A T cell behavioral manifold couples speed and turning to create heterogeneity *in vivo*

Elizabeth R. Jerison,[1] Stephen R. Quake[1,2,3*]

[1]Department of Applied Physics, Stanford University, Stanford, CA 94305,
[2]Department of Bioengineering, Stanford University, Stanford, CA 94305
[3]Chan Zuckerberg Biohub, San Francisco, CA 94158
[*]To whom correspondence should be addressed; E-mail: steve@quake-lab.org.

Abstract: Immune cells carry out dynamic search processes for antigen and pathogens. T cells *in vivo* migrate primarily via undirected random walks, but it remains unresolved how these random walks generate a characteristic search behavior. Here we use light sheet microscopy to observe the native population of T cells in the larval zebrafish tail and fin fold over long times. We find that cell-to-cell variability is amplified by a correlation between speed and directional persistence, generating a characteristic cell behavioral manifold that is preserved under a perturbation to cell speeds. Our results suggest that speed-persistence coupling may be an intrinsic feature of T cell migration that provides an alternative to Levy flight for accessing a broad range of length scales *in vivo*.

Many immune cells migrate through tissue in search of antigen or pathogens. In some cases, such as during extravasation from blood vessels and homing to target organs, this migration is guided by chemokine gradients (*1–3*). However, for naive T cells within T cell zones, *in situ* imaging studies have found that unguided random walk processes dominate ((*4–12*), reviewed in (*13*)). This observation creates a conceptual challenge: T cells must dwell at scales of microns to make contact with antigen presenting cells, yet migrate over scales of millimeters to find rare targets. A conventional diffusive random walk struggles to access these varied scales efficiently, since a walker that dwells near another cell for 1 minute would require several days to travel 1 mm. Several authors have suggested that T cells may have an intrinsic behavioral program that allows them to explore over different length scales (*9, 13, 14*). However, testing this hypothesis via *in situ* fluorescence microscopy raises inherent technical challenges: to observe a single cell accessing a broad range of spatial scales, it is necessary to have micron scale resolution over fields of view of millimeters, with low enough photodamage to observe the same cells at high

spatiotemporal resolution over long periods. For example, one intriguing proposal is that T cells perform Levy flight (*9*), an anomalous random walk characterized by a power-law distribution of step sizes. Such random walks have been described in detail in the physics and ecology literature (*15, 16*), and their scale-free behavior provides a natural way for foragers to accelerate searches in many contexts (*17*). However, observation over short periods cannot distinguish between Levy flight and heterogeneity amongst individual walkers (*18*), both of which can create a broad distribution of displacements. More generally, we would like to understand whether there is a statistically-consistent behavioral program carried out by individual cells.

To address this question, we used selective plane illumination microscopy (*19, 20*) to image the native population of T cells in the larval zebrafish over millimeter fields of view and periods of a few hours. In whole mount live imaging of Tg(*lck*:GFP, *nacre*-/-) zebrafish (*21*), we observed a population of motile cells concentrated in the tail fin and larval fin fold. (Figure 1A,F, Movie S1-S2). We chose this population for further study because of the simplicity of the tissue context: it enables us to make detailed measurements of the behavior of many individual members of a native population sampling a similar environment.

Previous studies have proposed two classes of models for T cell migration: Levy flight, and persistent random walks (*7–9, 12*). In the Levy-flight model, all cells follow the same migration program, but sample from a very broad distribution of step sizes. This model is conceptually appealing because cells have a consistent internal behavior that nonetheless allows them to explore across many length scales. In persistent random walk models, cells behave diffusively at long times, but speeds and persistence times may vary across the population. The distinction hinges on whether the statistics of individual trajectories are scale-free, so that super-diffusive behavior continues to long times; or if, alternatively, individual trajectories are diffusive at long times but there is heterogeneity across the population. To address this question in our system, we performed a standard analysis of mean squared displacement as a function of time interval. Consistent with previous measurements (*7–9, 12*), we observed a faster-than-linear increase in MSD at early times, indicating super-diffusive behavior, with a best-fit line in surprisingly good quantitative agreement with previous observations up through 10 minutes (*9*) (Figure 1C). However, we observed a transition at the scale of minutes (particularly evident on a linear scale, Figure 1C inset), consistent with the persistent random walk class of models, and inconsistent with Levy flight. To further test for an intermediate timescale, we computed the velocity-velocity power spectrum, using secant-approximated velocities along each trajectory. This quantity captures the timescale at which the velocities become decorrelated, if it exists. The function is flat for a simple random walk; for a Levy-flight process the same negative slope is observed at all frequencies (*22*). A persistent random walk-type model passes between these two behaviors at some intermediate frequency scale (*22*). Consistent with the MSD analysis, we observe two regimes, with a clear timescale on the order of minutes (Figure 1D). Finally, we computed the distribution of lengths between direction changes (bout lengths) within a trajectory (Materials and Methods), scaled by the average bout length as suggested in (18), and did not observe the characteristic Levy-flight power law (Figure 1E).

