Emergence of embryonic pattern through contact inhibition of locomotion

John R. Davis¹, Chieh-Yin Huang¹, Jennifer Zanet¹, Sam Harrison², Edward Rosten², Susan Cox¹, Daniel Y. Soong¹,², Graham A. Dunn¹,* and Brian M. Stramer¹,*

SUMMARY

The pioneering cell biologist Michael Abercrombie first described the process of contact inhibition of locomotion more than 50 years ago when migrating fibroblasts were observed to rapidly change direction and migrate away upon collision. Since then, we have gleaned little understanding of how contact inhibition is regulated and only lately observed its occurrence in vivo. We recently revealed that Drosophila macrophages (haemocytes) require contact inhibition for their uniform embryonic dispersal. Here, to investigate the role that contact inhibition plays in the patterning of haemocyte movements, we have mathematically analysed and simulated their contact repulsion dynamics. Our data reveal that the final pattern of haemocyte distribution, and the details and timing of its formation, can be explained by contact inhibition dynamics within the geometry of the Drosophila embryo. This has implications for morphogenesis in general as it suggests that patterns can emerge, irrespective of external cues, when cells interact through simple rules of contact repulsion.

KEY WORDS: Contact inhibition of locomotion, Drosophila, Hemocytes, Cell migration

INTRODUCTION

Drosophila haemocytes are highly migratory cells that are readily amenable to analysis of their developmental movements in vivo (Wood et al., 2006; Stramer et al., 2008; Siekhaus et al., 2010; Stramer et al., 2010). They are initially derived at stage 10-11 of Drosophila development and subsequently disperse evenly throughout the embryo, taking stereotypical migratory routes. One of these routes is along the ventral surface of the embryo, where a three-line cellular pattern is formed (Fig. 1A). Previously, we revealed that this uniform dispersal is driven, at least in part, by contact inhibition (Stramer et al., 2010). However, we still do not know the role that contact inhibition might be playing in the emergence of this pattern, and whether an external cue is required. We have therefore set out to model this process mathematically to determine whether contact inhibition dynamics can fully explain the pattern generation.

MATERIALS AND METHODS

Imaging

Srp-Gal4 was recombined with UAS-redStinger and UAS-ClipGFP to label the nucleus and microtubules, respectively, and mounted as previously described (Wood et al., 2006). Time-lapse images were collected on a Leica SP5 or a PerkinElmer UltraVIEW spinning disk microscope and cells tracked with Velocity (PerkinElmer) or Imaris (Bitplane) software. For Cyclin A overexpression, Srp-Gal4, UAS-ClipGFP, UAS-redStinger was crossed with UAS-CycA (Bloomingston Stock Center). To examine Collagen IV deposition, the GFP enhancer trap line Viking-GFP (Morin et al., 2001) was imaged on the ventral surface at stages 14 and 16.

Domain analysis

For domain analysis, UAS-GFPmoesin (Dutta et al., 2002) or UAS-LifeActGFP (Zanet et al., 2012) was driven in haemocytes to label actin, and cells were visualised for 2 hours at 1-minute intervals. The time-lapse series was then thresholded within ImageJ (NIH) and flattened with an average intensity projection. A contour plot was generated using the ImageJ contour plot plug-in. For analysis of moving domains, a 2-hour movie was generated and a map created in a 40-minute walking average using the walking average ImageJ plug-in.

Correlating migration to segments

To quantify the percentage of tracks that crossed segment boundaries, ~2-hour time-lapse movies were acquired of nuclei-labelled haemocytes along with brightfield images, which allowed us to visualise segments. Real segment maps were overlaid onto simulated tracks to approximate segment crossing in computer simulations. The simulation was run for 60 minutes, and in reality tracks of 50-70 minutes were used for subsequent analysis.

Simulation

For details of the model, see supplementary material Appendix S1 and Fig. S4.

Tracking and segmentation methods

To track both the nuclear movement and the cell body, segmentation of nuclear and tubulin data was performed using a graph theoretic algorithm (Boykov and Jolly, 2001). The algorithm was initialised with a user-selected threshold to identify areas of interest. To compute the solution, the max-flow was found, using a variant of Dinic’s algorithm, before the solution was transformed into the min-cut boundaries. The centres of the extracted boundaries were calculated by finding the largest inscribing circle that could fit within the boundary. This method was used because simpler methods, such as calculating centroids, could be altered by the extension and retraction of structures such as lamellipodia.

