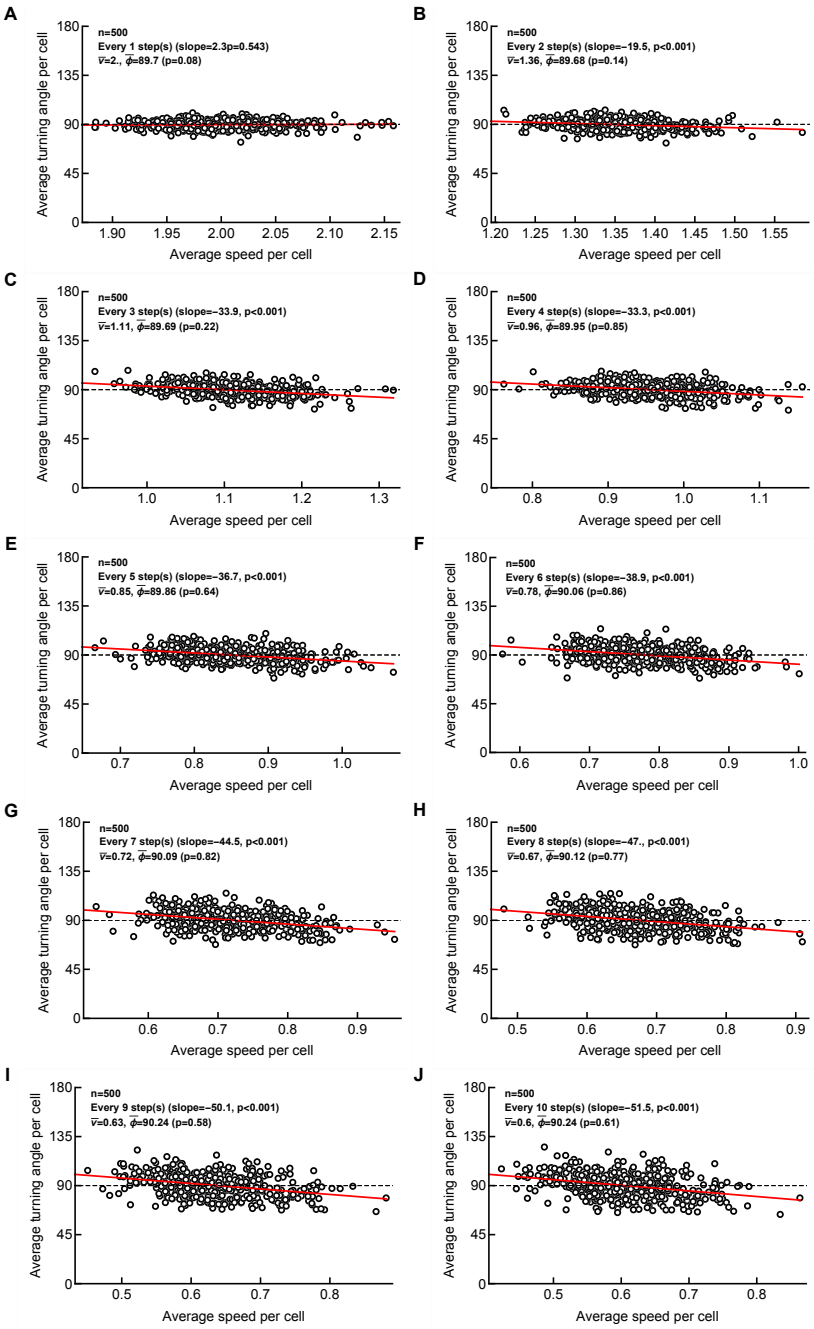

Figure S2: Correlation between average speed and average turning angle naturally arises for coarsely sampled data for uncorrelated random (Brownian) walk. Here all cells have the same concentration parameter $\kappa_t \to 0$. We simulated movements of 500 cells for 100 steps and calculated the average speed and average turning angle per cell when sampling the data at different frequencies, starting with every step (panel A) and finishing with every 10 steps (panel J). Each panel contrains information on the average speed for all cells ($\bar{v}$), average turning angle for all cells ($\bar{\phi}$), the result of linear regression of the average speed per cell and average speed per cell (denoted as "slope") with p value from the t-test. We also provide a test if the average turning angle of cells in the population is different from 90^0 (Mann-Whitney test).

