TRANSPARENT
PROCESS

THE EMBO JOURNAL

Transcriptome sequencing during mouse brain development identifies long non-coding RNAs functionally involved in neurogenic commitment

Julieta Aprea[1,5], **Silvia Prenninger**[1,5],
Martina Dori[1], **Tanay Ghosh**[2],
Laura Sebastian Monasor[1],
Elke Wessendorf[1], **Sara Zocher**[1],
Simone Massalini[1], **Dimitra Alexopoulou**[1],
Mathias Lesche[1], **Andreas Dahl**[1],
Matthias Groszer[2], **Michael Hiller**[3,4] and
Federico Calegari[1,*]

[1]DFG–Research Center and Cluster of Excellence for Regenerative Therapies, Dresden, Germany, [2]Institut du Fer à Moulin, INSERM UMR-S 839, Université Pierre and Marie Curie, Paris, France, [3]Max Planck Institute of Molecular Cell Biology and Genetics, Dresden, Germany and [4]Max Planck Institute for the Physics of Complex Systems, Dresden, Germany

Transcriptome analysis of somatic stem cells and their progeny is fundamental to identify new factors controlling proliferation versus differentiation during tissue formation. Here, we generated a combinatorial, fluorescent reporter mouse line to isolate proliferating neural stem cells, differentiating progenitors and newborn neurons that coexist as intermingled cell populations during brain development. Transcriptome sequencing revealed numerous novel long non-coding (lnc)RNAs and uncharacterized protein-coding transcripts identifying the signature of neurogenic commitment. Importantly, most lncRNAs overlapped neurogenic genes and shared with them a nearly identical expression pattern suggesting that lncRNAs control corticogenesis by tuning the expression of nearby cell fate determinants. We assessed the power of our approach by manipulating lncRNAs and protein-coding transcripts with no function in corticogenesis reported to date. This led to several evident phenotypes in neurogenic commitment and neuronal survival, indicating that our study provides a remarkably high number of uncharacterized transcripts with hitherto unsuspected roles in brain development. Finally, we focussed on one lncRNA, Miat, whose manipulation was found to trigger pleiotropic effects on brain development and aberrant splicing of Wnt7b. Hence, our study suggests that lncRNA-mediated alternative splicing of cell fate determinants controls stem-cell commitment during neurogenesis.

*Corresponding author. CRTD, Technische Universität Dresden, Fetscherstrasse 105, Dresden 01307, Germany. Tel.: +49 0 351 45882204; Fax: +49 0 351 45882209; E-mail: federico.calegari@crt-dresden.de
[5]Joint first authors.

Introduction

Sequencing technologies allow the analysis of genomes and transcriptomes at an unprecedented coverage and speed and were readily adopted in a number of studies including genome-wide profiling of epigenetic marks, personal genomics and sequencing of extinct species (Metzker, 2010). For stem-cell research, next-generation sequencing has been predominantly used for the study of pluripotent embryonic stem cells (Mikkelsen *et al*, 2007; Meissner, 2010) for which relatively homogeneous cell populations can be grown in culture. In contrast, and limiting the use of transcriptome sequencing in other tissues, somatic stem cells are intermingled with more differentiated progenitors and various types of terminally differentiated cells, making it difficult to isolate highly enriched pools of individual cell types.

Specifically, during embryonic development of the mammalian cortex neuroepithelial stem cells expand by undergoing mitosis at the apical boundary of the ventricular zone (VZ); hence, they are referred to as apical progenitors (APs). As development proceeds, an increasing proportion of APs switches from proliferative to differentiative divisions to generate either basal progenitors (BPs) that leave the VZ to form the subventricular zone (SVZ) or neurons. While most APs continue to proliferate, the majority of BPs undergo neurogenic divisions to generate two postmitotic neurons that migrate through the intermediate zone (IZ) to form the cortical plate (CP) (Götz and Huttner, 2005). Notably, both APs and BPs can undergo proliferative as well as differentiative divisions but to a different degree with studies indicating that at embryonic day (E) 14.5 about 60% of APs are proliferative progenitors (PPs) while only about 20% of BPs remain PPs (Attardo *et al*, 2008; Arai *et al*, 2011). Correspondingly, the remaining APs and BPs switch their fate to become differentiating progenitors (DPs) to generate neuronal-committed BPs or postmitotic neurons, respectively (Götz and Huttner, 2005). Thus, PPs represent the pool of symmetrically expanding cells generating daughters that are cell biologically identical to their mother. In contrast, DPs generate at least one daughter with a more restricted potential and depleting the progenitor pool. To understand the mechanisms controlling the transition from proliferation to differentiation, systems are required that allow the identification of progenitors from neurons while, at the same time, distinguishing between the two intermingled pools of PPs and DPs. Although conceptually simple, achieving this specificity was revealed to be a major challenge.

Several studies have addressed this problem by generating reporter mice in which, for instance, BPs (Kwon and Hadjantonakis, 2007), DPs (Haubensak *et al*, 2004) or neurons (Attardo *et al*, 2008) were identified by the expression of an endogeneous fluorescent protein (typically

GFP) under a marker-specific promoter (e.g., *Tbr2/Eomes*, *Btg2* or *Tubb3* respectively). However, the transient nature of these cell populations together with the inheritance of the reporter protein from a dividing mother cell to her progeny typically limited the analysis to tissue sections where location (VZ, SVZ or IZ/CP) was used as a proxy for cell identity. For these reasons, transcriptome analyses using single-reporter lines had to be complemented with various strategies to try to increase the cell homogeneity, for example, by limiting the comparison to different developmental stages (Matsuki *et al*, 2005; Hartl *et al*, 2008; Ling *et al*, 2009), microdissecting randomly selected cells to retrospectively deduce cell identity (Kawaguchi *et al*, 2008) or exclusively analyse cells in S phase (Arai *et al*, 2011). Moreover, previous expression profiles comparing stem and progenitor cells (Pinto *et al*, 2008), or neurons (Faux *et al*, 2010), during development were mostly derived from mRNA microarrays that are limited with regard to transcriptome coverage, sensitivity and quantification of transcripts. To our knowledge, only four studies have used next-generation sequencing during physiological corticogenesis by, again, adopting different strategies to try to enrich specific cell types including selecting small pools of microdissected cells (Ayoub *et al*, 2011), comparing developmental stages (Han *et al*, 2009; Yao *et al*, 2012) or different species (Fietz *et al*, 2012).

Here, we sought to combine direct and rigorous isolation of PPs, DPs and neurons with deep sequencing to interrogate transcriptomes for signatures specific to the onset of differentiation. To this aim, we generated a combinatorial RFP and GFP reporter mouse line and sequenced the transcriptomes of the three sub-populations of PPs (RFP–/GFP–), DPs (RFP+/GFP–) and neurons (GFP+) coexisting in time and space during corticogenesis.

Results

Generation of Btg2^RFP and Btg2^RFP/Tubb3^GFP reporter lines

To identify DPs by RFP expression, we chose the promoter of *Btg2* (also known as *Tis21* or *PC3*) because this gene is expressed in early G1 specifically in DPs but not in PPs or neurons (Iacopetti *et al*, 1999) and the use of a previous *Btg2*^GFP line proved to be instrumental in a number of studies of embryonic corticogenesis including lineage tracing of DPs by time lapse microscopy (Haubensak *et al*, 2004; Arai *et al*, 2005; Calegari *et al*, 2005; Attardo *et al*, 2008). We inserted the coding sequence of a nuclear-localized RFP into the gene's first exon encoded within a bacterial artificial chromosome (Figure 1A) and used the resulting construct for oocyte pronuclear injection.

Heterozygous *Btg2*^RFP mouse embryos displayed endogeneous RFP fluorescence along the neural tube with an onset at the level of the spinal cord/hindbrain at E9.5 extending to the midbrain at E10.5 and reaching the telencephalon at E11.5 and, thus, faithfully recapitulating the caudal-to-rostral gradients of neurogenesis (Figure 1B). *In situ* hybridization on E14.5 brain sections revealed that RFP transcripts were abundant in the VZ and the SVZ but virtually absent in the IZ/CP (Figure 1C). In contrast, fluorescence microscopy revealed RFP+ nuclei along the entire apico-basal axis of the E14.5 lateral cortex with scattered cells in the VZ, a denser distribution in the SVZ and most cells being RFP+

in the IZ/CP (Figure 1D and F, red). Using Pax6, Tbr2 (i.e., Eomes) and Tbr1 as markers of APs, BPs and neurons, respectively (Hevner *et al*, 2006), we found that *ca.* 60% of Pax6+/Tbr2– APs in the VZ were RFP–, *ca.* 80% of Tbr2+ BPs in the VZ and SVZ were RFP+ and essentially all (>95%) Tbr1+ neurons in the SVZ, IZ or CP were also RFP+ (Supplementary Figure S1A). The gradient of *Btg2*^RFP expression during development (Figure 1B) and the proportion of *Btg2*^RFP+ cells within APs, BPs and neurons (Supplementary Figure S1A) fit well with the known pattern of Btg2 mRNA and protein expression (Iacopetti *et al*, 1999) and data in *Btg2*^GFP reporter mice (Haubensak *et al*, 2004; Arai *et al*, 2011), suggesting that expression of *RFP* mRNA in *Btg2*^RFP embryos begins in DPs while the RFP protein is subsequently inherited by newborn neurons. Also consistent with the reported expression of *Btg2* in adult tissues (Terra *et al*, 2008; Attardo *et al*, 2010), scattered RFP+ cells were found in the adult hippocampus (Supplementary Figure S1B), subependymal zone and other organs including testis, skeletal muscle and kidney (Supplementary Figure S1C, and data not shown).

To further validate our *Btg2*^RFP line, we took advantage of the extensively characterized *Btg2*^GFP knock-in reporter (Haubensak *et al*, 2004; Calegari *et al*, 2005; Attardo *et al*, 2008; Arai *et al*, 2011) and crossed *Btg2*^RFP and *Btg2*^GFP mice to quantify the degree of colocalization of the two transgenes in the VZ and the SVZ, that is, where PPs and DPs reside. Cryosections of double heterozygous *Btg2*^RFP/*Btg2*^GFP E14.5 embryos revealed a very high degree of colocalization with the vast majority (*ca.* 90%) of fluorescent cells being positive for both reporters and the remaining RFP+/GFP– or RFP–/GFP+ cells being equally represented (*ca.* 5% each) (Figure 1D). Double RFP+/GFP+ cells were observed already in mitosis (Figure 1D′) and throughout the VZ and SVZ although intensity levels of the two reporters not always correlated. In contrast, the IZ/CP showed a substantial persistence of RFP inherited by newborn neurons (Figure 1F, red) that seemed to be more significant than *Btg2*-driven GFP. Clearly, differences in intensity and persistence of fluorescence in daughter cells can be ascribed to the different time required for the maturation/degradation of the two reporters and to the different strategies used to obtain the two mouse lines (pronuclear injection versus knock-in, respectively). Nevertheless, the high degree of colocalization of *Btg2*-driven RFP and GFP and the many reports validating the use of the *Btg2*^GFP line (Haubensak *et al*, 2004; Calegari *et al*, 2005; Attardo *et al*, 2008; Arai *et al*, 2011) led us to conclude that our new *Btg2*^RFP reporter is equally well suited to reliably identify DPs; at least when RFP+ cells were scored within the VZ and the SVZ.

