

## 1 **Adaptive Clonal Expansion Shapes Brain Development**

2

3 Giulia Di Muzio<sup>1,2,8</sup>, Sarah Benedetto<sup>3,4,8</sup>, Li-Chin Wang<sup>1</sup>, Lea Weber<sup>4</sup>, Franciscus van der  
4 Hoeven<sup>5</sup>, Brittney Armstrong<sup>5</sup>, Hsin-Jui Lu<sup>1,2</sup>, Jana Berlanda<sup>1,2</sup>, Verena Körber<sup>6</sup>, Nina  
5 Claudino<sup>4</sup>, Michelle Krogemann<sup>4</sup>, Thomas Höfer<sup>3,4,9</sup>, Pei-Chi Wei<sup>1,2,7,9</sup>

6

7

8 1. German Cancer Research Center (DKFZ), Brain Mosaicism and Tumorigenesis, Heidelberg, Germany

9 2. Ruprecht Karl University of Heidelberg, Faculty of Bioscience, Heidelberg, Germany

10 3. German Cancer Research Center (DKFZ), Division of Theoretical Systems Biology, Heidelberg, Germany

11 4. Ruprecht Karl University of Heidelberg, Combined Faculty of Mathematics, Engineering and Natural  
12 Sciences, Heidelberg, Germany

13 5. German Cancer Research Center (DKFZ), Transgenic Service, Heidelberg, Germany

14 6. MRC Molecular Haematology Unit, Weatherall Institute of Molecular Medicine, Oxford, United Kingdom

15 7. Ruprecht Karl University of Heidelberg, Faculty of Bioscience, Interdisciplinary Center for Neuroscience  
16 (IZN), Heidelberg, Germany

17 8. Equal contribution

18 9. Correspondence to: [t.hoefer@dkfz-heidelberg.de](mailto:t.hoefer@dkfz-heidelberg.de), [p.wei@dkfz-heidelberg.de](mailto:p.wei@dkfz-heidelberg.de)

### 19 **Abstract:**

20 Embryonic neural stem/progenitor cells (NSPCs) exhibit remarkable proliferative  
21 plasticity, allowing them to fully recover neuronal populations even after substantial cell loss  
22 <sup>1,2</sup>. However, it remains unclear whether all embryonic NSPCs respond to brain lesions. To  
23 address this, we developed a mouse model to investigate NSPC proliferation dynamics,  
24 hypothesizing that the loss of progenitor cells would induce fitness competition among NSPCs.  
25 In this model, half of the founder NSPCs were ablated using diphtheria toxin A at the onset of  
26 neurogenesis, yet the surviving cells regenerated a brain containing all neuronal types within  
27 five of the total twenty embryonic days. Analysis of allelic variants revealed overrepresented  
28 somatic variants, indicating that only a small fraction of NSPCs underwent significant clonal  
29 expansion during early neurogenesis. Modelling proliferation dynamics predicted that as few  
30 as 10% of NSPCs could produce 83% of neurons by the time of birth. Single nucleotide  
31 substitution analysis suggested a potential link to oxidative metabolism in some of the  
32 expanded clones. Moreover, single-cell transcriptomics showed delayed development and a  
33 reduced NSPC pool as consequences of adaptive clonal expansion. Our findings highlight  
34 that NSPC exhibit varying expansion potential and that adaptive clonal expansion indirectly  
35 altered neuronal cell composition in the brain.

36

37 **Keywords:** clonal expansion, brain development, modelling, cell cycle dynamics, neural stem  
38 cells, genome mosaicism.

39 **Data Availability:** The scRNA-seq data are deposited at the GEO (GSE281055) and the WGS  
40 datasets at NLM-NCBI (Bioproject PRJNA1209637).

41

42

## 1 **Background**

2 The proliferation of neural stem/progenitor cells (NSPCs) is crucial for laying the  
3 foundation of mammalian brain development. This process generates a large pool of neural  
4 progenitors that will differentiate into neurons and glial cells in the brain. It also shapes the  
5 basic structure and size of the developing brain<sup>3,4</sup>. Moreover, the early expansion of NSPCs  
6 influences the final size of specific brain regions<sup>5,6</sup>. In certain species, this expansion  
7 contributes to the formation of cortical folds, which increase the cortical surface area within  
8 the confined space of the skull<sup>3,7</sup>.

9 The proliferation dynamics of NSPCs directly determine the number of progenies each  
10 cell produces. As NSPCs replicate their DNA during each proliferation cycle, all progenies  
11 within the same neuronal clone inherit the DNA blueprint as their cell of origin. Errors in this  
12 blueprint can be passed down to neurons, potentially contributing to neurological disorders<sup>8,9</sup>.  
13 For example, in some cases, errors present in only 1% of neurons are sufficient to cause focal  
14 cortical dysplasia<sup>10</sup>. Therefore, investigating the expansion capacity of NSPCs is critical for  
15 understanding how erroneous genetic blueprints are distributed across different brain regions.

16 Recent studies on the neuronal clone size in the human brain have revealed that  
17 approximately 90-200 neural progenitor cells are present at the time of midline  
18 establishment<sup>11</sup>. In hindsight, one might expect an equal representation of neuronal clones  
19 derived from each individual founder NSPC. However, somatic variants-dependent lineage-  
20 tracing studies suggested otherwise. In up to a quarter of human brains, neuronal clones  
21 comprising, bias in proportion, for more than 8% of neurons were found<sup>12</sup>. Additionally, a study  
22 analysing over 130 human brains has suggested evidence for selective expansion of neuronal  
23 progenitors in normal human brains<sup>13</sup>. These findings indicate that NSPC possess intrinsic  
24 capacities to generate a larger number of neurons. However, within the tightly constrained  
25 period of embryonic neurogenesis, the extent to which this potential can be exploited during  
26 early neurogenesis remains unknown.

27 Although it has been speculated that NSPCs exhibit differential proliferative  
28 abilities<sup>14,15</sup>, answering this question remains challenging. Existing experiment approaches  
29 have either focused on collective population-level proliferative behavior<sup>14,15</sup> or imposed all-or-  
30 none conditions<sup>16-19</sup>, making it difficult to capture the heterogeneous nature of NSPC  
31 proliferation. Additionally, while alterations in the proliferation dynamics of neural progenitor  
32 cells can directly impact the spread of brain-specific mutations, it remains unclear whether  
33 clonal expansion dynamics can exceed pathological thresholds and what risk factors may  
34 trigger unbalanced clonal expansion. Therefore, an animal model that allows NSPCs to exhibit  
35 their natural proliferative dynamics is needed.

1 We present a mouse model in which a reduction in NSPCs at the onset of  
2 neurogenesis promotes selective clonal expansion. We analysed *in vivo* changes in cell  
3 proliferation rates and examined neuronal differentiation status at the end of embryonic  
4 neurogenesis. We used whole-genome sequencing to evaluate clonal expansion, by  
5 accessing variant allele frequency distributions to infer the clonal population size. Our analysis  
6 revealed the presence of dominant clones, with the largest comprising 50% of neuronal cells,  
7 contributing significantly to the overall neuronal population. Furthermore, we simulated clonal  
8 expansion dynamics throughout embryonic neurodevelopment using a pool-size-weighted  
9 mathematical model. Finally, through transcription factor activity profiling and co-regulated  
10 gene expression analysis in single cells, we concluded that adaptive clonal expansion reflects  
11 a developmental crisis, potentially posing risks to brain health.

12

## 13 **Results**

### 14 **Creating a mouse model allowing NSPCs exert their expansion potential**

15 The ability of the NSPCs to survive and proliferate can be influenced by stochastic  
16 developmental bottlenecks, such as high frequency of programmed cell death<sup>20</sup>, where unfit  
17 cells were eliminated through programmed cell death<sup>17,21</sup>. We hypothesize that embryonic cell  
18 death in brain creates a development bottleneck, encouraging NSPCs to expand.

19 To test this hypothesis, we created chimeric mice using host blastocysts encoding for  
20 the attenuated diphtheria toxin A (DTA)-expression cassette at the *ROSA26* locus (*R26-DTA*).  
21 The DTA expression was controlled by removing the 5' end floxed stop cassette by the Cre  
22 recombinase, which was controlled by a transgenic rat *Nestin* promoter (*NesCre*)<sup>22,23</sup>. Prior  
23 research has shown that the rat *Nestin* promoter only activates in NSPCs - radial glial (RG)  
24 cells and the intermediate progenitor (IP) cells - at embryonic day (E) 12.5, but not at time  
25 points before the emergence of RG cells<sup>22</sup>. In addition, although DTA leakiness was known<sup>24</sup>,  
26 it did not affect our ability to address the expansion capacity of surviving NSPCs. Thus,  
27 selectively eliminating NSPCs by internally expressing DTA, represents an appropriate  
28 approach to address our question. In the following context, we refer to blastocysts containing  
29 *NesCre* and *R26-DTA* expression cassettes as *NesCre::R26-DTA* embryos (**Fig. 1A**). The  
30 control blastocysts lacking Cre recombinase expression are referred to as *R26-DTA* embryos.

31 To benchmark DTA-induced NSPC death *in vivo*, we crossed male mice hemizygous  
32 for *NesCre* with female mice homozygous for the *R26-DTA* transgene (**Fig. S1A**). This  
33 approach produced about 50% *NesCre::R26-DTA* (n=16) and 50% control *R26-DTA* embryos  
34 (n=14), indicating that DTA expression did not affect embryonic mortality. Histological analysis  
35 revealed that the cerebral area was largely ablated in the *NesCre::R26-DTA* embryos at E12.5

1 (Fig. S1B) and E15.5 (Fig. 1B), confirming the failure of neurogenesis. Due to lack of brain  
2 tissue, the head structure of the *NesCre::R26-DTA* embryos collapsed during sample  
3 preparation for histology. In addition, immunofluorescent staining confirmed that a substantial  
4 lack of RG cells (PAX6-positive, Fig. S1C,D) and IP cells (TBR2-positive, Fig. S1C,D) in the  
5 brains of the E15.5 *NesCre::R26-DTA* embryos.

6 Next, to deplete part of the NSPCs, we created chimeric embryos containing a mixture  
7 of wild-type and *NesCre::R26-DTA* cells. This was achieved by blastocyst complementation,  
8 initially developed for immunology research<sup>25</sup> and was applied for telencephalon development  
9 studies<sup>1,2</sup>. In brief, we injected wild-type mouse embryonic stem (ES) cells, which carry the  
10 constitutively expressed mCherry transgene in their genomes, into blastocysts to create  
11 *NesCre::R26-DTA* chimeras (Fig. 1C). In these chimeric embryos, half of the NSPCs were  
12 wild-type, while the other half carried the *NesCre::R26-DTA* transgene cassettes and  
13 subsequently died upon DTA expression. These embryos were referred to as *NesCre::R26-*  
14 *DTA* chimeras. In parallel, we injected wild-type mouse ES cells into *R26-DTA* blastocysts,  
15 creating *R26-DTA* chimeras as controls (Fig. 1C).

16 We found that the injected mouse ES cells were incorporated into the central nervous  
17 system and reconstituted the brain and spinal cord tissue in the *NesCre::R26-DTA* chimeras  
18 (Fig. 1D). These data suggest that NSPCs in the *NesCre::R26-DTA* embryonic brains  
19 accelerate their proliferation rate, at least transiently. Given that our focus was on brain  
20 development, we did not investigate the spinal cord phenotype in this article. As expected,  
21 over  $91 \pm 5.3$  % (mean  $\pm$  SD) of forebrain cells were mCherry-positive in *Nestin-Cre::R26-DTA*  
22 chimeras, significantly higher than  $48 \pm 16.3$  % (mean  $\pm$  SD) mCherry-positive forebrain cells  
23 in *R26-DTA* chimeras (Fig. 1E). This bias was not due to differential ES cell contribution to  
24 the embryo itself, as the mCherry-positive cell proportion was comparable in tail fibroblasts  
25 isolated from *NesCre::R26-DTA* ( $67.5 \pm 12.3$ %; mean  $\pm$  SD) and *R26-DTA* chimeras ( $55.2 \pm$   
26  $11$ %; mean  $\pm$  SD) (Fig. S1E). We reasoned that the difference in the mCherry-positive ratio  
27 between tail fibroblasts and neuronal cells in the *Nestin-Cre::R26-DTA* chimeras resulted from  
28 complementation by the surviving wild-type NSPCs in the embryonic brains.

