Mathematical modelling of cortical neurogenesis reveals that the founder population does not necessarily scale with neuronal output.

Noemi Picco[1,2,3], Fernando García-Moreno[4,5], Thomas Woolley[6], Philip K. Maini[1,2] and Zoltán Molnár[1,3]

1. St John's College Research Centre, St John's College, St Giles Street, Oxford, OX1 3JP, UK

2. Wolfson Centre for Mathematical Biology, Mathematical Institute, University of Oxford, Woodstock Road, Oxford, OX1 6GG, UK

3. Department of Physiology, Anatomy and Genetics, University of Oxford, South Parks Road, Oxford, OX1 3QX, UK

4. Achucarro Basque Center for Neuroscience, Parque Científico UPV/EHU Edif. Sede, E-48940 Leioa, Spain

5. IKERBASQUE Foundation, María Díaz de Haro 3, 6th Floor, 48013 Bilbao, Spain

6. Cardiff School of Mathematics, Cardiff University, Senghennydd Road, Cardiff, CF24 4AG, UK

Corresponding Author: Noemi Picco, St John's College, St Giles Street, Oxford, OX1 3JP, UK. E-mail: picco@maths.ox.ac.uk. Phone: +44 (0) 7552162449. Fax: +44 (0) 1865 273583.

Running title: Mathematical Modelling informs scaling of founder population in cortical neurogenesis.

Abstract

The mammalian cerebral neocortex has a unique structure, composed of layers of different neuron types, interconnected in a stereotyped fashion. While the overall developmental program seems to be conserved, there are divergent developmental factors generating cortical diversity amongst mammalian species. In terms of neuronal output some of the determining factors are the size of the founder population, the duration of cortical neurogenesis, the proportion of different progenitor types, and the fine-tuned balance between self-renewing and differentiative divisions. We develop a mathematical model of neurogenesis that, accounting for these factors, aims at explaining the high diversity found across mammalian species. By framing our hypotheses in rigorous mathematical terms, we are able to identify paths of neurogenesis that match experimentally observed patterns in mouse, and to extrapolate the corresponding patterns in macaque and human. Additionally, we use the model to identify key parameters that would benefit from accurate experimental investigation. We find that the timing of switch to symmetric neurogenic divisions produces the highest

variation in neuronal output. Additionally, we find that an increase in neuronal output does not necessarily reflect a larger size of founder population, and we suggest the need for further experimental quantifications to test our predictions.

Keywords: Asymmetric division, cell cycle of cortical progenitors, cerebral cortex, cortical development, evolution, radial glial cells.

Introduction

The mammalian cerebral neocortex has a unique structure, composed of layers of different neuron types, interconnected in a stereotyped fashion (Brodmann 1908; Cajal 1909; Rakic 2008; Amunts and Zilles 2015). The cortex is linked to numerous cognitive functions such as language, voluntary movement, and episodic memory. Understanding its development and evolution is key to shedding light on the divergent mechanisms that give rise to inter-species differences or diseases, such as microcephaly. The most striking differences across species consist of final neuronal output and surface area (Krubitzer and Kaas 2005), while the variation in radial thickness of the cortex is less pronounced (Rakic 2009, see Table 1). The 1000-fold variation in cortical surface area between mouse and human results in a remarkable specialisation and diversification of cortical function, allowing faster processing of complex information (Kaas 2013). However, the early developmental program that gives rise to each of these cortices seems to be conserved (Bystron et al. 2008; Lui et al. 2011; Kwan et al. 2012). Among the many factors determining the development of a normal cortex in mammalian

species are: the size of the founder population present at the beginning of cortical neurogenesis, the length of neurogenesis, the dynamics of cell cycle length, the proportion of different cortical progenitor types, and the fine-tuned balance between self-renewing and differentiative divisions (Rakic 1995; Florio and Huttner 2014; Imayoshi and Kageyama 2014). Qualitative changes in the outcome of neurogenesis following variations in any of these factors can be intuitively determined. For example, many decades ago Rakic predicted that a prolongation of the time devoted to amplification of the founder population results in expanded neuronal number hence cortical surface (Rakic 1988). However, understanding how all of these factors work together to ensure the cortex is correctly formed is less amenable to verbal reasoning and more suitable for rigorous mathematical modelling.

An impressive amount of comparative studies, quantifying and analysing the main differences in cortical neurogenesis of mammalian species, is available (Molnár et al. 2006; Cheung et al. 2007; Herculano-Houzel 2009; Martínez-Cerdeño et al. 2012). Some comparative quantitative data are beginning to be available on neuronal and glial cell numbers in various species (Herculano-Houzel 2012). However, these studies are limited to finding scaling rules that explain the variation in neuronal output by means of other observable characteristics, such as length of neurogenesis and brain volume, and do not investigate the underlying mechanisms (Mota and Herculano-Houzel 2015).

Recently, focus has turned to the identification of the progenitor types involved in cortical neurogenesis, in an attempt to exhaustively classify the cell types, from neural stem cells to post-mitotic neurons, and all stages of differentiation in between, in

representative species such as mouse (Gal et al. 2006; Stancik et al. 2010), macaque (Wang et al. 2011; Betizeau et al. 2013), and human (Fietz et al. 2010; Hansen et al. 2010). It is possible to draw similarities in the process of cortical neurogenesis in mouse and human, such as the radial expansion of the cortex orchestrated by stage-specific dividing cells, resulting in the formation of the layers in an "inside-first, outside-last" fashion. However, the diversification of progenitor cell types, and non-apically constrained proliferation, are indicative of divergent processes adopted by different species in the course of evolution (Reillo et al. 2011; Pilz et al. 2013). The more we know about the processes involved in each species' cortex development, the further we seem to be from finding a unifying theory that could explain the key variation observed in various mammals with cerebral cortex.

We propose a theoretical framework, based on a first-principle approach, to build a general, yet accurate, representation of a minimal set of processes and players involved in the cortical neurogenesis of the mammalian brain. Building on the current understanding of the processes involved, and framing hypotheses of unknown processes in a mathematically rigorous way, we can make testable predictions and/or expose where there is a lack of biological data necessary for a systematic testing of these hypotheses.

The final goal of this study is to integrate mathematical modelling and experimental observations to highlight both the communal processes in the early development of, and the divergent factors that can explain cortical diversity amongst, mammalian species.

Materials and Methods

We define a mathematical model of cortical neurogenesis, describing the dynamics of proliferation and differentiation at the cell population level, during the development of a mammalian brain. For this model to be descriptive of any species considered, we classify cells into two compartments: progenitors (P) and neurons (N). The former includes all pre-mitotic neural progenitor cell types involved in neocortex development for a given species, e.g. neuroepithelial cells, apical and basal radial glial, intermediate progenitors (Rakic 1995; Huttner and Brand 1997; Noctor et al. 2004; Florio and Huttner 2014). The neuron compartment includes all post-mitotic and permanently differentiated cells. We only consider neurons locally generated by progenitors situated between the ventricular and pial surfaces. Migration of glutamatergic neurons from outside sources (Barber and Pierani 2016) or GABAergic neurons from subpallial ganglionic eminences (De Carlos et al. 1996; Parnavelas 2000; Anderson et al. 2001; Tamamaki et al. 2003) is not explicitly modelled, however it will be accounted for when integrating experimental quantifications into the model.

We consider three possible modes of progenitor cell division: self-amplifying division (SymP), generating two identical progenitors; asymmetric neurogenic division (AsymN), generating one progenitor and one neuron; symmetric neurogenic division (SymN), generating two neurons.

As we will see later, if these three types of division occurred with constant proportions, we would observe one of two possible outcomes: extinction or unlimited growth of the

progenitor population. It is known, instead, that these modes of proliferation are preferentially balanced at different stages of neocortical neurogenesis (Noctor et al. 2004; Götz and Huttner 2005), and this program seems to be conserved across species (Kornack 2000).

We, therefore, define time-dependent probabilities describing the preferred mode of division in the progenitor population during neurogenesis: $\alpha(t)$ is the probability of AsymN at a given time t, $\beta(t)$ is the probability of SymN; hence $1-\alpha(t)-\beta(t)$ is the probability of SymP. The modulation of these probability functions in time defines a strategy that can be represented as a trajectory in the three-dimensional space (AsymN, SymN, SymP) (Fig. 1).

