A new method shows that a minority of liver-localized CD8 T cells display hard-to-detect attraction to Plasmodium-infected hepatocytes

Viktor S. Zenkov[1*], James O'Connor[2,3], Ian Cockburn[2] and Vitaly V. Ganusov[4,5*]

[1] *Electrical Engineering and Computer Science, University of Tennessee, Knoxville, TN 37996, USA*
[2] *Department of Pathogens and Immunity, John Curtin School of Medical Research, Australian National University, Canberra, 2600, Australia*
[3] *The Australian National University Medical School, Canberra, ACT 2601, Australia*
[4] *Department of Microbiology, University of Tennessee, Knoxville, TN 37996, USA*
[5] *Department of Mathematics, University of Tennessee, Knoxville, TN 37996, USA*
[*] *Corresponding authors: vzenkov@vols.utk.edu, vitaly.ganusov@gmail.com*

August 3, 2021

Abstract

Malaria is a disease caused by *Plasmodium* parasites, resulting in over 200 million infections and 400,000 deaths every year. A critical step of malaria infection is when mosquito-injected sporozoites travel to the liver and form liver stages. Malaria vaccine candidates tested in mice which induce large numbers of malaria-specific CD8 T cells are able to eliminate all liver stages, preventing fulminant malaria. However, how CD8 T cells find all parasites in 48 hours of the liver stage lifespan is not well understood. Using intravital microscopy of murine livers, we generated unique data on T cell search for malaria liver stages within a few hours after infection. To detect attraction of T cells to an infection site, we used the Von Mises-Fisher distribution in 3D, similar to the 2D von Mises distribution previously used in ecology, to detect attraction of searching T cells towards the infection site. Our results suggested that the vast majority (70-95%) of malaria-specific CD8 T cells and non-specific T cells did not display attraction towards the infection site, suggesting that the search for malaria liver stages occurs randomly. However, a small fraction (15-20%) displayed weak but detectable attraction towards parasites which already had been surrounded by several T cells. We found that speeds and turning angles correlated with attraction, suggesting that understanding mechanisms determining the speed of T cell movement in the liver may improve T cell vaccine efficacy. Stochastic simulations suggested that a small movement bias towards the parasite dramatically reduces the number of CD8 T cells needed to eliminate all malaria liver stages, but to detect such attraction by individual cells requires extremely long imaging experiments which are not currently feasible. Importantly, this is the first demonstration that we know about of how activated/memory CD8 T cells search for the pathogen in nonlymphoid tissues few hours after infection. We have also established a framework for how attraction of individual T cells towards a location in 3D can be rigorously evaluated.

Abbreviations: vMF - von Mises-Fisher.

1 Introduction

Malaria is a disease caused by parasites of the genus *Plasmodium* that kills over 400,000 people every year[1]. Mosquitoes carrying malaria sporozoites inject sporozoites when searching for blood[2]. Sporozoites travel through the bloodstream to the liver, invade hepatocytes, and form liver stages. The liver stage of malaria infection is asymptomatic. The development of the liver stage takes 48 hours in mice and 7 days in humans[3–6]. Importantly, vaccines inducing exclusively malaria-specific CD8 T cells are capable of removing all liver stages in mice, thus preventing clinical malaria[7–9]. Intravital imaging experiments have shown that soon after infection, activated Plasmodium-specific CD8 T cells and CD8 T cells of irrelevant specificity cluster around Plasmodium-infected hepatocytes[10,11]. Mathematical modeling-based analysis of the T cell clustering data suggested that the formation of the clusters is best explained by a model in which the first T cell finds the liver stage randomly and the attraction of other T cells (including T cells with irrelevant specificity) to the parasite increases with the number of T cells per cluster[11]. While an earlier study suggested that there may be attraction of distant T cells to the clustered liver stage[12], it still remains unclear if attraction to the liver stage occurs prior to the formation of T cell cluster around the parasite or if attraction is exhibited by all or just a subset of cells.

Based on intravital imaging experiments, in many if not most analyzed cases, T cells appear to move randomly in tissues *in vivo*[13]. Few studies have accurately quantified if motile T cells exhibit biased migration. In part, this is because of the difficulty in quantifying bias in movement patterns of T cells and relating the movement bias to specific structures in the tissue. Previous studies on the movement of naive B cells in B cell zones of the lymph nodes showed a difference in migration towards the boundary of the follicle, or the plane separating light and dark zones of the germinal center reactions[14,15]. A more rigorous analysis of movement of virus-specific CD8 T cells in the skin demonstrated attraction of T cells to an infection site[16]. However, these studies employed simple metrics to measure attraction (e.g., percent of cells moving to a particular area or the distance between a cell and the boundary) and in most cases estimated bias involving a comparison group. Also, there was no analysis into if any of these metrics have biases. For example, T cells tend to undergo a persistent random walk in which a movement's direction tends to be similar to the previous movement's direction[17], which may generate an illusion of a bias towards a specific location. While methods to detect attraction to specific locations have been proposed and studied extensively in ecology[18–20], ecological movement data are typically 2D and whether previously proposed methods apply to 3D situations (e.g., for T cells moving in the liver) has not been thoroughly investigated.

At a more fundamental level, how vaccine-induced T cells search for in vivo sites of pathogen replication in peripheral tissues within hours after infection is largely unknown. This remains an experimental challenge because at early time points individual pathogens may have a low fluorescent signal to be easily detectable experimentally, and it is difficult to localize memory CD8 T cells near sites of infection. To address this fundamental knowledge gap and to understand how CD8 T cells search for the Plasmodium liver stages, we performed novel intravital microscopy-based experiments with murine livers in which we tracked positions of liver-localized fluorescently labeled malaria-specific CD8 T cells, CD8 T cells of irrelevant specificity, and malaria liver stages. In our experiments, Plasmodium sporozoites expressed GFP at a sufficiently high level so that individual parasites could be followed in the liver minutes to hours after intravenous inoculation[21].

To evaluate if T cells display attraction towards the infection, we reviewed previously-used metrics and develop a novel metric based on the von Mises-Fisher (vMF) distribution of angles to the parasite,

which is more powerful (requires less data to detect deviations from random movement) than several previously used metrics. We showed that one intuitive metric based on the change in distance between the T cells and the parasite is biased, and thus underestimates the degree of attraction unless corrected. Importantly, the vMF distribution-based metric can be also slightly biased for T cells with a correlated/persistent random walk, but this bias may be approximately corrected using simulations. Our results suggest that the majority of CD8 T cells do not display attraction towards the infection site, irrespective of the existence of a cluster of CD8 T cells near the liver stage. However, liver stages with an already formed T cell cluster tend to display strong attraction for a subset of cells. Stochastic simulations based on the vMF distribution suggest that the detection of weak attraction of individual cells towards the infection site requires amounts of data that are difficult-to-impossible to collect with current intravital imaging protocols. Our results thus establish a rigorous framework for how the attraction of moving cells towards a particular location can be evaluated.

2 Materials and Methods

2.1 Experimental design

Our data consists of 3D positions over time of CD8 T cells specific for malaria sporozoites; malaria liver stages; and in some cases, control CD8 T cells specific to irrelevant antigens. We analyzed datasets from five sets of experiments: three datasets generated for this analysis and two from a published study[10].

For the new **"unclustered/small clustered"** (dataset #1) and **"large clustered"** (dataset #2) datasets, mice were infected with GFP-expressing *Plasmodium berghei* (Pb) sporozoites, which carry the SIINFEKL epitope from OVA, Pb-CS^{5M}[22,23]. Activated CD8 T cells, specific to SIINFEKL epitope (OT1), were generated *in vitro* by co-culture of naive TCR transgenic CD8 T cells with SIINFEKL peptide as described previously[23]. As a control, we used TCR-transgenic CD8 T cells, specific to GP33 epitope from LCMV (P14) that were activated similarly to the OT1 cells. B6 mice received 5×10^6 activated Pb-specific (OT1) or LCMV-specific (P14) CD8 T cells and then 1.5-2 hours later, mice were infected with 10^5 Pb-CS^{5M} sporozoites (Figure S1). We performed imaging between 30 min and 2 hours after sporozoite infection using a two-photon microscope[23]. The unclustered/small clustered dataset experiments featured no or few T cells located within 40 μm from the parasite at the beginning of the experiment, and the large clustered dataset experiments featured several T cells located within 40 μm from the parasite (i.e., there was a T cell cluster[10]). We used the 40 μm diameter to distinguish closeness because it represents the average width of a standard murine hepatocyte[11]. The unclustered/small clustered dataset contains 3D coordinates over 3 hours (with timesteps of 1.5 or 2 minutes between recorded volumes) of Pb-specific (OT1) and LCMV-specific (P14) CD8 T cells in 4 mice (1 parasite/mouse, Figures S1 and S2). The large clustered dataset contains 3D coordinates over 3 hours (with timesteps of 1 or 2 minutes between recorded volumes) of Pb-specific and LCMV-specific CD8 T cells for 3 parasites in 3 mice. As a control, we also performed an experiment in which naive B6 mice received 5×10^6 activated Pb-specific (OT1) cells and we imaged 1.5-2 hours after T cell transfer (no infection was given). This **"control"** dataset (dataset #3) contains 3D coordinates over 30 minutes (with timesteps of 30 seconds between recorded volumes) for one parasite in one mouse.

We also analyzed datasets from our previous study[10]. To generate the **"Paris"** dataset (dataset

#4, named for the location of the experimentation), the following experimental set-up was used. Activated CD8 T cells specific for *Plasmodium yoelii* (Py) sporozoites (PyTCR) were generated *in vivo* by infecting Balb/c mice with the Vaccinia virus expressing the circumsporozoite (CS) protein from Py[10]. Balb/c were first infected with Py sporozoites, then 20 hours later, activated Py-specific (PyTCR) CD8 T cells (5×10^6 per mouse) were transferred to the infected mice intravenously. Imaging of the livers of infected mice was performed 6 hours after the T cell transfer[10]. To generate the **"co-clustered"** dataset (dataset #5), PyTCR cells were activated as described above. In addition, CD8 T cells specific to chicken ovalbumin (OT1) were activated by infecting B6 mice with the Vaccinia virus expressing OVA. CB6 mice (F1 progeny of B6 and Balb/c mice) were infected with 10^5 Py sporozoites, 20 hours later activated Py-specific (PyTCR, 5×10^6 per mouse) and OVA-specific (OT1, 5×10^6 per mouse) CD8 T cells were transferred into infected mice. Six hours later, livers of these mice were imaged. In both experiments, intravital imaging was done using spinning-disk confocal microscopy. The Paris dataset contains 3D coordinates over 5 hours (with timesteps of 1, 2, or 4 minutes between recorded volumes) of Py-specific CD8 T cells for 26 parasites in 4 mice. The co-clustered dataset contains 3D coordinates over 40 minutes (with timesteps of 2 minutes between recorded volumes) of Py-specific and OVA-specific CD8 T cells for one parasite in one mouse.

All 3D coordinate data are available as a supplement to this paper. We have also provided a Mathematica-based script that allows one to estimate the bias of moving agents to a point or a plane based on our newly developed vMF distribution-based metric, at `https://github.com/viktorZenkov/measuringAttraction/tree/master/Measuring`.

2.2 Metrics

To measure T cell *bias* (attraction or repulsion) towards the parasite, we define the following 3 metrics (Figure S3).

1. Angle metric (metric 1). For every movement of a cell, we calculate the angle between the cell's movement angle and the angle to the parasite (Figure S3A). An acute angle corresponds to the T cell "getting closer", and an obtuse angle corresponds to the T cell "getting farther". For an unbiased cell making n movements, the choices of closer/farther for the movements are given by a binomial distribution with $p = 0.5$. This metric has been extensively used in ecology[19].

2. Distance metric (metric 2, with improvements discussed below). For every movement of a cell, we calculate the change in the distance between the cell's first position and the parasite ($D0$) and the distance between the cell's next position and the parasite ($D1$, Figure S3B). A negative change in distance ($r = D0 - D1$) corresponds to "getting closer" and positive corresponds to "getting farther. For an unbiased cell making n movements, the choices of closer/farther for the movements are given by a Poisson Binomial distribution with $p = 0.5 - r/4x$, where r is the cell's movement length and $x = D0$ is the initial distance between the cell and the parasite (Figure 1)[24].

3. Angle distribution metric (metric 3). For every movement of a cell, we calculate the angle between the cell's movement vector and the vector to the parasite and compare all angles with the Von Mises-Fisher (vMF) distribution (Figure S3C and eqn. (1))[25,26]. By fitting the vMF distribution to angle data, we calculate a concentration parameter κ which indicates the strength of a cell's attraction towards the infection site, with $\kappa > 0$ indicating attraction and

$\kappa < 0$ indicating repulsion. A 2D version of a similar, von Mises distribution-based metric, also has been used in ecology[20].

4. Average angle metric (metric 4). For every movement of a cell, we calculate the angle between the cell's movement vector and the vector to the parasite[27]. We then use a Student T test to compare the mean of the angles to the expected average angle, 90 degrees[16,27].

2.3 Von Mises-Fisher distribution

To quantify the degree of bias (or absence thereof) we propose using the Von Mises-Fisher (vMF) distribution of angles towards the parasite[28]. The vMF distribution describes a probability distribution on an n-dimensional (we use $n = 3$) vector given a direction vector and a concentration parameter κ. The output vector is chosen randomly with a bias toward the given direction whose strength depends on κ. When $\kappa \to 0$, the vMF distribution approaches a uniform distribution; $\kappa > 0$ indicates positive bias (attraction); and $\kappa < 0$ indicates negative bias (repulsion, Figure S4). We simplify the vMF distribution from a vector to a single angle between the output vector and the given vector - this angle ϕ corresponds to our angle metric. The probability density function of the angles of the vMF distribution with respect to the point of attraction is

$$P(\phi|\kappa) = \frac{\kappa \sin(\phi) e^{\kappa \cos(\phi)}}{2 \sinh(\kappa)}, \qquad 0 \leq \phi \leq \pi, \tag{1}$$

where ϕ is the angle between the vector to the attraction point and the cell's movement vector. More intuitive parameters such as the fraction of acute angles f_a or the average angle towards infection $\bar{\phi}$ can be calculated as well if needed. Multiple biases, such as towards previous movement vectors or to the infection site, can be further incorporated using the vMF distribution, and the standard likelihood method can be used to estimate its parameters (see Supplemental Information for more detail).

We use a log likelihood test to determine the concentration parameter for the angle version of the Von Mises-Fisher distribution most likely to generate the directions corresponding to the angles eqn. (S5). If this distribution's test statistic is significantly different from a random test statistic, then that corresponds to biased movement, and if it is not significantly different that corresponds to unbiased movement. If a cell is biased with $\kappa_a > 0$, then the cell is attracted, and if a cell is biased with $\kappa_a < 0$, then the cells is repulsed.

2.4 Simulations

To understand various aspects of T cell movement with respect to the parasite, we performed stochastic simulations in Mathematica 12.0. In simulations, cell movements are characterized by a distance traveled per time step and the direction of cell movement (with respect to the previous movement and/or to the parasite). Movement lengths were chosen from a Generalized Pareto (Pareto type IV) distribution, and the direction of the movement was chosen from a vMF distribution with a given concentration parameter κ (see Supplemental Information for more detail).

3 Results

3.1 Measuring attraction using the distance metric is biased unless corrected

There have been few studies that rigorously addressed the question of how to detect attraction of a moving lymphocyte towards an infection site. Commonly, two metrics or their variations have been used: the angle-based metric and the distance-based metric[12,16,19]. According to these metrics, cells moving towards the parasite exhibit acute angles to the parasite and reduction in distance to the parasite (Figure S3A&B). The natural assumption is that randomly moving (unbiased) cells should exhibit a similar number of movements towards and away from infection. However, as far as we know this was not thoroughly explored previously.

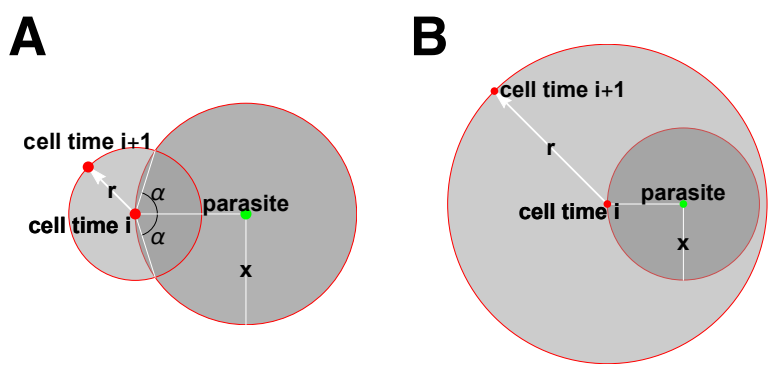

Figure 1: Graphical demonstration of the bias in the distance metric to measure attraction to the infection site. The probability that a T cell moves closer to the parasite is lower than the probability that the T cell moves away from the parasite, independently of if movement length r is shorter (panel A) or longer (panel B) than the distance x between T cell and the parasite. The probability that the distance between a randomly moving T cell and the parasite declines with the movement is $P(\Delta D < 0) = 1/2 - r/(4x)$ for $r < x$ and is $P(\Delta D < 0) = 0$ if $r > x$.

We performed stochastic simulations in which T cells searched for the infection site to evaluate the potential bias of T cells towards the infection. Simulations showed that while the angle metric is not biased (i.e., for randomly moving T cells, there are similar proportions of acute and obtuse angles to infection), the distance metric is biased (i.e., there are more movements away from the infection than towards the infection, results not shown). (Note that this usage of the word "bias" refers to a skewed result, unlike our other usage, which refers to cell movement being attracted or repulsed.) Graphing movement possibilities of T cells with respect to the infection site revealed that for a finite distance between the cell and the infection, there are more chances for the T cell to move away than to get closer (Figure 1). Defining r as the distance that the cell moves and x as the distance from the cell to the parasite, the probability that the distance decreases for a T cell movement is $\frac{1}{2} - \frac{r}{4x}$. The subset of the surface of the r-sphere located inside the x-sphere are the positions for which the cell's distance to the parasite gets smaller, and the subset of the surface of the r-sphere located outside the x-sphere are the positions for which the cell's distance to the parasite gets larger. Note that if $r > 2x$, causing the fraction to be less than 0, then the cell moves so far that none of its potential destinations are closer to the parasite than its current location, which means the probability that the distance gets smaller is 0 (Figure 1).

While the distance metric is biased, the bias can be corrected by taking into account the actual

probability that the T cell moves towards the infection. The probability that k movements (out of n) are towards the parasite can be calculated using the Poisson binomial distribution:

$$P(k) = \sum_{A \in F_k} \prod_{i \in A} p_i \prod_{j \in A^c} (1 - p_j), \tag{2}$$

where F_k is the set of all subsets of k integers that can be selected from $1, 2 \ldots n$ and A^c is the complement of A, $p_i = 1/2 - r_i/(4x_i)$ and r_i is the length of the ith cell movement and x_i is the distance between the T cell and the parasite before the ith movement (Figure 1). With this correction, the angle-based and distance-based metrics are numerically indistinguishable.

