

SUPPLEMENTARY MATERIALS

Supplementary Figures and Legends

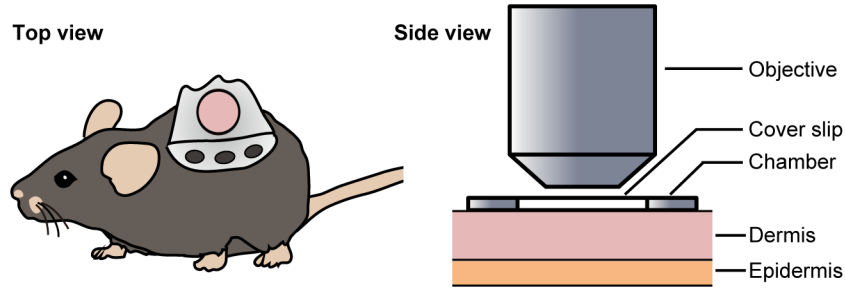


Figure S1. The top and side views of intravital microscopic imaging on mouse dorsal skinfold window chamber model used for tracking the searching behaviors of immunocytes *in vivo*.

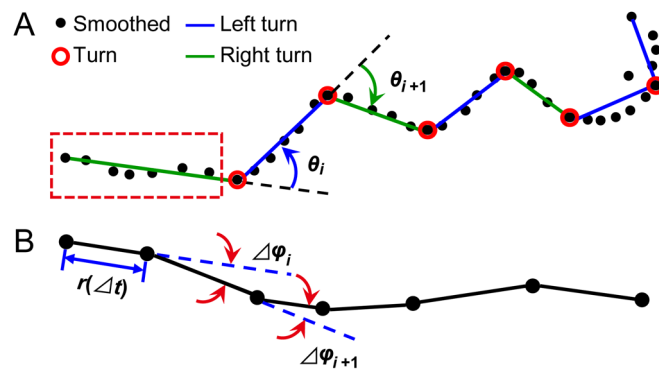


Figure S2. Recognition of turns within the cell trajectory. (A) Turn points (Turn) and turn angles (θ). Black dots indicate the smoothed cell trajectory (Smoothed). (B) Changes in directions of consecutive displacements $r(\Delta t)$, which were denoted as $\Delta\varphi$. The trajectory shown here is the magnified view of the area marked in (A) by the red rectangle.