The resulting system of Ordinary Differential Equations (ODEs), describing the evolution in time of the progenitor (P) and neuron (N) populations is:

$$\begin{cases} \frac{dP}{dt} = \rho\, P\, (1-\alpha(t)-2\beta(t)), & P(t_0) = P_0, \\ \frac{dN}{dt} = \rho\, P\, (\alpha(t) + 2\beta(t)), & N(t_0) = 0, \end{cases} \tag{1}$$

where ρ is the rate of division, linked to the cell cycle length T_C by: $\rho = \log 2/T_C$, and P_0 is the founder population. We define t_0 as the time of onset of neurogenesis, i.e. when the first neuron is produced, and t_F as the final time of neurogenesis, and consider the time interval $t \in (t_0, t_F)$.

Cell death of post-mitotic neurons is not explicitly modelled, but will be accounted for when integrating experimental estimates of final neuronal output into the model.

Similarly, the model can be refined to account for cell death of the progenitor cells. Gohlke et al. (2007) investigated the role of cell death in a computational model of neuronal acquisition.

In the special case of constant α and β, the P population either goes extinct, remains constant, or grows without bound, depending on the value of $(1-\alpha-2\beta)$, as shown in supplemental Figure S1. Our time-dependent probabilities will model the observed patterns of division (Takahashi et al. 1996), with SymP the preferred mode of division at early neurogenesis, AsymN peaking at mid-neurogenesis and finally decreasing to leave SymN as the preferred mode of division at late neurogenesis. Therefore, we choose the probability functions to be piece-wise linear functions of time and define the following parameters: α_0, the proportion of AsymN at onset of neurogenesis; α_S, the proportion of AsymN at the time of strategy switch; β_F, the proportion of SymN at the end of neurogenesis; and t_S, the time of the switch to a strategy favouring SymN. These four parameters uniquely determine the shape of the probability functions represented in Figure 2, hence the trajectory in parameter space (Fig. 1). For an analytical description of the probability functions, see the supplemental information. We only consider biologically relevant parameter values, namely:

$$0 < \alpha_0 < 1, \quad \alpha_0 < \alpha_S < 1, \quad 0 < \beta_F < 1, \quad t_0 < t_S < t_F\,. \tag{2}$$

Within these constraints, any combination of the four parameter values will correspond to a trajectory in the strategy space.

The model outcome, given a choice of parameter values, consists of the functions of

time $P(t)$ and $N(t)$, representing the cell counts over the course of neurogenesis, i.e. for $t \in (t_0, t_F)$.

We initially consider mouse embryonic cortical neurogenesis, occurring between E11 and E19 (E indicates embryonic day), and fixed cell cycle length, $T_C = 14.3$ hours (see Table 1). However, in a second version of the model, we introduce an age-dependent cell cycle time, $T_C(t)$, to account for the experimentally observed modulation of the cell cycle length throughout the course of neurogenesis, mostly due to variations in the S phase (Dehay and Kennedy 2007).

Results

Our study focuses on three species of interest: mouse, macaque monkey, and human. We intend to build a mathematical framework that can easily be extended to consider a larger number of species, with the aim of analysing and explaining cortical neurogenesis through the lens of evolution. Although data on a large number of other mammalian species exist, we focus on these three species that have been chosen in previous studies to effectively illustrate neocortex development across distinct branches of the phylogenetic tree (Rakic 1995).

Initially focussing on the mouse, a species whose neurogenesis has been best characterised, we search the strategy space to find the strategy, or strategies, matching the observed patterns of neuronal output. Finally, we translate the mouse strategy to macaque and human neurogenesis, to make predictions about the size of the founder population, currently unavailable in the experimental literature. The model we propose

can be used to compare different developmental strategies to determine the existence of evolutionary or developmental constraints. Finally, we can determine those parameters, to which the model outcome is most sensitive, suggesting the need for higher resolution data, and identify those parameters whose variation does not substantially affect model predictions.

The strategy to build a mouse neocortex

The laminar position of a newly generated neuron is determined by the time of birth in an "inside-first, outside-last" fashion (McConnell and Kaznowski 1991), therefore we can assume that the radial expansion of the neocortex is proportional to the number of neurons. In mice, specifically, the total number of neurons per column is approximately shared between upper and deeper cortical layers (Markram et al. 2004; Vasistha et al. 2015; Table 1).

Since we are interested in the ratio of population sizes at different times, we can rewrite equations (1) as:

$$\begin{cases} \frac{d\tilde{P}}{dt} = \rho\, \tilde{P}\, (1 - \alpha(t) - 2\beta(t)), & \tilde{P}(t_0) = 1, \\ \frac{d\tilde{N}}{dt} = \rho\, \tilde{P}\, (\alpha(t) + 2\beta(t)), & \tilde{N}(t_0) = 0, \end{cases} \tag{3}$$

where $\tilde{P}$ and $\tilde{N}$ are progenitor and neuron cell numbers, respectively, per initial progenitor P_0, so that the problem is independent of the size of the founder population.

Therefore, we define the *mouse optimal strategy* as the 4-tuple $(\alpha_0, \alpha_S, \beta_F, t_S)$ that best approximates the following:

$$\widetilde{N}(t_F) = 2\widetilde{N}(t_M) + \varepsilon, \quad (4)$$

with minimum error ε. Here, $t_M = t_0 + (t_F - t_0)/2$, is defined as the time when production of layer V is completed and production of layer IV starts. Hence, we look for the strategy that, to a first approximation, produces half of the total neuronal output (i.e. deeper layers) in the first half of neurogenesis and the remainder (i.e. upper layers) in the second half.

Given the small dimensionality of the problem, it is computationally feasible to perform a systematic and exhaustive search of the parameter space corresponding to all possible 4-tuples that satisfy equations (2). Doing so, we obtain the optimal strategy:

$$\alpha_0 = 0.1,\ \alpha_S = 0.9,\ \beta_F = 0.8,\ t_S = E13\,. \quad (5)$$

The trajectory corresponding to this mouse optimal strategy is shown in Figure 3(A). Note that, given the linearity of the system, the strategy satisfying equation (4) for $\varepsilon = 0$ will still be valid for any value of P_0, or for a different neurogenesis length $(t_F - t_0)$, provided that the time of switch t_S is rescaled. Supplemental Figure S2 shows the result of the full search of the 4-dimensional parameter space.

Our parameter sensitivity analysis (see supplemental information for details) reveals that the highest sensitivity of the output can be attributed to variations in the t_S parameter (Fig. 4). This means that the variation in model output observed when varying the time of switch (t_S) is more substantial when compared to the effect of varying the absolute prevalence of different types of division $(\alpha_0, \alpha_S, \beta_F)$. This sensitivity,

as well as lack of experimental quantification, illustrates a demand for future experimental research into quantifying, or at least characterising, this time scale.

Age-dependent cell cycle

A simplifying assumption in the model is that cell cycle length is kept constant during neurogenesis. However, species-specific age-dependencies could significantly affect the total number of divisions occurring during neurogenesis, and consequently the final neuronal output. For example, the linear increase of cell cycle length observed in mice determines a lower rate of division in late neurogenesis. In macaque, however, an initial increase in cell cycle length is followed by a decrease in mid-to-late neurogenesis (Kornack and Rakic 1998). This age-dependent modulation of frequency of division, combined with the temporal modulation of propensities of the modes of division, can substantially alter the dynamics and, hence, the outcome of neurogenesis. Our model (1) can be extended to account for this time dependency (i.e. $\rho = \rho(t)$) to give a more accurate representation of the system and describe what turn out to be non-intuitive implications, in terms of neuronal output.

For each species of interest we chose a linear description of $T_C(t)$, obtained by interpolation of the data from Kornack and Rakic (1998; see Fig. 5A top row, and supplemental information for the analytical definition). Repeating the search in the strategy space with the age-dependent model, calibrated to the mouse, we obtained a new optimal strategy:

$$\alpha_0 = 0.5,\ \alpha_S = 0.8,\ \beta_F = 0.9,\ t_S = E13.5, \tag{6}$$

represented in Figure 3(B). Intuitively, the delay in the time of switch to neurogenic divisions increases the pool of progenitors to compensate for a decrease in the frequency of division at late neurogenesis. Supplemental Figure S3 shows the result of the full search of the 4-dimensional parameter space.