3.2 Novel angle distribution-based metric to detect attraction

One of the limitations of the angle- and distance-based metrics is that the actual value of the angle/distance is ignored and converted into a boolean variable, thus perhaps reducing the power of detecting bias in cell movement. To account for the actual values of the angles to infection, we propose using the von Mises-Fisher (vMF) distribution (see Materials and Methods). The vMF distribution is a natural extension of the von Mises distribution used in ecology, which is restricted to describe bias in 2D[20]. In the vMF distribution (eqn. (1)), bias towards a particular direction is determined by the concentration parameter κ_a and both attraction ($\kappa_a > 0$) and repulsion ($\kappa_a < 0$) can be described. To test if this new metric has a higher power than the angle/distance-based metrics, we performed stochastic simulations in which the attraction of T cells towards an infection site was given by the vMF distribution with different concentration parameters. Analysis showed that for the same number of movements (data points) using the vMF distribution-based metric, we were able to detect a significantly smaller bias in movement, and for the same degree of attraction, the vMF distribution-based metric required less data (i.e., was more powerful, Figure S5). Additionally, the vMF distribution-based metric was more powerful than a metric based on the average angle to the parasite that has also been used previously[16].

The vMF distribution-based metric is not biased for cells moving randomly with respect to a cell's previous movement. However, in many situations T cells undergo a correlated random walk where the cell moves in a direction correlated with the previous movement direction[17]. To investigate if a correlated random walk may introduce bias in estimating attraction towards the infection site, we performed a set of simulations in which the cell's movement was described by a vMF distribution with the angle ϕ determining the cell's turning angle and the concentration parameter κ_t indicating the degree of the cell's persistence in the correlated random walk (see Supplemental Information for more detail). Analysis suggested that a correlated random walk can introduce a biased result by defining some cells as attracted to the infection site (or repulsed from the infection site) even though there was no actual attraction or repulsion (Figure S6). Importantly, the fraction of cells detected as biased depended on the degree of walk persistence and the size of the area in which the movement of T cells was considered, and in general, it was easier to detect cells as repulsed than as attracted (Figure S6). This result arose due to an artifact that cells moving away from the parasite tend to continue to move away, while cells moving towards the parasite tend to pass the parasite and then also move away. We could not derive an analytical expression to correct for the bias in detecting attraction of cells moving in a correlated random walk. However, by simulating the movement of cells using the cell's initial position, movement length distribution, and turning angles as observed in the data, but with no attraction, it is possible to determine the bias of T cell movement towards

the infection site in the absence of actual attraction, κ_a^0. We used such a correlation in our test of whether detected bias κ_a in actual data was significantly different from κ_a^0 found in stochastic simulations given cells' movement properties.

3.3 Only a small proportion of activated liver-localized CD8 T cells display attraction towards malaria liver stages

To determine if liver-localized CD8 T cells display attraction towards the malaria liver stages, we performed novel experiments in which we tracked positions of malaria-specific CD8 T cells, CD8 T cells with irrelevant specificity, and malaria liver stages over time in murine livers using intravital microscopy (see Materials and Methods for more details). We attempted to design our experiments to detect the event in which the first T cell locates the infection site, and such an event did occur in one experiment (see below). In all other experiments, by the starting time of imaging, a small (1-2) or large number (5-7) of T cells had already formed a cluster around the liver stage (Figure S9); in general imaging experiments performed later after the sporozoite infection resulted in larger clusters (results not shown). Notably, T cells displayed different movement characteristics depending on whether there were no/few or many cells in the cluster (Figure 2i-ii). Specifically, in unclustered/small clustered data, cell speeds were 2.42 and 2.63 μm/min for OT1 and P14 cells, respectively (Figure S10). And interestingly, in the large clustered data, both cell types were significantly slower, with speeds of OT1 and P14 cells being 1.55 and 1.59 μm/min, respectively (Mann-Whitney test, $p < 0.001$). In the presence of large clusters, T cells were also likely to have a larger arrest coefficient (fraction of cell movements with a speed below 1μm/min): 0.29 and 0.27 for OT1 and P14 cells, respectively, in unclustered/small cluster data vs. 0.47 for both cell types in large cluster data. Additional analysis based on the meandering index (the distance between the first and last recorded positions divided by the total length of the path) and turning angles for T cells showed that all cells tend to turn (Figure S11) suggesting an active search for an infection. In our other two independent datasets (the "Paris" and "co-clustered" datasets), we observed similar cell behaviors (results not shown).

To understand how T cells search for the infection site, we first pooled all track data from a given dataset into one set and determined if T cells display attraction towards the infection by fitting a vMF distribution to the data and estimating the concentration parameter κ_a (see Materials and Methods for more detail). We found that T cells displayed a statistically significant but extremely weak attraction towards the parasite ($\kappa_a = 0.077$ ($p = 3.8 \times 10^{-4}$) and $\kappa_a = 0.085$ ($p = 3.9 \times 10^{-5}$) for the unclustered/small clustered and large clustered datasets, respectively); concentration parameter $\kappa_a = 0.077$ corresponds to only 52% of movements being towards the parasite. There was stronger attraction of Plasmodium-specific CD8 T cells towards the infection site detected in the other two datasets ($\kappa_a = 0.19$ ($p = 0.003$) and $\kappa_a = 0.51$ ($p = 0.001$), in Figures S8 and S18).

To investigate if weak attraction towards the parasite comes from all cells exhibiting weak bias or a few cells exhibiting strong bias, we calculated κ_a for individual cells. There was a broad distribution of κ_a for individual cells, but most of these values were not statistically different from a random value with the average indicating attraction (Figure 2iii-iv and Figure S8iii-iv). It should be noted that because the vMF distribution-based metric is inherently slightly biased for T cells undergoing a correlated random walk, for every T cell track we calculated the null hypothesis concentration parameter κ_a^0 that would be expected given the cell's initial position (with respect to the parasite and imaging area), turning angle distribution, and movement lengths, and statistically compared the true κ_a to the null hypothesis value κ_a^0 (results not shown); in most cases the null hypothesis κ_a^0

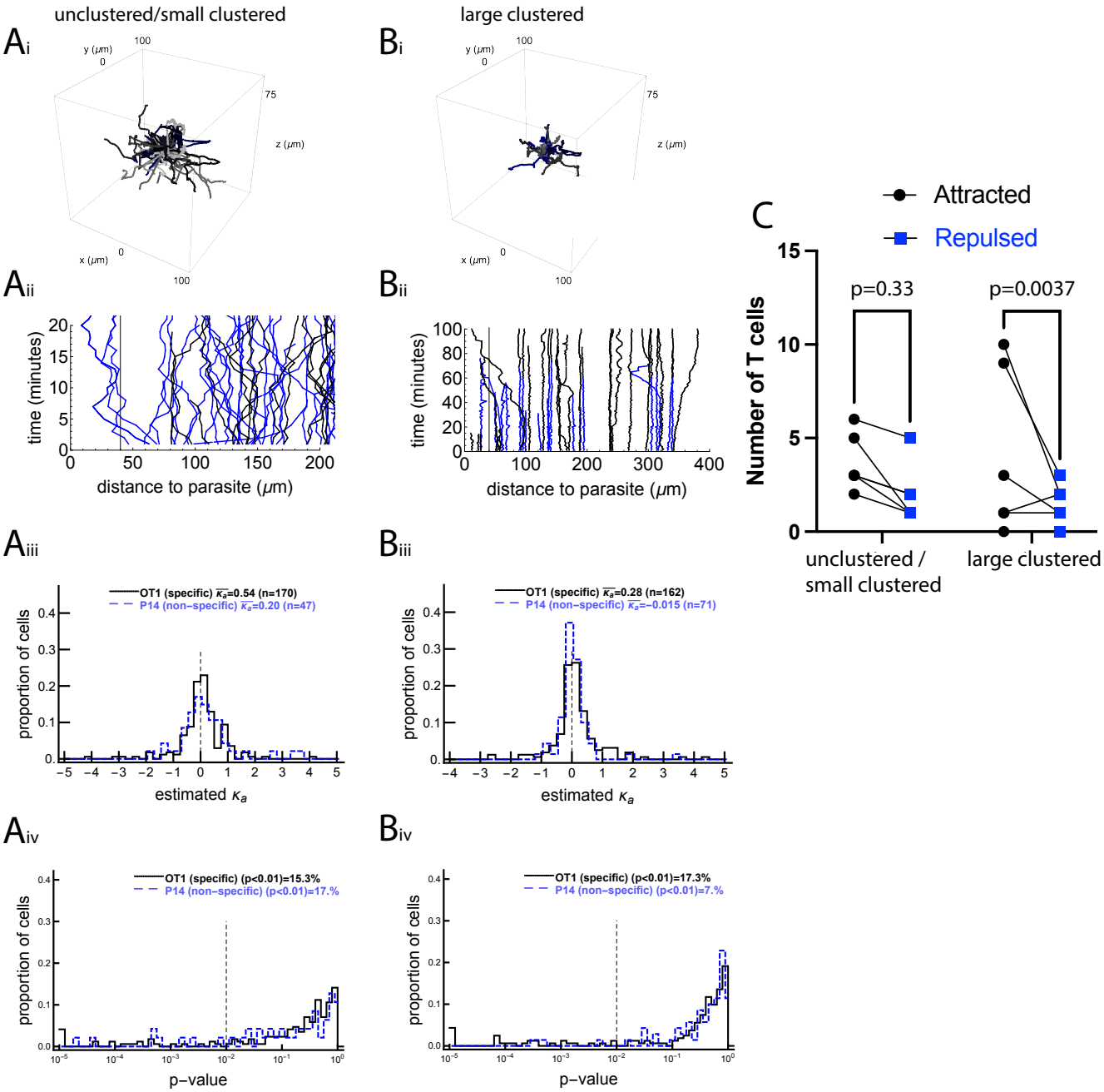

Figure 2: (Caption is on next page)

Figure 2: A minority of activated CD8 T cells display movement bias to the malaria liver stage. We performed four sets of experiments in which the movement of CD8 T cells with respect to the location of the malaria liver stage was recorded using intravital microscopy (see Materials and Methods), and performed tests on the data. In the unclustered/small clustered dataset, the movement of malaria-specific (OT1) CD8 T cells and T cells of irrelevant specificity (P14) was recorded around a single liver stage of Pb when no or few T cells were near the parasite (dataset #1, panels A, 4 movies in total). In the large clustered dataset, the movement of malaria-specific (OT1) CD8 T cells and T cells of irrelevant specificity (P14) was recorded around a single liver stage of Pb when several T cells were already near the parasite (dataset #2, panels B, 3 movies in total). Panels i show the T cell tracks from one of each set of experiments (the second infection site of the unclustered/small clustered data and the first infection site of the large clustered data) and panels ii show the distance from each T cell to the liver stage over time for the same infection sites. For each T cell we calculated the bias of T cell movement towards the parasite by estimating the concentration parameter κ_a from the vMF distribution (see eqn. (1) and Figure S3) using the maximum likelihood approach. Panels iii show the distribution of estimated κ_a and panels iv show p-values from the likelihood ratio test for the VFM distribution with $\kappa_a \to \kappa_a^0$, where κ_a^0 is the threshold value from the null distribution (see text for more detail of how κ_a^0 is calculated). The average concentration for all cells is shown on individual panels in iii, and the percent of T cells with a statistically significant bias to the parasite (with $p \leq 0.01$ from our log likelihood test) is shown in iv. Biased cells include both attracted ($\kappa_a > 0$) and repulsed ($\kappa_a < 0$) cells. In the unclustered/small clustered dataset, we detected 16 OT1 and 5 P14 cells as attracted and 10 OT1 and 3 P14 as repulsed. In the large clustered dataset, we detected 22 OT1 and 2 P14 cells as attracted and 6 OT1 and 3 P14 as repulsed. The number of OT1 cells detected as attracted was found to be significant for the large clustered dataset and nonsignificant for the unclustered/small clustered dataset (panel C). Results of the analysis for two other datasets are given in the Supplement (Figure S8).

was not very different from zero. Notably, only OT1 cells (specific for malaria liver stages) displayed statistically significant attraction towards the infection site for large clustered data ($\bar{\kappa}_a = 0.28$; $p = 0.0065$, signed rank test), while P14 cells in the large clustered data or both OT1 and P14 cells in the unclustered/small clustered data did not display significant attraction ($p > 0.05$, signed rank test). Similarly, only malaria-specific T cells (PyTCR) displayed attraction (or weak attraction) for the two other datasets ($\bar{\kappa}_a = 0.55$; $p = 0.0024$ in the co-clustered dataset and $\bar{\kappa}_a = 0.76$; $p = 0.059$ for the Paris dataset), while non-specific OT1 cells were not attracted to the infection site in the co-clustered dataset ($\bar{\kappa}_a = 0.41$; $p = 0.13$, Figure S8).

We found that about 15% of both malaria-specific and non-specific T cells displayed bias to the parasite movement ($p < 0.01$, Figure 2iv). Interestingly, while the fraction of Plasmodium-specific T cells (OT1) displaying bias to the infection was similar between the unclustered/small clustered and large clustered datasets (15% vs. 17%, respectively), there were more P14 T cells (specific to LCMV) that displayed bias to the parasite in the unclustered/small clustered data than in the large clustered data (17% versus 7%, Figure 2iv), suggesting that some "biased" T cells in the unclustered/small clustered dataset may be an artifact of the statistical analysis. Indeed, we found similar fractions of T cells detected as attracted to or repulsed from the parasite in all cases except for the OT1 cells in the large clustered data (22/28, binomial test $p = 0.004$), suggesting that only during larger clustering is there a strong bias in a minority of malaria-specific CD8 T cells towards the infection site (Figure 2).

3.4 Detecting bias to the infection is correlated with higher cell speed and persistence of movement

Our results suggested that the majority of CD8 T cells searching for malaria liver stages perform such a search without displaying a detectable bias towards the infection site. However, some malaria-specific CD8 T cells do display a strong bias towards the infection site when there are already some T cells near the parasite. We next sought to determine which T cell characteristics may be correlated

with bias towards the infection. To increase the power of the analysis, we pooled the data for malaria-specific CD8 T cells and CD8 T cells of irrelevant specificity. Previously it was suggested that the distance between an infection site and the T cell may determine the strength of attraction[16]. However, we found that the degree of attraction did not correlate with the starting distance between T cells and the infection (Figure 3A and Figure S12A). The distance to the closest T cell or the overall time per track also did not correlate with attraction (Figure 3D&E and Figure S12D&E). However, the average movement per time step, cell velocity, and walk persistence (determined by the concentration parameter κ_t) strongly correlated with the degree of attraction (Figures 3 and S12). These results therefore suggest that rapidly moving T cells are more likely to display bias towards the infection site. Surprisingly, we also found that for the unclustered/small clustered data, velocity and small turning angles (determined by the large concentration parameter κ_t) are correlated with cells displaying bias to the infection site (Figures S13 and S14), suggesting that some T cells may have been detected as biased due to a statistical artifact.

3.5 The amount of attraction does not change immediately upon formation of a cluster

Our results show a greater proportion of attraction toward parasites in data with large CD8 T cell clusters. This potentially suggests that the first cells to reach the parasite may do so randomly, and other cells then begin to show the bias after the environment around the parasite was changed by the first scout cells. Interestingly, in one of our experiments we observed an event (dubbed the "Attenborough experiment") when at the start of imaging no T cells were near the parasite but then one cell reached the parasite (Figures S9C and S15). However, no more cells reached the parasite after the initial cell and we found no evidence that other moving T cells display any additional attraction after the first encounter between the T cell and the parasite (Figure S15). This specific result suggests that if the first T cell that located the parasite changes the environment, such a change takes longer than 30-40 min to be detected by other moving T cells.

3.6 Detecting weak attraction is difficult with current experimental setups

In our analyses so far we found that, with some exceptions, the vast majority of liver-localized CD8 T cells search for the malaria liver stages randomly, with little evidence of bias towards the infection site. Yet, we know that with sufficient numbers of liver-localized CD8 T cells, all liver stages can be eliminated within 48 hours[7,29]. We reasoned that while our liver imaging experiments were sufficiently long ($\sim 1.5 - 2$ hours), it may be possible that they were still too short to detect weak attraction. Therefore we performed several sets of simulations to determine the length of experiments that would be required to detect a given degree of T cell attraction towards the infection site (Figure 4).

We simulated movement of 10^3 cells varying the strength of attraction towards the infection site (κ_a) and the number of cell movements (determining the length of the imaging experiments assuming that every movement was recorded in 2 min timesteps between recorded volumes), assuming that cells start their search at 150 μm from the parasite. For each combination of parameters, we determine the power to detect attraction as the proportion of cells which are detected as attracted based on the angle distribution metric. For our typical experiment with a length of 1.5 hours (about 50

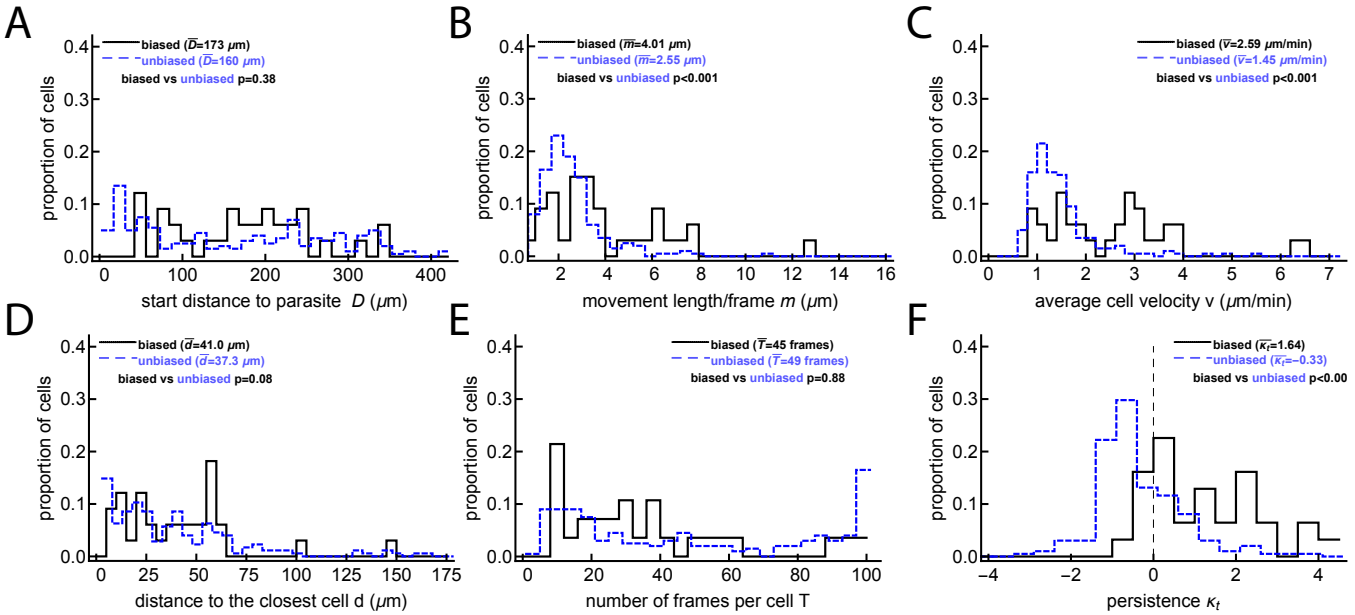

Figure 3: Biased T cells move more rapidly with smaller turning angles. For the large clustered dataset, in which some T cells have already located the parasite (shown previously in Figure 2B), we calculated multiple characteristics for cells that display no bias to the parasite's location ("unbiased" cells) and cells that do display bias to the parasite's position ("biased" cells). These characteristics include: starting distances D (panel A), movement lengths per frame m (panel B), average speed per T cell v (panel C), distance to the closest T cell d (panel D), the number of recorded positions T (panel E), and the estimated concentration of the vMF for the turning angles κ_t (panel F). Comparisons were done using the Mann-Whitney test and the p-values for the comparisons are shown on individual panels. The unbiased results are offset slightly to not directly overlap with the biased results for ease of viewing. Analyses when the data were divided into attracted, repulsed, or unbiased T cells are shown in Figure S12, and analysis of the data with no/small clusters (dataset #3) is shown in Figure S14. The majority of biased T cells in this dataset display attraction towards the infection (24/33, binomial test $p = 0.014$).

movements of T cells), the attraction strength κ_a must be more than 1 for the bias of T cell movement towards infection to be consistently detected (Figure 4A). Furthermore, to detect weaker attraction ($\kappa_a = 0.2$), 30 hours of imaging experiments (or 1000 movements per cell) would be needed, which is not currently possible. The latter result strongly suggests that our failure to detect weak attraction of T cells towards the infection site may be due to limited (but the best currently possible) data.