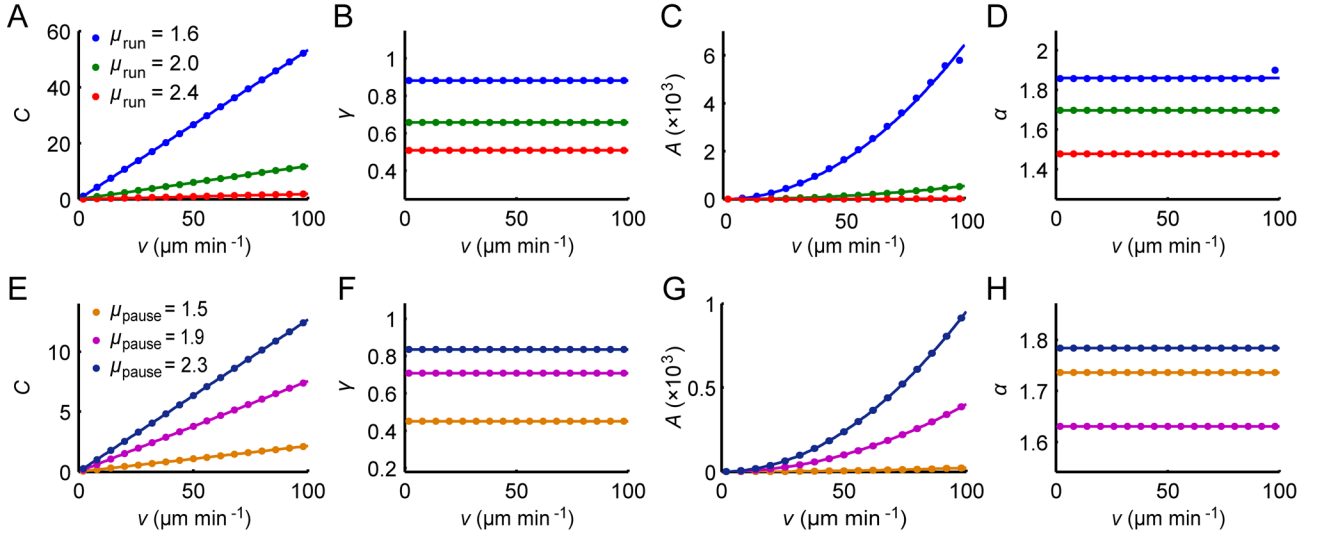


Figure S3. Dependence of the displacement scale factor $\zeta(t) = Ct^\gamma$ and the mean squared displacement m.s.d. $= At^\alpha$ on the running velocity v in the generalized Lévy walk model. (A, E) C linearly depended on v as $C = a_1 v$ and (C, G) A increased over v as $A = a_2 v^2$. a_1 and a_2 were constants for each specific Lévy exponent pair. Solid lines show the corresponding fitting results. (B, F) γ and (D, H) α were independent of v . Solid lines show the mean value over v from $1 \mu\text{m min}^{-1}$ to $100 \mu\text{m min}^{-1}$. Different colors represent different Lévy exponent pairs. For (A–D), $\mu_{\text{pause}} = 1.7$, whereas for (E–H), $\mu_{\text{run}} = 2.2$.

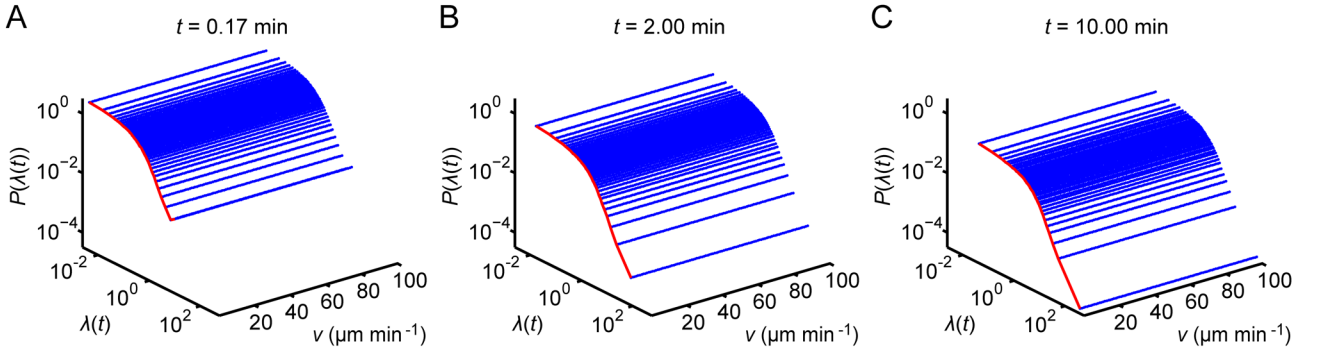


Figure S4. For the generalized Lévy walk model, the probability density distribution $P(\lambda(t))$ of scaled displacement $\lambda(t) = r(t) / \zeta(t)$, was independent of the running velocity v . Red solid lines outline the shape of $P(\lambda(t))$ at different time intervals: (A) $t = 0.17$ min, (B) $t = 2.00$ min and (C) $t = 10.00$ min.

