Mathematical Modeling of Cortical Neurogenesis Reveals that the Founder Population does not Necessarily Scale with Neurogenic Output

Noemi Picco, Fernando García-Moreno, Philip K. Maini, Thomas E. Woolley and Zoltán Molnár

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 species. In terms of cortical neuronal numbers, 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 in neuronal numbers found across species. By framing our hypotheses in rigorous mathematical terms, we are able to identify paths of neurogenesis that match experimentally observed patterns in mouse, macaque and human. Additionally, we use our model to identify key parameters that would particularly benefit from accurate experimental investigation. We find that the timing of a switch in favor of symmetric neurogenic divisions produces the highest variation in cortical neuronal numbers. Surprisingly, assuming similar cell cycle lengths in primate progenitors, the increase in cortical neuronal numbers does not reflect a larger size of founder population, a prediction that has identified a specific need for experimental quantifications.

Key words: asymmetric division, cell cycle of cortical progenitors, 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 interspecies differences or diseases, such as microcephaly. The most striking differences across species consist of final neurogenic output (i.e., the number of neurons produced by neurogenesis) and surface area (Krubitzer and Kaas 2005; Dehay and Kennedy 2007), 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 specialization 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; Dehay and Kennedy 2007; Florio and Huttner 2014; Imayoshi and Kagayama 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 and hence, cortical surface (Rakic 1988). However, understanding how all of these factors work together to ensure the cortex is correctly formed, and finding the key parameters that have the greatest impact on the neuronal numbers, is less amenable to verbal reasoning and more suitable for rigorous mathematical modeling.

An impressive number of comparative studies, quantifying and analyzing the main differences in cortical neurogenesis of mammalian species, is available (Molnár et al. 2006; Cheung et al. 2007; Herculeano-Houzel 2009; Martínez-Cerdeño et al. 2017). 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 neurogenic output by means of other observable characteristics, such as length of neurogenesis and brain volume, and do not investigate the underlying mechanisms (Mota and Herculeano-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 postmitotic 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; Betezau et al. 2013), and human (Fietz et al. 2010; Hansen et al. 2010). It is possible to identify 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 nonapically 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 a 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. Our model is minimal in the sense that we start from the simplest possible representation and introduce the time-dependency that is sufficient to capture the experimentally observed actual progenitor and neuronal numbers. 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 modeling 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 2 cellular populations: progenitors (P) and neurons (N). The former includes all premitotic neural progenitor cell types involved in neocortex development for a given species, for example, 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 population includes all postmitotic and permanently differentiated cells. In our model 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; García-Moreno et al. 2018) 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 modeled, however, it will be accounted for when integrating experimental quantifications into the model.

We consider 3 possible modes of progenitor cell division: self-amplifying division (SymP), generating 2 progenitors; asymmetric neurogenic division (AsymN), generating 1 progenitor and 1 neuron; symmetric neurogenic division (SymN), generating 2 neurons.

As we will see later, if these 3 types of division occurred with constant proportions, we would observe one of two possible outcomes: extinction or unlimited growth of the progenitor population. This type of model captures the eventual depletion of the progenitor population observed in adult cortical neurogenesis. Indeed, constant proportions of proliferation and differentiation were used to model similar dynamics in the adult hippocampus (Ziebell et al. 2014). It is known, instead, that these modes of proliferation are preferentially balanced at different stages of neocortical neurogenesis during development (Noctor et al. 2004; Götz and Huttner 2005), and this program seems to be conserved across species (Kornack 2000). Furthermore, the progenitor population is not depleted during embryonic neurogenesis, to allow for gliogenesis later in development (Kessaris et al. 2006; Mallamaci 2013).

We, therefore, define time-dependent probabilities describing the preferred mode of division in the progenitor population during neurogenesis: \( a(t) \) is the probability of AsymN at a given time \( t \), \( \beta(t) \) is the probability of SymN; hence \( 1 - a(t) - \beta(t) \) 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 3-dimensional space (AsymN, SymN, SymP) (Fig. 1).
Table 1 Quantities of interest for mouse macaque and human. E indicates embryonic day.