Estimate of founder population for mouse, macaque and human cortex

The number of cortical neurons present at the end of neurogenesis far exceeds the number in the adult neocortex, because of significant apoptosis of progenitors and newly born neurons (Kuan et al. 2000; Pompeiano et al. 2000). Additionally, in several instances, neurons migrate to the cortex after being generated elsewhere. Some examples are glutamatergic Cajal-Retzius and subplate neurons generated in the cortical hem or rostral medial telencephalic wall (Bielle et al. 2005; García-Moreno et al. 2007; Pedraza et al. 2014) or GABAergic interneurons generated in the medial, lateral and caudal ganglionic eminences (De Carlos et al., 1996; Lavdas et al., 1999; Anderson et al. 2001; Corbin et al. 2001; Tamamaki et al. 2003). Therefore, we estimate the neuronal output $N(t_F)$ adjusting the number of adult cortical neurons for 30% cell death and 25% of interneurons migrating to the cortex. Later we show that the results are robusts of variations in these quantities.

Given $N(t_F)$, we intend to use our neurogenesis model to estimate a key quantity of interest that lacks experimental quantification: the founder population, P_0. For each species of interest, we use Approximate Bayesian Computation (ABC, see supplemental information) to obtain an estimate of P_0. This method gives the best fit to

match the neuronal output $N(t_F)$, given the model of choice (constant, or age-dependent, cell cycle length), and the corresponding optimal strategy (Eq. 5 or Eq. 6, respectively). Figure 5(B) shows the estimates of P_0 for the three species and two models.

Both the constant and age-dependent cell cycle models predict that the exponential increase of cortical neurons from mouse to macaque is justified by a linear increase in the founder population (note the logarithmic scale on the y axis). This prediction agrees with the radial unit linear model (Rakic 2009), attributing an exponential expansion of the number of radial columns to the linear increase in the symmetrically dividing neural stem cell pool that, in our model, is represented by the founder population P_0.

Our estimate for human P_0, however, leads to a somewhat counterintuitive prediction. When moving from macaque to human, the founder population required to produce larger neuronal output is not increasing. Specifically, the founder population estimated for the human cortex is smaller than that for the macaque, despite producing a considerably larger neuronal output.

Solutions $N(t)$ and $P(t)$ for the parameterised models are shown in Figure 5(A), bottom row. It is worth noting that, despite the estimation being obtained by fitting one data point only, $N(t_F)$, for a given strategy, the time course of the model output does not vary significantly when comparing the two models (compare solid and dashed lines). Therefore, a more accurate representation of age-dependent cell cycle length does not change the time dynamics, but it does significantly affect quantitative estimations of the

size of the founder population, especially for larger species.

To test the robustness of our result and challenge such a counterintuitive prediction, we introduce variations in the assumptions that could have skewed the results. Specifically, we repeat the estimations for varied amounts of interneuronal migration and post-neurogenesis cell death. Even an unrealistically large rate of post-neurogenesis cell death does not alter the prediction of lower human founder population (Suppl. Fig. S4A). Similarly, the prediction still holds even when increasing the rate of interneuronal migration to 50% (Suppl. Fig. S4B), although a more realistic 13% rate has been suggested (Mountcastle 1997). Finally, due to the lack of quantification of cell cycle length for human progenitors at various developmental stages, we used the experimental data for the macaque cell cycle length. This assumption, adopted in previous comparative studies (Rakic 1995), is valid only as a first approximation, since the two species belong to the same order. However, an underestimation of the cell cycle length in humans could cause an underestimation of the corresponding founder population. Therefore, we repeat the estimations for longer cell cycles, up to twice the baseline value for macaque progenitor T_C (Suppl. Fig. S5A and B for age-dependent and constant cell cycle length models, respectively). For such large (biologically unrealistic) values, the model prediction breaks down: an increasingly large founder population is necessary to obtain increasingly larger neuronal outputs by the end of neurogenesis.

The neurogenesis simulator

We have developed a graphic user interface allowing the user to select and calibrate the neurogenesis model, observing how the temporal population dynamics vary when changing parameter values. The user can select and compare strategies of neocortical neurogenesis across different species. The neurogenesis simulator is available for download at www.dpag.ox.ac.uk/team/noemi-picco .

Discussion

We propose a simple mathematical model of cortical neurogenesis that aims at describing the dynamics of cell proliferation and differentiation in various species.

Specifically, we can think of the cortex in each species arising as the result of an optimisation of cell proliferation and differentiation, to obtain the optimal strategy. Such a strategy consists of the modulation in time of these processes during development (Fig. 1) so that, within the time allowed for cortical neurogenesis in each species, the required number of neurons is produced and cortical layers are formed.

In order to make meaningful comparisons between species, we defined a two-population model, describing the time dynamics of progenitors and neurons during cortical neurogenesis. The progenitor population is defined differently for any species of interest and its description can be refined to include different types of progenitor cells (e.g. neuroepithelial cells, radial glial, intermediate progenitors, etc.).

The main hypothesis of this model is that the three types of division (SymP, AsymN, SymN) are preferentially adopted at different stages of neurogenesis. Therefore, we

imposed a two-step linear strategy with the prevalent mode of division shifting from self-replication at early neurogenesis to neurogenic differentiation at later times (Fig. 2). We propose two variants of the model, with constant, and age-dependent, cell cycle length, respectively.

The model framework allows the systematic exploration of the strategy space by simultaneously varying all parameter values within biologically relevant ranges to find the strategy (strategies) matching the observed outcomes (Suppl. Fig. S2 and S3). We could obtain the optimal strategy for the mouse, experimentally the best characterised species (Fig. 3).

Despite imposing the shape of probability functions, to describe a shift in prevalence of mode of division, we do not make assumptions about the timing of the switch from mostly self-replication to mostly neurogenic differentiation. The model prediction is that, to match the pattern of neuronal output observed in mice, the time of switch must be around early/mid-neurogenesis.

We then used sensitivity analysis to determine which parameters have the highest influence on the model outcome. We found that the timing of switch of strategy from mostly self-replicating to mostly neurogenic divisions is more important, in terms of neuronal output, than the mere proportions of these types of division (Fig. 4). Although a number of cell-intrinsic factors have been found to regulate modes of division at the single-cell level, a full understanding of the mechanisms is lacking. For example, overexpression of $p27^{Kip1}$, a cyclin-dependent kinase inhibitor, was found to

substantially reduce neuronal output in supergranular layers (i.e. late-stage neuronal production; Tarui et al. 2005). Although this constitutes only circumstantial evidence, $p27^{Kip1}$ is a potential candidate for the experimental validation of the effects of a delay in the time of switch. Similarly, β –catenin was found to be a key regulator of cell cycle re-entry of neural precursors, promoting self-replicating divisions and generating aberrant cortices in mutant mice (Chenn and Walsh 2002). Primary microcephaly is an example of a developmental disorder associated with a mistimed switch between proliferative and differentiative divisions. Specifically, all microcephaly genes identified are related to centrosomal abnormalities affecting mitosis (Gilmore and Walsh 2013).

Having characterised the optimal mouse strategy, we used the model to estimate another quantity of interest, currently unavailable in the literature, namely, the size of the founder population of progenitors. Our model makes a counterintuitive prediction: the number of founder progenitors needed to produce the neurons present in the human cortex at the end of neurogenesis is lower than the corresponding estimate for macaque, despite the two-fold increase in neuronal output (Fig. 5). This prediction holds even for large variations in key parameters, such as rates of post-neurogenesis cell death and interneuronal migration (Suppl. Fig. S4).

Finally, we investigate the effects of variations in cell cycle length, assumed to be approximately similar in all primates, given the lack of data. We allowed some variation from the baseline value (macaque) and observed that the model prediction breaks down for large, unrealistic, values of human progenitor cell cycle length. This is an example of how, by integrating theoretical and experimental modelling, we can determine how other

observed phenomena affect our hypotheses, predictions and understanding of the process that is being investigated. This approach is particularly relevant in resolving questions arising from circumstantial evidence. Indeed, the theoretical model can suggest whether a given mechanism could lead to significantly different predictions, suggesting that further experimental investigation is necessary, or otherwise futile for the problem being investigated.

To the best of our knowledge, a quantification of the cell cycle length of human progenitors is not available from the current literature in vivo. Therefore, at this stage, we can only speculate on the implications of this counterintuitive prediction. However, our study suggests that further experimental studies should focus on the quantification of cell cycle duration, as our model shows that very different predictions can follow. Conversely, variations in other properties of the system, such as post-neurogenesis neuronal cell death and interneuronal migration, do not affect the qualitative predictions of the model.