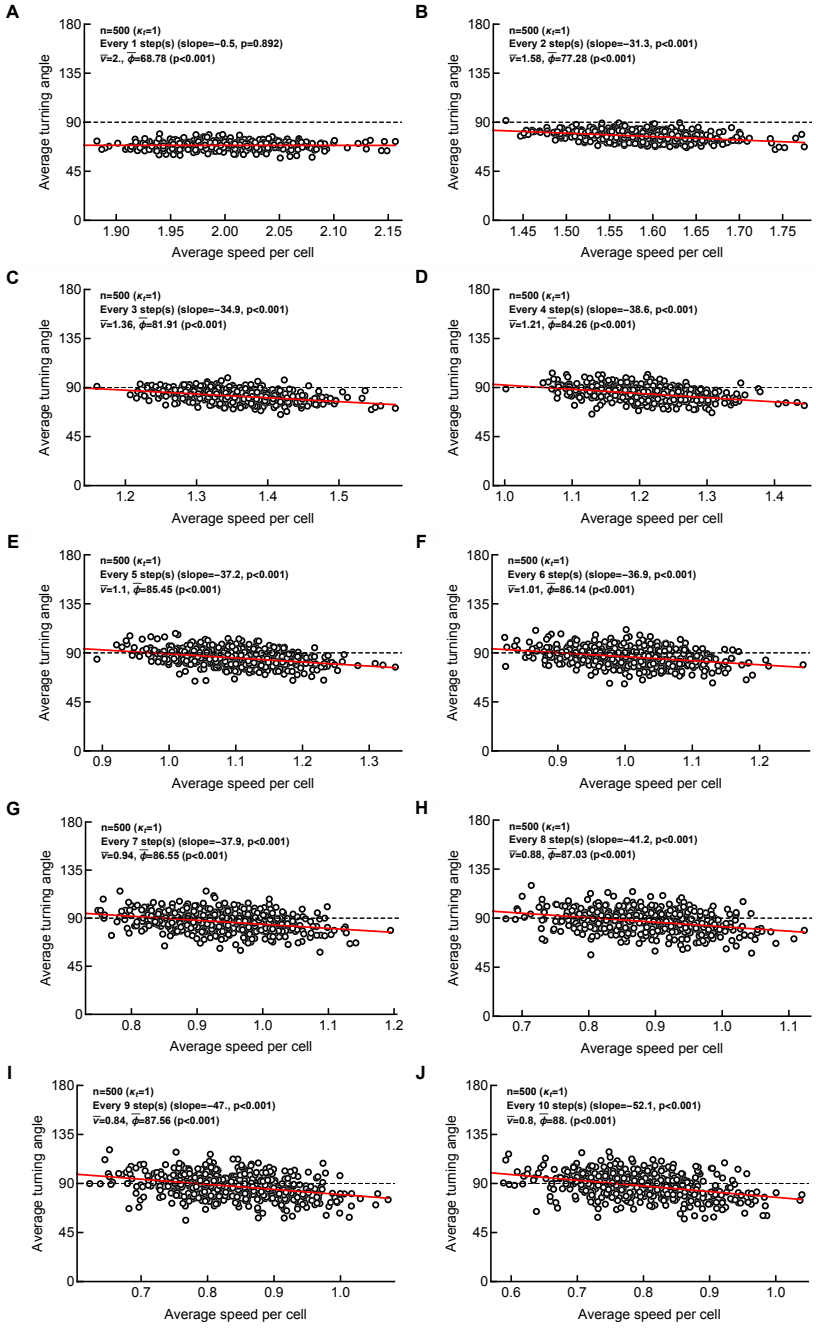

Figure S3: Correlation between average speed and average turning angle appears for coarsely sampled data when cell undergo a correlated random walk. Here all cells have the same concentration parameter $\kappa_t = 1$. See Figure S2 for other details.