While the population-averaged MSD and PSD are consistent with a persistent random walk

with a single average speed and persistence time across all the cells, examples of velocity traces suggest heterogeneity in speed between cells, that can persist over spans of a few hours (Figure 1E-F, Movie S2). In fact, 88% of trajectories have distributions of secant-approximated speeds that are inconsistent with the overall distribution (KS test, $p < .01$). Interestingly, we also found significant heterogeneity in cell turning behavior: 67% of cells had turn angle distributions inconsistent with the overall distribution (KS-test at $p < .01$).

The surprising degree of heterogeneity in random walk behavior amongst individual cells in this very simple tissue context led us to ask whether there are any underlying rules governing this variation. Are individual cells free to pick any turn and speed statistics, or are there behavioral constraints?

Since cells displayed a characteristic range of speeds over a period of hours, we investigated whether other migration behaviors were correlated with this property. To do so, we first divided the cells into quintiles based on speed. We refer to these groups as speed classes. We observed strong variation in the distribution of turn angles amongst speed classes (Figure 2A). All of the cells have a propensity to either turn shallowly or, occasionally, to turn around 180 degrees. However, the fastest cells turn the most shallowly and are the least likely to turn backwards, and this characteristic varies smoothly across the speed classes (Figure 2A).

A systematic dependence of the distribution of turn angles on the cell's average speed could be driven by a local coupling between speed and turn angle: cells tend to go straighter whenever they go fast, which the faster cells do more often. Alternatively, it could be driven by an overall behavioral difference between fast and slow cells: fast cells tend to go straighter even during periods where they happen to be traveling slowly, and vice-versa. To distinguish between these possibilities, we measured the average turn angle as a function of the size of the steps surrounding it (Figure 2B). We found that both of these effects contribute: all cells are more likely to continue straight during larger steps, but for a given step size, slow cells are more likely to turn sharply.

Given the relationship between speed and turning behavior, we would expect faster cells, which tend to make shallower turns, to have longer typical persistence times. We tested this by measuring the average persistence time along each trajectory (Materials and Methods), and found a strong positive correlation (Figure 2C). The relationship between speed and persistence time suggests that there may also be systematic differences in the scaling of the MSD at short times between cells. In particular, variation in speed alone amongst individuals would not change the shape of the MSD, which would collapse when appropriately scaled (Supplementary Text). On the other hand, the systematically longer persistence times of faster cells would be expected to boost the slope of their MSD at short times: faster cells appear more superdiffusive because they tend to wait longer to turn, an effect we observe in the data (Figure 2D).

The analysis at the level of speed classes suggested that there might be a single scalar variable, for which the cell's average speed is a good proxy, that determines a number of higher-order statistics characterizing the cell's migration behavior. To test this at the level of individual trajectories, we chose two summary statistics that capture the cell's turning behavior: the average of the cosine of the turn angles along the trajectory, and the correlation between speeds

and turn angles along the trajectory. The former is a summary of the overall distribution of turn angles for that cell, while the latter captures the degree of additional local coupling between speed and turn angle. Together with the cell's speed, these two summary statistics form a three-dimensional behavioral space. We observed that the cell trajectories fall close to a curve in this space (Figure 2E). In particular, 73% of the variance in the average cosine can be explained by cell speed, with some residual variance due to the stochasticity of the process (7%) and other unknown effects (20%) (Figure S1). Thus T cell migration statistics can be organized into a one-dimensional behavioral manifold, characterized by a strong dependence between speed and turning behavior.

Our observation of a behavioral manifold suggests that, despite the apparent heterogeneity in migration strategies, there may be a common program with a single underlying variable. In this view, a cell's location on the manifold reflects its internal value of this control variable, which in turn dictates its random walk behavior. We propose a speed-persistence coupling model (SPC), a particular constrained form of a persistent random walk, as a single-parameter toy model describing the migration program carried out by these cells (Supplementary Text).

As in other persistent random walk models, SPC walkers are diffusive at long times: the MSD scales linearly with time, and the ratio between these quantities defines an effective diffusion constant (Supplementary Text). The SPC model predicts a strong scaling of the effective diffusion constant with cell speed (Supplementary Text). We tested this prediction at several time intervals and found good quantitative agreement between the model and the data (Figure 2F). In particular, with a uniform persistence time (UPT), we would expect a 60-fold variation in diffusion constants across the trajectories (5th-95th speed percentiles), whereas the SPC model predicts and the data supports a 300-fold variation.