There is strong evidence of a link between cell cycle lengthening (specifically, of the G1 phase) and the progression to a higher propensity to differentiative divisions (Dehay and Kennedy 2007). However, the relationship of causality between these two factors has yet to be established. By proposing two versions of the cell cycle model, we offer an alternative interpretation of the developmental program, with no direct relationship between cell cycle length and preferred modes of division. Our constant, and age-dependent, cell cycle models lead to qualitatively similar predictions regarding the scaling of the founder population, but very different quantitative predictions for a given

species (Fig. 5B). Experimental quantification of P_0 would give the added benefit of validating either one of the two models, determining the need (or not) of age-dependent cell cycle to quantitatively recapitulate the population dynamics.

Conclusions

With this study we introduced a minimal theoretical framework which, informed by experimental observations, offers a mechanistic explanation of the development and evolution of the mammalian neocortex. This model allows a comparative study of mammalian species, and future modelling of human pathologies, such as microcephaly. In building our first minimal model we left out a number of properties of the system and mechanisms that could be integrated as they become relevant and necessary to answer specific biological questions, and as data becomes available. For example, radial migration is only modelled implicitly, by assuming that positioning in the radial direction is related to time of birth (Rakic 1974; McConnell et al. 1991) fundamental to our assumption of half neuronal output being produced by mid-neurogenesis, corresponding to approximately equal neuron numbers between upper and deeper cortical layers. Additionally, our model captures the dynamics of neuronal acquisition averaged over the entire neocortex, neglecting variations in cortical areas. Early regionalization determined by variations in cell-cycle kinetics is crucial (Polleux et al. 1997). However, such differences are relatively minor compared to the cross-species variation that we intended to address with this work. Furthermore, our model explains inter-species variation in neuronal output regardless of cortical surface and folding. To explain these differences (i.e. gyrencephalic vs lissencephalic) we would have to account for density,

cell size, tangential migration, etc., all of which are species-specific quantities and properties. Specifically, it was shown that cortical folding approximately scales with surface area, and thickness, and not with neuronal numbers (Mota and Herculano-Houzel 2015). Additionally, through integration of imaging and mathematical modelling it was demonstrated that variations in cortical thickness within the cortex are caused by the physical forces emerging during the folding process (Holland et al. 2017). However, the problem of neocortex folding is a separate avenue of research, decoupled from the mechanisms that we want to investigate here.

The aim of this work was to go beyond scaling rules, to identify the minimum set of players, properties and interactions sufficient to explain cross-species variation in cortical neurogenesis in terms of mechanisms.

Specifically, we can exploit the predictive and analytical power of mathematical modelling to discriminate between the many factors contributing to the formation of the cortex. Additionally, sensitivity analysis allows us to pinpoint those mechanisms where higher resolution data would be beneficial to improve our understanding of cortical neurogenesis. The usefulness of mathematical modelling is two-fold. It can be used to guide us on experiments that are necessary to quantify the key parameters that need to be systematically compared across species. To this end we introduced the *neurogenesis simulator*, a tool allowing the user to observe changes in the model outcome as model parameters and assumptions are varied. Mathematical modelling can also be helpful in formally proving (or falsifying) empirical hypotheses generated from experimental observations, such as the radial unit hypothesis formulated many

decades ago by Rakic. We found that the timing of shift from self-replicating to neurogenic divisions is key to obtaining the correct neuronal output by the end of neurogenesis. The lack of experimental quantification of a second key quantity, cell cycle duration, leaves an open question. Namely, whether the predicted reduction in founder population size of human versus macaque is real. If experimentally validated, this prediction would have a key implication in evolutionary terms. That is, species with increasingly larger numbers of cortical neurons could likely have adopted the same *optimal strategy* adjusting the duration of neurogenesis, thus requiring an approximately conserved size of founder population. Indeed, a founder population size that scales with the neuronal output would be prohibitively large and inefficient, compared to running the same developmental program (strategy) over a stretched window of time.

We envisage the use of our neurogenesis model to map evolutionary trajectories that describe neurogenesis in different species, as well as deviations in such trajectories corresponding to brain diseases characterised by a developmental program gone awry.

Acknowledgements

This work was supported by a St John's College Research Centre Grant to TW, PKM and ZM in collaboration with FG-M. The work in the laboratory of ZM was supported by grants from MRC, Wellcome Trust, Royal Society and St John's College Research Centre. PKM and TEW would like to thank the Mathematical Biosciences Institute (MBI) at Ohio State University, for partially supporting this research. MBI receives its funding

through the National Science Foundation (grant number DMS1440386). FG-M holds an IKERBASQUE Research Fellowship. The authors gratefully acknowledge useful discussions with members of the Brain Research And Inquiry Network (BRAIN) at St John's College Research Centre. We also thank Dr. Navneet Vasistha, Dr. Patricia Garcez, Prof. Bruno Mota, Prof. Alain Goriely for discussion, critical reading and feedback on earlier versions of this manuscript.

References

Amunts K, Zilles K. 2015. Architectonic Mapping of the Human Brain beyond Brodmann. Neuron 88:1086-1107.

Anderson SA, Marín O, Horn C, Jennings K, Rubenstein JL. 2001. Distinct cortical migrations from the medial and lateral ganglionic eminences. Development. 128:353-363.

Azevedo FA, Carvalho LR, Grinberg LT, Farfel JM, Ferretti RE, Leite RE, Jacob Filho W, Lent R, Herculano-Houzel S. 2009. Equal numbers of neuronal and nonneuronal cells make the human brain an isometrically scaled-up primate brain. J Comp Neurol. 513:532-541.

Badea A, Ali-Sharief AA, Johnson GA. 2007. Morphometric analysis of the C57BL/6J mouse brain. Neuroimage. 37:683-693.

Barber M, Pierani A. 2016. Tangential migration of glutamatergic neurons and cortical

patterning during development: lessons from Cajal-Retzius cells. Dev Neurobiol. 76:847-881.

Betizeau M, Cortay V, Patti D, Pfister S, Gautier E, Bellemin-Ménard A, Afanassieff M, Huissoud C, Douglas RJ, Kennedy H Dehay C. 2013. Precursor diversity and complexity of lineage relationships in the outer subventricular zone of the primate. Neuron. 80:442-457.

Bielle F, Griveau A, Narboux-Nême N, Vigneau S, Sigrist M, Arber S, Wassef M, Pierani A. 2005. Multiple origins of Cajal-Retzius cells at the borders of the developing pallium. Nat Neurosci. 8:1002-1012

Brodmann K. 1908. Beiträge zur histologischen Lokalisation der Grosshirnrinde, Sechste Mitteilung: Die Cortexgliederung des Menschen. J Psychol Neurol. 10:231–246.

Bystron I, Blakemore C, Rakic P. 2008. Development of the human cerebral cortex: Boulder Committee revisited. Nat Rev Neurosci. 9:110-122.

Cajal SR. 1909. Histologie dù Système Nerveux de l´homme et des vertébrés. París (FR): Maloine.

Charvet CJ, Cahalane DJ, Finlay BL. 2015. Systematic, cross-cortex variation in neuron numbers in rodents and primates. Cereb Cortex. 25:147-160.

Chenn A, Walsh CA. 2002. Regulation of cerebral cortical size by control of cell cycle exit in neural precursors. Science. 297:365-369.

Cheung AF, Pollen AA, Tavare A, DeProto J, Molnár Z. 2007. Comparative aspects of cortical neurogenesis in vertebrates. J Anat. 211:164-176.

Clark DA, Mitra PP, Wang SSH. 2001. Scalable architecture in mammalian brain. Nature. 411:189-193.

Corbin JG, Nery S, Fishell G. 2001. Telencephalic cells take a tangent: non-radial migration in the mammalian forebrain. Nat Neurosci. 4:1177-1182.

De Carlos JA, López-Mascaraque L, Valverde F. 1996. Dynamics of cell migration from the lateral ganglionic eminence in the rat. J Neurosci. 16:6146-6156.

Dehay C, Kennedy H. 2007. Cell-cycle control and cortical development. Nat Rev Neurosci 8:438-450.

Fietz SA, Kelava I, Vogt J, Wilsch-Bräuninger M, Stenzel D, Fish JL, Corbeil D, Riehn A, Distler W, Nitsch R, Huttner WB. 2010. OSVZ progenitors of human and ferret neocortex are epithelial-like and expand by integrin signaling. Nat Neurosci. 13:690-699.

Florio M, Huttner WB. 2014. Neural progenitors, neurogenesis and the evolution of the neocortex. Development. 141:2182-2194.

