Coordinated waves of gene expression during neuronal differentiation of embryonic stem cells as basis for novel approaches to developmental neurotoxicity testing

Marcel Leist, Bastian Zimmer, Philipp Balthasar Kuegler, Birte Baudis, Andreas Genewsky, Vivek Tanavde, Winston Koh, Betty Tan, Tanja Waldmann, Suzanne Kadereit

ABSTRACT

As neuronal differentiation of embryonic stem cells recapitulates embryonic neurogenesis, disturbances of this process may model developmental neurotoxicity (DNT). To identify the relevant steps of *in vitro* neurodevelopment, we implemented a differentiation protocol yielding neurons with desired electrophysiological properties. Results from focussed transcriptional profiling suggested that detection of non-cytotoxic developmental disturbances triggered by toxicants such as retinoic acid or cyclopamine was possible. Therefore, a broad transcriptional profile of the 20-day differentiation process was obtained. Cluster analysis of expression kinetics, and bioinformatic identification of overrepresented gene ontologies revealed waves of regulation relevant for DNT testing. We further explored the concept of superimposed waves as descriptor of ordered, but overlapping biological processes. The initial wave of transcripts indicated reorganisation of chromatin and epigenetic changes. Then, a transient upregulation of genes involved in the formation and patterning of neuronal precursors followed. Simultaneously, a long wave of ongoing neuronal differentiation started. This was again superseded towards the end of the process by shorter waves of neuronal maturation that yielded information on specification, extracellular matrix formation, disease-associated genes, and the generation of glia. Short exposure to lead during the final differentiation phase, disturbed neuronal maturation. Thus, the wave kinetics and the patterns of neuronal specification define the time windows and endpoints for examination of DNT.

INTRODUCTION

Pluripotent stem cells are able to form any cell type in an organism, including all different types of neurons. Ultimately, the entire complexity of the mammalian central nervous system (CNS) is generated during ontogenesis from a few single cells. This intricate process involves proliferation and cell fate determination events as well as migration, synaptogenesis, apoptosis and myelination. Many of these steps, in particular neuronal generation and differentiation, can be recapitulated by embryonic stem cells (ESC) under appropriate culture conditions[1-6].

ESC-based studies of neurodevelopment allow investigations and interventions not easily possible *in vivo*. However, known differentiation protocols differ in their suitability for such mechanistic or even toxicological studies. They were originally developed to generate large cell numbers e.g. for cell substitution therapies, and often focus on one defined neuronal subtype[4,7]. For instance, older protocols involve a step of embryoid body (EB) formation before neural induction is triggered[8]. EBs contain cells of all three germ layers, and neurogenesis is often enhanced by addition of the morphogen and patterning factor retinoic acid (RA)[7]. Frequently, only a small number of the initially-present ESC forms neurons, and the observation of individual cells is hardly possible. Other protocols use co-cultures with stromal cell lines like MS5[4] to differentiate ESC towards neurons, and would therefore introduce additional complexity into models for developmental neurotoxicity (DNT). A recently developed monolayer differentiation protocol allows monitoring of the differentiation procedure and of possible effects of different chemicals during the whole period of differentiation on a single cell level. Moreover, the use of fully defined media components allows for stringent system standardization[2,6,9].

DNT is the form of toxicity least examined and hardest to trace, as it is not necessarily related to cell loss. Behavioural pathology in the absence of cell loss is also known from disease

models, e.g. for Huntington's disease[10] or schizophrenia[11]. Toxicants, such as mercury, lead or polychlorinated biphenyls may trigger behavioral or cognitive deficits without histophathological hallmarks[12]. As the DNT endpoints are particularly difficult to test, less than 0.1% of frequently used industrial chemicals have been examined, and for the few known toxicants the mechanism of action is still elusive (reviewed in[12-14]). It has been suggested that cellular physiology (Ca^{2+} handling) may be affected during the period of exposure[15]. This may eventually lead to changes in differentiation and patterning in the CNS, which is the basis for long term effects that are observed after the exposure to toxicants has ceased.

CNS development is assumed to be orchestrated by waves of gene expression[16,17] that determine different intermediate cell phenotypes and form the basis for subsequent steps. Some periods may be more sensitive to certain toxicants than others. Epidemiological proof for such "windows of sensitivity" in organ development with long term consequences for the organism comes from thalidomide exposure in man[3] and various animal models[18].

The use of ESC may allow for new approaches to understand mechanisms of DNT and to evaluate the safety of chemicals. However, current test systems based on the differentiation of stem cells to either cardiomyocytes[19] or neural cells[14] neither yield mechanistic info, nor do they account for the complexity of CNS development, i.e. the establishment of a balance between multiple neuronal cell types[3,20]. The "toxicology for the 21st century" initiative[21,22] suggests the identification of pathways, and proposes a biologically-based explanation of toxicant effects, as opposed to the current black-box test systems. In the case of ESC-based models of DNT, this requires a detailed understanding of the developmental process leading to multiple different cell types.

Detailed knowledge on the waves of gene induction controlling different developmental steps would be an essential prerequisite. However, CNS development is proceeding at different paces. For instance, the anterior and posterior part of the neural tube

follow different kinetics, and some regions of the CNS continue neurogenesis, while in other regions cells have already reached fully postmitotic stages of their cells[20]. At present, it is not clear to which extent this is replicated in developing ESC cultures, but it is known that waves of neuronal development of different amplitudes may be detected *in vitro*[16,17]. These would need to be deconvoluted and characterized in depth in order to provide markers that comprehensively describe the neurodevelopmental process.

Our study was undertaken to analyze the wave-like expression pattern of mESC neurodevelopment as a basis for the definition of test windows and markers. This knowledge should help to identify non-cytotoxic, but neuroteratogenic compounds able to shift neuronal composition or phenotypes. Finally, the markers should distinguish multiple cell types and differentiation stages, and be able to indicate subpopulations of adherent cells that are inhomogeneously distributed.

Results

Monolayer differentiation of mESC to neurons

CGR8 were chosen as a widely available murine ESC line suitable for feeder-free culture maintenance and with established potential to develop along the neuroectodermal and neuronal lineage[23,27]. Differentiation was performed under defined serum-free and adherent conditions according to a published protocol[9]. The method was optimised for developmental neurotoxicity testing requirements, which demand high reproducibility between individual experiments. The efficiency of the final differentiation was confirmed by immunocytochemistry on day-of-differentiation 20 (DoD20). The majority of cells was positive for the pan-neuronal markers Tuj1 (neuronal form of betaIII-tubulin) and NeuN (neuronal-specific postmitotic nuclear antigen, encoded by *fox3*)[28]. Many cells also expressed synapse associated markers such as SV2 or PSD95, respectively (Fig. 1A). As a more quantitative overall measure for the robustness of the differentiation protocol to generate neurons from mESC, we chose mRNA expression, which we followed over time in five independent differentiation experiments, performed at different times, with different CGR8 cell batches, and by different operators. The kinetics of loss of the stemness marker *oct4*, transitional upregulation of the neural stem cell marker *nestin* and the delayed induction of the mature neuronal marker *synaptophysin* or the glial marker *gfap* were highly reproducible across different experiments (Fig. 1B). For functional characterization, neurons from at least three independent differentiations were tested for electrophysiological activity. Differentiation to mature, electrophysiologically-active neurons was shown by the presence of voltage-dependent Na^+ and K^+ and Ca^{2+} channels in individual patch-clamped neurons (Fig. 2A-C, Fig. S1). Further experiments also identified spontaneous neuronal electrical activity (Fig. 2D). Action potentials could be evoked by depolarization of individual neurons (Fig. S1), and currents were also evoked by exposure to N-methyl-D-aspartate (NMDA) or kainic acid. The

latter were blocked by the respective selective antagonists of NMDA and non-NMDA glutamate receptors (Fig. 2E). Alternatively, picrotoxin-sensitive currents were evoked by GABA (Fig. S1). Accordingly, the chosen neuronal differentiation protocol provides a solid basis for reproducible generation of *bona fide* neurons.

Transcription-based endpoints to identify disturbed neuronal differentiation

We next investigated whether we could detect subtle perturbations of the differentiation process below the cytotoxicity threshold. We tested whether mRNA-based readouts would fulfil the requirements of giving such information, e.g. altered neuronal patterning or shifts in subpopulations. Parallel mESC cultures were differentiated for 7, 15, and 20 days, and mRNA was prepared for quantitative RT-PCR analysis. These cells were treated during two different time windows (DoD1-7, DoD8-15) with two neuro-teratogens (Fig. 3A). With the concentrations used here cell death was not detectable (data not shown) and cells looked viable and were morphologically indistinguishable from untreated cells (Fig. 3B). We used the morphogen retinoic acid (RA) as a known *in vivo* and *in vitro* reproductive toxicant and cyclopamine for its ability to alter sonic hedgehog (Shh) signaling resulting in the disruption of patterning gradients responsible for floor plate and ventral neurons[2,20]. The analysis of neural patterning-associated marker genes in the chemical-treated cultures showed that their differentiation had been severely affected compared to untreated cultures. As expected from the literature[29], RA induced accelerated neuronal differentiation indicated by increased synaptophysin (marker of mature neurons) expression (Fig 3A, a,b) whereas cyclopamine reduced the expression of markers typical for more ventrally-located neurons like Shh, Nkx2.1 and Dlx1 (Fig 3A, a) but not overall neuronal differentiation (Fig. 3A). These data indicate that well chosen marker genes can indicate subtle shifts in differentiation patterns not visible morphologically. This was further analyzed in different variations of the incubation scheme with a set of stage-specific markers (Fig. 3A, b-d). For instance cultures were exposed

to cyclopamine from DoD1-7 (Fig. 3A, b) and immediately analyzed thereafter. Treatment did not affect the overall formation of NPC (as indicated by unchanged nestin expression), but again the reduced Shh expression suggested a reduced ventral development. When cells were released from cyclopamine after 7 days of treatment and left to differentiate further the compound effect on shh expression was still observable on day 15 (Fig. 3A, c). The known dorsalizing activity of cyclopamine would suggest a shift of neurotransmitter phenotype from GABAergic (Gad2 as marker) to glutamatergic (vglut1 as marker)[2]. This was not observed after treatment for the first seven days (Fig 3A, c), but a significant decrease of Gad2 (more ventrally prominent) was observable when the cells were treated during a short transient period between DoD8 and 15 (Fig. 3A, d).

In the case of RA, the acceleration of development (increased synaptophysin expression) was already significant at early stages (Fig 3A, b). In addition we examined its known caudalizing effect[29]. We found upregulation of markers usually expressed in caudal parts of the neural tube (hoxa6, hb9), and associated with the development of motor neuron precursors (isl1) (Fig. 3A, c,d).

Finally, we examined whether inhibited (instead of shifted) differentiation by non-cytotoxic chemicals was detectable by RNA markers. We used 3i, a kinase inhibitor mix known for inhibiting differentiation of mESC[26]. Early exposure (Dod0-7) resulted in cultures with an immature cellular phenotype (data not shown) and retarded neural differentiation indicated by a decreased expression of neural markers hes5, nestin and betaIIItubulin and an increased expression of the stemness factor Oct4 (Fig. S2, a). When cells were treated on DoD1-7, washed thoroughly, and then left to differentiate without interference until DoD15, 3i did not show effects on Oct-4 expression anymore, but neural differentiation was severely delayed as seen by significantly lowered expression of nestin, hes5 and the mature neuronal marker synaptophysin (Fig. S2, c). Treatment of cells with 3i after DoD7 (after neural differentiation had been initiated) did not return them to the stem cell state, but was rather cytotoxic. In

summary the usefulness of transcript profiling for detection of patterning disturbances was clearly demonstrated and confirmed.

The low variation of the data across different experiments suggests that transcript profiling is useful for quantitative assessments of disturbed neuronal differentiation. As the effects of chemicals are likely to be different, depending on the time window of exposure and on the differentiation stage assessed, it may be necessary to measure the impact on differentiation at different DoD, and specific markers need to be selected for each stage. In order to identify such potential markers, global, and kinetically-resolved transcription profiling was performed.

Identification of clusters of genes regulated during neuronal differentiation of mESC.

We analyzed changes in the transcriptome over time (DoD 0, 7, 15, 20) based on oligonucleotide microarrays. By using a few marker genes, we verified that the differentiation kinetics of the cultures used for microarray analysis matched the ones observed during many other well-controlled experiments performed earlier (Fig. 1B). Having established this important correlation, the kinetics of expression of each gene represented on the chip was used as input for an unbiased clustering analysis. This analysis yielded eight groups of regulation profiles (Clusters Ia, Ib, IIa, IIb, IIIa, IIIb, IV and V) (Fig. 4A, S3), besides the group of genes not regulated at all or not yielding any signal in this analysis (genes not shown). Cluster Ia was characterized by rapid (within the first 7 days), and cluster Ib by slow (continuous from DoD0 to DoD20) downregulation. These two clusters obtained by non-hierarchical analysis exemplify the principle of superimposed gene regulation waves with different amplitudes (see also clusters III-V). Clusters IIa and IIb contained genes that were transiently regulated at DoD7 (Fig. 4A) (IIa: transient upregulation, IIb: transient downregulation). Cluster IIIa and IIIb were characterized by a rapid increase of transcripts between day 0 and DoD7 maintained then at high levels (Fig 4A). Cluster IV contained genes,

which remained low until DoD7 and then reached high levels on DoD15, with no major changes until DoD20 (Fig 4A). The final cluster V comprised transcripts that were hardly upregulated until DoD15, and reached their maximum on DoD20 (Fig. 4A). Thus, large groups of genes appear to be regulated in a co-ordinated manner in defined, and partially overlapping, waves. Notably, the most-pronounced changes in expression occurred during culture periods not involving any manipulation of the cells, i.e. between DoD0 and DoD7 (cells are plated on day 0 and only replated on day 7), and between DoD15 and DoD20. This has important implications for the choice of relevant treatment windows in studies of developmental neurotoxicity.

To explore this further, the genes falling into the different clusters were subjected to a more detailed analysis. First, we controlled for the efficiency of differentiation, and examined the behaviour of about 40 genes that characterize the initial mESC stage[3]. We found 33 of them differentially regulated in our analysis and examined their distribution across the individual clusters. About 80% of these fell into cluster Ia and Ib (overall downregulation), while the remainder was found in cluster IIb (downregualtion from DoD0-DoD7). Thus, all mESC markers identified on the chip were confirmed to be downregulated upon initiation of the neuronal differentiation (Fig. 4B). Next, we proceeded to ask how neural precursor cell (NPC) and neuronal differentiation (N) markers clustered. About 130 N-markers were found in the different clusters. NPC markers (n = 73) were assembled from the literature[3], and 63 of them were found in our analysis. Concerning the relative distribution between the clusters, most NPC markers were found in clusters IIa and IIIa/b (both containing genes upregulated early) and about 15% were found in clusters IV and V (Fig. 4B). In contrast to this, most N markers were found in the clusters with increasing gene expression (clusters III-V), while about 20% were found in cluster IIa (transient upregulation on DoD7) (Fig. 4B). These findings strongly suggest that the clusters identified by unbiased bioinformatics methods, correspond to waves of real biological processes describing the differentiation process of mESC to neurons. To

explore this working hypothesis further, we continued with a detailed analysis of the biological significance of genes in individual clusters.

Loss of pluripotency is accompanied by progressive changes in transcripts responsible for chromatin organization and DNA/cell cycle functions

Initially, genes in cluster I were analyzed for GO categories significantly overrepresented. For instance, cell cycle-related processes were identified (Fig 5A, S4). More unexpectedly, also chromatin structure and epigenetic processes seemed to be affected (Fig 5A). We examined this in more detail. Gene lists of relevant processes were assembled both with the help of the GO data base and extensive literature search. The clusters were then queried for the presence of these genes (Fig. 5A, S4). Interestingly, of 14 genes known to be associated with chromosome structure, all were identified in clusters Ia and Ib (overall downregulation). The same applied to genes associated with DNA replication (n = 57), DNA repair (n = 62) and DNA methylation (n = 5). Also, most of the genes coding for histones, histone modifiers, chromatin remodelling and chromatin substructuring were found in cluster I (Fig. 5A, S4). We also examined whether the transcriptional changes gave clear evidence for global quantitative changes in chromatin structure. However, the observed pattern of regulated genes rather suggested a restructuring of the chromatin without an overall increase of DNA methylations or histone modifications (Fig. S4). Therefore restructuring of pericentric heterochromatin was examined by high resolution confocal microscopy. While the chromatin, distributed relatively homogeneously over the nucleus in mESC, it was organized entirely differently after 20 days of differentiation, with neurons characterized by conspicuously large (1-2 µm scale) dense areas of chromatin (Fig. 5B). In general, many genes of cluster Ia/b may be typical for the biological processes initiated in an actively cycling stem cell developing towards a post-mitotic and highly differentiated cell. As indicated by our data, this involves many different processes running in parallel. Among the identified genes, four (Smarca1,

Myst4, Jmjd3 and Hdac11) are known to be neurospecific, and five (Suz12, Ezh2, Bmi1, Cbx2 and Cbx8) are components of the polycomb repressor complexes (PRC), which play an important role in differentiation-related control of gene promoters. These genes could serve as sensitive markers to detect negative effects of compounds on early developmental processes (Fig. S4).