We note that the analyses above depend on measured cell speeds and turn angles, which are an imperfect proxy for the true instantaneous process. In particular, both noise in the cell locations and finite sampling intervals can introduce bias in the measured speeds, which could in principle generate spurious relationships between measured speed and turning behavior. We took two approaches to addressing the sensitivity of our conclusions to these issues. First, we addressed sensitivity to sampling rate by repeating the analyses above, subsampling timepoints by a factor of 2. This makes the turning behavior of the slowest two speed classes harder to distinguish, because they are rarely persistent over more than one timestep (Figure S2A,D), and introduces more noise in the local coupling (Figure S2B), but otherwise does not alter the structure of the correlations (Figure S2A-F). Second, we assessed the potential biases introduced by mislocation noise and finite sampling to the speed-persistence relationship in simulations (Supplementary Text, Figure S3). We found that mislocation noise can lead to spurious correlations between speed and persistence at the slow end of the speed spectrum, but cannot account for the consistent correlation we observe across speeds.

Our observations indicate that the unperturbed behavior of a population of T cells is heterogeneous, but that this heterogeneity may reflect different states of an underlying behavioral program. We next asked about the robustness of this behavioral program to a perturbation that affects cell speed. To do so, we first performed single-cell RNA sequencing on cells isolated

from the tail of 15 dpf Tg(*lck*:GFP) zebrafish to identify pathways that may be relevant to migration in this cell population. To assess the fidelity of the marker, we sorted GFP+ cells from an unbiased FSC/BSC gate (Materials and Methods, Figure S4). We used standard dimensional reduction and clustering methods (Materials and Methods) to identify 351 putative T cells (Figure 3A-B). Unexpectedly, we also identified a population of epithelial cells that may mis-express *lck* at low levels (Materials and Methods, Figures S5,S6).

The 50 most significantly differentially expressed genes (Wilcoxon rank-sum test) for the putative T cells include a number of T cell and immune markers (*23*) (Figure 3B). We note that we did not observe significant expression of markers associated with other motile non-T immune cells (B cells, macrophages, or neutrophils) (*23, 24*), and we do not find support for NK cells (Materials and Methods, Figure S6).

The 50 most significantly enriched genes also included several canonically involved in actin nucleation and remodeling in the leukocyte cytoskeleton (*25,26*) (WASP/ARP2/3 pathway; Figure 3B, Figure S6). Based on these results, we chose the drug Rockout, a known Rho kinase inhibitor affecting this pathway (*27*), as a candidate for perturbing cell speed. We repeated the measurements and analysis of cell migration behavior in the presence of the drug. We found that the distribution of cell speeds shifted downwards, but we still observed a quantitatively similar positive relationship between speed and turning behavior, and the turn angles of perturbed cells were similar to those of unperturbed cells of the same speed (Figure 3C, D). This is consistent with a model where the perturbation primarily shifted an internal cell state variable that determines location along the behavioral manifold, which in turn dictates both speed and turning behavior.

Finally, to address whether any of the features of cell migration behavior observed here might hold outside of our system, we used datasets from two published studies, one of mouse T cells *in situ* (*28*), and the other of *Dictyostelium* (*29*). While some of the analyses that depend on longer time traces and larger cell numbers are not possible with these datasets, we tested the relationship between average turn angle and cell speed, which drives many other differences in the dynamics. We found that this correlation held amongst the control cells in both studies (Figure 4). This suggests that, as in our dataset, there is heterogeneity in speed and turning behavior amongst the cells, and is consistent with a similar behavioral manifold. In the two published studies, genetic perturbations that knocked out or down one member of the actin remodeling machinery were used: a knockout of the non-canonical myosin Myo1g in one case, and a knockout of the Arp2/3 inhibitor Arpin in the other. In each case, the perturbation had a substantial effect on the distribution of cell speeds (Figure 4, C-D). However, in both cases, a quantitatively similar positive relationship between the speed and turning behavior amongst the perturbed cells was preserved.

Together with the small molecule perturbation used in this study, these observations suggest that there may be global coupling amongst some components of the actin remodeling machinery, so that diverse perturbations can lead to similar internal states.

We note that, in general, heterogeneity of motility behavior across a population could be caused by the tissue context rather than by cell-intrinsic factors. However, the effects of the

drug perturbation, as well as the effects of the genetic perturbations from (*28*) and (*29*), support a cell-intrinsic basis for the behavioral manifold we observe here. In particular, we performed trials in which the same regions of tissue were imaged and cells tracked before and after addition of the drug (Figure S7); the observed changes in migration statistics must then be caused by the drug's effect on the cell's internal state, not by the tissue context.

The genetic perturbations used in previous studies made the cells faster and more persistent on average (*28, 29*). Our results suggest that this connection may be general to the cells rather than specific to the perturbation. In particular, shifts in the average turning behavior have been used to argue that Arpin and Myo1g control cell steering. Our data suggests that increasing cell speed may in many cases increase straightness, and vice versa, so that the effect on cell steering may be indirect.