To identify neurons that inherited, but did not express, RFP we crossed *Btg2*^RFP mice with a characterized *Tubb3*^GFP reporter in which GFP is selectively expressed in newborn neurons as one of the earliest events upon mitosis of a neurogenic progenitor (Attardo *et al*, 2008). Double heterozygous *Btg2*^RFP/*Tubb3*^GFP embryos (Figure 1E and F) displayed a virtually complete (>95%) colocalization of GFP with the early neuronal markers Tubb3 and Tbr1 (not shown). Taken together, our results and previous reports (Haubensak *et al*, 2004; Attardo *et al*, 2008) validate the use of *Btg2*^RFP/*Tubb3*^GFP embryos to rigorously discriminate PPs (RFP–/GFP–), DPs (RFP+/GFP–) and neurons (RFP+/
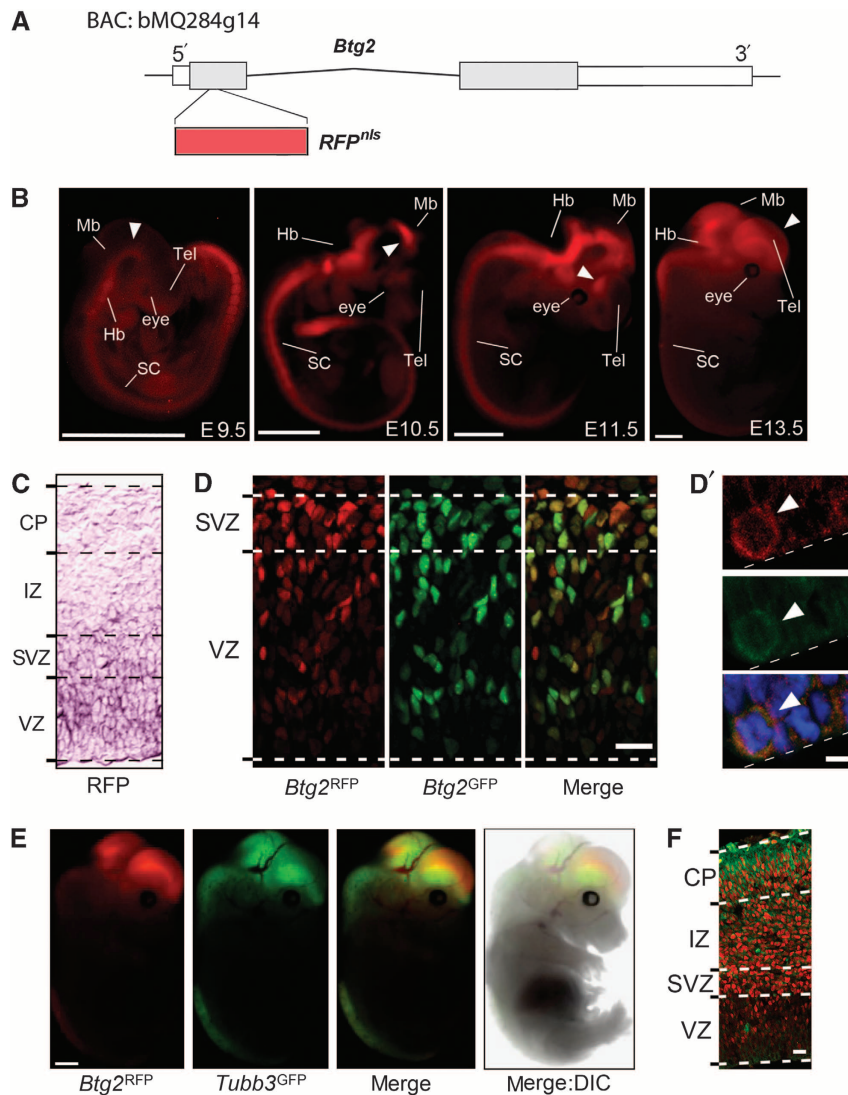
Figure 1 Generation of *Btg2*^{RFP} and *Btg2*^{RFP}/*Tubb3*^{GFP} mouse lines. (**A**) Map (exons not in scale) of the Btg2 locus within a bacterial artificial chromosome (BAC) showing the insertion of a nuclear-localized (nls) RFP in the gene first exon. (**B**) Whole-mount pictures of E10.5-E13.5 *Btg2*^{RFP} mice showing endogeneous RFP fluorescence. Note the gradients of RFP expression (arrowhead) along the forebrain. (**C**) Picture of a cryosection through the lateral cortex of an E14.5 *Btg2*^{RFP} mouse embryo after *in situ* hybridization using RFP probes. (**D–D′**) Fluorescence pictures of the E14.5 VZ/SVZ (**D**) and high-power view of a mitotic APs (**D′**; arrowhead) of a *Btg2*^{RFP}/*Btg2*^{GFP} embryo displaying (**D** = left to right; **D′** = top to bottom) single RFP or GFP fluorescence and their colocalization. (**E**) Whole-mount pictures of an E14.5 *Btg2*^{RFP}/*Tubb3*^{GFP} embryo showing individual and merged RFP, GFP and DIC channels. (**F**) Fluorescence picture of a cryosection through the lateral cortex of an E14.5 *Btg2*^{RFP}/*Tubb3*^{GFP} embryo. SC = spinal cord; Hb = hindbrain; Mb = midbrain; Tel = telecephalon; VZ = ventricular zone; SVZ = subventricular zone; IZ = intermediate zone; CP = cortical plate. Dashed lines in (**C**), (**D**) and (**F**) indicate boundaries between cortical layers. Scale bars = 1 mm (**B, E**), 20 μm (**D, F**) and 5 μm (**D′**).

GFP +) (henceforth referred to as RFP−, RFP + and GFP +, respectively).

Sorting and transcriptome sequencing of PPs, DPs and neurons

Cell sorting was performed after crossing double hetero-zygous *Btg2*^{RFP}/*Tubb3*^{GFP} with wild-type C57/Bl6 mice and selecting E14.5 embryos according to their colours as identified by whole-mount fluorescent stereomicroscopy. A Mendelian proportion of RFP + and/or GFP + embryos were observed and their brains were collected to obtain single-cell suspensions after dissection of the lateral cortex and removal of meninges. FAC sorting revealed a continuous gradient of RFP expression and two distinct populations of

GFP− and GFP + cells (Figure 2A). Therefore, thresholds for RFP were chosen to mimic the proportion of RFP− and RFP + progenitors (i.e., after excluding GFP + neurons) as judged by fluorescent microscopy in the VZ and the SVZ (Figure 1D; Supplementary Figure S1A) and corresponding to *ca.* 30 and 40% of RFP− and RFP + cells, respectively, while discarding the remaining 30% with intermediate levels of fluorescence as cells of dubious identity (Figure 2A; Supplementary data; Supplementary Figure S2A). Validating our gating parameters, western blot analyses on freshly sorted cells revealed that known markers of PPs, DPs and neurons (Sox2, Tbr2 and Tubb3) were significantly enriched in RFP−, RFP + and GFP + cells, respectively (Supplementary Figure S2C). Moreover, in line with previous reports showing a
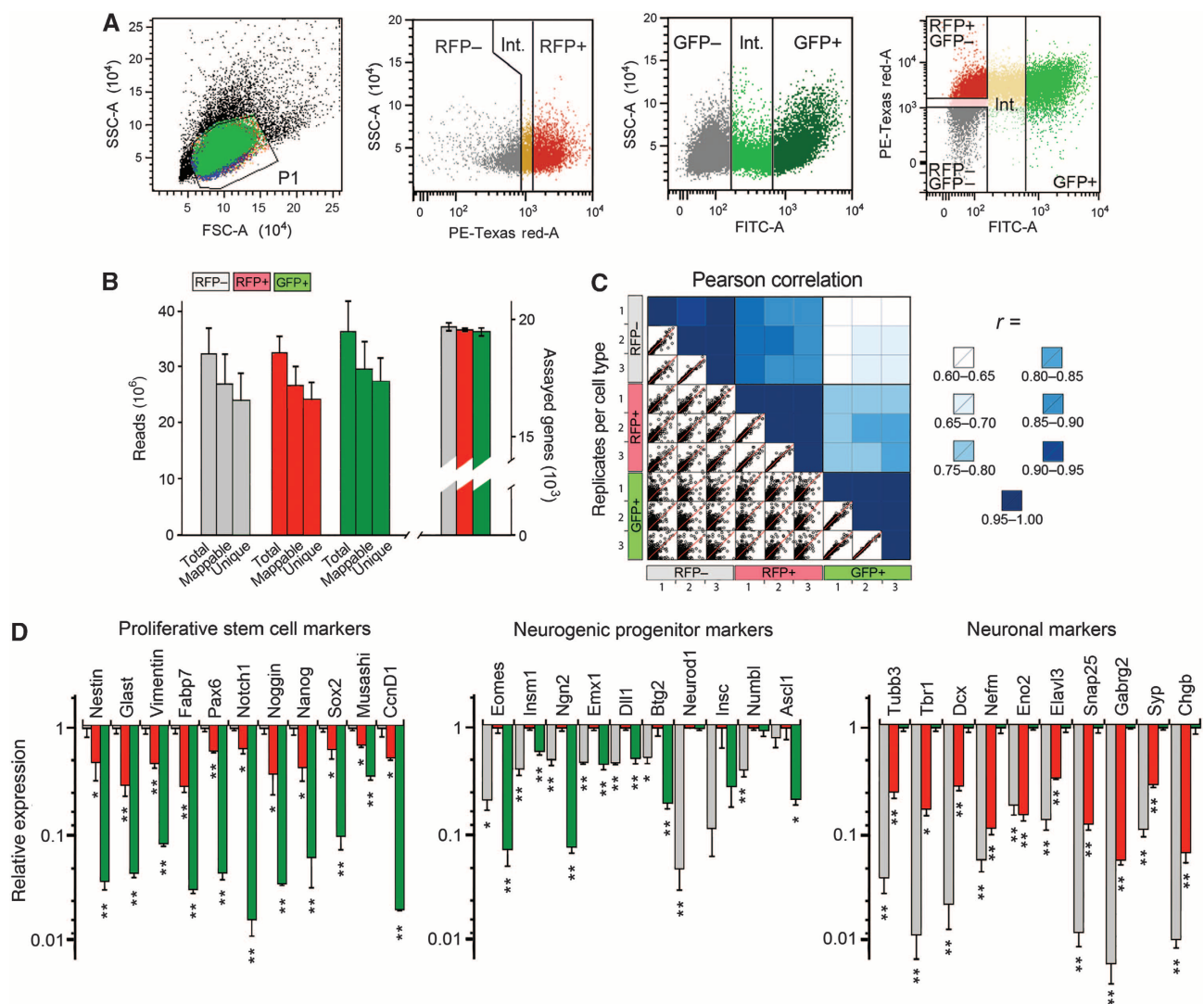
Figure 2 Isolation and sequencing of progenitors. (**A**, left to right) Gating parameters (lines) used to sort single (P1) and individual, or combined, RFP/GFP fluorescent (red, green and yellow as merge) cells. Interspaces (int.) used as thresholds are indicated. (**B**) Total, mappable and unique reads (left axis) plus assayed genes (right axis) of RFP– (grey), RFP + (red) and GFP + (green) cells. (**C**) Sample to sample Pearson's correlation (blue; top-right quadrants) and pairwise comparison (dotted boxes; bottom-left quadrants) of DESeq-normalized gene expression between three biological replicas (1–3) and cell types (colours). (**D**) Expression of established markers of undifferentiated neural stem cells (left), neurogenic progenitors (middle) and neurons (right). Normalized counts in cell types (colours) are represented on a logarithmic scale relative to the population of reference within each panel. By this, expression of markers of PPs, DPs and neurons is defined = 1 in the left, middle and right panels, respectively. Bars = s.d.; *$P < 0.05$; **$P < 0.005$.

longer G1 in DPs relative to PPs (Calegari *et al*, 2005; Arai *et al*, 2011), FACS analysis revealed that a higher proportion (*ca.* 10% increase) of RFP + relative to RFP– cells were in G1 (Supplementary Figure S2A and B).

Libraries for massive parallelized sequencing were prepared from each population in three biological replicates. Sequencing was performed on the Illumina HiSeq2000 platform, resulting in 30–40 million reads per sample, a depth sufficient to achieve high transcriptome coverage for robust differential analyses of gene expression (Tarazona *et al*, 2011). We uniquely mapped 70–80% of reads to the mouse genome, corresponding to ∼20 000 genes per sample (Figure 2B). Library diversity was additionally assessed by investigating redundancy within the mapped reads, which showed a high degree of coverage starting at almost 90% for a random subsample of 1 million reads (Supplementary Figure S2D). No difference in transcript length distributions

was found in the three populations (Supplementary Figure S2E and E').

Sample to sample correlation of normalized gene expression resulted in a range between biological replicates that was highly reproducible ($r = 0.98 \pm 0.02$) with RFP– cells being more closely associated to RFP + ($r = 0.89 \pm 0.05$) than each was to GFP + ($r = 0.65 \pm 0.02$ and $r = 0.79 \pm 0.02$, respectively) cells (Figure 2C), which is consistent with the lineage PPs→DPs→neurons. As a first validation of our sequencing data, we took advantage of an extensive literature and analysed the expression of several genes widely accepted to mark different cell types in the embryonic cortex (Götz and Huttner, 2005; Hevner *et al*, 2006; Guillemot, 2007). For APs/PPs, these included nestin, Glast (Slc1a3), vimentin, Fabp7, Pax6 as well as markers of proliferating neural (and other somatic) stem cells such as Notch1, noggin, Nanog, Sox2 and musashi that were all >2-fold down-regulated in RFP +

relative to RFP– cells and virtually absent (i.e., 10- to 100-fold down-regulated) in GFP+ cells (Figure 2D, left). Conversely, markers of BPs/DPs, including Tbr2, Insm1, Neurog2, Emx1, Dll1 and Btg2 itself as well as markers of early neurogenic commitment such as Neurod1, Insc, Numbl and Ascl1 (commonly known as Mash1) were also >2-fold down-regulated in both RFP– and GFP+ relative to RFP+ cells (Figure 2D, middle). Furthermore, well-characterized neuronal markers such as Tubb3, Tbr1, Dcx as well as neuronal-specific cytoskeletal and synaptic genes, pumps, channels and receptors (Nefm, Eno2, Elavl3, Snap25, Gabrg2, Syp and Chgb) were all virtually absent (i.e., 10- to 100-fold down-regulated) in both RFP– and RFP+ as compared to GFP+ cells (Figure 2D, right). As one additional feature, cell-cycle length of APs/PPs is known to be shorter than that of BPs/DPs (Salomoni and Calegari, 2010) and, accordingly, RFP+ cells displayed a decrease relative to RFP– cells in the expression of key cyclins, most prominently cyclin D1/D2 (Ccnd1/2), and increased levels of antiproliferative genes including Rb1, Cdkn1a (p21) and Cdkn1b (p27) (Figure 2D, left; Supplementary File 1).

