Is cell migration or proliferation dominant in the formation of linear arrays of oligodendrocytes?

Darragh M. Walsh[a,*], Philipp T. Röth[b], William R. Holmes[a,c], Kerry A. Landman[a], Tobias D. Merson[b], Barry D. Hughes[a]

[a]*School of Mathematics and Statistics, University of Melbourne, Victoria 3010, Australia*
[b]*The Florey Institute of Neuroscience and Mental Health, Parkville, Victoria 3010, Australia*
[c]*Department of Physics and Astronomy, Vanderbilt University, Tennessee 37240, USA*

Abstract

Oligodendrocytes are the myelin-producing cells of the central nervous system that are responsible for electrically insulating axons to speed the propagation of electrical impulses. A striking feature of oligodendrocyte development within white matter is that the cell bodies of many oligodendrocyte progenitor cells become organized into discrete linear arrays of three or more cells before they differentiate into myelin-producing oligodendrocytes. These linear arrays align parallel to the direction of the axons within white matter tracts and are believed to play an important role in the co-ordination of myelination. Guided by experimental data on the abundance and composition of linear arrays in the corpus callosum of the postnatal mouse brain, we construct discrete and continuous models of linear array generation to specifically investigate the relative influence of cell migration, proliferation, differentiation and death of oligodendroglia upon the genesis of linear arrays during early postnatal development. We demonstrate that only models that incorporate significant cell migration can replicate all of the experimental observations on number of arrays, number of cells in arrays and total cell count of oligodendroglia within a given area of the corpus callosum. These models are also necessary to accurately reflect experimental data on the abundance of linear arrays composed of oligodendrocytes that derive from progenitors of different clonal origins.

Keywords: Migration, proliferation, structure formation, myelination, oligodendrocyte and oligodendrocyte progenitor cell

1. Introduction

Oligodendrocytes (OLs) are the cells in the central nervous system (CNS) responsible for the myelination of axons. The long thin sheaths of myelin that wrap around segments of axons provide both electrical and physical insulation which enables fast conduction of electrical signals along axons, known as saltatory conduction (Baumann and Pham-Dinh 2001). OLs also provide critical trophic and metabolic support of neurons (Nave 2010; Saab *et al.* 2013). Interestingly, the abundance and topographic organization of myelin differs greatly among different parts of the CNS. In the mouse, for example, almost the entire length of every axon in the optic nerve is myelinated (Dangata and Kaufman 1997) whilst in the corpus callosum (CC), only around 30% of axonal profiles are myelinated (Sturrock 1980). In the cerebral cortex, the axons of pyramidal cells are composed of myelinated segments interspersed with long, unmyelinated tracts rather than series of uninterrupted myelin internodes (Tomassy *et al.* 2014). Collectively these findings raise the possibility that myelinated axons in white matter tracts such as the CC could also exhibit both myelinated and unmyelinated segments along their length.

---
[*]Tel: +613 83440584. Email: darragh.walsh@unimelb.edu.au

Email addresses: `darragh.walsh@unimelb.edu.au` (Darragh M. Walsh), `philipp.roth@florey.edu.au` (Philipp T. Röth), `william.holmes@vanderbilt.edu` (William R. Holmes), `kerryl@unimelb.edu.au` (Kerry A. Landman), `tobias.merson@florey.edu.au` (Tobias D. Merson), `barrydh@unimelb.edu.au` (Barry D. Hughes)

The implication that myelination is controlled in a context-dependent manner is supported by magnetic resonance imaging studies in humans which demonstrate that learning skills such as juggling or piano playing is associated with increased myelination of neuronal circuits required for that behaviour (Scholz *et al.* 2009; Bengtsson *et al.* 2005). Similarly, recent studies in mice indicate that the generation of new myelinating OLs is critical for learning a complex motor task in adulthood (McKenzie *et al.* 2014). A role for electrical activity within axons in regulating the process of selecting axons for myelination is becoming increasingly clear (Gibson *et al.* 2014; Hines *et al.* 2015; Mensch *et al.* 2015). Despite these recent advances, the cellular dynamics responsible for orchestrating the precise topographic distribution of myelin along axons within white matter remains poorly understood. Clarifying the mechanisms underlying the spatial organization of OLs could also shed light on how myelin is regenerated after myelin loss such as in multiple sclerosis, a demyelinating disease in which OLs are killed as a result of autoimmune attack. Consequently, there is considerable interest in understanding the cellular dynamics underlying remyelination after demyelinating injury (Xing *et al.* 2014), as well as during brain development and ongoing adult myelin plasticity.

How then is myelin topography established within white matter? Answering this question requires careful dissection of the sequential development of OLs from oligodendrocyte progenitor cells (OPC), a highly motile and proliferative cell population that differentiates into OLs through a series of tightly coordinated stages (Mitew *et al.* 2014). Whilst electrical activity within axons has been demonstrated to promote OPC differentiation (Gibson *et al.* 2014) and to bias the selection of axons that will be myelinated (Hines *et al.* 2015), additional clues underlying the establishment of myelin topography have arisen from studies examining how OL patterning emerges within white matter. Röth *et al.* (2016) have described that OLs in white matter tracts such as the CC tend to organize themselves into linear arrays typically composed of 3-8 cells that align in parallel with the longitudinal axis of axons in the tract. These linear arrays are similar to an earlier description of rows of interfascicular OLs in the rat fimbria (Suzuki and Raisman 1997).

Linear arrays could form, for example, through proliferation of individual, stationary OPCs prior to maturation into OLs, or through a process such as migration and attachment where unattached OPCs attach to another OPC or OL to generate arrays. Here we develop both discrete stochastic and continuous deterministic models to investigate the relative importance of cell migration, proliferation, differentiation and death for the development of these array structures in the mouse CC. Experimental data on the number of arrays, number of cells in arrays and the extent to which OLs within linear arrays in the CC derive from different clonal origins (from dorsal Emx1- specified or ventral Gsh2 - specified OPCs) are used to test the relative importance of cell migration and proliferation to array formation.

The modelling supports the hypothesis that linear arrays of oligodendrocytes in the CC are primarily generated as a result of the preferential attachment of migratory OPCs. *In situ* clonal expansion also plays a role, but the preference for arrays to be of mixed clonal origin establishes that *in situ* clonal expansion is a lesser contributor than migration and attachment.

2. Biological background

Here we briefly describe the structural properties and temporal dynamics of linear array development in the postnatal CC of the mouse. Röth *et al.* (2016) define a linear array of oligodendroglia (OPCs and/or OLs) using three criteria. An array contains at least three cells (OLs, OPCs or both), the distance between cells is less than the diameter of a cell nucleus and the alignment of the array must be parallel to the axonal orientation (here taken to be horizontal). Examples of linear arrays located in the CC of a postnatal day 14 (P14) mouse brain are indicated in Figure 1. The axons of neurons in this coronal tissue section are not visible but are orientated horizontally. A schematic of the region of interest and the alignment of OLs along axons is shown in Figure 2.

OPCs start to migrate into the CC as early as embryonic day 16 (E16) in the first of three distinct waves that are specified in different regions of the developing forebrain. By P30, the CC contains a heterogeneous mixture of two clonally distinct populations, namely Gsh2-specified progenitors derived from the ventral forebrain and Emx1-derived progenitors derived from the dorsal forebrain (Kessaris *et al.* 2006). OPCs in the mouse CC continue to proliferate extensively in the early postnatal period and from P7 start to exit the cell cycle, differentiate into pre-myelinating OLs and lose their migratory capacity (Mitew *et al.* 2014). Pre-myelinating OLs represent a highly transient stage wherein cells that receive the necessary trophic and maturation cues either terminally differentiate into myelinating OLs or undergo programmed cell death (apoptosis) (Barres *et al.* 1992; Trapp *et al.* 1997; Hughes *et al.* 2013). Estimates of the extent of programmed cell death range from 20% to 50% in the developing rat cerebral cortex and optic nerve
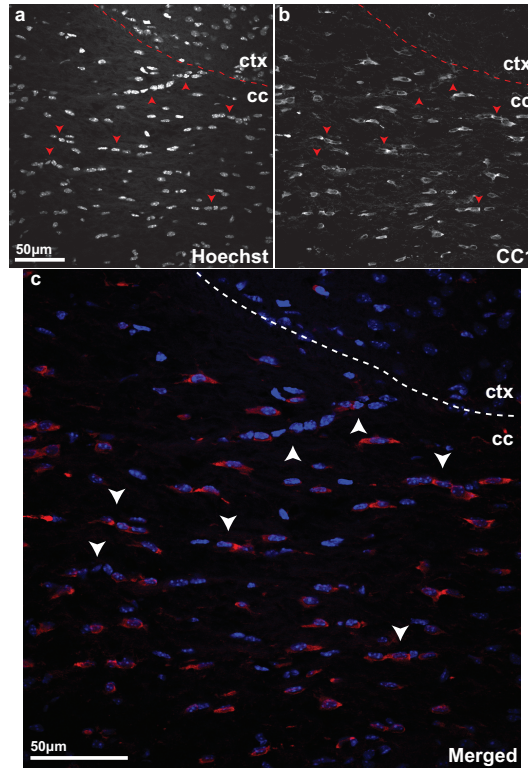
Figure 1: **Immunohistochemical image of linear arrays in the corpus callosum of a postnatal day 14 (P14) mouse sectioned in the coronal plane.** (a) Labelling of cell nuclei using hoechst dye reveals linear array structures of variable length (red arrowheads). (b) CC1-labelled cells to detect mature oligodendrocytes (red arrowheads). (c) Merged image reveals that many mature oligodendrocytes (CC1+ cells) reside within linear arrays (white arrowheads). Red dashed lines outline the border between the corpus callosum (cc) and the cortex (ctx). Scale bar = 50 μm.

.

respectively (Trapp *et al.* 1997; Barres *et al.* 1992). By P14 most OPCs have differentiated into pre-myelinating OLs (Röth *et al.* 2016) and differentiation into myelinating OLs is well advanced by P30 (Sturrock 1980); however a small population of OPCs remain throughout adulthood and continue to produce new OLs, albeit at a reduced rate (Dimou *et al.* 2008; Rivers *et al.* 2008; Zhu *et al.* 2011; Young *et al.* 2013). We use these experimental findings to inform our model parameters.