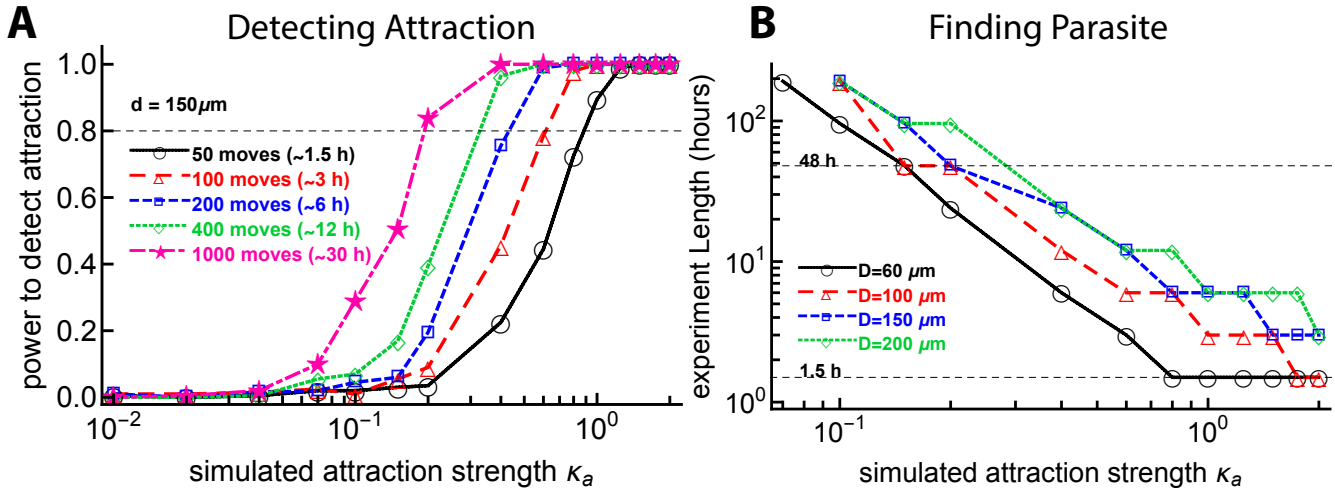

Figure 4: To detect weak T cell attraction to the parasite or to observe a T cell finding the parasite requires prohibitively long-term imaging experiments. In panel A we simulated T cell movement with some degree of attraction to the parasite defined by the concentration κ_a starting at a distance $D = 150$ μm. For 1000 simulated T cells, we varied the number of T cell movements (which changes the duration of the imaging movie). For every track we then determined the probability that the movement track would exhibit attraction (by estimating κ_a from the simulated tracks using maximum likelihood and using the likelihood ratio test to test if the estimated value is different from zero, i.e. from random). In panel B we performed simulations with 1000 cells for values of κ_a between 0 and 2, with starting distances of 60, 100, 150, and 200 μm from the parasite and T cell movement lengths chosen from a generalized Pareto distribution (eqn. (S7)) with randomly chosen turning angles. We calculated the time it took for the T cell to actually reach the parasite (i.e., reach the distance of 40 μm from the parasite. A typical time of imaging a movie (~ 1.5 h) and the time it takes for liver stages to develop (48 h) are indicated by dashed horizontal lines in panel B.

We wondered how the time to find the parasite may depend on the starting distance from the T cell to the liver stage and the degree of the T cell attraction towards the infection site. We performed another set of simulations (with 10^3 cells) by varying the starting distance between the T cells and the parasite and the degree of T cell attraction to the parasite (κ_a) and estimated the time when a cell found the parasite (reaches within 40 μm from the parasite) with 80% probability. A high degree of attraction ($\kappa_a \sim 1$) and short distance is required for T cells to find the parasite within 1.5 hours, and to reach the parasite in 48 hours still requires moderate attraction ($\kappa_a = 0.3$, Figure 4B). In addition, for a T cell with $\kappa_a = 1.0$ and a starting distance of 150 μm, the experiment must last for 6 hours for the T cell to reach the parasite; currently we are not able to perform experiments with live mice for this length of time. These results suggest that there may be something special about parasites that had T cell clusters formed very early after the infection, for example, such parasites may be entering the liver where there are already some T cells nearby, perhaps indicating that Plasmodium infection of the liver and T cell localization in the liver do not occur randomly. Furthermore, analyses also suggest that some level of weak attraction is needed to explain T cells finding the parasite in 48 hours after infection.

3.7 Challenging the conclusion that only a small fraction of T cells display bias towards the infection site

In our clustered data we found that only $\sim$ 20% of T cells display statistically significant bias in movement towards the infection site (Figure 2iv). An alternative interpretation of this result is that perhaps all T cells in these experiments exhibit small bias towards the infection site but due to random sampling only a small fraction of the cells display significant bias. To test this explanation we performed additional simulations. Specifically, we simulated movement of 500 cells by varying the strength of attraction towards the infection site (κ_a) for all cells and the number of cell movements (assuming that every movement was recorded with 2 min timesteps between recorded volumes), assuming that cells start their search at 150 μm from the parasite. For each combination of parameters, we fit the vMF distribution to each of 500 trajectories and, thus, generate a distribution of concentrations κ_a. One interesting result from these simulations is that as the bias towards infection site set in simulations increases, the mean of the estimated concentration parameters increases as well leading to a large difference between experimentally observed and simulated distributions (Figure S7A-C). The simulated distributions, however, typically had smaller tails than the experimental distribution implying less variation in the estimated concentrations κ_a (Figure S7A-B). This result does suggest that our data are not consistent with the idea that all cells exhibit some level of bias towards the infection site. Furthermore, power analysis suggests that to see $\sim$ 20% of cells exhibit statistically significant bias towards infection due to sampling, all cells must have a relatively substantial attraction ($\kappa_a \approx 0.5$, Figure S7D) which would result in high average attraction which is not observed in actual data. Thus our results are more consistent with the explanation that only a small fraction of T cells exhibit bias towards the parasite when there is already a cluster of T cells.

3.8 Attraction to random infection sites

In our analysis, we focused on estimating if T cells are attracted to the site of infection, and our tests did not involve a control because it is possible that both malaria-specific T cells and T cells with irrelevant specificity may be equally attracted to the infection site. However, our vMF distribution-based metric could be too sensitive, detecting attraction when it does not exist. Indeed, we showed that for T cells undergoing a correlated random walk, some cells may be detected as attracted to the infection in the absence of actual attraction (Figure S6). We reasoned that if T cells are truly attracted to the infection site, there must be something special about that specific location, and fewer cells (if any) should display attraction to areas distant from the actual parasite's position.

For this analysis we chose "fake parasite" positions equally spaced in the $500 \times 500 \times 50\mu$m imaging box (in total 216 positions) and determined the number of T cells that are attracted to or repulsed from each fake location (using our vMF distribution-based metric). For the control and unclustered/small clustered datasets, the number of cells detected as attracted to or repulsed from any position was similar, suggesting there was nothing special about the real parasite position in the unclustered/small clustered dataset (Figure S17i-ii). This result supports that in the absence of large cluster formation, T cells were searching for the parasite randomly (or with a weak bias that was not detectable). In contrast, in the large clustered dataset, more cells are detected as attracted and fewer as repulsed when the distance between the real parasite and "fake" parasite is short, suggesting that there is some T cell attraction to the real parasite position when there are other T cells already present near the parasite (Figure S17iii).

3.9 The number of liver-resident T cells required for sterilizing protection from malaria is reduced greatly if there is biased T cell movement towards the infection site

If CD8 T cells search for the malaria liver stages nearly randomly, we wondered how the presence of weak attraction (that is not detectable in current experiments) could change the ability of T cells to find the parasite (i.e., reach within 40 μm of the parasite). We therefore performed 1000 simulations with varying T cell attraction to the parasite (the concentration κ_a) and the starting distance from the cell to the parasite and counted the cells which reach the parasite (within 40 μm) within a defined time period. Our results suggest that a small change in attraction strength κ_a can dramatically increase the chance of T cells finding the parasite, especially when the initial distance between the T cell and parasite is large (Figure 5A-C). Importantly, a relatively weak bias towards the infection ($\kappa_a \approx 0.2 - 0.3$) is sufficient for T cells to locate the parasite within 48 hours after infection, even for a relatively large initial distance between the T cell and the parasite (Figure 5C).

In our experiments, the initial distances between T cells and the parasite varied dramatically; the average was around 150-200 μm (e.g., Figure 3). Given a mouse liver size of about 1 mL volume[30], we calculated that if there are 10^5 randomly distributed T cells in the liver (a number found experimentally to provide sterilizing protection against malaria in mice[31,32]), the shortest distance between a randomly positioned parasite and the nearest T cell is about 100 μm at 99% probability (Figure 5D). However, without biased movement towards the parasite, CD8 T cells located at 100 μm from the parasite are unlikely to locate the infection even within 48 hours (Figure 5B), suggesting that some weak attraction may be guiding T cell search for the infection.

Our simulations so far focused on scenarios where a single T cell reaches the parasite. However, when multiple T cells search for the parasite, the probability of at least one T cell finding the infection may be higher. To investigate this more rigorously, we performed 10^3 simulations to allow several T cells to search for the parasite by varying the strength of T cell attraction to the parasite (concentration κ_a,) the number of searching T cells, and the length of search, and fixing the starting distance of T cells from the parasite to 150 μm. For each combination of parameters, we determine the proportion of the 1000 simulations for which at least one T cell reaches the parasite (within 40 μm). Results suggest that for essentially all parasites to be found in 48 hours, we need around 10 T cells per parasite and a weak attraction defined by the concentration $\kappa_a = 0.2$ (Figure S16). These results suggest that when multiple T cells search for an infection, weak attraction greatly improves the chances of T cells finding the parasite.

4 Discussion

It has been well established that activated and memory CD8 T cells are capable of providing sterilizing protection against infection with Plasmodium sporozoites in mice[7,9,29]. However, rules by which T cells are able to locate and eliminate all Plasmodium-infected hepatocytes remain unclear. By generating unique data on movement of liver-localized activated CD8 T cells within few hours after infection with Plasmodium sporozoites and by rigorous analysis of these data using a newly developed metric based on the Von Mises-Fisher distribution, we found that most CD8 T cells search for the infection site randomly. Using stochastic simulations, we showed that randomly searching T cells have a high failure rate of finding rare parasites. A small bias towards the infection dramatically

improves the chance of T cells finding the infection, and thus reduces the number of liver-localized T cells needed for sterilizing protection. Power analyses showed, however, that current experimental set-ups allowing for intravital imaging of livers of live mice for only a few hours do not generate sufficient data to detect attraction to the infection site by individual T cells.

For all analyzed datasets, we consistently found a small fraction (about 15-20%) of T cells displaying strong movement bias to the infection site; such biased T cells included both T cells attracted to the parasite and T cells repulsed from the parasite. In the case where there were several cells near the parasite (large clustered data), the number of T cells attracted to the infection site was significantly higher than that of repulsed cells. The distance between the T cell and the infection did not correlate with the strength of attraction; however, the speed at which T cells moved was a strong predictor of attraction, that is, more rapidly moving T cells displayed stronger bias. This observation highlights the need for a better understanding of the mechanisms that regulate movement speed at which T cells survey peripheral tissues. In our case, however, we found that rapidly moving T cells having a higher bias towards the infection site may be an artifact of a correlated/persistent random walk.

The reasons why some T cells (e.g., most OT1 cells in the large clustered data) may display attraction towards the infection site while others (e.g., P14 cells in the large clustered data) search randomly are unclear. It is generally believed that movement patterns of T cells in tissues are regulated by chemokines, i.e. T cells with appropriate chemokine receptors follow chemokine gradients in the environment[33]. However, it is not yet understood whether T cells indeed follow chemokine gradients (which may be shallow in many circumstances) or chemokines simply regulate the cells' velocities[16,33,34]. For example, CXCL21 chemokine and CCR7 receptors on B cells regulate the localization of B cells near B-T zones of lymph nodes[34]. However, while the CXCR3 receptor is important for CD8 T cell movement in the skin or brain[16,35], if T cells in these tissues follow the chemokines' gradients has not been established. We recently found that the LFA-1 receptor on T cells is critical for T cell motility in the liver; however, even LFA-1-deficient T cells are capable of locating Plasmodium liver stages with somewhat less success[23]. Whether receptors regulate T cell attraction to the Plasmodium liver stages remains to be determined. It is possible, however, that such attraction is achieved by a combination of receptors, and the lack of any single receptor does not dramatically impact the ability of T cells to locate the infection. It was recently found that the migration of *Mycobacterium tuberculosis*-specific CD4 T cells from the lung vasculature into parenchyma is dependent on many receptors, and lacking a single receptor has a moderate to minimal impact on the rate at which T cells migrate from the blood to the lung[36].

The presented results have several limitations. Even though our intravital imaging experiments lasted for 2-3 hours, the amount of data collected was not sufficient to detect the weak attraction that T cells may exhibit when searching for the parasite. Experiments with a longer duration may be ethically difficult and long-term imaging of surgically exposed livers may also general artifacts due to local tissue damage. The imaging frequency in our data was relatively low (1.5-2min/frame) to allow for longer movies with lower laser exposure to the liver. While more frequent measurements could increase the number of data points, long movies may increase chances of tissue damage, require time-consuming processing steps, and may result in a bias of imaged cells remaining in the imaging volume for a long time. Therefore we aimed for a balance between the length of imaging and costs associated with longer movies. Due to these limitations of intravital imaging, we could only record cell positions in a limited 3D volume of approximate $500 \times 500 \times 50$ μm centered near a parasite. If a cell exits that volume we no longer record its positions, so there may be bias toward recording more movements from T cells that act attracted to the parasite than cells that are "repulsed", since repulsed cells are more likely to leave the imaging area. We performed simulations that partially

allow us to correct for such bias in our vMF distribution-based metric.

While it is not often mentioned, intravital imaging may induce local damage due to exposing tissues to lasers, especially in longer experiments. It is often difficult to evaluate the degree of liver damage any specific experiment involves. In our experiments we have observed that in Plasmodium sporozoite-infected livers, CD8 T cells move relatively slowly, with speeds of 1.5-2.5 μm/min. This was significantly slower than the 3.5 μm/min estimated in one control movie without an infection (control dataset), and lower than the average T cell speeds estimated in previous studies, which are in the range of 5-10 μm/min[23,37].

For most of our analyses, we pooled together the movements of different cells found in different mice to increase the analyses' power. In general we found similar results (although at times statistically non-significant) as when we analyzed the individual datasets separately. The Von Mises-Fisher distribution, while a straightforward representation of bias in 3D movement, may not perfectly approximate the distribution of angles towards the infection, although it is close (Figure S18). We found that in cases when vMF distribution does not fit the angle to infection data well, a mixture of vMF distributions can match the data nearly perfectly (Figure S18A-C). Such an approach also allows us to estimate the fraction of angles near 90^0 that deviate from the single vMF distribution; what such angles represent biologically remains to be determined. This pattern may arise if cells attempt to approach a particular location when they are unable to do so directly due to the absence of a direct path to the location, in which case the cells may thus meander.

All our simulations were concerned with the process of finding the parasite; however, for protection, T cells must also kill the parasite. The time it takes for CD8 T cells to kill Plasmodium liver stages has not been quantified, and therefore our estimates of the search time should be considered minimum estimates of the time it takes for T cells to find and kill the parasite.

One major limitation is that we ignored the fact that liver-localized CD8 T cells move in liver sinusoids[23,38] and are constrained in their movement by the structure of the sinusoids. The chance of randomly finding an infection may be influenced by the structure of the sinusoids[38], and it remains to be determined if a biased search is more successful in the liver than in open (3D) space. An argument could be made that our simulations ignoring liver sinusoids makes the simulations not useful for approximating cell movement; however, as the parameters are chosen to replicate real movement, especially in simulations with persistent random walks, we think our simulations, even though performed in open space, are a reasonable representations of cell movement in the liver. We plan to investigate the sinusoidal structure the cells are constrained to in greater detail in a future work.

One of the purposes of this work was to test predictions of our recently published density-dependent recruitment model of CD8 T cell cluster formation around Plasmodium liver stages, which states that the formation of clusters is driven by a positive feedback loop where larger clusters recruit more T cells[10,11]. While the data are supportive of a random search of T cells for the parasite, it remains unclear why some parasites already have T cells nearby at the start of imaging while other parasites have either no T cells nearby or there is no accumulation of T cells near other parasites in $\sim 1-2$ hours of imaging. Interestingly, even in situations in which the first T cell discovers the parasite, we found no evidence that other T cells in surrounding areas become more attracted to the parasite. The data thus do not yet allow us to discriminate between density-dependent recruitment and variability in parasite attractiveness models[11]. Interestingly, we found that a minority of Plasmodium-specific CD8 T cells (OT1) display attraction towards already formed T cell cluster, while cells of irrelevant specificity (P14) were not attracted. This could be due to lower numbers

of P14 cells generally retained in the liver; these cells tend to undergo less robust expansion and differentiation in *in vitro* culture (results not shown).

There are several ways the work outlined in this paper can be extended. We did not investigate if any specific chemokine receptors (e.g., CXCR3, CCR5, or CX3CXR1) impact the success rate of the T cell search for the infection site. This is a current focus of our research. Including the structure of liver sinusoids in simulations of T cell search for the infection would likely provide more realistic estimates of the time it takes for T cells to find the parasite. To measure T cell attraction to the infection in this work, we used the vector from the T cell to the parasite as the optimal route to reach the infection site. However, due to physical constraints from the liver sinusoids, the optimal route to the infection may not be the direct vector. By imaging the parasite, T cells, and liver sinusoids, it may be possible to quantify an optimal route, traveling through the sinusoidal structure, and to determine if the cells move with attraction along the structure..