Correlation of neural precursor formation with a strong, transient change of gene expression levels

We examined whether genes of cluster II (Fig. 4A) were specifically linked to the process of neural precursor cell (NPC) formation, an essential step in neurodevelopment. One of the genes to be expected, and actually found, in cluster IIa was nestin(Fig. 1B), an established marker for NPC[30]. We confirmed nestin expression by immunocytochemistry, with a majority of cells expressing this cytoskeletal protein on DoD7. Moreover, nestin-positive cells often seemed to be arranged in ring-like structures, reminiscent of so-called rosettes, structures resembling two-dimensional neural tubes[31,32] (Fig. 6A). Quantification by flow cytometry analysis showed that about 80% of all cells in the culture became nestin-positive (Fig. 6B). High synchronization at this stage was suggested by the very distinct and sharp expression profile of genes in cluster IIa (Fig. 6C). Some genes were upregulated several thousand-fold on DoD7, and strongly downregulated again on DoD15 and DoD20. Besides nestin, many other genes typically associated with neuroepithelial precursors, NPC, and neurogenesis were found in cluster IIa (S3). In addition, some genes associated with early, but definitive neuronal development, and with specification and patterning of neuronal subtypes were identified (e.g. Dll1, Hes3). However, cluster IIa also contained genes thought to be involved in multiple cellular and metabolic processes besides neurodevelopment (e.g. Janus kinase 2 (Jak2), forkhead box D4 (Foxd4), B-cell leukemia/lymphoma 2 (Bcl-2), kinesin family member 21A (Kif21a), angiotensin II receptor, type 1a (Agtr1a, also known as AT1),

monooxygenase-DBH-like 1 (Moxd1), acetoacetyl-CoA synthetase (Aacs), ADP-ribosylation factor-like 2 binding protein (Arl2bp), stearoyl-Coenzyme A desaturase 2 (Scd2). Therefore, we performed an unbiased and quantitative analysis of the biological significance of cluster IIa, and examined which of the 16000 gene ontology (GO) categories were statistically overrepresented by cluster IIa genes. The result confirmed the efficacy of the chosen differentiation protocol, as the GO "nervous system development" emerged with a p-value < 10^{-13}, and only neuronal/neurodevelopmental GOs were identified with the exception of eight (weakly significant) potential ossification genes (Table 1). Thus, genes of cluster IIa are associated with determination of early neuroectodermal/neuronal development and their expression could represent an important endpoint for testing of disturbed proliferation and differentiation during this early and crucial developmental time window.

Markers of regional fate decisions in the CNS

Overall *in vivo* neurodevelopment is a non linear sequential process, even at this initial stage. Early central nervous system formation is characterized by local differences in timing and patterning[20] that may be reflected in some *in vitro* systems, while others show already a clear regional fate decision[1]. To characterize our differentiation system more thoroughly, we examined region-specific markers. Expression of regional markers was relatively limited in cluster IIa. Amongst the few markers expressed, forebrain (e.g. Foxg1) and hindbrain (e.g. Hoxa2/b2)-related indicator genes were evenly distributed in cluster IIa (Fig. 6D, S5). More patterning markers were found in the clusters containing continuously upregulated genes (clusters IIIa/b and IV+V). Also here, forebrain (e.g. Reln), midbrain (e.g. En1/2), and hindbrain (e.g. Lmx1a or Hoxa1) markers were evenly distributed (S5). Accordingly, our experimental model appears to reflect several parallel lines of *in vivo* neural specification. In this system with broad developmental potential, the ratios of expression of different patterning

markers may provide very sensitive indicators of disturbed neurodevelopment worth an in depth exploration as piloted in figure 3.

Specificity for neuronal induction with respect to glial cells

Most neuronal differentiation systems generate a certain percentage of neurons, but frequently also less-characterized populations of additional neural cells *plus* non-neural cells. Characterization of these undesirable cell types within the culture is of particular importance for quantitative assays of disturbed neurodevelopment, where such populations may increase on cost of neurons. Moreover, such cells could confound (enhance or decrease) neurotoxicity of tested compounds. The now available broad transcriptional profile allowed us a more detailed analysis. First, we examined gliogenesis, as Gfap was sharply upregulated on DoD20 (Fig. 1B), and some small GFAP-positive cell areas were reproducibly identified by immunocytochemistry (Fig. 7A). These cells displayed typical astrocytic morphology and were found in 1-2 small islands/cm^2. An unbiased search for overrepresented GO categories did not result in any hits related to gliogenesis or glial function (Table 2). As alternative approach, we used a list of 25 astrocyte-related genes[3] and found 11 of them to be upregulated on DoD20 compared to DoD0, with 4 additional astrocytes-related genes transiently upregulated on DoD7 (Fig. 7B). Most likely, the early upregulation (e.g. vimentin) of apparent astrocytic markers is due to the generation of radial glia-like NPC at DoD7. This cell type, as exemplified by the upregulation of Fabp7 (brain lipid binding protein; in cluster IIIb)[33] or ascl1 (=Mash1)[34] shares many markers with astrocytes[5]. In contrast to this, the late upregulation of many mature astroglial markers is due to a real generation of astrocytes, as corroborated independently by immunocytochemistry (Fig. 7A) and by qPCR of a larger set of markers and (Fig 1B, 7C). Small increases of this normally minor cell population may affect toxicity testing during the later differentiation phases and requires good controls on the

basis of mRNA markers identified here. Any potential role of astrocytes appears to be insignificant for early phases of DNT testing.

We also examined genes known to be expressed in oligodendrocytes and microglia. The latter cell type appeared to be absent. Few oligodendrocyte markers were upregulated, while many functionally important genes were not (Fig. 7B). By immunocytochemistry, we were not able to identify any oligodendrocytes either (not shown). Thus their contribution to the DoD20 cultures appears to be negligible.

Specificity for neuronal induction with respect to other germ layer lineages

Finally, all GO categories significantly overrepresented by the genes of clusters III-V (upregulation on DoD20 vs DoD0) were determined bioinformatically, and searched for evidence of non-neural cell type formation. Individual clusters did not indicate any non-neural cell types while representation of neuronal GOs was highly significant (Table 2). Upon pooled analysis of clusters IV and V (late upregulation), the GOs “blood vessel development” and “muscle organ development” emerged as significant. Thus, it is possible that a subpopulation of cells present on DoD20 may have a smooth muscle phenotype or partial features of such cells.

Waves of clustered genes related to neuronal induction

The classic characterization of neuronal differentiations is heavily based on immunocytochemistry (Fig. 1A). This approach allows only for detection of a few antigens, and quantitative analysis has been difficult in cultures with heterogeneous cell distribution. Here, we used instead transcriptional profiling. For characterization of the cultures, we used groups of genes identified by non-biased clustering as basis to bioinformatically extract the main GO categories represented by them (Table 2). In addition, a complementary approach based on literature and expert judgement was chosen to select interesting groups of genes (Fig. 8). Novel information emerged from this combination of strategies. Most importantly,

we found that the differentiation did not proceed as a linear sequence of synchronized sequential steps, although clear and distinct waves of genes had been identified by cluster analysis. The process rather involved strongly overlapping processes with one underlying large wave (cluster IIIa/b) superseded by shorter waves (cluster IV and V). For instance, generation of neurons and axogenesis/growth cone formation seemed to be ongoing in the entire period from DoD7 to DoD20 as indicated by groups of neuroreceptors and growth cone/axon guidance-related genes in cluster III (Fig. 8A, B). A larger group of genes associated with synaptic vesicles or the transmission of nerve impulse only appeared later (cluster IV/V) (Fig 8A). In the latest phase, hardly any new genes associated with growth cones and neurodevelopment were induced. Instead, genes associated with responses to stress and hormonal stimuli were strongly up-regulated (Table 2). Interestingly, one frequently neglected group of genes also emerged significantly: the “regulation of extracellular matrix components” (Fig 8C). Such genes were strongly and distinctly up-regulated in late differentiation stages, similar to the more expected transporters, genes associated with neuronal projection and post-synaptic membranes. A further interesting group of genes was highly upregulated during the last step of differentiation and involved those known to be “associated with hereditary neurodegenerative diseases”. Reanalysis by quantitative PCR confirmed this finding and yielded regulation factors (DoD20 vs DoD0, n = 2 differentiations) of: 92-fold and 16-fold for the Alzheimer’s disease associated genes App and Mapt, 273-fold for the schizophrenia-associated gene Nrnx1, 91-fold for the prion protein Prnp, and 19-fold/ 56-fold for the Parkinson’s disease-related genes Pink1/ Snca.

We wondered whether toxicants may affect this very late phase of neuronal differentiation identified by our transcriptional profiling approach. We chose lead as well-established toxicant with a purported role in the developmental origins of neurodegenerative diseases (see Fig. S6). The cells were treated with a non cytotoxic concentration (assessed by resazurin reduction and LDH release, data not shown) of lead (1 µM) only during the last phase of

differentiation (DoD14-DoD20). The transcript levels of two neuronal markers and the set of disease associated genes identified above were used to examine differences in differentiation. The mRNA expression of Tubb3 and Syp was significantly reduced. Moreover lead exposure had also a dampening effect on the expression of App, Mapt, Nrnx1 and Prnp (Fig. 8D). Pink1 and Snca were not affected. Also their relative increase with respect to the pan-neuronal marker synaptophysin was not significant. The knowledge on markers of such different processes together with that of the expected timing of their expression provides an ideal toolkit for fine-mapping of subtle developmental disturbances.

DISCUSSION

We have here demonstrated a concept of overlapping waves of gene regulation and suggested its use to define protocols, test windows and endpoints for developmental neurotoxicity testing. Our findings should be helpful to close a gap between two highly developed, but isolated disciplines: experimental developmental neurobiology and toxicology. The former has been highly successful in defining the functional importance, regional expression and cell type/stage association of important genes. The latter has an urgent need for robust and sensitive marker genes to identify disturbances of development. As functionally important genes are not always suitable markers, and good markers sometimes lack a functional role, it was important to examine the complex system of mESC differentiation to neurons to define patterns of gene expression changes useful as toxicological endpoints. We confirmed here earlier findings[2,6,9,27] that the chosen differentiation protocol yields highly reproducible results and supports the development of multiple neuronal subtypes, while maintaining a high specificity for neurons as opposed to other cells. In this system we showed that subtle changes in the speed of differentiation, or in dorso-ventral or anterior-posterior patterning due to toxicants can be detected by using the right choice of mRNA markers. Such changes may be considered *in vitro* correlates of known teratogenic effects of the chosen compounds. For instance, cyclopamine is known to cause dramatic patterning disturbances in a very defined period of brain development, and a defined brain region (holoprosencephaly); and retinoic acid is well-known to cause shifts in the anterior-posterior axis organisation favouring the more posterior parts, as found here by transcript markers. Lead is known to subtly affect multiple neuronal types, which is in agreement with the broad pattern of disturbances found here for exposure to lead during late neuronal differentiation (see Fig. S6 for references). However, the data also suggest some warning on the limitations of in vitro – in vivo correlations. Although the cyclopamine data in this study suggest some type of disturbance in

patterning, they would not indicate a problem in the separation of the forebrain hemispheres, as observed in animal studies. Similarly, the *in vitro* test system would not be able to predict the non-neuronal teratogenic effects of retinoic acid. Thus, observations from stem cell systems will have a major value for raising alerts on certain compounds and pinpointing potential mechanisms, while complementary data from other systems may be required to predict specific effects on humans.

Transcriptional profiling studies frequently rely predominantly on stringent bioinformatic analysis. Although this approach appears to be unbiased, it strongly suffers from the weakness and errors of data bases and algorithms. An additional problem is the visualization of the large amount of data in a form that generates meaningful knowledge instead of long tables or unreadable heat maps. In more extensive studies using hundreds to thousands of chips, the generation of data bases with user-friendly interfaces for secondary individual analysis can be a solution e.g. FunGeneES for mESC differentiation[35]. This does, however, not solve the inherent bioinformatic problems. For example, assignment of genes to GO categories is not always perfect. For instance, the GO for gliogenesis contains ubiquitous signalling and metabolic molecules (such as Igf1, Sod1) as well as highly specific transcription factors. On the other hand, typical astrocyte markers such as Gfap and glutamine synthetase are not members of this GO. It is evident that bioinformatic analyses to identify markers are problematic on such a basis. First, certain important markers may not be present, second genes may be identified as markers, although they do not qualify biologically and third, the equal weight given to ubiquitous vs. specific genes in statistical analysis results in biological skewing. With these considerations in mind, we chose to combine bioinformatic analysis with classical knowledge-based approaches. The latter were used to control and extend findings based on non-biased approaches, but were also used independently, for the definition and compilation of stem cell and glial markers, and for the analysis of biological functions related to chromatin and neuronal function (Fig 5 and 8). During this procedure, the entire hit list of

several thousand genes was manually screened, sorted and annotated. For the former approach, a large consortium of experts was consulted, and results were compiled in an open access review format[3]. We strongly advocate such combined approaches for toxicological systems biology, which is at present strongly driven by computational methods[22,36].

We used different types of endpoints as framework for the transcriptional profiling. Electrophysiology was chosen as functional endpoint to verify that culture conditions were optimal. The studies show that cells with a distinct neuronal morphology indeed expressed the major types of channels and were electrically- or chemically excitable and able to generate action potentials (Fig 2, S1). Although we included replicates from various differentiations, these studies have a rather qualitative character, as the cells that were patched were chosen based on certain morphological criteria and may not represent the entire culture. However, our results fully corroborate earlier findings that functional neurons can be generated from mESC[7,8] and extend them to our particularly defined culture conditions. Immunostaining and quantitative RT-PCR were used as classical and established methods to link chip-based transcript profiling to other experiments that have been performed with much higher replicate number. Notably, although the shape of PCR and chip profiles was very similar, the induction levels differed sometimes. Such variations may be due to different hybridization efficiencies and background levels. Therefore, very extensive studies, involving RT-PCR controlled by internal standards, will be necessary for a quantitative definition of a final set of markers. Notably, we did not use differences in absolute numbers of regulations in the present study as basis for any of our conclusions. In addition, all major conclusions are built on groups of co-regulated and biologically linked genes as opposed to speculations based on the presence or absence of a single gene with possibly inappropriate oligonucleotide probe. In this context, we are also aware of the issue, that the levels of mRNA do not essentially correlate with the expression, modification and localization of protein, which eventually determines the cellular phenotype. Even if such correlations were given, our approach should not be interpreted as

phenotype definition on single cell resolution. This would require different approaches, that are able to measure several markers within one cell[37]. The genes grouped within the clusters described here are not necessarily expressed in the same cell (e.g. cluster V: Gfap (astrocytic) and syntaxin (neuronal)) and therefore do not automatically describe a single biological entity. However, with these caveats, we feel that markers that plausibly describe relevant biological processes, and that can be sensitive indicators of disturbances of the default development, can still be selected with confidence on the basis of our study.

In the area of developmental toxicology and especially in DNT, cause-effects relationships are still mostly unknown, and human epidemiological data are only available for a handful of industrial chemicals[12]. Rodent data based on the OECD test guideline 426 are at least available for around 200 substances[13]. With this lack of human-relevant information and the better animal data base, it appears reasonable to us to perform proof-of-principle experiments for the usefulness of a new approach in rodent cells first, and to validate human cells against these in case of a positive outcome. Moreover, mESC have been shown to model even complex aspects of rodent brain development faithfully, as far as neuronal specification is concerned[2].

At present, DNT studies are based on e.g. behavioral, cognitive or neuropathological endpoints, and the next step towards mechanistic information would be an understanding of changes on the level of cells and gene expression. The overlapping waves defined here would provide a conceptual framework for this. Such waves (i.e. spatially shifting activation) of gene expression are known from many pioneering studies of mammalian in vivo CNS development[20] and are for instance well-characterized in high density and resolution in the hippocampus[38]. Waves have also been defined in vitro in mESC[17,35] or differentiating embryonic carcinoma cells[16,39]. Here, we extended this concept, by relating regulation clusters to underlying biological processes important for toxicity testing. This translation from

developmental biology to the toxicological perspective defines the windows of sensitivity relevant for test protocols

In the field of cardiac development, the mESC based embryonic stem cell test (EST) has been thoroughly validated and frequently applied[19]. Exposure of cells during the entire test period can result in data that are hard to interpret and that are confounded by relatively unspecific toxicity. Therefore separation of exposure into the proliferation and differentiation phase has been suggested[40]. We want to expand this principle here by suggesting four relevant test periods. DoD1-7: testing of lineage commitment, efficiency of NPC formation, axon formation and guidance, and of epigenetic changes associated with the transition from pluripotent cells to more committed NPC. DoD8-15: major phase of neuronal patterning and vesicle development. DoD15-20: a more unexpected, but highly interesting and relevant phase, when most proliferation has ceased, and maturation becomes evident by expression of matrix components, important transporters and disease-associated genes. Our data on lead exposure during this phase show that it will be of high importance for future testing. DoD20+ has not been explored here. It requires, however, further investigation to determine whether this period can be used as stable reference for neurotoxicity *vs* DNT, or whether new processes such as synaptogenesis, gliogenesis, or myelination take a dominant role here.

Examples for markers of NPC and neuronal differentiation, and the issues associated with such a selection have already been compiled[3]. The major task for the future will be the validation of a larger set of such markers, first with known specific and mechanistically-defined disruptors of developmental pathways, then with known DNT compounds, in order to select the smallest group of final markers useful for a comprehensive description of toxicities triggered by the test compounds.

Conflict of interest

The authors declare no conflict of interest

Acknowledgement

The Work was supported in part by the Doerenkamp-Zbinden Foundation, the DFG, the EU FP7 project ESNATS (ML, SK), an IRTG1331 fellowship (BZ) and a fellowship from the KoRS-CB (PBK). We are grateful to Giovanni Galizia and Sabine Kreissl for help with the electrophysiological recordings and indebted to Bettina Schimmelpfennig for invaluable experimental support. The monoclonal antibodies Gad-6 developed by D.I. Gottlieb and SV2 developed by K.M. Buckley were obtained from the Developmental Studies Hybridoma Bank developed under the auspices of the NICHD and maintained by The University of Iowa, Department of Biology, Iowa City, IA 52242. We thank K.H. Krause for the CGR8 mESC-line and J. Vilo and S. Ilmjärv for help with bioinformatics analysis.