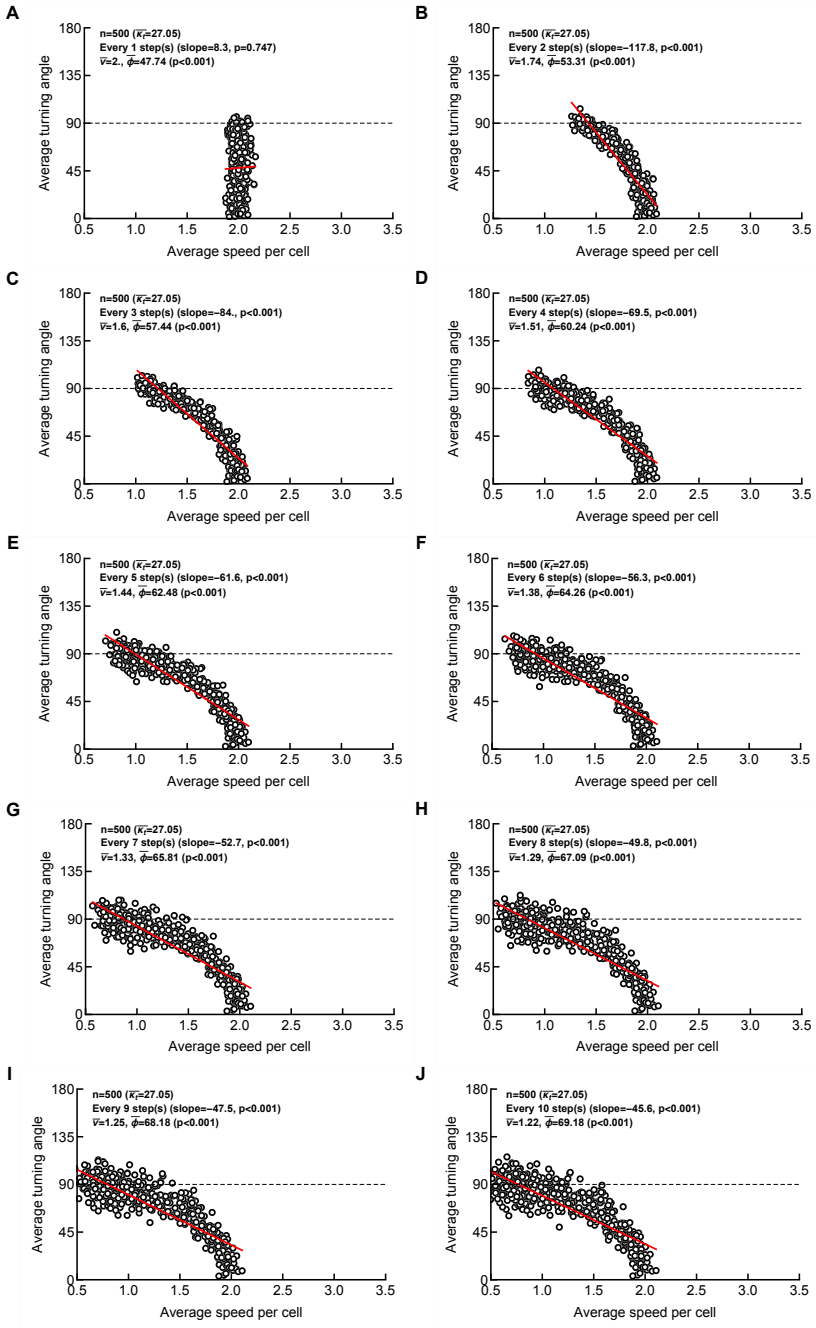

Figure S4: Strong negative correlation between average speed and average turning angle for heterogeneous cell populations with coarsely sampled data. Here we assume that every cell in the population has a different κ_t which was drawn from a lognormal distribution (eqn. (4) with $\mu = 1$ and $\sigma = 2$). Here $\bar{\kappa_t}$ is the average concentration parameter in the population. See Figure S2 for other details.