Gal JS, Morozov YM, Ayoub AE, Chatterjee M, Rakic P, Haydar TF. 2006. Molecular and morphological heterogeneity of neural precursors in the mouse neocortical proliferative zones. J Neurosci. 26:1045-1056.

García-Moreno F, López-Mascaraque L, De Carlos JA. 2007. Origins and migratory routes of murine Cajal-Retzius cells. J Comp Neurol. 500:419-432.

Gilmore EC, Walsh CA. 2013. Genetic causes of microcephaly and lessons for neuronal development. Wiley Interdiscip Rev Dev Biol. 2:461-478.

Gohlke JM, Griffith WC, Faustman EM. 2007. Computational models of neocortical neuronogenesis and programmed cell death in the developing mouse, monkey, and human. Cereb Cortex. 17:2433-2442.

Götz M, Huttner WB. 2005. The cell biology of neurogenesis. Nat Rev Mol Cell Biol 6:777-788.

Hansen DV, Lui JH, Parker PRL, Kriegstein AR. 2010. Neurogenic radial glia in the outer subvenmtricular zone of human neocortex. Nature. 464:554-561.

Haydar TF, Nowakowski RS, Yarowsky PJ, Krueger BK. 2000. Role of founder cell deficit and delayed neurogenesis in microencephaly of the trisonomy 16 mouse. J Neurosci. 20:4156-4164.

Herculano-Houzel S, Mota B, Lent R. 2006. Cellular scaling rules for rodent brains. PNAS 103:12138-12143.

Herculano-Houzel Collins CE, Wong P, Kaas JH. 2007, Cellular scaling rules for primate brains. Procl Natl Acad Sci USA. 104:3562-3567.

Herculano-Houzel S. 2009. The human brain in numbers: a linearly scaled-up primate

brain, Front Hum Neurosci. 3:1-11.

Herculano-Houzel S. 2012. The remarkable, yet not extraordinary, human brain as a scaled-up primate brain and its associated cost. Procl Natl Acad Sci USA. 109:10661-10668.

Holland M, Budday S, Goriely A, Kuhl E. 2017. Symmetry breaking in the developing brain: cortical thickness variations emerge from folding. Unpublished.

Huttner WB, Brand M. 1997. Asymmetric division and polarity of neuroepithelial cells. Curr Opin Neurobiol. 7:29-39.

Imayoshi I, Kageyama R. 2014. bHLH factors in self-renewal, multipotency, and fate choice of neural progenitor cells. Neuron. 82:9-23.

Kaas JH. 2013. The evolution of brains from early mammals to humans. Wiley Interdiscip Rev Cogn Sci. 4:33-45.

Kornack DR, Rakic P. 1998. Changes in cell-cycle kinetics during the development and evolution of primate neocortex. Proc Natl Acad Sci USA. 95:1242-1246.

Kornack DR. 2000. Neurogenesis and the evolution of cortical diversity: mode, tempo, and partitioning during development and persistence in adulthood. Brain Behav Evol. 55:336-344.

Krubitzer L, Kaas J. 2005. The evolution of the neocortex in mammals: how is phenotypic diversity generated? Curr Opin Neurobiol. 15:444-453.

Kuan C, Roth KA, Flavell RA, Rakic P. 2000. Mechanisms of programmed cell death in the developing brain. Trens Neurosci. 23:291-297.

Kwan KY, Sestan N, Anton ES. 2012. Transcriptional co-regulation of neuronal migration and laminar identity in the neocortex. Development. 139:1535-1546.

Lavdas AA, Grigoriou M, Pachnis V, Parnavelas JG. 1999. The medial ganglionic eminence gives rise to a population of early neurons in the developing cerebral cortex. J Neurosci. 19:7881-7888.

Lui JH, Hansen DV, Kriegstein AR. 2011. Development and evolution of the human neocortex. Cell. 146:18-36.

Markram H, Toledo-Rodriguez M, Wang Y, Gupta A, Silberberg G, Wu C. 2004. Interneurons of the neocortical inhibitory system. Nat Rev Neurosci. 5:793-807.

Martínez-Cerdeño V, Cunningham CL, Camacho J, Antczak JL, Prakash AN, Cziep ME, Walker AI, Noctor SC. 2012. Comparative analysis of the subventricular zone in rat, ferret and macaque: evidence for an outer subventricular zone in rodents. PLoS One. 7:e30178.

McConnell SK, Kaznowski CE. 1991. Cell cycle dependence of laminar determination in developing neocortex. Science. 254:282-285.

Molnár Z, Métin C, Stoykova A, Tarabykin V, Price DJ, Francis F, Meyer G, Dehay C, Kennedy H. 2006. Comparative aspects of cerebral cortical development. Eur J Neurosci. 23:921-934.

Mota B, Herculano-Houzel S. 2015. Cortical folding scales universally with surface area and thickness, not number of neurons. Science. 349:74-77.

Mountcastle VB. 1997. The columnar organization of the neocortex. Brain. 120:701-722.

Noctor SC, Martínez-Cerdeño V, Ivic L, Kriegstein AR. 2004. Cortical neurons arise in symmetric and asymmetric division zones and migrate through specific phases. Nat Neurosci. 7:136-44.

Parnavelas JG. 2000. The origin and migration of cortical neurons: new vistas. Trends Neurosci. 23:126-131.

Pedraza M, Hoerder-Suabedissen A, Albert-Maestro MA, Molnár Z, De Carlos JA. 2014. Extracortical origin of some murine subplate cell populations. Proc Natl Acad Sci U S A. 111:8613-8618.

Picco N, E Sahai, PK Maini, ARA Anderson. 2017. Integrating models to quantify environment mediated drug resistance. Cancer Res. 19:5409-5418.

Pilz GA, Shitamukai A, Reillo I, Pacary E, Schwausch J, Stahl R, Ninkovic J, Snippert HJ, Clevers H, Godinho L, Guillemot F, Borrell V, Matsuzaki F, Götz M. 2013. Amplification of progenitors in the mammalian telencephalon includes a new radial glial cell type. Nat Commun. 4:2125.

Polleux F, Dehay C, Moraillon B, Kennedy H. 1997. Regulation of neuroblast cell-cycle kinetics plays a crucial role in the generation of unique features of neocortical areas. J Neurosci. 17:7763-7783.

Pompeiano M, Blaschke AJ, Flavell RA, Srinivasan A, Chun J. 2000. J Comp Neurol. 423:1-12.

Rakic P. 1974. Neurons in rhesus monkey visual cortex: systematic relation between time of origin and eventual disposition. Science. 183:425-427.

Rakic P. 1988. Specification of cerebral cortical areas. Science. 241:170-176.

Rakic P. 1995. A small step for the cell, a giant leap for mankind: a hypothesis of neocortical expansion during evolution. Trends Neurosci. 18:383-388.

Rakic P. 2008. Confusing cortical columns. Procl Natl Acad Sci USA. 105:12099-12100.

Rakic P. 2009. Evolution of the neocortex: a perspective from developmental biology. Nat Rev Neurosci. 10:724-735.

Reillo I, de Juan Romero C, García-Cabezas MÁ, Borrell V. 2011. A role for intermediate radial glia in the tangential expansion of the mammalian cerebral cortex. Cereb Cortex. 21:1674-1694.

Rockel AJ, Hiorns RW, Powell PS. 1980. The basic uniformity in structure of the neocortex. Brain. 103:221-244.

Stancik EK, Navarro-Quiroga I, Sellke R, Haydar TF. 2010. Heterogeneity in ventricular zone neural precursors contributes to neuronal fate diversity in the postnatal neocortex. J Neurosci. 30:7028-36.

Stephan H, Frahm H, Baron G. 1981. New and revised data on volumes of brain

structures in insectivores and primates. Folia Primatol. 35:1-29.

Sultan F. 2002. Analysis of mammalian brain architecture. Nature. 415:133-134.

Takahashi T, Nowakowski RS, Caviness VS. 1996. The leaving or Q fraction of the murine cerebral proliferative epithelium: a general model of neocortical neuronogenesis. J Neurosc. 16:6183-6196.

Tamamaki N, Fujimori KE, Nojyo K, Kaneko Y, Takauji R. 2003. Evidence that Sema3A and Sema3F regulate the migration of GABAergic neurons in the developing neocortex. J Comp. Neurol. 455:238-248.