1. Conti, L and Cattaneo, E, (2010) Neural stem cell systems: physiological players or in vitro entities? Nat Rev Neurosci 11: 176-87.
2. Gaspard, N, Bouschet, T, Hourez, R, Dimidschstein, J, Naeije, G, van den Ameele, J et al., (2008) An intrinsic mechanism of corticogenesis from embryonic stem cells. Nature 455: 351-7.
3. Kuegler, PB, Zimmer, B, Waldmann, T, Baudis, B, Ilmjarv, S, Hescheler, J et al., (2010) Markers of murine embryonic and neural stem cells, neurons and astrocytes: reference points for developmental neurotoxicity testing. Altex 27: 17-42.
4. Barberi, T, Klivenyi, P, Calingasan, NY, Lee, H, Kawamata, H, Loonam, K et al., (2003) Neural subtype specification of fertilization and nuclear transfer embryonic stem cells and application in parkinsonian mice. Nat Biotech 21: 1200-1207.
5. Götz, M and Huttner, WB, (2005) The cell biology of neurogenesis. Nat Rev Mol Cell Biol 6: 777-88.
6. Abranches, E, Silva, M, Pradier, L, Schulz, H, Hummel, O, Henrique, D et al., (2009) Neural differentiation of embryonic stem cells in vitro: a road map to neurogenesis in the embryo. PLoS One 4: e6286.
7. Lee, S-H, Lumelsky, N, Studer, L, Auerbach, JM and McKay, RD, (2000) Efficient generation of midbrain and hindbrain neurons from mouse embryonic stem cells. Nat Biotech 18: 675-679.
8. Strübing, C, Ahnert-Hilger, G, Shan, J, Wiedenmann, B, Hescheler, J and Wobus, AM, (1995) Differentiation of pluripotent embryonic stem cells into the neuronal lineage in vitro gives rise to mature inhibitory and excitatory neurons. Mechanisms of Development 53: 275-287.
9. Ying, QL and Smith, AG, (2003) Defined conditions for neural commitment and differentiation. Methods Enzymol 365: 327-41.
10. Hansson, O, Petersén, A, Leist, M, Nicotera, P, Castilho, RF and Brundin, P, (1999) Transgenic mice expressing a Huntington's disease mutation are resistant to quinolinic acid-induced striatal excitotoxicity. Proceedings of the National Academy of Sciences of the United States of America 96: 8727-8732.
11. Penschuck, S, Flagstad, P, Didriksen, M, Leist, M and Michael-Titus, AT, (2006) Decrease in parvalbumin-expressing neurons in the hippocampus and increased phencyclidine-induced locomotor activity in the rat methylazoxymethanol (MAM) model of schizophrenia. Eur J Neurosci 23: 279-84.
12. Grandjean, P and Landrigan, PJ, (2006) Developmental neurotoxicity of industrial chemicals. Lancet 368: 2167-78.
13. Makris, SL, Raffaele, K, Allen, S, Bowers, WJ, Hass, U, Alleva, E et al., (2009) A retrospective performance assessment of the developmental neurotoxicity study in support of OECD test guideline 426. Environ Health Perspect 117: 17-25.
14. Bal-Price, AK, Hogberg, HT, Buzanska, L, Lenas, P, van Vliet, E and Hartung, T, (2009) In vitro developmental neurotoxicity (DNT) testing: Relevant models and endpoints. Neurotoxicology.
15. Rossi, AD, Larsson, O, Manzo, L, Orrenius, S, Vahter, M, Berggren, PO et al., (1993) Modifications of Ca2+ signaling by inorganic mercury in PC12 cells. FASEB J. 7: 1507-1514.
16. Wei, Y, Harris, T and Childs, G, (2002) Global gene expression patterns during neural differentiation of P19 embryonic carcinoma cells. Differentiation 70: 204-19.
17. Aiba, K, Sharov, AA, Carter, MG, Foroni, C, Vescovi, AL and Ko, MS, (2006) Defining a developmental path to neural fate by global expression profiling of mouse embryonic stem cells and adult neural stem/progenitor cells. Stem Cells 24: 889-95.

18. Jongen-Relo, AL, Leng, A, Luber, M, Pothuizen, HH, Weber, L and Feldon, J, (2004) The prenatal methylazoxymethanol acetate treatment: a neurodevelopmental animal model for schizophrenia? Behav Brain Res 149: 159-81.
19. Marx-Stoelting, P, Adriaens, E, Ahr, HJ, Bremer, S, Garthoff, B, Gelbke, HP et al., (2009) A review of the implementation of the embryonic stem cell test (EST). The report and recommendations of an ECVAM/ReProTect Workshop. Altern Lab Anim 37: 313-28.
20. Rao, MS, Jacobson, M., (2005) Developmental Neurobiology. (Kluwer Academic/Plenum Publishers, New York).
21. Collins, FS, Gray, GM and Bucher, JR, (2008) TOXICOLOGY: Transforming Environmental Health Protection. Science 319: 906-907.
22. Leist, M, Hartung, T and Nicotera, P, (2008) The dawning of a new age of toxicology. Altex 25: 103-14.
23. Suter, DM, Tirefort, D, Julien, S and Krause, KH, (2009) A Sox1 to Pax6 switch drives neuroectoderm to radial glia progression during differentiation of mouse embryonic stem cells. Stem Cells 27: 49-58.
24. Livak, KJ and Schmittgen, TD, (2001) Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the 2-[Delta][Delta]CT Method. Methods 25: 402-408.
25. Reimand, J, Kull, M, Peterson, H, Hansen, J and Vilo, J, (2007) g:Profiler--a web-based toolset for functional profiling of gene lists from large-scale experiments. Nucl. Acids Res. 35: W193-200.
26. Ying, QL, Wray, J, Nichols, J, Batlle-Morera, L, Doble, B, Woodgett, J et al., (2008) The ground state of embryonic stem cell self-renewal. Nature 453: 519-23.
27. Conti, L, Pollard, SM, Gorba, T, Reitano, E, Toselli, M, Biella, G et al., (2005) Niche-independent symmetrical self-renewal of a mammalian tissue stem cell. PLoS Biol 3: e283.
28. Kim, KK, Adelstein, RS and Kawamoto, S, (2009) Identification of Neuronal Nuclei (NeuN) as Fox-3, a New Member of the Fox-1 Gene Family of Splicing Factors. Journal of Biological Chemistry 284: 31052-31061.
29. Irioka, T, Watanabe, K, Mizusawa, H, Mizuseki, K and Sasai, Y, (2005) Distinct effects of caudalizing factors on regional specification of embryonic stem cell-derived neural precursors. Brain Res Dev Brain Res 154: 63-70.
30. Kawaguchi, A, Miyata, T, Sawamoto, K, Takashita, N, Murayama, A, Akamatsu, W et al., (2001) Nestin-EGFP transgenic mice: visualization of the self-renewal and multipotency of CNS stem cells. Mol Cell Neurosci 17: 259-73.
31. Elkabetz, Y, Panagiotakos, G, Al Shamy, G, Socci, ND, Tabar, V and Studer, L, (2008) Human ES cell-derived neural rosettes reveal a functionally distinct early neural stem cell stage. Genes Dev 22: 152-65.
32. Zhang, SC, Wernig, M, Duncan, ID, Brustle, O and Thomson, JA, (2001) In vitro differentiation of transplantable neural precursors from human embryonic stem cells. Nat Biotechnol 19: 1129-33.
33. Feng, L, Hatten, ME and Heintz, N, (1994) Brain lipid-binding protein (BLBP): a novel signaling system in the developing mammalian CNS. Neuron 12: 895-908.
34. Battiste, J, Helms, AW, Kim, EJ, Savage, TK, Lagace, DC, Mandyam, CD et al., (2007) Ascl1 defines sequentially generated lineage-restricted neuronal and oligodendrocyte precursor cells in the spinal cord. Development 134: 285-93.
35. Schulz, H, Kolde, R, Adler, P, Aksoy, I, Anastassiadis, K, Bader, M et al., (2009) The FunGenES database: a genomics resource for mouse embryonic stem cell differentiation. PLoS One 4: e6804.

36. Hartung, T and Hoffmann, S, (2009) Food for thought ... on in silico methods in toxicology. Altex 26: 155-66.
37. Kawaguchi, A, Ikawa, T, Kasukawa, T, Ueda, HR, Kurimoto, K, Saitou, M et al., (2008) Single-cell gene profiling defines differential progenitor subclasses in mammalian neurogenesis. Development 135: 3113-24.
38. Mody, M, Cao, Y, Cui, Z, Tay, KY, Shyong, A, Shimizu, E et al., (2001) Genome-wide gene expression profiles of the developing mouse hippocampus. Proc Natl Acad Sci U S A 98: 8862-7.
39. Przyborski, SA, Smith, S and Wood, A, (2003) Transcriptional profiling of neuronal differentiation by human embryonal carcinoma stem cells in vitro. Stem Cells 21: 459-71.
40. van Dartel, DA, Zeijen, NJ, de la Fonteyne, LJ, van Schooten, FJ and Piersma, AH, (2009) Disentangling cellular proliferation and differentiation in the embryonic stem cell test, and its impact on the experimental protocol. Reprod Toxicol 28: 254-61.

Materials and Methods

Materials

Unless otherwise mentioned, cell culture media and reagents were from Invitrogen (Carlsbad, USA) and accessory reagents from Sigma. **Antibodies**: anti-Tuj1 (cat. # MMS-435P; Covance), anti-NeuN (cat. # MAB377; Chemicon), anti-GAD65 (GAD-6; DSHB), anti-SV2 (SV2; DSHB), anti-PSD95 (cat. # 51-6900; Zymed), anti-Nestin (cat. # MAB353; Chemicon), anti-GFAP (clone: G-A-5; Sigma) anti-Nestin-647 (clone: 25/NESTIN; BD Biosciences).

mRNA Primer: Pou5f1-fw: CTCTTTGGAAAGGTGTTCAGCCAGAC, Pou5f1-re: CGGTTCTCAATGCTAGTTC-GCTTTCTC; Nestin-fw: CTGGAAGGTGGGCAGCAACT, Nestin-re: ATTAGGCAAGGGGGAAGAGAAGGATG; Synaptophysin-fw: GGGTCTTTGCCATCTTCGCCTTTG, Synaptophysin-re: CGAGGAGGAGTAGTCACCAACTAGGA; Gfap-fw: GCCCGGCTCGAGGTCGAG, Gfap-re: GTCTATACGCAGCCAGGTTGTTCTCT; Shh-fw: CAGCGGCAG-ATATGAAGGGAAGATCA, Shh-re: GTCTTTGCACCTCTGAGTCATCAGC; Hes5-fw: CCCAAGGAGAAAAACC-GACTGCG, Hes5-re: CAGCAAAGCCTTCGCCGC; Tubb3: GACAACTTTATCTTTGGTCAGAGTGGTGCTG, Tubb3-re: GATGCGGTCGGGGTACTCC; Nkx2.1-fw: TACCACATGACGGCGGCG, Nkx2.1-re: ATGAAGCGGGA-GATGGCGG; Dlx1-fw: TCACACAGACGCAGGTCAAGATATGG, Dlx1-re: AGATGAGGAGTTCGGATTCCAGCC; HoxA6-fw: CTGTGCGGGTGCCGTGTA, HoxA6-re: GCGTTAGCGATCTCGATGCGG ; Hb9-fw: CGAACCTCTTGGGGAAGTGCC, Hb9-re: GGAACCAAATCTTCACCTGAGTCTCGG; Vglut1-fw: GGTCACA-TACCCTGCTTGCCAT, Vglu1-re: GCTGCCATAGACATAGAAGACAGAACTCC; Gad2-fw: AAGGGGAC-TACTGGGTTTGAGGC, Gad2-re: AGGCGGCTCATTCTCTCTTCATTGT; Isl1-fw: ACCTTGCGGACCTGCTATGC, Isl1-re: CCTGGATATTAGTTTTGTCGTTGGGTTGC, Tubb3-fw: GACAACTTTATCTTTGGTCAGAGTGGTGCTG; Tubb3-re: GATGCGGTCGGGGTACTCC, Mapt-fw: ACACCCCGAACCAGGAGGA; Mapt-re: GCGTTGGAC GTGCCCTTCT ; App-fw: TCAGTGAGCCCAGAATCAGCTACG, App-re: GTCAGCCCAGAACCTGGTCG, Pink1-fw: GGGATCTCAAGTCCGACAACATCCT, Pink1-re: CTGTGGACACCTCAGGGGC; Snca-fw: ATGGAGTGACAACAGTGGCTGAGA, Snca-re: CACAGGCATGTCTTCCAGGATTCC; Prnp-fw: ACCATCAAGCAGCACACGGTC, Prnp-re: GACAGGAGGGGAGGAGAAAAGCA; Nrnx1-fw: GTGGGGAATGTGAGGCTGGTC, Nrnx1-re: TCTGTGGTCTGGCTGATGGGT Aqp4-fw: GCTCAGAAAACCCCTTACCTGTGG Aqp4-re: TTCCATGAACCGTGGTGACTCC Gjb6-fw: CGTACACCAGCAGCATTTTCTTCC Gjb6-re: AGTGAACACCGTTTTCTCAGTTGGC SparcL-fw: CCCAGTGACAAGGCTGAAAAACC SparcL-re: GTAGATCCAGTGTTAGTGTTCCTTCCG Slc1a3-fw: CTCTACGAGGCTTTGGCTGC Slc1a3-re: GAGGCGGTCCAGAAACCAGTC Pla2g7-fw:

GGGCTCTCAGTGCGATTCTTG Pla2g7-re: CAACTCCACATCTGAATCTCTGGTCC Aldh1l1-fw: CTCGGTTTGCTGATGGGGACG Aldh1l1-re: GCTTGAATCCTCCAAAAGGTGCGG Pygb-fw: GGACTGTTATGATTGGGGGCAAGG Pygb-re: GCCGCTGGGATCACTTTCTCAG Vim-fw: GAGATGGCTCGTCACCTTCGTG Vim-re: CCAGGTTAGTTTCTCTCAGGTTCAGG

Toxicants: Retinoic acid (cat. # R2625; Sigma), Cyclopamine: (cat. # 239803; Calbiochem), PD184352: (cat. # Axon 1368, axon medchem), SU5402 (cat. # 572631, Calbiochem), CHIR99021 (cat. # Axon 1386, axon medchem)

Cell culture and differentiation

The murine embryonic stem cell (mESC) line CGR8[23], kindly provided by K.-H. Krause (Geneva), was cultured in complete Glasgow's modified Eagles medium (GMEM), supplemented with 10% heat inactivated fetal bovine serum (FBS; PAA, Pasching, Austria), 2 mM Glutamax, 100 µM non-essential amino acids, 50 µM β-mercaptoethanol, 2 mM sodium pyruvate and 1000 U/ml leukemia inhibitory factor (Chemicon). Cells were kept at 37°C in 5% CO_2 on tissue culture plates coated with 0.1% gelatin, and were routinely passaged every 48 h.

The mESC were differentiated towards the neural lineage according to the protocol developed by Ying and colleagues[9]. At critical steps, we used the following parameters: cells were plated in the priming phase at 1.2 x 10^5 cells/cm^2 in complete GMEM on 0.1% gelatin coated Nunclon culture dishes (Nunc, Langenselbold, Germany). Next day, for neural induction, cells were plated on gelatin-coated Nunclon dishes at 10^4 cells/cm^2 in N2/B27 medium (composition as described in[9], for a detailed description of B27 see http://www.paa.com/cell_culture_products/reagents/growthsupplements/neuromix.html). On day 7 of differentiation (DoD7) for neuronal generation and maturation, cells were replated at 10^4 cells/cm^2 on poly-L-ornithin (10 µg/ml) and laminin (10 µg/ml) coated Nunclon dishes in N2/B27 medium. Cells were fed every other day with complete medium change with N2/B27 medium.

Immunostaining and FACS analysis

For immunocytochemical analysis, cells were fixed with methanol (-20°C) or 4% paraformaldehyde (PFA) in phosphate-buffered saline (PBS) and permeabilized with 0.1% Triton X-100 in PBS. After blocking with 10% FBS, cells were incubated with primary antibodies (Tuj1 1:1000, NeuN 1:200, GAD65 1:200, SV2 1:200, PSD95 1:500, Nestin 1:500, Nestin-647 1:40, GFAP 1:800) over night. After incubation with appropriate secondary antibodies nuclei were counterstained with Hoechst H-33342 dye. Images were taken on the original cell culture dishes using an IX81 inverted microscope (Olympus, Hamburg, Germany) equipped with a 40x, NA 0.6 long range lens and processed using CellP imaging software (Olympus). For confocal microscopy cells were grown on 4-well chamber slides (Nunc), fixed with 4% PFA/2% sucrose in PBS and permeabilized with 0.6% Triton X-100 in PBS. After blocking with 5% BSA/0.1% Triton X100 in PBS cells were incubated with Tuj1 antibody in blocking buffer for one hour at room temperature. After incubation with appropriate secondary antibodies nuclei were counterstained with DAPI. Confocal images were taken using a Zeiss LSM 510Meta confocal microscope equipped with a Plan Apochromat 63x, NA 1.4 oil DIC lens. Images were analyzed and processed using ImageJ.