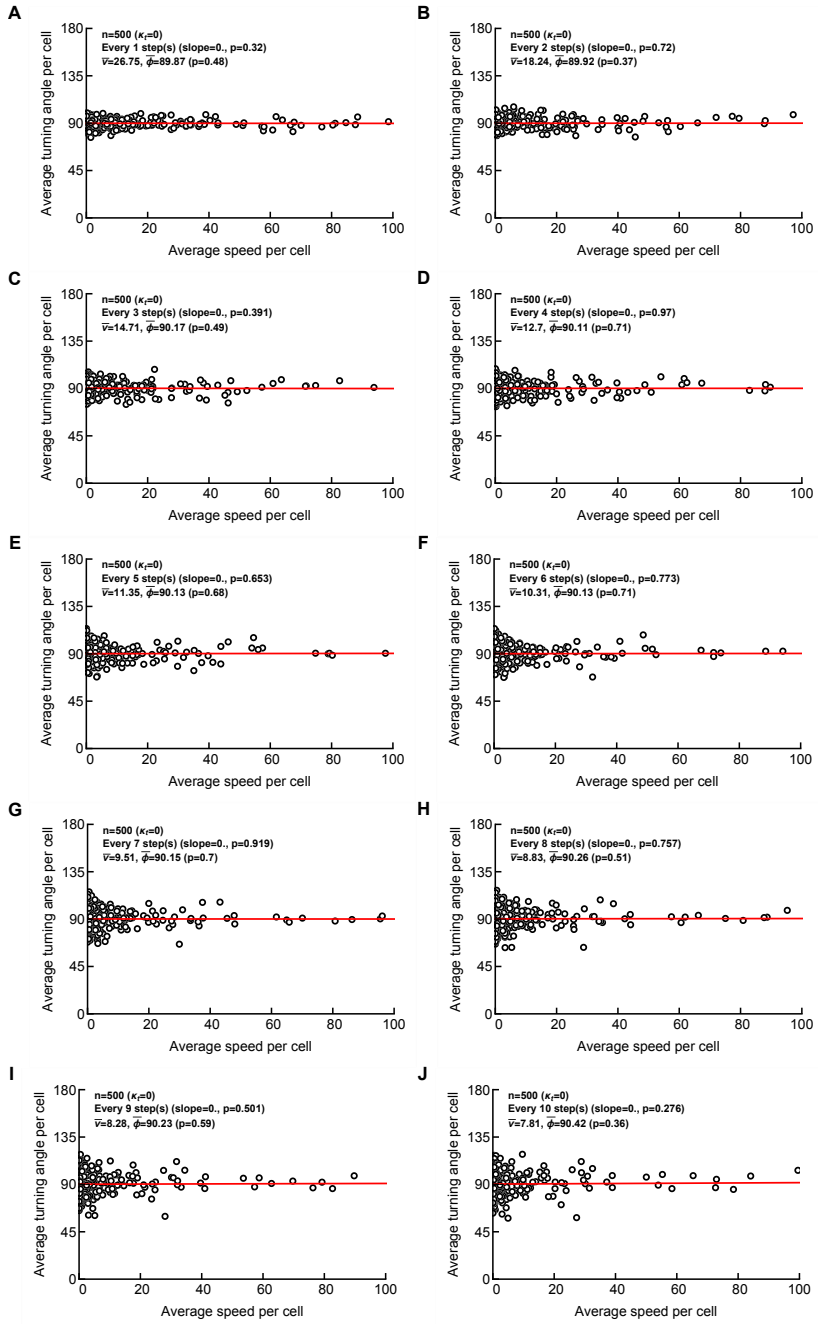

Figure S5: Variability in intrinsic movement speed does not result in correlation between average speed and average turning angle. Here we assume that every cell in the population has a different $\bar{r}$ which was drawn from a lognormal distribution (eqn. (4) with $\mu = 1$ and $\sigma = 2$). Turning angles are random with $\kappa_t \to 0$. See Figure S2 for other details.