29 Active complementation also reconstituted all neuronal cell types at the end of  
30 embryonic neurogenesis. We analysed the gene expression for 61,080 (*NesCre::R26-DTA*)  
31 and 41,788 (*R26-DTA*) mCherry-expressing single cells isolated from the forebrain and the  
32 mid- and hindbrain (Fig. S1F,G; Table S1). In the *NesCre::R26-DTA* chimeras, the mCherry-  
33 positive cells represented all the neuronal cell types. We concluded that, in response to  
34 induced cell death, the living RG cells exerted their expansion potential to meet the demands  
35 of brain development.

1

## 2 **Selected clonal expansion in the *NesCre::R26-DTA* chimera's brain**

3       When NSPCs expands faster, they produce more progeny. This expansion can be  
4 driven either by all neuronal clones equally or by a specific subset of clones during early  
5 neurogenesis. Somatic mutations provide a way to trace neuronal clones, as any mutation  
6 that arises in the NSPC will propagate to all the neurons derived from it, allowing to associate  
7 specific DNA alterations to specific neuronal clones. To investigate which of the scenarios  
8 occurred during brain complementation, we performed deep whole-genome sequencing  
9 (WGS; mean coverage of 198x; **Table S2**) on the genomes of mCherry-positive cells isolated  
10 from the forebrain of three *R26-DTA* and six *NesCre::R26-DTA* chimeras at E17.5. We then  
11 analysed the frequency of small variants, specifically single nucleotide variants (SNVs) and  
12 small insertions and deletions (indels), acquired after the germ-layer separation at around  
13 E6.5<sup>26</sup>. To exclude variants present in the ES cell genomes or those that arose before germ-  
14 layer separation, we sequenced the genomes of mCherry-positive ES cells (55X coverage)  
15 and embryo-matched mCherry-positive tail fibroblasts (mean coverage of 76X; Table S2),  
16 which are derived from the mesoderm. Variants shared between tail fibroblasts and forebrains  
17 or ES cells and forebrains were removed for downstream analyses (**Fig. 2A**). Considering that  
18 only fifteen ES cells were injected into each blastocyst and about 70% of the ES cells  
19 contributed to the embryo proper<sup>27,28</sup>, we expect to identify variants that are specific to each  
20 ES cell at a minimum frequency of 10%. With deep sequencing coverage of forebrain cells,  
21 we expect to identify and remove ES cell-specific variants from the forebrain genomes,  
22 including those originating from culture artifacts.

23       To track clonal evolution after germ-layer separation, we analysed the allele  
24 frequencies of variants that were absent in both tail fibroblasts and ES cell genomes. For  
25 variants with allele frequencies (VAF) above 2%, we observed a characteristic  $1/\text{VAF}^2$   
26 distribution, which is typical of subclonal variants<sup>29</sup> - mutations found in only a subset of the  
27 sequenced cells - regardless of the chimera type (**Fig. 2B** at lower frequencies; **Table S3**).  
28 These variants represent neutral mutations that accumulated in cells between E6.5 and E17.5.  
29 However, in *NesCre::R26-DTA* chimeras only, we identified an unusual cluster of subclonal  
30 variants at approximately 10-25% frequency, which was more pronounced in chimera m18  
31 and less so in chimera m7 (**Fig. 2B**, right panels). Since mCherry-positive cells had a diploid  
32 genome, clonal mutations (those present in the most recent common ancestor) were expected  
33 to appear at a VAF of 50%, reflecting a mutation in one of the two alleles. The presence of  
34 these subclones at a VAF of 25% suggested that a dominant neuronal clone made up half of  
35 the sequenced cerebral cells, which is extraordinary.

1           Given that these subclonal clusters of variants were found only in to the *NesCre::R26-*  
2 *DTA* chimeras, they likely represent an early clonal expansion triggered in response to the  
3 selectively induced NSPCs death. A complementary multivariate analysis with a variational  
4 Binomial mixture model (VIBER<sup>30</sup>) confirmed that four of the six *NesCre::R26-DTA* chimeras  
5 (m2, m7, m8 and m18) developed distinct clones (**Fig. S2A**, cluster C2), separated from the  
6 neutral low-frequency variants (**Fig. S2A**, cluster C1). To elucidate whether these subclones  
7 could be putatively driven by shared mutations, we analysed the nonsynonymous SNVs and  
8 indels at the coding regions. We found no known driver somatic mutations shared between  
9 *NesCre::R26-DTA* chimeras. In addition, the clonal mutations we found could not explain the  
10 expansion phenotype, suggesting they are passenger mutations (**Fig. 2B**). For the remaining  
11 chimeras, namely m10 and m17, VIBER did not recognize a subclone, likely due to the  
12 population being a mix of clonal lineages.

13           In addition to WGS, we also performed scRNA-seq on an *R26-DTA* chimera (m12) and  
14 an *NesCre::R26-DTA* chimera (m17). We analysed the somatic SNV in single cell transcripts  
15 in order to validate if subclonal SNVs propagated into the neuronal cell. Despite the limitations  
16 in sampling and sequencing coverage of the scRNA-seq assay, we found five clonal SNVs  
17 across neuronal lineages in the *R26-DTA* and two clonal SNVs in the *NesCre::R26-DTA*  
18 chimeras (**Fig. 2C**, **Table S4**). In the forebrain of the *NesCre::R26-DTA* chimera, a subclone  
19 containing an intronic SNV at the *Nav2* gene locus was found in glioblasts, neuroblasts, and  
20 GABAergic as well as glutamatergic neurons. The SNV at the first exon of *Ddx1* was found in  
21 glutamatergic neuronal populations in both forebrain and mid-/hindbrain, indicating this clone  
22 was established at the beginning of neurogenesis (**Fig. 2C**). The number of RG and IP cells  
23 expressing *Nfia*, *Nav2*, and *Ddx1* were low (**Fig. S2B**), implying under-sampling of the  
24 founders carrying these SNVs.

25           To eliminate the possibility that the observations resulted from intrinsic nucleotide  
26 variations present in the ES cells' genome, we performed reversed complementation  
27 experiments. Instead of injecting wild-type ES cells into the *NesCre::R26-DTA* blastocysts, we  
28 injected *NesCre::R26-DTA* ES cells into DsRed-expressing wild-type blastocysts (**Fig. S2C**).  
29 In the reversed complementation setting, the surviving cells were derived from the wild-type  
30 blastocysts and they globally expressed DsRed (see Material and Methods for more details).  
31 All three embryonic brains that were deeply whole-genome sequenced (mean coverage of  
32 180x for the forebrains and 31x for the tail fibroblasts; **Table S2**) displayed an  
33 overrepresentation of variants in the DsRed-positive wild-type neuronal cells at E17.5 (**Fig.**  
34 **S2D**). Hence, we concluded that the selective clonal expansion that arose during  
35 neurogenesis was independent of pre-existing nucleotide variants in the ES cells' genome.

1           In summary, upon induced cell death, the surviving RG cells compete with one another  
2 through active clonal expansion, independent of single nucleotide variation or indels present  
3 in genes or regulatory elements driving the process.

4

## 5 **Modelling cell cycle dynamics predicts neuronal compensation from few RG clones**

6           The observation that half of the brain originated from a single dominant clone was  
7 remarkable. This finding prompted us to develop a mathematical framework to predict clonal  
8 expansion dynamics throughout the embryonic neurogenesis period *in silico*.

9           Our model simulated the normal proliferation dynamics of NSPC populations,  
10 comprising of the neuroepithelial cells (NECs), RG cells, IP cells, and the neurons they  
11 produce. This model simulated three differentiation steps of the developing mouse brain. First,  
12 at E12.5, the model assumed that NECs maintain a constant population size while gradually  
13 differentiating into RGs. Second, from E12.5 to E18.5, RGs either self-renew or differentiate  
14 into IPs. Third, from E12.5 to E18.5, IPs either self-renew or produce two neurons. To simulate  
15 the transition from proliferation to differentiation, the model decreased the self-renewal rate of  
16 RGs with the total cell count, eventually approaching zero as neurogenesis progresses. This  
17 regulation, based on the total number of progenitors and neurons, ensured that brain  
18 development reaches its expected final cell counts (**Fig. 3A**).

19           The model was calibrated using observed cell count data<sup>31</sup> from the embryonic mouse  
20 cortex (**Fig. 3B**). It accurately captured NSPC dynamics and neuron output (**Fig. 3B**, solid  
21 lines). Using these rates, we then simulated cell count dynamics under varying levels of  
22 NSPCs ablation, ranging from mild (30% and 50%) to severe (90%). Under mild ablation  
23 conditions (**Fig. 3B**, left and central panel), the model predicted a half-day delay in RG and IP  
24 recovery, accompanied by an accelerated proliferation rate of the surviving progenitors.  
25 Despite this delay, neuron numbers closely matched observed values by E18.5. Even under  
26 severe ablation (**Fig. 3B**, right panel), the model predicted a delayed recovery of RG and IP  
27 cell populations and suggested neuron production could still reach 83.3% of normal levels by  
28 E18.5. These predictions indicate that a nearly complete mouse brain could still develop  
29 starting from just 10% of the original RG cell population at the onset of neurogenesis.

30           Next, we simulated how initial clone size and proliferation advantage affect cell  
31 competition during neurogenesis (**Fig. 3C**). We started with a clonal size comprising either  
32 8%, 10%, or 12% of the progenitor pool at E12.5 (**Fig. 3C**), adjusting for compensatory  
33 reductions in the proliferation rates of non-clonal cells to maintain overall progenitor numbers.  
34 Our model predicted that the proliferation advantage plays a critical role in clonal contribution.  
35 Clones with higher proliferation rates expanded exponentially, hence allowing even smaller

1 initial clones to achieve dominance. In contrast, lower proliferation advantages slowed clonal  
2 expansion, resulting in a more balanced distribution between clonal and non-clonal  
3 progenitors.

4 To experimentally evaluate the cell cycle length, we performed EdU-BrdU pulse-chase  
5 labelling *in vivo* at E15.5<sup>32,33</sup>. We chose E15.5 as time point because it exhibited the greatest  
6 differences in NSPC counts when ablation was simulated *in silico*. In the developing mouse  
7 brains at this stage, the majority proliferating cells were IPs. We found that the cell cycle length  
8 of IPs was reduced from 10 to 8 hours in the *NesCre::R26-DTA* E15.5 brains, primarily due to  
9 a shortening of the G1/G2/M phases (**Fig. S3A-C**). Considering that IPs in *NesCre::R26-DTA*  
10 embryonic brains proliferated two hours faster at E15.5, we predicted that the overrepresented  
11 sub-neuronal clones originated from 8-12% of RGs with a 20-30% growth advantage.

12 These simulations demonstrate that even small clones, when endowed with a  
13 proliferation advantage, can disproportionately grow to dominate the progenitor pool,  
14 highlighting the system's capacity for compensating neurogenesis while maintaining the  
15 overall cell population dynamics.

16

### 17 **C>A and C>T substitutions characterized the *NesCre::R26-DTA* chimera's genomes**

18 To investigate the underlying processes promoting clonal expansion, we used single-  
19 base substitution (SBS) analysis, which examines mutations where one nucleotide is replaced  
20 by another in the DNA sequence. SBS patterns predict distinct mutational signatures,  
21 providing insights into the processes contributing to genome alterations in both normal and  
22 cancer tissues<sup>34</sup>. In mCherry-positive neuronal cells isolated from the *R26-DTA* chimeric  
23 embryos, most of the SNVs are C>T or T>C substitutions, typically associated with aging and  
24 cell cycles<sup>34</sup> (**Figs. 4A, B**). In contrast, C>A substitutions were significantly enriched at the  
25 mCherry-positive neuronal genomes in the *NesCre::R26-DTA* chimeras. C>A substitutions  
26 are strongly linked to oxidative DNA damage<sup>34</sup>. In the reversed complemented chimeras, most  
27 of the SNVs were C>T substitutions, with fewer C>A substitutions (**Figs. S4A, B**).

28 In summary, SBS analysis indicated that most NSPCs underwent more cell cycles (as  
29 evidenced by the C>T substitutions) and, in some cases, engaged biological processes that  
30 generate oxidative stress, which possibly promoted clonal expansion.