Tarui T, Takahashi T, Nowakowski RS, Hayes NL, Bhide PG, Caviness VS. 2005. Overexpression of p27 Kip 1, probability of cell cycle exit, and laminar destination of neocortical neurons. Cerebral Cortex. 15:1343-1355.

Vasistha NA, García-Moreno F, Arora S, Cheung AF, Arnold SJ, Robertson EJ, Molnár Z. 2015. Cortical and Clonal Contribution of Tbr2 Expressing Progenitors in the Developing Mouse Brain. Cereb Cortex. 25:3290-3302.

Wang X, Tsai JW, LaMonica B, Kriegstein AR. 2011. A new subtype of progenitor cell in the mouse embryonic neoxortex. Nat Neurosci. 14:555-561.

Workman AD, Charvet CJ, Clancy B, Darlington RB, Finlay BL. 2013. Modeling transformations of neurodevelopmental sequences across mammalian species. J Neurosci. 33:7368-7383.

Table 1. Quantities of interest for mouse macaque and human. E indicates embryonic day.

	Mouse	Macaque	Human	Reference
Cortical surface area (cm^2)	1X 4.20	100X 287.68	1000X 2600.00	Rockel et al. 1980; Rakic 1995; Mountcastle 1997; Sultan 2002.
Brain volume (mm^3)	508.91	87896.00	1251847.00	Stephan et al. 1981; Badea et al. 2007.
Neocortex volume (mm^3)	169.61	63482.00	1006525.00	Stephan et al. 1981; Badea et al. 2007.
Neocortex/brain % volume	28.31 (rodentia)	72.22	80.40	Stephan et al. 1981; Clark et al. 2001.

Brain weight (mg)	416.00	87350.00	1508000.00	Herculano-Houzel 2009.
Cortical thickness (mm)	0.6 – 1	1.5 - 2.5	3-4	Rockel et al. 1980; Rakic 1995; Mountcastle 1997.
Supragranular layers neurons (%)	53 (hamster)	69 (owl monkey)	-	Charvet et al. 2015.
Adult brain neurons	70.89 million	6376.16 million	86.06 billion	Herculano-Houzel et al. 2006, 2007; Azevedo et al. 2009.
Adult neocortex neurons $\overline{N}$	13.69 million	1.71 billion	16.34 billion	Herculano-Houzel et al. 2006, 2007; Azevedo et al. 2009.
Ventricles (mm^3)	4.80	834.00	18732.00	Stephan et al. 1981; Badea et al. 2007.

Founder population P_0	375000[1]	-	-	Haydar et al. 2000.
Neurogenesis starts	E11	E40	E40	Rakic 1988; Rakic 1995; Dehay and Kennedy 2007.
Neurogenesis duration (days)	8	30-60 (60 in primary visual)	77	Rakic 1988; Rakic 1995; Dehay and Kennedy 2007.
Cell cycle length	E12: 10.2 hr; E16: 18.4 hr; average: 14.3 hr	E40: 23 hr; E60: 54 hr; E80: 27 hr; average: 34.67 hr	-	Takahashi et al. 1996; Kornack and Rakic 1998.
Gestational period (days)	18.5	165	270	Workman et al. 2013.

(1) 75% of pallial progenitors estimated by Haydar et al, 2000, discounting progenitors of the medial, lateral and ventral pallia.

Captions to Figures

Figure 1. Strategy space for modes of progenitor cell division defined on the right. P = progenitor. N = neuron. The progenitor cell population navigates the strategy space by balancing probabilities of committing to each of the three division types: symmetric proliferative (SymP), asymmetric neurogenic (aSymN) and symmetric neurogenic (SymN). The trajectory indicated by the arrow is an example of time-dependent strategy, initially increasing the prevalence of AsymN, while reducing SymP, and finally favoring SymN.

Figure 2. Time-dependent probability functions for the three types of division. A two-step linear strategy is hypothesized, with a shift from asymmetric to symmetric neurogenic division and concurrent reduction in self-amplifying divisions. Neurogenesis occurs in the time interval (t_0, t_F). Parameters α_0, α_S, β_F, t_S unequivocally determine the shape of the functions.

Figure 3. Representation in strategy space of the optimal mouse strategy for constant **(A)** and age-dependent **(B)** cell cycle length models.

Figure 4. Local sensitivity analysis of strategy parameters around the optimal mouse strategy (Eq. 5). Sensitivity is calculated according to the formula reported in the supplemental information. θ^* indicates the reference value for the parameter whose sensitivity is being calculated. By definition sensitivity corresponding to θ^* is zero.

Figure 5. (A) Top row: constant (solid line) and age-dependent (dashed line) cell cycle models for mouse, macaque, and human. Bottom row: Solutions $N(t)$ and $P(t)$ for the two alternative cell cycle models parameterised on the three species. **(B)** Founder population estimates for mouse, macaque, and human using the optimal mouse strategy for constant (o) and age-dependent (☆) cell cycle length models.

Figure 1

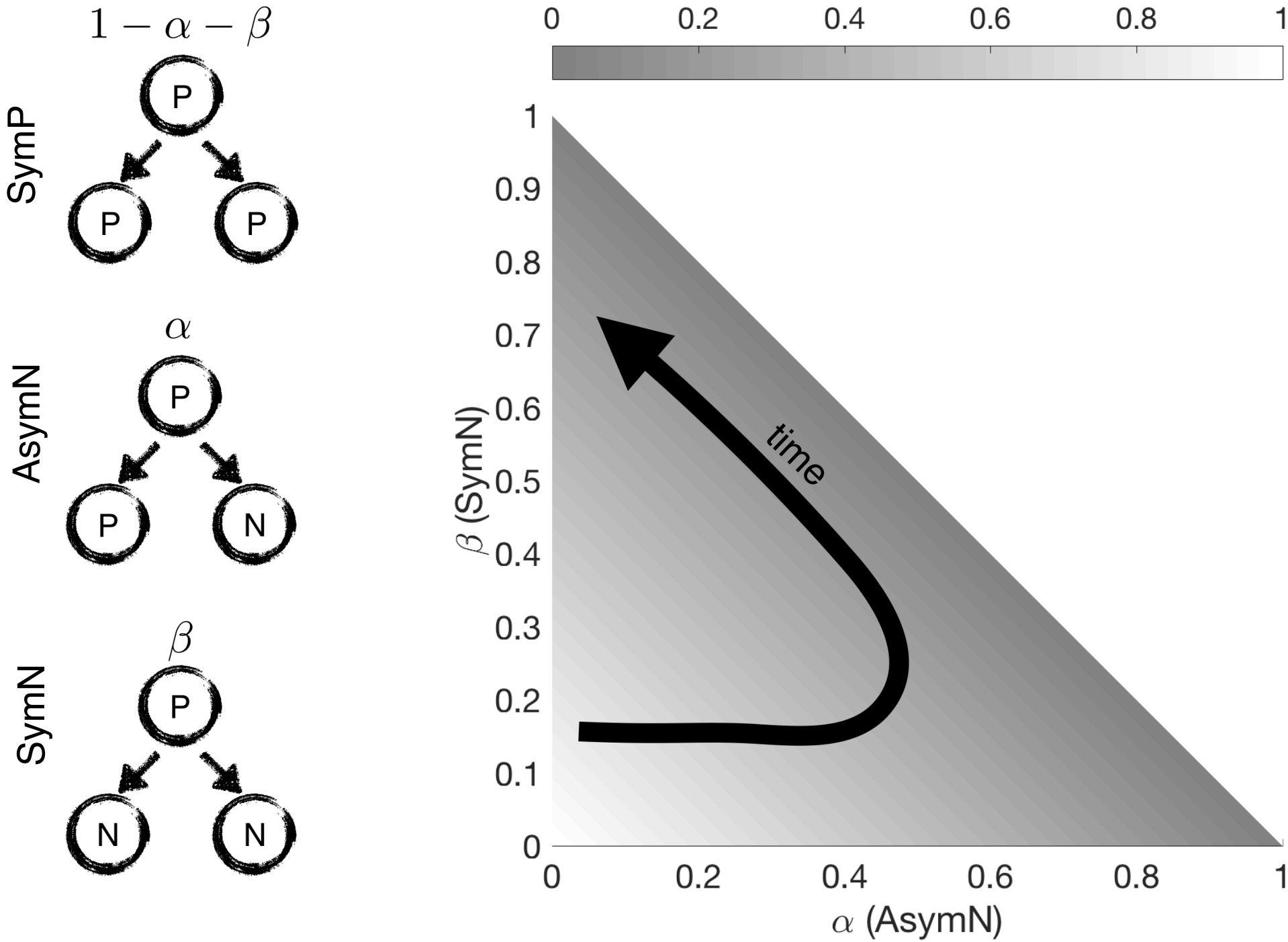

$1-\alpha-\beta$
SymP
P
P
P
α
AsymN
P
P
N
β
SymN
P
N
N
1-α-β (SymP)
0
0.2
0.4
0.6
0.8
1
β (SymN)
α (AsymN)
time