Taken together, the analysis of tissue sections (Figure 1D and F; Supplementary Figure S1A), sorted cells (Figure 2A; Supplementary Figure S2A–C) and transcript levels of over 50 well-established markers (Figure 2D) demonstrated that $Btg2^{RFP}/Tubb3^{GFP}$ embryos can be used to isolate highly enriched pools of PPs, DPs and neurons allowing us to provide a comprehensive description of their transcriptional profiles (Supplementary File 1; raw data deposited in GEO GSE51606).

Differentially expressed genes

We next sought to identify the genes that were up-regulated by >50% in one progenitor pool relative to the other i.e. in RFP – relative to RFP+ or, reciprocally, RFP+ relative to RFP – (the latter indicating down-regulation) (FDR 5%). A similar criterion was used to identify up-/down-regulated genes in DP versus neurons (Supplementary File 1). Considering that the lineage from PPs to neurons requires, per definition, the intermediate population of DPs, we thought that differences between PPs and neurons were not biologically relevant with regard to understanding stem cell commitment and, thus, these will not be further discussed.

We found that the expression of the vast majority of transcripts was not significantly changed during the transition neither from PPs to DPs (*ca.* 90%) nor from DPs to neurons (*ca.* 75%) (Figure 3A and B) and that among differentially expressed transcripts a similar proportion was either up- or down-regulated corresponding to about 6% of transcripts from PPs to DPs and 13% from DPs to neurons (Figure 3B). Interestingly, we also found that the vast majority of genes being up-regulated during the transition from PPs to DPs continued to be up-regulated, or remained constant, also in the transition from DPs to neurons while, conversely, genes down-regulated in DPs continued to be downregulated, or remained constant, in neurons (*ca.* 85% in either case) (Figure 3B). Patterns of up-/down-regulation in the three cell types were revealed to be remarkably symmetric (Figure 3B).

Gene ontology analysis (DAVID) for functional enrichment of differentially expressed genes in RFP+ relative to either RFP– or GFP+ cells revealed a preponderance of terms related to neuronal differentiation, axo-/dendro-genesis,

synaptic transmission, ion transport and cell cycle (Figure 3C and D, top; Supplementary File 2). Less expected was the finding that functional enrichment analysis could not identify any major distinction between genes up-regulated in the two groups of RFP– versus RFP+ and RFP+ versus GFP+ cells with most terms (neuronal differentiation, migration, axo-, dendro- and synapto-genesis) being either generic or common to both (Figure 3C and D, middle; Supplementary File 2), which is likely explained by the large number of transcripts that are commonly up-regulated in both DPs and neurons (discussed below). On the other hand, functional annotation of down-regulated genes was more distinct and specific with adhesion, polarity and extracellular matrix characterizing the transition from RFP– to RFP+ and cell cycle, DNA replication and cytoskeleton characterizing the one from RFP+ to GFP+ (Figure 3C and D, bottom; Supplementary File 2). These ontology terms are consistent with DPs down-regulating polarity to leave the VZ and form the SVZ and neurons down-regulating cell cycle and DNA replication to become postmitotic.

Switch genes identify the signature of neurogenic commitment and include numerous uncharacterized protein-coding transcripts

We next sought to identify the set of genes that were specifically up-/down-regulated in the transient DPs population as compared to both PPs and neurons. This was particularly important because the vast majority of transcripts enriched in GFP+ neurons, including Tubb3 and other axonal, synaptic and cytoskeletal markers, started to be up-regulated already at the level of DPs (Figure 2D, right). More generally, the consistent trend of up-/down-regulation displayed in the transition both from PPs to DPs and from DPs to neurons (i.e., PPs<DPs<neuron and PPs>DPs> neuron) (Figure 3B) suggested that a high proportion of differentially expressed genes is implicated in neuronal specification and maturation without being necessarily involved in the switch from PPs to DPs proper; which in turn explains the common and generic functional enrichment terms of up-regulated genes (Figure 3C and D). In such a case, comparing only two cell types (as typically done in previous transcriptome analyses) would likely lead to misleading conclusions. In contrast, comparing the expression profiles of the three populations together allowed us to derive a DPs-specific gene signature that is more specific in revealing stem-cell commitment from proliferation to neurogenesis.

To this end, we extracted genes that were up-regulated by >50% in DPs as compared to both PPs and neurons simultaneously (i.e.: PPs<DPs>neuron, for 'on-switch' genes) or, alternatively, up-regulated by >50% in PPs and neurons as compared to DPs (i.e.: PPs>DPs<neuron, for 'off-switch' genes) (FDR 5%). This yielded 415 genes (Figure 4A and B) representing <2% of all (21 210) transcripts and ~15% of all genes differentially expressed between PPs and DPs (2627). We validated the expression pattern of switch genes in tissue using genome-wide atlases of gene expression of the E14.5 mouse brain including Eurexpress and the Allen Brain Atlas (Lein *et al*, 2007; Diez-Roux *et al*, 2011). Essentially all switch genes included in these resources (*ca.* 80%) displayed a pattern of expression across the lateral cortex that was entirely consistent with their expression profiles in RFP–,
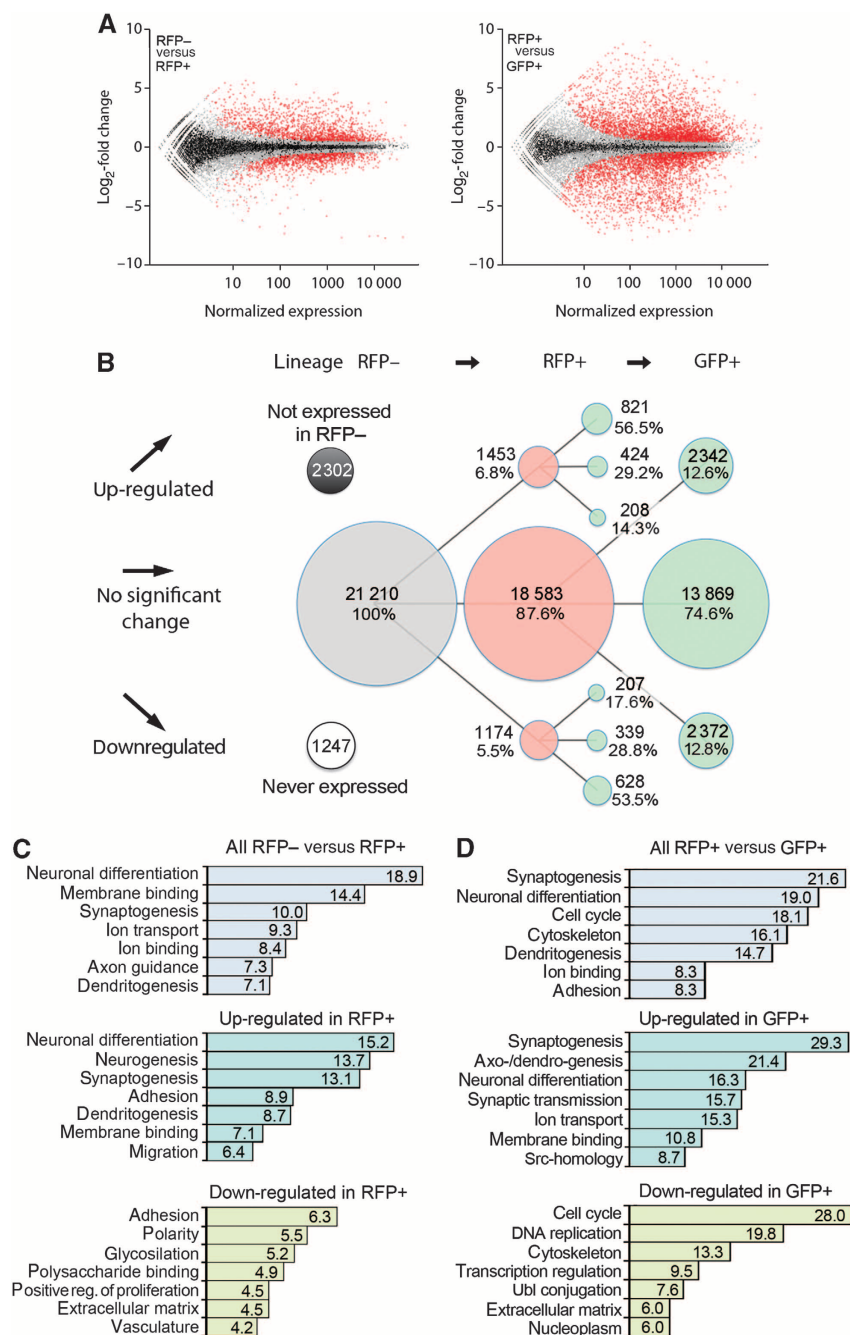
Figure 3 Differentially expressed genes. (**A**) Differential expression among RFP– and RFP+ (left) or RFP+ and GFP+ (right) cells. DESeq-normalized mean expression (*x* axis; $\log_{10}$ scale) and fold changes (*y* axis; $\log_2$ values) of differentially expressed (red) or unchanged (grey) genes (5% FDR) are indicated. Black dots indicate overlapping grey dots. (**B**) Representation of up- or down-regulation (lines pointing up or down, respectively) along the neurogenic lineage from PP to DPs and DPs to neurons (grey, red and green, respectively). The area of circles represents the number of transcripts detected in each population with number and proportion relative to the parental population (connecting lines) being indicated within each. Transcripts not expressed in RFP– or never detected in any population are also indicated (left; black and white, respectively). Horizontal lines include both not significantly changed transcripts and those changing <50%. (**C, D**) DAVID-functional annotation analysis of differentially expressed genes in RFP– versus RFP+ (**C**) or RFP+ versus GFP+ (**D**) cells. All (top), up- (middle) or down- (bottom) regulated transcripts were analysed separately and enrichment scores of the top 7 indicated (Supplementary File 2).

RFP+ and GFP+ cells with on-switches being detectable in the VZ and, more strongly, in the SVZ while, conversely, off-switches were found to be specifically depleted in the SVZ but enriched in both the VZ and the IZ/CP (Supplementary Figure S3).

We found 208 on-switch genes (Figure 4A) including most BPs/DPs markers mentioned above such as Tbr2, Neurog2,