<table>
<thead>
<tr>
<th></th>
<th>Mouse</th>
<th>Macaque</th>
<th>Human</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cortical surface area</td>
<td>1x</td>
<td>100x</td>
<td>1000x</td>
<td>Rockel et al. (1980), Rakic (1995), Mountcastle (1997), Sultan (2002)</td>
</tr>
<tr>
<td>(cm²)</td>
<td>4.20</td>
<td>287.68</td>
<td>2600.00</td>
<td>Stephan et al. (1981), Badea et al. (2007)</td>
</tr>
<tr>
<td>Brain volume (mm³)</td>
<td>508.91</td>
<td>87 896.00</td>
<td>1 251 847.00</td>
<td>Stephan et al. (1981), Badea et al. (2007)</td>
</tr>
<tr>
<td>Neocortex volume (mm³)</td>
<td>169.61</td>
<td>63 482.00</td>
<td>1 006 525.00</td>
<td>Stephan et al. (1981), Clark et al. (2001)</td>
</tr>
<tr>
<td>Neocortex/brain % volume</td>
<td>28.31</td>
<td>72.22</td>
<td>80.40</td>
<td>Herculano-Houzel (2009)</td>
</tr>
<tr>
<td>Brain weight (mg)</td>
<td>416.00</td>
<td>87 350.00</td>
<td>1 508 000.00</td>
<td>Rockel et al. (1980), Rakic (1995), Mountcastle (1997), Charvet et al. (2015)</td>
</tr>
<tr>
<td>Cortical thickness (mm)</td>
<td>0.6–1</td>
<td>1.5–2.5</td>
<td>3–4</td>
<td>Workman et al. (2013)</td>
</tr>
<tr>
<td>Supragranular layers neurons (%)</td>
<td>53 (hamster)</td>
<td>71 (capuchin monkey)</td>
<td>86.06 billion</td>
<td>Herculano-Houzel et al. (2006), Azevedo et al. (2009)</td>
</tr>
<tr>
<td>Start of layer IV production: tₘ</td>
<td>E17</td>
<td>E64</td>
<td>E93</td>
<td></td>
</tr>
<tr>
<td>Adult brain neurons</td>
<td>70.89 million</td>
<td>6376.16 million</td>
<td>16.34 billion</td>
<td>Herculano-Houzel et al. (2006), Azevedo et al. (2009)</td>
</tr>
<tr>
<td>Adult neocortex neurons: N</td>
<td>13.69 million</td>
<td>1.71 billion</td>
<td>18 732.00</td>
<td>Stephan et al. (1981), Badea et al. (2007).</td>
</tr>
<tr>
<td>Ventricles (mm³)</td>
<td>4.80</td>
<td>834.00</td>
<td>18 732.00</td>
<td></td>
</tr>
<tr>
<td>Founder population: P₀</td>
<td>375 000⁷</td>
<td>–</td>
<td>–</td>
<td>Haydar et al. (2000)</td>
</tr>
<tr>
<td>Cell cycle length: Tₐ</td>
<td>E12:10.2 h; E16:18.4 h; average: 14.3 h</td>
<td>E40:23 h; E60:54 h; E80:27 h; average: 34.67 h</td>
<td>–</td>
<td>Takahashi et al. (1996), Kornack and Rakic (1998)</td>
</tr>
<tr>
<td>Gestational period (days)</td>
<td>18.5</td>
<td>165</td>
<td>270</td>
<td>Workman et al. (2013)</td>
</tr>
<tr>
<td>Predicted strategy—constant cell cycle model: (α₀, rₛ, Fₛ, tₛ)</td>
<td>(0.4, 0.5, 0.7, 14)</td>
<td>(0.1, 1, 0.7, 65.5)</td>
<td>(0.3, 1, 0.5, 94)</td>
<td></td>
</tr>
<tr>
<td>Predicted strategy—age-dependent cell cycle model: (α₀, rₛ, Fₛ, tₛ)</td>
<td>(0.2, 0.7, 1, 16)</td>
<td>(0.5, 0.8, 1, 63.5)</td>
<td>Tₐ x 1: (0.1, 0.9, 0.8, 92)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Tₐ x 1.5: (0.4, 0.8, 0.7, 84.5)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Tₐ x 2: (0.1, 0.4, 0.7, 56)</td>
</tr>
</tbody>
</table>

⁷75% of pallial progenitors estimated by Haydar et al. (2000), discounting progenitors of the medial, lateral, and ventral pallia.
The resulting system of ordinary differential equations (ODEs), describing the evolution in time of the progenitor \( P \) and neuron \( N \) populations is as follows:

\[
\begin{align*}
\frac{dP}{dt} &= \rho P(1 - \alpha(t) - 2\beta(t)), \quad P(t_0) = P_0, \\
\frac{dN}{dt} &= \rho P(\alpha(t) + 2\beta(t)), \quad N(t_0) = 0,
\end{align*}
\]

where \( \rho \) is the rate of division, and \( P_0 \) is the founder population. We define \( t_0 \) as the time of onset of neurogenesis, that is, when the first neuron is produced in the neocortex, and \( t_F \) as the final time of neurogenesis, and consider the time interval \( \lbrack t_0, t_F \rbrack \).