For flow cytometry, cells were dissociated on DoD7 with accutase, fixed and permeabilized in Cytofix Buffer followed by Perm Buffer I (both BD Bioscience, Franklin Lakes, USA), and stained with anti-nestin antibody conjugated to Alexa-647, or isotype control. Cells were analyzed with an Accuri C6 flow cytometer (Accuri Cytometers, Ann Arbor, USA) and data processed with CFlow Plus (Accuri Cytometers).

Quantitative PCR and quality control of differentiation

Total RNA was isolated at indicated time points for marker gene expression analyses using Trizol, the RNA was retro-transcribed with SuperScript II reverse transcriptase, and the resultant cDNAs were amplified in a Biorad Light Cycler (Biorad, München, Germany) with

primers specific for the genes of interest and designed for a common melting temperature of 60°C. Real-time quantification for each gene was performed using SybrGreen and expressed relative to the amount of *gapdh* mRNA using the 2^(-Delta Delta C(T)) method[24]. For each run, the consistency of conditions and constancy of *gapdh* amounts in the samples was controlled by assessment of its absolute cycle number (= 18 ± 0.5).

Gene expression analysis

Cells were used for RNA preparation as undifferentiated mESC before the priming phase (day 0), on DoD7 (before replating), on DoD15 and on DoD20. RNA was extracted from Trizol preparations and purified using RNeasy Mini prep columns (Qiagen). The total RNA harvested was quantified using a Nanodrop device (Thermo Scientific, USA) and its integrity was assessed using Agilent Bioanalyser (Agilent, USA). Illumina TotalPrep RNA Amplification Kit (Ambion, USA) and 500 ng total RNA of each sample was used according to the manufacturer's protocol to produce biotin-labelled cRNAs. For hybridization onto Sentrix Mouse Ref.8 V2 mRNA microarray beadchips (Illumina), 750 ng labelled cRNA were incubated for 16 h at 58°C. After hybridization, chips were washed, blocked, and streptavadin-Cy3 stained. Fluorescence emission by Cy3 was quantitatively detected using BeadArray Reader Scan. Statistical analysis data is based on duplicate samples. Each of the samples contained pooled RNA from two differentiations to further increase robustness of results. Technical variation of the chip was minimal as tested by rerun of the same sample on two different arrays and by comparison of results from two beadchips within one array.

Data analysis

Original and processed data have been deposited for public access in the EBI Arrayexpress database (Accession Number to be supplied). For initial processing, data were uploaded to Beadstudio (Illumina) for background subtraction. Further processing (baseline

transformation and normalization to 75 percentile) and analysis was performed with Genespring 9.0 (Agilent, Santa Clara, CA), and all normalized expression kinetics data sets were used as input for an unsupervised non-hierarchical clustering with relation to the average of expression of all genes on the chip, using the K-means algorithm. The eight major clusters were selected for further analysis. Within these, significant gene expression profiles were selected, based on a minimum regulation of 2.0-fold on any of the time points and on two-way ANOVA taking into account the regulation range and the variation between different arrays.

Patch-clamp recording

Electrodes with a resistance of 2-5 MΩ were pulled of borosilicate glass (Clark, G150F, Warner Instruments, Hamden, CT, USA) on a Sutter Instruments (Novato, CA, USA) P-97 horizontal micropipette puller. All experiments were carried out using a custom built recording chamber (800 µl volume) made of Teflon within a temperature-controlled microscope stage (37°C). Whole cell voltage and current clamp recordings were obtained from cells at day of differentiation (DoD) 20-24. Cells were grown on coated glass cover slips (10 mm) from DoD7 on. Whole-cell currents were recorded using an L/M-EPC-7 amplifier (List Medical Electronic, Darmstadt, Germany), digitised at sampling frequencies between 10 kHz to 50 kHz using a DigiData 1320A AD/DA converter (Axon Instruments Inc.). The patch pipettes for spontaneous and evoked action potential measurements as well as for the neurotransmitter responses were filled with (in mM) 90 K^+-gluconate, 40 KCl, 1 $MgCl_2$, 10 NaCl, 10 EGTA, 4 Mg-ATP, 10 HEPES/KOH (pH 7.4 at 37°C), whereas the bath solution contained (in mM): 155 NaCl, 1 $CaCl_2$, 3 KCl, 10 D-(+)-glucose, 10 HEPES/NaOH (pH 7.4 at 37°C). The protocol for recording of Na^+ and K^+ channels was as follows: cells were hyperpolarized to -90 mV, and subsequently stepped to a defined voltage as indicated and returned to -70 mV, before the next cycle with a different voltage step was run. Each cycle

took 120 ms. For the neurotransmitter response measurements, the different substances were directly added as concentrated stock solutions to the recording chamber in amounts of 1-10 µL. Antagonists were added at least one min before the agonists. Recordings were initiated within 100 ms after addition of agonists. For the measurement of barium currents through calcium channels the pipette filling solution contained (in mM) 110 CsF, 10 NaCl, 20 TEA-Cl, 10 EGTA, 4 Na_2-ATP, 10 HEPES/CsOH (pH 7.4 at 37°C), whereas the bath solution contained (in mM) 130 NaCl, 10 $BaCl_2$, 10 D-(+)-glucose, 5 tetraethylammonium chloride, 10 4-aminopyridine, 0.5 tetrodotoxin, 10 HEPES/NaOH (pH 7.4 at 37°C).

All current signals were normalized against the individual cell capacitances (as a surrogate measure for cell size) and are expressed in current densities (current divided by cell capacitance). Liquid junction potentials (LJP) were measured and corrected, using the method described by Erwin Neher (1992) except for barium current measurements. Stimulation, acquisition and data analysis were carried out using pCLAMP 10.2 (Axon Instruments Inc.) and ORIGIN 8.0 (OriginLab Corp., MA, USA). Fast and slow capacitive transients were cancelled online by means of analogue circuitry. Residual capacitive and leakage currents were removed online by the P/4 method. Series Resistance Compensation was set to at least 50%. For analysis, traces were filtered offline at 5 kHz. Cells for measurements were chosen with respect to their morphological phenotype (small round highly elevated (phase-bright) cell bodies with projections of at least five times cell body diameter, growing in network-like clusters containing at least 20-30 similar cells). The patch pipette was approached to these cells perpendicular to the plane formed by the cell membrane in the patch region.

Statistics and data mining

The numbers of replicates of each experiment are indicated in figure legends. Data were presented, and statistical differences were tested by ANOVA with post-hoc tests as appropriate, using GraphPad Prism 4.0 (Graphpad Software, La Jolla, USA). Assignment of

significantly overrepresented gene ontology (GO) categories to different clusters, and calculation of probabilities of a false-positive assignment was performed by G-profiler (http://biit.cs.ut.ee/gprofiler/[25]). For coverage of biological domains without appropriate and well-controlled GO category, relevant genes were assembled from the literature and cross-checked by 2-3 independent specialists. The number of genes within these groups identified in this study was indicated in relation to the overall number of possible hits or in relation to their distribution over different clusters. The genes defined in this study as embryonic stem cell markers or neural stem cell (NPC) markers were derived from recent literature[3]. Neuronal (N) differentiation markers (n = 574) were defined as all members of gene ontology (GO) GO:0048699 (generation of neurons) corrected for those genes used as NPC markers. The graphical representation of identified genes (or groups) within their biological context is based on the major gene function as indicated on the NCBI-gene website and the literature. Importantly, members of each identified group were scored according to their suitability as markers for a PCR-based quality control of the differentiation pattern in toxicity experiments. Several selection rounds were run to identify the final set of markers displayed as example genes in the tables and some of the figures.

Toxicity experiments

Cells were exposed to chemicals during different phases of differentiation to test the suitability of the model system for neurotoxicity testing, and for testing of developmental neurotoxicity during defined time windows. Retinoic acid (1 µM), "3i" (a mixture of 0.8 µM PD184352, 2 µM SU5402, 3 µM CHIR99021)[26] or cyclopamine (1 µM) were added to cultures from DoD1-DoD7 or from DoD8-DoD15. Then the experiment was ended, or incubation continued in the absence of chemicals for additional 6 days. On the final day, RNA was prepared by the Trizol method for PCR analysis. For morphological observations, the monolayer regions within the culture wells were imaged. Genes were preselected before the

analysis as endpoints for initial proof-of-concept experiments, and results from all genes chosen are presented.

FIGURE LEGENDS

Figure 1. Protein and mRNA-based markers of robust neuronal differentiation of mESC.

A. Cultures of mESC were fixed and stained on day 20 of differentiation. DNA, (blue) was stained with H-33342. Proteins are indicated as text on the micrograph in the same color as used for the display of their staining pattern. Tuj1: neuronal form of beta-III tubulin; NeuN: nuclear neuron-specific antigen; GAD: glutamate decarboxylase; SV2: synaptic vesicle glycoprotein 2a; PSD95: post-synaptic density protein 95. Scale bars: 20 µm. B. mESC cultures (n = 5 biological experiments) were differentiated towards neurons, and RNA was prepared at the indicated days of differentiation. Gene expression was quantified by quantitative RT-PCR. The means ± SD of the relative expression compared to day 0 (set to 1 on each diagram) was calculated and displayed (dotted lines). Relative gene expression data were also obtained by chip analysis and the means (n = 2) are displayed (solid line). Note the different scaling of the axes for chip or RT-PCR analysis, respectively, which was chosen for reasons of better comparability of the overall curve shapes. The figures in the diagram indicate the relative expression level on DoD20 (DoD7 for nestin) vs DoD0, and thus define the axis scaling.

Figure 2. Electrophysiological evidence for successful neuronal development.

Cells were differentiated on glass cover slips towards the neuronal lineage for 20-24 days and then placed into a temperature controlled recording chamber for whole cell patch-clamp studies. A. Representative example for the currents observed during the 20 ms voltage steps of the whole cell voltage clamp recording protocol displayed in B. Note that Na^+ currents (downwards deflection) are observed at voltages $\geq$ -40 mV (solid line). Strong depolarizing and repolarizing (K^+ currents; upwards deflection) are observed at depolarization to 0 mV

(dashed line). C. For voltage clamp recording (voltage step from – 80 mV to 0 mV) of Ca^{2x} channels Na^{+} and K^{+} channels were blocked by addition of tetrodotoxin, tetraethylammoniumchloride (5 mM), 4-aminopyridine (10 mM), and substitution of intracellular K^{+} ions by 120 mM Cs^{+}. Moreover, the measurement of Ca-currents was favoured by a bath solution containing barium ions (10 mM) instead of calcium ions. Current traces were obtained without $Ca2^{+}$-channel blocker, or with the blockers nimodipine (1 µM) or Cd^{2+} (1 mM) added. Current data at 15 ms after the voltage step were corrected for cell capacitance (indirect measure for cell size) and displayed. Data represent means ± SD. ** $p < 0.01$. D. Spontaneous action potentials were recorded in current clamp mode (0 pA). At the time indicated by an arrow, tetrodotoxin was added. The dashed line indicates 0 mV membrane potential. The scale bars indicate the dimensions of the membrane potential and the time domain. E. Recordings at individual neurons excited with specific glutamate receptor agonists in the presence or absence of blockers. Current traces were recorded after application of N-methyl-D-aspartate (NMDA) or kainic acid. All agonists were also tested in the presence of their respective specific antagonist (traces with 5-aminophosphovalerate (AP-5), 6,7-dinitroquinoxalin-2,3-dione (DNQX)). The scale bars represent the current and time dimensions of the experiment. Data are representative for $n \geq 10$ neurons (for agonists) and $n = 3$ for antagonists (on neurons with positive agonist response).

Figure 3. Detection of non-cytotoxic developmental disturbances by transcriptional analysis

Cultures of mESC were neuronally differentiated for 7, 15 or 20 days as indicated in a-d. They were exposed to retinoic acid (RA) or cyclopamine (Cyclo) for the time periods indicated by the hatched boxes. A. RNA was isolated at the indicated days (diamond) and used for quantitative RT-PCR analysis of selected differentiation and patterning markers. Headings indicate the overall biological effect, such as accelerated neuronal differentiation

(e.g. Neuronal diff. (+)) or altered patterning (e.g. Caudalization). Names are the official gene names, apart from the following: Vglut1 = Slc17a7, HB9 = Mnx1. The data indicate relative expression levels in % compared to untreated controls at the same time point, and are means ± SD from two to three independent experiments for each treatment and exposure schedule. Significance levels (by ANOVA within a given experimental condition) are indicated (*: $p < 0.05$, **: $p < 0.01$, ***: $p > 0.001$). Red boxes: upregulation; green boxes: downregulation; grey boxes: non-significant changes. The complete data set with standard deviations is given in Figure S2 B. Representative images of cultures on DoD15 in condition a. RA and Cyclopamine-treated cultures were indistinguishable from controls (ctrl.).

Figure 4. Cluster analysis of mRNA time course profiles, and their association with distinct phases of differentiation.

A. Gene expression kinetics were determined for all genes represented on the chip. An unbiased clustering analysis of the kinetic profiles of all regulated genes was performed. For each cluster (named Ia, Ib, IIa, IIb, IIIa, IIIb, IV, V), the means of the absolute expression level of all genes in the respective cluster, for each analysis time point is displayed. and plotted on a logarithmic scale; n: number of genes in the cluster B. Number of genes expressed in mESC, NPC and developing neurons (N) were analyzed by extensive literature search (mESC, NPC) or GO-analysis (N). The relative distribution of these genes across the different clusters was calculated (in %) and displayed (e.g. 65% of all ESC markers were found in cluster Ia, 35% of all N markers in cluster III).

Figure 5. Indication of a progressive change in chromatin organization and epigenetic factors in waves of fast and slow downregulation

A. Processes linked to chromatin or DNA-repair and –replication are displayed, and for each of them the number of genes found to be regulated during neuronal differentiation of mESC is displayed in brackets. The individual genes are listed in Figure S4. For each process, the percentage of genes present in the different clusters is indicated by colour-coded pie charts. All green shades represent clusters of genes downregulated from DoD0 to DoD20. B. Changes in chromatin structure during differentiation were visualized by DNA staining with DAPI (green) and confocal microscopic analysis. Left panel: undifferentiated mESC; right panels: neuronally differentiated cells on DoD20 that were stained with neuron-specific betaIIItubulin antibody (red). Scale bar: 10 µm.

Figure 6. Correlation of neural precursor formation with a transiently upregulated group of genes.

A. On DoD7, cultures were immunostained for the neural stem cell marker nestin (green) and DNA (red). Scale bar: 100 µm B. For quantification of nestin-positive NPC, cells were immunostained for nestin on DoD7, and analyzed by flow cytometry. Data are means ± SD of 7 independent differentiations. ***: $p < 0.001$. C. Relative expression profiles of genes from cluster IIa were calculated by normalization of expression of each gene to DoD0 expression, which was arbitrarily set to 1. The expression kinetics for each gene within that cluster are displayed. D. Genes upregulated during neuronal differentiation of mESC were analyzed for their role in regional specification of the brain and classified accordingly (colour-coding). The number of genes associated with each of the three chosen subregions of the brain, are displayed separately for each regulation cluster. A detailed list of genes with their regional assignment is given in figure S5.

Figure 7. Analysis of glia-associated genes

A DoD20 cultures were fixed and stained for GFAP (green; to identify astrocytes) and Tuj1 (red; to identify neurons). The left image shows a representative overview with large neuronal areas and one typical astrocytic island. The right image shows an astrocytic island in greater detail. Scale bars = 100 µm. **B** The table in the bottom part indicates the glia-related genes identified in this study, sorted by the cluster of expression kinetics they fell into. Astrocyte-related genes searched for, but not identified here were glutamine synthetase (Glul), S100b, Slc1a2 (Glt-1, Eaat2), Connexin 30/43 (Gjb6/Gja1), NfiA (also found in oligodendrocytes). Oligodendrocyte-related genes not found here were ATP-binding cassette, sub-family A (Abca2), CNPase (Cnp1), a microtubule-associated protein (Mtap4), myelin-glycoproteins (Omg and Mog), Olig2/3 (Olig2, Olig3), myelin protein zero (Mpz), Ng2 (Cspg4), NfiA. **C** Expression of selected astrocyte-related genes was monitored by qPCR. on day 0, 7, 15 and 20 of two differentiations. Data for each differentiation are given individually. The lines indicate the respective mean values.

Figure 8. Functional assignment of neuronal genes up-regulated in different waves

A combination of bioinformatics tools and literature information was used to search all upregulated clusters for conspicuous biological themes and for genes associated with them. Themes are displayed, and corresponding genes (with original NCBI gene names) are colour-coded according to the clusters they were found in (displayed graphically besides the legend, with dots on the lines representing DoD0, DoD7, DoD15 and DoD20). A. Core neurochemical themes. Note a relatively early induction of receptors and channels, compared to late emergence of genes coding for transporters and synaptic vesicles, and those related to neurodegenerative disease. B. Themes related to neurite growth indicate an early focus on

growth cone formation and guidance. C. Genes related to extracellular matrix are displayed. D. The cultures from two differentiations were exposed to medium (untreated) or lead acetate (1 µM) from DoD14 to DoD20. RNA was isolated on DoD20 and used for quantitative RT-PCR analysis of genes associated with neurodevelopment and known to affect neuronal disease. The data indicate relative expression levels in % compared to the untreated controls of the first differentiation on DoD20, and are means ± SD (n = 2). Significance levels (ANOVA) are indicated (*: $p < 0.05$, **: $p < 0.01$, ***: $p > 0.001$).