In order to create a mathematical model of linear array genesis, it was necessary to define various parameters based on empirical data collected from experimental observations. Data for the number of linear arrays, the number of cells localized within linear arrays and total cell count (cells within and outside arrays) were obtained by examination of the postnatal mouse CC at P7, P14 and P30, as described in Röth *et al.* (2016). This necessitated sacrificing the mice involved at each time point. The proportion of OPCs in the cell cycle, defined by co-expression of PdgfRa and Ki67, was found not to differ between array localized and isolated OPCs, both at P7 and at P14 (Röth *et al.* 2016). Genetic fate-mapping experiments to track the anatomical origin of OLs in the CC revealed that approximately half of the linear arrays that contained fate-mapped oligodendroglia at P30 were composed of oligodendroglia derived from both ventral Gsh2-specified progenitors and dorsal Emx1-specified progenitors, providing definitive evidence of clonal mixing. These linear arrays were defined as 'mixed origin arrays'. The remaining half of the fate-mapped linear arrays were either Gsh2-specified or Emx1-specified oligodendroglia which we refer to as 'single origin arrays'. Finally, since the mouse brain and the CC itself increase substantially in size during early postnatal development, it was important to take account of the fact that linear array formation occurs on a growing tissue. Specifically, the mean measured area of the CC in these experiments increased by approximately 60% in size predominately in the horizontal direction (it is 0.089mm^2 at P7, 0.147mm^2 at P14 and there was no significant difference between the P14 and P30
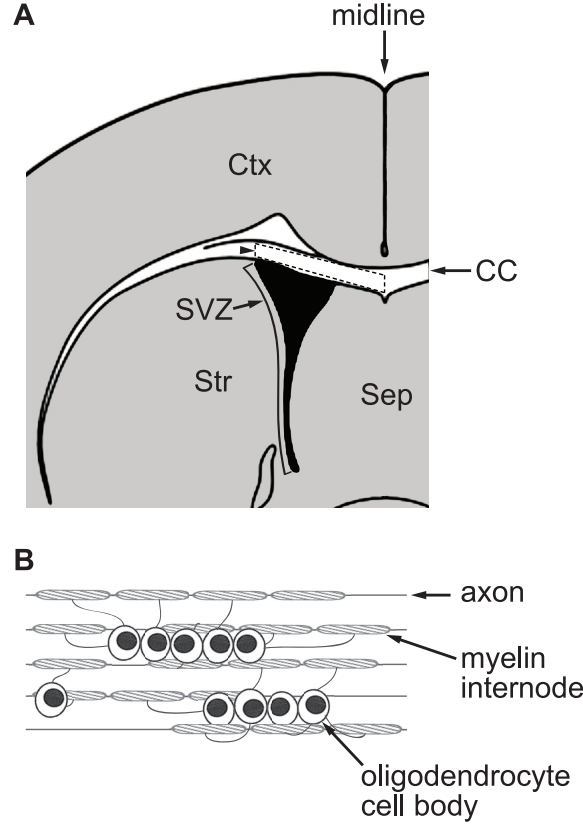
3

Figure 2: (A) Schematic view of a coronal slice through the adult mouse brain showing the region of interest (dashed rhombus) for analysis of linear array formation within the white matter of the corpus callosum. Black arrowhead indicates the direction of cell influx (from lateral to medial). CC, corpus callosum; Ctx, cerebral cortex; Sep, septum; Str, striatum; SVZ, subventricular zone. (B) Representation of 2 linear arrays of oligodendrocytes within a coronal section of the corpus callosum. Note the oligodendrocyte cell bodies align parallel to the axonal axis. Myelin internodes generated by oligodendrocytes are also illustrated. Depicted are two linear arrays comprising 4 or 5 oligodendrocytes and one isolated oligodendrocyte.

areas (Röth *et al.* 2016)). Our analysis focusses on the time period from P7 to P30 and our discrete stochastic model, described below, breaks this interval into 6 minute time-steps.

3. Discrete and continuous models: an overview

A number of cellular processes are known to be involved in the early development of the CNS. OPCs are known to move, proliferate, and differentiate to form OLs which myelinate axons (Mitew *et al.* 2014). We develop a discrete stochastic model that incorporates these processes. Despite the experimental observations discussed in Section 2, there is still considerable uncertainty about the mechanisms that lead to array formation. For example, can OPCs form arrays prior to maturation, and if so, are these arrays motile, or are only isolated OPCs motile? When two OPCs meet and contact, will they attach to each other to form nascent arrays? Most importantly, what combination of strengths (modeled through representative parameters) of migration, proliferation, differentiation and death lead to results consistent with the experimental observations? Acknowledging these uncertainties, our aim is to create a versatile simulation platform to test various assumptions and focus our attention on the general question of the extent to which (if any) cell migration or proliferation contributes to the formation of these array structures. To give further insight into the discrete stochastic model trends, we also consider a continuous deterministic model of agent array

formation based on the Becker–Döring model of coagulation (Becker and Döring 1934; Slemrod 2000). Note that the Becker–Döring model is usually referred to as a model of cluster formation. However, to be consistent with our analysis of linear arrays, we shall refer to it as a model of array formation.

In order to challenge these models, compare the efficacy of different assumptions, and determine the relative influence of different processes in the formation of these arrays, we utilize the following quantitative data for comparison with our simulation results:

R1: The number of arrays present at P14 and P30 (Röth *et al.* 2016).

R2: The number of cells observed to be in arrays at P14 and P30 (Röth *et al.* 2016).

R3: The total cell count at P14 and P30 (Röth *et al.* 2016).

R4: The clonal origin of cells in these arrays at P30(Röth *et al.* 2016).

R5: The observations that the majority of OPCs have started to differentiate by P14 and differentiation into CC1+ (mature) OLs is essentially complete by P30 (Röth *et al.* 2016), while approximately half of the newly generated cells die by P30 (Barres *et al.* 1992).

Our simulations begin at P7, using the experimental data at this time-point as our initial condition, and finish at P30.

3.1. Discrete stochastic model

A discrete-time agent-based simulation algorithm is developed, based on exclusion processes, to test ideas for the formation of linear arrays of OPCs and OLs. Each agent (cell) lies on a two-dimensional lattice with regularly spaced sites that can be occupied. A regular square lattice (site spacing the size of a cell) is used. Two agent types are introduced – OPC cell agents (two types with identical properties but labeled with a different anatomical origin, reflecting the ventral Gsh2 or Dorsal Emx1 origin, that does not influence cellular properties) and OL agents. A lattice site can only be occupied by at most one cellular agent at a time. This defines an exclusion process (Sander and Deisboeck 2002; Deutsch and Dormann 2005). The agents can perform cellular dynamic processes such as migration, attachment, proliferation, differentiation and death. These dynamic processes are governed by probabilities. In addition to these cellular processes, we account for the known growth in the area of the CC and also account for the continual migration of OPCs into the CC from other regions of the developing forebrain.

We seek to determine the relative importance of proliferation of OPCs versus migration and attachment as a means of building the observed linear arrays. To do so, we construct a sequence of three models that encode these mechanisms individually and in tandem.

1. Discrete stochastic model 1: OPC movement (of isolated cells) and attachment is included. In particular, OPC proliferation, differentiation and death, as well as an influx of OPCs, are not included in this model. Although this model is not biologically realistic, it isolates the influence of migration and attachment in array formation.
2. Discrete stochastic model 2: OPC proliferation, differentiation and death, as well as OPC influx are included. No OPC movement and attachment are included. In contrast to model 1, model 2 isolates the influence of OPC proliferation on array formation. Although this model is also not biologically realistic, it is still informative as it isolates the role of proliferation in creating predominantly single origin arrays. Influx was modelled by adding OPC cells (non-motile in this model) to the left third of the lattice (influx is lateral to medial, see the schematic in Figure 2 (A)).
3. Discrete stochastic model 3: This model combines all processes in models 1 and 2, to determine how well the combination of proliferation and migration can account for observations.

Using the data R1 – R5, we will determine which of these models best accounts for observations. To test the results of each of these *in silico* models, we measure the following outputs of each simulation: total cell number, array number, number of cells in arrays, distribution of array sizes and the composition of the arrays in terms of ventral Gsh2-specified versus dorsal Emx1-specified oligodendroglia. The model is simulated starting with empirical data collected from the CC of P7 C57Bl/6 mice as described in Röth *et al.* (2016) and evolved in time to P30, with model parameters determined from the literature where possible and fit to data (R1 – R5) where not.

3.2. Continuous deterministic model

To determine more rigorously the distribution of linear array lengths or array sizes that would result from a migration and attachment mechanism, we consider a variant of the Becker–Döring model of array coagulation (Becker and Döring 1934; Slemrod 2000), ignoring the array fragmentation terms. Array sizes take on discrete integer values corresponding to the number of cells in an array. To mimic the attachment of migrating isolated OPCs, which is the mechanism of array formation in the discrete stochastic model, we assume the system is well-mixed where only binary interactions involving single cells (array of size one) and larger arrays are considered. If we denote an array of size r by C_r then the process may be summarized as

$$C_r \oplus C_1 \to C_{r+1},$$

where $\oplus$ denotes the attachment of arrays due to random collisions. This matches a simplified version of discrete stochastic model 1 where cell densities are assumed to be homogeneous at all times.

The model assumes that arrays interact with each other according to the law of mass action (Murray 2002). The rate at which arrays react with one another is proportional to their concentrations. Furthermore, we can also include cell proliferation into this model by adding a term proportional to C_1. For simplicity, we have not accounted for the influx into the CC. The Becker–Döring equations (without fragmentation) with a source term due to cell proliferation are then

$$\frac{dC_1}{dt} = -2aC_1^2 - aC_1 \sum_{r=2}^{\infty} C_r + SC_1, \tag{1}$$

$$\frac{dC_r}{dt} = aC_1 C_{r-1} - aC_1 C_r \qquad (r \geq 2). \tag{2}$$

Here the constant a is the single cell/array interaction rate which has dimensions (time $\times$ concentration)$^{-1}$ and the constant S is a proliferation rate with dimension (time)$^{-1}$. Equation (2) says that the rate of change of the concentration of arrays of size r is equal to the rate of creation of arrays of size r from collisions between single cells and arrays of size $r-1$ minus the rate of depletion of arrays of size r from collisions between arrays of size r and single cells. Note that in the right-hand side of Eq. (1) the coefficient of C_1^2 is 2 because a joining of two single cells to form a array of size two removes two cells from the isolated cell population, whereas the joining of a single cell to a array of size $r \geq 2$ only reduces the single cell population by one. This system of equations is transformed (as described in Appendix 1) and the simplified version solved analytically, providing additional insight into the discrete stochastic model results.

