

# Using mathematical modelling to explore mRNA localization

Jonathan Harrison<sup>†</sup>, Richard Parton<sup>‡</sup>, Ilan Davis<sup>‡</sup>, Ruth Baker<sup>†</sup>

<sup>†</sup>Mathematical Institute, University of Oxford; <sup>‡</sup> Department of Biochemistry, University of Oxford

## Abstract

In early *Drosophila* development, mRNA localization plays a key role in determination of the body axes. Maternal mRNA is transported from the nurse cells into the oocyte itself, where it is localized to specific sites in order to target proteins to regions where they can function. We investigate the rate-limiting steps of mRNA localization and targets for regulation by the cell. By employing a simple velocity jump process model of the transport of mRNA, we describe this transport process and demonstrate how it depends on certain parameters of the model, such as the bias in the microtubule network. The model used represents the active transport and diffusion phases observed experimentally. Based on simulations from the model with inferred parameters, we estimate a mean first passage time from the nucleus of the nurse cell to the ring canal joining to the oocyte of 18 min. Since the timescales over which mRNA localization occurs during oogenesis are at least an order of magnitude longer than this, we suggest that production is the rate-limiting step in mRNA localization. Finally, we incorporate production of mRNA complexes directly into our model and are thus able to reproduce a spatial distribution of mRNA complexes similar to experimental observations. Mathematical modelling can offer mechanistic insight into this system by providing testable hypotheses and informing experimental design.

## 1 Introduction

The localization of mRNA is crucial in a variety of biological contexts for the targeting of proteins to their site of function. This method of gene expression is particularly relevant for polarized cells such as oocytes and early embryos. For example, the axes of *Drosophila Melongaster* are established through regulation of gradients in *bicoid* (*bcd*), *gurken* (*grk*), *oskar* (*osk*) and *nanos* (*nos*) mRNA (1). mRNA localization has been observed in a variety of species and cell types, including *Drosophila* and *Xenopus* oocytes, neurons, chicken fibroblasts, yeast and bacteria (2–6), thus demonstrating that this process is ubiquitous and not limited to large cells. *Drosophila* relies heavily on the asymmetric localization of mRNAs to coordinate early development processes both spatially and temporally. This makes *Drosophila* an ideal model system for the study of mRNA localization.

Recent advances in imaging techniques and image analysis technologies (7–9) have advanced our understanding of the mechanisms of mRNA localization. The use of fluorescence *in situ* hybridisation (FISH) enables single molecules of mRNA to be labelled with high levels of sensitivity and specificity. The development of the MS2-MCP system, which consists of a MS2 bacteriophage RNA stem loop bound by MS2 coat protein fusion to a fluorescent protein (10), has provided great benefits for the imaging of live cells *in vivo*. The MS2-MCP system has been used successfully in the visualisation of *nos*, *grk*, *bcd* and *osk* mRNAs

in *Drosophila* (11–14). Although there are limitations to these technologies dependent on cell type, they have permitted an improvement of our understanding of intracellular motility. What has become clear is that a variety of different mechanisms are used to ensure localization of mRNAs.

### 1.1 Motivation

After transcription and transport out of the nucleus, many mRNAs are localized and translated in the distinct cytoplasmic domains where they function (10, 15). This process is highly regulated and involves active transport by intricate molecular motors. mRNA localization is crucial for the establishment of polarity of cells and the formation of the basic animal body plan during development (1). It may have importance for a wide range of functions involving memory and learning in the nervous system. However, we are a long way from a complete understanding of the mechanisms governing the mRNA cargo transport process.

In particular, the mRNA transcripts that set up the *Drosophila* body axes originate in the nurse cells (see Figure 1) and are actively transported on microtubules into the oocyte through the ring canals (16). We will focus on this maternal transfer of transcripts into the oocyte, since this area has been less well studied and it is tractable both to modelling and data collection.

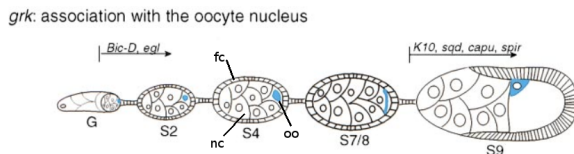


Figure 1: Stages of *Drosophila* development, showing nurse cells labelled as nc, follicle cells denoted by fc and the oocyte marked as oo. The localization of *grk* is shown in blue. Adapted from Lasko (17).

One of the challenges involved in live imaging of cells to examine this transport process is the need to scan a sample in three spatial dimensions as well as in time. This places restrictions on the resolution of the data that can be obtained temporally and in the  $z$  direction due to the limitations of current imaging technologies (18). Although there are physical restrictions on what can be achieved by improvements in imaging technologies, these could be overcome by modelling of the transport of mRNA cargoes, which may permit sparser sampling in time and finer sampling in space. By simulating from our model, assuming we have knowledge of the current state of a particle and appropriate values for model parameters, we can predict a likelihood for the state of a particle after a short time  $\Delta t$ . This could be used to inform the linkage step to identify particles between frames and may allow larger timesteps  $\Delta t$  to be taken between imaging frames, although a hidden Markov model (HMM) framework may be necessary to update the current state of the particle. An HMM approach has previously been applied to tracking in the context of bacteria (19) and for single-molecule diffusive protein data (20).

Through mathematical modelling of mRNA transport in the *Drosophila* nurse cell, we aim to identify the limiting factors in mRNA localization. This will improve our understanding of which parts of the process are important biologically for its function. It should give insight into the molecular mechanisms used and targets for regulation by the cell, through informing future experimental design. In addition, we see the potential for this modelling framework to be applied using an HMM to improve tracking of RNP complexes over longer time periods in three-dimensional data.

### 1.2 Outline of mRNA localization process

Once mRNA has been transcribed from DNA in the nucleus, the process by which it reaches its final site of localization, where it is then translated into protein, can be divided into four main stages: particle formation; nuclear export; transport; and anchoring. We will address each of these processes in turn. Understanding of these stages will enable the development of our model in Section 2.1.

mRNA does not exist in isolation inside the cell. Instead the mRNA molecules bind to proteins to form particle complexes, which are also known as ribonucleoproteins (RNPs). These proteins perform a range of different functions, including translational regulation, preventing translation while the mRNA is in transit, and may determine the final destination of the particle (21). There is also evidence that RNPs are dynamically remodelled during the transport process, to support translational regulation (22).

After the mRNA has assembled into RNPs, it must diffuse through the crowded interchromatin spaces in the nucleus until it reaches a nuclear pore complex (NPC). The RNP is then exported out of the NPC into the cytoplasm via interactions with co-factors. Certain NPCs are more active than others at different times (22), but the overall direction of export out of the nucleus is unbiased (2).

The transport stage of mRNA localization occurs by different mechanisms for different mRNAs. The most common method observed is transport via molecular motors moving along the cytoskeleton (either microtubules or actin filaments). However, *nos* mRNA localizes using a diffusion and trapping technique (11) rather than by active transport. Active transport results in fast, directed motion with velocities of the order of  $1\mu\text{ms}^{-1}$  (13, 14, 23), which is an order of magnitude faster than movement by free diffusion. Bidirectionality was observed by Vendra et al. (24) in the Dynein mediated motion of RNP complexes on microtubules in *Drosophila* blastoderm embryos. It should be noted that, conventionally, individual types of molecular motors are thought to move unidirectionally, with Dynein directed towards the minus end of microtubules and Kinesin directed towards the plus end (25). The number of motors required for a given RNP complex can vary and may be governed by cis-acting localization elements in the mRNA (26). Possible explanations for the bidirectionality include: a tug of war between different molecular motors moving in opposing directions on different microtubules; a tug of war between different motor species all bound to the same RNP complex; reversal of Dynein moving on a single microtubule, possibly due to regulation by microtubule associated proteins (MAPs) (27). The molecular motors must be correctly joined to their cargo and the motor cargo complex secured to the cytoskeleton. This is ensured in *Drosophila* oocytes by the linkers *Bicaudal D* (*BicD*) and *Egalitarian* (*Egl*) (10). The cytoskeleton was initially thought to be a stable static network, due to how it was analysed in fixed material. However, it has recently been revealed, in *Drosophila* oocytes, to be a dynamic network with a biased random orientation of microtubules (28). This underlying structure had been suggested by the work of Zimyanin et al. (14) who observed RNPs containing *osk* mRNA moving in a biased random walk.

Once the RNP complex has reached its destination, it must be maintained in position to keep the mRNA localized within the cell. In many cases an anchoring mechanism is

used; *grk* mRNA is anchored by the molecular motor Dynein (29) in the *Drosophila* oocyte and *nos* is trapped by actin (11). Dynein was observed by Delanoue and Davis (29) to act as a static motor without requiring further energy in the form of ATP to function. An alternative hypothesis is that continuous active transport is required to ensure the localization of mRNA, since Weil et al. (13) found that Dynein motor activity is required to ensure *bcd* localization in the anterior cortex of the *Drosophila* embryo. The transport process may ensure that RNP complexes are kept near their targeted region of localization by directing them back if they move away.

In this work, we present a simple model for transport of mRNA particles through nurse cells during localization and demonstrate how this model can be effectively parametrised. In Section 2, we construct the velocity jump process model and explain the approximate Bayesian computation (ABC) approach to inference of the model parameters. We exhibit the posteriors over the model parameters generated via our inference approach for both *in silico* and *in vivo* data in Section 3. Finally, in Section 4, we discuss some of the implications of our modelling and parameter inference, before exploring possible future directions.

## 2 Materials and Methods

### 2.1 Modelling approach

In this context, modelling involves breaking down the mRNA localization process into its key components and identifying the biologically important aspects. We put in place appropriate, simple modelling assumptions to describe the most important parts of the system biologically. These assumptions should as far as possible be supported by existing experimental observations, as described in Section 1.2.

Previous modelling approaches in this area have focused on descriptions of the mechanisms driving molecular motors at smaller spatial and temporal scales (30) or with continuous densities of mRNA (31) rather than representing single RNP particles which would be more appropriate at low densities. Other studies have generally considered mRNA dynamics in the oocyte in terms of cytoplasmic streaming or diffusion without direct reference to active transport (32, 33). Instead we shall describe the active transport of mRNA in the nurse cells surrounding the oocyte, which are connected to each other and to the oocyte via ring canals, using an individual-based stochastic description.