Table 1: GO categories significantly overrepresented in cluster IIa

Biological process (GO)*	**genes in IIa**	**p-value**	**examples of upregulated genes listed in the GO**
Nervous system development	51	3e-14	Neurod4, Nes (nestin), Cdh2 (N-cadherin), Fgf5, Sema5b, Efnb2
Regulation of nervous system development	17	7e-09	Nefm (neurofilament M), Chrna3 (cholinergic R.), Ntrk3, Isl1, Foxg1
Regulation of neurogenesis	16	9e-09	Hoxa2, Smo, Dll1 (delta-like 1), Hes3, Metrn, Ntrk3 (= Trkc)
Neuron projection morphogenesis	13	2e-05	Epha7, Mtap1b, Myh10 (myosin heavy chain), Egr2, Epha7, Isl1
Central nervous system development	21	1e-06	Mtap1b (microtubule-associated protein), Bmi, Foxg1, Isl1, Fgfr3
Neuron projection regeneration	5	2e-06	Mtap1b, Bcl2, Smo, Chst3 (carbohydrate sulfotransferase)
Parasympathetic nervous system development	4	4e-06	Hoxb2, Egr2, Smo (smoothened), Hes3 (hairy and enhancer of split)
Neuron development	20	5e-06	Mtab1b, Foxg1, Epha7 (Eph receptor A7), Isl1, Ulk2, Bmpr1b
Cranial nerve development	5	1e-05	Gli3, Hoxb2, Egr2 (early growth response), Smo, Hes3
Dorsal/ventral pattern formation	9	7e-07	SP8, Foxg1, Bmpr1a, Bmpr1b (bone morphogenic protein R.), Hoxa2
Tissue development	28	1e-06	Homer1, Prox1 (prospero-related homeobox 1), Fzd2, Sdc1 (syndecan)
MAPKKK cascade	12	2e-06	Mapk8, Fgf13, Jak2, Nrg1, Fgfr3, Tgfbr1, Mapk8 (=Jnk)
Anterior/posterior pattern formation	12	4e-06	Hoxb2, Hoxa2, Tgfbr1 (transforming growth factor, beta receptor)
Regulation of ossification	8	3e-05	Smad5, Calca (calcitonin), Sfrp1 (secreted frizzled-rel. protein 1), Egr2

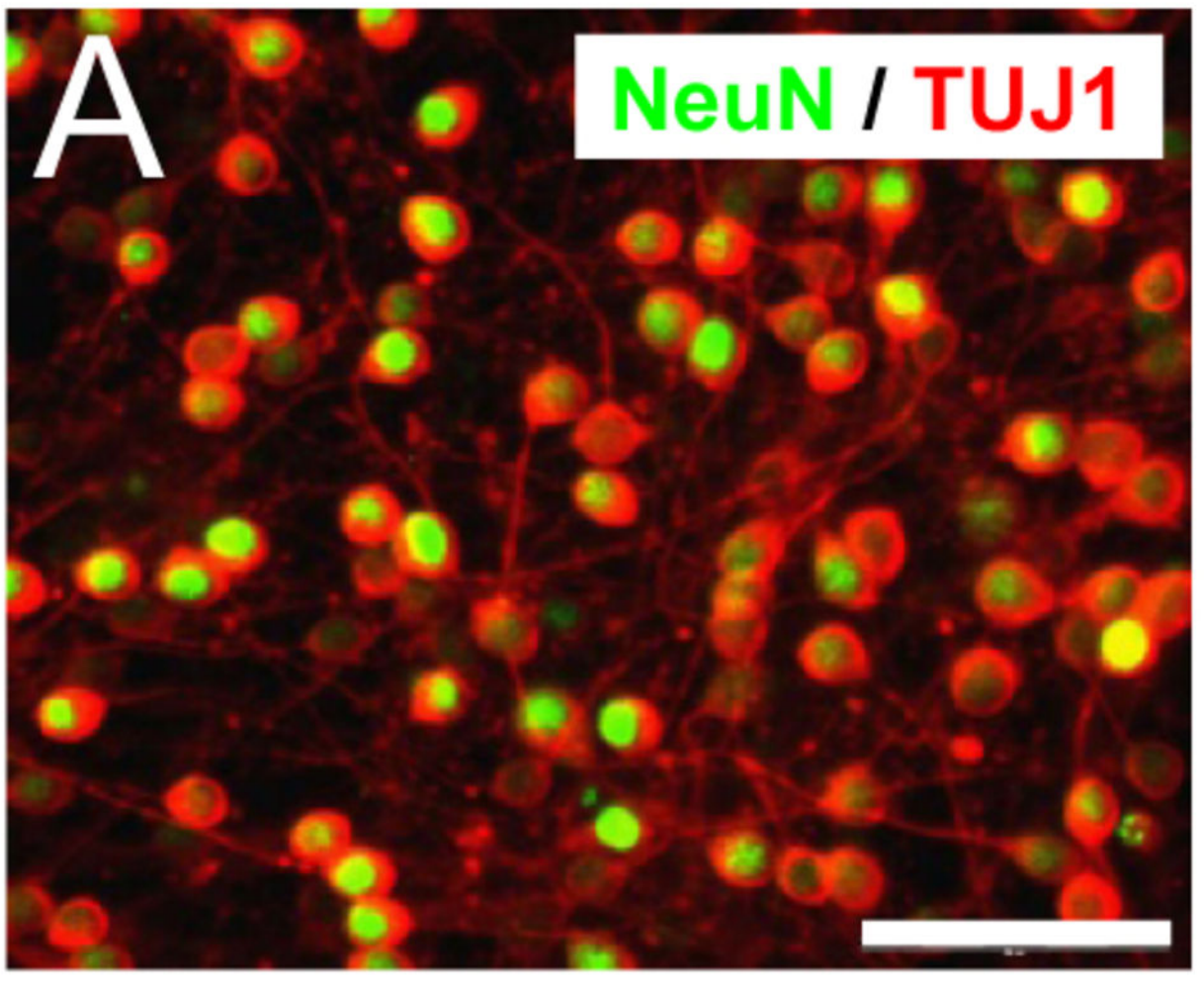
* All categories identified by gProfiler bioinformatics analysis, with their p-values indicated after correction by removal of “nervous system development” genes from non-neuronal GOs.

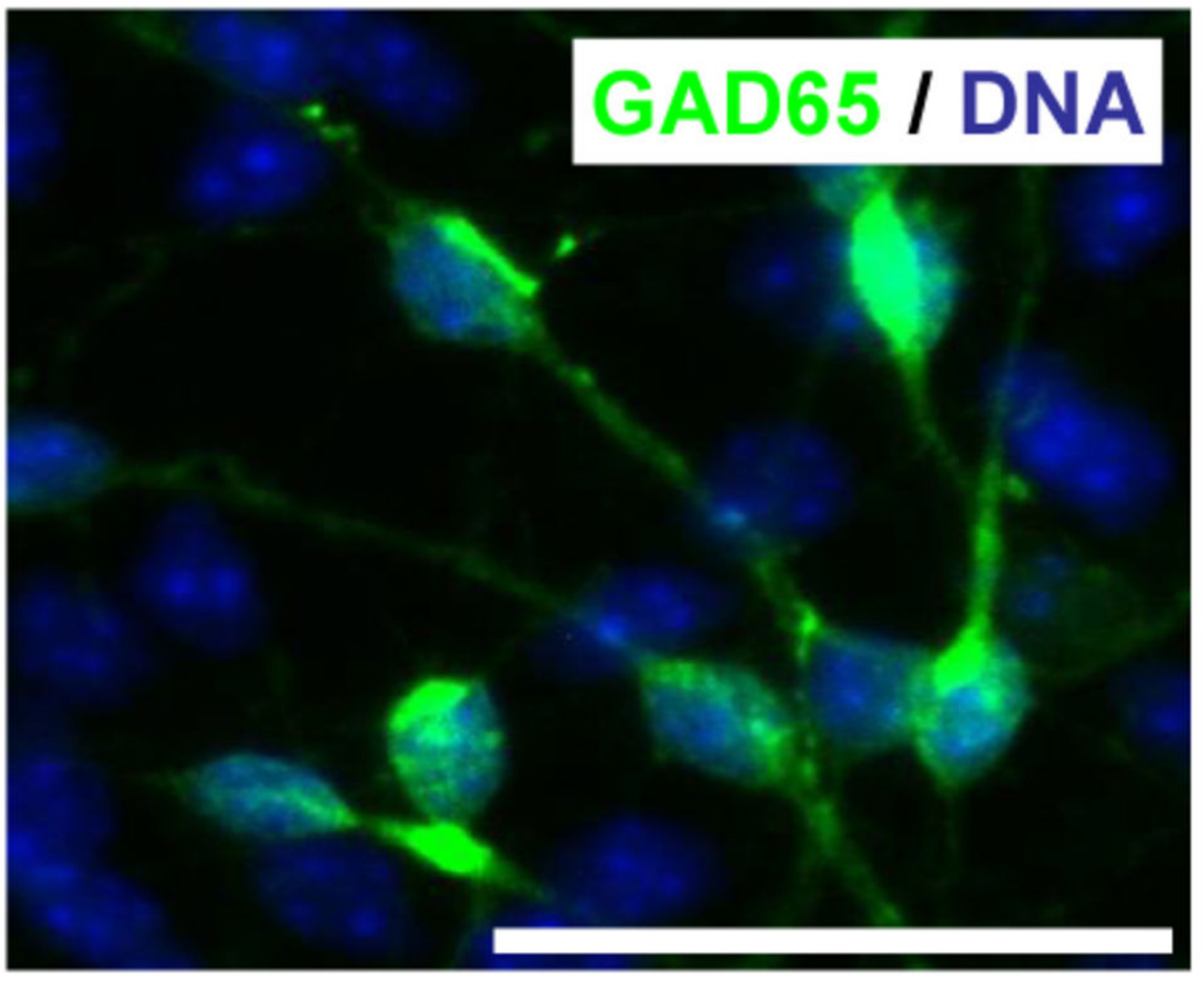
Table 2: GO categories that are overrepresented in the clusters comprising genes upregulated during differentiation

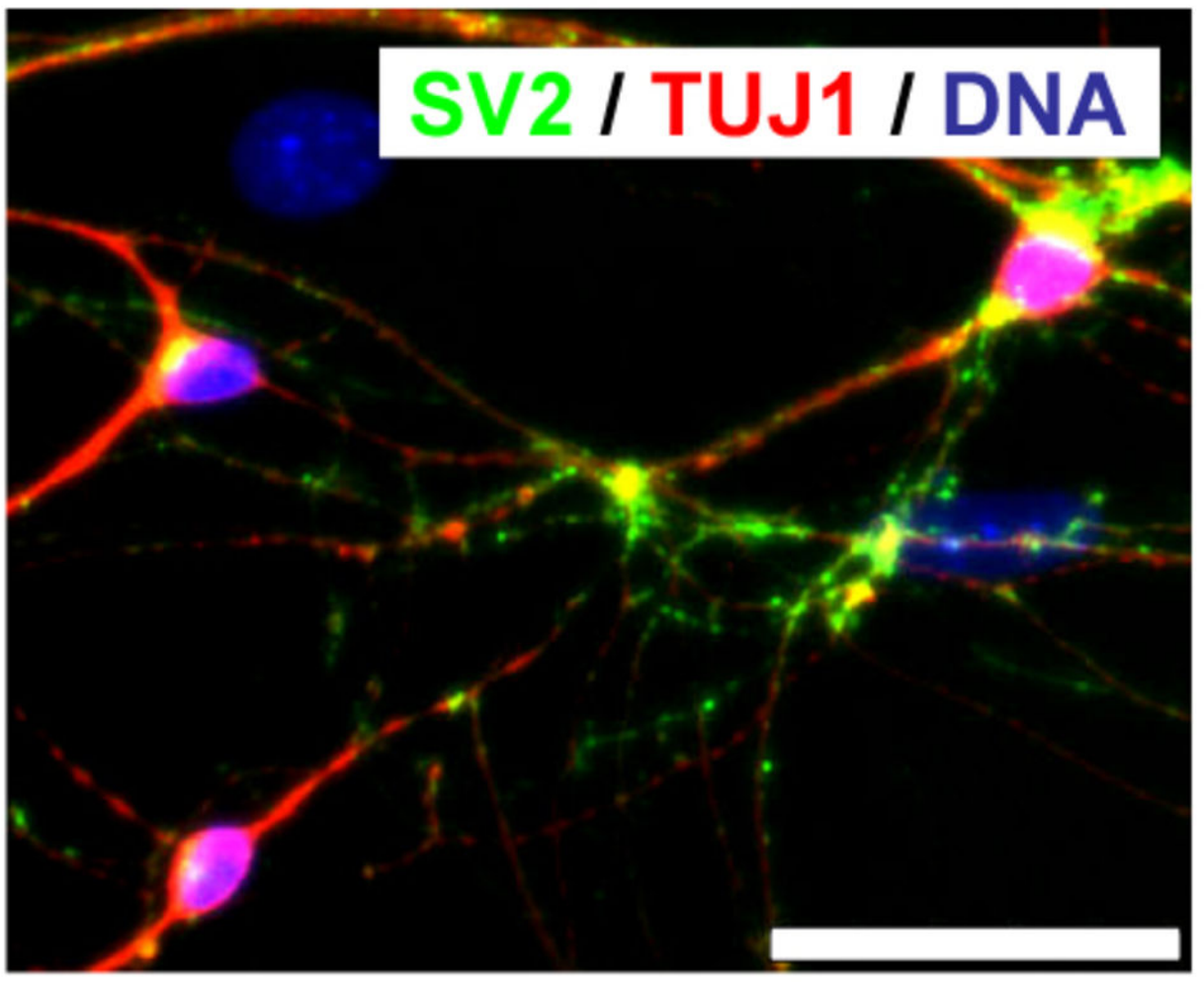
Cluster	Biological Process (GO)*	number of genes	p-value	examples of upregulated genes
IIIa/b	Nervous system development	107	2e-32	Hes5, Notch3, Otx1, FoxA2, Nkx2.2, Ntrk3, Nrxn2 (neurexin)
	Generation of neurons	69	6e-23	Sox5, Shh (sonic hedgehog), Wnt3a, Dcx (doublecortin), Nog (noggin)
	CNS development	49	1e-16	Zic1, Wnt7a, Fgf8, Pitx2
	Neuron development	39	7e-13	Gap43, Gprin2 (inducer of neurite outgrowth), App (Aβ precursor protein), Reln
	Axogenesis	28	3e-12	Cdk5r1 (kinase), EfnB1 (ephrin), Ntng1 (netrin), Stxbp1 (syntaxin binding protein)
	Axon guidance	19	3e-10	Apbb1 (APP-binding), Cxcr4, Slit2, Kif5C (kinesin), Ephb1 (ephrin-R)
	Neuron projection	32	4e-7	Grik5 (glutamate-R), Gria3 (glutamate-R), Cacna1g (Ca^{2+} channel), Mtap2 (map2)
IV	Vesicle	33	2e-7	Sv2a (synaptic vesicle glycoprotein), Syn2 (synapsin), Syt1 (synaptotagmin)
	Nervous system development	43	4e-7	Neurog2 (neurogenin), Unc5b (netrin-R), Bai2, FoxD1, Egfr, Dner, En1 (engrailed)
V	Extracellular matrix	24	10e-11	Dcn (decorin), Col1a1 (collagen), Spon2 (spondin), Lum (lumican), Tnc (tenascin)
	Lipid storage	5	4e-6	Apoa1 (apolipoprotein), Gm2a (ganglioside activator), Enpp1, Cav1 (caveolin)
	Response to stress	42	1e-5	Hspa2 (heat shock protein), Fas (fas), Fos, Pparg (PPAR-gamma). Pink1, Snca
IV+V	Extracellular matrix	39	3e-11	Col1a2 (collagen), Col3a1 (collagen), Ecm1 (extracellular matrix), Efemp2 (fibulin)
	Response to hormone stimulus	38	7e-10	Rbp4 (retinol BP), Rxra, Thra, Rgs9, Igfbp7 (insulin binding)
	Nervous system development	70	2e-10	Nrxn1 (neurexin), Mapt (tau), Tgfbr2, Dlx1
	Blood vessel development	29	5e-9	Cdh13 (cadherin-H), Prrx1, SphK1 (sphingosine kinase), Cul7 (cullin)
	Neuron projection	33	5e-7	Tubb4 (tubulin), Syt1 (synaptotagmin), Psd2, Syt4, Ttyh1 (tweety homolog)
	Neurogenesis	44	9e-7	Myo6 (myosin), Nrn1 (neuritin), En2 (engrailed), Hoxa1, Lhx5
	Synaptic vesicles	14	1e-6	Syp (synaptophysin), Slc17a6, Rabac1 (rab acceptor)
	Muscle organ development	22	3e-6	Gata6, Des (desmin), Myl2 (myosin ligh chain), Vamp5 (vesicle associated protein)
	Transmission of nerve impulse	23	2e-5	Gria2 (glutamate-R), Slc17a6 (vGlut), Chrnb1 (ACh-R), Kcnmb4 (K^+ channel)