Figure 2

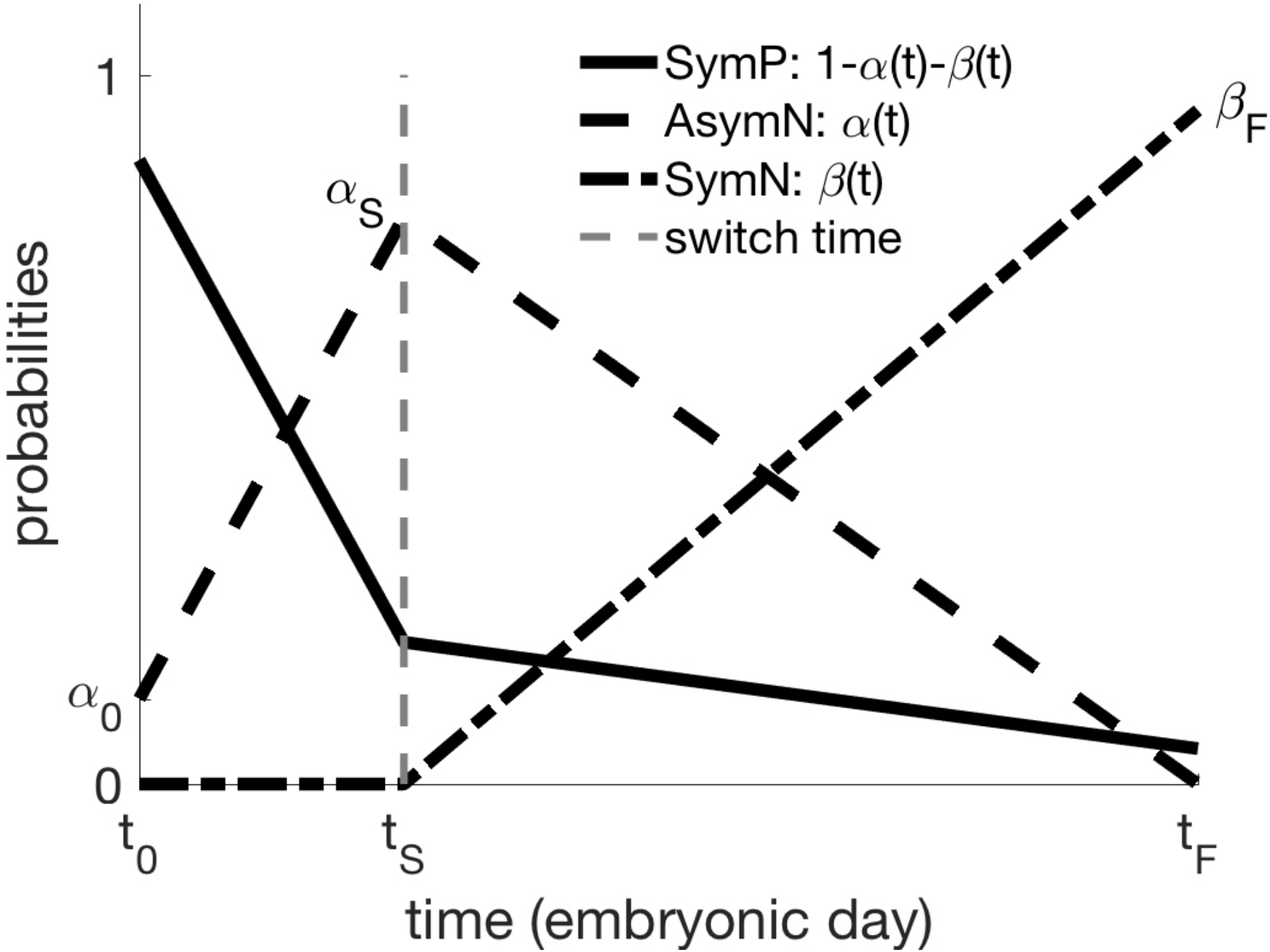

SymP: 1-α(t)-β(t)
AsymN: α(t)
SymN: β(t)
switch time
1
0
probabilities
α0
αS
βF
t0
tS
tF
time (embryonic day)