4. Simulation rules and implementation

For simplicity we have not created a separate class of cells representing pre-myelinating OLs. Instead, we have divided the cell population into just two categories. These are OPCs (which are motile, proliferate, differentiate and can die during the modelling timeframe) and mature OLs (which are not motile and do not proliferate or die up to P30). When we speak of an OPC differentiating into an OL, we mean that the OPC has transformed into a mature OL. The rules and their implementation governing the dynamics of OPCs and the growing CC are now outlined. Their schematic representation is presented in Figure 3. The alignment of the array must be parallel to the axonal orientation, which here is taken to be horizontal as may be seen in Figure 2.

4.1. Cell movement and attachment

We consider a computational domain consisting of a square lattice. A site **s** on the square lattice has coordinates $\mathbf{s} = (i, j)$, where i, j are integers. In any time-step of duration τ, suppose there are n OPCs. A random selection of n OPCs is chosen, with replacement, for motility. On average each OPC will be chosen once per time-step. Since arrays are aligned horizontally (parallel to the axonal orientation in the CC, see Figure 2), we define an isolated OPC at site (i, j) as one where lattice site (i, j) is occupied by an OPC agent while its horizontally aligned neighbouring lattice sites $(i-1, j)$ and $(i+1, j)$ are empty. A chosen isolated OPC may move, with probability P_m, to one of the four sites $(i \pm 1, j \pm 1)$. If the randomly chosen site is occupied the move is aborted. The migration algorithm then repeats
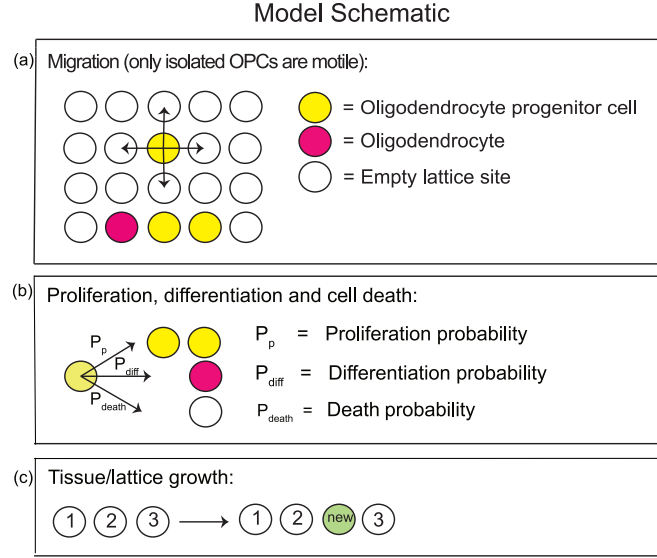
Figure 3: **Schematic of the discrete time agent based model of linear array formation in the mouse CC.** (a) Isolated OPCs are motile and can move once per time-step τ with probability P_m to one of four nearest neighbour sites. If an OPC collides with another cell, it attaches to that cell with probability P_a if the two cells are aligned with the horizontal axis. A linear array consisting of one OL and two OPCs is displayed. (b) An OPC in a time-step τ may proliferate with probability P_p (placing its daughter randomly on one of four nearest neighbour sites), differentiate with probability P_{diff} or die with probability P_{death}. (c) The area of the lattice grows to match the growth of the CC in the experiments of Röth *et al.* 2016.

this process for the next chosen OPC. This update scheme is known as a random sequential update (Chowdhury *et al.* 2005). Here we are assuming that only isolated OPCs are motile

If an OPC has the opportunity to move so that it is next to another cell, aligning them on a common horizontal axis, there is a probability P_a that it will attach to that neighboring cell and no longer be motile. Alternatively, there is a probability, $1-P_a$, that it will not attach, and in this case the move is aborted.

4.2. Proliferation and influx

There is an increase in OPC numbers through cell proliferation as well as through an influx of OPCs into the CC from the subventricular zone (Suzuki and Goldman 2003) and the neocortex (Kessaris *et al.* 2006), see Figure 2. During a time-step τ, on average each OPC attempts to proliferate with probability P_p. Again we use random sequential updating to choose the order of proliferation attempts. An OPC, at site **s** say, that has been chosen to proliferate, inspects its four nearest neighbour lattice sites for empty sites. If more than one site is vacant, one of the sites is chosen at random: call this site **v**. One daughter cell remains at the original site **s** whilst the other daughter cell now occupies **v**. This can be called a 'myopic proliferation' rule, since it is analogous to a myopic motility rule (Landman and Fernando 2011; Hackett-Jones *et al.* 2011). The influx of OPCs into the CC is modeled by adding OPCs to randomly selected empty sites, as described in Section 4.6.

4.3. Differentiation and death

Cell differentiation and death are treated as instantaneous processes, although biologically they are multi-staged processes. During a time step τ, any OPC agent may differentiate with probability P_{diff} into an OL agent. Once a cell has differentiated it is no longer motile, nor can it proliferate, differentiate or die. It remains on the lattice site where it was 'born'. However, a large proportion of pre-myelinating OLs die before they can differentiate into OLs (Barres *et al.* 1992; Trapp *et al.* 1997) which for simplicity we indicate here as death of OPCs. During a time-step τ, OPCs may die with probability P_{death}. Note that an attached OPC may still proliferate or die whilst the population of mature OLs is non-dynamic. Thus if differentiation of cells were not implemented too many cells would be created.

7

4.4. Domain growth

The mean measured area of the CC increases primarily in the horizontal direction, defining a uniaxially growing domain. Suppose our lattice is of size $L_x(t) \times L_y$, where the horizontal distance grows with time t. We model this growth by randomly choosing times for the lattice to grow by one space unit in the horizontal direction, such that the lattice has the required area at P14 and at P30. At such a time, we randomly choose an empty site, (i, j) say, and insert an unoccupied site at $(i + 1, j)$ pushing all sites to the right of $(i + 1, j)$ one lattice space to the right. This process is repeated for each row until we have a rectangular lattice of size $(L_x + 1) \times L_y$. We then repeat the whole process at the remaining growth times. By randomly choosing an empty site we ensure that no arrays are broken up by the growth. This process is related to those described in detail in Binder and Landman (2009) and Hackett-Jones et al. (2011) except that we require that arrays are not broken up by the growth.

4.5. Initial and boundary conditions

We require initial and boundary conditions. Initially at P7, the mean number of OPCs, OLs and the size and number of arrays are known (Röth et al. 2016). We choose our initial conditions to be consistent with this data. Our lattice is randomly seeded with 104 cells, made up of 94 OPCs and 10 OLs (Röth et al. 2016), and we assume ventral Gsh2-specified oligodendroglia and dorsal Emx1-specified oligodendroglia each represent 50% of the total oligodendroglial cell population in the CC (in Kessaris et al. (2009) this was found to be the case at P10). The mean experimental data at P7 tells us that there are approximately 4 arrays present at P7. We place two arrays, of length three cells, composed of cells exclusively from species one ('pure' single origin arrays) at random locations and repeat for two three-arrays composed of cells from species two. No OPCs or OLs are allowed to leave the CC through migration.

4.6. Implementation

At each time-step, each of the operations of domain growth, agent movement, proliferation and differentiation and death occurs in turn. Sequential independent random choices of cells (equal to the number of cells at that time step) are made for each operation as described above. We describe how the time-step of the simulations and the probability parameters are estimated in the next subsection.

4.7. Parameter values

We choose a diffusion rate typical of biological cells, as $D = 2.66\,\mu\mathrm{m}^2/\mathrm{min}$ (Cai et al. 2007; Johnston et al. 2015), a single lattice spacing Δ to represent a cell diameter (8 μm) and the probability of movement $\mathrm{P_m} = 1$. Then the time step τ can be determined using (Simpson et al. 2009)

$$D = \frac{\mathrm{P_m}\,\Delta^2}{4\tau}.$$

This provides an estimate of the time-step as $\tau = 6$ minutes.

Psachoulia et al. (2009) have determined the mean cell cycle time at P6 is 1.7 days increasing to 9 days by P60. Young et al. (2013) determined the cell cycle time to be 2.7 days at P21 and 9.5 days at P60. Röth et al. (2016) have found that expression of the Ki67 proliferation marker was low at P14 with only 5% of the entire OPC population identified to express Ki67, whereas 27% of OPCs expressed Ki67 at P7. This means that the proportion of cells in the cell cycle reduces with postnatal age. Combining these pieces of data, we choose the mean cycle time T_p (often denoted as T_C in the biological literature) as shown in Figure 4.

We set the mean time to differentiate $\mathrm{T_{diff}}$ equal to the cell cycle time since it was observed in Psachoulia et al. (2009) that the differentiation rate roughly parallels the cell division rate. Other data indicating that the time for OPCs to differentiate increases with postnatal age can be found in Kang et al. (2010) and Zhu et al. (2011). Since oligodendroglial cell death principally occurs at the pre-myelinating OL stage, reflecting the transition between OPC and mature OL, we assume that the time for cells to die increases roughly as a function of the time for OPCs to differentiate. We choose the mean time to die to be proportional to the differentiation time and perform a parameter search for the proportionality constant subject to the constraint that approximately 50% of newly generated cells die up to P30 (R5).
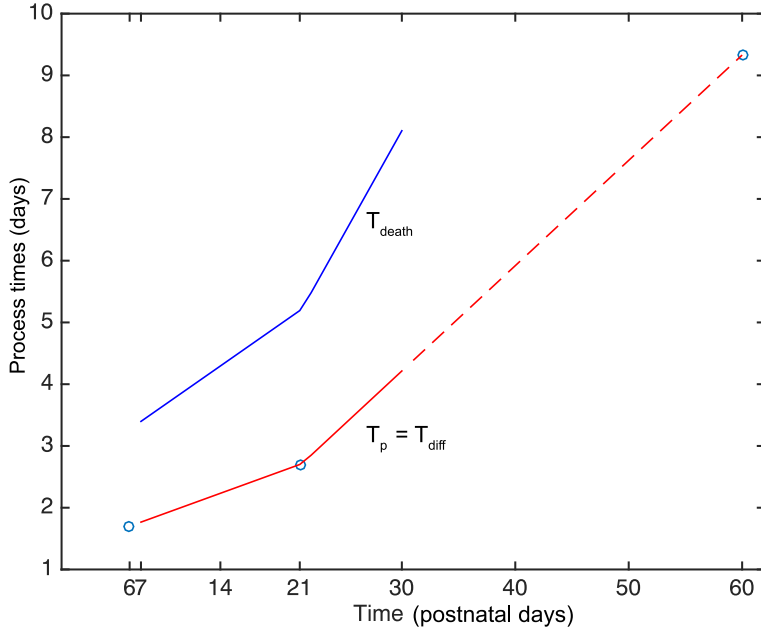
Figure 4: **Process times, namely cycle time T$_p$, mean time to differentiate T$_{diff}$ and mean time to death T$_{death}$.** Here T$_p$ is 1.7 days, at P6 and increases linearly to 2.7 days at P21 and then increases linearly to 9.5 days at P60 (Psachoulia *et al.* 2009; Young *et al.* 2013). Data points are denoted by circles. The dashed line is to highlight that the experimental observations at P6 and P60 lie outside the modelling timeframe. T$_{diff}$ is chosen to be equal to T$_p$ (Psachoulia *et al.* (2009) noted that the differentiation rate roughly paralleled the cell division rate) and the mean time for an OPC to die is chosen to be proportional to T$_{diff}$ (hence also T$_p$) with a constant of proportionality of 1.9 determined by a parameter search and R5. The percentage of cells that die up to P30 is taken from Barres *et al.* (1992).