31

### 32 **Adaptive clonal expansion slowed neural stem/progenitor cell fate transition**

33 Given that a few subclones formed an entire brain within just five embryonic days, we  
34 wondered whether the neuronal cells in the brains of *NesCre::R26-DTA* chimeras had a

1 biological age equivalent to those in *R26-DTA* chimeras. To address this, we analysed the  
2 frequency of various cell types in the chimeras. In the mCherry-positive cell populations of  
3 *NesCre::R26-DTA* chimeric embryos, we observed a reduced proportion of Cajal-Retzius  
4 cells, ependymal cells, RG cells, glioblasts, and neuroblasts in the brain at E17.5 (**Fig. 5A**).  
5 Specifically, the combination of RG cells, glioblasts, and neuroblasts - representing the  
6 NSPCs at this stage - decreased from 16% to 14.5%. We reasoned that this was because of  
7 the NSPCs actively producing neuronal progenies through asymmetric proliferation and  
8 differentiation, thereby consuming themselves. Indeed, 78% of cells in the *NesCre::R26-DTA*  
9 embryonic brains were neurons, which was six percent higher than in *R26-DTA* embryos. The  
10 other neuronal cell types did not show significant differences between genotypes (**Fig. S5A**).

11 In addition, the transition of cell fate can be inferred from the proportion of each NSPC  
12 type (highlighted in blue in **Fig. S5B**) that remains in the S phase. By E17.5, most RG cells  
13 had exited the cell cycle in both *NesCre::R26-DTA* and *R26-DTA* brains (**Fig. 5B**). However,  
14 we identified a significantly higher proportion of neuroblasts that were still in the S phase within  
15 the *NesCre::R26-DTA* brains, accompanied by a reduction in S phase activity in  
16 oligodendrocyte precursor cells and glioblasts (**Fig. 5B**). Since oligodendrogenesis and  
17 gliogenesis typically occur after neuroblast-dependent neurogenesis<sup>35</sup>, we speculated that  
18 embryonic neurogenesis was delayed.

19 Furthermore, we found the expression levels of genes associated with oxidative  
20 phosphorylation were reduced in the *NesCre::R26-DTA* NSPCs – namely, RG cells, IPs,  
21 glioblasts, neuroblasts, and oligodendrocyte precursor cells (**Fig. 5C, Fig. S5C-E**). As NSPCs  
22 shift from aerobic to oxidative metabolism during cell fate transitions<sup>36</sup>, low oxidative  
23 phosphorylation suggested a delay in the cell fate transition process. Lastly, the expression  
24 levels of genes involved in the G2/M checkpoint pathway in the *NesCre::R26-DTA* NSPCs  
25 were higher than in the *R26-DTA* controls. This indicated that the fast-cycling NSPCs produce  
26 more DNA damage during the S phase, further suggesting that cells experienced replication  
27 stress. Conversely, microglia cells, lymphoid-origin cells migrating to the mouse brain at  
28 E9.5<sup>37</sup>, did not show alterations in oxidative phosphorylation or in the G2/M checkpoint  
29 pathways, suggesting the overall phenotype was limited to the neuronal lineage (**Fig. 5C**).

30 In summary, although NSPCs reduced cell cycle length to meet the demands of brain  
31 development, they experienced a delayed cell fate transition at the end of embryonic  
32 neurogenesis. Adaptive clonal expansion also consumed the pool of NSPCs, leaving  
33 significantly fewer NSPC reserves for life.

34

35 **Limitations of the study**

1           Although *Nestin* mRNA is detectable in the central nervous system as early as E7.5<sup>38</sup>,  
2 the transgenic rat *Nestin* promoter and enhancer used to drive Cre expression in this study  
3 only become active from E12.5 onward<sup>22</sup>. Furthermore, the neuron output prediction model  
4 relies on existing neuronal cell count data, which is unavailable between E10.5 and E12.5. As  
5 a result, the model is constrained to address clonal competition starting from E12.5 and not  
6 earlier. Consequently, our analysis did not incorporate clonal expansion between the germ  
7 layer branching point and the onset of neurogenesis.

8

## 9 **Discussion**

10           In this study, we showed that the reduction in the NSPC population created a  
11 stochastic developmental bottleneck, which allowed NSPCs to unleash their proliferation  
12 potential. This scenario promoted adaptive expansion of a small subset of NSPCs, ultimately  
13 restoring the neuronal cell types despite significant brain depletion (**Fig. 1**). Variant allele  
14 frequency analysis provided evidence of selective clonal expansion, occurring independently  
15 of known driver genetic mutations (**Fig. 2**). We hypothesize that an epigenomic or  
16 metabolomic adaptation enabled NSPCs to compete and expand clonally when neighbouring  
17 cells are lost. Our findings also showed that clonal dominance during neurogenesis depends  
18 on both initial clone size and proliferation advantage, highlighting the neurogenic system's  
19 adaptability to balance clonal expansion with overall progenitor pool stability (**Fig. 3**). This  
20 expansion is associated to biological processes that generate oxidative DNA damage, a by-  
21 product of accelerated proliferation (**Fig. 4**). We further speculate that these clonal expansions  
22 occurred between E10.2 and E12.5, a timeframe for which we lacked accurate measurement  
23 data for fitting. While clonal expansion in the human brain has been reported<sup>12,13</sup>, the  
24 underlying reasons and the potential health consequences remain unclear. Our findings  
25 suggest that this driver-independent expansion may be associated with a developmental crisis  
26 affecting the NSPC pool size and may have impact on brain diseases.

27           While neuron numbers were restored through rapid NSPC expansion (**Figs. 1,3**), this  
28 process came at a cost. We observed delayed cell fate transitions and a significant depletion  
29 of NSPC reserves by the end of embryogenesis, implying consequences for postnatal brain  
30 development and long-term neurogenesis (**Fig. 5**). This observation aligns with earlier studies  
31 in adult mice, which showed that certain neural stem cells, typically dormant, become  
32 activated under stress but fail to re-enter the cell cycle after a second wave of loss, suggesting  
33 an irreversible dormancy state<sup>39</sup>. Our findings suggest that adaptive clonal expansion may  
34 consume these stem cell reservoirs, posing risks for impaired neurogenesis later in life.

1           Contrary to the lateral inhibition model, which proposes that Notch1-Dll1 interactions  
2 promote self-renewal and suppress asymmetric neurogenesis<sup>40</sup>, we found that losing  
3 neighboring cells facilitated RG cells to restore the progenitor pool. This shift likely occurs due  
4 to reduced competition for Notch1 ligands, enabling the remaining RG cells to re-enter active  
5 cycling. Alternative pathways, such as FGF2 signalling, hypoxia-induced cell cycle  
6 acceleration, or soluble Notch1 ligands, may complement this framework. Furthermore,  
7 progenitor cell pool reduction-driven clonal expansion can also be observed in real-world  
8 scenarios, such as hematopoietic cell transplantation. In this context, clonal diversity declines  
9 after allogeneic transplants due to selective pressures or loss of a suppressive niche<sup>41</sup>. Unlike  
10 the foreign environment hematopoietic stem cell encountered in the recipients, adaptive clonal  
11 expansion in the brain reflects a native stem cell response to developmental stress.

12

## 13 **Material and Methods**

### 14 **Mice**

15           All mice were bred in-house at the DKFZ Center for Preclinical Research and  
16 maintained under specific pathogen and opportunistic-free conditions on a 12-hour light-dark  
17 cycle in a temperature-controlled environment with *ad libitum* access to food and water. All  
18 procedures adhered to DKFZ guidelines and were approved by the Regierungspräsidium  
19 Karlsruhe (G154/20 and G241/20). To generate brain-ablated embryos, NesCre male mice  
20 (B6.Cg-Tg(Nes-cre)1Kln/J, Jackson Laboratory) were crossed with R26-DTA female mice  
21 (B6.129P2-Gt(ROSA)26Sortm1(DTA)Lky/J, Jackson Laboratory), producing *NesCre::R26-*  
22 *DTA* and *R26-DTA* embryos. Mating females were checked for vaginal plugs (VP) the  
23 following morning, and pregnant females were euthanized via cervical dislocation 12- or 15-  
24 days post-mating to collect E12.5 and E15.5 embryos. SW nursing females and vasectomized  
25 SW males were bred in-house at DKFZ for blastocyst complementation. DsRed (B6.Cg-  
26 Tg(CAG-DsRed\*MST)1Nagy/J) mice were used for reversed complementation experiments  
27 following the above mentioned procedures.

### 28 **Blastocysts Complementation**

29           Blastocyst injection and implantation followed procedures described in <sup>42</sup>, with minor  
30 modifications. In brief, 6- to 12-week-old R26-DTA females were used as blastocyst donors.  
31 They were superovulated by intraperitoneal injection of 5 I.U. pregnant mare serum  
32 gonadotropin (PMSG), followed by 5 I.U. human chorionic gonadotropin (hCG) 46-48h later.  
33 *R26-DTA* homozygous females and *Nestin* hemizygous male animals were set in the same  
34 cage at the earliest one hour after the last hormone administration. The next morning, the  
35 mated females were checked for VP. 3.5 days post-mating, which was equal to day 2.5 for

1 morulae collection, VP-positive females were killed by cervical dislocation. The oviducts and  
2 part of the uterus were removed, and the blastocysts, or morulae, were isolated by flushing  
3 the oviduct/uterus. On day 3.5, around 15 mCherry-positive mouse ES cells from the cell line  
4 JM8/A4 were injected into the blastocyst cavity. These blastocysts were then implanted  
5 bilaterally into the uteri of a pseudopregnant uteri of a 2.5-day pseudopregnant Swiss Webster  
6 (SW) female mouse. At E15.5 or E17.5, the pregnant SW females were killed by cervical  
7 dislocation, and the chimeric embryos were harvested.

8 For the reversed complementation experiments, 6- to 12-week-old *DsRed::R26-DTA*  
9 homozygous females were used as donors. These females were superovulated and mated  
10 with *Nestin* hemizygous males. Blastocysts were isolated, consisting of either *DsRed::R26-*  
11 *DTA* hemizygous or *DsRed::NesCre::R26-DTA* hemizygous genotypes. *NesCre::R26-DTA*  
12 mouse ES cells were then injected into all blastocysts. All *DsRed:: NesCre::R26-DTA*  
13 chimeras exhibited complete ablation phenotypes, leaving no brains for analysis. Deep WGS  
14 was performed on DsRed-positive cells derived from *DsRed::R26-DTA* chimeras.

15

## 16 **Measurement of Cell Cycle Length Using BrdU/EdU Pulse-Chase *in Vivo***

17 Pregnant female mice carrying embryos at E15.5 were used for the *in vivo* pulse-chase  
18 experiments. At embryonic day E15.5, the pregnant female mice received an intraperitoneal  
19 injection of 5-ethynyl-2'-deoxyuridine (EdU) at a dose of 50 mg/kg body weight followed by 5-  
20 bromo-2'-deoxyuridine (BrdU) injection at 100 mg/kg body weight after a chase period of one  
21 hour. EdU and BrdU were both dissolved in sterile PBS and administered separately to label  
22 S-phase cells at two distinct time points. 30 minutes after BrdU injection, the mice were  
23 euthanized, the embryos were isolated and fixed in 10% (w/v) neutral buffered formalin  
24 solution for at least 48 hrs at RT. The embryos were then FFPE processed and cut as  
25 described in the embryo preparation for histology and Immunofluorescence section below.

26 TBR2-, BrdU-, or EdU-positive cells were determined by immunofluorescence staining.  
27 The number of TBR2+ cells that were either positive or negative for BrdU or EdU, respectively,  
28 as well as triple positive cells, were determined by using the neural network-based cell  
29 segmentation software CellPose<sup>243</sup> in combination with a custom ImageJ Macro workflow for  
30 analysing marker expression and their colocalization. In Cellpose, we trained custom models  
31 and adapted the flow and the cell probability threshold for each marker of interest, based on  
32 Cellpose's integrated nuclei model<sup>44</sup>. To calculate cell cycle length and S phase length, we  
33 adapted the published approach<sup>32</sup>, using the following formulas:

- 34
- S-phase-length =  $\text{BrdU+}/\text{BrdU-EdU+}$
  - cell cycle length =  $\text{S-Phase-length}/(\text{BrdU+}/\text{all})$ .
- 35

## 1 **Cell Sorting**

2 To enrich mCherry-positive neuronal cells and tail fibroblasts, pseudopregnant SW  
3 females were euthanized by cervical dislocation at E17.5. After dissection, embryos were  
4 immediately placed in PBS 1X solution on ice. The forebrain was microdissected, the  
5 meninges removed, and the olfactory bulb was dissected out. For single-cell RNA sequencing  
6 experiments, the forebrain and mid hindbrain from each embryo were micro-dissected and  
7 used as one replicate. For WGS experiments, each forebrain isolated from one embryo was  
8 used as one replicate and tail embryonic fibroblasts from the corresponding embryo were  
9 isolated and used as a control. Brain tissue was digested using the papain dissociation kit  
10 (Worthington Biochemical Corporation, LK003150), and tail embryonic fibroblasts were  
11 obtained by digesting the embryonic tails using a digesting solution consisting of pronase and  
12 collagenase D, following the published protocol<sup>45</sup>.