RESULTS AND DISCUSSION

Automatic tracking of colliding and non-colliding haemocytes during their developmental dispersal

We have focused on the movement of haemocytes along the ventral surface of the Drosophila embryo as they spread out within an acellular space beneath the overlying epithelium. This is a thin and
approximately flat space, which has allowed us to model haemocyte movement in two dimensions (Stramer et al., 2010). To enhance the analysis, we have generated fly lines expressing a red nuclear marker along with GFP-labelled microtubules. This allowed us to automatically track haemocytes using the nucleus as a fiduciary marker (supplementary material Movie 1), while at the same time using microtubule alignment as a visual marker of a collision event as previously described (Fig. 1B; supplementary material Movie 2) (Stramer et al., 2010). We also compared this nuclear tracking method with tracking the cell centroid position. Comparison of these two methods revealed that nuclear and cell body tracks were nearly identical (supplementary material Fig. S1A) even after directional changes in response to collision events (supplementary material Movie 3), showing that the haemocyte nucleus does not move independently during changes in cell polarity as has been reported in other cell types in vitro (Gomes et al., 2005). Analysis of acceleration vectors during contact inhibition revealed a significant negative acceleration coincident with the time of collision, justifying our use of microtubule-alignment as a visual marker of a collision (Fig. 1C, D).

Haemocytes initially migrate down the ventral midline above the nerve cord. Subsequently, the cells spread laterally to adopt a three-line pattern along the ventral surface (Fig. 1A), which has been hypothesised to require an external cue (Wood et al., 2006). The lateral boundary for haemocytes within this ventral space is a dynamically changing border created by other haemocytes migrating within an underlying plane (supplementary material Movie 4). Cells at the boundary occasionally change planes; however, this only occurs rarely, when they can do so without colliding with another cell. For the remainder of development, haemocytes continue to migrate within this semi-enclosed two-dimensional space and yet the population maintains a uniform distribution and a three-line pattern (supplementary material Movie 5).

Mathematical modelling of non-colliding haemocyte motility

We first sampled the behaviour of freely moving haemocytes. Non-colliding haemocytes were tracked at 1-minute intervals for a total of 6 minutes. This time span was chosen to maximise the sample size, as cells are highly constrained and do not run freely for long before colliding. The speed (in actuality the mean magnitude of the 1-minute displacement) of these freely moving cells was 2.1±1.4 μm min⁻¹ (mean ± s.d.), which was slightly faster than the overall mean speed of haemocytes in the embryo (1.8±0.05 μm min⁻¹), suggesting that motility is decreased by frequent collisions. We extracted 50 pairs of consecutive 1-minute velocity vectors from the sample for fitting the free motility model.

The free motility model assumes that freely moving cells undergo a form of persistent two-dimensional random walk based on a simple recursive equation:

\[ \mathbf{v}_t = \phi \mathbf{v}_{t-1} + \sigma \mathbf{n}_t, \]

where \( \mathbf{v} \) are vector velocities, \( \phi \) is a scalar persistence factor, \( \sigma \) is a scalar standard deviation (s.d.) multiplier, and \( \mathbf{n} \) is a two-dimensional random noise vector with Gaussian distribution, mean \( = 0 \) and s.d. \( = 1 \) on both axes.