Our parameter choices are illustrated in Figure 4. Note that the biological literature refers to the cell cycle time, (Young *et al.* 2013), which we must convert to a probability for implementation in our discrete stochastic model. The probabilities used in the simulations associated with proliferation, differentiation and death are each determined in the following way (Simpson *et al.* 2007): P$_p$ = τ/T_p, P$_{diff}$ = τ/T_{diff} and P$_{death}$ = τ/T_{death}, where τ = 1/240 days.

The influx of OPCs into the CC is modeled by adding OPCs to randomly selected empty sites, in the left-hand third of the lattice bounded by the left boundary (since this is the region closest to the subventricular zone and neocortex, see Figure 2). The quantitative value of this influx is unknown from experimental data. The value chosen is governed by matching the known total cell count R3. Every 240 time steps (equal to one day), two OPCs, one of each species, are added to empty sites in this portion of the lattice.

The mean measured area of the CC at times P7 and P14 was 0.089mm^2 and 0.147mm^2 respectively. The photomicrographs of tissue sections used to quantify linear array counts (Röth *et al.* (2016) were of a fixed size within which the mean area of the CC was quantified. Reanalysis of the measured areas used to derive array counts at P14 and P30 revealed no significant differences between these two time-points and this is our justification for keeping the lattice size fixed between these times. Since a lattice spacing is Δ = 8 μm, a rectangular lattice with dimensions 70 $\times$ 20 corresponds to the initial measured area at P7 and grows to 115 $\times$ 20 at P14 and remains at this value until P30.

The attachment probability P$_a$ is first set to unity for Model 1. It is then reduced to 0.02 in Model 1 in order to suppress excess array formation (R1 and R2) and to match the biological data at P14. In Model 2, the attachment probability is not relevant. In Model 3, the value of P$_a$ is again reduced to 0.01 in order to further suppress the number of arrays and to match the observed data.
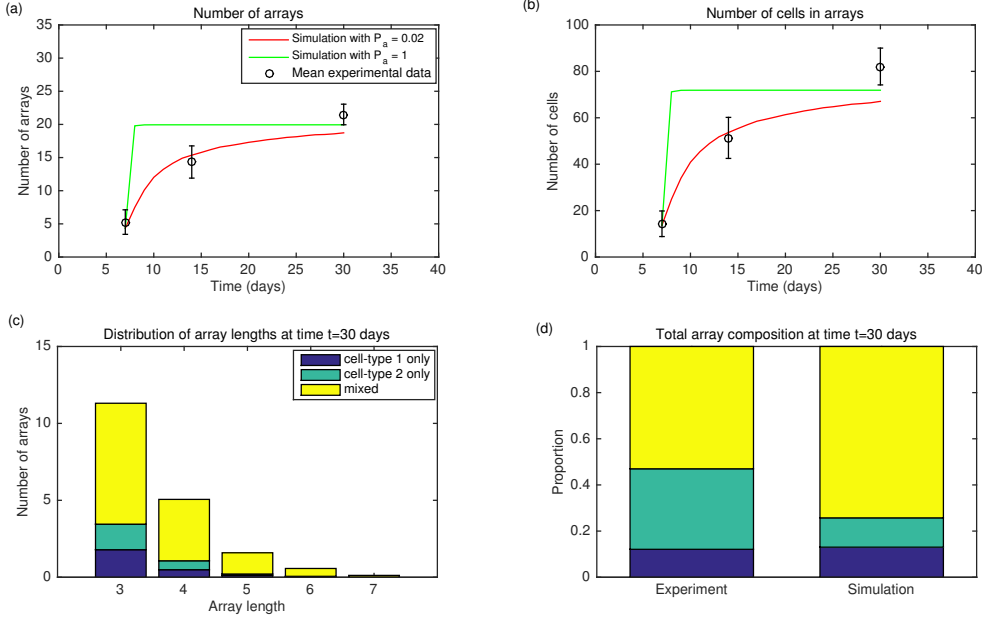
9

Figure 5: **Discrete stochastic model 1: Migration stimulates mixed array formation.** Simulation results with migration and attachment only are averaged over 100 realizations. (a), (b) Number of arrays and number of cells in arrays starting from P7 and evolving to P30. With probability of attachment $P_a = 1$, there is rapid growth in number of arrays with all cells attached by P14 and no array formation henceforth. There is poor agreement with experimental data at P14 but good agreement at P30. When $P_a = 0.02$, there is excellent agreement with experimental data at P14 but poor agreement at P30. Error bars correspond to one standard deviation above and below the mean value. (c), (d) Array length distribution and composition at simulation time corresponding to P30 with $P_a = 0.02$. (Note: The results at P30 are virtually insensitive to the value P_a). Arrays of length 6+ are rare and approximately 75% of all arrays are mixed origin arrays.

5. Results

The results from the three discrete stochastic models are presented and compared to data collected from animal studies. To complement the discrete model results, the results of the continuous deterministic model of array formation are discussed. For comparison purposes, we present results on the number of arrays, number of cells in arrays, total cell count (all as a function of time), distribution of array lengths, and composition of arrays (single or mixed origin), at P30. The model results are compared to experimental data at the observation times P7, P14 and P30. The data presented shows the best fit results after a parameter search over all parameters not constrained from literature, as discussed in Section 4.7.

5.1. Discrete stochastic model 1: Migration stimulates mixed origin array formation

In this model, the dynamic processes present are migration and attachment only. There is no increase in cell number, as cell proliferation and influx terms are both neglected. The purpose of this model is to determine if migration and attachment alone can generate the experimentally observed arrays in a manner consistent with all the data. While migration, which we assume to be random in space, is well defined, the process of attachment is not. We thus consider two possibilities. We first discuss the case where cells that come into contact are certain to attach to each other ($P_a = 1$). Figure 5 compares the results of this model with observations.

Figure 5(a), (b) shows the number of arrays and number of cells in arrays. The total cell count is a constant, equal to 104. The simulation results in a rapid increase and comes to a steady state and few isolated (motile) cells remain after P14. Figure 5(c)–(d) illustrates the distribution of array sizes and proportions of single origin versus mixed origin arrays. Comparison with experimental results shows this model overestimates the proportion of mixed origin arrays. Therefore, this model does not replicate the gradual increase in the number of arrays and number of cells in
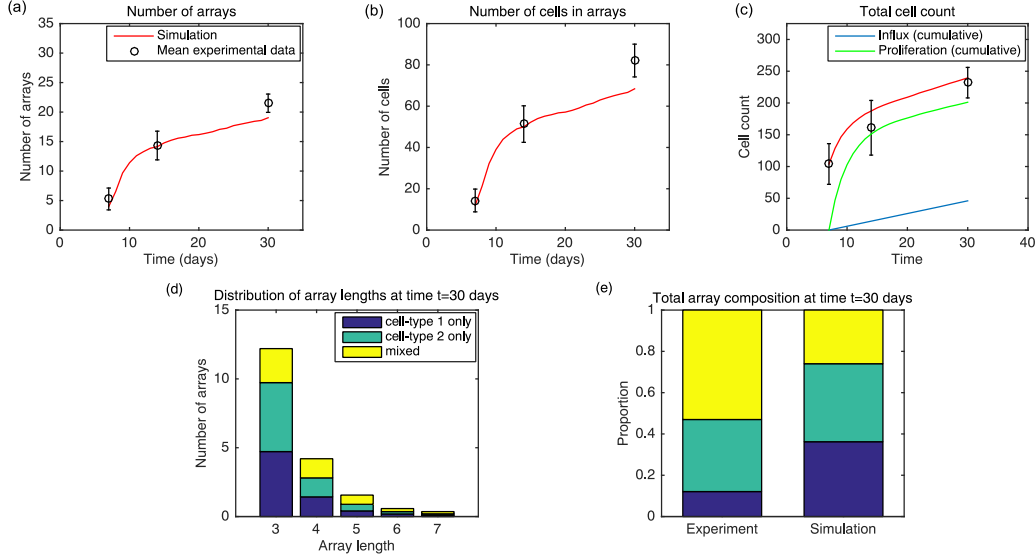
Figure 6: **Discrete stochastic model 2: Proliferation does not produce sufficiently mixed arrays.** Simulation results with proliferation, differentiation, cell death and cell influx, averaged over 100 realizations. Proliferation, differentiation and death are governed by values in Figure 4. No migration and attachment parameters are required. (a)–(c) Number of arrays, number of cells in arrays and total cell count starting from P7 and evolving to P30. There is good agreement with mean experimental data for the total cell count, but an insufficient number of arrays are formed at P30. Error bars correspond to one standard deviation above and below the mean value. (d)–(e) Array length distribution and composition at simulation time corresponding to P30. Now only 25% of arrays are mixed, indicating that without migration an insufficient number of mixed arrays are formed.

arrays. However there is reasonable agreement with experimental data by P30. We next reduced the probability that a random collision of cells leads to attachment (reducing $P_a = 0.02$). This lowers the rate of array formation, giving good agreement with experimental data at P14. However the rate of array formation is too low to account for data at P30. These results demonstrate that a migration and attachment mechanism alone cannot account for the full time course of array formation. In model 3, we will reconsider the influence of this mechanism in conjunction with OPC proliferation and influx.