Figure 3

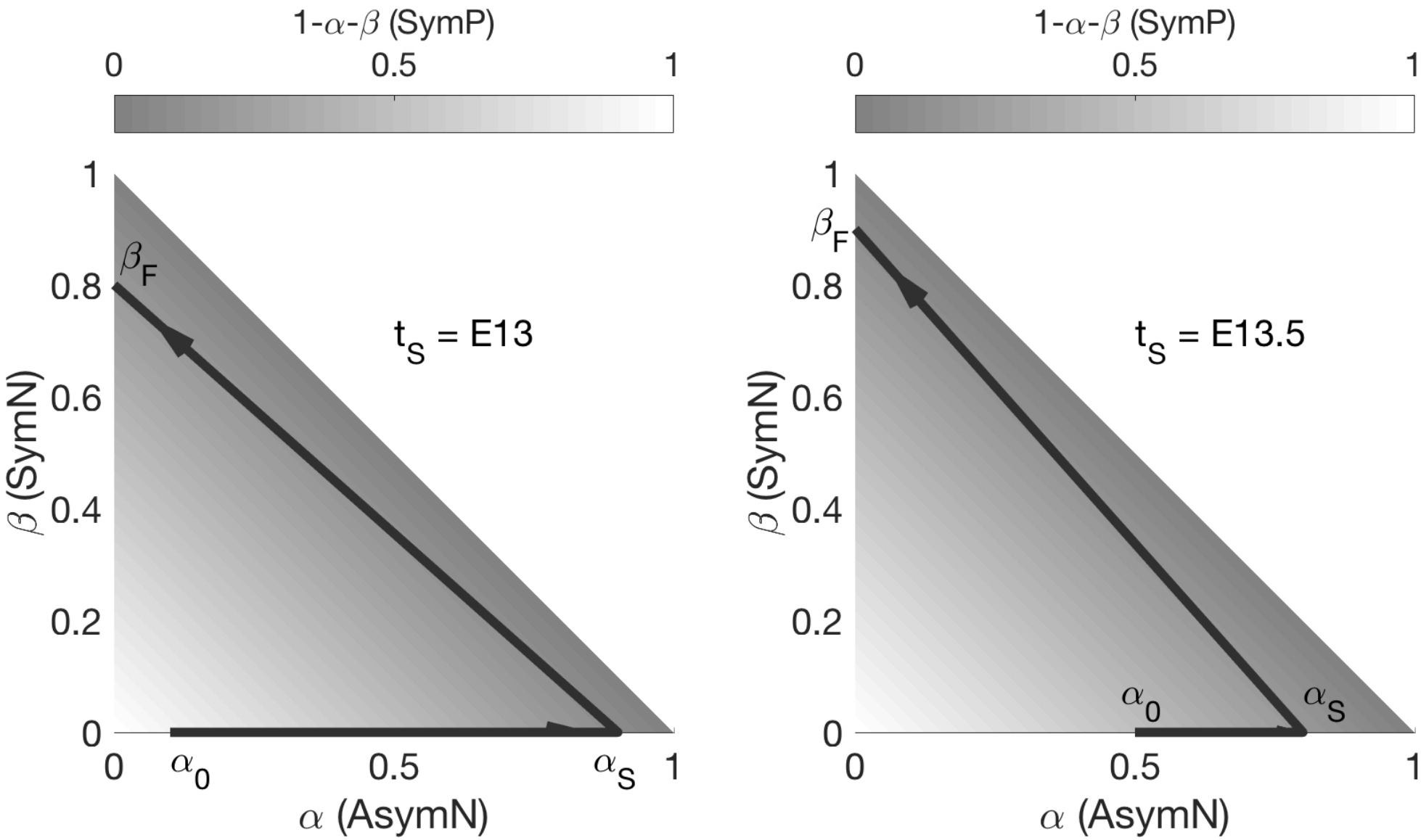


Figure 4

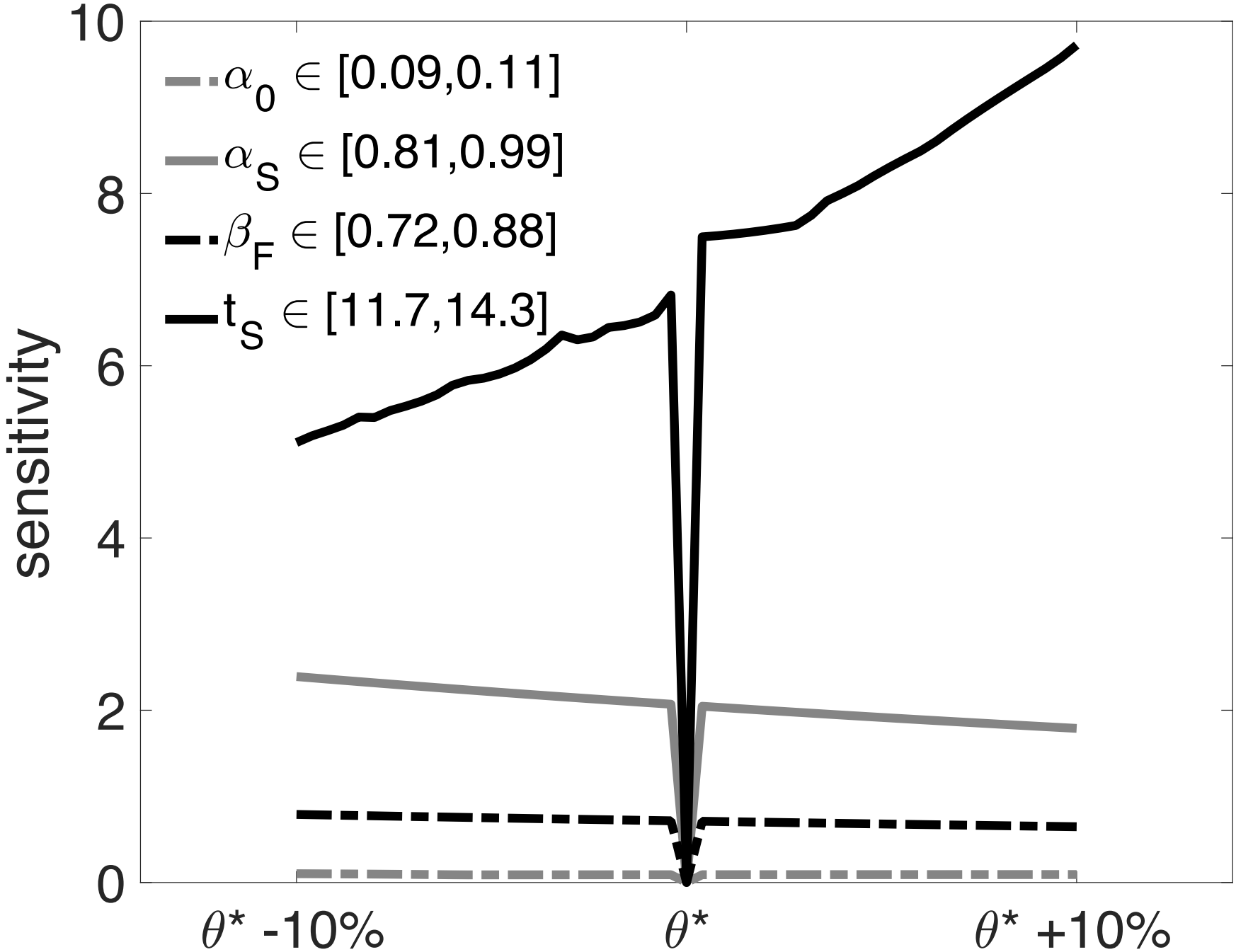
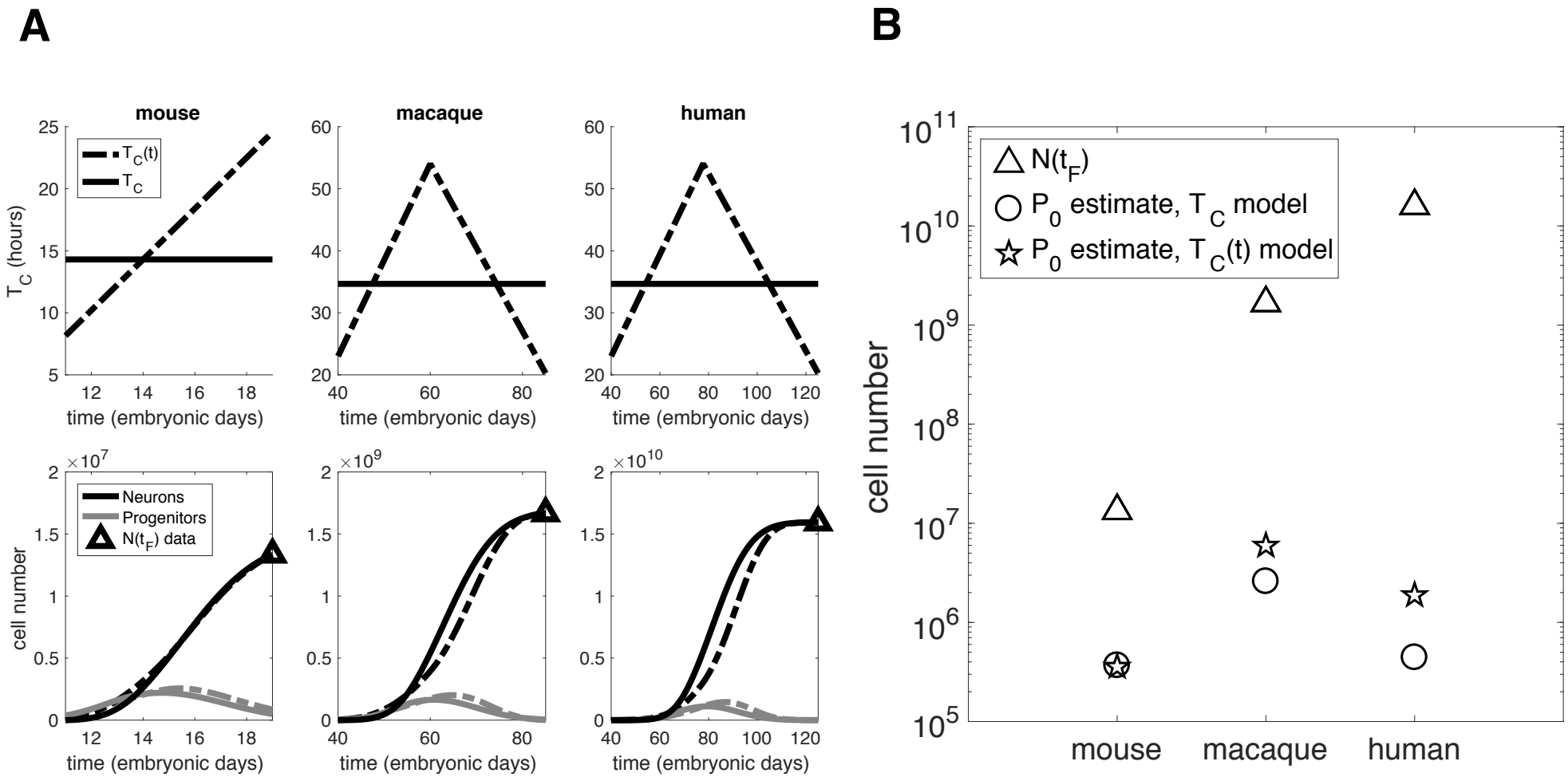

10
8
6
4
2
0
sensitivity
$\alpha_0 \in [0.09,0.11]$
$\alpha_S \in [0.81,0.99]$
$\beta_F \in [0.72,0.88]$
$t_S \in [11.7,14.3]$
θ^* -10%
θ^*
θ^* +10%

Figure 5


A
mouse
macaque
human
$T_C(t)$
T_C
T_C (hours)
time (embryonic days)
Neurons
Progenitors
$N(t_F)$ data
cell number
$\times 10^7$
$\times 10^9$
$\times 10^{10}$
B
$N(t_F)$
P_0 estimate, T_C model
P_0 estimate, $T_C(t)$ model
cell number
mouse
macaque
human