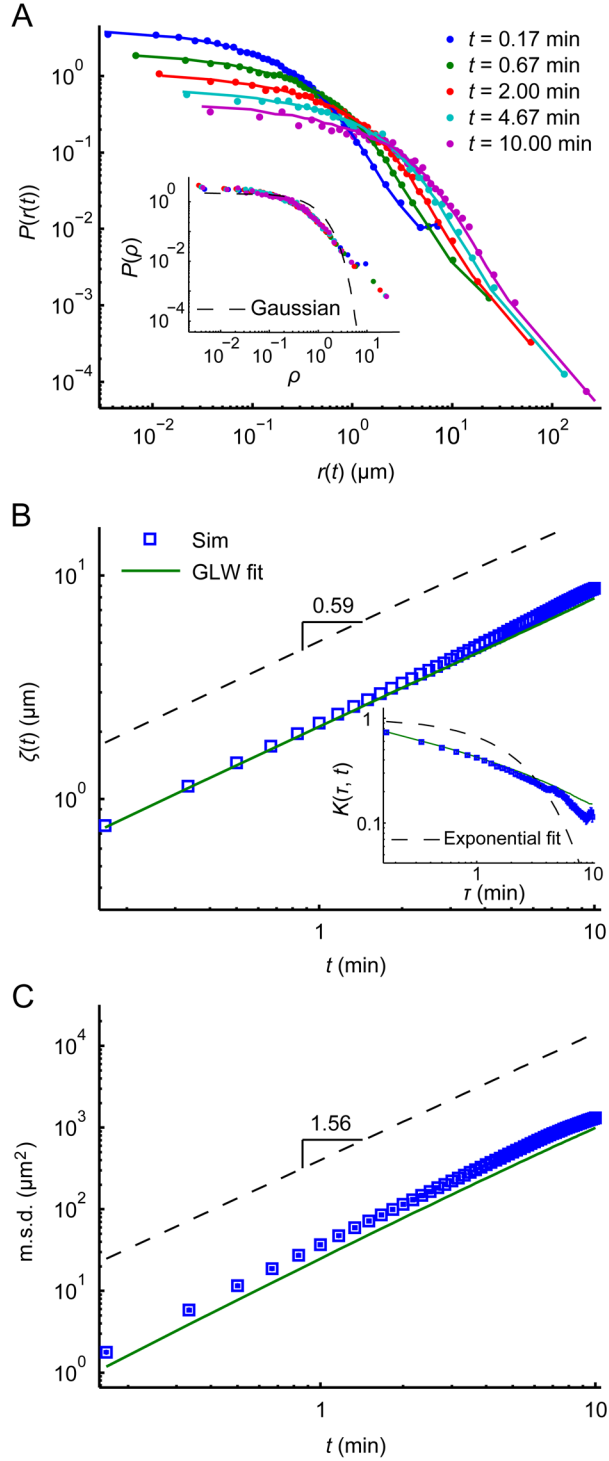


Figure S5. Displacement statistics of simulated cells (Sim) performing the generalized Lévy walk (GLW) compared with the corresponding GLW fitting. (A) The displacement probability density distribution $P(r(t))$ at different time intervals t . Colored dots indicate simulated data and solid lines indicate GLW fitting. The inset shows $P(r(t))$ at different t after normalizing the displacement $r(t)$ by the scale factor $\zeta(t)$ ($\rho = r(t) / \zeta(t)$). The tail of the normalized $P(r(t))$ was heavier than that of the Gaussian distribution (dashed line; Gaussian fitting of the normalized $P(r(t))$ at $t = 5$ min). (B) Displacement scale factor $\zeta(t)$ of the simulate data (squares) with the corresponding GLW fitting (solid line). $\zeta(t)$ increased approximately according to a power law t^ν ,

where $\gamma \approx 0.59$ (dashed line). The inset shows the normalized displacement correlation $K(\tau, t)$ of the simulate data (squares) with the corresponding GLW fitting (solid line). $K(\tau, t)$ decayed more slowly than exponentially (dashed line) over time τ . (C) Mean squared displacement m.s.d. of the simulated data (squares) with the corresponding GLW fitting (solid line). m.s.d. grew approximately according to t^α , where $\alpha \approx 1.56$ (dashed line). The error bars for $K(\tau, t)$ and m.s.d. denote the SEM. Simulated data: 600 cells performing GLW with $\mu_{\text{run}} = 2.2$, $\mu_{\text{pause}} = 1.7$ and $v = 50 \mu\text{m min}^{-1}$ and each trajectory covered a time length of 12.5 min.