5.2. Discrete stochastic model 2: Proliferation does not produce sufficient mixed origin arrays

An alternative hypothesis is that proliferation of individual cells may generate arrays. The idea here is that a single OPC will colonize a region of an axon and subsequent divisions of that OPC and its progeny (provided they do not move) generate an array. Model 2 considers the extreme case where cell motility is absent. The cell numbers increase through cell proliferation as well as a small steady influx into the domain. We include cell differentiation and death in this model to depress the growth in cell number due to proliferation and influx.

Results in Figure 6 indicate that whilst both the number of arrays and the number of cells in arrays agree with experimental data at P14, the agreement is not as good at P30, whereas the total cell count matches experimental results at P14 and P30. At early times, when the cell cycle time is relatively short, arrays are formed as OPCs divide. At later times, the cell cycle time has increased leading to fewer cell divisions and less frequent array formation.

Intuitively proliferation (without influx and migration) will give rise to single origin arrays that are clonally related. In this case, mixed origin arrays are rare and arise only when two separated cells or arrays, which are single origin and of different origin (one being Gsh2-specified, the other Emx1-specified), merge due to the elongation of arrays that occurs when cells proliferate along a common horizontal axis. If influx is included as it is here, Figure 6(d)–(e) shows that approximately 25% of arrays are of mixed type (compared to 75% of arrays being mixed type in the migration only model Figure 5(d)). Mixed origin arrays also arise if a randomly added cell through the influx process is of a different type to its nearest neighbor.
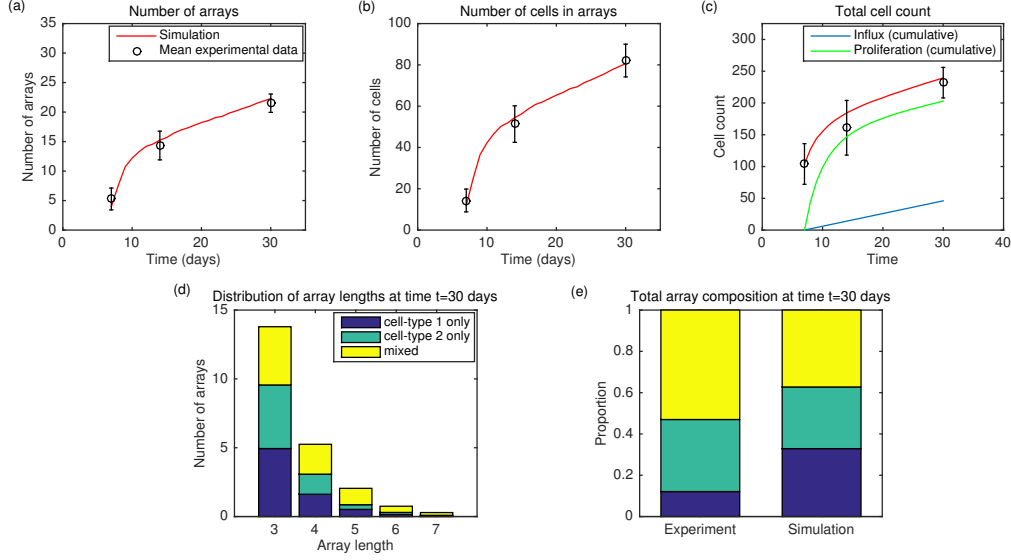
11

Figure 7: **Discrete stochastic model 3: Both migration and proliferation are required to satisfy all R1 – R5.** Simulation results with migration, attachment, proliferation, differentiation, cell death and cell influx, averaged over 100 realizations. Proliferation, differentiation and death are governed by values in Figure 4 and $P_a = 0.01$. (a)–(c) Number of arrays, number of cells in arrays and total cell count starting from P7 and evolving to P30. The simulation results provide an excellent fit to all three at P14 and P30. Error bars correspond to one standard deviation above and below the mean value. (d)–(e) Array length distribution and composition at simulation time corresponding to P30. The proportion of mixed origin arrays is close to that observed experimentally.

5.3. Discrete stochastic model 3: Both migration and proliferation are required to satisfy all observed data

Here the two previous models are combined and all the five key dynamic processes of OPCs are included, as well as a moderate cell influx. We can now fit all the data R1 – R5 well. In particular the evolution of the number of arrays, number of cells in arrays as well as the total cell count accurately match the experimental data at P14 and P30, as shown in Figure 7(a)–(c). Furthermore the array composition predicts a large number of mixed origin arrays at P30 (Figure 7(e)). The detailed distribution as determined from the model is given in Figure 7(d). All the observations R1 – R5 are met with this model. From the results of the restricted simulation models 1 and 2, migration and attachment naturally creates arrays, as does proliferation without migration (provided the cell cycle time is sufficiently short as is the case at early times). Only in a model with significant migration do we get close to the observed number of mixed origin arrays. Increasing the probability of attachment increases the percentage of mixed origin arrays but also increases the number of arrays and number of cells in arrays.

A feature of our array composition results from the three discrete stochastic models is that single origin ('pure') arrays and cell-type counts are in equal abundance at P30. The relative abundance at P30 will effectively mimic the relative abundance at P7 because proliferation dominates greatly over influx and our model treats these cells identically. In the discrete stochastic models above we assumed an equal abundance of these Gsh2-specified and Emx1-specified cells at P7 and also that the influx of cells was equally divided between the two cell types. Röth *et al.* (2016) observed these cell populations at P30, with 36% Gsh2-specified and 64% Emx1-specified. Using instead these proportions as our initial condition, and keeping the influx cells equally divided between the two cell types, we found that in the revised simulation models 2 and 3 that the final cell type proportions were 40% and 60% and with similar proportions of single origin arrays at P30.

5.4. Continuous deterministic model

First consider the deterministic coagulation model without the cell proliferation term ($S = 0$) in Eq. (1). This corresponds to the discrete stochastic model 1 where the only dynamic process is migration and attachment. This
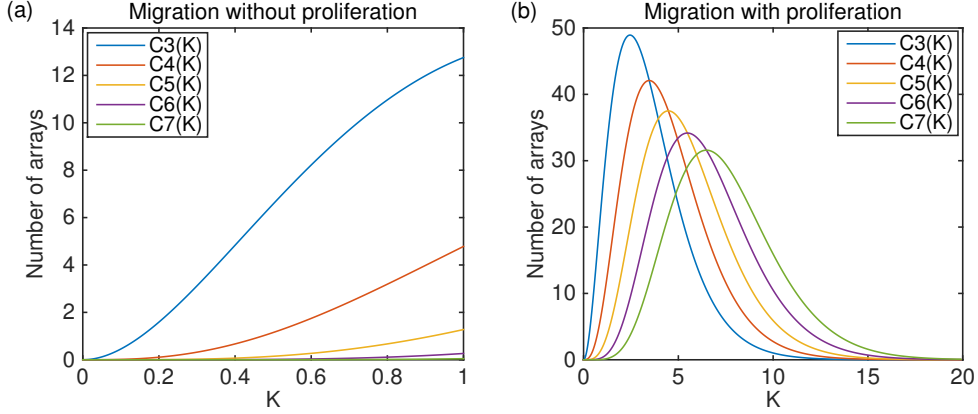
12

Figure 8: **Array formation derived from a variant of the Becker–Döring equations without fragmentation.** We choose $Q = 104$ (to match the number of cells at P7 in the discrete models) and $a = 1$. (a) No proliferation. (b) With proliferation. When we include sufficient cell proliferation (we choose $S = 2aQ$), the distribution of array sizes is now very different from (a): a transfer of cells to arrays of larger size occurs, eventually creating a array of infinite size.

system can be solved exactly in terms of a new time variable K, explained in the appendix, defined by

$$K(t) = a \int_0^t C_1(\tau) d\tau. \tag{3}$$

To enable us to distinguish between dependence of the array size distribution C_r on ordinary time t or on the modified time K we use the notations $C_r(t)$ and $C_r[K]$, respectively.

For an initial condition where there are only Q single cells and no larger arrays, C_1 and C_r can be determined exactly in terms of K as

$$C_1[K] = Q(1 - K)e^{-K}, \qquad C_r[K] = Q\left(\frac{K^{r-1}}{(r-1)!} - \frac{K^r}{r!}\right)e^{-K}. \tag{4}$$

Of course we would like to express the solutions in terms of our original time variable t rather than $K(t)$. From the definition of K, this transformation is defined by

$$t = \int_0^K \frac{dK'}{C_1[K']} = \int_0^K \frac{e^{K'} dK'}{Q(1 - K')}. \tag{5}$$

We see that t diverges as $K \to 1$, and therefore we obtain the limiting array distribution as

$$\lim_{t \to \infty} C_r(t) = \frac{Q}{r!e}(r - 1). \tag{6}$$

From this result, we deduce that there are no single cells (that is arrays of size $r = 1$) as $t \to \infty$. This distribution of array sizes is in good agreement with discrete stochastic model 1 (Figure 5) at P30, by which time no isolated motile cells remain. Furthermore, the number of arrays at $K = 1$ in this model is 19.1 and the number of cells in arrays is 65.7 (in close agreement with the results from discrete stochastic model 1). This simple continuous deterministic model replicates the late-time simulation results.

Next consider the case when there is single cell proliferation ($S > 0$), but no cell influx. This corresponds most closely to discrete stochastic model 3, since both migration, attachment and proliferation are present.

There is a critical proliferation rate $S = aQ$. For the subcritical case ($S < aQ$) the reproduction rate S is sufficiently small that the addition of reproduction to the coagulation model does not alter the array distribution qualitatively. We still obtain a steady state (at finite K-time) with no motile single cells at $t = \infty$, as in simulation model 1. For the supercritical case ($S > aQ$), the evolution of the array size distribution is now very different. All array sizes of

13

size three or larger display a turning point because the mass of the system is steadily transferred to arrays of greater size, with an infinite array in the limit as $K \rightarrow \infty$. One such distribution of array sizes for the choice of $S = 2aQ$ is illustrated in Figure 8(b). This should be compared to discrete stochastic model 3 results in Figure 7, where differentiation and death play an important role. Since we do not have a steady-state limiting distribution, it is more difficult to make the direct comparisons as could be made with the migration only model. However, this simplified analytic model does give insight into array formation when it is not suppressed by differentiation, cell death and the tendency of the cell cycle time to increase with postnatal age (Psachoulia *et al.* 2009; Kang *et al.* 2010; Zhu *et al.* 2011; Young *et al.* 2013).