We propose two different approaches to parameterize our model. The first will involve direct measurement of parameters, such as the average speed of cargo complexes during periods of active transport. The second method requires application of an ABC framework (34–36) to parameterize the model via comparison with certain summary statistics

of the data. This will enable comparison between the experimentally measured parameter values and the posterior for these parameters obtained via ABC.

We aim to produce a testable model in the sense that we are able to make predictions from the model based on varying certain parameters, such as the speed of the motor carrying the cargo. These could be tested experimentally using genetic mutants that have behaviour that corresponds to a different value of a given parameter. This should inform our understanding of how transcripts transport and localize.

### 2.2 Velocity jump process

The movement of RNP cargoes is often described as a biased random walk (14). Rather than modelling this transport process as a simple random walk in position, we let the direction of motion vary. This type of model is known as a velocity jump process and has been successfully applied to directed migration of animals and cells (37, 38).

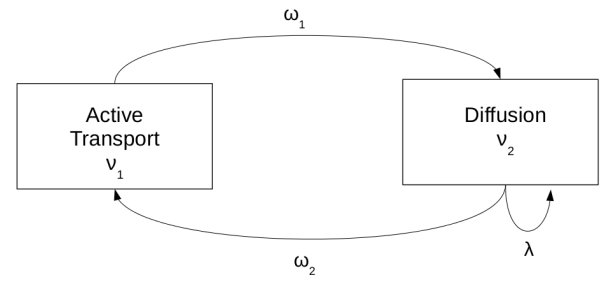


Figure 2: States and transitions between the states in the velocity jump process model. The speeds in active transport and diffusion are  $v_1$  and  $v_2$ , respectively. Transition rates between states are given by  $\omega_1$ ,  $\omega_2$  and  $\lambda$ , as shown.

We model a two-dimensional slice through a nurse cell as a rectangular region  $[0, L_x] \times [-L_y/2, L_y/2]$ , with the  $x$  axis aligned along the anterior-posterior axis of the oocyte. The cell nucleus is modelled as a circular region of given radius,  $R$ , placed centrally in the cell. This is excluding to the RNP complexes in the model which are reflected by it. The remainder of the nurse cell is accessible to the particles, which seems reasonable since experimentally particles are observed in all parts of the nurse cell (data not shown). Initially, RNP cargoes are placed randomly at time  $t = 0$  on the circumference of the nucleus to represent their export from NPCs. In the model therefore, all mRNAs are produced simultaneously rather than directly modelling continual production and export over time. Cargoes move randomly, as described below, until eventually they hit a small target at the posterior end of nurse cell representing the ring canal joining to the oocyte, where they are absorbed and removed from the system. Throughout all simulations in this work, we fix  $L_x = 52\mu\text{m}$ ,  $L_y = 37\mu\text{m}$ ,  $R = 10\mu\text{m}$  and the size

of the ring canal as  $2\mu\text{m}$ , based on measurements taken of nurse cells in the early stages of oogenesis. The position of the ring canals between the nurse cells and the oocyte can be clearly seen in Figure 3. The geometry of the model system can be seen in Figure 4 and contrasted with a real nurse cell.

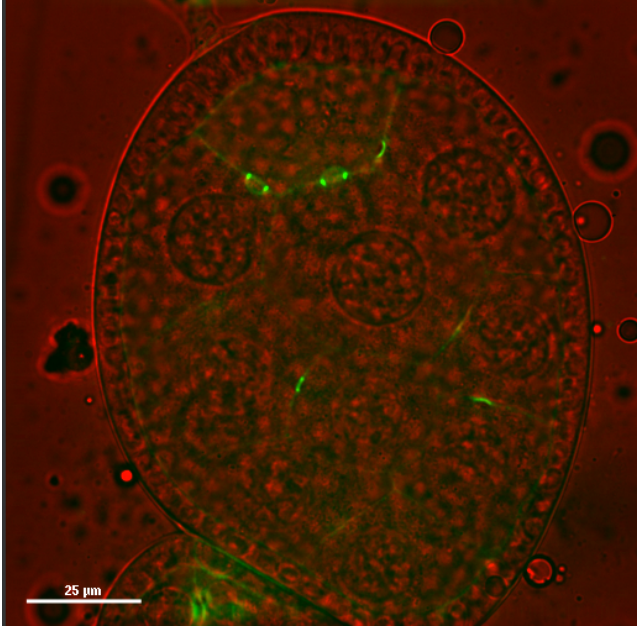
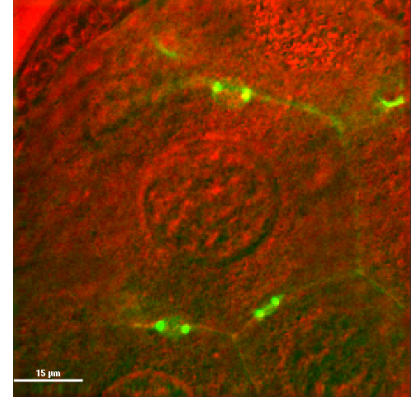
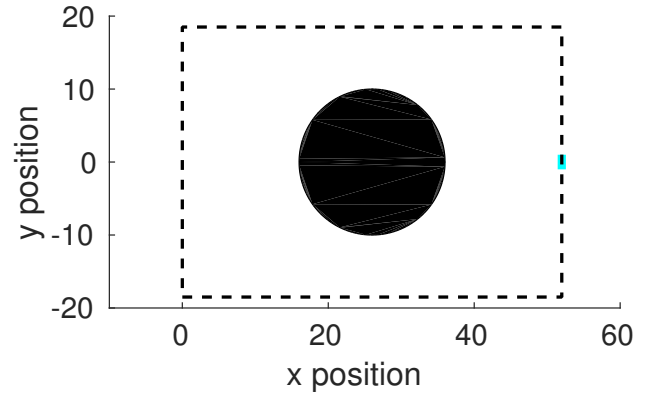


Figure 3: Position of the ring canals between the nurse cells and the oocyte for a stage 5 *Drosophila* oocyte. The ring canals are highlighted with Sqh Utrophin GFP and the red channel was obtained via differential interference contrast microscopy.

There are two phases of motion included in the model: an active transport phase when the cargo is attached to the motor and is moving on the microtubules with constant speed  $\nu_1$ ; and a slower diffusive phase with constant speed  $\nu_2$ . Switching occurs between these two phases of motion with exponential waiting times between switching events. Biologically this switching corresponds to the molecular motor falling off and reattaching to the microtubule. As illustrated in Figure 2, switching occurs from the active transport state to diffusion with rate  $\omega_1$ . From the diffusive phase, reorientations within the same phase occur with rate  $\lambda$  and switching to the active transport phase takes place at rate  $\omega_2$ . At a reorientation event, the particle remains in the diffusive phase at speed  $\nu_2$  and a new direction is picked uniformly on  $[0, 2\pi)$ . After each switching event from diffusion to the active transport state, the RNP complex also changes direction by picking a new direction of travel at random, where the angle  $\theta$  is drawn from some distribution  $T(\theta)$ . Here we choose to define  $T(\theta)$  empirically as uniform on  $[0, \pi)$  and  $[\pi, 2\pi)$  with a bias in favour of  $[0, \pi)$  based on biological data. We define a proportion of microtubules,  $\phi$ , aligned in the posterior direction (parallel to the  $x$  direction in our model) and thus take the



(a) *In vivo*.



(b) Model nurse cell domain.

Figure 4: Comparison of the geometrical structure of our model nurse cell and a nurse cell *in vivo*. In (a), Sqh Utrophin GFP was used to show the position of the ring canals in an oocyte in stage 8 of oogenesis. In (b), the dashed black rectangular box shows the boundary of the nurse cell, with the solid black circle representing the nucleus, and the thick cyan line at the posterior showing the position of the ring canal between the nurse cell and the oocyte.

following for  $T(\theta)$ :

$$T(\theta) = \begin{cases} \frac{\phi}{\pi} & \theta \in [0, \pi) \\ \frac{1-\phi}{\pi} & \theta \in [\pi, 2\pi). \end{cases}$$

### 2.3 Approximate Bayesian computation (ABC)

In a statistical inference context, the likelihood of data given a certain set of parameters is a central quantity, particularly in calculation of the posterior over the parameters given certain data. The posterior is desirable as it offers more information than just point estimates of the model parameters. Although for simple models it may be possible to evaluate the likelihood analytically, for more complex models the likelihood is often not tractable or is very expensive to compute. ABC techniques have been developed to address this issue (36). Instead of directly evaluating the likelihood, ABC

techniques assume it is possible to cheaply simulate from the model and use this to approximate the likelihood.

In an ABC rejection sampling approach we follow the procedure shown in Algorithm 1, where  $\pi(\theta)$  is a prior on the parameters,  $\rho$  is a distance metric,  $S(x)$  is a summary statistic used to summarize the data,  $x$ , and  $\epsilon$  is a maximum tolerance for acceptance.

---

**Algorithm 1** ABC Rejection Sampling
 

---

```

1: for  $i = 1$  to  $n$  {
2:   Sample parameters  $\theta$  from a prior on those parameters  $\pi(\theta)$ .
3:   Simulate data  $x$  from the model  $M(\theta)$  using those parameters.
4:   Calculate distance from observed data  $y$ :
5:   if  $\rho(S(x), S(y)) < \epsilon$  then accept the parameters  $\theta$ 
6:   else reject the parameters  $\theta$ .
7: }
  
```

---

One drawback of this algorithm is that it depends on appropriate choice of the distance metric,  $\rho$ , the summary statistic,  $S(x)$ , and the tolerance,  $\epsilon$ . Clearly the quality of the posterior will depend on the choice of these hyperparameters, as, for example, increasing  $\epsilon$  will decrease the quality of the corresponding posterior. In some settings it may be possible to take the full data, rather than a summary statistic, and to set  $\epsilon = 0$  which would give an exact sample from the posterior, but in general this is not possible computationally. We can eliminate the importance of the tolerance to some extent by simulating  $N$  samples from the prior, storing all the distances and keeping the closest  $\alpha$  quantile of the sampled parameters to the observed data. The quality of the posterior still depends on  $N$  and  $\alpha$  but the choice of  $\epsilon$ , which may be dependent on other model parameters, is removed.