Future work may need to investigate how quickly T cells form clusters around the parasites and if the degree of attraction we estimate in this work is sufficient to explain the formation of relatively large clusters (e.g., 5-10 CD8 T cells) near individual parasites[10,11]. Similarly to how we used statistics to predict distribution of angles to the parasite using vMF distribution, we also attempted to find an analytical description of a distribution of changes of distances to the parasite with T cell movement. However, that distribution is significantly more complicated, involving higher forms of mathematics, and we abandoned it for this paper. Nevertheless, we may investigate a distance distribution in future work. Our analysis suggests a provocative idea that the speed at which T cells search for an infection may be correlated with the ability of T cells to locate the infection site. However, there is a clear trade-off between the speed at which T cells may be moving in tissues and the ability of such T cells to process environmental signals to detect the infection - too rapid cells may miss many of the infected cells in their haste[39]. Whether there is an optimal movement pattern of T cells to locate all infections in a small enough time to cure infection remains to be determined[40]. Taken together, despite potential limitations, we provide novel data from innovative imaging experiments and rigorous mathematical analyses of how CD8 T cells search for the infection in complex tissue such as the liver.

Data sources

T cell movement/track data generated in this paper (unclustered/small clustered and large clustered datasets) are provided as a csv file (`https://github.com/viktorZenkov/measuringAttraction/blob/master/Data/AllPositionData.xlsx`.), some published previously (Paris and co-clustered datasets).

Code sources

Analyses have been primarily performed in Mathematica (ver 12) and codes used to generate most of the figures in the paper are provided on GitHub: `https://github.com/viktorZenkov/measuringAttraction/tree/master/Measuring`.

Ethics statement

All animal procedures were approved by the Animal Experimentation Ethics Committee of the Australian National University (Protocol numbers: A2016/17; 2019/36). All research involving animals was conducted in accordance with the National Health and Medical Research Council's Australian Code for the Care and Use of Animals for Scientific Purposes and the Australian Capital Territory Animal Welfare Act 1992.

Author contributions

VSZ, JOC, IAC, and VVG developed ideas for the study. JOC performed intravital imaging experiments and segmentation of the imaging data with Imaris (under supervision of IAC), with assistance from HM on supplemental newer data. VSZ performed analyses of the data and simulations of cell movements (under supervision of VVG). VSZ wrote the first draft of the paper, with modifications later by VVG and VSZ with some contributions from JOC. VSZ is the primary author of the paper.

Acknowledgments

We would like to thank Harshana Rajakaruna for discussion of various aspects of this work, and all members of GanusovLab for feedback on earlier versions of the paper. This work was supported by NIH grant (R01 GM118553) to VVG.

References

1. `https://www.who.int/malaria/publications/world-malaria-report-2018/report/en/`.
2. Aleshnick, M., Ganusov, V. V., Nasir, G., Yenokyan, G. & Sinnis, P. Experimental determination of the force of malaria infection reveals a non-linear relationship to mosquito sporozoite loads. **16,** e1008181. ISSN: 1553-7374. aheadofprint.
3. Murphy, J. R., Baqar, S., Davis, J. R., Herrington, D. A. & Clyde, D. F. Evidence for a 6.5-day minimum exoerythrocytic cycle for Plasmodi um falciparum in humans and confirmation that immunization with a synthetic peptide representative of a region of the circumsporozoite protein retards infection. *J Clin Microbiol* **27,** 1434–1437 (1989).
4. Hermsen, C. C. *et al.* Detection of Plasmodium falciparum malaria parasites in vivo by real-time quantitative PCR. *Mol Biochem Parasitol* **118,** 247–251 (2001).
5. Sturm, A. *et al.* Manipulation of host hepatocytes by the malaria parasite for delivery into liver sinusoids. *Science* **313,** 1287–1290 (2006).
6. Miller, L. H., Ackerman, H. C., Su, X.-z. & Wellems, T. E. Malaria biology and disease pathogenesis: insights for new treatmen ts. *Nat Med* **19,** 156–167 (2013).
7. Schmidt, N. W. *et al.* Memory CD8 T cell responses exceeding a large but definable threshold provide long-term immunity to malaria. *Proc Natl Acad Sci U S A* **105,** 14017–22 (2008).

8. Schmidt, N. W., Butler, N. S., Badovinac, V. P. & Harty, J. T. Extreme CD8 T cell requirements for anti-malarial liver-stage immunity following immunization with radiation attenuated sporozoites. *PLoS Pathog* **6,** e1000998 (2010).

9. Schmidt, N. W. & Harty, J. T. Cutting edge: attrition of Plasmodium-specific memory CD8 T cells results in decreased protection that is rescued by booster immunization. *J Immunol* **186,** 3836–3840 (2011).

10. Cockburn, I. A. *et al.* In vivo imaging of CD8+ T cell-mediated elimination of malaria liver stages. eng. *Proc Natl Acad Sci U S A* **110,** 9090–9095. `http://dx.doi.org/10.1073/pnas.1303858110` (May 2013).

11. Kelemen, R. K., Rajakaruna, H., Cockburn, I. A. & Ganusov, V. V. Clustering of activated CD8 T cells around malaria-infected hepatocytes is rapid and is driven by antigen-specific cells. *Frontiers in immunology* **10,** 2153. ISSN: 1664-3224 (2019).

12. Kelemen, R. K. *et al.* Classification of T Cell Movement Tracks Allows for Prediction of Cell Function. *International Journal of Computational Biology and Drug Design* (2014).

13. Krummel, M. F., Bartumeus, F. & Gérard, A. T cell migration, search strategies and mechanisms. *Nature Reviews Immunology* **16,** 193–201. ISSN: 1474-1733 (2016).

14. Okada, T. *et al.* Antigen-Engaged B Cells Undergo Chemotaxis toward the T Zone and Form Motile Conjugates with Helper T Cells (B Cell-T Cell Interaction Dynamics). eng. *PLoS Biology* **3,** e150. ISSN: 1544-9173 (2005).

15. Beltman, J. B., Allen, C. D. C., Cyster, J. G. & de Boer, R. J. B cells within germinal centers migrate preferentially from dark to light zone. eng. *Proceedings of the National Academy of Sciences of the United States of America* **108,** 8755–8760. ISSN: 0027-8424 (2011).

16. Ariotti, S. *et al.* Subtle CXCR3-Dependent Chemotaxis of CTLs within Infected Tissue Allows Efficient Target Localization. eng. *Journal of immunology (Baltimore, Md. : 1950)* **195.** ISSN: 1550-6606 (Dec. 2015).

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

18. Turchin, P., Odendaal, F. J. & Rausher, M. D. Quantifying Insect Movement in the Field. *Environ Entomol* **20,** 955–963. ISSN: 1938-2936 (1991).

19. Turchin, P. Quantitative Analysis of Movement: Measuring and Modeling Population Redistribution in Animals and Plants. 381. ISSN: 0033-5770 (1998).

20. Duchesne, T., Fortin, D. & Rivest, L.-P. Equivalence between Step Selection Functions and Biased Correlated Random Walks for Statistical Inference on Animal Movement. *PloS one* **10,** e0122947. ISSN: 1932-6203 (4 2015). epublish.

21. Wang, L. T. *et al.* A Potent Anti-Malarial Human Monoclonal Antibody Targets Circumsporozoite Protein Minor Repeats and Neutralizes Sporozoites in the Liver. *Immunity* **53,** 733–744.e8. ISSN: 1097-4180 (4 Oct. 2020). ppublish.

22. Cockburn, I. A., Tse, S.-W. & Zavala, F. CD8+ T cells eliminate liver-stage Plasmodium berghei parasites without detectable bystander effect. eng. *Infect Immun* **82,** 1460–1464. `http://dx.doi.org/10.1128/IAI.01500-13` (Apr. 2014).

23. McNamara, H. A. *et al.* Up-regulation of LFA-1 allows liver-resident memory T cells to patrol and remain in the hepatic sinusoids. *Science Immunology,* 1–10 (2 2017).

24. Hong, Y. On computing the distribution function for the Poisson binomial distribution. eng. *Computational Statistics and Data Analysis* **59,** 41–51. ISSN: 0167-9473 (2013).

25. Mardia, K. V. *Directional statistics* eng. Chichester ;

26. Fisher, N. I. *Statistical analysis of spherical data* eng. Cambridge [Cambridgeshire] ;

27. Castellino, F. *et al.* Chemokines enhance immunity by guiding naive CD8+ T cells to sites of CD4+ T cell–dendritic cell interaction. *Nature* **440,** 890–895. ISSN: 0028-0836 (2006).

28. Jakob, W. *Numerically stable sampling of the von Mises Fisher distribution on S2 (and other tricks)* `http://www.mitsuba-renderer.org/~wenzel/files/vmf.pdf`.

29. Fernandez-Ruiz, D. *et al.* Liver-Resident Memory CD8(+) T Cells Form a Front-Line Defense against Malaria Liver-Stage Infection. *Immunity* **45,** 889–902. ISSN: 1097-4180 (4 Oct. 2016).

30. Marino, D. J. Age-specific absolute and relative organ weight distributions for B6C3F1 mice. *Journal of toxicology and environmental health. Part A* **75,** 76–99. ISSN: 1528-7394 (2 2012). ppublish.

31. Olsen, T. M., Stone, B. C., Chuenchob, V. & Murphy, S. C. Prime-and-Trap Malaria Vaccination To Generate Protective CD8+ Liver-Resident Memory T Cells. *J Immunol* **201,** 1984–1993. ISSN: 1550-6606 (7 Oct. 2018).

32. Gola, A. *et al.* Prime and target immunization protects against liver-stage malaria in mice. *Science translational medicine* **10.** ISSN: 1946-6242 (460 Sept. 2018).

33. Krummel, M. F., Bartumeus, F. & Gérard, A. r. T cell migration, search strategies and mechanisms. *Nature reviews. Immunology* **16,** 193–201. ISSN: 1474-1741 (Mar. 2016).

34. Okada, T. *et al.* Antigen-engaged B cells undergo chemotaxis toward the T zone and form motile conjugates with helper T cells. *PLoS biology* **3,** e150. ISSN: 1545-7885 (June 2005).

35. Harris, T. H. *et al.* Generalized Lévy walks and the role of chemokines in migratio n of effector CD8+ T cells. *Nature* **486,** 545–548 (June 2012).

36. Hoft, S. G. *et al.* The Rate of CD4 T Cell Entry into the Lungs during Mycobacterium tuberculosis Infection Is Determined by Partial and Opposing Effects of Multiple Chemokine Receptors. *Infection and immunity* **87.** ISSN: 1098-5522 (6 June 2019).

37. Guidotti, L. G. *et al.* Immunosurveillance of the liver by intravascular effector CD8(+) T cells. *Cell* **161,** 486–500. ISSN: 1097-4172 (3 Apr. 2015).

38. Rajakaruna, H., O'Connor, J., Cockburn, I. A. & Ganusov, V. V. *Environment-imposed constraints make Brownian walkers efficient searchers* BioRxiv. 2020. `https://doi.org/10.1101/2020.11.06.371690`.

39. Mrass, P. *et al.* ROCK regulates the intermittent mode of interstitial T cell migration in inflamed lungs. eng. *Nat Commun* **8,** 1010–1010. ISSN: 2041-1723 (Dec. 2017).

40. Moses, M. E., Cannon, J. L., Gordon, D. M. & Forrest, S. Distributed Adaptive Search in T Cells: Lessons From Ants. **10,** 1357. ISSN: 1664-3224. epublish.

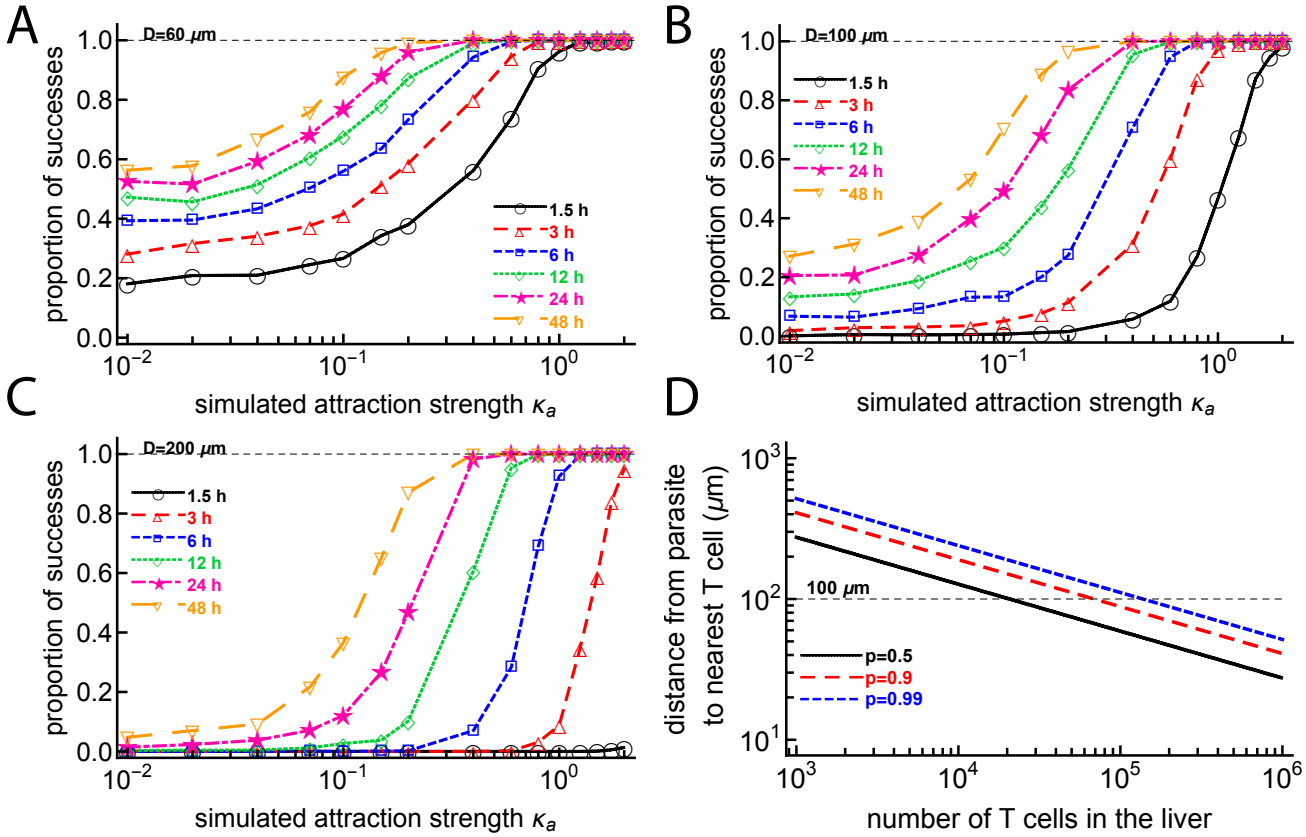

Figure 5: Small attraction dramatically increases the chances for a T cell to find the parasite within 48 hours. We simulated CD8 T cells searching for a malaria liver stage assuming variable levels of T cell attraction to the parasite, defined by the concentration parameter κ_a of the vMF distribution. Cells started their search for the infection 60 (A), 100 (B), or 200 (C) μm from the parasite, and the probability that the cells reach the parasite within 40μm distance at various times after infection was calculated. To simulate a T cell walk we assumed that T cell movement lengths follow a generalized Pareto distribution (eqn. (S7)) with randomly chosen turning angles. Simulations were done for 1000 cells for each value of κ_a. In panel D we calculated the distance between a parasite and nearest CD8 T cells if T cells are randomly distributed in the 1 mL = 1 cm^3 volume of the liver for different levels of certainty (defined by the probability p). The distance from the parasite to the nearest T cell is calculated as $d = \frac{r}{2}\sqrt[3]{\frac{\ln(1/(1-p))}{n}}$, where $r = \sqrt[3]{3/(4\pi)}$ cm is the radius of a sphere with volume 1 cm^3 approximating a liver and n is the number of T cells in the liver. For $p = 0.99$ when there are $n = 10^5$ T cells in the liver, the chance that one T cell is at the distance of $d = 100$ μm from the parasite is 99%.

A new method shows that a minority of liver-localized CD8 T cells display hard-to-detect attraction to Plasmodium-infected hepatocytes

Viktor Zenkov, James O'Connor, Ian A. Cockburn, and Vitaly V. Ganusov

Supplemental Information

Additional experimental details

Mice were prepared for two-photon microscopy as previously described[23]. Mice were anesthetized with a mix of Ketamine (100 mg/kg) and Xylazine (10 mg/kg). Throughout the surgery and imaging procedure the mouse temperature was maintained at 37°C using a heating mat attached to a feedback probe inserted in the mouse rectum. A lateral incision was made over the left lobe of the liver and any vessels cauterized by applying light pressure to the vessel until clotting occurred naturally. The mouse was then placed in a custom made holder. The liver was then exposed and directly adhered to a cover slip that was secured in the holder. Once stable the preparation was transferred to a Fluoview FVMPE-RS multiphoton microscope system (Olympus) equipped with a XLPLN25XWMP2 objective (25x; NA1.05; water immersion; 2mm working distance). For analysis of the motility of the sporozoites and liver-localized CD8 T cells a 50 μm z-stack (2 μm/slice) was typically acquired using a standard galvano-scanner at a frame rate between 0.5 to 1 frames per minute depending on the movie.

The selection of parasites was contingent upon clear, healthy sections of liver tissue, and the proximity of lymphocytes at the time of imaging. Each parasite chosen for imaging required unimpeded blood flow of surrounding sinusoids, bright autofluorescence and clear surrounding structures. In addition, the parasites chosen for imaging were also dependent upon what information we wanted. For the 0 hour timepoint, parasites were selected if they had healthy tissue and had no lymphocytes within a 40 μm diameter of the parasite. This allowed for analysis of attraction of lymphocytes with no primary lymphocyte signaling from an infected cell. At 2 hours post infection, parasites were selected if they had healthy tissue and had ≥ 1 lymphocytes within a 40 μm diameter of the parasite. This allowed for analysis of attraction of lymphocytes with at least 1 primary lymphocyte signaling from an infected hepatocyte. We used the 40 μm diameter because it represents the average width of a standard murine hepatocyte[11].

Imaging data processing details

Raw imaging data was analyzed using the Imaris x64 software (Bitplane) v9.2. Tracking of individual cells in a z-stack was performed using the "Spots" function in surpass mode, and in a single slice by the "Surfaces" function in surpass mode. Extensive manual adjustments were subsequently done to ensure the accuracy of the tracks. The detection of individual cells relied upon their relative fluorescence intensity and size (diameter ≥ 9 μm), and used the remaining default function settings for the calculation of the motion tracks using an autoregressive motion algorithm. Detection of fluorescence for each spot was manually adjusted to reduce background detection at each time point to ensure clear distinction of cell vs autofluorescence by the algorithm.

Datasets

1. Dataset #1: unclustered/small clustered data (OT1 T cells in Pb-CS^{5M}-infected B6 mice). No or few T cells were near the parasite.
2. Dataset #2: large clustered data (OT1 T cells in Pb-CS^{5M}-infected B6 mice). Several T cells were found near the parasite (i.e., with T cell cluster).
3. Dataset #3: control data (OT1 T cells in naive/uninfected B6 mice).
4. Dataset #4: Paris data (PyTCR cells in Py-infected Balb/c mice).
5. Dataset #5: co-clustered ("Kelemen") data (PyTCR and OT1 cells in Py-infected CB6 mice).

Metrics

We define three alternative metrics to measure T cell attraction towards the parasite (Figure S3).