Insm1, Emx1 and, clearly, Btg2 itself. Notably, on-switch genes included many other transcripts that were implicated in neurogenesis only recently such as Cbfa2t2, Chd7, Ezh2 and Ncor2 (Jepsen *et al*, 2007; Aaker *et al*, 2009; Pereira *et al*, 2010; Engelen *et al*, 2011). Conversely, the list of 207 off-switch genes (Figure 4B) was lacking widely used markers of cortical development and only very few were found to be
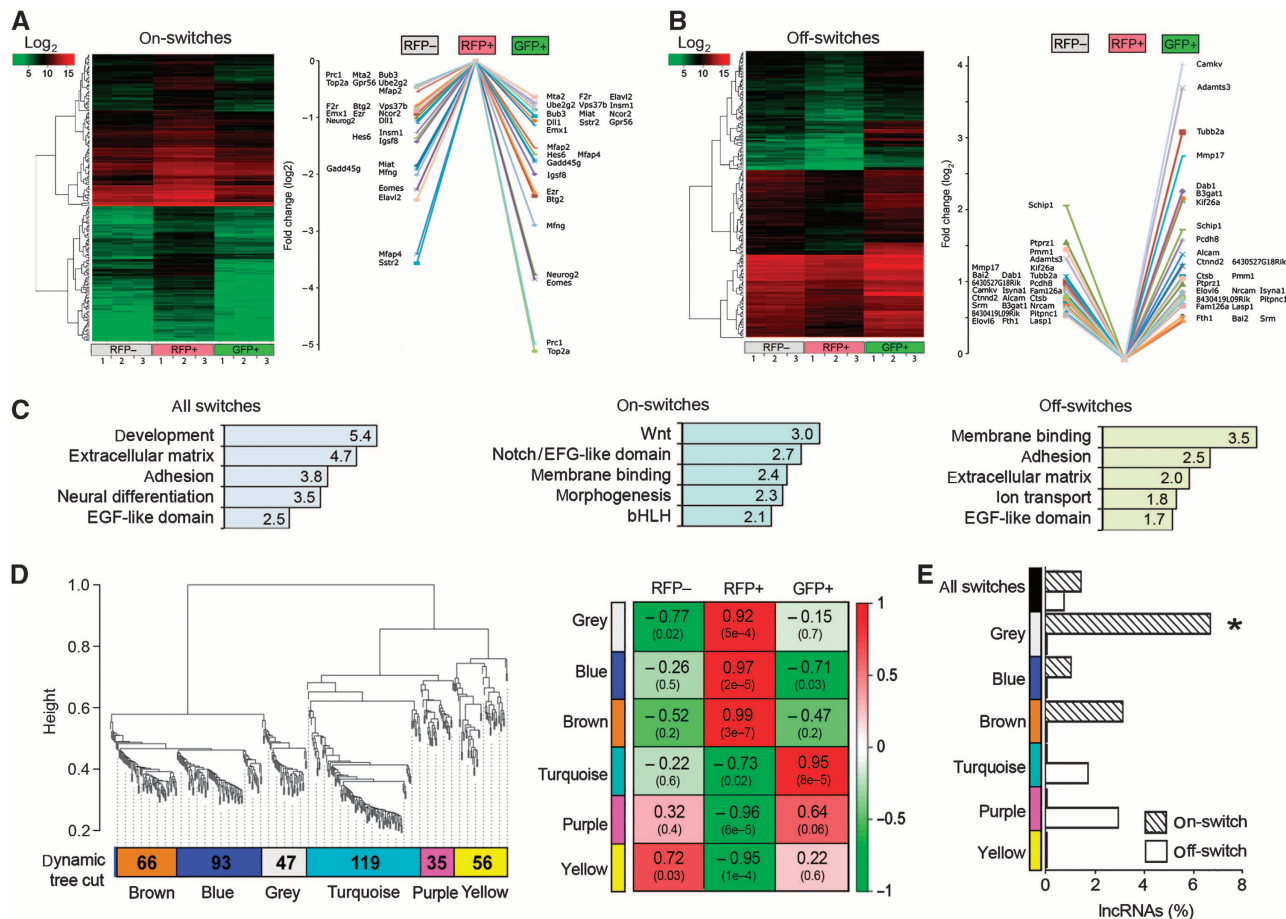
Figure 4 Switch genes identify the signature of DPs. (**A**, **B**) Heat maps representation of all on- (**A**) and off- (**B**) switch genes (left) and differential gene expression (y axis = $\log_2$ values) of 25 with the higher expression in RFP + cells (right) (Supplementary File 4). Note in (**A**) the presence of several markers of neurogenic commitment. (**C**) DAVID-functional annotation and enrichment scores of the top five terms of all (left), on- (middle) and off- (right) switch genes. (**D**) Hierarchical cluster tree (left) and association (right) of modules Eigengene (rows) with cell types (columns) are indicated. Numbers correspond to correlation and *P*-value (brackets). Negative values represent down-regulation. Colours represent modules (grey, blue, brown, etc.) or their correlation value (green-to-red gradient) (Supplementary File 3). (**E**) Proportion of switch lncRNAs within each module. Asterisk in grey indicates a significant overrepresentation ($P < 0.05$, hypergeometric distribution).

linked to neurogenesis to date such as Ptprz1, Alcam, Nrcam and Rorb (Diekmann and Stuermer, 2009; Lamprianou *et al*, 2011; Jabaudon *et al*, 2012; Sakurai, 2012). Functional annotation analysis of on-switch genes yielded terms closely associated with known cell fate determinants including Wnt, Notch and bHLH transcription factors, while off-switches were associated with epithelial polarity, membrane binding, cell adhesion and extracellular matrix (Figure 4C; Supplementary File 4).

So far our description was limited to genes known to play roles in neurogenesis as a means to validate our approach. However, the purpose of our study was to comprehensively uncover new factors and uncharacterized genes potentially involved in this process. To this aim, we noticed that several switch transcripts had only an automatic annotation number (i.e., *Rik, GM* or Fam* symbols) instead of a name, which is often the case for uncharacterized genes. Therefore, we identified switch genes that had no annotation in the Molecular Function and Biological Process Gene Ontology as a proxy for lack of detailed characterization and found 63 (15% of 415) without any descriptive GO term (Supplementary File 3). This proportion of uncharacterized genes seems to be remarkably high if one considers that

all displayed highly significant, robust and transient changes in expression levels in a cell type-specific manner. Yet, the vast majority of these transcripts have never been studied, in any context, to date.

Apart from single genes, the transition from proliferation to differentiation can be further defined by ensembles of genes (modules) having specific functional relations. To identify coregulated modules, we performed weighted gene coexpression network analysis (WGCNA) that proved to be useful in a number of neurodevelopmental and disease contexts (Geschwind and Konopka, 2009). Gene clusters were constructed based on expression counts in PPs, DPs and neurons yielding six modules comprised by genes with a more similar expression pattern (Figure 4D; Supplementary File 3). In analysing the distinct modules, we found that all 10 on-switch genes annotated with the GO Term 'Wnt signalling' were in 2 of the 6 modules (Figure 4D. brown and blue). Furthermore, we found that one module was significantly enriched in long non-coding RNAs (lncRNAs) where they were proportionally overrepresented (6%; $P < 0.05$) (Figure 4E, grey). The very limited knowledge on lncRNAs during cortical development prompted us to further investigate their roles in PPs, DPs and neurons.

Switch genes include several known and novel lncRNAs overlapping protein-coding, neurogenic switch genes

LncRNAs are a major class of transcripts in eukaryotic genomes (Okazaki *et al*, 2002) with roles in many biological processes including chromatin remodelling, transcriptional regulation, splicing and regulation of stem-cell dynamics during embryogenesis (Mercer *et al*, 2009; Nagano and Fraser, 2011; Pauli *et al*, 2011).

Identification of lncRNAs in our data sets was based on manually annotated databases (Amaral *et al*, 2011) and validation of lack of long open reading frames and signatures of protein-coding conservation according to RNAcode. Consistent with recent reports showing the abundance of lncRNAs in the developing and adult central nervous system (Derrien *et al*, 2012; Ramos *et al*, 2013), we found among switch transcripts two previously described lncRNAs, Miat and Rmst, both of which were suggested to play roles in neurogenesis (Blackshaw *et al*, 2004; Uhde *et al*, 2010; Ng *et al*, 2013). In addition, we found seven other lncRNAs annotated in Ensembl but still uncharacterized. Of these nine switch lncRNAs, six were on-switches (Miat, Rmst, Gm17566, Gm14207, Gm16758 and 2610307P16Rik) and three were off-switches (AC102815.1, C230034O21Rik and 9930014A18Rik) (Supplementary File 3). Moreover, three lncRNAs (Miat, Rmst and 2610307P16Rik) had ortho-logous human transcripts and are therefore clearly conserved while four lncRNAs (Gm14207, Gm17566, Gm16758 and 9930014A18Rik) showed sequence conservation of parts of their exons, potentially indicating a conserved RNA whose human transcript has not been annotated yet.

As the catalogue of lncRNAs is still incomplete, especially for those lncRNAs with a very specific expression pattern in defined biological processes, we inspected our sequencing data for indications of unknown switch lncRNAs. We found six genomic loci with a high number of transcriptome reads specifically in DPs cells (Supplementary File 3). Sequences of four of these loci (in chromosomes 9, 10, 12 and 18) partly overlapped some ESTs while for the remaining two (in chromosomes 1 and 7) we could not find any reference except for bioinformatic predictions (Lv *et al*, 2013) that appeared while our work was being revised for publication. Thus, we aimed to validate the two potentially new lncRNAs by RT–PCR and Sanger sequencing from lysates of the E13.5 cortex. The first lncRNA was intergenic and we named it as Cortical On-Switch lncRNA 1 (Cosl1) (Supplementary Figure S4). To our surprise, the second lncRNA was only 3 Kb from Btg2, the marker used in this study to identify DPs cells (Figure 5A; bottom, right) and therefore classifies as a new genic, Btg2-Antisense, switch lncRNA (Btg2-AS1). RT–PCR validation revealed two alternative partial lncRNA transcripts for Btg2-AS1 that differed in a 3′ splice site usage (Figure 5A; bottom, right).

Next, we found that six of the nine already known switch lncRNAs were either overlapping (genic) or in close proxi-mity to the promoter of another protein-coding gene (Figure 5A), which was also the case for the newly identified
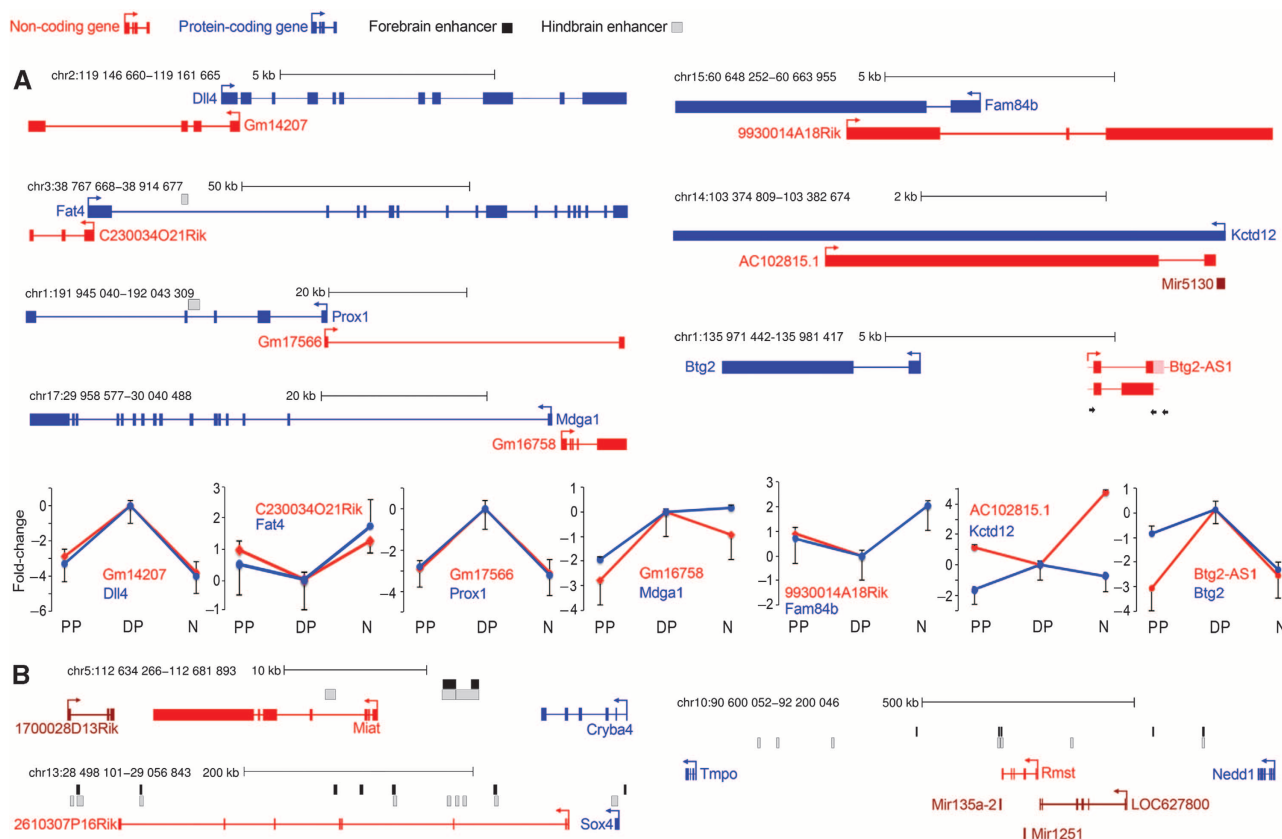

Figure 5 Switch lncRNAs loci of intergenic (**A**) and genic (**B**) lncRNAs (red), protein-coding genes (blue) and miRNAs (dark red) transcribed in sense or antisense (arrows) are depicted together with their pattern of expression in PPs, DPs and neurons (N) (**B**, bottom; log$_2$-fold change relative to DPs). Intronic and exonic regions, genomic mapping and p300 fore-/hind-brain enhancers (black/grey boxes) are shown. Note two Btg2-AS1 variants (**A**, bottom, right) differing in a 3′ splice site usage (red/pink = validated by sequencing/as PCR band).

Btg2-AS1. Surprisingly, protein-coding genes overlapping switch lncRNAs were revealed not only to be switch genes themself but also to play critical roles in neurogenesis. For example, the lncRNAs Gm14207 and Gm17566 overlapped in antisense direction the promoters of the Notch ligand Dll4 and of the neuronal homeobox transcription factor Prox1, respectively. Likewise, the two off-switch lncRNAs C230034O21Rik and 9930014A18Rik overlapped Fam84B and Fat4, each being implicated in fundamental processes controlling neurogenesis, namely, cell-cycle progression and extracellular matrix formation, respectively (Camps et al, 2009; Saburi et al, 2012). Moreover, Fat4 was recently shown to be involved in recessive human syndromes including periventricular neuronal heterotopia (Cappello et al, 2013). Finally, the on-switch lncRNA Gm16758 was adjacent (2 Kb) to Mdga1, whose knockout leads to abnormal neuronal migration (Ishikawa et al, 2011), while 2610307P16Rik was at 67 Kb from Sox4, a REST-regulated gene with important functions during neuronal maturation (Bergsland et al, 2006).