While our observations do not support Levy flight, the SPC model generates other ways for the cells to explore across a broad range of length scales. The steep scaling of the effective diffusion constant with speed leads to a more than 300-fold variation in effective diffusion constants across the unperturbed cell population. Together with extracellular signals that may shift cells along the behavioral manifold (*14, 30, 31*), this strategy could provide a regulated alternative to Levy flight that produces an efficient search *in vivo*.

References

1. M. Sarris, M. Sixt, *Current opinion in cell biology* **36**, 93 (2015).

2. C. M. Witt, S. Raychaudhuri, B. Schaefer, A. K. Chakraborty, E. A. Robey, *PLoS biology* **3**, e160 (2005).

3. R. N. Germain, E. A. Robey, M. D. Cahalan, *Science* **336**, 1676 (2012).

4. M. J. Miller, S. H. Wei, I. Parker, M. D. Cahalan, *Science* **296**, 1869 (2002).

5. M. J. Miller, S. H. Wei, M. D. Cahalan, I. Parker, *Proceedings of the National Academy of Sciences* **100**, 2604 (2003).

6. M. D. Cahalan, I. Parker, *Annu. Rev. Immunol.* **26**, 585 (2008).

7. J. B. Beltman, A. F. Marée, J. N. Lynch, M. J. Miller, R. J. de Boer, *Journal of Experimental Medicine* **204**, 771 (2007).

8. E. J. Banigan, T. H. Harris, D. A. Christian, C. A. Hunter, A. J. Liu, *PLoS computational biology* **11**, e1004058 (2015).

9. T. H. Harris, *et al.*, *Nature* **486**, 545 (2012).

10. T. Worbs, T. R. Mempel, J. Bölter, U. H. von Andrian, R. Förster, *Journal of Experimental Medicine* **204**, 489 (2007).

11. J. Textor, *et al.*, *Proceedings of the National Academy of Sciences* **108**, 12401 (2011).

12. C. Beauchemin, N. M. Dixit, A. S. Perelson, *The Journal of Immunology* **178**, 5505 (2007).

13. M. F. Krummel, F. Bartumeus, A. Gérard, *Nature Reviews Immunology* **16**, 193 (2016).

14. T. R. Mempel, S. E. Henrickson, U. H. Von Andrian, *Nature* **427**, 154 (2004).

15. M. F. Shlesinger, G. M. Zaslavsky, U. Frisch (1995).

16. F. Bartumeus, M. G. E. da Luz, G. M. Viswanathan, J. Catalan, *Ecology* **86**, 3078 (2005).

17. F. Bartumeus, J. Catalan, U. Fulco, M. Lyra, G. Viswanathan, *Physical Review Letters* **88**, 097901 (2002).

18. S. Petrovskii, A. Mashanova, V. A. Jansen, *Proceedings of the National Academy of Sciences* **108**, 8704 (2011).

19. P. G. Pitrone, *et al.*, *nature methods* **10**, 598 (2013).

20. R. M. Power, J. Huisken, *Nature methods* **14**, 360 (2017).

21. D. M. Langenau, *et al.*, *Proceedings of the National Academy of Sciences* **101**, 7369 (2004).

22. G. Viswanathan, E. Raposo, F. Bartumeus, J. Catalan, M. Da Luz, *Physical Review E* **72**, 011111 (2005).

23. N. Schaum, *et al.*, *Nature* **562**, 367 (2018).

24. Q. Tang, *et al.*, *Journal of Experimental Medicine* **214**, 2875 (2017).

25. M. Vicente-Manzanares, D. Sancho, M. Yáñez-Mó, F. Sánchez-Madrid, *International review of cytology* (Elsevier, 2002), vol. 216, pp. 233–289.

26. T. Takenawa, S. Suetsugu, *Nature reviews Molecular cell biology* **8**, 37 (2007).

27. F. Barros-Becker, P.-Y. Lam, R. Fisher, A. Huttenlocher, *J Cell Sci* **130**, 3801 (2017).

28. A. Gérard, *et al.*, *Cell* **158**, 492 (2014).

29. I. Dang, *et al.*, *Nature* **503**, 281 (2013).

30. H. D. Moreau, *et al.*, *Proceedings of the National Academy of Sciences* **112**, 12151 (2015).

31. M. L. Dustin, S. K. Bromley, Z. Kan, D. A. Peterson, E. R. Unanue, *Proceedings of the National Academy of Sciences* **94**, 3909 (1997).

32. C. Sommer, C. Straehle, U. Koethe, F. A. Hamprecht, *Biomedical Imaging: From Nano to Macro, 2011 IEEE International Symposium on* (IEEE, 2011), pp. 230–233.