1. Metric 1: the angle metric.
   (a) For a movement of a cell between timepoints, the angle of movement is defined as the angle between the vector from the cell's position before it moves to the parasite and the vector from the cell's position before it moves to the cell's position after it moves.
   (b) An acute angle corresponds to "getting closer" and obtuse corresponds to "getting farther". The probability of randomly getting closer is 0.5.
   (c) To test, we associate a Bernoulli distribution with $p = 0.5$ with each movement, then sum up the distributions for all the movements to get a binomial distribution. Our null distribution is then a binomial distribution with $n =$ the number of movements and 0.5.
2. Metric 2: the improved change of distance metric.
   (a) For a movement of a cell between timepoints, the change of distance is defined as the distance from the cell to the parasite after the cell moves minus the distance from the cell to the parasite before the cell moves.
   (b) A negative change of distance corresponds to "getting closer" and positive corresponds to "getting farther. The probability of randomly getting closer is $1/2 - r/4x$ (Figure 1).
   (c) To test, we associate a Bernoulli distribution with $p = 1/2 - r/4x$ with each movement, then sum up the distributions for all the movements to get a Poisson binomial distribution, and this is our null distribution.
3. Metric 3: the angle distribution metric.
   (a) The angle distribution metric uses the angle of movement from Metric 1 and the angle version of the Von Mises-Fisher distribution.
   (b) Taking a dataset of angles, we use a log likelihood test to determine the concentration parameter for the angle version of the Von Mises-Fisher distribution most likely to generate the directions corresponding to the angles.

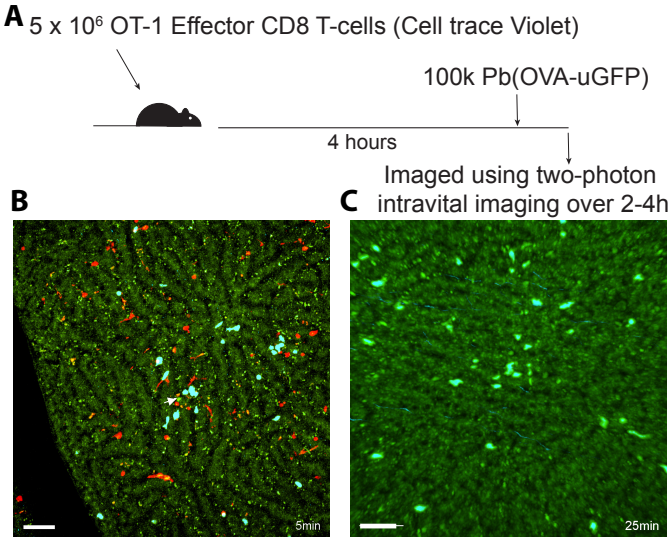

Figure S1: Investigating T cell dynamics in the liver following infection with Plasmodium sporozoites. Four hours after i.v. cell transfer of 5×10^6 Cell trace Violet labelled OT1 CD8 T-lymphocytes (blue), mice were injected with 1×10^5 *Plasmodium berghei* Pb(OVA-uGFP) and immediately prepared for intravital imaging. The mice were imaged using a two-photon standard galvanometer scanner in order to acquire a 50 μm deep Z-stack. Representative images of small clustering (panel B) and large clustering (panel C) image acquisitions are shown. Representative videos of small clustering (movieS1) and large clustering (movieS2) are in the supplemental materials. The green sporozoite is highlighted with a white arrow in both images. The scale bar is 20 μm.

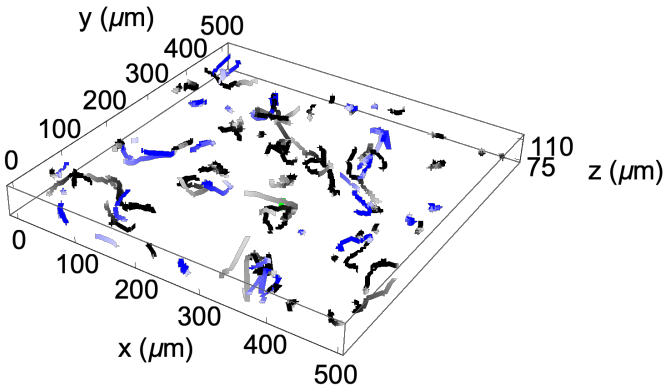

Figure S2: Example of CD8 T cell movements in the liver over time for the second parasite of the unclustered/small clustered dataset. Blue tracks are Plasmodium-specific T cells (OT1) and black are nonspecific T cells (P14), with the parasite shown in green (with coordinates (290, 210, 76)). Tracks are shown over time, with early timepoints lighter in color than later timepoints. The data can be explored in greater detail in an interactive Unity environment at `https://viktorzenkov.github.io/measuringAttraction`.

(c) If this distribution's test statistic is significantly different from a random test statistic, then that corresponds to "bias", and if it is not significantly different that corresponds to unbiased movement. If a cell is biased with $\kappa_a > 0$, then the cell is attracted, and if a cell is biased with $\kappa_a < 0$, then the cells is repulsed.

(d) A followup test uses the multinomial distribution. For a set of cells, each of which has had its bias ascertained as attracted, repulsed, or biased, we set our null distribution to a multinomial distribution with probabilities set to the proportion of cells detected as unbiased, p_u, and the other probabilities each set to $\frac{1-p_u}{2}$. This takes into account the proportion of cells that are unbiased while still allowing a test between attracted and repulsed quantities.

4. Metric 4: average angle metric.

   (a) The average angle metric uses the angle of movement from Metric 1.

   (b) To test, we use a Student T test to compare the mean of the angles for a cell to the expected average angle, 90 degrees.

Statistical/computational analyses

Von Mises-Fisher distribution. The Von Mises-Fisher distribution chooses an n-dimensional (we use $n = 3$) vector given a direction vector and κ, a concentration parameter[25,26]. The output vector is chosen randomly with a bias toward the given vector whose strength is dependent on κ, with $\kappa \to 0$ having no bias, $\kappa > 0$ having positive bias (attraction), and $\kappa < 0$ having negative bias (repulsion). The probability density function for a vector χ in the Von Mises-Fisher (vMF) distribution in 3D is

$$P(\chi|\mu,\kappa) = \frac{\kappa \mathrm{e}^{\kappa\mu^T\chi}}{2\pi(\mathrm{e}^{\kappa} - \mathrm{e}^{-\kappa})}, \tag{S1}$$

where μ is the direction vector toward which there is bias and $|\mu| = |\chi| = 1$. To calculate the angle ϕ between vectors χ and μ, we let $\mu = (0, 0, 1)$ and $\chi = (x, y, z)$. Given that $|\chi| = 1$, the angle ϕ then only depends on the value of z and can be calculated from the relationship $\cos(\phi) = z$. With that transformation and using Mathematica we found the following probability density function for the angle ϕ:

$$P(\phi|\kappa) = \frac{\kappa \sin(\phi)\mathrm{e}^{\kappa\cos(\phi)}}{2\sinh(\kappa)}, \qquad 0 \le \phi \le \pi, \tag{S2}$$

where ϕ is the angle between the vector to the attraction point and the cell movement vector. Note that $\lim_{\kappa\to 0} P(\phi|\kappa) = \frac{1}{2}\sin(\phi)$. Because the magnitude of the concentration parameter is not intuitive, another useful parameter is the fraction of acute angles towards the attraction site. This fraction is found by integrating the vMF distribution

$$f_a(\kappa) = \int_0^{\pi/2} P(\phi)\mathrm{d}\phi = \frac{1}{1+e^{-\kappa}}, \tag{S3}$$

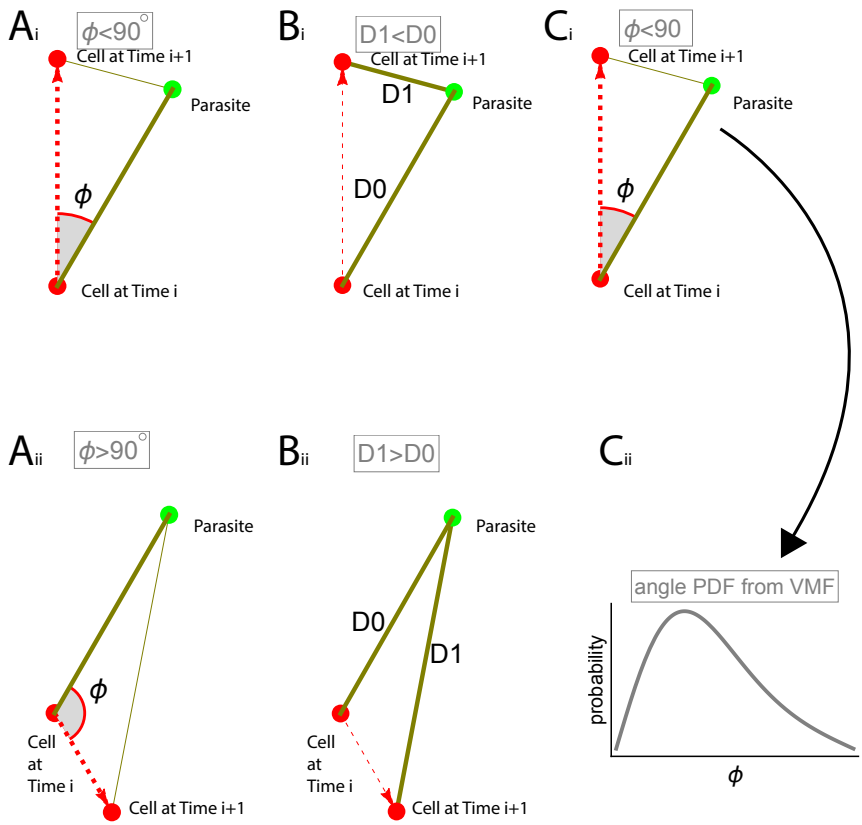

Figure S3: Three alternative metrics were used to test for attraction of CD8 T cells towards the infection site. Panels A: the angle metric is defined as the angle between the vector from the cell to the parasite and the vector of the cell's movement. An acute angle corresponds to "getting closer" (Ai) and obtuse corresponds to "getting farther" (Aii). Panels B: the distance metric is defined as the change between the distances from the cell to the parasite before and after the cell moves. A negative change of distance corresponds to "getting closer" (Bi) and positive corresponds to "getting farther" (Bii). Panels C: the angle distribution metric takes into account actual values of angles as in the angle metric and compares these values with the Von Mises-Fisher distribution (eqn. (1)) leading to an estimation of the concentration parameter κ_a (measure of attraction). If $\kappa_a > 0$ and this distribution's test statistic is significantly different from a random test statistic ($p < 0.01$), we consider the cell to have "attraction". If the test statistic is not significantly different then we consider the cell to not have bias. If $\kappa_a < 0$ and the test statistic is significantly different then we consider the cell to be repulsed.

where $f_a(0) = 1/2$ and $f_a(\infty) = 1$.

Modeling T cell attraction to the parasite. We use the vMF distribution to model T cell attraction to the parasite, and we define κ_a as the concentration parameter determining the strength of this attraction.

Modeling correlated (persistent) random walk. Previous studies employed various methods to simulate a persistent random walk, such as assuming a normal distribution of turning angles with mean 0 and some variance (e.g.,[40,41]). Other studies used Orenstein-Uhlenbeck model to simulate persistent random walks[42–44]. We propose using the vMF distribution to model persistent random walk using the concentration parameter κ_t for turning angles. In this case, turning angles are sampled

from the vMF distribution with concentration κ_t. There is a direct relationship between the average turning angle $\bar{\phi}$ of a cell moving in a persistent random walk and the concentration κ_t

$$\bar{\phi} = \int_0^{\pi} \phi P(\phi|\kappa_t) \mathrm{d}\phi = \frac{\pi \cosh(\kappa_t)\Big(I_0(\kappa_t) - \mathrm{e}^{-\kappa_t}\Big)}{2}, \tag{S4}$$

where $I_0(\kappa)$ is the Bessel function of the first kind of order 0.

Estimating the concentration parameter of the vMF distribution. We sum the logs of the probabilities of getting each experimentally determined angle ϕ_i from the vMF angle PDF (eqn. (1)) to get the negative log likelihood function

$$\mathcal{L}(\kappa_a) = -\sum_{i=1}^{n} \log\Big(P(\phi_i|\kappa_a)\Big), \tag{S5}$$

where n is the number of movements in the data. The maximum likelihood estimate κ_a is found by maximizing eqn. (S5). Whether the found estimate $\hat{\kappa}_a$ is different from zero is determined using a likelihood ratio test by comparing $\mathcal{L}(\hat{\kappa}_a)$ and $\lim_{\kappa_a \to 0} \mathcal{L}(\kappa_a)$[45].

It is straightforward to include a mixture of vMF distributions to explain angle distribution data. In particular we found that while a single vMF distribution describes the distribution of angles to parasite reasonably well, there is appears to be an enrichment of 90^0 angles in the data (Figure S18). Therefore, in this case we fitted the following mixture distribution to the data

$$P(\phi|\kappa, f, n) = f\frac{\kappa \sin(\phi)\mathrm{e}^{\kappa\cos(\phi)}}{2\sinh(\kappa)} + (1-f)\sin^n(\phi), \tag{S6}$$

where f is the fraction of cells that are biased towards the infection, $\sin(\phi)$ is proportional to vMF distribution with no bias, and n is used to explain the over-abundance of angles at 90^0. Improvements of this model fit over the simpler model, based on a single vMF distribution, can be tested using likelihood ratio test.

Stochastic simulations. We can simulate a moving cell by generating a set of movements, each of which is defined by a distance traversed and a direction vector of the cell's movement. To generate the movement distances, we picked a distribution that best represented the real data from our datasets using Mathematica to estimate the best distribution for a dataset. We picked the Pareto Type IV (generalized Pareto, GP) distribution

$$f(x) = \left(1 + ((x-\mu)/k)^{(1/\gamma)}\right)^{-\alpha}, \tag{S7}$$

with the following best fit parameters: minimum value $k = 7$, location parameter $\mu = 0$, scale parameter $\alpha = 2.7$, and inequality parameter $\gamma = 0.68$, assuming that these movements occur in 2 min intervals. These parameters were found by fitting the GP distribution to movement data from datasets #3 and #4 pooled together using the function `FindDistributionParameters` on a

`ParetoDistribution` in Mathematica. Of note, the GP was found to best fit the movements of malaria-specific CD8 T cells in the absence of the infection[38]. To simulate attraction of T cells towards the parasite each direction is chosen from the vMF distribution with the angle between T cell movement and the parasite given in eqn. (1). The concentration κ_a was chosen based on the desired amount of attraction ($\kappa_a > 0$ corresponds to attraction, $\kappa_a = 0$ corresponds to no bias, and $\kappa_a < 0$ corresponds to repulsion). To create a desired vector, we first choose a vector with bias toward direction {0,0,1}, which simplifies the process to choosing x and y randomly from a normal distribution $N(0, 1)$ and choosing z based on the von Mises-Fisher distribution, chosen by

$$z = 1 + (\ln(r) + \ln(1 + (1 - r)\frac{e^{-2\kappa_t}}{r}))/\kappa_t, \tag{S8}$$

where r is chosen uniformly between 0 and 1, exclusively[46]. Then we weight x and y to place the chosen vector on the unit sphere, and we then use a rotation transform to adjust the generated vector with respect to the desired bias direction.

To simulate a persistent random walk, each movement direction is chosen from a vMF distribution with a vector input of the vector from the cell's previous position to the cell's current position; the concentration κ_t is chosen based on the desired amount of persistence, with $\kappa_t > 0$ corresponding to persistence, $\kappa_t = 0$ corresponding to randomness, and $\kappa_a < 0$ corresponding to the cell turning back on itself. For a constant parasite position, a sequence of movements is simulated consecutively, resulting in a list of cell positions similar to those from the experimental datasets. We can also simulate a walk with both attraction to a parasite and persistence by choosing two directions with bias as described previously and summing each pair of directions to find a direction with bias influenced by both the parasite and persistence.

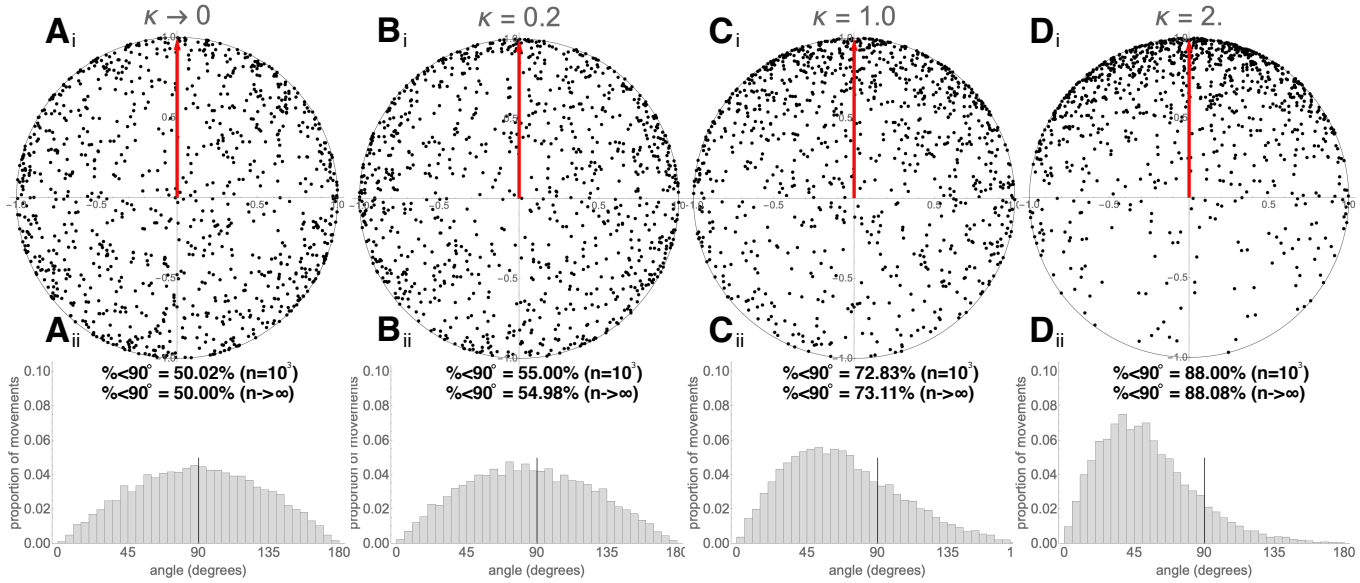

Figure S4: Illustration of the random vectors generated using the Von Mises-Fisher (vMF) distribution. The vMF distribution generates direction vectors with preference toward a given direction, with the strength of preference determined by the concentration parameter κ. We sampled 1000 directions from a vMF distribution for several values of κ (noted on individual panels) and plotted those in a 2D image by removing the z-coordinates (panels i). We also show histograms of the actual angles between the movement vector and the vector to the parasite (panels ii, angle metric defined in Figure S3). In panels ii we also show the percents of acute angles generated for the $n = 10^3$ randomly sampled vectors and calculated using the exact formula, eqn. (S3).