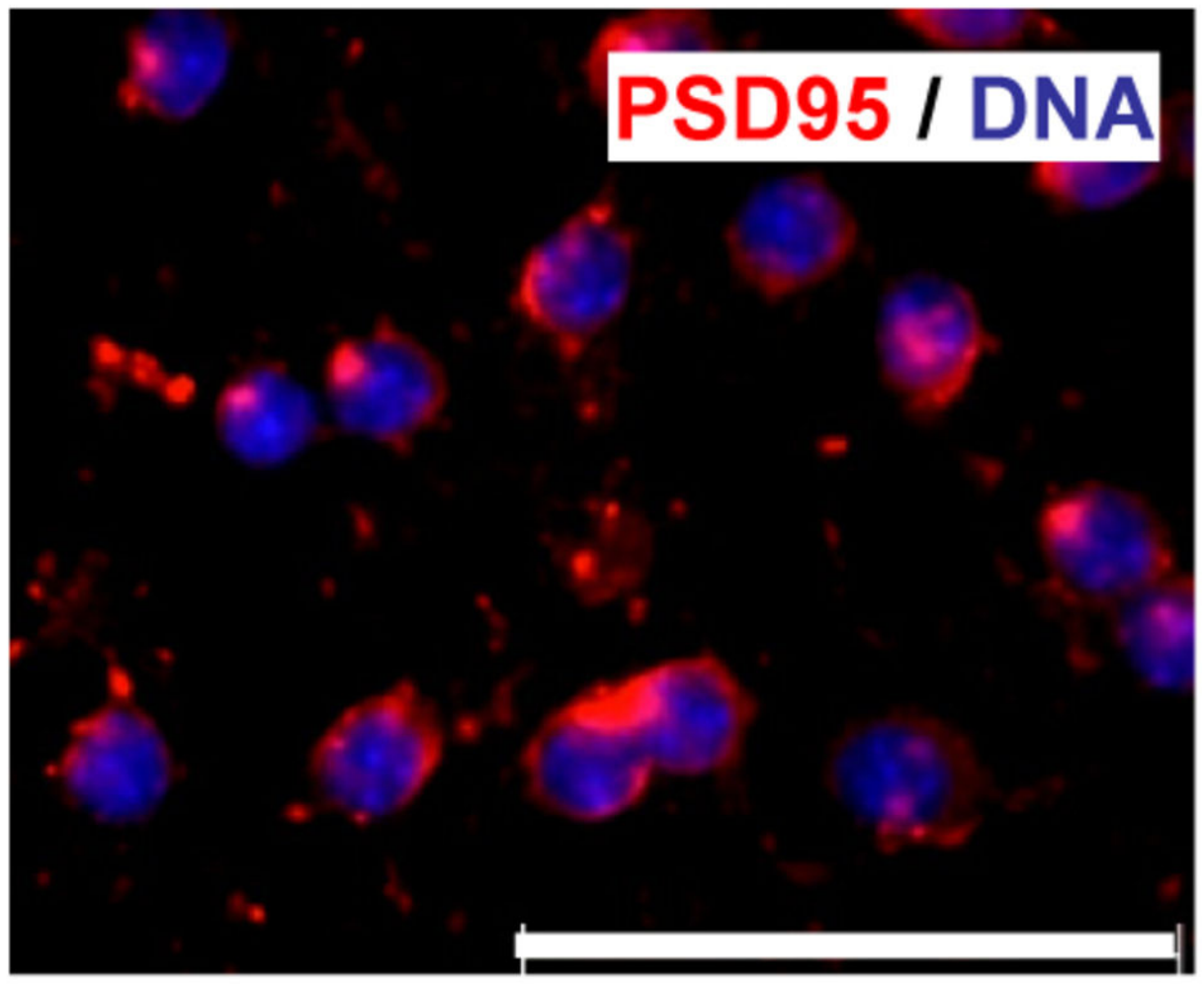
* All categories identified by gProfiler bioinformatics analysis, with their p-values indicated after correction by removal of “nervous system development” genes from non-neuronal GOs

A

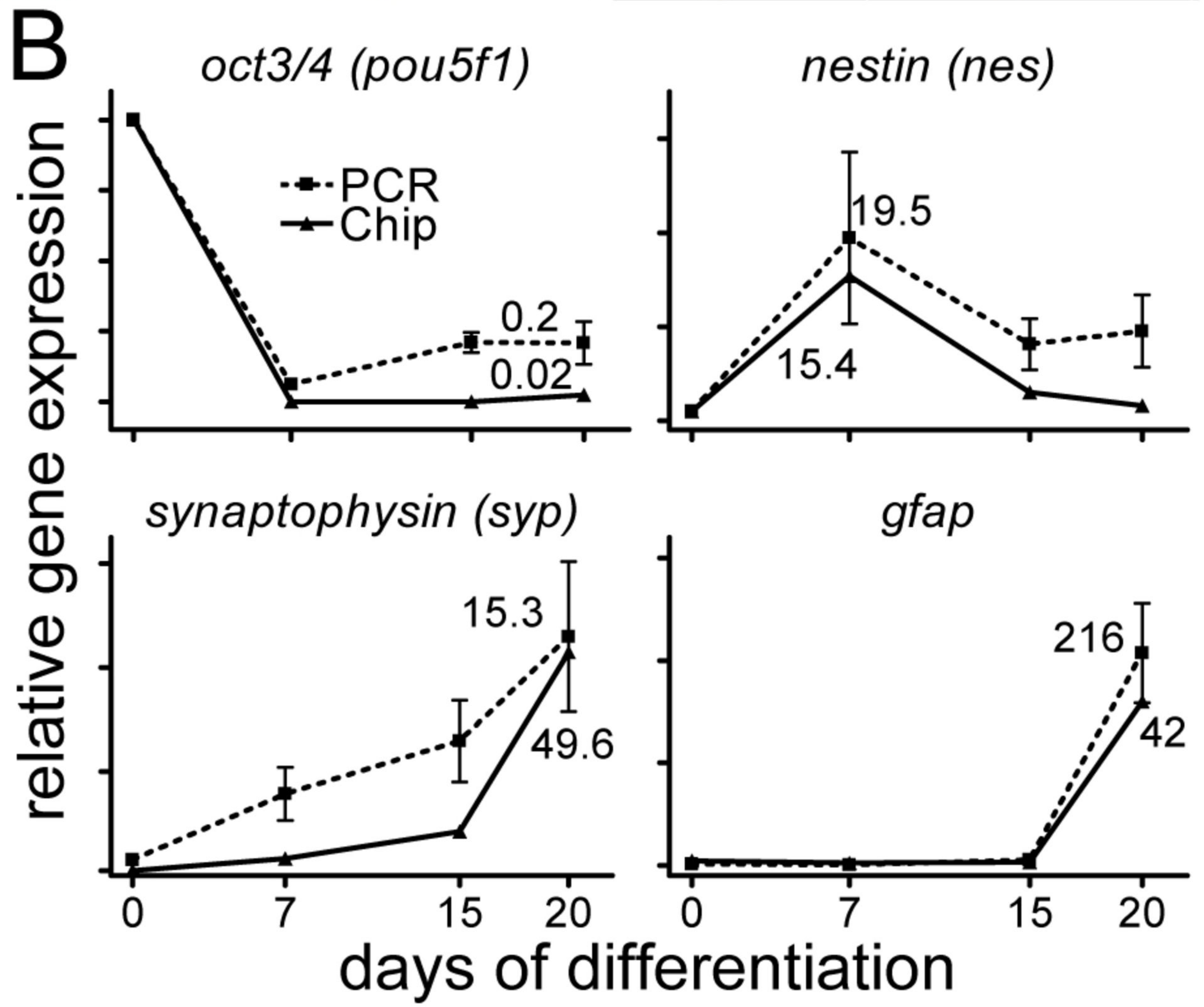

NeuN / TUJ1


GAD65 / DNA


SV2 / TUJ1 / DNA


PSD95 / DNA

B


oct3/4 (pou5f1)
nestin (nes)
synaptophysin (syp)
gfap
PCR
Chip
0.2
0.02
19.5
15.4
15.3
49.6
216
42
relative gen
e expression
0
7
15
20
days of differentiation

A

E_M (mV)

50
40
30
20
10
0
-10
-20
-30
-40
-50

1 nA
2 ms

+50 mV
-70 mV
-80 mV
-90 mV
20 ms

B

-70 mV

★★

C

Current Density (-pA/pF)

40
30
20
10
0

Control
Nimodipine
$CdCl_2$

D

V_M = 0 mV

50 mV
30 s

Tetrodotoxin (500 nM)

E

NMDA

NMDA + AP-5 (50 μM)

Kainic Acid

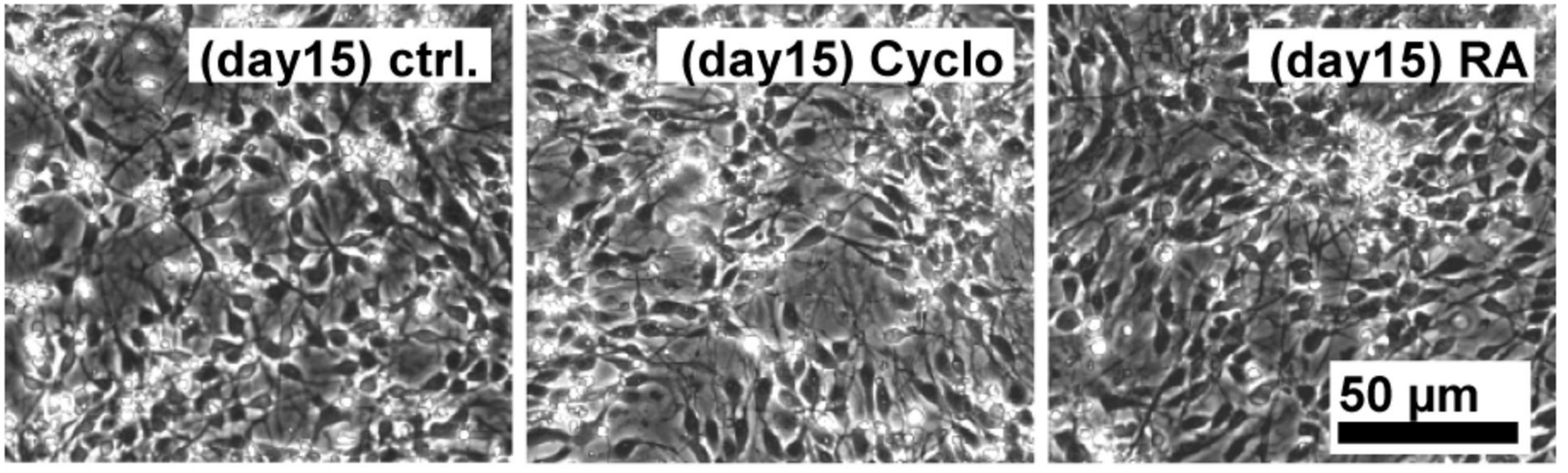
Kainic Acid + DNQX (40 μM)

250 pA
10 s

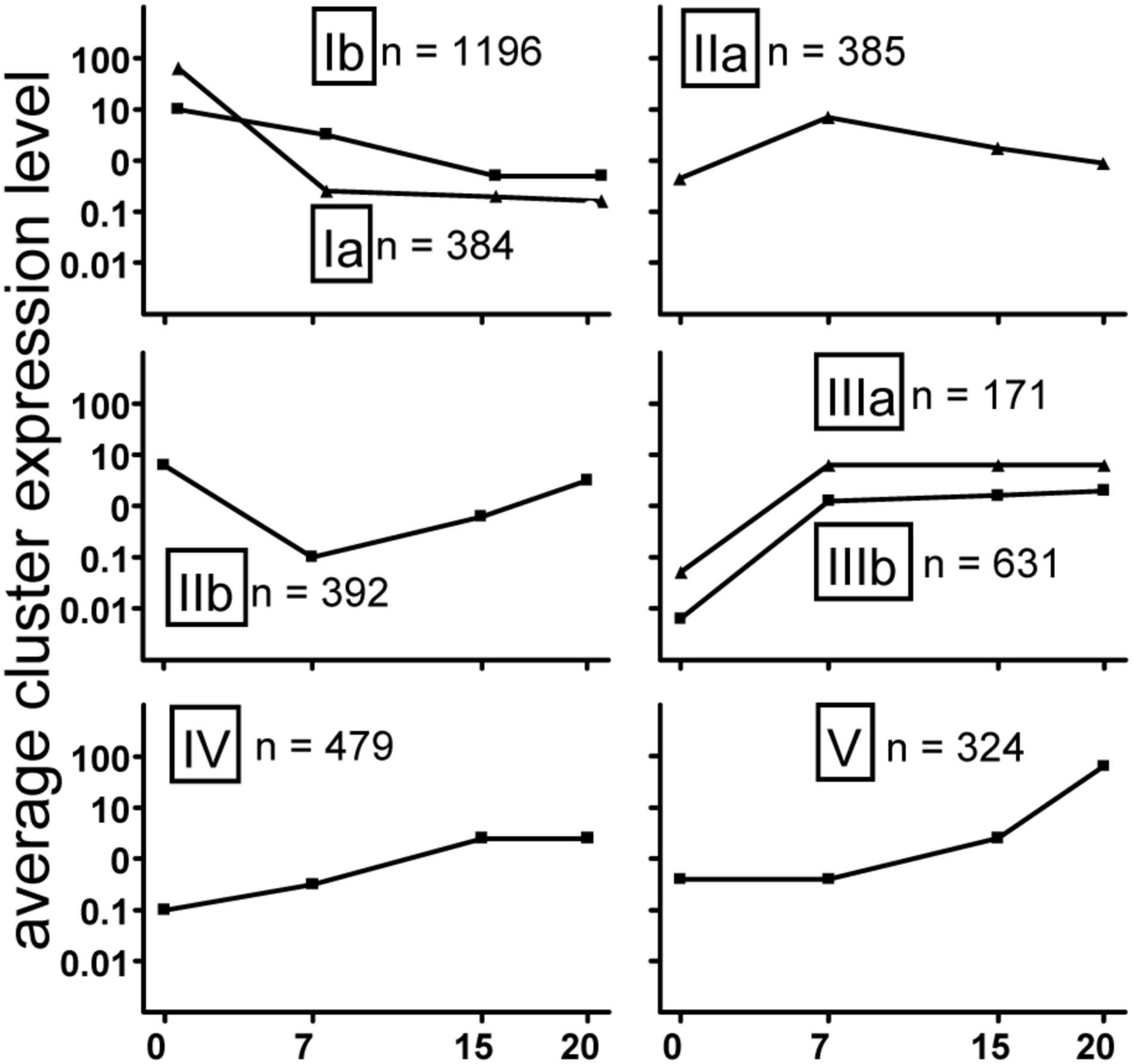
A

	Incubation scheme	RA	Cyclo
		regulation in % vs. control	
a	Overall effect examined:	Neuronal diff.(+)	Dorsalization
	0 — 7 — 15 days	Syp : 2400(***)	NKx2.1: 30 (***) Dlx1 : 13 (***) Shh : 40 (***)
b	Overall effect examined:	Neuronal diff.(+)	Dorsalization
	0 — 7 days	Syp : 450(**)	Nes : 80 (ns) Shh : 20 (***)
c	Overall effect examined:	Caudalization	Dorsalization
	0 — 7 — 15 days	HoxA6: 840000 (*) HB9 : 760(**)	Syp : 92 (ns) Vglut1:106 (ns) Gad2 : 57 (ns) Shh : 40 (***)
d	Overall effect examined:	Caudalization	Neurotransmitters
	0 — 7 — 15 — 20	HoxA6: 4500(***) Isl1 :39500 (**)	Vglut1:150 (ns) Gad2 : 17 (**)

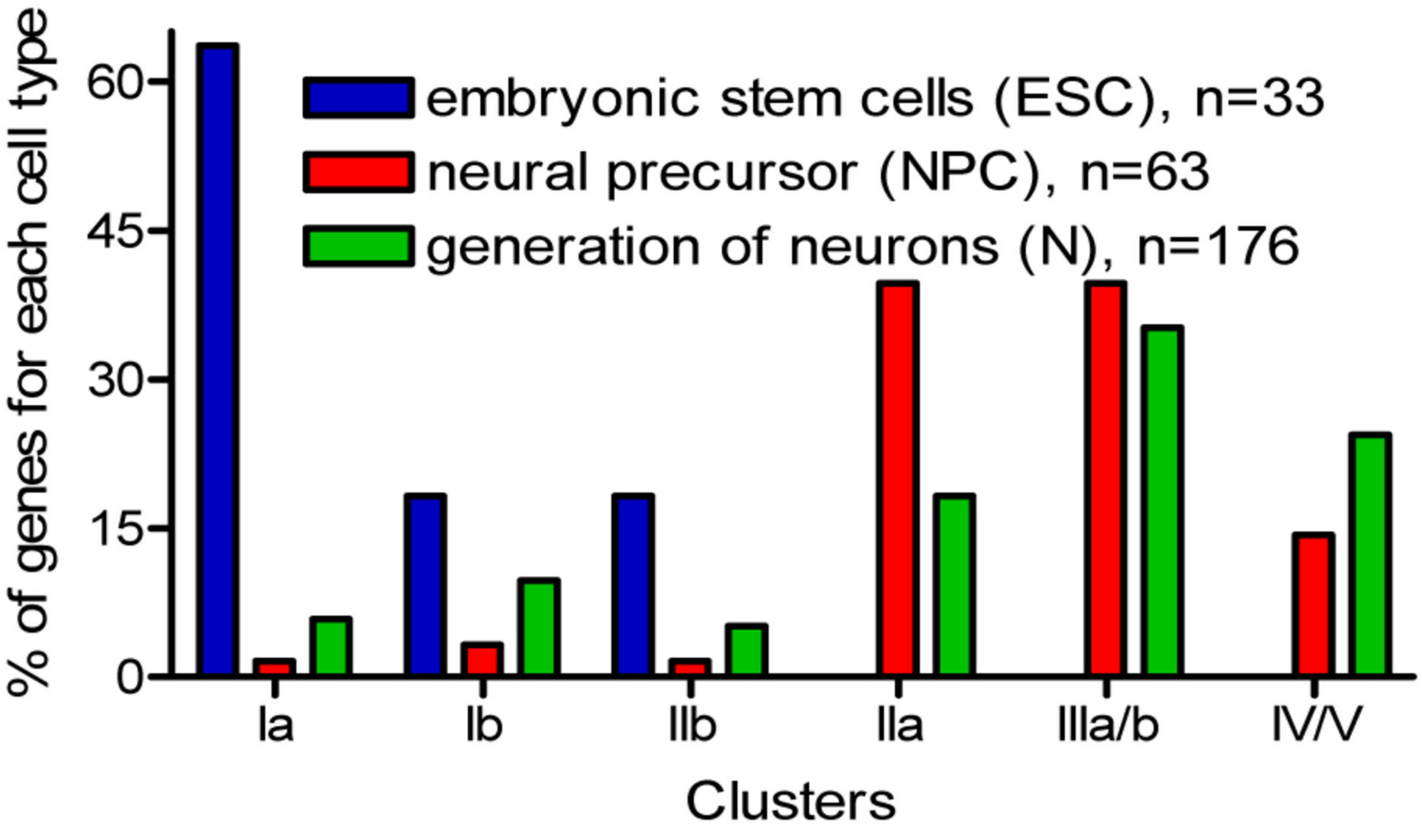
B



A


Ib n = 1196
Ia n = 384
IIa n = 385
IIb n = 392
IIIa n = 171
IIIb n = 631
IV n = 479
V n = 324
100
10
0
0.1
0.01
0
7
15
20
average cluster expression level
days of differentiation