Fig. 2A, B explains how the two parameters of the model, \( \phi \) (persistence) and \( \sigma \) (noise), were estimated from the sample data of velocity pairs \( \mathbf{v}_{t-1} \) and \( \mathbf{v}_t \). A justification for adopting this simple
model was that \( \sigma \) was not very different when measured along or across the direction of \( \mathbf{v}_{-1} \). Using these estimates, we ran the model as a simulation with an initial velocity of zero and found that the simulated cell speed reached a constant mean value (i.e. the time series became stationary) after a few minutes (supplementary material Fig. S1). We found that this mean speed and its variation as a simulation with an initial velocity of zero and found that the model was that \( \sigma \) was not very different when measured along or across the direction of \( \mathbf{v}_{-1} \). Using these estimates, we ran the model as a simulation with an initial velocity of zero and found that the simulated cell speed reached a constant mean value (i.e. the time series became stationary) after a few minutes (supplementary material Fig. S1C). We obtained an estimate for \( \sigma \), which enabled us to estimate \( \psi \) as \( \psi = \frac{\text{mean}(\mathbf{v}_{-1})}{\text{mean}(\mathbf{n}_{-1})} \). The estimated value of \( \psi \) (1.25) was next used to calculate new vectors, \( \mathbf{v}_{-1} = \psi \mathbf{n}_{-1} \), which have the influence of the direction of the target cell removed. We used these modified vectors to estimate \( \phi \) and \( \sigma \) by the methods used for the free motility model. We obtained an estimate for \( \phi \) of 0.51, and the two estimates for \( \sigma \) (1.73 and 1.24) were again not significantly different at the 1% level [two-sided Fisher’s variance ratio test: variance ratio=1.94; \( \text{df}(49,49), 0.05 > P > 0.01 \)].

**Mathematical modelling of haemocytes undergoing contact inhibition**

In contrast to free migrations, collisions are discrete events and cannot be modelled by a continuous stochastic process, so instead of the time variable \( t \) used in the free model we use a fixed \( t \) to denote the time of collision. Data were gathered by tracking 27 pairs of colliding cells for 6 minutes (t=2 to t+4), making sure that the two cells collided only with each other during the time period; for comparison, freely running cells were tracked for an identical time. The average distance between the two cell positions at time \( t \) was remarkably constant at \( 30.0 \pm 5.2 \, \mu\text{m} \). The mean speed of these 54 cells over the 4-minute period after collision (2.03 \pm 1.4 \, \mu\text{m} \, \text{min}^{-1}) did not differ substantially from that of their freely moving counterparts. Nevertheless, we found a slight decrease in speed at the time of collision when tracks were analysed with better temporal resolution (15-second intervals) (supplementary material Fig. S1D). In comparison to these singly colliding cells, the overall mean speed of haemocytes in the embryo at around the same time period was lower (1.8 \, \mu\text{m} \, \text{min}^{-1}). The likely reason is that the majority of haemocytes in the embryo are undergoing more collisions than our sample of colliding cells, which was selected to have only one collision during a 6-minute period.

To examine the influence of the colliding cell in single collisions, we measured the acceleration of each cell away from its target neighbour (Fig. 1C,D) (Dunn and Paddock, 1982; Paddock and Dunn, 1986). This acceleration only proved to be significant at time \( t \), so the rest of the data were discarded and only the 1-minute interval surrounding the collision was used for constructing the model.

For the collision model, again we adopted the simplest likely model:

\[
\mathbf{v}_t = \psi \mathbf{r}_t + \phi \mathbf{v}_{t-1} + \sigma \mathbf{n}_t ,
\]

where \( t \) is now the time of the onset of collision and \( \mathbf{v}_t \) is the velocity during the first time interval after collision. The new vector \( \mathbf{r}_t \) is a unit vector that points from the position of the colliding partner at time \( t \) towards that of the cell being modelled. We needed an extra parameter, \( \psi \), to describe the strength of influence of the colliding
partner. The rest of the model is the same as the free motility model, except that $\phi$ and $\sigma$ can take on new estimated values.

Fig. 2C,D explains how the three parameters of the collision model were estimated from the sample data of velocity pairs. The new estimate of $\phi$ (persistence) was reduced by 30% compared with the free model, indicating that the cell tends to ‘forget’ its initial direction of travel after a collision. The estimate of $\psi$ (collision sensitivity) was 1.25, comparable with the mean velocity, indicating that the position of the other cell is a significant factor. The estimate of $\sigma$ was not very different from that of freely moving cells.

Colliding and non-colliding haemocyte behaviour predicts the dispersal pattern

We next used the two models for simulating the behaviour of cells within a two-dimensional space that mimics the shape and dimensions of the ventral surface of the embryo. A trial simulation with cells starting from random positions gave a mean speed of $1.96\pm0.24 \mu m$ min$^{-1}$, which was close to the measured speeds for freely moving haemocytes and those from the collision sample ($2.03\pm1.4 \mu m$ min$^{-1}$). However, for the reason mentioned earlier, this speed was faster than the mean speed of haemocytes in the embryo and the emerging pattern was more chaotic than in reality.