The efficiency of the rejection sampling method is low, particularly for small values of  $\epsilon$  as they entail tiny acceptance rates resulting in many wasted samples. More efficient ways of sampling possible parameters  $\theta$  have been suggested including Sequential Monte Carlo and Population Monte Carlo techniques (39–42).

Inference on model parameters using ABC has been successfully applied to a variety of types of model in a biological context, often helping to offer biological insight into the model system. Cell migration models have used ABC to process data from scratch assays (34), *in vivo* data (43), and an individual-based ecological model with 14 parameters (44).

When using ABC methods, it is important to be aware also of their limitations. It has been highlighted via Robert et al. (45) that, although the error from the tolerance  $\epsilon$  can be eliminated in the limit  $\epsilon \rightarrow 0$ , the use of summary statistics of the data may have larger effects, particularly when comparing models using ABC. Some more systematic methods for choice of summary statistics have been suggested,

which use minimization of certain measures of entropy to determine which summary statistics are most informative (46).

## 2.4 Imaging

We observed *grk*-MS2 mRNA in *Drosophila* nurse cells tagged using MCP-Dendra and mCherry. Sqh Utrophin GFP was used to show the position of ring canals. Imaging was performed with a wide-field deconvolution system (Delta Vision CORE) from Applied Precision. Image stacks were deconvolved with the SoftWoRx Resolve 3D constrained iterative deconvolution algorithm (Applied Precision). Segmentation of images was performed using Felzenszwalb's algorithm (47) after denoising and deconvolution of the images, with coordinates of particles obtained from centroids of appropriately sized segments. The age of egg-chambers was determined by measuring oocyte size and comparing known sizes at different stages of development.

## 3 Results and Discussion

### 3.1 Direct parameter estimation

Based on direct measurements from manually tracked *grk*-GFP mRNA complexes in *Drosophila* nurse cells (48), we are able to make direct preliminary estimates of the parameters of our model. All data in Section 3.1 were obtained from Davidson (23). Particles are classed as static, paused or active. Paused and active particles will correspond to those in the diffusive and active transport phases of our model respectively. Static particles will be assumed to be either anchored or undergoing anomolous behaviour. The number of total particles counted was  $N = 340$  in a  $40 \times 40 \mu\text{m}$  area. The proportions of particles in each of the classes is 25% active, 50% paused and 21% static. Over the timescale observed, particles did not transition between states and these are proportions of particles. Particle speeds were assessed by observing a  $10 \times 10 \mu\text{m}$  area for 50 time-points, with images taken in a single  $z$  slice at three frames per second. The average speed in the active transport state  $\nu_1$  was  $1.163 \pm 0.08 \mu\text{ms}^{-1}$  from  $n = 33$ . The average run length was  $2.785 \pm 0.66 \mu\text{m}$ . In the paused phase, an average speed,  $\nu_2$ , of  $0.798 \pm 0.6 \mu\text{ms}^{-1}$  was observed from 58 particles with an average run length of  $0.84 \pm 0.06 \mu\text{m}$ .

Considering the process as a continuous time Markov chain, with active transport and diffusive states, we obtain that the steady state of the continuous time Markov chain is  $[\omega_2/(\omega_1 + \omega_2), \omega_1/(\omega_1 + \omega_2)]$ . Then assuming ergodicity and neglecting static particles, this approach gives us  $\omega_2/(\omega_1 + \omega_2) = 2/3$  and  $\omega_1/(\omega_1 + \omega_2) = 1/3$  by comparing to the proportions of particles observed in each state.

Assuming there are no internal transitions back to the active transport state, then we have  $\omega_1 = 0.42\text{s}^{-1}$  and hence  $\omega_2 = 0.84\text{s}^{-1}$ . We can also deduce that the rate

$\lambda = 0.95 - 0.84 = 0.11\text{s}^{-1}$ . In addition, we assume the distribution of microtubule orientations in the nurse cells is similar to that in the oocyte, since there is data available for the oocyte (28) but not for the nurse cell. Therefore we take  $\phi = 0.58$ . This gives parameter estimates as shown in Table 1.

Parameter	Value
$\nu_1$	$1.16\ \mu\text{ms}^{-1}$
$\nu_2$	$0.8\ \mu\text{ms}^{-1}$
$\omega_1$	$0.42\ \text{s}^{-1}$
$\omega_2$	$0.84\ \text{s}^{-1}$
$\lambda$	$0.11\ \text{s}^{-1}$
$\phi$	0.58
$\gamma$	$0.01\ \text{s}^{-1}$

Table 1: Parameter values for the velocity jump process model.

### 3.2 Timescales of the model

Here, we evaluate the timescales of the transport phase of the localization process. Based on simulations from the model, we obtain a mean first passage time (MFPT) from the nucleus to the posterior ring canal, averaged across 1000 particles, of  $1070 \pm 840\text{s} \approx 18\text{min}$ . We plot the evolution of a typical path in Figure 5.

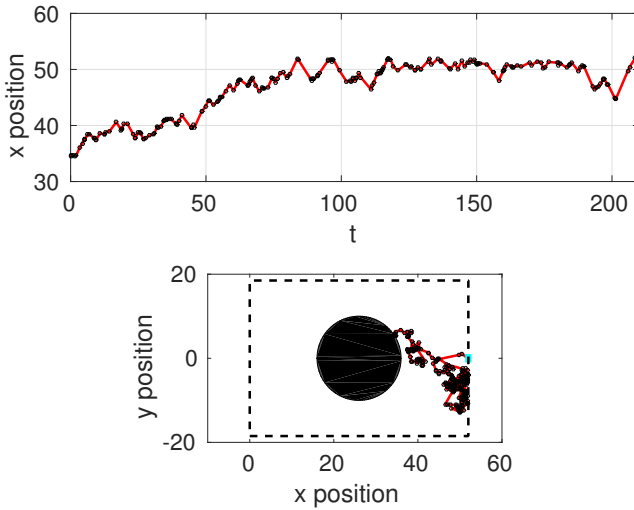


Figure 5: The evolution of a typical path until absorption at the posterior ring canal. Small black circles on the path show switching events between the active transport and diffusion states. The geometry of the domain is as shown previously in Figure 4.

Given that stage 5 in *Drosophila* oocytes lasts five hours, this timescale for the MFPT appears to be an order of magnitude shorter than the timescales over which *grk* mRNA transport is observed in oogenesis, suggesting that the limiting step in

the transport and localization of the mRNAs is their production from the nucleus rather than the transport step. We note that *grk* mRNA complexes can be observed in the nurse cells from the germarium stage up until stage 8, when production of *grk* appears to cease, as illustrated in Figure 6. The distribution of particles visible in Figure 6 also demonstrates that the difference in timescales is not due to a bottleneck elsewhere in the transport process, for instance at the ring canals, since no notable accumulation is seen there.

Furthermore, we are able to make predictions on what would happen if we halved the speed of the molecular motors via a genetic mutation. We assume that altering the speed of the molecular motor,  $\nu_1$ , would leave the speed in diffusion,  $\nu_2$ , unchanged. The resulting distribution is shown in Figure 7 and the MFPT would be  $2150 \pm 1600\text{s} \approx 36\text{min}$ .

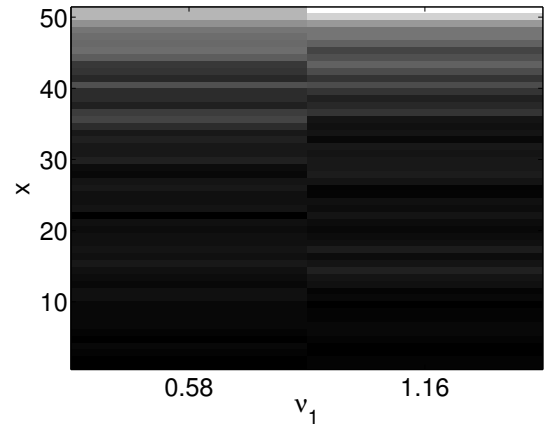


Figure 7: The spatial distribution resulting at time  $t = 270\text{s}$  from 1000 model simulations with parameters  $\phi = 0.58$ ,  $\nu_2 = 0.8\ \mu\text{ms}^{-1}$ ,  $\omega_1 = 0.42\text{s}^{-1}$ ,  $\omega_2 = 0.84\text{s}^{-1}$ ,  $\lambda = 0.11\text{s}^{-1}$ , varying  $\nu_1$  as  $0.58\ \mu\text{ms}^{-1}$  and  $1.16\ \mu\text{ms}^{-1}$ . This represents the effect of introducing a genetic mutation to alter the speed of molecular motors.

### 3.3 Sensitivity analysis

Our model is dependent on six parameters:  $\nu_1, \nu_2, \omega_1, \omega_2, \lambda, \phi$ . We are able to explore the dependence of our model on these parameters by performing a sensitivity analysis, varying each parameter over several orders of magnitude. We show in Figure 8 the effect of varying key parameters on the MFPT. The results of this sensitivity analysis demonstrate that the turning rate within the diffusive state,  $\lambda$ , has no noticeable effect on this summary statistics (Panel 8e). Decreasing the speed parameters,  $\nu_1$  and  $\nu_2$ , or the microtubule bias,  $\phi$ , leads to a sharp rise in the MFPT (Panels 8a, 8b). The transition rates  $\omega_1$  and  $\omega_2$  have opposite effects with an increase in  $\omega_1$  increasing the MFPT, and a decrease in  $\omega_2$  increasing the MFPT (Panels 8c, 8d). Similar effects are seen for two other