B


embryonic stem cells (ESC), n=33
neural precursor (NPC), n=63
generation of neurons (N), n=176
60
45
30
15
0
% of genes for each cell type
Ia
Ib
IIb
IIa
IIIa/b
IV/V
Clusters

A

Histones (20)

Histone modification (46)

Me

Ac

Heterochromatin (6)

DNA Methylation (5)

Remodeling (13)

Euchromatin (6)

TTGCGCGG

AACGCGC

GCGCTAG

CGCGATC

DNA repair (62)

Chromosome structure (14)

G1

Cell cycle (299)

M

G2

S

cluster

Ia

Ib

IIb

IIa

IIIa/b/IV/V

DNA replication (57)

B

mESC

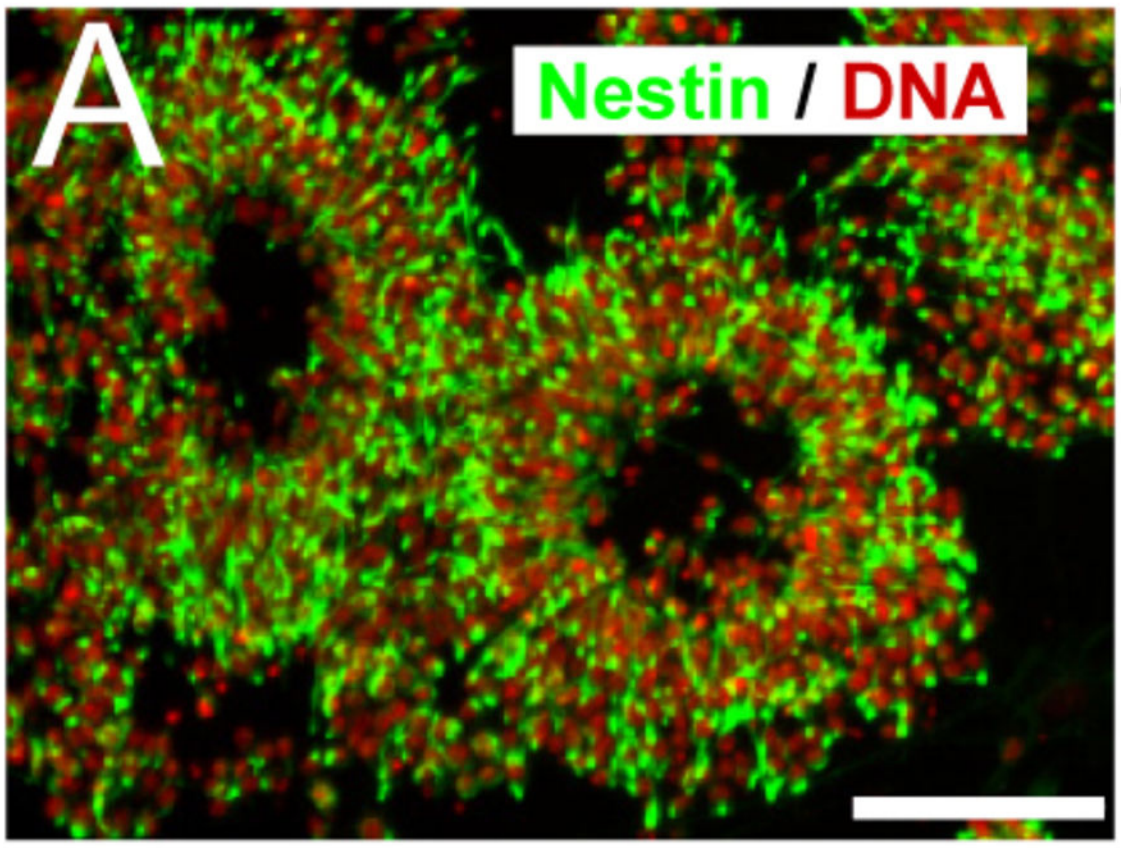
DNA

day 20 neurons

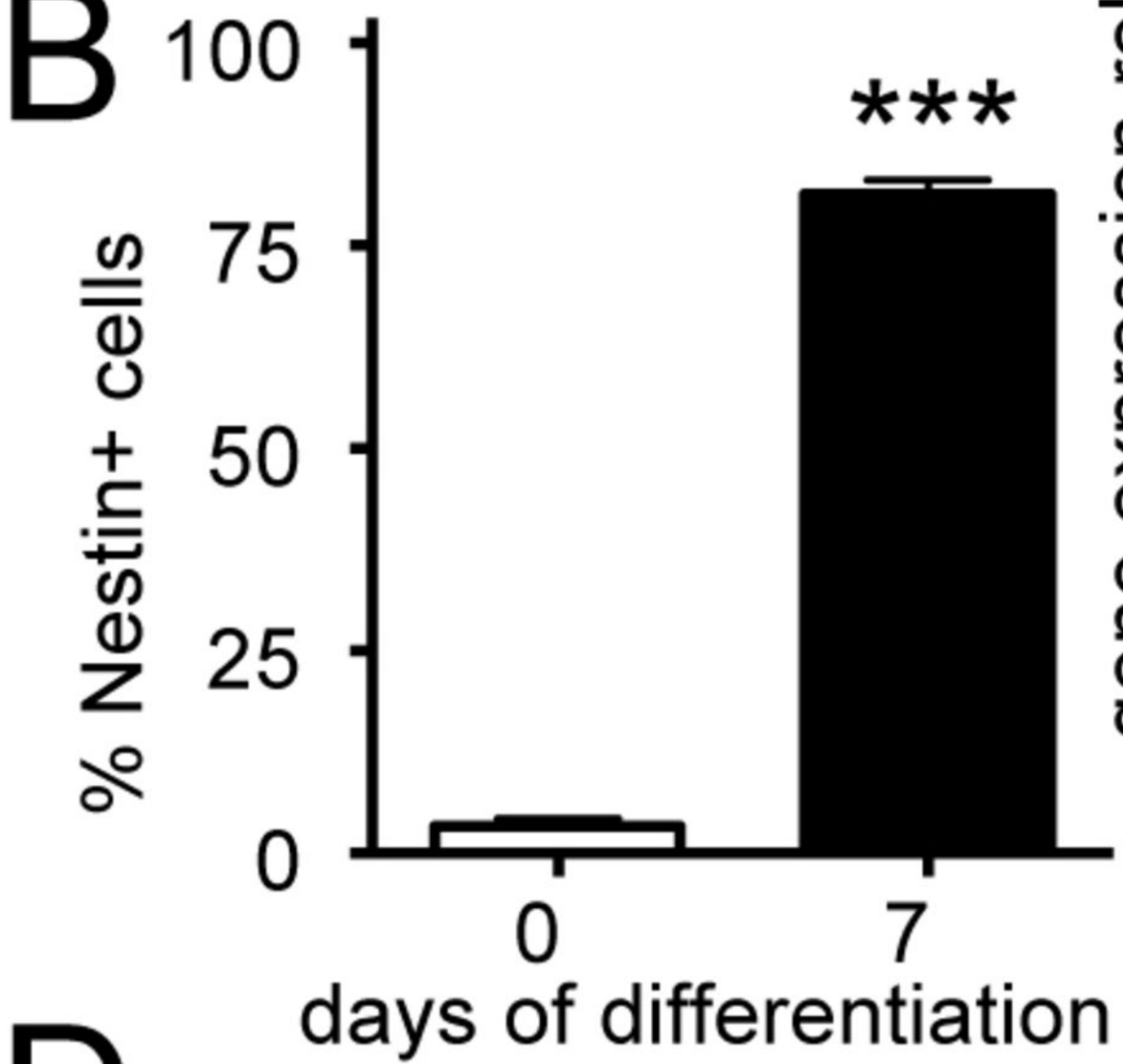
DNA / TUJ1

DNA / TUJ1

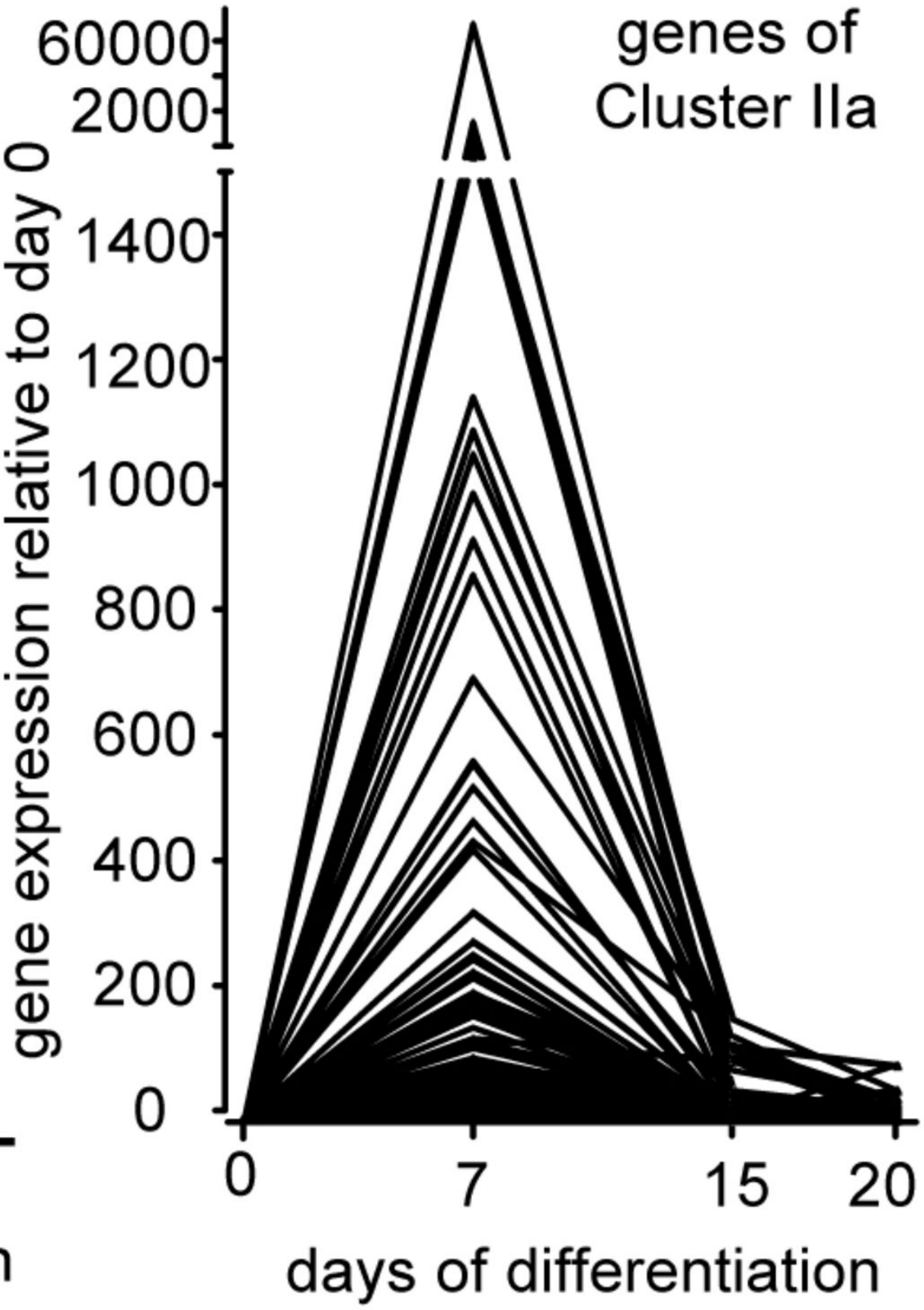
A


Nestin / DNA

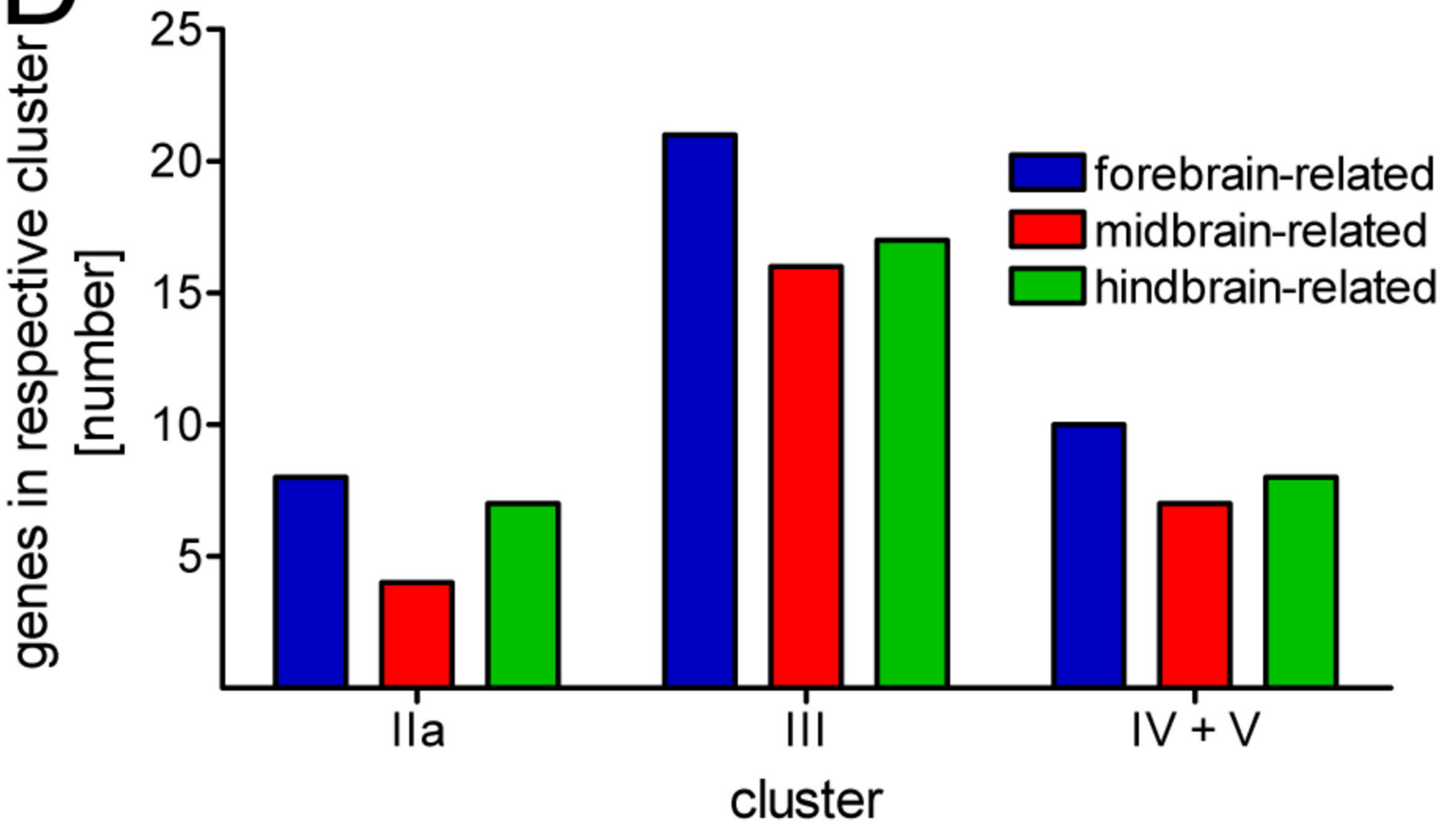
B


100
75
50
25
0
% Nestin+ cells
***
0
7
days of differentiation

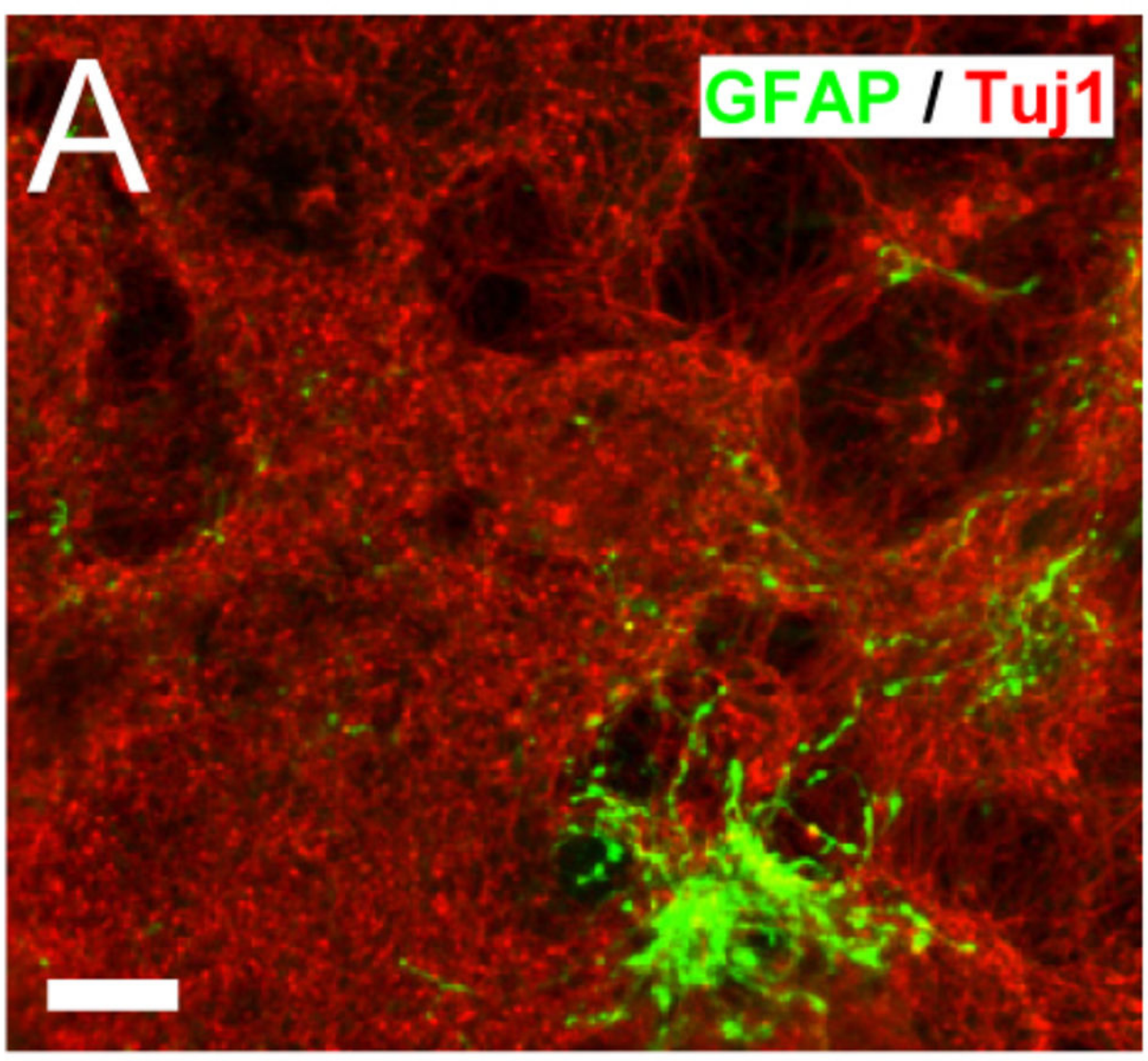
C

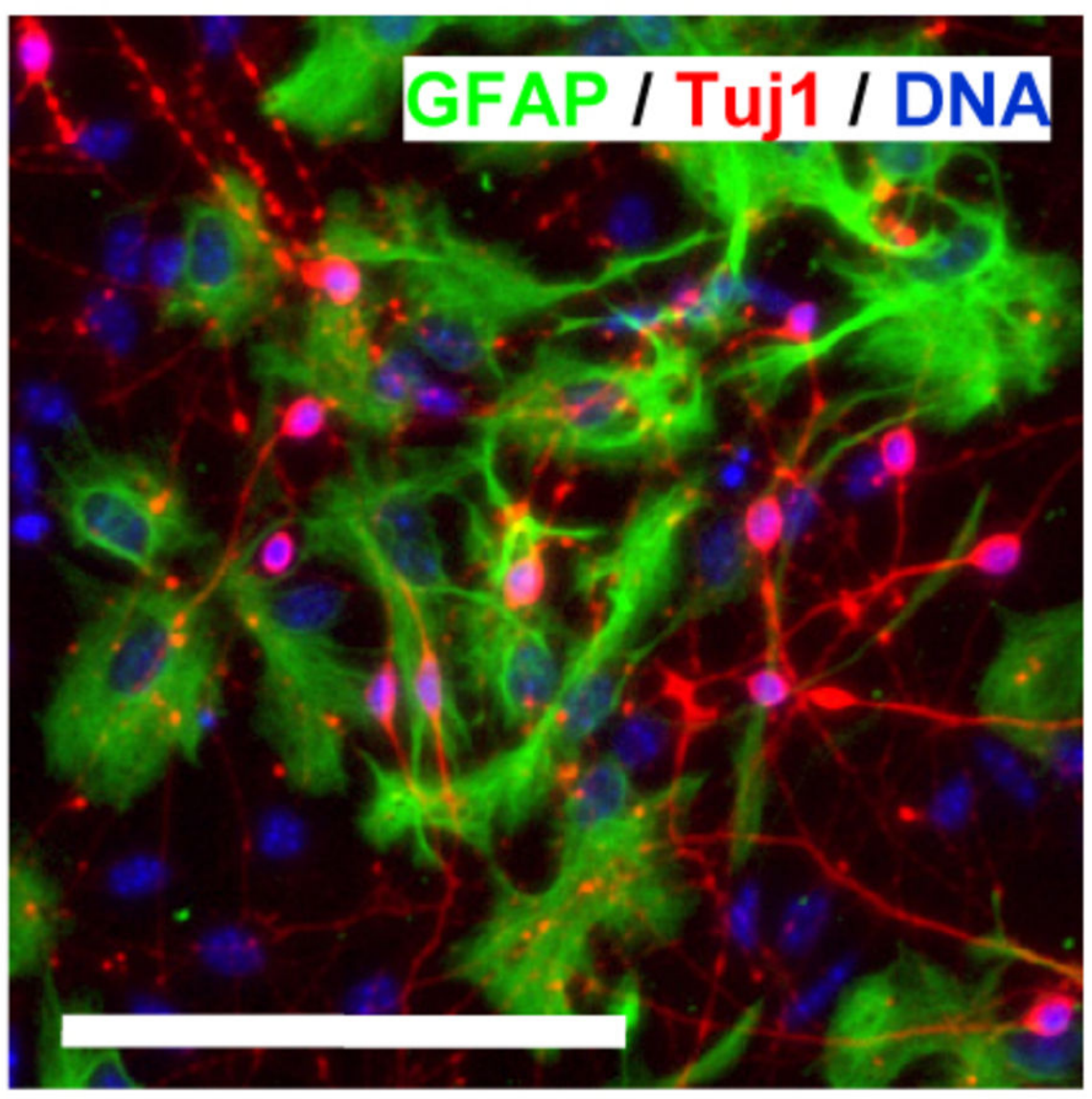

60000
2000
1400
1200
1000
800
600
400
200
0
gene expression relative to day 0
genes of
Cluster IIa
0
7
15
20
days of differentiation

D


25
20
15
10
5
genes in respective cluster
[number]
forebrain-related
midbrain-related
hindbrain-related
IIa
III
IV + V
cluster

A

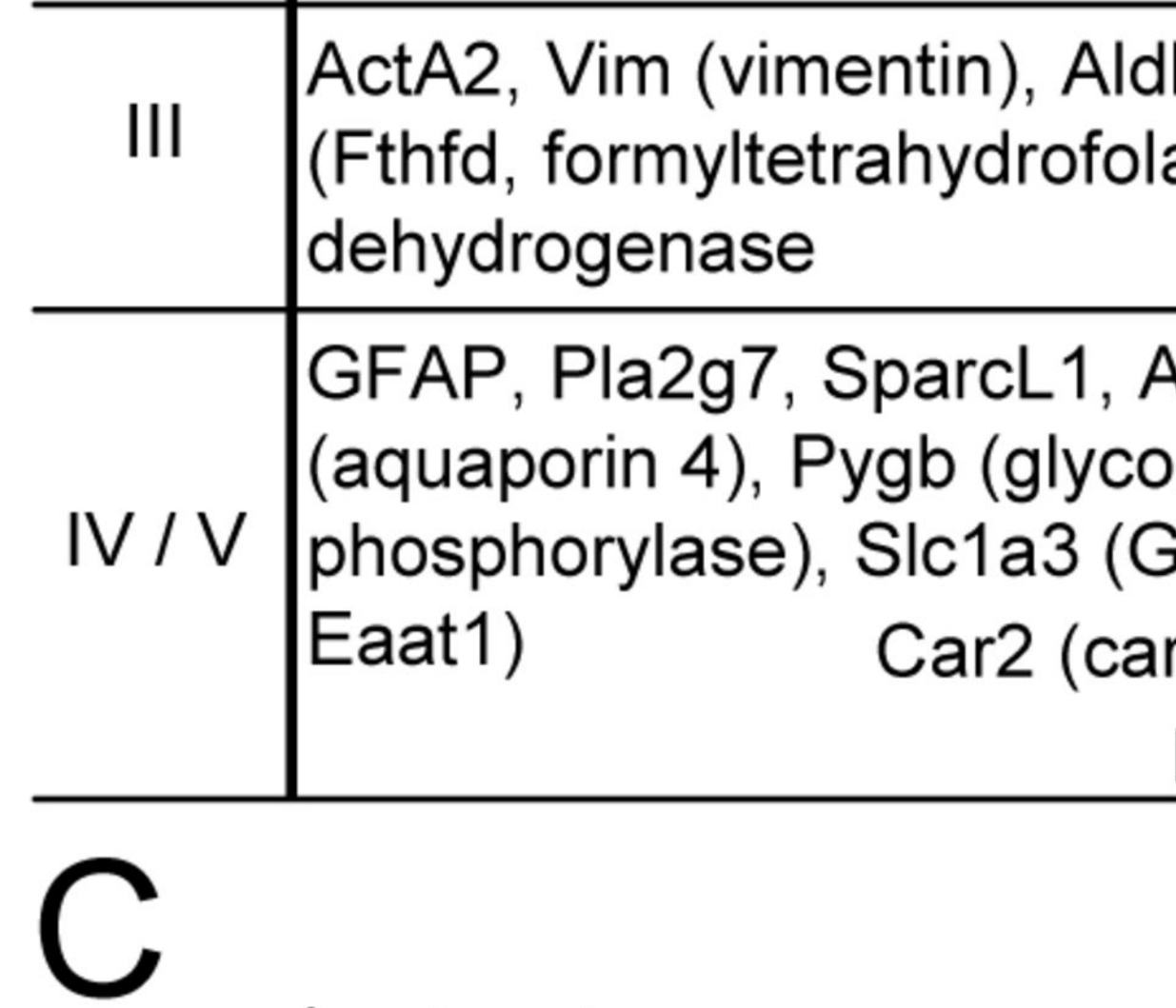


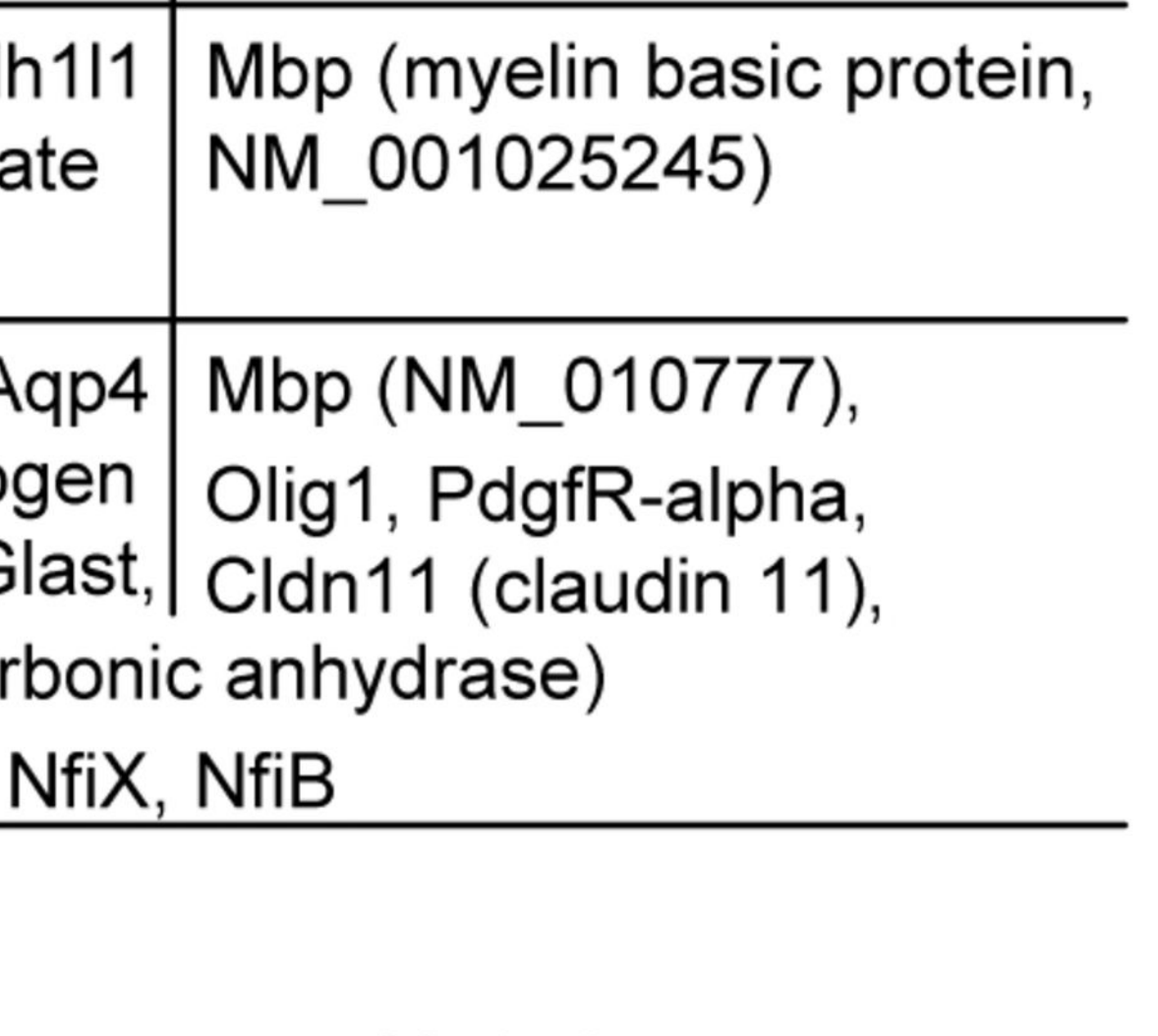


B

Cluster	Glia-associated genes identified on chip: Astrocytes	Oligodendrocytes
I	none	
II	ApoE, AldoC, Csad (cys-sulfinate-decarboxylase), Cbs (cystathionine synthase)	Trf (transferrin)