13

## 14 **Mouse Embryo Preparation for Histology and Immunofluorescence**

### 15 **Sample Collection**

16 Mouse embryos were collected at E12.5 and E15.5 after timed mating.  
17 Pregnant females were euthanized by cervical dislocation following institutional animal  
18 care guidelines. Embryos were dissected in cold phosphate-buffered saline (PBS) and  
19 immediately processed for fixation.

### 20 **Fixation**

21 Embryos were fixed in 10% (w/v) neutral buffered formalin solution for at least  
22 24 hours at room temperature. After fixation, embryos were then processed for  
23 dehydration on a STP 120 processing machine (Thermo Fisher) and embedded in  
24 paraffin using a Histo Star embedding station (Thermo Fisher) for histological  
25 sectioning. Sections (thickness: 6  $\mu$ m) were cut using a RM2255 microtome (Leica),  
26 mounted on Super-Frost Plus glass slides, and dried overnight at 37°C.

### 27 **H&E Staining**

28 Paraffin sections were deparaffinated and rehydrated as follows: 2x 10 min in  
29 xylene, 2x 10 min in 100% ethanol, 2x 5 min in 95% ethanol, 2x 5 min in 70% ethanol  
30 and 1x 5 min in ddH<sub>2</sub>O. The staining was performed with the deparaffinated and  
31 rehydrated sections by the following steps: 30 sec in hematoxylin solution, 20 sec  
32 rinsing with tap water, 10 sec in 0.1% hydrochloric acid, 6 min rinsing with warm tap  
33 water, 30 sec in eosin solution, 1 min rinsing with flowing tap water. After the staining,  
34 dehydration of the slides was performed as follows: 1x 3 min in 70% ethanol, 1x 3 min

1 in 85% ethanol, 2x 5 min in 100% ethanol and 2x 5 min in xylene. Using Eukitt  
2 mounting medium (ORSatec GmbH #202200401), the dehydrated and stained slides  
3 were mounted using glass coverslips. Imaging was performed afterwards by using an  
4 Axioscan 7 slide scanner (Zeiss).

## 5 **Immunofluorescence Staining and imaging acquisition**

6 Sections were deparaffinized and rehydrated as described in the H&E staining  
7 section, then washed three times in 1X PBS solution. Antigen retrieval was performed  
8 by boiling the sections at 80°C for 10 minutes in a solution of citrate buffer (BioGenex  
9 #HK088-9K; 25 mL of 10X citrate buffer diluted with 225 mL of ddH<sub>2</sub>O). The sections  
10 were blocked in 10% normal donkey serum in PBS-T (with 0.2% Triton X-100) for 1  
11 hour, RT. Click-IT reaction was performed using the Click-iT™ EdU Cell Proliferation  
12 Kit for Imaging, Alexa Fluor™ 647 dye (ThermoFisher Scientific cat. NO. C10340)  
13 according to the manufacture instruction to detect EdU. Sections were then incubated  
14 with primary antibody against BrdU (rat, Abcam ab6326) at 4°C overnight. After  
15 washing with PBS, sections were incubated with the fluorescently conjugated  
16 secondary antibody anti-rat Alexa Fluor® 750 (goat, Abcam ab175751) for 1 hour, RT.  
17 The next day, the same sections were incubated with primary antibody against PAX6  
18 (rabbit, Biolegend, cat #901301) or fluorophore-conjugated TBR2 (Alexa488, rat,  
19 eBioscience 53-4875-82) primary antibody, and a fluorophore-conjugated secondary  
20 antibody (Alexa 555, goat-anti-rabbit, ThermoFisher A32732) was used to label  
21 primary antibody against PAX6. After washing, nuclei were counterstained with DAPI.  
22 Slides were mounted with ProLong Gold Antifade Mounting Reagent using glass  
23 coverslips. Immunofluorescent images were obtained using THUNDER Imager 3D  
24 Live Cell (Leica), using an HC PL APO 20x/0.80 lens. All samples were blinded before  
25 the acquisition of the data.

## 26 **Whole-genome sequencing**

### 27 **Sample Preparation**

28 Neuronal cells were isolated from the E17.5 mouse embryonic brain following  
29 the procedure described in the Cell sorting section. For control comparisons, tail  
30 fibroblasts and mouse ES cells were harvested following established protocols.  
31 mCherry-positive cells were sorted by using Aria, Fusion, and Symphony S6 SE, and  
32 genomic DNA was extracted from all samples using the PureLink™ Genomic DNA  
33 Mini Kit (Invitrogen cat. N. K1820-00) according to the manufacturer's instructions.  
34 DNA quality and concentration were assessed using a NanoDrop spectrophotometer  
35 and Qubit dsDNA Assay Kit, respectively.

## 1 **Whole-Genome Sequencing**

2 Whole-genome sequencing libraries were prepared using the Illumina DNA  
3 Prep, (M) Tagmentation kit, and the libraries were mixed to create a multiplex. Pair-  
4 end sequencing was performed using the NovaSeq 6000 S4 V1.5 or the NextSeq 500  
5 Mid-Output sequencer to achieve a depth of approximately 50X for ES cells, 30X for  
6 tail fibroblasts, and 150X for neuronal cells (see **Table S2** for sequencing coverage  
7 details). Sequencing reads were 150bp in length, and a minimum of 52 million reads,  
8 140 and 255 million reads per sample were generated for tail, forebrain and ES cell  
9 line, respectively.

## 10 **Germline and Somatic SNV Calling**

11 BAM files were generated from the raw sequencing data using BWA (version  
12 0.7.15) to align reads to the reference genome GRCm38mm10\_PhiX with the help of  
13 the DKFZ ODCF team. Somatic single nucleotide variants (SNVs) and small  
14 insertions/deletions (indels) in the forebrain samples were identified using Strelka  
15 (v2.9.244) and Mutect2 (GATK v4.2.0.043), with matched tail fibroblast samples and  
16 ES cell sequencing, both serving as germline controls. Variants located in repeat and  
17 simple repeat regions were filtered out using Bedtools Intersect (v2.24.045). High-  
18 confidence variants were defined as those passing the default filters of both Strelka  
19 and Mutect2.

## 20 **Variant Allele Frequency**

21 Variant allele frequencies were calculated for each small variant (SNVs and  
22 indels) using GATK. Only variants with a VAF  $\geq 0.02$  in neuronal cells and absent in  
23 both tail fibroblasts and ES cells were included for further analysis. The neuronal-  
24 specific SNVs were then used in downstream mutation signature and single base  
25 substitution (SBS) analyses.

## 26 **Binomial clustering of VAF**

27 To perform binomial clustering we use VIBER, where we require the minimum  
28 cluster size to be 2% of the total number of points.

29

## 30 **Single Nucleotide Substitution and Mutation Signature**

31 To investigate the mutational processes in neuronal cells, SBS patterns were  
32 categorized into the six classes of base substitutions (C>A, C>G, C>T, T>A, T>C,  
33 T>G). Mutation signature analysis was performed using the COSMIC database.

1 Somatic SNVs were mapped to these signatures, and the de novo extraction of  
2 mutational signatures and relative contribution of each signature was calculated using  
3 SigProfilerExtractor<sup>46</sup> .

4

## 5 **Mathematical Modeling of Neural Progenitor Dynamics**

### 6 **Model Framework**

7 A system of ordinary differential equations (ODEs) was developed to simulate  
8 neural progenitor cell types' proliferation and differentiation dynamics during early  
9 brain development. The model tracks neuroepithelial cells (NECs), radial glial cells  
10 (RGs), intermediate progenitors (IPs), and the neurons (Neu) they produce.

11 This model assumes that NECs self-renew to maintain a constant population  
12 size. They then differentiate into RGs at rate  $\alpha_{\text{NEC}}$ . RGs self-renew at rate  $\lambda_{\text{RG}}$  and  
13 differentiate into IPs at rate  $\alpha_{\text{RG}}$ . The RGs' self-renewal rate,  $\lambda_{\text{RG}}$ , decreases over time,  
14 depending on the total cell count  $n_{\text{total}}$ . This aims at simulating the transition from  
15 proliferation to differentiation during neurogenesis as the cell population in the  
16 neocortex approaches a carrying capacity  $CC$ . Lastly, IPs self-renew at rate  $\lambda_{\text{IP}}$  and  
17 divide to produce two neurons at rate  $\alpha_{\text{IP}}$ . The system of ODEs is given by:

$$18 \quad \begin{cases} \frac{d\text{NEC}}{dt} = 0 \\ \frac{d\text{RG}}{dt} = (\lambda_{\text{RG}} - \alpha_{\text{RG}})\text{RG} \\ \frac{d\text{IP}}{dt} = \alpha_{\text{RG}}\text{RG} - \alpha_{\text{IP}}\text{IP} \\ \frac{d\text{Neu}}{dt} = 2 \alpha_{\text{IP}}\text{IP} \end{cases}$$

19 where

$$20 \quad \lambda_{\text{RG}} = \begin{cases} \lambda_0 \left(1 - \frac{n_{\text{total}}}{CC}\right) & \text{for } n_{\text{total}} < CC \\ 0 & \text{else} \end{cases}$$

21

### 22 **Model Calibration and Validation**

23 The model was calibrated using quantitative flow cytometry data<sup>31</sup>, which  
24 provided absolute counts of progenitor cell types from E12.5 to E18.5. The model  
25 predicts  $\alpha_{\text{NEC}} = \lambda_{\text{IP}} = 0$ , i.e. NECs do not proliferate into RGs after E12.5, and IPs do  
26 not self-renew. These predictions were validated by comparing simulated results to  
27 observed data across embryonic time points. The model accurately predicts progenitor  
28 population dynamics under non-ablation conditions.

## 1           **Ablation and Clonal Dynamic Simulations**

2           The model applied a 30%, 50%, or 90% reduction in NEC, RG, and IP  
3           populations at E12.5 to simulate ablation. To simulate clonal dynamics, we took the  
4           proliferation and differentiation rates from the model and ran a Gillespie simulation. We  
5           assumed an initial clone consisting of between 8%, 10%, and 12% of RG cells at  
6           E12.5, proliferating between 20% and 30% faster than the remaining cells, determined  
7           by a selective advantage. To keep the total cell count constant, the proliferation rate  
8           of the non-clonal cells was adjusted.

9

## 10          **Single-Cell RNA-sequencing**

### 11           **Sample Preparation**

12           The entire embryonic mouse brain was harvested at E17.5. The embryonic  
13           brains were dissected under a stereomicroscope, dividing into the forebrain (including  
14           the thalamus and the neocortex) and the mid-/hindbrain (including medulla oblongata,  
15           cerebellum, and midbrain). The dissected brain parts were dissociated into a single-  
16           cell suspension using the papain dissociation kit (Worthington Biochemical  
17           Corporation, LK003150) according to established protocols. Cells were filtered through  
18           a 70  $\mu$ m cell strainer to remove debris, and viability was assessed using trypan blue  
19           staining to ensure a minimum viability of 90%. Cells were counted, and the  
20           concentration was adjusted to 500,000 cells/mL for downstream applications.

### 21           **Library Preparation**

22           Single-cell RNA sequencing libraries were prepared using the 10x Genomics  
23           Chromium platform (Single Cell 5' v2 kit) following the manufacturer's protocol at the  
24           single-cell Open lab at the DKFZ. Briefly, 20,000 of single cells were loaded into the  
25           Chromium Controller to generate single-cell gel bead-in-emulsions (GEMs). Reverse  
26           transcription was carried out within the GEMs to capture mRNA and generate cDNA.  
27           In the reverse transcription step, a custom primer was added to the reaction at 10nM  
28           concentration to specifically enrich the mCherry transcript. Libraries were constructed  
29           through subsequent amplification, fragmentation, and attachment of Illumina  
30           sequencing adapters.