33. G. E. Uhlenbeck, L. S. Ornstein, *Physical review* **36**, 823 (1930).

34. M. Doi, S. F. Edwards, *The theory of polymer dynamics*, vol. 73 (oxford university press, 1988).

35. P.-H. Wu, A. Giri, S. X. Sun, D. Wirtz, *Proceedings of the National Academy of Sciences* **111**, 3949 (2014).

36. D. T. Gillespie, *Physical review E* **54**, 2084 (1996).

Acknowledgments

We acknowledge Aya Ludin-Tal and Leonard Zon for the generous gift of the Tg(*lck*:GFP) zebrafish line. We acknowledge Kiran Kocherlakota and the Stanford VSC for assistance with zebrafish management and husbandry. We acknowledge Saroja Korullu for assistance with library preparation. We also acknowledge Stanford Research Computing and the Sherlock2 computer cluster for computational support and resources. Finally, we acknowledge Louis Leung and Karen Mruk for invaluable advice regarding microscopy and zebrafish; Edward Marti for helpful discussions on imaging, analysis, and the manuscript, and Felix Hornes and Michael Swift for comments on the manuscript. **Funding**: This project was supported by the Chan Zuckerberg Biohub; Author contributions: E.J. and S.R.Q. designed the research; E.J. performed experiments and analysis; E.J. and S.R.Q. wrote the paper. **Competing interests**: Authors declare no competing interests. **Data and materials availability**: Sequencing data and the gene expression count table have been deposited on GEO (accession: GSE137770). Analysis code and trajectory data are available at https://github.com/erjerison/TCellMigration.

Supplementary materials

Materials and Methods
Supplementary Text
Figs. S1 to S7
Movies S1,S2
References (*32–36*)