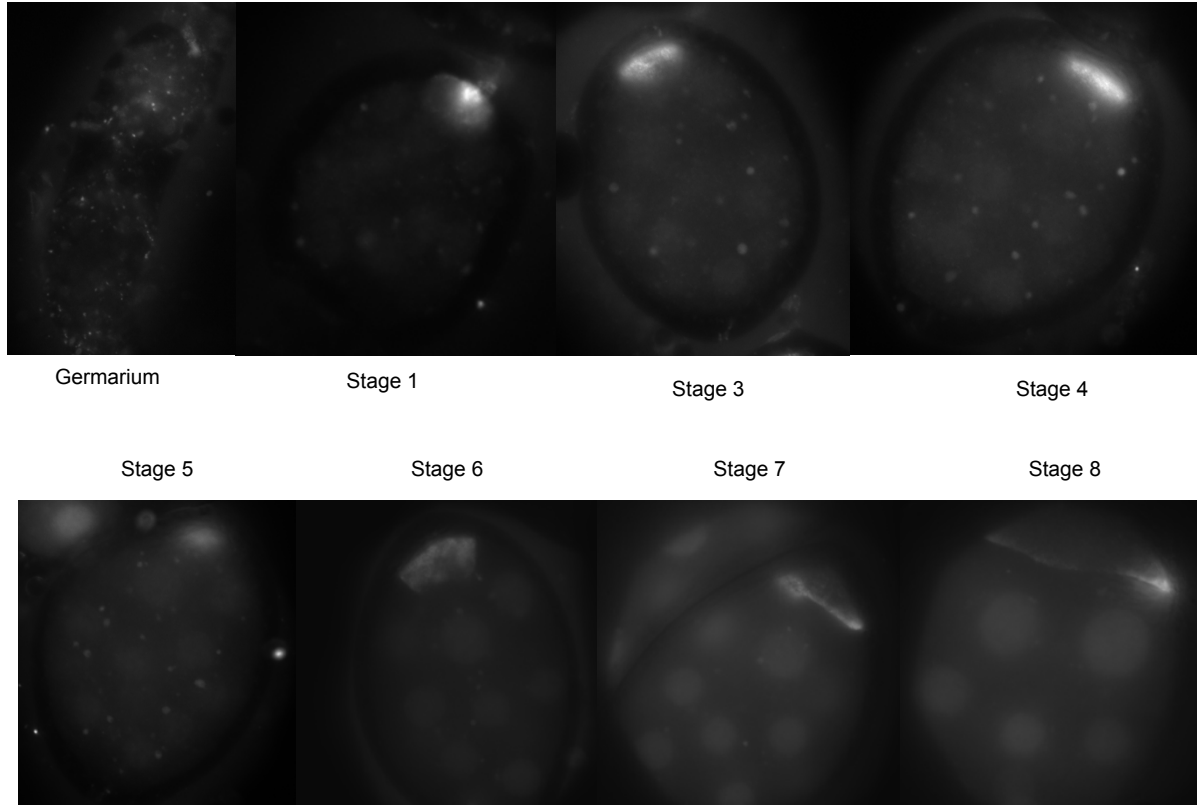


Figure 6: Localization of *grk* mRNA at different stages of *Drosophila* oogenesis. Increasing accumulation is seen in the oocyte over time. RNP particles can be seen in the nurse cell, until stage 8. The *grk* mRNA was tagged with mcherry for visualization.

summary statistics: the mean number of jumps on each path up to absorption, and the mean over different particles of the median jump distance on a single path, see Appendix A.

In addition, we exhibit the dependence of the spatial distribution parallel to the anterior-posterior axis of particles simulated from our model, as shown in Figure 9. For increased motor speeds,  $\nu_1$  and  $\nu_2$ , the particles spread out faster from the nucleus and are directed towards the posterior of the nurse cell. By increasing the microtubule bias,  $\phi$ , the motion of the cargoes becomes much more directed towards the posterior with accumulation near the posterior and absorption upon reaching the ring canal there. For high values of the switching rate,  $\omega_1$ , the distribution of particles is more spread and has not reached the posterior, as little time is spent in the active transport state. For low values of  $\omega_1$ , the particles do not change direction often in the active transport state so greater accumulation is seen for intermediate values of  $\omega_1$ . We note that, for large values of  $\omega_2$ , there is an accumulation of particles near the posterior, but that these are not quickly absorbed, since they spend little time in the diffusive phase of motion. The effect that the turning rate  $\lambda$  has on the spatial distribution appears to be unclear. Some

analytical analysis of the model can be found in Appendix B.

### 3.4 Evaluation of ABC methods for *in silico* data

We demonstrate the effectiveness of the ABC inference techniques on *in silico* data, before applying them directly to real datasets. We present first the results of applying ABC methods to our model with a broad prior and known parameter values, as shown in Figure 10.

For each experiment, we simulate a dataset from chosen parameters and evaluate its summary statistic. That summary statistic is compared to summaries of data generated from proposed parameters, as described above. We average the results obtained across a number of experiments denoted  $n_{repeats}$ . As summary statistics we take the spatial distribution of the cargoes averaged in the  $x$  direction at given times  $t_1, \dots, t_\kappa$  and use a symmetric version of the Kullback-Leibler divergence as our distance to compare distributions. Broad uniform priors are used for each parameter between

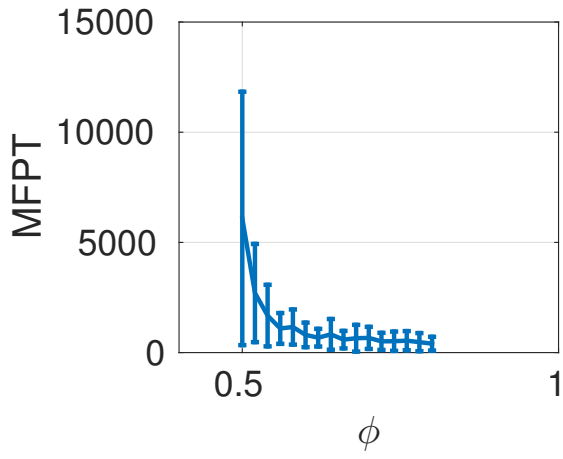
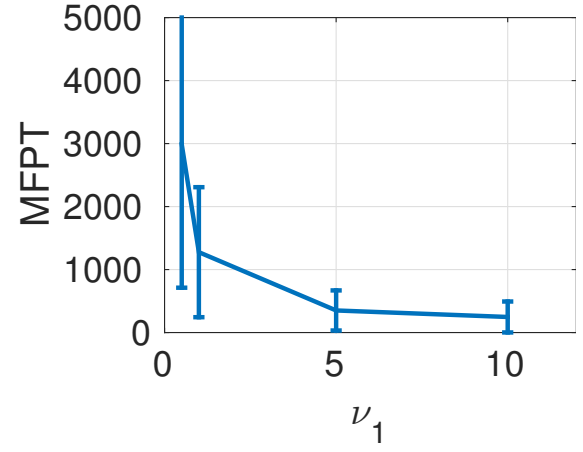
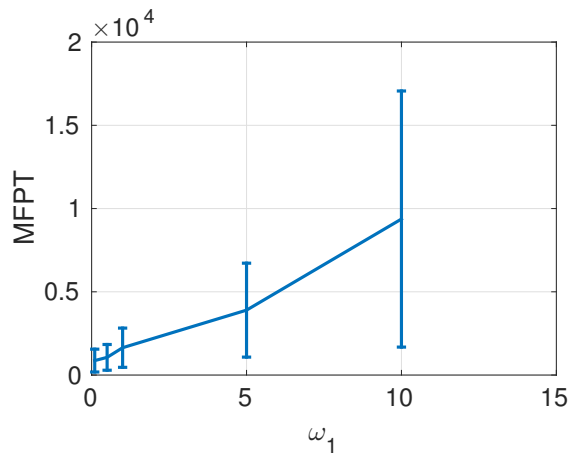
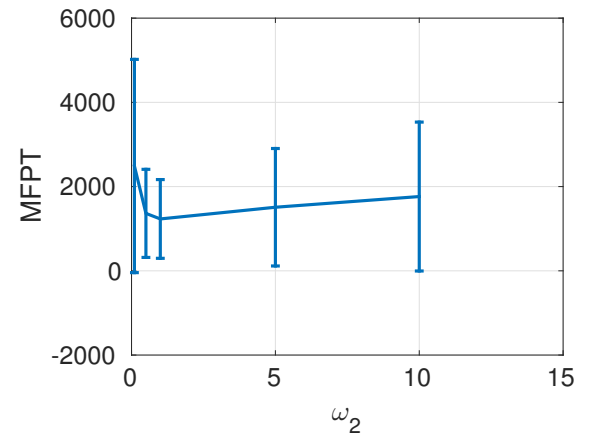
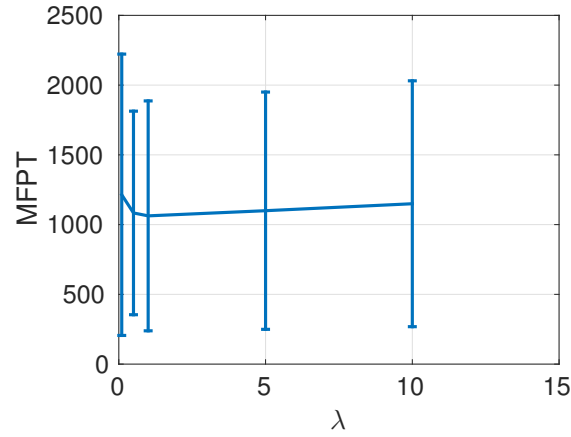
(a) Microtubule bias,  $\phi$ .(b) Motor speeds,  $\nu_1$  and  $\nu_2$ .(c) Switching rate to diffusion,  $\omega_1$ .(d) Switching rate to active transport,  $\omega_2$ .(e) Reorientation rate in diffusion,  $\lambda$ .

Figure 8: Dependence of MFPT on the model parameters. Parameters used for simulations were  $\phi = 0.58$ ,  $\nu_1 = 1.16\mu\text{ms}^{-1}$ ,  $\nu_2 = 0.8\mu\text{ms}^{-1}$ ,  $\omega_1 = 0.42\text{s}^{-1}$ ,  $\omega_2 = 0.84\text{s}^{-1}$ ,  $\lambda = 0.11\text{s}^{-1}$ . For each of (a) to (e), we vary a parameter in turn. For (b),  $\nu_2 = \nu_1/2$  while varying  $\nu_1$ . Results are averaged over 100 particles with standard deviation shown on the error bars.