To avoid the complexities of attempting to analyse multiple cell collisions, we modified our collision model by reducing the noise ($\sigma$) to match the average speed in the embryo. When the trial simulation was rerun this change proved to be crucial and the simulated cells quickly settled into a three-line pattern characteristic of haemocytes in the embryo (supplementary material Movie 6).

To mimic the actual haemocyte dispersal process more closely, we confined the movement of simulated cells to a region within 4 $\mu m$ of the midline for 40 minutes before allowing them to disperse (Fig. 3A; supplementary material Movie 7; see Materials and methods for more details). This led to further unexpected complexities that bore a striking resemblance to real haemocyte dispersal. The simulated cells left at similar 90° angles from the midline along ‘exit points’ in a segmented pattern in a manner that resembles living haemocytes (Fig. 3B,C) (Wood et al., 2006). Real haemocytes were observed to undergo a significant increase in speed and directionality during the lateral migration phase (Wood et al., 2006), which was not significantly different in the simulation (lateral speed of $2.4\pm0.18 \mu m$ min$^{-1}$ in reality and $2.6\pm0.37 \mu m$ min$^{-1}$ in simulations, mean $\pm$ s.d.; $P=0.05$, df=28, unpaired t-test). Subsequent to this lateral migration, both simulated and real haemocytes also showed a progressive decrease in mean speed over time as cells became more constrained by their neighbours (supplementary material Fig. S1E).

Interestingly, taking a different modelling approach, which assumes that haemocyte contact inhibition involves simple volume exclusion between cells, failed to recapitulate the details and timing of the dispersal pattern and resulted in a reduction of haemocyte patrolling behaviour (supplementary material Movie 8, Mathematica Simulation 2). This reveals that the repulsive dynamics between colliding cells plays a crucial role in their stereotyped migration.

Simulations reveal a geometrically close-packed haemocyte dispersal pattern

These data suggest that ‘three-lined’ might not be the most accurate description of the pattern, as the model predicts that haemocyte spacing is actually due to the geometry of close packing in a confined space. A consequence of this is that the mean distance between neighbouring cells along any of the three lines is the same as that between the lines. To directly test this, we developed a method for visualising the mean cell positions throughout the duration of their movement. We generated density maps of the...
frequency of cell occupancy by temporally averaging time-lapse movies of haemocytes. The pattern that emerged showed clearly defined residence domains, even though the whole of the available space was occupied by haemocytes at some time during the movie (Fig. 3D,E).

The domain maps of real and simulated cells were remarkably similar (Fig. 3D,F) (unlike the volume exclusion model, supplementary material Fig. S2), and, when cell tracks were overlaid onto the map, both simulated and real cells appeared to hop between domains (supplementary material Movie 9). Slightly more domain hopping occurred in reality as haemocytes occasionally dropped off the ventral surface laterally towards the plane below, which opened up a domain allowing cells to shift, thus creating a dynamic, yet stable, structure. When real cells jumped between domains they showed a sudden increase in speed, which was similar regardless of whether cells were moving laterally or vertically (2.4±0.18 μm min⁻¹ or 2.2±0.25 μm min⁻¹, respectively, mean ± s.d.; \( P > 0.05, \text{df} = 19, \text{unpaired t-test} \)). Therefore, the increased lateral movement of haemocytes reported previously (Wood et al., 2006) is comparable to a domain hop as cells move into free space.

In order to quantify the spacing of the haemocyte domains, a contour plot was taken of the domain map and the distance between contour peaks measured. Interestingly, the domains showed an even spacing of ~34 μm (Fig. 3G,H), which was consistent with a contact inhibition distance of ~30 μm. However, occasionally, the domain spacing was observed to break down, suggesting that acquisition of the pattern might be stochastic (Fig. 3H). To observe the reason for this aberrant domain, the time-lapse series and cell tracks were overlaid onto the domain map, which allowed us to visualise the offending cell. This analysis revealed that the aberrant domain was generated by a single cell that became accidentally stuck between the domains occupied by two other haemocytes (supplementary material Movie 10). A comparison of domain maps from multiple embryos and simulations also revealed the random nature of the final pattern (supplementary material Fig. S2).