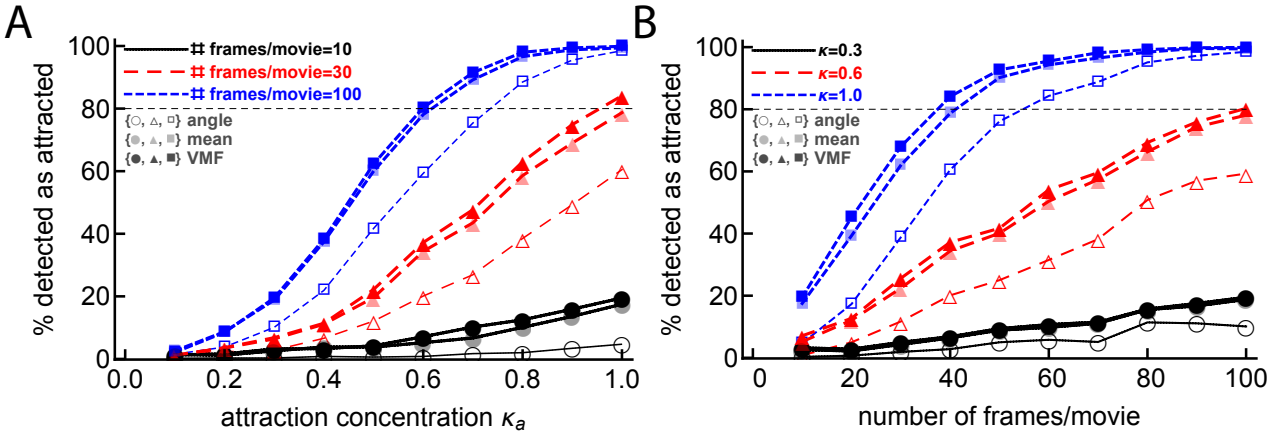

Figure S5: Third metric based on comparing distribution of angles between movement vector and the vector to the parasite is most powerful out of tested metrics. We simulated T cell movement with varying degrees of bias towards the parasite for different movement durations as described in Figure 4. Simulations were used because only with simulations can we make data which contains a small amount of attraction and for which we *know* there is a small inherent amount of attraction; for real data, we do not know the amount of attraction, that is our objective. For every track we determined which of the angle metric (metric 1, open symbols), mean angle metric (metric 4, transparent symbols), and the vMF distribution-based metric (metric 3, solid symbols) better detects attraction towards the infection site. In total, we simulated 1000 cells per every set of parameters. We note that the third metric is most sensitive - to detect 80% of cells as attracted with 100-frame movies, a simulated κ_a of 0.6 is necessary for the third metric, while the necessary κ_a for the mean angle test is around 0.62 and the necessary κ_a for the binomial angle test is around 0.74. We also note that as κ_a approaches 0, the percent of cells detected as attracted decreases, indicating a small false positive rate.

Measuring attraction using angles (in the first and third metric) is biased for correlated random walks

Our third metric test requires a null hypothesis (random) angle bias κ_a^0 to calculate a test statistic. The value of that random κ_a^0 is not immediately clear for our data. In general, $\kappa_a^0 = 0$ should be the null hypothesis. However, this does not hold in all circumstances. We found that for a persistent walk (cells tending to move in the same direction), the estimated $\hat{\kappa}_a$ will be less than 0. As a result, more cells will be detected as repulsed than attracted if the path is due to a persistent walk.

Demonstrating this, we simulated cells attempting to recreate the situation of the real cells as much as possible, with the assumption that cells do not display attraction. We simulated a persistent walk by choosing the direction of each movement with bias toward the previous direction (using a Von Mises-Fisher distribution with a constant chosen persistence bias κ_t and the previous direction to generate each direction). We therefore had a persistent walk scenario and observed a disparity between the detection of attraction and repulsion (Figure S6): the number of simulated cells detected as repulsed for a similar box size to our real experiments ($500 \times 500 \times 50$ μm) is larger than the number of cells detected as attracted. Since our real data is similar to a persistent walk due to the presence of sinusoids, a proper test for real data must use a proper random $\kappa_a^0 < 0$. Nearly equal numbers of cells were detected as attracted and repulsed for $\kappa_t = 0$, and about 6.5% and 9% of cells were detected as attracted and repulsed, respectively, for $\kappa_t = 2$ (Figure S6Bii).

Interestingly, the size of the imaging area did not impact the fraction of cells detected as attracted for any degree of persistence; however, larger imaging areas tended to increase the fraction of cells

detected as repulsed (Figure S6ii). For example, in simulations with unlimited imaging area (no box), 11% of cells were detected as repulsed for $\kappa_t = 2$ (Figure S6Cii). This is not especially surprising, as cells that are detected as repulsed tend to move away from the parasite and are more likely to be leave smaller imaging areas (i.e., smaller box).

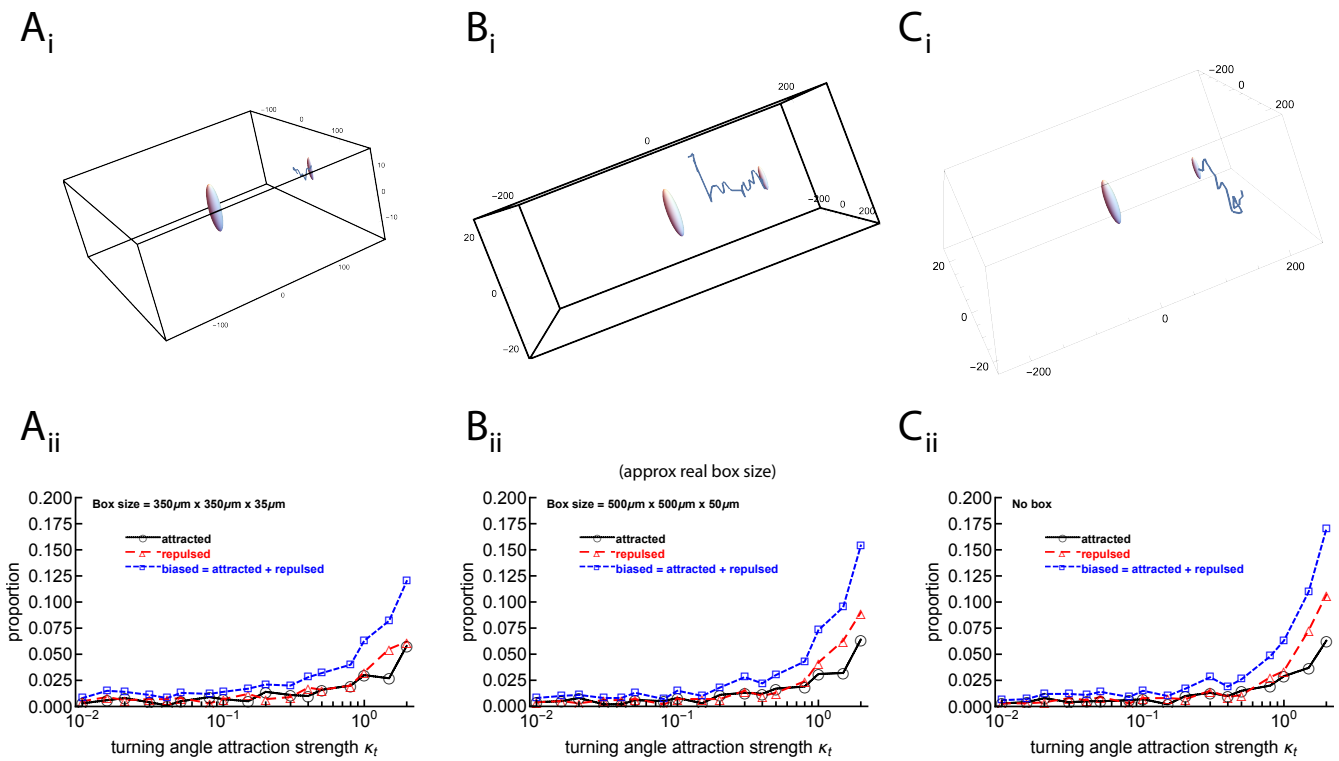

Figure S6: Quantifying bias in the angle distribution (vMF distribution-based) metric. We run stochastic simulations of T cells moving in a small "boxed" space ($350 \times 350 \times 35$ μm, panels A), regular boxed space as used in intravital imaging experiments ($500 \times 500 \times 50$ μm, panels B), or in open space (panels C). The parasite is assumed to be located in the center of the box and T cells start the movement at 150 μm from the parasite at coordinates $(x, y, z) = (150, 0, 0)$. We assume 50 movements per simulated T cell with movement lengths described by the Generalized Pareto distribution (eqn. (S7)) with parameters given in Materials and Methods. To simulate a persistent random walk, we assume that T cell turning angles are described by a vMF distribution with constant κ_t determining the bias of a T cell walk (see Materials and Methods for more detail). A cell is assumed to be out of the imaging area when the cell's position is outside of the defined imaging box (in A and B) which then results in fewer than 50 moves recorded. For each cell's track, we determine if the cell is attracted ($\kappa_a > 0$), repulsed ($\kappa_a < 0$), or unbiased ($\kappa_a = 0$) to the parasite's position by comparing the distribution of angles to the parasite to a vMF distribution (Figure S3) and estimating the concentration κ_a. For each concentration determining the persistence of the random walk κ_t, we simulate movement of 1000 cells. In Panels i are examples of a cell whose path is cut off by the small box (Ai), a cell detected as attracted (Bi), or a cell detected as repulsed (Ci). Note that in Panels i the figures are not to scale, resulting in the spheres representing the cells being elongated. In these examples, the parasite and cell's starting position and cell's moving track are shown. Parasites are depicted as a sphere with a radius of 40 μm representing infected hepatocyte[10,11].

We do not currently have an analytical method to determine analytically the proper null hypothesis κ_a^0. However, we can make an estimate for the random κ_a^0 using simulations. We simulate movements of 100 random cells for each real cell and calculate the average estimated κ_a of those cells, and we use that as our new null hypothesis κ_a^0 for the third metric tests. It is important to note, however, that in our experience the estimated null hypothesis κ_a^0 did not differ much from 0 (around -0.06).

We theorized that the original angle metric may also be biased in a persistent walk scenario. To correct, we used the same data simulated to replicate real data and, for each simulated cell, calculated the proportion of movements which are toward the parasite (acute). The average proportion for each real cell is then the p used in the binomial test.

Supplemental references

41. Ahmed, D. A. & Petrovskii, S. V. Analysing the impact of trap shape and movement behaviour of ground-dwelling arthropods on trap efficiency. *Methods Ecol Evol* **10,** 1246–1264. ISSN: 2041-210X (2019).

42. Uhlenbeck, G. E. & Ornstein, L. S. On the Theory of the Brownian Motion. *Phys. Rev.* **36,** 823–841. `https://link.aps.org/doi/10.1103/PhysRev.36.823` (5 Sept. 1930).

43. Wu, P.-H., Giri, A., Sun, S. X. & Wirtz, D. Three-dimensional cell migration does not follow a random walk. *Proceedings of the National Academy of Sciences* **111,** 3949–3954. ISSN: 0027-8424. eprint: `https://www.pnas.org/content/111/11/3949.full.pdf`. `https://www.pnas.org/content/111/11/3949` (2014).

44. Jerison, E. R. & Quake, S. R. Heterogeneous T cell motility behaviors emerge from a coupling between speed and turning in vivo. *eLife.* `https://elifesciences.org/articles/53933` (2020).

45. Pawitan, Y. *In All Likelihood: Statistical Modelling and Inference Using Likelihood* 544 (Oxford University Press, 2001).

46. *Vectors with a certain magnitude in Mathematica* `https://mathematica.stackexchange.com/questions/13038/vectors-with-a-certain-magnitude-in-mathematica/13042#13042`.

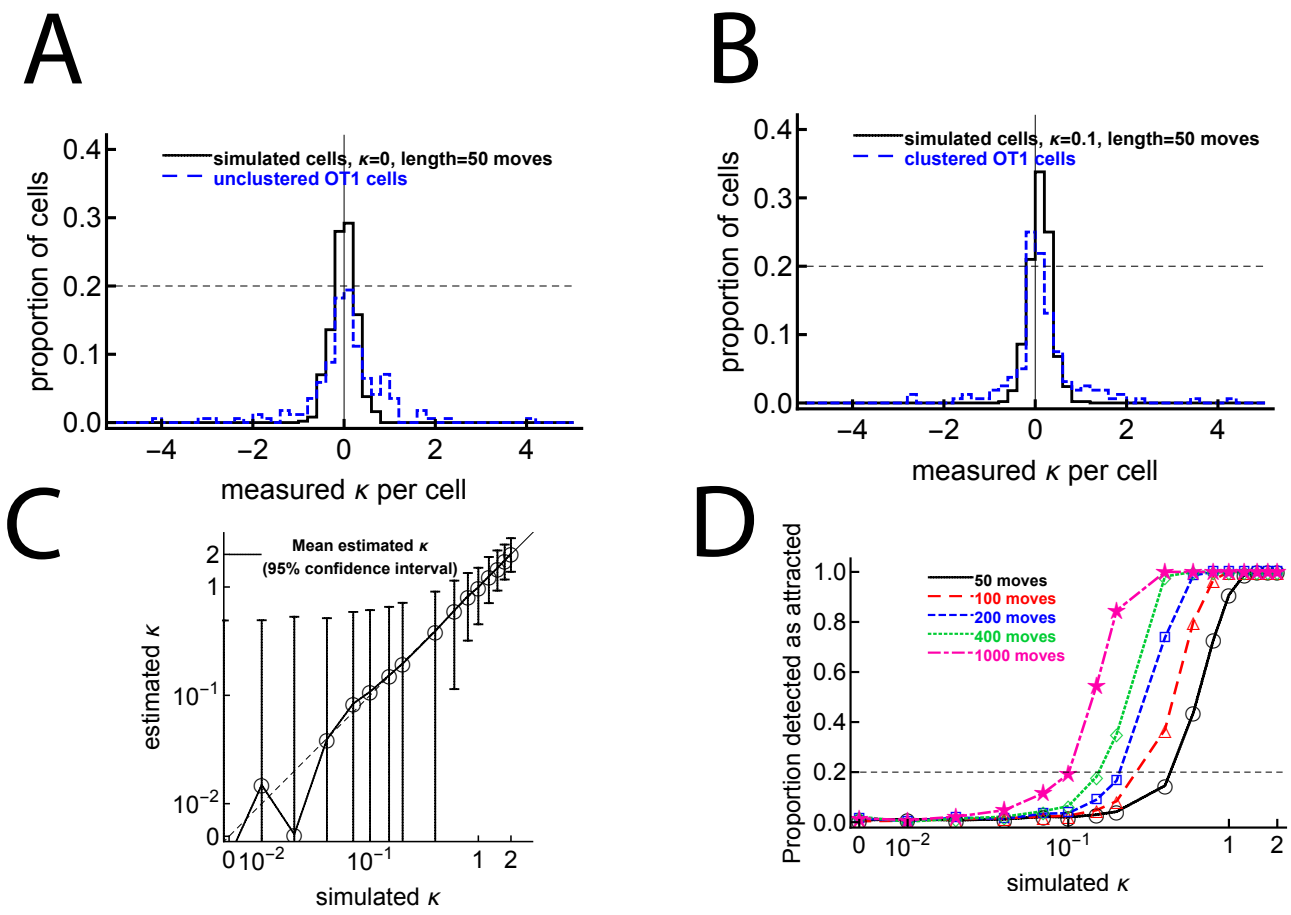

Figure S7: Analyzing the distribution of κ_a s estimated from cells simulated with a uniform bias. We simulated CD8 T cells searching for a malaria liver stage assuming variable levels of T cell attraction to the parasite, defined by the concentration parameter κ_a of the vMF distribution. We also simulated experiments of various lengths, between 50 and 10000 movements. All cells start 150 μm from the parasite. To simulate a T cell walk we assumed that T cell movement lengths follow a generalized Pareto distribution (eqn. (S7)) with randomly chosen turning angles. Simulations were done for 500 cells for each value of κ_a and each experiment length. In panels A and B we overlay distributions of detected κ_a's from real and simulated dataset; panel A shows data for unclustered cells and simulated cells with no inherent bias, while panel B shows data for clustered cells and simulated cells with a minor bias ($\kappa_a = 0.1$) Panel C shows the mean estimated κ_a from the 500 simulated cells for the simulations with 50 moves and each original κ_a; for perfectly simulated data and a perfect test, the datapoints would lie in the line $y = x$ (shown as dashed line), error bars denote 95% confidence intervals. Panel D shows the proportion of cells detected as attracted for each combination of simulated κ_a and number of movements. This is similar to Figure 4A, but emphasizing the threshold of $\kappa_a = 0.5$ at which percent of cells detect as attracted is similar to the clustered data but at this κ_a average attraction in simulations does not match that in the data.

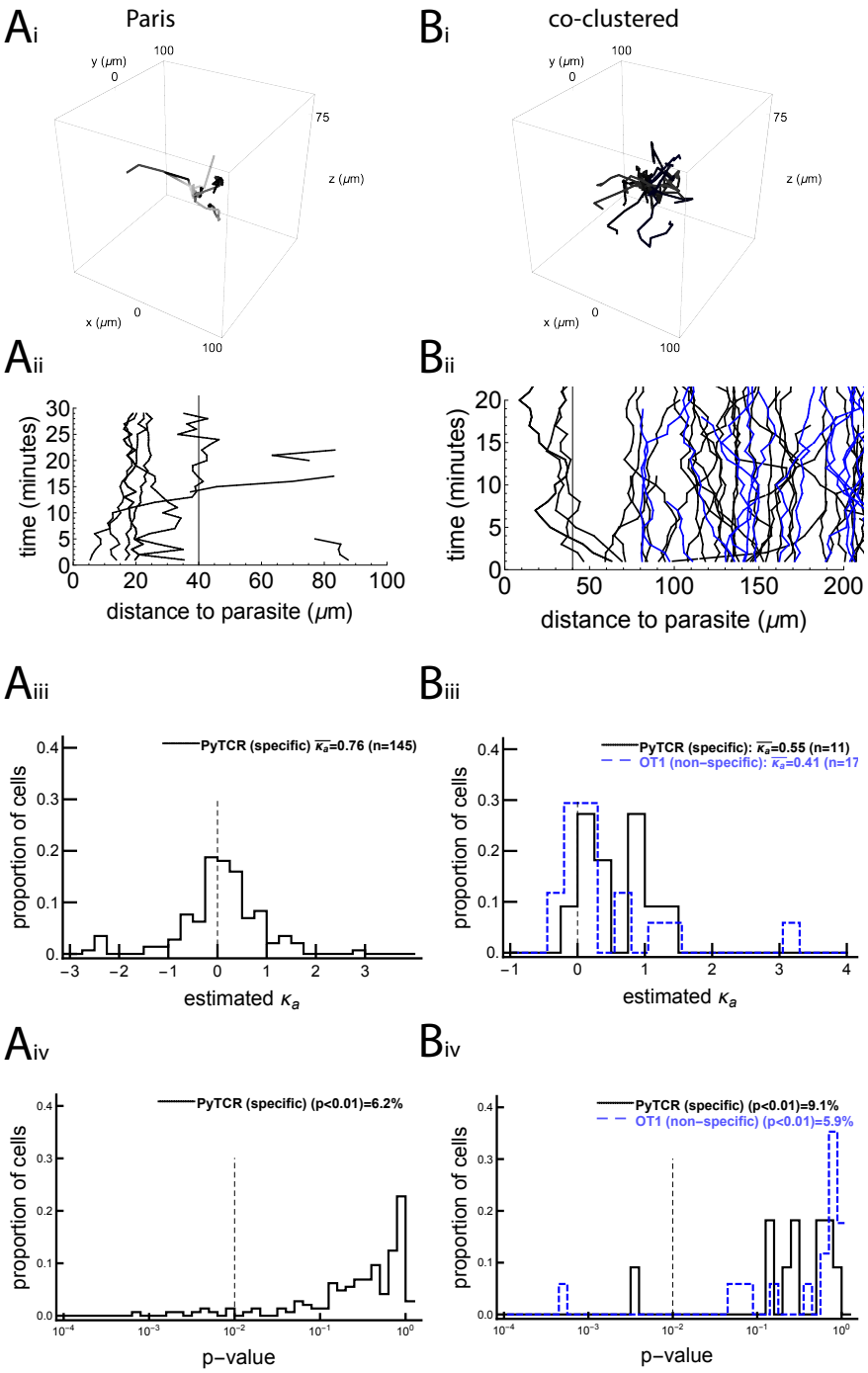

Figure S8: A minority of activated CD8 T cells display movement bias to the liver stage. The analysis was similar to Figure 2 for the previously published data[10,12]. In the co-clustered data, CD8 T cells, specific to Py (PyTCR) or specific to a irrelevant antigen (OT1), in one mouse with one Py liver stage were imaged (see dataset #1). In the Paris data there were no nonspecific cells, but there were four mice with several infection sites each. For the Paris data, we detect 5 PyTCR cells as attracted and 4 PyTCR cells as repulsed. For the co-clustered data, we detect 1 PyTCR cells and 1 OT1 cells as attracted and 0 PyTCR cells and 0 OT1 cells as repulsed.