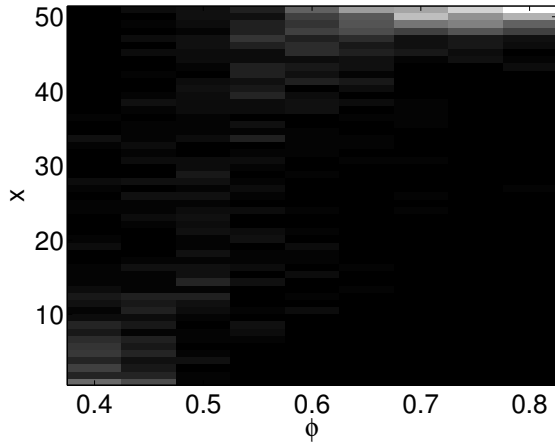
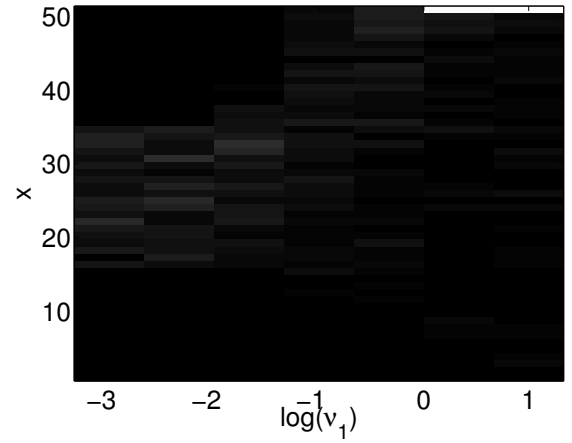
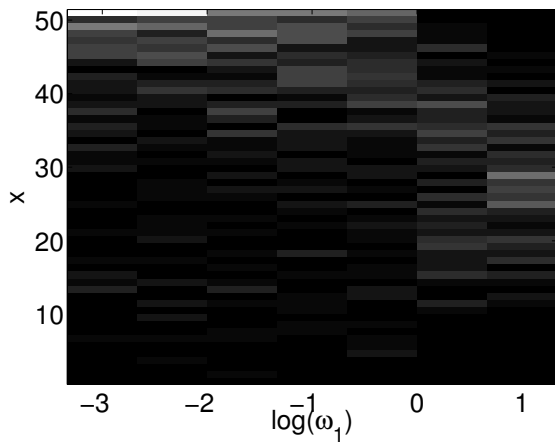
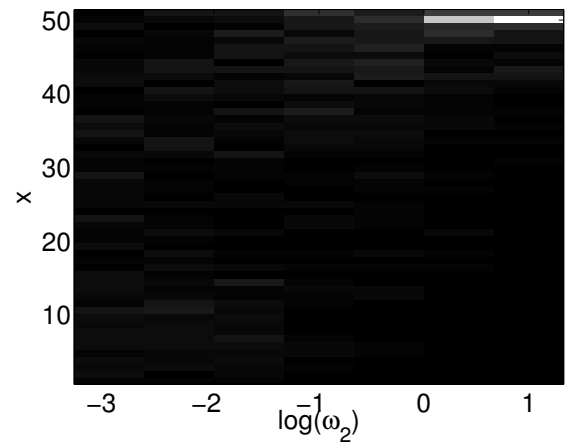
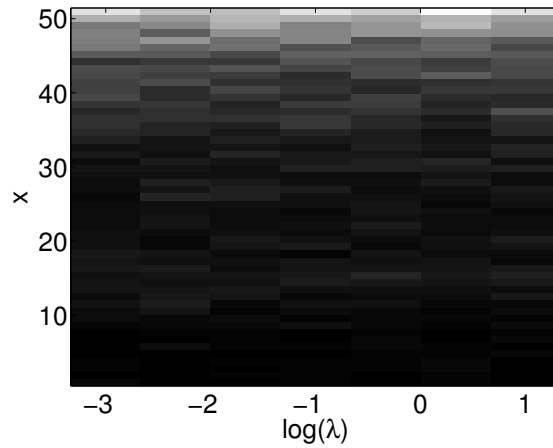
(a) Spatial distribution as a function of  $\phi$ .(b) Spatial distribution as a function of  $\nu_1$  and  $\nu_2$ .(c) Spatial distribution as a function of  $\omega_1$ .(d) Spatial distribution as a function of  $\omega_2$ .(e) Spatial distribution as a function of  $\lambda$ .

Figure 9: The spatial distribution in the  $x$  direction of particles simulated for parameters as in Figure 8. For each of (a) to (e), we vary a parameter in turn. For (b),  $\nu_2 = \nu_1/2$  while varying  $\nu_1$ . Results are shown for 100 particles at time  $t = 270s$ . Note the final spatial box (indexed as box 52) represents particles that have been absorbed and removed from the system.

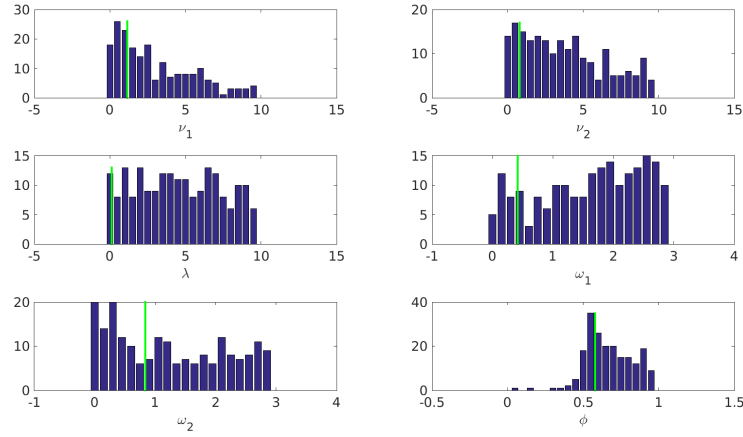
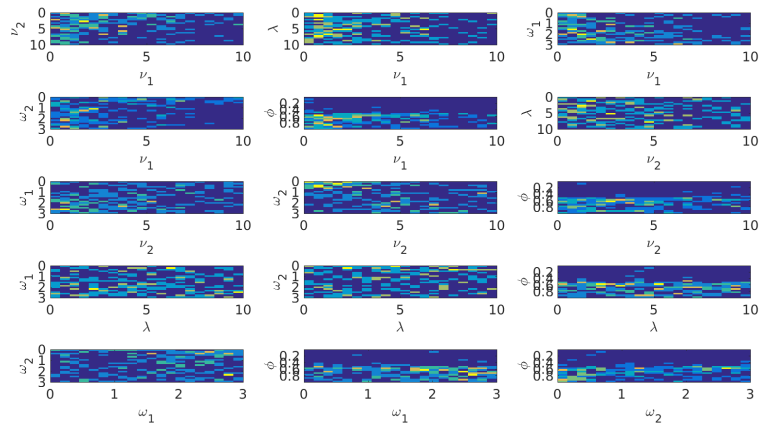
(a) Posterior obtained for  $N = 2000$ ,  $\alpha = 0.1$ .(b) Posterior obtained for  $N = 2000$ ,  $\alpha = 0.1$ .

Figure 10: Posterior for each parameter approximated via ABC rejection sampling, using  $N = 2000$ ,  $\alpha = 0.1$ ,  $n_{repeats} = 1$ . In 10a, histograms for each parameter are shown, with a green line for the real parameter values used for the simulated data. For 10b, heatmaps are shown pairwise for the posterior parameters.

biologically realistic maximum and minimum values. In Figure 10, heatmaps of the pairwise parameter distributions are shown, along with histograms for each parameter.

We obtain most information from the posterior for the microtubule bias,  $\phi$ , while the speed parameters,  $\nu_1$  and  $\nu_2$ , are also identifiable. The remaining rate parameters have posteriors similar to the uniform prior and it is harder to recover the original parameter values used. This suggests that the microtubule bias and the speeds have most effect on the summary statistic chosen, which agrees with our earlier sensitivity analysis.

### 3.5 Incorporation of production into the model

Since we have suggested that production is an important part of the mRNA localization process, we choose to examine the

effect of incorporating this into our model set out previously in Section 2.2. We assume a constant rate of production,  $\gamma$ , modelling the export of RNP complexes from the NPC as discrete events that occur as a Poisson process with rate  $\gamma$ . Although there is some evidence that the production rate varies during oogenesis, we will assume here that it remains constant for simplification. Upon export of a particle in our model, it is placed randomly on the circumference of the nucleus and the model evolves as outlined previously. We begin at time  $t = 0$  with a single particle that has just been exported.

By introducing the extra parameter,  $\gamma$ , into our model, there is now an extra parameter to estimate. To estimate this directly, we note that the rate of production of mRNA is limited by its rate of transcription in the nucleus. Since the rate of transcription for *Drosophila* at 22°C is 25 nt/s

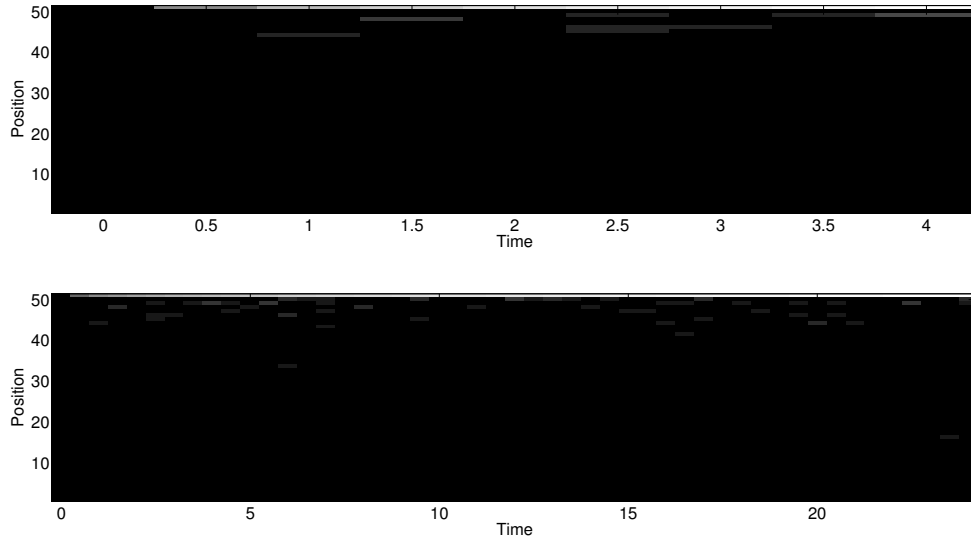


Figure 11: The spatial distribution evolving over time up to 25 hours from a single simulation of the model with production included and parameters of  $\phi = 0.58$ ,  $\nu_1 = 1.16\mu\text{ms}^{-1}$ ,  $\nu_2 = 0.8\mu\text{ms}^{-1}$ ,  $\omega_1 = 0.42\text{s}^{-1}$ ,  $\omega_2 = 0.84\text{s}^{-1}$ ,  $\lambda = 0.11\text{s}^{-1}$ ,  $\gamma = 0.01\text{s}^{-1}$ . A logarithmic scale for the colourmap is used to emphasise the presence of small numbers of particles in the nurse cell.