III	ActA2, Vim (vimentin), Aldh1l1 (Fthfd, formyltetrahydrofolate dehydrogenase	Mbp (myelin basic protein, NM_001025245)
IV / V	GFAP, Pla2g7, SparcL1, Aqp4 (aquaporin 4), Pygb (glycogen phosphorylase), Slc1a3 (Glast, Eaat1) Car2 (carbonic anhydrase) NfiX, NfiB	Mbp (NM_010777), Olig1, PdgfR-alpha, Cldn11 (claudin 11),

C





A

Transporter

Slc17a6, Slc17a8,
Slc6a12,
Slc6a13

Abat, Gad1

Neurodegenerative disease - associated

App, Prnp, Cln3,
Snca, Mapt,
Pink1

Synaptic vesicles

Syt11, Vamp4, Syn2,
Syt1, Vamp8, Sv2a, Syt14l,
Syp, Syt4

pre-synaptic active zone

Mapk8ip1, Caskin1

Synaptic adhesion

Nrnx2, Nlgn2, Pcdha6,
Pcdhb3, Pcdhb22, Nrnx1

Receptors

Chrna4, Grik5, Gria3,Grina,
Gabarap, Gabarapl1, Adra2a,
P2rx4, Chrnb1, Gria2,
Adra2b

Channels

Cacna1h, Cacna1g, Cacng4,
Ttyh3, Kcnab2, Cacnb3, Kcna5,
Kcnmb4, Ttyh1, Scn3b

post-synaptic proteins

Lrrc4b, Lrrc49, Nrgn, Gnao

B

Guidance cues

Slit2, Sema3f,
Ntng1, Sema6d

Cxcr4

axon

Regulators of actin cytoskeleton

Pak3, Dpysl2, Dpysl4,
Dpysl5, Cdk5R1, Rasl10b
Rasl11b, Rgs3, Cdc42ep5,
Dpysl3

Cytoskeleton

Pfn2, Dcx, Mtap2,
Tubb2a, Tubb4, Mapt

Unc5c, Unc5b

Ephb1, Ephb2, Epha3

growth cone

C

Collagens

Col4a1, Col4a2, Col4a5, Col5a1, Col2a1, Col23a, Col1a2,
Col16a1, Col1a1, Col9a2, Col3a1, Col6a1, Col6a2

Thrombospondins

Thbs3, Spon1,Thbs2

Tenascins

Tnc

Proteoglycan core proteins

Ptn, Sdc3 (Syndecan), Ncan (Neurocan),
Lum (Lumican), Ogn (Mimican), Prelp Dcn (Decorin)

Time profile:

Cluster III:

Cluster IV:

Cluster V:

expression

D