Cell death of postmitotic neurons is not explicitly modeled, but will be accounted for when integrating experimental estimates of final neurogenic 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.

A vast literature for cell cycle models is available (Weis et al. 2014), and different modeling approaches can be adopted for the description of the cell cycle at different scales of complexity. We are interested in the cell population dynamics, a higher scale level than the fine-grained study of progression through cell cycle phases and the molecular processes involved. Therefore, we simply account for the progenitor cell cycle length, \( T_C \), defining the growth rate as: \( \rho = \log 2/T_C \).

In the special case of constant \( \alpha \) and \( \beta \), 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: \( \alpha_0 \), the proportion of AsymN at onset of neurogenesis; \( \alpha_S \), the proportion of AsymN at the time of strategy switch; \( \beta_F \), the proportion of SymN at the end of neurogenesis; and \( t_s \), the time of the switch to a strategy favoring SymN. These 4 parameters unequivocally 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:

\[0 < \alpha_0 < 1, \quad 0 < \alpha_S < 1, \quad 0 < \beta_F < 1, \quad t_0 < t_s < t_F.\]
Within these constraints, any combination of the 4 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, that is, 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, \( T_c = 14.3 \) h (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 3 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 analyzing and explaining cortical neurogenesis through the lens of evolution. Specifically, our study focuses on cortical neuronal number, rather than brain size. This allows for a more straightforward comparison of proliferative and differentiative dynamics across species, without the need to introduce species-specific characterization of cell size and density.

Although data on a large number of other mammalian species exist, we focus on these 3 species that have been chosen in previous studies to effectively illustrate neocortex development across distinct branches of the phylogenetic tree (Rakic 1995).

We initially search the (AsymN, SymN, SymP) space to find the strategy, or strategies, matching the observed patterns of cerebral cortical neuronal numbers in each species. Then, we use each species’ estimated strategy 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

---

**Figure 3.** Representation in strategy space of the mouse, macaque, and human strategy for constant (A) and age-dependent (B) cell cycle length models. The lower panels, (C) and (D), respectively, show the corresponding estimates for the time of switch. Macaque and human times are rescaled on the equivalent mouse time scale, to facilitate a meaningful comparison of the timing of events across species.
and deeper cortical layers, while upper layers are significantly expanded in primates (Markram et al. 2004; Vasistha et al. 2015; Charvet et al. 2015, Table 1).

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

\[
\begin{align*}
\frac{d\vec{p}}{dt} &= \rho \vec{p}(1 - a(t) - 2b(t)), \\
\frac{d\vec{n}}{dt} &= \rho \vec{p}(a(t) + 2b(t)), \\
\end{align*}
\]

where \(\vec{p}\) and \(\vec{n}\) 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 strategy’ as the 4-tuple \((a_0, a_5, \beta_f, t_3)\) that best approximates the following:

\[
\tilde{N}(t_f) = \tilde{N}(t_0) + \varepsilon,
\]

with minimum error \(\varepsilon\). Here, \(\varphi\) is the ratio of neurons in deeper layers (V and VI) to give a more accurate representation \(P_{tt}, P_{t}t, N_t\) and \(f\), obtained by interpolation of the data from different types of division \((a_0, a_5, \beta_f)\). This sensitivity, as well as lack of experimental quantification, illustrates a demand for future experimental research into quantifying, or at least characterizing, this time scale.

Species-Specific Strategies

We now move from the mouse, where neurogenesis has been extensively studied, to primates, where many key processes are relatively poorly characterized. Despite the lack of experimental quantification, we proceed with the best approximation found in the literature, and later test the robustness of our results against these simplifying assumptions and approximations.

When evaluating the macaque strategy, we implement the method previously presented for the mouse, adjusting the timing of events \((t_0, t_3, t_0)\), the cell cycle length, \(T_c\), and the ratio of deeper layer neurons, \(\varphi\). Finally, in order to evaluate the human strategy, values for parameters \(T_c\) and \(\varphi\) are approximated using the macaque parameters, as they cannot be found in the literature. A summary of model parameters adopted, and resulting strategies for each species, are reported in Table 1.

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 neurogenic 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 nonintuitive implications, in terms of neurogenic 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) (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 each species, we obtained a new set of strategies, reported in Table 1 and represented in Figure 3B,D. Intuitively, the delay in the time of switch to neurogenic divisions predicted for the mouse strategy increases the pool of progenitors to compensate for a decrease in the frequency of division at late mouse neurogenesis. Conversely, the nonmonotonic cell cycle regulation throughout primate neurogenesis results in an earlier time of switch for both macaque and human.

 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.
Acquisition of founder population count

(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 neurogenic output \( N(t) \) 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 robust to variations in these quantities.