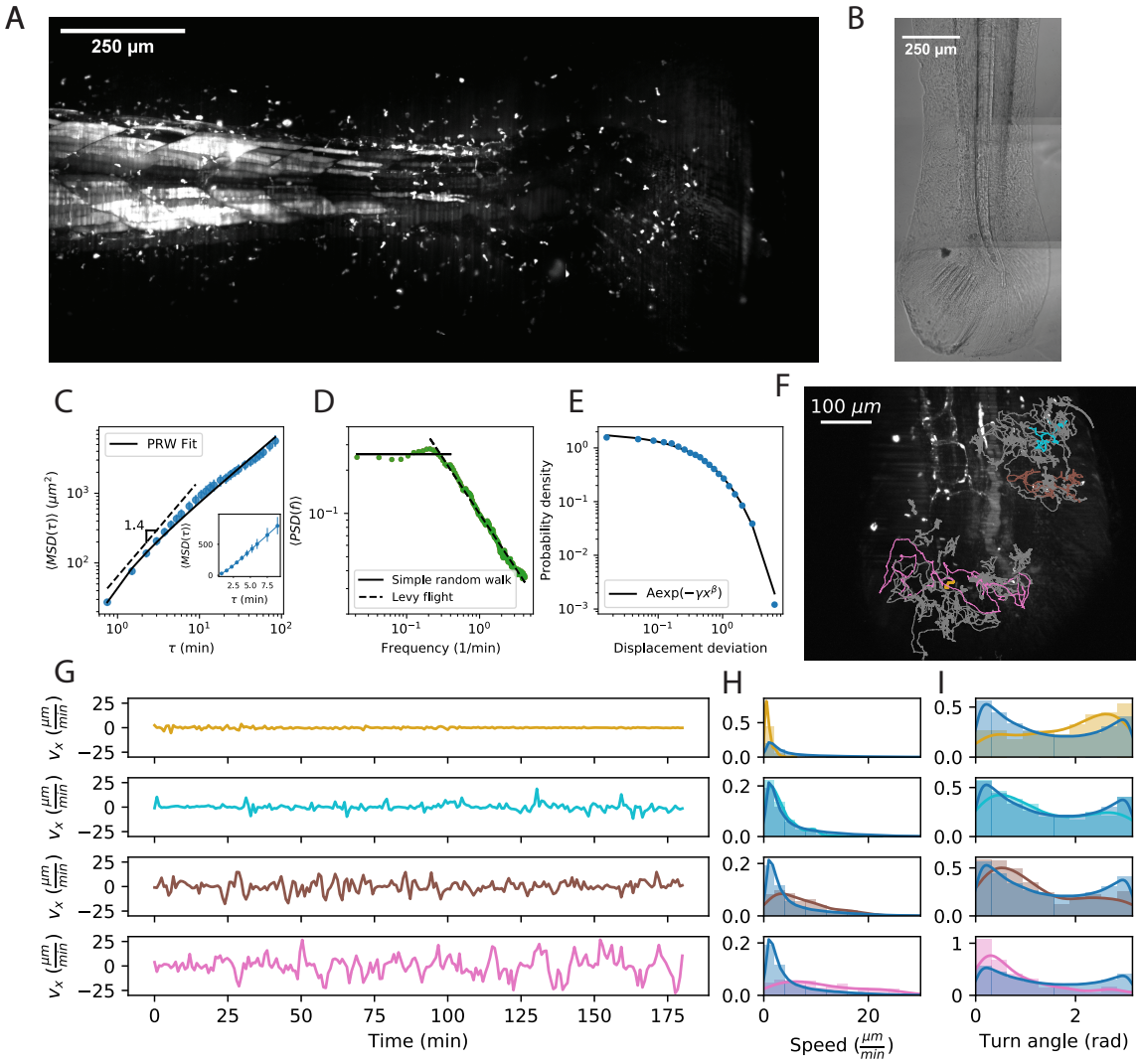

Figure 1: Heterogeneity in speed and turning angles in T cell migration dynamics. A. Maximum Z projection of a Tg(*lck*:GFP, *nacre*-/-) zebrafish at 12 dpf. This projection represents the first frame of a timecourse; see Movie S1. B. Brightfield of the region of tissue shown in A. C. Mean squared displacement as a function of time lag. The cells migrate super-diffusively on scales of a few minutes. The MSD for a persistent random walk is fit to the data (Materials and Methods, Supplementary Text). Error bars represent 95% confidence intervals on a bootstrap over trajectories. Inset: linear scale for the first 10 minutes. D. The velocity power spectrum, averaged across all trajectories. A Levy (scale-free) process consistent with the short time behavior would result in a continuation of the high frequency slope (dashed line). Instead, we observe a timescale at a few minutes. E. Distribution of bout lengths within a trajectory (Materials and Methods), fit with a stretched exponential. F. Example of trajectories recorded over 3 hours at a 12 second interval (Tg(*lck*:GFP, *nacre*-/-) zebrafish; 10 dpf). Here we show a maximum Z projection of the 900th frame with trajectories overlaid; see Movie S2 for the timecourse. Examples of four cell trajectories, with a range of characteristic speeds, are colored. G. Velocity traces for the four cells highlighted in E, along with the secant-approximated speed and turn angle distributions for each cell, compared with the distribution over all cells (blue).