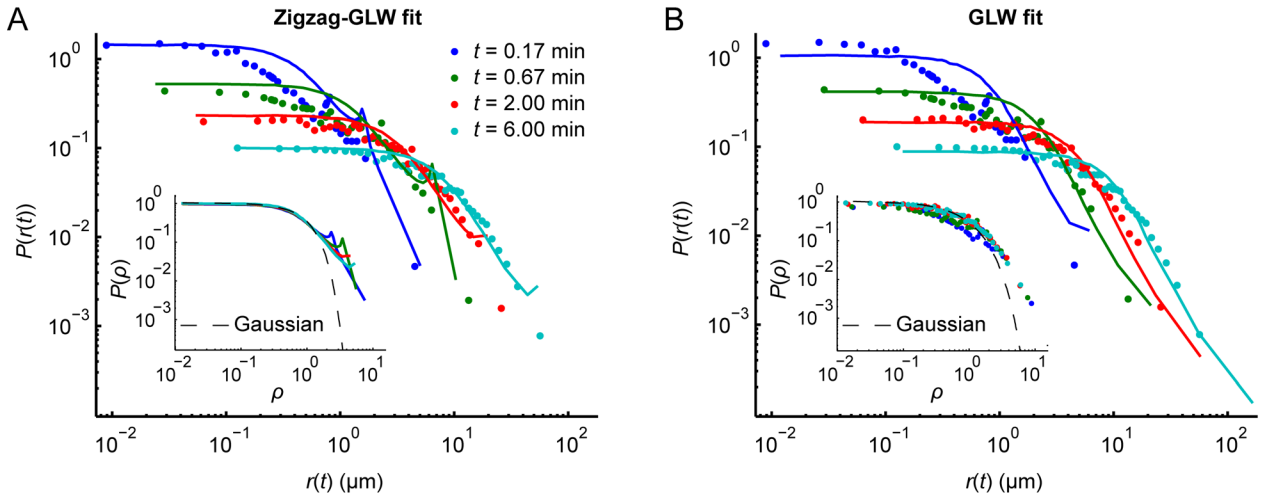


Figure S6. Displacement probability density distribution $P(r(t))$ at different time intervals t of leukocytes compared with the corresponding (A) zigzag generalized Lévy walk (Zigzag-GLW) fitting and (B) generalized Lévy walk (GLW) fitting. Colored dots indicate the experimental data. Solid lines indicate the model fittings. The insets show $P(r(t))$ at different t after normalizing the displacement $r(t)$ by the scale factor $\zeta(t)$ ($\rho = r(t) / \zeta(t)$). Normalized $P(r(t))$ in the insets of A and B were obtained from the Zigzag-GLW fitting and the experimental data, respectively. The Gaussian fittings of the normalized $P(r(t))$ at $t = 5$ min are indicated by dashed lines for comparison.

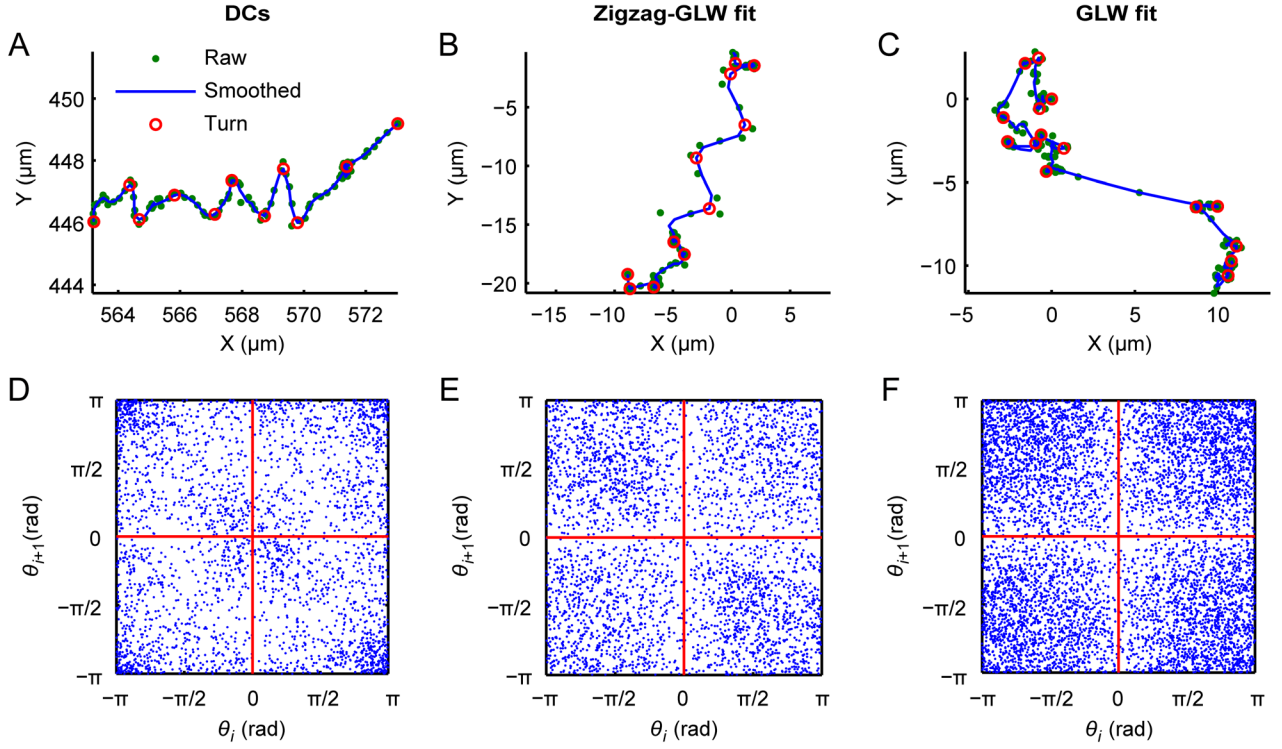


Figure S7. Turning characteristics of dendritic cells (DCs) compared with the corresponding zigzag generalized Lévy walk (Zigzag-GLW) fitting and generalized Lévy walk (GLW) fitting. (A–C) Typical migration trajectories of DCs (A) and the corresponding Zigzag-GLW fitting (B) and GLW-fitting (C) (green dots: raw trajectories (Raw); blue solid lines: smoothed trajectories (Smoothed); red circles: recognized turns (Turn)). (D–F) Return maps of turn angles θ of DCs (D) and the corresponding Zigzag-GLW fitting (E) and GLW-fitting (F). i denotes the sequence number of θ within each trajectory.