Interestingly, the expression level of all genic switch-lncRNAs and their protein-coding genes showed a strong positive correlation within each population of PPs, DPs and neurons (Figure 5A, bottom). The only exception was the negative correlation of AC102815.1 that completely overlaps the coding gene Kctd12, a K^+ channel subunit involved in neuronal differentiation (Aprea and Calegari, 2012). Similarly, the expression of the lncRNAs Gm16758 and 2610307P16Rik peaked in DPs and decreased in neurons and partly correlated with the expression of the neighbouring Mdga1 and Sox4, both peaking in DPs and remaining high in neurons (Figure 5A and B; Supplementary File 1). The finding that nearly all genes overlapping or adjacent to a switch lncRNA share with it a nearly identical expression pattern and are implicated in fundamental processes regulating neurogenesis makes it conceivable that switch lncRNAs may influence cell fate determination by controlling the expression of nearby, protein-coding genes.

Further supporting an important role of lncRNAs in brain development, five of the nine switch lncRNA genomic loci contained regions bound by the mouse fore- or mid-brain transcriptional enhancer component p300 (Figure 5), potentially highlighting regulatory elements driving the expression pattern of the lncRNA and/or its nearby gene. In the case of a clearly intergenic lncRNA, Miat, these enhancers are very likely to regulate the lncRNA itself rather than its neighbouring genes (Cryba4 and 1700028D13Rik) since the latter were not expressed at any significant level in our transcriptome data (Supplementary File 1).

Manipulation of switch genes affects corticogenesis: a powerful approach for the identification of novel players in brain development

To assess the proportion of uncharacterized and novel switch genes triggering an effect on cortical development, and thus the power of our approach, we acutely manipulated their expression during cortical development by in utero electroporation. To this aim, we selected four switch transcripts with the sole criterion that none had any reported function in cell fate specification of cortical progenitors. For protein-coding transcripts, we included (i) the uncharacterized on-switch 9630028B13Rik, a highly conserved transcript predicted to

encode for a transmembrane protein of 160 aa and (ii) the off-switch Schip1 whose only described role in the CNS is as a late component of the nodes of Ranvier of mature neurons (Martin et al, 2008). Similarly, lncRNAs included (iii) the uncharacterized Gm17566, a genic on-switch lncRNA antisense to Prox1 (Figure 5A) and (iv) Miat, an intergenic on-switch of unknown function that was until now exclusively described in the retina where it is thought to be primarily expressed in neurons (Blackshaw et al, 2004; Sone et al, 2007).

The cDNA of each candidate was cloned into expression vectors encoding also for a nuclear fluorescent reporter and in utero electroporation performed to target both PPs and DPs of the E13.5 lateral cortex. Brains were collected 2 days later and distribution of electroporated cells and their progeny quantified as a direct and fast assessment of neurogenesis, neuronal migration and/or survival. Overexpression of 9630028B13Rik, Schip1 and Gm17566 all led to an evident change in the distribution of electroporated cells and their progeny (Figure 6) with a most substantial change in the CP that was almost completely deficient in targeted cells (a reduction relative to controls by 84.2 ± 13.4, 74.8 ± 9.9 and $84.8 \pm 13.4\%$, respectively; $P < 0.01$). In the case of our fourth candidate, Miat, the change in the proportion of cells in the CP was less striking (Figure 6, right) ($38.0 \pm 14.1\%$); which did not exclude more subtle effects on progenitor subtypes.

Elucidating the molecular mechanism and cellular function of all these switch genes would extend beyond the purpose of the current study. Remains the fact that overexpression of four out of four randomly selected switch genes with no reported function in corticogenesis (Schip1 and Miat) or in any other context whatsoever (9630028B13Rik and Gm17566) was sufficient to trigger strong phenotypes in brain development. As such, it seems likely that a very large proportion of switch genes play hitherto unsuspected roles in neurogenesis providing a powerful new resource to the field. With regard to the modest effect of Miat over-expression, it was still possible that more subtle phenotypes would emerge following a more systematic quantification. This was also important because this lncRNA is not adjacent to any other gene implicated in neurogenesis (Figure 5B) and, thus, its manipulation may directly reveal its function rather than causing secondary effects through its neighbouring cell fate determinant, as it might be the case for the genic lncRNA Gm17566. Therefore, we decided to further investigate potential effects of Miat in neurogenic commitment.

Miat controls the differentiation of neural progenitors, the survival of newborn neurons and the splicing of Wnt7b

Miat was first described in the retina where it was suggested to be primarily expressed in neurons and localize in nuclear subdomains that do not overlap with any other nuclear body described to date including interchromatin granules, para-speckles, nucleoli and PLM or Cajal bodies (Blackshaw et al, 2004; Sone et al, 2007). Two conflicting studies proposed that Miat promotes the differentiation of embryonic pluripotent stem cells (Sheik Mohamed et al, 2010) or, conversely, inhibits the differentiation of retina precursors (Rapicavoli et al, 2010) but these effects were not assessed in any other cell type. Recent evidence using synthetic oligonucleotides on cell extracts suggested that Miat plays a role in splicing by
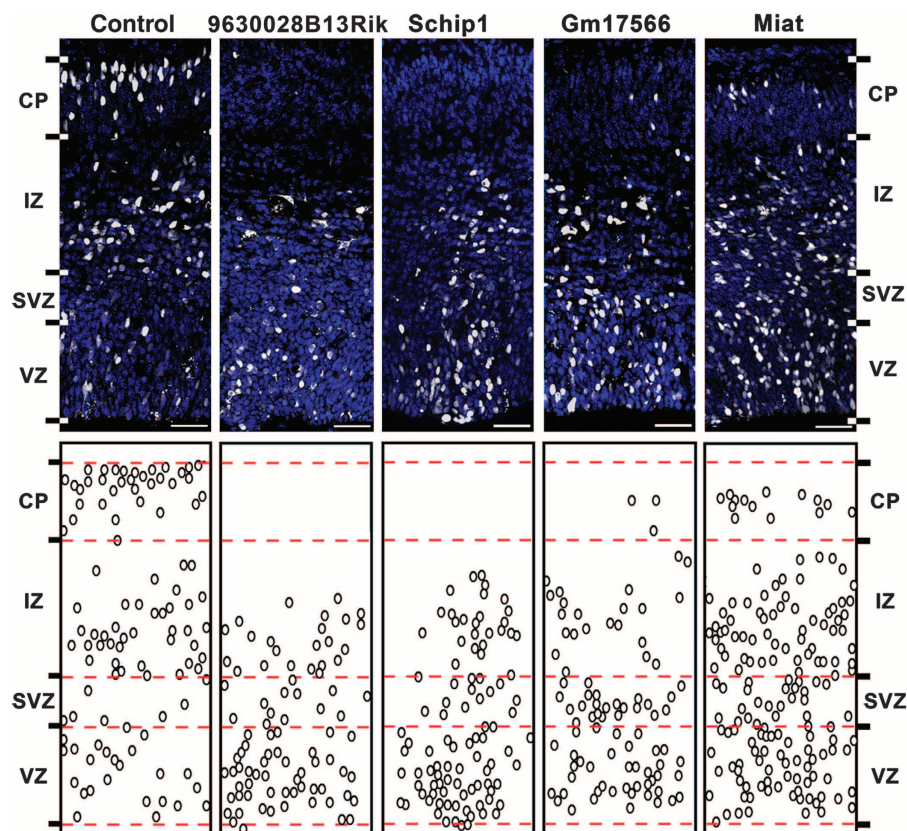
Figure 6 Switch genes influence corticogenesis Fluorescence pictures of cryosections through the E15.5 mouse cortex (top) and related diagrams of targeted cells (bottom) after 2 days overexpression of control or switch genes (as indicated). Targeted cells were identified by a coexpressed nuclear-fluorescent reporter (white; top) and outlined in the respective diagrams (circles; bottom). DAPI was used to label all nuclei (blue; top). Cortical layers are indicated (red lines; bottom). VZ = ventricular zone; SVZ = subventricular zone; IZ = intermediate zone; CP = cortical plate. All switch transcripts induced significant ($P < 0.05$; $n = 3$; Gm17566 = 2) changes in the proportion of electroporated cells in the CP relative to control. Scale bar = 40 μm.

competing with pre-mRNAs for binding to splicing factors such as splicing factor 1, quaking homologue and others (Ip and Nakagawa, 2011; Barry et al, 2013). Consistently, Miat overexpression in iPS-derived neurons was found to induce aberrant splicing (Barry et al, 2013) although this effect remains to be validated in vivo. In our study, surprisingly, we found that Miat was the single most abundant transcript of the whole transcriptome of DPs (Supplementary File 1).

We further studied the role of Miat in corticogenesis by in utero electroporation at E13.5. Quantifications 2 days after Miat overexpression revealed a 30% increase in the proportion of cells in the VZ relative to controls (Figure 7A) that alone accounted for the decrease in neurons in the CP (Figures 6 and 7A). Importantly, the increased abundance of targeted cells in the VZ appeared to be solely due to an increased generation of BPs since the proportion of cells immunoreactive for the BPs marker Tbr2 more than doubled in the VZ after Miat electroporation (Figure 7B).

Two possibilities may explain why a higher proportion of BPs after Miat overexpression did not correlate with increased neurogenesis despite the fact that in physiological conditions the majority of Tbr2 +, BPs are neurogenic; namely that sustained overexpression of Miat promoted neuronal death and/or induced the supernumerary BPs to remain PPs rather than switching their fate to become DPs. We investigated the former possibility by quantifying the number of caspase-3 + cells 1 day after electroporation and

found a 3-fold increase in the IZ/CP upon Miat overexpression (Figure 7C). Importantly, no increase in caspase-3 immunoreactivity was found in the VZ and the SVZ (Figure 7C), indicating that the effect of Miat on survival is specifically restricted to neurons. To also address the latter possibility that upon Miat overexpression a higher proportion of BPs remains PPs rather than switching their fate to become DPs, we electroporated E13.5 Btg3$^{\text{RFP}}$ mouse embryos using a nuclear-localized GFP as a reporter of targeted cells and found a 30% decrease in Btg2$^{\text{RFP}}$+, DPs in the VZ at E14.5 (Figure 7D) implying that a higher proportion of newborn BPs remained PPs.