(BNID 111484) (49), and the length of the *grk* gene is 2513 nt (50), we can obtain an estimate for the production rate of *grk* mRNA of  $\gamma = 0.01\text{s}^{-1}$ . Simulation from the extended model with production at rate  $\gamma = 0.01\text{s}^{-1}$  suggests that a spatial steady state is reached after a short initial transient, with a constant rate of absorption at the posterior ring canal. Results are shown in Figure 11 and exhibit spatial profiles similar to those seen in experimental data exhibited in Figure 12.

### 3.6 Application of ABC to *Drosophila* nurse cell data

Although comparative simulation results with ABC suggest that the detailed summary statistics using MFPT, number of jumps and jump distances result in a higher quality posterior, there are limitations to what data can be collected experimentally. In particular, these summary statistics would require detailed tracking data of individual particles over an extended period of time, which would be difficult to obtain. Instead, we use the spatial distribution of particles which is more tractable from a single time point. To ensure the timescales of the observed experimental data and simulated data match, we use the model with production of mRNA at rate  $\gamma$  and infer this parameter also.

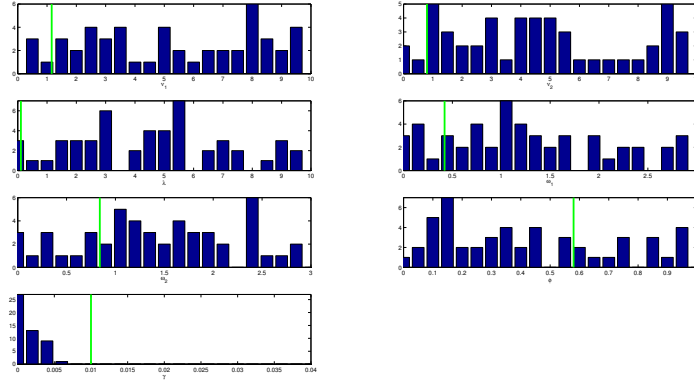
Applying inference with ABC to these data using a Euclidean distance measure and binned spatial histograms in the  $x$  direction as summary statistics, we obtain posteriors for the model parameters as shown in Figure 13. Since simulating from the model with production over longer timescales is more computationally intensive than simulating from the

original model, as in Section 3.4, we must sample fewer parameter values, taking  $N = 500$  and  $\alpha = 0.1$  in the ABC rejection sampling algorithm. The results of the ABC inference show a very tight distribution for  $\gamma$ , despite the low value of  $N$  used in these simulations. This suggests that the output of the production model depends most strongly on  $\gamma$  when summarised by our chosen summary statistics. Little extra information beyond the prior is gained for the other parameters, based on these results with  $N = 500$ .

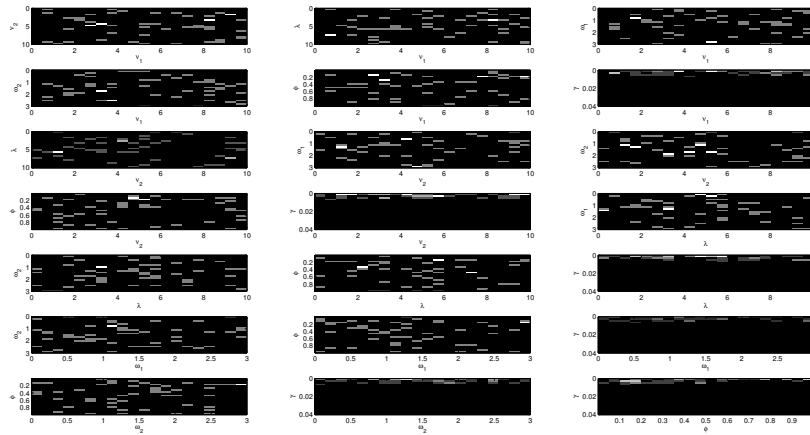
## 4 Conclusion

Through use of a model-based approach for explaining mRNA transport and localization, we have been able to represent the dynamics of this process and reproduce behaviour observed experimentally, showing that a biased network of microtubules can maintain localization of mRNA. We constructed a velocity jump process with two phases of motion to represent directed movement via active transport on microtubules and random motion in free diffusion. Since simulations of this model were accessible but the likelihood was intractable, we employed ABC methods to infer the parameters from our model from data collected via live *in vitro* imaging and compared the results of this with values obtained from direct measurement.

The timescales for MFPTs based on simulations from the model are on the order of tens of minutes, rather than hours, for the inferred parameters, suggesting that the transport step



(a) Histogram for each parameter.



(b) Heatmap for pairwise parameters.

Figure 13: Posterior for each parameter in the full model with production approximated via ABC rejection sampling, using  $N = 500$ ,  $\alpha = 0.1$ , based on real data from one single image frame. In 13a, histograms for each parameter are shown, with a green line for the directly estimated parameter values for comparison. For 13b, heatmaps are shown pairwise for the posterior parameters.

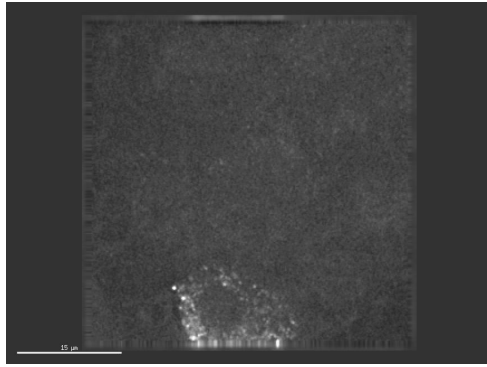
is not the rate-limiting stage of the mRNA localization process. Based on the modelling work presented here, we postulate that the production of mRNA in the nucleus of the nurse cells is instead the rate-limiting stage.

By incorporating production at a constant rate into our model, we recover transport on timescales of the order of hours seen during oogenesis. We are also able to simulate spatial distributions of mRNA particles similar to what is observed experimentally.

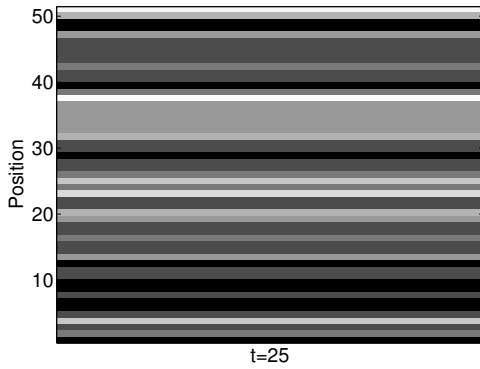
#### 4.1 Further work

The subcellular environment is in many ways drastically different to that assumed in our spatially homogeneous two-dimensional model. In reality, cells exist in three dimensions and are crowded inside with a heterogeneous environment.

One possible method of representing heterogeneity in the cell would be by use of a potential to account for volume exclusion effects, as used by Isaacson et al. (51) in the context of nuclear export of mRNAs. The three-dimensional nature of the environment presents further challenges for tracking, as RNPs move between frames. Although some progress has already been made in this area (52), we hope that incorporation of modelling into a tracking framework could enable tracking of particles over longer timescales at increased resolution. Motivated by a Kalman filter type approach (53), we hope to investigate further using our model combined with regular inputs of data over time to assist particle tracking algorithms in the linkage step of tracking. Therefore, in future, the model should be extended to enable us to capture behaviour in three dimensions. Since the oocyte more than triples in size during the early stages



(a) Raw data after denoising and deconvolution.



(b) Spatial distribution of *grk* particles seen in raw data.

Figure 12: Distribution of *grk* MCP Dendra in a stage 5 *Drosophila* oocyte and surrounding nurse cells. Stage 5 is equivalent to a time point of around  $t = 25$  hours, assuming *grk* mRNA production begins at stage 2. A logarithmic scale for the colourmap is used in (b).

of oogenesis, it may be relevant to include domain growth in our model in future, as the dimensions of the domain effect the timescales of the MFPT of the process.

If further time were available, we would investigate further the parameterization of the full model from experimental data via ABC, which may require use of other ABC algorithms, such as Adaptive Population Monte Carlo ABC (42), to overcome the inefficiencies of the rejection sampling method. We could compare more comprehensively a variety of summary statistics to guide the use of experiments in inference of our model parameters.

## SUPPLEMENTARY MATERIAL

Two Sections and one Figure are available as a supplement to this article below and can be found by visiting BJ Online at <http://www.biophysj.org>.

The authors would like to acknowledge valuable discussions with Dr S. Filippi and Dr J. Rittscher.

## A Further sensitivity analysis

Here, we show, in Figure 14, the sensitivity of the mean number of jumps per simulation and the mean of the median jump distance to the parameters of the basic model without production.

## B Analytics

For many models, analytical techniques can offer more general insight than simulating from the model with specific parameters. We show here what can be achieved with these techniques and give some indicative results. First, we must make some simplifying assumptions to ensure the model is analytically tractable. We will consider only one spatial dimension, parallel to the  $x$  axis, restricted to the interval  $[0, L_x]$ . Furthermore, we assume  $\nu_1 \gg \nu_2$  so that it is reasonable to ignore the diffusive phase of motion, leaving only reorientations with the active transport phase at rate  $\omega_1$ .

Suppose we have a velocity jump process with one mode of motion in one dimension where the speed in the  $x$  direction varies as  $\nu_1 \sin(\theta)$  when the velocity is oriented in direction  $\theta$  in two dimensions. Let  $p(x, \theta, t|y, \alpha, s)dt$  be the probability that the particle is at position  $x$  moving in direction  $\theta$  in the time interval  $[t, t+dt)$  having started previously at time  $s$  at position  $y$  moving in direction  $\alpha$ . We will now suppress the conditioning and write only  $p(x, \theta, t)$ .