Given \( N(t) \), 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 and Picco et al. 2017) to obtain an estimate of \( P_0 \). This method gives the best fit to match the neurogenic output \( N(t) \), given the model of choice (constant, or age-dependent, cell cycle length), and the corresponding strategy. Figure 5B shows the estimates of \( P_0 \) for the 3 species and 2 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 neurogenic output is not increasing. Specifically, the founder population estimated for the human cortex is smaller than those values predicted for both mouse and macaque, despite producing a considerably larger number of neurons.

Solutions \( N(t) \) and \( P(t) \) for the parameterized models are shown in Figure 5A, bottom row. It is worth noting that, despite the estimation being obtained by fitting one data point only, \( N(t_P) \), for a given strategy, the time course of the model output does not vary significantly when comparing the 2 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. The qualitative behavior of \( P(t) \) obtained with the parameterized model recapitulates the temporal dynamics expected of the progenitor population, as it reflects the observed expansion followed by shrinking of the proliferative zone (Bystron et al. 2008). More importantly, at the end of neurogenesis we expect a nondepleted progenitor population, since gliogenesis, occurring in a consecutive window of time, will be carried out by the remaining pool of progenitors, \( P(t_F) \).

To test the robustness of our result and challenge the counterintuitive prediction regarding the founder population, we introduce variations in the assumptions that could have skewed the results. Specifically, we repeat the estimations for varied amounts of interneuronal migration and postneurogenesis cell death. Even an unrealistically large rate of postneurogenesis cell death does not alter the prediction of lower human founder population (Suppl. Fig. S3A). Similarly, the prediction still holds even when increasing the rate of interneuronal migration to 50% (Suppl. Fig. S3B), although a more realistic 20% rate has been suggested (Petanjek et al. 2009). 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 2 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. Additionally, a different frequency of cell divisions will alter the strategy adopted. Therefore, we re-evaluate the human strategy for longer cell cycles, up to twice the baseline value for macaque progenitor \( T_C \).
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 neurogenic 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 p27Kip1, a cyclin-dependent kinase inhibitor, was found to substantially reduce neurogenic output in supergranular layers (i.e., late-stage neuronal production; Tarui et al. 2005). Although this constitutes only circumstantial evidence, p27Kip1 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 characterized the species-specific strategies, 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 2-fold increase in neurogenic output (Fig. 5). This prediction holds even for large variations in key parameters, such as rates of postneurogenesis cell death and interneuronal migration (Supplementary Fig. S3).

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 modeling, 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.

Mora-Bermúdez et al. (2016) estimate an approximately 1.3x amplification in progenitor cell cycle length in human organoids, with respect to macaque. To the best of our knowledge, however, 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 postneurogenesis 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 2 factors has yet to be established. By proposing 2 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 2 models, determining the need (or not) of age-dependent cell cycle to quantitatively recapitulate the population dynamics.

Conclusions

In 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 modeling 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 more data become available. For example, radial migration is only modeled implicitly, by assuming that positioning in the radial direction is related to time of birth (Rakic 1974; McConnell et al. 1991), fundamental to establishing a relation between the timing of neurogenesis and the distribution of neurons between upper and deeper cortical layers (equation 4). 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; Dehay and Kennedy 2007). However, such differences are relatively minor compared with the cross-species variation that we intended to address with this work. Furthermore, our model explains interspecies variation in neurogenic 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 modeling 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. 2018). 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 modeling 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 modeling 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 modeling 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 (1988). We found that the timing of shift from self-replicating to neurogenic divisions is key to obtaining the correct neuronal numbers 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 ‘developmental strategy’ adjusting the duration of neurogenesis, thus requiring an approximately conserved size of founder population. Indeed, a founder population size that scales with the neurogenic output would be prohibitively large and inefficient, compared with 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 that correspond to brain diseases.

Supplementary Material

Supplementary material is available at Cerebral Cortex online.

Funding

St John’s College Research Centre Grant to T.E.W., P.K.M. and Z.M. in collaboration with F.G.-M. The work in the laboratory of Z.M. was supported by grants from MRC, Wellcome Trust, Royal Society and St John’s College Research Centre. P.K.M. and T.E.W. would like to thank the Mathematical Biosciences Institute (MBI) at Ohio State University, for partially supporting this research. National Science Foundation (Grant number DMS1440386) to MBI F.G.-M. holds an IKERBASQUE Research Fellowship.

Notes

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 article. We would like to thank the 3 reviewers for their valuable feedback and constructive comments. Conflict of interest: none declared.

References

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...


Charvet CJ, Cahalane DJ, Finlay BL. 2015. Systematic, cross-