These experiments indicated the importance of Miat in neurogenic commitment and neuronal survival. Yet, we were surprised to observe that overexpression of an on-switch gene, that is, a gene physiologically overexpressed in DPs, decreased, rather than increased, the proportion of DPs. This seemingly counterintuitive effect might be explained by the peculiar localization of Miat in, hitherto uncharacterized, nuclear subdomains (Sone et al, 2007) in which its physiological function might be perturbed after overexpression of ectopic Miat as already shown by Miat-IRES-GFP constructs (Rapicavoli et al, 2010). We reasoned that, if this was to be the case, Miat RNAi should trigger effects that are similar to its overexpression. Thus, we performed electroporations with a previously characterized Miat shRNA vector (Rapicavoli et al, 2010) and assessed the distribution of
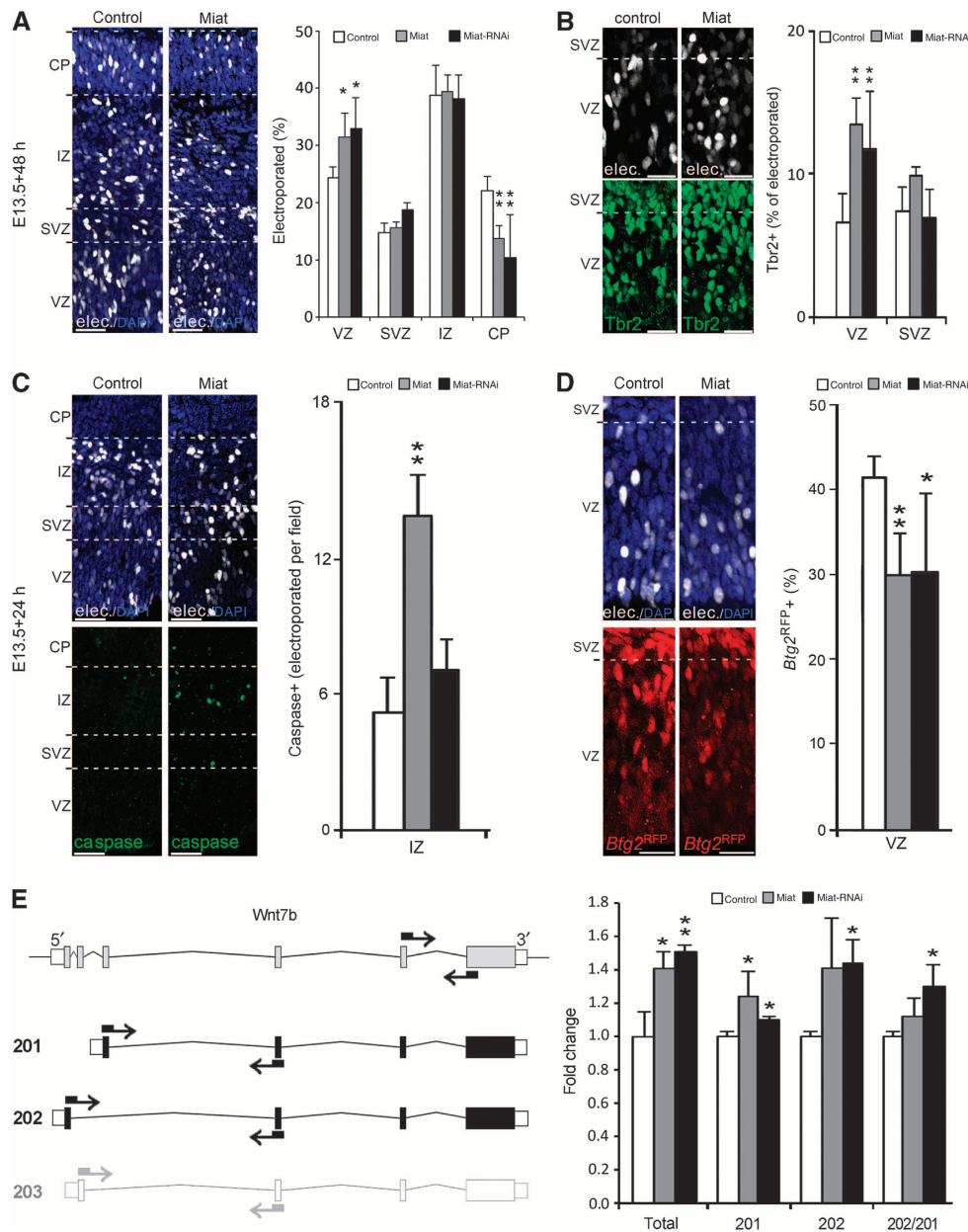
Figure 7 Miat influences corticogenesis and splicing. (**A–D**) Fluorescence pictures of cryosections through the mouse cortex (left) and quantifications of cell types (right) after *in utero* electroporation at E13.5 with control (white), Miat overexpression (grey) or RNAi (black) vectors. Embryos were dissected 48 (**A**, **B**) or 24 (**C**, **D**) h later to quantify distribution of cells (**A**) (dashed lines), Tbr2 immunoreactivity (**B**), number of caspase-3 + cells per optical field (i.e., independently if electroporated or not) (**C**) or $Btg2^{RFP}$ endogeneous fluorescence (**D**). (**E**) Maps (left) of Wnt7b Ensembl annotated variants (201, 202 and 203) and quantification of expression (right) upon Miat manipulation. Primers used (arrows) are indicated (not in scale). Wnt7b-203 (left, light grey) was not significantly detected neither by sequencing nor by qRT-PCR. Graph represents fold changes (left to right) of total, Wnt7b variants and their proportion. VZ = ventricular zone; SVZ = subventricular zone; IZ = intermediate zone; CP = cortical plate. Error bars = s.d.; $n \geqslant 3$ embryos; *$P < 0.05$; **$P < 0.005$. Scale bar = 40 (**A**, **C**) or 20 (**B**, **D**) μm.

electroporated cells and their progeny throughout cortical layers and proportion of Tbr2, caspase-3 or $Btg2^{RFP}$ + cells within this population, as described above. We found that in nearly all cases phenotypes after Miat overexpression or RNAi were virtually identical (Figure 7A–D) supporting the notion that either manipulation leads to a Miat loss of function as previously reported (Chen and Carmichael, 2010; Rapicavoli *et al*, 2010).

Another puzzling observation was the pleiotropic, and partly counterintuitive, effects induced by the manipulation of this single transcript resulting in a decrease in $Btg2^{RFP}$ +

cells but an increase in Tbr2 + progenitors and neuronal cell death without a change in progenitor survival.

Since the only proposed function of Miat is in regulating splicing (Ip and Nakagawa, 2011; Barry *et al*, 2013), it seems to be plausible that its pleiotropic effects are due to overall effects on splicing of different targets in different cell types. Since additional sequencing data gave us hints about exon usage in PPs, DPs and neurons under physiological conditions (manuscript in preparation), it was intuitive to investigate whether the splicing of known cell fate determinants was being altered by our manipulations. To

this aim, we performed *in utero* electroporation with Miat overexpression or RNAi plasmids together with vectors encoding a fluorescent reporter (as described above) and FAC-sorted targeted cells 48 h later at E15.5. RNA of sorted cells was extracted and qRT-PCR performed using primers that recognize specific splice variants of candidate genes (Figure 7E; Supplementary Table S1). These included one member each of the Wnt family of morphogens (Wnt7b) and Rho family of GTPases (Cdc42), both of which were shown to control the orientation of cell division in the neuroepithelium and other tissues (Cappello *et al*, 2006; Yu *et al*, 2009). Miat overexpression or RNAi triggered, once again, similar phenotypes with an increase in the total levels of Wnt7b as well as a change in the proportion of its splice variant Wnt7b-201 relative to 202 (Figure 7E). Suggesting a certain degree of specificity, no significant change was observed in the splicing of Cdc42-001 relative to 002 (not shown). To our knowledge, this represents the first validation of the role of Miat in splicing *in vivo*.

All together, our data show that Miat plays multiple and complex roles in different cell types at the level of (i) generation of BPs from APs, (ii) their switch from proliferative to neurogenic division and (iii) survival of newborn neurons. The pleiotropic effect of Miat is likely due to the aberrant splicing of several factors implicated in brain development, of which we identified one.

Discussion

The Btg2^{RFP} line provides a new tool to identify cell types in tissues

Several aspects of our study are novel and worth discussing. First, the generation of the *Btg2*^{RFP}/*Tubb3*^{GFP} line has allowed us to isolate the three coexisting pools of proliferating neural stem cells, differentiating progenitors and newborn neurons coexisting in space and time during brain development. Similar in purpose to recent reports during adult neurogenesis (Beckervordersandforth *et al*, 2010; Ramos *et al*, 2013), our study during embryonic development overcomes the obstacles inherent in the unambiguous identification of cell types in complex tissues due to the inheritance of a reporter protein from mother to daughter cells. More generally, Btg2 may serve as a marker not only during embryonic (Haubensak *et al*, 2004) and adult (Attardo *et al*, 2010) neurogenesis but also in a number of other tissues, stem-cell contexts or cancer, which is likely due to its role as an antiproliferative gene (Tirone, 2001; Lim, 2006; Terra *et al*, 2008). Considering the large number of mouse lines that use GFP as a reporter, our new RFP line may serve as an additional resource to identify cell types after crossing with any other appropriate GFP line. In the context of corticogenesis, it is relatively easy to imagine how our approach might be further optimized and extended for the comparison of proliferating versus differentiating APs or BPs (e.g., Prom1 + or Tbr2 + cells), neurogenic versus gliogenic progenitors (e.g., by selecting cells during neuro- versus gliogenesis) and/or mature versus immature neurons (e.g., by sorting cells with different levels of GFP and RFP fluorescence). It is also likely that future studies based on the Btg2^{RFP} line will extend our work by focussing on the many other aspects of gene expression not addressed here including the identification of non poly-A RNAs, short/miRNAs, circular RNAs, transcriptome-wide alternative splicing, promoter and poly-A site usage, transcription factor binding or chromatin modifications to better understand the many complex regulatory mechanisms of gene expression underlying the development of the mammalian brain.

Switch genes characterize the signature of neurogenic commitment: a powerful resource to identify novel regulators of corticogenesis

Beyond previous reports that compared different developmental stages, species, cortical layers or retrospectively deduced cell identity of individual cells (Han *et al*, 2009; Ayoub *et al*, 2011; Fietz *et al*, 2012; Yao *et al*, 2012), our study provides the community with a comprehensive and highly quantitative gene-expression profile of PPs, DPs and neurons coexisting in time and space. We believe that this is particularly important to discriminate between genes controlling the fate of a certain cell population from those being influenced by the many general systemic changes occurring during development, across species, tissue domains or even due to stochastic fluctuations within individual cells. Testifying to the quality of our preparations, our libraries revealed to be highly complex and reproducible with over 50 well-known markers and functional genes being enriched in the expected cell pool.

As perhaps the most important conceptual novelty of our approach, this allowed us to distinguish genes that, while being up-/down-regulated in DPs relative to PPs, were very unlikely to be implicated in the switch from PPs to DPs. Notable examples among this group were neuronal markers themself, including several cytoskeletal genes and transcripts implicated in axo-/dendro-genesis, synaptic transmission and regulated exocytosis, all of which play no role in DPs but that nonetheless start to be up-regulated in this population. This observation could be explained by considering a contamination of neuronal transcripts in DPs preparations and to a certain degree this could have contributed to the observed effect. However, it should be noted that expression of neuronal markers, such as Tubb3, has already been observed in mitotic DPs by time-lapse microscopy of intact cortical tissue (Attardo *et al*, 2008), thus, indicating that DPs are already '*primed*' during their last neurogenic division to express a number of neuronal genes even if these will become functionally relevant only in neuronal daughter cells. These observations led us to focus on the small pool of transcripts that specifically distinguished DPs; the switch genes.

Switch genes included essentially all known markers of DPs and genes that were only recently associated with cortical development such as, to mention a few, (i) Cbfa2t2, a downstream gene and feedback regulator of Neurog2 essential for neurogenesis (Aaker *et al*, 2009), (ii) Chd7, a Sox2 cofactor, chromatin remodelling ATPase involved in CHARGE syndrome (Engelen *et al*, 2011), (iii) Ezh2, a histone methyl transferase regulating self-renewal of cortical progenitors (Pereira *et al*, 2010) and (iv) Ncor2, a histone H3 trimethyl demethylase regulating the transition from stem cells to neurogenesis (Jepsen *et al*, 2007). Notably, most of these recently identified switch genes are implicated in chromatin remodelling confirming the emerging role of epigenetic marks in neurogenesis (Hirabayashi and Gotoh, 2010; Hu *et al*, 2012).

In focussing our attention on switch genes, we were surprised to observe two categories of transcripts that appeared to be at the opposite extremes of the spectrum of

previous investigations. On the one side, we found well-known and extensively characterized genes including Tbr2, Btg2, Neurog2, Dll1 and Wnt. On the other side, we found a similarly high number of genes that are completely uncharacterized such that many are still assigned only automatic annotation numbers instead of a gene name. We speculated that the existence of these two extreme categories might reflect a bias of the field in that genes already recognized to be 'important' in stem-cell commitment remain in the focus of research while uncharacterized genes remain uncharacterized. In this work, we wanted to challenge this conventional attitude and attempted to validate our approach by *in utero* electroporation of two completely uncharacterized switch transcripts: the protein-coding 9630028B13Rik and the lncRNA Gm17566. No other criterion has been used to select these two except that no study has been reported for neither of them in any context. Yet, their manipulation led to evident and immediate effects on brain development in both cases. Considering that a similar result was obtained with two additional transcripts studied in tissues other than the cortex, and in neurons rather than in stem cells (Schip1 and Miat), we conclude that a remarkably high proportion of switch genes identified in our study are prime candidates for hitherto unsuspected roles in brain development and, perhaps, in the context of somatic stem cells differentiation in other tissues.

Our study provides the community with the comprehensive list of these under-studied transcripts whose highly specific, robust and transient expression signature was unknown before.

LncRNAs as novel players in corticogenesis

A rapidly growing literature is highlighting the relevance of lncRNAs in a number of physiological processes (Mercer *et al*, 2009; Nagano and Fraser, 2011) but overall their characterization is still very incomplete to the point that, to our surprise, only one study has addressed their role by direct manipulation during corticogenesis (Onoguchi *et al*, 2012). In our study, we identified several known, as well as novel, genic and intergenic switch lncRNAs, including two that we named as Cosl1 and Btg2-AS1. Noteworthy, essentially all genic, switch lncRNAs overlapped switch protein-coding genes known to play major roles in corticogenesis (e.g., Dll1 and Prox1) or to be implicated in developmental brain syndromes (e.g., Fat4) (Cappello *et al*, 2013). To date, it is unclear whether a genic lncRNA controls the expression of an overlapping protein-coding gene or, alternatively, the former is a by-product of the latter but recent results show that at least in some case the former may occur even when the lncRNA is being ectopically expressed from another locus (Berghoff *et al*, 2013). Consistently, we found that episomal expression of a genic lncRNA, Gm17566, had a direct effect on neurogenesis. As such, the correlation in expression levels of genic lncRNAs and their protein-coding genes strongly suggests that lncRNAs influence stem-cell dynamics by controlling the expression of nearby cell fate determinants.

Finally, we investigated the role of one intergenic lncRNA in corticogenesis and found that Miat is involved in cell fate change of progenitors and survival of newborn neurons. Our results suggest that Miat controls proliferation versus differentiation by regulating splicing of cell fate determinants. In this study, we identified Wnt7b as one target but we are confident that more will emerge in future studies addressing

the role of lncRNA-mediated alternative splicing in the control of stemness.