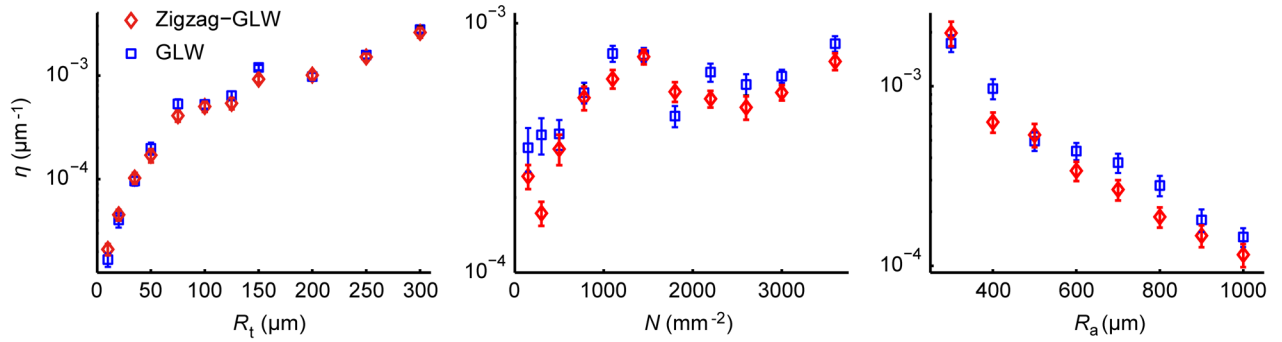


Figure S8. Shortest-capture-time-decided search efficiency η as functions of the target detectable radius R_t , the immunocyte density N , and the radius of the search area R_a of the zigzag generalized Lévy walk (Zigzag-GLW) compared with the generalized Lévy walk (GLW). The η values of the Zigzag-GLW and the GLW are similar. The error bars denote the SEM.