**Examination of crosstalk with extrinsic cues**

In order to further examine how well the model recapitulates reality we measured the dispersal of cells by computing the nearest neighbour distribution (the distance between each cell and its closest neighbour). In the model, the median nearest neighbour distance was slightly greater and the variance reduced compared with real embryos (25.08±1.45 μm versus 20.42±7.84 μm; \( P < 0.001, \text{df} = 219, \text{Mann-Whitney test} \)) (supplementary material Fig. S3A). This is unsurprising because in simulations cells automatically repel whenever they are within the defined contact distance, whereas in reality haemocytes are polarised and appear to only respond upon lamellar overlap; therefore, in real embryos, haemocytes can often be found closer than their defined contact distance if, for example, lamellae of possible colliding cells missed each other or cells were positioned back-to-back (this can be observed in supplementary material Movie 10).

Despite this increase in nearest neighbour variance in reality, the domain pattern closely mimicked simulations, and at times it even appeared that the domains were slightly more structured than simulations suggested. We therefore investigated obvious external factors that might be contributing to the stability of the pattern. As haemocytes are known to be major producers of extracellular matrix in the embryo (Yasothornsrikul et al., 1997), we first examined whether cells were laying down tracks of matrix as they dispersed, which might then behave as a haptotactic guidance cue for haemocytes, thus reinforcing the domain pattern. However, examination of a fly line expressing a GFP-Collagen IV fusion protein (Viking-GFP enhancer trap) revealed no obvious domain-like pattern to the matrix, suggesting that this hypothesis was unlikely (supplementary material Fig. S3B).

We next correlated cellular tracks to embryonic segments to determine whether, analogous to neural crest somite confinement (Gammill and Roffers-Agarwal, 2010), cells were unable to cross segment boundaries. When we compared haemocyte dispersal tracks to segments we observed no obvious restriction in movement (supplementary material Fig. S3C). When we compared the percentage of tracks per embryo that crossed segments in reality to simulated tracks that crossed estimated segmental boundaries, no significant difference was observed (supplementary material Fig. S3C). Finally, we examined haemocyte domain dynamics in relation to segments. Computing a walking average domain map over the duration of the dispersal process revealed that domains were not static structures; they regularly disappeared and reappeared as space for haemocyte movement became available (supplementary material Movie 11), and also often crossed segment boundaries (supplementary material Movie 12). However, as we could not precisely correlate domains with segments and
parasegments, we could not rule out the influence of more subtle segmentation effects. For example, the segmented epithelial grooves (Larsen et al., 2003) could create topography in the haemocoel that subtly biases haemocyte motility.

As we could not directly identify extrinsic factors modulating haemocyte dispersal, we took a closer look at the domain pattern in an attempt to infer their existence. It is clear from the dispersal timecourse that there must be some initial cue(s) that brings cells down the ventral midline and we therefore examined whether there was anything anomalous about its haemocyte domain structure. In real embryos, we first compared the distance between domains residing on the midline to domain distances within lateral ‘lines’ (supplementary material Fig. S3D). Although there was no difference in the spacing of domains between midline and lateral lines, we noticed a significant difference in the variance; in real embryos, the variance in the midline domain spacing began low and increased over time, whereas the variance in lateral domains was constant (supplementary material Fig. S3E). By contrast, in simulated embryos, midline domains showed the reverse; variance began high in the midline and subsequently decreased as the pattern stabilised. Furthermore, there were subtle differences in the spacing of domains, as real embryos showed an initial increased spacing, which decreased over time until it closely matched simulations (supplementary material Fig. S3D). These data suggest that there might be some midline effects that subtly modify the pattern. The PDGF/VEGF-like ligand Pvf2 is expressed along the ventral midline in a segmental fashion, and it has been speculated that cells might specifically cluster around these points (Wood et al., 2006). If this is the case, then the increase in the variance of domain spacing over time might be the result of Pvf2 expression decreasing along the midline over the timecourse of the dispersal process (Wood et al., 2006). It is also possible that these foci of Pvf2 act as anchor points in midline domains, which we hypothesise would have a knock-on effect that modifies lateral domains and helps to stabilise the overall embryonic pattern.