### 31           **Sequencing**

32           The prepared libraries were sequenced on an Illumina Novaseq 6000 S2 to a  
33           depth between 50,000 and 60,000 reads per cell. Sequencing was performed in  
34           paired-end mode with 28 cycles for Read 1 and 90 cycles for Read 2 to obtain high-

1 quality reads for downstream analyses. The QC plots of each single-cell sequencing  
2 sample can be found at the Github repository  
3 (<https://github.com/orgs/brainbreaks/repositories/NPC-Clone-Plasticity>)

#### 4 **Data pre-processing and clustering**

5 The raw sequencing data were processed using the CellRanger pipeline  
6 (v7.0.0, 10x Genomics), aligning reads to a custom genome reference built using the  
7 mkref function in CellRanger. This reference was based on the mouse genome and  
8 transcriptome (mm10 from 10X Genomics), with additional sequences for mCherry,  
9 DTA, and Cre (included in the GTF and FASTA files). Default parameters were used  
10 to generate gene-by-cell count matrices.

11 Raw count matrices were individually processed to remove doublet cells using  
12 Scrublet<sup>47</sup>. The Seurat package<sup>48</sup> (v4.3.0) was employed to generate sparse count  
13 matrices for downstream analyses. Cells were filtered to retain those with higher  
14 quality, defined by the detection of at least 500 genes and a mitochondrial gene  
15 expression percentage of  $\leq 5\%$ . Data normalization followed the LogNormalize function  
16 in Seurat, using a scaling factor of 10,000. Additionally, 3000 highly variable genes  
17 were identified, and the effects of number of detected features (nFeature) RNA and  
18 mitochondrial gene expression percentage were regressed out for subsequent  
19 analyses.

20 Principal component analysis (PCA) was performed on the top 3000 highly  
21 variable genes, retaining 50 principal components for downstream analyses. To  
22 integrate single-cell data across different batches, Harmony<sup>49</sup> (v1.0.1) was applied.  
23 The integrated data were subsequently used to construct a k-nearest neighbors' graph,  
24 followed by clustering with the Leiden algorithm, using a final resolution of four.

#### 26 **Cell Type Annotation**

27 To annotate cell types in the single-cell data, we performed high-resolution  
28 clustering within each cluster to generate meta-cells for annotation. We evaluated 23  
29 different resolutions, ranging from 1 to 10 in increments of 0.4, and selected the optimal  
30 resolution based on the size and nFeatures of the meta-cells. Canonical correlation  
31 analysis (CCA)-based label transfer, as implemented in Seurat, was used to transfer  
32 annotations from the reference atlas<sup>50</sup> utilizing the 3,000 highly variable genes  
33 identified. The final cluster annotations were determined by evaluating the highest  
34 predicted scores, consensus agreement of transferred labels, and the expression  
35 patterns of canonical marker genes.

## 1 **Cell Type Proportion Statistical Analyses**

2 We calculated the cell type composition across biological replicates for both  
3 the *NesCre::R26-DTA* and *R26-DTA* conditions. To account for the replicated  
4 information, we applied a Generalized Linear Model with a binomial distribution  
5 followed by an ANOVA using the chi-square test to assess statistically significant  
6 differences in cell type proportions between the conditions.

## 7 **S Phase Score Analyses**

8 We assigned cell cycle scores to each cell using the Seurat CellCycleScoring  
9 function and assessed the distribution differences between the *NesCre::R26-DTA* and  
10 *R26-DTA* conditions within each cell type using the Kolmogorov-Smirnov test.

## 11 **Regulon Analyses**

12 We used SCENIC<sup>51,52</sup> to infer gene regulatory networks, utilizing the latest  
13 motif database (mm10\_10kbp\_up\_10kbp\_down from SCENIC) and transcription  
14 factors from the cisTarget database for *Mus musculus*. To identify conserved regulons,  
15 we implemented the SCENIC algorithm using pySCENIC (v0.12.1) in Python (v3.8.19),  
16 running 78 iterations. In each iteration, we performed random subsampling with  
17 replacement to obtain 20,000 cells, applying GRNBoost2 with default settings. For  
18 consensus regulon identification, we retained regulons with a frequency of occurrence  
19 greater than 50% across iterations and used the average weight to perform AUCell  
20 scoring for all cells.

## 21 **Pathway Enrichment Analysis**

22 We performed Gene Set Enrichment Analysis (GSEA) using fgsea<sup>53</sup> (v1.30.0  
23 in R 4.3.0) and the Hallmark gene sets for *Mus musculus* from Msigdb (v7.5.1)  
24 (<https://github.com/igordot/msigdb>, also <sup>54</sup>) to compare the *NesCre::R26-DTA* and  
25 *R26-DTA* conditions within each cell type. The input for GSEA was derived from the  
26 results of a Wilcoxon rank-sum test (wilcoxauc) performed using Presto (v1.0.0.).

## 27 **Regulon Activity and Associated Pathways Analysis**

28 The activity of regulons showing significant differences ( $p$ -values  $\leq 0.01$ )  
29 between the *NesCre::R26-DTA* and *R26-DTA* conditions, as identified through a  
30 permutation test, was selected for further analysis. To explore pathways associated  
31 with each regulon, we used the same Hallmark gene sets as the GSEA analysis.  
32 Fisher's exact test was performed on the Odds Ratio, considering the overlap between  
33 regulon and pathway gene sets. As a background, we included gene sets with more  
34 than 5% expression in the single-cell population. Pathways and regulons with an Odds

1 Ratio  $\geq 1$  and p-values  $\leq 0.01$  were considered significantly associated. Regulons with  
2 gene set overlaps below the 10th percentile were excluded. To validate that these  
3 overlapping gene sets contributed to pathway differences, we applied the Kolmogorov-  
4 Smirnov test on module scores (calculated via Seurat's AddModuleScore function),  
5 retaining only those with significant differences (p-values  $\leq 0.01$ ) for the final results.

## 6 **Single-Cell Variant Calling**

7 We utilized CellSNP-lite to perform SNP calling at single-cell resolution using  
8 E17.5 10x Cell Ranger outputs. This analysis was restricted to paired animals with  
9 whole-genome sequencing (WGS) data (mouse12 and mouse17). SNP calling was  
10 conducted in Mode 1a with default settings, except for the --minMAF 0 and --  
11 minCOUNT 1. Additionally, we considered only SNPs that were also identified in the  
12 WGS data of paired animals and known to contribute to clonal expansion.

13

## 14 **Code availability**

15 R-based scripts for analyzing the scRNA-seq data were deposited under the  
16 brainbreaks GitHub repository (<https://github.com/orgs/brainbreaks/repositories/NPC-Clone-Plasticity>). WGS datasets have been uploaded as Bioproject PRJNA1209637 (BioProject and  
17 associated SRA metadata are available  
18 at [https://dataview.ncbi.nlm.nih.gov/object/PRJNA1209637?reviewer=ar8r2fidffnio7pl9rv8ne  
19 6m2m](https://dataview.ncbi.nlm.nih.gov/object/PRJNA1209637?reviewer=ar8r2fidffnio7pl9rv8ne6m2m) ) and the R script for analyzing it will be published on Github upon publication. For  
20 reviewers: <https://tsbnc.dkfz.de/index.php/s/W7RLZYr4WAZWG5w> Password: 7S4UrleRVm.

21

## 22 **Statistical Power**

23 Statistical power for all the experiments was mentioned in the associated text, figure legends,  
24 or methods.

25

## 26 **Acknowledgments**

27 This work is supported by the Helmholtz Young Investigator grant and an ERC starting grant  
28 BrainBreaks to P-C W.. G.D.M. received a Conference Travel Program from DAAD for  
29 presenting this work at FENS 2024. We thank Björn Schwer and Cameron M. Holman for  
30 sharing our mCherry-positive ES cells. We thank Yassin Harim for computational and  
31 technical support to improve ImageJ batch workflow performance, Lorenzo Corazzi for  
32 mapping the mCherry integration position and sample preparation, Boyu Ding and the FACS  
33

1 core facility for cell sorting, the light microscopy core facility, the Transgenic Core facility, the  
2 Single-cell open lab, the NGS core facility, the ODCF data management core facility, and the  
3 staff at the preclinic center at the DKFZ. We also thank the imaging facility at EMBL for  
4 establishing the BrdU/EdU imaging workflow. We extend our acknowledgments to P. Ginno  
5 and T. Martin for their valuable feedback. We also thank the team members at the Wei and  
6 Höfer labs for intellectual discussions and technical support.

7

## 8 **Author contributions**

9 Conceptualization: G.D.M., S.B., T.H., P.-C. W.; Study design: P.-C. W., G.D.M.;  
10 Supervision: P.-C. W., T.H.; Managing, planning, and generate chimeras: F. v.d. H., B.A.;  
11 Conducting histology, immunofluorescent experiments and imaging: G.D.M., H.-J. L., J.B.;  
12 Mathematical analysis and modeling: L.W., M.K.; Sample and library preparation: G.D.M.,  
13 N.C.; Whole-genome sequencing: G.D.M., N.C.; WGS data analysis: S.B., Single-Cell RNA-  
14 sequencing analysis: L.-C. W., Visualization: G.D.M., S.B., L. W., L.-C. W., P.-C. W. and  
15 T.H.; Writing (original draft): G.D.M, S.B., L.-C. W., L.W., P.-C.W. ; Writing (review and  
16 editing): V.K., T.H..

17

## 18 **References**

- 19 1. Chang, A. N. *et al.* Neural blastocyst complementation enables mouse forebrain  
20 organogenesis. *Nature* **563**, 126–130 (2018).
- 21 2. Huang, J. *et al.* Generation of rat forebrain tissues in mice. *Cell* **187**, 2129-2142.e17  
22 (2024).
- 23 3. Stahl, R. *et al.* Trnp1 Regulates Expansion and Folding of the Mammalian Cerebral  
24 Cortex by Control of Radial Glial Fate. *Cell* **153**, 535–549 (2013).
- 25 4. Fernández, V., Llinares-Benadero, C. & Borrell, V. Cerebral cortex expansion and folding:  
26 what have we learned? *EMBO J.* **35**, 1021–1044 (2016).
- 27 5. Rakic, P. Specification of Cerebral Cortical Areas. *Science* **241**, 170–176 (1988).
- 28 6. Rakic, P. The radial edifice of cortical architecture: From neuronal silhouettes to genetic  
29 engineering. *Brain Res. Rev.* **55**, 204–219 (2007).
- 30 7. Borrell, V. How Cells Fold the Cerebral Cortex. *J. Neurosci.* **38**, 776–783 (2018).
- 31 8. D’Gama, A. M. & Walsh, C. A. Somatic mosaicism and neurodevelopmental disease. *Nat.*  
32 *Neurosci.* **21**, 1504–1514 (2018).