Materials and methods

Animals

Mice were kept under standard housing conditions and experiments carried out according to local regulations. The $Btg2^{RFP}$ line was generated by Red/ET recombination (Genebridges) of a nuclear-localized RFP at base 102 from the start codon of Btg2 encoded within a bacterial artificial chromosome (BAC bmQ284g14, Sanger Institute). The recombined BAC (2 μg/ml) was injected in male pronuclei of fertilized oocytes (129 genetic background) and chimeras backcrossed into a C57Bl/6 background. Mice were kept as heterozygous and time-mated $Btg2^{RFP}$, eventually crossed with $Btg2^{GFP}$ or $Tubb3^{GFP}$ mice, defined as E0.5 the morning of vaginal plug. Genotypes were assessed by endogeneous RFP and/or GFP fluorescence and/or by PCR using RFP/GFP primers.

In utero electroporation

Plasmids were generated by inserting mCherry (Artegiani *et al*, 2011) in the MCS2 of pBI-CMV1 (Clontech) followed by cloning into MCS2 of cDNAs obtained either from FANTOM, RIKEN (9630028B13Rik: clone C230029O13) or from RT-PCR of E13.5 brains (Gm17566 and Schip1) (Supplementary data; Supplementary Table S1). Miat vectors were kindly provided by Dr Seth Blackshaw (Rapicavoli *et al*, 2010) and coelectroporated with pDSV-mRFP^{nls} (Lange *et al*, 2009). Pregnant mice were isofluorane anaesthetized at E13.5 and 1–3 ml of PBS containing 1–3 mg/ml of plasmids injected into the lumen of the telencephalon followed by 6 pulses of 30 V, 50 ms each at 1 s interval delivered through platinum electrodes using a BTX-830 electroporator (Genetronics) as previously described (Artegiani *et al*, 2012).

Histochemistry

Brains were fixed in 4% paraformaldehyde in 0.1 M phosphate buffer (pH 7.4) (PFA) at 4°C for 12 h, cryoprotected in 30% sucrose and cyosections (10 μm) assessed for endogenous fluorescence or immunohistochemistry as previously described (Lange *et al*, 2009) (Supplementary data). For *in situ* hybridization, dioxygenin-labelled (Roche) cRNA antisense and sense probes corresponding to the RFP sequence were used on 10 μm sections according to standard protocols (Supplementary data). Sections were analysed with a conventional Axioscope or confocal LSM510 Axiovert 200M (Carl Zeiss, Oberkochen, Germany) microscope and images acquired with a Zeiss LSM 4.2 camera (Carl Zeiss) and processed with ImageJ 1.33 (www.imagej.nih.gov) or Photoshop CS3 (Adobe, San Jose, CA, USA).

Sorting, RNA isolation and sequencing

E14.5 $Btg2^{GFP}$/$Tubb3^{GFP}$ cortices or E15.5 wild-type electroporated brains were dissociated using the papain-based neural dissociation kit (Milteney Biotec) after removal of meninges, ganglionic eminences and, eventually, the non-electroporated portion of the cortex. FACS was performed at 4°C in the 4-way purity mode with a flow rate of 20 μl/min using side and forward scatter light to eliminate debris and aggregates and gating established for green (488 nm) and red (561 nm) fluorescence. Prior to sorting, an aliquot of cells was stained by propidium iodide (0.3 μg/ml) to assess lethality (usually <1%) with a second aliquot re-sorted to determine purity (usually >99%). For sequencing, about 1×10^6 sorted cells from >3 embryos from different litters were immediately lysed using the μMACS™ mRNA Isolation Kit and lysates cleaned on LysateClear Colums (Miltenyi) resulting in *ca.* 1 μg of poly-A RNAs with an integrity number of >9.2. Libraries were prepared according to standard procedures and kits used according to the manufacturers' instructions including oligo(dT) for transcripts selection, first-strand cDNA synthesis by random primers, second-strand synthesis, end repair, adaptor ligation, dUTP cleavage and enrichment with indexed primers (detailed description in Supplementary data). After XP beads purification, libraries were quantified using the Qubit dsDNA HS Assay Kit (Invitrogen). For Illumina flowcell production, samples were pooled in three lanes for 75 bp single read sequencing on Illumina HiSeq 2000 resulting in *ca.*

30–40 million reads per sample. Sequencing raw data were deposited in GEO GSE51606.

Statistical and bioinformatic analyses

Characterization of the $Btg2^{RFP}$ line, bioinformatic assessment of RNA libraries and functional manipulations by *in utero* electroporation were performed by pulling together >3 embryos per sample using >3 litters as independent biological replicates. Phenotypes upon electroporation were assessed by two-tailed, unpaired *t*-test assuming normal distribution with $P < 0.05$ being considered as significant. For bioinformatic analyses of transcriptomes, a splice junction library of 120 nucleotides length was created with RSEQTools based on known exon–exon junctions according to the Ensembl Genes v. 61 annotation. Reads alignment to the mm9 transcriptome was performed by pBWA (http://pbwa.sourceforge.net) resulting in a mappability range for uniquely mapped reads of 70–80%. A table of counts per gene was created based on the overlap of uniquely mapped reads using BEDtools (v. 2.11). The DESeqR package (v.1.8.1) was used for normalization of raw counts and further testing of differential expression following negative binomial test. Benjamini–Hochberg (FDR method) was used for adjusting *P*-values with <0.05 being considered as significant. Genes with normalized counts = 0 (in any sample) were removed from the entire data set before analysis and mean counts from replicates used for fold change (FC) calculations. Switch genes were identified using the following criteria: $\log_2(DPs/PPs) \geqslant 0.58$ (FC 1.5) and $\log_2(DPs/N) \geqslant 0.58$ (on-switches) or $\log_2(DPs/PPs) \geqslant -0.58$ and $\log_2(DPs/N) \geqslant -0.58$ (off-switches). WGCNA was performed using the R package (WGCNA 1.20). Average linkage hierarchical clustering was performed from Topological Overlap-based dissimilarity matrix and modules were identified using cutreeDynamic function (Supplementary data). Ensembl annotation and RNAcode were used to distinguish between coding and non-coding genes and to assess the signature of coding sequence conservation. The UCSC genome browser was used to explore the genomic neighbourhood of lncRNAs and GO terms downloaded from Ensembl Version 61 after excluding the generic terms 'molecular function' and 'biological process'.

Validation of lncRNAs and splicing

Novel lncRNAs were identified by excluding reads associated or in close proximity (<2000 bp) to annotated transcripts and switch behaviour assessed with stronger selection criteria than for coding genes (FC>2; FDR<0.025). Validation was performed by RT–PCR sequencing using primers designed on Cufflinks prediction (Supplementary data; Supplementary Table S1). Alternative splicing of Miat targets was investigated by qRT-PCR as described in Supplementary data; Supplementary Table S1.

Supplementary data

Supplementary data are available at *The EMBO Journal* Online (http://www.embojournal.org).

Conflict of interest

The authors declare that they have no conflict of interest.

References

Aaker JD, Patineau AL, Yang HJ, Ewart DT, Gong W, Li T, Nakagawa Y, McLoon SC, Koyano-Nakagawa N (2009) Feedback regulation of NEUROG2 activity by MTGR1 is required for progression of neurogenesis. *Mol Cell Neurosci* **42:** 267–277

Amaral PP, Clark MB, Gascoigne DK, Dinger ME, Mattick JS (2011) lncRNAdb: a reference database for long noncoding RNAs. *Nucleic Acids Res* **39:** D146–D151

Aprea J, Calegari F (2012) Bioelectric state and cell cycle control of Mammalian neural stem cells. *Stem Cells Int* **2012:** 816049

Arai Y, Funatsu N, Numayama-Tsuruta K, Nomura T, Nakamura S, Osumi N (2005) Role of Fabp7, a downstream gene of Pax6, in the maintenance of neuroepithelial cells during early embryonic development of the rat cortex. *J Neurosci* **25:** 9752–9761

Arai Y, Pulvers JN, Haffner C, Schilling B, Nusslein I, Calegari F, Huttner WB (2011) Neural stem and progenitor cells shorten S-phase on commitment to neuron production. *Nat Commun* **2:** 154

Artegiani B, Lange C, Calegari F (2012) Expansion of embryonic and adult neural stem cells by in utero electroporation or viral stereotaxic injection. *J Vis Exp* **68:** e4093

Artegiani B, Lindemann D, Calegari F (2011) Overexpression of cdk4 and cyclinD1 triggers greater expansion of neural stem cells in the adult mouse brain. *J Exp Med* **208:** 937–948

Attardo A, Calegari F, Haubensak W, Wilsch-Brauninger M, Huttner WB (2008) Live imaging at the onset of cortical neurogenesis reveals differential appearance of the neuronal phenotype in apical versus basal progenitor progeny. *PLoS One* **3:** e2388

Attardo A, Fabel K, Krebs J, Haubensak W, Huttner WB, Kempermann G (2010) Tis21 expression marks not only populations of neurogenic precursor cells but also new postmitotic neurons in adult hippocampal neurogenesis. *Cereb Cortex* **20:** 304–314

Ayoub AE, Oh S, Xie Y, Leng J, Cotney J, Dominguez MH, Noonan JP, Rakic P (2011) Transcriptional programs in transient embryonic zones of the cerebral cortex defined by high-resolution mRNA sequencing. *Proc Natl Acad Sci USA* **108:** 14950–14955

Barry G, Briggs JA, Vanichkina DP, Poth EM, Beveridge NJ, Ratnu VS, Nayler SP, Nones K, Hu J, Bredy TW, Nakagawa S, Rigo F, Taft RJ, Cairns MJ, Blackshaw S, Wolvetang EJ, Mattick JS (2013) The long non-coding RNA Gomafu is acutely regulated in response to neuronal activation and involved in schizophrenia-associated alternative splicing. *Mol Psychiatry* (advance online publication 30 April 2013; doi:10.1038/mp.2013.45)

Beckervordersandforth R, Tripathi P, Ninkovic J, Bayam E, Lepier A, Stempfhuber B, Kirchhoff F, Hirrlinger J, Haslinger A, Lie DC, Beckers J, Yoder B, Irmler M, Götz M (2010) In vivo fate mapping and expression analysis reveals molecular hallmarks of prospectively isolated adult neural stem cells. *Cell Stem Cell* **7:** 744–758

Berghoff EG, Clark MF, Chen S, Cajigas I, Leib DE, Kohtz JD (2013) Evf2 (Dlx6as) lncRNA regulates ultraconserved enhancer methylation and the differential transcriptional control of adjacent genes. *Development* **140:** 4407–4416

Bergsland M, Werme M, Malewicz M, Perlmann T, Muhr J (2006) The establishment of neuronal properties is controlled by Sox4 and Sox11. *Genes Dev* **20:** 3475–3486

Blackshaw S, Harpavat S, Trimarchi J, Cai L, Huang H, Kuo WP, Weber G, Lee K, Fraioli RE, Cho SH, Yung R, Asch E, Ohno-Machado L, Wong WH, Cepko CL (2004) Genomic analysis of mouse retinal development. *PLoS Biol* **2:** E247

Calegari F, Haubensak W, Haffner C, Huttner WB (2005) Selective lengthening of the cell cycle in the neurogenic subpopulation of neural progenitor cells during mouse brain development. *J Neurosci* **25:** 6533–6538

Camps J, Nguyen QT, Padilla-Nash HM, Knutsen T, McNeil NE, Wangsa D, Hummon AB, Grade M, Ried T, Difilippantonio MJ (2009) Integrative genomics reveals mechanisms of copy number alterations responsible for transcriptional deregulation in colorectal cancer. *Genes, Chromosomes Cancer* **48:** 1002–1017

Cappello S, Attardo A, Wu X, Iwasato T, Itohara S, Wilsch-Brauninger M, Eilken HM, Rieger MA, Schroeder TT,

Huttner WB, Brakebusch C, Götz M (2006) The Rho-GTPase cdc42 regulates neural progenitor fate at the apical surface. *Nat Neurosci* **9:** 1099–1107

Cappello S, Gray MJ, Badouel C, Lange S, Einsiedler M, Srour M, Chitayat D, Hamdan FF, Jenkins ZA, Morgan T, Preitner N, Uster T, Thomas J, Shannon P, Morrison V, Di Donato N, Van Maldergem L, Neuhann T, Newbury-Ecob R, Swinkells M *et al* (2013) Mutations in genes encoding the cadherin receptor-ligand pair DCHS1 and FAT4 disrupt cerebral cortical development. *Nat Genet* **45:** 1300–1308

Chen LL, Carmichael GG (2010) Decoding the function of nuclear long non-coding RNAs. *Curr Opin Cell Biol* **22:** 357–364

Derrien T, Johnson R, Bussotti G, Tanzer A, Djebali S, Tilgner H, Guernec G, Martin D, Merkel A, Knowles DG, Lagarde J, Veeravalli L, Ruan X, Ruan Y, Lassmann T, Carninci P, Brown JB, Lipovich L, Gonzalez JM, Thomas M *et al* (2012) The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. *Genome Res* **22:** 1775–1789