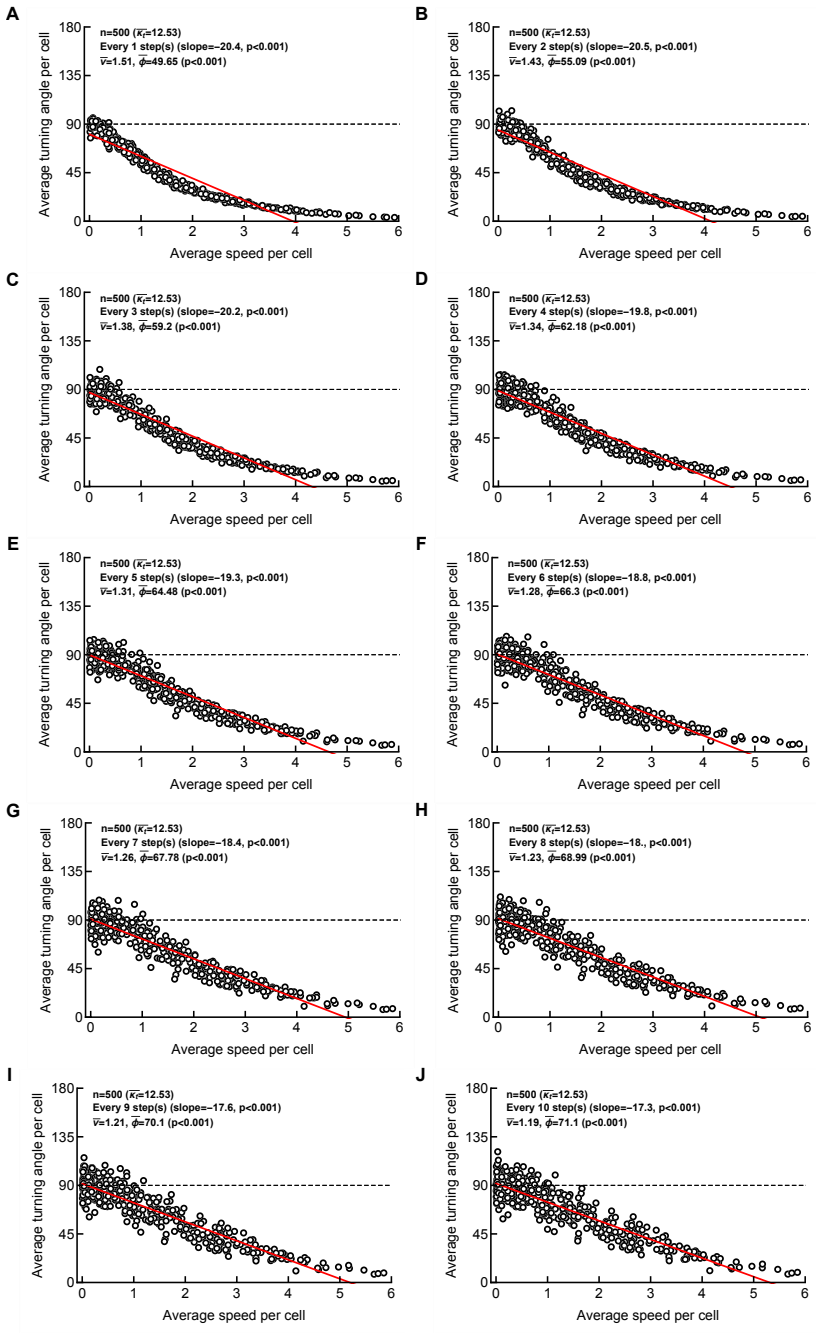

Figure S6: Intrinsic correlation between instantaneous cell speed and turning angles is largely insensitive to sampling frequency. Here we assume that every cell in the population has a different κ_t which was drawn from a lognormal distribution (eqn. (4) with $\mu = 1$ and $\sigma = 2$), and every cell has a speed determined by κ_t (i.e., assuming a relationship for each cell as $\bar{r} = \ln(1 + \kappa_t)$ for the Pareto distribution (eqn. (3))). See Figure S2 for other details.