The evolution of the probability is governed (54) by:

$$\frac{\partial p}{\partial t} + \nu_1 \sin \theta \frac{\partial p}{\partial x} = -\omega_1 p + \omega_1 \int_0^{2\pi} T(\theta) p(x, \theta', t) d\theta', \quad (1)$$

where the turning kernel  $T(\theta)$  gives the probability of reorienting into a new angle  $\theta$  and we assume this does not depend on the previous angle. We define  $\langle f(\theta) \rangle = \int_0^{2\pi} f(\theta) d\theta$  and  $\langle \langle f(x, \theta) \rangle \rangle = \int_0^{2\pi} \int_0^{2\pi} f(\theta) d\theta dx$ .

The orientation of the velocity,  $\theta$ , is updated via the reorientation kernel  $T(\theta)$  after every transition. It depends only on the initial condition and the kernel  $T(\theta)$ . After a sufficiently long time, the initial condition will have been forgotten and we should be able to decompose the probability  $p$  such that:

$$p(x, \theta, t) = T(\theta) q(x, t), \quad (2)$$

since whenever the angle  $\theta$  is updated, this is independent of time or position. For a particle to forget the initial condition, it must simply execute at least one jump. After a time  $t_{\min} = 1/\omega_1 \log(100)$ , then 99% of the particles should have executed a jump, since this is the 99th percentile of an exponential distribution with parameter  $\omega_1$ . Assuming a value of  $\omega_1 = 0.42s^{-1}$ , then  $t_{\min} = 11s$  which is much less than our timescales of interest.

Now integrating (1) over  $\theta$ , we obtain:

$$\frac{\partial}{\partial t} \langle p \rangle + \frac{\partial}{\partial x} \langle \nu_1 \sin(\theta) p \rangle = 0, \quad (3)$$

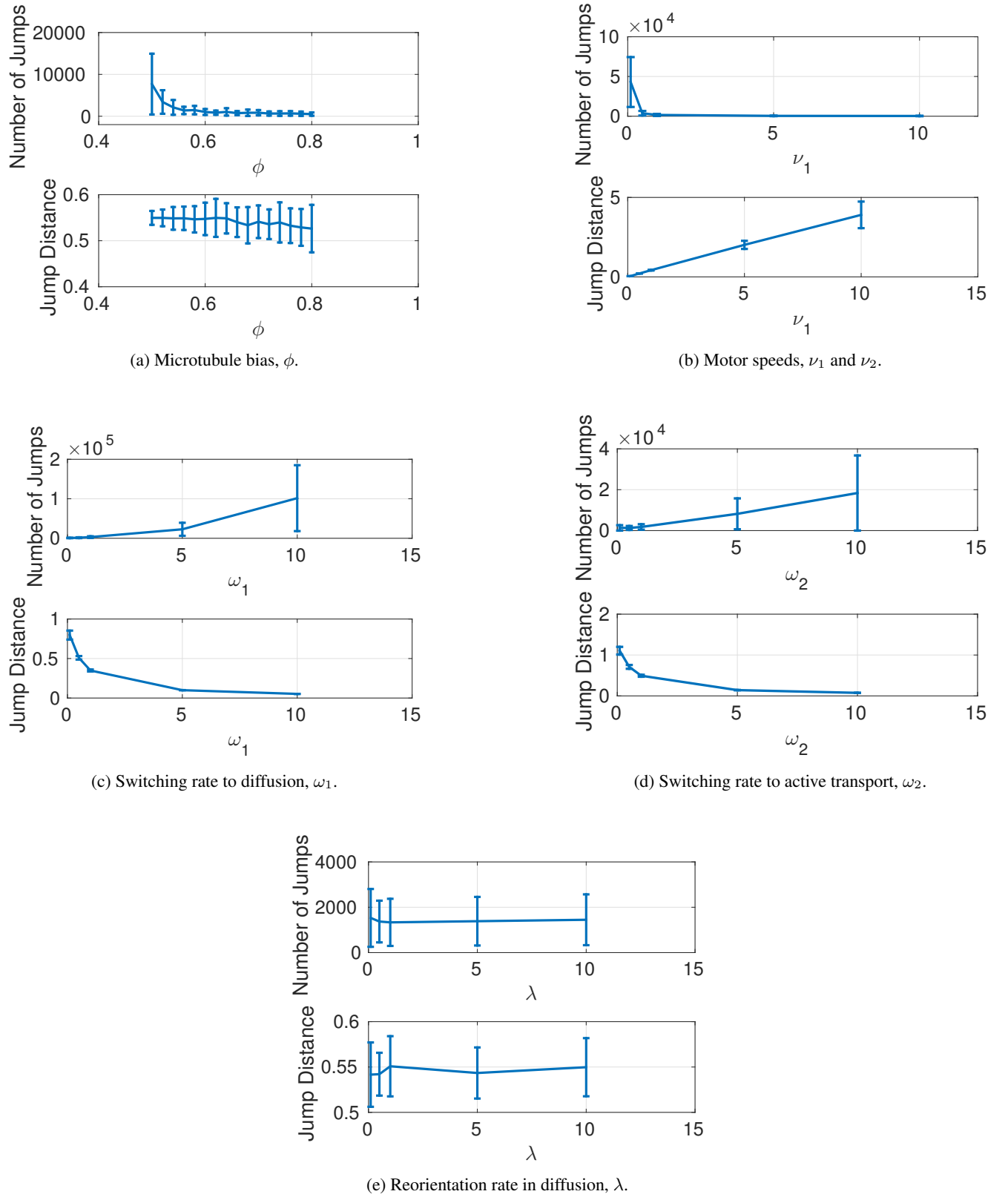


Figure 14: Dependence of the number of jumps and the length of jumps on the model parameters. Parameters used for simulations were  $\phi = 0.58$ ,  $\nu_1 = 1.16\mu\text{ms}^{-1}$ ,  $\nu_2 = 0.8\mu\text{ms}^{-1}$ ,  $\omega_1 = 0.42\text{s}^{-1}$ ,  $\omega_2 = 0.84\text{s}^{-1}$ ,  $\lambda = 0.11\text{s}^{-1}$ . For each of (a) to (e), we vary each parameter in turn. For (b),  $\nu_2 = \nu_1/2$  while varying  $\nu_1$ . Results are averaged over 100 particles with standard deviation shown on the error bars.

which becomes via our assumption (2):

$$\frac{\partial q}{\partial t} + \nu_1 \langle T(\theta) \sin(\theta) \rangle \frac{\partial q}{\partial x} = 0. \quad (4)$$

From multiplying (1) by  $x$  and considering the first spatial moment, we have:

$$\begin{aligned} \frac{\partial}{\partial t} (\langle xp \rangle) &= \nu_1 \langle \sin(\theta) p \rangle, \\ \frac{\partial}{\partial t} (\langle xT(\theta)q \rangle) &= \nu_1 \langle \sin(\theta)T(\theta)q \rangle, \\ \frac{\partial}{\partial t} \left( \int_0^L xq \, dx \right) &= \nu_1 \langle \sin(\theta)T(\theta) \rangle \int_0^L q \, dx. \end{aligned}$$

Suppose we let  $\mu^{(1)}(t) = \int_0^L xq(x, t)dx$ . Then we can write:

$$\begin{aligned} \frac{d\mu^{(1)}}{dt} &= \nu_1 \langle \sin(\theta)T(\theta) \rangle \int_0^L q \, dx, \\ \mu^{(1)}(t) &= \mu^{(1)}(0) + \nu_1 \langle \sin(\theta)T(\theta) \rangle \int_0^t \int_0^L q \, dxdt. \quad (5) \end{aligned}$$

We can interpret  $\int_0^L q(x, t) \, dx$  as a survival probability of a particle at time  $t$ . After very short times, when the particle cannot physically have exited the system yet, this will equal one, leading to a linear form for the average position,  $\mu^{(1)}(t)$ . At later times, the particle may have exited the system and the survival probability will decrease to 0. If we release an ensemble of particles, the average position of this ensemble changes according to the mean net velocity,  $\nu_1 \langle \sin(\theta)T(\theta) \rangle$ , until particles begin to be absorbed. We note that this mean net velocity depends on the microtubule bias. For our specified choice of  $T(\theta)$ , we have  $\nu_1 \langle \sin(\theta)T(\theta) \rangle = \nu_1(4\phi - 2)/\pi$ . Using our estimated parameters,  $\phi = 0.58$  and  $\nu_1 = 1.16\mu\text{ms}^{-1}$ , we find an estimate of the mean net velocity of  $0.12\mu\text{ms}^{-1}$ .

By similar arguments, expressions for higher spatial moments and angular moments can be obtained. For the  $j$ th spatial moment, we have

$$\frac{\partial \mu^{(j)}}{\partial t} = j\nu_1 \langle \sin(\theta)T(\theta) \rangle \mu^{(j-1)}, \quad (6)$$

leading to a hierarchy of equations. Extensions of these arguments to the full model with two phases of motion are also possible.