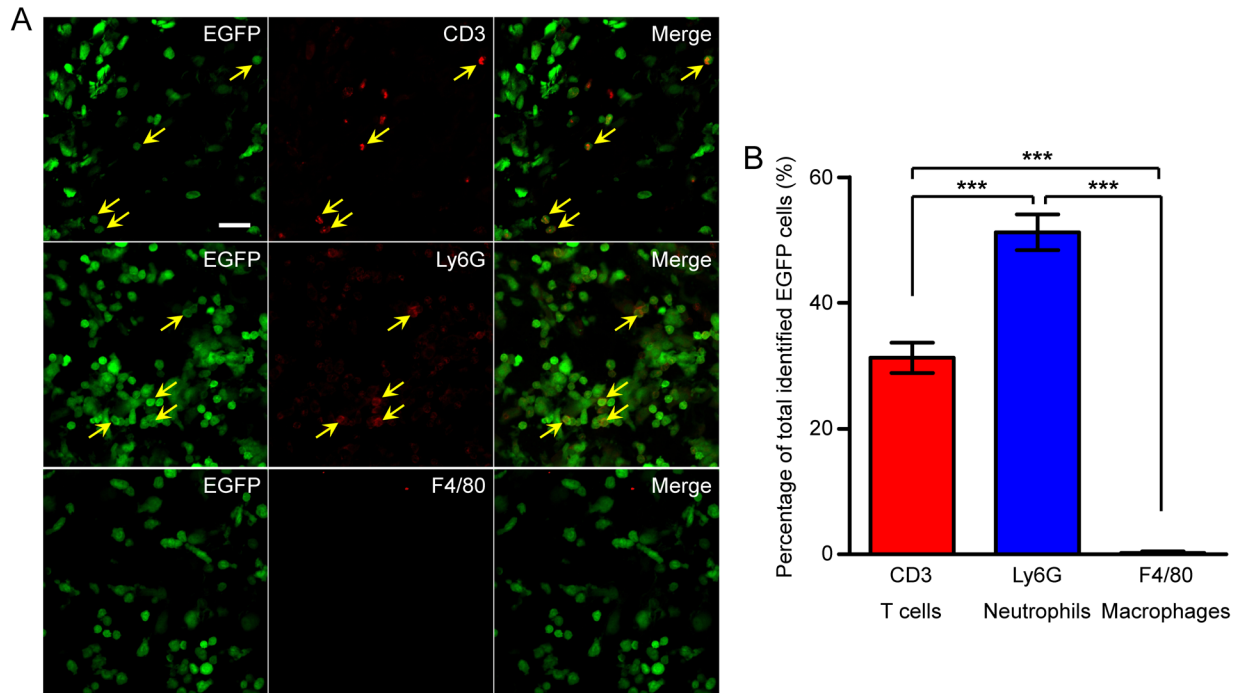


Figure S9. Immunofluorescence histological analysis revealing the cell types of the investigated leukocytes, which were identified from the EGFP cells in B6-EGFP mice, according to the size and shape of leukocytes described in previous studies [12, 34, 35]. (A) Skin cryosections from B6-EGFP mice were immunostained with CD3 (CD3-Alexa Fluor 594 (17A2)), Ly6G (Ly6G-Alexa Fluor 700 (1A8)), and F4/80 (F4/80-Alexa Fluor 647 (BM8)) to label T cells, neutrophils, and macrophages, respectively. The first two columns show images obtained from the EGFP channel and the immunofluorescence channel, respectively. The third column shows merged images of the first two columns. Yellow arrows indicate representative immunostained EGFP cells. Scale bar: 30 μ m. (B) The percentage of each immunostained cell type of the total identified EGFP cells. ***: $P < 0.001$, one-way ANOVA and Tukey's multiple comparison test. The error bars denote the SEM. The data were obtained from more than three independent experiments with 3–6 samples per group for each experiment.

Supplementary Methods and Discussion

Parameter estimation of the GLW model

In the GLW model, at a fixed velocity v , the walker runs along a straight path of random direction over a distance l_{run} , and then pauses for a time t_{pause} before executing the next run, after which the process is repeated. l_{run} and t_{pause} are drawn randomly from Lévy distributions with Lévy exponents μ_{run} and μ_{pause} , respectively [23]. Thus, to fit the GLW model to the experimental data, three parameters μ_{run} , μ_{pause} and v have to be estimated.

Here, a scaled displacement probability density distribution $P(\lambda(t))$, which is independent of the running velocity v , was constructed to separately estimate Lévy exponents and v . In detail, $P(\lambda(t))$ was constructed with the scaled displacement:

$$\lambda(t) = \frac{r(t)}{\zeta(\Delta t)}. \quad (1)$$

According to simulation studies (Fig. S3), we found that in the scale factor and the mean squared displacement of $r(t)$:

$$\zeta(t) = \langle r(t) \rangle = Ct^\gamma, \quad (2)$$

$$\text{m.s.d.} = \langle r^2(t) \rangle = At^\alpha, \quad (3)$$

$C = a_1 v$, $A = a_2 v^2$, a_1 and a_2 are constants for each specific Lévy exponent pair (Fig. S3A, C, E, G), and γ and α are independent of v (Fig. S3B, D, F, H). Therefore, we can easily conclude that the scale factor and the mean squared displacement of $\lambda(t)$:

$$\zeta_\lambda(t) = \langle \lambda(t) \rangle = \left(\frac{t}{\Delta t} \right)^\gamma, \quad (4)$$

$$\text{m.s.d.}_\lambda = \langle \lambda^2(t) \rangle = \frac{a_2}{a_1 \Delta t^{2\gamma}} t^\alpha, \quad (5)$$

are all independent of v . Furthermore, $P(\lambda(t))$ was also proven to be independent of v by simulation studies (Fig. S4). Thus, based on a series of GLW $P(\lambda(t))$ constructed using simulated data with different Lévy exponent pairs, the most likely μ_{run} and μ_{pause} were estimated independently of v by the least square estimation (LSE):

$$S_1 = \sum_t (P_{\text{exp}}(\lambda(t)) - P(\lambda(t, \mu_{\text{run}}, \mu_{\text{pause}}, v_{\text{max}})))^2, \quad (6)$$

$$\{\hat{\mu}_{\text{run}}, \hat{\mu}_{\text{pause}}\} = \arg \min_{1 < \mu \leq 3} S_1(\mu_{\text{run}}, \mu_{\text{pause}}), \quad (7)$$

where $P_{\text{exp}}(\lambda(t))$ is the experimental data and $P(\lambda(t))$ is the simulated data. In the simulation, we set v to the maximum instantaneous velocity of the experimental data v_{max} , because according to the model definition, v is actually the maximum instantaneous velocity of GLW walkers. To ensure that the statistics are good enough to obtain reliable displacement distributions, we set the simulation time unit $t_0 = \Delta t / 200$ and the simulation distance unit $l_0 = v t_0$. The scale factors of the Lévy distributions were set to one (in simulation units, l_0 or t_0).