6. Sensitivity analysis

We now outline the role of the parameters in determining the simulation results. The parameter we have most detailed data for is the cell cycle time/proliferation rate (Psachoulia *et al.* 2009; Young *et al.* 2013), so this is held fixed and the other parameters are varied. Recall that Psachoulia *et al.* (2009) found that the mean time for a cell to differentiate was approximately equal to the cell cycle time. Our death rate was chosen to agree with Psachoulia *et al.* (2009) and our migration rate was chosen to agree with Johnston *et al.* (2015) and Cai *et al.* (2007). The attachment probability we used was determined by a parameter search to fit to the experimental data of Röth *et al.* 2016. Using the parameter values used in discrete stochastic model 3 as our baseline, we compare the outcomes of varying parameter combinations (differentiation, death or attachment), looking for combinations that achieve a good fit to the experimental data.

If we increase the differentiation rate fewer cells will be created (since fewer OPCs will be actively proliferating), so to achieve results comparable to the baseline we would need to decrease the death rate. With P_{diff} increased to $1.2\times$ P_p (so fewer cells will be created) and P_{death} decreased to $0.45\times P_p$, a very good fit to all experimental observations is obtained. Alternatively, decreasing P_{diff} to $0.8\times P_p$ whilst increasing P_{death} to $0.6\times P_p$, we find that just under 40% of arrays are mixed and too many cells have died. There is a good fit to all the other experimental data. Similar conclusions may be drawn from the scenario with P_{diff} decreased to $0.9\times P_p$ and P_{death} increased to $0.55\times P_p$.

Increasing the attachment probability will create more arrays and more cells in arrays, so to achieve a good fit to the experimental observations we would need to either increase the probability that a cell will die or decrease the probability that a cell will be created in a particular timestep. We outline the results for varying the attachment probability and the probability of death first.

Increasing the probability of attachment to 0.02 (from 0.01 in the baseline) and increasing the likelihood that an OPC will die to $P_{death} = 0.6\times P_p$, it is found that the proportion of newly created cells that die is too high (all other outputs are within 1 standard deviation of the mean experimental observation). If we decrease the probability of attachment to 0.005 and decrease P_{death} to $0.45\times P_p$ there are too many cells in arrays, too many arrays and the final cell count is too high. The proportion of arrays that are mixed drops to approximately 32 %.

The second possibility to achieve results in agreement with the experimental observations, having increased the probability of attachment, is decreasing the probability of a cell being created or, since we hold the probability of proliferation fixed, increasing the probability that a cell will differentiate. If we increase the probability of attachment to 0.02 (from 0.01 in the baseline) and increase the probability of differentiation to $1.2\times P_p$ (from $1\times P_p$ in the baseline), then we achieve a very good fit to all experimental observations. The proportion of mixed origin arrays is approximately 45%.

Decreasing the probability of attachment to 0.005 (from 0.01 in the baseline) and reducing the probability of differentiation to $0.8\times P_p$ (from $1\times P_p$ in the baseline), so more cells will be created, it is found that there are too many cells in arrays, too many arrays and the final cell count is too high. Only approximately 32% of arrays are of mixed origin. Keeping this value of the differentiation probability and reducing the probability of attachment further to 0.001 improves the fit to the experimental observations except that the cell count is too high and mixing is now at only approximately 30%.

From this sensitivity analysis we see that the attachment parameter is important not just to array formation but also in creating mixed origin arrays since it determines the contribution of migration to array formation. A number of different parameter combinations are possible which give results in good agreement with the experimental data. These models have a low attachment parameter value (0.01 or 0.02). If a higher attachment probability is chosen there will

	R1		R2		R3		R4	R5
	P14	P30	P14	P30	P14	P30	P30	
Discrete stochastic model 1 ($P_a = 1$)	×	✓	×	✓	×	×	×	×
Discrete stochastic model 1 ($P_a = 0.02$)	✓	×	✓	×	×	×	×	×
Discrete stochastic model 2	✓	×	✓	×	✓	✓	×	✓
Discrete stochastic model 3 ($P_a = 0.01$)	✓	✓	✓	✓	✓	✓	✓	✓

Table 1: **Summary of the results of the discrete stochastic model and its agreement with the five biological experimental results R1–R5**. A ✓ denotes that the result is satisfied in this model whereas a × denotes that it is not satisfied. We required only a reasonable fit of within 15% of the experimentally observed value for R4 since we have no experimental data on the initial proportions of each OPC cell type at P7. Simulation model 3 is the only model to satisfy all five experimental results.

be an abundance of mixed arrays. If it is chosen lower then most arrays will be formed by proliferation rather than migration and attachment and the proportion of mixed origin arrays will be too low. From discrete stochastic model 1 we know that migration naturally creates mixed arrays. The overproduction of mixed arrays in discrete stochastic model 1 is tempered by the addition of proliferation (which creates predominantly single origin arrays) that occurs in discrete stochastic model 3.

7. Discussion

We have assessed the importance of migration and proliferation to OL linear array formation by comparing the results of an agent-based cellular automata algorithm to mean experimental data, as well as presented some complimentary analytic models. The biologically reasonable migration mechanism, where only isolated cells are motile, naturally forms linear arrays. Whereas proliferation (without migration or influx) elongates clonally derived arrays of single origin and occasionally causes clonally derived arrays of different origin to merge, migration leads to predominantly mixed origin arrays without any need to assume some attraction mechanism between cells of different anatomic origins.

In Table 1 we summarize the results of our discrete stochastic models to show which models agree with the experimental observations. Clearly, discrete stochastic model 3 is the only model to meet all five constraints, and therefore is the most plausible and convincing. Moreover, it is also the most biologically reasonable since all known dynamic processes of migration, proliferation, differentiation and death are included.

Adding even a small amount of migration leads to a significant increase in the number of arrays formed, number of cells in arrays and proportion of arrays that are mixed. Since we know that OPCs are motile, in order to fit our model to the experimental data it was necessary to introduce another biologically reasonable parameter, the small attachment probability P_a, which suppressed array formation.

The migration only model (discrete stochastic model 1) creates too many mixed origin arrays whilst the proliferation dominated model (discrete stochastic model 2) creates too few mixed arrays. It should be emphasised that these models are structurally different with regards to how arrays are formed and even though our combined model, discrete stochastic model 3, still does not quite create the precise proportions of mixed origin arrays observed experimentally, it does tell us that migration is the dominant process. The effect of migration on array formation is so strong that even when cell collisions have only a 1% chance of resulting in attachment, approximately twice as many mixed arrays are formed than through proliferation alone.

The migration model used allowed only isolated OPCs to move. Other assumptions were trialled, such as allowing the terminal cell on an array (that is, one on the end) to move or allowing all OPCs to move, but none of these provided satisfactory agreement with experimental results. Arrays in these migration models tended to break up before the cells differentiated. We also implemented a migration model where all OPCs could move and attachment only occurred when an OPC was adjacent to a differentiated cell. However, this model produces an over-abundance of mixed arrays since many more OPCs are motile compared to the model we presented in Section 4.1 where only isolated cells were motile.

In the discrete stochastic models 1 and 3 we chose a constant attachment probability. It is possible that cell adhesiveness is a dynamic property of the system. Notice in Figure 5 (a) that few new arrays are created after P14. The reason for this is that few motile cells remain after P14 as the system approaches a steady state. To achieve a

15

good fit to the experimental data with migration and attachment as the only dynamic processes, it is necessary to add cell influx and make the attachment probability dynamic. If the attachment probability is 0.01 from P7 to P14 and 0.001 thereafter and there is a constant influx of OPCs into the domain (two cells migrate into the domain every one hundred time-steps) then we obtain a good fit to all the experimental observations except that there is an over abundance of mixed origin arrays (about 80%). Since no proliferation, differentiation or cell death is present in this model it is biologically implausible but it once again reinforces our conclusion that migration naturally creates mixed origin arrays.

The continuous deterministic model presented sheds further light on the dynamics of linear array formation. It provides a complementary picture by ignoring spatial effects and focusing instead on concentrations in a well-mixed system evolving according to the law of mass action. With no cell proliferation, all cells are attached at late times. A steady state has been reached and the distribution of array sizes at late times in discrete stochastic model 1 agrees with the analytic formula. Adding sufficient reproduction changes the situation dramatically. With the increased number of isolated cells (C_1), array numbers increase to a maximum value before turning and decreasing towards zero as the mass of the system is shifted towards larger array sizes.

The modelling of linear array formation described in this paper provides insight into the cellular dynamics that could be responsible for the formation of linear arrays during postnatal development. The insights that we have gained provide the opportunity for validating these predictions against real-time video microscopy of linear array formation in living tissue slices. Furthermore, it provides the opportunity to explore the influence of genetic or environmental manipulations that modulate one or more of the parameters that we have defined. Since linear arrays are observed in the CNS white matter of most, if not all vertebrates, the principles that we have defined are also likely to be relevant across white matter tracts of most vertebrate species, including humans. Further refinement and experimental validation of our model will help determine which parameters to experimentally manipulate in order to optimise the regeneration of linear arrays following OL death, as occurs in diseases such as multiple sclerosis.

This work is a first step towards modelling the cellular dynamics of oligodendroglia that could underlie the generation of linear arrays during postnatal myelination of a major white matter tract. OLs are the cells responsible for myelinating the axons of neurons in the CNS and this work has shown that random dynamic processes can account for the observed spatial organization of OLs in the mouse CC. Future research will be directed at understanding how myelin internodes pattern the axons in the vicinity of the myelin-producing OLs within linear arrays. This will provide further insight into how myelination is co-ordinated at the population level and the functional significance of linear array structures for regulating electrical conduction within white matter tracts.