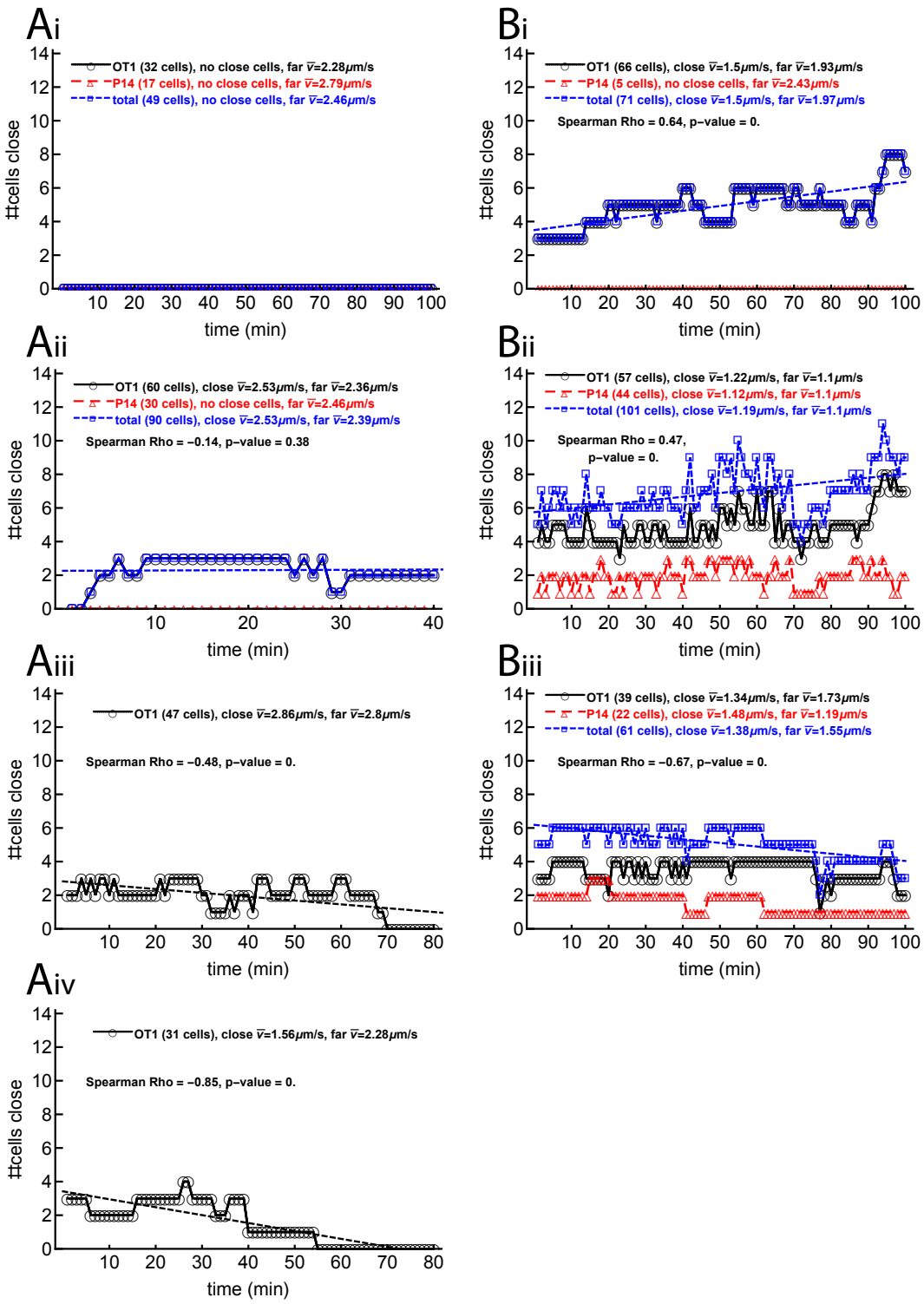

Figure S9: The numbers of cells close to the parasite change more over time for large clustered datasets than unclustered/small clustered datasets, according to the Spearman Rho test. The size of the cluster (the number of cells close to a parasite) is of interest, because if the cluster size increases, then cells are clearly coming to the parasite and not leaving. Shown are the number of cells close to the parasite (within 40μm from the parasite) at each time point, with one panel for each parasite. The left column is for the unclustered/small clustered dataset, and the right column is for the large clustered dataset.

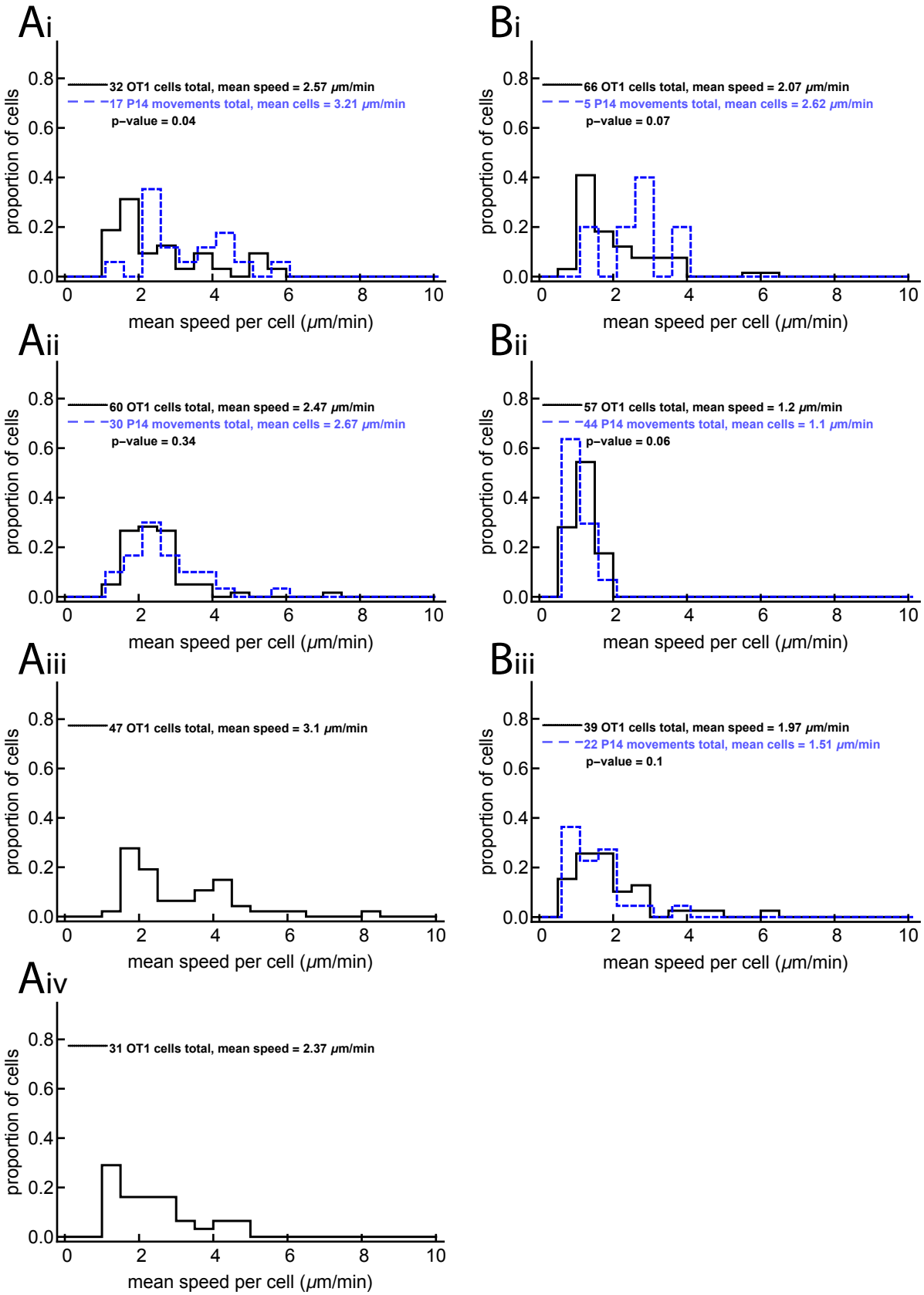

Figure S10: Distribution of average speed per cell in different imaging experiments. Left panels are for unclustered/small clustered dataset, and right panels are for large clustered dataset. The p-values come from the Mann Whitney test.

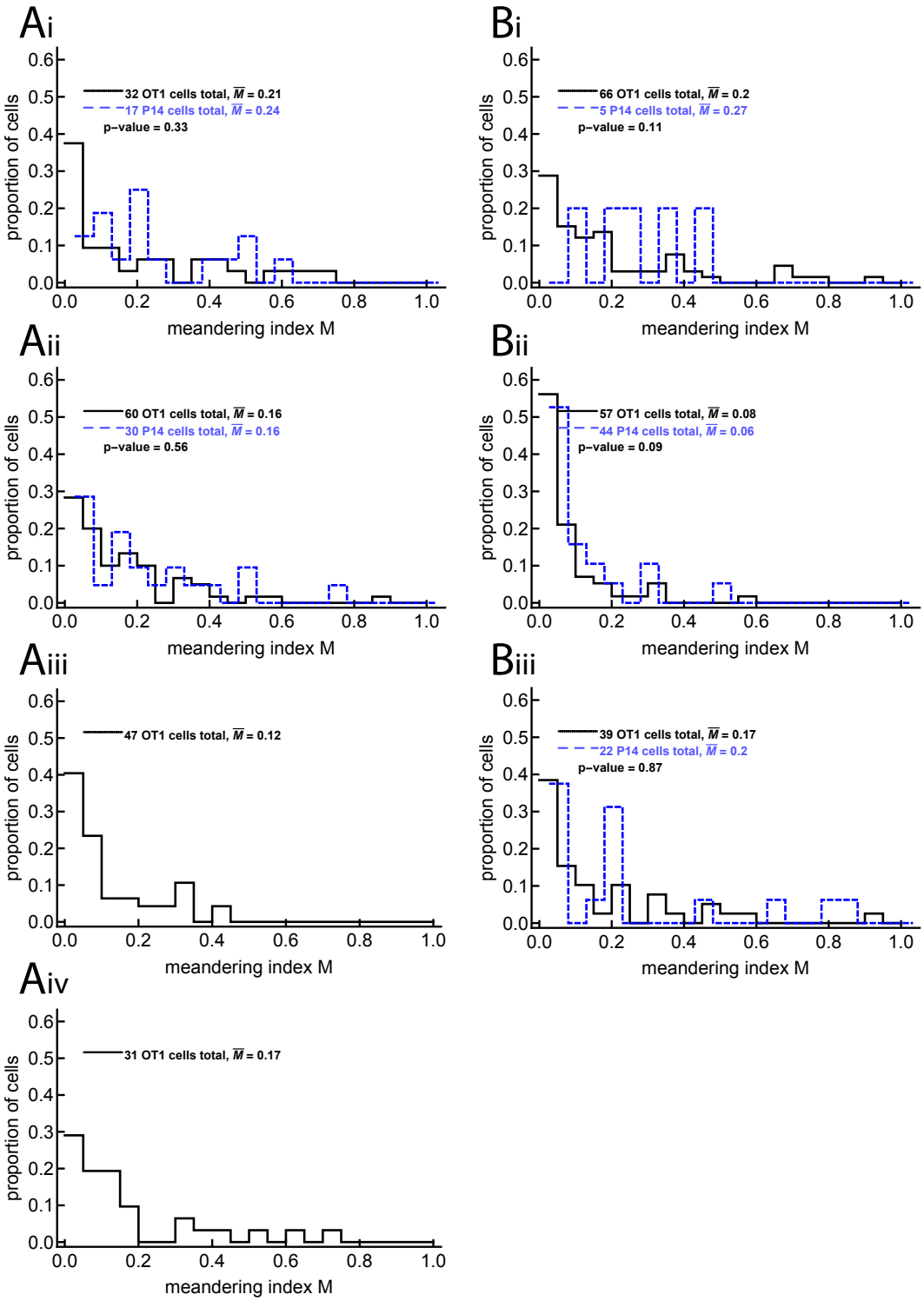

Figure S11: Distribution of meandering index for different cells in the experiments with dataset #1 (unclustered/small clustered dataset) and #2 (large cluster dataset). The meandering index M is calculated as the distance between the first and last recorded positions of the cell divided by the total length of the cell's movement in space; it is therefore between 0 and 1, with 0 equivalent to the cell returning to its original position and 1 meaning the cell moved in a straight line. In our real data, the meandering indices are mostly around 0.15 to 0.2, and identical between corresponding specific and non-specific datasets. The p-values come from the Mann Whitney test.

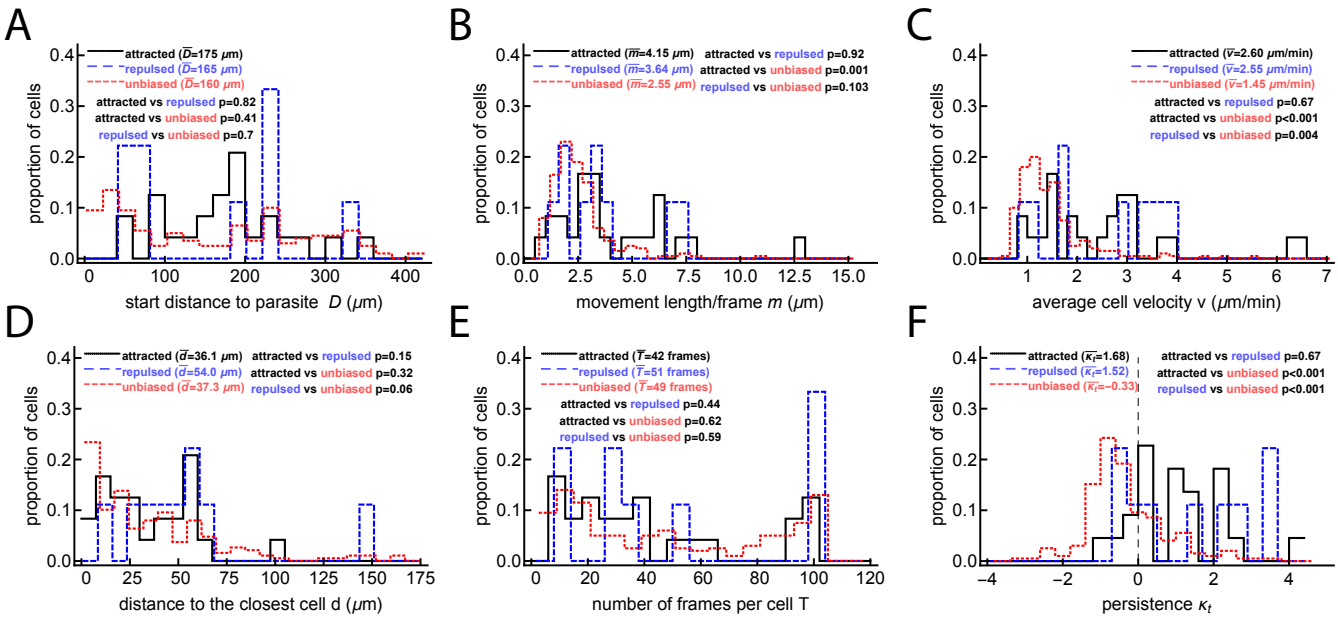

Figure S12: Speed and turning angle correlate with detection of attraction when several T cells have found the parasite. These are metadata results when data are divided into attracted, repulsed, and unbiased T cells for the large clustered dataset (dataset #2). For more details, see Figure 3.

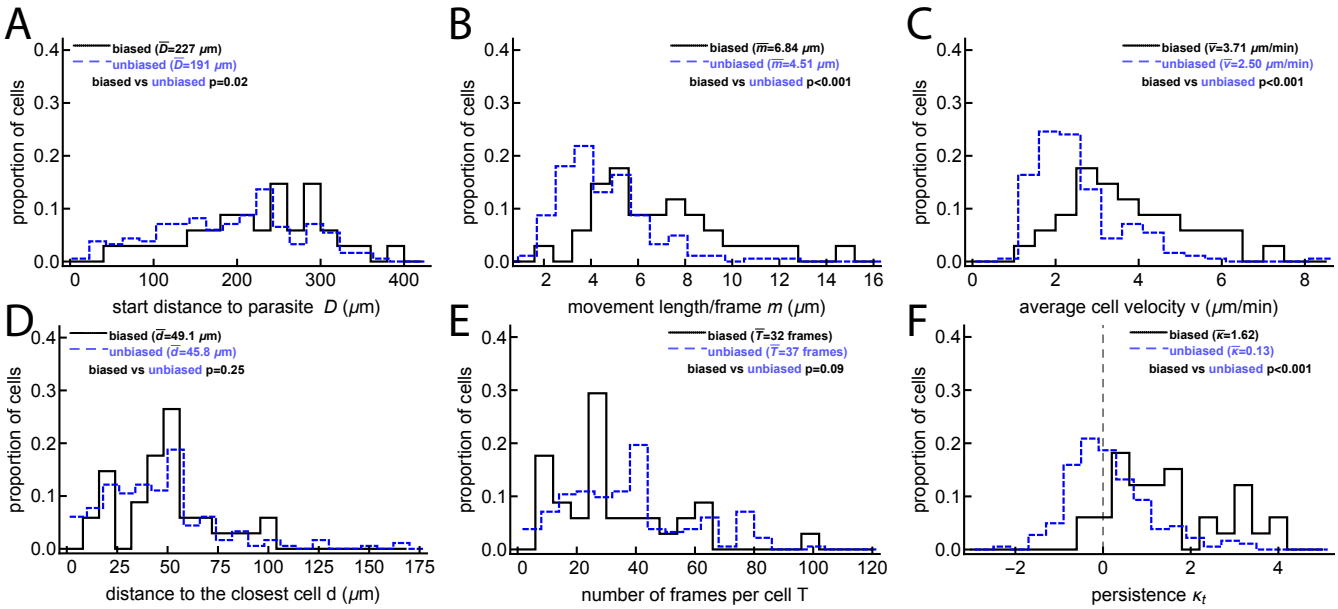

Figure S13: Speed and turning angle correlate with detection of attraction when no or few T cells have found the parasite. These are metadata results when data are divided into biased and unbiased T cells for the unclustered/small clustered dataset (dataset #1). For more details, see Figure 3.