- 1 9. S., J. S. *et al.* Somatic Mutations in Cerebral Cortical Malformations. *N. Engl. J. Med.* **371**,  
2 733–743 (2014).
- 3 10. D’Gama, A. M. *et al.* Somatic Mutations Activating the mTOR Pathway in Dorsal  
4 Telencephalic Progenitors Cause a Continuum of Cortical Dysplasias. *Cell Rep.* **21**, 3754–  
5 3766 (2017).
- 6 11. Breuss, M. W. *et al.* Somatic mosaicism reveals clonal distributions of neocortical  
7 development. *Nature* **604**, 689–696 (2022).
- 8 12. Körber, V. *et al.* Detecting and quantifying clonal selection in somatic stem cells. *bioRxiv*  
9 2021.12.15.472780 (2023) doi:10.1101/2021.12.15.472780.
- 10 13. Bae, T. *et al.* Analysis of somatic mutations in 131 human brains reveals aging-  
11 associated hypermutability. *Science* **377**, 511–517 (2022).
- 12 14. Gao, P. *et al.* Deterministic Progenitor Behavior and Unitary Production of Neurons in  
13 the Neocortex. *Cell* **159**, 775–788 (2014).
- 14 15. Noctor, S. C., Martínez-Cerdeño, V., Ivic, L. & Kriegstein, A. R. Cortical neurons arise in  
15 symmetric and asymmetric division zones and migrate through specific phases. *Nat.*  
16 *Neurosci.* **7**, 136–144 (2004).
- 17 16. Gaiano, N., Nye, J. S. & Fishell, G. Radial Glial Identity Is Promoted by Notch1 Signaling  
18 in the Murine Forebrain. *Neuron* **26**, 395–404 (2000).
- 19 17. Nonomura, K. *et al.* Local Apoptosis Modulates Early Mammalian Brain Development  
20 through the Elimination of Morphogen-Producing Cells. *Dev. Cell* **27**, 621–634 (2013).
- 21 18. Kawaue, T. *et al.* Lzts1 controls both neuronal delamination and outer radial glial-like cell  
22 generation during mammalian cerebral development. *Nat. Commun.* **10**, 2780 (2019).
- 23 19. Laporte, M. H. *et al.* Alix is required during development for normal growth of the mouse  
24 brain. *Sci. Rep.* **7**, 44767 (2017).
- 25 20. Ameisen, J. C. On the origin, evolution, and nature of programmed cell death: a timeline  
26 of four billion years. *Cell Death Differ.* **9**, 367–393 (2002).
- 27 21. Yamaguchi, Y. & Miura, M. Programmed Cell Death in Neurodevelopment. *Dev. Cell* **32**,  
28 478–490 (2015).
- 29 22. Liang, H., Hippenmeyer, S. & Ghashghaei, H. T. A Nestin-cre transgenic mouse is  
30 insufficient for recombination in early embryonic neural progenitors. *Biol. Open* **1**, 1200–  
31 1203 (2012).
- 32 23. Tronche, F. *et al.* Disruption of the glucocorticoid receptor gene in the nervous system  
33 results in reduced anxiety. *Nat. Genet.* **23**, 99–103 (1999).
- 34 24. Shafiee, F., Aucoin, M. G. & Jahanian-Najafabadi, A. Targeted Diphtheria Toxin-Based  
35 Therapy: A Review Article. *Front. Microbiol.* **10**, 2340 (2019).

- 1 25. Chen, J., Lansford, R., Stewart, V., Young, F. & Alt, F. W. RAG-2-deficient blastocyst  
2 complementation: an assay of gene function in lymphocyte development. *Proc. Natl. Acad.*  
3 *Sci.* **90**, 4528–4532 (1993).
- 4 26. Rivera-Pérez, J. A. & Hadjantonakis, A.-K. The Dynamics of Morphogenesis in the Early  
5 Mouse Embryo. *Cold Spring Harb. Perspect. Biol.* **7**, a015867 (2015).
- 6 27. Junyent, S. *et al.* The first two blastomeres contribute unequally to the human embryo.  
7 *Cell* **187**, 2838–2854.e17 (2024).
- 8 28. Polejaeva, I. & Mitalipov, S. Stem cell potency and the ability to contribute to chimeric  
9 organisms. *Reproduction* **145**, R81–R88 (2013).
- 10 29. Williams, M. J. *et al.* Quantification of subclonal selection in cancer from bulk sequencing  
11 data. *Nat. Genet.* **50**, 895–903 (2018).
- 12 30. Caravagna, G. *et al.* Subclonal reconstruction of tumors by using machine learning and  
13 population genetics. *Nat. Genet.* **52**, 898–907 (2020).
- 14 31. Jungas, T., Joseph, M., Fawal, M.-A. & Davy, A. Population Dynamics and Neuronal  
15 Polyploidy in the Developing Neocortex. *Cereb. Cortex Commun.* **1**, tgaa063 (2020).
- 16 32. Martynoga, B., Morrison, H., Price, D. J. & Mason, J. O. Foxg1 is required for  
17 specification of ventral telencephalon and region-specific regulation of dorsal telencephalic  
18 precursor proliferation and apoptosis. *Dev. Biol.* **283**, 113–127 (2005).
- 19 33. Harris, L., Zalucki, O. & Piper, M. BrdU/EdU dual labeling to determine the cell-cycle  
20 dynamics of defined cellular subpopulations. *J. Mol. Histol.* **49**, 229–234 (2018).
- 21 34. Alexandrov, L. B. *et al.* The repertoire of mutational signatures in human cancer. *Nature*  
22 **578**, 94–101 (2020).
- 23 35. Silbereis, J. C., Pochareddy, S., Zhu, Y., Li, M. & Sestan, N. The Cellular and Molecular  
24 Landscapes of the Developing Human Central Nervous System. *Neuron* **89**, 248–268  
25 (2016).
- 26 36. Zheng, X. *et al.* Metabolic reprogramming during neuronal differentiation from aerobic  
27 glycolysis to neuronal oxidative phosphorylation. *eLife* **5**, e13374 (2016).
- 28 37. Ginhoux, F. & Prinz, M. Origin of Microglia: Current Concepts and Past Controversies.  
29 *Cold Spring Harb. Perspect. Biol.* **7**, a020537 (2015).
- 30 38. Lendahl, U., Zimmerman, L. B. & McKay, R. D. G. CNS stem cells express a new class  
31 of intermediate filament protein. *Cell* **60**, 585–595 (1990).
- 32 39. Morshead, C. M. *et al.* Neural stem cells in the adult mammalian forebrain: A relatively  
33 quiescent subpopulation of subependymal cells. *Neuron* **13**, 1071–1082 (1994).
- 34 40. Kawaguchi, A. *et al.* Single-cell gene profiling defines differential progenitor subclasses  
35 in mammalian neurogenesis. *Development* **135**, 3113–3124 (2008).
- 36 41. Chapman, M. S. *et al.* Clonal dynamics after allogeneic haematopoietic cell  
37 transplantation. *Nature* **635**, 926–934 (2024).

- 1 42. Dai, H.-Q. *et al.* Direct analysis of brain phenotypes via neural blastocyst  
2 complementation. *Nat. Protoc.* **15**, 3154–3181 (2020).
- 3 43. Pachitariu, M. & Stringer, C. Cellpose 2.0: how to train your own model. *Nat. Methods*  
4 **19**, 1634–1641 (2022).
- 5 44. Stringer, C., Wang, T., Michaelos, M. & Pachitariu, M. Cellpose: a generalist algorithm  
6 for cellular segmentation. *Nat. Methods* **18**, 100–106 (2021).
- 7 45. Khan, M. & Gasser, S. Generating Primary Fibroblast Cultures from Mouse Ear and Tail  
8 Tissues. *J. Vis. Exp. : JoVE* (2016) doi:10.3791/53565.
- 9 46. Islam, S. M. A. *et al.* Uncovering novel mutational signatures by de novo extraction with  
10 SigProfilerExtractor. *Cell Genom.* **2**, 100179 (2022).
- 11 47. Wolock, S. L., Lopez, R. & Klein, A. M. Scrublet: Computational Identification of Cell  
12 Doublets in Single-Cell Transcriptomic Data. *Cell Syst.* **8**, 281-291.e9 (2019).
- 13 48. Hao, Y. *et al.* Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573-3587.e29  
14 (2021).
- 15 49. Korsunsky, I. *et al.* Fast, sensitive and accurate integration of single-cell data with  
16 Harmony. *Nat. Methods* **16**, 1289–1296 (2019).
- 17 50. Manno, G. L. *et al.* Molecular architecture of the developing mouse brain. *Nature* **596**,  
18 92–96 (2021).
- 19 51. Sande, B. V. de *et al.* A scalable SCENIC workflow for single-cell gene regulatory  
20 network analysis. *Nat. Protoc.* **15**, 2247–2276 (2020).
- 21 52. Aibar, S. *et al.* SCENIC: Single-cell regulatory network inference and clustering. *Nat.*  
22 *methods* **14**, 1083–1086 (2017).
- 23 53. Korotkevich, G. *et al.* Fast gene set enrichment analysis. *bioRxiv* 060012 (2021)  
24 doi:10.1101/060012.
- 25 54. Liberzon, A. *et al.* The Molecular Signatures Database Hallmark Gene Set Collection.  
26 *Cell Syst.* **1**, 417–425 (2015).
- 27

## 1 **Figure Legends, Main Text**

### 2 **Figure 1. Creating the adaptive NSPC expansion model, the *NesCre::R26-DTA*** 3 **chimeras.**

4 **(A)** Schematic representation of the transgenic expression cassettes in the genome of  
5 *NesCre::R26-DTA* embryos. Chr: chromosome. Nestin: rat transgenic Nestin promoter. Cre:  
6 Cre recombinase. ROSA26: ROSA26 locus on Chr6. DTA: diphtheria toxin subunit A. NSPCs:  
7 neural stem and progenitor cells. **(B)** A representation of hematoxylin and eosin (H&E) -  
8 stained sagittal brain sections of an *R26-DTA* embryo (left) and a *NesCre::R26-DTA* embryo  
9 (right) at embryonic day (E) 15.5. Insets showed higher magnification views of the rectangular  
10 areas of neocortex (ncx) or the choroid plexus (cp). Scale bars: 1 mm (overview), 50  $\mu$ m (ncx  
11 view), and 20  $\mu$ m (cp view). Brain structures, including midbrain (mb), neocortex (ncx),  
12 thalamus (th), medulla oblongata (mo), cerebellum (cb), and the spinal cord (sp) as well as  
13 craniofacial structures such as tongue (t), and nose cavity (n) were annotated. **(C)** Schematic  
14 of blastocyst complementation strategy. mCherry-positive embryonic stem cells (pink) were  
15 injected into host blastocysts (*R26-DTA* or *NesCre::R26-DTA*) to generate chimeras. The  
16 NSPCs in the *NesCre::R26-DTA* chimeras (dashed lined cell) died upon DTA expression,  
17 whereas the *R26-DTA* NSPCs (grey cell) were alive. **(D)** H&E-stained sagittal brain sections  
18 of an *R26-DTA* chimera (left) and a *NesCre::R26-DTA* chimera (right) at E15.5. **(E)** Barchart  
19 showed the proportion of mCherry+ cells in forebrain tissue of chimeras at E17.5. Data were  
20 presented as mean  $\pm$  standard deviation, and the statistical significance was determined by  
21 unpaired two-tail Student t test. \*\*\*  $p < 0.005$ .

22

### 23 **Figure 2. Adaptive NSPC clonal expansion in the *NesCre::R26-DTA* chimeras. (A)**

24 Schematic of the experimental workflow. At E17.5, tail fibroblasts and forebrain cells were  
25 isolated from *R26-DTA* or *NesCre::R26-DTA* chimeras. mCherry+ cells were sorted by FACS  
26 and analyzed by whole-genome sequencing (WGS) at a mean of 76x (tail fibroblasts) or 198x  
27 (forebrain cells) coverage. Shared allelic variants identified in tail fibroblasts and in ES cell  
28 genomes were excluded from analysing forebrain-specific variant frequencies. **(B)** Distribution  
29 of variant allele frequencies (VAFs) in mCherry+ forebrain cells from *R26-DTA* (left, n=3) and  
30 *NesCre::R26-DTA* (right, n=6) chimeras. Each panel represented an individual mouse (e.g.,  
31 m09, m12). Exonic and intronic mutations in genes and their frequencies were annotated for  
32 select samples. **(C)** The Uniform Manifold Approximation and Projection (UMAP) plots display  
33 the distribution of single cells carrying specific clonal SNVs. For each SNV, the VAF  
34 determined by WGS, the genomic position, and the substitution type were annotated. Single-  
35 cell types carrying intronic SNVs in *Nfia* (left), *Nav2* (middle), and *Ddx1* (right) were indicated.

1 The region of origin of these cells (fb: forebrain; m/hb: mid-/hindbrain), was also noted. The  
2 size of the icons represented the expression levels of the indicated genes.

3

4 **Figure 3. Modelling adaptive clonal expansion. (A)** Schematic representation of the  
5 framework used for modelling NSPC proliferative dynamics and neuron production. NEC:  
6 neuroepithelial cells, RG: radial glial cells, IP: intermediate progenitors, Neu: Neurons.  $\alpha$ :  
7 differentiation rate.  $\lambda$ : proliferation rate. **(B)** Simulation of progenitor and neuron dynamics  
8 under varying levels of RG and IP ablation (30%, 50%, and 90%). Colored dots with the  
9 standard deviation represented observed cell counts by Jungas et al.<sup>31</sup>, while lines  
10 represented modelled cell counts under non-ablation (solid) and RG-ablation conditions  
11 (dashed). **(C)** Simulation of clonal expansion dynamics with advanced (adv) clones starting at  
12 8% (left) or 12% (right) of the RG. Dashed lines represented cell counts under 50% ablation  
13 adjusted with clonal proliferation advantage, as shown in the middle plot in panel B. Light  
14 green solid lines represented cell counts derived from clones with proliferation advantages  
15 (20%, 25%, or 30%). The percentage of neurons derived from the advanced NSPCs (light  
16 green) and the rest of the NSPCs (dark green) at E18.5 was indicated for each simulation.