Simulations predict that cell number is critical for the haemocyte dispersal pattern

Our simulations suggest that the acquisition and stability of the pattern are strongly dependent on collision parameters (see supplementary material Mathematica Simulation 1) and governed by a number of variables, such as embryo size, collision distance and cell number. However, it was a challenge to individually examine many of these variables within our system in vivo with the exception of cell number, which simulations suggested would severely affect the migratory pattern (supplementary material Movie 13). Overexpression of Cyclin A specifically in haemocytes led to a ~50% decrease in cell number [as observed in other cell types (Keck et al., 2007)] without affecting their basic migratory properties. At stage 14 of development the cells still migrated down the ventral midline, subsequently dispersed, and displayed normal contact inhibition (Fig. 4A; supplementary material Movies 14, 15). Furthermore, the collision distance, which is a critical parameter in the model, in Cyclin A-expressing cells (31±4.7 µm) was not significantly different to that of wild-type cells (P>0.9, df=26, unpaired t-test). However, lateral movement off the midline failed to show the segmental periodicity and a three-line pattern failed to form (Fig. 4A; supplementary material Movie 14). Furthermore, the migratory domain map, when cell numbers were reduced in both reality and simulations, revealed a complete failure to acquire the close-packed geometry (Fig. 4B,C; supplementary material Movie 14).

Conclusion

Our aim was to model the dispersal of haemocytes in the simplest possible terms taking into account their contact inhibitory dynamics. Considering only two kinematic regimes – freely moving and singly colliding cells – our estimated parameters recapitulate the migratory pattern of cells within the embryo with surprising accuracy. This result has broader implications for morphological development, suggesting that Michael Abercrombie’s ‘social behaviour’ of cells can be a significant driving force for embryonic pattern formation.

Acknowledgements

We thank Marc Dionne, Roberto Mayor, Tanya Shaw and Eric Theveneau for helpful discussions.

Funding

This work is supported by a Biotechnology and Biological Research Council (BBBSC) grant [BB/R020635/1], the Medical Research Council (MRC), and the British Heart Foundation Centre of Research Excellence. Deposited in PMC for release after 6 months.

Competing interests statement

The authors declare no competing financial interests.

Supplementary material

Supplementary material available online at http://dev.biologists.org/lookup/suppl/doi:10.1242/dev.082248/-/DC1