Acknowledgments

Appendix: Becker–Döring equations without fragmentation

We first consider the equations (1)–(2) without a source term $S = 0$, so that the mass

$$M(t) = \sum_{r=1}^{\infty} rC_r(t) \tag{7}$$

is constant. Equations (1)–(2) form a coupled system of nonlinear ordinary differential equations. This system may be greatly simplified if we introduce new functions (Hendriks and Ernst 1984)

$$K(t) = a \int_0^t C_1(\tau)d\tau, \tag{8}$$

and

$$L(t) = \sum_{r=1}^{\infty} C_r(t). \tag{9}$$

16

We shall use K as a new time variable. To enable us to distinguish between dependence of the array size distribution C_r on ordinary time t or on the modified time K, we use the notations $C_r(t)$ and $C_r[K]$, respectively. Dividing Eq. (1) (with $S = 0$) and Eq. (2) by aC_1 we obtain the simplified system (in terms of K-time variable)

$$\frac{dC_1}{dK} = -C_1 - L, \tag{10}$$

$$\frac{dC_r}{dK} = C_{r-1} - C_r \qquad (r \geq 2). \tag{11}$$

We can simplify the system further by noting that

$$\frac{dL}{dK} = \sum_{r=1}^{\infty} \frac{dC_r}{dK} = \frac{dC_1}{dK} + \sum_{r=2}^{\infty} \frac{dC_r}{dK} = -L. \tag{12}$$

Since $K = 0$ corresponds to $t = 0$, we find that $L(t) = L(0)e^{-K}$. If our initial condition is that all arrays are initially single cells, then we can write it in terms of a Kronecker delta function as $C_r(0) = Q\delta_{r,1}$. Then $L(t) = Qe^{-K}$, and we can determine C_1 as a function of K by solving the initial value problem

$$\frac{dC_1}{dK} + C_1 = -Qe^{-K}, \tag{13}$$

where $C_1 = Q$ when $K = 0$. The solution is

$$C_1[K] = Q(1 - K)e^{-K}. \tag{14}$$

From Eq. (11), all array sizes r can be determined as

$$C_r[K] = Q\left(\frac{K^{r-1}}{(r-1)!} - \frac{K^r}{r!}\right)e^{-K}. \tag{15}$$

Of course we would like to express this in terms of our original variable t rather than $K(t)$. From the definition of K, this transformation is defined by

$$t = \int_0^K \frac{dK'}{aC_1[K']} = \int_0^K \frac{e^{K'}\,dK'}{aQ(1-K')}, \tag{16}$$

so that, on writing $x = 1 - K'$ we have

$$\frac{aQt}{e} = \int_0^K \frac{e^{K'-1}\,dK'}{1-K'} = \int_{1-K}^1 \frac{e^{-x}dx}{x} = \int_{1-K}^1 \frac{dx}{x} - \int_{1-K}^1 \frac{(1-e^{-x})dx}{x}$$

$$= -\ln(1-K) - \int_0^1 \frac{(1-e^{-x})dx}{x} + \int_0^{1-K} \frac{(1-e^{-x})dx}{x}$$

$$= -\ln(1-K) - \beta + \sum_{n=1}^{\infty} \frac{(-1)^{n-1}(1-K)^n}{n \cdot n!}, \tag{17}$$

where

$$\beta = \int_0^1 \frac{(1-e^{-x})dx}{x} \approx 0.796\,599\,599\,297 \tag{18}$$

(numerical evaluation in MATHEMATICA of an expression for this integral in terms of Euler's constant and the exponential integral). We see that t diverges as $K \to 1$, and therefore we obtain the limiting array distribution

$$\lim_{t\to\infty} C_r(t) = \frac{Q}{r!e}(r-1). \tag{19}$$

We see from Eq. (17) that we have $1 - K \sim \exp(-aQt/e - \beta)$ in the $K \to 1$ (and correspondingly $t \to \infty$) limit. For the limit $K \to 0$ (and so $t \to 0$), we can expand the integrand in Eq. (16) in a Taylor series, giving $aQt = K + K^2 + \cdots$. In Figure 9 we show the exact relation between t and K, together with excellent simple analytic approximations for small K or small $1 - K$.
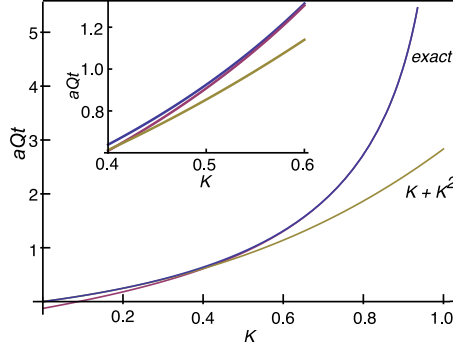
17

Figure 9: **Time coordinates for the zero proliferation case.** Here K tends to a finite value as t tends to infinity. The exact relation between t and K is shown in blue. The small-K approximation $aQt \approx K + K^2$ (shown in brown) is excellent for $K < 0.3$. The alternative approximation $aQt/e \approx -\ln(1-K) - \beta + 1 - K - (1-K)^2/4$, shown in red is indistinguishable from the exact (blue) curve for $K \geq 0.5$. The inset shows the transition region around $K = 0.5$.

Next consider the case when there is single cell proliferation ($S > 0$). The same approach that we have used for the $S = 0$ case leads to

$$C_1[K] = \left[Q(1-K) + \frac{SK}{a}\right]e^{-K}. \tag{20}$$

It follows that

$$t = \int_0^K \frac{e^{K'}dK'}{aQ + (S - aQ)K'} \tag{21}$$

and that for $r \geq 1$,

$$C_r[K] = \left[\frac{QK^{r-1}}{(r-1)!} - \left(Q - \frac{S}{a}\right)\frac{K^r}{r!}\right]e^{-K}. \tag{22}$$

There are three cases to consider.

1. *Subcritical reproduction rate*: $S < aQ$. Here $t \to \infty$ when $K \to K_*$, where

   $$K_* = \frac{aQ}{aQ - S}, \tag{23}$$

   and so

   $$\lim_{t \to \infty} C_r(t) = C_r[K_*] = \frac{QK_*^{r-1}(r-1)e^{-K_*}}{r!}. \tag{24}$$

   In this case the reproduction rate S is sufficiently small that we still obtain a steady state (at finite K-time) with $C_1 = 0$, that is, no motile single cells, at $t = \infty$.

2. *Supercritical reproduction rate*: $S > aQ$. In this case $t \to \infty$ corresponds to $K \to \infty$. We can obtain rigorous lower and upper bounds for the integral (21) by replacing K' in the denominator with 1 and 0, respectively, so that

   $$\frac{1}{S}(e^K - 1) < t < \frac{1}{aQ}(e^K - 1)$$

   and it follows that $K \sim \text{constant} \times \ln(t)$ as $t \to \infty$. The reproduction rate S is sufficiently high that a qualitatively different picture emerges. We no longer achieve a steady state in finite K-time and the evolution of the array size distribution is now very different, as discussed in Section 5.4.

3. *Critical reproduction rate*: $S = aQ$. This borderline case is qualitatively similar to the supercritical case, but since $K = \ln(1 + aQt)$ we have a simple solution for the time-dependent array size distribution:

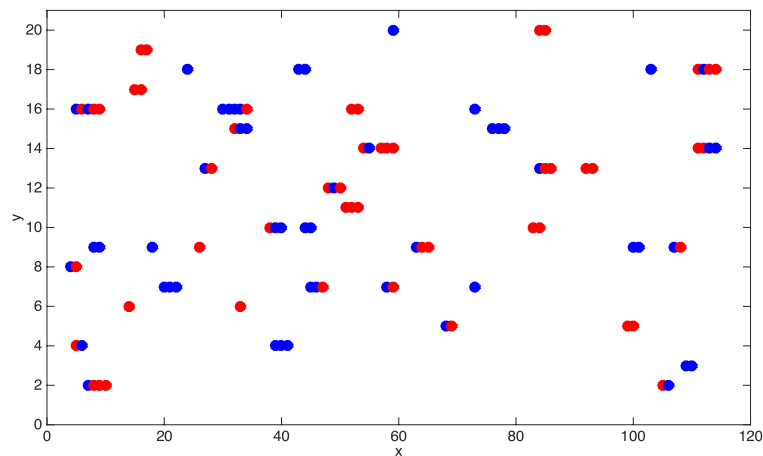 $$C_r(t) = \frac{Q[\ln(1 + aQt)]^{r-1}}{(r-1)!(1 + aQt)}. \tag{25}$$

18

Figure 10: **A realisation of the lattice at P30 from discrete model 1.** Blue agents denote OPCs and OLs of cell type 1 whereas red agents denote OPCs and OLs of cell type 2. Notice that there are more mixed origin arrays than single origin arrays.

Appendix 2: Typical lattice realisations at P30

We present three typical lattice realisations at P30, one for each of the discrete stochastic models. Note that no single realisation will necessarily be representative of the averaged results which we presented in Figures 5–7. Averages were take over 100 realisations for each of the discrete stochastic models.

References

Barres, B. A., Hart, I.K., Coles, H.S.R., Burne, J.F., Voyvodic, J.T., Richardson, W.D., Raff, M.C., 1992. Cell death and control of cell survival in the oligodendrocyte lineage. Cell 70, 31–46.

Baumann, N., Pham-Dinh, D., 2001. Biology of oligodendrocyte and myelin in the mammalian central nervous system. Physiol. Rev. 81, 871–927.

Becker, R. and Döring, W. 1935. Kinetische Behandlung in der Keimbildung übersättingen Dämpfen. Annalen der Physik. 24, 719–752.

Bengtsson, S. L., Nagy, Z., Skare, S., Forsman, L., Forssberg, H., Ullén, F. 2005. Extensive piano practicing has regionally specific effects on white matter development. Nat. Neurosci. 8(9), 1148–1150.

Binder B.J., Landman, K.A., 2009. Exclusion processes on a growing domain. J. Theor. Biol. 259, 541–551.

Cai, A.Q., Landman, K.A., Hughes, B.D. 2007. Multicale modelling of a wound-healing cell migration assay. J. Theor. Biol. 245, 576–594.

Chowdhury, D., Schadschneider, A., Nishinari, K., 2005. Physics of transport and traffic phenomena in biology: from molecular motors and cells to organizms. Phys. Life Rev. 2, 318–352.

Dangata, Y.Y., Kaufman, M.H., 1997. Myelinogenesis in the optic nerve of (C57BL x CBA) F1 hybrid mice: a morphometric analysis. Eur. J. Morphol. 35, 3–17.

Deutsch, A., Dormann, A. 2005. Cellular automaton modelling of biological formation: characterization, applications and analysis. Birkhauser, Boston.