17

18 **Figure 4. Mutational signatures in *R26-DTA* and *NesCre::R26-DTA* chimeras revealed**  
19 **distinct patterns of single base substitutions. (A)** Variant allele frequency (VAF)  
20 distributions for each substitution type (represented by different colors and stacked) in  
21 mCherry+ forebrain cells of *R26-DTA* chimeras (top: m09, m12, m16) and *NesCre::R26-DTA*  
22 chimeras (bottom: m02, m07, m08, m10, m17, m18). The *NesCre::R26-DTA* chimeras  
23 displayed broader VAF distributions. **(B)** Relative contribution of each substitution type across  
24 individual *R26-DTA* (top) and *NesCre::R26-DTA* (bottom) chimeras. *NesCre::R26-DTA*  
25 samples showed an increased proportion of C>T and T>A mutations compared to controls.

26

27 **Figure 5. Adaptive clonal expansion delayed NSPC cell fate transition and altered cell**  
28 **composition in *NesCre::R26-DTA* chimeras. (A)** Proportional analysis of brain cell types in  
29 E17.5 *NesCre::R26-DTA* and *R26-DTA* chimeras. Four biological repeats for each condition  
30 were analysed. Statistical significance was determined by a chi-square test. \*\*\*  $p \leq 0.001$ , \*\*\*\*  
31  $p \leq 0.0001$ . **(B)** S phase score in NSPCs. All mCherry+ single cells were considered for this  
32 analysis. The x-axis depicts the activity of S phase genes, determined by CellCycleScore of  
33 the S phase, and the y-axis indicates the proportion of single cells in each cell type. Non-  
34 cycling cells showed a peak centered at 0 on the x-axis (e.g. Cajal Retzius cells). Statistical  
35 significance was determined by a Kolmogorov–Smirnov test. \*\*\*\*  $p \leq 0.0001$ . n.s.: not

1 significant. **(C)** Gene set enrichment analysis was conducted to compare *NesCre::R26-DTA*  
2 and *R26-DTA* chimeras for each cell type based on the scRNA-seq dataset. The direction of  
3 the triangle indicates whether a pathway is up- ( $\blacktriangle$ ) or downregulated ( $\blacktriangledown$ ) in the *NesCre::R26-*  
4 *DTA* compared to the *R26-DTA* chimeras. The size of the triangle depicted the normalized  
5 enriched score (NES). The color indicated the adjusted p-value, shown on a scale of  $-\log_{10}$ .  
6 RG: radial glial cells, GB: glioblasts. NB: neuroblasts. OPC: oligodendrocyte precursor cells.  
7 IP: intermediate progenitors. MG: microglia cells.

## 8 9 **Supplementary Figure Legends**

10

11 **Figure S1. Characterization of tissue and cell type composition in the *NesCre::R26-DTA***  
12 **chimeras. (A)** Schematic representation of the genetic strategy. Males carrying the rat *Nestin*  
13 transgenic promoter at Chr12, which drives Cre recombinase expression, were crossed with  
14 females homozygous for the *R26-DTA* transgene at Chr6, containing a Cre-dependent  
15 diphtheria toxin A (DTA) expression cassette. **(B)** H&E-stained sagittal brain sections of an  
16 *R26-DTA* and a *NesCre::R26-DTA* embryo at E12.5. Insets provided higher magnification  
17 views of rectangular areas of neocortex (ncx) or and the choroid plexus (cp). Brain regions  
18 including midbrain (mb), hindbrain (hb), diencephalon (di), neocortex (ncx), and the spinal  
19 cord (sp) were annotated. Craniofacial structures including tongue (t), and nasal cavity (n)  
20 were also indicated. Scale bars: 500  $\mu\text{m}$  (overview), 50  $\mu\text{m}$  (ncx view), and 20  $\mu\text{m}$  (cp view).  
21 **(C)** Insets from neocortex of sagittal brain section from *R26-DTA* and *NesCre::R26-DTA*  
22 embryos at E15.5, stained against PAX6 (a marker for RG cells) and TBR2 (a marker for IP  
23 cells). Cell nuclei were counterstained with DAPI. Scale bars: 50  $\mu\text{m}$ . **(D)** Bar charts showing  
24 the absolute cell counts of DAPI+, PAX6+, and TBR2+ cells in the whole brains from three  
25 *R26-DTA* and five *NesCre::R26-DTA* embryos at E15.5. Seven sagittal planes were quantified  
26 per embryo. Statistical significance was determined using an unpaired two-tailed Student's t-  
27 test. Data are presented as mean  $\pm$  SD of each sagittal plan; \* $p < 0.05$ . **(E)** A barchart showed  
28 the proportion of mCherry+ cells in the tail fibroblasts from *NesCre::R26-DTA* or *R26-DTA*  
29 chimeras at E17.5. Data were presented as mean  $\pm$  standard deviation, and the statistical  
30 significance was determined by an unpaired two-tailed Student t test. n.s.: not significant.  
31 **(F,G)** Uniform Manifold Approximation and Projection (UMAP) plots showing the first and  
32 second reduced dimensions for single-cell RNA-seq datasets. Cells were visualized by brain  
33 regions **(F)**, forebrain in orange, mid-/hindbrain in blue) or by cell types **(G)** in *R26-DTA* and  
34 *NesCre::R26-DTA* chimeras, respectively.

35

1 **Figure S2. Identifying expanded subclones in the forebrain of the *NesCre::R26-DTA***  
2 **chimeras. (A)** Subclone enrichment analysed by VIBER. VAF shown in Figure 2 are color-  
3 coded for the small subclones (C1, orange) and the selected/expanded subclones (C2,  
4 crimson red). Each panel represented an individual embryo. **(B)** Corrected expression level  
5 for *Nfia* (top), *Nav2* (middle), and *Ddx1* (bottom) in the cells highlighted in Figure 2C. **(C)**  
6 Schematic of the reversed blastocyst complementation strategy. *NesCre::R26-DTA* ES cells  
7 (blue) were injected into host blastocysts (DsRed) to generate *NesCre::R26-DTA* chimeras.  
8 DsRed-positive tail fibroblasts and forebrain cells were sorted by FACS at E17.5 and analysed  
9 by whole-genome sequencing, as described in Figure 2A. **(D)** VAF distribution in DsRed+  
10 forebrain cells from three *NesCre::R26-DTA* chimeras. Each panel represented an individual  
11 mouse.

12

13 **Figure S3. Intermediate neuronal progenitors (TBR2+) display shortened cell cycles in**  
14 **the *NesCre::R26-DTA* chimera. (A)** EdU-BrdU pulse-chase experimental timeline. EdU was  
15 injected intraperitoneally (i.p.) at time t0, followed by BrdU at t1 (1 hour later), and embryos  
16 were collected at t2 (1.5 hours after the initial injection). **(B)** Representative  
17 immunofluorescence images of E15.5 cortical sections stained against TBR2+ (magenta) EdU  
18 (green), BrdU (red), and nuclei labelled with DAPI (grey). Images compared the isocortex  
19 region of *R26-DTA* (top) and *NesCre::R26-DTA* (bottom) chimeras. Scale bar: 50  $\mu$ m. **(C)**  
20 Quantification of S-phase duration and overall cell cycle length for TBR2+ progenitors at E15.5  
21 in *R26-DTA* and *NesCre::R26-DTA* chimeras (n = 3 embryos per group). Data were presented  
22 as mean  $\pm$  SD; \*p < 0.05, determined by a one-tailed unpaired Student t-test.

23

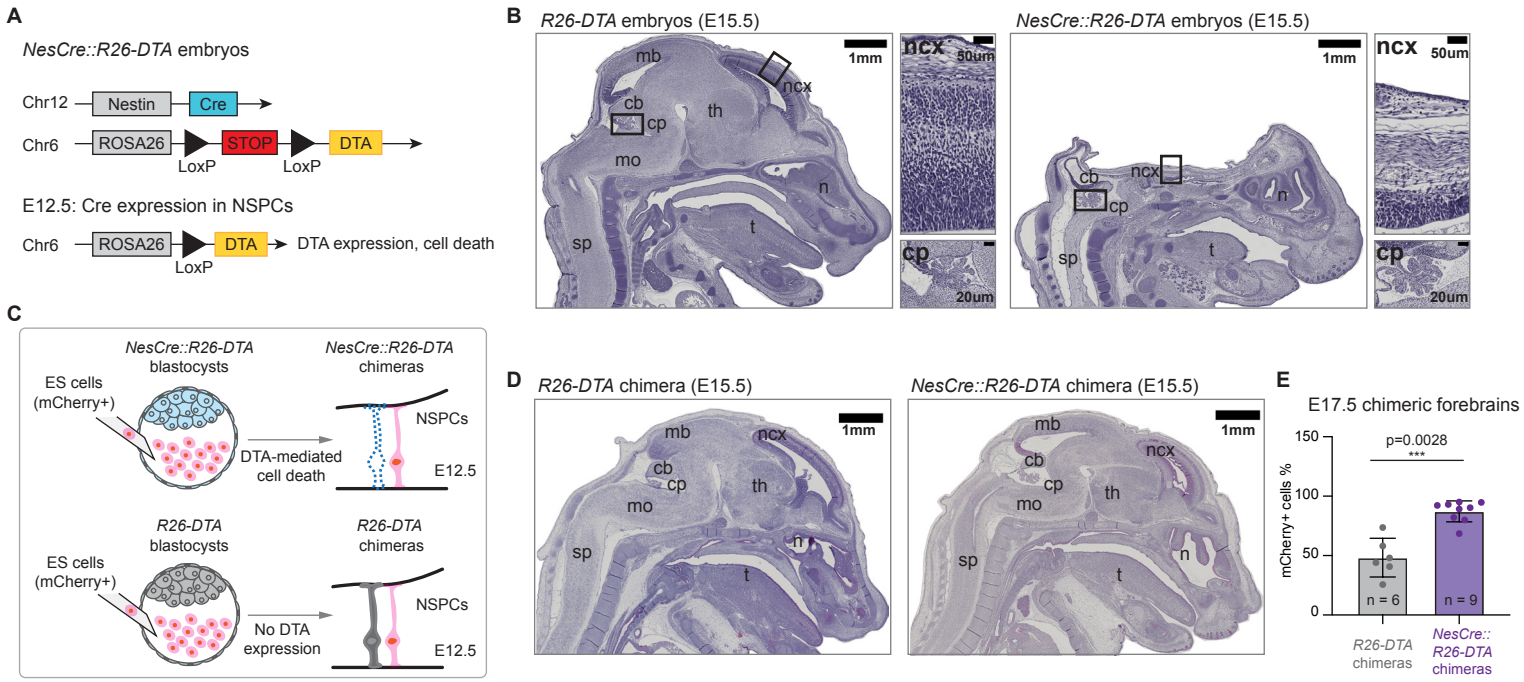
24 **Figure S4. Mutational signatures in *DsRed::R26-DTA* chimeras. (A)** Variant allele  
25 frequency (VAF) distributions for each substitution type (represented by different colors and  
26 stacked) in *DsRed::R26-DTA* chimeras. **(B)** Relative contribution of each substitution type  
27 across individual *DsRed::R26-DTA* chimeras.