References


Fig. S1. Kinematic testing of cell tracking and model algorithms. (A) A cell was tracked for ~30 minutes using cell centroid or nuclear tracking methods to reveal that there was little difference between the magnitude or direction of velocity vectors using either method. (B) Approach to stationarity of the free motility model. Mean and standard error of the speed during the first 10 minutes of 10,000 simulations starting from rest. (C) Six successive mean (± s.e.m.) cell speeds for samples of ten non-colliding cells from the model simulations (left) and real haemocytes. The mean speeds are a good match but the real cells show slightly more variation. (D) Haemocyte collisions were observed over 15-second intervals for a total of 13 time points (four timepoints before the collision and nine after). Each individual track was smoothed over a 30-second sliding window and normalised for its average speed. As can be seen from the graph, haemocytes significantly decrease their speeds immediately following a collision and rapidly return to normal. Error bars indicate s.d. (n=6). (E) Mean haemocyte speed over time in real and simulated embryos. Cells initially have a high speed as they leave the midline and subsequently decrease their speed as they become constrained by neighbouring cells. Graphs are an average of three real and simulated embryos. Error bars represent s.e.m.
Fig. S2. Example contour plots of domain maps from real and simulated embryos. Note the subtle variation in final pattern in real and simulated embryos compared with a complete failure of pattern formation in the volume exclusion model. Scale bars: 15 μm.
Fig. S3. Examination of crosstalk with extrinsic cues. (A) Probability distributions of nearest neighbour distances of real and simulated haemocytes, versus an equal number of randomly distributed cells. The median nearest neighbour distance in reality was slightly smaller than in simulations (20.43±7.85 versus 25.08±1.45 μm, respectively; $P<0.0001$, df=168, Mann-Whitney test), but significantly greater than a random distribution (20.43±7.85 versus 12.99±7.15 μm, respectively; $P<0.0001$, df=158, Mann-Whitney test). (B) Viking-GFP (Collagen IV) enhancer trap expression in stage 14 and 16 embryos. Arrows highlight haemocytes expressing high levels of the enhancer trap. Scale bar: 10 μm. (C) (Left) Haemocyte tracks in simulated and real embryos were overlaid onto a colour-coded map defining embryonic segments. (Right) Quantification of the percentage of tracks per embryo that crossed segment boundaries revealed that there is no significant difference between real and simulated embryos (35.70±4.40% versus 28.89±5.66%, respectively; $P>0.25$, df=4, unpaired t-test). (D) (Left) Schematic embryo domain map showing how midline and lateral line domain distances were calculated. (Right) Distances between midline and lateral line domains were measured and plotted over time in real and simulated embryos. In reality, during the early stages of haemocyte dispersal, the mean distance between the domains was greater than in the model (for the midline domains at 60 minutes 35.48±3.25 μm versus 29.17±6.62 μm, df=47, and 120 minutes 35.50±5.24 μm versus 28.15±3.28 μm, df=44. For the lateral domains at 60 minutes 35.72±7.38 μm versus 28.29±3.02 μm, df=87, and 120 minutes 35.44±7.67 μm versus 27.81±4.06 μm, df=93). However, at later stages the mean distances become less significant (for the midline at 180 minutes 31.95±11.06 μm versus 28.20±3.02 μm, df=49, and the lateral lines at 180 minutes 30.07±8.40 μm versus 27.55±2.91 μm, df=94). All tests used the Welch’s correction t-test. (E) The standard deviations of the mean distances between domains within the midline and within lateral lines in real and simulated embryos were plotted over time. In real embryos, the standard deviation in midline domains was significantly less than domains in lateral lines early in haemocyte dispersal [60 minutes; two-sided Fisher’s variance ratio test: variance ratio=0.16; df=(21,35), $P<0.0001$]. By contrast, in simulated embryos in early haemocyte dispersal (60 minutes) the standard deviation in midline domains was significantly greater than domains in lateral lines [two-sided Fisher’s variance ratio test: variance ratio=4.67; df=(26,52), $P<0.0001$]. However, at later stages of haemocyte dispersal (180 minutes), the standard deviation in reality and simulated embryos was not significantly different [two-sided Fisher’s variance ratio test: variance ratio=1.37; df=(26,43), $P>0.05$ and variance ratio=1.10; df=(23,51), $P>0.05$, respectively].
Fig. S4. Graphical representation of the free-running and collision models. (A) Free-running model. (B) Collision model. See Appendix S1 for details of the simulation.
Appendix S1

Simulation

The model consists of two alternatives and one of these is applied to each cell in turn at each time step of 1 min. The free-moving model is used when there are no other cells within collision distance of the cell being modeled and the collision model is used when the nearest neighboring cell is within the collision distance. In real cells the collision distance is remarkably constant at 30 µm. In the simulation, the motile cells start off by being confined to a narrow region located at the midline of the embryo. After a predetermined time interval, they are released to wander anywhere within the ventral embryo space. The boundary of this space is modeled as a chain of closely packed cells. These are stationary and do not take part in the simulation but they can act as neighboring cells during the collision response. This arbitrary boundary condition is partially justified by the observation that cells in the real embryo appear to be largely confined to the ventral space by contact interactions with other cells at the edge of the space.

The simplest 2-D discrete random process that we considered as a model for cell tracks, sampled at unit time intervals, is based on the recursive velocity process... \( v_t = \phi v_{t-1} + \sigma n_t \) (Eq.1)... where the \( v \) are vector velocities, \( \phi \) is a scalar persistence factor, \( \sigma \) is a scalar standard deviation multiplier and \( n \) is a 2-D random noise vector with Gaussian distribution, mean = 0 and standard deviation = 1 on both axes. For \( 0 \leq \phi < 1 \) the process becomes stationary and can be used to generate the successive position vectors, \( p \), in a 2-D random walk: \( p_{t+1} = p_t + v_t \). The model is represented graphically in supplementary material Fig. S4A, where the cloud of grey dots is a sample of the possible positions of \( p_{t+1} \).