Dimou, L., Simon, C., Kirchhoff, F., Takebayashi, H., Götz, M. 2008. Progeny of Olig2-expressing progenitors in the gray and white matter of the adult mouse cerebral cortex. J. Neurosci. 28(41), 10434–10442.

Gibson, E. M., Purger, D., Mount, C. W., Goldstein, A. K., Lin, G. L., Wood, L. S., Inema, I., Miller, S.E., Bieri, G., Zuchero, J.B., Barres, B.A., Woo, P.J., Vogel, H., Monje, M., 2014. Neuronal activity promotes oligodendrogenesis and adaptive myelination in the mammalian brain. Science 344(6183), 1252304.

Hackett-Jones, E.J., Landman, K.A., Newgreen, D.F., Zhang, D., 2011. On the role of differential adhesion in gangliogenesis in the enteric nervous system. J. Theor. Biol. 287, 148–159.

Hendriks, E. M., Ernst, M. H., 1984. Exactly soluble addition and condensation models in coagulation kinetics. J. Coll. Int. Sci. 97, 176–194.

Hines, J. H., Ravanelli, A. M., Schwindt, R., Scott, E. K., Appel, B., 2015. Neuronal activity biases axon selection for myelination in vivo. Nat. Neurosci. 18, 683–689.

Hughes, E.G., Kang, S.H., Fukaya, M., Bergles, D.E., 2013. Oligodendrocyte progenitors balance growth with self-repulsion to achieve homeostasis in the adult brain. Nat. Neurosci. 16, 668–676.

Johnston, S.T., Shah, E.T., Chopin, L.K., McElwain, D.L.S., Simpson, M. J. 2015. Estimating cell diffusivity and cell proliferation rate by interpreting IncuCyte Zoom assay data using the Fisher-Kolmogorov model. BMC Syst. Biol. 9, 38.
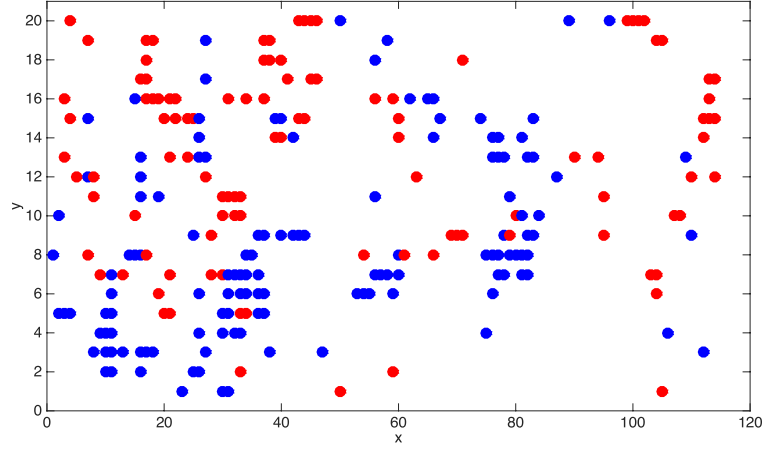
Figure 11: **A realisation of the lattice at P30 from discrete model 2.** Blue agents denote OPCs and OLs of cell type 1 whereas red agents denote OPCs and OLs of cell type 2. Notice that there are few mixed origin arrays.
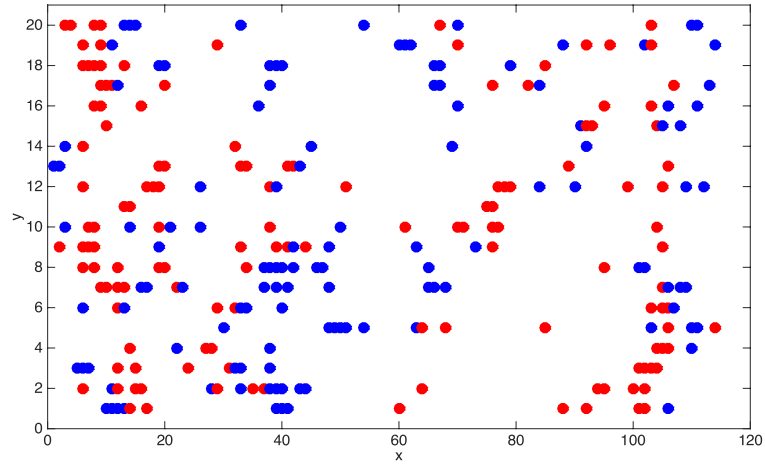

Figure 12: **A realisation of the lattice at P30 from discrete model 3.** Blue agents denote OPCs and OLs of cell type 1 whereas red agents denote OPCs and OLs of cell type 2. Notice that mixed origin arrays are much more frequent than in Figure 11 due to the addition of cell migration.

Kang, S. H., Fukaya, M., Yang, J. K., Rothstein, J. D., Bergles, D. E. 2010. NG2+ CNS glial progenitors remain committed to the oligodendrocyte lineage in postnatal life and following neurodegeneration. Neuron, 68(4), 668–681.

Kessaris, N., Fogarty, M., Iannarelli, P., Grist, M., Wegner, M., Richardson, W.D., 2006. Competing waves of oligodendrocytes in the forebrain and postnatal elimination of an embryonic lineage. Nat. Neurosci. 9, 173–179.

Landman, K.A., Fernando, A.E., 2011. Myopic random walkers and exclusion processes: Single and multispecies. Physica A 390, 3742–3753.

McKenzie, I. A., Ohayon, D., Li, H., Paes de Faria, J., Emery, B., Tohyama, K., Richardson, W.D., 2014. Motor skill learning requires active central myelination. Science 346 (6207), 318–322.

Mensch, S., Baraban, M., Almeida, R., Czopka, T., Ausborn, J., Manira, El, A., Lyons, D. A., 2015. Synaptic vesicle release regulates myelin sheath number of individual oligodendrocytes in vivo. Nat. Neurosci. 18, 628–630.

Mitew, S., Hay, C. M., Peckham, H., Xiao, J., Koenning, M., Emery, B., 2014. Mechanisms regulating the development of oligodendrocytes and central nervous system myelin. Neurosci. 276, 29–47.

Murray, J.D. 2002. Mathematical Biology I: An introduction. Springer, Berlin.

Nave, K.A., 2010. Myelination and support of axonal integrity by glia. Nature 468, 244–252.

Psachoulia, K., Jamen, F.,Young, K.M., Richardson, W.D. 2009. Cell cycle dynamics of NG2 cells in the postnatal and ageing brain. Neuron Glia Biol. 5, 57–67.

Rivers, L. E., Young, K. M., Rizzi, M., Jamen, F., Psachoulia, K., Wade, A., et al. 2008. PDGFRA/NG2 glia generate myelinating oligodendrocytes and piriform projection neurons in adult mice. Nat. Neurosci. 11(12), 1392–1401.

Röth, P., Mitew, S, Stratton J-A.S., Xing, Y.L., Chuang, B.H.A., Walsh, D.M., Holmes, W.R., Hughes, B.D., Landman, K.A., Finighan, R., Johnston, L., Tripathi, R.B., Richardson, W.D., Emery, B., Kilpatrick, T.J., Merson, T.D., 2016. Oligodendrocyte topography in white matter is regulated by axonal electrical activity and is correlated with remyelination efficiency. In preparation.

Saab, A. S., Tzvetanova, I. D., Nave, K.A., 2013. The role of myelin and oligodendrocytes in axonal energy metabolism. Curr. Opin. Neurobiol. 23, 1065–1072.

Sander, L.M., Deisboeck, T.S. 2002. Growth patterns of microscopic brain tumors. Phys. Rev. E 66, 051901.

Scholz, J., Klein, M.C., Behrens, T.E.J., Johansen-Berg, H., 2009. Training induces changes in white-matter architecture. Nat. Neurosci. 12 1370–137.

Simpson, M.J., Merrifield, A., Landman, K.A., Hughes, B.D., 2007. Simulating invasion with cellular automata: connecting individual and macroscopic properties. Phys. Rev. E 76, 021918.

Simpson, M.J., Landman, K.A., Hughes, B.D., 2009. Multi-species simple exclusion processes. Physica A 388, 399–406.

Slemrod, M., 2000. The Becker-Döring equations. In N. Bellomo and M. Pulvirenti (E.), Modelling in Applied Sciences: a Kinetic Theory Approach. Springer, New York.

Sturrock, R.R., 1980. Myelination of the mouse corpus callosum. Neuropathol. Appl. Neurobiol. 6, 415–420.

Suzuki, M., Raisman, G., 1992. The glial framework of central white matter tracts: segmented rows of contiguous interfascicular oligodendrocytes and solitary astrocytes give rise to a continuous meshwork of transverse and longitudinal processes in the adult rat fimbria. Glia 6, 222–235.

Suzuki S.O. and Goldman J.E. 2003. Multiple cell populations in the early post-natal subventricular zone take distinct migratory pathways: a dynamic study of glial and neuronal progenitor migration. J Neurosci. 23, 4240 – 4250.

Tomassy, G. S., Berger, D. R., Chen, H., Kasthuri, N., Hayworth, K. J., Vercelli, A., Seung, H.S., Lichtman, J.W., Arlotta, P., 2014. Distinct profiles of myelin distribution along single axons of pyramidal neurons in the neocortex. Science 344 (6181), 319–324.

Trapp, B.D., Nishiyama, A., Cheng, D., Macklin, W., 1997. Differentiation and death of premyelinating oligodendrocytes in developing rodent brain. J. Cell Biol. 137, 459 – 468.

Xing, Y.L., Röth, P.T., Stratton, J.A., Chuang, B.H.A., Danne, J., Ellis, S.L., Ng, S.W., Kilpatrick, T.J., Merson, T.D., 2014. Adult neural precursor cells from the subventricular zone contribute significantly to oligodendrocyte regeneration and remyelination. J. Neurosci. 34, 14128–14146.

Young, K. M., Psachoulia, K., Tripathi, R. B., Dunn, S.-J., Cossell, L., Attwell, D., 2013. Oligodendrocyte dynamics in the healthy adult CNS: evidence for myelin remodeling. Neuron 77, 873–885.

Zhu, X., Hill, R. A., Dietrich, D., Komitova, M., Suzuki, R., Nishiyama, A. (2011). Age-dependent fate and lineage restriction of single NG2 cells. Development 138(4), 745–753.