28

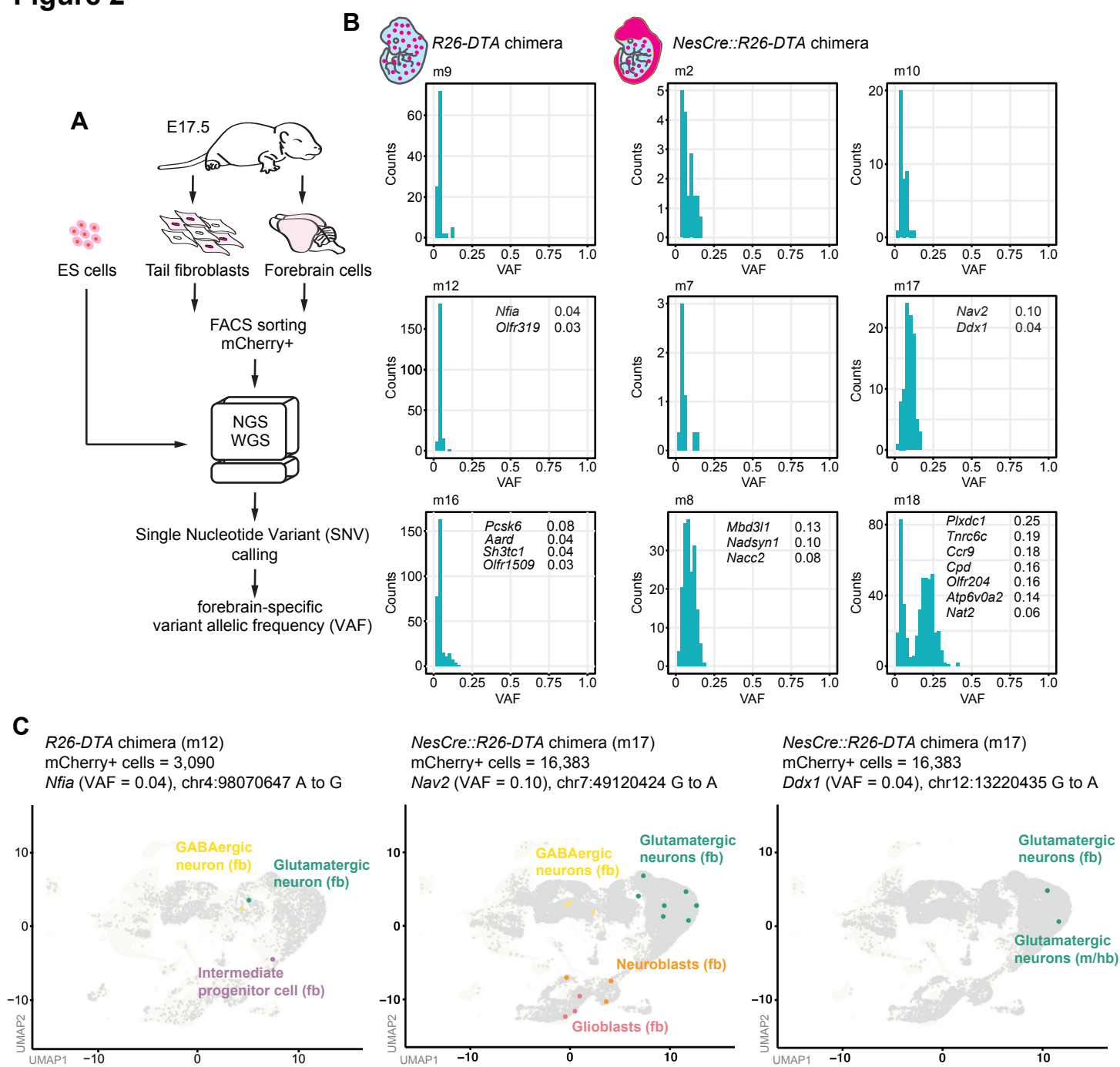
29 **Figure S5. Adaptive clonal expansion altered cell type-specific pathway activity in**  
30 **chimera brains. (A)** Box plots showing the proportions of various cell types, including  
31 glycinergic neurons (GLYC), intermediate progenitor (IP) cells, microglia, oligodendrocyte  
32 precursor cells (OPC), pericytes, and endothelial cells, in *R26-DTA* (gray) and *NesCre::R26-*  
33 *DTA* (purple) chimeras. No significant differences (n.s.) were observed between groups for  
34 most cell types, except for the endothelial cells. Statistical power was determined by general  
35 linear models using Chi-square. \*\*\*\*:  $p \leq 0.0001$ . **(B)** Developmental trajectory of

1 neuroepithelial-derived cell types. Neuroepithelial cells transition to radial glial cells, which  
2 give rise to intermediate progenitors, glioblasts, oligodendrocyte precursors and ependymal  
3 cells. Intermediate progenitors further differentiate into neuroblasts and eventually generate  
4 neurons (GLYC, GABA, GLUT). Key time points (E11.5–E17.5) for differentiation were  
5 indicated. **(C–E)** Pathway-regulon activity heatmaps showing changes in the relative  
6 expression of shared genes between regulons and pathways (y-axis) and associated specific  
7 regulon activity (x-axis) in radial glial cells **(C)**, glioblasts **(D)**, and neuroblasts **(E)** in *R26-DTA*  
8 and *NesCre::R26-DTA* chimeras. Triangles represented regulons associated with the  
9 corresponding pathways (relative gene expression levels of shared gene sets) that were  
10 significantly higher (**▲**) or lower (**▼**) in their activity in *NesCre::R26-DTA* chimeras. The color  
11 of the triangles indicated statistical significance, as represented by the  $-\log_{10}(p)$  values,  
12 shown in the vertical bar at the top-right corner. The color scale at the bottom indicated regulon  
13 activity, with red representing higher activity and blue representing lower activity in  
14 *NesCre::R26-DTA* compared to *R26-DTA* chimeras.

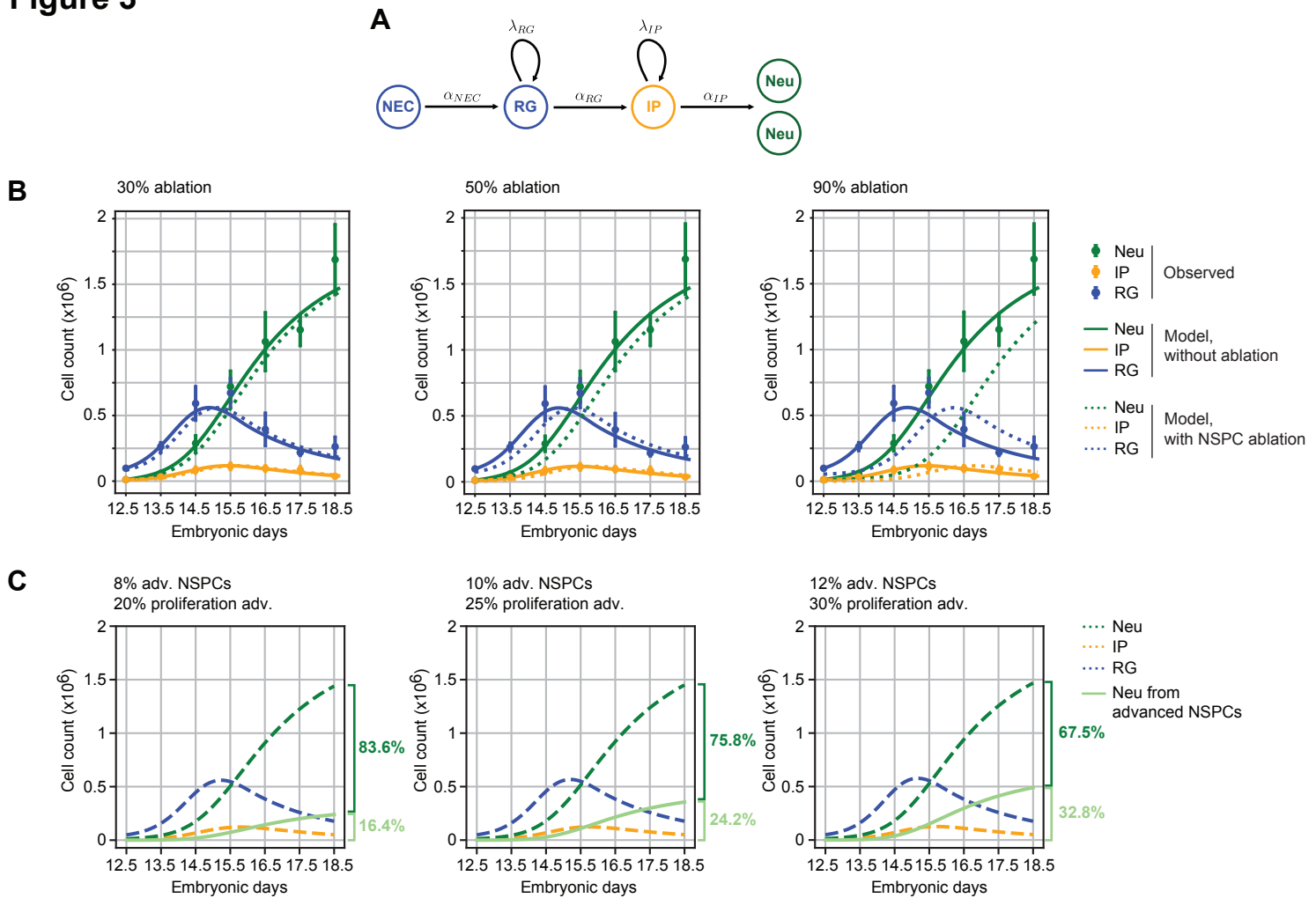
## Figure 1



## Figure 2

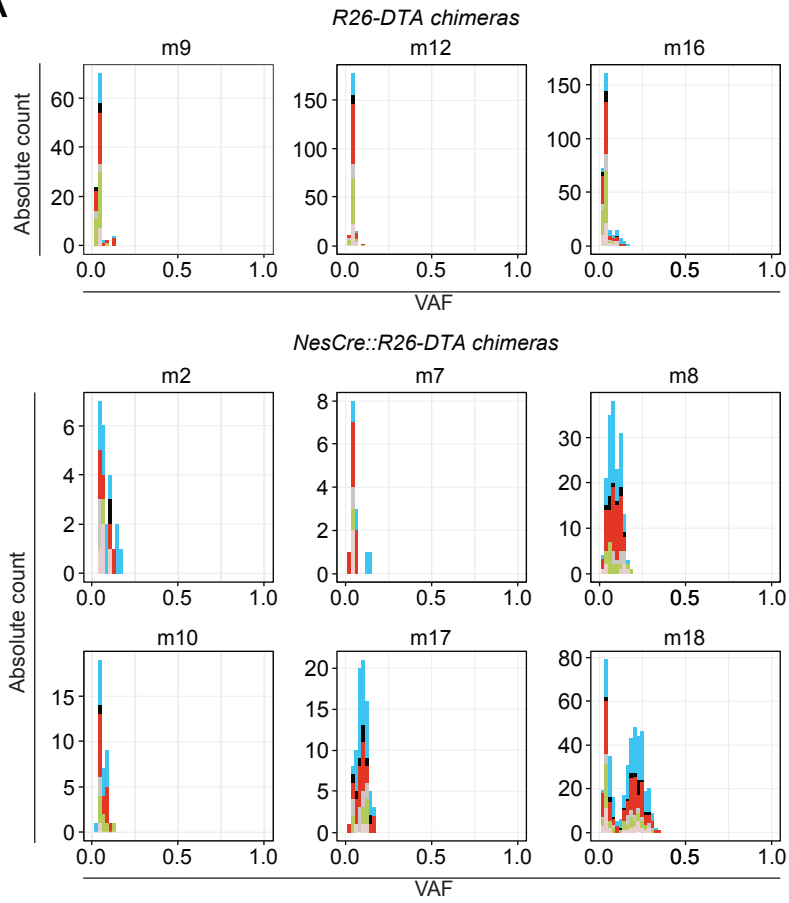


## Figure 3

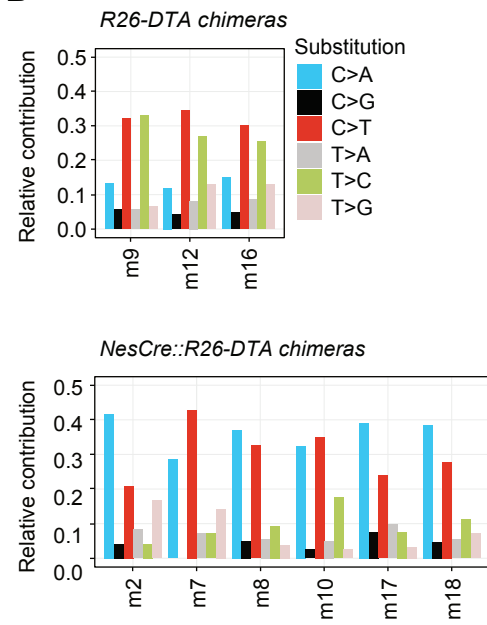


## Figure 4

**A**



**B**



**Figure 5**