## References

1. Wolpert, L., R. Beddington, J. Brockes, T. Jessell, P. Lawrence, and E. Meyerowitz, 1998. *Principles of Development*. Oxford University Press, Oxford.
2. Wilkie, G. S., and I. Davis, 2001. *Drosophila* wingless and pair-rule transcripts localize apically by dynein-mediated transport of RNA particles. *Cell* 105:209–219.
3. Bobola, N., R.-P. Jansen, T. H. Shin, and K. Nasmyth, 1996. Asymmetric accumulation of Ash1p in postanaphase nuclei depends on a myosin and restricts yeast mating-type switching to mother cells. *Cell* 84:699–709.
4. Mowry, K. L., and D. A. Melton, 1992. Vegetal messenger RNA localization directed by a 340-nt RNA sequence element in *Xenopus* oocytes. *Science* 255:991–994.
5. Rosbash, M., and R. H. Singer, 1993. RNA travel: tracks from DNA to cytoplasm. *Cell* 75:399–401.
6. Nevo-Dinur, K., A. Nussbaum-Shochat, S. Ben-Yehuda, and O. Amster-Choder, 2011. Translation-independent localization of mRNA in *E. coli*. *Science* 331:1081–1084.
7. Jeffery, W. R., C. R. Tomlinson, and R. D. Brodeur, 1983. Localization of actin messenger RNA during early ascidian development. *Developmental Biology* 99:408–417.
8. Bertrand, E., P. Chartrand, M. Schaefer, S. M. Shenoy, R. H. Singer, and R. M. Long, 1998. Localization of ASH1 mRNA particles in living yeast. *Molecular Cell* 2:437–445.
9. Hamilton, R. S., R. M. Parton, R. A. Oliveira, G. Vendra, G. Ball, K. Nasmyth, and I. Davis, 2010. ParticleStats: open source software for the analysis of particle motility and cytoskeletal polarity. *Nucleic Acids Research* 38:W641–W646.
10. Parton, R. M., A. Davidson, I. Davis, and T. T. Weil, 2014. Subcellular mRNA localisation at a glance. *Journal of Cell Science* 127:2127–2133.
11. Forrest, K. M., and E. R. Gavis, 2003. Live imaging of endogenous RNA reveals a diffusion and entrapment mechanism for nanos mRNA localization in *Drosophila*. *Current Biology* 13:1159–1168.
12. Jaramillo, A. M., T. T. Weil, J. Goodhouse, E. R. Gavis, and T. Schupbach, 2008. The dynamics of fluorescently labeled endogenous gurken mRNA in *Drosophila*. *Journal of Cell Science* 121:887–894.
13. Weil, T. T., K. M. Forrest, and E. R. Gavis, 2006. Localization of bicoid mRNA in late oocytes is maintained by continual active transport. *Developmental Cell* 11:251–262.
14. Zimyanin, V. L., K. Belaya, J. Pecreaux, M. J. Gilchrist, A. Clark, I. Davis, and D. St Johnston, 2008. In vivo imaging of oskar mRNA transport reveals the mechanism of posterior localization. *Cell* 134:843–853.
15. Jansen, R.-P., 2001. mRNA localization: message on the move. *Nature Reviews Molecular Cell Biology* 2:247–256.
16. Clark, A., C. Meignin, and I. Davis, 2007. A Dynein-dependent shortcut rapidly delivers axis determination transcripts into the *Drosophila* oocyte. *Development* 134:1955–1965.
17. Lasko, P., 1999. RNA sorting in *Drosophila* oocytes and embryos. *The FASEB journal* 13:421–433.
18. Weil, T. T., R. M. Parton, and I. Davis, 2010. Making the message clear: visualizing mRNA localization. *Trends in Cell Biology* 20:380–390.
19. Rosser, G., A. G. Fletcher, D. A. Wilkinson, J. A. de Beyer, C. A. Yates, J. P. Armitage, P. K. Maini, and R. E. Baker, 2013. Novel methods for analysing bacterial tracks reveal persistence in *Rhodobacter sphaeroides*. *PLOS Comp Biol* 9:1:18.
20. Persson, F., M. Lindén, C. Unoson, and J. Elf, 2013. Extracting intracellular diffusive states and transition rates from



- single-molecule tracking data. *Nature Methods* 10:265–269.
21. Hamilton, R. S., G. Ball, and I. Davis, 2013. A multi-disciplinary approach to RNA localisation. In *Biophysical approaches to translational control of gene expression*, Springer, 213–233.
  22. Weil, T. T., R. M. Parton, B. Herpers, J. Soetaert, T. Veenendaal, D. Xanthakis, I. M. Dobbie, J. M. Halstead, R. Hayashi, C. Rabouille, et al., 2012. Drosophila patterning is established by differential association of mRNAs with P bodies. *Nature Cell Biology* 14:1305–1313.
  23. Davidson, A., 2015. RNA localisation and remodelling in development. Ph.D. thesis, University of Oxford.
  24. Vendra, G., R. S. Hamilton, and I. Davis, 2007. Dynactin suppresses the retrograde movement of apically localized mRNA in Drosophila blastoderm embryos. *RNA* 13:1860–1867.
  25. Howard, J., 2001. Mechanics of motor proteins and the cytoskeleton. Sinauer, Sunderland, MA.
  26. Amrute-Nayak, M., and S. L. Bullock, 2012. Single-molecule assays reveal that RNA localization signals regulate dynein-dynactin copy number on individual transcript cargoes. *Nature Cell Biology* 14:416–423.
  27. Buxbaum, A. R., H. Gal, and S. R. H., 2015. In the right place at the right time: visualizing and understanding mRNA localization. *Nature Reviews Molecular Cell Biology* 16:95–109.
  28. Parton, R. M., R. S. Hamilton, G. Ball, L. Yang, C. F. Cullen, W. Lu, H. Ohkura, and I. Davis, 2011. A PAR-1-dependent orientation gradient of dynamic microtubules directs posterior cargo transport in the Drosophila oocyte. *The Journal of Cell Biology* 194:121–135.
  29. Delanoue, R., and I. Davis, 2005. Dynein anchors its mRNA cargo after apical transport in the Drosophila blastoderm embryo. *Cell* 122:97–106.
  30. Bressloff, P. C., and J. M. Newby, 2013. Stochastic models of intracellular transport. *Reviews of Modern Physics* 85:135.
  31. Szymańska, Z., M. Parisot, and M. Lachowicz, 2014. Mathematical modeling of the intracellular protein dynamics: The importance of active transport along microtubules. *Journal of Theoretical Biology* 363:118–128.
  32. Ganguly, S., L. S. Williams, I. M. Palacios, and R. E. Goldstein, 2012. Cytoplasmic streaming in Drosophila oocytes varies with kinesin activity and correlates with the microtubule cytoskeleton architecture. *Proceedings of the National Academy of Sciences* 109:15109–15114.
  33. Liu, W., and M. Niranjana, 2011. The role of regulated mRNA stability in establishing bicoid morphogen gradient in Drosophila embryonic development. *PLOS ONE* 6:e24896–e24896.
  34. Johnston, S. T., M. J. Simpson, D. S. McElwain, B. J. Binder, and J. V. Ross, 2014. Interpreting scratch assays using pair density dynamics and approximate Bayesian computation. *Open Biology* 4:140097.
  35. Turner, B. M., and T. Van Zandt, 2012. A tutorial on approximate Bayesian computation. *Journal of Mathematical Psychology* 56:69–85.
  36. Beaumont, M. A., W. Zhang, and D. J. Balding, 2002. Approximate Bayesian computation in population genetics. *Genetics* 162:2025–2035.
  37. Codling, E., and N. Hill, 2005. Calculating spatial statistics for velocity jump processes with experimentally observed reorientation parameters. *Journal of Mathematical Biology* 51:527–556.
  38. Taylor-King, J., E. van Loon, G. Rosser, and S. Chapman, 2015. From birds to bacteria: generalised velocity jump processes with resting states. *Bulletin of Mathematical Biology* 1–24.
  39. Toni, T., D. Welch, N. Strelkowa, A. Ipsen, and M. P. Stumpf, 2009. Approximate Bayesian computation scheme for parameter inference and model selection in dynamical systems. *Journal of the Royal Society Interface* 6:187–202.
  40. Sisson, S. A., Y. Fan, and M. M. Tanaka, 2007. Sequential Monte Carlo without likelihoods. *Proceedings of the National Academy of Sciences* 104:1760–1765.
  41. Beaumont, M. A., J.-M. Cornuet, J.-M. Marin, and C. P. Robert, 2009. Adaptive approximate Bayesian computation. *Biometrika* 96:983–990.
  42. Lenormand, M., F. Jabot, and G. Deffuant, 2013. Adaptive approximate Bayesian computation for complex models. *Computational Statistics* 28:2777–2796.
  43. Liepe, J., H. Taylor, C. P. Barnes, M. Huvet, L. Bugeon, T. Thorne, J. R. Lamb, M. J. Dallman, and M. P. Stumpf, 2012. Calibrating spatio-temporal models of leukocyte dynamics against in vivo live-imaging data using approximate Bayesian computation. *Integrative Biology* 4:335–345.
  44. van der Vaart, E., M. A. Beaumont, A. S. Johnston, and R. M. Sibly, 2015. Calibration and evaluation of individual-based models using approximate Bayesian computation. *Ecological Modelling* 312:182–190.
  45. Robert, C. P., J.-M. Cornuet, J.-M. Marin, and N. S. Pillai, 2011. Lack of confidence in approximate Bayesian computation model choice. *Proceedings of the National Academy of Sciences* 108:15112–15117.
  46. Nunes, M. A., and D. J. Balding, 2010. On optimal selection of summary statistics for approximate Bayesian computation. *Statistical Applications in Genetics and Molecular Biology* 9.
  47. Felzenszwalb, P. F., and D. P. Huttenlocher, 2004. Efficient graph-based image segmentation. *International Journal of Computer Vision* 59:167–181.
  48. Neuman-Silberberg, F. S., and T. Schüpbach, 1996. The Drosophila TGF- $\alpha$ -like protein Gurken: expression and cellular localization during Drosophila oogenesis. *Mechanisms of Development* 59:105–113.
  49. Milo, R., P. Jorgensen, U. Moran, G. Weber, and M. Springer, 2010. BioNumbers: the database of key numbers in molecular and cell biology. *Nucleic Acids Research* 38:750–753.
  50. Marygold, S. J., P. C. Leyland, R. L. Seal, J. L. Goodman, J. Thurmond, V. B. Strelets, R. J. Wilson, F. Consortium, et al., 2013. FlyBase: improvements to the bibliography. *Nucleic Acids Research* 41:751–757.
  51. Isaacson, S., D. McQueen, and C. S. Peskin, 2011. The influence of volume exclusion by chromatin on the time required to find specific DNA binding sites by diffusion. *Proceedings of the National Academy of Sciences* 108:3815–3820.
  52. Thompson, M. A., J. M. Casolari, M. Badieirostami, P. O. Brown, and W. Moerner, 2010. Three-dimensional tracking of single mRNA particles in *Saccharomyces cerevisiae* using a double-helix point spread function. *Proceedings of the National Academy of Sciences* 107:17864–17871.
  53. Faragher, R., et al., 2012. Understanding the basis of the

- Kalman filter via a simple and intuitive derivation. IEEE Signal Processing Magazine 29:128–132.
54. Othmer, H. G., S. R. Dunbar, and W. Alt, 1988. Models of dispersal in biological systems. Journal of Mathematical Biology 26:263–298.