After μ_{run} and μ_{pause} were determined, we further estimated v based on the displacement probability density distribution $P(r(t))$ in the same spirit:

$$S_2 = \sum_t (P_{\text{exp}}(r(t)) - P(r(t), \hat{\mu}_{\text{run}}, \hat{\mu}_{\text{pause}}, v))^2, \quad (8)$$

$$\hat{v} = \arg \min_{v \in V_{\text{range}}} S_2(v), \quad (9)$$

where $P_{\text{exp}}(r(t))$ is the experimental data and $P(r(t))$ is the simulated data. Before estimation, the value range of v was greatly narrowed down to V_{range} by comparing the scale factor $\zeta(t)$ of the GLW model to the experimental data.

$$flag_i = \left\langle \frac{\zeta(t, v_i) - \zeta_{\text{exp}}(t)}{|\zeta(t, v_i) - \zeta_{\text{exp}}(t)|} \right\rangle \quad (10)$$

was involved to determine if $\zeta(t)$ of a GLW model was close enough to the experimental data, where $\zeta_{\text{exp}}(t)$ is the experimental data, $\zeta(t)$ is the simulated data, $i \geq 1$, and

$$v_i = 2^{-(i-1)flag_i} v_{\text{max}}. \quad (11)$$

The value range of $flag_i$ is $[-1, 1]$. $flag_i = 1$ means $\zeta(t) > \zeta_{\text{exp}}(t)$ for all of the t values. Thus, v_{i+1} would be set to $1/2 v_i$ to further approach to the experimental data. $flag_i = -1$ means $\zeta(t) < \zeta_{\text{exp}}(t)$ for all of the t values and v_{i+1} would be set to $2 v_i$. And then, $flag_{i+1}$ was sequentially calculated. In such way, the value of $flag_i$ was continually updated until one of the flowing cases appears:

$$V_{\text{range}} = \begin{cases} (\frac{v_i}{2}, 2v_i), -1 < flag_i < 1 \\ (v_{i-1}, v_i), flag_i - flag_{i-1} = 2 \\ (v_i, v_{i-1}), flag_i - flag_{i-1} = -2 \end{cases}. \quad (12)$$

Then, V_{range} was determined.

Evaluation of the GLW parameter estimation method

To validate the accuracy of the proposed GLW parameter estimation method, we fitted simulated cells

performing GLW with $\mu_{\text{run}} = 2.2$, $\mu_{\text{pause}} = 1.7$ and $v = 50 \text{ } \mu\text{m min}^{-1}$ to the GLW model using the proposed parameter estimation method. The estimated values $\mu_{\text{run}} = 2.2$, $\mu_{\text{pause}} = 1.7$ and $v = 49.25 \text{ } \mu\text{m min}^{-1}$ are nearly equal to the corresponding true values. The displacement statistics including $P(r(t))$, $\zeta(t)$, m.s.d. and $K(\tau, t)$ of the simulated data are consistent with those resulting from the GLW fitting well (Fig. S5). Moreover, as expected, the simulated data show significantly different features from the Brown walk: the tail of the $P(r(t))$ was heavier than that of the Gaussian distribution (the inset of Fig. S5A); $K(\tau, t)$ decayed slowly than exponentially (the inset of Fig. S5B); for $\zeta(t) \sim t^\gamma$, $\gamma \approx 0.59$, not 0.5 (Fig. S5B); and $\alpha \approx 1.56$, not 1 for m.s.d. $\sim t^\alpha$ (Fig. S5C).

In addition, when constructing GLW $P(\lambda(t))$ with different v values to estimate the Lévy exponents (Equations S6 and S7), the estimated values remained constant. This result is in line with the fact that $P(\lambda(t))$ is independent of v (Fig. S4) and validates that the estimation of Lévy exponents is independent of v . The simulated data analyzed here are 600 simulated cells with trajectories covering a time length of 12.5 min. The data size is comparable to that of the experimental data of DCs and WBCs. Therefore, while validating the accuracy of the GLW parameter estimation method, these results also demonstrate that the data size of the experimental data of DCs and WBCs is sufficient to obtain accurate model-fitting results when using the proposed method.