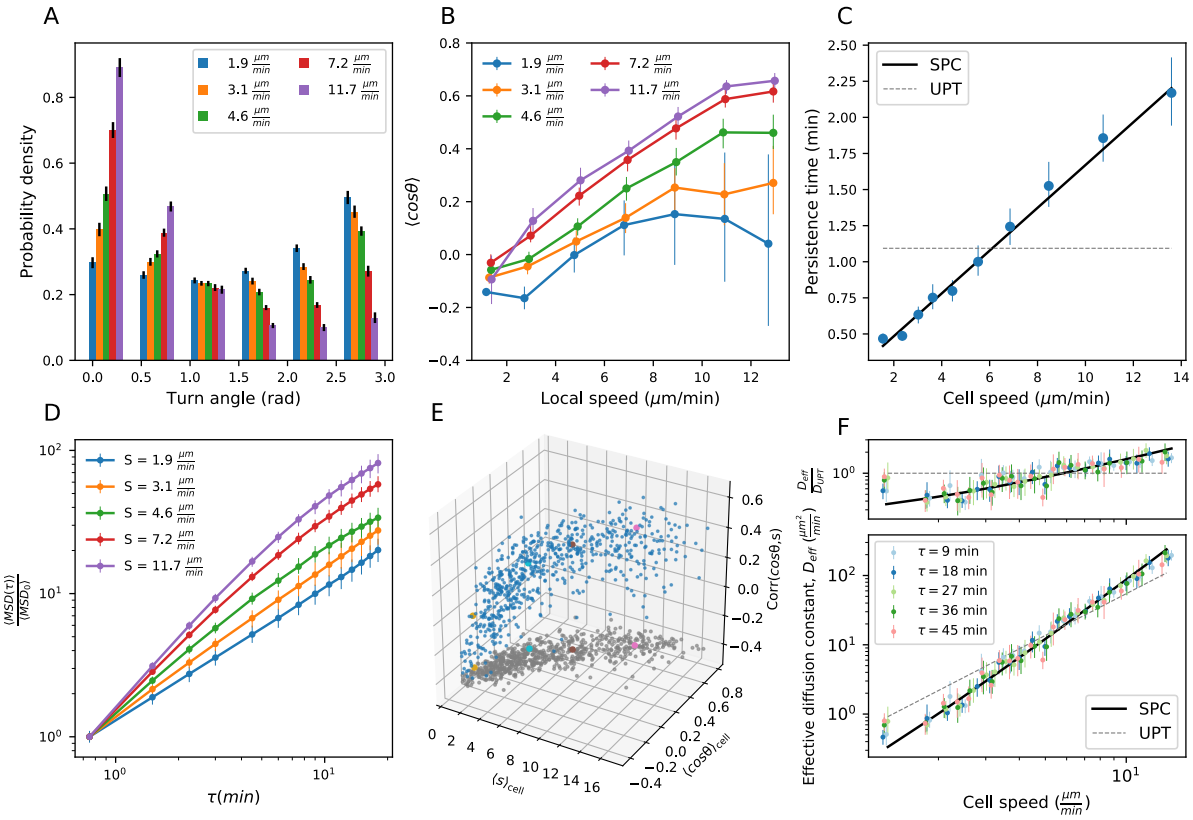

Figure 2: Heterogeneous migration strategies fall on a behavioral manifold A. Distribution of turn angles amongst cells grouped by speed class. The distribution varies smoothly from faster cells, which tend to go straighter, to slower cells, which tend to turn around more often. Error bars represent 95% confidence intervals from a bootstrap over trajectories in each speed class. The legend reports the mean speed for trajectories in each class. B. Turning behavior conditioned on current cell speed. The average of the cosine of the turn angle as a function of the average length of the steps on either side. Cells are grouped into speed classes as in A. Error bars represent 95% confidence intervals from a bootstrap over trajectories in each speed class. C. Mean persistence time as a function of cell speed, measured along trajectories. Error bars represent 95% confidence intervals from a bootstrap over trajectories. UPT: Uniform persistence time; SPC: Speed-persistence coupling (see Supplementary Text). D. Mean squared displacement by speed class. Due to the variation in turning behavior, the faster cells appear initially more superdiffusive. Error bars represent 95% confidence intervals from a bootstrap over trajectories in each speed class. E. Organization of cell behavior into a curve in a three dimensional behavioral space. Each point represents a trajectory, and we show the average speed, turn angle, and local speed-turn correlation. Grey: projection into the x-y plane. The trajectories shown in Figure 1F-H are colored. F. Scaling of the effective diffusion constant (see Materials and Methods) with cell speed. A UPT (uniform persistence time) model predicts an S2 scaling, whereas the SPC model predicts a steeper scaling. Except for a constant offset, parameters are fixed based on the speed-persistence relationship in C (see Materials and Methods). Error bars represent 95% confidence intervals on a bootstrap over cells.