To estimate \( \phi \), we extracted all pairs of successive velocities \( (v_{t-1}, v_t) \) and rotated each pair so that \( v_{t-1} \) points along the +x axis. We then take the x-components of the rotated vectors so we
can reduce Eq.1 to the scalar equation \( v_{\xi t} = \phi v_{\xi t-1} + \sigma n_{\xi t} \). Taking the expectations of these quantities we get \( E[v_{\xi t}] = \phi E[v_{\xi t-1}] + \sigma E[n_{\xi t}] \). Since we know \( E[n_{\xi t}] = 0 \), the estimate of \( \phi \) becomes \( \hat{\phi} = \frac{\text{Mean}[v_{\xi t}]}{\text{Mean}[v_{\xi t-1}]} \). To estimate \( \sigma \), we plot the noise \( n_{\xi t} = v_{\xi t} - \phi v_{\xi t-1} \) using the rotated vectors and the estimated \( \phi \) and take standard deviations on both axes. We thus get two estimates which should not be significantly different if the model is valid:

\[
\sigma_x = \text{SD}[v_{\xi t} - \phi v_{\xi t-1}],
\]

\[
\sigma_y = \text{SD}[v_{\eta t} - \phi v_{\eta t-1}].
\]

When a collision occurs at time \( t \) we need a new term to take account of the direction of the other cell (since we will show that this is significant): \( v_t = \psi r_t + \phi v_{\xi t-1} + \sigma n_{\xi t} \) ... (Eq.2)... Here \( r_t \) is a unit vector pointing to \( p_t \) from the position of the other cell at the time of collision, \( \psi \) is a new scalar factor that determines the influence of the other cell's position. \( v_t \) is now the vector velocity during the first time interval after collision and \( \phi \) and \( \sigma \) may take on new expected values. The model represented graphically in supplementary material Fig. S4B, where the cloud of grey dots is a sample of the possible position of \( p_{t+1} \). To estimate the parameters, we extracted all pairs of successive velocities \((v_{t-1}, v_t)\) from colliding cells that had first made contact at time \( t \) and again rotated each pair so that \( v_{t-1} \) points along the +x axis.

We then take the y-components of the rotated vectors so we can reduce Eq.2 to the scalar equation \( v_{\eta t} = \psi r_{\eta t} + \phi v_{\xi t-1} + \sigma n_{\eta t} \). Taking the expectations of these quantities we get \( E[v_{\eta t}] = \psi E[r_{\eta t}] + \phi E[v_{\xi t-1}] + \sigma E[n_{\eta t}] \) which becomes \( E[v_{\eta t}] = \psi E[r_{\eta t}] \) since two terms vanish. Now the expectations of \( r_{\eta t} \) and hence of \( v_{\eta t} \) are zero if the target cells are distributed equally to the right and left of the approaching cells. However, assuming the cell behavior is left/right symmetrical, we can bias the distribution by taking mirror images of all the tracks where the target cells is, say, on the left (so \( r_{\eta t} > 0 \) in all cases). The estimate of \( \psi \) then becomes \( \hat{\psi} = \frac{\text{Mean}[v_{\eta t}]}{\text{Mean}[r_{\eta t}]} \). Considering the x-components we have \( E[v_{\xi t}] = \psi E[r_{\xi t}] + \phi E[v_{\xi t-1}] + \sigma E[n_{\xi t}] \) which becomes \( E[v_{\xi t}] = \psi E[r_{\xi t}] + \phi E[v_{\xi t-1}] \). Using the estimate of \( \psi \), the estimate of \( \phi \)
then becomes: \( \phi_{\text{estimate}} = \frac{\text{Mean}\left[ v_x_t - \psi r_x_t \right]}{\text{Mean}\left[ v_x_{t-1} \right]} \). Finally, to estimate \( \sigma \), we plot the noise

\( \sigma_{\text{t}} = v_{\text{t}} - \psi r_{\text{t}} - \phi v_{\text{t-1}} \)

using the rotated vectors and estimated \( \phi \) and \( \psi \) and take standard deviations on both axes.

Thirty cells were chosen as the typical number of hemocytes on the ventral surface, and the cells were simulated in rotation at each 1-minute interval using Mathematica©. Each cell was simulated using the free motility model unless its nearest neighbor was within 30 \( \mu \text{m} \), in which case the collision model was used. The boundary of the simulated embryo consisted of 70 equally spaced points, which would also invoke the collision model.