Diekmann H, Stuermer CA (2009) Zebrafish neurolin-a and -b, orthologs of ALCAM, are involved in retinal ganglion cell differentiation and retinal axon pathfinding. *J Comp Neurol* **513:** 38–50

Diez-Roux G, Banfi S, Sultan M, Geffers L, Anand S, Rozado D, Magen A, Canidio E, Pagani M, Peluso I, Lin-Marq N, Koch M, Bilio M, Cantiello I, Verde R, De Masi C, Bianchi SA, Cicchini J, Perroud E, Mehmeti S *et al* (2011) A high-resolution anatomical atlas of the transcriptome in the mouse embryo. *PLoS Biol* **9:** e1000582

Engelen E, Akinci U, Bryne JC, Hou J, Gontan C, Moen M, Szumska D, Kockx C, van Ijcken W, Dekkers DH, Demmers J, Rijkers EJ, Bhattacharya S, Philipsen S, Pevny LH, Grosveld FG, Rottier RJ, Lenhard B, Poot RA (2011) Sox2 cooperates with Chd7 to regulate genes that are mutated in human syndromes. *Nat Genet* **43:** 607–611

Faux C, Rakic S, Andrews W, Yanagawa Y, Obata K, Parnavelas JG (2010) Differential gene expression in migrating cortical inter-neurons during mouse forebrain development. *J Comp Neurol* **518:** 1232–1248

Fietz SA, Lachmann R, Brandl H, Kircher M, Samusik N, Schroder R, Lakshmanaperumal N, Henry I, Vogt J, Riehn A, Distler W, Nitsch R, Enard W, Paabo S, Huttner WB (2012) Transcriptomes of germinal zones of human and mouse fetal neocortex suggest a role of extracellular matrix in progenitor self-renewal. *Proc Natl Acad Sci USA* **109:** 11836–11841

Geschwind DH, Konopka G (2009) Neuroscience in the era of functional genomics and systems biology. *Nature* **461:** 908–915

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

Guillemot F (2007) Cell fate specification in the mammalian telencephalon. *Prog Neurobiol* **83:** 37–52

Han X, Wu X, Chung WY, Li T, Nekrutenko A, Altman NS, Chen G, Ma H (2009) Transcriptome of embryonic and neonatal mouse cortex by high-throughput RNA sequencing. *Proc Natl Acad Sci USA* **106:** 12741–12746

Hartl D, Irmler M, Romer I, Mader MT, Mao L, Zabel C, de Angelis MH, Beckers J, Klose J (2008) Transcriptome and proteome analysis of early embryonic mouse brain development. *Proteomics* **8:** 1257–1265

Haubensak W, Attardo A, Denk W, Huttner WB (2004) Neurons arise in the basal neuroepithelium of the early mammalian telencephalon: a major site of neurogenesis. *Proc Natl Acad Sci USA* **101:** 3196–3201

Hevner RF, Hodge RD, Daza RA, Englund C (2006) Transcription factors in glutamatergic neurogenesis: conserved programs in neocortex, cerebellum, and adult hippocampus. *Neurosci Res* **55:** 223–233

Hirabayashi Y, Gotoh Y (2010) Epigenetic control of neural precursor cell fate during development. *Nat Rev Neurosci* **11:** 377–388

Hu XL, Wang Y, Shen Q (2012) Epigenetic control on cell fate choice in neural stem cells. *Protein Cell* **3:** 278–290

Iacopetti P, Michelini M, Stuckmann I, Oback B, Aaku-Saraste E, Huttner WB (1999) Expression of the antiproliferative gene TIS21 at the onset of neurogenesis identifies single neuroepithelial cells that switch from proliferative to neuron-generating division. *Proc Natl Acad Sci USA* **96:** 4639–4644

Ip JY, Nakagawa S (2011) Long non-coding RNAs in nuclear bodies. *Dev Growth Differ* **54:** 44–54

Ishikawa T, Gotoh N, Murayama C, Abe T, Iwashita M, Matsuzaki F, Suzuki T, Yamamoto T (2011) IgSF molecule MDGA1 is involved in radial migration and positioning of a subset of cortical upper-layer neurons. *Dev Dyn* **240:** 96–107

Jabaudon D, Shnider SJ, Tischfield DJ, Galazo MJ, Macklis JD (2012) RORbeta induces barrel-like neuronal clusters in the developing neocortex. *Cereb Cortex* **22:** 996–1006

Jepsen K, Solum D, Zhou T, McEvilly RJ, Kim HJ, Glass CK, Hermanson O, Rosenfeld MG (2007) SMRT-mediated repression of an H3K27 demethylase in progression from neural stem cell to neuron. *Nature* **450:** 415–419

Kawaguchi A, Ikawa T, Kasukawa T, Ueda HR, Kurimoto K, Saitou M, Matsuzaki F (2008) Single-cell gene profiling defines differential progenitor subclasses in mammalian neurogenesis. *Development* **135:** 3113–3124

Kwon GS, Hadjantonakis AK (2007) Eomes::GFP-a tool for live imaging cells of the trophoblast, primitive streak, and telence-phalon in the mouse embryo. *Genesis* **45:** 208–217

Lamprianou S, Chatzopoulou E, Thomas JL, Bouyain S, Harroch S (2011) A complex between contactin-1 and the protein tyrosine phosphatase PTPRZ controls the development of oligodendrocyte precursor cells. *Proc Natl Acad Sci USA* **108:** 17498–17503

Lange C, Huttner WB, Calegari F (2009) Cdk4/cyclinD1 overexpression in neural stem cells shortens G1, delays neurogenesis, and promotes the generation and expansion of basal progenitors. *Cell Stem Cell* **5:** 320–331

Lein ES, Hawrylycz MJ, Ao N, Ayres M, Bensinger A, Bernard A, Boe AF, Boguski MS, Brockway KS, Byrnes EJ, Chen L, Chen L, Chen TM, Chin MC, Chong J, Crook BE, Czaplinska A, Dang CN, Datta S, Dee NR *et al* (2007) Genome-wide atlas of gene expression in the adult mouse brain. *Nature* **445:** 168–176

Lim IK (2006) TIS21 (/BTG2/PC3) as a link between ageing and cancer: cell cycle regulator and endogenous cell death molecule. *J Cancer Res Clin Oncol* **132:** 417–426

Ling KH, Hewitt CA, Beissbarth T, Hyde L, Banerjee K, Cheah PS, Cannon PZ, Hahn CN, Thomas PQ, Smyth GK, Tan SS, Thomas T, Scott HS (2009) Molecular networks involved in mouse cerebral corticogenesis and spatio-temporal regulation of Sox4 and Sox11 novel antisense transcripts revealed by transcriptome profiling. *Genome Biol* **10:** R104

Lv J, Liu H, Huang Z, Su J, He H, Xiu Y, Zhang Y, Wu Q (2013) Long non-coding RNA identification over mouse brain development by integrative modeling of chromatin and genomic features. *Nucleic Acids Res* (advance online publication 13 September 2013; doi:10.1093/nar/gkt818)

Martin PM, Carnaud M, Garcia del Cano G, Irondelle M, Irinopoulou T, Girault JA, Dargent B, Goutebroze L (2008) Schwannomin-interacting protein-1 isoform IQCJ-SCHIP-1 is a late component of nodes of Ranvier and axon initial segments. *J Neurosci* **28:** 6111–6117

Matsuki T, Hori G, Furuichi T (2005) Gene expression profiling during the embryonic development of mouse brain using an oligonucleotide-based microarray system. *Brain Res Mol Brain Res* **136:** 231–254

Meissner A (2010) Epigenetic modifications in pluripotent and differentiated cells. *Nat Biotechnol* **28:** 1079–1088

Mercer TR, Dinger ME, Mattick JS (2009) Long non-coding RNAs: insights into functions. *Nat Rev Genet* **10:** 155–159

Metzker ML (2010) Sequencing technologies - the next generation. *Nat Rev Genet* **11:** 31–46

Mikkelsen TS, Ku M, Jaffe DB, Issac B, Lieberman E, Giannoukos G, Alvarez P, Brockman W, Kim TK, Koche RP, Lee W, Mendenhall E, O'Donovan A, Presser A, Russ C, Xie X, Meissner A, Wernig M, Jaenisch R, Nusbaum C *et al* (2007) Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* **448:** 553–560

Nagano T, Fraser P (2011) No-nonsense functions for long non-coding RNAs. *Cell* **145:** 178–181

Ng SY, Bogu GK, Soh BS, Stanton LW (2013) The long noncoding RNA RMST interacts with SOX2 to regulate neurogenesis. *Mol Cell* **51:** 349–359

Okazaki Y, Furuno M, Kasukawa T, Adachi J, Bono H, Kondo S, Nikaido I, Osato N, Saito R, Suzuki H, Yamanaka I, Kiyosawa H, Yagi K, Tomaru Y, Hasegawa Y, Nogami A, Schonbach C, Gojobori T, Baldarelli R, Hill DP *et al* (2002) Analysis of the

mouse transcriptome based on functional annotation of 60,770 full-length cDNAs. *Nature* **420:** 563–573

Onoguchi M, Hirabayashi Y, Koseki H, Gotoh Y (2012) A noncoding RNA regulates the neurogenin1 gene locus during mouse neo-cortical development. *Proc Natl Acad Sci USA* **109:** 16939–16944

Pauli A, Rinn JL, Schier AF (2011) Non-coding RNAs as regulators of embryogenesis. *Nat Rev Genet* **12:** 136–149

Pereira JD, Sansom SN, Smith J, Dobenecker MW, Tarakhovsky A, Livesey FJ (2010) Ezh2, the histone methyltransferase of PRC2, regulates the balance between self-renewal and differentia-tion in the cerebral cortex. *Proc Natl Acad Sci USA* **107:** 15957–15962

Pinto L, Mader MT, Irmler M, Gentilini M, Santoni F, Drechsel D, Blum R, Stahl R, Bulfone A, Malatesta P, Beckers J, Götz M (2008) Prospective isolation of functionally distinct radial glial subtypes-lineage and transcriptome analysis. *Mol Cell Neurosci* **38:** 15–42

Ramos AD, Diaz A, Nellore A, Delgado RN, Park KY, Gonzales-Roybal G, Oldham MC, Song JS, Lim DA (2013) Integration of genome-wide approaches identifies lncRNAs of adult neural stem cells and their progeny *in vivo*. *Cell Stem Cell* **12:** 616–628

Rapicavoli NA, Poth EM, Blackshaw S (2010) The long noncoding RNA RNCR2 directs mouse retinal cell specification. *BMC Dev Biol* **10:** 49

Saburi S, Hester I, Goodrich L, McNeill H (2012) Functional inter-actions between Fat family cadherins in tissue morphogenesis and planar polarity. *Development* **139:** 1806–1820

Sakurai T (2012) The role of NrCAM in neural development and disorders–beyond a simple glue in the brain. *Mol Cell Neurosci* **49:** 351–363

Salomoni P, Calegari F (2010) Cell cycle control of mammalian neural stem cells: putting a speed limit on G1. *Trends Cell Biol* **5:** 332–342

Sheik Mohamed J, Gaughwin PM, Lim B, Robson P, Lipovich L (2010) Conserved long noncoding RNAs transcriptionally regulated by Oct4 and Nanog modulate pluripotency in mouse embryonic stem cells. *RNA* **16:** 324–337

Sone M, Hayashi T, Tarui H, Agata K, Takeichi M, Nakagawa S (2007) The mRNA-like noncoding RNA Gomafu constitutes a novel nuclear domain in a subset of neurons. *J Cell Sci* **120:** 2498–2506

Tarazona S, Garcia-Alcalde F, Dopazo J, Ferrer A, Conesa A (2011) Differential expression in RNA-seq: a matter of depth. *Genome Res* **21:** 2213–2223

Terra R, Luo H, Qiao X, Wu J (2008) Tissue-specific expression of B-cell translocation gene 2 (BTG2) and its function in T-cell immune responses in a transgenic mouse model. *Int Immunol* **20:** 317–326

Tirone F (2001) The gene PC3(TIS21/BTG2), prototype member of the PC3/BTG/TOB family: regulator in control of cell growth, differentiation, and DNA repair? *J Cell Physiol* **187:** 155–165

Uhde CW, Vives J, Jaeger I, Li M (2010) Rmst is a novel marker for the mouse ventral mesencephalic floor plate and the anterior dorsal midline cells. *PLoS One* **5:** e8641

Yao MJ, Chen G, Zhao PP, Lu MH, Jian J, Liu MF, Yuan XB (2012) Transcriptome analysis of microRNAs in developing cerebral cortex of rat. *BMC Genomics* **13:** 232

Yu J, Carroll TJ, Rajagopal J, Kobayashi A, Ren Q, McMahon AP (2009) A Wnt7b-dependent pathway regulates the orientation of epithelial cell division and establishes the cortico-medullary axis of the mammalian kidney. *Development* **136:** 161–171