Reduction of the computation quantity during Zigzag-GLW fitting

GLW parameter estimation is the basis of fitting the Zigzag-GLW model to the experimental data. The commonly used estimation methods involve testing specific functional forms of $P(r(t))$ and determining which function is most likely to describe the experimental data. Because the analytical form of the GLW $P(r(t))$ is unknown, numerically constructed $P(r(t))$ is used. To estimate GLW parameters μ_{run} , μ_{pause} and v , a number of walkers performing the GLW have to be simulated with various combinations of the three parameters to construct different $P(r(t))$, thus leading to a costly computation. Moreover, in the previously used maximum likelihood estimation (MLE), the $P(r(t))$ at each experimental displacement has to be calculated by interpolating values between neighboring histogram points of the GLW $P(r(t))$ to calculate a likelihood. This approach leads to the dependence of the estimated results on the construction of $P(r(t))$. $P(r(t))$ is constructed with a constant number of data points per bin. To minimize such dependence and find the most likely parameters, $P(r(t))$ constructed with different combinations of the number of bins and the number of data points per bin have to be traversed to fit the data [23]. As a result, the computation quantity of the MLE is significantly costly.

To avoid the costly computation during parameter estimation, we took three measures. First, the $P(\lambda(t))$, which was independent of ν , was constructed to separately estimate Lévy exponents and ν . Second, before estimation, the value range of ν was greatly narrowed down by comparing $\zeta(t)$ of the GLW model to that of the experimental data. Third, the LSE was used rather than the MLE. Unlike MLE, the LSE only needs to calculate the value of GLW $P(r(t))$ corresponding to each histogram point of the experimental $P(r(t))$ to calculate a sum of squared residuals instead of the likelihood. Thus, the number of interpolated values between neighboring histogram points of the GLW $P(r(t))$ is reduced from the number of experimental displacements to the number of bins of the experimental $P(r(t))$. As a result, the error of interpolating greatly decreases, and we only need to construct the GLW $P(r(t))$ and the experimental $P(r(t))$ with the same number of bins, rather than traverse GLW $P(r(t))$ constructed with different combinations of the number of bins and the number of data points per bin. Since the computation quantity was reduced, we obtained a faster estimation of GLW parameters and in turn, a faster Zigzag-GLW fitting.

The search-and-capture model

To investigate the search efficiency of immunocytes, we simulated immunocytes' searching processes using a search-and-capture model [23]. In the model, immunocytes with a density N are placed in a search area of radius R_a to search for a target of a detectable radius R_t at the center (Fig. 6A). In the present study, we focused on DCs and WBCs consisting of mostly T cells and neutrophils in skin tissues that we didn't specifically involve any external stimulation into. In this case, the targets of DCs and T cells are antigens [6, 44–46] and cells bearing cognate antigens [6, 7], respectively, whereas the targets of neutrophils are tissue debris and pathogens [7–9]. Moreover, the targets are all rare. Thus, R_t of the targets of DCs and WBCs were estimated as 5–150 μm , since detection might occur upon direct contact or contact within a short distance. Moreover, R_a ranging from 300 μm to 1200 μm , which corresponded to a target density of 0.69 mm^{-2} to 11 mm^{-2} , were involved in the search-efficiency investigation. N were measured from intravital optical imaging results. For DCs, it ranged from 100 mm^{-2} to 550 mm^{-2} with a mean value of 180 mm^{-2} , and for WBCs, it ranged from 290 mm^{-2} to 1820 mm^{-2} with a mean value of 780 mm^{-2} . The standard setting was $R_t = 100 \mu\text{m}$, $R_a = 600 \mu\text{m}$ and N corresponded to the measured mean cell density.

Supplementary Movie Legends

Movie S1. Dendritic cells migrating *in vivo*. Red solid lines indicate cell trajectories. Gray balls mark cell centroids. The time is shown at the upper left corner as min:s.

Movie S2. Leukocytes migrating *in vivo*. Red solid lines indicate cell trajectories. Gray balls mark cell centroids. The time is shown at the upper left corner as min:s.

Movie S3. A typical migration trajectory (red) of dendritic cells. The centroid of the cell is marked by a gray ball. The time is shown at the upper right corner as min:s.

Movie S4. A typical migration trajectory (red) of leukocytes. The centroid of the cell is marked by a gray ball. The time is shown at the upper right corner as min:s.

Movie S5. Comparison of the typical search process of the zigzag generalized Lévy walk (Zigzag-GLW) and the generalized Lévy walk (GLW) to locate a target. By performing the Zigzag-GLW (red) or the GLW (blue), simulated immunocytes (green) were searching for a target (magenta) with a certain detectable radius (yellow boundary). The time is shown at the upper right corner as h:min:s.