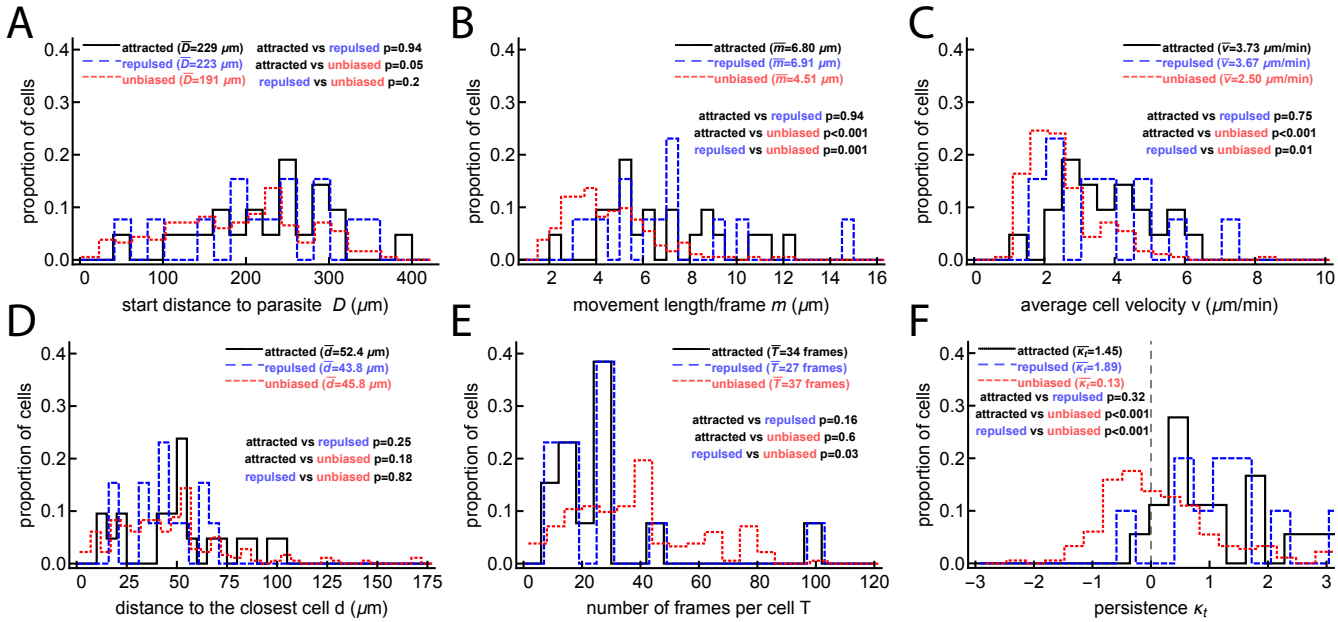

Figure S14: Speed and turning angle correlate with detection of attraction when no or few T cells have found the parasite. These are metadata results when data are divided into attracted, repulsed, and unbiased T cells for the unclustered/small clustered dataset. For more details, see Figure 3.

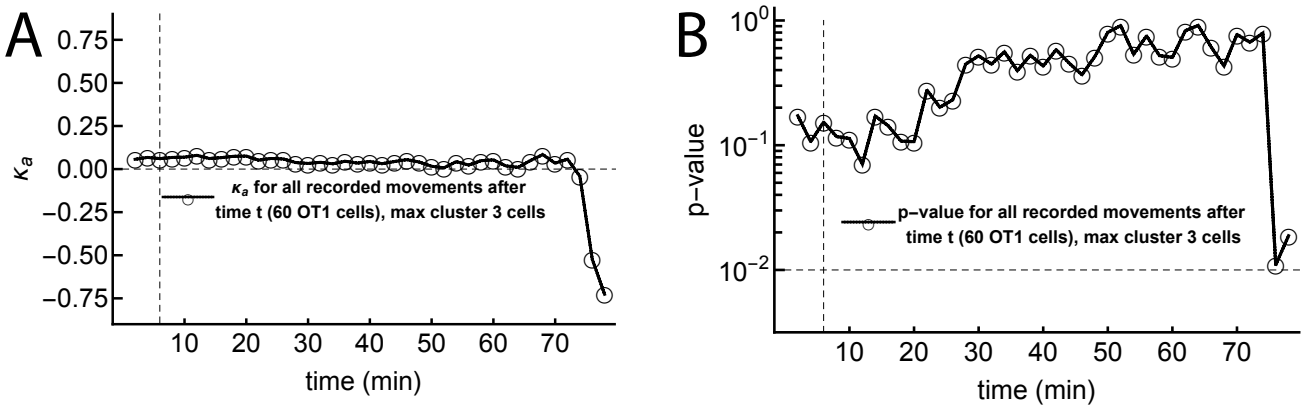

Figure S15: An "Attenborough" experiment, where the first cell to reach the parasite does so during the recorded positions, is of interest because it could demonstrate a changing amount of attraction over time. The second of the unclustered/small clustered datasets is such an Attenborough experiment, with the first cell finding the parasite six minutes into the recorded positions. After this cell and two others within a few minutes, no more cells reach the parasite.

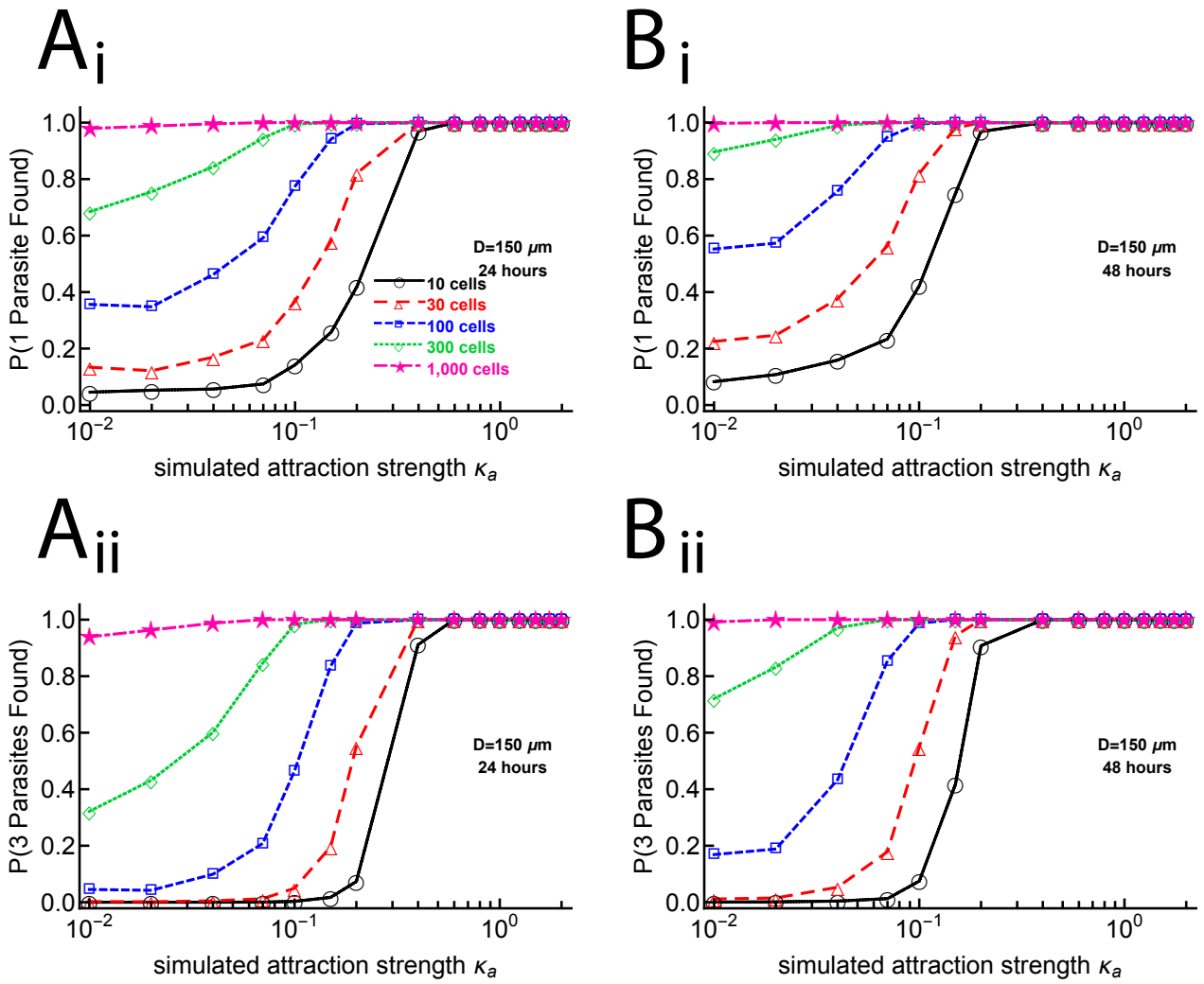

Figure S16: Many CD8 T cells per parasite are needed to ensure that all parasites are found within 48 hours after infection. Groups of 10, 30, 100, 300, and 1000 cells per parasite are simulated 1000 times for values of κ_a between 0 and 2 with a starting distances of 150 μm and are "sampled" every two minutes for lengths of 24 (panels A) or 48 (panels B) hours. Each parasite that has a cell reach within 40μm of it is considered a success. Panels i show the probability p of a parasite being found, while panels ii show the probabilities of three parasites being found (p^3). If there are 100 cells per parasite, a κ_a of 1.0 is required for essentially all parasites to be found within 48 hours.

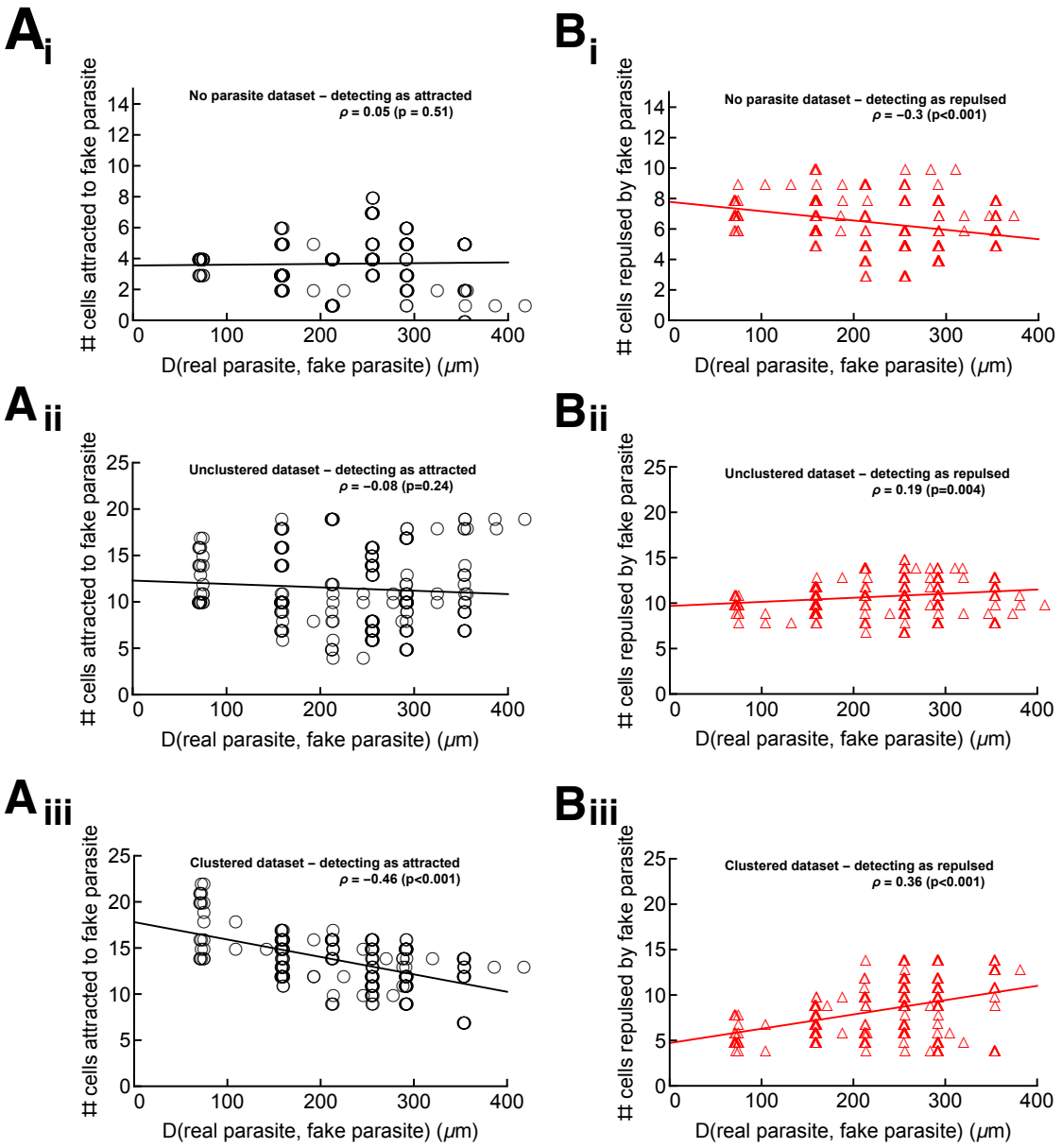

Figure S17: Tests for attraction were performed with other positions of potential attraction to determine if cells display attraction to points without significance. Positions were chosen on a grid extending 250μm away from the parasite, then our third metric test was applied using angles to that position instead of to the real parasite position. The distance to the real parasite is used as the x axis in each graph, and the number of cells total detected as attracted and repulsed are shown in panels A and B, respectively. Panels i correspond to the control data, panels ii correspond to the unclustered/small clustered dataset, and panels iii correspond to the large clustered dataset. The changes in the numbers of cells detected as attracted or repulsed are tested using the Spearman Rank test. For the unclustered/small clustered dataset, the number of cells detected to any position barely changes, suggesting that there is nothing special about the real parasite position, while for the large clustered dataset, there is a trend of more cells detected as attracted and less as repulsed for fake parasite positions close to the real parasite, suggesting that there is something special about the real parasite position. For the control data, the number of cells detected as attracted does not change, while the number detected as repulsed decreases for positions far away from the center of the data. In all cases there was heterogeneity about points with greater bias.

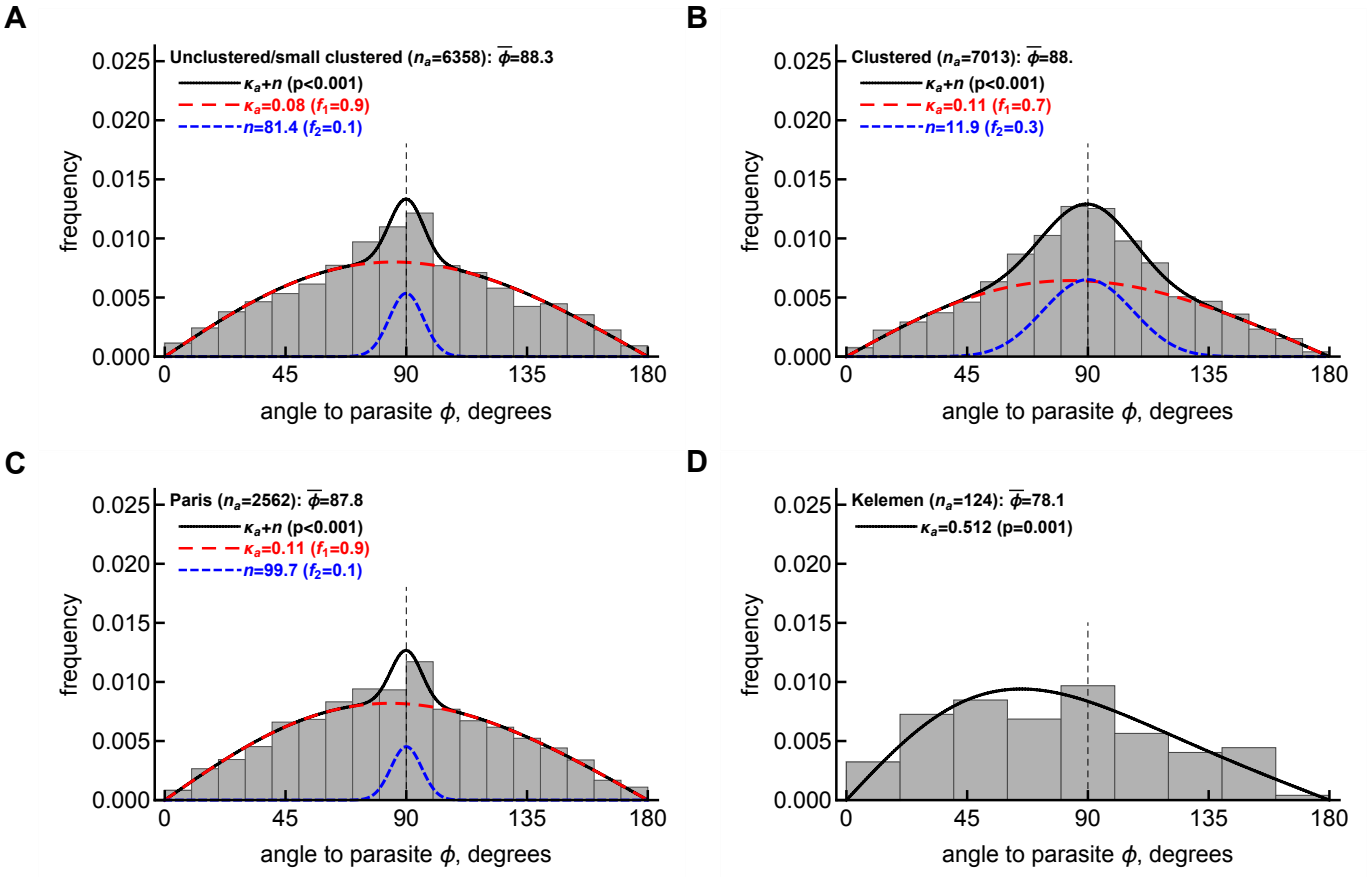

Figure S18: The von Mises-Fisher distribution describes relatively well the distribution of angles to the parasite from four independent datasets. We show the distribution of angles to the parasite by Plasmodium-specific liver-resident CD8 T cells as outlined in Materials and Methods for the unclustered/small clustered dataset (A), clustered dataset (B), Paris dataset (C), and co-clustered dataset (D). We fitted either a vMF distribution (given in eqn. (1)) or a mixture of distributions (eqn. (S6)) to these data and estimated the concentration parameter κ_a (noted on individual panels). Fits of the model are shown by lines. We compared every fit with the fit of a simpler model (a single vMF distribution fit in A-B or null distribution of angles with $\kappa_a \to 0$ in C-D) using a likelihood ratio test; resulting p-values are shown on individual panels. For each dataset we also plot the total number of angles in the dataset (n_a) and the average angle to the parasite ($\bar{\phi}$).