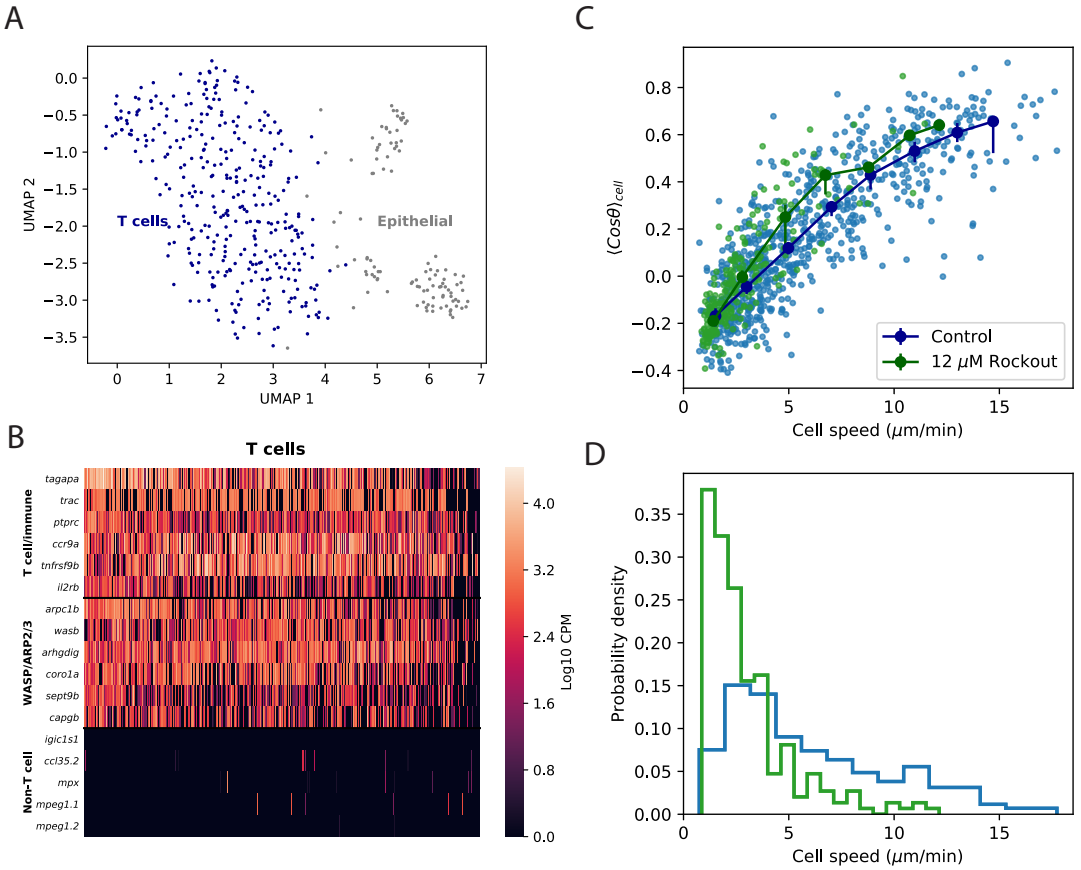

Figure 3: A small molecule perturbation preserves the speed-persistence relationship. A. Dimensional reduction via UMAP of scRNAseq gene expression profiles for cells sorted from 15 dpf Tg(*lck*:GFP) zebrafish. We called 351 putative T cells and 110 putative epithelial cells (Materials and Methods). B. Marker gene expression in the T cell cluster. Cells from the T cell cluster express a number of T cell and ubiquitous immune markers, including the T cell receptor light chain, *trac*, as well as genes involved in actin nucleation and remodeling. With the exception of *trac*, the genes shown in these first two categories were amongst the top 50 differentially expressed genes for the T cell cluster (Wilcoxon rank-sum test). Finally, we do not observe significant expression of immune markers associated with B cells, macrophages, or neutrophils. C. Correlation between average cosine of the turn angle along the trajectory and cell speed, for cells in control and Rockout-treatment conditions. Data for all cells is shown as well as a binned average. Error bars represent 95% confidence intervals on the binned average on a bootstrap over cells. D. The distribution of speeds amongst control and Rockout-treated trajectories. The treatment lowers cell speeds but maintains the relationship between speed and persistence. See also Figure S7.

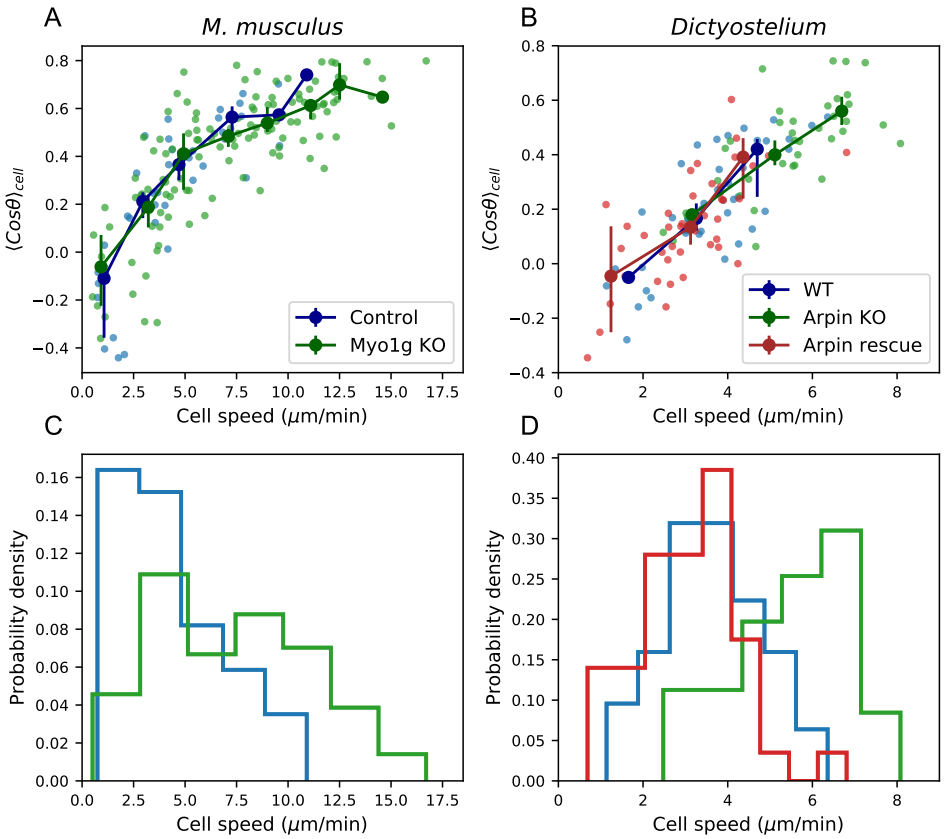

Figure 4: Angle-speed correlation in mouse T cells and *Dictyostelium*. A. As in 3B-C, for mouse T cells (data from (*28*)). The perturbation is a genetic knockout of a non-canonical myosin motor, Myo1g. B. Same as in A, for *Dictyostelium* (data from (*29*)). The perturbations are a knockout and rescue of the Arp2/3 inhibitor Arpin. C-D. Distributions of cell speeds for the control and treatment conditions shown in A-B. In each case, the distribution of speeds shifts, but the cells tend to move along the speed-turn curve.